USPC
Université Sorbonne
Paris Cité

université
PARIS
DIDEROT
PARIS 7

**Thèse de doctorat**
**de l'Université Sorbonne Paris Cité**
**Préparée à l'Université Paris Diderot**

Spécialité

**Histoire et Philosophie des Mathématiques**

École doctorale Savoirs scientifiques (ED 400) :
Épistémologie, histoire des sciences, didactique des disciplines

Présentée par

**Cédric VERGNERIE**

---

# La théorie des caractéristiques dans les *Vorlesungen über die Theorie der algebraischen Gleichungen* de Kronecker : la fin du "cycle d'idées sturmiennnes" ?

---

Dirigée par Ivahn SMADJA, Maître de conférences à l'université Denis Diderot

Soutenue le 2 décembre 2017 devant le jury composé de :

| | |
|---|---|
| Mme Hourya BENIS SINACEUR | Directrice de Recherche, IHPST |
| Mme Karine CHEMLA | Directrice de Recherche, SPHERE |
| M. Harold M. EDWARDS | Professeur, NYU, Rapporteur |
| M. Norbert SCHAPPACHER | Professeur, Université de Strasbourg, Rapporteur |
| M. Ivahn SMADJA | Maître de conférences, Université Paris VII, Directeur |
| M. Jean-Jacques SZCZECINIARZ | Professeur, Université Paris VII, Président |

CC BY NC ND

*Beim Rechnen hat man die Regeln, in der Algebra die Gesetze ;*
*es ist das beinahe wie der Schritt von der Kunst zur Wissenschaft.*

*Dans le calcul on a les règles, dans l'algèbre les lois ;*
*c'est presque comme un pas de l'art à la science.*

*Vorlesungen über die Theorie der algebraischen Gleichungen* [Kronecker, 1885b] p. 2

*Man strauchelt leichter auf ebenem Boden, wo man nicht Acht giebt,*
*als auf Bergpfaden, wo die Aufmerksamkeit geschärft ist.*

*On trébuche plus facilement sur un sol égal où l'on ne fait pas attention,*
*que sur un sentier de montagne où l'attention est aiguisée.*

*Vorlesungen über die Theorie der algebraischen Gleichungen* [Kronecker, 1885b] p. 3

**REMERCIEMENTS**

Je tiens tout d'abord à exprimer toute ma gratitude à mon directeur de thèse, Ivahn Smadja, dont les cours de Master ont très largement participé à mon envie de préparer cette thèse. Il a été un directeur attentif, généreux de son temps et partageant toujours avec passion sa grande connaissance de l'histoire des mathématiques.

Je souhaite remercier Norbert Schappacher et Harold M. Edwards d'avoir accepté d'être rapporteurs de cette thèse, ainsi que pour leurs nombreux commentaires qui m'ont aidé à améliorer grandement mon manuscrit. Norbert Schappacher a en particulier fait une relecture minutieuse de mes transcriptions des textes de Kronecker. Qu'il en soit vivement remercié ici. Je remercie également Hourya Benis Sinaceur, Karine Chemla et Jean-Jacques Szczeciniarz pour leur participation au jury.

Merci aux organisateurs du séminaire *Histoire et philosophie des mathématiques* et du groupe de travail *Mathématiques et Philosophie, 19e et 20e siècle*, Marie-José Durand-Richard, Simon Decaens, Emmylou Haffner, Eleonora Sammarchi, Sylvain Cabanacq, et David Rabouin qui m'ont non seulement permis d'exposer à plusieurs reprises les avancées de mes travaux, mais aussi d'être initié à de multiples pans de l'histoire des mathématiques qui ont nourri mes recherches. Merci aux membres de SPHERE, je pense en particulier à Karine Chemla, Jean-Jacques Szczeciniarz, David Rabouin et à Pascal Crozet, qui par leurs encouragements et leurs conseils m'ont aidé à parvenir au bout de ce travail. Un immense merci à Erwan Penchèvre pour sa relecture extrêmement précise et ses nombreuses remarques.

Merci aussi à tous les doctorants avec qui j'ai pu échanger lors de mes (trop rares) venues à Paris, et en particulier à Emmylou Haffner, qui m'a fourni tout au long de ces années de précieux conseils. Un moment important de ma formation a été la semaine passée au CIRM dans le cadre de l'École du GDR. Je souhaiterais ainsi remercier Norbert Schappacher d'avoir eu la patience de m'écouter lui exposer la théorie des caractéristiques de Kronecker lors d'une promenade dans la pinède environnante.

Un grand merci à ma famille pour son soutien, à Gillianne pour son aide dans mes traductions de l'allemand, à mon père pour ses relectures un peu trop « physiciennes », et surtout à Claire pour ses très nombreuses relectures, ses connaissances dans des langues très diverses et pour tout le reste. Enfin merci à mes filles, Suzanne et Luce, qui ont su m'éloigner de mon travail, et à ma mère, qui les en a régulièrement empêchées.

**La théorie des caractéristiques dans les *Vorlesungen über die Theorie der algebraischen Gleichungen* de Kronecker : la fin du "cycle d'idées sturmiennnes" ?**

## Résumé

Hourya Sinaceur présente dans son ouvrage *Corps et Modèles* la théorie des caractéristiques de Kronecker comme la fin d'« un cycle d'idées sturmiennes », la situant ainsi dans une histoire de l'algébrisation du théorème de Sturm. Pourtant, cette théorie est souvent aussi présentée comme le point de départ de certains des premiers travaux de topologie de la toute fin du dix-neuvième siècle. Nous souhaitons dans cette thèse rendre compte de la façon dont ces deux histoires se rencontrent au sein de la théorie de Kronecker. Pour cela, nous disposons d'un matériel particulièrement éclairant et jusqu'alors très peu exploité : les manuscrits des cours que Kronecker a donnés à l'Université de Berlin entre 1872 et 1891 sur la *Théorie des équations algébriques*. Nous commencerons par présenter ces manuscrits, leur contenu et leur contexte de rédaction. Nous nous intéresserons ensuite plus précisément à la reprise du théorème de Sturm par Kronecker, et nous montrerons que la théorie des caractéristiques n'est pas seulement un prolongement algébrique de ce théorème, mais qu'elle se transforme pour fournir certains des outils analytiques que Poincaré utilisera lors de la construction de son *Analysis Situs*. L'exposition de la théorie des caractéristiques dans ses cours est l'occasion pour Kronecker de reprendre trois des quatre démonstrations que Gauss a données du théorème fondamental de l'algèbre, et nous montrerons comment, dans la pratique de Kronecker, la notion même de racine est interrogée.

### Mots-clefs

Algèbre, Kronecker, Sturm, Topologie, histoire des mathématiques, Picard, Weber, théorie des caractéristiques

**The theory of characteristics in the *Vorlesungen über die Theorie der algebraischen Gleichungen* of Kronecker : the end of the "cycle of Sturmian ideas" ?**


## Abstract

In her book *Corps et Modèles*, Hourya Sinaceur presented Kronecker's theory of characteristics as the end of a "cycle of Sturmian ideas", making it a step in history of the algebraization of Sturm's theorem. However, this theory is often also introduced as the starting point of some of the early works of topology of the very end of the nineteenth century. In this PhD thesis, I will describe how these two stories are connected in Kronecker's theory. To achieve this, I used material which has seldom been discussed before : the manuscripts of the courses that Kronecker gave at the University of Berlin between 1872 and 1891 on the *Theory of algebraic equations*. I begin with the presentation of these manuscripts, their contents and their writing contexts. I then look more closely at Kronecker's rework of the theorem of Sturm and show that the theory of characteristics is not only an algebraic extension of this theorem but also that it is transformed in order to provide some of the analytic tools that Poincaré will use for the construction of his *Analysis Situs*. The exposition of the theory of characteristics in his courses is an opportunity for Kronecker to take up three of the four demonstrations from Gauss of the fundamental theorem of algebra, and I will show how, in Kronecker's practice, the very notion of root is questioned.



### Keywords

Algebra, Kronecker, Sturm, Topology, history of mathematics, Picard, Weber, theory of characteristics

# Table des matières

**Introduction générale** **17**

Quel corpus ? Quelle méthodologie ? . . . . . 19

Les manuscrits des cours de Kronecker . . . . . 22

Plan de la thèse . . . . . 25

**I La théorie des équations algébriques de Kronecker** **29**

**1 Introduction** **31**

1.1 L'université de Berlin dans la seconde moitié du XIX^e siècle . . . . . 32

1.1.1 La construction de l'université . . . . . 32

1.1.2 Les mathématiques à l'université de Berlin . . . . . 34

1.1.2.1 L'émergence d'une discipline . . . . . 34

1.1.2.2 Le rapport entre cours et recherche . . . . . 35

1.1.3 Le séminaire de mathématiques à l'université de Berlin . . . . . 36

1.1.3.1 Naissance du séminaire . . . . . 37

1.1.3.2 Organisation des séminaires . . . . . 38

1.2 Présentation des *Vorlesungen* de Kronecker . . . . . 39

1.2.1 Kronecker enseignant . . . . . 39

1.2.1.1 Une publication souhaitée . . . . . 40

1.2.1.2 Des cours difficiles . . . . . 41

1.2.1.3 Les atouts de l'oralité . . . . . 42

1.2.2 Les manuscrits des cours de Kronecker . . . . . 43

1.2.2.1 Comment était organisé le cycle des cours de Kronecker ? . 46

1.2.2.2 Qui a rédigé ces cours ? . . . . . 48

1.2.2.3 Une source crédible . . . . . 49

1.3 Les *Vorlesungen über die Theorie der algebraischen Gleichungen* . . . . . 51

1.3.1 La structure des manuscrits . . . . . 52

1.3.2 Permanence de la structure . . . . . 58

1.3.3 Évolution du contenu . . . . . 58

1.3.4 Une introduction historique . . . . . 59

**2 L'*Algèbre* de Kronecker** **63**

2.1 Quelle Algèbre ? . . . . . 63

2.1.1 L'algèbre avant Kronecker . . . . . 63
2.1.1.1 Lagrange et Vandermonde . . . . . 64
2.1.1.2 Gauss . . . . . 67
2.1.1.3 Dirichlet et Kummer . . . . . 69
2.1.2 L'algèbre dans les *Vorlesungen über die Theorie der algebraischen Gleichungen* . . . . . 70
2.1.2.1 L'algèbre comme formation à l'abstraction . . . . . 71
2.1.2.2 De la *résolution* à la *théorie* des équations . . . . . 71
2.1.2.3 Une algèbre poreuse . . . . . 74
2.1.2.4 L'arithmétisation chez Kronecker . . . . . 75
2.2 Les racines d'une équation . . . . . 78
2.2.1 La notion de racine . . . . . 79
2.2.2 La séparation des racines . . . . . 84
2.2.3 Le théorème de Bolzano . . . . . 91
2.2.3.1 Bolzano et la *Rein analytischer Beweis* . . . . . 93
2.2.3.2 La place de la philosophie dans les mathématiques . . . . . 94
2.2.3.3 Kant, Bolzano et Kronecker . . . . . 97
2.2.3.4 La place de la géométrie . . . . . 97

**3 Présentation et commentaire du cours de Kronecker de l'hiver 1880-1881 sur la théorie des équations algébriques 101**
3.1 Équations quadratiques et cubiques . . . . . 103
3.1.1 Équations quadratiques . . . . . 104
3.1.2 Équation cubique . . . . . 105
3.1.2.1 Expression de $\varphi$, $\psi$ et $\chi$ en fonction des racines $x_0, x_1$ et $x_2$ de l'équation initiale . . . . . 108
3.1.2.2 Expression d'une racine en fonction d'une autre . . . . . 110
3.1.2.3 Le *Postulat d'Abel* . . . . . 113
3.1.3 Conclusion . . . . . 114
3.2 Proposition fondamentale des équations générales et théorie des fonctions symétriques et alternées . . . . . 115
3.2.1 Quelques définitions et premières propriétés . . . . . 116
3.2.1.1 Une équation peut-elle être identiquement nulle ? . . . . . 119
3.2.1.2 Deux théorèmes fondamentaux de la théorie des équations algébriques . . . . . 120
3.2.2 Fonctions symétriques . . . . . 122
3.2.2.1 Définitions et polynômes symétriques élémentaires . . . . . 123
3.2.2.2 Construction et utilisations des fonctions $F_k$ . . . . . 124
3.2.2.3 Un cas particulier : étude de la fonction $F = \prod_{(i)}(z - \psi(x_{i_1}, ..., x_{i_n}))$ . . . . . 131
3.2.2.4 Un exemple de réduction . . . . . 134

3.2.2.5 Les « Types » de Vandermonde . . . 137
3.2.2.6 Théorème fondamental des polynômes symétriques . . . 138
3.3 Sur la divisibilité des fonctions entières . . . 144
3.3.1 Résultant et Discriminant . . . 144
3.3.1.1 Conditions pour avoir un facteur commun . . . 144
3.3.1.2 Décomposition en fractions continues . . . 146
3.3.2 Décomposition d'une fonction entière en facteurs . . . 147
3.3.2.1 Propriétés préliminaires . . . 147
3.3.2.2 Méthode de décomposition . . . 148
3.4 Conclusion . . . 149

**4 Les *Vorlesungen* de Kronecker et le *Lehrbuch* de Heinrich Weber 151**
4.1 Weber lecteur de Kronecker . . . 152
4.2 Présentation du *Lehrbuch der Algebra* . . . 155
4.2.1 Le *Lehrbuch der Algebra* . . . 155
4.2.2 Le premier Volume . . . 156
4.2.2.1 Le contenu . . . 156
4.2.2.2 La préface . . . 156
4.3 Comparaison . . . 157
4.3.1 La structure . . . 157
4.3.2 Des *Algèbres* différentes ? . . . 159
4.3.3 Les deux introductions . . . 160
4.4 La notion de Racine . . . 160
4.4.1 Introduction . . . 160
4.4.1.1 La notion de Racine chez Weber . . . 161
4.4.1.2 La notion de Racine chez Kronecker . . . 161
4.4.2 Le théorème fondamental de l'algèbre . . . 162
4.5 Un exemple : le théorème de Sturm . . . 163
4.5.1 Introduction . . . 163
4.5.2 Les présentations de Weber et Kronecker . . . 164
4.5.2.1 La présentation de Weber . . . 164
4.5.2.2 La présentation de Kronecker . . . 164
4.5.3 Une extension : la théorie des caractéristiques . . . 166
4.5.3.1 Présentation (très) rapide de cette théorie . . . 166
4.5.3.2 Chez Weber . . . 166
4.5.3.3 Chez Kronecker . . . 167
4.6 Conclusion . . . 167

**5 Conclusion 169**

**II Le théorème de Sturm dans les *Vorlesungen über die Theorie der algebraischen Gleichungen* 173**

**6 Introduction 175**

6.1 Méthodologie . . . 177
6.2 Le théorème de Sturm dans les *Vorlesungen über die Theorie der algebraischen Gleichungen* . . . 179
6.2.1 Les motivations de Kronecker . . . 179
6.2.2 Le théorème de Sturm : une place centrale dans les *Vorlesungen* . . . 180
6.2.3 Contenu du chapitre sur le théorème de Sturm . . . 182
6.2.3.1 Le théorème dans chacun des manuscrits . . . 183
6.2.3.2 Conclusion . . . 188
6.3 Énoncé et premier exemple . . . 189
6.4 Réception et diffusion du théorème de Sturm . . . 191

**7 Démonstration du théorème par Sturm 193**

7.1 La démonstration de 1835 . . . 193
7.2 Première démonstration du théorème de Sturm par Kronecker . . . 199
7.2.1 Introduction . . . 199
7.2.1.1 Une vision *géométrique* . . . 199
7.2.1.2 Racines, continuité géométrique et continuité arithmétique 201
7.2.2 Une démonstration sans considération de continuité ? . . . 204
7.2.2.1 Démonstration du théorème de Sturm . . . 205
7.2.2.2 Analogie avec la théorie de Cauchy . . . 208
7.2.2.3 Une première application de l'algorithme d'Euclide . . . 208
7.2.2.4 Un exemple . . . 212
7.2.2.5 Une seconde démonstration . . . 214
7.2.2.6 Interprétation géométrique de $\sum_{x_1<\xi<x_2} sgn\,(f'(\xi)f_1(\xi))$ . . 216
7.2.3 Recherche du nombre total de solutions d'une fonction entière . . . 219
7.2.4 Un exemple : l'équation biquadratique . . . 221
7.2.4.1 Application du théorème de Sturm . . . 221
7.2.4.2 Étude de la surface-discriminante (*Discriminantenfläche*) . 224
7.3 Conclusion . . . 230

**8 Sylvester et les fractions continues 233**

8.1 Les premiers travaux sur les suites de Sturm : 1839-1841 . . . 234
8.1.1 De nouvelles fonctions auxiliaires . . . 235
8.2 Démonstration par Sturm de la proposition de Sylvester . . . 237
8.2.1 Expression de $V_2, V_3, ...$ en fonction de $x$ et des racines de $V = 0$ . . 239
8.2.2 Cauchy . . . 240
8.2.3 Expression des $\lambda$ . . . 241
8.3 Retour à la note de 1841. . . . 243

8.4 *Digression* de Kronecker sur les développements en fractions continues . . . 244
8.4.1 Expression de $N(g_1...g_\nu)$ sous la forme d'un déterminant . . . . . 249
8.4.2 Conclusion . . . . . 250

**9 Le problème d'interpolation de Cauchy 253**
9.1 Introduction . . . . . 253
9.1.1 Cinq méthodes . . . . . 254
9.1.2 Exposé du problème . . . . . 254
9.1.2.1 Le problème . . . . . 254
9.1.2.2 Le rapport au théorème de Sturm . . . . . 255
9.1.3 Premières méthodes . . . . . 256
9.1.3.1 La méthode de Cauchy . . . . . 256
9.1.3.2 Méthode attribuée à Richard Baltzer . . . . . 256
9.2 Méthode de Liouville . . . . . 257
9.2.1 Description de la méthode . . . . . 258
9.2.2 Ensemble des solutions du problème de Cauchy . . . . . 259
9.3 Cayley et la Résolution du problème de Cauchy . . . . . 264
9.3.1 Les recherches de Cayley (1821-1895) sur les fonctions de Sturm . . . 264
9.3.1.1 L'article de 1846 . . . . . 264
9.3.1.2 L'article de 1848 . . . . . 269
9.3.2 La résolution par Kronecker . . . . . 272
9.3.2.1 Un premier point de vue . . . . . 272
9.3.2.2 Méthode de Jacobi . . . . . 274
9.3.2.3 Application au Théorème de Sturm . . . . . 284
9.3.2.4 Les expressions de Sylvester . . . . . 287
9.4 Dernière résolution du problème de Cauchy . . . . . 293
9.4.1 Redéfinition du problème . . . . . 293
9.4.2 Autre forme de $\mathcal{F}(x)$ . . . . . 297
9.4.3 Autre façon de mettre en évidence la relation qui unit $\mathcal{F}, \Psi$ et $\Phi$. . . 299
9.4.4 Quelques considérations sur la dimension de l'ensemble des solutions 301

**10 La méthode d'Hermite-Jacobi 305**
10.1 Le principe d'inertie de Sylvester . . . . . 307
10.2 Application de la loi d'inertie au théorème de Sturm . . . . . 310

**11 Conclusion 315**

**III La théorie des caractéristiques 319**

**12 Introduction 321**
12.1 Les sources . . . . . 323
12.2 Contenu des manuscrits . . . . . 325

12.3 Une première exposition de la théorie . . . 328
12.3.1 Remarques préliminaires . . . 329
12.3.2 Intersection de $\varphi$ et $\psi$ . . . 331
12.3.3 Caractéristique d'un système de trois fonctions . . . 332
12.3.4 Points d'intersection de deux courbes du plan dans un compact . . . 335
12.3.4.1 Rapport entre la notation pour deux courbes et celle pour trois courbes . . . 335
12.3.4.2 Caractère d'un point . . . 336
12.3.4.3 Extension à des cas plus généraux . . . 336
12.3.5 Racines complexes . . . 338
12.3.6 Détermination de la caractéristique . . . 341
12.4 Conclusion . . . 342

**13 La théorie des caractéristiques selon Kronecker 343**
13.1 Considérations introductives . . . 345
13.2 La caractéristique de deux fonctions d'une variable . . . 347
13.2.1 Introduction . . . 347
13.2.2 Définition de la caractéristique de deux fonctions . . . 348
13.2.2.1 Définition . . . 349
13.2.2.2 Propriétés . . . 350
13.2.2.3 Retour sur les Hypothèses . . . 351
13.2.3 Interprétation graphique . . . 351
13.2.4 Lien avec le théorème de Sturm . . . 352
13.2.5 Modification de la caractéristique en fonction des coefficients de $\varphi$ et $\psi$. . . . 353
13.2.5.1 Cas particulier ou cas général ? . . . 353
13.2.5.2 À propos du résultant de deux fonctions . . . 355
13.2.5.3 Modification de la caractéristique . . . 356
13.2.5.4 Application au discriminant . . . 358
13.2.5.5 Conclusion . . . 359
13.3 Le *Fortgangsprincip* . . . 359
13.3.1 Introduction . . . 359
13.3.1.1 Dans quel cadre ? En quoi consiste-il ? . . . 359
13.3.1.2 Un concept « fondamental », mais peu présent . . . 360
13.3.1.3 Les raisons de l'introduction de ce principe . . . 361
13.3.2 Explication du principe . . . 362
13.3.2.1 Dans l'article de 1869 . . . 362
13.3.2.2 Dans son cours : le cas $n = 2$ . . . 364
13.3.2.3 Le cas général . . . 369
13.3.3 Conclusion . . . 372
13.3.3.1 Postérité . . . 373

13.3.3.2 Pourquoi avoir introduit ce principe ? . . . . . 374
13.4 La caractéristique de trois fonctions à deux variables . . . . . 375
13.4.1 (*Fortgangsprincip*) : un résumé pratique . . . . . 375
13.4.2 Une proposition problématique . . . . . 376
13.4.3 Remarques à propos de l'argument utilisé par Kronecker . . . . . 377
13.4.4 Caractéristique de trois fonctions de deux variables . . . . . 380
13.4.4.1 Définition de la caractéristique de 3 fonctions de 2 variables 383
13.4.4.2 Interprétations graphique . . . . . 383
13.4.4.3 Extension aux courbes qui ne sont pas fermées . . . . . 385
13.4.4.4 Modification de la caractéristique . . . . . 387
13.5 Retour sur la notion de caractéristique . . . . . 388
13.5.1 La notion de racine d'une équation . . . . . 388
13.5.2 La notion de caractéristique . . . . . 390
13.5.2.1 *Rationalisation* du problème . . . . . 390
13.5.2.2 Un exemple . . . . . 392
13.6 Conclusion . . . . . 393

**14 Le passage d'une théorie *topologico-arithmétique* à une théorie *topologico-analytique* 395**
14.1 Le cours de Kronecker sur l'intégration . . . . . 397
14.1.1 Définitions des intégrales . . . . . 398
14.1.2 Le Grand et le Petit . . . . . 400
14.1.3 Interprétation géométrique . . . . . 401
14.1.4 La notion de limite . . . . . 402
14.1.5 Sur la continuité . . . . . 404
14.1.6 Kronecker et les *théories modernes de l'intégration* . . . . . 405
14.2 L'indice de Cauchy . . . . . 406
14.2.1 Présentation de la notion d'indice par Cauchy . . . . . 407
14.2.2 Les *Comptes rendus* de 1837 . . . . . 409
14.2.3 Une volonté d'éliminer le calcul intégral . . . . . 410
14.3 Le passage d'une théorie topologico-algébrique à une théorie topologico-analytique . . . . . 411
14.3.1 Expression analytique . . . . . 414
14.3.2 Intégrale de Cauchy . . . . . 416
14.3.3 Étude de l'intégrale $\frac{1}{2\pi}\int_{f(x,y)=0} d \operatorname{arctg} \frac{\psi}{\varphi}$ . . . . . 416
14.4 La caractéristique de $n+1$ fonctions à $n$ variables . . . . . 422
14.4.1 Rappels et notations . . . . . 422
14.4.2 Extension *topologico-algébrique* de la notion de caractéristique . . . . 423
14.4.3 Extension *topologico-analytique* de la notion de caractéristique . . . . 424
14.5 La controverse entre Picard et Kronecker . . . . . 426
14.5.1 Les premières publications de Picard . . . . . 428

14.5.1.1 L'article de 1889 et le traité de 1891 . . . . . 428
14.5.1.2 Les réserves de Picard . . . . . 429
14.5.2 Les *Comptes rendus* du second semestre de 1891 : l'objet de la controverse . . . . . 430
14.5.2.1 L'article du 7 septembre 1891 . . . . . 432
14.5.2.2 L'article du 16 novembre 1891 . . . . . 432
14.5.2.3 La réponse de Kronecker . . . . . 434
14.5.3 Une résolution posthume . . . . . 437
14.5.3.1 L'article de 1892 du *Journal de Liouville* . . . . . 437
14.5.3.2 Les volumes du *Traité d'analyse* . . . . . 438
14.5.3.3 L'intervention de Van Dyck . . . . . 440
14.5.4 Conclusion . . . . . 442

**15 Le théorème fondamental de l'algèbre 445**
15.1 Introduction . . . . . 445
15.1.1 Les quatre preuves de Gauss . . . . . 446
15.1.2 La seconde preuve de Gauss . . . . . 447
15.1.2.1 La démonstration de Gauss . . . . . 448
15.1.2.2 Dans les leçons de Kronecker . . . . . 453
15.2 Les première et quatrième démonstrations . . . . . 454
15.2.1 La première démonstration de Gauss de 1799 . . . . . 455
15.2.1.1 La première partie . . . . . 456
15.2.1.2 La seconde partie . . . . . 458
15.2.2 La quatrième démonstration de Gauss de 1849 . . . . . 468
15.2.3 La reprise par Kronecker . . . . . 478
15.2.3.1 Les coefficients du polynôme . . . . . 479
15.2.3.2 Géométrisation du problème . . . . . 480
15.2.3.3 Circonscription du problème . . . . . 481
15.3 La troisième démonstration . . . . . 483
15.3.1 La version de Gauss . . . . . 483
15.3.1.1 Notation . . . . . 484
15.3.1.2 Restriction à un intervalle . . . . . 484
15.3.1.3 Plusieurs calculs pour une même intégrale . . . . . 487
15.3.1.4 Démonstration du TFA . . . . . 489
15.3.2 La reprise par Kronecker . . . . . 491
15.3.2.1 Un outil commun . . . . . 491
15.3.2.2 Démonstration de Kronecker . . . . . 492
15.3.2.3 Conclusion . . . . . 493
15.4 Vers un nouvel énoncé . . . . . 493

**Conclusion générale** **499**
De l'« algébrisation » à la « généralisation » . . . . . 500
Postérité de la théorie . . . . . 504

**Annexes** **509**

**A Transcription et traduction partielles du cours de 1880-1881** **511**
A.1 Transcription de l'introduction . . . . . 513
A.2 Traduction de l'introduction . . . . . 521
A.3 Transcription de la seconde partie du cours . . . . . 529
A.4 Traduction de la seconde partie du cours . . . . . 597

**B Transcription et traduction du chapitre du cours de 1890-1891 sur la théorie des caractéristiques** **661**
B.1 Sommaire du cours de 1890-1891 . . . . . 663
B.2 Transcription . . . . . 668
B.3 Traduction . . . . . 714

**C Traduction des troisième et quatrième démonstrations de Gauss du théorème fondamental de l'algèbre** **759**
C.1 Traduction de la troisième démonstration de Gauss (1816) . . . . . 760
C.2 Traduction de la quatrième démonstration de Gauss (1849) . . . . . 766

**D Réseau de citation** **775**

**Bibliographie** **781**

# Introduction générale

Dans son histoire du théorème de Sturm, Hourya Sinaceur remarque que :

> Du côté de la recherche, [*le théorème de Sturm*] ne resta pas (...) lettre morte. Il fut suivi, en effet, par toute une série de mémoires ou d'articles visant à l'améliorer ou le transformer de la plume des plus grands mathématiciens comme Sylvester, Cayley, Hermite ou Kronecker, et aussi de moins grands comme Borchardt ou Darboux. Il engendra ainsi ce que Sylvester a appelé "un cycle d'idées sturmiennes". Ce cycle s'acheva pourtant, sinon avec les travaux de Sylvester comme celui-ci voulait bien le croire, du moins avec la difficile théorie des caractéristiques de Kronecker. ([Sinaceur, 1988], p. 103.)

Ainsi, la théorie des caractéristiques de Kronecker clôturerait le « cycle d'idées sturmiennes » [1], dont le cœur témoigne d'une volonté d'algébrisation du théorème de Sturm.

Par ailleurs, Jean-Claude Pont, dans son histoire de la topologie algébrique [2], montre que « les *Contributions* [3] de Dyck jouent le rôle de plaque tournante » dans l'histoire de la topologie et que ses travaux partent de ceux de « Mobius, Klein et Kronecker », et sont en particulier liés à « la caractéristique de Kronecker » [4]. De même, F.R. Brown explique dans [Brown, 1999] que

> Poincaré based the proof of the generalized form of the intermediate value theorem on the Kronecker index, introduced in 1869 [5]. The purpose of Hadamard's appendix to Tannery's text [6] was to demonstrate the usefulness of the same tool. Brouwer's proof instead used the notion of topological degree, which he introduced as a replacement for the Kronecker index that would not require the map to be differentiable. ([Brown, 1999] p. 274)

Dans le même recueil d'articles, Erhard Scholz ajoute que :

> Working with the Kronecker characteristic as a symbolical tool, Poincaré was able to sketch the idea of a high dimensional version of the index theorem for vector fields, the later Poincaré-Hopf index theorem, for the case of open submanifolds or hypersurfaces in $\mathbb{R}^n$. ([Scholz, 1999] pp. 40-41)

Et dans son célèbre ouvrage *Differential Topology*, Hirsch affirme enfin :

> The origin of the notion of the degree is Kronecker's characteristic of a "system of functions" defined in 1869. [Hirsch, 2012] p. 140

La théorie des caractéristiques de Kronecker semble donc être présentée à la fois comme l'aboutissement d'un mouvement d'algébrisation, et comme le point de départ de certains des premiers travaux de topologie de la toute fin du XIX$^{e}$ siècle. Si le travail de Kronecker fait très certainement partie du « cycle » dont parle H. Sinaceur, en est-il réellement l'« aboutissement » ? Quel changement de point de vue permet à Kronecker de produire

1. Hourya Sinaceur reprend ici une expression utilisée par Sylvester dans [Sylvester, 1853d] p. 486 : « the cycle of Sturmian ideas ».
2. Voir [Pont, 1974].
3. Voir [Dyck, 1888] et [Dyck, 1890].
4. [Pont, 1974] p. 152.
5. Voir [Kronecker, 1869c] et [Kronecker, 1869d].
6. Voir [Tannery et Hadamard, 1910].

un des résultats sur lesquels Poincaré bâtira son *Analysis situs* ? Comment ces deux histoires se rencontrent-elles au sein de cette théorie, et comment s'y articulent les différentes influences, les différentes traditions à l'œuvre ?

## Quel corpus ? Quelle méthodologie ?

La théorie des caractéristiques, qui est exposée par Kronecker pour la première fois en 1869 dans un essai intitulé *Über Systeme von Functionen mehrer Variabeln*, et qui est lu lors de deux séances à l'Académie de Berlin [7], semble donc s'étendre à travers une grande partie des mathématiques du XIX^e^ siècle. Un second texte, *Über die Charakteristik von Funktionen-Systemen*, publié en 1878, complète sa théorie. Pour commencer notre étude des sources et de l'origine de ce travail de Kronecker, nous avons dans un premier temps tenté de constituer un corpus. Devions-nous nous concentrer sur les ouvrages ayant trait à la « théorie des fonctions » ou à l'« algèbre » pour reprendre la classification du *Jahrbuch* [8] des articles de 1869 et 1878 ? Ou au contraire utiliser le classement du XX^e^ siècle [9] de ce même *Jahrbuch* et étudier ces travaux dans le cadre de la « géométrie différentielle classique » ? En effet, si l'article de 1869 est classé par le *Jahrbuch* dans la section *théorie des fonctions*, celui de 1878, qui prolonge pourtant le précédent, se trouve dans la section *algèbre* [10]. Bien sûr, le choix des *reviewers* dépend très certainement du contenu des articles, mais on peut tout de même remarquer que ce classement a été effectué en 1869 par Wangerin (1844-1933), qui a essentiellement travaillé sur les fonctions harmoniques et sphériques, et celui de 1878 par Netto (1848-1919), dont les recherches ont porté en grande partie sur la théorie des groupes [11].

Des études récentes [12] ont montré les difficultés liées à une telle approche, voire, dans de nombreux cas, le manque de pertinence de l'utilisation exclusive d'une classification historiquement située. Nous avons donc plutôt cherché à constituer, à partir des deux articles où Kronecker élabore cette théorie des caractéristiques, un réseau de références bibliographiques. En dépit de la base étroite dont nous sommes partis, nous avons recensé l'ensemble des références bibliographiques explicites de son article de 1878, et à partir des textes obtenus, nous avons de nouveau recensé les références, et ainsi de suite jusqu'à « épuisement ».

7. [Kronecker, 1869c] en mars et [Kronecker, 1869d] en août.

8. Le *Jahrbuch über die Fortschritte der Mathematik* a été édité, entre autres, par Carl Ohrtmann, Felix Müller, Albert Wangerin, etc. Cette revue, qui a été publiée entre 1868 et 1944, donnait des bibliographies annuelles des publications mathématiques.

9. Mathematics Subject Classification (MSC) 2000.

10. Pour des précisions sur le classement à l'œuvre dans le *Jahrbuch über die Fortschritte der Mathematik*, on pourra lire par exemple [Rollet et Nabonnand, 2002].

11. L'article de 1869 est ensuite classé en « géométrie différentielle classique » avec le passage à la classification (MSC) 2000.

12. Voir par exemple le travail de Frédéric Brechenmacher [Brechenmacher, 2006], celui de Catherine Goldstein [Goldstein, 1999] ou celui de Sébastien Gauthier sur la notion de discipline [Gauthier, 2007].

Cependant, nous avons choisi de borner cet « épuisement » [13]. Tout d'abord, Kronecker affirme, à propos de sa théorie des caractéristiques, que « dans les recherches dont [il a] esquissé ici le développement, [il est] parti du théorème de *Sturm* » [14]. Ainsi, toute référence antérieure à 1835, date à laquelle Sturm publie la démonstration de son théorème, sera prise en compte mais ne sera pas poursuivie. Ensuite, lorsqu'un ouvrage cité n'est pas publié dans un périodique, il est la plupart du temps trop vaste et contient des sujets trop différents pour que cela ait un sens d'en extraire les références bibliographiques. Enfin, cette méthode n'est plus pertinente si on réitère la méthode un trop grand nombre de fois : une valeur du diamètre du graphe obtenu à partir de notre réseau supérieure à huit n'apporte plus de renseignement exploitable. Notre but étant ici de travailler sur le contexte de l'élaboration par Kronecker de cette théorie, le choix de commencer notre réseau à partir de l'article de 1878 ([Kronecker, 1878b]) nous fixe une période s'étendant donc de 1835 à 1878.

Il ne faudrait pas pour autant faire de cette méthode de constitution d'un corpus une panacée [15]. Tout d'abord, elle pose techniquement quelques problèmes : les références bibliographiques sont, à cette époque, souvent implicites, quelquefois imprécises ou erronées, ou même passées sous silence. Nous nous autoriserons ainsi, au fur et à mesure de notre travail, à élargir ou à restreindre ce corpus si cela s'avère nécessaire. Ensuite, cette méthode, et surtout son analyse à l'aide de méthodes quantitatives et d'outils spécifiques est efficace pour des corpus de « grande taille ». Or notre corpus est bien trop « petit » [16] pour entrer dans cette catégorie.

Il ne s'agira donc pas ici d'avoir une approche quantitative comme ont pu le faire par exemple Hélène Gispert pour l'étude des sociétés savantes [17] ou Catherine Goldstein sur la théorie des nombres [18], mais cette méthode aura finalement ici un rôle essentiellement heuristique : la représentation graphique du réseau ainsi constitué permettra de repérer plusieurs groupements de textes qu'il faudra ensuite parcourir pour tenter de les caractériser. Ce travail de repérage rendra ensuite possible une analyse plus approfondie des motivations de Kronecker qui l'ont conduit à construire sa théorie.

Par ailleurs, la quasi totalité de notre corpus est antérieure au recensement du *Jahrbuch*, et nous ne pourrons donc utiliser ni sa classification, ni sa liste de références bibliographiques, comme le font par exemple F. Brechenmacher dans [Brechenmacher, 2011] et C. Goldstein dans [Goldstein, 1999]. Le tome du *Catalogue of scientific papers* [19] de

13. Dans son travail sur les *inégalités séculaires*, F. Brechenmacher semble obtenir un réseau fermé. Il ne nous a pas semblé possible, ni souhaitable, de développer notre réseau de citations au delà des critères de cloisonnement que nous allons définir.

14. [Kronecker, 1869c] p. 205 : « Bei den Untersuchungen, deren Entwickelungsgang ich hier in Umrissen dargelegt habe, bin ich vom *Sturm*'schen Satze ausgegangen. »

15. On pourra par exemple lire [Brechenmacher, 2012].

16. 113 nœuds et 154 liens.

17. Voir [Gispert, 1991].

18. Voir [Goldstein, 1999].

19. 19 tomes de ce *Catalogue* ont été publiés entre 1867 et 1925 par la *Royal Society of London*. Ils se proposaient de recenser l'ensemble des articles scientifiques publiés entre 1800 et 1900. Trois tomes thématiques, dont un sur les *mathématiques pures* ont été publiés à partir de 1908.

1908 donne un classement thématique d'un grand nombre d'articles de mathématiques publiés entre 1800 et 1900, mais reste très incomplet : il ne serait pas possible de classifier notre corpus à partir de celui-ci. Remarquons tout de même qu'une entrée de ce *Catalogue* est consacrée à la *théorie des caractéristiques*. Mais on trouve dans celle-ci des objets aussi différents que la caractéristique au sens de Kronecker et celle au sens de Chasles : l'auteur du classement s'est laissé tromper par un vocabulaire commun qui désigne pourtant des objets extrêmement différents [20]. Plutôt que de fabriquer des catégories artificielles, nous avons donc choisi de ne tenir compte d'aucune classification, quelle qu'elle puisse être.

Après avoir procédé ainsi, on obtient alors une représentation graphique de notre réseau [21]. Dans ces deux figures, obtenues à l'aide du logiciel *Gephi* [22], le principe de représentation est le suivant : un nœud est d'autant plus grand que le nombre d'articles qui s'y réfèrent est important. Ainsi par exemple, le nœud correspondant à l'article de Kronecker de 1878 est trop petit pour apparaître dans ces présentations, aucun autre article ne pouvant le citer.

On voit apparaître sur le premier des deux graphiques six grands ensembles :

— Le premier gravite autour des travaux de Sylvester et de Cayley.

— Le second porte sur un article de Kronecker de 1865 dans lequel il généralise la formule d'interpolation de Lagrange à des polynômes à plusieurs indéterminées [23] et en lien avec un article de Jacobi de 1835.

— Le troisième a pour centre le *Mémoire sur l'interpolation* de Cauchy de 1835.

— Le cours de Cauchy est la source d'un sous-ensemble qui traite de son indice.

— Le cinquième concerne la théorie du potentiel traitée par Gauss.

— Le dernier est en lien avec deux articles d'Hermite qui prolongent le travail de Sturm.

Le second graphique, qui accentue l'importance des articles les plus référencés en leur conférant une grande « force d'attraction », accorde une place centrale à l'article de Sylvester de 1839, qui semble « écraser » le reste des références. Si le travail algébrique de Sylvester, auteur, rappelons-le, de la formule « cycle d'idées sturmiennes », est bien présent, il s'accompagne cependant d'articles aux sujets très divers, touchant la théorie du potentiel, l'analyse complexe et des problèmes d'interpolation. Comme nous pouvions le soupçonner, les sources de la théorie des caractéristiques semblent bien multiples et nous aurons besoin d'un fil directeur qui puisse nous permettre de circuler à travers ce réseau.

20. Voir [Society, 1908] p. 571.

21. Voir les figures 1 et 2. Un tableau récapitulatif de l'ensemble des textes utilisés pour construire ces graphiques est donné en annexe p. 775.

22. Gephi est un logiciel libre d'analyse et de visualisation de réseaux, développé en Java.

23. [Kronecker, 1865] et [Jacobi, 1835]. On pourra lire pour plus de détails sur ces articles [Penchèvre, 2006] p. 255.

## Les manuscrits des cours de Kronecker

Car, ce qui nous intéresse ici finalement, c'est la lecture que Kronecker fait de ces textes lorsqu'il élabore sa théorie. Comme nous l'avons dit, celle-ci est développée dans deux articles [24], et on n'en trouve que quelques rares utilisations dans le reste de l'œuvre publiée de Kronecker [25]. C'est dans d'autres sources, encore inexploitées, que nous allons chercher des traces de ces lectures.

Nous disposons en effet de manuscrits contenant les notes de cours prises par des élèves de Kronecker entre 1872 et 1891. Ces manuscrits sont presque en totalité issus du *fonds Hensel* qui se trouve à la bibliothèque de Strasbourg, et quelques-uns ont fait l'objet d'une édition [26], mais la série de manuscrits intitulés *Theorie der algebraichen Gleichungen* n'a été ni transcrite ni éditée [27].

Dans ses *Éléments d'analyse de Karl Weierstrass* [28], Pierre Dugac avait utilisé les cours que ce dernier avait donnés à Berlin durant la même période pour en « dégager quelques notions essentielles », « en chercher l'origine » et en « suivre leur évolution et leurs transformations » [29]. Il y montrait l'usage pertinent que l'on peut faire de ces leçons pour l'étude de la conception des mathématiques de Weierstrass. Nous nous sommes proposé de procéder de la même manière avec les cours de Kronecker.

Dans une lettre à son professeur Hjalmar Holmgren (1822-1885), Gösta Mittag-Leffler [30] décrit les enseignements qu'il a suivis à Berlin [31] :

> D'un point de vue scientifique, je suis très satisfait de mon séjour à Berlin. Nulle

24. Celui de 1869 en deux parties : [Kronecker, 1869c], [Kronecker, 1869d] et celui de 1878 [Kronecker, 1878b].

25. [Kronecker, 1873b], [Kronecker, 1878c], [Kronecker, 1889c] et [Kronecker, 1891b].

26. Hensel, chargé du *Nachlass* de Kronecker, a publié ses cours sur l'arithmétique et les déterminants ([Kronecker, 1901] et [Kronecker, 1903]). Netto, ancien élève de Kronecker, a de son côté dirigé la publication de son cours sur les intégrales ([Kronecker, 1894]).

27. ll y a dix manuscrits accessibles de cette « théorie des équations algébriques », qui correspondent à sept sessions différentes : [Kronecker, 1873a] 72/73, [Kronecker, 1879] 78/79, [Kronecker, 1881a] 80/81, [Kronecker, 1885b], [Kronecker, 1885a] et [Kronecker, 1885c] 84/85, [Kronecker, 1887a] et [Kronecker, 1887b] 86/87, [Kronecker, 1889b] 88/89, [Kronecker, 1891a] 90/91.

28. Voir [Dugac, 1973].

29. [Dugac, 1973] p. 43.

30. Gösta Mittag-Leffler (1846-1927), qui créa et dirigea les célèbres *Acta Mathematica*, obtient sa thèse, dont le sujet a trait à l'analyse complexe, en 1872 à Upsalla. Ce n'est qu'ensuite, en 1875, après avoir passé quelques temps à Paris, qu'il décide, sur les conseils d'Hermite, de suivre les cours de Weierstraß à Berlin.

31. Lettre du 16 février 1875 traduite en allemand du suédois dans [Weierstrass *et al.*, 1966] p. 54 : « Mit meinen Aufenthalt in Berlin bin ich in wissenschaftlicher Hinsicht sehr zufrieden. Nirgends habe ich so vieles zu lernen gefunden wie hier. Weierstraß und Kronecker haben beide die in Deutschland ungewöhnliche Eigenschaft, gedruckte Publikationen, soweit es möglich ist, zu vermeiden. Weierstraß publiziert bekanntlich gar nicht, und Kronecker nur Resultate ohne Beweise.

In ihren Vorlesungen legen sie die Resultate ihrer Forschungen vor. Kaum dürfte wohl die Mathematik unserer Tage etwas aufzuweisen haben, was mit der Funktionentheorie von Weierstraß oder der Algebra von Kronecker wetteifern kann. (...)

Sowohl Weierstraß als auch Kronecker zeichnen sich übrigens durch vollständigste Klarheit und Schärfe beim Beweisen aus. Zugleich haben sie von Gauss die Furcht vor aller Art Metaphysik bei der Fixierung der mathematischen Grundbegriffe geerbt, und dies gibt ihren Deduktionen eine Einfachheit und Natürlichkeit, die man wohl kaum früher so systematisch ausgeführt und mit dem höchsten Grad an Schärfe gesehen hat. »

> part je n'ai trouvé autant de choses à apprendre qu'ici. Weierstraß et Kronecker ont tous deux la particularité inhabituelle en Allemagne d'éviter autant que possible les publications imprimées. Il est connu que Weierstraß ne publie presque rien, et Kronecker seulement des résultats sans démonstration.
>
> Les résultats de leurs recherches sont présentés dans leurs leçons. Les mathématiques de notre temps n'ont pas grand-chose qui puisse rivaliser avec la théorie des fonctions de Weierstraß ou l'algèbre de Kronecker. (...)
>
> Autant Weierstraß que Kronecker s'illustrent d'ailleurs par la plus complète clarté et la précision de leurs démonstrations. Ils ont tout deux hérité de Gauss la crainte de tout genre de métaphysique dans la fixation des concepts fondamentaux des mathématiques, et cela donne à leurs démonstrations une simplicité et un naturel que l'on a rarement vus conduits de façon aussi systématique avec un si grand degré de précision. ([Weierstrass *et al.*, 1966] p. 54)

Mittag-Leffler met l'accent sur la nécessité d'avoir accès aux cours de Weierstrass et de Kronecker pour comprendre le contenu de leurs recherches, car leurs articles publiés sont rares ou ne comportent qu'une suite de résultats sans démonstration. L'un des cours que cite Mittag-Leffler dans la suite de cette lettre est d'ailleurs celui dans lequel Kronecker expose sa théorie des caractéristiques, objet de notre étude, et qu'il professera régulièrement tout au long de sa carrière à l'université de Berlin sous l'intitulé : *Theorie der Algebraischen Gleichungen.*

Si ces cours se caractérisent, pour Mittag-Leffler, par une grande clarté, en revanche Harold Edwards écrit dans un article où il cherche pourtant à réhabiliter la place de Kronecker dans l'histoire des mathématiques :

> Surely one of the principal reasons Kronecker is so little studied today (except, it seems to me, by the best mathematicians) is that his works are so very difficult to read. Jordan very aptly called them in 1870 « l'envie et le désespoir des géomètres. » This shows that the difficulty we experience today in reading Kronecker's works is not merely the difficulty of reading an old text written in an outdated terminology. Kronecker's difficult style was difficult for his contemporaries, too. I have been working at it for many years, especially his Kummer Festschrift entitled "Grundzüge einer arithmetischen Theorie der algebraischen Grössen". I feel my efforts have been rewarded and that I will soon be able to publish papers that will convey something of what Kronecker was doing and that will, I hope, awaken the interest of others in studying Kronecker. ([Edwards, 1988] p. 143)

Disons-le, les travaux d'H. Edwards sur Kronecker ont en effet « éveillé notre intérêt pour l'étude de Kronecker ». Mais les difficultés [32] dont il parle en 1988 restent en partie présentes dans ses *Essays in Constructive Mathematics* [33], publiés en 2005, et confirment ce qu'il avait déjà écrit dans un article de 1987 lorsqu'il rappelait que Frobenius affirmait qu'à peine trois personnes comprenaient certains travaux de Kronecker [34]. H. Edwards ajoutait alors :

32. En particulier la notion de *Modul System*, qui devient très importante dans le travail algébrique de Kronecker à partir de la publication des *Grundzüge*. Voir [Edwards, 2005] p. 5.
33. [Edwards, 2005].
34. Voir [Frobenius, 1893] p. 8 : « Gab es unter seinen Fachgenossen kaum drei, die dem Fluge seiner

> My impression is that this estimate is generous. I believe there are important passages in Kronecker's work that *no one, ever* has fully understood, other than Kronecker himself. ([Edwards, 1987] p. 29)

Trouvant « le style de présentation de Kronecker dans ses papiers publiés obscur et sans motivation » [35], H. Edwards a recherché son *Nachlass* [36], espérant trouver dans celui-ci quelques éclaircissements. Malheureusement, comme Pierre Dugac avant lui [37], H. Edwards ne parvient pas à localiser ce *Nachlass*.

Le fin mot de cette énigme se trouve finalement exposé dans un article de 1978 intitulé *On the Kronecker Nachlass*, où H. Edwards retrace les péripéties de sa quête du *Nachlass* perdu. On sait que la tâche de s'occuper du *Nachlass* de Kronecker échut à Kurt Hensel, ancien élève de Kronecker [38]. Celui-ci, *Privatdozent* à l'université de Berlin après y avoir soutenu sa thèse, enseigna à partir de 1901 à l'Université de Marburg.

Dans un premier temps, grâce à Helmut Hasse, un ancien élève de Hensel, et Walter Hayman, petit fils de Hensel et professeur de mathématiques à l'Imperial College à Londres, H. Edwards met la main sur quelques lettres adressées à Kronecker, et acquiert la certitude qu'Hensel possédait bien le *Nachlass* complet à Marburg au moment de sa mort. Il découvre ensuite que ce dernier était en 1943 en la possession de Hasse, qui malheureusement, malade et âgé de 78 ans au moment où H. Edwards l'interroge, ne se souvient absolument pas de ce qu'il a pu faire de ces documents. H. Edwards réussit finalement à établir que Hasse confia en 1944 ses papiers les plus importants, et que ceux-ci furent mis en sécurité au fond d'une mine située à 20km de Göttingen qui servait aussi de réserve pour des munitions. L'ensemble explosa en 1945 et presque tous les papiers seront perdus. Il semble donc bien que la presque totalité des *Nachlass* de Kronecker et de Hensel ait été détruite de cette façon.

Pour étudier le travail de Kronecker, hors ses articles de recherche, nous ne disposons donc essentiellement que des manuscrits de ses leçons à l'Université de Berlin. Mais si ces leçons sont bien, comme le dit Mittag-Leffler, d'une « plus complète clarté », et si au contraire son œuvre publiée est, pour reprendre les mots de H. Edwards, « si difficile à lire », alors très certainement la lecture des premières devrait éclairer la seconde.

---

Gedanken zu folgen vermochten ».

35. Ma traduction de [Edwards, 1978] p. 425.

36. Voir [Edwards, 1978].

37. Pierre Dugac, dans un article consacré aux travaux de Weierstrass et Dedekind, écrit ceci : « Signalons que nous n'avons pas pu trouver trace du *Nachlass* de Kronecker qui serait très utile pour comprendre cette époque (la correspondance de Kronecker avec tous les grands mathématiciens de son temps serait essentielle à cette fin) ». ([Dugac, 1976] p. 11)

38. Hensel lui-même l'affirma dans les introductions des éditions des oeuvres de Kronecker. De plus, dans une lettre de Fuch publiée par Biermann dans son célèbre *Die Mathematik und ihre Dozenten an der Berliner Universität*, adressée au ministre von Zedlik et datant de 1892, on trouve la demande de promotion d'Hensel au titre de *Extraordinarius* mentionnant l'attribution de cette tâche à ce dernier.

## Plan de la thèse

Dans une première partie, nous présenterons les manuscrits des cours de Kronecker sur la *Theorie der algebraischen Gleichungen* – matériel original et qui jusqu'ici n'a été ni édité ni exploité. Après en avoir présenté le contexte de rédaction, nous nous attarderons sur la notion d'« Algèbre » qui émerge de ces leçons. Il nous faudra ensuite entrer plus précisément dans le contenu mathématique de ces cours, et nous consacrerons un long développement à l'analyse de la première partie de l'un d'entre eux, la seconde partie faisant l'objet de nos deuxième et troisième parties. Pour cela, nous tenterons de fournir une description aussi précise que possible des *Vorlesungen über die Theorie der algebraischen Gleichungen* : dans le cadre de notre travail, il n'est pas envisageable d'exposer minutieusement chaque manuscrit en notre possession, et nous avons donc choisi de suivre plus particulièrement l'un d'entre eux[39], en le comparant, autant que possible, aux autres versions de ce cours. Pour continuer notre travail de contextualisation et mettre en évidence l'originalité des leçons de Kronecker, nous les comparerons à un autre cours emblématique de la fin du XIX^e^ siècle : le *Lehrbuch der Algebra* de Weber.

D'après H. Sinaceur, mais aussi de l'aveu même de Kronecker, le théorème de Sturm tient une place centrale dans l'élaboration de la théorie des caractéristiques. Après avoir présenté notre source principale, nous poursuivrons son exploitation en examinant la lecture que fait Kronecker du théorème de Sturm et de ses prolongements par de nombreux mathématiciens, tels Sylvester, Cayley, Cauchy ou encore Hermite. À cette occasion, Kronecker revient sur les notions de racine et de continuité. Il nous faudra alors examiner comment ces concepts sont remaniés à la lumière de pratiques mathématiques spécifiques, celles qu'il expose dans ses cours.

Ayant analysé les soubassements et le contexte d'écriture de la théorie des caractéristiques, nous en donnerons une exposition complète dans notre troisième partie. Nous chercherons à comprendre alors comment la lecture du cours facilite, voire permet, l'appréhension de ses articles de recherche de 1869 et 1878, et nous ferons apparaître comment cette théorie répond, du point de vue de Kronecker, au problème de l'existence des racines d'un polynôme. On trouve en effet au cœur de cette théorie trois des quatre démonstrations que Gauss a données du théorème fondamental de l'algèbre, et nous serons ainsi amenés à examiner la façon dont Kronecker reprend celles-ci.

39. Ce sera le manuscrit du semestre d'hiver 1880/1881 : nous donnerons une justification de ce choix ultérieurement.

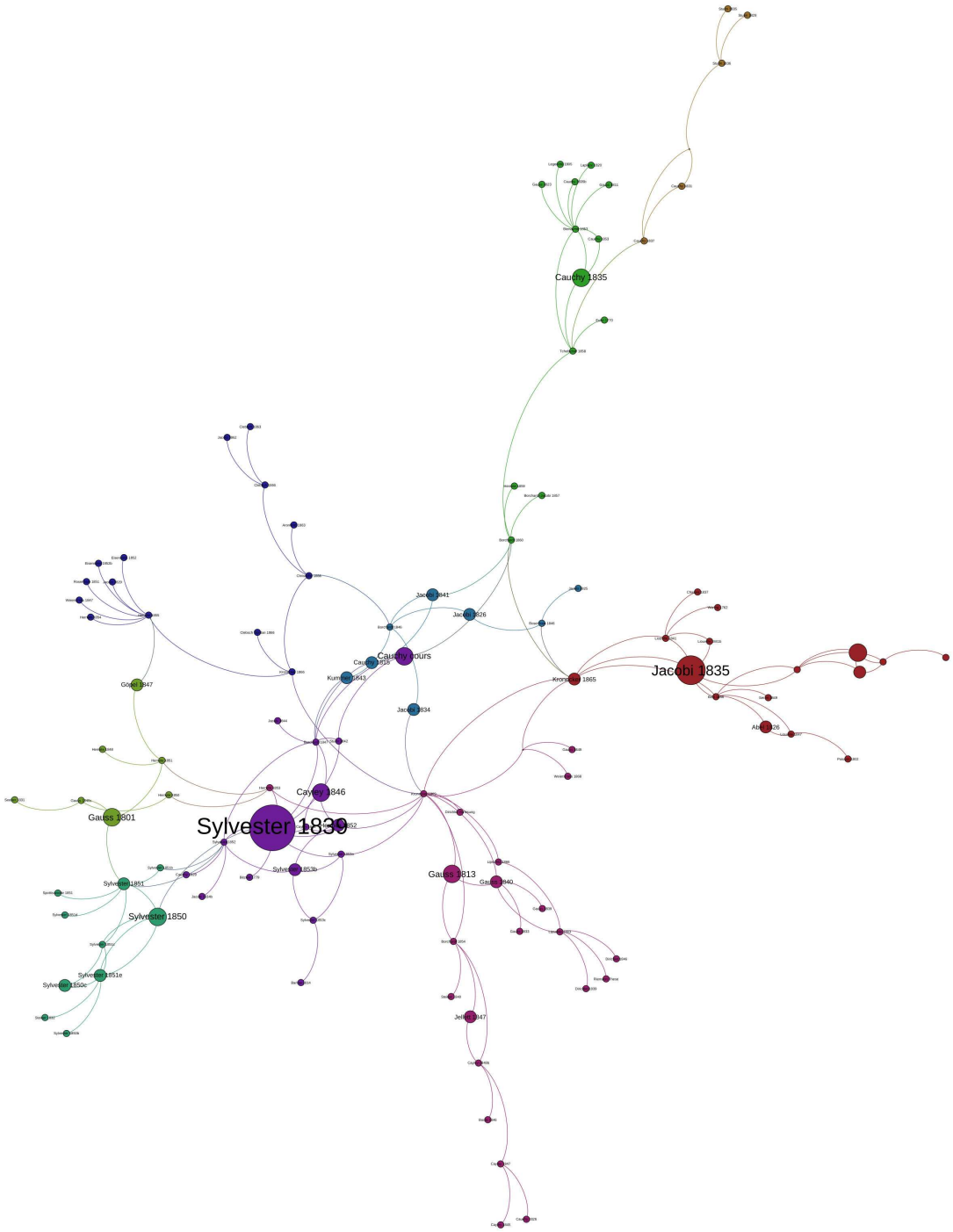



FIGURE 1 – Graphique obtenu à partir de la spatialisation *Yifan Hu*

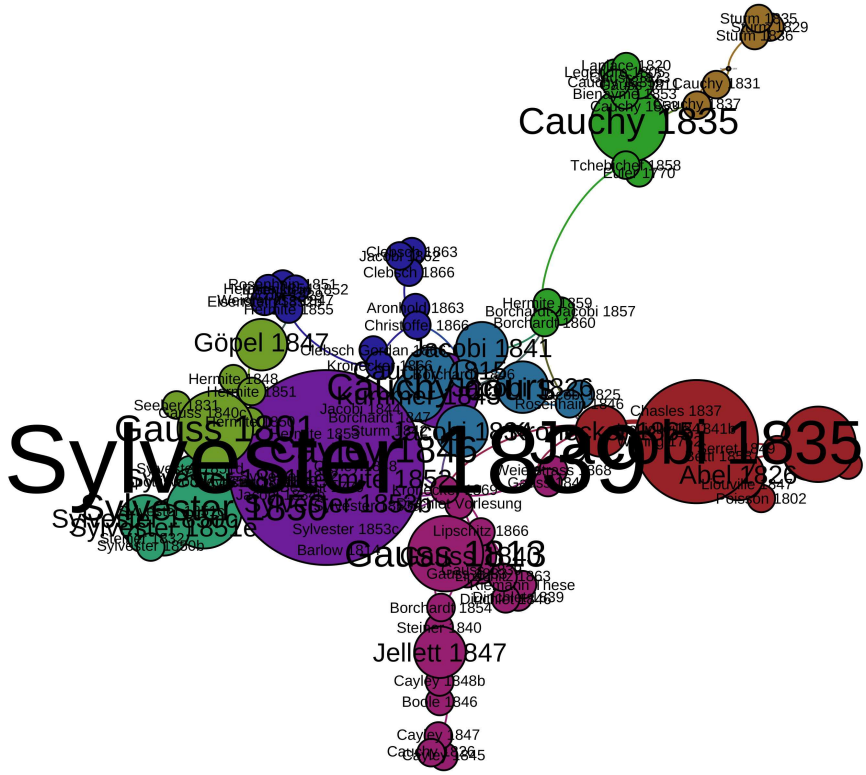



FIGURE 2 – Graphique obtenu à partir de la spatialisation *Force Atlas*

# Première partie

# La théorie des équations algébriques de Kronecker

# Chapitre 1

# Introduction


1.1 L'université de Berlin dans la seconde moitié du XIX^e siècle . . . . . **32**
1.1.1 La construction de l'université . . . . . 32
1.1.2 Les mathématiques à l'université de Berlin . . . . . 34
1.1.3 Le séminaire de mathématiques à l'université de Berlin . . . . . 36
1.2 Présentation des *Vorlesungen* de Kronecker . . . . . **39**
1.2.1 Kronecker enseignant . . . . . 39
1.2.2 Les manuscrits des cours de Kronecker . . . . . 43
1.3 Les *Vorlesungen über die Theorie der algebraischen Gleichungen* . . . . **51**
1.3.1 La structure des manuscrits . . . . . 52
1.3.2 Permanence de la structure . . . . . 58
1.3.3 Évolution du contenu . . . . . 58
1.3.4 Une introduction historique . . . . . 59


Dans un article [1] publié en 1998, Bruno Belhoste montre que l'enseignement « constitue un aspect essentiel de l'activité mathématique » [2] et que les historiens des mathématiques auraient tort de ne pas lui accorder « toute l'importance qu'il mérite ». Dans ce court essai, il présente essentiellement trois axes d'étude que selon lui la recherche en histoire des mathématiques devrait suivre. Le premier concerne l'histoire institutionnelle, le second cherche à décrire les liens entre pratiques pédagogiques et pratiques de recherches, et le dernier discute de la représentation des mathématiques. Nous allons, dans cette introduction, nous attacher aux deux premiers de ces trois thèmes. Pour présenter les *Vorlesungen* de Kronecker, nous commencerons par montrer comment ces leçons s'intègrent dans l'histoire de la construction et de l'évolution de l'université de Berlin depuis son ouverture en 1810. Ce faisant, les liens très forts qui existent entre les cours de Kronecker et sa recherche apparaîtront comme le témoignage d'une pratique commune, que l'étude plus détaillée des parties suivantes nous permettra de particulariser.

1. Voir [Belhoste, 1998].
2. Voir [Belhoste, 1998] p. 289.

## 1.1 L'université de Berlin dans la seconde moitié du XIX^e siècle

Dans ses célèbres *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, après avoir longuement présenté les mathématiques en France au XIX^e siècle et en particulier à l'école Polytechnique, Felix Klein poursuit avec l'Allemagne [3] :

> La nouvelle Allemagne du XIX^e siècle, qui se développe progressivement après les guerres Napoléoniennes, est dans son essence déterminée par des suggestions venues de France, et assimilées par l'esprit allemand. Comme Goethe dans un autre domaine, Gauss s'est tenu dans notre science en dehors des développements de son temps. Ce développement commence à Berlin, mais, comme nous l'avons déjà mentionné, un peu plus tard pour les sciences exactes que pour les autres domaines scientifiques. Pour les sciences humaines le point de départ a été la fondation de l'université de Berlin en 1810. Elles prospèrent, supportées par la doctrine néo-humaniste de construction libre de la personnalité, une doctrine qui se détourne des sciences exactes.
>
> Le nouveau mouvement dans les sciences commence à être notable en 1820, essentiellement à l'initiative d'Alexandre von Humboldt, comme je l'ai déjà mentionné. [Klein, 1926] p. 93

Klein date ainsi le début de l'« Aufblühen der reinen Mathematik in Deutschland » [4] vers 1820, après la création de l'université de Berlin, sous la double influence de l'« esprit allemand » et du modèle français. Et c'est dans cette université que Kronecker étudiera d'abord, et enseignera ensuite, participant activement à cet *Aufblühen*, en particulier pendant ce que Adolf Kneser qualifiera de « temps héroïques » (*heroische Zeitalter* [Kneser, 1925] p. 211). Nous allons montrer dans cette section comment la mise en place de la *Friedrich-Wilhelms-Universität*, de par les idées qui ont participé à sa formation, agit sur la structure de l'enseignement de Kronecker, et contribue ainsi à rendre pertinent l'accès à ses travaux les plus pointus par l'étude de ses *Vorlesungen*.

### 1.1.1 La construction de l'université

Plusieurs facteurs, qu'ils soient politiques ou philosophiques, amènent à la création de l'université de Berlin [5]. Le début du XIX^e siècle voit de grands changements dans la partie

3. Das neue Deutschland des 19. Jahrhunderts, das sich allmählich aus den Napoleonischen Kriegen heraus entwickelt, ist in seinem Wesen bestimmt durch die von Frankreich kommenden Anregungen, die im Sinne des deutschen Geistes verarbeitet werden. Wie auf anderem Gebiet Goethe, so steht in unserer Wissenschaft Gauß außerhalb der von der Zeitströmung getragenen Entwicklung. Diese Entwicklung setzt in Berlin ein, aber, wie schon früher bemerkt, für die exakten Wissenschaften etwas später als auf anderen wissenschaftlichen Gebieten. Fur die Geisteswissenschaften bildet die Gründung der Universität Berlin 1810 den Ausgangspunkt. Sie blühen auf, gestützt auf die neuhumanistische Lehre von der freien Bildung der Persönlichkeit, die sich vom Interesse für die exakten Wissenschaften direkt ab wandte.

Hier macht sich die neuzeitliche Regung erst von 1820 ab bemerkbar, wesentlich durch die Initiative Alexander von Humboldts, wie ich es früher ja bereits dargelegt habe.

4. Voir [Klein, 1926] p. 93.

5. Nous concentrerons ici nos efforts sur cette université, car notre étude porte sur le travail de Kronecker, dont l'ensemble de la carrière universitaire se déroulera à Berlin. Il serait cependant faux de croire que les changements que nous allons décrire ne prennent place qu'à Berlin, et qu'aucune des autres grandes universités n'y participe. Turner montre bien dans [Turner, 1980] l'importance du rôle de Halles, Königsberg,

germanique de l'Europe. 1806 marque, avec l'abdication de François II, la fin du Saint Empire romain germanique, et le congrès de Vienne quelques années plus tard permettra le rassemblement des états allemands au sein de la Confédération germanique. C'est dans cette période de reconstruction politique que Wilhelm von Humboldt (1767–1835), ministre de l'éducation, va, en Prusse, initier un renouveau dans l'enseignement. Il centralise les institutions, commence à recruter un corps professoral qui sera rémunéré par l'État, produit des « programmes » et finalement change complètement et en profondeur l'enseignement, et en particulier celui des mathématiques, d'abord en Prusse, puis dans la plupart des universités allemandes [6]. Ce mouvement sera facilité, au niveau politique, et donc financier, par l'importance que revêt la faculté comme un symbole de la « Prusse unie, prestigieuse et puissante » [7].

Dans son ouvrage *Neohumanism and the persistence of pure mathematics in Wilhelmian Germany*, Lewis Pyenson montre comment le néo-humanisme [8] influence les réformes éducatives en Prusse, et en particulier participe à la structuration de l'université de Berlin. Ce renouveau est en partie un néo-classicisme qui pose comme idéal la culture et la philosophie de la Grèce antique. Le but n'est pas ici de chercher les soubassements de ce courant de pensée, mais de voir comment il influence les réformes éducatives en Prusse, dont l'université de Berlin fait partie. Gert Schubring affirme que :

> The conception underlying the reforms was neo-humanism, which understood human knowledge as an interconnected organism so that intellectual development afforded an all-sided incitement of cognitive abilities in the students. [Karp et Schubring, 2014] p. 242

Et ainsi, à côté des « humanités » classiques, l'enseignement des sciences s'accroît, et en particulier celui des mathématiques, pour développer certaines de ces « habiletés cognitives ». Ainsi, contrairement à ce que semble affirmer Klein dans la citation précédente, le néo-humanisme sera l'un des facteurs du développement des sciences exactes, et non un obstacle qu'il faudra surmonter [9]. La position de Klein doit se comprendre dans le contexte de la fin du XIX$^{e}$ siècle, période durant laquelle il cherchera à lutter contre les héritiers, essentiellement berlinois, de cette vision des mathématiques. L'enseignement de ces dernières prend donc au contraire une place importante dans les neuf années du *Gymnasium* nouvellement créé : six heures par semaine leur seront consacrées et cela nécessitera la présence de deux enseignants de mathématiques. Un programme pour cet enseignement est proposé, mais on peut noter qu'une liberté totale est laissée à l'enseignant. Lorsque Kronecker suivra les cours de Kummer (1810-1893) au *Gymnasium*, ces cours auront été construits par Kummer lui-même, sur un programme qu'il aura décidé de suivre. Cela institue donc un rapport très étroit entre le professeur et le contenu de son cours, et ce, dès le secondaire.

Francfort, Bonn, Breslaü...

6. Cette évolution a une histoire, et il ne faudrait pas croire que que Wilhelm von Humboldt, sans aucun modèle, a créé une réforme entièrement nouvelle. On pourra lire à ce sujet [Karp et Schubring, 2014].

7. Voir [Turner, 1980] p. 74.

8. On pourra lire à ce sujet [Pyenson, 1983].

9. Voir [Karp et Schubring, 2014] p. 245.

En 1810, dans une Prusse relativement ruinée et dont la situation politique reste instable, W. von Humboldt crée une université à Berlin dans laquelle ces préceptes seront appliqués. Celle-ci sera donc, dans une certaine mesure, l'aboutissement de ces réformes, et en particulier dans la continuité du *Gymnasium* où

> After Greek and Latin, mathematics formed a third, crystalline language in the secondary schools of Central Europe. Mathematics came as a system of formal operation, lacking, as Greek and Latin lacked, practical utility. [Pyenson, 1983] p. 3

Cette affirmation de Pyenson fait écho à la célèbre phrase de Jacobi [10] à propos de « l'honneur de l'esprit humain », et correspond à un point de vue qui sera largement partagé par les enseignants de la faculté de philosophie de Berlin de la seconde moitié du XIX^e^ siècle, tous issus de ces *Gymnasium* [11]. À la fin du XIX^e^ siècle, un changement s'opère dans l'enseignement des mathématiques, en particulier sous l'égide de Felix Klein (1849-1925) à Göttingen, qui tente de mettre fin à cet *ethos* des mathématiques « pures ».

### 1.1.2 Les mathématiques à l'université de Berlin

Nous allons maintenant examiner plus précisément comment, à l'intérieur de cette université, l'enseignement des mathématiques s'organise.

#### 1.1.2.1 L'émergence d'une discipline

Au cours de la création de l'université de Berlin, les mathématiques acquièrent une certaine indépendance. D'abord, la faculté de philosophie, dans laquelle sont enseignées les mathématiques, n'a plus le rôle de propédeutique auprès des étudiants se destinant aux études de loi, de médecine et de théologie : il y a ainsi une plus grande spécialisation des étudiants présents. Par ailleurs, après un accroissement significatif du nombre d'*Ordinarius* et de *Privatdozenten* à l'université entre le début des réformes et 1830, il devient plus difficile d'y obtenir un statut. L'une des solutions sera alors de développer un nouveau domaine qui pourrait justifier l'ouverture d'une nouvelle chaire. C'est ainsi qu'une tendance à la spécialisation se développe chez les jeunes chercheurs, cette voie devenant la seule qui les mènera sur la « route rapide vers l'*Ordinariat* convoité » [12].

Ensuite, la faculté de philosophie hérite de la formation des enseignants du secondaire, et en particulier de ceux de mathématiques, qui suivent un cursus de quatre années. Or les réformes en matière d'éducation et la création des *Gymnasium* nécessitent la formation de

10. Jacobi écrit à Legendre en 1830 : « M. Fourier avait l'opinion que le but principal des mathématiques était l'utilité publique et l'explication des phénomènes naturels ; mais un philosophe comme lui aurait dû savoir que le but unique de la science, c'est l'honneur de l'esprit humain, et que sous ce titre, une question de nombres vaut autant qu'une question du système du monde. »[Jacobi, 1881] p. 454.

11. En 1869, Weierstrass et Kummer, interrogés sur la possibilité d'envoyer à l'université des élèves issus des *Realschule*, se prononcent farouchement contre, car, disent-ils, les sciences qui y sont enseignées « émoussent plutôt qu'elle ne stimulent le désir de connaissances » ([Pyenson, 1983] p. 28). Ils souhaiteraient d'ailleurs que les enseignants de ces écoles soient issus du *Gymnasium*.

12. Voir [Turner, 1980] p. 76.

nombreux enseignants de mathématiques : on voit donc arriver dans ces cours de mathématiques un contingent d'étudiants spécialisés, plus à même de suivre et de participer à des cours d'un niveau plus élevé.

#### 1.1.2.2 Le rapport entre cours et recherche

L'un des éléments essentiels de cette université est la volonté de mettre au premier plan la recherche, et non plus uniquement l'enseignement. F. Paulsen, présente ainsi l'université allemande du XIX$^{e}$ siècle :

> The 19th. century first introduced the requirement of independent learned research : only he is effective as a teacher of science who is himself productive in science. [...] It was the era of the highest intellectual achievement that the German people have ever known, the era of Kant and Goethe, which found the courage to advance this view. Fichte and Schleiermacher first enunciated it decisively in their treatises written upon the occasion of the founding of Berlin University : whoever wishes to enter a scholarly career, upon him is the demand to be placed that he not merely have learned the knowledge at hand, but rather that he also be capable of producing knowledge out of his own independent activity. [...] Under the dominion of this idea the German universities [...] have developed into what they are today : the workshops and the forges of the intellectual life of our people. [Paulsen, 1902] p. 205

C'est cet idéal de la fonction du professeur que Turner [13] appelle « the research imperative », et dont Weierstrass et Kronecker, comme l'affirme Dugac, seront des représentants :

> WEIERSTRASS et KRONECKER présentaient dans leurs leçons, les résultats de leurs recherches et, contrairement à ce qui se faisait habituellement en Allemagne, ils évitaient de faire imprimer leurs travaux. [Dugac, 1973] pp. 60-61

.

Que ce soit dans les cours de Weierstrass ou de Kronecker, la recherche prend donc une place importante. La construction et le contenu de ces leçons sont souvent directement liés à l'état de leur recherches personnelles : il s'agit donc toujours de cours où, contrairement à des leçons plus standardisées et soumises à des programmes comme par exemple à l'école polytechnique à Paris, les frontières entre les différents domaines des mathématiques sont plus floues, plus encore dans les exposés faits par les enseignants au séminaire de mathématiques, où leur recherche propre constitue l'essentiel du contenu. Ainsi, par exemple, après la création du séminaire, Kummer réserve la présentation de ses recherches aux cours assurés dans ce cadre.

Nous l'avons dit, l'une des charges de cette nouvelle université est la formation des enseignants des *Gymnasium* et des *Realschule*. Ces enseignants reçoivent non seulement une formation très poussée en mathématiques, mais doivent en plus, pour obtenir leur diplôme,

13. Voir [Turner, 1980] p. 68.

produire un travail original, et donc s'initier à la recherche. Les bienfaits pédagogiques de cette « méthode d'enseignement » sont décrits par Weierstrass [14] :

> Le succès réside en grande partie dans le fait que le professeur amène constamment l'étudiant à ses propres recherches. Mais cela ne se produit pas par quelque instruction pédagogique, mais d'abord et principalement du fait que le professeur laisse reconnaitre convenablement à l'étudiant lors de l'exposé d'une discipline, par la disposition elle-même de son contenu et la mise en relief des idées directrices, par quels moyens le penseur mûr qui maîtrise la recherche tout juste exposée accède de façon logique et progressive à de nouveaux résultats ou de meilleures justifications à celles déjà existantes. Puis il n'omet pas de lui décrire les limites actuelles non-franchies de la connaissance et d'indiquer les points pour lesquels une avancée future semble prochainement possible. Il ne lui refuse pas non plus un aperçu détaillé dans ses propres recherches, et ne tait pas même les erreurs commises et les attentes déçues. [Biermann, 1973] p. 107

Et c'est ainsi, comme le remarque Bruno Belhoste, qu'« à tous les niveaux, le contact avec la recherche est une condition pour entrer et progresser dans la fonction enseignante » [15].

Faisons ici une première remarque à propos du rapport entre les cours et la recherche chez Kronecker. Certains de ses articles de recherche sont notoirement difficiles, et cette difficulté est aussi liée aux ellipses dans les explications (lorsqu'il y a des explications) et à une grande pauvreté en exemples et en applications. La lecture en parallèle des notes des cours qu'il a donnés sur le même sujet est souvent très éclairante, voire nécessaire à la compréhension de ces articles. Comme on l'a vu, ces cours sont pris en note et mis en forme par les étudiants faisant partie des séminaires au profit des autres participants : les étudiants choisis ont ainsi les outils permettant d'appréhender des mathématiques telles qu'elles sont en train de se faire.

### 1.1.3 Le séminaire de mathématiques à l'université de Berlin

Norbert Elias, dans *La civilisation des mœurs*, alors qu'il compare les concepts de culture et de civilisation en France et en Allemagne, cite une lettre d'Eckermann à Goethe datant de 1827 :

> Du peuple proprement dit nous vient peu de culture, et nos talents et têtes intelligentes sont dispersés à travers toute l'Allemagne. L'un est à Vienne, l'autre à Berlin, un encore à Königsberg, un autre à Bonn ou à Düsseldorf, cinquante à cent milles séparant les uns et les autres, si bien que le contact personnel et l'échange personnel de pensées sont choses fort rares. Ce qu'elles nous apporteraient, je le sens quand des

14. Weierstrass cité par Biermann : « Der Erfolg [...] beruht [...] zum großen Teile darauf, daß der Lehrer den Lernenden fortwährend zu eigener Forschung leitet. Dies geschieht aber nicht etwa durch pädagogische Anweisung, sondern zunächst und hauptsächlich dadurch, daß der Lehrer beim Vortrag einer Disziplin in seiner Darstellung selbst durch Anordnung des Stoffes und Hervorhebung der leitenden Gedanken angemessen den lernenden erkennen läßt, auf welchem Wege der gereifte und das bereits Erforschte beherrschende Denker folgerichtig vorschreitend zu neuen Ergebnissen oder besserer Begründung schon vorhandener gelangt. Dann versäumt er es nicht, ihm die zurzeit nicht überschrittenen Grenzen der Wissenschaft zu bezeichnen und diejenigen Punkte anzudeuten, von denen aus ein weiteres Vordringen zunächst möglich erscheint. Auch einen tieferen Einblick in den Gang seiner eigenen Forschungen versagt er ihm nicht, verschweigt selbst nicht begangene Irrtümer und getäuschte Erwartungen. »

15. Voir [Belhoste, 1998] p. 294.

> hommes comme Alexander von Humbolt passent par ici et me font gagner, en une seule journée, là où je cherche et où j'ai besoin de savoir, ce que sur mon chemin solitaire je n'aurais pas trouvé en une année.
>
> Or imaginez une ville comme Paris où les esprits les plus remarquables du royaume tout entier sont réunis, où ils s'instruisent et s'enrichissent dans la fréquentation, la lutte et la compétition de tous les jours, où les meilleurs produits de tous les domaines de la nature et de l'art, de la terre entière, s'offrent à la méditation quotidienne ; imaginez cette métropole, où chaque promenade sur un pont ou une place évoque un grand passé. Et, ce qui plus est, n'imaginez pas le Paris d'une époque terne et inculte, mais le Paris du XIX^e siècle, où depuis trois générations des hommes comme Molière, Voltaire, Diderot et d'autres ont mis en circulation une somme d'esprit comme on n'en a jamais réunie à la fois en un seul endroit de la terre, et vous comprendrez qu'un homme talentueux comme Ampère, qui a grandi au milieu d'une telle plénitude, peut être quelqu'un en sa vingt-quatrième année. [Elias et Kamnitzer, 1973] p. 61-62

On voit donc la fascination que Paris peut exercer, par la possibilité d'y rencontrer un grand nombre de savants et d'y converser de sujets très divers, sur une partie des intellectuels allemands du début du XIX^e siècle. Norbert Elias rappelle [16] qu'à cette époque, à la différence de la France, le moyen principal de la communication scientifique est l'écrit, et que l'absence de lieu de discussion place le chercheur dans une certaine solitude. Les séminaires, mis en place à partir de ces réformes au début du XIX^e siècle, seront l'une des formes que prendra cet effort en faveur de la recherche à l'université, et permettront en partie de remédier à cette difficulté. Le séminaire de mathématiques de l'université de Berlin est directement issu des séminaires qui sont progressivement mis en place dans les universités allemandes, d'abord en sciences humaines, puis en sciences à partir de 1825.

#### 1.1.3.1 Naissance du séminaire

À l'aube des années 1860, toutes les conditions sont réunies pour la création d'un séminaire de mathématiques à Berlin. Tout d'abord, comme tous les séminaires mis en place depuis le début du siècle, il nécessite un financement. Or le début des années 1860 est marqué par une forte augmentation des sommes allouées à la création d'instituts et de séminaires : la période est donc politiquement propice à la création d'un nouvel enseignement de ce type.

Mais il faut aussi, pour qu'un tel séminaire soit mis en place, que sa demande soit portée à l'intérieur de l'université. Weierstrass et Kummer seront à l'origine de cette demande, et il faut bien dire que l'appui de ce dernier, professeur *ordinarius*, ayant déjà assumé la fonction de doyen de l'université et ayant reçu de multiples prix internationaux, a pu être déterminant dans l'obtention de la création de ce séminaire.

Enfin, Kummer a étudié à Halle entre 1828 et 1831, période durant laquelle est née l'association de mathématiques d'Heinrich Scherk [17], et dont le fonctionnement était proche

16. Voir [Elias et Kamnitzer, 1973] p. 63.
17. Heinrich Ferdinand Scherk (1798-1885) est un mathématicien allemand, qui a en particulier enseigné à l'université de Halle, dans laquelle il a encadré la thèse de Kummer.

de celui d'un séminaire de sciences, et en particulier de mathématiques. C'est d'ailleurs en participant à ces séances que Kummer se spécialise en mathématiques : il a ainsi l'expérience du fonctionnement d'un séminaire.

De plus, lorsque le séminaire de mathématiques de l'université de Berlin se crée en 1861, il a le potentiel pour devenir le lieu privilégié des échanges entre les divers domaines des mathématiques et de la physique au sein de la faculté. Comme on le sait, en 1861 Kronecker commence à enseigner à Berlin et rejoint ainsi Kummer (1855) et Weierstrass (1856) : ensemble, ces trois piliers du séminaire couvrent la plupart des domaines des mathématiques connus de l'époque.

Il y a une volonté d'extension des mathématiques de la part de ces chercheurs, et de multiplier les interactions entre les domaines : ils tentent ainsi d'étendre l'arithmétisation des mathématiques à l'ensemble de ces domaines.

#### 1.1.3.2 Organisation des séminaires

Ce séminaire se propose d'atteindre un double but : il doit servir à la fois à aider les étudiants à aiguiser leur habileté à enseigner et en même temps à développer leurs facultés de recherche. Par ailleurs, c'est le lieu privilégié, sinon le seul lieu, où il peut exister un contact direct entre étudiants et enseignants. Les exposés sont d'ailleurs tour à tour présentés par les uns ou les autres.

Le contenu des exposés, compléments des cours auxquels assistent ou ont assisté les étudiants du séminaire, est donc donné par les enseignants, et finalement, comme l'écrit Weierstrass en 1882, les étudiants y ont

> the opportunity to [begin] with a two-year series of lectures on the most important mathematical disciplines in measured succession. Among topics covered, not a few will be read either not at all at other universities or not regulary. This had the effect...especially since lectures in mathematical physics received more complete treatment here than elsewhere...[of drawing] young men... not in trifling numbers. [Calinger, 1996] p.166

Entre les cours et le séminaire, c'est l'ensemble des mathématiques de l'époque qui est discuté au sein de l'université.

L'admission est très sélective : seuls 12 étudiants par an ont le droit de s'inscrire à ce séminaire, et l'un des outils les plus importants que ce séminaire apporte à ses membres, c'est un accès privilégié à une bibliothèque spécialisée, qui contient la plupart des revues de mathématiques européennes, mais aussi de physique : chaque étudiant pourra ainsi avoir accès à l'ensemble des domaines des mathématiques qui sont en train de se faire.

Les membres du séminaire produisaient eux-mêmes des compléments à cette bibliothèque : ils mettaient en forme leurs notes de cours, souvent avec l'assistance de l'enseignant lui-même, et les rendaient ainsi à leurs condisciples.

Tous les mercredis après-midi, une séance de recherche de problèmes et d'exercices a lieu, et toujours en rapport avec les recherches de l'enseignant qui propose le problème : ces recherches systématiques de problèmes ouverts encouragent l'utilisation d'outils issus

de divers domaines des mathématiques. Sur une période plus longue, la mise en place d'un prix spécifique au séminaire encourage le même type de pratiques.

Le fait que les mathématiques se sont constituées au sein de l'université comme une discipline à part entière est l'une des raisons pour lesquelles le séminaire de mathématiques apparaît. Ce dernier a un pouvoir attractif fort auprès d'étudiants souhaitant se consacrer aux mathématiques, et c'est donc finalement une plus grande spécialisation des étudiants qui a permis la création de ce lieu qui sera l'un des endroits privilégiés des échanges entre les divers domaines des mathématiques à l'intérieur de l'université de Berlin.

## 1.2 Présentation des *Vorlesungen* de Kronecker

Les supports sur lesquels nous allons travailler seront essentiellement constitués de manuscrits : contrairement aux autres séries de cours que Kronecker a professées à Berlin, les *Vorlesungen über die Theorie der algebraischen Gleichungen* que nous examinerons dans cette partie n'ont pas fait l'objet d'une édition. Avant de parler plus particulièrement de ces leçons sur la théorie des équations algébriques, nous allons commencer par donner une présentation générale des cours de Kronecker, ainsi que des manuscrits qui nous les rendent accessibles. Nous venons de voir comment l'université de Berlin s'est constituée au cours du XIX$^{e}$ siècle : décrivons maintenant à grands traits la trajectoire que Kronecker y a effectuée.

### 1.2.1 Kronecker enseignant

L'université de Berlin a été le lieu de formation de Kronecker, mais aussi celui où il aura passé l'ensemble de sa carrière de chercheur [18] et d'enseignant. Il y professera ainsi ses *Vorlesungen*, pendant 30 ans, sous différents statuts, entre le *Wintersemester* 1862-1863 et le *Wintersemester* 1891-92. En 1861, il est élu, sur proposition de Kummer – proposition soutenue par Weierstraß et Borchardt – à l'Académie de Berlin, ce qui lui donne le droit d'enseigner à l'université de Berlin [19] comme *lesendes Akademiemitglied*. Encouragé par Kummer, il utilisera ce droit dès l'année suivante, et continuera jusqu'à sa mort en 1891. En 1864, il devient *Prädikat Professor*, et obtient enfin, en 1883, le statut d'*Ordentlicher Professor*. C'est finalement un lien fort qui unit Kronecker, la *Karl Friedrich-Wilhelm Universität*, les cours qu'il y a donnés et les étudiants qui ont pu y assister.

18. Nous avons vu que l'une des particularités de l'université berlinoise est d'avoir, lors de sa création, accru l'importance de la recherche par rapport à l'enseignement, entre autres en développant un statut d'enseignant qui laisse du temps à la recherche : c'est ce que Gert Schubring appelle « the "dual role" of the teacher-researcher » ([Mehrtens *et al.*, 1981] p. 116). Par ailleurs, entre sa thèse soutenue en 1845 et le début de ses enseignements en 1861, Kronecker aura eu une autre activité qui le retiendra dix ans à Liegnitz : il devra, à la mort de son oncle, s'occuper des affaires familiales. Il se mariera en 1848 avec sa cousine, et lorsqu'il retournera en 1855 à Berlin, Kronecker sera suffisamment à l'aise financièrement pour ne pas avoir à chercher un emploi rétribué (On pourra lire [Biermann, 1981]).

19. On pourra lire à ce sujet [Biermann, 1973] p. 100.

#### 1.2.1.1 Une publication souhaitée

Hensel, dans l'introduction qu'il rédige pour le second tome de l'édition des *Vorlesungen*, explique que Kronecker lui-même souhaitait que ses leçons servent de base à une publication [20] :

> Kronecker lui-même m'a souvent et de façon détaillée parlé du dessein qu'il avait de faire publier ses leçons ; mais il insistait toujours sur le fait qu'il fallait pour cela encore les retravailler entièrement et les organiser. La valeur et l'attrait des leçons académiques reposent ainsi dans de tout autres qualités que celles d'un livre de cours sur le même sujet. Ici on ne doit pas donner tous les moyens pour l'examen approfondi de domaines entiers, mais l'enseignant doit plutôt éveiller l'enthousiasme pour cette discipline, il doit en quelque sorte introduire les auditeurs à l'intérieur de l'atelier des hommes qui ont réellement fait avancer la science. Ici on peut se passer d'une disposition entièrement rigoureuse, d'une représentation complète de tous les domaines, car cela les étudiants les trouveront plus tard dans les traités et les publications ; ici l'enseignant doit aborder des problèmes stimulants et prometteurs, même si les recherches ne peuvent pas encore être conduites jusqu'à leur achèvement, car justement des esprits réceptifs sont stimulés par de telles questions plus profondément que par des recherches pleinement menées. De plus les auditeurs des leçons de Kronecker étaient déjà en grande partie si bien initiés qu'il pouvait supposer connus de nombreux prérequis qui dans une représentation systématique auraient nécessairement dû être traités en détail. [Kronecker, 1901] p. VIII-IX

On voit dans ce passage l'enthousiasme que semble susciter Kronecker dans ses leçons, et la place importante qu'il réserve à l'exposition de ses recherches les plus récentes. Nous possédons, grâce aux manuscrits des cours de Kronecker, un témoignage direct de son regard sur les notes que prenaient ses étudiants [21] :

> Je souhaite ensuite vous demander si l'un ou plusieurs d'entre vous ont l'envie et la possibilité de préparer d'abord un sténogramme, puis une mise au propre de mes leçons. Peut-être plusieurs peuvent-ils se partager le travail. Pour ceux qui souhaiteraient faire cela pour moi, je vous demande de venir me voir jeudi à 12 heures. Je leur indiquerai alors les conditions plus précises. [Kronecker, 1885b] p. 16

20. Kronecker selbst hat über den Plan, seine Vorlesungen herauszugeben, oft und eingehend mit mir gesprochen ; aber er betonte dabei stets, dass sie für diesen Zweck noch ganz wesentlich umgearbeitet und systematisiert werden müssten. Liegt doch der Wert und der Reiz akademischer Vorlesungen in ganz anderen Vorzügen, als der eines Lehrbuches über denselben Gegenstand. Hier sollen nicht alle Hülfsmittel zur Durchforstung des ganzen Gebietes gegeben werden, wohl aber soll der Lehrer die Begeisterung für jene Disziplin wecken, er soll die Hörer gewissermaßen in das Innere der Werkstatt der Männer einfuhren, welche die Wissenschaft wirklich gefördert haben. Hier kann man auf eine völlig strenge Disposition, auf eine erschöpfende Darstellung des ganzen Gebietes verzichten, denn dies findet der Lernende später in den Lehrbüchern und Abhandlungen ; hier darf der Lehrer auf anregende und aussichtsvolle Probleme eingehen, auch wenn die Untersuchung noch nicht zu einem vollen Abschluss geführt werden kann, denn gerade solche Fragen werden empfängliche Geister viel tiefer zu eigenen Problemstellungen anregen, als vollständig durchgeführte Untersuchungen. Außerdem waren die Zuhörer der Kroneckerschen Vorlesungen großenteils bereits so gut vorgebildet, dass er viele Voruntersuchungen als bekannt voraussetzen konnte, auf die bei einer systematischen Darstellung notwendig ausführlich eingegangen werden musste.

21. Ich möchte dann noch die Frage an Sie richten, ob nicht jemand oder mehrere Lust und Vermögen dazu hat, ein Stenogramm und danach eine Ausarbeitung meiner Vorlesung anzufertigen, vielleicht sich mehrere sie die Arbeit teilen. Diejenigen, welche die Idee haben sollten, das für mich zu machen, bitte ich, Donnerstag um 12 Uhr zu mir zu kommen. Ich werde Diesen dann die näheren Umstände mitteilen.

Nous n'avons malheureusement pas plus de témoignages sur ce qui a pu se dire à ce rendez-vous, mais il est certain qu'au moins trois des étudiants semblent avoir répondu présent [22]. Plusieurs témoignages [23] corroborent le fait que Kronecker ait en effet souhaité publier ses leçons, mais Hensel précise qu'« après sa mort aucun travail préliminaire pour la réalisation de ce projet n'a été trouvé dans son *Nachlass* » [24]. Dans ce paragraphe, Hensel décrit aussi la forme que devrait prendre à ses yeux un cours : on retrouve cette volonté d'allier enseignement et recherche qui caractérise l'université de Berlin au XIX$^{e}$ siècle et que nous décrirons dans le prochain chapitre.

Nous avons aussi là un témoignage qui nous donne une première information sur la qualité des auditeurs de Kronecker : il s'agit bien d'étudiants avancés dans leurs études et prêts à suivre un cours basé sur des prérequis étendus.

#### 1.2.1.2 Des cours difficiles

En effet, il leur fallait une connaissance assez vaste des mathématiques pour pouvoir suivre un cours dans lequel Kronecker exposait les résultats de ses recherches les plus récentes [25] :

> Ses leçons étaient exclusivement vouées aux objets auxquels il portait ses propres intérêts scientifiques, et ainsi les résultats de ses recherches étaient connus très tôt du cercle de ses étudiants les plus proches ; et ses leçons faisaient partie, même si elles n'étaient compréhensibles qu'aux étudiants avancés et scientifiquement intéressés, de ce que l'université de Berlin avait de plus précieux et de plus particulier à offrir au mathématicien. [Weber, 1893b] p. 15

Ces cours, par leur contenu, attiraient des mathématiciens venant de toute l'Europe. Adolf Kneser [26], dans un discours écrit pour les cent ans de la naissance de Kronecker, donne une description de l'enseignant qu'il fut, et le décrit comme un orateur enthousiaste et plein d'esprit [27] :

> Les traits d'esprit, c'était une chose à part entière. Kronecker aimait commencer avec une belle introduction préparée, dans laquelle un ensemble de traits aiguisés, plein d'esprit étaient entrelacés, qui tombaient dans l'oreille de chaque auditeur, et qui

22. Il y a trois rédacteurs différents du manuscrit des *Vorlesungen über die Theorie der algebraischen Gleichungen* du semestre d'été 1884-85 ([Kronecker, 1885b]).

23. Nous en verrons d'autres plus loin.

24. Nach seinem Tode fanden sich in seinem Nachlasse gar keine Vorarbeiten für die Ausführung dieses Planes. ([Kronecker, 1901] p. IX)

25. Seine Vorlesungen waren ausschliesslich den Gegenständen gewidmet, denen er sein eigenes wissenschaftlisches Interesse zugewandt hatte, und so waren seine Forschungsresultate dem Kreise seiner vertrauteren Schüler schon früh bekannt, und seine Vorlesungen gehörten, wenn auch nur den fortgeschrittenen und wissenschaftlich interessirten Schülern verständlich, zu dem Werthvollsten und Eingenartigsten, was die Berliner Universität dem Mathematiker bot.

26. Adolf Kneser (1862-1930) a soutenu en 1884 sa thèse dirigée par Kummer et Kronecker à Berlin où il a suivi les cours de Kronecker.

27. Ja die Aperçus, das war eine eigene Sache. Kronecker liebte es, mit einer schön gefaßten Einleitung zu beginnen, in die eine Menge scharfgeschliffener, geistvoller Aussprüche eingeflochten waren, die jedem Zuhörer ins Ohr fielen und bei allen die angenehmen Illusion erweckten, man verstünde etwas von der Sache.

> insufflaient à chacun l'agréable illusion qu'il comprenait quelque chose à la question. [Kneser, 1925] p. 213

Mais la difficulté de ses leçons décourageait beaucoup d'étudiants et [28]

> Les rangées d'auditeurs qui remplissaient au début la pièce s'éclaircissaient vite ; Kronecker supportait ce phénomène récurrent avec calme et sérénité, et il disait qu'il voulait faire accrocher un rideau tiré au milieu de la pièce, de façon à ce que, lorsque l'essaim s'était perdu, le cercle des initiés puisse s'assoir ensemble dans une intimité plus chaleureuse. Kronecker avait une joie infinie à communiquer ses pensées à ces initiés dans des conversations mathématiques. [Kneser, 1925] p. 213

Mittag-Leffler, alors étudiant à Berlin, écrit d'ailleurs en 1875 une lettre à Hermite donnant une idée du nombre d'étudiants qui constitue ce *Kreis der Eingeweihten* :

> Une chose étonnante, je trouve, c'est que Monsieur Weierstrass et Monsieur Kronecker peuvent trouver tant d'auditeurs – entre 15 et 20 – pour des cours si difficiles et si élevés... [Dugac, 1973] p. 153

C'est un enseignant passionné qui n'hésite pas à mettre dans ses cours les résultats de ses recherches les plus récentes qui est donc décrit par Kneser et Hensel. S'il perd une grande partie de son auditoire, ceux qui réussissent à le suivre et qui deviennent, pour la plupart, des mathématiciens à part entière, peuvent attendre de lui toute l'aide souhaitée [29] :

> Je voudrais encore noter, que je m'entretiens volontiers avec vous après les leçons, que je vous encourage expressément à venir me voir librement pour toute remarque ou tout doute, et ce directement après la leçon, de préférence les lundis et les jeudis mais cela est possible les autres jours aussi. [Kronecker, 1885b] p. 15

#### 1.2.1.3 Les atouts de l'oralité

Dans certains des manuscrits [30] des cours de Kronecker que nous allons étudier, on trouve un exposé de ces « belles introductions préparées », qui nous donnent quelquefois un aperçu du point de vue de Kronecker sur ce que doit constituer un cours à l'université [31] :

28. Die Reihen der Zuhörer, die anfangs den Raum füllten, lichteten sich schnell ; Kronecker ertrug diese sich immer wiederholende Erscheinung mit heiterer Gelassenheit und sagte, er wolle einen Vorhang anbringen lassen, der durch die Mitte des Raumes gezogen werden solle, damit, wenn sich der Schwarm verlaufen habe, der Kreis der Eingeweihten in gemütlicher Enge beisammensitzen könne. Den Eingeweihten gegenüber hatte Kronecker eine unendliche Freude an der Mitteilung seiner Gedanken, am mathematischen Gespräch.

29. Ich möchte noch erwähnen, daß ich mich gerne von Ihnen nach den Vorlesungen sprechen lasse, daß ich Sie ausdrücklich ersuche bei jedem Bedenken, bei jeder Unklarheit ganz frei [sich] an mich zu wenden und zwar stets nach der Vorlesung, am liebsten Montags und Donnerstags, aber es geht auch an den andere Tagen.

30. Tout particulièrement dans celui sur la théorie des équations algébriques de l'hiver 1884-85 ([Kronecker, 1885b]).

31. In den Lehrbüchern kann der Verfasser nicht alles sagen, was dem mündlichen Vortrag ~~vorbehalten~~ gestattet bleibt, und wäre es nur der Unterschied, daß ich hier das Recht und vielleicht die Pflicht habe, nicht nur die Methoden anzugeben, durch die man zum Ziele gelangt, sondern auch Sie vor denjenigen zu warnen, welche keinen Erfolg haben können. Diese negative Thätigkeit, dieses Bewahren vor den Abwegen ist das, was nur das mündliche Wort geben kann. Ja, man kann fast sagen, daß eine vollendete schriftliche Darstellung gerade das Gegenteil von dem bewirkt, was sie bezweckt. Kaum wird es einen Leser geben, welcher aus einem Satz, der genau so viel besagt, wie er besagen soll, nicht zuerst doch etwas Abgeändertes herausließt. Der mündliche Vortag hat den Vorzug denselben Gedanken-Inhalt unter verschiedenen Formen

> Dans un livre scolaire, l'auteur ne peut pas dire tout ce qui est autorisé à garder dans une présentation orale. Et s'il n'y avait qu'une différence, c'est que j'ai ici le droit et peut-être le devoir, de ne pas seulement donner les méthodes qui conduisent au but, mais aussi de vous mettre en garde contre celles qui ne peuvent pas réussir. C'est cette activité négative, cette mise en garde contre la déroute, que seule la parole orale peut transmettre. Oui, on peut même dire qu'une présentation écrite achevée est le contraire complet de ce qu'elle a pour but de servir. Il n'y a que peu de lecteurs qui, d'une phrase qui dit exactement ce qu'elle doit dire, ne lisent pas d'abord quelque chose de différent. L'exposé oral a l'avantage de pouvoir donner le même contenu de la pensée plusieurs fois sous différentes formes, qui comme un nombre de cercles ont une petite partie commune, et cette façon de délimiter est plus facile que par seulement une courbe. [Kronecker, 1885b] p. 2

La fonction du cours revêt ainsi pour Kronecker une grande importance, en particulier par la possibilité qu'il lui est laissée d'avancer d'un point à l'autre par de multiples chemins, dont certains, et il insiste sur ce point, n'ont pas d'issue. Il y a une difficulté intrinsèque à comprendre un texte mathématique écrit, que seule une explication orale permet de surmonter. Si une grande partie de cette oralité a certainement été perdue lors de la prise de notes [32], les manuscrits resteront une aide précieuse pour comprendre et ne pas donner de mauvaises interprétations à certains textes que Kronecker a publiés.

### 1.2.2 Les manuscrits des cours de Kronecker

Les manuscrits des cours de Kronecker se trouvent, pour l'essentiel, dans les collections patrimoniales des bibliothèques de l'Université de Strasbourg et sont issus pour la plupart du *fonds Hensel*. Il s'agit de notes de cours prises par certains de ses auditeurs qui ont par la suite été mises en forme, souvent à la demande de Kronecker lui-même [33]. Nous avons vu que beaucoup de documents appartenant à Hensel avaient été détruits en même temps que le *Nachlass* de Kronecker. Sa bibliothèque, qui contenait environ 1500 volumes, a quant à elle été vendue rapidement après sa mort, en 1941, par sa famille. La *Reichuniversität* de Strasbourg – généreusement dotée car relais de la propagande en territoire récemment conquis – a acquis tous ces volumes. Il semble que les notes de cours personnelles de Kronecker que possédait Hensel aient fini comme le reste de ses documents personnels : au fond de la mine [34].

On peut dénombrer 26 manuscrits à la bibliothèque de Strasbourg répartis ainsi :

---

mehrmals geben zu können, welche wie eine Anzahl Kreise ein Stückchen Gemeinsamer haben, und diese Art der Begrenzung ist leichter als die durch nur eine Curve.

32. Dans l'introduction du manuscrit [Kronecker, 1885b], le rédacteur semble avoir pris en note l'ensemble du discours de Kronecker. La plupart des passages que nous utilisons pour essayer de donner quelques éléments de contexte sur le déroulé de ces leçons ont d'ailleurs été barrés en vue très certainement de l'édition de ce cours. Si cette dernière avait vu le jour, tous ces éléments en auraient certainement été expurgés. Ce rédacteur, qui ne semble faire aucun tri dans sa prise de note, laissera d'ailleurs sa place à deux autres étudiants pour la suite du cours.
33. Voir [Kronecker, 1901] p. VI, ainsi que le paragraphe précédent.
34. Voir plus haut p. 24.

Cours professés par Kronecker à l'université de Berlin dont nous possédons au moins un manuscrit se trouvant à la bibliothèque de l'Université de Strasbourg.

1. Winter Semester 1872/73 Theorie der algebraischen Gleichungen (Mis en forme par Weltzien, certainement Carl ou Karl.)
2. Winter Semester 1875/76 Anwendung der Analysis des Unendlichen auf die Zahlentheorie (Mis en forme par Georg Hettner)
3. Winter Semester 1878/79 Theorie der algebraischen Gleichungen
4. Winter Semester 1880/81 Theorie der algebraischen Gleichungen
5. Winter Semester 1881/82 Höhere Zahlentheorie (Mis en forme par Carl Runge)
6. Winter Semester 1883/84 Vorlesungen über Zahlentheorie
7. Winter Semester 1883/84 Vorlesungen über einige bestimmte Integrale
8. Sommer Semester 1884 Vorlesungen über die Theorie der Determinanten
9. Sommer Semester 1884 Arithmetische Vorlesungen
10. Winter Semester 1884/85 Vorlesungen über die Theorie der algebraischen Gleichungen (donné par le Dr. Thiel, tampon du Dr. Schlesinger Lajos)
11. Winter Semester 1884/85 Vorlesungen über die Theorie der algebraischen Gleichungen : Theil I
12. Sommer Semester 1885 Theorie der algebraischen Gleichungen : Teil II
13. Sommer Semester 1885 Vorlesungen über die Theorie der einfachen und mehrfachen Integrale
14. Winter Semester 1885/86 Vorlesungen über Zahlentheorie (Mis en forme par Th. Lohnstein)
15. Winter Semester 1885/86 Über die Affecte der Gleichungen, welche in der Theorie der elliptischen Funktionen vorkommen (Mis en forme par Paul Stäckel).
16. Sommer Semester 1886 Vorlesungen über die Theorie der Determinanten
17. Sommer Semester 1886 Arithmetische Vorlesungen
18. Winter Semester 1886/87 Vorlesungen über die Theorie der algebraischen Gleichungen
19. Winter Semester 1887/88 Vorlesungen über Zahlentheorie
20. Sommer Semester 1888 Vorlesungen über die Theorie der Determinanten
21. Winter Semester 1888/89 Vorlesungen über die Theorie der algebraischen Gleichungen
22. Sommer Semester 1889 Theorie der einfachen und mehrfachen Integrale 2 tomes (August Gutzmer)
23. Winter Semester 1889/90 Allgemeine Arithmetik : I. Teil (Zahlentheorie)

24. Sommer Semester 1890 Vorlesungen über die Theorie der Determinanten
25. Winter Semester 1890/91 Theorie der algebraischen Gleichungen
26. Sommer Semester 1891 Über dem Begriff der Zahl in der Mathematik

On peut ajouter à ces manuscrits quelques notes de cours de Runge et Kneser, ainsi que deux manuscrits de cours supplémentaires présents à la bibliothèque de Göttingen :

— Un manuscrit du cours sur la théorie des équations algébriques du WS 1886/87, donné en 1962 à la bibliothèque de Göttingen par le docteur Ernst Lamla, physicien, qui semble être soit une copie améliorée de celui de l'université de Strasbourg, soit reconstitué à partir des mêmes notes.

— Un manuscrit du cours sur les déterminants du SS 1886, donné aussi en 1962 à la bibliothèque de Göttingen par le docteur Ernst Lamla [35].

— Des notes de cours de Karl Runge, prises vraisemblablement entre 1878 et 1881, sur la théorie des équations algébriques. Ces notes sont en Gabelsberger [36], donc très difficilement accessibles. Cependant, les seules parties du cours prises en notes sont les rappels historiques sur la théorie des équations et la théorie des caractéristiques.

— Les notes (incomplètes) du cours du WS 1880/81 sur la théorie des équations algébriques d' Adolf Kneser.

Ces manuscrits sont référencés dans la bibliographie. Nous ne les utiliserons pas tous, mais l'essentiel de ceux concernant la théorie des équations algébriques seront exploités.

Même si la liste précédente ne correspond qu'aux manuscrits dont nous avons connaissance, nous pouvons tout de même voir se détacher les quatre « grands » cours que donnera Kronecker au cours de sa carrière :

1. La théorie des équations algébriques
2. La théorie des nombres / Arithmétique
3. La théorie des déterminants
4. La théorie des intégrales simples et multiples (trois manuscrits sur les cinq cours professés entre 1883 et 1891)

On peut aussi ajouter à cela son cours *Sur le concept de nombre en mathématiques* qu'il a donné en 1891, et que Jacqueline Boniface et Norbert Schappacher ont retranscrit et commenté dans la *Revue d'histoire des mathématiques* [37]. Il n'assura qu'une fois cet enseignement, l'année de sa mort. Il faudrait peut-être aussi s'interroger plus précisément sur ce que recoupe exactement le bloc « théorie des nombres / Arithmétique », mais pour

35. Bien que physicien, Ernst Lamla a rédigé une histoire de l'association de mathématiques de Berlin (*Mathematische Verein*), association née en 1861, qui contenait un grand nombre de notes de cours : il est possible que ces deux manuscrits en soient issus. Voir [Lamla, 1911].

36. Le Gabelsberger – nommé d'après son créateur Franz Xaver Gabelsberger – est un système de sténographie qui a été mis en place en Allemagne en 1834.

37. Voir [Boniface et Schappacher, 2001]




FIGURE 1 – Ex-Libris du cours de Kronecker sur la théorie des équations algébriques [Kronecker, 1879]

notre propos, cette qualification est suffisante.

Eugen Netto et Kurt Hensel ont édité, sous le titre de *Vorlesungen über Mathematik*, les trois derniers cours. Ils sont tous deux des anciens élèves de Kronecker, et Hensel a été chargé, comme nous l'avons vu, de faire paraître ses œuvres complètes. Netto a publié le premier volume des *Vorlesungen* : *leçons sur les intégrales simples et multiples* en 1893. Les *Leçons sur l'arithmétique générale* seront elles publiées en 1901 et les *Leçons sur la théorie des déterminants* en 1903 par Hensel. C'est donc au premier de ces cours, le seul à ne pas avoir été édité, que nous nous intéresserons ici, celui sur la théorie des équations algébriques.

#### 1.2.2.1 Comment était organisé le cycle des cours de Kronecker ?

Commençons par examiner ce que dit Weierstrass de la période pendant laquelle lui-même, Kummer et Kronecker enseignaient à l'université. Il a été alors possible [38]

38. nach einem umfassenden [...] Plan, dem sich später auch herangezogene jüngere Kräfte willig und erfolgreich anschlossen, den mathematischen Unterricht in der Weise zu organisieren, daß den Studierenden Gelegenheit gegeben ist, in einem zweijährigen Cursus eine beträchtliche Reihe von Vorträgen über die wichtigsten mathematischen Disziplinen in angemessener Aufeinanderfolge zu hören, darunter nicht wenige,

> d'organiser les cours de mathématiques d'après un plan, auquel se joindra plus tard aussi volontairement et avec succès l'appui de forces plus jeunes, de sorte que les étudiants aient la possibilité dans un cursus de deux ans d'écouter une suite considérable d'exposés sur les plus importantes disciplines des mathématiques dans un ordre convenable, y compris un certain nombre qui n'était pas ou peu pratiqué dans les autres universités. [Biermann, 1973] p. 102

L'ensemble des cours sera donc organisé sur deux ans [39], et Weierstrass, comme d'ailleurs Kronecker organiseront leur propre cycle à l'intérieur de ce cursus.

Nous possédons quelques informations sur l'organisation des cours de Kronecker. Ainsi, le journal interne de la *Königliche Technische Hochschule zu Berlin* publie en 1910 la nécrologie de l'un de ses *Privatdozent*, Eugen Meyer, qui a suivi à Berlin les cours de Kronecker [40] :

> Il a étudié d'abord à Gießen, ensuite à Berlin, où Pasch puis Kronecker ont exercé une influence déterminante sur son développement intellectuel. Kronecker l'a rapidement remarqué et a chargé le jeune étudiant au troisième semestre de la rédaction de ses leçons sur les déterminants, au semestre suivant de celle de sa grande leçon sur les équations algébriques, rédactions qui ont fourni un matériel précieux lors de la publication des *Vorlesungen* de Kronecker [der Technische Hochschule zu Berlin, 1910] p. 152

Deux semestres consécutifs sont donc consacrés aux déterminants, puis à la théorie des équations algébriques. Nous remarquons au passage que nous avons là une fois de plus un témoignage montrant que ces manuscrits que nous étudions ont bien été rédigés, au moins dans certains cas, à la demande de Kronecker. Hensel précise un peu les choses dans l'introduction du second tome qu'il a consacré à l'édition des leçons de Kronecker [41] :

> Les leçons sur la théorie des déterminants dont je rends ici public le premier volume, forment la deuxième partie du cycle des conférences académiques sur l'arithmétique générale, que Léopold Kronecker a tenues entre 1883 et 1891 à l'université de Berlin. La trame était répartie dans ces trois leçons sur la théorie des nombres, la théorie des déterminants et l'algèbre de sorte que chacune d'entre elles fasse un tout et puisse être comprise sans la connaissance des deux autres. [Kronecker, 1903] p. V

Il y a ainsi une volonté de rendre indépendant chacun des cours qu'un même enseignant donne : cela autorisera à ne professer un cours qu'une année sur deux, tout en permettant

die an anderen Universitäten gar nicht oder doch nicht regelmäßig gelesen werden.

39. Parfois une troisième année sera nécessaire pour terminer un cycle.

40. Er studierte zuerst in Gießen, wo Pasch, dann in Berlin, wo Kronecker einen entscheidenden Einfluß auf seine geistige Entwickelung ausübte. Kronecker wurde bald auf ihn aufmerksam und beauftragte den jungen Studenten im dritten Semester mit der Ausarbeitung seiner Vorlesungen über Determinanten, im folgenden Semester mit der seiner großen Vorlesung über algebraische Gleichungen, Ausarbeitungen, die bei der Herausgabe der Kroneckerschen Vorlesungen das wertvollste Material geliefert haben.

41. Die Vorlesungen über Determinantentheorie, deren ersten Band ich hiermit der Öffentlichkeit übergebe, bildeten den zweiten Teil in dem Cyklus akademischer Vorträge über allgemeine Arithmetik, welche Leopold Kronecker in den Jahren 1883 — 1891 an der Berliner Universität zu halten pflegte. Der Stoff war in diesen drei Vorlesungen über Zahlentheorie, Determinantentheorie und Algebra so verteilt, dass jede von ihnen ein Ganzes für sich bildete und ohne Kenntnis der beiden anderen vollständig verstanden werden konnte.

à l'ensemble des étudiants de les suivre. Si l'on regarde la liste des manuscrits précédente, nous voyons que le cours sur les déterminants (SS 84/86/88/90), la théorie des nombres (WS 81/83/85/87/89), et le cours sur la théorie des équations algébriques (WS 78/80/84/86/88/90) semblent correspondre à ce schéma. Nous connaissons de plus, grâce à E. Netto [42], l'existence de manuscrits contenant les cours de Kronecker sur les intégrales (WS 83/ SS 85/87/89/91). Nous obtenons ainsi une image à peu près complète de ce en quoi devait consister l'enseignement de Kronecker, tout au moins les dix dernières années de sa vie : il enseignait en alternance tous les deux ans *Théorie des nombres/Théorie des déterminants* et *Théorie des équations algébrique/Théorie des intégrales*. À cela pouvait s'ajouter parfois un cours de moindre importance.

#### 1.2.2.2 Qui a rédigé ces cours ?

Les manuscrits ne sont pas systématiquement signés, mais parfois un nom apparaît. Nous pouvons ainsi citer comme rédacteurs :

— *Georg Hettner* (1854–1914) : il soutient sa thèse, intitulée *Über die Reduction der Integrale einer besonderen Classe von algebraischen Differentialen auf die hyperelliptischen Integrale*, en 1877 sous la direction de Weierstrass et Kummer. Il sera professeur à la Technischen Hochschule à Berlin.

— *Carl David Tolmé Runge* (1856-1927) : en 1880, lui aussi sous la direction de Weierstraß et Kummer, il soutient sa thèse *Über die Krümmung, Torsion und geodätische Krümmung der auf einer Fläche gezogenen Curven*. Il sera professeur à Hanovre à partir de 1886, puis professeur de mathématiques appliquées de l'université Georg-August de Göttingen à partir de 1904.

— *Paul Stäckel* (1862-1919) : il soutient sa thèse en 1885 sous la direction de Leopold Kronecker et Julius Weingarten intitulée *Über die Bewegung eines Punktes auf einer Fläche*. Il sera professeur à l'université de Königsberg (außerordentlicher Professor de 1895 à 1897), à l'université de Kiel (ordentlicher Professor de 1897 à 1905), à l'université de Hannover (de 1905 à 1908), à la TH Karlsruhe (de 1908 à 1913) et à la Ruprecht-Karls-Universität de Heidelberg (de 1913 à 1919).

— *Karl Weltzien* : un étudiant de ce nom [43] passe sa thèse (*Über die Zerlegung einer ganzen homogenen Funktion von mehreren Veränderlichen in lineare Faktoren*) sous la direction de Weierstrass et Kummer en 1882. L'a-t-il soutenue une dizaine d'années après ses études ? Ou a-t-il mis en forme des notes de cours ayant une

42. [Kronecker, 1894] p. v.
43. Tout au moins d'après le *Mathematics Genealogy Project* (`http://genealogy.math.ndsu.nodak.edu/`).

dizaine d'années ? Nous n'avons pas pu le vérifier.

— *Karl Friedrich August Gutzmer* (1860-1924) : thèse (*Über gewisse partielle Differentialgleichungen höherer Ordnung*) en 1892 sous la direction d'Albert Wangerin. Il a suivi les cours de Kronecker à Berlin, et deviendra *Privatdozent* à l'université de Halle-Wittenberg en 1896, et professeur à l'université de Jena en 1899.

— Dans l'exemplaire de 1884-85 donné par le Dr Thiel apparaît le tampon du Dr. Schlesinger Lajos (Ludwig Schlesinger : 1864-1933) qui a suivi les cours de Kronecker à Berlin, et a passé sa thèse intitulée *Über lineare homogene Differentialgleichungen vierter Ordnung, zwischen deren Integralen homogene Relationen höheren als ersten Grades bestehen* sous sa direction et celle de Lazarus Fuchs en 1887. Il est donc possible que ce soit ce mathématicien d'origine hongroise, futur *Privatdozent* à Berlin, qui soit à l'origine des notes qui ont permis la réalisation de ce manuscrit.

Nous pouvons donner quelques indications sur un dernier rédacteur potentiel : un document de la *Königliche Technische Hochschule zu Berlin*[44], que nous avons déjà cité, nous apprend que Kronecker a chargé le jeune Eugen Meyer, futur *Privatdozent* dans cette même *Hochschule*, de prendre en note ses leçons sur les déterminants et la théorie des équations algébriques. Meyer n'arrivant à Berlin qu'en 1890, il ne peut donc s'agir que des *Vorlesungen über die Theorie der Determinanten* du semestre d'été 1890 et des leçons sur la *Theorie der algebraischen Gleichungen* du semestre d'hiver 1890/1891.

Il s'agit à chaque fois d'élèves de Kronecker qui soutiendront leur thèse deux à trois ans plus tard, et qui auront une carrière universitaire. Ce sont donc de jeunes mathématiciens en devenir, compétents, et maîtrisant certainement le contenu mathématique de ces cours. Nous les avons appelés « rédacteurs », mais on ne peut en fait qu'affirmer que ces manuscrits sont tirés de leurs notes : l'identité du calligraphe est incertaine. En effet, il n'était pas rare que l'on fasse appel à des copistes professionnels pour mettre en forme un manuscrit. Pour Runge, on peut comparer l'écriture du cours qui lui est attribué (*Vorlesungen* du Wintersemester 1881/82) avec des notes de sa main et il est tout à fait possible qu'il ait lui-même copié le manuscrit.

Pour quelques-uns d'entre eux, on trouve plusieurs rédacteurs. Par exemple, dans [Kronecker, 1885b] il y a au moins trois rédacteurs, et les changements ont lieu à chaque nouveau cours. On imagine sans peine les étudiants de Kronecker s'organiser pour mettre au propre à tour de rôle les notes prises à chaque séance.

#### 1.2.2.3 Une source crédible

Quel statut pouvons-nous donner à ces manuscrits ? Nous ne possédons pas les notes de cours de Kronecker, et nous n'avons presque aucune de ses lettres et très peu d'écrits

44. [der Technische Hochschule zu Berlin, 1910] p. 152.

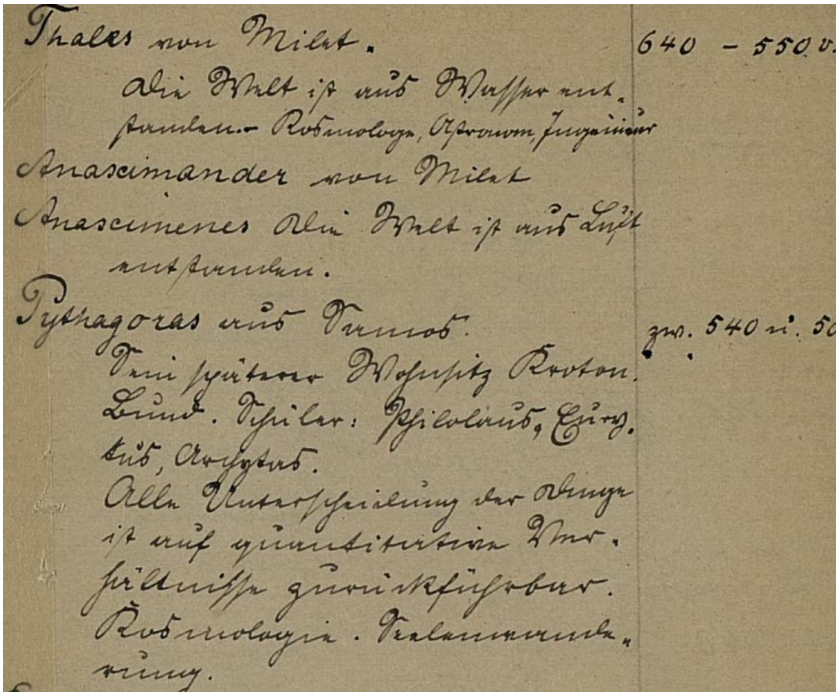


FIGURE 2 – Extrait des notes de Runge [Runge, 1881]

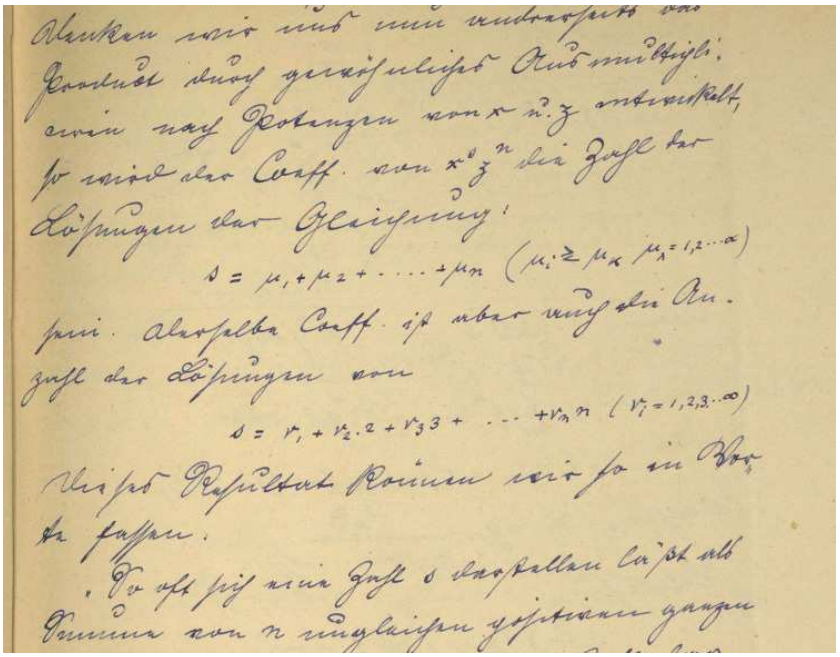


FIGURE 3 – Extrait d'un manuscrit du Wintersemester 1881/1882 [Kronecker, 1882a]

dans lesquels il a exprimé son point de vue sur ceux-ci [45]. Cependant, nous avons toutes les raisons de penser que ces manuscrits, s'ils ne sont pas de la main de Kronecker, correspondent à sa pensée. Déjà, les rédacteurs connus sont, comme nous l'avons vu, de futurs

45. Voir plus haut p. 24 pour tout ce qui concerne le *Nachlass* de Kronecker.

mathématiciens, tous aptes à retranscrire le déroulement des cours. Ensuite, la plupart de ceux-ci viennent du fonds Hensel : il n'y a donc presque aucun doute sur leur authenticité, et nous savons qu'ils ont servi, ou étaient destinés, à une édition des *Vorlesungen*. Enfin Kronecker a souhaité la rédaction de certains d'entre eux, en avait connaissance, et les prenait en compte dans la rédaction de ses articles [46] :

> Je voudrais dans ce qui suit associer quelques autres développements à mon essai publié dans le *Monatsbericht* de février 1873 et j'utiliserai les mêmes notations que dans ce mémoire sans les expliquer de nouveau. Sur ce point j'indique au préalable que j'ai déjà développé le contenu entier des trois premiers chapitres ainsi que l'essentiel de la dernière partie portant sur les équations du quatrième degré dans mes leçons à l'université en janvier et février 1875, et lors de la présentation se trouvant ici, je me suis bien tenu à la rédaction d'un de mes auditeurs de l'époque, le docteur *Hettner*. [Kronecker, 1878c] p. 39

Nous avons vu plus haut d'autres témoignages des interactions entre Kronecker et ces manuscrits [47] : ils sont rédigés à la demande de Kronecker, qui choisit ses étudiants les plus brillants pour effectuer cette tâche. Nous estimons donc que ces manuscrits doivent être très représentatifs, sinon du contenu réel, tout au moins du contenu que Kronecker souhaitait faire apparaître dans ses cours. C'est pourquoi, sauf cas particulier, nous choisissons d'attribuer à Kronecker le contenu de ces manuscrits.

## 1.3 Les *Vorlesungen über die Theorie der algebraischen Gleichungen*

Nous possédons dix manuscrits de la « théorie des équations algébriques », qui correspondent à sept sessions différentes : 72/73, 78/79, 80/81, 85/86, 87/88, 88/89, 90/91. Il s'agit donc de cours manuscrits rédigés à partir de notes de cours d'élèves de Kronecker. La plupart des rédactions, mais pas toutes, sont en *Kurrentschrift* [48], et font entre 200 et 700 pages chacune. Pour les passages que nous avons étudiés, il y a très peu d'erreurs, et rien qui puisse ressembler aux erreurs d'un copiste inattentif. Même si nous nous autoriserons à utiliser tous les manuscrits, nous choisirons, selon notre but, d'examiner plus précisément l'un ou l'autre d'entre eux. La quantité, plus ou moins 5000 pages manuscrites, ne nous permet pas ici de les exploiter en totalité. Nous allons tout de même en donner un aperçu général, auquel succédera une description plus détaillée de l'un d'entre eux [49].

Il existe un dernier manuscrit à la bibliothèque de Strasbourg intitulé *Theorie der algebraischen Gleichungen : Theil II* et daté du semestre d'été 1885. C'est d'ailleurs le seul de ces cours à avoir eu lieu au semestre d'été. Dans ces leçons, Kronecker développe une

46. Ich will im Folgenden an meinen im Monatsbericht, vom Februar 1873 veröffentlichten Aufsatz einige weitere Entwickelungen knüpfen und werde dabei die dort angewendeten Bezeichnungen benutzen, ohne dieselben von Neuem zu erklären. Hierzu schicke ich die Bemerkung voraus, dass ich den ganzen Inhalt der ersten drei Abschnitte sowie das Wesentlichste des auf die Gleichungen vierten Grades bezüglichen letzten Theils schon in meinen Universitäts-Vorlesungen im Januar und Februar 1875 ausgeführt und bei der hier vorliegenden Darstellung desselben mich genau an die Ausarbeitung eines meiner damaligen Zuhörer, des

théorie de l'élimination, dont la présentation générale consiste à observer des intersections de variétés et la notion centrale est celle de résultant. Nous verrons que ce cours correspond à la dernière partie des cours précédents (1872-73 et 1878-79).

Nous ne possédons pas de manuscrit de son cours sur la théorie des équations algébriques datant d'avant 1872. Cependant, dans un article de 1881 intitulé *Ueber die Discriminante algebraischer Functionen einer Variabeln*, Kronecker affirme [50] avoir déjà présenté une grande partie des résultats de cet article dans ses cours des *Wintersemester* 1865/66 et 1870-71. Or on retrouvera ces notions dans l'ensemble des manuscrits à notre disposition : Kronecker semble bien avoir commencé à travailler sur ces leçons dès ses premières années d'enseignement.

### 1.3.1 La structure des manuscrits

La structure des *Vorlesungen über die Theorie der algebraischen Gleichungen* évolue au fil des années, en fonction des expériences d'enseignement accumulées, des thèmes de recherche de Kronecker. Dans cette section, nous essayerons de conserver un point de vue assez large : après avoir exposé, sous la forme d'un sommaire, les grandes parties de chacun des sept cours, nous essayerons de donner quelques mots d'explication des changements qui apparaîtront. Nous allons présenter les grandes parties qui structurent ces cours. Ce sera aussi l'occasion pour nous d'exposer rapidement les caractéristiques de chacun des manuscrits.

**Le manuscrit du semestre d'hiver 1872-73 [Kronecker, 1873a]** Il s'agit du manuscrit le plus ancien que nous possédons des *Vorlesungen über die Theorie der algebraischen Gleichungen*, mis en forme par Carl Weltzien. Ce manuscrit, écrit en *Kurrentschrift*, comporte 494 pages. La fin des cours est datée du 13 mars 1873.

1. 1 Introduction historique
2. 4 Équations quadratiques et cubiques
3. 8 Équations du degré $n$ et fonctions symétriques
4. 122 Théorème de Sturm / Principe d'inertie de Sylvester
5. 168 Théorie des caractéristiques
6. 196 Théorie de l'élimination et Théorie des grandeurs algébriques (Résolvante de Galois,...)

Hrn. Dr. *Hettner*, gehalten habe.

47. Voir p. 40 et 47.

48. Le *Kurrentschrift* correspond à la forme manuscrite (et cursive) des diverses formes de gothique utilisées en Allemagne. Elle disparaît progressivement au XX$^{e}$ siècle.

49. À partir de maintenant, lorsque nous nous réfèrerons à ces manuscrits, nous utiliserons indifféremment pour les désigner : « *Vorlesungen* », « Vorlesungen über die Theorie der algebraischen Gleichungen », « *Theorie der algebraichen Gleichungen* » ou même « Théorie des équations algébriques ».

50. [Kronecker, 1881b] p. 301.

**Les manuscrits des semestres d'hiver 1878-79 et 1880-81** Ces deux manuscrits suivent un même plan [51]. Le manuscrit de 1878-79 est le plus détaillé et le plus long. Celui de de 1880-81, plus synthétique, plus court, se termine juste après la théorie des caractéristiques, et ne traite donc pas de tout ce qui concerne la théorie des grandeurs algébriques.

**Le manuscrit du semestre d'hiver 1878-79 [Kronecker, 1879]** Ce manuscrit n'est pas écrit en *Kurrentschrift* et contient 787 pages.

1. 1 Introduction historique
2. 2 Équations quadratiques et cubiques
3. 25 Équations du degré $n$ et fonctions symétriques
4. 296 Théorème de Sturm / Principe d'inertie de Sylvester
5. 482 Théorie des caractéristiques
6. 566 Caractéristique de $n+1$ fonctions de $n$ variables
7. 622 Développement autour de la théorie de l'élimination
8. 701 Théorie des grandeurs algébriques (Résolvante de Galois, ...)

**Le manuscrit du semestre d'hiver 1880-81 [Kronecker, 1881a]** Ce manuscrit est écrit en *Kurrentschrift* et ne contient que 170 pages. Sa rédaction, peut-être incomplète [52], se termine par la théorie des caractéristiques. Il y a un seul rédacteur.

1. 2 Introduction historique
2. 8 Équations quadratiques et cubiques
3. 17 Équations du degré $n$ et fonctions symétriques
4. 104 Théorème de Sturm / Principe d'inertie de Sylvester
5. 150 Théorie des caractéristiques

**Les manuscrits des semestres d'hiver 1884-85** Nous avons en notre possession deux manuscrits du cours de 1884-85. Ce cours est le premier qui suit la publication des *Grundzüge*. Le manuscrit est composé de deux grandes parties : la première qui se termine par le théorème de Sturm et la seconde qui décrit la théorie des *Gattungen*.

**Le manuscrit du semestre d'hiver 1884-85 du fonds Hensel [Kronecker, 1885b]** Ce cours commence le lundi 27 octobre 1884 et comporte 572 pages numérisées, dont 519 représentent le cours proprement dit rédigé sur les pages de droite, le reste correspondant aux ajouts sur la page de gauche de dessins ou de remarques, l'ensemble étant en *Kurrentschrift*. Il y a plusieurs rédacteurs [53] qui alternent d'une leçon à l'autre. Ce manuscrit a une particularité : son introduction est beaucoup

51. La numérotation du manuscrit de 1880-81 que l'on utilise débute à l'introduction : elle correspond donc à la page $n-2$ de la version PDF téléchargeable sur le site de la Bibliothèque numérique patrimoniale de Strasbourg.

52. Dans sa fiche descriptive de l'université de Strasbourg il est précisé qu'il « fait partie d'un ensemble de 2 tomes », bien que l'on n'ait aucune précision sur le second tome.

53. Le premier pp. 1-27, le second 47-75 ; 101-114 ; 146-162 ; 194-208 ; 236-248 ; 279-300 ; 334-369 ; 424-476 le troisième pp. 27-47 ; 75-101 ; 114-146 ; 163-193 ; 208-235 ; 248-279 ; 300-333 ; 369-424 ; 476-519.

plus longue. Celle-ci semble contenir l'ensemble du discours de Kronecker, sans qu'aucun tri n'ait été opéré : de long passages ont d'ailleurs été barrés, très certainement en vue de la constitution d'une édition.

1. 1 Introduction historique
2. 38 Définitions préalables
3. 57 Équations quadratiques et cubiques
4. 114 Équations du degré $n$ et fonctions symétriques
5. 236 Le théorème de Sturm / Principe d'inertie de Sylvester
6. 344 Théorie des *Gattungen*

**Le manuscrit du semestre d'hiver 1884-85 du fonds Hensel donné par le Dr. Thiel [Kronecker, 1885a]** Ce manuscrit de 407 pages, rédigé en *Kurrentschrift*, s'achève le 6 mars 1885. Deux notes sont ajoutées en fin d'ouvrage : la première le 22 octobre 1826 (signature indéchiffrable), la seconde le 5 mars 1889 (il est indiqué : *Bemerkung nach Kronecker* [Kronecker, 1885a]).

1. 1 Définitions préalables
2. 6 Équations quadratiques et cubiques
3. 52 Équations du degré $n$ et fonctions symétriques
4. 149 Théorème de Sturm / Principe d'inertie de Sylvester
5. 240 Théorie des *Gattungen*

Le plan reste très proche de celui des manuscrits de 1878-79 et 1880-81. Les notations sont d'ailleurs inchangées, ce qui facilite la comparaison. La principale différence est contenue dans la partie *Divers développements autour des formes quadratiques*, qui porte sur diverses techniques de transformations effectives des formes quadratiques, à rapprocher du travail de Kronecker sur les formes bilinéaires [54].

**Les manuscrits du semestre d'hiver 1886-87** Pour ce semestre aussi nous avons pu comparer deux manuscrits : celui conservé à Strasbourg et celui offert par le docteur Lamla à Göttingen. On trouve dans la marge la date de la prise des notes : cela nous permettra d'évaluer le temps consacré par Kronecker à chaque partie. Les cours ont lieu le lundi, le mardi et le jeudi.

**Le manuscrit du semestre d'hiver 1886-87 du fonds Hensel [Kronecker, 1887a]** Ce manuscrit composé de 325 pages écrites en *Kurrentschrift* débute le premier novembre 1886 et se termine le 3 mars 1887.

1. 1 Introduction historique
2. 22 Définitions préalables
3. 30 Équations quadratiques et cubiques
4. 62 Équations du degré $n$ et fonctions symétriques
5. 126 Théorème de Sturm / Principe d'inertie de Sylvester
6. 200 Théorie des *Gattungen*

54. Voir par exemple [Kronecker, 1866], [Kronecker, 1868], [Kronecker, 1872], [Kronecker, 1874c], [Kronecker, 1874a] ou [Kronecker, 1874b].

**Le manuscrit du semestre d'hiver 1886-87 donné à l'université de Gottingen par le Dr. Lamla [Kronecker, 1887b]** Unique exemplaire ne provenant pas de la Bibliothèque de Strasbourg, ce manuscrit contient 422 pages rédigées en alphabet latin.

1. 1 Introduction historique
2. 15 Définitions préalables
3. 21 Équations quadratiques et cubiques
4. 53 Équations du degré $n$ et fonctions symétriques
5. 131 Théorème de Sturm / Principe d'inertie de Sylvester
6. 232 Théorie des *Gattungen* et de leurs discriminants
7. 329 La résolution algébrique des équations
8. 382 Les équations *Abéliennes*
9. 418 Sur le remplacement des grandeurs algébriques par des fonctions rationnelles à l'aide des modules-systèmes

**Le manuscrit du semestre d'hiver 1888-89 [Kronecker, 1889b]** Les cours ont lieu cette année-là le lundi, le mardi, le jeudi et le vendredi. Il y a donc une séance de plus par semaine par rapport à l'année précédente. Ce cours est aussi muni d'une table des matières (pp. 299-301 [Kronecker, 1889b]) et contient 316 pages rédigées en *Kurrentschrift*.

1. 1 Introduction historique
2. 2 Définitions préalables
3. 38 Équations quadratiques et cubiques
4. 68 Équation du degré $n$ et fonctions symétriques
5. 120- Le théorème de Sturm (dont la théorie des caractéristiques)
6. 165 Théorie des *Gattungen*
7. 237 Les équations algébriques résolubles
8. 266 Les équations abéliennes

**Le manuscrit du semestre d'hiver 1890-91 [Kronecker, 1891a]** Les cours ont lieu le lundi, le mardi et le jeudi : un jour de moins que lors de la session précédente. Il s'agit de son dernier cours sur la théorie des équations algébriques, et le manuscrit – rédigé en *Kurrentschrift* – contient un sommaire détaillé qui permet de se repérer facilement dans l'ensemble du cours. Il est composé, pour sa version numérique, de 666 pages. Mais seules 420 d'entre elles contiennent le cours, qui a été écrit uniquement sur la page de droite des feuillets. La page de gauche a été utilisée, rarement, pour des schémas. On trouve aussi sur ces pages les commentaires d'un lecteur [55].

1. 1 Introduction historique
2. 5 Définitions préalables (avec en plus *Rationalitäts* et *Integritäts Bereich*)
3. 8 Racine d'une équation
4. 20 La théorie des équations quadratiques (Principe de Galois, fonctions symétriques, fractions continues)
5. 69 Théorie des Fonctions cubiques (Fonctions cycliques, abéliennes, symétriques)
6. 121 Théorie des équations algébriques de degré $n$ (Fonctions symétriques)
7. 170 Le théorème de Sturm
8. 208 Théorie des caractéristiques
9. 278 Théorie des *Gattungen*
10. 366 Les équations algébriques résolubles (équations abéliennes)

55. Nous ne connaissons pas l'identité du ou des auteurs de ces commentaires. Cependant, il est indiqué que l'une des figures est de *Dyck*. Nous savons (voir plus loin p. 440) que Walther Dyck a eu certains de ces manuscrits entre les mains.

**L'organisation interne de ces cours** Il est difficile de trouver des documents qui nous donneraient des indices sur la façon dont se déroulaient exactement les cours de Kronecker à l'université de Berlin. Cependant, dans l'un des manuscrits des *Vorlesungen über die Theorie der algebraischen Gleichungen*[56], les « apartés » de Kronecker sur l'organisation interne de ses leçons ont été retranscrits, nous apportant ainsi un précieux témoignage sur le déroulement de ses cours. Commençons par donner une traduction de ce passage[57] :

> Je vais introduire à cet endroit encore quelque chose d'externe. Je vais utiliser ici une assez grande plage de temps ; j'ai besoin de celle-ci car je souhaite non seulement en[58] donner un développement très vaste, mais aussi les approfondir et les éclairer sous différentes lumières. C'est pourquoi j'ai annoncé six heures pour cette leçon. Je donne habituellement trois fois deux heures. Cela n'est pas faisable cette fois à cause d'un accord que j'ai passé avec mon ami Weierstrass. Je souhaite enseigner maintenant ces deux fois deux heures en une fois, c'est à dire à la suite. Cela a des avantages majeurs mais aussi des inconvénients. Il ne faut pas ignorer la difficulté, non seulement pour

56. Voir [Kronecker, 1885b].

57. Ich werde an dieser Stelle noch etwas Äußerliches anführen. Ich gebrauche hier eine möglichst große Zeit ; ich brauche dieselbe, weil ich gern nicht bloß eine weit reichende Entwicklung geben möchte, sondern sie auch recht vertiefen und durch mannigfache Beleuchtung recht deutlich machen möchte. Darum habe ich diese Vorlesung sechsstündig angekündigt. Ich habe früher dreimal zwei Stunden gelesen. Das ließ sich diesmal nicht machen wegen der Uebereinkunft, welche ich mit meinem Freunde Weierstraß zu machen hatte. Diese zweimal je zwei Stunden möchte ich nun durchlesen d.h. hintereinander fort. Das hat wesentliche Vorteile, aber auch Nachteile. Die Schwierigkeit ist nicht zu verkennen, die es nicht nur für mich, sondern auch für den Hörer hat, längere Zeit die Aufmerksamkeit auf solche abstracten Dinge zu concentrieren. In allen Dingen der Welt muß man sich aber mit Compromissen begnügen. Ein solcher ist hier zu schlissen zwischen den Schwierigkeiten und den Vorteilen, die ich noch zu erwähnen habe. Ein Stunde ist noch den Gewohnheiten der Universität nicht eine Stunde, sondern höchstens 40 Minuten. Das ist ein etwas enger Raum für manche längeren Entwicklungen und nicht bloß, daß, wenn der Faden abgerissen wird, in der nächsten Stunde es schwierig ist, schwieriger ihn wieder aufzunehmen, als ihn fortzuspinnen, sondern es geht oft gar nicht und man muß ein ganzes Stück zurückgreifen, um ihn wieder anknüpfen zu können, so daß eigentlich von den 40 Minuten noch so weitere verloren gehen auf Wiederholungen. Ich weiß das weniger aus meiner Erfahrung, denn viele Jahre hindurch habe ich mich an diese Art des Durchlesens gewöhnt, vielmehr aus Erfahrungen meiner Studentenzeit. Meine berühmter Lehrer Dirichlet laß eine große zahlentheoretische Vorlesungen vierstündig in vier einzelnen Stunden. Er laß die eine, welche ich gehört habe und noch genau in der Erinnerung trage, ~~des morgens~~ morgens von sieben bis acht, eine sehr frühe Stunde für Berlin. Ich hatte nun im Sommer noch früher etwas zu thun und das bewirkte, daß ich erst um halbe acht ins Colleg eintrat. Es war ein Kleines ~~Colleg~~ Auditorium, wir waren im ganzen dreizehn Zuhörer. Ich ging mit Dirichlet stets nach Hause und er sagte mir, daß er es ganz bequem und richtig fände, daß ich erst um halb acht käme, denn dann gerade fange er das Neue an. Damit wurden aus den sogenannten vier Stunden zwei. Welcher Universitätslehrer würde es heute wagen, die ganze Zahlentheorie in dieser Zeit zu lesen. Da liegt freilich noch etwas anderes hinter. Die wenigen Mathematiker nämlich, die es damals gab, waren an ein rasches Lernen gewöhnt, oder sie mussten sich dann gewöhnen. Dirichlet gab ungefähr alles, was von Dedekind herausgegeben ist, wirklich in einem kurzen Sommersemester, aber er sprach sehr rasch und deutete die Formeln kaum an, alles auszuführen, war uns überlassen. Ganz anders ist man heute gewohnt, Vorlesungen zu hatten und zu hören und es ist gewiss ein Fortschritt, dass wir heute im Stande sind, bei längerer, ausführlicherer Methode denselben Stoff in viel längerer Zeit, aber einem viel größerem mathematischen Publikum klar zu machen, und auch ich muss mich in die jetzige Weise schicken. Ich gebrauche eine längere Zeit als Dirichlet, wozu nun auch kommt, daß der Inhalt dieser Theorie des algebraischen Gleichungen doch noch viel mehr umfasst, als diese Elemente des Zahlentheorie. Ich möchte mit deshalb in der That Ihre Zustimmung dazu erbitten, denn diese Zustimmung brauche ich, daß ich an diesen zwei Tagen durchlese. Wir haben dann noch kurze Stunden genug an den beiden andern Tagen. Sie werden sehe, daß man sich bald daran gewöhnt, ja und so mehr, als ich mir und Ihnen vorbehalten muss in den Zeiten und in den Gegenden, wo es für beide Teile besonders anstrengend ist, einmal früher abzuschließen.

58. Kronecker se réfère ici au contenu de son cours.

> moi mais aussi pour les auditeurs, de concentrer son attention pendant longtemps sur de telles abstractions. Dans toutes les choses de ce monde, il faut se contenter cependant de compromis. Ici il y en a un, entre des difficultés et des avantages que je dois encore présenter. Une heure fait habituellement à l'université non pas une heure mais plutôt au maximum 40 minutes. C'est un temps un peu restreint pour certains développements longs, et pas seulement parce que quand le fil est rompu il est difficile, voire même plus difficile de le reprendre que de le garder, mais plutôt parce que souvent cela ne va pas du tout et il faut reprendre toute une partie de ce qui précède pour pouvoir à nouveau relier le fil de la pensée, de telle sorte qu'une partie des 40 minutes est encore perdue dans des répétitions. Je sais cela moins de par mon expérience, car je me suis habitué pendant de longues années à cette sorte de leçons, mais plus par mon expérience d'étudiant. Mon fameux professeur donnait une grande leçon sur la théorie des nombres de quatre heures en quatre heures séparées. Il donnait un cours, que j'ai suivi et qui me revient précisément en mémoire, tôt le matin, de sept à huit, une heure très matinale pour Berlin. J'avais quelque chose à faire plus tôt en été, et cela avait pour effet que je n'arrivais au Collège qu'à sept heure trente. C'était un petit auditorium du Collège, nous étions treize auditeurs en tout. Je faisais toujours le chemin du retour avec Dirichlet, et il m'avait dit qu'il trouvait cela tout à fait commode et correct que je vienne seulement à sept heure trente, car il commençait alors ce qui était nouveau. Ainsi les prétendues quatre heures en devenaient deux. Quel enseignant d'université oserait aujourd'hui présenter la théorie des nombres en entier dans ce temps ? Il y a là cependant quelque-chose d'autre derrière. Le peu de mathématiciens de l'époque avaient l'habitude d'un apprentissage rapide, ou ils devaient s'y habituer. Dirichlet présentait à peu près tout ce que Dedekind a publié réellement en un court semestre d'été, mais il parlait très vite, indiquait à peine les formules, et il nous laissait les développements à faire. C'est différent aujourd'hui où l'on est habitué à avoir, et à écouter, des leçons que l'on est en mesure de présenter clairement et plus longuement, pour un même contenu, par des méthodes plus complètes à un public de mathématiciens beaucoup plus nombreux, et je dois moi aussi m'adapter à ces nouvelles méthodes. J'ai besoin d'un temps plus long que Dirichlet, le contenu de cette théorie des équations algébriques englobant bien plus que ces éléments de la théorie des nombres. C'est pourquoi je souhaite en effet demander votre accord, car j'en ai besoin pour donner une leçon ces deux jours. Nous avons encore assez de leçons courtes les deux autres jours. Vous verrez que l'on s'habitue vite, et ce d'autant plus que vous et moi gardons, lorsque cela sera particulièrement fatigant pour les deux parties, la possibilité de finir éventuellement plus tôt. [Kronecker, 1885b] pp. 7-8

Les cours de Kronecker sur la théorie des équations algébriques sont donc habituellement de six heures hebdomadaires, réparties sur trois jours de la semaine. D'après E. Netto, la durée des cours pouvait être de deux heures, quatre heures ou six heures [59] : il s'agit donc là du plus grand format classique de cours proposé par l'université de Berlin, et donc de l'un des « grands cours » que pouvait y suivre un étudiant.

59. [Kronecker, 1894] p. V.

### 1.3.2 Permanence de la structure

Lorsque l'on observe les grands chapitres des différents manuscrits, on trouve finalement peu de changements au cours des années. À une exception près[60], le premier chapitre, plus ou moins développé, consiste en une introduction retraçant l'histoire de l'algèbre. Nous en détaillerons un peu plus loin le contenu. Ensuite, Kronecker fait une introduction « pratique », qu'il appelle parfois *sachliche Einleitung*[61] dans laquelle il donne quelques définitions, mais exprime aussi son point de vue sur certaines notions. Si le contenu de cette courte partie évolue au fil des ans, elle sera toujours présente.

Une fois ces deux introductions terminées, Kronecker commence systématiquement par présenter les équations du second et du troisième degré. Cette exposition, qui s'étend de 4 à 100 pages, en expose évidemment la résolution, mais est aussi utilisée par Kronecker pour mettre en évidence une méthode, pour introduire un nouveau concept, etc.

Le chapitre suivant est consacré au passage aux équations de degré $n$. Le travail de Kronecker portera essentiellement sur la manipulation des racines des équations, et en particulier sur les polynômes symétriques des racines des équations. Les outils construits dans cette partie seront fondamentaux pour la suite de son cours, et lui permettront par exemple, comme l'a déjà fait Gauss en 1816, de produire une démonstration du théorème fondamental de l'algèbre. La première grande partie de son cours se termine alors par un chapitre sur le théorème de Sturm qui contient, lorsqu'elle est présente, la théorie des caractéristiques.

C'est à cette première partie que notre travail est consacré. La seconde, dont le volume et le contenu évoluent à chaque nouveau semestre, contient dès l'hiver 1878-79 des éléments de la théorie des *Gattungen*, théorie qui prendra toute sa place après la publication des *Grundzüge* en 1882. On voit donc immédiatement que la structure, lorsqu'on la regarde dans son ensemble, évolue très peu :

1. Introduction historique et définitions préalables
2. Équations quadratiques et cubiques
3. Équation de degré $n$ et fonctions symétriques
4. Le théorème de Sturm (avec ou sans la théorie des caractéristiques)
5. Théorie des *Gattungen* (ou théorie de l'élimination avant 1882, mais toujours avec les équations abéliennes et les principes de Galois)

### 1.3.3 Évolution du contenu

Si la structure semble immuable, ce n'est pas le cas du contenu. Comme nous l'avons vu, Kronecker expose ses recherches en cours dans les cours qu'il professe à Berlin. Ainsi, par exemple, le contenu de son article de 1878 sur la théorie des caractéristiques[62] n'est pas

60. Le manuscrit de 1884-85 donné par le Dr. Thiel [Kronecker, 1885a].
61. Voir par exemple [Kronecker, 1885b] p. 38.
62. Voir [Kronecker, 1878b].

présent dans le manuscrit de 1872-73, apparaît dans celui de 1878-79, disparaît à partir de 1884 pour réapparaître en 1890-91 : selon les préoccupations mathématiques de Kronecker, certains contenus peuvent être plus ou moins développés.

De même, la publication de son article de 1882 *Grundzüge einer arithmetischen Theorie der algebraischen Grössen* marque un tournant dans la présentation qu'il donne de sa théorie des équations algébriques : il introduit la théorie des *Gattungen*, et réinterprète tout son travail sur Galois et Abel dans les termes de ses nouveaux concepts.

### 1.3.4 Une introduction historique

La presque totalité des cours de Kronecker[63] commence par une introduction, plus ou moins longue, dans laquelle il donne une approche historique de la question qu'il souhaite traiter. Cette approche n'est pas systématique[64] chez les autres enseignants de la *Friedrich-Wilhelms-Universität*, ni d'ailleurs dans les cours publiés dans le reste de l'Europe à la même époque[65]. Ainsi, la présence d'un véritable *cours d'histoire des mathématiques* au début de ses leçons est particulièrement remarquable. Comme de plus, pour Kronecker, « il n'y a pas d'autre discipline qui contienne comme l'algèbre un enseignement aussi riche dans son histoire »[66], nous allons essayer de dire quelques mots de la place et du contenu de ses remarques historiques dans ses *Vorlesungen über die Theorie der algebraischen Gleichungen*.

Il n'y a pas d'évolution à proprement parler pour la partie historique de l'introduction, mais selon les années et les manuscrits, elle est plus ou moins développée et plus ou moins indépendante de la partie consacrée à l'introduction des notations et des concepts. Cette introduction, qui peut dépasser dans certains manuscrits les quarante pages, s'étale quelquefois sur plusieurs cours.

Ce que Kronecker souhaite y faire, c'est présenter « à grands traits l'histoire de l'algèbre, dans laquelle toute l'histoire des mathématiques non mécaniques et géométriques est résumée »[67].

Le cours de Kronecker commence par un passage en revue des travaux sur la théorie des équations algébriques depuis Euclide. Il met en avant deux périodes dans cette histoire : avant et après les années 1770. Cette date, nous l'avons vu plus haut[68], correspond à la publication de deux traités, l'un de Vandermonde et l'autre de Lagrange[69].

Avant cette date, il présente une histoire de l'algèbre qui est une histoire de la

63. Tout au moins pour tous ceux auxquels nous avons eu accès (Voir plus haut p. 43).

64. Ce n'est pas le cas, par exemple, pour les cours de Weierstrass [Weierstrass, 1881], [Weierstrass, 1882c], [Weierstrass, 1882d] et [Weierstrass, 1883].

65. Voir par exemple : [Francoeur, 1838], [Serret, 1849]), [Salmon, 1885], [Weber, 1895] ou [Netto, 1900].

66. [Kronecker, 1885b] p. 3 : « Es enthält nun keine andere Disziplin, wie die Algebra, Lehrreicher in ihrer Geschichte. »

67. Une transcription et une traduction de l'introduction de son cours sur les équations algébriques du semestre d'hiver 1880-81 sont données en annexe.

68. Voir p. 64.

69. [Vandermonde, 1771] et [Lagrange, 1771].

résolution algébrique des équations, chaque étape permettant la résolution générale de l'équation de degré supérieur, jusqu'à celle de degré quatre par Ferrari. Kronecker distingue deux problèmes majeurs qui, d'après lui, « ont occupé de préférence pendant 2000 ans les algébristes » [70]. Le premier, théorique, concerne la résolubilité par radicaux, résolue négativement d'après lui en 1826 par Abel. Un second, pratique, touche au nombre de racines réelles d'une équation entre deux bornes, problème auquel Sturm a répondu en 1829. Kronecker ne relève donc pas, même s'il en a parlé précédemment, le théorème fondamental de l'algèbre comme un problème qui aurait occupé « de préférence les algébristes » [71]. Et de fait, les premières tentatives de démonstration de ce théorème ont pour Kronecker à peine deux siècles.

Cette introduction historique est un exposé chronologique, qui commence en traitant les mathématiques de Euclide à la Renaissance. Kronecker montre une connaissance approfondie de l'histoire des mathématiques [72], et s'intéresse plus particulièrement aux recherches qui ont eu lieu au XVIII$^{e}$ et XIX$^{e}$ siècle. Il faut enfin noter la présence prédominante de Gauss, dont les travaux et les idées semblent avoir particulièrement marqué Kronecker.

Pour Kronecker, comme le montre la lecture de chacune des introductions des *Vorlesungen über die Theorie der algebraischen Gleichungen*, ce cours traite d'algèbre. Mais nous savons que le terme *algèbre* peut renfermer des domaines et concepts extrêmement variés. Nous allons donc nous efforcer maintenant de comprendre ce que recouvre ce terme pour Kronecker.

70. [Kronecker, 1881a] p. 5.
71. [Kronecker, 1881a] p. 5.
72. Ces deux manuscrits ne reflètent pas convenablement l'étendue de ses connaissances : celui de 1884-85 fait presque quarante pages.

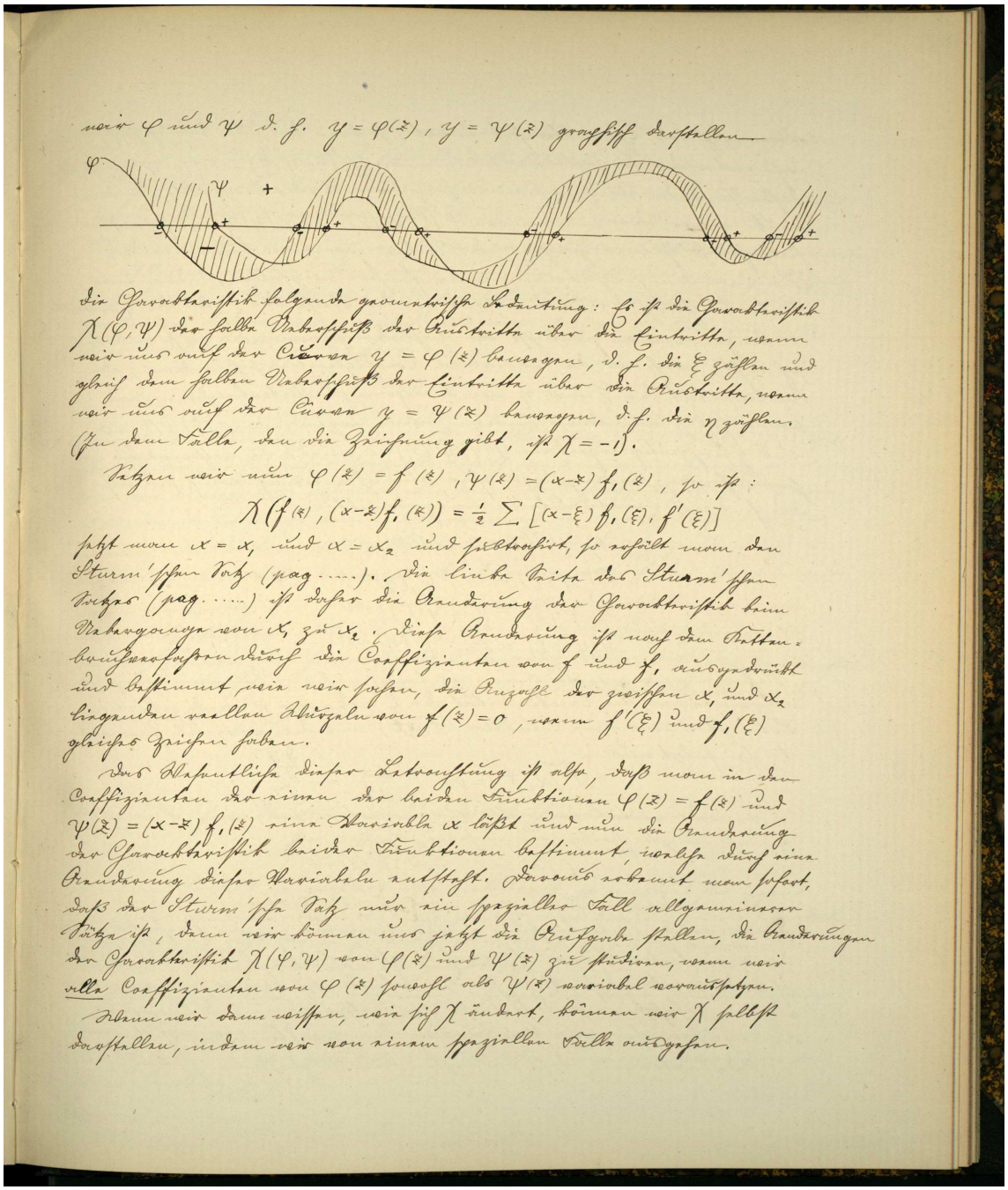


FIGURE 4 – Cours de Kronecker sur la théorie des équations algébriques [Kronecker, 1881a] p. 157

# Chapitre 2

# L'*Algèbre* de Kronecker


2.1 Quelle Algèbre ? . . . . . **63**
2.1.1 L'algèbre avant Kronecker . . . . . 63
2.1.2 L'algèbre dans les *Vorlesungen über die Theorie der algebraischen Gleichungen* . . . . . 70
2.2 Les racines d'une équation . . . . . **78**
2.2.1 La notion de racine . . . . . 79
2.2.2 La séparation des racines . . . . . 84
2.2.3 Le théorème de Bolzano . . . . . 91


On sait qu'entre le début et la fin du XIX^e^ siècle, de profonds changements s'opèrent dans ce qui représentera le terme d'*algèbre* pour les mathématiciens. Non seulement les méthodes, mais aussi les objets même de l'algèbre changent, au point qu'il est extrêmement délicat d'utiliser ce terme pour désigner un contenu spécifique sans tenter de le contextualiser. Au début du XIX^e^ siècle, l'algèbre correspond pour la plupart des mathématiciens à l'étude des équations polynomiales. Cependant, les formes que prennent ces études évoluent tellement au cours du siècle que l'algèbre de 1900 a finalement peu à voir, de par ses méthodes et ses objets, avec celui de 1800. Kronecker, qui produit ses mathématiques pendant la seconde moitié du XIX^e^ siècle, participe activement à cette évolution, et nous allons essayer de rendre compte de sa vision de l'algèbre en partant de celles, diverses, de certains mathématiciens qui l'ont influencé.

## 2.1 Quelle Algèbre ?

### 2.1.1 L'algèbre avant Kronecker

Il n'est pas question ici de donner une histoire de l'algèbre, tâche qui dépasserait bien évidemment le cadre de cette thèse, et nous ne souhaitons pas non plus exposer quelques propos généraux qui seraient, au mieux, imprécis. Ce qui constitue l'algèbre au XIX^e^ siècle dépend, comme l'ont montré C. Ehrhardt et F. Brechenmacher dans [Brechenmacher et Ehrhardt, 2010], des acteurs eux-mêmes et des réseaux dans lesquels ils

évoluent. Nous avons donc choisi de commencer par examiner chez certains des prédécesseurs de Kronecker ce que recouvre le terme d'*algèbre*. Pour cela, nous avons commencé par examiner la liste des mathématiciens qu'il cite dans l'introduction de son cours [1]. Cependant, l'« ère » à laquelle il se réfère comme étant celle qui précède sa propre génération correspond à la fin du XVIII^e siècle et au début du XIX^e, et nous avons ainsi réduit notre étude aux mathématiciens cités de cette période [2]. Kronecker propose d'ailleurs 1770 comme début de cette période [3], c'est-à-dire la publication des célèbres mémoires [4] de Lagrange et de Vandermonde sur la résolution des équations : ces deux mathématiciens feront l'objet de notre premier paragraphe. Nous connaissons l'importance de l'influence de Gauss sur le travail des mathématiciens du XIX^e siècle, et en particulier celle des *Disquisitiones Arithmeticae*, oeuvre que Kronecker qualifiait de « livre de tous les livres » [5]. Nous poursuivrons ainsi notre étude avec Gauss, et nous terminerons en exposant les positions de ses maîtres, Kummer et Dirichlet.

#### 2.1.1.1 Lagrange et Vandermonde

Dans une analyse du *traité de la résolution des équations numériques* de Lagrange (1736-1813), Louis Poinsot (1777-1859) affirme que « toute l'Algèbre se réduit, au fond, à l'analyse des équations » [6], et que ce traité forme donc un « Traité d'Algèbre » [7]. Il poursuit en donnant ce qu'il estime être une partition de l'Algèbre [8] :

1. « La théorie générale des équations, c'est-à-dire l'ensemble des propriétés qui leur sont communes à toutes ».
2. « Leur résolution générale, qui consiste à trouver une expression composée des coefficiens de la proposée, et qui, mise au lieu de l'inconnue, satisfasse identiquement à cette équation ».
3. « La résolution numérique ».

En formulant ces deux derniers domaines, il reprend bien les propos de Lagrange lorsque ce dernier affirmait qu'« il faut bien distinguer la résolution des équations numériques de ce que l'on appelle en Algèbre la résolution générale des équations ». En effet pour Lagrange, l'Algèbre ne cherche pas à donner « les valeurs individuelles des quantités qu'on cherche », mais elle porte sur les « opérations ». C'est à l'arithmétique que reviendra la résolution numérique des équations. L'Algèbre, *l'arithmétique universelle* selon Newton, est placée ainsi par Lagrange *au dessus* de l'arithmétique et de la géométrie. Il arrive alors à la définition suivante :

> L'Algèbre, prise dans le sens le plus étendu, est l'art de déterminer les inconnues par des fonctions des quantités connues, ou qu'on regarde comme connues ; et la résolution

1. Dans celles de [Kronecker, 1881a], [Kronecker, 1885b] et [Kronecker, 1887b].
2. Voir les deux citations de Kronecker p. 72 et p. 73.
3. [Kronecker, 1887b] p. 2.
4. [Lagrange, 1771] et [Vandermonde, 1771].
5. [Boniface et Schappacher, 2001] p. 219 : « Buch aller Bücher ».
6. [Lagrange, 1826] p. v.
7. *Ibid.*
8. *Ibid.* p. vi.

> générale des équations consiste à trouver pour toutes les équations d'un même degré, les fonctions des coefficiens de ces équations qui peuvent en représenter toutes les racines. [Lagrange, 1826] p. xxiii

Il n'est donc pas question, dans la science supérieure que représente l'Algèbre, de trouver *effectivement* les solutions d'une équation, mais bien d'indiquer les opérations à faire. C'est ainsi que lorsque Lagrange propose ses méthodes de résolution des équations à l'aide des *résolvantes* qui portent maintenant son nom ou lorsqu'il donne une méthode permettant de trouver le nombre de racines réelles d'une équation à l'aide de l'*équation aux carrés des différences des racines de la proposée*, l'aspect qu'il qualifie lui-même de « presque impraticable » n'est pas un problème majeur. L'Algèbre ne s'occupe pas de calculs effectifs : « c'est même pour prévenir toute imperfection (...) que, dans le calcul, au lieu de nombres, on emploie des lettres ; d'où il résulte qu'aucune opération ne peut se faire actuellement » [9].

Serret, en 1849, lorsqu'il publie son *cours d'Algèbre supérieur*, commence son ouvrage en donnant la définition suivante de l'algèbre :

> L'Algèbre est à proprement parler, l'*Analyse des équations* ; les diverses théories partielles qu'elle comprend se rattachent toutes, plus ou moins, à cet objet principal. A ce point de vue, l'Algèbre peut se diviser en trois Parties bien distinctes :
>
> 1°*La théorie générale des équations*, c'est-à-dire l'ensemble des propriétés qui sont communes à toutes les équations ;
>
> 2°*La résolution des équations numériques*, c'est-à-dire la détermination des valeurs exactes ou approchées des racines d'une équation dont les coefficients sont donnés en nombres ;
>
> 3°*La résolution algébrique des équations*, c'est-à-dire la détermination d'une expression composée avec les coefficients d'une équation donnée, et qui, substituée à l'inconnue, satisfasse identiquement à cette équation, soit que les coefficients de l'équation proposée soient *numériquement donnés*, soit que, étant simplement considérés comme connus, ils restent indéterminés et représentés par des lettres. [Serret, 1849] p. 1

Cette partition de l'Algèbre est identique à celle que Poinsot faisait ressortir du travail de Lagrange : pour Serret, Poinsot et Lagrange, l'Algèbre a pour objets les équations. Elle ne se propose pas seulement de résoudre ces dernières, mais elle cherche à établir des propriétés qui ont trait à elles. Le statut de la résolution numérique est en revanche discuté chez Lagrange : les méthodes appartiennent à l'algèbre mais l'exécution effective de ces méthodes à l'arithmétique.

Quant à Alexandre-Théophile Vandermonde (1735-1796), il ne s'est pas, dans ses écrits, exprimé sur ses conceptions de l'algèbre, et comme l'affirme Henri Lebesgue dans [Lebesgue, 1955], « nous ignorerions sans doute toute l'œuvre mathématique de Vandermonde sans les éloges enthousiastes que le grand algébriste Kronecker lui adressait dans ses Leçons. » [10]. Cependant, Kronecker affirme que [11] « l'essor moderne de l'algèbre com-

9. [Lagrange, 1826] p. 314.

10. [Lebesgue, 1955] p. 205

11. Ces propos son rapportés par Lebesgue dans [Lebesgue, 1955] p. 205. Il s'agirait, d'après lui, d'une partie de l'introduction de Kronecker à l'un de ses cours. Cependant, comme le montre Olaf Neumann dans [Goldstein *et al.*, 2007] p. 110, on trouve cette phrase dans l'introduction que Kronecker a rédigée pour la

mence avec le mémoire intitulé : *Sur la résolution numérique des équations*, la profondeur des conceptions, si clairement exprimées dans cet ouvrage, nous semble vraiment surprenante ». Qu'est-ce qui a, dans cet ouvrage, si particulièrement marqué Kronecker ? Il s'agit en fait dans un premier temps de la forme que Vandermonde utilise pour traiter les équations [12] :

> Vandermonde ne part pas de la forme habituelle des équations, qui apparaît dans son côté gauche comme un polynôme, mais sous la forme du produit $(x - x_1)(x - x_2)...$, ce qui peut sembler étrange à première vue, mais qui [fait] de nos jours le fondement de tout le traitement des équations. [Kronecker, 1887b] p. 8

Mais il s'agira aussi de la méthode, liée bien sûr à la forme précédente, que Vandermonde propose. Cette méthode est finalement plus un programme, une liste de problèmes à résoudre, que Vandermonde expose ainsi [13] :

> 1. Trouver une fonction des racines, de laquelle on puisse dire, dans un certain sens, qu'elle égale telle de ces racines que l'on voudra.
> 2. Mettre cette fonction sous une forme telle qu'il soit plus indifférent d'y changer les racines entr'elles.
> 3. Y substituer les valeurs en somme de ces racines, somme de leur produits deux à deux, etc.
>
> [Vandermonde, 1771] p. 370

Ne nous attardons pas sur le contenu de chacun de ces trois problèmes – nous reviendrons longuement dessus lorsque nous examinerons la partie du cours de Kronecker qui traite des équations quadratiques et cubiques [14] – mais retenons que Vandermonde propose comme objet de son traité des résultats sur les fonctions symétriques des racines.

Le but de l'algèbre reste donc pour tous ces auteurs la résolution des équations, et l'aspect numérique de la question est à la frontière de cette Algèbre, quelquefois dedans,

---

traduction allemande du mémoire de Vandermonde ([Vandermonde, 1771]). On trouve aussi dans l'introduction de son cours de 1886-87 : [Kronecker, 1887b] pp. 7-8 : « In ein neues Stadium tritt die Algebra mit Vandermonde (1735-1796) dem seine algebraischen Arbeiten 1771 die Mitgliedschaft der Pariser Akademie eintragen. »

12. Vandermonde geht nicht von der gewöhnlichen Gestalt der Gleichung aus, in der ihre linke Seite als Polynom erscheint, sondern von der Produktform $(x-x_1)(x-x_2)...$, was auf den ersten Anblick wunderlich erscheinen mag, aber in der That das Fundament der ganzen heutigen Behandlungsweise der Gleichungen [leistet].

13. On trouve dans ([Kronecker, 1887b] p. 8) une traduction presque mot à mot de ce passage de Vandermonde : « Die Aufgabe der algebraischen Auflösung der Gleichungen zerlegt Vandermonde in die folgenden drei getrennten Probleme :

> 1. Eine Function der Wurzeln zu finden, die in einem gewissen Sinne irgend eine der Wurzeln darstellen kann, d.h. die so vieldeutig ist, als Wurzeln vorhanden sind.
> 2. Diese Function unter eine solche Form zu setzen, daß eine Vertausch[ung] der Wurzeln sie umgeändert läßt, mit andern Worte, die in Rede stehende Function soll nur von den symmetrischen Functionen der Wurzeln abhängen.
> 3. Die symmetrischen Functionen durch diejenigen symmetrischen Functionen der Wurzeln auszudrücken, die unmittelbar durch die Coeff. der Gleichung gegeben sind, d.h. durch die sog. elementaren symmetrischen Functionen.

»

14. Voir p. 104.

quelquefois à l'extérieur. Cependant, et Kronecker le souligne [15], avec Vandermonde et Lagrange un travail plus systématique sur les méthodes commence, dont les résultats seront à la base des découvertes de Galois, et comme nous allons le voir, de certains travaux de Kronecker. Ce dernier note d'ailleurs à propos de ces deux mémoires de Vandermonde et Lagrange que « le sens de ces travaux n'émerge vraiment seulement qu'avec les recherches de *Gauss* » [16]. Et il est certain, comme le souligne A.N. Kolmogorov [17] à propos de la période qui succède à celle de Lagrange et Vandermonde, que « le premier événement majeur de cette période est la parution en 1801 par C.F. Gauss des *Disquisitiones Arithmeticae.* »

#### 2.1.1.2 Gauss

Dans sa thèse d'habilitation de 1799, Gauss propose un théorème qu'il nomme le « théorème fondamental de la théorie des équations. » [18]. Ce théorème correspond à ce que l'on appelle aujourd'hui le *théorème fondamental de l'algèbre*, mais Gauss n'utilise pas le terme d'Algèbre. On comprend d'ailleurs facilement, à la lecture par exemple de sa thèse, qu'il utilise, en tout cas autour des années 1800, indifféremment les termes d'« Algébriste » et d'« Analyste » [19], et par extension on voit qu'« analyse » et « algèbre » [20] sont pour lui synonymes.

L'un des principaux apports de Gauss dans la théorie des équations est son travail sur l'irréductibilité des polynômes (dans un certain corps), et en particulier la recherche d'un polynôme irréductible de plus bas degré ayant pour racine une racine $n^{\text{ième}}$ de l'unité. Dedekind, qui a contribué à l'élaboration des *œuvres complètes* de Gauss, affirme à propos des *Disquisitiones Arithmeticae* [21] :

> Ici se traite pour la première fois le concept d'irréductibilité (art. 341), qui est devenu crucial par la suite dans toute l'orientation de l'algèbre ; (...) En effet, à partir du germe laissé par Gauss, une science est née, que l'on peut (...) peut-être décrire comme la science des affinités algébriques entre les nombres, ou, si l'on souhaite utiliser une expression choisie par moi, comme la science des affinités entre les corps. [Dedekind, 1873]

La particularité de Gauss, quant à sa vision de l'algèbre, est bien à chercher dans le rapport entre nombre et algèbre, c'est-à-dire entre l'arithmétique et l'algèbre. En effet, si pour lui l'utilisation de la géométrie, sans être complètement interdite, semble pouvoir être

15. Une analyse tout à fait comparable est faite presque un siècle plus tard dans [Bourbaki, 2006a] p. 100.

16. [Kronecker, 1881a] p. 4 : « Die Bedeutung dieser Arbeiten tritt aber erst durch die Forschungen von *Gauss* (1777-1855) zu Tage. »

17. [Kolmogorov *et al.*, 2001] p. 36 : « The first event of this period was the appearance, in 1801, of C.F. Gauss *Disquisitiones Arithmeticae.* »

18. [Gauss, 1799] p.10 : « theorema hoc fundamentale doctrinae aequationum ».

19. *algebraistae* et *analystae.*

20. *analysis* et *algebraicis.*

21. Hierbei tritt zum ersten Male der Begriff der Irreduktibilität auf (art. 341), welcher entscheidend für die ganze Richtung der späteren Algebra geworden ist ; (...) In der Tat ist aus dem von Gauß gelegten Keime eine Wissenschaft entstanden, welche man (...) vielleicht als die Wissenschaft von der algebraischen Verwandtschaft der Zahlen oder, wenn man sich eines von mir gewählten Ausdruckes bedienen will, als die Wissenschaft von der Verwandtschaft der Körper bezeichnen könnte.

avantageusement remplacée dans tout ce qui a trait à l'algèbre [22], l'arithmétique aura, elle, un tout autre statut. Dans un des rares textes où il aborde la science qu'il pratique d'un point de vue philosophique, il donne une définition des mathématiques et précise les choses suivantes [23] :

> Une quantité en soi ne peut pas être l'objet d'une investigation scientifique : les mathématiques considèrent les grandeurs uniquement dans leur relation les unes avec les autres. [Gauss, 1800] p. 57

Ce à quoi il ajoute [24] :

> Maintenant les mathématiques apprennent réellement des vérités générales à propos des relations entre grandeurs (...). Ces différentes relations entre grandeurs sont les fondations des deux principales disciplines mathématiques. L'arithmétique considère les grandeurs en rapports arithmétiques, et les représente arithmétiquement ; la géométrie considère les grandeurs en rapports géométriques, et les représente géométriquement. Représenter géométriquement des grandeurs qui ont des rapports arithmétiques, comme cela était commun chez les anciens, n'est plus habituel de nos jours, car sinon nous regarderions cela comme une partie de la géométrie. Au contraire, on applique l'exposition arithmétique très souvent à des grandeurs en relation géométrique, comme par exemple dans la trigonométrie, et aussi dans la théorie des lignes courbes, que certains considèrent comme une discipline géométrique. Que les modernes préfèrent si fortement l'exposition arithmétique à la géométrique n'est pas sans raison, en particulier car notre méthode pour compter (par dizaines) est infiniment plus facile que celle des Anciens. [Gauss, 1800] p. 57-58

22. Gauss écrit ainsi, reprochant indirectement à d'Alembert d'avoir utilisé des conceptions géométriques là où des considérations analytiques conviendrait mieux : « Il serait bon d'observer que d'Alembert a appliqué des considérations géométriques dans l'exposé de sa preuve et a considéré $X$ comme l'abscisse et $x$ comme l'ordonnée d'une courbe (selon la coutume de tous les mathématiciens de la première partie de ce siècle à qui la notion de fonction était moins familière). Mais tout son raisonnement, si l'on en considère seulement l'essentiel, ne repose pas sur la géométrie mais sur des principes purement analytiques, et une courbe imaginaire ainsi que des coordonnées imaginaires sont des concepts assez difficiles et qui pourraient choquer un lecteur de notre époque. Par conséquent j'ai plutôt donné ici une représentation purement analytique. » ([Gauss, 1799] p. 8-9)
« Observare convenit, ill. d'Alembert in sua huius demonstrationis expositione considerationes geometricas adhibuisse, atque $X$ tamquam abscissam, $x$ tamquam ordinatam curvae spectavisse (secundum morem omnium geometrarum primae huius saeculipartis, apud quos notio functionum minus usitata erat). Quia vero omnia ipsius ratiocinia, si ad ipsornm essentiam solam respicis, nllis principiis geometricis, sed pure analyticis innituntur, et curva imaginaria, ordinataeque imaginariae expressiones duriores esse lectoremque hodiernum facilius offendere posse videntur, formam repraesentationis mere analyticam hic adhibere malui. »

23. Eine Grösse für sich kann noch kein Gegenstant einer wissenschaftlischen Untersuchung werden : die Mathématik betrachtet die Grössen nur in Beziehung auf einander.

24. Die Mathematik lehrt nun eingentlich allgemeine Wahrheiten, welche die Relationen der Grössen betreffen (...). Diese verschiedenen Beziehungen der Grössen sind die Grundlage der beiden Hauptdisciplinen der Mathematik. Die Arithmetik betrachtet Grössen in arithmetischen Beziehungen und stellt sie arithmetisch dar ; die Geometrie betrachtet Grössen in geometrieschen Beziehungen und stellt sie geometrisch dar. Grössen, die arithmetische Beziehungen haben, geometrisch darzustellen, was bei den Alten so gewöhnlich war, ist es gegenwärtig nicht mehr, sonst würde man dieses als einen theil der Geometrie anzuschen haben. Im Gegentheil wendet man die arithmetische Darstellungsart äusserst häufig auf Grössen in geometrischen Beziehungen an, zum Exempel in der Trigonometrie, auch in der Lehre von den krummen Linien, welche man als geometrische Disciplinen betrachtet. Dass die Neuern der arithmetischen Darstellungsart so sehr den Vorzug vor der geometrischen geben, geschicht nicht ohne Grund, besonders da unsere Methode zu zählen (nach der Dekadik) so unendlich leichter ist, als die der Alten.

Ainsi chez Gauss, les fondations des mathématiques sont la géométrie et l'arithmétique. Mais l'arithmétique, ou tout au moins l'« exposition arithmétique », peut se substituer à l'« exposition géométrique », et on voit poindre dans un texte écrit au tout début [25] du XIX^e siècle les prémisses de la période d'« arithmétisation » qui se développe tout au long de ce siècle. L'algèbre Gauss est donc une algèbre basée sur l'arithmétique, qui travaille sur les rapport entre les grandeurs : nous nous rappellerons cela lorsque nous parlerons de la place des congruences – rapport arithmétique entre grandeurs algébriques – dans les travaux de Kronecker.

#### 2.1.1.3 Dirichlet et Kummer

Dirichlet est l'un des premiers à comprendre les *Disquisitiones Arithmeticae* de Gauss [26], et surtout à les professer à l'université de Berlin, où il enseignera de 1829 à 1855. Il obtiendra ensuite la chaire de Gauss à l'université de Göttingen, qu'il occupera jusqu'à sa mort en 1859. Par ses enseignements, il aura, comme le montre D. Rowe, une profonde influence sur Kronecker et ses contemporains :

> Although he published relatively little in his lifetime, Dirichlet, unlike Gauss, placed a high value not only on the discovery of new ideas but on their dissemination as well. In the course of his career at Berlin and Göttingen his students included some of the outstanding figures of the next generation – Eisenstein, Kronecker, Dedekind, and Riemann – and his *Vorlesungen über Zahlentheorie*, which only became accessible to the mathematical public in 1863 through the efforts of Richard Dedekind, also exerted a profound influence on number theory and algebra in the late nineteenth century. [Rowe, 1988] p. 14

Si l'on examine la liste des cours qu'il a donnés à l'université de Berlin [27], on voit la multitude des sujets qu'il y a abordés : algèbre, analyse théorie des nombres, théorie des équations, arithmétique, théorie des dérivées partielles, intégration, suites, probabilité... Il est finalement impossible de trouver un domaine des mathématiques qu'il n'a pas enseigné. Mais plus encore, dès 1830, nombre de ses leçons s'intituleront « Application de ... à ... » [28]. Il donne ainsi des cours sur l'application des mathématiques à la physique, la mécanique, l'électricité ou la théorie de la chaleur, mais aussi des applications du calcul intégral à d'autre domaines des mathématiques. On peut par exemple citer « Chapitres choisis du calcul intégral avec application aux probabilités », « Application du calcul intégral à l'étude des suites » ou encore « Quelques applications du calcul intégral à la théorie des nombres » [29]. Car en effet, Dirichlet aime à établir des ponts entre les disciplines, et à propos de l'usage qu'il fait des séries infinies en théorie des nombres il affirme :

25. Le texte qui n'est pas daté, semble avoir été écrit à cette période : voir [Ewald, 2005] p. 293.
26. On pourra lire par exemple [Rowe, 1988] ou [Begehr, 1998] p. 33 et suivantes.
27. Voir [Biermann, 1959] pp. 34-39.
28. par exemple [Biermann, 1959] p. 37 : « Die Anwendung der Integralrechnung auf die Zahlenlehre ».
29. « Ausgewählte Kapitel der Integralrechnung mit anwendung auf die Probabilitätslehre », « Die Anwendung der Integralsrechnung auf die Reihenlehre » et « Einige Anwendungen der Integralsrechnung auf die Theorie der Zahlen »

> La méthode que j'emploie me paraît surtout mériter quelque attention par la liaison qu'elle établit entre l'Analyse infinitésimale et l'Arithmétique transcendante. [Dirichlet, 1889] p. 360

On sait que le travail de Dirichlet est à la base de la théorie analytique des nombres [30]. Si cette approche, en ce qui concerne les questions ayant trait de près ou de loin à l'algèbre, n'est pas celle que Kronecker suivra, il reprendra à son compte l'idée d'un décloisonnement des mathématiques, comme la prône Dirichlet.

Par ailleurs, Dirichlet, et Kummer après lui, s'intéressent au grand théorème de Fermat [31] : Dirichlet démontre le théorème pour les cas $n = 5$ et $n = 14$, et Kummer pour tous les nombres premiers inférieurs à 100. Plus que le théorème en soi, ce sont les méthodes qui seront utilisées à cette occasion qui auront une influence certaine sur le devenir de l'algèbre, et en particulier bien sûr la théorie des nombres idéaux de Kummer [32]. En effet, le point problématique sur lequel ce dernier se penche touchera à la notion de divisibilité dans certains ensembles d'entiers complexes, et en particulier aux problèmes d'unicité de la décomposition en facteurs irréductibles dans certains anneaux cyclotomiques. L'algèbre de Kummer, que Kronecker connaissait et appréciait [33], est une algèbre teintée d'arithmétique qui s'articule autour d'une théorie de la division. Kronecker lui aussi produira une *théorie de la division*, à laquelle Hermann Weyl et plus tard Harold Edwards consacreront de nombreuses pages [34]. Kronecker appliquera les outils issus de cette théorie à partir de 1882, et de la publication de ses *Grundzüge einer arithmetischen Theorie der algebraischen Grössen*, à la théorie des *Gattungen*, théorie qui constitue à partir de cette même année la seconde partie des *Vorlesungen über die Theorie der algebraischen Gleichungen*, et qui s'apparente, dans un langage plus moderne, à une théorie des extensions de corps [35].

### 2.1.2 L'algèbre dans les *Vorlesungen über die Theorie der algebraischen Gleichungen*

L'algèbre a pour Kronecker une place privilégiée dans son enseignement, car elle permet d'amener l'étudiant à un plus grand degré d'abstraction. En effet, d'après Kronecker, « nous avons affaire à l'abstraction non seulement dans les mathématiques, mais aussi dans toutes les sciences », et après le passage « des nombres aux calculs », un degré d'abstraction supplémentaire est obtenu par le passage « des calculs à l'algèbre » [36].

30. Voir [Shields, 1989].
31. Théorème qui affirme que l'équation $x^n + y^n = z^n$ n'a pas de solution entière pour $n > 2$.
32. Voir par exemple [Neumann, 2007] p.85 ou [Boniface, 2004] pp. 77-85.
33. Rappelons que Kummer sera à la fois l'enseignant de mathématiques au *Gymnasium* de Kronecker à Liegnitz, mais aussi plus tard son ami et collègue à Berlin.
34. Voir [Edwards, 1990] et [Weyl, 1940].
35. Un « domaine de genre $\mathfrak{G}'$ » (*Bereich der Gattung* $\mathfrak{G}'$ [Kronecker, 1882b] p. 252) est en fait une extension d'un corps de base $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', ...)$, que Kronecker nomme « un domaine de base » (*Stammbereich* [Kronecker, 1882b] p. 252). Pour plus de détails, on pourra lire [**?**].
36. [Kronecker, 1885b] pp. 1-2 : « dann kommt die weitere Abstraktion von der Bestimmtheit der Zahlen zu dem Unbestimmten. Diese beiden Schritte sind die zum Rechnen und vom Rechnen zur Algebra. »

#### 2.1.2.1 L'algèbre comme formation à l'abstraction

À propos de la grande difficulté de ses cours, Kronecker affirme [37] :

> Je devrai pour cela exposer votre endurance à un matériel qui semble très abstrait, voire sec, et peu habituel ; mais je sais, et pas seulement par ma propre expérience, que, une fois que l'on s'est habitué, on apprend à reconnaître les beautés partiellement cachées des équations algébriques, telles que celles des nouvelles théories mathématiques, et en même temps on acquiert le grand avantage d'un excellent apprentissage en abstraction mathématique. [Kronecker, 1885b] p.1

Pour Kronecker, l'algèbre fournit donc « un excellent exercice et une excellente formation en abstraction mathématique » [38], abstraction qui n'est pas le propre des mathématiques, mais qui au contraire est commune à toutes les sciences [39]. Si l'algèbre est particulièrement adaptée à cette tâche, c'est qu'il est a priori plus facile d'y être « pointu et complet » [40]. Mais si l'abstraction présente dans l'algèbre est commune à toutes les sciences, il ne faudrait pas comprendre que Kronecker estime que l'algèbre est une sorte de « logique mathématique » [41] :

> L'algèbre a été nommée la logique des mathématiques mais je ne souhaite pas y souscrire car toute la mathématique est logique et le comprendre et le faire comprendre est somme toute la tâche principale de la science. [Kronecker, 1885b] p. 2

Il est d'ailleurs possible que Kronecker fasse ici allusion à l'ouvrage que Frege vient juste de publier : *Die Grundlagen der Arithmetik* [42] paru en 1884 et dont le sous-titre est *Eine logisch mathematische Untersuchung über den Begriff der Zahl.*

#### 2.1.2.2 De la *résolution* à la *théorie* des équations

Pour Kronecker [43],

> Dans le calcul on a les règles, dans l'algèbre les lois ; c'est presque comme un pas de l'art à la science. [Kronecker, 1885b] p. 2

C'est le passage de la mise en place de techniques de résolution à une vraie théorie des équations que semble vouloir caractériser Kronecker par ces quelques mots. De

37. Ich werde dabei an Ihre Ausdauer bei ganz abstracten, fast trocken erscheinenden Materien Ansprüche, und zwar nicht gewöhnliche stellen müssen ; aber ich weiß und nicht bloß aus eigener Erfahrung, daß, wenn man sich hier hinein gewöhnt hat, man die etwas verborgenen Schönheiten der algebraischen Gleichungen erkennen lernt, ebenso wie die der jüngeren mathematischen Theorien, und dabei erwirbt man sich zugleich den großen Vorteil einer vortrefflichen Schulung in mathematischer Abstraction.

38. [Kronecker, 1885b] pp. 2-3 : « eine vortreffliche Uebung und Schulung in mathematischer Abstraction »

39. [Kronecker, 1885b] p. 1 : « Nous avons affaire à l'abstraction non seulement dans les mathématiques mais aussi dans toutes les sciences » « Wir haben es in der Mathematik nicht allein, sondern in aller Wissenschaften überhaupt mit der Abstraction zu thun. »

40. [Kronecker, 1885b] p. 3 : « scharf und vollständig ».

41. Die Algebra ist einmal die Logik der Mathematik genannt worden, doch will ich das nicht unterschreiben, denn die ganze Mathematik ist Logik und das Einsicht geben und finden ist wie die Hauptaufgabe der Wissenschaft überhaupt

42. Voir [Frege, 1884]

43. Beim Rechnen hat man die Regeln, in der Algebra die Gesetze ; es ist das beinahe wie der Schritt von der Kunst zur Wissenschaft.

plus, en examinant les spécificités du rapport à l'algèbre des quelques mathématiciens précédents dont le travail a particulièrement influencé Kronecker, on voit se dessiner une algèbre qui, par son rapport aux autres domaines des mathématiques, commence à s'échapper de la classique *résolution algébrique des équations*. Pourtant, les leçons de Kronecker ne s'intitulent pas *Vorlesungen über Algebra*, comme d'autres traités du XIX^e siècle [44], mais bien *Vorlesungen über die Theorie der algebraischen Gleichungen*. Si, comme nous l'avons déjà remarqué, les équations algébriques sont bien les objets centraux de ces cours, est-ce pour autant que Kronecker voit dans leur étude la totalité de l'algèbre ?

Il répond en partie à cette question dans l'introduction de son cours du semestre d'hiver 1884-85 [45] :

> Mes leçons doivent, en partant du plus simple, s'étendre aux parties de l'algèbre qui concernent les équations algébriques, soit qu'elles y aient leur fondement, soit qu'elles y trouvent leur but ; pourtant une grande partie de l'algèbre en reste exclue, même si de temps en temps une courte excursion dans un tel domaine plus lointain est entreprise. La restriction se trouvant dans le titre est imposée par la grande extension de l'algèbre dont les résultats et les méthodes si divers ne trouvent plus maintenant de place dans une leçon. Une telle leçon pourrait plutôt, même si elle ne contenait que l'essentiel, tant qu'elle n'est pas complétée par des principes entièrement généraux, manquer de cohérence. Déjà dans la partie que je veux aborder le systématisme n'est pas sans difficulté ; c'est pourquoi je n'annonce qu'un développement aussi systématique que possible. Ce n'est pas une construction comme dans la géométrie d'Euclide ou la théorie des nombres de Gauss, ce n'est qu'une certaine suite de choses ordonnées selon un certain point de vue réfléchi, méthodique et pédagogique, dans laquelle je vous donnerai des développements. [Kronecker, 1885b] p. 1

Pour Kronecker, l'algèbre ne se limite pas à l'étude des équations algébriques. Cependant, dans le cadre de ses leçons, il décide de se restreindre à ces dernières, que ce soit pour chercher à les résoudre ou pour les prendre comme objet d'étude. Ces restrictions ont deux raisons. Premièrement, l'immensité des domaines qui y trouveraient leur place est telle que le cadre d'un cours semestriel est bien trop étroit pour tous les contenir. Et ensuite, pour

44. Par exemple l'*Algèbre supérieure* de Francoeur ([Francoeur, 1838]), les *leçons d'algèbre* de Louis Etienne Lefébure de Fourcy ([de Fourcy, 1850]), le *Traité d'algèbre supérieure* de Serret ([Serret, 1849]) pour la première moitié du XIX^e siècle, et les *Lessons introductory to the modern higher algebra* de Salmon ([Salmon, 1885]), le *Lehrbuch der Algebra* de Weber ([Weber, 1895]) ou les *Vorlesungen über algebra* de Netto ([Netto, 1900]) pour la seconde moitié.

45. Meine Vorlesungen soll vom Einfachsten ausgehend sich über die Teile der Algebra erstrecken, welche sich auf die algebraischen Gleichungen beziehen, sei es, dass sie hier ihren Entstehungsgrund haben, sei es, dass sie hier ihren Zielpunkt finden ; dabei werden immer große Teile der Algebra ausgeschlossen bleiben, wenn auch hin und wieder ein kurzer Streifzug in ein solches, ferner liegendes Gebiet unternommen werden wird. Die im dem Titel liegende Beschränkung ist durch die große Ausdehnung der Algebra geboten, deren so verschiedenartige Resultate und Methoden jetzt nicht mehr in einer Vorlesung Platz finden. Vielmehr möchte eine solche Vorlesung, wenn sie auch nur das Wichtigste nimmt, so lange sie nicht durch ganz allgemeine Prinzipien ergänzt wird, des Zusammenhanges ermangeln. Schon in den Teilen, welche ich auseinandersetzen will, hat das Systematische nicht geringe Schwierigkeiten ; ich habe deshalb nur eine möglichst systematische Entwicklung angekündigt. Es ist nicht ein Aufbau, wie der der Geometrie bei Euclid oder der Zahlentheorie bei Gauß, es ist nur eine überlegte durch sachliche, methodische und pädagogische Gesichtspunkte bedingte Reihenfolge, in welcher ich Ihnen die Entwicklungen geben werde.

réussir à en donner une exposition systématique, il manque une théorie générale qui en serait le liant. Nous verrons que pour Kronecker, une solution pour dépasser cette dernière difficulté sera donnée par le développement d'une « arithmétique générale » (*allgemeine Arithmetik*). Kronecker, décrivant ce qu'il cherche à faire dans ses cours, ajoute [46] :

> Jusqu'à il y a environ cent ans, la résolution des équations algébriques, qui fut aussi bien le motif et le but de l'invention en algèbre, constituait le centre de l'intérêt algébrique. Toutes les recherches algébriques se rattachaient à la résolution des équations algébriques, et toutes étaient élaborées en vue de leur résolution. Avec les larges développements de l'algèbre de ces derniers temps, les points de vue de la recherche sont devenus essentiellement autres. Le titre : « Théorie des équations algébriques » n'est habituel que depuis environ une génération, et de même l'est son contenu. [Kronecker, 1891a] p. 1

Kronecker distingue dans cet extrait du cours du semestre d'hiver 1890-91 trois périodes dans l'algèbre : 1) il y a plus de cent ans, 2) depuis une génération et 3) entre les deux, c'est-à-dire de la fin du XVIII$^{e}$ siècle jusqu'au milieu du XIX$^{e}$ siècle. En effet, le début de ces grands changements au sein de l'algèbre est daté par Kronecker à la publication des deux ouvrages de Vandermonde et Lagrange que nous avons cités plus haut, c'est-à-dire vers 1770. La dernière génération, celle qui a su travailler à partir du « livre de tous les livres », a multiplié les « points de vue » sur l'algèbre. Ce changement est caractérisé, selon Kronecker, par le passage du terme de *résolution* à celui de *théorie*. Cette distinction ne permet pas d'englober tout l'algèbre [47] :

> Dans l'ensemble, ce contenu est si immense que ces leçons ne peuvent seulement que promettre d'en donner un extrait. Convient-il d'en distinguer dès le départ par une définition ce qui lui appartient ? C'est tout à fait impossible, de même que la parcourir de façon systématique. Une grande série d'exemples confirme les mots de *Dirichlet* disant qu'il n'y a pas une seule véritable méthode en analyse. [Kronecker, 1891a] pp. 1-2

Pour Kronecker, tenter d'en donner une définition est vain. Par contre, en rappelant comment Dirichlet a multiplié les méthodes dans son travail, sans poser de cloisons fictives entre les divers domaines des mathématiques, il donne l'une des clefs pour comprendre la façon dont il aborde lui-même l'« algèbre ».

46. Nach vor hundert Jahren bildete die Auflösung der algebraischen Gleichungen, die ja auch Anlass und Ziel der Erfindung der Algebra war, den Mittelpunkt alles algebraischen Interesses. An die Auflösung der Gleichungen knüpfen alle algebraischen Untersuchungen an, und alle arbeiteten auf sie hin. In der breiten Entwicklung der Algebra während der neusten Zeit sind die Gesichtspunkte der Forschung wesentlich andere geworden. Wie der Titel : « Theorie der algebraischen Gleichungen » seit einem Menschenalter etwa erst sich eingebürgert hat, so auch der Inhalt.

47. Im ganzen ist dieser Inhalt ein so gewaltiger, daß diese Vorlesungen daraus nur einen Ausschnitt zu geben versprechen können. Soll von vornherein durch eine Definition abgegrenzt werden, was in ihn hineingehört ? Das ist ebenso unmöglich, wie die Durchführung einer strengen Systematik in seinem Verlaufe. Eine Große Reihe von Beispielen bewahrheitet *Dirichlets* Ausspruch, daß es eine einzig wahre Methode in der Analysis nicht giebt.

#### 2.1.2.3 Une algèbre poreuse

S'interdire des interactions entre les domaines des mathématiques ne fait pas que freiner le travail du chercheur, mais les domaines eux-mêmes se retrouvent en « danger » [48] :

> La partie de l'algèbre à laquelle nous avons affaire ici s'est éloignée du danger d'isolement depuis 1770 ; depuis ce temps elle est en lien étroit avec les autres disciplines mathématiques ; à partir de ce moment-là il s'exerce une interaction plus active entre analyse, géométrie et théorie des nombres d'un côté, et théorie des équations algébriques de l'autre. [Kronecker, 1887b] p. 2

On retrouve la date de 1770, date approximative de la publication de [Lagrange, 1771] et [Vandermonde, 1771], comme début de ces interactions. Au début du XIX[e] siècle, l'analyse n'est pas seulement une branche des mathématiques particulièrement féconde, c'est aussi celle qui est en train de remplacer la géométrie comme paradigme. Et c'est son degré de généralité et d'abstraction qui lui vaut cette place. C'est ainsi qu'en 1796 dans son *Exposition du système du monde*, Laplace écrit : « L'analyse algébrique nous fait bientôt oublier l'objet principal, pour nous occuper de combinaisons abstraites ; et ce n'est qu'à la fin, qu'elle nous y ramène. Mais en s'isolant ainsi des objets, après en avoir pris ce qui est indispensable pour arriver au résultat que l'on cherche ; en s'abandonnant ensuite aux opérations de l'analyse, et réservant toutes ses forces pour vaincre les difficultés qui se présentent ; on est conduit par la généralité de cette méthode, et par l'inestimable avantage de transformer le raisonnement en procédés mécaniques, à des résultats souvent inaccessibles à la synthèse [49] » [50]. L'analyse et l'algèbre, selon les circonstances et les auteurs, peuvent se confondre, et on voit là que Laplace encourage, pour progresser en géométrie, l'utilisation des méthodes de l'analyse.

Mais si ces interactions sont vieilles d'un siècle, Kronecker bouscule encore un peu plus les frontières, dont le dépassement n'est pas seulement une source de progrès, mais un « bouleversement » dans la façon de faire des mathématiques, qui « a son fondement dans le rapport de ces leçons avec ceux des leçons précédentes sur la "théorie des nombres" et "la théorie des déterminants" » [51]

On a souvent considéré Kronecker comme un *algébriste* ou un *arithméticien* [52], sans savoir d'ailleurs exactement à quel contenu on lie ces termes. Ce serait une erreur de limiter son travail à un domaine des mathématiques, et d'ailleurs Frobenius ne s'y trompe

48. Der Teil der Algebra, mit dem wir hier zu thun haben, ist seit 1770 der Gefahr der Isolierung völlig entrückt ; seit jener Zeit steht sie in innigen Konnex mit den übrigen mathematischen Wissenschaften ; von jener Zeit macht sich die lebendigste Wechselwirkung zwischen Analysis, Geometrie, Zahlentheorie einerseits und der Theorie der algebraischen Gleichungen anderseits geltend.

49. Le terme « synthèse » signifie ici géométrie, et au début du XIX[e], il s'agit même de géométrie avec coordonnées.

50. Voir [Laplace, 1835] p. 340.

51. [Kronecker, 1891a] p. 2 : « Ihren Grund hat diese Umwälzung in dem Zusammenhange dieser Vorlesungen mit den ihnen vorangehenden über die "Zahlentheorie" und "die Theorie der Determinanten". »

52. Eric Temple Bell le décrit comme « l'un des plus importants algébristes et arithméticiens du XIX[e] siècle » (*one of the leading algebraists and arithmeticians of the nineteenth century* [Bell, 1945] p. 170) et Felix Klein affirmait qu'il « s'occupait surtout d'arithmétique et d'algèbre » (*er sich vorwiegend mit Arithmetik und Algebra* [Klein, 1926] p. 281).

pas [53] :

> Tous les mathématiciens louent les créations de Kummer en théorie des nombres [...]. Tous connaissent le travail fondationnel de Weierstrass dans la théorie des fonctions [...]. Mais il est bien plus difficile de caractériser rapidement et de façon pertinente la place de Kronecker dans la science, à cause des découvertes d'une grande portée qui lui assureront une gloire durable et qui ne restent pas dans le cadre d'une seule discipline mathématique. [Frobenius, 1893] p. 8

L'algèbre de Kronecker est tellement poreuse, qu'en dresser les limites est voué à l'échec. Finalement, le terme même d'algèbre perd son sens, et c'est pourquoi Kronecker utilise essentiellement la notion de « théorie des équations algébriques », partie des mathématiques plus facilement identifiable, et qu'il cherche à fondre dans une « arithmétique générale ». Car, comme l'affirme déjà Kronecker en 1861 dans son discours d'entrée à l'Académie des Sciences de Berlin [54],

> L'algèbre n'est pas vraiment une discipline en soi, mais un fondement et un outil pour l'ensemble des mathématiques, et son développement rapide dans les dernières années a été en fait suscité et dirigé par les besoins d'autres disciplines mathématiques. [Kronecker, 1861] p. 387

Dans l'introduction de son cours de 1890-91, il affirme aussi que l'ensemble de ses leçons « devraient être mises ensemble et elles doivent construire une "arithmétique générale" dans laquelle l'arithmétique des nombres entiers, comme elle est systématisée par *Euclide*, sera étendue aux fonctions de variables indéterminées » [55]. Essayons de donner maintenant un aperçu de ce que signifie pour Kronecker cette arithmétique générale.

#### 2.1.2.4 L'arithmétisation chez Kronecker

À propos de Lagrange, on trouve dans les *Vorlesungen über die Theorie der algebraischen Gleichungen* de Kronecker l'affirmation suivante [56] :

> Il a partout reconnu et mis en avant le côté algébrique ; si bien qu'il a voulu algébriser l'analyse. [Kronecker, 1885b] p. 26

Kronecker n'a pas cet « esprit algébrique » [57], et la direction qu'il suit, persuadé des bienfaits d'une algèbre poreuse, sera plutôt celle de l'*arithmétisation* de l'analyse.

53. Jeder Mathematiker rühmt Kummer's bahnbrechende Schöpfungen in der Zahlentheorie [...]. Jeder kennt die grundlegenden Arbeiten von Weierstrass in der Theorie der Functionen [...]. Ungleich schwerer ist es, Kronecker's Stellung in der Wissenschaft kurz und zutreffend zu kennzeichnen, weil die weittragenden Entdeckungen, die ihm dauernden Ruhm sichern, nicht in dem Rahmen einer einzelnen mathematischen Disciplin Platz finden.

54. La traduction, tirée de [Boniface, 2006] p. 141 est de J.Dieudonné : *Die Algebra ist insofern nicht eigentlich eine Disciplin für sich, sondern Grundlage und Werkzeug der gesammten Mathematik, und ihre neuere grossartige Entwickelung ist in der That durch das Bedürfnis anderer mathematischer Disciplinen erweckt und gefördert worden.*

55. [Kronecker, 1891a] p. 2 : « Mit diesen zusammen sollten und sollen sie bilden eine "allgemeine Arithmetik", in welcher die Arithmetik der ganzen Zahlen, wie sie systematisiert ist von *Euklid*, ausgedehnt wird auf Funktionen unbestimmter Variabeler. »

56. Er hat überall die algebraische Seite herauserkannt und hervorgehoben ; das geht so weit, dass er die Analysis algebraisieren wollte.

57. Voir [Sinaceur, 1991] p. 71.

**Arithmétisation de l'analyse** Le terme « arithmétisation » a été employé à la fin du XIX^e siècle pour décrire diverses tentatives de fondation des mathématiques, et en particulier de l'analyse, sur des bases qui n'étaient plus géométriques. N. Schappacher et J. Boniface rappellent d'ailleurs dans [Boniface et Schappacher, 2001] que selon Jules Molk [58] le verbe « arithmétiser » (*arithmetisieren*) aurait été employé pour la première fois par Kronecker en 1887 dans son article *Sur le concept de nombre* [59]. Cette définition négative a permis à des programmes très divers de coexister au sein de ce *mouvement*, auquel Kronecker, à côté de Weierstrass et Dedekind, a participé [60].

Cette arithmétisation est dans un premier temps, au cours du XIX^e siècle, celle de l'analyse. On en date en général les prémisses dans la *Rein analytischer Beweis* [61] de Bolzano et le cours de Cauchy de la même époque [62], mais son apogée a lieu dans la seconde moitié de ce siècle, avec les cours de Weierstrass et de Kronecker, ainsi que les constructions des réels par Dedekind et Cantor. Nous n'allons pas détailler ici les différents sens que peut revêtir ce terme, mais examiner ce que Kronecker entend par *arithmétique générale*, c'est-à-dire les mathématiques une fois qu'elles ont été « arithmétisées » [63].

**L'*arithmétique générale* selon Kronecker** L'algèbre, ou plutôt la « théorie des équations algébriques », est pour Kronecker, au même titre que l'analyse, une partie d'une « arithmétique générale » [64]. Ainsi celui-ci ne cherche pas, comme le fait par exemple Weierstrass, à fonder les mathématiques sur l'arithmétique, mais il veut plutôt construire

58. Jules Molk (1857-1914) a suivi les cours de Kronecker à Berlin et a passé sa thèse en 1884 sous sa direction. Cette thèse intitulée *Sur une notion qui comprend celle de la divisibilité et sur la théorie générale de l'élimination* et qui a été publiée dans les *Acta mathematica* se veut être une « introduction » au travail de Kronecker concernant les *Modulsystem* :

> Les voies nouvelles ouvertes à l'Algèbre par les travaux de Gauss, d'Abel et de Galois ont été le point de départ des recherches de M. Kronecker sur la théorie générale de l'élimination. Ces recherches sont intimement liées à celles qui ont pour objet l'étude des systèmes de diviseurs d'un système de fonctions entières. Je me suis proposé, dans ce mémoire, d'exposer les unes et les autres, en me plaçant du point de vue arithmétique de M. Kronecker. [Molk, 1884] p. 1

59. Voir [Boniface et Schappacher, 2001].

60. B. Petri et N. Schappacher notent que « the term "arithmetization" was in used around 1900 as a generic description of various programmes wich provided non-geometrical foundations of analysis, or other mathematical disciplines ».

61. Voir [Bolzano *et al.*, 1905].

62. [Cauchy, 1821]

63. On peut lire à ce sujet le chapitre *On Arithmetization* de Birgit Petri et Norbert Schappacher dans [Goldstein *et al.*, 2007] p. 343.

64. Kant utilise lui aussi le terme d'*arithmétique générale* pour désigner l'algèbre : « L'arithmétique générale (*algebra*) est une science si étendue que l'on ne peut nommer aucune science de la raison qui lui soit en cela semblable [...] de plus, les autres parties de la pure *mathesis* attendent leur croissance surtout de l'extension de cette doctrine générale des grandeurs. » ([Kant, 1991] p. 327 : lettre de Kant à J. Schultz du 25 novembre 1788). Cependant, si le terme est identique, il ne faudrait pas y voir une filiation : l'algèbre que connaît Kant est celui du milieu du XVIII^e siècle, et comme le montre Frank Pierobon dans [Pierobon, 2003], « l'algèbre représente, en quelque sorte, la limite de la conception kantienne des mathématiques ; à partir du moment où réellement et historiquement, l'écriture (algébrique) s'émancipe et devient autonome par rapport à l'image (géométrique) qu'elle est censée symboliser, de nouvelles mathématiques apparaissent que Kant n'a pas eu le temps et-ou l'occasion d'évaluer à leur juste mesure. »

ce qu'il appellera une « arithmétique générale », qui contiendra l'analyse et l'algèbre [65] :

> Et je crois aussi que l'on parviendra un jour à « arithmétiser » le contenu entier de ces disciplines mathématiques, c'est-à-dire à le fonder purement et simplement sur le concept de nombre pris dans son sens le plus étroit et donc à dépouiller à nouveau ce concept des modifications et élargissements le plus souvent provoqués par les applications à la géométrie et à la mécanique. [Boniface, 1999] p. 54

En 1883, il affirme d'ailleurs, dans une lettre datée du 7 août 1883 adressée à Rudolf Lipschitz (1832-1903) que ce travail est celui auquel il aura consacré sa vie [66] :

> À cette occasion [67] j'ai trouvé les fondations longtemps cherchées de toute ma théorie des formes entières qui achève en quelque sorte « l'arithmétisation de l'algèbre », à laquelle j'ai consacré ma vie de mathématicien. [Lipschitz et Scharlau, 1986] pp. 181-182

Kronecker remarque qu'« en considérant que l'objet de la théorie des nombres est essentiellement constitué par les seuls nombres entiers et en excluant en particulier les nombres irrationnels (numeri surdi), Gauss sépare peut-être un peu trop strictement l'objet de ses considérations des autres parties des mathématiques (quantitates) qui se rapportent aux propriétés des autres grandeurs mathématiques » [68]. Et c'est en cela qu'il s'oppose à Gauss, sur cette démarcation entre les différents domaines des mathématiques [69] :

> Il y a partout dans le cours de la science des moments où elle réfléchit sur elle-même, et revient à un point de vue abandonné depuis longtemps, mais plus riche en contenu et en expérience. Chaque science produit dans son développement un grand ensemble de nouvelles ressources qu'elle systématise, schématise et érige inconsciemment à des fins de recherche. L'algèbre doit aussi rester consciente qu'elle a produit par la mise en œuvre d'un schématisme uniforme des fictions conceptuelles et duquel elle renvoie à l'exploration des choses données par la nature elle-même. Même si elle est enrichie par

65. Traduction de J. Bonniface et N. Schappacher de : « Und ich glaube auch, dass es dereinst gelingen wird, den gesammten Inhalt aller dieser mathematischen Disciplinen zu „arithmetisiren" d. h. einzig und allein auf den im engsten Sinne genommenen Zahlbegriff zu gründen, also die Modificationen und Erweiterungen dieses Begriffs wieder abzustreifen, welche zumeist durch die Anwendungen auf die Geometrie und Mechanik veranlasst worden sind. » [Kronecker, 1887d] p.253.

66. Bei dieser Gelegenheit habe ich das lange gesuchte Fundament meiner ganzen Formentheorie gefunden, welches gewissermassen „die Arithmetisierung der Algebra" – nach der ich ja das Streben meines mathematischen Lebens gerichtet habe – vollendet

67. Il s'agit de la publication des *Grundzüge einer arithmetischen Theorie der algebraischen Grössen.*

68. Traduction d'Ivahn Smadja de [Kronecker, 1901] p. 2 : « Indem er so als das Objekt der Zahlentheorie wesentlich nur die ganzen Zahlen ansieht und insbesondere die irrationalen (numeri surdi) streng ausgeschlossen wissen will, trennt er den Gegenstand seiner Betrachtungen vielleicht etwas zu scharf von den Teilen der Mathematik, die sich auf die Eigenschaften anderer mathematischer Größen (quantitates) beziehen. »

69. Es ist überall der Gang der Wissenschaft, daß sie von Zeit zu Zeit, sich besinnend auf sichselbst, zurückkehrt zu einem lange zuvor verlassenen Standpunkt, aber reicher an Inhalt und an Erfahrung. In ihrer Entwicklung erzeugt jede Wissenschaft eine große Menge neuer Hilfsmittel, die sie systematisiert, schematisiert und unbewußt vom Mittel zum Zwecke der Untersuchung erhebt. Auch die Algebra muß sich dessen bewußt bleiben, daß sie zur Durchführung eines uniformierenden Schematismus begriffliche Fiktionen geschaffen hat und von denen zurückkehrt zur Ergründung der von der Natur selbst gegebenen Gebilde. Wenn sie auch durch die Vermischung mit der Analysis und Geometrie auf das Reichste befruchtet ist, so darf sie sich doch darum nicht den Begriffen und Vorstellungen dieser an und für sich ihr fremden Disciplinen unterwerfen. Löst sie aber sich wieder los von Analysis und Geometrie, so wird sie von selbst « arithmetisch ».

> le mélange de l'analyse et de la géométrie, elle ne doit pas se soumettre aux concepts et aux idées de ces disciplines étrangères. Mais si elle se détache de nouveau de l'analyse et de la géométrie, elle devient en soi « arithmétique ». [Kronecker, 1891a] p. 2

Dans ses cours sur la théorie des équations, c'est donc d'une algèbre centrée autour des notions d'équations et de racines que Kronecker construit. Mais il n'en dessine pas les contours – contours qui sont « si immenses que ces leçons ne peuvent seulement que promettre d'en donner un extrait » – et il fait appel à tous les outils des « mathématiques pures », c'est-à-dire de l'*arithmétique générale*. Ce cours s'inscrit bien dans un mouvement d'arithmétisation des mathématiques [70].

## 2.2 Les racines d'une équation

On sait que le concept de nombre (*Zahlbegriff*) tient une place importante dans les mathématiques de Kronecker – en particulier dans ce qu'il appellera l'« arithmétique générale ». Il lui consacre d'ailleurs un article : *Über den Zahlbegriff* ([Boniface, 1999]). Ce qu'il cherche principalement à éviter, c'est l'introduction de nouveaux concepts, soit qu'ils lui semblent superflus, soit qu'il estime qu'ils sont étrangers à l'arithmétique. Car si son algèbre est « poreuse », comme nous l'avons vu au chapitre précédent, c'est qu'elle est en fait dissoute – ou à dissoudre – dans l'arithmétique. Cette dernière forme un tout dans lequel certains concepts restent à proscrire. Il trouve une solution à ce problème avec l'introduction de l'indéterminée $X$, et en raisonnant modulo des polynômes à coefficients entiers :

> Mais avec l'introduction principielle de l' « indéterminée » (indeterminatae), qui remonte à Gauss, la théorie particulière des nombres entiers s'est élargie à la théorie arithmétique générale des fonctions entières d'indéterminées à coefficients entiers. Cette théorie générale permet de se passer de tous les concepts étrangers à l'arithmétique véritable : les nombres négatifs, fractionnaires, réels et algébriques imaginaires. [Boniface, 1999] p. 59

Il peut ainsi se passer des entiers négatifs en travaillant (*modulo* $x+1$). Pour les nombres rationnels, le système qu'il met en place est plus complexe, et utilise plusieurs indéterminées et plusieurs modules. Ces *Systèmes modulaires* (*Modulsystem*) sont introduits dans les *Grundzüge* et constituent l'un des outils essentiels de ce texte. Les irrationnels ne sont quant à eux pas considérés comme des nombres, au sens où ils ne font pas partie de l'arithmétique, mais de la géométrie. Cependant, on peut avoir accès aux irrationnels algébriques modulo un polynôme irréductible à coefficients entiers, et c'est de cette façon qu'il développera son *théorème fondamental de l'arithmétique* [71].

Dans son article *Sur le concept de nombre*, ce qui sera particulièrement intéressant pour notre étude, c'est le fait que la seconde moitié soit consacrée à la *séparation des racines*.

70. Au sens où nous venons de le voir.
71. Voir son article intitulé *Ein Fundamentalsatz der allgemeinen Arithmetik* [Kronecker, 1887c].

Le cadre dans lequel Kronecker *pense* le nombre est celui de la théorie des équations algébriques : les notions de nombre et de racine semblent donc ici se confondre.

### 2.2.1 La notion de racine

On l'a vu, si l'*Algèbre* de Kronecker ne se limite pas à la résolution des équations, l'objet central des *Vorlesungen über die Theorie der algebraischen Gleichungen* est l'équation, et la notion de racine est donc bien sûr primordiale, que l'on parle de leur recherche, de leur existence ou de leur nature, ces trois qualités étant intimement liées. Pour essayer d'avoir une idée précise de la façon dont Kronecker envisage cela, nous avons choisi de suivre ce qu'il en dit dans l'introduction du manuscrit de 1890-91.

Celle-ci se compose ainsi :

Einleitung / Introduction (1)
Rationalitäts und Integritäts Bereiche / Domaines de rationalité et d'intégrité (5)
Einige vorbereitenden Eigenschaften der Gleichungen / Quelques propriétés préliminaires sur les équations (8)

1. Der Begriff der Wurzeln / La notion de racine (8)
2. Abgrenzung der Gebietes (der Wurzeln) / Délimitation du domaine (des racines) (12)
3. Die zahl der Wurzeln / Le nombre de racines (15)
4. Die reinen Gleichungen oder die Wurzeln der Einheit / Les équations pures ou les racines de l'unité (17)

On voit donc que les propriétés préliminaires sur les équations tournent en fait autour de de celles des racines. Nous allons nous concentrer sur le paragraphe intitulé « La notion de racine » [72] :

72. Es lag nahe, die Operation bei der *potestas pura* vorerst umkehren und rückwärts wieder $x$ aus dem Wert

$$x^n = y$$

($x, n, y$ ganze positive Zahlen) bestimmen zu wollen.

Das Mittelalter schuf dafür das Zeichen : $x = \sqrt[n]{y}$, und weil sich diese Gleichungen angenähert auflösen lassen, so nahm man später das Zeichen für die durch dasselbe angedeutete Operation. Nun ist dieses Zeichen zwar ganz ausschließlich den reinen Gleichungen angepaßt. Als aber mit seiner hilfe die Wurzeln aller Gleichungen des zweiten und später gar des dritten und vierten Grades dargestellt waren, da war von der Mitte des sechzehnten Jahrhunderts an die mathematische Welt widerstandslos beherrschst von dem Aberglauben, daß jede *potestas affecta* beliebigen Grades umkehrbar sei durch Wurzelausziehung. *Vieta*, *Descartes*, *Fermat, Newton, Euler* waren fest überzeugt, daß jede Gleichung sich auflösen lassen müsse. Aber alle Beweise von der Existenz der Wurzeln jeder algebraischen Gleichung würden [Correction dans la marge : « waren »], wie *Gauss* aufdeckte, fehlerhaft. Das hatte seinen Grund darin, daß auch die größten Mathematiker sich über den Begriff der Wurzeln nicht genug Rechenschaft gegeben hatten. *Gauss* legt Gewicht darauf, daß er in den *Disquisitiones* die Kreisteilungsgleichungen aufzeigt als ein Paar von Gleichungen, welches durch Wurzelziehung wirklich auflösbar ist. Er knüpft daran die Bemerkung, daß die Bemühungen um die allgemeine Auflösung vielleicht gescheitert seien an der Unmöglichkeit der Sache.

> Il semblait logique de vouloir d'abord inverser les opérations dans les *potestas pura* et de retrouver ensuite $x$ à partir de la valeur
>
> $$x^n = y$$
>
> ($x, n, y$ sont des nombres entiers positifs).
>
> Le Moyen-Age créa pour cela le symbole : $x = \sqrt[n]{y}$, et comme ces équations se laissent résoudre de façon approchée, nous avons pris plus tard ce symbole pour l'opération en question elle-même. Ce symbole est maintenant certes adapté exclusivement aux équations pures. Mais comme avec son aide les racines du second, et même plus tard celles du troisième et du quatrième degré ont été représentées, au milieu du XVI[e] siècle domine dans le monde mathématiques, sans opposition, la croyance que toutes les *potestas affecta* de n'importe quel degré sont résolubles par extraction de racine. *Viète*, *Descartes*, *Fermat, Newton* et *Euler* étaient fermement convaincus que toute équation peut être résolue.
> Mais toutes les preuves de l'existence des racines d'une équation algébrique, comme *Gauss* l'a montré, étaient défectueuses. Cela est dû aussi au fait que les grands mathématiciens ne se sont pas suffisamment penchés sur le concept de racine. *Gauss* met l'accent sur le fait qu'il montre dans les *Disquisitiones* les équations cyclotomiques comme un ensemble d'équations qui sont réellement résolubles par extraction de racine. Il fait à ce propos remarquer que les efforts sur la résolution générale risquent d'échouer de par l'impossibilité de la chose. [Kronecker, 1891a] p. 8

Ici, dans le contexte du cours de Kronecker, nous avons vu qu'une *aequatio affecta* est une équation quelconque, alors qu'une *aequatio pura* est une équation dans laquelle il ne reste qu'un terme en $x$ [73]. Kronecker, dans ces premières phrases, fait déjà remarquer que le *symbole* racine fait seulement référence à la possibilité de donner, par une méthode qu'il reste à définir, une résolution approchée (*angenähert auflösen*) : il ne parle pas de valeur approchée, qui ferait ainsi référence à une valeur exacte. Il poursuit en faisant allusion à la première partie de la thèse de Gauss de 1799 où celui-ci revient sur les diverses démonstrations du théorème fondamental de l'algèbre [74] et montre leurs faiblesses. Mais surtout, c'est de son travail sur les polynômes cyclotomiques dont Kronecker parle : Gauss a montré que dans ce cas particulier une résolution par radicaux est possible [75], mais il a plusieurs fois remarqué que la possibilité d'une résolution *générale* par radicaux des équations algébriques de degré quelconque était douteuse. Ce travail de Gauss sur les polynômes cyclotomiques, Kronecker le poursuit et aboutit d'ailleurs au théorème sur les extensions abéliennes auquel son nom est associé. Le nom d'Abel n'apparaît pas ici alors qu'il aurait toute sa place. Voilà comment il poursuit [76] :

> Les ambiguïtés de la notion de racine proviennent de différentes sources. D'un

73. Nous reviendrons sur ces termes plus loin (Voir p. 104).

74. Voir [Gauss, 1799].

75. Le $n^{\text{ième}}$ polynôme cyclotomique est : $\Phi_n(X) = \prod_{0 \leq k < n,\ k \wedge n = 1} \left(X - \exp\left(\frac{2ik\pi}{n}\right)\right)$, et donc ses racines sont les racines $n^{\text{ième}}$ de l'unité.

76. Die Unklarheiten des Wurzelbegriffes entsprangen verschieden Quellen. Einerseits traten aus der Praxis heraus die Gleichungen auf als « Rätsel », die bestimmt eine Auflösung zuließen. Denn war eine

> côté, d'un point de vue pratique, les équations apparaissent comme une « énigme » qui permet de déterminer une solution. En effet un nombre rationnel $x$ s'élève à une *potestas affecta* :
>
> $$c_0 + c_1 x ... + c_n x^n = y,$$
>
> de sorte que puissent être retrouvés à partir du $y$ fourni les rationnels $x$. D'un autre côté les équations pures se laissent résoudre par approximations. Pour l'équation $x^2 - 2 = 0$ on peut trouver des nombres $\xi_1$ et $\xi_2$ de sorte que
>
> $$0 < \xi_1^2 - 2 < \tau \ , \ 0 > \xi_2^2 - 2 > -\tau',$$
>
> où $\tau, \tau'$ sont des rationnels positifs aussi petits que l'on veut. On dit alors que la « valeur réelle » de $|\sqrt{2}|$ se trouve entre $\xi_1$ et $\xi_2$ et on peut avec de tels nombres $\xi_1$ et $\xi_2$ approcher d'aussi près que l'on veut $\sqrt{2}$. On oublie ainsi complètement que $\sqrt{2}$ s'explique uniquement par l'équation : $(\sqrt{2})^2 - 2 = 0$ et que l'on peut s'approcher avec une suite de nombres d'un nombre existant (donc rationnel), mais pas d'un être qui est seulement « défini » par une exigence. [Kronecker, 1891a] p. 9

Une équation, dit Kronecker, peut être vue comme un problème à résoudre : il est alors naturel de chercher à en donner une solution. Et si l'on conçoit la résolution d'une équation comme synonyme de *trouver* $x$, alors une confusion peut survenir : $\sqrt{2}$ n'est pas un « un nombre existant ». Ce qui « existe », c'est l'équation $x^2 - 2 = 0$, ce n'est pas $\sqrt{2}$. On retrouve cette même idée dans les notes prises par Hilbert (1862-1943) lors de son voyage à Berlin en 1888 à la suite d'un long entretien avec Kronecker [77] :

> Une égalité est seulement $2 = 2$. Les nombres irrationnels et transcendants sont donnés 1) par la représentation implicite $\sin x = 0, x^2 = 5$ ou 2) par approximation.

---

rationale Grundzahl $x$ erhoben zu einer *potestas affecta* :

$$c_0 + c_1 x ... + c_n x^n = y,$$

so mußte rückwärts aus dem vorgelegten $y$ des rationale $x$ wieder zu erraten sein. Anderseits ließen die reinen Gleichungen sich angenähert lösen. Für die Gleichung $x^2 - 2 = 0$ ließen sich Zahlen $\xi_1$ und $\xi_2$ finden, so daß

$$0 < \xi_1^2 - 2 < \tau \ , \ 0 > \xi_2^2 - 2 > -\tau',$$

wo $\tau, \tau'$ beliebig kleine positiver rationale zahlen sind. Man sagte dann, der « wahre Wert » von $|\sqrt{2}|$ läge zwischen $\xi_1$ und $\xi_2$ und man könne sich mit solchen Zahlen $\xi_1$ und $\xi_2$ der $\sqrt{2}$ beliebig nähern. Man vergaß dabei ganz, daß $\sqrt{2}$ allein erklärt ist durch die Gleichung : $(\sqrt{2})^2 - 2 = 0$ und daß man sich mit einer Reihe von Zahlen wohl einer vorhandenen (also rationalen) Zahl nähern könne, nicht aber einem Wesen, das allein durch eine Forderung « definiert » ist.

77. Gleich sei nur $2 = 2$. Irrationale und transzendente Zahlen seien 1.) durch die implizite Darstellung $sinx = 0, x^2 = 5$ gegeben oder 2.) durch Annäherung. Im allgemeinen sei es gar nicht schwer auf dieser Grundlage bei Vermeidung des Weierstrass'schen Begriffs der Gleichheit und der Continuirlichkeit alles strenge aufzubauen. Aber bei gewissen kritischen Stellen sei es schwer und da könne man nicht genau und strenge genug sein. Nur das Diskrete und Singuläre habe Bedeutung. Das übrige könne man aber auch erhalten, nämlich durch Interpolation. So sei es bei den elliptischen Funktionen sein Ziel, nur die singulären Moduln zuzulassen und dementsprechend alles arithmetisch aufzubauen.

Bei irrationalen Betrachtungen vermenge man häufig 1) und 2) in dem man unter dem Zeichen $\sqrt{5}$ sowohl das implicit durch $x^2 = 5$ gegebene Ding, wie auch die rational gegebene Annäherung verstehe. Daraus entstehe eine entsetzliche Confusion. Man verlangt so z.B. die Gleichung $n + \sqrt{5}v = 0$ zu lösen, was ganz unmöglich ist. Man kann nur verlangen, dass $n + \sqrt{5}v$ möglich klein werde, d.h. $= \frac{1}{N}$.

> En général, il n'est pas du tout difficile de construire sur ces bases tout en évitant la notion de Weierstrass d'égalité et de continuité. Mais à un certain moment critique cela devient difficile et on ne peut pas être assez précis et rigoureux. Seuls le discret et le singulier ont du sens. Mais on peut aussi obtenir le reste, à savoir par interpolation. Ainsi son but est avec les fonctions elliptiques d'autoriser seulement les modules singuliers et en conséquence établir tout arithmétiquement.
>
> Par considérations irrationnelles on mélange souvent 1) et 2) en comprenant sous le signe de $\sqrt{5}$ ce qui est donné implicitement par $x^2 = 5$, comme aussi l'approximation rationnelle donnée. Il se forme de là une épouvantable confusion. On demande ainsi de résoudre par exemple l'équation $n + \sqrt{5}v = 0$, ce qui est totalement impossible. On peut seulement demander que $n + \sqrt{5}v$ devienne aussi petit que l'on veut, c'est-à-dire plus petit que $= \frac{1}{N}$. [Hilbert, 1888] pp. 1-2

Un irrationnel n'aura donc pas de valeur approchée, même rationnelle, étant donné qu'il n'existe pas en tant que nombre. Kronecker poursuit en analysant l'origine de cette *épouvantable confusion* (*entsetzliche Confusion*) [78] :

> Cette conception était alimentée par la géométrie analytique. On voit ou on sait en effet que la courbe
>
> $$y = x^3 + c_1x^2 + c_2x + c_3$$
>
> passe pour des grandeurs négatives $x$ sous l'axe des $x$, pour des grandeurs positives au dessus. Ne doit-elle pas au moins une fois, comme elle est en effet continue, couper l'axe des $x$ ? Chaque équation cubique ne possède-t-elle pas au moins une racine réelle ?
>
> C'est cette « déduction » qui est encore et toujours une *petitio principii* grossière, bien que *Gauss* aussi ait pris comme principe la proposition disant que toute équation de degré impair a une racine réelle. En réalité $y$ se laisse calculer uniquement pour des $x$ rationnels, mais n'est jamais réalisée entre les points ainsi obtenus une continuité, une « courbe ». C'est pourquoi sont aussi trompeuses les déductions de Bolzano, un ecclésiastique catholique de Prague, qui veut démontrer pour une fonction continue quelconque, que si elle a des signes différents dans son domaine de continuité, elle doit passer une fois par zéro. [Kronecker, 1891a] p. 10

C'est donc la conception géométrique d'une courbe qui nous leurre, et son pendant analytique, c'est-à-dire en fait le théorème des valeurs intermédiaires. Nous préciserons ensuite

78. Eine beständige Nahrung fand diese Vorstellung in der analytischen Geometrie. Wußte oder sah man doch, daß die Kurve

$$y = x^3 + c_1x^2 + c_2x + c_3$$

für großes negatives $x$ unterhalb der $x$ Achse, für großes positives oberhalb verläuft. Muß sie also nicht wenigstens einmal, da sie doch stetig ist, die $x$ Achse schneiden ? Besitzt nicht also jede kubische Gleichung mindestens eine reelle Wurzel ?

Es ist diese « Folgerung » immer wieder dieselbe grobe *petitio principii*, obgleich auch *Gauss* den Satz zur Grundlage nimmt, daß jede Gleichung ungeraden Grades eine reelle Wurzel hat. In Wahrheit läßt sich allein für rationales $x$ das zugehörige $y$ berechnen, niemals aber zwischen den so gewonnenen Punkten, seien ihrer auch noch so viele, eine Stetigkeit, eine « Kurve » erzielen. Deshalb sind auch trügerisch die Deduktionen Bolzanos, eines katholischen Geistlichen aus Prag, der nun gar von einer beliebigen stetigen Funktion beweisen will, daß, wenn sie im Bereiche der Stetigkeit verschiedenes Zeichen hat, sie einmal durch Null hindurchgehen muß.

le rapport de Kronecker avec le théorème de Bolzano, mais voyons comment il conclut cette partie [79] :

> Les résultats de l'algèbre doivent être rendus autant que possibles indépendants de toutes les fictions sur les racines des équations. Cela n'a pas encore été réussi dans l'important champ de la théorie de l'élimination. Certes ses résultats en tant que tels sont complètement indépendants de la notion de racine, mais une dérivation de ceux-ci sans l'aide des racines n'a pas été encore possible. [Kronecker, 1891a] p. 10

Plus encore, c'est le concept même de *racine* qu'il s'agit d'exclure de la théorie de l'élimination. On obtient d'importants résultats dans la théorie des équations en passant par la forme factorisée d'une équation, et donc en utilisant ses racines. Les propriétés que l'on obtient, comme par exemple celles relatives au discriminant, ne font pas toujours apparaître ces racines : il faut maintenant, dit Kronecker, s'évertuer à ne pas les utiliser lors de la démonstration de ces propriétés. Il poursuit ainsi [80] :

> La théorie de l'élimination montre entre autres qu'en principe seulement les propriétés d'une équation algébrique ont besoin d'être connues ; en effet le contenu entier d'un système d'équations algébriques peut toujours être réduit à une équation. Par un système de $\mu$ équations entre $\nu$ variables :
>
> $$F_k(x_1, x_2 ... x_\nu) = 0 \qquad (k = 1, 2 ... \mu)$$
>
> les $\nu$ variables ne sont ensuite pas définies autrement que comme des fonctions des coefficients présents dans les équations, comme cela peut se faire par une équation. Les fonctions ainsi définies se laissent ramener à des équations telles que celles qui

79. Die Resultate der Algebra sind nach Möglichkeit von allen Fiktionen über die Wurzeln der Gleichungen unabhängig zu machen. Noch nicht gelungen ist dies bisher auf dem wichtigen Felde der Eliminationstheorie. Zwar ihre Resultate als solche sind ganz unabhängig vom Begriff der Wurzeln, aber eine Herleitung derselben ohne Zuhilfenahme der Wurzeln ist noch nicht möglich geworden.

80. Die Eliminationstheorie zeigt unter anderem, daß prinzipiell nur die Eigenschaften einer algebraischen Gleichung gekannt zu werden brauchen, da sich der ganze Inhalt eines Systemes von algebraischen Gleichungen immer in eine Gleichung übertragen läßt. Durch ein System von $\mu$ Gleichungen zwischen $\nu$ Variabelen :

$$F_k(x_1, x_2 ... x_\nu) = 0 \qquad (k = 1, 2 ... \mu)$$

werden danach die $\nu$ Variabelen nicht anders als Funktionen der in den Gleichungen auftretenden Koefficienten definiert, wie es durch eine Gleichung geschehen kann. Die so definierten Funktionen lassen sich zurückführen auf Funktionen, wie sie erzeugt werden durch eine Gleichung :

$$c_0 + c_1 x + c_2 x^2 ... + c_n x^n = 0$$

Aber auch wenn $x$ betrachtet wird als mittels dieser Gleichung erklärt als Funktion von $\left(\frac{c_0}{c_n}, \frac{c_1}{c_n} ... \frac{c_{n-1}}{c_n}\right)$, so ist damit das Dunkel in des Existenz der Wurzeln nicht aufgehellt. Logisch ist der Begriff der Funktion durch eine solche Gleichung wohl festgelegt ; Die Mathematik aber braucht Größen, mit denen sie rechnen kann. Darum betont *Gauss* nicht ohne Grund, daß die imaginären Größen, die ja allein als algebraische Funktionen erklärt sind, sich den Gesetzen der Arithmetik fügen. (Freilich übersieht er dabei das sehr wichtige Gesetz, daß ein Produkt nur dann null werden kann, wenn ein Faktor verschwindet.) Den algebraischen Funktionen aber stellt sich von vornherein entgegen die gefährliche Eigenschaft der Vieldeutigkeit, und niemals ist bis in die Einzelheiten nachgewiesen worden, wie weit sich die algebraischen Funktionen der Arithmetik fügen. Alle Schwierigkeiten aber werden vermieden, wenn die Theorie der algebraischen Funktionen aufgeht in einer streng arithmetisch durchgeführten Theorie der ganzen oder wenigstens rationalen Funktionen.

sont engendrées par une équation :

$$c_0 + c_1 x + c_2 x^2 ... + c_n x^n = 0$$

Mais aussi si $x$ est considéré au moyen de cette équation comme fonction de $\left(\frac{c_0}{c_n}, \frac{c_1}{c_n} ... \frac{c_{n-1}}{c_n}\right)$, alors ainsi l'obscurité dans l'existence des racines n'est pas éclaircie. La notion de fonction est bien définie logiquement par une telle équation ; mais les mathématiques ont besoin de grandeurs avec lesquelles elles peuvent calculer. C'est pourquoi *Gauss* insiste non sans raison sur le fait que les grandeurs imaginaires, qui sont expliquées seulement comme des fonctions algébriques, se soumettent aux lois de l'arithmétique. (Toutefois il ne remarque pas la très importante propriété qui affirme qu'un produit de facteurs peut être nul seulement lorsque l'un des facteurs est nul). Mais les fonctions algébriques se tiennent a priori éloignées de la périlleuse propriété de multivalence, et il n'a jamais été précisé jusque dans les détails jusqu'à quel point les fonctions symétriques se soumettaient à l'arithmétique. Mais toutes les difficultés seront évitées si la théorie des fonctions algébriques se réalise dans une théorie proprement arithmétique des fonctions entières, ou tout au moins rationnelles. [Kronecker, 1891a] p. 10-11

Après avoir affirmé qu'un système d'équations peut toujours se réduire à l'étude des propriétés d'une unique équation, Kronecker relève l'importance d'avoir des « grandeurs avec lesquelles elles peuvent calculer » : si une grandeur algébrique n'est rien d'autre que l'équation qui la définit, il n'en faut pas moins déterminer des lois pour « calculer » avec ces grandeurs. C'est ce que se propose de faire Kronecker dans une « théorie proprement arithmétique des fonctions entières, ou tout au moins rationnelles », et donc dans son *arithmétique générale*. Il semble donc que considérer les racines comme des grandeurs ne soit pas toujours pertinent pour Kronecker : voyons comment il transforme le problème.

### 2.2.2 La séparation des racines

Comme nous l'avons dit, une partie essentielle de l'unique article [81] où Kronecker a publié son point de vue sur les mathématiques traite de la séparation des racines. Kronecker, dans son cours du semestre d'hiver de 1886-87 sur la théorie des équations algébriques – c'est-à-dire durant l'année qui précède la publication de cet article – affirme que le but essentiel du théorème de Sturm est la « détermination du nombre des racines réelles des équations algébriques » [82]. Il ajoute que la plupart des ouvrages considèrent ce problème comme définitivement résolu. Mais, ce qu'il reste à accomplir, c'est la *séparation des racines* [83]. Et c'est ce qu'il fait ici [84], en commençant par délimiter un domaine dans lequel se trouvera l'ensemble des racines.

81. *Über den Zahlbegriff* [Boniface, 1999].
82. [Kronecker, 1887b] p. 155 : « die numerische Bezeichnung der reellen Wurzeln der algebraischen Gleichungen ».
83. *Ibid.*
84. Nous rappelons que nous sommes en train de suivre l'introduction du cours de l'hiver 1890/91, et que nous passons maintenant au second chapitre : *Délimitation du domaine (des racines).*

**Le but :** Soit $F(x)$ une fonction réelle de la variable réelle $x$, continue, et qui prend des valeurs positives et négatives. Le but de Kronecker est de trouver un intervalle dans lequel cette fonction ne change pas plus d'une fois de signe.

**Un exemple :** Kronecker prend l'exemple suivant [85] :

$$F(z) = \left(\sum_{n=1}^{+\infty} \frac{\sin n^2 z\pi}{n^2}\right) - \frac{1}{16\sqrt{2}}.\pi^2$$

converge pour tout $z$, elle est continue, et prend pour $z = 0$ une valeur négative et pour $z = \frac{1}{4}$ une valeur positive. Elle devrait donc être nulle pour une valeur située dans l'intervalle $[0, \frac{1}{4}]$. La méthode que propose Bolzano, méthode par dichotomie, consiste à chercher si la fonction s'annule dans la première ou dans la seconde partie de l'intervalle. Mais non seulement, dit Kronecker, le calcul même des valeurs aux bornes est problématique, il est en plus impossible de répondre à cette question. Dans ce passage, nous sommes vraiment très proches des thèses que défendront les intuitionnistes, et en particulier Brouwer. Nous reviendrons plus en détail sur les difficultés que posent pour Kronecker le travail et les méthodes de Bolzano.

**Détermination d'un intervalle dans lequel il n'y a qu'une racine** On suppose maintenant que nous connaissons la notion de racine d'une équation algébrique, telle que nous l'avons décrite dans la première partie : les coefficients de l'équation que l'on considère sont donc des nombres rationnels. On trouve un intervalle dans lequel on est certain qu'il n'y ait qu'une seule racine de $F(x) = 0$. Soit $\xi_1, ..., \xi_n$ les racines de l'équation. On suppose construite l'équation au carré des différences, dont les racines sont donc

$$(\xi_1 - \xi_2)^2, (\xi_1 - \xi_3)^2, ..., (\xi_1 - \xi_n)^2, (\xi_2 - \xi_3)^2, ..., (\xi_{n-1} - \xi_n)^2$$

85. Die ganze Frage kommt offenbar darauf hinaus, ein Intervall zu finden, von dem man weiß, daß in ihm die Funktion nie mehr als einmal ihr Zeichen wechselt. Versucht man nun dieses Problem mit Hilfe der Bolzano'schen Methode bei einer beliebigen, stetigen, eindeutigen, ganz bestimmt definierten Function zu lösen, so wird man sofort deren Mängel erkennen. Durch die Reihe $\sum_{n=1}^{\infty} \frac{\sin n^2 z\pi}{n^2}$, die für jeden reellen Wert von $z$ convergiert, ist ein ganz bestimmte, eindeutige und stetige Function definiert ; für $z = 0$ hat sie den Wert Null ; für $z = \frac{1}{2}$ verschwinden alle Summanden, die zu einem geraden $n$ gehören, daher ist für $z = \frac{1}{2}$ der Wert der Function $\frac{\pi^2}{8}$. Für $z = \frac{1}{4}$ verschwinden jene Glieder ebenfalls ; für ungerades $n$ wird der Wert des Sinus $\sin \frac{\pi}{4} = \frac{1}{\sqrt{2}}$ ; für $z = \frac{1}{2}$ ist also der Wert der Function $\frac{\pi^2}{8\sqrt{2}}$. – Die Function

$$\sum_{n=1}^{\infty} \frac{\sin n^2 z\pi}{n^2} - \frac{1}{16\sqrt{2}}.\pi^2$$

ist also für $z = 0$ negativ, für $z = \frac{1}{4}$ positiv ; sie müßte also für einen Wert von $z$ in dem Intervall von 0 bis $\frac{1}{4}$ verschwinden. ([Kronecker, 1887b] p. 157)

On sait que l'on peut exprimer rationnellement les coefficients de cette équation avec les polynômes symétriques élémentaires des racines, et donc avec les coefficients de $F$. Soit

$$F(x) = x^n - c_1 x^{n-1} + ... \pm c_n$$

et l'équation de degré $\frac{n(n-1)}{2}$

$$\Phi(x, c_1, c_2, ..., c_n) = \prod_{1 \leq h < k \leq n} \left(x - (\xi_h - \xi_k)^2\right) = 0$$

On souhaite donc trouver la plus petite racine de cette dernière équation, ou, ce qui revient au même, trouver un intervalle du type $[0, \delta^2]$ dans lequel on est sûr que $\Phi = 0$ n'a aucune racine. On peut supposer sans perte de généralité que les racines sont toutes simples.

Soit

$$F(x) = a_0 z^n + a_1 z^{n-1} + ... + a_n = 0$$

On peut se limiter au cas où $F$ est à coefficients entiers et où toutes les racines sont simples. On peut ainsi trouver deux fonctions entières $\Phi$ et $\Psi$ de degrés respectifs $n-1$ et $n$ et n'ayant que des facteurs irréductibles simples tels que

$$F\Phi + F'\Psi = 1$$

Kronecker veut alors montrer que l'ensemble des racines se situe dans un intervalle borné. Cette propriété, il la décrit comme celle de la courbe associée à l'équation, et non de l'équation elle-même : l'ensemble des points d'intersection de la courbe $y = F(z)$ se situe sur un morceau fini (*ein endlichen Stück*) de l'axe des abscisses.

On écrit l'équation sous la forme :

$$z^n + \frac{a_1}{a_0} z^{n-1} + ... + \frac{a_n}{a_0} = 0$$

Soit $g$ un nombre tel que

$$\sup\left(\left|\frac{a_1}{a_0}\right|, ..., \left|\frac{a_n}{a_0}\right|\right) \leq g - 1$$

donc

$$\left|\frac{a_1}{a_0} z^{n-1} + ... + \frac{a_n}{a_0}\right| \leq (g-1)\frac{1 - |z|^n}{1 - |z|}$$

et pour $|z| > g$, on a :

$$\left|\frac{a_1}{a_0} z^{n-1} + ... + \frac{a_n}{a_0}\right| \leq |z|^n - 1$$

Finalement,

$$\left|z^n + \frac{a_1}{a_0} z^{n-1} + ... + \frac{a_n}{a_0}\right| \geq |z^n| - \left|\frac{a_1}{a_0} z^{n-1} + ... + \frac{a_n}{a_0}\right| > 1$$

Donc toutes les racines réelles de $F(z) = 0$ se trouvent dans l'intervalle $[-g, g]$. Kronecker

a donc réussi à borner son ensemble de recherche. Lorsqu'il cherchera à démontrer le théorème fondamental de l'algèbre, tout au moins lorsqu'il le fera en adaptant les première, troisième et quatrième démonstrations de Gauss, ce travail sera un préliminaire à ses démonstrations. Nous en reparlerons dans la troisième partie.

Il doit maintenant trouver, à l'intérieur de l'intervalle précédent, un intervalle plus petit dans lequel la fonction ne changera de signe qu'une fois. Soit

$$F(z+h) = F(z) + F'(z).h + \bar{F}(z,h).h^2$$

où $\bar{F}$ est une fonction entière de $z$ et $h$. En restant dans l'intervalle ainsi déterminé, on peut trouver $G$ telle que

$$\begin{aligned} |\Phi| &< G \\ |\Psi| &< G \\ |\Psi.\bar{F}| &< G \end{aligned}$$

On peut alors poser :

$$\begin{aligned} |\Phi(z)| &= G \cdot \gamma(z) \\ |\Psi(z)| &= G \cdot \gamma'(z) \\ |\Psi(z) \cdot \bar{F}(z,h)| &= G\gamma_0(z,h) \end{aligned}$$

où $\gamma(z), \gamma'(z)$ et $\gamma_0(z,h)$ sont en valeur absolue strictement inférieurs à 1 sur l'intervalle en question.

Si l'on revient maintenant sur l'équation

$$F\Phi + F'\Psi = 1$$

en considérant

$$F'(z) = \frac{F(z+h) - F(z)}{h} - h \cdot \bar{F}(z,h)$$

alors on obtient

$$F(z)\Phi(z) + \Psi(z) \cdot \frac{F(z+h) - F(z)}{h} - h \cdot \bar{F}(z,h)\Psi(z) = 1$$

donc

$$F(z)\gamma(z) + \gamma'(z) \cdot \frac{F(z+h) - F(z)}{h} = \frac{1 + h \cdot G \cdot \gamma_0(z,h)}{G}$$

Si maintenant on suppose de plus que $|h| \leq \frac{1}{2G}$ alors

$$|h \cdot G \cdot \gamma_0(z,h)| < \frac{1}{2}$$

Ainsi

$$F(z)\gamma(z) + \gamma'(z) \cdot \frac{F(z+h) - F(z)}{h} > \frac{1}{2G}$$

On peut donc écrire

$$|F(z)| + \left| \frac{F(z+h) - F(z)}{h} \right| > \frac{1}{2G}$$

où $|h| \leq \frac{1}{2G}$ et $-g < z < g$.

Et donc pour $z_1$ et $z_2$ dans $[-g; g]$ avec $|z_1 - z_2| < \frac{1}{2G}$ on a

$$(I) \qquad |F(z_1)| + \left| \frac{F(z_2) - F(z_1)}{z_2 - z_1} \right| > \frac{1}{2G}$$

De plus, on peut aussi écrire

$$F(z+h) - F(z) = h \cdot E(z,h)$$

Il existe $G'$ telle que dans les mêmes conditions que précédemment $E(z,h) < G'$ et donc

$$F(z+h) - F(z) < h \cdot G'$$

ou

$$(II) \qquad |F(z_2) - F(z_1)| < |z_2 - z_1| \cdot G'$$

S'il existe $z_0$ tel que

$$|z_2 - z_1| < \frac{|F(z_0)|}{G'}$$

alors

$$|F(z_2) - F(z_1)| < |F(z_0)|$$

Nous voulons montrer que dans les intervalles de longueur $\frac{1}{2G}$, $F(z)$ change soit une, soit aucune fois de signe. On suppose donc que l'intervalle $[-g, g]$ est découpé en intervalles de longueur $\frac{1}{2G}$ et nous nommons l'un d'entre eux $I$. On montre tout d'abord que (1) lorsque $F(z)$ a le même signe aux bornes de cet intervalle, le signe est constant sur l'intervalle, et que donc le comportement de la fonction aux bornes de l'intervalle nous donne une indication sur le comportement de la fonction sur l'intervalle entier. En faisant cette remarque, Kronecker nous montre que la qualité de cette méthode réside en grande partie dans la possibilité d'éviter l'examen d'un ensemble continu de points (l'intervalle

tout entier) et de se contenter d'un ensemble discret (les deux bornes de l'intervalle). Ensuite (2) on montrera que si la fonction a des signes opposés aux bornes de l'intervalle alors elle changera une et une fois seulement de signe à l'intérieur de celui-ci.

Commençons par montrer que, dans le cas où $F$ a des signes opposés aux bornes de $I$, pour $\sigma > 0$ aussi petit que l'on veut, on peut trouver $z$ dans cet intervalle tel que $|F(z)| < \sigma$. On prend de plus $\sigma$ suffisamment petit pour que

$$\frac{\sigma}{G'} < \frac{1}{2G}$$

D'après ce qui précède, on peut trouver $z_1$ et $z_2$ dans $I$ tels que $z_2 - z_1 = \dfrac{\sigma}{G'}$ et, $F$ ayant des signes opposés aux bornes de $I$, tels que $F(z_1)$ et $F(z_2)$ soient de signes opposés [86]. On peut alors appliquer l'inégalité $(II)$ :

$$|F(z_2) - F(z_1)| < |z_2 - z_1| \cdot G' = \frac{\sigma}{G'} \cdot G' = \sigma$$

Comme on a choisi $z_1$ et $z_2$ dans $I$ tels que $F(z_1)$ et $F(z_2)$ soient de signes opposés, on a

$$|F(z_2) - F(z_1)| = |F(z_2)| + |F(z_1)| < \sigma$$

Donc $|F(z_1)| < \sigma$ et $|F(z_2)| < \sigma$. On a donc montré que pour tout $\sigma > 0$, il existe $z \in I$ tel que $|F(z)| < \sigma$.

Supposons maintenant que $F$ a le même signe aux bornes de $I$, et on peut par exemple supposer que ce signe est positif. On pose $I = \left[z_0; z_0 + \dfrac{1}{2G}\right]$. Kronecker entreprend maintenant une démonstration par l'absurde. Supposons que pour $z' \in I$ on ait $F(z') < 0$. D'après $(II)$, on a :

$$|F(z_0) - F(z')| < |z_0 - z'| \cdot G'$$

donc, comme $F(z_0) > 0$ et $F(z') < 0$, on a $|F(z_0) - F(z')| > |F(z')|$, et donc

$$|F(z')| < |z_0 - z'| \cdot G'$$

ou

$$\frac{|F(z')|}{G'} < |z_0 - z'|$$

On a de même

$$\frac{|F(z')|}{G'} < \left|z_0 + \frac{1}{2G} - z'\right|$$

On peut donc découper l'intervalle $I$ en trois : $\left[z_0; z' - \dfrac{|F(z')|}{G'}\right]$,

86. Kronecker ne propose pas de démonstration de ce dernier fait.

$\left[z' - \frac{|F(z')|}{G'}; z' + \frac{|F(z')|}{G'}\right]$ et $\left[z' + \frac{|F(z')|}{G'}, z_0 + \frac{1}{2G}\right]$. Avec l'inégalité $(II)$, on obtient :

$$\left|F(z') - F\left(z' - \frac{|F(z')|}{G'}\right)\right| < G' \left|z' - \left(z' - \frac{|F(z')|}{G'}\right)\right| = |F(z')|$$

et de même

$$\left|F(z') - F\left(z' + \frac{|F(z')|}{G'}\right)\right| < G' \left|z' - \left(z' + \frac{|F(z')|}{G'}\right)\right| = |F(z')|$$

ce qui n'est possible que pour

$$sgn\left(F\left(z' - \frac{|F(z')|}{G'}\right)\right) = sgn(F(z'))$$

et

$$sgn\left(F\left(z' + \frac{|F(z')|}{G'}\right)\right) = sgn(F(z'))$$

ce qui n'est possible que si $sgn\left(F\left(z' - \frac{|F(z')|}{G'}\right)\right)$ et $sgn\left(F\left(z' + \frac{|F(z')|}{G'}\right)\right)$ sont négatifs. Ainsi, les intervalles $\left[z_0; z' - \frac{|F(z')|}{G'}\right]$ et $\left[z' + \frac{|F(z')|}{G'}, z_0 + \frac{1}{2G}\right]$ ont des extrémités de signes opposés : pour $\sigma > 0$, on peut trouver $z_1 \in \left[z_0; z' - \frac{|F(z')|}{G'}\right]$ et $z_2 \in \left[z' + \frac{|F(z')|}{G'}, z_0 + \frac{1}{2G}\right]$ tels que $|F(z_1)| < \sigma$ et $|F(z_2)| < \sigma$. On a de plus

$$|z_2 - z_1| > 2 \cdot \frac{|F(z')|}{G'}$$

D'après $(I)$,

$$|F(z_1)| + \left|\frac{F(z_2) - F(z_1)}{z_2 - z_1}\right| > \frac{1}{2G}$$

mais ici

$$|F(z_1)| + \left|\frac{F(z_2) - F(z_1)}{z_2 - z_1}\right| < \sigma + \frac{2\sigma}{2 \cdot \frac{|F(z')|}{G'}}$$

Comme on peut prendre $\sigma$ aussi petit que l'on veut, il y a une contradiction dans les deux inégalités précédentes : on peut conclure qu'il n'existe pas de valeur $z' \in I$ tel $F(z') < 0$. (1) vient donc d'être démontré. Remarquons que nous venons de démontrer qu'il *n'existe pas* de valeur $z'$ : une démonstration d'existence correcte pour Kronecker aurait supposé la construction de ce $z'$.

Reprenons notre intervalle $I = \left[z_0; z_0 + \frac{1}{2G}\right]$ et supposons cette fois que $F$ a des

signes opposés aux bornes de cet intervalle. Si on peut trouver deux valeurs distinctes pour lesquelles $F(z)$ peut être rendue aussi petite que l'on veut, alors on peut trouver deux intervalles disjoints $[z_1, z_2]$ et $[z'_1, z'_2]$ où $F$ prend aux bornes des signes opposés [87] avec $sgn(F(z_1)) = sgn(F(z_0))$ et $sgn(F(z'_2)) = sgn\left(F(z_0 + \frac{1}{2G})\right)$. On a alors :

$$sgn(F(z_0)) = sgn(F(z'_1))$$

et donc d'après (1) le signe de $F$ est constant sur $[z_0; z'_1]$. De même le signe de $F$ est constant sur $[z_2; z_0 + \frac{1}{2G}]$, donc finalement le signe de $F$ doit être constant sur $[z_0; z_0 + \frac{1}{2G}]$, ce qui contredit notre hypothèse. Il y aura donc une unique valeur dans $I$ pour laquelle $F$ peut être rendue aussi petite que l'on veut.

Kronecker résume alors ce qui vient d'être fait [88]

> Si nous regroupons les résultats des développements précédents, alors nous avons montré ce qui suit. Nous pouvons séparer un tel intervalle fini $-g...+g$ du domaine des grandeurs réelles, tel que en dehors de celui-ci $F(z)$ n'est pas nulle ; nous pouvons à nouveau découper cet intervalle en intervalles de taille $\frac{1}{2G}$, où $G$ est un nombre défini, de sorte que dans chacun de ces intervalles la fonction $F(z)$, soit garde le même signe, soit change seulement une fois de signe et donc la fonction devient aussi petite que l'on veut ; et le plus important de tout est que nous avons déterminé nos intervalles par des nombres rationnels. Habituellement on exprime cette propriété des équation en disant que dans un intervalle où la fonction $F(z)$ a des signes différents à ses bornes, il y a une « racine » de l'équation, que l'on ne peut certes pas donner sous forme de nombres rationnels, mais qui existe bien. [Kronecker, 1887b] p. 165

Kronecker a ainsi pu séparer les racines, et même donner une méthode effective pour les séparer, sans utiliser le concept de nombre réel, tout en se plaçant dans un « domaine de grandeurs réelles » (*aus dem Gebiete der reellen Grösse z*). On voit bien ici que pour lui, la notion de racine d'une équation est indépendante de celle de nombre réel, et que le travail de recherche des solutions exactes est donc remplacé par la séparation des racines, sous la forme d'une suite d'intervalles à bornes rationnelles.

### 2.2.3 Le théorème de Bolzano

Nous savons que Lagrange (1736-1813) a tenté de résoudre *algébriquement* le problème de l'existence d'une racine entre deux bornes où la fonction prend des valeurs de signes

87. Les hypothèses qui permettent à Kronecker d'affirmer cela sont les suivantes : $F$ est un polynôme n'ayant que des racines simples. Toute racine réelle signifie donc un changement de signe.

88. Fassen wir das Ergebnis der vorstehenden Entwicklungen noch einmal zusammen, so hat sich folgendes gezeigt. Wir können aus dem Gebiete der reellen Grösse $z$ ein solches endliches Intervall $-g...+g$ aussondern, dass ausserhalb desselben $F(z)$ von null verschieden ist ; wir können dies Intervall wieder in Teilintervalle von der Grösse $\frac{1}{2G}$, wo $G$ ein bestimmt definierte Zahl ist teilen, so dass in jedem dieser Teilintervalle die Function $F(z)$ entweder ihr Zeichen beibehält oder nur einmal wechselt und dass die Function dann beliebig klein wird ; und wesentlich bei Allem war, dass wir unsere Intervalle durch rationale Zahlen bestimmt haben. Gewöhnlich drückt man diese Eigenschaft der Gleichungen so aus, dass man sagt, in einem Intervall, an dessen Anfang und Ende die Function $F(z)$ verschiedenen Zeichen hat, liege eine « Wurzel » der Gleichung, die man zwar nicht als rationale Zahl angeben könne, die aber doch existiere.

contraires, mais ces tentatives se sont soldées par un constat d'échec. Voilà ce que Kronecker dit de cette proposition [89]

> Bolzano, un prêtre catholique de Prague, s'est beaucoup occupé de telles questions plus transcendantales et philosophiques dans les mathématiques ; c'était un mathématicien plein de talent, mais concernant la question précédente il était dans l'erreur. Cela est dû au fait qu'il n'a pas pu s'affranchir de l'erreur qui consiste à mélanger la géométrie et l'arithmétique. « Mais ce pont entre l'arithmétique et la géométrie ne réussira pas ».
>
> Le fait que les erreurs logiques des déductions de Bolzano n'apparaissent pas d'emblée vient de sa forme elle-même, celle d'un dialogue, telle qu'elle apparaît chez Bolzano. Il pose toujours des questions auxquelles on ne peut répondre que par « oui » ou « non », mais il n'a pas cherché au préalable à savoir si ses questions sont posées de sorte que l'on puisse y répondre. Si c'est le cas, alors sa démonstration est juste, mais pas nécessaire. Sinon la déduction entière est fausse. [Kronecker, 1887b] p. 156

Ce passage fait bien sûr allusion au théorème des valeurs intermédiaires démontré par Bolzano (1781-1848) dans la bien nommée *Rein analytischer Beweis des Lehrsatzes daß zwischen je zwei Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege*. Il faut bien remarquer que le théorème de Bolzano est faux si l'on se restreint au corps des rationnels. Prenons la fonction continue $f(x) = x^2 - 2$ : on a alors $f(0) = -2 < 0$ et $f(2) = 2 > 0$. Or $x^2 - 2 = 0$ n'a pas de solution rationnelle sur l'intervalle $]0; 2[$. La démonstration de Bolzano doit donc utiliser la « complétude » du corps des réels. Nous avons déjà croisé des propos similaires dans les paragraphes précédents lorsque nous avons parlé de la notion de racine [90]. Il y a ici une double affirmation : la géométrie n'a pas à intervenir dans l'arithmétique – donc dans l'arithmétique générale qui contient l'analyse – et il souligne la nécessaire effectivité des réponses que l'on doit donner en mathématiques. Si on cherche à savoir si une fonction s'annule, on ne peut pas se contenter d'un oui ou d'un non : on doit pouvoir *l'exhiber*. Kronecker poursuit en donnant l'exemple de :

$$F(z) = \sum_{n=1}^{+\infty} \frac{\sin n^2 z\pi}{n^2} - \frac{1}{16\sqrt{2}} \cdot \pi^2$$

que l'on a déjà étudié plus haut [91]. À partir de cette expression, Kronecker montre ce qui ne convient pas, d'après lui, dans le théorème des valeurs intermédiaires tel qu'il est

89. Bolzano, ein Katholischer Geistlicher in Prag, hat sich viel mit solchen mehr transcendentalen und philosophischen Fragen in der Mathematik beschäftigt ; er war ein talentvoller Mathematiker, aber in bezug auf die vorliegende Frage ist er im Irrtum. Das liegt besonders daran, dass er von dem Fehler, Arithmetik und Geometrie zu vermengen, sich nicht losmachen konnte. „Diese Brücke aber zwischen Arithmetik und Geometrie zu schlagen, wird nicht gelingen".

Dass die logischen Fehler der Bolzano'schen Deduction nicht so ohne Weiteres hervortreten, hat in der dialogischen Form derselben, wie sie bei Bolzano erscheint, seinen Grund. Er stellt immer Fragen, die nur mit „ja" oder „nein" beantwortet werden können, aber er untersucht nicht vorher, ob seine Fragen überhaupt so gestellt sind, dass sie der Beantwortung fähig sind. Sind sie dies, so ist der Beweis richtig, aber unnötig ; im entgegengesetzten Fall ist die ganz Deduction falsch.

90. [Kronecker, 1891a] pp. 9-10 : « C'est pourquoi sont aussi trompeuses les déductions de Bolzano, un ecclésiastique catholique de Prague, qui veut démontrer pour une fonction continue quelconque, que si elle a des signes différents dans son domaine de continuité, elle doit passer une fois par zéro. »

91. Voir p. 85.

énoncé et démontré par Bolzano : on connaît certaines valeurs de cette fonction qui nous permettent d'affirmer que la fonction est positive pour $z = 0$ et négative pour $z = \frac{1}{4}$. La méthode que Bolzano emploie à partir de là est une méthode par dichotomie. Mais, fait remarquer Kronecker, nous n'avons aucun moyen de déterminer $F(\frac{1}{8})$, et donc de connaitre son signe. La méthode générale présente dans la démonstration de Bolzano n'est donc pas effective dans le cas particulier exhibé par Kronecker. La question posée – c'est-à-dire $F(\frac{1}{8})$ est-il négatif – n'obtient pas ici comme réponse un « oui » ou un « non ». Kronecker propose la méthode alternative que nous venons de voir [92] : dans le cas des fonctions polynômiales à coefficients rationnels, il est tout à fait possible de déterminer un intervalle à l'intérieur duquel se trouve au plus une racine.

On retrouve déjà un témoignage de cette opposition aux conclusions de Bolzano dans une lettre que Schwarz a écrite à Cantor en juin 1870 [93] :

> Prof. Kronecker explique dans la lettre qu'il m'a adressée le (3./6. 70) que les conclusions de Bolzano sont évidemment de faux arguments ; il dit que mon beau-père [94], Borchardt et Heine se trouvent de son côté. Je suis content que toi, Thome et moi, soyons dans le camp de Weierstrass. [Meschkowski et Cantor, 1967] p. 68

Cette partition, forcément simpliste, souligne toutefois que l'héritage de Bolzano est plus à chercher dans le travail de Weierstrass [95] que dans celui de Kronecker. Nous allons essayer de voir si une étude un peu plus poussée du contexte d'écriture de la *Rein analytischer Beweis* et des soubassements philosophiques de Bolzano permet d'avancer dans l'analyse de cette opposition.

#### 2.2.3.1 Bolzano et la *Rein analytischer Beweis*

Comme beaucoup de grands penseurs du XVIII^e^ siècle, Bolzano possède un savoir encyclopédique : ses champs d'étude s'étendent ainsi des mathématiques à la philosophie en passant par la physique, la logique et la théologie. Ce grand scientifique tchèque de langue allemande est sans doute, comme l'écrit Jocelyn Benoist, le « dernier exemple dans l'histoire de la pensée d'une conjonction que l'évolution des sciences devait bientôt rendre définitivement improbable, à savoir celle d'une pratique scientifique créative de premier plan et d'une œuvre de grand philosophe » [96].

En 1805, ses études terminées et depuis peu ordonné prêtre, son « esprit universel » le pousse à briguer deux chaires à l'université de Prague : celle de mathématiques et celle

92. Voir p. 85.

93. Herr Prof. Kronecker erklärte in seinem an mich gerichteten Briefe (3./6. 70) die Bolzanoschen Schlüsse als offenbare Trugschlüsse ; er sagte, mein Schwiegervater, Borchardt und Heine befinden sich auf seiner Seite. Es ist mir lieb, daß Du, Thome und ich auf Weierstraß' Seite stehen.

94. Il s'agit de Kummer.

95. Nous pouvons, à ce propos, reprendre une lettre de Schwarz citée par p. Dugac dans [Dugac, 1976] p. 77 : « sans les conclusions, qui ont été développées par Weierstrass à partir des principes de Bolzano, on n'aurait pas pu réussir dans de nombreuses recherches », qui montre que dès 1870 Weierstrass connaissait le travail de Bolzano et que ce dernier était l'une des sources de son travail de refondation de l'analyse.

96. Jocelyn Benoist, *Bolzano et l'idée de Wissenschaftslehre* dans [Pierre, 2002] p. 659.

de science de la religion. C'est finalement cette dernière qu'il obtiendra et dont il sera destitué en 1819. Ainsi, interdit d'enseignement et de publication à l'âge de trente-huit ans, il se retire pour se consacrer à l'élaboration de ses deux grands systèmes : sa logique dans la *Wissenschaftslehre* et sa mathématique dans la *Grossenlehre*.

Très tôt dans sa carrière, il publie une méthode qu'il intitulera *Contributions pour un exposé mieux fondé des mathématiques*, dont la deuxième partie, *Sur la méthode mathématique* est l'exposé de la « bonne méthode » que doit suivre, selon lui, le scientifique désireux de produire de la « science véritable ». Quelques années plus tard, il fera suivre ses *Contributions* de trois essais d'analyse. Dans l'un de ces petits traités, il dit d'ailleurs son ambition d'utiliser lui aussi ses travaux en mathématiques pour illustrer la méthode qu'il a précédemment publiée. Ce mémoire, sans doute le plus marquant d'entre eux, est la *Rein analytischer Beweis*, publiée en 1817. Bolzano y démontre entre autres le théorème des valeurs intermédiaires et introduit pour la première fois notre définition moderne de la continuité pour une fonction [97]. De plus, comme le fait remarquer Steve Russ [98], dans ses deux traités d'analyse de 1816 et 1817, c'est-à-dire dans Le *binomische Lehrsatz* ([Bolzano, 1816]) et la *Rein analystisher Beweis*, Bolzano cite plus de vingt références sur des travaux majeurs, tous publiés dans les vingt années qui précèdent [99]. C'est donc en philosophe et en mathématicien accomplis qu'il s'attaque à une réflexion approfondie sur la méthode scientifique, et cherche à démontrer le théorème des valeurs intermédiaires.

Et c'est bien aussi ce mémoire que Kronecker critique, qu'il s'agisse de la méthode qui y est employée ou des théorèmes qui y sont énoncés. Le désaccord porte en particulier sur deux questions : la place de la philosophie dans les mathématiques, et celle de la géométrie.

#### 2.2.3.2 La place de la philosophie dans les mathématiques

L'addition d'une formation poussée en mathématiques et en philosophie amène Bolzano à s'interroger sur les fondements des mathématiques. Ce sont à la fois son opposition à Kant et ses lectures d'ouvrages mathématiques qui vont le convaincre de mettre en place un système d'exposition des sciences et à tenter de l'appliquer dans ses recherches.

Aristote voulait séparer strictement le rôle du mathématicien et celui du philosophe. Il affirme en effet que « en ce qui concerne les principes eux-mêmes, le géomètre, en tant que géomètre, ne doit pas en rendre raison » [100] et qu' « il est manifeste que l'examen de ces axiomes relève d'une seule science et que c'est celle du philosophe » [101]. Bolzano au contraire, dans les *Contributions*, fabrique sa propre méthode pour faire des mathématiques

97. À ce sujet, et en particulier sur la controverse de l'élaboration de cette notion arithmétique de la continuité, on pourra lire *Bolzano, Cauchy and the « new Analysis » of the Early Nineteenth Century* de I. Grattan-Guinness ([Grattan-Guinness, 1970]) ainsi que *Did Cauchy Plagiarize Bolzano ?* de H. Freudenthal ([Freudenthal, 1971]).
98. [Russ, 2004] p. 142.
99. Entre autre les travaux de : Lagrange, Lacroix, Gauss, Crelle, Newton, Euler, D'Alembert.
100. [Tricot, 1979] I.12, 77 b 5.
101. [Duminil et Jaulin, 2008] Γ, 3, 1005a.

et tente dans ses mémoires de sciences « pures » de mettre en pratique cette méthode.

Ces *Contributions* ne sont pas à proprement parler une œuvre de mathématique, mais plutôt une réflexion sur les mathématiques. Et comme le montre clairement le titre de leur appendice [102], elles sont à comprendre en partie comme une réponse à Kant, et en particulier à son esthétique transcendantale. Dans une note ajoutée à son autobiographie, on apprend que Bolzano a passé deux années, de dix-huit à vingt ans, à travailler sur la *Critique de la raison pure*. C'est donc une pensée qui l'a profondément marqué, et il fut d'ailleurs, à tort, accusé de « kantisme » lorsqu'il était jeune professeur à l'Université de Prague. Cependant, il semble bien que si Bolzano s'est formé à la lecture de la Critique, c'est plutôt contre elle qu'avec elle [103].

Il commence, dans l'appendice des *Contributions*, par rendre grâce à Kant pour sa distinction entre jugement analytique et synthétique, a priori et a posteriori. Il reviendra d'ailleurs dans ses travaux sur ces distinctions et en précisera les frontières. Il accepte aussi tout en la redéfinissant la partition des représentations en concepts et intuitions.

Ce qui est central dans la critique par Bolzano de l'Esthétique transcendantale, tout au moins dans l'appendice des Contributions, c'est sa réfutation de la possibilité d'existence d'intuitions a priori. En s'attaquant à ces intuitions, c'est en fait au fondement de la réponse qu'apporte Kant à la question : « Comment les mathématiques sont-elles possibles ? » que Bolzano s'attaque : sans intuition a priori, pas de jugement synthétique a priori, et donc pas de mathématiques [104].

Pour Kronecker, la relation que doit entretenir la philosophie avec les mathématiques est tout autre. Dans une lettre datée du 18 août 1884, Kronecker écrit à Cantor [105] :

102. « Sur la doctrine kantienne de la construction des concepts par les intuitions ».

103. On pourra lire à ce sujet [Laz, 1993].

104. Les attaques de Bolzano sont souvent violentes et rarement dénuées d'ironie : « Parce que nous ajoutons mes unités du nombre 5 au nombre 7 l'une après l'autre, c'est donc – conclut-il – l'intuition du temps qui fonde ces raisonnements. Quelle déduction ! », Wissenschaftslehre, III, § 305. ([Bolzano, 1973], notre traduction de l'anglais).

105. Da Sie vor mehr als 20 Jahren selber noch meine Vorlesungen gehört und auch seitdem, in fast ununterbrochenen Beziehungen zu mir stehend, oft genug meine Ansichten vernommen haben, so wissen Sie besser, als ich es Ihnen jetzt auseinanderzusetzen vermöchte, dass ich – sehr früh unter Kummers Anleitung in philosophische Studien vertieft – nachher gleich ihm, die Unsicherheit aller jener Spekulationen erkannt und mich in den sicheren Hafen der wirklichen Mathematik geflüchtet habe. Was natürlicher, als dass ich in dieser Mathematik selbst nun mich bemüht habe, ihre Erscheinungen oder ihre Wahrheiten möglichst frei von jeden philosophischen Begriffsbildungen zu erkennen. Ich bin deshalb darauf ausgegangen, Alles in der *reinen* Mathematik auf die Lehre von den ganzen Zahlen zurückzuführen, und ich *glaube*, dass dies durchweg gelingen wird. Indessen ist dies eben nur mein *Glaube*. Aber wo es gelungen ist, sehe ich darin einen wahren Fortschritt, obwohl – oder weil – es ein Rückschritt zum Einfachsten ist, noch mehr aber deshalb, weil es denn beweist, dass die neuen Begriffsbildungen wenigstens nicht *nothwendig* sind. Ich werde das Wesentlichste meiner Ansichten ja nächstens einmal im Druck bekannt zu geben haben und dabei noch meine Einwendungen gegen jene Stolz'sche Deduction – die Sie ja aus meinen mündlichen Mittheilungen kennen – formuliren. Dann mögen diese Dinge publice sine ira et studio erörtert werden ? Was aber in aller Welt soll eine solche Erörterung unsrer persönlichen Beziehung schaden ? Dass ich jene Einwendungen nur gelegentlich machen will, beruht darauf, dass ich denselben nur einen höchst secundären Werth beilege. Einen wahren wissenschaftlichen Werth erkenne ich – auf dem Felde der Mathematik – nur in concreten mathematischen Wahrheiten, oder schärfer ausgedrückt, „nur in mathematischen Formeln". Diese allein sind, wie die Geschichte der Mathematik zeigt, das Unvergängliche. Die verschiedenen Theorien für die Grundlagen der Mathematik (so die von Lagrange) sind von der Zeit weggeweht, aber die Lagrangesche Resolvente ist geblieben !

> Comme vous avez suivi mes cours il y a plus de 20 ans et que vous êtes depuis lors en contact avec moi de façon presque ininterrompue, vous avez entendu mes vues assez souvent pour comprendre mieux que je ne saurais l'exprimer ici que – ayant approfondi très tôt mes études de philosophie sous les conseils de Kummer – je reconnais, comme il le faisait, l'incertitude de telles spéculations, et je me réfugie dans le paradis des vraies mathématiques. Quoi de plus naturel alors que je me sois efforcé dans mon travail mathématique d'exprimer ses phénomènes (*Erscheinungen*) et ses vérités (*Wahrheiten*) dans une forme aussi indépendante que possible de concepts philosophiques. J'ai donc choisi de baser toutes les mathématiques pures sur la théorie des nombres entiers et je *pense* que cela peut être fait sans exception. Toutefois, il s'agit seulement de ma *croyance*. Mais partout où cela a fonctionné, j'y ai vu un vrai progrès, même si – ou parce que – c'est une régression aux principes les plus simples, d'autant plus que cela prouve que les nouveaux concepts qui sont introduits sont pour le moins *superflus*. Je pense publier l'essentiel de mes vues très bientôt, et par la même occasion formuler mes objections aux déductions de Stoltz, que vous connaissez déjà par mes communications orales. Alors toutes ces choses pourront être discutées publice sine ira et studio. Mais pourquoi diable une telle discussion devrait porter préjudice à nos relations personnelles ?
>
> Que je ne veuille exprimer mes objections seulement en passant provient du fait que je leur attribue au plus une valeur secondaire. Je reconnais la vraie valeur scientifique – dans le champ des mathématiques – seulement dans les vérités mathématiques concrètes, ou, pour l'exprimer plus nettement, seulement dans les formules mathématiques. Elles seules, comme l'histoire des mathématiques le montre, sont impérissables. Les diverses théories sur les fondations des mathématiques (comme celle de Lagrange) sont depuis longtemps tombées dans l'oubli, alors que la résolvante de Lagrange reste ! [Cantor *et al.*, 1991] p. 196

Kronecker exprime ici sa volonté de ne pas faire intervenir la philosophie dans les mathématiques, ce qui correspond d'ailleurs à la position qu'il tiendra dans l'article annoncé [106]. Mais finalement, cette affirmation est déjà en soi un positionnement philosophique. De plus, son idée du concept de nombre et plus largement son engagement dans la création d'une « arithmétique générale » a des incidences sur sa façon de faire des mathématiques, et nous en verrons quelques exemples dans les parties suivantes. Il commence d'ailleurs son article sur le concept de nombre par les phrases suivantes :

> C'est sur le plan libre d'un travail philosophique préparatoire, à partir duquel on entre dans les domaines clos des différentes sciences, qu'il faut aussi développer les concepts de nombre, d'espace et de temps, dont on fait usage dans les mathématiques. Et il paraît approprié d'y conduire le développement de ces concepts aussi loin qu'il faudra pour qu'ils soient déjà dotés de leurs propriétés fondamentales lorsque commence leur traitement spécifiquement scientifique. [Boniface, 1999] p. 53

Les concepts de nombre, comme ceux d'espace et de temps ne sont pas discutés à l'intérieur des mathématiques, mais seulement dans un travail philosophique qui le précède : contrairement à Bolzano, Kronecker fait sien les propos d'Aristote.

106. Il s'agit sans doute de l'article *Über den Zahlbegriff* publié trois ans plus tard en 1887 ([Kronecker, 1887d]).

#### 2.2.3.3 Kant, Bolzano et Kronecker

Kronecker a reçu, comme Bolzano, une formation avancée en philosophie, et on peut dire que lui aussi s'oppose à Kant lorsqu'il voit, comme Gauss [107], dans les nombres entiers une pure création de l'esprit. Cependant, rappelons-nous que Kant caractérise les mathématiques par rapport à la philosophie dans sa *Théorie transcendantale de la méthode* de la façon suivante :

> la connaissance *philosophique* est la *connaissance rationnelle* par *concepts* et la connaissance *mathématique* la connaissance rationnelle par *construction* de concepts. Or, *construire* un concept, c'est représenter *a priori* l'intuition qui lui correspond. La construction d'un concept exige donc une intuition *non empirique*, qui par conséquent comme intuition est un objet *singulier*, mais qui n'en est pas moins, comme construction d'un concept (d'une représentation générale) exprimer dans la représentation une validité universelle, pour toutes les intuitions possibles qui appartiennent au même concept. [Kant, 1980] p.604

Si on retrouve ici une différence fondamentale entre la conception kantienne des mathématiques et celle de Bolzano, on s'aperçoit par contre que ce point de vue sur les mathématiques rejoint en partie celui de Kronecker. En effet, pour Bolzano, les objets, les propositions ou les représentations mathématiques sont en soi, c'est-à-dire indépendants de toute subjectivité. Au contraire pour Kant les mathématiques sont une œuvre humaine de construction, et il en est de même pour Kronecker, à l'exception notable des nombres entiers. Il semble bien pour Kant que ce soit l'intuition qui guide le mathématicien, et que le passage à l'universalité ne puisse se faire que par le vecteur d'une intuition pure. Une conséquence de cela porte sur les conditions d'existence des objets mathématiques. Alors que pour Bolzano « il suffit de montrer la possibilité de chaque combinaison de concepts que l'on introduit » , c'est-à-dire simplement de prouver leur non-contradiction, il faut pour Kronecker, comme pour Kant, une démonstration de leur existence. Chez ce dernier on est finalement très proche de la géométrie d'Euclide où tout objet n'existe que parce que sa construction, au sens propre du terme, a été effectuée, et c'est cette idée que Kronecker prolonge.

Nous sommes donc face à une opposition profonde entre les conceptions des mathématiques de Bolzano et de Kronecker, conceptions qui chez ces deux mathématiciens ont été nourries pourtant par leur lecture de l'œuvre de Kant. Nous allons maintenant regarder comment cela se fixe sur la place que chacun d'entre eux attribue à la géométrie.

#### 2.2.3.4 La place de la géométrie

Dans *Über der Zahlbegriff*, Kronecker exprime son point de vue sur quelques notions, et, se plaçant en dehors des mathématiques, il en esquisse le schéma suivant :

> En réalité, l'arithmétique est aux deux autres disciplines mathématiques, géométrie et mécanique, dans un rapport semblable à celui de l'ensemble des mathématiques

107. Voir la citation de Gauss p. 97.

> à l'astronomie et aux autres sciences de la nature ; l'arithmétique rend aussi tout un éventail de services à la géométrie et à la mécanique et reçoit en retour de ses disciplines-sœurs une foule d'inspirations. Le mot « arithmétique » ne doit cependant pas être compris dans son sens usuel limité, mais on doit comprendre sous ce terme toutes les disciplines mathématiques à l'exception de la géométrie et de la mécanique, donc nommément, l'algèbre et l'analyse. Et je crois aussi que l'on parviendra un jour à « arithmétiser » le contenu entier de ces disciplines mathématiques, c'est-à-dire à le fonder purement et simplement sur le concept de nombre pris dans son sens le plus étroit et donc à dépouiller à nouveau ce concept des modifications et élargissements le plus souvent provoqués par les applications à la géométrie et à la mécanique. La différence principielle entre la géométrie et la mécanique d'une part et d'autre part l'ensemble des autres disciplines mathématiques rassemblées ici sous la dénomination d'arithmétique, tient selon Gauss en ce que l'objet de la dernière, le nombre, est une production de notre esprit seul, alors que l'espace, de la même manière que le temps, a aussi une réalité en dehors de notre esprit, réalité à laquelle nous ne pouvons a priori dicter entièrement ses lois [108]. [Boniface, 1999] p. 54-55

Kronecker place donc la géométrie *à côté* de l'arithmétique, et finalement à côté des « mathématiques pures ». Il n'y a pas d'inclusion de l'un dans l'autre, et lorsqu'il rêve que « l'on parviendra un jour à "arithmétiser" le contenu entier de ces disciplines mathématiques », la géométrie n'en fait pas partie. Il n'y a pas de rapport de dépendance, mais plutôt des échanges : la géométrie utilise l'arithmétique comme outil, et l'arithmétique utilise la géométrie comme un moyen heuristique dans son élaboration [109].

Pour Bolzano, « l'arithmétique est encore de loin la discipline mathématique la plus parfaite, tandis que la géométrie a des défauts bien plus importants et bien plus difficiles à résoudre » [110]. C'est ainsi que la première publication d'un de ses mémoires de mathématiques, en 1804, porte sur les fondements de la géométrie. Il s'agit de ses *Considérations sur quelques objets de géométrie élémentaire* [111], composées de deux parties et d'une préface. La première partie consiste en une étude sur les triangles et les parallèles où Bolzano tente entre autres de résoudre ce que Jean Le Rond d'Alembert a appelé le « scandale de la géométrie », c'est à dire le cinquième postulat d'Euclide. La seconde partie, quant à elle, traite de la théorie de la ligne droite. Remarquons qu'ici encore le but de Bolzano n'est pas la « découverte » de nouvelles propriétés, mais bien de repenser les fondations des mathématiques en général et celles de la géométrie en particulier [112]. Il construit dans ces

108. J. Boniface rappelle la citation de Gauss dont il est question dans ce passage :
« Selon ma plus intime conviction la théorie de l'espace a une position par rapport à notre savoir a priori tout à fait différente de la théorie pure des grandeurs ; il manque tout à fait à notre connaissance de la première *cette* entière conviction de sa nécessité (donc aussi de son absolue vérité) qui est propre à la *dernière* ; nous devons avouer en toute humilité que si le nombre est une production de notre esprit *seul*, l'espace a aussi *en dehors* de notre esprit une *réalité* à laquelle on ne peut a priori entièrement dicter ses lois. »

109. Ce dernier point sera mis en avant dans les deux parties suivantes à l'occasion de notre étude du théorème de Sturm et de la théorie des caractéristiques.

110. Préface des *Contributions*, p 87.

111. [Russ, 2004] p. 24 : « Betrachtungen über einige Gegenstande der Elementargeometrie ».

112. Rappelons ce que Bolzano dit des *Éléments* d'Euclide dans son autobiographie :

> I should like to relate how the famous mathematician Euclid once became my doctor.

*Contributions* une classification des mathématiques dans laquelle la géométrie est exclue des mathématiques pures, mais sera, en tant que « mathématiques appliquées » dans une forte relation de dépendance vis-à-vis des mathématiques pures qui sont composées de l'arithmétique, la combinatoire, l'analyse et l'algèbre.

Le passage de la *Critique de la raison pure* précédemment cité peut expliquer en grande partie la place que Kant réserve à la géométrie : c'est pour lui, contrairement à Bolzano ou à Kronecker, une branche des mathématiques pures. Le passage d'une figure singulière à une propriété universelle est bien plus convaincant pour illustrer sa théorie que tout exemple pouvant être issu de l'arithmétique ou de l'algèbre, même s'il rappelle que dans cette dernière, les symboles peuvent remplacer les figures. Mais si Kronecker et Bolzano s'opposent tous deux à la place que réserve Kant à la géométrie, le statut de cette dernière n'est pas la même, comme nous venons de le voir, chez ces deux mathématiciens.

C'est une opposition profonde qui existe finalement entre Bolzano et Kronecker, et le refus de ce dernier d'accepter le théorème des valeurs intermédiaires tel qu'il est élaboré dans la *Rein analystischer Beweis* a son fondement non seulement dans leurs visions extrêmement différentes des mathématiques, mais aussi dans leurs pratiques : si le premier travaille essentiellement, seul, sur les fondements des mathématiques, le second produit des résultats nouveaux dans le cadre de la recherche institutionnelle et internationale.

---

It was during a holiday period which as usual, I was spending in Prague, when I was struck by an illness I had never had before that caused a shivering and shaking of all my limbs. In order to keep up my spirits as much as possible I took the Elements of Euclid and read the theory of proportion which I found treated here in a way completely new to me. The ingenuity of the Euclidean presentation gave me such vivid pleasure that I soon felt better again. (Traduction de S. Russ dans [Russ, 2004] p. 15)

Cependant, des différences fondamentales apparaissent entre Bolzano et Euclide, principalement sur l'existence des objets mathématiques.

# Chapitre 3

# Présentation et commentaire du cours de Kronecker de l'hiver 1880-1881 sur la théorie des équations algébriques


3.1 Équations quadratiques et cubiques . . . . . **103**
3.1.1 Équations quadratiques . . . . . 104
3.1.2 Équation cubique . . . . . 105
3.1.3 Conclusion . . . . . 114
3.2 Proposition fondamentale des équations générales et théorie des fonctions symétriques et alternées . . . . . **115**
3.2.1 Quelques définitions et premières propriétés . . . . . 116
3.2.2 Fonctions symétriques . . . . . 122
3.3 Sur la divisibilité des fonctions entières . . . . . **144**
3.3.1 Résultant et Discriminant . . . . . 144
3.3.2 Décomposition d'une fonction entière en facteurs . . . . . 147
3.4 Conclusion . . . . . **149**


Il serait illusoire de penser pouvoir donner ici une idée précise du contenu de l'ensemble des cours sur la théorie des équations algébriques de Kronecker. Les deux parties suivantes de ce travail seront d'ailleurs centrées autour de l'étude de deux chapitres particuliers de ces *Vorlesungen* : celui sur le théorème de Sturm et celui sur la théorie des caractéristiques. Cependant, nous aurons tout de même besoin de savoir ce que Kronecker expose avant de présenter ces deux notions, et pour cela, nous allons consacrer ce chapitre à la lecture détaillée de la première moitié des manuscrits des semestres d'hiver 1878-79 et 1880-81.

Commençons par exposer les raisons qui nous ont conduit à choisir ces manuscrits. Tout d'abord, il s'agit des deux derniers qui précèdent la publication des *Grundzüge einer arithmetischen Theorie der algebraischen Grössen*, publication qui représente un tournant

dans l'œuvre de Kronecker. Comme nous le verrons, c'est dans cette période qui précède les *Grundzüge* que se développe la théorie des caractéristiques comme prolongement du théorème de Sturm, et il était donc nécessaire de choisir des manuscrits datant de ces années pour exposer les fondations sur lesquelles Kronecker construit sa théorie. Ensuite, ces deux manuscrits suivent un plan presque identique. Il y a très peu d'évolution entre ces deux cours. Des changements fondamentaux apparaissent par contre dans celui de 1884-85 (nous ne possédons pas d'exemplaire de l'hiver 1882-83, et nous ne sommes même pas sûr, bien que cela soit fort probable, que Kronecker ait professé son cours sur les équations algébrique cette année-là). En ce sens ces deux manuscrits forment un tout. Rappelons leur structure générale :

**Le Manuscrit du semestre d'hiver 1878-79 [Kronecker, 1879]**

1. 1 Introduction historique
2. 2 Équations quadratiques et cubiques.
3. 25 Équation du degré $n$ et fonctions symétriques
4. 296 Théorème de Sturm / Principe d'inertie de Sylvester
5. 482 Théorie des caractéristiques
6. 566 Caractéristique de $n+1$ fonctions de $n$ variables.
7. 622 Développement autour de la théorie de l'élimination
8. 701-787 Théorie des grandeurs algébriques (*Rationalitätsbereich*, *Gattung*, Résolvante de Galois,...)

**Le Manuscrit du semestre d'hiver 1880-81 [Kronecker, 1881a]**

1. 2 Introduction historique
2. 8 Équations quadratiques et cubiques
3. 17 Équation du degré $n$ et fonctions symétriques
4. 104 Théorème de Sturm / Principe d'inertie de Sylvester
5. 150-170 Théorie des caractéristiques

Le manuscrit de 1880-81 correspond donc aux 565 premières pages de celui de 1878-79, et il semble qu'il ait existé un second tome qui le complète [1]. Si la structure des deux manuscrits est identique, celui de 1880-81 est beaucoup plus synthétique, et celui de 1878-79 plus développé. Un aller-retour entre les deux nous permettra donc à la fois de facilement dégager la structure tout en nous gardant la possibilité, lorsque cela sera nécessaire, d'obtenir plus de détails. Par ailleurs, nous ne nous interdirons pas de regarder de temps à autre dans d'autres manuscrits, mais à chaque fois, nous le préciserons. Notre présentation suivra au plus près la structure indiquée dans le manuscrit de 1880-81. Celui-ci est découpé en deux chapitres : le premier s'intitule

*Proposition fondamentale des équations générales ; théorie des fonctions symétriques et alternées*

et le second

1. Voir note p. 53.

*Sur la divisibilité des fonctions entières ; Déduction et proposition de Sturm ; Conclusions tirées de celles-ci ; théorie des caractéristiques et preuve de l'existence des racines*

Nous nous proposons ici de rendre compte du contenu du premier chapitre, ainsi que de la première partie du second qui traite de la divisibilité des fonctions entières. L'exposition du théorème de Sturm et celle de la théorie des caractéristiques seront l'objet respectivement de la seconde et troisième partie de notre travail.

Nous ne chercherons pas à détailler l'ensemble du contenu, même si nous souhaitons que tout ce que Kronecker aborde soit présent dans notre travail, ne serait-ce, pour certaines parties, que sous la forme d'un résumé rapide. Ainsi, si certaines notions seront simplement mentionnées, d'autres au contraire seront particulièrement approfondies. Notre but est, outre la présentation générale des *Vorlesungen über die Theorie der algebraischen Gleichungen*, de nous préparer à l'étude de la théorie des caractéristiques : c'est à partir de ce critère que nous choisirons de développer quelques notions de ce premier chapitre. Cependant, si un passage nous semble intéressant pour progresser dans la compréhension du point de vue de Kronecker sur les mathématiques, nous nous y arrêterons, même si son contenu mathématique est indépendant de la suite de notre exposé. Enfin, nous noterons le passage d'un paragraphe à l'autre en indiquant entre parenthèse le symbole **§** suivi de son numéro.

## 3.1 Équations quadratiques et cubiques

Après quelques rappels historiques, Kronecker commence par un chapitre qu'il intitule *Perspectives sur la théorie des équations algébriques* (**§ 2**), et dans lequel il donne quelques exemples et propriétés de la théorie des équations quadratiques et cubiques. Il considère que [2] :

> Pour montrer quels sont les points de vue à prendre en compte dans la théorie des équations algébriques, nous voulons continuer avec une présentation rapide de la théorie des équations quadratiques et cubiques. [Kronecker, 1881a] p. 6

Et ce chapitre sert donc, dans le cours de 1880-81, d'« introduction mathématique » à la suite de son exposé. Dans ce cours du semestre 1880-81, Kronecker ne s'étend pas autant sur les équations cubiques qu'il le fera dix ans plus tard lors de son denier cours sur la théorie des équations algébriques. Cependant, nous allons voir tout de même comment Kronecker utilise l'étude de ces équations de petit degré pour illustrer des concepts très généraux.

2. Um zu zeigen, welche Gesichtspunkte in der Theorie der Algebraischen Gleichungen ins Auge zu fassen sind, wollen wir mit wenigen Worten auf die Theorie der quadratischen und kubischen Gleichungen zu sprechen kommen.

### 3.1.1 Équations quadratiques

Kronecker aborde l'étude des équations du second degré en termes d'*aequatio pura* et d'*aequatio affecta*. Cette distinction, Kronecker l'attribue à François Viète [3] :

> L'apparition notable suivante est celle de *François Viète* (*Vieta*) de 1540 à 1603. Il était avocat à *Poitiers* et a accompagné plus tard *Henri IV* à *Paris*. Il a été le premier à introduire le calcul littéral et à en développer les conséquences. Ainsi désigna-t-il par exemple les grandeurs connues par des consonnes et les inconnues par les voyelles. Les expressions "*potestas pura*" (puissance) et "*potestas affecta*" (fonction entière) viennent de lui, expressions qui seront conservées jusqu'à *Gauss*. C'est *Viète* qui a reconnu comme fondement de la théorie des fonctions entières la décomposition en facteurs linéaires. Par contre il évitait encore les racines négatives, et il ignorait encore le signe de la multiplication et de l'égalité. [Kronecker, 1881a] pp. 2-3

Et lorsque Kronecker utilise ces termes, c'est aussi à Gauss qu'il se réfère, car d'après lui ce dernier « désigne comme la nature de la résolution des équations la *Reductio aequationum affectarum ad aequationes puras* ». Dans la septième section des *Disquisitiones Arithmeticae* [4], il est en effet question de cette « réduction à des équations pures » (*affectarum reductionem ad puras* [Gauss, 1801] art. 359).

Kronecker reprend en fait ici, pour la résolution de l'équation du second degré, exactement l'exposé que donne Vandermonde dans [Vandermonde, 1771], et suit ainsi la méthode proposée par ce dernier pour le cas général [5] :

> 1. *Trouver une fonction des racines, de laquelle on puisse dire, dans un certain sens, qu'elle égale telle de ces racines que l'on voudra.*
> 2. *Mettre cette fonction sous une forme telle qu'il soit de plus indifférent d'y échanger les racines entr'elles.*
> 3. *Y substituer les valeurs en somme de ces racines, somme de leurs produits deux-à-deux, etc. [Vandermonde, 1771] p. 370*

Pour cela, il pose

$$x^2 - f_1 x + f_2 = 0$$

donc

$$4(x^2 - f_1 x + f_2) = 0$$

et par *Reductio ad aequationem puram*, c'est-à-dire ici ce que l'on appelle maintenant la forme canonique de l'équation, on obtient :

3. Die nächste bedeutende Erscheinung ist der Franzose *François Viète* (*Vieta*) von 1540-1603. Er war Advokat in *Poitou* und begleitete später *Henri IV.* nach *Paris*. Er war der erste, der die Buchstabenrechnung einführte und mit Consequenz ausbildete. So bezeichnete er z.B. die bekannten Größen durch Consonanten, die Unbekannten durch Vokale. Von ihm rühren die Ausdrücke „*potestas pura*" (Potenz) und „*potestas affecta*" (ganze Funktion) her, Ausdrücke, die sich bis zu *Gauss* erhielten. *Viète* war es auch, der als das Fundament der Theorie der ganzen Funktionen die Zerlegung derselben in Linearfaktoren erkannte. Dagegen vermied er noch die negativen Wurzeln und war noch unbekannt mit den Zeichen der Multiplikation und der Gleichheit.

4. Deuxième partie, articles 359 à 364.

5. Cas général qui permet bien sûr de ne résoudre une équation que dans certains cas très particuliers.

$$(2x - f_1)^2 + (4f_2 - f_1^2) = 0$$

Si $x_1$ et $x_2$ sont les solutions de l'équation, alors on a :

$$x^2 - f_1 x + f_2 = (x - x_1)(x - x_2)$$

donc : $f_1 = x_1 + x_2$ et $f_2 = x_1 x_2$.

D'après ce qui précède, on a :

$$\begin{aligned} 2x &= f_1 \pm \sqrt{f_1^2 - 4f_2} \\ &= x_1 + x_2 \pm \sqrt{(x_1 - x_2)^2} \\ &= x_1 + x_2 \pm (x_1 - x_2) \\ &= \begin{cases} 2x_1 \\ 2x_2 \end{cases} \end{aligned}$$

Ainsi, **l'idée générale** de la résolution de l'équation quadratique est la suivante : on se ramène à une équation de la forme $x^2 = A$, avec $A$ qui ne contient que des fonctions symétriques des racines de l'équation. Ce rôle d'auxiliaire que jouent ici les fonctions symétriques s'amplifiera par la suite. Il fait ensuite de même avec les équations cubiques, et retrouve ainsi les formules de Cardan.

### 3.1.2 Équation cubique

Après avoir suivi la démarche de Vandermonde pour l'équation quadratique, Kronecker expose maintenant la résolution des équations cubiques telle que Lagrange la propose dans ses *Réflexions sur la résolution algébrique des équations* ([Lagrange, 1771]). Par rapport à ce dernier, on peut remarquer que Kronecker évite autant qu'il le peut les nombres rationnels, auxquels il préfère les nombres entiers. Pour cela, il est amené à choisir certaines formes pour ses coefficients (par exemple le $3x$ dans la substitution au lieu de $x$).

Soit

$$x^3 - f_1 x^2 + f_2 x - f_3 = 0$$

Pour se ramener au cas traité par Lagrange, Kronecker commence par faire disparaître le second terme avec la substitution, par la méthode de Tschirnhaus [6] :

$y = 3x - f_1$, puis vient
$y^3 - 3\varphi y - \psi = 0$, où

6. [Kronecker, 1885b] p. 19 : « die Tschirnhausen'sche Methode ».

$$\varphi = f_1^2 - 3f_2 \ , \ \psi = 2f_1^3 - 9f_1f_2 + 27f_3$$

car on a en effet :

$$y^3 - 3\varphi y - \psi = 27(x^3 - f_1x^2 + f_2x - f_3)$$

$$\boxed{\text{On a donc maintenant à résoudre } y^3 = \psi + 3\varphi y}$$

Kronecker utilise ensuite une méthode que Lagrange attribue à Johan Hudde [7] « qui consiste à représenter la racine par la somme de deux indéterminées qui permettent de partager l'équation en deux parties propres à faire en sorte que les deux indéterminées ne dépendent que d'une équation résoluble à la manière de celles du second degré. » [8]. On pose donc :

$$y = z + t$$

et on obtient

$$y^3 = (z+t)^3 = z^3 + t^3 + 3zt(z+t)$$

et par identification on a $\psi = z^3 + t^3$ et $\varphi = zt$, et donc $t = \frac{\varphi}{z}$.

On a donc

$$\psi = z^3 + \frac{\varphi^3}{z^3}$$

Comme le fait remarquer Lagrange, cette équation [9] est de degré 6.

Mais $z^3$ est ainsi déterminé par une équation quadratique ( car $\varphi$ et $\psi$ ne dépendent que des coefficients de l'équation de départ).

On a ainsi

$$z^3 = \frac{1}{2}\psi \pm \frac{1}{2}\sqrt{\psi^2 - 4\varphi^3}$$

Pour pouvoir poursuivre, Kronecker a besoin de la notion de fonction multivalente.

**Fonctions multivalentes** Kronecker utilise les termes de fonction bivalente (*doppeldeutig*), univalente (*eindeutig*) ou multivalente (*mehrdeutig*). $\sqrt{A}$ signifie donc à la fois la racine positive et la racine négative de $A$. Ces termes sont utilisés couramment [10] à partir du XVIII$^{\text{e}}$ siècle : pour les fonctions racines, lors de la construction du théorème des fonctions implicites comme Renaud Chorlay le montre dans [Chorlay, 2003] et surtout lors de la construction des logarithmes complexes. Les fonctions multivalentes [11] apparaissent

7. Voir [Lagrange, 1771] p. 207. Johan Hudde (1628-1704) est un mathématicien hollandais. Lagrange fait référence ici à [Hudde, 1661].

8. Voir [Lagrange, 1771] p. 207.

9. Dans [Lagrange, 1771], il s'agit de l'équation en $y$ :

$$y3 - \frac{n^3}{27y^3} + p = 0.$$

10. Voir par exemple l'utilisation qu'en fait Weierstrass dans [Nabonnand, 1999] p. 56.

11. Ou, selon les auteurs et les traduction : multivoque / multiforme.

naturellement, comme l'affirme Euler, lors du passage du logarithme réel au logarithme complexe, et ce dernier n'est compréhensible que par l'existence de ces fonctions :

> Lorsqu'on dit que le logarithme d'un nombre proposé est l'exposant de la puissance d'un certain nombre pris à volonté, laquelle devient égale au nombre proposé, il semble qu'il ne manque rien à la justesse de cette idée. Cela est aussi bien vrai ; mais on accompagne communément cette idée d'une circonstance qui ne lui convient point : c'est qu'on suppose ordinairement presque sans qu'on s'en aperçoive qu'à chaque nombre il ne répond qu'un seul logarithme & pour peu qu'on y réfléchisse on trouvera que toutes les difficultés & contradictions dont la doctrine des logarithmes sembloit embarassée, ne subsistent qu'entant qu'on suppose, qu'à chaque nombre ne répond qu'un seul logarithme. Je dis donc pour faire disparoitre toutes ces difficultés & contradictions, qu'en vertu même de la définition donnée il répond à chaque nombre une infinité de logarithmes [Académie de Berlin, 1751] pp. 155-156

D'après les relations coefficients-racines [12], la somme des deux racines de cette équation quadratique vaut $\psi$. Or $\psi = z^3 + \frac{\varphi^3}{z^3}$, donc si $z^3$ est la première racine, la seconde est $\frac{\varphi^3}{z^3}$.

Soit maintenant $\rho$ une racine primitive troisième de l'unité, par exemple $\rho = \frac{-1+\sqrt{-3}}{2}$. On a alors

$$\rho^3 = 1 \ , \ 1 + \rho + \rho^2 = 0$$

et les racines troisièmes de l'unité sont $1, \rho$ et $\rho^2$.

Soit maintenant $\chi$ tel que $\psi + \sqrt{\psi^2 - 4\varphi^3} = \psi + 3\chi\sqrt{-3}$. C'est-à-dire en fait $\chi = \sqrt{\frac{4\varphi^3 - \psi^2}{27}}$. Kronecker appellera $\chi^2$ le discriminant de l'équation du troisième degré, et cela correspond à la définition actuelle du discriminant [13].

Comme $\rho = \frac{-1+\sqrt{-3}}{2}$, on a

$$\sqrt{-3} = \rho + \rho + 1 = \rho - \rho^2 - 1 + 1 = \rho - \rho^2$$

et alors

$$z^3 = \frac{1}{2}\psi + \frac{1}{2}\sqrt{\psi^2 - 4\varphi^3} = \frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi$$

et

$$\frac{\varphi^3}{z^3} = \frac{1}{2}\psi + \frac{3}{2}(\rho^2 - \rho)\chi$$

On avait posé $y = 3x - f_1$, et d'autre part on a $y = z + t = z + \frac{\varphi}{z}$ et donc

$$3x = f_1 + z + \frac{\varphi}{z} = f_1 + \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi} + \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho^2 - \rho)\chi}$$

ce que l'on trouve ici, c'est bien la formule de Cardan : c'est une expres-

12. Ou *formules de Viète*, que l'on trouve d'ailleurs aussi dans le travaux d'Harriot ([Harriot, 1631]).
13. voir [Kronecker, 1887b] p. 43.

sion trivalente, car la racine cubique peut prendre 3 valeurs, et le fait que $\sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi}.\sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho^2 - \rho)\chi} = \varphi$ limite les possibilités à trois couples.

**La résolvante** On peut en fait développer l'équation du départ (*aequatio affecta*) en trois équations pures (*aequationes purae*) :

$$\begin{cases} \chi = \sqrt{\dfrac{4\varphi^3 - \psi^2}{27}} \\ z = \sqrt[3]{\dfrac{1}{2}\psi + \dfrac{3}{2}(\rho - \rho^2)\chi} \\ 3x = f_1 + z + \dfrac{\varphi}{z} \end{cases}$$

Kronecker remarque que l'on procède ici à l'inverse de l'élimination classique : on part d'une équation pour en obtenir trois[14]. A priori, $\chi$ peut prendre deux valeurs et pour chacune d'entre elles $z$ peut en prendre trois. On a ainsi six valeurs possibles pour $x$. Mais en fait, seules trois de ces six valeurs sont distinctes.

Kronecker donne la définition suivante : l'équation $z^3 + \frac{\varphi^3}{z^3} = \psi$ est appelée la *résolvante* de l'équation cubique initiale. Elle est de degré 6, et elle correspond à ce que Lagrange nomme la « *réduite* de l'équation initiale »[15]. Il poursuit en donnant une expression de certains coefficients en fonction des racines de l'équation initiale.

#### 3.1.2.1 Expression de $\varphi$, $\psi$ et $\chi$ en fonction des racines $x_0, x_1$ et $x_2$ de l'équation initiale

Le but de Kronecker est maintenant de montrer que l'on peut exprimer les diverses grandeurs qu'il utilise *rationnellement* en fonction des racines de son équation de départ.

Pour cela, Kronecker pose :

$$w_1 = x_0 + \rho x_1 + \rho^2 x_2 \ , \ w_2 = x_0 + \rho^2 x_1 + \rho x_2$$

et on a donc $w_1^3$ et $w_2^3$ qui sont les racines de la résolvante de l'équation initiale. On rappelle aussi que si $x_0, x_1$ et $x_2$ sont les racines de l'équation $x^3 - f_1 x^2 + f_2 x - f_3 = 0$, on a alors

$$x^3 - f_1 x^2 + f_2 x - f_3 = (x - x_0)(x - x_1)(x - x_2)$$

et

$$f_1 = x_0 + x_1 + x_2 \ , \ f_2 = x_0 x_1 + x_1 x_2 + x_2 x_0 \ , \ f_3 = x_0 x_1 x_2$$

**Expression de $\psi$** Comme $3x = y + f_1 = y + x_0 + x_1 + x_2$

14. [Kronecker, 1881a] p. 8 : « Wir schlagen also den umgekehrten Weg ein, wie bei der Elimination, wo man aus mehreren Gleichungen eine einzige Endgleichung bildet. » « Nous prenons donc le chemin inverse, comme avec l'élimination où l'on construit à partir de plusieurs équations une seule équation finale. »

15. Voir [Kronecker, 1881a] : « Die in $z^3$ quadratisch Gleichung hat von Lagrange den Namen : *Resolvente* der ursprunglischen kubischen Gleichung erhalten. » et [Lagrange, 1771] pp. 213 et 215.

On a :

$$\begin{aligned} y^3 - 3\varphi y - \psi &= 27(x^3 - f_1x^2 + f_2x - f_3) \\ &= 27(x - x_0)(x - x_1)(x - x_2) \\ &= (3x - 3x_0)(3x - 3x_1)(3x - 3x_2) \\ &= (y + x_1 + x_2 - 2x_0)(y + x_0 + x_2 - 2x_1)(y + x_0 + x_1 - 2x_2) \end{aligned}$$

Pour $y = 0$ on obtient :

$$\psi = (2x_0 - x_1 - x_2)(2x_1 - x_2 - x_0)(2x_2 - x_0 - x_1)$$

**Expression de $\varphi$** Pour déterminer $\varphi = f_1^2 - 3f_2$ nous posons :

On a alors :

$$\begin{aligned} w_1w_2 &= x_0^2 + x_1^2 + x_2^2 + \rho^2(x_0x_1 + x_1x_2 + x_2x_0) + \rho(x_0x_1 + x_1x_2 + x_2x_0). \\ &= (x_0 + x_1 + x_2)^2 + (\rho^2 + \rho - 2)(x_0x_1 + x_1x_2 + x_2x_0) \\ &= (x_0 + x_1 + x_2)^2 - 3(x_0x_1 + x_1x_2 + x_2x_0) = f_1^2 - 3f_2 = \varphi. \end{aligned}$$

donc finalement

$$\varphi = w_1w_2 = (x_0 + \rho x_1 + \rho^2 x_2)(x_0 + \rho^2 x_1 + \rho x_2).$$

**Expression de $\chi$** On pose $y_0 = 3x_0 - f_1$ et on regarde l'équation $y_0 = z + \frac{\varphi}{z}$, c'est-à-dire $z^2 - y_0z + \varphi = 0$.

Par ailleurs, $w_1 + w_2 = x_0 + \rho x_1 + \rho^2 x_2 + x_0 + \rho^2 x_1 + \rho x_2 = 2x_0 - x_1 - x_2 = 3x_0 - f_1 = y_0$

Ainsi, on a $\begin{cases} w_1 + w_2 = y_0 \\ w_1w_2 = \varphi \end{cases}$

Donc $w_1$ et $w_2$ sont les racines de l'équation $z^2 - y_0z + \varphi = 0$.

On rappelle que $\chi = \sqrt{\dfrac{4\varphi^3 - \psi^2}{27}}$, et que ici on pourra poser $z^3 = w_1^3$, et que donc $\frac{\varphi^3}{z} = \frac{w_1^3w_2^3}{w_1^3} = w_2^3$. On a alors :

$$\begin{aligned} -27\chi^2 &= \psi^2 - 4\varphi^3 = (z^3 + \frac{\varphi^3}{z^3})^2 - 4w_1^3w_2^3 \\ -27\chi^2 &= (w_1^3 + w_2^3)^2 - 4w_1^3w_2^3 = (w_1^3 - w_2^3)^2. \\ 3\sqrt{-3}\chi &= w_1^3 - w_2^3 = (w_1 - w_2)(w_1 - \rho w_2)(w_1 - \rho^2 w_2) \text{ (au signe près)} \\ 3\sqrt{-3}\chi &= (\rho - \rho^2)(x_1 - x_2)(1 - \rho)(x_0 - x_1)(1 - \rho^2)(x_0 - x_2) \\ 3\sqrt{-3}\chi &= 3(\rho - \rho^2)(x_0 - x_1)(x_1 - x_2)(x_0 - x_2) \end{aligned}$$

Or, $\rho - \rho^2 = \sqrt{-3}$, donc, au signe près,

$$\chi = (x_0 - x_1)(x_1 - x_2)(x_0 - x_2)$$

On remarque ici que l'on a bien, comme nous l'avions annoncé, l'expression désormais classique du discriminant [16]. On a de plus :

$$\chi^2 = f_1^2 f_2^2 - 4f_1^3 f_3 + 18 f_1 f_2 f_3 - 4f_2^3 - 27 f_3^2$$

**Conclusion** On a donc obtenu :

$$\begin{aligned} \varphi &= (x_0 + \rho x_1 + \rho^2 x_2)(x_0 + \rho^2 x_1 + \rho x_2) = w_1 w_2 \\ \psi &= (2x_0 - x_1 - x_2)(2x_1 - x_2 - x_0)(2x_2 - x_0 - x_1) = w_1^3 + w_2^3 \\ \chi &= (x_0 - x_1)(x_1 - x_2)(x_2 - x_0) \end{aligned}$$

C'est-à-dire que Kronecker a exprimé $\varphi, \psi, \chi$ et $z$ rationnellement en fonction des racines $x_0, x_1$ et $x_2$.

#### 3.1.2.2 Expression d'une racine en fonction d'une autre

Kronecker souhaite maintenant exprimer rationnellement une racine en fonction d'une autre. Dans un essai [17] de 1853, dont nous reparlerons par la suite, il donne la propriété suivante [18] lorsque $\mu$ est un nombre premier :

> Les $\mu$ racines d'une équation résoluble sont toujours représentables l'une en fonction de l'autre, de sorte que :
>
> $$z_2 = f(z_1, r_1), z_3 = f(z_2, r_1), ...z_1 = f(z_\mu, r_1)$$
>
> où $f(z, r_1)$ une fonction rationnelle de $z, r_1$ et $A, B, C, etc.$ et $r_1$ représente la racine d'une équation abélienne dont les coefficients sont des fonctions rationnelles de $A, B, C, etc.$. [Kronecker, 1853] p. 7

Il ajoute que « cette relation des racines de toute équation résoluble est d'ailleurs la vraie source de la propriété d'Abel et Galois donnée comme critère caractéristique des racines d'une équation résoluble de degré premier » [19]. On peut ainsi voir le développement qui suit comme une illustration de cette propriété caractéristique dans le cas où $\mu = 3$.

16. Par exemple [Stewart, 2003] p. 255.
17. Voir [Kronecker, 1853].
18. die $\mu$ Wurzeln einer auflösbaren Gleichung sind immer dergestalt unter einander verbunden, daß :

$$z_2 = f(z_1, r_1), z_3 = f(z_2, r_1), ...z_1 = f(z_\mu, r_1)$$

wo $f(z, r_1)$ eine rationale Function von $z, r_1$ und von $A, B, C, etc.$ bedeutet und $r_1$ die Wurzel einer *Abel*schen Gleichung ist, deren Coëfficienten rationale Functionen von $A, B, C, etc.$ sind.

19. Voir [Kronecker, 1853] p. 7 : « Diese Relation der Wurzeln einer jeden auflösbaren Gleichung ist übrigens die wahre Quelle jener von Abel und Galois als charakteristisches Merkmal der Wurzeln auflösbarer Gleichungen von Primzahlgraden angegebenen Eigenschaft ».

Les 3 racines de l'équation :

$$y^3 - 3\varphi y - \psi = 0$$

sont

$$y_0 = w_1 + w_2 \ , \ y_1 = \rho^2 w_1 + \rho w_2 \ , \ y_2 = \rho w_1 + \rho^2 w_2$$

Kronecker cherche ici à exprimer rationnellement $y_1$ en fonction de $y_0$.

On a $w_2 = \frac{\varphi . w_1^2}{w_1^3}$. On a alors :

$$w_2 = \frac{\varphi}{w_1} = \frac{\varphi \cdot w_1^2}{w_1^3}$$

donc

$$w_1^2 = w_2 \frac{w_1^3}{\varphi} = w_2 \cdot \frac{\psi - 3(\rho - \rho^2)\chi}{2\varphi}$$

et

$$w_2 = w_1^2 \frac{\varphi}{w_1^3} = w_1^2 \cdot \frac{1}{\varphi^2} \cdot \frac{\varphi^3}{w_1^3} = w_1^2 \frac{\psi + 3(\rho - \rho^2)\chi}{2\varphi^2} = w_1^2 \cdot \frac{w_2^3}{(w_1 \cdot w_2)^2}$$

On a ainsi $w_2$ en fonction de $w_1$, grâce au lien qui les relie. On a par ailleurs vu que la formule de Cardan s'écrit :

$$3x = f_1 + \sqrt[3]{w_1^3} + \sqrt[3]{w_2^3}$$

donc

$$3x = f_1 + \sqrt[3]{w_1^3} + \sqrt[3]{w_2^3} = f_1 + \sqrt[3]{w_1^3} + \sqrt[3]{w_1^2 \cdot \frac{w_2^3}{(w_1 \cdot w_2)^2}}$$

On pose maintenant :

$$v_1 = \frac{w_1^2}{w_2} \text{ et } v_2 = \frac{w_2^2}{w_1}$$

ce qui donne :

$$3x = f_1 + \rho^{2k} \sqrt[3]{v_1^2 v_2} + \rho^k \sqrt[3]{v_1 v_2^2}$$

On a $v_1 \cdot v_2 = \varphi$ et $v_1 + v_2 = \frac{\psi}{\varphi}$, donc $v_1$ et $v_2$ sont deux racines d'une équation quadratique.

On a par ailleurs :

$$\begin{aligned}
3x - f_1 &= \rho^{2k} \sqrt[3]{v_1^2 v_2} + \rho^k \sqrt[3]{v_1 v_2^2} \\
(3x - f_1)^3 &= (\rho^{2k} \sqrt[3]{v_1^2 v_2} + \rho^k \sqrt[3]{v_1 v_2^2})^3 \\
(3x - f_1)^3 &= v_1^2 v_2 + v_1 v_2^2 + 3\rho^{5k} \sqrt[3]{v_1^5 v_2^4} + 3\rho^{4k} \sqrt[3]{v_1^4 v_2^5} \\
(3x - f_1)^3 &= v_1 v_2 (v_1 + v_2) + 3 v_1 v_2 (3x - f_1)
\end{aligned}$$

$$(3x - f_1)^3 - v_1v_2(v_1 + v_2) - 3v_1v_2(3x - f_1) = 0$$

**$y_1$ en fonction de $y_0$** On a :

$$\begin{aligned} y_0^2 &= w_1^2 + w_2^2 + 2w_1w_2 \\ -2\varphi + y_0^2 &= w_1^2 + w_2^2 = v_1w_2 + v_2w_1 \end{aligned}$$

Or $y_0 = w_1 + w_2$, donc

$$\begin{aligned} &(\rho^2v_1 - \rho v_2)y_0 + (\rho - \rho^2)(y_0^2 - 2\varphi) \\ &= \rho v_1w_2 + \rho v_2w_1 - \rho^2v_1w_2 - \rho^2v_2w_1 + \rho^2v_1(w_1 + w_2) - \rho v_2(w_1 + w_2) \\ &= (v_1 - v_2)y_1 \end{aligned}$$

Comme

$$v_1 + v_2 = \frac{\psi}{\varphi} \text{ et } v_1 - v_2 = 3(\rho - \rho^2)\frac{\chi}{\varphi}$$

et donc que

$$v_1 = \frac{1}{2}\left(\frac{\psi}{\varphi} + \frac{3(\rho - \rho^2)\chi}{\varphi}\right)$$

on a

$$\begin{aligned} \rho^2v_1 - \rho v_2 &= \rho^2v_1 - \rho v_1 + \rho v_1 - \rho v_2 \\ &= (\rho^2 - \rho)v_1 + \rho(v_1 - v_2) \\ &= (\rho^2 - \rho)v_1 + 3\rho(\rho - \rho^2)\frac{\chi}{\varphi} \\ &= (\rho^2 - \rho)\frac{1}{2}\left(\frac{\psi}{\varphi} + \frac{3(\rho - \rho^2)\chi}{\varphi}\right) + 3\rho(\rho - \rho^2)\frac{\chi}{\varphi} \end{aligned}$$

Si on divise cela par $v_1 - v_2 = 3(\rho - \rho^2)\frac{\chi}{\varphi}$ on obtient :

$$\begin{aligned} &= -\frac{1}{2}\left(\frac{\psi}{3\chi} + \rho - \rho^2\right) + \rho \\ &= -\frac{1}{2}\left(\frac{\psi}{3\chi} + \rho - \rho^2 - 2\rho\right) \\ &= -\frac{1}{2}\left(\frac{\psi}{3\chi} + 1\right) \end{aligned}$$

donc

$$\begin{aligned} (\rho^2v_1 - \rho v_2)y_0 + (\rho - \rho^2)(y_0^2 - 2\varphi) &= (v_1 - v_2)y_1 \\ \frac{(\rho^2v_1 - \rho v_2)y_0}{(v_1 - v_2)} + \frac{(\rho - \rho^2)(y_0^2 - 2\varphi)}{(v_1 - v_2)} &= y_1 \end{aligned}$$

$$-\frac{1}{2}\left(\frac{\psi}{3\chi}+1\right)+\frac{\varphi}{3\chi}(y_0^2-2\varphi)=y_1$$

Finalement $y_1$ est exprimé rationnellement en fonction de $y_0$, car les coefficients sont eux-mêmes des expressions rationnelles de $\varphi, \psi, \chi$.

#### 3.1.2.3 Le *Postulat d'Abel*

Nous venons de voir qu'une grande partie du travail de Kronecker consiste ici à exhiber des expressions rationnelles qui existent entre certaines quantités liées à ces équations [20] :

> Dans notre chaîne apparaissent les valeurs annexes $\chi$ et $z$. Il est d'une signification fondamentale que les valeurs annexes $\chi$ et $z$ soient des fonctions rationnelles des racines $x_0x_1x_2$ de l'équation cubique initiale ($z^3$ est bien égal à $w_1^3$ resp. $w_2^3$). Cette connaissance, une des observations les plus précieuses de la théorie des équations algébriques, apparut pour la première fois chez *Abel*. [Kronecker, 1881a] p. 9

Mais quelle est cette « signification fondamentale » à laquelle Kronecker fait allusion, et comment intercepte-t-elle les travaux d'Abel ?

Dans un article issu du livre *The legacy of Niels Henrik Abel* ([Abel *et al.*, 2004]), Norbert Schappacher et Brigit Petri transcrivent un passage d'un cours de Kronecker où celui-ci décrit ce qu'il entend par *Postulat d'Abel* [21] :

> Résoudre une équation, c'est lui substituer une chaîne d'équations dont la nature est prescrite par avance ; l'essence de la "chaîne" consiste en ceci qu'une grandeur tirée d'une équation est toujours déterminée pour la suivante. Dans le cas de la résolution au moyen d'équations pures, la proposition découverte par Abel – sur laquelle sa démonstration d'impossibilité repose essentiellement – est vraie, et énonce que toutes les grandeurs introduites successivement sont des fonctions rationnelles des racines de l'équation à résoudre. L'exigence que dans toute résolution les grandeurs qui se présentent successivement soient des fonctions rationnelles des racines de l'équation à résoudre, doit être désignée comme le "postulat d'Abel". Sa justification tient au fait que, si l'on renonce à ce postulat, des équations du même genre ont besoin de méthodes de résolution différentes, et donc que par l'entremise des fonctions irrationnelles, la division des fonctions algébriques en genres est rompue. Ce postulat Abélien

20. In unserer Kette traten die Hülfsgrößen $\chi$ und $z$ auf. Von fundamentaler Bedeutung nun ist der Umstand, daß die Hülfsgrößen $\chi$ und $z$ rationale Funktionen der Wurzeln $x_0x_1x_2$ der ursprünglichen kubischen Gleichung sind ($z^3$ ist ja gleich $w_1^3$ resp. $w_2^3$). Diese Erkenntnis, eine der kostbarsten Bemerkungen in der Theorie der algebraischen Gleichungen, stammt erst von *Abel*.

21. Eine Gleichung auflösen heißt sie ersetzen durch eine Kette von Gleichungen vorgeschriebener Beschaffenheit ; das Wesen der „Kette" besteht darin, dass immer aus einer Gleichung eine Größe for die folgenden bestimmt wird. Bei der Auflösung durch reine Gleichungen gilt nun der von Abel entdeckte Satz, – auf dem Abels Unmoglichkeitsbeweis wesentlich beruht -, dass alIe successive eingeführten Größen rationale Funktionen der Wurzeln der aufzulösenden Gleichung sind. Die Forderung, dass bei jeder Auflösung die successive auftretenden Großen rationale Funktionen der Wurzeln der aufzulösenden Gleichung seien, soll als das „Abel'sche Postulat" bezeichnet werden. Die Berechtigung desselben liegt darin, daß, sowie man das Postulat aufgiebt, Gleichungen derselben Gattung verschiedene Auflösungsmethoden nötig haben, dass also bei der Vermittlung durch irrationale Funktionen die Einteilung der algebraischen Funktionen in Gattungen durchbrochen wird. Das Abelsche Postulat ist daher identisch mit der Forderung, dass man keine Methoden der Auflösung algebraischer Gleichungen einführen darf, welche Gleichungen die derselben Gattung angehören auseinander reißen.

> est par conséquent identique à l'exigence que l'on n'introduise aucune méthode de résolution des équations algébriques qui sépare les équations qui appartiennent au même genre. ([Kronecker, 1886] p. 164, retranscrit par B. Petri et N. Schappacher dans [Petri et Schappacher, 2004] pp. 229-230, traduit par Ivahn Smadja)

Ce postulat est donc une exigence de méthode lors de la résolution d'équation : toutes les grandeurs introduites doivent être des fonctions rationnelles des racines. Et on comprend enfin les longs développements que Kronecker expose sur l'équation cubique : en exprimant $\varphi$, $\chi$ et $\psi$ rationnellement en fonction de $x_0, x_1$ et $x_2$, il montre en fait que le postulat d'Abel est respecté dans son étude.

### 3.1.3 Conclusion

Kronecker termine ensuite cette première partie en étudiant certains cas particuliers, comme par exemple les équations cubiques dont le discriminant [22] est égal à 1. Finalement, ce travail sur les équations de degré 2 et 3 est pour Kronecker une introduction, dans laquelle il a pu commencer à exposer certains des outils et certaines méthodes qu'il met en place pour les équations de degré supérieur, équations dont il est question dans la suite de son exposé : l'utilisation de la résolvante, du discriminant, et finalement des polynômes symétriques des racines. Il présente la résolution des équations quadratiques *à la façon de Vandermonde* et les cubiques comme Lagrange les a lui-même présentées : il fait ainsi ressortir le rôle de ces deux mathématiciens, dont l'importance a été soulignée dans son introduction historique, dans la métamorphose que l'étude des équations a subie au court du XIX$^{\text{e}}$ siècle, et se place ainsi dans leur lignée.

Kronecker présente successivement trois expressions différentes de la forme générale de la solution de l'équation cubique :

$$(1)\ 3x = f_1 + z + \frac{\varphi}{z} = f_1 + \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi} + \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho^2 - \rho)\chi}$$

$$(2)\ 3x = f_1 + \sqrt[3]{w_1^3} + \sqrt[3]{w_2^3}$$

$$(3)\ 3x = f_1 + \rho^{2k}\sqrt[3]{v_1^2 v_2} + \rho^k \sqrt[3]{v_1 v_2^2}$$

La première est la forme classique de la formule attribuée à Cardan. Kronecker l'obtient rapidement, mais il ne s'arrête pas là, et il poursuit son travail pour nous amener à la seconde formule. En faisant cela, Kronecker fait explicitement référence à un travail qu'il a publié très jeune, en 1853, et qui s'intitule *Über die algebraisch auflösbaren Gleichungen* ([Kronecker, 1853]). Dans ce cours article, où Kronecker énonce le théorème que l'on appelle aujourd'hui le théorème de Kronecker-Weber [23], il reprend [24] le résultat obtenu par Abel dans son mémoire *Sur la résolution algébrique des équations* :

22. Il attribue d'ailleurs le nom de *discriminant* à Sylvester ([Kronecker, 1881a] p. 14).
23. Ce théorème affirme que toute extension abélienne du corps $\mathbb{Q}$ est cyclotomique.
24. Dans [Kronecker, 1853] p. 5.

> Si une équation irréductible d'un degré premier $\mu$ est résoluble algébriquement, les racines auront la forme suivante :
>
> $$y = A + \sqrt[\mu]{R_1} + \sqrt[\mu]{R_2} + ... + \sqrt[\mu]{R_{\mu-1}}$$
>
> où $A$ est une quantité rationnelle, et $R_1, R_2, ...R_{\mu-1}$ les racines d'une équation du degré $\mu - 1$. [Abel, 1839] p. 190

Dans la seconde formule de Kronecker, les racines de l'équation sont exprimées en fonction de $w_1^3$ et $w_2^3$ qui sont, rappelons-le, les solutions d'une équation quadratique, et nous avons donc là bien un cas particulier de la formule générale donnée par Abel avec $A = f_1$ et $\mu = 3$.

Pour Kronecker la troisième « forme est préférable, car maintenant on a juste à ajouter que $v_1$ et $v_2$ sont les racines d'une équation quadratique et l'expression pour les racines de l'éq. cubique est complètement définie »[25]. Et effet, $v_1$ et $v_2$ sont les racines de $v_1 v_2 = \varphi$ et $v_1 + v_2 = \dfrac{\psi}{\varphi}$, c'est-à-dire d'une équation quadratique dont les coefficients s'expriment rationnellement en fonction des polynômes symétriques élémentaires[26]. En fait, cette forme est celle qui l'amènera à poursuivre un travail sur la forme générale des équations abéliennes de degré 3 (voir par exemple [Kronecker, 1891a] p.92 et suivante), mais nous ne développerons pas cet aspect dans notre travail.

Cette partie sur les équations quadratiques et cubiques qui, dans le manuscrit de 1880-81 tout au moins, conclut l'introduction de son cours, permet à Kronecker de montrer, sur un exemple simple sa façon de travailler. Ainsi, il fait apparaître ici, par la manipulation d'expressions algébriques, transformant celle de Cardan pour d'autres équivalentes mais formellement différentes, sa filiation avec Vandermonde, Lagrange, Abel et Galois, tout en apportant des résultats nouveaux sur les equations abéliennes[27].

## 3.2 Proposition fondamentale des équations générales et théorie des fonctions symétriques et alternées

Voilà comment commence ce premier chapitre[28] (*§3*) :

> Le contraste entre la façon de voir l'algèbre ancienne et actuelle se trouve déjà dans les différentes façons de s'exprimer. On parlait avant d'une *potestas affecta*, laquelle

25. diese Form ist vorzuziehen, denn jetzt hat man nur noch zu sagen, daß $v_1$ und $v_2$ Wurzeln einer quadratischen Gleichung sind und der Ausdrück für die Wurzeln der kubischen Gl. ist vollständig definirt.

26. $\varphi = f_1^2 - 3f_2$ et $\dfrac{\psi}{\varphi} = \dfrac{2f_1^3 - 9f_1 f_2 + 27f_3}{f_1^2 - 3f_2}$.

27. Ces résultats sont à peine esquissés dans le manuscrit de 1880-81, mais ils prendront toute leur mesure dans son dernier cours de 1890-91.

28. Der Gegensatz der früheren Anschauungen in der Algebra zu den heutigen liegt bereits in der verschiedenen Ausdrucksweise ausgesprochen. Früher sprach man von einer *potestas affecta*, welche gleich 0 gesetzt eine *aequatio affecta* darstellte. Man hob also ausschließlich das Operative hervor. Heute ist es der Begriff der Variabilität, der im Vordergrund steht und erst nach der *Cartesi*'schen Entdeckung entstanden ist. Während man früher Unbekannte, Coefficienten und *potestas affecta* unterschied, spricht man jetzt vorzugsweise von Variabeln, Constanten und ganzen Funktionen.

> représentait une *aequatio affecta* posée égale à 0. On a donc souligné exclusivement le fonctionnement. Aujourd'hui c'est la notion de variabilité qui prime et qui s'est formée depuis la découverte de Descartes. Alors que l'on distinguait avant inconnues, coefficients et *potestas affecta*, on parle maintenant de préférence de variables, constantes et fonctions entières. [Kronecker, 1881a] p. 16

Kronecker se positionne résolument dans l'algèbre « actuelle », mais surtout il présente immédiatement ce qui constitue l'une des spécificités de son travail, même s'il ne s'agit ici que d'une allusion : l'absence de distinction entre coefficients et fonctions entières. C'est l'utilisation des fonctions entières en tant que coefficients qui l'amène à proposer les notions *Rationalitäts-Bereich* dans les *Grundzüge* qui seront publiés en 1882.

### 3.2.1 Quelques définitions et premières propriétés

Commençons par essayer de comprendre le vocabulaire employé par Kronecker. L'expression générale d'une fonction entière (*ganze Funktion*) de plusieurs variables est :

$$\sum_{k_1 \ldots k_\nu} c_{k_1 \ldots k_\nu} x_1^{k_1} \ldots x_\nu^{k_\nu} \text{ où } k_1, k_2 \ldots = 0, 1, 2, \ldots ; \; k_1 + \ldots + k_\nu \leq n$$

et où $n$ est la dimension (*Dimension*) de la fonction. Le degré (*Grad*) en $x_k$ est le plus grand degré en $x_k$. Si de plus pour chaque terme on a $\sum k_i = n$, on dit que la fonction est homogène (*homogen*). Une fonction quadratique homogène est appelée une forme quadratique (*quadratische Form*). Une fonction entière est donc un polynôme d'une ou plusieurs variables, dont la nature des coefficients reste à préciser. Kronecker donne en fait une définition inductive des fonctions entières de plusieurs variables : une fonction entière de deux variables est une fonction entière d'une variable dont les coefficients sont des fonctions entières d'une variable.

Cette approche n'est pas celle qu'emploie Weber [29], qui préfère définir directement les fonctions de plusieurs variables par leur forme, mais ce n'est pas non plus celle de Netto [30] – qui a pourtant suivi les cours de Kronecker à l'université de Berlin – qui part aussi d'une forme générale pour les définir avant de les interpréter comme fonctions dont les coefficients sont des fonctions de plusieurs variables [31]. Serret [32] définit bien ces fonctions entières de plusieurs variables comme étant « un polynôme entier et rationnel relativement à chacune des variables », mais il s'agit plutôt là d'une formule pratique que l'énoncé d'une structure qui sera ensuite exploitée. Cette structure, c'est-à-dire le fait de regarder les coefficients comme une fonction entière de plusieurs variables, est l'un

29. [Weber, 1895] p. 64.
30. [Netto, 1900] p. 1.
31. Après Avoir donner une définition très proche de celle de Kronecker, il ecrit :

> Eine ganze Function von $m$ Variabeln kann in die Form einer einzigen dieser Variablen gebracht werden, deren Coefficienten ganze Functionen der $(m-1)$ anderen Variablen sind,
>
> $$(1^a) \; f(z_1, z_2, z_3, \ldots) = \sum_\alpha \varphi_\alpha(z_2, z_3, \ldots) z_1^\alpha, \qquad (\alpha = 0, 1, \ldots r_1);$$

32. [Serret, 1849] p. 143.

des piliers de l'*arithmétique générale* de Kronecker. Toutes ces définitions de fonctions entières, de fonctions entières rationnelles, de fonctions rationnelles fractionnaires, de fonctions entières à coefficients entiers, etc. trouvent une unité dans les cours de Kronecker lorsqu'il introduit les notions de domaine d'intégrité (*Integritäts-Bereich*) et de domaine de rationalité (*Rationalitäts-Bereich*). Cependant, même dans les cours postérieurs à ses *Grundzüge einer arithmetischen Theorie der algebraischen Grössen*, la première partie conserve ces définitions multiples, et ce n'est qu'ensuite, lorsque ce dernier est défini, qu'il revient précisément sur la notion de fonction entière [33].

Kronecker remarque que l'ensemble des fonctions entières forment un groupe fermé (*abgeschlossene Gruppe*), car une fonction entière d'une fonction entière reste une fonction entière [34].
Une fonction rationnelle [35] (*rationale Funktion*) est [36]

> un agrégat construit à partir de grandeurs constantes et variables à l'aide des opérations d'addition, de soustraction, multiplication et division. L'expression générale d'une fonction rationnelle peut être ainsi représentée par le quotient de deux fonctions rationnelles entières. [Kronecker, 1881a] p. 17

Une équation algébrique (*algebraisch Gleichung*) est une fonction entière égalée à 0. Soit l'équation algébrique :

$$c_0 x^n + c_1 x^{n-1} + ... + c_n = 0$$

On la met sous la forme :

$$x^n + \frac{c_1}{c_0} x^{n-1} + ... + \frac{c_n}{c_0} = 0$$

Les $\frac{c_k}{c_0}$ sont appelés les coefficients de l'équation (*Gleichungskoeffizienten*). Si tous ces coefficients sont des nombres entiers, alors on parle de fonction entière à coefficients entiers (*ganz ganzzahligen Funktion*) et d'équation à coefficients entiers (*ganzzahligen Gleichung*).

Kronecker définit ensuite trois grandes catégories d'égalité :

1. Les identités (*Identität*) comme $2.2 = 4$ , $ab = ba, ....$

33. Par exemple dans [Kronecker, 1887a], les Rationalitäts-Bereich et Integritäts-Bereich sont définis p. 121, à peu près au tiers du cours, dans un chapitre intitulé *Über die aritmetischen Grundlagen der ganzen ganzzahligen Funktionen* où il est bien question de redéfinir le domaine des fonctions entières. Pourtant Kronecker a bien sûr utilisé les fonctions entières et les a donc définies dès le début du cours (p. 15-17).

34. On peut lire dans [Kronecker, 1881a] p. 16 : « Schließlich erkennt man, daß alle ganzen Funktionen eine in sich abgeschlossene Gruppe bilden ; denn ganze Funktionen von ganzen Funktionen sind wieder ganze Funktionen. » Kronecker met ainsi en avant le fait que pour les fonctions entières, la composition est, en termes contemporains, une *loi de composition interne*.

35. Remarquons qu'il abandonne ici la distinction qu'il fait dans son cours de 1878/79 entre fonction rationnelle (sommes, produits et différences de fonctions entières) et fonction rationnelle fractionnaire (*rational gebroche Funktion* : sommes, produits, différences et quotient de fonctions entières) pour ne considérer que les fonctions rationnelles, qui sont en fait ses anciennes fonctions rationnelles fractionnaires.

36. Eine rationale Funktion ist ein Aggregat, gebildet aus konstanten und variabeln Größen mittelst der Operationen der Addition, Subtrakion, Multiplikation und Division. Der allgemeine Ausdruck einer rationalen Funktion ist daher durch den Quotienten zweier ganzen rationalen Funktionen dargestellt.

2. Les identités partielles (*teilweise Identität*). Par exemple $\sqrt{4} = 2$ : la racine carrée est pour Kronecker une fonction multivalente, et on pourrait aussi avoir $\sqrt{4} = -2$. Dans le même esprit, il donne aussi l'exemple suivant : $4^{\frac{1}{2}} = 8^{\frac{1}{3}}$.
$\sqrt{1-x} = 1 - \frac{1}{2}X + ...$ est aussi une identité partielle, car elle n'est valable que pour un ensemble de valeurs de $x$.

3. La dernière catégorie est l'équation ordinaire $c_0x^n + c_1x^{n-1} + ... + c_n = 0$. Elle contient une exigence que l'on n'a a priori aucun moyen de satisfaire. Voilà comment la décrit Kronecker [37]

> La troisième catégorie d'égalité comprend les équations ordinaires.
>
> $$c_0x^n + c_1x^{n-1} + ... + c_n = 0$$
>
> exprime l'exigence, dont on ne peut être nullement convaincu *a priori*, qu'elle peut être satisfaite. On demande bien une grandeur $x$ avec laquelle on peut effectuer les opérations habituelles, c'est-à-dire que l'on demande une grandeur qui ait une existence arithmétique. La recherche d'une telle grandeur $x$ qui satisfait l'équation est équivalente à retrouver le nombre de départ. On doit parcourir à l'envers le chemin par lequel on obtient la fonction entière, comme on devait pour la condition $x^n = a$ retrouver $x$. Ces opérations inverses (soustraire, diviser, extraire une racine, résolution d'équations) sont de loin plus difficiles que les opérations directes, mais aussi les plus instructives. Par elles, le domaine des grandeurs connues est amené progressivement à être prolongé aux grandeurs négatives, fractionnaires et complexes. Lors de chacun de ces prolongements des propriétés sont perdues : la notion de nombre cardinal pour les grandeurs négatives, la notion de plus grand et plus petit pour les imaginaires. Nous ne pouvons maintenant toujours pas savoir *a priori* quelles propriétés se perdent par l'exigence de la résolution d'une équation algébrique ou si les grandeurs demandées se laissent joindre aux grandeurs connues. Nous appelons chaque grandeur qui appartient au domaine des grandeurs connues une racine de l'équation, et nous ne savons donc toujours pas si chaque équation a une racine. [Kronecker, 1881a] p. 18

37. Die dritte Categorie von Gleichheiten sind die gewöhnlichen Gleichungen.

$$c_0x^n + c_1x^{n-1} + ... + c_n = 0$$

drückt eine Forderung aus, von der man *a priori* keineswegs überzeugt sein kann, daß sie erfüllbar ist. Man verlangt ja eine Größe $x$ von der Eigenschaft, daß man die gewöhnlichen Operationen mit ihr ausführen kann, d.h. man verlangt eine Größe, welche eine arithmetische Existenz hat. Das Aufsuchen einer solchen Größe $x$, welche die Gl. befriedigt, ist gleichbedeutend mit einem Zurückgehen auf die Grundzahl. Man soll den Weg, auf dem man zu der ganzen Funktion gelangt ist, rückwärts verfolgen, ebenso wie bei der Forderung aus $x^n = a$ rückwärts $x$ zu finden. Diese umgekehrten Operationen (Subtrahiren, dividiren, Radiziren, Auflösung von Gleichungen) sind bei weitem schwieriger als die direkten, aber auch um so lehrreicher. Durch sie veranlaßt ist das Gebiet der bekannten Größen allmählich zu den negativen, gebrochenen und komplexen Größen erweitert worden. Bei einer jeden solchen Erweiterung gehen Eigenschaften verloren, so bei den negativen Größen der Begriff der Anzahl, bei den imaginären das Größer – und Kleinersein. Wir können nun *a priori* noch nicht wissen, welche Eigenschaften durch die Forderung der Auflösung einer algebr. Gl. verlorengehen oder ob die verlangten Größen sich den bekannten anreihen lassen. Wir nennen jede Größe, welche dem Gebiete der bekannten Größen angehört, eine Wurzel der Gl. und wissen also noch nicht, ob jede Gl. eine Wurzel hat.

Revenons sur le sens de cette dernière valeur du signe = en mathématiques. La racine d'une équation n'est pas a priori contenue dans l'équation. Il faut pour cela quelque fois accepter de prolonger le « domaine des grandeurs connues ». On peut voir ici apparaître les prémices du principe, rappelé plus haut [38], qui mettent en jeu ses *Modulsystem*, qui lui permettront d'étendre les domaines de nombres sur lesquels il travaille, et qu'il formalisera dans ses *Grundzüge*.

#### 3.2.1.1 Une équation peut-elle être identiquement nulle ?

À quelles conditions $c_0x^n + c_1x^{n-1} + ... + c_n$ peut-elle être nulle pour toute valeur de $x$ (***§4***) ? Pour cela, on suppose que les $c_k$ sont des réels, $c_0 \neq 0$ et on pose

$$c_0x^n + c_1x^{n-1} + ... + c_n = c_0\left(x^n + \frac{c_1}{c_0}x^{n-1} + ... + \frac{c_n}{c_0}\right).$$

Soit maintenant un réel $M$ plus grand que tous les $\frac{c_k}{c_0}$. Ainsi, pour tout $x > 0$, on a :

$$-M(x^{n-1} + ... + 1) < \frac{c_1}{c_0}x^{n-1} + ... + \frac{c_n}{c_0} < M(x^{n-1} + ... + 1)$$

Pour tout $x > 0$, il existe donc $-1 < \Theta < 1$ tel que

$$\frac{c_1}{c_0}x^{n-1} + ... + \frac{c_n}{c_0} = \Theta M(x^{n-1} + ... + 1) = \Theta M\frac{x^n - 1}{x - 1},$$

On pose maintenant

$$\eta = \frac{M}{x - 1}$$

et donc on a

$$x = 1 + \frac{M}{\eta}$$

Ainsi :

$$\begin{aligned} c_0x^n + c_1x^{n-1} + ... + c_n &= c_0x^n\left(1 + \frac{c_1}{c_0x^n}x^{n-1} + ... + \frac{c_n}{c_0x^n}\right) \\ &= c_0x^n\left(1 + \Theta M\frac{x^n - 1}{x^n(x - 1)}\right) \\ &= c_0x^n\left(1 + \Theta\eta(1 - \frac{1}{x^n})\right) \end{aligned}$$

**Ächte Bruch** Kronecker utilise la distinction *ächte Bruch*/*unächte Bruch* [39] qui correspondent respectivement aux fractions dont le numérateur est inférieur au dénominateur et aux fractions dont le numérateur est supérieur au dénominateur. Par exemple, $\Theta$ est une *ächte Bruch*.

38. Voir p. 78.

39. Ce qui signifie en allemand du XXI^e^ siècle (et déjà quelquefois à la toute fin du XIX^e^ siècle) *echter Bruch*/*unechter Bruch*.

Pour $x$ suffisamment grand, $0 < \eta < 1$, et donc $-1 < \Theta\eta(1 - \frac{1}{x^n}) < 1$. Ainsi :

$$1 + \Theta\eta(1 - \frac{1}{x^n}) > 0$$

Finalement, pour $x$ suffisamment grand, $c_0 x^n \left(1 + \Theta\eta(1 - \frac{1}{x^n})\right)$ est du signe de $c_0$, et il est donc différent de zéro. Donc pour $c_0 \neq 0$ la fonction ne peut pas être identiquement nulle. Kronecker ne conclut pas, mais la conséquence de cela est évidente : pour qu'une fonction soit identiquement nulle il faut que tous ses coefficients soient nuls.

Il faut souligner ici que Kronecker donne un moyen effectif de trouver ce $x$ suffisamment grand : il suffit de prendre une *ächte Bruch* positive et de poser $x = 1 + \frac{M}{\eta}$. Il fait ensuite remarquer que la preuve classique de ce théorème, « Si l'on pose x=0 etc. », est bien juste (*richtig*) sur le fond, mais pas sur la forme : il veut que sa démonstration puisse s'appliquer aux fonctions entières transcendantes. Il finit en donnant la même démonstration en partant cette fois de l'expression :

$$c_{n-k}x^k + c_{n-k-1}x^{k+1} + ...$$

#### 3.2.1.2 Deux théorèmes fondamentaux de la théorie des équations algébriques

Kronecker s'attaque ensuite (***§5***) à deux théorèmes qu'il qualifie de « fondamentaux » (*Fundamentalsätze*) :

**Théorème 1** : toute fonction entière de la racine d'une équation algébrique de degré $n$ peut être réduite à une fonction entière de degré $n - 1$ de cette racine.

En effet de

$$c_0x^n + c_1x^{n-1} + ... + c_n = c_0\left(x^n + \frac{c_1}{c_0}x^{n-1} + ... + \frac{c_n}{c_0}\right) = 0,$$

on tire

$$x^n + \frac{c_1}{c_0}x^{n-1} + ... + \frac{c_n}{c_0} = 0,$$

c'est-à-dire

$$x^n = -\frac{c_1}{c_0}x^{n-1} - ... - \frac{c_n}{c_0}$$

ce qui nous permet d'abaisser tous les monômes de degré supérieur ou égal à $n$ de un. En répétant la même opération on arrive en un nombre fini d'étapes à une fonction entière de degré $n - 1$. On appelle cette forme la forme réduite (*reduzirte Form*) de la fonction entière. De plus, les coefficients de la forme réduite sont des fonctions linéaires entières des coefficients de la fonction à réduire et des fonctions entières à coefficients entiers des coefficients de l'équation.

Soit $a_0x^r + ... + a_r$ la fonction à réduire et $c_0x^n + c_1x^{n-1} + ... + c_n = 0$ l'équation

algébrique. On peut faire la division euclidienne :

$$a_0x^r + ... + a_r = \left(x^n + \frac{c_1}{c_0}x^{n-1} + ... + \frac{c_n}{c_0}\right) Q(x) + R(x)$$

Si $x$ est une racine de l'équation, alors $x^n + \frac{c_1}{c_0}x^{n-1} + ... + \frac{c_n}{c_0} = 0$. Dans ce cas on a

$$a_0x^r + ... + a_r = R(x)$$

avec le degré de $R$ inférieur ou égal à $n-1$. Ici Kronecker conclut à l'égalité entre $R$ et la forme réduite cherchée. L'unicité de la forme réduite, certes évidente par sa construction, n'est pas explicitement utilisée.

**Exemple :** On réduit $x^4$ avec l'aide de l'équation cubique $x^3 - f_1x^2 + f_2x - f_3 = 0$.

On fait la division euclidienne et on obtient :

$$x^4 = (x + f_1)(x^3 - f_1x^2 + f_2x - f_3) + (f_1^2 - f_2)x^2 + (f_3 - f_1f_2)x + f_1f_3$$

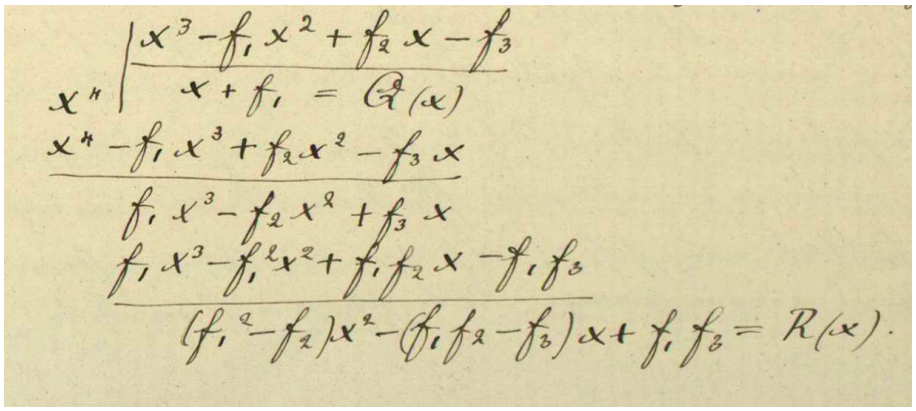



FIGURE 1 – [Kronecker, 1881a] p. 20

On a ainsi la forme réduite :

$$x^4 = (f_1^2 - f_2)x^2 + (f_3 - f_1f_2)x + f_1f_3$$

**Théorème 2** : Soit $\xi_1$ une racine de $c_0x^n + c_1x^{n-1} + ... + c_n = 0$, alors la fonction entière $c_0x^n + c_1x^{n-1} + ... + c_n$ reste algébrique lors de la division par $x - \xi_1$.

On a en effet $c_0\xi_1^n + c_1\xi_1^{n-1} + ... + c_n = 0$, donc

$$\begin{aligned} c_0x^n + c_1x^{n-1} + ... + c_n &= c_0x^n + c_1x^{n-1} + ... + c_n - (c_0\xi_1^n + c_1\xi_1^{n-1} + ... + c_n) \\ &= c_0(x^n - \xi_1^n) + c_1(x^{n-1} - \xi_1^{n-1}) + ... + c_{n-1}(x - \xi_1) \end{aligned}$$

D'où le résultat. On a donc

$$c_0x^n + c_1x^{n-1} + ... + c_n = (x - \xi_1)(c_0x^{n-1} + c'_1x^{n-2} + ... + c'_{n-1})$$

Si on a une seconde racine $\xi_2$ :

$$c_0x^n + c_1x^{n-1} + ... + c_n = (x - \xi_1)(x - \xi_2)(c_0x^{n-2} + c''_1x^{n-3} + ... + c''_{n-2})$$

Si on a $n$ racines $\xi_1, ..., \xi_n$, alors :

$$c_0x^n + c_1x^{n-1} + ... + c_n = (x - \xi_1)(x - \xi_2)...(x - \xi_n)Q(x)$$

et dans ce cas on a $Q(x) = c_0$. On a ainsi démontré que si l'équation algébrique $c_0x^n + c_1x^{n-1} + ... + c_n = 0$ de degré $n$ a $n$ racines $\xi_1, ..., \xi_n$, alors on a la factorisation linéaire :

$$c_0x^n + c_1x^{n-1} + ... + c_n = c_0(x - \xi_1)(x - \xi_2)...(x - \xi_n).$$

À partir de cette forme, la proposition suivante est évidente :

**Proposition :** Une fonction entière de degré $n$ a au plus $n$ racines distinctes.

Kronecker termine en faisant quelques remarques :

1. On peut étendre ces résultats aux fonctions de plusieurs variables. Si une fonction de deux variables $x$ et $y$ de degrés respectifs $m$ et $n$ qui s'annule pour plus de $m$ valeurs de $x$ ou plus de $n$ valeurs de $y$, alors elle est identiquement nulle.
2. Il est équivalent de supposer que toute équation de degré $n$ a une racine ou que toute équation de degré $n$ a $n$ racines (certaines d'entre elles pouvant être identiques).
3. On ne sait toujours pas si toute équation algébrique a des racines : ce sera l'objet de la toute fin de son cours [40]. Cependant, on travaille maintenant avec

   $$F(x) = (x - x_1)(x - x_2)...(x - x_n)$$

   pour déterminer un certain nombre des propriétés des polynômes symétriques des $x_1, ..., x_n$.

### 3.2.2 Fonctions symétriques

Kronecker passe maintenant aux fonctions symétriques et à leurs propriétés (***§6***), ce qui constitue une partie importante de son travail.

40. En tout cas en ce qui concerne [Kronecker, 1881a].

#### 3.2.2.1 Définitions et polynômes symétriques élémentaires

Soit

$$F(x) = (x - x_1)(x - x_2)...(x - x_n)$$

dont les racines sont $x_1, ..., x_n$.

$$\begin{aligned} F(x) = (x - x_1)(x - x_2)...(x - x_n) &= \prod_{k=1}^{n}(x - x_k) \\ &= x^n - f_1 x^{n-1} + ... + (1)^n f_n \\ &= \sum_{h=0}^{n}(-1)^h f_h x^{n-h} = F(x, f_1, ..., f_n) \end{aligned}$$

avec

$$\begin{aligned} f_0 &= 1 \\ f_1 &= x_1 + x_2 + ... + x_n \\ f_2 &= x_1x_2 + ... + x_{n-1}x_n \\ ... &= ... \\ f_n &= x_1x_2...x_n \end{aligned}$$

Chaque fonction $f_h$ est une fonction homogène entière à coefficients entiers des variables $x_1, ..., x_n$ de dimension $h$.

**Définition :** Une fonction de $n$ variables $x_1, ..., x_n$ est dite symétrique si pour chacun des $n - 1$ échanges (*Vertauschungen*) $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\begin{pmatrix} x_2 \\ x_3 \end{pmatrix}$,..., $\begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix}$ elle reste inchangée.

On peut alors montrer qu'une telle fonction reste inchangée par toute transposition (*Paarvertauschungen*) $\begin{pmatrix} x_h \\ x_k \end{pmatrix}$ et toute permutation (*Permutation*) $\begin{pmatrix} x_1...x_n \\ x_{i_1}...x_{i_n} \end{pmatrix}$.

On peut démontrer que les fonctions $f_i$ sont symétriques :

$$\prod_{k=1}^{n}(x - x_k) = \sum_{h=0}^{n}(-1)^h f_h(x_1, ..., x_n)x^{n-h}$$

Soit la permutation $\begin{pmatrix} x_1...x_n \\ x_{i_1}...x_{i_n} \end{pmatrix}$. Comme $\prod_{k=1}^{n}(x - x_k)$ reste inchangée par cette permutation, on a :

$$\prod_{k=1}^{n}(x - x_k) = \sum_{h=0}^{n}(-1)^h f_h(x_{i_1}, ..., x_{i_n})x^{n-h}$$

donc

$$\sum_{h=0}^{n}(-1)^h f_h(x_1,...,x_n)x^{n-h} = \sum_{h=0}^{n}(-1)^h f_h(x_{i_1},...,x_{i_n})x^{n-h}$$

et par identification des coefficients, on a pour tout $h$ :

$$f_h(x_1,...,x_n) = f_h(x_{i_1},...,x_{i_n})$$

**Définition :** Les fonctions $f_i$ sont appelées les fonctions symétriques élémentaires (*elementaren symmetrischen Funktionen*).

Kronecker termine ce paragraphe en précisant qu'il préférerait donner aux fonctions symétriques le nom de *functiones invariabiles*, d'après les « fonctions invariables » de Vandermonde.

#### 3.2.2.2 Construction et utilisations des fonctions $F_k$

Kronecker construit des fonctions qui seront particulièrement importantes dans la partie suivante de notre travail.

**Construction** Soit (***§7***)

$$F(x) = (x - x_1)(x - x_2)...(x - x_n)$$

On peut, d'après le théorème 1 ci-dessus, réduire toute fonction entière en $x_k$ en une fonction entière de degré en $x_k$ inférieur ou égal à $n - 1$ dont les coefficients sont des fonctions entières à coefficients entiers des coefficients de $F(x)$, c'est-à-dire des $f_i$.

Rappelons que cette façon d'envisager une fonction entière de plusieurs variables comme une fonction entière d'une seule variable dont les coefficients sont des fonctions entières des autres variables est l'une des clefs pour lire nombre de ses travaux ultérieurs, comme par exemple sa démonstration du théorème fondamental de l'arithmétique de 1887 ([Kronecker, 1887c]).

Soit $\varphi(x_k)$ une fonction entière de $x_k$, on peut alors la réduire sous la forme :

$$\varphi(x_k) = \varphi_0(f_1,...,f_n) + \varphi_1(f_1,...,f_n)x_k + ... + \varphi_{n-1}(f_1,...,f_n)x_k^{n-1}$$

où, dit Kronecker, les $\varphi_i$ sont des fonctions entières à coefficients entiers des $f_i$. Il y a là une erreur ou une imprécision : soit $\varphi(x_k)$ est elle-même à coefficients entiers et la proposition est vraie, soit ce n'est pas le cas et alors les $\varphi_i$ ne sont elles-mêmes que des fonctions entières dont on ne sait rien des coefficients.

Kronecker revient sur la lacune que l'on avait relevée précédemment : l'unicité de la forme réduite. Supposons qu'il en existe deux, alors leur différence est une

fonction de degré $n-1$ qui s'annule pour les $n$ valeurs $x_1, ..., x_n$. D'après une des propositions précédentes, c'est donc la fonction nulle : les deux formes réduites sont identiques.

Du fait que $F(x_1) = 0$, on tire :

$$\begin{aligned}\frac{F(x)}{x-x_1} &= \frac{F(x)-F(x_1)}{x-x_1} \\ &= x^{n-1} + (x_1 - f_1)x^{n-2} + (x_1^2 - f_1x_1 + f_2)x^{n-3} + ... + (x_1^{n-1} - f_1x_1^{n-2} + ... + f_{n-1}) \\ &= F_1(x, x_1, f_1, ..., f_{n-1}) \\ &= \frac{\sum(-1)^h f_h(x^{n-h} - x_1^{n-h})}{x-x_1} = \sum(-1)^h f_h \sum_{(\mu+\nu=n-h-1)} x^\mu x_1^\nu\end{aligned}$$

On obtient ainsi une fonction $F_1(x, x_1, f_1, ..., f_{n-1})$ qui a les propriétés suivantes :

1. Elle est symétrique et de degré $n-1$ en $x$ et $x_1$.
2. Elle est de dimension $n-1$.
3. Elle est linéaire en $f_1, ..., f_{n-1}$.
4. $F_1(x, x_1) = (x-x_2)...(x-x_n)$, donc elle possède les racines $x_2, ..., x_n$.
5. $F_1(x_1, x_1) = F'(x_1) = (x_1-x_2)...(x_1-x_n)$

On pose maintenant :

$$\begin{aligned}F_2(x, x_1, x_2, f_1, ..., f_{n-2}) &= \frac{F_1(x, x_1)}{(x-x_2)} \\ &= \frac{F_1(x, x_1) - F_1(x_2, x_1)}{(x-x_2)} \\ &= x^{n-2} + x^{n-3}(x_1 + x_2 - f_1) + x^{n-4}(x_1^2 + x_1x_2 + x_2^2 - f_1(x_1+x_2) + f_2) + ...\end{aligned}$$

On obtient ainsi une fonction $F_2(x, x_1, x_2, f_1, ..., f_{n-2})$ qui a les propriétés suivantes :

1. Elle est symétrique et de degré $n-2$ en $x$, $x_1$ et $x_2$.
2. Elle est de dimension $n-2$.
3. Elle est linéaire en $f_1, ..., f_{n-2}$.
4. $F_2(x, x_1, x_2) = (x-x_3)...(x-x_n)$, donc elle possède les racines $x_3, ..., x_n$.

En continuant à procéder ainsi, on obtient :

$$\begin{aligned}F(x, f_1, ...f_n) &= (x-x_1)(x-x_2)...(x-x_n) \\ F_1(x, x_1, f_1, ...f_{n-1}) &= (x-x_2)(x-x_3)...(x-x_n) \\ F_2(x, x_1, x_2, f_1, ...f_{n-2}) &= (x-x_3)(x-x_4)...(x-x_n) \\ ... &\quad ... \\ F_{n-2}(x, x_1, ...x_{n-2}, f_1, f_2) &= (x-x_{n-1})(x-x_n) \\ F_{n-1}(x, x_1, ...x_{n-1}, f_1) &= (x-x_n)\end{aligned}$$

De $F(x, f_1, ...f_n) = (x - x_1)(x - x_2)...(x - x_n)$ on déduit que les fonctions symétriques élémentaires des $n$ racines sont égales aux coefficients de l'équation, et de $F_k(x, x_1, ...x_k, f_1, ..., f_{n-k}) = (x - x_{k+1})...(x - x_n)$ on déduit que les fonctions symétriques élémentaires de $n - k$ racines s'expriment comme fonctions symétriques entières à coefficients entiers des $k$ racines restantes avec l'aide des fonctions symétriques élémentaires de toutes les racines.

**Propriétés des fonctions $F_k$.** Nous allons démontrer les propriétés suivantes :

1. $F_k(x_r) \equiv 0$ pour $r > k$.
2. $F_k$ est une fonction entière à coefficients entiers des $n + 1$ grandeurs $x, x_1, ..., x_k, f_1...f_{n-k}$ et les coefficients des plus hautes puissances de $x, x_1, ..., x_k$ sont égaux à 1.
3. Par rapport à $x, x_1, ...x_k$, $F_k$ est de degré $n - k$ et de dimension $n - k$.
4. $F_k$ est linéaire en les $f_i$.
5. Par rapport à $x, x_1, ...x_k$, $F_k$ est symétrique.

Nous allons essentiellement faire une démonstration par induction, et donc supposer ces propriétés vraies pour $F_k$ et tenter de les déduire pour $F_{k+1}$. On a par ailleurs déjà montré ces propriétés pour $k = 1$ et partiellement pour $k = 2$.

On a la formule

$$F_{k+1} = \frac{F_k(x, x_1, ...x_k, f_1, ..., f_{n-k}) - F_k(x_{k+1}, x_1, ...x_k, f_1, ..., f_{n-k})}{x - x_{k+1}}$$

ce qui rend les propositions 1), 2), 3) et 4) évidentes. Il reste à démontrer la dernière.

Déjà, $F_{k+1}$ reste inchangée pour les échanges $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\begin{pmatrix} x_2 \\ x_3 \end{pmatrix}$,..., $\begin{pmatrix} x_{k-1} \\ x_k \end{pmatrix}$, $\begin{pmatrix} x_{k+1} \\ x \end{pmatrix}$. Il reste à montrer que l'échange $\begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix}$ ne l'altère pas. Cela est évidemment le cas car

$$F_{k+1} = (x - x_{k+2})...(x - x_n)$$

ne contient ni $k$, ni $k + 1$, ce qui conclut notre démonstration.

Les propriétés ont donc été démontrées de façon indirecte en utilisant un raisonnement par induction.

Kronecker veut ensuite en donner une preuve directe (*direkten Beweis*). Pour cela, il utilise la forme suivante de $F_k$ :

On obtient $F_k$ à partir de $F$ en divisant par $(x-x_1)...(x-x_k)$. On a ainsi :

$$F(x) = \sum_{h=0}^{n}(-1)^h f_h x^{n-h} \qquad = (x-x_1)...(x-x_k)Q(x,x_1,...,x_k,f_1...) + R(x_1,...)$$

Et il redémontre chaque propriété, sauf la 5), en utilisant le fait que $F_k = Q(x)$, les propriétés de la division, et en choisissant parfois d'ignorer le sens des $f_i$.

Pour démontrer la 5), on écrit, en posant $x_0 = x$ :

$$\begin{aligned} F_k &= \frac{F(x)}{(x-x_1)...(x-x_k)} + \frac{F(x_1)}{(x_1-x)(x_1-x_2)...(x_1-x_k)} + ... + \frac{F(x_k)}{(x_k-x)...(x_k-x_{k-1})} \\ &= \sum_{r=0}^{r=k} \frac{F(x_r)}{\prod_{t\neq r}(x_r-x_t)} = \sum_{h=0}^{n}(-1)^h f_h \sum_{r=0}^{r=k} \frac{x_r^{n-h}}{\prod_{t\neq r}(x_r-x_t)} \end{aligned}$$

et on voit ainsi immédiatement sous cette forme que $F$ est symétrique par rapport à $x, x_1, ..., x_k$.

C'est la première fois que Kronecker propose de revenir sur une démonstration déjà établie pour en donner une autre qu'il estime *meilleure*. Ici, on voit bien que l'intérêt de la seconde démonstration est de fournir une construction effective des fonctions $F_k$ qui permet de « voir » la propriété. Ce retour sur des propriétés démontrées, s'il n'est pas systématique, apparaît souvent dans ses leçons.

**Réduction grâce aux $F_k$** On veut réduire la fonction des $n$ racines de l'équation $F(x) = 0$ :

$$\sum_{\alpha,\beta,\gamma,...} C_{\alpha,\beta,\gamma,...} x_1^\alpha x_2^\beta x_3^\gamma ...$$

à l'aide des fonctions $F_k$.

Ici (***§8***), une remarque s'impose. Dans les travaux de Kronecker, le « lieu naturel » de l'étude de ce type de fonctions est le *Rationalitäts-Bereich*, notion qu'il introduit dans ses *Grundzüge* en 1882. Et de fait, dans ses cours de 1885 et 1890, c'est bien l'outil qu'il emploiera pour présenter cette même réduction. Néanmoins, au moment où il professe son cours, durant cet hiver 1880-1881, il doit déjà avoir à l'esprit cette notion.

De plus, la nature des coefficients $C$ est la plupart du temps passée sous silence. Cette négligence est peut-être l'œuvre du rédacteur, mais il est plus probable que le contexte doit suffire, pour un étudiant de cette époque, à déterminer cette nature. Peut-être même cette nature n'a pas encore l'importance qu'elle revêtira plus tard. Ce ne sera plus le cas dans

son cours de 1890, où les coefficients sont d'abord des nombres entiers, puis sont étendus à d'autres ensembles.

— $x_n$ est solution de $F_{n-1}(x) = x_n - x = 0$. Avec $x_n = f_1 - x_1 - ... - x_{n-1}$ on peut retirer tous les $x_n$.

— $x_{n-1}$ est solution de l'équation du second degré $F_{n-2}(x) = 0$, on peut alors réduire les $x_{n-1}$ à des termes du premier degré.

— $x_{n-2}$ est solution de l'équation du troisième degré $F_{n-3}(x) = 0$, on peut alors réduire les $x_{n-2}$ à des termes du second degré.

— On continue ainsi jusqu'à réduire les $x_1$ à des termes de degré $n-1$.

Ainsi réduite, $\sum_{\alpha,\beta,\gamma,...} C_{\alpha,\beta,\gamma,...} x_1^\alpha x_2^\beta x_3^\gamma ...$ prend la forme

$$\sum_{r_1,r_2,...} A_{r_1 r_2 ...} x_1^{r_1} x_2^{r_2} ... x_n^{r_n}$$

où $r_n = 0$, $r_{n-1} \leq 1$,..., $r_k \leq n-k$,...,$r_1 \leq n-1$.

Remarque importante : Par construction, et d'après les propriétés des fonctions $F_k$ déjà démontrées, les coefficients $A$ sont des fonctions linéaires des coefficients $C$ et par ailleurs aussi des fonctions rationnelles entières à coefficients entiers des fonctions symétriques élémentaires $f_i$ : ces coefficients sont donc des fonctions symétriques des $x_i$.

**Réduction des fonctions symétriques** Supposons que la fonction que l'on cherche à réduire soit une fonction entière symétrique par rapport aux variables $x_1, ..., x_n$.

Dans la forme réduite, il n'y a plus de $x_n$, et $x_{n-1}$ n'apparaît qu'à la puissance 1 : on peut donc écrire cette forme en ordonnant suivant les $x_{n-1}$ :

$$\Phi x_{n-1} + \Psi$$

où $\Phi$ et $\Psi$ ne contiennent ni $x_n$ ni $x_{n-1}$, et ne changent donc pas si l'on permute $x_n$ et $x_{n-1}$. On a donc

$$\Phi x_{n-1} + \Psi = \Phi x_n + \Psi$$

et donc $\Phi(x_{n-1} - x_n) = 0$, c'est-à-dire $\Phi = 0$. Il n'y a donc pas de terme en $x_{n-1}$.

Après avoir ordonné la forme réduite selon les $x_{n-2}$, on obtient :

$$\Phi_1 x_{n-2}^2 + \Psi_1 x_{n-2} + \chi_1$$

où $\Phi_1, \Psi_1$ et $\chi_1$ ne contiennent ni $x_n$ ni $x_{n-1}$ ni $x_{n-2}$, et ne changent donc pas si l'on permute $x_n$, $x_{n-1}$ et $x_{n-2}$. On a ainsi :

$$\Phi_1 x_{n-2}^2 + \Psi_1 x_{n-2} + \chi_1 = \Phi_1 x_{n-1}^2 + \Psi_1 x_{n-1} + \chi_1 = \Phi_1 x_n^2 + \Psi_1 x_n + \chi_1 = A$$

et donc l'équation $\Phi_1 x^2 + \Psi_1 x + \chi_1 - A = 0$ a trois racines distinctes. Ses coefficients sont donc nuls, ce qui signifie que

$$\Phi_1 = \Psi_1 = 0$$

Il n'y a donc pas de termes en $x_{n-2}$ dans la forme réduite.

On continue ainsi et on s'aperçoit qu'il n'y a finalement aucun $x_i$ dans la forme réduite, et que celle-ci se résume donc au terme $A_{00\ldots0}$. Kronecker passe ensuite à la partie *existence* du théorème que l'on trouve aujourd'hui sous le nom de *théorème fondamental des polynômes symétriques* : tout polynôme symétrique peut s'exprimer de façon unique comme polynôme des fonctions symétriques élémentaires.

**Unicité de la forme réduite** Supposons que pour

$$\sum_{\alpha,\beta,\gamma,\ldots} C_{\alpha,\beta,\gamma,\ldots} x_1^\alpha x_2^\beta x_3^\gamma \ldots$$

il existe deux formes réduites

$$\sum_{r_1,r_2,\ldots} A_{r_1 r_2 \ldots} x_1^{r_1} x_2^{r_2} \ldots x_n^{r_n}$$

et

$$\sum_{r_1,r_2,\ldots} A'_{r_1 r_2 \ldots} x_1^{r_1} x_2^{r_2} \ldots x_n^{r_n}$$

où $r_n = 0$, $r_{n-1} \leq 1$,..., $r_k \leq n-k$,...,$r_1 \leq n-1$.

Ces deux formes doivent être égales, et donc égales terme à terme, c'est-à-dire que pour tout $r_1, \ldots, r_n$ on doit avoir

$$A_{r_1 r_2 \ldots} - A'_{r_1 r_2 \ldots} = 0$$

Les deux formes réduites sont donc identiques.

**Résumé et conséquences** Kronecker fait alors un résumé de ce qu'il vient de démontrer, en y ajoutant quelques remarques (***§9***) :

**Proposition** : Toute fonction entière des $n$ racines d'une équation algébrique peut être réduite de manière unique à une forme dont les coefficients sont des fonctions linéaires des coefficients de la fonction originale à réduire et des fonctions entières à coefficients entiers des coefficients de l'équation.

De là il suit la proposition :

**Proposition** : Toute fonction entière à coefficients entiers des $n$ racines d'une équation à coefficients entiers a dans sa forme réduite pour coefficients les $A_{r_i}$, qui sont des nombres entiers.

Si la fonction à réduire est symétrique, seul le terme $A_{000...}$ compte, donc :

**Proposition** : Toute fonction entière symétrique des $n$ racines d'une équation algébrique est une fonction entière à coefficients entiers des coefficients de l'équation.

On a comme cas particulier :

**Proposition** : Toute fonction entière symétrique à coefficients entiers des $n$ racines d'une équation algébrique à coefficients entiers est un nombre entier.

**Proposition** : Pour toute fonction symétrique des $n$ racines d'une équation algébrique entière on a pour chacune des $n!$ permutations :

$$S(x_1, ..., x_n) = S(x_{i_1}...x_{i_n})$$

Kronecker montre cette dernière proposition en reprenant la définition qu'il a donnée d'une fonction symétrique dans le cas $n = 3$ et en utilisant le fait que les permutations sont engendrées par les transpositions.

**Passages aux fonctions rationnelles** On veut maintenant (***§10***) réduire la fonction rationnelle

$$\frac{\varphi(x_1)}{\psi(x_1)}$$

où $x_1$ est une des $n$ racines d'une équation algébrique

$$\frac{\varphi(x_1)}{\psi(x_1)} = \frac{\varphi(x_1)\psi(x_2)...\psi(x_n)}{\psi(x_1)\psi(x_2)...\psi(x_n)}$$

Le dénominateur, $\psi(x_1)\psi(x_2)...\psi(x_n)$ est une fonction entière symétrique de $x_1, x_2, ..., x_n$, c'est donc une fonction entière des $f_i$.

Dans le numérateur, $\psi(x_2)...\psi(x_n)$ est une fonction entière symétrique par rapport aux racines de $F_1(x)$. On peut donc la réduire comme fonction entière de $x_1$ et des coefficients de $F_1(x)$, c'est-à-dire d'après ce que l'on a vu précédemment, puisque les coefficients de $F_1(x)$ sont eux-mêmes des fonctions entières de $x_1, f_1, ..., f_{n-1}$, comme fonction entière de $x_1, f_1, ..., f_{n-1}$. Pour obtenir le numérateur, on multiplie maintenant par $\varphi(x_1)$ et on conserve ainsi une fonction entière de $x_1, f_1, ...f_{n-1}$, que l'on réduira pour obtenir une

fonction entière de $x_1$ de degré $n-1$.

Conclusion : On obtient ainsi une fonction entière de $x_1$ dont les coefficients sont des fonctions rationnelles de $f_1, ...f_n$.

On peut prolonger cette proposition aux fonctions rationnelles de $n$ variables :

$$\frac{\varphi(x_1, ..., x_n)}{\psi(x_1, ..., x_n)}$$

et obtenir la proposition suivante :

Proposition : Toute fonction rationnelle des $n$ racines d'une équation algébrique de degré $n$ se laisse représenter comme une fonction entière de $x_1, ..., x_n$ dont les coefficients sont des fonctions rationnelles de $f_1, ..., f_n$.

#### 3.2.2.3 Un cas particulier : étude de la fonction $F = \prod_{(i)}(z - \psi(x_{i_1}, ..., x_{i_n}))$

Soit la fonction (***§11***) :

$$F = \prod_{(i)}(z - \psi(x_{i_1}, ..., x_{i_n}))$$

où $\psi(x_{i_1}, ..., x_{i_n})$ est une fonction entière à coefficients entiers [41]. C'est une fonction entière de degré $n!$ en $z$ symétrique par rapport à $x_1, ..., x_n$. Les coefficients de cette fonction entière sont donc, d'après ce qui précède, des fonctions entières à coefficients entiers des $f_i$.

**Remarque :** Dans le cours de 1890/91 (p. 132) Kronecker profite de cette étude pour donner la définition d'un nombre algébrique (*algebraische Zahlen*) : c'est une racine d'une équation algébrique à coefficients entiers. Si le polynôme est de plus unitaire, ce que semble supposer Kronecker la plupart du temps, cela correspond à la définition actuelle d'un *entier algébrique*.

On a ainsi la propriété suivante :

**Propriété :** Toute fonction entière à coefficients entiers ***des*** $n$ ***racines*** d'une équation algébrique de degré $n$ est la racine d'une équation de degré $n!$ dont les coefficients sont des fonctions entières à coefficients entiers de $f_1, ..., f_n$.

41. La nature des coefficients de $\psi$ n'est pas précisée. Ce type d'oubli est particulièrement gênant lorsque comme ici l'essentiel des propositions porte sur cette nature. Dans les cours plus tardifs (1886-87 p. 77, 1890-91 p. 131), la nature des coefficients est systématiquement explicitée ; par contre ce n'est pas le cas dans le précédent (1878-79 p. 75).

Regardons le cas particulier :

$$\prod_{k=1}^{n}(z-\psi(x_k))=0$$

On a donc la propriété :

**Propriété :** Toute fonction entière à coefficients entiers ***d'une racine*** d'une équation algébrique de degré $n$ est la racine d'une équation de degré $n$ dont les coefficients sont des fonctions entières à coefficients entiers de $f_1, ..., f_n$.

Si on prend maintenant par exemple $\psi(x) = x^r$, à partir de

$$\prod_{k=1}^{n}(z-x_k^r)=0$$

on a :

**Propriété :** Les puissances r$^{\text{ème}}$ des $n$ racines d'une équation algébrique de degré $n$ sont les racines d'une équation de degré $n$ dont les coefficients sont des fonctions entières à coefficients entiers de $f_1, ..., f_n$.

**Avec des racines de module 1** Kronecker reprend ici le contenu d'un article de 1857 paru dans le *Journal de Crelle*[42]. On remarquera aussi que la seconde partie de cet article remplace cette étude dans les manuscrits des *Wintersemester* 1888/89 et 1890/91. Si l'on suppose maintenant que les racines sont toutes complexes de module 1, on peut les représenter dans le plan complexe sur un cercle ayant pour centre l'origine et de rayon 1, et elles ont toutes une expression de la forme :

$$e^{vi} = \cos v + i \sin v$$

Si $\xi$ est l'une de ces racines, alors c'est aussi le cas de $\frac{1}{\xi}$ et on dit que ces racines sont conjuguées. De plus, dans ce cas, pour tout $r \in \mathbb{N}$

$$\prod_{k=1}^{n}(z-x_k^r)=0$$

est une équation dont les racines sont aussi des nombres complexes de module 1. Nous allons montrer la proposition suivante :

42. Voir [Kronecker, 1857].

**Propriété :** Il y a un nombre fini d'équations de la forme $\prod_{k=1}^{n}(z-x_k^r)=0$.

Soit

$$x_1 = e^{v_1 i}\ ,\ x_2 = e^{v_2 i}\ ,\ ...x_n = e^{v_n i}$$

et on a donc

$$f_k = \sum e^{(v_1+...+v_k)i}$$

où $f_k$ a $\frac{n(n-1)...(n-k+1)}{k!}$ termes. Si l'on note

$$f_{k,r} = \sum x_{i_1}^r ... x_{i_k}^r$$

Alors les $f_{k,r}$ contiennent le même nombre de termes que les $f_k$ et sont les coefficients de l'équation $\prod_{k=1}^{n}(z-x_k^r)=0$.

Par ailleurs rappelons que les $f_k$ sont des nombres entiers réels (*reelle ganze Zahlen*) :

$$\sum e^{(v_1+...+v_k)i} \in \mathbb{R}$$

et donc $f_k$ est une somme de $\frac{n(n-1)...(n-k+1)}{k!}$ cosinus, tous inférieurs à 1 en valeur absolue :

$$|f_k| \leq \frac{n(n-1)...(n-k+1)}{k!}$$

Pour tout $k$, si $|x_{i_1}...x_{i_k}| \leq 1$, alors $|x_{i_1}^r...x_{i_k}^r| \leq 1$, et donc

$$|f_{k,r}| \leq \frac{n(n-1)...(n-k+1)}{k!}$$

Or les coefficients de l'équation $\prod_{k=1}^{n}(z-x_k^r)=0$ doivent être des nombres entiers, il y a donc pour chacun des $n$ coefficients $f_{k,r}$ au plus $2\frac{n(n-1)...(n-k+1)}{k!}+1$ possibilités, il y a donc au plus

$$\mathcal{N} = \prod_{k=1}^{n}\left\{2\frac{n(n-1)...(n-k+1)}{k!}+1\right\}$$

équations de la forme $\prod_{k=1}^{n}(z-x_k^r)=0$. C'est-à-dire que pour certains $r$ distincts, on obtient des équations identiques.

Voilà ce que dit Kronecker de ce dernier raisonnement [43] :

> C'est ici un principe apparemment si trivial et pourtant si important et fondamental à appliquer en mathématiques, qui a notamment été assez souvent utilisé par Dirichlet [44] , et qui dit que si un certain nombre fini de propriétés est considéré et qu'il y a ici plus

43. Es kommt hier das scheinbar so triviale und doch in der Mathematik so wichtige und fundamentale Prinzip zur Anwendung, welches namentlich von *Dirichlet* öfters benutzt wurde und welches aussagt, daß, wenn eine gewisse endliche Anzahl von Eigenschaften betrachtet wird und mehr Gegenstände da sind, die diese Eigenschaften besitzen, nothwendigerweise hinsichtlich dieser Eingenschaften Identitäten entstehen müssen.

44. On appelle aujourd'hui ce principe le *principe des tiroirs de Dirichlet*.

> d'objets qui possèdent ces propriétés, nécessairement des identités doivent se former au niveau de ces propriétés. [Kronecker, 1881a] p. 32

Dans notre cas : Prenons $h\mathcal{N}+1$ de ces équations. Il y en a alors au plus $\mathcal{N}$ distinctes, et on peut en trouver $h+1$ identiques. On en choisit deux identiques :

$$\prod_{k=1}^{n}(z-x_k^{\lambda})=\prod_{k=1}^{n}(z-x_k^{\mu})$$

ainsi pour tout $k\in[\![1,n]\!]$ il existe $k'\in[\![1,n]\!]$ tel que $x_k^{\lambda}=x_{k'}^{\mu}$. ce qui s'écrit aussi :

$$x_1^{\lambda}=x_{k_1}^{\mu}\ ,\ x_2^{\lambda}=x_{k_2}^{\mu}\ ,\ \dots\ ,\ x_n^{\lambda}=x_{k_n}^{\mu}$$

On a ainsi $n!$ permutations possibles. Si l'on choisit $h>n!$, nécessairement une permutation devra apparaitre deux fois, et il existe donc $\mu\neq\mu'$ tel que

$$x_{k_1}^{\mu'}=x_{k_1}^{\mu}\ ,\ x_{k_2}^{\mu'}=x_{k_2}^{\mu}\ ,\ \dots\ ,\ x_{k_n}^{\mu'}=x_{k_n}^{\mu}$$

et donc pour tout $k\in[\![1,n]\!]$ on a $x_k^{\mu'-\mu}=1$. Finalement, on vient de démontrer que les $x_k$ étaient des racines de l'unité, et on a donc la propriété suivante [45] :

> Si une équation à coefficients entiers de degré $n$ n'a que des racines de module 1, alors ces racines doivent nécessairement être des racines de l'unité. Le nombre de ces équations à coefficients entiers est
>
> $$\mathcal{N}=\prod_{k=1}^{n}\left\{2\frac{n(n-1)...(n-k+1)}{k!}+1\right\}.$$
>
> Nous pouvons aussi formuler cette proposition ainsi : Si les racines d'une équation à coefficients entiers se trouvent sur un cercle de rayon 1, alors elles doivent être à une place telle qu'elles aient un rapport rationnel avec la circonférence. [Kronecker, 1881a] p. 33

#### 3.2.2.4 Un exemple de réduction

Le discriminant de l'équation cubique ***§12***

$$x^3-f_1x^2+f_2x-f_3=0$$

est

$$\chi=(x_1-x_2)(x_2-x_3)(x_3-x_1)$$

45. Wenn eine ganzzahlige Gleichung $n$ten Grades lauter Wurzeln hat mit dem Modul 1, so müssen diese Wurzeln nothwendig Wurzeln der Einheit sein. Die Anzahl dieser ganzzahligen Gleichungen ist

$$\mathcal{N}=\prod_{k=1}^{n}\left\{2\frac{n(n-1)...(n-k+1)}{k!}+1\right\}.$$

Wir können den Satz auch so aussprechen : Wenn die Wurzeln einer ganzzahligen Gleichung auf der Kreisperipherie mit den Radius 1 liegen, so müssen sie an solchen Stellen liegen, die ein rationales Verhälteiß zur ganzen Kreisperipherie haben.

1) On a $x_3 = f_1 - x_1 - x_2$, donc

$$\begin{aligned}\chi &= (x_1 - x_2)(x_2 - (f_1 - x_1 - x_2))(f_1 - x_1 - x_2 - x_1)\\ \chi &= (x_2 - x_1)(2x_1 + x_2 - f_1)(2x_2 + x_1 - f_1)\\ \chi &= (x_2 - x_1)(2x_1^2 + 2x_2^2 + 5x_1x_2 - 3f_1(x_1 + x_2) + f_1^2)\end{aligned}$$

2) De plus, on a $0 = F_1(x_2) = x_1^2 + x_1x_2 + x_2^2 - f_1(x_1 + x_2) + f_2$ (voir plus haut), ce qui nous permet d'éliminer les $x_2^2$. À partir de

$$2x_1^2 + 2x_2^2 = 2f_1(x_1 + x_2) - 2x_1x_2 - 2f_2$$

on obtient :

$$\begin{aligned}\chi &= (x_2 - x_1)(3x_1x_2 - f_1(x_1 + x_2) + f_1^2 - 2f_2)\\ &= (x_2 - x_1)(x_2(3x_1 - f_1) - 3f_1(x_1 + x_2) + f_2)\end{aligned}$$

et il reste à éliminer un $x_2^2(3x_1 - f_1)$. Pour cela on ajoute $(f_1 - 3x_1)F_1(x_2) = 0$, c'est-à-dire

$$(f_1 - 3x_1)(x_1^2 + x_1x_2 + x_2^2 - f_1(x_1 + x_2) + f_2)$$

Il apparaît alors un $-3x_1^3$, qui s'éliminera en ajoutant $3F(x_1) = 0$. Ainsi

$$\begin{aligned}\chi &= (x_2 - x_1)(x_2(3x_1 - f_1) - 3f_1(x_1 + x_2) + f_2) + (f_1 - 3x_1)F_1(x_2) + 3F(x_1)\\ &= -6x_1^2x_2 + 2f_1x_1^2 + 4f_1x_1x_2 - 2f_2x_2 + 2(f_2 - f_1^2)x_1 + f_1f_2 - 3f_3\\ &= -2x_2(3x_1^2 - 2f_1x_1 + f_2) + 2f_1x_1^2 + 2(f_2 - f_1^2)x_1 + f_1f_2 - 3f_3\\ &= -2x_2F'(x_1) + 2f_1x_1^2 + 2(f_2 - f_1^2)x_1 + f_1f_2 - 3f_3\end{aligned}$$

On utilise maintenant cela pour exprimer $\chi^2$ en fonction uniquement de $f_1, f_2$ et $f_3$.

On a :

$$\begin{aligned}\chi^2 &= \big(-2x_2F'(x_1) + 2f_1x_1^2 + 2(f_2 - f_1^2)x_1 + f_1f_2 - 3f_3\big)^2\\ &= 4x_2^2F'(x_1)^2 + \big(2f_1x_1^2 + 2(f_2 - f_1^2)x_1 + f_1f_2 - 3f_3\big)^2 - 4x_2[...]\end{aligned}$$

Par ailleurs,

$-3x_1F(x_1) = -9x_1^4 + 9x_1^3f_1 - 9x_1^2f_2 + 9f_3x_1 = 0$ et $3f_1F(x_1) = 3x_1^3f_1 - 3x_1^2f_1^2 + 3f_1f_2x_1 - 3f_1f_3 = 0$, donc :

$$\begin{aligned}F'(x_1)^2 =&(3x_1^2 - 2f_1x_1 + f_2)^2\\ =&9x_1^4 - 12x_1^3f_1 + (4f_1^2 + 6f_2)x_1^2 - 4f_1f_2x_1 + f_2^2\\ &- 9x_1^4 + 9x_1^3f_1 - 9x_1^2f_2 + 9f_3x_1\\ &+ 3x_1^3f_1 - 3x_1^2f_1^2 + 3f_1f_2x_1 - 3f_1f_3\\ =&(f_1^2 - 3f_2)x_1^2 + (9f_3 - f_1f_2)x_1 + (f_2^2 - 3f_1f_3)\end{aligned}$$

et donc

$$\chi^2 = 4x_2^2\left((f_1^2 - 3f_2)x_1^2 + (9f_3 - f_1f_2)x_1 + (f_2^2 - 3f_1f_3)\right) + \left(2f_1x_1^2 + 2(f_2 - f_1^2)x_1 + f_1f_2 - 3f_3\right)^2 - 4x_2[...]$$

Par ailleurs, on a vu précédemment que :

$$0 = F_1(x_2) = x_1^2 + x_1x_2 + x_2^2 - f_1(x_1 + x_2) + f_2$$

donc

$$x_2^2 = -x_1^2 - x_1x_2 + f_1(x_1 + x_2) - f_2 = x_2(f_1 - x_1) - (x_1^2 - f_1x_1 + f_2)$$

Ainsi

$$\begin{aligned}\chi^2 = &-4(x_1^2 - f_1x_1 + f_2)\left((f_1^2 - 3f_2)x_1^2 + (9f_3 - f_1f_2)x_1 + (f_2^2 - 3f_1f_3)\right)\\ &+ \left(2f_1x_1^2 + 2(f_2 - f_1^2)x_1 + f_1f_2 - 3f_3\right)^2 - 4x_2[...]\end{aligned}$$

Or, on a vu que si une fonction symétrique entière de $x_1, x_2, x_3$ se met sous la forme $\Phi x_2 + \Psi$ où $\Phi$ et $\Psi$ ne contiennent ni $x_3$, ni $x_2$, alors dans ce cas $\Phi = 0$.

Et ainsi :

$$\begin{aligned}\chi^2 = &-4(x_1^2 - f_1x_1 + f_2)\left((f_1^2 - 3f_2)x_1^2 + (9f_3 - f_1f_2)x_1 + (f_2^2 - 3f_1f_3)\right)\\ &+ \left(2f_1x_1^2 + 2(f_2 - f_1^2)x_1 + f_1f_2 - 3f_3\right)^2\end{aligned}$$

D'autre part, à partir de $x^3 - f_1x^2 + f_2x - f_3 = 0$, on obtient :

$$\begin{aligned}x^4 &= f_1x^3 - f_2x^2 + f_3x\\ 0 &= f_1x^3 - f_1^2x^2 + f_2f_1x - f_3f_1\end{aligned}$$

et donc

$$x_1^4 = f_1 x_1^3 - f_2 x_1^2 + f_3 x_1$$
$$0 = f_1 x_1^3 - f_1^2 x_1^2 + f_2 f_1 x_1 - f_3 f_1$$

On obtient par soustraction

$$x_1^4 = (f_1^2 - f_2) x_1^2 + (f_3 - f_1 f_2) x_1 + f_1 f_3$$

$$\begin{aligned} -x_2^2.x_1 &= (x_1^2 - f_1 x_1 + f_2) x_1 \\ &= x_1^3 - f_1 x_1^2 + f_2 x_1 \\ &= f_1 x_1^2 - f_2 x_1 + f_3 - f_1 x_1^2 + f_2 x_1 \\ &= f_3 \end{aligned}$$

On poursuit les calculs et on arrive laborieusement à obtenir

$$\chi^2 = f_1^2 f_2^2 + 18 f_1 f_2 f_3 - 27 f_3^2 - 4 f_1^3 f_3 - 4 f_2^3$$

À partir du cours de 1884/85 cet exemple n'est plus traité comme une application de cette forme réduite, mais il devient une illustration de la méthode de Gauss-Waring [46] (nom donné par Kronecker). C'est la méthode qui utilise l'ordre lexicographique et qui sert de base aux démonstrations maintenant classiques du théorème fondamental des polynômes symétriques [47]. On trouve une présentation de cette méthode par Gauss dans [Gauss, 1816a], à l'occasion de sa seconde démonstration du théorème fondamental de l'algèbre et Kronecker la présentera dans ce cours un peu plus loin. Kronecker termine en remarquant que lorsque l'on dit que la forme réduite ne doit pas être nulle, cela ne signifie pas qu'elle ne peut pas s'annuler pour certaines valeurs particulières des $x_k$ et il en donne un exemple.

#### 3.2.2.5 Les « Types » de Vandermonde

Kronecker commence ***§13*** par expliquer que les fonctions symétriques existent sous plusieurs formes, par exemple [48] $\prod_{(i)} \phi(x_{i_1}....x_{i_n})$, $\prod_{(i)}(c_0 + c_1 x_{i_1} + ....c_n x_{i_n})$. Dans son *Mémoire sur la résolution des équations* [49], Vandermonde introduit la notion de *type*, qui sert à distinguer les différents éléments qui composent une fonction symétrique. C'est cette notion que Kronecker reprend ici.

46. Edward Waring (1736-1798), voir en particulier [Waring, 1762]. Nous retrouverons ce mathématicien dans la seconde partie à propos de son travail sur l'équation au carré des différences.
47. Par exemple dans des livres scolaires comme [Gozard, 2009] ou [Calais, 2006b].
48. Il nomme la seconde, qui est de degré $n!$, *équation de Galois*.
49. Voir [Vandermonde, 1771].

On peut classer avec un ordre défini selon les variables toutes les fonctions entières, donc en particulier les fonctions entières symétriques données sous la forme :

$$\sum_{(k)} C_{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n}$$

Les fonctions symétriques doivent rester identiques par permutation des variables, donc pour toute permutation $(i_1, ..., i_n)$ de $(1, ..., n)$, on a :

$$\sum_{(k)} C_{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n} = \sum_{(k)} C_{k_1 \ldots k_n} x_{i_1}^{k_1} \ldots x_{i_n}^{k_n}$$

Si on réécrit à droite les $x_{i_k}$ pour rétablir l'ordre $x_1, x_2, ...x_n$, on obtient :

$$\sum_{(k)} C_{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n} = \sum_{(k)} C_{k_1 \ldots k_n} x_1^{k_{h_1}} \ldots x_n^{k_{h_n}}$$

et donc

$$\sum_{(k)} C_{k_{h_1} \ldots k_{h_n}} x_1^{k_{h_1}} \ldots x_n^{k_{h_n}} = \sum_{(k)} C_{k_1 \ldots k_n} x_1^{k_{h_1}} \ldots x_n^{k_{h_n}}$$

et donc

$$C_{k_{h_1} \ldots k_{h_n}} x_1^{k_{h_1}} \ldots x_n^{k_{h_n}} = C_{k_1 \ldots k_n} x_1^{k_{h_1}} \ldots x_n^{k_{h_n}}$$

Comme $(h_1, ..., h_n)$ est une permutation quelconque de $(1, ..., n)$, tous les coefficients sont en fait égaux, et on a :

$$\sum_{(k)} C_{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n} = C \sum_{(k)} x_1^{k_1} \ldots x_n^{k_n}$$

C'est-à-dire que les polynômes symétriques sont des agrégats de fonctions de la forme

$$C \sum_{(i)} x_{i_1}^{k_1} \ldots x_{i_n}^{k_n}$$

Ces derniers sont nommés *Type* par Vandermonde. Par symétrie, chaque $\sum_{(i)} x_{i_1}^{k_1} \ldots x_{i_n}^{k_n}$ a ainsi $\dfrac{n!}{\alpha_1! \alpha_2! \ldots \alpha_m!}$ termes où $\alpha_1, \alpha_2, \ldots \alpha_m$ comptent les puissances identiques.

#### 3.2.2.6 Théorème fondamental des polynômes symétriques

Le paragraphe suivant (***§14***) est consacré à une démonstration du théorème fondamental des polynômes symétriques. Ainsi, de nouveau, Kronecker veut proposer une méthode « directe » (*direkt*) pour une proposition qu'il a déjà démontrée.

Examinons le double produit :

$$\prod_{k=1}^{n} \prod_{h=1}^{\nu} (1 + x_k \xi_h)$$

On a d'une part

$$\prod_{k=1}^{n}\prod_{h=1}^{\nu}(1+x_k\xi_h)=\prod_{h=1}^{\nu}\prod_{k=1}^{n}(1+x_k\xi_h)=\prod_{h=1}^{\nu}(1+f_1\xi_h+....+f_n\xi_h^n)$$

et d'autre part

$$\prod_{k=1}^{n}\prod_{h=1}^{\nu}(1+x_k\xi_h)=\prod_{k=1}^{n}(1+\varphi_1x_k+....+\varphi_\nu x_k^\nu)$$

où $f_1,...,f_n$ sont comme précédemment les $n$ polynômes symétriques élémentaires des $n$ variables $x_1,...x_n$ et $\varphi_1,...,\varphi_\nu$ les $\nu$ polynômes symétriques élémentaires des $\nu$ variables $\xi_1,...\xi_n$.

Par ailleurs, si on développe le membre de droite de cette seconde égalité, on obtient :

$$\sum_{\alpha,\beta,\gamma...}S_{\alpha,\beta,\gamma...}\varphi_1^\alpha\varphi_2^\beta\varphi_3^\gamma...\quad(\alpha+\beta+\gamma+...\leq n)$$

où les $S_{\alpha,\beta,\gamma...}$ sont les fonctions symétriques dont les termes sont toutes les permutations de :

$$x_1x_2...x_\alpha x_{\alpha+1}^2....x_{\alpha+\beta}^2x_{\alpha+\beta+1}^3...x_{\alpha+\beta+\gamma}^3...$$

On obtient donc une expression génératrice de la plus générale des fonctions symétriques des $x_i$. De même $\prod_{h=1}^{\nu}(1+f_1\xi_h+....+f_n\xi_h^n)$ est une expression génératrice de la plus générale des fonctions symétriques des $\xi_i$.

On suppose maintenant que toutes les fonctions entières symétriques des $\nu$ variables $\xi_i$ est une fonction entière à coefficients entiers des $\nu$ coefficients $\varphi_i$, et on le démontre pour les $x_i$.

Si l'on suppose que $\nu=n-1$, alors Kronecker affirme qu'il est simple de faire une démonstration par induction, le cas $n=1$ et le passage de $n-1$ à $n$ étant faciles.

Il reste seulement à montrer que notre hypothèse $\nu=n-1$ n'a pas de conséquence pour la généralité de notre démonstration. En effet, cette hypothèse suppose que dans toutes les fonctions symétriques de $x_1,...x_n$ qui sont les coefficients de $\varphi_1^\alpha\varphi_2^\beta\varphi_3^\gamma...$ n'apparaît aucune puissance de $x$ supérieure à $n-1$. Ainsi lorsque l'on suppose que « toutes » les fonctions symétriques entières des $n-1$ racines d'une équation de degré $n-1$ sont des fonctions entières à coefficients entiers des coefficients de l'équation, nous avons démontré en fait la validité de cette proposition pour les $n$ racines d'une équation de degré $n$. Mais cette démonstration n'est valable que dans le cas où, dans les fonctions symétriques, les $n$ variables n'apparaissent pas à un degré supérieur à $n-1$. Mais ce cas est le cas général, car si l'on avait une fonction symétrique entière des $n$ racines $x_1,...,x_n$ d'une équation de

degré $n$ avec $x_k$ de degré $n$, on pourrait poser :

$$x_k^n = f_1 x_k^{n-1} - f_2 x_k^{n-2} + .... \pm f_n$$

Passons à la conclusion de l'induction :

Selon l'hypothèse, le produit

$$\prod_{h=1}^{\nu} (1 + f_1\xi_h + .... + f_n\xi_h^n)$$

se transforme en

$$\sum_{\alpha,\beta,\gamma...} V_{\alpha,\beta,\gamma...}\varphi_1^\alpha\varphi_2^\beta\varphi_3^\gamma...$$

où $V_{\alpha,\beta,\gamma...}$ est une fonction à coefficients entiers des $f$. Nous avons donc :

$$\prod_{h=1}^{n-1}(1 + f_1\xi_h + .... + f_n\xi_h^n) = \prod_{k=1}^{n}\prod_{h=1}^{n-1}(1 + x_k\xi_h) = \prod_{k=1}^{n}(1 + \varphi_1 x_k + .... + \varphi_{n-1}x_k^\nu)$$

et donc

$$\sum_{\alpha,\beta,\gamma...} V_{\alpha,\beta,\gamma...}\varphi_1^\alpha\varphi_2^\beta\varphi_3^\gamma... = \sum_{\alpha,\beta,\gamma...} S_{\alpha,\beta,\gamma...}\varphi_1^\alpha\varphi_2^\beta\varphi_3^\gamma...$$

Si nous pouvons montrer que dans cette égalité les termes correspondent un à un (c'est-à-dire que $V_{\alpha,\beta,\gamma...} = S_{\alpha,\beta,\gamma...}$), alors nous aurons démontré que <u>toute</u> fonction symétrique entière des $x_1, ..., x_n$ est une fonction entière à coefficients entiers des $f_1, ...f_n$. Pour cela il ne reste plus qu'à démontrer qu'une fonction entière des fonctions symétriques élémentaires des racines d'une équation ne peut jamais être identiquement nulle et que ces polynômes symétriques élémentaires sont indépendants les uns des autres.

Nous devons donc (***§15***) montrer qu'une fonction

$$\sum_{q_1...q_n} C_{q_1...q_n} f_1^{q_1}...f_n^{q_n}$$

n'est pas identiquement nulle à moins que tous les $C_{q_1...q_n}$ le soient.

Par exemple : Soit les trois fonctions :

$$x^2 + y^2 + y^2\ ,\ x + y + z\ ,\ f(x^2 + y^2 + z^2, x + y + z)$$

que l'on organise selon les variables $x, y, z$, alors on peut difficilement reconnaitre le rapport entre les 3 fonctions, c'est à dire le fait que la troisième s'exprime en fonction des deux autres.

Il n'est donc pas évident que si

$$\sum_{q_1 \ldots q_n} C_{q_1 \ldots q_n} f_1^{q_1} \ldots f_n^{q_n} = 0$$

les $C$ doivent être nuls, même si on sait que ce serait le cas si les $f_i$ étaient remplacées par les $x_i$.

Les $f_i$ sont bien les coefficients d'une équation, et ceux-ci sont à l'évidence indépendants les uns des autres.

Mais en faisant abstraction du fait que jusque-là les $f_i$ ne sont que les coefficients des équations factorisées et que si l'on veut supposer pouvoir dans toute équation (dans laquelle en effet les coefficients sont indépendants) partir des $x_i$ pour trouver les coefficients, c'est-à-dire utiliser les équations :

$$\frac{C_k}{C_0} = x_1 x_2 \ldots x_k + \ldots$$

alors la preuve de l'existence des racines doit être présupposée. Mais cette preuve reposant sur la théorie des fonctions symétriques, il est donc nécessaire de montrer que l'on ne peut pas trouver d'équation liant les $f_i$.

On ordonne $\sum_{q_1 \ldots q_n} C_{q_1 \ldots q_n} f_1^{q_1} \ldots f_n^{q_n} = 0$ selon les puissances de $x_1$. Par ailleurs, on a la propriété suivante $f_k = x_1 g_{k-1} + g_k$ où $g_k$ est le $k^{\text{ième}}$ polynôme symétrique élémentaire de $x_2, \ldots x_n$ et où l'on pose $g_0 = 1$ et $g_n = 0$ (démonstration par récurrence assez simple). On obtient ainsi :

$$\sum_{q_1 \ldots q_n} C_{q_1 \ldots q_n} f_1^{q_1} \ldots f_n^{q_n} = \sum_{q_1 \ldots q_n} C_{q_1 \ldots q_n} (x_1 g_0 + g_1)^{q_1} \ldots (x_1 g_{n-1})^{q_n}$$

Ce qui nous permettra, en raisonnant par récurrence (on a maintenant affaire aux $(n-1)$ fonctions symétriques élémentaires des $n-1$ variables $x_2, \ldots, x_n$) sur le nombre de variable de conclure que les coefficients sont tous nuls.

**Indépendance linéaire de polynômes de degrés différents** La proposition selon laquelle une fonction entière de $z$, c'est-à-dire une fonction linéaire des différentes puissances de $z$ ne peut être nulle à moins que tous ses coefficients ne le soient peut être étendue (***§16***) de la façon suivante : entre des fonctions entières de degrés différents il n'existe pas d'équations linéaires (c'est-à-dire que des fonctions entières de $z$ de degrés différents sont linéairement indépendantes). Soit en effet $\psi_1(z), \psi_2(z), \ldots$ des fonctions entières de $z$ de degrés respectifs $h_1, h_2, \ldots$ où on suppose $h_1 > h_2 > \ldots$. Posons

$$C_1 \psi_1(z) + C_2 \psi_2(z) + \ldots + C_n \psi_n = 0.$$

Alors l'unique coefficient de la plus grande puissance $z^{h_1}$ est $C_1$, donc $C_1 = 0$. On a de même $C_2 = \ldots = C_n = 0$.

**Définition** : Des grandeurs sont dites linéairement indépendantes s'il n'existe pas d'équation linéaire les liant.

On a donc la

**proposition** : Des fonctions entières de degrés différents sont linéairement indépendantes.

Il continue (***§19***) à utiliser cette méthode pour quelques fonctions symétriques particulières, et ensuite (***§20***) établir les *Formules de Newton*. Kronecker affirme que cette formule se trouve dans les travaux de *Girard* de 1633, redécouverte par *Newton*, mais démontrée pour la première fois par *Euler*. Il y a ici une légère confusion au niveau de la date donnée par Kronecker : 1633 est la date de décès de Girard, et l'ouvrage auquel Kronecker fait allusion est *Invention nouvelle en Algebre* ([Girard, 1629]) publié en 1629.

Pour $k$ allant de 1 à $n$ et en posant :

$$x_1^k + ... + x_n^k = s_k$$

on a les « formules de Newton » :

$$s_k - f_1 s_{k-1} + f_2 s_{n-2} - ... \pm n f_n = 0$$

que l'on peut aussi écrire

$$s_{n+r} - f_1 s_{n+r-1} + ... \pm s_r f_n = 0$$

Kronecker propose deux démonstrations de ces formules, l'une utilisant les fonctions $F_1(x, x_h)$ vues précédemment, et la seconde reprenant le travail d'Euler [50]. Il illustre ces formules (***§21*** et ***§22***) en étudiant les nombres de Bernouilli.

Kronecker démontre maintenant (***§23***), tout à fait classiquement, la très célèbre *formule d'interpolation de Lagrange*, et montrer que celle-ci est équivalente à ce qu'il appelle la *formule d'Euler*. Nous ne nous étendrons pas sur cette dernière : son importance est telle que nous préférons l'exposer plus longuement dans la seconde partie au moment de leur utilisation. Donnons-en tout de même une idée :

50. Cette dernière consiste à développer

$$\frac{F'(x)}{F(x)}$$

en série et à montrer que les termes de cette série sont tous nuls à part les $n + 1$ premiers. Voir [Kronecker, 1881a] p. 51.

**Formules d'Euler**

Si les $\xi$ sont les racines (simples) d'un polynôme $f$ de degré $n$, alors

$$\sum_{\xi} \frac{\xi^k}{f'(\xi)} = \delta_{k,n-1}$$

On a donc aussi

$$\sum_{\xi} \frac{F(\xi)}{f'(\xi)} = 0$$

si $\deg(F) < n-1$ et

$$\sum_{\xi} \frac{F(\xi)}{f'(\xi)} = A$$

si $\deg(F) = n-1$ avec $A$ le coefficient de plus haut degré de $F$.

Après quelques applications (***§24***) de ces dernières formules, il poursuit en définissant les fonctions alternées [51] (***§25***). Cela lui permet de construire le déterminant (***§26***) et d'en présenter les propriétés (***§27*** et ***§28***). Il utilise ces propriétés pour étudier la résolution des systèmes d'équations linéaires (***§29***). Il commence ainsi par prendre un système de $n+1$ équations à $n+1$ inconnues, en faisant remarquer que cela ne constitue en rien une limitation des hypothèses étant donné que s'il y a en fait moins d'équations, il suffit de réécrire plusieurs fois l'une d'entre elles. Il démontre (***§30***) que le déterminant, comme fonction homogène, linéaire et alternée d'une ligne horizontale ou verticale est déterminé à un facteur constant près.

Il revient (***§31***) sur un déterminant qui tiendra une place particulièrement importante dans la suite :

$$\prod_{k>h} (x_k - x_h) = |x_h^{k-1}|_{h,k=1,\ldots n}$$

Ce déterminant, que l'on appelle aujourd'hui le déterminant de Vandermonde, est introduit naturellement lors du calcul de discriminants.

Kronecker poursuit (***§32***) en donnant une nouvelle démonstration du fait que les fonctions symétriques élémentaires sont indépendantes les unes des autres et introduit à cette occasion le déterminant fonctionnel (*Funktionaldeterminante* [Kronecker, 1881a] p. 78). Dans les deux derniers paragraphes de cette première partie (***§33*** et ***§34***), Kronecker complète son travail sur la construction de formes réduites.

51. Nom que Kronecker attribue d'ailleurs à Cauchy.

## 3.3 Sur la divisibilité des fonctions entières

Le second chapitre s'intitule [52] :

> Sur la divisibilité des fonctions entières ; Déductions du théorème de Sturm ; conséquences de celui-ci ; Théorie des caractéristiques et preuve de l'existence des racines. [Kronecker, 1881a] p. 86

Les deux derniers points seront l'objet de notre seconde et troisième partie. Nous ne présenterons donc ici que le premier point sur la divisibilité des fonctions entières.

### 3.3.1 Résultant et Discriminant

Du chapitre §35 au chapitre §39, Kronecker développe sa démonstration du fait que deux fonctions entières ont un facteur commun si et seulement si leur résultant est nul. Pour cela, il s'inspire du travail que Gauss a produit dans sa deuxième démonstration du théorème fondamental de l'algèbre [53]

> Cette preuve ingénieuse vient de *Gauss*, qui a élaboré sur les bases de cette proposition démontrée la preuve de l'existence des racines. [Kronecker, 1881a] p. 94

Il donnera aussi dans ces chapitres une description et une démonstration de l'algorithme d'Euclide.

#### 3.3.1.1 Conditions pour avoir un facteur commun

Soit (***§35***)

$$F(x) = (x - x_1)...(x - x_n) = x^n - f_1 x^{n-1} + f_2 x^{n-2} - ... \pm f_n$$
$$\Phi(x) = (x - \xi_1)...(x - \xi_\nu) = x^\nu - \varphi_1 x^{\nu-1} + ... \pm \varphi_\nu$$

deux fonctions entières telles que $m$ des $\xi$ soient égaux à $x_1, ..., x_m$ $(m < n)$. Ces deux fonctions ont donc pour facteur commun

$$(x - x_1)...(x - x_m)$$

Prenons par ailleurs une fonction symétrique entière $\mathcal{G}(x_1, ..., x_m)$. On a :

$$\sum_{(i)} \mathcal{G}(x_{i_1}...x_{i_m})\Phi(x_{i_{m+1}})...\Phi(x_{i_n}) = \mathcal{G}(x_1...x_m)\Phi(x_{m+1})...\Phi(x_n)$$

On peut maintenant poser $\mathcal{G} = (x - x_1)...(x - x_n)$ et on obtient :

$$\sum_{(i)} \prod_{h=1}^{m}(x - x_{i_h}) \prod_{k=m+1}^{n} \Phi(x_{i_k}) = \prod_{h=1}^{m}(x - x_h) \prod_{k=m+1}^{n} \Phi(x_k)$$

52. Ueber die Theilbarkeit ganzer Funktionen ; Ableitung des Sturm'schen Satzes ; Folgerungen aus demselben ; Charakteristikentheorie und Existenzbeweis der Wurzeln.

53. Dieser sinnreiche Beweis rührt von Gauss, der auf den bewiesenen Satz den Existenzbeweis der Wurzeln gründete.

En divisant cette égalité par

$$\sum_{(i)} \prod_{k=m+1}^{n} \Phi(x_{i_k}) = \prod_{k=m+1}^{n} \Phi(x_k)$$

on obtient :

$$\frac{\sum_{(i)} \prod_{h=1}^{h=m}(x - x_{i_h}) \prod_{k=m+1}^{n} \Phi(x_{i_k})}{\sum_{(i)} \prod_{k=m+1}^{n} \Phi(x_{i_k})} = \prod_{h=1}^{m}(x - x_h)$$

Le facteur commun de $F$ et $\Phi$ est donc représenté par le quotient de deux fonctions symétriques de $x_1, ..., x_n$.

On a alors :

$$\sum_{(i)} \prod_{k=2}^{n} \Phi(x_{i_k}) = 0, ..., \sum_{(i)} \prod_{k=m}^{n} \Phi(x_{i_k}) = 0$$

Car dans chaque produit il y a au moins $m - n$ $x_i$, donc au moins l'un d'entre eux annule $\Phi$. Pour que ce facteur commun soit le plus grand, il faut en plus que :

$$\sum_{(i)} \prod_{k=m+1}^{n} \Phi(x_{i_k}) \neq 0$$

Pour que $(x - x_1)...(x - x_m)$ soit le plus grand facteur commun, il y a donc $m$ équations et 1 inéquation à remplir.
Par ailleurs, ce facteur commun a été représenté comme un quotient de fonctions symétriques en $x_1, ..., x_n$. Le numérateur et le dénominateur sont donc des fonctions entières de $f_1, ..., f_n$ et des coefficients $\varphi_i$ de la fonction $\Phi$. Kronecker fait remarquer que le cas $m = 1$ a été donné par *Abel*[54].

Kronecker considère maintenant (***§36***) le cas où il n'y a pas de facteur commun : s'il n'y a aucun facteur commun, on a alors $\prod_{k=1}^{n} \Phi(x_k) \neq 0$. On peut alors écrire :

$$\frac{1}{\Phi(x_k)} = \frac{\Phi(x_1)....\Phi(x_{k-1})\Phi(x_{k+1})...\Phi(x_n)}{\prod_{k=1}^{n} \Phi(x_k)} = \frac{\Psi(x_k)}{R(f_1...f_n; \varphi_1...\varphi_\nu)}$$

On a donc :

$$\Phi(x_k)\Psi(x_k) = \prod_{k=1}^{n} \Phi(x_k) = R(f_1...f_n; \varphi_1...\varphi_\nu) = R(F, \Psi)$$

Pour $x = x_1, ..., x_n$, la fonction $\Phi(x)\Psi(x) - R(F, \Psi)$ s'annule, et elle est donc divisible par $F(x)$. Ainsi il existe une fonction entière $\mathcal{G}(x)$ telle que

$$\Phi(x)\Psi(x) - R(F, \Psi) = -\mathcal{G}(x)F(x)$$

54. C'est-à-dire :

$$(x - x_1) = \frac{\sum_{(i)}(x - x_{i_h}) \prod_{k=2}^{n} \Phi(x_{i_k})}{\sum_{(i)} \prod_{k=2}^{n} \Phi(x_{i_k})}.$$

c'est-à-dire telle que :

$$\Phi(x)\Psi(x) + \mathcal{G}(x)F(x) = R(F, \Psi)$$

où $\Phi, \Psi, F$ sont de degré respectif $\nu, n-1, n$ et il suit que $\mathcal{G}(x)$ doit être de degré $\nu - 1$.

**Résumé :** Kronecker vient donc de démontrer que si $F$ et $\Phi$ n'ont pas de facteur commun, alors $\Phi(x)\Psi(x) + \mathcal{G}(x)F(x) = R(F, \Psi)$ est indépendante de $x$.

Dans le paragraphe suivant (*§37*), Kronecker considère le cas particulier où $\Phi = F'$, et définit le discriminant de Sylvester (*Die Diskriminante von Sylvester*). Il obtient ainsi la condition classique pour l'existence d'une racine double.

#### 3.3.1.2 Décomposition en fractions continues

Kronecker souhaite qu'avant toute chose, nous puissions [55]

> d'abord trouver une méthode à partir de laquelle nous pourrons décider si un facteur commun existe et qui nous apprendra à déterminer un tel facteur. [Kronecker, 1881a] p. 90

C'est-à-dire qu'il justifie (*§38*) la possibilité de parler d'un plus grand facteur commun en exhibant une méthode qui permet de le construire.

On veut déterminer le plus grand facteur commun entre deux fonctions entières $f(x)$ et $f_1(x)$ (on supposera que le degré de $f_1$ est inférieur ou égal à celui de $f$).

On a successivement :

$$\begin{aligned} f(x) &= g_1(x)f_1(x) - f_2(x) \\ f_1(x) &= g_2(x)f_2(x) - f_3(x) \\ &etc... \end{aligned}$$

On obtient ainsi le classique algorithme d'Euclide. Remarquons que Kronecker rapproche cet algorithme de la décomposition en fractions continues et écrit ainsi le quotient $\frac{f_1(x)}{f(x)}$ sous la forme :

$$\frac{f_1(x)}{f(x)} = \frac{1}{g_1 - \frac{f_2}{f_1}} = \cfrac{1}{g_1 - \cfrac{1}{g_2 - \cfrac{1}{g_3 - \cfrac{1}{\ldots \cfrac{1}{g_\nu}}}}}$$

55. Wir müssen nun zunächst eine Methode kennen lernen, nach welcher man entscheiden kann, ob ein gemeinschaftlicher Theiler existirt und welche einen solchen aufzufinden lehrt.

Ce qui est mis en avant par Kronecker, c'est l'aspect *rationnel* du procédé. À partir de là, il lui reste à montrer (***§39***) que si deux fonctions entières n'ont pas de facteur linéaire commun alors leur résultant est non nul (la réciproque vient d'être faite). Et Kronecker insiste sur le point suivant : cette propriété a été démontrée « sans supposer l'existence des racines » [56]. Toute cette partie reproduit en fait le travail de Gauss lors de sa seconde démonstration du théorème fondamental de l'algèbre. Nous verrons que ce dernier produit dans son traité deux preuves de cette propriété : l'une supposant l'existence des racines, l'autre pas [57].

### 3.3.2 Décomposition d'une fonction entière en facteurs

#### 3.3.2.1 Propriétés préliminaires

Kronecker donne ici, par analogie avec la décomposition d'un nombre en produit de facteurs premiers, une décomposition pour les fonctions entières à coefficients entiers [58] :

> Nous étudions maintenant la décomposition d'une fonction entière à coefficients entiers en facteurs. Cette étude importante est analogue à la décomposition d'un nombre en facteurs premiers. [Kronecker, 1881a] pp. 94-95

Kronecker commence (***§40***) par démontrer ce théorème :

Théorème : Si le produit $\Psi(x)\chi(x)$ de deux fonctions entières à coefficients entiers est divisible par le nombre premier $p$ et que $\chi$ n'est pas divisible par $p$, alors $\Psi$ doit être divisible par $p$.

dont il attribue à Gauss le mérite d'avoir été le premier à démontrer ce théorème « simple » [59].

Kronecker poursuit en montrant :

Théorème : Si une fonction entière à coefficients entiers $\Phi(x)$ est décomposable en fonctions entières à coefficients rationnels, alors elle est aussi décomposable en facteurs entiers à coefficients entiers.

On peut donc se contenter d'examiner les fonctions entières à coefficients entiers.

56. [Kronecker, 1881a] p. 95 : « ohne die Existenz der Wurzeln vorauszusetzen ». Il le dit d'ailleurs dès le début de la démonstration : « sans qu'il soit supposé que $A(x)$ et $B(x)$ sont décomposables en facteurs linéaires » (*(...) nicht vorausgesetzt werden, daß $A(x)$ und $B(x)$ in linearfaktoren zerlegbar seien*).

57. Nous regarderons précisément ces preuves dans notre troisième partie.

58. Wir untersuchen nun die Zerlegbarkeit ganzer ganzzahliger Funktionen in Faktoren. Diese wichtige Untersuchung ist analog der Zerlegung einer Zahl in ihre Primfaktoren.

59. [Kronecker, 1881a] p. 96 : « Aussi simple que cette proposition semble être, on ne la trouve pour la première fois étonnamment que chez *Gauss* » (*So einfach dieser Satz zu sein scheint, so findet er sich merkwürdigerweise doch erst bei Gauss.*)

#### 3.3.2.2 Méthode de décomposition

Kronecker propose ensuite un algorithme de décomposition des fonctions entières qu'Edwards expose dans deux de ses ouvrages ([Edwards, 1990] et [Edwards, 2005]). Il commence par établir la méthode de décomposition pour les fonctions entières d'une variable. C'est seulement une fois cette méthode exposée que l'on est « autorisé à établir le concept d'indécomposabilité ou d'irréductibilité » [60]. Il remarque qu'il est toujours possible de résoudre la question grâce à un système d'équations linéaires [61].

Mais cette méthode pose pour Kronecker deux problèmes : elle est « trop compliquée et suppose ensuite la résolution d'un système d'équations » [62]. Il expose donc une autre méthode qui consiste en un algorithme de décomposition des fonctions entières. Cette méthode, que nous n'allons pas décrire ici, utilise les fonctions $F(x, x_k)$ définies précédemment.

Maintenant, Kronecker est autorisé à parler de fonctions irréductibles [63] :

> Après que nous avons obtenu une méthode qui permet dans chaque cas de décider si une fonction entière à coefficients entiers de $x$ donnée a des diviseurs entiers à coefficients entiers, nous sommes autorisé à établir le concept d'indécomposibilité ou d'irréductibilité. [Kronecker, 1881a] p. 98

On voit ici l'un des exemples les plus représentatifs de l'exigence de Kronecker d'avoir un algorithme effectif pour montrer l'existence d'une notion. On ne peut parler d'un polynôme irréductible seulement parce qu'on a un moyen, si cela est possible, de décomposer un polynôme. La réductibilité effective est la condition pour la *possibilité* de la notion d'irréductibilité.

Il démontre maintenant l'unicité de la décomposition de la forme [64] :

60. [Kronecker, 1881a] p. 98. Voir la citation suivante.

61. Soit la fonction

$$\Phi(x) = \varphi_0 + \varphi_1 x + \varphi_2 x^2 + ...$$

On cherche les nombres entiers $\Psi_0, \Psi_1, ..., \chi_0, \chi_1, ...$ tels que :

$$(\Psi_0 + \Psi_1 x + \Psi_2 x^2 + ...)(\chi_0 + \chi_1 x + \chi_2 x^2 + ...) = \varphi_0 + \varphi_1 x + \varphi_2 x^2 + ...$$

On peut déterminer ces entiers, s'ils existent, par le système :

$$\begin{aligned} \Psi_0\chi_0 &= \varphi_0 \\ \Psi_0\chi_1 + \Psi_1\chi_0 &= \varphi_1 \\ \Psi_0\chi_2 + \Psi_1\chi_1 + \Psi_2\chi_0 &= \varphi_2 \\ ... &= ... \end{aligned}$$

On reconnait la méthode des coefficients indéterminés mise en place par Descartes dans *La géométrie* ([Descartes, 1987]), essentiellement dans le cas des équations biquadratiques.

62. [Kronecker, 1881a] p. 96 : « Aber diese Methode ist erstens sehr umständlich und setzt sie die Auflösbarkeit eines Systems von Gleichungen voraus. »

63. Nachdem wir nun eine Methode gewonnen haben, die in jedem einzelnen Falle erlaubt, zu entscheiden, ob eine vorgelegte ganze ganzzahlige Funktion von $x$ ganze ganzzahlige Divisoren hat, sind wir berechtigt, den Begriff der Unzerlegbarkeit oder Irreduzibilität aufzustellen.

64. Jede ganze ganzzahlige Funktion läßt sich also jetzt darstellen in der Form :

$$F(x) = p_1^{h_1} p_2^{h_2} ... p_\mu^{h_\mu} f_1^{k_1} f_2^{k_2} ... f_\nu^{k_\nu},$$

> Chaque fonction entière à coefficients entiers peut se représenter désormais par la forme :
>
> $$F(x) = p_1^{h_1} p_2^{h_2} ... p_\mu^{h_\mu} f_1^{k_1} f_2^{k_2} ... f_\nu^{k_\nu},$$
>
> où $p_1 ... p_\mu$ sont des nombres premiers, $f_1, ..., f\mu$ des fonctions irréductibles et $h_1 ... h_\mu, k_1 ... k_\nu$ des nombres entiers positifs. De la même façon que pour les nombres, on a la proposition disant que cette décomposition est unique. [Kronecker, 1881a] pp. 98-99

Cette proposition est démontrée grâce au théorème suivant :

Théorème : Lorsque le produit $\Phi.\Psi$ de deux fonctions entières de $x$ à coefficients entiers $\Phi$ et $\Psi$ est divisible par une fonction irréductible $F(x)$, alors soit $\Phi$ soit $\Psi$ doit être divisible par $F(x)$.

À partir de son travail sur les fonctions entières à coefficients entiers, Kronecker étend par induction (*§42*) sa méthode de décomposition aux fonctions de plusieurs variables. Il applique ensuite son travail au discriminant de deux fonctions $F$ et $\Phi$ et au résultant de $F$, et montre que ces derniers sont irréductibles. Kronecker, toujours sensible à ce type de question, termine cette partie sur la décomposition des fonctions entières en montrant que « le résultant est un invariant » [65]

## 3.4 Conclusion

Nous venons donc de parcourir l'ensemble des résultats que Kronecker présente avant d'aborder le théorème de Sturm, en nous centrant sur le manuscrit de 1880-81. Les propriétés exposées sont pour la plupart « classiques », tout du moins depuis le travail de Gauss, même si, comme nous l'avons vu pour l'équation cubique, l'exposition peut renvoyer aux préoccupations les plus profondes de Kronecker. Dans la seconde partie du cours, à partir du théorème de Sturm, apparaissent de façon plus systématique et plus directe des références à ses articles de recherche.

Nous pouvons observer une présentation qui va du singulier au général, et où les hypothèses sur les objets qu'il manipule n'apparaissent qu'au fur et à mesure des besoins. Il commence ainsi par les équations quadratiques, puis les équations cubiques avant de passer au cas $n$. Son travail sur les équations de degré 2 et 3 est d'ailleurs, on l'a vu, l'occasion pour Kronecker de construire en petite dimension un objet qui joue un rôle essentiel dans la suite de son cours : le discriminant. Tout aussi importante pour lui est

---

wo $p_1 ... p_\mu$ Primzahlen, $f_1, ..., f\mu$ irreduzible Funktionen und $h_1 ... h_\mu, k_1 ... k_\nu$ ganze positive Zahlen sind. Es gilt nun wie bei der Zerlegung der Zahlen der Satz, daß diese Zerlegung eine eindeutige ist.

65. Es möge hier noch der Beweis hinzugefügt werden, daß die Resultante eine Invariante ist, d. h. ungeändert bleibt, wenn wir in den Funktionen $F(x)$ und $\Phi(x)$ für $x$ die gebrochene lineare Funktion $\frac{ax+b}{cx+d}$ setzen, die Funktionen ganz machen und dann aus den Coeffizienten die Resultante bilden. [Kronecker, 1881a] p. 103.

la présentation de la résolvante de Lagrange – essentiel pour comprendre le travail de Galois – dont l'esprit est finalement proche de ce qu'il fera lorsqu'il produira son *théorème fondamental de l'arithmétique*[66].

Dans cette exposition, souvent très technique, comme par exemple sa méthode de réduction[67], ou quelquefois presque entièrement conceptuelle comme lorsqu'il énumère les différents « types » d'égalité que l'on peut utiliser dans les mathématiques, Kronecker propose souvent plusieurs explications à un même phénomène, plusieurs démonstrations à un même théorème. Cela lui permet de multiplier ses approches et de les comparer. L'exemple du théorème fondamental des polynômes symétriques est particulièrement éclairant, car il explicite les raisons qui l'amènent à le redémontrer : tout d'abord pour éviter de supposer l'existence de racines et ensuite pour fournir un algorithme permettant d'exprimer tout polynôme symétrique en fonction des polynômes symétriques élémentaires. Cette pratique qui consiste à aborder sous plusieurs angles une même notion, à ne pas « borner » le champ des méthodes possibles, est un exemple supplémentaire de, comme nous l'avons déjà exprimé plus haut, la « porosité de son algèbre »[68].

Pour continuer à tenter de dégager l'originalité de l'approche de Kronecker, nous allons maintenant donner quelques éléments de comparaison entre ces *Vorlesungen* et le fameux *Lehrbuch* de Heinrich Weber.

66. Voir [Kronecker, 1887c].
67. Il y aurait beaucoup à dire sur cette méthode, qui mériterait une étude bien plus approfondie que ce que nous avons fait dans ce chapitre.
68. Voir p. 74.

# Chapitre 4

# Les *Vorlesungen* de Kronecker et le *Lehrbuch* de Heinrich Weber


4.1 Weber lecteur de Kronecker . . . . . **152**
4.2 Présentation du *Lehrbuch der Algebra* . . . . . **155**
4.2.1 Le *Lehrbuch der Algebra* . . . . . 155
4.2.2 Le premier Volume . . . . . 156
4.3 Comparaison . . . . . **157**
4.3.1 La structure . . . . . 157
4.3.2 Des *Algèbres* différentes ? . . . . . 159
4.3.3 Les deux introductions . . . . . 160
4.4 La notion de Racine . . . . . **160**
4.4.1 Introduction . . . . . 160
4.4.2 Le théorème fondamental de l'algèbre . . . . . 162
4.5 Un exemple : le théorème de Sturm . . . . . **163**
4.5.1 Introduction . . . . . 163
4.5.2 Les présentations de Weber et Kronecker . . . . . 164
4.5.3 Une extension : la théorie des caractéristiques . . . . . 166
4.6 Conclusion . . . . . **167**


Nous venons de présenter la *théorie des équations algébriques*, et donc ce qui constitue en fait pour Kronecker son cours d'« algèbre », même si celui-ci, comme nous l'avons vu, sera de plus en plus pensé, année après année, comme une partie du *tout* que constitue l'*Arithmétique générale*. Pour mieux en saisir l'originalité, mais aussi ce qui en fait un cours parfaitement ancré dans la seconde moitié du XIX$^{e}$ siècle, nous allons le comparer au *Lehrbuch der Algebra* de Weber. Pour rendre notre comparaison la plus pertinente possible, nous utiliserons essentiellement le manuscrit du semestre d'hiver 1890-91 des *Vorlesungen über die Theorie der algebraischen Gleichungen* de Kronecker. C'est en effet le plus contemporain de la publication du premier tome du *Lehrbuch*. Ce chapitre sera aussi pour nous l'occasion de rencontrer pour la première fois les deux théories qui seront au cœurs des deux parties suivantes : le théorème de Sturm et la théorie des caractéristiques.

## 4.1 Weber lecteur de Kronecker

Commençons par situer Weber dans la communauté mathématique de son temps, et par rapport à Kronecker en particulier. Weber appartient à la génération qui suit celle de Kronecker : il naît en 1842 lorsque Kronecker est déjà étudiant à l'université de Berlin. Issu d'une famille dans laquelle l'éducation avait une place primordiale [1], Weber commence ses études à l'université d'Heidelberg en 1860 et obtient son doctorat en 1863. Il aura alors [2] comme professeurs Gustav Kirchhoff (1824-1887), Otto Hesse (1811-1874) et Moritz Cantor (1829-1920). Les deux premiers sont d'anciens étudiants de Jacobi, et il n'est donc pas étonnant que, sur leurs conseils, Weber poursuive ses études à Königsberg. Ce sera sous la direction de Franz Neumann (1798-1895) et Friedrich Julius Richelot (1808-1875) qu'il passera sa thèse d'habilitation [3] en 1866 et il deviendra *Privatdozent* la même année à l'université de Heidelberg. Neumann et Richelot ont eux aussi étudié les mathématiques avec Jacobi, et si Weber ne l'a jamais croisé, ce dernier a tout de même eu une influence considérable sur sa formation.

On sait que Kronecker, lui, après une retraite de dix ans pour régler des affaires familiales (1845-1855), retourne à la vie universitaire et commence à enseigner en 1862 à l'université de Berlin : Weber aurait pu faire partie des premiers étudiants de Kronecker. Ainsi, malgré une différence d'âge de vingt ans, ils commencent donc leur carrière d'enseignant à peu près à la même époque. Si l'enseignement de Kronecker sera essentiellement tourné vers l'algèbre et l'arithmétique, celui de Weber au contraire s'étendra à tous les domaines des mathématiques. Autre différence notable : alors que Weber sera le mathématicien allemand du XIX$^{e}$ siècle qui « changea le plus souvent de poste [4] », Kronecker au contraire passera l'ensemble de sa carrière universitaire à Berlin, allant même jusqu'à refuser en 1868 un poste de professeur à Göttingen, statut qu'il n'obtiendra à Berlin, nous le rappelons, qu'en 1883.

Par ailleurs, deux des raisons pour lesquelles on connaît généralement Kronecker sont liées à Weber. Il y a tout d'abord la fameuse citation [5] :

> Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.
> [Weber, 1893b] p. 15

dont la seule preuve d'authenticité est le témoignage de Weber en 1893 dans le *Jahresbericht der Deutschen Mathematiker-Vereinigung*.

Il y a ensuite le non moins célèbre théorème de Kronecker-Weber, qui affirme que toute extension abélienne du corps $\mathbb{Q}$ des nombres rationnels est un sous-corps du corps

1. Son père, Georg Weber (1808-1888), était historien et directeur d'une *höheren Schule* à Heidelberg.
2. Durant ses trois ans à Heidelberg, il passera aussi un semestre à l'université de Leipzig.
3. *Singuläre Auflösungen partieller Differentialgleichungen erster Ordnung.*
4. [Schappacher et Volkert, 1997] p. 2.
5. D'après Weber, cette phrase aurait été prononcée lors d'une conférence à la *Berliner Naturforscher-Versammlung* en 1886.

cyclotomique $\mathbb{Q}(\xi_m)$ où $\xi_m$ est une racine $m^{\text{ième}}$ de l'unité [6].

Weber, qui n'a pas étudié à Berlin mais à Heidelberg, Leipzig et Königsberg, n'a donc pas été un étudiant de Kronecker, et comme il le dit lui-même [7] :

> Je n'ai pas suivi en tant qu'étudiant les cours de Kronecker, et si j'ai aussi beaucoup étudié les écrits du maître et que je lui dois beaucoup, je ne suis cependant pas pénétré également de toutes ses recherches si diverses, et je doute qu'il me soit possible d'en apprécier correctement leur universalité. ([Weber, 1893b])

Mais s'il n'a pas suivi les cours de Kronecker, il l'a beaucoup lu, et il l'a régulièrement rencontré et a pu à plusieurs reprises s'entretenir avec lui lors de réceptions dans sa maison de Bellevue-Strasse.

Les deux citations précédentes sont tirées du *Festschrift* en l'honneur de Kronecker que Weber a publié en 1893. Pourtant, comme le note Norbert Schappacher [8], Weber est principalement lié au réseaux des mathématiciens du « Sud de l'Allemagne », qui se caractérisent plus par une opposition à l'école Berlinoise que par leur origine géographique. Et de plus, cette nécrologie est publiée dans les *Mathematische Annalen*, qui est la revue de cette « École du Sud ». Alors que dans le journal de Kronecker, le *Journal für die reine und angewandte Mathematik*, Fuchs n'écrit qu'une demi-page lors de sa mort, Weber publie lui une longue nécrologie de presque trente pages montrant une connaissance et une admiration certaine de l'homme et de son travail. Cette appartenance à des réseaux opposés à l'université à laquelle appartient Kronecker ne semble pas avoir influencé de manière décisive le regard de Weber sur le travail de ce dernier.

Examinons maintenant les références citées par Weber dans le premier tome du *Lehrbuch* [9] :

6. Voir plus haut p.114.

7. Ich habe nicht als Student zu Kronecker's Füssen gesessen, und wenn ich auch die Schriften des Meisters viel studirt und ihm viel zu verdanken habe, so bin ich doch nicht in alle Seiten seiner so vielseitigen Forschung gleichmässig eingedrungen, und ich zweifle, ob es mir gelingen wird, sie allseitig richtig zu würdigen.

8. Voir [Schappacher et Volkert, 1997].

9. Ce travail reste imparfait du fait des problèmes techniques liés à sa réalisation (OCR). Nous avons par ailleurs réalisé un tri : une occurrence sera supprimée s'il s'agit par exemple du nom d'un principe ou d'un théorème, mais conservée s'il s'agit d'une « vraie » référence.

Références données par Weber dans le premier tome du *Lehrbuch der Algebra*

| | | | |
|---|---|---|---|
| Galois 77 | Bézout 6 | Hesse 2 | Heine 1 |
| Gauss 55 | Cantor 4 | Clebsch 2 | Lipschitz 1 |
| Sturm 40 | Descartes 4 | Di Bruno 2 | Waring 1 |
| Kronecker 26 | Weierstrass 3 | Gordan 2 | Baur 1 |
| Newton 19 | Cauchy 3 | Dyck 2 | Dziobkk 1 |
| Jacobi 15 | Eisenstein 3 | Laplace 2 | Harriot 1 |
| Lagrange 15 | Bernoulli 3 | Hurwitz 2 | Cohn 1 |
| Hermite 14 | Jordan 3 | Laguerre 2 | Gräffe 1 |
| Budan 14 | Hesse 3 | Holder 2 | Enke 1 |
| Dedekind 13 | Euclide 3 | Hilbert 1 | Mehmke 1 |
| Fourier 11 | Meyer 3 | Borchard 1 | Degen 1 |
| Abel 10 | Gundelfinger 3 | Hensel 1 | Liouville 1 |
| Klein 9 | Kummer 2 | Betti 1 | Dupuy 1 |
| Frobenius 7 | Serret 2 | Sylow 1 | Maser 1 |
| Cayley 7 | Netto 2 | Petersen 1 | Schönemann 1 |
| Sylvester 7 | Brioschi 2 | Study 1 | Bachmann 1 |
| Euler 6 | Rolle 2 | Von Staudt 1 | Mertens 1 |
| Hurwitz 6 | Legendre 2 | Arndt 1 | Ruffini 1 |
| | Horner 2 | Wiener 1 | Buhrkhardt 1 |
| | Dirichlet 2 | Mehmke 1 | Runge 1 |

Si l'on regarde un peu plus précisément ces résultats, on peut faire les remarques suivantes. Tout d'abord le grand nombre d'occurrences pour Galois, Sturm et Newton s'explique par la présentation de leur travaux : ce sont donc des références bien localisées dans des chapitres précis.

Ensuite, on voit se dessiner un réseau qui est très loin de se limiter à son entourage (celui que Norbert Schappacher appelle les « mathématiciens du sud »), ni à l'Allemagne, mais qui semble au contraire s'étendre à toute l'Europe. On retrouve ici certes cette universalité si souvent décrite chez Weber, mais certainement aussi un exemple parmi d'autres des très nombreux échanges qui avaient lieu dans l'Europe des mathématiques de cette fin du XIX$^{e}$ siècle.

Enfin, si Gauss est bien sûr extrêmement présent, on voit Kronecker apparaître juste après. On peut donc affirmer sans risque de se tromper, que le travail de Kronecker a influencé la rédaction du *Lehrbuch* de Weber.

Car c'est bien en effet dans ce sens que semble avoir lieu la connexion : l'influence de

Kronecker sur le travail de Weber.

Une dernière remarque : en 1892, après la mort soudaine de Kronecker, une commission s'organise pour lui chercher un remplaçant. Même si, à la lecture du compte rendu [10] de la séance, Weber n'a pas l'air d'avoir déclenché des passions, il sera tout de même finalement classé second, après Schwarz : on a donc pu voir en Weber un successeur potentiel de Kronecker.

Nous allons maintenant présenter le *Lehrbuch der Algebra*, ce qui nous permettra ensuite d'en comparer la structure avec celle des *Vorlesungen*. Ensuite, sur quelques points particuliers, nous essayerons de mettre en évidence ce qui les rend à la fois proches et différents.

## 4.2 Présentation du *Lehrbuch der Algebra*

La période s'étendant de la fin du XIX$^{e}$ siècle au début du XX$^{e}$ siècle est, on le sait, une période charnière pour l'algèbre, qui subit d'importants changements, à la fois dans ses techniques et ses objets, et on sait aussi que l'un des moteurs de ces changements sera la diffusion des idées de Galois à partir de leur publication par Liouville en 1846.

Deux des manuels d'algèbre les plus utilisés au cours du XIX$^{e}$ siècle sont le *Cours d'algèbre supérieure* de Joseph Serret (publiés en 1849-1854-1866) et le *Traité des substitutions et des equations algébriques* de Camille Jordan (publié en 1870). Le premier incorporera, au fur et à mesure de ses rééditions, les résultats de Galois, et le second présentera une théorie des groupes assez complète. Cependant, ces questions ne seront étudiées dans ces deux ouvrages que dans le cas de la théorie des équations algébriques, et on peut considérer que l'un des premiers manuels qui sera centré sur les structures algébriques abstraites (groupes, corps, anneaux) est le *Moderne Algebra* de L. van der Waerden.

### 4.2.1 Le *Lehrbuch der Algebra*

Le *Lehrbuch der Algebra*, que Weber publie en trois tomes entre 1895 et 1908, se situe entre ces deux traditions. Il contient, en particulier dans les tomes 2 et 3, de nombreuses pistes sur ce que deviendront les outils et les objets de l'algèbre du XX$^{e}$ siècle, mais en reprenant essentiellement les idées développées au XIX$^{e}$ siècle. Dans sa structure et sa présentation, il s'agit réellement d'un texte du XIX$^{e}$ siècle.

Pourtant, lorsqu'il publie le premier tome du *Lehrbuch*, Weber est parfaitement familier avec les dernières avancées en algèbre, auxquelles il a lui-même participé, par exemple lors de ses recherches en collaboration avec Dedekind sur la notion d'idéal, et il est ainsi capable de décrire ces nouveaux objets en terme de structure. En 1893, il publie

10. Voir [Biermann, 1973] pp. 305-306.

d'ailleurs un article [11] dans lequel il donne, et utilise, les définitions de groupe et de corps. Il s'agit donc bien d'un *choix*, lorsqu'il décide de ne réserver dans son cours d'algèbre qu'un rôle marginal à ces structures.

Nous allons essentiellement regarder le premier tome. Présentons donc en quelques mots le contenu des tomes 2 et 3. Le second volume commence par une définition abstraite de la notion de groupe et des premières propriétés de cet objet. Dans une seconde partie il regarde certains groupes particuliers (groupe des substitutions linéaires, groupes polyédriques,...) et présente certaines applications (invariant, théorie de Galois,...). Il termine par une théorie des nombres algébriques. Le dernier volume, publié en 1908, est en fait à peu de choses près une extension de son livre écrit en 1891 sur les fonctions elliptiques.

### 4.2.2 Le premier Volume

#### 4.2.2.1 Le contenu

Ce premier tome se constitue de 3 livres. Le premier, intitulé *les principes*, est essentiellement une somme de prérequis à avoir pour aborder les autres parties (Arithmétique, polynômes et fractions rationnelles, déterminant, équations quadratiques et cubiques, fonctions symétriques, transformation). Il contient de plus une présentation de la théorie des invariants inspirée du travail de Clebsch et de ses propres travaux, c'est-à-dire l'exposé de théories à la pointe de ce qu'il appelle les « mathématiques modernes ». Il est intéressant de noter que cette première partie est résumée par Weber lui-même par le terme de « Calcul littéral ».

La seconde partie est entièrement centrée autours des *racines*, et en particulier sur les racines réelles. Seul le dernier chapitre sur les racines $n^{\text{ièmes}}$ de l'unité envisage essentiellement les racines complexes.

Enfin, la dernière partie est une présentation de la théorie de Galois et de ses applications, en particulier à la résolution des équations.

#### 4.2.2.2 La préface

Dans la préface de la première édition, Weber justifie l'écriture de cet ouvrage, et en désigne les buts.

1. Il s'agit d'un ouvrage issu de ses leçons à l'université : Weber souhaitait publier le fruit de longues années d'enseignement.
2. Il était nécessaire de refaire un ouvrage qui expose complètement l'algèbre, à la vue des avancées des dernières décennies.
3. Il veut participer à la mise au clair du vocabulaire récemment introduit en algèbre.

11. Voir [Weber, 1893a].

4. Il veut donner une présentation de l'algèbre allant du plus élémentaire aux dernières avancées.

Weber exprime donc assez clairement sa double volonté de faire le point sur ce qui a jusque-là constitué l'algèbre et d'ouvrir sur les dernières recherches de cette « science en voie de développement rapide, et sur laquelle on travaille des côtés les plus divers » [12]. Après cette présentation rapide du *Lehrbuch* de Weber, donnons maintenant quelques éléments de comparaison de ces deux cours.

## 4.3 Comparaison

### 4.3.1 La structure

En 1975, Van der Waerden publie un article dans lequel il revient sur les sources de son *Moderne Algebra* :

> I studied mathematics and physics at the University of Amsterdam from 1919 to 1924. A very nice course in classical algebra was given by Hendrik de Vries. It included subjects like :
>
> Determinants and linear equations,
> Symmetric functions,
> Resultants and discriminants,
> Sturm's theorem on real roots,
> Sylvester's "index of inertia" for real quadratic forms,
> Solution of cubic and biquadratic equations by radicals.
>
> I supplemented this course by studying Galois theory and other subjects in Heinrich Weber's admirable three-volume textbook on algebra. [Van Der Waerden, 1975] pp. 31-32

On peut immédiatement remarquer deux choses : premièrement, nous y reviendrons, la structure du cours est étonnamment proche de celles des deux cours que l'on étudie qui lui sont pourtant antérieurs de trente ans. Ensuite on voit que la première référence qu'il donne comme ouvrage de formation est bien le *Lehrbuch der Algebra* de Weber, ce qui confirme, comme on l'a vu précédemment, que cet ouvrage a un rôle charnière entre l'algèbre considérée comme *théorie des équations algébriques* et l'algèbre abstraite décrite par le manuel de Van der Waerden. Voyons maintenant la structure des deux cours.

**La table des matières** Regardons les grands chapitres du *Lehrbuch* (volume 1) et des *Vorlesungen*. Pour plus de clarté, nous avons résumé, fusionné ou même éliminé les titres de plusieurs chapitres de ces dernières, mais nous avons joint en annexe la totalité du sommaire [13].

12. Voir [Weber, 1898] p. 8.
13. Voir p. 663.

***Lehrbuch der Algebra***

*Livre I : Principes*

Fonctions rationnelles
Déterminants
Racines des équations algébriques
Fonctions symétriques
Transformations linéaires – invariants
transformation de Tschirnhaus

*Livre II : Racines*

Réalités des racines
Théorème de Sturm
Encadrement du nombre et des valeurs des racines
Approximations des racines
Fractions continues
Racine de l'unité

*Livre III : Quantités algébriques*

Théorie de Galois
Application des groupes de permutation aux équations
Équations cycliques
Divisions du cercle en parts égales
Résolution algébriques des équations
Racines des équations méta-cycliques

**les *Vorlesungen über die Theorie der algebraischen Gleichungen***

*Partie I Introduction*
Domaine de rationalité et d'intégrité
Quelques propriétés préliminaires des équations (notion de racine)

*Partie II Théorie des équations quadratiques*

Résolution et discussion des équations quadratiques
La réduction des fonctions de deux variables
Fonctions symétriques de deux variables et digression sur la théorie des fractions continues

*Partie III Théorie des équations cubiques*
Résolution
Classification des équations cubiques selon le nombre de leurs racines réelles
Fonctions rationnelles de trois variables : cas particulier des fonctions cycliques et des équations abéliennes

*Partie IV Théorie des équations algébriques de degré n*
Réduction et calcul approché des racines
Fonctions symétriques de $n$ variables
Discriminant et invariants
Formules de *Newton* et d'*Euler*
Théorème de *Sturm*

*Partie V Sur la théorie des caractéristiques des systèmes de fonctions*
*Partie VI Théorie du genre*
*Partie VII Sur les équations algébriques résolubles* (Équation *abélienne)*
*Partie VIII Supplément sur la congruence fondamentale de l'arithmétique générale*

On voit déjà que l'ensemble des chapitres qui seront abordés dans le cours que Van der Waerden a suivi à l'université vers 1920 est présent, d'une façon ou d'une autre, dans les deux cours. Ainsi, leur contenu n'est pas éloigné de ce que l'on enseignera encore une trentaine d'années plus tard. De plus, dans les deux cas, une place extrêmement importante est donnée aux racines des équations.

Weber semble procéder systématiquement de la façon suivante : il commence toujours par donner les définitions, les propriétés les plus générales, puis développe les exemples, les cas particuliers. Il s'agit donc d'une démarche *synthétique*. Par exemple, il fournit la définition et les méthodes de calcul des déterminants dans le cas $n$ avant de traiter certains cas particuliers. Au contraire Kronecker commence par explorer longuement des cas particuliers, ou les cas des petites dimensions avant de généraliser à la dimension $n$. Ainsi, il traite longuement les équations quadratiques et cubiques, sur lesquelles il développe tout un arsenal d'outils : polynômes symétriques, équations abéliennes, etc. Il procède donc plutôt par *généralisation*.

Si le développement du cours est différent, les contenus semblent eux très proches, à l'exception notable du dernier tiers de chacun d'entre eux. D'un côté Weber développe une théorie de Galois, pas tellement éloignée dans sa structure de la façon dont on a pu la présenter au cours du XX$^{e}$ siècle. Cette théorie semble absente du cours de Kronecker, tout au moins si l'on se contente d'un examen rapide de la table des matières. Ce n'est en fait pas du tout le cas : la théorie des genres (*Theorie der Gattungen*), qui est exposée à la place de la théorie de Galois dans le cours de Kronecker remplit la même fonction. En fait, la théorie de Galois, au sens où la voit Kronecker, découle de sa théorie des genres [14].

### 4.3.2 Des *Algèbres* différentes ?

Nous avons déjà vu ce que Kronecker dit dans ses leçons de l'algèbre [15]. Dans ce premier volume, Weber termine l'introduction en donnant quelques pistes sur sa vision de l'algèbre :

> Il nous reste à dire un mot du calcul littéral, le procédé le plus important de l'Algèbre. Son emploi est si général qu'on le prend souvent comme synonyme d'Algèbre. [Weber, 1898] p. 23

et en parlant des égalités, et en particulier des valeurs numériques qui peuvent se substituer aux variables, il ajoute que

> L'Algèbre a pour but de déterminer ces valeurs numériques, de *résoudre ces équations*. [Weber, 1898] p. 24

L'algèbre a donc pour Weber un but, résoudre les équations, et un procédé, le calcul littéral. Nous sommes encore une fois bien loin de ce qui caractérisera l'algèbre au XX$^{e}$ siècle.

Pourtant, dans la préface de la première édition, il note bien que :

14. On pourra voir par exemple les chapitre 11 et 12 des *Grundzüge einer arithmetischen Theorie der algebraischen Grössen* ainsi que le commentaire qu'Erwan Penchèvre en fait dans [Penchèvre, 2007]. Un important travail de clarification, trop long pour avoir sa place ici, reste à faire sur cette théorie. Et si celle-ci, en première lecture, peut s'apparenter à ce que l'on appelle maintenant la théorie des extensions de corps, il faudrait regarder de plus près comment s'articule en fait la lecture que fait Kronecker des travaux de Galois, et surtout d'Abel. Sur la cubique (voir p. 105), nous avons fait, très succinctement, quelques remarques à ce sujet.

15. Voir p. 72.

> Deux théories ont acquis une importance particulière pour le progrès de l'Algèbre moderne : d'une part, la théorie des groupes, qui tend de plus en plus à dominer les sujets les plus divers, et contribue à répandre partout l'ordre et la lumière ; en second lieu, la théorie des nombres. [Weber, 1898] p. VIII

Cette contradiction apparente n'est finalement rien d'autre que la différence entre ce que Weber pense être l'algèbre au moment où il écrit son *Lehrbuch*, et ce qu'il pense qu'elle deviendra dans les décennies qui suivront.

On a vu que Kronecker aussi, dans les premières phrases de son cours, décrit ce domaine comme une partie des mathématiques que les évolutions récentes ont complètement modifiée : ils ont donc l'un et l'autre parfaitement conscience des changements fondamentaux que subit l'algèbre en cette fin de XIX$^{e}$ siècle. Ils en sont eux-mêmes d'ailleurs les acteurs. Mais au contraire de Weber qui voit une transformation des objets : déplacement des équations vers les groupes, Kronecker propose plutôt une extension des objets, qui d'ailleurs par étapes successives devrait former ce qu'il appelle l'« arithmétique générale ». Si le constat sur l'algèbre est largement partagé, la vision de ce qu'elle doit devenir n'est pas la même.

### 4.3.3 Les deux introductions

Un dernier point : regardons comment Weber et Kronecker ont rédigé leurs introductions. Le long préambule du premier tome du *Lehrbuch* est une exposition de la notion d'ensemble à partir de laquelle Weber construit les nombres réels en faisant appel au travail de Dedekind. On sait bien sûr que la conception des nombres de Kronecker est bien différente de celle que présente Weber, mais ce n'est pas ce qui nous importe de faire remarquer ici. Weber choisit de faire une sorte d'introduction fondationnelle, basée sur la théorie des ensembles. Kronecker, au contraire, préfère s'interroger sur la notion de racine, aborder l'histoire de la théorie des équations, et finalement se pose plutôt des questions sur les concepts et leur légitimité dans les mathématiques. Examinons maintenant justement la façon dont ces deux mathématiciens abordent la notion de racine [16].

## 4.4 La notion de Racine

### 4.4.1 Introduction

Le concept de racine (*Wurzel*) est central, à plusieurs titres, dans ces deux cours. Il l'est dans la structure même des cours, mais il l'est aussi dans l'ensemble des théories développées dans ces deux cours : nous parlons en effet bien ici de « théorie des équations ». De plus, on trouve chez Weber comme chez Kronecker un paragraphe intitulé « Notion de racine » (*Begriff der Wurzeln*).

16. Nous avons déjà, pour Kronecker, longuement parlé de cette notion. Voir p. 78.

#### 4.4.1.1 La notion de Racine chez Weber

Preuve de son ancrage dans le XIX^e siècle, les racines occupent une place importante dans le tome I du *Lehrbuch der Algebra* de H. Weber. Il suffit pour s'en apercevoir de consulter la table des matières : non seulement le chapitre 3 du premier livre sur les *principes* (qui contient six chapitres) est consacré aux « Racines d'une équation algébrique », mais en plus le second livre (sur les trois qui forment ce premier tome) est intitulé « Les Racines ». Weber introduit la notion de racine lorsque les lettres qui étaient jusqu'alors « considérées comme symboles de grandeurs variables » (*variable Grössen*) deviennent maintenant des « grandeurs numériques » (*Zahlengrössen*). C'est l'occasion pour lui de rappeler ce qu'il entend par nombre (*Zahlen*) : ce sont des grandeurs de la forme $a + bi$, forme qu'il a définie dans son introduction. Dans celle-ci, il étend la notion de nombre à partir de celle d'ensemble, qu'il désigne comme étant un

> système d'objets ou d'éléments d'espèce quelconque, limité et achevé de telle manière que l'on puisse dire si un objet déterminé appartient à l'ensemble ou non ; peu importe d'ailleurs, que dans chaque cas particulier nous ayons réellement la possibilité de prendre cette décision ou non. [Weber, 1898] p. 4

Il construit ensuite l'ensemble des nombres réels à partir des coupures de Dedekind, et justifie ainsi non seulement, sinon l'existence, tout au moins la non contradiction d'un ensemble continu, mais aussi en géométrie la continuité de la droite. Les nombres complexes sont eux introduits en définissant sur $\mathbb{R}^2$ les quatre opérations : il montre finalement que l'ensemble qu'il a ainsi défini est un corps.

Voilà la définition qu'il donne de la *racine d'une équation algébrique* :

> Un nombre $\alpha$ qui, mis à la place de $x$, fait acquérir la valeur 0 à la fonction $f(x)$ s'appelle une racine de l'équation $f(x) = 0$. [Weber, 1898] p. 123

Weber poursuit en donnant les premières propriétés de ces racines sans s'interroger sur leur nature. En fait, tout le travail à propos de cette nature a finalement déjà été effectué dans l'introduction : une fois acceptée la construction des nombres réels et son extension au corps des complexes, la nature du nombre $\alpha$ est claire.

#### 4.4.1.2 La notion de Racine chez Kronecker

Le problème est ici bien différent, et nous allons rappeler en quelques mots la position de Kronecker [17]. On le sait, les nombres réels en tant qu'ensemble ne sont pas acceptables pour Kronecker, et il ne peut donc pas « piocher » à l'intérieur de ce dernier pour définir ce qu'il entend par racine. Il n'est donc pas question pour lui de revenir, comme Weber, à une définition formelle de racine : celle-ci est acquise. Par contre ce qui est discuté, ce sont la nature des racines et les méthodes qui y conduisent. L'utilisation du symbole $\sqrt[n]{a}$, qui n'est que formelle, et on connaît des méthodes qui permettent

17. Nous avons déjà traité de cette notion p. 78.

d'en calculer une valeur « approchée ». De cela découle, comme nous l'avons déjà vu, une première fausse piste : chercher à résoudre toutes les équations par extraction de racine.

Pour Kronecker, la « racine » d'une équation est d'autant plus difficile à définir que son existence n'est pas assurée. La notion même est à remplacer : ce n'est pas toujours un nombre que l'on peut atteindre, mais plutôt la limite d'une suite d'intervalles de plus en plus petits. Finalement, ce qui définit par exemple la racine (ou les racines) de l'équation $x^2-2$ est exactement l'équation $x^2-2$. Pour manipuler ces « racines », Kronecker mettra en place une théorie de la division, où par exemple dans le cas qui nous intéresse et en simplifiant dangereusement, il se placerait dans le corps $\mathbb{Q}$ quotienté par $x^2-2$. Mais Kronecker va plus loin : la notion même de continuité est remise en cause. En fait, comme on le verra ensuite, il n'hésite pas à s'en servir, lorsqu'il se place d'un point de vue géométrique. Ce qui est important ici pour Kronecker, c'est que dans le but de construire une arithmétique générale, il faut évacuer le plus possible toute notion de continuité. Cela aura par exemple une influence sur sa façon de traiter le théorème fondamental de l'algèbre.

### 4.4.2 Le théorème fondamental de l'algèbre

Maintenant que nous avons une idée plus précise de ce que la notion de racine représente pour chacun de ces deux mathématiciens, examinons comment ils appréhendent le théorème fondamental de l'algèbre. En effet, une fois que l'on a cerné la nature et la définition des racines, vient la question de leur existence, question à laquelle répond le théorème fondamental de l'algèbre. Ce théorème apparaît assez tôt chez Weber. Il en donne deux démonstrations assez semblables et presque complètement indépendantes de ce qu'il a fait jusqu'alors, et de ce qu'il fera ensuite : ce sont des démonstrations classiques basées sur le principe du maximum, et assez peu « algébriques ». Ce qui prime ici pour Weber, c'est d'établir le résultat, en évitant toute forme de complication. Une fois le théorème démontré, la question de l'existence de ces racines est complètement évacuée : il travaille sur $\mathbb{C}[X]$, et donc tout polynôme de degré $n$ a $n$ racines.

Kronecker, dès les premières pages, note l'importance de ce théorème [18] :

> La proposition, que *Gauss* nomme le théorème fondamental de l'algèbre, qui forme aussi véritablement le fondement de l'algèbre, doit apparaître dans ces leçons plutôt au début qu'à la fin. [Kronecker, 1891a] pp. 1-2

En fait, pour parvenir à démontrer ce théorème, Kronecker devra attendre les deux tiers de son cours, et son développement du théorème de Sturm et de sa théorie des caractéristiques : ce théorème sera pour lui un but à atteindre, qui justifiera en soi une partie de ce qu'il développe pour y parvenir, alors que pour Weber il s'agit d'un résultat à établir sur lequel il pourra développer son travail. Ce dernier reprend la démonstration

18. Der Satz, den *Gauss* den Fundamentalsatz der Algebra nennt, der gemeinhin auch wirklich das Fundament der Algebra bildet, soll in diesen Vorlesungen erst gegen den Schluß hin erscheinen

qu'en donne Kronecker dans son chapitre sur le théorème de Sturm, mais elle n'est présentée qu'à titre d'exemple.

Par ailleurs on se souvient que Kronecker cherche à échapper si possible à la notion de continuité : c'est ainsi que dans sa dernière partie, celle où il introduit la notion des genres (*Gattugen*), il transforme ce théorème fondamental de l'algèbre en un *théorème fondamental de l'arithmétique*, qui est en fait ce que Christian Gilain appelle le théorème de factorisation linéaire [19], c'est-à-dire un théorème qui affirme que l'on peut factoriser un polynôme dans la bonne extension de corps.

Nous allons pouvoir regarder maintenant le traitement que Weber et Kronecker ont réservé au théorème de Sturm, que nous explorerons plus longuement dans la partie suivante, ce qui nous permettra d'en donner une présentation succincte.

## 4.5 Un exemple : le théorème de Sturm

### 4.5.1 Introduction

Pourquoi avoir choisi cet exemple ?

Tout d'abord, ce théorème est extrêmement populaire à la fin du XIX$^{e}$ siècle, et il a fait l'objet, depuis sa publication en 1829, d'un grand nombre de recherches : en utilisant la bibliographie faite par H. Sinaceur [20], on peut compter 15 mémoires en rapport avec le théorème de Sturm entre 1832 et 1850, 45 entre 1851 et 1880, 28 entre 1881 et 1900 14 entre 1901 et 1950, et enfin 20 entre 1951 et 1982. Ainsi, quand ces deux cours sont rédigés, ce théorème est bien connu. Ensuite, ce théorème semble être un parfait représentant de l'algèbre du XIX$^{e}$ siècle, car après un succès que l'on pourrait presque qualifier de populaire dans la seconde moitié du XIX$^{e}$ siècle, il disparaît progressivement des livres d'enseignement au XX$^{e}$ siècle.
Le théorème de Sturm a aussi une place importante dans les deux cours : son traitement par chacun des deux auteurs aura une postérité notable au XX$^{e}$ siècle. D'ailleurs, le nom de Sturm est l'un des plus cités par Weber (40 fois), et cela toujours en rapport avec ce théorème. La notion de racine est centrale dans le théorème de Sturm, et comme on l'a vu elle est aussi centrale dans ces deux cours.
Enfin, notre seconde partie, étape obligée pour la compréhension de la théorie des caractéristiques, objet essentiel de notre travail, consiste en l'analyse de l'interprétation de ce théorème par Kronecker.

19. Voir [Gilain, 1991] et nos développement sur ce sujets dans la troisième partie p. 445.
20. Voir [Sinaceur, 1991] p. 438.

### 4.5.2 Les présentations de Weber et Kronecker

#### 4.5.2.1 La présentation de Weber

La présentation de ce théorème dans le *Lehrbuch* s'étend sur environ 40 pages, et a la structure suivante :

1. Présentation du problème
2. Définition abstraite des conditions que doivent respecter les suites de Sturm
3. Exemples
4. Extensions par Hermite
5. Théorie des caractéristiques de Kronecker
6. Application

Ce qui est ici remarquable, par rapports aux autres présentations de ce théorème qui ont cours à cette époque, c'est le second point. En effet, comme le fait remarquer Hourya Sinaceur :

> Heinrich Weber [...] introduit directement les « fonctions de Sturm », faisant totalement l'économie de l'approche naïve et ne conservant plus de la démonstration que son épure abstraite. Son exposé, le plus indépendant de celui de Sturm dans la mesure où il ne le reproduit pas mimétiquement, permet de mesurer la transformation éprouvée par le théorème après un demi-siècle d'existence et sous l'effet de la mathématique abstraite. [Sinaceur, 1991] p. 91

Weber commence donc non seulement par extraire les conditions que doivent respecter les fonctions de la suite, mais il en propose de plus une présentation basée sur ces conditions. Les exemples, applications et extensions, ne viendront qu'après. Weber fonctionne ainsi bien par abstraction, isolant les propriétés de la série de fonctions que Sturm avait fournies, pour en faire la base sur laquelle reposera son étude.

#### 4.5.2.2 La présentation de Kronecker

Kronecker, au contraire, choisit dans son cours de 1890 une présentation très différente. Il en donne successivement quatre démonstrations, ou tout au moins quatre cadres dans lesquels on peut faire vivre ce théorème. Pour trois d'entre elles, il s'agit de montrer que ce théorème se déduit de résultats qu'il a précédemment démontrés : des résultats sur les déterminants, le jacobien ou les fractions continues. La preuve que l'on retrouve dans tous ses cours entre 1880 et 1890, celle à partir de laquelle il fournit des exemples, est finalement l'une des moins abstraite, au sens où elle n'utilise plus un gros arsenal d'algèbre linéaire, mais au contraire elle repose sur une conception « géométrique » de la notion de racine [21] :

21. Wir knüpfen an die geometrische Anschauung an, welche uns lehrt, daß wenn die Curve $y = F(x)$, bald oberhalb, bald unterhalb der Abszissenaxe verläuft, sie dieselbe schneiden muß, falls die Curve stetig verläuft.

Der Begriff dieser geometrischen Stetigkeit ist aus der Anschauung klar, analytisch werden wir die Funktion $F(x)$ als an der Stelle $x$ stetig bezeichnen, wenn $F(x_1) - F(x_2)$, wo $x_1 < x < x_2$ sei, kleiner

> Nous renouons avec la conception géométrique, laquelle nous apprend que si une courbe $y = F(x)$ passe tantôt en dessous, tantôt au dessus de l'axe des abscisses, elle doit le couper, dans le cas où la courbe est continue.
>
> La notion de continuité géométrique est, vue ainsi, claire. Nous désignerons analytiquement la fonction $F(x)$ comme continue en $x$, si $F(x_1) - F(x_2)$, où $x_1 < x < x_2$, peut être faite plus petite que n'importe quelle grandeur, lorsque l'on rapproche $x_1$ et $x_2$ aussi près de $x$ que l'on veut et lorsque cette différence reste ensuite aussi plus petite que cette grandeur choisie aussi petite que l'on veut. Ce concept de continuité est si simple, mais aussi tellement moderne, qu'on ne le trouve pas encore chez Gauss. Il avait seulement la conception géométrique et le concept de continuité géométrique.
>
> Nous voulons maintenant, pour fonder la proposition de *Sturm* sur cette conception géométrique des racines de $F(x) = 0$, définir un symbole, que nous désignerons par $[A]$. [Kronecker, 1881a] p.104

On peut s'étonner, alors que l'on a vu comment il qualifiait de « trompeuses » les déductions de Bolzano, que Kronecker présente une démonstration utilisant une « conception géométrique de la continuité », dont la définition est exactement celle qu'utilise Bolzano dans la *Rein analystischer Beweiss*. Concernant ce paradoxe apparent, il y a deux remarques à faire. Tout d'abord, il ne faut pas oublier que Kronecker expose un cours à des étudiants, et qu'il doit, selon le but à atteindre, faire reposer sa leçon sur le socle de leurs connaissances communes. Et si pour cela dans un premier temps il doit utiliser des nombres « réels », ou une notion géométrique qui n'aurait pas sa place dans dans son *arithmétique générale*, le pédagogue pragmatique qu'il semble être n'hésite pas à les employer. Dans un second temps, il cherche à extraire de son exposé la notion de racine, mais seulement une fois que le contexte a été acquis par l'auditoire. Ensuite, la définition arithmétique de la continuité lui permet, nous l'avons vu, de passer de l'« existence des racines », qui est problématique, à la « séparation des racines ». Kronecker nous donne finalement une conclusion claire de ce que signifie pour lui le théorème de Sturm [22] :

> Le théorème de *Sturm* ne détermine par là rien d'autre que le nombre des intervalles caractéristiques précédents $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$ dans lesquels a lieu un changement de signe. [Kronecker, 1881a] p. 120

On détermine un nombre d'intervalles rationnels, et rien d'autre. Notre seconde partie consistera justement à démêler les différents points de vue que Kronecker porte sur ce théorème.

---

gemacht werden kann als jede beliebig kleine Größe, dadurch daß man $x_1$ und $x_2$ beliebig nahe an $x$ heranrückt und wenn diese Differenz dann auch kleiner als diese beliebig kleine Größe bleibt. So einfach dieser Begriff der Stetigkeit ist, so ist er doch so modern, daß er bei *Gauss* noch nicht vorkommt. Er besaß nur die geometrische Anschauung und den Begriff der geometrische Stetigkeit.

Wir wollen nun, um auf diese geometrische Anschauung der Wurzeln von $F(x) = 0$ den *Sturm*'schen Satz aufzubauen, ein Zeichen definiren, welches wir durch $[A]$ bezeichnen wollen.

22. Der *Sturm*'sche Satz bestimmt daher eigentlich nichts anderes als die Anzahl der oben charakterisirten Intervalle $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, zwischen welchen ein Zeichenwechsel stattfindet.

**Conclusion** Ainsi, lorsque l'on regarde un chapitre précis, que l'on retrouve dans tous les ouvrages d'algèbre de l'époque, d'importantes différences de traitement apparaissent. D'un côté, on note chez Weber une forte tendance à présenter de façon abstraite les notions. De l'autre côté Kronecker montre une volonté de faire vivre une même notion dans des cadres mathématiques très différents, en essayant d'atteindre une plus grande généralité, tout en conservant une position philosophique forte, ici sur la notion de nombre.

Autant le théorème de Sturm est un passage obligé des cours d'algèbre du XIX^e siècle, autant son extension à des systèmes de fonctions, que Kronecker développe sous le nom de théorie des caractéristiques, est rarement présentée. Rien d'étonnant à ce qu'elle fasse partie du cours de Kronecker, mais le fait que Weber choisisse lui aussi de la présenter est tout à fait remarquable.

### 4.5.3 Une extension : la théorie des caractéristiques

Cette théorie des caractéristiques a été élaborée par Kronecker et diffusée par une série d'articles publiés en 1869 et 1878. Elle a très peu été présentée dans des manuels d'algèbre, et généralement par des anciens élèves de Kronecker : exposer cette théorie est donc pour Weber un engagement fort. Si l'on reprend la structure générale des deux cours, cette théorie se situe dans les deux cas après la présentation du théorème de Sturm, au cœur de leurs cours.

#### 4.5.3.1 Présentation (très) rapide de cette théorie

Ce que l'on appelle la caractéristique [23] d'un système de $n+1$ fonctions à $n$ inconnues est un nombre entier attribué à ce système. Elle permet, entre autres, de trouver le nombre de solutions d'un système de $n$ équations à $n$ inconnues situées dans un compact de $\mathbb{R}^n$. Weber et Kronecker donnent des présentations de cette théorie extrêmement différentes l'une de l'autre.

#### 4.5.3.2 Chez Weber

Cette théorie est composée de deux versants. On pourrait qualifier le premier de topologique et d'algébrique [24] et le second d'analytique. Weber choisit de le présenter seulement sous la première de ces deux formes, en évacuant la partie topologique et en s'appuyant sur des évidences géométriques et un peu de géométrie différentielle. Les notions sont présentées à partir de schémas, desquels découlent des formules. La théorie a été élaborée pour la dimension $n$, mais Weber ne la présente que pour la dimension 2, c'est-à-dire le plan. En fait, sa présentation dépend bien trop du cas particulier du plan pour pouvoir être généralisée.

23. Il n'est pas question de s'étendre ici sur le contenu de la théorie : ce sera l'objet de notre troisième partie.

24. Nous développerons plus tard le choix de ces qualificatifs. Voir p. 393.

#### 4.5.3.3 Chez Kronecker

Kronecker aussi présente sa théorie dans le cas $n = 2$. Mais contrairement à Weber, il utilise des outils et des procédures issus de ses travaux de recherches en dimensions $n$ : ceux-ci pourrait donc être généralisés sans difficulté majeure à des dimensions supérieures. Prenons un exemple : dans les deux cas l'exposition de la théorie commence par la détermination d'un sens de parcours sur une courbe algébrique d'équation $f(x) = 0$. Weber détermine ce sens ainsi : on se déplace sur la courbe $f = 0$ dans le sens positif si on laisse à gauche la partie du plan où $f < 0$. On voit immédiatement la difficulté pour étendre cette définition à des dimensions supérieures. Kronecker donnera lui un critère de parcours qu'il qualifiera d'« arithmétique », et dont l'interprétation graphique, dans le cas $n = 2$, correspondra au critère donné par Weber. Bien sûr, au niveau pédagogique, le texte de Weber est bien plus abordable, mais il présente une théorie close et limitée, alors que celui de Kronecker est une invitation à la poursuite de la réflexion et à la généralisation.

## 4.6 Conclusion

Commençons par une remarque. Nous l'avons finalement assez peu mentionné, mais il convient de garder en tête la différence de conception des deux cours étudiés : l'un est un produit fini, revu et corrigé par son rédacteur, et destiné à être diffusé. L'autre au contraire n'a pas forcément été complètement relu par Kronecker, est peut-être incomplet, et n'avait pour destination qu'un cercle proche du rédacteur. Néanmoins, comme le cours de Weber reste issu de ceux qu'il a effectivement donnés devant ses étudiants, et que l'on sait [25] que Kronecker demandait à certains de ses étudiants de prendre des notes soigneuses dans l'idée d'une future publication de ses cours, la comparaison reste valable si l'on prend quelques précautions.

L'examen de la structure, ainsi que celle, plus fine, de certains concepts présents dans ces deux cours permet donc comme on vient de le voir de commencer une comparaison des deux cours. Cependant, pour affiner et confirmer les résultats de cette étude, il faudrait entrer bien plus précisément dans le contenu. Cela nous permettrait en effet de voir, au plus près de leur pratique mathématique, leurs ressemblances et dissemblances. Nous avons en partie effectué ce travail pour le théorème de Sturm, et les différences que l'on vient de voir en ressortent avec encore plus de force.

Essayons de résumer nos conclusions et de donner quelques éléments de comparaison entre ces deux expositions. Tout d'abord, nous sommes bien face à deux cours d'algèbre du XIX$^{e}$ siècle, centrés autour de la résolution des équations. Tous deux, bien que très ancrés dans leur époque, contiennent en germe quelques-uns des objets de l'algèbre du XX$^{e}$ siècle, comme l'algèbre abstraite dans le cas de Weber, et, nous en reparlerons plus

25. Nous avons vu que certains d'entre eux ont été rédigés à la demande de Kronecker en vue d'une publication : cependant, Kronecker lui-même le disait, un immense travail de synthèse aurait été à faire avant d'envisager l'édition d'un manuel.

tard, la topologie algébrique pour Kronecker.

Par ailleurs, nous avons vu se dessiner aussi de nombreuses dissemblances. Si le *Lehrbuch der Algebra* reste, tout au moins en ce qui concerne le premier tome, un manuel d'algèbre du XIX^e^ siècle centré autour de ce que Poinsot appelle en 1826 l'« analyse des équations »[26], l'utilisation plus systématique de la notion de groupe, et la direction que prennent les livres II et III le distinguent tout de même des *Vorlesungen* de Kronecker et le projettent dans les travaux sur les structures du XX^e^ siècle.

Ces deux cours montrent une volonté d'extension des concepts du XIX^e^ siècle. Chez Weber, cette extension semble s'apparenter plutôt à une montée en abstraction. Si l'on voulait simplifier la différence des deux positions sur les racines, nous pourrions dire que pour Weber, leur nature est acquise, mais il faut les définir formellement, et pour Kronecker, au contraire, ce qui est important, ce sont leur nature et les méthodes qui nous permettent de les appréhender.

Chez Kronecker, il y a plutôt une volonté de *généralisation* – avec par exemple l'extension du théorème de Sturm aux espaces de dimension $n$ – et d'interactions des divers champs des mathématiques. On trouve en effet souvent plusieurs démonstrations pour un même théorème abordant des domaines différents des mathématiques. Kronecker cherche finalement à étendre le domaine de l'algèbre, ou plutôt de ce qu'il appelle l'*allgemein Arithmetik*, l'arithmétique générale.

Cette comparaison donne ainsi du relief aux constats que l'on avait pu faire plus haut, lorsqu'on avait étudié la notion d'algèbre, telle que Kronecker l'entend dans ses *Vorlesungen über die Theorie der algebraischen Gleichungen*[27].

26. Voir l'introduction de la troisième édition du *traité de la résolution des équations numériques de tous les degrés* de Lagrange [Lagrange, 1826] p. V.
27. Voir p. 72.

# Chapitre 5

# Conclusion

Au même titre que le *Lehrbuch* pour Weber, les *Vorlesungen über die Theorie der algebraischen Gleichungen* constituent bien, pour Kronecker, son cours d'« algèbre ». Et si l'on en croit les propos de Gösta Mittag-Leffler que nous avons cités dans notre introduction, à savoir que « les mathématiques de [son] temps n'ont pas grand-chose qui puisse rivaliser avec la théorie des fonctions de Weierstraß ou l'algèbre de Kronecker », alors ces *Vorlesungen* représentent aussi une source importante pour comprendre ce qu'est l'« algèbre » de la seconde moitié du XIX$^{e}$ siècle.

Cette source est d'autant plus nécessaire que, comme nous l'avons vu, les résultats des recherches de Kronecker ont été certes pour la plupart publiés, mais souvent sans démonstration ni explication : leur lecture seule ne permet pas toujours d'en saisir le contenu, et presque jamais l'origine. Les cours qu'il donnait à Berlin étaient le lieu de la mise en contexte de ces résultats, et ils contiennent ainsi de nombreux exemples et justifications qui complètent les articles de recherche. Les manuscrits qui ont été l'objet de notre première partie, sont, comme nous l'avons vu, le moyen privilégié pour y accéder. La partie des mathématiques que Kronecker expose ici est, il le reconnaît, très ardue, mais c'est justement pour lui le lieu où il peut exposer ses méthodes de travail, « car on trébuche plus facilement sur un sol égal où l'on ne fait pas attention, que sur un sentier de montagne où l'attention est aiguisée » [1].

On voit se construire, dans les manuscrits des *Vorlesungen über die Theorie der algebraischen Gleichungen*, une Algèbre dont les contours, de l'aveu même de Kronecker, sont encore à définir, et qui parfois, peut sembler se propager à l'ensemble des mathématiques. Si le domaine des objets considérés s'étend pour tendre vers ce que Kronecker appellera son « arithmétique générale », une notion semble poser problème : celle de continuité. Et par là même, c'est le concept de racine, a priori central dans la théorie des équations, qu'il remet en cause. Car il est faux, pour Kronecker, d'interpréter le théorème de Bolzano

1. [Kronecker, 1885b] p. 3 : *denn man strauchelt leichter auf ebenem Boden, wo man nicht Acht giebt, als auf Bergpfaden, wo die Aufmerksamkeit geschärft ist.*

comme un théorème d'*existence*, tout au moins comme la preuve de l'existence d'une racine. Au mieux, on peut *séparer les racines*, au sens où l'on peut obtenir des intervalles, aussi petit que l'on veut, dans lesquels la fonction que l'on étudie change de signe une fois et une seule. La notion *géométrique* de racine, qui correspondrait à l'existence d'un point d'intersection entre le graphe d'une fonction et l'axe des abscisses, n'a pas sa place en algèbre, ni dans son *arithmétique générale*.

Car Kronecker propose clairement une partition des mathématiques en trois domaines : Arithmétique / Géométrie / Mécanique [2]. Il ne les hiérarchise pas, comme le font par exemple Gauss ou Bolzano – ce sont des « disciplines sœurs » – mais leur interactions ne sont pas non plus symétriques. On peut fonder des notions géométriques avec l'arithmétique, mais non l'inverse. Ainsi, alors que Weber interprète des schémas pour en déduire des formules, Kronecker, lui, démontre des formules puis les interprète géométriquement [3] :

> Certes les rapports arithmétiques sont transmissibles et applicables aux concepts géométriques clairs et [...] certes la géométrie a contribué et contribuera toujours aux progrès des connaissances arithmétiques, mais [...] jamais un pur concept « arithmétique » ne pourra et ne devra être défini par une illustration géométrique. [Kronecker, 1891a] p. 250

La géométrie est par contre une source d'inspiration extrêmement féconde pour l'arithmétique. Kronecker affirme, reprenant une affirmation de Gauss, que « les mathématiques sont une science pour les yeux bien plus que pour les oreilles » [4]. Bien sûr le contexte de cette phrase est celui de l'écriture algébrique, mais finalement, nous verrons que Kronecker propose un cours qui contient un nombre important de « visuels ».

Si l'on revient maintenant sur la structure générale des *Vorlesungen über die Theorie der algebraischen Gleichungen*, nous avons vu [5] qu'elles étaient constituées de deux parties principales, et que nous allions essentiellement nous occuper de la première, qui contient l'objet principal de notre recherche : la théorie des caractéristiques. Cette première partie est elle-même composée de plusieurs sections : 1) la théorie des équations quadratiques

2. Voir plus haut p. 97.

3. Die Mathematiker, welche eine solche, uns sehr bedenklich scheinende Ansicht zu der ihrigen machen, lassen dabei völlig außer Acht, daß zwar die arithmetischen Beziehungen auf die anschaulichen geometrischen Begriffe übertragbar und anwendbar sind, und daß die Geometrie freilich stets ganz wesentlich zum Fortschritt der arithmetischen Wissenschaften beigetragen hat und beitragen wird, daß aber niemals der rein « arithmetische » Begriff durch geometrische Veranschaulichung definiert werden kann und definiert werden darf.

4. Kronecker affirme que ces propos lui ont été rapportés par Johann Franz Encke (1791-1865), un ancien élève de Gauss :

> Die technische Ausbildung aber dient dem Fortschritt und hierfür citiere ich eine Äusserung von Gauss. Mir ist dieselbe von Encke citiert worden. Er sagt, die Mathematik sei weit mehr eine Wissenschaft für das Auge, als für das Ohr ; durch eine glückliche Zeichensprache mache sie es möglich, ganze Gedankenreihen auf einmal zu übersehen. Encke war ein ganz spezieller Schüler von Gauss und daher ein Zuverlässiger Berichterstatter. [Kronecker, 1885b] p. 11

5. Voir p. 58.

et cubiques, 2) les équations de degré $n$ et les fonctions symétriques, 3) Le théorème de Sturm et enfin 4) la théorie des caractéristiques. Nous avons, à l'aide essentiellement du manuscrit de 1880-81, donné dans cette partie un aperçu des sections 1) et 2). Nous allons maintenant consacrer la partie suivante au théorème de Sturm.

Deuxième partie

# Le théorème de Sturm dans les *Vorlesungen über die Theorie der algebraischen Gleichungen*

# Chapitre 6

# Introduction


6.1 Méthodologie . . . . . **177**
6.2 Le théorème de Sturm dans les *Vorlesungen über die Theorie der algebraischen Gleichungen* . . . . . **179**
6.2.1 Les motivations de Kronecker . . . . . 179
6.2.2 Le théorème de Sturm : une place centrale dans les *Vorlesungen* 180
6.2.3 Contenu du chapitre sur le théorème de Sturm . . . . . 182
6.3 Énoncé et premier exemple . . . . . **189**
6.4 Réception et diffusion du théorème de Sturm . . . . . **191**


Kronecker l'affirme [1], la « difficile théorie des caractéristiques » [2] prend sa source dans le travail algébrique de Sturm (1803-1855). Avant de pouvoir aborder cette théorie, nous allons donc devoir examiner ce que l'on appelle le *théorème de Sturm*, théorème qui, rappelons-le, compte pour un intervalle donné le nombre de racines réelles d'une équation. Nous allons voir comment ce théorème, entre les mains de Kronecker, évolue et subit finalement de profondes transformations.

Les difficultés rencontrées au XVIII$^{e}$ siècle pour donner une résolution algébrique générale (c'est-à-dire trouver les racines par une formule explicite) des équations polynomiales [3] poussent la communauté mathématique du début du XIX$^{e}$ siècle à développer des recherches sur les méthodes numériques. Ce théorème que Sturm expose [4] en 1829 à l'Académie des Sciences de Paris et dont il ne publie [5] une démonstration qu'en 1835 est directement issu de ces travaux.

Un travail important a été fait par Hourya Sinaceur sur l'histoire de ce théorème, et nous nous réfèrerons très souvent à ses articles *Deux moments dans l'histoire du théorème d'algèbre de CH. F. Sturm* [Sinaceur, 1988] et *Cauchy, Sturm et les racines des équations*

1. Voir plus loin p. 322.
2. Voir [Sinaceur, 1988] p. 103.
3. Voir par exemple le travail de Lagrange de 1870-71 : *Réflexions sur la résolution algébrique des équations*.
4. Voir [Sturm, 1829a] et [Sturm, 1829a].
5. Voir [Sturm, 1835].

[Sinaceur, 1992], à son ouvrage *Corps et modèles* [Sinaceur, 1991] ainsi qu'au chapitre *L'œuvre algébrique de Charles François Sturm* issu du livre édité par Jean-Claude Pont et Flavia Padovani [Padovani et Pont, 2009]. On trouve en effet dans ces travaux une analyse globale des travaux d'algèbre de Sturm, de leur enracinement dans les travaux de Fourier et de Lagrange. Ce que H. Sinaceur montre dans *Corps et modèles*, c'est le mouvement d'*algébrisation* qui a lieu autour de ce théorème, essentiellement par Hermite et Sylvester. Cependant, le chemin qu'elle choisit d'emprunter conduit à la construction des corps réels par Emil Artin (1898-1962) et Otto Schreier (1901-1923) ainsi qu'à l'utilisation que fera Tarski du théorème de Sturm dans ses travaux de logique. L'angle sous lequel nous l'aborderons ici est différent. Ce sont les lectures de Kronecker de ce théorème qui nous intéressent : sa lecture mathématique, mais aussi sa lecture historique. Cependant, pour comprendre le travail de Kronecker, il sera indispensable de rappeler les grandes étapes de la construction de ce théorème, et nous nous efforcerons de présenter le plus précisément possible les travaux sur lesquels Kronecker s'appuie.

Nous savons [6] que H. Sinaceur inscrit le travail de Kronecker sur le théorème de Sturm dans ce processus d'algébrisation, processus qui aboutit d'après elle à la théorie des caractéristiques. Pourtant, ce qui frappe, lorsque l'on parcourt les *Vorlesungen über die Theorie der algebraischen Gleichungen*, c'est la présence de dessins, de schémas, de graphiques, en nombre bien supérieur à ce que l'on peut trouver dans d'autres cours à la même époque [7], en particulier dans un cours d'algèbre. Et ces représentations graphiques ne sont pas uniformément réparties, mais concentrées dans les chapitres traitant du théorème de Sturm et de la théorie des caractéristiques. En examinant plus précisément ces figures, nous verrons qu'il ne s'agit pas de simples illustrations, mais d'éléments décisifs qui participent, au moins à première vue, à une *interprétation* graphique de la situation. Nous devrons ainsi examiner la cohérence de cette exposition avec le discours de Kronecker concernant les rapports entre algèbre et géométrie que nous avons traités dans la partie précédente [8].

Nous allons chercher à comprendre, dans cette partie, comment le travail de Kronecker sur le théorème de Sturm s'inscrit bien dans un processus d'algébrisation, et quel rôle y jouent la continuité et la géométrie. Cela nous permettra, dans la partie suivante, de découvrir que la théorie des caractéristiques n'est pas issue seulement de ce *mouvement d'algébrisation* du théorème de Sturm mais aussi de la confrontation de ces deux points

6. Voir p. 18.

7. Par exemple, on peut constater qu'il y a autant de schémas de tout ordre (à peu près une vingtaine) dans les chapitres consacrés à Sturm et à la théorie des caractéristiques des *Vorlesungen* de Kronecker du semestre d'hiver 1880-1881 qu'il y en a dans l'ensemble des trois cours de Weierstrass : *Theorie der analytischen Functionen* de 1880-1881 [Weierstrass, 1881], *Theorie der elliptischen Functionen* de 1881-1882 [Weierstrass, 1882c] et *Theorie der Abelschen Funktionen* de 1881-1882 [Weierstrass, 1882a] et [Weierstrass, 1882b]. Si l'on regarde le cours d'algèbre publié de Weber, tout au moins le premier tome, on constate qu'il contient seulement 28 figures, dont une dizaine se trouvent justement dans la partie où il traite de la théorie des caractéristiques de Kronecker.

8. Voir p. 97.

de vue : le premier *algébrique*, issu de ce que nous présenterons dans cette partie, le second *analytique*, qui permettra à Kronecker de *généraliser* ses résultats. Nous montrerons d'ailleurs, lors de notre exposition de la théorie des caractéristiques, comment cette dualité de point de vue est présente dans les démonstrations de Gauss du théorème fondamental de l'algèbre, et comment Kronecker, grand lecteur de Gauss, reprend lui-même ces démonstrations. Le travail de Kronecker apparaîtra comme issu d'une volonté de *généralisation*, volonté que l'on peut voir comme une des origines de cette approche double.

Dans cette introduction, après avoir décrit la méthode que nous suivrons dans notre seconde partie et avoir exposé la place du théorème de Sturm dans le cours de Kronecker, nous montrerons en quoi ce théorème consiste et nous donnerons quelques indications sur sa réception.

## 6.1 Méthodologie

Nous souhaitons présenter l'évolution du théorème de Sturm entre sa première publication en 1829 et sa reprise par Kronecker dans ses articles *Sur le théorème de Sturm* [Kronecker, 1869b], *Über die verschiedenen Sturm'schen Reihen und ihre gegenseitigen Beziehungen* [Kronecker, 1873b] et *Über Sturm'sche Functionen* [Kronecker, 1878c], tout en gardant à l'esprit le but de cette présentation : le cœur de notre étude est la théorie des caractéristiques, et en particulier son exposition dans les *Vorlesungen über die Theorie der algebraischen Gleichungen*. Il ne s'agira donc pas ici de prétendre à une étude exhaustive de l'ensemble des travaux algébriques de Sturm pendant la période citée, mais d'examiner, *du point de vue de Kronecker*, quelles ont été ces évolutions. Nous allons ainsi essayer de mettre en relation la théorie présentée par Kronecker avec les travaux sur lesquels ce dernier affirme s'appuyer.

La méthode employée sera la suivante : nous allons suivre pas à pas la trame du chapitre que Kronecker consacre au travail de Sturm, ou plutôt, comme nous le verrons, à *son* travail sur le théorème de Sturm. Pour éclairer ce chapitre, bien plus vaste que la simple exposition d'une théorie déjà constituée, nous expliciterons les principaux travaux auxquels Kronecker se réfère. Nous avons ainsi choisi de découper cette présentation en fonction des principaux acteurs : c'est ainsi que les noms de Sturm, Sylvester, Cauchy, Cayley et Hermite se succèdent dans les titres de nos chapitres. À l'intérieur de chacun d'entre eux apparaîtront bien sûr nombre d'autres mathématiciens, et cela nous permettra de mettre en évidence, en parallèle du contenu, une histoire du théorème de Sturm et de son appropriation par Kronecker.

Dans chaque chapitre, nous commencerons généralement par exposer la théorie à laquelle Kronecker fait référence, que ce soit d'ailleurs de façon explicite ou implicite, puis nous aborderons la partie de son cours qui correspond à cette théorie : cela nous permettra de faire ressortir l'originalité de l'apport de Kronecker. Ce dernier fait preuve d'une grande

technicité dans son cours et il serait illusoire, voire dangereux, d'essayer de la gommer. Rappelons en effet ici l'une des hypothèses que nous avons formulées dans l'introduction et qui sera l'un des fils directeurs de notre travail : choisir d'aborder les mathématiques de Kronecker à travers les manuscrits de son cours est une porte d'entrée vers ses travaux les plus complexes. En effet l'expérience montre qu'une simple lecture de ses articles de recherche ne permet pas toujours d'appréhender correctement [9] ses théories : pour rendre compte de cela il nous faudra mettre en évidence ces difficultés de lecture, dont certaine sont liées à des problèmes d'une grande technicité.

Ensuite, nous souhaitons confronter le point de vue sur les mathématiques de Kronecker avec sa pratique. Comme nous l'avons déjà dit, le cours qu'il professe à l'université de Berlin est un lieu particulièrement propice pour examiner cette question. Mais cette confrontation sera d'autant plus fine qu'elle portera sur des exemples précis dont les mécanismes auront été entièrement démontés. C'est pourquoi nous essayerons de donner le plus de détails, tout au moins lorsque l'on exposera le travail de Kronecker.

Pour être au plus près du texte de Kronecker, notre analyse de son travail sur Sturm suivra celle de son cours sur *la théorie des équations algébriques*. Bien sûr, ce cours évolue au fils des ans et sa structure n'est pas toujours la même, c'est pourquoi nous expliciterons l'ordre de notre exposition, ainsi que les manuscrits auquel il se rapporte, dans les sections suivantes. Les notations employées sont assez proches d'une année sur l'autre, mais pas toujours identiques. Pour faciliter la compréhension, nous uniformiserons les notations, en prenant bien sûr garde à chaque fois d'examiner la signification potentielle de ces différences. Cette partie ne sera donc pas organisée par thème, et nous nous permettrons de développer certaines de nos remarques au fur et à mesure de notre exploration du travail mathématique de Kronecker. Ces allers-retours entre explication mathématique et commentaires permettront aussi de montrer la grande proximité qui existe entre pratique et philosophie des mathématiques chez Kronecker. Ce dernier n'hésite en effet pas à interrompre le cours d'un exposé complexe pour parler de la *notion de racine*, du *concept de nombre*, de la validité de *la notion de continuité* ou même pour rappeler l'histoire d'un théorème ou d'une notion.

Enfin, tout au long de cette partie, nous essayerons de nous reporter le plus souvent, et le plus précisément possible aux manuscrits de Kronecker. En effet, l'un des buts de notre travail, et non le moindre, est aussi de commencer à rendre accessibles ces cours qui n'ont pour le moment jamais été exploités.

9. Nous pourrons nous reporter aux propos de H. Sinaceur et H. M. Edwards p. 18 et p. 23.

## 6.2 Le théorème de Sturm dans les *Vorlesungen über die Theorie der algebraischen Gleichungen*

Nous suivrons essentiellement les cours des *Wintersemester* 1878-1879 ([Kronecker, 1879], 1880-1881 [Kronecker, 1881a], 1884-1885 [Kronecker, 1885b] et [Kronecker, 1885a], et 1886-1887 [Kronecker, 1887b] et [Kronecker, 1887a]. Nous commencerons par décrire succinctement les plans que suivent les chapitres consacrés à Sturm dans chacun des manuscrits [10].

On peut déjà faire deux remarques au sujet de l'importance que revêt ce théorème dans le cours de Kronecker. La première est que la partie qui lui est consacrée est, dans la plupart des manuscrits, clairement identifiée, quelquefois par un titre, [11] et cela fait figure d'exception, quelquefois par un changement de rédacteur [Kronecker, 1885b]. Ensuite, sous une forme ou sous une autre, plus ou moins développée, ce chapitre est présent sans exception dans les manuscrits à notre disposition. Nous allons donc nous interroger sur la place que prend le théorème de Sturm dans les *Vorlesungen* et sur les raisons de sa présence.

### 6.2.1 Les motivations de Kronecker

Voici comment Kronecker introduit son chapitre sur le théorème de Sturm dans le manuscrits [12] de 1884-1885 :

> Le théorème de Sturm offre une autre application de la méthode d'Euclide. Dans l'introduction, j'ai beaucoup relaté les efforts pour résoudre la question de savoir combien il y a de racines réelles parmi les racines d'une équation à coefficients réels. [Kronecker, 1885b] pp. 234-235

et dans celui [13] de 1886-1887 :

> Nous allons nous occuper dans la suite de façon détaillée du théorème de Sturm. Celui-ci trouve sa place ici pour au moins deux raisons : premièrement parce qu'on y trouve une application importante de la méthode d'Euclide pour trouver le plus grand facteur commun de deux fonctions entières ; ensuite, parce qu'il nous conduit

10. Les manuscrits de 1880-1881 et 1878-1879, en ce qui concerne la partie sur le théorème de Sturm, suivent exactement le même plan. La différence réside essentiellement dans le fait que l'exemplaire de 1878-1879 est beaucoup plus développé. Celui de Göttingen de 1886-1887 et celui de Strasbourg de la même année sont comparables et ne nécessitent pas ici de traitement différencié. Il en est de même pour les deux manuscrits de 1884-1885. Par ailleurs, le cours de 1872-73 est antérieur à son principal article de 1873 sur le théorème de Sturm, et les deux derniers de 1888-89 et 1890-91 n'abordent le théorème que de manière assez succincte : ces trois cours ne prendront donc dans cette partie qu'une place secondaire.

11. Voir [Kronecker, 1881a] p. 85, [Kronecker, 1885b] p. 236, [Kronecker, 1887b] p. 131, [Kronecker, 1889b] p. 120 et [Kronecker, 1891a] p. 170.

12. Weitere Anwendung des Euklidischen Verfahrens bietet der Sturmsche Satz. Ich habe in der Einleitung viel von den Bemühungen erzählt, die Frage zu lösen, wieviel reelle Wurzeln zwischen den Wurzeln einer Gleichung mit reellen Koeffizienten sind.

13. Wir werden uns im Folgenden eingehend mit dem Sturm'schen Satz beschäftigen. Aus zwei Gründen findet derselbe am besten hier seine Stelle : einmal, weil bei ihm das im vorigen Kapitel auseinandergesetzte Euclidische Verfahren, den größten gemeinschaftlichen Teiler zweier ganzen Funktionen aufzufinden, eine wichtige Anwendung findet ; sodann, weil er uns zu besonders wichtigen und eleganten symmetrischen Functionen führen und dadurch an die Entwicklungen der beiden ersten Kapitel wieder anknüpfen wird.

> aux particulièrement importantes et élégantes fonctions symétriques et ce faisant il renoue avec le développement des deux premiers chapitres. [Kronecker, 1887b] p. 131

Bien sûr, les raisons de la présence de ce chapitre dans le cours de Kronecker sur les équations algébriques sont nombreuses et dépassent largement la simple nécessité de trouver des applications aux chapitres précédents. Mais il est cependant notable de constater que Kronecker attache une certaine importance à ce que son propos soit suivi d'applications (*Anwendung*), ce qui dénote de sa part de réelles préoccupations pédagogiques. Il s'agit donc pour lui ici de donner des applications de ce dont il a traité précédemment : les polynômes symétriques et l'algorithme d'Euclide. La première partie de son cours contient en effet un très long travail sur les polynômes symétriques : l'algorithme d'Euclide a en revanche été traité dans ses *Vorlesungen über Zahlentheorie*[14]. En effet, comme nous l'avons déjà vu les *Vorlesungen über die Theorie der algebraischen Gleichungen* font partie d'un cycle de cours que Kronecker donne à l'université de Berlin : on y trouve donc des applications de ses autres cours.

Pour Kronecker, le théorème de Sturm ne permet pas « seulement » de compter les racines[15] :

> Pourtant le théorème de *Sturm* ne permet pas seulement de déterminer le nombre de racines réelles comprises entre $x_1$ et $x_2$, mais il donne aussi une méthode pour calculer numériquement ces racines. Si on laisse en effet $x_1$ et $x_2$ se rapprocher jusqu'à ce qu'il n'y ait qu'une racine comprise entre les deux, alors nous pouvons enfermer celle-ci entre des limites toujours plus étroites $x_1$ et $x_2$, jusqu'à ce que nous les ayons déterminées à la précision voulue. [Kronecker, 1881a] p. 107

Nous avons déjà vu, dans la partie précédente, que cette notion de racine d'une équation, à savoir un intervalle rationnel que l'on peut rendre aussi petit que l'on veut, est fondamentale dans la conception des mathématiques de Kronecker. Ce que dit Kronecker ici, c'est que le théorème de Sturm, et donc en fait la formule qui pour lui représente ce théorème, fournit un algorithme (une méthode) qui permet un calcul numérique de ces solutions[16].

Cependant, il ne faut pas se contenter des raisons explicites que fournit Kronecker. Considérer la partie sur le théorème de Sturm comme une simple illustration des chapitres précédents serait une erreur. Au contraire, il s'agit, comme nous allons le voir, d'un point d'articulation important de son cours.

### 6.2.2 Le théorème de Sturm : une place centrale dans les *Vorlesungen*

Nous allons maintenant regarder de plus près comment s'articule la partie consacrée à Sturm dans l'ensemble des cours qu'il a donnés à Berlin sur la théorie des équations

14. Voir l'édition de K. Hensel [Kronecker, 1901] p. 170.

15. Aber der *Sturm*'sche Satz erlaubt nicht nur die Anzahl der zwischen $x_1$ und $x_2$ liegenden reellen Wurzeln zu bestimmen, sondern gibt auch eine Methode, diese Wurzeln numerisch zu berechnen. Lassen wir nämlich $x_1$ und $x_2$ zusammenrücken, bis nur noch eine Wurzel zwischen ihnen liegt, so können wir diese in immer engern Grenzen $x_1$ und $x_2$ einschließen, bis wir sie mit jeder gewünschten Genauigkeit bestimmt haben.

16. Solutions qui, rappelons-le, *sont* ces intervalles rationnels.

algébriques. On peut déjà examiner l'espace occupé par ce chapitre dans les manuscrits :

| Semestre | 72-73 | 78-79 | 80-81 | 84-85 | 84-85(Thiel) |
|---|---|---|---|---|---|
| Nombre de pages total | 493 | 787 | 170 | 519 | 411 |
| localisation des pages consacrées au théorème de Sturm | 97-167 | 296-481 | 104-150 | 236-333 | 149-239 |
| Nombre de pages consacrées au théorème de Sturm | 70 | 185 | 46 | 97 | 90 |
| Pourcentage | 14 | 24 | 27 | 19 | 22 |

| Semestre | 87-88 | 87-88 (Lamla) | 88-89 | 90-91 |
|---|---|---|---|---|
| Nombre de pages total | 319 | 422 | 298 | 420 |
| localisation des pages consacrées au théorème de Sturm | 126-200 | 131-232 | 120-131 | 170-201 |
| Nombre de pages consacrées au théorème de Sturm | 74 | 101 | 11 | 31 |
| Pourcentage | 23 | 24 | 4 | 7 |

Ce tableau fait apparaître trois périodes :

1. La première [17], avant la publication de son plus important article [18] consacré exclusivement au théorème de Sturm en 1873.
2. La seconde, du manuscrit de 78-79 à ceux de 87-88 durant laquelle la part consacrée à Sturm reste à peu près stable.
3. Et enfin la dernière, entre 88 et 91, année de sa mort, où Kronecker semble avoir supprimé une partie de son enseignement sur le théorème de Sturm.

Nous verrons qu'une lecture attentive du manuscrit de 1872-1873 montre qu'il s'agit plus d'une ébauche de ce qu'il fera les années suivantes que d'une version réellement différente. En revanche, les deux derniers manuscrits montrent des changements dont il faudra tenir compte. C'est pourquoi notre exposition se concentrera sur la période centrale et les six manuscrits déjà cités. On voit donc apparaître déjà une certaine forme de centralité temporelle dans l'évolution de ce chapitre au fil des manuscrits.

Ce que le tableau montre aussi, c'est l'endroit du cours où le théorème est traité. On

17. Il aurait été bien sûr préférable de pouvoir comparer avec d'autres manuscrits de cette période, mais cela n'a pas été possible : nous n'avons pas connaissance de l'existence de notes sur les cours de Kronecker avant 1872. Par ailleurs, il est tout à fait envisageable que le semestre d'hiver 1872-73 soit le premier durant lequel Kronecker enseigne la théorie des équations algébriques.

18. Voir [Kronecker, 1873b].

voit que la place occupée par le théorème de Sturm est au sens propre, dans la période qui nous intéresse, au *centre* de son cours [19]. Le décalage vers la première moitié du cours pour les deux derniers manuscrits est dû à la volonté croissante de Kronecker, à partir de la publication de ses *Grundzüge*, de faire entrer dans ses leçons sur les équations algébriques sa théorie des *Gattungen*. Comme nous l'avons déjà vu, certains manuscrits contiennent des indications sur les dates auxquelles les notes sont prises. Ces informations nous permettent de voir que le chapitre consacré à Sturm est traité de la fin décembre à la mi-janvier, c'est-à-dire au *centre* du semestre d'hiver qui s'étend d'octobre à mars. Pour poursuivre plus avant notre analyse, il nous faut maintenant examiner d'un peu plus près ce qui constitue ce chapitre.

### 6.2.3 Contenu du chapitre sur le théorème de Sturm

Nous allons commencer par essayer de lister ce que l'on trouve dans ce chapitre consacré au théorème de Sturm. Il faut noter qu'il n'est pas toujours facile, dans le cas des manuscrits où le découpage en chapitres n'a pas été effectué, de distinguer ce qui appartient ou n'appartient pas à cette partie. Cependant, en utilisant les sommaires ou le découpage en chapitres des manuscrits [Kronecker, 1881a], [Kronecker, 1887b], [Kronecker, 1889b] et [Kronecker, 1891a] et après lecture des passages concernés, on obtient la liste suivante :

- Énoncé du Théorème de Sturm
- Démonstration avec la notion géométrique de continuité : obtention de la formule
$$\sum_{f(\xi)=0} sgn(f'(\xi)f_1(\xi))$$
pour $x_1 < \xi < x_2$.
- Interprétation géométrique du théorème de Sturm
- Exemple $x^n + nax + b = 0$
- Passage du cas $]-\infty;+\infty[$ au cas $[x_1;x_2]$
- Nombre total de racines réelles
- Application du théorème de Sturm aux équations biquadratiques
- la notion de racine d'une équation
- Travail préliminaire sur les fractions continues
- Définition du problème de Cauchy
- 1re méthode de résolution du Problème de Cauchy. Attribuée à Liouville : utilisation directe de la décomposition en fraction continue de $\frac{f}{f_1}$
- Résolution du problème de Cauchy sous la forme d'un déterminant
- 2e méthode de résolution du problème de Cauchy. Attribuée à Baltzer [20] : résoudre un système en prenant dans l'équation
$$F = f_1\Psi - f\Phi$$
$\Psi$ et $\Phi$ comme inconnues.
- 3e méthode de résolution du problème de Cauchy. Résoudre un système en prenant

19. Le chapitre est un peu décalé dans le cours de 1880-81, mais les notes de cours de Runge qui semblent être issues de cette session [Runge, 1881] montrent qu'il manque peut-être la fin du dernier chapitre sur la théorie des caractéristiques. De plus, nous l'avons vu, ce manuscrit semble être la première partie d'une série de deux tomes.

20. [Kronecker, 1881a] p. 178 : « Die Ausführeung findet sich bei Baltzer ». Il s'agit de Richard Heinrich Baltzer (1818-1887) qui a enseigné les mathématiques à l'université de Grießen. On sait par ailleurs que Kronecker correspondait avec lui, en particulier à propos de la théorie des déterminants ([Kronecker, 1869a]).

dans l'équation

$$F = f_1\Psi - f\Phi$$

$F$ et $\Psi$ comme inconnues.

— 4e méthode de résolution du problème de Cauchy. Attribuée à Jacobi : grâce aux formules d'Euler, on se retrouve avec un système de $\mu$ équations à résoudre. Dans cette méthode apparaît le déterminant $D_\mu$

— Le déterminant $|s_{p+q}|$

— Utilisation des fonctions $D_\mu$ pour le théorème de Sturm.

— 5e méthode de résolution du problème de Cauchy. Utilisation d'un développement selon les puissances décroissantes.

— Discussion sur la forme des solutions du problème de Cauchy

— Discussion sur la dimension de la variété représentant l'ensemble des solutions du problème de Cauchy

— Problème d'approximation de Tchebicheff

— Démonstration du Théorème de Sturm par la Méthode de Jacobi-Hermite

— Loi d'inertie de Sylvester

— Lien entre le théorème de Sturm et les inégalités séculaires du mouvement des planètes

L'ordre dans lequel ces points sont abordés peut changer d'une année à l'autre. Par exemple, les manuscrits de 78-79 et 80-81 sont traités de façon identique, mais celui de 90-91 sera au contraire rédigé à rebours. Nous allons examiner les changements qui apparaissent dans l'ordre du traitement de ces questions, et essayer d'en rendre compte. Pour cela nous allons commencer par donner un aperçu de cet ordre, année par année. Nous en profiterons pour donner, ou rappeler, certaines informations sur chacun de ces manuscrits qui nous semblent pertinentes dans le cadre de notre étude du théorème de Sturm.

#### 6.2.3.1 Le théorème dans chacun des manuscrits

**Manuscrit du semestre d'hiver 1872-73** Dans ce manuscrit, l'énoncé du théorème de Sturm arrive tard. Kronecker commence par construire tous les prérequis dont il a besoin, sans donner d'indication quant au but de son travail. Même si l'on retrouve la plupart des résultats qui seront énoncés les années suivantes, le théorème n'est pas mis en avant et l'ensemble des résultats ne semble pas former un tout cohérent. Ce qui manque ici, c'est le *liant algébrique* qui est présent dans toutes les autres présentations : le *problème de Cauchy*.

1. Travail préliminaire sur les fractions continues (p. 97)
2. Définition et construction des déterminants $D_\mu$ (p. 108)
3. Résumé de l'ensemble des résultats trouvés (p. 119)
4. Loi d'inertie (p. 122)
5. Interprétation graphique (p. 137)
6. Énoncé du Théorème de Sturm (p. 145)
7. Exemple : $x^n - p = 0$ (p. 147)
8. Formules de Sylvester (p. 150)
9. Méthode de Hermite
10. Exemple : $x^7 - 28x^4 + 480 = 0$ avec recherche numérique des solutions (p. 164)

Le théorème de Sturm à proprement parler n'est donc énoncé que dans un second temps, après que tout le travail de Kronecker sur les fractions continues, les déterminants

$D_\mu$ et même la loi d'inertie ont été exposés. Cependant, les grandes démonstrations du théorème, celle à partir de la formule $\sum_{f(\xi)=0} sgn(f'(\xi)f_1(\xi))$ et celle à partir des formes quadratiques sont déjà présentes.

**Les manuscrits des semestres d'hiver 1878-79 et 1880-81** Ces deux manuscrits suivent un même plan, tout au moins pour ce qui concerne le théorème de Sturm, et peuvent donc être traités simultanément [21]. Nous mettrons entre parenthèses d'abord la référence au manuscrit de 1878-79, puis celle de 1880-81. Le manuscrit de 1878-79 est le plus détaillé et le plus long, tandis que celui de 1880-81, le plus synthétique, est le plus court. De plus, il s'agit des deux derniers manuscrits qui précèdent la publication des *Grundzüge*, publication qui représente un tournant dans l'œuvre de Kronecker. C'est pourquoi, à peu de choses près, nous avons choisi de suivre globalement le plan de ces manuscrits.

1. Démonstration avec la notion géométrique de continuité : obtention de la formule
$$\sum_{f(\xi)=0} sgn(f'(\xi)f_1(\xi))$$
pour $x_1 < \xi < x_2$. (p. 296 / p. 104)
2. Interprétation géométrique du théorème de Sturm (p. 312 / p. 110)
3. Exemple : $x^n + nax + b = 0$ (p. 317 / p. 111)
4. Nombre total de racines. Obtention de la formule $A - E = \sum sgn(c_{k-1}c_k)$ (p. 322 / p. 112)
5. Application aux équations biquadratiques et surface-discriminante (p. 324 / p. 113)
6. La notion de racine (p. 338 / p. 116)
7. Travail préliminaire sur les fractions continues (p. 349 / p. 121)
8. Le problème d'interpolation de Cauchy (p. 360 / p. 125)
9. 1re méthode de résolution du Problème de Cauchy. Attribuée à Liouville : utilisation directe de la décomposition en fraction continue de $\frac{f}{f_1}$ (p. 362 / p. 126)
10. 2e méthode de résolution du problème de Cauchy. Attribuée à Baltzer : Résoudre un système en prenant dans l'équation
$$F = f_1\Psi - f\Phi$$
$\Psi$ et $\Phi$ comme inconnues.(p. 416 / Annexe p. 178)
11. 3e méthode de résolution du problème de Cauchy. Résoudre un système en prenant dans l'équation
$$F = f_1\Psi - f\Phi$$
$F$ et $\Psi$ comme inconnues.(p. 416 / Annexe p. 178)
12. 4e méthode de résolution du problème de Cauchy. Attribuée à Jacobi : grâce aux formules d'Euler, on se retrouve avec un système de $\mu$ équations à résoudre. Dans cette méthode apparaît le déterminant $D_\mu$. (p. 370 / p. 127)
13. Utilisation des fonctions $D_\mu$ pour le théorème de Sturm. (p. 386 / p. 131)
14. Discussion sur la forme des solutions du problème de Cauchy (p. 398 / Annexe p. 176)
15. 5e méthode de résolution du problème de Cauchy. Utilisation d'un développement selon les puissances décroissantes. (p. 419 / p. 136)

21. La numérotation du manuscrit de 1880-81 que l'on utilise débute à l'introduction : elle correspond donc à la page $n-2$ de la version PDF téléchargeable sur le site de la Bibliothèque numérique patrimoniale de Strasbourg.

16. Problème d'approximation de Tchebicheff (p. 425 / p. 139)
17. Démonstration du Théorème de Sturm par la Méthode de Jacobi-Hermite (p. 433 / p. 141)
18. Loi d'inertie de Sylvester (p. 440 / p. 142)
19. Transformation de Jacobi
20. Lien avec les inégalités séculaires du mouvement des planètes (p. 467 / p. 151)

Ces deux cours sont certainement les plus complets que Kronecker a donnés sur le théorème de Sturm. Nous essayerons de présenter l'ensemble de ces points, à l'exception du problème d'approximation de Tchebicheff et de la transformation de Jacobi, présents dans peu de manuscrits, qui demanderaient de trop longs développements et qui n'ont que peu d'importance pour le but que nous nous sommes fixé.

**Les manuscrits des semestres d'hiver 1884-85** Nous avons en notre possession deux manuscrits du cours de 1884-85. Nous mettrons entre parenthèses d'abord la référence au manuscrit [Kronecker, 1885b], puis celle de [Kronecker, 1885a]. Ce cours est le premier (en notre possession) qui suit la publication des *Grundzüge*. Le manuscrit est composé de deux grandes parties : la première qui se termine par le théorème de Sturm et la seconde qui décrit la théorie des *Gattungen*. Le théorème de Sturm est donc ici une sorte de conclusion de cette partie pré-*Gattungen*.

1. Démonstration avec la notion géométrique de continuité : obtention de la formule
$$\sum_{f(\xi)=0} sgn(f'(\xi)f_1(\xi))$$
pour $x_1 < \xi < x_2$. (p. 236 / p. 149)
2. Interprétation géométrique du théorème de Sturm (p. 242 / p. 154)
3. Passage du cas $]-\infty;+\infty[$ au cas $[x_1;x_2]$ (p. 244 / p. 157)
4. Nombre total de racines. Obtention de la formule $A-E=\sum sgn(c_{k-1}c_k)$ (p. 246 / p. 159)
5. Travail préliminaire sur les fractions continues (p. 249 / p. 161)
6. Le problème d'interpolation de Cauchy (p. 256 / p. 166)
7. Énoncé 2$^{e}$ et 3$^{e}$ méthode (p. 257 / p. 167)
8. 1$^{re}$ méthode de résolution du Problème de Cauchy. Attribuée à Liouville : utilisation directe de la décomposition en fraction continue de $\frac{f}{f_1}$ (p. 257 / p. 167)
9. Exemple : $f(x)=(x-1)(x-2)(x-3)(x-\xi)$ et $f_1(x)=x^2-6x+11$ (p. 262 / p. 172)
10. Résolution du problème de Cauchy sous la forme d'un déterminant (p. 266 / p. 175)
11. 4$^{e}$ méthode de résolution du problème de Cauchy. Attribuée à Jacobi : grâce aux formules d'Euler, on se retrouve avec un système de $\mu$ équations à résoudre. Dans cette méthode apparaît le déterminant $D_\mu$. (p. 268 / p. 176)
12. Utilisation des fonctions $D_\mu$ pour le théorème de Sturm. (p. 275 / p. 182)
13. Formules de Sylvester (p. 279 / p. 185)
14. Application aux équations biquadratiques et surface-discriminante (p. 286 / p. 192)
15. Démonstration du Théorème de Sturm par la Méthode de Jacobi-Hermite (p. 293 / p. 201)
16. Loi d'inertie de Sylvester (p. 295 / p. 202)
17. Divers développements autour des formes quadratiques (p. 302 / p. 212)
18. 5$^{e}$ méthode de résolution du problème de Cauchy. Utilisation d'un développement

selon les puissances décroissantes. (p. 332 / p. 233)

19. la notion de racine d'une équation (p. 334 / p. 235)

Le plan reste très proche de celui des manuscrits de 1878-79 et 1880-81. Les notations restent d'ailleurs inchangées, ce qui facilite la comparaison. La principale différence est contenue dans la partie *Divers développements autour des formes quadratiques*, qui en fait ne nous intéresseront pas : ils portent sur diverses techniques de transformations effectives des formes quadratiques et sont plus à rapprocher du travail – extrêmement fécond – de Kronecker sur les formes bilinéaires [22]. Il s'agit du seul cours dont nous connaissons la durée hebdomadaire [23] : Kronecker enseignait ce cours en deux séances, l'une de quatre heures et la seconde de deux heures.

**Les manuscrits des semestres d'hiver 1886-87** Pour ce semestre aussi nous avons pu comparer deux manuscrits : celui conservé à Strasbourg et celui offert par le docteur Lamla à Göttingen. Nous mettrons entre parenthèses d'abord la référence au manuscrit [Kronecker, 1887b], puis celle de [Kronecker, 1887a]. Dans [Kronecker, 1887a], on trouve dans la marge la date de la prise des notes : cela nous permettra d'évaluer le temps consacré par Kronecker à chaque partie. Les cours ont lieu le lundi, le mardi et le jeudi. La partie sur le théorème de Sturm entre le mardi 14 décembre 1886 et le 25 janvier 1887.

1. Démonstration avec la notion géométrique de continuité : obtention de la formule
$$\sum_{f(\xi)=0} sgn(f'(\xi)f_1(\xi))$$
pour $x_1 < \xi < x_2$. (p. 131 / p. 126)
2. Interprétation géométrique du théorème de Sturm (p. 143 / p. 133)
3. Nombre total de racines. Obtention de la formule $A-E = \sum sgn(c_{k-1}c_k)$ (p. 145 / p. 135)
4. Passage du cas $]-\infty;+\infty[$ au cas $[x_1;x_2]$ (p. 147 / p. 137)
5. La notion de racine d'une équation (p. 155 / p. 143)
6. Application aux équations biquadratiques et surface-discriminante (p. 166 / p. 152)
7. Travail préliminaire sur les fractions continues (p. 180 / p. 164)
8. Le problème d'interpolation de Cauchy (p. 187 / p. 166)
9. Énoncé 2$^{e}$ et 3$^{e}$ méthodes (p. 189 / p. 169)
10. 1$^{re}$ méthode de résolution du Problème de Cauchy. Attribuée à Liouville : utilisation directe de la décomposition en fraction continue de $\frac{f}{f_1}$ (p. 189 / p. 169)
11. Exemple : $f(x) = (x-1)(x-2)(x-3)(x-\xi)$ et $f_1(x) = x^2-6x+11$ (p. 194 / p. 172)
12. Résolution du problème de Cauchy sous la forme d'un déterminant (p. 195 / p. )
13. 4$^{e}$ méthode de résolution du problème de Cauchy. Attribuée à Jacobi : grâce aux formules d'Euler, on se retrouve avec un système de $\mu$ équations à résoudre. Dans cette méthode apparaît des déterminants équivalents au déterminant $D_\mu$. (p. 198 / p. 173)
14. 5$^{e}$ méthode de résolution du problème de Cauchy. Utilisation d'un développement selon les puissances décroissantes. (p. 209 / p. 183)

22. Voir par exemple [Kronecker, 1866], [Kronecker, 1868], [Kronecker, 1872], [Kronecker, 1874c], [Kronecker, 1874a] ou [Kronecker, 1874b].
23. Voir p. 56.

15. Utilisation de la 4$^{e}$ méthode pour le théorème de Sturm (p. 210 / p. 184)
16. Formules de Sylvester (p. 220 / p. 193)
17. Démonstration du Théorème de Sturm par la Méthode de Jacobi-Hermite (p. 224 / p. 195)
18. Loi d'inertie de Sylvester (p. 225 / p. 196)

Ce cours contient, à quelques détails près, l'ensemble du contenu des cours précédents. Mais ici tout a été entièrement réécrit : l'ordre d'exposition est globalement inchangé, mais les notations, l'introduction des notions et même certaines démonstrations sont complètement différentes. L'une des différences fondamentales est l'utilisation systématique des variétés (*Mannigfaltigkeiten*) pour décrire l'ensemble des solutions du problème de Cauchy, donc pour décrire un ensemble de fonctions.

**Le manuscrit du semestre d'hiver 1888-89** Les cours ont lieu cette année-là le lundi, le mardi, le jeudi et le vendredi, donc il y a une séance de plus par semaine par rapport à l'année précédente, où la part réservée à Sturm est la plus importante, tout au moins en nombre de pages. Il s'agit pourtant du cours dans lequel la partie consacrée à Sturm est proportionnellement la plus petite. À cela plusieurs explications : tout d'abord, la partie qui traite de la théorie des *Gattungen* et des équations abéliennes a dans ce manuscrit une place prépondérante. Ensuite, si on le compare à celui de l'année précédente, on voit que Kronecker y a réintroduit la théorie des caractéristiques, tout au moins celle de deux fonctions. Nous verrons plus tard que cette dernière peut « remplacer » certaines propriétés qu'il démontre habituellement dans son chapitre sur le théorème de Sturm. Enfin, il est possible que les quatre cours de 1888-89 correspondent à des séances d'une heure, alors que celui de 1886-87 à des séances de deux heures.

1. Démonstration avec la notion géométrique de continuité : obtention de la formule
$$\sum_{f(\xi)=0} sgn(f'(\xi)f_1(\xi))$$
pour $x_1 < \xi < x_2$ (p. 120)
2. Interprétation géométrique du théorème de Sturm (p. 124)
3. Nombre total de racines. Obtention de la formule $A - E = \sum sgn(c_{k-1}c_k)$ (p. 125)
4. Démonstration du Théorème de Sturm par la Méthode de Jacobi-Hermite (p. 126)
5. Loi d'inertie de Sylvester (p. 128)
6. Formules de Sylvester (p. 129)

On a bien ici une version minimale du chapitre. Ce que l'on peut tout de même constater, c'est la présence des deux démonstrations ; celle reposant sur la définition *géométrique* de la continuité, qui est d'ailleurs accompagnée d'une interprétation graphique du théorème, et celle utilisant les formes quadratiques et qu'il déduit de la loi d'inertie.

**Le manuscrit du semestre d'hiver 1890-91** Les cours ont lieu le lundi, le mardi et le jeudi : un jour de moins que l'année précédente. Il est difficile d'en tirer une quelconque

conclusion, car nous ignorons la durée de ces cours [24], mais il s'agit très probablement d'un cours de six heures. Il s'agit là de son dernier cours sur la théorie des équations algébriques, et le manuscrit contient un sommaire détaillé qui permet de se repérer facilement dans l'ensemble du cours. La partie consacrée à Sturm est abordée le 19 décembre 1890 et se termine le 12 janvier 1891. Cette partie n'est ici qu'une sous-partie de la partie intitulée *Die Theorie der algebraischen Gleichungen vom $n^{ten}$ Grade*, mais son importance est mise en avant par le fait qu'elle est la seule sous-partie à être soulignée dans le sommaire. La structure même de la première partie de ce cours – découpée en fonction du degré de l'équation étudié : quadratique, cubique, puis de degré $n$ – ne permettait pas d'isoler le théorème de Sturm. Ce dernier « remplace » en effet dans cette partie les théorèmes donnant de manière effective les racines pour les équations de degré deux et trois.

1. Le déterminant $|s_{p+q}|$ (p. 170)
2. Transformation de Jacobi (p. 171)
3. Loi d'inertie de Sylvester (p. 175)
4. Démonstration avec la notion géométrique de continuité : obtention de la formule
$$\sum_{f(\xi)=0} sgn(f'(\xi)f_1(\xi))$$
pour $x_1 < \xi < x_2$ (p. 184)
5. Travail préliminaire sur les fractions continues (p. 190)
6. Le problème d'interpolation de Cauchy (p. 191)
7. Énoncé 2$^{e}$ et 3$^{e}$ méthode (p. 192)
8. 1$^{re}$ méthode de résolution du Problème de Cauchy. Attribuée à Liouville : utilisation directe de la décomposition en fraction continue de $\frac{f}{f_1}$ (p. 192)
9. 4$^{e}$ méthode de résolution du problème de Cauchy. Attribuée à Jacobi : grâce aux formules d'Euler, on se retrouve avec un système de $\mu$ équations à résoudre. Dans cette méthode apparaissent des déterminants équivalents au déterminant $D_\mu$. (p. 194)
10. Nombre total de racines. Obtention de la formule $A - E = \sum sgn(c_{k-1}c_k)$ (p. 197)
11. Passage du cas $]-\infty;+\infty[$ au cas $[x_1;x_2]$ (p. 198)

Nous voyons qu'ici l'ordre dans lequel Kronecker traite le chapitre est à l'inverse de ce qu'il a proposé les années précédentes. Il donne dans un premier temps le théorème à l'aide des formes quadratiques, puis il fournit la démonstration à l'aide de la continuité. On voit bien ainsi l'indépendance des deux expositions : l'une ne suit pas forcément l'autre. La seconde différence notable, c'est l'absence de toute interprétation géométrique. En fait, celle-ci est présente, mais elle sera inclue dans le chapitre suivant sur la théorie des caractéristiques.

#### 6.2.3.2 Conclusion

En examinant ainsi, de façon globale, le contenu et l'ordre d'exposition du chapitre sur le théorème de Sturm dans le cours de Kronecker sur la théorie des équations algébriques, on voit apparaître des permanences, mais aussi des modifications importantes.

24. Il semble que classiquement, la durée des cours pouvait être de deux heures, quatre heures ou six heures par semaine : voir les propos de E. Netto dans [Kronecker, 1894] p. V et plus haut p. 56.

Dans tous les manuscrits, il y a une double exposition du théorème ; d'une part, celle qui repose sur la notion géométrique de continuité, et qui amène Kronecker à donner une interprétation graphique du théorème de Sturm ; d'autre part deux présentations plus *algébriques*, l'une reprenant le travail de Sylvester et conduisant aux formules que ce dernier avait annoncées dès 1839, l'autre étant issue des travaux de Jacobi et de Hermite sur les formes quadratiques.

L'ordre dans lequel il expose la version « géométrique » et les versions « algébriques » s'inverse lors du dernier cours. Jusque-là, il avait toujours commencé par l'exposé du théorème à partir de considérations sur la notion de *continuité géométrique*. Nous verrons que, plus ou moins explicitement, Kronecker justifie cet ordre par des raisons pédagogiques (l'aspect heuristique de cette présentation), mais considère aussi que la présentation de Hermite est, pour des raisons que nous exposerons plus tard, « meilleure » que les autres. Cette inversion dans le dernier cours conduit à mettre l'accent sur la version *géométrique*, qui est en fait, comme nous le verrons dans la partie suivante, issue d'une vision plus *topologique* du théorème. Ce dernier point de vue est celui qui permettra à Kronecker d'introduire la théorie des caractéristiques : la particularité de ce dernier cours est justement le retour à une exposition complète de cette théorie. Nous reviendrons dans la conclusion de cette partie sur la comparaison entre ces versions. En effet la « supériorité » de la démonstration de Hermite-Jacobi sur celle basée sur les fractions continues n'est en fait pas évidente pour Kronecker : même si elle lui permet de s'émanciper de toute référence à la continuité, nous verrons qu'il lui reproche d'être moins *générale*.

## 6.3 Énoncé et premier exemple

Il est maintenant temps de regarder précisément ce que dit le théorème de Sturm, qui permet donc de calculer le nombre de racines réelles comprises entre deux bornes $a$ et $b$ d'un polynôme.
Soit

$$P(x) = x^n + a_1 x^{n-1} + ... + a_n \in \mathbb{R}[x]$$

On construit la suite $(P_n)$ de polynômes de la façon suivante : on note $P_0 = P$, $P_1 = P'$ et ensuite on effectue la division euclidienne de $P_0$ par $P_1$. On notera $P_2$ l'opposé du reste obtenu. On recommence ainsi le processus jusqu'à obtenir un reste de degré 0 (ce qui arrive en moins de $n$ itérations). On a ainsi obtenu une suite finie, appelée *suite de Sturm*, $S(x) = (P_0(x), P_1(x), ..., P_k(x))$. Si on note $V(x)$ le nombre de changements de signes dans la suite $S(x)$, alors le nombre de racines de $P$ sur $[a; b]$ est $V(a) - V(b)$. Ce théorème s'étend d'ailleurs facilement au cas où $a = -\infty$ et $b = +\infty$.

Reprenons l'exemple que donne Sturm lui-même dans son mémoire de 1835 :

$$P(x) = x^3 - 2x - 5$$

On a alors les deux premiers termes de la suite :

$$P_0(x) = x^3 - 2x - 5$$
$$P_1(x) = 3x^2 - 2$$

On effectue la division euclidienne de $3P_0$ par $P_1$. Le coefficient 3 est là pour que l'on puisse conserver des coefficients entiers : comme le justifie Sturm, la multiplication d'un des termes de la suite par un coefficient positif ne change rien aux signes de ces termes. Or seul le signe de ces termes importe ici. On obtient alors :

$$3x^3 - 6x - 15 = (3x^2 - 2)x - (4x + 15)$$

et ainsi $P_2(x) = 4x + 15$. On effectue ensuite la division euclidienne de $4V_1$ par $V_2$ et on obtient :

$$12x^2 - 8 = (4x + 15)(3x - 45) - (-643)$$

et ainsi $P_3(x) = -643$. Comme $P_3$ est un réel différent de zéro, on peut déjà affirmer que $P_3$ n'a pas de racines doubles. Pour obtenir toutes les racines de $P$, on applique le théorème précédent avec $a = -\infty$ et $b = +\infty$. On obtient les deux suites de signes suivantes :

$$S_{-\infty} = \{-, +, -, -\}$$
$$S_{+\infty} = \{+, +, +, -\}$$

Ainsi le nombre de racines réelles est :

$$V_{-\infty} - V_{+\infty} = 2 - 1 = 1$$

Il y a donc une unique racine réelle.

Sturm commence ainsi l'article de 1829 qui expose pour la première fois ce théorème :

> La résolution des équations numériques est une question qui n'a cessé d'occuper les géomètres depuis l'origine de l'algèbre jusqu'à nos jours. La véritable difficulté de ce problème se réduit, comme on le sait, à trouver, pour chaque racine réelle de l'équation proposée, deux limites, entre lesquelles cette racine soit seule comprise. [Sturm, 1829a]

On voit donc qu'il s'agit d'un procédé effectif, algorithmique, qui répond à une question qui s'énonce simplement et qui a été longtemps débattue. Voyons maintenant comment la communauté scientifique a reçu ce théorème.

## 6.4 Réception et diffusion du théorème de Sturm

Comme le montre H. Sinaceur, ce théorème a une popularité immédiate : il est commenté et repris dès 1829, et donc bien avant que Sturm n'en donne une preuve [25], et apparaît au XIX^e siècle dans les ouvrages généraux d'algèbre, aussi bien en France que dans le reste de l'Europe et aux États-Unis [26]. En utilisant la bibliographie faite par H. Sinaceur [27], on voit qu'il se publie 15 mémoires en rapport avec le théorème de Sturm entre 1832 et 1850, 45 entre 1851 et 1880, 28 entre 1881 et 1900 14 entre 1901 et 1950, et enfin 20 entre 1951 et 1982. Si l'on regarde plus précisément la distribution des mémoires entre 1830 et 1930, on obtient le tableau suivant :

| Période | 1830-40 | 1841-50 | 1851-60 | 1861-70 | 1871-80 |
|---|---|---|---|---|---|
| Nombre de Mémoires | 7 | 8 | 23 | 11 | 12 |

| Période | 1881-90 | 1891-00 | 1901-10 | 1911-20 | 1921-47 | 1948-86 |
|---|---|---|---|---|---|---|
| Nombre de Mémoires | 11 | 15 | 12 | 1 | 0 | 21 |

On voit, à partir du tableau précédent, l'intérêt pour ce théorème croître jusqu'aux milieu des années 1850, pour ensuite se stabiliser jusqu'à la fin du XIX^e siècle, et s'évanouir complètement dans les années 1930. Il s'agit ici des mémoires de recherche : nous aurions une vision très différente de la présence du théorème de Sturm si l'on avait regardé les ouvrages généraux, à portée plus pédagogique, qui présentent ce théorème. Il aurait d'ailleurs fallu dans ce cas ne pas se contenter des première éditions, mais prendre en compte les rééditions régulières de ces ouvrages et regarder plus précisément leurs usages dans les diverses grandes institutions de formation. Par ailleurs, cet argument quantitatif n'est utile que si l'on regarde plus finement le contenu de ces articles. On peut distinguer quatre périodes.

La première, étudiée par H. Sinaceur dans [Sinaceur, 1991], s'étend du début des travaux algébriques de Sturm à l'algébrisation du théorème par Cayley (2 mémoires), Hermite (5 mémoires) et Sylvester (13 mémoires) dans les années 1860. Nous allons en donner les grandes lignes dans les prochains chapitres.

La seconde, qui est au centre de notre étude, commence avec les articles de 1869 de Kronecker et se termine à la fin du XIX^e siècle. Kronecker publie pendant cette période pas moins de 6 articles, la plupart présentant ou faisant intervenir sa théorie des caractéristiques, qui est par ailleurs présente, ou au moins citée, dans certains des articles publiés sur le théorème de Sturm.

La troisième correspond au déclin de l'intérêt de la communauté mathématique pour le théorème, tout au moins en ce qui concerne la recherche, du fait d'un déplacement de la

25. Il ne sera d'ailleurs pas le premier à en publier une démonstration. Crelle propose en 1835 à partir de [Sturm, 1829a] une preuve qui par ailleurs ressemble en de nombreux points à celle que Sturm publiera la même année. Voir [Crelle, 1835].

26. Par exemple dès 1832 [Mayer et Choquet, 1832] en France, en 1844 aux États-Unis [Davies, 1844] ou en 1895 [Weber, 1895] en Allemagne.

27. [Sinaceur, 1991] p. 438.

notion d'algèbre comme « théorie des équations » à une algèbre plus abstraite.

La dernière montre, à partir des années 1950, un regain d'intérêt qui s'explique essentiellement par son utilisation en logique et en algèbre réelle [28].

Examinons maintenant la démonstration du théorème que Sturm a donnée de son théorème et la façon dont Kronecker l'a reprise.

28. Voir [Sinaceur, 1991].

# Chapitre 7

# Démonstration du théorème par Sturm


7.1 La démonstration de 1835 . . . . . **193**
7.2 Première démonstration du théorème de Sturm par Kronecker . . . . . **199**
7.2.1 Introduction . . . . . 199
7.2.2 Une démonstration sans considération de continuité ? . . . . . 204
7.2.3 Recherche du nombre total de solutions d'une fonction entière . 219
7.2.4 Un exemple : l'équation biquadratique . . . . . 221
7.3 Conclusion . . . . . **230**


En 1835, six ans après après avoir énoncé son théorème, et quelques années après la publication de plusieurs démonstrations d'autres mathématiciens [1], Sturm donne sa propre preuve dans un *Mémoire sur la résolution des équations numériques* [Sturm, 1835]. Comme point de départ à notre travail, nous allons donner cette démonstration de Sturm, avant de regarder ensuite comment Kronecker la transcrit et l'interprète dans son cours.

## 7.1 La démonstration de 1835

Nous allons donc reprendre, point par point, la démonstration que donne Sturm de son théorème en 1835 dans son mémoire [Sturm, 1835]. En voici le plan général :

1. Les six premiers chapitres du mémoire, qui constitue l'essentiel de ce qui nous intéresse ici, énoncent le théorème et donnent sa démonstration.
2. Les six chapitres suivants présentent quelques astuces calculatoires pour abréger les applications du théorème.
3. Les chapitre 13 et 14 traitent, dans certains cas, du dénombrement des racines imaginaires : il s'agit ici aussi de simplifier, lorsque cela est possible, la détermination du nombre de racines, réelles et imaginaires.

1. Voir plus haut.

4. Le chapitre 17 donne 3 exemples sur lesquels Sturm fait fonctionner l'ensemble de ce qu'il a présenté.
5. Le chapitre 18 étend le théorème au cas des racines multiples.

Nous allons donc regarder en détail les six premiers paragraphes.

Pour faciliter la lecture de cette démonstration en comparant avec le texte original, nous noterons le paragraphe du mémoire de Sturm qui correspond à la partie traitée par le symbole §1 suivi du numéro du chapitre.

§1 Prenons comme le fait Sturm l'équation suivante :

$$V = Nx^m + Px^{m-1} + Qx^{m-2} + ... + Tx + U = 0$$

où $V$ est un polynôme dont les coefficients sont réels (*équation numérique)*. Soit $V_1 = V'$, on effectue la division euclidienne en modifiant les restes comme on l'a vu précédemment et on obtient :

$$\begin{aligned} V &= q_1 V_1 - V_2 \\ V_1 &= q_2 V_2 - V_3 \\ ... &= ... \\ V_{r-2} &= q_{r-1} V_{r-1} - V_r \end{aligned}$$

La suite des degrés de $V_n$ étant strictement décroissante, on a facilement que $V_r$ est un nombre réel. §2 Pour éviter le cas de racines multiples, Sturm choisit dans un premier temps de supposer $V_r \neq 0$.

Il énonce ensuite son théorème : le nombre de racines réelles de $V$ comprises entre $A$ et $B$ est la différence :

$$\begin{aligned} &\textit{Nombre de changements de signe dans la suite } (V(A), V_1(A), ..., V_r(A)) \\ -&\textit{Nombre de changements de signe dans la suite } (V(B), V_1(B), ..., V_r(B)) \end{aligned}$$

§3 Le principe consiste à examiner comment varie le nombre de changements de signe dans la suite $(V(x), V_1(x), ..., V_r(x))$ lorsque $x$ varie [2].

Un changement apparaît dans la suite des signes uniquement si $V$ ou l'un des $V_i$ change de signe, et donc s'annule. Sturm utilise ici implicitement le fait que les $V$ soient

2. Nous ferons plus loin quelques remarques sur ce principe que Kronecker reprendra à son compte dans sa présentation du théorème de Sturm et de la théorie des caractéristiques.

continues et le théorème de Bolzano. Il distingue deux cas : soit $V = 0$, soit $V_i = 0$.

Si $V(c) = 0$

On a alors par hypothèse §4 que $V_1(c) \neq 0$. Cette même continuité nous permet d'affirmer que l'on peut trouver $u > 0$ suffisamment petit pour que $V_1$ reste de signe constant sur $[c-u; c+u]$. On peut alors appliquer la formule de Taylor :

$$\begin{aligned} V(c+u) &= V(c) + V'(c)u + \frac{V''(u)}{2}u^2 + \frac{V'''(c)}{6}u^3 + ... \\ &= u\left[V'(c) + \frac{V''(c)}{2}u + \frac{V'''(c)}{6}u^2 + ...\right] \end{aligned}$$

et conclure que pour $u > 0$ suffisamment petit $V(c+u)$ aura le même signe que $V'(c)$, et donc que $V'(c+u)$.

Ainsi $V$ a le même signe que $V_1$ pour $c+u$. De même, en posant $u = -u$ on voit que $V$ a le signe contraire de celui de $V_1$ pour $c-u$. On peut résumer la situation avec les deux tableaux suivants :

| | $V$ | $V_1$ |
|---|---|---|
| $c-u$ | - | + |
| $c$ | 0 | + |
| $c+u$ | + | + |

ou

| | $V$ | $V_1$ |
|---|---|---|
| $c-u$ | + | - |
| $c$ | 0 | - |
| $c+u$ | - | - |

On voit que dans les deux cas on perd une « variation de signe » lorsque l'on passe de $c-u$ à $c+u$. Si aucun autre $V_i$ $(i = 2, ..., r)$ ne s'annule en $c$, il y a donc lorsque $x$ croit de $c-u$ à $c+u$ la perte d'un changement de signe.

Si $V_i(b) = 0$ où $1 \leq i \leq r-1$

§5 Comme on a l'égalité

$$(E) \qquad V_{i-1} = q_i V_i - V_{i+1}$$

on ne peut avoir $V_{i-1}(b) = 0$ ni $V_{i+1}(b) = 0$. En effet, par construction on aurait tous les $V_i(b)$ égaux à 0, et donc aussi $V_r$, ce qui n'est pas possible. Par $(E)$ on a de plus :

$$V_{i-1}(b) = -V_{i+1}(b)$$

donc $V_{i-1}(b)$ et $V_{i+1}(b)$ sont de signes contraires. Pour $u > 0$ suffisamment petit on a

| | $V_{i-1}$ | $V_i$ | $V_{i+1}$ |
|---|---|---|---|
| $b-u$ | + | $\pm$ | - |
| $b$ | + | 0 | - |
| $b+u$ | + | $\mp$ | - |

Dans tous les cas, lors du passage de $b-u$ à $b+u$ il y aura pas de variation dans le nombre de changements de signe.

**Remarque** Le résultat serait identique si plusieurs $V_i$ intermédiaires s'annulaient pour un même $b$. Si une valeur annule $V$ et un $V_i$ intermédiaire, alors il y a comme précédemment *une* perte de changements de signes.

**Conclusion** Si $x$ croît de $A$ à $B$, la suite $(V(x), V_1(x), ..., V_r(x))$ perdra autant de changements de signe qu'il y aura de valeurs pour lesquelles $V$ s'annulera. §6

Voilà donc la démonstration que Sturm propose en 1835. Sur quoi repose-t-elle ? En grande partie sur le théorème de Bolzano [3] et sur une méthode consistant à effectuer de « petits accroissements », tous deux appliqués sur une suite de fonctions auxiliaires que Sturm construit à partir de l'algorithme d'Euclide, mais dont il n'extrait pas explicitement les propriétés fondamentales qui permettent à sa méthode de fonctionner, même s'il est, comme le seront aussi ses principaux lecteurs, parfaitement conscient de ces conditions. Nous verrons d'ailleurs que certaines de ces conditions sont à l'origine de la mise au point de son algorithme d'Euclide modifié. Ces conditions sont données, par exemple, par Weber dans le premier tome du *Lehrbuch* :

Soit une suite de polynômes $f(x), f_1(x), ..., f_m(x)$ où $f$ est supposé sans racine multiple et $[a; b]$ un intervalle tel que ni $a$ ni $b$ ne soit une racine de $f$. Alors pour être utilisé dans le théorème de Sturm cette suite doit respecter les conditions suivantes [4] :

1. Deux fonctions consécutives ne s'annulent pas pour une même valeur de l'intervalle $[a; b]$.
2. La fonction $f_m(x)$ ne s'annule pour aucune valeur de l'intervalle $[a; b]$.
3. Si $f_i(\alpha) = 0$, alors $f_{i-1}(\alpha)$ et $f_{i+1}(\alpha)$ sont de signes contraires.
4. Si $f(\alpha) = 0$, alors $f_1(\alpha)$ a le même signe que $f'(\alpha)$.

À partir de ces conditions, un grand nombre de *suites de Sturm* peuvent être construites. Sturm le savait dès 1829, et il affirmait que l'on pouvait « en former une infinité d'autres qui jouissent des même propriétés » [5]. On peut voir dans la seconde partie du mémoire de 1835, que nous ne développerons pas ici, d'autres considérations sur ces *suites de Sturm*, et

3. Nous avons vu le rapport de Kronecker avec l'utilisation du travail de Bolzano p. 91.
4. Une analyse de ces conditions est faite dans [Sinaceur, 1991] p. 98-99.
5. Voir [Sturm, 1829a] p. 421.

c'est *sur* ces suites, comme nous le verrons, que se concentre, avant les travaux d'Hermite et de Kronecker, l'algébrisation du théorème.

Le choix de Sturm de prendre comme suite de fonctions les restes modifiés de l'algorithme d'Euclide peut sembler, à première vue, miraculeux. Et de fait, comme le rapporte H. Sinaceur, certains de ses contemporains ont pu qualifier son travail de « divination », « d'illumination de l'esprit » [6]. Mais réduire ce choix à quelque inspiration « géniale » est pour le moins trompeur. Hourya Sinaceur le montre bien : le travail de Sturm trouve sa source dans ceux de Fourier et de Lagrange.

**Origines** Les origines du travail de Sturm sont multiples, mais la règle de Descartes [7] peut être considérée comme un « ancêtre » du théorème de Sturm. En effet ce dernier ne se contente pas que rendre plus précise cette règle, il en conserve la forme : il s'agit bien encore d'étudier le nombre de changements dans une suite de signes.

Par ailleurs, Sturm est bien évidemment un lecteur de Lagrange, et il a une très bonne connaissance du *Traité de la résolution des équations numériques* [8]. Il commence d'ailleurs son mémoire de 1835 ainsi :

> La résolution des équations numériques est une question qui n'a pas cessé d'occuper les géomètres, depuis l'origine de l'algèbre jusqu'à nos jours. Nous ne rappelons pas tous les procédés qui ont été proposés pour la détermination des racines réelles des équations. Lagrange, le premier, a donné pour cet objet une méthode rigoureuse ; [Sturm, 1835] p. 273

Pour Lagrange, l'objet de l'algèbre est la résolution *générale* des équations :

> Newton a appelé l'Algèbre Arithmétique universelle. Cette dénomination est exacte à quelques égards ; mais elle ne fait pas assez connaître la véritable différence qui se trouve entre l'Arithmétique et l'Algèbre. Le caractère essentiel de celle-ci consiste en ce que les résultats de ses opérations ne donnent pas les valeurs individuelles des

6. Voir [Sinaceur, 1991] p. 108. Ces qualificatifs ont été donnés au théorème de Sturm par J.M. Duhamel, mathématicien et physicien qui deviendra professeur à l'école polytechnique.

7. On trouve cette règle dans la *Géométrie* de Descartes :

> On connoist aussy de cecy combien il peut y auoir de vrayes racines, & combien de fausses en chasque Equation. A sçauoir il y en peut auoir autant de vrayes, que les signes + & − s'y trouuent de fois estre changés ; & autant de fausses qu'il s'y trouve de fois le signe +, ou deux signes − qui s'entresuiuent.[Descartes, 1987] p. 446

C'est-à-dire que nombre de racines positives (les racines « vrayes ») d'une équation $P(x) = 0$ où $P \in \mathbb{R}[X]$ est au plus égal aux nombre de changements de signes dans $P$, une fois bien sûr que $P$ a été réduit et ordonné. La partie de la règle concernant les racines négatives (les racines « fausses ») est fausse, à moins de faire apparaître les termes manquants et d'attribuer un signe (peu importe lequel) aux coefficients nuls.

Pour une démonstration de cette règle, il faudra attendre Gauss, qui par ailleurs l'attribue à Harriot : ce théorème, tout au moins la partie concernant les racines positives, peut se deviner dans le traité d'algèbre d'Harriot *Artis analyticae praxis ad aequationis algebraicas resolvendas* paru en 1631. (Voir [Gagneux, 1989]). Gauss a en effet proposé dans le *Journal de Crelle* de 1828 une démonstration rigoureuse de ce théorème, en l'accompagnant du résultat complémentaire suivant : le nombre de racines positives a la même parité que le nombre de changements de signes.

8. Voir [Lagrange, 1826].

> quantités qu'on cherche, comme ceux des opérations arithmétiques ou des constructions géométriques, mais représentent seulement les opérations, soit arithmétiques ou géométriques, qu'il faudra faire sur les premières quantités données pour obtenir les valeurs cherchées ; je dis arithmétiques ou géométriques, car on connaît depuis Viète les constructions géométriques par lesquelles on peut faire sur les lignes les mêmes opérations que l'on fait en Arithmétique sur les nombres. L'Algèbre plane pour ainsi dire également sur l'Arithmétique et la Géométrie ; son objet n'est pas de trouver les valeurs mêmes des quantités cherchées, mais le système d'opérations à faire sur les quantités données pour en déduire les valeurs des quantités qu'on cherche, d'après les conditions du problème. [Lagrange, 1826] p. xxii

En ce sens, un « esprit algébrique » [9] souffle sur son travail et la méthode qu'il emploiera pour résoudre le problème du nombre de racines et qui consistera en des *opérations* sur les coefficients (les *premières quantités données*). Cette méthode, qui utilise l'équation aux carrés des différences [10], permet, avant le théorème de Sturm, d'apporter une réponse complète au problème. Mais cette méthode présente un défaut majeur :

> Cette méthode, considérée sous un point de vue purement théorique, ne laisse rien à désirer du côté de la rigueur. Mais l'application, la longueur des calculs nécessaires pour former l'équation aux carrés des différences, et la multitude des substitutions qu'on peut avoir à effectuer, la rendent presque impraticable ; [Sturm, 1835] p. 274

En effet, dans cette méthode, les calculs deviennent rapidement « monstrueux », et Lagrange lui-même renonce à l'utiliser pour des équations de degré supérieur à cinq. Ce problème d'effectivité est le premier que Sturm cherchera à résoudre. Pour cela, il s'inspire du travail de Fourier :

> Les différentes méthodes qui ont été proposées pour arriver à ce but sont trop connues pour qu'il soit nécessaire de les rappeler ici. Aucune ne peut être comparée, sous le double rapport de la simplicité et de l'exactitude, à celle que M. Fourier a depuis long-temps découverte, et qui est fondée sur une proposition générale, dont la règle des signes de Descartes n'est qu'un cas particulier.(...) C'est en m'appuyant sur les principes qu'il a posés, et en imitant ses démonstrations, que j'ai trouvé les nouveaux théorèmes que je vais énoncer. [Sturm, 1829a] p. 419

Les méthodes que Fourier emploie sont très différentes de celles de Lagrange [11] et font partie de ce qu'il appelle l'*analyse algébrique* [12]. En fait c'est l'objet (les équations algébriques) qui constitue le liant de cette « méthode » [13], et pour aborder cet objet, il a recours aussi bien à des méthodes issues de l'algèbre au sens où Lagrange l'entend, qu'à des méthodes mettant

9. L'expression est de H. Sinaceur [Sinaceur, 1991] p. 71.

10. Pour plus de détails sur cette méthode, qui trouve sa source dans les travaux de Waring, on pourra consulter par exemple [Mayer et Choquet, 1832] p. 478. Il est intéressant de constater que ce travail de Lagrange inspiré par Waring fait l'objet, comme on l'a vu, d'une section importante du cours de Kronecker sur les équations algébriques.

11. Il s'agit même d'une réponse au *Traité de la résolution des équations numériques*.

12. Voir par exemple [Fourier, 1830] p. 9-11.

13. Au XIX$^{e}$ siècle, l'*analyse algébrique* est quelquefois appelé *analyse des équations*. Pour Fourier, qui la décrit et en donne une trajectoire historique dans l'introduction de son *Analyse des équations déterminées*, il s'agit en fait d'un ensemble de méthodes, algébriques ou analytiques, mises au service de la résolution des équations.

en jeu des notions telles que la continuité. Mais le manque de *généralité* des théorèmes obtenus constitue pour Sturm le second écueil à éviter. Les commentaires dithyrambiques de son article de 1829 sont remplacés en 1835 par la mise en avant des limites du travail de Fourier :

> Mais ce théorème ne donnant pas précisément le nombre de ces racines, on peut être exposé, en l'appliquant, à chercher des racines dans des intervalles où il n'en existe pas, de sorte que de nouvelles règles sont nécessaires pour faire disparaître cette incertitude. [Sturm, 1835] p. 274

Le théorème de Sturm peut ainsi finalement être vu comme le résultat de la confrontation entre les travaux de Lagrange portés par une *volonté de généralité* et de *pureté algébrique* avec ceux de Fourier, utilisant aussi des méthodes analytiques, et recherchant *l'effectivité*. Voyons maintenant comment Kronecker reprend ce théorème dans son cours sur la théorie des équations algébriques.

## 7.2 Première démonstration du théorème de Sturm par Kronecker

### 7.2.1 Introduction

Le théorème de Sturm est, quelle que soit l'année, toujours présent dans les *Vorlesungen* de Kronecker. À partir de la liste du contenu de ces leçons [14], nous voyons que Kronecker propose trois présentations différentes du théorème de Sturm :

1. La première est, selon lui, basée sur des considérations « géométriques ».
2. La seconde utilise la décomposition en fraction continue. Couplée avec ce que l'on appellera le problème de Cauchy, elle compose la plus grande partie de son chapitre.
3. La dernière est basée sur les travaux de Hermite et de Jacobi sur les formes quadratiques.

Dans ce chapitre, nous allons présenter la première qui, le plus souvent [15], constitue l'introduction de Kronecker au théorème de Sturm. Les deux autres seront traitées dans les chapitres suivants.

#### 7.2.1.1 Une vision *géométrique*

La démonstration que nous allons présenter conserve, au cours des années, une grande stabilité. Nous allons regarder plus précisément les versions des années 1880-81, 1884-85, 1887-88 et 1890-91.

14. Voir p. 182.
15. C'est le cas dans tous les manuscrits sauf celui de 1890-91.

Kronecker commence systématiquement par se placer d'un point de vue qu'il qualifie de « géométrique ». Dans le manuscrit de 1880-81, Kronecker affirme qu'il renoue [16] :

> avec l'intuition géométrique [17], laquelle nous apprend que si une courbe $y = F(x)$ passe tantôt en dessous, tantôt au dessus de l'axe des abscisses, elle doit le couper, dans le cas où la courbe se déroule continuement. [Kronecker, 1881a] p. 104

Et ce passage est illustré par la figure suivante :

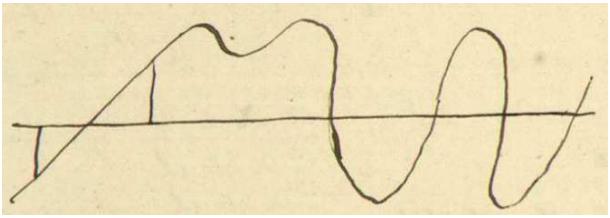


FIGURE 1 – [Kronecker, 1881a] p. 104

Kronecker souhaite répondre à la question : combien y-a-t-il de racines réelles entre deux bornes données ? Dans le manuscrit de 1886-87, il commence par rappeler que le théorème de Sturm n'est pas indispensable pour y répondre, mais qu'il s'agit sans doute de la réponse la plus *élégante* [18]. Nous pouvons en effet, dit-il, répondre à la question grâce à ce qu'il appelle des *interpolations géométriques* [19] : on remplace la résolution de l'équation $f(x) = 0$ par la recherche des points d'intersection entre la « courbe parabolique $y = f(x)$ » [20] et l'axe des abscisses. Il développe cette conception sur les cas simples des équations quadratiques et des équations cubiques. Il utilise le fait que la dérivée s'annule entre deux racines réelles, et il construit ainsi le discriminant : le signe de ce dernier donnera le nombre de racines réelles. (Figures 2 et 3).

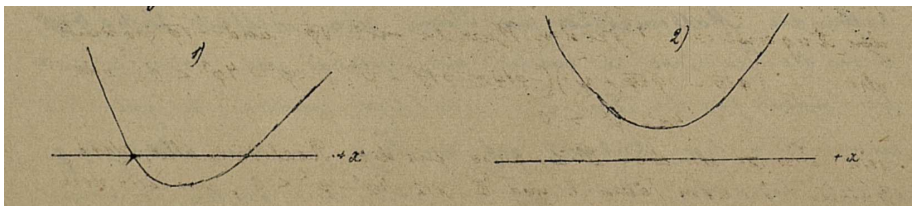


FIGURE 2 – [Kronecker, 1887a] p. 132

Il appuie ici sa démonstration sur les différentes représentations possibles de la

16. Wir knüpfen an die geometrische Anschauung an, welche uns lehrt, daß wenn die Curve $y = F(x)$, bald oberhalb, bald unterhalb der Abszissenaxe verläuft, sie dieselbe schneiden muß, falls die Curve stetig verläuft.

17. Le terme utilisé par Kronecker est celui de *Anschauung*, c'est-à-dire l'*intuition* dans la terminologie Kantienne. On peut aussi penser au concept d'*intuition* (*Anschauung*) chez le pédagogue allemand du début du XIX$^{e}$ siècle Johann Heinrich Pestalozzi (1746-1827).

18. « Eleganz » [Kronecker, 1887b] p. 131.

19. [Kronecker, 1887b] p. 132 : « Man kann aber die oben genannten Fragen, ohne von dem Sturmschen Satz Gebrauch zu machen, auf anderen weniger eleganten, aber theoretisch einfacheren Wege beantworten, wenn man sich der geometrischen Interpolation bedient ».

20. [Kronecker, 1887b] p. 132 : « die parabolische Curve $y = f(x)$ ».

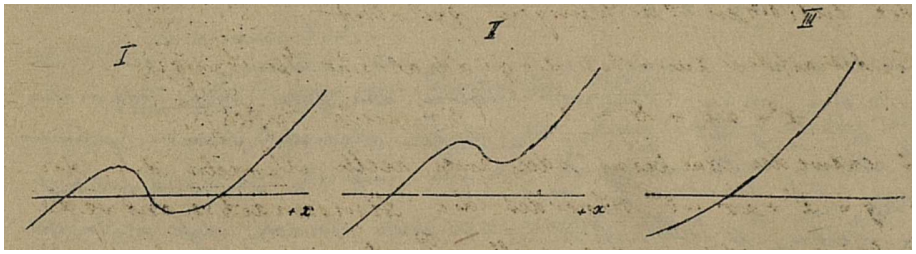



FIGURE 3 – [Kronecker, 1887a] p. 133

fonction, et il ne souhaite d'ailleurs pas prolonger ce processus au-delà du troisième degré.

Il ne s'agit pas là d'une méthode « géométrique » – Il reprend en grande partie des méthodes algébriques que l'on trouve chez Lagrange – mais plutôt d'une présentation géométrique, au sens où elle fait appel à la représentation des courbes dans un repère du plan. Plus encore, c'est la « présentation géométrique » des racines qui est au cœur de la preuve qu'il propose du théorème [21] :

> Nous voulons maintenant, pour construire la proposition de *Sturm* sur cette intuition géométrique des racines de $F(x) = 0$, définir un symbole, que nous désignerons par $[A]$ et qui doit représenter $-1, 0$ ou $+1$, suivant que la grandeur supposée réelle $A$ est négative, égale à zéro, ou positive. [Kronecker, 1881a] p. 104

#### 7.2.1.2 Racines, continuité géométrique et continuité arithmétique

Nous l'avons vu, le concept de racine est fondamental dans le travail de Kronecker sur la théorie des équations, et nous avons vu dans la partie précédente que dans une première approche, il considérait en fait que la donnée d'un intervalle aussi petit que l'on veut dont les bornes soient rationnelles et où la fonction prend des signes opposés en chacune de ces bornes pouvait correspondre à la notion de racine.

L'utilisation de celle-ci évolue au cours des manuscrits, mais même lorsque cela n'est pas explicité, Kronecker aborde toujours cette démonstration dans l'état d'esprit suivant [22] :

> Nous voulons supposer dans la suite que la notion de racine d'une équation est déjà définie. Nous reviendrons encore sur les justifications rigoureuses et les fondements de ces notions plus tard indépendamment des développements qui suivent maintenant. [Kronecker, 1881a] p. 104

Quel que soit le manuscrit, il cherche bien à trouver les racines réelles (*reellen Wurzeln* [23]) de l'équation. Le sens et la nature de ce qu'il entend par *racine* d'une équation seront,

21. Wir wollen nun, um auf diese geometrische Anschauung der Wurzele von $F(x) = 0$ den *Sturm*'schen Satz aufzubauen, ein Zeichen definiren, welches wir durch $[A]$ bezeichnen wollen und welches $-1, 0$ oder $+1$ vorstellen soll, je nachdem die als reell vorausgesetzte Größe $A$ negativ, gleich null, oder positiv ist.

22. Wir wollen in Folgenden annehmen, der Begriff der Wurzeln einer Gleichung sei bereits definirt. Auf die strenge Begründung und Fundamentirung dieses Begriffes werden wir später noch unabhängig von den jetzt folgenden Entwicklungen zurückkommen.

23. [Kronecker, 1887b] p. 131.

selon les cours, reportés à un chapitre particulier [24] ou incorporés à un autre chapitre [25].

Kronecker considère ensuite la notion de continuité. Celle-ci est bien sûr liée fortement à la notion de racine, mais cependant Kronecker cherche, au fur et à mesure des manuscrits, à dissocier le plus possible ces deux concepts. Dans ses leçons du semestre d'hiver 1880-81, il donne deux conceptions de la continuité. La première, il la qualifie de « géométrique ». Gauss lui-même, dit-il, [26] « n'avait que l'intuition géométrique et le concept de la continuité géométrique ». Cette continuité semble correspondre à la *possibilité* de l'intersection de deux courbes, avec dans le cas des racines d'un polynôme, à l'intersection des deux courbes algébriques $y = 0$ et $y = f(x)$.

On retrouve ici en partie le point problématique de la démonstration de Gauss du théorème fondamental de l'algèbre de 1799, lorsque ce dernier, parle d'une *ligne* qui pourrait sortir d'un cercle sans le couper [27] :

> L'absurdité de cela est évidente du fait qu'au point où l'on est entré dans le cercle, on a eu la première surface au dessus de soi, mais à la sortie au dessous ; c'est pourquoi nécessairement on a dû quelque part rencontrer la première surface, à savoir en un point de la première ligne. Par ailleurs, d'après ce raisonnement reposant sur les principes de la géométrie de position, lesquels ne sont pas moins valides que ceux de la géométrie des grandeurs, il s'ensuit seulement que si on entre dans le cercle par une branche de la première ligne, on peut sortir du cercle à un autre endroit, toujours en restant sur la première ligne. Mais il ne s'ensuit pas que le chemin suivi soit une ligne continue au sens qui lui est donné dans la géométrie supérieure. Mais il suffit ici que le chemin soit une ligne continue au sens commun, c'est-à-dire interrompue nulle part mais cohérente partout. [Gauss, 1799] p. 27

Nous verrons dans la dernière partie, lorsque nous examinerons la reprise par Kronecker des démonstrations de Gauss du théorème fondamental de l'algèbre, qu'il utilisera, tout au moins de façon implicite, ce type d'arguments. Ceux-ci sont en fait des arguments de type *topologique*, avec des *intérieurs* et des *extérieurs*.

La continuité *analytique* [Kronecker, 1881a] p. 104 : *analytisch*) est quant à elle définie [28]

24. Par exemple dans [Kronecker, 1881a].

25. Par exemple dans le chapitre sur le théorème de Sturm dans [Kronecker, 1887b] p. 134.

26. [Kronecker, 1881a] p. 104 : « Er besaß nur die geometrische Anschauung und den Begriff der geometrischen Stetigkeit. »

27. Ma traduction de « Hoc vero absurdum esse, inde patet, quod in puncto, ubi in circulum ingressus es, superficiem primam supra te habuisti, in egressu, infra ; quare necessario alicubi in superficiem primam ipsam incidere debulsti, sive in punctum lineae primae. — Ceterum ex hoc ratiocinio principiis geometriae situs innixo, quae haud minus valida sunt, quam principia geometriae magnitudinis, sequitur tantummodo, si in aliquo ramo lineae primae in circulum intres, te alio loco ex circulo rursus egredi posse, semper in linea prima manendo, neque vero, viam tuam esse lineam continuam in eo sensu, quo in geometria sublimiori accipitur. Sed hic sufficit, viam esse lineam continuam in sensu communi, i. e. nullibi interruptam sed ubique cohaerentem. »

28. analytisch werden wir die Funktion $F(x)$ als an der Stelle $x$ stetig bezeichnen, wenn $F(x_1) - F(x_2)$, wo $x_1 < x < x_2$ sei, kleiner gemacht werden kann als jede beliebig kleine Größe, dadurch daß man $x_1$ und $x_2$ beliebig nahe an $x$ heranrückt und wenn diese Differenz dann auch kleiner als diese beliebig kleine Größe bleibt. So einfach dieser Begriff der Stetigkeit ist, so ist er doch so modern, daß er bei Gauss noch nicht vorkommt.

par Kronecker dans son cours :

> Nous désignerons analytiquement la fonction $F(x)$ comme continue en $x$, si $F(x_1) - F(x_2)$, où $x_1 < x < x_2$, peut être faite plus petite que n'importe quelle grandeur, lorsque l'on rapproche $x_1$ et $x_2$ aussi près de $x$ que l'on veut et lorsque cette différence reste ensuite aussi plus petite que cette grandeur choisie aussi petite que l'on veut. Ce concept de continuité est si simple, mais aussi tellement moderne, qu'on ne le trouve pas encore chez Gauss. [Kronecker, 1881a] p. 104

La continuité analytique correspond donc pour Kronecker à la définition que Bolzano en donne dans la *Rein analytischer Beweis*[29] ou Cauchy dans son *Cours d'analyse*[30]. Une lecture rapide de ces passages peut laisser quelque peu perplexe : nous avons vu dans la partie précédente le point de vue de Kronecker sur le théorème de Bolzano, et nous savons donc déjà qu'il en rejette les conclusions.

Cette définition analytique est pourtant qualifiée de « simple » et « moderne », et Kronecker n'émet donc aucune critique à son encontre : il ne s'agit pas d'un rejet, comme il le fera du théorème de Bolzano[31]. Pourtant, à partir du manuscrit de 1884-85, Kronecker propose une démonstration sans « considération de continuité » (*Kontinuitätsbetrachtungen*). Si le problème n'est pas dans la définition du concept de fonction continue (*stetig*), comment et pourquoi Kronecker souhaite-t-il s'en passer ?

Dans ses leçons sur l'intégration[32], Kronecker décrit la notion de continuité de la façon suivante[33] :

> Cette notion n'est pas initialement extraite de l'arithmétique mais des applications de l'analyse à la géométrie et à la physique. Mais sa signification géométrique et physique est très obscure. Une courbe n'est, ni comme on la trace, ni comme on l'imagine en principe continue ; on ne peut toujours envisager – physiquement et mentalement – qu'un seul point déterminé à la fois. Et dans la nature d'un côté chaque action à distance semble inconcevable, mais d'un autre côté aucun changement de lieu dans l'espace, c'est-à-dire aucun mouvement, ne se laisse en général imaginer sans l'hypothèse d'une quelconque discontinuité dans le remplissage de l'espace. ([Kronecker, 1901] p 11)

Comme nous l'avons déjà vu[34], cette notion est bien une notion extérieur à l'« arithmétique générale » telle que la conçoit Kronecker. Pour lui, même Gauss se trompe, ou tout au moins travaille avec une version inadéquate de la continuité[35] :

29. Voir [Bolzano *et al.*, 1905].
30. Voir [Cauchy, 1821]. Pour plus de détails, voir p. 404.
31. Voir la partie précédente p. 91.
32. Voir [Kronecker, 1894].
33. Dieser Begriff ist kein ursprünglich arithmetischer, sondern er ist aus den Anwendungen der Analysis auf die Geometrie und Physik entnommen. Seine geometrische und physikalische Bedeutung ist aber sehr dunkel. Eine Curve ist weder, wie man sie zeichnet, noch wie man sie denkt, eigentlich continuirlich ; man kann immer nur einzelne bestimmte Punkte — körperlich wie geistig — ins Auge fassen. Und in der Natur scheint einerseits freilich jede Fernwirkung unfassbar, andererseits aber lässt sich ohne Annahme irgend welcher Unstetigkeit in der Raumerfüllung überhaupt keine Ortsveränderung im Raume, d. h. Bewegung, denken.
34. Voir section 2.2.3, p. 91.
35. In der Analysis handelt es sich immer nur um die Stetigkeit von Functionen. Dabei spielte aber

> Dans l'analyse il s'agit toujours de la continuité des fonctions. Mais la conception géométrique de la continuité y a joué depuis longtemps un rôle. Ainsi chez Gauss la continuité d'une fonction $y$ de $x$ est introduite dans le sens suivant
>
> "Si $x$ va de $x_0$ à $x_1$ , et que l'on prend pour $y$ en ces deux valeurs de la variable les valeurs de $y_0$ et $y_1$, alors il y a chaque fois entre $x_0$ et $x_1$ un $x'$, pour lequel la fonction obtient n'importe quelle valeur $y'$ choisie entre $y_0$ et $y_1$."
>
> On pense à une courbe qui connecte les deux points $(x_0, y_0)$ et $(x_1, y_1)$ et qui représente dans son cours ininterrompu le tracé des valeurs de la fonction. [Kronecker, 1901] pp. 11-12

Cette définition, qui est finalement le théorème des valeurs intermédiaires, Kronecker la rejette. Bolzano ne souhaitait pas non plus que cette propriété soit utilisée comme définition de la continuité, mais il voulait au contraire la *démontrer*[36]. Comme nous l'avons vu[37], pour Kronecker le problème est ailleurs : le théorème des valeurs intermédiaires n'a aucun sens étant donné que seul le dénombrable est accessible. Nous reviendrons sur cette notion dans la troisième partie lorsque nous examinerons l'intégration chez Kronecker[38], mais il est déjà claire que pour Kronecker, une démonstration *algébrique*, au sens où elle ferait partie de son arithmétique générale, devrait se passer de toute considération de continuité.

### 7.2.2 Une démonstration sans considération de continuité ?

Dans son cours de l'hiver 1884-85, Kronecker affirme au début de son chapitre sur le théorème de Sturm[39] :

> Pour ce théorème Sturm a donné une démonstration très simple, qui, comme tout ce traitement, repose essentiellement sur des considérations de continuité. Je vais vous en donner une autre tout à fait différente qui renonce à ces considérations de continuité, et ce en la construisant palier par palier, pourrait-on dire, sur des fondements triviaux. [Kronecker, 1885b] p. 237

Pourtant, Kronecker affirme qu'il souhaite « construire la proposition de *Sturm* sur cette intuition géométrique des racines de $F(x) = 0$ »[40].

---

lange Zeit hindurch die geometrische Auffassung der Stetigkeit eine Rolle. So kommt bei Gauss der Begriff Stetigkeit einer Function $y$ von $x$ nur in folgendem Sinne vor :

„Geht $x$ von $x_0$ bis $x_1$ , und nimmt $y$ für diese beiden Werthe der Variablen die Werthe $y_0$ und $y_1$ an, dann giebt es zwischen $x_0, x_1$ jedesmal ein $x'$ , für welches die Function den zwischen $y_0$ und $y_1$ beliebig gewählten Werth $y'$ erhält". Es ist dabei an eine Curve gedacht, welche die beiden Punkte $(x_0, y_0)$ und $(x_1, y_1)$ mit einander verbindet und in ihrem ununterbrochenen Laufe den Verlauf der Functionswerthe darstellt.

36. C'est ce qu'il fera en 1817 dans la *Rein analytischer Beweis* ([Bolzano *et al.*, 1905]).
37. Voir section 2.2.3, p. 91.
38. Voir p. 404.
39. Für diesen Satz hat Sturm auch eine sehr einfache Deduktion gegeben, die, wie alle diese Behandlungen, wesentlich an Kontinuitätsbetrachtungen anknüpft. Ich gebe Ihnen eine ganz andere, die von dieser Kontinuitätsbetrachtung absieht, und zwar habe ich allmählich diese auf, man kann sagen, trivialen Grund gebaut.
40. [Kronecker, 1881a] p. 104 : « um auf diese geometrische Anschauung der Wurzele von $F(x) = 0$ den *Sturm*'schen Satz aufzubauen ».

L'intuition géométrique dont il parle ne semble donc pas toucher directement à la continuité, mais plutôt aux *racines*, car il conclut, à la fin de sa démonstration [41] :

> Nous n'avons pas vraiment besoin dans notre démonstration de considération de continuité. [Kronecker, 1885b] p. 244

Examinons maintenant cette démonstration.

#### 7.2.2.1 Démonstration du théorème de Sturm

Kronecker commence par prendre une équation

$$x^n - c_1 x^{n-1} + ... \pm c_n = 0$$

où il fait bien remarquer que les $c_i$ sont des nombres réels (*reelle Zahlen* [42]).

**Une notation** Après cette introduction, Kronecker donne sa première démonstration du théorème de Sturm.

Il commence par définir la notation suivante :

$$sgn(a) = \begin{cases} \dfrac{a}{|a|} \text{ si } a \neq 0 \\ 0 \text{ si } a = 0 \end{cases}$$

En fait, cette notation est légèrement différente dans les cours de 1872-73, 1878-79 et 1880-81 :

$$[a] = \begin{cases} -1 \text{ si } a < 0 \\ 0 \text{ si } a = 0 \\ 1 \text{ si } a > 0 \end{cases}$$

Ce changement de notation est certainement une conséquence directe de son article de 1884 *Beweis des Reciprocitätsgesetzes für die quadratischen Reste* [43]. Il introduit en effet dans cet article trois notations :

Tout d'abord, « à la manière de Gauss » [44], il désignera par $[a]$ l'entier tel que

$$a - 1 < [a] \leq a$$

Ensuite par $|a|$, « à la manière de Weierstrass » [45], la valeur absolue de $a$.

Et enfin, nous pourrions ajouter « à la manière de Kronecker », par *sgn* $a$ l'entier $\dfrac{a}{|a|}$. Le cas $a = 0$, qui donne à chaque fois 0, n'est pas traité lors de la définition.

41. Wir haben in unserer Deduktion Kontinuitätsbetrachtungen eigentlich nicht gebraucht.
42. Voir [Kronecker, 1887b] p. 131.
43. Voir [Kronecker, 1884a] p. 500.
44. *Ibid.*
45. *Ibid.*

Il explique ainsi ces choix [46] :

> Dans mes leçons à l'université sur les équations algébriques et dans mes deux articles du *Monatsbericht* de février 1878 j'ai désigné le signe d'une grandeur réelle $a$ par $[a]$ ; Mais pour pouvoir donner au signe [ ] le sens que *Gauss* avait lui-même fixé, j'ai dû dans cet essai modifier le sens de cette notation. [Kronecker, 1884a] p. 500

On a les deux propriétés suivantes :

1. Si $a$ et $b$ sont de même signe alors

$$sgn(a+b) = \frac{1}{2}(sgn(a) + sgn(b))$$

2. Si $a$ et $b$ sont de signes contraires alors

$$sgn(a-b) = \frac{1}{2}(sgn(a) - sgn(b))$$

et

$$(1 - sgn(ab))sgn(a-b) = sgn(a) - sgn(b)$$

**Une première formule** Soit maintenant une fonction $F(x)$ :

— réelle de la variable réelle $x$ sur l'intervalle $[x_1; x_2]$

— finie et continue sur ce même intervalle

— dérivable et dont la dérivée a les mêmes propriétés sur ce même intervalle

Ces propriétés de la fonction $F$ sont présentes dans l'ensemble des manuscrits.

Kronecker remarque alors que cette fonction a sur cet intervalle le « caractère d'une fonction entière » [47], c'est-à-dire que ce qui suit pourra donc s'appliquer essentiellement (mais pas uniquement) aux polynômes.

Supposons maintenant que

$$sgn(F(x_1)) = -sgn(F(x_2))$$

On obtient facilement que :

$$sgn\left(\frac{F(x_2) - F(x_1)}{x_2 - x_1}\right) = -\frac{1}{2}sgn(F(x_1)) + \frac{1}{2}sgn(F(x_2))$$

On suppose de plus qu'entre $x_1$ et $x_2$ la fonction $F$ s'annule pour une unique valeur $\zeta$, et posons :

$$x_1 < \zeta_1 < \zeta < \zeta_2 < x_2$$

46. In meinen Üniversitätsvorlesungen über algebraische Gleichungen und in meinen beiden Aufsätzen im Monatsbericht vom Februar 1878 habe ich das Vorzeichen einer reellen Grösse $a$ mit $[a]$ bezeichnet ; um aber hier das Zeichen [ ] in derjenigen Bedeutung brauchen zu können, welche *Gauss* demselben beigelegt hat, musste ich in diesem Aufsatze jene Bezeichnungsweise abändern.

47. [Kronecker, 1887b] p. 136 : « Charakter einer ganzen Function ».

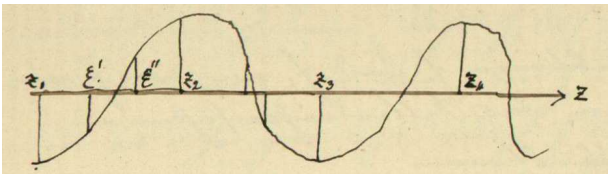


FIGURE 4 – [Kronecker, 1881a] p. 105

On a bien sûr :

$$sgn\left(\frac{F(\zeta_2)-F(\zeta_1)}{\zeta_2-\zeta_1}\right)=-\frac{1}{2}sgn(F(x_1))+\frac{1}{2}sgn(F(x_2))$$

et pour $\zeta_1$ et $\zeta_2$ suffisamment proches de $\zeta$ :

$$sgn\left(F'(\zeta)\right)=-\frac{1}{2}sgn(F(x_1))+\frac{1}{2}sgn(F(x_2))$$

Si maintenant on suppose que pour

$$x_1<z_2<z_3<...<x_2$$

on ait dans chaque intervalle $[x_1,z_2]$, $[z_2,z_3]$,... une unique racine de $F$ que l'on désignera par $\zeta',\zeta'',...$, alors

$$sgn\left(F'(\zeta')\right)=-\frac{1}{2}sgn(F(x_1))+\frac{1}{2}sgn(F(z_2))$$
$$sgn\left(F'(\zeta'')\right)=-\frac{1}{2}sgn(F(z_2))+\frac{1}{2}sgn(F(z_3))$$
$$.....$$

et donc

$$(1)\qquad \sum_{F(\zeta)=0,x_1<\zeta<x_2} sgn\left(F'(\zeta)\right)=-\frac{1}{2}sgn(F(x_1))+\frac{1}{2}sgn(F(x_2))$$

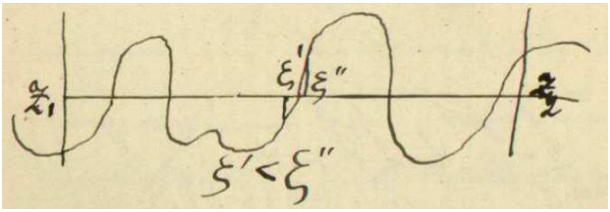


FIGURE 5 – [Kronecker, 1881a] p. 105(2)

Kronecker précise que l'on a supposé ici que $F$ et $F'$ ne s'annulent pas ensemble sur l'intervalle $[x_1,x_2]$, et il en précise l'interprétation graphique (la courbe n'est à aucun moment tangente à l'axe des abscisses).

#### 7.2.2.2 Analogie avec la théorie de Cauchy

Kronecker fait alors remarquer que le fait que l'on ait des informations sur l'ensemble d'un intervalle grâce à ses bornes est à rapprocher de l'intégrale de Cauchy. On a en effet [48] dans son manuscrit de 1880-81 :

> Ce résultat a le caractère de la proposition de Cauchy, le comportement de la fonction aux limites nous donne des éclaircissements sur son comportement entre les deux. [Kronecker, 1881a] p. 106

et [49] dans celui de 1886-87 il affirme :

> On a ici d'une certaine manière un cas analogue à l'intégrale de Cauchy. [Kronecker, 1887b] p. 137

Enfin, dans celui de 1884-85, il ajoute [50] :

> On peut ici connaitre quelque chose sur une somme algébrique de signe à partir de la valeur de la fonction à son point de départ et à son point d'arrivée ; cela n'est en fait rien d'autre que le prolongement au cas de la ligne du théorème de Green sur le potentiel ou de l'intégral de Cauchy. [Kronecker, 1885b] p. 240

Cette analogie est particulièrement importante pour notre propos. Un des grands ressorts de sa théorie des caractéristiques, nous le verrons dans la partie suivante, sera justement de faire le lien entre théorie du potentiel et extension du théorème de Sturm à des systèmes d'équations à plusieurs variables. Dans son dernier cours, Kronecker ajoute d'ailleurs qu'« il est très curieux que le théorème de Sturm soit capable à partir du comportement d'une fonction sur les limites d'un intervalle d'en tirer des conséquences sur le comportement de la fonction à l'intérieur de cet intervalle. Le théorème de Sturm rappelle en cela la théorie du potentiel » [51].

#### 7.2.2.3 Une première application de l'algorithme d'Euclide

On pose maintenant

$$f(x) = x^n - c_1 x^{n-1} + ... \pm c_n$$

et $f_1(x)$ une fonction entière (un polynôme dans le vocabulaire de Kronecker) de degré $n-1$ telle que $f$ et $f_1$ n'aient pas de racine commune. Kronecker applique la division euclidienne, modifiée par Sturm :

48. Dieses Resultat hat den Charakter des *Cauchy*'schen Satzes, das Verhalten der Funktion an den Grenzen gibt uns Aufschluß über ihr Verhalten zwischen denselben.

49. Es liegt hier somit in gewisser Weise ein Analogen zum Cauchy'schen Integral vor.

50. Man kann hier etwas über eine algebraische Summe von Zeichen wissen durch die Werte der Funktion am Anfangspunkt und Endspunkt ; das ist eigentlich nichts anderes als auf den Fall der Linie ausgedehnt, der Greensche Potentialsatz oder das Cauchysche Integral.

51. [Kronecker, 1891a] p. 184 : « Es ist eine große Merkwürdigkeit, daß der Sturmsche Satz aus dem Verhalten einer Funktion an den Grenzen eines Intervalles Folgerungen zu ziehen vermag über das Verhalten im Innern. Die Sturmsche Satz erinnert darin an die Theorie der Potentiale. »

$$\begin{aligned} f - g_1 f_1 + f_2 &= 0 \\ f_1 - g_2 f_2 + f_3 &= 0 \\ ... &= 0 \\ f_{\nu-1} - g_\nu f_\nu &= 0 \end{aligned}$$

où $f_\nu$ est une constante.

On pose

$$F_h(x) = f_{h-1}(x) f_h(x)$$

donc

$$F'_h(x) = f_{h-1}(x) f'_h(x) + f'_{h-1}(x) f_h(x)$$

Si on note $\xi_h$ les racines de $f_h(x) = 0$ et $\xi_{h-1}$ les racines de $f_{h-1}(x) = 0$, alors

$$F'_h(\xi_h) = f_{h-1}(\xi_h) f'_h(\xi_h)$$

et

$$F'_h(\xi_{h-1}) = f'_{h-1}(\xi_{h-1}) f_h(\xi_{h-1})$$

On peut appliquer (1) à $F_h$ et on obtient :

$$\begin{aligned} & \sum_{F_h(\zeta)=0, x_1<\zeta<x_2} sgn\left(F'_h(\zeta)\right) \\ = & \sum_{f_{h-1}(\xi)=0, x_1<\xi<x_2} sgn\left(F'_h(\xi)\right) + \sum_{f_h(\xi)=0, x_1<\xi<x_2} sgn\left(F'_h(\xi)\right) \\ = & \sum_{f_{h-1}(\xi)=0, x_1<\xi<x_2} sgn\left(f'_{h-1}(\xi) f_h(\xi)\right) + \sum_{f_h(\xi)=0, x_1<\xi<x_2} sgn\left(f_{h-1}(\xi) f'_h(\xi)\right) \\ = & -\frac{1}{2} sgn(f_{h-1}(x_1) f_h(x_1)) + \frac{1}{2} sgn(f_{h-1}(x_2) f_h(x_2)) \end{aligned}$$

donc

$$\begin{aligned} & \sum_{h=1}^{\nu} \left( \sum_{f_{h-1}(\xi)=0, x_1<\xi<x_2} sgn\left(f'_{h-1}(\xi) f_h(\xi)\right) + \sum_{f_h(\xi)=0, x_1<\xi<x_2} sgn\left(f_{h-1}(\xi) f'_h(\xi)\right) \right) \\ = & -\frac{1}{2} \sum_{h=1}^{\nu} sgn(f_{h-1}(x_1) f_h(x_1)) + \frac{1}{2} \sum_{h=1}^{\nu} sgn(f_{h-1}(x_2) f_h(x_2)) \end{aligned}$$

or

$$f_{h-1}(x) - g_h(x) f_h(x) + f_{h+1}(x) = 0$$

donc

$$f_{h-1}(\xi_h) + f_{h+1}(\xi_h) = 0$$

c'est-à-dire que

$$sgn(f_{h-1}(\xi_h)) = -sgn(f_{h+1}(\xi_h))$$

et par conséquent :

$$\begin{aligned}
&\sum_{h=1}^{\nu} \left( \sum_{f_{h-1}(\xi)=0, x_1<\xi<x_2} sgn\left(f'_{h-1}(\xi) f_h(\xi)\right) - \sum_{f_h(\xi)=0, x_1<\xi<x_2} sgn\left(f_{h+1}(\xi) f'_h(\xi)\right) \right) \\
&= -\frac{1}{2} \sum_{h=1}^{\nu} sgn(f_{h-1}(x_1) f_h(x_1)) + \frac{1}{2} \sum_{h=1}^{\nu} sgn(f_{h-1}(x_2) f_h(x_2))
\end{aligned}$$

on peut simplifier le membre de gauche et on obtient :

$$\begin{aligned}
&\sum_{f(\xi)=0, x_1<\xi<x_2} sgn\left(f'(\xi) f_1(\xi)\right) - \sum_{f_\nu(\xi_\nu)=0, x_1<\xi_\nu<x_2} sgn\left(f_{\nu+1}(\xi_\nu) f'_\nu(\xi_\nu)\right) \\
&= -\frac{1}{2} \sum_{h=1}^{\nu} sgn(f_{h-1}(x_1) f_h(x_1)) + \frac{1}{2} \sum_{h=1}^{\nu} sgn(f_{h-1}(x_2) f_h(x_2))
\end{aligned}$$

Or $f_{\nu+1} = 0$ et $f'_\nu = 0$ donc on a la formule :

$$\begin{aligned}
(2) \quad &\sum_{x_1<\xi<x_2} sgn\left(f'(\xi) f_1(\xi)\right) \\
&= -\frac{1}{2} \sum_{h=1}^{\nu} sgn(f_{h-1}(x_1) f_h(x_1)) + \frac{1}{2} \sum_{h=1}^{\nu} sgn(f_{h-1}(x_2) f_h(x_2))
\end{aligned}$$

Pour Kronecker, cette dernière formule est la « généralisation du théorème initial de Sturm » [52], et donc certainement la formule la plus importante de la partie sur le théorème de Sturm. On peut déjà noter que Kronecker associe ce théorème à une *formule* : nous reviendrons plus tard sur ce sujet.

On peut donc prendre pour $f_1(x)$ toute fonction de degré $n-1$ qui n'a pas de racine commune avec $f$. Si toutes les racines de cette dernière sont simples, on peut prendre $f_1 = f'$. Dans ce cas, on a

$$sgn\left(f'(\xi) f_1(\xi)\right) = sgn\left(f'(\xi)^2\right) = +1$$

et donc le nombre de racines réelles de l'équation $f(x) = 0$ comprises entre $x_1$ et $x_2$

52. [Kronecker, 1887b] p. 139 : « die Verallgemeinerung des ursprünglichen Sturm'schen Satzes ».

(c'est-à-dire le nombre de $\xi$ compris entre $x_1$ et $x_2$) est

$$-\frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_1)f_h(x_1)) + \frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_2)f_h(x_2))$$

Considérons les suites

$$f(x_1); f_1(x_1); ...; f_\nu(x_1)$$

et

$$f(x_2); f_1(x_2); ...; f_\nu(x_2)$$

On note $m_1$ et $m_2$ le nombre de changements de signes dans ces suites ($m$ comme *mutare*) et $s_1$ et $s_2$ le nombre de permanences de signes ($s$ comme *servare*). Alors on a :

$$\sum_{h=1}^{\nu} sgn(f_{h-1}(x_1)f_h(x_1)) = s_1 - m_1$$
$$\sum_{h=1}^{\nu} sgn(f_{h-1}(x_2)f_h(x_2)) = s_2 - m_2$$

Comme de plus on a

$$s_1 + m_1 = s_2 + m_2$$

et ainsi

$$s_1 - s_2 = -(m_1 - m_2)$$

alors

$$\begin{aligned}
&-\frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_1)f_h(x_1)) + \frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_2)f_h(x_2)) \\
=&-\frac{1}{2}(s_1 - m_1) + \frac{1}{2}(s_2 - m_2) \\
=&\frac{1}{2}(m_1 - m_2) + \frac{1}{2}(m_1 - m_2) \\
=&m_1 - m_2
\end{aligned}$$

Et sous cette forme, il s'agit bien du théorème de Sturm. C'est donc un cas particulier de la formule qu'il vient de démontrer :

$$\begin{aligned}
&\sum_{x_1<\xi<x_2} sgn\left(f'(\xi)f_1(\xi)\right) \\
=&-\frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_1)f_h(x_1)) + \frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_2)f_h(x_2))
\end{aligned}$$

Pour Kronecker, c'est sous cette forme plus générale que se trouve le « vrai » théorème.

Nous venons de voir la première présentation de ce théorème par Kronecker : il s'agit déjà d'une généralisation du théorème, tout au moins au sens où la version de Sturm peut être déduite de celle-ci. Mais Kronecker ne s'attribue pas, à raison, le mérite de cette généralisation : c'est dans le travail de Sylvester que l'on trouve cette extension [53]. Ce qui est marquant dans cette version, c'est la présentation du théorème sous la forme d'une *formule*. On retrouve cette idée dans une lettre [54] que Kronecker a écrit à Cantor en 1884 :

> Je reconnais la vraie valeur scientifique – dans le champ des mathématiques – seulement dans les vérités mathématiques concrètes, ou, pour l'exprimer plus nettement, seulement dans les formules mathématiques. Elles seules, comme l'histoire des mathématiques le montre, sont impérissables. Les diverses théories sur les fondations des mathématiques (comme celle de Lagrange) sont depuis longtemps tombées dans l'oubli, alors que la résolvante de Lagrange reste ! [Cantor *et al.*, 1991] p. 196

#### 7.2.2.4 Un exemple

Dans certains de ses manuscrits [55], Kronecker illustre son travail avec un exemple. Ce ne sera pas toujours le même, mais l'idée restera de prendre un polynôme de degré $n$, et de discuter selon la parité de $n$.

Il considère l'équation trinomiale :

$$f(x) = x^n + nax + b = 0$$

et il pose

$$f_1(x) = \frac{1}{n} f'(x) = x^{n-1} + a$$

Cela permet de calculer tous les termes de la suite de *Sturm* :

$$\begin{aligned} f(x) &= x^n + nax + b \\ f_1(x) &= x^{n-1} + a \\ f_2(x) &= -(n-1)ax - b \\ f_3(x) &= \frac{(-1)^n b^{n-1} - (n-1)^{n-1} a^n}{a^{n-1}} = \frac{\Theta}{a^{n-1}}. \end{aligned}$$

avec donc

$$\Theta = (-1)^n b^{n-1} - (n-1)^{n-1} a^n$$

53. Nous le verrons dans le chapitre suivant.

54. La lettre est reproduite dans [Edwards, 1995] p. 45 et provient de [Cantor *et al.*, 1991] p. 196 et a déjà été commenté dans notre première partie p. 95. « Einen wahren wissenschaftlichen Werth erkenne ich – auf dem Felde der Mathematik – nur in concreten mathematischen Wahrheiten, oder schärfer ausgedrückt, 'nur in mathematischen Formeln.' Diese allein sind, wie die Geschichte der Mathematik zeigt, das Unvergängliche. Die verschiedenen Theorien für die Grundlagen der Mathematik (so die von Lagrange) sind von der Zeit weggeweht, aber die Lagrangesche Resolvente ist geblieben ! »

55. Par exemple dans [Kronecker, 1881a] p. 112.

En fait, si l'on fait le calcul on trouve pour $f_3$ :

$$f_3(x) = \frac{(-1)^n b^{n-1} - (n-1)^{n-1} a^n}{a^{n-1}} = \frac{\Theta}{(n-1)^{n-1} a^{n-1}}$$

Il y a donc un facteur supplémentaire au dénominateur. Mais comme l'écrivait déjà Sturm en 1835 :

> Dans les divisions successives qui servent à former les fonctions $V_2, V_3, etc.$, on peut, avant de prendre un polynome pour dividende ou pour diviseur, le multiplier ou le diviser par tel nombre positif qu'on voudra. [Sturm, 1835] p. 282

Car dans la suite construite, seul le signe des expressions nous importe.

On pose $x_1 = -\infty, x_2 = +\infty$. Deux cas se présentent : $n$ est pair ou impair.

1. Soit $n$ est impair
   pour $x = -\infty$ on a la suite de signes :

   $$-, +, sgn(a), sgn(a^{n-1}\Theta)$$

   et pour $x = +\infty$ on a la suite de signes :

   $$+, +, -sgn(a), sgn(a^{n-1}\Theta)$$

   finalement, en appliquant la formule fondamentale que Kronecker vient de démontrer :

   $$\frac{1}{2}\sum_{h=1}^{3} sgn(f_{h-1}(x_2)f_h(x_2)) - \frac{1}{2}\sum_{h=1}^{3} sgn(f_{h-1}(x_1)f_h(x_1)) = 1 - sgn(a) - sgn(a^n\Theta)$$

   Deux nouveaux cas se présentent :

   (a) Si $a > 0$ alors le nombre de racines réelles est $1 - 1 - sgn(\Theta) = -sgn(\Theta)$, et donc, le nombre de racines ne pouvant être négatif, on doit avoir $\Theta < 0$ et l'équation possède ainsi 1 racine.

   (b) Si $a < 0$ alors le nombre de racines réelles est $1 + 1 + sgn(\Theta) = 2 + sgn(\Theta)$, et donc l'équation possède 1 ou 3 racines.

2. Soit $n$ est pair et on a pour $x_1 = -\infty, x_2 = +\infty$ :

   $$\frac{1}{2}\sum_{h=1}^{3} sgn(f_{h-1}(x_2)f_h(x_2)) - \frac{1}{2}\sum_{h=1}^{3} sgn(f_{h-1}(x_1)f_h(x_1)) = 1 - sgn(a^n\Theta))) = 1 - sgn(\Theta)$$

   et donc le nombre de racines réelles est égal à 0 ou à 2. Si l'on désigne le nombre de racines réelles de l'équation trinomiale par $\mathcal{A}$, alors on a pour $n$ pair et impair

   $$\mathcal{A} = \frac{3 + (-1)^{n-1}}{2} + (-1)^{n-1} sgn(\Theta)$$

#### 7.2.2.5 Une seconde démonstration

La continuité a-t-elle été utilisée dans la démonstration précédente ? Kronecker suppose qu'une sorte de séparation des racines a été effectuée : l'existence de petits intervalles les contenant et dont les bornes rationnelles ont des signes opposés est admis. À partir de là, il développe son calcul sans aucune considération pour la continuité de la fonction, « même le passage de quotients de différences aux quotients différentiels qui s'y produit n'entre pas en ligne de compte car il est question seulement des signes et non des valeurs de la fonction » [56].

Le fait de ne manipuler qu'un ensemble discret qui contient trois éléments (nul, positif et négatif) semble l'exempter de toute considération de continuité. Il en est de même de la notion de *différentiel*. Pour bien saisir où se trouve le point en discussion, nous allons regarder la seconde démonstration qu'il fournit, et qui correspond, comme il le rappelle lui-même [57], à celle de Sturm de 1835.

Dans le manuscrit de 1886-87, Kronecker poursuit son travail en faisant ce qu'il appelle une *vérification du théorème de Sturm au moyen de considérations de continuité* [58]. En fait, sous une forme à peu près stable, cette démonstration se trouve dans la plupart des manuscrits : il s'agit donc de celle que Sturm produit en 1835, et qui utilise – sans que cela soit nécessaire d'après Kronecker – des considérations de continuité.

Voici comment il procède :

La dérivée de $\frac{f}{f_1}$ est

$$d(\frac{f}{f_1}) = \frac{f'f_1 - ff_1'}{f_1^2}$$

donc pour toute racine $\xi$ de $f$ on a

$$d(\frac{f}{f_1})(\xi) = \frac{f'(\xi)f_1(\xi)}{f_1^2(\xi)}$$

et ainsi (2), c'est-à-dire ce que Kronecker appelle le théorème de Sturm, devient

$$\sum_{f(\xi)=0, x_1<\xi<x_2} sgn\left(d(\frac{f}{f_1})(\xi)\right)$$
$$= -\frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_1)f_h(x_1)) + \frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_2)f_h(x_2))$$

car $sgn(f_1^2(\xi)) = 1$. En posant

56. [Kronecker, 1887b] p. 140 : « Selbst der Übergang von Differenzenquotienten zum Differenzialquotienten, der dabei vorkam, ist nicht als solcher zu betrachten, weil es sich nur um Vorzeichen, nicht um Werte der Function selbst handeln. »

57. [Kronecker, 1881a] p. 108 : « Wir lassen jetzt den Beweis folgen, welchen Sturm zu seinem Satze angewendet hat. / Nous allons maintenant suivre la preuve que *Sturm* a utilisée pour son théorème. »

58. [Kronecker, 1887b] p. 140 : « Verification des Sturm'schen Satzes mit Hilfe von Continuitätsberachtungen ».

$$\sigma(x) = \sum_{h=1}^{\nu} sgn(f_{h-1}(x)f_h(x))$$

on a

$$-\sum_{x_1<\xi<x_2} sgn\left(d(\frac{f}{f_1})(\xi)\right) = \frac{1}{2}\sigma(x_1) - \frac{1}{2}\sigma(x_2)$$

Lorsque $x$ varie, la fonction $\sigma$ ne peut se modifier que lorsque $f_{h-1}(x)f_h(x)$ change de signe, et donc s'annule [59]. Cela n'arrive que lorsque $f_{h-1}(x) = 0$ ou $f_h(x) = 0$. Par ailleurs, ces deux fonctions ne peuvent s'annuler ensemble, car cela entrainerait l'existence d'une racine commune entre $f$ et $f_1$.

Soit maintenant $\delta > 0$ et examinons pour $h = 1, 2, ...\nu - 1$ (le cas $h = 0$ sera traité plus loin) :

$$sgn(f_{h+1}(\xi_h - \delta)f_h(\xi_h - \delta)) + sgn(f_h(\xi_h - \delta)f_{h-1}(\xi_h - \delta))$$

et

$$sgn(f_{h+1}(\xi_h + \delta)f_h(\xi_h + \delta)) + sgn(f_h(\xi_h + \delta)f_{h-1}(\xi_h + \delta))$$

Comme

$$f_{h-1} - g_h f_h + f_{h+1} = 0$$

et que $f_h(\xi_h) = 0$ on peut trouver $\delta$ suffisamment petit pour que

$$sgn(f_{h+1}(\xi_h - \delta)) = -sgn(f_{h-1}(\xi_h - \delta))$$

et

$$sgn(f_{h+1}(\xi_h + \delta)) = -sgn(f_{h-1}(\xi_h + \delta))$$

et donc

$$sgn(f_{h-1}(\xi_h - \delta))sgn(f_h(\xi_h - \delta)) = -sgn(f_h(\xi_h - \delta)sgn(f_{h+1}(\xi_h - \delta))$$
$$sgn(f_{h-1}(\xi_h + \delta))sgn(f_h(\xi_h + \delta)) = -sgn(f_h(\xi_h + \delta)sgn(f_{h+1}(\xi_h + \delta))$$

On peut passer, pour la suite $f_{h-1}, f_h, f_{h+1}$

— soit de $-++$ à $--+$

— soit de $--+$ à $-++$

— soit de $++-$ à $+--$

— soit de $+--$ à $++-$

59. Kronecker utilise donc ici, comme Sturm, la méthode des variations.

Il n'y a donc aucune modification du nombre de changements de signes. Si ce dernier peut changer, ce sera ainsi uniquement pour $x = \xi$. On a pris $f$ de sorte à ce qu'elle n'ait pas de racine double, donc

$$sgn(f(\xi - \delta)f_1(\xi - \delta)) - sgn(f(\xi + \delta)f_1(\xi + \delta)) = \pm 2$$

Or

$$f(\xi \pm \delta) = \pm f'(\xi)\delta + ...$$

et pour $\delta$ suffisamment petit, on a

$$sgn(f(\xi \pm \delta)) = sgn(\pm f'(\xi)) = \pm sgn(f'(\xi))$$

Si $f$ passe du négatif au positif, alors $f'$ est positive en $\xi$, et on passe ainsi de $-+$ à $++$. Si en revanche $f$ passe du positif au négatif, alors $f'$ est négative en $\xi$, et on passe ainsi de $+-$ à $--$. Il y a donc lors du passage par chaque racine $\xi$ la perte d'un changement de signes. Entre $x_1$ et $x_2$, il y a donc finalement autant de racines à l'équation $f(x) = 0$, qu'il y a de pertes de changements de signes, c'est-à-dire $m_2 - m_1$.

Kronecker fait remarquer que cette forme généralisée, c'est-à-dire avec $f_1$ qui n'est plus forcément égale à $f'$ est à attribuer à Sylvester [60].

Quelle est la différence avec la démonstration précédente ? Et bien essentiellement l'utilisation de ces $\delta$, caractéristique de la continuité *analytique*, dans, comme la nomme H. Sinaceur, la « méthode des variations et des permanences de signes », c'est-à-dire lorsque l'on examine les changements de signes que produit la variation des valeurs de l'inconnue lors du *passage* par les racines de l'équation donnée. C'est ce *passage* par la racine que Kronecker préfère éviter : nous avons vu [61] que pour Kronecker ce passage n'a aucun sens *arithmétique*, même s'il accepte de lui en donner un *géométrique*. Nous allons d'ailleurs maintenant examiner en détail une interprétation graphique qu'il donne de sa *formule-théorème*.

#### 7.2.2.6 Interprétation géométrique de $\sum_{x_1<\xi<x_2} sgn\,(f'(\xi)f_1(\xi))$

Une fois de plus à propos du théorème de Sturm, Kronecker choisit délibérément un point de vue géométrique, et c'est ainsi qu'il interprète cette somme. Cette partie est particulièrement importante pour notre propos, car elle peut constituer, comme on le verra plus tard, une sorte d'introduction à la théorie des caractéristiques, et montre le lien très fort de cette théorie avec le travail de Kronecker sur le théorème de Sturm.

60. Die obigen allgemeineren Betrachtungen rühren von Sylvester her. [Kronecker, 1887b] p. 142. Cette idée est bien présente dans les écrits de Sylvester dès [Sylvester, 1839], mais, comme le remarque Hourya Sinaceur ([**?**] p.99), Sturm lui même le dit expressément dans son article de 1835.

61. Voir p. 78.

On représente donc dans le plan muni d'un repère les deux courbes associées à l'équation :

$$F(x,y) = (y - f(x))(y - f_1(x)) = 0$$

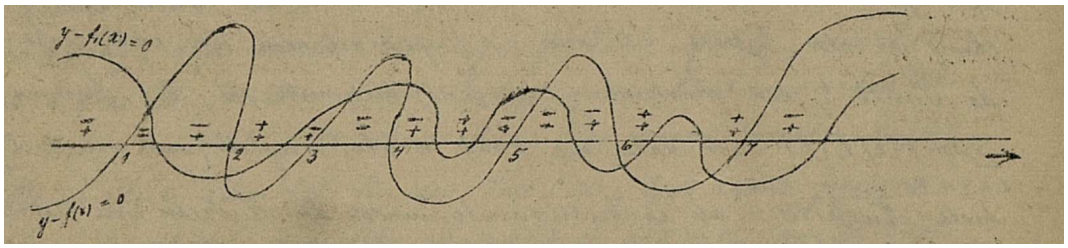


FIGURE 6 – [Kronecker, 1887b] p. 144

Et Kronecker définit, « par analogie avec le cercle », *l'intérieur* (*Innere*) de cette courbe comme étant la partie du plan telle que

$$(y - f(x))(y - f_1(x)) < 0$$

et l'*extérieur* (*Äussere*) l'autre partie[62]. Kronecker poursuit en définissant les :

**Points d'entrée et de sortie** : lorsque l'on se déplace dans le sens croissant sur l'axe des abscisses, on considère les points d'intersection avec la courbe d'équation $y - f(x) = 0$. « Entrer » signifie entrer dans la surface définie par $(y - f(x)).(y - f_1(x)) < 0$ et « Sortir » signifie sortir de cette même surface. Ainsi dans le dessin précédent, les points $1, 2, 3, 5, 6$ sont des points de sortie (*Austrittspunkte*) tandis que $4, 7$ sont des points d'entrée (*Eintrittspunkte*).

Soit $\xi$ une racine de $f(x) = 0$ : si $f'(\xi)f_1(\xi) > 0$, $F$ est croissante lorsque l'on passe par le point $(\xi, 0)$ en parcourant l'axe des abscisses dans le sens croissant, alors $F(x,0) = f(x).f_1(x)$ passe des négatifs aux positifs, et $(\xi, 0)$ est donc un point de *sortie*.

De même, si $f'(\xi)f_1(\xi) < 0$, $F$ est décroissante lorsque l'on passe par le point $(\xi, 0)$ en parcourant l'axe des abscisses dans le sens croissant, alors $F(x,0) = f(x).f_1(x)$ passe des positifs aux négatifs, et $(\xi, 0)$ est donc un point d'*entrée*.

On a

- $sgn(f'(\xi)f_1(\xi)) = +1$ si $(\xi, 0)$ est un point de *sortie*
- $sgn(f'(\xi)f_1(\xi)) = -1$ si $(\xi, 0)$ est un point d'*entrée*

62. Cette comparaison avec le cercle peut faire penser au travail de Gauss dans la première et quatrième démonstration du théorème fondamental de l'algèbre [Gauss, 1799] et [Gauss, 1849].

et finalement, lorsque l'on parcourt l'axe des abscisses de $x_1$ à $x_2$ et que l'on regarde les points de $y - f(x) = 0$ que l'on intersecte, alors $\sum_\xi sgn(f'(\xi)f_1(\xi))$ est la différence du nombre de points de sortie avec celui des points d'entrée. Si l'on note $\mathfrak{a}$ le nombre de points de sortie et $\mathfrak{e}$ le nombre de points d'entrée, on a en reprenant les notations précédentes :

$$\mathfrak{a} - \mathfrak{e} = m_1 - m_2$$

Dans le cas du travail de Sturm, où $f_1 = f'$, on a pour tout $\xi$ :

$$sgn(f'(\xi)f_1(\xi)) = sgn(f'(x)^2) = +1$$

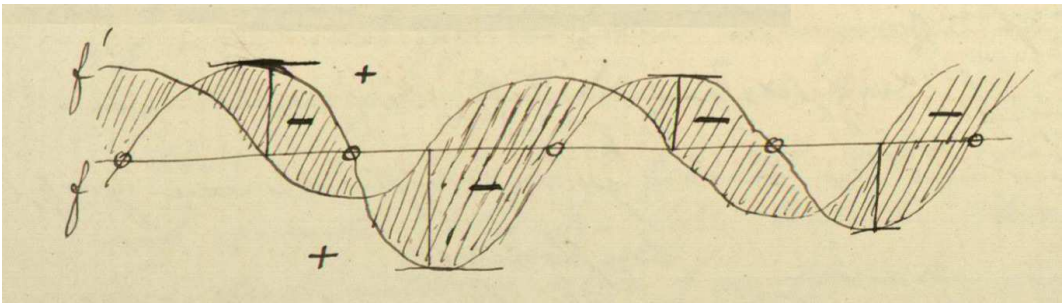



FIGURE 7 – [Kronecker, 1881a] p. 110

Il n'y a donc que des sorties et $\sum_\xi sgn(f'(\xi)f_1(\xi))$ correspond ainsi au nombre de racines réelles de $f(x) = 0$.

**Topologie** Kronecker propose ici une interprétation « géométrique » d'une formule qu'il a donnée dans un cadre algébrique. Les notions d'intérieur et d'extérieur y sont définies, et l'on pense tout de suite au théorème de Jordan qui affirme que

> toute courbe continue C divise le plan en deux régions, l'une extérieure, l'autre intérieure, cette dernière ne pouvant se réduire à zéro, car elle contient un cercle de rayon fini. [Jordan, 1893] p. 99

ou, en des termes plus contemporains, que le complémentaire d'une courbe de Jordan [63] a exactement deux composantes connexes, dont exactement l'une est bornée. Nous connaissons bien, depuis le travail de F. Brechenmacher [64], la controverse qui a eu lieu entre Jordan et Kronecker à propos des formes bilinéaires, et il ne faudrait pas confondre ce théorème de Jordan avec celui dont il est question dans cette controverse, à savoir celui sur la réduction des endomorphismes. Il s'agit ici d'un théorème de topologie [65], et ce qui est remarquable, c'est que ce théorème se déduira finalement assez naturellement de la théorie des caracté-

63. Rappelons ici qu'une courbe de Jordan est l'image continue, dans le plan affine, du cercle trigonométrique.

64. Voir par exemple [Brechenmacher, 2008a].

65. On peut le démontrer dans plusieurs contextes, dont celui de la géométrie différentielle, de l'analyse complexe...

ristiques de Kronecker[66], dont l'idée première, comme nous le verrons, est issue de cette interprétation graphique du théorème de Sturm.

### 7.2.3 Recherche du nombre total de solutions d'une fonction entière

Rappelons que

$$f(x) = x^n - c_1 x^{n-1} + ... \pm c_n$$

Kronecker se pose ensuite la question suivante : que se passe-t-il si $x_1 = -\infty$ et $x_2 = +\infty$ ? Pour y répondre, il commence par poser $f_1 = f'$ et

$$f_h(x) = c_h x^{n_h} + ...$$

et donc

$$sgn(f_h(x_1)) = (-1)^{n_h} sgn(c_h) \qquad sgn(f_h(x_2)) = sgn(c_h)$$

Le théorème de Sturm devient donc dans ce cas :

$$\begin{aligned} &\sum_{-\infty<\xi<+\infty} sgn\left(f'(\xi)f_1(\xi)\right) \\ &= -\frac{1}{2}\sum_{h=1}^{\nu}(-1)^{n_{h-1}+n_h} sgn(c_{h-1}c_h) + \frac{1}{2}\sum_{h=1}^{\nu} sgn(c_{h-1}c_h) \end{aligned}$$

Maintenant, si $n_h + n_{h-1} \equiv 0[2]$ (ce qui revient à dire que $n_{h-1} - n_h \equiv 0[2]$) alors

$$-(-1)^{n_{h-1}+n_h} sgn(c_{h-1}c_h) + sgn(c_{h-1}c_h) = 0$$

Si au contraire $n_{h-1} - n_h \equiv 1[2]$ alors

$$-(-1)^{n_{h-1}+n_h} sgn(c_{h-1}c_h) + sgn(c_{h-1}c_h) = 2sgn(c_{h-1}c_h)$$

donc finalement

$$\sum_{\substack{f(\xi)=0 \\ -\infty<\xi<+\infty}} sgn\left(f'(\xi)f_1(\xi)\right) = \sum_{n_{h-1}-n_h\equiv 1[2]} sgn(c_{h-1}c_h)$$

Dans le cas régulier (*regulären*) les degrés des fonctions $f_h$ décroissent d'une unité, et donc on a $n_{h-1} - n_h = 1$. Ce cas est très important pour la suite : nous verrons que ce que Kronecker met en avant lorsqu'il commente les mérites de sa propre exposition du théorème, c'est justement de ne pas s'être contenté de ce cas régulier.

66. La démonstration est basée sur l'indice de Kronecker, comme le fait remarquer Hadamard en 1910 : « La démonstration, d'après M. Ames, du théorème de M. Jordan sur les courbes fermées sans point double (n 306, 307) repose sur la considération de l'ordre d'un point ou, si l'on veut, sur la considération d'une variation d'argument. La généralisation, dans le cas où le nombre des dimensions dépasse deux, est fournie par l'indice de Kronecker ». [Tannery et Hadamard, 1910] p. 437.

On a alors :

$$\sum_{\substack{f(\xi)=0 \\ -\infty<\xi<+\infty}} sgn\left(f'(\xi)f_1(\xi)\right) = \sum_{h=1}^{\nu} sgn(c_{h-1}c_h)$$

et si on note $s$ le nombre de permanences de signes (*Zeichenfolgen*) et $m$ le nombre de changements de signes (*Zeichenwechsel*) dans la suite des plus hautes puissances de $x$ des fonctions $f_h$, alors on a :

$$\mathfrak{a} - \mathfrak{e} = s - m$$

Comme nous sommes dans le cas régulier, donc dans le cas où on doit avoir $n+1$ fonctions $f_h$, on a $s + m = n$ et on obtient facilement

$$n - \mathfrak{a} + \mathfrak{e} = 2m$$
$$n + \mathfrak{a} - \mathfrak{e} = 2s$$

Soit maintenant $v$ le nombre de racines réelles de l'équation $f(x) = 0$, et il est sûr que

$$-v \leq \mathfrak{a} - \mathfrak{e} \leq +v$$

et donc que

$$\frac{n-v}{2} \leq s, m \leq \frac{n+v}{2}$$

Si en particulier $f_1 = f'$, on a vu qu'il n'y avait que des sorties et que donc $\mathfrak{a} - \mathfrak{e} = \mathfrak{a} = s - m = v$. Donc comme bien sûr $s + m \leq n$, on a

$$m \leq \frac{n-v}{2}$$

Ce nombre, $\dfrac{n-v}{2}$ est par définition le nombre de paires de racines imaginaires conjuguées de l'équation $f(x) = 0$.

Kronecker termine ce chapitre en montrant que l'on peut considérer le cas où $x_1 = -\infty$ et $x_2 = +\infty$ comme le cas général : on revient au cas de l'intervalle fini[67] par des substitutions bien choisies.

On a donc finalement ici obtenu la formule :

67. [Kronecker, 1887b] p. 148 : « Wir sehen also, dass auf mehrere Weisen der algemeine Fall des Sturm'schen Satzes auf den speciellen $x_1 = -\infty$, $x_2 = +\infty$ zurückgeführt werden kann. »

$$\sum_{\substack{f(\xi)=0 \\ -\infty<\xi<+\infty}} sgn\left(f'(\xi)f_1(\xi)\right) = \sum_{h=1}^{\nu} sgn(c_{h-1}c_h)$$

qui nous permet, dans le cas régulier, de donner le nombre de solutions réelles en fonction des coefficients des termes de plus haut degré des fonctions d'une suite de Sturm. Cela a une importance effective, car il est en effet possible de calculer ces termes sans avoir à connaître complètement les fonctions de cette suite [68].

### 7.2.4 Un exemple : l'équation biquadratique

Kronecker développe un long travail sur l'équation biquadratique. Cette étude prend place, selon les années, soit juste après la démonstration qu'il attribue à Sturm, comme nous le ferons ici, soit après que les idées de Sylvester ont été développées, voire dans le cadre de la théorie des caractéristiques [69]. En fait, l'intérêt de ce dernier enchaînement est essentiellement de lui permettre d'utiliser plus facilement le théorème de Sturm grâce aux matrices $|s_{p+q}|$ dont nous verrons l'importance. Mais ce faisant, la place du discriminant est partiellement cachée, et c'est pourquoi nous avons choisi de présenter cette application maintenant [70].

#### 7.2.4.1 Application du théorème de Sturm

Comme le fait remarquer Kronecker, l'application du théorème de Sturm aux équations des deux premiers degrés n'a pas vraiment d'intérêt [71] :

> L'étude des équations linéaires et quadratiques est tellement triviale que l'on considère donc en premier les équations cubiques. Celles-ci ont déjà été étudiées en détail et nous avons vu que les deux seuls cas possibles (3 racines réelles ou seulement une seule) se distinguent par le signe de leur discriminant. [Kronecker, 1881a] p. 113

Et les équations cubiques ont en effet été étudiées dans la première partie : c'est donc un travail détaillé sur l'équation biquadratique que propose Kronecker. On retrouve cette étude dans tous les manuscrits entre 1878 et 1888, c'est-à-dire dans ce que l'on a défini comme étant la seconde période. Durant celle-ci, le théorème de Sturm occupe d'ailleurs une place prépondérante dans les *Vorlesungen*.

On pose

$$f(u) = u^4 - 6zu^2 + 4yu + 3x = 0$$

68. On peut voir par exemple [Cayley, 1846a] ou [Koralek, 1852].

69. Ce sera le cas dans le manuscrit de 1890-91.

70. On retrouve cet exemple traité dans son article [Kronecker, 1878c] p. 68. Dans ses cours, il apparaît par exemple dans [Kronecker, 1881a] p.114, [Kronecker, 1887a] p. 152 ou encore [Kronecker, 1887b] p.166.

71. Die Untersuchung der linearen und quadratischen Gleichungen ist zu trivial, es wären also zunächst die kubischen Gleichungen zu betrachten. Diese sind bereits eingehend studirt und wir haben gesehen, daß die beiden einzigen möglichen Fälle (3 reelle Wurzeln oder nur eine) durch das Vorzeichen der Diskriminante von einander unterschieden werden.

où $x, y, z$ sont des variables indépendantes (et les coefficients sont là pour simplifier les formules).

Kronecker introduit, une fois encore, dans son travail un point de vue géométrique. Il souhaite en effet considérer le triplet $(x, y, z)$ comme les coordonnées d'un point de l'espace dans un système de coordonnées rectangulaires. Ainsi chaque point de l'espace est associé à une équation, et inversement. On peut ainsi diviser l'espace en trois parties ou domaines (*Gebiete*) $\mathcal{G}_0, \mathcal{G}_1, \mathcal{G}_2$, selon que l'équation a aucune, une, ou deux paires de racines imaginaires (*imaginärer Wurzeln*). Le but de Kronecker est de caractériser chacun de ces domaines à l'aide du théorème de Sturm.

Posons

$$f_1(u) = \frac{1}{4}f'(u) = u^3 - 3zu + y$$

et appliquons l'algorithme modifié d'Euclide, en simplifiant les coefficients des fonctions obtenues :

$$\begin{aligned}
g_1(u) &= u \\
f_2(u) &= zu^2 - yu - x \\
g_2(u) &= \frac{u}{z} + \frac{y}{z^2} \\
f_3(u) &= \mathcal{D}_1 u - y(x + z^2) \\
g_3(u) &= 3z^2\left(z\mathcal{D}_1 u + y(2xz + y^2 - 2z^3)\right) \\
f_4(u) &= \frac{3z^2\mathcal{D}}{\mathcal{D}_1^2}
\end{aligned}$$

où

$$\mathcal{D}_1 = 3z^3 - xz - y^2$$

et

$$\mathcal{D} = (x + z^2)^3 - (3xz + y^2 - z^3)^2 = x(x - 3z^2)^2 - 2y^2z(3x - z^2) - y^4$$

Pour obtenir ces fonctions, on a comme précédemment éliminé les facteurs positifs [72].

Pour chercher le nombre de racines réelles de l'équation, on considère l'intervalle $]-\infty; +\infty[$. Nous sommes dans le cas que Kronecker qualifie de régulier, c'est-à-dire dans le cas où les fonctions $f_h$ décroissent d'un degré à chaque étape. On a donc, comme on l'a vu

$$\sum_{\substack{f(\xi)=0 \\ -\infty<\xi<+\infty}} sgn\left(f'(\xi)f_1(\xi)\right) = \sum_{h=1}^{4} sgn(c_{h-1}c_h)$$

où $c_0, c_1, .., c_4$ sont les coefficients des termes de plus haut degré des fonctions $f, f_1, ..., f_4$.

72. $z^2$ pour $f_3$ et $D_1^2$ pour $g_3$.

Or, comme de plus ici $f_1 = f'$, on a :

$$\sum_{\substack{f(\xi)=0 \\ -\infty<\xi<+\infty}} sgn\left(f'(\xi)^2\right) = \sum_{h=1}^{4} sgn(c_{h-1}c_h)$$

c'est-à-dire que le nombre de racines réelles de $f$ est $\sum_{h=1}^{4} sgn(c_{h-1}c_h)$, et donc :

$$sgn(1.1) + sgn(1.z) + sgn(z.\mathcal{D}_1) + sgn(\mathcal{D}_1.\mathcal{D})$$

C'est la suite

$$1 \ , \ z \ , \ z.\mathcal{D}_1 \ , \ \mathcal{D}_1.\mathcal{D}$$

que Kronecker désigne comme étant la suite de Sturm [73] et il fait remarquer que $\mathcal{D}$ est, à un facteur près, le discriminant de notre équation biquadratique. On a donc :

| $\mathcal{D} > 0$ | | | | |
|---|---|---|---|---|
| $sgn(\mathcal{D}_1)$ | +1 | +1 | -1 | -1 |
| $sgn(z)$ | +1 | -1 | +1 | -1 |
| Racines réelles | 4 | 0 | 0 | 0 |

| $\mathcal{D} < 0$ | | | | |
|---|---|---|---|---|
| $sgn(\mathcal{D}_1)$ | +1 | +1 | -1 | -1 |
| $sgn(z)$ | +1 | -1 | +1 | -1 |
| Racines réelles | 2 | -2 | 2 | 2 |

On voit que le cas $(\mathcal{D}) < 0$, $(\mathcal{D}_1) > 0$ et $z < 0$ pose problème, étant donné que l'on obtient un nombre négatif de racines. Cependant, si l'on suppose

$$\mathcal{D}_1 = 3z^3 - xz - y^2 > 0$$

alors

$$x > \frac{y^2 - 3z^3}{-z}$$

et donc comme $z < 0$, on a $x > 0$. On a par ailleurs

$$3z^3 - xz - y^2 > 0$$
$$3z^2 - x - \frac{y^2}{z} < 0$$
$$x + z^2 > 4z^2 - \frac{y^2}{z}$$

Considérons maintenant l'expression

$$\mathcal{D} = x(x - 3z^2)^2 - 2y^2z(3x - z^2) - y^4$$

Déjà,

$$x(x - 3z^2)^2 > 0$$

73. [Kronecker, 1887b] p. 167 : « Sturm'sche Reihe ».

On a d'après ce qui précède

$$\begin{aligned} x - 3z^2 = x + z^2 - 4z^2 &> -\frac{y^2}{z} \\ 3x - z^2 &> 8z^2 - \frac{3y^2}{z} \\ -2y^2z(3x - z^2) &> -16y^2z^3 + 6y^4 \\ -2y^2z(3x - z^2) - y^4 &> -16y^2z^3 + 5y^4 > 0 \end{aligned}$$

Donc $\mathcal{D} > 0$, donc le cas $(\mathcal{D}) < 0$, $(\mathcal{D}_1) > 0$ et $z < 0$ n'arrive jamais. Nous avons donc déterminé grâce au théorème de Sturm le fait que :

1. L'équation biquadratique a 4 racines réelles lorsque $(\mathcal{D}) > 0$, $(\mathcal{D}_1) > 0$ et $z > 0$.
2. L'équation biquadratique a 2 racines réelles lorsque $(\mathcal{D}) < 0$.
3. L'équation biquadratique a 0 racine réelle lorsque $(\mathcal{D}) > 0$ et soit $z < 0$, soit $(\mathcal{D}_1) < 0$ et $z > 0$.

Ces trois cas correspondent aux trois domaines $\mathcal{G}_0, \mathcal{G}_1, \mathcal{G}_2$ que Kronecker étudie plus en détail dans le chapitre suivant.

#### 7.2.4.2 Étude de la surface-discriminante (*Discriminantenfläche*)

À propos des trois domaines $\mathcal{G}_0, \mathcal{G}_1, \mathcal{G}_2$, Kronecker se demande par quel « genre de variété est représenté leurs frontières » [74].

Commençons par exemple par examiner l'intersection de $\mathcal{G}_1$ et $\mathcal{G}_0$. Un point est sur cette frontière si l'équation correspondante a 4 racines réelles et une paire de racines complexes conjuguées. Cela n'est possible que lorsque l'équation possède une racine réelle double [75]. Ainsi, le discriminant est nul pour tout point situé sur la frontière. On a donc :

$$(x + z^2)^3 - (3xz + y^2 - z^3)^2 = 0$$

et la frontière entre $\mathcal{G}_1$ et $\mathcal{G}_0$ est donc une partie de la surface définie par l'équation précédente. Il en est de même pour les frontières entre $\mathcal{G}_1$ et $\mathcal{G}_2$ ou entre $\mathcal{G}_0$ et $\mathcal{G}_2$ : l'existence d'une racine réelle double est nécessaire, et donc ces frontières sont sur la même surface. On désigne ainsi cette surface par le terme de surface-discriminante (*Discriminantenfläche*).

Ici, Kronecker affirme que cette surface est *développable* : en effet, $\mathcal{D}$ est le résultat de l'élimination de $u$ dans $f(u) = u^4 - 6zu^2 + 4yu + 3x = 0$ et $f'(u) = u^3 - 3zu + 4y = 0$,

74. [Kronecker, 1887b] p. 169 : « durch eine Mannichfaltigkeit von welcher Art ihre Grenze dargestellt wird. »

75. On aura alors deux racines conjuguées de la forme $a \pm 0 \cdot i$.

c'est donc l'intersection de deux plans (en considérant $u$ comme un paramètre), ce qui correspond bien à la définition de surface développable.

En fait, la surface-discriminante est le lieu de tous les points qui correspondent à une équation ayant des racines doubles ou d'ordre supérieur. Dans notre exemple, trois cas se présentent :

1. L'équation a une racine réelle double.
2. L'équation a une racine réelle triple.
3. L'équation a une racine imaginaire double $a + ib$. Dans ce cas la racine conjuguée $a - ib$ doit aussi être double. Or le coefficient en $x^3$ de l'équation est nul, donc la somme des racines est nulle, c'est à dire que $4a = 0$ et donc les racines sont des imaginaires purs.
4. L'équation a deux racines réelles doubles $\xi_1$ et $-\xi_1$.
5. L'équation a une racine réelle quadruple, qui doit donc être nulle (la somme des racines étant nulle).

**Cas (1)** Il est représenté par la surface-discriminante entière, à l'exception d'une courbe et de l'origine, qui correspond au cas (5).

**Cas (2)** Il est représenté par une courbe contenue dans la surface-discriminante dont on trouve facilement les équations. Soit $\xi$ la racine triple, et donc $3\xi$ la quatrième, alors :

$$f(u) = (u - \xi)^3(u + 3\xi) = u^4 - 6\xi^2u^2 + 8\xi^3u - 3\xi^4$$

et donc

$$z = \xi^2 \ , \ y = 2\xi^3 \ , \ x = -\xi^4$$

ce qui nous donne en éliminant $\xi$, l'équation de la surface qui correspond au cas (2) :

$$\begin{cases} x + z^2 = 0 \\ y^2 - 4z^3 = 0 \end{cases}$$

**Cas (3)** Soit $\eta i$ et $-\eta i$ les deux racines imaginaires doubles. On a alors :

$$f(u) = (u^2 + \eta^2)^2 = u^4 + 2u^2\eta^2 + \eta^4$$

Donc

$$z = -\frac{1}{3}\eta^3 \ , \ y = 0 \ , \ x = -\frac{1}{3}\eta^4$$

et ainsi l'équation de la courbe est :

$$\begin{cases} y = 0 \\ x - 3z^2 = 0 \end{cases}$$

**Cas (4)** Soit $\xi$ et $-\xi$ les deux racines

$$f(u) = (u^2 - \xi^2)^2 = u^4 - 2u^2\xi^2 + \xi^4$$

Donc

$$z = \frac{1}{3}\xi^3 \ , \ y = 0 \ , \ x = -\frac{1}{3}\xi^4$$

et ainsi l'équation de la courbe est :

$$\begin{cases} y & =0 \\ x - 3z^2 & =0 \end{cases}$$

**Recherche des singularités** Nous allons maintenant chercher les singularités de la surface d'équation :

$$F(x, y, z) = (x + z^2)^3 - (3xz + y^2 - z^3)^2 = x(x - 3z^2)^2 - 2y^2z(3x - z^2) - y^4$$

c'est-à-dire de la surface-discriminante.

Pour cela, nous avons besoin des expressions suivantes :

$$\begin{aligned} \frac{\partial F}{\partial x} &= 3(x + z^2)^2 - 6z(3xz + y^2 - z^3) \\ \frac{\partial F}{\partial y} &= -4y(3xz + y^2 - z^3) \\ \frac{\partial F}{\partial z} &= 6x(x + z^2)^2 - 6(3xz + y^2 - z^3)(x - z^2) \end{aligned}$$

Et nous cherchons $x, y, z$ tels que

$$\begin{cases} F(x, y, z) & = 0 \\ \dfrac{\partial F}{\partial x} & = 0 \\ \dfrac{\partial F}{\partial y} & = 0 \\ \dfrac{\partial F}{\partial z} & = 0 \end{cases}$$

On a $\dfrac{\partial F}{\partial y} = 0$ si $y = 0$ ou $3xz + y^2 - z^3 = 0$

1. Si $y = 0$, alors
$$\begin{cases} (x+z^2)^3 - (3xz - z^3)^2 & = 0 \\ 3(x+z^2)^2 - 6z(3xz - z^3) & = 0 \\ y & = 0 \\ 6x(x+z^2)^2 - 6(3xz - z^3)(x - z^2) & = 0 \end{cases}$$
et on obtient
$$\begin{cases} 6z(3x - z^2)(3z^2 - x) & = 0 \\ y & = 0 \\ (x+z^2)32 - (3xz - z^3)^2 & = 0 \end{cases}$$
   (a) Si $z = 0$, alors $x = 0$ et on obtient le point $(0, 0, 0)$, c'est-à-dire l'origine du repère.
   (b) Si $3x - z^2 = 0$ on obtient de même le point $(0, 0, 0)$.
   (c) Si $(3z^2 - x) = 0$ on obtient la courbe d'équation
$$\begin{cases} y & = 0 \\ x - 3z^2 & = 0 \end{cases}$$
   qui est une parabole dans le plan $(xz)$, dont l'axe est l'axe des $x$ et le sommet est l'origine du repère.
2. Si $3xz + y^2 - z^3 = 0$, alors on obtient la courbe d'équation
$$\begin{cases} 3xz + y^2 - z^3 & = 0 \\ x + z^2 & = 0 \end{cases}$$

C'est-à-dire

$$\begin{cases} x + z^2 & = 0 \\ y^2 - 4z^3 & = 0 \end{cases}$$

Ainsi Kronecker affirme que la courbe est l'intersection des deux surfaces cylindriques d'équation $x+z^2 = 0$ et $y^2-4z^3 = 0$. Si la première est bien la surface de ce que l'on appelle maintenant un cylindre parabolique, c'est-à-dire un cylindre dégénéré dont la directrice est une parabole, la seconde n'est pas une surface quadratique, mais cubique. Kronecker utilise en fait une analogie en désignant cette cubique comme un cylindre ayant comme directrice une parabole de Niel [76].

— le premier a pour équation $x + z^2 = 0$, sa directrice est la parabole d'équation $x + z^2 = 0$ dans le plan $xy$ et sa génératrice est parallèle à l'axe des $y$.
— le second a pour équation $y^2-4z^3 = 0$, sa directrice est la parabole de Niel d'équation $y^2 - 4z^3 = 0$ dans le plan $yz$ et sa génératrice est parallèle à l'axe des $x$.

Selon les manuscrits, Kronecker poursuit une étude plus précise de certains points d'intersections (par exemple dans [Kronecker, 1887b]) ou au contraire donne une description

76. Une parabole de Niel est une parabole semi-cubique d'équation $y = \pm a x^{\frac{3}{2}}$.

générale de la surface-discriminante (par exemple dans [Kronecker, 1881a]). Dans tous les cas il aboutit à sa représentation graphique « vue de dessus » (Figure 8).

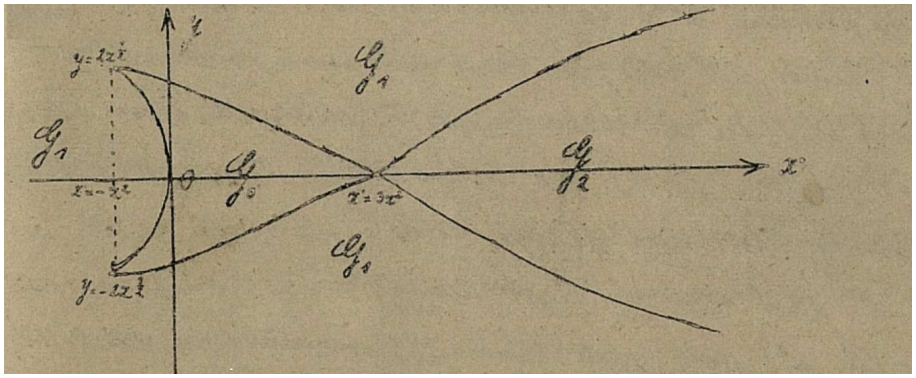


FIGURE 8 – [Kronecker, 1887b] p. 175

et dans le manuscrit [Kronecker, 1881a] il donne aussi l'allure de la surface dans l'espace (Figure 9).

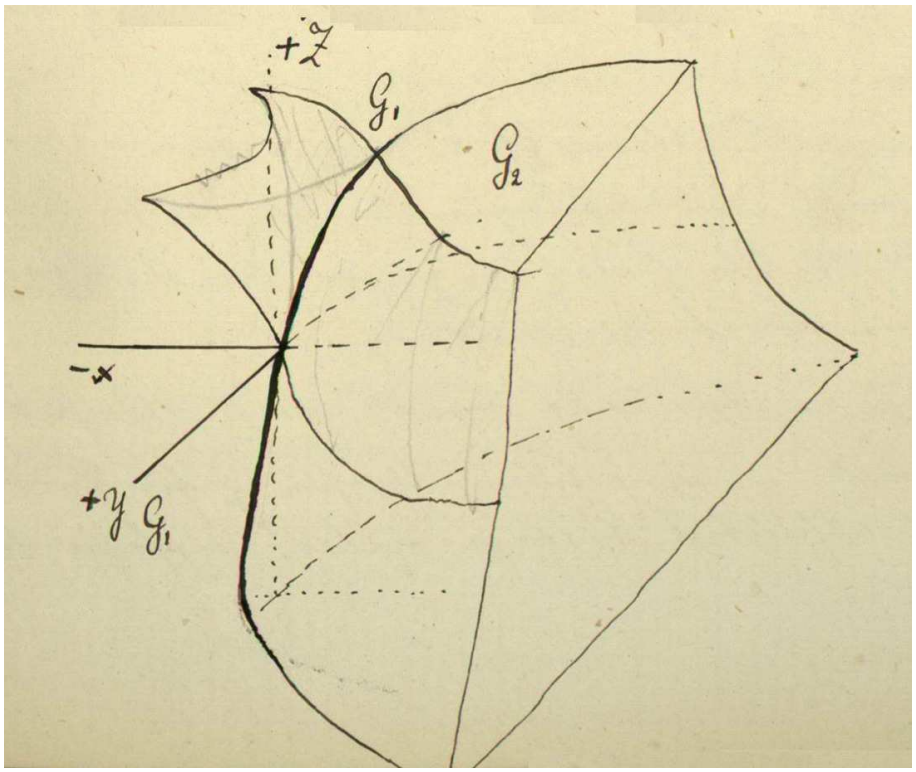



FIGURE 9 – [Kronecker, 1881a] p. 115

Encore une fois, alors que Kronecker traite le plus long exemple du chapitre sur le théorème de Sturm, voire le plus long exemple de l'ensemble des manuscrits, nous rencontrons une description *géométrique* de son travail. Plus encore, il ne s'agit pas seulement d'une représentation graphique d'une fonction de trois variables, mais bien de la recherche des lieux de l'espace qui correspondent à certains types de solutions d'une famille d'équations.

Des considérations du même ordre seront reprises par Kronecker lorsqu'il regarde comment évolue la *caractéristique* de plusieurs fonctions [77] et comme nous le verrons, ces développements inspireront Poincaré dans son *Analysis Situs*.

Pour terminer, nous présentons deux représentations graphiques de la *surface-discriminante* réalisées à l'aide du logiciel *Grapher* et correspondant aux schémas du cours de Kronecker (Figure 10 et 11).

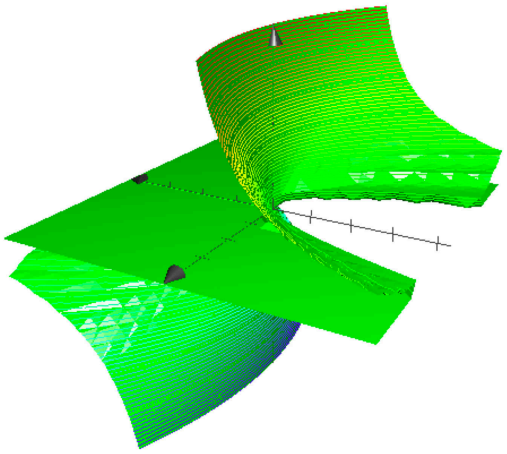


FIGURE 10 –

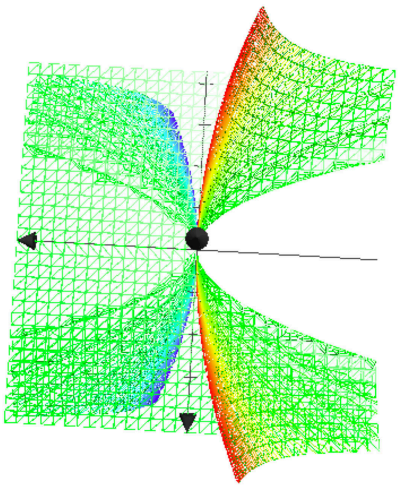


FIGURE 11 –

77. Nous verrons cela en détail dans la partie suivante.

## 7.3 Conclusion

Nous avons affaire à des leçons qui s'adressent à des étudiants, certes brillants, mais en cours d'apprentissage. Kronecker, nous l'avons vu, a des objectifs pédagogiques qu'il souhaite atteindre, et s'il faut pour cela un vocabulaire classique et connu des étudiants, il n'hésite pas à l'utiliser, même si celui-ci est problématique dans sa propre conception des mathématiques.

Cependant, il ne faudrait pas non plus voir dans *Kontinuitätbetrachtungen*, *Stetigkeit* ou *geometrische Anschauung* des termes que Kronecker emploierait indistinctement. Ainsi, pour sa première approche du théorème de Sturm, il utilise l'« intuition géométrique » (*geometrische Anschauung*), et considére les racines réelles comme étant les points d'intersection de la courbe représentant la fonction avec l'axe des abscisses. le terme *réel*, que Kronecker utilise, est plus à comprendre comme *qui n'est pas complexe*. Il n'y a en fait pas de concept sous-jacent de nombre réel, et dans un second temps Kronecker montre que l'on peut toujours éviter dans le théorème de Sturm de considérer « l'existence » des racines, au sens où il n'est pas nécessaire d'assimiler racine et nombre : « le théorème de *Sturm* ne détermine par là rien d'autre que le nombre des intervalles (...) dans lesquels a lieu un changement de signe » [78]. Voilà ce qu'il dit dans son dernier cours sur la théorie des équations algébriques [79] :

> Il est aussi surprenant que la détermination des racines réelles ne soit pas du tout altérée par les imaginaires, et la fiction des racines imaginaires n'est donc pas du tout nécessaire. Mais les prétendues racines réelles sont aussi à éviter. Le théorème de Sturm se laisse complètement déduire à partir de faits dans le comportement de la fonction que pour résumer, peut-être avec une abstraction trop sommaire, on appelle communément l'« existence des racines ». [Kronecker, 1891a] p. 184

Kronecker remarque donc, dans son dernier cours, que la notion d'imaginaire n'est pas nécessaire pour compter le nombre de « racines réelles ». Mais que le théorème de Sturm peut être modifié de sorte que celle d'« racine réelle » n'intervienne plus. Car ces « faits dans le comportement de la fonction », c'est-à-dire ici l'existence d'un nombre fini d'intervalles de longueur déterminable dans lesquels a lieu un unique changement de signe, suffit à la construction du théorème de Sturm. C'est donc la racine, en tant que nombre, que Kronecker exclut ici de ses considérations.

Il ne rejette pas l'« intuition géométrique », à travers laquelle il conçoit dans un premier temps la continuité et les racines d'une équation, et la continuité géométrique (*geometrische*

78. [Kronecker, 1881a] p. 120 : « Der *Sturm*'sche Satz bestimmt daher eigentlich nichts anderes als die Anzahl der oben charakterisirten Intervalle $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, zwischen welchen ein Zeichenwechsel stattfindet. »

79. Sodann ist erstaunlich, daß die Ermittlung der reellen Wurzeln durch die imaginären gar nicht beeinträchtigt wird, die Fiktion imaginärer Wurzeln also gar nicht nötig ist. Aber auch die sogenannten reellen Wurzeln sind zu umgehen. Der Sturmsche Satz läßt sich vollständig herleiten aus den Thatsachen im Verhalten der Funktion, welche in kurzer, vielleicht zu kurzer Abstraktion gemeinhin unter der „Existenz der Wurzeln" zusammengefaßt zu werden pflegen.

*Stetigkeit*)[80] d'une courbe, qui correspond à un tracé sans interruption, en est le produit. C'est sur elle que Kronecker construit sa première démonstration du théorème de Sturm.

Par contre, Kronecker exclut la *continuité arithmétique* de ses démonstrations, ou tout au moins les conclusions que Bolzano en a tirées pour démontrer son théorème des valeurs intermédiaires. Il n'y a pas vraiment d'*arithmétisation* de la continuité, comme c'est le cas chez Bolzano ou Cauchy. Il souhaite plutôt se passer de toute « considération de continuité » (*Continuitätsbetrachtungen*), et il affirme d'ailleurs avoir donné une démonstration qui pouvait s'en affranchir[81] :

> Lors de la démonstration que nous venons de donner du théorème de Sturm, il doit être considéré comme essentiel le fait que nous nous soyons servi somme toute d'aucune considération de continuité. [Kronecker, 1887b] p. 140

En fait, ce concept n'a tout simplement pas sa place dans son « arithmétique générale ». S'il y a bien *arithmétisation* de sa démonstration, il ne faut pas la chercher dans la notion de continuité, mais plutôt dans celle de racine, qu'il expose à l'aide d'une partition en un nombre fini d'intervalles. il extrait progressivement le noyau algébrique de ces « intuitions géométriques » jusqu'à obtenir une démonstration qui pourra s'exempter de toute notion de continuité[82]. En ce sens, la suppression progressive de considérations de continuité, Kronecker participe à l'*algébrisation* du théorème de Sturm.

Terminons en faisant remarquer que, sans doute, s'il souhaite commencer par cette démonstration basée sur la conception géométrique des racines, c'est pour des raisons pédagogiques. Il s'agit de celle de Sturm, et donc historiquement de l'une des premières démonstrations du théorème[83]. La place de l'histoire des mathématiques dans les cours de Kronecker est loin d'être négligeable, et la connaissance de l'histoire des concepts semble lui tenir à cœur. Le théorème de Sturm est, comme l'a montré H. Sinaceur, un « classique » pour tout étudiant en mathématiques de la seconde moitié du XIX$^{e}$ siècle. Ainsi, plutôt que de présenter uniquement son point de vue, il préfère commencer par donner la preuve « classique » habituelle, directement abordable par ses étudiants. Après leur avoir permis d'en maitriser le contenu, Kronecker construit un autre exposé sur la base de ces connaissance préalables. Ce nouvel exposé, qui repose sur un travail approfondi sur les fractions continues, est l'objet de notre prochain chapitre.

80. Voir [Kronecker, 1881a] p. 104.

81. Bei der Herleitung, welche wir von dem Sturm'schen Satz gegeben haben, ist als wesentlich der Umstand anzusehen, dass wir uns dabei überhaupt keiner Continuitätsbetrachtungen bedienten.

82. C'est tout au moins ce qu'il affirmera de la dernière version du théorème de Sturm que nous exposerons ici. Voir plus loin p. 312.

83. On se souvient que Sturm n'est pas le premier à en publier une démonstration.

# Chapitre 8

# Sylvester et les fractions continues


8.1 Les premiers travaux sur les suites de Sturm : 1839-1841 . . . . . **234**
8.1.1 De nouvelles fonctions auxiliaires . . . . . 235
8.2 Démonstration par Sturm de la proposition de Sylvester . . . . . **237**
8.2.1 Expression de $V_2, V_3, ...$ en fonction de $x$ et des racines de $V = 0$ 239
8.2.2 Cauchy . . . . . 240
8.2.3 Expression des $\lambda$ . . . . . 241
8.3 Retour à la note de 1841. . . . . **243**
8.4 *Digression* de Kronecker sur les développements en fractions continues **244**
8.4.1 Expression de $N(g_1...g_\nu)$ sous la forme d'un déterminant . . . 249
8.4.2 Conclusion . . . . . 250


Kronecker pose la question suivante [1] : la méthode consistant à appliquer l'algorithme d'Euclide modifié est-elle la seule qui permette de trouver une suite de Sturm ? À cette question, indique Kronecker, Sylvester est le premier à avoir répondu, grâce à des fonctions symétriques, et ce d'« une façon élégante » (*eleganter Weise* [Kronecker, 1887b] p. 181). Il ajoute que ce travail de Sylvester [2] englobe aussi la suite de fonctions donnée par Sturm. La *généralisation* du théorème est une fois encore mise en avant. Mais auparavant Kronecker souhaite présenter un travail préalable sur les fractions rationnelles qui est issu, comme nous allons le voir, d'une étude du théorème de Sturm initiée par Sylvester.

Sylvester (1814-1897) inaugure ce que l'on peut appeler, à la suite de Hourya Sinaceur, l'« algébrisation du théorème de Sturm », et il publie entre 1839 et 1853 une dizaine d'articles ayant trait au travail de Sturm sur le nombre de racines réelles d'une équation. C'est donc au tout début de sa carrière, alors qu'il oscille entre l'enseignement des mathématiques, le travail d'actuaire et les études de droit, entre l'Angleterre et les États-Unis, qu'il se consacre à ces travaux.

1. Es fragt sich nun : Sind diese durch das Euclidische Verfahren zu gewinnenden Restfunctionen die einzigen, welche zur Aufstellung einer solchen Sturm'schen Reihe führen, oder giebt es noch andere auf andere Weise zu erlangende Functionen, welche dasselbe leisten ? [Kronecker, 1887b] p. 181

2. Kronecker fait bien évidemment allusion à [Sylvester, 1839],[Sylvester, 1841] et [Sylvester, 1853d].

En 1839, avec un article intitulé *On rational derivation from equations of coexistence, that is to say, a new and extended theory of elimination*, Sylvester commence donc cette longue série d'articles. Il travaille essentiellement sur les fonctions auxiliaires qui sont utilisées pour compter le nombre de racines réelles. Il les modifiera pour, dit-il, en trouver la « vraie nature » [3].

Dans un article de 1844 intitulé *On the existence of absolute criteria for determining the roots of numerical equations*, Sylvester cherche même à étendre le théorème de Sturm à la détermination des racines *rationnelles* d'une équation [4] :

> Furthermore, viewing the striking analogy of the general nature of the conditions of rationality already obtained, to those which serve to determine the reality of the roots of equations, I am strongly of opinion that a theorem remains to be discovered, which will enable us to pronounce on the existence of integer, as Sturm's theorem on that of possible roots of a complete equation of any degree : the analogy of the two cases fails however in this respect, that while imaginary roots enter an equation in pairs, irrational roots are limited to entering in groups, each containing *two* or MORE. [Sylvester, 1844] p. 445.

Nous allons maintenant examiner plus précisément les deux articles qui fondent l'ensemble de ce travail.

## 8.1 Les premiers travaux sur les suites de Sturm : 1839-1841

L'intérêt que Sylvester porte à ce théorème apparaît dès 1839 en annexe de l'un de ses premiers articles. Il s'agit d'un essai traitant de la théorie de l'élimination et qu'il publie dans le *Philosophical Magazine*, où il cherche des conditions pour que deux équations polynomiales d'une variable aient une racine commune : cet article représente aussi le début de ses travaux sur le résultant [5]. Il faut noter que cet article, qui arrive très tôt dans sa carrière, est le premier à traiter exclusivement d'algèbre.

Il s'agit pour Sylvester d'exposer un travail sur l'élimination dans le cas de deux équations à une inconnue, et il cherche à trouver une condition sur les coefficients des polynômes pour caractériser l'existence d'une racine commune. Ce faisant, il obtient des conditions qu'il exprime en fonction des racines des deux équations initiales. Il énonce alors sans démonstration, comme conséquence du travail qu'il vient de présenter, des formules qui peuvent servir de fonctions auxiliaires dans la méthode de Sturm.

En 1839, Sylvester rentre d'un séjour à Paris. S'il est difficile de savoir s'il a pu durant ce court voyage discuter avec Sturm, il affirmera en 1869 à propos de l'origine de la découverte de Sturm :

3. [Sylvester, 1841] p. 23 : « real nature ».

4. Cette citation conclue l'article dont elle est issue, et dans lequel Sylvester a cherché, pour certaines équations, à trouver un critère de rationalité des racines. Son travail n'aboutit dans son article que dans des cas où il est possible d'exprimer rationnellement les solutions de l'équation en fonction des coefficients.

5. Pour plus de détail sur les travaux de Sylvester sur la théorie de l'élimination, on pourra consulter [Penchèvre, 2006].

> Sturm's theorem about the roots of equations, which, as he informed me with his own lips, stared him in the face in the midst of some mechanical investigations connected with the motion of compound pendulums [Sylvester, 1908] p. 655.

Il semble donc que Sylvester, soit en 1839, soit plus tard, ait eu l'occasion de discuter avec Sturm de son théorème sur les racines réelles d'une équation. Ce qu'il introduit de nouveau dans cet article à propos du théorème de Sturm, il l'écrit à John Herschel [6] dans une lettre datée du 27 mai 1841 :

> Sturm's theorem is presented under a totally novel aspect and the relation between his functions and the squaring of the differences of the roots pointed out for the first time. [Parshall, 1998] p. 8

Nous allons maintenant examiner ce que fait, ou plutôt ce qu'affirme Sylvester dans cet article, et nous regarderons la démonstration que Sturm lui-même donne de ce résultat en 1842 dans le *Journal de Liouville*.

### 8.1.1 De nouvelles fonctions auxiliaires

Nous allons en fait examiner conjointement ce premier article de 1839 avec celui, très court, qu'il publie deux ans plus tard dans le *Plymouth British Association Report 1841* et qu'il intitule *On the relation of Sturm's auxiliary functions to the roots of an algebraic equation*. En effet, ce second article ne fait que reprendre la note de son article précédent, en modifiant quelques notations et en nommant les nouvelles fonctions auxiliaires décrites *Déterminateurs de Sturm* [7].

Sylvester propose des fonctions auxiliaires dans lesquelles interviennent les racines de l'équation proposée. Ce qu'affirme le théorème de Sylvester, et que les *Nouvelles Annales de mathématiques* proposent comme *théorèmes à démontrer* [8] en 1842, c'est que les fonctions auxiliaires de Sturm peuvent se mettre sous une *autre forme*. Il ne s'agit pas pour Sylvester de redémontrer les propriétés des fonctions de Sturm, considérées comme acquises. Il y a bien, comme le montre [9] Hourya Sinaceur, une volonté d'« algébrisation » du théorème de Sturm. Sylvester, qui parle de lui à la troisième personne, conclut ainsi son article de 1841 :

> He would conclude by calling attention to the importance to the general interests of algebraical and arithmetical science that a searching investigation should be instituted for showing, *a priori*, how, when a set of quantities is known to be made up partly of possible and partly of *pairs* of impossible values, symmetrical functions of these, one less in number than the quantities themselves, may be formed, from the signs of the ratios of which to unity and to one another the respective amounts of possible and impossible quantities may at once be inferred : in short, we ought not to rest satisfied, until, from the very *form* of Sturm's Determinators, without caring to know how they

6. John Herschel (1792-1871) est un astronome et mathématicien britannique.
7. [Sylvester, 1841] p. 23 : « Sturm's Determinators ».
8. Voir [Nouvelles-Annales, 1842].
9. Voir [Sinaceur, 1991] pp. 124 et suivantes.

> have been obtained, we are able to pronounce upon the uses to which they may be applied. [Sylvester, 1841] p. 59

Mais Sylvester ne propose pas une nouvelle méthode pour *démontrer* le théorème de Sturm. Le soubassement sur lequel il s'appuie reste la démonstration initiale de Sturm et donc la notion de continuité et le théorème de Bolzano.

Reprenons maintenant les notations de Sylvester dans [Sylvester, 1839]. Si $h_1, ..., h_n$ sont les solutions de l'équation, on notera :

$$\begin{aligned} SPD(h_1, ..., h_k) &= \text{le produit des carrés des différences de ces racines} \\ &= \prod_{1 \leq i < j \leq k} (h_i - h_j)^2 \\ S_i(h_1, ..., h_k) &= \text{le i}^{\text{ème}} \text{ polynôme symétrique élémentaire des variables } h_1, ..., h_k \end{aligned}$$

Sylvester propose alors de prendre comme fonctions auxiliaires de Sturm [10] :

$$\begin{aligned} V_n &= SPD(h_1, ..., h_n) \\ V_{n-1} &= \sum_i SPD(h_1, ..., h_{i-1}, h_{i+1}, ...h_n)x \\ &\quad + \sum_i S_1(h_1, ..., h_{i-1}, h_{i+1}, ...h_n) SPD(h_1, ..., h_{i-1}, h_{i+1}, ...h_n) \\ V_{n-2} &= \sum_{i,j} SPD(h_1, ..., h_{i-1}, h_{i+1}, ..., h_{j-1}, h_{j+1}, ...h_n)x^2 \\ &\quad + \sum_{i,j} S1(h_1, ..., h_{i-1}, h_{i+1}, ..., h_{j-1}, h_{j+1}, ...h_n) \\ &\quad \times SPD(h_1, ..., h_{i-1}, h_{i+1}, ..., h_{j-1}, h_{j+1}, ...h_n)x \\ &\quad + \sum_{i,j} S2(h_1, ..., h_{i-1}, h_{i+1}, ..., h_{j-1}, h_{j+1}, ...h_n) \\ &\quad \times SPD(h_1, ..., h_{i-1}, h_{i+1}, ..., h_{j-1}, h_{j+1}, ...h_n) \\ &\quad ... \end{aligned}$$

où l'on a noté $V_1, V_2, ...$ les fonctions en question. La notation $V_n$ est absente de l'article de Sylvester de 1839 : ce sont celles que Sturm emploiera en 1842. En fait, la notion même de fonction est absente du travail de Sylvester sur Sturm : il privilégie ce que H. Sinaceur appelle « une conception combinatoire de l'algèbre » [11].

10. Il s'agit des notations utilisées dans [Sylvester, 1839] et non [Sylvester, 1841].
11. Voir [Sinaceur, 1991] p. 125.

## 8.2 Démonstration par Sturm de la proposition de Sylvester

Dans un article qu'il intitulera *Démonstration d'un théorème d'algèbre de M. Sylvester*[12], Sturm reprend donc les résultats énoncés par Sylvester et en donner une démonstration. Le titre annonce clairement que Sturm attribue à ce dernier la paternité de ces fonctions auxiliaires, et nous allons voir lors de cette démonstration que le théorème est profondément modifié de manière à en permettre l'« algébrisation ».

Sturm formule différemment les fonctions auxiliaires données par Sylvester. Voici le théorème tel qu'il cherche à le démontrer :

*Notations de Sylvester*

Si l'on pose $V = 0$ une équation de degré $n$ à une inconnue $x$ et que l'on note $h_1, ..., h_n$ ses racines (supposées toutes distinctes), on a donc

$$V = (x - h_1)...(x - h_n)$$

Soit $V_1 = V'$ et $V_2, ..., V_n$ la suite habituelle des fonctions que Sturm obtient avec son algorithme d'Euclide modifié. On a alors :

$$\begin{aligned}
V_1 &= \sum_i \prod_{j \neq i} (x - h_j) \\
V_2 &= \frac{1}{n^2} \sum_{i<j} (h_i - h_j)^2 \prod_{k \neq i,j} (x - h_k) \\
V_3 &= \frac{1}{\lambda_3} \sum_{i_1<i_2<i_3} \prod_{1 \leq r<s \leq 3} (h_{i_r} - h_{i_s})^2 \prod_{k \neq i_1,i_2,i_3} (x - h_k) \\
V_4 &= \frac{1}{\lambda_4} \sum_{i_1<i_2<i_3<i_4} \prod_{1 \leq r<s \leq 4} (h_{i_r} - h_{i_s})^2 \prod_{k \neq i_1,i_2,i_3,i_4} (x - h_k) \\
... &= ... \\
V_l &= \frac{1}{\lambda_l} \sum_{i_1<i_2<...<i_l} \prod_{1 \leq r<s \leq l} (h_{i_r} - h_{i_s})^2 \prod_{k \neq i_1,i_2,...,i_l} (x - h_k) \\
... &= ... \\
V_n &= \frac{1}{\lambda_n} \prod_{1 \leq r<s \leq n} (h_r - h_s)^2
\end{aligned}$$

où les quantités $\lambda_3, ..., \lambda_n$ sont *indépendantes* de $x$ et sont des fonctions *symétriques* des racines $h_1, ..., h_n$ de l'équation $V = 0$.

Voyons maintenant comment Sturm s'y prend pour le démontrer. Soit $q_1, q_2, ..., q_{n-1}$ les quotients obtenus par la division euclidienne modifiée. Commençons déjà par remarquer

12. Voir [Sturm, 1842].

que ces quotients sont des polynômes de la forme $ax + b$. On a alors $V = V_1q_1 - V_2$ donc $V_2 = V_1q_1 - V$. De même, $V_1 = V_2q_2 - V_3$, donc $V_3 = V_2q_2 - V_1$. On obtient alors :

$$V_3 = V_1(q_1q_2 - 1) - Vq_2$$

et de la même façon

$$V_4 = V_1[(q_1q_2 - 1)q_3 - q_1] - V(q_2q_3 - 1)$$

On peut ainsi mettre tous les $V_k$ sous la forme :

$$V_k = V_1N_k - VP_k$$

avec $\deg(N_k) = k - 1$ et $\deg(P_k) = k - 2$.

En fait, les $\frac{N_k}{P_k}$ sont les réduites de la fraction continue

$$\frac{V}{V_1} = q_1 - \frac{1}{q_2 - \frac{1}{q_3 - \ldots}}$$

On obtient le résultat suivant :

Pour deux polynômes $V$ et $V_1$ de degré respectif $n$ et $n-1$ on peut toujours trouver des polynômes $T, Y$ et $Z$ de degré $n-k$, $k-1$ et $k-2$ tels que

$$T = V_1Y - VZ$$

Pour cela, il suffit de choisir $T = \lambda V_k$, $Y = \lambda N$ et $Z = \lambda P$ où $\lambda$ est indépendante de $x$

On veut alors montrer que tout triplet de polynômes qui conviendra sera toujours de cette forme. En effet, si l'on a :

$$\begin{cases} T = V_1Y - VZ \\ V_k = V_1N_k - VP_k \end{cases}$$

On a ainsi à partir de ces deux équations :

$$NT - V_kY = V(PY - NZ)$$

Si les deux membres de l'égalité ne sont pas nuls, on a :

$$\deg(NT - V_kY) \leq \sup((k-1) + (n-k), (n-k) + (k-1)) = n - 1$$

et

$$\deg(V(PY - NZ)) \geq n$$

ce qui est absurde, donc on a :

$$\begin{cases} NT - V_k Y = 0 \\ PY - NZ = 0 \end{cases}$$

et donc

$$\frac{T}{V_k} = \frac{Y}{N} = \frac{Z}{P} = \lambda$$

comme deg($T$)=deg($V_k$) et que $T = \lambda V_k$, le degré de $\lambda$ en $x$ est nul, et ainsi $\lambda$ ne dépend pas de $x$.

### 8.2.1 Expression de $V_2, V_3, ...$ en fonction de $x$ et des racines de $V = 0$

L'expression de $V_1$ est évidente, et s'obtient bien sûr par une simple dérivation de l'expression factorisée de $V$. Sturm ne fait pas une démonstration générale de l'expression de $V_k$, mais, pour « fixer les idées » [13] il s'occupe de $V_4$. Pour cela, on pose

$$Y = \sum_{i_1<i_2<i_3} \left( \prod_{1\leq r<s\leq 3} (h_{i_r} - h_{i_s})^2 \right) (x - h_{i_1})(x - h_{i_2})(x - h_{i_3})$$

Cette fonction est un polynôme symétrique des racines de $V = 0$ du troisième degré par rapport à $x$. $V_1 Y$ a alors pour degré $n + 2$. Si l'on effectue la division euclidienne de $V_1 Y$ par $V$, on obtient :

$$V_1 Y = VZ + T$$

où deg($Z$) = 2 et deg($T$) $\leq n - 1$. Observons maintenant la décomposition en éléments simples de $\frac{T}{V}$. Nous n'avons affaire qu'à des pôles simples, les racines de $V$ étant supposées toutes différentes. On a donc :

$$\frac{T}{V} = \frac{\alpha_1}{x - h_1} + ... + \frac{\alpha_n}{x - h_n}$$

La technique usuelle dans ce cas extrêmement simple nous donne :

$$\alpha_s = \frac{V_1(\alpha_r)Y(\alpha_r)}{V'(\alpha_r)} = \frac{V_1(\alpha_r)Y(\alpha_r)}{V_1(\alpha_r)} = Y(\alpha_r)$$

donc

$$\frac{T}{V} = \sum_{i=1}^{n} \frac{Y(\alpha_i)}{x - h_i}$$

Or, si l'on décompose en éléments simples la fraction

$$\frac{\sum_{i_1<i_2<i_3<i_4} \prod_{1\leq r<s\leq 4} (h_{i_r} - h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3,i_4} (x - h_k)}{(x - h_1)(x - h_2)...(x - h_n)}$$

on obtient aussi

$$\sum_{i=1}^{n} \frac{Y(\alpha_i)}{x - h_i}$$

13. [Sturm, 1842] p. 360.

donc

$$\frac{T}{V} = \frac{\sum_{i_1<i_2<i_3<i_4} \prod_{1\leq r<s\leq 4}(h_{i_r} - h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3,i_4}(x - h_k)}{V}$$

et donc

$$T = \sum_{i_1<i_2<i_3<i_4} \prod_{1\leq r<s\leq 4} (h_{i_r} - h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3,i_4} (x - h_k)$$

qui est bien de degré $n-4$. Or, on a

$$V_1 Y = VZ + T$$

c'est-à-dire que

$$T = V_1 Y - VZ$$

avec $T, Y$ et $Z$ de degrés respectifs $m-4$, 3 et 2. D'après ce qui précède, on doit donc avoir $T = \lambda V_4$, et donc :

$$V_4 = \frac{1}{\lambda} \sum_{i_1<i_2<i_3<i_4} \prod_{1\leq r<s\leq 4} (h_{i_r} - h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3,i_4} (x - h_k)$$

ce qui est bien l'égalité proposée par Sylvester. Sturm donne ensuite la forme des $Y$ à choisir pour obtenir les autres fonctions $V_k$.

### 8.2.2 Cauchy

Sturm poursuit en soulignant le lien qui existe entre ces résultats et le travail que Cauchy expose dans son cours sur les équations algébriques [14]. Sturm utilise en effet une formule issue de cet ouvrage qui prolonge la formule d'interpolation de Lagrange aux fractions rationnelles : la résolution de ce problème conduirait ici en effet à obtenir les expressions de $T$ et $Y$ à partir de celle de $\frac{T}{Y}$, et d'en déduire ensuite l'expression de $V_4$. Cette formule sera particulièrement importante pour Kronecker qui en fournit une étude poussée lors de la résolution de ce qu'il appelle le « problème de Cauchy » [15]. Les différentes réponses à ce problème, qu'il attribuera à Liouville, Jacobi ou Sylvester, seront au centre de notre prochain chapitre [16].

En fait, comme nous l'avons déjà dit, ce qui est exposé dans cette partie a pour but d'éclairer la source centrale de notre recherche, c'est-à-dire les *Vorlesungen über die Theorie der algebraischen Gleichungen* de Kronecker. Et comme nous pourrons le vérifier, nous retrouverons dans celles-ci la presque totalité des travaux algébriques produits sur Sturm en Europe au moment de leur rédaction, ainsi qu'une bonne partie des éléments techniques sur lesquels sont basés ces travaux.

14. [Cauchy, 1821] p. 528.
15. Voir par exemple [Kronecker, 1881a] p. 125.
16. Voir p. 253.

### 8.2.3 Expression des $\lambda$

On commence par poser les notations suivantes :

$$\begin{aligned}
T_2 &= \sum_{i<j}(h_i-h_j)^2 \prod_{k\neq i,j}(x-h_k)\\
T_3 &= \sum_{i_1<i_2<i_3}\prod_{1\leq r<s\leq 3}(h_{i_r}-h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3}(x-h_k)\\
T_4 &= \sum_{i_1<i_2<i_3<i_4}\prod_{1\leq r<s\leq 4}(h_{i_r}-h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3,i_4}(x-h_k)\\
\ldots &= \ldots\\
T_n &= \prod_{1\leq r<s\leq n}(h_r-h_s)^2
\end{aligned}$$

Pour chacun de ces polynômes $T_l$ de degré $n-l$, on a vu qu'il existait deux polynômes $Y_l$ et $Z_l$ de degré $l-1$ et $l-2$ tels que

$$T_l = V_1 Y_l - V Z_l$$

avec, comme on l'a vu :

$$\begin{aligned}
Y_2 &= \sum_{i=1}^{n}(x-h_i)\\
Y_3 &= \sum_{i_1<i_2}(h_{i_1}-h_{i_2})^2(x-h_{i_1})(x-h_{i_2})\\
Y_4 &= \sum_{i_1<i_2<i_3}\left(\prod_{1\leq r<s\leq 3}(h_{i_r}-h_{i_s})^2\right)(x-h_{i_1})(x-h_{i_2})(x-h_{i_3})\\
Y_5 &= \sum_{i_1<i_2<i_3<i_4}\left(\prod_{1\leq r<s\leq 4}(h_{i_r}-h_{i_s})^2\right)(x-h_{i_1})(x-h_{i_2})(x-h_{i_3})(x-h_{i_4})\\
&\ldots\ldots
\end{aligned}$$

Si maintenant on désigne par $p_l$ le coefficient du terme de plus haut degré en $x$ dans $T_l$, alors

$$\begin{aligned}
p_2 &= \sum_{i<j}(h_i-h_j)^2\\
p_3 &= \sum_{i_1<i_2<i_3}\prod_{1\leq r<s\leq 3}(h_{i_r}-h_{i_s})^2\\
p_4 &= \sum_{i_1<i_2<i_3<i_4}\prod_{1\leq r<s\leq 4}(h_{i_r}-h_{i_s})^2
\end{aligned}$$

.....

Par ailleurs, on a :

$$\begin{cases} T_l = V_1 Y_l - V Z_l \\ T_{l+1} = V_1 Y_{l+1} - V Z_{l+1} \end{cases}$$

donc

$$T_l Y_{l+1} - T_{l+1} Y_l = V(Y_l Z_{l+1} - Y_{l+1} Z_l)$$

En désignant comme précédemment par $N_l$ et $P_l$ les nominateurs et dénominateurs de la $l^{\text{ième}}$ réduite, on a vu que l'on avait :

$$T_l = \lambda_l V_l \qquad Y_l = \lambda_l N_l \qquad Z_l = \lambda_l P_l$$

et

$$Y_{l+1} = \lambda_{l+1} N_{l+1} \qquad Z_{l+1} = \lambda_{l+1} P_{l+1}$$

En faisant appel à la propriété classique des fractions continues affirmant que

$$N_l P_{l+1} - P_l N_{l+1} = 1$$

on obtient

$$Y_l Z_{l+1} - Y_{l+1} Z_l = \lambda_l \lambda_{l+1}(N_l P_{l+1} - P_l N_{l+1}) = \lambda_l \lambda_{l+1}$$

Donc on a

$$\begin{cases} T_l Y_{l+1} - T_{l+1} Y_l = V(Y_l Z_{l+1} - Y_{l+1} Z_l) \\ Y_l Z_{l+1} - Y_{l+1} Z_l = \lambda_l \lambda_{l+1} \end{cases}$$

ainsi

$$T_l Y_{l+1} - T_{l+1} Y_l = V \lambda_l \lambda_{l+1}$$

Quels sont les termes de plus haut degré dans les deux membres de cette égalité ? Le terme de plus haut degré de $T_l Y_{l+1}$ est $p_l^2 x^{n-l} x^l = p_l^2 x^n$, celui de $T_{l+1} Y_l$ est $p_{l+1} x^{n-l-1} p_{l-1} x^{l-1} = p_{l+1} p_{l-1} x^{n-2}$, donc celui de $T_l Y_{l+1} - T_{l+1} Y_l$ est $p_l^2 x^n$. Par ailleurs, celui de $V \lambda_l \lambda_{l+1}$ est $\lambda_l \lambda_{l+1} x^n$, donc on a, par identification,

$$p_l^2 = \lambda_l \lambda_{l+1}$$

Comme $\lambda_2 = n^2$, on peut facilement calculer les $\lambda_l$ de proche en proche :

$$\lambda_2 = n^2$$

$$p_2^2 = \lambda_2\lambda_3 = n^2\lambda_3 \text{ donc } \lambda_3 = \left(\frac{p_2}{n}\right)^2$$

...

On a finalement démontré que :

$$V_1 = \sum_i \prod_{j\neq i}(x-h_j)$$

$$V_2 = \frac{1}{n^2}\sum_{i<j}(h_i-h_j)^2 \prod_{k\neq i,j}(x-h_k)$$

$$V_3 = \left(\frac{n}{p_2}\right)^2 \sum_{i_1<i_2<i_3} \prod_{1\leq r<s\leq 3}(h_{i_r}-h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3}(x-h_k)$$

$$V_4 = \left(\frac{p_2}{np_3}\right)^2 \sum_{i_1<i_2<i_3<i_4} \prod_{1\leq r<s\leq 4}(h_{i_r}-h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3,i_4}(x-h_k)$$

$$V_5 = \left(\frac{np_3}{p_2p_4}\right)^2 \sum_{i_1<i_2<i_3<i_4<i_5} \prod_{1\leq r<s\leq 5}(h_{i_r}-h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3,i_4,i_5}(x-h_k)$$

.....

$$V_n = \left(\frac{p_2p_4...p_{n-2}}{np_3p_5...p_{n-1}}\right)^2 p_n \text{ si } n \text{ est pair}$$

$$= \left(\frac{np_3p_5...p_{n-2}}{p_2p_4...p_{n-1}}\right)^2 p_n \text{ si } n \text{ est impair}$$

Ce qui, comme l'écrit Sturm, « est là le théorème complet de M. Sylvester [17]. ».
On remarque, comme le fait Sturm, que les $\lambda$ étant positifs, on peut les négliger lors de la recherche du nombre de racines positives entre deux réels.

## 8.3 Retour à la note de 1841.

Maintenant que la « nouvelle » forme de ces fonctions a été exposée, examinons ce que Sylvester remarque à leur propos [18] :

> It may be here remarked, that the work of assigning the total number of real and of imaginary roots falls exclusively upon the coefficients of the leading terms, which the author proposes to call "Sturm's Superiors" : these superiors are only *partial* symmetric functions of the *squared differences*, but *complete* symmetric functions of the *roots themselves*, differing in the former respect from those other (at first sight similar-looking) functions of the squared differences of the roots, in which, from the time of Waring downwards, the conditions of reality have been sought for.

On retrouve dans ces expressions des fonctions sturmiennes les carrés des différences

17. Voir [Sturm, 1842] p. 364.
18. Voir [Sylvester, 1841] p. 24.

des racines telles que l'on peut les trouver dans la méthode de Waring-Lagrange. Par ailleurs, si l'on cherche à déterminer le nombre total de racines de l'équation, seul compte le terme de plus haut degré, qui est prépondérant au voisinage de $-\infty$ et de $+\infty$. Or ce terme est une fonction symétrique des racines, donc un nombre entier dans le cas présent, le polynôme $V$ ayant ses coefficients entiers [19]. On pourra même comme Cayley [20] exprimer ces fonctions grâce aux coefficients de la fonction $V$. C'est d'ailleurs finalement ce que cherche à faire Sturm dans la dernière partie de son article de 1842 lorsqu'il explique comment retrouver ces formules sans connaître les racines.

Ce que Sturm n'essaye donc pas de faire, c'est de redémontrer son théorème à l'aide de la nouvelle forme de ses fonctions : il tient cela pour acquis et il ne démontre que l'égalité des fonctions proposées par Sylvester avec les siennes. On retrouve aussi le même souci d'effectivité et de clarté que celui qui domine dans sa démonstration de 1835. L'étape suivante, qui concerne toujours essentiellement la forme de ces fonctions auxiliaires, sera faite, comme on vient d'y faire allusion, par Cayley, et l'examen de celle-ci sera l'objet du prochain chapitre. Mais pour le moment nous allons suivre un travail préliminaire sur lequel Kronecker s'attarde dans son chapitre sur le théorème de Sturm et dans lequel nous allons retrouver les résultats précédents sur les fractions continues.

## 8.4 *Digression* de Kronecker sur les développements en fractions continues

Dans *tous* les manuscrits, Kronecker propose un paragraphe, plus ou moins long, sur le développement en fractions continues. Cette notion, classique au XIX$^{e}$ siècle, lui permet de dérouler tous les résultats dont il aura besoin ensuite pour travailler sur les fonctions $f_h$, c'est-à-dire sur les éléments de la suite de Sturm qu'il a construite, redémontrant les formules que Sturm utilise dans sa démonstration du théorème de Sylvester (comme par exemple $N_l P_{l+1} - P_l N_{l+1} = 1$). C'est bien ainsi que Kronecker justifie l'introduction de ces « digressions » [21] :

> Avant que nous nous occupions plus en détail des recherches de Sylvester, qui suivent le

19. Il faut rappeler ici l'importance de l'utilisation de cette propriété qui découle du théorème fondamental des polynômes symétriques. Il s'agit en effet d'un outil puissant fréquemment utilisé au XIX$^{e}$ siècle et nous verrons d'ailleurs plus tard la virtuosité avec laquelle Gauss l'utilise lors de la seconde démonstration qu'il donne du théorème fondamental de l'algèbre. La démonstration qu'il a construite pour l'occasion du théorème fondamental des polynômes symétriques, basée sur un ordre lexicographique des monômes, constitue la démonstration classique enseignée aujourd'hui. Dans le cours sur les équations algébriques de Kronecker, la place accordée aux fonctions symétriques est particulièrement importante, et il présente le théorème de Sturm comme un exemple d'utilisation de ces fonctions symétriques.

20. Voir [Cayley, 1846a], sa démonstration dans [de Zeipel, 1860] ainsi que le travail de Koralek dans [Koralek, 1852] qui utilise les formules obtenues par Cayley pour les polynômes de degré deux, trois et quatre.

21. Bevor wir auf diese Sylvester'schen Untersuchungen, welche den doppelten Zweck verfolgen, erstens, den Sturm'schen Satz zu begründen, zweitens, eine Anwendung auf die symmetrischen Functionen zu machen, näher eingehen, erweist es sich als notwendig, tiefer in das Euclidische Kettenbruchverfahren und in die Theorie der Kettenbrüche überhaupt einzudringen.

> double but de, premièrement, fonder le théorème de Sturm, et ensuite de donner une application des fonctions symétriques, il se révèle nécessaire du reste d'approfondir la décomposition en fraction continues par la méthode d'Euclide et la théorie des fractions continues. [Kronecker, 1887b] p. 181

Examinons ce qu'il fait :

Soit donc deux fonctions entières $f$ et $f_1$ avec $f_1$ de degré moindre que celui de $f$. La division euclidienne modifiée de $f$ par $f_1$ donne :

$$\begin{aligned} f - g_1 f_1 + f_2 &= 0 \\ f_1 - g_2 f_2 + f_3 &= 0 \\ ... &= 0 \\ f_{\nu-1} - g_\nu f_\nu &= 0 \end{aligned}$$

On obtient ainsi le développement en fraction continue :

$$\frac{f_1}{f} = \cfrac{1}{g_1 - \cfrac{1}{g_2 - \cfrac{1}{... \cfrac{1}{g_\nu}}}}$$

Jusque là, Kronecker a essentiellement considéré le cas où $f_1$ est de degré $n-1$, et où donc chaque $g_h$ est une fonction linéaire. Mais, toujours dans un souci de *généralisation*, il suppose maintenant que les $g_h$ sont des fonctions entières à coefficients entiers quelconques.

Kronecker considère à partir de maintenant les $g_h$ comme des variables indépendantes les unes des autres et souhaite leur appliquer les résultats sur les fonctions entières à coefficients entiers de plusieurs variables qu'il a obtenus dans les chapitres précédents.

Posons $N(g_k, ..., g_h)$ le dénominateur de

$$\cfrac{1}{g_k - \cfrac{1}{g_{k+1} - \cfrac{1}{... \cfrac{1}{g_h}}}}$$

après simplification sous la forme d'une fraction irréductible. On veut de plus, ce qui est toujours possible, que le terme de plus haut degré de $N(g_1, ..., g_\nu)$ soit égal à 1.

Par ailleurs,

$$\frac{f_1}{f} = \frac{1}{g_1 - \frac{f_2}{f_1}}$$

et

$$\frac{f_2}{f_1} = \cfrac{1}{g_2 - \cfrac{1}{g_3 - \cfrac{1}{\ldots \cfrac{1}{g_\nu}}}}$$

donc

$$\frac{f_1}{f} = \frac{1}{g_1 - \frac{(\ldots)}{N(g_2 \ldots, g_\nu)}} = \frac{N(g_2 \ldots, g_\nu)}{g_1 N(g_2 \ldots g_\nu) - (\ldots)} = \frac{N(g_2 \ldots g_\nu)}{N(g_1 \ldots g_\nu)}$$

Car on voit facilement que $N(g_1, \ldots, g_\nu)$ et $N(g_2, \ldots, g_\nu)$ n'ont pas de facteurs communs.

En effet, comme

$$\cfrac{1}{g_2 - \cfrac{1}{g_3 - \cfrac{1}{\ldots \cfrac{1}{g_\nu}}}} = \frac{(\ldots)}{N(g_2 \ldots g_\nu)}$$

Par définition $(\ldots)$ et $N(g_2 \ldots g_\nu)$ sont premiers entre eux. Ainsi, si $g_1 N(g_2 \ldots g_\nu) - (\ldots)$ et $N(g_2 \ldots, g_\nu)$ ont un facteur commun, celui-ci divise $N(g_2, \ldots, g_\nu)$, donc aussi $(\ldots)$, ce qui est impossible : $g_1 N(g_2, \ldots g_\nu) - (\ldots)$ et $N(g_2 \ldots, g_\nu)$ sont premier entre eux et donc $g_1 N(g_2 \ldots g_\nu) - (\ldots) = N(g_1 \ldots g_\nu)$

On voit de même que

$$(\ldots) = N(g_3, \ldots, g_\nu)$$

et donc en égalisant les dénominateurs :

$$N(g_1, \ldots, g_\nu) = g_1 N(g_2, \ldots, g_\nu) - N(g_3, \ldots, g_\nu)$$

On a en fait la relation, pour tous $p, q$ tels que $q > p$ :

$$N(g_p, \ldots, g_{q-1}) - g_p N(g_{p+1}, \ldots, g_{q-1}) + N(g_{p+2}, \ldots, g_{q-1}) = 0$$

Kronecker propose alors d'étendre la formule, d'en donner une plus « générale » [22] :

> Mais on peut encore en déduire une formule sur les fractions continues plus générale, dans laquelle sont contenues, en tant que cas particulier, toutes les relations connues jusqu'alors entre fractions continues, et qui ne se trouve pas dans les livres. [Kronecker, 1887b] p. 183

Cette formule, qui correspond donc à l'un des apports de Kronecker sur cette partie des fractions continues, et que nous n'allons pas démontrer ici, est la suivante :

22. Man kann aber noch eine viel allgemeinere Kettenbruchformel ableiten, in welcher alle bisher bekannten Kettenbruchrelationen als Specialfälle enthalten sind, und die sich nicht in den Lehrbüchern findet.

$$N(g_{p+1},...,g_{q-1})N(g_{r+1},...,g_{s-1}) - N(g_{p+1},...,g_{r-1})N(g_{q+1},...,g_{s-1})$$
$$+N(g_{p+1},...,g_{s-1})N(g_{q+1},...,g_{r-1}) = 0$$

où $p \leq q \leq r \leq s$.

Pour simplifier l'écriture, on peut poser $N_{p,q} = N(g_{p+1},...,g_{q-1})$ et on peut écrire :

$$(*) \qquad N_{p,q}N_{r,s} - N_{p,r}N_{q,s} + N_{p,s}N_{q,r} = 0$$

**Remarque :** Dans la plupart des manuscrits il n'y a aucun traitement particulier pour le cas problématique $p = q$. En fait, Kronecker pose implicitement [23] $N_{p,p} = 0$. On a aussi $N_{p,p+1} = 1$

À partir de cette formule, il propose d'en déduire des propriétés connues des développements en fraction continue :

Nous allons prendre, pour $p, q, r$ et $s$ des valeurs particulières qui nous permettront d'obtenir les trois formules $(a), (b)$ et $(c)$, dont on tirera de nouvelles relations.

1. Posons tout d'abord $p = 0$, $q = h$, $r = k+1$, $s = k+2$, avec $h \leq k$. On obtient alors :

$$N(g_1,...,g_{h-1}) = N(g_1,...,g_k)N(g_{h+1},...,g_{k+1}) - N(g_1,...,g_{k+1})N(g_{h+1},...,g_k)$$

On a par définition,

$$\frac{N_{p,q}}{N_{p-1,q}} = \cfrac{1}{g_p - \cfrac{1}{g_{p+1} - \cfrac{1}{\ldots \cfrac{1}{g_{q-1}}}}}$$

Donc, en prenant $p = q - 1$

$$\frac{N_{q-1,q}}{N_{q-2,q}} = \frac{1}{g_{q-1}}$$

et ainsi $N_{q-1,q} = 1$ et $N_{q-2,q} = g_{q-1}$

Donc ici en posant $h = k$ on obtient

$$(a) \qquad N(g_1...g_{k-1}) = g_{k+1}N(g_1...g_k) - N(g_1...g_{k+1})$$

23. Dans [Kronecker, 1891a] p. 190, ceci est explicité : « Will man auch $p = q$ zulassen, so ist zu setzen $Z_{p,p} = 0$, $N_{p,p} = 0$. »

**Introduction des notations $\psi_k$, $\varphi_k$** Kronecker pose maintenant des notations qui seront particulièrement importantes pour la suite de l'exposé :

$$\cfrac{1}{g_1 - \cfrac{1}{g_2 - \cfrac{1}{\ldots \cfrac{1}{g_{k-1}}}}} = \frac{\varphi_k}{\psi_k} = \frac{N(g_2, ..., g_{k-1})}{N(g_1, ..., g_{k-1})}$$

où $\varphi_k$ et $\psi_k$ sont des « fonctions entières », donc des polynômes, des variables $g_1, ..., g_\nu$.

Comme par ailleurs $N(g_2, ..., g_{k-1})$ et $N(g_1, ..., g_{k-1})$ n'ont pas de facteur commun, on peut écrire :

$$\boxed{\varphi_k = N(g_2...g_{k-1}) = N_{1,k} \text{ et } \psi_k = N(g_1...g_{k-1}) = N_{0,k}}$$

On a de plus

$$\boxed{f_k = f_\nu N(g_{k+1}...g_\nu)}$$

On obtient ainsi, à partir de $(*)$ :

$$N_{0,k-1}N_{k,k+1} - N_{0,k}N_{k-1,k+1} + N_{0,k+1} = 0$$

et

$$N_{1,k-1}N_{k,k+1} - N_{1,k}N_{k-1,k+1} + N_{1,k+1} = 0$$

C'est-à-dire

$$\psi_{k+1} = g_k\psi_k - \psi_{k-1}$$

et

$$\varphi_{k+1} = g_k\varphi_k - \varphi_{k-1}$$

En considérant que $\varphi_1 = 0, \varphi_2 = 1$ et $\psi_1 = 1, \psi_2 = g_1$, on obtient ainsi deux formules de récurrence qui nous permettent de calculer les $\varphi_k$ et $\psi_k$. Notons $(R)$ l'ensemble de ces deux formules de récurrence muni des conditions initiales précédentes :

$$\boxed{(R) \qquad \begin{cases} \psi_{k+1} = g_k\psi_k - \psi_{k-1} \\ \varphi_{k+1} = g_k\varphi_k - \varphi_{k-1} \end{cases}}$$

2. Posons maintenant $h = 1$ dans

$$N(g_1, ..., g_{h-1}) = N(g_1, ..., g_k)N(g_{h+1}, ..., g_{k+1}) - N(g_1, ..., g_{k+1})N(g_{h+1}, ..., g_k)$$

on obtient alors

$$(b) \qquad 1 = N(g_1, ..., g_k)N(g_2, ..., g_{k+1}) - N(g_1, ..., g_{k+1})N(g_2, ..., g_k)$$

De cette dernière formule, qualifiée par Kronecker de « très connue » [24], on obtient :

$$\boxed{(R') \qquad \psi_k\varphi_{k+1} - \varphi_k\psi_{k+1} = 1}$$

3. Posons ici $p = 0$, $q = 1$, $r = k + 1$ et $s = \nu + 1$ on obtient :

$$(c) \qquad N(g_{k+2}, ..., g_n) = N(g_1...g_k)N(g_2...g_\nu) - N(g_1....g_\nu)N(g_2...g_k)$$

et donc en multipliant par $f_\nu$ :

$$f_\nu N(g_{k+2}, ..., g_\nu) = N(g_1...g_k)f_\nu N(g_2...g_\nu) - f_\nu N(g_1....g_\nu)N(g_2...g_k)$$

donc comme $f_k = f_\nu N(g_{k+1}...g_\nu)$ et $f_0 = f$, on a finalement :

$$\boxed{(R'') \qquad \boxed{f_k = \psi_k f_1 - \varphi_k f}}$$

**Conclusion** À partir de la formule originale de Kronecker $(*)$ et de l'égalité $f_k = f_r N(g_{k+1}...g_r)$, on a obtenu les trois formules $(R), (R')$ et $(R'')$ dont l'utilisation sera fondamentale dans la résolution du problème d'interpolation de Cauchy que nous aborderons dans la section suivante.

### 8.4.1 Expression de $N(g_1...g_\nu)$ sous la forme d'un déterminant

On rappelle que :

$$\begin{aligned} f - g_1 f_1 + f_2 &= 0 \\ f_1 - g_2 f_2 + f_3 &= 0 \\ ... &= 0 \\ f_{\nu-1} - g_\nu f_\nu &= 0 \end{aligned}$$

on obtient

$$\begin{aligned} -1 + g_1\frac{f_1}{f} - \frac{f_2}{f} &= 0 \\ -\frac{f_1}{f} + g_2\frac{f_2}{f} - \frac{f_3}{f} &= 0 \\ ... &= 0 \end{aligned}$$

24. [Kronecker, 1887b] p. 185 : « Sehr bekannte ».

$$-\frac{f_{\nu-1}}{f} - g_\nu \frac{f_\nu}{f} = 0$$

C'est donc un système linéaire de $\nu$ équations dont les inconnues sont $\frac{f_1}{f}, ...., \frac{f_\nu}{f}$. Il s'agit d'un système de Cramer, ce qui nous permet d'écrire :

$$\frac{f_1}{f} = \frac{\begin{vmatrix} 1 & -1 & ... & & & \\ 0 & g_2 & -1 & & & \\ & -1 & g_3 & -1 & & \\ & ... & ... & ... & ... & ... \\ & & & & -1 & g_\nu \end{vmatrix}}{\begin{vmatrix} g_1 & -1 & ... & & & \\ -1 & g_2 & -1 & & & \\ & -1 & g_3 & -1 & & \\ & ... & ... & ... & ... & ... \\ & & & & -1 & g_\nu \end{vmatrix}} = \frac{\begin{vmatrix} g_2 & -1 & & & & \\ -1 & g_3 & -1 & & & \\ ... & ... & ... & ... & ... & \\ & & & -1 & g_\nu & \end{vmatrix}}{\begin{vmatrix} g_1 & -1 & ... & & & \\ -1 & g_2 & -1 & & & \\ & -1 & g_3 & -1 & & \\ & ... & ... & ... & ... & ... \\ & & & & -1 & g_\nu \end{vmatrix}}$$

Le coefficient du terme du plus haut degré du dénominateur, donc celui de $g_1...g_\nu$ est égal à 1, et on a donc :

$$N(g_1...g_\nu) = \begin{vmatrix} g_1 & -1 & ... & & & \\ -1 & g_2 & -1 & & & \\ & -1 & g_3 & -1 & & \\ & ... & ... & ... & ... & ... \\ & & & & -1 & g_\nu \end{vmatrix}$$

Kronecker explique ici s'être inspiré du travail de *Fürstenau*. Il s'agit certainement de Eduard Fürstenau, auteur de l'ouvrage[25] *Darstellung der reellen Wurzeln algebraischer Gleichungen durch Determinanten der Coefficienten* paru en 1860.

Cette représentation n'est pas vraiment exploitée par Kronecker, et elle n'est d'ailleurs présente que dans les exemplaires du semestre d'hiver 1886-87. On la retrouve dans un contexte légèrement différent dans son article de 1878 sur le théorème de Sturm[26].

### 8.4.2 Conclusion

Ces idées ont une portée extrêmement générale : les $g_i$ sont ici des « variables » qui ne doivent posséder que les quelques propriétés arithmétiques classiques. Cette idée de voir des quantités comme des polynômes à plusieurs variables, et de considérer ce mécanisme comme la source d'une *arithmétique générale*, est à la base du travail de Kronecker que l'on trouve, par exemple, dans les *Grundzüge*.

On a obtenu de ce qui précède les degrés des diverses fonctions :

25. Voir [Fürstenau, 1860].
26. Voir [Kronecker, 1878c] p. 49.

$$deg(f_k) = n_k \,;\, deg(\varphi_k) = n_1 - n_{k-1} \,;\, deg(\psi_k) = n - n_{k-1} \,;\, deg(g_k) = n_{k-1} - n_k$$

Kronecker insiste sur le fait de ne pas se contenter du cas où $n_{k-1} = n_k + 1$, mais de prendre le *cas général*, qui est source de nouveaux aperçus qui se révèlent féconds [27] :

> On traite habituellement seulement le cas spécial où les $g$ sont des fonctions linéaires, c'est-à-dire où $n_{k-1} = n_k + 1$. Mais justement, si on ne se limite pas à cela, on obtient des résultats intéressants. [Kronecker, 1885b] p. 256

et de cette généralisation, il est déjà question dans son article de 1873 sur le théorème de Sturm [28] :

> À ma connaisance on s'est limité jusqu'ici au cas le plus simple, dans lequel toutes les différences $n_k - n_{k-1}$ sont égales à 1. [Kronecker, 1873b] p. 308

Comme Sturm le fait dans son article [29] de 1842 où il démontre les formules données par Sylvester [30] en 1839, Kronecker a procédé à un examen des degrés de chacune des fonctions qu'il a introduites. L'une des principales différences que l'on peut noter entre son travail et celui de Sturm réside dans le fait qu'il ne considère pas ici que $n_k = k$, et donne donc une version plus générale de ses résultats, au sens où ceux-ci doivent s'appliquer à une plus grande classe de fonctions.

Outre cette volonté, toujours présente, de donner à ses résultats la forme la plus générale possible, cette partie du cours, où Kronecker pose les bases qui lui permettront d'aborder les travaux de Sylvester puis de Cayley, participe bien de cette algébrisation du théorème de Sturm. En effet, comme le remarque Max Noether, dans les articles de Sylvester « la notion de *fonction* d'une grandeur variable continue ne joue aucun rôle et y est à peine mentionnée » [31]. C'est bien le cas dans cette partie, Kronecker ayant lui-même momentanément évacué la notion de continuité. Il faut maintenant, sa digression achevée, qu'il revienne au théorème de Sturm. Nous allons voir tout de suite comment, à partir d'un problème purement algébrique qu'il attribue à Cauchy, il peut, dans une certaine mesure, poursuivre cette algébrisation.

27. Man hat gewöhnlich nur den speciellen Fall behandelt, in dem die $g$ lineare Functionen sind, d.h $n_{k-1} = n_k + 1$. Aber gerade, wenn man diese Beschränkung nicht einführt, erhält man interessante Resultate.

28. man sich meines Wissens bisher auf den einfachsten Fall beschränkt hat, in welchem sämmtliche Differenzen $n_k - n_{k-1}$ gleich Eins sind.

29. Voir [Sturm, 1842].

30. Voir [Sylvester, 1839].

31. [Noether, 1898] p. 134-135 : « der Begriff der *Function* einer stetig variablen Grösse, keine Rolle spielt, ja kaum genannt wird. »

# Chapitre 9

# Le problème d'interpolation de Cauchy


9.1 Introduction . . . . . **253**
9.1.1 Cinq méthodes . . . . . 254
9.1.2 Exposé du problème . . . . . 254
9.1.3 Premières méthodes . . . . . 256
9.2 Méthode de Liouville . . . . . **257**
9.2.1 Description de la méthode . . . . . 258
9.2.2 Ensemble des solutions du problème de Cauchy . . . . . 259
9.3 Cayley et la Résolution du problème de Cauchy . . . . . **264**
9.3.1 Les recherches de Cayley (1821-1895) sur les fonctions de Sturm 264
9.3.2 La résolution par Kronecker . . . . . 272
9.4 Dernière résolution du problème de Cauchy . . . . . **293**
9.4.1 Redéfinition du problème . . . . . 293
9.4.2 Autre forme de $\mathcal{F}(x)$ . . . . . 297
9.4.3 Autre façon de mettre en évidence la relation qui unit $\mathcal{F}, \Psi$ et $\Phi$. 299
9.4.4 Quelques considérations sur la dimension de l'ensemble des solutions . . . . . 301


## 9.1 Introduction

Le problème dit d'interpolation de Cauchy – nous y avons fait allusion dans le travail de Sylvester que nous avons examiné en détail plus haut [1] – est utilisé par Kronecker pour approfondir encore un peu plus son travail sur le théorème de Sturm. Kronecker expose, à travers plusieurs présentations de la résolution de ce problème, des travaux qu'il attribuera à Cauchy, Liouville, Jacobi ou Cayley. Cette *méthode d'interpolation* constitue la *note V* du cours d'*analyse algébrique* de Cauchy [2]. Dans cette note, intitulée *Sur la formule*

1. Voir p. 240.
2. Voir [Cauchy, 1821] p. 525.

*de Lagrange relative à l'interpolation*, Cauchy commence par montrer comment on peut obtenir les formules d'interpolation de Lagrange à l'aide des déterminants de Vandermonde, et donc en fait de résultats sur les polynômes symétriques.

### 9.1.1 Cinq méthodes

Kronecker exposera exactement cinq méthodes distinctes pour résoudre ce problème [3] :

1. Deux de ces méthodes consistent en la résolution directe et classique de systèmes d'équations linéaires. Il les attribue à Cauchy et Baltzer.
2. La troisième, dite de Liouville, utilise les résultats sur les fractions continues que nous avons exposées dans le chapitre précédent.
3. La quatrième, qu'il nomme d'après Jacobi, est inspirée des travaux de Cayley et utilise la théorie des déterminants.
4. La dernière utilise, comme Cayley le fait aussi, un développement en puissances décroissantes.

Celles-ci ne seront pas toutes présentes dans l'ensemble des manuscrits, mais le problème de Cauchy sera toujours abordé à partir de l'exemplaire de 1880-1881, et constituera le *liant*, de la partie que Kronecker réservera à Sturm. S'il est absent du manuscrit du semestre d'hiver [4] 1872-1873, et aussi d'ailleurs de l'article de Kronecker de 1873 [Kronecker, 1873b], c'est qu'il n'apparaît pour la première fois, à notre connaissance, dans l'œuvre de Kronecker qu'en 1878 dans son article *über Sturmsche Functionen* [5].

### 9.1.2 Exposé du problème

#### 9.1.2.1 Le problème

Avant de poursuivre, examinons déjà comment Kronecker définit *le problème d'interpolation de Cauchy* dans son cours [6] du semestre d'hiver 1880-1881 :

3. Voir par exemple [Kronecker, 1881a] p. 126 : « Nous donnons pour la résolution du problème de *Cauchy* 5 méthodes différentes. » (*Wir geben für die Lösung des Cauchy'schen Problems 5 verschiedene Methoden.*)

4. Le cours de 1872-73 propose d'ailleurs une présentation particulière du théorème de Sturm : il n'est introduit qu'une fois tous les préliminaires techniques établis. Cette partie manque ainsi finalement d'unité, et il est particulièrement difficile dans ce cours de décider ce qui appartient ou non au chapitre sur Sturm.

5. Voir [Kronecker, 1878c].

6. Nach diesen vorbereitenden Sätzen über die Kettenbrüche und die *Sturm*'schen Entwicklungsfunktionen, fassen wir ein Problem ins Auge, welches mit der Darstellung der *Sturm*'schen Funktionen als symmetrische Funktionen in engem Zusammenhange steht, nämlich das Problem der *Cauchy*'schen Interpolationsformel. Dasselbe lautet : Es soll der Quotient zweier ganzen rationalen Funktionen $\frac{F(x)}{\Psi(x)}$ so bestimmt werden, daß er für $x = \xi_h$ $(h = 1, ...n)$ den Werth $\mathcal{C}_h$ $(h = 1, ...n)$ annimmt. Sei $F(x)$ von Grade $\mu$, $\Psi(x)$ vom Grade $\nu$, so soll also :

$$\frac{\sum_{r=0}^{\mu} a_r \xi_h^r}{\sum_{s=0}^{\nu} b_s \xi_h^s} = \mathcal{C}_h \ (h = 1, ...n) \text{ sein.}$$

Es sollen also die $\mu + \nu + 2$ unbekannten Coeffizienten $a_0, ...a_\mu, b_0 ...b_\nu$ so bestimmt werden, daß die $n$ Gleichungen bestehen. Wir haben es also mit einem Problem der Auflösung homogener linearer Gleichungen

> Après ces propositions préliminaires sur les fractions continues et les fonctions de *Sturm*, nous envisageons un problème qui est en étroit rapport avec la représentation des fonctions de *Sturm* comme fonctions symétriques, à savoir le problème de la formule d'interpolation de *Cauchy*. Celui-ci s'énonce ainsi : le quotient de deux fonctions rationnelles entières $\frac{F(x)}{\Psi(x)}$ doit être déterminé de sorte qu'il prenne pour $x = \xi_h$ $(h = 1, ...n)$ la valeur $\mathcal{C}_h$ $(h = 1, ...n)$. Soit $F(x)$ de degré $\mu$, $\Psi(x)$ de degré $\nu$, alors on doit donc avoir :
>
> $$\frac{\sum_{r=0}^{\mu} a_r \xi_h^r}{\sum_{s=0}^{\nu} b_s \xi_h^s} = \mathcal{C}_h \ (h = 1, ...n)$$
>
> Les $\mu + \nu + 2$ coefficients inconnus $a_0, ...a_\mu, b_0...b_\nu$ doivent ainsi être déterminés de sorte que les $n$ équations soient satisfaites. Nous avons donc à faire à un problème de résolution d'équations linéaires homogènes et nous savons que le nombre de ces dernières, si ce problème doit avoir une solution, doit être au moins égal à $\mu + \nu + 1$. Comme première condition pour $\mu$ et $\nu$, nous obtenons de là que $\mu + \nu + 1 \leq n$ ou $\underline{\mu + \nu < n}$. [Kronecker, 1881a] p. 125

Il s'agit donc d'une « généralisation de la formule d'interpolation de Lagrange »[7]. La formule d'interpolation de Lagrange[8], rappelons-le, permet de déterminer un polynôme $P(x)$ de degré au plus $n-1$ qui prend $n$ valeurs déterminées pour $n$ valeurs de $x$ distinctes[9]. Kronecker veut étendre cette « idée » (*Gedanke*) : une fois encore, il s'attache à mettre en avant dans son exposition du cours sur le théorème de Sturm une certaine forme de généralisation. Ici, c'est l'« idée » qui est généralisée, au sens où il souhaite, par *analogie*, accroître l'ensemble des fonctions auxquelles cette formule s'applique. Dans la définition précédente, on voit facilement que si $\nu = 0$, alors la proposition s'apparente au théorème de Lagrange : on est bien dans le cas classique de la généralisation par inclusion d'une proposition. Mais Kronecker précise que c'est l'*idée* qui est généralisée. Nous verrons que si la théorie des caractéristiques est bien une généralisation du théorème de Sturm, il s'agit essentiellement d'une généralisation de son *idée*.

#### 9.1.2.2 Le rapport au théorème de Sturm

Avant de poursuivre et d'entrer dans le détail des principales méthodes que Kronecker développe, examinons le rapport qu'il y a entre ce problème et le théorème de Sturm. Le but de la résolution du problème de Cauchy est la détermination de deux fonctions $F$ et $\Psi$ dont on connaît le quotient pour un certain nombre de valeurs de la variable. Nous montrerons que, sous certaines conditions, les fonctions $F$ construites sont en fait les fonctions d'une

zu thun und wissen, daß die Anzahl der letzteren, falls das Problem lösbar sein soll, mindestens gleich $\mu + \nu + 1$ sein muß. Als erste Bedingung für $\mu$ und $\nu$ erhalten wir daher $\mu + \nu + 1 \leq n$ oder $\underline{\mu + \nu < n}$.

7. [Kronecker, 1887b] p. 188 : « Verallgemeinerung der Lagrange'schen Interpolationsformel ».

8. On la trouve dans [Lagrange, 1867] pp. 284-286.

9. Il ne s'agit pas ici du problème d'interpolation lié à ce que l'on nomme aujourd'hui les *déterminants de Cauchy* et qui consiste à chercher une fraction rationnelle qui admet exactement $n$ pôles simples distincts connus.

suite de Sturm (méthode de Liouville). Cela nous donnera une un moyen efficace d'utiliser le théorème de Sturm, sans avoir à calculer la suite de fonctions $f_k$ (Méthode de Jacobi).

### 9.1.3 Premières méthodes

#### 9.1.3.1 La méthode de Cauchy

Kronecker propose donc plusieurs méthodes pour résoudre ce problème, et dans le dernier paragraphe de la citation précédente, il présente finalement une première méthode : on peut ramener le problème à la recherche de solution d'un système de plusieurs équations linéaires à plusieurs inconnues. Avant de voir ce qu'il pense de cette procédure, on peut déjà remarquer qu'elle apporte tout de même une condition sur les degrés du numérateur et du dénominateur de la fonction à trouver :

$$deg(\text{numérateur}) + deg(\text{dénominateur}) < n$$

Kronecker remarque [10] que l'on peut chercher à résoudre le problème de Cauchy directement, en considérant les coefficients des fonctions à déterminer comme des inconnues et en formant ainsi un système d'équations linéaires. Il s'agit finalement de la méthode employée par Cauchy lui-même dans sa *Note V*, et qui consiste à résoudre le système formé par les égalités

$$F(\xi_h) = C_h \Psi(\xi_h) \text{ pour } h = 1, .., n$$

#### 9.1.3.2 Méthode attribuée à Richard Baltzer

La seconde méthode, elle aussi qualifiée par Kronecker de « directe », est issue du livre intitulé *Theorie und Anwendung der Determinanten* [11] et publié par Richard Baltzer en 1857. Baltzer cite d'ailleurs à propos de sa démonstration un article [12] de Johann Georg Rosenhain [13] dans lequel ce dernier utilise la formule de Cauchy pour ses travaux sur le résultant.

*Construction des fonctions $f$ et $f_1$*

Dans tout ce chapitre, nous utiliserons les deux fonctions suivantes : on commence par construire, grâce à la formule d'interpolation de Lagrange une fonction $f_1(x)$ de degré au plus $n-1$ qui prend les valeurs $\mathcal{C}_1, .., \mathcal{C}_n$ pour $x = \xi_1, ...\xi_n$. On doit donc avoir :

$$\frac{F(\xi_h)}{\Psi(\xi_h)} = \mathcal{C}_h = f_1(\xi_h) \text{ (h=1,...n)}$$

On pose $f(x) = (x-\xi_1)(x-\xi_2)...(x-\xi_n)$.

10. On pourra voir par exemple [Kronecker, 1879] p. 416 ou [Kronecker, 1881a] p. 178.
11. Voir [Baltzer, 1857] p. 119.
12. Voir [Rosenhain, 1846].
13. Johann Georg Rosenhain (1816-1887) est un mathématicien allemand qui a passé sa thèse d'habilitation à Breslau sous la direction de Jacobi.

Il existe alors $\Phi$ tel que

$$F(x) = f_1(x)\Psi(x) + f(x)\Phi(x)$$

Remarquons que

$$\frac{F}{\Psi} = f_1 + f\frac{\Phi}{\Psi}$$

Avec $deg(F) = \mu$ et $deg(\Psi) = \nu$, on a $\mu < n$, et ainsi $deg(\Phi) < deg(\Psi)$, c'est-à-dire $deg(\Phi) \leq \nu - 1$.

Ainsi on a :

$$deg(f_1(x)\Psi(x) + f(x)\Phi(x)) = n + \nu$$

et aussi, comme on vient de le voir, il faut que $\mu + \nu < n$, donc que $n - \mu > \nu$

$F$ étant au plus de degré $n - \nu - 1$, on a $\mu < n - \nu - 1$, et ainsi il y a

$$n + \nu - \mu > 2\nu$$

coefficients de $f_1(x)\Psi(x)+f(x)\Phi(x)$ qui sont nuls, ce qui nous donne le nombre d'équations linéaires nécessaires pour trouver les $2\nu+1$ coefficients des fonctions $\Psi$ et $\Phi$. Cette méthode est qualifiée par Kronecker de la « méthode la plus maladroite » [14].

Mais, remarque Kronecker, lorsque le déterminant du système est nul, ces deux méthodes ne conviennent plus. Il dit à ce propos de celles-ci qu'elles sont « "générales" en apparence » (*scheinbar „allgemeine"*) [15] au sens ou le cas *général*, qui est celui où tout fonctionne correctement, n'est en fait qu'un cas particulier. Kronecker fait là la distinction entre ce qui fonctionne en général et le *cas général*. Mais, avec l'aide du travail qu'il a déjà effectué sur les fractions continues, il peut contourner le problème, et c'est ce que nous examinerons dans la méthode dite de Liouville.

## 9.2 Méthode de Liouville

*Résumé*

Kronecker utilise le travail sur les fractions continues qu'il a réalisé précédemment à partir des fonctions $f$ et $f_1$ pour montrer que, sous certaines conditions et pour $k$ bien choisi, $\frac{f_k}{\psi_k}$ est une solution du problème de Cauchy. Il prolonge son travail en s'interrogeant sur l'ensemble des solutions du problème de Cauchy.

Si Kronecker donne ce nom à cette méthode, c'est dû au fait que [16] :

14. [Kronecker, 1879] p. 417 : « die ungeschickteste Methode »

15. [Kronecker, 1887b] p. 189 : « man erkennt hier, daß der scheinbar „allgemeine" Fall in Wirklichkeit nur ein sehr specieller ist. »

16. Die Idee der Verbindung des *Cauchy*'schen mit der Kettenbruchentwicklung rührt von *Liouville* her. Sie findet sich veröffentlicht in dem Aufsatze von *Sturm* : « *Démonstration d'un théorème d'Algèbre de M. Sylvester* » (*Liouville's Journal Bd.7, pag. 356* im Jahre 1842).

> L'idée de la relation du travail de *Cauchy* avec la décomposition en fraction continue vient de *Liouville*. Elle se trouve publiée dans l'essai de *Sturm* : « *Démonstration d'un théorème d'Algèbre de M. Sylvester* » (*Journal de Liouville Bd.7, pag. 356* en 1842. [Kronecker, 1881a] p. 126

En effet, dans cet article que nous avons déjà présenté [17], Sturm utilise ce problème de Cauchy et affirme que :

> M. Liouville m'a fait voir que les résultats précédents se déduisent aussi d'une formule que M. Cauchy a donné dans son *Cours d'Analyse algébrique*, note 5, page 528, et par laquelle on détermine une fraction rationnelle, dont on connaît un certain nombre de valeurs particulières. [Sturm, 1842] p. 361

### 9.2.1 Description de la méthode

On commence par construire, grâce à la formule d'interpolation de Lagrange une fonction $f_1(x)$ de degré au plus $n-1$ qui prend les valeurs $\mathcal{C}_1, .., \mathcal{C}_n$ pour $x = \xi_1, ...\xi_n$. On doit donc avoir :

$$\frac{F(\xi_h)}{\Psi(\xi_h)} = \mathcal{C}_h = f_1(\xi_h) \ (h{=}1,...\text{n})$$

On pose $f(x) = (x-\xi_1)(x-\xi_2)...(x-\xi_n)$ et on construit comme on l'a vu précédemment grâce à la décomposition en fraction continue les suites de fonctions $f_k, \varphi_k$ et $\psi_k$.

Mais on a vu dans le chapitre précédent p. 249 que

$$f_k = \psi_k f_1 - \varphi_k f$$

On a ainsi pour $x = \xi_h$ : $f_k(\xi_h) = \psi_k(\xi_h) f_1(\xi_h)$

et donc, en supposant que $f_k$ et $\psi_k$ n'ont aucun facteur commun, on a

$$f_1(\xi_h) = \frac{f_k(\xi_h)}{\psi_k(\xi_h)} \ (h = 1, ...n)$$

Car en effet, comme

$$f_1(x)\psi_k(x) - f(x)\varphi_k(x) = f_k(x)$$

si $f_k$ et $\psi_k$ ont un facteur commun, c'est encore le cas pour $\psi_k$ et $f(x)\varphi_k(x)$. Mais $\psi_k$ et $\varphi_k$ sont premiers entre eux, donc $f$ et $\psi_k$ ont un facteur commun et ainsi il existe $h$ tel que $\psi_k(\xi_h) = 0$.

On a donc trouvé une fonction rationnelle $\frac{f_k(x)}{\psi_k(x)}$ qui prend pour $x = \xi_h$ les valeurs $\mathcal{C}_h = f_1(\xi_h)$. De plus la somme des degrés de $\frac{f_k(x)}{\psi_k(x)}$ est $n + n_k - n_{k-1} < n$ car $n_k < n_{k-1}$. Le problème de Cauchy est donc résolu dans le cas où $f_k$ et $\psi_k$ n'ont aucun facteur commun.

17. Voir p. 237.

Nous allons maintenant examiner de plus près ce qu'il se passe si finalement $f_k$ et $\psi_k$ ont un facteur commun, et nous regarderons ensuite comment Kronecker réussit à caractériser l'*ensemble* des solutions du problème de Cauchy.

### 9.2.2 Ensemble des solutions du problème de Cauchy

Dans une annexe au cours du semestre d'hiver 1880-1881 [18], mais aussi dans le corps du texte de celui de 1878-1879 [19] ou encore celui de 1886-1887 [20], Kronecker se pose la question suivante : quel est l'*ensemble* des solutions du problème de Cauchy ? Ce qu'il fait pour répondre à cette question ne varie pas au cours des années, en revanche ce qu'il tirera de ses calculs change de forme : la notion de *variété* (*Mannigfaltigkeit*) apparaît.

**Un problème équivalent** La première chose à noter est que Kronecker utilise une formulation de ce problème légèrement différente. Ce changement lui permet de donner une version *entière* à ce problème de fractions rationnelles, et dans les cours les plus tardifs, d'utiliser la notion de congruence qui joue un rôle si important après les *Grundzüge*. Ce problème équivalent est le suivant :

Soient deux fonctions $f(x)$ et $f_1(x)$ de degré $n$ et $n_1$. On cherche à déterminer trois fonctions entières $\Psi$, $\Theta$ et $F$ telles que

$$(E) \qquad f_1(x)\Psi(x) + f(x)\Theta(x) = F(x)$$

avec $F$ de degré $\mu$ et $\Psi$ de degré $\nu$ et $\mu + \nu < n$.

En effet, si $\dfrac{F}{\Psi}$ est une solution du problème de Cauchy, alors pour $\xi_1, ..., \xi_n$,

$$\frac{F(\xi_i)}{\Psi(\xi_i)} = f_1(\xi_i)$$

donc pour $\xi_1, ..., \xi_n$,

$$F(\xi_i) - f_1(\xi_i)\Psi(\xi_i) = 0$$

et ainsi on peut trouver $\Phi$ telle que

$$F(x) - f_1(x)\Psi(x) = f(x)\Phi(x)$$

On a bien ainsi

$$f_1(x)\Psi(x) + f(x)\Theta(x) = F(x)$$

en posant $\Theta = \Phi$.

On fixe $k$

18. Voir [Kronecker, 1881a] p. 177.
19. Voir [Kronecker, 1879] p. 365.
20. Voir [Kronecker, 1887b] p. 189.

La suite $n_k$, qui représente, on le rappelle [21], le degré de la fonction $f_k$, est décroissante avec $n_0 = n$ et $n_\nu = 0$, donc comme $\mu < n$, il existe $k$ tel que

$$n_k \leq \mu < n_{k-1}$$

On vient de voir que $F = f_k$, $\Psi = \psi_k$ et $\Theta = -\varphi_k$ était une solution à ce problème. Si $F, \Psi, -\Phi$ en est une autre on a :

$$f_k = \psi_k f_1 - \varphi_k f$$
$$F = \Psi f_1 - \Phi f$$

Ce qui donne, en éliminant $f_1$ des deux équations,

$$\psi_k.F - f_k\Psi = f(\varphi_k.\Psi - \psi_k\Phi)$$

À partir de là, le $k$ étant fixé, et comme le degré de $\psi_k F$ est $n - n_{k-1} + \mu$ et celui de $f_k\Psi$ est $n_k + \nu$, on a

$$deg(\psi_k.F) = n - n_{k-1} + \mu < n - n_{k-1} + n_{k-1} = n$$

et

$$deg(f_k\Psi) = n_k + \nu \leq \mu + \nu < n$$

donc

$$deg(\psi_k.F - f_k\Psi) < n$$

Comme $\psi_k.F - f_k\Psi$ est divisible par $f$ qui est de degré $n$, on a

$$\psi_k.F - f_k\Psi = 0$$

et donc

$$\frac{F}{\Psi} = \frac{f_k}{\psi_k}$$

On a aussi bien sûr

$$\frac{\Phi}{\Psi} = \frac{\varphi_k}{\psi_k}$$

On rappelle que l'on a construit $\varphi_k$ et $\psi_k$ tels qu'ils n'aient pas de facteur commun : on peut donc poser

21. Voir p. 250.

$$\Phi(x) = \Theta(x)\varphi_k(x)$$
$$\Psi(x) = \Theta(x)\psi_k(x)$$

et ainsi

$$F(x) = \Psi(x)f_1(x) - \Phi(x)f(x) = \Theta(x)(\psi_k(x)f_1(x) - \varphi_k(x)f(x)) = \Theta(x)f_k(x)$$

Donc si $F$ et $\Psi$ répondent au problème de Cauchy, elles sont de la forme :

$$F(x) = \Theta(x)f_k(x)$$
$$\Psi(x) = \Theta(x)\psi_k(x)$$

où $\Theta$ est le PGCD (*der größte gemeinschaftliche Teiler* [Kronecker, 1887b] p. 190) de $\Psi$ et $\Phi$.

**Cas où $f_k$ et $\psi_k$ ont un facteur commun** On a vu que dans le cas où $f_k$ et $\psi_k$ n'ont pas de facteur commun, alors $\frac{f_k}{\psi_k}$ est une solution du problème de Cauchy. Voyons ce qu'il se passe si ce n'est plus le cas.

On cherche donc une fraction rationnelle $\frac{F}{\Psi}$ où $deg(F) = \mu$ et $\deg(\Psi) = \nu$ telles que

1. $\frac{F}{\Psi}(\xi_i) = \mathcal{C}_i$ pour tout $i \in [\![1\ ;\ n]\!]$.
2. $\nu + \mu < n$

Pour cela on prend donc $f(x) = (x-\xi_1)....(x-\xi_n)$ et $f_1$ le polynôme de degré $n-1$ obtenu par la méthode d'interpolation de Lagrange tel que $f_1(\xi_i) = \mathcal{C}_i$ pour tout $i \in [\![1\ ;\ n]\!]$. Notre travail précédent montre que si des fonctions rationnelles conviennent, elles seront de la forme $\frac{\Theta(x)f_k(x)}{\Theta(x)\psi_k(x)}$. Nous avons vu que si $f_k$ et $\psi_k$ ont un facteur commun, alors il existe $h$ tel que $\psi_k(\xi_h) = 0$.

Alors $\psi_k$, $f_k$ et $f$ ont un facteur commun. En notant $PGCD(\psi_k, f)$ ce facteur, on a :

$$f_1(x)\frac{\psi_k}{PGCD(\psi_k,f)}(x) - \frac{f}{PGCD(\psi_k,f)}(x)\varphi_k(x) = \frac{f_k}{PGCD(\psi_k,f)}(x)$$

et donc, comme maintenant $\frac{\psi_k}{PGCD(\psi_k,f)}(\xi_h) \neq 0$ :

$$f_1(x) = \frac{\frac{f_k}{PGCD(\psi_k,f)}(x)}{\frac{\psi_k}{PGCD(\psi_k,f)}(x)} + \varphi_k(x)\frac{\frac{f}{PGCD(\psi_k,f)}(x)}{\frac{\psi_k}{PGCD(\psi_k,f)}(x)}$$

où bien sûr ces fractions sont en fait des fonctions entières. Soit $\xi$ une racine commune à

$f(x) = 0$ et $PGCD(\psi_k, f)(x) = 0$. Cette racine existe car on a supposé l'existence d'une racine commune entre $f(x) = 0$ et $\psi_k(x) = 0$. Dans ce cas

$$\frac{f}{PGCD(\psi_k, f)}(\xi) \neq 0$$

et donc $\dfrac{\dfrac{f_k}{PGCD(\psi_k, f)}(\xi)}{\dfrac{\psi_k}{PGCD(\psi_k, f)}(\xi)} = \dfrac{f_k(\xi)}{\psi_k(\xi)} \neq f_1(\xi)$, et dans ce cas le problème de Cauchy n'a donc pas de solution.

**Conclusion** Le problème de Cauchy n'a pas de solution lorsque $\psi_k(x)$, et donc aussi $f_k(x)$ n'est pas premier avec $f(x)$. Kronecker fait remarquer que ce cas particulier est traité par la méthode qu'il propose, même si la réponse est une absence de solution, contrairement à ce qui se passe lorsque l'on cherche directement à résoudre un système d'équations linéaires dont les inconnues sont les coefficients du numérateur et du dénominateur de la fraction rationnelle.

**Exemple** Kronecker poursuit en donnant un exemple où le problème de Cauchy *n'a pas* de solution. Soit

$$f(x) = (x-1)(x-2)(x-3)(x-a) = x^4 - (6+a)x^3 + (11+6a)x^2 - (6+11a)x + 6a$$

et donc $\xi_1 = 1, \xi_2 = 2, \xi_3 = 3, \xi_4 = a$. Soit de plus

$$f_1(x) = x^2 - 6x + 11$$

avec donc $f_1(\xi_1) = 6, f_1(\xi_2) = 3, f_1(\xi_3) = 2, f_1(\xi_4) = a^2 - 6a + 11$. On veut déterminer une fraction rationnelle $\dfrac{\mathcal{F}}{\Psi}$ qui prend les valeurs $6, 3, 2, a^2 - 6a + 11$ pour $1, 2, 3, a$.

et on obtient facilement :

$$\begin{aligned}
f_2(x) &= 6(x-a) \\
f_3(x) &= (6-a)a - 11 \\
g_1(x) &= x(x-a) \\
g_2(x) &= \frac{1}{4}(x-a-b) \\
g_3(x) &= \frac{6(x-a)}{(6-a)a-11}
\end{aligned}$$

La formule[22] de récurrence $(R)$ nous permet de calculer les $\psi_k$, et on obtient $\psi_2 = g_1 =$

22. Voir plus haut p. 248.

$x(x-a)$. On veut que la somme des degrés du numérateur et du dénominateur soit strictement inférieure à 4, et $deg(f_2) + deg(\psi_2) = 3$, donc une solution du problème de Cauchy devrait être de la forme :

$$\frac{f_2}{\psi_2} = \frac{6(x-a)}{x(x-a)} = \frac{6}{x}$$

Cette dernière correspond bien à $f_1(x)$ pour $x = 1, 2, 3$ mais pas pour $x = a$ en général. Il y a tout de même des cas particuliers où cela fonctionne, c'est-à-dire si $\frac{6}{a} = a^2 - 6a + 11$, donc si $a = 1, 2$ ou 3.

Regardons maintenant ce qu'il se passe lorsque l'on cherche à passer par la résolution d'un système d'équations.

Pour des raisons de degré, la fraction rationnelle cherchée peut se mettre sous la forme :

$$\frac{\mathcal{F}(x)}{\Psi(x)} = \frac{\alpha + \alpha_1 x}{\beta + \beta_1 x + \beta_2 x^2}$$

Grâce aux différentes valeurs que doit prendre cette fonction, on obtient

$$\frac{\mathcal{F}(x)}{\Psi(x)} = \frac{6\beta_1 + 6\beta_2 x}{x(\beta_1 + \beta_2 x)} = \frac{6}{x}$$

donc le même résultat que précédemment.

**Conclusion** Nous avons démontré que :

1. les fonctions $f_k$ et $\psi_k$ sont des solutions du problème de Cauchy si et seulement si elles n'ont pas de facteur commun.
2. Toutes fonctions $F$ et $\Psi$ qui répondent au problème de Cauchy sont de la forme :

$$F(x) = \Theta(x) f_k(x)$$
$$\Psi(x) = \Theta(x) \psi_k(x)$$

**Quelques considérations sur l'ensemble des solutions du problème de Cauchy** Ce que Kronecker vient en fait de montrer, c'est que si une fonction rationnelle $\frac{F}{\Psi}$ répond au problème, alors elle sera de la forme $\frac{F}{\Psi} = \frac{\Theta f_k}{\Theta \psi_k}$ On ne trouvera donc pas une solution unique au problème de Cauchy, mais leur ensemble constituera une *variété* (*Mannigfaltigkeit*) dont le degré dépendra des conditions sur celui de $\Theta$.

Examinons la fonction $\Theta$, et en particulier son degré que l'on notera $\lambda$. De l'inégalité $deg(F) + deg(\Psi) < n$ on tire

$$n_k + \lambda + n - n_{k-1} + \lambda < n$$

et donc on a la condition

$$\lambda < \frac{1}{2}(n_{k-1} - n_k)$$

Ainsi, dans le cas que Kronecker appelle « régulier » (*regulären*), c'est-à-dire où $n_{k-1} - n_k = 1$, on voit que $\lambda = 0$ et donc que $\Theta$ est une constante et l'ensemble des solutions est une « variété simple » (*einfache Mannigfaltigkeit*).

Nous venons ainsi de décrire la méthode, issue de l'article de Sturm de 1842, que Kronecker utilise pour résoudre le problème de Cauchy. Un travail sur les déterminants, inspiré de Cayley, lui permet de poursuivre son étude.

## 9.3 Cayley et la Résolution du problème de Cauchy

### 9.3.1 Les recherches de Cayley (1821-1895) sur les fonctions de Sturm

On sait que Cayley et Sylvester se sont rencontrés en 1846, alors qu'ils étudiaient tous deux le droit à Lincoln's Inn :

> we are told that during the following years they might often be found walking together round the Courts, of Lincoln's Inn, discussing no doubt many things but among them assuredly the Theory of Invariants. [Baker, 1912] p. XIII

S'ils ont certainement discuté de la théorie des invariants – Cayley venait de publier ses deux articles fondateurs sur cette théorie [Cayley, 1845] et [Cayley, 1846b] – il est aussi probable que Sylvester ait exposé à son ami ses résultats sur les fonctions de Sturm. Il n'est donc pas étonnant que Cayley, immergé dans son travail sur les déterminants, ait été amené à réécrire ces fonctions avec les outils sur lesquels il travaillait. C'est ainsi qu'en 1846 et 1848 il publie, en français et dans le *Journal de mathématiques pures et appliquées*, deux articles sur le théorème de Sturm.

#### 9.3.1.1 L'article de 1846

*Résumé*

Dans cet article, Cayley se propose d'exprimer les fonctions proposées par Sylvester avec les coefficients de la fonction initiale $f$. Pour cela, il utilise de façon judicieuse des déterminants dont les éléments sont des sommes de puissances des racines de $f$, que l'on sait exprimer avec les coefficient de $f$.

L'article de 1846, intitulé *Note sur les fonctions de Sturm*, est assez court et ne cite à aucun moment Sylvester, dont les fonctions auxiliaires sont simplement introduites par un « On sait que » [23]. Cayley commence donc par poser :

$$f(x) = (x - h_1)(x - h_2)...(x - h_n)$$

23. [Cayley, 1846a] p. 297.

$$
\begin{aligned}
f_1(x) &= \sum_i \prod_{j\neq i}(x-h_j)\\
f_2(x) &= \sum_{i<j}(h_i-h_j)^2 \prod_{k\neq i,j}(x-h_k)\\
f_3(x) &= \sum_{i_1<i_2<i_3} \prod_{1\leq r<s\leq 3}(h_{i_r}-h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3}(x-h_k)\\
f_4(x) &= \sum_{i_1<i_2<i_3<i_4} \prod_{1\leq r<s\leq 4}(h_{i_r}-h_{i_s})^2 \prod_{k\neq i_1,i_2,i_3,i_4}(x-f_k)\\
&... = ...\\
f_l(x) &= \sum_{i_1<i_2<...<i_l} \prod_{1\leq r<s\leq l}(h_{i_r}-h_{i_s})^2 \prod_{k\neq i_1,i_2,...,i_l}(x-h_k)\\
&... = ...\\
f_n(x) &= \prod_{1\leq r<s\leq n}(h_r-h_s)^2
\end{aligned}
$$

Le but que Cayley se propose d'atteindre est d'exprimer ces fonctions avec les coefficients de $f$. Pour cela, on pose pour $l \in [\![1;n]\!]$ et tout $1 \leq i_1 < ... < i_l \leq n$

$$P^l_{i_1,...,i_l} = (-1)^{\frac{l(l-1)}{2}} \prod_{1\leq r<s\leq l}(h_{i_r}-h_{i_s})$$

On obtient ainsi

$$
\begin{aligned}
\frac{f_l(x)}{f(x)} &= \frac{\sum_{1\leq i_1<i_2<...<i_l\leq n}\prod_{1\leq r<s\leq l}(h_{i_r}-h_{i_s})^2 \prod_{k\neq i_1,i_2,...,i_l}(x-h_k)}{(x-h_1)(x-h_2)...(x-h_n)}\\
&= \sum_{i_1<i_2<...<i_l} \frac{\left(P^l_{i_1,...,i_l}\right)^2}{\prod_{j=i_1,...,i_l}(x-h_j)}
\end{aligned}
$$

C'est ici que Cayley réécrit les $P^l_{i_1,...,i_l}$ qu'il vient d'introduire en termes de déterminant : on reconnait en effet un déterminant de Vandermonde :

$$P^l_{i_1,...,i_l} = \begin{vmatrix} 1 & h_{i_1} & h_{i_1}^2 & ... & h_{i_1}^{l-1}\\ 1 & h_{i_2} & h_{i_2}^2 & ... & h_{i_2}^{l-1}\\ 1 & ... & ... & ... & ...\\ 1 & h_{i_l} & h_{i_l}^2 & ... & h_{i_l}^{l-1} \end{vmatrix}$$

Ce qui est particulièrement remarquable dans cet article de Cayley, c'est l'utilisation systématique des déterminants. Comme nous allons le voir, la pratique du calcul de Cayley démontre une virtuosité certaine et il nous faut exposer moult détails pour réussir à le suivre.

On a :

$$\frac{P^l_{i_1,\dots,i_l}}{(x-h_{i_1})...(x-h_{i_l})} = \frac{(-1)^{\frac{l(l-1)}{2}}\prod_{1\le r<s\le l}(h_{i_r}-h_{i_s})}{(x-h_{i_1})...(x-h_{i_l})}$$

Ici, Cayley suppose implicitement que les $h_i$ sont tous différents, ce qui lui permet de procéder à la décomposition en éléments simples :

$$(1)\qquad \frac{P^l_{i_1,\dots,i_l}}{(x-h_{i_1})...(x-h_{i_l})} = (-1)^{\frac{l(l-1)}{2}}\left[\sum_{j=1}^{n}\frac{\prod_{1\le r<s\le l;r,s\neq j}(h_{i_r}-h_{i_s})}{(x-h_{i_j})}\right]$$

Ce que semble faire ensuite Cayley, c'est une sorte de développement en série entière, ou plutôt en série formelle, la convergence de la série n'étant pas considérée. Mais une série formelle d'un type particulier :

$$\sum_{n\ge 0} X^{-n}$$

c'est-à-dire en fait selon les puissances décroissantes de $x$ (ou croissantes de $\frac{1}{x}$).

En effet, il cherche dans l'égalité (1) le « coefficient de $x^{-r}$ ». Ce coefficient s'obtient certainement de la façon suivante. Considérons

$$\frac{1}{x-h_l} = \frac{1}{x}\frac{1}{1-\frac{h_l}{x}}$$

et utilisons, comme semble le faire Cayley, l'égalité suivante

$$\frac{1}{1-X} = 1+X+X^2+X^3+...$$

On obtient

$$\frac{1}{x-h_l} = \frac{1}{x}\left(1+\frac{h_l}{x}+\frac{h_l^2}{x^2}+...\right) = x^{-1}+h_l x^{-2}+...+h_l^{r-1}x^{-r}+...$$

En introduisant ce résultat dans (1), le coefficient en $x^{-r}$, sous réserve qu'il existe, est égal à :

$$(-1)^{\frac{l(l-1)}{2}}\left(\sum_{j=1}^{n}h_{i_j}^{r-1}\prod_{1\le r<s\le l;r,s\neq j}(h_{i_r}-h_{i_s})\right)$$

Ces coefficients se mettent aussi facilement sous la forme d'un déterminant :

$$\begin{vmatrix} 1 & h_{i_1} & h_{i_1}^2 & ... & h_{i_1}^{l-2} & h_{i_1}^{r-1} \\ 1 & h_{i_2} & h_{i_2}^2 & ... & h_{i_2}^{l-2} & h_{i_2}^{r-1} \\ 1 & ... & ... & ... & ... & \\ 1 & h_{i_l} & h_{i_l}^2 & ... & h_{i_l}^{l-2} & h_{i_l}^{r-1} \end{vmatrix}$$

Pour obtenir $\frac{f_m}{f}$ à partir de (1), il faut multiplier par $P^l_{i_1,\dots,i_l}$ puis sommer sur $1 \le i_1 <$

$i_2 < ... < i_l \leq n$. Le coefficient de $x^{-r}$ dans $\frac{f_m}{f}$ est donc [24]

$$\sum_{1\leq i_1<i_2<...<i_l\leq n} \begin{vmatrix} 1 & 1 & ... & 1 \\ h_{i_1} & h_{i_2} & ... & h_{i_l} \\ ... & ... & ... & ... \\ h_{i_1}^{l-1} & h_{i_2}^{l-1} & ... & h_{i_l}^{l-1} \end{vmatrix} \begin{vmatrix} 1 & h_{i_1} & h_{i_1}^2 & ... & h_{i_1}^{l-2} & h_{i_1}^{r-1} \\ 1 & h_{i_2} & h_{i_2}^2 & ... & h_{i_2}^{l-2} & h_{i_2}^{r-1} \\ 1 & ... & ... & ... & ... & \\ 1 & h_{i_l} & h_{i_l}^2 & ... & h_{i_l}^{l-2} & h_{i_l}^{r-1} \end{vmatrix}$$

$$= \begin{vmatrix} S_0 & S_1 & S_2 & ... & S_{l-2} & S_{r-1} \\ S_1 & S_2 & S_3 & ... & S_{l-1} & S_r \\ ... & ... & ... & ... & ... & \\ S_{l-1} & S_l & S_{l+1} & ... & S_{2l-3} & S_{r+l-2} \end{vmatrix}$$

Où $S_q = h_1^q + h_2^q + ... + h_n^q$. Pour obtenir $\frac{f_l(x)}{f(x)}$, nous pouvons écrire (1) sous la forme d'une matrice :

$$\frac{P^l_{i_1,...,i_l}}{(x-h_{i_1})...(x-h_{i_l})} = \begin{vmatrix} 1 & h_{i_1} & h_{i_1}^2 & ... & h_{i_1}^{l-2} & \frac{1}{x-h_{i_1}} \\ 1 & h_{i_2} & h_{i_2}^2 & ... & h_{i_2}^{l-2} & \frac{1}{x-h_{i_2}} \\ 1 & ... & ... & ... & ... & \\ 1 & h_{i_l} & h_{i_l}^2 & ... & h_{i_l}^{l-2} & \frac{1}{x-h_{i_l}} \end{vmatrix}$$

puis multiplier par $P^l_{i_1,...,i_l} = \begin{vmatrix} 1 & 1 & ... & 1 \\ h_{i_1} & h_{i_2} & ... & h_{i_l} \\ ... & ... & ... & ... \\ h_{i_1}^{l-1} & h_{i_2}^{l-1} & ... & h_{i_l}^{l-1} \end{vmatrix}$ et sommer sur les $1 \leq i_1 < i_2 < ... < i_l \leq n$ :

$$\frac{f_l(x)}{f(x)} = \sum_{1\leq i_1<i_2<...<i_l\leq n} \begin{vmatrix} 1 & 1 & ... & 1 \\ h_{i_1} & h_{i_2} & ... & h_{i_l} \\ ... & ... & ... & ... \\ h_{i_1}^{l-1} & h_{i_2}^{l-1} & ... & h_{i_l}^{l-1} \end{vmatrix} \begin{vmatrix} 1 & h_{i_1} & h_{i_1}^2 & ... & h_{i_1}^{l-2} & \frac{1}{x-h_{i_1}} \\ 1 & h_{i_2} & h_{i_2}^2 & ... & h_{i_2}^{l-2} & \frac{1}{x-h_{i_2}} \\ 1 & ... & ... & ... & ... & \\ 1 & h_{i_l} & h_{i_l}^2 & ... & h_{i_l}^{l-2} & \frac{1}{x-h_{i_l}} \end{vmatrix}$$

donc

$$\text{(2)} \qquad \frac{f_l(x)}{f(x)} = \begin{vmatrix} S_0 & S_1 & S_2 & ... & S_{l-2} & T_0 \\ S_1 & S_2 & S_3 & ... & S_{l-1} & T_1 \\ ... & ... & ... & ... & ... & \\ S_{l-1} & S_l & S_{l+1} & ... & S_{2l-3} & T_{l-1} \end{vmatrix}$$

où $T_q = \sum_{i=1}^n \frac{h_i^q}{x-h_i}$.

24. On utilise la *multiplication de deux systèmes rectangulaires* : voir plus loin p. 288.

Posons $f(x) = x^n - p_1 x^{n-1} + ... + (-1)^n p_n$ et multiplions (2) par $f(x)$. On obtient :

$$\begin{vmatrix} S_0 & S_1 & S_2 & ... & S_{l-2} & T_0 \\ S_1 & S_2 & S_3 & ... & S_{l-1} & T_1 \\ ... & ... & ... & ... & ... & \\ S_{l-1} & S_l & S_{l+1} & ... & S_{2l-3} & T_{l-1} \end{vmatrix} f(x)$$

$$= \begin{vmatrix} S_0 & S_1 & S_2 & ... & S_{l-2} & \sum_{i=1}^n \frac{1}{x-h_i}(x^n - p_1 x^{n-1} + ... + (-1)^n p_n) \\ S_1 & S_2 & S_3 & ... & S_{l-1} & \sum_{i=1}^n \frac{h_i}{x-h_i}(x^n - p_1 x^{n-1} + ... + (-1)^n p_n) \\ ... & ... & ... & ... & ... & \\ S_{l-1} & S_l & S_{l+1} & ... & S_{2l-3} & \sum_{i=1}^n \frac{h_i^{l-1}}{x-h_i}(x^n - p_1 x^{n-1} + ... + (-1)^n p_n) \end{vmatrix}$$

$$= \begin{vmatrix} ... & \sum_{i=1}^n (x^{-1} + h_i x^{-2} + ... + h_i^{r-1} x^{-r} + ...)(x^n - p_1 x^{n-1} + ... + (-1)^n p_n) \\ ... & \sum_{i=1}^n h_i (x^{-1} + h_i x^{-2} + ... + h_i^{r-1} x^{-r} + ...)(x^n - p_1 x^{n-1} + ... + (-1)^n p_n) \\ ... & ... \\ ... & \sum_{i=1}^n h_i^{l-1} (x^{-1} + h_i x^{-2} + ... + h_i^{r-1} x^{-r} + ...)(x^n - p_1 x^{n-1} + ... + (-1)^n p_n) \end{vmatrix}$$

On obtient pour la première ligne de la dernière colonne :

$$\begin{aligned} &\sum_{i=1}^n [(x^{n-1} + h_i x^{n-2} + ... + h_i^{r-1} x^{n-r} + ...) \\ &-p_1(x^{n-2} + h_i x^{n-3} + ... + h_i^{r-1} x^{n-r-1} + ...) \\ &+p_2(x^{n-3} + h_i x^{n-4} + ... + h_i^{r-1} x^{n-r-2} + ...) \\ &\quad ..... \\ &+(-1)^n p_n(x^{-1} + h_i x^{-2} + ... + h_i^{r-1} x^{-r} + ...)] \end{aligned}$$

Le coefficient en $x^{n-l+r}$ est :

$$\sum_{i=1}^n h_i^{r+l-1} - p_1 h_i^{r+l-2} + p_2 h_i^{r+l-3} - ... + (-1)^r p_r h_i^{l-1}$$

c'est-à-dire :

$$Q_{l,r} = S_{r+l-1} - p_1 S_{r+l-2} + p_2 S_{r+l-3} - ... + (-1)^r p_r S_{l-1}$$

En conservant cette notation, on obtient ainsi :

$$f_l(x) = \sum_{r=0}^{n-l} \begin{vmatrix} S_0 & S_1 & S_2 & ... & S_{l-2} & Q_{l,r} \\ S_1 & S_2 & S_3 & ... & S_{l-1} & Q_{l+1,r} \\ ... & ... & ... & ... & ... & \\ S_{l-1} & S_l & S_{l+1} & ... & S_{2l-3} & Q_{2l-1,r} \end{vmatrix} x^{n-l-r}$$

où, comme $f_l(x)$ est un polynôme, $r$ ne peut s'étendre que de 0 à $n-l$. On obtient ainsi une expression des fonctions $f_m$ qui dépend des sommes de Newton (les $S_k$), ainsi que des coefficients de la fonction de départ $f$. Cayley termine son article en donnant une « simplification » qui permet de réduire le nombre de calculs qui conduisent à ces fonctions. On voit dans cet article une technicité au niveau des opérations sur les déterminants – technicité qu'il double d'un certain sens de l'ellipse, ce qui rend l'article difficile à lire – qui permettra à Cayley de produire ses célèbres travaux en algèbre linéaire. On sait bien sûr que les $S_k$ peuvent eux-mêmes s'exprimer en fonction des coefficients de la fonction $f(x)$ : Cayley a donc réussi, et c'était là son but premier, à déterminer une méthode opératoire permettant de donner les fonctions de Sturm en fonction des coefficients de la fonction de départ [25].

#### 9.3.1.2 L'article de 1848

*Résumé*

Cayley poursuit son travail en exhibant une suite de fonctions exprimées sous la forme de déterminants qui dépendent cette fois *directement* des coefficients de la fonction initiale. Il souhaite démontrer que cette suite a la propriété des suites de Sturm, mais il n'atteindra pas son but. C'est encore une fois un travail sur la *forme* des fonctions sturmiennes.

Mais pourtant Cayley n'est pas entièrement satisfait. À cela deux raisons : il souhaite tout d'abord obtenir une expression de ces fonctions « au moyen des coefficients mêmes ; ce qui convient mieux que d'exprimer ces fonctions (...) par les sommes des puissances » [26], ce qui permettrait d'échapper au calcul laborieux des $S_k$ lorsque l'on ne connaît pas les racines. La seconde raison, qui n'est pas d'ordre calculatoire, mais qui change la nature de son travail, il la donne ainsi [27] :

> D'ailleurs il ne m'est plus nécessaire de parler des expressions de M. Sylvester, ou même des divisions successives de M. Sturm ; mais ma méthode fait voir directement que les fonctions que je vais définir sont douées de la propriété fondamentale sur laquelle se repose la théorie de M. Sturm

Cette propriété « fondamentale » correspond à celle que nous avons caractérisée plus haut comme la troisième [28] : « en considérant trois fonctions successives, la première et la

25. On pourra, comme nous l'avons déjà mentionné, consulter une application de ces formules faite par Koralek dans [Koralek, 1852].
26. Voir [Cayley, 1848] p. 269.
27. *Ibid.*
28. Voir p. 196.

dernière fonctions sont de signe contraire pour toute valeur de la variable qui fait évanouir la fonction intermédiaire ».

Examinons maintenant le contenu de cet article. Cayley choisit deux polynômes de même degré $n$ :

$$\begin{cases} V = a_0x^n + a_1x^{n-1} + ... + a_n \\ V' = b_0x^n + b_1x^{n-1} + ... + b_n \end{cases}$$

et on construit la suite de déterminants suivante :

$$F_1 = -\begin{vmatrix} V & V' \\ a_0 & b_0 \end{vmatrix}$$

$$F_2 = \begin{vmatrix} xV & V & xV' & V' \\ a_0 & 0 & b_0 & 0 \\ a_1 & a_0 & b_1 & b_0 \\ a_2 & a_1 & b_2 & b_1 \end{vmatrix}$$

...

$$F_i = (-1)^i \begin{vmatrix} x^{i-1}V & x^{i-2}V & ... & V & x^{i-1}V' & x^{i-2}V' & ... & V' \\ a_0 & 0 & ... & 0 & b_0 & 0 & ... & 0 \\ a_1 & a_0 & ... & 0 & b_1 & b_0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... \\ a_{2i-2} & a_{2i-3} & ... & a_{i-1} & b_{2i-2} & b_{2i-3} & ... & b_{i-1} \end{vmatrix}$$

La construction de cette suite de déterminants est faite pour que la suite des polynômes qui en découle soit de degré $n-1, n-2, ...$, donc de degré décroissant. Cayley construit alors deux autres suites auxiliaires qui lui permettront de montrer, mais pas de démontrer, que la suite des fonctions $F_i$ a la propriété des fonctions de Sturm citée précédemment :

$$P_1 = -\begin{vmatrix} a_0 & b_0 \\ a_1 & b_1 \end{vmatrix}$$

$$P_2 = \begin{vmatrix} a_0 & 0 & b_0 & 0 \\ a_1 & a_0 & b_1 & b_0 \\ a_2 & a_1 & b_2 & b_1 \\ a_3 & a_2 & b_3 & b_2 \end{vmatrix}$$

.....

$$P_i = (-1)^i \begin{vmatrix} a_0 & 0 & ... & 0 & b_0 & 0 & ... & 0 \\ a_1 & a_0 & ... & 0 & b_1 & b_0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... \\ a_{2i-1} & a_{2i-2} & ... & a_i & b_{2i-1} & b_{2i-2} & ... & b_i \end{vmatrix}$$

et

$$P'_1 = -\begin{vmatrix} a_0 & b_0 \\ a_2 & b_2 \end{vmatrix}$$

$$P'_2 = \begin{vmatrix} a_0 & 0 & b_0 & 0 \\ a_1 & a_0 & b_1 & b_0 \\ a_2 & a_1 & b_2 & b_1 \\ a_4 & a_3 & b_4 & b_3 \end{vmatrix}$$

.....

$$P'_i = (-1)^i \begin{vmatrix} a_0 & 0 & ... & 0 & b_0 & 0 & ... & 0 \\ a_1 & a_0 & ... & 0 & b_1 & b_0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... \\ a_{2i-2} & a_{2i-3} & ... & a_{i-1} & b_{2i-2} & b_{2i-3} & ... & b_{i-1} \\ a_{2i} & a_{2i-1} & ... & a_{i+1} & b_{2i} & b_{2i-1} & ... & b_{i+1} \end{vmatrix}$$

Cayley affirme alors que l'on a la relation suivante :

$$P_{i-1}^2 F_{i+1} + (xP_{i-1}P_i + P_{i-1}P'_i + P'_{i-1}P_i)F_i + P_i^2 F_{i-1} = 0$$

Cette relation montre clairement que les fonctions $F_i$ ont bien la propriété souhaitée, mais Cayley ne réussit pas à le prouver, et reporte à plus tard sa démonstration. La suite de l'article consiste à appliquer ce qui précède au cas où $V'$ est la dérivée de $V$. Il obtient ainsi des expressions des fonctions de Sturm (il s'agit bien, à un coefficient près, des mêmes fonctions que celles utilisées par Sturm) à l'aide des coefficients de la fonction de départ et donne ces expressions pour les quatre premières d'entre elles. On a une fois de plus un travail sur la *forme* de ces fonctions : c'est de l'expression et du calcul de ces

fonctions dont il s'agit ici, et non d'une modification de la démonstration du théorème de Sturm. Ainsi après les premiers travaux de Sylvester et les remarques de Cayley, on voit un déplacement de l'objet de l'étude des travaux autour du théorème de Sturm vers l'élaboration des fonctions de Sturm.

Ces articles de Cayley, Kronecker les reprend dans son cours [29] :

> Les déterminants de la forme de $D_\mu(x)$ ont été utilisé pour la première fois par Cayley dans une publication du journal de Liouville. [Kronecker, 1885b] p. 270

Nous allons maintenant voir comment, à l'occasion de cette reprise, Kronecker en donne une version *généralisée*.

### 9.3.2 La résolution par Kronecker

#### 9.3.2.1 Un premier point de vue

Après avoir utilisé son travail sur les fractions continues pour résoudre le problème de Cauchy, Kronecker cherche maintenant à retrouver ses résultats en utilisant, à la façon de Cayley, des déterminants. Pour cela, il prend $u_0$ et $\xi_0$ deux *indéterminées*, $\frac{F(x)}{\Psi(x)}$ une fraction rationnelle à déterminer (avec $deg(F) = \mu$ et $deg(\Psi) = \nu$) et $\xi_1, ...., \xi_n, u_1, ..., u_n$ définis par :

$$\frac{F(\xi_h)}{\Psi(\xi_h)} = u_h$$

pour tout $h \in [\![1\ ;\ n]\!]$ où on a pris $n = \mu + \nu + 1$.

On retrouve donc les conditions du problème de Cauchy, avec en plus les deux indéterminées $\xi_0$ et $u_0$.

On pose alors

$$F(\xi_0) - u_0\Psi(\xi_0) = \begin{vmatrix} 1 & \xi_0 & \xi_0^2 & ... & \xi_0^\mu & u_0 & u_0\xi_0 & ... & u_0\xi_0^\nu \\ 1 & \xi_1 & \xi_1^2 & ... & \xi_1^\mu & u_1 & u_1\xi_1 & ... & u_1\xi_1^\nu \\ ... & ... & ... & ... & ... & ... & ... & ... & ... \\ 1 & \xi_n & \xi_n^2 & ... & \xi_n^\mu & u_n & u_n\xi_n & ... & u_n\xi_n^\nu \end{vmatrix}$$

où comme précédemment $\mu$ et $\nu$ sont les degrés de $F$ et $\Psi$, et donc $\mu + \nu = n - 1$.

Le développement de ce déterminant par la première ligne montre que $F$ et $\Psi$ sont bien des fonctions entières de la variable $\xi_0$ de degré au plus $\mu$ et $\nu$.

Ce déterminant devient bien nul pour $\xi_0 = \xi_h$ et $u_0 = u_h$ où $h \in [\![1\ ;\ n]\!]$ car alors deux lignes du déterminant sont identiques. Dans le cas où l'ensemble des sous-déterminants d'ordre $n$ sont nuls, nous ne pouvons rien dire de plus : Kronecker qualifie alors la solu-

29. Determinanten von der Form wie die $D_\mu(x)$ besitzt hat zuerst Cayley in einem Aufsatze im Liouvilleschen Journal betrachtet.

tion d'*illusoire*[30]. Dans le cas contraire, nous obtenons donc une solution au problème de Cauchy, tout au moins dans sa forme modifiée : trouver $F$ et $\Psi$ tels que $F(x_h) = u_h\Psi(x_h)$ pour tout $h \in [\![1\ ;\ n]\!]$. Nous ne pourrons écrire

$$\frac{F(x_h)}{\Psi(x_h)} = u_h$$

pour tout $h \in [\![1\ ;\ n]\!]$, que dans les cas où $\Psi(x_h) \neq 0$, et donc $F(x_h) \neq 0$.

Ainsi $\dfrac{F(x)}{\Psi(x)}$ est une solution du problème de Cauchy dans la plupart des cas. Mais Kronecker ne peut pas se contenter de ce manque de généralité, et il présente ensuite la façon de résoudre complètement, ou plutôt « en général »[31] le problème sous la forme d'un déterminant[32] :

> Je veux maintenant montrer comment on peut aussi donner une solution sous la forme d'un déterminant de sorte que tous les cas soient épuisés. [Kronecker, 1885b] p. 268

Cette méthode, que l'on va examiner maintenant, si Kronecker l'attribue à Jacobi, c'est bien aux travaux de Cayley qu'il se réfère, comme on le voit dans le passage suivant[33] :

> En ce qui concerne l'historique des développements donnés ci-dessus, il convient de noter que les déterminants $\Psi(x)$ et $\Phi(x)$ ont été traités en premier par Cayley dans le tome 11 du Journal de Liouville dans un essai aussi court que son contenu est riche. Cependant il se limite au cas « régulier » où $f_1(x)$ est de degré $n-1$ et les nombres $n_1, n_2, ...n_k, ...$ sont égaux à $1, 2, ...k, ....$ Dans ce cas, comme nous le verrons encore, les fonctions $\mathcal{F}(x)$ correspondent à un facteur constant près aux fonctions de Sturm, c'est-à-dire aux fonctions de reste $f_k(x)$. Le cas non régulier se trouve traité dans la littérature pour la première fois par Hattendorf dans sa monographie „über die Sturm'schen Funktionen" ; il y est montré que l'expression du théorème de Sturm subsiste de la façon dont nous l'avons formulé précédemment aussi pour le cas non régulier. Ce fait a été découvert longtemps avant Hattendorf par Kronecker à l'occasion de ses leçons, mais il a vu en même temps que Hattendorf comme Borchardt, qui s'est également beaucoup occupé de cet objet, opéraient partiellement avec des déterminants qui sont identiquement nuls – Nous reviendrons par la suite sur ces

30. [Kronecker, 1885b] p. 267 : « Jedoch wird das Resultat illusorisch, wenn die Determinante identisch verschwindet, d.h wenn alle Unterdeterminanten $n^{\text{ter}}$ Ordnung verschwinden. »

31. [Kronecker, 1887b] p. 198 : « allgemeinen Lösung ».

32. Ich will Ihnen jetzt zeigen, wie man auch in Determinantenform die Lösung so geben kann, dass alle Fälle erschöpft werden.

33. Was das Historische der vorstehend gegebenen Entwicklungen anlangt, so ist zu bemerken, daß die Determinanten $\Psi(x)$ und $\Phi(x)$ zuerst von Cayley im 11. Band des Liouville'schen Journals behandelt worden sind, in einem ebenso kurzen, wie inhaltsreichen Aufsatz. Allerdings beschränkte er sich auf den „regulären" Fall, wo $f_1(x)$ vom $n-1$ ten Grade ist, und die Zahlen $n_1, n_2, ...n_k, ...$ bez. mit $1, 2, ...k, ...$ identisch sind. In diesem Fall stimmen, wie wir noch sehen werden, die Functionen $\mathcal{F}(x)$ bis auf constante Factoren mit den Sturm'schen Functionen überein, d.h. mit den Restfunctionen $f_k(x)$. In der Litteratur findet sich der nicht reguläre Fall zuerst von Hattendorf in dessen Monographie „über die Sturm'schen Funktionen" behandelt ; es ist dort gezeigt, daß der Ausspruch des Sturm'schen Satzes in der früher von uns formulierten Art auch für den nicht regulären Fall bestehen bleibt. Dieselbe Thatsache hat Kronecker schon lange vor Hattendorf bei Gelegenheit seiner Vorlesungen gefunden, zugleich aber dabei gesehen, daß sowohl Hattendorf wie Borchardt, der sich ebenfalls viel mit diesen Gegenständen beschäftigt hat, teilweise dabei mit Determinanten operieren, welche identisch Null sind – auf diese Determinanten werden wir im Folgenden noch zurückkommen. Allerdings bleiben trotz dieses Umstandes die von beiden gegeben Entwicklungen richtig.

> déterminants. Cependant les développements qu'ils ont donnés restent malgré tout corrects. [Kronecker, 1887b] p. 202-203

Dans le manuscrit que l'on vient de citer, l'usage implicite est de souligner le nom des auteurs lorsque l'on présente leur travaux pour la première fois. C'est ici le cas pour Kronecker, qui, dans une certaine mesure, se cite lui-même. On entre ainsi dans le cœur de son travail de recherche : l'introduction du théorème de Sturm dans les *Vorlesungen* sera en effet le début de la présentation d'un travail plus personnel et plus original de Kronecker. Avant cela il s'agissait surtout pour lui d'introduire les notions qui lui seront utiles pour développer les théories sur lesquelles il a travaillé et qui ont la plupart du temps donné lieu à la rédaction de plusieurs articles[34]. Il cite donc bien ici Cayley, dont nous avons examiné le travail, mais aussi Karl Hattendorf (1834-1882). On trouve peu de traces de ce dernier, qui a fait ses études à Göttingen et a passé l'essentiel de sa carrière à la *Rheinisch-Westfälische Technische Hochschule Aachen*. Le livre[35] auquel se réfère Kronecker est publié en 1874, et correspond à la publication de la thèse que Hattendorf a soutenue en 1864. Ces résultats sont donc publiés par Hattendorf après que Kronecker les a introduits dans ses leçons sur la théorie des équations algébriques[36]. Cependant, Kronecker ne débutant ses cours à l'université de Berlin qu'en 1862, il a pu avoir connaissance de la thèse qu'Hattendorf soutient en 1864.

Cette première introduction des déterminants n'est pas satisfaisante, et Kronecker développe une méthode plus « générale » qu'il attribue à Jacobi[37].

#### 9.3.2.2 Méthode de Jacobi

*Résumé*

Utilisant les formules d'Euler, Kronecker exhibe un système d'équations auquel devra satisfaire la fonction $F$. Il montre ensuite qu'une certaine fonction $D_\mu$, définie par un déterminant qui ne dépend que de $f_1$ et $f'$, satisfait ces équations. Ces nouvelles fonctions seront particulièrement intéressantes car elles permettront, sous certaines conditions, de remplacer avantageusement les fonctions $f_k$ dans l'utilisation du théorème de Sturm.

Dans son article *Über die Darstellung einer Reihe gegebener Werthe durch eine ge-*

34. Cette présentation de son cours sur la théorie des équations algébriques est un peu trop générale pour être complètement exacte, mais correspond bien au moins à celui de 1880-1881 et dans une large mesure à celui dont il est plus particulièrement question ici, c'est-à-dire celui du Wintersemester 1886-1887. Sa présentation n'est cependant pas identique tous les ans : par exemple durant le semestre d'hiver 1890-1891, quelques mois avant sa mort, la première partie de son cours, consacré à la théorie des équations de degré deux et trois, sera l'occasion pour lui d'introduire sur des cas simples des notions qui touchent de près ses recherches les plus récentes, comme les équations abéliennes.

35. Voir [Hattendorff, 1874].

36. On trouve en effet dans son article de 1873 [Kronecker, 1873b] un traitement du cas général dont il parle, et très certainement ce cas a été aussi exposé dans ses cours avant la rédaction de cet article. Cependant, la seule source que l'on ait, c'est-à-dire son cours du *Wintersemester* 1872-1873 [Kronecker, 1873a] n'en fait pas mention.

37. Dans [Kronecker, 1881a] p. 127, il note « Methode von Jacobi : (Crelle Bd. 30) ».

*brochene rationale Function*[38] publié en 1846, Jacobi se propose en effet de résoudre le problème de Cauchy, et c'est bien ce travail que Kronecker expose dans ses leçons sous l'intitulé « méthode de Jacobi ».

La méthode

On commence comme précédemment, mais on considère le fait que l'on a alors

$$\frac{F(\xi_h)}{f_1(\xi_h)} = \Psi(\xi_h) \text{ (h=1,...n)}$$

On suppose donc que les $f_1(\xi)$ sont connus et on intervertit les $f_1(\xi)$ et $\Psi(\xi)$.

Il s'agit donc maintenant de trouver une fonction $F$ de degré $\mu < n$ telle que $\frac{F}{f_1}$ coïncide avec $\Psi$ pour $x = \xi_1, ..., \xi_n$. Supposons que la fonction $\Psi$ soit construite à l'aide des polynômes d'interpolation de Lagrange de sorte que :

$$\frac{F(\xi_h)}{f_1(\xi_h)} = \Psi(\xi_h) \text{ (h=1,...n)}$$

On obtient donc

$$\Psi(x) = b_0 + b_1 x + ... + b_{n-\mu-1}x^{n-\mu-1} + b_{n-\mu}x^{n-\mu} + ... + b_{n-1}x^{n-1}$$

Mais en fait, une des conditions du problème de Cauchy montre que la fonction $\Psi$ doit être de degré inférieur à $n - \mu - 1$, et ainsi

$$b_{n-\mu} = 0, ...b_{n-1} = 0$$

Le début est donc similaire à la méthode que lui a inspirée Baltzer.

*Les formules d'Euler*

Kronecker utilise alors ce qu'il nomme les *formules d'Euler* (*Euler'schen Formeln*) et que Brioschi appelle « les formules connues de la décomposition des fractions »[39]. Nous les avons déjà rencontrées dans notre première partie[40], mais rappelons ici ce qu'elles nous disent[41] :

Si les $\xi$ sont les racines (simples) d'un polynôme $f$ de degré $n$, alors

$$\sum_{\xi} \frac{\xi^k}{f'(\xi)} = \delta_{k,n-1}$$

38. Voir [Jacobi, 1846].
39. Voir [Brioschi, 1854] p. 75.
40. Voir p. 143.
41. On pourra par exemple regarder aussi [Liouville, 1841] p. 375, où l'auteur attribue à Jacobi l'extension de cette formule aux fonctions de plusieurs variables. On remarquera par ailleurs que dans le cadre de ses cours, ainsi que dans son article de 1873, Kronecker utilise fréquemment le symbole $\delta_{i,j}$ qui porte maintenant son nom.

On a donc aussi

$$\sum_{\xi} \frac{F(\xi)}{f'(\xi)} = 0$$

si $\deg(F) < n-1$ et

$$\sum_{\xi} \frac{F(\xi)}{f'(\xi)} = A$$

si $\deg(F) = n-1$ avec $A$ le coefficient de plus haut degré de $F$.

On utilise maintenant ces formules pour obtenir un système de $\mu$ équations à $\mu+1$ inconnues.

D'après ces formules d'Euler, on a ainsi

$$\begin{cases} \sum_{h=1}^{n} \frac{\xi_h^{n-1}}{f'(\xi_h)} = 1 \\ \sum_{h=1}^{n} \frac{\xi_h^{k}}{f'(\xi_h)} = 0 \quad \text{si } k < n-1 \end{cases}$$

Grâce à l'utilisation astucieuse de celles-ci, on a

$$b_{n-1} = \sum_{h=1}^{n} \frac{\Psi(\xi_h)}{f'(\xi_h)} \ , \ b_{n-2} = \sum_{h=1}^{n} \frac{\xi_h \Psi(\xi_h)}{f'(\xi_h)} \ ,...b_{n-\mu} = \sum_{h=1}^{n} \frac{\xi^{\mu-1}\Psi(\xi_h)}{f'(\xi_h)}$$

et comme $\Psi(\xi_h) = \dfrac{F(\xi_h)}{f_1(\xi_h)}$ on a

$$(E_1) \qquad \sum_{h=1}^{n} \frac{\xi_h^t F(\xi_h)}{f_1(\xi_h) f'(\xi_h)} = 0 \ ,(t = 0, 1, ...\mu - 1)$$

Calcul des coefficients de $F$

Soit

$$F(x) = C_0 + C_1 x + ... + C_\mu x^\mu$$

On peut donc écrire, d'après ce qui précède :

$$\sum_{\nu=1}^{n} \sum_{h=0}^{\mu} \frac{C_h \xi_\nu^{t+h}}{f_1(\xi_\nu) f'(\xi_\nu)} = 0 \ (t = 0, ...\mu - 1)$$

Posons maintenant :

$$s_{h+t} = \sum_{\nu=1}^{n} \frac{\xi_\nu^{t+h}}{f_1(\xi_\nu) f'(\xi_\nu)}$$

Ainsi :

$$\sum_{h=0}^{\mu} C_h s_{h+t} = 0 \ (t = 0, ...\mu - 1)$$

On obtient donc $\mu$ équations homogènes des $\mu + 1$ coefficients de la fonction cherchée $F$ :

$$(S) = \begin{cases} s_0C_0 + s_1C_1 + ... + s_\mu C_\mu = 0 \\ s_1C_0 + s_2C_1 + ... + s_{\mu+1}C_\mu = 0 \\ ... = 0 \\ s_{\mu-1}C_0 + s_\mu C_1 + ... + s_{2\mu-1}C_\mu = 0 \end{cases}$$

Kronecker introduit alors la fonction $D_\mu(x)$ :

$$D_\mu(x) = \begin{vmatrix} s_0 & s_1 & ... & s_\mu \\ s_1 & s_2 & ... & s_{\mu+1} \\ ... & ... & ... & ... \\ s_{\mu-1} & s_\mu & ... & s_{2\mu-1} \\ 1 & x & ... & x^\mu \end{vmatrix}$$

Si le système $(S)$ admet une solution non triviale, il en admettra une infinité. L'écriture du déterminant $D_\mu$ est une façon astucieuse de noter une fonction dont les sous-déterminants seront les coefficients de la fonction. C'est une méthode que Kronecker utilisera de nouveau lorsqu'il exposera sa théorie des caractéristiques.

Cette construction apparaît dans son article de 1869 *Sur le théorème de Sturm*[42] et son article de 1878 *Über Sturm'sche Functionen*[43], mais dans une forme – comme toujours dans ses articles publiés – bien plus dense que ce que l'on peut trouver dans son cours. Ce qui est remarquable, c'est que ces deux articles précèdent ou succèdent à chaque fois de quelques mois la publication de l'un des articles fondamentaux sur la théorie des caractéristiques ([Kronecker, 1869c] et [Kronecker, 1878b]) : celle-ci est imbriquée dans ses réflexions sur le travail de Sturm.

Kronecker, dans le manuscrit de 1880-1881, pose directement le résultat en se référant à une partie précédente du cours[44]. Dans celui de 1886-1887 en revanche, il propose une démonstration, qui d'ailleurs diffère selon l'exemplaire. La version de Göttingen du manuscrit du Wintersemester 1886-1887 ([Kronecker, 1887b] p. 200) utilise des équations pour aboutir à une sorte de construction de la fonction. Mais elle pose problème dans son utilisation des coefficients à déterminer : la version de Strasbourg[45] utilise une méthode moins directe, mais qui dans une certaine mesure découle plus facilement de ce que Kronecker vient de mettre en place. C'est celle que l'on reproduit ici. Cette différence entre les deux versions, pourtant globalement très proches, nous apporte de précieux

42. Voir [Kronecker, 1869b].
43. Voir [Kronecker, 1878c].
44. Voir [Kronecker, 1881a] p. 128.
45. [Kronecker, 1887a] p. 176.

renseignements sur la constitution de ces manuscrits. Leur proximité, même dans les formulations, laisse penser à une origine commune, c'est-à-dire à deux mises en forme différentes d'une même prise de note. Cependant, des différences telles que celle qui vient d'être citée montrent que dans ce cas ces notes restent incomplètes, et il faut compléter, au moment de la mise en forme, certaines justifications qui n'y figuraient pas. Il est aussi possible que les deux versions soient issues de deux prises de notes différentes, et que Kronecker, pour les parties qui lui semblaient triviales ait laissé aux étudiants, comme « exercices à la maison », certaines démonstrations.

Montrons maintenant, comme Kronecker le fait, que $D_\mu(x)$ vérifie les équations $(E_1)$.

$$\sum_{h=1}^{n} \frac{\xi_h^t D_\mu(\xi_h)}{f_1(\xi_h) f'(\xi_h)}$$

$$= \sum_{h=1}^{n} \begin{vmatrix} s_0 & s_1 & \dots & s_\mu \\ s_1 & s_2 & \dots & s_{\mu+1} \\ \dots & \dots & \dots & \dots \\ s_{\mu-1} & s_\mu & \dots & s_{2\mu-1} \\ 1 & \xi_h & \dots & \xi_h^\mu \end{vmatrix} \frac{\xi_h^t}{f_1(\xi_h) f'(\xi_h)}$$

$$= \begin{vmatrix} s_0 & s_1 & \dots & s_\mu \\ s_1 & s_2 & \dots & s_{\mu+1} \\ \dots & \dots & \dots & \dots \\ s_{\mu-1} & s_\mu & \dots & s_{2\mu-1} \\ \sum_{h=1}^{n} \frac{\xi_h^t}{f_1(\xi_h) f'(\xi_h)} & \sum_{h=1}^{n} \frac{\xi_h^{t+1}}{f_1(\xi_h) f'(\xi_h)} & \dots & \sum_{h=1}^{n} \frac{\xi_h^{t+\mu}}{f_1(\xi_h) f'(\xi_h)} \end{vmatrix}$$

$$= \begin{vmatrix} s_0 & s_1 & \dots & s_\mu \\ s_1 & s_2 & \dots & s_{\mu+1} \\ \dots & \dots & \dots & \dots \\ s_{\mu-1} & s_\mu & \dots & s_{2\mu-1} \\ s_t & s_{t+1} & \dots & s_{t+\mu} \end{vmatrix}$$

Pour tout $t \in [\![1 \ ; \ \nu - 1]\!]$ ce dernier déterminant est nul (car deux lignes sont alors identiques), et $D_\mu$ vérifie donc bien les $\mu$ équations précédentes.

Ainsi la fonction $D_\mu(x)$ convient comme fonction $F$.

On a donc

$$F(x) = D_\mu(x) = \sum_{h=0}^{\mu} C_h x^h = \begin{vmatrix} s_0 & s_1 & ... & s_\mu \\ s_1 & s_2 & ... & s_{\mu+1} \\ ... & ... & ... & ... \\ s_{\mu-1} & s_\mu & ... & s_{2\mu-1} \\ 1 & x & ... & x^\mu \end{vmatrix}$$

On peut encore simplifier ce déterminant :

$$\begin{vmatrix} ... & ... & ... & ... \\ s_p & s_{p+1} & ... & s_{p+\mu} \\ ... & ... & ... & ... \\ 1 & x & ... & x^\mu \end{vmatrix}$$

$$=(-1)^\mu \begin{vmatrix} ... & ... & ... & ... & ... \\ s_p & xs_p - s_{p+1} & xs_{p+1} - s_{p+2} & ... & xs_{p+\mu-1} - s_{p+\mu} \\ ... & ... & ... & ... & ... \\ 1 & 0 & 0 & ... & 0 \end{vmatrix}$$

$$= |xs_{p+q} - s_{p+q+1}| \ (p, q = 0, 1, ...\mu - 1)$$

Cette dernière égalité nous permet d'introduire une notation que Kronecker affectionne tout particulièrement pour écrire les déterminants et qu'il conservera dans l'ensemble de son œuvre.

On notera

$$D_\mu(x) = |xs_{p+q} - s_{p+q+1}| \ (p, q = 0, 1, ...\mu - 1)$$

Nous allons maintenant examiner plus précisément les fonctions $D_\mu$ selon la valeurs de $\mu$. Tout d'abord pour les cas $\mu = n_k$ et $\mu = n_{k-1} - 1$, puis pour $n_k < \mu < n_{k-1}$.

**Définition de $\sigma_k$ et $\tau_k$** Ces fonctions $D_\mu$ que Kronecker construit avec tant de contraintes techniques représentent en fait l'outil principal de la *généralisation* de son travail. Elles lui permettent de considérer des $n_k \neq n - k$, ce qui est, nous l'avons vu [46], une source de « résultats intéressants ».

On peut [47] écrire $D_\mu(x)$ sous la forme :

$$D_\mu(x) = f_k(x)\Theta(x)$$

où $k$ est déterminé comme précédemment [48] tel que $n_k \leq \mu < n_{k-1}$. Ce changement de notation est dû au fait que Kronecker veut insister sur le lien entre la fonction et le nombre

46. Voir p. 250.
47. Voir p. 259.
48. Voir p. 259.

$\mu$.

Pour les deux cas limites, c'est-à-dire pour $\mu = n_k$ ou $\mu = n_{k-1} - 1$ :

On a $deg(\Theta) = \lambda$, $deg(\varphi_k) = n_k$, $deg(\psi_k) = n - n_{k-1}$, $deg(F) = \mu$ et $deg(\Psi) \leq n - \mu - 1$ . Donc :

$$\begin{aligned} \mu &= \lambda + n_k \\ \lambda + n - n_{k-1} &\leq n - \mu - 1 \end{aligned}$$

ainsi

$$\begin{aligned} \lambda &= \mu - n_k \\ \lambda &\leq n_{k-1} - \mu - 1 \end{aligned}$$

Ces deux conditions devant être vérifiées, $\lambda = 0$ lorsque $\mu = n_k$ ou $\mu = n_{k-1} - 1$. C'est-à-dire que $\Theta(x)$ est en fait une constante. Si on pose $c_k$ le terme de plus haut degré de $f_k$, on définit $\sigma_k$ et $\tau_k$ par :

$$\frac{D_\mu(x)}{\sigma_k} = \frac{f_k(x)}{c_k} \text{ pour } \mu = n_k$$

$$\frac{D_\mu(x)}{\tau_k} = \frac{f_k(x)}{c_k} \text{ pour } \mu = n_{k-1} - 1$$

Où dans le cas $\mu = n_{k-1} - 1$, $\Theta(x)$ étant une constante, les coefficients des $x^\mu$, $x^{\mu-1}$, ..., $x^{n_k+1}$ sont nuls. Et donc $\sigma_k$ est le coefficient de plus haut degré de

$$D_\mu(x) = \begin{vmatrix} s_0 & s_1 & ... & s_\mu \\ s_1 & s_2 & ... & s_{\mu+1} \\ ... & ... & ... & ... \\ s_{\mu-1} & s_\mu & ... & s_{2\mu-1} \\ 1 & x & ... & x^\mu \end{vmatrix}$$

pour $\mu = n_k$, on a donc :

$$\sigma_k = \begin{vmatrix} s_0 & s_1 & ... & s_{n_k-1} \\ s_1 & s_2 & ... & s_{n_k} \\ ... & ... & ... & ... \\ s_{n_k-2} & s_{n_k-1} & ... & s_{2n_k-3} \\ s_{n_k-1} & s_{n_k} & ... & s_{2n_k-2} \end{vmatrix}$$

De même, mais cette fois dans le cas régulier, on a pour $\mu = n_{k-1} - 1$ :

$$\tau_k = \begin{vmatrix} s_0 & s_1 & \dots & s_{n_{k-1}-2} \\ s_1 & s_2 & \dots & s_{n_{k-1}-1} \\ \dots & \dots & \dots & \dots \\ s_{n_k-2} & s_{n_k-1} & \dots & s_{2n_{k-1}-3} \\ s_{n_k-1} & s_{n_k} & \dots & s_{2n_{k-1}-2} \end{vmatrix}$$

**Expression des coefficients $c_k$ en fonction des $\sigma_k$ et des $\tau_k$** Nous allons maintenant exprimer les coefficients $c_k$ en fonction des $\sigma_k$ et des $\tau_k$. On a vu [49] que :

$$D_\mu(\xi_h) = \begin{vmatrix} s_0 & s_1 & \dots & s_\mu \\ s_1 & s_2 & \dots & s_{\mu+1} \\ \dots & \dots & \dots & \dots \\ s_{\mu-1} & s_\mu & \dots & s_{2\mu-1} \\ 1 & \xi_h & \dots & \xi_h^\mu \end{vmatrix}$$

Donc

$$\frac{\xi_h^\mu . D_\mu(\xi_h)}{f_1(\xi_h) f'(\xi_h)} = \begin{vmatrix} s_0 & s_1 & \dots & s_\mu \\ s_1 & s_2 & \dots & s_{\mu+1} \\ \dots & \dots & \dots & \dots \\ s_{\mu-1} & s_\mu & \dots & s_{2\mu-1} \\ \dfrac{\xi_h^\mu}{f_1(\xi_h) f'(\xi_h)} & \dfrac{\xi_h^{\mu+1}}{f_1(\xi_h) f'(\xi_h)} & \dots & \dfrac{\xi_h^{2\mu}}{f_1(\xi_h) f'(\xi_h)} \end{vmatrix}$$

et ainsi

$$\sum_{h=1}^{n} \frac{\xi_h^\mu . D_\mu(\xi_h)}{f_1(\xi_h) f'(\xi_h)} = \begin{vmatrix} s_0 & s_1 & \dots & s_\mu \\ s_1 & s_2 & \dots & s_{\mu+1} \\ \dots & \dots & \dots & \dots \\ s_{\mu-1} & s_\mu & \dots & s_{2\mu-1} \\ \sum_{h=1}^n \dfrac{\xi_h^\mu}{f_1(\xi_h) f'(\xi_h)} & \sum_{h=1}^n \dfrac{\xi_h^{\mu+1}}{f_1(\xi_h) f'(\xi_h)} & \dots & \sum_{h=1}^n \dfrac{\xi_h^{2\mu}}{f_1(\xi_h) f'(\xi_h)} \end{vmatrix}$$

$$\sum_{h=1}^{n} \frac{\xi_h^\mu . D_\mu(\xi_h)}{f_1(\xi_h) f'(\xi_h)} = \begin{vmatrix} s_0 & s_1 & \dots & s_\mu \\ s_1 & s_2 & \dots & s_{\mu+1} \\ \dots & \dots & \dots & \dots \\ s_{\mu-1} & s_\mu & \dots & s_{2\mu-1} \\ s_\mu & s_{\mu+1} & \dots & s_{2\mu} \end{vmatrix}$$

Ce déterminant correspond au coefficient en $x^{\mu+1}$ dans $D_{\mu+1}$, et on pose

49. Voir p. 259.

$$\sigma_{k-1} = \sum_{h=1}^{n} \frac{\xi_h^{n_{k-1}-1}.D_{n_{k-1}-1}(\xi_h)}{f_1(\xi_h)f'(\xi_h)}$$

On exprime maintenant cette somme d'une autre façon :

comme

$$c_k D_{n_{k-1}-1}(x) = \tau_k f_k(x) \text{ et } \psi_k(\xi_h) = \frac{f_k(\xi_h)}{f_1(\xi_h)}$$

on a

$$\begin{aligned} D_{n_{k-1}-1}(x) &= \frac{\tau_k}{c_k} f_k(x) \\ \frac{D_{n_{k-1}-1}(x)}{f_1(\xi_h)} &= \frac{\tau_k}{c_k} \psi_k(x) \\ \sum_{h=1}^{n} \frac{\xi_h^{n_{k-1}-1}.D_{n_{k-1}-1}(\xi_h)}{f_1(\xi_h)f'(\xi_h)} &= \frac{\tau_k}{c_k} \sum_{h=1}^{n} \frac{\xi_h^{n_{k-1}-1}.\psi_k(\xi_h)}{f'(\xi_h)} \end{aligned}$$

La fonction $\psi_k$ est de degré $n - n_{k-1}$, et donc $\xi_h^{n_{k-1}-1}.\psi_k(\xi_h)$ est de degré $n-1$. Or d'après les formules d'Euler $\sum_{h=1}^{n} \frac{\xi_h^i}{f'(\xi_h)}$ est égal à 0 ou 1 selon que $i < n-1$ ou égal à $n-1$ :

$$\sum_{h=1}^{n} \frac{\xi_h^{n_{k-1}-1}.\psi_k(\xi_h)}{f'(\xi_h)}$$

est donc le coefficient de la plus haute puissance de $\psi_k$. Par ailleurs, à partir du travail préalable sur les fractions rationnelles, on obtient la formule :

$$f(x) = f_{k-1}(x)\psi_k(x) - \psi_{k-1}(x)f_k(x)$$

or $deg(f_{k-1}(x)\psi_k(x)) = n_{k-1} + n - n_{k-1} = n$ et $deg(\psi_{k-1}(x)f_k(x)) < n$ donc le terme de plus haut degré se trouve dans $f_{k-1}(x)\psi_k(x)$. Comme le coefficient du terme de plus haut degré de $f$ est 1, on obtient :

$$\begin{aligned} 1 &= c_{k-1} \sum_{h=1}^{n} \frac{\xi_h^{n_{k-1}-1}.\psi_k(\xi_h)}{f'(\xi_h)} \\ 1 &= c_{k-1} \frac{c_k}{\tau_k} \sigma_{k-1} \\ c_{k-1}c_k &= \frac{\tau_k}{\sigma_{k-1}} \end{aligned}$$

**Remarque** L'argument précédent nous permet d'affirmer que le terme de plus haut degré de $\psi_k$ est $\frac{1}{c_{k-1}}$

Comme nous savons que $c_0 = 1$, cette formule nous permet de trouver par récurrence

tous les $c_k$, c'est-à-dire tous les termes de plus haut degré des fonctions $f_k$. Kronecker obtient les formules suivantes [50] :

$$c_{2k} = \frac{\sigma_0.\sigma_2....\sigma_{2_k-2}.\tau_2.\tau_4....\tau_{2k}}{\sigma_1.\sigma_3....\sigma_{2_k-1}.\tau_1.\tau_3....\tau_{2k-1}}$$

et

$$c_{2k+1} = \frac{\sigma_1.\sigma_3....\sigma_{2_k-1}.\tau_1.\tau_3....\tau_{2k+1}}{\sigma_0.\sigma_2....\sigma_{2_k}.\tau_2.\tau_4....\tau_{2k}}$$

Nous venons donc d'établir un lien entre les fonctions $f_k$, c'est-à-dire celles que Kronecker appelle les *fonctions de reste* de Sturm, et les déterminants $D_\mu$ avec $\mu = n_k$ et $\mu = n_{k-1} - 1$.

Mais qu'en est-il des déterminants intermédiaires ? Soit $\mu = n_k + \lambda$ où l'on a supposé que $n_{k-1} > n_k + 2$ et $n_k < n_k + \lambda < n_{k-1} - 1$. On a

$$D_{n_k+\lambda} = \begin{vmatrix} s_0 & s_1 & ... & s_{n_k+\lambda} \\ s_1 & s_2 & ... & s_{n_k+\lambda+1} \\ ... & ... & ... & ... \\ s_{n_k+\lambda-1} & s_{n_k+\lambda} & ... & s_{2(n_k+\lambda)-1} \\ 1 & x & ... & x^{n_k+\lambda} \end{vmatrix}$$

D'après les formules d'Euler,

$$\sum_{h=1}^{n} \frac{\xi_h^t.\psi_k(\xi_h)}{f'(\xi_h)} = 0$$

si $t + n - n_{k-1} < n - 1$, donc si $t < n_{k-1} - 1$.

Or $\dfrac{f_k(\xi_h)}{f_1(\xi_h)} = \psi_k(\xi_h)$, donc

$$\sum_{h=1}^{n} \frac{\xi_h^t.f_k(\xi_h)}{f_1(\xi_h)f'(\xi_h)} = 0$$

si $t < n_{k-1} - 1$. On pose maintenant de nouveau

$$f_k(x) = c_k x^{n_k} + c'_k x^{n_k-1} + ... + c_k^{(n_k)}$$

et on obtient ainsi

$$\sum_{\xi} \frac{c_k\xi^{n_k+t} + c'_k\xi^{n_k+t-1} + ... + c_k^{(n_k)}}{f_1(\xi)f'(\xi)} = 0 \text{ pour } t = 0, 1, ...n_{k-1} - 2$$

c'est-à-dire

$$c_k s_{n_k+t} + c'_k s_{n_k+t-1} + c''_k s_{n_k+t-2} + ... + c_k^{(n_k)} s_t = 0 \text{ pour } t = 0, 1...n_{k-1} - 2$$

50. Voir par exemple [Kronecker, 1885b] p. 273-274.

On peut faire la somme de l'ensemble des égalités telles que $t \leq n_k + \lambda$. On obtient alors :

$$c_k^{(n_k)}(s_0 + s_1 + ... + s_{n_k+\lambda}) + c_k^{(n_k-1)}(s_1 + ... + s_{n_k+\lambda} + 1) + ... + c_k(s_{n_k} + ... + s_{2n_k+\lambda}) = 0$$

Comme $\lambda \geq 1$, on a $n_k \leq n_k + \lambda - 1$ et donc ce que montre l'égalité précédente, c'est que certaines lignes de $D_\mu(x)$ sont des combinaisons linéaires d'autres lignes de $D_\mu(x)$ : dans ce cas, on a donc $D_\mu(x) = 0$.

**Conclusion** On vient donc de voir que si $n_k < \mu < n_{k-1} - 1$ alors $D_\mu(x) = 0$, et qu'en revanche si $\mu = n_k$ ou $\mu = n_{k-1} - 1$ alors $D_\mu(x)$ n'est pas nulle et est égale à un facteur près à l'une des *fonctions de reste* de Sturm. On peut donc obtenir la suite de fonctions $D_\mu(x)$ sans utiliser les développements en fractions continues.

#### 9.3.2.3 Application au Théorème de Sturm

*Résumé*

Les fonctions $D_\mu$ qui viennent d'être construites peuvent être utilisées comme suite de Sturm. Kronecker étudie en particulier dans certains cas leur terme de plus haut degré, qu'il note $\sigma_k$, pour appliquer le théorème de Sturm dans le cas où l'on cherche le nombre de racines réels sur $\mathbb{R}$. Ces coefficients $\sigma_k$ sont exprimés sous forme de déterminants que l'on a déjà rencontrés dans les travaux de Cayley.

L'égalité fondamentale, qui représente pour Kronecker le théorème de Sturm est, rappelons-le, la suivante :

$$\sum_{x_1<\xi<x_2} sgn\left(f'(\xi)f_1(\xi)\right) = \frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_2)f_h(x_2)) - \frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_1)f_h(x_1))$$

On veut maintenant remplacer les $f_k$ par les $D_\mu$ dans cette égalité. On a alors :

1. Si $\mu = n_k$ alors comme $c_{k-1}c_k = \dfrac{\tau_k}{\sigma_{k-1}}$

$$\begin{aligned} D_{\mu-1}.D_\mu &= \frac{\tau_{k+1}}{c_{k+1}}\frac{\sigma_k}{c_k}f_k(x)f_{k+1}(x) \\ &= \sigma_k^2 f_k(x)f_{k+1}(x) \end{aligned}$$

   donc

$$sgn(D_{\mu-1}D_\mu) = sgn(f_k.f_{k+1})$$

2. Si $\mu$ n'est pas l'un des entiers : $n_1, n_2, ...n_\nu$, alors $D_{\mu-1}D_\mu = 0$, sauf lorsque $\mu =$

$n_k + 1 = n_{k-1} - 1$. Dans ce dernier cas :

$$D_{\mu-1}.D_\mu = D_{n_k}.D_{n_{k-1}-1} = \frac{\tau_k}{c_k}\frac{\sigma_k}{c_k}f_k(x)f_k(x)$$
$$= \frac{\sigma_k\tau_k}{c_k^2}f_k^2(x)$$

ainsi le signe de $D_{\mu-1}.D_\mu$ est constant, et donc

$$sgn(D_{\mu-1}(x_2).D_\mu(x_2)) - sgn(D_{\mu-1}(x_1).D_\mu(x_1)) = 0$$

On a finalement

$$\sum_{x_1<\xi<x_2} sgn\left(f'(\xi)f_1(\xi)\right) = \frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_2)f_h(x_2)) - \frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_1)f_h(x_1))$$
$$= \sum_{h=1}^{n} sgn(D_{h-1}(x_2)D_h(x_2)) - \frac{1}{2}\sum_{h=1}^{n} sgn(D_{h-1}(x_1)D_h(x_1))$$

Kronecker cite ici le travail d'Hattendorf sur les fonctions de Sturm [51], dans lequel ce dernier traite des mêmes $D_\mu$. Mais, affirme Kronecker, Hattendorf n'a pas remarqué la nullité de certains de ces déterminants [52].

**Conclusion** Cette dernière formule remplace *avantageusement* la précédente [53], car si les $f_k$ doivent être recalculés pour chaque cas, les $D_\mu$ bénéficient de formules plus *générales*, car [54]

> On ne peut établir une fois pour toute les fonctions $f_1, f_2, ...f_\nu$ et remplacer dans chaque cas particulier des valeurs particulières, alors que l'on peut le faire avec les fonctions $D_0, D_1, ...D_n$. C'est pourquoi la formule que l'on vient de prouver est plus avantageuse que la précédente, dans laquelle on rencontre des produits des fonctions $f$. [Kronecker, 1885b] p. 277

**Un cas particulier** Si l'on suppose que $f_1$ et $f$ ont le même signe pour tous les $\xi$, par exemple en prenant $f_1(\xi) = f'(\xi)$ comme le fait Sturm, ou mieux encore, comme le fait Kronecker, en prenant $f_1(\xi) = \dfrac{1}{f'(\xi)}$, on peut construire $f_1$ avec la formule d'interpolation de Lagrange, et on obtient :

$$f_1(x) = \sum_{(\xi)} \frac{f(x)}{x-\xi}\frac{1}{f'(\xi)^2}$$

51. Voir [Hattendorff, 1874].
52. Voir [Kronecker, 1885b] p. 277.
53. Voir [Kronecker, 1885b] p. 277.
54. Die Functionen $f_1, f_2, ...f_\nu$ kann man nicht ein für allemal herstellen und in jedem speciellen Falle die speciellen Werte einsetzen, wohl aber kann man es mit den Functionen $D_0, D_1, ...D_n$ machen. Infolgedessen ist die jetzt abgeleitete Formel vorteilhafter als die frühere, in welcher die Producte der $f$ auftreten.

Par ailleurs, on a

$$\sum_{(\xi)} \frac{\xi^t}{f_1(\xi)f'(\xi)} = s_t = \sum_{(\xi)} \xi^t$$

et le nombre $\sum_{x_1<\xi<x_2} sgn\,(f'(\xi)f_1(\xi))$ représente le nombre de racines réelles comprises entre $x_1$ et $x_2$.

Dans ce cas, on peut donc exprimer le nombre de racines réelles de l'équation $f(x) = 0$ entre $x_1$ et $x_2$ à l'aide des $s_t$, qui sont ici des sommes de puissances des racines, sommes que Kronecker qualifie [55] d'« habituelles » (*gewöhnlich*), et qui sont donc connues.

On a vu plus haut [56] que

$$\sum_{-\infty<\xi<+\infty} sgn\,(f'(\xi)f_1(\xi)) = -\frac{1}{2}\sum_{h=1}^{\nu}(-1)^{n_{h-1}+n_h} sgn(c_{h-1}c_h) + \frac{1}{2}\sum_{h=1}^{\nu} sgn(c_{h-1}c_h)$$

et

$$\sum_{\substack{f(\xi)=0 \\ -\infty<\xi<+\infty}} sgn\,(f'(\xi)f_1(\xi)) = \sum_{n_{h-1}-n_h\equiv 1[2]} sgn(c_{h-1}c_h)$$

comme de plus $c_{h-1}c_h = \dfrac{\tau_h}{\sigma_{h-1}}$, on a

$$sgn(c_{h-1}c_h) = sgn(\tau_h\sigma_{h-1})$$

et donc

$$\sum_{\substack{f(\xi)=0 \\ -\infty<\xi<+\infty}} sgn\,(f'(\xi)f_1(\xi)) = \sum_{n_{h-1}-n_h\equiv 1[2]} sgn(\tau_h\sigma_{h-1})$$

Si l'on considère un cas particulier de ce cas particulier, en supposant que l'on se trouve dans le cas régulier où $n_k = n - k$, alors $\tau_k = \sigma_k$, alors on a toujours $n_{h-1} - n_h = 1 \equiv 1[2]$ et $\nu = n$. On a donc le résultat suivant :

**Proposition** Le nombre de racines réelles de l'équation $f(x) = 0$ est $\sum_{h=1}^{n} sgn(\sigma_h\sigma_{h-1})$ avec

$$\sigma_k = \begin{vmatrix} s_0 & s_1 & \dots & s_{n-k-1} \\ s_1 & s_2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ s_{n-k-1} & s_{n-k} & \dots & s_{2n-2k-2} \end{vmatrix} = |s_{p+q}| \ (p, q = 0, 1...n-k-1)$$

Kronecker ne le précise pas, mais $\sigma_k$ ne peut pas être défini de cette façon pour $k = n$. Dans ce dernier cas, en revenant à la définition de $\sigma_k$ et en utilisant les formules précédentes, on obtient $\sigma_n = s_0$. Or $\sigma_{n-1} = s_0$, donc $sgn(\sigma_n\sigma_{n-1}) = sgn(s_0^2) = 1$. C'est ainsi qu'il trouve

55. [Kronecker, 1885b] p. 278.
56. Voir p. 220.

la formule [57] :

$$\sum_{h=1}^{n} sgn(\sigma_h \sigma_{h-1}) = 1 + s - m$$

où $s$ est le nombre de « successions » (*Folgen*) et $m$ celui de « changements » (*Wechsel*) dans la suite $\sigma_0, ..., \sigma_{n-1}$.

Si on note maintenant $\mathcal{A}$ le nombre de racines réelles et $2\mathcal{B}$ celui des racines imaginaires, alors :

$$\mathcal{A} = 1 + s - m$$
$$2\mathcal{B} = n - \mathcal{A} = n - 1 - s + m.$$

Or $s + m = n - 1$, donc

$$\mathcal{B} = m$$

C'est-à-dire que [58]

> le nombre de changements de signes dans la suite indiquée est égal au nombre de paires de racines imaginaires. Cette proposition est valable, comme on le remarque, seulement si le développement de $f$ et $f_1$ est régulier. [Kronecker, 1881a] p. 133

Ces déterminants $D_\mu$ sont, comme nous venons de le voir et comme nous le verrons de nouveau, très importants dans l'exposition par Kronecker de la théorie de Sturm. L'origine de ces déterminants se trouve dans les travaux de Jacobi, et non pas dans ceux de Cayley [59] :

> Ces fonctions $D_\mu$ ont été étudiées en premier par *Jacobi* (*Cr. Bd. 30). Cayley* a montré en premier leur utilisation pour le théorème de *Sturm*, c'est pourquoi on lui attribue fréquemment faussement la découverte de ces fonctions. [Kronecker, 1881a] p. 132

La caractéristique de cette méthode réside bien dans les fonctions $D_\mu$, dont nous allons maintenant montrer le lien avec le théorème de Sturm.

#### 9.3.2.4 Les expressions de Sylvester

*Résumé*

Grâce à un travail préliminaire sur les déterminants, Kronecker obtient une formule extrêmement générale qui lui permet de retrouver les formules données par Sylvester en 1839 en reliant les déterminants $D_\mu$ aux racines de la fonction $f$.

Kronecker s'attaque maintenant aux *expressions de Sylvester*, c'est-à-dire aux formules que ce dernier donne dans son article de 1839. Mais comme dans tout ce qui précède, le

57. Voir [Kronecker, 1881a] p. 133 ou [Kronecker, 1879] p. 393.

58. Die Anzahl der Zeichenwechsel in der genannten Reihe ist gleich der Anzahl der Paare imaginärer Wurzeln. Dieser Satz gilt, wie bemerkt, nur wenn die Entwicklung von $f$ und $f_1$ regulär ist.

59. Diese Funktionen $D_\mu$ sind zuerst von *Jacobi* (*Cr. Bd. 30)* studirt worden. Ihre Anwendung auf den *Sturm*'schen Satzt zeigte zuerst *Cayley*, weßhalb man ihm auch häufig fälschlich die Auffindung dieser Funktionen zugeschrieben hat.

but que se fixe Kronecker n'est pas seulement de présenter ce travail déjà constitué, mais d'en donner une forme plus *générale*[60] :

> Je vais de plus, pour transformer nos déterminants dans les expressions de Sylvester, utiliser à la place des fonctions symétriques $s$ les racines de l'équation $f(x) = 0$ elles-mêmes. Sylvester ne les a mentionnées au départ que pour un cas particulier qui a ensuite été démontré par Sturm. Je vais lors de ma présentation plus générale travailler de façon constructive. [Kronecker, 1885b] p. 279

Sans sur-interpréter, ni faire d'anachronisme, on peut tout de même remarquer l'utilisation du terme « constructive » (*konstruktiv*) pour décrire le procédé qu'il emploie : lorsqu'elles sont possibles, ces méthodes sont en effet privilégiées[61].

**Préliminaire** Kronecker, avant de poursuivre, présente la multiplication de deux *Systèmes rectangulaires*, c'est-à-dire en fait la multiplication de deux matrices[62]

$$\left\|\begin{matrix} a_{00} & a_{01} & ... & a_{0n} \\ a_{10} & a_{11} & ... & a_{1n} \\ ... & ... & ... & ... \\ a_{m,0} & a_{m,1} & ... & a_{m,n} \end{matrix}\right\| \left\|\begin{matrix} b_{00} & b_{01} & ... & b_{0n} \\ b_{10} & b_{11} & ... & b_{1n} \\ ... & ... & ... & ... \\ b_{m,0} & b_{m,1} & ... & b_{m,n} \end{matrix}\right\|$$

$$= |c_{gh}|$$

où $m \leq n$ et $c_{gh} = \sum_{k=0}^{n} a_{gk} b_{hk}$.

Kronecker utilise la propriété suivante :

$$|c_{gh}|_{g,h=0,1,...,m} = \sum \left\|\begin{matrix} a_{0k_0} & a_{1k_1} & ... & a_{mk_m} \\ ... & ... & ... & ... \\ a_{m,k_0} & a_{mk_1} & ... & a_{mk_m} \end{matrix}\right\| \left\|\begin{matrix} b_{0k_0} & b_{1k_1} & ... & b_{mk_m} \\ ... & ... & ... & ... \\ b_{m,k_0} & b_{mk_1} & ... & b_{mk_m} \end{matrix}\right\|$$

où la somme est étendue à l'ensemble des $\frac{(n+1)!}{(n-m)!(m+1)!}$ combinaisons $(k_0, ..., k_m)$ des $n+1$ nombres $0, 1, ..., n$. Cette proposition, qui utilise en fait un développement selon les colonnes et les lignes, est soit rapidement justifiée [Kronecker, 1881a], [Kronecker, 1879], soit simplement énoncée [Kronecker, 1885b]. Dans [Kronecker, 1887b] il y a un renvoi à son cours sur le déterminant. Dans tous les cas, Kronecker fait allusion à la proposition de Laplace sur les déterminants, qui fixe les règles du développement selon les colonnes et les lignes.

60. Ich gehe dazu über, unsere Determinanten in die Sylvesterschen Ausdrücke zu verwandeln, welche an Stelle der symmetrischen Funktionen $s$ die Wurzeln der Gleichung $f(x) = 0$ selber benutzen. Sylvester hat sie ursprünglich nur für specielle Fälle angegeben, sie wurden dann von Sturm bewiesen. Ich gehe bei meiner allgemeineren Darstellung konstruktiv zu Werke.

61. On a vu par exemple les multiples façons dont Kronecker traite le théorème fondamental des polynômes symétriques, et la place qu'il donne à la méthode *constructive*. Voir p. 138.

62. Nous conserverons ici la notation que Kronecker utilise dans [Kronecker, 1881a] : une double barre permet de différencier le *système rectangulaire* du déterminant.

**Application** Ces propriétés sont appliquées à

$$a_{i,g} = u_i \xi_i^g \text{ et } b_{i,h} = v_i \eta_i^h$$

où

$$\begin{matrix} i = 0, & ... & n \\ g, h = 0, & ... & m \end{matrix}$$

et où $u_i, v_i, \xi_i, \eta_i$ sont des *variables*.

On obtient

$$|\sum_{k=0}^{n} u_k v_k \xi_k^g \eta_k^h|_{g,h=0,1,...,m} = \sum_{(i)} |u_i \xi_i^g|_{g=0,1,...,m} |v_i \eta_i^g|_{g=0,1,...,m}$$

où la somme est étendue à l'ensemble des $\frac{(n+1)!}{(n-m)!(m+1)!}$ combinaisons $(i_0, ..., i_m)$ des $n+1$ nombres $0, 1, ..., n$.

$|u_i \xi_i^g| \begin{pmatrix} i = i_0, & ... & i_m \\ g = 0, & ... & m \end{pmatrix}$ et $|v_i \eta_i^g| \begin{pmatrix} i = i_0, & ... & i_m \\ g = 0, & ... & m \end{pmatrix}$ sont des déterminants de Vandermonde, et on a donc :

$$|u_i \xi_i^g| \begin{pmatrix} i = i_0, & ... & i_m \\ g = 0, & ... & m \end{pmatrix} = \prod_{\alpha = i_0, ..., i_m} u_\alpha \prod_{\substack{\alpha, \beta = i_0, ..., i_m \\ \alpha < \beta}} (\xi_\alpha - \xi_\beta)$$

et

$$|v_i \eta_i^g| \begin{pmatrix} i = i_0, & ... & i_m \\ g = 0, & ... & m \end{pmatrix} = \prod_{\alpha = i_0, ..., i_m} v_\alpha \prod_{\substack{\alpha, \beta = i_0, ..., i_m \\ \alpha < \beta}} (\eta_\alpha - \eta_\beta)$$

donc on obtient

$$(E) \qquad \boxed{|\sum_{k=0}^{n} u_k v_k \xi_k^g \eta_k^h|_{g,h=0,1,...,m} = \sum_{(i)} \prod_{\alpha = i_0, ..., i_m} u_\alpha v_\alpha \prod_{\substack{\alpha, \beta = i_0, ..., i_m \\ \alpha < \beta}} (\xi_\alpha - \xi_\beta)(\eta_\alpha - \eta_\beta)}$$

À propos de cette dernière formule, Kronecker affirme qu'elle est [63]

> beaucoup plus générale que toutes les formules de Sylvester, Sturm, Cauchy, Jacobi.

De cette formule très générale, Kronecker tire maintenant un cas particulier : on pose tout d'abord $\eta_0 = 0$ et $\eta_k = \xi_k$ pour $k = 1, ..., n$. On a alors :

63. [Kronecker, 1885b] p. 280 : « Diese Formel ist viel allgemeiner als alle die Formeln von Sylvester, Sturm, Cauchy, Jacobi. »

$$
\begin{aligned}
&|\sum_{k=0}^{n} u_k v_k \xi_k^g \eta_k^h|_{g,h=0,1,...,m} \\
=&|u_0 v_0 \xi_0^g + \sum_{k=1}^{n} u_k v_k \xi_k^g, \sum_{k=1}^{n} u_k v_k \xi_k^{g+1}, ..., \sum_{k=1}^{n} u_k v_k \xi_k^{g+m}|_{g=0,1,...,m} \\
=&u_0 v_0 |\xi_0^g, \sum_{k=1}^{n} u_k v_k \xi_k^{g+1}, ..., \sum_{k=1}^{n} u_k v_k \xi_k^{g+m}|_{g=0,1,...,m} + |\sum_{k=1}^{n} u_k v_k \xi_k^g, \sum_{k=1}^{n} u_k v_k \xi_k^{g+1}, ..., \sum_{k=1}^{n} u_k v_k \xi_k^{g+m}|_{g=0,1,...,m}
\end{aligned}
$$

On poursuit en posant $v_k \xi_k = 1$ pour $k = 1, ..., n$ et $u_k = \dfrac{1}{f_1(\xi_k) f'(\xi_k)}$. On a alors

$$\sum_{k=1}^{n} u_k v_k \xi_k^{g+h} = \sum_{k=1}^{n} u_k \xi_k^{g+h-1} = s_{g+h-1}$$

et donc avec $u_0 = 1$

$$
\begin{aligned}
&u_0 v_0 |\xi_0^g, \sum_{k=1}^{n} u_k v_k \xi_k^{g+1}, ..., \sum_{k=1}^{n} u_k v_k \xi_k^{g+m}|_{g=0,1,...,m} + |\sum_{k=1}^{n} u_k v_k \xi_k^g, \sum_{k=1}^{n} u_k v_k \xi_k^{g+1}, ..., \sum_{k=1}^{n} u_k v_k \xi_k^{g+m}|_{g=0,1,...,m} \\
&= v_0 |\xi_0^g, s_g, ..., s_{g+m-1}|_{g=0,1,...,m} + |s_{g-1}, s_g, ..., s_{g+m-1}|_{g=0,1,...,m}
\end{aligned}
$$

et le coefficient de $v_0$ est ainsi, en posant de plus $\xi_0 = x$ d'après la définition de $D_\mu$ p. 278 :

$$(-1)^m |s_g, ..., s_{g+m-1}, x^g|_{g=0,1,...,m} = (-1)^m D_m(x)$$

Par ailleurs, sous les mêmes conditions, le coefficient de $v_0$ du membre de droite de l'équation $(E)$ est :

$$
\begin{aligned}
&\sum_{(i)} v_0 \prod_{\alpha=i_1,...,i_m} u_\alpha v_\alpha \prod_{\alpha=i_1,...,i_m} (x - \xi_\alpha)(-\xi_\alpha) \prod_{\substack{\alpha,\beta = i_1,...,i_m \\ \alpha<\beta}} (\xi_\alpha - \xi_\beta)^2 \\
&= v_0 (-1)^m \sum_{(i)} \prod_{\alpha=i_1,...,i_m} u_\alpha \prod_{\alpha=i_1,...,i_m} (x - \xi_\alpha) \prod_{\substack{\alpha,\beta = i_1,...,i_m \\ \alpha<\beta}} (\xi_\alpha - \xi_\beta)^2 \\
&= v_0 (-1)^m \sum_{(i)} \prod_{\alpha=i_1,...,i_m} \frac{1}{f_1(\xi_\alpha) f'(\xi_\alpha)} \prod_{\alpha=i_1,...,i_m} (x - \xi_\alpha) \prod_{\substack{\alpha,\beta = i_1,...,i_m \\ \alpha<\beta}} (\xi_\alpha - \xi_\beta)^2
\end{aligned}
$$

Donc finalement, en égalisant les coefficients de $v_0$, on obtient :

$$D_m(x) = \sum_{(i)} \prod_{\alpha=i_1,...,i_m} \frac{1}{f_1(\xi_\alpha) f'(\xi_\alpha)} \prod_{\alpha=i_1,...,i_m} (x - \xi_\alpha) \prod_{\substack{\alpha,\beta = i_1,...,i_m \\ \alpha<\beta}} (\xi_\alpha - \xi_\beta)^2$$

où la somme s'étend à toutes les combinaisons $(i_1, ..., i_m)$ des $n$ nombres $1, ..., n$.

**Remarque** Dans le travail précédent, on aurait pu ne pas remplacer immédiatement $u_\alpha$ par $\dfrac{1}{f_1(\xi_\alpha)f'(\xi_\alpha)}$ et on aurait obtenu :

$$D_m(x) = \sum_{(i)} \prod_{\alpha=i_1,...,i_m} u_\alpha \prod_{\alpha=i_1,...,i_m} (x-\xi_\alpha) \prod_{\substack{\alpha,\beta = i_1,...,i_m \\ \alpha<\beta}} (\xi_\alpha - \xi_\beta)^2$$

Ce travail a donc permis de relier la fonction $D_\mu$ qui est au coeur de son travail sur le problème de Cauchy avec les racines de $f = 0$. Pour obtenir maintenant les formules de Sylvester il lui suffit de particulariser encore un peu plus son travail. Il pose $f_1(\xi_\alpha) = \dfrac{1}{f'(\xi_\alpha)}$ et ainsi obtenir pour $m = 1, ..., n$ :

$$\begin{vmatrix} s_0 & s_1 \\ 1 & x \end{vmatrix} = \textstyle\sum_{\alpha=1}^{n}(x-\xi_\alpha).$$

$$\begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ 1 & x & x^2 \end{vmatrix} = \sum_{(\alpha,\beta=1,...n)} (x-\xi_\alpha)(x-\xi_\beta)(\xi_\alpha-\xi_\beta)^2 \quad \text{etc. etc.}$$

et ainsi de suite jusqu'à :

$$\begin{vmatrix} s_0 & s_1 & ... & s_n \\ s_1 & s_2 & ... & s_{n+1} \\ ... & ... & ... & ... \\ s_{n-1} & ... & ... & s_{2n-1} \\ 1 & x & ... & x^n \end{vmatrix} = \prod_{\alpha=1}^{n}(x-\xi_\alpha) \prod_{\alpha,\beta=1...n;\alpha<\beta} (\xi_\alpha - \xi_\beta)^2$$

qui correspondent bien aux formules que présente Sylvester dans son article de 1839.

À partir de cela, Kronecker produit un autre développement. On sait que

$$D_m(x) = |xs_{p+q} - s_{p+q+1}| \ (p, q = 0, 1, ...m-1)$$

Posons maintenant $x_k = \xi_k$ et $u_k = \mathcal{E}(x_k) = \dfrac{1}{f_1(\xi_\alpha)f'(\xi_\alpha)}$. On a alors :

$$\sum_{(i)} \prod_{i=i_1,...,i_m} \mathcal{E}(x_i)(x-x_i) \prod_{\substack{i,j = i_1,...,i_m \\ i<j}} (x_i - x_j)^2$$

$$= |xs_{p+q} - s_{p+q+1}| \ (p, q = 0, 1, ...m-1)$$

$$=|x\sum_{k=1}^{n}\mathcal{E}(x_k)x_k^{p+q}-\sum_{k=1}^{n}\mathcal{E}(x_k)x_k^{p+q+1}|\ (p,q=0,1,...m-1)$$
$$=|\sum_{k=1}^{n}(x-x_k)\mathcal{E}(x_k)x_k^{p+q}|\ (p,q=0,1,...m-1)$$

donc

$$\sum_{(i)}\prod_{i=i_1,...,i_m}\mathcal{E}(x_i)(x-x_i)\prod_{\substack{i,j=i_1,...,i_m\\ i<j}}(x_i-x_j)^2=|\sum_{k=1}^{n}(x-x_k)\mathcal{E}(x_k)x_k^{p+q}|\ (p,q=0,1,...m-1)$$

Si dans l'égalité précédente on égalise les termes de plus haut degré on obtient :

$$\sum_{(i)}\prod_{i=i_1,...,i_m}\mathcal{E}(x_i)\prod_{\substack{i,j=i_1,...,i_m\\ i<j}}(x_i-x_j)^2=|\sum_{k=1}^{n}\mathcal{E}(x_k)x_k^{p+q}|\ (p,q=0,1,...m-1)$$

Si l'on pose maintenant

$$\boxed{\mathcal{E}(x_k)=1}$$

on obtient :

$$\sum_{(i)}\prod_{\substack{i,j=i_1,...,i_m\\ i<j}}(x_i-x_j)^2=|s_{p+q}|\ (p,q=0,1,...m-1)$$

Voilà ce que dit Kronecker à propos de cette dernière égalité[64] :

> Nous avons à gauche une fonction symétrique des $x$, et celle-ci est représenté par notre déterminant dans lequel apparaissent des sommes de puissances habituelles.
> Sylvester a eu cette heureuse pensée d'exprimer les fonctions de Sturm, qui se trouvent ici à droite, par les racines de la fonction, les $x$ eux-mêmes. Les fonctions de Sturm ont été trouvées alors presque simultanément par Sturm avec les fractions continues, Cayley avec les déterminants et Sylvester à l'aide de la dernière formule. L'unification vient en fait de Liouville. Sturm a alors déduit cette formule directement par les fractions continues. Jacobi a publié dans le volume 30 du *Journal de Crelle* une dissertation dans laquelle, avec sa virtuosité, le calcul de toutes ces formules avec des

64. Wir haben links eine symmetrische Funktion der $x$, und diese ist in unserer Determinantenform dargestellt, in der dies die gewöhnlichen Potenzsummen sind.

Sylvester hatte den glücklichen Gedanken : die Sturmschen Funktionen welche hier rechts stehen, durch die Wurzeln der Gleichung, die $x$ selber, auszudrücken. Es fanden damals fast gleichzeitig Sturm durch sein Kettenbruchsverfahren, Cayley durch die Determinanten und Sylvester vermöge der letzten Formel die Sturmschen Funktionen ; die Vereinigung rührt eigentlich von Liouville her. Sturm hat dann diese Ausdrücke direkt durch das Kettenbruchsverfahren hergeleitet. Jacobi hat Bd. 30 des Crelleschen Journals einen Aufsatz veröffentlicht in dem er mit seiner Virtuosität der Rechnung alle diese Determinatenformeln entwickelte.

déterminants. [Kronecker, 1885b] p. 282

Kronecker reprend ainsi le plan de son travail, qui correspond à notre exposition des différentes méthodes, et dont l'aboutissement est en quelques sorte la formule précédente. La suite des matrices $|s_{p+q}|$ $(p, q = 0, 1, ...m-1)$ a un rôle particulièrement important : il s'agit bien de celle que l'on retrouve dans les travaux de Cayley, Sylvester, Borchard, Brioschi [65]. Nous allons voir comment Kronecker utilise ces derniers pour examiner la « taille » de l'ensemble des solutions du problème de Cauchy.

## 9.4 Dernière résolution du problème de Cauchy

*Résumé*

Le but est ici d'utiliser un développement de $\frac{f}{f_1}$ selon les puissances décroissantes de $x$ pour dépasser le problème des racines simples. Grâce à cela, nous obtenons de nouveaux des déterminants semblables à ceux utilisés par Cayley dans [Cayley, 1846a]. Pour une application au théorème de Sturm, se seront cette fois les $\Psi_k$ qui joueront le rôle des $f_k$.

Kronecker développe dans son cours un dernier moyen pour résoudre le problème de Cauchy, qui utilise en partie le travail que l'on vient de présenter, mais qui est « essentiellement différent » [66]. Quelle que soit la présentation qu'il en donne, ce qui caractérise cette méthode, c'est l'utilisation d'un développement selon les puissances décroissantes identique à celui dont se sert Cayley [67]. Nous allons suivre celle du manuscrit de 1886-87 qui a l'avantage de présenter cette méthode dans son ensemble, quitte à répéter et redémontrer certaines propriétés déjà acquises à ce niveau du cours de Kronecker. Le problème que permet de corriger cette méthode est, comme nous allons le voir, celui des racines multiples.

### 9.4.1 Redéfinition du problème

Kronecker commence par présenter un autre moyen de définir le problème de Cauchy. Soit $f$ une fonction entière de degré $n$. On veut déterminer $\mathcal{F}(x)$ et $\Psi(x)$ tels que

$$f_1(x)\Psi(x) \equiv \mathcal{F}(x)[mod.\ f(x)]$$

où $\mathcal{F}$ est de degré $n-\nu-1$ et $\Psi$ de degré $\nu$. Cette notation *modulo* $f$ n'est présente qu'à partir du moment où les *Grundzüge* ont été publiées.

Pour $\xi_1, ..., \xi_n$ racines de $f(x) = 0$, on a alors $f_1(\xi_h)\Psi(\xi_h) = \mathcal{F}(\xi_h)$, et on peut appliquer la formule d'interpolation de Lagrange :

65. Voir [Cayley, 1848], [Sylvester, 1853d],[Borchardt, 1847] et [Brioschi, 1854].
66. [Kronecker, 1881a] p. 136 : « gänzlich verschiedener »
67. Voir p. 266.

$$\mathcal{F}(x) = \sum_{h=1}^{n} \Psi(\xi_h) f_1(\xi_h) \frac{f(x)}{x - \xi_h} \frac{1}{f'(\xi_h)}$$

Or on a[68] :

$$\begin{aligned}
\frac{f(x)}{(x - x_h)} &= (x - x_1)...(x - x_{h-1})(x - x_{h+1})...(x - x_n) \\
&= x^{n-1} + (x_h - \lambda_1)x^{n-2} + (x_h^2 - \lambda_1 x_h + \lambda_2)x^{n-3} + ...(x_h^{n-1} - \lambda_1 x_h^{n-2} + ... \pm \lambda_{n-1}) \\
&= \sum_{i=0}^{n-1} (-1)^i \lambda_i \sum_{g=0}^{n-i-1} x_h^g \\
&= \sum_{g=0}^{n-1} x_h^g f^{(g)}(x) \\
&= F_1(x, x_h)
\end{aligned}$$

où $\lambda_i$ est le $i^{\text{ième}}$ polynôme symétrique élémentaire, $x_1, ...x_n$ les $n$ racines de $f$ et les $f^{(g)}(x)$ sont des polynômes de degré $n - g - 1$ linéairement indépendants. On peut donc écrire :

$$\begin{aligned}
\mathcal{F}(x) &= \sum_{h=1}^{n} \Psi(\xi_h) f_1(\xi_h) \frac{f(x)}{x - \xi_h} \frac{1}{f'(\xi_h)} \\
&= \sum_{h=1}^{n} \Psi(\xi_h) f_1(\xi_h) \sum_{g=0}^{n-1} \xi_h^g f^{(g)}(x) \frac{1}{f'(\xi_h)} \\
&= \sum_{g=0}^{n-1} f^{(g)}(x) \sum_{h=1}^{n} \frac{\Psi(\xi_h) f_1(\xi_h) \xi_h^g}{f'(\xi_h)}
\end{aligned}$$

Comme $\mathcal{F}$ est de degré $n - \nu - 1$ et que les $f^{(g)}$ sont linéairement indépendants, on doit avoir pour $t = 0, 1, 2, ..., \nu - 1$ :

$$\sum_{h=1}^{n} \frac{\Psi(\xi_h) f_1(\xi_h) \xi_h^t}{f'(\xi_h)} = 0$$

En effet, pour $t = 0, 1, 2, ..., \nu - 1$, $f^{(t)}$ est de degré $n - 1 - t > deg(\mathcal{F}) = n - \nu - 1$.

On obtient ainsi $\nu$ équations liant les coefficients de $\Psi$. Posons

$$c_p = \sum_{h=1}^{n} \frac{f_1(\xi_h) \xi_h^p}{f'(\xi_h)}$$

On voit que la principale différence avec ce que nous avons fait précédemment vient du fait que $f_1$ est au numérateur et pas au dénominateur. En effet, on rappelle que la définition des $s_p$ était :

68. Kronecker a démontré cette forme par exemple dans [Kronecker, 1887b] p. 59. Voir aussi la partie précédente p. 124.

$$s_p = \sum_{h=1}^{n} \frac{\xi_h^p}{f_1(\xi_h) f'(\xi_h)}$$

Pourquoi représenter ce travail sous cette forme ? La réponse apparaîtra clairement lorsque nous utiliserons le développement suivant les puissances décroissantes de $\frac{f_1}{f}$.

Montrons que

$$\Psi(x) = \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & 1 \\ c_1 & c_2 & \dots & c_\nu & x \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & x^\nu \end{vmatrix}$$

convient. Pour cela, il nous suffit de montrer, comme nous l'avons fait précédemment avec la fonction $|D_\mu(x)|$ qui jouait un rôle équivalent [69] que la fonction $\Psi$ vérifie les $\nu$ équations précédentes. Or on a

$$\begin{aligned} &\sum_{h=1}^{n} \frac{\Psi(\xi_h) f_1(\xi_h) \xi_h^t}{f'(\xi_h)} \\ =&\sum_{h=1}^{n} \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & 1 \\ c_1 & c_2 & \dots & c_\nu & \xi_h \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & \xi_h^\nu \end{vmatrix} \frac{f_1(\xi_h) \xi_h^t}{f'(\xi_h)} \end{aligned}$$

$$= \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & \sum_{h=1}^{n} \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^t \\ c_1 & c_2 & \dots & c_\nu & \sum_{h=1}^{n} \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^{t+1} \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & \sum_{h=1}^{n} \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^{t+\nu} \end{vmatrix}$$

$$= \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & c_t \\ c_1 & c_2 & \dots & c_\nu & c_{t+1} \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & c_{t+\nu} \end{vmatrix}$$

Pour tout $t \in [\![1 \ ; \ \nu - 1]\!]$ ce dernier déterminant est nul, et $\Psi$ vérifie donc bien les $\nu$ équations précédentes. Maintenant que l'on a trouvé un $\Psi$ convenable, il n'est pas difficile de déterminer $\mathcal{F}$. On a en effet :

69. Voir p. 278.

$$\mathcal{F}(x) = \sum_{h=1}^{n} \Psi(\xi_h) f_1(\xi_h) \frac{f(x)}{x-\xi_h} \frac{1}{f'(\xi_h)}$$
$$= \sum_{h=1}^{n} \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & 1 \\ c_1 & c_2 & ... & c_\nu & \xi_h \\ ... & ... & ... & ... & ... \\ c_\nu & c_{\nu+1} & ... & c_{2\nu-1} & \xi_h^\nu \end{vmatrix} f_1(\xi_h) \frac{f(x)}{x-\xi_h} \frac{1}{f'(\xi_h)}$$

$$= \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & \sum_{h=1}^{n} f_1(\xi_h) \frac{f(x)}{x-\xi_h} \frac{1}{f'(\xi_h)} \\ c_1 & c_2 & ... & c_\nu & \sum_{h=1}^{n} f_1(\xi_h) \frac{f(x)}{x-\xi_h} \frac{\xi_h}{f'(\xi_h)} \\ ... & ... & ... & ... & ... \\ c_\nu & c_{\nu+1} & ... & c_{2\nu-1} & \sum_{h=1}^{n} f_1(\xi_h) \frac{f(x)}{x-\xi_h} \frac{\xi_h^\nu}{f'(\xi_h)} \end{vmatrix}$$

$$= \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & \sum_{g=0}^{n-1} f^{(g)}(x) \sum_{h=1}^{n} f_1(\xi_h) \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^g \\ c_1 & c_2 & ... & c_\nu & \sum_{g=0}^{n-1} f^{(g)}(x) \sum_{h=1}^{n} f_1(\xi_h) \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^{g+1} \\ ... & ... & ... & ... & ... \\ c_\nu & c_{\nu+1} & ... & c_{2\nu-1} & \sum_{h=1}^{n} f_1(\xi_h) \frac{f(x)}{x-\xi_h} \frac{\xi_h^\nu}{f'(\xi_h)\xi_h^{g+\nu}} \end{vmatrix}$$

$$= \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & \sum_{g=0}^{n-1} c_g f^{(g)}(x) \\ c_1 & c_2 & ... & c_\nu & \sum_{g=0}^{n-1} c_{g+1} f^{(g)}(x) \\ ... & ... & ... & ... & ... \\ c_\nu & c_{\nu+1} & ... & c_{2\nu-1} & \sum_{g=0}^{n-1} c_{g+\nu} f^{(g)(x)} \end{vmatrix}$$

$$= \sum_{g=0}^{n-1} f^{(g)}(x) \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & c_g \\ c_1 & c_2 & ... & c_\nu & c_{g+1} \\ ... & ... & ... & ... & ... \\ c_\nu & c_{\nu+1} & ... & c_{2\nu-1} & c_{g+\nu} \end{vmatrix}$$

Or, pour $g = 0, 1, ..., \nu - 1$ ces déterminants sont nuls (deux colonnes identiques), donc :

$$\mathcal{F}(x) = \sum_{h=\nu}^{n} f^{(g)}(x) \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & c_g \\ c_1 & c_2 & ... & c_\nu & c_{g+1} \\ ... & ... & ... & ... & ... \\ c_\nu & c_{\nu+1} & ... & c_{2\nu-1} & c_{g+\nu} \end{vmatrix}$$

qui est bien une fonction entière de degré $n-\nu-1$.

### 9.4.2 Autre forme de $\mathcal{F}(x)$

Kronecker poursuit en donnant une autre forme à la fonction $\mathcal{F}(x)$. On a :

$$\frac{f(x)}{(x-\xi_h)} = \sum_{g=0}^{n-1} \xi_h^g f^{(g)}(x) \text{ pour } h = 1, ...n$$
$$f(x) = \sum_{g=0}^{n-1} (x\xi_h^g - \xi_h^{g+1}) f^{(g)}(x) \text{ pour } h = 1, ...n$$

ainsi

$$\begin{aligned} c_p f(x) &= \sum_{h=1}^{n} \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^p f(x) \\ &= \sum_{g=0}^{n-1} \sum_{h=1}^{n} (x \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^{g+p} - \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^{g+p+1}) f^{(g)}(x) \\ &= \sum_{g=0}^{n-1} (c_{p+g} x - c_{g+p+1}) f^(g)(x) \end{aligned}$$

On a par ailleurs

$$(-1)^\nu \mathcal{F}(x) = \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & \sum_{g=0}^{n-1} c_g f^{(g)}(x) \\ -c_1 & -c_2 & ... & -c_\nu & -\sum_{g=0}^{n-1} c_{g+1} f^{(g)}(x) \\ ... & ... & ... & ... & ... \\ -c_\nu & -c_{\nu+1} & ... & -c_{2\nu-1} & -\sum_{g=0}^{n-1} c_{g+\nu} f^{(g)(x)} \end{vmatrix}$$

et en ajoutant à chaque ligne la précédente multipliée par $x$ ou $-x$ et en utilisant l'égalité précédente [70] :

$$(-1)^\nu \mathcal{F}(x) = \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & f_1(x) \\ xc_0 - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu & c_0 f(x) \\ ... & ... & ... & ... & ... \\ xc_{\nu-1} - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} & c_{\nu-1} f(x) \end{vmatrix}$$

Donc

$$(-1)^\nu \mathcal{F}(x) = \begin{vmatrix} 0 & 0 & ... & 0 & f_1(x) \\ xc_0 - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu & c_0 f(x) \\ ... & ... & ... & ... & ... \\ xc_{\nu-1} - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} & c_{\nu-1} f(x) \end{vmatrix}$$

70. Le $f_1$ de la première ligne, dernière colonne vient de la formule d'interpolation de Lagrange.

$$+\begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & 0 \\ xc_0 - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu & c_0 f(x) \\ ... & ... & ... & ... & ... \\ xc_{\nu-1} - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} & c_{\nu-1} f(x) \end{vmatrix}$$

$$=(-1)^{\nu+2} f_1(x) \begin{vmatrix} xc_0 - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu \\ ... & ... & ... & ... \\ xc_{\nu-1} - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} \end{vmatrix}$$

$$+(-1)^{\nu} f(x) \begin{vmatrix} 0 & c_0 & c_1 & ... & c_{\nu-1} \\ c_0 & xc_0 - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu \\ ... & ... & ... & ... & ... \\ c_{\nu-1} & xc_{\nu-1} - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} \end{vmatrix}$$

Or on a

$$\Psi(x) = \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & 1 \\ c_1 & c_2 & ... & c_\nu & x \\ ... & ... & ... & ... & ... \\ c_\nu & c_{\nu+1} & ... & c_{2\nu-1} & x^\nu \end{vmatrix}$$

$$= \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & 1 \\ c_1 - xc_0 & c_2 - xc_1 & ... & c_\nu - xc_{\nu-1} & 0 \\ ... & ... & ... & ... & ... \\ c_\nu - xc_{\nu-1} & c_{\nu+1} - xc_\nu & ... & c_{2\nu-1} - xc_{2\nu-2} & 0 \end{vmatrix}$$

$$= (-1)^{\nu} \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & 1 \\ xc_0 - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu & 0 \\ ... & ... & ... & ... & ... \\ xc_{\nu-1} - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} & 0 \end{vmatrix}$$

$$= (-1)^{\nu}(-1)^{\nu} \begin{vmatrix} xc_0 - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu \\ ... & ... & ... & ... \\ xc_{\nu-1} - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} \end{vmatrix}$$

$$= \begin{vmatrix} xc_0 - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu \\ ... & ... & ... & ... \\ xc_{\nu-1} - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} \end{vmatrix}$$

Donc finalement

$$\mathcal{F}(x) = f_1(x)\Psi(x) - f(x)\Phi(x)$$

où

$$\Phi(x) = -\begin{vmatrix} 0 & c_0 & c_1 & ... & c_{\nu-1} \\ c_0 & xc_0x - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu \\ ... & ... & ... & ... & ... \\ c_{\nu-1} & xc_{\nu-1}x - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} \end{vmatrix}$$

**Racines simples** Kronecker précise ensuite que l'on a supposé implicitement que les racines étaient toutes distinctes lorsque l'on a défini les $c_p$ (en effet, on suppose alors que $f'(\xi) \neq 0$ si $\xi$ est une racine de $f$). Or nous pouvons, dit-il, les définir différemment, grâce à l'égalité :

$$\frac{f_1(x)}{f(x)} = \sum_{h=1}^{n} \frac{f_1(\xi_h)}{f'(\xi_h)} \frac{1}{x - \xi_h} = \sum_{h=1}^{n} \sum_{g=0}^{\infty} \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^g x^{-g-1} = \sum_{g=0}^{\infty} c_g x^{-g-1}$$

Cela permet donc à Kronecker de définir ces $c_p$ comme les coefficients d'une sorte de développement en fonction des puissances de $\frac{1}{x}$. On retrouve ici une technique utilisée par Cayley [71] dans son article de 1846 sur les fonctions de Sturm, article auquel Kronecker d'ailleurs se réfère [72].

### 9.4.3 Autre façon de mettre en évidence la relation qui unit $\mathcal{F}, \Psi$ et $\Phi$.

On peut maintenant remarquer que les trois fonctions $\mathcal{F}, \Psi$ et $\Phi$ dépendent de l'entier $\nu$, et on peut donc les nommer respectivement $\mathcal{F}_\nu, \Psi_\nu$ et $\Phi_\nu$. Kronecker travaille ensuite sur la forme de $\Phi_\nu$. On a

$$-\Phi(x) = \begin{vmatrix} 0 & c_0 & c_1 & ... & c_{\nu-1} \\ c_0 & xc_0 - c_1 & xc_1 - c_2 & ... & xc_{\nu-1} - c_\nu \\ ... & ... & ... & ... & ... \\ c_{\nu-1} & xc_{\nu-1} - c_\nu & xc_\nu - c_{\nu+1} & ... & xc_{2\nu-2} - c_{2\nu-1} \end{vmatrix}$$

En faisant $L_2 = L2 - xL_1$, puis $L_3 = L_3 + xL_2$, puis $L_4 = L_4 + xL_3$, etc. on obtient

71. Voir [Cayley, 1846a] p. 298.
72. Voir [Kronecker, 1887b] p. 202-203. Voir citation p. 273.

$$-\Phi(x) = \begin{vmatrix} 0 & c_0 & c_1 & ... & c_{\nu-1} \\ c_0 & -c_1 & -c_2 & ... & -c_\nu \\ ... & ... & ... & ... & ... \\ c_0x^{\nu-1} + c_1x^{\nu-2} + ... + c_{\nu-1} & -c_\nu & -c_{\nu+1} & ... & -c_{2\nu-1} \end{vmatrix}$$

donc

$$-\Phi(x) = (-1)^{2\nu+1} \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & 0 \\ c_1 & c_2 & ... & c_\nu & c_0 \\ ... & ... & ... & ... & ... \\ c_\nu & c_{\nu+1} & ... & c_{2\nu-1} & c_0x^{\nu-1} + c_1x^{\nu-2} + ... + c_{\nu-1} \end{vmatrix}$$

c'est-à-dire

$$\boxed{\Phi(x) = \begin{vmatrix} c_0 & c_1 & ... & c_{\nu-1} & 0 \\ c_1 & c_2 & ... & c_\nu & c_0 \\ c_2 & c_3 & ... & c_{\nu+1} & c_0x + c_1 \\ ... & ... & ... & ... & ... \\ c_\nu & c_{\nu+1} & ... & c_{2\nu-1} & c_0x^{\nu-1} + c_1x^{\nu-2} + ... + c_{\nu-1} \end{vmatrix}}$$

La dernière colonne de ce déterminant est composé de polynômes de la forme $c_0x^{h-1} + c_1x^{h-2} + ... + c_{h-1}$, que l'on peut voir aussi comme la partie entière du développement

$$\frac{x^h f_1(x)}{f(x)} = \sum_{g=0}^{\infty} c_g x^{h-g-1}$$

Kronecker introduit ensuite la notation suivante, qui, dit-il, est directement issue des travaux de Gauss [73] :

$$\left[\frac{x^h f_1(x)}{f(x)}\right] = c_0x^{h-1} + c_1x^{h-2} + ... + c_{h-1}$$

c'est-à-dire que $\left[\frac{x^h f_1(x)}{f(x)}\right]$ représente une sorte de « partie entière » de ce développement infini, et la partie du développement conservée est celle contenant les puissances de $x$ positives. On a ainsi

73. Il utilise déjà cette notation, mais dans son cadre originel, dans un article de 1884 *Beweis des Reciprocitätsgesetzes für die quadratischen Reste* ([Kronecker, 1884a] p. 500) : « à la manière de Gauss » [74], il désignera par $[a]$ l'entier tel que

$$a - 1 < [a] \leq a$$

$$\Phi_\nu(x) = \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & \left[\frac{f_1(x)}{f(x)}\right] \\ c_1 & c_2 & \dots & c_\nu & \left[\frac{xf_1(x)}{f(x)}\right] \\ c_2 & c_3 & \dots & c_{\nu+1} & \left[\frac{x^2f_1(x)}{f(x)}\right] \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & \left[\frac{x^\nu f_1(x)}{f(x)}\right] & \end{vmatrix}$$

Avec

$$\Psi_\nu(x) = \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & 1 \\ c_1 & c_2 & \dots & c_\nu & x \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & x^\nu \end{vmatrix}$$

et

$$\mathcal{F}_\nu(x) = \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & \sum_{g=0}^{n-1} c_g f^{(g)}(x) \\ c_1 & c_2 & \dots & c_\nu & \sum_{g=0}^{n-1} c_{g+1} f^{(g)}(x) \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & \sum_{g=0}^{n-1} f^{(g)(x)} \end{vmatrix}$$

Kronecker montre assez facilement, par une méthode directe (*direct veriefiriert* [Kronecker, 1887b] p. 205), l'égalité

$$f_1(x)\Psi_\nu(x) - f(x)\Phi_\nu(x) = \mathcal{F}(x)$$

### 9.4.4 Quelques considérations sur la dimension de l'ensemble des solutions

Maintenant, Kronecker discute la dimension de la variété (*Mannigfaltigkeit*) qui contient l'ensemble des solutions de notre problème. Ces solutions ont été trouvées à partir des $\nu$ équations : pour $t = 0, 1, 2, ...\nu - 1$

$$\sum_{h=1}^{n} \frac{\Psi_\nu(\xi_h) f_1(\xi_h) \xi_h^t}{f'(\xi_h)} = 0$$

et il faut regarder le rang de la matrice :

$$M = \begin{matrix} c_0 & c_1 & \dots & c_{\nu-1} \\ c_1 & c_2 & \dots & c_\nu \\ \dots & \dots & \dots & .... \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} \end{matrix}$$

Kronecker l'explique en passant par le déterminant des sous-matrices : la dimension de la variété sera $\nu - \text{rang}(M) + 1$. De plus, on peut remarquer que dans le cas où le rang de

la matrice est $\nu$, c'est-à-dire lorsque l'ensemble des solutions du problème de Cauchy est une variété d'ordre 1, et comme

$$\Psi_\nu(x) = \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & 1 \\ c_1 & c_2 & \dots & c_\nu & x \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & x^\nu \end{vmatrix}$$

alors $\Psi_\nu$ n'est pas identiquement nulle. La réciproque est évidemment juste. Il y a donc équivalence entre le fait que $\Psi_\nu \neq 0$ et le fait que l'ensemble des solutions du problème de Cauchy soit de dimension 1. Nous avons vu précédemment que :

1. Pour $\nu = n_k$, la fonction $\Psi_\nu$ est de la forme $C.\psi_\nu$ où $C$ est une constante : on a bien affaire à une variété de dimension 1.
2. En revanche, si $\nu = n_k + \tau$ où $\tau \in \mathbb{N}^*$ et $n_k < \nu < n_{k-1} - 1$, alors $\Psi_\nu$ est de la forme $C(x).\psi_\nu$ où $C$ est de degré $\tau$.
3. Enfin, si $\nu = n_{k-1} - 1$, on a comme dans le premier cas une variété simple (en fait, c'est le cas précédent avec $\tau = 0$).

L'utilisation que nous avons faite du terme de *rang* n'est pas simplement une transposition contemporaine : le terme apparaît dans le cours de Kronecker, même si celui de matrice n'est pas présent [75]. Le mathématicien tchèque Eduard Weyr (1852-1903), dans un article [76] de 1890 intitulé *Zur theorie der bilinearen formen*, attribue d'ailleurs à Kronecker cette notion de rang [77] :

> D'après Kronecker, une matrice A est dite de rang $r$ $(0 < r \leq n)$ lorsque $r$ est l'ordre maximal des mineurs non nuls parmi tous les mineurs de $|A|$ ; $A$ a le rang 0 lorsque tous les éléments de $A$ sont identiquement nuls.

Comme on le voit, cette définition [78] correspond bien à l'utilisation que fait ici Kronecker du terme « rang ».

**Utilisation d'un système « illimité »** On a

$$\Psi_\nu(x) = \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & 1 \\ c_1 & c_2 & \dots & c_\nu & x \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & x^\nu \end{vmatrix}$$

75. *„Range“* [Kronecker, 1887b] p. 206.
76. [Weyr, 1890] p. 172.
77. Nach Kronecker heißt $A$ vom Range $r$ $(0 < r \leq n)$, wenn es unter den nicht verschwindenden Minoren von $|A|$, welche höchstens vom Grade $r$ gibt ; $A$ hat ferner den Rang 0, wenn alle Elemente von $A$ gleich Null sind. La traduction proposée est de Frédéric Brechenmacher [Brechenmacher, 2006] p. 351
78. Ce terme, Kronecker l'a déjà utilisé dans deux articles de 1884 : *Die Periodensysteme von Functionen reeller Variabeln* [Kronecker, 1884b] et *Näherungsweise ganzzahlige Auflösung linearer Gleichungen* [Kronecker, 1884c].

donc le coefficient de $x^\nu$ dans $\psi_n u$ est proportionnel à

$$\begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} \\ c_1 & c_2 & \dots & c_\nu \\ \dots & \dots & \dots & \dots \\ c_{\nu-1} & c_\nu & \dots & c_{2\nu-2} \end{vmatrix}$$

Dans le cas où $\nu = n_k$, c'est-à-dire dans le cas où $\Psi_\nu$ est « vraiment » de degré $n_k$, alors ce déterminant, que l'on pourra appeler le « déterminant principal » (*die „Hauptdeterminante"*) n'est pas nul. Kronecker considère ensuite un système « illimité » (*„unbegrenzte System"* [Kronecker, 1887b] p. 207) :

$$\begin{pmatrix} c_0 & c_1 & c_2 & \dots \\ c_1 & c_2 & c_3 & \dots \\ c_2 & c_3 & c_4 & \dots \\ & & & \end{pmatrix}$$

où les $c_i$ sont définis, comme on l'a vu, par le développement

$$\frac{f_1(x)}{f(x)} = \sum_{h=1}^{n} \frac{f_1(\xi_h)}{f'(\xi_h)} \frac{1}{x - \xi_h} = \sum_{h=1}^{n} \sum_{g=0}^{\infty} \frac{f_1(\xi_h)}{f'(\xi_h)} \xi_h^g x^{-g-1} = \sum_{g=0}^{\infty} c_g x^{-g-1}$$

La notation pourrait faire penser ici à une matrice de dimension infinie, considérée en tant qu'objet, et donc en complet désaccord avec l'idée d'« infini » que Kronecker a souvent proclamée. Mais en fait il n'en est rien. Tout d'abord, il est bien précisé que la série est une *fortschreitende Reihe*, qui donc se forme pas à pas, et n'est pas considérée comme entièrement construite. Le but est en effet cette fois de *calculer* le nombre $n_k$ en calculant l'ordre (*die Ordnungszahlen*) de la suite de déterminants :

$$|c_0|, \begin{vmatrix} c_0 & c_1 \\ c_1 & c_2 \end{vmatrix}, \begin{vmatrix} c_0 & c_1 & c_2 \\ c_1 & c_2 & c_3 \\ c_2 & c_3 & c_4 \end{vmatrix}, \dots$$

On obtient ainsi une suite de nombres entiers $n_1, n_2, n_3, \dots$ qui bien sûr sont nuls à partir d'un certain rang, les « déterminants principaux » devenant eux-même nuls. Kronecker fait ici remarquer que la représentation des fonctions $\mathcal{F}_\nu$ et $\Psi_\nu$ à l'aide de déterminant, c'est-à-dire le fait que l'on ait :

$$\Psi_\nu(x) = \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & 1 \\ c_1 & c_2 & \dots & c_\nu & x \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & x^\nu \end{vmatrix}$$

et

$$\mathcal{F}_\nu(x) = \begin{vmatrix} c_0 & c_1 & \dots & c_{\nu-1} & \sum_{g=0}^{n-1} c_g f^{(g)}(x) \\ c_1 & c_2 & \dots & c_\nu & \sum_{g=0}^{n-1} c_{g+1} f^{(g)}(x) \\ \dots & \dots & \dots & \dots & \dots \\ c_\nu & c_{\nu+1} & \dots & c_{2\nu-1} & \sum_{g=0}^{n-1} f^{(g)(x)} \end{vmatrix}$$

montre que si $\Psi_\nu = 0$, il en est de même pour $\Phi_\nu$ et $\mathcal{F}_\nu$. Rechercher une solution au problème de Cauchy *à l'aide des déterminants* n'a donc de sens que lorsque $\nu = n_k$ ou $\nu = n_{k+1} - 1$.

La suite de déterminants précédente qui, comme l'affirme Kronecker, joue un rôle si important [79] ont reçu, de la part de certains mathématiciens, un nom particulier. Par exemple Hankel [80] utilise l'expression *orthosymetrische Determinanten* et Schering [81] les *einreihige Determinanten*. Ces déterminants sont complètement équivalents aux déterminants $|s_{p,q}|$, et ils jouent ici le même rôle.

79. [Kronecker, 1887b] p. 208 : « eine so wichtige Rolle spielen »

80. Hermann Hankel (1839-1873), qui a commencé ses études à Leipzig en 1857, les poursuit à Berlin à partir de 1861. Il travaille alors avec Weierstrass et Kronecker. Il soutient son doctorat intitulé *Über eine besondere Classe symmetrischer Determinanten* en 1862 sous la direction de Möbius.

81. Ernst Christian Julius Schering (1824-1897) a été l'éditeur des papiers de Gauss.

# Chapitre 10

# La méthode d'Hermite-Jacobi

10.1 Le principe d'inertie de Sylvester . . . . . **307**
10.2 Application de la loi d'inertie au théorème de Sturm . . . . . **310**

Dans son article [1] de 1873, Kronecker commence par se référer à un article de Borchardt [2] pour l'aspect *historique* de cette méthode, en précisant que sa propre exposition sera plus *générale*. L'article de Borchardt fait déjà en quelque sorte le point sur les méthodes existantes, mais manque, du point de vue de Kronecker, de *généralité*.

Cette méthode, basée sur les travaux de Hermite sur les formes quadratiques, mais aussi, comme le souligne Kronecker, sur ceux de Sylvester et de Jacobi [3], est exposée dans le manuscrit de Göttingen du cours de l'hiver 1886-1887, sous le titre : *La méthode d'Hermite-Jacobi pour déterminer le nombre de racines réelles d'une équation* [4]. On ne trouve donc dans l'intitulé aucune allusion au théorème de Sturm, même si Kronecker cherche bien ici à donner « l'exposé d'une nouvelle méthode, laquelle nous conduira au théorème de *Sturm* » [5]. La méthode employée ici a en effet un fondement très différent. Elle est introduite ainsi [6] :

> Nous voulons clore nos développement concernant le théorème de Sturm avec l'exposé d'une méthode qui détermine le nombre de racines réelles d'une équation à coefficients réels, découverte simultanément par Hermite et Jacobi. Cette méthode repose sur la théorie des formes quadratiques ; d'un point de vue théorique, elle a l'inconvénient de

1. Voir [Kronecker, 1873b].
2. Voir [Borchardt, 1857].
3. En particulier [Hermite, 1852], [Hermite, 1853], [Sylvester, 1853d] et [Jacobi, 1857].
4. [Kronecker, 1887b] p. 224 : « Hermite-Jacobi'sche Methode die Anzahl der reellen Wurzeln einer Gleichung zu bestimmen ».
5. [Kronecker, 1881a] p. 141 : « Wir wollen nun noch der regulären Fall der Kettenbruchentwicklung ins Auge fassen, um daran die Auseinandersetzung einer neuen Methode zu knüpfen, welche uns zum *Sturm*'schen Satze führt. »
6. Wir wollen unsere den Sturm'schen Satz betreffenden Entwicklungen mit der Darlegung einer gleichzeitig von Hermite und Jacobi gefunden Methode, die Anzahl der reellen Wurzeln einer Gleichung mit reellen Coefficienten zu bestimmen, beschließen. Diese Methode beruht auf der Theorie des quadratischen Formen ; vom theoretischen Gesichtspunkt betrachtet, hat sie den Nachteil, daß sie nur für des regulären Fall giltig ist.
On pourra aussi se reporter à [Kronecker, 1887a] p. 195.

n'être valable que pour le cas régulier. [Kronecker, 1887b] p. 224

Kronecker expose donc cette méthode. Il part de deux fonctions $f(x)$ et $f_1(x)$, définies comme précédemment, mais il suppose cette fois que le développement de $\frac{f}{f_1}$ en fraction continue est régulier, c'est-à-dire que $n_k = n - k$. Rappelons que le degré de $\psi_k$ est donc ici égal à $k$. On a vu que [7]

$$\sum_{h=1}^{n} \frac{\xi_h^t f_k(\xi_h)}{f_1(\xi_h) f'(\xi_h)} = 0$$

pour $t = 0, 1, 2, ...n - k - 1$.

Ainsi, par linéarité, pour $f_i$ tel que $n - i < n - k$, c'est-à-dire si $k < i$ on a :

$$\sum_{h=1}^{n} \frac{f_i(\xi_h) f_k(\xi_h)}{f_1(\xi_h) f'(\xi_h)} = 0$$

Pour $k > i$, on peut inverser les rôles de $f_i$ et $f_k$ et on obtient aussi une somme nulle. Finalement, en notant

$$S_k = \sum_{h=1}^{n} \frac{f_k(\xi_h)^2}{f_1(\xi_h) f'(\xi_h)}$$

on a

$$\boxed{\sum_{h=1}^{n} \frac{f_i(\xi_h) f_k(\xi_h)}{f_1(\xi_h) f'(\xi_h)} = \delta_{ik} S_k}$$

où comme précédemment $\delta_{ik}$ est le symbole de Kronecker.

Soit maintenant $n$ variables indéterminées (*unbestimmte Variablen*) $y_1, ..., y_n$. on peut écrire :

$$\begin{aligned}\sum_{i,k=1}^{n} y_i y_k \sum_{h=1}^{n} \frac{f_i(\xi_h) f_k(\xi_h)}{f_1(\xi_h) f'(\xi_h)} &= \sum_{h=1}^{n} \frac{1}{f_1(\xi_h) f'(\xi_h)} \sum_{i,k=1}^{n} f_i(\xi_h) f_k(\xi_h) y_i y_k \\ &= \sum_{h=1}^{n} \frac{1}{f_1(\xi_h) f'(\xi_h)} \left( \sum_{k=1}^{n} f_k(\xi_h) y_k \right)^2\end{aligned}$$

car lorsque l'on développe le dernier carré, la plupart des termes s'annulent comme on vient de le voir.

Par ailleurs,

$$\begin{aligned}\sum_{i,k=1}^{n} y_i y_k \sum_{h=1}^{n} f_i(\xi_h) f_k(\xi_h) \frac{1}{f_1(\xi_h) f'(\xi_h)} &= \sum_{i,k=1}^{n} \delta_{ik} S_k y_i y_k \\ &= \sum_{k=1}^{n} S_k y_k^2\end{aligned}$$

7. Voir p. 283.

Si on pose

$$\sum_{k=1}^{n} f_k(\xi_h) y_k = z_h$$

Alors on obtient

$$\sum_{k=1}^{n} S_k y_k^2 = \sum_{h=1}^{n} \frac{1}{f_1(\xi_h) f'(\xi_h)} z_h^2$$

On voit que l'on a ainsi obtenu deux formes quadratiques, et que la transformation linéaire :

$$\sum_{k=1}^{n} f_k(\xi_h) y_k = z_h \text{ pour } h = 1, 2, ...n$$

permet de passer de l'une à l'autre. On passe donc d'une somme de carrés à une autre somme de carrés [8].

Kronecker utilise bien sûr à partir de là le principe d'inertie de Sylvester, et il affirmera même que [9]

> le théorème de *Sturm* est seulement une autre façon d'exprimer le principe d'inertie des formes quadratiques. [Kronecker, 1881a] p. 145

## 10.1 Le principe d'inertie de Sylvester

Ce principe – le *Trägheitsgesetz der quadratischen Formen* – Kronecker le présente d'abord et le démontre d'une manière qu'il qualifie d'« intellectuellement extrêmement simple » (*gedanklich höchst einfache* [Kronecker, 1887b] p. 226). Dans son cours de 1880-1881, il attribue cette proposition essentiellement à Jacobi [10] :

> Cette invariance du nombre de signes positifs dans une somme de carrés positifs était connue depuis longtemps, notamment en géométrie. Mais *Jacobi* le premier (*Crelle Bd.* 53) a découvert que dans une forme quadratique quelconque le nombre de signes positifs et négatifs reste inchangé lors d'une transformation linéaire. Presque au même moment *Hermite* et *Sylvester* se sont occupés de cette question. Du dernier vient l'expression : principe d'inertie des formes quadratiques. [Kronecker, 1881a] p. 146

On remarque tout d'abord qu'une fois de plus Kronecker met en avant, même si c'est ici de façon implicite, l'aspect heuristique de la géométrie : elle nous donne la propriété, tout au moins en partie, dont on peut ensuite démontrer une généralisation.

Mais si sa version géométrique, dans un cas particulier, est depuis longtemps utilisée, il attribue bien la première démonstration analytique du cas général à Jacobi. En effet,

8. Si d'ailleurs $f_1(\xi_h) f'(\xi_h)$ est positif, c'est une somme de carrés positifs que l'on a transformé en somme de carrés positifs.

9. der *Sturm*'sche Satz nur ein andrer Ausdruck für das Trägheitsprinzip der quadratischen Formen ist.

10. Diese Invarianz der Anzahl der positiven Zeichen bei einer Summe von positiven Quadraten, war schon frühe bekannt, namentlich in der Geometrie. Erst *Jacobi* aber entdeckte (*Crelle Bd.* 53), daß für eine beliebige quadratische Form die Anzahl der positiven und der negativen Zeichen bei einer linearen Transformation ungeändert bleibt. Fast gleichzeitig haben sich *Hermite* und *Sylvester* mit dieser Frage beschäftigt. Von letzterem stammt der Ausdruck : Trägheitsprinzip der quadratischen Formen.

Borchardt publie en 1857 dans son journal trois articles issus des travaux de Jacobi. L'un d'entre eux s'intitule *Über einen algebraischen Fundamentalsatz und seine Anwendungen*[11] et propose une démonstration de la loi d'inertie. Bien que n'étant pas datée, cette contribution de Jacobi a été écrite avant 1851, année de sa mort, tandis que Sylvester publie sa démonstration dans l'article *A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares* en 1852. À Sylvester ne revient que le nom de la loi. Kronecker ajoute d'ailleurs dans le cours de 1886-1887, à la fin de la démonstration de la loi d'inertie[12]

> En ce qui concerne l'histoire du théorème de Sylvester il faut encore remarquer que Jacobi connaissait déjà le théorème en 1847, mais n'a rien publié à ce sujet ; Brioschi, entre autre, en a aussi donné une preuve – celle-ci est basée sur l'observation directe des formules de transformation – qui est cependant soumise à certaines conditions. [Kronecker, 1887b] p. 228

Ici Kronecker date les recherches de Jacobi et les situe en 1847. Dans l'un des trois articles mentionnés précédemment et publiés en 1857 par Borchardt, il est bien mentionné cette date, mais il ne s'agit pas de celui dont il est ici question. L'article[13] de Brioschi cité dans le passage précédent a été publié en 1856 dans les *Nouvelles Annales de Mathématiques* et s'intitule *Sur les séries qui donnent le nombre de racines réelles des équations algébriques à une ou à plusieurs inconnues.* Celui-ci est le second en lien avec le théorème de Sturm que publie Brioschi dans ce journal[14]. Dans ce dernier, après avoir redémontré le principe d'inertie, il l'utilise pour constituer de nouvelles fonctions de Sturm et présenter les résultats de Hermite sur l'extension du théorème de Sturm à deux variables[15]. La démonstration de la loi d'inertie présentée dans cet article semble à première vue tout aussi générale que celle de Kronecker. Nous essayerons plus tard de voir ce que Kronecker entend par « certaines conditions ».

Voyons comment cette loi est énoncée par Kronecker[16] :

> Si une somme algébrique de carrés de variables réelles à coefficient réels se change en une nouvelle somme de carrés par une transformation linéaire réelle des variables, alors les nombres de carrés positifs et négatifs restent invariants. [Kronecker, 1891a] p. 175

11. Voir [Jacobi, 1857].

12. Hinsichtlich der Geschischte des Sylvester'schen Satzes sei noch bemerkt, daß Jacobi den Satz schon 1847 kannte, aber nichts darüber publiciert hat ; einen Beweis für ihn hat u.A. auch Brioschi gegeben, dieser gründet sich auf directe Betrachtung der Transformation formeln, ist aber an gewisse Bedingungen geknüpft.

13. Voir [Brioschi, 1856].

14. Le premier, publié en 1854, a pour titre *Sur les fonctions de Sturm* [Brioschi, 1854], et donnait une nouvelle démonstration, plus *directe* d'après l'auteur, des formules fournies par Cayley dans son article [Cayley, 1848] de 1848 qui fournissait les fonctions de Sturm à l'aide des coefficients de l'équation.

15. Voir [Hermite, 1852] et [Hermite, 1853].

16. Wenn eine algebraische Summe von quadraten reeller Variabeler mit reellen Koefficienten durch reelle linear Transformation der Variabelen wieder in eine Summe von Quadraten übergeht, so bleibt die Anzahl der positiven, wie der negativen Vorzeichen invariant.

Dans celui [17] du cours de l'hiver 1886-1887, il en donne une version plus explicite qu'il peut donc démontrer :

> (...) toute expression d'une forme quadratique "ordinaire" (c'est-à-dire d'une forme dont le déterminant n'est pas nul) de $n$ carrés et de $n$ variables réelles à coefficients réels, possède le même de nombre de carrés d'un même signe. Pour la forme quadratique $\sum_{k=1}^{n} p_k u_k^2$, $\sum_{k=1}^{n} sgn(p_k)$ est donc un invariant. Si l'on passe ainsi de la forme
>
> $$\sum_{k=1}^{n} p_k u_k^2$$
>
> par la transformation
>
> $$u_h = \sum_{k=1}^{n} \tau_{hk} v_k$$
>
> à $\sum_{k=1}^{n} q_k v_k^2$, alors le théorème de Sylvester signifie :
>
> $$\sum_{k=1}^{n} sgn(p_k) = \sum_{k=1}^{n} sgn(q_k)$$
>
> [Kronecker, 1887b] p. 225

Il est sous-entendu ici que le nombre total de carrés est bien le même dans les deux expressions. Il commence par traiter le cas d'une forme avec deux variables

$$ax^2 + by^2$$

où $a, b$ sont des grandeurs réelles et on regarde ce qui se passe avec la substitution

$$\xi = x + bcy \ , \ \eta = y - acx$$

où $c$ est un nombre réel quelconque. Il n'y a pas ici de perte de généralité : toute transformation linéaire peut être modifiée de façon à se mettre sous cette forme. Dans [Kronecker, 1881a] p. 143, la démonstration est faite sans cette astuce qui permet de simplifier les calculs. On voit là un exemple du soin qui est mis par Kronecker dans la rédaction

17. (...) bei allen Ausdrücken einer „ordinären" quadratischen Form (d.h. einer solchen mit nicht verschwinden der Determinante) von $n$ reellen Variablen, und mit reellen Coefficienten durch $n$ Quadrate, sich dieselbe Anzahl Quadrate eines Zeichens ergiebt. Für die quadratische Form $\sum_{k=1}^{n} p_k u_k^2$ ist also $\sum_{k=1}^{n} sgn(p_k)$ eine Invariante. Führen wir also die Form

$$\sum_{k=1}^{n} p_k u_k^2$$

durch die Transformation

$$u_h = \sum_{k=1}^{n} \tau_{hk} v_k$$

über in $\sum_{k=1}^{n} q_k v_k^2$, so besagt der Sylvester'sche Satz, daß

$$\sum_{k=1}^{n} sgn(p_k) = \sum_{k=1}^{n} sgn(q_k)$$

de ses cours, et du travail qu'il fournit pour les mettre à jour, année après année. Ces changements ne sont pas fortuits et méritent d'être analysés. Alors on a

$$ax^2 + by^2 = \alpha\xi^2 + \beta\eta^2$$

où

$$\alpha = \frac{a}{1+abc^2} \qquad \beta = \frac{b}{1+abc^2}$$

Ici, le manuscrit [Kronecker, 1887b] propose ceci :

$$sgn(\alpha) + sgn(\beta) = sgn(1+abc^2)(sgn(a) + sgn(b)) = sgn(a) + sgn(b)$$

Une explication supplémentaire est nécessaire : clairement, si $a$ et $b$ sont de même signe, alors $\alpha$ est du signe de $a$ et $\beta$ du signe de $\beta$ et la formule précédente est correcte. Si $a$ et $b$ sont de signe contraires, alors

$$\alpha\beta = \frac{a}{1+abc^2}.\frac{b}{1+abc^2} < 0$$

donc $\alpha$ et $\beta$ sont aussi de signes contraires et

$$sgn(\alpha) + sgn(\beta) = sgn(a) + sgn(b) = 0$$

Le théorème est ainsi démontré pour deux variables. Après avoir traité ce cas particulier, Kronecker poursuit en donnant une démonstration par récurrence du théorème qu'il n'est pas nécessaire de retranscrire ici.

## 10.2 Application de la loi d'inertie au théorème de Sturm

Kronecker le dit [18], « le théorème de *Sturm* est seulement une autre façon d'exprimer le principe d'inertie des formes quadratiques ». Pour montrer le lien qui unit ces deux propositions, il considère la forme quadratique suivante :

$$\sum_{h=1}^{n} \frac{1}{f_1(\xi_h)f'(\xi_h)} \left( \sum_{k=1}^{n} y_k f_k(\xi_h) \right)^2$$

Tous les termes de cette somme pour lesquels $\xi_h$ est réel sont aussi réels, car $f(x)$ est supposée à coefficients réels. Soit maintenant $\xi_h$ une racine complexe [19], alors on peut noter

$$\sum_{k=1}^{n} y_k f_k(\xi_h) = Y_h + iY'_h \qquad \frac{1}{f_1(\xi_h)f'(\xi_h)} = \alpha_h + \beta_h.i$$

Si maintenant on note $\xi_{h'}$ la racine conjuguée de $\xi_h$, alors

18. [Kronecker, 1881a] p. 145 : « Wir werden sehen, daß der *Sturm*'sche Satz nur ein andrer Ausdruck für das Trägheitsprinzip der quadratischen Formen ist. »

19. Il s'agit bien ici toujours des racines de l'équation $f(x) = 0$.

$$\sum_{k=1}^{n} y_k f_k(\xi_{h'}) = Y_h - iY'_h \qquad \frac{1}{f_1(\xi_{h'})f'(\xi_{h'})} = \alpha_h - \beta_h.i$$

donc

$$\begin{aligned}
&\frac{1}{f_1(\xi_h)f'(\xi_h)}\left(\sum_{k=1}^{n} y_k f_k(\xi_h)\right)^2 + \frac{1}{f_1(\xi_{h'})f'(\xi_{h'})}\left(\sum_{k=1}^{n} y_k f_k(\xi_{h'})\right)^2 \\
=&(\alpha_h + \beta_h.i)(Y_h^1 - Y_h'^2 + 2iY_hY'_h) + (\alpha_h - \beta_h.i)(Y_h^1 - Y_h'^2 - 2iY_hY'_h) \\
=&2\alpha_h(Y_h^2 - Y_h'^2) - 4\beta_h Y_h Y'_h = \frac{2}{\alpha_h}\left((\alpha_h Y_h - \beta_h Y'_h)^2 - (\alpha_h^2 + \beta_h^2)Y_h'^2\right)
\end{aligned}$$

La dernière ligne montre que l'on peut passer de

$$\sum_{h=1}^{n} \frac{1}{f_1(\xi_h)f'(\xi_h)}\left(\sum_{k=1}^{n} y_k f_k(\xi_h)\right)^2$$

par la substitution linéaire (à coefficients *réels*) :

$$\begin{aligned}
v_i &= \sum_{k=1}^{n} y_k f_k(\xi_i) \text{ pour } i \neq h, h' \\
v_h &= \alpha_h Y_h - \beta_h Y'_h \\
v_{h'} &= Y'_h
\end{aligned}$$

à une somme où la racine complexe et sa conjuguée ne changent rien à la somme des signes des carrées, ces dernières produisant deux signes opposés. La loi d'inertie nous permet donc de négliger, pour le calcul de cette somme, les racines qui ne sont pas réelles. La somme des signes de cette forme quadratique est donc, en utilisant toujours la loi d'inertie,

$$\sum_{\xi} sgn\left(f_1(\xi)f'(\xi)\right) = \sum_{k=1}^{n} sgn(S_k)$$

où la somme est prise sur l'ensemble des racines *réelles* de l'équation $f(x) = 0$. Or :

$$S_k = \sum_{h=1}^{n} \frac{f_k(\xi_h)^2}{f_1(\xi_h)f'(\xi_h)} = \sum_{h=1}^{n} \frac{f_k(\xi_h)\psi_k(\xi_k)}{f'(\xi_h)}$$

Comme $deg(f_k(x)\psi_k(x) = n - k + k - 1 = n - 1)$, les formules d'Euler [20] nous permettent de dire que $S_k$ est égal au terme de plus haut degré de $f_k(x)\psi_k(x)$, c'est-à-dire d'après ce que l'on a vu précédemment [21] $\frac{c_k}{c_{k-1}}$. On obtient finalement

20. Voir p. 143.
21. Voir p. 282.

$$\sum_{\xi} sgn\left(f_1(\xi)f'(\xi)\right) = \sum_{k=1}^{n} sgn(c_k c_{k-1})$$

c'est-à-dire le théorème de Sturm, tel que nous l'avons déjà vu [22], lorsque l'on cherche l'*ensemble* des racines réelles de l'équation $f(x) = 0$.

Dans le cours du *Wintersemester* 1886-1887, Kronecker termine son chapitre sur Sturm sur cette note, mais dans ceux de 1878-1879, 1880-1881 et 1884-1885, il prolonge son travail de la façon suivante, en constatant que [23]

> Cette démonstration du théorème de *Sturm* a le défaut d'utiliser d'abord la preuve de l'existence de racines complexes, lesquelles sont pourtant complètement éliminées du résultat final. Nous voulons ainsi tout de suite envisager le cas général du théorème de *Sturm*, lequel nous conduit en même temps à la formule de *Jacobi*. Ce chemin révèle aussi que le principe de conservation des signes est pour ainsi dire identique au théorème de *Sturm*. [Kronecker, 1881a] p. 146

Une dernière modification qu'il cherche à apporter consiste donc à éviter de supposer l'existence d'une racine complexe, c'est-à-dire de supposer démontré le théorème fondamental de l'algèbre. En effet, quel que soit le manuscrit, ce théorème sera démontré dans la suite du cours, avec un rapport plus ou moins direct au théorème de Sturm [24].

Le dernier paragraphe du chapitre consacré au théorème de Sturm contenu dans le manuscrit de Göttingen du semestre d'hiver 1886-1887 se termine ainsi [25] :

> Cette démonstration du théorème de Sturm semblait à Hermite particulièrement précieuse car elle est libre de toute considération de continuité. Toutefois, cet avantage est très affecté par le fait que la preuve ne fonctionne que pour le cas régulier et que la notion de racine imaginaire d'une équation y est complètement utilisée. Kronecker a libéré la méthode de Hermite de la première de ces limitations. [Kronecker, 1887b] p. 230-231

Kronecker affirme ici qu'Hermite se rangeait donc à ses côtés pour considérer que la continuité n'est pas la base sur laquelle il faut construire ce travail sur les racines. On voit que selon les cas, il peut y avoir une opposition entre exigence de *généralité* et utilisation de concepts extérieurs à l'*arithmétique générale*. Il est alors difficile de choisir, et la bonne

22. Voir p. 220.

23. Diese Deduktion des *Sturm*'schen Satzes hat jedoch den Mangel, daß wir zuerst den Beweis der Existenz komplexer Wurzele zu bringen haben, welche doch bei dem Endresultat wieder vollständig herausfallen. Wir wollen daher sofort den allgemeinen Fall des *Sturm*'schen Satzes ins Auge fassen, welcher uns zugleich zu den *Jacobi*'schen Formeln führt. Auch dieser Weg wird erkennen lassen, daß das Prinzip der Erhaltung der Vorzeichen geradezu identisch mit dem *Sturm*'schen Satze ist.

24. Nous verrons que dans [Kronecker, 1881a] ce théorème est démontré à partir de la théorie des caractéristiques, qui est elle-même issue du travail de Kronecker sur Sturm.

25. Dieser Beweis des Sturm'schen Satzes erschien Hermite darum besonders wertvoll, weil er von Continuitätsbetrachtungen vollständig frei ist. Indes wird dieser Vorzug dadurch sehr beeinträchtigt, daß der Beweis nur für den regulären Fall gilt, und daß der Begriff der imaginären Wurzeln einer Gleichung dabei vollständig verbraucht wird. Von der ersteren Beschränkung hat Kronecker die Hermite'sche Methode frei gemacht.

solution consiste souvent à retravailler la preuve, comme le fait ici Kronecker, pour tenter de lever cette opposition.

# Chapitre 11

# Conclusion

Nous venons donc de voir la richesse des chapitres des *Vorlesungen* consacrés au théorème de Sturm, que ce soit au niveau du contenu, mais aussi de ce qu'ils nous apprennent sur la façon dont Kronecker appréhendait les mathématiques. Nous aurions pu aborder son travail sur ce théorème en partant des articles de recherche que Kronecker lui a consacrés : nous y aurions alors retrouvé une grande partie des résultats que nous venons de présenter. Cependant, l'inclusion serait plutôt dans le sens inverse à celui auquel nous aurions pu penser : les articles sont contenus dans le cours. Nous l'avons déjà souligné, pour Kronecker, comme d'ailleurs pour Weierstrass, les *Vorlesungen* sont l'occasion de présenter leurs travaux de recherches les plus récents, et il n'est pas rare qu'une théorie ne soit publiée qu'après avoir été « testée » et « polie » plusieurs années pendant leurs cours. Kronecker le sous-entend d'ailleurs lorsqu'il propose une formule sur les fractions continues qui, dit-il, ne se trouve pas dans les livres de cours (*die sich nicht in den Lehrbüchen findet* : [Kronecker, 1887b] p. 183). Et en cela nous l'avons vu : la partie sur le théorème de Sturm est le point d'articulation entre sa présentation de résultats plus ou moins classiques et l'exposé de ses recherches les plus récentes. Si Kronecker donne un exposé complet des méthodes algébriques qui permettent, à la suite de Sylvester, Jacobi ou Hermite, de parvenir au théorème de Sturm, ce qui caractérise son travail n'est peut-être pas la poursuite de l'« algébrisation » de ce théorème : la diversité des méthodes utilisées dans ce chapitre semble au contraire marquer un tournant dans cette histoire du théorème de Sturm. Comme l'affirme K.R. Biermann [1] à propos de Kronecker :

> Incroyablement agile et polyvalent, il n'a pas gagné d'importance dans les disciplines prises individuellement, comme Cauchy et Jacobi dans l'analyse, Riemann et Weierstaß dans la théorie des fonctions, Dirichlet et Kummer dans la théorie des nombres, Abel et Galois dans l'algèbre. Sa grandeur repose dans la polyvalence, dans le lien entre la théorie des nombres, l'algèbre et l'analyse. [Biermann, 1973] p. 111-112

Kronecker affirme que le but premier du théorème de Sturm est la « détermination du

1. Unglaublich beweglich und vielseitig, hat er in den einzelnen Disziplinen nicht die Bedeutung erlangt, wie sie Cauchy und Jacobi in der Analysis, Riemann und Weierstraß in der Funktionentheorie, Dirichlet und Kummer in der Zahlentheorie, Abel und Galois in der Algebra erreicht haben. Seine Größe liegt in der Vielseitigkeit, in der Verbindung von Zahlentheorie, Algebra und Analysis.

nombre des racines réelles des équations algébriques » [2], et, dit-il, la plupart des ouvrages le considèrent comme définitivement résolu. Mais ce qu'il reste à accomplir, d'après lui, c'est la *séparation des racines*. Et c'est bien cette vision du théorème qui lui permet comme on l'a vu, de le déconnecter de la continuité.

Le travail que nous venons de faire sur le théorème de Sturm dans les *Vorlesungen über die Theorie der algebraischen Gleichungen* nous a aussi permis de mettre en évidence la volonté de *généralisation* certaine de Kronecker. Et c'est dans les détails techniques que celle-ci se révèle : généralisation des formules sur les fractions rationnelles, extension au cas où ce développement n'est pas régulier, généralisation de la formule d'interpolation de Lagrange aux cas des fractions rationnelles, etc. Nous avons dû, à chaque fois, entrer dans le cœur des calculs pour mettre en avant la recherche de généralité de Kronecker.

La forme que Kronecker donne au théorème de Sturm est elle-même une forme généralisée, au sens où le calcul du nombre de racines réelles d'une équation, et donc le résultat de Sturm, n'est qu'un cas particulier de la formule qu'il donne. Car la forme est bien une *formule*, qui est pour Kronecker, comme nous l'avons vu, la quintessence des mathématiques. Et s'il en attribue la paternité à Sylvester, il faut bien remarquer que cet objet – la formule – n'apparaît que dans son travail.

Il y a donc souvent une extension des cas traités, mais la *généralisation* que semble chercher Kronecker s'apparente davantage à la constitution de diverses méthodes qui, ensemble, permettent de résoudre le problème de la façon la plus large, qu'à l'utilisation d'*une* méthode unique qui serait la plus englobante. Tout en conservant le problème qu'il s'est proposé de résoudre, il s'agit plutôt pour Kronecker de changer de cadre et de proposer des chemins différents : l'interprétation géométrique donne une vision claire des choses, mais qui semble malgré tout profondément liée à la notion de continuité ; les fractions continues lui permettront de partiellement résoudre ce problème, et la méthode de Hermite-Jacobi, avec l'utilisation des formes quadratiques, exclura toute référence à la continuité. Mais si cette dernière méthode, par le plus haut degré d'abstraction qu'elle nécessite, peut sembler à première vue plus générale, il ne faudrait pas s'y tromper : seul le cas régulier [3] est traité, et, autre limitation, ce procédé implique la connaissance de l'existence de racines complexes pour toute équation. Cette dernière remarque, qui est, comme nous l'avons vu, une remarque faite par Kronecker lui-même, nous permet de préciser ce que Kronecker entend par « théorie plus générale » : elle ne doit pas seulement s'appliquer à un plus grand nombre d'objets, mais elle doit aussi supposer le moins de résultats pré-établis possible. C'est bien la confrontation de ces différents points de vue – fractions continues, formes quadratiques, continuité, etc...- qui caractérise le travail de Kronecker sur le théorème de Sturm, et nous allons voir dans la partie suivante que c'est cette même confrontation qui amène Kronecker à donner à sa théorie des caractéristiques

2. *die numerische Bezeichnung der reellen Wurzeln der algebraischen Gleichungen* [Kronecker, 1887b] p. 155.

3. Voir p. 219.

toute son ampleur.

L'une des composantes essentielles des articles d'Hermite auxquels Kronecker se réfère dans le dernier chapitre est totalement absente : la généralisation du théorème de Sturm à des dimensions supérieures, c'est-à-dire, si nous reprenons ce que Hermite fait en 1853, la possibilité de trouver le nombre de couples de réels qui sont solutions de deux équations à deux inconnues, ou de triplet de réels solutions de trois équations à trois inconnues dans certains *rectangles* ou *parallélépipèdes droits*[4]. Kronecker étend ainsi sa réflexion à des dimensions de degré supérieur, et ce faisant, il met en place sa *théorie des caractéristiques*. Nous allons maintenant examiner cette théorie, et nous montrerons qu'elle n'est pas une simple généralisation du théorème de Sturm. Elle ne se contente pas de chercher simplement à étendre la recherche des solutions d'un système d'équations, comme le faisait Hermite, mais elle produit des outils qui permettront de travailler sur les variétés que définissent ces systèmes.

4. Voir [Hermite, 1853].

# Troisième partie

# La théorie des caractéristiques

# Chapitre 12

# Introduction


12.1 Les sources . . . . . **323**
12.2 Contenu des manuscrits . . . . . **325**
12.3 Une première exposition de la théorie . . . . . **328**
12.3.1 Remarques préliminaires . . . . . 329
12.3.2 Intersection de $\varphi$ et $\psi$ . . . . . 331
12.3.3 Caractéristique d'un système de trois fonctions . . . . . 332
12.3.4 Points d'intersection de deux courbes du plan dans un compact 335
12.3.5 Racines complexes . . . . . 338
12.3.6 Détermination de la caractéristique . . . . . 341
12.4 Conclusion . . . . . **342**


Comme nous l'avons vu, le théorème de Sturm, incontournable pour toute personne s'intéressant aux mathématiques de la seconde moitié du XIX^e siècle, sombra ensuite dans l'oubli au cours du XX^e siècle avant sa redécouverte par Tarski au XX^e siècle. La *théorie des caractéristiques* de Kronecker, en revanche n'a pas connu pareille éclipse. À la suite de Kronecker, seul un très petit nombre d'ouvrages expose cette théorie. Elle ne donna pas lieu, contrairement au théorème de Sturm, à un « cycle d'idées » qui porterait son nom, et aujourd'hui encore, elle est complètement ignorée de l'essentiel de la communauté mathématique, ainsi que de celle des historiens des mathématiques.

Pourtant cette théorie revêt pour Kronecker une grande importance : nous allons donc essayer de comprendre les raisons d'un tel décalage. En effet, presque vingt ans après le premier exposé qu'il en donne, Kronecker voit toujours dans cette théorie l'une des sources de ses idées les plus récentes. Lors d'un voyage à Berlin en 1888, Hilbert a un long entretien avec Kronecker, en partie à propos de ce qui deviendra son seizième problème [1]. À cette occasion, et concernant le travail de Harnack sur les configurations topologiques possibles des courbes algébriques, Hilbert rapporte les propos de Kronecker [2] :

1. Ce problème, tout au moins sa première partie, demande de compter le nombre de branches réelles, c'est-à-dire d'ovales, d'une courbe algébrique, et de donner des indications à propos de leur disposition.
2. Er könne Zahl und Lage der Züge bestimmen, mit der Charakteristik sei Alles gemacht (Arbeit vom

> [il dit] qu'il peut déterminer le nombre et la position des branches, qu'avec la caractéristique tout est fait (Travail de l'année 1878). Aussi le travail d'Harnack et mes futures recherches sont inclus dans cette théorie. [Hilbert, 1888]

Ce point de vue sur sa théorie, que D. Rowe qualifie de « bien trop optimiste » [3], montre donc que Kronecker pense pouvoir résoudre d'importantes questions à l'aide de celle-ci. Mais quelles sont ces questions que Kronecker résout ? Car le but que se propose d'atteindre Kronecker lorsqu'il élabore cette théorie est clairement énoncé : il s'agit bien pour lui de prolonger le travail de Sturm sur la détermination du nombre de racines réelles d'une équation. Il écrit d'ailleurs dans son article de mars 1869 qui constitue sa présentation de la théorie [4] :

> Dans les recherches dont j'ai esquissé ici le développement je suis parti du théorème de *Sturm*. Une extension de celui-ci aux systèmes d'équations a déjà été donnée depuis longtemps par M. *Hermite*, mais il m'importait en outre de généraliser le procédé même des fractions continues qui est au fondement des développements de *Sturm* et, après que cela fut fait, d'interpréter naturellement les résultats généraux qui sont ainsi obtenus. [Kronecker, 1869c] p. 205

Mais ce point de départ, c'est-à-dire la seule extension du théorème de Sturm à la recherche du nombre de solutions d'un système d'équations, est-il vraiment le cœur de son travail ? Dans une certaine mesure Hermite, et Kronecker le cite [5], a déjà traité cette question [6] : ce que Kronecker souhaite ajouter, c'est une méthode, les fractions continues, et une interprétation des résultats. Nous avons vu dans la partie précédente comment Kronecker utilisait les fractions continues pour son travail sur le théorème de Sturm, et comment, selon le contexte, il en multipliait les interprétations graphiques. Nous allons maintenant présenter la poursuite de ce travail à travers sa théorie des caractéristiques.

Nous verrons que cette théorie sera ainsi l'« acte ultime de généralisation » du théorème de Sturm, et que finalement elle en changera profondément la nature. Car cette généralisation ne s'inscrit pas seulement dans le mouvement d'algébrisation du théorème de Sturm tel que le décrit H. Sinaceur. Par un changement de point de vue, Kronecker est amené à en transformer radicalement le contenu, ce qui le conduit à se placer dans un contexte que l'on pourrait qualifier, nous en verrons les raisons, de topologico-analytique.

---

Jahre 1878). Auch die Harnack'sche Arbeit und meine künftigen Untersuchungen seien der Theorie nach darin enthalten.

3. [Rowe, 2005] p. 75 : « overly optimistic assessment ».

4. Bei den Untersuchungen, deren Entwickelungsgang ich hier in Umrissen dargelegt habe, bin ich vom *Sturm*'schen Satze ausgegangen. Eine Ausdehnung desselben auf Systeme von Gleichungen ist schon vor längerer Zeit von Hrn. *Hermite* angegeben worden, aber es kam mir überdies darauf an, das den *Sturm*'schen Entwickelungen zu Grunde liegende Kettenbruchs-Verfahren selbst zu verallgemeinern und nachdem dies geschehen war die dadurch erhaltenen allgemeineren Resultate naturgemäss zu interpretiren.

5. Kronecker donne la note suivante : « *Ch. Hermite*, Sur l'extension du théorème de Sturm à un système d'équations simultanées. Comptes rendus t. 35. 1852. II. p. 52.

*Ch. Hermite*, Remarques sur le théorème de M. *Sturm*. Comptes rendus t. 36. 1853. I. p. 294. » [Kronecker, 1869c] p. 205.

6. Voir la partie précédente p. 305, [Hermite, 1852] et [Hermite, 1853].

Enfin, cette théorie est réputée difficile, comme d'ailleurs la plupart des travaux de Kronecker. L'un des objectifs que nous souhaitons atteindre sera justement de la rendre abordable. Nous allons ainsi montrer comment l'utilisation de son cours sur la théorie des équations algébriques rend possible cet accès, et dans une certaine mesure, lui est même nécessaire.

Dans cette partie, nous souhaitons donc présenter la théorie des caractéristiques de Kronecker, et pour cela, nous allons de nouveau, comme pour le théorème de Sturm, nous plonger dans les manuscrits des cours sur la théorie des équations algébriques. C'est en effet dans ce cours que Kronecker présente cette théorie. Mais contrairement au théorème de Sturm qui est présent dans chacun des manuscrits, la caractéristique d'un système de plusieurs fonctions n'apparaît que dans la moitié d'entre eux. Nous allons donc exposer la théorie telle que Kronecker le fait d'abord dans ses *Vorlesungen*, puis dans ses articles de recherche. Cela nous permettra de montrer comment cette théorie achève de *généraliser* – dans un sens sur lequel nous reviendrons – le théorème de Sturm.

Nous allons commencer, dans cette introduction, par présenter les sources qui nous permettront d'aborder cette théorie, puis nous en donnerons un premier aperçu.

## 12.1 Les sources

Kronecker développe principalement dans deux articles[7] publiés entre 1869 et 1878 la notion de caractéristique d'un système de $n+1$ fonctions à $n$ variables. Cela lui permet, entre autres, de déterminer le nombre de solutions d'un système d'équations à plusieurs inconnues, tout au moins à l'intérieur d'un domaine qui sera dans un premier temps ce qui correspond à ce que nous appellerions aujourd'hui un compact. Pour ce qui est de son œuvre publiée, ce sera donc à ces articles que nous nous référerons essentiellement. Nous pouvons ajouter à cette liste les articles *Über die verschiedenen Sturm'schen Reihen und ihre gegenseitigen Beziehungen*[8] de 1873, *Über Sturm'sche Functionen*[9] de 1878, *Über symmetrische Systeme*[10] de 1889, ainsi qu'une lettre de 1891 intitulée *Sur le nombre de racines communes à plusieurs équations simultanées*[11] qui constitue une réponse à un article d'Émile Picard (1856-1941) sur la théorie des caractéristiques. Nous reviendrons plus loin sur la controverse que cet échange a suscitée. Mais comme nous l'avons dit, les outils qui nous permettent de vraiment entrer dans cette théorie sont les manuscrits des cours que Kronecker a consacrés à la théorie des équations algébriques.

7. Celui de 1869, en deux parties :[Kronecker, 1869c], [Kronecker, 1869d] et celui de 1878 [Kronecker, 1878b].
8. Voir [Kronecker, 1873b].
9. Voir [Kronecker, 1878c].
10. Voir [Kronecker, 1889c].
11. Voir [Kronecker, 1891b].

Cette théorie est présente, sous une forme plus ou moins développée, dans cinq des dix manuscrits de son cours sur la théorie des équations algébriques en notre possession. En fait, elle disparaît après l'exemplaire de 1880-81, pour réapparaître, d'abord succinctement, puis sous sa forme la plus développée à partir de 1888. Il y a donc une période de presque dix ans pendant laquelle Kronecker ne l'enseigne plus. Nous en donnerons, dans le cours de cette partie, quelques éléments d'explication. En plus des articles et des manuscrits de Kronecker mentionnés, nous utiliserons les documents suivants qui développent ou utilisent explicitement la théorie des caractéristiques de Kronecker, et qui ont été publiés entre 1878, année de la publication du second article de Kronecker, et la fin du premier quart du XX$^{e}$ siècle.

Liste de l'ensemble des articles

— *Beiträge zur Analysis situs* de W. Dyck en 1888 ([Dyck, 1888])

— *Beiträge zur Analysis situs II.* de W. Dyck en 1890 ([Dyck, 1890])

— *Sur le nombre de racines communes à plusieurs équations simultanées* d'E. Picard en 1891 ([Picard, 1891c])

— *Sur la recherche du nombre de racines communes à plusieurs équations simultanées* d'E. Picard en 1891 ([Picard, 1891b])

— *Du nombre de racines communes à plusieurs équations simultanées* d'E. Picard en 1891 ([Picard, 1891a])

— *Traité d'analyse vol 1.* d' E. Picard en 1891 ([Picard, 1891d])

— *Sur le nombre de racines communes à plusieurs équations simultanées* d'E. Picard en 1892 ([Picard, 1892])

— *Sur une extension du théorème de Sturm* de Lars Edvard Phragmén en 1892 ([Phragmén, 1892])

— *Traité d'analyse vol 2.* d'E. Picard en 1893 ([Picard, 1893])

— *Bemerkungen zu Kronecker's Theorie der Charakteristiken von Functionen-Systemen.* de W. Dyck en 1894 ([Dyck, 1895a])

— *Sur la détermination du nombre de racines communes à un système d'équations simultanées et sur le calcul de la somme des valeurs d'une fonction en ces points.* de W. Dyck en 1891 ([Dyck, 1894])

— *Ueber Systeme von Functionen reeller Variabeln* de p. Stäckel en 1893 ([Stäckel, 1893])

— *Sur les racines communes à plusieurs équations* de W. Dyck en 1895 ([Dyck, 1895b])

— *Lehrbuch Der Algebra* de H. Weber en 1895 ([Weber, 1895])

— *Beiträge zur Potentialtheorie* de W. Dyck ([Dyck, 1896])

— *Vorlesungen über algebra* de Netto en 1900 ([Netto, 1900])

— *Sur le nombre des racines communes à plusieurs équations* de Davidoglou ([Davidoglou, 1901a] et [Davidoglou, 1901b])

— Deuxième édition du Tome 1 du *Traité d'analyse* de Picard en 1901([Picard, 1891d])

— *Theorie der algebraischen Funktionen einer Variablen : und ihre Anwendung auf algebraische Kurven und Abelsche Integrale* de K. Hensel et G. Landsberg en 1902 ([Hensel et Landsberg, 1902])

— *Encyclopédie des sciences mathématiques, pure et appliquées : v. 1. Arithmétique* de J. Molk en 1904 ([Molk, 1904])

— *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen.* de Runge en 1904 ([Meyer, 1904])

— *Introduction à la théorie des fonctions d'une variable* de J. Tannery en 1910 : article de Hadamard ([Tannery et Hadamard, 1910])

— *Beiträge zur Topologie geschlossener Kurven mit Knotenpunkten und zur Kroneckerschen Charakteristikentheorie* de G. Landsberg en 1911 ([Landsberg, 1911])

— *An unpublished theorem of Kronecker respecting numerical equations* par H.B. Fine en 1914 ([Fine, 1914])

— *Teorema di Sturm e sue estensioni* de Gaspare Mignosi en 1925 ([Mignosi, 1925])

## 12.2 Contenu des manuscrits

Comme nous l'avons fait pour le théorème de Sturm, regardons, sans entrer dans les détails, ce que contient chacun des manuscrits des *Vorlesungen* en rapport avec cette théorie.

**Le manuscrit du semestre d'hiver 1872-73**

— p. 168 Définition intérieur / extérieur.

— p. 169 Point d'entrée et de sortie.

— p. 170 Définition *géométrique* de la caractéristique de trois fonctions à deux inconnues.

— p. 172 Construction du déterminant

$$\Delta(f,\varphi,\psi) = \begin{vmatrix} f & \varphi & \psi \\ f_1 & \varphi_1 & \psi_1 \\ f_2 & \varphi_2 & \psi_2 \end{vmatrix}$$

et introduction des notations $\dfrac{\partial\Delta}{\partial f} = \Delta_1$...

— p. 173 : *Fortgangsprincip.*

— p. 174 : Démonstration des différentes formules donnant la caractéristique d'un système de trois fonctions à deux inconnues.

— p. 178 : Cas où deux des fonctions sont les parties réelles et imaginaires d'une fonction complexe.

— p. 179 : Passage à la forme intégrale de la caractéristique

$$\chi(f,\varphi,\psi) = \frac{1}{2i\pi}\int_{f(x,y)=0} d\log\Theta(x+yi) + \frac{i}{4\pi}\int_{f(x,y)=0} d\log(\varphi^2+\psi^2)$$

— p. 182 : Démonstration du théorème de Cauchy.

— p. 183 : 1re démonstration du théorème fondamental de l'algèbre (3e de Gauss). (énoncé du théorème p. 185)

— p. 190 : 2e démonstration du théorème fondamental de l'algèbre(1re et 4e de Gauss).

**Les manuscrits des semestres d'hiver 1878-79 et 1880-81**

— p. 155 et p. 483 Caractéristique de deux fonctions à une variable.

— p. 156 et p. 486 Stabilité de la caractéristique par substitutions linéaires.

— p. 157 et p. 487 Interprétation graphique de la caractéristique.

— p. 157 et p. 488 Lien avec le théorème de Sturm.

— p. 157 et p. 489 Modification de la caractéristique en fonction des coefficients de $\varphi$ et $\psi$.

— p. 163 et p. 510 Caractéristique d'un système de trois fonctions $f, \varphi, \psi$ à deux variables.

— p. 164 et p. 511 Construction du déterminant

$$\Delta(f,\varphi,\psi) = \begin{vmatrix} f & \varphi & \psi \\ f_1 & \varphi_1 & \psi_1 \\ f_2 & \varphi_2 & \psi_2 \end{vmatrix}$$

et introduction des notations $\dfrac{\partial\Delta}{\partial f} = \Delta_1$...

— p. 164 et p. 512 *Fortgangsprincip.*

— p. 165 et p. 515 Application à deux fonctions $U$ et $V$.

— p. 165 et p. 516 $U = f$ et $V = \varphi\psi$.

— p. 166 et p. 518 Démonstration des différentes formules donnant la caractéristique d'un système de trois fonctions à deux inconnues.

— p. 167 et p. 522 Cas où deux des fonctions sont les parties réelles et imaginaires d'une fonction complexe.

— p. 168 et p. 526 Passage à la forme intégrale de la caractéristiques

$$\chi(f,\varphi,\psi) = \frac{1}{2i\pi}\int_{f(x,y)=0} d\log\Theta(x+yi) + \frac{i}{4\pi}\int_{f(x,y)=0} d\log(\varphi^2+\psi^2)$$

— p. 535 Démonstration du théorème de Cauchy.

— p. 542 Modification de la caractéristique par transformation linéaire.

— p. 545 Modification de la caractéristique en fonction de la modification de ses coefficients.

— p. 549 $2^{\text{e}}$ démonstration du théorème fondamental de l'algèbre ($1^{\text{re}}$ et $4^{\text{e}}$ de Gauss).

— p. 558 Cas particulier : $\Theta(x+iy) = (x+iy)^2 - p$.

— p. 566 Caractéristique de $n+1$ fonctions de $n$ variables.

**Le manuscrit du semestre d'hiver 1888-89** Perdu au milieu du court chapitre sur le théorème de Sturm, on trouve un passage décrivant la caractéristique de deux fonctions à une variable.

— p. 132 Définition de la caractéristique de deux fonctions à une variable.

— p. 133 Interprétation graphique de la caractéristique de deux fonctions à une variable.

**Le manuscrit du semestre d'hiver 1890-91**

— p. 209 Définition intérieur / extérieur.

— p. 210 Point d'entrée et de sortie.

— p. 211 Définition de la caractéristique de deux fonctions à une variable.

— p. 224 Application aux équations du $4^{\text{e}}$ degré.

— p. 229 Caractéristique d'un système de trois fonctions $f, \varphi, \psi$ à deux variables.

— p. 230 Construction du déterminant

$$\Delta(f,\varphi,\psi) = \begin{vmatrix} f & \varphi & \psi \\ f_1 & \varphi_1 & \psi_1 \\ f_2 & \varphi_2 & \psi_2 \end{vmatrix}$$

et introduction des notations $\dfrac{\partial \Delta}{\partial f} = \Delta_1 ...$

— p. 231 *Fortgangsprincip.*

— p. 233 Démonstration des différentes formules donnant la caractéristique d'un système de trois fonctions à deux inconnues.

— p. 237 Extension au cas où les courbes ne sont pas fermées.

— p. 239 Modification de la caractéristique lors d'une variation des coefficients des fonctions.

— p. 241 Cas où deux des fonctions sont les parties réelles et imaginaires d'une fonction complexe : $1^{\text{re}}$ démonstration du théorème fondamental de l'algèbre.

— p. 243 Passage à la forme intégrale de la caractéristiques

$$\chi(f,\varphi,\psi) = \frac{1}{2i\pi}\int_{f(x,y)=0} d\log\Theta(x+yi) + \frac{i}{4\pi}\int_{f(x,y)=0} d\log(\varphi^2+\psi^2)$$

— p. 246 Intégrale de Cauchy

— p. 247 $2^{\text{e}}$ démonstration du théorème fondamental de l'algèbre($3^{\text{e}}$ de Gauss).

— p. 248 Une histoire du théorème fondamental de l'algèbre.

Avant de donner une première exposition de cette théorie, nous pouvons déjà, à partir des sommaires précédents, faire quelques remarques. Tout d'abord, comme nous l'avions annoncé, cette théorie est absente des *Vorlesungen* entre 1881 et 1890, à l'exception des deux pages que Kronecker lui consacre durant l'hiver 1888-1889. Nous reviendrons sur ce point plus loin. Ensuite, la théorie n'est généralement exposée que pour le cas de deux ou trois fonctions. Le cas général n'est présent que dans le manuscrit de l'hiver 1878-1879, c'est-à-dire celui qui suit immédiatement la parution de son article *Über die Charakteristik von Funktionen-Systemen*. Nous exposerons nous aussi cette théorie principalement en petite dimension, mais nous essaierons tout de même de rendre compte du passage à $n+1$ fonctions de $n$ variables, tel que Kronecker le fait dans ses articles de recherche.

Les sommaires précédents nous permettent aussi de faire apparaître la structure de cette théorie, dont nous devons rendre compte. On a ainsi :

1. La caractéristique de deux fonctions d'une variable.
2. Lien avec le théorème de Sturm – Interprétation géométrique
3. La caractéristique de trois fonctions de deux variables
4. Modification de la caractéristique lors de la variation continue de ses coefficients.
5. Application de la théorie
   (a) Théorème fondamental de l'algèbre (sous deux formes)
   (b) Théorie de Cauchy (Intégrale)

L'ordre d'exposition reste globalement le même, et nous examinerons ponctuellement les petites variations qui peuvent intervenir. Le contenu en revanche évolue fortement, et ce sera là l'un des points que nous étudierons.

## 12.3 Une première exposition de la théorie

Nous allons, dès l'introduction, donner un aperçu de cette théorie des caractéristiques. Pour cela, nous allons commencer par l'exposer, non pas telle que la présente Kronecker, mais telle que l'expose Heinrich Weber dans son *Lehrbuch der Algebra*. Pourquoi procéder ainsi ? Tout d'abord, comme nous l'avons vu dans la première partie, comparer le cours de Kronecker avec celui de Weber nous permet de mieux comprendre l'originalité des *Vorlesungen* : il sera donc intéressant d'examiner en détail la présentation de Weber. Nous aurions pu exposer d'abord celle de Kronecker, et conserver ainsi une cohérence chronologique. Cependant, le talent de pédagogue de Weber nous a convaincu qu'exposer sa version est une bonne introduction à cette théorie. Comme nous l'avons rappelé au

début de l'introduction, la théorie des caractéristiques de Kronecker n'est pas connue par la communauté scientifique : c'est pourquoi nous allons insister, tout au long de cette partie, sur le contenu mathématique et en multiplier les approches. Même si le but de notre exposé n'est pas de tenir lieu de « cours » sur la théorie, il nous semble nécessaire d'en donner une description mathématique suffisamment précise pour permettre au lecteur d'en comprendre les enjeux.

Comme le fait la plupart du temps Kronecker dans ses cours, Weber se place en petite dimension et ne traite que le cas de trois fonctions numériques de deux variables. Pour ce dernier, l'objet de cette théorie est extrêmement clair :

> Les propositions de Sturm se généralisent d'une façon remarquable par la *Théorie des caractéristiques* de Kronecker ; elle se propose le même objet que le théorème de Sturm, pour les systèmes d'équations à plusieurs inconnues. [Weber, 1898] p. 339

C'est d'ailleurs dans le chapitre sur le théorème de Sturm qu'il expose cette théorie, et dans le cadre d'un travail sur les racines. Il ne s'agit pour Weber que de prolonger le théorème de Sturm, et donc de répondre à la question suivante : combien de solutions a un système d'équations à plusieurs inconnues ? Ce n'est pourtant là que l'une des applications de cette théorie, et nous verrons plus tard comment Kronecker, en tentant de répondre à cette question, en change la nature et comment « ce même objet » ne sera finalement plus tout à fait le même.

### 12.3.1 Remarques préliminaires

Nous avons déjà présenté le *Lehrbuch der Algebra* dans la première partie [12]. Précisons simplement que la théorie des caractéristiques est exposée dans le livre II du premier tome, donc dans la partie qu'il consacre aux racines d'une équation, et plus précisément dans le second chapitre de ce livre II qui développe le théorème de Sturm et ses applications.

Weber commence par présenter le matériel sur lequel il travaille. Soient $\varphi(x, y)$ et $\psi(x, y)$ deux fonctions réelles (de $\mathbb{R}^2$ dans $\mathbb{R}$). On appellera $\varphi$ et $\psi$ les courbes du plan rapporté à un repère orthonormal d'équations $\varphi(x, y) = 0$ et $\psi(x, y) = 0$.

On commence par supposer que $\varphi$ et $\psi$ sont « fermées, sans branches infinies » [13], c'est-à-dire en fait que l'ensemble des points tels que $\varphi(x, y) = 0$ et $\psi(x, y) = 0$ soient des lacets [14].

Faisons deux remarques sur la façon de se déplacer sur la courbe :

— On instaure un *sens de parcours* sur $\varphi$ : on « avance » (sens positif) en laissant le domaine où $\varphi < 0$ sur la gauche.

— Quitte à remplacer $\varphi$ par $-\varphi$, on a $\varphi > 0$ à « l'extérieur ».

12. Voir pp. 155 et suivantes.
13. Voir [Weber, 1898] p. 339.
14. Au sens où nous l'entendons aujourd'hui, comme par exemple dans [Dieudonné, 1964].

Gauss, dans sa première et dernière démonstration du Théorème fondamental de l'algèbre instaure la même convention [15], et nous verrons par la suite que ce sens de parcours, et la possibilité de l'instaurer, constituent l'un des fondements de notre théorie.

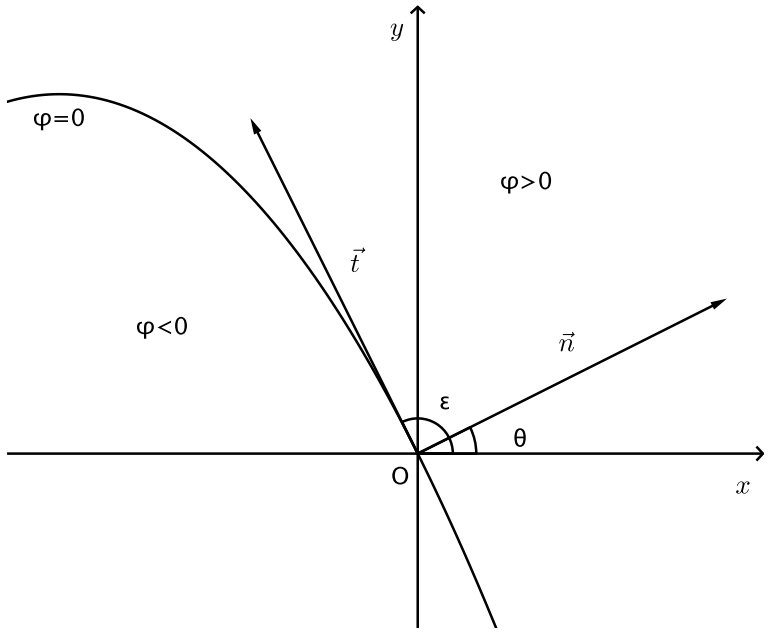



FIGURE 1 –

Prenons un point quelconque de la courbe $\varphi$. Pour simplifier les calculs, plaçons l'origine de notre repère en ce point. On notera [16] $\varphi'_x = \frac{\partial \varphi}{\partial x}(0,0)$ et $\varphi'_y = \frac{\partial \varphi}{\partial y}(0,0)$. En ce point, le vecteur tangent est

$$\vec{t}\begin{pmatrix} -\varphi'_y \\ \varphi'_x \end{pmatrix}$$

et le vecteur normal est

$$\vec{n}\begin{pmatrix} \varphi'_x \\ \varphi'_y \end{pmatrix}$$

On prend en $O$ un arc $ds$ ayant $\vec{t}$ comme direction et on pose $dx$ et $dy$ ses projections sur les axes. Alors le vecteur de coordonnées $(dx, dy)$ sera colinéaire à $\vec{t}$ et de même sens que $\vec{t}$. Il existe donc $\lambda > 0$ tel que

$$\begin{cases} \lambda dx = -\varphi'_y \\ \lambda dy = \varphi'_x \end{cases}$$

Soit $\Phi$ une autre fonction. On pose :

$$d\Phi = \Phi'_x dx + \Phi'_y dy$$

15. Voir [Gauss, 1799] ainsi que plus loin le chapitre consacré au *Fortgangsprincip* (p. 359).
16. On veut donc déjà que les fonctions aient des dérivées partielles.

et

$$[\varphi, \Phi] = \begin{vmatrix} \varphi'_x & \varphi'_y \\ \Phi'_x & \Phi'_y \end{vmatrix}$$

Alors on a :

$$[\varphi, \Phi] = \varphi'_x \Phi'_y - \varphi'_y \Phi'_x = \lambda \Phi'_y dy + \lambda \Phi'_x dx = \lambda d\Phi$$

Comme $\lambda > 0$, on la propriété suivante :

$$\boxed{d\Phi \text{ et } [\varphi, \Phi] \text{ ont le même signe}}$$

Le travail qui vient d'être fait porte donc en fait sur le signe du *Jacobien* de la fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$ qui à $(x, y)$ associe $(\varphi(x, y), \Phi(x, y))$. Remarquons tout de suite que ce signe dépend du sens de parcours choisi.

### 12.3.2 Intersection de $\varphi$ et $\psi$

Observons maintenant les points d'intersection de $\varphi$ et $\psi$ : on décide de parcourir $\varphi$ dans le sens positif (c'est-à-dire en laissant l'intérieur de la courbe sur la gauche) et on regarde les points d'entrée (c'est-à-dire lorsque l'on passe de l'extérieur à l'intérieur de $\psi$) et les points de sortie (c'est-à-dire lorsque l'on passe de l'intérieur à l'extérieur de $\psi$).

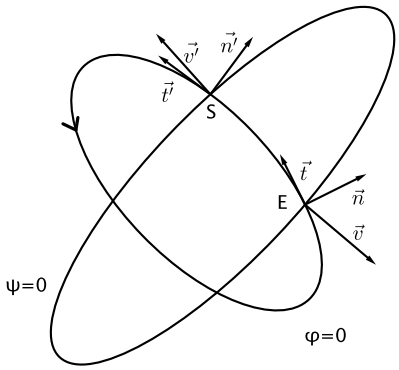



FIGURE 2 –

On notera $E(\varphi, \psi)$ un point d'entrée et $S(\varphi, \psi)$ un point de sortie et, comme précédemment, $\vec{t}$ (resp. $\vec{t'}$) le vecteur tangent à $\varphi$ en $E(\varphi, \psi)$ (resp. en $S(\varphi, \psi)$), $\vec{n}$ (resp. $\vec{n'}$) le vecteur normal à $\varphi$ en $E(\varphi, \psi)$ (resp. en $S(\varphi, \psi)$), et $\vec{v}$ (resp. $\vec{v'}$) le vecteur normal à $\psi$ en $E(\varphi, \psi)$ (resp. en $S(\varphi, \psi)$).

Plaçons-nous au point d'entrée $E(\varphi, \psi)$. On a :

$$d\psi = \psi'_x dx + \psi'_y dy = \frac{1}{\lambda} \vec{v} \cdot \vec{t}$$

Or $(\vec{v}, \vec{t}) \in ]\frac{\pi}{2}; \frac{3\pi}{2}[$ (sinon $\varphi$ n'« entre » pas dans $\psi$), et donc $\vec{v} \cdot \vec{t} < 0$, c'est-à-dire que $d\psi < 0$, et finalement d'après ce qui précède, $[\varphi, \psi] < 0$. Les cas limites, qui correspondent

en fait au cas de deux courbes tangentes – ce qui provoque l'annulation du déterminant fonctionnel – ne sont pas traités ici. On a de même, au point de sortie $S(\varphi, \psi)$, $[\varphi, \psi] > 0$.

On a donc finalement :

$$\begin{cases} [\varphi, \psi] < 0 \text{ aux points d'entrée } E(\varphi, \psi) \\ [\varphi, \psi] > 0 \text{ aux points de sortie } S(\varphi, \psi) \end{cases}$$

Remarques

1. Il y a, les courbes étant fermées, autant de points d'entrée que de sortie.
2. Chaque point d'entrée $E(\varphi, \psi)$ est aussi un point de sortie $S(\psi, \varphi)$, et inversement.
3. On donne la définition suivante : la partie du plan où $\psi \cdot \varphi < 0$ est appelée l'*enclos* (*Binnenraum*) de $\psi\varphi$.

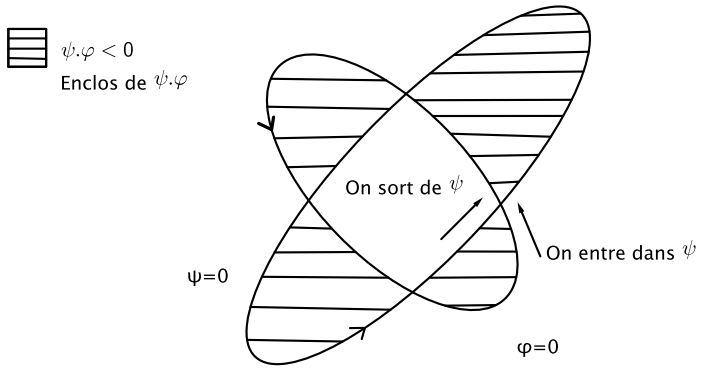



FIGURE 3 –

Nous avons ainsi pu caractériser, lorsque l'on regarde l'intersection de deux courbes en ayant choisi de parcourir l'une dans le sens que l'on appelle positif, les points d'entrée et de sortie par le signe du jacobien $[\varphi, \psi]$. Ce travail préliminaire étant terminé, nous allons pouvoir définir la caractéristique d'un système de trois fonctions.

### 12.3.3 Caractéristique d'un système de trois fonctions

On considère maintenant trois fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$, $\varphi$, $\psi$ et $f$ dont les courbes correspondantes délimitent des compacts (fermés et bornés).

Avec ces 3 fonctions, on peut faire 2 cycles différents :

1. $f, \varphi, \psi, f, \varphi, \psi, f, ...$ que l'on qualifiera de positifs.
2. $f, \psi, \varphi, f, \psi, \varphi, f, ...$ que l'on qualifiera de négatifs.

À partir de ces deux cycles, on construit 6 systèmes différents :

1. à $(f, \varphi, \psi)$ correspondra le sens de parcours positif sur $f$ et le décompte des points d'intersection avec $\varphi$ en considérant l'enclos $\varphi\psi$.

2. à $(\varphi, \psi, f)$ correspondra le sens de parcours positif sur $\varphi$ et le décompte des points d'intersection avec $\psi$ en considérant l'enclos $\psi f$.
3. à $(\psi, f, \varphi)$ correspondra le sens de parcours positif sur $\psi$ et le décompte des points d'intersection avec $f$ en considérant l'enclos $f\varphi$.
4. à $(f, \psi, \varphi)$ correspondra le sens de parcours négatif sur $f$ et le décompte des points d'intersection avec $\psi$ en considérant l'enclos $\psi\varphi$.
5. à $(\psi, \varphi, f)$ correspondra le sens de parcours négatif sur $\psi$ et le décompte des points d'intersection avec $\varphi$ en considérant l'enclos $f\varphi$.
6. à $(\varphi, f, \psi)$ correspondra le sens de parcours négatif sur $\varphi$ et le décompte des points d'intersection avec $f$ en considérant l'enclos $f\psi$.

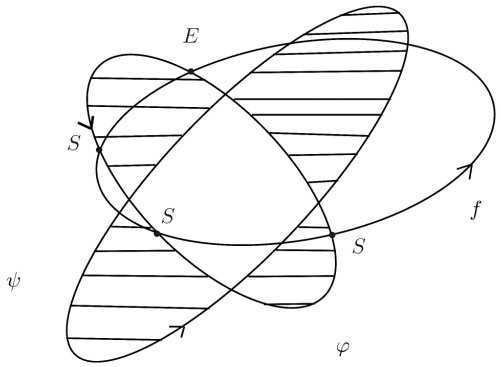



FIGURE 4 – Parcours $(f, \varphi, \psi)$

Considérons le premier système $(f, \varphi, \psi)$. Un point d'intersection de $f$ et $\varphi$ sera un point de sortie, noté $S(f, \varphi, \psi)$, lorsqu'en parcourant $f$ dans le sens positif on passe de l'intérieur vers l'extérieur de l'enclos $\varphi\psi$. Dans le cas contraire, ce sera un point d'entrée noté $E(f, \varphi, \psi)$.

Notons $E[f, \varphi, \psi]$ le nombre de points d'entrée et $S[f, \varphi, \psi]$ le nombre de points de sortie. $E[f, \varphi, \psi] + S[f, \varphi, \psi]$ est le nombre d'intersections entre $\varphi$ et $f$, c'est donc un nombre pair, et il en est de même pour $E[f, \varphi, \psi] - S[f, \varphi, \psi]$.

On donne alors la définition suivante : l'entier $\frac{E[f,\varphi,\psi]-S[f,\varphi,\psi]}{2}$ est appelé *la caractéristique* du système $f, \varphi, \psi$.

Pour que la définition soit valide, il faut donc montrer que cette caractéristique est indépendante du système choisi, et donc que les valeurs suivantes sont égales :

1. $\dfrac{E[f, \varphi, \psi] - S[f, \varphi, \psi]}{2}$
2. $\dfrac{E[\varphi, \psi, f] - S[\varphi, \psi, f]}{2}$
3. $\dfrac{E[\psi, f, \varphi] - S[\psi, f, \varphi]}{2}$
4. $\dfrac{E[f, \psi, \varphi] - S[f, \psi, \varphi]}{2}$

5. $\dfrac{E[\psi,\varphi,f]-S[\psi,\varphi,f]}{2}$ 6. $\dfrac{E[\varphi,f,\psi]-S[\varphi,f,\psi]}{2}$

**Montrons que** (1) = (4) Pour $E(f,\varphi,\psi)$ et $S(f,\psi,\varphi)$ on a le même enclos $\psi\varphi < 0$. Parcourons $f$ dans le sens positif. $f$ entre autant de fois dans l'enclos qu'il en sort.

Or elle y entre soit par $\varphi$ en $E(f,\varphi,\psi)$, soit par $\psi$ en $S(f,\psi,\varphi)$. En effet, dans le système $(f,\psi,\varphi)$ le sens de parcours de $f$ est négatif ; à un point où $f$ sort de $\psi$ dans le sens négatif correspond un point où $f$ entre dans $\psi$ dans le sens positif.

Le nombre de points d'entrée est donc :

$$E[f,\varphi,\psi]+S[f,\psi,\varphi]$$

On montre de même que le nombre de points de sortie de l'enclos est :

$$E[f,\psi,\varphi]+S[f,\varphi,\psi]$$

Donc

$$E[f,\varphi,\psi]+S[f,\psi,\varphi]=E[f,\psi,\varphi]+S[f,\varphi,\psi]$$

Et on a bien ainsi (1) = (4).
On montrerait exactement de la même façon que (2) = (6) et (3) = (5).

**Montrons que** (1) = (6)

On note $\alpha$ un point d'intersection des deux courbes $f$ et $\varphi$. On considère successivement les deux systèmes $(f,\varphi,\psi)$ et $(\varphi,f,\psi)$.

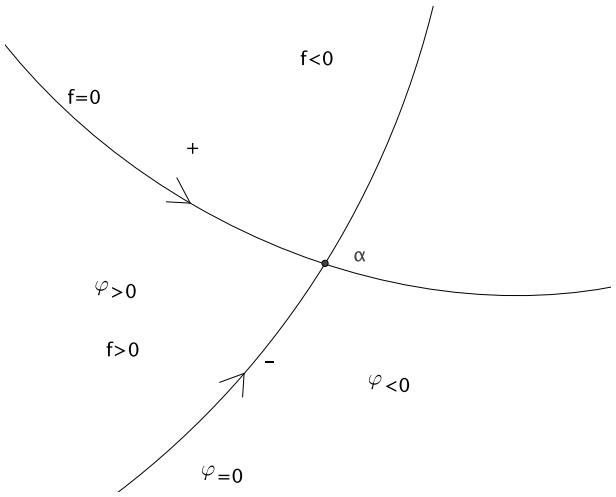



FIGURE 5 –

$(f,\varphi,\psi)$ : on parcourt $f$ dans le sens positif

— Si $\psi(\alpha) > 0$, alors en $\alpha$ on a $\varphi\psi$ qui passe de $> 0$ à $< 0$, donc on entre dans l'enclos, et $\alpha$ est du type $E(f, \varphi, \psi)$.

— Si $\psi(\alpha) < 0$, alors en $\alpha$ on a $\varphi\psi$ qui passe de $< 0$ à $> 0$, donc on sort de l'enclos, et $\alpha$ est du type $S(f, \varphi, \psi)$.

$(\varphi, f, \psi)$ : on parcourt $\varphi$ dans le sens négatif

— Si $\psi(\alpha) > 0$, alors en $\alpha$ on a $f\psi$ qui passe de $> 0$ à $< 0$, donc on entre dans l'enclos, et $\alpha$ est du type $E(\varphi, f, \psi)$.

— Si $\psi(\alpha) < 0$, alors en $\alpha$ on a $f\psi$ qui passe de $< 0$ à $> 0$, donc on sort de l'enclos, et $\alpha$ est du type $S(\varphi, f, \psi)$.

On a donc $E[f, \varphi, \psi] = E[\varphi, f, \psi]$ et $S[f, \varphi, \psi] = S[\varphi, f, \psi]$

Conclusion : on a donc (1) = (6). On montre de même que (2) = (5) et (3) = (4). Donc finalement ces 6 entiers sont égaux, ce qui justifie la définition de la caractéristique d'un système de 3 fonctions à deux inconnues.
Dans la figure 4 on a par exemple la caractéristique qui est égale à : $\frac{1}{2}(1 - 3) = -1$.

Les arguments utilisés ici sont de deux types : (1) étude du signe au point d'intersection ou (2) la somme des entrées et des sorties d'un lacet fermé dans un lacet fermé est constant. Sur ce dernier point, nous reviendrons plus tard. On a donc un argument local, point par point, avec une correspondance directe entre point d'entrée et de sortie, et une correspondance globale sur l'ensemble des points d'entrée et de sortie.

### 12.3.4 Points d'intersection de deux courbes du plan dans un compact

Weber souhaite maintenant caractériser comme précédemment à l'aide d'un calcul de signe le fait qu'un point d'intersection soit un point d'entrée ou de sortie dans le cas de trois fonctions. Il pourra ensuite étendre son travail au cas des courbes qui ne sont plus ni bornées ni fermées.

#### 12.3.4.1 Rapport entre la notation pour deux courbes et celle pour trois courbes

En un point de type $E(\varphi, \psi)$, la courbe $\varphi$ entre dans la courbe fermée $\psi$ (on rappelle que le sens de parcours sur $\varphi$ est positif). En ce point :

— Soit $f$ est positif, et avec un raisonnement analogue à celui du paragraphe précédent on s'aperçoit que $\varphi$ entre en fait dans l'enclos $\psi f < 0$, donc que ce point de type $E(\varphi, \psi)$ correspond à un point du type $E(\varphi, \psi, f)$.

— Soit $f$ est négatif, et donc $\varphi$ sort de l'enclos $\psi f < 0$, donc ce point de type $E(\varphi, \psi)$ correspond à un point du type $S(\varphi, \psi, f)$.

On pourrait faire de même avec un point de type $S(\varphi, \psi)$. On a donc en résumé :

$$E(\varphi, \psi) = \begin{cases} E(\varphi, \psi, f) \text{ si } f > 0 \\ S(\varphi, \psi, f) \text{ si } f < 0 \end{cases}$$

$$S(\varphi, \psi) = \begin{cases} E(\varphi, \psi, f) \text{ si } f < 0 \\ S(\varphi, \psi, f) \text{ si } f > 0 \end{cases}$$

Or $[\varphi, \psi] > 0$ en $S(\varphi, \psi)$ et $[\varphi, \psi] < 0$ en $E(\varphi, \psi)$

Donc un point d'intersection de $\varphi$ et $\psi$ est de type :

1. $E(\varphi, \psi, f)$ si $[\varphi, \psi]f < 0$
2. $S(\varphi, \psi, f)$ si $[\varphi, \psi]f > 0$

On a donc la propriété suivante :

*La caractéristique est la moitié de l'excès du nombre de points d'intersection de $\varphi$ et $\psi$ pour lesquels $[\varphi, \psi]f < 0$ sur celui du nombre d'intersections pour lesquelles $[\varphi, \psi]f > 0$.*

#### 12.3.4.2 Caractère d'un point

On associe[17] à chaque point d'intersection entre les courbes $\varphi = 0$ et $\psi = 0$ l'une des deux valeurs $+1$ et $-1$. Pour cela, on utilise le tableau suivant :

| | $E(\varphi, \psi)$ | $S(\varphi, \psi)$ |
|---|---|---|
| Extérieur de $f$ | $+1$ | -1 |
| Intérieur de $f$ | -1 | $+1$ |

En fait, un point prendra la valeur $+1$ s'il est du type $E(\varphi, \psi, f)$ et $-1$ s'il est du type $S(\varphi, \psi, f)$. Cela nous permet de définir la caractéristique d'un système de 3 fonctions comme la moitié de la somme des caractères.

#### 12.3.4.3 Extension à des cas plus généraux

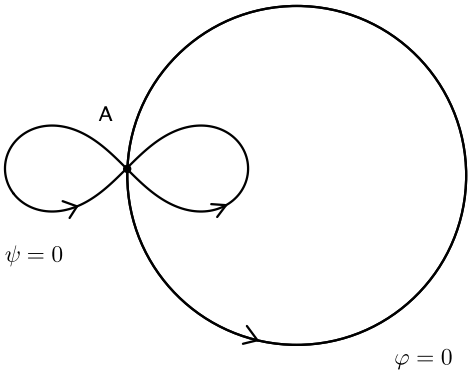
Cas à inclure :

1. Dans le cas où les courbes sont constituées de plusieurs courbes fermées, on applique ce qui précède sans aucun changement.
2. Les points doubles ne sont pas problématiques, sous réserve que l'intérieur et l'extérieur de chaque partie fermée soient de signes opposés, permettant ainsi de déterminer un sens de parcours.

17. Il s'agit ici d'une notation qu'introduit Weber, librement inspirée de Kronecker, mais légèrement différente de ce que celui-ci appelle un caractère.

Cas à exclure :

1. Si une courbe passe par un point double de l'autre :




   Dans l'exemple ci-dessus, le point $A$ est bien un point d'intersection de $\varphi$ et $\psi$, mais ne peut être qualifié de point d'entrée ou de sortie.

2. Si deux courbes sont tangentes :

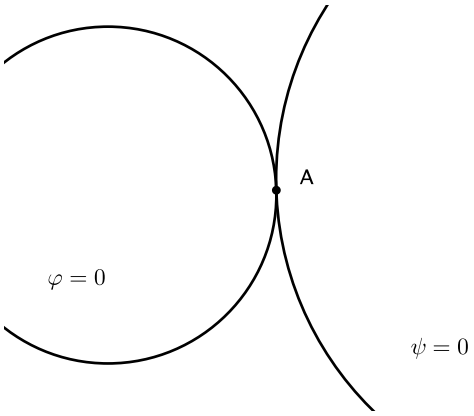



FIGURE 6 –

   Encore une fois le point $A$ est bien un point d'intersection de $\varphi$ et $\psi$, mais de nouveau il ne peut être qualifié de point d'entrée ou de sortie.

3. En un point où les trois courbes s'intersectent.

   On a vu précédemment qu'au point d'intersection de deux courbes, le signe de la troisième joue un rôle déterminant. Ici le fait qu'elle soit nulle provoque une indétermination et ne permet pas de classer ce point comme l'un des six types.

Cas à inclure après modification :

On ne peut appliquer ce qui précède à des courbes ayant des branches infinies qu'après avoir limité arbitrairement ces branches pour en faire des courbes fermées (Figure 12.3.4.3)

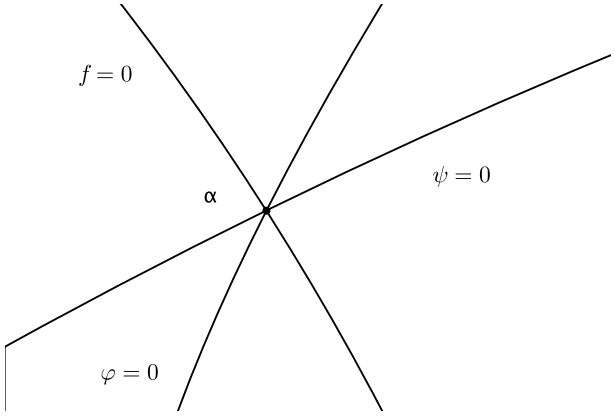



FIGURE 7 –

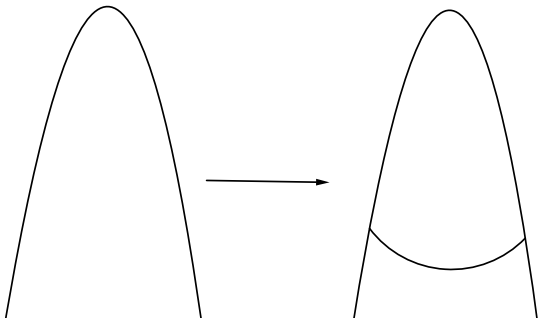


FIGURE 8 –

### 12.3.5 Racines complexes

On pose $z = x + iy \in \mathbb{C}$ et $F(z) = \varphi(x, y) + i\psi(x, y) \in \mathbb{C}[z]$ avec $\varphi(x, y)$ et $\psi(x, y)$ sont des éléments de $\mathbb{R}[x, y]$.

On suppose que $F(z)$ et $F'(z)$ ne s'annulent pas ensemble. On remarque que $F(z)$ s'annule si $\varphi(x, y)$ et $\psi(x, y)$ s'annulent ensemble, et donc aux points où les deux courbes $\varphi$ et $\psi$ s'annulent ensemble.

Comme $F$ est un polynôme en $z$, on a (Cauchy-Riemann) :

$$\varphi'_x = \psi'_y \ , \ \varphi'_y = -\psi'_x$$

et donc le déterminant fonctionnel :

$$[\varphi, \psi] = \varphi'_x\psi'_y - \varphi'_y\psi'_x = (\varphi'_x)^2 + (\varphi'_y)^2 = (\psi'_x)^2 + (\psi'_y)^2.$$

Ainsi $[\varphi, \psi] \geq 0$. De plus, si $[\varphi, \psi] = 0$ en un point d'intersection de $\varphi = 0$ et $\psi = 0$, alors en ce point $F(z) = F'(z) = 0$, ce que l'on a supposé impossible. Finalement, aux points d'intersection de $\varphi = 0$ et $\psi = 0$ $[\varphi, \psi] > 0$. Les courbes n'étant pas fermées, on

détermine, comme on l'a vu précédemment [18], le sens de parcours de la façon suivante : on laisse sur la « gauche » la partie du plan où la fonction est négative.

Ainsi, tous les points d'intersection peuvent être considérés comme des points de sortie : aucune des deux courbes ne peut donc être fermée.

Prenons une troisième courbe $f(x, y) = 0$ supposée cette fois fermée.

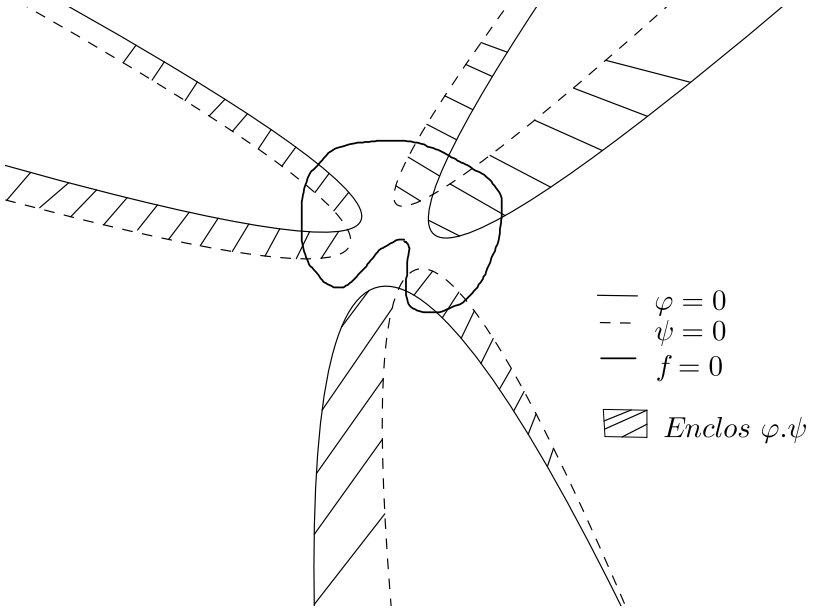



FIGURE 9 –

Maintenant « refermons » les courbes $\varphi = 0$ et $\psi = 0$ par des portions de courbes à *l'extérieur* de $f = 0$.

Mais en faisant cela, on crée des points d'intersection supplémentaires entre les courbes $\varphi = 0$ et $\psi = 0$. Soit $S'[\varphi, \psi]$ et $E'[\varphi, \psi]$ le nombre de points d'entrée et de sortie créés ainsi à l'extérieur de $f = 0$.

Soit maintenant $S_1(\varphi, \psi)$ le nombre de points de sortie initiaux se trouvant *à l'intérieur* de $f = 0$ et $S_2(\varphi, \psi)$ le nombre de points de sortie initiaux se trouvant *à l'extérieur* de $f = 0$. On a :

$$S = S_1 + S_2$$

et comme pour des courbes fermées il doit y avoir autant d'entrées que de sorties :

$$S + S' = E'$$

On peut attribuer à chaque type de point un caractère comme défini précédemment :

$$S_1 : +1$$

18. Voir p. 329.

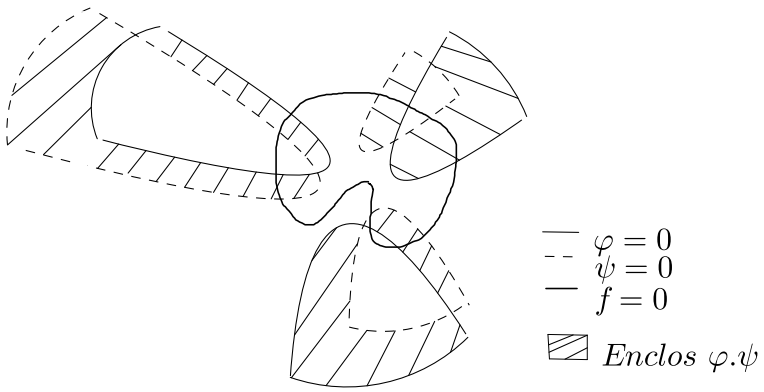



FIGURE 10 –

$$S_2 : -1$$
$$S' : -1$$
$$E' : +1$$

En notant $k$ la caractéristique du système $f, \varphi, \psi$ on a

$$2k = S_1 - S_2 - S' + E'$$

En ajoutant terme à terme avec

$$S + S' = E'$$

on obtient

$$2k = S_1 - S_2 + S = 2S_1$$

donc

$$k = S_1$$

La caractéristique est donc égale au nombre de points de sortie situés à l'intérieur de $f = 0$, c'est-à-dire en fait au nombre de points d'intersection de $\varphi$ et $\psi$ situés à l'intérieur de la courbe fermée $f = 0$.

*La caractéristique permet donc de déterminer le nombre de solutions d'une équation polynômiale $F(z) = 0$ à l'intérieur d'un compact dont le bord est suffisamment régulier.*

Cette façon de voir les racines d'une fonction de la variable complexe comme les points d'intersection de deux courbes réelles planes est le point de départ des première et qua-

trième démonstrations du théorème fondamental de l'algèbre, mais aussi de leur reprise par Kronecker, comme nous le verrons plus loin.

### 12.3.6 Détermination de la caractéristique

Weber décrit maintenant un moyen d'obtenir, pour une fonction $f$ suffisamment régulière, la caractéristique du système $f, \varphi, \psi$ à l'aide du théorème de Sturm.

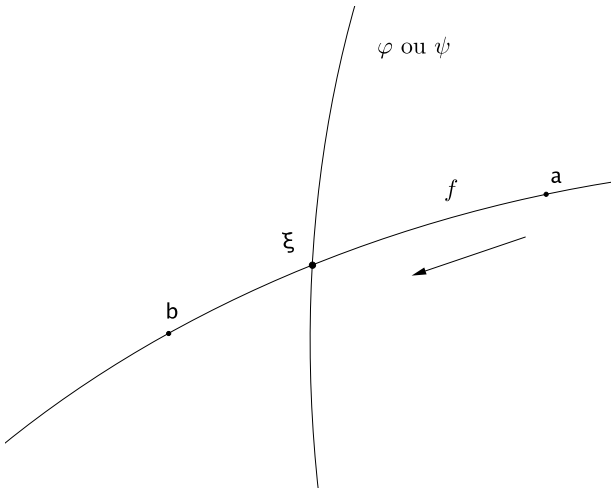



FIGURE 11 –

Supposons que l'on ait pu isoler chaque point d'intersection de $f$ avec $\varphi$ et $\psi$, de sorte qu'en chacun de ces points d'intersection $\xi$ l'on dispose d'un arc de la courbe de $f = 0$ qui ait un et un seul point d'intersection avec $\{\varphi = 0\} \cup \{\psi = 0\}$. Soient $a$ et $b$ l'origine et l'extrémité de cet arc. Prenons l'exemple illustré par la figure 11 : Le signe de $\varphi(a)\psi(a)$ nous permet de déterminer s'il s'agit d'un point de type $E(f, \varphi, \psi)$ ou de type $S(f, \varphi, \psi)$.

$$\begin{cases} E(f, \varphi, \psi) \text{ si } \varphi(a)\psi(a) > 0 \\ S(f, \varphi, \psi) \text{ si } \varphi(a)\psi(a) < 0 \end{cases}$$

Si de plus, on peut paramétriser la courbe $f = 0$, de sorte qu'il existe un intervalle $I$ tel que cette courbe soit l'ensemble : $\{(x(t), y(t)) \,/ t \in I\}$ où $x(t)$ et $y(t)$ sont tels que l'on puisse utiliser le théorème de Sturm pour déterminer les zéros des deux fonctions :

$$\begin{cases} t \mapsto \varphi\left(x(t), y(t)\right) \\ t \mapsto \psi\left(x(t), y(t)\right) \end{cases}$$

Et donc en fait on doit déterminer les zéros de la fonction $\varphi\left(x(t), y(t)\right) \cdot \psi\left(x(t), y(t)\right)$.

## 12.4 Conclusion

Cette dernière remarque, Weber semble l'avoir faite pour relier son exposé au théorème de Sturm. En fait la relation entre les deux théories, celle de Sturm et celle de Kronecker, est bien plus forte que la simple utilisation de l'une dans l'autre. Plus encore, présentée ainsi, on pourrait croire que la théorie des caractéristiques est une conséquence du théorème de Sturm, ou que le théorème de Sturm n'est qu'un artifice calculatoire dans la théorie des caractéristiques. Et, dans une certaine mesure, cela correspond bien à la présentation que Weber en fait. Ce premier exposé, qui facilite notre compréhension de la théorie des caractéristiques et nous permet d'aborder le travail de Kronecker, est donc loin d'être suffisant pour en saisir les enjeux.

# Chapitre 13

# La théorie des caractéristiques selon Kronecker


13.1 Considérations introductives . . . . . **345**
13.2 La caractéristique de deux fonctions d'une variable . . . . . **347**
13.2.1 Introduction . . . . . 347
13.2.2 Définition de la caractéristique de deux fonctions . . . . . 348
13.2.3 Interprétation graphique . . . . . 351
13.2.4 Lien avec le théorème de Sturm . . . . . 352
13.2.5 Modification de la caractéristique en fonction des coefficients de $\varphi$ et $\psi$. . . . . 353
13.3 Le *Fortgangsprincip* . . . . . **359**
13.3.1 Introduction . . . . . 359
13.3.2 Explication du principe . . . . . 362
13.3.3 Conclusion . . . . . 372
13.4 La caractéristique de trois fonctions à deux variables . . . . . **375**
13.4.1 (*Fortgangsprincip*) : un résumé pratique . . . . . 375
13.4.2 Une proposition problématique . . . . . 376
13.4.3 Remarques à propos de l'argument utilisé par Kronecker . . . . 377
13.4.4 Caractéristique de trois fonctions de deux variables . . . . . 380
13.5 Retour sur la notion de caractéristique . . . . . **388**
13.5.1 La notion de racine d'une équation . . . . . 388
13.5.2 La notion de caractéristique . . . . . 390
13.6 Conclusion . . . . . **393**


Dans la présentation précédente, nous venons de voir que la caractéristique d'un système de fonctions est un nombre entier qui donne une information globale sur le système. Lors de l'entretien entre Hilbert et Kronecker dont nous avons déjà parlé [1], ce dernier affirme que [2]

1. Voir p. 321.
2. Die wahren Invarianten seien nicht die « litteralen » (die später vergessen sein würden, wie die

> Les vrais invariants ne sont pas les invariants « littéraux » (qui seront plus tard oubliés, comme les travaux combinatoires sont aujourd'hui oubliés), mais certains nombres (comme le nombre de signes positifs des carrés est un invariant simultané d'une forme quadratique ; Cf. les invariants simultanés de deux formes quadratiques que Kronecker avait développés à partir des travaux les plus difficiles). [Hilbert, 1888]

Dans son cours sur le concept de nombre de 1891, Kronecker définit ce qu'il entend par invariant [3] :

> Y-a-t-il des fonctions des nombres $(a, b, c)$ faites de telle sorte qu'elles soient identiques pour tous les systèmes de la même classe ? Si une telle fonction existe, alors nous la désignerons par le terme de
>
> 'invariant'
>
> choisi par Sylvester de façon si exceptionnellement heureuse, bien que pour un cas très spécial.

La caractéristique d'un système d'équations n'est pas un *invariant*, au sens où beaucoup l'entendent à la fin du XIX$^{e}$ siècle [4], mais s'en rapproche tout de même beaucoup : elle est bien, « fonction » des coefficients des équations, mais cette fonction n'est pas un polynôme. La définition de Kronecker est plus générale, et ne donne pas de condition explicite sur la nature de la fonction en question. Dans la suite de cet essai, Kronecker fait une distinction entre invariants *arithmétiques*, algébriques et *analytiques* [5] :

> Je distingue les invariants entre invariants arithmétiques, algébriques et analytiques selon la méthode avec laquelle on les déduit des éléments d'un système. Chaque forme réduite est par exemple un invariant arithmétique car elle nécessite un algorithme qui ressemble à un développement en fractions continues pour se déduire des éléments du système. Ce sont des invariants algébriques dans le cas spécial où ils sont en effet des fonctions rationnelles des éléments. Dans les invariants analytiques, et cela vaut en fait pour chaque équivalent, il apparaît une limite. [Boniface et Schappacher, 2001] p. 238

---

Combinatorischen Arbeiten heute vergessen sind), sondern gewisse Zahlen (wie die Zahl der positiven Vorzeichen der Quadrate eine gleichzeitige Invariante einer quadratischen Form ist ; vegl. die gleichzeitigen Invarianten zweier quadratischer Formen, die Kronecker von allen Arbeiten die schwerste geworden sei).

3. Giebt es Funktionen der Zahlen $(a, b, c)$, welche so beschaffen sind, daß sie für sämtliche Systeme derselben Klasse identisch sind ? Ist eine solche Funktion vorhanden, so legen wir ihr die von Sylvester so außerordentlich glücklich gewählte, obgleich von ihm nur für einen ganz speziellen Fall gebrauchte Bezeichnung

'Invariante'

zu.

4. Si l'on se reporte par exemple au traité d'algèbre de Weber (Voir [Weber, 1898] p. 226), un invariant est une fonction polynomiale des coefficients d'une ou plusieurs « formes », c'est-à-dire de polynômes à une ou plusieurs inconnues.

5. Ich unterscheide die Invarianten als arithmetische, algebraische und analytische Invarianten nach der Methode, durch welche sie aus den Elementen eines Systems hergeleitet werden. Jede reduzierte Form ist z.B. eine arithmetische Invariante, weil es eines Algorithmus, welcher mit einer Kettenbruchentwicklung Ähnlichkeit hat, bedarf, um sie aus den Elementen der Systeme herzuleiten. Algebraische Invarianten sind in speziellen Fällen rational, wenn sie nämlich rationale Funktionen der Elemente sind. Bei analytischen Invarianten, und solche giebt es eigentlich für jede Äquivalenz, — tritt ein Limes auf.

C'est donc la méthode d'obtention des invariants, et non leur nature, qui permet à Kronecker de les classer. La méthode qui permettra de trouver la caractéristique d'un système d'équations pourra, nous le verrons, selon le point de vue choisi, être qualifiée d'« arithmétique » ou d'« analytique ». Mais gardons à l'esprit que Kronecker lui-même n'a à aucun moment présenté la caractéristique d'un système de fonctions comme un « invariant ».

Pour étudier la théorie des caractéristiques telle que la conçoit Kronecker, nous avons donc à notre disposition les trois articles mentionnés plus haut [6], ainsi que les *Vorlesungen* des années 1872-73, 1878-79, 1880-81, 1888-89 (seulement deux pages) et 1890-91. Nous verrons que la présentation qu'il en donne dans ses articles de recherche est bien trop succincte pour permettre à elle seule une compréhension réelle de la théorie. Cependant, à l'exception du manuscrit de 1878-79, seul le cas $n = 2$ est traité dans ses cours. Nous présenterons donc nous aussi ici en détail ce cas particulier, et nous le prolongerons ensuite au cas $n$. Nous avons choisi de ne pas suivre l'un ou l'autre de ces manuscrits, mais d'essayer d'en donner une présentation thématique en utilisant l'ensemble de nos sources. Cependant, pour faciliter notre approche et ne pas multiplier les notations, nous nous référerons essentiellement au manuscrit de 1880-81 et à celui de 1890-91. Cela nous permettra de plus de faire apparaître les changements qui ont pu s'opérer lors des dix dernières années de l'enseignement de Kronecker. Nous préciserons à chaque fois la source que nous suivrons.

## 13.1 Considérations introductives

Dans les deux derniers manuscrits que nous venons de citer, Kronecker se réfère au théorème de Sturm lorsqu'il commence son chapitre sur la théorie des caractéristiques, en citant ce théorème sous la forme qu'il qualifie de « plus générale » [7] :

$$\sum_{x'<\xi<x''} sgn\ f'(\xi)f_1(\xi) = \frac{1}{2}\sum_{h=1}^{v}(sgn\ f_{h-1}(x'').f_h(x'') - sgn\ f_{h-1}(x') \cdot f_h(x'))$$

Nous avons déjà vu que cette généralisation, Kronecker l'attribue, à raison, à Sylvester. Mais il insiste ici sur le fait que Sylvester n'a pas réussi à donner d'interprétation géométrique [8] correcte de cette formule [9] :

> Cette généralisation de la formule de Sturm vient de *Sylvester*, qui a fait paraître plusieurs essais sur les expressions combinatoires précédentes, expressions qu'il a formées

6. Voir [Kronecker, 1869c], [Kronecker, 1869d] et [Kronecker, 1878b].

7. Voir [Kronecker, 1891a] p. 208.

8. L'interprétation dont parle Kronecker dans la citation suivante est bien géométrique : nous verrons qu'il précise ce point dans [Kronecker, 1869c].

9. Diese Verallgemeinerung der Sturmschen Formel rührt von *Sylvester* her, welchen darüber und über die in vorhergehenden besprechenen kombinatorischen Ausdrücke, die er unter Voraussetzung der Zerlegung einer Funktion als Produkt von Linearfaktoren bildete, einige Aufsätze in den *Philosophical Transactions* z. Teil 1853 erscheinen ließ. Während aber für den Speziellen Fall, wo $f_1(x) = f'(x)$ gesetzt wird, die Bedeutung der Zeichensumme auf der Hand liegt, bemühte sich *Sylvester* vergeblich, diese allgemeineren Zeichenverbindungen, für welche er den Namen « intercallations » einführte, in Klarer und erschöpfend Weise zu interpretieren.

> sous l'hypothèse de la décomposition d'une fonction en facteurs linéaires dans les *Philosophical Transactions* de 1853. Mais alors que pour le cas particulier où $f_1(x) = f'(x)$, la signification de la somme de signes est évidente, *Sylvester* s'est efforcé en vain d'interpréter de façon claire et exhaustive cette relation entre les signes, pour laquelle il a introduit le terme d'« intercalations ». [Kronecker, 1891a] p. 208

Rappelons la définition que Sylvester donne de ces « intercalations » :

> Intercalations – The theory of intercalations is the theory of the relative distribution of the real roots, or point-roots, of two or more equations, but in this theory the number of roots mutually interposed is to be taken only with reference to the number 2 as a modulus. (Définition du glossaire joint par Sylvester dans [Sylvester, 1853d])

Nous sommes donc, pour le moment, toujours dans le cadre de la *recherche du nombre de solutions d'un système d'équations*. Comme nous le verrons, ce cadre n'est que l'un de ceux qui permettent une bonne interprétation de la théorie des caractéristiques.

**Interprétation géométrique** Dans le manuscrit de 1890-91, Kronecker décide de commencer son cours sur cette théorie par une interprétation géométrique de cette formule dans le cas particulier où l'on considère les courbes du plan données par $y = f(x)$ et $y = f_1(x)$. Nous avons déjà vu cette interprétation dans la partie précédente [10] : en se déplaçant le long de l'axe des abscisses entre $x'$ et $x''$ cette formule peut être interprétée « comme le nombre de points de sortie qui se trouvent dans l'intervalle $x'...x''$, diminué du nombre de points d'entrée se trouvant dans le même intervalle » [11] où les points d'entrée et de sortie sont définis comme précédemment. Pour nous remémorer cette interprétation nous pouvons regarder l'exemple que nous propose Kronecker sur la figure 1.

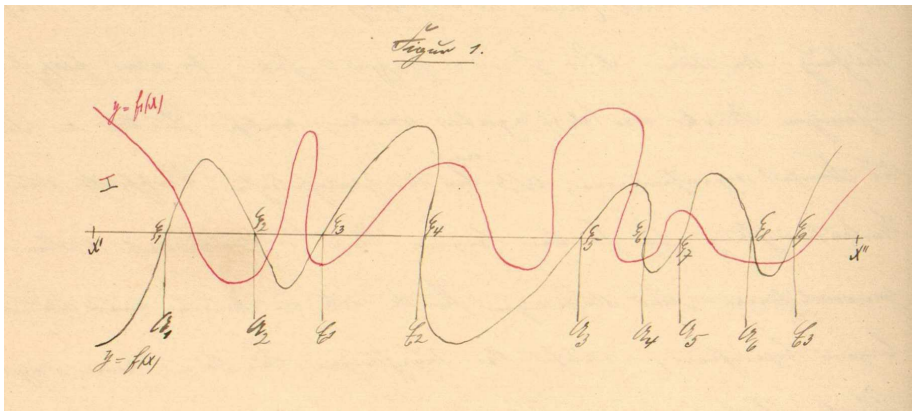


FIGURE 1 – [Kronecker, 1891a] p. 210

Kronecker conclut cette introduction de la manière suivante [12] :

10. Voir p. 216.

11. [Kronecker, 1891a] p. 210 : « als die Anzahl der Austrittspunkte, welche sich in dem Intervall $x'...x''$ befinden, vermindert um die Anzahl der in demselben Intervall liegenden Eintrittspunkte. »

12. Diese Betrachtungen können verallgemeinert werden, indem man anstatt der $x$ =Axe eine beliebige dritte Kurve zu den beiden andere hinzunimmt ; jedoch ist es, um diese Untersuchungen arithmetisch

> Ces remarques peuvent être généralisées en prenant à la place de l'axe des $x$ une troisième courbe en plus des deux autres ; il est pourtant nécessaire, pour poursuivre arithmétiquement ces recherches, d'introduire une nouvelle notion : la notion de caractéristique d'un système de fonctions. [Kronecker, 1891a] p. 211

C'est donc à partir de considérations géométriques qu'il introduit ici son chapitre. Nous trouverons ainsi la généralisation de ce travail lorsque nous étudierons la caractéristique d'un système de trois fonctions de deux variables. Plus encore, la poursuite de ce travail sera « arithmétique », et nous verrons ce que cela signifie pour Kronecker. Mais commençons par examiner comment Kronecker traite le cas de deux fonctions d'une variable.

## 13.2 La caractéristique de deux fonctions d'une variable

### 13.2.1 Introduction

Nous allons commencer par présenter la caractéristique de deux fonctions, telle que Kronecker la définit dans son cours sur la théorie des équations algébriques de 1880-81 et 1890-91. Ce cas n'est en effet pas traité dans l'article de présentation de Kronecker de la théorie des caractéristiques de 1869. Il correspond en fait au cas $n = 1$, auquel nous ne pouvons appliquer directement le travail fait dans ces articles : nous allons donc le traiter à part. Remarquons aussi que Weber ne traite pas ce cas.

La caractéristique de deux fonctions, qui n'est pas présente dans le manuscrit de 1872-73, représente 27 pages dans le manuscrit de 1878-79 (pp 483-510), 9 pages dans le manuscrit de 1880-81 (pp. 155-164), la totalité de celui de 1888-89 (pp. 132-142) et 19 pages dans celui de l'hiver 1890-91 (pp. 211-230).

Il ne faudrait pas s'y tromper : il ne s'agit pas du cas le plus simple de la théorie, dont nous pourrions nous servir pour généraliser à des systèmes possédant plus de fonctions : c'est en effet à partir du cas d'un système de trois fonctions de deux variables que la généralisation au cas $n$ se fait. Ce cas particulier est développé pour deux raisons. Tout d'abord, comme nous venons de le voir, il faut attendre le manuscrit de 1878-79 pour le voir apparaitre. En effet, la première fonction du cas $n = 1$ est d'appliquer les résultats que Kronecker développe la même année dans son article intitulé *Über die Charakteristik von Funktionen-Systemen*[13] sur la modification de caractéristique[14].

Ensuite, ce cas particulier permet, comme nous le verrons, de mettre en évidence le lien entre caractéristique et théorème de Sturm. Nous allons donc suivre le plan suivant pour présenter le contenu de ce que Kronecker fait sur la caractéristique de deux fonctions :

1. Définition de la caractéristique de deux fonctions
2. Interprétation graphique

durchführen zu können, nötig, einen neuen Begriff einzuführen : den Begriff der Charakteristik eines Funktionsystems.

13. Voir [Kronecker, 1878b].
14. Voir plus loin p. 353.

3. Lien avec le théorème de Sturm
4. Stabilité de la caractéristique lors d'une variation continue des coefficients

Dans cette partie, nous prendrons les notations suivantes utilisées par Kronecker dans [Kronecker, 1881a] et [Kronecker, 1891a] :

$$[a] = \begin{cases} 1 \text{ si } a \text{ est positif} \\ -1 \text{ si } a \text{ est négatif} \end{cases}$$

D'autre part, en particulier lorsque nous citerons [Kronecker, 1891a], nous noterons $sgn\ a$ à la place de $[a]$.

### 13.2.2 Définition de la caractéristique de deux fonctions

Revenons sur la somme

$$\sum_{x_1<\xi<x_2} [f_1(\xi)f'(\xi)]$$

qui intervient dans le théorème de Sturm où $f$ et $f_1$ sont deux fonctions polynômiales.

Prenons ici deux fonctions polynômiales $\varphi$ et $\psi$ que l'on supposera toutes les deux de degré $n$. Notons $\xi_1, ..., \xi_n$ les racines de $\varphi$ et $\eta_1, ..., \eta_n$ celles de $\psi$. Soit $F$ une fonction polynôme de degré $n$ : d'après les résultats de la page 207, on a :

$$\sum_{z_1<\xi<z_2} [F'(\xi_i)] = \frac{1}{2}([F(z_2)] - [F(z_1)])$$

Si l'on pose $F(x) = F_0x^n + ...$ et que l'on suppose que $z_1 = -\infty$ et $z_2 = +\infty$, on obtient

$$\sum_{z_1<\xi<z_2} [F'(\xi_i)] = \begin{cases} 0 \text{ si } n \text{ est pair} \\ sgn(F_0) \text{ si } n \text{ est impair} \end{cases}$$

Posons $F(x) = \varphi(x)\psi(x)$, qui est de degré pair $2n$ et dont les racines sont $\xi_1, ..., \xi_n, \eta_1, ..., \eta_n$. On obtient

$$\sum_{z_1<\xi<z_2} [\varphi'(\xi)\psi(\xi)] + \sum_{z_1<\eta<z_2} [\varphi(\eta)\psi'(\eta)] = -\frac{1}{2}[\varphi(z_1)\psi(z_1)] + \frac{1}{2}[\varphi(z_2)\psi(z_2)]$$

Si l'on pose maintenant

$$\Delta(z) = \begin{vmatrix} \varphi(z) & \psi(z) \\ \varphi'(z) & \psi'(z) \end{vmatrix}$$

alors on a $\Delta(z) = \varphi(z)\psi'(z) - \varphi'(z)\psi(z)$, et donc $\Delta(\xi) = -\varphi'(\xi)\psi(\xi)$ et $\Delta(\eta) = \varphi(\eta)\psi'(\eta)$. On a donc finalement :

$$-\sum_{z_1<\xi<z_2} [\Delta(\xi)] + \sum_{z_1<\eta<z_2} [\Delta(\eta)] = -\frac{1}{2}[\varphi(z_1)\psi(z_1)] + \frac{1}{2}[\varphi(z_2)\psi(z_2)].$$

Pour prendre en compte toutes les racines, on pose $z_1 = -\infty$ et $z_2 = +\infty$. $\varphi$ et $\psi$ sont du même degré, donc $\varphi(x)\psi(x)$ est de degré pair, et son signe est le même en $-\infty$ et en

$+\infty$. Ainsi

$$-\frac{1}{2}[\varphi(z_1)\psi(z_1)] + \frac{1}{2}[\varphi(z_2)\psi(z_2)] = 0$$

et donc

$$\sum_{(\xi)}[\varphi'(\xi)\psi(\xi)] + \sum_{(\eta)}[\varphi(\eta)\psi'(\eta)] = 0$$

et

$$-\sum_{\xi}[\Delta(\xi)] + \sum_{\eta}[\Delta(\eta)] = 0.$$

on a ainsi l'égalité

$$\sum_{\xi}[\Delta(\xi)] = \sum_{\eta}[\Delta(\eta)]$$

qui nous permet donner la définition suivante :

#### 13.2.2.1 Définition

Le nombre ou demi-nombre :

$$-\frac{1}{2}\sum_{\xi}[\Delta(\xi)] = -\frac{1}{2}\sum_{\eta}[\Delta(\eta)] = \frac{1}{2}\sum_{z=\xi}[\frac{d(\varphi(x)\psi(x))}{dx}(z)] = -\frac{1}{2}\sum_{z=\eta}[\frac{d(\varphi(x)\psi(x))}{dx}(z)]$$

est appelé *caractéristique* du système de fonctions $\varphi$ et $\psi$. On le note $\chi(\varphi, \psi)$.

**Supposons que $\varphi$ et $\psi$ sont de degré pair**

Dans le cas où $\varphi$ et $\psi$ sont de degré pair, on peut montrer que la caractéristique est un nombre entier.

On a :

$$\sum_{\xi}[\varphi'(\xi)] = \left\{-\frac{1}{2}[\varphi(z_1)] + \frac{1}{2}[\varphi(z_2)]\right\}_{\substack{z_1=-\infty \\ z_2=+\infty}} = 0$$

et

$$\sum_{\eta}[\psi'(\eta)] = \left\{-\frac{1}{2}[\psi(z_1)] + \frac{1}{2}[\psi(z_2)]\right\}_{\substack{z_1=-\infty \\ z_2=+\infty}} = 0$$

On peut alors écrire :

$$\begin{cases} \sum\limits_{\xi\ ,\ \psi(\xi)>0} [\varphi'(\xi)] = -\sum\limits_{\xi\ ,\ \psi(\xi)<0} [\varphi'(\xi)] \\ \sum\limits_{\eta\ ,\ \varphi(\eta)>0} [\psi'(\eta)] = -\sum\limits_{\eta\ ,\ \varphi(\eta)<0} [\varphi'(\eta)] \end{cases}$$

Par ailleurs :

$$\begin{aligned} -\sum[\Delta(\xi)] &= \sum[\varphi'(\xi)\psi(\xi)] = \sum_{\psi(\xi)>0} [\varphi'(\xi)] - \sum_{\psi(\xi)<0} [\varphi'(\xi)] \\ &= 2\sum_{\psi(\xi)>0} [\varphi'(\xi)] = -2\sum_{\psi(\xi)<0} [\varphi'(\xi)]. \end{aligned}$$

et ainsi $-\frac{1}{2}\sum_\xi[\Delta(\xi)] = -\frac{1}{2}\sum_\eta[\Delta(\eta)]$ est un nombre entier.

#### 13.2.2.2 Propriétés

On peut maintenant établir rapidement quelques propriétés.

**Propriété 1 :** $\chi(\varphi,\psi) = -\chi(\psi,\varphi)$.

Cela vient simplement du fait que $\begin{vmatrix} \varphi & \psi \\ \varphi' & \psi' \end{vmatrix} = -\begin{vmatrix} \psi & \varphi \\ \psi' & \varphi' \end{vmatrix}$.

Cette propriété est un cas particulier de :

**Propriété 2 :** Pour tout $\alpha,\beta,\gamma,\delta$, on a

$$\chi(\alpha\varphi+\beta\psi,\gamma\varphi+\delta\psi) = [\alpha\delta-\beta\gamma]\chi(\varphi,\psi)$$

Cela est une conséquence de :

$$\begin{vmatrix} \varphi & \psi \\ \varphi' & \psi' \end{vmatrix} \cdot \begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} = \begin{vmatrix} \alpha\varphi+\beta\psi & \gamma\varphi+\delta\psi \\ \alpha\varphi'+\beta\psi' & \gamma\varphi'+\delta\psi' \end{vmatrix}$$

On a aussi en particulier :

**Propriété 3 :** $\chi(\varphi,\psi) = \chi(\varphi,a\varphi+\psi) = \chi(\varphi+a\psi,\psi)$

Donc par une substitution linéaire seul le signe de la caractéristique peut changer, mais sa valeur absolue reste la même. C'est, dans une certaine mesure, une forme d'*invariant* du système.

#### 13.2.2.3 Retour sur les Hypothèses

Kronecker, dans ses *Vorlesungen*, pose deux conditions sur $\phi$ et $\psi$ : les deux fonctions doivent avoir le même degré, et ce degré doit être pair. Il ne revient pas sur celles-ci, mais prolonge pourtant la définition aux autres cas. Dans son article de 1878 sur le théorème de Sturm [15], Kronecker justifie cette extension : si $\varphi$ et $\psi$ n'ont pas le même degré, avec $\deg(\psi) < \deg(\varphi)$, on pose

$$\chi(\varphi, \psi) = \chi(\varphi, \varphi + \psi)$$

Dans le cas où les fonctions $\varphi$ et $\psi$ sont de degré impair, on peut toujours ajouter des facteurs $x - u$ et $x - v$ en choisissant $u$ et $v$ suffisamment grands.

### 13.2.3 Interprétation graphique

Kronecker passe ensuite, tout au moins dans le manuscrit de 1880-81, à une interprétation graphique de la caractéristique de deux fonctions. Il revient sur le travail qu'il a déjà effectué [16], où il avait montré que la somme

$$\sum_{x_1 < \xi < x_2} [f_1(\xi) f'(\xi)]$$

représente la différence entre le nombre de points de sortie et le nombre de points d'entrée, tels que ceux-ci ont été définis précédemment (voir p. 217) entre les bornes $x_1$ et $x_2$. On a alors [17] :

> la caractéristique a la signification géométrique suivante : La caractéristique $\chi(\varphi, \psi)$ vaut la moitié de l'excédent des sorties sur les entrées lorsque nous nous déplaçons sur la courbe $y = \varphi(z)$, c'est-à-dire en prenant en compte les $\xi$ et vaut de même la moitié de l'excédent des entrées sur les sorties lorsque l'on parcourt la courbe $y = \psi(z)$, c'est-à-dire en prenant en compte les $\eta$. [Kronecker, 1881a] p. 157

**Exemple :** Dans le manuscrit de 1880-81, on a la figure 2 où les « - » sont les entrées et les « + » sont les sorties. On a donc ici

$$\chi(\varphi, \psi) = \frac{1}{2}(2 - 4) = -1$$

15. Voir [Kronecker, 1878c] p. 59.

16. Voir le §45 p. 110 de [Kronecker, 1881a] sur l'extension par Sylvester du Théorème de Sturm et notre chapitre sur ce sujet dans la partie précédente p. 217.

17. Die Charakteristik folgende geometrische Bedeutung [hat] : Es ist die Charakteristik $\chi(\varphi, \psi)$ der halbe Ueberschuß der Austritte über die Eintritte, wenn wir uns auf der Curve $y = \varphi(z)$ bewegen, d.h. die $\xi$ zählen und gleich dem halben Ueberschuß der Eintritte über die Austritte, wenn wir uns auf der Curve $y = \psi(z)$ bewegen, d.h. die $\eta$ zählen.

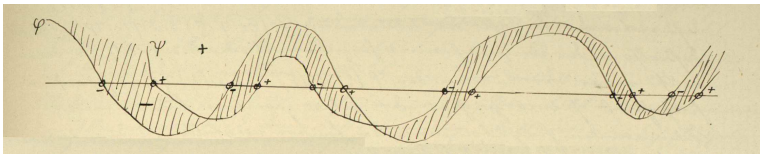


FIGURE 2 – [Kronecker, 1881a] p. 157

### 13.2.4 Lien avec le théorème de Sturm

On pose maintenant $\varphi(z) = f(z)$ et $\psi(z) = (x - z)f_1(z)$. On a alors

$$\chi(f(z), (x-z)f_1(z)) = \frac{1}{2}\sum_{\xi}[(x-\xi)f_1(\xi).f'(\xi)]$$

On obtient

$$\begin{aligned}
&\chi(f(z), (x_2-z)f_1(z)) - \chi(f(z), (x_1-z)f_1(z)) \\
=&\frac{1}{2}\sum_{\xi}[(x_2-\xi)f_1(\xi).f'(\xi)] - \frac{1}{2}\sum_{\xi}[(x_1-\xi)f_1(\xi).f'(\xi)] \\
=&\sum_{x_1<\xi<x_2}[f_1(\xi).f'(\xi)]
\end{aligned}$$

car :

— si $\xi \in [x_1; x_2]$ alors $(x_2 - \xi)$ et $(x_1 - \xi)$ sont de signes opposés et on a

$$\frac{1}{2}[(x_2-\xi)f_1(\xi).f'(\xi)] - \frac{1}{2}[(x_1-\xi)f_1(\xi).f'(\xi)] = [f_1(\xi).f'(\xi)]$$

— si $\xi \notin [x_1; x_2]$ alors $(x_2 - \xi)$ et $(x_1 - \xi)$ sont de même signe et on a

$$\frac{1}{2}[(x_2-\xi)f_1(\xi).f'(\xi)] - \frac{1}{2}[(x_1-\xi)f_1(\xi).f'(\xi)] = 0$$

On peut exprimer à l'aide de la caractéristique de deux fonctions et avec les bonnes conditions sur $f_1$, le nombre de racines de l'équation $f(x) = 0$ situées entre $x_1$ et $x_2$ : nous avons bien là le théorème de Sturm. Le lien décrit ici entre Sturm et la théorie des caractéristiques est beaucoup plus fort que celui que Weber a exposé. Le théorème de Sturm n'est plus un simple outil calculatoire que l'on applique *en plus* de la théorie des caractéristiques. Au contraire il est compris dans la cette dernière et n'en est en fait qu'un cas particulier. On a ici le premier exemple de ce que souhaite réellement faire Kronecker avec sa théorie, c'est-à-dire *englober*, *généraliser* d'autres travaux sur le même sujet.

### 13.2.5 Modification de la caractéristique en fonction des coefficients de $\varphi$ et $\psi$.

#### 13.2.5.1 Cas particulier ou cas général ?

Nous allons aborder maintenant l'une des propriétés fondamentales de la caractéristique d'un système de deux fonctions : la façon dont cette dernière se modifie lors d'une variation continue des coefficients des deux fonctions. Kronecker traite de cette modification pour un système de $n$ fonctions dans son article de 1878. Dans celui-ci, il montre d'ailleurs que [18]

> la modification de la caractéristique lors du passage d'un système de fonctions à un autre et donc lorsque seulement la caractéristique d'un seul système de fonctions est connue, permet de déterminer la caractéristique de tous les systèmes de fonctions correspondant aux différents points. En rendant la modification dépendante d'une seule variable $x$, on ramène cette détermination à l'évaluation de la caractéristique d'un système de deux fonctions d'une variable qui peut s'effectuer dans le cas des fonctions algébriques au moyen du procédé de Sturm. [Kronecker, 1878b] p. 77

Il ne s'agit donc pas d'un cas particulier : on peut toujours se ramener au cas de la caractéristique d'un système de deux fonctions d'une variable. Le contexte de la rédaction de cet article est donné par Kronecker lui-même en introduction [19] :

> À la suite des recherches que j'ai évoquées dans mon exposé d'il y a huit jours, je suis arrivé à la découverte d'une nouvelle propriété fondamentale de la caractéristique d'un système de fonctions de plusieurs variables que j'ai introduite dans ma communication du 4 mars 1869 et que j'ai alors exprimée par une intégrale multiple. La nouvelle propriété que je veux expliquer ici est obtenue par la variation du système de fonctions et peut être utilisée convenablement pour la détermination numérique de la caractéristique. [Kronecker, 1878b] p. 73

L'exposé auquel Kronecker fait allusion s'intitule *Über Sturm'sche Functionen* [20]. Il s'agit d'un long article qui complète celui [21] qu'il a écrit en 1873, et dans lequel il expose la plupart des méthodes que nous avons décrites dans la seconde partie de notre travail, ainsi que la formulation de ce qu'il appelle le *problème d'interpolation de Cauchy* [22]. Cet article se termine par la description du lien qui existe entre la caractéristique de deux fonctions et le théorème de Sturm tel que nous venons de le voir et une application aux équations

18. Hiernach lässt sich die Veränderung der Charakteristik beim Uebergang von einem Functionen-System zum andern und also, wenn nur die Charakteristik eines einzigen Functionen-Systems bekannt ist, die Charakteristik aller den verschiedenen Punkten entsprechenden Functiouen-Systeme bestimmen. Diese Bestimmung findet sich dabei, indem der Uebergang von einer einzigen Variabein $x$ abhängig gemacht wird, auf die Ermittelung der Charakteristik eines Systems von zwei Functionen einer Variabeln zurückgeführt, welche im Falle algebraischer Functionen mittels des *Sturm*'schen Verfahrens erfolgen kann.

19. Im Verfolg der Untersuchungen, welche ich in meinem vor acht Tagen gehaltenen Vortrage erwähnt habe, bin ich zur Auffindung einer neuen Fundamental-Eigenschaft jener Charakteristik der Systeme von Functionen mehrer Variabeln gelangt, welche ich in meiner Mittheilung vom 4. März 1869 eingeführt und dort durch ein vielfaches Integral ausgedrückt habe. Die neue Eigenschaft, welche ich hier auseinandersetzen will, wird durch Variirung der Functionen-Systeme erlangt und kann füglich zur numerischen Bestimmung der Charakteristik benutzt werden.

20. Voir [Kronecker, 1878c].

21. Voir [Kronecker, 1873b].

22. Voir p. 253.

biquadratiques : nous sommes donc toujours dans des travaux issus du « cycle des idées sturmiennes ».

Rappelons que la caractéristique de $\varphi$ et $\psi$ est égale à :

$$\chi(\varphi,\psi)=\frac{1}{2}\sum_{\xi}[\varphi'(\xi)\psi(\xi)]=-\frac{1}{2}\sum_{\eta}[\varphi(\eta)\psi'(\eta)].$$

Dans le paragraphe précédent, nous avons regardé comment se modifie la caractéristique $\chi(f(z),(x-z)f_1(z))$ lorsque $x$ évolue continûment. Ce que Kronecker souhaite faire, c'est examiner ces modifications lorsque l'ensemble des coefficients de $\varphi$ et $\psi$ varie.

On remarque que pour que la caractéristique change, il faut que $\Delta$ change de signe, et donc il faut que pour un $\eta$ ou un $\xi$ on ait

$$\Delta(x)=\varphi'(x)\psi(x)-\varphi(x)\psi'(x)=0.$$

Pour cela, il faut que l'une des conditions suivantes soit remplie :

1. $\varphi(x)\psi'(x)-\varphi'(x)\psi(x)=0$ et $\varphi(x)=0$
2. $\varphi(x)\psi'(x)-\varphi'(x)\psi(x)=0$ et $\psi(x)=0$

Kronecker, après un examen rigoureux de ces deux cas, en déduit le résultat suivant [23] :

> La valeur de la caractéristique est modifiée lors d'une variation continue du système des coefficients de ses fonctions si et seulement si le résultant des deux fonctions est nul. [Kronecker, 1891a] p. 215

Comme $\chi(\varphi,\psi)=-\frac{1}{2}\sum_{\xi}[\Delta(\xi)]$, cette modification est d'une unité et correspond à la disparition d'un terme

$$\frac{1}{2}[\varphi'(\xi)\psi(\xi)].$$

Nous noterons $R(\varphi,\psi)$ le résultant de $\varphi$ et $\psi$. Kronecker applique ensuite ses travaux sur le résultant, et en particulier ceux qu'il a développés dans le cadre du théorème de Sturm. Kronecker montre donc que $\chi(\varphi,\psi)$ ne change que lorsque $R(\varphi,\psi)=0$. Il s'intéresse alors à la somme :

$$\sum_{\xi}\frac{1}{\varphi'(\xi)\psi(\xi)}$$

dont l'un des termes, au voisinage de la valeur qui annule le résultant, tendra vers l'infini [24] :

> L'expression passera de $+\infty$ à $-\infty$ ou inversement selon que le terme en question $\varphi'(\xi)\psi(\xi)$ passera du positif au négatif en passant par le nul ou inversement. [Kronecker, 1881a] pp. 159-160

23. Die Charakteristik ändert ihren Wert bei einer stetigen Variirung des Koeffizientensystems ihrer Funktionen dann und nur dann, wenn die Resultante der beiden Funktionen verschwindet.

24. Der Ausdruck wird von $+\infty$ zu $-\infty$ oder umgekehrt übergehen, je nachdem der betreffende Term $\varphi'(\xi)\psi(\xi)$ von positiven durch die Null in negative oder umgekehrt geht.

Cette façon d'appréhender les choses, c'est-à-dire de regarder et de compter les « sauts » de $-\infty$ à $+\infty$ ou inversement, n'est pas sans rappeler la façon dont Cauchy pratique son calcul d'indice, lorsqu'il cherche à le faire sans l'aide du calcul intégral [25]. Nous verrons plus tard le lien étroit qui existe entre l'indice de Cauchy et la caractéristique de Kronecker.

#### 13.2.5.2 À propos du résultant de deux fonctions

Nous l'avons vu, le résultant tient une place importante dans les leçons de Kronecker, et son rôle est essentiel lorsque l'on examine la modification de la caractéristique. Avant de poursuivre, nous allons nous arrêter sur les résultats que Kronecker démontre sur le résultant. Pour cela, nous allons suivre le court chapitre qu'il y consacre dans son manuscrit de 1890-91 et qu'il intitule *Sur le résultant de deux fonctions entières* [26]. Dans cette sous-section, nous admettrons la plupart des résultats, le but étant d'exposer les formules et les notations dont Kronecker se servira par la suite.

Soit deux fonctions

$$F(x) = x^n - f_1 x^{n-1} + ... \pm f_n = \prod_{k=1}^{n} (x - x_k)$$

et

$$G(x) = x^{n'} - g_1 x^{n'-1} + ... \pm g_{n'} = \prod_{k=1}^{n'} (x - y_k).$$

On pose [27]

$$R_1(\varphi, \psi) = \sum_{\xi} \frac{R(\varphi, \psi)}{\varphi'(\xi)\psi(\xi)}$$

$$S_m = \sum_{i=1}^{n} x_i^m \frac{G(x_i)}{F'(x_i)}$$

et

$$T_{nn} = |S_{p+q}|_{p,q=0,1,...,n-1}.$$

On a alors [28] :

$$T_{nn} = \frac{\prod_{i=1}^{n} G(x_i) \prod_{k \neq i} (x_i - x_k)^2}{\prod_{i=1}^{n} F'(x_i)}$$

25. Voir par exemple le résumé de son exposé fait à Turin en 1831 ([Cauchy, 1831b] p. 171) ou son article sur le *calcul des indices des fonctions* dans le *Journal de l'école polytechnique* ([Cauchy, 1837a] p. 177).

26. [Kronecker, 1891a] pp. 202-208 : « über die Resultante zweier ganzen Funktionen ».

27. Selon les années, Kronecker a déjà défini ces expressions.

28. Kronecker utilise des résultats démontrés dans les chapitres précédents, et en particulier ce qu'il a appelé la formule d'Euler (voir p. 143).

$$
\begin{aligned}
&=(-1)^{\frac{n(n-1)}{2}}\prod_{i=1}^{n}G(x_i)\\
&=(-1)^{\frac{n(n-1)}{2}}\prod_{1\leq i,r\leq n}(x_i-y_r)
\end{aligned}
$$

Le résultant de $F$ et $G$, qui est, comme le précise Kronecker, symétrique par rapport à $x_1, ..., x_n$ et $y_1, ..., y'_n$ et rationnel par rapport à $f_1, ..., f_n, g_1, ..., g_{n'}$ est alors défini par :

$$
\begin{aligned}
R(F,G) =&R(f_1, ..., f_n, g_1, ..., g_{n'}) =\\
&=\prod_{i=1}^{n}G(x_i)=(-1)^{n\cdot n'}\prod_{i=1}^{n'}F(y_i)\\
&=\prod_{i=1}^{n}\prod_{r=1}^{n'}(x_i-y_r)\\
&=(-1)^{n\cdot n'}R(G,F)\\
&=(-1)^{\frac{n(n-1)}{2}}T_{nn}
\end{aligned}
$$

Une condition nécessaire et suffisante pour que les deux polynômes aient une racine commune est donc que le résultant de ces deux polynômes soit nul. Mais, remarque Kronecker, nous avons alors supposé que l'on pouvait écrire ces polynômes comme produit de facteurs linéaires. Il redémontre donc cette propriété sans l'aide de cette hypothèse et en utilisant les congruences modulo plusieurs fonctions. Pour cette démonstration, il crédite Gauss et sa seconde démonstration du théorème fondamental de l'algèbre. Nous verrons plus loin que ce dernier souhaite lui aussi modifier sa démonstration de cette propriété pour éviter de supposer l'existence de facteurs linéaires.

#### 13.2.5.3 Modification de la caractéristique

Ces propriétés sur le résultant étant établies, Kronecker se pose la question suivante : si les coefficients des fonctions $\varphi$ et $\psi$ varient continûment, comment se comporte la caractéristique ?

Pour répondre à cette question, il établit le cadre suivant : il étudie le cas où les $2(n+1)$ coefficients de $\varphi$ et $\psi$ « sont dépendants des $\nu$ variables $x_1, x_2, ...x_\nu$, lesquelles sont mutuellement indépendantes et forment ainsi une variété de dimension $\nu$. »[29] où $\nu$ est un entier choisi arbitrairement.

29. [Kronecker, 1891a] pp. 218-219 : *Denken wir uns jetzt die* $2(n+1)$ *Koeffizienten* $v_0...v_n; w_0....w_n$ *abhängig von insgesamt* $\nu$ *Variabeln* $x_1, x_2, ...x_\nu$, *welche von einander unabhängig sind und daher eine* $\nu$*-fache Mannigfaltigkeit bilden.*

Dans le cours de 1880-81, Kronecker commence par traiter le cas où $\nu = 1$, c'est-à-dire où la variété est de dimension 1, et donc le cas où les coefficients ne dépendent que d'une variable $x$. $R$ et $R_1$ sont alors des fonctions entières de $x$ et il montre que le changement de la caractéristique est alors égal à

$$\chi(\varphi^{(2)}, \psi^{(2)}) - \chi(\varphi^{(1)}, \psi^{(1)}) = \sum_{R(x)=0\ ,\ x_1<x<x_2} [R_1(x)R'(x)]$$

où $\chi(\varphi^{(1)}, \psi^{(1)})$ correspond au système $\chi(\varphi, \psi)$ pour $x = x_1$ et $\chi(\varphi^{(2)}, \psi^{(2)})$ correspond au système $\chi(\varphi, \psi)$ pour $x = x_2$

Nous retombons sur la généralisation du théorème de Sturm, et nous pourrions donc déterminer cette somme comme nous l'avons fait dans la seconde partie, par exemple à partir de la décomposition en fraction continue de $R$ et $R_1$.

Si maintenant nous supposons que les coefficients dépendent de $\nu$ variables $x_1, x_2, ...x_\nu$, alors $R = 0$ est une sous-variété d'ordre $\nu - 1$, et on a de même :

$$\chi(\varphi^{(2)}, \psi^{(2)}) - \chi(\varphi^{(1)}, \psi^{(1)}) = \sum_{R(x_1,....,x_\nu)=0} [R_1(x_1, ...., x_\nu)\delta R(x_1, ...., x_\nu)]$$

Comme nous l'avons fait précédemment dans le cas du plan [30], on définit des points d'*entrée* et de *sortie*. Pour cela, on considère la sous-variété de dimension $\nu - 1$ définie par $R \cdot R_1 = 0$. Kronecker remarque que cette sous-variété « sépare notre variété de dimension $\nu$ en un domaine positif et un domaine négatif, et nous pouvons avec l'utilisation de la notion de point de sortie et d'entrée pour la variété $R \cdot R_1 = 0$ exprimer la formule [précédente] de la manière suivante » [31] :

> La caractéristique ajoute sur le trajet d'un point à un autre de la variété de dimension $\nu$ à un autre autant d'unités que le nombre des points de sortie surpasse celui des points d'entrée. [Kronecker, 1891a] p. 219

Pour que la caractéristique change, il faut, lorsque nous nous déplaçons d'un point $(x_1, ..., x_\nu)$ de notre variété à un autre point $(x'_1, ..., x'_\nu)$, que nous traversions la sous-variété $R = 0$.

30. Voir p. 217 et p. 351.

31. Betrachten wir nun die $(\nu - 1)$-fache Mannigfaltigkeit $R \cdot R_1 = 0$. Dieselbe teilt unsern $\nu$-fache Mannigfaltigkeit in positive und negative Gebiete, und wir können mit Benutzung der Begriffe der Austrittspunkte und Eintrittspunkte für die Mannigfaltigkeit $R \cdot R_1 = 0$ die Formel 15.), welche sich auch in der Form

$$\chi(v''_0, w''_0) - \chi(v'_0, w'_0) = \sum_{(h_1)} sgn\ \delta[R(v_k^{(h_1)}, w_k^{(h_1)}).\delta R_1(v_k^{(h_1)}, w_k^{(h_1)})]$$

schreiben läßt, folgendermaßen drücken :

Die Charakteristik nimmt auf dem Wege von einem Punkte der $\nu$-fachen Mannigfaltigkeit zu einem andern nur so viel Einheiten zu, als die Anzahl der Austrittspunkte die Anzahl der Eintrittspunkte übertrifft.

#### 13.2.5.4 Application au discriminant

Kronecker applique ensuite ces résultats aux fonctions $\varphi$ et $\varphi + \varphi'$ [32] pour étudier le nombre de racines de l'équation $\varphi = 0$. Le résultant sera dans ces conditions, à une constante près, le discriminant $D$ de $\varphi$. Dans les mêmes conditions que précédemment, c'est-à-dire en considérant que $\varphi(x)$ dépend de $\nu$ variables indépendantes, c'est donc la variété $D = 0$ de dimension $\nu - 1$ qui sépare « les domaines dans lesquels se trouvent les équations avec un nombre différent de racines réelles et ce de sorte que l'on doive traverser inévitablement le discriminant pour aller d'une équation à une autre qui n'a pas le même nombre de racines réelles que la première » [33]. Cette variété, il la nomme variété-discriminante (*Diskriminanten-Mannigfaltigkeit*) ou surface-discriminante (*Diskriminantenfläche*), et nous l'avons déjà rencontrée lors de notre étude de l'application du théorème de Sturm à l'équation biquadratique [34].

La place que réserve d'ailleurs Kronecker à l'étude de l'équation de degré 4 dépend des manuscrits, et celle-ci se trouve soit dans le chapitre consacré à Sturm, soit dans celui sur la théorie des caractéristiques. Dans son dernier cours, Kronecker la propose comme application de ce qui vient d'être fait. L'équation biquadratique se réduisant à

$$\varphi(x) = u^4 - 6zu^2 + 4yu + 3x = 0$$

cela lui permet de faire correspondre une équation à un point de l'espace de coordonnées $(x, y, z)$. Kronecker obtient ainsi une approche *géométrique* de l'exemple détaillé qu'il donne dans son cours, mais aussi dans son article *Über Sturm'sche Funktionen* [35]. C'est dans cet article que Kronecker présente la modification de la caractéristique de deux fonctions d'une variable ; l'article de la même année *Über die Charakteristik von Functionen-Systemen* [36] étend l'étude de la modification de la caractéristique à un système de $n + 1$ fonctions de $n$ variables et il est directement inspiré du précédent.

Kronecker montre, dans cet exemple, mais aussi dans le cas général, que [37]

> les domaines auxquels appartiennent les équations ayant le même nombre de racines réelles sont séparés les uns des autres par la variété-discriminante ; dans la suite nous allons montrer que l'on peut aller continument d'une équation à une autre qui a le

32. On se souvient qu'il faut que $\varphi$ et $\psi$ soient de même degré.

33. [Kronecker, 1891a] p. 220 : *diejenigen Gebiete von einander, in welchen die Gleichungen mit verschiedener Anzahl der reellen Wurzeln liegen und zwar so, daß man, um von einer Gleichung zu einer andern zu kommen, welche nicht dieselbe Anzahl reeller Wurzeln hat wie die erste, unumgänglich die Diskriminante durchschreiten muß.*

34. Voir p. 221.

35. Voir [Kronecker, 1878c].

36. Voir [Kronecker, 1878b].

37. die Gebiete, welchen die Gleichungen mit derselben Anzahl reeller Wurzeln zugehören, durch die Diskriminanten-Mannig. von einander getrennt sind ; im folgenden werden wir zeigen, daß man in stetigem Übergang von jeder Gleichung zu jeder andern mit derselben Anzahl reeller Wurzeln gelangen kann, ohne die Diskriminanten-Mannigf. zu passieren ; m.a.W dass also diese Gebiete auch zusammenhängend sind.

> même nombre de racines réelles sans passer par la variété-discriminante ; c'est-à-dire que ces domaines sont donc aussi connexes. [Kronecker, 1891a] p. 222

Kronecker montre en fait ici que par des déformations continues, on peut aller d'une équation du quatrième degré à une autre, sous réserve qu'elles aient le même nombre de racines réelles.

#### 13.2.5.5 Conclusion

À partir de renseignements sur les racines d'une famille d'équations, Kronecker a donc exhibé les composantes connexes d'un espace associé (dont la dimension dépend en partie du degré de l'équation) : une étude *topologique* d'un espace est ainsi associée à une étude algébrique d'une famille d'équations. Avec cette première étude du travail de Kronecker sur la notion de caractéristique, nous commençons à voir apparaître une théorie qui fait intervenir, dans le cadre de travaux algébriques, des éléments que nous qualifierions aujourd'hui de topologique. La caractéristique de deux fonctions réapparaitra lorsque nous étudierons, plus loin, un chapitre que Kronecker ajoute en conclusion de sa partie sur la théorie des équations algébriques en 1891. Ce dernier concernera la *notion de caractéristique*, notion qu'il tentera d'expurger de tout usage des racines d'une équation. Nous allons maintenant passer à ce qui constitue le cœur de ce chapitre : la théorie des caractéristiques de 3 fonctions de 2 variables.

## 13.3 Le *Fortgangsprincip*

### 13.3.1 Introduction

Lorsque l'on commence à lire la partie du cours de Kronecker sur la caractéristique de 3 fonctions de 2 variables, la première notion qui apparaît est celle de *Fortgangsprincip*. Nous pouvons traduire le terme de « *Fortgangsprincip* » ou « *Fortgangsprinzip* » par *principe de parcours*, même si la traduction littérale serait plutôt *principe de continuation*. Il ne faudra pas oublier cependant l'aspect local de « continuation », au sens où ce que l'on définit devra répondre à la question : « comment poursuivre mon chemin lorsque je suis en un certain point d'une courbe ? ». Bien sûr, cela correspondra la plupart du temps à un sens de parcours global sur la courbe, mais pour certains cas limites ce sera cet aspect local sur lequel il faudra se focaliser. Car nous verrons qu'il s'agit bien simplement de cela : dans quel sens doit-on parcourir une courbe ?

#### 13.3.1.1 Dans quel cadre ? En quoi consiste-il ?

Essayons déjà d'avoir une première idée de ce principe. Son cadre d'application est le suivant : à partir de $n+1$ fonctions $F_0, F_1, ..., F_n$ de $n$ variables à valeurs réelles suffisamment régulières[38] on peut former $\frac{n(n+1)}{2}$ sous-

38. Dans le seul cours où Kronecker traite le cas général $n$, il nomme les fonctions $F_{00}$,...,$F_{n0}$ et il suppose implicitement qu'elles admettent des dérivées partielles (Voir [Kronecker, 1879] p. 466).

variétés de dimension 1. En effet l'ensemble des points de $(z_1, ..., z_n)$ tels que $F_0 = F_1 = ... = F_{h-1} = F_{h+1} = ... = F_{k-1} = F_{k+1} = ... = F_n = 0$ est une sous-variété de dimension 1 que Kronecker qualifie de suite continue de points (*stetige Folge*) ou de ligne (*Linie*). C'est sur ces lignes qu'il définit son *Fortgangsprincip*, son *sens de parcours*.

Ainsi, dans le cas $n = 2$, on a la figure 3 où l'on cherche à déterminer un sens de parcours sur $F_0 = 0$, $F_1 = 0$ et $F_2 = 0$.

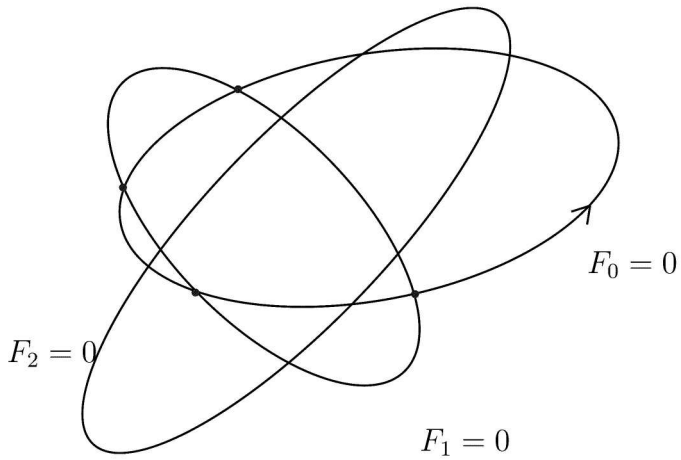



FIGURE 3 –

#### 13.3.1.2 Un concept « fondamental », mais peu présent

Ce terme apparaît dans les écrits de Kronecker pour la première fois dans son article de mars 1869 *Über Systeme von Functionen mehrer Variabeln*, où il introduit la théorie des caractéristiques. À son propos Kronecker affirme que [39]

> Le "principe de parcours" expliqué ici est le véritable fondement de mes recherches sur les systèmes de fonctions de plusieurs variables. [Kronecker, 1869c] p. 178

On ne trouve pourtant ce « principe fondamental » que quatre fois dans son œuvre publiée : trois fois dans [Kronecker, 1869c] et une fois dans [Kronecker, 1878b], et donc uniquement dans deux des quatre articles où il présente la théorie des caractéristiques. Par ailleurs, tout au moins sous le terme de *Fortgangsprincip*, ce principe n'a quasiment pas été repris par ses contemporains. À part dans un article de Paul Stäckel [Stäckel, 1893] et dans [Landsberg, 1911] nous n'avons trouvé aucun document le mentionnant. Par ailleurs, ce principe est cité dans un article de Yvon Gauthier [Gauthier, 2009] et [Gauthier, 2002], mais celui-ci n'en donne aucune description. Il apparaît aussi quelquefois, mais sans être

39. Das hier auseinandergesetzte „Fortgangsprincip" bildet die eigentliche Grundlage meiner Untersuchungen über Systeme von Functionen mehrer Variabeln.

nommé, dans la présentation de la théorie des caractéristiques (par exemple dans le cours d'algèbre de Netto [Netto, 1900]).

Bien que nous ayons décidé, pour les raisons que nous avons vues, de présenter essentiellement cette théorie dans le cas de trois fonctions de deux variables, nous allons exposer ce principe dans le cadre plus général de $n+1$ fonctions à $n$ variables. Pour cela, nous utiliserons l'article de 1869 [Kronecker, 1869c], les trois cours donnés à Berlin en 1878-1879, 1880-1881 et 1890-1891 sur la théorie des équations algébriques ([Kronecker, 1879], [Kronecker, 1881a] et [Kronecker, 1891a]) et les rares commentaires qui portent sur celui-ci ([Stäckel, 1893], [Netto, 1900] et [Landsberg, 1911]).

#### 13.3.1.3 Les raisons de l'introduction de ce principe

Examinons ce qui amène Kronecker à introduire ce principe. Comme Gauss [40] avant lui, lorsque Kronecker doit parcourir une courbe dans le plan, il détermine son sens de parcours comme étant celui qui laisse l'« intérieur » à gauche. Mais, comme il le remarque dans ses cours à l'université de Berlin, le passage à des dimensions supérieures devient problématique. Il écrit ainsi [41] en 1881 :

> En règle générale on parcourt ainsi une courbe de sorte que l'intérieur reste à gauche, mais cette considération se limite aux figures planes. Nous voulons ainsi établir un critère purement analytique.[Kronecker, 1881a] p. 164

et dans son dernier cours de 1891, il ajoute [42]

> Cette détermination de la direction – nous l'appellerons le principe de parcours – est dans le cas du plan équivalent à la règle fixant le sens de parcours de la courbe de sorte que l'intérieur reste à gauche. (...) Cependant cette dernière détermination n'est déjà plus applicable pour un déplacement dans l'espace, car ici une « droite » ou une « gauche » ne pourraient pas être différenciées, et pour entrer dans l'étude de variétés de plus grande dimension, fixer *arithmétiquement* un principe de parcours est bien une condition nécessaire préalable. [Kronecker, 1891a] p. 232

L'argument que développe Kronecker dans son cours porte donc sur la difficulté à étendre la notion de sens de parcours lorsque l'on est en dimension supérieure à deux. Pourtant, dans le cours de 1881 et dans celui de 1891, Kronecker ne traitera que des courbes du plan [43].

40. Voir la traduction de [Gauss, 1849] en annexe.

41. In der Regel durchläuft man nun eine Curve so, daß das Innere links bleibt, doch ist diese Betrachtung auf ebene Gebilde beschränkt. Wir wollen daher ein rein analytisches Criterium aufstellen

42. Diese Festlegung der Richtung – wir wollen es das Fortgangsprinzip nennen – ist für den Fall der Ebene gleichbedeutend mit der Vorschrift, die Kurve so zu durchlaufen, daß das Innere zur Linken liegt.(...) Jedoch ist diese letztere Festsetzung schon für eine Bewegung in Raume nicht mehr anwendbar, da hier ein « rechts » oder « links » nicht unterschieden werden kann, und für den Eintritt in die Untersuchungen höherer Mannigfaltigkeiten wird es erst recht eine unerlässliche Bedingung, ein Fortgansprinzip *arithmetisch* zu fixieren.

43. Cependant il semble ici nécessaire de rappeler l'aspect spéculatif de ces remarques : nous ne sommes en possession que d'un manuscrit décrivant, peut-être de façon incomplète, la théorie dans le cas général. La prise de notes de Kneser ([Kneser, 1881]) de 1880-1881 pourrait laisser penser que la théorie des caractéristiques n'est pas abordée, alors que celle de Runge de la même année (mais qui contient peut-être aussi celle de 79) se termine par quelques notes sur la théorie pour $n+1$ équations.

En fait, un seul des manuscrits des cours sur la théorie des équations algébriques présents à la bibliothèque de Strasbourg traite du cas $n$ : il s'agit de celui de 1878/79, qui suit donc immédiatement la parution de son article *Über die Charakteristik von Funktionen-Systemen*. Il faut donc chercher aussi ailleurs que dans ses cours les raisons de l'introduction du *Fortgangsprincip*.

### 13.3.2 Explication du principe

Nous allons commencer par suivre l'article de 1869 où Kronecker expose pour la première fois la théorie des caractéristiques.

#### 13.3.2.1 Dans l'article de 1869

Voici la définition que donne Kronecker dans son premier article de 1869 du *Fortgangsprincip*[44] :

44. Es seien

$$F_0, F_1, F_2, ..., F_n$$

eindeutige reelle Functionen der $n$ reellen unbeschränkt veränderlichen Grössen

$$z_1, z_2, ...z_n$$

und zwar solche, die sowohl eine $n$fach unendliche Anzahl positiver als negativer Werthe annehmen. Ueberdies werden die Functionen $F$ als im Allgemeinen stetig und nach den einzelnen Variabeln differentiirbar vorausgesetzt und es soll endlich angenommen werden, dass keine der $(n+1)$ Functionaldeterminanten gleichzeitig mit den betreffenden Functionen für unendlich viele Werthsysteme $z$ verschwindet Um diese $(n+1)$ Functionaldeterminanten auch ihrem Vorzeichen nach zu fixiren, stelle ich den aus den $n$ partiellen Ableitungen gebildeten $n$ Verticalreihen eine voran, deren $(n+1)$ Elemente resp. mit

$$F_{00}, F_{10}, ...F_{n0}$$

bezeichnet werden mögen, und bilde aus den auf diese Weise resultirenden und nach der Folge der Indices zu ordnenden $(n+1)^2$ Elementen die Determinante. Alsdann ist der nach $F_{k0}$ genommene partielle Differentialquotient dieser Determinante die auch dem Vorzeichen nach bestimmte Functionaldeterminante der $n$ Functionen :

$$F_0, F_1, ...F_{k-1}, F_{k+1}, ...F_n$$

welche mit $\Delta_k$ bezeichnet werden soll.
Wenn man irgend welche $(n-1)$ der Functionen $F$ gleich Null setzt, so wird hierdurch die Veränderlichkeit der Variabeln $z$ auf eine einfache Unendlichkeit oder Mannigfaltigkeit eingeschränkt. Die derselben angehörigen Werthsysteme der $n$ Grössen $z$ bilden eine stetige Folge, und man kann jene Werthsysteme füglich als „Punkte“ und deren stetige Folge als „Linie“ bezeichnen. Um an einem beliebigen Punkte der Linie :

$$F_i = 0 \text{ für } i \text{ alle Indices mit Ausnahme zweier } (h \text{ und } k) \text{ gesetzt,}$$

den Sinn des Fortgangs in derselben zu fixiren, setze ich entweder an die Stelle von $F_h$ oder an die von $F_k$. irgend eine eindeutige Function $\Phi$ und bestimme den Fortgang so, dass an der betrachteten Stelle $d\Phi$ dasselbe Vorzeichen erhält, welches die aus den $(n^-1)$ Functionen $F_i$ unter Hinzunahme von gebildete Functionaldeterminante in dem bezeichneten Punkte hat. Nach dieser Bestimmung ist der Sinn des Fortgangs verschieden, je nachdem $\Phi$ an die Stelle von $F_h$ oder an die von $F_k$ getreten ist, und es soll demgemäss die Linie selbst mit $[hk]$ oder mit $[kh]$ bezeichnet werden ; aber der Sinn des Fortgangs ist in allen Punkten der Linie, welche nicht Doppelpunkte sind, genau fixirt und übrigens von der Wahl der Function $\Phi$ unabhängig. Da bis zu beliebiger Nähe der Doppelpunkte der Sinn des Fortgangs noch festzustellen ist, so schliesst auch das Vorkommen von Doppelpunkten die Anwendbarkeit jener Bestimmung nicht aus.

Soient

$$F_0, F_1, F_2, ..., F_n$$

des fonctions réelles univoques des $n$ grandeurs variables réelles illimitées

$$z_1, z_2, ...z_n$$

lesquelles prennent $n$ fois un nombre infini de valeurs négatives et positives. De plus les fonctions $F$ sont considérées généralement continues et différentiables par rapport à chaque variable et on doit enfin supposer qu'aucun des $(n+1)$ déterminants fonctionnels ne s'annule simultanément avec les fonctions concernées pour un nombre infini de systèmes de valeurs $z$. Pour fixer ces $n+1$ déterminants fonctionnels quant à leur signe, je mets alors en tête des $n$ dérivées partielles sur les $n$ rangs verticaux les $n+1$ éléments que je désignerai respectivement par

$$F_{00}, F_{10}, ...F_{n0}$$

et je forme le déterminant construit à partir des $(n+1)^2$ éléments ainsi obtenus et ordonnés selon les indices. Le quotient différentiel partiel pris selon $F_{k0}$ de ce déterminant est bien au signe près le déterminant fonctionnel des $n$ fonctions

$$F_0, F_1, ...F_{k-1}, F_{k+1}, ...F_n$$

que l'on désignera par $\Delta_k$. Si l'on fixe $(n-1)$ des fonctions $F$ égales à zéro, alors de cette façon la variable $z$ se trouve dans une infinité simple, ou variété. Dans celle-ci les systèmes de valeurs de $n$ grandeurs $z$ forment une suite continue ; on peut désigner chaque système de valeur par "point" et cette suite continue par "ligne". Pour fixer le sens de parcours en l'un quelconque des points d'une ligne

$$F_i = 0 \text{ pour } i \text{ prenant toutes les indices à l'exception de deux } (h \text{ et } k),$$

je mets à la place de $F_h$ ou à la place de $F_k$ une fonction $\Phi$ et je détermine le parcours de façon à ce qu'à l'endroit considéré $d\Phi$ conserve le même signe que celui du déterminant fonctionnel formé à partir des $n-1$ fonctions $F_i$ et de $\Phi$ au point considéré. D'après cela, le sens de parcours est différent selon que l'on substitue $\Phi$ à $F_h$ ou à $F_k$, et l'on devra en conséquence désigner la ligne par $[hk]$ ou $[kh]$ ; mais le sens de parcours est fixé en tous les points de la ligne qui ne sont pas des points doubles, et d'ailleurs de façon indépendante du choix de $\Phi$. Comme jusqu'à n'importe quel voisinage d'un point double, le sens d'un parcours peut être déterminé, alors la présence d'un point double n'exclut pas non plus la possibilité d'application de cette détermination. [Kronecker, 1869c] p. 178-179

En première lecture, cette définition est très difficilement compréhensible : nous allons voir que ce passage, qui sert d'introduction à [Kronecker, 1869c], est en effet d'une densité remarquable, et réclame pour être compris un important travail d'explicitation. Nous donnerons cet éclairage en reprenant les notations de cet article, à partir de ce que fait Kronecker dans son cours.

#### 13.3.2.2 Dans son cours : le cas $n = 2$

Comme nous l'avons dit, Kronecker ne traite essentiellement dans son cours que le cas des courbes du plan, c'est-à-dire le cas $n = 2$.

On a alors trois fonctions numériques $F_0, F_1$ et $F_2$ que Kronecker suppose différentiables par rapport aux deux variables $x$ et $y$. On notera, comme Kronecker

$$F_{01} = \frac{\partial F_0}{\partial x} \qquad F_{02} = \frac{\partial F_0}{\partial y}$$
$$F_{11} = \frac{\partial F_1}{\partial x} \qquad F_{12} = \frac{\partial F_1}{\partial y}$$
$$F_{21} = \frac{\partial F_2}{\partial x} \qquad F_{22} = \frac{\partial F_2}{\partial y}$$

Soit de plus trois variables $F_{00}, F_{10}$ et $F_{20}$. Alors on a le déterminant :

$$\Delta = \begin{vmatrix} F_{00} & F_{10} & F_{20} \\ F_{01} & F_{11} & F_{21} \\ F_{02} & F_{12} & F_{22} \end{vmatrix}$$

et les sous déterminants

$$\Delta_1 = \begin{vmatrix} F_{11} & F_{21} \\ F_{12} & F_{22} \end{vmatrix} = F_{11}.F_{22} - F_{21}.F_{12}$$

$$\Delta_2 = -\begin{vmatrix} F_{01} & F_{21} \\ F_{02} & F_{22} \end{vmatrix} = F_{21}.F_{02} - F_{01}.F_{22}$$

$$\Delta_3 = \begin{vmatrix} F_{01} & F_{11} \\ F_{02} & F_{12} \end{vmatrix} = F_{01}.F_{12} - F_{11}.F_{02}$$

$F_0 = 0$, $F_1 = 0$ et $F_2 = 0$ sont les trois courbes – les 3 *lignes* dans le vocabulaire de Kronecker – sur lesquelles nous souhaitons nous déplacer. Le principe est le suivant : on parcourt une courbe en regardant les points d'intersection avec l'une des deux autres. On obtient alors 6 parcours différents : $[01], [10], [02], [20], [12], [21]$.

1. $[01]$ correspond à un parcours de $F_2 = 0$ en regardant les point d'intersection avec $F_0$
2. $[10]$ correspond à un parcours de $F_2 = 0$ en regardant les point d'intersection avec $F_1$
3. $[02]$ correspond à un parcours de $F_1 = 0$ en regardant les point d'intersection avec $F_0$
4. $[20]$ correspond à un parcours de $F_1 = 0$ en regardant les point d'intersection avec $F_2$

5. [12] correspond à un parcours de $F_0 = 0$ en regardant les point d'intersection avec $F_1$
6. [21] correspond à un parcours de $F_0 = 0$ en regardant les point d'intersection avec $F_2$

Regardons par exemple plus précisément [12] et [21] :

**[12]** : On parcourt $F_0 = 0$ en regardant les points d'intersection avec $F_1 = 0$. Soit $M(x', y')$ un point de $F_0 = 0$ à partir duquel on veut se déplacer et $\Phi(x, y)$ une nouvelle fonction répondant aux mêmes conditions que les $F_i$ et telle que la courbe $\Phi = 0$ passe par $M$.

Ce que souhaite Kronecker, c'est que

$$\frac{\partial \begin{vmatrix} F_{00} & \Phi & F_{20} \\ F_{01} & \Phi_1 & F_{21} \\ F_{02} & \Phi_2 & F_{22} \end{vmatrix}}{\partial F_{20}} . d\Phi > 0,$$

car il « détermine le parcours de façon à ce qu'à l'endroit considéré $d\Phi$ conserve le même signe que celui du déterminant fonctionnel formé à partir des $n-1$ fonctions $F_i$ et de $\Phi$ au point considéré. » [45]. Bien sûr, $\Phi_1$ désigne $\dfrac{\partial \Phi}{\partial x}$ et $\Phi_2 = \dfrac{\partial \Phi}{\partial y}$

c'est-à-dire que

$$\begin{vmatrix} F_{01} & \Phi_1 \\ F_{02} & \Phi_2 \end{vmatrix} . d\Phi > 0$$

donc que [46]

$$(F_{01}\Phi_2 - F_{02}\Phi_1).d\Phi > 0.$$

**Notation** Pour comprendre ce qui suit, il faut se rappeler ce que représentent $dx$ et $dy$ pour Kronecker dans le cadre de cette étude. On est en train de se déplacer sur une courbe. Si l'on pose $ds$ l'élément d'arc au point $M$, $dx$ et $dy$ sont les projections orthogonales de $ds$. Ainsi, on peut voir $\vec{t}(dx, dy)$ comme un vecteur tangent à la courbe $f = 0$. Comme $\vec{n}(f_1(x', y'), f_2(x', y'))$ est normale à cette même courbe en $M$, ces deux vecteurs sont orthogonaux et leur produit scalaire est donc nul :

$$f_1 dx + f_2 dy = 0$$

45. Voir citation précédente.
46. Kronecker écrit plutôt

$$sgn\ d\Phi = sgn\ F_{01}\Phi_2 - F_{02}\Phi_1$$

Cela ne semble pas nécessiter d'explications supplémentaires pour un lecteur de la fin du XIX^e siècle : dans les quatre manuscrits où Kronecker présente sa théorie des caractéristiques il pose sans aucune justification l'égalité précédente ([Kronecker, 1873a] p. 173, [Kronecker, 1879] p. 513, [Kronecker, 1882a] p. 164 et [Kronecker, 1891a] p. 232). Même dans le cours de 1878-79, où la plupart des arguments sont explicités, cette égalité n'est qu'affirmée : il la tient donc pour acquise. Et en effet, cette conception est l'un des outils de la géométrie différentielle du XIX^e siècle, et fait partie des connaissances du calcul différentiel enseigné dans la seconde moitié du XIX^e siècle. On pourra lire par exemple le cours de Sturm à l'école polytechnique ([Sturm *et al.*, 1901] p. 19) ou le cours de calcul infinitésimal pour ingénieur de Rouché-Levy ([Rouché et Lévy, 1900] p. 64).

Une autre façon d'interpréter cette égalité est la suivante : nous nous déplaçons sur une courbe où $f$ est constamment nulle. Si l'on suppose l'existence d'un paramétrage $(x(t), y(t))$ de cette courbe, alors $F(t) = f(x(t), y(t)) = 0$, donc $dF = x'(t)f_1(x(t), y(t))dt + y'(t)f_2(x(t), y(t))dt = 0$. En posant $x'(t)dt = dx$ et $y'(t)dt = dy$ on obtient l'égalité voulue.

Dans les deux cas, Kronecker utilise des résultats de la géométrie différentielle lorsqu'il cherche à « arithmétiser » son principe de parcours.

Il faut maintenant montrer que cette condition est indépendante du choix de la fonction $\Phi$. Voici comment Kronecker procède dans ses cours de 1880/81 et 1890/91.

Supposons que pour toute fonction $\Phi$ satisfaisant les hypothèses on ait $(F_{01}\Phi_2 - F_{02}\Phi_1).d\Phi > 0$ et posons $\Phi = y + a$ où $a \in \mathbb{R}$ est choisi tel que $\Phi$ passe par $M$. Alors $d\Phi = dy$, $\Phi_1 = 0$ et $\Phi_2 = 1$. Donc on a

$$F_{01}.dy > 0$$

Réciproquement, supposons que $F_{01}.dy > 0$ (et donc $F_{01} \neq 0$). On circule sur $F_0 = 0$, donc

$$F_{01}dx + F_{02}dy = 0$$

et donc

$$dx = -\frac{F_{02}dy}{F_{01}}.$$

Ainsi on a

$$\begin{aligned} d\Phi &= \Phi_1 dx + \Phi_2 dy \\ &= \left(\Phi_2 - \Phi_1 \frac{F_{02}}{F_{01}}\right) dy \\ &= (\Phi_2 F_{01} - \Phi_1 F_{02}) \frac{dy}{F_{01}} > 0 \end{aligned}$$

Or $F_{01}.dy > 0$, donc $\frac{dy}{F_{01}} > 0$ et finalement $d\Phi$ et $(\Phi_2 F_{01} - \Phi_1 F_{02})$ sont du même signe, c'est-à-dire que $(F_{01}\Phi_2 - F_{02}\Phi_1).d\Phi > 0$.

Donc pour toute fonction $\Phi$, il y a équivalence entre $F_{01}.dy > 0$ et $(F_{01}\Phi_2 - F_{02}\Phi_1).d\Phi > 0$, ce qui montre que la condition fixant le sens de parcours est indépendante du choix de $\Phi$. Examinons ce qui se passe dans les deux illustrations suivantes :

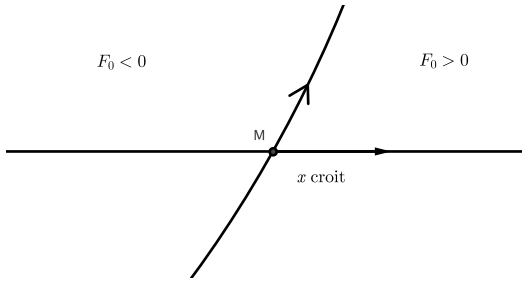



FIGURE 4 –

Dans la figure 4, lorsque $x$ croît, on passe de $F_0 < 0$ à $F_0 > 0$, donc $F_0$ croît et $F_{01} > 0$. Or $F_{01}.dy > 0$, donc $dy > 0$. On doit donc, à partir de $M$, « monter ».

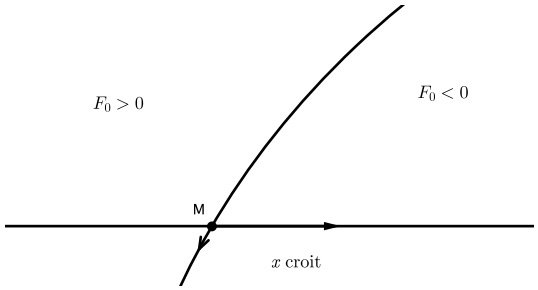



FIGURE 5 –

Dans la figure 5, lorsque $x$ croît, on passe de $F_0 > 0$ à $F_0 < 0$, donc $F_0$ décroît et $F_{01} < 0$. Or $F_{01}.dy > 0$, donc $dy < 0$. On doit donc, à partir de $M$, « descendre ».

**Inversion du sens** On passe maintenant à **[21]**. On souhaite parcourir $F_0$ en regardant les points d'intersection avec, non plus $F_1$, mais $F_2$. On veut ainsi que :

$$\frac{\partial \begin{vmatrix} F_{00} & F_{10} & \Phi \\ F_{01} & F_{11} & \Phi_1 \\ F_{02} & F_{12} & \Phi_2 \end{vmatrix}}{\partial F_{10}} d\Phi > 0$$

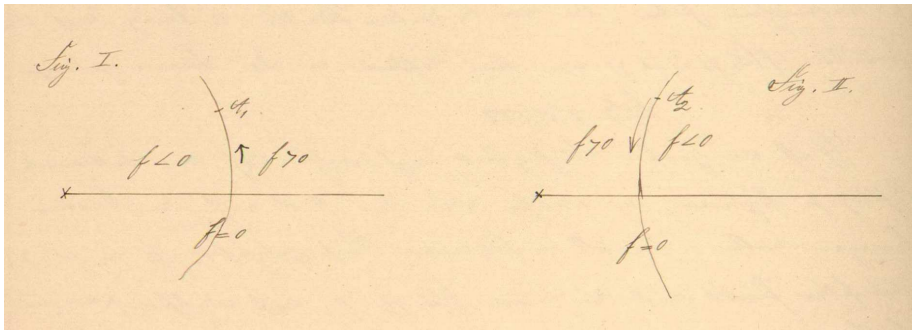



FIGURE 6 – [Kronecker, 1891a] p. 231

c'est-à-dire que

$$(-F_{01}\Phi_2 + F_{02}\Phi_1)d\Phi > 0$$

On voit bien que la condition que l'on obtient conduit à l'inversion du sens de parcours.

**Points doubles :** En ce qui concerne le cas du passage par un point double, où $F_{01} = F_{02} = 0$, la détermination précédente n'est plus valable. Nous nous contenterons de la remarque de Kronecker : on peut toujours dans ce cas utiliser le principe jusque dans un voisinage aussi petit que l'on veut du point en question. Il suffit ensuite de parcourir séparément chacune des deux « boucles » ainsi formées. L'exemple qui est souvent donné est celui du parcours de la lemniscate [47], qui permet de donner une idée claire de ce qu'il se passe [48] :

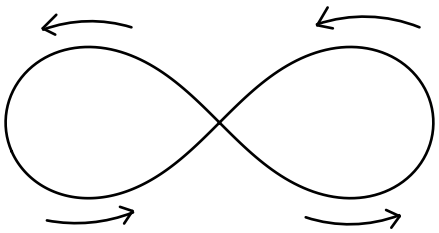


FIGURE 7 –

> De là il résulte que ce principe de parcours n'est pas applicable aux points pour lesquels $f_1 = f_2 = 0$, si bien que les singularités, comme par exemple les points doubles, ne peuvent pas être traversées, et on doit les parcourir séparément comme sur les deux parties d'une lemniscate. [Kronecker, 1891a] p. 233

47. Par exemple [Kronecker, 1881a] p. 170 ; [Kronecker, 1891a] p. 233 ; [Landsberg, 1911] p. 573.

48. Hieraus geht hervor, daß dieses Fortgangsprinzip nicht anwendbar ist an Punkten, für welche $f_1 = f_2 = 0$ ist, sodaß Singularitäten, wie z.B. Doppelpunkte, nicht durchlaufen werden können, und man z.B. den Weg um die beiden Teile einer Lemniskate einzeln machen muß.

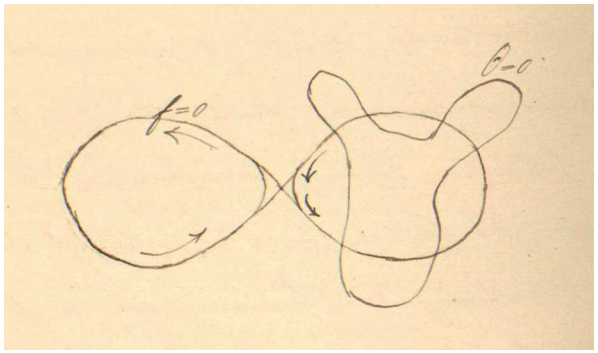



FIGURE 8 – [Kronecker, 1891a] p. 233

Après avoir examiné de près le cas $n = 2$ traité dans son cours, nous pouvons passer maintenant au cas général qui est l'objet de son article de 1869. L'interprétation géométrique de ce cas simple nous a permis en effet de suivre le raisonnement de Kronecker et de comprendre ses notations.

#### 13.3.2.3 Le cas général

Commençons par préciser à quel type de « courbes » s'applique ce principe. Soient $(n+1)$ fonctions numériques $F_0, F_1, ..., F_n$ dépendant de $n$ variables ayant les propriétés suivantes :

1. Elles sont continues et dérivables par rapport à chacune de leurs variables.
2. Elles doivent prendre une infinité de valeurs positives et négatives.
3. Les domaines définis par $F_i < 0$ sont *fermés* (*geschlossene*), c'est-à-dire ici qu'ils ne contiennent que des valeurs finies (ils sont bornés).
4. Ces $n+1$ fonctions n'ont pas de zéro commun.
5. Si l'on prend $n$ de ces fonctions, leur déterminant fonctionnel ne s'annule en aucun de leurs zéros communs.

Soit $h, k \in [\![0, n]\!]$. l'ensemble des points $(z_1, ..., z_n)$ tels que $F_1 = F_2 = ... = F_{h-1} = F_{h+1} = ... = F_{k-1} = F_{k+1} = ... = F_m = 0$ est une sous-variété de dimension 1 que Kronecker qualifie de suite continue de points (*stetige Folge*) ou de ligne (*Linie*). On pourra remarquer l'utilisation d'un vocabulaire issu de la géométrie qui peut rappeler la façon dont il traite ces notions pour les petites dimensions à l'intérieur de son cours. C'est sur ces lignes qu'il définit son *Fortgangsprincip*, son sens de parcours.

Introduisons les notations de Kronecker :

1. Pour tout $k \in [\![0, n]\!]$ et tout $h \in [\![1, n]\!]$ on note $F_{kh} = \frac{\partial F_k}{\partial z_h}$.
2. Soit $F_{00}, F_{10}, ..., F_{n0}$ $n+1$ nouvelles indéterminées.

3. Soit

$$\Delta = \begin{vmatrix} F_{00} & F_{10} & F_{20} & ... & F_{n0} \\ F_{01} & F_{11} & F_{21} & ... & F_{n1} \\ F_{02} & F_{12} & F_{22} & ... & F_{n2} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ F_{0n} & F_{1n} & F_{2n} & ... & F_{nn} \end{vmatrix}$$

et

$$\Delta_h = \frac{\partial \Delta}{\partial F_{h0}} = (-1)^{h+2} \begin{vmatrix} F_{01} & F_{11} & F_{21} & ... & F_{h-1,1} & F_{h+1,1} & ... & F_{n1} \\ F_{02} & F_{12} & F_{22} & ... & F_{h-1,2} & F_{h+1,2} & ... & F_{n2} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ F_{0n} & F_{1n} & F_{2n} & ... & F_{h-1,n} & F_{h+1,n} & ... & F_{nn} \end{vmatrix}$$

qui est, au signe près[49], le déterminant fonctionnel des fonctions $F_0, F_1, F_2, ..., F_{h-1}, F_{h+1}, ..., F_n$.

4. on désignera par $[hk]$ ou $[kh]$ la *ligne* définie par $F_1 = F_2 = ... = F_{h-1} = F_{h+1} = ... = F_{k-1} = F_{k+1} = ... = F_m = 0$, selon justement le sens de parcours choisi.

Pour déterminer le sens de parcours sur $[hk]$ en un point $M(z_0, ..., z_n)$, on choisit une fonction $\Phi$ répondant aux mêmes conditions que les $F_i$ et passant par ce point. On forme ensuite le déterminant fonctionnel suivant :

$$\Delta_{\Phi,k} = \begin{vmatrix} F_{00} & F_{10} & F_{20} & ... & F_{h-1,0} & \Phi & F_{h+1,0} & ... & F_{k0} & ... & F_{n0} \\ F_{01} & F_{11} & F_{21} & ... & F_{h-1,1} & \Phi_1 & F_{h+1,1} & ... & F_{k1} & ... & F_{n1} \\ F_{02} & F_{12} & F_{22} & ... & F_{h-1,2} & \Phi_2 & F_{h+1,2} & ... & F_{k2} & ... & F_{n2} \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & .... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ F_{0n} & F_{1n} & F_{2n} & ... & F_{h-1,n} & \Phi_n & F_{h+1,n} & ... & F_{kn} & ... & F_{nn} \end{vmatrix}$$

on a alors :

49. Dans la formule précédente, $(-1)^{h+2}$ peut être remplacé par $(-1)^{h+2}$. Cependant, pour souligner que la formule provient d'un développement du déterminant selon la première ligne, nous avons souhaité laisser $(-1)^{h+2}$, qui est issu du coefficient $(-1)^{h+2}F_{h0}$ qui apparaît lors de ce développement.

$$\frac{\partial \Delta_{\Phi,k}}{\partial F_{k0}}$$

$$= (-1)^{k+2} \begin{vmatrix} F_{01} & F_{11} & F_{21} & ... & F_{h-1,1} & \Phi_1 & F_{h+1,1} & ... & F_{k-1,1} & F_{k+1,1} & ... & F_{n1} \\ F_{02} & F_{12} & F_{22} & ... & F_{h-1,2} & \Phi_2 & F_{h+1,2} & ... & F_{k-1,2} & F_{k+1,2} & ... & F_{n2} \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ F_{0n} & F_{1n} & F_{2n} & ... & F_{h-1,n} & \Phi_n & F_{h+1,n} & ... & F_{k-1,n} & F_{k+1,n} & ... & F_{nn} \end{vmatrix}$$

Le sens de parcours à partir du point $M$ sera celui qui rend $\frac{\partial \Delta_{\Phi,k}}{\partial F_{k0}}.d\Phi > 0$. Comme précédemment, le sens de parcours est indépendant du choix de $\Phi$. Kronecker ne fait que mentionner cette propriété, mais n'en donne aucune démonstration.

Si l'on veut maintenant déterminer le sens de parcours sur $[kh]$, on forme :

$$\Delta_{\Phi,h} = \begin{vmatrix} F_{00} & F_{10} & F_{20} & ... & F_{h0} & ... & F_{k-1,0} & \Phi & F_{k+1,0} & ... & F_{n0} \\ F_{01} & F_{11} & F_{21} & ... & F_{h1} & ... & F_{k-1,1} & \Phi_1 & F_{k+1,1} & ... & F_{n1} \\ F_{02} & F_{12} & F_{22} & ... & F_{h2} & ... & F_{k-1,2} & \Phi_2 & F_{k+1,2} & ... & F_{n2} \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & .... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ F_{0n} & F_{1n} & F_{2n} & ... & F_{hn} & ... & F_{k-1,n} & \Phi_n & F_{k+1,n} & ... & F_{nn} \end{vmatrix}$$

et on a ici :

$$\frac{\partial \Delta_{\Phi,h}}{\partial F_{h0}}$$

$$= (-1)^{h+2} \begin{vmatrix} F_{01} & F_{11} & F_{21} & ... & F_{h-1,1} & F_{h+1,1} & ... & F_{k-1,1} & \Phi_1 & F_{k+1,1} & ... & F_{n1} \\ F_{02} & F_{12} & F_{22} & ... & F_{h-1,2} & F_{h+1,2} & ... & F_{k-1,2} & \Phi_2 & F_{k+1,2} & ... & F_{n2} \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ F_{0n} & F_{1n} & F_{2n} & ... & F_{h-1,n} & F_{h+1,n} & ... & F_{k-1,n} & \Phi_n & F_{k+1,n} & ... & F_{nn} \end{vmatrix}$$

On souhaite bien sûr toujours avoir $\frac{\partial \Delta_{\Phi,h}}{\partial F_{h0}}.d\Phi > 0$. Comparons $\frac{\partial \Delta_{\Phi,k}}{\partial F_{k0}}$ et $\frac{\partial \Delta_{\Phi,h}}{\partial F_{h0}}$. On voit que pour passer du premier au second, la colonne des $\Phi$ se déplace de $k-h+1$ colonnes, donc

$$\frac{\partial \Delta_{\Phi,k}}{\partial F_{k0}}$$

$$= (-1)^{k+2} \begin{vmatrix} F_{01} & F_{11} & F_{21} & ... & F_{h-1,1} & \Phi_1 & F_{h+1,1} & ... & F_{k-1,1} & F_{k+1,1} & ... & F_{n1} \\ F_{02} & F_{12} & F_{22} & ... & F_{h-1,2} & \Phi_2 & F_{h+1,2} & ... & F_{k-1,2} & F_{k+1,2} & ... & F_{n2} \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ F_{0n} & F_{1n} & F_{2n} & ... & F_{h-1,n} & \Phi_n & F_{h+1,n} & ... & F_{k-1,n} & F_{k+1,n} & ... & F_{nn} \end{vmatrix}$$

$$= (-1)^{k+2}(-1)^{k-h+1}$$

$$\times \begin{vmatrix} F_{01} & F_{11} & F_{21} & ... & F_{h-1,1} & F_{h+1,1} & ... & F_{k-1,1} & \Phi_1 & F_{k+1,1} & ... & F_{n1} \\ F_{02} & F_{12} & F_{22} & ... & F_{h-1,2} & F_{h+1,2} & ... & F_{k-1,2} & \Phi_2 & F_{k+1,2} & ... & F_{n2} \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ F_{0n} & F_{1n} & F_{2n} & ... & F_{h-1,n} & F_{h+1,n} & ... & F_{k-1,n} & \Phi_n & F_{k+1,n} & ... & F_{nn} \end{vmatrix}$$

$$= (-1)^{k+2}(-1)^{k-h+1}(-1)^{h+2}\frac{\partial \Delta_{\Phi,h}}{\partial F_{h0}}$$

$$= (-1)^{2k+5}\frac{\partial \Delta_{\Phi,h}}{\partial F_{h0}}$$

$$= -\frac{\partial \Delta_{\Phi,h}}{\partial F_{h0}}$$

$[hk]$ et $[kh]$ représentent donc une même ligne qui sera parcourue dans des sens opposés. $[hk]$ correspondra en fait au cas où l'on cherchera les intersections de la ligne $F_1 = F_2 = ... = F_{h-1} = F_{h+1} = ... = F_{k-1} = F_{k+1} = ... = F_m = 0$ avec la sous-variété $F_h = 0$, et $[kh]$ au cas où l'on cherchera les intersections de la même ligne avec la sous-variété $F_k = 0$.

### 13.3.3 Conclusion

Le travail que nous venons de voir est valable pour une très grande variété de courbes, et pas seulement les courbes algébriques, bien que Kronecker considère uniquement ces dernières. Finalement, ce qui est utile ici, c'est plutôt leur aspect *lisse*[50], et donc leurs propriétés *analytiques* plutôt que *algébriques*.

50. Au sens anglais du terme (*smooth)* : fonction extrêmement régulière qui a des différentielles de tout ordre, c'est-à-dire de classe $C^\infty$.

#### 13.3.3.1 Postérité

Les références explicites à ce principe de parcours de Kronecker que nous avons pu trouver sont rares [51]. On le trouve tout d'abord dans un article de Georg Landsberg publié en 1911 : *Beiträge zur Topologie geschlossener Kurven mit Knotenpunkten und zur Kroneckerschen Charakteristikentheorie* ([Landsberg, 1911]). Il est d'ailleurs remarquable que ce principe soit expliqué dans un article comportant un paragraphe intitulé *L'Analysis situs des courbes fermées* [52] : nous verrons que la théorie des caractéristiques peut en effet avoir sa place dans ce champ des mathématiques du début du XX^e^ siècle. Landsberg a travaillé avec Kurt Hensel pour la rédaction en 1902 de la *Theorie der algebraischen Funktionen einer Variablen : und ihre Anwendung auf algebraische Kurven und Abelsche Integrale* ([Hensel et Landsberg, 1902]) dans lequel il est question [53] de la caractéristique, au sens de Kronecker, d'un système de deux fonctions, et cette collaboration a certainement dû lui permettre d'avoir accès aux notes de cours de Kronecker. Dans ce mémoire, Landsberg cite Paul Stäckel : ce dernier a effectivement utilisé ce principe dans son article de 1893 *Ueber Systeme von Functionen reeller Variabeln* ([Stäckel, 1893]). Rappelons que Stäckel a suivi les cours de Kronecker à Berlin et a pris les notes à l'origine d'au moins l'un des manuscrits présents à Strasbourg ([Kronecker, 1886]) : il a donc eu un accès direct et privilégié aux cours de Kronecker. Enfin, Johannes Knoblauch (1855-1915), dans ses *Grundlagen der differentialgeometrie* ([Knoblauch, 1913]), utilise le terme de *Fortgangsprincip*. La notion est assez proche mais cependant différente : elles ne se rejoignent qu'accidentellement.

Pourtant, au début du XX^e^ siècle, le lieu où vit le *Fortgangsprincip*, c'est-à-dire la théorie des caractéristiques, est d'après Hadamard « une notion qui est maintenant classique » [54], et il en donne même un certain nombre d'applications [55]. Mais nous verrons que ce à quoi Hadamard fait allusion, c'est en fait à la seconde partie du premier article de Kronecker, où il développe ce qui est aujourd'hui appelé *l'intégrale de Kronecker*. C'est sous cette version que cette extension de l'indice de Cauchy trouve une postérité à la fin du XIX^e^ siècle et au début du XX^e^, et cela ne correspond pas à ce que Kronecker expose dans ses cours. Nous reviendrons sur cela plus tard, ainsi que sur le travail de Hadamard. Pour le moment, ce que l'on a mis en évidence, c'est que ce principe semble avoir été repris uniquement par des mathématiciens ayant eu accès au cours de Kronecker sur la théorie des équations algébriques.

Enfin, le *Fortgangsprincip* de Kronecker apparaît dans des articles [56] récents d'Yvon

51. Un seul résultat sur *Google schlolar* pour le terme utilisé par Kronecker (En fait l'article de Kronecker) *Fortgangsprincip* et huit résultats pour *Fortgangsprinzip* (La plupart dans des articles de Kronecker). Recherche effectuée le 10 février 2016.
52. [Landsberg, 1911] p. 564 : « Die Analysis situs der geschlossenen Kurven ».
53. Voir [Hensel et Landsberg, 1902] p. 609.
54. Voir [Tannery et Hadamard, 1910] p. 437.
55. Ibid.
56. Par exemple dans [Gauthier, 2002], [Gauthier, 2013] ou [Gauthier, 2015].

Gauthier, mais celui-ci ne semble pas correspondre à ce que Kronecker fait réellement dans ses travaux [57]. Ce principe, qui permet de s'orienter sur les sous-variétés de dimension 1, n'a été que très peu commenté et repris. Nous pouvons donc raisonnablement nous poser la question : pourquoi avoir introduit ce principe ?

#### 13.3.3.2 Pourquoi avoir introduit ce principe ?

Dans le premier tome de son traité d'algèbre où la théorie de la caractéristique de Kronecker est exposée [58], Weber la décrit, comme nous l'avons vu, uniquement pour le cas d'un système de 3 fonctions et en donne d'ailleurs l'exposé le plus détaillé que l'on puisse trouver de cette version : nous en verrons plus tard une seconde qui repose sur des principes très différents, et qui utilise des intégrales. Pourtant, dans cet exposé très complet, le *Fortgangsprincip* n'est pas décrit en tant que tel, et le fameux

$$(F_{01}\Phi_2 - F_{02}\Phi_1).d\Phi > 0.$$

est regardé comme une conséquence du fait de laisser à « gauche » le domaine $F_0 < 0$. Mais lorsqu'il examine un système de trois courbes et regarde comme Kronecker les intersections de l'une d'entre elles avec une autre, il donne sans justification une recette pour effectuer ce parcours.

Une comparaison minutieuse du manuel de Weber et du cours de Kronecker nous permet de supposer que Weber a eu dans les mains le cours de Kronecker sur les caractéristiques. Si dans ce livre, qualifié encore aujourd'hui de « mine de résultats » [59] et que Leo Corry décrit comme le « standard German textbook on algebra » [60], Weber a décidé de ne pas développer ce *Fortgangsprincip*, c'est certainement en grande partie parce que son exposition n'est pas nécessaire à la compréhension de la théorie des caractéristiques dans le cas $n = 2$.

Ce principe a donc eu une très faible postérité, tout au moins sous cette forme. Weber, qui donne l'une des présentations de la théorie des caractéristiques la plus proche de celle de Kronecker dans ses cours, évite de l'exposer.

Revenons à la question du début : pour quelles raisons Kronecker expose le *Fortgangsprincip* dans ses cours ?

On voit bien la nécessité pour Kronecker de donner cette définition « analytique » ou « arithmétique » dans l'article de 1869 où il manipule des variétés de dimension $n$, mais pourquoi a-t-il introduit le *Fortgansprincip* dans ses cours ? On ne peut noter aucune

57. Il décrit le *Fortgansprinzip* comme quelque chose de proche du théorème du point fixe ([Gauthier, 2002] p. 147), ce qui semble, à première vue, très étrange. Il l'assimile aussi quelques fois à un *principe de progression* pas à pas qui ne ressemble pas à première vue avec le contenu réel des textes de Kronecker.

58. Voir [Weber, 1898] p. 341.

59. Expression utilisée par J.-P. Serre : voir [Schappacher et Volkert, 1997] p.6.

60. Dans [Grattan-Guinness, 2005] pp. 690-699.

différence fondamentale entre l'exposition qu'il en fait en 1881 et celle qu'il présente en 1891 : à part dans les termes employés, il n'y a donc pas ici, comme on peut le voir dans d'autres parties du cours, une réécriture qui prendrait en compte les nouveaux outils qu'il a construits dans les *Grundzüge einer arithmetischen Theorie der algebraischen Grössen*[61], ce qui aurait justifié au moins sa présence en 1891.

Une première explication pourrait être la suivante. Comme on l'a vu, le court paragraphe de l'article de Kronecker qui présente ce principe ne suffit pas à en saisir le fonctionnement : il permet à la rigueur de convaincre le lecteur de l'existence d'un tel principe. On le sait, Kronecker présentait dans ses cours le résultat de ses recherches[62], et il est donc tout à fait envisageable de voir le développement dans son cours du *Fortgangsprincip* dans le cas $n = 2$ comme un complément nécessaire pour la compréhension de la théorie générale des caractéristiques qu'il publie en parallèle. Nous verrons que l'accès à cette théorie n'est pas simplement facilité par la lecture du cours, mais que l'étude des *Vorlesungen* la rend *possible*.

Ce principe est qualifié d'« analytique »en 1880 et d'« arithmétique » en 1890, et ces qualificatifs s'opposent tous deux à celui de « géométrique », qui désigne la méthode (à la Gauss) consistant à choisir un sens de parcours en utilisant les termes « gauche » et « droite » : en effet, on se rappelle que le mot « arithmétique » comprend pour Kronecker toutes les disciplines mathématiques sauf la géométrie et la mécanique. On peut aussi voir une illustration de cette position dans la volonté de faire une description analytique de ce *Fortgangsprincip* dans le cas des courbes du plan. Ce principe étant établi, nous allons pouvoir maintenant examiner comment Kronecker présente la théorie des caractéristiques.

## 13.4 La caractéristique de trois fonctions à deux variables

Kronecker procède dans son cours, selon les années, à trois types de présentation : la caractéristique d'un système de deux fonctions à une variable, celle de trois fonctions à deux variables et celle de $n + 1$ fonctions à $n$ variables. Nous avons déjà traité de la première, qui est fondamentalement différente des deux autres. Les deux dernières ne présentent pas de différences conceptuelles notables, la troisième étant une extension de la seconde à des espaces de dimensions supérieures. Nous allons ici traiter, comme le fait Kronecker dans la majorité des cours qui contiennent la théorie des caractéristiques, le cas $n = 2$.

### 13.4.1 (*Fortgangsprincip*) : un résumé pratique

Ce principe joue un rôle essentiel dans la théorie des caractéristiques et nous allons donner, pour faciliter la lecture, un résumé de la section précédente dans le cas de trois fonctions à deux variables :

61. Voir [Kronecker, 1882b].
62. Voir p. 35.

1. Il découle d'un choix préalable de l'ordre des fonctions étudiées : soit le déterminant

$$\Delta = \begin{vmatrix} f & \varphi & \psi \\ f_1 & \varphi_1 & \psi_1 \\ f_2 & \varphi_2 & \psi_2 \end{vmatrix}$$

Si l'on parcourt $f$ en regardant les points d'intersection avec $\varphi$, le sens de parcours aux points d'intersection sera celui tel que :

$$\frac{\partial \Delta}{\partial \psi}.d\varphi > 0$$

De même, si l'on parcourt $f$ en regardant les points d'intersection avec $\psi$, le sens de parcours aux points d'intersection sera celui tel que :

$$\frac{\partial \Delta}{\partial \varphi}.d\psi > 0$$

et sera donc l'opposé du précédent (antisymétrie du déterminant).

Il y a donc en tout 6 façons de voir l'ensemble de ces trois courbes.

### 13.4.2 Une proposition problématique

Kronecker commence par donner la proposition suivante :

**Proposition :** Si $U(x, y)$ et $V(x, y)$ sont deux fonctions numériques suffisamment régulières (le plus raisonnable serait de les prendre algébriques) et fermées (au sens que lui donne Kronecker). Alors :

$$\sum_{U=V=0} sgn(dV) = 0 = \sum_{U=V=0} sgn\ (U_1V_2 - U_2V_1)$$

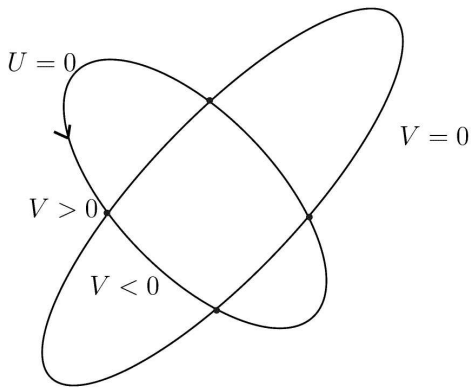



FIGURE 9 –

**Remarque :** Lors des sommes du type $\sum_{U=V=0}$ on placera la courbe parcourue en premier et la courbe intersectée en second.

**Preuve :** l'argument de Kronecker est finalement assez géométrique, même si l'adjectif « topologique » pourrait être avancé aujourd'hui. Il s'agit de dire que, les courbes étant fermées, si l'on parcourt entièrement la courbe $U = 0$, on sera « entré » autant de fois dans la courbe $V = 0$ qu'on en sera « sorti ». Je rappelle qu'« intérieur » et « extérieur » de $V = 0$ sont définis respectivement par $V < 0$ et $V > 0$. Lorsque l'on entre dans $V = 0$, on passe de $V > 0$ à $V < 0$, donc $dV < 0$. Au contraire si l'on sort de $V = 0$, on passe de $V < 0$ à $V > 0$, donc $dV > 0$. Finalement

$$\sum_{U=V=0} sgn(dV) = 0$$

Or, d'après le *Fortgangsprincip*, $dV$ est soit toujours du même signe que $U_1V_2 - U_2V_1$, soit toujours du même signe que $-(U_1V_2 - U_2V_1)$. On a donc toujours

$$\sum_{U=V=0} sgn\ (U_1V_2 - U_2V_1) = 0.$$

### 13.4.3 Remarques à propos de l'argument utilisé par Kronecker

Kronecker a utilisé ici – et d'ailleurs on retrouve cet argument à de nombreuses reprises dans son œuvre, de façon plus ou moins explicite – une argumentation dont Gauss s'est servi dans la première et la quatrième démonstration du théorème fondamental de l'algèbre [63] :

> Alors d'après les mathématiques supérieures, il est établi que n'importe quelle courbe algébrique (ou bien chacune des branches de n'importe quelle courbe algébrique si elle se trouve être composée de plusieurs) ou bien revient sur elle-même, ou bien s'étend de part et d'autre vers l'infini, et donc si une branche d'une courbe algébrique entre dans un espace fini, cette même branche doit nécessairement sortir quelque part de cet espace. [Gauss, 1799] p. 27

Ce à quoi il ajoute la note suivante [64] :

63. Ma traduction de « Iam ex geometria sublimiori constat, quamvis curvam algebraicam (sive singulas cuiusvis curvae algebraicae partes, si forte e pluribus composita sit) aut in se redeuntem aut utrimque in infinitum excurrentem esse, adeoque si ramus aliquis curvae algebraicae in spatium definitum intret, eundem necessario ex hoc spatio rursus alicubi exire debere. »

64. Ma traduction de « Satis bene certe demonstratum esse videtur, curvam algebraicam neque alicubi subito abrumpi posse (uti e. g. evenit in curva transscendente, cuius aequatio $y = \frac{1}{\log x}$), neque post spiras infinitas in aliquo puncto se quasi perdere (ut spiralis logarithmica), quantumque scio, nemo dubium contra hanc rem movit. Attamen si quis postulat, demonstrationem nullis dubiis obnoxiam alia occasione tradere suscipiam. In casu praesenti vero manifestum est, si aliquis ramus e.g. 2, ex circulo nllibi exiret (fig. 3), te in circulum inter 0 et 2 intrare, postea circa totum hunc ramum (qui in circuli spatio se perdere deberet) circummeare, et tandem inter 2 et 4 rursus ex circulo egredi posse, ita ut nullibi in tota via in lineam primam incideris. Hoc vero absurdum esse, inde patet, quod in puncto, ubi in circulum ingressus es, superficiem primam supra te habuisti, in egressu, infra ; quare necessario alicubi in superficiem primam ipsam incidere debulsti, sive in punctum lineae primae. — Ceterum ex hoc ratiocinio principiis geometriae situs innixo, quae haud minus valida sunt, quam principia geometriae magnitudinis, sequitur tantummodo, si in aliquo ramo lineae primae in circulum intres, te alio loco ex circulo rursus egredi posse, semper in linea

Certes, il semble avoir été suffisamment bien démontré qu'une courbe algébrique ne peut ni s'interrompre brutalement quelque part (comme par exemple dans la courbe transcendante dont l'équation est $y = \frac{1}{\log x}$) ni, pour ainsi dire, se perdre après des spirales en un point (comme la spirale logarithmique) ; autant que je sache, personne n'a émis de doute à ce sujet. Cependant, si quelqu'un le demande, je me chargerai à l'occasion de donner une démonstration dont on ne pourra douter. Mais dans le cas présent, il est manifeste que si une branche, par exemple 2, ne pouvait sortir nulle part du cercle (fig. 3), on pourrait entrer entre 0 et 2, ensuite faire le tour de toute cette branche (qui devrait se perdre dans l'espace du cercle) et enfin on pourrait de nouveau sortir du cercle entre 2 et 4 de telle sorte que nulle part sur tout le chemin on n'aurait intersecté la ligne première. L'absurdité de cela est évidente du fait qu'au point où l'on est entré dans le cercle, on a eu la première surface au dessus de soi, mais à la sortie au dessous ; c'est pourquoi nécessairement on a dû quelque part rencontrer la première surface, à savoir en un point de la première ligne. Par ailleurs, d'après ce raisonnement reposant sur les principes de la géométrie de position, lesquels ne sont pas moins valides que ceux de la géométrie des grandeurs, il s'ensuit seulement que si on entre dans le cercle par une branche de la première ligne, on peut sortir du cercle à un autre endroit, toujours en restant sur la première ligne. Mais il ne s'ensuit pas que le chemin suivi soit une ligne continue au sens qui lui est donné dans la géométrie supérieure. Mais il suffit ici que le chemin soit une ligne continue au sens commun, c'est-à-dire nulle part interrompue mais cohérente partout. [Gauss, 1799] p. 27

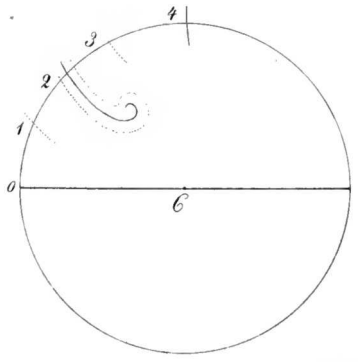



FIGURE 10 – (Fig. 3)[Gauss, 1799] p. 30

Ainsi toute courbe algébrique, et donc aussi chaque branche d'une telle courbe qui entre à l'intérieur d'une courbe *fermée* (au sens que lui donne Kronecker) doit en ressortir : en effet, « ou bien [elle] revient sur elle-même, ou bien [elle] s'étend de part et d'autre vers l'infini ».

Il s'agit là du point problèmatique des deux démonstrations de Gauss. C'est en fait une propriété des courbes algébriques qui n'est pas du tout évidente et il faudra

prima manendo, neque vero, viam tuam esse lineam continuam in eo sensu, quo in geometria sublimiori accipitur. Sed hic sufficit, viam esse lineam continuam in sensu communi, i. e. nullibi interruptam sed ubique cohaerentem. »

attendre 1920 pour que cette difficulté soit mise en évidence par Alexander Ostrowski dans [Ostrowski, 1920], Gauss se contentant de souligner le caractère trivial de cette proposition, et Kronecker s'en servant la plupart du temps de façon implicite.

On peut finalement trouver dans de nombreux commentaires sur les preuves de Gauss du théorème fondamental de l'algèbre cette remarque sur les manques des preuves que l'on peut qualifier de « topologiques » [65]. En revanche il est difficile de trouver une explication de ce que dit exactement Ostrowski [66]. Voyons ce qu'il dit [67] :

> Dans la preuve de cela se trouve la principale difficulté ; les autres conclusions de la *Dissertation* de Gauss suivent alors de façon irréprochable.
>
> Pour justifier cela, Gauss dit qu'il est connu à partir de la géométrie supérieure que si une branche d'une courbe algébrique entre dans un espace limité, elle doit nécessairement en ressortir, et il remarque dans une explication donnée dans une note en bas de page (*Werke* III, p. 27) qu'une courbe algébrique « ne peut ni s'interrompre brutalement quelque part, ni, pour ainsi dire, se perdre après des spirales en un point ». Il poursuit alors : « ... et autant que je sache, personne n'a émis de doute à ce sujet. Cependant, si quelqu'un le demande, je me chargerai à l'occasion de donner

65. Par exemple [Van der Waerden, 1985] p. 97 ou dans le plus récent [Dhombres et Alvarez, 2013] p. 107.

66. Nous n'avons tout au moins rien lu de précis sur ce sujet.

67. Im Nachweis dieser Tatsache liegt die Hauptschwierigkeit ; die weiteren Schlüsse der Gaussschen *Dissertation* folgen dann in durchaus einwandfreier Weise.

Zur Begründung der genannten Tatsache sagt nun Gauss, es sei aus der höheren Geometrie bekannt, dass, wenn ein Zweig einer algebraischen Kurve in einen begrenzten Raum eintritt, er notwendig wieder aus demselben heraustreten müsse, und er bemerkt zur Erläuterung in einer Fussnote (Werke III, S. 27), dass eine algebraische Kurve » weder plötzlich abbricht, noch sich nach unendlich vielen Umläufen gewissermassen in einem Punkt verlieren kann«. Er fügt dann hinzu : »... und soviel ich weiss, hat noch Niemand hiergegen einen Zweifel vorgebracht. Doch werde ich, wenn es jemand fordert, bei anderer Gelegenheit unternehmen, einen keinem Zweifel unterworfenen Beweis zu liefern. « Weiter sagt Gauss, dass im vorliegenden Falle das Nichtabbrechen der hier vorkommenden Kurven auch schon daraus hervorgeht, dass sie Gebiete verschiedenen Vorzeichens von $Re(f(z))$ bezw. $Im(f(z))$ voneinander trennen.

Hierzu ist nun zu bemerken, dass die genaue Durchführung der zuletzt angegebenen Überlegung einer sorgfältigen topologischen Ausführung bedarf. Vor allem wäre aber auch der Beweis, dass ein solches Verhalten unmöglich ist, insofern nicht genügend, als daraus die Möglichkeit des » Fortschreitens « längs der im Innern von $|z| = R$ liegenden Zweige unserer Kurven, d. h. die Existenz einer Parameterdarstellung, noch nicht folgt, wie denn überhaupt der Beweis für die Existenz einer Parameterdarstellung *durch die Fortsetzung im Kleinen* unter Umständen auf große Schwierigkeiten stoßen kann, falls man als einen solchen Parameter nicht eine der Koordinaten oder die Bogenlänge der Kurve benutzen kann. Und die Tatsache, dass es sich bei den Zweigen einer algebraischen Kurve um einfache Linien handelt, die sich also etwa aus endlich vielen regulären Bogen zusammensetzen, darf selbstverständlich nicht den Erfahrungen entnommen werden, die man bei der Untersuchung von Kurven etwa der ersten sechs Ordnungen gesammelt hat.

Demnach hat man zur Ergänzung des ersten Gaussschen Beweises zu zeigen, dass man jede algebraische Kurve oder auch nur die im Gaussschen Beweis vorkommenden Kurven in endlich viele reguläre Bogen zerlegen kann, die sich durch eine Gleichung von einer der beiden Formen

$$y = g(x) \ , \ x = g(y)$$

darstellen lassen, wobei die Funktion $g$ im Innern des betrachteten Intervalls stetig und beschränkt differenzierbar ist.

> une démonstration dont on ne pourra douter. » Plus loin Gauss dit que dans le cas présent si les courbes ne s'interrompent pas, cela provient aussi du fait qu'elles séparent l'un de l'autre des domaines de signes différents de $Re(f(z))$ resp. de $Im(f(z))$.
>
> Il faut maintenant remarquer ici que la mise en place exacte de cette dernière considération a besoin de conceptions topologiques minutieuses. Mais surtout, la preuve qu'un tel comportement est impossible ne serait pas suffisante dans la mesure où la possibilité d'une « progression » le long des branches de notre courbe se trouvant à l'intérieur de $|z| = R$, c'est-à-dire l'existence d'une représentation paramétrique, ne s'en déduit pas, comme en général la preuve de l'existence d'une représentation paramétrique *par le parcours local* peut éventuellement rencontrer de grandes difficultés dans le cas où l'on ne peut pas utiliser comme paramètre l'une des coordonnées ou la longueur d'arc de la courbe. Et le fait qu'il s'agisse dans le cas des courbes algébriques de lignes simples qui se composent d'un nombre fini quelconque d'arcs réguliers, ne doit pas être déduit comme allant de soi de l'expérience que l'on a accumulée lors de recherches sur les courbes des six premiers degrés.
>
> Par conséquent on doit montrer en complément de la première preuve de Gauss que l'on peut découper toute courbe algébrique, ou du moins seulement celles présentes dans la démonstration de Gauss, en un nombre fini d'arcs réguliers qui peuvent être représentés par l'une des deux formes
>
> $$y = g(x) \ , \ x = g(y)$$
>
> où la fonction $g$ est continue et différentiable dans les intervalles considérés. [Ostrowski, 1920] p. 5-6

Ostrowski pointe là le problème principal de ce raisonnement géométrique : il s'agit en fait d'un problème de géométrie différentielle qu'il résout par l'application d'une forme de théorème des fonctions implicites.

### 13.4.4 Caractéristique de trois fonctions de deux variables

À partir du *Fortgangsprincip* et de la proposition précédente, Kronecker déroule un certains nombre de calculs qui serviront de fondements à sa définition de la caractéristique. On est donc dans la situation suivante : soient $f, \varphi$ et $\psi$ trois fonctions algébriques de deux variables $x$ et $y$ formant des courbes fermées. Nous allons parcourir l'une de ces trois fonctions en examinant les points d'intersection avec une seconde. On rappelle les notations suivantes :

$$\Delta = \begin{vmatrix} f & \varphi & \psi \\ f_1 & \varphi_1 & \psi_1 \\ f_2 & \varphi_2 & \psi_2 \end{vmatrix}$$

où $f_1 = \dfrac{df}{dx}$ et $f_2 = \dfrac{df}{dy}$.

Dans le déterminant précédent, si l'on considère $f, \varphi$ et $\psi$ comme des variables et que l'on développe ce déterminant par rapport à la première ligne, on obtient :

$$\Delta = f\begin{vmatrix}\varphi_1 & \psi_1\\ \varphi_2 & \psi_2\end{vmatrix} - \varphi\begin{vmatrix}f_1 & \psi_1\\ f_2 & \psi_2\end{vmatrix} + \psi\begin{vmatrix}f_1 & \varphi_1\\ f_2 & \varphi_2\end{vmatrix}$$

et donc

$$\frac{\partial\Delta}{\partial f} = \Delta_1 = \begin{vmatrix}\varphi_1 & \psi_1\\ \varphi_2 & \psi_2\end{vmatrix}$$

$$\frac{\partial\Delta}{\partial \varphi} = \Delta_2 = -\begin{vmatrix}f_1 & \psi_1\\ f_2 & \psi_2\end{vmatrix}$$

$$\frac{\partial\Delta}{\partial \psi} = \Delta_3 = \begin{vmatrix}f_1 & \varphi_1\\ f_2 & \varphi_2\end{vmatrix}$$

On applique la propriété précédente pour $U = f$ et $V = \varphi.\psi$

Déjà, on voit que cela est possible car si $\varphi = 0$ et $\psi = 0$ sont des courbes fermées, il en est bien sûr de même pour $\varphi.\psi$. On obtient :

$$\begin{aligned}
&\sum_{f=\varphi.\psi=0} \text{sgn } (f_1(\varphi_2\psi + \varphi\psi_2) - f_2(\varphi_1\psi + \varphi\psi_1)) = 0\\
\Leftrightarrow &\sum_{f=\varphi.\psi=0} \text{sgn } (\psi(f_1\varphi_2 - f_2\varphi_1) + \varphi(f_1\psi_2 - f_2\psi_1)) = 0\\
\Leftrightarrow &\sum_{f=\varphi.\psi=0} \text{sgn } (\psi\Delta_3 - \varphi\Delta_2) = 0\\
\Leftrightarrow &\sum_{f=\varphi=0} \text{sgn } (\psi\Delta_3) - \sum_{f=\psi=0} \text{sgn } (\varphi\Delta_2) = 0
\end{aligned}$$

donc

$$\boxed{\sum_{f=\varphi=0} \text{sgn } (\psi\Delta_3) = \sum_{f=\psi=0} \text{sgn } (\varphi\Delta_2)}$$

Par ailleurs, on a

$$\sum_{f=\varphi\psi=0} \text{sgn } (d(\varphi.\psi)) = 0$$

On peut séparer en deux cette somme : la partie où $\varphi = 0$ et la partie où $\psi = 0$. Mais il faut se souvenir que le *Fortgangsprincip* doit être appliqué, et donc que les sens de circulation sur $f = 0$ dans $\sum_{f=\varphi=0}$ et $\sum_{f=\psi=0}$ sont inverses l'un de l'autre. Ainsi :

$$\sum_{f=\varphi=0} \text{sgn } (d(\varphi\psi)) - \sum_{f=\psi=0} \text{sgn } (d(\varphi\psi)) = 0$$

et donc

$$\boxed{\sum_{f=\varphi=0} \text{sgn } (d(\varphi\psi)) = \sum_{f=\psi=0} \text{sgn } (d(\varphi\psi))}$$

Ensuite, dans le cas d'une somme du type $\sum_{f=\varphi}$, le *Fortgangsprincip* nous dit que $\text{sgn } (d\varphi) = \text{sgn } (f_1\varphi_2 - f_2\varphi_1) = \text{sgn } \Delta_3$. Donc

$$\begin{aligned}
&\sum_{f=\varphi=0} \text{sgn } (\psi.\Delta_3) \\
&= \sum_{f=\varphi=0} \text{sgn } (\psi.d\varphi) \\
&= \sum_{f=\varphi=0} \text{sgn } (\psi.d\varphi + \varphi.d\psi) \\
&= \sum_{f=\varphi=0} \text{sgn } (d(\psi.\varphi))
\end{aligned}$$

et finalement

$$\boxed{\sum_{f=\varphi=0} \text{sgn } (d(\psi.\varphi)) = \sum_{f=\varphi=0} \text{sgn } (\psi.\Delta_3)}$$

Comme $\Delta = f\Delta_1 + \varphi\Delta_2 + \psi\Delta_3$, on a aussi

$$\boxed{\sum_{f=\varphi=0} \text{sgn } (\psi.\Delta_3) = \sum_{f=\varphi=0} \text{sgn } (\Delta)}$$

et

$$\boxed{\sum_{f=\psi=0} \text{sgn } (\varphi.\Delta_2) = \sum_{f=\psi=0} \text{sgn } (\Delta)}$$

On obtient aussi des formules symétriques aux précédentes en appliquant la propriété précédente pour $U = \psi$ et $V = f\varphi$.

On applique la propriété précédente pour $U = f$ et $V = \varphi$

On a alors :

$$\begin{aligned}
&\sum_{f=\varphi=0} \text{sgn } \Delta_3 = 0 \\
&\sum_{\substack{f=\varphi=0 \\ \psi<0}} \text{sgn } \Delta_3 + \sum_{\substack{f=\varphi=0 \\ \psi>0}} \text{sgn } \Delta_3 = 0
\end{aligned}$$

D'autre part,

$$\sum_{f=\varphi=0} \text{sgn } \Delta = \sum_{f=\varphi=0} \text{sgn } \Delta_3\psi = \sum_{\substack{f=\varphi=0\\ \psi>0}} \text{sgn } \Delta_3 - \sum_{\substack{f=\varphi=0\\ \psi<0}} \text{sgn } \Delta_3.$$

Et ainsi :

$$\sum_{f=\varphi=0} \text{sgn } \Delta = \sum_{f=\varphi=0} \text{sgn } \psi\Delta_3 = 2\sum_{\substack{f=\varphi=0\\ \psi>0}} \text{sgn } \Delta_3 = -2\sum_{\substack{f=\varphi=0\\ \psi<0}} \text{sgn } \Delta_3$$

#### 13.4.4.1 Définition de la caractéristique de 3 fonctions de 2 variables

La caractéristique du système $(f, \varphi, \psi)$ est l'entier relatif :

$$\chi(f,\varphi,\psi) = -\frac{1}{2}\sum_{f=\varphi=0} \text{sgn } \Delta$$

En rassemblant toutes les égalités précédentes, on peut donner l'ensemble des formes que prend la caractéristique du système $(f, \varphi, \psi)$ par le tableau que donne Kronecker [68] :

$$\chi(f,\varphi,\psi) = \left\{\begin{array}{|c|c|c|}\hline -\frac{1}{2}\sum\limits_{f=\varphi=0} sgn\ \Delta & -\frac{1}{2}\sum\limits_{\varphi=\psi=0} sgn\ \Delta & -\frac{1}{2}\sum\limits_{\psi=f=0} sgn\ \Delta \\ \hline \sum\limits_{\substack{\psi<0\\ f=\varphi=0}} sgn\ (f_1\varphi_2 - f_2\varphi_1) & \sum\limits_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1) & \sum\limits_{\substack{\varphi<0\\ \psi=f=0}} sgn\ (\psi_1 f_2 - \psi_2 f_1) \\ \hline -\frac{1}{2}\sum\limits_{\varphi=\psi=0} sgn\ d(f.\varphi) & -\frac{1}{2}\sum\limits_{\psi=f=0} sgn\ d(\varphi.\psi) & -\frac{1}{2}\sum\limits_{f=\varphi=0} sgn\ d(\psi.f) \\ \hline -\frac{1}{2}\sum\limits_{f=\psi=0} sgn\ d(f.\varphi) & -\frac{1}{2}\sum\limits_{\varphi=f=0} sgn\ d(\varphi.\psi) & -\frac{1}{2}\sum\limits_{\psi=\varphi=0} sgn\ d(\psi.f) \\ \hline\end{array}\right.$$

**Remarque** On arrive ainsi à une définition de la caractéristique que l'on pourrait presque qualifier d'arithmétique au regard de ce que fait Weber – qui découle d'un long enchaînement d'égalités. L'interprétation géométrique ne sera donnée qu'a posteriori.

#### 13.4.4.2 Interprétations graphique

Kronecker donne trois interprétations graphiques, ou plutôt *géométriques*, de ce qui précède

**Interprétation graphique 1** Considérons maintenant la partie du plan définie par $\varphi\psi < 0$, c'est-à-dire celle que Weber nomme « l'enclos » [69] et parcourons $f$ en regardant les points d'intersection avec $\varphi$ : un tel point sera qualifié *d'entrée* si on passe de l'extérieur à l'intérieur de l'enclos, c'est-à-dire si en ce point $d(\varphi.\psi) < 0$. De même, on le qualifiera de *sortie* si on passe de l'intérieur à l'extérieur de l'enclos, c'est-à-dire si en ce point

68. Voir [Kronecker, 1891a] p. 236.
69. *Binnenraum* [Weber, 1898] p. 341

$d(\varphi.\psi) > 0$. Si on note $\mathcal{E}$ le nombre d'entrées et $\mathcal{A}$ le nombre de sorties, on a

$$\begin{aligned}\chi(f,\varphi,\psi) &= -\frac{1}{2}\sum_{\varphi=f=0} sgn\ d(\varphi.\psi) \\ &= \frac{1}{2}\sum_{\varphi=f=0} sgn\ -d(\varphi.\psi) \\ &= \frac{1}{2}(\mathcal{E}-\mathcal{A})\end{aligned}$$

C'est-à-dire que la caractéristique du système $f, \varphi, \psi$ est la demi-somme de la différence entre le nombre d'entrées et de sorties de l'enclos. Les formules données par Kronecker montrent par ailleurs que ce nombre est indépendant du choix des deux fonctions pour former l'enclos.

**Interprétation graphique 2** Kronecker regarde la caractéristique du système $f\Delta_1, \varphi, \psi$. On sait, d'après le tableau précédent, que

$$\chi(f,\varphi,\psi) = \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1) = \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ \Delta_1$$

donc on a :

$$\begin{aligned}\chi(\Delta_1 f,\varphi,\psi) &= \sum_{\substack{\Delta_1 f<0\\ \varphi=\psi=0}} sgn\ \Delta_1 \\ &= \sum_{\substack{\Delta_1<0 \text{ et } f>0\\ \varphi=\psi=0}} sgn\ \Delta_1 + \sum_{\substack{\Delta_1>0 \text{ et } f<0\\ \varphi=\psi=0}} sgn\ \Delta_1 \\ &= \sum_{\substack{f>0\\ \varphi=\psi=0}} (-1) + \sum_{\substack{f<0\\ \varphi=\psi=0}} (+1)\end{aligned}$$

On peut donc dire que « la caractéristique de ce système donne l'excès des points d'intersection de $\varphi$ et $\psi$ à l'intérieur de $f = 0$ sur ceux à l'extérieur de $f = 0$ ».[70].

**Interprétation graphique 3** De

$$\begin{aligned}\chi(f,\varphi,\psi) &= \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1) \\ &= \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\Delta_1)\end{aligned}$$

70. Voir [Kronecker, 1881a] p. 166.

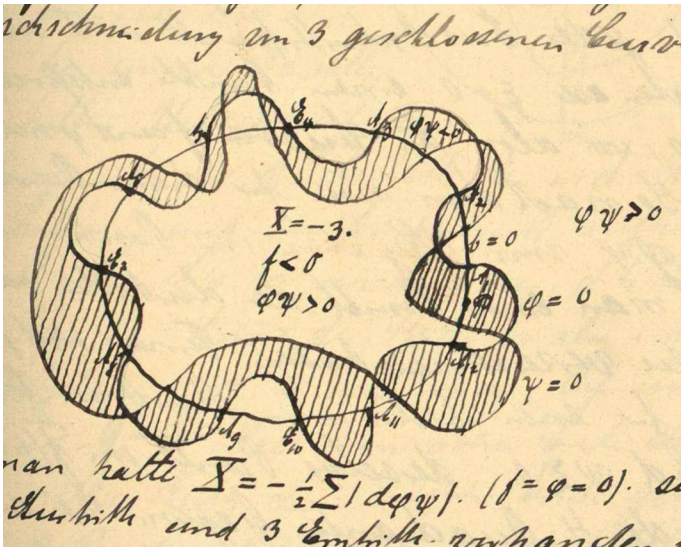



FIGURE 11 – Caractéristique de 3 courbes [Kronecker, 1879] p. 526

$$
\begin{aligned}
&= \sum_{\substack{f<0,\Delta_1>0\\ \varphi=\psi=0}} sgn\ (\Delta_1) + \sum_{\substack{f<0,\Delta_1<0\\ \varphi=\psi=0}} sgn\ (\Delta_1) \\
&= \sum_{\substack{f<0,\Delta_1>0\\ \varphi=\psi=0}} (+1) + \sum_{\substack{f<0,\Delta_1<0\\ \varphi=\psi=0}} (-1)
\end{aligned}
$$

on déduit que la caractéristique est le nombre de points d'intersection de $\varphi = 0$ et $\psi = 0$ à l'intérieur de $f = 0$ pour lesquels $\Delta_1$ est positif moins celui pour lesquels $\Delta_1$ est négatif. Autrement dit, la caractéristique est le nombre de solutions à l'intérieur du domaine $f = 0$ du système d'équations $\varphi = 0, \psi = 0$ pour lesquelles le déterminant fonctionnel est positif moins celui pour lesquelles le déterminant fonctionnel est négatif.

On mesure donc l'importance que revêt pour Kronecker l'interprétation graphique de sa théorie. On voit apparaitre deux types d'interprétation : les deux premières qui s'expriment uniquement en termes d'intersections de points, et la seconde dans laquelle intervient le signe du discriminant. Sur l'importance de ce dernier, nous renvoyons à la section sur la controverse entre Picard et Kronecker [71].

#### 13.4.4.3 Extension aux courbes qui ne sont pas fermées

Soient trois courbes d'équation $f = 0, \varphi = 0, \psi = 0$ qui ne sont pas supposées fermées. Comme le remarque Kronecker [72] :

71. Voir pp. 426 et suivantes.

72. Die vorstehenden Untersuchungen würden kaum nennenswerte Anwendungen zulassen, wenn sie nicht auf Funktionen übertragbar wären, welche der Bedingung, im Endlichen geschlossene Kurven zu repräsentieren, nicht mehr genügen. Es ist daher nötig, den Begriff der Charakteristik auch für solche Funktionensysteme als gültig nachzuweisen, welchen ausschließlich die Erfüllung gewisser Stetigkeits-Bedingungen auferlegt wird.

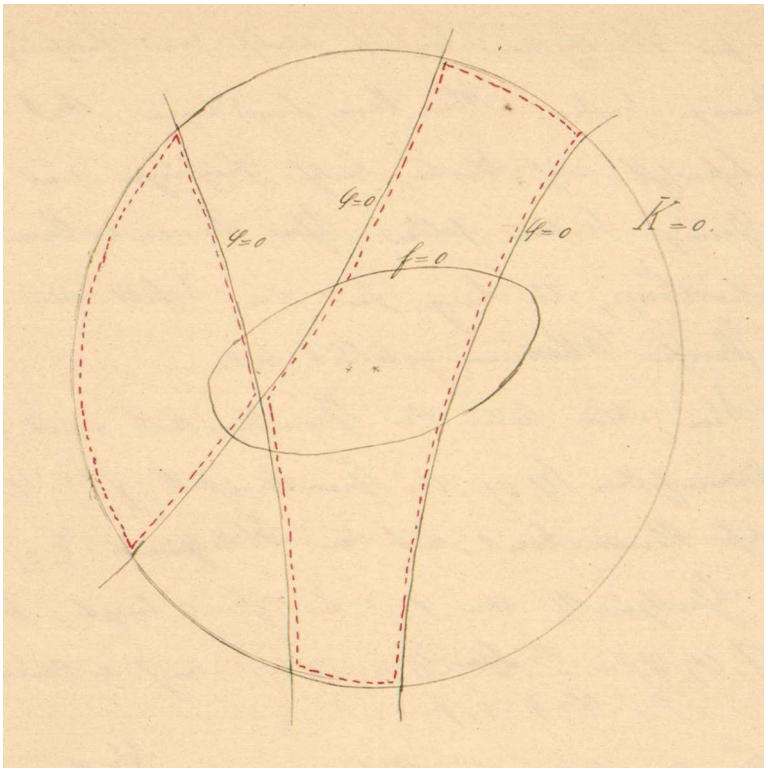



FIGURE 12 – Exemple de *fermeture* de $\varphi$ [Kronecker, 1891a] p. 238

> Les recherches ci-dessus ne permettraient guère d'emploi notable si elles n'étaient pas transférables aux fonctions qui ne sont pas représentables par des courbes finies fermées. Il est ainsi nécessaire de montrer que la notion de caractéristique est aussi valable pour un système de fonctions à qui on impose exclusivement le respect de certaines conditions de continuité. [Kronecker, 1891a] p. 237.

Contrairement à Weber, Kronecker donne une justification *arithmétique*. On nomme $K$ le cercle d'équation $K(x, y) = R^2 - x^2 - y^2 = 0$ (l'« intérieur » du cercle est ici la partie où $K > 0$). On choisit $R$ de façon à ce que le cercle renferme tous les points d'intersection de ces courbes.

On considère maintenant les courbes fermées $F = f.K = 0$, $\Phi = \varphi.K = 0$ et $\Psi = \psi.K = 0$. On définit $= \chi(f, \varphi, \psi)$ par :

$$\chi(F, \Phi, \Psi) = \chi(f, \varphi, \psi)$$

On a

$$\chi(F, \Phi, \Psi) = \sum_{\substack{F<0 \\ \Phi=\Psi=0}} sgn\ (\Phi_1\Psi_2 - \Phi_2\Psi_1) = \sum_{\substack{f<0 \\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1).K^2$$

Car les points d'intersections de $\Phi$ et $\Psi$ sont à situer par construction à l'intérieur de $K = 0$, et que dans ce même domaine $F < 0$ équivaut à $f < 0$.

Finalement

$$\begin{aligned}\chi(F,\Phi,\Psi) &= \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2-\varphi_2\psi_1).K^2\\ &= \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2-\varphi_2\psi_1)\end{aligned}$$

Par permutation des fonctions, on voit qu'au moins trois des énoncés du tableau précédent restent valables (en fait, l'ensemble des égalités reste valable).

#### 13.4.4.4 Modification de la caractéristique

Kronecker s'occupe de deux types d'« invariance » de la caractéristique :

Tout d'abord celle qui résulte d'une « transformation linéaire » du système $f,\varphi,\psi$. On a en effet, pour $a$ un réel :

$$\begin{aligned}\chi(af+b\varphi+c\psi,\varphi,\psi) &= -\frac{1}{2}\sum_{\varphi=\psi=0} sgn\ \begin{vmatrix} af+b\varphi+c\psi & \varphi & \phi\\ af_1+b\varphi_1+c\psi_1 & \varphi_1 & \phi_1\\ af_2+b\varphi_2+c\psi_2 & \varphi_2 & \phi_2\end{vmatrix}\\ &= -\frac{1}{2}\sum_{\varphi=\psi=0} sgn\ a\begin{vmatrix} f & \varphi & \phi\\ f_1 & \varphi_1 & \phi_1\\ f_2 & \varphi_2 & \phi_2\end{vmatrix}\\ &= sgn\ (a)\chi(f,\varphi,\psi)\end{aligned}$$

et il obtient finalement :

$$\chi(f,\varphi,\psi) = sign\ (|A|)\chi(A(f,\varphi,\psi))$$

où $A=\begin{vmatrix}\alpha & \beta & \gamma\\ \alpha' & \beta' & \gamma'\\ \alpha'' & \beta'' & \gamma''\end{vmatrix}$ est la matrice d'une application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$

Kronecker arrive donc à la propriété suivante [73] :

> Une caractéristique ne change pas sa valeur lors d'une transformation linéaire, et son signe change seulement lorsque le déterminant de substitution est négatif. [Kronecker, 1879] p. 544

Il utilise ensuite cette propriété pour examiner la façon dont se transforme la caractéristique lors d'une d'une variation continue des coefficients des fonctions $f,\varphi,\psi$ : il

73. Eine Characteristik ändert bei lineare Substitution ihnen Werth nie, und ihr Vorzeichen nur denn, wenn die Substitutiondeterminante negativ ist.

s'agit d'ailleurs de l'objet de l'article de Kronecker de 1878 *über die Charakteristik von Functionen-Systemen.*

Nous ne le démontrerons pas ici, mais ce qu'il faut retenir, c'est que pour que la caractéristique du système $f, \varphi, \psi$ se modifie lors d'un changement continu de ses coefficients, il faut que l'on *passe* par une valeur pour laquelle où $f = \varphi = 0$. Cette propriété est fondamentale dans l' utilisation qui sera faite de la caractéristique par Poincaré. Car nous verrons que cette invariance, qui est en fait une invariance par homotopie, est fondamentale dans ses articles traitant d'*analysis situs.*

## 13.5 Retour sur la notion de caractéristique

Dans son dernier cours sur la théorie des équations algébriques ([Kronecker, 1891a]), Kronecker termine son exposé sur la théorie des caractéristiques par un chapitre intitulé *Clarification de la notion de caractéristique, des racines réelles et de la décomposition d'une fonction entière en fonctions réelles.* [74]. Celui-ci commence par une « histoire du théorème fondamental de l'algèbre » que nous commenterons plus tard, dans laquelle il revient sur la notion de racine d'une équation.

### 13.5.1 La notion de racine d'une équation

Cette notion, Kronecker l'a déjà abordée dans son introduction, et nous en avons donné un aperçu [75] dans la première partie de notre travail. Après avoir expliqué en quoi Gauss avait résolu un grand nombre de questions liées à ce concept dans ses travaux sur le théorème fondamental, Kronecker pose le problème suivant qui, selon lui, n'a pas encore été résolu : « mais une chose reste non élucidée : la nature de la racine réelle. » [76]. La racine a-t-elle un statut géométrique ou arithmétique ? ces statuts sont-ils équivalents ? Kronecker affirme à propos de la notion de racine que Gauss lui-même « s'en tient à illustrer géométriquement sa signification et avait la ferme conviction d'avoir donné ainsi une définition suffisante. » [77]. Et il cite alors une remarque que fait Netto dans la traduction en allemand des preuves de Gauss qu'il venait à peine de publier : la première et quatrième preuve que Gauss donne du théorème fondamental de l'algèbre, consiste à [78]

> transférer sans plus d'examen la continuité géométrique sur des grandeurs arithmétiques, et donc, pour l'esquisser rapidement, faire correspondre à chaque distance entre deux points un nombre arithmétique. [Netto, 1876]

Ce qui intéresse ici plus particulièrement Kronecker, c'est le rapport entre le point de vue

74. [Kronecker, 1891a] p. 247 : « Klärung des Begriffs der Charakteristik, der reellen Wurzeln und der Zerlegung einer ganzen Funktion in reellen Funktionen. »

75. Voir p. 79.

76. [Kronecker, 1891a] p. 249 : « Aber eines blieb unaufgeklärt : die Natur der reellen Wurzeln ».

77. [Kronecker, 1891a] p. 249 : « textitGauss begnügte sich, ihre Bedeutung geometrisch zu veranschaulichen und war der festen Überzeugung, damit eine ausreichend Definition derselben gegeben zu haben. »

78. die geometrische Stetigkeit ohne weiteres auf die arithmetischen Größen übertragen, und also, um es kurz anzudeuten, einer jeden Streckenlänge eine arithmetische Zahl entsprechen lassen.

géométrique et arithmétique du nombre. Il rappelle d'ailleurs, dans une note en bas de page, ce que serait dans ce contexte un « nombre arithmétique » [79] :

> On doit comprendre par « nombres arithmétiques » des grandeurs sur lesquelles on peut appliquer les opérations arithmétiques de l'addition, la multiplication, la division, etc. et pour lesquelles la proposition suivante reste valable : si un produit est nul, alors l'un des facteurs doit toujours être nul. [Kronecker, 1891a] p. 250

On voit au passage l'importance de la notion d'intégrité dans l'arithmétique pour Kronecker. Ce qu'il dénonce, et ce sera finalement ce qui le différencie de Gauss, c'est-à-dire l'identification entre la notion géométrique de distance et la notion arithmétique de nombre [80] :

> Tous ces mathématiciens utiliseront cela avec d'autant peu de scrupules qu'ils considéreront aussi les grandeurs irrationnelles et imaginaires comme définies par des quantités géométriques. Que Gauss ait aussi été de ce dernier avis ressort le plus clairement d'une remarque qu'il a faite dans une leçon à la société de Göttingen. (imprimé dans le Göttinger gelehrten Anzeigen de mars 1831, étude 64.) [81]. Il cherche, à la fin de ses essais, à donner aux grandeurs imaginaires un fondement solide : « Ici la possibilité de donner une signification intuitive à $\sqrt{-1}$ est donc complètement justifié, et il n'en faut pas plus pour autoriser ces grandeurs dans le domaine des objets de l'arithmétique. » [Kronecker, 1891a] p. 250

Et finalement, la confusion s'accroît concernant le statut de ces nombres lorsque l'on est trompé par leur représentation géométrique (le plan pour les nombres complexes et la droite pour les nombres réels). C'est l'occasion pour Kronecker de rappeler son point de vue sur les interactions possibles entre géométrie et arithmétique [82] :

79. Unter « arithmetischen Zahlen » sind Größen zu verstehen, auf welche die arithmetischen Operationen der Addition, Multiplikation, Division u.s.w anwendbar sind und für welche der Satz gültig bleibt : Ist ein Produkt gleich Null, so muß jedenfalls einer der Faktoren gleich Null sein.

80. Das werden alle diejenigen Mathematiker mit um so weniger Bedenklichkeit thun, welche auch die irrationalen und imaginären Größen durch geometrische Quantitäten als « definiert » erachten. Daß Gauss auch dieser letzteren Ansicht war, geht aufs deutlichste aus einem Ausspruch hervor, welchen er in einer Vorlesung der Göttinger Societät gethun hat. (abgedruckt in den Göttinger gelehrten Anzeigen 1831, 64. März, Werke V). Er sagt dort am Schlusse dieses seines Versuchs, dem imaginären Größen eine feste Unterlage zu geben : « Hier ist also die Nachweisbarkeit einer anschaulichen Bedeutung von $\sqrt{-1}$ vollkommen gerechtfertigt, und mehr bedarf es nicht, um diese Größe in das Gebiet der Gegenstände der Arithmetik zuzulassen. »

81. Il s'agit en fait des *Göttingische gelehrte Anzeigen* du 23 avril 1831.

82. Die Mathematiker, welche eine solche, uns sehr bedenklich scheinende Ansicht zu der ihrigen machen, lassen dabei völlig außer Acht, daß zwar die arithmetischen Beziehungen auf die anschaulichen geometrischen Begriffe übertragbar und anwendbar sind, und daß die Geometrie freilich stets ganz wesentlich zum Fortschritt der arithmetischen Wissenschaften beigetragen hat und beitragen wird, daß aber niemals der rein « arithmetische » Begriff durch geometrisch Veranschaulichung definiert werden kann und definiert werden darf.

So glauben viele z.B. den Begriff der Quadratwurzel aus 2 als eine wirkliche arithmetische Zahl hinlänglich fixiert zu haben, wenn sie sagen, sie entspräche der Länge der Diagonale eines Quadrats, denn Seite als der Einheit gleich angenommenen würde. Sie vergessen dabei, daß sie auch hier schon vorausgesetzt haben, daß einem jeden Punkte der graden Linie bei willkürlich gewählter Maßeinheit eine arithmetische Zahl entspricht. Und um die Richtigkeit dieser letzteren Voraussetzung aufzuzeigen, ist für sie doch unbedingt wieder jene geometrische Begriffsbestimmung des Irrationalen unentbehrlich. Aber wie kann denn überhaupt ein geometrische Strecke einen arithmetischen Begriff « definieren » ? Richtet sich denn der Verstand nach den Sinnen, oder ist es nicht vielmehr der Verstand, welcher durch die Sinnenwelt ordnende Fäden legt, um sie so zu beherrschen ?

> Les mathématiciens qui font leur ce point de vue qui nous parait douteux ne tiennent aucun compte du fait que certes les rapports arithmétiques sont transférables et applicables aux concepts géométriques intuitifs et que certes la géométrie a contribué et contribuera toujours aux progrès des sciences arithmétiques, mais que jamais un pur concept « arithmétique » ne pourra et ne devra être défini par une illustration géométrique.
> Ainsi nombreux sont ceux qui pensent avoir par exemple suffisamment fixé la notion de racine carrée de 2 comme un nombre véritablement arithmétique lorsqu'ils disent qu'elle correspond à la longueur de la diagonale d'un carré dont le côté est égal à l'unité. Ils oublient à cette occasion qu'ici aussi ils ont déjà supposé qu'à tout point de la ligne graduée avec une unité arbitrairement choisie correspond à un nombre arithmétique. Et pour montrer la justesse de cette dernière supposition, cette définition géométrique des irrationnels leur est à nouveaux absolument nécessaire. Mais alors comment au juste un segment de droite géométrique peut « définir » un concept arithmétique ? Car la raison se conforme-t-elle aux sens, ou n'est-ce pas plutôt la raison qui tend à travers le monde sensible des fils qui permettent de le mettre en ordre ?
> [Kronecker, 1891a] p. 250

Pour Kronecker, arithmétique et géométrie ne sont pas incluses l'une dans l'autre, et si la géométrie est une « source d'inspiration » [83], il n'est pas question de définir un concept arithmétique à partir de celle-ci.

### 13.5.2 La notion de caractéristique

C'est pourquoi, au vu de l'importance que revêt cette notion pour Kronecker, il souhaite la rendre « indépendante de l'hypothèse de racines réelles des fonctions entières » [84]. Pour cela, Kronecker examine de nouveau le cas de la caractéristique de deux fonctions à une inconnue $\chi(V(x), W(x))$.

#### 13.5.2.1 *Rationalisation* du problème

Nous avons vu p. 349 que

$$\begin{aligned}\chi(V(x),W(x)) &= -\frac{1}{2}\sum_{V(x)=0} sgn \begin{vmatrix} V(x) & W(x) \\ V'(x) & W'(x) \end{vmatrix} \\ &= -\frac{1}{2}\sum_{V(x)=0} sgn \lim_{h=0} \begin{vmatrix} V(x) & W(x) \\ \dfrac{V(x+h)-V(x)}{h} & \dfrac{W(x+h)-W(x)}{h} \end{vmatrix}\end{aligned}$$

En simplifiant le déterminant, nous obtenons :

83. Voir p. 97.
84. [Kronecker, 1891a] p. 253 : « Wir wollen zunächst den Begriff der Charakteristik unabhängig machen von der Voraussetzung reeller Wurzeln ganzer Funktionen. »

$$\chi(V(x),W(x)) = -\frac{1}{2}\sum_{V(x)=0} sgn \lim_{h=0}\frac{1}{h}\begin{vmatrix} V(x) & W(x) \\ V(x+h) & W(x+h)\end{vmatrix}$$

et en sommant non seulement sur les $x$ pour lesquels $V(x) = 0$, mais aussi sur ceux pour lesquels $W(x) = 0$, on obtient :

$$\chi(V(x),W(x)) = \frac{1}{4}\sum_{V(x)=0,W(x)=0} sgn \lim_{h=0}\frac{1}{h}\left(W(x).V(x+h) - V(x).W(x+h)\right)$$

Le problème, pour le moment, c'est que nous nous servons des racines de $V(x)$ et $W(x)$. Mais Kronecker souhaite « supprimer l'utilisation de la notion de racine réelle dans cette dernière expression de la caractéristique » [85]. Il faut bien remarquer que c'est la notion de racine *réelle* qui pose problème, et que la difficulté reste celle de la nature de ces nombres réels.

Pour cela, Kronecker somme non plus sur les « racines », mais sur « toutes les valeurs rationnelles de $x$ » [86] pour lesquelles on a $V(x).V(x+h) < 0$ ou $W(x).W(x+h) < 0$. En fait, il part d'une valeur rationnelle $x_0$ inférieure à la plus petite des racines, puis il l'augmente de $h$ en $h$, où $h$ est une grandeur rationnelle. Tout l'art de Kronecker sera dans la détermination de cette grandeur.

On cherche ensuite à supprimer dans l'expression précédente le passage à la limite : il nous suffit de prendre $h$ suffisamment petit pour que dans tout intervalle de longueur $h$ il n'y ait au plus qu'une seule valeur pour laquelle $V(x).V(x+h) < 0$ et une seule valeur pour laquelle $W(x).W(x+h) < 0$. Cela n'a rien de difficile et Kronecker a déjà procédé ainsi lorsqu'il cherchait à *séparer* les racines [87].

Mais ne serions-nous pas en train d'effectuer une somme infinie, qui certes peut peut-être converger, mais n'est certainement pas calculable de façon effective ?

Non, car Kronecker a déjà montré dans son introduction [88] que nous pouvons nous restreindre à un intervalle borné lorsque nous examinons les racines d'un polynôme, ou plutôt, comme Kronecker le présente, lorsque nous regardons les changements de signe d'un polynôme. Il est intéressant de noter ici que Kronecker, pour décrire la méthode utilisant les racines du polynôme, parle de « langage habituel » ([Kronecker, 1891a] p. 255 : *gewöhnlichen Sprache*). Le langage qu'il utilise est donc *inhabituel*, et c'est donc dans celui-ci qu'il faut trouver l'originalité de son travail.

Cet intervalle borné est donc découpé en intervalles de longueur $h$ ne contenant qu'une seule valeur pour laquelle $W(x).V(x+h) < 0$ ou une seule valeur pour laquelle

85. [Kronecker, 1891a] p. 254 : « Wir können nun die Benutzung des Begriffs der reellen Wurzeln aus diesen letzten Ausdrücken für die Charakteristik ausmerzen ».
86. [Kronecker, 1891a] p. 254 : « über alle diejenigen rationalen Werte von $x$ ».
87. Voir p. 85.
88. Ibid.

$V(x).W(x+h) < 0$ : il n'y a donc qu'un nombre fini de ces valeurs, et Kronecker utilisera cette propriété pour justifier que la somme qu'il calcule est finie. Pour une telle valeur de $x$, on voit facilement que l'on a :

$$\frac{-\ sgn\ V(x).W(x+h) + \ sgn\ W(x).V(x+h)}{2}$$

$$= \ sgn\ \ (-V(x).W(x+h) + W(x).V(x+h))$$

Si maintenant on note $x_0, x_1, ..., x_l$ les rationnels $x_0, x_0+h, x_0+2h, ..., x_0+lh$, on obtient :

$$\chi(V(x), W(x)) = -\frac{1}{8}\sum_{k=0}^{l-1}\begin{vmatrix} sgn\ V(x_k) & sgn\ V(x_{k+1}) \\ sgn\ W(x_k) & sgn\ W(x_{k+1}) \end{vmatrix}$$

Kronecker a donc discrétisé le problème en le réduisant à une recherche parmi un nombre fini d'intervalles rationnels, dont la longueur sera aussi petite que l'on veut.

#### 13.5.2.2 Un exemple

Kronecker présente un exemple de ce qu'il vient de montrer, dans le cas, très simple, où $h = 1$ convient. Il donne la figure

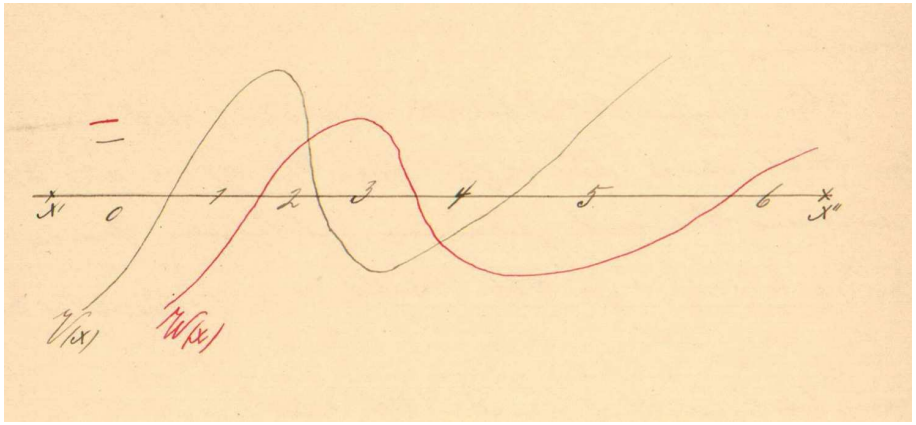



FIGURE 13 – [Kronecker, 1891a] p. 255

On obtient donc le tableau suivant :

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | |
|---|---|---|---|---|---|---|---|---|
| $V(x)$ | - | + | + | - | - | + | + | ; |
| $W(x)$ | - | - | + | + | - | - | + | |

et donc

$$\chi(V,W) = -\frac{1}{8}\left(\begin{vmatrix} -1 & +1 \\ -1 & -1 \end{vmatrix} + \begin{vmatrix} +1 & +1 \\ -1 & +1 \end{vmatrix} + ... \begin{vmatrix} +1 & +1 \\ -1 & +1 \end{vmatrix}\right)$$
$$= -\frac{1}{8}(2+2+2+2+2+2)$$

$$= -\frac{3}{2}$$

Kronecker termine ce paragraphe en montrant comment on peut en effet trouver un tel $h$. Ce long travail lui permet de conclure qu'il a déterminé [89]

> une expression de la caractéristique qui se laisse déterminer de façon strictement arithmétique, sans l'utilisation de racines réelles [Kronecker, 1891a] p. 259

## 13.6 Conclusion

Dans les *Vorlesungen über die Theorie der algebraischen Gleichungen*, Kronecker développe donc sa théorie des caractéristiques à la suite du théorème de Sturm. Il le fait tout d'abord à partir de la formule qu'il attribue à Sylvester et qui constitue, comme on l'a vu dans la partie précédente, l'élément central de son cours sur Sturm. Cette première exposition de sa théorie, nous la qualifierons de « topologico-algébrique » ou de « topologico-arithmétique ». D'un côté, on utilise le terme « topologique », car en effet, ce que Kronecker observe, ce sont les positions relatives de certaines courbes. Plus encore, à partir de 1878, il étudie le devenir de ces positions lorsque l'on produit une déformation continue de ces courbes. De l'autre, la notion d'« intersection » est considérée par Kronecker comme un concept qui appartient à la géométrie, et qui n'a pas sa place dans un travail qu'il veut « arithmétique » [90]. C'est le travail que nous avons décrit dans la section précédente, et qui lui permet d'*arithmétiser* la notion de racine [91], et par conséquent celle d'intersection. On voit cette arithmétisation de la notion de racine, année après année, avoir une place de plus en plus importante. Cependant, même dans le dernier cours de 1891, la construction est en deux temps : tout d'abord une présentation du contenu de la théorie en utilisant ce qu'il appelle le « langage habituel » [92], et ensuite des éclaircissements sur les notions de racine et de caractéristique [93]. Rendre son travail compatible avec son point de vue arithmétique ne doit pas gêner la compréhension : il a, dans ses cours, une position pragmatique qui lui permet de travailler avec des *coefficients réels* [94] ou des *racines réelles* [95]. À propos du théorème fondamental de l'algèbre, il affirme d'ailleurs dans [Kronecker, 1891a] qu'« il est pourtant aussi encore permis lors des recherches suivantes d'employer par souci de brièveté l'expression habituelle concernant les racines réelles » [96] : ce qui finalement compte pour Kronecker, c'est la *possibilité* de l'arithmétisation de la notion de racine. Il faut aussi voir dans cette affirmation une fois de plus le pédagogue pragmatique.

89. (...)einen Ausdruck für die Charakteristik erlangt haben, welcher sich streng arithmetisch(...)

90. On a vu dans la première partie son refus d'accepter le théorème de Bolzano et les conséquences sur la présentation de son travail. Voir p. 91.

91. En fait Kronecker montre comment se *passer* de la notion de racine. On retrouve une fois de plus la volonté de ce dernier de n'introduire uniquement des concepts nécessaires.

92. [Kronecker, 1891a] p. 255 : *gewöhnlichen Sprache*. Voir plus haut p. 391.

93. Ces remarques sont dans la sixième des sept sections qui composent son chapitre sur la théorie des caractéristiques dans son dernier cours de l'hiver 90-91.

94. *reellen Coeffizienten* [Kronecker, 1881a] p113 et [Kronecker, 1891a] p. 223.

95. « die reellen Wurzeln » [Kronecker, 1881a] p. 106 et [Kronecker, 1891a] p. 227.

96. [Kronecker, 1891a] p. 252 : « Es sei jedoch gestattet, auch noch bei den folgenden Untersuchungen der Kürze halber die gewöhnliche Ausdrucksweise in betreff der reellen Wurzeln zu verwenden ».

Sur cette première exposition, il n'y a pas de différences fondamentales entre la version du cours de 1878 et celle de celui de 1891 : il s'agit plus d'une différence formelle quant à l'utilisation, ou pas, du concept de racine. La théorie est dans ce sens le lieu du cours où Kronecker développe son point de vue sur les notions de nombre et de grandeur. Cependant, l'importance que revêtent certaines propriétés des caractéristiques change au cours des années. En particulier, la partie sur la modification de la caractéristique en fonction des coefficients – propriété centrale de son article de 1878 – finit par être la partie centrale de son cours. Nous allons voir au contraire que la seconde exposition de la théorie, qui occupe pourtant une partie importante de son article de 1869, est presque absente de ses derniers cours.

# Chapitre 14

# Le passage d'une théorie *topologico-arithmétique* à une théorie *topologico-analytique*


14.1 Le cours de Kronecker sur l'intégration . . . . . **397**
14.1.1 Définitions des intégrales . . . . . 398
14.1.2 Le Grand et le Petit . . . . . 400
14.1.3 Interprétation géométrique . . . . . 401
14.1.4 La notion de limite . . . . . 402
14.1.5 Sur la continuité . . . . . 404
14.1.6 Kronecker et les *théories modernes de l'intégration* . . . . . 405
14.2 L'indice de Cauchy . . . . . **406**
14.2.1 Présentation de la notion d'indice par Cauchy . . . . . 407
14.2.2 Les *Comptes rendus* de 1837 . . . . . 409
14.2.3 Une volonté d'éliminer le calcul intégral . . . . . 410
14.3 Le passage d'une théorie topologico-algébrique à une théorie topologico-analytique . . . . . **411**
14.3.1 Expression analytique . . . . . 414
14.3.2 Intégrale de Cauchy . . . . . 416
14.3.3 Étude de l'intégrale $\frac{1}{2\pi}\int_{f(x,y)=0} d \operatorname{arctg} \frac{\psi}{\varphi}$ . . . . . 416
14.4 La caractéristique de $n+1$ fonctions à $n$ variables . . . . . **422**
14.4.1 Rappels et notations . . . . . 422
14.4.2 Extension *topologico-algébrique* de la notion de caractéristique . 423
14.4.3 Extension *topologico-analytique* de la notion de caractéristique . 424
14.5 La controverse entre Picard et Kronecker . . . . . **426**
14.5.1 Les premières publications de Picard . . . . . 428
14.5.2 Les *Comptes rendus* du second semestre de 1891 : l'objet de la controverse . . . . . 430
14.5.3 Une résolution posthume . . . . . 437

14.5.4 Conclusion . . . . . 442

Dans son article de mars 1869, après avoir présenté la notion de caractéristique d'un système de fonctions comme nous venons de le voir, mais pour un système de $n+1$ fonctions à $n$ variables, Kronecker fait la remarque suivante à propos du travail de généralisation, et surtout d'interprétation du théorème de Sturm par Sylvester [1] :

> Les difficultés de la généralisation dont M. *Sylvester* fait mention doivent en grande partie être dues au fait qu'il s'est limité aux formes algébriques qu'il avait établies. Dès que je compris que toutes les considérations afférentes relevaient exclusivement de ce domaine général qui, pour le cas où $n = 2$ ou 3, est appelé "géométrie de position", les moyens les plus simples pour surmonter les difficultés rencontrées me sont apparus. [Kronecker, 1869c] p. 185

En effet, lorsque Sylvester expose sa théorie des *intercalations* [2], il affirme que sa théorie « peut être étendue à partir d'un système de deux fonctions générales de $x$ à un système de trois fonctions générales de $x, y$, de quatre fonctions générales de $x, y, z$, et en général à un système de $\epsilon$ fonctions générales de $\epsilon - 1$ variables » [3]. Mais pour justifier cette extension, il a recours aux « positions relatives » (*relative positions*) des courbes représentatives des fonctions, de plans d'équation du type $f(x, y) = 0$, de cônes, aux intersections de ces objets, etc. et il conclut que :

> And so, in like manner, but with a difficulty increasing at each step (as at next step we should have to pass into *quasi-space* of four dimensions), a theory of intercalations may be conjectured to exist for any $(n)$ general functions of any $(n - 1)$ variables. [Sylvester, 1853d] p. 496

Comme Kronecker le constate, Sylvester ne réussit pas à prolonger son travail de façon satisfaisante. Ces considérations géométriques qui semblent entraver sa progression sont qualifiées par Kronecker comme ne relevant finalement que de la « géométrie de position » (*Geometrie der Lage*). Il est peu probable que Kronecker fasse ici allusion aux travaux de 1847 de Karl von Staudt [4] ou à ceux qui en sont issus, et à sa géométrie basée sur les propriétés d'incidence des points [5]. L'expression est déjà utilisée par Gauss, par exemple dans [Gauss, 1900], où il est essentiellement question des nœuds que peuvent former des courbes fermées.

Pour résoudre ces difficultés, Kronecker propose donc de changer de point de vue, et

1. Die Schwierigkeiten der Verallgemeinerung, deren Herr *Sylvester* Erwähnung thut, dürften wohl zumeist in der Beschränkung auf algebraische Gebilde liegen, die er festgehalten hat. Sobald ich die Einsicht gewann, dass alle bezüglichen Betrachtungen ausschliesslich jenem allgemeinen Gebiete angehören, welches für den Fall, wo n = 2 oder 3 ist, als „Geometrie der Lage" bezeichnet wird, ergaben sich mir die einfachsten Mittel zur Bewältigung der entgegenstehenden Schwierigkeiten.

2. Voir [Sylvester, 1853d]. Il s'agit en fait de la théorie qui l'amène à obtenir une formule équivalente à celle que Kronecker lui attribue et sur laquelle, comme nous venons de le voir, il construit sa première exposition de la théorie des caractéristiques.

3. [Sylvester, 1853d] p. 495 : *(...)being extend from a system of two general functions of $x$, to a system of three general functions of $x, y$, four general functions of $x, y, z$, and in general to a system of $\epsilon$ general functions of $\epsilon - 1$ variables(...)*

4. Voir [von Staudt, 1847], sa traduction par Philippe Nabonnand [Nabonnand, 2011].

5. La *Geometrie der Lage*, qui s'apparente à ce que l'on appelle aujourd'hui la géométrie projective est présentée par exemple dans [Nabonnand, 2008].

il s'éloigne en fait de l'algèbre, pour travailler avec des variétés et des intégrales sur ces variétés. Avant de voir comment Kronecker s'échappe du mouvement d'« algébrisation du théorème de Sturm » pour prolonger sa théorie des caractéristiques, nous allons examiner la façon dont il envisage cette notion d'intégrale.

## 14.1 Le cours de Kronecker sur l'intégration

La notion d'intégrale est particulièrement importante lorsque nous abordons la théorie des caractéristiques. Il est alors indispensable pour l'appréhender correctement d'avoir quelques notions de la façon dont Kronecker la conçoit. C'est d'autant plus nécessaire que la théorie de l'intégration est à la fin du XIX$^{e}$ siècle dans ce que Thomas Hawkins [6] appelle une « période de transition » entre l'intégrale de Riemann et celles qui feront intervenir une théorie de la mesure.

Il y a ici deux questions principales auxquelles nous souhaiterions répondre : où se place Kronecker dans cette période de transition et comment sa philosophie des mathématiques, et en particulier la conception qu'il a du nombre et des grandeurs s'articulent-elles avec les concepts d'intégrale et d'intégration ?

Pour cela, nous avons à notre disposition une documentation particulièrement éclairante : son cours sur l'intégration. En effet, l'un des principaux cours que Kronecker donne tout au long de sa carrière d'enseignant à l'université de Berlin est une série de leçons sur l'intégration qu'il nommera *Vorlesungen über die Theorie der einfachen und der Vielfachen integrale*. Ce cours a été édité par E. Netto en 1894. Voyons d'ailleurs ce que ce dernier en dit dans l'introduction à cette édition [7] :

> Du cycle de leçons que Kronecker a présentées avec enthousiasme jusqu'à ses derniers jours à l'université de Berlin, nous exposerons ici celui sur la théorie des intégrales simples et multiples. Kronecker a exposé ce sujet à cinq reprises ; une première fois lors du semestre d'hiver 1883/84 pour des cours de deux heures intitulés "Sur quelques intégrales déterminées" ; ensuite durant le semestre d'été 1885 pour des cours de quatre heures, et en 1887, 1889, 1891 pour des cours de six heures. Pour la publication, les notes de Kronecker de la totalité des cinq cours ainsi que les notes complètes prises à son instigation en 1883/84, 1885, 1889 et la moitié de celles de l'année 1891 sont disponibles. [Kronecker, 1901] pp V-VI

Nous allons donc regarder la façon dont Kronecker introduit la notion d'intégrale. Pour cela, nous allons nous concentrer sur la première leçon dans laquelle il définit les intégrales simples. Voici le contenu de cette leçon [8]

6. Voir [Hawkins, 2001].

7. Aus dem Cyklus von Vorlesungen, welche der bis zum letzten Lebenstage schaffensfreudige Leopold Kronecker an der Berliner Universität gehalten hat, erscheint hier diejenige über die Theorie der einfachen und der vielfachen Integrale. Fünfmal hat Kronecker über diesen Gegenstand gelesen ; zuerst im Wintersemester 1883/84 zweistündig „über einige bestimmte Integrale" ; dann in den Sommersemestern 1885 vierstündig und 1887, 1889, 1891 sechsstündig. Für die Herausgabe lagen Kronecker's Notizen zu sämmtlichen fünf Collegien, sowie die auf seine Veranlassung angefertigten Nachschriften aus den Jahren 1883/84, 1885, 1889 völlig und die aus dem Jahre 1891 zur Hälfte vor.

8.

**Première leçon**

Deux définitions de l'intégrale. — Remarques historiques. — Partition quelconque de l'intervalle. — Notion de limite. — Continuité. — Condition pour la convergence des sommes. — Différentiation de l'intégrale. — Règles fondamentales. — Exemples. [Kronecker, 1901] pp VII

### 14.1.1 Définitions des intégrales

Kronecker donne deux définitions des intégrales. Il commence par rappeler celle que propose Euler du calcul intégral dans [Euler, 1768] [9] :

> Le calcul intégral est la méthode pour trouver une relation des quantités à partir d'une relation donnée des différentielles des mêmes quantités : et l'opération par laquelle cela est effectué est appelée habituellement intégration. [Kronecker, 1901] p. 1

Kronecker pose ensuite le problème qu'il cherche à résoudre dans son cours :

Soit $f(x)$ une fonction : on veut trouver $F(x)$ telle que

$$\frac{dF(x)}{dx} = f(x)$$

C'est-à-dire trouver une primitive de $F$.

De même, à partir de $f(x, y)$, on veut trouver $F(x, y)$ telle que

$$\frac{d^2F(x, y)}{dxdy} = f(x, y)$$

et ainsi de suite pour $n$ variables. On a donc un premier point de vue de l'intégration, qui consiste en une recherche de primitive : c'est un point de vue entièrement algébrique, qui évacue complètement la notion de grandeur.

Kronecker affirme ensuite que l'application de cette « partie de l'analyse » a été traitée par

— Gauss pour l'algèbre

— Dirichlet pour la théorie des nombres

— Cauchy pour la géométrie

Par contre, le titre que Dirichlet donne à ses leçons de 1842 *Theorie der bestimmen Integrale* ne convient pas à Kronecker. Il pourrait laisser croire que les bornes sont toujours des grandeurs déterminées (*bestimmte Grössen*), ce qui n'est pas le cas. Les contre-exemples qu'il donne montrent que Kronecker exclut l'infini et les variables de ce

**Erste Vorlesung**

Zwei Definitionen des Integrals. — Historische Bemerkungen. — Beliebige Theilung des Intervalles. — Grenzbegriff. — Stetigkeit. — Bedingung für die Convergenz der Summen. — Differentiation des Integrals. — Grundregeln. — Beispiele.

9. Calculus integralis est methodus, ex data differentialium relatione inveniendi relationem ipsarum quantitatum : et operatio, qua hoc praestatur, integratio vocari solet.

qu'il entend par « quantités déterminées ».

Kronecker propose ensuite un second point de vue. Soit une fonction $f(x)$ donnée, alors on peut chercher comme condition pour $F(x)$ :

$$(1) \qquad \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} = f(x)$$

Ce qui peut s'écrire :

$$(2) \qquad \frac{F(x+h)-F(x)}{h} = f(x)+\varphi(x,h)$$

où

$$\lim_{h\to 0} \varphi(x,h) = 0$$

Kronecker voit dans cette dernière égalité « une sorte de résolution algébrique » du problème posé : on a évacué le passage à la limite. Il rappelle aussi ici qu'une hypothèse indispensable sur la fonction $f$ est qu'elle doit être univoque (*eindeutig*).

Kronecker se place ensuite dans un **intervalle de valeurs réelles** (*Intervall reeller Werthe*) $[x_0; x]$ où l'égalité précédente sera valable. On pose alors

$$x - x_0 = nh$$

donc

$$h = \frac{x-x_0}{n}$$

On applique (2) à $x_0, x_0+h, ..., x_0+(n-1)h$ et on obtient

$$\begin{aligned} F(x_0+h)-F(x_0) &= hf(x_0)+h\varphi(x_0,h) \\ F(x_0+2h)-F(x_0+h) &= hf(x_0+h)+h\varphi(x_0+h,h) \\ &..... \\ F(x_0+nh)-F(x_0+(n-1)h) &= hf(x_0+(n-1)h)+h\varphi(x_0+(n-1)h,h) \end{aligned}$$

et donc

$$(4) \qquad F(x)-F(x_0) = \frac{x-x_0}{n}\left(\sum_{i=0}^{n-1} f\left(x_0+i\frac{x-x_0}{n}\right)+\sum_{i=0}^{n-1}\varphi\left(x_0+i\frac{x-x_0}{n},\frac{x-x_0}{n}\right)\right)$$

On peut ensuite prendre $h$ suffisamment petit (ou $n$ suffisamment grand) pour que les $\varphi$ restent plus petits qu'une valeur $\tau$ fixée à l'avance :

$$(4*) \quad F(x) - F(x_0) = \frac{x - x_0}{n} \sum_{i=0}^{n-1} f\left(x_0 + i\frac{x - x_0}{n}\right) + (x - x_0)\varepsilon\tau$$

où $0 \leq \varepsilon < 1$, et donc

$$(5) \quad F(x) - F(x_0) = \lim_{n \to +\infty} \frac{x - x_0}{n} \sum_{i=0}^{n-1} f\left(x_0 + i\frac{x - x_0}{n}\right)$$

On obtient ainsi une seconde définition de l'intégrale comme limite d'une somme.

### 14.1.2 Le Grand et le Petit

Kronecker rappelle que, historiquement, la conception de l'intégrale comme somme apparaît en premier (Archimède). Il s'agit de découper un intervalle en *petites parties*.

Kronecker s'interroge ensuite sur la signification du sens des termes « petit » et « grand » en mathématiques [10]

> Mais qu'est-ce pour les mathématiciens que "grand" ou "petit" ? Il est surprenant que l'on accepte si souvent comme évident le "petit" dans les sciences de la nature, lorsque l'on croit pouvoir expliquer par là le "grand". Cela rappelle les mots de Goethe :
>
> > Tu ne peux rien accomplir dans le grand
> > Et tu le commences alors dans le petit
>
> C'est ainsi que pour rendre l'attraction des masses compréhensible on suppose un éther d'attraction et on fait agir la force de petites parties sur petites parties ; ainsi la théorie de Darwin "explique" les grandes divergences qui apparaissent chez les individus d'un genre de nature organique, en enseignant comment elles résultent de petits changements.
> Mais comme lors des applications pratiques des mathématiques il importe toujours pour connaître un nombre qu'il se trouve à l'intérieur d'un intervalle déterminé dont la longueur dépend de la précision demandée, alors la méthode présentée nous sera utile aussi, comme nous le verrons, pour les intégrales. [Kronecker, 1901] pp. 3-4

10. Was ist aber für den Mathematiker „gross" oder „klein" ? Es ist überraschend, dass man in den Naturwissenschaften so oft das „Kleine" gern in den Kauf nimmt, wenn man sich das „Grosse" damit erklären zu können glaubt. Das erinnert an das Goethe'sche Wort :

> „Du kannst im Grossen nichts verrichten
> „Und fängst es nun im Kleinen an."

So meint man, die Massenattraction begreiflicher zu machen, wenn man einen Attractionsäther annimmt und die Kraft nun von Theilchen zu Theilchen wirken lässt ; so „erklärt" die Darwin'sche Theorie die grossen Abweichungen, welche bei den Individuen einer Gattung organischer Wesen auftreten, indem sie lehrt, wie dieselben aus kleinen Aenderungen hervorgehen.
Da es aber bei der praktischen Anwendung der Mathematik stets nur darauf ankommt, von einer Zahl zu wissen, dass sie innerhalb eines bestimmten Intervalles liegt, dessen Grösse von der zu verlangenden Genauigkeit abhängt, so wird uns auch die dargelegte Methode, wie wir sehen werden, in brauchbarer Weise zur Integralfunction verhelfen.

En effet, la méthode d'intégration peut se voir comme un découpage en intervalles de plus en plus petits que l'on examine séparément pour en déduire un résultat sur un intervalle plus grand. Il enchaine d'ailleurs en donnant quelques repères historiques sur l'origine de cette méthode de sommation [11]

> C'est d'après la méthode de sommation qui fournit l'intégrale que Leibniz l'a nommée "functio summatoria" et qu'il a introduit pour la désigner le signe somme $\int$. Chez Leibniz on peut suivre formellement dans chacun de ses stades la présentation du calcul intégral ; la grandeur $h$, le morceau $EF$ de la ligne de base deviennent chez lui plus petits d'année en année.
>
> Johann Bernoulli le premier a mis l'accent sur l'opération faisant passer d'un quotient différentiel donné à la fonction d'origine et l'a appelée ainsi "integrale" de "integer", l'"origine".
>
> Au fond, le calcul intégral vient essentiellement avant calcul différentiel ; seulement il offre de plus grandes difficultés que celui-ci et c'est pour cela que l'on suppose de préférence la connaissance du calcul différentiel.
>
> Dirichlet commence ses leçons par la définition de l'intégrale comme limite d'une somme comme nous l'avons déjà donnée en (5). Chez Riemann apparaît une définition en apparence autre, dans laquelle la distance $(x - x_0)$ est décomposée en parties qui ne sont plus supposées égales. D'ailleurs Riemann n'est pas le premier à avoir utilisé de telles fonctions, elles se trouvent déjà chez Gauss dans son traité sur la quadrature mécanique. [Kronecker, 1901] p. 4

Après avoir donc placé le calcul différentiel après le calcul intégral, il veut en donner une interprétation géométrique.

### 14.1.3 Interprétation géométrique

Il affirme en effet que [12] :

> Nous voulons maintenant discuter de cette sorte de définition et y appliquer une représentation géométrique. [Kronecker, 1901] p. 4

On se place ainsi dans un repère orthogonal et on considère la courbe d'équation

$$y = f(x)$$

11. Nach der Methode der Summation, die das Integral liefert, nannte es Leibniz „functio summatoria" und führte zur Bezeichnung derselben das Summenzeichen $\int$ ein. Bei Leibniz kann man die Erfindung der Integralrechnung förmlich in ihren einzelnen Stadien verfolgen ; die Grösse $h$, das Stück $EF$ der Grundlinie, wird bei ihm von Jahr zu Jahr kleiner.

Erst Johann Bernoulli legte den Hauptnachdruck auf die Operation des Zurückgehens von einem gegebenen Differentialquotienten auf die ursprüngliche Function und nannte diese daher „integrale", von „integer" das „Ursprüngliche".

Der Sache nach gehört die Integralrechnung eigentlich vor die Differentialrechnung ; nur bietet sie grössere Schwierigkeiten als diese, und daher setzt man die Kenntniss der Differentialrechnung lieber voraus.

Dirichlet beginnt seine Vorlesungen mit der Definition des Integrals als Grenzwerth einer Summe, wie wir sie in (5) gegeben haben. Bei Riemann tritt eine scheinbar noch weitere Definition auf, indem die Theile, in welche die Strecke $(x - x_0)$ zerlegt wird, nicht mehr als gleich vorausgesetzt werden. Uebrigens war Riemann nicht der erste, welcher solche Theilungen benutzte ; sie finden sich schon bei Gauss in der Abhandlung über mechanische Quadratur.

12. Wir wollen diese Art der Definition jetzt besprechen und uns dabei der geometrischen Repräsentation bedienen.

On découpe l'axe des abscisses de $x_0$ à $x_{2n}$ en $n$ parties $[x_0; x_2], [x_2; x_4], ..., [x_{2n-2}; x_{2n}]$. On peut construire $n$ trapèzes ayant pour sommets les points de coordonnées $(x_{2i}; 0); (x_{2i}; f(x_{2i})); (x_{2i+2}; f(x_{2i+2})); (x_{2i+2}; 0)$ pour $i = 0...n-1$.

On remarque que Kronecker ne suppose pas ici la fonction $f$ positive, ce qui pourrait entraîner quelques difficultés. Ce ne sont que des difficultés apparentes qui disparaissent par un choix correct des $x_i$, mais la suite de l'exposé montre que Kronecker suppose vraiment la courbe située au dessus de l'axe des abscisses.

On peut trouver entre $x_{2i}$ et $x_{2i+2}$ un $x_{2i+1}$ tel que l'aire du trapèze $(x_{2i}; 0); (x_{2i}; f(x_{2i})); (x_{2i+2}; f(x_{2i+2})); (x_{2i+2}; 0)$ soit :

$$(x_{2i+2} - x_{2i})f(x_{2i+1})$$

L'aire de l'ensemble des trapèzes ainsi construite est

$$\sum_{i=0}^{n-1}(x_{2i+2} - x_{2i})f(x_{2i+1})$$

Il fera ensuite un « passage à la limite », dont il démontrera la convergence. Cette notion de *limite* nécessite bien sûr quelques explications, et nous allons les donner immédiatement.

### 14.1.4 La notion de limite

En effet, Kronecker annonce que [13] :

> À propos de tous les développements précédents on doit remarquer que l'hypothèse sous-jacente à notre façon de voir, c'est-à-dire que l'aire d'une surface se laisse correctement évaluer par un nombre, n'est pas exempte d'objection. C'est seulement lorsque nous connaissons d'emblée une fonction $F(x)$ dont la dérivée est égale à la fonction donnée $f(x)$ que nous pouvons affirmer l'existence d'une telle valeur limite. L'intuition géométrique ne doit pas nous faire considérer cette existence comme allant de soi. Si nous ne connaissons pas de fonction $F(x)$, alors
>
> $$\left[\sum_{0}^{n-1}(-x_{2\chi} + x_{2\chi+2)}f(x_{2\chi+1})\right]$$
>
> conduit seulement à des suites de valeurs numériques qui convergent l'une vers l'autre, et nous ne pouvons compter que sur celles-ci. [Kronecker, 1901] p. 7

Ainsi, le fait d'avoir une surface – même une surface compacte du plan – ne suffit pas à affirmer que l'intégrale, c'est-à-dire la limite de la somme en question, puisse être représen-

13. Zu allen den bisherigen Entwickelungen ist zu bemerken, dass die unserer Auffassung zu Grunde liegende Annahme, der Inhalt einer Fläche lasse sich genau durch Zahlen auswerthen, durchaus nicht frei von Bedenken ist. Nur wenn wir von vornherein eine Function $F(x)$ kennen, deren Ableitung gleich der vorgelegten Function $f(x)$ ist, können wir die Existenz eines solchen Grenzwerthes behaupten. Die geometrische Anschauung darf uns nicht dazu verleiten, diese Existenz als selbstverständlich anzusehen. Kennen wir eine Function $F(x)$ nicht, so führt (6) lediglich auf gegen einander convergirende Reihen von Zahlenwerthen, und mit diesen allein dürfen wir rechnen.

tée par un nombre (*Zahlen*). La série aura beau converger, cela ne signifie pas que sa limite *existera*, au sens où Kronecker l'entend : ce ne sera le cas que lorsque nous connaîtrons une primitive, qui nous donnera donc un moyen *effectif* de calculer cette limite. Il poursuit ainsi [14] :

> $\psi(m)$ s'appelle une fonction du nombre positif $m$, si une procédure est fixée, grâce à laquelle pour chaque nombre $1, 2, 3, ...m, ...$ la valeur de $\psi(m)$ peut être trouvée. L'équation
>
> $$\lim_{m=\infty} \psi(m) = 0$$
>
> a alors la signification suivante : "Soit un nombre positif $\tau$ aussi petit que l'on veut donné, alors il est possible de choisir un nombre $M$ suffisamment grand de sorte que pour chaque valeur de $m$ qui est supérieure à $M$ on ait $|\psi(m)| < \tau$." Le trait vertical signifie à la suite de Herr Weierstrass, que l'on prend la valeur absolue de la grandeur qui y est enfermée. Comme tout autre limite se ramène à zéro, on peut concevoir chaque limite sur la base des explications données ici. Ainsi par exemple pour
>
> $$\lim_{m=\infty} m \sin \frac{\nu\pi}{m} = \nu\pi$$
>
> en posant
>
> $$m \sin \frac{\nu\pi}{m} - \nu\pi = \psi(m)$$
>
> on obtient
>
> $$\lim_{m=\infty} \psi(m) = \lim_{m=\infty} \left( m \sin \frac{\nu\pi}{m} - \nu\pi \right) = \lim_{m=\infty} m \sin \frac{\nu\pi}{m} - \nu\pi = 0$$

14. $\psi(m)$ heisse eine Function der positiven Zahl $m$, wenn ein bestimmtes Rechnungsverfahren festgestellt ist, mittels dessen für jede Zahl $1, 2, 3, ...m, ...$ der Werth von $\psi(m)$ gefunden werden kann. Es besitzt dann die Gleichung

$$\lim_{m=\infty} \psi(m) = 0$$

folgende Bedeutung : „Wird eine beliebig kleine positive Zahl $\tau$ gegeben, so ist es möglich, eine Zahl $M$ so gross zu wählen, dass für jeden Werth von $m$, der $> M$ ist, $|\psi(m)| < \tau$ wird." Dabei bedeuten die Verticalstriche, wie immer im Folgenden, nach dem Vorgange des Herrn Weierstrass, dass der absolute Werth der eingeschlossenen Grösse zu nehmen ist. Da sich jeder andere Grenzwerth auf Null zurückführen lässt, so kann man bei jeder Grenze die hier gegebene Erklärung zu Grunde legen. So z. B. erhält man bei

$$\lim_{m=\infty} m \sin \frac{\nu\pi}{m} = \nu\pi$$

wenn man

$$m \sin \frac{\nu\pi}{m} - \nu\pi = \psi(m)$$

setzt,

$$\lim_{m=\infty} \psi(m) = \lim_{m=\infty} \left( m \sin \frac{\nu\pi}{m} - \nu\pi \right) = \lim_{m=\infty} m \sin \frac{\nu\pi}{m} - \nu\pi = 0$$

Es braucht wohl kaum erwähnt zu werden, dass unsere Definition der Grenze keineswegs ein beständiges Abnehmen der Function mit wachsendem $m$ voraussetzt. So ist

$$\lim_{m=\infty} \frac{\sin mx}{m} = 0,$$

obwohl die Function bei wachsendem $m$ auch zunehmen kann.

Die Zurückführung der Grenzwerthe auf 0 geschieht deshalb, um anzuzeigen, dass wir keine neuen Begriff einfuhren wollen. Durch den Limes soll keine neue Grösse definirt werden ; wir gebrauchen ihn nur, wenn er gleich einer bekannten Grösse ist.

> Il est à peine nécessaire de mentionner que notre définition de la limite ne suppose nullement une décroissance de la fonction avec un accroissement de $m$.
> Ainsi
> $$\lim_{m=\infty} \frac{\sin mx}{m} = 0,$$
> alors que la fonction peut aussi augmenter avec des $m$ croissants.
> L'attribution d'une limite à 0 permet donc de montrer que nous ne voulons introduire aucun nouveau concept. Par la limite aucune nouvelle grandeur ne devra être définie ; nous ne l'utiliserons que lorsque qu'elle sera égale à une grandeur connue. [Kronecker, 1901] pp. 8-9

Le passage à la limite n'est pas source de création de nouvelles « grandeurs ». Il s'oppose ici bien sûr à la définition des réels comme limites d'un processus, tel que le fait par exemple Cantor.

### 14.1.5 Sur la continuité

Kronecker suggère ensuite que « pour ne pas être contraint dans les prochains développements de nous interrompre de nouveau, nous abordons ici également encore quelques discussions sur la notion de continuité. » [15]. Et après avoir distingué [16] *continuité géométrique* et *continuité analytique*, il fournit donc une définition *analytique* de la continuité [17] :

> À la place de cette illustration géométrique, nous choisissons une définition purement analytique : " $f(x)$ est dite continue en une valeur déterminée $x$ si l'on peut déterminer une valeur $h$ non nulle, aussi petite que l'on veut mais finie, pour laquelle on a l'inéquation
> $$|f(x + h.\varepsilon) - f(x)| < \tau$$
> $$(-1 \leq \varepsilon \leq 1),$$
> où $\tau$ représente une grandeur donnée aussi petite que l'on veut." [Kronecker, 1901] p. 12

C'est bien la définition de la continuité que Bolzano donne dans [Bolzano *et al.*, 1905] ou Cauchy dans son cours d'analyse [18]. Si on a vu que Kronecker s'opposait sur de nombreux

15. [Kronecker, 1901] p. 11 : « Damit wir nicht gezwungen sind, spätere Entwickelungen wieder zu unterbrechen, schalten wir gleich hier noch einige Erörterungen über den Begriff der Stetigkeit ein ».

16. Pour plus de détail, voir p. 201.

17. Statt dieser geometrischen Veranschaulichung Wählen wir eine rein analytische Definition : „$f(x)$ heisst bei einem bestimmten Werthe $x$ stetig, wenn ein von Null verschiedener, beliebig kleiner, aber endicher Werth von $h$ bestimmt werden kann, für welchen die Ungleichung gilt :
$$|f(x + h.\varepsilon) - f(x)| < \tau$$
$$(-1 \leq \varepsilon \leq 1),$$
wo $\tau$ eine beliebig kleine, gegebene Grösse bedeutet."

18. La définition de la continuité de Cauchy dans son traité publié en 1821, *le Cours d'Analyse de l'École Royal Polytechnique* : « la fonction $f(x)$ sera, entre les deux limites assignées à la variable $x$, fonction continue de cette variable, si, pour chaque valeur de $x$ intermédiaire entre ces limites, la valeur numérique $f(x+a) - f(x)$ décroît indéfiniment avec celle de $a$. En d'autres termes, la fonction $f(x)$ restera continue par rapport à $x$ entre les limites données, si, entre ces limites, un accroissement infiniment petit de la variable produit toujours un accroissement infiniment petit de la fonction elle-même. »

points au travail de Bolzano, tous les deux participent à l'« arithmétisation de l'analyse », dont l'un des éléments fondateurs est justement cette définition de la continuité. Il poursuit en donnant la définition de la continuité uniforme [19] :

> Nous utilisons de plus les concepts et définitions suivants :
> "Une fonction $f(x)$ est dite uniformément continue dans un intervalle si, après l'adoption du $\tau$, un et un seul $h$ fini remplit les conditions ci-dessus pour tous les $x$ de l'intervalle.
> Une fonction est dite généralement uniformément continue si la taille totale de tout intervalle dans lequel la condition de continuité uniforme n'est pas remplie s'approche en même temps que $\tau$ de zéro, donc devient plus petite que n'importe quelle grandeur prédéterminée". L'expression "uniformément continue" n'est certes pas très heureuse, car en effet seule $y = ax+b$ est uniformément continue, mais nous voulons la conserver par habitude. [Kronecker, 1901] p. 12

On voit que Kronecker raffine ici sa définition de la continuité uniforme en l'élargissant à des cas limites, ce qui préfigure son travail sur les fonctions uniformément bornées. Il utilisera d'ailleurs cette notion dans ses *Vorlesungen* comme condition pour la convergence de certaines sommes [20]. Pour aborder la notion d'intégrale, Kronecker examine de près les concepts de limite, de convergence et de continuité et les modifie de façon à ce qu'ils soient en adéquation avec sa philosophie des mathématiques. Mais son approche reste très classique et sa définition de l'intégrale est à première vue la même que celle que l'on trouve dans la plupart des manuels de la seconde moitié du XIX$^{e}$ siècle. Examinons comment, à l'occasion d'un article publié en 1878, Kronecker lance pourtant une piste qui permettra à ceux qui la suivront d'avancer dans la construction de l'intégrale de Lebesgue.

### 14.1.6 Kronecker et les *théories modernes de l'intégration*

Comme nous le savons, l'une des origines des « théories modernes de l'intégration » [21] est le travail de Fourier sur les séries, qui repose en effet sur le fait que toute fonction $f(x)$ définie sur un intervalle $[-l; l]$ peut s'écrire :

$$f(x) = \frac{1}{2}a_0 + \sum_{n=0}^{\infty} \left\{ a_n \cos\left(\frac{n\pi x}{l}\right) + b_n \sin\left(\frac{n\pi x}{l}\right) \right\}$$

Pour obtenir les coefficients $a_n$ et $b_n$, Fourier intègre cette égalité, puis intervertit les signes $\sum$ et $\int$, c'est-à-dire les signes lim et $\int$, sans d'ailleurs se poser de questions sur la validité de cette interversion. La convergence de cette série vers la fonction $f(x)$ est

19. Wir benutzen ferner die folgenden Begriffe und Definitionen :
„Eine Function $f(x)$ heisst in einem Intervalle gleichmässig stetig, wenn nach Annahme des $\tau$ ein und dasselbe endliche $h$ die obige Bedingung für jedes $x$ des Intervalles erfüllt.
Eine Function heisst in einem Intervalle im Allgemeinen gleichmässig stetig, wenn die Gesammtgrösse aller Intervalle, in denen die Bedingung gleichmässiger Stetigkeit nicht erfüllt ist, sich mit $\tau$ gleichzeitig der Null nähert, also kleiner wird, als eine vorgegebene, beliebig kleine Grösse." Der Ausdruck „gleichmässig stetig" ist zwar nicht glücklich gewählt, weil eigentlich nur $y = ax + b$ gleichmässig stetig ist, doch wollen wir ihn als eingebürgert beibehalten

20. Voir [Kronecker, 1901] p 13.

21. On pourra lire *The origins of modern theories of integration* de T. Hawkins [Hawkins, 1980].

valable, d'après Fourier, pour toute fonction *arbitraire*. La généralité d'une telle fonction, que Fourier d'ailleurs ne définit pas vraiment, ne lui permet plus de conserver la définition de l'intégrale grâce aux primitives. Le concept dominant sera alors celui de l'« aire sous la courbe ». Ce passage d'un concept à l'autre est repris par Kronecker, comme nous l'avons vu, dans les premières pages de son cours.

Comme le montre Hawkins [22], c'est un travail sur les ensembles, qui conduira à la théorie de la mesure, laquelle modifie la façon dont certains mathématiciens abordent la théorie de l'intégration au début des années 1880. En 1892, un an après la mort de Kronecker, Camille Jordan donnera à la théorie de la mesure sa forme définitive [23], et lors des dix dernières années du XIX$^{e}$ siècle, la théorie de l'intégration elle-même subit des changements radicaux, qui finiront par aboutir à l'intégrale de Lebesgue.

Kronecker ne prendra évidemment pas part à cette partie de l'histoire de l'intégration : les questions qui y sont abordées concernent des ensembles qui n'ont aucun sens pour lui.

Mais revenons toutefois sur cette possibilité d'intervertir les signes lim et $\int$. Nous savons, tout au moins après avoir pu la comparer à l'intégrale de Lebesgue, que l'un des défauts principaux de l'intégrale de Riemann, tient à la non-préservation de l'intégrabilité d'une suite de fonctions lors du passage à la limite, sauf condition de convergence uniforme. Un travail sur ce phénomène, et en particulier la recherche de conditions moins fortes que la convergence uniforme est entrepris en parallèle du travail autour de la mesure des ensembles. Ces travaux tournent autour de la notion de *suites uniformément bornées*, et Kronecker exhibe [24] dès 1878 une condition sur ces suites lui permettant d'intervertir les signes lim et $\int$. Cet article est le début d'une réflexion qui conduit Lebesgue à exposer son fameux théorème de convergence dominée [25].

## 14.2 L'indice de Cauchy

Dans la partie de son cours sur le théorème de Sturm, Kronecker avait déjà remarqué que les résultats qui y sont exposés avaient « le caractère de la proposition de Cauchy » [26]. Dans la partie suivante consacrée à la théorie des caractéristiques, Kronecker donne alors deux applications de son travail : le théorème fondamental de l'algèbre, que nous regarderons en détail dans le prochain chapitre, et l'indice de Cauchy [27]. Il considère d'ailleurs celui-ci comme *faisant partie* de sa théorie, et il présente donc sa théorie comme une généralisation des résultats de Cauchy. Ce mouvement de généralisation, entamé après le très algébrique théorème de Sturm, se poursuit ainsi grâce au changement de point de vue dont nous venons de parler. Pour bien en saisir la substance, nous allons présenter cette notion d'indice.

22. Voir [Hawkins, 1980].
23. Voir [Jordan, 1892].
24. Voir [Kronecker, 1878a].
25. Pour plus de détails on pourra consulter le chapitre 5 de [Hawkins, 2001].
26. Voir p. 208.
27. Voir [Kronecker, 1873a] p. 182, [Kronecker, 1879] p 535 et [Kronecker, 1891a] p. 246.

### 14.2.1 Présentation de la notion d'indice par Cauchy

La notion d'indice apparaît pour la première fois dans un mémoire lu à Turin à l'Académie des sciences le 27 novembre 1831 ([Cauchy, 1831a]). Ce mémoire, essentiellement accessible sous une forme manuscrite [28], est résumé dans un article publié la même année dans le *Bulletin des Sciences de Férussac* [29]. Dans ce mémoire, Cauchy fait le lien entre calcul de résidus et indices, et propose ainsi, comme l'analysent Sturm et Liouville, une approche « fondée sur l'emploi des intégrales définies » [30]. Pour présenter rapidement [31] la notion d'indice, nous utiliserons ce mémoire ([Cauchy, 1831a]), son résumé ([Cauchy, 1831b]) et le mémoire publié à Turin en 1833 auquel nous avons accès par sa publication en 1837 dans le *Journal de l'école polytechnique* sous le titre *Calcul des indices des fonctions* ([Cauchy, 1837b]). Nous utiliserons aussi les deux notes de Cauchy du printemps 1837 dans les *Comptes rendus* ([Cauchy, 1837c] et [Cauchy, 1837d]) dans lesquelles Cauchy revient sur ces deux mémoires.

Voilà comment en mai 1837 dans les *Comptes rendus* [32], Cauchy rappelle sa définition de l'indice :

> Soient $x$ une variable réelle, et $f(x)$ une fonction de cette variable qui devienne infinie pour $x = a$. Si l'on fait croître $x$, la fonction $f(x)$ passera en devenant infinie du négatif au positif, ou du positif au négatif, ou bien elle ne changera pas de signe. La quantité +1 dans le premier cas, -1 dans le deuxième, 0 dans le troisième, est ce que l'on nomme l'indice de la fonction pour la valeur donnée $a$ de la variable $x$.
> J'appelle Indice intégral pris entre deux limites données $x = x_0$, $x = X$, la somme des indices correspondant à toutes les valeurs de $x$ qui rendent la fonction infinie entre ces limites, et je désigne par la notation $\mathcal{I}_{x_0}^{X}((f(x)))$. [Cauchy, 1837c] pp. 672-673

Nous verrons que cette notion d'indice est très proche de celle de *caractère* que Kronecker introduit dans son article de 1869, et que dans certains cas, l'indice intégral correspondra à la caractéristique d'un système de fonctions. Cauchy donne une seconde définition, qui est celle qui sera reprise par Sturm et Liouville dans [Liouville et Sturm, 1836] :

> L'indice intégral est aussi l'excès $\Delta$ du nombre de fois où la fonction $f(x)$ en s'évanouissant pour les différentes valeurs de $x$ entre les limites $x_0, X$, passe du positif au négatif sur le nombre de fois où elle passe en s'évanouissant du négatif au positif.

Au moyen de son *calcul sur les résidus*, Cauchy obtient alors :

Soit un contour [33] fermé $C$ et $f(z) = \varphi(x, y) + i\chi(x, y)$ alors le nombre $m$ de racines

28. Une traduction italienne d' Antonio Lombardi a été publiée en 1839 dans les *Memorie di Matematica e di Fisica della Societd Italiana delle Scienze residente in Modena.*
29. Voir [Cauchy, 1831b].
30. Voir [Liouville et Sturm, 1836] p. 280.
31. Pour approfondir la question, on peut lire [Smithies, 1997] qui donne une exposition très détaillée du mémoire de 1831.
32. Voir [Cauchy, 1837c].
33. Cauchy le note O,O',O"...

de $f(z) = 0$ à l'intérieur du contour sera égal à [34]

$$\frac{1}{2i\pi}\int_C \frac{f'(z)}{f(z)}dz$$

Ici Cauchy ne s'embarrasse pas de suppositions sur $f$, mais il faut bien sûr qu'elle soit dérivable par rapport à la variable complexe, donc extrêmement régulière. Comme par ailleurs il démontre que, si l'on écrit $f(z) = \varphi(s) + i\chi(s)$, alors en posant $c$ la longueur du contour il obtient :

$$m = -\frac{1}{2}\mathcal{I}_0^c\left(\left(\frac{\chi(s)}{\varphi(s)}\right)\right) = \frac{1}{2}\mathcal{I}_0^c\left(\left(\frac{\varphi(s)}{\chi(s)}\right)\right)$$

on peut exprimer l'indice intégral en fonction d'une intégrale.

Ensuite, dans l'article de 1837, mais aussi dans le mémoire de 1833, Cauchy définit l'indice d'une fonction de deux variables : l'indice de $f(x, y)$ entre $x_0$ et $X$ et entre $y_0$ et $Y$ est :

$$\mathop{}_{x_0}^{X}\mathcal{I}_{y_0}^{Y}\left((f(x,y))\right) = \mathcal{I}_{x_0}^{X}\left((f(x,Y))\right) - \mathcal{I}_{x_0}^{X}\left((f(x,y_0))\right) - \mathcal{I}_{y_0}^{Y}\left((f(X,y))\right) + \mathcal{I}_{y_0}^{Y}\left((f(x_0,y))\right)$$

Et le théorème auquel Cauchy aboutit [35] dans son mémoire de 1833 donne le nombre de solutions d'un système de deux équations à deux inconnues :

> Soit $f(x, y)$, $F(x, y)$ deux fonctions de $x, y$, qui restent continues entre les limites $x = x_0$, $x = X$ ; $y = y_0$, $y = Y$. Nommons $\varphi(x, y)$, $\Phi(x, y)$ les dérivées de ces fonctions relatives à $x$, et $\chi(x, y)$, $X(x, y)$ leurs dérivées relatives à $y$. Enfin soit $N$ le nombre des différents systèmes de valeurs de $x, y$, propres à vérifier les équations simultanées $f(x, y) = 0$, $F(x, y) = 0$, et comprises entre les limites ci-dessus énoncées. On aura
>
> $$N = \frac{1}{2}\left[\mathcal{I}_{x_0}^{X}\left((\Psi(x,Y))\right) - \mathcal{I}_{x_0}^{X}\left((\Psi(x,y_0))\right) - \mathcal{I}_{y_0}^{Y}\left((\Psi(x_0,y))\right) + \mathcal{I}_{y_0}^{Y}\left((\Psi(X,y))\right)\right]$$
>
> En supposant
>
> $$\Psi(x,y) = \frac{\Phi(x,y)\chi(x,y) - \varphi(x,y)X(x,y)}{F(x,y)}f(x,y)$$
>
> [Cauchy, 1837b] p. 432

Ce théorème est une extension de celui qui se trouve dans le mémoire de Turin de 1831 et qui concerne le cas particulier où $f$ et $F$ sont les parties réelles et imaginaires d'une fonction complexe.

34. Cauchy donne la formule

$$m = \frac{1}{2\pi\sqrt{-1}}\int_0^c \frac{f'(z)}{f(z)}\frac{dz}{ds}ds$$

et fait ainsi intervenir un élément de longueur $ds$. ([Cauchy, 1831b] p. 175)

35. Le huitième théorème du mémoire de 1833. [Cauchy, 1837b] p. 432.

Il est intéressant de noter que

$$\Phi(x,y)\chi(x,y) - \varphi(x,y)X(x,y) = \frac{\partial F}{\partial x}\frac{\partial f}{\partial y} - \frac{\partial F}{\partial y}\frac{\partial f}{\partial x} = \begin{vmatrix} \dfrac{\partial F}{\partial x} & \dfrac{\partial F}{\partial y} \\ \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \end{vmatrix},$$

c'est-à-dire le déterminant fonctionnel, apparaît dans ce calcul. On connait, comme Kronecker l'a bien montré lors de l'exposition du *Fortgangsprincip*, l'importance du signe de celui-ci lorsque l'on calcule la caractéristique d'un système de trois fonctions à deux inconnues, et donc lorsque l'on cherche à résoudre le problème auquel souhaite répondre le théorème précédent. Cauchy n'en tient pas compte et c'est l'une des raisons pour lesquelles ce théorème est faux dans le cas général.

### 14.2.2 Les *Comptes rendus* de 1837

Sturm et Liouville rédigent un article en 1836, *Démonstration d'un théorème de M. Cauchy, relatif aux racines imaginaires des équations*, qui traite le cas particulier du mémoire de 1831. Ils n'ont pas connaissance de celui de 1833, et ignorent donc le théorème précédent proposé par Cauchy. Ce dernier envoie donc une lettre à l'Académie [36], rappelant le contenu de ce mémoire et donnant le théorème que nous venons de citer.

Très rapidement [37] Sturm et Liouville rédigent une note sur ce théorème. Ils y présentent trois contre-exemples du théorème précédent, et montrent que ce qui fait fonctionner le cas particulier du mémoire de 1831, c'est le fait que dans ce cas on a

$$\frac{\partial F}{\partial x}\frac{\partial f}{\partial y} - \frac{\partial F}{\partial y}\frac{\partial f}{\partial x} = \left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2$$

qui est positif.

Ils proposent ensuite « un autre théorème qu'on peut substituer à celui de M. Cauchy » où ils ont posé

$$R = \frac{\partial P}{\partial x}\frac{\partial Q}{\partial y} - \frac{\partial P}{\partial y}\frac{\partial Q}{\partial x}$$

> Considérons un contour fermé ABC sur lequel $P$ et $Q$ ne s'annullent [38] jamais à la fois, et admettons que, dans l'intérieur de ce contour, les valeurs de $x, y$, qui annullent(sic) $P$ et $Q$, donnent à $R$ une valeur positive ou négative, mais différente de zéro.
> Parmi les solutions $(x, y)$ des équations $P(x, y) = 0$, $Q(x, y) = 0$, contenues dans l'intérieur de ABC, les unes pourront correspondre à une valeur positive, les autres à une valeur négative de $R$. Nous désignerons par $\mu_1$ le nombre de solutions de la

36. Lettre qui est publiée dans les *Comptes rendus* : [Cauchy, 1837c]
37. La lettre de Cauchy est publiée le 22 avril, la réponse de Sturm et Liouville le 15 mai.
38. Voir [Liouville et Sturm, 1837] p. 723. Les deux "l" lorsque l'on conjugue le verbe annuler semblent courants dans les articles de mathématiques de la fin de la première moitié du XIX$^{e}$ siècle.

première espèce, et par $\mu_2$ le nombre des solutions de la seconde espèce.
Cela posé, on aura

$$\Delta = 2(\mu_1 - \mu_2),$$

$\Delta$ représentant l'excès du nombre de fois où la fraction $\frac{P}{Q}$ passe du positif au négatif sur le nombre de fois où elle passe du négatif au positif, en s'évanouissant, quand on parcourt d'un mouvement continu le contour entier ABC, en allant des $x$ positives aux $y$ positives. [Liouville et Sturm, 1837] pp. 723-724

Dans ce théorème, on voit que les solutions du système sont comptées différemment en fonction du signe du Jacobien, comme c'est le cas dans la théorie des caractéristiques de Kronecker. Ce dernier considère d'ailleurs que faire apparaître ce signe lorsque l'on cherche à compter le nombre de solutions d'un système donne la « vraie nature des choses » [39]. Nous reparlerons de cela lorsque nous examinerons la controverse sur ce sujet qui opposera en 1891 Émile Picard à Kronecker.

Dans ses derniers travaux, en particulier dans son article de 1855 intitulé *Sur les compteurs logarithmiques* [40], Cauchy établit [41] les prémisses de ce que l'on appelle aujourd'hui le « principe de l'argument », principe qui permet de construire l'indice d'un point par rapport à un lacet, et d'exprimer cet indice en fonction du nombre de zéros d'une fonction à l'intérieur d'un lacet [42].

### 14.2.3 Une volonté d'éliminer le calcul intégral

Dans le mémoire de Turin 1831, Cauchy détermine donc le nombre de racines d'une fonction complexe grâce à des calculs d'intégrales. Sturm et Liouville en 1836, dans leur article concernant la *démonstration d'un théorème de M. Cauchy, relatif aux racines imaginaires des équations* [43] proposent eux une démonstration de ce théorème qui ne repose que sur « les premiers principes de l'Algèbre » [44]. Comme le montre Hourya Sinaceur, ces *principes premiers* [45] correspondent pour Sturm à « des notions *minimalement nécessaires* [...] à l'"analyse des équations" : continuité, dérivée, théorème de Bolzano, formule de Taylor, règle de Descartes, algorithme d'Euclide,etc... » [46]. L'expression *premiers principes de*

39. Voir [Kronecker, 1891b] p. 1012.
40. Voir [Cauchy, 1855b].
41. On pourra lire à ce sujet [Outerelo et Ruiz, 2009] p. 6 ainsi que le chapitre 3.6.4 du livre de Umberto Bottazzini et Jeremy Gray intitulé *Hidden Harmony—Geometric Fantasies The Rise of Complex Function Theory* ([Bottazzini et Gray, 2013]).
42. L'indice d'un lacet $\gamma$ est défini par

$$\mathrm{Ind}_\gamma(z) = \frac{1}{2\mathrm{i}\pi}\int_\gamma \frac{\mathrm{d}\zeta}{\zeta - z}.$$

et représente géométriquement le nombre de tours que fait $\gamma$ autour de $z$.
43. Voir [Liouville et Sturm, 1836].
44. Voir [Liouville et Sturm, 1836] p. 280.
45. À propos de ceux-ci, et en particulier dans le cadre du lien entre la théorie de Cauchy et le théorème de Sturm, on pourra lire [Sinaceur, 1991] pp. 111-123.
46. Voir [Sinaceur, 1991] p. 123.

*l'algèbre* apparaît aussi dans les écrits de Cauchy [47]. C'est une Algèbre synonyme de *simplicité*, de *méthodes élémentaires* qu'envisagent ici Sturm et Cauchy. Et donc, lorsque Sturm et Liouville souhaiteront *algébriser* le travail de Cauchy, ils essayent d'éliminer tout ce qui a trait aux indices, aux résidus, et en fait au calcul intégral.

L'article des *Comptes rendus* de 1837 où Cauchy réagit à la démonstration par Sturm et Liouville de son théorème a deux propos. Tout d'abord Cauchy souhaite montrer qu'il a déjà, dans un mémoire de 1833 publié à Turin, *généralisé* ce théorème, mais surtout il rappelle la chose suivante :

> Dans un mémoire présenté à l'Académie des Sciences de Turin le 17 novembre 1831, j'ai fait connaître un nouveau calcul qui peut être fort utilement employé dans la résolution des équations de tous les degrés. Mais, dans le mémoire dont il s'agit, les principes de ce calcul, que je nomme calcul des indices, se trouvent déduits de la considération des intégrales définies. Je me propose ici de démontrer comment on peut établir directement ces mêmes principes sans recourir à des formules de calcul intégral. [Cauchy, 1837c] p. 673

Cauchy revendique donc le fait d'avoir lui aussi utilisé dans sa démonstration de 1833 – démonstration que Sturm et Liouville ne connaissent pas – ces *principes premiers de l'algèbre*, et il ajoute même dans [Cauchy, 1837d] qu'il est « complètement d'accord avec eux sur l'utilité de résoudre par des principes élémentaires les questions relatives à la détermination du nombre des racines réelles ou imaginaires » [48].

Comme nous l'avons vu, la vision de l'Algèbre de Kronecker n'est certainement pas un retour aux « principes premiers ». Lorsqu'il examine le théorème de Sturm, ce qu'il met en avant, ce n'est pas sa simplicité mais son manque de généralité : il s'évertue, comme nous l'avons vu dans la partie précédente, à développer des méthodes de plus en plus complexes, à multiplier les points de vue dans le but d'accroître cette généralité. Nous allons voir que ce qui lui permet d'étendre le théorème de Cauchy dont nous venons de parler, c'est justement un retour au calcul intégral.

## 14.3 Le passage d'une théorie topologico-algébrique à une théorie topologico-analytique

Au chapitre XII de l'article de 1869 où Kronecker établit sa théorie des caractéristiques, il revient sur les étapes qui lui ont permis sa découverte [49] :

47. [Cauchy, 1829] p. 89 : « Quoi qu'il en soit, j'ai pensé qu'il serait utile, pour ceux qui se proposent de cultiver les sciences mathématiques, d'offrir ici des méthodes simples et générales, à l'aide desquelles on puisse déterminer le nombre des racines, soit réelles, soit imaginaires, d'une équation de degré quelconque, et les calculer approximativement, sans recourir à l'équation auxiliaire dont l'inconnue a pour valeurs les carrés des différences entre ces racines, et sans employer des notations étrangères à ceux qui ne possèdent que les premiers principes de l'Algèbre. »

48. Voir [Cauchy, 1837d] p. 6.

49. Bei den Untersuchungen, deren Entwicklungsgang ich hier in Umrissen dargelegt habe, bin ich vom *Sturm*'schen Satze ausgegangen. Eine Ausdehnung desselben auf Systeme von Gleichungen ist schon vor längerer Zeit von Hrn. *Hermite* angegeben worden, aber es kam mir überdies darauf an, das den Stürmischen Entwickelungen zu Grunde liegende Kettenbruchs- Verfahren selbst zu verallgemeinern und nachdem

> Dans les recherches dont j'ai esquissé ici le développement je suis parti du théorème de *Sturm*. Une extension de celui-ci aux systèmes d'équations a déjà été donnée depuis longtemps par M. *Hermite*, mais il m'importait en outre de généraliser le procédé même des fractions continues qui est au fondement des développements de *Sturm*, et après avoir fait cela, d'interpréter d'une manière naturelle les résultats plus généraux qui sont ainsi obtenus. Ce faisant, j'ai été conduit aux considérations qui forment les quatre premières sections de la présente notice et que j'ai autrefois communiquées de vive voix à mon ami *Weierstrass*. À cette occasion il m'a incité à poursuivre plus avant l'objet de ma recherche à partir du nouveau point de vue obtenu et en particulier dans la direction dans laquelle on obtiendrait non seulement une extension du théorème de *Sturm* mais aussi de la proposition de *Cauchy*. Les travaux que j'ai entrepris en ce sens et les résultats que j'ai obtenus se trouvent expliqués dans les parties V à X, et pour l'essentiel dans le même ordre que celui dans lequel ils me sont apparus dans le cours de ma recherche. [Kronecker, 1869c] pp. 205-206

Les recherches de Kronecker ont bien leur point de départ dans le travail sur le théorème de Sturm. Nous avons d'ailleurs vu dans la partie précédente ce que Kronecker entend par « généraliser le procédé même des fractions continues qui est au fondement des développements de *Sturm* ». Mais au cours d'une discussion avec Weierstrass, un nouveau prolongement, un nouveau point de vue se fait jour : étendre le travail de Cauchy et de Sturm dans une même théorie. Pour cela, l'outil que Kronecker utilise, et que l'on a présenté plus haut, c'est l'intégration. Il n'est pas surprenant que ce prolongement *analytique* de la théorie des caractéristiques ait pour origine un échange avec Weierstrass, lequel était particulièrement à même de faire un lien entre les deux théories.

Nous allons, essentiellement à partir du cours de 1890-91, étudier comment Kronecker jette un pont entre la caractéristique d'un système de trois fonctions à deux variables et la théorie de Cauchy.

Examinons le cas particulier du $a$) dans la figure 1. La caractéristique du système est ici :

$$\chi = \frac{1}{2}(E - A) = \frac{7-3}{2} = 2$$

On fait l'équivalent d'un changement de coordonnées :

$$\begin{cases} \xi = \varphi(x, y) \\ \eta = \psi(x, y) \end{cases}$$

dies geschehen war die dadurch erhaltenen allgemeineren Resultate naturgemäß zu interpretiren. Ich wurde hierbei auf jene Betrachtungen geführt, welche den Inhalt der ersten vier Abschnitte vorliegender Notiz bilden, und welche ich damals in mündlichen Unterhaltungen meinem Freunde Weierstrass mittheilte. Dabei wurde ich von ihm angeregt unter den erlangten neuen Gesichtspunkten den Gegenstand meiner Untersuchung weiter und namentlich in derjenigen Richtung zu verfolgen, in welcher nicht bloss eine Ausdehnung des Stürmischen sondern auch des *Cauchy*'schen Satzes erhalten würde. Die Arbeiten, welche ich darauf hin unternommen, und die Resultate, welche ich dabei erlangt habe, finden sich in den Abschnitten V bis X auseinandergesetzt und zwar im Wesentlichen in derselben Reihenfolge, wie 'sie sich mir bei der Untersuchung ergeben haben.

Kronecker ne pose aucune hypothèse dans son cours, mais il suppose implicitement que $F : (x, y) \mapsto (\varphi(x, y), \psi(x, y))$ est une bijection (voire un $C^1-$difféomorphisme pour respecter les hypothèses de l'utilisation du changement de variable dans le calcul intégral). On peut alors trouver deux fonctions $g$ et $h$ telles que :

$$\begin{cases} x = g(\xi, \eta) \\ y = h(\xi, \eta) \end{cases}$$

**Explication de la figure** *b*) On peut alors exprimer $f$ en fonction de $\xi$ et $\eta$ et la représenter dans le nouveau repère $(Q, \xi, \eta)$.

1. Dans ce nouveau repère, la courbe $\varphi = 0$ se retrouve sur l'axe des ordonnées et $\psi = 0$ sur l'axe des abscisses.
2. La partie qui correspond à $\varphi\psi < 0$ correspond aux quadrants II et IV.
3. Commençons par placer le point $O$ : on a $\varphi < 0$ et $\psi > 0$, il doit donc être dans le quadrant II.
4. On continue dans le sens défini par le *Fortgangsprincip* : on coupe $\psi = 0$, c'est-à-dire l'axe des abscisses.
5. On coupe ensuite $\varphi = 0$, c'est-à-dire l'axe des ordonnées, au point $E_1$.
6. On continue jusqu'à rejoindre le point $O$.

On peut maintenant compter le nombre de tours effectués autour de $Q$. Pour cela :

1. Il faut deux points d'entrée pour faire un tour
2. Chaque point de sortie « annule » un point d'entrée

Le nombre de tours autour de $Q$ est donc :

$$\frac{1}{2}(E - A) = \chi(f, \varphi, \psi)$$

Ici, on effectue deux tours autour de $Q$, qui est bien la valeur de la caractéristique du système.

Kronecker, pour aider l'auditeur à appréhender la notion de « nombre de tours », a une fois de plus recours à une « image » [50]

> Nous pouvons illustrer la définition du nombre de tours de la manière suivante. Soit un clou vertical au plan fixé en $Q$ et soit posé un fil dans la position de la courbe $f(x, y) = 0$ ; Si nous attrapons alors ce dernier à un endroit quelconque et que nous le tendons, alors il fera deux fois le tour du clou. [Kronecker, 1891a] p. 245

50. Wir können uns die Definition der Anzahl der Windungen in folgender Weise veranschaulichen. Es sei in $Q$ ein Nagel vertikal zur Ebene angebracht und in die Lage der Kurve $f(x, y) = 0$ ein Faden gelegt ; greifen wir dann den letzteren an einer beliebigen Stelle und ziehen ihn an, so wird er zwei Mal um den Nagel gewunden bleiben.

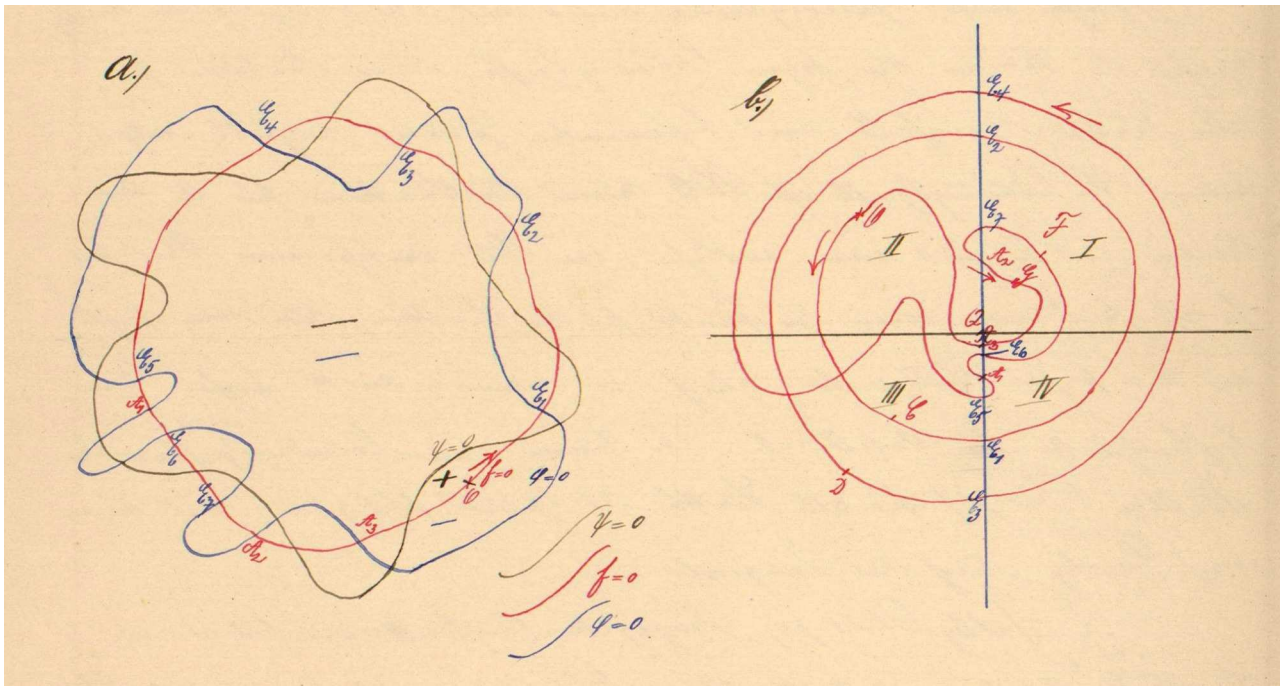



FIGURE 1 – [Kronecker, 1891a] p. 244

### 14.3.1 Expression analytique

Posons :

$$\begin{cases} \xi = r(v)\cos v \\ \eta = r(v)\sin v \end{cases}$$

Or on sait que le nombre de tours (*die Anzahl der Windungen*) est donné par

$$\chi(f,\varphi,\psi) = \frac{1}{2\pi}\int_{f(x,y)=0} dv$$

Par ailleurs :

$$\frac{\eta}{\xi} = \text{tg } v$$

et ainsi

$$v = \text{arctg}\,\frac{\eta}{\xi} = \text{arctg}\,\frac{\psi}{\varphi}$$

on a donc

$$\chi(f,\varphi,\psi) = \frac{1}{2\pi}\int_{f(x,y)=0} d\ \text{arctg}\,\frac{\psi}{\varphi} = \frac{1}{2\pi}\int_{f(x,y)=0}\frac{\varphi.d\psi - \psi.d\varphi}{\varphi^2+\psi^2}$$

Kronecker propose alors une méthode pour calculer, dans certains cas, la caractéristique

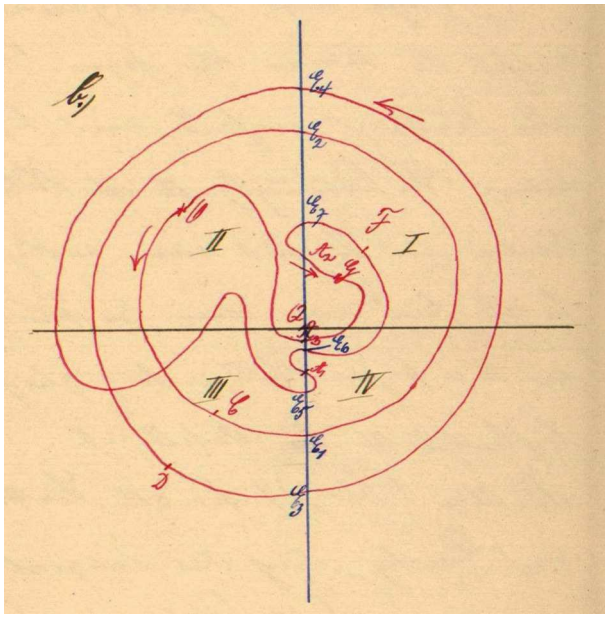



FIGURE 2 –

du système de trois fonctions à deux inconnues. Pour cela, on peut passer en coordonnées polaires en posant :

$$x = \rho.\cos\omega \ , \ y = \rho.\sin\omega$$

Et comme la caractéristique ne dépend que de « la position des courbes les unes par rapport aux autres » [51], on a alors $\varphi(x, y) = \Phi(\omega)$ et $\psi(x, y) = \Psi(\omega)$, c'est-à-dire que les deux fonctions ne dépendent plus que de $\omega$, dans le sens où $\rho = \rho(\omega)$.

Nous verrons dans la section suivante une utilisation de ce principe. On a donc :

$$\chi(f, \varphi, \psi) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\Phi\Psi' - \Psi\Phi'}{\Phi^2 + \Psi^2} d\omega.$$

Mais [52] :

> Comme la caractéristique est un nombre entier, alors il est seulement nécessaire de calculer l'intégrale dans la dernière formule exactement jusqu'à la première décimale pour obtenir la valeur de la caractéristique. [Kronecker, 1891a] p. 246

Cette méthode permettant de calculer de façon effective la caractéristique est simplement citée par Kronecker. Il est important pour lui qu'une méthode existe, mais ses qualités d'*effectivité* le sont beaucoup moins. Nous reviendrons sur ce sujet lorsque nous étudierons la controverse qui l'opposera à Picard.

51. [Kronecker, 1891a] p. 245 : « es nur auf die Lage der Kurven zu einander ankommt ».
52. Da die Charakteristik eine ganze Zahl ist, so ist es nur nötig, das Integral in der letzten Formel bis auf die erste Dezimale genau zu berechnen, um den Wert der Charakteristik zu erhalten.

### 14.3.2 Intégrale de Cauchy

Un cas particulier de ce qui vient d'être fait intéresse beaucoup Kronecker.

Posons

$$\Theta(x+iy) = \varphi(x,y) + i\psi(x,y)$$

Calculons

$$-i.d\log\Theta(x+iy)$$

$$\begin{aligned} -i.d\log\Theta(x+iy) &= -i\frac{d\varphi + id\psi}{\varphi + i\psi} = -i\frac{(d\varphi + id\psi)(\varphi - \psi i)}{\varphi^2+\psi^2} \\ &= -i.\frac{\varphi d\varphi + \psi d\psi}{\varphi^2+\psi^2} + \frac{\varphi d\psi - \psi d\varphi}{\varphi^2+\psi^2} \end{aligned}$$

donc

$$\frac{\varphi d\psi - \psi d\phi}{\varphi^2+\psi^2} = -i.d\log(\varphi+\psi i) + \frac{i}{2}d\log(\varphi^2+\psi^2)$$

et ainsi on a l'intégrale

$$\chi(f,\varphi,\psi) = \frac{1}{2i\pi}\int_{f(x,y)=0} d\log\Theta(x+yi) + \frac{i}{4\pi}\int_{f(x,y)=0} d\log(\varphi^2+\psi^2)$$

### 14.3.3 Étude de l'intégrale $\frac{1}{2\pi}\int_{f(x,y)=0} d\ \mathbf{arctg}\ \frac{\psi}{\varphi}$

Dans cette partie, on s'inspire en grande partie du travail d'Émile Picard [53].

**Étude d'une première intégrale** On calcule l'intégrale :

$$\frac{1}{2\pi}\int \frac{xdy - ydx}{x^2+y^2}$$

le long d'un contour fermé.

Intégrons-la par exemple le long du cercle $\mathcal{C}$ de centre $O$ et de rayon $R$ dans le sens positif :

$$\begin{cases} x = R\cos\theta \\ y = R\sin\theta \end{cases}, \theta \in [0;2\pi]$$

Alors :

$$\frac{xdy - ydx}{x^2+y^2} = \frac{(R\cos\theta)^2 + (R\sin\theta)^2}{R^2}d\theta = \frac{R^2}{R^2}d\theta = d\theta$$

53. Voir essentiellement [Picard, 1892], [Picard, 1891d], [Picard, 1891c], [Picard, 1891b] et [Picard, 1891a].

Et donc :

$$\boxed{\int_{\mathcal{C}} \frac{xdy - ydx}{x^2 + y^2} = \frac{1}{2\pi} \int_0^{2\pi} d\theta = 1}$$

Si maintenant on pose :

$$f(x, y) = \arctan\left(\frac{y}{x}\right)$$

On a

$$\begin{aligned} df &= \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \\ &= -\frac{y}{x^2} \left( \frac{1}{1 + \left(\frac{y}{x}\right)^2} \right) dx + \frac{1}{x} \left( \frac{1}{1 + \left(\frac{y}{x}\right)^2} \right) dy \\ &= \frac{-ydx}{x^2 + y^2} + \frac{xdy}{x^2 + y^2} \end{aligned}$$

donc

$$\frac{1}{2\pi} \int \frac{xdy - ydx}{x^2 + y^2} = \frac{1}{2\pi} \int d\left(\arctan \frac{y}{x}\right)$$

**Étude d'une seconde intégrale** Soit maintenant

$$I = \frac{1}{2\pi} \int d\left(\arctan \frac{\psi}{\varphi}\right)$$

où $\varphi(x, y)$ et $\psi(x, y)$ sont des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ développables en série de puissances croissantes au voisinage de tout point [54].

Posons $f(x, y) = \arctan\left(\frac{\psi(x,y)}{\varphi(x,y)}\right)$. On a

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

Or

$$\frac{\partial f}{\partial x} = \left( \frac{\frac{\partial \psi}{\partial x}\varphi - \frac{\partial \varphi}{\partial x}\psi}{\varphi^2} \right) \frac{1}{1 + \frac{\psi^2}{\varphi^2}} = \frac{\frac{\partial \psi}{\partial x}\varphi - \frac{\partial \varphi}{\partial x}\psi}{\varphi^2 + \psi^2}$$

et

$$\frac{\partial f}{\partial y} = \left( \frac{\frac{\partial \psi}{\partial y}\varphi - \frac{\partial \varphi}{\partial y}\psi}{\varphi^2} \right) \frac{1}{1 + \frac{\psi^2}{\varphi^2}} = \frac{\frac{\partial \psi}{\partial y}\varphi - \frac{\partial \varphi}{\partial y}\psi}{\varphi^2 + \psi^2}$$

donc

$$I = \frac{1}{2\pi} \int d\left(\arctan \frac{\psi}{\varphi}\right) = \frac{1}{2\pi} \int Pdx + Qdy$$

où

$$P = \frac{\frac{\partial \psi}{\partial x}\varphi - \frac{\partial \varphi}{\partial x}\psi}{\varphi^2 + \psi^2}$$

54. Par exemple une fonction $C^\infty$ dont la série de Taylor converge en tout point.

et

$$Q = \frac{\frac{\partial \psi}{\partial y}\varphi - \frac{\partial \varphi}{\partial y}\psi}{\varphi^2 + \psi^2}$$

On a alors

$$\boxed{\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}}$$

Rappelons que :

1. si $\gamma$ est un lacet à l'intérieur duquel $P$ et $Q$ sont continues et si $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ alors

$$\int_\gamma P dx + Q dy = 0$$

2. si $\gamma_1$ et $\gamma_2$ sont homotopes sur des domaines où $P$ et $Q$ sont continues, alors $\int_{\gamma_1} = \int_{\gamma_2}$

Soit $C$ un lacet. D'après ce qui précède, $P$ et $Q$ sont continues, sauf aux points où $\varphi = \psi = 0$. Si ce n'est jamais le cas dans $C$, c'est-à-dire qu'il n'y a aucune solution dans $C$ au système

$$(S) \begin{cases} \varphi(x, y) = 0 \\ \psi(x, y) = 0 \end{cases}$$

alors

$$\boxed{I = 0}$$

S'il y a une ou plusieurs solutions à ce système, on peut trouver des lacets qui contiennent une et une seule solution et tels que l'intégrale sur le lacet initial soit égale à la somme des intégrales sur ces lacets (figure 3)

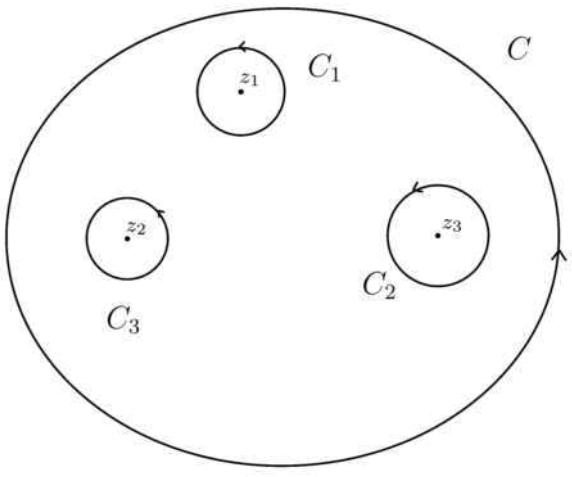



FIGURE 3 – $\int_C = \int_{C_1} + \int_{C_2} + \int_{C_3}$

Regardons ce qui se passe lorsque l'on intègre sur l'un de ces lacets, et supposons pour simplifier les calculs que $z_1 = (x_1, y_1) = (0, 0)$ est une solution du système. On intègre sur

un cercle de centre $O$ et de rayon $\rho$ suffisamment petit pour répondre aux conditions.

**Une propriété simplificatrice** On a choisi $\varphi$ et $\psi$ telles qu'elles puissent être développées en série des puissances croissantes de $x - x_0$ et $y - y_0$, donc ici de $x$ et $y$. De plus, comme $\varphi(0,0) = \psi(0,0) = 0$, on a :

$$\begin{cases} \varphi(x,y) = a_1 x + b_1 y + \sum_{r\geq 1, s\geq 1} c^1_{rs} x^r y^s \\ \psi(x,y) = a_2 x + b_2 y + \sum_{r\geq 1, s\geq 1} c^2_{rs} x^r y^s \end{cases}$$

On montre que le calcul de $I$ ne dépend que des termes de premier degré.

On a :

$$\begin{aligned} \frac{\partial \varphi}{\partial x}(x,y) &= a_1 + \sum_{r\geq 1, s\geq 1} r c^1_{rs} x^{r-1} y^s \\ \frac{\partial \psi}{\partial x}(x,y) &= a_2 + \sum_{r\geq 1, s\geq 1} r c^2_{rs} x^{r-1} y^s \\ \frac{\partial \varphi}{\partial y}(x,y) &= b_1 + \sum_{r\geq 1, s\geq 1} s c^1_{rs} x^r y^{s-1} \\ \frac{\partial \psi}{\partial y}(x,y) &= b_2 + \sum_{r\geq 1, s\geq 1} s c^2_{rs} x^r y^{s-1} \end{aligned}$$

On a alors

$$\frac{\partial \psi}{\partial x}\varphi - \frac{\partial \varphi}{\partial x}\psi = a_2 b_1 y - b_2 a_1 y + \text{termes de degré} \geq 2$$

et

$$\frac{\partial \psi}{\partial y}\varphi - \frac{\partial \varphi}{\partial y}\psi = a_1 b_2 x - b_1 a_2 x + \text{termes de degré} \geq 2$$

Or :

$$\begin{cases} \varphi^2 = a_1^2 x^2 + b_1^2 y^2 + 2a_1 b_1 xy + \sum_{r\geq 1, s\geq 1} d^1_{rs} x^r y^s \\ \psi^2 = a_2^2 x^2 + b_2^2 y^2 + 2a_2 b_2 xy + \sum_{r\geq 1, s\geq 1} d^2_{rs} x^r y^s \end{cases}$$

donc

$$\varphi^2 + \psi^2 = a_1^2 x^2 + b_1^2 y^2 + 2(a_1 b_1 + a_2 b_2)xy + \text{termes de degré} \geq 3$$

Ainsi

$$\begin{aligned} P dx &= \frac{a_2 b_1 y - b_2 a_1 y + \text{termes de degré} \geq 2}{a_1^2 x^2 + b_1^2 y^2 + 2(a_1 b_1 + a_2 b_2)xy + \text{termes de degré} \geq 3} dx \\ &= \frac{(a_2 b_1 y - b_2 a_1 y)dx + (\text{termes de degré} \geq 2)dx}{a_1^2 x^2 + b_1^2 y^2 + 2(a_1 b_1 + a_2 b_2)xy + \text{termes de degré} \geq 3} \end{aligned}$$

Si on fait le changement de variables

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \end{cases}$$

On a alors $dx = -\rho\sin\theta d\theta$ et :

$$Pdx = -\frac{\rho^2\sin^2\theta(a_2b_1 - b_2a_1)d\theta + \rho^3(\text{termes de degré} \geq 0)d\theta}{\rho^2(a_1^2\cos^2\theta + b_1^2\sin^2\theta + 2(a_1b_1 + a_2b_2)\cos\theta\sin\theta) + \rho^3(\text{termes de degré} \geq 0)}$$

donc

$$Pdx = -\frac{\sin^2\theta(a_2b_1 - b_2a_1)d\theta + \rho(\text{termes de degré} \geq 0)d\theta}{a_1^2\cos^2\theta + b_1^2\sin^2\theta + 2(a_1b_1 + a_2b_2)\cos\theta\sin\theta + \rho(\text{termes de degré} \geq 0)}$$

On montre de même que

$$Qdy = \frac{\cos^2\theta(a_1b_2 - b_1a_2)d\theta + \rho(\text{termes de degré} \geq 0)d\theta}{a_1^2\cos^2\theta + b_1^2\sin^2\theta + 2(a_1b_1 + a_2b_2)\cos\theta\sin\theta + \rho(\text{termes de degré} \geq 0)}$$

Or l'intégrale $I$ doit être indépendante[55] de $\rho$. Si on fait tendre $\rho$ vers 0, on obtient :

$$\int Pdx + Qdy = \int \frac{-\sin^2\theta(a_2b_1 - b_2a_1) + \cos^2\theta(a_1b_2 - b_1a_2)}{a_1^2\cos^2\theta + b_1^2\sin^2\theta + 2(a_1b_1 + a_2b_2)\cos\theta\sin\theta}d\theta$$

L'intégrale $I$ ne dépend que des termes de degré 1 des fonctions $\varphi$ et $\psi$.

**Calcul de $I$.** On calcule donc $I$ pour :

$$\begin{cases} \varphi(x,y) = a_1x + b_1y \\ \psi(x,y) = a_2x + b_2y \end{cases}$$

On obtient

$$I = \frac{D}{2\pi}\int \frac{xdy - ydx}{(a_1x + b_1y)^2 + (a_2x + b_2y)^2}$$

avec

$$\begin{aligned} D &= a_1b_2 - a_2b_1 \\ &= \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \end{aligned}$$

On calcule cette intégrale sur l'ellipse $(E)$ d'équation

$$(a_1x + b_1y)^2 + (a_2x + b_2y)^2 = 1$$

qui contient le point $O$ et qui est donc homotope à tout cercle de centre $O$[56].

55. En effet $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ et tous les cercles de centre $O$ sont homotopes sur $\mathbb{R}^2 - \{O\}$.

56. Si $(E)$ contient une autre solution du système $(S)$, on peut toujours trouver $k > 0$ tel que l'ellipse d'équation $k[(a_1x + b_1y)^2 + (a_2x + b_2y)^2] = 1$ n'en contienne qu'une.

On a alors $I = \frac{D}{2\pi} \int x dy - y dx$, et par passage aux coordonnées polaires

$$\begin{cases} x = \rho(\theta) \cos\theta \\ y = \rho(\theta) \sin\theta \end{cases}$$

on obtient

$$I = \frac{D}{2\pi} \int_0^{2\pi} \rho^2(\theta) d\theta$$

Or l'aire d'une ellipse ainsi paramétrée est

$$\frac{1}{2} \int_0^{2\pi} \rho^2(\theta) d\theta = \frac{\pi}{|D|}$$

Finalement

$$\boxed{I = \frac{D}{|D|} = \pm 1}$$

selon le signe de $D$, c'est-à-dire celui de

$$\begin{vmatrix} \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial y} \end{vmatrix}$$

On a :

$$\begin{cases} I = +1 \text{ si } D > 0 \\ I = -1 \text{ si } D < 0 \end{cases}$$

Et finalement : $\frac{1}{2\pi} \int_C \frac{\varphi d\psi - \psi d\varphi}{\varphi^2 + \psi^2}$ *est égale à : (nombre de solutions de* $(S)$ *dans* $C$ *où* $D > 0$*)* – *(nombre de solutions de* $(S)$ *dans* $C$ *où* $D < 0$*)*

où $C$ est un lacet fermé pris dans le sens positif.

S'il existe une fonction $f$ telle que $C = \{(x, y)/f(x, y) = 0\}$, le nombre $\frac{1}{2\pi} \int_C \frac{\varphi d\psi - \psi d\varphi}{\varphi^2 + \psi^2}$ est celui auquel Kronecker aboutit dans son cours [57] donné durant l'hiver 1881 à Berlin lorsqu'il cherche à calculer la caractéristique du système $f, \varphi, \psi$.

Cette intégrale est assez proche de celle utilisée par Gauss dans sa seconde démonstration [58] du théorème fondamental de l'algèbre, mais il lui manque les intégrales curvilignes [59] pour aboutir à une bonne interprétation de ses calculs [60].

57. Cours du semestre d'hiver 1881, manuscrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.
58. Voir [Gauss, 1816b].
59. Même s'il semble bien qu'il ait eu, avant que Cauchy ne l'introduise, des idées bien précises sur ce type d'intégrale ([Dieudonné, 1978] p. 133).
60. Voir plus loin p. 483.

## 14.4 La caractéristique de $n+1$ fonctions à $n$ variables

La notion de caractéristique de $n+1$ fonctions à $n$ variables n'est présentée que dans un seul des manuscrits en notre possession, celui du *Wintersemester* de 1878-1879. Cependant, c'est bien sous cette forme tout à fait générale que Kronecker la présente dans son article de 1869, et nous allons maintenant en donner un aperçu. Comme nous l'avons vu, cet article comporte deux grandes parties : la première présente la notion de caractéristique sous une forme que l'on a qualifiée de topologico-algébrique, et qui prolonge en quelque sorte l'algébrisation du théorème de Sturm ; la seconde généralise de son côté le travail de Cauchy. Nous allons donner les résultats auxquels Kronecker arrive.

### 14.4.1 Rappels et notations

Nous suivrons essentiellement le cours de 1878-1879 ([Kronecker, 1879]) et l'article de 1869 : nous sommes donc avant la publication des *Grundzüge*, et cela ne pose pas, à cette époque, de difficultés à Kronecker de prendre des fonctions « réelles univoques des $n$ grandeurs variables réelles illimitées » [61]. Kronecker parle dans son cours des « coefficients des fonctions », et considère donc que l'on souhaite travailler essentiellement avec des polynômes. Ces coefficients sont d'ailleurs « réels » [62].

Soit $F_{00}, F_{10}, ..., F_{n0}$ $(n+1)$ fonctions réelles définies sur $\mathbb{R}^n$, de classe $C^1$ et ayant les propriétés de la page 369. Pour tout $k \in [\![0, n]\!]$ et tout $h \in [\![1, n]\!]$ on note

$$F_{kh} = \frac{\partial F_{k0}}{\partial z_h}$$

Soit

$$\Delta = \begin{vmatrix} F_{00} & F_{01} & F_{02} & ... & F_{0n} \\ F_{10} & F_{11} & F_{12} & ... & F_{1n} \\ F_{20} & F_{21} & F_{22} & ... & F_{2n} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ F_{n0} & F_{n1} & F_{n2} & ... & F_{nn} \end{vmatrix}$$

61. [Kronecker, 1869c] p. 177 : « eindeutige reelle Functionen der $n$ reellen unbeschränkt veränderlichen Grössen ».

62. [Kronecker, 1879] p. 566 : « sind die coefficienten der Funktionen als reel angenommen. »

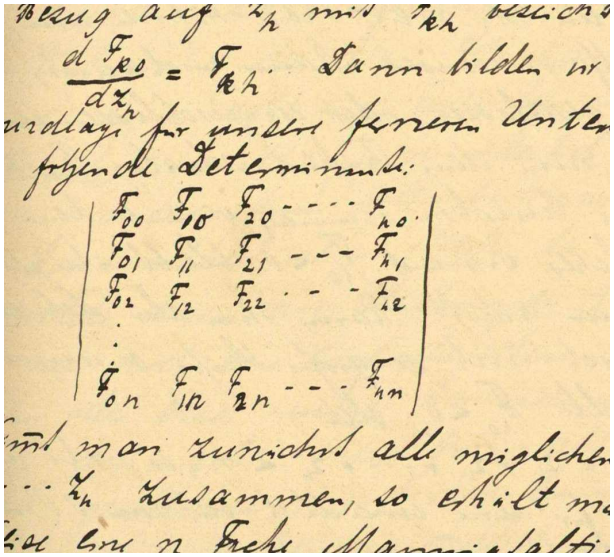


FIGURE 4 – [Kronecker, 1879] p. 566

et

$$\Delta_h = \frac{\partial \Delta}{\partial F_{h0}} = (-1)^{h+2} \begin{vmatrix} F_{10} & F_{11} & F_{12} & ... & F_{1,h-1} & F_{1,h+1} & ... & F_{1n} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ F_{n0} & F_{n1} & F_{n2} & ... & F_{n,h-1} & F_{n,h+1} & ... & F_{nn} \end{vmatrix}$$

On désignera par $[hk]$ ou $[kh]$ la *ligne* définie par

$$F_{00} = F_{10} = ... = F_{h-1,0} = F_{h+1,0} = ... = F_{k-1,0} = F_{k+1,0} = ... = F_{n0} = 0$$

selon le sens de parcours choisi.

### 14.4.2 Extension *topologico-algébrique* de la notion de caractéristique

Kronecker remarque que l'on travaille dans une variété de dimension $n$ [63] et que, pour tout $k$, $F_{k0} = 0$ forme une sous variété de dimension $n-1$ [64] qui sépare la variété en deux parties : l'intérieur (*Innere*) si $F_{k0} < 0$ et l'extérieur si $F_{k0} > 0$.

Kronecker suppose de plus que, pour tout $k$, $F_{k0} = 0$ est fermée (*geschlossen*), c'est-à-dire bornée, et donc il suppose qu'il existe $r > 0$ tel que pour tout $(z_1, ....z_n)$ où

63. [Kronecker, 1879] p. 566 : « eine $n$ Fache Mannigfaltigkeit ».
64. [Kronecker, 1879] p. 566 : « eine $n-1$ Fache Mannigfaltigkeit ».

$F_{k0}(z_1, ...., z_n) < 0$ on ait

$$z_1^2 + z_2^2 + .... + z_n^2 < r^2$$

De plus, l'ensemble $F_{i0} = 0/i \neq h, k$ forme une variété simple ou une courbe [65], qui sera donc fermée. Il s'agit pour Kronecker de regarder les points d'intersection de cette courbe avec la sous-variété $F_{h0} = 0$ de dimensions $n - 1$. Dans le manuscrit du semestre d'hiver 1878-1879, Kronecker poursuit en exposant, comme nous l'avons fait, le *Fortgangsprincip*.

On peut définir l'*intérieur* de la sous-variété $F_{h0}.F_{k0} = 0$ par $F_{h0}.F_{k0} < 0$. Cela correspond à ce que Weber a appelé l'enclos (*Binnenraum*). Si l'on parcourt $[hk]$, qui est fermée, il y aura autant d'« entrées » que de « sorties » de l'enclos. La somme des entrées et sorties de cet enclos est donc un nombre pair, et la caractéristique du système de fonctions sera en fait

$$\frac{1}{2}(\text{Nombre d'entrées} - \text{Nombre de sorties}).$$

Toute la difficulté sera de prouver que ce nombre est constant quels que soient le $h$ et le $k$ choisi. Pour cela, Kronecker montre, de façon similaire à ce qui a été fait précédemment dans le cas $n = 2$, que le nombre

$$\sum_{F_{g0}=\delta_{gm}.F_{m0} \text{ pour } g=0,1,...,n} sgn(\Delta)$$

est un invariant [66], quel que soit le nombre $m$ choisi. On remarque ici l'utilisation particulièrement rusée du « symbole de Kronecker ».

Il donne ensuite, pour la caractéristique du système, la définition suivante :

$$\chi(F_{00}, F_{10}, ..., F_{n0}) = -\frac{1}{2} \sum_{F_{g0}=\delta_{gm}.F_{m0} \text{ pour } g=0,1,...,n} sgn(\Delta)$$

et il obtient des formules équivalentes à celles trouvées précédemment, en particulier :

$$\chi(F_{00}, F_{10}, ..., F_{n0}) = \sum_{F_{00}=F_{10}=...=F_{m-1,0}=F_{m+1,0}=F_{n0}=0 \text{ avec } F_{m0<0}} sgn(|F_{ik}|)$$

où $i = 0, ...., m - 1, m + 1, ..., n$ et $k = 1, ..., n$, ce nombre étant invariant pour tout $m$ compris entre 1 et $n$.

### 14.4.3 Extension *topologico-analytique* de la notion de caractéristique

Ce que l'on vient de faire prolonge sans difficulté, dans le cadre que l'on a qualifié de topologico-algébrique, la caractéristique d'un système de trois fonctions à deux inconnues. Qu'en est-il pour la version que l'on a qualifiée d'« analytique » de cette théorie ?

65. [Kronecker, 1879] p. 568 : « einfache Mannigfaltigkeit / Curve ».
66. [Kronecker, 1879] p. 574 : « Invariant ».

Son prolongement au cas de $n$ fonctions est tout aussi naturel, et sera d'ailleurs, au contraire du reste de la théorie, largement repris par d'autres mathématiciens, quelquefois sous le nom de *Formule de Kronecker*. Celle-ci n'est pas exposée dans ses cours sur la théorie des équations algébriques, et même si elle constitue une bonne partie de ce que Kronecker expose dans ses articles de 1869, nous nous bornerons ici à rapidement exposer le résultat auquel Kronecker parvient [67]. La caractéristique peut ainsi s'exprimer en fonction d'une intégrale :

$$\chi(F_{00}, F_{10}, ..., F_{n0}) = -\frac{1}{\overline{\omega}_n} \int_{F_{m0}=0} \Delta \frac{dw}{\varphi^n \gamma}$$

où

$$\varphi = \sqrt{F_{10}^2 + ... + F_{n0}^2} \text{ et } \gamma = \sqrt{F_{m0}^2 + ... + F_{mn}^2}$$

et avec aussi

$$\overline{\omega}_n = \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}$$

$\Gamma$ étant la fonction gamma. On rappelle que

$$\Delta = \begin{vmatrix} F_{00} & F_{01} & F_{02} & ... & F_{0n} \\ F_{10} & F_{11} & F_{12} & ... & F_{1n} \\ F_{20} & F_{21} & F_{22} & ... & F_{2n} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ F_{n0} & F_{n1} & F_{n2} & ... & F_{nn} \end{vmatrix}$$

$\overline{\omega}_n$ est donc la surface de l'hypersphère dans un espace de dimension $n$, c'est-à-dire

$$\int_{z_1^2+z_2^2+...+z_n^2=1} dw$$

Par ailleurs, Kronecker précise dans son article de 1869 ou dans sa réponse à Picard dans les *Comptes rendus* [68] que « l'élément » $dw$ est défini par l'égalité :

$$|F_{mk}|dw = \gamma dz_1 ... dz_{k-1} dz_{k+1} ... dz_n$$

Cette élément $dw$ est défini par Poincaré, lorsqu'il rappelle la définition de la caractéristique de Kronecker, comme étant « un élément quelconque » de la « multiplicité », c'est-à-dire de la variété de dimension $n-1$ définie par $F_{00}(x_1, ..., x_n) = 0$. En terme moderne, on peut

67. Nous prendrons ici les notations de Kronecker dans [Kronecker, 1891b] qui nous serviront par la suite.

68. Voir [Kronecker, 1869c] et [Kronecker, 1891b].

donc écrire cette intégrale sous la forme [69] :

$$-\frac{1}{vol(S^{n-1})}\int_{F_{00}=0}\sum_{j=1}^{n}(-1)^{j-1}\frac{F_{j0}dF_{10}\wedge...\wedge dF_{j-1,0}\wedge dF_{j+1,0}\wedge...\wedge dF_{n0}}{(F_{10}^2+...+F_{n0}^2)^{\frac{n}{2}}}$$

Ce qui correspond, à un coefficient près, à la forme de l'intégrale que donne Picard dans [Picard, 1892] :

$$\int\int...\int_{F_{00}=0}\sum_{i=1}^{n}\frac{A_i dx_1...dx_{i-1}dx_{i+1}...dx_n}{\gamma^n}$$

où

$$A_i=(-1)^{n(i-1)}\begin{vmatrix} F_{10} & F_{11} & F_{12} & ... & F_{1,i-1} & F_{1,i+1} & ... & F_{1n}\\ ... & ... & ... & ... & ... & ... & ... & ...\\ ... & ... & ... & ... & ... & ... & ... & ...\\ ... & ... & ... & ... & ... & ... & ... & ...\\ ... & ... & ... & ... & ... & ... & ... & ...\\ F_{n0} & F_{n1} & F_{n2} & ... & F_{n,i-1} & F_{n,i+1} & ... & F_{nn}\end{vmatrix}$$

et correspond donc, au signe près, aux $\Delta_i$ précédents. Il n'est pas difficile, sous cette forme, de retrouver le cas $n=2$.

Ainsi, la caractéristique du système est exprimée sous la forme d'une intégrale multiple d'ordre $n-1$. Nous avons choisi d'intégrer sur $F_{00}=0$, mais cette intégrale, appelée par Hadamard *l'indice de Kronecker* [70], serait la même si l'on avait permuté le rôle de $F_{00}$ avec n'importe quel autre $F_{m0}$ $(m=1,...,n)$. C'est pourquoi le nombre entier auquel est égal cette intégrale *caractérise*, en un certain sens, le système des $n+1$ fonctions $F_{00}, F_{10}, ..., F_{n0}$.

C'est sous la forme de cette formule que la théorie des caractéristiques obtient donc une certaine postérité, et c'est à propos de celle-ci que s'engage, l'année de la mort de Kronecker, une controverse avec le mathématicien français Émile Picard.

## 14.5 La controverse entre Picard et Kronecker

Émile Picard, qui accède très jeune à une forte reconnaissance des instances scientifiques françaises [71] a produit dans les premières années de sa très longue carrière de nombreux travaux en analyse complexe. Proche de Charles Hermite [72], il connaît bien l'œuvre de Kronecker, et il expose en 1889, dans un article intitulé *Sur le nombre des racines communes*

69. Voir par exemple [Mawhin, 2000] p. 129 et [Dinca et Mawhin, 2009] p. 7.
70. Voir [Tannery et Hadamard, 1910] p. 437.
71. Normalien à 18 ans, agrégé à 21 ans, titulaire de la chaire de Calcul Différentiel de la Sorbonne à 30 ans, âge minimal requis pour cette fonction. Sa première nomination pour devenir membre de l'Académie doit même être reportée à 1889 en raison de son trop jeune âge.
72. Il est son gendre depuis 1881.

*à plusieurs équations simultanées* des résultats qu'il décrit comme directement issus des travaux de ce dernier :

> La recherche du nombre de racines communes à plusieurs équations simultanées a fait autrefois l'objet des travaux de Cauchy, et de Liouville et de Sturm. Plus récemment M. Kronecker a rencontré aussi cette question dans ses profondes recherches sur les caractéristiques des systèmes de fonctions de plusieurs variables (*Monatsberichte, 1869*). Je viens de traiter ce problème dans mon Cours à la Faculté des Sciences ; un résumé de la Leçon que j'ai faite à ce sujet pourra peut-être intéresser les lecteurs de ce recueil. [Picard, 1889] p. 5

Cet article est publié dans *les Nouvelles Annales de Mathématiques*, dont le public visé est essentiellement les élèves qui préparent les concours et les enseignants de mathématiques [73]. Il y a dans cette revue très peu d'articles de recherche, et celui de Picard n'en est pas un : il expose simplement partiellement, et pour résoudre une question particulière, la théorie des caractéristiques de Kronecker telle qu'il la présente dans son cours à la Sorbonne. On la retrouve d'ailleurs dans deux chapitres de son cours publié deux ans plus tard [74] intitulés « Exemples d'intégrales effectuées le long d'un contour fermé. Racines communes à deux équations » et « Racines communes à trois équations ». En 1891, l'année où est publié le *Traité d'analyse* de Picard, une controverse survient entre ce dernier et Kronecker au sujet de la théorie des caractéristiques. Commençons, avant de rentrer dans les détails, par donner une « chronologie » de cette controverse.

Nous donnons ci-dessous la liste des articles qui font référence à cette controverse sous une forme chronologique :

1. *Sur le nombre de racines communes à plusieurs équations simultanées* de Picard dans les *Comptes rendus* du 7 septembre 1891 ([Picard, 1891c])
2. *Sur la recherche du nombre de racines communes à plusieurs équations simultanées* de Picard dans les *Comptes rendus* du 16 novembre 1891 ([Picard, 1891b])
3. *Sur le nombre des racines communes à plusieurs équations simultanées* de Kronecker dans les *Comptes rendus* du 28 décembre 1891 ([Kronecker, 1891b])
4. *Du nombre de racines communes à plusieurs équations simultanées* de Picard dans les *Comptes rendus* du 28 décembre 1891 ([Picard, 1891a])
5. *Sur le nombre de racines communes à plusieurs équations simultanées.* de Picard en 1892 dans le *journal de Mathématiques pures et appliquées* ([Picard, 1892]) item *Sur une extension du théorème de Sturm* de Lars Edvard Phragmén dans les *Comptes rendus* en 1892 ([Phragmén, 1892])
6. Tome 2 du *Traité d'analyse* de Picard publié en 1893. ([Picard, 1891d])

73. On pourra lire à ce propos les articles de Laurent Rollet et Philippe Nabonnand, et en particulier [Nabonnand et Rollet, 2009].

74. Voir [Picard, 1891d] p. 80 pour deux équations à deux inconnues et p. 126 pour trois équations à trois inconnues.

7. *Sur la détermination du nombre de racines communes à un système d'équations simultanées et sur le calcul de la somme des valeurs d'une fonction en ces points.* de Walter Dyck dans les *Comptes rendus* en 1894 ([Dyck, 1894])
8. *Sur les racines communes à plusieurs équations* de Walter Dyck dans les *Comptes rendus* en 1895 ([Dyck, 1895b])
9. Deuxième édition du Tome 1 du *Traité d'analyse* de Picard en 1901([Picard, 1891d])
10. *Teorema di Sturm e sue estensioni* de Gaspare Mignosi en 1925 ([Mignosi, 1925])

Comme nous le voyons, l'échange entre Kronecker et Picard est très court, celui-ci commençant seulement quelques mois avant le décès de Kronecker, et quelques semaines après la mort de la femme de ce dernier. Ainsi, comme on ne peut parler de controverse qu'à partir du moment où les deux acteurs se sont exprimés, on peut même dire que celle-ci ne commence, au sens littéral, qu'à la veille de la mort de Kronecker : sa lettre qui répond aux commentaires de Picard sur sa théorie des caractéristiques est publiée dans les *Comptes rendus* le 28 décembre 1891, et Kronecker s'éteint le 29. Pourtant, celle-ci se poursuit quelques années, entretenue par Picard, et avec Walter Dyck qui dans une certaine mesure prêtera sa voix à Kronecker.

### 14.5.1 Les premières publications de Picard sur la théorie des caractéristiques

#### 14.5.1.1 L'article de 1889 et le traité de 1891

Picard présente donc la théorie de Kronecker, tout au moins la partie que l'on a qualifiée d'analytique, dans le premier tome de son cours sur d'analyse, et plus précisément dans ses chapitres [75] sur l'intégrale curviligne et les intégrales doubles. Dans son article de 1889 et dans la première édition de son traité, il n'est pas encore question de controverse : dans le cadre de son exposition de la théorie de l'intégration, Picard utilise les résultats que Kronecker expose dans son article de 1869 pour, en quelque sorte, illustrer son chapitre. Il y présentera, de la même façon que nous l'avons nous-même présentée plus haut [76], cette théorie dans le cadre de la théorie de l'intégration, pour le cas des fonctions de deux et trois variables.

Dans l'article de 1889, Picard cherche simplement à donner un « résumé de la Leçon » [77] qu'il a donnée à la Faculté des sciences, et il se place dans la lignée de Cauchy, Liouville, Sturm et Kronecker et dans le cadre de la « recherche du nombre de racines communes à plusieurs équations simultanées ». Il attribue d'ailleurs à Kronecker la paternité de ce travail :

> Plus récemment M. Kronecker a rencontré aussi cette question dans ses profondes recherches sur les caractéristiques des systèmes de fonctions de plusieurs variables

75. Chapitres III et IV de [Picard, 1891d].
76. Voir p. 411.
77. Voir [Picard, 1889] p. 5.

(*Monatsberichte, 1869*). [Picard, 1889] p. 5

Ainsi, dès sa première présentation, il ne décrira la théorie de Kronecker que par le biais des intégrales, et par leur interprétation dans le cas de la recherche des solutions d'un système d'équations. Il développe le cas de deux et trois variables, mais précise que « l'extension de la méthode à un nombre quelconque d'équations ne présenterait aucune difficulté. » [78], et il utilisera d'ailleurs le cas des fonctions de quatre variables.

Cette présentation est assez proche de celle qu'il donnera en 1891 dans la première édition de son *Traité*. Dans celui-ci, la présentation se fait en deux temps. Tout d'abord, comme application de l'intégrale curviligne, il obtient que la différence du nombre de solutions du système $\varphi = \psi = 0$ à l'intérieur d'un contour fermé $C = 0$ pour lesquelles le déterminant fonctionnel est positif avec celui des solutions pour lesquelles il est négatif est égale à :

$$\frac{1}{2\pi}\int_C \frac{\varphi d\psi - \psi d\varphi}{\varphi^2 + \psi^2}$$

et il poursuit en en déduisant, comme nous l'avons fait, le théorème de Cauchy en considérant $\varphi$ et $\psi$ comme les parties réelles et imaginaires d'un polynôme de $\mathbb{C}[X]$.

Dans le chapitre sur les intégrales doubles, il poursuivra son étude en fournissant [79] une formule équivalente dans le cas de trois équations à trois variables.

#### 14.5.1.2 Les réserves de Picard

Dans les deux exposés qui précèdent la controverse, Picard émet deux réserves. La première a trait à l'importance du rôle que joue le signe du déterminant fonctionnel et à ses conséquences sur la formule obtenue [80] :

> On voit l'importance du signe du déterminant fonctionnel : on ne connaît pas, du moins en général, de formule donnant pour les deux équations précédentes le nombre des racines contenues dans un contour ; mais si, quels que soient $x$ et $y$, le déterminant a un signe invariable, l'énoncé précédent permettra de trouver le nombre des racines. [Picard, 1891d] p. 87

Ces remarques à propos de la dépendance de l'intégrale par rapport au signe du déterminant fonctionnel se répètent à plusieurs reprises dans son cours. Picard regrette donc l'absence d'une *formule* qui donnerait le nombre exact de solutions : même s'il ne l'exprime pas ainsi, le travail de Kronecker lui semble donc incomplet. Il poursuit en montrant les cas particuliers où ce problème n'apparaît pas, c'est-à-dire lorsque le déterminant fonctionnel garde un signe constant.

78. *Ibid.*
79. Voir [Picard, 1891d] p. 123.
80. Il ajoute d'ailleurs, p. 127 du même ouvrage, que l'on « voit le rôle essentiel joué dans cette question par le signe du déterminant fonctionnel. »

Le signe du déterminant fonctionnel est en effet constant dans le cas particulier où $\varphi(x, y)$ et $\psi(x, y)$ seront les parties réelles et imaginaires d'un polynôme de la variable complexe. C'est en fait le cas pour toute fonction holomorphe : les équations de Cauchy-Riemann nous permettent en effet d'écrire ce déterminant sous la forme d'une somme de deux carrés [81].

La seconde réserve, Picard l'exprime ainsi dans la conclusion de son chapitre traitant du nombre de racines communes à trois équations :

> Remarquons aussi que l'intégrale $I$ prend une forme toute différente si, au lieu des trois équations primitives, on forme avec elles une combinaison linéaire. Par suite, l'intégrale qui, au signe près, ne devrait pas changer, ne se présente pas sous une forme invariante. Il semble donc que le résultat précédent, si intéressant qu'il soit, appelle encore de nouvelles recherches. [Picard, 1891d] p. 127

Évidemment, si l'on prend un système équivalent à celui étudié, par exemple en remplaçant une équation par une combinaisons linéaire d'elle-même et des autres, le nombre de solutions du système doit rester le même. Et une intégrale qui donnerait ce nombre devrait donc rester la même. Cette réserve, Kronecker aussi en fait part dans son article de 1869 [82] :

> C'est ce qui ressort déjà du fait que toutes les intégrales considérées ici prennent une forme complètement différente si l'on prend à la place du système d'équations $F_k = 0$ ou $f_k = 0$ un autre système équivalent, c'est-à-dire un système qui définit exactement les mêmes points $\zeta$ ou $\eta$ (y compris leurs caractères), tandis que bien sûr l'intégrale ne doit pas changer de valeur. Mais d'un autre côté on doit remarquer les facilités que l'on peut se permettre lors de l'évaluation directe des valeurs de ces intégrales. On peut en effet pour une fonction-frontière donnée $F_0$ en substituer une autre dont la représentation sera la plus commode possible et analytiquement appropriée, étant donné qu'il importe bien seulement que celles-ci concordent au niveau des points intérieurs $\eta$. [Kronecker, 1869c] p. 202

Mais la facilité calculatoire d'un choix pertinent du domaine cherché n'est pas primordiale pour Kronecker. Nous allons maintenant voir comment en 1891 Picard résout, de son point de vue, ces deux difficultés, et pour quelles raisons Kronecker affirme les avoir déjà résolues.

### 14.5.2 Les *Comptes rendus* du second semestre de 1891 : l'objet de la controverse

Souvenons-nous que tout le travail de Kronecker sur la théorie des caractéristiques est issu de la recherche du nombre de solutions de certaines équations (ou système d'équations)

81. Nous avons déjà vu les détails de cette application p. 418.

82. Dies erhellt schon aus dem Umstände, dass sämmtliche hier betrachteten Integrale eine ganz andere Form annehmen, wenn man für die Gleichungs Systeme : $F_k = 0$ oder $f_k = 0$ andere aber äquivalente nimmt, d. h. solche, welche genau dieselben Punkte $\zeta$ oder $\eta$ (sammt deren Charakter) bestimmen, während natürlich die Integrale ihrem Werthe nach dabei unverändert bleiben müssen. Andrerseits sollen aber auch die Erleichterungen hervorgehoben werden, welche man sich bei der directen Ermittelung jener Integralwerthe erlauben darf. Man kann nämlich für gegebene Begrenzungs-Functionen $F_0$ andere einer möglichst bequemen und geeigneten analytischen Darstellung zugängliche Begrenzungen substituiren, da es ja nur darauf ankommt, dass diese mit jenen in Bezug auf die umschlossenen Punkte ($\eta$) übereinstimmen.

algébriques. Ainsi, l'intégrale

$$\int\int ... \int_{F_{00}=0} \sum_{i=1}^{n} \frac{A_i dx_1 ... dx_{i-1} dx_{i+1} ... dx_n}{\gamma^n}$$

que l'on a vue dans la section précédente n'est rien d'autre que la différence du nombre de solutions du système

$$F_{10} = ... = F_{n0} = 0$$

« contenues » dans le domaine $F_{00} < 0$ pour lesquelles le déterminant fonctionnel est positif et celui pour lesquelles le déterminant fonctionnel est négatif. Et c'est sur ce dernier point – la condition portant sur le signe du déterminant fonctionnel – que porte la controverse. Le cas particulier où le déterminant fonctionnel s'annule est évacué par Picard, mais pris en compte par Kronecker. Ce que dit Picard, et ce sur quoi Kronecker le contredira, à propos du problème précédent, c'est la chose suivante :

> Une formule a été donnée à ce sujet par M. Kronecker, dans ses célèbres recherches sur les caractéristiques des systèmes de fonctions ; malheureusement l'intégrale de M. Kronecker, intégrale multiple d'ordre $n-1$ étendue à la *surface* du domaine $D$, ne représente pas le nombre cherché des racines. Le déterminant fonctionnel du système des fonctions $f$ joue, dans cette théorie, un rôle capital, et on obtient seulement la différence entre le nombre des racines pour lesquelles ce déterminant fonctionnel est positif et celui des racines pour lesquelles il est négatif.
> Je voudrais montrer ici, en quelques lignes, qu'il est possible de représenter par une intégrale convenable le nombre *exact* des racines. [Picard, 1891c] p. 356

C'est donc dans ce texte, en septembre 1891, que Picard résout le problème, ou tout au moins le manquement, qu'il avait relevé dans ses écrits précédents : représenter par une unique intégrale indépendante du signe du déterminant fonctionnel le nombre de solutions d'un système de $n$ équations à $n$ inconnues à l'intérieur d'un domaine borné.

Avant de poursuivre, précisons que Picard, comme Kronecker, compte les solutions *réelles* d'un système : ce sont dans les deux cas les intersections de courbes réelles que l'on étudie.

On doit noter que la « célébrité » de la théorie des caractéristiques dont parle Picard est très relative : il est certainement, avec Molk, l'un des seuls mathématiciens français à l'avoir exposée. Plus tard, lorsque Hadamard la qualifiera [83] de « classique », c'est aux travaux de Picard qu'il attribuera le mérite de sa diffusion. Picard est certainement familier avec cette théorie grâce à son beau-père Hermite [84], et nous verrons d'ailleurs que cette connaissance reste très partielle.

83. Voir [Tannery et Hadamard, 1910] p. 437.
84. Nous verrons que ce sera aussi par Hermite que Poincaré prendra connaissance des articles de Kronecker sur la théorie des caractéristiques.

#### 14.5.2.1 L'article du 7 septembre 1891

Dans cet article très court, Picard montre comment on peut astucieusement donner une intégrale qui fournit le nombre exact de solutions d'un système de deux équations à deux inconnues

$$\begin{cases} f(x,y)=0 \\ \varphi(x,y)=0 \end{cases}$$

à l'intérieur d'un contour $C$. Pour cela, il transforme en fait son système de façon à obtenir un déterminant positif. Il aura besoin de deux variables supplémentaires, $z \in [-\epsilon;\epsilon]$ et $t \in [-\eta;\eta]$, où $\epsilon$ et $\eta$ sont des réels positifs. Picard propose alors le système suivant :

$$\begin{cases} f(x,y)=0 \\ \varphi(x,y)=0 \\ z\dfrac{\partial f}{\partial x}+t\dfrac{\partial f}{\partial y}=0 \\ z\dfrac{\partial \varphi}{\partial x}+t\dfrac{\partial \varphi}{\partial y}=0 \end{cases}$$

qui a de façon évidente le même nombre de solutions que le système de départ, mais dont le déterminant fonctionnel est

$$\begin{vmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & 0 & 0 \\ \dfrac{\partial \varphi}{\partial x} & \dfrac{\partial \varphi}{\partial y} & 0 & 0 \\ A_1 & A_2 & \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ B_1 & B_2 & \dfrac{\partial \varphi}{\partial x} & \dfrac{\partial \varphi}{\partial y} \end{vmatrix} = \left(\frac{\partial f}{\partial x}\frac{\partial \varphi}{\partial y}-\frac{\partial f}{\partial y}\frac{\partial \varphi}{\partial x}\right)^2$$

où $A_1, B_1, A_2$ et $B_2$ n'interviennent pas dans le calcul [85]. On obtient ainsi une intégrale triple qui donnera en effet le nombre exact de solutions du système. Alors que Kronecker proposait une intégrale d'ordre $n-1$ où $n$ est le nombre de variables, Picard obtient une intégrale d'ordre $n+1$. On peut certes, dans le cas décrit, passer à une intégrale double. Picard le fait d'ailleurs remarquer. Mais dans le cas général, la manipulation n'est pas possible.

#### 14.5.2.2 L'article du 16 novembre 1891

La note dont on vient de parler est rapidement complétée par une autre dans les *Comptes rendus* deux mois plus tard. Le but de Picard est de « simplifier » [86] la méthode sans en « modifier le principe » [87]. Il remplace ici le système

85. Déterminant par blocs.
86. Voir [Picard, 1891b] p. 670.
87. *Ibid.*

$$\begin{cases} f_1(x_1, x_2, ..., x_n) = 0 \\ f_2(x_1, x_2, ..., x_n) = 0 \\ \qquad ... \\ f_n(x_1, x_2, ..., x_n) = 0 \end{cases}$$

par

$$\begin{cases} f_1(x_1, x_2, ..., x_n) = 0 \\ f_2(x_1, x_2, ..., x_n) = 0 \\ \qquad ... \\ f_n(x_1, x_2, ..., x_n) = 0 \\ \qquad zD = 0 \end{cases}$$

avec comme précédemment $z \in [-\epsilon; \epsilon]$ et où $D$ est le déterminant fonctionnel du premier système. Le déterminant fonctionnel du second système est alors égal à $D^2$ et son signe est donc constamment positif. Une fois de plus, la difficulté liée au signe de ce déterminant est évacuée. Il applique ensuite de nouveau ce résultat aux systèmes de deux équations à deux inconnues, et il obtient l'intégrale curviligne que l'on a déjà étudiée [88].

Comme Picard le fait remarquer lorsqu'il donne une expression de l'intégrale en question, celle-ci semble dépendre de la quantité $\epsilon$. Ce n'est évidemment pas le cas, et l'étude des cas limites – $\epsilon$ tend vers 0 ou vers $+\infty$ – permet même dans certains cas de simplifier son expression.

Dans ces deux notes, Picard présente son travail comme la résolution de la « difficulté principale » [89] dans la recherche du nombre des racines communes à plusieurs équations simultanées. Si, du point de vue de Picard, la dépendance du nombre de solutions d'un système au signe d'un déterminant fonctionnel est problématique, la résolution qu'il en donne n'est qu'une application astucieuse du travail colossal que Kronecker a fourni dans son article de 1869. On perçoit donc, dès ces deux premières notes, l'envie de Picard, non pas de s'attribuer le travail de Kronecker – il ponctue tous ses textes de commentaires élogieux sur le travail et le génie du mathématicien allemand – mais de mettre en avant sa participation dans la résolution de la question. Finalement, il affirme qu'il est le premier à avoir résolu cette question, comme Sturm un demi-siècle plus tôt a été le premier à le résoudre dans le cas $n = 1$. Il aurait d'ailleurs souhaité répondre à « la question très intéressante de savoir si, pour obtenir cette limite, on peut trouver un mode de calcul, qui serait l'analogue du théorème de Sturm » [90]. Nous allons voir que cette volonté de Picard d'associer son nom à une formule donnant le nombre de solutions d'un système d'équations perdure bien après la mort de Kronecker.

88. Voir p. 416.
89. Voir [Picard, 1891b] p. 669.
90. Voir [Picard, 1891c] p. 357.

#### 14.5.2.3 La réponse de Kronecker

**La lettre de Kronecker adressée à Picard** Kronecker, pourtant récemment endeuillé par la mort de sa femme, répond très rapidement à l'article de Picard. Ainsi, dans les *Comptes rendus* datés du 28 décembre 1891, la veille de la mort de Kronecker, est publiée en intégralité sa réponse à la première note de septembre [91]. Cette réponse constituera finalement la seule intervention directe de Kronecker.

Dans cette réponse, il commence par rappeler l'ensemble des notations qu'il a utilisées dans ses articles de 1869 et 1878 sur la théorie des caractéristiques, puis donne la définition de la caractéristique d'un système de $n+1$ fonctions à $n$ variables sous la forme d'une somme de signes de déterminant et d'une intégrale (il s'agit de l'égalité qu'il nommera $(I)$). Il fait ensuite « remarquer que la somme, dans l'égalité $(I)$, prend la même valeur pour tous les $n+1$ nombres $m = 0, 1, ..., n$ » [92], c'est-à-dire que la caractéristique du système de $n+1$ équations est indépendante du choix de celle qui servira de domaine d'intégration.

Ce point de vue est complètement ignoré par Picard, qui lui part d'un système de $n$ équations auquel il adjoint un domaine. Kronecker construit un invariant pour un système de $n+1$ équations à $n$ inconnues qu'il applique ensuite, entre autres, à la détermination du nombre de racines d'un système de $n$ équations, alors que Picard a pour but principal la recherche du nombre de solutions.

Kronecker poursuit alors en exprimant le nombre $N$ de solutions du système

$$\begin{cases} F_{10}(x_1, x_2, ..., x_n) = 0 \\ F_{20}(x_1, x_2, ..., x_n) = 0 \\ \qquad ... \\ F_{n0}(x_1, x_2, ..., x_n) = 0 \end{cases}$$

dans le domaine $G < 0$ par la différence de deux caractéristiques :

$$\chi(G\Delta_0, F_{10}, ...., F_{n0}) - \chi(\Delta_0, F_{10}, ...., F_{n0})$$

où $\Delta_0$ est le déterminant fonctionnel des $n$ fonctions $F_{10}, ...., F_{n0}$.

En remplaçant chacune des deux caractéristiques par l'intégrale correspondante, Kronecker obtient bien une expression de $N$ grâce à la différence de deux intégrales d'ordre $n-1$. Il traduit ensuite, en termes de *caractéristique*, la formule donnée par Picard, et montre que celle ci n'est rien d'autre que la caractéristique du système de $n+2$ fonctions à $n+1$ variables :

$$N = \chi\left((a-z)(b+z)G, z\Delta_0, F_{10}, ...., F_{n0}\right)$$

91. Kronecker ne semble pas avoir lu la seconde note datée de novembre.
92. Voir [Kronecker, 1891b] p. 1007.

où l'intervalle $[a, b]$ correspond à l'intervalle $[-\epsilon, \epsilon]$ de Picard. Il termine en examinant les cas limites, comme le fait Picard, où $a$ et $b$ sont nuls ou infinis [93]. Examinons comment Kronecker conclut sa réponse :

> Permettez-moi d'ajouter encore une remarque. Dans mes recherches sur les systèmes de fonctions de plusieurs variables, j'ai été forcé de distinguer les systèmes de valeurs réelles communes à $n$ équations selon leurs *caractéristiques*. Cette distinction entre, il est vrai, dans toutes les formules que j'ai développées dans mes divers Mémoires publiés sur ce sujet, principalement dans la Note sur le théorème de Sturm que j'ai présentée moi-même à l'Académie le 6 mai 1869 ( voir t. LXVIII). Mais cette distinction, bien loin d'être regrettable, me semble plutôt révéler la vraie nature des choses, cachée jusque-là, parce qu'on s'est borné à considérer cette espèce particulière de fonctions provenant des fonctions de variables complexes, pour laquelle la caractéristique conserve toujours la même valeur. [Kronecker, 1891b] p. 1012

En effet, Picard, comme Cauchy avant lui, montre et répète que le signe du déterminant fonctionnel est primordial dans l'examen de ces systèmes : alors, dit Kronecker, pourquoi le passer sous silence ? Nous avons vu que dans le cas des fonctions holomorphes, le déterminant fonctionnel est positif. De même, comme Kronecker le fait dans le chapitre IX de son article de 1869 ([Kronecker, 1869c]), si l'on prend le système formé des $2m$ parties réelles et imaginaires de polynômes de la variable complexe, on obtient aussi un déterminant fonctionnel dont le signe est constant, et on a alors une formule dans laquelle ce signe n'intervient pas. Mais pour Kronecker, la solution, astucieuse, de Picard, ne fait que « cacher la véritable nature » de ce système, ce qui le *caractérise*. Et ce qui dévoile cette nature pour Kronecker, c'est la recherche de *généralité*. On ne se borne pas à des « espèces particulières de fonctions », mais au contraire, « au lieu de la restriction (...) au cas particulier où $\Delta_0$ ne s'annule pas dans le domaine $G < 0$ » [94], le résultat qu'il donne est « général » [95]. Kronecker ne développe ici rien de nouveau : la formule donnant le nombre de racines en fonction de la différence de deux caractéristiques est bien présente [96] dans son article de 1869. Kronecker l'utilisera d'ailleurs pour obtenir le nombre de racines réelles d'un polynôme d'une variable, et revenir ainsi sur son travail sur le théorème de Sturm [97]. Par contre, il est vrai que Kronecker n'a pas développé dans ses articles, comme il le fait ici, l'expression de l'intégrale qui en découle.

C'est donc encore sur le front de la généralité, comme nous l'avons déjà vu dans les parties précédentes, que Kronecker se bat ici. Non seulement il ne reconnait à Picard aucune nouveauté, mais il lui reproche d'avoir caché, inutilement, la « vraie nature » de son travail et d'avoir ainsi perdu une certaine forme de généralité.

L'article de Kronecker se termine par le paragraphe suivant :

93. En fait, seul le cas infini est traité.
94. Voir [Kronecker, 1891b] p. 1012.
95. *Ibid.*
96. Voir [Kronecker, 1869c] p. 183.
97. *Ibid.*, chapitre IV.

> Rien, peut-être, ne montre plus clairement la portée de la distinction que j'ai introduite, que la comparaison de la formule (VI) pour le cas $n = 2$ avec celle que vous avez citée dans votre Article du 16 novembre. Au lieu de la restriction, nécessaire pour celle-ci, au cas particulier où $\Delta_0$ ne s'annule pas dans le domaine $G < 0$, ma formule (VI) donne un résultat général, mais en ajoutant une intégrale étendue sur le domaine $(\Delta_0 = 0, G < 0)$, qui n'existe pas dans la cas particulier. [Kronecker, 1891b] p. 1012

**Les commentaires de Picard** Picard, membre de l'Académie, ajoute un commentaire immédiatement à la suite de la lettre de Kronecker. Il utilisera systématiquement cette méthode lorsqu'une publication concernant son travail sur la théorie des caractéristiques paraîtra dans les *Comptes rendus*.

Picard commence par affirmer que, finalement, on trouve bien une « solution du problème » [98], c'est-à-dire une expression donnant le nombre de solutions d'un système d'équations. Mais, dit-il, il « est impossible de considérer cette solution comme satisfaisante ». L'argument de Picard est le suivant :

> On doit, en effet, dans ce problème, chercher à exprimer le nombre des racines par une formule dont l'application numérique ne nécessite aucune discussion spéciale relative au système particulier des équations $f = 0$, et les intégrations à effectuer doivent dépendre uniquement, au point de vue des limites, du domaine $\Delta$ : c'est ce qui n'a pas lieu dans l'analyse de M. Kronecker qui est obligé de partager le domaine $\Delta$ en plusieurs autres dépendant des équations particulières que l'on a à étudier. [Picard, 1891a] p. 1013

Le reproche de Picard est donc le suivant : la formule proposée par Kronecker n'est pas adéquate pour un calcul *pratique* du nombre en question, car il nécessite en fait un travail préalable, qui consiste à séparer les espaces où le déterminant fonctionnel est positif de ceux où il est négatif. Ce travail, proprement topologique, Picard n'en veut pas, et il préfère ajouter une condition (déterminant fonctionnel différent de zéro) et un degré de plus dans l'ordre des intégrales qui interviennent dans la formule. C'est bien pourtant sur ce point que Kronecker insiste, et ce point constitue le cœur même de son travail : sa théorie des caractéristiques repose, comme nous l'avons vu, sur des considérations que l'on peut qualifier de *topologiques*. Et lorsqu'il revient sur cette théorie en 1878, c'est en considérant la façon dont la caractéristique change lorsque les coefficients des fonctions du système évoluent. Ces derniers dépendent de variables qui prennent leurs valeurs dans une variété, et Kronecker observe les lieux de cette variété où la caractéristique est constante.

Picard insiste sur le fait que l'intégrale qu'il propose donne immédiatement (c'est-à-dire avant même d'étudier les cas limites de $\epsilon$) un moyen sûr pour obtenir le nombre de solutions. En effet, la valeur de cette intégrale étant un nombre entier, une approximation du résultat au dixième suffit pour en déduire la valeur exacte. Il souligne l'importance de cela par le recours aux italiques dans la phrase suivante :

98. Voir [Picard, 1891a] p. 1013.

> Elle [l'intégrale] renferme une indéterminée $\epsilon$, et c'est en faisant $\epsilon$ distinct de zéro et de l'infini, que *l'on n'a aucune difficulté dans l'application de la formule*, par exemple dans son application numérique. [Picard, 1891a] p. 1013

Cette note précise encore un peu plus le but visé par Picard : il s'agit d'obtenir un moyen effectif et général pour trouver le nombre de solutions d'un système d'équations donné. Mais pour lui, c'est la compacité de la formule, au sens où son application ne demande aucune discussion, qui fait office de généralité. Cependant, il ne faudrait pas croire que ces considérations soient étrangères à Kronecker, et il affirme d'ailleurs lui aussi dans son dernier cours sur le théorie des équations algébriques [99] :

> Comme la caractéristique est un nombre entier, alors il est seulement nécessaire de calculer l'intégrale dans la dernière formule exactement jusqu'à la première décimale pour obtenir la valeur de la caractéristique. [Kronecker, 1891a] p. 246

### 14.5.3 Une résolution posthume

#### 14.5.3.1 L'article de 1892 du *Journal de Liouville*

Malgré la mort de Kronecker, Picard revient de nombreuses fois sur le sujet. En 1892, il publie ainsi, dans le *Journal de Mathématiques pures et appliquées*, dont il est membre du Comité de rédaction, un article qui combine l'exposition des *Annales* de 1889 et le contenu de ses notes aux *Comptes rendus*.

Dans cet article, il commence par exposer ce qu'il nomme la « formule de Kronecker » [100], puis il en donne une démonstration dans le cas où $n = 3$. il ajoute, comme dans ses *Notes*, une équation au système qu'il examine, et regarde en détail les cas des systèmes de une, deux et trois équations. Il termine en étudiant les cas particuliers de recherche du nombre de solutions d'un système à l'intérieur de certains « contours » particuliers. Il s'attarde aussi, pour un système de deux équations à deux inconnues, au cas d'un contour carré. Cet exemple lui permet de faire le lien avec le travail de Hermite sur le théorème de Sturm, tel que nous l'avons décrit dans la partie précédente. De plus, il exprime la formule qu'il obtient en utilisant l'indice de Cauchy, et c'est finalement un moyen pour lui de faire, comme Kronecker avant lui, un lien avec la théorie de Cauchy. Mais si, pour Kronecker, la théorie des caractéristiques est une généralisation des théories de Sturm et de Cauchy, Picard se contente de faire apparaître des liens entre son travail et ces théories.

Dans le cas de trois équations à trois inconnues, Picard décrit le cas où le contour est un parallélépipède parallèle aux axes, puis il clôt son article par un long travail qui lui permet une fois de plus de « trouver le nombre des racines en effectuant simplement des recherches d'indices » [101].

99. Da die Charakteristik eine ganze Zahl ist, so ist es nur nötig, das Integral in der letzten Formel bis auf die erste Dezimale genau zu berechnen, um den Wert der Charakteristik zu erhalten.
100. Voir [Picard, 1892] p. 6.
101. Voir [Picard, 1892] p. 24.

Terminons en examinant une note que Picard ajoute à son article au moment où il rappelle que « le problème de la recherche du nombre *exact* de racines ne semble donc pas résolu par la formule de M. Kronecker » [102] :

> À la suite des deux notes que j'ai publiées sur cette question (*Comptes rendus*, 7 septembre et 16 novembre 1891), M. Kronecker m'a fait l'honneur de m'écrire une lettre où il me représente qu'on trouve dans son Mémoire de 1878 la solution de la question. Je ne puis partager l'opinion de l'illustre auteur ; il me semble que, dans ce problème, on doit chercher à exprimer le nombre des racines par une formule dont l'application ne nécessite aucune discussion spéciale relative au système particulier des équations $f = 0$ et que les intégrales à effectuer doivent dépendre uniquement, au point de vue des limites, du domaine $\Delta$ : c'est ce qui n'a pas lieu dans l'analyse de M. Kronecker, qui partage le domaine $\Delta$ en plusieurs autres dépendant des équations particulières que l'on a à étudier. [Picard, 1892] p. 6

Cette citation contient quelques imprécisions. Tout d'abord il est peu probable que Kronecker ait lu, lorsqu'il écrit sa lettre, la seconde note de Picard (celle du 16 novembre). Il ne fait d'ailleurs allusion qu'à la première dans sa réponse. Ensuite, Picard semble penser que Kronecker lui répond que si la réponse ne se trouve pas dans son article de 1869, c'est qu'elle est présente dans son article de 1878. Or non seulement ce n'est pas ce que Kronecker dit [103], mais de plus, l'article de 1878 traite du nombre de solutions d'un système d'équations uniquement dans le cas particulier de $2m$ équations à $2m$ inconnues qui représentent en fait les parties réelles et imaginaires de $m$ fonctions rationnelles. On sait dans ce cas que le déterminant fonctionnel est positif, et donc que le problème soulevé par Picard n'a pas lieu d'être.

#### 14.5.3.2 Les volumes du *Traité d'analyse*

Le *Traité d'analyse* de Picard se poursuit avec la publication en 1893 et 1896 de deux autres tomes. Il ne sera pas question ici du dernier tome, presque entièrement consacré à l'étude des équations différentielles. le second volume, en revanche, contient comme le premier une section sur le nombre de racines d'un système d'équations. De plus, une nouvelle édition, revue et augmentée, paraît en 1901 et 1904. Après avoir examiné la première édition du second volume, nous verrons comment Picard modifie dans la nouvelle édition les chapitres consacrés à l'intégrale de Kronecker pour prendre en compte son propre travail.

La première édition du second volume contient donc un chapitre intitulé *Nombre de racines communes à plusieurs équations*. Dans celui-ci, Picard commence par montrer comment on peut obtenir l'intégrale de Kronecker. Observons comment il introduit cette explication dans les deux éditions. En 1893, on lit :

> Dans cette étude, nous avons posé a priori une certaine intégrale ; il est intéressant de montrer comment on peut y être conduit : c'est par quoi nous allons commencer.

102. Voir [Picard, 1892] p. 6.
103. Il ne fait allusion à l'article de 1878 que pour dire que l'on y retrouve les notations de celui de 1869.

> La formule de M. Kronecker se déduit immédiatement d'une propriété élémentaire des fonctions $V$ [Picard, 1893] p. 193

Puis en 1904, il écrit :

> Dans cette étude, nous avons posé, *a priori*, une certaine intégrale, considérée par Kronecker. Il est intéressant de montrer comment l'illustre géomètre y a été conduit, c'est par quoi nous allons commencer. [Picard, 1904] p. 193

Entre 1893 et 1904, l'explication qu'il donne de la construction de l'intégrale en question correspond ainsi à celle de Kronecker. Il semble avoir tellement fait sienne la théorie qu'il s'était attaché à développer qu'il s'en est approprié aussi l'histoire. On a vu que dans son article de 1869 Kronecker donne une origine à sa théorie des caractéristiques [104], ainsi qu'à l'intégrale dont parle Picard : c'est son travail sur Sturm qui l'a dans un premier temps amené à construire la caractéristique d'un système, et c'est une discussion avec Weierstrass et le travail de Cauchy qui l'ont poussé à introduire cette intégrale. Ce à quoi Picard fait allusion n'est finalement rien d'autre que le travail de Gauss [105] qui est en effet cité par Kronecker, mais qui ne correspond qu'au cas particulier que traite Picard.

Cependant, lorsque l'on regarde plus précisément le contenu de ces deux versions, on voit que Picard a modifié en profondeur son exposition. Dans celle de 1893, il ne fait finalement que reproduire le contenu de son article de 1892, sans donner toutes les applications à certains contours. Dans la version de 1904, il s'éloigne au contraire de ses travaux antérieurs, et il calque réellement son exposé sur les chapitres V à VIII de l'article de 1869 de Kronecker. Plus encore, le discours de Picard sur le travail de Kronecker a évolué : il n'est plus question pour lui de mettre en avant son apport mais bien d'exposer la théorie de Kronecker et ses modifications. En fait, le titre même du chapitre a changé : de *Nombre de racines communes à deux équations simultanées* il passe à *Nombre de racines communes à plusieurs équations*. Ce changement vient du fait que Picard a déjà traité dans le premier volume de la seconde édition le cas de deux équations : il n'a donc plus à le faire dans le second volume. Dans cette seconde édition du premier volume, la citation que l'on a donnée plus haut p. 429 est reprise pour cette fois annoncer son résultat :

> On voit l'importance du signe du déterminant fonctionnel : nous verrons au Chapitre IV une formule donnant pour les deux équations précédentes le nombre des racines contenues dans un contour. Remarquons seulement pour le moment que, si, quels que soient $x$ et $y$ le déterminant fonctionnel a un signe invariable, l'énoncé précédent permettra de trouver le nombre des racines. [Picard, 1901] p. 99

Dans le second volume publié en 1904, Picard découpe son chapitre en deux parties, et la première s'intitule *les intégrales de Kronecker* : la reconnaissance du travail de Kronecker est complète, et surtout il met beaucoup moins en avant son apport personnel. En fait, nous verrons que ces changements sont dus en grande partie à un article que Walther von Dyck a publié en décembre 1894 dans les *Comptes rendus*.

104. Voir p. 411.
105. Voir [Gauss, 1840] p. 209.

#### 14.5.3.3 L'intervention de Van Dyck

Walther von Dyck (1856-1934) a soutenu sa thèse sur les surfaces de Riemann intitulée *Über regulär verzweigte Riemannsche Flächen und die durch sie definierten Irrationalitäten* sous la direction de Felix Klein en 1879 à Munich. On connaît la mésentente qui a pu exister entre Klein et certains mathématiciens berlinois, en particulier Kronecker. Pourtant, il conseille à son étudiant qui l'a suivi à Leipzig de passer quelques semaines à Berlin pour rencontrer le célèbre triumvirat composé de Leopold Kronecker, Ernst Eduard Kummer et Karl Weierstraß [106]. Van Dyck passe ainsi quelques semaines à Berlin au printemps 1881, durant lesquels des liens durables qui le lieront à Kronecker [107], mais aussi à l'un de ses élèves, Eugen Netto, vont se créer. Non seulement il assistera aux cours de Kronecker et lui rendra régulièrement visite chez lui, mais on sait de plus qu'il passera un certain temps à étudier les notes de cours des *Vorlesungen über der algebraïschen Gleichungen* [108].

Walther Dyck connaît donc très bien le travail de Kronecker sur la théorie des caractéristiques, et contrairement à Picard, il a pris connaissance de ses leçons sur le sujet. On a vu que celles-ci étaient souvent nécessaires pour réussir à comprendre les articles de recherche de Kronecker, et c'est ce qui permettra à Von Dyck de faire la part des choses entre ce que contiennent vraiment les articles de Kronecker et l'originalité du travail de Picard.

Lorsqu'il publie en 1894 dans les *Comptes rendus* un article intitulé *Sur la détermination du nombre de racines communes à un système d'équations simultanées et sur le calcul de la somme des valeurs d'une fonction en ces points* [109], cela fait dix ans qu'il travaille sur des questions de topologie, ou plutôt d'Analysis situs [110]. Dans cet article [111], qui est en fait un extrait d'une lettre qu'il a envoyée à Émile Picard, il reprend en effet le travail de Picard et donne son opinion sur la controverse que nous étudions ici. Après avoir affirmé que Kronecker a « donné une généralisation des deux théorèmes fondamentaux de Cauchy

106. Voir [Hashagen, 2003] p. 105.
107. *Ibid.* p. 112 et [Pont, 1974] p. 131.
108. Dans [Hashagen, 2003], Ulf Hashagen montre que Dyck avait déjà eu accès à ces notes de cours alors qu'il était étudiant à la TH de Munich.
109. Voir [Dyck, 1894].
110. Voir [Dyck, 1884], [Dyck, 1888], [Dyck, 1890]. Pour une description détaillée du travail de Dyck dans ce domaine, on pourra lire [Hashagen, 2003] pp. 265 et suivantes.
111. C'est d'ailleurs Hilbert qui en fait la revue dans le *Jachbuch* (JFM 25.0145.01 dans le Jahrbuch Database.) :

> Hr. Picard hat mit Hülfe einer Formel von Kronecker die Anzahl der in einem Gebiete liegenden gemeinsamen Wurzeln eines Gleichungssystems durch bestimmte, über das Innere und die Berandung dieses Gebietes zu erstreckende Integrale ausgedrückt. Das nämliche Resultat leitet der Verfasser auf einem einfacheren Wege ab, indem er nicht die von Hrn. Picard benutzte Formel, sondern eine andere Formel von Kronecker der Betrachtung zu Grunde legt.

qui se rapportent à la détermination du nombre des racines d'une fonction d'une variable complexe et au calcul de la somme des valeurs d'une seconde fonction en ces points » [112], il rappelle les résultats obtenus par Kronecker pour les fonctions de deux variables. À propos des articles de 1891 et 1892 de Picard, il ajoute :

> Quoique je ne puisse partager l'opinion de M. Picard, qui ne veut pas reconnaître comme une solution à la question la détermination de Kronecker (Mémoire cité,p. 6), la solution de M. Picard est fort satisfaisante sans doute, parce qu'elle ne contient que des intégrales étendues le long de $D = 0$, ce que fait ressortir avec raison M. Picard dans les remarques qu'il ajoute à la Communication de Kronecker. [Dyck, 1894] p. 1256

Ainsi, Walther Dyck tente de donner une sorte de synthèse : la solution proposée par Kronecker au problème en est bien une, mais celle donnée par Picard a l'avantage de ne pas provoquer de discussion sur le domaine d'intégration.

Il utilise une partie de l'article de Kronecker de 1869 ([Kronecker, 1869c]) dont nous n'avons pas parlé, et qui donne la somme algébrique (selon le signe du déterminant fonctionnel) des valeurs d'une fonction prises aux points qui annulent le système étudié, grâce à la somme de deux intégrales. Voyons exactement ce qu'il en est : soient trois fonctions $\varphi, \psi$ et $L$ des variables $x$ et $y$. Dyck les souhaite « continues et monodromes » [113], c'est-à-dire dérivables et n'ayant qu'une seule image [114]. Soit aussi $D$ le déterminant fonctionnel de $\varphi$ et $\psi$ : si on note $\xi_1, ...\xi_m$ les points du domaine $L < 0$ pour lesquels $\varphi = \psi = 0$, alors on suppose que $D$ ne s'annule pas en ces points. Soit de plus ne autre fonction $\mathfrak{F}$ qui est supposée « uniforme », c'est-à-dire *suffisamment régulière*. Kronecker la suppose « partout finie et continue à l'intérieur du domaine d'intégration » et « différentiable selon toutes les variables » [115]. Alors on a l'égalité suivante :

$$\frac{1}{2\pi}\left(\int_{L<0}\frac{1}{\varphi^2+\psi^2}\begin{vmatrix}0 & \mathfrak{F}_1 & \mathfrak{F}_2\\ \varphi & \varphi_1 & \varphi_2\\ \psi & \psi_1 & \psi_2\end{vmatrix}dxdy-\int_{L=0}\frac{\mathfrak{F}}{\varphi^2+\psi^2}[(\varphi\psi_1-\psi\varphi_1)dx+(\varphi\psi_2-\psi\varphi_2)dy]\right)$$
$$=\sum_{i=1}^{m} sgn(D(\xi_i))\cdot\mathfrak{F}(\xi_i)$$

Cette formule est plus générale que ce que nous avons appelé jusqu'à présent *l'intégrale de Kronecker*. En effet, en choisissant $\mathfrak{F} = 1$, on retrouve :

$$-\frac{1}{2\pi}\left(\int_{L=0}\frac{1}{\varphi^2+\psi^2}[(\varphi\psi_1-\psi\varphi_1)dx+(\varphi\psi_2-\psi\varphi_2)dy]\right)$$

112. Voir [Dyck, 1894] p. 1254.
113. Voir [Dyck, 1894] p. 1254.
114. C'est ce que Kronecker appelle univalentes.
115. [Kronecker, 1869c] p. 191 : « innerhalb des Integrations-Gebietes überall endlich und stetig » / « sämmtlichen Variabein differentiirbar ».

$$= \sum_{i=1}^{m} sgn(D(\xi_i))$$

qui est la formule dont se sert [116] Kronecker pour trouver le nombre exact de solutions du système $\varphi = \psi = 0$ dans le domaine $L < 0$.

Picard utilise cette dernière égalité pour mettre en place sa formule. Mais pour cela, comme le fait remarquer Dyck, il « introduit un élément étranger à la question, parce qu'il emploie un espace d'une dimension de plus » [117]. Cet élément étranger, c'est la dimension supplémentaire, et Picard multipliera les exemples [118] où après simplification l'intégrale étudiée sera de nouveau d'ordre $n-1$ : Walther Dyck montre alors comment, à l'aide de la seconde formule de Kronecker, on peut retrouver les résultats de Picard sans utiliser cette dimension supplémentaire. C'est cette méthode qui est reprise par Picard en 1904 dans le deuxième volume de la seconde édition de son *traité d'analyse*, et qui représente une certaine forme de synthèse du travail de Kronecker appliqué à la recherche du nombre de solutions d'un système de $n$ équations à $n$ inconnues.

### 14.5.4 Conclusion

Picard s'est donc efforcé, au moins jusqu'à la publication de la deuxième édition de son *Traité*, de revendiquer la découverte de la formule donnant le nombre de solutions d'un système d'équations à plusieurs inconnues, formule directement issue des travaux de Kronecker sur la théorie des caractéristiques, et acceptée d'ailleurs comme telle par Picard. Depuis 1885, ce dernier fait partie du Comité de rédaction du *Journal de Mathématiques pures et appliquées* que Jordan a créé en arrivant à la tête de la revue [119], et on sait qu'il est aussi, depuis 1889, membre de l'Académie des sciences. Il a donc un accès facile immédiat aux *Comptes Rendus*, comme le montrent les notes qu'il ajoute aux différents articles qui ont trait à son travail. Il utilise cet accès pour rappeler systématiquement que c'est lui, et non Kronecker, qui a donné une résolution définitive au problème. Nous l'avons vu dans le cas de la réponse de Kronecker ou dans l'article de Dyck, mais il est intéressant de remarquer que Picard utilise ce procédé la plupart du temps lorsqu'un article ayant trait à ce sujet est publié dans les *Comptes rendus*. Ainsi, dans un article de 1892 intitulé *Sur une extension du théorème de Sturm* [120], Lars Edvard Phragmén (1863-1937) reprend le travail d'Hermite sur Sturm [121] et propose une méthode alternative pour la détermination du nombre de solutions d'un système d'équations. Cette méthode n'utilise pas les intégrales qui sont au cœur des articles de Picard, mais ce dernier ajoute une note indiquant que dans son travail on « trouvera aussi la démonstration directe de ce résultat, qui est immédiate ;

116. Nous rappelons qu'il l'utilise deux fois : d'abord sur le domaine d'intégration $L.D = 0$ puis $L = 0$.
117. Voir [Dyck, 1894] p. 1256.
118. Voir par exemple [Picard, 1892] p. 14, [Picard, 1904] p. 215 pour les fonctions de deux variables.
119. On pourra lire [Brechenmacher, 2010].
120. Voir [Phragmén, 1892].
121. Voir [Hermite, 1852] et [Hermite, 1853].

l'intéressant article de M. Phragmén sur cette généralisation facile du théorème de Sturm ne renferme donc rien d'essentiellement nouveau ». En 1901 M. A. Davidoglou publie deux notes aux *Comptes rendus* « sur le nombre de racines communes à plusieurs équations » [122]. Picard, qui présente ces notes, ne réagit pas lorsque ce dernier qualifie les intégrales utilisées dans cette recherche non plus d'*intégrales de Kronecker* mais d'« intégrales de Picard » [123]. Il n'est donc pas étonnant qu'en 1910, lorsque Hadamard donne quelques applications de la théorie des caractéristiques, il indique que la notion d'indice de Kronecker est classique « surtout depuis le *Traité d'Analyse* de M. Picard » : c'est le prisme par lequel la théorie a été le plus souvent appréhendée.

En 1916, Picard publie un document intitulé *L'histoire des sciences et les prétentions de la science allemande*. Dans ce long article, qui est clairement un texte de propagande ayant pour but d'exacerber les sentiments patriotiques du lecteur, il réécrit l'histoire des sciences, et en particulier l'histoire des mathématiques. Pendant plus de cinquante pages, il s'évertue à montrer que les scientifiques allemands non seulement n'ont rien inventé, mais ont une tendance certaine à s'approprier les résultats des chercheurs des autres pays, et en particulier des Français. Si l'on s'en tient aux mathématiciens du XIX$^{e}$ siècle – et on ne peut pas soupçonner Picard de ne pas connaître leur travail – il ne citera que deux mathématiciens allemands : Weierstrass qui « conduit [ses lecteurs] en tenant une lanterne sourde » [124] et Gauss, dont « plus d'un passage constitue une énigme à déchiffrer » [125]. Le contexte de la Première Guerre mondiale, et donc le conflit entre la France et l'Allemagne impose bien sûr une lecture prudente de ce texte, qui est avant tout un texte de propagande. Cependant, le point de vue que Picard y défend doit être pris en compte, au moins partiellement, pour éclairer cette controverse : si l'attribution de telle ou telle qualité ou défaut en fonction de la nationalité est fantaisiste, ce qu'il y décrit comme étant de la bonne ou de la mauvaise mathématique doit correspondre à son propre Ethos de la science. Ainsi, aux mathématiciens allemands, même à ceux que « personne ne se donnerait le ridicule de vouloir diminuer » [126] il reproche un certain obscurantisme. Il est toutefois vrai que les articles de Kronecker, s'ils ne sont pas complétés par la lecture de ses cours à l'université, sont, par leur densité, d'un accès malaisé. Un autre reproche que Picard adresse à la « science allemande » peut être résumé par la phrase suivante :

> La difficulté à juger de l'importance réelle des problèmes fait parfois attacher un grand prix à des questions purement formelles sans intérêt pour le fond. [Picard, 1916] p. 23

Que la « science allemande », quelle que soit la signification de l'expression, témoigne ou non de ce travers n'a aucune importance ici. Ce qui est par contre primordial pour nous, c'est que Picard voit cette difficulté comme l'un des problèmes à surmonter lorsque l'on pratique les sciences pures. On ne peut donc pas se contenter d'expliquer la controverse par la différence, très formelle, entre les formules données par Kronecker et celles produites

122. Voir [Davidoglou, 1901a] et [Davidoglou, 1901b].
123. Voir [Davidoglou, 1901a] p. 784.
124. Voir [Picard, 1916] p. 22.
125. *Ibid.*
126. Voir [Picard, 1916] p. 19.

par Picard, ni plaider uniquement la mauvaise foi de Picard, qui est certes réelle, mais pas suffisante.

Ce que Picard reproche essentiellement à Kronecker, c'est-à-dire la dépendance de la formule donnant le nombre de racines par rapport au signe du déterminant fonctionnel, montre surtout la différence entre leurs deux approches. Picard cherche à résoudre un problème particulier, alors que Kronecker veut montrer comment la résolution de ce problème n'est que l'une des multiples facettes que décrit sa théorie des caractéristiques. En effet, il a construit cette théorie pour qu'elle contienne à la fois le théorème de Sturm et les travaux de Cauchy. On a vu que cette double inclusion l'a amené à produire deux parties dans sa théorie : l'une que l'on a qualifié de topologico-algébrique, l'autre de topologico-analytique. Picard ne fait allusion qu'au travail analytique de Kronecker. Cette façon d'aborder la théorie de Kronecker est celle que l'on retrouve majoritairement dans les rares articles qui la citent, et on peut par exemple citer Phragmén qui, parlant du travail de Picard sur la recherche du nombre de solutions d'un système d'équations, affirme que ce dernier « l'a traité du point de vue de la théorie des *caractéristiques* de Kronecker, c'est-à-dire à l'aide d'intégrales définies ». Et c'est pourtant cette dualité qui, aux yeux de Kronecker, confère à sa théorie une telle généralité. Il ne se permet donc pas de perdre cette généralité par un artifice calculatoire tel que celui décrit par Picard : en comparant l'approche de Kronecker à celle de Picard, on voit mieux comment le besoin de *généralité* chez le premier s'exprime par la pluralité des points de vue.

# Chapitre 15

# Le théorème fondamental de l'algèbre


15.1 Introduction . . . . . **445**
15.1.1 Les quatre preuves de Gauss . . . . . 446
15.1.2 La seconde preuve de Gauss . . . . . 447
15.2 Les première et quatrième démonstrations . . . . . **454**
15.2.1 La première démonstration de Gauss de 1799 . . . . . 455
15.2.2 La quatrième démonstration de Gauss de 1849 . . . . . 468
15.2.3 La reprise par Kronecker . . . . . 478
15.3 La troisième démonstration . . . . . **483**
15.3.1 La version de Gauss . . . . . 483
15.3.2 La reprise par Kronecker . . . . . 491
15.4 Vers un nouvel énoncé . . . . . **493**


## 15.1 Introduction

Dans un article publié en 1991, Christian Gilain [Gilain, 1991] réécrit l'histoire[1] du théorème fondamental de l'algèbre en remplaçant sa description ternaire usuelle, *conjecture – essai de preuve – preuve*, par une distinction entre deux théorèmes : le théorème fondamental de l'algèbre et le théorème de factorisation linéaire[2]. C'est-à-dire entre l'existence d'un corps de décomposition d'un polynôme, et donc la *possibilité* de le factoriser dans ce corps par un produit de facteurs simples, et le fait que tout polynôme de $\mathbb{C}[x]$ admette une racine complexe : différence fondamentale entre la preuve de l'existence de racines d'un polynôme et la détermination de la nature de celles-ci pour un polynôme à coefficients réels. Il y a deux théorèmes, et donc deux histoires distinctes.

Ces deux histoires se croisent plusieurs fois, en particulier lors des quatre démonstrations que donne Gauss du théorème fondamental de l'algèbre, et comme nous le verrons,

1. Pour une histoire détaillée de ce théorème, voir [Gilain, 1991], [Remmert, 1991], [Petrova, 1973] ou [Dhombres et Alvarez, 2011] et [Dhombres et Alvarez, 2013].
2. TFA et TFL pour reprendre les abréviations de C. Gilain.

plus encore lors de leur reprise par Kronecker. Les démonstrations de Gauss ne sont pas forcément [3] les premières à être « justes », mais ce sont certainement les premières à mettre en évidence l'existence de deux théorèmes : il ne démontrera pas, comme ses prédécesseurs, le TFA en supposant admis le TFL, mais regardera toujours séparément l'existence d'une racine et sa nature. Cependant, même s'il ne voit pas le TFL comme un théorème admis ou une évidence qui n'a pas à être démontrée, ce n'est qu'un moyen pour accéder au TFA [4].

### 15.1.1 Les quatre preuves de Gauss

Gauss a fourni, entre 1799 et 1849 quatre preuves du théorème fondamental de l'algèbre, essentiellement différentes [5]. Le qualificatif de *fondamental* apparaît d'ailleurs dès [6] sa première démonstration de 1799 [7] :

> Pour le reste, il me semble que d'Alembert n'a pas convenablement recouru aux séries infinies et qu'elles ne sont pas du tout appropriées pour établir ce théorème fondamental de la théorie des équations. [Gauss, 1799] p. 10

Mais Gauss l'utilise à propos de la « théorie des équations », et non de l'*algèbre*. Traditionnellement, on considère que la première et la quatrième sont de nature topologique, tandis que la seconde est très fortement algébrique (Gauss dirait analytique), et la troisième utilise l'intégration des fonctions de deux variables et l'analyse complexe. On en tire ainsi, en première approche, une classification en trois grands types de preuves [8] :

— Le premier type (première et quatrième preuves de Gauss) regroupe toutes les preuves qui étudient les propriétés des polynômes complexes, et en particulier leur continuité sur un compact (c'est-à-dire en fait le principe du maximum) ou celles, plus récentes, qui utilisent les notions de topologie algébrique comme l'homotopie et

3. Voir à ce sujet [Gilain, 1991], [Houzel, 1989] et [Bachmacova, 1960].

4. Il ne faudrait pas croire que le TFA ait été accepté comme vrai pour les mathématiciens du XVIII$^{e}$ siècle : Leibniz, en cherchant la solution de l'équation $x^4 + a^4$ se trouve confronté à des facteurs du type $x \pm a\sqrt{\sqrt{-1}}$ qu'il n'arrive pas à combiner pour en faire des facteurs réels et conclut que le TFA est faux.

5.

— *Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse*. PhD thesis, Universität Helmstedt, 1799. In Werke III, pp. 1-30.

— Demonstratio nova altera theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse. Comm. Recentiores (Gottingae), 1816. In Werke III, pp 31–56.

— *Theorematis de resolubilitate functionum algebraicarum integrarum in factores reales Demonstratio tertia supplentum commentationis praecedentis* (Gottingae), 1816. In Werke III, pp 57–64.

— Beiträge zur Theorie der algebraischen Gleichungen, 1849. In Werke III, pp 73–85.

6. [Gauss, 1799] p. 10 : « Ceterum mihi videtur, ill. d'A. hic non recte ad series infinitas confugisse, hasque ad stabiliendum theorema hoc fundamentale doctrinae aequationum haud idoneas esse. »

7. Et non pas dans la seconde démonstration de 1815, comme cela est affirmé dans [Dhombres et Alvarez, 2013] p. 201.

8. S.S. Petrova ([Petrova, 1973]) divise les preuves du TFA en deux types : les preuves "analytiques" et les preuves "algébriques". les premières se subdivisent ensuite en deux groupes : celles qui utilisent les propriétés topologiques des courbes algébriques et celles qui utilisent l'analyse complexe.

le groupe fondamental. Elles se rattachent historiquement à la preuve de d'Alembert.

— Le second type (seconde preuve de Gauss) rassemble les preuves qui recherchent une utilisation minimale des notions étrangères à l'algèbre. On sait qu'une preuve qui ne prendrait pas en compte, ne serait-ce que de manière très lointaine, la construction des réels aurait peu de chance d'aboutir. Les démonstrations de ce groupe utilisent ainsi en général l'existence d'un corps de décomposition pour tout polynôme de $\mathbb{C}[x]$. On en trouve les prémices dans les démonstrations d'Euler (1749) et de Lagrange (1772) ; et les démonstrations plus récentes qui utilisent la théorie de Galois sont parfois incluses dans ce groupe.

— Le dernier type enfin (troisième preuve de Gauss) est celui qui regroupe les preuves qui utilisent des résultats de l'analyse complexe, que ce soit le théorème de Liouville ou celui de Rouché.

Kronecker, à la fin de la première partie de son cours sur la théorie des équations algébriques du semestre d'hiver 1890-91, dit du théorème fondamental de l'algèbre qu'il est « le plus important de cette première partie, on pourrait presque dire (...) le seul théorème de la théorie des équations algébriques » [9], et en effet il réserve une place toute particulière à ce théorème dans ses *Vorlesungen*. Il affirme en effet que [10] :

> La proposition, que *Gauss* nomme le théorème fondamental de l'algèbre, qui forme aussi véritablement le fondement de l'algèbre, doit apparaitre dans ces leçons plutôt au début qu'à la fin. [Kronecker, 1891a] pp. 1-2

Cependant, le nombre de travaux préliminaires pour arriver à ce théorème est tel qu'il n'apparaîtra en fait jamais avant le milieu de son cours.

Nous avons déjà vu à de nombreuses reprises Kronecker citer ou se référer à divers travaux de Gauss. Cette influence est encore une fois déterminante dans la façon dont il aborde dans son cours le théorème fondamental de l'algèbre. Cependant, la nature des racines d'une équation revêt chez Kronecker un statut très particulier qui l'amène à modifier les démonstrations de Gauss. Ces dernières sont toutes reprises dans son cours sur la théorie des équations algébriques, et occupent une place spéciale à l'intérieur de celui-ci. Plus encore, les diverses natures de ces preuves, tout au moins pour trois d'entre elles, sont à mettre en parallèle avec les différentes facettes de la théorie des caractéristiques.

### 15.1.2 La seconde preuve de Gauss

La seconde des quatre démonstrations de Gauss est celle que l'on peut qualifier de maximalement algébrique. Mais, comme toute démonstration du TFA, elle utilise la continuité de $\mathbb{R}$, c'est-à-dire ici le théorème de Bolzano. Or, on a vu la position de Kronecker

9. [Kronecker, 1891a] p. 268 : (...) *für diesen ersten Teil des wichtigsten, ja, man könnte fast behaupten, einzigen Theorems der Lehre von den algebraischen Gleichungen* (...)
10. On a déjà vu cette citation p. 162.

par rapport à ce théorème : il ne présentera donc que la partie factorisation linéaire de ce théorème. Cette preuve ne sera pas au centre dans notre travail, car elle ne trouve pas sa place, contrairement au trois autres, dans la théorie des caractéristiques. Cependant, nous allons voir que nous la retrouvons, sous plusieurs formes, dans les *Vorlesungen über die Theorie der algebraischen Gleichungen*, et nous allons donc en donner rapidement les grandes lignes et examiner la place qu'elle occupe dans le travail de Kronecker.

#### 15.1.2.1 La démonstration de Gauss

Cette seconde démonstration de Gauss, datée du 7 décembre 1815 et publiée l'année suivante, est celle dans laquelle la distinction entre TFA et TFL est la plus mise en avant. Dans celle-ci, il répond en fait aux critiques qu'il avait émises dans sa thèse de 1799 à propos des travaux d'Euler et de d'Alembert, lorsqu'il montrait que ces deux mathématiciens avaient plus ou moins implicitement supposé la validité du TFL pour ensuite démontrer le TFA, c'est-à-dire qu'ils avaient supposé l'existence des racines pour ensuite en déterminer la forme. Regardons maintenant, dans les grandes lignes, comment s'articule cette démonstration. Le but est simplement ici de faire apparaitre les grandes étapes de l'article de Gauss et de mettre en avant les articulations de sa démonstration.

**Préliminaires** Gauss utilise, à la place des racines supposées de l'équation étudiée, les indéterminées $x_1, ..., x_n$ et il travaille donc dans des anneaux du type $\mathbb{R}[x, y, ..., z, x_1, ..., x_n]$. Il a donc besoin de certains résultats préliminaires ayant trait à ces objets mathématiques. Tout d'abord sur la divisibilité : il reprend la division euclidienne dans $\mathbb{R}[x]$ et le théorème de Bézout.

Une partie non négligeable de son travail sera ensuite consacrée aux polynômes symétriques. Il démontre, à l'aide de l'ordre lexicographique, le théorème que l'on appelle aujourd'hui le *Théorème fondamental des polynômes symétriques* et montre que les polynômes symétriques élémentaires $\lambda_1, ..., \lambda_n$ forment une famille libre. Comme il lui faut « compter » le nombre de facteurs des polynômes symétriques qu'il introduit, il donne aussi quelques résultats de dénombrement.

Pour plus de clarté, posons tout de suite les notations qui interviennent tout au long de la démonstration :

$$\begin{aligned} Y(x) &= x^n - l_1 x^{n-1} + l_2 x^{n-2} - ... + (-1)^n l_n \in \mathbb{R}[x] \text{ où } l_1, ..., l_n \in \mathbb{R} \\ y(x, x_1, ..., x_n) &= x^n - x_1 x^{n-1} + x_2 x^{n-2} - ... + (-1)^n x_n \in \mathbb{R}[x, x_1, ..., x_n] \\ v(x, x_1, ..., x_n) &= (x - x_1)...(x - x_n) = x^n - \lambda_1 x^{n-1} + \lambda_2 x^{n-2} - ... + (-1)^n \lambda_n \end{aligned}$$

Le principe général qu'il utilisera tout au long de son travail consistera à transférer des propriétés démontrées pour $v$ sur $Y$ en passant par $y$. Pour cela il emploiera « à la chaîne »

le fait que si

$$\varphi(x, y, ...z, \lambda_1, ..., \lambda_n) = \psi(x, y, ...z, \lambda_1, ..., \lambda_n)$$

alors

$$\varphi(x, y, ...z, x_1, ..., x_n) = \psi(x, y, ...z, x_1, ..., x_n)$$

**Le discriminant** Comme dans le cours de Kronecker, les polynômes symétriques et le discriminant sont les outils essentiels dans ce travail de Gauss. Il définit alors, à l'aide des résultats qu'il a démontrés sur les polynômes symétriques, le discriminant des polynômes $y$ et $Y$, que l'on notera $D(y)$ et $D(Y)$, et démontre les deux propriétés suivantes :

**Propriété 1** $Y$ et $Y'$ ont un facteur commun si et seulement si le discriminant de $Y$ est nul.

**Propriété 2** $Y$ peut se décomposer en facteurs de discriminant non nul.

C'est ici que Gauss, de façon particulièrement explicite, distingue le TFA et le TFL. En effet, il produit une démonstration en deux temps : d'abord en supposant comme ses prédécesseurs que $Y$ peut se décomposer en produit de facteurs du premier degré, puis en utilisant les transferts de propriétés mentionnées plus haut.

**Démonstration du théorème fondamental de l'algèbre** Gauss introduit une fonction auxiliaire :

$$\varsigma(x, u, x_1, ..., x_n) = \prod_{i<j} [u - (x_i + x_j)x + x_i x_j]$$

qui est un polynôme de degré $\dfrac{n(n-1)}{2}$ en u.

$\varsigma$ est symétrique en les indéterminées $x_1, ..., x_n$, donc d'après le théorème fondamental des polynômes symétriques $\exists z \in \mathbb{R}[x, u, x_1, ..., x_n]$ telle que

$$z(x, u, \lambda_1, ..., \lambda_n) = \varsigma(x, u, x_1, ..., x_n)$$

On pose ensuite

$$Z(x, u) = z(x, u, l_1, ..., l_n)$$

**Pourquoi introduire cette fonction auxiliaire ?**

**Lien entre $Z$ et $Y$** On a

$$\begin{aligned} & z(x, x^2, \lambda_1, ..., \lambda_n) \\ = \; & \varsigma(x, x^2, x_1, ..., x_n) \end{aligned}$$

$$\begin{aligned} &= \prod_{i\leq j} x^2 - (x_i + x_j)x + x_i x_j \\ &= \prod_{i\leq j} (x - x_i)(x - x_j) \\ &= v^{m-1}(x, x_1, ..., x_n) \\ &= y^{m-1}(x, \lambda_1, ..., \lambda_n) \end{aligned}$$

Donc, grâce au fait que les polynômes symétriques élémentaires $\lambda_1, ..., \lambda_n$ forment une famille libre. :

$$\begin{aligned} & z(x, x^2, x_1, ..., x_n) \\ = \; & y^{m-1}(x, x_1, ..., x_n) \end{aligned}$$

et ainsi

$$\begin{aligned} & Z(x, x^2) \\ = \; & z(x, x^2, l_1, ..., l_n) \\ = \; & y^{m-1}(x, l_1, ..., l_n) \\ = \; & Y^{m-1}(x) \end{aligned}$$

**Lien entre le degré de $Z$ en u et celui de $Y$** Si le degré de $Y$ est $n = 2^m r$ avec $r$ impair, alors celui de $Z$ est

$$\frac{n(n-1)}{2} = 2^{m-1} r'$$

où $r'$ est impair. La démonstration de Gauss se faisant par récurrence sur l'entier $m$ du degré de $Y$, cela lui permettra d'utiliser l'hypothèse de récurrence.

**Utilisation du discriminant**

**Propriété 3** Si $D(Y) \neq 0$ alors $D(Z) \neq 0$.

Par ailleurs, on a :

**Propriété 4** Si

$$U_0 = \frac{dZ}{du}(x_0, u_0) \neq 0$$

Alors il existe $\psi \in \mathbb{R}[x]$ telle que

$$Z(x, u_0 + \frac{A}{U_0} - \frac{B}{U_0} x) = Z(x_0, u_0)\psi(x)$$

De plus, **si** $D(Y) \neq 0$ alors $D(Z) \neq 0$ et il existe $x_0$ tel que $D(Z)(x_0) \neq 0$. Ainsi $\frac{dZ}{du}(x_0, u)$ et $Z(x_0, u)$ n'ont pas de facteur commun.

**Si de plus** $\exists u_0$ tel que $Z(x_0, u_0)$, alors $(u - u_0)$ divise $Z$ et ne divise pas $\frac{dZ}{du}(x_0, u)$ donc $U_0 = \frac{dZ}{du}(x_0, u_0) \neq 0$.

Ainsi, par la proposition précédente,

$$Z(x, u_0 + \frac{A}{U_0} - \frac{B}{U_0}x) = Z(x_0, u_0)\psi(x)$$

et $Z(x, u)$ est divisible par $u - u_0 - \frac{A}{U_0} + \frac{B}{U_0}x$.

Donc $Z(x, x^2)$ est divisible par $x^2 - u_0 - \frac{A}{U_0} + \frac{B}{U_0}x$, et donc a deux racines complexes.

**Démonstration par récurrence du théorème fondamental de l'algèbre**
Soit $Y$ de degré $n = 2^m r$ où $r$ est impair.

Si $m = 0$, alors le degré de $Y$ est impair et le théorème des valeurs intermédiaires nous affirme l'existence d'une racine : c'est ici qu'intervient donc le théorème que Bolzano. On peut d'ailleurs noter que c'est en travaillant sur cette démonstration de Gauss que Bolzano proposera une démonstration du théorème des valeurs intermédiaires [11].

Supposons la proposition vraie au rang $2^s r$ où $r$ est impair et $s < m$ et montrons-le pour $s = m$.

Si $D(Y) = 0$, on peut toujours décomposer $Y$ en produit de facteurs de discriminants non nuls. On peut donc supposer que $D(Y) \neq 0$. On a vu que le degré de $Z(x_0, u)$ est $2^{m-1}r'$ où $r'$ est impair : on peut ainsi appliquer l'hypothèse de récurrence et affirmer qu'il existe $u_0 \in \mathbb{C}$ tel que $Z(x_0, u_0) = 0$.

Comme $D(Y) \neq 0$ et qu'il existe $u_0$ tel que $Z(x_0, u_0) = 0$, alors $Z(x, x^2) = Y^{m-1}(x)$ a

11. Lorsque Bolzano publie en 1817 sa *Rein analytischer Beweis*, Gauss vient de rédiger deux nouvelles démonstrations du théorème fondamental de l'algèbre. Bien plus que dans sa première preuve, il essaye de donner des arguments uniquement algébriques, débarrassés de tout appel à la géométrie. Le seul présupposé auquel il fait appel dans la première de ces deux preuves récentes est justement le théorème sur lequel la *Rein analytischer Beweis* porte. Bolzano rappelle l'énoncé du théorème démontré par Gauss dans sa préface : « Toute fonction algébrique rationnelle entière d'une variable peut être décomposée en facteurs réels du premier ou du second degré »([Bolzano *et al.*, 1905] p. 136). Il note aussi que pour parfaire la démonstration de Gauss, et en exclure tout appel à l'évidence géométrique, il reste encore à démontrer la proposition qu'il se propose justement de prouver dans sa *Rein analytischer Beweis.* De plus, du théorème de Gauss, on peut bien évidemment déduire celui que veut démontrer Bolzano. C'est ce que fait Lagrange dans son traité *Sur la résolution des équations numériques.*

une racine complexe [12]. Ce qui démontre le théorème.

**Conclusion** Gauss travaille dans le cadre de sa seconde preuve dans l'anneau $\mathbb{R}[x, u, x_1, ..., x_n]$. $\mathbb{R}$ étant un corps, c'est un anneau intègre.

L'anneau $\mathbb{R}[x, u, x_1, ..., x_n]$ est donc lui aussi intègre et l'on peut donc construire son corps des fractions. $\mathbb{R}[x, u, x_1, ..., x_n]/(P)$ est lui aussi intègre si $P$ est irréductible. Son corps des fractions sera un corps de rupture pour $P$.

Ce que Gauss fait, c'est construire le corps de rupture d'un polynôme dans lequel son polynôme de départ, $Y$, possédera un facteur simple [13].

De l'égalité

$$Z(x, u_0 + \frac{x_0 - x}{U_0} X_0) = Z(x_0, u_0)\psi(x_0, u_0, \frac{x_0 - x}{U_0})$$

il tirera ainsi

$$Z(x, v + \frac{x_0 - x}{U_0} X_0) = Z(x_0, v)\psi(x_0, v, \frac{x_0 - x}{U_0}) = \varphi(u, x, v, x_0) \in \mathbb{R}[u, x, v]$$

On pose $Z_{x_0}(u) = Z(x_0, u)$. Le polynôme $Z_{x_0}$ construit par Gauss a deux caractéristiques : son discriminant n'est pas nul et l'hypothèse de récurrence de sa démonstration du théorème fondamental de l'algèbre lui permet d'affirmer que ce polynôme se décompose en produit de facteurs simples. Or il a montré que pour toute racine $u_0$ de $Z_{x_0}$, $Z(x, u)$ est divisible par $u - \left(u_0 + \frac{x_0 - x}{U_0} X_0\right)$.

Donc $Z(x, u) - \left(u - \left(u_0 + \frac{x_0 - x}{U_0} X_0\right)\right) \varphi(u, x, v, x_0)$ s'annule pour tout $v = u_0$ racine de $Z_{x_0}$. Toutes ces racines étant simples (le discriminant n'est pas nul), cette différence est divisible par $Z_{x_0}(v)$.

Finalement,

$$Z(x, u) \equiv \left(u - \left(u_0 + \frac{x_0 - x}{U_0} X_0\right)\right) \varphi(u, x, v, x_0) \mod (Z_{x_0}(v))$$

et donc

$$Y^{m-1}(x) = Z(x, x^2) \equiv \left(x^2 - \left(u_0 + \frac{x_0 - x}{U_0} X_0\right)\right) \varphi(x^2, x, v, x_0) \mod (Z_{x_0}(v))$$

Dans le corps de rupture de $(Z_{x_0}(v))$, $Y$ a donc bien un facteur quadratique, donc deux simples.

12. Et en fait on a démontré qu'il y en avait même deux.
13. Il s'agira d'un facteur quadratique, ce qui lui procurera évidemment deux facteurs simples.

#### 15.1.2.2 Dans les leçons de Kronecker

La seconde démonstration de Gauss se trouve, comme les trois autres, dans les *Vorlesungen* de Kronecker : avant 1882, dans le cadre de la théorie des *Gattungen*, et après comme exemple d'utilisation des *Modulsystem* et de ce qu'il nomme le *principe de Galois*[14]. On voit donc que dans tous les cas elle ne fait pas partie de la première partie du cours de Kronecker, et l'étude approfondie de sa reprise par Kronecker dépasse le cadre que nous nous sommes fixé. Néanmoins, à défaut d'en étudier le contenu et la forme, nous allons tout de même examiner la place qu'elle occupe dans les *Vorlesungen*.

Ce théorème de factorisation linéaire, Kronecker l'appelle à partir de 1887 *Théorème fondamental de l'arithmétique générale* et en donne une démonstration dans un article[15] intitulé *Ein Fundamentalsatz der allgemeinen Arithmetik*. C'est ce travail utilisant la notion de *Modulsystem* qui conclut ses *Vorlesungen über die Theorie der algebraischen Gleichungen* des semestres 1884-1885 ([Kronecker, 1885a] pp. 408-409), 1886-1887 ([Kronecker, 1887a] p. 319 et [Kronecker, 1887b] p. 422), 1888-1889 ([Kronecker, 1889b] p. 315. Il termine en donnant l'exemple de la détermination d'un *Modulsystem* permettant la factorisation de $x^2 - 1$.), et 1890-1891 ([Kronecker, 1891a] p. 342). Cependant, même avant la publication de ses *Grundzüge* et sa théorie des *Modulsystem*, une démonstration inspirée de la preuve de *Gauss* de 1815 est présente dans ses leçons[16] :

> La propriété du *Gattung* de fonctions algébriques de $\mathfrak{f}_1, \mathfrak{f}_2, ...\mathfrak{f}_n$ développée à la fin du paragraphe précédent constitue le véritable fondement dans l'une des preuves de *Gauss* de 1815 sur l'existence des racines des équations algébriques, que j'ai déjà, dans leçons tenues dans cette Université lors du semestre d'hiver 1870/71 et depuis dans mes leçons sur la théorie des équations algébriques, donné toujours de la même façon. [Kronecker, 1882b] p. 294

Ainsi, on trouve dans l'ensemble des manuscrits, dans la partie consacrée à la théorie des *Gattungen*, un travail qui correspondra en fait à ce que l'on décrirait aujourd'hui comme étant la construction du corps de rupture d'un polynôme, et qui pour Kronecker est le fondement de « l'existence des racines d'une équation algébrique ».

Jusqu'à la publication des *Grundzüge* en 1883, Kronecker donne ainsi dans son cours deux démonstrations du théorème fondamental de l'algèbre : celle qu'il tire de la théorie des caractéristiques, et celle qu'il construit à partir de la seconde démonstration de Gauss[17], tout en employant le vocabulaire des *Gattungen*. De ce travail de Gauss, il dit d'ailleurs

14. « Galoissche Princip ».
15. Voir [Kronecker, 1887c].
16. Die am Schlüsse des vorigen Paragraphen entwickelte Eigenschaft der Gattungen algebraischer Functionen von $\mathfrak{f}_1, \mathfrak{f}_2, ...\mathfrak{f}_n$ bildet die wesentlichste Grundlage bei einem dem *Gauss*'schen Beweise von 1815 nachgebildeten Existenzbeweise für die Wurzeln algebraischer Gleichungen, welchen ich schon in meiner an der hiesigen Universität gehaltenen Wintervorlesung von 1870/71 vorgetragen und seitdem in meinen Vorlesungen über die Theorie der algebraischen Gleichungen stets fast in derselben Weise gegeben habe.
17. On peut par exemple voir [Kronecker, 1885b] pp. 382 et suivantes.

ceci [18] :

> Il convient de souligner que dans ce développement, comme dans le traité cité de *Gauss* de 1815, un genre très spécial de l'existence des grandeurs algébriques est représenté par rapport à toutes les autres méthodes de preuve. Cette différence ressort particulièrement fortement lorsque l'on analyse précisément pour les équations de degré impair la notion ci-dessus de racine réelle d'une équation algébrique, comme je l'ai fait régulièrement dans mes cours à l'Université ; et je compte aussi publier cette analyse à une prochaine occasion. La preuve développée ici contient en fait bien un fondement de l'existence *arithmétique* des grandeurs algébriques et c'est pourquoi elle s'insère dans la représentation systématique que j'ai essayé de donner dans le présent ouvrage. [Kronecker, 1882b] p. 296

Cette « analyse » sera bien publiée comme nous l'avons vu quelques années plus tard en 1887 sous le titre *Ein Fundamentalsatz der allgemeinen Arithmetik*, et s'intégrera en effet, bien plus que ce qu'il fait dans ses cours qui précèdent la publication de cet article, dans la présentation d'une *Arithmétique générale*. Le contenu de cet article prendra d'ailleurs place, à partir du cours de 1886-1887, à la fin de son cours sur les équations algébriques.

Finalement, le théorème fondamental de l'algèbre et le théorème de factorisation linéaire constitueront, dans l'ensemble de ses *Vorlesungen* une conclusion : conclusion de la première partie et de la théorie des caractéristiques, conclusion de la seconde partie et de la théorie des *Gattungen* et des *Modulsystem*.

Dans ce chapitre nous allons revenir sur les trois autres démonstrations, et nous les mettrons en perspective avec la façon dont Kronecker les traite dans ses cours sur la théorie des équations algébriques, et en particulier dans la section qu'il consacre à la théorie des caractéristiques.

## 15.2 Les première et quatrième démonstrations

Kronecker affirme dans son cours de l'hiver 1890-91, à propos du théorème fondamental de l'algèbre, que [19]

> Le cœur de la première et la quatrième preuve de ce théorème de *Gauss* (1799 resp. 1849) n'est au fond rien de plus que la détermination de la caractéristique. Concernant

18. Es verdient hervorgehoben zu werden, dass in dieser Entwickelung, wie in der citirten *Gauss*'schen Abhandlung von 1815, eine viel *speciellere* Art der Existenz algebraischer Grössen dargelegt wird, als bei allen anderen Beweismethoden. Dieser Unterschied tritt besonders deutlich hervor, wenn man den oben für Gleichungen ungraden Grades vorausgesetzten Begriff der reellen Wurzeln algebraischer Gleichungen genauer analysirt, wie ich es in meinen Universitäts-Vorlesungen regelmässig gethan habe ; und ich gedenke auch diese Analyse bei einer nächsten Gelegenheit durch den Druck zu veröffentlichen. Der hier entwickelte Beweis enthält recht eigentlich eine Begründung der *arithmetischen* Existenz der algebraischen Grössen und fügt sich deshalb in die systematische Darstellung, welche ich in der vorliegenden Arbeit zu geben versucht habe, vollständig ein.

19. Der Kern des ersten und vierten Beweises dieses Satzes von *Gauss* (1799 bzw. 1849) ist im Grunde genommen nichts anderes als die Bestimmung der Charakteristik. In Bezug auf die Bedeutung, welche *Gauss* dem Theorem beilegte, und seinem algebraischen Äquivalente müssen wir auf die später folgenden Erörterungen verweisen.

> la signification que *Gauss* joint au théorème, et à son équivalent algébrique, nous renvoyons aux discussions qui suivront plus tard. [Kronecker, 1891a] p. 243

Kronecker fait bien la distinction entre le TFA, tel que Gauss le démontre dans ses versions topologiques de 1799 et 1849, et la version algébrique de la seconde démonstration, qui prendra sa place dans la seconde partie de ses leçons. Car le Théorème fondamental de l'algèbre est finalement pour Kronecker, nous allons le voir, un théorème profondément topologique.

Gauss l'affirmait à la fin de sa première démonstration [20] :

> Enfin, j'observe qu'il n'est pas tout à fait impossible de présenter également la démonstration précédente que j'ai bâtie sur des principes de géométrie, dans une forme purement analytique : mais la représentation que j'ai développée ici m'a paru s'avérer moins abstraite et j'ai montré ici la vraie nouveauté de ce qui devait être prouvé beaucoup plus clairement que ce que l'on peut attendre d'une démonstration analytique. [Gauss, 1799] p. 29

La seconde démonstration, qui est certes *analytique*, n'apporterait pas, d'après Gauss, la même clarté que la première démonstration, d'inspiration plus *géométrique*. Cet argument, on le retrouvera chez Kronecker, et cette recherche de la « vraie nouveauté », de la vraie nature est commune au deux mathématiciens : pour pouvoir les atteindre, pour pouvoir en atteindre le « coeur », il est nécessaire, tout au moins dans une première approche, de passer par des interprétations géométriques. Avant d'étudier la façon dont Kronecker produit ses démonstrations du TFA à l'aide de la théorie des caractéristiques, nous allons donc examiner en détail les première et la quatrième preuve de Gauss.

### 15.2.1 La première démonstration de Gauss de 1799

Cette première démonstration constitue en fait la *dissertation* que Gauss, alors âgé de vingt-deux ans, soutiendra et publiera. Elle se compose de deux parties : dans la première, il revient sur les démonstrations du TFA par ses prédécesseurs, et dans la seconde il propose sa propre preuve, « rigoureuse », du théorème [21] :

> L'essai présenté en 1799, *Demonstratio nova theorematis, omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse*, avait un double objectif ; celui de montrer en effet premièrement l'insuffisance et l'aspect illusoire de toutes les preuves tentées jusqu'alors de cet important théorème de la théorie des équations algébriques, et deuxièmement d'en donner une nouvelle preuve parfaitement rigoureuse. [Gauss, 1849] p. 73

20. Denique observo, minime impossibile esse, ut demonstratio praecedens, quam hic principiis geometricis superstruxi, etiam in forma mere analytica exhibeatur : sed eam repraesentationem, quam hic explicavi, minus abstractam evadere credidi, verumque nervum probandi hic multo clarius ob oculos poni, quam a demonstratione analytica exspectari possit.

21. Die im Jahre 1799 erschienene Denkschrift, *Demonstratio nova theorematis, omnem functionem algehraicam rationalem integrum unius variabiUs in factores reales primi vel secundi gradus resolvi posse*, hatte einen doppelten Zweck, nemlich erstens, zu zeigen, dass sämmtliche bis dahin versuchte Beweise dieses wichtigsten Lehrsatzes der Theorie der algebraischen Gleichungen ungenügend und illusorisch sind, und zweitens, einen neuen vollkommen strengen Beweis zu geben.

#### 15.2.1.1 La première partie

Gauss commence par rappeler l'équivalence entre l'existence d'une valeur $\alpha$ qui satisfait l'équation

$$X = x^m + a_1 x^{m-1} + a_2 x^{m-2} + etc \cdot + a_{m-2} x^2 + a_{m-1} x + a_m$$

et la possibilité de factoriser $X$ par $x - \alpha$ et par clarifier la notion de racine multiple. Gauss s'interroge ensuite sur la nature de ce $\alpha$ : dans le cas des équations de degré inférieur ou égal à quatre, il affirme que « les algébristes ont été forcés d'inventer une quantité imaginaire dont le carré est -1, et [qu']ils ont alors remarqué, [que] si les quantités de la forme $a + b\sqrt{-1}$ étaient admises de la même manière que les quantités réelles, (...) le théorème était vrai » [22]. En fait, Gauss fait ici allusion à la possibilité de résoudre par radicaux ces équations, possibilité qu'il dit ne pas pouvoir étendre aux équations de degrés supérieurs : ce refus, plus encore que celui d'une induction non démontrée, semble montrer qu'il préfigure cette impossibilité, telle que la présentera bien plus tard Abel. Mais ces *quantités imaginaires*, Gauss refuse de les utiliser [23] :

> Du reste, depuis le temps où les analystes ont découvert qu'il y a une infinité d'équations qui n'ont aucune racine à moins d'admettre les quantités de la forme $a + b\sqrt{-1}$ comme un genre particulier de quantités appelées quantités imaginaires et distinctes des *réels*, ce qui supposerait que ces quantités aient été étudiées et introduites dans toute l'analyse ; avec quelles justifications ? Je ne vais pas en discuter ici. Je vais retirer de ma preuve toute aide des quantités imaginaires, même si je pourrais me servir de cette liberté que tous ceux qui s'occupent d'analyse utilisent. [Gauss, 1799] p. 4

Et ce sera donc sans faire appel au nombre de la forme $a + b\sqrt{-1}$ qu'il donnera sa démonstration. En rejetant les nombres imaginaires, Gauss évite de « choquer » une partie de la communauté mathématique, ce qui pourra lui permettre de mieux diffuser son travail. Cependant, il souhaite pouvoir les utiliser [24], mais seulement une fois qu'elles seront « établies de façon solide ».

Cela étant posé, il revient alors sur les démonstrations de ses prédécesseurs, et il montre en cela, malgré son jeune âge, une culture étendue sur les mathématiques de son époque.

**Un argument circulaire** Le premier problème que relève Gauss, est l'argument circulaire qui est le plus souvent employé : on suppose implicitement que l'équation admet $m$ racines, pour en démontrer ensuite l'existence. Même lorsque l'auteur « semble avoir

22. Voir [Gauss, 1799] p. 4.

23. Ceterum ex eo tempore, quo analystae comperti sunt, infinite multas aequationes esse, quae nuUam omnino radicem haberent, nisi quantitates formae $a + b\sqrt{-1}$ admittantur, tales quantitates fictitiae tamquam peculiare quantitatum genus, quas *imaginarias* dixerunt, ut a *realibus* distinguerentur, consideratae et in totam analysin introductae sunt ; quonam iure ? hoc loco non disputo. – Demonstrationem meam absque omni quantitatum imaginariarum subsidio absolvam, etsi eadem libertate, qua omnes recentiores analystae usi sunt, etiam mihi uti liceret.

24. [Gauss, 1799] p. 6 : « Si l'on doit conserver les quantités imaginaires en analyse (ce qui semble pour certaines raisons plus souhaitable que de les abolir, une fois qu'elles sont établies de façon solide), alors elles doivent être considérées comme tout aussi possible que les quantités réelles. »

perçu la difficulté »[25], il suppose alors l'existence d'« une racine, sinon possible, tout au moins impossible »[26]. Le problème se trouve alors dans le qualificatif *impossible (impossibiles)* : ces racines *impossibles* ne sont généralement pas définies. Si on les considère comme étant des quantités imaginaires, alors l'argument redevient circulaire, sinon on peut même se « demander s'il ne peut pas exister d'équation qui n'ait même pas de racine impossible »[27]. Dans tous les cas, la *nature* de ces racines reste problématique, et un travail devra être fait sur celle-ci. C'est pourquoi la décision de Gauss de ne travailler qu'avec des nombres réels lui permet d'éviter toutes ces difficultés.

**La démonstration de d'Alembert** Le premier mathématicien auquel s'intéresse Gauss est Jean Lerond d'Alembert, qui envoie en 1746 à l'académie de Berlin un mémoire intitulé *Recherches sur le calcul intégral*[28] dans la première partie duquel on trouve une démonstration du théorème fondamental de l'algèbre. Cette démonstration[29], qui utilise un développement en série de puissances fractionnaires, est souvent considérée comme la première « tentative sérieuse de démonstration »[30]. Gauss commence par formuler des reproches qui touchent essentiellement à la forme, mais qui peuvent de son propre aveu « facilement être corrigés de telle manière que l'objection soit complètement levée »[31]. Les deux problèmes principaux touchent pour Gauss premièrement à l'utilisation excessive et injustifiée des infiniment petits, et surtout à l'utilisation d'une série dont la convergence n'est pas certaine. Il s'abstiendra donc d'utiliser ces outils dans ses propres démonstrations.

**La démonstrations d'Euler** Les démonstrations d'Euler, essentiellement algébriques, sont publiées en 1851 dans son travail intitulé *Recherches sur les racines imaginaires des équations*. Dans ce mémoire Euler donne deux démonstrations du TFA. Dans la première, pour démontrer que toute racine d'une équation est de la forme $a + b\sqrt{-1}$, il utilise le fait que tout polynôme de degré $n$ a $n$ racines : il retombe lui aussi sur l'argument circulaire que l'on a décrit plus haut. Dans la seconde, il travaille sur les grandeurs de la forme $a + b\sqrt{-1}$, en montre la stabilité, et prouve que toute racine d'un polynôme est de cette forme. Son argument principal consiste à affirmer que toute équation peut être résolue par radicaux. Cette affirmation, Gauss a bien compris, dès 1799, qu'elle ne reposait sur aucun fondement[32] :

> Contre ce raisonnement on peut objecter qu'après tout le travail fourni par de si grands mathématiciens il reste très peu d'espoirs d'arriver à une solution générale des

25. [Gauss, 1799] §3.

26. [Gauss, 1799] p. 5 : « Quidam auctores, qui debilitatem huius methodi percepisse videntur, tamquam axioma assumunt , quamvis aequationem revera habere radices, si non possibiles, impossibiles. »

27. Voir [Gauss, 1799] §4.

28. Voir [d'Alembert, 1746].

29. Pour plus de détails, on pourra lire par exemple [Petrova, 1973] ou [Houzel, 1989].

30. Voir [Gilain, 1991].

31. [Gauss, 1799] p. 9 : « Quamvis vero haec obiectio per se gravissima sit, tarnen hie ad solam dictionis formam pertinet, quae facile ita corrigi potest, ut illa penitus destruatur. »

32. Contra hoc ratiocinium obiici potest, post tot tantorum geometrarum labores perexiguam spem superesse, ad resolutionem generalem aequationum algebraicarum umquam perveniendi, ita ut magis magisque verisimile fiat, talem resolutionem omnino esse impossibilem et contradictoriam.

> équations algébriques. Il apparaît de plus en plus probable qu'une telle solution soit tout à fait impossible et contradictoire. [Gauss, 1799] p. 17

Et il donne ainsi la bonne impulsion pour les recherches à venir sur la résolubilité des équations algébriques de degré supérieur à quatre. Gauss termine en revenant sur les démonstrations de Foncenex et de Lagrange[33] qui comblent d'après lui une grande partie des lacunes de la démonstration d'Euler, mais qui, chacune, suppose acquis le fait qu'une équation de degré $n$ a $n$ racines.

#### 15.2.1.2 La seconde partie

Après avoir donc lors des douze premiers chapitres de son mémoire, qui en compte vingt-quatre, mis en évidence les failles des démonstrations de ses prédécesseurs, Gauss commence sa démonstration par deux lemmes :

**Lemme 1** $\forall m \in \mathbb{N}^*$,
$f_m(x) = \sin\varphi x^m - \sin m\varphi \cdot r^{m-1}x + \sin(m-1)\varphi \cdot r^m$ est divisible par $x^2 - 2\cos\varphi \cdot rx + r^2$

**Démonstration du lemme 1**

**Pour m=1** $f_1(x) = \sin\varphi \cdot x - \sin\varphi x = 0$ qui est divisible par n'importe quel polynôme.

**Pour m=2**

$$\begin{aligned} f_2(x) &= \sin\varphi \cdot x^2 - \sin 2\varphi \cdot rx + \sin\varphi \cdot r^2 \\ &= \sin\varphi \cdot x^2 - 2\sin\varphi\cos\varphi \cdot rx + \sin\varphi \cdot r^2 \\ &= \sin\varphi \cdot (x^2 - 2\cos\varphi \cdot rx + r^2) \end{aligned}$$

**Pour m>2** Effectuons le produit :

$$\begin{aligned} & (\sin\varphi \cdot x^{m-2} + \sin 2\varphi \cdot rx^{m-3} + \sin 3\varphi \cdot r^2x^{m-4} + ... + \sin(m-1)\varphi \cdot r^{m-2}) \\ \times\ & (x^2 - 2\cos\varphi \cdot rx + r^2) \\ =\ & \sin\varphi \cdot x^m + (\sin 2\varphi - 2\cos\varphi\sin\varphi)rx^{m-1} \\ & +(\sin 3\varphi - 2\cos\varphi\sin 2\varphi + \sin\varphi)r^2x^{m-2} \\ & +(\sin 4\varphi - 3\cos\varphi\sin 2\varphi + \sin 2\varphi)r^3x^{m-3} \\ & +... \\ & +(\sin(m-1)\varphi - 2\cos\varphi\sin(m-2)\varphi + \sin(m-3)\varphi)r^{m-2}x^2 \\ & +(\sin(m-2)\varphi - 2\cos\varphi\sin(m-1)\varphi)r^{m-1}x \end{aligned}$$

33. Démonstrations issues respectivement de [de Foncenex, 1759] et [Lagrange, 1772].

$$+\sin(m-1)\varphi \cdot r^m$$

Or $2\sin a\cos b = \sin(a+b) + \sin(a-b)$ donc

$$\begin{aligned} & -2\cos\varphi\sin(k-1)\varphi + \sin(k-2)\varphi \\ = \; & -(\sin k\varphi + \sin(k-2)\varphi) + \sin(k-2)\varphi \\ = \; & -\sin(k-2)\varphi \end{aligned}$$

Donc

$$\begin{aligned} & (\sin\varphi \cdot x^{m-2} + \sin 2\varphi \cdot rx^{m-3} + \sin 3\varphi \cdot r^2x^{m-4} + ... + \sin(m-1)\varphi \cdot r^{m-2}) \\ \times \; & (x^2 - 2\cos\varphi \cdot rx + r^2) \\ = \; & \sin\varphi x^m - \sin m\varphi \cdot r^{m-1}x + \sin(m-1)\varphi \cdot r^m \end{aligned}$$

Ce qui conclut la démonstration du lemme.

**Lemme 2** Si[34] l'on a $r \in \mathbb{R}$ et $\varphi \in \mathbb{R}$ tels que

$$\begin{aligned} r^m\cos m\varphi \quad & +a_1r^{m-1}\cos(m-1)\varphi + a_2r^{m-2}\cos(m-2)\varphi + ... \\ & +a_{m-2}r^2\cos 2\varphi + a_{m-1}r\cos\varphi + a_m = 0 \;\; (1) \\ r^m\sin m\varphi \quad & +a_1r^{m-1}\sin(m-1)\varphi + a_2r^{m-2}\sin(m-2)\varphi + ... \\ & +a_{m-2}r^2\sin 2\varphi + a_{m-1}r\sin\varphi = 0 \;\; (2) \end{aligned}$$

alors le polynôme $X = x^m + a_1x^{m-1} + a_2x^{m-2} + etc. + a_{m-2}x^2 + a_{m-1}x + a_m$ sera divisible par :

1. $x^2 - 2\cos\varphi \cdot rx + r^2$ si $r\sin\varphi \neq 0$.
2. $x - r\cos\varphi$ si $r\sin\varphi = 0$.

**Démonstration du lemme 2** Dans

$$\begin{array}{rcl} \sin\varphi \cdot rx^m & -\sin m\varphi \cdot r^m x & +\sin(m-1)\varphi \cdot r^{m+1} \\ a_1\sin\varphi \cdot rx^{m-1} & -a_1\sin(m-1)\varphi \cdot r^{m-1}x & +a_1\sin(m-2)\varphi \cdot r^m \\ a_2\sin\varphi \cdot rx^{m-2} & -a_2\sin(m-2)\varphi \cdot r^{m-2}x & +a_2\sin(m-3)\varphi \cdot r^{m-1} \\ \text{etc...} & \text{etc...} & \text{etc...} \\ a_{m-2}\sin\varphi \cdot rx^2 & -a_{m-2}\sin 2\varphi \cdot r^2x & +a_{m-2}\sin\varphi \cdot r^3 \\ a_{m-1}\sin\varphi \cdot rx & -a_{m-1}\sin\varphi \cdot rx & * \\ a_m\sin\varphi \cdot r & * & +a_m\sin(-\varphi) \cdot r \end{array}$$

34. Gauss fait remarquer que ce théorème a été démontré par Euler dans son *Introduction à l'Analyse infinitésimale.*

chaque ligne est divisible par $x^2 - 2\cos\varphi \cdot rx + r^2$, donc leur somme l'est aussi. De plus la somme des termes de la seconde colonne est égale à 0 d'après (2) et on a :

$$\begin{aligned}
& \sin\varphi \times (1) - \cos\varphi \times (2) \\
= \; & \sin\varphi[r^m \cos m\varphi + a_1 r^{m-1}\cos(m-1)\varphi + a_2 r^{m-2}\cos(m-2)\varphi + ... \\
& + a_{m-2} r^2 \cos 2\varphi + a_{m-1} r\cos\varphi + a_m] \\
& - \cos\varphi[r^m \sin m\varphi + a_1 r^{m-1}\sin(m-1)\varphi + a_2 r^{m-2}\sin(m-2)\varphi + ... \\
& + a_{m-2} r^2 \sin 2\varphi + a_{m-1} r\sin\varphi] \\
= \; & r^m(\sin\varphi\cos m\varphi - \sin(m)\varphi\cos\varphi) \\
& + r^{m-1} a_1(\sin\varphi\cos(m-1)\varphi - \sin(m-1)\varphi\cos\varphi) \\
& + ... \\
& + r a_{m-1}(\sin\varphi\cos\varphi - \sin\varphi\cos\varphi) \\
& + \sin\varphi a_m \\
= \; & \sin\varphi[r^m \cos m\varphi + a_1 r^{m-1}\cos(m-1)\varphi + a_2 r^{m-2}\cos(m-2)\varphi + ... \\
& + a_{m-2} r^2 \cos 2\varphi + a_{m-1} r\cos\varphi + a_m] \\
& - \cos\varphi[r^m \sin m\varphi + a_1 r^{m-1}\sin(m-1)\varphi + a_2 r^{m-2}\sin(m-2)\varphi + ... \\
& + a_{m-2} r^2 \sin 2\varphi + a_{m-1} r\sin\varphi] \\
= \; & -r^m \sin(m-1)\varphi \\
& - r^{m-1} a_1 \sin(m-2)\varphi \\
& - ... \\
& - r a_{m-1} \times 0 \\
& - \sin(-\varphi) a_m
\end{aligned}$$

Ainsi, la somme des termes de la troisième colonne est égale à $-r \times (\sin\varphi \times (1) - \cos\varphi \times (2))$, c'est-à-dire à 0.

Or, la somme des termes de la première colonne est égale à $\sin\varphi \cdot rX$, donc $\sin\varphi \cdot rX$ est divisible par $x^2 - 2\cos\varphi \cdot rx + r^2$.

1. 1er cas $r\sin\varphi \neq 0$ : alors d'après ce qui précède, $X$ est divisible par $x^2 - 2\cos\varphi \cdot rx + r^2$.
2. 2e cas $r\sin\varphi = 0$.
   (a) Si $r = 0$ alors avec (1) on a $a_m = 0$ et $X$ est divisible par $x = x - r\cos\varphi$.
   (b) Si $\sin\varphi = 0$ alors $\varphi \equiv 0[\pi]$ et donc $\cos\varphi = \pm 1$ et ainsi $\cos k\varphi = (\cos\varphi)^k$. on a :

$$\begin{aligned}
X \; = \; & x^m + a_1 x^{m-1} + a_2 x^{m-2} + ... \\
& + a_{m-2} x^2 + a_{m-1} x + a_m
\end{aligned}$$

$$\begin{aligned} X(r\cos\varphi) &= (\cos\varphi)^m + a_1(\cos\varphi)^{m-1} + a_2(\cos\varphi)^{m-2} + ... \\ &\quad + a_{m-2}(\cos\varphi)^2 + a_{m-1}(\cos\varphi) + a_m \\ &= \cos m\varphi + a_1\cos(m-1)\varphi + a_2\cos(m-2)\varphi + ... \\ &\quad + a_{m-2}\cos 2\varphi + a_{m-1}\cos\varphi + a_m \\ &= 0 \end{aligned}$$

d'après (1). Ainsi $X$ est divisible par $x - \cos\varphi$.

On voit donc qu'il suffit maintenant de prouver qu'il existe $r, \varphi$ qui permettent de se retrouver dans la situation du lemme 2.

Il faut remarquer que ces deux lemmes techniques permettent à Gauss d'éviter le calcul avec les nombres complexes : ces derniers sont donc remplacés par de brillants calculs trigonométriques. Nous allons maintenant aborder le cœur même de la démonstration de Gauss : ce qu'il fait ici, c'est simplement considérer $X$ comme un polynôme de la variable complexe, et de poser $x = r(\cos\varphi + i\sin\varphi)$. $U$ sera alors la partie réelle obtenue par cette substitution et $U$ la partie imaginaire. Cependant, comme nous l'avons déjà dit, et très certainement pour rendre sa démonstration plus convaincante, Gauss se refuse à faire appel aux nombres complexes. On verra que lors de sa quatrième démonstration, cinquante ans plus tard, en partie parce que ces nombres sont mieux acceptés par la communauté scientifique, Gauss n'hésitera pas à les utiliser. Ce type de précautions sera aussi pris lorsqu'il s'agira dans l'article 18 d'utiliser un cercle de rayon « infini », et il en proposera deux démonstrations, dont l'une est exempte de grandeurs infinies. Par contre, sur ce dernier point, le début de l'« arithmétisation » de l'analyse poussera certainement Gauss à ne plus utiliser ces grandeurs lors de sa démonstration de 1849.

On pose donc

$$\begin{aligned} U &= r^m\cos m\varphi + a_1 r^{m-1}\cos(m-1)\varphi + a_2 r^{m-2}\cos(m-2)\varphi + ... \\ &\quad + a_{m-2}r^2\cos 2\varphi + a_{m-1}r\cos\varphi + a_m \\ T &= r^m\sin m\varphi + a_1 r^{m-1}\sin(m-1)\varphi + a_2 r^{m-2}\sin(m-2)\varphi + ... \\ &\quad + a_{m-2}r^2\sin 2\varphi + a_{m-1}r\sin\varphi \end{aligned}$$

Gauss interprète géométriquement [35] $T$ et $U$. Dans l'espace muni des coordonnées cylindriques $(r, \varphi, z)$, on considère les surfaces d'équation $z = T$, $z = U$ et le plan d'équation $z = 0$, et on montre que

$$\{z = T\} \cap \{z = U\} \cap \{z = 0\} \neq \emptyset$$

35. Pour construire nos figures et avoir ainsi un appui visuel, nous allons prendre l'exemple choisi par Gauss, c'est à dire l'équation $X = x^4 - 2x^2 + 3x + 10$.

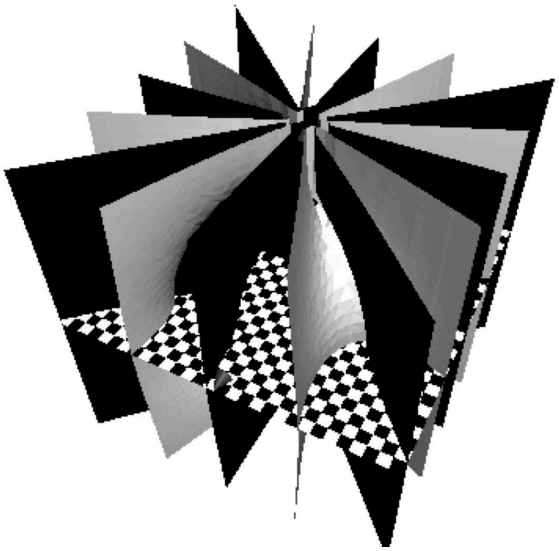


FIGURE 1 – $\{z = T\} \cap \{z = U\} \cap \{z = 0\} \neq \emptyset$

Gauss affirme que les courbes $T = 0$ et $U = 0$ sont algébriques. En effet, en posant $x = r\cos\varphi$ et $y = \sin\varphi$, on a [36] pour tout $n \in \mathbb{N}^*$ :

$$\cos(n\varphi) = \sum_{k=0}^{E\left[\frac{n}{2}\right]} (-1)^k \binom{n}{2k} (\cos\varphi)^{n-2k} (\sin\varphi)^{2k}$$

et

$$\sin(n\varphi) = \sin\varphi \sum_{k=0}^{E\left[\frac{n-1}{2}\right]} (-1)^k \binom{n}{2k+1} (\cos\varphi)^{n-2k-1} (\sin\varphi)^{2k}$$

donc

$$r^n \cos(n\varphi) = \sum_{k=0}^{E\left[\frac{n}{2}\right]} (-1)^k \binom{n}{2k} x^{n-2k} y^{2k}$$

et

$$r^n \sin(n\varphi) = y \sum_{k=0}^{E\left[\frac{n-1}{2}\right]} (-1)^k \binom{n}{2k+1} x^{n-2k-1} y^{2k}$$

et ainsi $T$ et $U$ sont bien composés de monômes de la forme $\alpha x^r y^t$, et $T = 0$ et $U = 0$ sont bien des courbes algébriques.

On est maintenant ramené à trouver l'intersection des deux courbes algébriques $\{T = 0\} \cap \{U = 0\}$.

36. La formule qui donne $\cos n\varphi$ en fonction de $\cos\varphi$ est obtenue à partir des formules de Moivre. Le polynôme $T_n$ tel que $\cos n\varphi = T_n(\cos\varphi)$ est connu aujourd'hui sous le nom de polynôme de Tchebichev de première espèce, en l'honneur de ce mathématicien contemporain de Gauss.

Gauss remarque que chaque terme de $T$ contient un $y$ : la droite d'équation $y = 0$ est donc une branche de la courbe $T = 0$.

**Théorème 1** $\exists R > 0$ tel que $T$ et $U$ s'annulent exactement en $2m$ points du cercle de centre l'origine du repère et de rayon $R$. De plus, chaque point tel que $U = 0$ est compris entre deux points où $T = 0$ et inversement.

**Démonstration du théorème 1 (première version)**

Pour $r$ « très grand »,

$$T \sim r^m \sin m\varphi$$

Or, $\sin m\varphi = 0$ correspond à $m$ droites $(t_k)$ du plan passant par l'origine et formant un angle $\dfrac{k\pi}{m}[\pi]$ avec l'axe des abscisses pour $k \in [\![0; m-1]\!]$. La courbe algébrique $\{T = 0\}$ aura donc $2m$ branches infinies correspondant à ces $m$ droites.

De même, $\{U = 0\}$ aura $2m$ branches infinies ayant pour asymptotes les $m$ droites $u_k$ du plan passant par l'origine et formant un angle $\dfrac{(2k+1)\pi}{2m}[\pi]$ avec l'axe des abscisses pour $k \in [\![0; m-1]\!]$. Ainsi, pour $r$ « infini », on se retrouve dans la situation suivante :

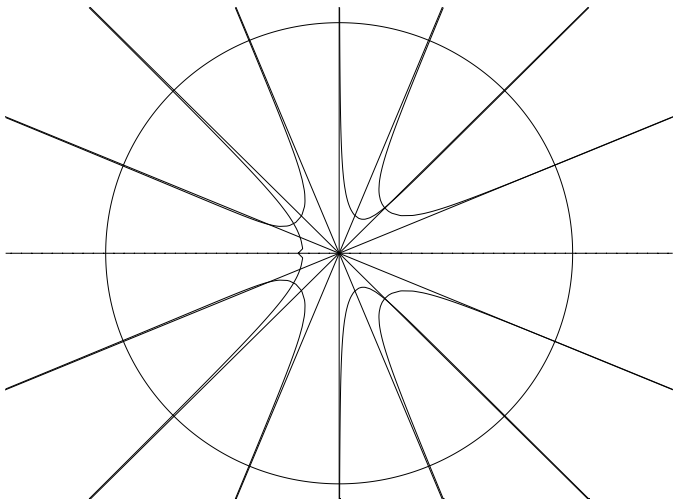


FIGURE 2 – $\{T = 0\}$ et $\{U = 0\}$ pour $r$ « infini ».

Et on est bien sur ce cercle dans la situation du théorème 1

Mais Gauss ne se satisfait pas de cette démonstration intuitive qui fait appel aux infinis. Il en propose donc une autre :

**Démonstration du théorème 1 (deuxième version)**

**Existence** Soit $R > \sup\{(a_1 + ... + a_m); 1\}$. Notons $\mathcal{C}$ le cercle de centre l'origine du repère et de rayon $R$. Pour $k \in [\![1; 4m]\!]$, on notera $(2k-1)$ le point de $\mathcal{C}$ dont les coordonnées polaires sont $(R, \frac{2k-1}{m}\frac{\pi}{4}$.

Entre les points (1) et (3), on a

$$\frac{\pi}{4} \leq m\varphi \leq \frac{3\pi}{4}$$

et alors

$$\frac{1}{\sqrt{2}} \leq \sin m\varphi \leq 1$$

donc

$$\begin{aligned} T &= R^{m-1}\left(R\sin m\varphi + a_1\sin(m-1)\varphi + ... + \frac{a_{m-1}}{R^{m-2}}\sin(\varphi)\right) \\ &\geq R^{m-1}\left(R\frac{1}{\sqrt{2}} + a_1\sin(m-1)\varphi + ... + \frac{a_{m-1}}{R^{m-2}}\sin(\varphi)\right) \\ &> R^{m-1}\left(S + a_1\sin(m-1)\varphi + ... + \frac{a_{m-1}}{R^{m-2}}\sin(\varphi)\right) \\ &> 0 \end{aligned}$$

Ainsi $T > 0$ entre les points (1) et (3).

De même, entre les points (8k+1) et (8k+3), on a

$$\frac{\pi}{4} + 2k\pi \leq m\varphi \leq \frac{3\pi}{4} + 2k\pi$$

c'est-à-dire

$$\frac{(8k+1)\pi}{4} \leq m\varphi \leq \frac{(8k+3)\pi}{4}$$

$$\frac{1}{\sqrt{2}} \leq \sin m\varphi \leq 1$$

et donc

$$\boxed{T > 0 \text{ entre les points } (8k+1) \text{ et } (8k+3) \text{ pour } k \in [\![0; m-1]\!]}.$$

Par un raisonnement semblable, on montre[37] que

$$\boxed{T < 0 \text{ entre les points } (8k+7) \text{ et } (8k+9) \text{ pour } k \in [\![0; m-1]\!]}.$$

37. On conviendra que le point $(8m+1)$ est le point (1).

Donc, la fonction $T$ étant continue en la variable $\varphi$, le théorème des valeurs intermédiaires nous permet d'affirmer que pour tout $k \in [\![0; m-1]\!]$ il existe un point compris strictement entre $(8k+3)$ et $(8k+5)$ et un autre compris strictement entre $(8k+7)$ et $(8k+9)$ tels que $T = 0$ en ces points. On obtient ainsi $2m$ points tels que $T = 0$.

**Unicité** Si $T$ s'annule, ce ne peut être qu'entre deux points du type $(8k+3)$, $(8k+5)$ ou entre deux points du type $(8k+7)$, $(8k+9)$. En effet, on a montré que partout ailleurs $T > 0$ ou $T < 0$.

On a [38]

$$\frac{dT}{d\varphi} = mR^{m-1}\left(R\cos m\varphi + \frac{m-1}{m}a_1\cos(m-1)\varphi + ... + \frac{a_{m-1}}{mR^{m-2}}\cos\varphi\right)$$

Entre les points $(8k+3)$ et $(8k+5)$, on a

$$\frac{(8k+3)\pi}{4} \leq m\varphi \leq \frac{(8k+5)\pi}{4}$$

$$-1 \leq \cos m\varphi \leq -\frac{1}{\sqrt{2}}$$

et donc

$$\begin{aligned} -\frac{dT}{d\varphi} &> mR^{m-1}\left(\frac{1}{\sqrt{2}}S\sqrt{2} - \frac{m-1}{m}a_1\cos(m-1)\varphi - ... - \frac{a_{m-1}}{mR^{m-2}}\cos\varphi\right) \\ &\geq mR^{m-1}(S-S) = 0 \end{aligned}$$

Finalement, entre ces points, on a $\frac{dT}{d\varphi} < 0$. Donc, d'après le théorème des valeurs intermédiaires, $T$ étant continue en la variable $\varphi$, la fonction $T$ s'annule au plus une fois entre les points $(8k+3)$ et $(8k+5)$.

On montre exactement de la même façon qu'entre les points $(8k+7)$ et $(8k+9)$ on a $\dfrac{dT}{d\varphi} > 0$ et que $T$ ne s'y annule qu'au plus une fois. Or on a vu que $T$ s'annulait en $2m$ points de $\mathcal{C}$. On conclut que $T$ s'annule *exactement* en $2m$ points de $\mathcal{C}$.

**Conclusion** Un raisonnement similaire sur $U$ nous montre que $U$ s'annule exactement en $2m$ points de $\mathcal{C}$, et que ces points sont situés entre les points $(8k+1)$ et $(8k+3)$, et entre les points $(8k+5)$ et $(8k+7)$. On a donc bien alternance entre les points où $T = 0$ et ceux où $U = 0$.

38. Il semble y avoir une erreur dans le calcul de la dérivée de Gauss, tout au moins dans l'édition [Gauss, 1799] et la traduction allemande de E.Netto [Netto, 1876] : il indique $R^{m-2}$ à la place de $R^{m-1}$.

**Remarque** $T = 0$ est une courbe algébrique de degré $m$ et $\mathcal{C}$ est un cercle, donc une courbe algébrique de degré $m$. On sait donc, et Gauss le savait aussi, étant donné qu'il en fait lui-même la remarque, que le nombre maximum d'intersections entre ces deux courbes est $2m$. $R$ a justement été choisi suffisamment grand pour qu'il y ait exactement $2m$ intersections. Mais une fois de plus, Gauss semble privilégier le point de vue géométrique.

**Théorème 2** À l'intérieur de $\mathcal{C}$, $\{T = 0\} \cap \{U = 0\} \neq \emptyset$

**Démonstration du théorème 2**

On désigne par (0) l'intersection [39] de $\mathcal{C}$ avec la partie négative de l'axe des abscisses. Ce point appartient à la courbe $T = 0$ comme on l'a remarqué précédemment. on numérote ensuite, en tournant dans le sens des aiguilles d'une montre, les points d'intersection de $\mathcal{C}$ et des courbes $T = 0$ et $U = 0$. D'après le théorème 2, les zéros de $T$ et de $U$ alternent, et on a donc :

— les points $(k)$ avec $k$ pair qui sont les zéros de $T$.

— les points $(k)$ avec $k$ impair qui sont les zéros de $U$.

Ainsi, on est dans la situation représentée sur la figure 3.

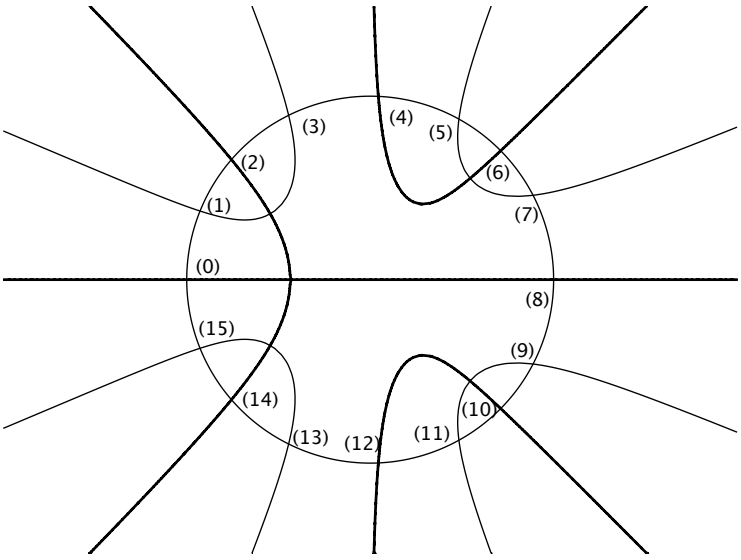



FIGURE 3 – $\{T = 0\} \cap \{U = 0\}$. ($T$ est en gras)

L'argumentation de Gauss est la suivante : La courbe $T = 0$ (de même que la courbe $U = 0$) est algébrique, et donc chaque branche de cette courbe qui entre dans $\mathcal{C}$ doit en

39. On fera attention à ne pas confondre les points $(k)$ de cette démonstration et ceux de la précédente.

ressortir, étant donné que « ou bien [elle] revient sur elle-même, ou bien [elle] s'étend de part et d'autre vers l'infini » [40]. Il s'agit là en fait d'une propriété des courbes algébriques qui n'est pas du tout évidente à démontrer. Nous avons vu [41] qu'Alexender Ostrowski en a donné une démonstration dans [Ostrowski, 1920] plus d'un siècle plus tard, et que cet argument est central dans la théorie des caractéristique de Kronecker.

Une fois cette propriété acceptée, il s'ensuit que chaque branche qui entre dans $\mathcal{C}$ par un point pair sort par un point pair, et chaque branche qui entre dans le cercle par un point impair sort par un point impair. Donc les points pairs sont reliés entre eux par des branches de la courbe $T = 0$ et les points impairs par des branches de la courbe $U = 0$.

**Supposons que** $\{T = 0\} \cap \{U = 0\} = \emptyset$ , alors le point (0) est relié au point $(2m)$, donc le point (1) doit être relié à un autre point impair du demi-plan supérieur, c'est-à-dire à un point $(n_1)$ tel que $n_1 < 2m$.

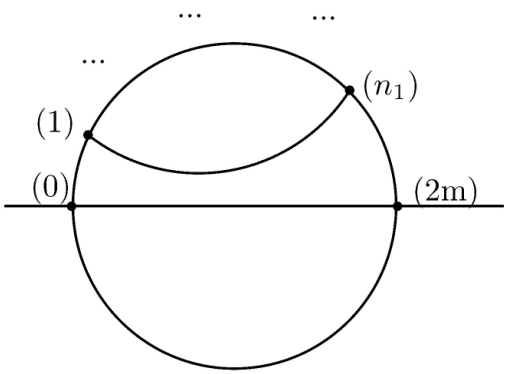



FIGURE 4 –

Le point (2) doit être relié à un autre point pair inférieur à $(n_1)$ (sinon les deux courbes s'intersectent), c'est-à-dire à un point $(n_2)$ tel que $n_2 < n_1$.
On remarque qu'il y a un nombre impair de points entre deux points pairs ou entre deux points impairs et qu'en continuant de procéder ainsi on obtient une suite d'entiers impairs strictement décroissante : $n_{k+1} - n_k - 1$ qui correspond aux nombres de points restants entre les points $(n_k)$ et $(n_{k+1})$.

Chacun des points devant être relié à un autre, le procédé précédant arrivera à une impasse lorsqu'il ne restera plus qu'un seul point : de ce dernier point $(h + 1)$ il partira nécessairement une branche qui intersectera la branche passant par les points $(h)$ et $(h+2)$.

40. Voir [Gauss, 1799] art. 21.
41. Voir p. 378.

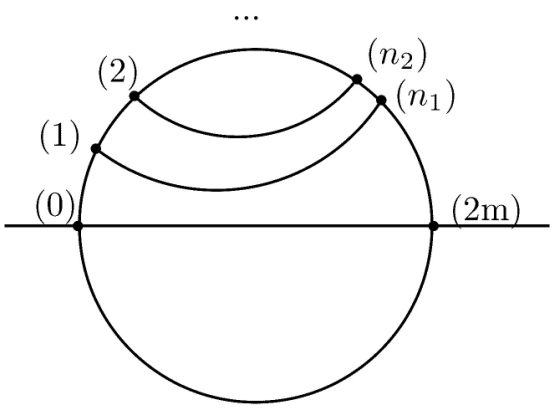



FIGURE 5 –

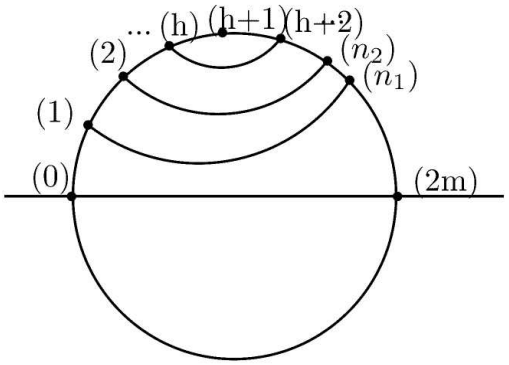



FIGURE 6 –

**Conclusion** Les courbes $T = 0$ et $U = 0$ ayant un point d'intersection, on se trouve donc dans les hypothèses du lemme 2, et le polynôme $X$ est donc bien factorisable par un polynôme du premier ou du second degré.

### 15.2.2 La quatrième démonstration de Gauss de 1849

Gauss donne dans un mémoire intitulé *Contribution à la théorie des équations algébriques* et présenté à la Société Royale des sciences le 16 juillet 1849 une quatrième et dernière preuve du TFA. Cette dernière, la seule des quatre à être rédigée en allemand, reprend celle qu'il a publiée cinquante ans plus tôt, mais sous « une forme nouvelle et avec de nombreux ajouts » [42].

Après avoir soigneusement évité toute référence aux nombres complexes dans sa première démonstration, et s'être contenté de polynômes appartenant à $\mathbb{R}[x]$ dans les deux suivantes, Gauss démontre son théorème dans le cadre plus large des polynômes à coefficients complexes. Il n'y a aucune difficulté mathématique à étendre le théorème de $\mathbb{R}[x]$ à

42. [Gauss, 1849] p. 73 : « in einer veränderten Gestalt und mit erheblichen Zusätzen ».

$\mathbb{C}[x]$, mais comme l'écrit Gauss [43] :

> A l'heure actuelle où le concept de grandeur complexe est courante, il semble plus approprié d'utiliser cette forme et d'exprimer ainsi le théorème : toute fonction peut être décomposée en $n$ facteurs simples, où la partie constante de ces facteurs n'est pas forcément des grandeurs réelles, mais peut aussi être une valeur complexe. Par cet habillage, le théorème gagne en généralité, car la restriction à des valeurs réelles pour les coefficients $A, B, ...$ n'a plus besoin d'être supposée, et pour toutes les valeurs la même permission est accordée. [Gauss, 1849] p. 74

Gauss semble donc bien avoir attendu que la communauté scientifique de son époque soit prête pour se placer dans le cadre des nombres complexes. On obtient donc les deux formulations suivants du TFA :

1. dans les trois premières démonstrations : *toute fonction algébrique rationnelle entière à une variable peut se décomposer en facteurs réels de premier ou de second degré.*
2. dans la quatrième démonstration : *toute fonction peut être décomposée en n facteurs simples, où la partie constante de ces facteurs n'est pas forcément des grandeurs réelles, mais peut aussi être une valeur complexe.*

**Quelques remarques** Gauss commence son traité par quelques remarques préliminaires

1. Pour démontrer que tout polynôme se décompose en produit de facteurs réels du premier ou du second degré, il suffit de montrer que tout polynôme possède *un* facteur du premier ou du second degré : c'est, affirme-t-il, ce que lui-même et Cauchy ont fait dans des preuves déjà existantes du théorème.
2. Une preuve sera plus claire, plus achevée si elle montre directement la présence de tous les facteurs. C'est ce qu'il fait ici.
3. La notion de nombre imaginaire étant en ce milieu du XIX$^{e}$ siècle bien connue, il montre le théorème pour tout $P \in \mathbb{C}[X]$.
4. $P(a) = 0$ ssi $(x - a)/P$

**Corps de la démonstration**

**Théorème 1** Soit $X = x^n + A_1x^{n-1} + ...A_{n-1}x + A_n \in \mathbb{C}[X]$. Alors $X$ peut se décomposer en produit de $n$ facteurs simples

Pour démontrer ce théorème, on commence par démontrer la propriété suivante :

43. Gegenwärtig, wo der Begriff der complexen Grössen jedermann geläufig ist, scheint es angemessener, jene Form fahren zu lassen und den Satz so auszusprechen, dass jene Function sich in $n$ einfache Factoren zerlegen lasse, wo dann die constanten Theile dieser Factoren nicht eben reelle Grössen zu sein brauchen, sondern für dieselben auch jede complexen Werthe zulässig sein müssen. Bei dieser Einkleidung gewinnt selbst der Satz noch an Allgemeinheit, weil dann die Beschränkung auf reelle Werthe auch bei den Coefficienten A, B u.s.w. nicht vorausgesetzt zu werden braucht, vielmehr jedwede Werthe für dieselben zulässig bleiben.

**Propriété 1** $\exists R > 0$ tel que $\forall |x| > R$ on a $X \neq 0$.

Ce qui termine la démonstration. On pose

$$\begin{aligned} x &= r(\cos\rho + i\sin\rho) \\ A_1 &= a_1(\cos\alpha_1 + i\sin\alpha_1) \\ ... &= ... \\ A_n &= a_n(\cos\alpha_n + i\sin\alpha_n) \end{aligned}$$

**Lemme 1** L'équation $r^n - \sqrt{2}(a_1 r^{n-1} + ... + a_n) = 0$ a une unique solution positive.

Pour démontrer cela, Gauss cite un théorème qu'il attribue à Harriot. Ce théorème, tout au moins la partie concernant les racines positives, est présenté dans son traité d'algèbre *Artis analyticae praxis ad aequationis algebraicas resolvendas* paru en 1631 [44]. Il s'agit là en quelque sorte de la règle des signes de Descartes. Gauss a proposé dans le *Journal de Crelle* de 1828 une démonstration rigoureuse de ce théorème qui affirme entre autres que le nombre de racines positives d'une équation $P(x) = 0$ où $P \in \mathbb{R}[X]$ est au plus égal aux nombre de changements de signes dans $P$ [45]. De plus, le théorème démontré par Gauss affirme aussi que ce nombre a la même parité que le nombre de changement de signes.

Ici, il y a un unique changement de signe (de + à -). Comme $r^n - \sqrt{2}(a_1 r^{n-1} + ... + a_n) \in \mathbb{R}[X]$ le nombre de racines positives de notre équation est un nombre impair inférieur ou égal à 1. Ce nombre est donc 1, et le lemme est démontré. On peut d'ailleurs en donner assez facilement une preuve directe [46].

Appelons $R$ cette racine.

**Lemme 2** $\forall r > R$, $r^n - \sqrt{2}(a_1 r^{n-1} + ... + a_n) > 0$

Gauss ne fait que l'affirmer, mais donnons-en une démonstration.

On peut supposer $a_n \neq 0$. En effet, si $a_n = 0$, on peut factoriser l'expression par $r^{n-i}$, où $a_i$ est le coefficient non nul ayant la plus petite puissance, et se ramener ainsi à une équation dont le terme constant est non nul.
Si tous les $a_i$ sont nuls, $X = r^n$ est déjà factorisé en produit de $n$ facteurs simples, et le théorème 1 est démontré.

44. Voir [Gagneux, 1989].
45. Une fois bien sûr que $P$ ait été réduit et ordonné.
46. Voir [Uspensky, 1948].

Pour $r = 0$, $r^n - \sqrt{2}(a_1 r^{n-1} + ... + a_n) < 0$. Comme il n'y a qu'une racine positive $R$, on a donc le tableau de signe suivant :

| $r$ | $0$ | | $R$ | | $+\infty$ |
|---|---|---|---|---|---|
| $r^n - \sqrt{2}(a_1 r^{n-1} + ... + a_n)$ | $-$ | | $0$ | $+$ | |

**Corollaire 1** $\forall r > R$, on a

$$r^n - \sqrt{2}((n-1)a_1 r^{n-1} + ... + a_{n-1} r) > 0$$

**Démonstration** Cette démonstration n'est pas donnée par Gauss.

$\forall r > R > 0$, on a

$$\begin{aligned} & r^n \geq r^n - \sqrt{2}(a_1 r^{n-1}) \\ \geq \; & r^n - \sqrt{2}(a_1 r^{n-1} + a_2 r^{n-2}) \\ \geq \; & ... \\ \geq \; & r^n - \sqrt{2}(a_1 r^{n-1} + ... + a_{n-1} r + a_n) > 0 \end{aligned}$$

Donc

$$\begin{aligned} r^n & > 0 \\ r^n - \sqrt{2}(a_1 r^{n-1}) & > 0 \\ r^n - \sqrt{2}(a_1 r^{n-1} + a_2 r^{n-2}) & > 0 \\ ... & \\ r^n - \sqrt{2}(a_1 r^{n-1} + ... + a_{n-1} r) & > 0 \end{aligned}$$

Et en sommant on obtient :

$$r^n - \sqrt{2}((n-1)a_1 r^{n-1} + ... + a_{n-1} r) > 0$$

**Démonstration de la propriété 1**

On pose $X = T + iU$ et on prend $x$ tel que $|x| > R$
On pourra dans la suite considérer $T$ et $U$ comme des fonctions des deux variables $r, \rho$ et on pourra noter $T(r, \rho)$ ou $U(r, \rho)$. Prendre $|x| > R$, c'est en fait prendre $r > R$.

On a alors :

$$\begin{aligned}
X &= x^n + A_1 x^{n-1} + ... + A_n \\
&= r^n(\cos n\rho + i\sin n\rho) + A_1 r^{n-1}(\cos(n-1)\rho + i\sin(n-1)\rho) + ... + A_n \\
&= r^n(\cos n\rho + i\sin n\rho) \\
&\quad +a_1(\cos\alpha_1 + i\sin\alpha_1)r^{n-1}(\cos(n-1)\rho + i\sin(n-1)\rho) \\
&\quad +... \\
&\quad +a_n(\cos\alpha_n + i\sin\alpha_n) \\
&= [a_1(\cos\alpha_1\cos(n-1)\rho - \sin\alpha_1\sin(n-1)\rho)r^{n-1} \\
&\quad +ia_1(\sin\alpha_1\cos(n-1)\rho + \sin(n-1)\rho\cos\alpha_1)r^{n-1}] \\
&\quad +[...] \\
&\quad +[a_n\cos\alpha_n + ia_n\sin\alpha_n] \\
&= \left[r^n\cos n\rho + a_1\cos((n-1)\rho+\alpha_1)r^{n-1} + a_2\cos((n-2)\rho+\alpha_2)r^{n-2}\right. \\
&\quad \left.+... + a_n\cos\alpha_n\right] \\
&\quad +i\left[r^n\sin n\rho + a_1\sin((n-1)\rho+\alpha_1)r^{n-1} + a_2\sin((n-2)\rho+\alpha_2)r^{n-2}\right. \\
&\quad \left.+... + a_n\sin\alpha_n\right]
\end{aligned}$$

Ainsi on a par identification :

$$T = r^n\cos n\rho + a_1\cos((n-1)\rho+\alpha_1)r^{n-1} + a_2\cos((n-2)\rho+\alpha_2)r^{n-2} + ... + a_n\cos\alpha_n$$

et

$$U = r^n\sin n\rho + a_1\sin((n-1)\rho+\alpha_1)r^{n-1} + a_2\sin((n-2)\rho+\alpha_2)r^{n-2} + ... + a_n\sin\alpha_n$$

Écrivons $T$ sous la forme

$$\begin{aligned}
T &= \sqrt{\frac{1}{2}}r^n - a_1 r^{n-1} - a_2 r^{n-2} - ... - a_n \\
&\quad +\left[\cos n\rho - \sqrt{\frac{1}{2}}\right]r^n \\
&\quad +\left[1+\cos((n-1)\rho+\alpha_1)\right]a_1 r^{n-1} \\
&\quad +\left[1+\cos((n-2)\rho+\alpha_2)\right]a_2 r^{n-2} \\
&\quad +... \\
&\quad +\left[1+\cos\alpha_n\right]a_n
\end{aligned}$$

D'après le lemme 2,

$$\sqrt{\frac{1}{2}}r^n - (a_1 r^{n-1} + ... + a_n) > 0$$

De plus,

$$\begin{aligned}
&+ [1 + \cos((n-1)\rho + \alpha_1)]\, a_1 r^{n-1} \geq 0 \\
&+ [1 + \cos((n-2)\rho + \alpha_2)]\, a_2 r^{n-2} \geq 0 \\
&+ ... \\
&+ [1 + \cos \alpha_n]\, a_n \geq 0
\end{aligned}$$

Donc

$$\boxed{\text{Si } \sqrt{\frac{1}{2}} \leq \cos n\rho \leq 1 \text{ alors } T > 0}$$

Écrivons maintenant $-T$ sous la forme

$$\begin{aligned}
-T \;=\; & \sqrt{\frac{1}{2}}r^n - a_1 r^{n-1} - a_2 r^{n-2} - ... - a_n \\
& + \left[ -\cos n\rho - \sqrt{\frac{1}{2}} \right] r^n \\
& + [1 - \cos((n-1)\rho + \alpha_1)]\, a_1 r^{n-1} \\
& + [1 - \cos((n-2)\rho + \alpha_2)]\, a_2 r^{n-2} \\
& + ... \\
& + [1 - \cos \alpha_n]\, a_n
\end{aligned}$$

et on obtient :

$$\boxed{\text{Si } -1 \leq \cos n\rho \leq -\sqrt{\frac{1}{2}} \text{ alors } -T > 0 \text{ et donc } T < 0}$$

On montre de même que

$$\boxed{\text{Si } \sqrt{\frac{1}{2}} \leq \sin n\rho \leq 1 \text{ alors } U > 0}$$

et que

$$\boxed{\text{Si } -1 \leq \sin n\rho \leq -\sqrt{\frac{1}{2}} \text{ alors } U < 0}$$

Or, quelque soit $\rho$, on est toujours dans l'un des quatre cas précédents. On a donc toujours $U \neq 0$ ou $T \neq 0$. Ce qui démontre la propriété 1.

On peut d'ailleurs présenter cela par un tableau

| $n\rho$ | | $-\frac{\pi}{4}$ | | $\frac{\pi}{4}$ | | $\frac{3\pi}{4}$ | | $\frac{5\pi}{4}$ | | $\frac{7\pi}{4}$ | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $T$ | | \| | + | \| | | \| | − | \| | | \| | ... |
| $U$ | − | \| | | \| | + | \| | | \| | − | \| | ... |

Et voici une représentation du cercle trigonométrique séparé en quatre quartiers

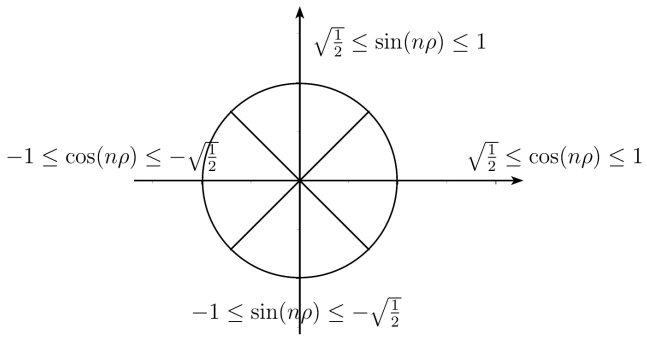



FIGURE 7 –

**Propriété 2** Soit $r > R$ fixé et $\omega = \dfrac{45^\circ}{n} = \dfrac{\pi}{4n}$.
Alors $T = 0$ a $2n$ solutions dans les $2n$ intervalles de la forme$[(4k+1)\omega; (4k+3)\omega]$
De même, $U = 0$ a $2n$ solutions dans les $2n$ intervalles de la forme$[(4k+3)\omega; (4k+5)\omega]$

**Démonstration** Entre $-\omega$ et $(8n-1)\omega$ il y a $2\pi$. On partage l'intervalle $[-\omega; (8n-1)\omega]$ en $4n$ parties égales de longueur $2\omega$.

De plus :

$$\begin{aligned}\frac{dT}{d\rho} &= -nr^n \sin n\rho \quad -(n-1)a_1 r^{n-1}\sin((n-1)\rho+\alpha_1)\\ &\quad -(n-1)a_1 r^{n-1}\sin((n-1)\rho+\alpha_1)\\ &\quad -(n-2)a_2 r^{n-2}\sin((n-2)\rho+\alpha_2)\\ &\quad -...\\ &\quad -a_{n-1}\sin(\rho+\alpha_{n-1})\end{aligned}$$

$$\begin{aligned} = -n\sqrt{\frac{1}{2}}\sin n\rho \quad & +(n-1)a_1 r^{n-1} + ...a_{n-1} \\ & -n(\sin n\rho - \sqrt{\frac{1}{2}})r^n \\ & -(n-1)(1+\sin((n-1)\rho+\alpha_1))a_1 r^{n-1} \\ & -(n-2)(1+\sin((n-2)\rho+\alpha_2))a_2 r^{n-2} \\ & -... \\ & -a_{n-1}(1+\sin(\rho+\alpha_1)) \end{aligned}$$

Donc, comme précédemment, on a

$$\boxed{\text{Si } \sqrt{\frac{1}{2}} \leq \sin n\rho \leq 1 \text{ alors } \frac{dT}{d\rho} > 0}$$

De même

$$\boxed{\text{Si } -1 \leq \sin n\rho \leq -\sqrt{\frac{1}{2}} \text{ alors } \frac{dT}{d\rho} < 0}$$

Par un raisonnement similaire sur $U$, on obtient

$$\boxed{\text{Si } \sqrt{\frac{1}{2}} \leq \cos n\rho \leq 1 \text{ alors } \frac{dU}{d\rho} > 0}$$

et

$$\boxed{\text{Si } -1 \leq \cos n\rho \leq -\sqrt{\frac{1}{2}} \text{ alors } \frac{dU}{d\rho} < 0}$$

**Exemple 1** Si $\rho \in [-\omega;\omega]$, alors $n\rho \in [-\frac{\pi}{4};\frac{\pi}{4}]$
donc

$$\sqrt{\frac{1}{2}} \leq \cos n\rho \leq 1$$

ainsi $\frac{dU}{d\rho} > 0$ et $U$ est donc strictement croissante sur cet intervalle.

Par ailleurs, si $-\frac{3\pi}{4} \leq n\rho \leq -\frac{\pi}{4}$ alors $-1 \leq \sin n\rho \leq -\sqrt{\frac{1}{2}}$ et donc d'après ce qui précède $U < 0$. En particulier $U(r, -\frac{\pi}{4n}) < 0$.

si $\frac{\pi}{4} \leq n\rho \leq \frac{3\pi}{4}$ alors $\sqrt{\frac{1}{2}} \leq \sin n\rho \leq 1$ et donc $U > 0$. En particulier $U(r, \frac{\pi}{4n}) > 0$.

D'après le théorème des valeurs intermédiaires, il existe donc une unique valeur $\rho_1$ sur l'intervalle $[-\omega, \omega]$ telle que $U(r, \rho_1) = 0$.

**Exemple 2** Si $\rho \in [\omega; 3\omega]$, alors $n\rho \in [\frac{\pi}{4}; \frac{3\pi}{4}]$
donc

$$\sqrt{\frac{1}{2}} \leq \sin n\rho \leq 1$$

ainsi $\frac{dT}{d\rho} < 0$ et $T$ est donc strictement décroissante sur cet intervalle.

Or $T(r, \omega) > 0$ et $T(r, 3\omega) < 0$.

D'après le théorème des valeurs intermédiaires, il existe donc une unique valeur $\rho_1'$ sur l'intervalle $[\omega, 3\omega]$ telle que $T(r, \rho_1') = 0$.

Gauss poursuit ses explications pour les deux intervalles qui suivent

On a ainsi

| $\rho$ | | $-\omega$ | | $\omega$ | | $3\omega$ | | $5\omega$ | | $7\omega$ | | $9\omega$ | | $11\omega$ | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $T$ | | | $+$ | | $0$ | | $-$ | | $0$ | | $+$ | | $0$ | | ... |
| $U$ | $-$ | | $0$ | | $+$ | | $0$ | | $-$ | | $0$ | | $+$ | ... | |

| $\rho$ | ... | | $(4k+1)\omega$ | | $(4k+3)\omega$ | | $(4k+5)$ | | $(4k+7)\omega$ | | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $T$ | ... | $+$ | | $0$ | | $-$ | | $0$ | | $+$ | ... |
| $U$ | ... | | | $+$ | | $0$ | | $-$ | | $0$ | ... |

*Remarque* : dans la seconde ligne, l'ordre des signes est arbitraire et peut changer en fonction de $k$.

Ce qui conclut la démonstration de la propriété 2

**Théorème 2** $\exists r < R$ et $\rho \in [0; 2\pi]$ tels que $U(r, \rho) = T(r, \rho) = 0$

**Démonstration** On se place dans le plan complexe. Soit $K$ le cercle de centre $O$ et de rayon [47] $R' > R$.

D'après la propriété 2, on peut trouver $2n$ points : $(1), (2), ..., (i), ...(2n)$ sur $K$ dont l'argument $\rho_i$ est compris dans l'intervalle $](4i-3)\omega; (4i-1)w[$ et tels que pour tout $i$, $T(R', \rho_i) = 0$.

47. Gauss explique en note que l'on peut en fait, sous certaines conditions, utiliser $R' = R$, mais que c'est inutile ici.

En outre, gardons à l'esprit que l'on a vu que

$$U(R', \rho_i) > 0 \text{ pour } i \text{ impair}$$

et

$$U(R', \rho_i) < 0 \text{ pour } i \text{ pair}$$

Comme $T$ est continue, l'ensemble des points du plan tels que $T > 0$ (de même que l'ensemble des points tels que $T < 0$) forme un certain nombre de surfaces du plan séparées par des lignes continues représentant les points tels que $T = 0$.

On prend comme exemple le graphique qui correspond à l'équation

$$r^5 \cos 5\rho - 3r^4 \cos 4\rho + r^3 \cos 3\rho + 4r \cos \rho = 0$$

et qui correspond à l'équation de degré 5

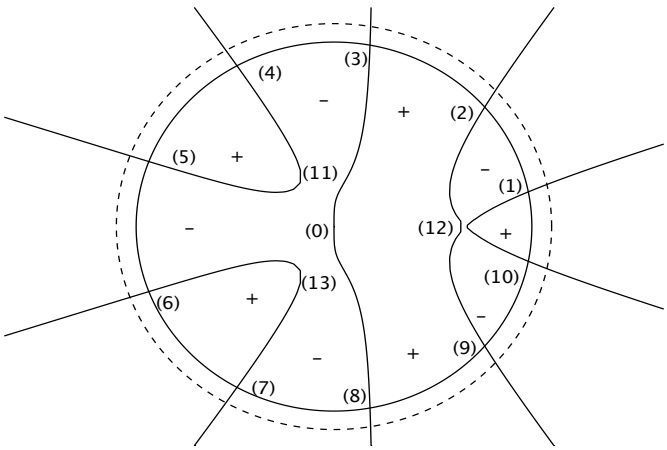
$$x^5 - 3x^4 + x^3 + 4x = 0$$

c'est-à-dire

$$x(x-2)^2(x^2 + x + 1) = 0$$

Le graphique obtenu est très proche de l'exemple fourni par Gauss, et il faut remarquer qu'il permet d'illustrer à la fois le cas où il y a des racines multiples et le cas où il y a des racines complexes.

On se retrouve donc dans la situation suivante à **l'extérieur de** $K$.




En effet, pour chaque arc de cercle $\widehat{(i)(i+1)}$, le signe de $T$ est constant, et $T = 0$ en $(i)$ et $(i+1)$. D'après ce qui précède, la situation doit rester la même pour

$R'' > R'$ (en pointillés sur la figure) et donc $T$ vaut 0 en $(i)'$ et en $(i+1)'$, et son signe reste constant (et identique à celui qu'il prend sur $\widehat{(i)(i+1)}$) sur l'arc de cercle $\widehat{(i)'(i+1)'}$).

On voit ainsi s'étendre $n$ surfaces à l'extérieur du cercle $K$ où $T > 0$ qui alternent avec $n$ surfaces où $T < 0$.

À l'**intérieur de** $K$ :

Lorsque l'on parcourt la courbe $T = 0$ en décidant de « laisser à droite » la surface $T > 0$, cette surface doit *toujours* être à droite. Ainsi, si l'on entre dans dans le cercle $K$ par un point impair, on devra ressortir par un point pair.

Ces $2n$ surfaces, par continuité, s'étendent à l'intérieur du cercle. Gauss distingue 3 cas différents :

1. La surface entrant dans $K$ se referme sur elle-même, comme dans 5.12.4/7.13.6/1.11.10.
2. La surface traverse $K$ comme dans 2.3.8.9.11.
3. La surface se sépare en 2 ou plus, comme dans 3.4.12.5.6.13.7.8.0.

Gauss affirme ensuite que l'on peut envisager d'autres configurations, mais qu'elles ne se produiront jamais. Ce qui compte ici, c'est que l'ensemble des points tels que $T = 0$ forme $n$ lignes continues séparant les surfaces où $T$ est négatif des surfaces où $T$ est positif. Ainsi, d'après la remarque précédente, ces lignes entrent par un point pair et sortent par un point impair. Or, $U > 0$ aux points impairs et $U < 0$ aux points pairs. Par continuité, il y a un point entre l'entrée et la sortie de $K$ où $U = 0$. En ce point, $U = T = 0$ et donc $X = 0$. Il y a $n$ lignes, il y a donc $n$ racines distinctes ou égales.

Gauss termine en démontrant la propriété suivante

**Propriété 3** Si $m$ lignes correspondant à $T = 0$ se rencontrent au point $P$ représentant la racine $p$ de $X$, alors $(x-p)^m$ est un facteur de $X$

### 15.2.3 La reprise par Kronecker

Nous allons maintenant voir comment Kronecker présente dans ses *Vorlesungen* cette démonstration du théorème fondamental de l'algèbre. Il reprend explicitement cette démonstration de Gauss [48] :

> Nous voulons donner encore une seconde preuve de notre théorème qui se déduit de la géométrie, et plus tard une troisième preuve algébrique ; les fondements des trois

48. Wir wollen noch einen zweiten Beweis unseres Satzes geben, der aus der Geometrie hergeleitet wird, später noch einen Dritten algebraischen ; alle 3 Beweise sind in ihren Grundzügzn von Gauß gegeben.

> démonstrations ont été donnés par Gauss. [Kronecker, 1873a] p. 190

Et plus encore, *toutes* les démonstrations qu'il donnera de ce théorème dans ses leçons seront inspirées du travail de Gauss. Il était donc bien nécessaire de décrire ces dernières avant d'examiner celles de Kronecker pour pouvoir en juger l'originalité. C'est également pourquoi il nous semble aussi important de permettre au lecteur de consulter les textes originaux de Gauss que nous avons largement cités des troisième et quatrième démonstrations : nous en avons ainsi joint en annexe une traduction.

Examinons les grandes étapes qui constituent la démonstration de Gauss :

1. Gauss travaille toujours dans un plan repéré : en 1799 ce plan est muni de coordonnées polaires, et en 1849 de coordonnées cartésiennes.
2. Il détermine un cercle de centre l'origine du repère et de rayon suffisamment grand pour contenir l'ensemble des points d'intersection des deux courbes qu'il étudie. L'existence d'un tel rayon est l'objet de la première partie de son travail [49].
3. Il détermine à l'intérieur de ce cercle les intersections entre les deux courbes étudiées.

On a donc commencé par « géométriser » le problème, puis on le restreint à l'intérieur d'un compact, et enfin on compte des intersections. Nous allons suivre ces grandes étapes à l'intérieur du travail de Kronecker. Pour cela, nous allons plus particulièrement étudier ce qu'il fait dans ses manuscrits de 1872-73, 1878-79, 1880-81 et 1890-91.

#### 15.2.3.1 Les coefficients du polynôme

Dans la démonstration de 1849, Gauss décide de regarder un polynôme à coefficients complexes, et il affirme qu'ainsi il « gagne en généralité » [50]. Kronecker, dans le manuscrit de 1872-73 affirme exactement l'inverse [51] :

> Soit
> $$\Theta(z) = z^n + c_1 z^{n-1} + c_2 z^{n-2} + ...$$
> où les $c$ sont réels (G. suppose dans son travail de l'année 1849 aussi les $c$ complexes, pourtant cela n'est pas le plus général, car si le procédé est donné pour les $c$ réels, alors s'ensuit celui pour les imaginaires). [Kronecker, 1873a] p. 190

Dans le manuscrit de 1878-79, Kronecker continue à ne traiter que le cas des polynômes à coefficients réels, mais dans son dernier cours sur les équations algébriques il prend une fonction « $\Theta(x + iy)$ une fonction entière de degré $n$ de la variable complexe $(x + iy)$ avec des coefficients complexes ». Il n'y a en effet pas de difficultés supplémentaires dans le choix de coefficients complexes. Le problème pour Kronecker, en particulier dans ses

49. Dans le théorème 1 de la première démonstration et dans la proposition 1.
50. Voir la traduction de la preuve 4. de Gauss.
51. Sei
$$\Theta(z) = z^n + c_1 z^{n-1} + c_2 z^{n-2} + ...$$
Darin seinen die $c$ reell (G. nimmt in der Arbeit aus dem Jahre 1849 die $c$ auch als complex an, jedoch ist das nicht allgemeiner, da, wenn das Verfahren für reele $c$ geg. ist, das für imaginäre sich ergiebt).

derniers cours, est plutôt dans l'utilisation des grandeurs irrationnelles. C'est ainsi que nous verrons dans la dernière section de ce chapitre comment en 1891 il revient sur la signification du théorème fondamental et les formes qu'il doit prendre.

Revenons maintenant sur l'affirmation précédente de Kronecker du manuscrit de 1872-73 : le cas des coefficients complexes n'est pas plus général que celui des coefficients réels, car ce dernier nous permet de traiter le premier. Évidemment, l'inverse est vrai, mais on retrouve ici la volonté de Kronecker de limiter l'introduction de notions nouvelles lorsque cela est inutile. Nous reviendrons plus loin sur cette notion de *généralité* dans le cadre de la démonstration du théorème fondamental de l'algèbre.

Dans une première approche de ce théorème – approche qui se trouve dans la partie qu'il consacre à la théorie des caractéristiques – Kronecker examine donc un polynôme de la variable complexe avec des coefficients réels dans les premiers manuscrits et des coefficients complexes dans les derniers.

#### 15.2.3.2 Géométrisation du problème

Cette démonstration que Kronecker propose dans son chapitre sur la théorie des caractéristiques, et qui est issue de la première et quatrième démonstration de Gauss est de nature profondément topologique. Elle s'appuie sur une représentation graphique de la situation et sur l'intersection des deux courbes représentant la partie réelle et la partie imaginaire du polynôme dont on cherche à trouver le nombre de racines. Le fait que nous cherchions finalement le nombre d'intersections de deux courbes algébriques permet à Kronecker de traiter à part le sens qu'il donne aux racines du polynôme, sa démonstration portant bien sur des intersections.

On doit maintenant distinguer les manuscrits qui précèdent son article de 1878 sur la théorie des caractéristiques et ceux qui la suivent. En effet, dans ses leçons du semestre d'hiver de l'année 1872-73, Kronecker propose une démonstration qui reprend presque exactement les arguments de Gauss et n'utilise pas directement le travail qu'il vient de faire sur la théorie des caractéristiques. Bien sûr, les techniques employées sont très proches, et le *Fortgangsprincip* apparaît d'ailleurs, sans être nommé, dans le cœur de la preuve. Dans ce manuscrit, ce théorème est en fait utilisé par Kronecker comme transition : il donne à la suite trois démonstrations du théorème [52] en terminant par celle inspirée de la seconde preuve de Gauss qui lui permet d'introduire des fonctions dont les coefficients sont pris dans un un *domaine d'intégrité*. En 1872, il n'utilise pas encore ce terme, mais il on voit déjà dans son travail apparaître certains des objets qu'il construira dans les *Grundzüge*.

En 1878, il publie donc son article *Über die Charakteristik von Funktionen-Systemen* :

52. Celle que nous décrivons ici est la seconde.

il en utilisera les résultats dans ses leçons cette même année. Sa démonstration se fera alors dans le cadre de sa théorie des caractéristiques, et il profitera ainsi de toutes les interprétations graphiques qu'il a pu en donner. Cependant, il choisira de brosser à grands traits cette preuve, et donnera en détail celle qui suit exactement le travail de Gauss et qu'il avait exposée dans son cours précédent.

Lorsque la théorie des caractéristiques réapparaît en 1890 sous une forme suffisamment développée, il n'exposera alors plus la réplique de la preuve de Gauss, mais celle qu'il a traduite dans le langage de sa théorie.

#### 15.2.3.3 Circonscription du problème

Dans tous les cas, et comme Gauss lui-même l'avait fait, il commence par chercher à construire un cercle suffisamment grand à l'intérieur duquel se trouveront toutes les intersections cherchées.

Soit donc une fonction $\Theta$ entière de degré $n$ de la variable complexe, c'est-à-dire un polynôme à coefficients complexes :

$$\Theta(x+iy) = \varphi(x,y) + i.\psi(x,y) = (x+iy)^n - c_1(x+iy)^{n-1} + ... \pm c_n$$

On pose

$$x + iy = R.e^{vi}$$

et on obtient

$$\Theta(x+yi) = R^n.e^{nvi}\left\{1 - c_1.\frac{e^{-vi}}{R} + c_2.\frac{e^{-2vi}}{R^2} - ....\right\}$$

Par ailleurs

$$\left|c_1.\frac{e^{-vi}}{R} + c_2.\frac{e^{-2vi}}{R^2} - ....\right| \leq \frac{|c_1|}{R} + \frac{|c_2|}{R^2} + ...$$

Si maintenant on choisit $u > 0$ tel que $u > c_k$ pour $k = 1, 2, ...n$ alors

$$\begin{aligned}\left|c_1.\frac{e^{-vi}}{R} + c_2.\frac{e^{-2vi}}{R^2} - ....\right| &< u.\left(\frac{1}{R} + \frac{1}{R^2} + ... + \frac{1}{R^n}\right)\\ &\leq u.\frac{1}{R}.\frac{1-\frac{1}{R^n}}{1-\frac{1}{R}}\\ &\leq \frac{u}{R-1}\end{aligned}$$

Donc pour $R > u+1$, on a $\Theta(R.e^{vi}) = \varphi + i.\psi = R^n.e^{nvi}\alpha$, où $0 < |\alpha| < 1$, et donc $|\Theta(R.e^{vi})| > 0$. Il nous suffit donc de choisir un $R > u+1$. On retrouve ici une étape qui est parfaitement similaire à celle que nous avons vu p. 463 pour la première démonstration

de Gauss et p. 470 pour la quatrième.

**Démonstration du théorème** Kronecker démontre alors que les courbes $\varphi = 0$ et $\psi = 0$ ont toujours $n$ points d'intersection.

Tout d'abord, les relations que l'on nomme aujourd'hui de Cauchy-Riemann permettent d'écrire :

$$\varphi_1 = \psi_2 \ , \ \varphi_2 = -\psi_1$$

Soit par ailleurs un cercle de rayon $R$ suffisamment grand pour que tous les points d'intersection de $\varphi$ et $\psi$ soient à l'intérieur de ce cercle et d'équation $f(x, y) = 0$. Calculons maintenant la caractéristique du système $(f, \varphi, \psi)$.

$$\chi(f,\varphi,\psi) = \sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn \begin{vmatrix} \varphi_1 & \varphi_2 \\ \psi_1 & \psi_2 \end{vmatrix} = \sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn\ (\varphi_1^2 + \varphi_2^2) = \sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn\ 1$$

c'est-à-dire que $\chi(f, \varphi, \psi)$ donne le nombre de points d'intersection des courbes $\varphi = 0$ et $\psi = 0$ qui se trouvent dans le domaine $f < 0$, c'est-à-dire ici le nombre de tous les points d'intersection.

La caractéristique ne change que lorsque l'on a simultanément $f = \varphi = \psi = 0$. Si les coefficients $c_1, ...c_n$ s'approchent de 0, la fonction $f$ conserve sa propriété, et donc nous n'obtenons jamais $f = \varphi = \psi = 0$. Comme l'explique Kronecker dans son article de 1878 sur la théorie des caractéristiques [53] :

> Mais dans le cas $m = 1$ le développement précédent conduit directement au « théorème fondamental de la théorie des équations algébriques » et explique la véritable essence du cheminement employé par *Gauss* dans son mémoire de 1849 en montrant que pour deux systèmes de courbes représentés par une équation algébrique $f(x + yi) = 0$ la configuration concernant leurs points d'intersection à l'intérieur d'un cercle choisi suffisamment grand n'est rien d'autre que celle des systèmes de courbes qui sont issus d'une « pure [54] » équation du même degré. [Kronecker, 1878b] p. 81

‘

Ainsi, si l'on pose

$$\varphi_0 + i.\psi_0 = (x + yi)^n - 1 = 0$$

53. Im Falle m = 1 aber führt die vorstehende Entwickelung direct zu dem „Grundlehrsatz der Theorie der algebraischen Gleichungen" und legt das eigentliche Wesen der von Gauss in seiner Abhandlung von 1849 gegebenen Herleitung dar, indem sie zeigt, dass für die zwei durch irgend eine algebraische Gleichung $f(x + yi) = 0$ dargestellten Curvensysteme die Configuration in Bezug auf deren Schnittpunkte innerhalb eines hinreichend gross gewählten Kreises nicht anders ist, wie für diejenigen Curvensysteme, welche aus einer „reinen" Gleichung desselben Grades hervorgehen.

54. On rappelle qu'une *aequatio affecta* est une équation quelconque, alors qu'une *aequatio pura*, une équation "pure" est une équation dans laquelle il ne reste qu'un terme en $x$, donc du type $x^n = a$.

On a

$$\chi(f, \varphi, \psi) = \chi(f, \varphi_0, \psi_0)$$

et il est évident que $\chi(f, \varphi_0, \psi_0) = n$, donc on a bien $\chi(f, \varphi, \psi) = n$, c'est-à-dire que :

> le nombre de tous les points d'intersection des courbes $\varphi = 0$ et $\psi = 0$ est aussi grand que le degré de cette fonction entière. [Kronecker, 1891a] p. 242

la méthode qu'il utilise ici, l'invariance de la caractéristique lorsque les coefficients des fonctions étudiées évolue à l'intérieur d'une variété déterminée, est profondément topologique. C'est cette propriété que l'on retrouvera d'ailleurs chez Poincaré dans ses articles sur l'*Analysis Situs*.

## 15.3 La troisième démonstration

Nous l'avons vu, la théorie des caractéristiques a deux pendants : l'un topologico-algébrique et qui peut s'interpréter en termes d'intersections de courbes, et l'autre plus analytique qui s'exprime à l'aide d'intégrales. Dans la section précédente, nous avons vu comment Kronecker utilisait la version topologico-algébrique de sa théorie pour reprendre les première et quatrième démonstrations du TFA de Gauss. Nous allons maintenant voir que Kronecker expose aussi une démonstration du TFA qui utilise la théorie des caractéristiques sous sa forme analytique et qui s'apparente à la troisième preuve de Gauss.

### 15.3.1 La version de Gauss

Quelques semaines après avoir publié sa seconde démonstration, Gauss propose une troisième preuve du TFA [55]. Il introduit celle-ci de la façon suivante [56] :

> Après que mon commentaire précédent fut déjà imprimé, encore une fois des réflexions sur le même sujet me conduisirent à une nouvelle démonstration du théorème, démonstration qui, de même que la précédente, est purement analytique, mais repose sur des principes totalement différents et au regard de sa simplicité, lui est de loin préférable. C'est pourquoi les pages qui suivent seront consacrées à cette troisième démonstration. [Gauss, 1816b] p. 59

Si cette démonstration est pour Gauss « purement analytique » c'est qu'elle n'utilisera pas les arguments plus géométriques de la démonstration de 1799. Cependant, il ne faudrait pas croire pour autant que l'on peut la rapprocher de la seconde preuve : nous allons voir qu'elle « repose sur des principes totalement différents ».

55. La seconde est datée du 15 décembre 1815 et la troisième du 30 janvier 1816.

56. Postquam commentatio praecedens typis iam expressa esset, iteratae de eodem argumento meditationes ad novam theorematis demonstrationem perduxerunt, quae perinde quidem ac praecedens pure analytica est, sed principiis prorsus diversis innititur, et respectu simplicitatis illi longissime praeferenda videtur. Huic itaque *tertiae* demonstrationi pagellae sequentes dicatae sunto.

#### 15.3.1.1 Notation

Gauss commence en donnant les nombreuses notations dont il aura ensuite besoin. Soit donc

$$P(x) = x^n + a_1 x^{n-1} + ... + a_{n-1}x + a_n \in \mathbb{R}[x]$$

Et soit de plus $r \in \mathbb{R}$ et $\varphi \in \mathbb{R}$. On a alors les fonctions :

$$\begin{aligned}
t(r,\varphi) &= r^n \cos n\varphi + a_1 r^{n-1}\cos(n-1)\varphi + ... + a_{n-1} r\cos\varphi + a_n \\
u(r,\varphi) &= r^n \sin n\varphi + a_1 r^{n-1}\sin(n-1)\varphi + ... + a_{n-1} r\sin\varphi \\
t'(r,\varphi) &= nr^n \cos n\varphi + (n-1)a_1 r^{n-1}\cos(n-1)\varphi + ... + a_{n-1} r\cos\varphi \\
u'(r,\varphi) &= nr^n \sin n\varphi + (n-1)a_1 r^{n-1}\sin(n-1)\varphi + ... + a_{n-1} r\sin\varphi \\
t''(r,\varphi) &= n^2 r^n \cos n\varphi + (n-1)^2 a_1 r^{n-1}\cos(n-1)\varphi + ... + a_{n-1} r\cos\varphi \\
u''(r,\varphi) &= n^2 r^n \sin n\varphi + (n-1)^2 a_1 r^{n-1}\sin(n-1)\varphi + ... + a_{n-1} r\sin\varphi
\end{aligned}$$

De nouveau Gauss utilise la forme trigonométrique des polynômes et se restreint aux coefficients réels : nous sommes bien dans des conditions similaires à celles de sa première preuve.

On peut alors poser

$$y = \frac{(t^2+u^2)(tt''+uu'') + (tu'-ut')^2 - (tt'+uu')^2}{r(t^2+u^2)^2}$$

Soit enfin $R$ tel que

$$R > \sup(n|a_1|\sqrt{2}, \sqrt{n|a_2|\sqrt{2}}, \sqrt[3]{n|a_3|\sqrt{2}}, ..., \sqrt[n]{n|a_n|\sqrt{2}})$$

#### 15.3.1.2 Restriction à un intervalle

Gauss commence, comme dans les première et quatrième démonstrations, par réduire la recherche des racines à un intervalle compact. Le $R$ qu'il détermine correspond au rayon du cercle dans lequel se trouvent les intersections des courbes dans les démonstrations de 1799 et 1849. Mais ici tout le vocabulaire de la géométrie, et donc toute interprétation géométrique sont soigneusement évités, car bien que $t$ et $u$ puissent être interprétées comme des courbes dont on examine les points d'intersection, Gauss veut donner une démonstration « analytique ».

**Propriété 1** Pour tout $\varphi \in \mathbb{R}$ et pour $r = R$, $tt' + uu' > 0$.

**Lemme 1** Les fonctions :

$$\begin{aligned}
T(\varphi) = \quad & R^n \cos\frac{\pi}{4} + a_1 R^{n-1}\cos(\frac{\pi}{4}+\varphi) + ... \\
& + a_{n-1} R\cos(\frac{\pi}{4}+(n-1)\varphi) + a_n cos(\frac{\pi}{4}+n\varphi) \\
U(\varphi) = \quad & R^n \sin\frac{\pi}{4} + a_1 R^{n-1}\sin(\frac{\pi}{4}+\varphi) + ... \\
& + a_{n-1} R\sin(\frac{\pi}{4}+(n-1)\varphi) + a_n sin(\frac{\pi}{4}+n\varphi) \\
T'(\varphi) = \quad & nR^n \cos\frac{\pi}{4} + (n-1)a_1 R^{n-1}\cos(\frac{\pi}{4}+\varphi) + ... \\
& + a_{n-1} R\cos(\frac{\pi}{4}+(n-1)\varphi) \\
U'(\varphi) = \quad & nR^n \sin\frac{\pi}{4} + (n-1)a_1 R^{n-1}\sin(\frac{\pi}{4}+\varphi) + ... \\
& + a_{n-1} R\sin(\frac{\pi}{4}+(n-1)\varphi)
\end{aligned}$$

sont toutes positives pour n'importe quelle valeur de $\varphi$.

**Démonstration du lemme 1**

Soit $\varphi \in \mathbb{R}$. On a de plus $R^n \cos\frac{\pi}{4} = \frac{R^n}{\sqrt{2}}$.

$$\begin{aligned}
T(\varphi) &= R^n \cos\frac{\pi}{4} + a_1 R^{n-1}\cos(\frac{\pi}{4}+\varphi) + ... \\
&\quad + a_{n-1} R\cos(\frac{\pi}{4}+(n-1)\varphi) + a_n cos(\frac{\pi}{4}+n\varphi) \\
&= n\frac{R^n}{n\sqrt{2}} + a_1 R^{n-1}\cos(\frac{\pi}{4}+\varphi) + ... \\
&\quad + a_{n-1} R\cos(\frac{\pi}{4}+(n-1)\varphi) + a_n cos(\frac{\pi}{4}+n\varphi) \\
&= \frac{R^n}{n\sqrt{2}} + a_1 R^{n-1}\cos(\frac{\pi}{4}+\varphi) \\
&\quad + \frac{R^n}{n\sqrt{2}} + a_2 R^{n-2}\cos(\frac{\pi}{4}+2\varphi) \\
&\quad + ... \\
&\quad + \frac{R^n}{n\sqrt{2}} + a_n \cos(\frac{\pi}{4}+n\varphi) \\
&= \frac{R^{n-1}}{n\sqrt{2}}\left[R + na_1\sqrt{2}\cos(\frac{\pi}{4}+\varphi)\right] \\
&\quad + \frac{R^{n-2}}{n\sqrt{2}}\left[R^2 + na_2\sqrt{2}\cos(\frac{\pi}{4}+2\varphi)\right] \\
&\quad + ... \\
&\quad + \frac{1}{n\sqrt{2}}\left[R^n + na_n\sqrt{2}\cos(\frac{\pi}{4}+n\varphi)\right]
\end{aligned}$$

Or

$$R > \sup(n|a_1|\sqrt{2}, \sqrt{n|a_2|\sqrt{2}}, \sqrt[3]{n|a_3|\sqrt{2}}, ..., \sqrt[n]{n|a_n|\sqrt{2}})$$

Donc , pour tout $i$ compris entre 0 et $n$, on a

$$R > |na_i\sqrt{2}\cos(\frac{\pi}{4} + i\varphi)|$$

et ainsi

$$R + na_i\sqrt{2}\cos(\frac{\pi}{4} + i\varphi) > 0$$

Finalement $T > 0$. On démontrerait de même que $U > 0$, $U' > 0$ et $T' > 0$.

**Lemme 2** Pour $r = R$,
$t, t', u, u'$ deviennent respectivement :
$t(R,\varphi) = T\cos(\frac{\pi}{4} + n\varphi) + U\sin(\frac{\pi}{4} + n\varphi)$
$u(R,\varphi) = T\sin(\frac{\pi}{4} + n\varphi) - U\cos(\frac{\pi}{4} + n\varphi)$
$t'(R,\varphi) = T'\cos(\frac{\pi}{4} + n\varphi) + U'\sin(\frac{\pi}{4} + n\varphi)$
$u'(R,\varphi) = T'\sin(\frac{\pi}{4} + n\varphi) - U'\cos(\frac{\pi}{4} + n\varphi)$

**Démonstration du lemme 2**
On a :

$$\begin{aligned}
&T\cos(\frac{\pi}{4} + n\varphi) + U\sin(\frac{\pi}{4} + n\varphi)\\
&= [R^n\cos\frac{\pi}{4}\cos(\frac{\pi}{4} + n\varphi) + a_1R^{n-1}\cos(\frac{\pi}{4} + \varphi)cos(\frac{\pi}{4} + n\varphi) + ...\\
&+ a_{n-1}R\cos(\frac{\pi}{4} + (n-1)\varphi)cos(\frac{\pi}{4} + n\varphi) + a_ncos(\frac{\pi}{4} + n\varphi)cos(\frac{\pi}{4} + n\varphi)]\\
&+ [R^n\sin\frac{\pi}{4}\sin(\frac{\pi}{4} + n\varphi) + a_1R^{n-1}\sin(\frac{\pi}{4} + \varphi)sin(\frac{\pi}{4} + n\varphi) + ...\\
&+ a_{n-1}R\sin(\frac{\pi}{4} + (n-1)\varphi)sin(\frac{\pi}{4} + n\varphi) + a_nsin(\frac{\pi}{4} + n\varphi)sin(\frac{\pi}{4} + n\varphi)]
\end{aligned}$$

Or $\cos a\cos b + \sin a\sin b = \cos(a - b)$.

Donc, pour tout $i$ compris entre 0 et $n$ :

$$\begin{aligned}
& a_iR^{n-i}\cos(\frac{\pi}{4} + i\varphi)\cos(\frac{\pi}{4} + n\varphi) + a_iR^{n-i}\sin(\frac{\pi}{4} + i\varphi)\sin(\frac{\pi}{4} + n\varphi)\\
= \quad & a_iR^{n-i}\cos(n-i)\varphi
\end{aligned}$$

Ainsi $a_iR^{n-i}\cos(\frac{\pi}{4} + i\varphi) = t(R,\varphi)$.

Les trois autres égalités se démontrent de la même manière.

**Démonstration de la propriété 1**

D'après le lemme 2, pour $r = R$, on a :

$$\begin{aligned}
tt' + uu' =&(T\cos(\frac{\pi}{4}+n\varphi) + U\sin(\frac{\pi}{4}+n\varphi))(T'\cos(\frac{\pi}{4}+n\varphi) + U'\sin(\frac{\pi}{4}+n\varphi))\\
&+ (T\sin(\frac{\pi}{4}+n\varphi) - U\cos(\frac{\pi}{4}+n\varphi))(T'\sin(\frac{\pi}{4}+n\varphi) - U'\cos(\frac{\pi}{4}+n\varphi))\\
=&TT'\cos^2(\frac{\pi}{4}+n\varphi) + UU'\sin^2(\frac{\pi}{4}+n\varphi)\\
&+ TU'\sin(\frac{\pi}{4}+n\varphi)\cos(\frac{\pi}{4}+n\varphi) + T'U\sin(\frac{\pi}{4}+n\varphi)\cos(\frac{\pi}{4}+n\varphi)\\
&- TU'\sin(\frac{\pi}{4}+n\varphi)\cos(\frac{\pi}{4}+n\varphi) - T'U\sin(\frac{\pi}{4}+n\varphi)\cos(\frac{\pi}{4}+n\varphi)\\
&+ TT'\sin^2(\frac{\pi}{4}+n\varphi) + UU'\cos^2(\frac{\pi}{4}+n\varphi)\\
=&TT' + UU' > 0
\end{aligned}$$

De la même manière, on montre que pour $r = R$, on a :

$$t^2 + u^2 = T^2 + U^2 > 0$$

Ainsi, pour $r = R$, on ne peut avoir $t(R, \varphi) = u(R, \varphi) = 0$.

Par conséquent, $R$ étant arbitrairement choisi tel que

$$R > \sup(n|a_1|\sqrt{2}, \sqrt{n|a_2|\sqrt{2}}, \sqrt[3]{n|a_3|\sqrt{2}}, ..., \sqrt[n]{n|a_n|\sqrt{2}})$$

on peut conclure que pour $r \geq R$, on ne peut avoir $t(R, \varphi) = u(R, \varphi) = 0$. On pourra donc à partir de maintenant limiter notre recherche pour $r \in [0; R]$.

#### 15.3.1.3 Plusieurs calculs pour une même intégrale

Pour montrer l'existence d'une racine commune aux deux fonctions $t$ et $u$, Gauss suppose qu'il n'en existe aucune, et calcule ensuite l'intégrale double

$$\Omega = \int_0^R \int_0^{2\pi} y d\phi dr = \int_0^{2\pi} \int_0^R y dr d\phi$$

de deux façons différentes. On obtiendra alors deux résultats incompatibles et l'existence de la racine sera ainsi assurée. C'est donc, contrairement aux précédentes, une démonstration par l'absurde que propose ici Gauss.

**Théorème 1** $\exists r_0 \in [0; R]$ et $\varphi_0 \in [0; 2\pi]$ tels que $t(r_0, \varphi_0) = u(r_0, \varphi_0) = 0$.

**Démonstration du théorème 1** On le démontre par l'absurde. On suppose donc que pour $r \in [0; R]$ et $\varphi \in [0; 2\pi]$ , $t(r, \varphi)$ et $u(r, \varphi)$ ne s'annulent jamais ensemble, et donc que $t^2 + u^2 > 0$.

**Remarque 1** $y = \frac{(t^2+u^2)(tt''+uu'')+(tu'-ut')^2-(tt'+uu')^2}{r(t^2+u^2)^2}$, et comme $t$ et $u$ sont divisibles par $r$ et que l'on a supposé $t^2+u^2>0$ , $y$ est bien définie et continue pour $r \in [0;R]$ et $\varphi \in [0;2\pi]$. L'intégrale $\Omega = \int_0^R \int_0^{2\pi} yd\phi dr = \int_0^{2\pi} \int_0^R ydrd\phi$ existe donc bien.

**Lemme 3**

$$\int_0^{2\pi} yd\phi = \left[\frac{tu'-ut'}{r(t^2+u^2)}\right]_0^{2\pi} = 0$$

**Démonstration du lemme 3**

On remarque que :

$$\begin{cases} \dfrac{dt}{d\phi} = -u' \\ \dfrac{du}{d\phi} = t' \\ \dfrac{dt'}{d\phi} = -u'' \\ \dfrac{du'}{d\phi} = t'' \end{cases}$$

On a alors

$$\begin{aligned} &\frac{d}{d\phi}\left(\frac{tu'-ut'}{r(t^2+u^2)}\right) = \frac{(t^2+u^2)(-u'^2+tt''-t'^2+uu'')-(-2tu'+2t'u)(tu'-ut')}{r(t^2+u^2)^2} \\ =&\frac{(t^2+u^2)(tt''+uu'')-(t^2+u^2)(u'^2+t'^2)+2(tu'-ut')^2}{r(t^2+u^2)^2} \end{aligned}$$

Or

$$-(t^2+u^2)(u'^2+t'^2)+(tu'-ut')^2 = -(u'u+t't)^2$$

Donc

$$\frac{d}{d\phi}\left(\frac{tu'-ut'}{r(t^2+u^2)}\right) = \frac{(t^2+u^2)(tt''+uu'')+(tu'-ut')^2-(u'u+t't)^2}{r(t^2+u^2)^2} = y$$

D'où

$$\int_0^{2\pi} yd\phi = \left[\frac{tu'-ut'}{r(t^2+u^2)}\right]_0^{2\pi}$$

De plus $u(r,0) = u'(r,0) = u''(r,0) = u(r,2\pi) = u'(r,2\pi) = u''(r,2\pi) = 0$ Ainsi

$$\int_0^{2\pi} yd\phi = \left[\frac{tu'-ut'}{r(t^2+u^2)}\right]_0^{2\pi} = 0$$

**Lemme 4**

$$\int_0^R ydr = \left[\frac{tt'+uu'}{t^2+u^2}\right]_0^R = \frac{TT'+UU'}{T^2+U^2} > 0$$

**Démonstration du lemme 4** On remarque que

$$\begin{cases} \dfrac{dt}{dr} = \dfrac{1}{r}t' \\ \dfrac{du}{dr} = \dfrac{1}{r}u' \\ \dfrac{dt'}{dr} = \dfrac{1}{r}t'' \\ \dfrac{du'}{dr} = \dfrac{1}{r}u'' \end{cases}$$

Ainsi

$$\begin{aligned} \frac{d}{dr}\left(\frac{tt'+uu'}{t^2+u^2}\right) =& \frac{(t^2+u^2)(\frac{1}{r}t'^2+\frac{1}{r}tt''+\frac{1}{r}u'^2+\frac{1}{r}uu'')-(\frac{2}{r}tt'+\frac{2}{r}u'u)(tt'+uu')}{(t^2+u^2)^2} \\ =& \frac{(t^2+u^2)(tt''+uu'')+(t^2+u^2)(t'^2+u'^2)-2(tt'+u'u)^2}{r(t^2+u^2)^2} \\ =& y \end{aligned}$$

car $(t^2+u^2)(t'^2+u'^2)-(tt'+uu') = (tu'-ut')^2$.
On a vu que

$$\frac{t(R,\phi)t'(R,\phi)+u(R,\phi)u'(R,\phi)}{t^2(R,\phi)+u^2(R,\phi)} = \frac{TT'+UU'}{T^2+U^2}$$

De plus

$$\frac{t(0,\phi)t'(0,\phi)+u(0,\phi)u'(0,\phi)}{t^2(0,\phi)+u^2(0,\phi)} = 0$$

Donc

$$\int_0^R y dr = \left[\frac{tt'+uu'}{t^2+u^2}\right]_0^R = \frac{TT'+UU'}{T^2+U^2} > 0$$

**Reprise de la preuve du théorème 1**

Comme $\int_0^{2\pi} y d\phi = \left[\frac{tu'-ut'}{r(t^2+u^2)}\right]_0^{2\pi} = 0$, on a $\Omega = \int_0^R \int_0^{2\pi} y d\phi dr = 0$.
De plus, comme

$$\int_0^R y dr = \left[\frac{tt'+uu'}{t^2+u^2}\right]_0^R = \frac{TT'+UU'}{T^2+U^2} > 0$$

on a

$$\Omega = \int_0^{2\pi}\int_0^R y dr d\phi = \int_0^{2\pi} \frac{TT'+UU'}{T^2+U^2} > 0$$

Ainsi $\Omega = 0$ et $\Omega > 0$, ce qui est absurde. Il existe donc bien $r_0 \in [0;R]$ et $\varphi_0 \in [0;2\pi]$ tels que $t(R_0,\varphi_0) = u(R_0,\varphi_0) = 0$.

#### 15.3.1.4 Démonstration du TFA

**Théorème 2 (Théorème fondamental de l'algèbre)** Soit $P = x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_{n-1}x + a_n \in \mathbb{R}[x]$. Alors $P$ est divisible par un polynôme de $\mathbb{R}[x]$ de degré

1 ou 2.

**Démonstration du théorème 2**

$$P(r(\cos\phi + i\sin\phi)) = t(r,\phi) + iu(r,\phi) \text{ et } P(r(\cos\phi - i\sin\phi)) = t(r,\phi) - iu(r,\phi).$$

Ainsi $P(r_0(\cos\phi_0 + i\sin\phi_0)) = t(r_0,\phi_0) + iu(r_0,\phi_0) = 0$ et $P(r_0(\cos\phi_0 - i\sin\phi_0)) = t(r_0,\phi_0) - iu(r_0,\phi_0) = 0$.

— Si $\sin\phi_0 = 0$, alors $P$ est divisible par $x - r_0\cos\varphi_0 \in \mathbb{R}[x]$

— Si $\sin\phi_0 \neq 0$, alors $P$ est divisible par $(x - r_0(\cos\phi_0 + i\sin\phi_0))(x - r_0(\cos\phi_0 - i\sin\phi_0)) = x^2 - 2r_0(\cos\phi)x + r_0^2 \in \mathbb{R}[x]$

Le théorème est ainsi démontré.

Le mémoire de Gauss ne s'arrête pas sur cette démonstration : il ajoute un court chapitre dans lequel il tente d'interpréter graphiquement des intégrales qu'il vient de manipuler [57] :

> Et pour illustrer cela davantage, nous observons que pour de telles valeurs on a $tt + uu = 0$, et en outre que $y$ lui-même est indéfini, d'où il ne sera pas davantage permis de traiter l'intégrale double $\int\int ydrd\varphi$ comme une quantité définie. En effet, de manière générale, en notant $\xi, \eta, \zeta$ les coordonnées de points dans l'espace, l'intégrale $\int\int ydrd\varphi$ représente le volume d'un solide qui est contenu entre cinq plans dont les équations sont
>
> $$\xi = 0, \eta = 0, \zeta = 0, \xi = R, \eta = 360$$
>
> et par une surface dont l'équation est $\zeta = y$, en considérant comme négatives les parties dans lesquelles les coordonnées $\zeta$ sont négatives. Mais on peut comprendre que la condition tacite de la *continuité* de la sixième surface cesse lorsque $y$ est indéfini, et qu'il peut advenir que cette représentation n'ait plus de signification. Dans un tel cas, il ne peut y avoir de discussion sur l'interprétation de l'intégrale $\int\int ydrd\varphi$. [Gauss, 1816b] p. 63

C'est finalement, en ce qui concerne l'interprétation graphique de ce qu'il vient de faire, un constat d'échec que Gauss fait ici. Le problème vient du fait qu'il cherche à illustrer

57. Quod ut magis illustretur, observamus, pro talibus valoribus fieri $tt + uu = 0$, adeoque ipsam $y$ infinitam, unde haud amplius licebit, integrale duplex $\int\int ydrd\varphi$ tamquam quantitatem assignabilem tractare. Generaliter quidem loquendo, denotantibus $\xi, \eta, \zeta$ indefinite coordinatas punctorum in spatio, integrale $\int\int ydrd\varphi$ exhibet volumen solidi, quod continetur inter quinque plana, quorum aequationes sunt

$$\xi = 0, \eta = 0, \zeta = 0, \xi = R, \eta = 360$$

atque superficiem, cuius aequatio $\zeta = y$ considerando eas partes tamquam negativas, in quibus coordinatae $\zeta$ sunt negativae. Sed tacite hic subintelligitur, superficiem sextam esse *continuam*, qua conditione cessante, dum $y$ evadit infinita, utique fieri potest, ut conceptus ille sensu careat. In tali casu de integrali $\int\int ydrd\varphi$ clligendo sermo esse nequit.

l'hypothèse qui lui permettra d'aboutir à une contradiction : c'est de méthode dont il parle ici, et ce qu'il veut montrer c'est qu'« il ne faut pas en outre s'étonner que les opérations analytiques appliquées par un calcul aveugle à des choses vaines conduisent à des absurdités » [58].

### 15.3.2 La reprise par Kronecker

Maxime Bôcher, dans un article [59] intitulé *Gauss's third proof of the fundamental theorem of algebra* et publié en 1895 montre comment cette troisième preuve peut être réinterprétée à partir de la théorie du potentiel, théorie sur laquelle Gauss a travaillé, et en particulier dans son mémoire *Theoria Attractionis Corporum Sphaeroidicorum Ellipticorum Homogeneorum* de 1822. Or nous avons vu que l'intégrale que propose Kronecker pour calculer la caractéristique d'un système de fonctions est une généralisation de celle qu'étudie Gauss dans cet article [60].

#### 15.3.2.1 Un outil commun

En utilisant le calcul de la caractéristique d'un système de trois fonctions à deux variables, Kronecker propose une nouvelle démonstration du théorème fondamental de l'algèbre. Celle-ci, contrairement à la précédente, ne suit pas exactement la version de Gauss. Si l'on peut rapprocher ces deux démonstrations, c'est par les outils utilisés, c'est-à-dire ici l'utilisation du calcul intégral. Le cœur de la méthode de Gauss, c'est-à-dire la *reductio ad absurdum*, n'est pas celle employée par Kronecker, qui fait au contraire un calcul direct pour montrer que tout polynôme de degré $n$ a $n$ racines complexes.

Kronecker a montré, rappelons-le, que la caractéristique de trois fonctions $f, \varphi$ et $\psi$ prenait la valeur

$$\xi(f,\varphi,\psi) = \frac{1}{2\pi}\int_{f(x,y)=0} \frac{\varphi.d\psi - \psi.d\varphi}{\varphi^2+\psi^2}$$

Cette intégrale, on le voit maintenant, est bien similaire à celle que Gauss a introduite dans sa démonstration, même si ce dernier travaille avec des intégrale doubles lorsque Kronecker utilise l'intégrale curviligne de Cauchy.

Cette démonstration n'est en fait présente que dans deux manuscrits : ceux de 1872-73 et de 1890-91. Dans le premier, Kronecker déduit, comme on l'a vu, de la théorie des caractéristiques l'intégrale de Cauchy, puis il utilise cette dernière pour démontrer, comme Cauchy l'a fait, le TFA [61]. Regardons plus précisément comment, dans le cours [62]

58. [Gauss, 1816b] p. 63 : « neque adeo mirandum est, operationes analyticas coeco calculo ad inania applicatas ad absurda perducere. »
59. Voir [Bôcher, 1895].
60. Voir [Kronecker, 1869c] chapitre V.
61. Il donnera aussi cette démonstration dans son cours sur le calcul intégral : on peut à se propos voir le manuscrit de ce cours du semestre d'été de 1889 ([Kronecker, 1889a] p. 108-110).
62. Voir [Kronecker, 1891a] p. 247 et suivantes.

de 1890-91, il construit sa démonstration dans le cadre de son chapitre sur la théorie des caractéristiques.

#### 15.3.2.2 Démonstration de Kronecker

Kronecker a obtenu [63] le fait que lorsque l'on se place dans le cas particulier où

$$\Theta(x+iy) = \varphi(x,y) + i\psi(x,y)$$

on a :

$$\chi(f,\varphi,\psi) = \frac{1}{2i\pi}\int_{f(x,y)=0} d\log\Theta(x+yi)$$

c'est-à-dire

$$\chi(f,\varphi,\psi) = \frac{1}{2i\pi}\int_{f(x,y)=0} \frac{\Theta'(x+yi)}{\Theta(x+yi)}$$

Il l'applique avec $\Theta = \varphi + i\psi = z^n + c_1 z^{n-1} + ... \pm c_n$, un polynôme de degré $n$, et $f = 0$ un cercle de rayon $R$ contenant toutes les intersections des courbes $\varphi = 0$ et $\psi = 0$. En posant $z = x + iy = Re^{iw}$, on obtient :

$$\begin{aligned}
\chi(f,\varphi,\psi) &= \frac{1}{2\pi i}\int_0^{2\pi} \frac{n.R^{n-1}.e^{(n-1)\omega i} - (n-1)R^{n-2}.e^{(n-2)\omega i} + ...}{R^n.e^{n\omega i} - c_1 R^{n-1}.e^{(n-1)\omega i} + ...}.d(R.e^{\omega i}) \\
&= \frac{1}{2\pi}\int_0^{2\pi} \frac{n.R^n.e^{n\omega i} - (n-1)R^{n-1}.e^{(n-1)\omega i} + ...}{R^n.e^{n\omega i} - c_1 R^{n-1}.e^{(n-1)\omega i} + ...}.d\omega \\
&= \frac{n}{2\pi}\int_0^{2\pi} d\omega.\frac{1 - \frac{n-1}{n.R.e^{\omega i}} + ...}{1 - \frac{c_1}{R.e^{\omega i}} + ...}
\end{aligned}$$

La caractéristique ne se modifie pas lorsque le rayon $R$ du cercle augmente, car elle ne peut se modifier que lorsque l'on a simultanément $\varphi = \psi = 0$. Or lorsque $R$ devient très grand, l'intégrale

$$\frac{n}{2\pi}\int_0^{2\pi} d\omega.\frac{1 - \frac{n-1}{n.R.e^{\omega i}} + ...}{1 - \frac{c_1}{R.e^{\omega i}} + ...}$$

est de plus en plus proche de

$$\frac{n}{2\pi}\int_0^{2\pi} d\omega = n$$

Comme la caractéristique est un nombre entier, elle est donc égale à $n$ : le théorème est ainsi démontré.

63. Voir plus haut p. 416.

#### 15.3.2.3 Conclusion

Kronecker utilise ici à la fois l'interprétation graphique de la caractéristique comme intersection de courbes, l'intégrale qui en donne la valeur, mais aussi dans son dernier cours, et c'est là le point essentiel qui lui permet de conclure sa preuve, l'invariance de la caractéristique lorsque l'on modifie continument les coefficients, sans « passer » par $f = \varphi = \psi = 0$.

Cette démonstration est très proche de celle donnée par Cauchy dans son cours et ce qui la place dans la partie du cours de Kronecker consacrée à la théorie des caractéristiques, c'est qu'il déduit les résultats de Cauchy à partir de cette dernière : ces résultats ne deviennent finalement que l'un des cas particuliers de la théorie. On a vu que le rapprochement que l'on peut faire avec la troisième preuve de Gauss est plus ténu que dans le cas de ses trois autres démonstrations. Cependant, la proximité des objets utilisés nous a permis de mettre en évidence une parenté entre ces deux travaux.

## 15.4 Vers un nouvel énoncé

On retrouve donc bien les trois différentes démonstrations de Gauss dans le cours de Kronecker sur la théorie des équations algébriques. Selon les années, elles ne sont pas toutes présentes : lorsque la théorie des caractéristiques n'est pas exposée, la démonstration « topologique » de Gauss ne l'est pas non plus. Par contre, quel que soit le manuscrit, on trouve toujours une version de la seconde preuve de Gauss.

Kronecker fournit, dans ses premiers cours, un travail très proche de celui de Gauss. Mais au fur et à mesure de ses publications, son exposition du TFA évolue. Les démonstrations de 1799 et 1849, ainsi que la seconde de 1816, s'appuient sur des phénomènes d'invariance de la caractéristique, et s'intègrent ainsi de plus en plus dans son chapitre consacré à cette théorie. Celle bien plus « algébrique » de 1816 va dans un premier temps suivre presque exactement le mémoire de Gauss, pour devenir dans les dernier cours un théorème assez différent qui portera, on le sait, le nom de théorème fondamental de l'arithmétique.

L'évolution des termes de l'énoncé du théorème que Kronecker propose nous apporte aussi des renseignements précieux sur les changements qui sont à l'œuvre dans son cours. Nous avons déjà vu que dans le cours de 1872-73, Kronecker discute, en se reportant à Gauss, l'utilisation de coefficients complexes. Cependant, la formulation de l'énoncé est parfaitement classique[64] :

> Toute équation de degré $n$ a précisément $n$ racines. [Kronecker, 1873a] p. 185

Dans le manuscrit de 1878-79, le théorème est démontré, mais n'est pas énoncé. Cependant, il n'y a aucune différence quant à ce que Kronecker souhaite prouver. Les changements

64. Jede Glch $n^{\text{ten}}$ Grades hate gerade $n$ complexe Wurzeln.

apparaissent à partir du cours de 1884-85, c'est-à-dire, rappelons-le, après la publication des des *Grundzüge*. Voilà comment Kronecker y énonce ce qu'il souhaite démontrer [65] :

> Gauss a seulement mis en évidence que l'on doit juste montrer qu'il y a des grandeurs telles que si on utilise sur elles les opérations requises par l'expression $ax^n+bx^{n-1}+...$, elles rendent cette expression aussi petite que l'on veut, ou exprimer autrement, si une équation est donnée, alors on doit montrer qu'il y a des grandeurs de la forme $\xi+\eta i$, où $\xi$ et $\eta$ sont des nombres rationnels pour lesquels l'expression devient en valeur absolue aussi petite que l'on veut. [Kronecker, 1885b] p. 382

La discussion sur les coefficients a complètement changé : ceux-ci ne sont ni réels ni complexes, mais rationnels. La notion de racine a elle-même disparu et il s'agit, comme nous l'avons déjà vu lorsque nous avons étudié [66] la notion de racine, plutôt de trouver des intervalles dans lesquels la fonction change exactement une fois de signe.

À partir de 1887 et de son *théorème fondamental de l'arithmétique*, on trouvera aussi un énoncé complètement différent du théorème. En effet, il construira alors un *Primmodulsystem*

$$(f_1 - c_1, f_2 - c_2, ...., f_n - c_n, y' - c', y'' - c'', ...)$$

avec lequel on a la congruence [67] :

$$\mathcal{F}(x) \equiv \prod_{k=1}^{n}(x - x_k)\ (modd. f_1 - c_1, f_2 - c_2, ...., f_n - c_n, y' - c', y'' - c'', ...)$$

Il n'est pas question ici de s'étendre sur le terme de *Primmodulsystem* ou de décrire exactement à quel type de congruence on a affaire [68], mais nous voyons que le TFA change alors de nature pour devenir un *théorème de factorisation linéaire*.

Néanmoins, dans le manuscrit de 1890-91, Kronecker ne traite pas uniquement du théorème de factorisation linéaire, mais il consacre aussi une partie de son cours sur la théorie des caractéristiques au théorème fondamental de l'algèbre. Dans celui-ci, il donne un énoncé du théorème dans lequel le terme de racine n'est pas employé [69] :

65. Gauss erst hat hervorgehoben, dass man gerade beweisen müsse, dass es Grössen von der Art gebe, dass wenn man auf sie die durch den Ausdruck $ax^n + bx^{n-1} + ...$ geforderte Operation anwende, sie diesen Ausdruck beliebig klein machen, oder andern ausgedrückt, wenn eine Gleichung gegeben ist $x^n + ax^{n-1} + ... = 0$, so soll man zeigen, dass es Grössen von der Form $\xi + \eta i$ giebt, wo $\xi$ und $\eta$ rationale Zahlen sind, welche für $x$ eingesetzt, den Ausdruck dem absoluten Betrage nach beliebig klein machen.

66. Dans la section 13.5.2 de cette partie et la section 2.2 de la première partie.

67. Voir par exemple [Kronecker, 1887a] p. 319.

68. Les *Vorlesungen* de Kronecker nous permettraient d'approfondir sa théorie des *Modulsystem* et finalement le contenu des *Grundzüge*, mais cela demanderait de tels développements que nous le renvoyons à un travail ultérieur.

69. Es sei $\Theta(x+yi)$ eine ganze Funktion $n-$ten Grades der komplexen Variablen $(x+yi)$ mit komplexen Koeffizienten und nach Sammlung aller reellen und komplexen Glieder :

$$\Theta(x+iy) = \varphi(x,y) + i.\psi(x,y);$$

> Soit $\Theta(x+yi)$ une fonction entière de degré $n$ de la variable complexe $(x+yi)$ avec des coefficients complexes et soit après le rassemblement de tous les termes réels et complexes :
>
> $$\Theta(x+iy) = \varphi(x,y) + i.\psi(x,y);$$
>
> alors les courbes définies par $\varphi = 0$ et $\psi = 0$ ont toujours $n$ points d'intersection. [Kronecker, 1891a] p. 241

Dans cet énoncé, l'objet de départ est bien un polynôme de degré $n$ [70] Cependant, Kronecker ne le donne ni en termes de racines, ni en termes d'intervalles. C'est l'aspect géométrique qui est mis en avant : on regarde le nombre d'intersections de certaines courbes, et donc une propriété plus topologique que métrique. Ce qui explique l'évolution de l'énoncé du TFA dans les cours de Kronecker, c'est en fait le « statut de la racine réelle » [71] :

> Mais une chose reste non élucidée : le statut de la racine réelle. *Gauss* s'en tient à illustrer géométriquement sa signification et il avait la ferme conviction d'avoir donné ainsi une définition suffisante. Ce qu'il voulait prouver et ce qu'il a rigoureusement démontré a été donné dans une partie précédente avec précision. On aurait ainsi pu s'en contenter comme preuve pour le contenu purement analytique du théorème fondamental – que l'on énonce communément ainsi : toute équation de degré $n$ a $n$ racines – si l'on avait voulu « appliquer d'emblée la continuité géométrique sur des grandeurs algébriques, et donc, pour l'esquisser rapidement, faire correspondre à chaque distance entre deux points un nombre arithmétique ». [Kronecker, 1891a] pp. 249-250

Kronecker, qui comme on l'a vu cite ici Netto, distingue en fait plusieurs contenus dans le théorème fondamental de l'algèbre. Le premier, « purement analytique », est basé sur une notion « géométrique » de racine. Ce contenu, avec de plus en plus de précautions (intervalles plutôt que racines, coefficients rationnels,...), il l'expose dans la partie de son cours dédiée à la théorie des caractéristiques et qui correspond en fait aux première, troisième et quatrième preuves de Gauss. Mais si on l'énonce ainsi, on utilise une correspondance géométrique entre nombre et longueur. Kronecker pose la question suivante à laquelle il refuse de répondre : « Mais alors comment au juste un segment de droite géométrique peut "définir" un concept arithmétique [72] » ? Plus encore, d'un point de vue arithmétique le théorème devient faux [73] :

dann haben die durch $\varphi = 0$ und $\psi = 0$ definierten Kurven stets $n$ Schnittpunkte.

70. On rappelle que dans le vocabulaire de Kronecker une fonction entière (*eine ganze Funktion*) est un polynôme. Ici Kronecker ne précise pas la nature des coefficients, mais dans le cadre de son cours du semestre d'hiver 1890-91, ces coefficients peuvent être considérés comme des nombres rationnels.

71. Aber eines blieb unaufgeklärt : die Statur der reellen Wurzeln. *Gauss* begnügte sich, ihre Bedeutung geometrisch zu veranschaulichen und war der festen Überzeugung, damit eine ausreichend Definition derselben gegeben zu haben. Was er beweisen wollte und wirklich aufs strengste bewiesen hat, ist im vorhergehenden Abschnitt des Näheren angegeben worden. Man könnte es auch damit als Beweis für den rein analytischen Inhalt des Fundamentaltheorems, daß nämlich, wie man gewöhnlich sagt, jede Gleichung $n-$ten Grades $n$ Wurzeln habe, sein Bewenden haben lassen, wenn man « die geometrische Stetigkeit ohne weiteres auf die arithmetischen Größen übertragen, und also, um es kurz anzudeuten, einer jeden Streckenlänge eine arithmetische Zahl entsprechen lassen ».

72. Voir [Kronecker, 1891a] p. 251.

73. Wenn wir den sogen Fundamentalsatz der Algebra « streng arithmetisch fassen und begründen » wollen ohne jegliche geometrische Stütze, so können wir behaupten, daß der eigentliche Fundamentalsatz, daß nämlich jede ganze, rationale Funktion $n-$ten Grades in ein Product von $n$ linearen, reellen oder imaginären Faktoren Zerlegbar sei, überhaupt nicht richtig ist.

> Si nous voulons « appréhender et fonder de façon strictement arithmétique le soi-disant théorème fondamental de l'algèbre » sans aucun soutien géométrique, alors nous pouvons prétendre que le véritable théorème fondamental, à savoir que chaque fonction entière rationnelle de degré $n$ est décomposable en un produit de $n$ facteurs linéaires réels ou imaginaires, n'est pas du tout correct. [Kronecker, 1891a] p. 251

Kronecker propose alors un nouvel énoncé dans son dernier cours sur la théorie des équations algébriques [74] :

> Ce que nous pouvons démontrer et ce que nous démontrerons, et ce qu'en substance Gauß a seulement démontré, c'est que à partir de chaque fonction rationnelle entière de degré $n$ on peut en fabriquer une autre par une variation du système de coefficient de la première aussi petite que l'on veut, cette dernière pouvant être découpée en un produit de $m$ facteurs linéaires et $\frac{n-m}{2}$ facteurs quadratiques et ce de sorte que les coefficients de tous les facteurs soient des nombres rationnels. [Kronecker, 1891a] pp. 251-252

Plutôt que des racines dont l'existence, pour Kronecker, n'est pas assurée, plutôt que des intervalles comme il l'a fait précédemment, Kronecker utilise des petites variations des coefficients. Ces dernières lui permettent d'utiliser des *approximations* de polynôme : il sous-entend donc une métrique sur l'ensemble des polynômes. On peut rapprocher cet énoncé du travail de Kronecker de 1878 sur la théorie des caractéristiques : là aussi, il fait varier les coefficients d'un polynôme, pour ensuite examiner la façon dont la caractéristique du système correspondant évolue.

Finalement, cet énoncé ne sera pas pour Kronecker le seul énoncé pertinent en arithmétique [75] :

> Mais cela n'est qu'un aspect de la version arithmétique du théorème fondamental. Pour les recherches algébriques, nous allons prouver la possibilité d'une deuxième décomposition fondamentalement différente de celle donnée, qui n'est pas valable au

74. Was wir beweisen können und beweisen werden und was, im Grunde genommen, auf Gauß nur beweisen hat, ist, daß sich aus jeder ganzen rationalen Funktion $n-$ten Grades eine andere durch ein Variirung des Koeffizientensystems der ersteren um beliebig kleine Größen herstellen läßt, welche in ein Produkt von $m$ linear-und $\frac{n-m}{2}$ quadratischen Faktoren zerlegt werden kann und zwar so, daß die Koeffizienten dieser sämtlichen Faktoren rationale Zahlen sind.

75. Aber dies ist auch nur die eine Seite, die arithmetische Fassung des Fundamentaltheorems. Für algebraische Untersuchungen werden wir die Möglichkeit einer zweiten von der angeführten grundverschiedenen Zerlegung nachweisen, welche nicht im Sinne der Gleichheit, sondern nur als Kongruenz für ein Primmodulsystem besteht. Wir werden nämlich die Richtigkeit der folgenden Kongruenz aufzeigen :

$$F(x) = x^n - c_1x^{n-1} + c_2x^{n-2} - ... \pm c_n \equiv (x-x_1)(x-x_2)...(x-x_n)$$
$$(modd.\ f_1 - c_1, f_2 - c_2, ..., f_n - c_n, y - c),$$

wenn

$$(x-x_1)(x-x_2)...(x-x_n) = x^n - f_1x^{n-1} + f_2x^{n-2} - ... \pm f_n;$$

$x_1...x_n$ unbestimmte Variable, $y$ eine gewisse von der algebraischen Beschaffenheit der Funktion $F(x)$ abhängende, nicht symmetrische Funktion von $x_1, x_2...x_n$ und $c$ eine rationale Zahl ist, die zu $y$ in derselben Beziehung steht, wie die Größen $f_1, f_2, ...f_n$ bezw. zu $c_1, c_2..c_n$. Nur wenn man so den arithmetischen und den algebraischen Teil des Fundamental-Theorems in aller Strenge sondert, ist es möglich, seinen Inhalt wirklich scharf zu fassen.

sens de l'égalité mais seulement en tant que congruence pour un *Primmodulsystem*. Nous allons en effet montrer la validité de la congruence suivante :

$$F(x) = x^n - c_1x^{n-1} + c_2x^{n-2} - ... \pm c_n \equiv (x - x_1)(x - x_2)...(x - x_n)$$
$$(modd.\ f_1 - c_1, f_2 - c_2, ..., f_n - c_n, y - c),$$

lorsque

$$(x - x_1)(x - x_2)...(x - x_n) = x^n - f_1x^{n-1} + f_2x^{n-2} - ... \pm f_n;$$

$x_1...x_n$ sont des variables indéterminées, $y$ une certaine variable dépendant de la qualité algébrique de la fonction $F(x)$, qui n'est pas une fonction symétrique de $x_1, x_2...x_n$, et $c$ est un nombre rationnel qui est pour $y$ dans le même rapport que $f_1, f_2, ...f_n$ pour resp. $c_1, c_2..c_n$. C'est seulement lorsque l'on sépare ainsi en toute rigueur la partie arithmétique et algébrique du théorème fondamental, qu'il est possible d'appréhender véritablement précisément son contenu. [Kronecker, 1891a] p. 252

On trouve donc finalement deux énoncés que Kronecker qualifie d'*arithmétiques*. Le premier, lié à la théorie des caractéristiques, est obtenu en remplaçant le concept de racine par celui d'intervalle rationnel. Le second est en fait le théorème de factorisation linéaire qu'il présentera dans son article de 1887.

# Conclusion générale

Lorsque l'on examine les cours sur la théorie des équations algébriques qui sont à notre disposition, on s'aperçoit qu'entre 1882 et 1890 la théorie des caractéristiques a presque complètement disparu [76]. Avant 1882, elle constituait pourtant une partie importante du cours de Kronecker, la seule d'ailleurs que Runge semble avoir prise en note [77] lors de son passage à Berlin entre 1879 et 1881. L'explication de cette disparition doit très probablement être cherchée dans les remaniements qui ont conduit à la remplacer par la théorie des *Gattungen* et des *Modulsysteme* : ces derniers sont en effet exposés dans les *Grundzüge einer arithmetischen Theorie der algebraischen Grössen* dont la date de publication, 1883, coïncide justement avec le retrait de la théorie des caractéristiques des *Vorlesungen*.

Dans les cours qui contiennent la théorie des caractéristiques, l'une de ses applications importantes, nous venons de le voir, est le théorème fondamental de l'algèbre. Cette théorie lui permet en effet de retrouver à la fois les première et quatrième preuves de Gauss, mais aussi la troisième. Au contraire, les *Modulsysteme*, largement inspirés par Gauss et qui consistent en fait en une extension des congruences, constituent le cadre dans lequel Kronecker développe et reprend une grande partie de ce que Gauss met en place dans sa seconde démonstration du théorème fondamental. Kronecker transforme ainsi ce théorème en « théorème fondamental de l'arithmétique », c'est-à-dire en ce que C. Gilain [78] appelle le théorème de factorisation linéaire [79]. Cela souligne une fois encore l'importance que revêtent pour Kronecker les discussions autour de la notion de racine : la façon dont il l'aborde, théorie des caractéristiques ou *Modulsysteme*, orientera le *contenu* même de son cours. Par ailleurs, les développements de la théorie des caractéristiques conduisent Kronecker vers des méthodes et des questions portant sur l'intégration, et en particulier l'intégration sur des variétés de dimension $n$ : la place de ces méthodes dans un cours sur la théorie des équations algébriques se justifie donc moins. Ainsi, au fil des ans, Kronecker développe dans ses cours une *théorie des équations algébriques* qu'il cherche à intégrer de plus en plus dans son *arithmétique générale*, et dans laquelle la théorie des caractéristiques ne semble plus avoir sa place.

76. Lors du semestre d'hiver 1888-89 la théorie des caractéristiques pour deux fonctions est toutefois rapidement abordée. Voir [Kronecker, 1889b] p. 132

77. [Runge, 1881]

78. [Gilain, 1991]

79. Rappelons que Kronecker en donne une démonstration dans *Ein Fundamentalsatz der allgemeinen Arithmetik* ([Kronecker, 1887c]).

Pourquoi cette théorie revient-elle alors dans le dernier cours de Kronecker après dix ans d'absence ?

Nous avons déjà mentionné la retranscription d'un entretien avec Hilbert sur la théorie des invariants en 1888, dans lequel ce dernier affirme que Kronecker décrivait sa théorie des caractéristiques comme pouvant pour ainsi dire « tout faire ». Cette discussion concerne en fait essentiellement la classification topologique des courbes algébriques, et le théorème de Harnack (1876) dont il est question dit que le nombre de composantes connexes d'une courbe algébrique de degré $d$ est borné par :

$$\frac{1-(-1)^d}{2} < c < \frac{(d-1)(d-2)}{2} + 1$$

Kronecker semble ici suggérer que sa théorie des caractéristiques serait capable de résoudre ce qui devrait devenir le 16ème problème de Hilbert sur la classification topologique des courbes algébriques. Même si David Rowe juge cette remarque très (sans doute même beaucoup trop) optimiste, elle révèle la vision profondément topologique que Kronecker avait alors de sa théorie. Et de fait, elle sera principalement exploitée à la fin du XIX$^{e}$ siècle et au début du XX$^{e}$ siècle dans l'élaboration de la topologie, et plus particulièrement de la topologie algébrique.

L'importance de la partie consacrée à la théorie des caractéristiques dans son dernier cours témoigne aussi du lien qui unit cours et recherche à l'université de Berlin : s'inscrivant de nouveau dans l'actualité de ses recherches [80], la théorie des caractéristiques réintègre le cœur des leçons de Kronecker. Dans le cas de son œuvre algébrique, il semble bien que la méthode consistant à exploiter les cours pour tenter de mieux comprendre les recherches s'avère extraordinairement efficace. En ce qui concerne la théorie des caractéristiques, c'est maintenant acquis, mais il serait possible – et souhaitable – de prolonger ce travail aux chapitres plus notoirement difficiles, comme par exemple ceux relevant de l'exploitation des *Modulsysteme*.

## De l'« algébrisation » à la « généralisation »

Nous avons tenté de montrer que la théorie des caractéristiques, loin d'être seulement la conclusion d'un « cycle », est un point de convergence entre différentes théories élaborées au cours du XIX$^{e}$ siècle. Plus encore, nous l'avons vu, cette théorie est le produit d'une volonté de *généralisation* de la part de Kronecker, car elle cherche à englober le travail de Sturm, de Sylvester, de Cayley et une partie de la théorie des indices de Cauchy. Comment ces deux mouvements s'articulent-ils ? Notre seconde partie témoigne d'un travail sur le théorème de Sturm qui participe bien de l'« algébrisation » qu'analysait H. Sinaceur. Cette « algébrisation », qui consiste à écarter les notions issues de l'analyse – on a vu que Sylvester proposait un travail sur Sturm dans lequel la notion même de

80. Et aussi d'autres mathématiciens, comme par exemple Poincaré, Picard ou Van Dyck.

fonction est absente – est poursuivie par Kronecker lorsqu'il propose, démonstration après démonstration, d'exclure tout concept de continuité – voire de racine – du théorème de Sturm. Dans son cours sur la théorie des équations algébriques, l'algébrisation de Kronecker devient au fur et à mesure des années une *arithmétisation* au sens de son *arithmétique générale*. La théorie des caractéristiques est introduite pour « poursuivre arithmétiquement » [81] les recherches sur le théorème de Sturm, et si en 1880 Kronecker qualifie de « purement analytique » [82] la mise en place du *Fortgangsprincip*, en 1890 il veut formuler « arithmétiquement » [83] ce même principe.

Pourtant, au bout de la route, par exemple lorsqu'il utilise la méthode d'Hermite-Jacobi, cette démarche entre en tension avec une autre grande volonté qui traverse tout son travail, la volonté de « généralisation ». Plus encore, nous avons montré comment, en construisant sa théorie des caractéristiques, Kronecker se rend compte dès son article de 1869 que le chemin de l'« algébrisation » ne peut plus être poursuivi. Il constate en effet que travailler, comme Sylvester, uniquement sur les « formes algébriques » [84] amène à un travail d'interprétation de plus en plus complexe au fur et à mesure que la dimension des espaces considérés augmente. Plus encore, chaque étape demande des considérations spécifiques, ce qui va à l'encontre de la volonté de généralisation qui traverse tout son travail.

Dans le chapitre intitulé *Algebraic generality vs arithmetic generality in the 1874 controversy between C. Jordan and L. Kronecker* du livre [Chemla *et al.*, 2016], F. Brechenmacher revient sur la controverse qui a opposé Kronecker et Jordan au sujet de la théorie des formes bilinéaires [85]. Il montre que Kronecker reproche à Jordan sa « fausse » généralité lorsqu'il présente un cas « générique » – il reprend ici l'expression de Hawkins (Hawkins 1977 : 122) -, c'est-à-dire ici un cas où l'équation n'a pas de racines multiples. Ce « raisonnement générique », F. Brechenmacher rappelle que Hawkins le définit ainsi :

> Analysis became a method for reasoning with, manipulating, expressions involving symbols with "general" values and a tendency developed to think almost exclusively in terms of the "general" case with little, if any, attention given to potential difficulties or inaccuracies that might be caused by assigning certain specific values to the symbols. Such reasoning with "general" expressions I shall refer to for the sake of brevity as generic reasoning. [Hawkins, 1977] p.122

C'est ainsi une généralité dans la *forme*, qui ne s'attache pas autant qu'elle le devrait aux cas particuliers. Ce que Kronecker qualifie de « fausse généralité », c'est une pratique algébrique du XVIII^e^ et du XIX^e^ siècle qui se focalise sur le cas « en général » ou « générique », mais qui ne s'occupe pas des valeurs problématiques que peuvent prendre certaines

81. [Kronecker, 1891a] p. 211 : *arithmetisch durchführen*
82. [Kronecker, 1881a] p. 164 : *rein analytisches*
83. *[Kronecker, 1891a] p. 232 : arithmetisch*
84. [Kronecker, 1869c] p. 185 : *algebraische Gebilde*
85. Voir aussi [Brechenmacher, 2006] et [Brechenmacher, 2008a].

variables ou certains paramètres [86] :

> Car on est habitué – en particulier dans les questions algébriques – à trouver des difficultés largement nouvelles, si on veut se détacher de la restriction à ces cas, que l'on a coutume de désigner comme généraux. Aussitôt que l'on perce la surface de la prétendue généralité, excluant chaque particularité, on pénètre l'intérieur de la vraie généralité qui embrasse toutes les singularités, et l'on trouve généralement ainsi seulement les difficultés réelles de l'étude, ainsi que les abondants nouveaux points de vue et phénomènes qu'elle contient dans ses profondeurs. ([Kronecker, 1874c] p.405 et [Brechenmacher, 2006] p. 57.)

La généralité de Kronecker est celle qui ne néglige pas les cas particuliers : il affirme que seule la prise en considération de ces cas limites nous permet de voir « la richesse des nouveaux points de vue et des phénomènes ». Mais ce n'est pas tout, la méthode qui permet de traiter les cas particuliers doit découler du cas qui est *véritablement général*.

Cette notion est très présente [87] dans les cours de Kronecker, et il met à chaque fois en avant l'extrême généralité des formules qu'il obtient, et qui lui permettent de déduire, en prenant certains cas particuliers, le théorème de Sturm, les formules de Sylvester, les déterminants de Cayley, ou même des résultats classiques sur les fractions continues.

Une grande partie de son travail sur le théorème de Sturm dépend de la formule :

$$\sum_{x_1<\xi<x_2} sgn\left(f'(\xi)f_1(\xi)\right) = -\frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_1)f_h(x_1)) + \frac{1}{2}\sum_{h=1}^{\nu} sgn(f_{h-1}(x_2)f_h(x_2))$$

qui est pour Kronecker la « généralisation du théorème initial de Sturm » [88]. Ensuite, lorsqu'il appliquera la formule précédente aux $\xi$ compris entre $-\infty$ et $+\infty$ pour compter l'*ensemble* des racines réelles, il montrera [89] que la formule obtenue permet de revenir à la formule initiale : Kronecker cherche à avancer en conservant la plus grande généralité possible.

De même, son travail sur les fractions continues lui permet d'aboutir à la formule :

$$\begin{aligned} &N(g_{p+1},...,g_{q-1})N(g_{r+1},...,g_{s-1}) - N(g_{p+1},...,g_{r-1})N(g_{q+1},...,g_{s-1}) \\ &+N(g_{p+1},...,g_{s-1})N(g_{q+1},...,g_{r-1}) = 0 \end{aligned}$$

86. Denn man ist es gewohnt – zumal in algebraischen Fragen – wesentlich neue Schwierigkeiten anzutreffen, wenn man sich von der Beschränkung auf diejenigen Fälle losmachen will, welche man als die allgemeinen zu bezeichnen pflegt. Sobald man von der Oberfläche der sogenannten, jede Besonderheit ausschliessenden Allgemeinheit in das Innere der wahren Allgemeinheit eindringt, welche alle Singularitäten mit umfasst, findet man in der Regel erst die eigentlichen Schwierigkeiten der Untersuchung, zugleich aber auch die Fülle neuer Gesichtpunkte und Erscheinungen, welche sie in ihren Tiefen enthält.

87. Par exemple, dans la seconde partie du cours du semestre d'hiver de 1880-81, des dérivées du mot *allgemein* apparaissent 33 fois (sous la forme *allgemein / Verallgemeinerung*).

88. *die Verallgemeinerung des ursprünglichen Sturm'schen Satzes* [Kronecker, 1887b] p. 139

89. Voir p. 220.

qu'il juge « plus générale », et « dans laquelle sont contenues, en tant que cas particulier, toutes les relations connues jusqu'alors entre fractions continues, et qui ne se trouvent pas dans les livres » [90]. Cette formule lui permettra, lorsqu'il cherchera à résoudre le *problème de Cauchy* – problème qui amène d'ailleurs à une généralisation de la formule d'interpolation de Lagrange -, d'améliorer les méthodes usuelles, qu'il qualifie de « "générales" en apparence » (*scheinbar „allgemeine"*) [91]. Enfin, la formule

$$|\sum_{k=0}^{n} u_k v_k \xi_k^g \eta_k^h|_{g,h=0,1,\ldots,m} = \sum_{(i)} \prod_{\alpha=i_0,\ldots,i_m} u_\alpha v_\alpha \prod_{\substack{\alpha,\beta=i_0,\ldots,i_m \\ \alpha<\beta}} (\xi_\alpha - \xi_\beta)(\eta_\alpha - \eta_\beta)$$

qu'il obtient à la fin de son travail sur Sturm est « beaucoup plus générale que toutes les formules de Sylvester, Sturm, Cauchy, Jacobi » [92] car ces dernières n'en sont que des cas particuliers.

Ces formes sont emblématiques de ce qu'entend Kronecker par *généralisation* : il exhibe une *formule* dont les *cas particuliers*, c'est-à-dire – et c'est très important – les cas obtenus par substitution d'objets particuliers aux variables de cette formule, seront le théorème de Sturm, des formules connues sur les fractions continues, la formule d'interpolation de Lagrange, etc. Cette idée qui consiste à considérer dans des formules certains éléments comme des variables est d'ailleurs au cœur de ses recherches sur les *Gattungen* et représente pour Kronecker la forme de la *vraie* généralisation.

Pour Kronecker, plus qu'un travail fait à partir de notions « générales », le *général* est plutôt une action : c'est la *généralisation*. Dans son travail sur le théorème de Sturm, il poursuit l'algébrisation entamée en 1839 par Sylvester. Il expose ainsi dans son cours la plupart des travaux de ses prédécesseurs en tentant de les généraliser, et lorsqu'il le peut, sous la forme d'une *formule*. Mais Kronecker l'affirme à propos du théorème de Sturm : « Les difficultés de la généralisation dont M. *Sylvester* fait mention doivent en grande partie être dues au fait qu'il s'est limité aux formes algébriques qu'il avait établies » [93] : pour étendre son travail, pour le *généraliser*, il doit changer de point de vue et apporter des outils qui n'entrent plus dans l'« algèbre », qui sont hors de son « arithmétique générale ». Que dire alors de l'assertion de E.T. Bell affirmant que « Kronecker n'avait aucune sensibilité pour la géométrie » [94] ?

Tout au long de notre lecture des *Vorlesungen über die Theorie der algebraischen Gleichungen*, nous avons suivi un mathématicien qui utilise l'intuition géométrique de ses objets algébriques, et qui loin d'« interdire », multiplie les approches des objets qu'il

90. [Kronecker, 1887b] p. 183 : « Man kann aber noch eine viel allgemeinere Kettenbruchformel ableiten, in welcher alle bisher bekannten Kettenbruchrelationen als Specialfälle enthalten sind, und die sich nicht in den Lehrbüchen findet. »

91. [Kronecker, 1887b] p. 189

92. [Kronecker, 1885b] p. 280 : « Diese Formel ist viel allgemeiner als alle die Formeln von Sylvester, Sturm, Cauchy, Jacobi. »

93. Voir p. 396.

94. [Bell, 1953b] p. 517 : « Kronecker had no feeling for geometry. »

présente, les démonstrations des théorèmes qu'il énonce, tout en présentant « en contexte » ses idées sur la nature des nombres, la continuité, la divisibilité...

Ainsi, la formule suivante, écrite par Poincaré, peut s'appliquer aux méthodes utilisées par Kronecker dans ses cours :

> Mais pourquoi, dira-t-on, ne pas conserver le langage analytique et le remplacer par un langage géométrique, qui perd tous ses avantages dès que les sens ne peuvent plus intervenir. C'est que ce langage nouveau est plus concis ; c'est ensuite que l'analogie avec la Géométrie ordinaire peut créer des associations d'idées fécondes et suggérer des généralisations utiles. ([Poincaré, 1895] p. 1)

L'utilisation de cette fonction heuristique de la géométrie, partagée par d'autres mathématiciens de la seconde moitié du XIX^e siècle [95] est très présente [96] dans le cours de Kronecker sur la théorie des équations algébriques, en particulier dans la partie qu'il réserve au théorème de Sturm et à la théorie des caractéristiques. Cette « porosité » de l'algèbre lui permet de proposer une autre façon de produire des mathématiques *plus générales* : il multiplie les approches pour obtenir une compréhension plus globale de l'objet étudié. Lorsqu'il construit sa théorie des caractéristiques, c'est la somme des points de vue topologique, algébrique et analytique qui éclaire son objet, et cet objet sera, selon l'angle choisi, tour à tour équation, système d'équations, et pour finir il pourra être considéré comme une variété définie par des équations [97]. Et c'est sous la forme d'une intégrale sur des variétés que cette théorie connaîtra une certaine postérité.

## Postérité de la théorie

On peut voir le degré topologique, et les théories qui lui sont liées, comme des outils donnant des informations sur l'existence de solutions à l'équation

$$f(x) = y.$$

95. Par exemple, dans son travail de thèse sur Hermann Weyl, Christophe Eckes montre que Dedekind, lorsqu'il construit sa fonction *valence*, très abstraite, s'appuie sur une vision géométrique de cet objet :

> Ainsi, Dedekind s'appuie sur un raisonnement purement géométrique pour construire la fonction $J$ dont Klein donnera une formulation analytique après coup. Les représentations géométriques utilisées ici par Dedekind ont une fonction indéniablement heuristique, puisqu'elles permettent de construire un objet mathématique, à savoir la fonction $J$. Ce passage montre qu'il est inexact de voir en Dedekind le promoteur d'une mathématique abstraite qui ne reposerait sur aucune visée intuitive. En réalité, il demeure attaché à une pratique que Gauss illustre de manière exemplaire dans sa théorie des formes quadratiques et qui consiste à s'appuyer sur des représentations géométriques pour appréhender intuitivement certains objets et concepts relevant de l'« arithmétique ». Ainsi, l'écrit à Borchardt souligne l'inscription des travaux de Dedekind dans des traditions et des pratiques qui doivent nous prévenir contre toute tendance à ne voir dans son œuvre qu'une préfiguration des mathématiques structurales. ([Eckes, 2011] p.113)

96. Voir p. 176.

97. Et donc, dans le langage du XX^e siècle, une sous-variété.

Bien sûr, les espaces dans lesquels vivent $x$ et $y$ ainsi que la nature de la fonction $f$ et des éléments que l'on tente de trouver sont à préciser. Par exemple, si l'on choisit de prendre $y = x$, on est conduit vers des problèmes de points fixes et donc vers les travaux de Brouwer.

Avec $y = 0$ et $f$ un polynôme, on se retrouve dans le cadre de la recherche d'informations sur les racines, et ainsi du théorème fondamental de l'algèbre ou du théorème de Sturm. Les travaux de Gauss, comme nous l'avons vu, sont une étape importante dans ce domaine.

Si maintenant $y = 0$, $x \in \mathbb{R}^n$ et $f$ est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^n$, on est cette fois dans le cadre de la théorie des caractéristiques de Kronecker. Ces distinctions *a posteriori*, si elles ne nous permettent aucune interprétation historique, nous amènent tout de même à nous poser la question de l'existence ou non d'une filiation entre ces théories. Nous avons déjà vu que la théorie des caractéristiques, qui peut être interprétée comme une extension du théorème de Sturm à des espaces de dimensions supérieures, est en partie issue d'un mouvement de *généralisation* du travail de Sturm. Il nous reste maintenant à traiter la question de la postérité de la théorie de Kronecker.

Comme le fait remarquer Hadamard dans une note ajoutée au second tome de *la théorie des fonctions d'une variable* de Jules Tannery [98] :

> La démonstration, d'après M. Ames, du théorème de M. Jordan sur les courbes fermées sans point double (n ° 306, 307) repose sur la considération de l'ordre d'un point ou, si l'on veut, sur la considération d'une variation d'argument. La généralisation, dans le cas où le nombre des dimensions dépasse deux, est fournie par l'indice de Kronecker. C'est là une notion qui est maintenant classique (Surtout depuis le *Traité d'Analyse* de M Picard (T. I, p.123 ; T. II, p. 193).). Elle a reçu, dans plusieurs travaux contemporains, des applications nouvelles. [99]

Cette note, que Hadamard intitule *Note sur quelques applications de l'indice de Kronecker*, est intéressante à plusieurs titres. Elle permet d'abord de donner une indication sur la circulation de la théorie des caractéristiques au début du XX$^{e}$ siècle, au moins en France : elle est connue et le vecteur essentiel de sa diffusion est le cours de Picard (1901). Hadamard décrit ensuite deux grands types d'application : la recherche du nombre de solutions d'un système de $n$ équations à $n$ inconnues, comme Picard l'a traitée, et les développements de l'*Analysis Situs* de Poincaré.

La question que Picard se propose de résoudre est suffisamment étroite pour lui permettre d'y apporter une réponse complète. Examinons maintenant la familiarité de Poincaré avec la théorie des caractéristiques et l'utilisation qu'il en fait dans son travail.

Nous avons vu que Kronecker présentait sa théorie sous deux formes : la première, que l'on a qualifiée de « topologico-algébrique », est celle qui consiste à examiner les signes que prend un determinant fonctionnel pour certaines valeurs. Cette exposition utilise les

98. Voir [Tannery et Hadamard, 1910] p. 437.
99. *Ibid.*

points d'intersection entre courbes algébriques, et finalement, comme le dit Poincaré de son *Analysis Situs*, « c'est là qu'intervient vraiment l'intuition géométrique. » [100]. C'est aussi dans ce contexte que Kronecker utilise la *proposition problématique* [101] qui repose sur l'affirmation de Gauss : « si une branche d'une courbe algébrique entre dans un espace fini, cette même branche doit nécessairement sortir quelque part de cet espace ». La validité de cette proposition n'est pas discutée par Kronecker. Mais finalement, en donnant une seconde exposition de la caractéristique d'un système, qui cette fois sera construite *analytiquement* sur les bases de la théorie du potentiel, et en montrant le passage de l'une à l'autre [102], il met en quelque sorte le problème de côté. Ses contemporains n'ont pas émis de réserve sur les justifications que Kronecker a fournies pour construire sa théorie, mais peut-être est-ce dû au fait que seule sa seconde exposition, qui est celle que présente Picard, a eu un réel auditoire.

Comme le montre Jean Mawhin dans son article *Poincaré's early use of Analysis situs in nonlinear differential equations : Variations around the theme of Kronecker's integral* [103], Poincaré utilise dès ses premiers travaux la théorie des caractéristiques. Bien sûr Poincaré, en tant qu'élève d'Hermite à l'École Polytechnique, est familier du théorème de Sturm et de ses premiers prolongements. Plus encore, dans une lettre datée du 27 novembre 1880, Hermite l'encourage à lire les articles de Kronecker publiés dans les *Monatsberichte* de l'Académie des Sciences de Berlin [104]. Kronecker lui-même conseille à Poincaré en 1883 de lire son travail sur la théorie des caractéristiques :

> Ayant lu votre dernière communication dans les Comptes Rendus je désirerais appeler votre attention à un mémoire que j'ai publié en 1869 et que je prends la liberté de vous envoyer. J'y ai ajouté quelques autres de mes mémoires et en outre je me suis permis d'adresser à Mr Tannery un petit paquet destiné pour vous, contenant tous les exemplaires de mes mémoires, que j 'ai à ma disposition. Le mémoire cité de 1869 (mois de mars) est intitulé : *Systèmes de fonctions de plusieurs variables.* J'y ai développé la généralisation de cet important théorème de Cauchy, qui me semble contenir le vrai fondement de la théorie des fonctions. [Poincaré, 1989] p. 146

Dans son *Mémoire sur les courbes définies par une équation différentielle* publié en 1881, Poincaré étudie des cycles sur la surface d'une sphère, et introduit l'indice d'un cycle, dont la définition est presque identique à celle que Cauchy donne en 1831. En 1883 dans son article sur les fonctions $\Theta$, il écrit :

> Pour le démontrer, on se sert de la formule de Cauchy qui sert à calculer le nombre des racines d'une équation qui sont intérieures à un contour donné, et l'on applique cette formule au parallélogramme des périodes. Depuis, M. Kronecker a généralisé la formule de Cauchy de manière à exprimer sous forme d'une intégrale multiple définie le

100. Voir [Poincaré, 1912a] p. 484.
101. Voir p. 376.
102. Tout au moins dans ses cours pour la dimension deux.
103. Voir [Mawhin, 2000].
104. Voir [Poincaré, 1986] p. 170.

> nombre des solutions communes à un système de $n$ équations à $n$ inconnues, lorsqu'on assujettit ces solutions à être comprises dans un domaine donné. [Poincaré, 1883] p.130

Il s'agit, d'après Hadamard [105], de la première fois où Poincaré utilise ce travail de Kronecker. Il redémontre et utilise l'invariance de la caractéristique – celle-ci étant définie sous la forme de ce qu'il appelle « l'intégrale de Kronecker » – lorsque les équations, supposées dépendre de certains paramètres, varient continument. On est donc en présence d'une invariance par homotopie, que Kronecker avait déjà montrée et utilisée dans son article de 1878. Dans son article de 1884, *Sur certaines solutions particulières du problème des trois corps* [106] Poincaré exploite encore un peu plus cette intégrale et en étudie certaines propriétés :

> M. Kronecker a présenté à l'Académie de Berlin, en 1869, un Mémoire sur les fonctions de plusieurs variables ; on y trouve un important théorème d'où il est aisé de déduire le résultat suivant : Soient $\xi_1, \xi_2, ..., \xi_n$ $n$ fonctions continues de $n$ variables $x_1, x_2, ..., x_n$ ; la variable $x_i$ est assujettie à varier entre les limites $+a_i$ et $-a_i$. Supposons que, pour $x_i = a_i$, $\xi_i$ soit constamment positif, et pour $x_i = -a_i$ constamment négatif ; je dis qu'il existera un système de valeurs des $x$ pour lequel tous les $\xi$ s'annuleront.

Poincaré obtient ainsi une version en dimension $n$ du théorème des valeurs intermédiaires. On est donc très loin de l'esprit dans lequel Kronecker, pour qui le théorème de Bolzano était une aberration, a construit l'intégrale qu'utilise Poincaré. En 1885, dans un article intitulé *Sur l'équilibre d'une masse fluide animée d'un mouvement de rotation* [107], et en 1886 dans *Sur les courbes définies par les équations différentielles* [108] Poincaré emploie de nouveau l'intégrale de Kronecker et son invariance par des transformations continues. Sans continuer l'examen détaillé de la présence de la théorie des caractéristiques dans l'œuvre de Poincaré, on peut tout de même noter qu'elle est encore présente dans un de ses derniers articles *Sur un théorème de géométrie* [109]. Hadamard, qui a beaucoup commenté l'œuvre de Poincaré, a aussi utilisé dans ses travaux l'indice de Kronecker, en particulier dans sa *note sur quelques applications de l'indice de Kronecker*, qu'il conclut en redémontrant le théorème du point fixe de Brouwer. Le degré topologique de ce dernier est d'ailleurs bien une extension et une généralisation de l'indice de Kronecker, et on peut se demander s'il y a eu une transmission des idées de Kronecker à Brouwer.

Dans une lettre de 1909, Hadamard conseille à Brouwer la lecture des articles de Poincaré. Hadamard et Brouwer se rencontrent l'année suivante, année durant laquelle Brouwer formule sa définition du degré d'une application continue sur des variétés. Si on peut lire parfois que Brouwer reconnaissait l'influence de Kronecker sur son travail sur le degré topologique [110], ces affirmations sont faites sans référence. Au contraire, si l'on en croit Dirk

105. Voir [Hadamard, 1921].
106. Voir [Poincaré, 1884] p.251.
107. Voir [Poincaré, 1885] p. 269.
108. Voir [Poincaré, 1886] p. 177.
109. Voir [Poincaré, 1912b].
110. Par exemple dans [Gauthier, 2002] p.147 : « Brouwer acknowledges that he has been inspired by Kronecker's work dating back in 1869 on systems of functions of several variables (...) where Kronecker defines the notions of index and characteristic for functional systems and curved surfaces (or closed curves

van Dalen, dont l'avis repose sur une étude consciencieuse des textes de Brouwer

> the influence of Kronecker on Brouwer is negligible. Neither in Brouwer's dissertation nor in his inaugural address is Kronecker mentioned. In the Berlin Lectures there is a reference to Kronecker in the historic introduction, but in the Vienna Lectures his name does not appear. ([van Dalen, 1999])

Cependant, si l'influence de Kronecker n'est pas directe, celle de Poincaré est évidente. Il est fort possible que Brouwer ait eu connaissance du travail de Kronecker, et il est certain qu'il connaissait l'utilisation que faisait Poincaré de la théorie des caractéristiques. C'est donc par le biais de l'« Analysis Situs » de Poincaré que la théorie des caractéristiques de Kronecker se prolongera dans le travail de Brouwer.

in the plane). »

# Annexes

# Annexe A

# Transcription et traduction partielles du cours de 1880-1881

Dans cette annexe, nous allons donner la transcription et la traduction de certains passages du manuscrit du cours de Kronecker du semestre d'hiver 1880-1881 ([Kronecker, 1881a]) intitulé *Theorie der algebraischen Gleichungen*, et dont l'original se trouve dans le fond *Hensel* de la bibliothèque de Strasbourg [1].

Ce manuscrit est écrit en *Kurrentschrift*. Nous avons identifié les mots figurant en caractère latin en les notant en italique. Chaque début et fin de page sont représentés par le symbole "|" et on trouvera dans la marge la pagination correspondante. Dans ce manuscrit, aucun numéro de page n'est indiqué : nous avons choisi de faire débuter la numérotation à la première page rédigée, c'est-à-dire la troisième page du PDF mis à disposition sur le site de la bibliothèque de Strasbourg. Toutes les illustrations figurant dans le manuscrit ont été insérées dans cette retranscription.

Très souvent, on trouve dans le manuscrit des renvois, que ce soit sous la forme de notes en bas de page ou dans le corps du texte, à d'autre parties du manuscrits. Ses renvois n'ont pas été complétés par le rédacteur, et nous trouverons donc, comme cela est le cas dans le manuscrit, de nombreux « siehe ... ». Dans le cadre d'une transcription complète du manuscrit, il serait intéressant, et nécessaire, d'exhiber les passages auxquels ils se réfèrent.

1. Une version numérique est consultable à l'adresse `http://docnum.u-strasbg.fr/cdm/compoundobject/collection/coll7/id/54535/rec/20`

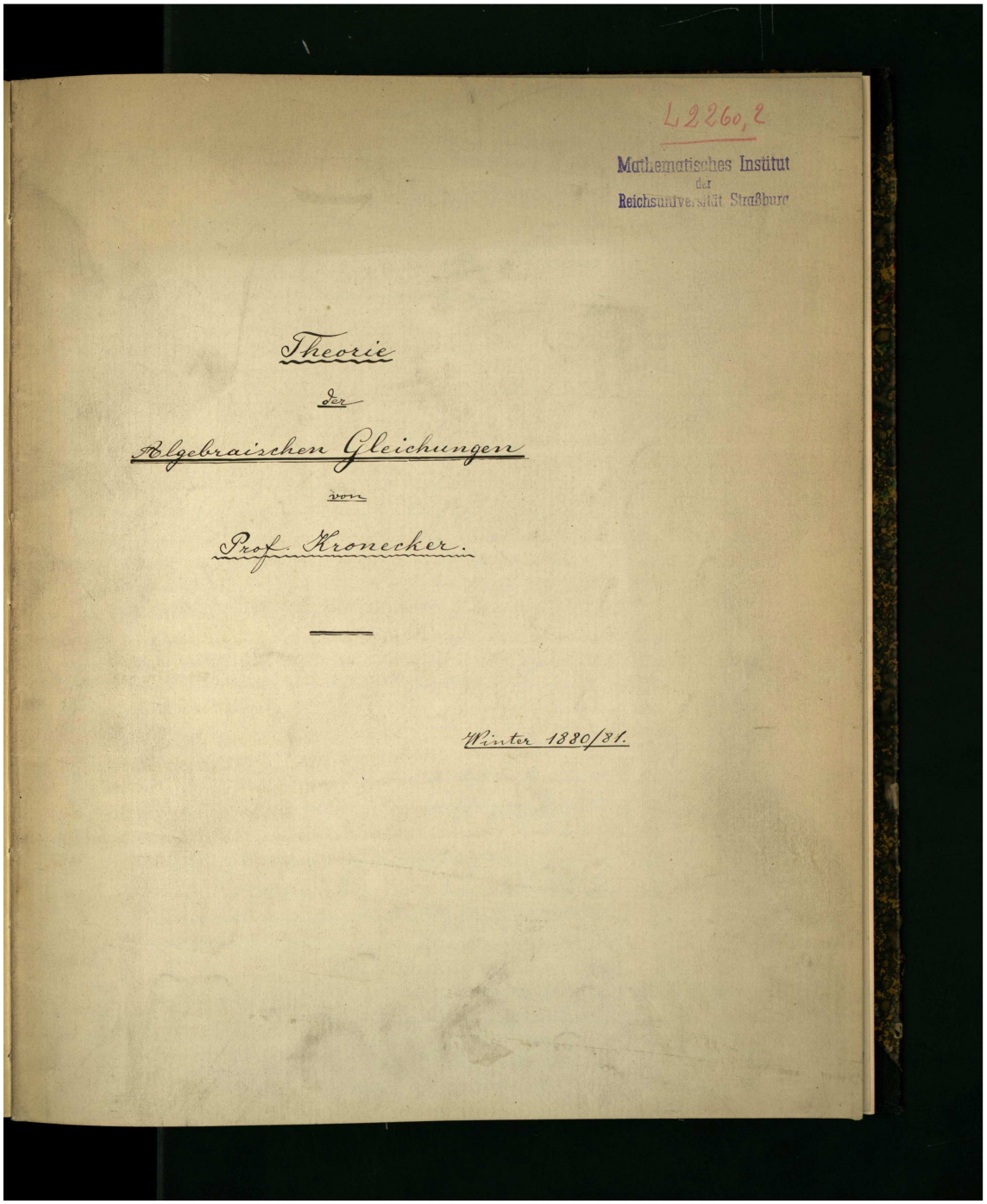
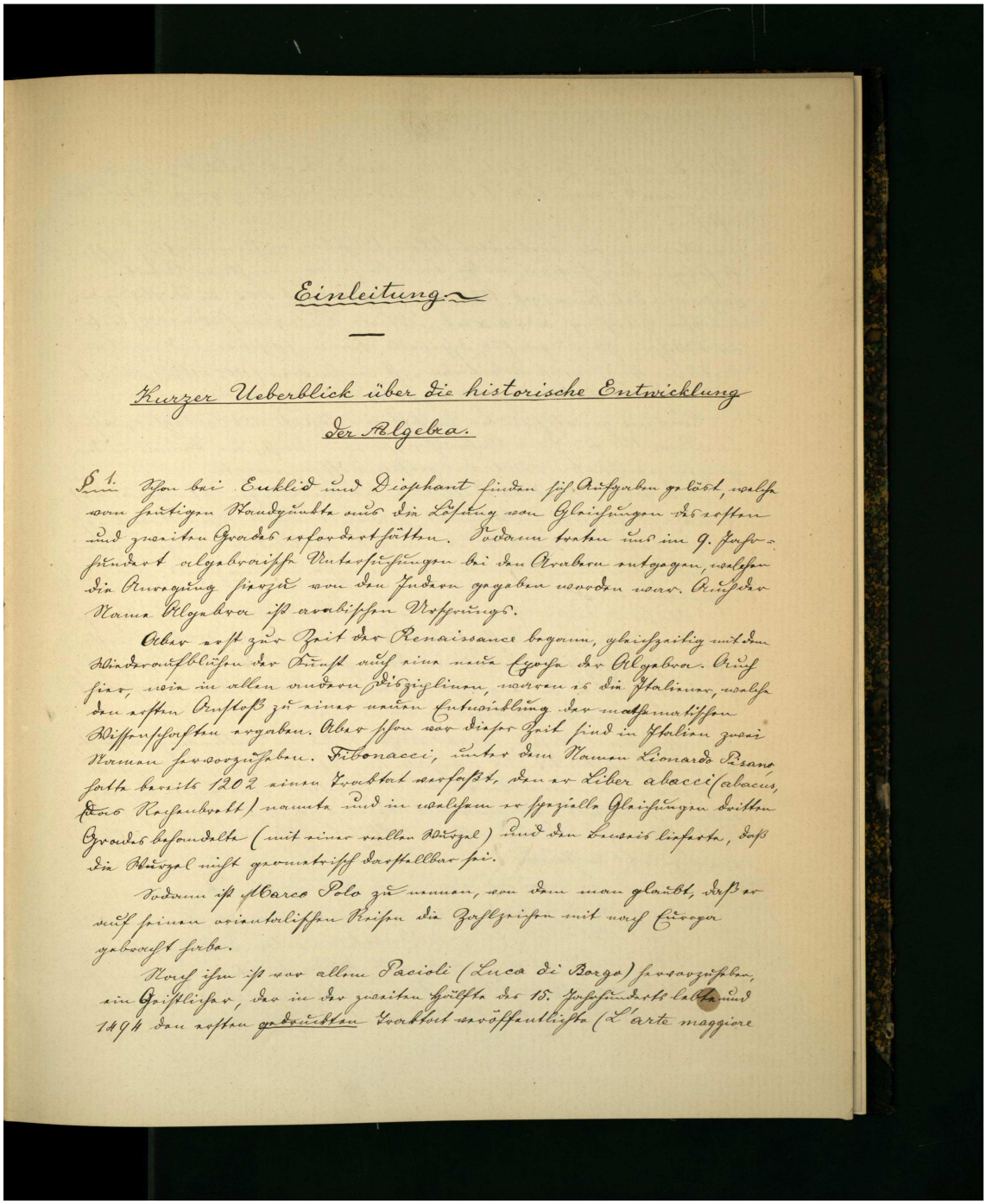



FIGURE 1 – Page de garde du manuscrit de 1880-81 ([Kronecker, 1881a])

## A.1 Transcription de l'introduction



Einleitung.

Kurzer Ueberblick über die historische Entwicklung der Algebra.

§ 1. Schon bei Euklid und Diophant finden sich Aufgaben gelöst, welche vom fertigen Standpunkte aus die Lösung von Gleichungen des ersten und zweiten Grades erfordert hätten. Sodann traten uns im 9. Jahrhundert algebraische Untersuchungen bei den Arabern entgegen, welchen die Anregung hierzu von den Indern gegeben worden war. Ersterer Name Algebra ist arabischen Ursprungs.

Aber erst zur Zeit der Renaissance begann, gleichzeitig mit dem Wiedererblühen der Kunst auch eine neue Epoche der Algebra. Auch hier, wie in allen andern diesbezüglichen, waren es die Italiener, welche den ersten Anstoß zu einer neuen Entwicklung der mathematischen Wissenschaften gaben. Aber schon vor dieser Zeit sind in Italien zwei Namen hervorzuheben. Fibonacci, unter dem Namen Leonardo Pisano selbst bereits 1202 einen Traktat verfaßt, den er Liber abacci (abacus, das Rechenbrett) nannte und in welchem er spezielle Gleichungen dritten Grades besonders (mit einer neuen Methode) und den Satz lieferte, daß die Wurzel nicht geometrisch darstellbar sei.

Sodann ist Marco Polo zu nennen, von dem man glaubt, daß er auf seinen orientalischen Reisen die Zahlzeichen mit nach Europa gebracht habe.

Noch ihm ist vor allem Pacioli (Luca di Borgo) hervorzuheben, ein Geistlicher, der in der zweiten Hälfte des 15. Jahrhunderts lebte und 1494 den ersten gedruckten Traktat veröffentlichte (L'arte maggiore

FIGURE 2 – Première page du manuscrit de 1880-81 ([Kronecker, 1881a])

# Einleitung

## Kurzer Ueberblick über die historische Entwicklung der Algebra

§1. Schon bei *Euklid* und *Diophant* finden sich Aufgaben gelöst, welche vom heutigen Standpunkte aus die Lösung von Gleichungen des ersten und zweiten Grades erfordert hätten. Sodann treten uns im 9. Jahrhundert algebraische Untersuchungen bei den Arabern entgegen, welchen die Anregung hierzu von den Indern gegeben worden war. Auch der Name Algebra ist arabischen Ursprungs.

Aber erst zur Zeit der *Renaissance* begann, gleichzeitig mit dem Wiederaufblühen der Kunst auch eine neue Epoche der Algebra. Auch hier, wie in allen andern Disziplinen, waren es die Italiener, welche den ersten Anstoß zu einer neuer Entwicklung der mathematischen Wissenschaften ergaben. Aber schon vor dieser Zeit sind in Italien zwei Namen hervorzuheben. *Fibonacci*, unter dem Namen *Lionardo Pisano*, hatte bereits 1202 einen Traktat verfaßt, den er *Liber abacci* [abaci] (*abacus*, das Rechenbrett) nannte und in welchem er spezielle Gleichungen dritten Grades behandelte (mit einer reellen Wurzel) und den Beweis lieferte, daß die Wurzel nicht geometrisch darstellbar sei.

Sodann ist *Marco Polo* zu nennen, von den man glaubt, daß er auf seinen orientalischen Reisen die Zahlzeichen mit nach *Europa* gebracht habe.

Nach ihm ist vor allem *Pacioli* (*Luca di Borgo*) hervorzuheben, ein Geistlicher, der in der zweiten Hälfte des 15. Jahrhunderts lebte und 1494 den ersten gedruckten
1|2 Traktat veröffentlichte (*L'arte maggiore | ditta da vulgo la regulo della cosa*). Der Ausdruck „*regula della cosa*“ (*cosa* bedeutet die Unbekannte) bleibt im ganzen Mittelalter bestehen.

Von nun an verbreitet sich die Algebra außerordentlich rasch. Es folgen kurz hintereinander eine Reihe von wichtigen Arbeiten. *Scipione del Ferro* (gest. 1524 in *Bologna*) gab 1505 die Auflösung der kubischen Gleichung $x^3 + ax = b$, *Nicolo Tartaglia* (1506-1559) löste die kubischen Gleichung für 3 spezielle Formen (positive und negative Coeffizienten), welche Arbeit aber erst von *Cardano* (1501-1576) publiziert und auf alle Fälle ausgedehnt wurde.

*Cardano* gerieth [geriet] hierüber in heftigen Streit mit *Tartaglia*, welcher die Priorität für sich in Anspruch nahm, ein Zwist, der damals in der ganzen zivilisirten Welt das größte Aufsehen erregte.

Nun folgte sofort die Auflösung der Gleichungen 4 ten Grades durch *Ferrari* (1522-1566 in *Bologna*). Publizirt wurde dessen Arbeit erst durch *Rafaello Bombelli*, der zuerst die Realität der Wurzeln einer kubischen Gleichung erkannte, auch wenn sie in imaginärer Form erscheinen.

Während dieser Zeit waren auch in Deutschland einige tüchtige Arbeiten erschienen. *Christoph Rudolfs* veröffentlichte 1524 die erste algebraische Schrift in Deutschland. Von ihm rührt das Wurzelzeichen her ; gleichzeitig führte der englische Algebraiker *Robert Ricord* (gest. 1558 in *London*) das Gleichheitszeichen ein.

*Michael Stifel*, welcher *Luther* 1523 in *Wittenberg* aufsuchte, um zum Protestantismus überzutreten, schrieb eine *Arithmetica integra* mit einer Vorrede von *Melanchthon*. Das dritte Buch derselben ist betitelt : *De numeris cossicis* und hier findet sich die erste Idee zu den Logarithmen. *Stifel* lebte von 1487-1567 und war zuletzt in *Iena*.

Von nun an wird die algebraische Wissenschaft international.

Die nächste bedeutende Erscheinung ist der Franzose *François Viète* (*Vieta*) von 1540-1603. Er war Advokat in *Poitou* und begleitete später *Henri IV.* nach *Paris*. Er war der erste, der die Buchstabenrechnung einführte und mit Consequenz ausbildete.

So bezeichnete er z.B. die bekannten Größen durch Consonanten, die Unbekannten durch Vokale. Von ihm rühren | die Ausdrücke „*potestas pura*" (Potenz) und „*potestas affecta*" (ganze Funktion) her, Ausdrücke, die sich bis zu *Gauss* erhielten. *Viète* war es auch, der als das Fundament der Theorie der ganzen Funktionen die Zerlegung derselben in Linearfaktoren erkannte. Dagegen vermied er noch die negativen Wurzeln und war noch unbekannt mit den Zeichen der Multiplikation und der Gleichheit. 2|3

*Thomas Harriot* (geb. 1560 in *Oxford*, gest. 1621 in *London*) war eigentlich Geometer und begleitete in seiner Eigenschaft als Vermesser den bekannten *Walter Raleigh* nach *Virginien* (1585). Das von ihm stammende Werk : *Artis analyticae praxis ad aequationes algebraicas resolvendas* wurde erst nach seinem Tode veröffentlicht.

*Albert Girard* (gest. 1633 in *Holland*) veröffentlichte 1629 eine Algebra, in der er die sog. *Newton*'schen Formeln begründete.

Nun folgt der große *Descartes* (1596-1650). Von 1624-1649 in Holland, wo er seine mathematischen und phylosophischen [sic] Werke publizirte. Von ihm rührt die Bestimmung der Anzahl der positiven und negativen Wurzeln (Cartes. Zeichenreg.) und vor allem die Einführung der kontinuirlichen Größen in die Algebra durch die große Entdeckung der analytischen Geometrie.

*Fermat* (1608-1665), Parlamentsrath in *Toulouse*. Von ihm stammt die sog. Reduktion der Gl., d. h. die Umformung derselben, wenn irrationale Größen darin vorkommen. Viel wichtiger aber ist sein indirekter Einfluß auf die Algebra durch seine zahlentheoretischen Untersuchungen.

*Newton* (1642-1726) *Arithmetica universalis*. Regel über die Anzahl der reellen Wurzeln. *Newton*'sche Formeln. Näherungsmethode zur Auflösung der Gleichungen.

Graf *Tschirnhausen* (geb. 1651, gest. 1708 in Kieslingswalde bei Görlitz). *Tschirnhausen*'s Methode zur Lösung der Gleichungen vierten Grades. 1682 wurde *Tsch.* auswärtiges Mitglied der Pariser Akademie.

*Euler* ( geb. 1707 in Basel, gest. 1783 in St. Petersburg). Euler war von 1741-1766 in Berlin. Von ihm rührt eine neue Auflösung der Gl. vierten Grades. Hervorzuheben ist seine Algebra (einem jungen Schneider von dem erblindeten Euler diktirt), 1774 ins Französische übersetzt und von Lagrange mit den kostbarsten Noten versehen.

*Bezout* (1730-1783). Von ihm rührt fast Alles, was man heute unter dem Namen Elimination und Theorie der symmetrischen Funktionen begreift.

| *Vandermonde* (1735-1796), Mitglied der Pariser Akademie, und *Lagrange* (1736-1813, geb. *Turin*, gest *Paris*), veröffentlichen fast gleichzeitig (1771 und 1772) ihre algebraischen Arbeiten, jener in der Pariser, dieser in der Berliner Akademie. *Lagrange* war von 1766-1787 in Berlin. Die Bedeutung dieser Arbeiten tritt aber erst durch die Forschungen von *Gauss* (1777-1855) zu Tage. *Gauss* promovirte 1799 zu Helmstädt und veröffentliche 1801 seine *Disquisitiones arithmeticae*. Die siebente Sektion enthält die Untersuchungen über die Kreistheilung, die größte mathematische Entdeckung seit dem Alterthum. Von *Euklid* bis *Gauss* war die lockende Aufgabe der Kreistheilung um kein Jota gefördert worden. Erst *Gauss* war es gegeben, den 2000 Jahre alten Zauber zu lösen. *Gauss* gibt als Datum seiner Entdeckung den 30. März 1796 an. Der hohe Werth der *Gauss*'schen Entdeckung ist das Beispiel der Auflösbarkeit von Gleichungen aller Grade. Seine Entdeckung machte auch klar, warum die Kreistheilung keine raschen Fortschritte gemacht hatte. Geometrisch (d. h. durch Zirkel und Lineal) konstruiren lassen sich nämlich nur die durch $2^{2^\lambda}+1$ definirten Theilungen des Kreises in $3(\lambda=0)$, $5(\lambda=1)$, $17(\lambda=2)$, $257(\lambda=3)$, etc. gleiche Theile. Die Theilung in 5 gleich Theile kannte schon *Euklid*. Erst *Vandermonde* untersuchte die Theilung in 7, 11, 13 Theile und begann grade mit der Theilung in 17 Theile, da erlahmte seine Kraft und er ließ die Untersuchung liegen. Nur ein kleiner Schritt hätte gefehlt und *Vandermonde* hätte *Gauss*' Entdeckung vorweggenommen. Gauss bezeichnet als das Wesen der Auflösung der Gleichungen die : *Reductio aequationum affectarum ad aequationes puras*. 3|4

Bereits in seiner Dissertation, in der er den ersten strengen Beweis der Wurzelexistenz gibt, spricht er die Vermuthung aus, daß Gleichungen von höherem als dem vierten Grade nicht mehr auflösbar seien. Im Anfange seiner Dissertation gibt er eine

Kritik der bisher geführten Beweise der Wurzelexistenz und zeigt namentlich, daß der von *Euler* herrührende Beweis unhaltbar ist, da der letztere geführt wird unter der Voraussetzung, daß alle Gleichungen auflösbar seien.

Dieses Vorurtheil, daß alle Gleichungen auflösbar seien, war vor *Gauss* allgemein verbreitet. *Gauss* fing an daran zu zweifeln, aber erst *Abel* wies die Unrichtigkeit dieses Vorurtheils strenge nach. Das zweite Vorurtheil aber, in dem alle Mathematiker
4|5 bis *Gauss* befangen | waren, daß nämlich jede Gleichung $n$ ten Grades $n$ Wurzeln habe, wurde zuerst durch die Dissertation von *Gauss* als Vorurtheil beseitigt, denn er lieferte den ersten strengen Beweis dieses Satzes, den man als selbstverständlich angesehen hatte. Er ist in der That nichts weniger als selbstverständlich, daß jene 2 Curven $n$ ten Grades, die in dem *Gauss*'schen Beweise auftreten, sich in $n$ reellen Punkten schneiden.

*Abel* (1802-1829) gab 1826 im 1. Band von *Crelle's Journal* den Beweis für die Unmöglichkeit der Auflösung von Gleichungen von höherem als dem 4. Grade. Im 4. Band von *Crelle* begründet er die Theorie der Gleichungen, welche von *Kronecker* als *Abel*'she Gleichungen bezeichnet wurden. Nicht minder hervorzuheben ist *Abel*'s indirekter Einfluß auf die Algebra durch seine funktionentheoretischen Arbeiten.

Bereits in den Jahren 1798 und 1806 hatte der Italiener *Ruffini* (1765-1822) einen Beweis für die Unmöglichkeit der Aufl. v. Gl. von höherem als dem 4. Grade versucht, doch sind seine Beweise nicht streng.

*Evariste Galois* (1811-1832) am 26. Okt. im Duell gefallen, einer der größten Algebraiker aller Zeiten, veröffentlichte in *Liouville's Journal* seine bedeutenden Arbeiten über algebraische Gleichungen, in denen er, von andern Gesichtspunkten als *Abel* ausgehend, vielfach den *Abel*'shen Arbeiten begegnete, theils dieselben erweiterte.

*Fourier* (1768-1830). Seit 1817 Sekretär der Akademie. Er stand an der Spitze der Akademie in Aegypten. Von ihm stammen Regeln zur Bestimmung der Anzahl der Wurzeln.

Die vollständige Lösung dieses Problems : die Anzahl der reellen Wurzeln einer algebr. Gl. zwischen gegebenen Grenzen zu bestimmen, gelang aber erst 1829 dem Genfer *Charles Sturm* (1803-1855). Er lebte größtentheils in *Paris*, wo er in engem Verkehr mit *Dirichlet* stand.

Blicken wir zurück. Zwei Probleme sind es vorzugsweise, die 2000 Jahre lang die Algebraiker beschäftigten :

Eine theoretische, die Auflösbarkeit der Gleichungen betreffend, wurde 1826 von *Abel* in negativem Sinne entschieden ;

eine praktische, die Anzahl der reellen Wurzeln zwischen zwei gegebenen Grenzen betreffend, wurde 3 Jahre später (1829) von *Charles Sturm* beantwortet.

Die Aufgabe, die *Abel*'sche Untersuchung zu einer positiven zu machen, d. h. die
5|6 Frage zu stellen nach den auflösbaren Gleichungen, | ist der Gegenstand der neueren Untersuchungen, zu denen die Arbeiten von *Gauss*, *Abel* und *Galois* den Grund gelegt haben.

## §2. Perspectiven in die Theorie der algebr. Gleichungen

Um zu ziegen welche Geschichtspunkte in der Theorie der Gleichungen ins Auge zu fassen sind, wollen wir mit wenigen Worten auf die Theorie der quadratischen und kubischen Gleichungen zu sprechen kommen. Es sei zunächst die quadratische Gl. gegeben :

$$x^2 - f_1 x + f_2$$

Die *Reductio ad aequationem puram* ist geleistet durch die Form :

$$(2x - f_1)^2 + (4f_2 - f_1^2) = 0,$$

welche sich von der *aequatio affecta* nur durch eine andere Zusammenfassung unterscheidet. Bezeichnet man mit $x_1$ und $x_2$ die Wurzeln der quadratischen Gleichung, so ist :

$$x^2 - f_1 x + f_2 = (x - x_1)(x - x_2)$$

also : $f_1 = x_1 + x_2$ und $f_2 = x_1 x_2$.

Setzt man dies für $f_1$ und $f_2$ in der *aequatio pura* $2x = f_1 + \sqrt{4f_2 - f_1^2}$ ein, so erhält man

$$\begin{aligned} 2x = &\quad x_1 + x_2 + \sqrt{(x_1 - x_2)^2} \\ 2x = &\quad x_1 + x_2 \pm (x_1 - x_2) = \begin{cases} 2x_1 \\ 2x_2 \end{cases} \end{aligned}$$

Wir erkennen also : das Wesen der Auflösung der quadratischen Gleichung $x^2 - f_1 x + f_2 = 0$ besteht in der Herstellung eines zweiwerthigen Ausdrucks, der nur eine Quadratwurzel enthält und in welchem nur symmetrische Funktionen von $x_1$ und $x_2$ vorkommen.

Nehmen wir die kubische Gleichung :

$$0 = x^3 - f_1 x^2 + f_2 x - f_3.$$

Schon in der alten Algebra war die Wegschaffung des zweiten Gliedes ein Hauptziel. Dieselbe erfolgt durch die Substitution :

$$\begin{aligned} &y = 3x - f_1, \text{ dann kommt} \\ &y^3 - 3\varphi y - \psi = 0, \text{ wo} \\ &\varphi = f_1^2 - 3f_2 \ , \ \psi = 2f_1^3 - 9f_1 f_2 + 27f_3. \end{aligned}$$

6|7

| Wir haben also jetzt die Gleichung $y^3 = \psi + 3\varphi y$ aufzulösen. Dieselbe stimmt mit der Identität

$$(z + t)^3 = z^3 + t^3 + 3zt(z + t)$$

überein, wenn man setzt :

$$\begin{aligned} &\psi = z^3 + t^3 \\ &\varphi = zt, \text{ also : } t = \frac{\varphi}{z} \\ &y = z + t. \end{aligned}$$

Setzen wir also $y = z + \frac{\varphi}{z}$, so wird der Gleichung $y^3 = \psi + 3\varphi y$ genügt, sofern wir $z$ so bestimmen, daß $\psi = z^3 + \frac{\varphi^3}{z^3}$ ist. Es ist also $z^3$ durch eine quadratische Gleichung zu bestimmen. Die Wurzeln derselben sind :

$$\begin{matrix} z^3 \\ \frac{\varphi^3}{z^3} \end{matrix} = \left\{ \frac{1}{2}\psi + \frac{1}{2}\sqrt{\psi^2 - 4\varphi^3} \right.$$

In Folge der Quadratwurzel ist der Ausdruck $\frac{1}{2}\psi + \frac{1}{2}\sqrt{\psi^2 - 4\varphi^3}$ doppeldeutig. Bezeichnen wir den einen seiner Werthe mit $z^3$, so ist der andre $\frac{\varphi^3}{z^3}$. Diese Doppeldeutigkeit drücken wir am passendsten so aus : Es bedeute $\rho$ einen der beiden Werthe von $\frac{-1+\sqrt{-3}}{2}$, gleichgültig welchen, dann ist aber, sobald der Werth einmal gewählt ist, $\rho^2$ ein ganz bestimmter Werth und es ist, da 1, $\rho$ und $\rho^2$ die Wurzeln der kubischen

Gleichung

$$\rho^3 = 1 \text{ sind, } 1 + \rho + \rho^2 = 0.$$

Wir schreiben nun : $\psi + \sqrt{\psi^2 - 4\varphi^3} = \psi + 3\chi\sqrt{-3}$ indem wir dem hierdurch definirten $\chi$ ein beliebiges, aber bestimmtes Zeichen geben. Vermöge der Gleichung : $\frac{-1+\sqrt{-3}}{2} = \rho$ läßt sich dann $\sqrt{-3} = \rho - \rho^2$ setzen. Indem wir $\rho$ einen der genannten Werthe sein lassen, läßt sich dann die eine der Wurzeln von $z^3 + \frac{\varphi^3}{z^3} = \psi$ darstellen durch

$$\begin{cases} z^3 = \frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi \text{ und die andre durch} \\ \frac{\varphi^3}{z^3} = \frac{1}{2}\psi + \frac{3}{2}(\rho^2 - \rho)\chi \end{cases}$$

Daher ist jetzt, da $3x - f_1 = y = z + \frac{\varphi}{z}$ war :

$$3x = f_1 + \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi} + \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho^2 - \rho)\chi}.$$

Darin bedeutet also $\rho$ eine beliebige aber bestimmte Wurzel der Gleichung $\rho^3 = 1$ (die 1 ausgeschlossen).

Wir haben in dieser Formal, welche die *Cardan*'sche heißt, einen 3-werthigen Ausdruck gefunden, welcher die 3 Wurzeln der kubischen Gleichung darstellt. In der hier befolgten Methode ist nun die geforderte *Reductio ad aequationes puras*, welche ja das Wesen der Auflösung ist, geleistet. Wir haben nämlich die folgende Kette von reinen Gleichungen :

$$\begin{cases} \chi = \sqrt{\frac{4\varphi^3 - \psi^2}{27}} \\ z = \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi} \\ 3x = f_1 + z + \frac{\varphi}{z} \end{cases}$$

7|8 | Diese Kette von reinen Gleichungen ersetzt die kubische Gleichung, sie bietet die Möglichkeit durch bloßes Wurzelausziehen den Werth von $x$ zu finden.

Es treten hier bereits merkwürdige algebraische Begriffe auf. Wir haben eine einfache Gleichung (*aequatio affecta*), wir lösen sie, indem wir 3 reine Gleichung (*aequ. purae*) aufstellen, welche zusammen als Kette aufgefaßt jene ersetzen. Wir schlagen also den umgekehrten Weg ein, wie bei der Elimination, wo man aus mehreren Gleichungen eine einzige Endgleichung bildet. Es scheint allerdings die Kette auf 6 Werthe von $x$ zu führen, da ja $\chi$ sowohl mit dem einen als dem andern Zeichen genommen werden kann. Aber man erkennt leicht, daß von jenen 6 Werthen immer 2 einander gleich sind. Aus der Diskussion über die Wurzeln des Gleichung $z^3 + \frac{\varphi^3}{z^3} = \psi$ geht nämlich hervor, daß $z^3$ in $\frac{\varphi^3}{z^3}$ übergeht, wenn man $\chi$ das entgegengesetzte Zeichen gibt. Man kann also, wie dies auch bereits oben angegeben ist, dem $\chi$ ein ganz bestimmtes Zeichen beigelegt denken, doch ist dies nicht erforderlich, da, wie bemerkt, in der Kette das Zeichen von $\chi$ keine Rolle spielt. Die in $z^3$ quadratische Gleichung hat von Lagrange den Namen : *Resolvente* der ursprünglichen kubischen Gleichung erhalten. In Bezug auf $z$ ist sie vom 6. Grade, aber die sechs Wurzeln zerfallen in 3 Gruppen von der Beschaffenheit, daß das Produkt der 2 Wurzeln einer jeden Gruppe gleich $\varphi$ ist. Der Ausdruck $z + \frac{\varphi}{z}$ ist also nur 3 deutig. Damit ist die Auflösung der kubischen Gleichungen vollständig geleistet. Wir wollen nun aber ähnlich wie bei den quadratischen Gleichungen einen Blick in das Innere, in den Bau dieser Auflösung werfen. Zu diesem Zwecke suchen wir die Hülfsgroßen durch die Wurzeln $x_0x_1x_2$ der ursprünglichen kubischen Gleichung auszudrücken.

Es ist $x^3 - f_1x^2 + f_2x - f_3 = (x - x_0)(x - x_1)(x - x_2)$. Dies ist eine reine Identität. Dabei ist :

$$f_1 = x_0 + x_1 + x_2 \ , \ f_2 = x_0x_1 + x_1x_2 + x_2x_0 \ , \ f_3 = x_0x_1x_2$$

Wir wollen nun $\varphi, \psi$ und $\chi$ durch $x_0, x_1, x_2$ ausdrücken. Da wir $3x - f_1 = y$ gesetzt hatten, so erhält man $\psi$, indem man in der Identität

$$(y + x_1 + x_2 - 2x_0)(y + x_2 + x_0 - 2x_1)(y + x_0 + x_1 - 2x_2) = y^3 - 3\varphi y - \psi$$

$y$ gleich 0 setzt. Es ist also | 8|9

$$\psi = (2x_0 - x_1 - x_2)(2x_1 - x_2 - x_0)(2x_2 - x_0 - x_1)$$

Zur Bestimmung von $\varphi = f_1^2 - 3f_2$ setzen wir :

$$w_1 = x_0 + \rho x_1 + \rho^2 x_2 \ , \ w_2 = x_0 + \rho^2 x_1 + \rho x_2. \text{ Dann ist :}$$

$$w_1 w_2 = x_0^2 + x_1^2 + x_2^2 + \rho^2(x_0x_1 + x_1x_2 + x_2x_0) + \rho(x_0x_1 + x_1x_2 + x_2x_0). \text{ Oder :}$$
$$w_1 w_2 = (x_0 + x_1 + x_2)^2 - 3(x_0x_1 + x_1x_2 + x_2x_0) = f_1^2 - 3f_2 = \varphi. \text{ Also ist :}$$
$$\varphi = (x_0 + \rho x_1 + \rho^2 x_2)(x_0 + \rho^2 x_1 + \rho x_2) = w_1 w_2.$$

Zur Berechnung von $\chi$ gehen wir aus von der Gleichung :

$$y_0 = 3x_0 - f_1 = Z + \frac{\varphi}{z} \ , \text{ oder } z^2 - (3x_0 - f_1)Z + \varphi = 0.$$

Nun ist $w_1 w_2 = \varphi$

$$w_1 + w_2 = 2x_0 - x_1 - x_2 = 3x_0 - f_1$$

Also sind $w_1$ und $w_2$ die Wurzeln jener quadratischen Gleichung. Es ist nun :

$$\begin{aligned}
-27\chi^2 &= \psi^2 - 4\varphi^3 = (z^3 + \frac{\varphi^3}{z^3})^2 - 4w_1^3 w_2^3 \\
\text{Also : } -27\chi^2 &= (w_1^3 + w_2^3)^2 - 4w_1^3 w_2^3 = (w_1^3 - w_2^3)^2. \\
\text{Also ist : } 3\sqrt{-3}\chi &= w_1^3 - w_2^3 = (w_1 - w_2)(w_1 - \rho w_2)(w_1 - \rho^2 w_2) \\
3\sqrt{-3}\chi &= (\rho - \rho^2)(x_1 - x_2)(1 - \rho)(x_0 - x_1)(1 - \rho^2)(x_0 - x_2) \\
3\sqrt{-3}\chi &= 3(\rho - \rho^2)(x_0 - x_1)(x_1 - x_2)(x_0 - x_2)
\end{aligned}$$

Da nun $\sqrt{-3} = \pm(\rho - \rho^2)$ ist und wir dem $\chi$ ein beliebiges Zeichen geben können, so ist also :

$$\chi = (x_0 - x_1)(x_1 - x_2)(x_2 - x_0)$$

Wir wollen hier kurz noch den Ausdruck für $\chi^2$ anmerken, da wir denselben später gebrauchen werden :

$$\chi^2 = f_1^2 f_2^2 - 4f_1^3 f_3 + 18 f_1 f_2 f_3 - 4f_2^3 - 27f_3^2$$

Wir haben also $\varphi, \psi, \chi$ in lineare Faktoren der $x_0, x_1, x_2$ zerlegt.

$$\begin{aligned}
\varphi &= (x_0 + \rho x_1 + \rho^2 x_2)(x_0 + \rho^2 x_1 + \rho x_2) \\
\psi &= (2x_0 - x_1 - x_2)(2x_1 - x_2 - x_0)(2x_2 - x_0 - x_1) \\
\chi &= (x_0 - x_1)(x_1 - x_2)(x_2 - x_0)
\end{aligned}$$

An diese Resultate lassen sich nun verschiedene Betrachtungen knüpfen. In unserer Kette traten die Hülfsgroßen $\chi$ und $z$ auf. Von fundamentaler Bedeutung nun ist der Umstand, daß die Hülfsgroßen $\chi$ und $z$ rationale Funktionen der Wurzeln $x_0x_1x_2$ der ursprünglichen kubischen Gleichung sind ($z^3$ ist ja gleich $w_1^3$ resp. $w_2^3$). Diese Erkenntnis, eine der kostbarsten Bemerkungen in der Theorie der algebraischen

Gleichungen, stammt erst von *Abel*. Auf dieser einfachen Bemerkung beruht die ganze Auflösbarkeit der kubischen Gleichungen.

9|10 | Wir haben nun, ähnlich wie bei den quadratischen Gleichungen, jetzt einen 3 werthigen Ausdruck herzustellen, der uns in Form einer Identität zu den Wurzeln $x_0 x_1 x_2$ führt.

## A.2 Traduction de l'introduction

# Introduction

## Un court aperçu du développement historique de l'algèbre

§1 : Déjà chez *Euclide* et *Diophante*, on trouve résolus les problèmes qui ont nécessité, de notre point de vue actuel, la résolution des équations du premier et du second degré. Ensuite nous trouvons au 9$^{\text{ième}}$ siècle les recherches algébriques des Arabes, dont l'impulsion a été donnée par les Juifs. Le nom Algebra est d'origine arabe.

Mais c'est d'abord au temps de la Renaissance que commence, simultanément avec le renouvellement de l'art, une nouvelle époque de l'algèbre. Sur cela, comme dans toutes les autres disciplines, ce sont les Italiens qui ont les premiers donné l'impulsion d'un nouveau développement des sciences mathématiques. Mais déjà avant cette époque en Italie deux noms sortent du lot. *Fibonacci*, sous le nom de *Lionardo Pisano*, avait déjà en 1202 rédigé un traité qu'il avait nommé le *Liber abaci* (*abacus*, le boulier) et dans lequel il traite du cas particulier des équations du troisième degré (avec une racine réelle) et où il fournit la preuve que la racine n'est pas représentable géométriquement.

Il faut ensuite nommer *Marco Polo*, dont on pense qu'il a ramené en Europe les chiffres de ses voyages orientaux.

Après lui il faut souligner en particulier *Pacioli* (*Luca di Borgo*), un ecclésiastique, qui a vécu dans la seconde moitié du 15$^{\text{ième}}$ siècle et qui a publié en 1494
1|2 le premier traité imprimé (*L'arte maggiore | ditta da vulgo la regulo della cosa*). L'expression « *regula della cosa* » (*cosa* signifie l'inconnue) demeure dans tout le Moyen Âge. À partir de là, l'algèbre se propage exceptionnellement vite. Une série de travaux importants s'enchaîne rapidement. *Scipione del Ferro* (mort en 1524 à *Bologna*) donne en 1505 la résolution de l'équation cubique $x^3 + ax = b$, *Nicolo Tartaglia* (1506-1559) résout les équations cubiques pour 3 formes particulières (coefficients positifs et négatifs), travail qui a été publié par contre en premier par *Cardano* (1501-1576) et étendu à tous les cas.

*Cardano* s'est fortement disputé à ce propos avec *Tartaglia*, lequel a fait valoir sa priorité, dispute qui a provoqué à l'époque un énorme scandale dans tout le monde civilisé.

Immédiatement après suit la résolution des équations du 4$^{\text{ième}}$ degré par *Ferrari* (1522-1566 à *Bologne*). *Rafaello Bombelli*, dont le travail a d'abord été publié, a le premier reconnu la forme réelle des racines de l'équation cubique, même si elles apparaissent sous forme imaginaire.

Pendant ce temps paraissent en Allemagne quelques bons travaux. *Christoph Rudolfs* a publié en 1524 le premier écrit d'algèbre en Allemagne. Le signe des racines vient de lui ; dans le même temps l'algébriste *Robert Ricord* (mort en 1558 à *Londres*) a amené le signe d'égalité.

*Michael Stifel*, lequel a été rendre visite à *Luther* en 1523 à Wittenberg pour se convertir au protestantisme, a écrit une *Arithmetica integra* avec un avant-propos de *Melanchthon*. Le troisième livre est lui-même intitulé : *De numeris cossicis* et on y trouve la première idée du logarithme. *Stifel* a vécu de 1487 à 1567 et passa la fin de sa vie à *Iena*.

Dès lors, la science algébrique devient internationale.

L'apparition notable suivante est celle de *François Viète* (*Vieta*) de 1540 à 1603. Il était avocat à *Poitiers* et a accompagné plus tard *Henri IV* à *Paris*. Il a été le premier à introduire le calcul littéral et à en développer les conséquences. Ainsi

désigna-t-il par exemple les grandeurs connues par des consonnes et les inconnues par les voyelles. Les expressions " *potestas pura*" (puissance) et "*potestas affecta*" (fonction entière) viennent de lui, expressions qui seront conservées jusqu'à *Gauss*. C'est *Viète* qui a reconnu comme fondement de la théorie des fonctions entières la décomposition en facteurs linéaires. Par contre il évitait encore les racines négatives, et il ignorait encore le signe de la multiplication et de l'égalité.

*Thomas Harriot* (né en 1560 à *Oxford*, mort en 1621 à *Londres*) était un vrai géomètre et a accompagné en tant que métreur le célèbre *Walter Raleigh* en *Virginie* (1585). Son œuvre : *Artis analyticae praxis ad aequationes algebraicas resolvendas* a été publiée après sa mort.

*Albert Girard* (mort en 1633 en *Hollande*) a publié en 1629 une algèbre dans laquelle il justifie la formule dite de *Newton*.

Maintenant vient le grand *Descartes* (1596-1650). Il a vécu de 1624 à 1649 en *Hollande*, où il a publié son œuvre mathématique et philosophique. De lui vient la détermination du nombre de racines positives et négatives (Règle des signes de *Descartes*) et avant tout l'introduction des grandeurs continues dans l'algèbre par la grande découverte de la géométrie analytique.

*Fermat* (1608-1665), parlementaire à *Toulouse*. De lui vient la réduction des équations, c'est-à-dire la transformation de celles-ci lorsque l'on y rencontre des grandeurs imaginaires. Son influence indirecte en algèbre par ses recherches en théorie des nombres est très importante.

*Newton* (1642-1726) *Arithmetica universalis*. Règle sur le nombre de racines réelles. Formules de Newton. Méthode d'approximation des solutions d'une équation.

*Graf Tschirnhausen* (né en 1651, mort en 1708 à *Kieslingswalde* près de *Görlitz*). La méthode de *Tschirnhaus* pour la recherche des solutions d'une équation du quatrième degré. En 1682 *Tschirnhaus* devient un membre externe de l'*Académie parisienne*.

*Euler* (Né en 1707 à *Basel*, mort en 1783 à *St. Petersburg*). *Euler* a vécu de 1741 à 1766 à *Berlin*. De lui provient une nouvelle résolution des équations du quatrième degré. Pour la diffuser, son algèbre (dictée par Euler aveugle à un jeune couturier) est traduite en 1774 en français et augmentée de précieuses notes par *Lagrange*.

*Bézout* (1730-1783). De lui provient à peu près tout ce que l'on comprend aujourd'hui sous le nom d'élimination et de théorie des fonctions symétriques.

| *Vandermonde* (1735-1796), Membre de l'*Académie de Paris*, et *Lagrange* (1736- 3|4
1813, né à *Turin*, mort à *Paris*), publient presque simultanément (1771 et 1772) leurs travaux algébriques, l'un à *Paris*, l'autre à l'*Académie de Berlin*. *Lagrange* était de 1766 à 1787 à *Berlin*. Le sens de ces travaux n'émerge vraiment seulement qu'avec les recherches de *Gauss* (1777-1855). *Gauss* passe son doctorat en 1799 à *Helmstädt* et publie en 1801 ses *Disquisitiones arithmeticae*. La septième section contient les recherches sur la construction de polygones réguliers, la plus grande découverte mathématique depuis l'antiquité. D'*Euclide* à *Gauss*, la tâche alléchante de la construction de polygones réguliers n'avait pas avancé d'un iota. C'est *Gauss* le premier qui a donné une solution à ce problème vieux de 2000 ans. *Gauss* a donné comme date de sa découverte le 30 Mars 1796. La haute valeur de la découverte de *Gauss* est l'exemple de la résolvabilité d' équations de tout degré. Sa découverte a rendu claires les raisons pour lesquelles la construction de polygones réguliers n'a fait aucun progrès rapide. En effet on peut construire géométriquement (c'est-à-dire par des droites et des cercles) seulement la division du cercle en $2^{2^\lambda}+1$ parties, c'est-à-dire en $3(\lambda=0)$, $5(\lambda=1)$, $17(\lambda=2)$, $257(\lambda=3)$, etc. parties égales. La division en 5 parties égales se trouve déjà chez *Euclide*. *Vandermonde* a le premier examiné les divisions en 7, 11, 13 parties et a commencé la division en 17 parties qui a épuisé ses forces et il a laissé de côté ses recherches. Seule une petite étape a manqué et *Vandermonde* avait anticipé la découverte de *Gauss*. *Gauss* désigne comme la nature de la résolution des équations la *Reductio aequationum affectarum ad aequationes puras*.

Déjà dans sa dissertation, dans laquelle il a donné la première preuve rigoureuse

de l'existence des racines, il a exprimé la supposition que les équations de degré de plus de quatre ne sont pas résolvables. Au début de sa dissertation il a donné une critique des preuves de l'existence des racines conduites jusqu'à présent et montre en particulier que la preuve provenant d'*Euler* n'est pas soutenable, car cette dernière est conduite sous l'hypothèse que toutes les équations sont résolvables.

Cette supposition que toutes les équations sont résolubles était assez généralement répandue avant *Gauss*. *Gauss* commence à en douter, mais c'est *Abel* le premier qui a prouvé l'incorrection de cette supposition. Mais la seconde supposition que tous les
4|5 mathématiciens jusqu'à *Gauss* regardaient comme évidente, | le fait qu'en effet chaque équation de degré $n$ a $n$ racines, a été écartée comme idée reçue dans la dissertation de *Gauss*, car il livra la première preuve rigoureuse de cette proposition. Il n'est en effet rien moins qu'évident que deux courbes de degré $n$, qui dans la démonstration de *Gauss* se rencontrent, se coupent en $n$ points réels.

*Abel* (1802-1829), a donné en 1826 dans le tome 1 du *Journal de Crelle* la démonstration de l'impossibilité de la résolution des équations de degré supérieur à 4. Dans le tome 4 de *Crelle* il fonde la théorie des équations qui seront désignées par *Kronecker* par équations *abéliennes*. Il ne faut pas manquer de souligner l'influence indirecte d'*Abel* sur l'algèbre par son travail sur la théorie des fonctions.

Déjà en 1798 et 1806 l'italien *Ruffini* (1765-1822) avait tenté de donner une preuve de l'impossibilité de la résolution des équations de degré supérieur à 4, bien que ses preuves ne soient pas rigoureuses.

*Evariste Galois* (1811-1832), qui meurt un 26 octobre en duel [2], un des plus grands algébristes de tous les temps, publie dans le *journal de Liouville* son travail sur les équations algébriques, dans lequel il traite, d'un autre point de vue, du travail d'*Abel*, et même l'élargit.

*Fourier* (1768-1830). Depuis 1817 Secrétaire de l'*Académie*. Il est à là tête de l'*Académie* en *Égypte*. De lui viennent des règles sur la détermination du nombre de racines.

La solution complète de ce problème : déterminer le nombre de racines réelles d'une équation algébrique entre deux valeurs réelles, est donnée en premier en 1829 par le genevois *Charles Sturm* (1803-1855). Il vécut la plupart du temps à *Paris*, où il était en étroite relation avec *Dirichlet*.

Jetons un regard en arrière. Deux problèmes ont occupé de préférence pendant 2000 ans les algébristes :

L'un théorique, concernant la résolubilité des équations, a été tranché par *Abel* en 1826 de manière négative ; l'autre pratique, concernant le nombre de racines réelles d'une équation entre deux valeurs données, a été résolu 3 ans plus tard (1829) par *Charles Sturm*.

Le problème de rendre positives les recherches d'*Abel*, c'est-à-dire la question de
5|6 déterminer les équations résolubles, | est l'objet des recherches les plus récentes, pour lesquelles les travaux de *Gauss*, *Abel* et *Galois* ont donné les fondations.

## §2 : Perspectives sur la théorie des équations algébriques

Pour montrer quels sont les points de vue à prendre en compte dans la théorie des équations algébriques, nous voulons continuer avec une présentation rapide de la théorie des équations quadratiques et cubiques. Soit d'abord l'équation quadratique donnée par :

$$x^2 - f_1 x + f_2$$

2. *N.d.T.* En fait, E. Galois meurt un 31 mai et nait un 26 octobre.

La *Reductio ad aequationem puram* est effectuée par la forme :

$$(2x - f_1)^2 + (4f_2 - f_1^2) = 0,$$

laquelle se distingue seulement de l'*aequatio affecta* par une autre réduction. Si l'on désigne par $x_1$ et $x_2$ les racines de l'équation quadratique, alors on a :

$$x^2 - f_1 x + f_2 = (x - x_1)(x - x_2)$$

donc : $f_1 = x_1 + x_2$ et $f_2 = x_1 x_2$.

Si l'on pose ceci pour $f_1$ et $f_2$ dans l' *aequatio pura* $2x = f_1 + \sqrt{4f_2 - f_1^2}$, alors on établit :

$$\begin{aligned} 2x = &\quad x_1 + x_2 + \sqrt{(x_1 - x_2)^2} \\ 2x = &\quad x_1 + x_2 \pm (x_1 - x_2) = \left\{ \begin{array}{l} 2x_1 \\ 2x_2 \end{array} \right. \end{aligned}$$

On le voit donc : l'essence de la résolution de l'équation quadratique $x^2 - f_1 x + f_2 = 0$ consiste en la fabrication d'une expression bivalente qui contient seulement une racine quadratique et dans laquelle n'interviennent que des fonctions symétriques de $x_1$ et $x_2$.

Considérons l'équation cubique :

$$0 = x^3 - f_1 x^2 + f_2 x - f_3$$

Déjà, dans la haute algèbre, faire disparaître le second terme était un des objectifs principaux. Cela a lieu par la substitution :

$$\begin{aligned} &y = 3x - f_1, \text{ puis vient} \\ &y^3 - 3\varphi y - \psi = 0, \text{ où} \\ &\varphi = f_1^2 - 3f_2 \ , \ \psi = 2f_1^3 - 9f_1 f_2 + 27f_3 \end{aligned}$$

6|7

| Nous avons donc maintenant l'équation $y^3 = \psi + 3\varphi y$. Celle-ci coïncide avec l'identité

$$(z + t)^3 = z^3 + t^3 + 3zt(z + t)$$

lorsque l'on pose :

$$\begin{aligned} &\psi = z^3 + t^3 \\ &\varphi = zt, \text{ donc : } t = \frac{\varphi}{z} \\ &y = z + t \end{aligned}$$

Si l'on pose donc $y = z + \dfrac{\varphi}{z}$, alors on a l'équation $y^3 = \psi + 3\varphi y$, dans la mesure où nous avons fixé $z$ tel que $\psi = z^3 + \dfrac{\varphi^3}{z^3}$. Alors $z^3$ est ainsi déterminé par une équation quadratique. Les racines elles-mêmes sont :

$$\begin{matrix} z^3 \\ \dfrac{\varphi^3}{z^3} \end{matrix} = \left\{ \frac{1}{2}\psi + \frac{1}{2}\sqrt{\psi^2 - 4\varphi^3} \right.$$

La racine quadratique consécutive est l'expression bivalente $\frac{1}{2}\psi + \frac{1}{2}\sqrt{\psi^2 - 4\varphi^3}$. Si l'on désigne l'une de ces valeurs par $z^3$, alors l'autre est $\frac{\varphi^3}{z^3}$. Nous exprimons donc cette bivalence de la façon la plus appropriée : que $\rho$ désigne donc l'une des deux valeurs de $\frac{-1+\sqrt{-3}}{2}$, n'importe laquelle, puis, une fois la valeur choisie, $\rho^2$ est une valeur entièrement déterminée, et comme les racines de l'équation cubique $\rho^3 = 1$ sont 1, $\rho$ et $\rho^2$, on a $1 + \rho + \rho^2 = 0$.

Nous écrivons maintenant : $\psi + \sqrt{\psi^2 - 4\varphi^3} = \psi + 3\chi\sqrt{-3}$ dans laquelle nous donnons par là à $\chi$ un signe quelconque, mais déterminé. En vertu de l'équation : $\frac{-1+\sqrt{-3}}{2} = \rho$ on peut alors poser $\sqrt{-3} = \rho - \rho^2$. Tandis que $\rho$ est l'une des valeurs en question, l'une des racines de $z^3 + \frac{\varphi^3}{z^3} = \psi$ se laisse ensuite représenter par :

$$\begin{cases} z^3 = \frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi \text{ et l'autre par} \\ \frac{\varphi^3}{z^3} = \frac{1}{2}\psi + \frac{3}{2}(\rho^2 - \rho)\chi \end{cases}$$

De là on a maintenant que $3x - f_1 = y = z + \frac{\varphi}{z}$ devient :

$$3x = f_1 + \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi} + \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho^2 - \rho)\chi}$$

où $\rho$ représente donc une racine quelconque mais fixée de l'équation $\rho^3 = 1$ (1 exclu).

Nous avons dans cette formule, laquelle s'appelle formule de *Cardan*, trouvé une expression trivalente, laquelle représente les 3 racines de l'équation cubique. Dans la méthode suivie ici est maintenant développée la *Reductio ad aequationes puras* demandée, laquelle est bien l'essence de la résolution. Nous avons en effet la succession suivante de pures équations :

$$\begin{cases} \chi = \sqrt{\frac{4\varphi^3 - \psi^2}{27}} \\ z = \sqrt[3]{\frac{1}{2}\psi + \frac{3}{2}(\rho - \rho^2)\chi} \\ 3x = f_1 + z + \frac{\varphi}{z} \end{cases}$$

7|8 | Cette série de pures équations remplace l'équation cubique qui offre la possibilité de trouver la valeur de $x$ par extraction de racine.

Il se présente ici déjà une étrange notion algébrique. Nous avons une équation simple (*aequatio affecta*), nous la résolvons en posant 3 pures équations (*aequ. purae*), lesquelles, conçues ensemble comme une série, remplacent celle-ci. Nous prenons donc le chemin inverse, comme avec l'élimination où l'on construit à partir de plusieurs équations une seule équation finale. Cela semble toutefois conduire la série à 6 valeurs de $x$, puisque $\chi$ peut être pris avec un signe ou l'autre. Mais on peut s'apercevoir facilement que parmi ces 6 valeurs toujours 2 sont égales. De la discussion sur les racines de l'équation $z^3 + \frac{\varphi^3}{z^3} = \psi$ il ressort en effet que $z^3$ devient $\frac{\varphi^3}{z^3}$ lorsque l'on donne à $\chi$ le signe opposé. On peut aussi, comme on l'a déjà vu plus haut, penser adjoindre à $\chi$ un signe entièrement fixé, mais cela n'est pas nécessaire, comme on le remarque, car dans la série le signe de $\chi$ ne joue aucun rôle. L'équation quadratique en $z^3$ a reçu de Lagrange le nom de *résolvante* de l'équation cubique initiale. Elle est de degré 6 en $z$, mais les six racines se séparent en 3 groupes définis par le fait que le produit de deux racines de l'un de ces groupes est égal à $\varphi$. L'expression $z + \frac{\varphi}{z}$ a donc seulement 3 valeurs différentes. Ainsi la résolution de l'équation cubique est

complètement effectuée. Mais nous voulons maintenant, par analogie avec les équations quadratiques, jeter un œil à l'intérieur de la construction de cette résolution. Dans ce but nous cherchons à exprimer les grandeurs annexes par les racines $x_0x_1x_2$ de l'équation cubique initiale.

On a : $x^3 - f_1x^2 + f_2x - f_3 = (x-x_0)(x-x_1)(x-x_2)$. Ceci est une pure identité. En même temps on a :

$$f_1 = x_0 + x_1 + x_2 \ , \ f_2 = x_0x_1 + x_1x_2 + x_2x_0 \ , \ f_3 = x_0x_1x_2$$

Nous voulons maintenant exprimer $\varphi, \psi$ et $\chi$ avec $x_0, x_1, x_2$. Là nous avons posé $3x - f_1 = y$, donc nous obtenons $\psi$, où nous avons posé dans l'identité

$$(y + x_1 + x_2 - 2x_0)(y + x_2 + x_0 - 2x_1)(y + x_0 + x_1 - 2x_2) = y^3 - 3\varphi y - \psi$$

$y$ égal à 0. On a donc | 8|9

$$\psi = (2x_0 - x_1 - x_2)(2x_1 - x_2 - x_0)(2x_2 - x_0 - x_1)$$

Pour déterminer $\varphi = f_1^2 - 3f_2$ nous posons :

$$w_1 = x_0 + \rho x_1 + \rho^2 x_2 \ , \ w_2 = x_0 + \rho^2 x_1 + \rho x_2 \text{ et ainsi :}$$

$$w_1w_2 = x_0^2 + x_1^2 + x_2^2 + \rho^2(x_0x_1 + x_1x_2 + x_2x_0) + \rho(x_0x_1 + x_1x_2 + x_2x_0). \text{ ou :}$$
$$w_1w_2 = (x_0 + x_1 + x_2)^2 - 3(x_0x_1 + x_1x_2 + x_2x_0) = f_1^2 - 3f_2 = \varphi. \text{ Donc on a :}$$
$$\varphi = (x_0 + \rho x_1 + \rho^2 x_2)(x_0 + \rho^2 x_1 + \rho x_2) = w_1w_2.$$

Pour le calcul de $\chi$ nous partons de l'équation :

$$y_0 = 3x_0 - f_1 = z + \frac{\varphi}{z} \ , \text{ ou } z^2 - (3x_0 - f_1)z + \varphi = 0.$$

Maintenant on a $w_1w_2 = \varphi$

$$w_1 + w_2 = 2x_0 - x_1 - x_2 = 3x_0 - f_1$$

Donc $w_1$ et $w_2$ sont les racines de cette équation quadratique. On a maintenant :

$$\begin{aligned}
-27\chi^2 &= \psi^2 - 4\varphi^3 = (z^3 + \frac{\varphi^3}{z^3})^2 - 4w_1^3w_2^3 \\
\text{Donc : } -27\chi^2 &= (w_1^3 + w_2^3)^2 - 4w_1^3w_2^3 = (w_1^3 - w_2^3)^2. \\
\text{Ainsi on a : } 3\sqrt{-3}\chi &= w_1^3 - w_2^3 = (w_1 - w_2)(w_1 - \rho w_2)(w_1 - \rho^2 w_2) \\
3\sqrt{-3}\chi &= (\rho - \rho^2)(x_1 - x_2)(1 - \rho)(x_0 - x_1)(1 - \rho^2)(x_0 - x_2) \\
3\sqrt{-3}\chi &= 3(\rho - \rho^2)(x_0 - x_1)(x_1 - x_2)(x_0 - x_2)
\end{aligned}$$

On a maintenant ici $\sqrt{-3} = \pm(\rho - \rho^2)$ et nous pouvons donner à $\chi$ un signe quelconque, on a donc :

$$\chi = (x_0 - x_1)(x_1 - x_2)(x_2 - x_0)$$

Nous voulons encore remarquer ici rapidement que l'expression de $\chi^2$ dont nous aurons besoin plus tard est :

$$\chi^2 = f_1^2f_2^2 - 4f_1^3f_3 + 18f_1f_2f_3 - 4f_2^3 - 27f_3^2$$

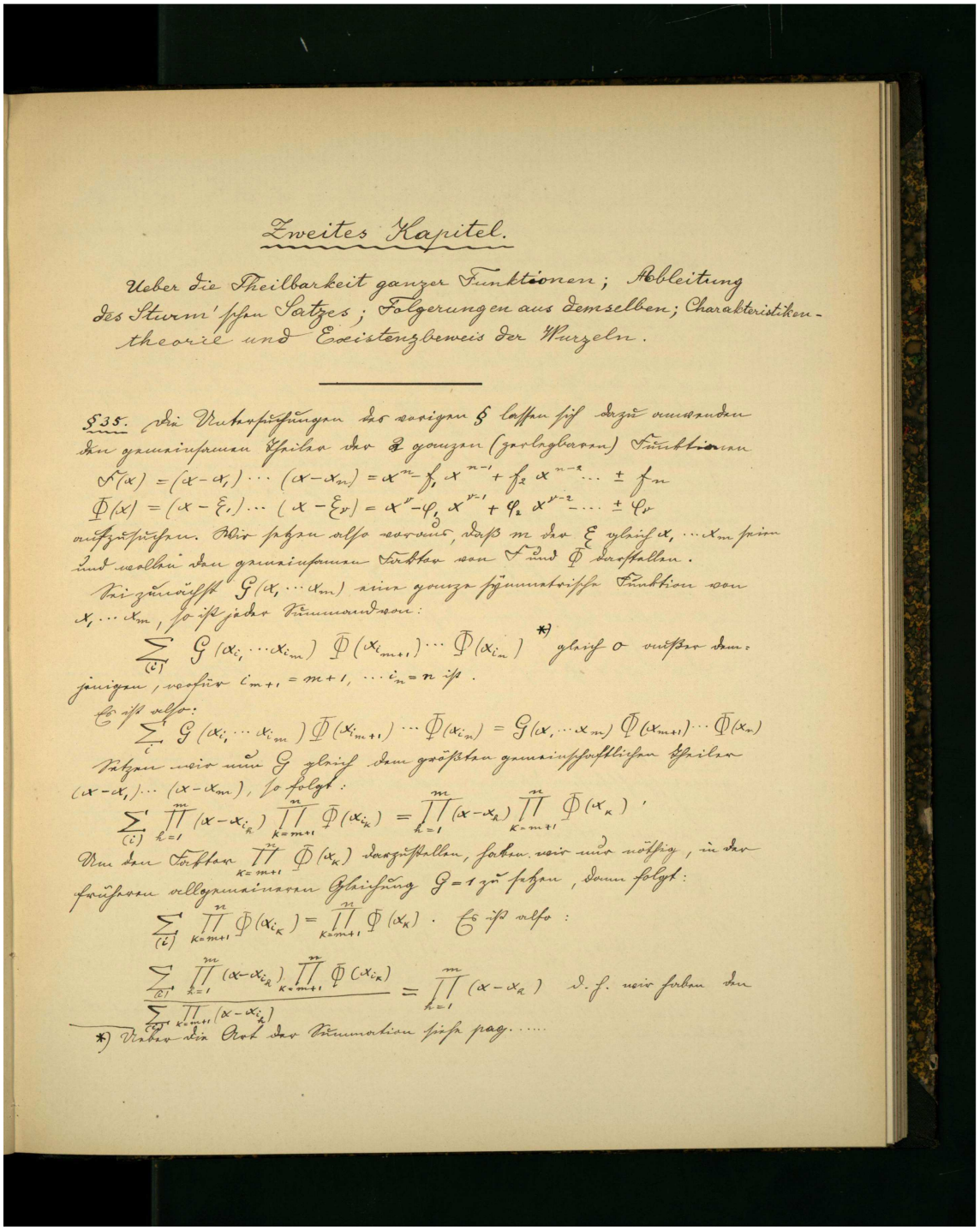
Nous avons donc décomposé $\varphi, \psi, \chi$ en facteurs linéaires de $x_0, x_1, x_2$.

$$\begin{aligned}\varphi &= (x_0 + \rho x_1 + \rho^2 x_2)(x_0 + \rho^2 x_1 + \rho x_2)\\ \psi &= (2x_0 - x_1 - x_2)(2x_1 - x_2 - x_0)(2x_2 - x_0 - x_1)\\ \chi &= (x_0 - x_1)(x_1 - x_2)(x_2 - x_0)\end{aligned}$$

Quelques observations sont à lier à ces résultats. Dans notre chaîne apparaissent les valeurs annexes $\chi$ et $z$. Il est d'une signification fondamentale que les valeurs annexes $\chi$ et $z$ soient des fonctions rationnelles des racines $x_0x_1x_2$ de l'équation cubique initiale ($z^3$ est bien égal à $w_1^3$ resp. $w_2^3$). Cette connaissance, une des observations les plus précieuses de la théorie des équations algébriques, apparut pour la première fois chez *Abel*. La résolution complète de l'équation cubique repose sur ces remarques simples.

9|10 | Nous avons déjà établi, comme dans l'équation quadratique, une expression à 3 valeurs qui nous conduit sous la forme d'une identité aux racines $x_0x_1x_2$.

## A.3 Transcription de la seconde partie du cours



Zweites Kapitel.

Ueber die Theilbarkeit ganzer Funktionen; Ableitung des Sturm'schen Satzes; Folgerungen aus demselben; Charakteristikentheorie und Existenzbeweis der Wurzeln.

§35. Die Untersuchungen des vorigen § lassen sich dazu verwenden den gemeinsamen Theiler der 2 ganzen (ganzbaren) Funktionen

$$f(x) = (x-\alpha_1)\cdots(x-\alpha_n) = x^n - f_1 x^{n-1} + f_2 x^{n-2} \cdots \pm f_n$$
$$\varphi(x) = (x-\xi_1)\cdots(x-\xi_\nu) = x^\nu - \varphi_1 x^{\nu-1} + \varphi_2 x^{\nu-2} - \cdots \pm \varphi_\nu$$

aufzusuchen. Wir setzen also voraus, daß m der $\xi$ gleich $\alpha_1 \cdots \alpha_m$ seien und wollen den gemeinsamen Theiler von $f$ und $\varphi$ darstellen.

Bezeichnet $G(\alpha_1 \cdots \alpha_m)$ eine ganze symmetrische Funktion von $\alpha_1 \cdots \alpha_m$, so ist jeder Summand von

$$\sum_{(i)} G(\alpha_{i_1} \cdots \alpha_{i_m})\, \varphi(\alpha_{i_{m+1}}) \cdots \varphi(\alpha_{i_n}) \quad *)$$

gleich 0 mit Ausnahme desjenigen, wofür $i_{m+1} = m+1, \cdots i_n = n$ ist.

Es ist also:

$$\sum_{(i)}' G(\alpha_{i_1} \cdots \alpha_{i_m})\, \varphi(\alpha_{i_{m+1}}) \cdots \varphi(\alpha_{i_n}) = G(\alpha_1 \cdots \alpha_m)\, \varphi(\alpha_{m+1}) \cdots \varphi(\alpha_n)$$

Setzen wir nun $G$ gleich dem größten gemeinschaftlichen Theiler $(x-\alpha_1)\cdots(x-\alpha_m)$, so folgt:

$$\sum_{(i)}' \prod_{k=1}^{m} (x-\alpha_{i_k}) \prod_{k=m+1}^{n} \varphi(\alpha_{i_k}) = \prod_{k=1}^{m} (x-\alpha_k) \prod_{k=m+1}^{n} \varphi(\alpha_k),$$

Um den Faktor $\prod_{k=m+1}^{n} \varphi(\alpha_k)$ darzustellen, haben wir nur nöthig, in der früheren allgemeinen Gleichung $G = 1$ zu setzen, dann folgt:

$$\sum_{(i)}' \prod_{k=m+1}^{n} \varphi(\alpha_{i_k}) = \prod_{k=m+1}^{n} \varphi(\alpha_k) \cdot \text{ Es ist also:}$$

$$\frac{\sum_{(i)}' \prod_{k=1}^{m} (x-\alpha_{i_k}) \prod_{k=m+1}^{n} \varphi(\alpha_{i_k})}{\sum_{(i)}' \prod_{k=m+1}^{n} (x-\alpha_{i_k})} = \prod_{k=1}^{m} (x-\alpha_k) \quad \text{d. h. wir haben den}$$

*) Wegen der Art der Summation siehe pag. ....

FIGURE 3 – Première page de la seconde partie du manuscrit de 1880-81 ([Kronecker, 1881a])

# Theorie

# der

# Algebraischen Gleichungen

# von

# Prof. Kronecker

---

Winter 1880/81

## | Zweites Kapitel

84|85

*Ueber die Theilbarkeit ganzer Funktionen ; Ableitung des Sturm'schen Satzes ; Folgerungen aus Demselben ; Charakteristiken-theorie und Existenzbeweis der Wurzeln.*

---

§35 : Die Untersuchungen des vorigen § lassen sich dazu anwenden den gemeinsamen Theiler der 2 ganzen (zerlegbaren) Funktionen

$$F(x) = (x - x_1)...(x - x_n) = x^n - f_1 x^{n-1} + f_2 x^{n-2} - ... \pm f_n$$
$$\Phi(x) = (x - \xi_1)...(x - \xi_\nu) = x^\nu - \varphi_1 x^{\nu-1} + ... \pm \varphi_\nu$$

aufzusuchen. Wir setzen also voraus daß $m$ der $\xi$ gleich $x_1, ..., x_m$ seien und wollen den gemeinsamen Faktor von $F$ und $\Phi$ darstellen. Sei zunächst $\mathcal{G}(x_1, ..., x_m)$ eine ganze symmetrische Funktion von $x_1, ..., x_m$, so ist jeder Summand von :

$$\sum_{(i)} \mathcal{G}(x_{i_1}...x_{i_m})\Phi(x_{i_{m+1}})...\Phi(x_{i_n}) \quad [3]$$

gleich 0 außer demjenigen, wofür $i_{m+1} = m + 1, ..., i_n = n$ ist.

Es ist also :

$$\sum_{(i)} \mathcal{G}(x_{i_1}...x_{i_m})\Phi(x_{i_{m+1}})...\Phi(x_{i_n}) = \mathcal{G}(x_1...x_m)\Phi(x_{m+1})...\Phi(x_n)$$

Setzen wir nun $\mathcal{G}$ gleich dem größten gemeinschaftlichen Theiler $(x-x_1)...(x-x_m)$, so folgt :

$$\sum_{(i)} \prod_{h=1}^{m}(x - x_{i_h}) \prod_{k=m+1}^{n} \Phi(x_{i_k}) = \prod_{h=1}^{m}(x - x_h) \prod_{k=m+1}^{n} \Phi(x_k).$$

Um den Faktor $\prod_{k=m+1}^{n} \Phi(x_k)$ darzustellen, haben wir nur nöthig, in der früheren allgemeineren Gleichung $\mathcal{G} = 1$ zu setzen, dann folgt :

$$\sum_{(i)} \prod_{k=m+1}^{n} \Phi(x_{i_k}) = \prod_{k=m+1}^{n} \Phi(x_k).$$

Es ist also :

$$\frac{\sum_{(i)} \prod_{h=1}^{h=m}(x - x_{i_h} \prod_{k=m+1}^{n} \Phi(x_{i_k})}{\sum_{(i)} \prod_{k=m+1}^{n}(x - x_{i_h})} = \prod_{h=1}^{m}(x - x_h)$$

d.h. wir haben den | gemeinsamen Theiler der beiden ganzen Funktionen $F$ und $\Phi$ als Quotienten zweier symmetrischer Funktionen der $n$ Größen $x$ dargestellt. 85|86

Haben $F$ und $\Phi$ den gemeinsamen Theiler $(x - x_1)...(x - x_m)$, so ist :

$$\prod_{k=1}^{n} \Phi(x_k) = 0, \prod_{k=2}^{n} \Phi(x_k), ..., \prod_{k=m}^{n} \Phi(x_k).$$

3. Ueber die Art der Summation siehe pag...

Dann ist auch stets :

$$\sum_{(i)} \prod_{k=2}^{n} \Phi(x_{i_k}) = 0, ..., \sum_{(i)} \prod_{k=m}^{n} \Phi(x_{i_k}) = 0.$$

Damit also ein Theiler $m$ ten Grades vorhanden sei, sind $m$ Bedingungen zu erfüllen. Nehmen wir an, daß $\sum_{(i)} \prod_{k=m+1}^{n} \Phi(x_{i_k}) \neq 0$ sei, so ist sicher kein größerer Theiler als der vom $m$ ten Grade vorhanden. In diesem letzteren Falle würde auch unsere Vorstellung vom gemeinschaftlichen Theiler $m$ ten Grades hinfällig. Damit also $(x-x_1)...(x-x_m)$ der größte gemeinschaftliche Theiler von $F$ und $\Phi$ sei, sind $m$ Gleichungen und eine Ungleichung zu erfüllen.

Wir haben den größten gemeinschaftlichen Theiler dargestellt als Quotienten zweier symmetrischer Funktionen. Daraus folgt, daß Zähler und Nenner dieser Darstellung ganze ganzzahlige Funktionen von $f_1...f_n$ und $\varphi_1, ..., \varphi_\nu$ sind.

Für $m = 1$ ist diese Darstellung zuerst von *Abel* gegeben worden.

§36 : Von besonderen Interesse ist nun der Fall, wo gar kein gemeinschaftlicher Theiler existirt. Alsdann ist schon die erste unsere Gleichungen, nämlich $\prod_{k=1}^{n} \Phi(x_k)$ von 0 verschieden. Wir können daher $\frac{1}{\Phi(x_k)}$ als ganze Funktion von $x_k$ darstellen. Es ist nämlich :

$$\frac{1}{\Phi(x_k)} = \frac{\Phi(x_1)....\Phi(x_{k-1})\Phi(x_{k+1})...\Phi(x_n)}{\prod_{k=1}^{n} \Phi(x_k)} = \frac{\Psi(x_k)}{R(f_1...f_n; \varphi_1...\varphi_\nu)}.$$

Es ist also :

$$\Phi(x_k)\Psi(x_k) = \prod_{k=1}^{n} \Phi(x_k) = R(f_1...f_n; \varphi_1...\varphi_\nu) = R(F, \Psi).$$

$R$ ist eine ganze ganzzahlige Funktion der $f$ und $\varphi$. Da nun $\Phi(x)\Psi(x) - R(F, \Psi)$ für $x = x_1, ...x_n$ verschwindet, so ist es durch $F(x)$ theilbar und wir erhalten die Gleichung :

$$\Phi(x)\Psi(x) + F(x)\mathcal{G}(x) = R(F, \Psi).$$

86|87 | Da $\Phi, \Psi, F$ von den Graden resp. $\nu, n-1, n$ sind, so folgt, daß $\mathcal{G}(x)$ vom Grade $\nu - 1$ sein muß.

Wir erhalten also das Resultat :

Haben $F(x)$ und $\Phi(x)$ keinen gemeinsamen Theiler, so lassen sich stets zwei Multiplikatoren $\Psi$ und $\mathcal{G}$ von den Graden resp. $n-1$ und $\nu - 1$ finden, so daß der Ausdruck $\Phi(x)\Psi(x) + F(x)\mathcal{G}(x) = R(F, \Phi)$ also unabhängig von $x$ ist. Wenn ein gemeinschaftlichen Theiler existirt, so muß $R = 0$ sein und wenn $R \neq 0$ ist, so kann kein gemeinschaftlicher Theiler existiren ; und zwar gilt dieser Satz für beliebige $F$ und $\Phi$, deren Zerlegbarkeit in Linearfaktoren nicht bewiesen zu sein braucht. Wir haben zwar die Gleichung $\Phi\Psi + F\mathcal{G} = R$ unter der Voraussetzung der Wurzelexistenz für $F$ und $\Phi$ hergeleitet, aber dieselbe gilt, weil identisch, für alle $f$ und $\varphi$.

Dagegen ist noch nicht der Satz bewiesen, daß, wenn $R = 0$ ist, ein Theiler existirt, oder daß umgekehrt $R \neq 0$ sein muß, wenn kein Theiler existirt. Dies ist nämlich nur bewiesen für $R(F, \Phi) = \prod_{k=1}^{n} \Phi(x_k)$ d.h. unter Voraussetzung der Wurzeln $x_1...x_n$ von $F$.

Setzen wir $\Phi(\xi) = \prod_{h=1}^{\nu}(\xi - \xi_h)$, so folgt :

$$R(F, \Phi) = \prod_{k=1}^{n} \prod_{h=1}^{\nu} (x_k - \xi_h) = (-1)^{n.\nu} \prod_{k=1}^{n} \prod_{h=1}^{\nu} (\xi_h - x_k) = (-1)^{n.\nu} \prod_{h=1}^{\nu} F(\xi_h).$$

Daraus folgt :

$$\prod_{k=1}^{n} \Phi(x_k) = R = (-1)^{n.\nu} \prod_{h=1}^{\nu} F(\xi_h).$$

Wir können dies so ausdrücken :

$$R(F, \Phi) = (-1)^{n.\nu} R(\Phi, F)$$

Es ist :

$$\Phi(x_k)\Psi(x_k) = R(F, \Phi) = F(\xi_h)\mathcal{G}(\xi_h).$$

Es ist daher : $\Psi(x) = \frac{R}{\Phi(x)}$ für $x = x_1...x_n$ und ebenso :

$$\mathcal{G}(\xi) = \frac{\mathcal{R}}{\mathcal{F}(\xi)} \text{ für } \xi = \xi_1, ..., \xi_\nu.$$

Wir können daher $\Psi(x)$ und $\mathcal{G}(x)$ nach der *Lagrange*'schen Formel herstellen und erhalten [4] :

$$\Psi(x) = RF(x) \sum_{k=1}^{n} \frac{1}{(x - x_k)F(x_k)\Phi(x_k)}$$

$$\mathcal{G}(x) = R\Phi(x) \sum_{h=1}^{\nu} \frac{1}{(x - \xi_h)\Phi'(\xi_h)F(\xi_h)}$$

| Es ist in der That $\Psi(x)$ vom Grade $n-1$ und $\mathcal{G}(x)$ vom Grade $\nu - 1$. Nun können 87|88
wir auch leicht direkt die Gleichung $\Phi(x)\Psi(x) + F(x)\mathcal{G}(x) = R$ beweisen, denn es ist :

$$\Phi(x)\Psi(x) + F(x)\mathcal{G}(x) = RF(x)\Phi(x) \left\{ \sum_{k=1}^{n} + \sum_{h=1}^{\nu} \right\}.$$

Um nun zu zeigen, daß der Faktor von $R$ gleich 1 ist, setzen wir $z = x_1...x_n, \xi_1...\xi_\nu$ als die Wurzeln von $F\Phi = 0$ und konstruiren nach Lagrange den Werth 1. Dann kommt :

$$1 = F\Phi \sum_{r=1}^{n+\nu} \frac{1}{z - z_r} \frac{1}{F(z_R)\Phi'(z_n) + F'(z_r)\Phi(z_r)} =$$

$$1 = F(z)\Phi(z) \left\{ \sum_{k=1}^{n} \frac{1}{(z - x_k)F'(x_k)\Phi(x_k)} + \sum_{h=1}^{\nu} \frac{1}{(z - \xi_h)\Phi'(\xi_h)F(\xi_h)} \right\} \text{ q.e.d.}$$

Man nennt $R(F, \Phi)$ die Resultante der Elimination von $x$ aus den beiden Gleichungen $F = 0$ und $\Phi = 0$.

Sei $F(y)$ eine ganze Funktion vom Gerade $n - \nu$. Dann können wir nach §34 das Produkt $\prod_{h=1}^{\nu} F(y_k)$, welches eine symmetrische Funktion der $y$ ist, interpolatorisch darstellen, indem $\prod F(y_h)$ eine symmetrische Funktion in reduzirter Form ist und wir erhalten :

$$\prod_{h=1}^{\nu} F(y_h) = \sum_{(i)} \prod_{h=1}^{\nu} F(x_{i_h}) \prod_{h=1}^{\nu} \prod_{k=\nu+1}^{n} \frac{y_h - x_{i_k}}{x_{i_h} - x_{i_k}}$$

4. *N.d.T.* Erreur manifeste :

$$\Psi(x) = RF(x) \sum_{k=1}^{n} \frac{1}{(x - x_k)F'(x_k)\Phi(x_k)}$$

Setzen wir $y_h = \xi_h$, so ist :

$$\prod_{h=1}^{\nu}\prod_{k=\nu+1}^{n}(\xi_h - x_{i_k}) = \prod_{k=\nu+1}^{n}(-1)^{\nu}\Phi(x_{i_k} = (-1)^{\nu.(n-\nu)}\prod_{k=\nu+1}^{n}\Phi(x_{i_k})$$

Also ist :

$$\sum_{(i)}\frac{\prod_{h=1}^{\nu}F(x_{i_h})\prod_{k=\nu+1}^{n}\Phi(x_{i_k})}{\prod_{h=1}^{\nu}\prod_{k=\nu+1}^{n}(x_{i_h}-x_{i_k})} = (-1)^{\nu.(n-\nu)}\prod_{h=1}^{\nu}F(\xi_h) = R(F,\Phi)$$

Damit ist also die Resultante von $F$, welches vom $n-\nu$ ten Grade ist und von $\Phi$, welches vom $\nu$ ten Grade ist, interpolatorisch dargestellt, durch die Werthe von $F$ und $\Phi$ für $n-\nu+\nu = n$ Werthe $x_1...x_n$. Diese elegante Darstellung rührt von *Rosenhain (Crelle Bd. 30).*

88|89 | §37 Wir hatten

$$R(F,\Phi) = \prod\Phi(x_k) = \prod_{h=1}^{\nu}\prod_{k=\nu+1}^{n}(x_k - \xi_h) = (-1)^{n.\nu}\prod_{h=1}^{\nu}F(\xi_h).$$

Betrachten wir nun gleichzeitig die beiden Gleichungen $F(x) = 0$ und $\Phi(x) = 0$ und setzen voraus, daß beide zusammen bestehen, d.h. gleichzeitig durch denselben Werth von $x$ erfüllt werden, so muß $R = 0$ sein. Anderseits, wenn $R = 0$ ist, so werden $F(x) = 0$ und $\Phi(x) = 0$ durch einen und denselben Werth von $x$ befriedigt. Man nennt deßwegen $R = 0$ die resultirende Gleichung von $F(x) = 0$ und $\Phi(x) = 0$. Sie ist das vollständige Resultat des gleichzeitigen Bestehens von $F(x) = 0$ und $\Phi(x) = 0$, d.h. die nothwendige und hinreichende Bedingung hierfür. Man hat diesen Begriff von den Gleichungen auch auf die Funktionen übertragen und nennt $R$ die Resultante von $F$ und $\Phi$.

Nehmen wir als speziellen Fall $\Phi(x) = F'(x)$, so ist :

$$R = \prod_{k=1}^{n}\Phi(x_k) = \prod_{k=1}^{n}F'(x_k) = \prod_{k,h}(x_k - x_h) = (-1)^{\frac{n(n-1)}{2}}\mathcal{P}^2.$$

Diese spezielle Resultante von $F(x)$ und $F'(x)$ nennt man nach Sylverster die Diskriminante von $F(x)$. Sie ist ein symmetrische Funktion von $x_1...x_n$ und folglich eine ganze ganzzahlige Funktion der elementaren symmetrischen Funktionen $f_1...f_n$. Ihr Verschwinden, oder was dasselbe ist, das gleichzeitige Bestehen von $F(x) = 0$ und $F'(x) = 0$ drückt aus, daß $F(x) = 0$ doppelte Wurzeln habe.

§38 : Wir hatten unabhängig von der Zerlegbarkeit von $F$ und $\Phi$ in Linearfaktoren bewiesen, daß wenn ein gemeinschaftlicher Theiler existirt, $R = 0$ sein muß, woraus von selbst folgt, daß für $R \neq 0$ kein gemeinschaftlicher Theiler existiren kann. Dieser Satz gilt also für beliebige ganze Funktionen $F$ und $\Phi$ mit beliebigen Coeffizienten $f$ und $\varphi$. Dagegen ist bis jetzt nur für die zerlegbaren Funktionen $F$ und $\Phi$ bewiesen, daß wenn kein gemeinschaftlicher Theiler existirt, $R \neq 0$ sein muß, woraus folgt, daß, wenn $R = 0$ ist, ein gemeinschaftlicher Theiler existirt. Wir wollen nun diesen letzteren Satz für beliebige ganze Funktionen $F$ und $\Phi$ beweisen, ohne deren Zerlegbarkeit in Linearfaktoren vorauszusetzen. Ist dieser Beweis gelungen, so ist damit für beliebige ganze Funktionen $F$ und $\Phi$ [5] bewiesen, daß $R = 0$ die nothwendige und hinreichende Bedingung für einen gemeinsamen Theiler ist, woraus dann von selbst
89|90 folgt, daß $R \neq 0$ die nothwendige und | hinreichende Bedingung dafür ist, daß kein gemeinschaftlicher Theiler existirt.

5. d. h. ohne Voraussetzung der Existenz der Wurzeln von $F_n.\Phi$.

Wir müssen nun zunächst eine Methode kennen lernen, nach welcher man entscheiden kann, ob ein gemeinschaftlicher Theiler existirt und welche einen solchen aufzufinden lehrt.

Diese Methode ist das alte, schon vor [von] Euklid bekannte Kettenbruchverfahren, mittelst dessen man den größten gemeinschaftlichen Theiler zweier Zahlen findet.

Es soll der größte gemeinschaftliche Theiler zwischen $f(x)$ und $f_1(x)$ bestimmt werden.

Wir nehmen an, $f_1$ sei von niedrigerem oder von gleichem Grade wie $f$ und dividiren successive :

$$\frac{f(x)}{f_1(x)} = g_1(x) - \frac{f_2(x)}{f_1(x)} \ , \ \frac{f_1(x)}{f_2(x)} = g_2(x) - \frac{f_3(x)}{f_2(x)} \text{ etc.}$$

Wir erhalten die Kette :

$$0 = f - g_1 f_1 + f_2 = f_1 - g_2 f_2 + f_3 = f_2[-]g_3 f_3 + f_4 = ... = f_{\nu-2} - g_{\nu-1} f_{\nu-1} + f_\nu.$$

Da sich bei der Division die Grade immer erniedrigen, so muß man einmal zum Grade 0 gelangen, d. h. zu einer Constanten. Nun kann dieselbe entweder von 0 verschieden sein, oder sie kann gleich 0 sein. Im ersteren Falle sei $f_\nu$ diese Constante, sodaß die letzte Gleichung der Kette lautet : $f_{\nu-1} - g_\nu f_\nu = 0$. Im letzteren Falle sei $f_{\nu+1}$ die Constante, welche verschwindet, sodaß auch dann $f_{\nu-1} - g_\nu f_\nu = 0$ die letzte Gl. der Kette ist.

Wir sehen, daß diese successive Division identisch ist mit der Verwandlung der rationalen Funktion $\frac{f_1(x)}{f(x)}$ in einen Kettenbruch. Es ist nämlich

$$\frac{f_1(x)}{f(x)} = \frac{1}{g_1 - \frac{f_2}{f_1}} = \frac{1}{g_1 - \cfrac{1}{g_2 - \cfrac{1}{g_3 - \cfrac{1}{\cdots \frac{1}{g_\nu}}}}}$$

Nun sind ohne Weiteres folgende 2 Sätze ersichtlich :

1. Der gemeinschaftliche Theiler zweier auf einander folgenden $f$ ist Theiler von jedem $f$.
2. Jedes $f$ ist eine homogene lineare Funktion von je 2 auf einander folgenden $f$.

Es ist also der gemeinschaftliche Theiler von irgend zwei auf einander | folgenden 90|91
$f$ z. B. von $f_1$ und $f$ auch Theiler von $f_\nu$. Also muß auch der größte gemeinschaftliche Theiler von $f$ und $f_1$ Theiler von $f_\nu$ sein. Andererseits ist $f_\nu$ Theiler von $f_{\nu-1}$, also von jedem $f$, woraus folgt, daß $f_\nu$ der größte gemeinschaftliche Theiler von $f$ und $f_1$ ist. Da jedes $f$ eine homogene lineare Funktion von je 2 auf einander folgenden $f$ ist, so folgt : $f_\nu = \varphi_1 f + \varphi f_1$. Auch hieraus ist ersichtlich, daß $f$ und $f_1$ keinen größern gemeinschaftlichen Theiler haben können. Drücken wir den Satz, daß $f_\nu$ größter gemeinschaftlicher Theiler von $f$ und $f_1$ ist, durch das Symbol : $f_\nu = Th(f, f_1)$ aus, so löst sich dasselbe in die 3 Gleichungen auf :

$$\left.\begin{array}{l} f = \Psi f_\nu \\ f_1 = \Psi_1 f_\nu \\ f_\nu = \varphi_1 f + \varphi f_1 \end{array}\right\}$$

welche zusammen ausdrücken, daß $f_\nu$ größter gemeinschaftlicher Theiler ist. Diese 3 Gleichungen sind äquivalent der Gleichung $f_\nu = Th(f, f_1)$, von jenen gelangt man zu dieser und auch umgekehrt von dieser zu jenen. Nun ist entweder $f_\nu$ eine ganze Funktion von $x$, sodaß $f_{\nu+1} = 0$ ist, in diesem Falle haben $f$ und $f_1$ wirklich einen größten gemeinschaftlichen Theiler $f_\nu$ ; oder aber $f_\nu$ ist eine von 0 verschiedene Constante, in diesem Falle existirt kein gemeinschaftlicher Theiler von $f$ und $f_1$. Es existirt also folgendes Criterium : Um zu entscheiden, ob $f$ und $f_1$ einen gemeinschaftlichen Theiler haben, dividire man successive nach dem Kettenbruchverfahren, bis man zu einer Funktion 0 ten Grades gelangt. Ist diese gleich 0, so ist die vorhergehende Funktion

der größte gemeinschaftliche Theiler von $f$ und $f_1$, ist sie eine von 0 verschieden Constante, so existirt ein solcher Theiler nicht. Die Coeffizienten von $f_\nu$ sind rational aus den Coeffizienten von $f$ und $f_1$ gebildet, das Kettenbruchverfahren ist überhaupt ein rationales Verfahren. Wenn wir dasselbe nun auf $F$ und $\Phi$ anwenden, so kommen wir sicher zu einer Gleichung : $\Phi\Psi' + F\mathcal{G}' = R'$. Es fragt sich nun, ob diese Gleichung mit der früheren : $\Phi\Psi + F\mathcal{G} = R$ übereinstimmt, nachdem wir $\Psi', \mathcal{G}', R'$ durch Multiplikation mit einem etwaigen Generalnenner zu ganzen ganzzahligen Funktionen auch der $f_i$ und $\varphi_i$ gemacht haben. Um dies zu entscheiden, beweisen wir, daß, wenn wir von $f$ und $f_1$ ausgehend zu $f_\nu$ gelangt sind, die Multiplikatoren $\varphi$ und $\varphi_1$, welche die Gleichung : $\varphi_1 f + \varphi f_1 = f_\nu$ befriedigen, vollständig bestimmt sind. Angenommen
91|92 nämlich auch | die Multiplikatoren $\varphi'$ und $\varphi_1'$ leisteten dies, dann wäre :

$$\varphi_1' f + \varphi' f_1 = f_\nu \text{ , also : } (\varphi_1 - \varphi_1')f + (\varphi - \varphi')f_1 = 0$$

$$\varphi(\varphi_1 - \varphi_1')f + (\varphi - \varphi')(f_\nu - \varphi_1 f) = 0.$$

Es müßte also $(\varphi - \varphi')f_\nu$ durch $f$ theilbar sein. Nun kann man ohne Beschränkung der Allgemeinheit annehmen, daß $\varphi$ von niedrigerem Grade ist als $f$ [6]. Es kann also $\varphi - \varphi'$ nur durch $f$ theilbar sein, wenn $\varphi = \varphi'$ ist. Ebenso zeigt man, daß $\varphi_1 = \varphi_1'$ ist. In der Gleichung $\varphi_1 f + \varphi f_1 = f_\nu$, wo $\varphi_1$ von niedrigerem Grade als $f_1$ und $\varphi$ von niedrigerem Grade als $f$ ist, sind also die Multiplikatoren vollständig eindeutig bestimmt. Wenn wir daher mit Hülfe des Kettenbuchverfahrens von $F$ und $\Phi$ ausgehend zu der Gleichung

$$\Phi\Psi' + F\mathcal{G}' = R'$$

gelangen, so kann sich diese nur durch eine multiplikative Constante von $\Phi\Psi + F\mathcal{G} = R$ unterscheiden, wenn wir $\Psi'\mathcal{G}', R'$ auch in den $f$ und $\varphi$ ganz und ganzzahlig gemacht haben ; und zwar kann $\Phi\Psi + F\mathcal{G} = R$ nur durch Multiplikation mit einer solchen Constanten nicht durch Division in $\Phi\Psi' + F\mathcal{G}' = R'$ übergehen. Nicht durch Division, denn sonst müßte $R$ in Faktoren zerlegbar sein, und daß dies nicht der Fall ist, sieht man leicht folgendermaßen ein : Angenommen :

$$R(f_1 ... f_n; \varphi_1 ... \varphi_\nu) = \prod(x_h - \xi_k)$$

sei in 2 Faktoren :

$$F(f_1 ... f_n; \varphi_1 ... \varphi_\nu) \text{ und } S(f_1 ... f_n; \varphi_1 ... \varphi_\nu)$$

so muß $F$ mindestens einen der Linearfaktoren $x_K - \xi_h$ enthalten, also : $F = (x_k - \xi_h)F(x1..., \xi_1....)$, da aber $F$ in Bezug auf alle $x$ und alle $\xi$ symmetrisch ist, so muß es alle Faktoren $(x_k - \xi_h)$ enthalten, d. h. es muß $F$ durch $R$ theilbar sein, d. h es ist

$$F = R \text{ , } S = 1.$$

Wenn wir daher mit Hülfe des Kettenbruchverfahrens zu der Gl. :

$$\Phi\Psi' + F\mathcal{G}' = R'$$

gelangt sind, in welchen $\Psi'\mathcal{G}', R'$ in Bezug auf die $f$ und $\varphi$ ganz und ganzzahlig
92|93 sind und wir heben einen etwa vorhandenen gemein- | -schaftlichen Theiler zwischen $\Psi'\mathcal{G}', R'$ weg, so kommen wir zu der früheren Gleichung :

$$\Phi\Psi + F\mathcal{G} = R$$

Wir wollen das Vorhergehende nun durch ein Beispiel illustriren. Sei $f = ax^3 +$

6. Siehe hierzu Anhang *pag.*...

$bx^2 + cx + d$, $f_1 = \alpha x^2 + \beta x + \gamma$, dann erhält man :

$$f_2 = (a\alpha\gamma + b\alpha\beta - c\alpha^2 - a\beta^2)\frac{x}{\alpha^2} - \frac{a\beta\gamma - b\alpha\gamma + d\alpha^2}{\alpha^2} = px - q \text{ (br. c.)}$$

$$f_3 = -\frac{1}{p^2}(pq\beta + p^2\gamma + q^2\alpha) \ , \ g_1 = \frac{a}{\alpha}x - \frac{a\beta - b\alpha}{\alpha^2} \ , \ g_2 = \frac{\alpha}{p}x + \frac{p\beta + q\alpha}{p^2}.$$

Setzen wir für $p$ und $q$ ihre Werthe, so erhalten wir für $f_3$ :

$$\begin{aligned} -p^2\alpha^2 f_3 \quad = \quad & a^2\gamma^3 + b^2\alpha\gamma^2 + c^2\alpha^2\gamma + d^2\alpha^3 - ab\beta\gamma^2 + ac\gamma(\beta^2 - 2\alpha\gamma) \\ & + ad\beta(3\alpha\gamma - \beta^2) - bc\alpha\beta\gamma + bd\alpha(\beta^2 - 2\alpha\gamma) - cd\alpha^2\beta. \end{aligned}$$

Es ist ferner : $f_3 = (g_1 g_2 - 1)f_1 - g_2 f$, daraus folgt, daß $-\alpha^2 p f_3 = R$ ist. Nehmen wir den speziellen Fall, daß $a = 1$, $b = -f_1$, $c = f_2$, $d = -f_3$, $\alpha = 3$, $\beta = -2f_1$, $\gamma = f_2$ ist, daß also $f_1$ die Ableitung von $f$ ist, so erhalten wir die Diskriminante und zwar wird diese zu :

$$-p^3\alpha^2 f_3 = R = (-1)^3(-f_1^2 f_2^2 + 4f_1^3 f_3 + 4f_2^3 - 18f_1 f_2 f_3 + 27f_3^2)$$

Diesem Werthe der Diskriminante der kubischen Gleichung $x^3 - f_1 x^2 + f_2 x - f_3 = 0$ sind wir in der That schon öfters begegnet.

§39 : Wir wollen jetzt den Beweis liefern, daß wenn 2 ganze Funktionen $A(x) = x^n - a_1 x^{n-1} + ... \pm a_n$ und $B(x) = x^\gamma - b_1 x^{\gamma-1} + ... \pm b_\gamma$ keinen gemeinschaftlichen Theiler haben, alsdann die Resultante $R(a_1...a_n; b_1...b\gamma) \neq 0$ sein muß und zwar soll nicht vorausgesetzt werden, daß $A(x)$ und $B(x)$ in Linearfaktoren zerlegbar seien.

Das Kettenbruchverfahren führt uns, wenn wir keinen gemeinschaftlichen Theiler voraussetzen zu der Gleichung :

$$P(x)A(x) + Q(x)B(x) = 1,$$

wo $P$ und $Q$ ganze Funktionen von $x$ sind resp. vom Grade $\gamma - 1$ und $n - 1$, deren Coeffizienten rationale Funktionen der $a$ und der $b$ sind. Die Gleichung ist eine reine Identität, wir können daher $x$ durch $x_k$ ersetzen, wo $x_k$ Wurzel von $F(x) = 0$ sei. Dann ist :

$$\begin{aligned} Q(x_k)B(x_k) \quad &= \quad 1 + P(x_k)(F(x_k) - A(x_k)) && \text{(A.1)} \\ &= \quad 1 + P(x_k)\sum_{h=1}^{n}(-1)^h(f_h - a_h)x_k^{n-h} && \text{(A.2)} \end{aligned}$$

| Multipliziren wir jetzt über $k$, so erhalten wir : 93|94

$$\prod_{k=1}^{n} Q(x_k) \prod_{k=1}^{n} B(x_k) = 1 + \sum_{h=1}^{n}(-1)^h(f_h - a_h)\mathcal{G}_h(f_1...f_n).$$

Auf der rechten Seite erhielten wir nämlich, abgesehen von der 1, ein Aggregat von Ausdrücken, von denen jeder ein $f_h - a_h$ enthielt, dessen Coeffizient dann eine symmetrische Funktion der $x$ war.

$\mathcal{G}_h(f_1...f_n)$ ist daher eine ganze ganzzahlige Funktion der $f$, welche außerdem noch die $a$ enthält und auch in Bezug auf diese ganz ist.

Nun ist $\prod_{k=1}^{n} Q(x_k)$ eine symmetrische Funktion der $x$ und eine rationale Funktion der $a$ und $b$. Es ist daher $\prod_{k=1}^{n} Q(x_k) = F(f_1...f_n)$. Ferner ist :

$$\prod_{k=1}^{n} B(x_k) = R(f_1...f_n; b_1...b\gamma).$$

Wir haben daher

$$F(f_1...f_n)R(f_1...f_n; b_1...b\gamma) = 1 + \sum_{h=1}^{n}(-1)^h(f_h - a_h)\mathcal{G}_h(f_1...f_n)$$

Dies ist eine reine Identität, wir können daher setzen $f_1 = a_1, ...f_n = a_n$ und erhalten :

$$F(a_1...a_n)R(a_1...a_n; b_1...b\gamma) = 1.$$

Da nun $F$ jedenfalls eine endliche Größe ist, so folgt :

$$R(a_1...a_n; b_1...b_\gamma) \neq 0. \text{ q. e. d.}$$

Dieser sinnreiche Beweis rührt von *Gauss*, der auf den bewiesenen Satz den Existenzbeweis der Wurzeln gründete. *Gauss* stellt diesse Betrachtung an für den speziellen Fall, wo die eine der beiden ganzen Funktionen die Ableitung der andern ist, dann erhält er das Resultat, daß wenn eine ganze Funktion mit ihrer Ableitung keinen gemeinsamen Theiler hat, alsdann die Diskriminante $R = (-1)^{\frac{n.n-1}{2}} P^2$ von Null verschieden sein muß.

Wir haben jetzt, ohne die Existenz der Wurzeln vorauszusetzen, bewiesen, daß für beliebige ganze Funktionen $F(x)$ und $\Phi(x)$ das Verschwinden der Resultante $R(F, \Phi)$ die nothwendige und hinreichende Bedingung dafür ist, daß $F$ und $\Phi$ einen gemeinschaftlichen Theiler haben, woraus von selbst sich ergibt, daß die nothwendige und hinreichende Bedingung dafür, daß 2 beliebige ganze Funktionen keinen gemeinschaftlichen Theiler haben, die ist, daß $R(F, \Phi)$ von 0 verschieden ist.

94|95 §40 [7] Wir untersuchen nun die Zerlegbarkeit ganzer ganzzahliger | Funktionen in Faktoren. Diese wichtige Untersuchung ist analog der Zerlegung einer Zahl in ihre Primfaktoren. Ist eine Zahl $n$ gegeben, so können wir nach allen Zahlen fragen, die in $n$ ohne Rest aufgehen. Die Methode besteht einfach darin, alle Zahlen von 1 bis $n-1$ inkl. darauf zu untersuchen, ob ein Vielfaches von ihnen gleich $n$ ist. Diese eventuellen Theiler können wir dann wieder in ihre Theilfaktoren zerlegen und gelangen schließlich zu der Darstellung $n = p_1^{k_1} p_2^{k_2}...$, wo $p_1, p_2, ....$ nicht weiter zerlegbar sind. Solche nur durch 1 und sich selbst theilbare Zahlen nennen wir Primzahlen. Es gilt dann der Satz, daß die Zerlegung einer Zahl $n$ in ihre Primfaktoren eine eindeutige sei. Der Beweis stützt sich auf den Satz, daß wenn ein Produkt von mehreren Faktoren durch eine Primzahl theilbar ist, nothwendigerweise einer dieser Faktoren durch diese Primzahl theilbar sein muß ; es ist also für die Eindeutigkeit der Zerlegung die Methode der größten gemeinschaftlichen Theilers erforderlich.

Uebertragen wir nun diese Untersuchengen auf ganze ganzzahlige Funktionen, so kann eine solche zunächst Divisoren haben, die bloße Zahlen sind. Dies tritt ein, wenn die Coeffizienten der nach Potenzen von $x$ geordneten ganzen ganzzahligen Funktion einen gemeinsamen Theiler haben. Den größten gemeinschaftlichen Theiler aller Coeffizienten heben wir heraus und untersuchen also nur noch solche ganze ganzzahlige Funktionen, welche keine Zahlendivisoren haben.

Wir beweisen jetzt zunächst folgenden Satz :

> Ist das Produkt $\Psi(x)\chi(x)$ zweier ganzen ganzzahligen Funktionen durch die Primzahl $p$ theilbar und ist $\chi$ nicht durch $p$ theilbar, so muß $\Psi$ durch $p$ theilbar sein.

7. *N.d.T.* Nous ajoutons ce numéro de chapitre, absent du manuscrit (la numérotation passe directement de 39 à 41) car il nous semble bien qu'il y a ici un changement de sujet

Es sei :

$$\begin{aligned}\Psi(x) &= a_0 + a_1x + a_2x^2 + ...\\ \chi(x) &= b_0 + b_1x + b_2x^2 + ...\end{aligned}$$

Da $\chi(x)$ nicht durch $p$ theilbar ist, so muß es Coeffizienten $b$ geben, die nicht durch $p$ theilbar sind. Angenommen $b_0, b_1, b_2, ..., b_k - 1$ seien durch $p$ theilbar, $b_k$ aber nicht, so muß :

$$(a_0 + a_1x + a_2x^2 + ...)(b_kx^k + b_{k+1}x^{k+1} + ...)$$

durch $p$ theilbar sein. Multipliziren wir aus, so sind Coefficienten von $x^k, x^{k+1}, ...$ resp. [8] :

$$a_0b_k, a_1b_k + a_0b_{k+1}, a_2b_k + a_1b_{k+1} + a_2b_{k+2}...$$

| Da $b_k$ nicht durch $p$ theilbar ist, so muß es $a_0$ sein ; daraus folgt weiter, daß 95|96
$a_1, a_2, ...$ durch $p$ theilbar sind und da aus der Theilbarkeit von $a_h$ die von $a_{h+1}$ folgt, so ist damit bewiesen, daß $\Psi(x)$ durch $p$ theilbar ist.

So einfach dieser Satz zu sein scheint, so findet er sich merkwürdigerweise doch erst bei *Gauss*. Angenommen nun, wir hätten eine ganze Funktion $\Phi(x)$ in 2 Faktoren $\Psi(x)$ und $\chi(x)$ zerlegt, welche zwar ganze Funktionen von $x$ sind, deren Coeffizienten aber nicht ganze, sondern nur rationale Zahlen sind. Dann bringen wir $\Psi$ auf seinen Generalnenner $n$, also etwa : $\Psi(x) = \frac{\Psi_1(x)}{m}\frac{\chi_1(x)}{n}$, dann müssen, weil $\Psi_1\chi_1$ durch $mn$ theilbar ist, die Primfaktoren von $mn$ sich so vertheilen, daß jeder derselben entweder in $\Psi_1$ oder in $\chi_1$ enthalten ist. Wir erhalten also zwei ganze ganzzahlige Funktionen $\Psi(x)$ und $\chi(x)$, deren Produkt gleich $\Phi(x)$ ist. Also :

Wenn eine ganze ganzzahlige Funktion $\Phi(x)$ in ganze Funktionen mit rationalen Coeffizienten zerlegbar ist, so ist sie auch in ganze ganzzahlige Faktoren zerlegbar. Wir brauchen also nur zu untersuchen, ob eine ganze ganzzahlige Funktion $\Phi(x)$ ganze ganzzahlige Theiler hat. Nun könnten wir einfach für die gegebene Funktion :

$$\Phi(x) = \varphi_0 + \varphi_1x + \varphi_2x^2 + ...$$

ansetzen :

$$(\Psi_0 + \Psi_1x + \Psi_2x^2 + ...)(\chi_0 + \chi_1x + \chi_2x^2 + ...) = \varphi_0 + \varphi_1x + \varphi_2x^2 + ...$$

und dann $\Psi_0, \Psi_1, \Psi_2, ..., \chi_0, \chi_1, \chi_2, ...$ als ganze Zahlen aus den Gleichungen bestimmen :

$$\begin{aligned}\Psi_0\chi_0 &= \varphi_0\\ \Psi_0\chi_1 + \Psi_1\chi_0 &= \varphi_1\\ \Psi_0\chi_2 + \Psi_1\chi_1 + \Psi_2\chi_0 &= \varphi_2\end{aligned}$$

Aber diese Methode ist erstens sehr umständlich und zweitens setzt sie die Auflösbarkeit eines Systems von Gleichungen voraus, was wir jetzt noch nicht annehmen können. Es gibt aber eine sehr einfache Methode, welche erlaubt, alle Coeffizienten in derselben Weise zu bestimmen wie $\Psi_0$ und $\chi_0$ aus $\Psi_0\chi_0 = \varphi_0$ bestimmt werden, nämlich einfach als Theiler von $\varphi_0$.

Um zu untersuchen, ob eine vorgelegte Funktion $F(x)$ vom $m$ ten Grade ganze ganzzahlige Divisoren z. B. vom $n-1$ ten Grade besitzt, (wir haben um alle diese Theiler zu erhalten nur nöthig $n-1$ alle Werthe | von 1 bis $\frac{m}{2}$ resp. $\frac{m-1}{2}$, je nachdem 96|97
$m$ grade oder ungrade ist, annehmen zu lassen), denken wir uns nämlich diesen Di-

8. *N.d.T.* Dans le manuscrit, on a :

$$a_0b_k, a_1b_k + a_0b_{k+1}, a_2b_k + a_1b_{k+1} + a_2b_{k+2})...$$

visor $n-1$ ten Grades nicht nach Potenzen von $x$, sondern nach ganzen Funktionen von $x$ geordnet. Sei $\Psi(x)$ der Divisor, so können wir nach der *Lagrange*'schen Formel schreiben :

$$\Psi(x) = \sum_{k=1}^{n} C_k \frac{F(x, x_k)}{F'(x_k)},$$

indem wir unter $x_1, x_2, ...x_n$ irgend welche ganze Zahlen verstehen, solche jedoch ausgenommen, für welche etwa $F'(x)$ verschwindet. Wir wollen diese Zahlen durch $r_1 ... r_n$ bezeichnen. Dann müssen wir die $C_1, ...C_n$ so bestimmen, daß für $x = r_k \Psi(r_k)$ ein Theiler von $F(r_k)$ wird. Nun ist $\Psi(r_k) = C_k$, also muß $C_k$ ein Theiler von $F(r_k)$ sein. Damit erhalten wir für $C_k$ eine ganz beschränkte Anzahl von Zahlenwerthen. Indem wir alle Combinationen $C_k$ in $\Psi(x) = \sum_{k=1}^{n} C_k \frac{F(x,r_k)}{F(r_k)}$ einsetzen, erhalten wir alle ganzen Funktionen $n-1$ ten Grades, welche überhaupt Divisoren sein können. Von diesen sind sodann alle auszuschließen, welche nicht auch ganzzahlige Funktionen von $x$ sind, und nur die übrig bleibenden ganzen ganzzahligen Funktionen von $x$ können Divisoren von $F(x)$ sein, indem wir einen Divisor so verstehen, daß $F(x)$ getheilt durch einen solchen Divisor wieder eine ganze ganzzahlige Funktion von $x$ ist. Wir hätten natürlich den Divisor $\Psi(x)$ auch nach den $g_h(x)$ ordnen können (*pag...*) nämlich :

$$\Psi(x) = \sum_{h=0}^{h=1} \sum_{k=1}^{n} \varphi_k \frac{{x_k}^{n-h-1}}{F'(x_k)} g_h(x),$$

wo jetzt die $\varphi_k$ als die Divisoren von $F(x_k) = F(r_k)$ zu bestimmen sind. Auch in dieser Darstellung sind natürlich nur die ganze ganzzahligen $\Psi(x)$ zu berücksichtigen. Damit ist eine vollkommen elementare und in jedem vorkommenden Falle unbedingt zum Ziele führende Methode gegeben, um zu entscheiden, ob eine vorgelegte ganze ganzzahlige Funktion $F(x)$ Theiler besitzt oder nicht.
Beispiel :

$$F(x) = x^5 - x^4 + x^3 - x^2 - 3x - 1.$$

Wir wollen untersuchen, ob ein Theiler 3 ten Grades existirt. Wir setzen also $n = 4$ und wählen $r_1 = 0, r_2 = 1, r_3 = -1, r_4 = -2$, dann ist : $F(0) = -1, F(1) = -4, F(-1) = -2, F(-2) = -55$. $C_1, C_2, C_3, C_4$ müssen also Theiler dieser Zahlen $F(r_k)$ sein.

97|98 | Nun ist :

$$\begin{aligned}
\tfrac{F(x,r_1)}{F'(r_1)} &= -\tfrac{1}{2}(x^3 + 2x^2 - x - 2) \\
\tfrac{F(x,r_2)}{F'(r_2)} &= \tfrac{1}{6}(x^3 + 3x^2 + 2x) \\
\tfrac{F(x,r_3)}{F'(r_3)} &= \tfrac{1}{2}(x^3 + x^2 - 2x) \\
\tfrac{F(x,r_4)}{F'(r_4)} &= \tfrac{1}{6}(x^3 - x)
\end{aligned}$$

Wählen wir zu $c_1, c_2, c_3, c_4$ resp. 1,4,-2,-11, so folgt :

$$\Psi(x) = \sum_{k=1}^{n} c_k \frac{F(x, x_k)}{F'(x_k)} = x^3 + 2x + 1.$$

Diese Wahl liefert also eine ganze ganzzahlige Funktion. Sie ist in der That ein Theiler von $F(x)$, denn durch Division von $x^3 + 2x + 1$ in $F(x)$ erhalten :

$$x^5 - x^4 + x^3 - x^2 - 3x - 1 = (x^3 + 2x + 1)(x^2 - x - 1).$$

§41. Nachdem wir nun eine Methode gewonnen haben, die in jedem einzelnen Falle erlaubt, zu entscheiden, ob eine vorgelegte ganze ganzzahlige Funktion von $x$ ganze ganzzahlige Divisoren hat, sind wir berechtigt, den Begriff der Unzerlegbarkeit oder Irreduzibilität aufzustellen. Unter einer unzerlegbaren oder irreduziblen Funktion von

$x$, verstehen wir eine ganze ganzzahlige Funktion von $x$, welche nicht in ganze ganzzahlige Divisoren zerlegt werden kann, also auch keinen Zahlenfaktor enthält. Eine irreduzible Funktion spielt unter den Funktionen eine ähnliche Rolle wie die Primzahlen bei den Zahlen. Jede ganze ganzzahlige Funktion lößt sich also jetzt darstellen in der Form :

$$F(x) = p_1^{h_1} p_2^{h_2} ... p_\mu^{h_\mu} f_1^{k_1} f_2^{k_2} ... f_\nu^{k_\nu},$$

wo $p_1 ... p_\mu$ Primzahlen, $f_1, ..., f\mu$ irreduzible Funktionen und $h_1 ... h_\mu, k_1 ... k_\nu$ ganze positive Zahlen sind. Es gilt nun wie bei der Zerlegung der Zahlen der Satz, daß diese Zerlegung eine eindeutige | ist. Um diese Eindeutigkeit zu beweisen, haben wir 98|99
zunächst den folgenden Satz darzuthun.

Wenn das Produkt $\Phi.\Psi$ zweier ganzer ganzzahliger Funktionen $\Phi$ und $\Psi$ von $x$ durch eine irreduzible Funktion $F(x)$ theilbar ist, so muß entweder $\Phi$ oder $\Psi$ durch $F(x)$ theilbar sein.

Angenommen nämlich, $\Phi$ sei nicht durch $F$ theilbar, so können $\Phi$ und $F$ auch keinen gemeinsamen Faktor haben, weil ja $F$ irreduzibel ist. Wir können also 2 ganze Funktionen $\Phi_1$ und $F_1$ mit rationalen Coeffizienten so bestimmen, daß : $\Phi_1 F + F_1 \Phi = 1$ ist. Dann folgt :

$$\Psi \Phi_1 F + F_1 \Phi \Psi = \Psi.$$

Nun ist die linke Seite durch $F$ theilbar, also auch $\Psi$, aber möglicherweise ist der Quotient nicht ganzzahlig. Setzen wir daher $\Psi = F.\chi$ und bezeichnen den Generalnenner in $\chi$ mit $m$, so müssen, weil $\Psi$ ganzzahlig ist und $F$ auch keine Zahlenfaktoren enthält, alle Primfaktoren von $m$, also $m$ selbst, in $\chi$ auch gehen, d. h. $\Psi$ ist durch $F$ auch ganzzahlig theilbar, so daß der Quotient $\chi$ eine ganze ganzzahlige Funktion von $x$ ist. q.e.d.

Aus diesem Satze ergibt sich ohne Weiteres die Eindeutigkeit der Zerlegnug einer ganzen ganzzahligen Funktion in irreduzible Faktoren. Denn angenommen es gäbe eine zweite Zerlegung $q_1^{r_1} ... q_{\mu'}^{r'_\mu} \varphi_1^{s_1} ... \varphi_{\nu'}^{s_{\nu'}}$ so wäre also diese gleich $p_1^{h_1} ... p_\mu^{h_\mu} f_1^{k_1} ... f_\nu^{k_\nu}$ und keiner der Faktoren der einen Seite ist durch irgend einen Faktor der andern Seite theilbar, während doch das Produkt durch ihn theilbar ist. Dies widerspricht also dem eben bewiesenen Satze, folglich ist die Zerlegung eindeutig.

§ 43[§42] : Wir wollen nun diese Resultate auf Funktionen mehrerer Variabeln übertragen.

Wir sagen, eine ganze ganzzahlige Funktion der Variabeln $x, y, z, ...$ ist ein Theiler der ganzen ganzzahligen Funktion $F(x, y, z, ...)$, wenn der Quotient wieder eine ganze ganzzahlige Funktion von $x, y, z, ...$ ist. Wir wenden das Induktionsverfahren an und setzen voraus, die Zerlegbarkeit von allen Funktionen, welche eine Variable weniger enthalten als $F(x, y, z, ...)$ | sei untersucht und die Eindeutigkeit der Zerlegung sei 99|100
nachgewiesen.

Wir untersuchen dann erstens, ob $F(x, y, z, ...)$ Theiler hat, die von $x$ unabhängig sind. Ein solcher Theiler müßte Theiler sein von allen Coeffizienten $f_k(y, z, ...)$ der nach $x$ geordneten Funktion $F(x, y, z...)$. Diese Untersuchung können wir ohne Weiteres erledigen, da wir ja bereits über die Theilbarkeit von Funktionen $f(y, z, ...)$ nach Voraussetzung alles Erforderliche wissen.

Zweitens können wir jetzt untersuchen, ob $F(x, y, z...)$ eine ganze ganzzahlige Funktion von $x, y, z, ...$ zum Theiler hat, indem wir alle etwaigen von $x$ freien Theiler aus $F(x, y, z, ...)$ bereits wegdividirt denken. Wir denken dann einen solchen präsumtiven Theiler nach ganzen Funktionen $G_k(x)$ geordnet, dann müssen die Coeffizienten $C_k$ Theiler von $F(r_k, y, z, ...)$ sein. Diese können wir alle auffinden und erhalten für die $C_k$ eine endliche Anzahl von ganzen ganzzahligen Funktionen von $y, z, ...$ (pag. 244). Wir erhalten also eine endliche Anzahl von Versuchen um zu entscheiden, ob $F(x, y, z, ...)$ einen Faktor $f(x, y, z...)$ besitzt.

Drittens beweisen wir, nachdem wir jetzt berechtigt sind, von ganzen ganzzahligen irreduzibeln Funktionen mehrerer Variabeln zu reden, daß wenn ein Produkt

$\Phi.\Psi$ zweier ganzen ganzzahligen Funktionen der Variabeln $x, y, z, ...$ zunächst durch eine irreduzible Funktion $\varphi(y, z, ...)$ theilbar ist, daß dann entweder $\Phi$ oder $\Psi$ durch $\varphi(y, z...)$ theilbar sein muß. Sei also :

$$\Phi.\Psi = (a_0 + a_1x + a_2x^2 + ...)(b_0 + b_1x + b_2x^2 + ...)$$

durch $\varphi(y, z, ...)$ theilbar (pag. 2...), wo die $a$ und $b$ ganze ganzzahlige Funktionen von $y, z, ...$ sind und sei $\Psi = b_0 + b_1x + b_2x^2 + ...$ nicht durch $\varphi$ theilbar, so dürfen also nicht alle $b$ durch $\varphi$ theilbar sein. Angenommen $b_k$ sei der erste nicht durch $\varphi$ theilbare Coeffizient, dann muß :

$$(a_0 + a_1x + a_2x^2 + ...)(b_kx^k + b_{k+1}x^{k+1} + ...)$$

durch $\varphi$ theilbar sein. Durch $\varphi$ müssen also theilbar sein :

$$a_0b_k, a_0b_{k+1} + a_1b_k, ...$$

Da für Funktionen der Variabeln $y, z, ...$ alle Sätze über die Theilbarkeit als bewiesen angenommen sind, so folgt nach der Reihe, daß $a_0, a_1, a_2...$ durch $\varphi$ theilbar sind, d.h.
100|101 daß $\Psi$ durch $\varphi$ theilbar ist. | Hieraus folgt als Corollar : Wenn eine ganze ganzzahlige Funktion $\Psi(x, y, z...)$ durch eine ebensolche Funktion $f(x, y, z, ...)$, welche keinen von $x$ unabhängigen Theiler hat, so theilbar ist, daß der Quotient eine ganze Funktion von $x$ ist, so muß der Quotient auch eine ganze ganzzahlige Funktion der übrigen Variabeln sein. Denn sei :

$$\Psi(x, y, z...) = f(x, y, z...)\frac{Q(x, y, z...)}{R(y, z...)}$$

indem wir unter $R$ den Generalnemmer von $Q$ verstehen, so ist $f.Q$ durch $R$ theilbar und da $f$ nicht durch $R$ theilbar ist, so muß $Q$ durch $R$ theilbar sein, d. h. der Quotient ist eine ganze ganzzahlige Funktion von $x, y, z....$ Die Frage also nach der Theilbarkeit einer ganzen ganzzahligen Funktion aufgefaßt als Funktion von $x$ allein, löst auch die Frage nach der Theilbarkeit dieser Funktion aufgefaßt als Funktion von allen Variabeln.

Nun wollen wir endlich den Satz beweisen, daß wenn das Produkt $\Phi.\Psi$ zweier ganzen ganzzahligen Funktionen durch die irreduzible Funktion $F(x, y, z, ...)$ theilbar ist, entweder $\Phi$ oder $\Psi$ durch $F$ theilbar sein muß.

Angenommen $\Phi$ sei nicht durch $F$ theilbar, so haben, weil $F$ irreduzibel ist, $\Phi$ und $F$ auch als Funktionen von $x$ keinen gemeinsamen Theiler. Wir können also setzen : $\Phi F_1 + F\Phi_1 = G(y, z, ...)$, wo $F_1\Phi_1$ und $G$ ganze ganzzahlige Funktionen von $y, z, ...$ seien. Dann folgt :

$$\Psi\Phi F_1 + \Psi F\Phi_1 = \Psi G(y, z, ...)$$

Die linke Seite ist jetzt durch $F$ theilbar, also auch die rechte und es muß $\frac{\Psi G(y,z,....)}{F(x,y,z...)}$ eine ganze ganzzahlige Funktion sein, welche wir mit $H(x, y, z....)$ bezeichnen. Dann ist also $H.F$ durch $G$ theilbar, und da $F$ nicht theilbar ist, so ist es $H$. Wir können daher schreiben $\Psi = H_1.F$, wo $H_1$ eine ganze ganzzahlige Funktion von $x, y, z[...]$ ist, d. h. es ist $\Psi$ durch $F$ theilbar. Aus diesem Satze folgt dann endlich die Eindeutigkeit der Zerlegung einer ganzen ganzzahligen Funktion in irreduzible Faktoren.

101|102

| §43. Wenden wir die erhaltenen Resultate auf die Untersuchung der Diskriminante und der Resultante an.

Es ist $P^2 = (-1)^{\frac{n.n-1}{2}} \prod(x_h - x_k)^2 = D(f_1, f_2, ...f_n)$ eine ganz ganzzahlige Funktion von $f_1...f_n$. Faßt man $P^2$ als Funktion von $f_1...f_n$ auf, so ist es auch irreduzibel. Denn enthielte es einen Faktor $F(f_1, ..., f_n)$ und ersetzen wir die $f$ durch die $x$, so

muß $F$ mindestens einen der Faktoren $x_h - x_k$ enthalten und weil $F$ symmetrisch ist, so enthält es alle und zwar wegen der Symmetrie alle in Quadrat, d. h. $D(f_1...f_n)$ ist irreduzibel.

Dies ist zwar nur bewiesen für die speziellen $f$, welche sich symmetrisch aus den $x$ zusammensetzen. Aber wenn eine Funktion für ein spezielles Werthsystem unzerlegbar ist, so kann sie auch nicht in Allgemeinen zerlegbar sein. Wir ersehen also hieraus, daß $D(f_1...f_n)$ irreduzibel ist, auch wenn wir die $f$ als ganz beliebige Variabeln ansehen.

Daß auch die Resultante $R(f_1...f_n;\varphi_1...\varphi_\nu)$ irreduzibel ist, wenn die $f$ und $\varphi$ durch die $x$ resp. $\xi$ definirt sind, haben wir bereits *pag....* bewiesen. Wir wollen dies etwas allgemeiner fassen. Seien die beiden Funktionen, welche wir zunächst noch als zerlegbar voraussetzen wollen, dargestellt durch

$$F(x) = a_0x^n + a_1x^{n-1} + ... + a_n = a_0(x^n + \tfrac{a_1}{a_0}x^{n-1} + ... + \tfrac{a_n}{a_0})$$

$$\Phi(x) = \alpha_0x^\nu + \alpha_1x^{\nu-1} + ... + \alpha_\nu = a_0(x^\nu + \tfrac{\alpha_1}{\alpha_0}x^{\nu-1} + ... + \tfrac{\alpha_\nu}{\alpha_0})$$

dann ist : $R = \prod_{k=1}^{n}\Phi(x_k)$ ein Quotient, dessen Zähler eine ganze ganzzahlige Funktion der $a$ und $\alpha$ ist und dessen Nenner das Produkt irgend einer Potenz von $a_0$ mit irgend einer Potenz von $\alpha_0$ ist. Wir wollen zeigen, daß in dieser Darstellung schon eine ganze ganzzahlige Funktion aller $a$ und $\alpha$ wird, wenn wir mit $a_0^\nu$ multipliziren. Denn es ist :

$$a_0^\nu\prod_{k=1}^{n}\Phi(x_k) = a_0^\nu\prod_{k,h}\alpha_0(\alpha_k - \xi_h) = a_0^\nu\alpha_0^n\prod_{k,h}(\alpha_k - \xi_h)$$

In dieser Darstellung erkennt man, daß die höchsten in $\prod_{k,h}(\alpha_k - \xi_h)$ vorkommenden Nenner $a_0^\nu$ und $\alpha_0^n$ sind, daß also :

$$a_0^\nu\prod_{k=1}^{n}\Phi(x_k) = a_0^\nu\alpha_0^n\prod_{k,h}(\alpha_k - \xi_h) = (-1)^{n\nu}\alpha_0^n\prod_{h=1}^{\nu}F(\xi_h)$$

eine ganze ganzzahlige Funktion von $a_0...a_n$ und $\alpha_0...\alpha_\nu$ ist.

| Bezeichnen wir dieselbe also durch $R(a_0...a_n;\alpha_0...\alpha_\nu)$ so ist zu zeigen, daß diese Funktion, welche wir jetzt als die Resultante von $F(x)$ und $\Phi(x)$ bezeichnen, irreduzibel ist. Nun folgt aber aus : 102|103

$$R(a_0...a_n;\alpha_0...\alpha_\nu) = a_0^\nu\alpha_0^n\prod_{h,k}(x_k - \xi_h),$$

daß $R(a_0....a_n;\alpha_0...\alpha_\nu)$ gleich ist $a_0^\nu\alpha_0^n$, multiplizirt mit einer irreduzibeln Funktion von $\frac{a_1}{a_0}, ...\frac{a_n}{a_0}; \frac{\alpha_1}{\alpha_0}, ...\frac{\alpha_\nu}{\alpha_0}$. Es könnte also $R$ höchstens eine Potenz von $a_0$ oder $\alpha_0$ zum Faktor haben ; daß dem aber nicht so ist, zeigt der Umstand, daß $R$ ein Glied $a_n^\nu\alpha_0^n$ enthält, in welchem kein $a_0$ vorkommt und ebenso ein Glied $\alpha_\nu^n a_0^\nu$, welches kein $\alpha_0$ enthält. Damit ist bewiesen, daß $R(a_0...a_n;\alpha_0...\alpha_\nu)$ irreduzibel ist, wenn die $\frac{a_k}{a_0}$ und $\frac{\alpha_k}{[\alpha_0]a_0}$ als symmetriche Funktionen der $x$ resp. der $\xi$ definirt sind. Da also $R(a_0...a_n;\alpha_0...\alpha_\nu)$ für spezielle $a_0...a_n;\alpha_0...\alpha_\nu$ irreduzibel, so ist es auch in Allgemeinen irreduzibel.

Die Resultante $R(a_0...a_n;\alpha_0...\alpha_\nu)$ zweier Funktionen

$$a_0x^n + a_1x^{n-1} + .... + a_n \text{ und } \alpha_0x^\nu + ... + \alpha_\nu,$$

über deren Zerlegbarkeit wir nichts voraussetzen, ist also irreduzibel.

Es möge hier noch der Beweis hinzugefügt werden, daß die Resultante eine Invariante ist, d. h. ungeändert bleibt, wenn wir in den Funktionen $F(x)$ und $\Phi(x)$ für $x$ die gebrochene lineare Funktion $\frac{ax+b}{cx+d}$ setzen, die Funktionen ganz machen und dann aus den Coeffizienten die Resultante bilden. Wir nehmen zunächst wieder an, $F(x)$

und $\Phi(x)$ seien zerlegbar. Der Beweis für die Invarianz wird successive geführt, indem wir zunächst zeigen,daß $R$ nicht geändert wird, wenn wir $x$ durch $-\frac{1}{2}$ ersetzen. Wir erhalten dann, indem wir die Funktionen ganz machen :

$$z^n F(-\tfrac{1}{2}) = F_1(z),\ z^\nu \Phi(-\tfrac{1}{2}) = \Phi_1(z)$$

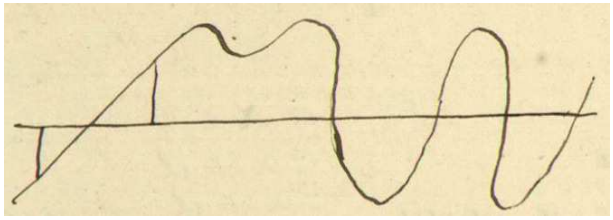
$$R(F_1, \Phi_1) = a_n^\nu \prod_{k=1}^n \Phi_1(-\tfrac{1}{x_k}) = a_n^\nu \prod_{k=1}^n (-1)^\nu x_k^{-\nu} \Phi(x_k)$$

$$= (-1)^{n\nu} a_n^\nu \tfrac{1}{(x_1 \ldots x_n)^\nu} \prod_{k=1}^n \Phi(x_k) = a_0^\nu \prod_{k=1}^n \Phi(x_k) = R(F, \Phi)$$

Ebenso erkennen wir aus der darstellung : $R = a_0^\nu \alpha_0^n \prod_{k,h}(x_k - \xi_h)$, daß $R$ ungeändert bleibt, wenn wir in $F$ und $\Phi$ $x$ um eine konstante Größe vermehren. Durch successive Anwendung dieser beiden Operationen können wir aber eine gebrochene
103|104 lineare Funktion $z = \frac{ax+b}{cx+d}$ herstellen, | welche für $x$ gesetzt $R$ nicht ändert, wenn man die Substitutionsdeterminante $ad - bc = 1$ setzt. Bezeichnen wir für diese Substitution, nachdem wir $F$ und $\Phi$ wieder ganz gemacht haben, die Ceffizienten mit $a'_0 \ldots a'_n; \alpha'_0 \ldots \alpha'_\nu$, so ist : $R(a_0 \ldots a_n; \alpha_0 \ldots \alpha_n) = R(a'_0 \ldots a'_n; \alpha'_0 \ldots \alpha'_n)$.

Die $a'_0 \ldots a'_n; \alpha'_0 \ldots \alpha'_n$ setzen sich aus den $a_0 \ldots a_n; \alpha_0 \ldots \alpha_n$ zusammen, sodaß diese Gleichung als eine reine Identität aufzufassen ist. Sie gilt daher auch, wenn die $a$ und $\alpha$ beliebige Gleichungskoeffizienten sind, d. h. auch wenn wir die Zerlegbarkeit von $F$ und $\Phi$ nicht voraussetzen.

§44. Wir kehren nun zu der Kettenbruchentwicklung des §38 zurück, um hieraus auf den Sturm'schen Satz zu gelangen. Wir wollen im Folgenden annehmen, der Begriff der Wurzeln einer Gleichung sei bereits definirt. Auf die strenge Begründung und Fundamentirung dieses Begriffes werden wir später noch unabhängig von den jetzt folgenden Entwicklungen zurückkommen. Wir knüpfen an die geometrische Anschauung an, welche uns lehrt, daß wenn die Curve $y = F(x)$, bald oberhalb, bald unterhalb der Abszissenaxe verläuft, sie dieselbe schneiden muß, falls die Curve stetig verläuft.

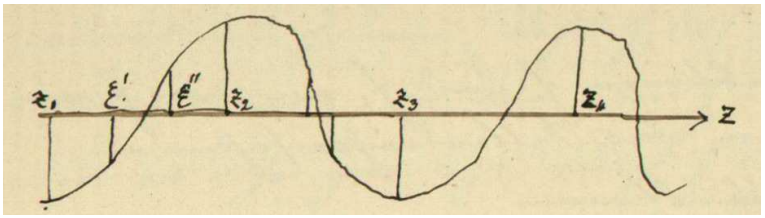


Der Begriff dieser geometrischen Stetigkeit ist aus der Anschauung klar, analytisch werden wir die Funktion $F(x)$ als an der Stelle $x$ stetig bezeichnen, wenn $F(x_1) - F(x_2)$, wo $x_1 < x < x_2$ sei, kleiner gemacht werden kann als jede beliebig kleine Größe, dadurch daß man $x_1$ und $x_2$ beliebig nahe an $x$ heranrückt und wenn diese Differenz dann auch kleiner als diese beliebig kleine Größe bleibt. So einfach dieser Begriff der Stetigkeit ist, so ist er doch so modern, daß er bei *Gauss* noch nicht vorkommt. Er besaß nur die geometrische Anschauung und den Begriff der geometrischen Stetigkeit.

Wir wollen nun, um auf diese geometrische Anschauung der Wurzeln von $F(x) = 0$ den *Sturm*'schen Satz aufzubauen, ein Zeichen definiren, welches wir durch $[A]$ bezeichnen wollen und welches $-1, 0$ oder $+1$ vorstellen soll, je nachdem die als reell vorausgesetzte Größe $A$ negativ, gleich null, oder positiv ist. Haben dann $a$ und $b$
104|105 gleiches Vorzeichen, | so folgt $\frac{1}{2}\{[a] + [b]\} = [a + b]$. Setzen wir $b = -c$ und berücksichtigen, daß : $-[c] = [-c]$ ist, so folgt : $\frac{1}{2}\{[a] - [c]\} = [a-c]$, wobei also vorausgesetzt wird, daß $a$ und $c$ entgegengesetztes Zeichen haben.

Angenommen nun, wir wüßten, gleichviel woher, daß zwischen $z_1$ und $z_2$, wo

$z_1 < z_2$ sei, die Curve $y = F(z)$ einmal, aber nur einmal die Achse passirt, daß also zwischen $z_1$ und $z_2$ eine Wurzel von $F(z) = 0$ läge.




Dann sind also $F(z_1)$ und $F(z_2)$ von verschiedenem Zeichen und es ist :

$$\frac{1}{2}\{[F(z_2)] - [F(z_1)]\} = [F(z_2) - F(z_1)],$$

einerlei ob $F(z_1)$ positiv und $F(z_2)$ negativ ist oder umgekehrt. Diese Gleichung bleibt bestehen, wie nahe wir auch $z_1$ und $z_2$ zusammenrücken lassen, wenn nur die zugehörigen Funktionswerthe noch verschiedenes Zeichen haben.

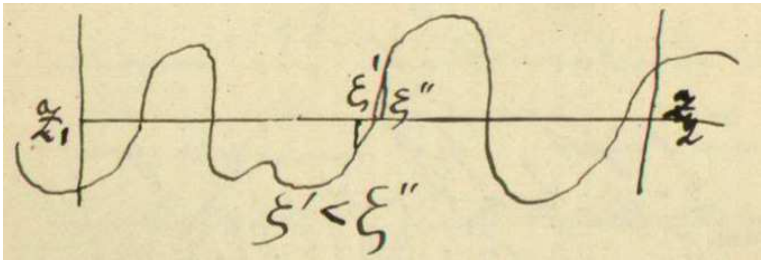
$\zeta'$ und $\zeta''$ mögen 2 beliebig nahe Werthe der Variabeln sein, so daß noch $F(\zeta')$ das Zeichen von $F(z_1)$ und $F(\zeta'')$ das von $F(z_2)$ besitzt. $(z_1 < \zeta' < \zeta'' < z_2)$. Dann ist auch :

$$[F(\zeta'') - F(\zeta')] = \frac{1}{2}\{[F(z_2)] - [F(z_1)]\}.$$

Wir nehmen nun ein zweites Intervall $z_2$ bis $z_2[z_3]$, wo $z_2 < \zeta_1' < \zeta_1'' < z_3$ sei und wo zwischen $\zeta_1'$ und $\zeta_1''$ wiederum eine Wurzel liegen möge, dann folgt ebenso :

$$[F(\zeta_1'') - F(\zeta_1')] = \frac{1}{2}\{[F(z_3)] - [F(z_2)]\}.$$

Fahren wir für folgende Intervalle so fort und summiren alle diese Gleichungen, so haben sich der rechten Seite alle $F(z)$ weg, außer den beiden, welche den Grenzen entsprechen, welche wir jetzt mit $z_1$ und $z_2$ $(z_1 < z_2)$ bezeichnen wollen.




Wir erhalten :

$$\sum_{\zeta}[F(\zeta'') - F(\zeta')] = \frac{1}{2}\{[F(z_2) - F(z_1)]\}.$$

Die Summe auf der linken Seite ist zu erstrecken über alle $\zeta'$ und $\zeta''$, zwischen denen die Funktion $F(\zeta)$ einmal ihr Zeichen gewechselt hat. Dabei ist $z_2 >$ als das größte $\zeta''$ und $z_1$ kleiner als das kleinste $\zeta'$. Nehmen wir nun $\zeta'$ und $\zeta''$ beliebig nahe an einander an, doch so weit natürlich, daß $F(\zeta')$ und $F(\zeta'')$ verschiedenes Zeichen haben, so kommt :

$$[F(\zeta'') - F(\zeta')] = \left[\frac{F(\zeta'') - F(\zeta')}{\zeta'' - \zeta'}\right] = [F'(\zeta)],$$

wenn wir mit $\zeta$ den zwischen den unendlich nahe aneinander liegenden Punkt bezeichnen, wofür $F(\zeta) = 0$ wird.

105|106 | Wir haben also :

$$\sum_{\zeta}[F'(\zeta)] = \frac{1}{2}\{[F(z_2)] - [F(z_1)]\}$$

wo die Summe zu erstrecken ist über alle $\zeta$, für welche $z_2 > \zeta > z_1$ ist und für welche $F(\zeta) = 0$ ist. Dieses Resultat hat den Charakter des *Cauchy*'schen Satzes, das Verhalten der Funktion an den Grenzen gibt uns Aufschluß über ihr Verhalten zwischen denselben.

Wir nehmen nun zwei ganze Funktionen $f$ und $f_1$ an, wo $f$ von höherem Grade sei als $f_1$ und wenden das Kettenbruchverfahren an. Wir erhalten dann die Reihe der Restfunktionen

$$f, f_1, f_2, f_3, ..., f_{h-1}, f_h, f_{h+1}, ..., f_\nu,$$

wo $f_\nu$ vom Grade 0 ist. Setzen wir jetzt, indem wir $x$ für $z$ schreiben :

$$F(x) = f_{h-1}(x)f_h(x)$$

und bezeichnen die reellen Wurzeln von $f_{h-1} = 0$ resp. $f_h = 0$ mit $\xi^{h-1}$ resp. $\xi^h$, so ist : $F'(x) = f'_{h-1}f_h + f_{h-1}f'_h$ und da wir in obiger Gleichung nur solche Werthe der Variabeln haben, für welche $F$ verschwindet, so folgt :

$$\sum_{\xi}[F'(\xi)] = \sum_{\xi^{h-1}}[f'_{h-1}(\xi^{h-1})f_h(\xi^{h-1})] + \sum_{\xi^h}[f_{h-1}(\xi^h)f'_h(\xi^h)].$$

Also ist :

$$\sum_{\xi^{h-1}}[f'_{h-1}(\xi^{h-1})f_h(\xi^{h-1})]+\sum_{\xi^h}[f_{h-1}(\xi^h)f'_h(\xi^h)] = \frac{1}{2}[f_{h-1}(x_2)f_h(x_2)]-\frac{1}{2}[f_{h-1}(x_1)f_h(x_1)]$$

Summiren wir nun über alle $h$ von $h = 1$ bis $h = r$, indem wir $f_0 = f$ setzen. Nun ist doch : $f_{h-1}(x) - g_h(x)f_h(x) + f_{h+1}(x) = 0$. Setzen wir hierin die Wurzeln $\xi^h$ von $f_h$ ein, so folgt :

$$f_{h-1}(\xi^h) + f_{h+1}(\xi^h) = 0.$$

Man sieht also, daß in der Summe auf der linken Seite, die man erhält, wenn man die Summation nach $h$ ausführt, sich immer zwei Summanden wegheben, so daß nur das Anfangsglied $\sum_\xi[f'(\xi)f_1(\xi)]$ und das Endglied $\sum_{\xi^r}[f_{r+1}(\xi^r)f'_r(\xi^r)]$ dieser Summe übrig bleibt. Es ist also [9] :

$$\sum_{\xi}[f'(\xi)f_1(\xi)]-\sum_{\xi^r}[f_{r+1}(\xi^r)f'_r(\xi^r)] = \sum_{h=1}^{r}\left\{\frac{1}{2}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}[f_{h-1}(x_1)f_h(x_1)]\right\}.$$

Angenommen nun wir kennen die Wurzeln $\xi^r$ von $f_r(x) = 0$ (es ist hier immer nur von reellen Wurzeln die Rede), so können wir die zweite Summe auf der linken Seite
106|107 ausführen. Aber dies können wir immer erreichen, | denn da die Grade von $f_r$ mit wachsendem $r$ abnehmen, so können wir einfach $r$ so groß wählen , daß $f_r(x) = 0$ eine Gleichung (etwa eine quadratische oder lineare) ist, deren Wurzeln wir kennen. Die rechte Seite läßt sich für gegebene $x_1$ und $x_2$ ebenfalls ausführen. Richten wir es nun so ein, daß $f_1(x)$ gleiches Vorzeichen mit $f'(x)$ hat, was wir ja können, da wir über $f_1(x)$ willkürlich zu disponiren haben und wir es also gradezu gleich $f'(x)$ setzen können, so stellt

$$\sum_{\xi}[f'(\xi)f_1(\xi)]$$

die Anzahl der zwischen $x_1$ und $x_2$ liegenden reellen Wurzeln von $f(x) = 0$ dar. Da

9. *N.d.T.* Il s'agit en fait ici de $\sum_\xi[f'(\xi)f_1(\xi)] + \sum_{\xi^r}[f_{r-1}(\xi^r)f'_r(\xi^r)]$

alles übrige in der Gleichung bekannt ist, so haben wir diese Anzahl jetzt bestimmt. Dies ist der Inhalt des *Sturm*'schen Satzes.

Aber der *Sturm*'sche Satz erlaubt nicht nur die Anzahl der zwischen $x_1$ und $x_2$ liegenden reellen Wurzeln zu bestimmen, sondern gibt auch eine Methode, diese Wurzeln numerisch zu berechnen. Lassen wir nämlich $x_1$ und $x_2$ zusammenrücken, bis nur noch eine Wurzel zwischen ihnen liegt, so können wir diese in immer engern Grenzen $x_1$ und $x_2$ einschließen, bis wir sie mit jeder gewünschten Genauigkeit bestimmt haben.

Wenn wir oben $f_1$ noch willkürlich ließen und nur so annahmen, daß sein Zeichen mit $f'$ übereinstimmt für alle zu betrachtenden $\xi$, so wurde dabei die nothwendige Voraussetzung gemacht, daß $f_1(x)$ keinen gemeinschaftlichen Theiler mit $f(x)$ besitze, weil sonst ja $f_1(\xi)$ für gewisse $\xi$ verschwindet. Da :

$$[f'(\xi)f_1(\xi)] = \left[\frac{f'(\xi)f_1(\xi)}{f_1^2(\xi)}\right] = \left[d\frac{f(\xi)}{f_1(\xi)}\right]$$

ist, so können wir auch die erste Summe der linken Seite unserer Gleichung durch $\sum_\xi \left[d\frac{f(\xi)}{f_1(\xi)}\right]$ ersetzen. Insbesondere nun können wir $r = \nu$ setzen, sodaß wir alle Restfunktionen $f, f_1, f_2, ...f_\nu$ haben. Die Wurzeln von $f_\nu = 0$ existiren dann nicht, denn $f_\nu$ ist eine Constante, es fällt daher die zweite Summe der linken Seite unserer Gleichung weg und $\sum_\xi \left[d\frac{f}{f_1}\right]$ ist durch die Summe auf der rechten Seite der Gleichung ausgedrückt, die jetzt von $h = 1$ bis $h = \nu$ zu nehmen ist.

Wir wollen nun noch dem Satze die Fassung geben, die ihm *Sturm* | ge- 107|108
geben hat. Es ist : $\sum_{h=1}^{\nu}[f_{h-1}(x)f_h(x)]$ die Anzahl der Fälle, wo in der Reihe $f, f_1, f_2, ...f_{h-1}, f_h, f_{h+1}, ...f_\nu$, welche man die *Sturm*'sche Reihe nennt, zwei auf einander folgende Funktionen $f_{h-1}$ und $f_h$ gleiches Zeichen haben, vermindert um die Anzahl der Fälle, wo zwei solcher Funktionen ungleiche Zeichen haben, oder kürzer, jene Summe ist gleich der Differenz der Anzahl der Zeichenfolgen und der Anzahl der Zeichenwechsel in der *Sturm*'schen Reihe. Bezeichnen wir die Anzahl der Zeichenfolgen durch $s$, die der Zeichenwechsel durch $m$, so ist :

$$\sum_{h=1}^{\nu}[f_{h-1}(x)f_h(x)] = s - m \ , \ s + m = \nu + 1.$$

Bezeichnen wir für $x = x_1$ resp. $x = x_2$ die Auswahl der Zeichenfolgen und Wechsel in der *Sturm*'schen Reihe durch $s_1$ und $m_1$ resp. $s_2$ und $m_2$, so folgt :

$$\sum_{h=1}^{\nu}\left\{\frac{1}{2}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}[f_{h-1}(x_1)f_h(x_1)]\right\} = \frac{1}{2}(s_2 - m_2 - (s_1 - m_1)).$$

Nun ist aber $s_2 + m_2 = \nu + 1 = s_1 + m_1$ also ist :

$$\frac{1}{2}(s_2 - m_2 - s_1 + m_1) = m_1 - m_2.$$

Es ist daher : $\sum_{x_1<\xi<x_2}\left[d\frac{f}{f'}\right] = m_1 - m_2$, d. h. die Anzahl der zwischen den Grenzen $x_1$ und $x_2$ liegenden reellen Wurzeln von $f(x) = 0$ ist gleich dem Verluste an Zeichenwechseln, den die *Sturm*'sche Reihe erleidet, wenn man von $x = x_1$ zu $x = x_2$ übergeht. Dies ist die Fassung, welche *Sturm* seinem berühmten Satze gegeben hat. Wir lassen jetzt den Beweis folgen, welchen *Sturm* zu seinem Satze angewendet hat. Wir denken uns die Reihe der *Sturm*'schen Funktionen $f, f_1, f_2, ...f_r$ gebildet, wo wir $r$ so groß wählen, daß $f_r$ keine reellen Wurzeln hat (nöthigenfalls können wir ja $r = \nu$ setzen).

Lassen wir nun $x$ von $x_1$ bis $x_2$ wachsen, so bietet für $x = x_1$ die *Sturm*'sche Reihe $m_1$ Zeichenwechsel für $x = x_2$ solcher $m_2$ dar. Es fragt sich nun, wie sich bei dem Uebergange von $x_1$ zu $x_2$ die Vorzeichen der *Sturm*'schen Funktionen ändern.

Aus $f_{h-1} - g_h f_h + f_{h+1} = 0$ geht zunächst hervor, daß keine zwei auf einander folgenden $f$ für denselben Werth von $x$ verschwinden können, denn sonst müßten alle $f$ verschwinden und $f$ und $f_1$ hâtten dann einen gemeinschaftlichen Theil-
108|109 bar, was wir ausgeschlossen haben. da nun für | $x = \xi^h$ die Gleichung besteht, $f_{h-1}(\xi^h) = -f_{h+1}(\xi^h)$, so haben auch noch unmittelbar vor $\xi^h$ und unmittelbar nach $\xi^h$ $f_{h-1}$ und $f_{h+1}$ entgegengesetztes Zeichen. Die Zeichen von $f_{h-1}f_h$ und $f_{h+1}$ gehen also beim Durchgang durch $\xi^h$ entweder aus $--+$ in $-++$ (oder aus $-++$ in $--+$) über, oder aber aus $+--$ in $++-$ über (oder auch ungekehrt aus $++-$ in $+--$).

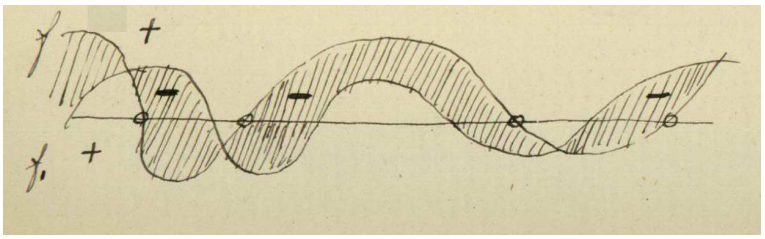
In jedem Falle sehen wir also, daß beim Durchgang von $x$ durch $\xi^h$ ein Zeichenwechsel in der *Sturm*'schen Reihe weder gewonnen noch verloren geht. Ein solcher kann also in der Reihe $f, f_1, f_2, ...f_r$, da $f_r$ fest bleibt, nur beim Passiren durch eine Wurzel $\xi$ von $f = 0$ eintreten. Die Herstellung einer Reihe, in der die Anzahl der Zeichenwechsel, trotz aller Zeichenschwankungen in Innern, nur von den beiden Anfangsgliedern abhängen, ist der geistige Inhalt der *Sturm*'schen Satzes. Bei dem Durchgang durch ein $\xi$ geht $f$ entweder aus dem negativen ins positive oder aus dem positiven ins negative über. Im ersteren Falle ist das Vorzeichen von $f'$ positiv, im letzteren negativ. Wir haben also vor und nach einem $\xi$ entweder den Uebergang von $-+$ in $++$ oder von $+-$ in $++$ [10], in jedem Falle geht ein Zeichenwechsel verloren. Damit ist der Satz bewiesen, daß zwischen $x_1$ und $x_2$ so viele reelle Wurzeln liegen, als Zeichenwechsel bei dem Uebergange von $x_1$ zu $x_2$ verloren gehen, diese Anzahl ist $m_1 - m_2$.

Da wir bei den *Sturm*'schen Funktionen immer nur die Vorzeichen brauchen, so gibt es natürlich unendlich viele Reihen, die alle dasselbe leisten, wie die Reihe $f, f_1, f_2, f_3, ...f_r$ wenn nur die Vorzeichen in allen diesen Reihen übereinstimmen.

§45. Wir wollen anknüpfend an die geometrische Anschauung die linke Seite der Gleichung :

$$\sum_{\xi}[f'(\xi)f_1(\xi)] = \frac{1}{2}\sum_{h=1}^{r}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}\sum_{h=1}^{r}[f_{h-1}(x_1)f_h(x_1)]$$

interpretiren. Wir denken uns die beiden parabolischen Curven $y = f_1(x)$ [und $y = f(x)$] konstruirt.




109|110 Wir können für jede der beiden Curven einen äußern und einen | inneren Theil der Ebene unterscheiden, je nachdem in einem solchen $y - f(x)$ resp. $y - f_1(x)$ positiv oder negativ ist. Ebenso können wir für die Funktion $(y - f(x))(y - f_1(x))$ einen äußern Theil von einem innern Theil der Ebene unterscheiden, je nachdem das Produkt positiv oder negativ ist. In unserer Zeichnung ist der schraffierte Theil der Ebene als negativ, alles übrige als positiv zu betrachten.

Gehen wir nun auf der $x$-Achse vorwärts, so ist $f'(\xi)f_1(\xi)$ der Ausdruck für die

10. *N.d.T.* Il semble y avoir une erreur : de $+-$ à $--$

Zunahme der Funktion :

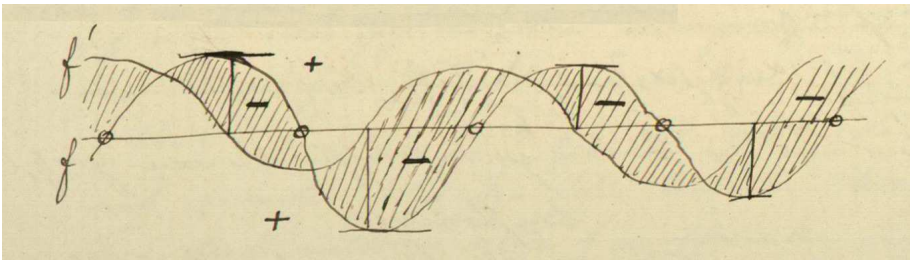
$$F(x,y) = (y - f(x))(y - f_1(x))$$

in dem Punkte $\xi$, für welchen $f(\xi) = 0$ ist. Ist daher $f'(\xi)f_1(\xi)$ positiv, so nimmt $(F(x,y))_{y=0} = ff_1$ zu, ist der Ausdruck negativ, so nimmt $ff_1$ ab.

Es ist daher $[f'(\xi)f_1(\xi)] = +1$ oder $= -1$, je nachdem wir auf der Achse beim Durchschreiten von $\xi$ aus dem Innern ins Aeußere oder aus dem Aeußern ins Innere gelangen.

So ist daher $\sum_\xi [f'(\xi)f_1(\xi)]$ gleich der Anzahl der Austritte vermindert um die Anzahl der Eintritte. Diese Differenz ist also andrerseits gleich dem Verluste an Zeichenwechseln d.h. gleich $m_1 - m_2$.

Ist im Speziellen $f_1 = f'$ so ist $[f'^2(\xi)]$ immer positiv, es gibt also nur Austritte und diese Anzahl ist dann natürlich gleich der Zahl der zwischen $x_1$ und $x_2$ befindlichen Wurzeln $\xi$ von $f(\xi) = 0$. Der *Sturm*'sche Satz ist also nur ein spezieller Fall des vorhergehenden allgemeineren Satzes, den man Sylvester verdankt, welcher ihn gleich nach dem *Sturm*'schen Satze entdeckte, indem er statt der speziellen aus $f$ und $f'$ entstehenden Reihe die allgemeinere aus dem Kettenbruchverfahren für $f$ und $f_1$ hergeleitete setzte. Für $f_1 = f'$ folgt die Interpretation des *Sturm*'schen Satzes auf direkt auch der Figur :




| §46. Wir lassen nun einige Anwendungen des *Sturm*'schen Satzes folgen und 110|111
betrachten zunächst die trinomische Gleichung :

$$f(x) = x^n + nax + b = 0.$$

Wir setzen, da ja $n$ positiv ist, $f_1(x) = \frac{1}{n}f'(x) = x^{n-1} + a$. Dann lautet die *Sturm*'sche Reihe :

$$\begin{aligned} f(x) &= x^n + nax + b \\ f_1(x) &= x^{n-1} + a \\ f_2(x) &= -(n-1)ax - b \\ f_3(x) &= \tfrac{(-1)^n b^{n-1} - (n-1)^{n-1}a^n}{a^{n-1}} = \tfrac{\Theta}{a^{n-1}}. \end{aligned}$$

Um alle reellen Wurzeln zu erhalten, setzen wir $x_1 = -\infty, x_2 = +\infty$ und unterscheiden, ob $n$ grade oder ungrade ist.

1. $n$ sei ungrade, so erhalten wir :
   für $x = -\infty$ die Reihe der Vorzeichen $-, +, [a], [a^{n-1}\Theta]$
   für $x = +\infty$ die Reihe der Vorzeichen $+, +, -[a], [a^{n-1}\Theta]$
   Dann ist $\frac{1}{2}\sum_{h=1}^{3}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}\sum_{h=1}^{3}[f_{h-1}(x_1)f_h(x_1)] = 1 - [a] - [a^n\Theta]$

Ist nun $a > 0$ so liefert dies $-[\Theta]$ und zwar muß dann $\Theta < 0$ sein, wie nicht nur aus $f_3$ folgt, sondern auch daraus, weil sonst die Anzahl der reellen Wurzeln negativ wäre, was keinen Sinn hat. Für $a > 0$ ist also die Anzahl der reellen Wurzeln gleich 1. Für $a < 0$ erhalten wir $2 + [\Theta]$ und da jetzt $\Theta$ sowohl positiv als negativ sein kann, so ist die Anzahl der reellen Wurzeln gleich 1 oder gleich 3. Fassen wir die beiden Fälle zusammen, so folgt, daß für ungrades $n$ die Anzahl aller reellen Wurzeln von $x^n + nax + b = 0$ gleich 1 oder gleich 3 ist.

2. Ist $n$ grade, so folgt für $x_1 = -\infty, x_2 = +\infty$ :
$\frac{1}{2}\sum_{h=1}^{3}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}\sum_{h=1}^{3}[f_{h-1}(x_1)f_h(x_1)] = 1 - [a]$ d.h. für grades $n$ ist die Anzahl der reellen Wurzeln der trinomischen Gleichungen gleich 0 oder gleich 2. Bezeichnen wir die Anzhal der reellen Wurzeln der trinomischen Gleichungen mit $\mathcal{A}$, so ist für grades und ungrades $n$ : $\mathcal{A} = \frac{3+(-1)^{n-1}}{2} + (-1)^{n-1}[\Theta]$.

---

111|112 | §47. Wenn wir $x_1 = -\infty$ und $x_2 = +\infty$ setzen, so können wir den erweiterten *Sturm*'schen Satz (*pag...*) in einfacherer Form darstellen. Bezeichnen wir die Anzahl der Eintritte und der Austritte mit $E$ resp. $A$, so hatten wir :

$$A - E = \frac{1}{2}\sum_{h=1}^{r}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}\sum_{h=1}^{r}[f_{h-1}(x_1)f_h(x_1)]$$

Nun sei

$$f(x) = cx^n + ...$$
$$f_k(x) = c_k x^{n_k}$$

Da wir $x_1 = -\infty$ und $x_2 = +\infty$ setzen, so brauchen wir nur die höchsten Potenzen zu berücksichtigen. Für $x = -\infty$ erhalten wir dann die Reihe :

$$(-1)^n c, (-1)^{n_1} c_1, (-1)^{n_2} c_2, ..., (-1)^{n_k} c_k, ...$$

Für $x = +\infty$ kommt :

$$c, c_1, c_2, ..., c_k, ...$$

Die obige Summe ist also zu erstrecken über Glieder der Form :

$$\frac{1}{2}[c_{k-1}c_k] - \frac{1}{2}(-1)^{n_{k-1}+n_k}[c_{k-1}c_k].$$

Nun fallen alle Glieder, für welche $n_{k-1} + n_k$ grade ist, weg und wir erhalten daher :

$$A - E = \sum[c_{k-1}c_k.]$$

Die Summe erstreckt über alle $k$ für welche $n_{k-1} + n_k \equiv 1 \bmod. 2$ ist.

Die Frage nach den zwischen $x_1$ und $x_2$ liegenden Wurzeln ist durchaus nicht spezieller als die Frage nach den zwischen $-\infty$ und $+\infty$ liegenden Wurzeln, da wir jederzeit die Strecke von $x_1$ bis $x_2$ auf die ganze Grade von $-\infty$ bis $+\infty$ abbilden können. Setzen wir nämlich in $f(x)$ für $x$ die Funktion $\frac{x_1+x_2z^2}{1+z^2}$ und schaffen die Nenner weg, so erhalten wir eine Funktion $F(z)$. Diese verschwindet für alle $\xi$, welche den Nullstellen $\xi$ von $f(x)$ entsprechen. Läßt man nun $z$ von $-\infty$ bis $+\infty$ wandern, so beschreibt $x$ die Strecke von $x_2$ bis $x_1$ und wieder zurück nach $x_2$. Die Anzahl der reellen Wurzeln von $F(z) = 0$ ist daher gleich der doppelten Anzahl der zwischen $x_1$ und $x_2$ liegenden Wurzeln von $f(x) = 0$. In Folge dieser Abbildbarkeit dieser Strecke auf die ganze Grade ist es daher gleichgültig, ob wir nach allen reellen Wurzeln von $F(z) = 0$ oder nach den zwischen $x_1$ und $x_2$ liegenden von $f(x) = 0$ fragen. Beide Fragen sind, wenn man vermöge der angegebenen Substitution von $f(x)$ zu $F(z)$ oder umgekehrt übergeht, vollständig äquivalent.

112|113 | §48. Der *Sturm*'sche Satz gibt uns ein einfaches Mittel an die Hand, die Realität der Wurzeln der Gleichungen aller Grade mit reellen Coeffizienten zu untersuchen.

Die Untersuchung der linearen und quadratischen Gleichungen ist zu trivial, es wären also zunächst die kubischen Gleichungen zu betrachten. Diese sind bereits eingehend studirt und wir haben gesehen, daß die beiden einzigen möglichen Fälle (3 reelle Wurzeln oder nur eine) durch das Vorzeichen der Diskriminante von einander unterschieden werden (siehe *pag....*). Wir wenden uns daher jetzt zu den biquadratischen Gleichungen mit reellen Coeffizienten.

Diese können stets auf die Form gebracht werden :

$$f(u) = u^4 - 6zu^2 + 4yu + 3x.$$

Wir betrachten dann $x, y, z$ als die Coordinaten eines Punktes im Raume und können dadurch jedem Punkte eine ganz bestimmte Gleichung eindeutig zuordnen und umgekehrt. Wir müssen dann 3 von einander getrennte Gebiete des Raumes erhalten [11], so daß für alle Punkte des ersten Gebietes die entsprechende biquadratische Gleichung 4 reelle Wurzeln hat, für alle Punkte des zweiten Gebietes die entsprechende Gl. deren 2 und für alle Punkte des dritten Gebietes die entsprechende Gl. gar keine reellen Wurzeln besitzt.

Diese 3 Gebiete werden von einander durch Flächen getrennt und da der Uebergang von einem Paar imaginärer Wurzeln zu einem Paare reeller Wurzeln, nur durch ein Paar gleicher Wurzeln stattfinden kann, so erkennt man, daß alle Punkte dieser Grenzflächen Gleichungen mit gleichen Wurzeln entsprechen müssen und umgekehrt, daß allen Gleichungen mit gleichen Wurzeln Punkte entsprechen, die auf deren Grenzflächen liegen. Alle Punkte der Grenzflächen entsprechen daher Gleichungen mit verschwindender Diskriminante. Diese ist eine ganze rationale Funktion der Coordinaten $x, y, z$ ; setzt man dieselbe gleich Null, so erhalten wir also eine Fläche, welche die Diskriminantenfläche der biquadratischen Gl. genannt wird. Die Diskriminantenfläche ist es also, welche allein die 3 genannten Gebiete des Raumes von einander trennen muß. Untersuchen wir nun mit Hülfe des *Sturm*'schen | Satzes diese 3 Gebiete und 113|114
die Diskriminantenfläche. Es ist :

$$f(u) = u^4 - 6zu^2 + 4yu + 3x.$$

Wir wählen : $f_1 = \frac{1}{4}f' = u^3 - 3zu + y$.
Wenn wir in den folgenden Restfunktionen immer an und für sich positive Faktoren weglassen, so erhalten wir folgende Sturm'she Reihe :

$$\begin{aligned} f &= u^4 - 6zu^2 + 4yu + 3x \\ f_1 &= u^3 - 3zu + y. \\ f_2 &= zu^2 - yu - x \\ f_3 &= D_1 u - y(x + z^2) \\ f_4 &= D. \end{aligned}$$

Dabei ist $D = -y^4 - 2y^2z(3x - z^2) + x(x - 3z^2)^2 = (x + z^2)^3 - (3xz + y^2 - z^3)^2 = 0$
Ferner ist :

$$D_1 = 3z^3 - xz - y^3.$$

Nach *pag....* ist die Anzahl der reellen Wurzeln gleich :

$$\sum [c_{k-1}c_k] = 1 + [z] + [zD_1] + [D_1D].$$

Wir sehen zunächst von dem Falle $D = 0$ ab, weil wir es ja in diesem Falle nur mit einer uneigentlichen biquadratischen Gleichung zu thun haben.

1. Ist $D < 0$, so ist :

$$1 + [Z] + [ZD_1] + [D_1D] = 1 + [z] + [zD_1] - [D_1] = 1 + [z] - [D_1]\{1 - [z]\}.$$

Ist nun $z > 0$, so ist diese Anzahl gleich 2, ist $z < 0$, so ist die Anzahl ebenfalls gleich 2, denn in diesem letzteren Falle muß $D_1$ nothwendig ebenfalls negativ sein, weil sonst die Anzahl der reellen Wurzeln eine negative Zahl wäre. Daß $D_1 < 0$ ist, wenn $D < 0$ und $z < 0$, ergibt sich natürlich auch aus den Ausdrücken für $D$ und $D_1$. Ist also $D < 0$, so hat die biquadratische Gleichung stets zwei reelle und ein Paar imaginärer Wurzeln.

11. *N.d.T* Dans le manuscrit il est écrit *enthalten*.

2. Ist $D > 0$, so ist :

$$1 + [Z] + [ZD_1] + [D_1 D] = \{1 + [z]\}\{1 + [D_1]\}.$$

Diese Zahl repräsentirt 0 oder 4.

114|115 Ist $z > 0$ und $D_1 > 0$, so hat die biquadratische Gl. 4 reelle Wurzeln, | ist dagegen $z < 0$ oder $D_1 < 0$, so hat die Gl. gar keine reellen Wurzeln. Bezeichnen wir nun die 3 Gebiete des Raumes, welche Gleichungen mit keinem, einem oder zwei Paaren imaginärer Wurzeln entsprechen, mit $G_0, G_1, G_2$, so ist also :

$G_0$ charakterisirt durch $D > 0$ und entweder $z < 0$ oder $D_1 < 0$.
$G_1$ charakterisirt durch $D < 0$.
$G_2$ charakterisirt durch $D > 0$ und $z > 0$ und $D_1 > 0$.

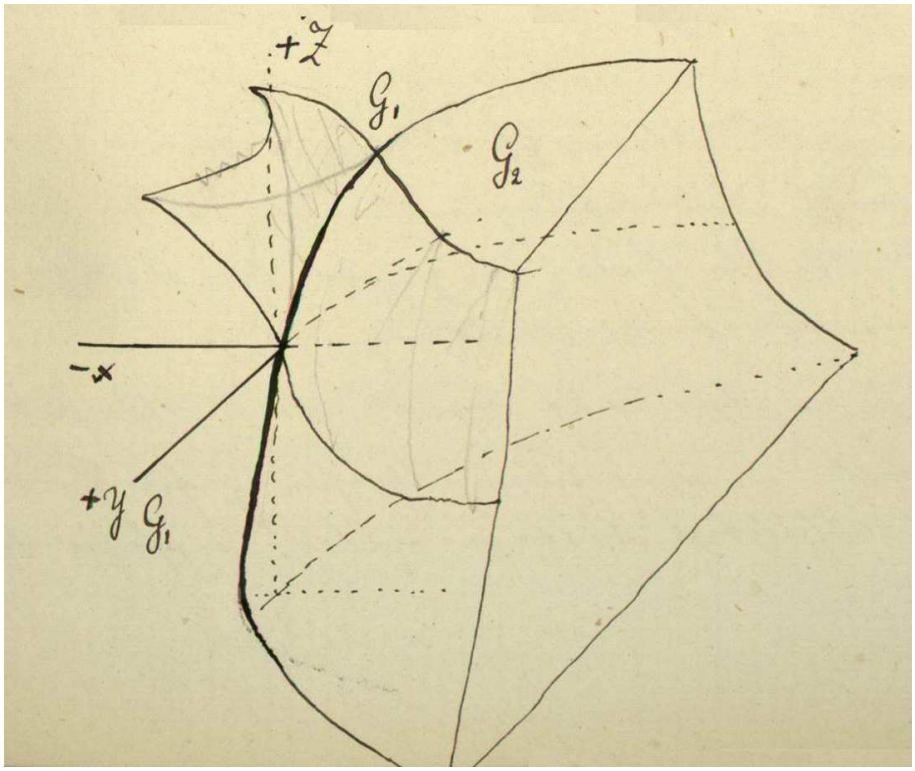
Wenn wir die Fläche $D = 0$ passiren, so gelangen wir von $G_1$ nach $G_0$ resp. $G_2$ oder umgekehrt von $G_0$ resp. $G_2$ nach dem Gebiete $G_1$. Wir bezeichnen $G_1$ als den äußern Raum, $G_0$ und $G_2$ als die beiden inneren Räume. Die Diskriminante $D$ ist das Resultat der Elimination von $u$ aus $f(u)$ und $f'(u)$. $f(u) = 0$ und $f'(u) = 0$ sind aber aufzufassen als eine Ebene mit dem Parameter $u$ und die Nachbarebene, die man erhält, wenn man $u$ unendlich wenig variirt. Es sind daher $f(u) = 0$ und $f'(u) = 0$ die Gleichungen der Schnittlinie dieser beiden Ebenen, und $D = 0$ ist daher der Ort dieser Schnittlinie und folglich eine abwickelbare Fläche. Ihre Rückkehrkurve hat die Gleichungen :

$$x + z^2 = 0 \text{ und } y^2 - 4z^3 = 0.$$

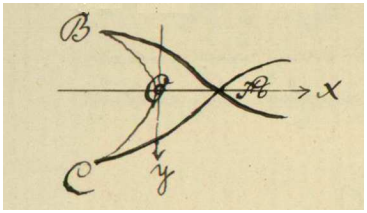
Ferner bezitzt sie eine Doppelkurve mit den Gleichungen :

$$y = 0 \text{ und } x - 3z^2 = 0.$$

Die Rückkehrkurve und die Doppelkurve bilden den vollständigen Durchschnitt der beiden Flächen $D = 0$ und $D_1 = 0$. Die Fläche $D = 0$ scheidet also das äußere Gebiet $G_1$ von den beiden inneren Gebieten $G_0$ und $G_2$. Die Diskriminantenfläche hat ungefähr folgende Gestalt :



| Unterhalb der $x-z$ Ebene bleibt die Fläche ganz auf der positiven Seite der $y-z$ Ebene und schmiegt sich derselben immer enger an, indem sie sich mit wachsendem negativem $z$ immer weiter öffnet. 115|116




Oberhalb der $x-y$ Ebene zeigt ein horizontaler Schnitt ungefähr die Gestalt der nebenstehenden Figur. Dabei beschreibt $O$ den positiven Theil der $z$ Achse, $\mathcal{A}$ beschreibt eine Parabel, nämlich die in der $x-z$ Ebene liegende Doppelkurve und $\mathcal{B}$ und $\mathcal{C}$ endlich beschreiben die Rückkehrkurve. In der $x-y$ Ebene fällt der Scheitel der Parabel mit den beiden Spitzen $\mathcal{B}$ und $\mathcal{C}$ zusammen. Nur der oberhalb dieser Ebene befindliche Theil der Parabel gehört der Diskriminantenfläche an, der untere Theil verläuft isolirt in dem Raume $G_2$, ist aber als mit zu der Diskriminantenfläche gehörig zu betrachten. Dem oberen Theile der Parabel entsprechen, da $z > 0$ ist, die 4 reellen Wurzeln $+\sqrt{3z}, +\sqrt{3z}, -\sqrt{3z}, -\sqrt{3z}$, wo die Quadratwurzel immer mit demselben Zeichen zu nehmen ist, während dem untern Theile, wo $z < 0$ ist, die 4 imaginären Wurzeln : $+\sqrt{3z}, +\sqrt{3z}, -\sqrt{3z}, -\sqrt{3z}$, entsprechen. Die Doppelkurve (die allerdings zur Hälfte eine isolirte Curve ist) entspricht also den Gleichungen mit 2 Paaren gleicher Wurzeln. Die Rückkehrkurve ist der Ort der Punkte, denen Gleichungen mit 3 gleichen Wurzeln entsprechen.

§49. Wir wollen nun dazu übergehen, nachträglich die Begriffe streng zu begründen, von denen wir im *Sturm*'schen Satze Gebrauch gemacht haben. Es handelt sich zunächst darum, den Begriff der Wurzeln einer Gleichung festzustellen.

Wir nehmen eine ganze Funktion $f(x)$ von $x$ vom Grade $m-1$ an, deren Coeffizienten reell und auf irgend welche Weise definirt seien, dann fragen wir : Welche Bedeutung hat der Ausspruch eine bestimmte Größe $x$ erfüllt die Gleichung ?

Wenn die reellen Coeffizienten auf irgend welche Weise definirt sind, so müssen wir jedenfalls eine positive Große $G$ finden können, welche größer ist als die absoluten Beträge aller Coeffizienten. | Bezeichnen wir den absoluten Betrag irgend einer Größe $a$ durch $|a|$, so ist für irgend ein beliebiges $x$ : 116|117

$$\left\{|x|^{m-1} + ... + |x| + 1\right\} G > |f(x)|$$

Bezeichnen wir mit $\delta$ einen positiven ächten Bruch, so ist also $f(x)$ selbst gleich :

$$f(x) = \delta G \frac{|x|^m - 1}{|x| - 1}.$$

Wir können dies noch verallgemeinern. Liegt eine gewisse Anzahl von Größen zwischen resp. $\pm a', \pm a'', ...$ so liegt die Summe derselben zwischen $\pm(a' + a'' + ...)$. Oder in Zeichen, indem das Symbol $\delta$ immer einen ächten Bruch bedeute :

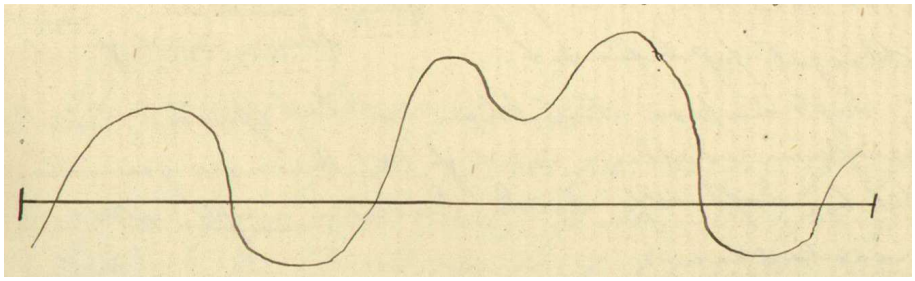
$$\delta' a' + \delta'' a'' + ... = \delta(a' + a'' + ...)$$

Die $a', a'', ...$ sind dabei als positive reelle Größen. Betrachten wir nun alle Funktionswerthe, bei welchen $|x| < \xi$ ist, so ergibt sich, daß die Summe dieser Funktionswerthe gleich ist :

$$\Sigma = \delta G \frac{\xi^m - 1}{\xi - 1}.$$

Dies vorausgeschickt, können wir nun ein solches Intervall abgrenzen, dass außerhalb dasselben sich keine Wurzel mehr befinden. Wir bedienen uns auch hier wieder der geometrischen Anschauung, obwohl wir wissen, daß diese schon das zu beweisende voraussetzt. Wir wollen aber dadurch auch nur die Anschauung unterstützen, ohne

den Beweis darauf zu stützen. Es handelt sich also darum, ein Intervall abzugrenzen,



so daß außerhalb desselben die Abszissenachse nicht mehr von der Curve geschnitten wird

Wir setzen jetzt unsere Funktion als vom Grade $m$ voraus, und vollen also ein Intervall finden, für welches $x^m+....$ das Vorzeichnen von $x^m$ hat. Die auf $x^m$ folgenden Glieder bilden eine Funktion vom Grade $m-1$, wie wir sie ober betrachtete, für welche also eine positive Größe $G$ existirt, so daß der Betrag dieser ganzen Funktion vom Grade $m-1$ kleiner ist als $G\frac{|x|^m-1}{|x|-1}$.

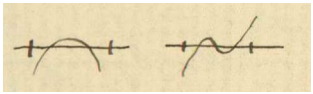
117|118 Wählen wir daher $|x| > G+1$, indem wir z. B. setzen $|x| = G+1+p$, | wo $p$ eine positive Größe ist, so ist der Betrag der auf $x^m$ folgenden Funktion $m-1$ ten Grades kleiner als :

$$\frac{G}{G+p}\left((G+1+p)^m-1\right)$$

also kleiner als der Betrag von $x^m$, denn dieser ist gleich $(G+1+p)^m$.

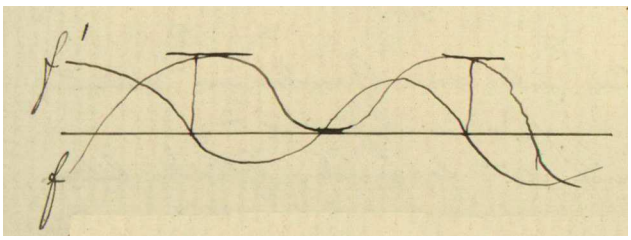
Ist also $x^m + ... = 0$ gegeben, so läßt sich stets eine solche Größe $G$ finden, von der Beschaffenheit, daß die Funktion $x^m + ...$ das Zeichen von $x^m$ hat, sobald $|x| > G+1$ ist. Damit haben wir jetzt den Vortheil gewonnen, unsern Betrachtung auf das endliche Intervall :
$-(G+1)... + (G+1)$ beschränken zu können.

Nachdem nun die Möglichkeit ein solches Intervall aufzufinden, gezeigt ist, handelt es sich darum, solche kleineren Intervalle zu finden, in welchen nur eine Wurzel sich befindet, und solche Punkte dann auch zu isoliren. Dies letztere kann unter Umständen sehr schwierig sein, wie die nebenstehenden Zeichnungen lehren.



Wir müssen also darauf ausgehen, von vornherein aus der Beschaffenheit der Funktion heraus die Grenzen für ein kleines Intervall zu bestimmen, in welchem nur eine oder keine Wurzel liegt. Um diese Untersuchung jetzt in Angriff zu nehmen, d. h. um zu erforschen, wie nahe aneinander müssen die Grenzen eines Intervalls gerückt werden, um sicher zu sein, daß in demselben nur eine oder keine Wurzel sich befindet.

Wir setzen voraus, daß unsere Funktion $m$ ten Grades, welche wir jetzt mit $f(x)$ bezeichnen wollen, mit $f'(x)$ keinen gemeinschaftlichen Theiler besitzt, daß also beide Curven nicht etwa in der folgenden Weise verlaufen :




Diese Voraussetzung läßt sich stets durch Abtrennen eines solchen gemeinschaftlichen Theiler realisiren. Alsdann lassen sich 2 funktionen $\varphi$ und $\psi$ so bestimmen, daß $f\varphi - f'\psi = 1$ ist. Wir sehen also, daß unsere jetzt folgende Methode sich lediglich auf algebraische, nicht aber auf

118|119 transzendente Gleichungen beschränkt. Bilden wir nun die folgenden Ausdrücke : |

$$\frac{f(x[_1]) - f(x_2)}{x_1 - x_2} = f'(x_1) + \frac{x_1 - x_2}{2} f''(x_1) + \ldots \text{ und}$$

$\left\{\frac{f(x[_1]) - f(x_2)}{x_1 - x_2} - f'(x_1)\right\} \frac{2}{x_1 - x_2}$, so bleiben, wenn wir uns auf das oben definirte endliche Intervall von $-(G+1)$ bis $+(G+1)$ beschränken, beide Ausdrücke unter gewissen endlichen Grenzen, wie nahe wir auch $x_1$ und $x_2$ an einander annehmen mögen. Dasselbe gilt für die ganzen Funktionen $\varphi$ und $\psi$. Bezeichnen wir nun die Grenzen von $\varphi$ und $\psi$ mit resp. $\mathcal{A}$ und $\mathcal{B}$ und diejenigen der obigen Ausdrücke mit resp. $\mathcal{C}$ und $\mathcal{D}$ und lassen die kleinen griechischen Buchstaben $\alpha, \beta, \gamma, \delta$ ächte positive oder negative Brüche bedeuten, so daß z. B. $\varphi = \alpha\mathcal{A}, \psi = \beta\mathcal{B}$ wird, und setzen endlich $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, wo $\Theta$ ein ächter Bruch und $\mathcal{M}$ noch unbestimmt sei, so ist $\mathcal{M}$ jetzt so zu bestimmen, daß in dem Intervalle $x_1 - x_2$ sich entweder keine oder nur eine Wurzel von $f(x) = 0$ befindet.

Nun ist :

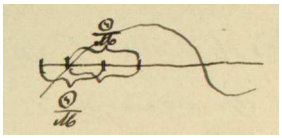
$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(x_1) + \frac{x_1 - x_2}{2} \delta\mathcal{D}.$$

Multipliziren wir mit $\psi$ und berücksichtigen die Gleichungen $f\varphi - f'\psi = 1$, $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, $\varphi = \alpha\mathcal{A}$, $\psi = \beta\mathcal{B}$, so kommt :

1. $-(f(x_1) - f(x_2))\frac{\beta\mathcal{B}\mathcal{M}}{\Theta} + f(x_1)\alpha\mathcal{A} = 1 - \frac{\Theta}{2}\delta\beta\frac{\mathcal{D}\mathcal{B}}{\mathcal{M}}$. Ferner ist :
2. $f(x_1) - f(x_2) = \Theta\gamma\frac{\mathcal{C}}{\mathcal{M}}$.

Die einzige Größe in diesen beiden Gleichungen, über welche wir noch verfügen können, ist $\mathcal{M}$. Wählen wir daher $\mathcal{M}$ zunächst größer als $\mathcal{D}\mathcal{B}$, so liegt die rechte Seite der ersten Gleichung jedenfalls zwischen $\frac{1}{2}$ und $\frac{3}{2}$, ist aber jedenfalls größer als $\frac{1}{2}$.

Betrachten wir also ein solches Intervall $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, wo $\Theta$ von 0 verschieden sei, so können, wenn $\mathcal{M} > \mathcal{B}\mathcal{D}$ ist, nicht gleichzeitig $f(x_2)$ und $f(x_1)$ beliebig klein werden, vorausgesetzt eben, daß $x_1 - x_2$ eine <u>bestimmte</u> Größe hat. Man kann daher wohl Das Intervall $x_1 - x_2$ (ohne seine Größe zu ändern) beliebig auf der $x-$Achse verschieben, dann muß aber stets $f(x_2)$ <u>über</u> einer | gewissen Größe bleiben, wenn 119|120
$f(x_1)$ beliebig klein wird. Dies ist aber das Charakteristische Dafür, daß in jedem so definirten Intervalle $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$ höchstens eine Stelle sich befindet, an welcher $f(x)$ verschwindet.



Unsern zweite Gleichung lehrt nun, daß wir das Intervall $\frac{\Theta}{\mathcal{M}}$ so klein annehmen können, daß, vorausgesetzt daß $f(x_1)$ und $f(x_2)$ entgegengesetztes Zeichen haben, der absolute Betrag von $f(x_1)$ und $f(x_2)$ beliebig klein gemacht werden können. Damit ist also folgendes Resultat gewonnen, welches uns den Begriff der Wurzeln einer Gleichung feststellt : Sei $\Theta$ ein beliebiger ächter Bruch, so können wir stets $\mathcal{M}$ so groß wählen, daß in <u>jedem</u> Intervalle $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, die Funktion $f(x)$ höchstens einmal beliebig klein wird und daß, vorausgesetzt, daß in dem Intervalle (d. h. also an den Endpunkten desselben) ein Zeichenwechsel stattfindet, der Betrag der Funktion an den Endpunkten beliebig klein gemacht werden kann.

Hiernach hat also z. B. die Gleichung $x^2 - 2 = 0$ keine andern Bedeutung als : Es gibt rationale Intervalle, an deren Endpunkten $x^2 - 2$ beliebig klein positiv, das andern Mal beliebig klein negativ gemacht werden kann.

Damit ist also die Bedeutung der Forderung erklärt, es soll eine Gleichung $f(x) = 0$ durch eine gewisse Größe $x$ erfüllt werden. Diese Größe $x$ ist als eine Grenze aufzufassen, der man sich nähert, wenn man das obige Intervall immer kleiner werden läßt. Der *Sturm*'sche Satz bestimmt daher eigentlich nichts anderes als die Anzahl der oben charakterisirten Intervalle $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, zwischen welchen ein Zeichenwechsel stattfindet.

Wenn wir uns, nicht zum beweise, sondern nur zur Erleichterung der Vorstellung, der geometrischen Anschauung bedienten und $f(x) = y$ als Curve aufzeichneten, so setzt dies bereits die Wurzelexistenz voraus. Denn zunächst können wir nur die rationalen Werthe konstruiren. Um nun die dadurch entstehenden Punkte zu einer

Curve zu verbinden, bedarf es der Voraussetzung, daß ein $x$ existire, welches dem zwischen $f(x_1)$ und $f(x_2)$ gelegenen $f(x)$ entspricht. Also schon daß man eine Curve zeichnet, implizirt die Voraussetzung, daß eine Gleichung Wurzeln hat.

120|121

| §50. Wir gehen nun dazu über, die Bedeutung der *Sturm*'schen Funktionen als symmetrische Funktionen zu studiren.

Wir knüpfen an die Kettenbruchentwicklung an und leiten zunächst einige Sätze über die Kettenbrüche ab. Es war :

$$\frac{f_1}{f} = \frac{1}{g_1 - \frac{1}{g_2 - \ldots - \frac{1}{g_\nu}}}$$

Wir setzen nun :

$$\frac{1}{g_1 - \frac{1}{g_2 - \ldots - \frac{1}{g_{k-1}}}} = \frac{\varphi_k}{\psi_k}$$

Einen etwaigen gemeinschaftlichen Theiler von $\varphi_k$ und $\psi_k$ haben wir weg. $\varphi_k$ und $\psi_k$ sind ganze Funktionen der $g$, die Dimension von $\varphi_k$ wird durch $g_2 g_3 \ldots g_{k-1}$, die von $\psi_k$ durch $g_1 g_2 \ldots g_{k-1}$ bestimmt.

Wir beweisen jetzt, daß die Gleichung besteht :

$$\varphi_{k+1} - g_k \varphi_k + \varphi_{k-1} = 0 \text{ und ebenso :}$$
$$\psi_{k+1} - g_k \psi_k + \psi_{k-1} = 0.$$

Wir beweisen dies durch Induktion. Wir erhalten $\frac{\varphi_{k+1}}{\psi_{k+1}}$ indem wir in $\frac{\varphi_k}{\psi_k}$ $g_{k-1}$ ersetzen durch $g_{k-1} - \frac{1}{g_k}$. Es ist daher :

$$\frac{\varphi_{k+1}}{\psi_{k+1}} = \frac{(g_{k-1} - \frac{1}{g_k})\varphi_{k-1} - \varphi_{k-2}}{(g_{k-1} - \frac{1}{g_k})\psi_{k-1} - \psi_{k-2}} = \frac{g_k \varphi_k - \varphi_{k-1}}{g_k \psi_k - \psi_{k-1}},$$

indem die Gleichungen :

$$\varphi_k - g_{k-1}\varphi_{k-1} + \varphi_{k-2} = 0 \text{ und } \psi_k - g_{k-1}\psi_{k-1} + \psi_{k-2}[= 0]$$

vorausgesetzt wurden. Da nun in $\frac{g_k \varphi_k - \varphi_{k-1}}{g_k \psi_k - \psi_{k-1}}$ Zähler und Nenner keinen gemeinschaftlichen Faktor (auch keinen Zahlenfaktor) besitzen, so ist :

$$\varphi_{k+1} = g_k \varphi_k - \varphi_{k-1}$$
$$\psi_{k+1} = g_k \psi_k - \psi_{k-1}.$$

Die rekurrirenden Formeln, aus denen man die Funktionen $\varphi$ und $\psi$ erhält, sind also für $\varphi$ und $\psi$ dieselben. Nur durch die Anfangswerthe wird eine Verschiedenheit bedingt, denn es ist :

$$\varphi_1 = 0, \varphi_2 = 1 \text{ und } \psi_1 = 1, \psi_2 = g_1$$

121|122 | Wir können aber Zähler und Nenner dieser Näherungsbrüche durch dieselbe Funktion Darstellen, wenn wir ein beliebiges Stück des Kettenbruches aus der Mitte heraus auffassen. Setzen wir nämlich :

$$\frac{1}{g_p - \frac{1}{g_{p+1} - \ldots - \frac{1}{g_{q-1}}}} = \frac{Z_{p,q}}{\mathcal{N}},$$

so wird es gelingen, den Nenner ebenfalls durch ein $Z$ auszudrücken, welches nach demselben Gesetze gebildet ist wie der Zähler. Es ist nämlich :

$$\frac{1}{g_{p-1} - \frac{Z_{p,q}}{\mathcal{N}}} = \frac{Z_{p-1,q}}{\mathcal{N}_1} = \frac{\mathcal{N}}{\mathcal{N} g_{p-1} - Z_{p,q}}$$

Da wir bereits oben voraussetzen, daß $Z_{pq}$ und $\mathcal{N}$ keinen gemeinschaftlichen Theiler haben, so folgt :

$$\mathcal{N} = Z_{p-1,q}.$$

Also ist :

$$\frac{1}{g_p - \frac{1}{g_{p+1} - \ldots - \frac{1}{g_{q-1}}}} = \frac{Z_{p,q}}{Z_{p-1,q}}.$$

Nun hatten wir *pag...* für $\varphi$ und $\psi$ eine rekurrirende Formel, die wir auch in die $Z$ übersetzen können, indem wir $q = k$ gesetzt denken, und $g_p$ eben als den ersten Nenner ansehen. Dann lautet die Gleichung in den $Z$ :

I) $Z_{p,q+1} - g_q Z_{pq} + Z_{p,q-1} = 0$

Wir wollen nun ein allgemeine Gleichung zwischen den $Z$ herleiten, welche alle Satze über die Kettenbrüche umfaßt. Dieselbe lautet :

II) $Z_{pq}Z_{rs} - Z_{pr}Z_{qs} + Z_{ps}Z_{qr} = 0$ wo $p \leq q \leq r \leq s$ sei.

Wir beweisen zunächst einen speziellen Fall dieser Gleichung und daraus als dann die allgemeine Gl.

Wir setzen nämlich $p = 0, q = 1, r = r, s = s$. Nun ist $Z_{pp} = 0, Z_{p,p+1} = 1$. Wir erhalten also zunächst die spezialle Gl. : $Z_{01}Z_{rs} - Z_{0r}Z_{1s} + Z_{0s}Z_{1r} = 0$, da aber $Z_{01} = 1$ ist, so folgt $Z_{rs} = Z_{0r}Z_{1s} - Z_{0s}Z_{1r}$. Diese spezialle Gl. wollen wir jetzt durch Induktion beweisen. Angenommen sie gälte für $s$, | so beweisen wir sie für $s+1$, d.h. 122|123
wir zeigen, die Richtigkeit der Gl. :

$$Z_{r,s+1} = Z_{0r}Z_{1,s+1} - Z_{0,s+1}Z_{1r}.$$

Nun können wir $Z_{1s+1}$ und $Z_{0s+1}$ mit Hülfe der Gl. I, *pag...* wegschaffen und erhalten :

$$\begin{aligned} Z_{r,s+1} &= Z_{0r}(g_s Z_{1s} - Z_{1,s-1}) - Z_{1r}(g_s Z_{0s} - Z_{0,s-1}) \\ Z_{r,s+1} &= g_s(Z_{0r}Z_{1s} - Z_{1r}Z_{0s}) - (Z_{0r}Z_{1,s-1} - Z_{1r}Z_{0,s-1}) \end{aligned}$$

Mit Berücksichtigung der Voraussetzung folgt hieraus :

$$Z_{r,s+1} = g_s Z_{rs} - Z_{r,s-1}.$$

Da die Richtigkeit dieser Gl. bewiesen ist, (Gl. I.) so ist damit also die Richtigkeit der Gl. : $Z_{rs} = Z_{0r}Z_{1s} - Z_{0s}Z_{1r}$ nachgewiesen. Mit Hülfe dieser Gl. läßt sich dann ohne Weiteres die allgemeine Gl. II. (*pag...*) beweisen, indem man einfach vermöge der Gl. für $Z_{rs}$ alle 6 Größen, $Z_{rs}$, $Z_{p,q}$, ... ausdrückt ; man erhält dann die Gl. $0 = 0$, wodurch die Richtigkeit der allgemeinen Relation II. Dargethan ist.

Wenn die Gl. I. uns eine Relation zwischen 3 auf einander folgenden $Z$ mit demselben ersten Index lieferte, so können wir dieser mit Hülfe der Gl. II. eine andern gegenüberstellen, wo der zweite Index derselbe ist und der erste sich ändert. Setzen wir nämlich in Gl. II. $p = q - 1, r = q + 1$ und berücksichtigen, daß :

$Z_{q-1,q} = 1, Z_{q-1,q+1} = g_q, Z_{q,q+1} = 1$ ist, so erhalten wir die gewünschte Relation, welche wir mit Gl. I. zusammenstellen wollen :

$$\begin{cases} Z_{p,q+1} - g_q Z_{p,q} + Z_{p,q-1} = 0. \\ Z_{q+1,s} - g_q Z_{q,s} + Z_{q-1,s} = 0. \end{cases}$$

Wenden wir nun diese Resultate auf die Kettenbruchentwicklung an, welche dem *Sturm*'schen Satze zu Grunde liegt.

Wir erkennen, daß für diese Kettenbruchentwicklung $Z_{k,\nu+1} = \frac{f_k}{f_\nu}$ ist, denn für $\frac{f_k}{f_\nu}$

besteht dieselbe Rekursionsformel wie für $Z_{k,\nu+1}$, nämlich die Gl. :

$$\frac{f_{k+1}}{f_\nu} - g_k \frac{f_k}{f_\nu} + \frac{f_{k-1}}{f_\nu} = 0$$

123|124 | und die Anfangswerthe sind dieselben, nämlich :

$$Z_{\nu,\nu+1} = 1 = \frac{f_\nu}{f_\nu}, Z_{\nu-1,\nu+1} = g_\nu = \frac{f_\nu - 1}{f_\nu}.$$

Setzen wir nun $q = p + 1, s = r + 1$, so folgt aus II. :

$$1 = Z_{pr} Z_{p+1,r+1} - Z_{p,r+1} Z_{p+1,r}.$$

Für $p = 0$ wird $Z_{0k} = \psi_k, Z_{1,k} = \varphi_k$ (indem wir $r = k$ setzen)
und wir erhalten :

1. $\varphi_{k+1}\psi_k - \varphi_k\psi_{k+1} = 1.$

Eine zweite einfache Relation folgt, wenn wir
$p = 0, q = 1, r = k, s = \nu + 1$ setzen, dann kommt :
$Z_{k,\nu+1} = Z_{0k} Z_{1,\nu+1} - Z_{0,\nu+1} Z_{1,k}$.
Oder :

2. $\psi_k f_1 - \varphi_k f_0 = f_k.$

Endlich erhalten wir noch eine dritte Formel, indem wir setzen :

$$p = 0, q = k, r = k + 1, s = \nu + 1$$

nämlich : $Z_{0,\nu+1} = Z_{0,k+1} Z_{k,\nu+1} - Z_{0k} Z_{k+1,\nu+1}$.
Oder :

3. $f_0 = \psi_{k+1} f_k - \psi_k f_{k+1}$

In den Gl. 2) und 3) haben wir wieder die früheren Resultate, diesmal abgeleitet aus der allgemeinen Kettenbruchrelation II, *pag...*, nämlich, daß jedes $f_k$ eine homogene lineare Funktion der beiden ersten Funktionen, $f_0$ und $f_1$, ist, daß aber auch $f_0$ eine homogene lineare Funktion von irgend zwei auf einander folgenden $f_k$ und $f_{k+1}$ ist. Hier erfahren wir aber überdies, daß in beiden Fälle die auftretenden Multiplikatoren Zähler resp. Nenner von Näherungsbrüchen der *Sturm*'schen Entwicklung sind.

Untersuchen wir noch, welches die Grade der bei der Kettenbruchentwicklung auftretenden Funktionen sind. Wir bezeichnen den Grad von $f_k$ mit $n_k$, so bilden die Zahlen $n_1, n_2 ... n_k, ... n_\nu$ eine abnehmende Reihe, wo $n_\nu = 0$ ist. Im sogenannten regulären Falle ist es die Reihe $n - 1, n - 2, ...1, 0$, so daß also $n_k = n - k$ und $\nu = n$ ist. Aus $f_{k-1} - g_k f_k + f_{k+1} = 0$ folgt, daß der Grad von $g_k$ gleich $n_{k-1} - n_k$ sein
124|125 muß, dem $f_{k+1}$ von niedrigerem Grade als $f_{k-1}$. | Den Grad von $\psi_k$ erhalten wir aus : $\psi_{k-1} - g_k\psi_k + \psi_{k-1} = 0$. Der Grad von $\psi$ folgt hieraus gleich $n = n_{k-1}$, denn setzen wir diesen Grad für $\psi_k$ voraus, so folgt, da $\phi_{k+1}$ von höherem Grade ist als $\psi_{k-1}$, für $\psi_{k+1}$ der Grad : $n_{k-1} - n_k + n - n_{k-1} = n - n_k$, so daß wir durch diesen Induktionsschluß in der That $n - n_{k-1}$ als den Grad von $\psi_k$ erhalten. Der Grad von $\psi_k$ ist nur den Grad von $g_1$ d.h. um $n - n_1$ kleiner als der von $\psi_k$, also gleich $n_1 - n_{k-1}$, wie man auch durch Induktion leicht bestätigt.

§51. Nach diesen vorbereitenden Sätzen über die Kettenbrüche und die *Sturm*'schen Entwicklungsfunktionen, fassen wir ein Problem ins Auge, welches mit der Darstellung der *Sturm*'schen Funktionen als symmetrische Funktionen in engem Zusammenhange steht, nämlich das Problem der *Cauchy*'schen Interpolationsformel. Dasselbe lautet : Es soll der Quotient zweier ganzen rationalen Funktionen $\frac{F(x)}{\Psi(x)}$ so

bestimmt werden, daß er für $x = \xi_h$ $(h = 1, ...n)$ den Werth $\mathcal{C}_h$ $(h = 1, ...n)$ annimmt. Sei $F(x)$ von Grade $\mu$, $\Psi(x)$ vom Grade $\nu$, so soll also :

$$\frac{\sum_{r=0}^{\mu} a_r \xi_h^r}{\sum_{s=0}^{\nu} b_s \xi_h^s} = \mathcal{C}_h \ (h = 1, ...n) \text{ sein.}$$

Es sollen also die $\mu + \nu + 2$ unbekannten Coeffizienten $a_0, ...a_\mu, b_0...b_\nu$ so bestimmt werden, daß die $n$ Gleichungen bestehen. Wir haben es also mit einem Problem der Auflösung homogener linearer Gleichungen zu thun und wissen, daß die Anzahl der letzteren, falls das Problem lösbar sein soll, mindestens gleich $\mu + \nu + 1$ sein muß. Als erste Bedingung für $\mu$ und $\nu$ erhalten wir daher $\mu + \nu + 1 \leq n$ oder $\mu + \nu < n$.

Ist der Nenner vom 0 ten Grade, also $\nu = 0$, so haben wir das Problem bereits durch die *Lagrange*'sche Interpolationsformel gelöst. Wir wissen nun, daß, wenn wir das *Cauchy*'sche Problem als ein Problem der Theorie der linearen Gleichungen ansehen, möglicherweise die Lösung von einer höheren Mannigfaltigkeit sein kann, so daß also gerade die nächstliegende Methode die Aufgabe möglicheweise sehr | kompliziren 125|126
kann. Wir geben für die Lösung des *Cauchy*'schen Problems 5 verschiedene Methoden.

§52. 1) *Methode von Liouville.*

Wir denken uns vermittelst der *Lagrange*'schen Interpolationsformel eine Funktion $f_1(x)$ vom Grade $n - 1$ hergestellt, welche für $x = \xi_1, ...x = \xi_n$ die Werhe $\mathcal{C}_1...\mathcal{C}_n$ annimmt. Dann ist :

$$\frac{F(\xi_h)}{\Psi(\xi_h)} = \mathcal{C}_h = f_1(\xi_h) \text{ (h=1,...n).}$$

Wir bilden nun die Funktion $f(x) = (x - \xi_1)(x - \xi_2)...(x - \xi_h)$ und wenden auf $f$ und $f_1$ das Kettenbruchverfahren an, so ist :

$$f_k = \psi_k f_1 - \varphi_k f \text{ ; für } x = \xi_h \text{ folgt :}$$
$$f_k(\xi_h) = \psi_k(\xi_h) f_1(\xi_h) \text{ oder :}^{12}$$
$$f_1(\xi_h) = \frac{f_k(\xi_h)}{\psi_k(\xi_h)} \ (h = 1, ...n).$$

Wir haben also eine rationale Funktion $\frac{f_k(x)}{\psi_k(x)}$ hergestellt, welche für $x = \xi_h$ den Werth $\mathcal{C}_h = f_1(\xi_h)$ annimmt, also unsern Aufgabe löst, wenn noch die Summe der Grade von Zähler und Nenner kleiner als $n$ ist. Der Grad des Zählers ist $n_k$, der des Nenners ist $n - n_{k-1}$, also ist diese Summe $= n + n_k - n_{k-1}$ und diese ist $< n$, weil $n_k < n_{k-1}$ ist. Alle Brüche $\frac{f_k(x)}{\psi_k(x)}$, bei welchen $f_k$ und $\psi_k$ keinen gemeinschaftlichen Theiler haben, von $k = 1$ bis $k = \nu$ genügen also dem *Cauchy*'schen Interpolationsproblem. Es fragt sich jetzt, ob es noch andern Lösungen dieses Problems gibt. (Siehe *Nachtrag pag...*).

---

Die Idee der Verbindung des *Cauchy*'schen mit der Kettenbruchentwicklung rührt von *Liouville* her. Sie findet sich veröffentlicht in dem Aufsatze von *Sturm* : „*Démonstration d'un théorème d'Algèbre de M. Sylvester*“ (*Liouville's Journal Bd.7, pag. 356* im Jahre 1842 [13]

126|127

| 53. 4. *Methode von Jacobi : (Crelle Bd. 30).*

Man kann das *Cauchy*'sche Problem auf folgendermaßen anfassen. Nachdem man mittelst der *Lagrange*'schen Formel die Funktion $f_1(x)$ hergestellt hat, folgt aus der Gl. :

12. Dabei wird jedoch vorausgesetzt, daß $f_k$ und $\psi_k$ keinen gemeinschaftlichen Theiler haben, siehe Anfang *pag.....*

13. Methode 2 und 3 siehe Anfang *pag....*

$\frac{F(\xi_h)}{\Psi(\xi_h)} = f_1(\xi_h)$ die andern Gl. :

$$\frac{F(\xi_h)}{f_1(\xi_h)} = \Psi(\xi_h).$$

Die Aufgabe lautet also jetzt so : Man soll eine Funktion $F(x)$ vom Grade $\mu < n$ so bestimmen, daß sie, dividirt durch $f_1(x)$, für $x = \xi_h$ (h=1,...n) mit der Funktion $\Psi(x)$ übereinstimmt, welche höchstens vom Grade $n - \mu - 1$ sein soll.

Wir denken uns, nach der *Lagrange*'schen Formel, die Funktion $\Psi$ so hergestellt, daß sie für $x = \xi_h$ der obigen Gl. genügt. Um dann auszudrücken, daß sie nur vom Grade $n-\mu-1$ sein soll, setzen wir die Coeffizienten, welche den höheren Exponenten entsprechen $= 0$. Hätte also die *Lagrange*'sche Formel für $\Psi(x)$ ergeben :

$$\Psi(x) = b_0 + b_1 x + ... + b_{n-\mu-1}x^{n-\mu-1} + b_{n-\mu}x^{n-\mu} + ... + b_{n-1}x^{n-1}$$

(die $b$ enthalten natürlich die unbekannten und noch zu bestimmenden $\mu + 1$ Coeffizienten von $F(x)$), so wäre zu setzen :

$$b_{n-\nu} = 0, ...b_{n-1} = 0.$$

Nun ist noch den *Euler*'schen Formeln (pag...) :

$$b_{n-1} = \sum_{h=1}^{n} \frac{\Psi(\xi_h)}{f'(\xi_h)} \text{ , } b_{n-2} = \sum_{h=1}^{n} \frac{\xi_h \Psi(\xi_h)}{f'(\xi_h)} \text{ ,...} b_{n-\mu} = \sum_{h=1}^{n} \frac{\xi^{\mu-1}\Psi(\xi_h)}{f'(\xi_h)}.$$

Setzt man diese Coefficienten gleich 0, so erhalten wir vermöge der Gl. für $\Psi(xi_h)$ (*pag 3....)* die folgenden Gleichungen :

$$\textstyle\sum_{h=1}^{n} \frac{\xi_h^t F(\xi_h)}{f_1(\xi_h) f'(\xi_h)} = 0 \text{ ,} (t = 0, 1, ...\mu - 1) \ ^{14}$$

Dies sind $\mu$ homogene, in den Coefficienten von $F$ lineare Gleichungen, welche also diese Coeffizienten zu bestimmen erlauben.

127|128

| Wir wollen nun diese Auflösung wirklich ausführen. Es sei :

$$F(x) = C_0 + C_1 x + ... + C_\mu x^\mu$$

Dann lauten unsern Gleichungen :

$$\sum_{\nu=1}^{n} \sum_{h=0}^{\mu} \frac{C_h \xi_\nu^{t+h}}{f_1(\xi_\nu) f'(\xi_\nu)} = 0 \ (t = 0, ...\mu - 1)$$

Setzen wir nun :

$$\sum_{\nu=1}^{n} \frac{\xi_\nu^{t+h}}{f_1(\xi_\nu) f'(\xi_\nu)} = s_{h+t},$$

so lauten unsern Gleichungen :

$$\sum_{h=0}^{\mu} C_h s_{h+t} = 0 \ (t = 0, ...\mu - 1) :$$

14. Eine einfachere Herleitung dieser Bedingungsgleichungen siehe Anhang *pag. ...*

Daraus folgt, abgesehen von einer multiplikativen Constanten :

$$F(x) = \begin{vmatrix} s_0 & s_1 & ... & s_\mu \\ s_1 & s_2 & ... & s_{\mu+1} \\ ... & ... & ... & ... \\ s_{\mu-1} & s_\mu & ... & s_{2\mu-1} \\ 1 & x & ... & x \end{vmatrix} \quad (S.\ pag.\ ...\text{u. }...)$$

Diese Determinante Läßt sich noch vereinfachen und auf den $\mu$ ten Grad reduziren. Es ist nämlich :

$$\begin{vmatrix} ... & ... & ... & ... \\ s_p & s_{p+1} & ... & s_{p+\mu} \\ ... & ... & ... & ... \\ 1 & x & ... & x^\mu \end{vmatrix} = (-1)^\mu \begin{vmatrix} ... & ... & ... & ... & ... \\ s_p & xs_p - s_{p+1} & xs_{p+1} - s_{p+2} & ... & xs_{p+\mu-1} - s_{p+\mu} \\ ... & ... & ... & ... & ... \\ 1 & 0 & 0 & ... & 0 \end{vmatrix} =$$

$$D_\mu(x) = |xs_{p+q} - s_{p+q+1}| \ (p, q = 0, 1, ...\mu - 1).$$

Wir haben nun oben erfahren, daß jede Lösung $\frac{F(x)}{\Psi(x)}$ in der Form $\frac{f_k(x)\Theta(x)}{\psi_k(x)\Theta(x)}$ enthalten sein muß. Es ist daher $F(x)$, oder, wie wir es zuletzt nannten, $D_\mu(x) \neq$ nichts anderes als :

$$D_\mu(x) = f_k(x)\Theta(x)$$

wo $k$ dadurch bestimmt ist, daß $\mu$ in dem Intervalle liegt : $n_k \leq \mu < n_{k-1}$. (Siehe Anfang *pag.*...).

Für die beiden Grenzfälle haben wir also :

$$C_k D_\mu(x) = \sigma_k f_k(x) \text{ für} \mu = n_k$$
$$C_k D_\mu(x) = \tau_k f_k(x) \text{ für} \mu = n_{k-1} - 1$$

128|129

| *Hattendorf* war die erste, der diese Untersuchung allgemein aufgefaßt hat, allerdings ohne die endgültige lösung des Problems zu finden.

Wir hatten oben die Ausdrücke gebildet :

$$\sum_\xi \frac{\xi^t D_\mu(\xi)}{f_1(\xi)f'(\xi)} \ (t = 0, ...\mu - 1);$$

bilden wir jetzt noch den Ausdruck, welcher auf jene folgt, nämlich :

$\sum_\xi \frac{\xi^\mu D_\mu(\xi)}{f_1(\xi)f'(\xi)}$. Zunächst entsteht der Ausdruck : $\frac{\xi^\mu D_\mu(\xi)}{f_1(\xi)f'(\xi)}$, indem man in der Determinante $\begin{vmatrix} s_0 & s_1 & ... & s_\mu \\ ... & ... & ... & ... \\ s_{\mu-1} & s_\mu & ... & s_{2\mu-1} \\ 1 & x & ... & x^\mu \end{vmatrix}$ die letzte Horizontalzeile durch die Ausdrücke $\frac{\xi^\mu}{f_1(\xi)f'(\xi)}, ..., \frac{\xi^{2\mu}}{f_1(\xi)f'(\xi)}$ ersetzt. Summirt man dann die so entstehenden Determinanten, so erhält man dasselbe, als wenn man gleich die Summation über $\xi$ in der letzten Horizontalreihe ausführt. Diese letzte Reihe wird dann zu $s_\mu, s_{\mu+1}, ...s_{2\mu}$. Analog der Definition von $\sigma_k$ können wir daher diese Determinante durch $\sigma_{k-1}$ bezeichnen, wenn wir $\mu = n_{k-1} - 1$ setzen. Nun war für $\mu = n_{k-1} - 1$ :

$$C_k D_{n_{k-1}-1}(x) = \tau_k f_k(x).$$

Hieraus folgt gemäß der Entstehung von $\sigma_{k-1}$ :

$$\sigma_{k-1} = \frac{\tau_k}{C_k} \sum_{\xi} \frac{f_k(\xi)\xi^{n_{k-1}-1}}{f_1(\xi)f'(\xi)} = \frac{\tau_k}{C_k} \sum_{\xi} \frac{\psi_k(\xi)}{f'(\xi)} \xi^{n_{k-1}-1} \text{ (Siehe } pag....)$$

Da nun $\psi_k$ vom Grade $n - n_{k-1}$ ist, so folgt aus den *Euler*'schen Formeln, daß in diese Summe alle Glieder bis auf das höchste verschwinden und die Summe nichts anderes ist als der Coeffizient des höchsten Gliedes von $\psi_k(x)$. Um diesen zu ermitteln haben wir :

$$f = \psi_k f_{k-1} - \psi_{k-1} f_{k-1}.$$

Da auf des rechten Seite $\psi_k f_{k-1}$ den Grad bestimmen, und $f_{k-1}$ als höchsten Coeffizienten $C_{k-1}$ besitzt, so muß der höchste Coeffizient von $\psi_k$ gleich $\frac{1}{C_{k-1}}$ sein, weil derjenige der linken Seite $= 1$ ist. Daraus folgt also jetzt :

$$\sigma_{k-1} = \frac{\tau_k}{C_k C_{k-1}} \text{ oder } C_k C_{k-1} \sigma_{k-1} = \Sigma_k.$$

Durch diese Rekursionsformel kann man also die höchsten Coeffizienten aller Restfunktionen durch Diese eigenthümlichen Determinanten ausdrücken.

Wir wollen nun zeigen, daß für die mittleren Fälle, wo also :

$$n_k < \mu < n_{k-1} - 1$$

129|130 | ist, die Determinanten $D_\mu(x)$ identisch verschwinden. Diese Funktionen $D_\mu$ waren durch die Gleichungen :

$$\sum_{\xi} \frac{\xi^t F(\xi)}{f_1(\xi)f'(\xi)} = 0 \ (t = 0, 1, ...\mu - 1)$$

bestimmt worden, wo $F$ vom $\mu$ ten Grade sein sollte. Nun kann es doch vorkommen, daß diesen $\mu$ linearen Gl. auch eine Funktion von niedrigeren als dem $\mu$ ten Grade genügt, und dies tritt gerade ein für den Fall, daß die Kettenbruchentwicklung eine nicht reguläre ist, wie wir ja grade jetzt durch die Annahme jener Intervalle voraussetzen. Der reguläre Fall ist durch $n_k = n - k$ definirt.

Nun ist gemäß den *Euler*'schen Formeln :

$$\sum_{\xi} \frac{\xi^t f_k(\xi)}{f_1(\xi)f'(\xi)} = \sum_{\xi} \frac{\xi^t \psi_k(\xi)}{f_1(\xi)f'(\xi)} = 0$$

so lange $t$ die Werthe von $0, 1...n_{k-1} - 2$ annimmt, denn es muß für das Verschwinden jener Summe die Bedingung erfüllt sein, $n - n_{k-1} + t < n - 1$. Wir haben also hier $n_{k-1} - 1$ homogene Gleichungen, welchen die Funktion $f_k$ vom Grade $n_k$ genügt. Ist daher $n_{k-1} - 1 = n_k$, so ist die Kettenbruchentwicklung eine reguläre und diesen Gleichungen wird durch $f_k$ genügt. Es sei nun :

$$f_k = C_k x^{n_k} + C'_k x^{n_k-1} + C''_k x^{n_k-2} + ...$$

so bestehen die Gleichungen :

$$\sum_{\xi} \frac{C_k \xi^{n_k+t} + C'_k \xi^{n_k+t-1} + ...}{f_1(\xi)f'(\xi)} = 0 \text{ für } t = 0, 1, ...n_{k-1} - 2$$

oder

$$C_k s_{n_k+t} + C'_k s_{n_k+t-1} + C''_k s_{n_k+t-2} + ... = 0 \ (t = 0, 1...n_{k-1} - 2$$

Diese Gleichungen liefern folgendes System :

$$\left\| \begin{matrix} s_0 & s_1 & \dots & s_{n_{k-1}-2} \\ s_1 & s_2 & \dots & s_{n_{k-1}+1-2} \\ \dots & \dots & \dots & \dots \\ s_{n_k} & s_{n_k+1} & \dots & s_{n_{k-1}+n_k-2} \end{matrix} \right\|$$

Im regulären Falle also ist $n_{k-1} = n_k + 1$, wir haben dan[n] $n_{k-1} - 1$ homogene Gleichungen mit $n_{k-1} = n_k + 1$ Unbekannten, deren Verhältnisse also bestimmt werden können. Die Funktion $f_k$ löst das Gleichungssystem, ohne daß noch weiteren Gleichungen, als erforderlich ist, genügt würde. Ist aber $n_{k-1} \geq n_k + 1$, so daß wir in obigem Systeme gleich viel resp. mehr Vertikal- als Horizontalreihen erhalten, so sind in Folge des | Bestehens der $n_{k-1} - 1$ homogenen linearen Gleichungen unter den 130|131
$n_k + 1$ Coeffizienten alle Determinanten $n_k + 1$ ter Ordnung, die man aus jenem System durch Unterdrückung von Vertikalreihen bilden kann, gleich 0.

Wir nehmen nun die Zahl $\mu$ in der Form an : $\mu = n_k + \lambda \leq n_{k-1} - 2$ (den Fall $\mu = n_{k-1} - 1$ haben wir schon behandelt und gefunden, daß dann durch das Verschwinden gewisser Determinanten der Grad von $\mu$ auf $n_k$ sinkt.) Dann ist :

$$\left| \begin{matrix} s_0 & s_1 & \dots & s_{n_k+\lambda} \\ s_1 & s_2 & \dots & s_{n_{k-1}+\lambda+1} \\ \dots & \dots & \dots & \dots \\ s_{n_k+\lambda-1} & s_{n_k+\lambda} & \dots & s_{2(n_k+\lambda)-1} \\ 1 & x & \dots & x^{n_k+\lambda} \end{matrix} \right|$$

Da nun aber $\lambda \geq 1$ ist und außerdem $n_{k-1} - 1 > n_k + 1$ anzunehmen ist, wenn wir nicht etwa einen der beiden schon behandelten Grenzfälle erhalten wollen, so folgt aus dem Verschwinden aller jener Determinanten $n_k + 1$ ter Ordnung, daß $D_\mu(x)$ identisch verschwindet.

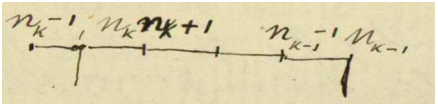
Dadurch also, daß die $f_k$ von kleineren Grade sind, als sie dem *Cauchy*'schen Probleme nach nöthig hätten, werden gewisse Bedingungen erfüllt, welche sich im Verschwinden der $D_\mu$ kundgeben. Das Verschwinden dieser Determinanten ist von *Hattendorf* nicht bemerkt werden, sondern wurde erst von *Kronecker* (Monatsberichte der Akad. 1873) nachgewiesen.

Wir wollen nun zeigen, daß diese Funktionen $D_\mu$ ebenso für den *Sturm*'schen Satz verwerthet werden können, wie die Restfunktionen der Kettenbruchentwicklung.

$$\text{Es war : } \sum_\xi [f_1(\xi) f'(\xi)] = \frac{1}{2} \sum_h [f_{h-1}(x_2) f_h(x_2)] - \frac{1}{2} \sum_h [f_{h-1}(x_1) f_h(x_1)].$$

Wollen wir nun die $f$ durch die Funktionen $D_\mu$ ersetzen, so müssen wir zunächst das Produkt zweier auf einander folgender $D_\mu$ bilden. Hierbei sind drei Fälle zu unterscheiden :

1. $D_{n_k-1} D_{n_k}$



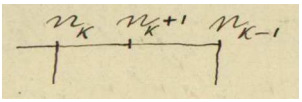
Hierfür haben wir :

$$C_k D_{n_k} = \sigma_k f_k$$
$$C_k D_{n_k-1-1} = \tau_k f_k \text{ also } C_{k+1} D_{n_k-1} = \tau_{k+1} f_{k+1}$$

131|132

| Hieraus folgt : $D_{n_k-1} D_{n_k} = \sigma_k^2 f_k f_{k+1}$.

2. $D_{n_k} D_{n_k+1}$




Hierfür brauchen wir nun den Fall zu berücksichtigen, der in der nebenstehenden Figur angegeben ist, wo also $n_k + 1 = n_{k-1} - 1$ ist, denn in einem andern Falle, wie er etwa in der dem erste Falle beigegebenen Figur angenommen ist, würde $D_{n_k+1}$ verschwinden. Für den Fall $n_k + 1 = n_{k-1} - 1$ ist jetzt :

$$C_k D_{n_{k-1}-1} = \tau_k f_k. \text{ Also :}$$
$$D_{n_k} D_{n_k+1} = \tfrac{\sigma_k \tau_k}{C_k^2} f_k^2$$

3. In jedem andern Falle ist das Produkt $D_{\mu-1} D_\mu$ gleich 0.

Wenn nun auch zunächst der Ausdruck :

$$\frac{1}{2}[D_{\mu-1}(x_2) D_\mu(x_2)] - \frac{1}{2}[D_{\mu-1}(x_1) D_\mu(x_1)]$$

mit der rechten Seite der *Sturm*'schen Gleichung (*pag....*) nur für $\mu = n_k$ übereinstimmt, so erkannt jedoch sofort, daß er es auch thut, wenn wir $\mu = h$ setzen und wie früher von 1 bis $n$ summiren, denn die letztgenannte Differenz verschwindet für den zweiten und den dritten Fall, d.h. für $\mu \neq n_k$. Damit ist also gezeigt, daß die *Sturm*'schen Restfunktionen sich durch diese eigenthümlichen Determinanten $D_\mu$ in dem *Sturm*'schen Satze volkommen ersetzen lassen. Diese Funktionen $D_\mu$ sind zuerst von *Jacobi* (*Cr. Bd. 30)* studirt worden. Ihre Anwendung auf den *Sturm*'schen Satzt zeigte zuerst *Cayley*, weßhalb man ihm auch häufig fälschlich die Auffindung dieser Funktionen zugeschrieben hat.

Wenn wir nun die Reihe der Funktionen $D_1, D_2, ...$ für $x = x_1$ und $x = x_2$ in Bezug auf ihre Zeichen untersuchen, so ergibt sich, wie früher, daß die Anzahl $m_1 - m_2$ der beim Uebergange von $x_1$ zu $x_2$ verlorenen Zeichenwechsel gleich der rechten Seite der *Sturm*'schen Gleichung (*pag....*) ist. Da einzelne der $D$ auch Null sind, so fragt es sich, ob die Null als positiv oder negativ gelten soll. es stellt sich aber heraus, daß dies gleichgültig ist ; wenn man sich nur konsequent entweder für das eine oder für das andern entscheidet.

Nehmen wir nun an $f_1$ habe immer dasselbe Zeichen wie $f'$. Setzen wir z.B.
132|133 $f_1 = f'$, oder $f_1 = \frac{1}{f'(x)}$. Im letzteren Falle | wäre $f_1$ nach *Lagrange* aus der Formel bestimmt :

$$\sum_\xi \frac{f(x)}{(x-\xi)} \frac{1}{f'(\xi)^2} = f_1(x).$$

Dann ist

$$\sum_\xi \frac{\xi^t}{f(\xi) f'(\xi)} = s_t = \sum_\xi \xi^t,$$

unsere $s$ gehen also in die Potenzsummen über.

Wir wollen dies nur für den regulären Fall weiter verfolgen. Wir nehmen an $x_1 = -\infty, x_2 = +\infty$, um alle reellen Wurzeln zuerhalten. Für den regulären Fall ist der Exponent der höchsten Potenz in dem Produkte $D_{h-1} D_h$ stets eine ungerade Zahl. Die Coeffizienten Dieser höchsten Potenzen in allen $D_{h-1} D_h$ sind die Produkte zweier aufeinander folgender $\sigma$, sodaß wir erhalten :

$$\sum [f.f'] = \sum [1] = \frac{1}{2} \sum_k [\sigma_k \sigma_{k-1}] - \frac{1}{2} \sum_k [-\sigma_k \sigma_{k-1}] = \sum_k [\sigma_k \sigma_{k-1}].$$

Die Anzahl aller reellen Wurzeln wird also durch die Summe

$$\sum_{k=1}^{n}[\sigma_k\sigma_{k-1}]$$

ausgedrückt, wo $\sigma_k$ die Determinante darstellt.

$$\sigma_k\begin{vmatrix} s_0 & s_1 & ... & s_{n-k-1} \\ s_1 & s_2 & ... & ... \\ ... & ... & ... & ... \\ s_{n-k-1} & s_{n-k} & ... & s_{2n-2k-2} \end{vmatrix} = |s_{p+q}| \ (p,q = 0,1...n-k-1)$$

Bezeichnet man mit $s$ und $m$ die Anzahl der Folgen und Wechsel in der Reihe $\sigma_0, \sigma_1, ...$ so ist :

$$\sum[\sigma_k\sigma_{k-1}] = 1 + s - m.$$

Sei nun $\mathcal{A}$ die Anzahl der reellen, $2\mathcal{B}$ die der imaginären Wurzeln, so ist :

$$\mathcal{A} = 1 + s - m$$
$$2\mathcal{B} = n - \mathcal{A} = n - 1 - s + m.$$

Da nun $s + m = n - 1$ ist, so folgt $\underline{\mathcal{B} = m}$, d.h. die Anzahl der Zeichenwechsel in der genannten Reihe ist gleich der Anzahl der Paare imaginärer Wurzeln. Dieser Satz gilt, wie bemerkt, nur wenn die Entwicklung von $f$ und $f_1$ regulär ist.

Wenn sich die Determinanten $|s_{p+q}|$ als Summen von Quadraten darstellen lassen, so hat man daraus geschlossen, daß die Gleichung nur reelle Wurzeln habe. Aber dieser Schluß ist falsch, denn eine Summe | von Quadraten kann ja auch Null 133|134
werden. Der Schluß gilt also nur, wenn man nachweist, daß die Determinaten von Null verschieden sind.

Da $\mathcal{B}$ höchstens $\frac{n}{2}$ sein kann, so können in jener Reihe $\sigma_0, \sigma_1, ...$ auf nur höchstens $\frac{n}{2}$ Zeichenwechsel eintreten, obgleich man deren von vornherein $n - 1$ erwarten könnte. Es rührt diese Beschränkung daher, daß die Potenzsummen $s_{n+1}...s_{2n}$ durch die früheren $s_0...s_n$ ausgedrückt werden können, daß also zwischen den $\sigma$ eine Abhängigkeit besteht. Diese ist derart, daß wenn in der ersten Hälfte $\sigma_0...\sigma_{\frac{n}{2}}$ $\frac{n}{2}$ Zeichenwechsel vorkommen sollten, in der zweiten Hälfte der Reihe deren keine mehr auftreten können.

§55. Bevor wir die dritte Methode, die *Cauchy*'sche Aufgabe zu lösen, besprechen, wollen wir einen dem Multiplikationssatze der Determinanten entsprechenden Satz über rechteckige Systeme aufstellen. Seien die beiden Systeme gegeben :

$$\left\|\begin{matrix} a_{00} & a_{01} & ... & a_{0n} \\ a_{10} & a_{11} & ... & a_{1n} \\ ... & ... & ... & ... \\ a_{m,0} & a_{m,1} & ... & a_{m,n} \end{matrix}\right\| \left\|\begin{matrix} b_{00} & b_{01} & ... & b_{0n} \\ b_{10} & b_{11} & ... & b_{1n} \\ ... & ... & ... & ... \\ b_{m,0} & b_{m,1} & ... & b_{m,n} \end{matrix}\right\| \quad n > m$$

so können wir hieraus folgendes quadratische System komponiren

$$C_{g,h} = \sum_{k=0}^{n} a_{g,k}b_{hk} \ , \text{ g h } = 0...m.$$

Dann ist :

$$|C_{g,h}| = \sum_{(h)} \mathcal{E}\begin{pmatrix} h_0 & ... & h_m \\ 0 & ... & m \end{pmatrix} C_{0,h_0} C_{1,h_1} ... C_{m,h_m}$$
$$= \sum_{(k)} \sum_{(h)} \mathcal{E}\begin{pmatrix} h_0 & ... & h_m \\ 0 & ... & m \end{pmatrix} a_{0,k_0} ... a_{mk_m} b_{h_0,k_0} ... b_{h_m k_m}.$$

Die Summe über $k$ ist so zu verstehen, daß aus den Zahlen $0, ...n$ ein Werthsystem $r_0 ... r_m$ ausgewählt werde und daß dann zunächst über alle $m!$ Permutationen $r_{i_1} ... r_{i_m}$ summirt und endlich über alle Werthsysteme $r_0, ... r_m$ die sich aus $0, ...n$ auswählen lassen, die Summe erstreckt werde. Dann folgt weiter, da :

$$\mathcal{E}\begin{pmatrix} h_0 & ... & h_m \\ 0 & ... & m \end{pmatrix} = \mathcal{E}\begin{pmatrix} h_0 & ... & h_m \\ k_0 & ... & k_m \end{pmatrix} \mathcal{E}\begin{pmatrix} k_0 & ... & k_m \\ 0 & ... & m \end{pmatrix} \text{ ist,}$$
$$|C_{g,h}| = \sum_{(k)} \mathcal{E}\begin{pmatrix} k_0 & ... & k_m \\ 0 & ... & m \end{pmatrix} a_{0k_0} ... a_{mk_m} |b_{hk_0} ... b_{hk_m}| \text{ oder :}$$
$$|C_{g,h}| = \sum_{(k)} |a_{gk_0} ... a_{gk_m}||b_{hk_0} ... b_{hk_m}|$$

134|135

| Jetzt ist die Summation nur noch zu erstrecken über alle Werthsysteme $k_0 ... k_m$, die man aus den Zahlen $0...n$ herausgreifen kann. Wir erhalten also $|C_{g,h}|$ indem man in dem ersten der beiden rechteckigen Systeme durch Ausscheiden von Vertikalreihen eine Determinante $m$ ter Ordnung herausgreift, mit der entsprechenden Determinante des zweiten Systems multiplizirt und über alle Produkte die auf diese Weise entstehen, summirt.

Wenden wir diesen Satz jetzt auf ein besonderes System von Größen an, $v[u], v, \xi, \eta$ seien beliebige Variabeln. Die beiden rechteckigen Systeme mögen dann lauten :

$$u_i \xi_i^g \text{ und } v_k \eta_k^h \begin{pmatrix} i,k=0, & ... & h_n \\ g,h=0, & ... & k_m \end{pmatrix}$$

Dann ist :

$$|C_{g,h}| = |\sum_{k=0}^{n} u_k v_k \xi_k^g \eta_k^h| = \sum_{(k)} |u_{k_0} \xi_{k_0}^g ... u_{k_m} \xi_{k_m}^g||v_{k_0} \eta_{k_0}^h ... v_{k_m} \eta_{k_m}^h| \ (g,h = 0...m)$$

Nun ist :

$$\begin{matrix} |u_i \xi_i^g| \\ \begin{pmatrix} i,k=0, & ... & h_n \\ g,h=0, & ... & k_m \end{pmatrix} \end{matrix} = \prod u |\xi_i^g| = \prod u_\alpha \prod (\xi_\alpha - \xi_\beta) \ \alpha, \beta = k_0 ... k_m$$

Der Index von $\beta$ sei in $\prod(\xi_\alpha - \xi_\beta)$ immer größer als der von $\alpha$.
Dann folgt :

$$|C_{g,h}| = \sum_{(k)} \prod u_\alpha v_\alpha \prod u_\alpha v_\alpha \prod (\xi_\alpha - \xi_\beta)(\eta_\alpha - \eta_\beta) \ \alpha, \beta = k_0 ... k_m.$$

Nun setzen wir $u_0 = 1$, $\xi_0 = x$, $\eta_0 = 0$, während für $k = 1, ...n$ sei :

$$\xi_k = \eta_k, v_k \eta_k = 1, \sum_{k=1}^{n} u_k \xi_k^h = s_h.$$

Dann wird unsere Determinante $|C_{gh}|$ zu :

$$|\sum_{k=0}^{n} u_k v_k \xi_k^g \eta_k^h| = |v_0 x^g + s_{g-1}, s_g, s_{g+1} ... s_{g+m-1}|.$$

Bestimmen wir in beiden Darstellungen für $|C_{g,h}|$ den Faktor von $v_0$, so folgt :

$$v_0|x^g, s_g, s_{g+1}...s_{g+m-1}| = v_0 \sum_{(k)} u_0 \prod_{(\alpha,\beta=k_1...k_m)} (\xi_\alpha - \xi_\beta)(\eta_\alpha - \eta_\beta) \prod (\xi_0 - \xi_\beta)(\eta_0 - \eta_\beta) \prod u_\alpha v_\alpha$$

Die Summe ist zu erstrecken über alle Werthsysteme $k_1...k_m$, die man aus der Zahlenreihe $1, ...n$ auswählen kann. Berücksichtigt man die Spezialisirung für $u, v, \xi, \eta$ so folgt :

$$|x^g, s_g, s_{g+1}...s_{g+m-1}| = \sum_{(k)} (-1)^m \prod u_\alpha \prod_{(\alpha,\beta=k_1...k_m)} (x - \xi_\alpha) \prod (\xi_\alpha - \xi_\beta)^2$$

Setzen wir nun $u_k = \frac{1}{f_1(\xi_k) f'(\xi_k)}$, so wird die letzte Gl. zu : 135|136

| $D_m(x) = \sum_{(k)} \prod \frac{1}{f_1(\xi_\alpha) f'(\xi_\alpha)} \prod (x - \xi_\alpha) \prod (\xi_\alpha - \xi_\beta)^2$.

Durch diese Voraussetzung über $u_k$ erhalten wir also $D_m$ dargestellt als symetrische Funktionen der $\xi$ und haben dadurch den Zusammenhang dieser kombinatorischen Ausdrücke mit dem *Cauchy*'schen Probleme hergestellt. Setzen wir $u_k = 1$ oder $f_1 = \frac{1}{f'}$, so erhalten wir die *Sylvester*'schen Formeln, nämlich für $m = 1, 2, ...n$ :

$$\begin{vmatrix} s_0 & s_1 \\ 1 & x \end{vmatrix} = \sum_{\alpha=1}^{n} (x - \xi_\alpha).$$

$$\begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ 1 & x & x^2 \end{vmatrix} = \sum_{(\alpha,\beta=1,...n)} (x - \xi_\alpha)(x - \xi_\beta)(\xi_\alpha - \xi_\beta)^2 \quad \text{etc. etc.}$$

Die letzte der Formeln lautet :

$$\begin{vmatrix} s_0 & s_1 & ... & s_n \\ s_1 & s_2 & ... & s_{n+1} \\ ... & ... & ... & ... \\ s_{n-1} & ... & ... & s_{2n-1} \\ 1 & x & ... & x^n \end{vmatrix} = \prod_{\alpha=1}^{n} (x - \xi_\alpha) \prod_{\alpha,\beta=1...n;\alpha<\beta} (\xi_\alpha - \xi_\beta)^2$$

In diesen Sylvester'schen Formeln bedeuten die $s$ die Potenzsummen. Die erste der Formeln ist evident, denn $\sum(x - \xi - \xi_\alpha)$ ist gleich $nx - s_1$ und da $s_0 = n$ ist, so folgt auch $\begin{vmatrix} s_0 & s_1 \\ 1 & x \end{vmatrix} = nx - s_1$.

Die Letzte der Formeln gibt an, daß Produkt der ursprünglichen Funktion $f(x) = \prod(x - \xi_\alpha)$ in ihre Diskriminante $\prod(\xi_\alpha - \xi_\beta)^2$ sich durch diese eigenthümliche Determinante darstellen läßt. (Siehe *Joachimsthal, Crelle Bd. 48.*)

Nachdem wir den Zusammenhang der abgeleiteten kombinatorischen Ausdrücke mit der *Jacobi*'schen Methode nachgewiesen haben, werden wir jetzt noch eine weitere von den früheren Methoden wesentlich verschiedene Auflösungsmethode des *Cauchy*'schen Problems geben, bei welcher wiederum die *Sylvester*'schen Formeln zur Anwendung kommen. Wir geben also jetzt die

5 te Methode zur Auflösung des *Cauchy*'schen Problems. Es ist :

$$\frac{f_1}{f} - \frac{\Phi}{\Psi} = \frac{F}{\Psi}.$$

Wir können dann die folgende Aufgabe stellen : Es soll ein Bruch $\frac{\Phi}{\Psi}$, mit dem Nenner vom Grade $\nu$, so bestimmt werden, daß die Differenz $\frac{f_1}{f} - \frac{\Phi}{\Psi}$ für den gegebenen Grad $\nu < n$, für unendlich | großes $x$ von möglichst hoher Ordnung unendlich klein werde. 136|137
$\frac{F}{f\Psi}$ wird nämlich für unendlich großes $x$ von der Ordnung $-(n+\nu-\mu)$ unendlich groß.

Es soll also $\frac{\Phi}{\Psi}$ sich dem $\frac{f_1}{f}$ für unendlich großes $x$ möglichst annähern. Entwickeln wir nun $\frac{f_1}{f}$ nach fallenden Potenzen von $x$, so kommt :

$$\frac{f_1}{f} = \sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} \frac{1}{x-\xi} = \sum_{k=0}^{\infty} \sum_{\xi} \frac{f_1(\xi)}{f'(\xi)\frac{\xi^k}{x^{k+1}}}.$$

Setzen wir : $\sum \frac{f_1(\xi)}{f'(\xi)} \xi^k = s'_k$, so ist also : $\frac{f_1}{f} = \sum_{k=0}^{\infty} \frac{s'_k}{x^{k+1}}$.

Zerlegen wir nun zunächst $\frac{\Phi}{\Psi}$ in Partialbrüche und nennen $\eta_k$ $(k = 1, ...\nu)$ die Wurzeln von $\Psi$, so ist :

$$\frac{\Phi}{\Psi} = \sum_{h=1}^{\nu} \frac{v_h}{x-\eta_k} = \sum_{h=1}^{\nu} \sum_{k=0}^{\infty} \frac{v_h \eta_k^k}{x^{k+1}} = \sum_{k=0}^{\infty} \frac{\sum_{h=1}^{[\nu]} v_h \eta_h^k}{x^{k+1}}.$$

Nun haben wir $2\nu$ Variabeln $v_h, \eta_h$ $(h = 1, ...\nu)$ und erfüllen also die Aufgabe, indem wir setzen :

$$\sum_{h=1}^{\nu} v_h \eta_h^k = s'_k \ (k = 0, ...2\nu - 1)$$

Dann wird für $x = \infty$ die Annäherung von $\frac{\Phi}{\Psi}$ an $\frac{f_1}{f}$ eine möglichst große. Benutzen wir nun die Formel *pag.*..., aus welcher für $u_k = 1$ die *Sylvester*'schen Formeln flossen und setzen sofort $\sum_{h=1}^{\nu} v_h \eta_h^k = s'_k$, so kommt :

$$\begin{vmatrix} s'_0 & s'_1 & ... & s'_\nu \\ s'_1 & ... & ... & ... \\ ... & ... & ... & ... \\ s'_{\nu-1} & ... & ... & s'_{2\nu-1} \\ 1 & y & ... & y^\nu \end{vmatrix} = \prod v_h \prod (y - \eta_k) \prod (\eta_h - \eta_k)^2 = D.\Psi(y).$$

Damit haben wir also $\Psi(y)$, abgesehen von einem Proportionalitätsfaktor, in Determinantenform dargestellt und wir erkennen, daß diese Auffassung des *Cauchy*'schen Problems zu demselben Resultate führt, wie die *Jacobi*'sche Methode, obwohl der Weg ein gänzlich verschiedener ist.

Wir können jetzt auch die Größen $v$ in Determinantenform darstellen. Denken wir in der Determinante, welche $= D\Psi(y)$ war, den Coeffizienten von $y^\nu$ herausgezogen, und bezeichnen durch $V(y) = y^\nu + c_1 y^{\nu-1} + ...$ die Funktion, deren Wurzeln $\eta_1 ... \eta_\nu$ sind und die sich also nur durch einen konstanten Faktor von unserer Determinante
137|138 $D\Psi(y)$ unterscheidet. Führen wir dann ebenfalls analog wie früher bei $F(x)$ | (siehe *pag.*...) die Funktionen $g_r(y)$ ein, definirt durch $g_r(y) = y^r - c_1 y^{r-1} + ... \pm c_r$ so ist in Folge dieser Definition :

$$g_r(y) = \begin{vmatrix} s'_0 & s'_1 & ... & ... & ... & ... & s'_\nu \\ s'_1 & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... \\ s'_{\nu-1} & ... & ... & ... & ... & ... & s'_{2\nu-1} \\ 0 & 0 & ... & 1 & y & ... & y^\nu \end{vmatrix}$$

Nun ist, wenn $V'(y)$ die Ableitung von $V(y)$ bedeutet :

$$\sum_{h=1}^{\nu} \sum_{r=0}^{\nu-1} g_r(\eta_k) v_h \eta_h^{\nu-r-1} = v_k V'(\eta_k) = \sum_{r=0}^{\nu-1} g_r(\eta_k) s'_{\nu-r-1}.$$

Schreiben wir nun den Ausdruck $\sum_{r=0}^{\nu-1} g_r(\eta_k) s'_{\nu-r-1}$ in Determinantenform, indem wir mit Hülfe der Form von $g_r(y)$ die Summe dadurch bilden, daß wir die Summation

in der letzten Horizontalreihe ausführen, so wird diese Reihe :

$$0, s'_0, s'_1 + ys_0, s'_2 + y's'_1 + y^2 s'_0, ..., s'_{\nu-1} + ys'_{\nu-2} + ... + y^{\nu-1} s'_0$$

Eine einfach Umformung führt uns zu der folgenden Determinante :

$$(-1)^\nu \begin{vmatrix} 0 & s'_0 & s'_1 & ... & s'_{\nu-1} \\ s'_0 & s'_1 - ys'_0 & s'_2 - ys'_1 & ... & s'_\nu - ys'_{\nu-1} \\ s'_1 & s'_2 - ys'_1 & s'_3 - ys'_2 & ... & s'_{\nu+1} - ys'_\nu \\ ... & ... & ... & ... & ... \\ s'_{\nu-1} & s'_\nu - ys'_{\nu-1} & s'_{\nu+1} - ys'_\nu & ... & s'_{2\nu-1} - ys'_{2\nu-2} \end{vmatrix}$$

Dieser Ausdruck stellt für $y = \eta_k$ den Werth $v_k V'(\eta_k)$ dar, so daß wir also $v_k$ als Quotienten zweier Determinanten bestimmt haben. Die letzte Determinante stellt abgesehen von konstanten Multiplikatoren die Funktion $\Phi(y)$ dar, denn es ist, abgesehen von diesem Multiplikator :

$$\Phi(y) = \sum_{h=1}^{\nu} v_k \frac{V(y)}{y - \eta_k} = \sum_{h=1}^{\nu} \sum_{r=0}^{\nu-1} g_r(y) v_h \eta_h^{\nu-r-1} = \sum_{r=0}^{\nu-1} g_r(y) s'_{\nu-r-1}$$

---

138|139

| §56. Wir fügen diesem Problem, welches eine möglichst große Annäherung eines Brüches an einen gegebenen Bruch verlangte, eine verwandte Aufgabe hinzu, welche von *Tchebicheff* herrührt und als eine Verallgemeinerung des *Lagrange*'schen Interpolationsproblems zu betrachten ist. Die Aufgabe lautet : Es soll eine ganze Funktion $\mathcal{G}(x)$ vom Grade $n - n_k - 1$ so bestimmt werden, daß sie für $n$ Werthe $\xi_1 ... \xi_n$ einer gegebenen ganzen Funktion $T(x)$ von höherem als dem $n - n_k - 1$ ten Grade möglichst nahe kommt.

Es ist also nach der Methode der kleinsten Fehlerquadrate der Ausdruck :

$$\sum_{h=1}^{n} \Theta_h^2 (\mathcal{G}(\xi_h) - T(\xi_h))^2$$

zu einem Minimum zu machen, wenn die positive Größe $\Theta_h^2$ das Gewicht des Fehlers $\mathcal{G}(\xi_h) - T(\xi_h)$ bedeutet.

Bilden wir $f(x) = (x - \xi_1)...(x - \xi_n)$ und hieraus $f'(\xi)$ und setzen wir dann $f_1(x) = \Theta^2 f'(x)$, so wäre :

$$\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} (\mathcal{G}(\xi) - T(\xi))^2$$

zu einem Minimum zu machen. Die nothwendigen Bedingungen hierfür sind das Bestehen der Gleichungen :

$$\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} (\mathcal{G}(\xi) - T(\xi)) \xi^t = 0 \ , \ (t = 0, 1, ...n - n_k - 1).$$

Die zweite Differentiation liefert als hinreichende Bedingung für das Minimum, daß :

$$\sum_{r,s} \sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} \xi^{r+s} z_r z_s$$

eine definite positive Form sei. Es ist aber :

$$\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} \sum_{r,s} \xi^{r+s} z_r z_s = \sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} (\sum_{r} \xi^r z_r)^2$$

und da $\frac{f_1(\xi)}{f'(\xi)} = \Theta^2$ ist, so sind die nothwendigen und hinreichenden Bedingungen für das Minimum bereits erfüllt, sofern man nur nachweisen kann, daß $\mathcal{G}$ stets so
139|140 bestimmt werden kann, | daß :

$$\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} (\mathcal{G}(\xi) - T(\xi))\xi^t = 0 \text{ ist, für } t = 0, 1, ...n - n_k - 1.$$

Diesen Nachweis verbindet man am besten mit der Kettenbruchsentwicklung von $f$ und $f_1$. Es ist :

$$f_1\psi_k - f\varphi_k = f_k \ , \ f_k\psi_{k+1} - f_{k+1}\psi_k = f.$$

Hieraus leitet man ab :

$$f_k\varphi_{k+1} - f_{k+1}\varphi_k = f_1.$$

Sei nun $T(x)$ eine beliebige ganze Funktion, so können wir dieselbe vermittelst $f(x) = 0$ in der reduzirten Form, also vom Grade $n-1$ höchstens voraussetzen. Dividiren wir nun $f_k(x)T(x)$ durch $f(x)$, so kommt :

$$f_k(x)T(x) = Q_k(x)f(x) + R_k(x).$$

$R_k$ ist nach der *Lagrange*'schen Formel :

$$R_k(x) = \sum_{\xi} \frac{T(\xi)f_k(\xi)}{f'(\xi)} \frac{f(x)}{x - \xi}.$$

Bilden wir nun hieraus den Ausdruck :

$$R_k(x)f_{k+1}(x) - R_{k+1}(x)f_k(x)$$

so erkennt man leicht, daß er durch $f(x)$ theilbar ist. Wir setzen ihn daher gleich $f(x).\ R(x)$, wo $R$ eine ganze Funktion ist. Nun ist $R$ höchstens vom Grade $n_k - 1$, wie aus der Bildung von $R$ hervorgeht.

Bilden wir ebenso :

$$R_k(x)\psi_{k+1}(x) - R_{k+1}(x)\psi_k(x)$$

so ist auch dieser Ausdruck durch $f$ theilbar. Wir setzen ihn daher gleich $f(x)\mathcal{G}(x)$, wo $\mathcal{G}(x)$ eine ganze Funktion vom Grade $n - n_k - 1$ [ist]. Dann ist jetzt :

$$\mathcal{G}(x)f(x) = \sum_{\xi} \frac{T(\xi)(f_k(\xi)\psi_{k+1}(x) - f_{k+1}(\xi)\psi_k(x))}{f'(\xi)} \frac{f(x)}{x - \xi}.$$

Ist nun $T$ vom Grade $n - n_k - 1$, so ist $S(x)$ mit $T(x)$ identisch. Wir wollen aber voraussetzen $T$ sei von höherem Grade als $n - n_k - 1$. Dann ist :

$$ff_1\mathcal{G} = R_k(f_{k+1} + f\varphi_{k+1}) - R_{k+1}(f_k + f\varphi_k) =$$

140|141 |

$$R_k f_{k+1} - R_{k+1}f_k + f(f_kT - Q_kf)\varphi_{k+1} - f(f_{k+1}T - Q_{k+1}f)\varphi_k$$
$$= Rf + ff_1T + f^2(\varphi_kQ_{k+1} - \varphi_{k+1}Q_k) = ff_1\mathcal{G}.$$

Es ist also : $f_1(\mathcal{G} - T) \equiv R \text{ mod. } f.$

Um dieses Resultat in Gleichungen auszudrücken, haben wir nach *Euler*:

$$\sum_{\xi} \frac{R(\xi)}{f'(\xi)} \xi^t = 0 \ , \ t = 0, 1, ...n - n_k - 1, \text{ folglich :}$$

$$\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} (\mathcal{G}(\xi) - T(\xi)) \xi^t = 0 \ , \ t = 0, 1, ...n - n_k - 1$$

Damit ist also die Möglichkeit der Herstellung einer Funktion $\mathcal{G}(x)$, wie sie im *Tchebicheff*'schen Probleme verlangt wird, bewiesen. $\mathcal{G}(x)$ stellt für $x = \xi_1 ... \xi_n$ eine gegeben Funktion $T(x)$ (oder gegebene Werthe $T(\xi_1), ...T(\xi_n)$) möglichst genau dar.

§ 57. Wir wollen nun noch der regulären Fall der Kettenbruchentwicklung ins Auge fassen, um daran die Auseinandersetzung einer neuen Methode zu knüpfen, welche uns zum *Sturm*'schen Satze führt. Für den regulären Fall ist $n_k = n - k$. Anknüpfend an die Bezeichnung von *pag....* haben wir $\sigma_k = \tau_k$, folglich :

$$\frac{\tau_{k-1}}{\tau_k} C_k C_{k-1} = 1 \ , \text{ wenn } C_k \text{ und } C_{k-1}$$

die coeffizienten der höchsten Potenzen von $f_k$ und $f_{k-1}$ sind. Wesentlich für den regulären Fall ist nun der Umstand, daß man jede reduzirte Form einer Funktion von $x$ linear ausdrücken kann durch $f_1(x), ...f_{n-1}(x) f_n(x)$, da die Grade dieser Funktionen die auf einander folgenden Werthe $n - 1, ...1, 0$ besitzen. Jede reduzirte Form von $x$ kann also auf die Form $y_1 f_x + y_2 f_2(x) + ... + y_n f_n(x)$ gebracht werden. Wir hatten nun (*pag....*) :

$$\sum_{\xi} \frac{\xi^t f_k(\xi)}{f_1(\xi) f'(\xi)} = 0 \ , \ t = 0, ...n - k - 1.$$

also ist auch :

$$\sum_{\xi} \frac{f_i(\xi) f_k(\xi)}{f_1(\xi) f'(\xi)} = 0 \ , \text{ wenn} n - i < n - k,$$

d.h. wenn $k < i$ bleibt.

| Da aber $i$ und $k$ gleichberechtigt sind, so hat man überhaupt : 141|142

$$\sum_{\xi} \frac{f_i(\xi) f_k(\xi)}{f_1(\xi) f'(\xi)} = 0 \text{ für} i \neq k.$$

Setzen wir für $i = k : S_k = \sum_{\xi} \frac{(f_k(\xi))^2}{f_1(\xi) f'(\xi)}$, so ist zu setzen :

$$\sum_{\xi} \frac{f_i(\xi) f_k(\xi)}{f_1(\xi) f'(\xi)} = \delta_{i,k} S_k = \left\{ \begin{array}{l} \delta_{i,k} = 0 \text{ für } i \neq k \\ \delta_{i,k} = 0 \text{ für } i = k \end{array} \right\}.$$

Diese $n^2$ Gleichungen lassen sich in eine zusammenfassen mit Hülfe der $n$ beliebigen Variabeln $y_1 ... y_n$. Wir erhalten :

$$\sum_{i,j} \sum_{\xi} y_i y_k \frac{f_i(\xi) f_k(\xi)}{f_1(\xi) f'(\xi)} = \sum_{\xi} \frac{1}{f_1(\xi) f'(\xi)} \left( \sum_{k=1}^{n} y_k f_k(\xi) \right)^2 = \sum_{k=1}^{n} S_k y^2.$$

Setzen wir nun $\sum_{k=1}^{n} y_k f_k(\xi_h) = z_h$, so haben wir :

$$\sum_{h=1}^{n} \frac{1}{f_1(\xi_h) f'(\xi_h)} z_h^2 = \sum_{k=1}^{n} S_k y^2.$$

Mit Hülfe der Substitution :

$$z_h = \sum_{k=1}^{n} y_k f_k(\xi_h)$$

haben wir somit eine Summe von Quadraten wieder in die Summe von Quadraten transformirt.
Sind nun $f_1(\xi_h)$ und $f'(\xi_h)$ für $h = 1, ...n$ von demselben Zeichen, so haben wir hier eine Summe von positiven Quadraten wieder in eine Summe von positiven Quadraten verwandelt.

Diese Invarianz der Anzahl der positiven Zeichen bei einer Summe von positiven Quadraten, war schon frühe bekannt, namentlich in der Geometrie. Erst *Jacobi* aber entdeckte (*Crelle Bd.* 53), daß für eine beliebige quadratische Form die Anzahl der positiven und der negativen Zeichen bei einer linearen Transformation ungeändert bleibt. Fast gleichzeitig haben sich *Hermite* und *Sylvester* mit dieser Frage beschäftigt. Von letzterem stammt der Ausdruck : Trägheitsprinzip der quadratischen Formen.

Wir wollen nun die Richtigkeit dieses Prinzips beweisen, indem wir den Satz auf den einfachsten Fall, nämlich auf die Transformation der Summe zweier Quadrate zurückführen.

142|143 | Es soll also $pu^2 + qv^2 = p'u'^2 + q'v'^2$ sein und nachgewiesen werden, daß bei dieser Transformation die Summe der positiven Vorzeichen ungeändert bleibt. Die Transformation sei durch die allgemeine lineare Substitution ausgeführt :

$$u' = m(u + qtv) \ , \ v' = n(v - ptu),$$

wo $m, n, t$ beliebig zu wählen sind. Es ist :

$$p' = \frac{p}{m^2(1 + pqt^2)} \ , \ q' = \frac{q}{n^2(1 + pqt^2)}.$$

Wir haben also eine dreifache Mannigfaltigkeit von möglichen Substitutionen. Haben nun $p$ und $q$ derselbe Zeichen, also $pq > 0$, so folgt aus :

$$p'q' = \frac{pq}{m^2n^2(1 + pqt^2)^2},$$

daß auch $p'q' > 0$ ist. Ebenso folgt aus $pq < 0$ auch $p'q' < 0$. Damit ist für diese einfache Transformation das Prinzip bewiesen.

Es werde nun allgemein die Form :

$$\sum_{i=1}^{n} a_i z_i^2 \text{ in die andere } \sum_{h=1}^{n} b_h y_h^2$$

übergeführt vermittelst der Substitution :

$$z_i = \sum_{k=1}^{n} C_{ik} y_k$$

so daß haben :

$$\sum_{i=1}^{n} a_i z_i^2 = \sum_{i,r,s=1}^{n} a_i C_{ir} C_{is} y_r y_s = \sum_{h=1}^{n} b_h y_h^2.$$

Dabei müssen die $n^2$ Coeffizienten die $\frac{n.n-1}{2}$ Bedingungen erfüllen :

$$\sum_{i} a_i C_{ir} C_{is} = \delta_{r,s} b_r.$$

In folge dessen hängen die $n^2$ Coeffizienten von $\frac{n.n-1}{2}$ unabhängigen Variabeln ab. Bei unserer elementaren Transformation konnten wir sofort die Coeffizienten ($n^2 = 4$) durch $\frac{n.n+1}{2} = 3$ beliebige Variabeln $m, n, t$ rational ausdrücken.

Analog mit der von *Clebsch* eingeführten Bezeichnung, nach welchen eine Curve $F(x, y) = 0$ als rational oder vom Geschlechte 0 bezeichnet wird, wenn $x$ und $y$ sich rational durch einen Parameter $t$ sich ausdrücken lassen, können wir auch unsere Transformation als vom Geschlechte Null bezeichnen. Es wird sich nämlich herausstellen, daß wir die $n^2$ Coeffizienten rational durch $\frac{n.n+1}{2}$ Variabeln auszudrücken im Stande sind. Wir lösen die Aufgabe schrittweise | durch successives Anwenden der 143|144
elementaren Transformationen.

Sei also : $\sum_{i=1}^{n} a_i z_i^2 = \sum_{h=1}^{n} b_h y_h^2$ vermöge $z_i = \sum_{k=1}^{n} C_{ik} y_k$.

Wir beginnen bei der Transformation von $\sum_{i=1}^{n} a_i z_i^2$ mit $a_1 z_1^2 + a_2 z_2^2$. Diese Summe können wir transformiren in $a_1' z_1'^2 + a_2' z_2'^2$ mit Hülfe eines $t$, welches wir so wählen können, daß in einer der linearen Funktionen der $y$, z.B. in $z_1$ das Glied $y_n$ fehlt. Ist dies erreicht, so nehmen wir $z_2$ und $z_3$ zusammen und transformiren so, daß auch in $z_2$ das Glied $y_n$ fehlt und so fahren wir fort. Dies geht so lange als ein $z$ zur Disposition steht, das wir mit dem vorhergehenden zusammennehmen können, d.h. wir können erreichen, daß in $n-1$ Fct. $z$ das Glied $y_n$ fehlt.

Wir haben jetzt :

$$\sum_{i=1}^{n} a_i' z_i'^2 = \sum_{h=1}^{n} b_h y_n^2,$$

wo $z_1' ... z_{n-1}'$ Fct. von $y_1 ... y_{n-1}$ allein sind und $z_n'$ noch eine lineare Fct. von $y_1 ... y_{n-1} y_n$ sein könnte, also $z_n' = \mathcal{A} y_n + \mathcal{B} y_{n-1} + ....$ Wir erhielten also : $a_n'(\mathcal{A} y_n + ...)^2 +$ Fct. von $y_1 ... y_{n-1}$ allein.

Führen wir das Quadrat aus, so kommen Produkte von $y_n$ mit den übrigen $y$ vor, welche sich nicht heben können, also 0 sein müssen. Hieraus folgt :

$$\sum_{i=1}^{n} a_i z_i^2 = \sum_{h=1}^{n-1} a_h' z_h'^2 + \mathcal{A} y_n'.$$

Jetzt haben wir es genau so mit einer Summe von $n-1$ Quadraten zu machen und erhalten :

$$\sum_{i} a_i z_i^2 = \sum_{g=1}^{n-2} a_g'' z_g'' + \mathcal{B} y_{n-1}^2 + \mathcal{A} y_n^2.$$

Gehen wir so weiter fort, so gelangen wir auf diese Weise schließlich auf eine Summe von Quadraten der $y$ allein, die natürlich mit unserer Summe $\sum_i a_i z_i^2$ identisch ist, so daß $\mathcal{A} = b_n$, $\mathcal{B} = b_{n-1}$ u.s.w.

Wir haben also auf diese Weise rückwärts eine Summe von Quadraten linearer Funktionen der $y$ verwandelt in eine Summe der Quadrate der $y$ selbst und zwar immer nur durch unsere elementaren Transformationen. Diese hängen nur ab von einem Parameter $t$, so daß wir im Ganzen $\frac{n.n-1}{2}$ der $t$ haben. Da wir aber außerdem für jedes $y$ noch einen Faktor $b$ brauchen, so erhäht sich die Anzahl | der nothwendigen Con- 144|145
stanten auf $\frac{n.n+1}{2}$ und wir erkennen, daß unser Verfahren wirklich aus $\frac{n.n+1}{2}$ Schritten besteht, daß es also als vom Geschlechte 0 zu bezeichnen ist.

Da wir nun erkannt haben, daß jede Transformation sich aus elementaren zusammensetzen läßt und bei den elementaren die Constanze der Vorzeichen nachgewiesen ist, so wissen wir, daß der Satz der Constanz der Vorzeichen auch für die allgemeine Transformation $\sum a_i z_i^2 = \sum b_h y_h^2$ gilt, so daß $\sum[a_i] = \sum[b_i]$ ist.

Mithin ist die Anzahl der Vorzeichen der Coeffizienten einer quadrat. Form eine Invariante.

§58. Wir wollen nun das gefundene Resultat, daß, wie man auch eine Summe

von Quadraten in eine andre Summe von Quadraten verwandeln möge, die Summe der positiven Zeichen, ebenso wie die der negativen, stets dieselbe bleibt, auf den *Sturm*'schen Satz anwenden. Wir werden sehen, daß der *Sturm*'sche Satz nur ein andrer Ausdruck für das Trägheitsprinzip der quadratischen Formen ist. Wir haben nämlich vermöge dieses Prinzips :

$$\sum_i [a_i z_i^2] = \sum_h [b_h y_h^2]$$

oder :

$$\sum_i [a_i] = \sum_h [b_h]$$

Wenden wir dies auf die quadratische Form :

$$\sum_\xi \frac{1}{f_1(\xi) f'(\xi)} (y_1 f_1(\xi) + ... + y_n f_n(\xi))^2 = \sum_k S_k^2 \text{ an.}$$

Wir nehmen an, $f$ habe reelle Coeffizienten, unter den $\xi$ befänden sich aber auch imaginäre und zwar konjugirt imaginäre Werthe. Die $y$ seien reell. Dann erscheint :

$$\frac{y_1 f_1(\xi) + ... + y_n f_n(\xi)}{\sqrt{f_1(\xi) f'(\xi)}}$$

unter der Form $u + vi$, wo $u$ und $v$ Funktionen von $y$ mit reellen Coeffizienten sind. Der konjugirte Werth von $\xi$ liefert dann $u - vi$. Wir erhalten daher für jedes Paar imaginärer Wurzeln als Beitrag zu unserer Summe den Ausdruck :

$$(u + vi)^2 + (u - vi)^2 = 2u^2 - 2v^2$$

145|146 | Für die Summe der Vorzeichen ist also :

$$\sum [2u^2] + \sum [-2v^2] = 0$$

d.h. sämmtliche imaginären Wurzeln zerstören sich, wenn wir für die quadratischen Formen nur die Summen der Vorzeichen ins Auge fassen. Wir haben also jetzt :

$$\sum_\xi [f_1(\xi) f'(\xi)] = \sum_k [S_k]$$

wo jetzt die Summe nur noch über die reellen $\xi$ zu erstrecken ist. Dies ist aber der *Sturm*'sche Satz für den Fall, daß man die Grenzen ins Unendliche rückt. Denn es ist :

$$S_k = \sum_\xi \frac{f_k(\xi)}{f_1(\xi) f'(\xi)} = \sum_\xi \frac{f_k(\xi) \psi_k(\xi)}{f'(\xi)}.$$

Dies reduzirt sich nach den *Euler*'schen Formeln auf den Coeffizienten des höchsten Gliedes, da der Grad des Zählers $= n - 1$ ist, d.h. auf :

$$C_k \frac{1}{C_{k-1}} \text{ folglich ist : } \sum_k [S_k] = \sum_k [C_k C_{k-1}] \text{ (Siehe } pag....)$$

Diese Deduktion des *Sturm*'schen Satzes hat jedoch den Mangel, daß wir zuerst den Beweis der Existenz komplexer Wurzeln zu bringen haben, welche doch bei dem Endresultat wieder vollständig herausfallen. Wir wollen daher sofort den allgemeinen Fall des *Sturm*'schen Satzes ins Auge fassen, welcher uns zugleich zu den *Jacobi*'schen Formeln führt. Auch dieser Weg wird erkennen lassen, daß das Prinzip der Erhaltung der Vorzeichen geradezu identisch mit dem *Sturm*'schen Satze ist.

Bestimmen wir jetzt allgemeinere Funktionen $F_i(\xi)$ $(i = 1, ...n)$, welche so be-

schaffen seien, daß :

$$\sum_{\xi} \frac{F_i(\xi)F_k(\xi)}{f_1(\xi)f'(\xi)} = \delta_{i,k} \sum_{\xi} \frac{(F_k(\xi))^2}{f_1(\xi)f'(\xi)} = \delta_{ik} S_k.$$

Daß es solche Funktionen $F_i$ gibt, zeigen die Restfunktionen $f_i$. Setzen wir : $\sum \xi \frac{F_i(\xi)F_k(\xi)}{f_1(\xi)f'(\xi)} = S_i \mathcal{A}_{ik}$ und vereinigen diese Gleichung mit der vorhergehenden, nachdem wir letztere mit $x$ multiplizirt haben, so kommt :

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{i=1}^{n} F_i(\xi) v_i \right)^2 = \sum_{i,k=1}^{n} (x\delta_{ik} - \mathcal{A}_{ik}) S_i v_i v_k.$$

Diese Gleichung stellt die Transformation einer quadratischen Form in die Summe von Quadraten dar. Wir hatten nun :

$$f_{k-1} - g_k f_k + f_{k+1} = 0.$$

| $g_k$ ist für den regularen Fall, den wir ja hier immer voraussetzen, eine lineare 146|147
Funktion von $x$. Sei $g'_k$ die Ableitung von $g_k(x)$, so setzen wir jetzt $F_k(x) = g'_k f_k$, wodurch unsern Voraussetzung über $F_k(x)$ erfüllt ist.

Wir wandeln nun unter dieser Annahme für $F_i$ die Summe der Quadrate um, indem wir die Glieder mit $i = k$ von den Gliedern $i \neq k$ absondern.

Unter steter Benutzung der *Euler*'schen Formeln erhalten wir dann ohne Weiteres :

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{i=1}^{n} F_i(\xi) v_i \right)^2 = \sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{i=1}^{n} g'_i f_i(\xi) v_i \right)^2 =$$
$$\sum_{\xi,k} \frac{x-\xi}{f_1(\xi)f'(\xi)} (f_i(\xi))^2 g'^2_k v_i^2 - 2 \sum_{\xi,k=1,\ldots n-1} \xi \frac{f_k(\xi)f_{k+1}(\xi)}{f_1(\xi)f'(\xi)} g'_k g'_{k+1} v_k v_{k+1}$$

Nun ist : $(x-\xi)g'_k = g_k(x) - g_k(\xi)$,
ferner : $f_k(\xi) = f_1(\xi)\psi(\xi)$. Daher wird unsere erste Summe zu :

$$\sum_{\xi,k} (g_k(x) - g_k(\xi)) g'_k f_k(\xi) \psi_k(\xi) \frac{v_k^2}{f'(\xi)}.$$

Zerlegt man diese Summe, durch Auflösen der Klammer in zwei Theile, berücksichtigt die bekannten Grade der Funktionen $f, g, \psi$, ferner den Umstand, daß der höchste Coeffizient von $\psi_k$ gleich $\frac{1}{C_{k-1}}$ ist, und daß endlich die Gleichung $g_k f_k = f_{k-1} + f_{k+1}$ besteht, so erhält man mittelst der *Euler*'schen Relationen für den ersten Fall unserer Summe :

$$[\sum_{\xi,k} (g_k(x) - g_k(\xi)) g'_k f_k(\xi) \psi_k(\xi) \frac{v_k^2}{f'(\xi)} = \sum_{k=1}^{n} g_k(x) v_k^2.$$

Der zweite Theil wird zu :

$$-2 \sum_{k=1}^{n-1} v_k v_{k+1} \sum_{\xi} \frac{\xi f_k(\xi) f_{k+1}(\xi)}{f_1(\xi)f'(\xi)} g'_k g'_{k+1}. \text{ Nun ist : } \frac{f_k(\xi)}{f_1(\xi)} = \psi_k(\xi)$$

und wir erkennen jetzt, daß der Grad des Zählers in $\sum_\xi$ gleich $n-1$ ist, daß also nach den *Euler*'schen Formeln diese Summe sich auf den Coeffizienten des höchsten Gliedes reduzirt. Der *Euler*'schen Formeln wegen können wir nun $\xi g'_{k+1}$ durch $g_{k+1}(\xi)$

ersetzen und da

$$g_{k+1}(\xi)f_{k+1}(\xi) = f_k(\xi) + f_{k+2}(\xi) \text{ ist,}$$

147|148 | so bleibt nur der Ausdruck :

$$g'_k f_k(\xi)\psi_k(\xi)$$

zu berücksichtigen ; der Coeffizient seines höchsten Gliedes aber ist gleich 1, denn aus $g_k f_k = f_{k-1} + f_{k+1}$ folgt $g'_k C_k = C_{k-1}$.

Wir haben also jetzt folgendes Endresultat :

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{k=1}^{n} g'_k f_k(\xi) v_k \right)^2 = \sum_{k=1}^{n} g_k(x) v_k^2 - 2 \sum_{k=1}^{n-1} v_k v_{k+1}.$$

Diese eigenthümliche Transformation einer Summe von Quadraten :

$$\sum_{h=1}^{n} \frac{x-\xi}{f_1(\xi)f'(\xi)} z_h^2 \text{ in den Ausdruck :}$$

$$\sum_{k=1}^{n} g_k(x) v_k^2 - 2 \sum_{k=1}^{n-1} v_k v_{k+1} \text{ mittelst der Substitution :}$$

ist zuerst von *Jacobi (Crelle Bd. 39*, Seite 290) gegeben werden, man nennt daher nach *Kronecker* diese Formen die *Jacobi*'schen Formen [15].

Von den *Jacobi*'schen Formen zu einer Summe von Quadraten überzugehen, ist höchst einfach.

Es ist nämlich :

$$g_k(x) = \frac{f_{k-1}(x)}{f_k(x)} + \frac{f_{k+1}(x)}{f_k(x)}.$$

Summiren wir daher über $k$, so können wir immer

$$\frac{v_{k-1}^2 f_k}{f_{k-1}} + \frac{v_k^2 f_{k-1}(x)}{f_k(x)} - 2v_k v_{k-1}$$

zusammenfassen und erkennen, daß wir die folgende Summe von Quadraten erhalten :

$$\sum_{k=1}^{n} f_k f_{k-1} \left( \frac{v_{k-1}}{f_{k-1}} - \frac{v_k}{f_k} \right)^2$$

indem wir $v_0 = 0$ setzen.

Wenn wir also eine quadratische Form auf die *Jacobi*'sche Form gebracht haben, so ist die weitere Transformation zu einer Summe von Quadraten außerordentlich einfach. Wenden wir nun das Prinzip von der Erhaltung der Vorzeichen an, so kommt :

$$\sum_{\xi} [(x-\xi)f_1(\xi)f'(\xi)] = \sum_{k=1}^{n} [f_k f_{k-1}].$$

148|149 | Hierin werden wir aber sofort den *Sturm*'schen Satz erkennen. $x$ sei nämlich eine reelle Größe ; setzen wir jetzt $x = x_1$ und $x = x_2, x_1 < x_2$ und ziehen die beiden Gleichungen von einander ab, so verschwinden diejenigen Terme, welche sich auf $\xi$ beziehen, die außerhalb des Intervalles $x_1 ... x_2$ liegen und wir erhalten :

$$2 \sum_{\xi} [f_1(\xi)f'(\xi)] = \sum_{k=1}^{n} [f_k(x_2) f_{k-1}(x_2)] - \sum_{k=1}^{n} [f_k(x_1) f_{k-1}(x_1)]$$

15. Ueber den Inhalt der *Jacobi*'schen Abhandlung s. Anhang *pag....*

die Summe bezogen auf alle in dem Intervalle $x_1 ... x_2$ liegenden $\xi$. Damit haben wir also den *Sturm*'schen Satz als identisch mit dem Trägheitsprinzip der quadratischen Formen nachgewiesen.

§59. Wenn wir in der Gleichung :

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{k=1}^{n} g'_k v_k f_k(\xi) \right)^2 = \sum (g_k x + g^0_k) v_k^2 - 2 \sum v_k v_{k+1}$$

die mit $x$ behafteten von den von $x$ freien Gliedern trennen, so kommt, da ja $x$ ganz beliebig ist :

$$\sum_{h=1}^{n} \frac{1}{f_1(\xi_h)f'(\xi_h)} z_h^2 = \sum_{k=1}^{n} g'_k v_k^2$$

und

$$\sum_{h=1}^{n} \frac{\xi_h}{f_1(\xi_h)f'(\xi_h)} z_h^2 = -\sum_{k=1}^{n} g^0_k v_k^2 - 2 \sum_{k=1}^{n-1} v_k v_{k+1}$$

mit Hülfe der Substitution :

$$z_h = \sum_{k=1}^{n} g'_k v_k f_k(\xi)$$

d.h. wir haben mit Hülfe einer orthogonalen Substitution (d.h. einer solchen, welche eine gegebene Summe von Quadraten, nämlich : $\sum_h \frac{1}{f_1(\xi_h)f'(\xi_h)} z_h^2$ wieder in eine Summe von Quadraten, nämlich in : $\sum g'_k v_k^2$ verwandelt.) eine Summe von Quadraten in eine *Jacobi*'sche Form verwandelt. Wir können nun ganz allgemein die Aufgabe lösen, eine beliebige quadratische Form mittelst einer orthogonalen Substitution in eine *Jacobi*'sche Form zu verwandeln. Nehmen wir nämlich für die Funktionen $F_i$ die Bedingung an, daß :

$$\sum_{\xi} \frac{F_i(\xi)F_k(\xi)}{f_1(\xi)f'(\xi)} = \delta_{ik} S_k \text{ sei und setzen :}$$

$$\sum \frac{\xi F_i(\xi)F_k(\xi)}{f_1(\xi)f'(\xi)} = S_i \mathcal{A}_{ik} \text{ so ist :}$$

149|150

|

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{k=1}^{n} v_k F_k(\xi) \right)^2 = \sum_{i,k} (x\delta_{ik} - \mathcal{A}_{ik}) S_i v_i v_k$$

d.h. wir haben mittelst der Substitution : $z_h = \sum_{k=1}^{n} v_k F_k(\xi_h)$ eine Summe von Quadraten in eine gewöhnlich quadratische Form verwandelt. Hierbei sind aus den $\xi_h$ und $F_k$, welche gegeben sind, die $\mathcal{A}_{ik}$ und $S_i$ hergeleitet, mittelst der obigen Relationen.

Wir wollen aber jetzt umgekehrt die $S_i$ und $\mathcal{A}_{ik}$ als gegeben betrachten, so daß auf der rechten Seite der letzten Gleichung eine beliebige quadratische Form steht. Wir wollen nun beweisen, daß sich die $\xi_h$ und $F_k$ dann aus den $S_i$ und $\mathcal{A}_{ik}$ so bestimmen lassen, daß obige Gleichung besteht, so daß wir jede beliebige quadratische Form in jene Summe von Quadraten verwandeln können. Ist dieser Nachweis gelungen, so liefert die Zerlegung der Gleichung nach $x$ die beiden folgenden Gleichungen :

$$\sum_{k=1}^{n} S_k v_k^2 = \sum_{h=1}^{n} \frac{1}{f_1(\xi_h)f'(\xi_h)} z_h^2 = \sum_{k=1}^{n} g'_k v_k^2$$

$$\sum_{i,k} S_i \mathcal{A}_{ik} v_i v_k = \sum_{h=1}^{n} \frac{\xi_h}{f_1(\xi_h) f'(\xi_h)} z_h^2 = \sum_{k=1}^{n} g_k^0 v_k^2 - 2 \sum_{k=1}^{n-1} v_k v_{k+1}$$

d.h. wir haben alsdann die beliebige quadratische Form in eine *Jacobi*'sche verwandelt, unter simultaner Verwandlung der gegebeben Summe von Quadraten $\sum_{k=1}^{n} S_k v_k^2$ in eine andere Summe von Quadraten, damit wäre dann unsere Behauptung erwiesen. Wir bemerken, daß diese orthogonale Transformation der beliebigen quadratischen Form in eine *Jacobi*'sche rational ist, während der erste Schritt, die Verwandlung der quadratischen Form in eine Summe von Quadraten, die Irrationale $\xi$ einführt, welche bei dem Uebergange von der Summe von Quadraten in die *Jacobi*'sche Form wieder herausfällt.

Wir haben also jetzt zu zeigen, daß bei gegebenen $S_i$ und $A_{ik}$ sich $\xi_h$ und $F_k$ so bestimmen lassen, daß die Gleichung stattfindet :

$$\sum \frac{x-\xi}{f_1(\xi) f'(\xi)} \left( \sum v_k F_k(\xi) \right)^2 = \sum (x\delta_{ik} - \mathcal{A}_{ik}) S_i v_i v_k.$$

Nun gibt der Satz, daß, wenn man eine quadratische Form linear transformirt, die Determinante der neuen Form gleich der der alten ist, multiplizirt mit dem Quadrate der Substitutionsdeterminante.

150|151 | Wenden wir dies hier an, so folgt :

$$|F_k(\xi_h)|^2 \frac{\prod(x-\xi)}{\prod f_1(\xi) f'(\xi)} = |x - \delta_{ik} - \mathcal{A}_{ik}| \prod S_k.$$

Dividiren wir durch den Coeffizienten der höchsten Potenz von $x$, so folgt :

$$f(x) = \prod (x-\xi) = |x\delta_{ik} - \mathcal{A}_{ik}|.$$

Wir wollen dabei von der Determinante voraussetzen, daß sie keine gleichen Wurzeln habe. Damit ist also $f(x)$ d.h. die $\xi$ vollständig bestimmt, es handelt sich also nur noch darum die $F_k$ obiger Gleichung gemäß zu bestimmen.

Wir erfahren nun andrerseits, daß jede Gleichung $f(x) = 0$ sich in diese eigenthümliche Determinantenform bringen läßt. Zu diesem Zwecke können wir $f_1(x)$ willkürlich wählen und bestimmen dann $F_i$, so daß $\sum_\xi \frac{F_i(\xi) F_k(\xi)}{f_1(\xi)(\xi)} = \delta_{ik} S_k$ ist. Bezeichnet man dann $\sum_\xi \frac{\xi F_i(\xi) F_k(\xi)}{f_1(\xi)(\xi)} = \delta_{ik} S_k$ mit $S_i \mathcal{A}_{ik}$, so ist die gewünschte Determinantendarstellung von $f(x)$ die folgende :

$$f(x) = |x\delta_{ik} - \mathcal{A}_{ik}| \ (i,k = 1, ...h).$$

Diese Darstellung kann, da $f_1$ und abgesehen von einer Beschränkung auch die $F_i$ willkürlich sind, auf unendlich viele Arten hergestellt werden. Es besteht nun die Gleichung $S_i \mathcal{A}_{ik} = S_k \mathcal{A}_{ki}$ ; wäre nun $\mathcal{A}_{ik} = \mathcal{A}_{ki}$, so wäre $f(x) = 0$ die Gleichung der Säkularstörungen. In dem speziellen Falle :

$$\sum_\xi \frac{x-\xi}{f_1(\xi) f'(\xi)} \left( \sum_k g_k' v_k f_k(\xi) \right)^2 = \sum g_k v_k^2 - 2 \sum v_k v_{k+1}$$

nimmt $f(x)$ die Gestalt an :

$$g_1' ... g_n' f(x) = \begin{vmatrix} g_1 & -1 & 0 & ... & ... & ... & 0 \\ -1 & g_2 & -1 & 0 & ... & ... & 0 \\ 0 & -1 & g_3 & -1 & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & ... & ... & -1 & g_{n-1} & -1 \\ 0 & ... & ... & ... & 0 & -1 & g_n \end{vmatrix}$$

| Hieraus ergibt sich sofort, daß alle Restfunktionen $f_k(x)$ sich als Theildeterminanten darstellen, welche aus jener entstehen, wenn man die ersten $k$ Horizontal- und Vertikalreihen wegstreicht. 151|152

Nehmen wir nach diesen Zwischenbemerkungen unsern Untersuchung wieder auf. Bezeichnen wir durch $f_{ik}(x)$ die Unterdeterminanten von $f(x)$, so ist :

$$\sum_{i=1}^{n}(x_i\delta_{ik} - \mathcal{A}_{ik})f_{ri}(x) = \delta_{kr}f(x) \text{ und ebenso :}$$
$$\sum(x\delta_{ik} - \mathcal{A}_{ki})f_{ir}(x) = \delta_{kr}f(x).$$

Nun ist : $\frac{S_1 \ldots S_n}{S_k} f_{i,k}(x)$ die Unterdeterminante das symmetrischen Systems $|(x\delta_{ik} - \mathcal{A}_{ki})S_k|$, folglich selbst symmetrisch ; es ist also :

$$\frac{f_{ir}(x)}{S_r} = \frac{f_{ri}(x)}{S_i} \text{ oder } S_r f_{ri}(x) = S_i f_{ir}(x).$$

Nun brauchen wir folgenden Determinantensatz : Sei $D$ die Determinante der $a_{ik}$ und $\alpha_{ki}$ der Faktor von $a_{ik}$, so ist zunächst $\sum_i \alpha_{hi}a_{ik} = \delta_{hk}D$. Setzen wir nun $\alpha'_{hi} = \alpha_{hi}$ für $h = 1$ und $h = 2$ und $\alpha'_{hi} = \delta_{hi}$ für $h = 3, \ldots n$, so ist :

$$\sum_i \alpha'_{hi}a_{ik} = \begin{cases} \delta_{h,k}D \text{ für } h = 1 \text{ und } h = 2 \\ a_{h,k} \text{ für } h = 3, ..n. \end{cases}$$

Also ist nach dem Multiplikationsgesetz :

$$|\alpha'_{hi}|D = \left|\begin{array}{cc|cc} D & 0 & \ldots & \ldots \\ 0 & D & 0 & \ldots \\ \hline a_{31} & a_{32} & .. & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{array}\right| = D^2\frac{\partial^2 D}{\partial a_{11}\partial a_{22}}.$$

Also ist :

$$\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} = D\frac{\partial^2 D}{\partial a_{11}\partial a_{22}}.$$

Wenden wir diesen Satz an, so folgt :

$$f_{ir}(\xi)f_{rk}(\xi) - f_{ik}(\xi)f_{rr}(\xi), \text{ da ja } f(x) = D_{x=\xi} = 0 \text{ ist.}$$

Hieraus folgt :

$$S_r f_{ri}(\xi)f_{rk}(\xi) = S_i f_{ik}(\xi)f_{rr}(\xi).$$

Nun ist

$$\sum_\xi S_r\frac{x-\xi}{f_{rr}(\xi)f'(\xi)}\left(\sum_k f_{rk}(\xi)v_k\right)^2 = \sum_{\xi,i,k}\frac{(x-\xi)S_i f_{ik}(\xi)v_i v_k}{f'(\xi)}.$$

| Summiren wir erst nach $\xi$, so findet man mit Hülfe der *Euler*'schen Formeln : 152|153

$$\sum_\xi \frac{f_{ik}}{f'(\xi)} = \delta_{ik}$$
$$\sum_\xi \frac{\xi f_{ik}}{f'(\xi)} = \mathcal{A}_{ik}.$$

Wir erhalten daher für unsere Ausdruck geradezu $\sum_{i,k}(x\delta_{ik} - \mathcal{A}_{ik})S_i v_i v_k$. Dem-

nach ist diese allgemeine quadratische Form in die gewünschte Form dadurch zu bringen, daß man setzt :

$$f(x) = |x\delta_{ik} - \mathcal{A}_{ik}| \ , \ f_1(x) = \frac{f_{rr}(x)}{S_r} \text{ und } F_k(x) = f_{rk}(x).$$

Bestimmt man $f, f_1, F_k$ in dieser Weise, so ist :

$$\sum_{i,k}(x\delta_{ik} - \mathcal{A}_{ik})S_i v_i v_k = \sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left(\sum_{k=1}^{n} F_k(\xi)v_k\right)^2 .$$

Wir können das gefundene Resultat auch so aussprechen : Es sei die beliebige quadratische Form $\sum_{i,k} \mathcal{A}_{ik}S_i v_i v_k$ gegeben und außerdem die Summe von Quadraten $\sum_k S_k v_k^2$. Dann haben wir jetzt die simultane Transformation :

$$\begin{cases} \sum_k S_k v_k^2 = \sum_h \frac{S_r}{f_{rr}(\xi_h)f'(\xi_h)} z_h^2 \\ \sum_{i,k} \mathcal{A}_{ik}S_i v_i v_k = \sum_h \frac{\xi_h S_r}{f_{rr}(\xi_h)f'(\xi_h)} z_h^2 \end{cases}$$

ausgeführt mit Hülfe der Substitution :

$$z_h = \sum_k f_{rk}(\xi_h)v_k.$$

Dabei sind die Funktionen $f_{ik}, f', f$ erklärt durch $f(x) = |x\delta_{ik} - \mathcal{A}_{ik}|$. Diese Transformation ist indessen, wie schon bemerkt, nicht rational, denn die Coeffizienten enthalten die Irrationalen $\xi$. Gehen wir nun aber zu der *Jacobi*'schen Form über, so verschwindet diese Irrationalität. Wir hatten :

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left(\sum_k v_k g'_k f_k(\xi)\right)^2 = \sum g_k v_k^2 - 2\sum v_k v_{k+1}.$$

Wir erhalten daher $\sum_{i,k}(x\delta_{ik} - \mathcal{A}_{ik})S_i v_i v_k$ in der *Jacobi*'schen Form, wenn wir die Summe von Quadraten, in welchen wir diese allgemeine quadratische Form transfor-
153|154 mirt haben, identifiziren mit | der linken Seite der letzten Gleichung durch :

$$\frac{S_r \left(\sum_k f_{rk}(\xi)v_k\right)^2}{f_{rr}(\xi)} = \frac{\left(\sum_k g'_k v_k f_k(\xi)\right)^2}{f_1(\xi)}.$$

Es muß daher $\frac{S_r}{f_{rr}(\xi)} f_1(\xi)$ gleich dem Quadrate einer ganzen Funktion $\varphi(\xi)$ sein. Bestimmen wir daher nach der *Lagrange*'schen Formel eine Funktion $n-1$ ten Grades :

$$S_r f_1(\xi) = (\varphi(\xi))^2 f_{rr}(\xi)$$

wo $\varphi(\xi)$ eine beliebige ganze Funktion ist, die wir auf den $n-1$ ten Grad reduzirt denken können, so führt die Substitution :

$$\varphi(\xi)\sum_k f_{rk}(\xi)v_k = \sum_k g'_k v_k f_k(\xi).$$

zu der simultanen Transformation :

$$\begin{cases} \sum_k S_k v_k^2 = \sum_k g'_k v_k^2 \\ \sum_{i,k} \mathcal{A}_{ik}S_i v_i v_k = -\sum_k g_k^0 v_k^2 + 2\sum_k v_k v_{k+1} \end{cases}$$

Fassen wir diesen ganzen Transformationsprozeß zusammen. Aus den gegeben $S_i$ und $\mathcal{A}_{ik}$ bilde man $f(x) = |x\delta_{ik} - \mathcal{A}_{ik}|$ und hieraus $f_{ik}$ und $f'$, so können wir nach den Formeln *pag....* mittelst der Substitution $z_h = \sum_k f_{rk}(\xi_h)v_k$ die beiden gegebenen Formen in zwei Summen von Quadraten verwandeln. Diese Transformation ist nicht

rational. Wählen wir nun, um zur *Jacobi*'schen Form zu gelangen, eine beliebige ganze Funktion $\varphi(x)$ von Grade $n-1$ und bestimmen nach *Lagrange* $f_1(x)$ so, daß

$$S_r f_1(\xi_h) = (\varphi(\xi_h))^2 f_{rr}(\xi_h) \ (h = 1, ...n)$$

ist und wenden sodann auf $f$ und $f_1$ das Kettenbruchverfahren an, welches uns die $f_k$ und $g_k$ liefert, so geht durch die Substitution :

$$\varphi(\xi_h) \sum_k f_{rk}(\xi_h) v_k = \sum_k v_k g'_k f_k(\xi_h) \ (h = 1, ...n)$$

folgende rationale simultane Transformation vor sich :

$$\begin{cases} \sum_k S_k v_k^2 = \sum_k g'_k v_k^2 \\ \sum_{i,k} \mathcal{A}_{ik} S_i v_i v_k = -\sum_k g_k^0 v_k^2 + 2 \sum_k v_k v_{k+1} \end{cases}$$

| Da $\varphi$ $n$ willkürliche Constanten besitzt, so ist eine solche Transformation durch 154|155
eine $n$ fache Mannigfaltigkeit von Substitutionen zu erreichen.

Wir können hieran ohne Weiteres den Beweis für die Realität der Wurzeln der Gleichung für die *Säcular*störungen knüpfen, welche Gleichung außordentlich oft behandelt wurde. Eigentlich bedarf es kaum eines Beweises. Die Wurzeln sind bereits reell, wenn alle $S_i$ positiv sind. Für die Gl. der säkularen Störungen ist $S_i = 1$. Wir haben nun :

$$\sum_k [S_k] = \sum_\xi [f_{rr}(\xi) f'(\xi)].$$

Sind alle $S$ positiv, so steht $n$ auf der linken Seit. Rechts ist zu summiren über alle reellen $\xi$, deren Anzahl folglich $n$ ist.

§60. Bei dem *Sturm*'schen Satze hatten wir die Summe :

$$\sum_{x_1 < \xi < x_2} [f_1(\xi) f'(\xi)] \text{ zu betrachten.}$$

Führen wir nun statt $f_1$ und $f$ 2 Funktionen $\varphi(z), \psi(z)$ ein, welche wir von gleichen Grade voraussetzen wollen. Die reellen Wurzeln von $\varphi(z) = 0$ seien durch $\xi$ und diejenigen von $\psi(z) = 0$ durch $\eta$ bezeichnet. $y = \varphi(z)$ und $y = \psi(z)$ sind dann 2 parabolische Curven gleichen Grades. Nun hatten wir für eine beliebige stetige Funktion $F(z)$ die Relation :

$$\sum_{z_1 < \xi < z_2} [F'(\xi)] = -\frac{1}{2}[F(z_1)] + \frac{1}{2}[F(z_2)] \ (pag....)$$

Setzen wir jetzt $F(z) = \varphi(z)\psi(z)$, so kommt :

$$\sum_{z_1 < \xi < z_2} [\varphi'(\xi)\psi(\xi)] + \sum_{z_1 < \eta < z_2} [\varphi(\eta)\psi'(\eta)] = -\frac{1}{2}[\varphi(z_1)\psi(z_1)] + \frac{1}{2}[\varphi(z_2)\psi(z_2)]$$

Setzen wir nun $\Delta(z) = \begin{vmatrix} \varphi & \psi \\ \varphi' & \psi' \end{vmatrix}$ so ist :

$$-\sum [\Delta(\xi)] + \sum [\Delta(\eta)] = -\frac{1}{2}[\varphi(z_1)\psi(z_1)] + \frac{1}{2}[\varphi(z_2)\psi(z_2)].$$

Setzen wir jetzt $z_1 = -\infty$, $z_2 = +\infty$, summiren also über alle reellen $\xi$ und $\eta$, so

verschwindet die rechte Seite, da $\varphi$ und $\psi$ von demselben Grade sind und wir haben :

$$-\sum_{\xi}[\Delta(\xi)]+\sum_{\eta}[\Delta(\eta)]=0.$$

155|156 | Nehmen wir nun an, daß $\varphi$ und $\psi$ von gradem Grade seien, so ist :

$$\sum_{\xi}[\varphi'(\xi)]=\left\{-\frac{1}{2}[\varphi(z_1)]+\frac{1}{2}[\varphi(z_2)]\right\}\begin{matrix} z_1=-\infty \\ z_2=+\infty \end{matrix}=0,$$

ebenso ist $\sum_{\eta}[\psi'(\eta)]=0$.

Theilen wir nun beide Gleichungen, so kommt :

$$\begin{cases} \sum\limits_{\xi\ ,\ \psi(\xi)>0}[\varphi'(\xi)]=-\sum\limits_{\xi\ ,\ \psi(\xi)<0}[\varphi'(\xi)] \\ \sum\limits_{\eta\ ,\ \varphi(\eta)>0}[\psi'(\eta)]=-\sum\limits_{\eta\ ,\ \varphi(\eta)<0}[\varphi'(\eta)] \end{cases}$$

Wir erhalten daher :

$$\begin{aligned} -\sum[\Delta(\xi)]&=\sum[\varphi'(\xi)\psi(\xi)]=\sum_{\psi(\xi)>0}[\varphi'(\xi)]-\sum_{\psi(\xi)<0}[\varphi'(\xi)] \\ &=2\sum_{\psi(\xi)>0}[\varphi'(\xi)]=-2\sum_{\psi(\xi)<0}[\varphi'(\xi)]. \end{aligned}$$

Es ist also $-\frac{1}{2}\sum_{\xi}[\Delta(\xi)]=-\frac{1}{2}\sum_{\eta}[\Delta(\eta)]$ eine ganze Zahl, welche wir durch $\chi(\varphi,\psi)$ bezeichnen wollen. Wir nennen dieselbe die Charakteristik des Funktionensystems $\varphi$ und $\psi$, sie ist gleich der halben Summe der Vorzeichen der Determinante $\Delta$, für alle Wurzeln $\xi$ von $\varphi(z)=0$ oder für alle Wurzeln $\eta$ von $\psi(z)=0$. Die Charakteristik ist abhängig von der Reihenfolge der beiden Funktionen $\varphi$ und $\psi$, denn aus $\begin{vmatrix}\varphi & \psi\\ \varphi' & \psi'\end{vmatrix}=-\begin{vmatrix}\psi & \varphi\\ \psi' & \varphi'\end{vmatrix}$ folgt $\chi(\varphi,\psi)=-\chi(\psi,\varphi)$. Es ist auch :

$$\chi=\frac{1}{2}\sum_{z=\xi}[d(\varphi\psi)]=-\frac{1}{2}\sum_{z=\eta}[d(\varphi\psi)]$$

Aus

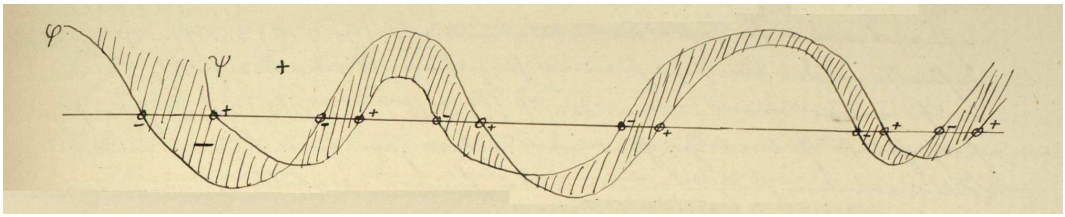
$$\begin{vmatrix}\varphi & \psi\\ \varphi' & \psi'\end{vmatrix}\cdot\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}=\begin{vmatrix}\alpha\varphi+\beta\psi & \gamma\varphi+\delta\psi\\ \alpha\varphi'+\beta\psi' & \gamma\varphi'+\delta\psi'\end{vmatrix}$$

folgt, wenn man unter $\alpha,\beta,\gamma,\delta$ beliebige Constanten versteht :

$$\chi(\alpha\varphi+\beta\psi,\gamma\varphi+\delta\psi)=[\alpha\delta-\beta\gamma]\chi(\varphi,\psi).$$

Für eine positive Substitutionsdeterminante ändert sich also $\chi$ durch eine lineare Substitution nicht. Im Allgemeinen ändert die Charakteristik also durch eine lineare Substitution höchstens ihr Zeichen, nicht aber ihren Werth.

156|157 Nach den Untersuchungen des 45, *pag....* hat, wenn | wir $\varphi$ und $\psi$ d.h. $y=\varphi(z)$, $y=\psi(z)$ graphisch darstellen. Die Charakteristik folgende geometrische Bedeutung : Es ist die Charakteristik $\chi(\varphi,\psi)$ der halbe Ueberschuß der Austritte über die Eintritte, wenn wir uns auf der Curve $y=\varphi(z)$ bewegen, d.h. die $\xi$ zählen und gleich dem halben Ueberschuß der Eintritte über die Austritte, wenn wir uns auf der Curve $y=\psi(z)$ bewegen, d.h. die $\eta$ zählen (In dem Falle, den die Zeichnung gibt, ist $\chi=-1$).




Setzen wir nun $\varphi(z) = f(z)$, $\psi(z) = (x - z)f_1(z)$, so ist :

$$\chi(f(z), (x - z)f_1(z)) = \frac{1}{2}\sum[(x - \xi)f_1(\xi).f'(\xi)]$$

setzt man $x = x_1$ und $x = x_2$ und subtrahirt, so erhält man den *Sturm*'schen Satz (*pag....*). Die linke Seite des *Sturm*'schen Satzes (*pag....*) ist daher die Aenderung der Charakteristik beim Uebergange von $x_1$ zu $x_2$. Diese Aenderung ist nach dem Kettenbruchverfahren durch die Coeffizienten von $f$ und $f_1$ ausgedrückt und bestimmt, wie wir sahen, die Anzahl der zwischen $x_1$ und $x_2$ liegenden reellen Wurzeln von $f(z) = 0$, wenn $f'(\xi)$ und $f_1(\xi)$ gleiches Zeichen haben.

Das Wesentliche dieser Betrachtung ist also, daß man in den Coeffizienten der einen der beiden Funktionen $\varphi(z) = f(z)$ und $\psi(z) = (x - z)f_1(z)$ eine Variable $x$ läßt und nun die Aenderung der Charakteristik beider Funktionen bestimmt, welche durch eine Aenderung dieser Variabeln entsteht. Daraus erkennt man sofort, daß der *Sturm*'sche Satz nur ein spezieller Fall allgemeineren Sätze ist, denn wir können uns jetzt die Aufgabe stellen, die Aenderungen der Charakteristik $\chi(\varphi, \psi)$ von $\varphi(z)$ und $\psi(z)$ zu studiren, wenn wir alle Coeffizienten von $\varphi(z)$ sowohl als $\psi(z)$ variabel voraussetzen.

Wenn wir dann wissen, wie sich $\chi$ ändert, können wir $\chi$ selbst darstellen, indem wir von einem speziellen Falle ausgehen.

| Wenn sich die Charakteristik ändert, so kann dies nur dadurch geschehen, daß 157|158
entweder die Zahl der Summanden $\varphi'(\xi)\psi(\xi)$ und $\varphi(\eta)\psi'(\eta)$ bleibt, daß also weder ein $\xi$ noch ein $\eta$ hinzukommt oder wegfällt, und dann muß einer der Summanden $\varphi'(\xi)\psi(\xi)$ und $\varphi(\eta)\psi'(\eta)$ sich ändern hinsichtlich seines Zeichens oder aber zweitens können neue Summanden hinzukommen oder frühere wegfallen, indem eine Unterbrechung der Continuität in der Reihe de $\xi$ oder der $\eta$ eintritt.

Angenommen zunächst die Zahl der Summanden in $\chi$ ändere sich nicht, während die variabeln Coeffizienten sich ändern ; die Continuität der $\xi$ und $\eta$ werde also nicht unterbrochen. Dann kann sich $\chi$ nur dann ändern, wenn irgend ein $\varphi'(\xi)\psi(\xi) = 0$ und gleichzeitig irgendein $\varphi(\eta)\psi(\eta) = 0$ wird. Die nothwendigen und hinreichenden Bedingungen für eine Aenderung von $\chi$ sind also durch die 4 Eventualitäten ausgedrückt :

$$\begin{aligned}
&1.)\psi(\xi) = 0\ ,\ \varphi(\eta) = 0\\
&2.)\psi(\xi) = 0\ ,\ \psi'(\eta) = 0\\
&3.)\varphi'(\xi) = 0\ ,\ \varphi(\eta) = 0\\
&4.)\varphi'(\xi) = 0\ ,\ \psi'(\eta) = 0
\end{aligned}$$

Nehmen wir zweitens an, es trete nach ein Summand hinzu, oder allgemeiner, die Continuität der $\xi$ oder der $\eta$ werde gestört. Betrachten wir eine Funktion $F(x, t)$ mit dem variabeln Parameter $t$, so folgt aus

$$F(x, t) = 0 \qquad \frac{dx}{dt} = -\frac{\frac{\partial F}{\partial t}}{F'(x)}.$$

Die Continuität kann also nur zerstört werden, wenn $F'(x) = 0$ ist. Es muß also

gleichzeitig in unsern Falle

$$\varphi(z) = 0 \text{ und } \varphi'(z) = 0$$

oder

$$\psi(z) = 0 \text{ und } \psi'(z) = 0 \text{ sein,}$$

damit eine Unterbrechung der Continuität in der Reihe der $\xi$ oder der $\eta$ eintreten könne. Es müßte also entweder $\varphi'(\xi)$ oder $\psi'(\eta)$ verschwinden, aber diese Möglichkeiten befinden sich bereits unter den vier oben angeführten, so dass diese die nothwendigen und hinreichenden Bedingungen für eine Aenderung der Charakteristik enthalten.

158|159 Wir erkennen, daß die ersten drei Möglichkeiten sich darauf | reduziren, daß für einen und denselben Werth von $x$ gleichzeitig $\varphi(z) = 0$ und $\psi(z) = 0$ ist, d.h. daß die Resultante $R(\varphi, \psi)$ verschwindet.

In dem 4 ten Falle nun, welcher nur eintritt, wenn die Continuität der $\xi$ oder der $\eta$ verloren geht, können wir statt $\varphi$ und $\psi$ lineare Funktionen von $\varphi$ und $\psi$ einfuhren mit positiver Substitutionsdeterminante. In Allgemeinen wird dann dieser ungünstige Fall, daß nämlich gleichzeitig diese linearen Funktionen und ihre Ableitungen verschwinden, nicht eintreten. Setzen wir nämlich $\varphi(z) - w\psi(z)$ an die Stelle von $\psi(z)$. Dann müßte gleichzeitig sein :

$$\varphi(z) = w\psi(z)$$
$$\varphi'(z) = w\psi'(z).$$

Dann müßte aber für dieses $z$ die Gleichung $\varphi(z)\psi'(z) = \varphi'(z)\psi(z)$ stattfinden, d.h. dieser spezielle Werth von $z$ wäre unabhängig von $w$. Für $w = 0$, resp. $= \infty$ erhielten wir dann, daß gleichzeitig $\varphi$ und $\psi$ verschwinden müßten. Es kann also demnach die vierte Möglichkeit, wo nämlich die Continuität gestört wird, zwar für spezielle $\varphi$ und $\psi$ eintreten, für alle linearen Funktionen von $\varphi$ und $\psi$, aber nur, wenn gleichzeitig $\varphi(z) = 0$ und $\psi(z) = 0$ ist. Damit haben wir also alle Möglichkeiten auf $R(\varphi, \psi) = 0$ reduzirt. Diese Gleichung ist die nothwendige und hinreichende Bedingung dafür, daß die Charakteristik von $\varphi$ und $\psi$ sich ändern. Wir haben nun die Art der Aenderung der Charakteristik zu studiren.

§61. Die Charakteristik wird durch das Verschwinden eines Termes :

$$\frac{1}{2}[\varphi'(\xi)\psi(\xi)]$$

um eine Einheit vergrößert oder verkleinert, je nachdem für irgendein $\xi$ der genannte Term von $-\frac{1}{2}$ zu $+\frac{1}{2}$ oder von $+\frac{1}{2}$ zu $-\frac{1}{2}$ sich ändert. Es handelt sich also darum zu entscheiden, ob $\varphi'(\xi)\psi(\xi)$ aus dem positiven ins negative oder umgekehrt gehe.

Bilden wir nun $\sum_\xi \frac{1}{\varphi'(\xi)\psi(\xi)}$, so wird in dieser Summe derjenige Term überwiegen [16], dessen Aenderung wir eben untersuchen wollen. Der Ausdruck wird von $+\infty$

159|160 zu $-\infty$ oder umgekehrt übergehen, je | nachdem der betreffende Term $\varphi'(\xi)\psi(\xi)$ von positiven durch die Null in negative oder umgekehrt geht.

Nun ist aber der ganze Ausdruck eine symmetrische Funktion der $\xi$ und folglich eine rationale Funktion der Coeffizienten von $\varphi$ und $\psi$ und speziell eine ganze Funktion der Coeffizienten von $\varphi$. Multipliziren wir mit $R = \prod_\xi \psi(\xi)$ so erhalten wir :

$$R_1 = \sum_\xi \frac{R(\varphi, \psi)}{\varphi'(\xi)\psi(\xi)}$$

als ganze Funktion der Coeffizienten von $\varphi$ und $\psi$. Nun lassen sich zwei Multiplikatoren

16. Siehe Anhang *pag....*

$\varphi_1(z)$ und $\psi_1(z)$ vom Grade $n-1$ herleiten, so daß

$$R = \varphi_1\psi - \varphi\psi_1 \text{ ist.}$$

Wir erhalten dann zur Bestimmung von $\varphi_1$ und $\psi_1$ :

$$\varphi_1 = \frac{R}{\psi(\xi)} \ , \ \psi_1 = \frac{R}{\varphi(\eta)}.$$

Die *Lagrange*'sche Formel liefert dann :

$$\varphi_1(z) = \sum \frac{R}{\psi(\xi)\varphi'(\xi)} \frac{\varphi(z)}{z-\xi}.$$

Indem wir nun in $\varphi(z)$ und $\psi(z)$ die Coeffizienten der $n$ ten Potenz als gleich 1 voraussetzen, also :

$$\varphi(z) = z^n + ... \ , \ \psi(z) = z^n + ...$$

erhalten wir

$$\varphi_1(z) = R_1 z^{n-1+...}$$
$$\psi_1(z) = R_1 z^{n-1} + ...$$

d.h. $R_1$ ist der höchste Coeffizient sowohl in $\varphi_1$ als in $\psi_1$. Wenn nun $\varphi$ und $\psi$ einen gemeinsamen Theiler haben, so wird $\frac{R_1}{R}$ unendlich groß und damit auch $\Phi$ und $\Psi$. In der That darf ja für diesen Fall die Gleichung $\Phi\psi - \Psi\varphi = 1$ nicht mehr bestehen. Die Art des Durchgangs durchs Unendliche, ob aus dem positiven ins negative oder umgekehrt, bestimmt die Art der Aenderung der Charakteristik beim Durchschreiten der Stelle $R(\varphi,\psi) = 0$. Diese Aenderung ist daher gleich $[R_1(\varphi,\psi)\delta R(\varphi,\psi)]$. Bezeichnen wir daher für 2 Werthsysteme der Coeffizienten von $\varphi$ und $\psi$ die zugehörigen Charakteristiken mit $\chi^{(2)}$ und $\chi^{(1)}$ so ist :

$$\sum_{R(\varphi,\psi)=0} [R_1(\varphi,\psi)\delta R(\varphi,\psi)] = \chi^{(2)}(\varphi,\psi) - \chi^{(1)}(\varphi,\psi).$$

Angenommen $\varphi$ und
*psi* hätten Coeffizienten, die nur von einen | Variabeln $x$ abhängen. $R$ und $R_1$ sind 160|161
dann ganze Funktionen von $x$. Die Aenderung der Charakteristik von $x = x_1$ bis $x = x_2$ ist dann :

$$\sum_{R(x)=0 \ , \ x_1<x<x_2} [R_1(x)R'(x)].$$

Dies ist die Verallgemeinerung des *Sturm*'schen Satzes. Die Summe $\sum[R_1(x)R'(x)]$ läßt sich wie früher aus den Restfunktionen der zu $R$ und $R_1$ gehörigen Kettenbruchentwicklung bestimmen. Der spezielle *Sturm*'sche Satz löst also gleichzeitig die Verallgemeinerung desselben.

Nehmen wir nun an, die Coeffizienten von $\varphi$ und $\psi$ seien nicht von einer, sondern von $\gamma$ Variabeln abhängig und wir wollen die Aenderung der Charakteristik von einem Werthsysteme zu einem andern kennen lernen. Aber der Uebergang von $u_1'...u_\gamma$ zu $u_1''...u_\gamma''$ kann stets auf einer Linie, die durch die Endstellen geht, d.h. auf einem Gebilde einfacher Mannigfaltigkeit, ausgeführt werden, indem man nämlich $u_1...u_\gamma$ von einem Parameter $t$ abhängig denkt und nun von $t'$ zu $t''$ übergeht.

Wenn die Coeffizienten von $\varphi$ und $\psi$ von 3 Variabeln abhängen, so ist $R = 0$ und $R_1 = 0$ eine Fläche. Die Charakteristik von $\varphi$ und $\psi$ ändert sich dann nur, wenn man die Fläche $R = 0$ überschreitet. Indem Punkte des Raumes entspricht eine ganz bestimmte Charakteristik und alle Punkte des Raumes, die einen gegenseitigen stetigen Uebergang gestatten, ohne daß man die Fläche $R = 0$ passirt, besitzen dieselbe Charakteristik. Die Fläche $R = 0$ muß daher mindestens so viele Gebiete abschließen,

als es Charakteristiken gibt. Setzen wir $\psi = \varphi'$, so ist $R(\varphi, \varphi') = 0$ die Diskriminantenfläche und $\chi(\varphi, \varphi')$ gibt die halbe Anzahl der reellen Wurzeln von $\varphi = 0$ an. Wir erhalten z.B. bei den biquadratischen Gleichungen $\varphi = 0$ die 3 Charakteristiken 0,1,2. (Siehe *pag....*)

§62. Wir wollen nun dazu übergehen, die Funktionen $R$ u. $R_1$ in einer gewissen Determinantenform darzustellen.

Wir haben gesehen, daß $\frac{R_1}{R}$ der Coeffizient der höchsten Potenz der Multiplikatoren $\Phi$ und $\Psi$ ist, für welche die Gleichung $\Phi\psi - \Psi\varphi = 1$ | gilt. Führen wir die $\xi$ ein, so ist $\Psi$ so zu bestimmen, daß $\Phi(\xi)\psi(\xi) = 1$ ist. Die $n$ Coeffizienten von $\Phi$ werden mitteslst der *Euler*'schen Formeln : 161|162

$$\sum_{(\xi)} \frac{\xi^h \Phi(\xi)\Psi(\xi)}{\varphi'(\xi)} = \delta_{h,n-1}.$$

Setzen wir :

$$\sum_{(\xi)} \frac{\xi^h \Psi(\xi)}{\varphi'(\xi)} = s_k,$$

zur Bestimmung der Coeffizienten :

$$\begin{aligned} &c_0 s_0 + ... + c_{n-1} s_{n-1} = 0 \\ &... \\ &c_0 s_{n-1} + ... + c_{n-1} s_{2n-2} = 0 \end{aligned}$$

Hieraus : $\Phi(z) = \frac{|zs_{g+h} + s_{g+h+1}| \ (g,h=0...n-2)}{|s_{i+k}| \ (i,k=0...n-1)}$ woraus folgt :

$$\frac{R_1}{R} = \frac{|s_{g+h}|}{|s_{i+k}|}.$$

Nun hatten wir :

$$|zs_{i+k} - s_{i+k+1}| \underset{(i,k=0...n-1)}{} = \prod_{\alpha} u_\alpha (z - \xi_\alpha) \prod_{\substack{\alpha,\beta \\ (\alpha,\beta=0...n-1)}} (\xi_\alpha - \xi_\beta)^2.$$

Setzen wir daher $u_k = \frac{\psi(\xi_k)}{\varphi'(\xi_k)}$ so folgt : $|s_{i+k}| = \pm \prod \psi(\xi)$.

Abgesehen vom Zeichen ist also $R = \underset{(i,k=0...n-1)}{|s_{i+k}|}$ und $R_1 = \underset{(g,h=0...n-2)}{|s_{g+h}|}$. Wir wollen jetzt noch zeigen, daß wenn man für $R$ und $R_1$ wirklich diese Determinanten setzt :

$$\frac{R_1}{R} = \sum_{\xi} \frac{1}{\varphi'(\xi)\psi(\xi)} \text{ wird. Es ist :}$$

$$|s_{g+h}| = \sum_i \prod_\alpha u_\alpha \prod_{\alpha<\beta} (\xi_\alpha - \xi_\beta)^2 = \prod_\alpha u_\alpha \prod_{\alpha<\beta} (\xi_\alpha - \xi_\beta)^2 \sum_k \frac{1}{u_k \varphi'(\xi_k)^2}.$$

Nun ist $u_k = \frac{\psi(\xi_k)}{\varphi'(\xi_k)}$, also kommt :

$$\frac{|s_{g+h}|}{|s_{i+k}|} = \frac{R_1}{R} = \sum_{(\xi)} \frac{1}{\psi(\xi_k)\varphi'(\xi_k)} \ \textit{q.e.d.}$$

Um den Zusammenhang zwischen der Charakteristik und den *Sturm*'schen Funktio-

nen zu erkennen, gehen wir auf die Gleichung zurück :

$$\sum_{(\xi)} \frac{1}{f_1(\xi)f'(\xi)} \left\{ \sum_{k=1}^{n} y_k F_k(\xi) \right\}^2 = \sum_{y=1}^{n} y_k S_k^2$$

wo $S_k\delta_{ik} = \sum_\xi \frac{F_i(\xi)F_k(\xi)}{f_1(\xi)f'(\xi)}$ und die $S$ die *Sturm*'schen Funktionen waren. Wir haben dann den speziellen Fall betrachtet, wo die $F$ die Restfunktionen der Kettenbruchentwicklung sind und gefunden, daß diese gewissen Determinanten proportional sind. Hieraus folgte dann, daß wir setzen könnten : $\begin{vmatrix} s_0 & \dots & s_m \\ \dots & \dots & \dots \\ s_{m-1} & \dots & s_{2m-1} \\ x^0 & \dots & x^m \end{vmatrix}$ und es wurde dann : $\sum \xi^t . \frac{F_k(\xi)}{f_1(\xi)f'(\xi)} = \delta_{t,m}|s_{i+k}|$.

| Hiernach ist also : 162|163

$$S_{n-m} = \sum_{\xi} \frac{(F_{n-m}(\xi))^2}{f_1(\xi)f'(\xi)} = |s_{g+h}||s_{i+k}| \ (i,k = 0,...m \ g,h = 0,...m-1)$$

wo $s_g = \sum \frac{\xi^g}{f_1(\xi)f'(\xi)}$ ist. Nehmen wir jetzt $f(x) = \varphi(x)$ an und $f_1(x) = \frac{1}{\psi(x)}$ so wird $s_g$ mit dem auf *pag.*... definirten identisch. Es wird daher unsere Charakteristik :

$$\chi(\varphi,\psi) = \frac{1}{2}\sum[\varphi'(\xi)\psi(\xi)] = \frac{1}{2}\sum_{k=1}^{n}[S_k] = \frac{1}{2}\sum[f'(\xi)f_1(\xi)].$$

Wir hatten nun für die Aenderung der Charakteristik gefunden :

$$\chi'' - \chi' = \frac{1}{2}\sum_{k=1}^{n}[S_k''] - \frac{1}{2}\sum_{k=1}^{n}[S_k'] = \sum[R_1(\varphi,\psi\delta R(\varphi,\psi)],$$

die letzte Summe erstreckt über alle Stellen, wo $R = 0$ ist. Also nur wenn $R = 0$ tritt eine Zeichenänderung der *Sturm*'schen Reihe eine. Nun ist :

$$R = \underset{(i,k=0\dots n-1)}{|s_{i+k}|} \ , \ R_1 = \underset{(g,h=0\dots n-2)}{|s_{g+h}|}$$

Definiren wir noch

$$R_2 = \underset{(g,h=0\dots n-3)}{|s_{g+h}|} \ , \ R_3 = \underset{(g,h=0\dots n-4)}{|s_{g+h}|} \quad etc.$$

so bilden die Determinanten $R, R_1, R_2$... eine Reihe *Sturm*'schen $F$, welche die Eigenthümlichkeit hat, daß die Summe ihrer Vorzeichen sich nur ändern kann, wenn $R$ sein Zeichen ändert.

Hier erkennen wir die Analogie mit den *Sturm*'schen Restfunktionen $f, f_1, f_2, ...$ bei denen wir fanden, daß die Anzahl der Zeichenwechsel dieser Reihe sich nur ändert, wenn $f$ sein Zeichen ändert. Bei den Restfunktionen lag dies an den Gleichungen $f_{k-1} - g_k f_k + f_{k+1} = 0$. Man kann daher vermuthen, daß auch zwischen den $R$ analoge Gleichungen bestehen. In der That findet man

$$P_k^2 R_{k-1} - Q_k R_k + P_{k+1}^2 R_{k+1} = 0.$$

§63. Wir können nun die bisherigen Untersuchungen, bei welchen wir die Schnitt-

punkte der Curven $y = \varphi(z)$ und $y = \psi(z)$ mit der Curve $y = 0$ d.h. der Achse betrachteten, dahin verallgemeinern, daß wir drei geschlossene Curven $f = 0$, $\varphi = 0$ und $\psi = 0$ betrachten. Wir untersuchen z. B. die Schnittpunkte von $f = 0$ und $\varphi = 0$ und sehen zu, ob für diese Stellen $\psi$ positiv oder negativ ist, d.h. wir suchen Werthsysteme $x, y$, welche 2 Gleichungen $f = 0$ und $\varphi = 0$ und gleichzeitig einer Ungleichheit
163|164 $\psi < 0$ resp. $> 0$ genügen. Bei | geschlossenen Curven nennt man denjenigen Theil der Ebene der Innere, welcher keine unendlich fernen Punkte enthält. Das Innere ist also durch $f < 0$ charakterisirt. In der Regel durchläuft man nun eine Curve so, daß das Innere links bleibt, doch ist diese Betrachtung auf ebene Gebilde beschränkt. Wir wollen daher ein rein analytisches Criterium aufstellen. Bilden wir die Determinante :

$$\Delta = \begin{vmatrix} f & \varphi & \psi \\ f_1 & \varphi_1 & \psi_1 \\ f_2 & \varphi_2 & \psi_2 \end{vmatrix}$$

Gehen wir nun auf der Curve $f = 0$, so wollen wir folgende Bestimmung des Fortgangs treffen : In einem Schnittpunkte von $f$ und $\varphi$ wollen wir auf $f$ in einer solchen Richtung vorwärts gehen, daß, wenn $\Theta$ eine beliebige durch den fraglichen Punkt gelegte Curve ist, der Ausdruck :

$$\frac{\partial \begin{vmatrix} f & \Theta & \psi \\ f_1 & \Theta_1 & \psi_1 \\ f_2 & \Theta_2 & \psi_2 \end{vmatrix}}{\partial \psi} d\Theta > 0 \text{ ist.}$$

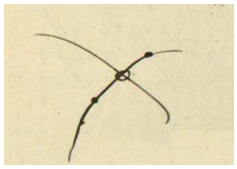
Es soll also mit andern Worten :

$$(f_1\Theta_2 - f_2\Theta_1)d\Theta > 0 \text{ sein.}$$

Zunächst ist zu zeigen, daß dieses Fortgangsprinzip unabhängig von der willkürlichen Curven $\Theta$ ist.

Nun ist :

$$d\Theta = \Theta_1 dx + \Theta_2 dy$$
$$0 = f_1 dx + f_2 dy$$



Es soll also durch unsere Bestimmung festgesetzt werden,
in welcher von beiden Richtungen von $\frac{dy}{dx}$ wir auf $f = 0$ vorwärts schreiten sollen. Nun ist $d_\Theta = (f_1\Theta_2 - f_2\Theta_1)\frac{dy}{f_1}$.

Unsere Bestimmung reduzirt sich also darauf, daß $\frac{dy}{f_1}$ oder $f_1 dy$ positiv sein soll.

Damit ist also $\Theta$ weggeschafft. Man erhält das letzte Criterum natürlich auch, indem man $\Theta = y$ setzt. Die beiden Figuren :




164|165 | zeigen, daß bei dem so definirten Fortgange das Innere immer links bleibt.

Wir betrachten nun 2 Curven $U = 0$ und $V = 0$ und gehen auf $U$ in dem durch unser Criterium Definirten Sinne. Es soll also

$$(U_1V_2 - U_2V_1)dV > 0 \text{ sein.}$$

Dann ist Klar, daß bei dieser Art des Fortgangs auf $U = 0$ die Summe : $\sum_{U=0=V}[dV] = 0$ sein muß, wenn $V = 0$ eine geschlossene Curven ist, weil man ja ebenso oft in das Innere von $V = 0$ hinein als aus demselben heraustreten muß, wenn man die ganze Curve $U = 0$ durchläuft. Daraus folgt jetzt :

$$\sum_{U=0=V} [U_1V_2 - U_2V_1] = 0$$

und in dieser Gleichung ist von einem Fortgangsprinzip gar nicht mehr die Rede.

Wir setzen dabei aber im Folgenden zunächst voraus, daß die Curven $U = 0$ und $V = 0$ so zu einander liegen, daß die Funktionaldeterminante nicht verschwindet.

Nun setzen wir $f = U$, $\varphi\psi = V$, so ist :

$$U_1V_2 - U_2V_1 = f_1(\varphi_2\psi + \varphi\psi_2) - f_2(\varphi_1\psi + \varphi\psi_1)$$

$$= \psi(f_1\varphi_2 - f_2\varphi_1) + \varphi(f_1\psi_2 - f_2\psi_1) = \psi\Delta_3 - \varphi\Delta_2,$$

wenn man mit $\Delta_2$ und $\Delta_3$ die Unterdeterminanten von $\varphi$ und $\psi$ in $\Delta$ bezeichnet.

Es ist also :

$$\sum_{f=0,\varphi\psi=0} [\psi\Delta_3 - \varphi\Delta_2] = 0. \text{ Oder :}$$

$$\sum_{f=0,\varphi=0} [\psi\Delta_3] - \sum_{f=0,\psi=0} [\varphi\Delta_2] = 0.$$

Auf diese Weise erhalten wir, da $\Delta = f\Delta_1 + \varphi\Delta_2 + \psi\Delta_3$ ist :

$$\sum_{f=0=\varphi} [\Delta] = \sum_{\varphi=0=\psi} [\Delta] = \sum_{\psi=0=f} [\Delta].$$

Dies ist die Verallgemeinerung des früheren Resultates :

$$\sum[\Delta(\xi)] = \sum[\Delta(\eta)].$$

Wir zeigen nun, daß $\sum[\Delta]$ immer eine grade Zahl ist. Für $U = f$ | und $V = \varphi$ folgt : 165|166

$$\sum_{f=0=\varphi} [\Delta_3] = 0 = \sum_{\substack{f=0=\varphi \\ \psi>0}} [\Delta_3] + \sum_{\substack{f=0=\varphi \\ \psi<0}} [\Delta_3].$$

Ferner ist:

$$\sum_{f=0=\varphi} [\Delta] = \sum_{f=0=\varphi} [\Delta_3\psi] = \sum_{\substack{f=0=\varphi \\ \psi>0}} [\Delta_3] - \sum_{\substack{f=0=\varphi \\ \psi<0}} [\Delta_3].$$

Es ist daher :

$$\sum_{f=0=\varphi} [\Delta] = \sum_{f=0=\varphi} [\psi\Delta_3] = 2 \sum_{\substack{f=0=\varphi \\ \psi>0}} [\Delta_3] = -2 \sum_{\substack{f=0=\varphi \\ \psi<0}} [\Delta_3].$$

Wir setzen jetzt :

$$\chi(f, \varphi, \psi) = -\frac{1}{2} \sum_{f=0=\varphi} [\Delta]$$

und bezeichnen $\chi$, welches also eine ganz Zahl ist, als die Charakteristik von $f, \varphi, \psi$. Aus $d\varphi = (\varphi_2 f_1 - \varphi_1 f_2)\frac{dy}{f_1} = \frac{\Delta_3 dy}{f_1}$ folgt, da beim Fortgang auf $f$ $\frac{dy}{f_1}$ positiv bleibt :

$$\chi = -\frac{1}{2} \sum_{f=0=\varphi} [\Delta] = -\frac{1}{2} \sum_{f=0=\varphi} [\psi\Delta_3] = -\frac{1}{2} \sum_{f=0=\varphi} [\psi d\varphi]$$

also :

$$\chi = -\frac{1}{2}\sum_{f=0=\varphi}[d\varphi\psi].$$

Wenn wir also auf der Curven $f = 0$ in dem gegebenen Sinne gehen, und die Schnittpunkte mit $\varphi = 0$ betrachten, so ist eine solche Schnittstelle in Bezug auf $\varphi\psi$ als Eintritts- oder Austrittsstelle zu bezeichnen, je nachdem $d\varphi\psi$ negativ oder positiv ist. Alsdann sagt die letzte Formel, daß die Charakteristik des Systems $f, \varphi, \psi$ gleich ist dem halben Ueberschuß der Eintritte über die Austritte. Da wir 3 Curven haben, so können wir diese Betrachtung auf 6 verschiedene Arten anstellen, denn auf einer der Curven fortgehend, können wir die Schnittpunkte derselben mit der einen oder der andern der beiden übrigen Curven betrachten. Die Invarianz der Charakteristik sagt nun aus, daß der Ueberschuß der Eintritte über die Austritte in allen 6 fällen derselbe ist.

Die Charakteristik $\chi(f, \varphi, \psi)$ läßt sich auch schreiben :

$$\chi(f,\varphi,\psi) = \sum_{\substack{\varphi=0=\psi\\ f<0}}[\Delta_1] = \sum_{\substack{\varphi=0=\psi\\ f<0}}[\varphi_1\psi_2 - \varphi_2\psi_1]$$

Wenden wir dies auf das System $f\Delta_1, \varphi, \psi$ an, so kommt :

166|167 |

$$\chi(f\Delta_1,\varphi,\psi) = \sum_{f\Delta_1<0}[\Delta_1] = \sum_{\substack{\varphi=0=\psi\\ f>0}}(-1) + \sum_{\substack{\varphi=0=\psi\\ f<0}}(+1)$$

d.h. die Charakteristik dieses Systems gibt den Ueberschuß der Schnittpunkte von $\varphi$ und $\psi$ im Innern von $f = 0$ über diejenigen außerhalb $f = 0$.

§64. Wenden wir diese Untersuchungen nun an auf die speziellen Funktionen $\varphi$ und $\psi$, welche entstehen, wenn man die Funktion $\Theta(x + yi)$ in ihren reellen und imaginären Theil zerlegt. Wir setzen :

$$2\Theta(x + yi) = \varphi(x, y) + i\psi(xy)$$

Hieraus folgt :

$$\varphi = \Theta(x + yi) + \Theta(x - yi)\ ,\ \psi(xy) = i\Theta(x - yi) - i\Theta(x + yi)$$
$$\varphi_1 = \Theta'(x + yi) + \Theta(x - yi)\ ,\ \psi_1 = i\Theta'(x - yi) - i\Theta'(x + yi)$$
$$\varphi_2 = -\psi_1\ ,\ \psi_2 = \varphi_1$$

$$\Delta_1 = \varphi_1\psi_2 - \varphi_2\psi_1 = \varphi_1^2 + \psi_1^2.$$

Also :

$$\chi = \sum_{\substack{\varphi=0=\psi\\ f<0}}[\varphi_1\psi_2 - \varphi_2\psi_1] = \sum_{\substack{\varphi=0=\psi\\ f<0}}(1).$$

Hierüber ist aber folgendes zu bemerken. $\varphi$ und $\psi$ sind nicht, wie wir es voraussetzen, geschlossene Curven. Alle Folgerungen aber, die sich auf das Innere von $f = 0$ beziehen, bleiben bestehen, wenn wir sie durch Curvenstücke schließen, die ganz außerhalb $f = 0$ verlaufen.

Dann gibt also die Charakteristik die Anzahl der Schnittpunkte von $\varphi = 0$ und $\psi = 0$, d.h. der Nullwerthe von $\Theta(x + yi) = 0$ an, welche in Innere von $f = 0$ liegen.

Bewegen wir uns jetzt auf der Curve $\varphi = 0$, so ist (*pag....*)

$$\chi = -\frac{1}{2} \sum_{\varphi=0=\psi} [df\psi] \text{ ; nun ist :}$$

$$df\psi = \frac{\partial f\psi}{\partial x}dx + \frac{\partial f\psi}{\partial y}dy = f(\psi_1 dx + \psi_2 dy) + \psi(...).$$

Da wir auf $\varphi = 0$ vorwärts gehen, so ist :

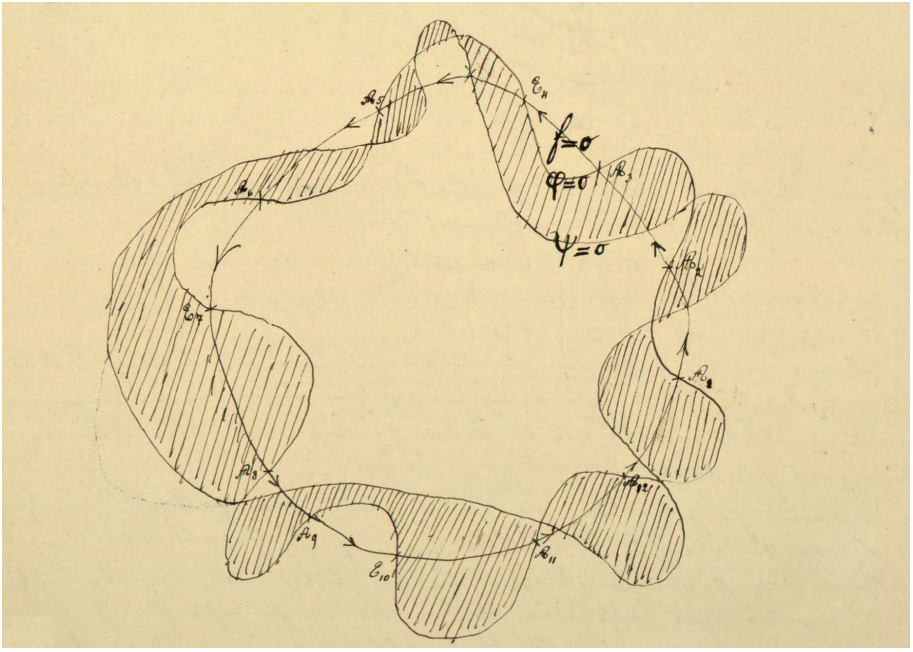
$$\varphi_1 dx + \varphi_2 dy = 0 \text{ , folglich :}$$

$$df\psi = f(\varphi_1\psi_2 - \varphi_2\psi_1)\frac{dy}{\varphi_1} + \psi(...).$$

| Nun zählen wir die Schnittpunkte von $\varphi$ mit $\psi$, der zweite mit $\psi$ multiplizirte Theil fällt also weg, ferner ist $\frac{dy}{\varphi_1}$ gemäß unserm Fortgangsprinzip positiv, ebenso $\varphi_1\psi_2 - \varphi_2\psi_1 = \varphi_1^2 + \psi_1^2$. 167|168

Also ist : $\chi = -\frac{1}{2}\sum_{\varphi=0=\psi}[f]$. Daraus folgt, daß im Innern von $f$, also für $f < 0$, nur Eintrittspunkte liegen. Das Produkt $f\psi$ entscheidet dabei über das Innere und Aeußere.

§65. Wir wollen nun durch eine Abbildung zu einer neuen Auffassung der Charakteristik gelangen. Zur Erleichterung der Anschauung wählen wir gleichzeitig ein spezielles Beispiel, welches durch die folgende Figur dargestellt werde :



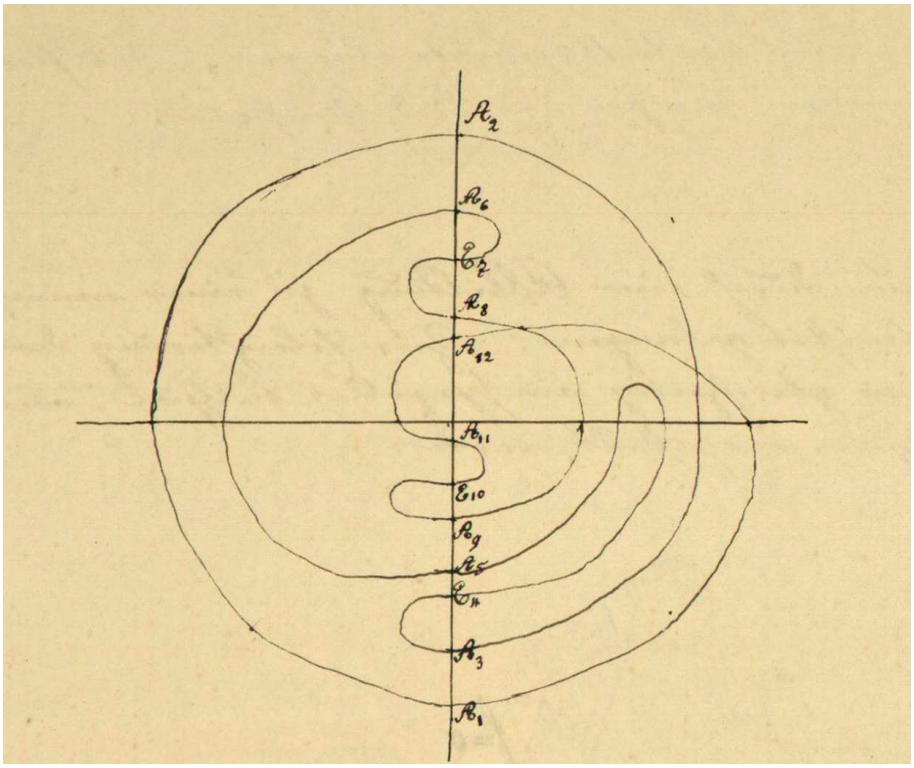

Bilden wir nun die ganze Figur dadurch ab, daß wir setzen :

$$\xi = \varphi(x, y) \text{ , } \eta = \psi(x, y).$$

Dann geht $f(x, y) = 0$ über in eine Curve $F(\xi, \eta) = 0$.

| Die Curven $F(\xi, \eta)$ bietet dann ungefähr die nachstehende Gestalt dar : 168|169

Da bei unserer Abbildung sich die im Innern und Aeußern gelegenen Theile folgendermaßen abbilden :



| | |
|---|---|
| $\varphi\psi < 0$ | $\varphi\psi > 0$ |
| $\varphi\psi > 0$ | $\varphi\psi < 0$ |

so erkennt man, daß ein Eintritt sich in der Abbildung dadurch markirt, daß wir die $y-$Achse im entgegengesetzten Sinne eines Uhrzeigers scheiden und ein Austritt dadurch, daß wir sie im Sinne eines Uhrzeigers schneiden.

Daraus folgt nun, da ja $\chi$ gleich dem halben Ueberschusse der Eintritte über die Austritte ist, daß die Charakteristik grade die richtige Anzahl von Windungen um den Nullpunkt angibt. Die Curve $f = 0$ repräsentirt also einen Faden, der in sich geschlossen ist und einen im Nullpunkt gedachten Stift so oft mal umwindet als die Charakteristik angibt. Wir denken den Faden in Polarkoordinaten ausgedrückt durch $\xi = r\cos v$ , $\eta = r\sin v$, dann erstreckt sich $v$ von 0 bis zu einem Vielfachen $2\pi$ und
169|170 dieses Vielfache ist die | Charakteristik. Daher ist $\chi(f, \varphi, \psi) = \frac{1}{2\pi}\int dv$, das Integral über die ganz Curven erstreckt. Nun ist :

$$v = arctg\,\frac{\eta}{\xi} = arctg\,\frac{\psi(x,y)}{\varphi(x,y)},\ \text{also :}$$

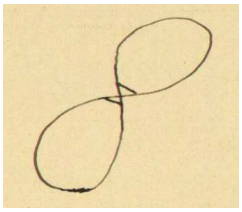
$\chi = \int d\ arctg\,\frac{\psi(x,y)}{\varphi(x,y)}$, das Integral erstreckt über alle $x, y$, wofür $f = 0$ ist. $f(x, y) = 0$ ist nun eine geschlossen Curven, welche wir ebenfalls in Polarkoordinaten ausdrücken wollen : $x = \rho\cos w$ , $y = \rho\sin w$. Wir erhalten dann $f(\rho\cos w, \rho\sin w) = 0$ und hieraus : $\rho = \Theta(w)$.

Nun können wir annehmen, daß die Curven $f(xy) = 0$ sich nicht selbst durchdringe, denn sonst hätten wir zwei geschlossene Curven :

Ferner können wir annehmen, daß $\rho$ eine eindeutige Funktion von $w$ ist. Denn wir können immer die vorkommenden Gebiete in solche Bezirke theilen, in welchen dies stattfindet. Existiren mehrere solcher Bezirke, so nehmen wir sie nach einander vor.

Wir können also annehmen wir hätten es jetzt mit <u>einem</u> Bezirke zu thun von der Beschaffenheit, daß die Darstellung :

$$x = \rho\cos w\ ,\ y = \rho\sin w\ ,\ \rho = \Theta(w)$$



eine eindeutige Funktion $\rho$ von $w$ ergibt. Dann gehen gleichzeitig $\varphi$ und $\psi$ in Funktionen von $w$ über, nämlich :

$$\varphi(x,y) = \Phi(w) \ , \ \psi(x,y) = \Psi(w)$$

und wir erhalten :

$$\chi = \frac{1}{2\pi}\int \frac{\varphi d\psi - \psi d\varphi}{\varphi^2 + \psi^2} = \frac{1}{2\pi}\int_0^{2\pi} \frac{\Phi\Psi' - \Psi\Phi'}{\Phi^2 + \Psi^2} dw$$

# Anhang.

# Ergänzungen und Verbesserungen

| Es sei $\frac{F(x)}{\Psi(x)}$ eine Lösung des Cauchy'schen Problems, $F(x)$ vom Grade $\mu$ und $\Psi$ 176|177
vom Grade $\nu$ und $\mu+\nu<n$. Man bilde $f(x)$ und $f_1(x)$ von den Graden $n$ und $n-1$. Dann lassen sich zwei Multiplikatoren finden $\Phi(x)$ und $\Psi(x)$, so daß :

$$\Psi(x)f_1(x)-\Phi(x)f(x)=F(x) \text{ ist,}$$

d.h. daß die Differenz von Grade $\mu$ wird. $\Psi(x)$ sei also vom Grade $\nu$, dann muß $\Phi(x)$ von Grade $\nu-1$ sein. Die Anzahl der zu bestimmenden Unbekannten, D.h. der Coeffizienten von $\Psi$ und $\Phi$ sind $2\nu+1$. Setzen wir nun $\mu+\nu=n-\rho$, $\rho\geq 1$, so folgt $\mu=n-\nu-\rho$. Damit die genannte Differenz auf den $\mu$ten Grad herabsinke, sind also $n+\nu-1-\mu=2\nu+\rho-1$ Gleichungen zu erfüllen. Für $\rho=1$, D.h; für $\mu+\nu=n-1$, hat man also $2\nu$ homogene Gleichungen, welche gerade zur Bestimmung der $2\nu$ Unbekannten ausreichen. Die Lösung dieser eigentlich der Theorie des linearen Gleichungen angehörigen Aufgabe geschieht nun am einfachsten durch die Kettenbruchentwicklung, indem man nämlich nachweisen kann, daß $F,\Phi$ und $\Psi$ Vielfache von $f_k,\varphi_k$ und $\psi_k$ sind.

$\mu$ befindet sich jedenfalls in einem der Intervalle :

$$n...n_1, n_1....n_2, ....n_2...n_3, ...n_k-1...n_k, ..., n_{\nu-1}...n_\nu$$

dann ist $k$ eindeutig bestimmt, wenn wir setzen :

$$n_{k-1}>\mu\geq n_k$$

dann muß $\nu<n-n_k$ sein. Jetzt ist :

$$\psi_k f_1-\varphi_k=f_k$$
$$\Psi f_1-\Phi=F. \text{ Hierausfolgt :}$$

$$f_k\Psi-\psi_k F=f\{\psi_k\Phi-\varphi_k\Psi\}$$

Zu Folge der oberen Grenzen von $\mu$ und $\nu$ hat die linke Seite dieser Gleichung einen kleineren Grade als $n$, sie verschwindet für alle $n$ Wurzeln von $f$, es muß also identisch die Gleichung bestehen :

$$f_k(x)\Psi(x)-\psi_k(x)F(x)=0.$$

Um die Untersuchung allgemein zu führen, sei $\chi(x)$ der größte gemeinschaftliche Theiler von $f_k(x)$ und $\psi_k(x)$, dann ist:

$$\Psi(x)=\frac{\psi_k(x)}{\chi(x)}\Theta(x)\ ,\ F(x)=\frac{f_k(x)}{\chi(x)}\Theta(x).$$

Nun ist :

$$\Phi(x)f(x)=\Psi(x)f_1(x)-F(x)=\frac{\Theta(x)}{\chi(x)}(\psi_k(x)f_1(x)-f_k(x))=\frac{\varphi_k(x)}{\chi(x)}f(x)\Theta(x)$$

folglich

$$\Phi(x)=\frac{\varphi_k(x)}{\chi(x)}\Theta(x).$$

Da nun $\varphi_k$ und $\psi_k$ keinen gemeinschaftlichen Theiler haben, so kann $\chi$ nicht in $\varphi_k$ enthalten sein, muß also ein Theiler von $\Theta$ sein, da $\Phi(x)$ eine ganze Funktion ist. Wir können also $\chi(x)$ in $\Theta(x)$ dividiert deken und erhalten, indem wir den Quotienten, der eine ganze Funktion ist, wieder mit $\Theta(x)$ bezeichnen, die Gleichungen :

$$F(x)=f_k(x)\Theta(x)\ ,\ \Phi(x)=\varphi_k(x)\Theta(x)\ ,\ \Psi(x)=\psi_k(x)\Theta(x).$$

Alle $F(x)$ und $\Psi(x)$ also, deren Quotient eine Losung des Cauchy'schen Problems ist, mussen ganze Vielfache von $f_k(x)$ und $\psi_k(x)$ sein. Diese Vorstellung gibt denn auch die unteren Grenzen für $\mu$ und $\nu$ an, es ist $\mu \geq n_k$ und $\nu \geq n - n_{k-1}$.

Ist $\mu = n_k$ so ist $\Theta(x)$ vom Grade 0, ebeso wenn $\nu = n - n_{k-1}$ ist. Allgemeine ist für den Grad $\lambda$ von $\Theta(x)$ zu setzen :

$$\mu = \lambda + n_k \ , \ \nu = \lambda + n - n_{k-1}.$$

Die Grenzwerthe $\mu = n_k$ und $\nu = n - n_{k-1}$ sind also simultane. Da $\mu + \nu < n$ ist, so folgt :

$$\begin{aligned} \lambda + n_k + \lambda - n_{k-1} &< n \\ 2\lambda &< n_{k-1} - n_k \\ \lambda &< \frac{1}{2}(n_{k-1} - n_k). \end{aligned}$$

Dieses ist die Grenze, unter welscher $\lambda$ stets bleiben muß.

Wir wissen also jetzt, daß jedes $F$ und $\Psi$, welsches die Cauchy'sche Aufgabe löst, in der Form $F = f_k\Theta$ , $\Psi = \psi_k\Theta$ enthalten sein muß. Dabei ist der Grad $\lambda$ von $\Theta$ kleiner als $\frac{1}{2}(n_{k-1} - n_k)$, wenn $\mu$ und $\nu$ in den Intervallen

$$n_{k-1} > \mu \geq n_k \ , \ n - n_k > \nu \geq n - n_{k-1} \text{ liegen.}$$

Aber nicht umgekehrt sind alle die so erhaltenen $F$ und $\Psi$ Lösungen der Cauchy'schen Aufgabe, denn es tritt nach die Bedingung hinzu, daß $\frac{F(x)}{\Psi(x)} = \frac{f_k(x)}{\psi_k(x)}$ mit $f_1$ für $x = \xi_h$ übereinstimmen. Die Lösung ist also nur für gewisse $\mu$ und $\nu$ möglich, denn es sind alle $\frac{f_k}{\psi_k}$ auszuscheiden, bei denen $f_k$ und $\psi_k$ gemeinschaftliche Faktoren haben.

## A.4 Traduction de la seconde partie du cours

# Théorie

# des

# équations algébriques

# du

# Prof. Kronecker

---

hiver 1880/81

84|85

# | Chapitre deux

*Sur la divisibilité des fonctions entières ; Déduction de la proposition de Sturm ; Conclusions tirées de celles-ci ; théorie des caractéristiques et preuve de l'existence des racines.*

---

§35 : On peut appliquer les études des § précédents pour trouver les facteurs communs des deux fonctions entières (décomposables) :

$$F(x) = (x - x_1)...(x - x_n) = x^n - f_1 x^{n-1} + f_2 x^{n-2} - ... \pm f_n$$
$$\Phi(x) = (x - \xi_1)...(x - \xi_\nu) = x^\nu - \varphi_1 x^{\nu-1} + ... \pm \varphi_\nu$$

Nous supposons aussi que $m$ des $\xi$ sont égaux aux $x_1, ..., x_m$ et nous voulons représenter le facteur commun de $F$ et $\Phi$. Soit ensuite $\mathcal{G}(x_1, ..., x_m)$ une fonction entière symétrique de $x_1, ..., x_m$, de sorte que chaque terme de la somme :

$$\sum_{(i)} \mathcal{G}(x_{i_1}...x_{i_m})\Phi(x_{i_{m+1}})...\Phi(x_{i_n}) \quad [17]$$

soit égale à 0, exceptées ceux pour lesquelles $i_{m+1} = m + 1, ..., i_n = n$. On a donc :

$$\sum_{(i)} \mathcal{G}(x_{i_1}...x_{i_m})\Phi(x_{i_{m+1}})...\Phi(x_{i_n}) = \mathcal{G}(x_1...x_m)\Phi(x_{m+1})...\Phi(x_n)$$

On pose maintenant $\mathcal{G}$ égale au facteur commun $(x - x_1)...(x - x_m)$, et il suit donc :

$$\sum_{(i)} \prod_{h=1}^{m}(x - x_{i_h}) \prod_{k=m+1}^{n} \Phi(x_{i_k}) = \prod_{h=1}^{m}(x - x_h) \prod_{k=m+1}^{n} \Phi(x_k)$$

Pour représenter le facteur $\prod_{k=m+1}^{n} \Phi(x_k)$, nous avons seulement besoin de poser dans l'ancienne équation plus générale $\mathcal{G} = 1$, et on a ensuite :

$$\sum_{(i)} \prod_{k=m+1}^{n} \Phi(x_{i_k}) = \prod_{k=m+1}^{n} \Phi(x_k)$$

On a donc :

$$\frac{\sum_{(i)} \prod_{h=1}^{h=m}(x - x_{i_h}) \prod_{k=m+1}^{n} \Phi(x_{i_k})}{\sum_{(i)} \prod_{k=m+1}^{n}(x - x_{i_h})} = \prod_{h=1}^{m}(x - x_h)$$

85|86 C'est-à-dire que nous avons | représenté le facteur commun des deux fonctions entières $F$ et $\Phi$ comme un quotient de deux fonctions symétriques des $n$ valeurs $x$.

Si $F$ et $\Phi$ ont comme facteur commun $(x - x_1)...(x - x_m)$, alors on a :

$$\prod_{k=1}^{n} \Phi(x_k) = 0, \prod_{k=2}^{n} \Phi(x_k), ..., \prod_{k=m}^{n} \Phi(x_k)$$

17. Sur le genre de sommation voir ...

On a ensuite aussi :

$$\sum_{(i)} \prod_{k=2}^{n} \Phi(x_{i_k}) = 0, ..., \sum_{(i)} \prod_{k=m}^{n} \Phi(x_{i_k}) = 0$$

Pour qu'il existe donc un facteur commun du $m^{\text{ième}}$ degré, il y a $m$ conditions à remplir. Si nous supposons que $\sum_{(i)} \prod_{k=m+1}^{n} \Phi(x_{i_k}) \neq 0$ alors il n'existera sûrement aucun facteur commun de degré plus grand que $m$. Dans ce dernier cas, notre représentation d'un facteur commun du $m^{\text{ième}}$ degré est caduque. Afin que $(x-x_1)...(x-x_m)$ soit donc le plus grand facteur commun de $F$ et $\Phi$, $m$ équations et une inéquation sont à satisfaire.

Nous avons représenté le plus grand facteur commun comme quotient de deux fonctions symétriques. Il en résulte que le numérateur et le dénominateur de cette représentation sont des fonctions entières de $f_1...f_n$ et $\varphi_1, ..., \varphi_\nu$.

Pour $m = 1$, cette représentation a été donnée en premier par *Abel*.

§36 : Mais maintenant le cas où il n'existe absolument aucun facteur commun est d'un intérêt notable. La première de nos équations, c'est-à-dire $\prod_{k=1}^{n} \Phi(x_k)$, est alors différente de 0. Nous pouvons par conséquent représenter $\dfrac{1}{\Phi(x_k)}$ comme une fonction entière de $x_k$. C'est en effet :

$$\frac{1}{\Phi(x_k)} = \frac{\Phi(x_1)....\Phi(x_{k-1})\Phi(x_{k+1})...\Phi(x_n)}{\prod_{k=1}^{n} \Phi(x_k)} = \frac{\Psi(x_k)}{R(f_1...f_n; \varphi_1...\varphi_\nu)}$$

On a donc :

$$\Phi(x_k)\Psi(x_k) = \prod_{k=1}^{n} \Phi(x_k) = R(f_1...f_n; \varphi_1...\varphi_\nu) = R(F, \Psi)$$

où $R$ est une fonction entière des $f$ et des $\varphi$. Ainsi pour $x = x_1, ...x_n$, $\Phi(x)\Psi(x) - R(F, \Psi)$ disparaît, et il est donc divisible par $F(x)$ et nous obtenons donc l'équation :

$$\Phi(x)\Psi(x) + F(x)\mathcal{G}(x) = R(F, \Psi)$$

| $\Phi, \Psi, F$ sont de degré resp. $\nu, n-1, n$ et il suit que $\mathcal{G}(x)$ doit être de degré $\nu - 1$. 86|87

Nous obtenons donc le résultat suivant :

Si $F(x)$ et $\Phi(x)$ n'ont aucun facteur commun, alors on trouve toujours deux multiplicateurs $\Psi$ et $\mathcal{G}$ de degré resp. $n-1$ et $\nu - 1$ tels que l'expression $\Phi(x)\Psi(x) + F(x)\mathcal{G}(x) = R(F, \Phi)$ soit indépendante de $x$. Lorsqu'un facteur commun existe, alors $R = 0$ et lorsque $R \neq 0$, alors aucun facteur commun ne peut exister ; à savoir que cette proposition est vraie pour n'importe quelles $F$ et $\Phi$, dont la possibilité d'être décomposées en facteurs linéaires n'a pas besoin d'être prouvée. Nous avons certes dérivé l'équation $\Phi\Psi + F\mathcal{G} = R$ sous l'hypothèse de l'existence de racines pour $F$ et $\Phi$, mais cela fonctionne, car identique pour tout $f$ et $\varphi$.

En revanche, il n'est pas encore démontré que lorsque $R = 0$, un facteur existe, ou que à l'inverse $R \neq 0$ lorsque aucun facteur commun n'existe. Cela est en effet seulement démontré pour $R(F, \Phi) = \prod_{k=1}^{n} \Phi(x_k)$ c'est-à-dire sous l'hypothèse de l'existence des racines $x_1...x_n$ de $F$.

Nous posons $\Phi(\xi) = \prod_{h=1}^{\nu}(\xi - \xi_h)$, il suit donc :

$$R(F, \Phi) = \prod_{k=1}^{n}\prod_{h=1}^{\nu}(x_k - \xi_h) = (-1)^{n.\nu} \prod_{k=1}^{n}\prod_{h=1}^{\nu}(\xi_h - x_k) = (-1)^{n.\nu} \prod_{h=1}^{\nu} F(\xi_h)$$

Il s'ensuit :

$$\prod_{k=1}^{n} \Phi(x_k) = R = (-1)^{n.\nu} \prod_{h=1}^{\nu} F(\xi_h)$$

Nous pouvons donc exprimer ceci :

$$R(F, \Phi) = (-1)^{n.\nu} R(\Phi, F)$$

On a :

$$\Phi(x_k)\Psi(x_k) = R(F, \Phi) = F(\xi_h)\mathcal{G}(\xi_h)$$

De là on a : $\Psi(x) = \dfrac{R}{\Phi(x)}$ pour $x = x_1...x_n$ et de même :

$$\mathcal{G}(\xi) = \frac{\mathcal{R}}{\mathcal{F}(\xi)} \text{ pour } \xi = \xi_1, ..., \xi_\nu$$

Nous pouvons de là encore établir la formule de *Lagrange* et obtenir [18] :

$$\Psi(x) = RF(x) \sum_{k=1}^{n} \frac{1}{(x - x_k)F(x_k)\Phi(x_k)}$$
$$\mathcal{G}(x) = R\Phi(x) \sum_{h=1}^{\nu} \frac{1}{(x - \xi_h)\Phi'(\xi_h)F(\xi_h)}$$

87|88 | $\Psi(x)$ est en effet de degré $n-1$ et $\mathcal{G}(x)$ de degré $\nu - 1$. Nous pouvons maintenant facilement prouver directement l'égalité $\Phi(x)\Psi(x) + F(x)\mathcal{G}(x) = R$, car on a :

$$\Phi(x)\Psi(x) + F(x)\mathcal{G}(x) = RF(x)\Phi(x) \left\{ \sum_{k=1}^{n} + \sum_{h=1}^{\nu} \right\}$$

Pour montrer maintenant que le facteur de $R$ est égal à 1, nous posons $z = x_1...x_n, \xi_1...\xi_\nu$ les racines de $F\Phi = 0$ et nous construisons d'après *Lagrange* la valeur 1 :

$$1 = F\Phi \sum_{r=1}^{n+\nu} \frac{1}{z - z_r} \frac{1}{F(z_R)\Phi'(z_n) + F'(z_r)\Phi(z_r)} =$$
$$1 = F(z)\Phi(z) \left\{ \sum_{k=1}^{n} \frac{1}{(z - x_k)F'(x_k)\Phi(x_k)} + \sum_{h=1}^{\nu} \frac{1}{(z - \xi_h)\Phi'(\xi_h)F(\xi_h)} \right\} \text{ q.e.d.}$$

On appelle $R(F, \Phi)$ le résultant d'élimination en $x$ des deux équations $F = 0$ et $\Phi = 0$.

Soit $F(y)$ une fonction entière de degré $n - \nu$. Nous pouvons ensuite d'après le §34 représenter le produit $\prod_{h=1}^{\nu} F(y_k)$, lequel est une fonction symétrique de $y$, à l'aide d'une interpolation, dans laquelle $\prod F(y_h)$ est un polynôme symétrique sous forme réduite et on établit :

$$\prod_{h=1}^{\nu} F(y_h) = \sum_{(i)} \prod_{h=1}^{\nu} F(x_{i_h}) \prod_{h=1}^{\nu} \prod_{k=\nu+1}^{n} \frac{y_h - x_{i_k}}{x_{i_h} - x_{i_k}}$$

18. *N.d.T.* Erreur manifeste :

$$\Psi(x) = RF(x) \sum_{k=1}^{n} \frac{1}{(x - x_k)F'(x_k)\Phi(x_k)}$$

Si l'on pose $y_h = \xi_h$, alors on a :

$$\prod_{h=1}^{\nu}\prod_{k=\nu+1}^{n}(\xi_h - x_{i_k}) = \prod_{k=\nu+1}^{n}(-1)^{\nu}\Phi(x_{i_k} = (-1)^{\nu.(n-\nu)}\prod_{k=\nu+1}^{n}\Phi(x_{i_k})$$

Et donc :

$$\sum_{(i)}\frac{\prod_{h=1}^{\nu}F(x_{i_h})\prod_{k=\nu+1}^{n}\Phi(x_{i_k})}{\prod_{h=1}^{\nu}\prod_{k=\nu+1}^{n}(x_{i_h}-x_{i_k})} = (-1)^{\nu.(n-\nu)}\prod_{h=1}^{\nu}F(\xi_h) = R(F,\Phi)$$

Ainsi donc le résultant de $F$, lequel est de degré $n-\nu$, et de $\Phi$, qui est de degré $\nu$, est représenté par interpolation par les valeurs de $F$ et $\Phi$ pour les $n-\nu+\nu = n$ valeurs $x_1...x_n$. Cette représentation élégante vient de *Rosenhain (Crelle Bd. 30).*

| §37 : Nous avons obtenu 88|89

$$R(F,\Phi) = \prod\Phi(x_k) = \prod_{h=1}^{\nu}\prod_{k=\nu+1}^{n}(x_k - \xi_h) = (-1)^{n.\nu}\prod_{h=1}^{\nu}F(\xi_h)$$

Si nous considérons maintenant en même temps les deux équations $F(x) = 0$ et $\Phi(x) = 0$ et que nous supposons que les deux se réalisent ensemble, c'est-à-dire qu'elles sont satisfaites en même temps par la même valeur de $x$, alors $R$ doit être égal à 0. D'un autre côté, lorsque $R = 0$, alors $F(x) = 0$ et $\Phi(x) = 0$ sont satisfaites par une seule et même valeur de $x$. On appelle pour cela $R = 0$ l'équation résultante de $F(x) = 0$ et $\Phi(x) = 0$. Elle est le résultat complet de la réalisation simultanée de $F(x) = 0$ et $\Phi(x) = 0$, c'est-à-dire la condition nécessaire et suffisante à cela. On a reporté cette notion des équations aux fonctions et on appelle $R$ le résultant de $F$ et $\Phi$.

Si l'on prend le cas particulier de $\Phi(x) = F'(x)$, alors on a :

$$R = \prod_{k=1}^{n}\Phi(x_k) = \prod_{k=1}^{n}F'(x_k) = \prod_{k,h}(x_k - x_h) = (-1)^{\frac{n(n-1)}{2}}\mathcal{P}^2.$$

Ce résultant particulier de $F(x)$ et $F'(x)$ est nommé d'après *Sylvester* le discriminant de $F(x)$. C'est une fonction symétrique de $x_1...x_n$ et par conséquent une fonction entière des fonctions symétriques élémentaires $f_1...f_n$. Leur disparition, ou ce qui est pareil, l'existence simultanée de $F(x) = 0$ et $F'(x) = 0$ exprime le fait que $F(x) = 0$ a une racine double.

§38 : Nous avons démontré indépendamment de la décomposition de $F$ et $\Phi$ en facteurs linéaires que lorqu'un facteur commun existe, on doit avoir $R = 0$, d'où il suit que pour $R \neq 0$ aucun facteur commun ne peut exister. Cette proposition est donc vraie pour n'importe quelles fonctions $F$ et $\Phi$ avec n'importe quels coefficients $f$ et $\varphi$. Par contre, il n'est prouvé jusqu'à maintenant que pour les fonctions décomposables $F$ et $\Phi$, que lorsqu'aucun facteur commun n'existe, alors $R$ doit être différent de 0, d'où il suit que si $R = 0$, un facteur commun existe. Nous voulons maintenant démontrer cette dernière proposition pour n'importe quelles fonctions entières $F$ et $\Phi$, sans supposer leur décomposition en facteurs linéaires. Si cette preuve est faite, alors il sera prouvé pour n'importe quelles fonctions $F$ et $\Phi$ [19] que $R = 0$ est une condition nécessaire et suffisante pour qu'il y ait un facteur en commun ; il vient donc ensuite que $R \neq 0$ est la condition nécessaire | et suffisante pour qu'il n'existe aucun 89|90 facteur commun.

Maintenant nous devons d'abord trouver une méthode à partir de laquelle nous

19. C'est-à-dire sans la supposition de l'existence des racines de $F_n.\Phi$.

pourrons décider si un facteur commun existe et qui nous apprendra à déterminer un tel facteur.

Cette méthode est ancienne, déjà *Euclide* connaissait le procédé de décomposition en fraction continue, méthode grâce à laquelle on trouve le plus grand facteur commun de deux nombres.

On doit déterminer le plus grand facteur commun entre $f(x)$ et $f_1(x)$.

Nous supposons que $f_1$ est d'un degré inférieur ou égal à $f$ et nous divisons successivement :

$$\frac{f(x)}{f_1(x)} = g_1(x) - \frac{f_2(x)}{f_1(x)} \text{ , } \frac{f_1(x)}{f_2(x)} = g_2(x) - \frac{f_3(x)}{f_2(x)} \text{ etc.}$$

Nous obtenons la chaîne :

$$0 = f - g_1 f_1 + f_2 = f_1 - g_2 f_2 + f_3 = f_2[-]g_3 f_3 + f_4 = ... = f_{\nu-2} - g_{\nu-1} f_{\nu-1} + f_\nu$$

Ici, par la division, le degré s'abaisse toujours, de sorte que l'on doive à un moment arriver au degré 0, c'est-à-dire à une constante. Maintenant celle-ci peut être soit différente de 0, soit égale à 0. Dans le premier cas, soit $f_\nu$ cette constante, de sorte que la dernière équation de la chaîne soit : $f_{\nu-1} - g_\nu f_\nu = 0$. Dans le dernier cas soit $f_{\nu+1}$ la constante, laquelle disparaît, de sorte qu'ici aussi la dernière équation de la chaîne soit $f_{\nu-1} - g_\nu f_\nu = 0$.

Nous voyons que cette division successive est identique à la transformation de la fonction $\dfrac{f_1(x)}{f(x)}$ en une fraction continue. On a en effet :

$$\frac{f_1(x)}{f(x)} = \frac{1}{g_1 - \dfrac{f_2}{f_1}} = \cfrac{1}{g_1 - \cfrac{1}{g_2 - \cfrac{1}{g_3 - \cfrac{1}{...\cfrac{1}{g_\nu}}}}}$$

Maintenant il ressort aisément les deux propositions suivantes :

1. Le facteur commun de deux $f$ consécutifs est un facteur de chaque $f$.
2. Chaque $f$ est une fonction homogène linéaire de chaque paire de $f$ qui se succède.

90|91 Le facteur commun de deux | $f$ successives par exemple de $f_1$ et $f$ est aussi un facteur de $f_\nu$. Donc le plus grand facteur commun de $f$ et $f_1$ doit être aussi un facteur de $f_{\nu-1}$, donc de chaque $f$, d'où il suit que $f_\nu$ est le plus grand facteur commun de $f$ et $f_1$. Ici chaque $f$ est une fonction linéaire homogène de chaque paire de $f$ qui se succède, il suit donc : $f_\nu = \varphi_1 f + \varphi f_1$. Il ressort de là que $f$ et $f_1$ ne peuvent avoir de facteur commun plus grand. Si nous exprimons la proposition disant que $f_\nu$ est le plus grand facteur commun de $f$ et $f_1$ par le symbole : $f_\nu = F_h(f, f_1)$, alors celui-ci se décompose en 3 équations :

$$\left.\begin{array}{l} f = \Psi f_\nu \\ f_1 = \Psi_1 f_\nu \\ f_\nu = \varphi_1 f + \varphi f_1 \end{array}\right\}$$

qui expriment ensemble que $f_\nu$ est le plus grand facteur commun. Ces 3 équations sont équivalentes à $f_\nu = F_h(f, f_1)$, de laquelle on parvient à celles-ci et inversement de celles-ci à celle-là. Maintenant $f_\nu$ est soit une fonction entière de $x$, de sorte que $f_{\nu+1} = 0$, et dans ce cas $f$ et $f_1$ ont véritablement un facteur commun $f_\nu$ ; soit $f_\nu$ est une constante différente de 0, et dans ce cas il n'existe aucun facteur commun à $f$ et $f_1$. Il existe ainsi le critère suivant : pour décider si $f$ et $f_1$ ont une partie commune, on divise successivement d'après le procédé des fractions continues, jusqu'à ce que l'on atteigne une fonction de degré 0. Si celle-ci est égale à 0, alors la fonction

précédente est le plus grand facteur commun entre $f$ et $f_1$ ; si c'est une constante différente de 0, alors il n'existe pas de tel facteur. Les coefficients de $f_\nu$ sont formés rationnellement à partir des coefficients de $f$ et $f_1$, le développement en fraction continue étant véritablement un procédé rationnel. Si nous appliquons maintenant cela à $F$ et $\Phi$, alors nous arrivons sûrement à l'équation : $\Phi\Psi' + F\mathcal{G}' = R'$. Il reste maintenant à savoir si cette équation est conforme avec la précédente : $\Phi\Psi + F\mathcal{G} = R$, après que nous avons fait, par la multiplication de $\Psi', \mathcal{G}', R'$ avec un éventuel dénominateur général, des $f_i$ et $\varphi_i$ des fonctions entières. Pour décider cela nous démontrons que, lorsque nous sommes parvenus de $f$ et $f_1$ à $f_\nu$, les multiplicateurs $\varphi$ et $\varphi_1$, qui satisfont les équations : $\varphi_1 f + \varphi f_1 = f_\nu$, sont complètement déterminés. À supposer en effet que | les multiplicateurs $\varphi'$ et $\varphi'_1$ ont effectué cela, on a ensuite : 91|92

$$\varphi'_1 f + \varphi' f_1 = f_\nu \text{ , donc : } (\varphi_1 - \varphi'_1)f + (\varphi - \varphi')f_1 = 0$$

$$\varphi(\varphi_1 - \varphi'_1)f + (\varphi - \varphi')(f_\nu - \varphi_1 f) = 0$$

$(\varphi - \varphi')f_\nu$ doit donc être divisible par $f$. Maintenant on peut accepter sans perte de généralité que $\varphi$ soit de degré plus bas que $f$ [20]. Il est donc possible que $\varphi - \varphi'$ soit divisible par $f$ seulement lorsque $\varphi = \varphi'$. On montre tout aussi bien que $\varphi_1 = \varphi'_1$. Dans l'équation $\varphi_1 f + \varphi f_1 = f_\nu$, où $\varphi_1$ est de degré plus bas que $f_1$ et $\varphi$ que $f$, les multiplicateurs sont donc complètement et clairement déterminés. Si de là nous arrivons finalement à l'équation

$$\Phi\Psi' + F\mathcal{G}' = R'$$

avec l'aide de la décomposition en fraction continue de $F$ et $\Phi$, alors celle-ci peut se distinguer de $\Phi\Psi + F\mathcal{G} = R$ seulement par une constante multiplicative, lorsque nous avons fait de $\Psi'\mathcal{G}', R'$ aussi des fonctions entières de $f$ et $\varphi$ ; à savoir $\Phi\Psi + F\mathcal{G} = R$ peut devenir $\Phi\Psi' + F\mathcal{G}' = R'$ seulement par une multiplication avec une telle constante et non par une division, car sinon $R$ devrait être décomposable en facteurs, et cela n'est pas le cas, on le voit facilement de la manière suivante. Supposons que :

$$R(f_1 ... f_n; \varphi_1 ... \varphi_\nu) = \prod(x_h - \xi_k)$$

soit décomposée en deux facteurs :

$$F(f_1 ... f_n; \varphi_1 ... \varphi_\nu) \text{ et } S(f_1 ... f_n; \varphi_1 ... \varphi_\nu)$$

Alors $F$ doit contenir au moins un des facteurs linéaires $x_K - \xi_h$, donc : $F = (x_k - \xi_h)F(x1..., \xi_1....)$, mais comme $F$ est symétrique en tous les $x$ et tous les $\xi$, alors ce sont tous les facteurs $(x_k - \xi_h)$ qu'il doit contenir, c'est-à-dire que $F$ doit être divisible par $R$, c'est-à-dire que

$$F = R \text{ , } S = 1.$$

Si nous voulons ainsi arriver avec l'aide du développement en fraction continue à l'équation :

$$\Phi\Psi' + F\mathcal{G}' = R'$$

dans laquelle $\Psi'\mathcal{G}', R'$ sont des fonctions entières des $f$ et des $\varphi$, et si nous avons quelque facteur commun | existant entre $\Psi'\mathcal{G}'$ et $R'$, alors nous parvenons à l'équation précédente : 92|93

$$\Phi\Psi + F\mathcal{G} = R$$

Nous voulons maintenant illustrer ce qui précède par un exemple. Soit $f = ax^3 + bx^2 + cx + d$, $f_1 = \alpha x^2 + \beta x + \gamma$, on établit ensuite :

$$f_2 = (a\alpha\gamma + b\alpha\beta - c\alpha^2 - a\beta^2)\frac{x}{\alpha^2} - \frac{a\beta\gamma - b\alpha\gamma + d\alpha^2}{\alpha^2} = px - q \text{ (br. c.)}$$

20. Voir sur ce point l'appendice page....

$$f_3 = -\frac{1}{p^2}(pq\beta + p^2\gamma + q^2\alpha) \ , \ g_1 = \frac{a}{\alpha}x - \frac{a\beta - b\alpha}{\alpha^2} \ , \ g_2 = \frac{\alpha}{p}x + \frac{p\beta + q\alpha}{p^2}$$

Si nous remplaçons $p$ et $q$ par leur valeur, alors nous établissons pour $f_3$ :

$$\begin{aligned} -p^2\alpha^2 f_3 &= a^2\gamma^3 + b^2\alpha\gamma^2 + c^2\alpha^2\gamma + d^2\alpha^3 - ab\beta\gamma^2 + ac\gamma(\beta^2 - 2\alpha\gamma) \\ &\quad + ad\beta(3\alpha\gamma - \beta^2) - bc\alpha\beta\gamma + bd\alpha(\beta^2 - 2\alpha\gamma) - cd\alpha^2\beta \end{aligned}$$

On a de plus : $f_3 = (g_1g_2 - 1)f_1 - g_2 f$, d'où il s'ensuit que : $-\alpha^2 p f_3 = R$. Si l'on prend le cas particulier où $a = 1$, $b = -f_1$, $c = f_2$, $d = -f_3$, $\alpha = 3$, $\beta = -2f_1$, $\gamma = f_2$ et où aussi $f_1$ est la dérivée de $f$, alors on établit le discriminant, à savoir que :

$$-p^3\alpha^2 f_3 = R = (-1)^3(-f_1^2 f_2^2 + 4f_1^3 f_3 + 4f_2^3 - 18f_1 f_2 f_3 + 27f_3^2),$$

est la valeur du discriminant de l'équation cubique $x^3 - f_1x^2 + f_2x - f_3 = 0$ que nous avons déjà effectivement assez souvent rencontrée.

§39 : Nous voulons désormais fournir la preuve que lorsque 2 fonctions entières $A(x) = x^n - a_1x^{n-1} + ... \pm a_n$ et $B(x) = x^\gamma - b_1x^{\gamma-1} + ... \pm b_\gamma$ n'ont pas de facteur commun, alors le résultant $R(a_1...a_n; b_1...b\gamma)$ doit être différent de 0, et sans qu'il ne soit supposé que $A(x)$ et $B(x)$ sont décomposables en facteurs linéaires.

La décomposition en fraction continue nous conduit, lorsque nous ne supposons aucun facteur commun, à l'équation :

$$P(x)A(x) + Q(x)B(x) = 1,$$

où $P$ et $Q$ sont des fonctions entières de $x$ de degré resp. $\gamma - 1$ et $n - 1$, dont les coefficients sont des fonctions rationnelles de $a$ et de $b$. L'équation est une pure identité, nous pouvons y remplacer $x$ par $x_k$, où $x_k$ est une racine de $F(x) = 0$. On a ensuite :

$$\begin{aligned} Q(x_k)B(x_k) &= 1 + P(x_k)(F(x_k) - A(x_k)) && \text{(A.3)} \\ &= 1 + P(x_k)\sum_{h=1}^{n}(-1)^h(f_h - a_h)x_k^{n-h} && \text{(A.4)} \end{aligned}$$

93|94 | Si l'on multiplie maintenant suivant $k$, alors on établit :

$$\prod_{k=1}^{n} Q(x_k) \prod_{k=1}^{n} B(x_k) = 1 + \sum_{h=1}^{n}(-1)^h(f_h - a_h)\mathcal{G}_h(f_1...f_n)$$

Du côté droit on obtient en effet, à part le 1, un agrégat d'expressions qui contiennent chacune un $f_h - a_h$, et dont les coefficients sont des fonctions symétriques des $x$.

De là $\mathcal{G}_h(f_1...f_n)$ est une fonction entière des $f$, laquelle contient en plus encore des $a$ et est aussi entière en ces $a$.

Maintenant $\prod_{k=1}^{n} Q(x_k)$ est une fonction symétrique des $x$ et une fonction rationnelle des $a$ et des $b$. De là on a : $\prod_{k=1}^{n} Q(x_k) = F(f_1...f_n)$. De plus :

$$\prod_{k=1}^{n} B(x_k) = R(f_1...f_n; b_1...b\gamma)$$

Ainsi :

$$F(f_1...f_n)R(f_1...f_n;b_1...b\gamma) = 1 + \sum_{h=1}^{n}(-1)^h(f_h - a_h)\mathcal{G}_h(f_1...f_n)$$

C'est une pure identité, et l'on peut poser $f_1 = a_1, ...f_n = a_n$ et obtenir :

$$F(a_1...a_n)R(a_1...a_n;b_1...b\gamma) = 1$$

Comme maintenant dans tous les cas $F$ est une grandeur finie, il s'ensuit que :

$$R(a_1...a_n;b_1...b_\gamma) \neq 0 \text{ q. e. d.}$$

Cette preuve ingénieuse vient de *Gauss*, qui a élaboré sur les bases de cette proposition démontrée la preuve de l'existence des racines. *Gauss* emploie ces considérations pour le cas particulier où l'une des deux fonctions est la dérivée de l'autre, puis il établit le résultat disant que si une fonction entière n'a aucun facteur commun avec sa dérivée, alors le discriminant $R = (-1)^{\frac{n.n-1}{2}}P^2$ doit être différent de zéro.

Nous avons maintenant, sans supposer l'existence des racines, prouver que pour n'importe quelles fonctions entières $F(x)$ et $\Phi(x)$, la disparition du résultant $R(F,\Phi)$ est la condition nécessaire et suffisante pour que $F$ et $\Phi$ aient un facteur commun, d'où il va de soi que la condition nécessaire et suffisante pour que deux fonctions entières quelconques n'aient aucun facteur commun est que $R(F,\Phi)$ soit différent de 0.

§40 [21] Nous étudions maintenant la décomposition d'une fonction | entière en 94|95
facteurs. Cette étude importante est analogue à la décomposition d'un nombre en facteurs premiers. Si l'on se donne un nombre $n$, alors nous pouvons nous demander pour tout nombre s'il est contenu sans reste dans $n$. La méthode consiste simplement à regarder pour les nombres compris entre 1 et $n-1$ inclus si un certain nombre de fois celui-ci est égal à $n$. Nous pouvons ensuite de nouveau décomposer cette partie éventuelle en facteurs et nous arrivons finalement à la représentation $n = p_1^{k_1}p_2^{k_2}...$, où $p_1, p_2, ....$ ne sont plus décomposables. Nous appelons nombres premiers ceux divisibles seulement par 1 et eux-mêmes. On a ensuite le fait que la décomposition d'un nombre $n$ en facteurs premiers est unique. La preuve s'appuie sur la proposition disant que si un produit de plusieurs facteurs est divisible par un nombre premier, alors nécessairement un de ses facteurs doit être divisible par un nombre premier ; la méthode de recherche du plus grand facteur commun est donc nécessaire pour la clarté de la méthode de décomposition.

Si nous transposons maintenant cette étude aux fonctions entières, alors nous pouvons dans un premier temps avoir des diviseurs qui sont des purs nombres. Cela se produit lorsque les coefficients des fonctions entières ordonnées selon les puissances de $x$ ont un facteur commun. Nous avons retiré le plus grand facteur commun de tous les coefficients et nous étudions donc maintenant seulement les fonctions entières qui n'ont aucun diviseur entier.

Nous démontrons tout d'abord la proposition suivante :

> Si le produit $\Psi(x)\chi(x)$ de deux fonctions entières est divisible par le nombre premier $p$ et que $\chi$ n'est pas divisible par $p$, alors $\Psi$ doit être divisible par $p$.

Soit :

$$\begin{aligned}\Psi(x) &= a_0 + a_1x + a_2x^2 + ...\\ \chi(x) &= b_0 + b_1x + b_2x^2 + ...\end{aligned}$$

Ici $\chi(x)$ n'est pas divisible par $p$, donc il doit y avoir des coefficients $b$ qui ne sont

21. N.d.T. Nous ajoutons ce numéro de chapitre, absent du manuscrit (la numérotation passe directement de 39 à 41) car il nous semble bien qu'il y a ici un changement de sujet.

pas divisibles par $p$. Si l'on suppose que $b_0, b_1, b_2, ..., b_k - 1$ sont divisibles par $p$, mais que $b_k$ ne l'est pas, alors :

$$(a_0 + a_1x + a_2x^2 + ...)(b_kx^k + b_{k+1}x^{k+1} + ...)$$

doit être divisible par $p$. Si nous effectuons le produit, alors les coefficients de $x^k, x^{k+1}, ...$ sont resp. :

$$a_0b_k, a_1b_k + a_0b_{k+1}, a_2b_k + a_1b_{k+1} + a_2b_{k+2})...$$

95|96 | Ici $b_k$ n'est pas divisible par $p$, donc $a_0$ doit l'être ; il en résulte que $a_1, a_2, ...$ est divisible par $p$ et ici, de la divisibilité de $a_h$ suit celle de $a_{h+1}$, il est ainsi démontré que $\Psi(x)$ est divisible par $p$.

Aussi simple que cette proposition semble être, on ne la trouve pour la première fois que chez *Gauss*. On suppose maintenant que l'on a une fonction entière $\Phi(x)$ décomposable en deux facteurs $\Psi(x)$ et $\chi(x)$, lesquels sont certes des fonctions entières de $x$, dont les coefficients ne sont pas des nombres entiers mais seulement des nombres rationnels. Ensuite nous faisons concorder $\Psi$ avec son dénominateur $n$, donc par exemple : $\Psi(x) = \frac{\Psi_1(x)}{m}\frac{\chi_1(x)}{n}$. Ensuite, comme $\Psi_1\chi_1$ est divisible par $mn$, les facteurs premiers de $mn$ se répartissent, de sorte que chacun soit contenu soit dans $\Psi_1$ soit dans $\chi_1$. Nous obtenons donc deux fonctions entières $\Psi(x)$ et $\chi(x)$, dont le produit est égal à $\Phi(x)$. Donc :

Si une fonction entière à coefficients entiers $\Phi(x)$ est décomposable en fonctions entières à coefficients rationnels, alors elle est aussi décomposable en facteurs entiers à coefficients entiers. Nous avons donc seulement besoin de chercher si une fonction entière à coefficients entiers $\Phi(x)$ a un facteur entier à coefficients entiers. Maintenant nous pouvons simplement pour la fonction donnée :

$$\Phi(x) = \varphi_0 + \varphi_1x + \varphi_2x^2 + ...$$

fixer :

$$(\Psi_0 + \Psi_1x + \Psi_2x^2 + ...)(\chi_0 + \chi_1x + \chi_2x^2 + ...) = \varphi_0 + \varphi_1x + \varphi_2x^2 + ...$$

et ensuite déterminer les nombres entiers $\Psi_0, \Psi_1, \Psi_2, ..., \chi_0, \chi_1, \chi_2, ...$ à partir des équations :

$$\begin{aligned}\Psi_0\chi_0 &= \varphi_0\\ \Psi_0\chi_1 + \Psi_1\chi_0 &= \varphi_1\\ \Psi_0\chi_2 + \Psi_1\chi_1 + \Psi_2\chi_0 &= \varphi_2\end{aligned}$$

Mais cette méthode est d'abord trop compliquée et suppose ensuite la résolution d'un système d'équations, ce que l'on ne peut pas accepter pour le moment. Mais il y a une méthode plus simple qui permet de déterminer tous les coefficients de la même façon que l'on a déterminé ceux de $\Psi_0$ et $\chi_0$ à partir de $\Psi_0\chi_0 = \varphi_0$, et ce simplement comme une partie de $\varphi_0$.

Pour chercher si une fonction donnée $F(x)$ de degré $m$ possède un diviseur entier à coefficients entiers, par exemple de degré $n-1$ (nous avons nécessairement seulement
96|97 à considérer pour les diviseurs les $n-1$ | pris entre 1 et $\frac{m}{2}$ resp. $\frac{m-1}{2}$, selon que $m$ est supposé pair ou impair), nous supposons en effet ce diviseur de degré $n-1$ ordonné non pas d'après les puissances de $x$, mais suivant les fonctions entières de $x$. Soit $\Psi(x)$ le diviseur, de façon à ce que nous puissions écrire selon la formule de *Lagrange* :

$$\Psi(x) = \sum_{k=1}^{n} C_k \frac{F(x, x_k)}{F'(x_k)},$$

où par $x_1, x_2, ...x_n$ on entend n'importe quel nombre entier, à l'exception cependant de ceux pour lesquels $F'(x)$ disparaît. Nous allons désigner ces nombres par $r_1...r_n$. Ensuite nous pouvons ainsi déterminer $C_1, ...C_n$ de façon à ce que pour $x = r_k$ on ait $\Psi(r_k)$ qui soit une partie de $F(r_k)$. Maintenant on a $\Psi(r_k) = C_k$, donc $C_k$ doit être un facteur de $F(r_k)$. Nous obtenons ainsi pour $C_k$ une limite entière des valeurs numériques. Tandis que nous mettons toutes les combinaisons $C_k$ dans $\Psi(x) = \sum_{k=1}^{n} C_k \frac{F(x,r_k)}{F(r_k)}$, nous obtenons toutes les fonctions entières de degré $n-1$ qui peuvent être vraiment un diviseur. De celles-ci sont ensuite à exclure celles qui ne sont pas des fonctions de $x$ à coefficients entiers, et seulement les fonctions entières de $x$ à coefficients entiers restantes peuvent être des diviseurs de $F(x)$, en comprenant comme diviseur une fonction entière à coefficients entiers de $x$ faisant partie de $F(x)$. Nous aurions pu naturellement ordonner aussi le diviseur $\Psi(x)$ selon les $g_h(x)$ (voir...), en effet :

$$\Psi(x) = \sum_{h=0}^{h=1} \sum_{k=1}^{n} \varphi_k \frac{{x_k}^{n-h-1}}{F'(x_k)} g_h(x),$$

où maintenant les $\varphi_k$ sont à déterminer comme diviseurs de $F(x_k) = F(r_k)$. Naturellement aussi dans cette représentation les $\Psi(x)$ sont à considérer entières à coefficients entiers. Ainsi nous avons une méthode parfaitement élémentaire et conduisant dans chaque cas à notre but pour décider si la fonction entière à coefficients entiers $F(x)$ est divisible ou pas.

Par exemple :

$$F(x) = x^5 - x^4 + x^3 - x^2 - 3x - 1.$$

Nous voulons rechercher si un diviseur de degré 3 existe. Nous posons donc $n = 4$ et appliquons $r_1 = 0, r_2 = 1, r_3 = -1, r_4 = -2$, ainsi on a : $F(0) = -1, F(1) = -4, F(-1) = -2, F(-2) = -55$. $C_1, C_2, C_3, C_4$ doivent donc être des diviseurs de ces nombres $F(r_k)$.

| Maintenant on a : 97|98

$$\begin{aligned}
\frac{F(x,r_1)}{F'(r_1)} &= -\frac{1}{2}(x^3 + 2x^2 - x - 2) \\
\frac{F(x,r_2)}{F'(r_2)} &= \frac{1}{6}(x^3 + 3x^2 + 2x) \\
\frac{F(x,r_3)}{F'(r_3)} &= \frac{1}{2}(x^3 + x^2 - 2x) \\
\frac{F(x,r_4)}{F'(r_4)} &= \frac{1}{6}(x^3 - x)
\end{aligned}$$

Si l'on choisit pour $c_1, c_2, c_3, c_4$ resp. 1,4,-2,-11, alors il suit :

$$\Psi(x) = \sum_{k=1}^{n} c_k \frac{F(x,x_k)}{F'(x_k)} = x^3 + 2x + 1$$

Ce choix apporte donc une fonction entière à coefficients entiers. Elle est effectivement une partie de $F(x)$, car par la division par $x^3 + 2x + 1$ de $F(x)$ on obtient :

$$x^5 - x^4 + x^3 - x^2 - 3x - 1 = (x^3 + 2x + 1)(x^2 - x - 1)$$

§41. Après que nous avons obtenu une méthode qui permet dans chaque cas de décider si une fonction entière à coefficients entiers de $x$ donnée a des diviseurs entiers à coefficients entiers, nous sommes autorisé à établir le concept d'indécomposibilité ou d'irréductibilité. Par fonction de $x$ indécomposable ou irréductible nous comprenons une fonction entière de $x$ à coefficients entiers qui ne peut pas être décomposée en diviseurs entiers à coefficients entiers, et qui ne contient donc aucun nombre entier comme facteur. Les fonctions irréductibles jouent auprès des fonctions un rôle

semblable à celui des nombres premiers auprès des nombres entiers. Chaque fonction entière à coefficients entiers peut se représenter désormais par la forme :

$$F(x) = p_1^{h_1} p_2^{h_2} ... p_\mu^{h_\mu} f_1^{k_1} f_2^{k_2} ... f_\nu^{k_\nu},$$

où $p_1 ... p_\mu$ sont des nombres premiers, $f_1, ..., f\mu$ des fonctions irréductibles et $h_1 ... h_\mu, k_1 ... k_\nu$ des nombres entiers positifs. Il est encore vrai que cette décomposi-
98|99 tion, comme celle des nombres premiers, est unique. | Pour démontrer cette unicité, nous avons d'abord à prouver la proposition suivante.

Lorsque le produit $\Phi.\Psi$ de deux fonctions entières de $x$ à coefficients entiers $\Phi$ et $\Psi$ est divisible par une fonction irréductible $F(x)$, alors soit $\Phi$ soit $\Psi$ doit être divisible par $F(x)$.

Supposons en effet que $\Phi$ ne soit pas divisible par $F$, alors $\Phi$ et $F$ ne peuvent avoir aucun facteur commun, car $F$ est irréductible. Nous pouvons donc déterminer 2 fonctions entières $\Phi_1$ et $F_1$ à coefficients rationnels de sorte que : $\Phi_1 F + F_1 \Phi = 1$. Il suit alors :

$$\Psi\Phi_1 F + F_1 \Phi\Psi = \Psi$$

Maintenant le côté gauche est divisible par $F$, donc aussi $\Psi$, mais éventuellement le quotient n'est pas entier. De là si l'on pose $\Psi = F.\chi$ et si l'on désigne le dénominateur commun en $\chi$ par $m$, alors, car $\Psi$ est entière et $F$ ne contient non plus aucun nombre entier comme facteur, tous les facteurs premiers de $m$, donc $m$ lui-même, peuvent s'absorber dans $\chi$, c'est-à-dire que $F$ est aussi un diviseur entier de $\Psi$, de sorte que le quotient $\chi$ est une fonction entière de $x$ à coefficients entiers. q.e.d.

Il résulte d'emblée de cette proposition l'unicité de la décomposition d'une fonction entière à coefficients entiers en facteurs irréductibles. Supposons donc donnée une deuxième décomposition $q_1^{r_1} ... q_{\mu'}^{r'_\mu} \varphi_1^{s_1} ... \varphi_{\nu'}^{s_{\nu'}}$ qui soit donc égale à $p_1^{h_1} ... p_\mu^{h_\mu} f_1^{k_1} ... f_\nu^{k_\nu}$ et qu'aucun des facteurs d'un côté ne soit divisible par l'un des facteurs de l'autre côté, tandis que cependant le produit, lui, soit divisible. Cela contredit la proposition qui vient d'être démontrée, par conséquent la décomposition est unique.

§ 43[§42] : Nous voulons maintenant étendre ce résultat aux fonctions de plusieurs variables.

Nous disons qu'une fonction entière à coefficients entiers des variables $x, y, z, ...$ est un facteur d'une fonction entière à coefficients entiers $F(x, y, z, ...)$ lorsque leur quotient est encore une fonction entière à coefficients entiers des variables $x, y, z, ....$ Nous appliquons le principe d'induction et nous supposons la décomposabilité de toutes
99|100 les fonctions ayant une variable de moins que $F(x, y, z, ...)$ étudiée | et son unicité démontrée.

Nous cherchons ensuite d'abord si $F(x, y, z, ...)$ a des facteurs qui sont indépendants de $x$. Un tel facteur doit être un facteur de tous les coefficients $f_k(y, z, ...)$ des fonctions $F(x, y, z...)$ ordonnées selon $x$. Nous pouvons tout de suite régler cette étude, étant donné que nous savons déjà par hypothèse tout ce qui est requis sur la divisibilité des fonctions $f(y, z, ...)$.

Deuxièmement nous pouvons maintenant rechercher si $F(x, y, z...)$ a pour facteur une fonction entière de $x, y, z, ...$ à coefficients entiers, dans lequel nous supposons déjà éliminées par division toutes les parties libres de $x$ de $F(x, y, z, ...)$. Nous supposons ensuite un tel facteur ordonné selon les fonctions entières $G_k(x)$, alors les coefficients $C_k$ peuvent être des facteurs de $F(r_k, y, z, ...)$. Nous pouvons tous les retrouver et obtenir pour les $C_k$ un nombre fini de fonctions entières de $y, z, ...$ à coefficients entiers (pag. 244). Nous obtenons donc un nombre fini de tentatives pour décider si $F(x, y, z, ...)$ possède un facteur $f(x, y, z...)$.

Troisièmement nous démontrons, puisque nous sommes déjà autorisé à parler de fonctions entières à coefficients entiers de plusieurs variables irréductibles, que lorsqu'un produit $\Phi.\Psi$ de deux fonctions entières à coefficients entiers des variables $x, y, z, ...$ est tout d'abord divisible par une fonction irréductible $\varphi(y, z, ...)$, alors soit

$\Phi$ soit $\Psi$ doit être divisible par $\varphi(y, z...)$. Soit donc :

$$\Phi.\Psi = (a_0 + a_1x + a_2x^2 + ...)(b_0 + b_1x + b_2x^2 + ...)$$

divisible par $\varphi(y, z, ...)$ (pag. 2...), où les $a$ et les $b$ sont des fonctions entières de $y, z, ...$ à coefficients entiers et soit $\Psi = b_0 + b_1x + b_2x^2 + ...$ qui n'est pas divisible par $\varphi$, alors tous les $b$ ne doivent pas être divisibles par $\varphi$. Supposons que $b_k$ soit le premier coefficient à ne pas être divisible par $\varphi$, alors :

$$(a_0 + a_1x + a_2x^2 + ...)(b_kx^k + b_{k+1}x^{k+1} + ...)$$

doit être divisible par $\varphi$. On doit aussi avoir :

$$a_0b_k, a_0b_{k+1} + a_1b_k, ...$$

qui soit divisible par $\varphi$, car pour les fonctions des variables $y, z, ...$ toutes les propositions sur la divisibilité sont supposées être démontrées, il vient donc de cette suite que $a_0, a_1, a_2...$ sont divisibles par $\varphi$, c'est-à-dire que $\Psi$ est divisible par $\varphi$. | De là 100|101
il suit comme corollaire : si une fonction entière à coefficients entiers $\Psi(x, y, z...)$ est divisible par une même fonction $f(x, y, z, ...)$, laquelle n'a aucun facteur indépendant de $x$, de sorte que le quotient soit une fonction entière de $x$, alors ce quotient doit aussi être une fonction entière à coefficients entiers des variables. Car soit :

$$\Psi(x, y, z...) = f(x, y, z...)\frac{Q(x, y, z...)}{R(y, z...)}$$

où nous désignons par $R$ le dénominateur général de $Q$, alors $f.Q$ est divisible par $R$ et les $f$ ne le sont pas, donc $Q$ doit être divisible par $R$, c'est-à-dire que le quotient est une fonction entière de $x, y, z...$ à coefficients entiers. Donc la question de la divisibilité d'une fonction entière à coefficients entiers conçue comme fonction de $x$ seulement résout aussi la question de la divisibilité de cette fonction conçue comme fonction de toutes les variables.

Maintenant nous voulons finalement prouver la propriété selon laquelle lorsque le produit $\Phi.\Psi$ de deux fonctions entières à coefficients entiers est divisible par la fonction irréductible $F(x, y, z, ...)$, alors soit $\Phi$ soit $\Psi$ doit être divisible par $F$.

Si l'on suppose que $\Phi$ n'est pas divisible par $F$, alors, comme $F$ est irréductible, $\Phi$ et $F$ n'ont pas non plus, comme fonction de $x$, de facteur commun. Nous pouvons donc poser : $\Phi F_1 + F\Phi_1 = G(y, z, ...)$, où $F_1\Phi_1$ et $G$ sont des fonctions entières de $y, z, ...$ à coefficients entiers. Il s'ensuit que :

$$\Psi\Phi F_1 + \Psi F\Phi_1 = \Psi G(y, z, ...)$$

Le côté gauche est maintenant divisible par $F$, donc aussi le côté droit et $\frac{\Psi G(y, z, ....)}{F(x, y, z...)}$ doit être une fonction entière à coefficients entiers, que l'on désignera par $H(x, y, z....)$. Donc ensuite $H.F$ est divisible par $G$, et comme $F$ n'est pas divisible, c'est $H$ qui l'est. Nous pouvons de là écrire $\Psi = H_1.F$, où $H_1$ est une fonction entière de $x, y, z[...]$ à coefficients entiers, c'est-à-dire que $\Psi$ est divisible par $F$. De cette proposition suit finalement l'unicité de la décomposition d'une fonction entière à coefficients entiers en facteurs irréductibles.

101|102

| §43. Nous appliquons le résultat établi à la recherche des discriminants et des résultants.

Soit $P^2 = (-1)^{\frac{n.n-1}{2}}\prod(x_h - x_k)^2 = D(f_1, f_2, ...f_n)$ une fonction entière de $f_1...f_n$ à coefficients entiers. Si l'on conçoit $P^2$ comme une fonction de $f_1...f_n$, alors elle est

irréductible. Elle contient donc un facteur $F(f_1, ..., f_n)$ et si l'on remplace les $f$ par les $x$, alors $F$ doit au moins contenir un facteur $x_h - x_k$ et comme $F$ est symétrique, alors elle les contient tous, et ce, tous au carré à cause de la symétrie, c'est-à-dire que $D(f_1...f_n)$ est irréductible.

Cela est certes seulement démontré pour le cas spécial de $f$ qui est symétrique par rapport aux $x$. Mais si une fonction est indécomposable pour un système de valeur particulier alors elle ne peut pas être décomposable de façon générale. Nous voyons donc ici que $D(f_1...f_n)$ est irréductible, même si l'on considère les $f$ comme n'importe quelle variable entière.

Que le résultant $R(f_1...f_n; \varphi_1...\varphi_\nu)$ soit aussi irréductible lorsque les $f$ et les $\varphi$ sont définies par $x$ resp. $\xi$, nous l'avons déjà prouvé pag..... Nous voulons rendre cela un peu général. Soit les deux fonctions, que nous supposerons tout d'abord décomposables, représentées par

$$F(x) = a_0x^n + a_1x^{n-1} + ... + a_n = a_0(x^n + \frac{a_1}{a_0}x^{n-1} + ... + \frac{a_n}{a_0})$$

$$\Phi(x) = \alpha_0x^\nu + \alpha_1x^{\nu-1} + ... + \alpha_\nu = a_0(x^\nu + \frac{\alpha_1}{\alpha_0}x^{\nu-1} + ... + \frac{\alpha_\nu}{\alpha_0})$$

Ensuite on a : $R = \prod_{k=1}^{n} \Phi(x_k)$ un quotient dont le numérateur est une fonction entière à coefficients entiers des $a$ et des $\alpha$ et dont le dénominateur est le produit d'une puissance de $a_0$ par une puissance de $\alpha_0$. Nous voulons montrer que dans cette représentation, on obtient déjà une fonction entière à coefficients entiers de tous les $a$ et les $\alpha$ lorsque nous multiplions par $a_0^\nu$. On a donc :

$$a_0^\nu \prod_{k=1}^{n} \Phi(x_k) = a_0^\nu \prod_{k,h} \alpha_0(\alpha_k - \xi_h) = a_0^\nu\alpha_0^n \prod_{k,h}(\alpha_k - \xi_h)$$

À partir de cette représentation on s'aperçoit que les dénominateurs les plus hauts se trouvant dans $\prod_{k,h}(\alpha_k - \xi_h)$ sont $a_0^\nu$ et $\alpha_0^n$ ; on a ainsi :

$$a_0^\nu \prod_{k=1}^{n} \Phi(x_k) = a_0^\nu\alpha_0^n \prod_{k,h}(\alpha_k - \xi_h) = (-1)^{n\nu}\alpha_0^n \prod_{h=1}^{\nu} F(\xi_h)$$

est une fonction entière à coefficients entiers de $a_0...a_n$ et $\alpha_0...\alpha_\nu$.

102|103 | Si l'on désigne celui-ci par $R(a_0...a_n; \alpha_0...\alpha_\nu)$, alors on doit montrer que cette fonction que nous avons jusqu'à présent désignée comme le résultant de $F(x)$ et $\Phi(x)$ est irréductible. Mais il suit maintenant de :

$$R(a_0...a_n; \alpha_0...\alpha_\nu) = a_0^\nu\alpha_0^n \prod_{h,k}(x_k - \xi_h),$$

que $R(a_0....a_n; \alpha_0...\alpha_\nu$ est égale à $a_0^\nu\alpha_0^n$ multipliée par une fonction irréductible de $\frac{a_1}{a_0}, ...\frac{a_n}{a_0}; \frac{\alpha_1}{\alpha_0}, ...\frac{\alpha_\nu}{\alpha_0}$. $R$ pourrait donc avoir au plus pour facteur une puissance de $a_0$ ou de $\alpha_0$ ; qu'il n'en est pas ainsi est montré par le fait que $R$ contient un terme $a_n^\nu\alpha_0^n$ dans lequel il n'y a aucun $a_0$ et également un terme $\alpha_\nu^n a_0^\nu$ qui ne contient aucun $\alpha_0$. Ainsi il est démontré que $R(a_0...a_n; \alpha_0...\alpha_\nu$ est irréductible lorsque $\frac{a_k}{a_0}$ et $\frac{\alpha_k}{[\alpha_0]a_0}$ sont définies comme fonctions symétriques des $x$ resp. des $\xi$. Comme donc $R(a_0...a_n; \alpha_0...\alpha_\nu)$ est irréductible pour des $a_0...a_n; \alpha_0...\alpha_\nu$ particuliers, alors il est aussi irréductible en général.

Le résultant $R(a_0...a_n; \alpha_0...\alpha_\nu)$ de deux fonctions

$$a_0x^n + a_1x^{n-1} + .... + a_n \text{ et } \alpha_0x^\nu + ... + \alpha_\nu,$$

que l'on n'a pas supposées décomposables est donc irréductible.

On peut ajouter ici à la preuve que le résultant est un invariant, c'est-à-dire qu'il reste inchangé, lorsque l'on remplace dans les fonctions $F(x)$ et $\Phi(x)$ les $x$ par la fonction homographique $\frac{ax+b}{cx+d}$, les fonctions permettent complètement de construire les résultants à partir des coefficients. Nous supposons de nouveau en premier que $F(x)$ et $\Phi(x)$ sont décomposables. La preuve de l'invariance sera conduite successivement en montrant tout d'abord que $R$ n'est pas modifié lorsque l'on remplace $x$ par $-\frac{1}{2}$. Nous obtenons ensuite en formant les fonctions :

$$z^n F(-\frac{1}{2}) = F_1(z),\ \ z^\nu \Phi(-\frac{1}{2}) = \Phi_1(z)$$

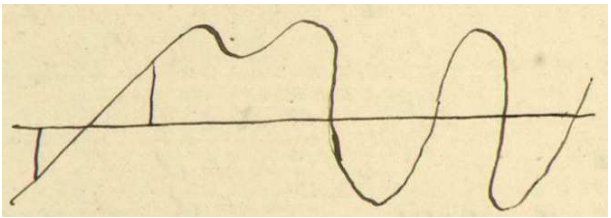
$$R(F_1, \Phi_1) = a_n^\nu \prod_{k=1}^n \Phi_1(-\frac{1}{x_k}) = a_n^\nu \prod_{k=1}^n (-1)^\nu x_k^{-\nu} \Phi(x_k)$$

$$= (-1)^{n\nu} a_n^\nu \frac{1}{(x_1 ... x_n)^\nu} \prod_{k=1}^n \Phi(x_k) = a_0^\nu \prod_{k=1}^n \Phi(x_k) = R(F, \Phi)$$

Nous nous apercevons également à partir de la représentation : $R = a_0^\nu \alpha_0^n \prod_{k,h}(x_k - \xi_h)$ que $R$ reste inchangé, lorsque l'on augmente $x$ dans $F$ et $\Phi$ d'une grandeur constante. Par cette application successive de ces deux opérations, nous pouvons construire une fonction homographique $z = \frac{ax+b}{cx+d}$, | qui, posée à 103|104
la place de $x$, ne change pas $R$, lorsque l'on pose la substitution du déterminant $ad - bc = 1$. Si nous désignons pour cette substitution, après que nous avons complètement construit à nouveau $F$ et $\Phi$, les coefficients avec $a'_0 ... a'_n; \alpha'_0 ... \alpha'_\nu$, alors on a : $R(a_0 .... a_n; \alpha_0 ... \alpha_n) = R(a'_0 .... a'_n; \alpha'_0 ... \alpha'_n)$.

Les $a'_0 .... a'_n; \alpha'_0 ... \alpha'_n$ se composent des $a_0 .... a_n; \alpha_0 ... \alpha_n$, de sorte que cette équation soit conçue comme une pure identité. De là elle est valable aussi lorsque les $a$ et les $\alpha$ sont n'importe quel coefficient d'équation, c'est-à-dire aussi lorsque la possibilité de décomposer $F$ et $\Phi$ n'est pas supposée.

§44. Nous revenons maintenant au développement en fraction continue du §38 pour parvenir à la proposition de *Sturm*. Nous voulons supposer dans la suite que la notion de racine d'une équation est déjà définie. Nous reviendrons encore sur les justifications rigoureuses et les fondements de ces notions plus tard indépendamment des développements qui suivent maintenant. Nous renouons avec avec l'intuition géométrique, laquelle nous apprend que si une courbe $y = F(x)$ passe tantôt en dessous, tantôt au dessus de l'axe des abscisses, elle doit le couper, dans le cas où la courbe se déroule continuement.

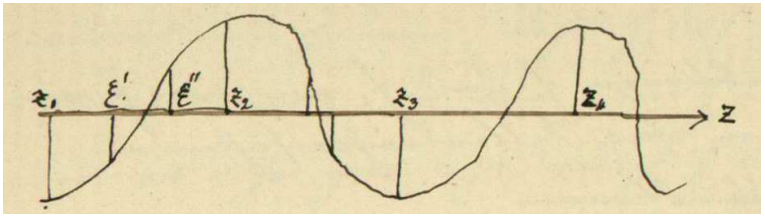


La notion de continuité géométrique est, vue ainsi, claire. Nous désignerons analytiquement la fonction $F(x)$ comme continue en $x$, si $F(x_1) - F(x_2)$, où $x_1 < x < x_2$, peut être faite plus petite que n'importe quelle grandeur, lorsque l'on rapproche $x_1$ et $x_2$ aussi près de $x$ que l'on veut et lorsque cette différence reste ensuite aussi plus petite que cette grandeur choisie aussi petite que l'on veut. Ce concept de continuité est si simple, mais aussi tellement moderne, qu'on ne le trouve pas encore chez *Gauss*. Il avait seulement la conception géométrique et le concept de continuité géométrique.

Nous voulons maintenant, pour fonder la proposition de *Sturm* sur cette conception géométrique des racines de $F(x) = 0$, définir un symbole, que nous désignerons

par $[A]$ et qui doit représenter $-1, 0$ ou $+1$, suivant que la grandeur supposée réelle
104|105 $A$ est négative, égale à zéro, ou positive. Si $a$ et $b$ ont les mêmes signes, | alors il suit $\frac{1}{2}\{[a]+[b]\} = [a+b]$. Si nous posons $b = -c$ et que l'on considère que : $-[c] = [-c]$, alors il suit : $\frac{1}{2}\{[a]-[c]\} = [a-c]$, où il sera donc supposé que $a$ et $c$ sont de signes contraires.

Si l'on suppose maintenant, qu'entre $z_1$ et $z_2$, où $z_1 < z_2$, la courbe $y = F(z)$ passe une fois, et une fois seulement par l'axe, alors entre $z_1$ et $z_2$ se trouve une unique racine de $F(z) = 0$.




Ainsi $F(z_1)$ et $F(z_2)$ sont donc de signes différents et on a :

$$\frac{1}{2}\{[F(z_2)]-[F(z_1)]\} = [F(z_2)-F(z_1)],$$

si $F(z_1)$ est positif et $F(z_2)$ est négatif, ou l'inverse. Cette équation reste valable, quand $z_1$ et $z_2$ se rapprochent, seulement si les valeurs des fonctions correspondantes ont encore des signes différents.

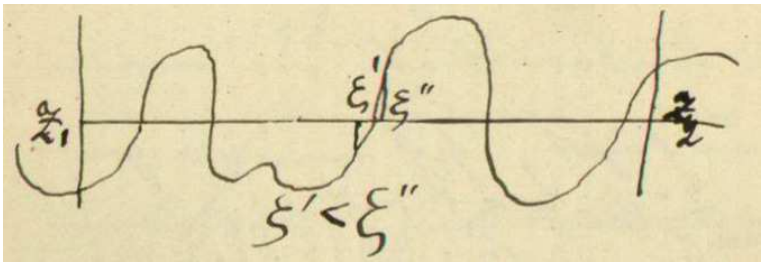
$\zeta'$ et $\zeta''$ peuvent être 2 valeurs proches de la variable, de sorte que encore $F(\zeta')$ reste du signe de $F(z_1)$ et $F(\zeta'')$ du signe de $F(z_2)$. $(z_1 < \zeta' < \zeta'' < z_2)$. Ensuite on a aussi :

$$[F(\zeta'')-F(\zeta')] = \frac{1}{2}\{[F(z_2)]-[F(z_1)]\}.$$

Nous prenons maintenant un second intervalle de $z_2$ à $z_2[z_3]$, où $z_2 < \zeta_1' < \zeta_1'' < z_3$ et où entre $\zeta_1'$ et $\zeta_1''$ peut de nouveau se trouver une racine, puis il suit :

$$[F(\zeta_1'')-F(\zeta_1')] = \frac{1}{2}\{[F(z_3)]-[F(z_2)]\}.$$

Si l'on poursuit ainsi pour l'intervalle suivant et que l'on somme toutes ces équations, alors on a réussi à faire partir du côté droit tous les $F(z)$, sauf les deux qui correspondent aux limites que l'on a déjà désignées par $z_1$ et $z_2$ $(z_1 < z_2)$.




Nous obtenons :

$$\sum_{\zeta}[F(\zeta'')-F(\zeta')] = \frac{1}{2}\{[F(z_2)-F(z_1)]\}.$$

La somme du côté gauche est à étendre à tous les $\zeta'$ et $\zeta''$, entre lesquels la fonction $F(\zeta)$ a changé une fois son signe. En même temps $z_2$ est $>$ au plus grand $\zeta''$ et $z_1$

inférieur au plus petit $\zeta'$. Si l'on suppose maintenant $\zeta'$ et $\zeta''$ quelconques et proches l'un de l'autre, mais suffisamment éloignés naturellement pour que $F(\zeta')$ et $F(\zeta'')$ aient des signes différents, alors il vient :

$$[F(\zeta'') - F(\zeta')] = \left[\frac{F(\zeta'') - F(\zeta')}{\zeta'' - \zeta'}\right] = [F'(\zeta)],$$

lorsque nous désignons par $\zeta$ le point se trouvant entre deux points infiniment près l'un de l'autre, pour lequel $F(\zeta) = 0$.

| Nous avons donc : 105|106

$$\sum_{\zeta}[F'(\zeta)] = \frac{1}{2}\{[F(z_2)] - [F(z_1)]\}$$

où la somme est à étendre à tous les $\zeta$ pour lesquels $z_2 > \zeta > z_1$ et $F(\zeta) = 0$. Ce résultat a le caractère de la proposition de *Cauchy*, c'est-à-dire que le comportement de la fonction aux limites nous donne des éclaircissements sur son comportement entre les deux.

Nous prenons maintenant deux fonctions entières $f$ et $f_1$, où $f$ est d'un degré plus grand que $f_1$ et nous appliquons le développement en fraction continue. Nous obtenons ainsi la suite de restes :

$$f, f_1, f_2, f_3, ..., f_{h-1}, f_h, f_{h+1}, ..., f_\nu,$$

où $f_\nu$ est de degré 0. Si l'on pose maintenant, en écrivant $x$ pour $z$ :

$$F(x) = f_{h-1}(x)f_h(x)$$

et que l'on désigne les racines réelles de $f_{h-1} = 0$ resp. $f_h = 0$ par $\xi^{h-1}$ resp. $\xi^h$, alors on a : $F'(x) = f'_{h-1}f_h + f_{h-1}f'_h$ et comme nous avons dans l'équation ci-dessus seulement de telles valeurs de la variable pour lesquelles $F$ disparaît, il suit donc :

$$\sum_{\xi}[F'(\xi)] = \sum_{\xi^{h-1}}[f'_{h-1}(\xi^{h-1})f_h(\xi^{h-1})] + \sum_{\xi^h}[f_{h-1}(\xi^h)f'_h(\xi^h)].$$

Donc on a :

$$\sum_{\xi^{h-1}}[f'_{h-1}(\xi^{h-1})f_h(\xi^{h-1})] + \sum_{\xi^h}[f_{h-1}(\xi^h)f'_h(\xi^h)] = \frac{1}{2}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}[f_{h-1}(x_1)f_h(x_1)]$$

On somme maintenant $h$ de $h = 1$ à $h = r$, en posant $f_0 = f$. On a donc maintenant : $f_{h-1}(x) - g_h(x)f_h(x) + f_{h+1}(x) = 0$. Si l'on introduit les racines $\xi^h$ de $f_h$, alors il suit :

$$f_{h-1}(\xi^h) + f_{h+1}(\xi^h) = 0.$$

On voit donc que lorsque l'on fait la sommation par rapport à $h$ dans la somme du côté gauche on obtient toujours la disparition de deux termes, de sorte que seulement le terme initial de la suite $\sum_\xi[f'(\xi)f_1(\xi)]$ et le terme final de la suite $\sum_{\xi^r}[f_{r+1}(\xi^r)f'_r(\xi^r)]$ de cette somme restent. On a donc [22] :

$$\sum_{\xi}[f'(\xi)f_1(\xi)] - \sum_{\xi^r}[f_{r+1}(\xi^r)f'_r(\xi^r)] = \sum_{h=1}^{r}\left\{\frac{1}{2}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}[f_{h-1}(x_1)f_h(x_1)]\right\}.$$

Si l'on suppose maintenant connues les racines $\xi^r$ de $f_r(x) = 0$ (il est toujours ici seulement question de racines réelles), alors on peut exécuter la deuxième somme du côté gauche. Mais nous pouvons toujours faire cela, | car ici le degré de $f_r$ baisse selon 106|107

22. *N.d.T.* Il s'agit en fait ici de $\sum_\xi[f'(\xi)f_1(\xi)] + \sum_{\xi^r}[f_{r-1}(\xi^r)f'_r(\xi^r)]$

les $r$ croissants, donc nous pouvons facilement choisir $r$ suffisamment grand pour que $f_r(x) = 0$ soit une équation (par exemple quadratique ou linéaire) dont on connaît les racines. On peut également effectuer le côté droit pour $x_1$ et $x_2$ donnés. Si l'on s'arrange maintenant pour que $f_1(x)$ ait le même signe que $f'(x)$, ce que l'on peut toujours faire, comme nous pouvons disposer arbitrairement de $f_1(x)$ et nous pouvons donc la poser de même degré que $f'(x)$, alors

$$\sum_{\xi}[f'(\xi)f_1(\xi)]$$

représente le nombre de racines réelles de $f(x) = 0$ se trouvant entre $x_1$ et $x_2$. Comme tout le reste dans l'équation est connu, ce nombre est donc maintenant déterminé. C'est le contenu de la proposition de *Sturm*.

Pourtant le théorème de *Sturm* ne permet pas seulement de déterminer le nombre de racines réelles comprises entre $x_1$ et $x_2$, mais il donne aussi une méthode pour calculer numériquement ces racines. Si on laisse en effet $x_1$ et $x_2$ se rapprocher jusqu'à ce qu'il n'y ait qu'une racine comprise entre les deux, alors nous pouvons enfermer celle-ci entre des limites toujours plus étroites $x_1$ et $x_2$, jusqu'à ce que nous les ayons déterminées à la précision voulue.

Si nous arrêtons encore arbitrairement le $f_1$ ci-dessus et que nous supposons ainsi que son signe correspond à $f'$ pour tous les $\xi$ à considérer, alors on aura la condition nécessaire, que $f_1(x)$ n'a pas de facteur commun avec $f(x)$, car sinon $f_1(\xi)$ disparaîtrait pour certains $\xi$. Comme on a :

$$[f'(\xi)f_1(\xi)] = \left[\frac{f'(\xi)f_1(\xi)}{f_1^2(\xi)}\right] = \left[d\frac{f(\xi)}{f_1(\xi)}\right],$$

nous pouvons aussi remplacer la première somme du membre de gauche de notre équation par $\sum_\xi \left[d\frac{f(\xi)}{f_1(\xi)}\right]$. En particulier nous pouvons maintenant poser $r = \nu$ de sorte que nous ayons tous les restes $f, f_1, f_2, ...f_\nu$. Les racines de $f_\nu = 0$ n'existent donc pas, car $f_\nu$ est une constante, de là la deuxième somme du côté gauche de notre équation est supprimée et $\sum_\xi \left[d\frac{f}{f_1}\right]$ est exprimée par la somme du côté droit de l'équation, qui est maintenant à prendre de $h = 1$ à $h = \nu$.

Nous voulons maintenant encore donner à la proposition la version que *Sturm*
107|108 | lui a donnée. $\sum_{h=1}^{\nu}[f_{h-1}(x)f_h(x)]$ est le nombre de fois où dans la suite $f, f_1, f_2, ...f_{h-1}, f_h, f_{h+1}, ...f_\nu$, que l'on nomme la suite de *Sturm*, deux fonctions consécutives $f_{h-1}$ et $f_h$ ont le même signe, moins le nombre de cas où deux telles fonctions ont des signes différents ; ou de façon plus courte, cette somme est égale à la différence du nombre de signes avec le nombre de changements de signes dans la suite de *Sturm*. Si l'on désigne le nombre de signes par $s$, le nombre de changements de signes par $m$, alors on a :

$$\sum_{h=1}^{\nu}[f_{h-1}(x)f_h(x)] = s - m \ , \ s + m = \nu + 1.$$

Si l'on désigne, pour $x = x_1$ resp. $x = x_2$, le nombre de signes et des changements dans la suite de *Sturm* par $s_1$ et $m_1$ resp. $s_2$ et $m_2$, alors il suit :

$$\sum_{h=1}^{\nu}\left\{\frac{1}{2}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}[f_{h-1}(x_1)f_h(x_1)]\right\} = \frac{1}{2}(s_2 - m_2 - (s_1 - m_1)).$$

Or on a $s_2 + m_2 = \nu + 1 = s_1 + m_1$ donc :

$$\frac{1}{2}(s_2 - m_2 - s_1 + m_1) = m_1 - m_2.$$

De là on a : $\sum_{x_1<\xi<x_2}\left[d\frac{f}{f'}\right] = m_1 - m_2$, c'est-à-dire le nombre de racines réelles de $f(x) = 0$ comprises entre les limites $x_1$ et $x_2$ est égal à la perte en changement de signes que la suite de *Sturm* subit lorsque l'on passe de $x = x_1$ à $x = x_2$. C'est la version que *Sturm* a donnée à son célèbre théorème. Nous allons maintenant suivre la preuve que *Sturm* a utilisée pour son théorème. Supposons que la suite de fonctions de *Sturm* $f, f_1, f_2, ...f_r$ soit construite, où nous avons choisi $r$ suffisamment grand pour que $f_r$ n'ait aucune racine réelle (si besoin est, on peut toujours poser $r = \nu$).

Si l'on fait maintenant croître $x$ de $x_1$ à $x_2$ alors, pour $x = x_1$, la suite de *Sturm* présente $m_1$ changements de signes et pour $x = x_2$ elle en présente $m_2$. Il reste à savoir maintenant comment changent les signes des fonctions de *Sturm* lors du passage de $x_1$ à $x_2$.

De $f_{h-1} - g_h f_h + f_{h+1} = 0$ il ressort tout d'abord que deux $f$ quelconques consécutives ne peuvent pas s'annuler pour une même valeur de $x$, car sinon tous les $f$ devraient disparaître et $f$ et $f_1$ auraient alors un facteur commun, ce que nous avons exclu. Comme pour $| \; x = \xi^h$ l'équation devient $f_{h-1}(\xi^h) = -f_{h+1}(\xi^h)$, alors 108|109
$f_{h-1}$ et $f_{h+1}$ ont aussi encore immédiatement avant $\xi^h$ et immédiatement après $\xi^h$ des signes opposés. Les signes de $f_{h-1}f_h$ et $f_{h+1}$ changent donc lors du passage par $\xi^h$ soit de $--+$ à $-++$ (ou de $-++$ à $--+$), soit de $+--$ à $++-$ (ou aussi à l'inverse de $++-$ à $+--$).

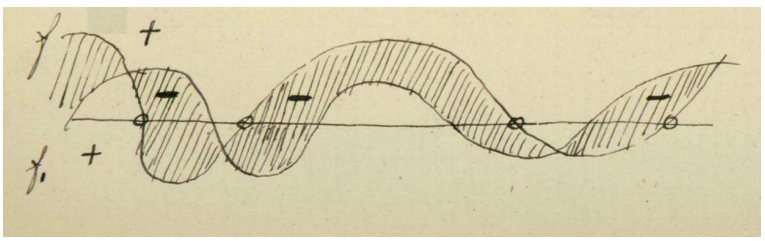
Dans tous les cas nous voyons donc que lors du passage de $x$ par $\xi^h$ un changement de signe dans la suite de *Sturm* n'est ni gagné ni perdu. Cela ne peut donc arriver dans la suite $f, f_1, f_2, ...f_r$, comme $f_r$ reste stable, seulement lors du passage par une racine $\xi$ de $f = 0$. La fabrication d'une suite, dans laquelle le nombre de changements de signes, malgré tous les changements de signes internes, dépend seulement des deux termes initiaux, est l'esprit interne du théorème de *Sturm*. Lors du passage par un $\xi$, $f$ passe soit du négatif au positif, soit du positif au négatif. Dans le premier cas le signe de $f'$ est positif, dans le second cas négatif. Nous avons donc avant et après un $\xi$ soit le passage de $-+$ à $++$ soit celui de $+-$ à $++$ [23], dans tous les cas un changement de signe disparaît. Ainsi est démontrée la proposition affirmant qu'entre $x_1$ et $x_2$ se trouvent autant de racines réelles qu'il y a de pertes de changements de signes lors du passage de $x_1$ à $x_2$, et ce nombre est $m_1 - m_2$.

Comme nous avons seulement besoin des signes dans les fonctions de *Sturm*, il y a naturellement une infinité de suites qui apporte la même chose que la suite $f, f_1, f_2, f_3, ...f_r$ , du moment que les signes de ces suites concordent.

§45. Nous voulons interpréter le côté gauche de l'équation :

$$\sum_{\xi}[f'(\xi)f_1(\xi)] = \frac{1}{2}\sum_{h=1}^{r}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}\sum_{h=1}^{r}[f_{h-1}(x_1)f_h(x_1)]$$

en renouant avec les conceptions géométriques. Nous supposons construites les deux courbes paraboliques $y = f_1(x)$ [et $y = f(x)$].




23. *N.d.T.* Il semble y avoir une erreur : de $+-$ à $--$

Nous pouvons pour chacune des deux courbes différencier dans le plan une partie
109|110 intérieure et | une partie extérieure, suivant que $y - f(x)$ resp. $y - f_1(x)$ est positif ou négatif. Nous pouvons également pour la fonction $(y-f(x))(y-f_1(x))$ distinguer une partie extérieure d'une partie intérieure du plan, suivant que le produit est positif ou négatif. Dans notre dessin la partie hachurée du plan est négative, tout ce qui reste est à considérer comme positif.

Si on avance maintenant sur l'axe des $x$, alors $f'(\xi)f_1(\xi)$ est l'expression de l'augmentation de la fonction :

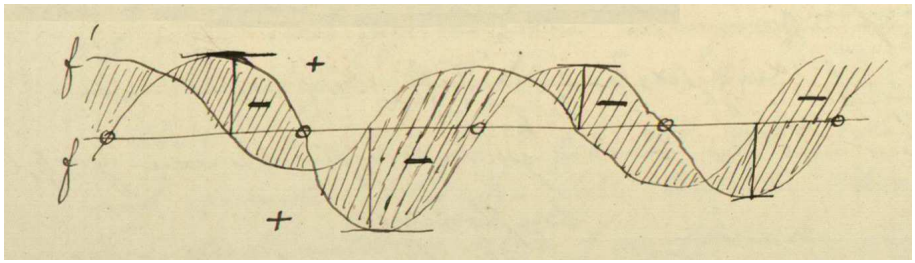
$$F(x,y) = (y - f(x))(y - f_1(x))$$

au point $\xi$ pour lequel $f(\xi) = 0$. Si de là $f'(\xi)f_1(\xi)$ est positif, alors $(F(x,y))_{y=0} = ff_1$ augmente, si l'expression est négative, alors $ff_1$ diminue.

On a de là $[f'(\xi)f_1(\xi)] = +1$ ou $= -1$, selon que nous arrivions lors de la traversée de $\xi$ sur l'axe de l'intérieur à l'extérieur ou de l'intérieur vers l'extérieur.

De là donc $\sum_\xi [f'(\xi)f_1(\xi)]$ est égale au nombre de sorties moins le nombre d'entrées. Cette différence est donc d'un autre côté égale à la perte de changements de signes, c'est-à-dire égale à $m_1 - m_2$.

Dans le cas particulier $f_1 = f'$ où $[f'^2(\xi)]$ est donc toujours positif, il y a alors seulement des sorties, et ce nombre est donc naturellement égal au nombre de racines $\xi$ de $f(\xi) = 0$ se trouvant entre $x_1$ et $x_2$. Le théorème de *Sturm* est donc seulement un cas particulier de la proposition plus générale précédente, que l'on doit à *Sylvester*, qui a découvert une proposition similaire à celle de *Sturm*, en mettant à la place du cas particulier de la suite se formant à partir de $f$ et $f'$ la suite plus générale déduite de la décomposition en fraction continue de $f$ et $f_1$. L'interprétation géométrique pour $f_1 = f'$ du théorème de *Sturm* suit directement de la figure :




110|111 | §46. Nous allons maintenant donner quelques applications du théorème de *Sturm* et nous considérons tout d'abord l'équation trinomiale :

$$f(x) = x^n + nax + b = 0.$$

Nous posons, puisque $n$ est positive, $f_1(x) = \dfrac{1}{n}f'(x) = x^{n-1} + a$. Ensuite, la suite de *Sturm* s'écrit :

$$\begin{aligned} f(x) &= x^n + nax + b \\ f_1(x) &= x^{n-1} + a \\ f_2(x) &= -(n-1)ax - b \\ f_3(x) &= \frac{(-1)^n b^{n-1} - (n-1)^{n-1}a^n}{a^{n-1}} = \frac{\Theta}{a^{n-1}}. \end{aligned}$$

Pour obtenir toutes les racines réelles, nous posons $x_1 = -\infty, x_2 = +\infty$ et nous distinguons si $n$ est pair ou impair.

1. Soit $n$ impair, alors on obtient :
   pour $x = -\infty$ la suite de signes : $-, +, [a], [a^{n-1}\Theta]$
   pour $x = +\infty$ la suite de signes $+, +, -[a], [a^{n-1}\Theta]$
   Donc on a : $\frac{1}{2}\sum_{h=1}^{3}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}\sum_{h=1}^{3}[f_{h-1}(x_1)f_h(x_1)] = 1 - [a] - [a^n\Theta]$

Si maintenant $a > 0$, alors on connaît $-[\Theta]$ à savoir que l'on doit avoir $\Theta < 0$, d'après $f_3$ et le fait que sinon le nombre de racines réelles serait négatif, ce qui n'aurait aucun sens. Pour $a > 0$ le nombre de racines réelles est donc égal à 1. Pour $a < 0$ nous obtenons $2+[\Theta]$ et comme maintenant $\Theta$ peut être aussi bien positif que négatif, alors le nombre de racines réelles est égal à 1 ou à 3. Si l'on résume les deux cas, il suit donc que pour $n$ impair le nombre de racines réelles de $x^n + nax + b = 0$ est égal à 1 ou à 3.

2. Si $n$ est pair, on a pour $x_1 = -\infty, x_2 = +\infty$ :
   $\frac{1}{2}\sum_{h=1}^{3}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}\sum_{h=1}^{3}[f_{h-1}(x_1)f_h(x_1)] = 1 - [a]$ c'est-à-dire que pour $n$ pair, le nombre de racines réelles de l'équation trinomiale est égal à 0 ou à 2. Si l'on désigne le nombre de racines réelles de l'équation trinomiale par $\mathcal{A}$, alors on a pour $n$ pair et impair : $\mathcal{A} = \frac{3+(-1)^{n-1}}{2} + (-1)^{n-1}[\Theta]$.

---

| §47. Lorsque nous posons $x_1 = -\infty$ et $x_2 = +\infty$, alors nous pouvons représenter 111|112
le théorème étendu de *Sturm* (pag...) par une forme simple. Si l'on désigne le nombre d'entrées et de sorties par $E$ resp. $A$, alors on a :

$$A - E = \frac{1}{2}\sum_{h=1}^{r}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}\sum_{h=1}^{r}[f_{h-1}(x_1)f_h(x_1)]$$

Soit maintenant

$$f(x) = cx^n + ...$$
$$f_k(x) = c_k x^{n_k}$$

Comme on pose $x_1 = -\infty$ et $x_2 = +\infty$, alors nous avons seulement besoin de prendre en compte la plus haute puissance. Pour $x = -\infty$ nous obtenons ainsi la suite :

$$(-1)^n c, (-1)^{n_1} c_1, (-1)^{n_2} c_2, ..., (-1)^{n_k} c_k, ...$$

Pour $x = +\infty$ il vient :

$$c, c_1, c_2, ..., c_k, ...$$

La somme ci-dessus est donc à étendre aux membres de la forme :

$$\frac{1}{2}[c_{k-1}c_k] - \frac{1}{2}(-1)^{n_{k-1}+n_k}[c_{k-1}c_k].$$

Maintenant tous les membres pour lesquels $n_{k-1} + n_k$ est pair sont supprimés et on obtient donc :

$$A - E = \sum[c_{k-1}c_k].$$

La somme s'étend à tous les $k$ pour lesquels on a $n_{k-1} + n_k \equiv 1 \bmod. 2$.

La question des racines se trouvant entre $x_1$ et $x_2$ n'est absolument pas plus particulière que la question des racines se trouvant entre $-\infty$ et $+\infty$, puisque nous pouvons toujours représenter la distance de $x_1$ à $x_2$ par la droite complète de $-\infty$ à $+\infty$. Si nous remplaçons en effet dans $f(x)$ le $x$ par la fonction $\frac{x_1 + x_2 z^2}{1 + z^2}$ et que nous enlevons le dénominateur, alors on obtient une fonction $F(z)$. Celle-ci s'annule pour tout $\xi$ qui correspond à un zéro $\xi$ de $f(x)$. Si l'on promène $z$ de $-\infty$ à $+\infty$, alors $x$ décrit le chemin de $x_2$ à $x_1$ et revient ensuite à $x_2$. Le nombre de racines réelles de $F(z) = 0$ est donc égal au double du nombre de racines de $f(x) = 0$ se trouvant entre $x_1$ et $x_2$. Suite à cette possibilité de parcours sur la droite entière, il est indifférent de nous poser la question de toutes les racines réelles de $F(z) = 0$ ou de celles de $f(x) = 0$ comprises entre $x_1$ et $x_2$. Les deux questions sont, si on passe, en vertu de la substitution donnée, de $f(x)$ à $F(z)$ ou l'inverse, complètement équivalentes.

| §48. Le théorème de *Sturm* nous met dans les mains un moyen simple de 112|113
rechercher la réalité des racines des équations de tout degré à coefficients réels.

L'étude des équations linéaires et quadratiques est tellement triviale que l'on considère donc en premier les équations cubiques. Celles-ci ont déjà été étudiées en détail et nous avons vu que les deux seuls cas possibles (3 racines réelles ou seulement une seule) se distinguent par le signe de leur discriminant (voir pag....). Nous nous tournons de là maintenant vers l'équation biquadratique à coefficients réels.

Elle peut toujours être rapportée à la forme :

$$f(u) = u^4 - 6zu^2 + 4yu + 3x.$$

Nous considérons ensuite $x, y, z$ comme les coordonnées d'un point dans l'espace et nous pouvons ainsi affecter clairement à chaque point une équation complètement déterminée et inversement. Nous devons ensuite obtenir 3 régions séparées, de sorte que pour chaque point du premier domaine l'équation biquadratique correspondante ait 4 racines réelles, pour chaque point du second domaine l'éq. correspondante ait 2 racines réelles, et pour chaque point du troisième domaine l'éq. correspondante n'ait aucune racine réelle.

Ces trois domaines sont séparés par des surfaces et comme le passage d'une paire de racines imaginaires à une paire de racines réelles ne peut avoir lieu seulement que par une paire de racines égales, on s'aperçoit donc que tous les points du bord de ces surfaces doivent correspondre à des équations avec des racines égales, et inversement que toute équation avec des racines égales correspond à un point qui se trouve au bord de la surface. Tous les points du bord de la surface correspondent donc à des équations avec un discriminant qui s'annule. Celui-ci est une fonction entière à coefficients rationnels des coordonnées $x, y, z$ ; Si on le pose nul, alors on obtient donc une surface, laquelle est ainsi la surface-discriminante, laquelle doit séparer uniquement les 3 domaines de l'espace cités. Nous étudions maintenant à l'aide du théorème de
113|114 *Sturm* | ces 3 domaines et la surface-discriminante. On a :

$$f(u) = u^4 - 6zu^2 + 4yu + 3x.$$

Nous choisissons : $f_1 = \frac{1}{4}f' = u^3 - 3zu + y.$
Si nous omettons toujours les facteurs positifs dans les restes de fonctions suivants, alors nous obtenons la suite de *Sturm* suivante :

$$\begin{aligned} f &= u^4 - 6zu^2 + 4yu + 3x \\ f_1 &= u^3 - 3zu + y. \\ f_2 &= zu^2 - yu - x \\ f_3 &= D_1 u - y(x + z^2) \\ f_4 &= D. \end{aligned}$$

En même temps on a $D = -y^4 - 2y^2z(3x - z^2) + x(x - 3z^2)^2 = (x + z^2)^3 - (3xz + y^2 - z^3)^2 = 0$
De plus :

$$D_1 = 3z^3 - xz - y^3.$$

D'après la page.... le nombre de racines réelles est égal à :

$$\sum [c_{k-1}c_k] = 1 + [z] + [zD_1] + [D_1D].$$

Nous renonçons tout d'abord au cas $D = 0$, car nous aurions à faire dans ce cas avec seulement une équation biquadratique impropre.

1. Si $D < 0$, alors :

$$1 + [Z] + [ZD_1] + [D_1D] = 1 + [z] + [zD_1] - [D_1] = 1 + [z] - [D_1]\{1 - [z]\}.$$

Si maintenant $z > 0$, alors ce nombre est égal à 2, si $z < 0$, alors le nombre est également égal à 2, car dans ce dernier cas $D_1$ doit également nécessairement être négatif, car sinon le nombre de racines réelles serait un nombre négatif. Que $D_1 < 0$ lorsque $D < 0$ et $z < 0$, cela résulte aussi naturellement de l'expression de $D$ et $D_1$. Si donc $D < 0$, alors l'équation biquadratique a toujours deux racines réelles et une paire de racines imaginaires.

2. Si $D > 0$, alors :

$$1 + [Z] + [ZD_1] + [D_1 D] = \{1 + [z]\}\{1 + [D_1]\}.$$

Ce nombre représente 0 ou 4.

Si $z > 0$ et $D_1 > 0$, alors l'éq. biquadratique a 4 racines réelles, | si par contre 114|115
$z < 0$ ou $D_1 < 0$, alors l'éq. n'a aucune racine réelle. Si l'on désigne maintenant les 3 domaines de l'espace, lesquels correspondent aux équations avec aucune, une ou deux paires de racines imaginaires, par $G_0, G_1, G_2$, on a donc :

$G_0$ caractérisé par $D > 0$ et soit $z < 0$ soit $D_1 < 0$.
$G_1$ caractérisé par $D < 0$.
$G_2$ caractérisé par $D > 0$ et $z > 0$ et $D_1 > 0$.

Lorsque nous arrivons à la surface $D = 0$, alors nous passons de $G_1$ à $G_0$ resp. $G_2$ ou inversement de $G_0$ resp. $G_2$ au domaine $G_1$. Nous désignons $G_1$ comme l'espace extérieur, $G_0$ et $G_2$ comme les deux espaces intérieurs. Le discriminant $D$ est le résultat de l'élimination de $u$ dans $f(u)$ et $f'(u)$. Mais $f(u) = 0$ et $f'(u) = 0$ sont à concevoir comme un plan avec le paramètre $u$ et un plan voisin que l'on obtient en faisant varier $u$ infiniment peu. $f(u) = 0$ et $f'(u) = 0$ sont donc les équations de l'intersection de ces deux plans, et $D = 0$ est donc le lieu de cette intersection et par conséquent une surface développable. Sa courbe cuspidale a pour équations :

$$x + z^2 = 0 \text{ et } y^2 - 4z^3 = 0.$$

De plus elle possède une courbe double avec les équations :

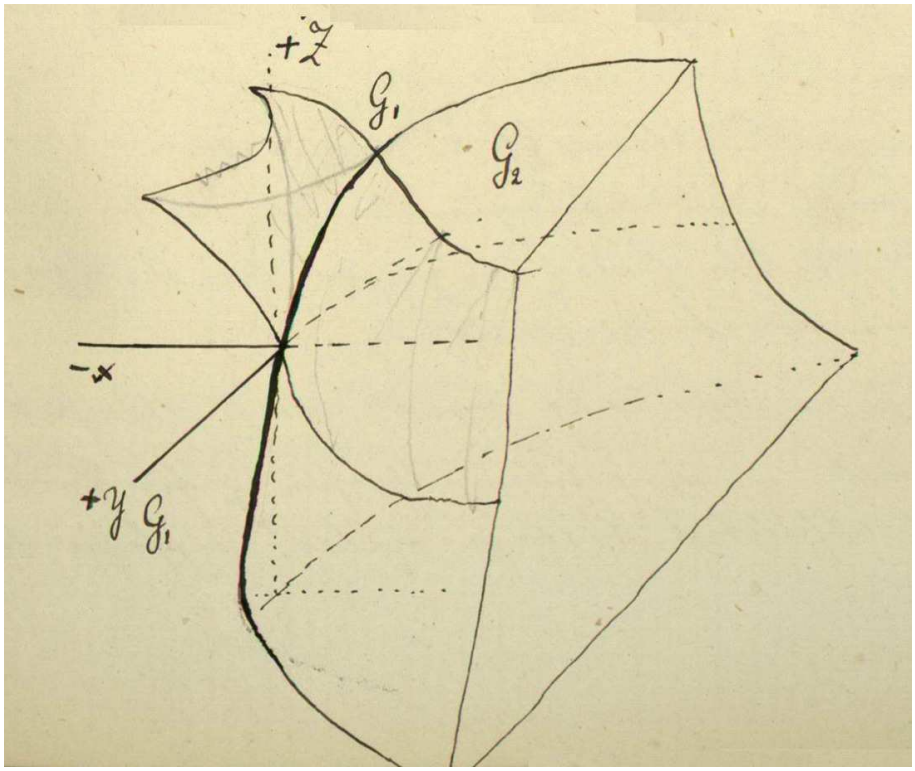
$$y = 0 \text{ et } x - 3z^2 = 0.$$

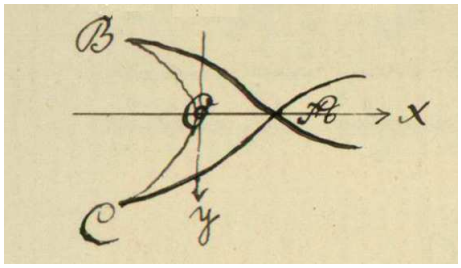
La courbe cuspidale et la courbe double forment l'intersection complète des deux surfaces $D = 0$ et $D_1 = 0$. La surface $D = 0$ sépare donc le domaine extérieur $G_1$ des deux domaines intérieurs $G_0$ et $G_2$. La surface-discriminante a à peu près la forme suivante : 115|116

| Au dessous du plan $x - z$, la surface reste complètement sur le côté positif du plan $y - z$ et se moule toujours elle-même étroitement en s'ouvrant encore et encore selon les $Z$ négatifs croissants.

Au dessus du plan $x - y$, une intersection horizontale est montrée approximativement par la forme de la figure ci-contre.

En même temps $O$ représente la partie positive de l'axe des $z$, $\mathcal{A}$ représente une parabole, à savoir la courbe double se trouvant dans le plan $x - z$ et finalement $\mathcal{B}$ et $\mathcal{C}$ représentent les courbes d'entrée. Dans le plan $x - y$ le sommet de la parabole coïncide avec les deux pointes $\mathcal{B}$ et $\mathcal{C}$. Seulement la partie de la parabole se trouvant au dessus de ce plan fait partie de la surface-discriminante, qui passe sous la partie isolée dans l'espace $G_2$, mais est à considérer comme appartenant à la surface-discriminante. À la partie du dessus de la parabole correspondent, comme $z > 0$, les 4 racines réelles $+\sqrt{3z}, +\sqrt{3z}, -\sqrt{3z}, -\sqrt{3z}$, où la racine quadratique est toujours à prendre avec le même signe, tandis qu' à la partie du dessous, où $z < 0$, correspondent les 4 racines imaginaires :
$+\sqrt{3z}, +\sqrt{3z}, -\sqrt{3z}, -\sqrt{3z}$. La courbe double (qui est en effet pour moitié une courbe isolée) correspond donc aux équations avec 2 paires de racines égales. La courbe d'entrée est le lieu des points auxquels correspondent les équations avec 3 racines




égales.

§49. Nous voulons maintenant en venir à justifier après coup rigoureusement la notion dont nous avons fait usage dans le théorème de *Sturm*. Il s'agit d'abord de fixer la notion de racine d'une équation.

Nous prenons une fonction entière $f(x)$ de $x$ de degré $m-1$, dont les coefficients sont réels et définis d'une façon quelconque, puis nous nous demandons : quelle signification a l'expression "une grandeur déterminée $x$ satisfait l'équation" ?

Lorsque les coefficients réels sont définis de façon quelconque, alors nous devons pouvoir trouver en tout cas une grandeur positive $G$, laquelle est plus grande que
116|117 les valeurs absolues de tous les coefficients. | Si l'on désigne la valeur absolue d'une grandeur $a$ par $|a|$, alors pour n'importe quel $x$ :

$$\left\{|x|^{m-1}+\ldots+|x|+1\right\}G>|f(x)|$$

Si l'on désigne par $\delta$ une fraction positive inférieure à 1, alors $f(x)$ est donc elle-même égale à :

$$f(x)=\delta G\frac{|x|^m-1}{|x|-1}.$$

Nous pouvons encore généraliser cela. Si un certain nombre de grandeurs se trouve

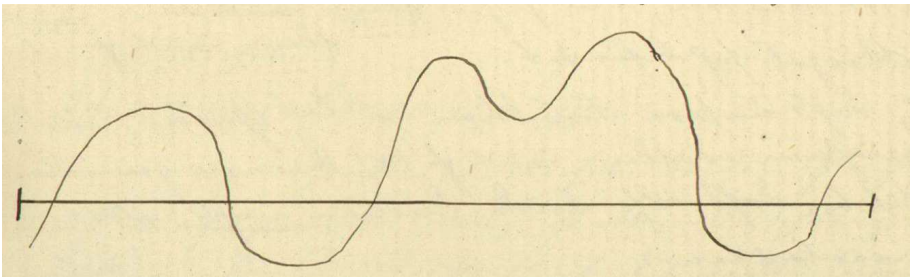
entre resp. $\pm a', \pm a'', ...$, alors la somme se trouve elle-même entre $\pm(a' + a'' + ...)$. C'est-à-dire, en entendant toujours le symbole $\delta$ par une fraction :

$$\delta' a' + \delta'' a'' + ... = \delta(a' + a'' + ...)$$

Les $a', a'', ...$ sont des grandeurs réelles positives. Si l'on considère maintenant toutes les valeurs de fonctions par lesquelles on a $|x| < \xi$, alors il s'avère que la somme de ces valeurs de fonction est égale à :

$$\Sigma = \delta G \frac{\xi^m - 1}{\xi - 1}.$$

Cela étant indiqué, nous allons maintenant délimiter un tel intervalle en dehors duquel on ne peut trouver aucune racine supplémentaire. Nous nous servons aussi ici de l'interprétation géométrique, bien que nous sachions que cela suppose déjà la proposition à démontrer. Mais nous voulons par là aussi seulement renforcer cette conception, sans pour autant l'appuyer d'une preuve.



Il s'agit donc de délimiter un intervalle de sorte que en dehors de celui-ci l'axe des abscisses ne sera plus coupé par la courbe.

Nous supposons maintenant notre fonction de degré $m$, et nous voulons donc trouver un intervalle pour lequel $x^m + ....$ ait le signe de $x^m$. Les termes qui suivent $x^m$ forment une fonction de degré $m-1$, comme nous l'avons considéré plus haut, pour laquelle il existe donc une grandeur positive $G$ telle que la valeur de cette fonction entière de degré $m-1$ soit plus petite que $G\frac{|x|^m - 1}{|x| - 1}$.

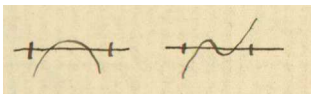
Si l'on choisit $|x| > G + 1$, en posant par exemple $|x| = G + 1 + p$, | où $p$ est une 117|118
grandeur positive, alors la valeur de la fonction de degré $m-1$ qui suit $x^m$ est plus petite que :

$$\frac{G}{G+p}\left((G + 1 + p)^m - 1\right)$$

donc plus petite que la valeur de $x^m$, car celle-ci est égale à $(G + 1 + p)^m$.

Si donc $x^m + ... = 0$ est donnée, alors on trouve toujours une telle grandeur $G$, à condition que la fonction $x^m + ...$ait le même signe que $x^m$ dès que $|x| > G + 1$. Nous avons ainsi maintenant gagné l'avantage de pouvoir limiter nos considérations à l'intervalle fini :$-(G + 1)... + (G + 1)$.

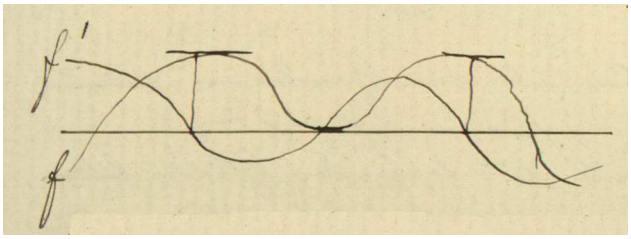
Comme la possibilité de trouver un tel intervalle est maintenant montrée, il s'agit de trouver un plus petit, dans lequel ne se trouvera qu'une seule racine, et d'isoler ensuite un tel point. Cela peut éventuellement être très délicat, comme le montre le dessin ci-contre.



Nous devons donc pour en sortir déterminer d'emblée hors des conditions de la fonction les limites pour un petit intervalle dans lequel se trouve seulement une ou aucune racine. Attaquons-nous maintenant à cette étude, c'est-à-dire cherchons

comment les limites d'un intervalle doivent être déplacées près l'une de l'autre pour être certain qu'une seule ou aucune racine ne s'y trouve.

Nous supposons que notre fonction de degré $m$, que nous avons déjà désignée par $f(x)$, n'a aucun facteur commun avec $f'(x)$, que donc les deux courbes ne se présentent pas de la manière suivante :



Cette hypothèse peut toujours être réalisée par la séparation d'un tel facteur commun. On peut ainsi déterminer 2 fonctions $\varphi$ et $\psi$ telles que $f\varphi - f'\psi = 1$. Nous voyons donc que notre méthode se limite seulement aux équations algébriques, et ne traite pas des équations transcendantes. Si nous
118|119 formons maintenant les expressions suivantes : |

$$\frac{f(x[_1]) - f(x_2)}{x_1 - x_2} = f'(x_1) + \frac{x_1 - x_2}{2} f''(x_1) + \text{... et}$$

$\left\{\frac{f(x[_1]) - f(x_2)}{x_1 - x_2} - f'(x_1)\right\} \frac{2}{x_1 - x_2}$, alors, lorsque l'on se limite à l'intervalle de $-(G+1)$ à $+(G+1)$ défini ci-dessus, les deux expressions restent entre les limites finies connues, comme nous voulons aussi supposer $x_1$ et $x_2$ proches l'un de l'autre. La même chose est valable pour les fonctions entières $\varphi$ et $\psi$. Si nous désignons maintenant les limites de $\varphi$ et $\psi$ par resp. $\mathcal{A}$ et $\mathcal{B}$ et celles des expressions ci-dessus avec resp. $\mathcal{C}$ et $\mathcal{D}$ et que nous symbolisons par les lettres grecques minuscules $\alpha, \beta, \gamma, \delta$ des fractions positives ou négatives inférieures à 1, de sorte que par exemple $\varphi = \alpha\mathcal{A}, \psi = \beta\mathcal{B}$, et si nous posons enfin $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, où $\Theta$ est une fraction inférieure à 1 et $\mathcal{M}$ est encore indéterminée, alors $\mathcal{M}$ est à déterminer de sorte que dans l'intervalle $x_1 - x_2$ se trouve soit une, soit aucune racine de $f(x) = 0$.

Maintenant on a :

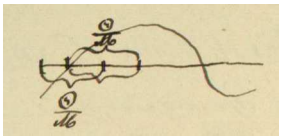
$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(x_1) + \frac{x_1 - x_2}{2} \delta\mathcal{D}.$$

Si on le multiplie par $\psi$ et que l'on prend en compte les équations : $f\varphi - f'\psi = 1$, $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, $\varphi = \alpha\mathcal{A}$, $\psi = \beta\mathcal{B}$, alors il vient :

1. $-(f(x_1) - f(x_2))\frac{\beta\mathcal{B}\mathcal{M}}{\Theta} + f(x_1)\alpha\mathcal{A} = 1 - \frac{\Theta}{2}\delta\beta\frac{\mathcal{D}\mathcal{B}}{\mathcal{M}}$. De plus on a :
2. $f(x_1) - f(x_2) = \Theta\gamma\frac{\mathcal{C}}{\mathcal{M}}$.

La seule valeur dans ces deux équations dont nous pouvons disposer est $\mathcal{M}$. Si nous choisissons donc ensuite $\mathcal{M}$ plus grand que $\mathcal{D}\mathcal{B}$, alors le côté droit de la première équation se trouve toujours entre $\frac{1}{2}$ et $\frac{3}{2}$, et il est donc toujours plus grand que $\frac{1}{2}$.

Si l'on considère donc un tel intervalle $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, où $\Theta$ est différent de 0, alors, lorsque $\mathcal{M} > \mathcal{B}\mathcal{D}$, $f(x_2)$ et $f(x_1)$ ne peuvent pas être en même temps aussi petites que l'on veut, si l'on suppose que $x_1 - x_2$ a une grandeur déterminée. On peut de là tout à fait déplacer l'intervalle $x_1 - x_2$ (sans changer sa taille) à volonté sur l'axe
119|120 des $x$, mais ensuite $f(x_2)$ devra toujours rester au-dessus d'une | certaine grandeur lorsque $f(x_1)$ deviendra aussi petite que l'on veut. Mais cela caractérise le fait que dans chaque intervalle $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$ défini ainsi, il y ait au maximum un endroit où $f(x)$ s'annule.



Notre deuxième équation nous apprend maintenant que l'intervalle $\frac{\Theta}{\mathcal{M}}$ peut être pris suffisamment petit pour que, en supposant $f(x_1)$ et $f(x_2)$ de signes opposés, les valeurs absolues de $f(x_1)$ et $f(x_2)$ puissent devenir aussi petites que l'on veut. On obtient donc le résultat suivant, qui établit la notion de racine d'une équation : soit $\Theta$ une fraction quelconque, de sorte que l'on puisse toujours choisir $\mathcal{M}$ suffisamment grand pour que dans chaque intervalle $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$, la fonction $f(x)$ soit arbitrairement petite et que, à supposer que dans l'intervalle (c'est-à-dire donc aux extrémités elles-même) un changement de signe ait lieu, la valeur de la fonction peut être faite aussi petite que l'on veut aux extrémités.

Ainsi, par exemple, l'équation $x^2 - 2 = 0$ n'a pas d'autre signification que : il y a un intervalle rationnel, sur les extrémités duquel $x^2 - 2$ peut être rendue positive et aussi petite que l'on veut, et d'un autre côté négative et aussi petite que l'on veut.

Le sens du fait qu'une équation $f(x) = 0$ soit satisfaite par une certaine grandeur $x$ est à expliquer ainsi : cette grandeur $x$ est à concevoir comme une limite que l'on approche lorsque l'on rend l'intervalle précédent toujours plus petit. Le théorème de *Sturm* ne détermine par là rien d'autre que le nombre des intervalles caractéristiques précédents $x_1 - x_2 = \frac{\Theta}{\mathcal{M}}$ dans lesquels a lieu un changement de signe.

Si, non pas pour démontrer, mais seulement pour la facilité de la présentation, nous nous servons de la conception géométrique et nous notons $f(x) = y$ comme une courbe, alors cela suppose déjà l'existence des racines. Nous pouvons alors d'abord construire seulement des grandeurs rationnelles. Pour relier les points se constituant de cette façon en une courbe, on a besoin de supposer que un $x$ existe, lequel correspond à un $f(x)$ se trouvant entre $f(x_1)$ et $f(x_2)$. Donc dessiner une courbe implique déjà la supposition qu'une équation a une racine.

120|121

| §50. Nous en venons à étudier le sens des fonctions de *Sturm* comme fonction symétrique.

Nous renouons avec le développement en fraction continue et nous allons d'abord déduire plusieurs propositions sur les fractions continues. Soit :

$$\frac{f_1}{f} = \frac{1}{g_1 - \frac{1}{g_2 - \ldots - \frac{1}{g_\nu}}}$$

Nous posons maintenant :

$$\frac{1}{g_1 - \frac{1}{g_2 - \ldots - \frac{1}{g_{k-1}}}} = \frac{\varphi_k}{\psi_k}$$

On a fait disparaître un éventuel facteur commun entre $\varphi_k$ et $\psi_k$. $\varphi_k$ et $\psi_k$ sont des fonctions entières de $g$, la dimension de $\varphi_k$ sera déterminée par $g_2 g_3 \ldots g_{k-1}$, celle de $\psi_k$ par $g_1 g_2 \ldots g_{k-1}$.

Nous allons maintenant démontrer que les équations vérifient :

$$\varphi_{k+1} - g_k \varphi_k + \varphi_{k-1} = 0 \text{ et de même :}$$
$$\psi_{k+1} - g_k \psi_k + \psi_{k-1} = 0.$$

Nous démontrons cela par induction. Nous obtenons $\frac{\varphi_{k+1}}{\psi_{k+1}}$ en remplaçant $g_{k-1}$ par $g_{k-1} - \frac{1}{g_k}$ dans $\frac{\varphi_k}{\psi_k}$. On a donc :

$$\frac{\varphi_{k+1}}{\psi_{k+1}} = \frac{(g_{k-1} - \frac{1}{g_k})\varphi_{k-1} - \varphi_{k-2}}{(g_{k-1} - \frac{1}{g_k})\psi_{k-1} - \psi_{k-2}} = \frac{g_k \varphi_k - \varphi_{k-1}}{g_k \psi_k - \psi_{k-1}},$$

en supposant les égalités :

$$\varphi_k - g_{k-1}\varphi_{k-1} + \varphi_{k-2} = 0 \text{ et } \psi_k - g_{k-1}\psi_{k-1} + \psi_{k-2}[= 0].$$

Comme maintenant dans $\frac{g_k\varphi_k-\varphi_{k-1}}{g_k\psi_k-\psi_{k-1}}$ numérateur et dénominateur n'ont aucun facteur commun (ni facteur numérique commun), alors :

$$\varphi_{k+1} = g_k\varphi_k - \varphi_{k-1}$$
$$\psi_{k+1} = g_k\psi_k - \psi_{k-1}.$$

Les formules de récurrence à partir de laquelle on obtient les fonctions $\varphi$ et $\psi$ sont donc les mêmes pour $\varphi$ et $\psi$. La différence sera due seulement aux valeurs initiales, car on a :

$$\varphi_1 = 0, \varphi_2 = 1 \text{ et } \psi_1 = 1, \psi_2 = g_1$$

121|122 | Mais nous pouvons représenter le numérateur et le dénominateur de cette réduite par la même fonction, lorsque nous concevons un morceau quelconque de la fraction continue sortie du milieu. Si nous posons en effet :

$$\frac{1}{g_p - \frac{1}{g_{p+1}-\ldots-\frac{1}{g_{q-1}}}} = \frac{Z_{p,q}}{\mathcal{N}},$$

alors on aura réussi à exprimer également le dénominateur par un $Z$, lequel est construit d'après la même loi que le numérateur. On a en effet :

$$\frac{1}{g_{p-1} - \frac{Z_{p,q}}{\mathcal{N}}} = \frac{Z_{p-1,q}}{\mathcal{N}_1} = \frac{\mathcal{N}}{\mathcal{N}g_{p-1} - Z_{p,q}}$$

Comme nous avons déjà supposé plus haut que $Z_{pq}$ et $\mathcal{N}$ n'ont aucun facteur commun, alors il suit :

$$\mathcal{N} = Z_{p-1,q}.$$

Donc on a :

$$\frac{1}{g_p - \frac{1}{g_{p+1}-\ldots-\frac{1}{g_{q-1}}}} = \frac{Z_{p,q}}{Z_{p-1,q}}.$$

Maintenant nous avons eu *pag* ... une formule de récurrence pour $\varphi$ et $\psi$, que nous pouvons aussi traduire dans $Z$ en posant $q = k$ et en considérant $g_p$ comme le premier dénominateur. Ensuite l'équation en $Z$ s'énonce :

I) $Z_{p,q+1} - g_qZ_{pq} + Z_{p,q-1} = 0$

Nous voulons maintenant faire dériver une équation générale entre les $Z$, qui comprenne toutes les propositions sur les fractions continues. Celle-là même s'énonce :

II) Soit $Z_{pq}Z_{rs} - Z_{pr}Z_{qs} + Z_{ps}Z_{qr} = 0$ où $p \leq q \leq r \leq s$.

Nous démontrons d'abord un cas particulier de cette équation et de là ensuite l'éq. générale.

Nous posons en effet $p = 0, q = 1, r = r, s = s$. Maintenant on a $Z_{pp} = 0, Z_{p,p+1} = 1$. Nous obtenons donc d'abord l'éq. particulière : $Z_{01}Z_{rs} - Z_{0r}Z_{1s} + Z_{0s}Z_{1r} = 0$, mais comme $Z_{01} = 1$, alors il suit $Z_{rs} = Z_{0r}Z_{1s} - Z_{0s}Z_{1r}$. Nous voulons maintenant
122|123 démontrer cette éq. particulière par induction. Supposons qu'elle s'applique pour $s$, | et démontrons-la pour $s + 1$, c'est-à-dire montrons la justesse de l'éq. :

$$Z_{r,s+1} = Z_{0r}Z_{1,s+1} - Z_{0,s+1}Z_{1r}.$$

Nous pouvons maintenant enlever $Z_{1s+1}$ et $Z_{0s+1}$ avec l'aide de l'éq. I. pag... et nous obtenons :

$$Z_{r,s+1} = Z_{0r}(g_sZ_{1s} - Z_{1,s-1}) - Z_{1r}(g_sZ_{0s} - Z_{0,s-1})$$
$$Z_{r,s+1} = g_s(Z_{0r}Z_{1s} - Z_{1r}Z_{0s}) - (Z_{0r}Z_{1,s-1} - Z_{1r}Z_{0,s-1})$$

Avec la prise en considération de l'hypothèse il vient :

$$Z_{r,s+1} = g_s Z_{rs} - Z_{r,s-1}.$$

Comme la véracité de cette éq. est prouvée (Eq. I.), alors il en est donc ainsi de même pour l'éq. : $Z_{rs} = Z_{0r}Z_{1s} - Z_{0s}Z_{1r}$. Avec l'aide de cette éq., l'éq. II. (pag...) se laisse ensuite aisément démontrer, en exprimant facilement en vertu de l'éq. pour $Z_{rs}$ toutes les 6 grandeurs $Z_{rs}$, $Z_{p,q}$, ... ; on obtient ensuite l'éq. $0 = 0$, ce qui explique que la véracité de la relation générale II. soit démontrée.

Si l'éq. I. nous livre une relation entre 3 $Z$ se suivant avec le même premier indice, alors nous pouvons confronter celle-ci avec l'aide de l'éq. II. à une autre où le deuxième indice est le même et le premier différent. Si nous posons en effet dans l'éq. II. $p = q - 1, r = q + 1$ et si on prend compte le fait que :

$Z_{q-1,q} = 1, Z_{q-1,q+1} = g_q, Z_{q,q+1} = 1$, alors nous obtenons la relation voulue, que nous voulons combiner avec l'éq. I. :

$$\begin{cases} Z_{p,q+1} - g_q Z_{p,q} + Z_{p,q-1} = 0. \\ Z_{q+1,s} - g_q Z_{q,s} + Z_{q-1,s} = 0. \end{cases}$$

Appliquons maintenant ce résultat à la décomposition en fraction continue, résultat qui est à la base du théorème de *Sturm*.

On s'aperçoit que cette décomposition en fraction continue est $Z_{k,\nu+1} = \frac{f_k}{f_\nu}$, car pour $\frac{f_k}{f_\nu}$ convient la même formule de récurrence que pour $Z_{k,\nu+1}$, en effet l'éq. :

$$\frac{f_{k+1}}{f_\nu} - g_k \frac{f_k}{f_\nu} + \frac{f_{k-1}}{f_\nu} = 0$$

| et les valeurs initiales sont les mêmes, à savoir : 123|124

$$Z_{\nu,\nu+1} = 1 = \frac{f_\nu}{f_\nu}, Z_{\nu-1,\nu+1} = g_\nu = \frac{f_\nu - 1}{f_\nu}.$$

Si nous posons maintenant $q = p + 1, s = r + 1$, alors il suit de II. :

$$1 = Z_{pr}Z_{p+1,r+1} - Z_{p,r+1}Z_{p+1,r}.$$

Pour $p = 0$ on a $Z_{0k} = \psi_k, Z_{1,k} = \varphi_k$ (en posant $r = k$)
et on obtient :

1. $\varphi_{k+1}\psi_k - \varphi_k\psi_{k+1} = 1.$

Une seconde relation simple suit, lorsque l'on pose
$p = 0, q = 1, r = k, s = \nu + 1$, ensuite il vient :
$Z_{k,\nu+1} = Z_{0k}Z_{1,\nu+1} - Z_{0,\nu+1}Z_{1,k}$.
Ou :

2. $\psi_k f_1 - \varphi_k f_0 = f_k.$

On obtient finalement encore une troisième formule en posant :

$$p = 0, q = k, r = k + 1, s = \nu + 1$$

en effet : $Z_{0,\nu+1} = Z_{0,k+1}Z_{k,\nu+1} - Z_{0k}Z_{k+1,\nu+1}$.
Ou :

3. $f_0 = \psi_{k+1}f_k - \psi_k f_{k+1}$

Dans les éq. 2) et 3) nous avons de nouveau le résultat précédent, *pag...*, cette fois dérivé de la relation entre fractions continues II. *pag....*, à savoir que chaque $f_k$ est une fonction linéaire homogène des deux premières fonctions $f_0$ et $f_1$, mais que $f_0$ aussi est une fonction linéaire homogène d'une paire quelconque de fonctions consécutives $f_k$ et $f_{k+1}$. Mais nous apprenons ici en outre que dans les deux cas les multiplicateurs qui

apparaissent sont le numérateur resp. le dénominateur des réduites du développement de *Sturm*.

Nous recherchons encore quels sont les degrés des fonctions apparaissant dans le développement en fraction continue. Nous désignons le degré de $f_k$ par $n_k$, et nous construisons ainsi la suite décroissante $n_1, n_2 ... n_k, ... n_\nu$, où $n_\nu = 0$. Dans le cas prétendument régulier, la suite $n-1, n-2, ...1, 0$ est telle qu'alors $n_k = n-k$ et $\nu = n$. De $f_{k-1} - g_k f_k + f_{k+1} = 0$ il suit que le degré de $g_k$ doit être égal à $n_{k-1} - n_k$ et
124|125 celui de $f_{k+1}$ plus bas que celui de $f_{k-1}$. | Nous obtenons le degré de $\psi_k$ à partir de : $\psi_{k-1} - g_k \psi_k + \psi_{k-1} = 0$. Le degré de $\psi$ est ainsi égal à $n = n_{k-1}$, car nous supposons que $\psi_k$ a ce degré, d'où il suit, comme $\phi_{k+1}$ est d'un degré plus haut que $\psi_{k-1}$, que $\psi_{k+1}$ a le degré : $n_{k-1} - n_k + n - n_{k-1} = n - n_k$, de sorte que nous obtenons en conclusion de l'induction que $n - n_{k-1}$ est bien le degré de $\psi_k$. Le degré de $\psi_k$ est seulement le degré de $g_1$, c'est-à-dire $n - n_1$ plus petit que celui de $\psi_k$, donc égal à $n_1 - n_{k-1}$, comme on le vérifie facilement par induction.

§51. Après ces propositions préliminaires sur les fractions continues et le développement des fonctions de *Sturm*, nous envisageons un problème qui est en étroit rapport avec la représentation des fonctions de *Sturm* comme fonctions symétriques, à savoir le problème de la formule d'interpolation de *Cauchy*. Autrement dit : le quotient de deux fonctions rationnelles entières $\dfrac{F(x)}{\Psi(x)}$ doit être déterminé de sorte qu'il prenne pour $x = \xi_h$ $(h = 1, ...n)$ la valeur $\mathcal{C}_h$ $(h = 1, ...n)$. Soit $F(x)$ de degré $\mu$, $\Psi(x)$ de degré $\nu$, alors on doit avoir :

$$\frac{\sum_{r=0}^{\mu} a_r \xi_h^r}{\sum_{s=0}^{\nu} b_s \xi_h^s} = \mathcal{C}_h \ (h = 1, ...n)$$

Les $\mu + \nu + 2$ coefficients inconnus $a_0, ...a_\mu, b_0 ... b_\nu$ doivent être déterminés de sorte que les $n$ équations soient satisfaites. Nous avons donc à faire avec un problème de résolution d'équations linéaires homogènes et nous savons que le nombre de ces dernières, dans le cas où ce problème peut être résolu, doit être au moins égal à $\mu+\nu+1$. Comme première condition pour $\mu$ et $\nu$, nous obtenons ainsi que $\mu + \nu + 1 \leq n$ ou $\mu + \nu < n$.

Si le dénominateur est de degré 0, donc $\nu = 0$, alors nous avons déjà résolu le problème par la formule d'interpolation de *Lagrange*. Nous savons maintenant, lorsque l'on considère le problème de *Cauchy* comme un problème de la théorie des équations linéaires, que la résolution à partir d'une variété de plus haut degré peut être éventuellement telle que justement la méthode la plus évidente puisse peut-être
125|126 compliquer beaucoup le problème. | Nous donnons pour la résolution du problème de *Cauchy* 5 méthodes différentes.

§52. 1) *Méthode de Liouville.*

Nous nous représentons grâce à la formule d'interpolation de *Lagrange* une fonction $f_1(x)$ de degré $n-1$ qui prend pour $x = \xi_1, ...x = \xi_n$ les valeurs $\mathcal{C}_1 ... \mathcal{C}_n$. Ensuite on a :

$$\frac{F(\xi_h)}{\Psi(\xi_h)} = \mathcal{C}_h = f_1(\xi_h) \text{ (h=1,...n).}$$

Nous construisons maintenant la fonction $f(x) = (x - \xi_1)(x - \xi_2)...(x - \xi_h)$ et nous appliquons sur $f$ et $f_1$ la décomposition en fraction continue ; on a alors :

$f_k = \psi_k f_1 - \varphi_k f$ ; pour $x = \xi_h$, il suit :

$$f_k(\xi_h) = \psi_k(\xi_h) f_1(\xi_h) \text{ ou :}^{24}$$
$$f_1(\xi_h) = \frac{f_k(\xi_h)}{\psi_k(\xi_h)} \ (h = 1, ...n).$$

Nous avons donc construit une fonction rationnelle $\frac{f_k(x)}{\psi_k(x)}$ qui prend pour $x = \xi_h$ les valeurs $\mathcal{C}_h = f_1(\xi_h)$, notre problème est ainsi résolu lorsque la somme des degrés du numérateur et du dénominateur est encore plus petite que $n$. Le degré du numérateur est $n_k$, celui du dénominateur est $n - n_{k-1}$, cette somme est donc $= n + n_k - n_{k-1}$, qui est $< n$, car $n_k < n_{k-1}$. L'ensemble des fractions $\frac{f_k(x)}{\psi_k(x)}$ dans lesquelles $f_k$ et $\psi_k$ n'ont pas de facteur commun, pour $k = 1$ à $k = \nu$, satisfait donc au problème d'interpolation de *Cauchy*. Il reste à savoir maintenant s'il y a encore d'autres solutions à ce problème. (Voir *annexe pag...*).

---

L'idée de la relation du travail de *Cauchy* avec la décomposition en fraction continue vient de *Liouville*. Elle se trouve publiée dans l'essai de *Sturm* : « *Démonstration d'un théorème d'Algèbre de M. Sylvester* » (*Journal de Liouville Bd.7, pag. 356* en 1842 [25]

126|127

| 53. 4. *Méthode de Jacobi : (Crelle Bd. 30).*
On peut aborder le problème de *Cauchy* de la manière suivante. Après que l'on a construit la fonction $f_1(x)$ au moyen de la formule de *Lagrange*, il suit de l'éq. :

$$\frac{F(\xi_h)}{\Psi(\xi_h)} = f_1(\xi_h) \text{ l'autre éq. :}$$

$$\frac{F(\xi_h)}{f_1(\xi_h)} = \Psi(\xi_h).$$

Le problème s'énonce donc maintenant ainsi : on doit déterminer une fonction $F(x)$ de degré $\mu < n$ de sorte que, divisée par $f_1(x)$, elle coïncide pour $x = \xi_h$ (h=1,...n) avec la fonction $\Psi(x)$, laquelle doit être au maximum de degré $n - \mu - 1$.

Nous nous représentons, d'après la formule de *Lagrange*, la fonction $\Psi$ établie de sorte qu'elle satisfasse pour $x = \xi_h$ l'éq. ci-dessus. Pour encore exprimer qu'elle doit être seulement de degré $n - \mu - 1$, nous posons les coefficients $= 0$ pour les exposants les plus hauts. Si l'on avait donné la formule de *Lagrange* pour $\Psi(x)$ :

$$\Psi(x) = b_0 + b_1 x + ... + b_{n-\mu-1} x^{n-\mu-1} + b_{n-\mu} x^{n-\mu} + ... + b_{n-1} x^{n-1}$$

(les $b$ contiennent naturellement les $\mu + 1$ coefficients inconnus et encore à déterminer de $F(x)$), alors on aurait à poser :

$$b_{n-\nu} = 0, ...b_{n-1} = 0.$$

On a encore maintenant la formule d'*Euler* (*pag...*) :

$$b_{n-1} = \sum_{h=1}^{n} \frac{\Psi(\xi_h)}{f'(\xi_h)} \ , \ b_{n-2} = \sum_{h=1}^{n} \frac{\xi_h \Psi(\xi_h)}{f'(\xi_h)} \ ,...b_{n-\mu} = \sum_{h=1}^{n} \frac{\xi^{\mu-1} \Psi(\xi_h)}{f'(\xi_h)}.$$

Si l'on pose ces coefficients égaux à 0, alors on obtient en vertu de l'éq. concernant $\Psi(\xi_h)$ (*pag 3....)* les équations suivantes :

24. Sachant qu'il a pourtant été supposé que $f_k$ et $\psi_k$ n'ont aucun facteur commun, voir annexe *pag.....*

25. Methode 2 et 3 : voir annexe *pag....*

$$\sum_{h=1}^{n} \frac{\xi_h^t F(\xi_h)}{f_1(\xi_h) f'(\xi_h)} = 0 \ ,(t = 0, 1, ...\mu - 1) \ ^{26}$$

On a $\mu$ équations homogènes, linéaires en les coefficients de $F$, qui nous permettent donc de déterminer ces coefficients.

127|128

| Nous voulons maintenant vraiment montrer cette résolution. Soit :

$$F(x) = C_0 + C_1 x + ... + C_\mu x^\mu$$

Ensuite notre équation s'énonce :

$$\sum_{\nu=1}^{n} \sum_{h=0}^{\mu} \frac{C_h \xi_\nu^{t+h}}{f_1(\xi_\nu) f'(\xi_\nu)} = 0 \ (t = 0, ...\mu - 1)$$

Nous posons maintenant :

$$\sum_{\nu=1}^{n} \frac{\xi_\nu^{t+h}}{f_1(\xi_\nu) f'(\xi_\nu)} = s_{h+t},$$

donc nos équations s'énoncent :

$$\sum_{h=0}^{\mu} C_h s_{h+t} = 0 \ (t = 0, ...\mu - 1) :$$

De là il suit, à une constante multiplicative près :

$$F(x) = \begin{vmatrix} s_0 & s_1 & ... & s_\mu \\ s_1 & s_2 & ... & s_{\mu+1} \\ ... & ... & ... & ... \\ s_{\mu-1} & s_\mu & ... & s_{2\mu-1} \\ 1 & x & ... & x^\mu \end{vmatrix} \ (V. \ pag. \ ...\text{et. } ...)$$

Ce déterminant se laisse encore simplifier et réduire au degré $\mu$. On a en effet :

$$\begin{vmatrix} ... & ... & ... & ... \\ s_p & s_{p+1} & ... & s_{p+\mu} \\ ... & ... & ... & ... \\ 1 & x & ... & x^\mu \end{vmatrix} = (-1)^\mu \begin{vmatrix} ... & ... & ... & ... & ... \\ s_p & xs_p - s_{p+1} & xs_{p+1} - s_{p+2} & ... & xs_{p+\mu-1} - s_{p+\mu} \\ ... & ... & ... & ... & ... \\ 1 & 0 & 0 & ... & 0 \end{vmatrix} =$$

$$D_\mu(x) = |xs_{p+q} - s_{p+q+1}| \ (p, q = 0, 1, ...\mu - 1).$$

Nous avons maintenant vu plus haut que chaque solution $\frac{F(x)}{\Psi(x)}$ doit être contenue dans la forme $\frac{f_k(x)\Theta(x)}{\psi_k(x)\Theta(x)}$. On a de là $F(x)$, ou, comme on l'a nommée en dernier, $D_\mu(x)$ qui n'est rien d'autre que :

$$D_\mu(x) = f_k(x)\Theta(x)$$

où $k$ est déterminé par le fait que $\mu$ se trouve dans l'intervalle : $n_k \leq \mu < n_{k-1}$. (Voir annexe *pag.*...).

Nous avons donc pour les deux cas limites :

$$C_k D_\mu(x) = \sigma_k f_k(x) \text{ pour } \mu = n_k$$
$$C_k D_\mu(x) = \tau_k f_k(x) \text{ pour } \mu = n_{k-1} - 1$$

128|129

26. Pour une dérivation plus simple de ces équations de conditions, voir annexe *pag.* ...

| *Hattendorf* a été le premier à généraliser ces recherches, sans toutefois trouver la solution définitive du problème.

Nous avons construit ci-dessus l'expression :

$$\sum_{\xi} \frac{\xi^t D_\mu(\xi)}{f_1(\xi) f'(\xi)} \ (t = 0, ...\mu - 1);$$

et l'on construit encore l'expression suivante, à savoir :

$\sum_\xi \frac{\xi^\mu D_\mu(\xi)}{f_1(\xi)f'(\xi)}$. D'abord on forme l'expression : $\frac{\xi^\mu D_\mu(\xi)}{f_1(\xi)f'(\xi)}$, en remplaçant la dernière ligne horizontale du déterminant $\begin{vmatrix} s_0 & s_1 & ... & s_\mu \\ ... & ... & ... & ... \\ s_{\mu-1} & s_\mu & ... & s_{2\mu-1} \\ 1 & x & ... & x^\mu \end{vmatrix}$ par les expressions $\frac{\xi^\mu}{f_1(\xi)f'(\xi)}, ..., \frac{\xi^{2\mu}}{f_1(\xi)f'(\xi)}$. Si l'on somme ensuite les déterminants se constituant ainsi, on obtient la même chose que lorsqu'on réalise de même la sommation sur les $\xi$ dans la dernière rangée horizontale. Cette dernière rangée se transforme ensuite en $s_\mu, s_{\mu+1}, ...s_{2\mu}$. De façon analogue à la définition de $\sigma_k$, nous pouvons de là désigner ce déterminant par $\sigma_{k-1}$ lorsque nous posons $\mu = n_{k-1} - 1$. Maintenant pour $\mu = n_{k-1} - 1$ on a :

$$C_k D_{n_{k-1}-1}(x) = \tau_k f_k(x).$$

Il s'ensuit conformément à la construction de $\sigma_{k-1}$ :

$$\sigma_{k-1} = \frac{\tau_k}{C_k} \sum_{\xi} \frac{f_k(\xi)\xi^{n_{k-1}-1}}{f_1(\xi)f'(\xi)} = \frac{\tau_k}{C_k} \sum_{\xi} \frac{\psi_k(\xi)}{f'(\xi)} \xi^{n_{k-1}-1} \text{ (Voir } pag....)$$

Comme alors $\psi_k$ est de degré $n - n_{k-1}$, il résulte de la formule d'*Euler* que dans cette somme tous les termes jusqu'au plus haut disparaissent et que la somme n'est rien d'autre que le coefficient du terme le plus haut de $\psi_k(x)$. Pour déterminer cala nous avons :

$$f = \psi_k f_{k-1} - \psi_{k-1} f_{k-1}.$$

Comme du côté droit $\psi_k f_{k-1}$ détermine le degré, et que $f_{k-1}$ possède comme plus haut coefficient $C_{k-1}$, alors le plus haut coefficient de $\psi_k$ doit être égal à $\frac{1}{C_{k-1}}$, car celui du côté gauche est $= 1$. Il en résulte donc maintenant :

$$\sigma_{k-1} = \frac{\tau_k}{C_k C_{k-1}} \text{ ou } C_k C_{k-1} \sigma_{k-1} = \Sigma_k.$$

Par cette formule de récurrence on peut donc exprimer les plus hauts coefficients de toutes les fonctions de reste par ces déterminants caractéristiques.

Nous voulons montrer que pour le cas moyen, où on a donc :

$$n_k < \mu < n_{k-1} - 1$$

| les déterminants $D_\mu(x)$ sont identiquement nuls. Ces fonctions $D_\mu$ ont été déterminées par les équations : 129|130

$$\sum_{\xi} \frac{\xi^t F(\xi)}{f_1(\xi) f'(\xi)} = 0 \ (t = 0, 1, ...\mu - 1)$$

où $F$ doit être de degré $\mu$. Maintenant il peut encore se produire qu'une fonction de degré inférieur à $\mu$ satisfasse aussi $\mu$ éq. linéaires, et cela s'applique justement au cas où la décomposition en fraction continue n'est pas régulière, comme nous l'avons justement bien supposé maintenant par l'hypothèse sur cet intervalle. Le cas régulier est défini par $n_k = n - k$.

On a maintenant conformément à la formule d'*Euler* :

$$\sum_{\xi} \frac{\xi^t f_k(\xi)}{f_1(\xi) f'(\xi)} = \sum_{\xi} \frac{\xi^t \psi_k(\xi)}{f_1(\xi) f'(\xi)} = 0$$

tant que $t$ prend les valeurs $0, 1 ... n_{k-1} - 2$, car pour la disparition de cette somme il faut que la condition $n - n_{k-1} + t < n - 1$ soit remplie. Nous avons donc ici $n_{k-1} - 1$ équations homogènes, que satisfait la fonction $f_k$ de degré $n_k$. Si ainsi $n_{k-1} - 1 = n_k$, alors la décomposition en fraction continue est régulière et ces équations sont satisfaites par $f_k$. On a maintenant :

$$f_k = C_k x^{n_k} + C'_k x^{n_k - 1} + C''_k x^{n_k - 2} + ...$$

de sorte que soient valables les équations :

$$\sum_{\xi} \frac{C_k \xi^{n_k + t} + C'_k \xi^{n_k + t - 1} + ...}{f_1(\xi) f'(\xi)} = 0 \text{ pour } t = 0, 1, ... n_{k-1} - 2$$

ou

$$C_k s_{n_k + t} + C'_k s_{n_k + t - 1} + C''_k s_{n_k + t - 2} + ... = 0 \ (t = 0, 1 ... n_{k-1} - 2)$$

Ces équations donnent le système suivant :

$$\left\| \begin{matrix} s_0 & s_1 & ... & s_{n_{k-1} - 2} \\ s_1 & s_2 & ... & s_{n_{k-1} + 1 - 2} \\ ... & ... & ... & ... \\ s_{n_k} & s_{n_k + 1} & ... & s_{n_{k-1} + n_k - 2} \end{matrix} \right\| \ (V.\ pag.\ ...\text{et}\ ...)$$

Dans le cas régulier on a $n_{k-1} = n_k + 1$, et nous avons ensuite $n_{k-1} - 1$ équations homogènes avec $n_{k-1} = n_k + 1$ inconnues dont on peut déterminer les relations. La fonction $f_k$ résout le système d'équations, sans qu'elle satisfasse, comme cela est nécessaire, encore d'autres équations. Mais si $n_{k-1} \geq n_k + 1$, de sorte que nous obtenons dans le système précédent autant resp. plus de rangées verticales que d'horizontales,
130|131 à la suite de | l'existence des $n_{k-1} - 1$ équations linéaires homogènes sous les $n_k + 1$ coefficients, tous les déterminants d'ordre $n_k + 1$ que l'on peut construire à partir de ce système par la suppression de rangées verticales sont égaux à 0.

Nous supposons maintenant le nombre $\mu$ mis sous la forme : $\mu = n_k + \lambda \leq n_{k-1} - 2$ (nous avons déjà traité le cas $\mu = n_{k-1} - 1$ et nous avons trouvé qu'alors par la disparition de certains déterminants le degré baisse de $\mu$ à $n_k$). Ensuite on a :

$$\left| \begin{matrix} s_0 & s_1 & ... & s_{n_k + \lambda} \\ s_1 & s_2 & ... & s_{n_{k-1} + \lambda + 1} \\ ... & ... & ... & ... \\ s_{n_k + \lambda - 1} & s_{n_k + \lambda} & ... & s_{2(n_k + \lambda) - 1} \\ 1 & x & ... & x^{n_k + \lambda} \end{matrix} \right|$$

Mais comme maintenant $\lambda \geq 1$ et qu'en outre on doit supposer $n_{k-1} - 1 > n_k + 1$ si nous ne voulons pas obtenir par hasard un des deux cas limites déjà traités, alors il résulte de la disparition de tous ces déterminants d'ordre $n_k + 1$ que $D_\mu(x)$ est identiquement nul.

Du fait donc que les $f_k$ sont de degré plus bas, comme cela est nécessaire pour le problème de *Cauchy*, certaines conditions sont remplies pour lesquelles on a la disparition des $D_\mu$. La disparition de ces déterminants n'a pas été remarquée par *Hattendorf*, mais a été prouvée en premier par *Kronecker* (Monatsberichte der Akad. 1873).

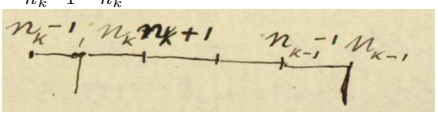
Nous voulons maintenant montrer que ces fonctions $D_\mu$ peuvent être aussi utilisées pour le théorème de *Sturm*, comme fonction de reste du développement en fraction

continue.

$$\text{On a : } \sum_{\xi}[f_1(\xi)f'(\xi)] = \frac{1}{2}\sum_{h}[f_{h-1}(x_2)f_h(x_2)] - \frac{1}{2}\sum_{h}[f_{h-1}(x_1)f_h(x_1)].$$

Nous voulons maintenant remplacer les $f$ par les fonctions $D_\mu$, et pour cela nous devons d'abord former le produit de deux $D_\mu$ successifs. Ici trois cas sont à différencier :

1. $D_{n_k-1}D_{n_k}$



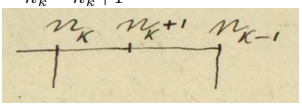
Pour cela nous avons :

$$C_k D_{n_k} = \sigma_k f_k$$
$$C_k D_{n_k-1-1} = \tau_k f_k \text{ donc } C_{k+1}D_{n_k-1} = \tau_{k+1}f_{k+1}$$

131|132

| Il s'ensuit : $D_{n_k-1}D_{n_k} = \sigma_k^2 f_k f_{k+1}$.

2. $D_{n_k}D_{n_k+1}$



Pour cela nous avons besoin maintenant de considérer que comme dans la figure ci-contre, on a $n_k + 1 = n_{k-1} - 1$, car dans un autre cas, comme cela est supposé par exemple dans la figure du premier cas, $D_{n_k+1}$ serait nul. Pour le cas $n_k + 1 = n_{k-1} - 1$ on a maintenant :

$$C_k D_{n_{k-1}-1} = \tau_k f_k. \text{ donc :}$$
$$D_{n_k}D_{n_k+1} = \frac{\sigma_k\tau_k}{C_k^2}f_k^2$$

3. Dans tous les autres cas le produit $D_{\mu-1}D_\mu$ est égal à 0.

Si maintenant l'expression :

$$\frac{1}{2}[D_{\mu-1}(x_2)D_\mu(x_2)] - \frac{1}{2}[D_{\mu-1}(x_1)D_\mu(x_1)]$$

correspond au côté droit de l'équation de *Sturm* (*pag....*) seulement pour $\mu = n_k$, alors on s'aperçoit pourtant tout de suite que c'est aussi le cas lorsque nous posons $\mu = h$ et que l'on somme comme plus tôt de 1 à $n$, car cette dernière différence disparaît pour le deuxième et le troisième cas, c'est-à-dire pour $\mu \neq n_k$. Ainsi il est donc montré que les fonction de reste de *Sturm* se laissent complètement remplacer par ces déterminants particuliers $D_\mu$ dans le théorème de *Sturm*. Ces fonctions $D_\mu$ ont été étudiées en premier par *Jacobi* (*Cr. Bd. 30). Cayley* a montré en premier leur utilisation pour le théorème de *Sturm*, c'est pourquoi on lui attribue fréquemment faussement la découverte de ces fonctions.

Lorsque l'on étudie maintenant les signes de la suite de fonctions $D_1, D_2, ...$ pour $x = x_1$ et $x = x_2$, alors il s'avère que, comme plus tôt, le nombre $m_1 - m_2$ de changements de signes perdus lors du passage de $x_1$ à $x_2$ est égal au côté droit de l'équation de *Sturm* (*pag....*). Comme certains $D$ sont effectivement nuls, alors il reste à savoir si le nul doit valoir comme positif ou comme négatif. Mais il s'avère que c'est indifférent, tant que l'on se décide pour l'un ou pour l'autre de façon constante.

Supposons maintenant que $f_1$ a toujours le même signe que $f'$. Posons par exemple $f_1 = f'$, ou $f_1 = \frac{1}{f'(x)}$. Dans le dernier cas | $f_1$ a été déterminée d'après *Lagrange* par la formule : 132|133

$$\sum_{\xi}\frac{f(x)}{(x-\xi)}\frac{1}{f'(\xi)^2} = f_1(x).$$

Ensuite on a

$$\sum_{\xi} \frac{\xi^t}{f_1(\xi) f'(\xi)} = s_t = \sum_{\xi} \xi^t,$$

nos $s$ se changent donc en somme de puissances.

Nous voulons poursuivre cela plus avant pour le cas régulier. Nous prenons $x_1 = -\infty, x_2 = +\infty$, pour obtenir toutes les racines réelles. Dans le cas régulier, l'exposant de la plus grande puissance dans le produit $D_{h-1}D_h$ est toujours un nombre impair. Les coefficients de ces plus hautes puissances dans tous les $D_{h-1}D_h$ sont les produits de deux $\sigma$ successifs, de sorte que l'on obtient :

$$\sum[f.f'] = \sum[1] = \frac{1}{2}\sum_k[\sigma_k\sigma_{k-1}] - \frac{1}{2}\sum_k[-\sigma_k\sigma_{k-1}] = \sum_k[\sigma_k\sigma_{k-1}].$$

Le nombre de toutes les racines réelles sera donc exprimé par la somme

$$\sum_{k=1}^{n}[\sigma_k\sigma_{k-1}],$$

où $\sigma_k$ représente le déterminant.

$$\sigma_k \begin{vmatrix} s_0 & s_1 & \ldots & s_{n-k-1} \\ s_1 & s_2 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ s_{n-k-1} & s_{n-k} & \ldots & s_{2n-2k-2} \end{vmatrix} = |s_{p+q}| \ (p, q = 0, 1 \ldots n-k-1)$$

Si l'on désigne par $s$ et $m$ le nombre de successions et de changements dans la suite $\sigma_0, \sigma_1, \ldots$ alors on a :

$$\sum[\sigma_k\sigma_{k-1}] = 1 + s - m.$$

Soit maintenant $\mathcal{A}$ le nombre de racines réelles, $2\mathcal{B}$ celui des racines imaginaires, alors :

$$\begin{aligned} \mathcal{A} &= 1 + s - m \\ 2\mathcal{B} &= n - \mathcal{A} = n - 1 - s + m. \end{aligned}$$

Comme maintenant $s + m = n - 1$, alors il suit $\underline{\mathcal{B} = m}$, c'est-à-dire que le nombre de changements de signes dans la suite indiquée est égal au nombre de paires de racines imaginaires. Cette proposition est valable, comme on le remarque, seulement si le développement de $f$ et $f_1$ est régulier.

Lorsque les déterminants $|s_{p+q}|$ se laissent représenter par des sommes de carrés, alors on en a conclu que l'équation avait seulement des racines réelles. Mais cette
133|134 conclusion est fausse, car une somme | de carrés peut bien aussi devenir nulle. La conclusion est donc seulement valable lorsque l'on prouve que les déterminants ne sont pas nuls.

Comme $\mathcal{B}$ peut être au maximum égal à $\frac{n}{2}$, alors il peut se produire dans cette suite $\sigma_0, \sigma_1, \ldots$ au plus seulement $\frac{n}{2}$ changements de signes, bien que l'on puisse a priori en espérer $n-1$. Cette limitation provient du fait que les sommes de puissance $s_{n+1} \ldots s_{2n}$ peuvent être exprimées par les précédentes $s_0 \ldots s_n$, et que donc entre les $\sigma$ il existe une dépendance. Ainsi, si dans la première moitié $\sigma_0 \ldots \sigma_{\frac{n}{2}}$, $\frac{n}{2}$ changements de signes devaient avoir lieu, alors dans la seconde moitié de la suite il ne peut plus s'en produire.

§55. Avant que nous discutions de la troisième méthode pour résoudre le problème de *Cauchy*, nous voulons établir une proposition correspondant à la multiplication des déterminants pour des systèmes rectangulaires. Soit les deux systèmes :

$$\left\| \begin{matrix} a_{00} & a_{01} & ... & a_{0n} \\ a_{10} & a_{11} & ... & a_{1n} \\ ... & ... & ... & ... \\ a_{m,0} & a_{m,1} & ... & a_{m,n} \end{matrix} \right\| \left\| \begin{matrix} b_{00} & b_{01} & ... & b_{0n} \\ b_{10} & b_{11} & ... & b_{1n} \\ ... & ... & ... & ... \\ b_{m,0} & b_{m,1} & ... & b_{m,n} \end{matrix} \right\| \quad n > m$$

alors, à partir de cela, nous pouvons composer le système quadratique suivant :

$$C_{g,h} = \sum_{k=0}^{n} a_{g,k} b_{hk} \text{ , g h } = 0...m.$$

Ensuite on a :

$$\begin{aligned} |C_{g,h}| &= \textstyle\sum_{(h)} \mathcal{E} \begin{pmatrix} h_0 & ... & h_m \\ 0 & ... & m \end{pmatrix} C_{0,h_0} C_{1,h_1} ... C_{m,h_m} \\ &= \textstyle\sum_{(k)} \sum_{(h)} \mathcal{E} \begin{pmatrix} h_0 & ... & h_m \\ 0 & ... & m \end{pmatrix} a_{0,k_0} ... a_{mk_m} b_{h_0,k_0} ... b_{h_m k_m}. \end{aligned}$$

La somme sur $k$ est à comprendre ainsi : à partir des nombres $0, ...n$ un système de valeurs $r_0 ... r_m$ est choisi, ensuite on somme sur toutes les $m!$ permutations $r_{i_1} ... r_{i_m}$ et finalement la somme est étendue à tous les systèmes de valeurs $r_0, ... r_m$ que l'on peut choisir sur $0, ...n$. Ensuite cela continue ainsi :

$$\mathcal{E} \begin{pmatrix} h_0 & ... & h_m \\ 0 & ... & m \end{pmatrix} = \mathcal{E} \begin{pmatrix} h_0 & ... & h_m \\ k_0 & ... & k_m \end{pmatrix} \mathcal{E} \begin{pmatrix} k_0 & ... & k_m \\ 0 & ... & m \end{pmatrix},$$

$$|C_{g,h}| = \textstyle\sum_{(k)} \mathcal{E} \begin{pmatrix} k_0 & ... & k_m \\ 0 & ... & m \end{pmatrix} a_{0k_0} ... a_{mk_m} |b_{hk_0} ... b_{hk_m}| \text{ ou :}$$

$$|C_{g,h}| = \textstyle\sum_{(k)} |a_{gk_0} ... a_{gk_m}| |b_{hk_0} ... b_{hk_m}|$$

134|135

| Maintenant la sommation est seulement encore à étendre à tous les systèmes de valeur $k_0 ... k_m$ que l'on peut choisir parmi $0...n$. Nous obtenons donc $|C_{g,h}|$ en choisissant dans le premier des deux systèmes rectangulaires un déterminant d'ordre $m$ par le retrait d'une rangée verticale, en multipliant avec le déterminant correspondant du deuxième système et en sommant tous les produits constitués de cette manière.

Appliquons maintenant cette proposition à un système de grandeurs particulier. Soit $u, v, \xi, \eta$ des variables quelconques. Les deux systèmes rectangulaires doivent ensuite s'énoncer :

$$u_i \xi_i^g \text{ et } v_k \eta_k^h \begin{pmatrix} i, k = 0, & ... & h_n \\ g, h = 0, & ... & k_m \end{pmatrix}$$

Ensuite on a :

$$|C_{g,h}| = |\sum_{k=0}^{n} u_k v_k \xi_k^g \eta_k^h| = \sum_{(k)} |u_{k_0} \xi_{k_0}^g ... u_{k_m} \xi_{k_m}^g| |v_{k_0} \eta_{k_0}^h ... v_{k_m} \eta_{k_m}^h| \ (g, h = 0...m)$$

Et donc :

$$\begin{matrix} |u_i \xi_i^g| \\ \begin{pmatrix} i, k = 0, & ... & h_n \\ g, h = 0, & ... & k_m \end{pmatrix} \end{matrix} = \prod u |\xi_i^g| = \prod u_\alpha \prod (\xi_\alpha - \xi_\beta) \ \alpha, \beta = k_0 ... k_m$$

L'indice de $\beta$ est toujours plus grand que celui de $\alpha$ dans $\prod(\xi_\alpha - \xi_\beta)$.
Il s'ensuit :

$$|C_{g,h}| = \sum_{(k)} \prod u_\alpha v_\alpha \prod u_\alpha v_\alpha \prod (\xi_\alpha - \xi_\beta)(\eta_\alpha - \eta_\beta)\ \alpha, \beta = k_0 ... k_m.$$

Nous posons maintenant $u_0 = 1$, $\xi_0 = x$, $\eta_0 = 0$, tandis que pour $k = 1, ...n$ on a :

$$\xi_k = \eta_k, v_k \eta_k = 1, \sum_{k=1}^{n} u_k \xi_k^h = s_h.$$

Ensuite notre déterminant $|C_{gh}|$ devient :

$$|\sum_{k=0}^{n} u_k v_k \xi_k^g \eta_k^h| = |v_0 x^g + s_{g-1}, s_g, s_{g+1} ... s_{g+m-1}|.$$

Si l'on fixe dans les deux représentations pour $|C_{g,h}|$ les facteurs $v_0$, il s'ensuit :

$$v_0 |x^g, s_g, s_{g+1} ... s_{g+m-1}| = v_0 \sum_{(k)} u_0 \prod (\xi_\alpha - \xi_\beta)(\eta_\alpha - \eta_\beta) \prod_{(\alpha,\beta = k_1 ... k_m)} (\xi_0 - \xi_\beta)(\eta_0 - \eta_\beta) \prod u_\alpha v_\alpha$$

La somme est à étendre sur tous les systèmes de valeurs $k_1 ... k_m$, que l'on peut choisir parmi la suite de nombres $1, ...n$. Si l'on considère le cas spécial de $u, v, \xi, \eta$ alors il suit :

$$|x^g, s_g, s_{g+1} ... s_{g+m-1}| = \sum_{(k)} (-1)^m \prod u_\alpha \prod_{(\alpha,\beta = k_1 ... k_m)} (x - \xi_\alpha) \prod (\xi_\alpha - \xi_\beta)^2$$

135|136 Si l'on pose maintenant $u_k = \frac{1}{f_1(\xi_k) f'(\xi_k)}$, alors la dernière éq. devient :

$| D_m(x) = \sum_{(k)} \prod \frac{1}{f_1(\xi_\alpha) f'(\xi_\alpha)} \prod (x - \xi_\alpha) \prod (\xi_\alpha - \xi_\beta)^2.$

Par cette hypothèse sur $u_k$ nous obtenons donc $D_m$ représentée comme une fonction symétrique des $\xi$ et nous avons ainsi établi le rapport de ces expressions combinatoires avec le problème de *Cauchy*. Si nous posons $u_k = 1$ ou $f_1 = \frac{1}{f'}$, alors nous obtenons les formules de *Sylvester*, à savoir pour $m = 1, 2, ...n$ :

$$\begin{vmatrix} s_0 & s_1 \\ 1 & x \end{vmatrix} = \sum_{\alpha=1}^{n} (x - \xi_\alpha).$$

$$\begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ 1 & x & x^2 \end{vmatrix} = \sum_{(\alpha,\beta=1,...n)} (x - \xi_\alpha)(x - \xi_\beta)(\xi_\alpha - \xi_\beta)^2 \quad \text{etc. etc.}$$

La dernière formule s'énonce :

$$\begin{vmatrix} s_0 & s_1 & ... & s_n \\ s_1 & s_2 & ... & s_{n+1} \\ ... & ... & ... & ... \\ s_{n-1} & ... & ... & s_{2n-1} \\ 1 & x & ... & x^n \end{vmatrix} = \prod_{\alpha=1}^{n} (x - \xi_\alpha) \prod_{\alpha,\beta=1...n;\alpha<\beta} (\xi_\alpha - \xi_\beta)^2$$

Dans les formules de Sylvester les $s$ désignent les sommes de puissances. La première des formules est évidente, car $\sum (x - \xi - \xi_\alpha)$ est égal à $nx - s_1$ et comme $s_0 = n$, il s'ensuite que $\begin{vmatrix} s_0 & s_1 \\ 1 & x \end{vmatrix} = nx - s_1$.

La dernière des formules indique que le produit de la fonction initiale $f(x) = \prod (x - \xi_\alpha)$ par son discriminant $\prod (\xi_\alpha - \xi_\beta)^2$ se laisse représenter par ce déterminant caractéristique. (Voir *Joachimsthal, Crelle Bd. 48.*)

Nous avons maintenant démontré le rapport de l'expression combinatoire dérivée avec la méthode de *Jacobi* et nous allons donner maintenant encore une autre méthode de résolution du problème de *Cauchy* essentiellement différente des méthodes précédentes. Dans celle-ci les formules de *Sylvester* sont à nouveau utilisées. Nous donnons donc maintenant la

5ième Méthode de résolution du problème de *Cauchy*. On a :

$$\frac{f_1}{f} - \frac{\Phi}{\Psi} = \frac{F}{\Psi}.$$

Nous pouvons ensuite poser le problème suivant : on doit déterminer une fraction $\frac{\Phi}{\Psi}$, dont le dénominateur est de degré $\nu$, de sorte que la différence $\frac{f_1}{f} - \frac{\Phi}{\Psi}$ pour le degré $\nu < n$ donné devienne infiniment petite pour | $x$ infiniment grand d'un ordre 136|137
aussi grand que possible. $\frac{F}{f\Psi}$ devient en effet infiniment grand pour des $x$ d'ordre $-(n+\nu-\mu)$ infiniment grand. $\frac{\Phi}{\Psi}$ devrait donc se rapprocher autant que possible de $\frac{f_1}{f}$ pour des $x$ infiniment grands. Si l'on développe maintenant $\frac{f_1}{f}$ selon les puissances décroissantes de $x$, alors il vient :

$$\frac{f_1}{f} = \sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} \frac{1}{x-\xi} = \sum_{k=0}^{\infty} \sum_{\xi} \frac{f_1(\xi)}{f'(\xi) \frac{\xi^k}{x^{k+1}}}.$$

Si l'on pose : $\sum \frac{f_1(\xi)}{f'(\xi)} \xi^k = s'_k$, on a alors : $\frac{f_1}{f} = \sum_{k=0}^{\infty} \frac{s'_k}{x^{k+1}}$.

Si l'on décompose d'abord $\frac{\Phi}{\Psi}$ en fraction partielle et que l'on appelle $\eta_k$ $(k = 1, ...\nu)$ les racines de $\Psi$, alors on a :

$$\frac{\Phi}{\Psi} = \sum_{h=1}^{\nu} \frac{v_h}{x-\eta_k} = \sum_{h=1}^{\nu} \sum_{k=0}^{\infty} \frac{v_h \eta_k^k}{x^{k+1}} = \sum_{k=0}^{\infty} \frac{\sum_{h=1}^{[\nu]} v_h \eta_h^k}{x^{k+1}}.$$

Nous avons maintenant $2\nu$ variables $v_h, \eta_h$ $(h = 1, ...\nu)$ et le problème est satisfait en posant :

$$\sum_{h=1}^{\nu} v_h \eta_h^k = s'_k \ (k = 0, ...2\nu - 1)$$

Alors l'approximation de $\frac{\Phi}{\Psi}$ par $\frac{f_1}{f}$ est la plus grande possible pour $x = \infty$. Si nous utilisons maintenant la formule *pag....* à partir de laquelle, pour $u_k = 1$, découle la formule de *Sylvester* et nous posons aussitôt $\sum_{h=1}^{\nu} v_h \eta_h^k = s'_k$, alors il vient :

$$\begin{vmatrix} s'_0 & s'_1 & ... & s'_\nu \\ s'_1 & ... & ... & ... \\ ... & ... & ... & ... \\ s'_{\nu-1} & ... & ... & s'_{2\nu-1} \\ 1 & y & ... & y^\nu \end{vmatrix} = \prod v_h \prod (y - \eta_k) \prod (\eta_h - \eta_k)^2 = D.\Psi(y).$$

Ainsi nous avons représenté $\Psi(y)$ sous la forme d'un déterminant, à un facteur de proportionnalité près, et nous nous apercevons que cette conception du problème de *Cauchy* conduit au même résultat que la méthode de *Jacobi* bien que le moyen en soit complètement différent.

Nous pouvons maintenant aussi représenter la grandeur $v$ sous la forme d'un déterminant. Supposons extraits du déterminant, qui est $= D\Psi(y)$, les coefficients de $y^\nu$ et désignons par $V(y) = y^\nu + c_1 y^{\nu-1} + ...$ la fonction dont les racines sont $\eta_1 ... \eta_\nu$ et qui se différencie donc seulement par un facteur constant de notre déterminant $D\Psi(y)$. Si l'on introduit ensuite également de manière analogue à ce qui a été fait plus tôt pour $F(x)$ | (voir *pag....*) les fonctions $g_r(y)$ définies par $g_r(y) = y^r - c_1 y^{r-1} + ... \pm c_r$, 137|138
alors on a en conséquence cette définition :

$$g_r(y) = \begin{vmatrix} s'_0 & s'_1 & ... & ... & ... & ... & s'_\nu \\ s'_1 & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... \\ s'_{\nu-1} & ... & ... & ... & ... & ... & s'_{2\nu-1} \\ 0 & 0 & ... & 1 & y & ... & y^\nu \end{vmatrix}$$

Maintenant, lorsque $V'(y)$ désigne la dérivée de $V(y)$, on a :

$$\sum_{h=1}^{\nu} \sum_{r=0}^{\nu-1} g_r(\eta_k) v_h \eta_h^{\nu-r-1} = v_k V'(\eta_k) = \sum_{r=0}^{\nu-1} g_r(\eta_k) s'_{\nu-r-1}.$$

Si l'on écrit maintenant l'expression $\sum_{r=0}^{\nu-1} g_r(\eta_k) s'_{\nu-r-1}$ sous la forme d'un déterminant en formant la somme avec l'aide de la forme de $g_r(y)$, et en exécutant la sommation dans la dernière rangée horizontale, alors cette rangée devient :

$$0, s'_0, s'_1 + y s_0, s'_2 + y' s'_1 + y^2 s'_0, ..., s'_{\nu-1} + y s'_{\nu-2} + ... + y^{\nu-1} s'_0$$

Une simple transformation nous conduit au déterminant suivant :

$$(-1)^\nu \begin{vmatrix} 0 & s'_0 & s'_1 & ... & s'_{\nu-1} \\ s'_0 & s'_1 - y s'_0 & s'_2 - y s'_1 & ... & s'_\nu - y s'_{\nu-1} \\ s'_1 & s'_2 - y s'_1 & s'_3 - y s'_2 & ... & s'_{\nu+1} - y s'_\nu \\ ... & ... & ... & ... & ... \\ s'_{\nu-1} & s'_\nu - y s'_{\nu-1} & s'_{\nu+1} - y s'_\nu & ... & s'_{2\nu-1} - y s'_{2\nu-2} \end{vmatrix}$$

Cette expression représente pour $y = \eta_k$ la valeur $v_k V'(\eta_k)$, de sorte que nous avons donc déterminé $v_k$ comme quotient de deux déterminants. Le dernier déterminant représente à une constante multiplicative près la fonction $\Phi(y)$, car on a, à cette constante multiplicative près :

$$\Phi(y) = \sum_{h=1}^{\nu} v_k \frac{V(y)}{y - \eta_k} = \sum_{h=1}^{\nu} \sum_{r=0}^{\nu-1} g_r(y) v_h \eta_h^{\nu-r-1} = \sum_{r=0}^{\nu-1} g_r(y) s'_{\nu-r-1}$$

---

138|139

| §56. Nous ajoutons à ce problème, qui demandait de trouver la plus grande approximation possible d'une fraction à une fraction donnée, un problème analogue qui vient de *Tchebicheff* et qui est à considérer comme une généralisation du problème d'interpolation de *Lagrange*. Le problème s'énonce ainsi : on doit déterminer une fonction entière $\mathcal{G}(x)$ de degré $n - n_k - 1$ de sorte qu'elle devienne aussi proche que possible pour $n$ valeurs $\xi_1 ... \xi_n$ d'une fonction entière donnée $T(x)$ d'un degré plus haut que $n - n_k - 1$.

On a donc pour obtenir un minimum, d'après la méthode des moindres carrés, l'expression :

$$\sum_{h=1}^{n} \Theta_h^2 (\mathcal{G}(\xi_h) - T(\xi_h))^2$$

lorsque la grandeur positive $\Theta_h^2$ représente l'importance de l'erreur $\mathcal{G}(\xi_h) - T(\xi_h)$.

Si nous formons $f(x) = (x - \xi_1)...(x - \xi_n)$ et de là $f'(\xi)$, et que nous posons ensuite $f_1(x) = \Theta^2 f'(x)$, alors on a :

$$\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} (\mathcal{G}(\xi) - T(\xi))^2$$

pour obtenir un minimum. La condition nécessaire pour cela est la concomitance des équations :

$$\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)}(\mathcal{G}(\xi) - T(\xi))\xi^t = 0 \ , \ (t = 0, 1, ...n - n_k - 1).$$

La deuxième différenciation apporte comme condition suffisante pour le minimum que :

$$\sum_{r,s}\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)}\xi^{r+s} z_r z_s$$

soit une forme définie positive. Mais on a :

$$\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} \sum_{r,s} \xi^{r+s} z_r z_s = \sum_{\xi} \frac{f_1(\xi)}{f'(\xi)} (\sum_{r} \xi^r z_r)^2$$

et comme $\frac{f_1(\xi)}{f'(\xi)} = \Theta^2$, alors les conditions nécessaires et suffisantes pour un minimum sont déjà remplies, dans la mesure où l'on peut seulement prouver que $\mathcal{G}$ peut toujours être déterminée de sorte | que : 139|140

$$\textstyle\sum_{\xi} \frac{f_1(\xi)}{f'(\xi)}(\mathcal{G}(\xi) - T(\xi))\xi^t = 0 \text{ pour } t = 0, 1, ...n - n_k - 1.$$

Cette preuve se relie de préférence avec le développement en fraction continue de $f$ et $f_1$. On a :

$$f_1\psi_k - f\varphi_k = f_k \ , \ f_k\psi_{k+1} - f_{k+1}\psi_k = f.$$

De là on déduit :

$$f_k\varphi_{k+1} - f_{k+1}\varphi_k = f_1.$$

Soit maintenant $T(x)$ une fonction entière quelconque de sorte que nous puissions développer $f(x) = 0$ dans sa forme réduite, donc la supposer au plus de degré $n - 1$. Divisons maintenant $f_k(x)T(x)$ par $f(x)$, de sorte qu'il vienne :

$$f_k(x)T(x) = Q_k(x)f(x) + R_k(x).$$

$R_k$ est d'après la formule de *Lagrange* :

$$R_k(x) = \sum_{\xi} \frac{T(\xi)f_k(\xi)}{f'(\xi)} \frac{f(x)}{x - \xi}.$$

Si nous formons maintenant de là l'expression :

$$R_k(x)f_{k+1}(x) - R_{k+1}(x)f_k(x)$$

alors on peut s'apercevoir facilement qu'elle est divisible par $f(x)$. Nous la posons de là égale à $f(x).R(x)$, où $R$ est une fonction entière. Maintenant $R$ est au moins de degré $n_k - 1$, comme il ressort de la construction de $R$.

Si nous construisons également :

$$R_k(x)\psi_{k+1}(x) - R_{k+1}(x)\psi_k(x)$$

alors cette expression est aussi divisible par $f$. Nous la posons de là égale à $f(x)\mathcal{G}(x)$, où $\mathcal{G}(x)$ est une fonction entière de degré $n - n_k - 1$. Ensuite on a :

$$\mathcal{G}(x)f(x) = \sum_{\xi} \frac{T(\xi)(f_k(\xi)\psi_{k+1}(x) - f_{k+1}(\xi)\psi_k(x))}{f'(\xi)} \frac{f(x)}{x - \xi}.$$

Si maintenant $T$ est de degré $n - n_k - 1$, alors $S(x)$ est identique à $T(x)$. Mais nous voulons supposer que $T$ soit d'un degré plus haut que $n - n_k - 1$. Ensuite on a :

$$ff_1\mathcal{G} = R_k(f_{k+1} + f\varphi_{k+1}) - R_{k+1}(f_k + f\varphi_k) =$$

140|141 |

$$R_k f_{k+1} - R_{k+1}f_k + f(f_k T - Q_k f)\varphi_{k+1} - f(f_{k+1}T - Q_{k+1}f)\varphi_k$$
$$= Rf + ff_1 T + f^2(\varphi_k Q_{k+1} - \varphi_{k+1}Q_k) = ff_1\mathcal{G}.$$

On a donc : $f_1(\mathcal{G} - T) \equiv R \bmod. f$.
Pour exprimer ce résultat en équations, nous avons d'après *Euler* :

$$\sum_\xi \frac{R(\xi)}{f'(\xi)}\xi^t = 0 \ , \ t = 0, 1, ...n - n_k - 1, \text{ et par conséquent :}$$

$$\sum_\xi \frac{f_1(\xi)}{f'(\xi)}(\mathcal{G}(\xi) - T(\xi))\xi^t = 0 \ , \ t = 0, 1, ...n - n_k - 1$$

Il est ainsi démontré la possibilité de construire une fonction $\mathcal{G}(x)$, comme le demande le problème de *Tchebicheff*. $\mathcal{G}(x)$ peut représenter justement pour $x = \xi_1 ... \xi_n$ une fonction donnée $T(x)$ (ou des grandeurs données $T(\xi_1), ...T(\xi_n)$).

§ 57. Nous voulons maintenant encore envisager le cas régulier du développement en fraction continue, pour faire ensuite l'exposé d'une nouvelle méthode, laquelle nous conduira au théorème de *Sturm*. Pour le cas régulier on a $n_k = n - k$. En reprenant les désignations de la *pag....* nous avons $\sigma_k = \tau_k$, en conséquence :

$$\frac{\tau_{k-1}}{\tau_k}C_k C_{k-1} = 1 \ , \text{ si } C_k \text{ et } C_{k-1}$$

sont les coefficients des plus hautes puissances de $f_k$ et $f_{k-1}$. On a maintenant, essentiellement pour le cas régulier, le fait que l'on peut exprimer linéairement chaque forme réduite d'une fonction de $x$ par $f_1(x), ...f_{n-1}(x)f_n(x)$, étant donné que le degré de ces fonctions sont les valeurs consécutives $n - 1, ...1, 0$. Chaque forme réduite de $x$ peut donc être donnée par $y_1 f_x + y_2 f_2(x) + ... + y_n f_n(x)$. Nous avions (*pag....*) :

$$\sum_\xi \frac{\xi^t f_k(\xi)}{f_1(\xi)f'(\xi)} = 0 \ , \ t = 0, ...n - k - 1.$$

on a donc aussi :

$$\sum_\xi \frac{f_i(\xi)f_k(\xi)}{f_1(\xi)f'(\xi)} = 0 \ , \text{ lorsque } n - i < n - k,$$

c'est-à-dire lorsque $k < i$.
141|142 | Mais comme $i$ et $k$ sont égaux en droit, alors on a en fait :

$$\sum_\xi \frac{f_i(\xi)f_k(\xi)}{f_1(\xi)f'(\xi)} = 0 \text{ pour } i \neq k.$$

Si nous posons pour $i = k$ : $S_k = \sum_\xi \frac{(f_k(\xi))^2}{f_1(\xi)f'(\xi)}$, alors on doit poser :

$$\sum_\xi \frac{f_i(\xi)f_k(\xi)}{f_1(\xi)f'(\xi)} = \delta_{i,k}S_k = \left\{ \begin{array}{l} \delta_{i,k} = 0 \text{ pour } i \neq k \\ \delta_{i,k} = 0 \text{ pour } i = k \end{array} \right\}.$$

Ces $n^2$ équations se laissent assembler en une seule avec l'aide de $n$ variables quelconques $y_1 ... y_n$. Nous obtenons :

$$\sum_{i,j}\sum_{\xi} y_i y_k \frac{f_i(\xi) f_k(\xi)}{f_1(\xi) f'(\xi)} = \sum_{\xi} \frac{1}{f_1(\xi) f'(\xi)} \left( \sum_{k=1}^{n} y_k f_k(\xi) \right)^2 = \sum_{k=1}^{n} S_k y^2.$$

Si l'on pose maintenant $\sum_{k=1}^{n} y_k f_k(\xi_h) = z_h$, alors on a :

$$\sum_{h=1}^{n} \frac{1}{f_1(\xi_h) f'(\xi_h)} z_h^2 = \sum_{k=1}^{n} S_k y^2.$$

Avec l'aide de la substitution :

$$z_h = \sum_{k=1}^{n} y_k f_k(\xi_h)$$

nous avons par conséquent transformé une somme de carrés en une autre somme de carrés.
Si maintenant $f_1(\xi_h)$ et $f'(\xi_h)$ sont du même signe pour $h = 1, ...n$, alors on a ici transformé une somme de carrés positifs en une autre somme de carrés positifs.

Cette invariance du nombre de signes positifs dans une somme de carrés positifs était connue depuis longtemps, notamment en géométrie. Mais *Jacobi* le premier (*Crelle Bd.* 53) a découvert que dans une forme quadratique quelconque le nombre de signes positifs et négatifs reste inchangé lors d'une transformation linéaire. Presque au même moment *Hermite* et *Sylvester* se sont occupés de cette question. Du dernier vient l'expression : principe d'inertie des formes quadratiques.

Nous voulons maintenant démontrer la justesse de ce principe, en ramenant la proposition au cas simple, à savoir la transformation de la somme de deux carrés.

| On doit donc avoir $pu^2 + qv^2 = p'u'^2 + q'v'^2$ et on doit démontrer que lors de 142|143
cette transformation la somme de signes positifs reste inchangée. La transformation est réalisée par la substitution linéaire générale :

$$u' = m(u + qtv) \ , \ v' = n(v - ptu),$$

où $m, n, t$ sont à choisir quelconques. On a :

$$p' = \frac{p}{m^2(1 + pqt^2)} \ , \ q' = \frac{q}{n^2(1 + pqt^2)}.$$

Nous avons donc une variété d'ordre 3 de substitutions possibles. Si maintenant $p$ et $q$ ont le même signe, et donc que $pq > 0$, alors il résulte de :

$$p'q' = \frac{pq}{m^2 n^2 (1 + pqt^2)^2},$$

que l'on a aussi $p'q' > 0$. Il résulte également de $pq < 0$ qu'aussi $p'q' < 0$. Ainsi ce principe est démontré pour cette transformation simple.

La forme :

$$\sum_{i=1}^{n} a_i z_i^2 \text{ se transforme généralement en } \sum_{h=1}^{n} b_h y_h^2$$

au moyen de la substitution :

$$z_i = \sum_{k=1}^{n} C_{ik} y_k$$

de sorte que l'on a :

$$\sum_{i=1}^{n} a_i z_i^2 = \sum_{i,r,s=1}^{n} a_i C_{ir} C_{is} y_r y_s = \sum_{h=1}^{n} b_h y_h^2.$$

En même temps les $n^2$ coefficients doivent remplir les $\frac{n.n-1}{2}$ conditions :

$$\sum_i a_i C_{ir} C_{is} = \delta_{r,s} b_r.$$

En conséquence de quoi les $n^2$ coefficients dépendent des $\frac{n.n-1}{2}$ variables indépendantes. Dans notre transformation élémentaire nous pouvons tout de suite exprimer les coefficients ($n^2 = 4$) par $\frac{n.n+1}{2} = 3$ variables rationnelles quelconques $m, n, t$.

De façon analogue à la qualification introduite par *Clebsch*, d'après laquelle une courbe $F(x, y) = 0$ sera qualifiée de rationnelle ou de genre 0 si $x$ et $y$ se laissent exprimer rationnellement par un paramètre $t$, nous pouvons aussi qualifier notre transformation de genre nul. Il s'avère en effet que nous sommes susceptibles d'exprimer rationnellement les $n^2$ coefficients par $\frac{n.n+1}{2}$ variables. Nous résolvons progressive-
143|144 ment le problème | par l'utilisation successive des transformations élémentaires.
Soit donc : $\sum_{i=1}^{n} a_i z_i^2 = \sum_{h=1}^{n} b_h y_h^2$ en vertu de $z_i = \sum_{k=1}^{n} C_{ik} y_k$.
Nous commençons lors de la transformation de $\sum_{i=1}^{n} a_i z_i^2$ avec $a_1 z_1^2 + a_2 z_2^2$. Nous pouvons transformer ces sommes dans $a'_1 z'^2_1 + a'_2 z'^2_2$ avec l'aide d'un $t$ que nous pouvons choisir tel que dans une des fonctions linéaires des $y$, par exemple dans $z_1$, le terme $y_n$ soit absent. Si cela arrive, alors nous assemblons $z_2$ et $z_3$ et nous les transformons de sorte que aussi dans $z_2$ le terme $y_n$ soit absent et on continue ainsi. Cela continue aussi longtemps qu'il reste un $z$ à disposition que nous pouvons assembler avec le précédent, c'est-à-dire aussi longtemps que nous pouvons avoir dans les $n-1$ fct. $z$ le terme $y_n$ qui manque.

Nous avons maintenant :

$$\sum_{i=1}^{n} a'_i z'^2_i = \sum_{h=1}^{n} b_h y_n^2,$$

où $z'_1 ... z'_{n-1}$ sont uniquement des fct. de $y_1 ... y_{n-1}$ et $z'_n$ peut être encore une fonction linéaire de $y_1 ... y_{n-1} y_n$, et donc $z'_n = \mathcal{A} y_n + \mathcal{B} y_{n-1} + ....$ Nous avons donc obtenu : $a'_n(\mathcal{A} y_n + ...)^2 +$ fct. de $y_1 ... y_{n-1}$ uniquement.

Si nous développons ce carré, on obtient des produits de $y_n$ avec les autres $y$, lesquelles ne peuvent pas apparaitre, et donc doivent être égaux à 0. Il s'ensuit :

$$\sum_{i=1}^{n} a_i z_i^2 = \sum_{h=1}^{n-1} a'_h z'^2_h + \mathcal{A} y'_n.$$

Nous avons à faire précisément avec une somme de $n-1$ carrés et nous obtenons :

$$\sum_i a_i z_i^2 = \sum_{g=1}^{n-2} a''_g z''_g + \mathcal{B} y_{n-1}^2 + \mathcal{A} y_n^2.$$

Si l'on continue ainsi, alors nous obtenons finalement de cette manière uniquement une somme de carrés de $y$, qui est naturellement identique à notre somme $\sum_i a_i z_i^2$, de sorte que $\mathcal{A} = b_n$, $\mathcal{B} = b_{n-1}$ etc.

Nous avons donc ainsi changé une somme de carrés de fonctions linéaires de $y$ en une somme de carrés de $y$ [27], et ce toujours simplement par nos transformations élémentaires. Celles-ci dépendent uniquement d'un paramètre $t$, de sorte que nous avons globalement $\frac{n.n-1}{2}$ $t$. Mais comme nous avons en plus besoin pour chaque $y$

27. *N.d.T* Il s'agit très certainement d'un $z$.

d'un facteur $b$, alors le nombre | de constantes nécessaires se maintient à $\frac{n.n+1}{2}$ et 144|145
nous nous apercevons que notre procédé fonctionne effectivement en $\frac{n.n+1}{2}$ pas, et qu'il est donc à qualifier de genre 0.

Comme nous nous sommes aperçus que chaque transformation se compose de transformations élémentaires et que lors des transformations élémentaires la constance des signes est démontrée, alors nous savons que le théorème sur la constance des signes est vrai aussi pour la transformation générale $\sum a_i z_i^2 = \sum b_h y_h^2$, de sorte que $\sum[a_i] = \sum[b_i]$.

Par conséquent le nombre des signes des coefficients d'une forme quadratique est un invariant.

§58. Nous voulons maintenant appliquer le résultat trouvé disant que lorsque l'on peut ainsi transformer une somme de carrés en une autre somme de carrés, la somme des signes positifs, tout comme celle des négatifs, reste elle-même constante, au théorème de *Sturm*. Nous allons voir que le théorème de *Sturm* est seulement une autre façon d'exprimer le principe d'inertie des formes quadratiques. Nous avons en effet en vertu de ce principe :

$$\sum_i [a_i z_i^2] = \sum_h [b_h y_h^2]$$

ou :

$$\sum_i [a_i] = \sum_h [b_h]$$

Appliquons cela à la forme quadratique :

$$\sum_\xi \frac{1}{f_1(\xi) f'(\xi)} (y_1 f_1(\xi) + ... + y_n f_n(\xi))^2 = \sum_k S_k^2.$$

Nous supposons que $f$ a des coefficients réels, mais que les $\xi$ désignent aussi des imaginaires, à savoir des valeurs imaginaires conjuguées. Les $y$ sont réelles. Ensuite :

$$\frac{y_1 f_1(\xi) + ... + y_n f_n(\xi)}{\sqrt{f_1(\xi) f'(\xi)}}$$

apparaît sous la forme $u + vi$, où $u$ et $v$ sont des fonctions de $y$ avec des coefficients réels. La valeur conjuguée de $\xi$ est ensuite $u - vi$. Nous obtenons de là pour chaque paire de racines imaginaires comme contribution à notre somme l'expression :

$$(u + vi)^2 + (u - vi)^2 = 2u^2 - 2v^2$$

| Pour la somme des signes on a : 145|146

$$\sum[2u^2] + \sum[-2v^2] = 0$$

c'est-à-dire que toutes les racines imaginaires s'annihilent lorsque l'on prend en compte pour les formes quadratiques seulement les sommes de signes. Nous avons donc maintenant :

$$\sum_\xi [f_1(\xi) f'(\xi)] = \sum_k [S_k]$$

où à présent la somme est à étendre seulement sur les réels $\xi$. Mais c'est le théorème de *Sturm* dans le cas où l'on a repoussé les limites à l'infini. Car on a :

$$S_k = \sum_\xi \frac{f_k(\xi)}{f_1(\xi) f'(\xi)} = \sum_\xi \frac{f_k(\xi) \psi_k(\xi)}{f'(\xi)}.$$

Cela se réduit d'après la formule d'*Euler* aux coefficients des termes de plus haut degré, où le degré du numérateur est $= n-1$, c'est-à-dire à :

$$C_k \frac{1}{C_{k-1}} \text{ et par conséquent : } \sum_k [S_k] = \sum_k [C_k C_{k-1}] \text{ (Voir } pag....)$$

Cette démonstration du théorème de *Sturm* a pourtant le défaut d'utiliser d'abord la preuve de l'existence d'une racine complexe, preuve qui sortira complètement à nouveau lors du résultat final. Nous voulons ainsi tout de suite envisager le cas général du théorème de *Sturm*, lequel nous conduit en même temps à la formule de *Jacobi*. Ce chemin révèle aussi que le principe de conservation des signes est pour ainsi dire identique au théorème de *Sturm*.

Déterminons maintenant les fonctions plus générales $F_i(\xi)$ $(i = 1, ...n)$ telles que :

$$\sum_\xi \frac{F_i(\xi)F_k(\xi)}{f_1(\xi)f'(\xi)} = \delta_{i,k} \sum_\xi \frac{(F_k(\xi))^2}{f_1(\xi)f'(\xi)} = \delta_{ik} S_k.$$

Qu'il y ait de telles fonctions $F_i$, les fonctions de reste $f_i$ le montrent. Si l'on pose : $\sum \xi \frac{F_i(\xi)F_k(\xi)}{f_1(\xi)f'(\xi)} = S_i \mathcal{A}_{ik}$ et que l'on réunit cette égalité avec la précédente après avoir multiplié la dernière par $x$, alors il vient :

$$\sum_\xi \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{i=1}^n \right)^2 = \sum_{i,k=1}^n (x\delta_{ik} - \mathcal{A}_{ik}) S_i v_i v_k.$$

Cette égalité représente la transformation d'une forme quadratique en la somme de carrés. Nous avons maintenant obtenu :

$$f_{k-1} - g_k f_k + f_{k+1} = 0.$$

146|147 | $g_k$ est dans le cas régulier, ce que nous supposons bien toujours ici, une fonction linéaire de $x$. Soit $g'_k$ la dérivée de $g_k(x)$, de sorte que l'on pose maintenant $F_k(x) = g'_k f_k$, par quoi notre supposition sur $F_k(x)$ est réalisée.

Nous transformons maintenant, avec cette hypothèse sur $F_i$, la somme des carrés en isolant le terme où $i = k$ des termes où $i \neq k$.

Avec l'utilisation pas à pas de la formule d'*Euler*, on obtient alors immédiatement :

$$\sum_\xi \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{i=1}^n F_i(\xi) v_i \right)^2 = \sum_\xi \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{i=1}^n g'_i f_i(\xi) v_i \right)^2 =$$
$$\sum_{\xi,k} \frac{x-\xi}{f_1(\xi)f'(\xi)} (f_i(\xi))^2 g'^2_k v_i^2 - 2 \sum_{\xi,k=1,...n-1} \xi \frac{f_k(\xi)f_{k+1}(\xi)}{f_1(\xi)f'(\xi)} g'_k g'_{k+1} v_k v_{k+1}$$

Or on a : $(x-\xi)g'_k = g_k(x) - g_k(\xi)$,
et de plus : $f_k(\xi) = f_1(\xi)\psi(\xi)$. De là notre première somme devient :

$$\sum_{\xi,k} (g_k(x) - g_k(\xi)) g'_k f_k(\xi) \psi_k(\xi) \frac{v_k^2}{f'(\xi)}.$$

Si l'on découpe cette somme en deux parties par la suppression des parenthèses, en considérant les degrés connus des fonctions $f, g, \psi$ ainsi que le fait que le plus haut coefficient de $\psi_k$ est égal à $\frac{1}{C_{k-1}}$ et enfin que l'équation $g_k f_k = f_{k-1} + f_{k+1}$ est valide, alors on obtient au moyen des relations d'*Euler* pour le premier terme de

notre somme :

$$[\sum_{\xi,k}(g_k(x)-g_k(\xi))g'_k f_k(\xi)\psi_k(\xi)\frac{v_k^2}{f'(\xi)} = \sum_{k=1}^{n} g_k(x)v_k^2.$$

La deuxième partie devient :

$$-2\sum_{k=1}^{n-1} v_k v_{k+1} \sum_{\xi} \frac{\xi f_k(\xi) f_{k+1}(\xi)}{f_1(\xi)f'(\xi)} g'_k g'_{k+1}. \text{ Or on a : } \frac{f_k(\xi)}{f_1(\xi)} = \psi_k(\xi)$$

et on s'aperçoit maintenant que le degré du numérateur dans $\sum_\xi$ est égal à $n-1$, et que donc d'après les formules d' *Euler* cette somme se réduit à son terme de plus haut degré. En raison des formules d'*Euler* nous pouvons maintenant remplacer $\xi g'_{k+1}$ par $g_{k+1}(\xi)$ et comme

$$g_{k+1}(\xi)f_{k+1}(\xi) = f_k(\xi) + f_{k+2}(\xi)$$

| il reste alors seulement à prendre en compte l'expression : 147|148

$$g'_k f_k(\xi)\psi_k(\xi).$$

Mais le coefficient de son terme de plus haut degré est égal à 1, car de $g_k f_k = f_{k-1} + f_{k+1}$ il suit que $g'_k C_k = C_{k-1}$.

Nous avons donc maintenant le résultat final suivant :

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left(\sum_{k=1}^{n} g'_k f_k(\xi) v_k\right)^2 = \sum_{k=1}^{n} g_k(x)v_k^2 - 2\sum_{k=1}^{n-1} v_k v_{k+1}.$$

Cette transformation particulière d'une somme de carrés :

$$\sum_{h=1}^{n} \frac{x-\xi_h}{f_1(\xi_h)f'(\xi_h)} z_h^2 \text{ en l'expression :}$$
$$\sum_{k=1}^{n} g_k(x)v_k^2 - 2\sum_{k=1}^{n-1} v_k v_{k+1} \text{ grâce à la substitution :}$$
$$z_h = \sum_{k=1}^{n} g'_k f_k(\xi_h) v_k$$

a été donnée en premier par *Jacobi (Crelle Bd. 39*, Seite 290), et d'après *Kronecker*, on appelle ainsi cette forme la forme de *Jacobi*[28].

Passer de la forme de *Jacobi* à une somme de carrés est extrêmement simple.

On a en effet :

$$g_k(x) = \frac{f_{k-1}(x)}{f_k(x)} + \frac{f_{k+1}(x)}{f_k(x)}.$$

Si de là on somme sur $k$, alors on peut toujours réunir

$$\frac{v_{k-1}^2 f_k}{f_{k-1}} + \frac{v_k^2 f_{k-1}(x)}{f_k(x)} - 2v_k v_{k-1}$$

et on s'aperçoit que l'on obtient la somme de carrés suivante :

$$\sum_{k=1}^{n} f_k f_{k-1} \left(\frac{v_{k-1}}{f_{k-1}} - \frac{v_k}{f_k}\right)^2$$

28. Sur le contenu du mémoire de *Jacobi* v. annexe *pag.*...

en posant $v_0 = 0$.

Si nous avons donc exprimé une forme quadratique sous la forme de *Jacobi*, alors la nouvelle transformation en une somme de carrés est extrêmement facile. Si nous appliquons le principe de la conservation des signes, alors il vient :

$$\sum_{\xi} [(x-\xi) f_1(\xi) f'(\xi)] = \sum_{k=1}^{n} [f_k f_{k-1}].$$

148|149

| Mais ici nous allons reconnaître immédiatement le théorème de *Sturm*. $x$ est en effet une grandeur réelle ; si l'on pose maintenant $x = x_1$ et $x = x_2, x_1 < x_2$ et que l'on soustrait les deux équations l'une avec l'autre, alors disparaissent les termes qui se rapportent aux $\xi$ se trouvant à l'extérieur des intervalles $x_1 ... x_2$ et on obtient :

$$2\sum_{\xi} [f_1(\xi) f'(\xi)] = \sum_{k=1}^{n} [f_k(x_2) f_{k-1}(x_2)] - \sum_{k=1}^{n} [f_k(x_1) f_{k-1}(x_1)]$$

la somme se rapportant à tous les $\xi$ se trouvant dans l'intervalle $x_1 ... x_2$. Ainsi nous avons donc démontré que le théorème de *Sturm* était identique au principe d'inertie des formes quadratiques.

§59. Si dans l'équation :

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi) f'(\xi)} \left( \sum_{k=1}^{n} g'_k v_k f_k(\xi) \right)^2 = \sum (g_k x + g_k^0) v_k^2 - 2 \sum v_k v_{k+1}$$

nous séparons les termes avec $x$ de ceux sans $x$, alors il vient, puisque $x$ est totalement quelconque :

$$\sum_{h=1}^{n} \frac{1}{f_1(\xi_h) f'(\xi_h)} z_h^2 = \sum_{k=1}^{n} g'_k v_k^2$$

et

$$\sum_{h=1}^{n} \frac{\xi_h}{f_1(\xi_h) f'(\xi_h)} z_h^2 = -\sum_{k=1}^{n} g_k^0 v_k^2 - 2 \sum_{k=1}^{n-1} v_k v_{k+1}$$

avec l'aide de la substitution :

$$z_h = \sum_{k=1}^{n} g'_k v_k f_k(\xi)$$

c'est-à-dire que nous avons transformé avec l'aide d'une substitution orthogonale (c'est-à-dire une substitution telle qu'elle transforme une somme de carrés, à savoir : $\sum_h \frac{1}{f_1(\xi_h) f'(\xi_h)} z_h^2$ en une autre somme de carrés, à savoir en : $\sum g'_k v_k^2$.) une somme de carrés en une forme de *Jacobi*. Nous pouvons maintenant résoudre en toute généralité le problème qui consiste à transformer une forme quadratique quelconque au moyen d'une substitution orthogonale en une forme de *Jacobi*. Si l'on suppose maintenant pour les fonctions $F_i$ que :

$$\sum_{\xi} \frac{F_i(\xi) F_k(\xi)}{f_1(\xi) f'(\xi)} = \delta_{ik} S_k \text{ et que l'on pose :}$$

$$\sum \frac{\xi F_i(\xi) F_k(\xi)}{f_1(\xi) f'(\xi)} = S_i \mathcal{A}_{ik} \text{ alors on a :}$$

149|150

|

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_{k=1}^{n} v_k F_k(\xi) \right)^2 = \sum_{i,k} (x\delta_{ik} - \mathcal{A}_{ik}) S_i v_i v_k$$

c'est-à-dire que nous avons transformé au moyen de la substitution : $z_h = \sum_{k=1}^{n} v_k F_k(\xi_h)$ une somme de de carrés en une forme quadratique ordinaire. Ici les $\mathcal{A}_{ik}$ et les $S_i$ sont déduits de $\xi_h$ et $F_k$, lesquels sont donnés, au moyen de la relation ci-dessus.

Nous voulons maintenant inversement considérer les $S_i$ et les $\mathcal{A}_{ik}$ comme donnés, de sorte que du côté droit de la dernière équation se trouve une forme quadratique quelconque. Nous voulons maintenant démontrer que les $\xi_h$ et les $F_k$ se laissent ensuite déterminer à partir des $S_i$ et des $\mathcal{A}_{ik}$ tels que l'exige l'équation ci-dessus, de sorte que nous puissions transformer chaque forme quadratique quelconque en cette somme de carrés. Si cette démonstration est faite, alors la décomposition de l'équation selon $x$ fournit les deux équations suivantes :

$$\sum_{k=1}^{n} S_k v_k^2 = \sum_{h=1}^{n} \frac{1}{f_1(\xi_h)f'(\xi_h)} z_h^2 = \sum_{k=1}^{n} g'_k v_k^2$$

$$\sum_{i,k} S_i \mathcal{A}_{ik} v_i v_k = \sum_{h=1}^{n} \frac{\xi_h}{f_1(\xi_h)f'(\xi_h)} z_h^2 = \sum_{k=1}^{n} g_k^0 v_k^2 - 2\sum_{k=1}^{n-1} v_k v_{k+1}$$

c'est-à-dire que nous avons bien transformé la forme quadratique quelconque en une forme de *Jacobi*, sous la transformation simultanée d'une somme donnée de carrés $\sum_{k=1}^{n} S_k v_k^2$ en une autre somme de carrés afin d'établir notre affirmation. Nous remarquons que cette transformation orthogonale d'une forme quadratique quelconque en une forme de *Jacobi* est rationnelle, alors que la première étape, la transformation de la forme quadratique en une somme de carrés introduit l'irrationnel $\xi$, lequel sort de nouveau lors du passage de la somme de carrés en la forme de *Jacobi*.

Nous avons donc maintenant à montrer que pour les $S_i$ et $A_{ik}$ donnés, $\xi_h$ et $F_k$ se laissent déterminer de sorte qu'on ait l'équation :

$$\sum \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum v_k F_k(\xi) \right)^2 = \sum (x\delta_{ik} - \mathcal{A}_{ik}) S_i v_i v_k.$$

On a maintenant la proposition : lorsque l'on transforme linéairement une forme quadratique, le déterminant de la nouvelle forme est égal à celui de l'ancienne, multiplié par le carré du déterminant de la substitution.

| Si nous appliquons cela ici, alors il suit : 150|151

$$|F_k(\xi_h)|^2 \frac{\prod(x-\xi)}{\prod f_1(\xi)f'(\xi)} = |x - \delta_{ik} - \mathcal{A}_{ik}| \prod S_k.$$

Si nous divisons par le coefficient de la plus haute puissance de $x$, alors on obtient :

$$f(x) = \prod(x-\xi) = |x\delta_{ik} - \mathcal{A}_{ik}|.$$

Nous supposons que le déterminant n'a pas de racines égales. Ainsi $f(x)$, c'est-à-dire les $\xi$, est complètement déterminée, il reste donc seulement encore à déterminer les $F_k$ conformes à l'égalité ci-dessus

Nous allons voir maintenant que par ailleurs chaque équation $f(x) = 0$ se laisse transformer sous cette forme de déterminant caractéristique. Dans ce but nous pouvons choisir arbitrairement $f_1(x)$ et déterminer ensuite $F_i$, de sorte que $\sum_{\xi} \frac{F_i(\xi)F_k(\xi)}{f_1(\xi)(\xi)} = \delta_{ik} S_k$. Si on désigne ensuite $\sum_{\xi} \frac{\xi F_i(\xi)F_k(\xi)}{f_1(\xi)(\xi)} = \delta_{ik} S_k$ par $S_i \mathcal{A}_{ik}$, alors

la représentation en déterminant souhaitée pour $f(x)$ est la suivante :

$$f(x) = |x\delta_{ik} - \mathcal{A}_{ik}| \ (i,k = 1,...h).$$

Cette représentation peut, comme $f_1$ et aussi les $F_i$ à une limitation près sont arbitraires, être établie d'une infinité de manières. Maintenant l'égalité $S_i\mathcal{A}_{ik} = S_k\mathcal{A}_{ki}$ est valide ; si donc $\mathcal{A}_{ik} = \mathcal{A}_{ki}$, alors $f(x) = 0$ serait l'équation des perturbations séculaires. Dans les cas particuliers :

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left( \sum_k g'_k v_k f_k(\xi) \right)^2 = \sum g_k v_k^2 - 2\sum v_k v_{k+1}$$

$f(x)$ prend la forme :

$$g'_1...g'_n f(x) = \begin{vmatrix} g_1 & -1 & 0 & ... & ... & ... & 0 \\ -1 & g_2 & -1 & 0 & ... & ... & 0 \\ 0 & -1 & g_3 & -1 & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & ... & ... & -1 & g_{n-1} & -1 \\ 0 & ... & ... & ... & 0 & -1 & g_n \end{vmatrix}$$

151|152 | Il en résulte immédiatement que toutes les fonctions de reste $f_k(x)$ se représentent comme des déterminants partiels de celui-ci lorsque l'on raye les $k$ premières rangées horizontales et verticales.

Après cette parenthèse, reprenons notre étude.
Si l'on désigne par $f_{ik}(x)$ le sous-déterminant de $f(x)$, alors on a :

$$\sum_{i=1}^{n} (x_i\delta_{ik} - \mathcal{A}_{ik}) f_{ri}(x) = \delta_{kr} f(x) \text{ et également :}$$
$$\sum (x\delta_{ik} - \mathcal{A}_{ki}) f_{ir}(x) = \delta_{kr} f(x).$$

Ainsi $\frac{S_1...S_n}{S_k} f_{i,k}(x)$ est le sous-déterminant du système symétrique $|(x\delta_{ik} - \mathcal{A}_{ki})S_k|$, et par conséquent lui-même symétrique ; on a donc :

$$\frac{f_{ir}(x)}{S_r} = \frac{f_{ri}(x)}{S_i} \text{ ou } S_r f_{ri}(x) = S_i f_{ir}(x).$$

Maintenant nous avons besoin de la propriété suivante sur les déterminants : soit $D$ le déterminant des $a_{ik}$ et $\alpha_{ki}$ les facteurs de $a_{ik}$, alors on a tout d'abord $\sum_i \alpha_{hi} a_{ik} = \delta_{hk} D$. Si l'on pose maintenant $\alpha'_{hi} = \alpha_{hi}$pour $h = 1$et $h = 2$, et $\alpha'_{hi} = \delta_{hi}$ pour $h = 3,...n$, alors on a :

$$\sum_i \alpha'_{hi} a_{ik} = \begin{cases} \delta_{h,k} D \text{ pour } h = 1 \text{ et } h = 2 \\ a_{h,k} \text{pour } h = 3,..n. \end{cases}$$

Donc on a d'après la loi multiplicative :

$$|\alpha'_{hi}| D = \left| \begin{array}{cc|cc} D & 0 & ... & ... \\ 0 & D & 0 & ... \\ \hline a_{31} & a_{32} & .. & ... \\ ... & ... & ... & ... \end{array} \right| = D^2 \frac{\partial^2 D}{\partial a_{11} \partial a_{22}}.$$

Et donc :

$$\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} = D \frac{\partial^2 D}{\partial a_{11} \partial a_{22}}.$$

Si on applique cette propriété, il s'ensuit :

$$f_{ir}(\xi)f_{rk}(\xi) - f_{ik}(\xi)f_{rr}(\xi), \text{ da ja } f(x) = D_{x=\xi} = 0.$$

Il en résulte :

$$S_r f_{ri}(\xi)f_{rk}(\xi) = S_i f_{ik}(\xi)f_{rr}(\xi).$$

Donc on a

$$\sum_{\xi} S_r \frac{x-\xi}{f_{rr}(\xi)f'(\xi)} \left(\sum_k f_{rk}(\xi)v_k\right)^2 = \sum_{\xi,i,k} \frac{(x-\xi)S_i f_{ik}(\xi)v_i v_k}{f'(\xi)}.$$

| Si l'on somme d'abord sur les $\xi$, alors on trouve avec l'aide des formules d' *Euler* : 152|153

$$\sum_{\xi} \frac{f_{ik}}{f'(\xi)} = \delta_{ik}$$
$$\sum_{\xi} \frac{\xi f_{ik}}{f'(\xi)} = \mathcal{A}_{ik}.$$

Nous obtenons de là tout simplement pour notre expression $\sum_{i,k}(x\delta_{ik} - \mathcal{A}_{ik})S_i v_i v_k$. En conséquence cette forme quadratique générale est à amener sous la forme voulue en posant :

$$f(x) = |x\delta_{ik} - \mathcal{A}_{ik}| \ , \ f_1(x) = \frac{f_{rr}(x)}{S_r} \text{ et } F_k(x) = f_{rk}(x).$$

Si l'on détermine $f, f_1, F_k$ de cette manière, alors on a :

$$\sum_{i,k}(x\delta_{ik} - \mathcal{A}_{ik})S_i v_i v_k = \sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left(\sum_{k=1}^{n} F_k(\xi)v_k\right)^2.$$

Nous pouvons exprimer aussi le résultat trouvé ainsi : soit données la forme quadratique quelconque $\sum_{i,k} \mathcal{A}_{ik}S_i v_i v_k$ et la somme de carrés $\sum_k S_k v_k^2$. Ensuite nous avons réalisé les transformations simultanées :

$$\begin{cases} \sum_k S_k v_k^2 = \sum_h \frac{S_r}{f_{rr}(\xi_h)f'(\xi_h)} z_h^2 \\ \sum_{i,k} \mathcal{A}_{ik}S_i v_i v_k = \sum_h \frac{\xi_h S_r}{f_{rr}(\xi_h)f'(\xi_h)} z_h^2 \end{cases}$$

avec la substitution :

$$z_h = \sum_k f_{rk}(\xi_h)v_k.$$

Les fonctions $f_{ik}, f', f$ y sont exprimées par $f(x) = |x\delta_{ik} - \mathcal{A}_{ik}|$. Cependant cette transformation, comme nous l'avons déjà remarqué, n'est pas rationnelle car les coefficients contiennent les irrationnels $\xi$. Mais si nous passons maintenant à la forme de *Jacobi*, alors cette irrationalité disparaît. Nous avons :

$$\sum_{\xi} \frac{x-\xi}{f_1(\xi)f'(\xi)} \left(\sum_k v_k g'_k f_k(\xi)\right)^2 = \sum g_k v_k^2 - 2\sum v_k v_{k+1}.$$

Nous obtenons ainsi $\sum_{i,k}(x\delta_{ik} - \mathcal{A}_{ik})S_i v_i v_k$ dans la forme de *Jacobi*, lorsque nous avons identifié la somme de carrés, dans laquelle nous avons transformé cette forme quadratique générale, avec | le côté gauche de la dernière équation grâce à : 153|154

$$\frac{S_r \left(\sum_k f_{rk}(\xi) v_k\right)^2}{f_{rr}(\xi)} = \frac{\left(\sum_k g'_k v_k f_k(\xi)\right)^2}{f_1(\xi)}.$$

Ainsi $\frac{S_r}{f_{rr}(\xi)} f_1(\xi)$ doit être égale au carré d'une fonction entière $\varphi(\xi)$. Si de là on détermine d'après la formule de *Lagrange* une fontion de degré $n-1$ :

$$S_r f_1(\xi) = (\varphi(\xi))^2 f_{rr}(\xi)$$

où $\varphi(\xi)$ est une fonction entière quelconque que nous pouvons supposer réduite au degré $n-1$, alors la substitution :

$$\varphi(\xi) \sum_k f_{rk}(\xi) v_k = \sum_k g'_k v_k f_k(\xi).$$

conduit à la transformation simultanée :

$$\begin{cases} \sum_k S_k v_k^2 = \sum_k g'_k v_k^2 \\ \sum_{i,k} \mathcal{A}_{ik} S_i v_i v_k = -\sum_k g_k^0 v_k^2 + 2 \sum_k v_k v_{k+1} \end{cases}$$

Résumons ce processus de transformation entier. Si à partir de $S_i$ et de $\mathcal{A}_{ik}$ donnés, on construit $f(x) = |x\delta_{ik} - \mathcal{A}_{ik}|$ et de là $f_{ik}$ et $f'$, alors on peut transformer d'après la formule *pag....* au moyen de la substitution $z_h = \sum_k f_{rk}(\xi_h) v_k$ les deux formes données en deux sommes de carrés. Cette transformation n'est pas rationnelle. Si l'on choisit maintenant, pour arriver à la forme de *Jacobi*, une fonction quelconque $\varphi(x)$ de degré $n-1$ et que l'on détermine $f_1(x)$ d'après *Lagrange*, de sorte que

$$S_r f_1(\xi_h) = (\varphi(\xi_h))^2 f_{rr}(\xi_h) \ (h = 1, ...n)$$

et que l'on applique ensuite le procédé de fractions continues à $f$ et $f_1$, lequel nous donne les $f_k$ et les $g_k$, alors on a par la substitution :

$$\varphi(\xi_h) \sum_k f_{rk}(\xi_h) v_k = \sum_k v_k g'_k f_k(\xi_h) \ (h = 1, ...n)$$

la transformation rationnelle simultanée suivante :

$$\begin{cases} \sum_k S_k v_k^2 = \sum_k g'_k v_k^2 \\ \sum_{i,k} \mathcal{A}_{ik} S_i v_i v_k = -\sum_k g_k^0 v_k^2 + 2 \sum_k v_k v_{k+1} \end{cases}$$

154|155 | Comme $\varphi$ possède $n$ constantes arbitraires, alors une telle transformation est dans une variété d'ordre $n$ de substitutions.

Nous pouvons ici immédiatement relier cela à la preuve de la réalité des racines de l'équation pour les perturbations *séculaires*, équation qui est traitée exceptionnellement souvent. En principe cela a à peine besoin de preuve. Les racines sont déjà réelles si tous les $S_i$ sont positifs. Pour l'équation des perturbations séculaires on a $S_i = 1$. Nous obtenons maintenant :

$$\sum_k [S_k] = \sum_\xi [f_{rr}(\xi) f'(\xi)].$$

Si tous les $S$ sont positifs, alors on obtient $n$ du côté gauche. À droite on doit sommer sur tous les réels $\xi$, et cela vaut par conséquent $n$.

§60. Dans le théorème de *Sturm* nous avons eu à considérer la somme :

$$\sum_{x_1 < \xi < x_2} [f_1(\xi) f'(\xi)].$$

Introduisons à la place de $f_1$ et $f$ 2 fonctions $\varphi(z), \psi(z)$ que nous supposons de même degré. Les racines réelles de $\varphi(z) = 0$ seront désignées par $\xi$ et celles de $\psi(z) = 0$ par $\eta$. $y = \varphi(z)$ et $y = \psi(z)$ sont ensuite deux courbes paraboliques de même degré. Maintenant nous avons pour une fonction continue quelconque $F(z)$ la relation :

$$\sum_{z_1<\xi<z_2} [F'(\xi)] = -\frac{1}{2}[F(z_1)] + \frac{1}{2}[F(z_2)] \ (pag....)$$

Si l'on pose maintenant $F(z) = \varphi(z)\psi(z)$, alors il vient :

$$\sum_{z_1<\xi<z_2} [\varphi'(\xi)\psi(\xi)] + \sum_{z_1<\eta<z_2} [\varphi(\eta)\psi'(\eta)] = -\frac{1}{2}[\varphi(z_1)\psi(z_1)] + \frac{1}{2}[\varphi(z_2)\psi(z_2)]$$

Si l'on pose maintenant $\Delta(z) = \begin{vmatrix} \varphi & \psi \\ \varphi' & \psi' \end{vmatrix}$ alors on a :

$$-\sum[\Delta(\xi)] + \sum[\Delta(\eta)] = -\frac{1}{2}[\varphi(z_1)\psi(z_1)] + \frac{1}{2}[\varphi(z_2)\psi(z_2)].$$

Si on pose alors $z_1 = -\infty$, $z_2 = +\infty$, et que l'on somme donc sur tous les $\xi$ et $\eta$, alors le côté droit disparaît car $\varphi$ et $\psi$ sont du même ordre, et nous avons :

$$-\sum_{\xi}[\Delta(\xi)] + \sum_{\eta}[\Delta(\eta)] = 0.$$

| Supposons maintenant que $\varphi$ et $\psi$ sont de degré pair, alors on a : 155|156

$$\sum_{\xi}[\varphi'(\xi)] = \left\{-\frac{1}{2}[\varphi(z_1)] + \frac{1}{2}[\varphi(z_2)]\right\}_{\substack{z_1=-\infty \\ z_2=+\infty}} = 0,$$

et également $\sum_\eta[\psi'(\eta)] = 0$.

Si l'on partage maintenant en deux égalités, il vient :

$$\begin{cases} \displaystyle\sum_{\xi\ ,\ \psi(\xi)>0} [\varphi'(\xi)] = -\sum_{\xi\ ,\ \psi(\xi)<0} [\varphi'(\xi)] \\ \displaystyle\sum_{\eta\ ,\ \varphi(\eta)>0} [\psi'(\eta)] = -\sum_{\eta\ ,\ \varphi(\eta)<0} [\varphi'(\eta)] \end{cases}$$

Nous obtenons de là :

$$\begin{aligned} -\sum[\Delta(\xi)] &= \sum[\varphi'(\xi)\psi(\xi)] = \sum_{\psi(\xi)>0} [\varphi'(\xi)] - \sum_{\psi(\xi)<0} [\varphi'(\xi)] \\ &= 2\sum_{\psi(\xi)>0} [\varphi'(\xi)] = -2\sum_{\psi(\xi)<0} [\varphi'(\xi)]. \end{aligned}$$

Alors $-\frac{1}{2}\sum_\xi[\Delta(\xi)] = -\frac{1}{2}\sum_\eta[\Delta(\eta)]$ est un nombre entier que nous désignerons par $\chi(\varphi, \psi)$. Nous l'appelons la caractéristique du système de fonctions $\varphi$ et $\psi$, elle est égale à la moitié de la somme des signes que prend le déterminant $\Delta$ pour toutes les racines $\xi$ de $\varphi(z) = 0$ ou pour toutes les racines $\eta$ de $\psi(z) = 0$. La caractéristique dépend de l'ordre des deux fonctions $\varphi$ et $\psi$, car de $\begin{vmatrix} \varphi & \psi \\ \varphi' & \psi' \end{vmatrix} = -\begin{vmatrix} \psi & \varphi \\ \psi' & \varphi' \end{vmatrix}$ il s'ensuit $\chi(\varphi, \psi) = -\chi(\psi, \varphi)$. On a aussi :

$$\chi = \frac{1}{2}\sum_{z=\xi}[d(\varphi\psi)] = -\frac{1}{2}\sum_{z=\eta}[d(\varphi\psi)]$$

À partir de

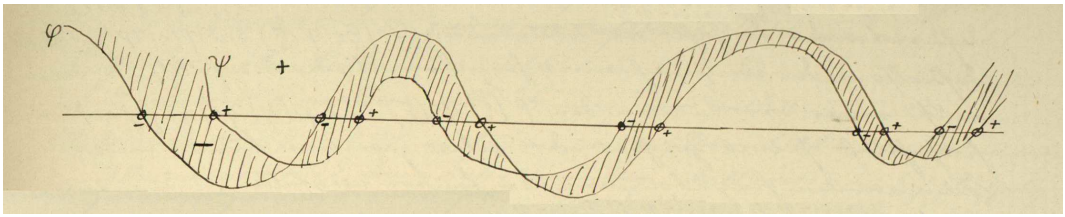
$$\begin{vmatrix}\varphi & \psi\\ \varphi' & \psi'\end{vmatrix}\cdot\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}=\begin{vmatrix}\alpha\varphi+\beta\psi & \gamma\varphi+\delta\psi\\ \alpha\varphi'+\beta\psi' & \gamma\varphi'+\delta\psi'\end{vmatrix}$$

il suit, lorsque l'on comprend par $\alpha,\beta,\gamma,\delta$ des constantes quelconques :

$$\chi(\alpha\varphi+\beta\psi,\gamma\varphi+\delta\psi)=[\alpha\delta-\beta\gamma]\chi(\varphi,\psi).$$

Donc pour un déterminant de substitution positif, $\chi$ ne change pas lors d'une substitution linéaire. Donc en général lors d'une substitution linéaire, au plus seul le signe de la caractéristique change, pas sa valeur.

156|157 D'après les recherches du 45, *pag.*..., si | nous représentons graphiquement $\varphi$ et $\psi$, c'est-à-dire $y=\varphi(z)$, $y=\psi(z)$ la caractéristique a la signification géométrique




suivante : La caractéristique $\chi(\varphi,\psi)$ vaut la moitié de l'excédent des sorties sur les entrées lorsque nous nous déplaçons sur la courbe $y=\varphi(z)$, c'est-à-dire en prenant en compte les $\xi$ et vaut de même la moitié de l'excédent des entrées sur les sorties lorsque l'on parcourt la courbe $y=\psi(z)$, c'est-à-dire en prenant en compte les $\eta$ (dans le cas du dessin on a $\chi=-1$).

Si l'on pose maintenant $\varphi(z)=f(z)$, $\psi(z)=(x-z)f_1(z)$, alors on a :

$$\chi(f(z),(x-z)f_1(z))=\frac{1}{2}\sum[(x-\xi)f_1(\xi).f'(\xi)]$$

Si l'on pose $x=x_1$ et $x=x_2$ et que l'on soustrait, alors on obtient le théorème de *Sturm* (*pag.*...). Le côté gauche du théorème de *Sturm* (*pag.*...) est ainsi le changement de la caractéristique lors du passage de $x_1$ à $x_2$. Ce changement est d'après la décomposition en fraction continue exprimé par les coefficients de $f$ et $f_1$ et détermine, comme nous l'avons vu, le nombre de racines réelles de $f(z)=0$ se trouvant entre $x_1$ et $x_2$, lorsque $f'(\xi)$ et $f_1(\xi)$ ont les mêmes signes.

Voici l'essentiel de cette remarque : on laisse dans les coefficients de l'une des deux fonctions $\varphi(z)=f(z)$ et $\psi(z)=(x-z)f_1(z)$ une variable $x$ et on détermine alors le changement de la caractéristique des deux fonctions qui se forme lors de la modification de cette variable. On s'aperçoit tout de suite que le théorème de *Sturm* est seulement un cas particulier du cas général, car nous pouvons nous mettre maintenant à étudier le problème de la modification de la caractéristique $\chi(\varphi,\psi)$ de $\varphi(z)$ et $\psi(z)$ lorsque nous supposons variable <u>tous</u> les coefficients de $\varphi(z)$ tout comme ceux de $\psi(z)$.

Quand nous savons ensuite comment $\chi$ change, nous pouvons représenter $\chi$ lui-même en partant d'un cas particulier.

157|158 | Lorsque la caractéristique change, alors cela peut seulement se produire soit si le nombre des termes des sommes $\varphi'(\xi)\psi(\xi)$ et $\varphi(\eta)\psi'(\eta)$ reste le même, et que donc ni un $\xi$ ni un $\eta$ n'est ajouté ou supprimé, et donc que le signe d'un des termes des sommes $\varphi'(\xi)\psi(\xi)$ et $\varphi(\eta)\psi'(\eta)$ se modifie, ou soit si de nouveaux termes de la somme s'ajoutent ou sont éliminés en introduisant une discontinuité dans la suite des $\xi$ ou des $\eta$.

Supposons d'abord que le nombre de termes de la somme dans $\chi$ ne change pas, tandis que les coefficients variables se modifient ; la continuité des $\xi$ et des $\eta$ n'est donc pas interrompue. $\chi$ ne peut alors se modifier que si l'on a un quelconque $\varphi'(\xi)\psi(\xi)=0$ et en même temps un quelconque $\varphi(\eta)\psi(\eta)=0$. Les conditions nécessaires et

suffisantes pour une modification de $\chi$ peuvent donc s'exprimer par les 4 éventualités :

$$\begin{aligned}
&1.)\psi(\xi) = 0\ ,\ \varphi(\eta) = 0\\
&2.)\psi(\xi) = 0\ ,\ \psi'(\eta) = 0\\
&3.)\varphi'(\xi) = 0\ ,\ \varphi(\eta) = 0\\
&4.)\varphi'(\xi) = 0\ ,\ \psi'(\eta) = 0
\end{aligned}$$

Supposons dans un second temps que l'on a ajouté un terme à la somme, ou plus généralement, que la continuité de $\xi$ ou de $\eta$ a été perturbée. Si l'on considère une fonction $F(x,t)$ avec le paramètre variable $t$, alors il s'ensuit

$$F(x,t) = 0 \qquad\qquad \frac{dx}{dt} = -\frac{\frac{\partial F}{\partial t}}{F'(x)}.$$

La continuité peut donc seulement être brisée lorsque $F'(x) = 0$. On doit avoir en même temps dans notre cas :

$$\varphi(z) = 0 \text{ et } \varphi'(z) = 0$$

ou

$$\psi(z) = 0 \text{ et } \psi'(z) = 0,$$

pour qu'une interruption de la continuité dans la suite des $\xi$ ou des $\eta$ puisse se produire. Donc soit $\varphi'(\xi)$ soit $\psi'(\eta)$ doit disparaître, mais ce fait est déjà présent dans les quatre possibilités ci-dessus, de sorte que celles-ci contiennent les conditions nécessaires et suffisantes pour une modification de la caractéristique.

Nous nous apercevons que les trois premières possibilités se réduisent, | et que 158|159
pour une même valeur de $x$ on a [29] simultanément $\varphi(z) = 0$ et $\psi(z) = 0$, c'est-à-dire que le résultant $R(\varphi,\psi)$ disparaît.

Maintenant dans le $4^{\text{ième}}$ cas, qui ne survient que lorsque la continuité des $\xi$ ou des $\eta$ s'interrompt, nous pouvons introduire à la place de $\varphi$ et $\psi$ des fonctions linéaires de $\varphi$ et $\psi$ avec un déterminant de substitution positif. En général, ce cas défavorable où ces fonctions linéaires et leurs dérivées s'annulent simultanément ne se produit pas. Posons en effet $\varphi(z) - w\psi(z)$ à la place de $\psi(z)$. Alors on doit avoir simultanément :

$$\begin{aligned}
&\varphi(z) = w\psi(z)\\
&\varphi'(z) = w\psi'(z).
\end{aligned}$$

Mais alors pour ce $z$ l'égalité $\varphi(z)\psi'(z) = \varphi'(z)\psi(z)$ doit avoir lieu, c'est-à-dire que cette valeur particulière de $z$ sera indépendante de $w$. Pour $w = 0$, resp. $= \infty$ on obtient ensuite que simultanément $\varphi$ et $\psi$ doivent disparaître. Il peut donc par conséquent se produire la quatrième possibilité où en effet la continuité peut être interrompue, à savoir pour un $\varphi$ et un $\psi$ particulier, et pour toutes les fonctions linéaires de $\varphi$ et $\psi$, mais seulement lorsque simultanément $\varphi(z) = 0$ et $\psi(z) = 0$. Ainsi nous avons donc réduit toutes les possibilités à $R(\varphi,\psi) = 0$. Cette équation est la condition nécessaire et suffisante pour que la caractéristique de $\varphi$ et $\psi$ se modifie. Nous avons maintenant le genre de changement de la caractéristique à étudier.

§61. Par la disparition d'un terme :

$$\frac{1}{2}[\varphi'(\xi)\psi(\xi)]$$

29. *N.d.T* Il s'agit très certainement d'un $z$ et non pas d'un $x$.

la caractéristique sera augmentée ou diminuée d'une unité selon que pour un tel $\xi$ le terme cité change de $-\frac{1}{2}$ à $+\frac{1}{2}$ ou de $+\frac{1}{2}$ à $-\frac{1}{2}$. Il s'agit donc de décider si $\varphi'(\xi)\psi(\xi)$ passe du positif au négatif, ou bien le contraire.

Si l'on forme maintenant $\sum_\xi \frac{1}{\varphi'(\xi)\psi(\xi)}$, alors dans cette somme prédomine le terme [30], dont nous venons d'étudier la modification. L'expression passera de $+\infty$
159|160 à $-\infty$ ou inversement selon | que le terme en question $\varphi'(\xi)\psi(\xi)$ passera du positif au négatif en passant par le nul ou inversement.

Mais maintenant l'expression complète est une fonction symétrique des $\xi$ et par conséquent une fonction rationnelle des coefficients de $\varphi$ et $\psi$ et en particulier une fonction entière des coefficients de $\varphi$. Si on multiplie par $R = \prod_\xi \psi(\xi)$ alors on obtient :

$$R_1 = \sum_\xi \frac{R(\varphi,\psi)}{\varphi'(\xi)\psi(\xi)}$$

qui est une fonction entière des coefficients de $\varphi$ et $\psi$. On obtient ainsi maintenant deux multiplicateurs $\varphi_1(z)$ et $\psi_1(z)$ de degré $n-1$ de sorte que

$$R = \varphi_1\psi - \varphi\psi_1$$

On peut déterminer $\varphi_1$ et $\psi_1$ par :

$$\varphi_1 = \frac{R}{\psi(\xi)}\ ,\ \psi_1 = \frac{R}{\varphi(\eta)}.$$

La formule de *Lagrange* nous fournit alors :

$$\varphi_1(z) = \sum \frac{R}{\psi(\xi)\varphi'(\xi)}\frac{\varphi(z)}{z-\xi}.$$

En supposant maintenant que dans $\varphi(z)$ et $\psi(z)$ les coefficients de la puissance $n$ sont égaux à 1, donc que :

$$\varphi(z) = z^n + \dots\ ,\ \psi(z) = z^n + \dots$$

nous obtenons

$$\varphi_1(z) = R_1 z^{n-1+\dots}$$
$$\psi_1(z) = R_1 z^{n-1} + \dots$$

c'est-à-dire que $R_1$ est le plus haut coefficient non seulement dans $\varphi_1$ mais aussi dans $\psi_1$. Si maintenant $\varphi$ et $\psi$ ont un facteur commun, alors $\frac{R_1}{R}$ devient infiniment grand et il en est ainsi de même pour $\Phi$ et $\Psi$ [31]. En effet dans ce cas l'égalité $\Phi\psi - \Psi\varphi = 1$ n'est plus valable. Le genre de passage par l'infini, s'il est du positif au négatif ou le contraire, détermine le genre de changement de la caractéristique lors de la traversée du lieu $R(\varphi,\psi) = 0$. Ce changement est ainsi égal à $[R_1(\varphi,\psi)\delta R(\varphi,\psi)]$. Si l'on désigne ainsi pour deux systèmes de valeurs des coefficients de $\varphi$ et $\psi$ les caractéristiques correspondantes par $\chi^{(2)}$ et $\chi^{(1)}$ alors on a :

$$\sum_{R(\varphi,\psi)=0} [R_1(\varphi,\psi)\delta R(\varphi,\psi)] = \chi^{(2)}(\varphi,\psi) - \chi^{(1)}(\varphi,\psi).$$

160|161 Supposons que $\varphi$ et $\psi$ aient des coefficients qui ne dépendent | que d'une variable $x$. $R$ et $R_1$ sont alors des fonctions entières de $x$. Le changement de la caractéristique de

30. Voir Annexe *pag.*...

31. *N.d.T.* Il manque ici la définition de $\Psi$ et $\Phi$. On a en fait $\Phi = \frac{\varphi_1}{R}$ et $\Psi = \frac{\psi_1}{R}$.

$x = x_1$ à $x = x_2$ est alors :

$$\sum_{R(x)=0\ ,\ x_1<x<x_2} [R_1(x)R'(x)].$$

C'est la généralisation du théorème de *Sturm*. La somme $\sum[R_1(x)R'(x)]$ se laisse déterminer comme plus tôt à partir des fonctions de reste de la décomposition en fraction continue de $R$ et $R_1$. Le cas particulier du théorème de *Sturm* résout donc le cas général lui-même.

Supposons maintenant que les coefficients de $\varphi$ et $\psi$ ne dépendent plus de une, mais de $\nu$ variables et nous souhaitons connaître le changement de la caractéristique d'un système de valeurs à un autre. Mais le changement de $u'_1...u_\nu$ à $u''_1...u''_\nu$ peut toujours être réalisé sur une ligne qui passe par les bornes, c'est-à-dire sur une structure de variété plus simple, en imaginant en effet $u_1...u_\nu$ dépendant d'un paramètre $t$ et en passant maintenant de $t'$ à $t''$.

Quand les coefficients de $\varphi$ et $\psi$ dépendent de 3 variables, alors $R = 0$ et $R_1 = 0$ sont des surfaces. La caractéristique de $\varphi$ et $\psi$ change alors seulement lorsque l'on traverse la surface $R = 0$. À un point de l'espace correspond une caractéristique complètement déterminée et tous les points de l'espace qui permettent un passage continu réciproque sans passer par la surface $R = 0$ possèdent la même caractéristique. La surface $R = 0$ doit ainsi fermer au moins autant de régions qu'il y a de caractéristiques. Si l'on pose $\psi = \varphi'$, alors $R(\varphi, \varphi') = 0$ est la surface discriminante et $\chi(\varphi, \varphi')$ la moitié du nombre de racines réelles de $\varphi = 0$. Nous obtenons par exemple pour l'équation biquadratique $\varphi = 0$ les 3 caractéristiques 0,1,2. (Voir *pag....*)

§62. Nous voulons maintenant en venir à représenter les fonctions $R$ et $R_1$ sous la forme d'un déterminant connu.

Nous avons vu que $\frac{R_1}{R}$ est le coefficient de plus haut degré des multiplicateurs $\Phi$ et $\Psi$, pour lesquels on a l'égalité $\Phi\psi - \Psi\varphi = 1$. | Si l'on introduit les $\xi$, alors $\Psi$ est à déterminer de sorte que $\Phi(\xi)\psi(\xi) = 1$. Les $n$ coefficients de $\Phi$ deviennent au moyen de la formule d'*Euler* : 161|162

$$\sum_{(\xi)} \frac{\xi^h \Phi(\xi)\Psi(\xi)}{\varphi'(\xi)} = \delta_{h,n-1}.$$

Posons :

$$\sum_{(\xi)} \frac{\xi^h \Psi(\xi)}{\varphi'(\xi)} = s_k,$$

pour la détermination des coefficients :

$$\begin{aligned} &c_0 s_0 + ... + c_{n-1} s_{n-1} = 0 \\ &... \\ &c_0 s_{n-1} + ... + c_{n-1} s_{2n-2} = 0 \end{aligned}$$

De là on a : $\Phi(z) = \frac{|zs_{g+h} + s_{g+h+1}|\ (g,h=0...n-2)}{|s_{i+k}|\ (i,k=0...n-1)}$ D'où :

$$\frac{R_1}{R} = \frac{|s_{g+h}|}{|s_{i+k}|}.$$

Nous avons ainsi obtenu :

$$|zs_{i+k} - s_{i+k+1}| = \prod_{\alpha} u_\alpha (z - \xi_\alpha) \prod_{\substack{\alpha,\beta \\ (\alpha,\beta=0...n-1)}} (\xi_\alpha - \xi_\beta)^2.$$

Si l'on pose de là $u_k = \frac{\psi(\xi_k)}{\varphi'(\xi_k)}$ alors il s'ensuit que : $|s_{i+k}| = \pm \prod \psi(\xi)$.

Au signe près on a donc $R = \underset{(i,k=0\ldots n-1)}{|s_{i+k}|}$ et $R_1 = \underset{(g,h=0\ldots n-2)}{|s_{g+h}|}$. Nous voulons encore montrer que lorsque l'on pose réellement pour $R$ et $R_1$ ces déterminants, on a :

$$\frac{R_1}{R} = \sum_{\xi} \frac{1}{\varphi'(\xi)\psi(\xi)}. \text{ On a :}$$

$$|s_{g+h}| = \sum_i \prod_\alpha u_\alpha \prod_{\alpha<\beta} (\xi_\alpha - \xi_\beta)^2 = \prod_\alpha u_\alpha \prod_{\alpha<\beta} (\xi_\alpha - \xi_\beta)^2 \sum_k \frac{1}{u_k \varphi'(\xi_k)^2}.$$

Maintenant $u_k = \frac{\psi(\xi_k)}{\varphi'(\xi_k)}$, donc il vient :

$$\frac{|s_{g+h}|}{|s_{i+k}|} = \frac{R_1}{R} = \sum_{(\xi)} \frac{1}{\psi(\xi_k)\varphi'(\xi_k)} \ \textit{q.e.d.}$$

Pour rendre compte du rapport entre la caractéristique et les fonctions de *Sturm*, nous allons revenir sur l'équation :

$$\sum_{(\xi)} \frac{1}{f_1(\xi) f'(\xi)} \left\{ \sum_{k=1}^{n} y_k F_k(\xi) \right\}^2 = \sum_{y=1}^{n} y_k S_k^2$$

où $S_k \delta_{ik} = \sum_\xi \frac{F_i(\xi) F_k(\xi)}{f_1(\xi) f'(\xi)}$ et les $S$ sont les fonctions de *Sturm*. Nous avons ensuite considéré le cas particulier où les $F$ sont les fonctions de reste du développement en fraction continue et nous avons trouvé que ces déterminants sont proportionnels. De là il suit alors que nous pouvons poser : $\begin{vmatrix} s_0 & \ldots & s_m \\ \ldots & \ldots & \ldots \\ s_{m-1} & \ldots & s_{2m-1} \\ x^0 & \ldots & x^m \end{vmatrix}$ et on a alors : $\sum \xi^t . \frac{F_k(\xi)}{f_1(\xi) f'(\xi)} = \delta_{t,m} |s_{i+k}|$.

162|163 | D'après cela on a donc :

$$S_{n-m} = \sum_{\xi} \frac{(F_{n-m}(\xi))^2}{f_1(\xi) f'(\xi)} = |s_{g+h}||s_{i+k}| \ (i,k = 0, \ldots m \ g,h = 0, \ldots m-1)$$

où $s_g = \sum \frac{\xi^g}{f_1(\xi) f'(\xi)}$. Si l'on suppose que $f(x) = \varphi(x)$ et que $f_1(x) = \frac{1}{\psi(x)}$ alors $s_g$ est définie comme *pag.*... . Notre caractéristique devient donc :

$$\chi(\varphi, \psi) = \frac{1}{2} \sum [\varphi'(\xi)\psi(\xi)] = \frac{1}{2} \sum_{k=1}^{n} [S_k] = \frac{1}{2} \sum [f'(\xi) f_1(\xi)].$$

Or nous avions trouvé pour le changement de la caractéristique :

$$\chi'' - \chi' = \frac{1}{2} \sum_{k=1}^{n} [S_k''] - \frac{1}{2} \sum_{k=1}^{n} [S_k'] = \sum [R_1(\varphi, \psi \delta R(\varphi, \psi)],$$

la dernière somme s'étendant à tous les endroits où $R = 0$. Il apparaît donc un changement de signe dans la suite de *Sturm* seulement lorsque $R = 0$. Maintenant on a :

$$R = \underset{(i,k=0\ldots n-1)}{|s_{i+k}|}, \ R_1 = \underset{(g,h=0\ldots n-2)}{|s_{g+h}|}$$

Si nous définissons encore

$$R_2 = \underset{(g,h=0\ldots n-3)}{|s_{g+h}|}, \; R_3 = \underset{(g,h=0\ldots n-4)}{|s_{g+h}|} \quad etc.$$

alors les déterminants $R, R_1, R_2 \ldots$ forment une suite de *Sturm* $F$, qui a pour propriété : la somme de ses signes ne pourra changer que si $R$ change son signe.

Nous nous rendons compte ici de l'analogie avec les fonctions de reste de *Sturm* $f, f_1, f_2, \ldots$ chez lesquelles nous avons trouvé que le nombre de changements de signes de cette suite ne peut changer que lorsque $f$ change son signe. Cela nous a conduit pour les fonctions de reste aux équations $f_{k-1} - g_k f_k + f_{k+1} = 0$. On peut ainsi supposer qu'il existe aussi entre les $R$ des équations analogues. Et en effet on trouve

$$P_k^2 R_{k-1} - Q_k R_k + P_{k+1}^2 R_{k+1} = 0.$$

§63. Nous pouvons maintenant généraliser les recherches effectuées jusqu'à présent, dans lesquelles nous avons considéré les points d'intersection des courbes $y = \varphi(z)$ et $y = \psi(z)$ avec la courbe $y = 0$, c'est-à-dire l'axe, en considérant trois courbes fermées $f = 0$, $\varphi = 0$ et $\psi = 0$. Nous recherchons par exemple les points d'intersection de $f = 0$ et $\varphi = 0$ et nous regardons si à cet endroit $\psi$ est positive ou négative, c'est-à-dire que nous cherchons des systèmes de valeurs $x, y$, pour lesquels on a les deux égalités $f = 0$ et $\varphi = 0$ et simultanément l'inégalité $\psi < 0$ resp. $> 0$. Par
| courbes fermées on entend les parties du plan à l'intérieur desquelles il n'y a aucun 163|164
point infiniment loin. L'intérieur est donc caractérisé par $f < 0$. En règle générale on parcourt ainsi une courbe de sorte que l'intérieur reste à gauche, mais cette considération se limite aux figures planes. Nous voulons ainsi établir un critère purement analytique. Construisons le déterminant :

$$\Delta = \begin{vmatrix} f & \varphi & \psi \\ f_1 & \varphi_1 & \psi_1 \\ f_2 & \varphi_2 & \psi_2 \end{vmatrix}$$

Si nous allons maintenant sur la courbe $f = 0$, alors nous voulons trouver la détermination de la poursuite suivante. À un point d'intersection de $f$ et $\varphi$ nous voulons aller en avant sur $f$ dans une direction telle que,lorsque $\Theta$ est une courbe quelconque passant par le point en question, on a l'expression :

$$\frac{\partial \begin{vmatrix} f & \Theta & \psi \\ f_1 & \Theta_1 & \psi_1 \\ f_2 & \Theta_2 & \psi_2 \end{vmatrix}}{\partial \psi} d\Theta > 0$$

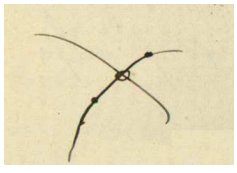
En d'autres termes, on doit avoir :

$$(f_1\Theta_2 - f_2\Theta_1)d\Theta > 0$$

Tout d'abord on doit montrer que ce principe de parcours est indépendant du choix de la courbe $\Theta$.

Ainsi on a :

$$d\Theta = \Theta_1 dx + \Theta_2 dy$$
$$0 = f_1 dx + f_2 dy$$



Notre détermination doit donc fixer dans laquelle

des deux directions de $\frac{dy}{dx}$ nous devons avancer sur $f = 0$. Maintenant on a $d_\Theta = (f_1\Theta_2 - f_2\Theta_1)\frac{dy}{f_1}$.

Notre détermination se réduit donc à ce que $\frac{dy}{f_1}$ ou $f_1 dy$ soit positif.

Ainsi $\Theta$ n'est plus présent. On obtient naturellement le dernier critère aussi en posant $\Theta = y$. Les deux figures :




164|165

| montrent que lors du parcours ainsi défini l'intérieur reste toujours à gauche.

Nous considérons maintenant deux courbes $U = 0$ et $V = 0$ et nous parcourons $U$ dans le sens défini par notre critère. On doit ainsi avoir :

$$(U_1V_2 - U_2V_1)dV > 0$$

Il est ensuite clair que lors de ce genre de parcours sur $U = 0$ on a l'égalité : $\sum_{U=0=V}[dV] = 0$ lorsque $V = 0$ est une courbe fermée, car on doit bien sortir de l'intérieur de $V = 0$ aussi souvent que l'on y entre quand on parcourt toute la courbe $U = 0$. Il en résulte :

$$\sum_{U=0=V}[U_1V_2 - U_2V_1] = 0$$

et dans cette équation il n'est plus du tout question d'un principe de parcours.

Nous supposons d'abord que les courbes $U = 0$ et $V = 0$ sont telles que le déterminant fonctionnel n'est pas nul.

On pose alors $f = U$, $\varphi\psi = V$, de sorte que l'on a :

$$U_1V_2 - U_2V_1 = f_1(\varphi_2\psi + \varphi\psi_2) - f_2(\varphi_1\psi + \varphi\psi_1)$$

$$= \psi(f_1\varphi_2 - f_2\varphi_1) + \varphi(f_1\psi_2 - f_2\psi_1) = \psi\Delta_3 - \varphi\Delta_2,$$

lorsque l'on désigne par $\Delta_2$ et $\Delta_3$ les sous-déterminants de $\varphi$ et $\psi$ dans $\Delta$.

On a donc :

$$\sum_{f=0,\varphi\psi=0}[\psi\Delta_3 - \varphi\Delta_2] = 0. \text{ ou :}$$

$$\sum_{f=0,\varphi=0}[\psi\Delta_3] - \sum_{f=0,\psi=0}[\varphi\Delta_2] = 0.$$

De cette façon nous obtenons, comme $\Delta = f\Delta_1 + \varphi\Delta_2 + \psi\Delta_3$ :

$$\sum_{f=0=\varphi}[\Delta] = \sum_{\varphi=0=\psi}[\Delta] = \sum_{\psi=0=f}[\Delta].$$

Voici la généralisation des résultats précédents :

$$\sum[\Delta(\xi)] = \sum[\Delta(\eta)].$$

165|166 Nous montrons maintenant que $\sum[\Delta]$ est toujours un nombre pair. Pour $U = f$ | et $V = \varphi$ il suit :

$$\sum_{f=0=\varphi}[\Delta_3] = 0 = \sum_{\substack{f=0=\varphi \\ \psi>0}}[\Delta_3] + \sum_{\substack{f=0=\varphi \\ \psi<0}}[\Delta_3].$$

De plus on a :

$$\sum_{f=0=\varphi}[\Delta]=\sum_{f=0=\varphi}[\Delta_3\psi]=\sum_{\substack{f=0=\varphi\\ \psi>0}}[\Delta_3]-\sum_{\substack{f=0=\varphi\\ \psi<0}}[\Delta_3].$$

Et ainsi :

$$\sum_{f=0=\varphi}[\Delta]=\sum_{f=0=\varphi}[\psi\Delta_3]=2\sum_{\substack{f=0=\varphi\\ \psi>0}}[\Delta_3]=-2\sum_{\substack{f=0=\varphi\\ \psi<0}}[\Delta_3].$$

Nous posons maintenant :

$$\chi(f,\varphi,\psi)=-\frac{1}{2}\sum_{f=0=\varphi}[\Delta]$$

et nous désignons par $\chi$, qui est donc un nombre entier, la caractéristique de $f,\varphi,\psi$. De $d\varphi=(\varphi_2 f_1-\varphi_1 f_2)\frac{dy}{f_1}=\frac{\Delta_3 dy}{f_1}$, il résulte, comme lors du parcours de $f$, que $\frac{dy}{f_1}$ reste positif :

$$\chi=-\frac{1}{2}\sum_{f=0=\varphi}[\Delta]=-\frac{1}{2}\sum_{f=0=\varphi}[\psi\Delta_3]=-\frac{1}{2}\sum_{f=0=\varphi}[\psi d\varphi]$$

donc :

$$\chi=-\frac{1}{2}\sum_{f=0=\varphi}[d\varphi\psi].$$

Lorsque l'on parcourt donc la courbe $f=0$ dans le sens donné et que l'on considère le point d'intersection avec $\varphi=0$, alors une telle intersection avec $\varphi\psi$ est à désigner comme entrée ou sortie selon que $d\varphi\psi$ est négatif ou positif. La dernière formule dit bien que la caractéristique du système $f,\varphi,\psi$ est égale à la moité de l'excès des entrées sur les sorties. Comme nous avons 3 courbes, nous pouvons faire ces considérations de 6 façons différentes, car en partant de l'une des courbes, nous pouvons considérer son intersection avec l'une ou l'autre des deux courbes restantes. D'après l'invariance de la caractéristique l'excès des entrées sur les sorties doit être le même dans les 6 cas.

La caractéristique $\chi(f,\varphi,\psi)$ peut aussi s'écrire :

$$\chi(f,\varphi,\psi)=\sum_{\substack{\varphi=0=\psi\\ f<0}}[\Delta_1]=\sum_{\substack{\varphi=0=\psi\\ f<0}}[\varphi_1\psi_2-\varphi_2\psi_1]$$

Si on applique cela au système $f\Delta_1,\varphi,\psi$ , alors il vient :

| 166|167

$$\chi(f\Delta_1,\varphi,\psi)=\sum_{f\Delta_1<0}[\Delta_1]=\sum_{\substack{\varphi=0=\psi\\ f>0}}(-1)+\sum_{\substack{\varphi=0=\psi\\ f<0}}(+1)$$

c'est-à-dire que la caractéristique de ce système donne l'excès des points d'intersection de $\varphi$ et $\psi$ à l'intérieur de $f=0$ sur ceux à l'extérieur de $f=0$.

§64. Si l'on applique maintenant ces études au cas particulier des fonctions $\varphi$ et $\psi$ que l'on forme en décomposant la fonction $\Theta(x+yi)$ en sa partie réelle et imaginaire. Nous posons :

$$2\Theta(x+yi)=\varphi(x,y)+i\psi(xy)$$

De là, il suit :

$$\varphi = \Theta(x+yi) + \Theta(x-yi)\ ,\ \psi(xy) = i\Theta(x-yi) - i\Theta(x+yi)$$
$$\varphi_1 = \Theta'(x+yi) + \Theta(x-yi)\ ,\ \psi_1 = i\Theta'(x-yi) - i\Theta'(x+yi)$$
$$\varphi_2 = -\psi_1\ ,\ \psi_2 = \varphi_1$$

$$\Delta_1 = \varphi_1\psi_2 - \varphi_2\psi_1 = \varphi_1^2 + \psi_1^2.$$

Donc :

$$\chi = \sum_{\substack{\varphi=0=\psi \\ f<0}} [\varphi_1\psi_2 - \varphi_2\psi_1] = \sum_{\substack{\varphi=0=\psi \\ f<0}} (1).$$

Mais à ce propos il faut faire la remarque suivante. $\varphi$ et $\psi$ ne sont pas, comme nous l'avons supposé, des courbes fermées. Mais toutes les conclusions sur ce qui est à l'intérieur de $f = 0$ restent valables lorsqu'on les referme par des bouts de courbes qui passent complètement à l'extérieur de $f = 0$.

Ensuite la caractéristique donne donc le nombre de points d'intersection de $\varphi = 0$ et $\psi = 0$, c'est-à-dire les zéros de $\Theta(x+yi) = 0$ qui se trouvent à l'intérieur de $f = 0$.

Si l'on se déplace maintenant sur la courbe $\varphi = 0$, alors on a (*pag....*)

$$\chi = -\frac{1}{2}\sum_{\varphi=0=\psi} [df\psi]\,;\ \text{ainsi :}$$

$$df\psi = \frac{\partial f\psi}{\partial x}dx + \frac{\partial f\psi}{\partial y}dy = f(\psi_1 dx + \psi_2 dy) + \psi(...).$$

Comme nous nous déplaçons en avant sur $\varphi = 0$, on a :

$$\varphi_1 dx + \varphi_2 dy = 0\ ,\ \text{par conséquent :}$$

$$df\psi = f(\varphi_1\psi_2 - \varphi_2\psi_1)\frac{dy}{\varphi_1} + \psi(...).$$

167|168 | Nous comptons maintenant les points d'intersection de $\varphi$ avec $\psi$, la seconde partie multipliée avec $\psi$ étant donc supprimée. De plus $\frac{dy}{\varphi_1}$ est positif d'après notre principe de parcours, et on a également $\varphi_1\psi_2 - \varphi_2\psi_1 = \varphi_1^2 + \psi_1^2$.

Donc on a : $\chi = -\frac{1}{2}\sum_{\varphi=0=\psi}[f]$. Il en résulte qu'à l'intérieur de $f$, donc pour $f < 0$, se trouvent seulement des points d'entrée. Le produit $f\psi$ décide en même temps de l'intérieur et de l'extérieur.

§65. Nous voulons maintenant par une illustration parvenir à une nouvelle conception de la caractéristique. Pour faciliter notre façon de voir, nous choisissons également un exemple particulier qui sera représenté par la figure suivante :

Nous représentons alors la figure entière en posant :

$$\xi = \varphi(x,y)\ ,\ \eta = \psi(x,y).$$

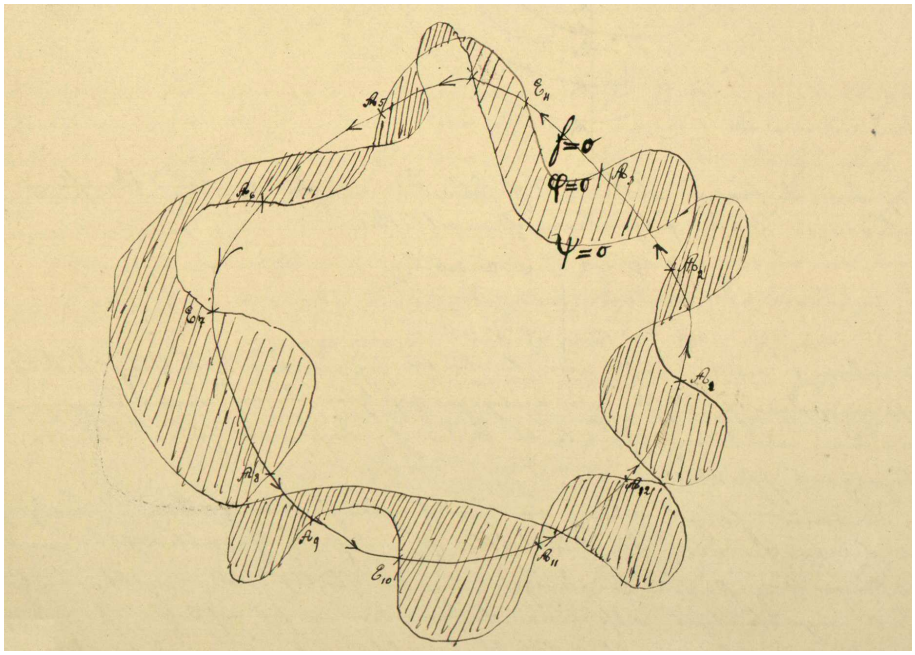
Ensuite $f(x,y) = 0$ devient une courbe $F(\xi,\eta) = 0$.

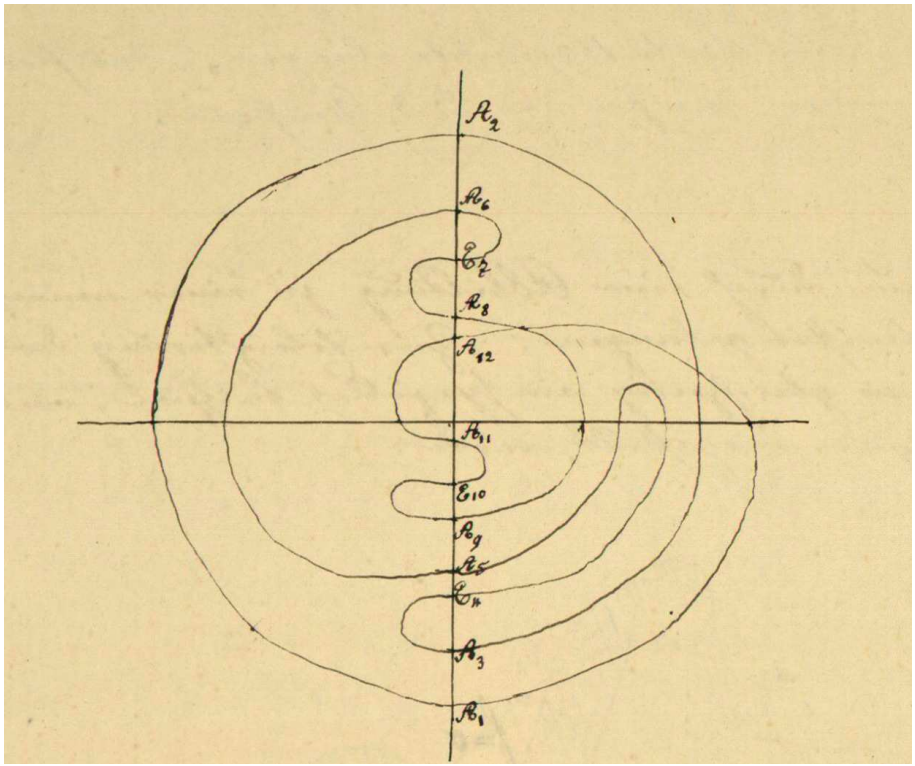
168|169 | La courbe $F(\xi,\eta)$ présente ensuite approximativement la forme ci-dessous :

Comme dans notre figure la partie se trouvant à l'intérieur et à l'extérieur se représente de la manière suivante :

| | |
|---|---|
| $\varphi\psi < 0$ | $\varphi\psi > 0$ |
| $\varphi\psi > 0$ | $\varphi\psi < 0$ |

alors on se rend compte que dans la figure une entrée est marquée par le fait que nous coupons l'axe des $y$ dans le sens contraire des aiguilles d'une montre et une sortie



est marquée par le fait que nous le coupons dans le sens des aiguilles d'une montre.

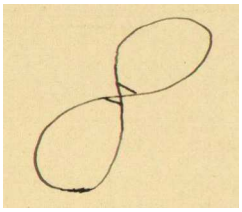
Il en résulte ainsi, comme $\chi$ est bien égal à la moitié de l'excès des entrées sur les sorties, que la caractéristique indique justement le bon nombre de lacets autour du point zéro. La courbe $f = 0$ représente donc un fil qui se referme sur lui-même et qui tourne autour d'une pointe imaginaire fixée au point zéro autant de fois que l'indique la caractéristique. Nous supposons le chemin exprimé en coordonnées polaires par $\xi = r\cos v$ , $\eta = r\sin v$, ensuite $v$ s'étend de 0 à un multiple de $2\pi$ et ce multiple est la | caractéristique. Ainsi on a $\chi(f, \varphi, \psi) = \frac{1}{2\pi}\int dv$, l'intégrale s'étendant à toute la 169|170

courbe. Maintenant on a :

$$v = arctg\ \frac{\eta}{\xi} = arctg\ \frac{\psi(x,y)}{\varphi(x,y)},\ \text{donc :}$$

$\chi = \int d\ arctg\ \frac{\psi(x,y)}{\varphi(x,y)}$, l'intégrale s'étendant à tous les $x, y$, pour lesquels $f = 0$. $f(x,y) = 0$ est une courbe fermée que nous voulons également exprimer en coordonnées polaires : $x = \rho\cos w$ , $y = \rho\sin w$. Nous obtenons ensuite $f(\rho\cos w, \rho\sin w) = 0$ et de là : $\rho = \Theta(w)$.

Nous pouvons maintenant supposer que la courbe $f(xy) = 0$ ne se traverse pas elle-même, car sinon on aurait deux courbes fermées :



Nous pouvons supposer de plus que $\rho$ est une fonction univoque de $w$. Nous pouvons donc toujours découper les régions en des secteurs dans lesquels cela a lieu. S'il existe plusieurs de ces secteurs, alors nous les traitons l'un après l'autre.

Nous pouvons donc supposer que nous avons maintenant à le faire avec un secteur à partir du fait que la représentation :

$$x = \rho\cos w\ ,\ y = \rho\sin w\ ,\ \rho = \Theta(w)$$

produit une fonction univoque $\rho$ de $w$. Ensuite $\varphi$ et $\psi$ se changent simultanément en des fonctions de $w$, à savoir :

$$\varphi(x,y) = \Phi(w)\ ,\ \psi(x,y) = \Psi(w)$$

et nous obtenons :

$$\chi = \frac{1}{2\pi}\int \frac{\varphi d\psi - \psi d\varphi}{\varphi^2 + \psi^2} = \frac{1}{2\pi}\int_0^{2\pi} \frac{\Phi\Psi' - \Psi\Phi'}{\Phi^2 + \Psi^2} dw$$

# Annexe B

# Transcription et traduction du chapitre du cours de 1890-1891 sur la théorie des caractéristiques

Nous allons donner ici la transcription et la traduction du sommaire et du chapitre sur la théorie des caractéristiques du manuscrit du cours de Kronecker du semestre d'hiver 1890-1891 ([Kronecker, 1891a]) intitulé *Die Theorie der algebraischen Gleichungen*, et dont l'original se trouve dans le fond *Hensel* de la bibliothèque de Strasbourg [1]. La version sur laquelle nous avons travaillé est un fichier PDF constitué de 666 pages.

Comme celui du semestre d'hiver de 1880-81, ce manuscrit est écrit en *Kurrentschrift*, et comme précédemment, nous avons identifié les mots figurant en caractères latin en les notant en italique. Chaque début et fin de page sont représentés par le symbole "|" et on trouvera dans la marge la pagination correspondante. Dans ce manuscrit, le cours est rédigé sur la page de droite, celle de gauche servant aux schémas. On peut aussi noter la présence sur ces pages de gauche d'annotations qui apportent correction et références à d'autres manuscrits du même cours. Les deux dernières pages sont d'ailleurs consacrées à la reconstitution d'un sommaire prenant en compte ce manuscrit et celui de 1890-91. La pagination utilisée est celle indiquée dans le manuscrit.

1. Une version numérique est consultable à l'adresse `http://docnum.u-strasbg.fr/cdm/compoundobject/collection/coll7/id/53004/rec/15`

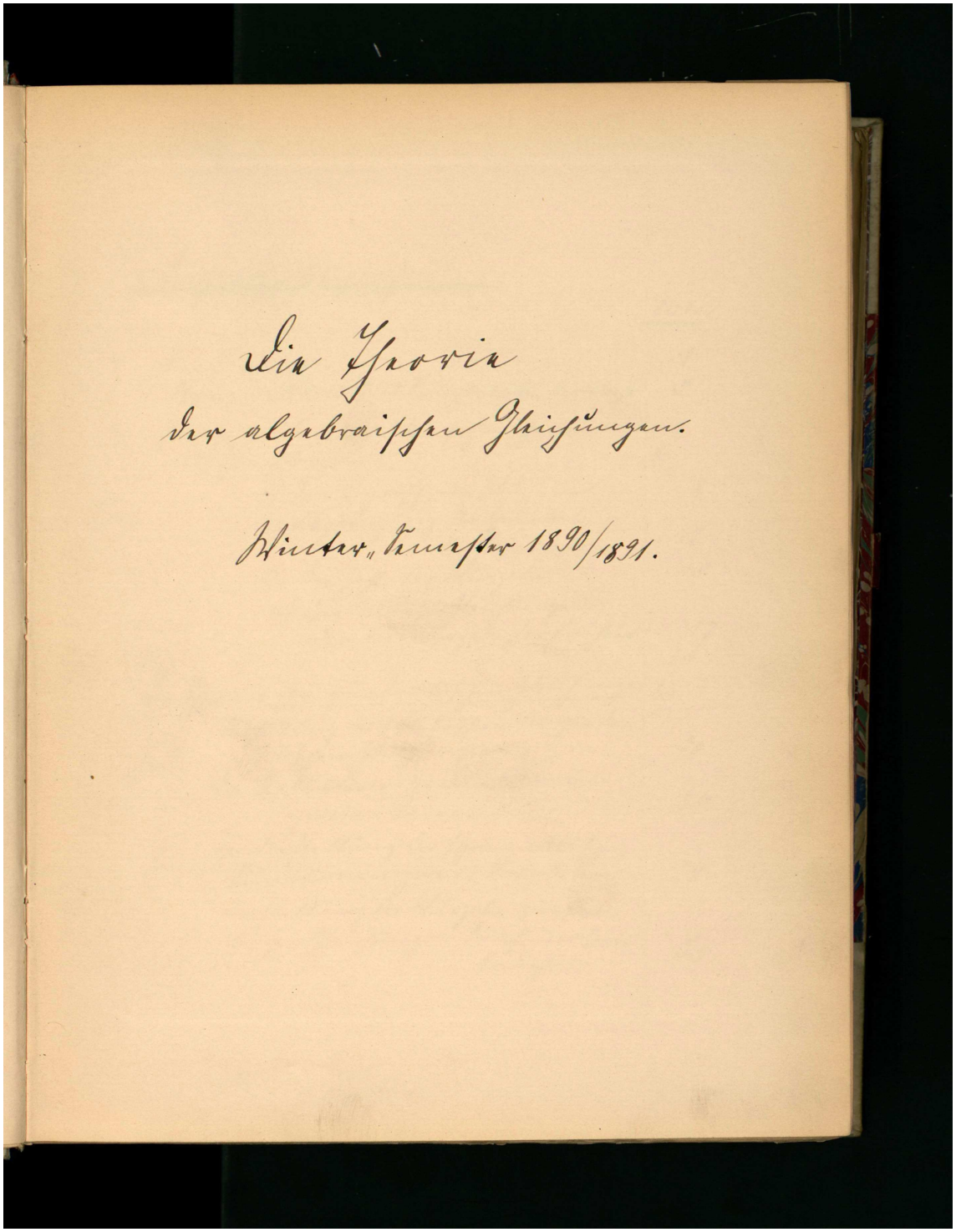



FIGURE 1 – Page de garde du manuscrit de 1890-91 ([Kronecker, 1891a])

## B.1 Sommaire du cours de 1890-1891

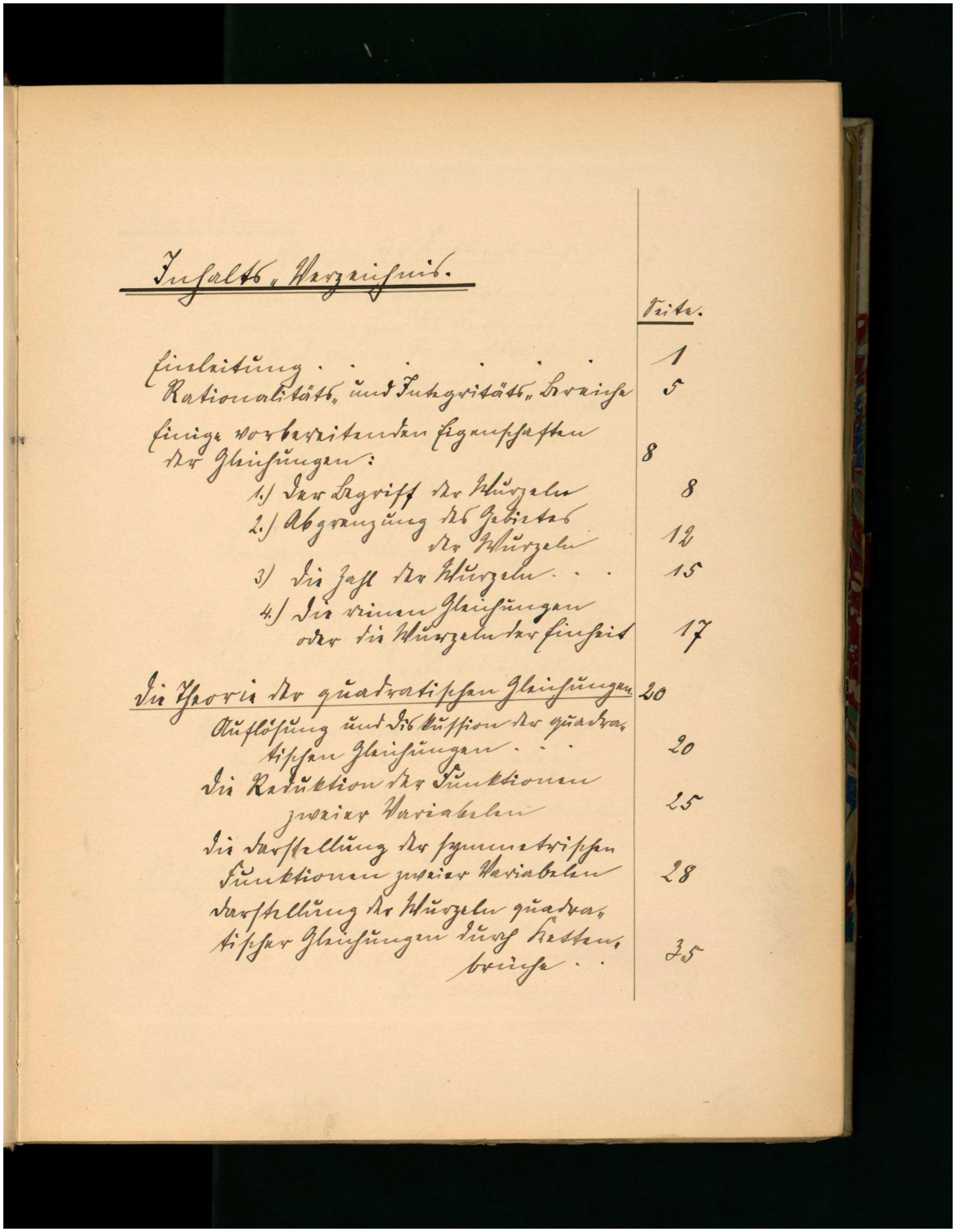



FIGURE 2 – Page 2 du manuscrit de 1890-91 ([Kronecker, 1891a])

## Annexe : Sommaire des *Vorlesungen über der algebraischen Gleichungen* du semestre d'hiver 1890/91

Inhalts Verzeichnis / Liste des contenus

Einleitung / Introduction (1)
Rationalitäts und Integritäts Bereiche / Domaines de rationalité et d'intégrité (5)
Einige vorbereitenden Eigenschaften der Gleichungen / Quelques propriétés préliminaires sur les équations (8)

1. Der Begriff der Wurzeln / La notion de racine (8)
2. Abgrenzung der Gebietes (der Wurzeln) / Délimitation du domaine (des racines) (12)
3. Die zahl der Wurzeln / Le nombre de racines (15)
4. Die reinen Gleichungen oder die Wurzeln der Einheit / Les équations pures ou les racines de l'unité (17)

Theorie der quadratischen Gleichungen / La théorie des équations quadratiques (20)

1. Auflösung und Diskussion der quadratischen Gleichung / Résolution et discussion de l'équation quadratique (20)
2. Die Reduktion der Funktionen zweier Variablen / La réduction des fonctions de deux variables (25)
3. Die Darstellung der symmetrischen Funktionen zweier Variablen / La représentation des fonctions symétriques de deux variables (28)
4. Darstellung der Wurzeln quadratischer Gleichungen durch Kettenbrüche / Représentation des racines des équations quadratiques par des fractions continues (35)

Digression : Theorie der Kettenbrüche / Digression : Théorie des fractions continues (37)

1. Einleitendes, Definitionen u.s.w. / Préliminaires, définitions, etc. (37)
2. Darstellungen der Kettenbruches und seiner Partialbrüche / Représentations des fractions continues et de leurs fractions partielles (41)
3. Die Grenzen der Annäherung durch Kettenbrüche / Les limites de convergence avec des fractions continues (45)
4. Der Zusammenhang der Kettenbrüche mit den quadratischen Gleichungen / Le rapport des fractions continues avec les équations quadratiques (48)
5. Über die Annäherung rationaler Brüche an die Wurzeln ganzzahligen quadratischen Gleichungen / Sur la convergence des fractions continues dans les racines des équations quadratiques à coefficients entiers (51)
6. Divergenz des einer quadratischen Gleichung entsprechenden Kettenbruches für den Fall imaginärer Wurzeln / Divergence d'une fraction continue correspondant à une équation quadratique pour le cas des racines imaginaires (52)
7. Nachtrag über die arithmetische Behandlung des Irrationalen / Supplément sur le traitement arithmétique des irrationnels (56)
8. Die « Abstandsfunktionen » für einige Irrationalitäten / Les « fonctions distance » pour certains irrationnels (61)

Abschluß der Theorie der Quadratischen Gleichungen / Conclusion de la théorie des equations quadratiques

Auflösung derselben durch beliebige zweiwertige Funktionen / La résolution de celles-ci par des fonctions bivalentes quelconques (66)

Die Theorie der kubischen Gleichungen / La théorie des équations cubiques (69)

1. Die Cardanische Formel / La formule de Cardan (69)

2. Die trigonometrische Auflösung oder die Auflösung durch Bogenteilung / La résolution trigonométrique ou la résolution par des parties d'arc (75)
3. Die Einteilung der kubischen Gleichungen nach der Anzahl ihrer reellen Wurzeln / Le classement des équations cubiques par le nombre de ses racines réelles (77)
4. Die rationalen Funktionen von drei Variablen / Les fonctions rationnelles de trois variables (80)
5. Die cyclischen Funktionen im besonderen / Cas particulier des fonctions cycliques (83)
6. Die Adjunktion der cyclischen Funktionen zum Rationalitätsbereich / L'adjonction des fonctions cycliques au domaine de rationalité (85)
7. Die Abelschen Gleichungen / Les équations abéliennes (92)
   (a) Definition der Abelschen Gleichungen / Définition des équations abéliennes (92)
   (b) Die allgemein Form / La forme générale (95)
   (c) Die Komposition der Abelschen Gleichungen dritten Grades / La composition des équations abéliennes du troisième degré (98)

Die Theorie der algebraischen Gleichungen vom n$^{\text{ten}}$ Grades / La théorie des équations algébriques du n$^{\text{ième}}$ degré (121)

1. Die Reduktion der rationalen Funktionen $n$ Variabler vermittelst der aus einer Gleichung abgeleiteten Gleichungen / La réduction des fonctions rationnelles de $n$ variables au moyen d'équations dérivée d'une équation (121)
2. Folgerungen aus der Reduktion / Conséquences de la réduction (131)
3. Eine Eigenschaft der ganzzahligen Gleichungen, deren Wurzeln sämtlich reell und ihrem absoluten Betrage nach nicht größer als 2 sind / Une propriété des équations à coefficients entiers dont les racines sont complètement réelles et leur valeur absolue n'est pas plus grande que 2 (134)
4. Die angenäherte Berechnung der Wurzeln / Le calcul approché des racines (136)
5. Die ganzen, im besonderen die symmetrischen Funktionen von $n$ Variabelen, nach Grad, dimension, Ordnung, Darstellung (Typen) / La représentation complète, en particulier les fonctions symétriques de $n$ variables, d'après le degré, la dimension, l'ordre, la présentation (Type) (140)
6. Beispiel für die *Gauss-Waring*sche Methode (zur Reduktion symmetrischer Funktionen) / Exemple pour la méthode de *Gauss-Waring* (pour la réduction des fonctions symétriques) (147)
7. Die Unabhängigkeit der symmetrischen Funktionen $(f_1 ... f_n)$ von einander / L'indépendance des fonctions symétriques $(f_1 ... f_n)$ (148)
8. Die diskriminante und die mit ihr zusammenhangenden Invarianten / Le discriminant et les invariants qui lui sont liés (153)
9. Die *Newton*schen und die *Euler*schen Formeln / Les formules de *Newton* et d'*Euler* (165)
10. Der *Sturm*sche Satz / Le théorème de *Sturm* (170)
    (a) Beweis (mit vorangehenden Determinanten Entwicklungen, welche mit einer Darstellung der Diskriminante beginnen) / Démonstration (avec les développements sur les déterminants précédents qui commencent avec une représentation du discriminant) (170)
    (b) Digression über die Bedeutung der verallgemeinerten *Sturm*schen Funktionen / Digression sur la signification des fonctions généralisées de *Sturm* (178)
    (c) Zweite Herleitung des *Sturm*schen Satzes / Deuxième déduction du théorème de *Sturm* (180)

(d) Dritte Herleitung (mit dem arithmetisch präzisierten Begriff der Wurzeln) / Troisième déduction (avec une notion arithmétique précise de racine) (184)

(e) Vierte Ableitung, durch Kettenbruchsentwicklung / Quatrième démonstration, par les développements en fractions continues (188)

(f) Diskussion derselben (Bestimmung einer Bruches mit gegebenen Werten) / Discussion (Détermination d'une fraction avec des valeurs données) (190)

(g) Rückblick auf den Charakter der verallgemeinerten *Sturm*schen Funktionen / Retour sur le caractère des fonctions généralisées de *Sturm* (199)

Über die Resultante zweier ganzen Funktionen / Sur le *résultant* de deux fonctions entières (202)

Über die Charakteristik eines Funktionensystems / Sur la théorie des caractéristiques d'un système de fonctions (208)

1. Einführende Betrachtung / Considérations introductives (208)
2. Die Charakteristik eines Systems von zwei Funktionen einer Variabelen / La caractéristique d'un système de deux fonctions d'une variable (211)
3. Anwendung auf die Gleichung vierten Grades / Application aux équations du quatrième degré (224)
4. Charakteristik eines Systems von drei Funktionen von je zwei Variabeln / Caractéristique d'un système de trois fonctions de deux variables (229)
5. Anwendung der Resultate zur Herleitung des *Gauss*'schen Fundamentalsatzes der Algebra und des Integrals von *Cauchy* / Application des résultats à la déduction du théorème fondamental de l'algèbre de *Gauss* et des intégrales de *Cauchy* (241)
6. Klärung des Begriffes der Charakteristik, der reellen Wurzeln, wie der zerlegung einer ganzen Funktion in reelle Faktoren / Clarification de la notion de caractéristique, des racines réelles et de la décomposition d'une fonction entière en fonctions réelles (247)
7. Über ein Eigenschaft der Gebiete, in welchen die Funktionen mit gleicher Charakteristik liegen / Sur une propriété du domaine dans lequel se trouvent les fonctions ayant la même caractéristique (268)

Theorie der Gattungen / Théorie des genres (278)

1. Definitionen ; die Permutationen und die Ordnung einer Gattung / Définitions ; les permutation et l'ordre d'un genre (278)
2. Die Gattung der zweiwertigen Funktionen / Le genre des fonctions bivalentes (286)
3. Gattungen, welche andere Gattungen unter sich enthalten (Gattungsbereich) / Les genres qui contiennent d'autres genres (domaine de genre) (291)
4. Fundamentalsystem und Fundamentalgleichung einer Gattung / Système fondamental et équation fondamental d'un genre (296)
5. Eigentliche und uneigentliche Gattungen ; vollständige Zerlegung einer ganzen rationalen Funktion in lineare und quadratische Faktoren / Genres propres et impropres ; Décomposition complète d'une fonction rationnelle entière en facteurs linéaires et quadratiques (313)
6. Die Fundamental Kongruenz der allgemeinen Arithmetik / La congruence fondamentale de l'arithmétique générale (325)
7. Die Gattungen der Funktionen von drei Elementen / Le genre des fonctions à trois éléments (346)

8. Die Gattungen der Funktionen von vier Elementen. Die Auflösungsmethode und die Klassen der Gleichungen vierten Grads / Le genre des fonctions à quatre éléments. La méthode de résolution et la classification des équations du quatrième degré (348)
9. Eigentliche Gattungen der Funktionen von 5 und 7 Elementen. Allgemeine Bemerkungen über die Auffindung von Gattungen / Genres propres des fonctions de 5 et 7 éléments. Remarques générales sur la découverte des genres (359)

Über die algebraisch auflösbaren Gleichungen / Sur les équations algébriques résolubles (366)

1. Ein substitutionen-theoretischer Lehrsatz / Un théorème de substitution théorique (366)
2. Das *Abel*sche Theorem / Le théorème d'*Abel* (373)
3. *Abel*sche Gleichungen / Les équations *abeliennes* (383)
4. Aufstellung aller *Abel*schen Gleichungen eines Rationalitätsbereiches / Formation de toutes les équations *abeliennes* d'un domaine de rationalité (410)
5. Der allgemeinste Ausdruck, welcher einer *Abel*schen Gleichung genügen kann / L'expression la plus générale qui peut satisfaire une équation *abelienne* (415)

Nachtrag zur Fundamental Kongruenz der allgemein Arithmetik / Supplément sur la congruence fondamentale de l'arithmétique générale (418)

## B.2 Transcription du chapitre du cours de 1890-1891 sur la théorie des caractéristiques

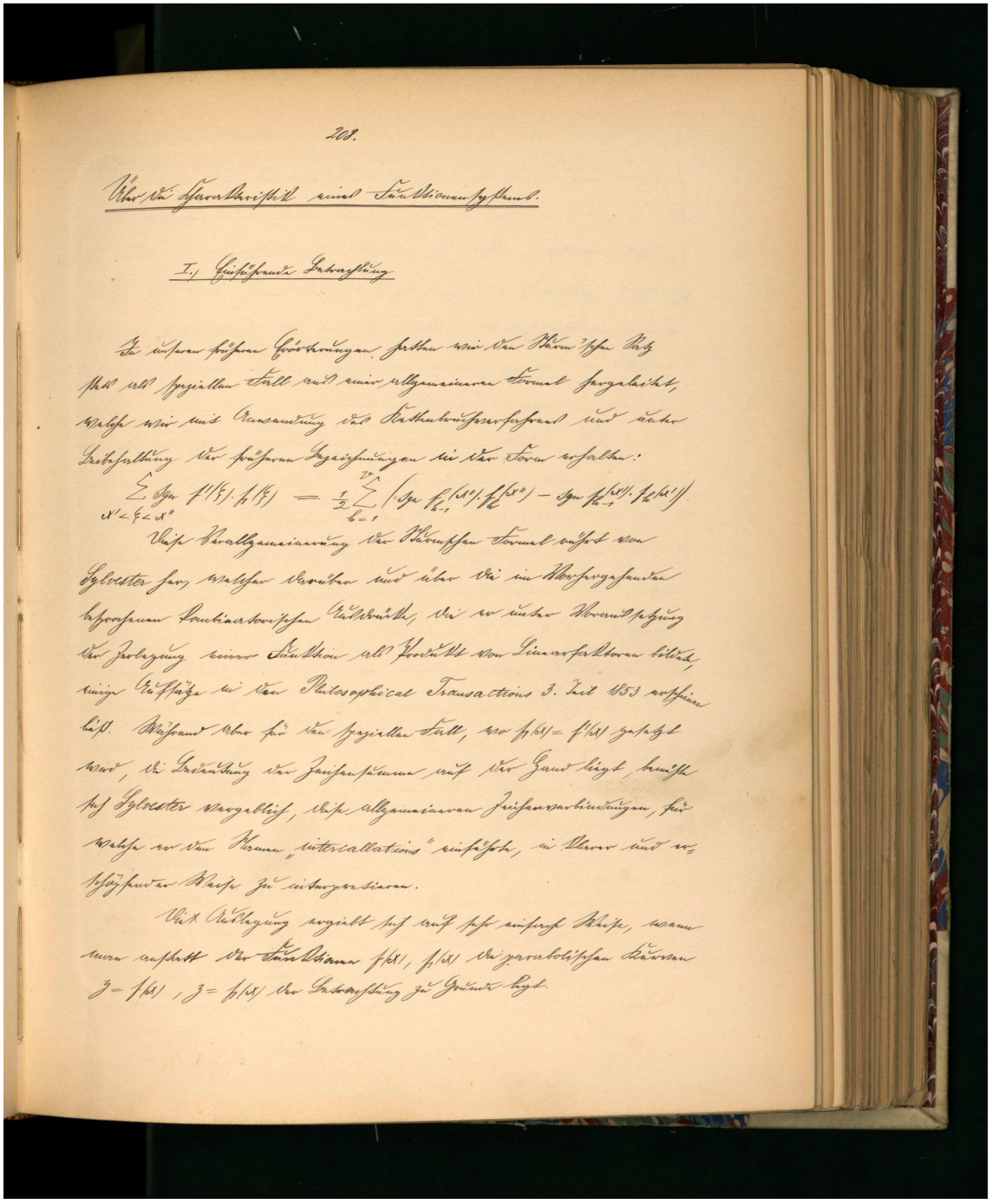


FIGURE 3 – Page 208 du manuscrit de 1890-91 ([Kronecker, 1881a])

# Über die Charakteristik eines Funktionensystems

## I.) Einführende Betrachtung

| In unseren früheren Erörterungen hatten wir den Sturm'schen Satz stets als 207|208
speziellen Fall aus einer allgemeinen Formel hergeleitet, welche wir mit Anwendung des Kettenbruchverfahrens und unter Beibehaltung der früheren Bezeichnungen in der Form erhalten :

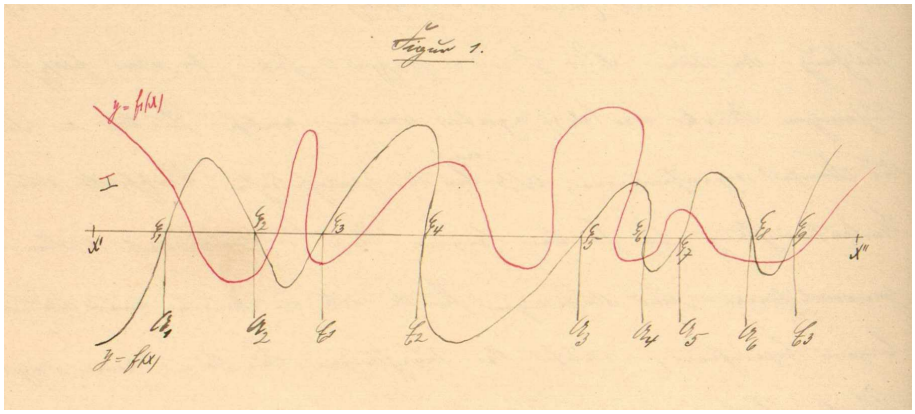
$$\sum_{x'<\xi<x''} sgn\ f'(\xi)f_1(\xi) = \frac{1}{2}\sum_{h=1}^{v}(sgn\ f_{h-1}(x'').f_h(x'') - sgn\ f_{h-1}(x').f_h(x'))$$

Diese Verallgemeinerung der Sturmschen Formel rührt von *Sylvester* her, welchen darüber und über die in vorhergehenden besprechenen kombinatorischen Ausdrücke, die er unter Voraussetzung der Zerlegung einer Funktion als Produkt von Linearfaktoren bildete, einige Aufsätze in den *Philosophical Transactions* z. Teil 1853 erscheinen ließ. Während aber für den Speziellen Fall, wo $f_1(x) = f'(x)$ gesetzt wird, die Bedeutung der Zeichensumme auf der Hand liegt, bemühte sich *Sylvester* vergeblich, diese allgemeineren Zeichenverbindungen, für welche er den Namen „intercallations“ einführte, in Klarer und erschöpfend Weise zu interpretieren.

Die Auslegung ergiebt sich auf sehr einfache Weise, wenn man anstatt der Funktionen $f(x), f_1(x)$ die parabolischen Kurven $y = f(x)$, $y = f_1(x)$ der Betrachtung zu Grunde legt.

| Jede Kurve teilt sämtliche Punkte der Ebene in drei Gruppen, solche, welche 208|209
auf der kurve selbst liegen, solche, die bei einer Bewegung auf der Kurve in einer bestimmten Richtung zur rechten Hand und solche, welche zur linken Hand liegen. Spricht man diese Einteilung arithmetisch aus, so heißt das : Alle endlichen, reellen Wertsystem $(x, y)$ können wir in Beziehung auf eine beliebige ganze Funktion $f(x, y)$ in drei Klassen sondern, je nachdem sie die Funktion $f(x, y)$ zum Verschwinden bringen oder ihr einen positiven oder negativen Wert erteilen. Dabei muß natürlich die Funktion $f(x, y)$ in einer unabänderlichen Form gegeben sein, ebenso wie wir bei unserer Bewegung auf der Kurve die einmal angenommene Richtung stets beibehalten müssen. Betrachten wir z.B. sämtliche Punkte einer Ebene in Beziehung auf einen Kreis, dessen Gleichung die Form $x^2 + y^2 - 1 = f(x, y) = 0$ habe, so wird $f(x, y)$ für diejenigen Werte $(x, y)$ negativ werden, welche Punkte in Innere des Kreises repräsentieren, also nach der gewöhnlichen Vorschrift der Umlaufsrichtung zur Linken liegen. Diese Ausdrucksweise wollen wir generalisieren und diejennigen Punkte als in Innern einer Kurve liegend bezeichnen, welche der zugehörigen Funktion einen negativen Wert erteilen.

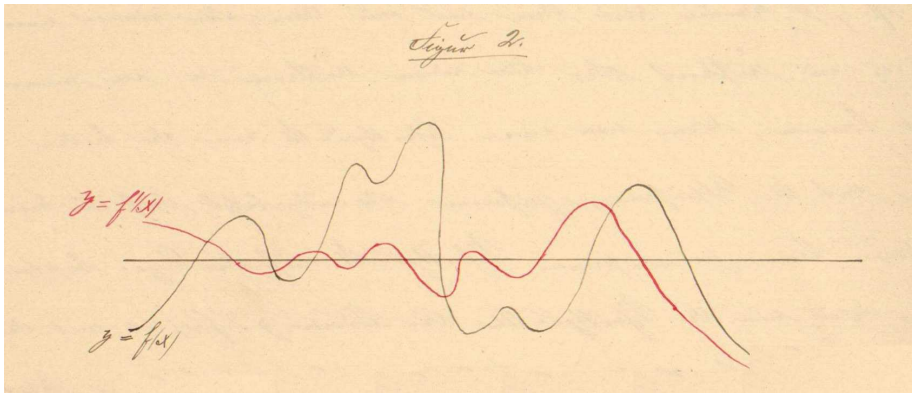
Betrachten wir unter diesem Gesichtspunkte unsere Kurven $y - f(x) = 0$ und $y - f_1(x) = 0$ oder vielmehr die Gesamtheit der beiden Kurven, d.h. das Produkt $(y - f(x))(y - f_1(x))$, so liegen diejenigen | Punkte $(x, y)$ in Innere der Kurve $(y -$ 209|210
$f(x)).(y - f_1(x)) = \varphi(x, y) = 0$, für welche das Produkt einen negativen Wert erhält.



Bewegen wir uns nun auf der Abscissenaxe (d.h. auf einer dritten kurve $y = 0$) und betrachten also nur die Wertsysteme $(x, y)$, für welche $y$ konstant gleich Null ist, so werden wir dann und nur dann von einen innere Gebiet in ein äußeres oder von einen äußeren in ein inneres Gebiet kommen, wenn wir einen Schnittpunkt einer der beiden Kurven mit der Abscissenaxe passieren, vorausgesetzt, daß die beiden Funktionen keinen gemeinsamen Schnittpunkt auf der Axe besitzen. Fassen wir nun die Schnittpunkte der Kurve $y - f(x) = 0$ mit der Axe ins Auge, so wird $sgn\ d(y - f(x))(y - f_1(x))$ für einen Wert $x = \xi$, für welchen $f(\xi) = 0$ ist, gleich $+1$ oder $-1$ sein, je nachdem wir in dem diesem Werte entsprechenden Punkte der Axe vom Innere ins Aüßere kommen (Austrittspunkt), oder vom Aüßere ins Innere eintreten (Eintrittspunkt). Nun ist aber

$$[sgn\ d(y - f(x))(y - f_1(x))]_{x=\xi} = sgn\ f'(\xi)f_1(\xi),$$

woraus also ersichtlich, daß die Zeichensumme $\sum_{x'<\xi<x''} sgn\ f'(\xi).f_1(\xi)$ kurz erklärt werden kann als die Anzahl der Austrittspunkte, welche sich in dem Intervall $x'...x''$ befinden, vermindert um die Anzahl der in demselben Intervall liegenden Eintrittspunkte.



Die umstehende Figur 1 führt ein Beispiel zu dem Vorhergehenden vor Augen. Die
210|211 Funktion $y = f(x)$ ist hier vom ungraden Grade ; wir | erhalten also für sehr große negative Werte der unabhängigen Variablen $x$ ebensolche Funktionswerte $y$, während $f_1(x)$ vom graden Grade ist. Das erste, mit I bezeichnete Gebiet ist in Beziehung auf

das Produkt $(y-f(x))(y-f_1(x))$ als ein inneres aufzufassen. Dieses Gebiet verlassen wir bei einer Bewegung auf der Axe in dem Punkte $\xi_1$ und gelangen alsdann in ein äußeres Gebiet ; $\xi_1$ ist also Austrittspunkt und als solcher mit $A_1$ bezeichnet. Für den aufgezeichneten speziellen Fall ist

$$\sum sgn\ f'(\xi).f_1(\xi) = 6-3 = 3.$$

In dem besonderen Falle, wo $f'(x) = f_1(x)$, erhalten wir nur Austrittspunkte, da zwischen zwei Nullstellen der Funktion stets eine ungrade Anzahl von Nullstellen der Ableitung liegt. (Figur 2).

Diese Betrachtungen können verallgemeinert werden, indem man anstatt der $x$ =Axe eine beliebige dritte Kurve zu den beiden andere hinzunimmt ; jedoch ist es, um diese Untersuchungen arithmetisch durchführen zu können, nötig, einen neuen Begriff einzuführen : den Begriff der Charakteristik eines Funktionsystems.

## II.) Die Charakteristik eines Systems von zwei Funktionen einer Variablen

15 Jan.

Es seien $V(z)$ und $W(x)$ zwei ganze Funktionen von beliebigen Grade [2] und die Determinante

$$\begin{vmatrix} V(z) & W(z) \\ V'(z) & W'(z) \end{vmatrix} = \Delta(z)$$

gesetzt, sodaß $\Delta(z)$ eine Funktion in $z$ vom Grade $(2n-1)$ ist.

| Bezeichnen dann $\xi$ die Werte, für welche die Funktion $V(z)$ verschwindet, so nennen wir die mit $-\frac{1}{2}\sum\ sign\ \Delta(\xi)$ bezeichnete ganze oder halbe Zahl die Charakteristik des Funktionensystems $(V(z), W(z))$ oder in arithmetischer Sprache : 211|212

1. $\chi(V(z), W(z)) = -\frac{1}{2}\sum_{(\xi)}\ sign\ \Delta(\xi)$

wo die Summe erstreckt ist über sämtliche $\xi$, für welche die Funktion $V(z)$ verschwindet.

Es kommt zunächst darauf an, zu zeigen, daß wenn $V(z)$ und $W(z)$ von gleichen Grade sind, an Stelle der Wurzeln von $V(z) = 0$ diejenigen von $W(z) = 0$ in die Definitionsgleichung 1) eingesetzt werden können. Zu diesem Zwecke benützen wir die früher (S. 186) abgeleitete Formel :

$$\sum_{z_1<\xi<z_2} sign\ \Phi'(\xi) = -\frac{1}{2}\ sgn\ \Phi(z_1) + \frac{1}{2}\ sgn\ \Phi(z_2)$$

Hierin ist $\Phi(z) = \Phi_0 z^n - \Phi_1 z^{n-1} + ...$ eine ganze Funktion $n$-ten Grades. Die Summe wird erstreckt über alle Werte $\xi$, welche zwischen $z_1$ und $z_2$ liegen und für welche, kurz ausgedrückt $\Phi(\xi)$ verschwindet. Insbesondere ist, wenn $z_1$ kleiner als die

2. *N.d.T.* Correction sur la page de gauche : $n$

kleinste reelle Wurzel, $z_2$ größer als die größte ist [3] :

$$\sum_{z_1<\xi<z_2} sgn\ \Phi'(\xi) = -\frac{1}{2}\ sgn\ z_1^n\Phi_0 + \frac{1}{2}\ sgn\ z_2^n\Phi_0$$
$$= 0\ ,\ \text{wenn } n \text{ grade ist}$$
$$= 1\ ,\ \text{wenn } n \text{ ungrade ist}$$

Da $V(z)$ und $W(z)$ als von gleichen Grade vorausgesetzt werden, so ist ihr Produkt
212|213 $V(z).W(z)$ von gradem Grade, also die Zeichen- | summe der Ableitung gleich Null, d.h.

2. $\sum_{(\xi)}\ sgn\ V'(\xi).W(\xi) + \sum_{(\eta)}\ sgn\ V(\eta).W'(\eta) = 0$, wenn $\eta$ diejenigen Werte sind, für welche $W(\eta) = 0$ ist, also :

$$\frac{1}{2}\sum_{(\xi)}\ sgn\ V'(\xi).W(\xi) = -\frac{1}{2}\sum_{(\eta)}\ sgn\ V(\eta).W'(\eta)$$

Mithin ist, da

$$\Delta(\xi) = -V'(\xi).W(\xi) \qquad \Delta(\eta) = V(\eta).W'(\eta) :$$

3. $\chi(V,W) = -\frac{1}{2}\sum_{(\xi)}\ sgn\ \Delta(\xi) = \frac{1}{2}\sum_{(\xi)}\ sgn\ V'(\xi).W(\xi)$
$= -\frac{1}{2}\sum_{(\eta)}\ sgn\ V(\eta).W'(\eta) = -\frac{1}{2}\sum_{(\eta)}\ sgn\ \Delta(\eta).$

Aus der Determinantenform von $\Delta(z)$ geht ferner hervor, daß

4. $\chi(V,W) = -\chi(W,V)$ und daß

5. $\chi(V, aV+W) = -\frac{1}{2}\sum_{(\xi)}\ sgn \begin{vmatrix} V & aV+W \\ V' & aV'+W' \end{vmatrix} = \chi(V,W)$

$$= -\frac{1}{2}\sum_{(\eta)}\ sgn \begin{vmatrix} V+aW & W \\ V'+aW' & W' \end{vmatrix} = \chi(V+aW, W)\ \text{ist.}$$

Wir können die vielbesprochene Zeichensumme $\sum_{x'<\xi<x''}\ sgn\ f'(\xi).f_1(\xi)$ durch die Differenz zweier Charakteristiken ausdrücken. Es ist nämlich :

$$\chi(f(z),(x''-z).f_1(z)) = \frac{1}{2}\sum_{(\xi)}\ sgn\ (x''-\xi).f_1(\xi).f'(\xi) = -\frac{1}{2}\sum_{(\xi)}\ sgn\ \Delta(\xi)$$

$$\chi(f(z),(x'-z).f_1(z)) = \frac{1}{2}\sum_{(\xi)}\ sgn\ (x'-\xi).f_1(\xi).f'(\xi) = -\frac{1}{2}\sum_{(\xi)}\ sgn\ \overline{\Delta}(\xi),$$

die Summe erstreckt über sämtliche Wurzeln $\xi$ der Gleichung $f(z) = 0$. Da nun $sgn\ (x''-\xi) = \pm sgn\ (x'-\xi)$, je nachdem $\xi$ außerhalb oder innerhalb des von $x'$ und $x''$ begrenzten Intervalls liegt, so ist :

$$\chi(f(z),(x''-z).f_1(z)) - \chi(f(z),(x'-z).f_1(z)) = \sum_{x'<\xi<x''}\ sgn\ f'(\xi).f_1(\xi)$$
$$= -\frac{1}{2}\sum_{(\xi)}\ sgn\ \Delta(\xi)$$

3. Correction dans la page de gauche :

$$\sum_{z_1<\xi<z_2} sgn\ \Phi'(\xi) = -\frac{1}{2}\ sgn\ z_1^n\Phi_0 + \frac{1}{2}\ sgn\ z_2^n\Phi_0$$
$$= 0\ ,\ \text{si } n \text{ est pair}$$
$$= sgn\Phi_0\ ,\ \text{si } n \text{ est impair}$$

$$+\frac{1}{2}\sum_{(\xi)} sgn\ \overline{\Delta}(\xi)$$

Die Zeichensumme ist also der Zuwachs, welchen die Charakteristik | des Funktionensystems $(f(z),(x-z)f_1(z))$ beim Übergang von $x'$ nach $x''$ erhält. 213|214

Dieses letztere Funktionensystem, in welchem die Koeffizienten der zweiten Funktion von einer veränderlichen Größe $x$ abhängen, kann aufgefaßt werden als ein besonderer Fall des Funktionensystems $(V(z),W(z))$, wo $V(z)=v_0+v_1z+...+v_nz^n$, $W(z)=w_0+w_1z+...w_nz^n$ und die $2n$ Koeffizienten $v_0,v_1...v_n;w_0,w_1...w_n$ Variable sind. Wir fragen, wann ändert sich bei Variierung der Koeffizienten die Charakteristik $\chi(V,W)$, und welches ist der Wert der Änderung, wenn wir von einem bestimmten Koeffizienten-System $(v_0'..v_n';w_0'...w_n')$ zu einem andern $(v_0''..v_n'';w_0''...w_n'')$ übergehen ?

Die Charakteristik des Funktionensystems kann sich offenbar nur dann ändern, wenn die Determinante $\Delta(z)$ für irgend eine der Wurzele $\xi$ von $V(z)=0$ oder der Wurzele $\eta$ von $W(z)=0$ ihr Zeichen ändert, also durch Null hindurchgeht. Notwendige Bedingung dafür, daß sich die Charakteristik ändert, ist also, daß entweder die beiden Gleichungen

$$\text{I.)}\ V(z).W('(z)-V'(z).W(z)=0 \text{ und } V(z)=0 \text{ oder}$$
$$\text{II.)}\ V(z').W('(z')-V'(z').W(z')=0 \text{ und } W(z')=0$$

während der stetigen Veränderung der Koeffizienten einmal eine gemeinschaftliche Wurzel haben.

Es ist dies für I.) dann der Fall, wenn entweder gleichzeitig 1.) $V(z)=0$, $V'(z)=0$ oder 2.) $V(z)=0$, $W(z)=0$ ; für II.), wenn entweder gleichzeitig 1a) $W(z')=0$, $W'(z')=0$ oder 2.) $W'(z')=0$, $V(z')=0$ ist.

Es ist jedoch 1.) oder 1a) (das Verschwinden der Diskriminante einer | der beiden Funktionen) keine ausreichende Bedingung für eine Änderung der Charakteristik, nur das Bestehen der Gleichung 2.) (das Verschwinden der Resultante) hat eine solche zur notwendigen Folge. Denn $\chi(V(z'),W(z'))$ verschwindet dann und nur dann, wenn $\chi(V(z'),W(z'))-x.V(z'))$ ebenfalls gleich Null wird (Gl. 5.)) also wenn z.B. nach Gl. II : 6) $W(z')-xV(z')=0$ und 7.) $V(z')(W'(z')-x.V'(z')=0$ also 8.) $V(z').W(z')-W(z').V'(z')=0$. Diese letzte Gleichung 8.) bestimmt uns Werte von $z'$, welche auf die Gleichung 6.) befriedigen sollen und zwar für jeden Wert der variablen Größe $x$ ; es muß also $V(z')=0$ und $W(z')=0$ sein. Das ist die einzige aus II.) folgende Bedingung, während aus I.) sich noch die Bedingung $V(z)=0$, $V'(z)=0$ ergeben würde. Daß aber auch diese Bedingung für eine Änderung der Charakteristik nicht ausreichend ist, folgt wieder daraus, daß dieselbe nur gleichzeitig mit einer Änderung von $\chi(V(z)-xW(z),W(z))$ stattfinden kann, für welche letztere jedoch als notwendige Bedingungen nur $V(z)=0$, $W(z)=0$ oder $W(z)=0$, $W'(z)=0$ bestehen. Damit ist also beweisen : 214|215

Die Charakteristik ändert ihren Wert bei einer stetigen Variirung des Koeffizientensystems ihrer Funktionen dann und nur dann, wenn die Resultante der beiden Funktionen verschwindet.

Da $\chi(V(z),W(z))=-\frac{1}{2}\sum\ sgn\ \Delta(\xi)$, so kann sich die Charakteristik bei stetigem Fortgang nur eine Einheit vermehren oder vermindern, je nachdem das betreffende $\Delta(\xi)$, welches sich ändert, von $+1$ zu $-1$ oder von $-1$ zu $+1$ übergeht. Wenn wir von einem Koeffizientensystem $(v_k^{(1)},w_k^{(1)})$ $(k=0,1,2,...n)$ übergehen zu einem andern $(v_k^{(2)},w_k^{(2)})$, so ist die Änderung der Charakteristik :

$$-\frac{1}{2}\ sgn\ \Delta(v_k^{(1)},w_k^{(1)})+\frac{1}{2}\ sgn\ \Delta(v_k^{(2)},w_k^{(2)}),$$

also die Gesamt- | änderung 215|216

9. $\chi(v_0'...;w_0'...)-\chi(v_0''...;w_0''...)$
$=\sum_{(h)}\left(-\frac{1}{2}\ sgn\ \Delta(v_k^{(h)},w_k^{(h)})+\frac{1}{2}\ sgn\ \Delta(v_k^{(h+1)},w_k^{(h+1)})\right)$

indem sich die Summation in der ersten Summe bezieht auf alle Werte $h$, welche so beschaffen sind, daß zwischen dem System $(v_0^{(h)}...v_n^{(h)}; w_0^{(h)}...w_n^{(h)})$ und $(v_0^{(h+1)}...v_n^{(h+1)}; w_0^{(h+1)}...w_n^{(h+1)})$ ein System von Koeffizienten $(v_0^{(h_1)}...v_n^{(h_1)}; w_0^{(h_1)}...w_n^{(h_1)})$ sich befindet, für welches $R = 0$ wird und entsprechend die Summation der zweiten Summe über alle Werte $\xi_h$, für welche $V(\xi_h) = 0$ und $W(\xi_h) = 0$ ist.

---

Um nun die Änderung der Charakteristik wirklich zu berechnen, können wir genau dieselben Deduktionen anwenden, deren wir uns bei der Herleitung des verallgemeinerten Sturm'schen Satzes bereits bedient haben ; wir wollen das Verfahren mit entsprechender Veränderung der Bezeichnungen kurz rekapitulieren.

Es sei

$$V(z) = v_0 + v_1 z + v_2 z^2 + ... + v_n z^n = v_n \prod(z - \xi)$$
$$W(z) = w_0 + w_1 z + w_2 z^2 + ... + w_n z^n = w_n \prod(z - \eta).$$

Alsdann ist $R(V(z), W(z))$ nicht mehr eine ganze Funktion der Koeffizienten, sondern nur ganz in den Quotienten $\frac{v_0}{v_n}, ... \frac{v_{n-1}}{v_n}$. Wir definieren ferner :

$R_{n-m} = (-1)^{\frac{n(n-1)}{2}} T_{m,m}$ $(m = 0, 1...n)$ und erhalten damit eine Reihe : $R_0, R_1...R_n$, in welcher $R_0 = R$, $R_n = 1$ zu setzen ist.

Da nun :

$$(-1)^{\frac{n(n-1)}{2}} T_{nn} = R(V, W) \text{ (S. 202) und}$$
$$(-1)^{\frac{n(n-1)}{2}} T_{n-1,n-1} = R(V, W) \sum_{i=1}^{n} \frac{1}{W(\xi_i).V'(\xi_i)} \text{ (S. 205),}$$

so erhalten wir :

216|217 10. $\frac{R_1}{R} = \frac{T_{n-1,n-1}}{T_{n,n}} = \sum_{i=1}^{n} \frac{1}{W(\xi_i).V'(\xi_i)}$ oder : |

$$R_1 = \frac{R}{W(\xi_i).V'(\xi_1)} + \frac{R}{W(\xi_2).V'(\xi_2)} + ...$$

Es mögen nun $V(z)$ und $W(z)$ durch stetige Variirung der Koeffizienten eine gemeinsame Wurzel erhalten, z.B. $\xi_1 = \eta_1$ wurden, so werden, wie man aus der letzten Entwicklung sieht, in der Umgebung von $\xi_1$ alle Glieder auf der rechten Seite beliebig klein mit Ausnahme des ersten, in welchem Zähler und Nenner beliebig klein werden. Daraus folgt, daß

$$sgn\ R_1 = \ sgn\ \frac{\delta R}{\delta(W(\xi_1).V'(\xi_1))}$$

oder

$$sgn\ R_1.\delta R = \ sgn\ \delta(W(\xi_1).V'(\xi_1)).$$

Nun folgt aus 9) :

11. $\chi(v_0'...; w_0'...) - \chi(v_0''...; w_0''...) = -\frac{1}{2} \sum_{(h)}\ sgn\ \delta[W(x).V'(x)]_{x=\xi_h}$
$= -\sum_{(h_1)}\ sgn\ R_1(v_k^{(h_1)}.w_k^{(h_1)}).\delta R(v_k^{(h_1)}.w_k^{(h_1)})$ oder :

12. $\chi(v_0'...; w_0'...) - \chi(v_0''...; w_0''...) = \sum_{(h_1)}\ sgn\ R_1(v_k^{(h_1)}.w_k^{(h_1)}).\delta R(v_k^{(h_1)}.w_k^{(h_1)})$

indem $\chi(v_0''...; w_0''...)$ die Charakteristik des Funktionensystems

$$V^{(2)}(x) = v_0'' + v_1'' x + ... + v_n'' x^n \ , \ W^{(2)}(x) = w_0'' + w_1'' x + ... + w_n'' x^n;$$

$\chi(v_0'...; w_0'...)$ diejenige des Funktionsystems

$$V^{(1)}(x) = v'_0 + v'_1 x + ... + v'_n x^n \ , \ W^{(1)}(x) = w'_0 + w'_1 x + ... + w'_n x^n$$

bezeichnet. Nun ist (S. 180) : $T_{m+1,m+1}.T_{m-1,m-1} \equiv -(\ )^2 (mod\ T_{m,m})$ also

13. $R_{h-1}.R_{h+1} \equiv -(\ )^2 (mod\ R_h)$.

Dasselbe galt für die Reihe

14. $T_0(x), T_1(x), ...T_n(x)$, d.h. es war auch

$$T_{m+1}(x).T_{m-1}(x) \equiv -(\ )^2 (mod\ T_m(x)).$$

Daraus wurde gefolgert, daß die Anzahl der Zeichenwechsel der | Reihe 14) sich nicht 217|218
ändert, wenn bei stetiger Veränderung des $x$ eine der Funktionen außer $T_n(x)$ durch Null hindurchgeht und daß die Anzahl der Zeichenwechsel, welche verloren gehen, wenn man $x$ den Weg von $x'$ nach $x''$ machen läßt, sich durch die Summe (S. 187)

$$\begin{aligned} w' - w'' &= \sum_{x'<\xi<x''} sgn\ T'_n(\xi).T_{n-1}(\xi) \\ &= -\frac{1}{2}\sum_{h=1}^{n} sgn\ T_h(x').T_{h-1}(x') + \frac{1}{2}\sum_{h=1}^{n} sgn\ T_h(x'').T_{h-1}(x'') \end{aligned}$$

ausdrücken laßt, wo sich die erste Summe über alle Werte von $\xi$ erstreckt, für welche $T_n(\xi) = 0$ ist und welche beim stetigen Übergange von $x'$ nach $x''$ passiert werden.

Diese Schlußfolgerungen können wir unmittelbar auf unsere Reihe

$$R_0, R_1, ...R_n$$

übertragen, da als einziger Unterschied der Umstand zu beachten ist, daß $R_0, R_1 ... R_n$ nicht von einer veränderlichen Größe $x$ abhängen, sondern in Bezug auf die $2(n+1)$ Koeffizienten von $V(x)$ und $W(x)$ variabel sind. Man hat deshalb auch bei der Deduktion anstatt der Ableitung $T'_n(x)$ das vollständige Differential $\delta R(v_k, w_k)$ zu setzen. Nach dieser Überlegung sind wir berechtigt, die Gleichung 12) in folgender Weise zu vervollständigen :

15. $\chi(v''_0, w''_0) - \chi(v'_0, w'_0) = \sum_{(h_1)} sgn\ R_1(v_k^{(h_1)}, w_k^{(h_1)}).\delta R(v_k^{(h_1)}, w_k^{(h_1)}) = \frac{1}{2}\sum_{h=0}^{n-1} sgn\ R_h(v''_0, w''_0)R_{h+1}(v''_0; w''_0) - \frac{1}{2}\sum_{h=0}^{n-1} sgn\ R_h(v'_0, w'_0)R_{h+1}(v'_0; w'_0)$.

Kennen wir für ein bestimmtes System von Koeffizienten der Wert der Charakteristik, so kennen wir denselben für alle Koeffizientensysteme.

Denken wir uns jetzt die $2(n+1)$ Koeffizienten $v_0 ... v_n; w_0 .... w_n$ abhängig von insgesamt $\nu$ Variabeln $x_1, x_2, ...x_\nu$, welche von einander | unabhängig sind und daher eine 218|219
$\nu$-fache Mannigfaltigkeit bilden. Dann gehört zu jedem Punkte derselben ein Funktionsystem $(V(z), W(z))$ und dessen Charakteristik und umgekehrt. Die Gleichung $R = 0$ fixiert eine $(\nu-1)$-fache Mannigfaltigkeit, welche aus der $\nu$-fachen ausgeschieden ist. Bewegen wir uns mit unserer Charakteristik von einem Punkte $(x'_1, x'_2, ...x'_\nu)$ zu einem andern $(x''_1, x''_2, ...x''_\nu)$, so wird sich stets dann und nur dann der wert derselben ändern, wenn wir die $(\nu-1)$-fache Mannigfaltigkeit $R = 0$ passieren. Betrachten wir nun die $(\nu-1)$-fache Mannigfaltigkeit $R.R_1 = 0$. Dieselbe teilt unsern $\nu$-fache Mannigfaltigkeit in positive und negative Gebiete, und wir können mit Benutzung der Begriffe der Austrittspunkte und Eintrittspunkte für die Mannigfaltigkeit $R.R_1 = 0$ die Formel 15.), welche sich auch in der Form

$$\chi(v''_0, w''_0) - \chi(v'_0, w'_0) = \sum_{(h_1)} sgn\ [\delta R(v_k^{(h_1)}, w_k^{(h_1)}).\delta R_1(v_k^{(h_1)}, w_k^{(h_1)})]$$

schreiben läßt, folgendermaßen drücken :

Die Charakteristik nimmt auf dem Wege von einem Punkte der $\nu$-fachen Mannigfaltigkeit zu einem andern nur so viel Einheiten zu, als die Anzahl der Austrittspunkte

die Anzahl der Eintrittspunkte übertrifft.

Demnach lässt sich nun auch der Inhalt des von *Sylvester* verallgemeinerten Sturm'schen Satze (S. 213) in folgender Form aussprechen.

Der von *Sylvester* verallgemeinerte Sturmsche Satz giebt die Veränderung der Charakteristik eines Systems von zwei Funktionen an, wenn man die eine Funktion als konstant festhält und die Koeffizienten der andern als lineare Funktionen einer
219|220 einzigen Variabeln $x$ auffaßt. | Fassen wir nunmehr den speziellen Fall der Charakteristik des Funktionensystems $(V(z), V(z) + V'(z))$ ins Auge. Für dieses ist

$$\Delta(z) = \begin{vmatrix} V & V + V' \\ V' & V' + V'' \end{vmatrix} = V.V'' - V'^2.$$

Die Resultante von $V(z)$ und $V'(z)$ ist die Diskriminante von $V(z)$ und ihre Charakteristik ist $-\frac{1}{2}\sum_{(\xi)} \; sgn\; \Delta(\xi) = +\frac{1}{2}\sum_{(\xi)} \; sgn\; (V'(\xi))^2$, also gleich der halben Anzahl der reellen Wurzeln von $V(x) = 0$.

Aus 11) folgt außerdem die Änderung der Charakteristik :

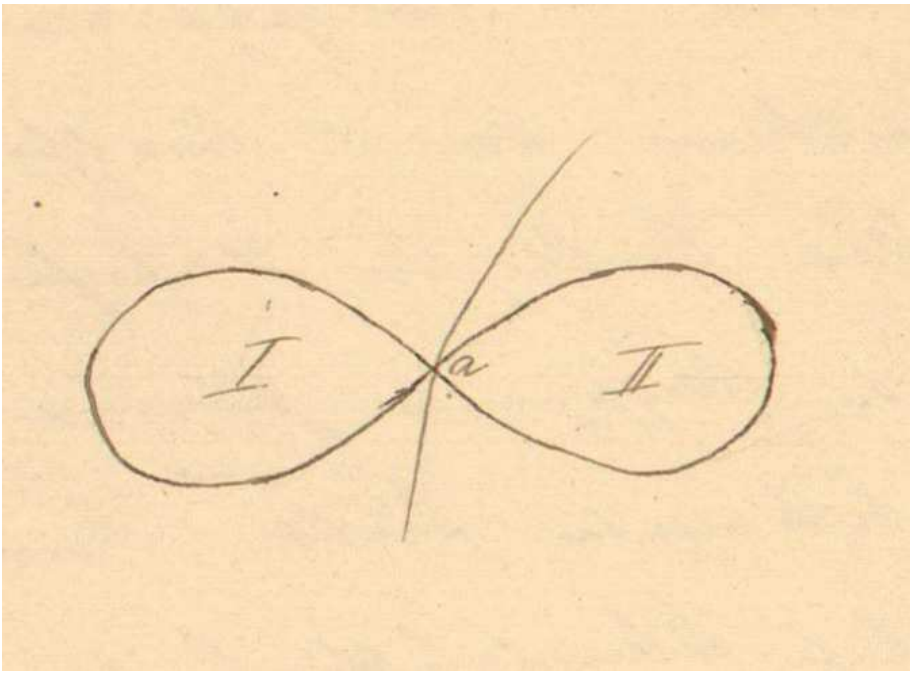
$$\chi(v_0''..., w_0''...) - \chi(v_0'..., w_0'...) = \frac{1}{2}\sum_{(h)} \; sgn\; [V''(x).\delta V(x)]_{x=\xi_h},$$

wenn mit $\xi_h$ diejenigen Werte bezeichnet werden, für welche zugleich $V(x)$ und $V'(x)$ verschwinden.

Denken wir uns jetzt weiter die Koeffizienten der Gleichung $V(x) = 0$ als Funktionen von $\nu$ unabhängigen Variabeln $x_1, x_2...x_\nu$, so können wir jedem Punkt der $\nu-$fachen Mannigfaltigkeit eine Gleichung mit bestimmten Koeffizienten zuordnen, und umgekehrt fixiert jede Gleichung eine bestimmten Punkt. Dann scheidet nach dem Vorhergehenden die $(\nu - 1)-$fache Mannigfaltigkeit, deren Gleichung $D = 0$ ist, diejenigen Gebiete von einander, in welchen die Gleichungen mit verschiedener Anzahl der reellen Wurzeln liegen und zwar so, daß man, um von einer Gleichung zu einer andern zu kommen, welche nicht dieselbe Anzahl reeller Wurzeln hat wie die erste, unumgänglich die Diskriminante durchschreiten muß.

Aus diesen Untersuchungen geht hervor, daß eigentlich schon die Diskriminante allein uns denselben Dienst leistet, welchen wir bisher nur von einer Reihe Sturm'schen Funktionen erwarten durften : die Anzahl der reellen Wurzeln einer Gleichung
220|221 zu | bestimmen. Es bedarf nämlich nur einer erschöpfenden Diskussion der Lage

der Diskriminanten-Mannigfaltigkeit und der Teile, in welche sie die höhere Mannigfaltigkeit sondert, um jeder Gleichung ihren Platz in der letzteren anweisen und damit die Anzahl ihrer reellen Wurzeln bestimmen zu können. Wir werden dies speziell an dem Beispiel der Gleichung vierten Grades durchführen und sehen, daß es dabei vor allen darauf ankommt, alle Singularitäten der Mannigfaltigkeit ausfindig zu machen. Die Sturm'schen Funktionen ergeben Mannigfaltigkeiten, welche diese singulären Mannigfaltigkeiten niederen Ordnung, die auf der Diskriminanten-Mannigf. oder in allgemeineren Fall auf der Resultanten-Mannigf. liegen, enthalten und durch ihr Zeichen das Gebiet, in welchem die vorgelegte Gleichung, (bez. das vorgelegte Funktionensystem) sich befindet, genau bestimmen. Wenn z.B., um das Gesagte zu veranschaulichen, die Diskriminante die Form einer Lemniskate hat, so ist es nötig, um die Punkte der beiden Gebiete I und II durch arithmetische Angaben von einander sondern zu können, noch durch den Doppelpunkt $a$ eine Kurve zu legen, und erst, wenn man die Lage eines Punktes der Ebene in Beziehung auf jede der beiden Kurven durch das Zeichen fixiert hat, welches seine Koordinaten den beiden Funktionen erteilen, ist auch arithmetisch genau bestimmt, welchem der drei durch die Lemniskate




ausgesonderten Gebiete zugehört.

Man beachte ferner, daß die Sturmschen Funktionen für ein bestimmte Gleichung durchaus nicht absolut bestimmt sind und daß sie auch in Bezug auf ihre Anzahl lediglich durch eine untere Grenze beschränkt werden.

| Wir hatten bewiesen, daß die Gebiete, welchen die Gleichungen mit derselben 221|222
Anzahl reeller Wurzeln zugehören, durch die Diskriminanten-Mannig. von einander getrennt sind ; im folgenden werden wir zeigen, daß man in stetigem Übergang von jeder Gleichung zu jeder andern mit derselben Anzahl reeller Wurzeln gelangen kann, ohne die Diskriminanten-Mannigf. zu passieren ; m.a.W dass also diese Gebiete auch zusammenhängend sind. Dasselbe gilt auch für den allgemeineren Fall der Charakteristik eines Funktionensystems und der zugehörigen Resultanten-Mannig.

Wir wollen zeigen, daß wir von der Gleichung 1.) $\prod_{k=1}^{n}(w-w'_k)=0$ zu der andern 2.) $\prod_{k=1}^{n}(w-w''_k)=0$ „reisen“ können ohne Überschreitung der Diskriminante, unter der Annahme, daß beide Gleichungen gleich viel reelle Wurzeln haben und sämtliche Wurzele jeder der beiden Gleichungen von einander verschieden sind. Zu diesem Zwecke denken wir uns die Wurzeln $w_k = u_k + v_k i$ nach der Größe ihrer imaginären Teile $v_k$ (jedoch diese nicht dem absoluten Werte nach, sondern mit Berücksichtigung des Vorzeichens vom $i$ geordnet und die Wurzeln mit gleichen imaginären Teilen nach der Größe der reellen Teile $u_k$. Wir bilden alsdann :

$$3.)\quad \prod_{k=1}^{n}(w-(1-t)w'_k - t.w''_k)=0$$

für $0 \leq t \leq 1$, sodaß wir in stetigem Übergange von der ersten Gleichung zu der zweiten gelangen, wenn $t$ alle Werte von 0 bis 1 durchläuft. Es erübrigt, zu zeigen, daß für keinen Wert von $t$ die Discriminante dieser Gleichung :

$$\prod_{h<k}\left((w'_k-w'_h)(1-t)+(w''_k-w''_h)t\right)^2$$

verschwinden kann.

| Wenn dies der Fall wäre, so müßte einer der Faktoren, d.h. der reelle und 222|223
der imaginär Teil desselben einzeln verschwinden. Setzen wir aber $w'_k - w'_h = u'_k + v'_k i - (u'_h + v'_h i) = w'_k - w_h + i(v'_k - v'_h)$, so ist nach unserer Anordnung entweder

$v'_k - v'_h > 0$, oder wenn $v'_k = v'_h$, so ist $u'_k - u'_h > 0$, dasselbe ist der Fall für $w''_k - w''_h$. Da nun $t$ und $(1-t)$ ebenfalls positive Größen sind, so ist der imaginäre Teil des Faktors positiv oder doch, wenn dieser verschwinden sollte, der reelle Teil größer als Null. Beide Teile können nur dann verschwinden, wenn $w'_k = w'_h$ und $w''_k = w''_h$ ; aber diese Gleichungen widersprechen beide unserer Annahmen.

Es ist jedoch noch der Einwand zu entkräften, der gegen diese Argumentation leicht erhoben werden kann : es sei mit derselben der Beweis erbracht, daß man von jeder Gleichung zu jeder beliebigen andern ohne Durchschreitung der Diskriminanten-Mannigf. gelangen könne, da ja die Voraussetzung der gleichen Anzahl reeller Wurzeln, welche für die vorgelegten Gleichungen gemacht sei, im Beweise durchaus nicht aufgebraucht worden wäre. Daraus ist zu erwidern, daß nur unter dieser Annahme die Gleichung 3) reelle Koeffizienten hat. Denn wenn z.B. die Gleichung 1) mehr komplexe Wurzeln besitzt als 2), so wird es in dem Produkt auf der rechten Seite von 3) sicher zwei Faktoren geben, in welchen ein komplexes $w'_k$ und ein reelles $w''_k$ vorkommt und die beiden komplexen Wurzeln $w'_k$ einander konjugiert sind. Das Produkt dieser beiden in $w$ linearen Faktoren kann aber nur dann reelle Koeffizienten haben, wenn die beiden reellen Wurzeln $w''_k$ einander gleich sind, - und dieser Fall hatten wir ausgeschlossen.

223|224

## | III.) Anwendung auf die Gleichung vierten Grades.

Wir wollen nunmehr unsere allgemeinen Erörterungen anwenden auf den speziellen Fall, wo $V(z)$ eine Gleichung vierten Grades ist und dafür die Lage und die Grenzen der Gebiete bestimmen, in welche unsere Mannigf. zerlegt wird. Wir benutzen zu diesem zwecke die früher (S. 176) abgeleitete Formel :

$$1.) \quad \sum_{(\xi_i)} sgn\ \Theta(\xi_i) = \sum_{m=0}^{n-1} sgn\ T_{m,m}.T_{m+1,m+1},$$

wo die erste Summe sich über alle reellen Wurzele der Gleichung $V(z) = 0$ erstreckt und $\Theta$ eine beliebig Funktion bezeichnet. Für $\Theta = 1$ mögen die $T$ übergehen in $\overline{T}$ und die linke Seite bedeutet dann die Anzahl $t$ der reellen Wurzeln der Gleichung $V(y) = 0$. Bezeichnet $w$ die Anzahl der Zeichenwechsel, $\varphi$ diejenige der Zeichenfolgen in der Reihe

$$2.) \quad T_{00}, T_{11}, ...T_{nn},$$

so folgt aus 1.) :

$$A = \sum_{m=0}^{n-1} sgn\ \overline{T}_{m,m}.\overline{T}_{m+1,m+1} = \varphi - w = n - 2w,$$

woraus hervorgeht, daß $w$ die Anzahl der Paare komplexer Wurzele der Gleichung $V(z) = 0$ bezeichnet. Ferner ist :

$$\bar{T}_{m,m} = |s_{p+q}| \ (p, q = 0, 1, 2, ...m-1)$$

wo $s_\alpha$ wie immer die Summe der $\alpha-$ten Potenzen aller Wurzeln bezeichnet. Wir haben also eine Reihe von $(n+1)$ solchen Determinanten, da $m$ die Werte $0, 1, ...n$ haben kann, jedoch kann die Anzahl der Zeichenwechsel nicht größen sein als $\frac{1}{2}n$, bezw $\frac{1}{2}(n-1)$, je nachdem $n$ grade oder ungrade ist. Es ist eine bisher noch nicht beantworte Frage, ob es möglich ist, durch die ersten $\frac{1}{2}n$, bezw. $\frac{1}{2}(n-1)$ Determinanten
224|225 alle übrigen so linear | darzustellen, daß man daran diese Beschränkung in der Anzahl

der Zeichenwechsel klar zu durchschauen vermag. Für die Fälle $n = 3, 4, 5$ ist die Darstellung in dieser Weise möglich und zwar mit Koeffizienten, welche abwechselnd negative und positive Summen von Quadraten sind, derart, daß wenn die ersten $\frac{n}{2}$ bezw. $\frac{n-1}{2}$ Glieder der Reihe nur Zeichenwechsel darbieten, durch die übrigen die Anzahl der Zeichenwechsel nicht mehr vergrößert werden kann.

Für unsere speziellen Fall nehmen wir nun als Gleichung :

$$3.)\ F(w) = 3x + 4y.w \star -6z.v^2 + w^4 = 0,$$

nachdem wir dieselbe von dem Glied befreit haben, welches die dritte Potenz enthält. Legen wir diese Form der Gleichung unserer Betrachtung zu Grunde, so können wir die drei variablen Koeffizienten $x, y, z$ als Koordinaten in Raume auffassen, bezogen auf ein rechtwinkliges Axensystem $X, Y, Z$. Dann erfüllen sämtliche Gleichungen vierten Grades den dreidimensionaten Raum. Für uns handelt es sich darum, die Gebiete von einander zu trennen, in welchen die Gleichungen mit 0, 1 und 2 Paaren imaginärer Wurzeln liegen. Wir wollen dieselben bezw. mit $G_0, G_1$ und $G_2$ bezeichnen. Die Diskriminante der Gleichung, welche gleich Null gesetzt, eine Fläche darstellt, ist

$$\begin{aligned} 4.a)\ D &= |s_{p+q}|_{p,q=0,1,2,3} \\ &= (x + z^2)^3 - (3xz + y^2 - z^3)^2, \text{ (abgesehen von einen Zahlenfaktor)} \\ 4.b)\ &= x(x - 3z^2)^2 - 2y^2z(3x - z^2) - y^4, \end{aligned}$$

da $s_0 = 4$, $s_1 = 0$, $s_2 = 12z$, $s_3 = -12y$, $s_4 = 72z^2 - 12x$. Zusammen mit der Diskriminante $D$ bilden | 225|226

$$D_1 = \frac{1}{4.12^2}\begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix} = -xz - y^2 + 3z^3,$$

$$D_2 = \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix} = z$$

$$\frac{1}{4}s_0 = 1$$

die Funktionen unserer Sturmschen Reihe. Die Anzahl der reellen Wurzeln der Gleichung ist demnach :

$$5.)\ A = \ sgn\ D.D_1 + \ sgn\ D_1.z + \ sgn\ z + \ sgn\ 1.$$

Die Gleichung $F(w) = 0$ mit dem willkürlichen Parameter $w$ stellt für das Koordinatensystem $X, Y, Z$ eine Schaar von Ebenen dar, und da, von einer multiplikativen numerischen Konstanten abgesehen

$$D = R\left(F(w), \frac{\partial F(w)}{\partial w}\right),$$

so ist $D = 0$ die Enveloppe der Ebenenschaar und mithin eine abwickelbare Fläche.

Die Diskriminatenfläche enthält alle Gleichungen mit einem Paar gleicher Wurzeln. Ihre Wendelkurve laßt sich darstellen als Durchschnitt zweier Flächen, deren Gleichungen aus $F(w) = 0$, $F'(w) = 0$, $F''(w) = 0$ durch Elimination von $w$ hervorgehen, und enthält die Gleichungen mit einer dreifachen Wurzel. Daher erscheint sie auch als gemeinsame Kurve der beiden Flächen $D = 0$ und $D_1 = 0$. Am einfachsten

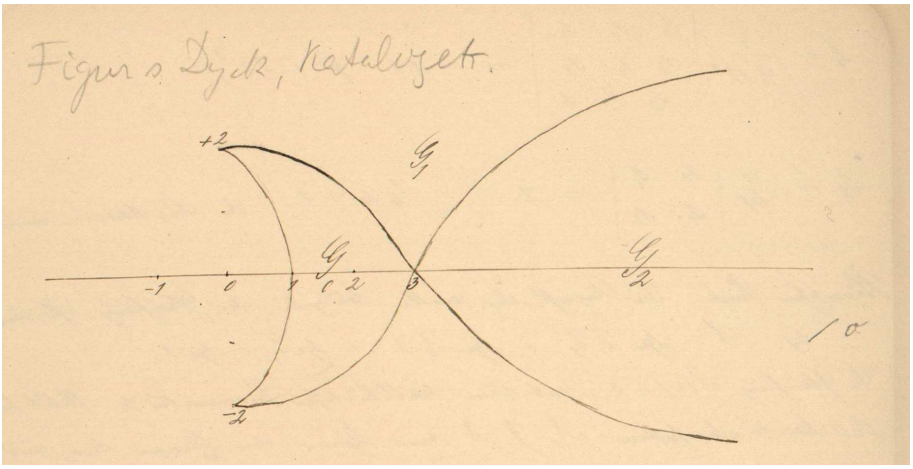
ergeben sich ihre Gleichungen aus der Darstellung von $D$ in 4.$a$) :

$$6.)\ x + z^2 = 0 \qquad y^2 - 4z^3 = 0.$$

Die Diskriminantenfläche besitzt jedoch noch eine zweite singuläre Kurve, eine Doppelkurve, deren Gleichungen sich aus 4.$b$) in der Form :

$$7.)\ x - 3z^2 = 0 \qquad y = 0.$$

226|227 ergeben. Dieselbe ist eine Parabel |




in der $XZ-$Ebene und der Ort aller Gleichungen mit zwei Paaren gleicher Wurzeln.

Um uns von der Gestalt der Fläche $D = 0$ ein Bild zu machen, wollen wir den Durchschnitt zeichnen für die Ebene $z = 1$. Die Gleichung des Schnitt ist :

$$(x+1)^3 - (3x + y^2 - 1)^2 = 0$$

und es wird für :

$$x = -1\ ,\ y = \pm 2\ ;\ x = +3\ ,\ y = 0\ ;\ x = +1\ ,\ y = 0.$$

Das Gebiet $G_0$, in welchem die Punkte liegen, welche die Gleichungen mit nur reellen Wurzeln repräsentieren, ist das Kleinste. Es spitzt sich sackförmig nach der $XY-$Ebene zu und besitzt in dieser nur den Anfangspunkt der Koordinaten. Auf der negativen Seite der $XY-$Ebene kann das Gebiet, da es zusammenhängend sein muß, keine Fortsetzung besitzen. Dagegen zieht sich das Gebiet $G_2$ ebenso wie $G_1$ durch die $XY-$Ebene hinunter, ersteres an der $Z-$Axe entlang, welche letztere die Diskriminanten-Fläche in ihrer ganzen Länge enthält. Jedoch liegt die Fläche $D = 0$ im Bereich der negativen $Z-$Koordinaten ausschließlich auf der rechten (positiven) Seite der $YZ-$Ebene. Wir erhalten nämlich z.B. für $z = -1$ aus 4.$a$) :[4]

$$8.)\ (x+1)^3 - (-3x + y^2 + 1)^2 = 0 \text{ oder } x^3 - 6y^2 + 9x - y^4 + 6xy^2 - 2y^2 = 0$$

eine Gleichung, welche für negative Werte von $x$ nur imaginäre Werte von $y$ liefert. Die Gebiete $G_0$ und $G_2$ haben als gemeinsame Grenze die Parabel 7.), welche ihren Scheitel im Koordinaten-Anfangspunkt besitzt. Variirt man einen Punkt des

4. Correction dans la page de gauche :

$$8.)\ (x+1)^3 - (-3x + y^2 + 1)^2 = 0 \text{ oder } x^3 - 6x^2 + 9x - y^4 + 6xy^2 - 2y^2 = 0$$

et commentaire : nur eine Hälfte

Gebiet $G_0$ in stetigem Fortschritt bis zu einem Punkte von $G_1$, so werden bei der den einzelnen Punkten zugeordneten Gleichung im | Augenblick der Überschreitung der 227|228
Diskriminanten-Fläche zwei Wurzeln einander gleich und sodann komplex, bei weiterem Fortschreiten bis zu einem Punkt von $G_2$ werden auch die beiden andern Wurzeln im Übergangspunkte auf der Fläche einander gleich und alsdann komplex. Geht man dagegen direkt von einem Punkte in $G_0$ durch die Parabel Hindurch zu einen solchen in $G_2$, so werden in dem Punkte der Parabel die vier Wurzeln paarweise einander gleich und dann sämtlich komplex. Nun befinden sich aber innerhalb des Gebiet $G_2$ Gleichungen mit zwei gleichen Paaren komplexen Wurzeln, und es scheint, als ob wir hier Gleichungen mit Doppelwurzeln passieren könnten ohne Überschreitung der Diskriminanten-Fläche. Jedoch erstreckt sich mitten in dieses Gebiet hinein der andere Zweig unserer Parabel, und auf dieser liegen die Gleichungen mit zwei gleichen Paaren komplexen Wurzeln. Dieser Parabelzweig zeigt sich z.B in dem durch Gleichung 8.) charakterisierten Schnitt, indem diese Gleichung für $x = 3$ außer den Werten $y = \pm 4$ auch noch $y = 0$ ergiebt. Dieser letzte Punkt $(x = 3, y = 0, z = -1)$ ist ein Parabelpunkt, ebenso wie in dem aufgezeichneten Schnitt der Punkt $x = 3, y = 0, z = +1$. Alle vier Wurzeln sind nur gleich im Nullpunkt der Koordinaten, wo doppelkurve und Wendelkurve zusammentreffen und die dem Punkte entsprechende Gleichung $w^4 = 0$ ist.

Aus dieser Diskussion folgt, daß die drei Gebiete $G_0, G_1, G_2$ durch folgende Zeichenbestimmungen charakterisiert werden :

| | | | | | |
|---|---|---|---|---|---|
| $G_0$ | $D > 0$ | $z > 0$ | $3z^2 > x$ | | |
| $G_1$ | $D < 0$ | | | | |
| $G_2$ | $D > 0$ | $z > 0$ | $3z^2 < x$ | $D > 0$ | $z < 0$ |

228|229
|

Die Anzahl der komplexen Wurzeln wird bestimmt durch die Zahl der Zeichenwechsel der Reihe

$D, D_1, z, 1$ also z.B. für das Gebiet $G_1$ und
$z > 0$ - $+sgn\ D_1$ + +
$z < 0$ - $+sgn\ D_1$ - +

Da nur ein Zeichenwechsel vorhanden sein darf, so muß im ersten Falle $D_1 > 0$, im zweiten dagegen $D_1 < 0$ sein.- Für $z < 0$ wird durch das Vorzeichen von $D$ allein entschieden, in welschem Gebiete (ob in $G_1$ oder in $G_2$) wir uns befinden ; für $z > 0$ dagegen ist außerdem noch zu bestimmen, ob $3z^2 > x$ oder $3z^2 < x$ ist.

Die Flächen, welche eine Sonderung der Gebiete bewirken sollen, sind nach Angabe unserer Funktionenreihe außer $D = 0$ :

$$D_1 = 0 \text{ und } z = 0.$$

Die Fläche $D_1 = 0$ enthält, wie bereits angegeben, die Doppelkurve. An Stelle dieser Fläche können wir einfacher den Cylinder nehmen, welcher die Parabel als Grundkurve enthält. Keinesfalls aber können die beiden Gebilde $D_1 = 0$ und $z = 0$ durch ein einziges ersetzt werden, wie *Clebsch* in seinen Buche „Theorie der linären algebraischen Formen“ fälschlicher Weise ergiebt.

## IV.) Characteristik eines Systems von 3 Funktionen von je zwei Variabeln.

Wir hatten der Betrachtung, welche uns zu dem Begriff der Charakteristik eines Funktionensystems führte, die drei Funktionen

$$y - f(x)\ ,\ y - f_1(x)\ ,\ y$$

zu Grunde gelegt und die Lage der Kurven zu einander studiert, welche sich durch 229|230
Nullsetzen dieser Funktionen ergeben. Am Schluss jener Einleitung (S. 211) hatten wir bereits darauf hingewiesen, daß die Frage eine allgemeinere würde, wenn man drei ganz beliebige Funktionen von $x$ und $y$ :

$$f(x,y)\ ,\ \varphi(x,y)\ ,\ \psi(x,y)$$

und die entsprechenden Kurven unter dem angeführten Gesichtspunkte untersucht.

Wir setzen voraus, daß diese drei Funktionen eindeutig, stetig und differentiirbar sind, und daß die Kurven, deren analytischer Ausdruck die Gleichungen $f(x,y)=0$, $\varphi(x,y)=0$, $\psi(x,y)=0$ sind, eine geschlossene Gestalt haben. Diese letzere Eigenschaft kann arithmetisch folgendermaßen gekennzeichnet werden : Es soll möglich sein, wie bestimmte endliche Größe $R$ anzugeben, welche so beschaffen ist, daß für alle Werte $x,y$, welche einer der drei Gleichungen $f=0$, $\varphi=0$, $\psi=0$ Genüge leisten, die Ungleichung $x^2+y^2<R^2$ besteht.

Wir bilden nun die Determinante :

1. $\begin{vmatrix} f & \varphi & \phi \\ f_1 & \varphi_1 & \phi_1 \\ f_2 & \varphi_2 & \phi_2 \end{vmatrix} = \Delta$, wo der Index 1 die Ableitung nach $x$, der Index 2 diejenige nach $y$ bezeichnen möge. Wenn ferner $\Delta_1, \Delta_2, \Delta_3$ die Funktionaldeterminanten von bzw. $\varphi,\psi;\psi,f;f,\varphi$ bezeichnen, so lässt sich :
2. $\Delta = f\Delta_1 + \varphi\delta_2 + \psi\Delta_3$ schreiben, und $\Delta$ selbst ist, die Funktionaldeterminante von $(z.f, z.\varphi, z.\psi$ für $z=1$.

230|231 Untersuchen wir nun die durch unsere drei Kurven gebildete allgemeinere Konfiguration in derselben Weise wie oben die speziellere, so zeigt sich, daß der damaligen zweifachen Betrachtungsweise hier eine sechsfache gegenübersteht. Während wir nämlich a.a.O. entweder den Lauf der Kurve $f=0$ oder denjenigen der Kurve $f_1=0$ verfolgten und ihre Schnittpunkte mit der Absissenaxe betrachteten, können wir uns hier entweder auf der Kurve $f=0$ oder auf $\varphi=0$ oder schließlich auf der Kurve $\psi=0$ fortbewegen und die Schnittpunkte der betreffenden Kurve mit einer der beiden andere studieren. Verfolgen wir z.B. in der nebenstehenden Figur den Lauf der Kurve $f(x,y)=0$ von $O$ an, und beachten die Schnittpunkte mit der Kurve $\psi(x,y)=0$, so werden wir in $\xi_1$ einen Eintrittspunkt zu verzeichnen haben, da wir in diesem Punkte in Bezug auf das Produkt $\varphi(x,y).\psi(x,y)$ von dem Äußere in das Innere gelangen.

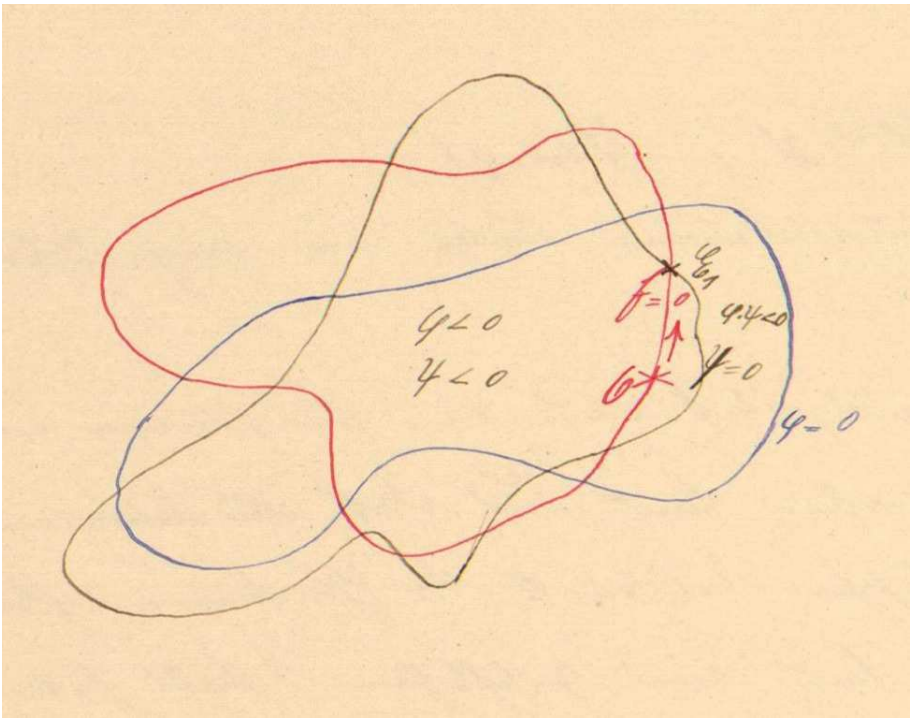
## Fortgangsprinzip

Es ist von großen Wichtigkeit, auf arithmetischen eine bestimmte Richtung zu fixieren, in welcher wir uns stets auf den Kurven bewegen wollen. Das soll in folgender Weise geschehen. Es sei $(x',y')$ ein fester Punkt auf der Kurve $f(x,y)=0$, auf welcher wir uns fortbewegen wollen und durch denselben eine beliebige Kurve $\Theta(x,y)=0$ gelegt. Dann soll für den Fall, wo wir die Schnittpunkte mit der Kurve $\varphi=0$ ins Auge fassen, $\frac{\partial|f,\Theta,\psi|}{\partial\psi}d\Theta>0$, d.h. :

$$3.)\ sgn\ d\Theta = sgn\ \frac{\partial|f,\Theta,\psi|}{\partial\psi} = \ sgn\ (f_1\Theta_2 - f_2\Theta_1)$$

sein. Durch die Gleichung 3) wird festgesetzt, ob wir im Punkte $(x',y')$ in Bezug auf die Kurve $\Theta(x,y)=0$ von Innere ins Äußere oder von Äußere ins Innere gelangen. Betrachten wir die Schnittpunkte mit der Kurve $\psi=0$, so wird die Richtung die
231|232 entgegengesetzte, da alsdann

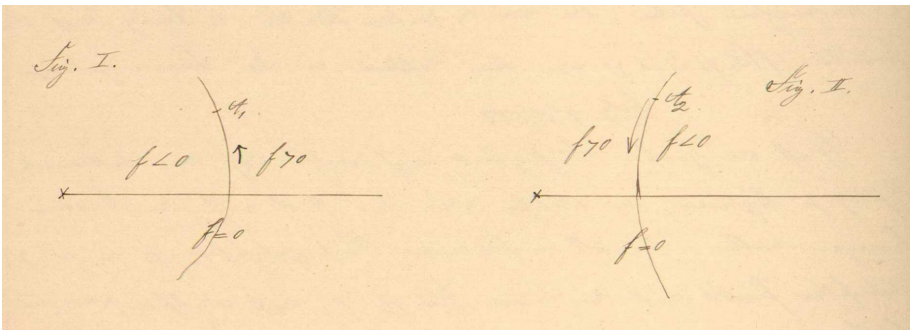
$$sgn\ d\Theta = sgn\ \frac{\partial|f,\varphi,\Theta|}{\partial\varphi}\ \text{sein muß.}$$




Die so fixierte Richtung bleibt dieselbe, was für eine Kurve $\Theta = 0$ wir auch nehmen. Dann, wenn z.B. $\Theta = y$ ist, so wird $d\Theta = dy$, $f_1\Theta_2 - f_2\Theta_1 = f_1$, also unsere Gleichung 3) :

$$4.)\ f_1.dy > 0.$$

Nun ist, da wir uns auf der Kurve $f = 0$ fortbewegen, für jeden Punkt derselben :




$f_1dx + f_2dy = 0$, $dx = -\dfrac{f_2dy}{f_1}$ und durch Einsetzung dieses Wertes in $d\Theta = \Theta_1dx + \Theta_2dy$ wird $d\Theta = (f_2\Theta_2 - f_2\Theta_1)\frac{dy}{f_1}$, d.h.

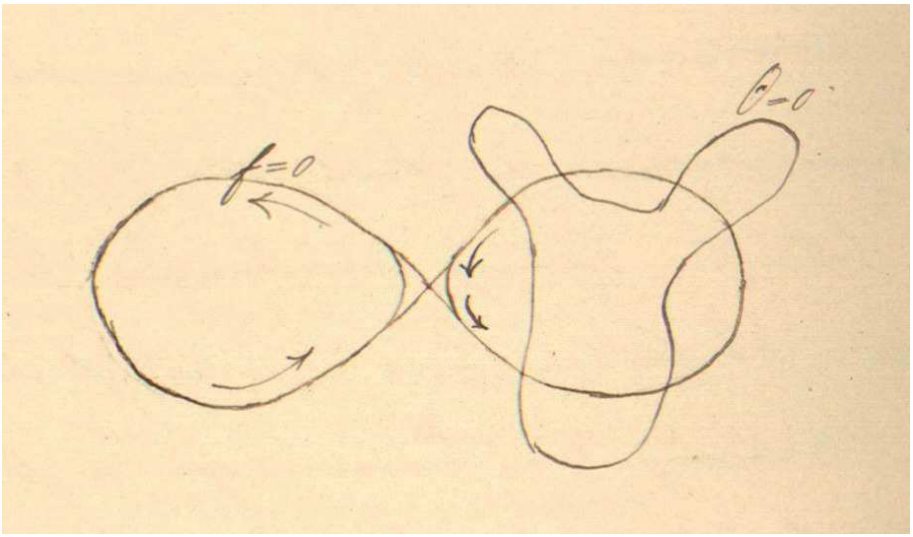
$$sgn\ d\Theta = sgn\ (f_1.\Theta_2 - f_2.\Theta_1).sgn\ \frac{dy}{f_1},$$

sodaß also unter Voraussetzung von $dy.f_1 > 0$ auch

$$sgn\ d\Theta = sgn\ (f_1.\Theta_2 - f_2.\Theta_1)$$

für beliebiges $\Theta$ ist. Diese Festlegung der Richtung - wir wollen es das Fortgangsprinzip nennen - ist für den Fall der Ebene gleichbedeutend mit der Vorschrift, die Kurve so zu durchlaufen, daß das Innere zur Linken liegt. So ist z.B. in Figur I für die Koor-

dinaten $x_1, y_1$ des Punktes $\mathcal{A}_1$ $f_1 > 0$, weil bei wachsenden $x$ und Konstanten $y$ die Werte der Funktion $f(x, y)$ zunehmen, also ist nach Gl. 4) $dy > 0$, die $y-$Koordinaten müssen zunehmen und die Kurve steigen. In Figur II dagegen nehmen die Werte von $f(x, y)$ ab, also müssen wir die Kurve in fallender Richtung durchlaufen. In beiden Fällen aber liegt das Innere zur linken Hand. Jedoch ist diese letztere Festsetzung schon für eine Bewegung in Raume nicht mehr anwendbar, da hier ein „rechts“ oder „links“ nicht unterschieden werden kann, und für den Eintritt in die Untersuchungen höherer Mannigfaltigkeiten wird es erst recht eine unerlässliche Bedingung, ein
232|233 Fortgansprinzip *arithmetisch* | zu fixieren.




Denken wir uns nun in der Funktion $f(x, y)$ die Variabeln $x$ und $y$ als Funktionen einer Variabeln $z$, $x = V(z)$, $y = W(z)$ und fragen, wie ich dann die Gleichung

3) $sgn \begin{vmatrix} f_1 & f_2 \\ \Theta_1 & \Theta_2 \end{vmatrix} = sgn\ d\Theta$, welche das Fortgangsprinzip definiert, ändert. Aus $f(V(z), W(z))$ folgt :

$f_1 V' + f_2 W' = 0$ , $d\Theta = (V'.\Theta_1 + W'\Theta_2)dz$, sodaß also 3) übergeht in

$(f_1\Theta_2 - f_2\Theta_1)(V'.\Theta_1 + W'.\Theta_2)dz > 0$. Setzen wir hierin $f_2 = -f_1.\frac{V'}{W'}$, so erhalten wir :

$$(f_1\Theta_2 + f_1.\frac{V'}{W'}.\Theta_1)(V'\Theta_1 + W'\Theta_2)dz = \frac{f_1}{W'}(V'\Theta_1 + W'\Theta_2)^2 dz > 0,$$

woraus als die gewünschte Bestimmung

$$sgn\ f_1.W' = sgn\ dz = -sgn\ f_2.V'$$

hervorgeht, während für die ursprüngliche Funktion $sgn\ f_1 = sgn\ dz$, oder $-sgn\ f_2 = sgn\ dx$ sein mußte. Hieraus geht hervor, daß dieses Fortgangsprinzip nicht anwendbar ist an Punkten, für welche $f_1 = f_2 = 0$ ist, sodaß Singularitäten, wie z.B. Doppelpunkte, nicht durchlaufen werden können, und man z.B. den Weg um die beiden Teile einer Lemniskate einzeln machen muß.

## Wert und Bedeutung der Charakteristik.

Sind $f(x, y) = 0$ und $\Theta(x, y) = 0$ zwei geschlossene Kurven, so wird die Anzahl der Schnittpunkte der beiden Kurven, an dann die Kurve $\Theta = 0$ steigt, ebenso groß

sein wie die Anzahl derjenigen, an welchen sie fällt. Es findet also dann die Gleichung statt :

6.) $\sum_{f=0,\Theta=0} sgn\ d\Theta = \sum_{f=0,\Theta=0} sgn \begin{vmatrix} f_1 & f_2 \\ \Theta_1 & \Theta_2 \end{vmatrix} = 0$. Dieser einfache Satz | 233|234
leistet uns für die Theorie der Charakteristik sehr wesentliche Dienste. Zuvörderst machen wir von ihn eine Anwendung auf den Fall, daß $\Theta = \varphi.\psi$ ist, indem wir voraussetzen, daß $\varphi(x,y) = 0$ und $\psi(x,y) = 0$ zwei geschlossen kurven darstellen und daher die durch das Produkt $\varphi.\psi = 0$ dargestellte kurve diese Eigenschaft habe.

Also muß nach Gleichung 6.) :

$$\sum_{f=0,\phi.\psi} sgn\ [f_1(\varphi_2.\psi + \varphi.\psi_2) - f_2(\varphi_1.\psi + \psi_1.\varphi)] = \sum_{f=0,\phi.\psi} sgn\ (\psi.\Delta_3 - \varphi.\Delta_2) = 0,$$

eine Gleichung, die auch in der Form :

7.) $\sum_{f=0,\varphi=0} sgn\ \psi.\Delta_3 - \sum_{f=0,\psi=0} sgn\ \varphi.\Delta_2 = 0$ geschrieben werden kann. Ebenso folgt aus 6.) :

8.) $\sum_{f=0,\varphi=0} sgn\ d(\varphi.\psi) = \sum_{f=0,\psi=0} sgn\ d(\varphi\psi)$, da in der letzten Summe die Richtung, in welcher man die Kurve $f = 0$ umkreist derjenigen der ersten Summe entgegengesetzt angenommen werden muß.

Nun ist ferner mit Benutzung der vorausgesetzten Bedingung $\frac{dy}{f_1} > 0$ :

$$\sum_{f=0,\varphi=0} sgn\ \psi.\Delta_3 = \sum_{f=0,\varphi=0} sgn\ \psi(f_1.\varphi_2 - f_2\varphi_1)$$

$$= \sum_{f=0,\varphi=0} sgn\ \psi.(f_1.\varphi_2 - f_2\varphi_1)\frac{dy}{f_1} = \sum_{f=0,\varphi=0} sgn\ \psi(\varphi_1 dx - \varphi_2 dy)$$

$$= \sum_{f=0,\varphi=0} sgn\ d(\varphi.\psi).$$

Also :

$$9.) \sum_{f=0,\varphi=0} sgn\ d(\varphi.\psi) = \sum_{f=0,\psi=0} sgn\ d(\varphi.\psi) = \sum_{f=0,\psi=0} sgn\ \varphi.\Delta_2$$

$$= \sum_{f=0,\varphi=0} sgn\ \psi.\Delta_3 = \sum_{f=0,\varphi=0} sgn\ \Delta = \sum_{f=0,\psi=0} sgn\ \Delta \text{ (Gl. 2))}.$$

Auf analoge Weise würden wir erhalten für $\Theta = f.\varphi$ : | 234|235

$$10) \sum_{\varphi=0,\psi=0} sgn\ d(\psi.f) = \sum_{\varphi=0,f=0} sgn\ d(\psi.f) = \sum_{\varphi=0,f=0} sgn\ \psi.\Delta_3$$

$$= \sum_{\varphi=0\psi=0} sgn\ f.\Delta_1 = \sum_{\varphi=0\psi=0} sgn\ \Delta = \sum_{\varphi=0,f=0} sgn\ \Delta$$

und da das letzte Glied der Gleichungen 10.) auch in 9.) vorkommt, so sind die Glieder der Reihe 9.) und diejenigen von 10.) unter einander gleich. Eine dritte Serie von Gliedere wurden wir erhalten für $\Theta = f.\varphi$. Insbesondere ist also :

$$11.) \sum_{f=0,\varphi=0} sgn\ \Delta = \sum_{\varphi=0,\psi=0} sgn\ \Delta = \sum_{\psi=0,f=0} sgn\ \Delta.$$

Ein zweite Anwendung der Formel 6.) machen wir für den Fall, daß $\Theta = \varphi$. Dann

erhalten wir aus 6.) :

$$12.) \sum_{f=\varphi=0} sgn\ \Delta_3 = 0 = \sum_{f=\varphi=0,\psi>0} sgn\ \Delta_3 + \sum_{f=\varphi=0,\psi<0} sgn\ \Delta_3$$

und da nach 9.) :

$$\sum_{f=\varphi=0} sgn\ \Delta = \sum_{f=\varphi=0,\psi>0} sgn\ \Delta_3 - \sum_{f=\varphi=0,\psi<0} sgn\ \Delta_3,$$

so erhalten wir durch Addition und Subtraktion dieser beiden Gleichungen :

$$\sum_{f=\varphi=0} sgn\ \Delta = 2 \sum_{f=\varphi=0,\psi>0} sgn\ \Delta_3 = -2 \sum_{f=\varphi=0,\psi<0} sgn\ \Delta_3,$$

und also :

$$13.)\ -\frac{1}{2} \sum_{f=\varphi=0} sgn\ \Delta = \sum_{f=\varphi=0,\psi<0} sgn\ \Delta_3$$

$$= \sum_{f=\varphi=0,\psi<0} sgn\ (f_1\varphi_2 - f_2\varphi_1) = -\frac{1}{2} \sum_{f=\varphi=0} sgn\ d(\varphi.\psi)\ \text{(Gl. 9.)).}$$

$$= \chi(f,\varphi,\psi).$$

Die in der Reihe 13.) enthalten Glieder sind verschiedene Ausdrücke für die Charakteristik der drei Funktionen $f, \varphi$ und $\psi$. Der Vollständigkeit halber wollen wir
235|236 alle möglichen verschiedenen Formen derselben zusammenstellen :
|

$$14.)\ \chi(f,\varphi,\psi) =$$

$$\begin{cases} -\frac{1}{2}\sum_{f=\varphi=0} sgn\ \Delta & -\frac{1}{2}\sum_{\varphi=\psi=0} sgn\ \Delta & -\frac{1}{2}\sum_{\psi=f=0} sgn\ \Delta \\ \sum_{\substack{\psi<0\\ f=\varphi=0}} sgn\ (f_1\varphi_2 - f_2\varphi_1) & \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1) & \sum_{\substack{\varphi<0\\ \psi=f=0}} sgn\ (\psi_1 f_2 - \psi_2 f_1) \\ -\frac{1}{2}\sum_{\varphi=\psi=0} sgn\ d(f.\varphi) & -\frac{1}{2}\sum_{\psi=f=0} sgn\ d(\varphi.\psi) & -\frac{1}{2}\sum_{f=\varphi=0} sgn\ d(\psi.f) \\ -\frac{1}{2}\sum_{f=\psi=0} sgn\ d(f.\varphi) & -\frac{1}{2}\sum_{\varphi=f=0} sgn\ d(\varphi.\psi) & -\frac{1}{2}\sum_{\psi=\varphi=0} sgn\ d(\psi.f) \end{cases}$$

Dabei ist für die letzten 6 Ausdrücke zu beachten, daß man, wie sich aus Gl. 8.) ergiebt, stets auf der Kurve sich fortzubewegen hat, welche unter dem Differentialzeichen nicht vorkommt und von uns unter dem Summenzeichen an die Zweite Stelle geschrieben ist. Nimmt man z.B. $\chi(f,\varphi,\psi) = -\frac{1}{2}\sum_{\psi=f=0}\ sgn\ d(\varphi.\psi)$, so hat man die Kurve $f = 0$ zu durchlaufen und ihre Schnittpunkte mit der Kurve $\psi = 0$ aufzusuchen. Die geometrische Bedeutung dieser Charakteristik ergibt sich dann sehr einfach aus der Betrachtung auf S.231. Für den Eintrittspunkt $\xi_1$ der dort gezeichneten Figur würde nämlich $d(\varphi.\psi) < 0$, für einen Austrittspunkt dagegen $d(\varphi.\psi) > 0$ sein, sodaß also, wenn $\mathcal{E}$ die Anzahl der Eintrittspunkte, $\mathcal{A}$ diejenige der Austrittspunkte bezeichnet :

$$\chi(f,\varphi,\psi) = \frac{1}{2}(\mathcal{E} - \mathcal{A})\ \text{ist.}$$

Man kann hieraus mit Hülfe der in der zweiten Zeile von 14.) aufgeführten Ausdrücke schließen, dasß die Differenz $(\mathcal{E} - \mathcal{A})$ stets eine grade Zahl ist und ferner, daß sie dieselbe bleibt, welche der mehrfach erwähnten sechs Kombinationen man auch der Beobachtung zu Grunde legt. Faßt man die Ausdrücke der zweiten Zeile von 14.)
236|237 näher ins Auge, so erkennt man, daß die Charakteristik auch eine algebraische | Zeichensumme für Wertsysteme liefert, welche zwei Gleichungen und einer Ungleichung

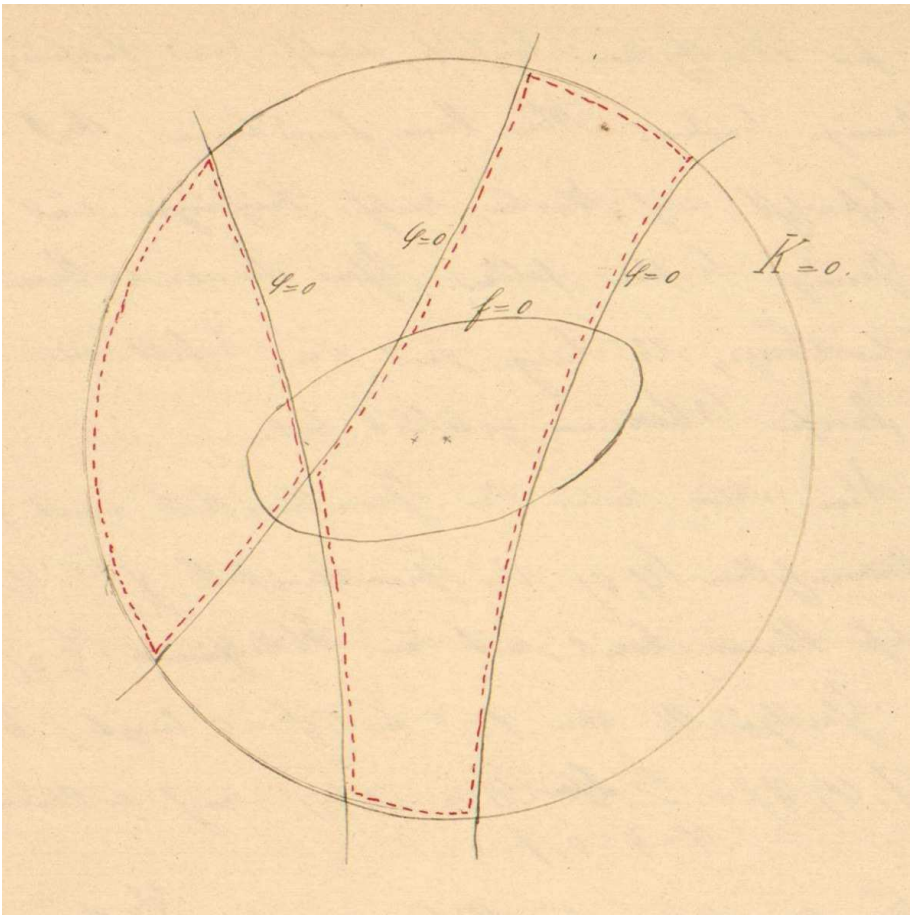
Genüge leisten. Man kann hinzufügen, daß die Charakteristik alle Fragen beherrscht nach Werten, welche Gleichungen und zugleich Ungleichungen Genüge leisten sollen. Schon *Fourier* erkannte die große Tragweite dieser Frage, die lange Zeit den Inhalt einer großen Preisaufgabe der Pariser Akademie gebildet hat.

Wenn wir unter der Charakteristik eines Punktes $(\xi, \eta)$ für ein Funktionensystem $(\varphi, \psi)$ die Charakteristik $\chi(\theta, \varphi, \psi)$ verstehen, wo $\theta$ ein sehr kleiner Kreis mit dem Mittelpunkt $(\varphi, \psi)$ ist, in welchem höchstens ein Schnittpunkt von $\varphi = 0$ und $\psi = 0$ liegt, so können wir

$$\chi(f, \varphi, \psi) = \sum_{\substack{f<0 \\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1)$$

auch erklären als die Summe der Charakteristiken sämtlicher einzelner Punkte, welche in Innern der Kurve $f = 0$ liegen.

Ausdehnung der Resultate auf den Fall nicht geschlossener Kurven.




Die vorstehenden Untersuchungen würden kaum nennenswerte Anwendungen zulassen, wenn sie nicht auf Funktionen übertragbar wären, welche der Bedingung, im Endlichen geschlossene Kurven zu repräsentieren, nicht mehr genügen. Es ist daher nötig, den Begriff der Charakteristik auch für solche Funktionensysteme als gültig nachzuweisen, welchen ausschließlich die Erfüllung gewisser Stetigkeits-Bedingungen auferlegt wird. Dazu verfahren wir in folgender Weise.

Wir denken uns einen Kreis $K = R^2 - x^2 - y^2 = 0$ von solchen Umfange, daß er sämtliche Schnittpunkte von je Zweien der Kurven umschließt und betrachten an Stelle der nicht geschlossenen Kurven $f = 0$, $\varphi = 0$ und $\psi = 0$ die Kurven $F = f.K = 0$, $\Phi = \varphi.K = 0$, $\Psi = \psi.K = 0$.

237|238 | Wenn wir dieselben in Sinne des Fortgansprinzips durchlaufen, so wird jede von ihnen in eine Anzahl von geschlossenen Kurven zerfallen. So stellen in der nebenstehenden Figur z.B. die zwei punktierten geschlossenen Kurven das für uns allein in Betracht kommende geometrische Bild von $\varphi.K = 0$ dar, welches an die Stelle der drei Zweige $\varphi = 0$ tritt. Nun definieren wir :

$$\chi(f, \varphi, \psi) = \chi(F, \Phi, \Psi)$$

und weil :

$$\chi(F, \Phi, \Psi) = \sum_{\substack{F<0 \\ \Phi=\Psi=0}} sgn\ (\Phi_1\Psi_2 - \Phi_2\Psi_1) = \sum_{\substack{f<0 \\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1).K^2$$

(denn die Schnittpunkte der Kurven $\Phi = 0$ und $\Psi = 0$ sind innerhalb das Kreises $K = 0$ genau dieselben wie diejenigen von $\varphi = 0$ und $\psi$ und die Forderung $F < 0$ reduziert sich, da für alle in Betracht kommenden Punkte $K > 0$ ist, auf $f < 0$), so ist :

$$\begin{aligned} 15.)\ \chi(f, \varphi, \psi) &= \sum_{\substack{f<0 \\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1) \\ &= \sum_{\substack{\varphi<0 \\ \psi=f=0}} sgn\ (\psi_1 f_2 - \psi_2 f_1) \\ &= \sum_{\substack{\psi<0 \\ f=\varphi=0}} sgn\ (f_1\varphi_2 - f_2\varphi_1), \end{aligned}$$

da wir je auch in dem Ausdruck $\chi(F, \Phi, \Psi)$ die Funktionen unter einander vertauschen können. Für die Charakteristik des allgemeineren Funktionensystems $\chi(f, \varphi, \psi)$ bleiben also drei von den zwölf Erklärungen in 14.) in Kraft.

Die in der Einleitung zu diesen Betrachtungen interpretierte Zeichensumme können wir jetzt auch als Charakteristik eines Systems von drei Funktionen erklären. Wir bilden nämlich aus den drei Funktionen $f(x) - y, f_1(x) - y, y$ für welche wir die
238|239 Konfiguration der entsprechenden Kurven a.a.O. betrachtet hatten : |

$$\chi(f(x) - y, f_1(x) - y, y) = -\frac{1}{2} \sum_{\substack{f(x)-y=0 \\ y=0}} sgn \begin{vmatrix} f(x) - y & f_1(x) - y & y \\ f'(x) & f_1'(x) & 0 \\ -1 & -1 & 1 \end{vmatrix}$$

$$= -\frac{1}{2} \sum_{\substack{f(x)-y=0 \\ y=0}} sgn\ (f(x).f_1'(x) - f'(x).f_1(x)) = +\frac{1}{2} \sum_{(\xi)} f'(\xi).f_1(\xi).$$

Der letztere Ausdruck ist dieselbe Zeichensumme, welche wir a.a.O. zu deuten unternommen hatten.

Nach der für alle stetigen Funktionen gültigen Erklärung der Charakteristik ist [5] :

$$16.)\ \chi(f_1, \varphi - u.f, \psi - v.f) = \sum_{\varphi - u.f = \psi - v.f = 0, f < 0} sgn \begin{vmatrix} \varphi_1 - u.f_1 & \varphi_2 - u.f_2 \\ \psi_1 - v.f_1 & \psi_2 - v.f_2 \end{vmatrix}$$

$$\sum_{f = \psi - v.f = 0, \varphi - u.f < 0} sgn \begin{vmatrix} f_1 & f_2 \\ \psi_1 - v.f_1 & \psi_2 - v.f_2 \end{vmatrix} = \sum_{f = \psi = 0, \varphi < 0} \begin{vmatrix} f_1 & f_2 \\ \psi_1 & \psi_2 \end{vmatrix} = \chi(f, \varphi, \psi)$$

Denken wir uns jetzt die Koeffizienten der Funktionen $f, \varphi$ und $\psi$ abhängig von einem variablen Parameter $t$, so erkennt man aus dem zweiten Glied der Gleichung 16.), daß die Charakteristik $\chi(f, \varphi, \psi)$ sich nur dann ändern kann, wenn für mindestens ein Wertpaar $(x, y)$ gleichzeitig :

$$\varphi - u.f = 0\ ,\ \psi - v.f = 0\ ,\ \begin{vmatrix} \varphi_1 - u.f_1 & \varphi_2 - u.f_2 \\ \psi_1 - v.f_1 & \psi_2 - v.f_2 \end{vmatrix} = \Delta_1 + u.\Delta_2 + v.\Delta_3 = 0$$

ist. Führt man in den letzten Ausdruck die Wert $u = \frac{\varphi}{f}$, $v = \frac{\psi}{f}$ ein, so ergeben sich die drei Bedingungen :

$$17.)\ \varphi - u.f = 0\ ,\ \psi - v.f = 0\ ,\ \Delta = 0.$$

Aus $\Delta = 0$ erhalten wir für $x$ und $y$ nur Werte, welche von $u$ und $v$ | unabhängig 239|240
sind ; es müß also, da $u$ und $v$ Unbestimmte sind, für jedes solches Wertepaar

$$f = 0\ ,\ \varphi = 0\ ,\ \psi = 0 \text{ sein.}$$

Setzen wir nun voraus, daß wir, ähnlich wie früher für zwei Funktionen von einer Variablen eine symetrische Funktion der Wurzeln von $f = 0$ , $\varphi = 0$ , $\psi = 0$ finden können, deren Verschwinden notwendige und hinreichende Bedingung für die Existenz eines gemeinsamen Teilers der drei Funktionen ist, und daß wir dieser symmetrischen Funktion der Wurzeln auch den Namen „Resultante“ beigelegt hätten, so können wir sagen :

Die Charakteristik eines Systems von drei Funktionen ändert ihren Wert bei einer stetigen Variirung des Koeffizientensystems ihrer Funktionen dann und nur dann, wenn die Resultant der drei Funktionen verschwindet.

Denken wir uns ebenso wie früher die Koeffizienten von $f$, $\varphi$ und $\psi$ als Funktionen von $\nu$ unabhängigen Variablen, so entspricht jeden Punkte der von den letzteren fixierten $\nu-$fachen Mannigfaltigkeit ein Funktionensystem mit seiner Charakteristik. Die $(\nu - 1)-$fache Mannigf. $R = 0$, wenn $R$ wieder die Resultante bezeichnet, trennt die Gebiete von einander, denen Funktionensystem mit verschiedener Charakteristik angehören, sodaß man sicher niemals, ohne die Mannigf. $R = 0$ zu überschreiten, von einem Funktionensysteme zu einem andere gelangen kann, dessen Charakteristik von

5. Version corrigée sur la page de gauche :

$$\sum_{f = \psi - v.f = 0, \varphi - u.f < 0} sgn \begin{vmatrix} \psi_1 - v.f_1 & \psi_2 - v.f_2 \\ f_1 & f_2 \end{vmatrix} = \sum_{f = \psi = 0, \varphi < 0} \begin{vmatrix} \psi_1 & \psi_2 \\ f_1 & f_2 \end{vmatrix} = \chi(f, \varphi, \psi)$$

Note sur la page de gauche :

> Es wird sich zeigen, dass dieses Fall ohne $f \neq 0$ nicht eintreten kann.
> Der Fall $f = 0$ wird nachher behandelt werden.

240|241 der des ersten verschieden ist.
|

## V.) Anwendung der Resultate zur Herleitung des *Gauss*'schen Fundamentalsatzes der Algebra und des Integrals von *Cauchy*.

Unsere Untersuchungen über die Charakteristik eines Funktionensystems setzen uns in den Stand, ohne große Mühe diejenige Eigenschaft ganzer Funktionen zu beweisen, welche die Inhalt des von *Gauss* als den „Fundamentalsatz der Algebra" bezeichneten Theorems ausmacht. Dasselbe sagt folgendes aus :

Es sei $\Theta(x+yi)$ eine ganze Funktion $n-$ten Grades der komplexen Variablen $(x+yi)$ mit komplexen Koeffizienten und nach Sammlung aller reellen und komplexen Glieder :

$$\Theta(x+iy) = \varphi(x,y) + i.\psi(x,y);$$

dann haben die durch $\varphi = 0$ und $\psi = 0$ definierten Kurven stets $n$ Schnittpunkte.

Es sei zunächst $\Theta(x+yi)$ eine beliebige stetige und monogene Funktion, so ist bekanntermaßen :

$$\varphi_1 = \psi_2 \ , \ \varphi_2 = -\psi_1.$$

Wenn nun $f(x,y) = 0$ eine geschlossene Kurve darstellt, so ist die Charakteristik :

$$1.)\ \chi(f,\varphi,\psi) = \sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn \begin{vmatrix} \varphi_1 & \varphi_2 \\ \psi_1 & \psi_2 \end{vmatrix} = \sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn\ (\varphi_1^2+\varphi_2^2) = \sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn\ 1;$$

d.h. $\chi(f,\varphi,\psi)$ gibt uns die Anzahl der Schnittpunkte der Kurven $\varphi = 0$ und $\psi = 0$ an, welche in dem Gebiete $f < 0$ liegen.

Den Inhalt des ausgesprochenen Theorems erhalten wir für den besonderen Fall, daß $\Theta(x+yi)$ eine ganze Funktion von der Form

$$2.)\ \Theta(x+yi) = \varphi(x,y) + i.\psi(x,y) = (x+yi)^n - c_1(x+yi)^{n-1} + ... \pm c_n$$

241|242 und | $f(x,y) = 0$ ein Kreis ist mit so großen Radius, daß alle Schnittpunkte der Kurven $\varphi = 0$ und $\psi = 0$ in seinem Innern liegen.
Die Möglichkeit einen solchen Kreis zu finden, wollen wir zunächst aufzeigen. Wenn wir $x + yi = R.e^{vi}$ setzen, so wird

$$\Theta(x+yi) = R^n.e^{nvi}\left\{1 - c_1.\frac{e^{-vi}}{R} + c_2.\frac{e^{-2vi}}{R^2} - ....\right\},$$

also :

$$\left|c_1.\frac{e^{-vi}}{R} + c_2.\frac{e^{-2vi}}{R^2} - ....\right| \leq \frac{|c_1|}{R} + \frac{|c_2|}{R^2} + ...$$

und wenn $c_k < u$ (für $k = 1,2,...n$) : $< u.\frac{1}{R}.\frac{1-\frac{1}{R^n}}{1-\frac{1}{R}} < \frac{u}{R-1} < 1$, wenn $R$ so bestimmt wird, daß : $u < R-1$, d.h. $R > u+1$ oder $R > |c_k|+1$ $(k = 1,2,...n)$. Es ist also : $\Theta(R.e^{vi}) = \varphi + i.\psi = R^n.e^{nvi}\alpha$, wo $0 < |\alpha| < 1$ ist.

Der Kreis mit dem Radius $R$ von der angegebenen Größe ist also so beschaffen, daß nur in seinem Innere (nicht außerhalb und auf der Peripherie) die Kurven $\varphi = 0$ und $\psi = 0$ sich schneiden können.

Nun wissen wir, daß die Charakteristik $\chi(f,\varphi,\psi)$ sich nur dann ändert, wenn gleichzeitig $f = 0$, $\varphi = 0$ und $\psi = 0$ ist. In unseren Falle kann also eine Änderung

derselben keinesfalls dann eintreten, wenn wir die Werte der Koeffizienten $c_1, c_2 ... c_n$ verkleinere, denn der Kreis $f = 0$ behält dann *a fortiori* die vorgeschriebene Eigenschaft, und es kann niemals zu einem gemeinsamen Schnitte der Drei Kurven kommen. Wenn wir also :

$$3.)\ \varphi_0 + i.\psi_0 = (x + yi)^n - 1 = 0$$

setzen, so ist auch :

$$\chi(f, \varphi, \psi) = \chi(f, \varphi_0, \psi_0),$$

und da die Charakteristik $\chi(f, \varphi_0, \psi_0)$ ihrer in 1.) angegebenen Bedeutung gemäß offenbar gleich $n$ ist, so ist auch für jede ganze Funktion 2.) :

$$\sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn\ 1 = n,$$

d.h. die Anzahl aller Schnitt- |-punkte der kurven $\varphi = 0$ und $\psi = 0$ ist so groß als der 242|243
Grad jener ganzen Funktion.

Der Kern des ersten und vierten Beweises dieses Satzes von *Gauss* (1799 bzw. 1849) ist im Grunde genommen nichts anderes als die Bestimmung der Charakteristik. In Bezug auf die Bedeutung, welche *Gauss* dem Theorem beilegte, und seinem algebraischen Äquivalente müssen wir auf die später folgenden Erörterungen verweisen.

Nicht mehr Schwierigkeiten bietet uns das Integral von Cauchy, welches die Summe der sogen. Ordnungszahlen einer Funktion für alle Punkte innerhalb eines bestimmten Gebietes angiebt, abzuleiten. Die Ordnungszahl ist Null, wenn die Funktion für die Koordinaten des Punktes einen endlichen, von Null verschiedenen Wert hat, sie ist gleich $+a$, wenn der Punkte für die Funktion, wie man sagt, ein $a-$facher Nullpunkt, gleich $-a$, wenn er ein $a-$facher Unendlichkeitspunkt für dieselbe ist. Die Summe dieser Ordnungszahlen der Funktion $\Theta(x + yi)$ für alle Punkte des Gebietes $f(x, y) < 0$ wird dann nach *Cauchy* durch das Integral[6]

$$\int d \log\ \Theta(x + yi)$$

angegeben (erstreckt über $f(x, y) = 0$). Es ist zu beachten, daß Funktionen $\Theta(x+yi)$, welche für endlich Werte des Arguments unendlich groß werden, in unseren Betrachtungen nicht vorkommen können, da eine solche Eigenschaft den geforderten Stetigkeitsbedingungen widersprechen würde.

Betrachten wir nunmehr die Charakteristik $\chi(f, \varphi, \psi) = -\frac{1}{2} \sum_{f=\varphi=0} sgn\ d(\varphi, \psi)$ für den speziellen Fall der mit a.) bezeichneten Figur. (s. umstehend ).

Wir bewegen uns auf $f = 0$ und erhalten in Beziehung auf das Produkt $\varphi.\psi$ die Eintrittspunkte $E_1, E_2, ... E_7$ und die | Austrittspunkte $A_1, A_2, A_3$, sodaß 243|244

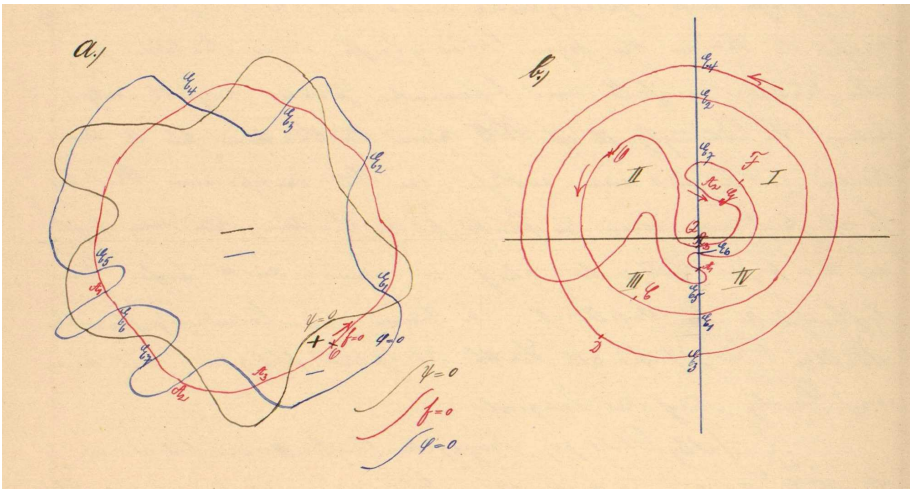
$$\chi = \frac{1}{2}(E - A) = \frac{7 - 3}{2} = 2$$

wird. Diese Zahl ist einer höchst einfachen Interpretation fähig. Bilden wir nämlich die Figur a) nach Maßgabe der Gleichungen

$$4.)\ \xi = \varphi(x, y)\ ,\ \eta = \psi(x, y)$$

ab, so transformiert sich die Kurve $\varphi = 0$ in die $y-$axe, $\psi = 0$ in die $x-$axe und $\varphi.\psi = 0$ hat in Figur b.) nur zwei innere Gebiete, nämlich die Quadranten II und IV sowie zwei äußere, die Quadranten I und III. Bestimmen Wir dann aus 4.) $x$ und $y$ als Funktionen von $\xi$ und $\eta$ und transformieren $f(x, y)$ in eine Funktion von $\xi$ und $\eta$, so

6. Correction sur la page de gauche : $\frac{1}{2\pi i} \int d \log\ \Theta(x + yi)$




bildet sich die Kurve $f(x,y)=0$ in Figur a.) in eine anders gestaltete in Figur b.) ab. Der Punkt $O$ der ersteren liegt in b.) in zweiten Quadranten und beim Fortgang von diesem Punkte in der Richtung des Pfeils muß die Kurve in b.) entsprechend ihrem Laufe in Figur a.) zuerst die $x-$axe, dann die $y-$axe, alsdann wieder die $x-$axe u.s.w. überschreiten. Die Punkte $E_1 ... E_7$; $A_1, A_2, A_3$ liegen natürlich sämtlich auf der $y-$axe. Ferner ist ersichtlich, daß in Fig. b.) für alle Eintrittspunkte (d.h. beim Übergang von I nach II oder von III nach IV) die Bewegungsrichtung derjenigen des Uhrzeigers entgegengesetzt ist, für die Austrittspunkte dagegen (d.h. beim Übergang von II nach I oder von IV nach III) dieselbe mit derjenigen des Uhrzeigers übereinstimmt. Will man die vollständigen Windungen der Kurve $f(x,y)=0$ um den Nullpunkt in Fig. b.) zählen, so hat man, wenn man zwei Mal hinter einander einen Eintrittspunkt passiert, eine Windung zu zählen und zu beachten, daß jeder Austrittspunkt den
244|245 vorhergehenden Eintrittspunkt aufhebt. Denn, wenn ich z.B. (Fig b.)) | von $\mathcal{C}$ aus mich über $E_1$ und $E_2$ fortbewege, so kommen ich in $\mathcal{D}$ in den Quadranten III zurück und habe den Anfangspunkt der Koordinaten ein Mal umkreist ; bewege ich mich dagegen von $\mathcal{F}$ über $E_7$ und $A_2$ nach $\mathcal{G}$, so gelange ich zwar auch den Quadranten I zurück, aber ohne den Nullpunkt umlaufen zu haben. Es ist also die Anzahl $\mathcal{A}$ der Windungen :

$$5.)\ \mathcal{A}=\frac{1}{2}(E-A)=\chi(f,\varphi,\psi).$$

Wir können uns die Definition der Anzahl der Windungen in folgender Weise veranschaulichen. Es sei in $Q$ ein Nagel vertikal zur Ebene angebracht und in die Lage der Kurve $f(x,y)=0$ ein Faden gelegt ; greifen wir dann den letzteren an einer beliebigen Stelle und ziehen ihn an, so wird er zwei Mal um den Nagel gewunden bleiben.

Setzen wir jetzt, um diese Bedeutung der Charakteristik analytisch zu verwerten :

$$\xi=r.\cos v\ ,\ y=r.\sin v,$$

so ist also :

$$\chi(f,\varphi,\psi)=\frac{1}{2\pi}\int dv,$$

das Integral erstreckt über die Kurve $f(x,y)=0$, und da

$$\frac{y}{\xi}=\operatorname{tg} v\ ,\ v=\operatorname{arctg}\frac{y}{\xi},$$

so wird

$$6.)\ \chi(f,\varphi,\psi) = \frac{1}{2\pi}\int_{f(x,y)=0} d\ \text{arctg}\ \frac{\psi}{\varphi} = \frac{1}{2\pi}\int_{f(x,y)=0} \frac{\varphi.d\psi - \psi.d\varphi}{\varphi^2+\psi^2}.$$

Führen wir auch für $x$ und $y$ Polarkoordinaten ein, indem wir setzen :

$$x = \rho.\cos\omega\ ,\ y = \rho.\sin\omega,$$

sodaß $\varphi(x,y) = \Phi(\omega)$, $\psi(x,y) = \Psi(\omega)$, da es nur auf die Lage der Kurven zu einander ankommt, so ist also auch :

$$\chi(f,\varphi,\psi) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\Phi\Psi' - \Psi\Phi'}{\Phi^2+\Psi^2}dw.$$

| Da die Charakteristik eine ganze Zahl ist, so ist es nur nötig, das Integral in der letzten Formel bis auf die erste Dezimale genau zu berechnen, um den Wert der Charakteristik zu erhalten. 245|246

Das Integral von *Cauchy* ergiebt sich für speziellen Fall der besonderen Beschaffenheit von $\varphi$ und $\psi$, wenn

$$\varphi(x,y) + i.\psi(x,y) = \Theta(x+yi)$$

eine Funktion der Komplexen Variablen $x+yi$ ist. Da nämlich :

$$-i.d\log(\varphi+\psi i) = -i\frac{d\varphi + id\psi}{\varphi+i\psi} = -i\frac{(d\varphi+id\psi)(\varphi-\psi i)}{\varphi^2+\psi^2}$$

$$= -i.\frac{\varphi d\varphi + \psi d\psi}{\varphi^2+\psi^2} + \frac{\varphi d\psi - \psi d\varphi}{\varphi^2+\psi^2},$$

so ist :

$$\frac{\varphi d\psi - \psi d\phi}{\varphi^2+\psi^2} = -i.d\log(\varphi+\psi i) + \frac{i}{2}d\log(\varphi^2+\psi^2)$$

und nach 6.) :

$$\chi(f,\varphi,\psi) = \frac{1}{2\pi}\int_{f(x,y)=0} d\log\Theta(x+yi) + \frac{i}{4\pi}\int_{f(x,y)=0} d\log(\varphi^2+\psi^2)$$

da $d\log(\varphi^2+\psi^2)$ eine eindeutige Funktion ist, so verschwindet das zweite Integral der letzten Gleichung nach einen bekannten Satze von *Cauchy* und es bleibt :

$$7.)\ \chi(f,\varphi,\psi) = \frac{1}{2\pi}\int_{f(x,y)=0} d\log\Theta(x+yi),$$

ein Resultat, welches wegen der Bedeutung der linken Seite von 7.) mit dem eingangs sitierten Satze von *Cauchy* übereinstimmt.

Spezialisieren wir die Formel 7.) noch weiter, indem wir annehmen, $\Theta(x+yi) = \varphi + i\psi$ sei eine ganze Funktion von $n-$ten Grade und $f(x,y) = 0$ ein Kreis, welcher alle Schnittpunkte der Kurven $\varphi = 0$ und $\psi = 0$ einschließt, so erhalten wir als Korrolar den Satz von *Gauss*. Es sei nämlich : 246|247

|

$$\Theta(x+yi) = (x+yi)^n - c_1(x+yi)^{n-1} + ... \pm c_n$$

und daher, wenn man [7]

$$x + yi = R.e^{\omega i}$$

setzt :

$$8.)\ \chi(f,\varphi,\psi) = \frac{1}{2\pi i}\int_0^{2\pi} \frac{n.R^{n-1}.e^{(n-1)\omega i} - (n-1)R^{n-2}.e^{(n-2)\omega i} + ...}{R^n.e^{n\omega i} - c_1R^{n-1}.e^{(n-1)\omega i} + ...}.d(R.e^{\omega i})$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \frac{n.R^n.e^{n\omega i} - (n-1)R^{n-1}.e^{(n-1)\omega i} + ...}{R^n.e^{n\omega i} - c_1R^{n-1}.e^{(n-1)\omega i} + ...}.d\omega$$

$$= \frac{n}{2\pi}\int_0^{2\pi} d\omega.\frac{1 - \frac{n-1}{n.R.e^{\omega i}} + ...}{1 - \frac{c_1}{R.e^{\omega i}} + ...}$$

Lassen wir den Radius $R$ immer größer werden, so nähert sich der Ausdruck auf der rechten Seite mehr und mehr der Zahl $n$, $\chi(f,\varphi,\psi)$ bleibt ungeändert, und da es eine ganze Zahl sein müß, so haben wir hier auf eine zweite, von der früheren unabhängigen Weise den Inhalt des *Gauss*'schen Theorems in aller Strenge bewiesen.

Der Begriff der Charakteristik und die daran angeknüpften Untersuchungen lassen sich ausdehnen auf ein System von $(n+1)$ Funktionen von $n$ Variabeln, bezw. auf $n-$fache Mannigfaltigkeiten. Die Windungen einer Spirale hat schon *Gauss* durch ein Integral ausgedrückt, *Tait* hat die Theorie der Windungen und ihre Berechnung in einem Buche ausführlich erörtert.

## VI.) Klärung des Begriffs der Charakteristik, der reellen Wurzeln und der Zerlegung einer ganzen Funktion in reellen Funktionen.

### Zur Geschichte des Fundamentalsatzes der Algebra

Die Geschichte des Fundamentalsatzes der Algebra ist außerordentlich lehrreich für die Entwicklungsgeschichte der Mathematik überhaupt. Nicht in grader Linie
247|248 findet der Fortschritt statt : durch mancherlei Krümmungen | und Windungen führt die mathematische Wissenschaft den Forscher stets zu den ersten Quellen zurück und läßt ihm nicht mehr und nicht weniger übrig als eine liefern Erkenntnis ihrer wenigen einfachen Grundwahrheiten. An sehr viele Probleme hat sich der Ausbau ganz neuer Theorien, haben sich neue Begriffsbildungen angeschlossen, die kühnsten Hoffnungen wurden darauf gesetzt, bis man zu der Erkenntnis gelangte, daß auch mit den früheren einfachen Mitteln dasselbe Ziel erreichbar sei und zwar jetzt mit der höheren Einsicht, welche der Durchgang durch die neuen Theorieen verliehen hatte. Eine solche Entwicklung hat auch unser Problem erfahren müssen. Während die griechischen Mathematiker nur rationale Zahlen als Wurzele von Gleichungen zuließen, führten die Inder auch Irrationalitäten als Wurzeln ein. Cardano betrachtete auch die imaginären Größen und führte sie als „unmögliche“ Wurzeln von Gleichungen ein.

Die Methode, die Gleichungen zweiten, dritten und vierten Grades aufzulösen, indem man „die Unbekannte“ *explicite* als irrationale Funktion der Koeffizienten darstellte, verleitete zu dem Glauben, daß jede Gleichung beliebigen Grades auf solche

7. Correction sur la page de gauche :

$$8.)\ \chi(f,\varphi,\psi) = \frac{1}{2\pi i}\int_0^{2\pi} \frac{n.R^{n-1}.e^{(n-1)\omega i} - c_1(n-1)R^{n-2}.e^{(n-2)\omega i} + ...}{R^n.e^{n\omega i} - c_1R^{n-1}.e^{(n-1)\omega i} + ...}.d(R.e^{\omega i})$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \frac{n.R^n.e^{n\omega i} - c_1(n-1)R^{n-1}.e^{(n-1)\omega i} + ...}{R^n.e^{n\omega i} - c_1R^{n-1}.e^{(n-1)\omega i} + ...}.d\omega = \frac{n}{2\pi}\int_0^{2\pi} d\omega.\frac{1 - c_1\frac{n-1}{n.R.e^{\omega i}} + ...}{1 - \frac{c_1}{R.e^{\omega i}} + ...}$$

Art durch Wurzelzeichen auflösbar sein müssen. Man glaubte zu der Annahme berechtigt zu sein, daß sich jede ganze rationale Funktion einem Produkt von ebenso viel linearen Faktoren gleichsetzen lasse als ihr Grad eingäbe, und in der That konnte auch nur diese zwiefache Art der Darstellung zur Auffindung von Eigenschaften der Wurzeln algebraischer Gleichungen und zu deren Beweise führen. Die bedeutenden Mathematiker, wie *Euler* und *D'Alembert* sahen zwar sehr wohl ein, daß der Satz, welcher aussagt, eine jede algebr. Gleichung $n-$ten Grades | habe höchstens $n$ Wurzeln von der Form $a + bi$, eines Beweises unbedingt bedürftig sei. Aber sie hatten sich in Betreff der Beschaffenheit der Wurzeln algebraischer Gleichungen so phantastischen Fiktionen hingegeben, daß ihre Beweise gänzlich haltlos sein mußten, und der Aberglaube, daß jede Gleichung durch Wurzelzeichen auflösbar sei, hatte sich so fest eingebürgert, daß *Euler* mit Hülfe dieses unrichtigen Satzes jenen richtigen Satz zu beweisen unternahm. 248|249

Da war es wieder *Gauss*, der das erlösende Wort sprach. In seiner berühmten Dissertation (*Demonstratio nova theorematis 1799*) deckte er die Schwächen der bisherigen Beweise auf, erschütterte den Glauben an die Auslösbarkeit der allgemeinen Gleichung durch Wurzelzeichen und gab einen neuen Beweis für das Theorem, welchem er später noch drei andere (1815, 1816, 1849) hinzufügte. Diese Beweise sind zwar alle von epochemachender Wichtigkeit gewesen, und besonders der zweite derselben ist von der weitesttragenden Bedeutung für die Theorie der algebraischen Gleichungen geworden. Aber eines blieb unaufgeklärt : die Natur der reellen Wurzeln. *Gauss* begnügte sich, ihre Bedeutung geometrisch zu veranschaulichen und war der festen Überzeugung, damit eine ausreichend Definition derselben gegeben zu haben. Was er beweisen wollte und wirklich aufs strengste bewiesen hat, ist im vorhergehenden Abschnitt des Näheren angegeben worden. Man könnte es auch damit als Beweis für den rein analytischen Inhalt des Fundamentaltheorems, daß nämlich, wie man gewöhnlich sagt, jede Gleichung $n-$ten Grades $n$ Wurzeln habe, sein Bewenden haben lassen, wenn man „die geometrische Stetigkeit ohne weiteres auf die arithmetischen Größen übertragen, und also, um es kurz | anzudeuten, einer jeden Streckenlänge eine arithmetische Zahl entsprechen lassen“ [8] wollte. Das werden alle diejenigen Mathematiker mit um so weniger Bedenklichkeit thun, welche auch die irrationalen und imaginären Größen durch geometrische Quantitäten als „definiert“ erachten. Daß Gauss auch dieser letzteren Ansicht war, geht aufs deutlichste aus einem Ausspruch hervor, welchen er in einer Vorlesung der Göttinger Societät gethun hat. (abgedruckt in den Göttinger gelehrten Anzeigen 1831, 64. März, Werke V) [9]. Er sagt dort am Schlusse dieses seines Versuchs, dem imaginären Größen eine feste Unterlage zu geben : „Hier ist also die Nachweisbarkeit einer anschaulichen Bedeutung von $\sqrt{-1}$ vollkommen gerechtfertigt, und mehr bedarf es nicht, um diese Größe in das Gebiet der Gegenstände der Arithmetik zuzulassen.“ Die Mathematiker, welche eine solche, uns sehr bedenklich scheinende Ansicht zu der ihrigen machen, lassen dabei völlig außer Acht, daß zwar die arithmetischen Beziehungen auf die anschaulichen geometrischen Begriffe übertragbar und anwendbar sind, und daß die Geometrie freilich stets ganz wesentlich zum Fortschritt der arithmetischen Wissenschaften beigetragen hat und beitragen wird, daß aber niemals der rein „arithmetische“ Begriff durch geometrische Veranschaulichung definiert werden kann und definiert werden darf. 249|250

So glauben viele z.B. den Begriff der Quadratwurzel aus 2 | als eine wirkliche arithmetische Zahl hinlänglich fixiert zu haben, wenn sie sagen, sie entspräche der Länge der Diagonale eines Quadrats, denn Seite als der Einheit gleich angenommenen würde. Sie vergessen dabei, daß sie auch hier schon vorausgesetzt haben, daß einem jeden Punkte der graden Linie bei willkürlich gewählter Maßeinheit eine arithmetische 250|251

8. Netto's Deutsche Ausgabe der Gauss'schen Beweise.

Unter „arithmetischen Zahlen“ sind Größen zu verstehen, auf welche die arithmetischen Operationen der Addition, Multiplikation, Division u.s.w anwendbar sind und für welche der Satz gültig bleibt : Ist ein Produkt gleich Null, so muß jedenfalls einer der Faktoren gleich Null sein.

9. Il s'agit en fait des *Göttingische gelehrte Anzeigen* du 23 avril 1831

Zahl entspricht. Und um die Richtigkeit dieser letzteren Voraussetzung aufzuzeigen, ist für sie doch unbedingt wieder jene geometrische Begriffsbestimmung des Irrationalen unentbehrlich. Aber wie kann denn überhaupt ein geometrische Strecke einen arithmetischen Begriff „definieren“? Richtet sich denn der Verstand nach den Sinnen, oder ist es nicht vielmehr der Verstand, welcher durch die Sinnenwelt ordnende Fäden legt, um sie so zu beherrschen ?

Nein, $\sqrt{2}$ wird einzig und allein definiert durch die Gleichung $x^2 - 2 = 0$, und wenn man sagt, $\sqrt{2}$ ist eine arithmetische Zahl, welche zwar durch rationale Zahlen nicht völlig genau, aber darf mit beliebig großer Annäherung bestimmt werden kann, so heißt das nichts weiter, als daß es beliebig kleine Zahlen $\delta$ giebt, für welche die Gleichung

$$x^2 - 2 = \pm\delta$$

durch rationale Zahlen befriedigt wird. Dies führt uns auf unsere Betrachtungen über die Wurzel-Existenz-Beweise zurück. Wenn wir den sogen Fundamentalsatz der Algebra „streng arithmetisch fassen und begründen“ wollen ohne jegliche geometrische Stütze, so können wir behaupten, daß der eigentliche Fundamentalsatz, daß nämlich jede ganze, rationale Funktion $n-$ten Grades in ein Product von $n$ linearen, reellen oder imaginären Faktoren Zerlegbar sei, überhaupt nicht richtig ist. Was wir beweisen können | und beweisen werden und was, im Grunde genommen, auf Gauß nur 251|252
beweisen hat, ist, daß sich aus jeder ganzen rationalen Funktion $n-$ten Grades eine andere durch ein Variirung des Koeffizientensystems der ersteren um beliebig kleine Größen herstellen läßt, welche in ein Produkt von $m$ linear-und $\frac{n-m}{2}$ quadratischen Faktoren zerlegt werden kann und zwar so, daß die Koeffizienten dieser sämtlichen Faktoren rationale Zahlen sind. Dabei ist $m$ eine Anzahl, welche für jede Funktion mit bestimmten Koenffizienten angegeben werden kann [10].

Man erkennt jedoch, daß eine wenn auch noch so sehr benachbarte Funktion, was ihre algebraischen Eigenschaften betrifft, von der ursprünglich vorgelegten im allgemeinen *toto genere* verschieden sein wird, und daß also eine solche Zerlegung für theoretisch-algebraische Untersuchungen keinerlei Nutzen verspricht. Aber dies ist auch nur die eine Seite, die arithmetische Fassung des Fundamentaltheorems. Für algebraische Untersuchungen werden wir die Möglichkeit einer zweiten von der angeführten grundverschiedenen Zerlegung nachweisen, welche nicht im Sinne der Gleichheit, sondern nur als Kongruenz für ein Primmodulsystem besteht. Wir werden nämlich [11] die Richtigkeit der folgenden Kongruenz aufzeigen :

$$F(x) = x^n - c_1x^{n-1} + c_2x^{n-2} - ... \pm c_n \equiv (x - x_1)(x - x_2)...(x - x_n)$$

$$(modd.\ f_1 - c_1, f_2 - c_2, ..., f_n - c_n, y - c),$$

wenn

$$(x - x_1)(x - x_2)...(x - x_n) = x^n - f_1x^{n-1} + f_2x^{n-2} - ... \pm f_n;$$

$x_1...x_n$ unbestimmte Variable, $y$ eine gewisse von der algebraischen Beschaffenheit der Funktion $F(x)$ abhängende, nicht symmetrische | Funktion von $x_1, x_2...x_n$ und $c$ eine 252|253
rationale Zahl ist, die zu $y$ in derselben Beziehung steht, wie die Größen $f_1, f_2, ...f_n$ bezw. zu $c_1, c_2..c_n$. Nur wenn man so den arithmetischen und den algebraischen Teil des Fundamental-Theorems in aller Strenge sondert, ist es möglich, seinen Inhalt wirklich scharf zu fassen.

10. Es sei jedoch gestattet, auch noch bei den folgenden Untersuchungen der Kürze halber die gewöhnliche Ausdrucksweise in betreff der reellen Wurzeln zu verwenden

11. S. 333 ffs

Der Begriff der Charakteristik. 26. Jan.

Wir wollen zunächst den Begriff der Charakteristik unabhängig machen von der Voraussetzung reeller Wurzeln ganzer Funktionen.

Es ist nach unserer früheren Erklärung :

$$1.)\ \chi(V(x),W(x)) = -\frac{1}{2}\sum\ sgn\ \lim_{h=0}\begin{vmatrix} V(x) & W(x) \\ \dfrac{V(x+h)-V(x)}{h} & \dfrac{W(x+h)-W(x)}{h}\end{vmatrix}$$

$$= \frac{1}{2}\sum_{\xi}\ sgn\ V'(\xi)W(\xi),$$

die Summen erstreckt über alle Werte $\xi$, für welche $V(\xi)=0$ wird oder auch die erste Summe über alle Werte $\eta$, für welche $W(x)$ gleich Null wird. Vorausgesetzt wird, daß $V(x)$ und $W(x)$ Funktionen gleichen Grades sind, und daß weder $V(x)$ und $W(x)$, noch $V(x)$ und $V'(x)$, noch $W(x)$ und $W'(x)$ einen gemeinsamen Teiler besitzen. Aus der Gleichung 1.) geht hervor :

$$2.)\ \chi(V(x),W(x)) = -\frac{1}{2}\sum_{(\xi)}\ sgn\ \lim_{h=0}\frac{1}{h}\begin{vmatrix} V(x) & W(x) \\ V(x+h) & W(x+h)\end{vmatrix}$$

$$2^a.)\ = \frac{1}{4}\sum_{(\xi),(\eta)}\ sgn\ \lim_{h=0}\frac{1}{h}\left(W(x).V(x+h)-V(x).W(x+h)\right),$$

wenn die Summe in 2.) erstreckt wird über Werte $\xi$, die in $2^a$.) sowohl über die Werte $\xi$ als auch über die Werte $\eta$.

| Wir können nun die Benutzung des Begriffs der reellen Wurzeln aus diesen letzten Ausdrücken für die Charakteristik ausmerzen, indem wir die Summe in 2.) nicht erstrecken über die Werte $\xi$, für welche $V(\xi)=0$ ist, sondern über alle diejenigen rationalen Werte von $x$, für welche $V(x)$ und $V(x+h)$ entgegengesetzte Zeichen haben, und die Summe in $2^a$) über alle diejenigen rationalen Werte von $x$, für welche sowohl [12] $V(x).V(x+h)<0$ als auch [13] $W(x).W(x+h)<0$ ist. Denn es ist z.B. für alle Werte von $x$, über welche dann in der Summe der Formel 2.) zu summieren ist : 253|254

$sgn\ V(x+h) = sgn\ V'(\xi)$, und da $sgn\ W(x) = sgn\ \lim W(x+h)$, so ist auch :

$$sgn\ \lim\begin{vmatrix} V(x) & W(x) \\ V(x+h) & W(x+h)\end{vmatrix} = sgn\ \lim\left(V(x).W(x+h)-V(x+h).W(x)\right)$$

$$= sgn\ \lim V(x+h).W(x) = sgn\ V'(\xi).W(\xi).$$

Den Grenzübergang in den Ausdrücken 2) und $2^a$) beseitigen wir folgendermaßen.

Wir machen das Intervall $h$ so klein, daß niemals zwei beliebig Werte von $x$, für welche $V(x).V(x+h)<0$ oder $W(x).W(x+h)<0$ ist, in einem Intervalle zusammen sich befinden können. Dann ist :

$$\frac{-\ sgn\ V(x).W(x+h) + \ sgn\ W(x).V(x+h)}{2}$$

$$= sgn\ \left(-V(x).W(x+h)+W(x).V(x+h)\right)$$

12. corrigé dans la page de gauche par *entweder*
13. corrigé dans la page de gauche par *oder*

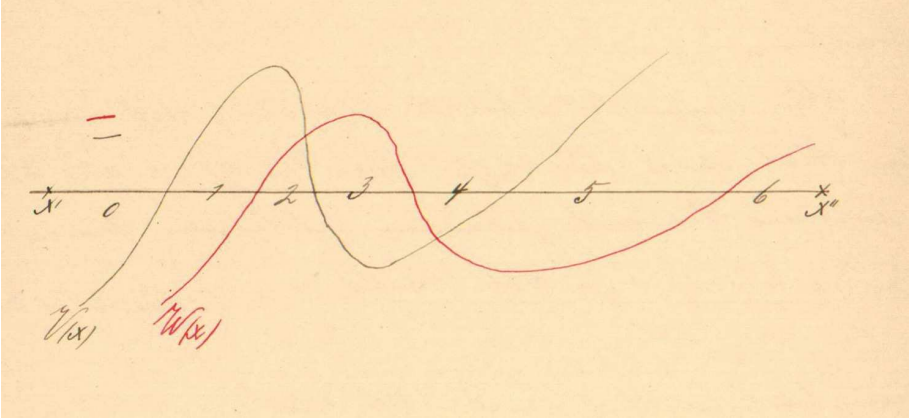
für jeden beliebigen Wert von $x$ [14]. Also wird :

$$3.)\ \chi(V(x),W(x)) = -\frac{1}{8}\sum [\ sgn\ V(x).W(x+h) - \ sgn\ W(x).V(x+h)]\,,$$

wo $h$ die angegeben Größe hat und die Summe erstreckt wird sowohl über alle Werte von $x$, für welche $V(x).V(x+h) < 0$ als auch über alle, für welche $W(x).W(x+h) < 0$
254|255 ist. Diese Gleichung 3.) | können wir, wenn anstatt $x$ $x_k$, anstatt $x+h$ $x_{k+1}$ gesetzt wird auch in der Form

$$4.)\ -8\chi(V(x),W(x)) = \sum_{(k)} \begin{vmatrix} sgn\ V(x_k) & sgn\ V(x_{k+1}) \\ sgn\ W(x_k) & sgn\ W(x_{k+1}) \end{vmatrix}$$

schreiben, indem die Summe wie vorher sowohl über alle Werte von $k$ erstreckt wird, für welche $V(x_k).V(x_{k+1}) < 0$ als auch über diejenigen, für welche $W(x_k).W(x_{k+1}) < 0$ ist. Da aber, wie wir wissen, ein bestimmtes endliches Intervall $x'...x''$ abgegrenzt werden kann, in welchem alle diese Werte $x_k$ und $x_{k+1}$ liegen (d.h. in der gewöhnlichen Sprache : in welchem alle Wurzele von $V(x) = 0$ und $W(x) = 0$ liegen), so können wir auch die Summe in 4.) erstrecken über alle Werte $x_k$, welche dadurch entstehen, daß wir das ganze Intervall $x'...x''$ in kleinere Intervalle von der angegebenen Größe $h$ einteilen. Denn die bei dieser Summierung neu hinzutretenden Determinanten haben sämtlich den Wert Null wegen der Gleichheit ihrer Kolonnen.




So ist z.B. die Charakteristik der beiden Funktionen dritten Grades, welche den nebenstehend gezeichneten Kurven entsprochen, gleich $\frac{-3}{2}$, wie sich aus der beigefügten Tabelle ergiebt :

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $V(x)$ | - | + | + | - | - | + | + |
| $W(x)$ | - | - | + | + | - | - | + |

;

$$-8\chi(V,W) = \begin{vmatrix} -1 & +1 \\ -1 & -1 \end{vmatrix} + \begin{vmatrix} +1 & +1 \\ -1 & +1 \end{vmatrix} + ...$$
$$= +2+2+2+2+2+2$$
$$=12$$

Wir haben jetzt noch zu zeigen, daß sich die Größe $h$ der Intervalle unserer Forderung gemäß so bestimmen läßt, daß niemals zwei Werte $x_k$ in einem und demselben,
255|256 Intervalle | angetroffen werden können, oder mit andern Worten, daß nicht zwei Wurzeln der Gleichung $V(x) = 0$ oder zwei Wurzeln der Gleichung $W(x) = 0$ oder je

14. Ajout dans la page de gauche : „für den $V(x).V(x+h) < 0$ oder $W(x)W(x+h) < 0$

eine Wurzel von $V(x) = 0$ und $W(x) = 0$ in demselben Intervalle liegen. Diese an die Größe von $h$ zu knüpfenden Bedingungen können wir jedoch reduzieren auf die eine, daß nicht die beiden Gleichungen $V(x) = 0$ und $W(x) = 0$ je eine Wurzel in demselben Intervalle besitzen. Denn wenn z.B. $V(x) = 0$ in einen Intervall $a...b$, in welchem eine Wurzel von $W(x) = 0$ nicht vorhanden ist, mehr als eine Wurzel besitzt, so werden sich immer in der Gleichung 4.) die Determinanten, welche zwei auf einander folgenden Wurzeln entsprechen, gegenseitig aufheben. Ist z.B. in einem Intervall $W(x)$ konstant positiv und geht $V(x)$ an einer Stelle desselben von Positiven zum Negativen, alsdann wieder vom Negativen zum Positiven über, so heben sich die entsprechenden Determinanten gegen einander auf :

$$\begin{vmatrix} + & - \\ + & + \end{vmatrix} + \begin{vmatrix} - & + \\ + & + \end{vmatrix} = 0$$

Wenn die Anzahl der Wurzeln von $V(x) = 0$ ungrade ist, so bliebt uns eine einzige Determinante übrig, welche, wie leicht zu durchschauen ist, denselben Wert hat wie die Determinante

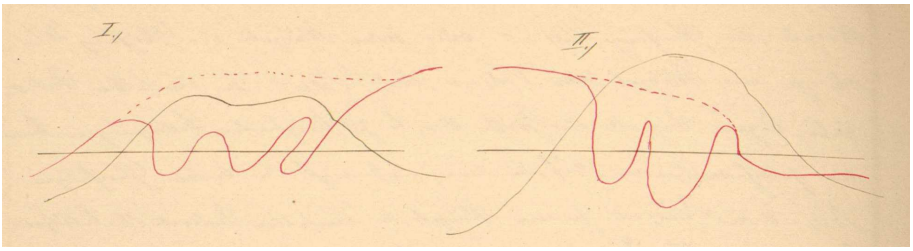
$$5.) \quad \begin{vmatrix} sgn\ V(a) & sgn\ V(b) \\ sgn\ W(a) & sgn\ W(b) \end{vmatrix} ;$$

ist die Anzahl grade, so ist die Summe sämtlichen Determinanten Null und die Determinante 5.) hat denselben Wert, da dann

$$sgn\ V(a) = \ sgn\ V(b)\ ;\ sgn\ W(a) = \ sgn\ W(b)$$

ist. Man kann daher in jedem der beiden Falle die ganze Anzahl von Intervallen durch das eine Intervall $a...b$ ersetzen, ohne der Richtigkeit | der Formel 4.) Eintrag 256|257 zu thun und z.B. in den Figuren $I$) und $II$) den Lauf der ursprünglichen Kurve durch die punktiert gezeichnete ersetzen.

Nunmehr wollen wir die geforderte Größe des Intervalls bestimmen. Es ist ein wesentliches Erfordernis, daß wir ins unserer Deduktion die Voraussetzung aufbrauchen, daß nicht beide Kurven $V(x) = 0$ und $W(x) = 0$ in einem und demselben Punkte durch die Axe gehen, - denn ohne diese Voraussetzung ist es unmöglich, die Bedingung zu erfüllen, welche wir an die Größe des Intervalls geknüpft haben.



Nach S. 204 Gleichung 3.) ist die Resultante :

$$6.)\ R(V, W) = P(x).W(x) + Q(x).V(x)$$

und zwar ist für unsere Fall also $R(V, W) \neq 0$. Wenn $V(x)$ und $W(x)$ bezw. vom Grade $n$ und $n'$ sind (denn für die folgenden Schlußfolgerungen ist es nicht nötig, die Einschränkung, daß $V(x)$ und $W(x)$ vom gleichen Grade sind, beizubehalten), so können wir $P(x)$ und $Q(x)$ in der Form schreiben :

$$P(x) = p_0 + p_1x + p_2x^2 + ... + p_nx^{n-1}$$
$$Q(x) = q_0 + q_1x + q_2x^2 + ... + q_{n'}x^{n'-1}$$

ist und $= 0$ für alle anderen Werte von $x$.“

Geben nun die rationalen Zahlen $-r$ und $+r$ die Endpunkte des Intervalls an, innerhalb dessen alle Wurzele von $V(x) = 0$ und $W(x) = 0$ liegen, so ist für jeden Wert von $x$ welcher diesem Intervall angehört [15] :

$$\begin{aligned}|P(x)| &< |p_0| + |p_1|.r + |p_2|.r^2 + ... + |p_{n-1}|.r^{n-1} \\ &< |p_k|.(1 + r + r^2 + ... + r^{n-1}) \\ &< p_k.\frac{r^n - 1}{r - 1},\end{aligned}$$

wenn $|p_k| > |p_i|$ $(i = 0, 1, 2, ...k - 1; k + 1...n - 1)$.

257|258 | Und auch :

$$\frac{P(x)}{|R|} = (s_1 - 1)\varphi(x),$$

wo $s_1$ eine bestimmte ganze Zahl ist und $\varphi(x)$ eine Funktion, welche für jeden dieser Werte $x$ kleiner als 1 ist ; denn der Quotient auf der linken Seite bleibt ebenfalls unterhalb einer festen Grenze.

Ferner können wir setzen :

$$Q(x) = (s_2 - 1)|R|.\psi(x),$$

wo $s_2$ ebenfalls eine bestimmte ganze Zahl und $\psi(x)$ für alle in Betracht kommenden Werte von $x$ kleiner als 1 ist, und drittens :

$$P(x).\frac{W(x + \sigma) - W(x)}{\sigma} = (s - 1)|R|.\chi(x),$$

da der Ausdruck auf der linken Seite eine ganze Funktion wie $x$ und $\sigma$ ist und also für Werte von $\sigma$, für welche $|\sigma| < 1$ ist, auch eine feste obere Grenze haben müß.

Wenn um $s$ eine endliche ganze Zahl ist, welche $s_1, s_2$ und $s_3$ an Größe übertrifft, so können wir setzen :

$$\begin{aligned}&7.)\ P(x) = (s - 1)|R|.\varphi(x) \\ &8.)\ Q(x) = (s - 1)|R|.\psi(x) \\ &7.)\ P(x)\frac{W(x + \sigma) - W(x)}{\sigma} = (s - 1)|R|.\chi(x),\end{aligned}$$

und die Funktionen $\varphi(x), \psi(x), \chi(x)$ haben die Eigenschaft, daß sie für alle Werte von $x$, welche in dem Intervall $-r.... + r$ liegen, einen Wert haben, welcher kleiner ist als 1. Aus der Gleichung 6.) folgt [16] :

$$10.)\ R(V, W) = P(x_1).(W(x_1) - W(x_2)) + P(x_1).W(x_2) + Q(x_1)V(x_1)$$

258|259 und unter der Voraussetzung, daß $|x_2 - x_1| < 1$ ist, aus 9.) : |

$$11.)\ P(x_1).\frac{W(x_2) - W(x_1)}{x_2 - x_1} = (s - 1).|R|.\chi(x_1).$$

15. Correction dans la page de gauche :

$$\begin{aligned}|P(x)| &< |p_0| + |p_1|.r + |p_2|.r^2 + ... + |p_{n-1}|.r^{n-1} \\ &\leq |p_k|.(1 + r + r^2 + ... + r^{n-1}) \\ &= |p_k|.\frac{r^n - 1}{r - 1},\end{aligned}$$

wenn $|p_k| \geq |p_i|$ $(i = 0, 1, 2, ...k - 1; k + 1...n - 1)$.

16. Ajout dans la page de gauche : „da die linke Seite von $x$ unabhängig ist,“

Mit Benutzung der Gleichungen 7.), 8.) und 11.) ergiebt sich aus 10.) :

12.) $|R| = |(s-1).|R|.(x_1-x_2).\chi(x_1)+(s-1).|R|.\varphi(x_1).W(x_2)+(s-1).|R|.\psi(x_1).V(x_1)|$

und nach Division beiden Seiten durch $(s-1).|R|$ :

$$\frac{1}{s-1} = |(x_1-x_2)\chi(x_1)+\varphi(x_1).W(x_2)+\psi(x_1).V(x_1)|$$

also wegen der Eigenschaft der Funktionen $\varphi(x), \psi(x)$ und $\chi(x)$ :

$$\frac{1}{s-1} < |x_1-x_2|+|V(x_1)|+|W(x_2)|$$

Es sei ferner $|x_2-x_1| \leq \frac{1}{s}$ gemacht, dann folgt durch Einsetzen dieses Wertes in die letzte Gleichung :

$$13.)\ -\frac{1}{s}+\frac{1}{s-1} = \frac{1}{s(s-1)} < |V(x_1)|+|W(x_2)|$$

Damit haben wir unsere Aufgabe gelöst. Denn machen wir $h \leq \frac{1}{s}$, so kann sicher niemals eine Wurzel von $V(x) = 0$ und eine solche von $W(x) = 0$ in demselben Intervalle liegen, da wenn $V(x_1)$ und $W(x_2)$ die Funktionswerte für zwei Stellen $x_1$ und $x_2$ des Intervalls sind, die Summe ihrer absoluten Werte stets ein bestimmte endliche Größe übersteigen muß (Gl. 13.)).

Nach Ausführung dieser Bestimmung können wir sagen, daß wir in der Formel 4.) einen Ausdruck für die Charakteristik erlangt haben, welcher sich streng arithmetisch, ohne Benutzung der reellen Wurzele bestimmen läßt und uns der Berechnung der Charakteristik durch den von *Sylvester* verallgemeinerten Sturm'schen Satz enthebt.

Teilen wir insbesondere das allein in Betracht kommende Intervall $-r...+r$ in $2s.r$ gleich Seite von der Größe $\frac{1}{s}$ und | bezeichnen die Teilpunkte mit 259|260

$$x_0 = -r\ ,\ x_1 = -r+\frac{1}{s}\ ,\ x_2 = -r+\frac{2}{s}\ ,\ ...\ ,\ x_{2s.r-1} = -r+\frac{2s.r-1}{s}\ ,\ x_{2s.r} = +r,$$

so wird :

$$14.)\ -8\chi(V(x),W(x)) = \sum_{k=0}^{k=2s.r} \begin{vmatrix} sgn\ V(-r+\frac{k}{s}) & sgn\ V(-r+\frac{k+1}{s}) \\ sgn\ W(-r+\frac{k}{s}) & sgn\ W(-r+\frac{k+1}{s}) \end{vmatrix}$$

Für diese Darstellung sei jedoch noch Folgendes bemerkt. Es kann bei der so bewirkten Einteilung des Intervalls vorkommen, daß wir zufällig mit einem Teilpunkt auf eine rationale Wurzel der Gleichung $V(x) = 0$ oder $W(x) = 0$ stoßen, sodaß etwa $V(-r+\frac{k}{s}) = 0$ ist : Es scheint, als ob in einem solchen Falle die Formel 14.) ihre Gültigkeit verlieren würde, da nach der Herleitung derselben (s. Formel 4.)) für alle Werte von $k$ :

$$sgn\ V(-r+\frac{k}{s}).V(-r+\frac{k+1}{s} = \pm 1$$

sein soll. In Wirklichkeit aber bleibt auch für dieser Fall die Richtigkeit der Formel gewahrt, denn, wenn $V(-r+\frac{h}{s}) = 0$ ist, so braucht man nur die beiden Intervalle $-r+\frac{h-1}{s}....-r+\frac{h}{s}$ und $-r+\frac{h}{s}....-r+\frac{h+1}{s}$ sich zu dem einen Intervall $-r+\frac{h-1}{s}....-r+\frac{h+1}{s}$ vereinigt zu denken, um die Formel 14.) wieder mit der Gleichung 4.) in Einklang zu bringen. Nach der Bedingung der Gleichung 13.) muß nämlich die Entfernung einer Wurzel von $V(x) = 0$ von einen Wurzel von $W(x) = 0$ größer sein als $\frac{1}{s}$ und daher kann sich in unseren Doppelintervall außer der Wurzel $(-r+\frac{h}{s})$ von

$V(x) = 0$ nicht noch eine von $W(x) = 0$ vorfinden, selbst nicht auf den Grenzen des Intervalls.

Wir wollen nunmehr diese Untersuchungen anwenden zur Bestimmung der bereits
260|261 oben S. 213 berechneten Differenz zweier Charakteristiken : |

$$15.)\ \chi(f(x),(x''-x)f_1(x)) - \chi(f(x),(x'-x).f_1(x)) = \sum_{x'<\xi<x''} sgn\ f_1(\xi).f'(\xi),$$

aber jetzt ohne Anwendung reeller Wurzele, indem wir für die folgenden Untersuchungen voraussetzen, was sich ja stets erreichen läßt, daß $f_1(x)$ eine ganze Funktion ist, deren Grad um 1 niedriger ist als der von $f(x)$.

Zunächst haben wir die Identität abzuleiten :

$$\begin{aligned} 16.)\ &\frac{1}{2}\sum_{(\xi)} sgn\ f_1(\xi).f'(\xi) + \frac{1}{2}\sum_{(\eta)} sgn\ f(\eta).f_1'(\eta) \\ &- \frac{1}{4}\ sgn\ f(x'').f_1(x'') + \frac{1}{4}\ sgn\ f(x').f_1(x') = 0 \end{aligned}$$

wo $\xi$ die wurzele von $f(x) = 0$, $y$ diejenigen von $f_1(x) = 0$ [17] bezeichnen mögen. Es sei :

$$\begin{vmatrix} f(x) & (x''-x).f_1(x) \\ f'(x) & (x''-x).f_1'(x) - f_1(x) \end{vmatrix} = \Delta(x)$$

gesetzt, so ist :

$$17.)\ -\frac{1}{2}\sum\ sgn\ \Delta(\xi) = \frac{1}{2}\ sgn\ (x''-\xi).f_1(\xi).f'(\xi)$$

und

$$18.)\ -\frac{1}{2}\sum\ sgn\ \Delta(\eta) = -\frac{1}{2}\ sgn\ (x''-\eta).f(\eta).f_1'(\xi) + \frac{1}{2}\ sgn\ f(x'').f_1(x'')$$

und ferner sei :

$$\begin{vmatrix} f(x) & (x'-x).f_1(x) \\ f'(x) & (x'-x).f_1'(x) - f_1(x) \end{vmatrix} = \Delta_1(x)$$

gesetzt, so wird :

$$19.)\ -\frac{1}{2}\sum\ sgn\ \Delta_1(\xi) = \frac{1}{2}\ sgn\ (x'-\xi).f_1(\xi).f'(\xi)$$

und

$$20.)\ -\frac{1}{2}\sum\ sgn\ \Delta_1(\eta) = -\frac{1}{2}\ sgn\ (x'-\eta).f(\eta).f_1'(\xi) + \frac{1}{2}\ sgn\ f(x').f_1(x')$$

da nun $x' - \xi \neq -(x'' - \xi) < 0$, [18] so ist :

$$\begin{aligned} \frac{1}{2}(17.) - 19.) - 18.) + 20.)) &= \frac{1}{2}\sum\ sgn\ f_1(\xi).f'(\xi) + \frac{1}{2}\sum\ sgn\ f(\eta).f_1'(\eta) \\ &= -\frac{1}{4}\ sgn\ f(x'').f_1(x'') + \frac{1}{4}\ sgn\ f(x').f_1(x') = 0 \end{aligned}$$

17. Ajout dans la page de gauche : „zwischen x' und x''“
18. Ajout dans la page de gauche : „und desgl. $x' - \eta$ und $-(x' - x) < 0$“

Denken wir uns nun für die Funktionen $f(x)$ und $f_1(x)$ die ganze Zahl $s$ berechnet und darauf für das Intervall $x'..x''$ ebenso wie | in Formel 14.) für das Intervall $-r...+r$ 261|262 die Teilpunkte $x_0, x_1...$ festgelegt, alsdann die algebraische Zeichensumme

$$21.) \quad -\frac{1}{4}\sum_{(k)}\begin{vmatrix} sgn\ f(x_k) & sgn\ f(x_{k+1}) \\ sgn\ f_1(x_k) & sgn\ f_1(x_{k+1}) \end{vmatrix} + \frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}\ sgn\ f(x').f_1(x')$$

gebildet, so können wir deren Wert auch in der Form

$$22.) \quad \frac{1}{2}\sum_{x'<\xi<x''} sgn\ f_1(\xi).f'(\xi) + \frac{1}{2}\sum_{x'<\eta<x''} sgn\ f(\eta).f_1'(\eta) + \frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}\ sgn\ f(x').f_1(x')$$

schreiben, wie sich nach den früheren Erörterungen (S. 253 4)) leicht ergibt. Dieser letztere Ausdruck ist aber nichts anderes als $\sum_{x'<\xi<x''} sgn\ f_1(\xi)f'(\xi)$, wie sofort erhellt, wenn man zu 22.) die Identität 16.) addirt. Daher ist 21.) der gewünschte Wert der Charakteristik-Differenz 15.)

Die Zeichensumme [19]

$$23.) \quad -\frac{1}{2}\sum\begin{vmatrix} sgn\ f(x_k) & sgn\ f(x_{k+1}) \\ sgn\ f_1(x_k) & sgn\ f_1(x_{k+1}) \end{vmatrix} + \frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}f(x').f_1(x'),$$

wenn die summe nur erstreckt wird über diejenigen Intervalle $x_k...x_{k+1}$, für welche $f(x_k).f(x_{k+1}) < 0$ ist, während das Zeichen von $f_1(x)$ für die Endpunkte des Intervalls und alle Punkte innerhalb desselben konstant bleibt, ist ebenfalls dem Ausdruck 22.) gleich, da in demselben :

$$\frac{1}{2}\sum_{x'<\xi<x''} sgn\ f_1(\xi).f'(\xi) = -\frac{1}{2}\sum_{x'<\eta<x''} sgn\ f_1(\eta).f'(\eta).$$

Diese Gleichheit können wir in der Form ausdrücken [20] :

$$24.) \quad \chi(f(x),(x''-x).f_1(x)) - \chi(f(x),(x'-x)f_1(x)) = \frac{1}{2}\sum_{f(x_k).f(x_{k+1})<0} sgn\ f_1(x_k).(\ sgn\ f(x_{k+1}) - \ sgn\ f(x_k))$$

19. Correction dans la page de gauche :

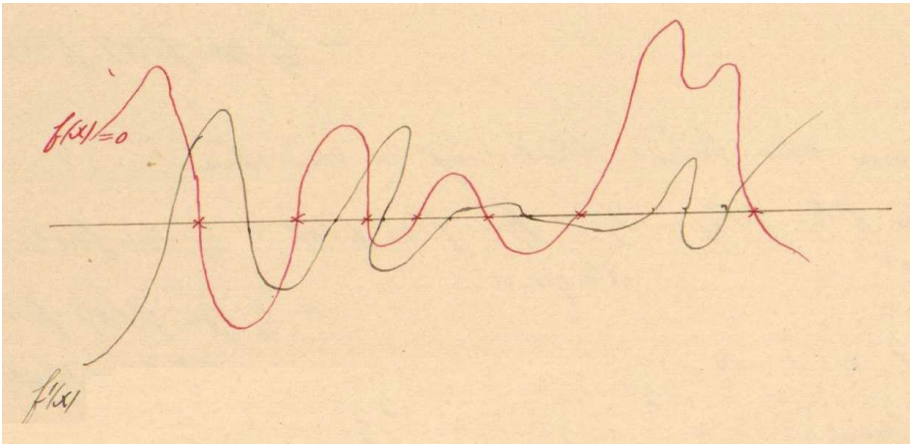
$$23.) \quad -\frac{1}{4}\sum\begin{vmatrix} sgn\ f(x_k) & sgn\ f(x_{k+1}) \\ sgn\ f_1(x_k) & sgn\ f_1(x_{k+1}) \end{vmatrix} + \frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}f(x').f_1(x')$$

20. Correction dans la page de gauche :

$$24.) \quad \chi(f(x),(x''-x).f_1(x)) - \chi(f(x),(x'-x)f_1(x)) = -\frac{1}{4}\sum_{f(x_k).f(x_{k+1})<0} sgn\ f_1(x_k).(\ sgn\ f(x_{k+1}) - \ sgn\ f(x_k)) + \frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}\ sgn\ f(x')f_1(x').$$

$$+\frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}\ sgn\ f(x')f_1(x').$$

262|263 | Für den speziellen Fall $f_1(x) = f'(x)$ erhalten wir in 23.) und 24.) diejenige Zahl, welche angiebt, wieviel „reelle Wurzeln“ zwischen $x'$ und $x''$ liegen.




Beim Zeichnen der Figur ist darauf zu achten, dass
1) $f_1$ auch eine parabolische Curve ist
2) Die Nullstellen von $f'$ unter den Max und Min von $f$
3) Die Max und Min von $f'$ unter den Wendepunkten von $f$ liegen.

Diese Anzahl läßt sich aber noch auf eine einfachere Art ermitteln. Bestimmt man nämlich für das Funktionensystem $(f(x), f'(x))$ die Zahl $s$ und teilt darauf das Intervall $x'...x''$ in der bekannten Art in kleinere Intervalle, so wird für das besondere kurvensystem $f(x) = 0, f'(x) = 0$ der eigentümliche Umstand eintreten, daß in keinem Intervall die Kurve $f(x) = 0$ mehr als ein Mal die Absissen-axe schneidet, da zwischen zwei solchen Schnittpunkten, stets mindestens ein Schnittpunkt der kurve $f'(x) = 0$ liegen muß, vorausgesetzt, daß die Diskriminante der beide Gleichungen $f(x) = 0$ und $f'(x) = 0$ nicht verschwindet. Daher können wir die Anzahl der reellen Wurzeln, welche in dem Intervall $x'...x''$ liegen, einfach der Zeichensumme

$$\frac{1}{2}\sum_{(k)} |\ sgn\ f(x_k) -\ sgn\ f(x_{k+1}|$$

gleichsetzen, wo sich die Summe über sämtliche Werte von $x_k$ erstreckt, welche Teilpunkte jenes Intervalls sind.

29. Jan. Die Natur der reellen Wurzeln und die Zerlegung ganzer Funktionen.

Wir wollen nun unserer ganzen Zahl $s$ außer den Bedingungen, welchen wir sie bereits unterworfen hatten (S. 258 Gl. 7.) 8.) 9.) ) für unsere besonderen Fall des Funktionensystems $(f(x), f'(x))$ mit ganzzahligen Koeffizienten noch die eine auferlegen,
263|264 daß : |

$$\left|\frac{f(x_2) - f(x_1)}{x_2 - x_1}\right| < (s-1).|D|,$$

wenn $x_2 - x_1 < \frac{1}{s}$ ; $x_2 > x_1$ ist, oder :

$$1.)\ |f(x_2) - f(x_1)| < (x_2 - x_1)(s-1).|D|.$$

Teilen wir jetzt jedes der Intervalle von der Größe $\frac{1}{s}$ wieder in $r.|D|$ gleiche Teile, ($D$ ist eine ganze Zahl), so wird, wenn $x'$ und $x''$ rationale Werte sind, welche einem und

demselben dieser Unterintervalle angehören,

$$x'' - x' \leq \frac{1}{r.s.|D|}$$

sein, unter der Annahme, daß $x'' > x'$ ist ; und nach Gleichung 1.) :

$$2.)\ |f(x'') - f(x')| < \frac{1}{r},$$

wo $r$ eine völlig beliebige ganze Zahl ist.

Greifen wir unter den ursprünglichen Intervallen von der Größe $\frac{1}{s}$ eines von der Beschaffenheit heraus, daß die Funktion an beiden Endpunkten desselben entgegengesetzte Zeichen hat, so wird es auch unter seinen Teilintervallen eines von derselben Beschaffenheit geben. Sind $x'$ und $x''$ die rationalen Endwerte dieses Intervalls, so muß

$$|f(x')| < \frac{1}{r} \text{ und } |f(x'')| < \frac{1}{r}$$

sein, da nach 2.) sogar die Summe dieser absoluten Werte kleiner als $\frac{1}{r}$ ist.

Dieselbe Ungleichung gilt für alle rationalen Werte, welche innerhalb des Intervalls liegen. Auf diese Weise erhalten wir also anstatt einer reellen Wurzel $\xi$ von $f(x) = 0$ ein Intervall $x'...x''$ von der Art, daß

$$a.)\ f(x').f(x'') < 0 \text{ und } b.)\ |f(x' + (x'' - x')\delta)| < \frac{1}{r},$$

wo $0 \leq \delta \leq 1$ und $x', x''$ bestimmbare, rationale Zahlen sind.

Wieviel solcher Intervalle es giebt, können wir nach den Angaben des vorhergehenden Abschnitts ermitteln. Ist ihre Anzahl $v$, so mögen die Intervalle mit | 264|265

$$\mathfrak{x}'_1\mathfrak{x}''_1, \mathfrak{x}'_2\mathfrak{x}''_2 ... \mathfrak{x}'_v\mathfrak{x}''_v$$

bezeichnet sein und

$$\mathfrak{x}_1, \mathfrak{x}_2, ...\mathfrak{x}_v$$

rationale Werte, welche in den einzelnen Intervallen liegen. Dann können wir setzen :

$$3.)\ f(x) = (x - \mathfrak{x}_1)(x - \mathfrak{x}_2)...(x - \mathfrak{x}_v)Q(x) + R(x),$$

wo der Rest $R(x)$ in allgemeinen von Null verschieden ist. Jedoch läßt sich zeigen, daß seine Koeffizienten bei passender Bestimmung der unbestimmt gebliebenen Zahl $r$ beliebig klein gemacht werden können.

Wir benutzen zu diesem Zwecke die Lagrange'sche Interpolationsformel, nach welcher :

$$4.)\ R(x) = \sum_{k=1}^{v} \frac{f(\mathfrak{x}_k)}{\varphi'(\mathfrak{x}_k)} \frac{\varphi x}{x - \mathfrak{x}_k},$$

da $f(\mathfrak{x}_k) = R(\mathfrak{x}_k)$ für $k = 1, 2...v$. $\varphi'(\mathfrak{x}_k)$ ist das Produkt von $(v-1)$ Differenzen $(\mathfrak{x}_k - \mathfrak{x}_i)$. In zwei verschiedene $\mathfrak{x}_i, \mathfrak{x}_k$ sind aber um mindestens $\frac{1}{s}$ von einander getrennt. Dann, da eine Wurzel von $f(x) = 0$ von der nächstbenachbarten Wurzel von $f'(x) = 0$ stets um mehr als $\frac{1}{s}$ entfernt liegt, (S.259 Gl. 13.)), zwischen zwei Wurzele von $f(x) = 0$ aber immer eine solche von $f'(x) = 0$ liegt [21], so muß der Abstand zwischen zwei Wurzeln

21. Commentaire dans la page de gauche : Es ist zu beachten, dass hier das Rollesche Theorem angewandt wird

von $f(x) = 0$ jedenfalls mehr als $\frac{2}{s}$ betragen. Die Entfernung zwischen den beiden diesen Wurzeln zunächst liegenden rationalen Werte $\mathfrak{x}_i, \mathfrak{x}_k$ kann jedoch um höchstens $\frac{2}{r.s.|D|}$ geringer sein, da jedes $\mathfrak{x}_h$ von der zugehöriger Wurzel von $f(x) = 0$ um nicht mehr als $\frac{1}{r.s.|D|}$ entfernt sein kann. Immerhin ist also :

$$\begin{aligned} |\mathfrak{x}_i - \mathfrak{x}_k| &> \frac{2}{s} - \frac{2}{r.s.|D|} \\ &> \frac{1}{s}, \end{aligned}$$

wenn nur $r.|D| \geq 2$ ist und [22] :

$$|\prod(\mathfrak{x}_i - \mathfrak{x}_k)| > \frac{1}{s^{v-1}} \ , \ \frac{1}{|\prod(\mathfrak{x}_i - \mathfrak{x}_k)|} > s^{v-1}.$$

265|266 Da ferner $|\ |f(\mathfrak{x}_k)| < \frac{1}{r}$ ist, so ist der Koeffizient eines jeden Gliedes der Summe in Gl. 4.) :

$$5.) \quad \left|\frac{f(\mathfrak{x}_k)}{\varphi'(\mathfrak{x}_k)}\right| < \frac{s^{v-1}}{r}.$$

Betrachten wir ferner die noch vorkommenden ganzen Funktionen $(v-1)$ten Grades $\frac{\varphi(x)}{x-\mathfrak{x}_k}$, so werden deren Koeffizienten, da $|\mathfrak{x}_k| < r$ $(k = 1, 2, ...v)$, wenn $r$ die bekannte Bedeutung hat, der Reihe nach kleiner als :

$$(v-1).r \ , \ \frac{(v-1)(v-2)}{2}.r^2 ... \ , \ r^{v-1}.$$

Ist $t$ eine Zahl, welche größer ist als jeder dieser Werte, so wird schließlich jeder beliebige Koeffizient der ganzen Funktion $(v-1)$ten Grades $R(x)$ dem absoluten Werte nach kleiner sein als $\frac{v.s^{v-1}.t}{r}$. Also ist :

$$6.) \ R(x) = \sum_{k=1}^{v} \frac{f(\mathfrak{x}_k)}{\varphi'(\mathfrak{x}_k)}.\frac{\varphi(x)}{x-\mathfrak{x}_k} = \frac{v.s^{v-1}.t}{r}.\Theta(x)$$

und

$$7.) \ \prod(x - \mathfrak{x}_k).Q(x) + \frac{v.s^{v-1}.t}{r}.\Theta(x),$$

wo $\Theta(x)$ eine ganze Funktion von höchstens $(v-1)$ten Grade ist, deren Koeffizienten sämtlich kleiner als 1 sind. $s, v, t$ sind ganze Zahlen, welche aus den Koeffizienten der vorgelegten Funktion $f(x)$ fest bestimmbar sind, $r$ dagegen ist beliebig und kann so groß gemacht werden, daß die Koeffizienten der Restfunktion, welche wir nunmehr wegen dieser Abhängigkeit von $r$ mit $f_r(x)$ bezeichnen wollen, sämtlich kleiner sind als ein beliebig vorgeschriebene noch so kleine Größe. Das Resultat, welches wir also erhalten haben, ist das folgende :

Es ist im allgemeinen nicht möglich, eine vorgelegte ganze Funktion $f(x)$ vom $n$-ten Grade selbst in ein Produkt von linearen Faktoren zu zerlegen, deren Koeffizienten
266|267 einen | wirklich genauen und erschöpfenden arithmetischen Deutung fähig wären. Dagegen sind wir imstande, wenn wir den Wert von $2\chi(f, f')$ mit $v$ bezeichnen, die Differenz der beiden Funktionen $f(x) - f_r(x)$, wo $f_r(x)$ eine wohl bestimmbar ganze Funktion von höchstens $(v-1)$-ten Grade mit beliebig kleinen Koeffizienten ist, in ein Produkt zu zerlegen, welches aus einer ganzen Funktion $Q(x)$ vom Grade $(n - v)$ mit rationalen Koeffizienten und $v$ rationalen, linearen Faktoren besteht, sodaß :

22. Correction dans la page de gauche :

$$|\prod(\mathfrak{x}_i - \mathfrak{x}_k)| > \frac{1}{s^{v-1}} \ , \ \frac{1}{|\prod(\mathfrak{x}_i - \mathfrak{x}_k)|} < s^{v-1}.$$

8.) $f(x) - f_r(x) = \prod_{k=1}^{v}(x - \mathfrak{x}_k)Q(x)$ ist.

Die Zerlegung von $f(x) - f_r(x)$ können wir, wenn wir uns $f(x)$ von sämtlichen rationalen Wurzeln befreit denken, unter ausschließlicher Benutzung der Endwerte $x_k, x_{k+1}$ der Intervalle von der Größe $\frac{1}{s}$, in welche wir das Intervall $-r ... + r$ geteilt haben, in folgender Weise darstellen :

$$f(x) - f_r(x) = \prod_{(k)} \left\{(x - x_k - 1).\frac{1}{2}\ sgn\ f'(x_k).[\ sgn\ f(x_{k+1} -\ sgn\ f(x_k)] + 1\right\},$$

wo sich das Produkt über sämtliche Werte von $k$ erstreckt. Denn, wenn $x_{k+1} > x_k$ ist und $f(x) = y$ in dem Intervall $(x_k, x_{k+1})$ sein Zeichen wechselt, so muß $sgn\ f'(x_k) = sgn\ f(x_{k+1})$ sein, also ist der Ausdruck $\frac{1}{2}\ sgn\ f'(x_k)[\ sgn\ f(x_{k+1}) -\ sgn\ f(x_k)]$ gleich $+1$ oder 0, je nachdem das Intervall $(x_{k+1}, x_k)$ ein „reelle Wurzel“ einschließt oder nicht. Ein andere Art der Darstellung ist :

$$f'(x) - f_r(x) = \prod_{(k)}(x - \delta_k.x_k - (1 - \delta_k)x_{k+1})^{\frac{1}{2}\ sgn\ f'(x_k).[\ sgn\ f(x_{k+1}) - sgn\ f(x_k)]}.Q(x),$$

wo $\delta_k$ (für $k = 1, 2, ...$) beliebige positive echte Brüche sind, nach deren Fixierung auch $f_r(x)$ eine bestimmte Funktion ist.

Den Inhalt der Gl. 8.) können wir auch folgendermaßen | aussprechen : Zu jeder 267|268
beliebigen ganzen Funktion $f(x)$, für welche $2\chi(f, f') = v$ ist, läßt sich eine beliebig benachbarte Funktion ermitteln, welche $v$ rationale Wurzeln hat.

Für den besonderen Fall, in welchem $v = n$ ist, sind wir hier am Ziele, da dann $Q(x) = 1$ wird. Ist aber $v < n$, so können wir die Restfunktion $f_r(x)$ so abändern, daß die Differenz $f(x) - f_r(x)$ sich als Produkt von $v$ linearen und $\frac{n-v}{2}$ quadratischen Faktoren mit rationalen Koeffizienten darstellen läßt. Haben wir dies gezeigt und eine Methode zur Bestimmung auch der Koeffizienten der quadratischen Faktoren angegeben, so ist der Inhalt des Fundamentaltheorems in seiner rein arithmetischen Fassung und Begründung erschöpfend zur Darstellung gebracht und damit für diesen ersten Teil des wichtigsten, ja, man könnte fast behaupten, einzigen Theorems der Lehre von den algebraischen Gleichungen „der Weisheit letzter Schluß“ gezogen.

Es ist zu beachten daß für andere als ganze rationale Funktionen diese Art der Zerlegung nur dann möglich ist, wenn man für dieselben die Isolierungs-Intervalle von bestimmbaren Größe ermitteln kann. Daß dafür die Bedingung der Stetigkeit nicht hinreichend ist, lehren einige von *Riemann* und *Weierstrass* angegeben Funktionen, z.B. $\sum_{n=1}^{\infty} \frac{\sin n^2 x\pi}{n^2}$, welche zwar stetig sind, aber die Bestimmung der Isolierungsintervalle ebenso wenig zulassen wie diejenige eines Differentialquotienten.

## VII.) Über eine Eigenschaft der Gebiete, in welchen die Funktionen mit gleicher Charakteristik liegen.

Wenn wir uns jedem Punkte einer $v-$fachen Mannigf. ein | bestimmtes Funk- 268|269
tionensystem $(f(x), f_1(x))$ in einer wiederholt angegebenen Weise zugeordnet denken, so hatten wir gesehen, daß wir, um von einem System zu einem andern zu gelangen, dessen Charakteristik von der des ersteren verschieden ist, unbedingt die ResultantenMannigf. überschreiten mußten. Es läßt sich aber auch der Beweis dafür erbringen, daß man stets ohne Durchgang durch die ResultantenMannigf. von einem Funktionsystem zu einem andern gelangen kann, dessen Charakteristik derjenigen des ersteren gleich ist. Diesen Beweis haben wir zwar für den speziellen Fall, wo $f_1(x) = f'(x)$ ist, bereits auf S. 222 erbracht, jedoch damals unter der Voraussetzung der Zerlegung einer ganzen Funktion $n-$ten Grades in $n$ reelle oder complexe

Linearfaktoren, sodaß wir nach unsere letzten Untersuchungen demselben völlige arithmetische Strenge nicht mehr zusprechen können. An dieser Stelle wollen wir seine Richtigkeit für den allgemeineren Fall ohne Benutzung jener Fiktion erhärten.

Dazu schlagen wir folgendem Weg ein. Wir zeigen, daß wir von dem vorgelegten, durchaus beliebigen Funktionensystem [23]

$$1.)\ (f(x), f_1(x))$$

stets zu dem System :

$$2.)\ \begin{cases} \varphi(x) = (x-1)(x-3)(x-5)...(x-2\mu+1)(x^2+1^2)(x^2+3^2)...(x^2+(n-\mu-1)^2) \\ \psi(x) = (x-2)(x-4)(x-6)...(x-2\mu)(x^2+2^2)(x^2+4^2)...(x^2+(n-\mu)^2) \end{cases}$$

gelangen können, (in welchem einzig und allein $\mu$ von der Charakteristik des Systems 1.) abhängig ist), ohne die Resultanten-Mannigf. zu überschreiten. Denn dann sind wir auch in der Lage, von jedem Funktionensystem zu jedem andern, dessen Charak-
269|270 teristik denselben Wert hat, zu gelangen, indem wir stets unsere Weg über jenen | einen festen Punkt des Gebiets nehmen.

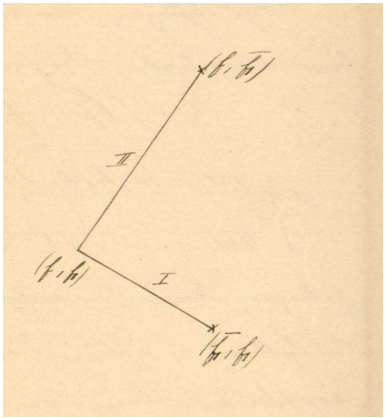
Die Überführung des Funktionensystems 1.) in das System 2.) nehmen wir Schritt für Schritt vor. Da $\chi(f, f_1) = \chi(f, t.f + f_1)$ für alle Werte von $t$ ist, so können wir durch stetige Variirung des Parameters $t$ von 0...1 zu einem System gelangen, dessen Funktionen von gleichen Grade sind. Wir verändern weiter das Funktionensystem, welches wir wieder $(f, f_1)$ nennen wollen, so, daß

$$\chi(f, f') = \chi(f_1, f_1')$$

ist, d.h. daß die Anzahl der reellen Wurzeln einer jeden der beiden Funktionen dieselbe ist. Das kann auch folgende Weise geschehen. Wir gehen von dem System $(f, f_1)$ zu dem benachbarten $(\overline{f_1}, f_1)$ wo $\overline{f_1} = (1-t)f + f.f_1$ und $t = \tau$ ist, $\tau$ aber so klein, daß auf dem Wege von $t = 0$ nach $t = \tau$ die Disktiminanten-Mannigf. $R(f, f') = 0$ nicht überschritten wird, d.h. daß also $\chi(f, f') = \chi(f_1, f_1')$ ist [24]. Dies ist immer zu erwirken, da sich nach Voraussetzung der Punkt $(f, f_1)$ in endlichem Abstand von dieser Diskriminanten-Mannigf. befinden soll. Nun ist aber :

$$\chi(f, \overline{f_1}) = \chi(f, (1-\tau).f + \tau.f_1) = \chi(f, f_1),$$

also haben wir, von $(f, f_1)$ ausgehend, das Funktionensystem $(f, \overline{f_1})$ mit der gewünschten Eigenschaft erlangt.




23. Das Einzige, was wir über dasselbe voraussetzen, ist $R(f, f_1) \neq 0$, $R(f, f') \neq 0$ und $R(f_1, f_1') \neq 0$.

24. Correction dans la page de gauche : $\chi(f, f') = \chi(\overline{f_1}, f_1')$

Zur Erläuterung diene die nebenstehende Figur. Den Weg $I$ konnten wir machen ohne Mannigf. $R(f, f') = 0$ zu überschreiten, den Weg $II$ ohne Durchgang durch die Mannigf. $R(f, f_1) = 0$.

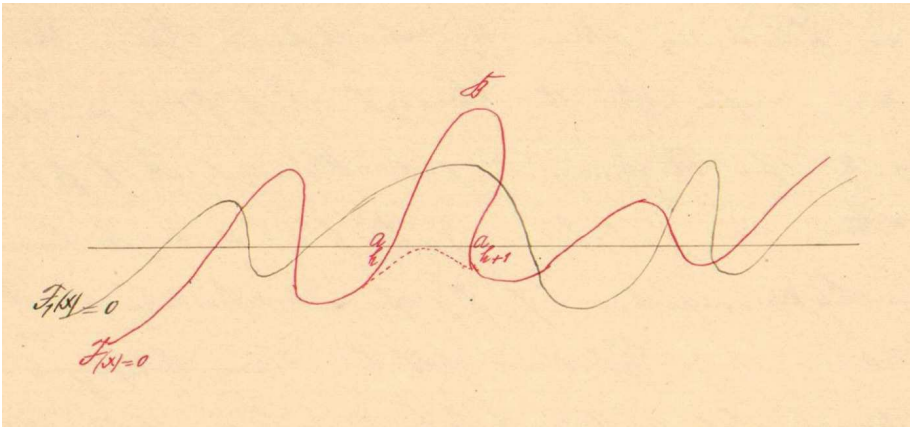
Wenn nun $\chi(f, f') = \chi(\overline{f_1}, \overline{f_1'}) = v$ ist, so können wir uns von dem Funktionensystem $(f, \overline{f_1})$ ohne Überschreitung der Mannigf. $R(f, f_1) = 0$, - dieses Erfordernis soll bei unsern künftigen Schritten als selbstverständlich nicht mehr besonders hervorgehoben werden, - | zu einem beliebig nahe benachbarten begeben, welches mit $(F, F_1)$ bezeichnet sei und so beschaffen ist, daß beide Funktionen je $v$ rationale Wurzel haben : 270|271

$$3.) \quad \begin{cases} F(x) = (x - a_1)...(x - a_v)\Phi(x) \\ F_1 = (x - b_1)...(x - b_v)\Psi(x). \end{cases}$$

Die Resultante von $F(x)$ und $F_1(x)$ ist, da allgemein

$$R(\varphi, \Theta.\psi) = R(\varphi, \psi).R(\varphi, \Theta) :$$

$$\begin{aligned} R(F, F_1) &= R(\Phi, \Psi).R(\Phi, \prod(x - b)).R(\Psi, \prod(x - a)).\prod(a_i - b_k) \\ &= R(\Phi, \Psi).\prod\Phi(b).\prod\Psi(a).\prod(a_i - b_k). \end{aligned}$$



Der Wert dieser Resultante kann bei einer Variirung des Funktionensystems nur durch Null hindurchgehen, wenn $\prod(a_i - b_k)$ die Null passiert ; Denn $\Phi(x)$ und $\Psi(x)$ können für keinen reellen Wert von $x$ verschwinden. Wir können ferner annehmen, daß zwischen je zwei rationalen Wurzeln von $F(x) = 0$ eine und nur eine Wurzel von $F_1(x) = 0$ liegt. Denn wenn z.B. die beiden Wurzeln $a_h$ und $a_{h+1}$ von $F(x) = 0$ durch keine Wurzel von $F_1(x) = 0$ getrennt werden, so führen wir in $F(x)$ an Stelle von $(x - a_h)(x - a_{h+1})$ das Produkt [25] $[(x - \frac{1}{2}(2 - t)a_h + t.a_{h+1}][x - \frac{1}{2}(2 - t)a_{h+1} + t.a_h]$ ein und lassen $t$ von 0...1 variiren, wodurch wir von $(x - a_h)(x - a_{h+1})$ nach $(x - \frac{1}{2}(a_h + a_{h+1}))^2$ gelangen. Entwickeln wir das Quadrat und ändern seine Koeffizienten nach einer bestimmten Richtung beliebig wenig, so wird die quadratische Funktion, gleich Null gesetzt, zwei komplexe Wurzeln erhalten (s. die Diskussion der Gleichung vierten Grades). Da sich also zwei solche Wurzeln stets beseitigen lassen (indem ich - anschaulich gesprochen, - in der nebenstehenden Figur den Bogen $\mathfrak{K}$ erst so weit hinunterziehe, daß er die Axe berührt und alsdann unter die Axe ziehe), so können wir 3.) auch in der Form : | 271|272

25. Correction dans la page de gauche : $[(x - \frac{1}{2}((2 - t)a_h + t.a_{h+1})][x - \frac{1}{2}((2 - t)a_{h+1} + t.a_h)]$

$$4.)\ \begin{cases} (x-\alpha_1)...(x-\alpha_\mu).\Phi(x) \\ (x-s_1)...(x-s_v)\Psi(x). \end{cases}$$

annehmen, wo $\alpha_1 < s_1 < \alpha_2 < s_2 < ... < \alpha_\mu < s_\mu$, und daher auch die Anzahl der rationalen Wurzeln beider Gleichungen wieder dieselbe sein muß. Das Kurvensystem, welches die Funktionen 4.) repräsentierten, hat also dieselben Eigenschaften wie das Kurvensystem in dem speziellen Falle der Charakteristik $\chi(f, f')$, mit dem einzigen Unterschiede, daß für 4.) nur Austrittspunkte vorhanden sind, während für das System $(f, f')$ nur Eintrittspunkte vorhanden waren. Da nun $\chi(f, f') = \frac{1}{2}(E - A)$, so ist für unsern Fall [26]

$$\chi(F, F') = -\frac{1}{2}\mu.$$

2. Febr.

Jetzt wandern wir von dem Funktionensystem 4.) aus in folgender weise : Wir gehen mit dem System

$$\prod_{h=1}^{\mu}[x - (\alpha_h(1-t) + (2h-1).t)].\Phi(x); \prod_{h=1}^{\mu}[x - (s_h(1-t) + 2h.t)].\Psi(x)$$

von $t = 0$ bis $t = 1$, wobei die Resultanten-Mannigf. $R(F, F_1) = 0$ nicht überschritten wird, da die Ungleichung für die Wurzeln :

$$\alpha_1(1-t)+t < s_1(1-t)+2t < \alpha_2(1-t)+3t < ... < \alpha_\mu(1-t)+(2\mu-1)t < s_\mu(1-t)+2\mu.t$$

für alle vorkommenden Werte von $t$ erhalten bleibt. Auf diese Weise gelangen wir zu dem Funktionensystem :

$$5).\ \prod_{h=1}^{\mu}(x-(2h-1))\Phi(x)\ ;\ \prod_{h=1}^{\mu}(x-2h).\Psi(x).$$

Es erübrigt jetzt nur noch, das Funktionensystem 5.) überzuführen in das System 2.), welches wir dem System 5.) analog schreiben wollen :

$$\varphi(x) = \prod_{h=1}^{\mu}(x-(2h-1))\Phi_0(x)\ ,\ \psi(x) = \prod_{h=1}^{\mu}(x-2h).\psi_0(x)$$

wo also die Bedeutung von $\Phi_0(x)$ und $\Psi_0(x)$ einleuchtet.
Da die Resultante von [27] 5.) $\prod\Phi(2h)\prod\Psi(2h-1).R(\Phi,\Psi)$ nur verschwindet, wenn $R(\Phi, \Psi)$ gleich Null wird, (vorausgesetzt, daß bei der Variirung stets $\chi(\Phi, \Phi') = 0$ und
272|273 $\chi(\Psi, \Psi') = 0$ bleibt), so können wir die | verbleibunde Aufgabe dahin präcisieren :

Es soll ein Weg gefunden werden, welcher die Funktionsysteme $(\Phi, \Psi)$ und $(\Phi_0, \Psi_0)$ mit einander verbindet, ohne durch die Mannigf. $R(\Phi, \Psi) = 0$ hindurchzugehen.

Wenn wir nun wie immer annehmen, daß die Koeffizienten der höchsten Glieder der Funktionen 1.) positiv sind, so werden die Werte der beiden Funktionen $\Phi(x)$ und $\Psi(x)$, in welchen dann auch die Koeffizienten der höchsten Potenzen positiv sind, für sehr große reelle Werte von $x$ und daher für alle reellen Werte des Arguments positiv sein, ebenso wie die Werte der Funktionen $\Phi_0$ und $\Psi_0$. Anstatt des Systems $(\Phi, \Psi)$ können wir auch ein sehr nahe benachbartes System $(\Phi_1, \Psi_1)$ nehmen, wenn wir nur darauf achten, daß wir auf dem Wege von $(\Phi, \Psi)$ nach $(\Phi_1, \Psi_1)$ weder die Diskriminanten-Mannigfaltigkeiten $R(\Phi, \Phi') = 0$, $R(\Psi, \Psi') = 0$ noch die Resultanten-Mannigf. $R(\Phi, \Psi) = 0$ überschreiten. Der gewünschte Weg von dem Punkte $(\Phi, \Psi)$ zu dem andern $(\Phi_0, \Psi_0)$ ist dann der folgende :

Wir begeben uns von $(\Phi, \Psi)$ zu $(\Phi_1, \Psi_1)$ (dies System wird noch näher bestimmt

26. Correction dans la page de gauche : $\chi(F, F_1) = -\frac{1}{2}\mu$
27. Correction dans la page de gauche : 5.) $\prod\Phi(2h)\prod\Psi(2h-1).R(\Phi,\Psi)\prod(2h-1-2k)$

werden) und von $(\Phi_1, \Psi_1)$ zu $(\Phi_0, \Psi_0)$ auf einer graden Linie

$$6.) \quad \begin{array}{l} \Phi_t(x) = (1-t).\Phi_1(x) + t.\Phi_0(x) \\ \Psi_t(x) = (1-t).\Psi_1(x) + t.\Psi_0(x) \text{ für } t \text{ von } 0...1 \end{array}$$

Diese Funktionen sind für alle Werte von $t$, welche wir durchlaufen, und alle reellen Werte von $x$ positiv, können also beide nur komplexe Linearfaktoren besitzen. Wenn wir also unterwegs die Resultanten-Mannigf. $R(\Phi, \Psi) = 0$ überschreiten, so muß die Überschreitungsstelle ein Punkt einer Mannigf. niederer Ordnung sein, welche sich auf | $R(\Phi, \Psi) = 0$ befindet, weil, wenn die Funktionen 6.) einen gemeinschaftlichen 273|274
Divisor besitzen, dieser mindestens von zweiten Grade sein muß. Sind die Funktionen $\Phi_1, \Psi_1$ also auch $\Phi_0, \Psi_0$ z.B. von vierten Grade, so heißt das in unseren früheren Anschauungen und Bezeichnungen (S. 227) : Ein grade Linie zwischen zwei Punkte des Gebiets $G_2$ kann die Diskriminantenfläche höchstens in einen Punkte der in $G_2$ hineinragenden zu der Fläche gehörenden Parabel durchschreiten. Die diesem Parabelpunkte zugehörenden Funktionen besitzen einen gemeinsamen Divisor vom zweiten Grade.

Um nun zu zeigen, daß man eine solche Mannigfaltigkeit, welche in ihrem grade um mehr als eine Einheit niedriger ist als diejenige, in welcher wir uns bewegen, stets umgehen kann, geht man sonst am zweckmäßigsten in der Art vor, daß man bis an die Mannigf. in grader Linie herangeht und dann, wenn es nötig ist, derselben ausweicht, - hier ist es leichter, den Ausgangspunkt $(\Phi_1, \Psi_1)$, welcher ja eine gewisse Variabilität besetzt, so zu fixieren, daß die grade Linie, welche $(\Phi_1, \Psi_1)$ mit $(\Phi_0, \Psi_0)$ verbindet, die Mannigf. nicht schneidet.

Diese Mannigf. haben wir zunächst zu bestimmen oder mit andern Werten, die beiden Gleichungen aufzusuchen, welche unter den unbestimmten Koeffizienten von

$$V(x) = v_0 + v_1 x + v_2 x^2 + ... + v_n x^n \; ; \; W(x) = w_0 + w_1 x + w_2 x^2 + ... + w_n x^n$$

stattfinden müssen, damit dieselben einen gemeinsamen Teiler von zweiten Grade besitzen. Die Bedingung dafür, daß die beiden Funktionen $V$ und $W$ überhaupt einen gemeinsamen Divisor haben, ist :

$$R(V, W) = R(v_0, v_1, ...v_n; w_0, w_1 ...w_n) = 0$$

| Daher haben wir identisch : 274|275

$$R(v_0 - V(u), v_1 ...v_n; w_0 - W(u), w_1 ...w_n) = 0,$$

da die Funktionen $V(x) - V(u), W(x) - W(u)$ den gemeinsamen Teiler $x - u$ besitzen. Da $v_0 - V(u) = -v_1 u - v_2 u^2 - ... - v_n u^n$ ist, so ergiebt die Differentiation dieser Identität [28] :

$$-u.\frac{\partial R(v_0 - V(u), v_1, v_2 ...)}{\partial (v_0 - V(u))} + \frac{\partial R(v_0, v_1, v_2 ...)}{\partial v_1} = 0$$

und wenn wir nach Potenzen von $V(u)$ und $W(u)$ entwickeln :

$$-u.\frac{\partial R(v_0, v_1, ...; w_0 ...)}{\partial v_0} + \frac{\partial R(v_0, v_1 ...; w_0 ...)}{\partial v_1} \equiv 0 \text{ (modd. } V(u), W(u)).$$

Schreiben wir anstatt der Unbestimmten $u$ die Variable $x$, so wird :

$$7) \quad -x.\frac{\partial R(V, W)}{\partial v_0} + \frac{\partial R(V, W)}{\partial v_1} \equiv 0 \text{ (modd. } V(u), W(u)).$$

Wenn nun $V(x)$ und $W(x)$ einen gemeinsamen Teiler haben, so muß auch die linke

28. Ajout dans la page de gauche : nach $v_1$

Seite von 7.) durch denselben teilbar sein. Ist diese Teilerfunktion linear, so kann sie nichts anderes als

$$-x.\frac{\partial R(V,W)}{\partial v_0}+\frac{\partial R(V,W)}{\partial v_1}$$

selbst sein ; ist sie jedoch quadratisch, so muß

$$8.)\ \frac{\partial R(V,W)}{\partial v_0}=0,\ \frac{\partial R(V,W)}{\partial v_1}=0$$

identisch bestehen.

Diese beiden Gleichungen sind notwendige Bedingungen für die Existenz eines gemeinsamen quadratischen Teilers der Funktionen $V(x)$ und $W(x)$ ; notwendig und hinreichend ist, das gleichzeitig

$$9.)\ R=0 \text{ und } \frac{\partial R(V,W)}{\partial v_0}=0 \text{ ist.}$$

Die gleichung 9.) ziehen die zweite der Gleichungen 8.) als notwendige Folgerung nach sich, denn da wir aus 7.) und 9.) den Schluß ziehen können, daß $\frac{\partial R(V,W)}{\partial v_1}$ durch eine Funktion von $x$ von mindestens ersten Grade teilbar ist, so muß $\frac{\partial R(V,W)}{\partial v_1}=0$ sein.

275|276 |

Für unsere speziellen Fall haben wir also zu zeigen, daß die Gleichung

$$10.)\ R(\Phi_t(x),\Psi_t(x))=0\ ,\ \frac{\partial R(\Phi_t(x),\Psi_t(x))}{\partial \varphi_{t,0}}=0,$$

wo $\varphi_{t,0}$ das absolute Glied der Funktion $\Phi_t(x)$ ist, für keinen Wert von $t$ zusammen bestehen können. Betrachten wir diese beiden in $t$ variablen Ausdrücke als Funktionen dieser einen Größe, so soll also die Resultante derselben

$$11.)\ R\left(R(\Phi_t(x),\Psi_t(x)),\frac{\partial R(\Phi_t(x),\Psi_t(x))}{\partial \varphi_{t,0}}\right)\neq 0 \text{ sein.}$$

Diese Resultante ist von $x$ und $t$ nicht mehr abhängig. Dagegen wird sie bei Variirung des Funktionensystems $\Phi_1(x),\Psi_1(x))$ andere und andere Werte annehmen. Wenn wir uns insbesondere in einer $v$-fachen Mannigf. bewegen und die Koeffizienten dieses Funktionensystems als ganze Funktionen von $v$ Parametern $x_1,x_2,...x_v$ auffassen, welche wir innerhalb bestimmter Grenzen variiren dürfen, so sind für 11.) zwei Fälle möglich. Entweder der Ausdruck auf der linken Seite von 11.), welcher also auch eine ganze Funktion von $x_1...x_v$ ist, verschwindet identisch oder nur für gewisse Werte dieser Veränderlichen. In zweiten Falle ist unser Wunsch erfüllt : wir können ein System $(\Phi_1,\Psi_1)$ finden, für welches die Ungleichung 11.) besteht und welches daher als Ausgangspunkt für unsern Weg (Gl. 6.)) dienen kann. Die Annahmen der Möglichkeit des ersten Falls führt dagegen auf einen Widerspruch. Denn wenn der Ausdruck in 11.) für alle Werte von $x_1,x_2,...x_v$ verschwindet, so muß er auch für diejenigen speziellen Werte von $x_1,x_2...x_v$ gleich Null werden, für welche aus dem Funktionensystem

276|277 | $(\Phi_1,\Psi_1)$ das system $(\Phi_0,\Psi_0)$ hervorgeht ; denn wir können durch unbeschränkte Veränderung der $v$ von einander unabhängigen Parameter zu jedem Funktionensystem der $v$-fachen Mannigf. gelangen. Da nun aber $R(\Phi_0,\Psi_0)\neq 0$ ist, so stehen wir von dem Widerspruch, daß zwei von $t$ unabhängige Größen $R(\Phi_0(x),\Psi_0(x))$ und $\partial R(\Phi_0(x),\Psi_0(x))\partial\varphi_0$, von denen jedenfalls die erste von Null verschieden ist, einen von $t$ abhängigen gemeinsamen Teiler haben sollen.

Damit haben wir bewiesen, daß die durch die Mannigf. $R(f,f_1)=0$ abgesonderten Gebiete, welchen die Funktionensysteme mit gleicher Charakteristik angehören, in dem oft angegebenen Sinne zusammenhängend sind.

Für den besonderen Fall $f_1=f'$ besagt dies Resultat, daß wir von jeder Gleichung

$f(x) = 0$ mit $\mu$ reellen Wurzeln durch Variirung ihrer Koeffizienten zu der Gleichung

$$(x-1)(x-2)...(x-\mu)(x^2+1^2)(x^2+2^2)...(x^2+\left(\left(\frac{n-\mu}{2}\right)^2\right)=0$$

gelangen können, ohne die Diskriminanten-Mannigf. zu überschreiten und daher auch zu jeder beliebigen andern Gleichung $g(x) = 0$, welche gleichfalls $\mu$ reelle Wurzeln besitzt (d.h für welche $\chi(g(x), g'(x)) = \frac{\mu}{2}$ ist).

## B.3 Traduction du chapitre du cours de 1890-1891 sur la théorie des caractéristiques

théorie arithméti

# Sur la caractéristique d'un système de fonctions

## I.) Considération introductive

| Dans notre discussion précédente nous avons déduit le théorème de *Sturm* 207|208
à partir d'un cas particulier d'une formule générale que nous avions obtenue, par l'application de la décomposition en fraction continue, sous la forme :

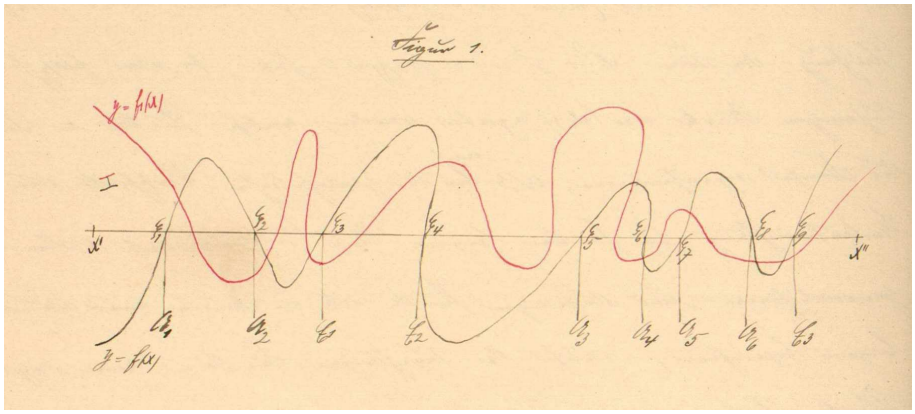
$$\sum_{x'<\xi<x''} sgn\ f'(\xi)f_1(\xi) = \frac{1}{2}\sum_{h=1}^{v}(sgn\ f_{h-1}(x'').f_h(x'') - sgn\ f_{h-1}(x').f_h(x'))$$

où les notations précédentes ont été conservées.

Cette généralisation de la formule de *Sturm* vient de *Sylvester*, qui a fait paraître plusieurs essais sur les expressions combinatoires précédentes, expressions qu'il a formées sous l'hypothèse de la décomposition d'une fonction en facteurs linéaires dans le *Philosophical Transactions* de 1853. Mais alors que pour le cas particulier où $f_1(x) = f'(x)$, la signification de la somme de signes est évidente, *Sylvester* s'est efforcé en vain d'interpréter de façon claire et exhaustive cette relation entre les signes, pour laquelle il a introduit le terme d'« intercallations ».

L'interprétation apparaît facilement lorsque l'on met à la place des fonctions $f(x), f_1(x)$ comme fondement de nos considérations les courbes paraboliques $y = f(x)$, $y = f_1(x)$.

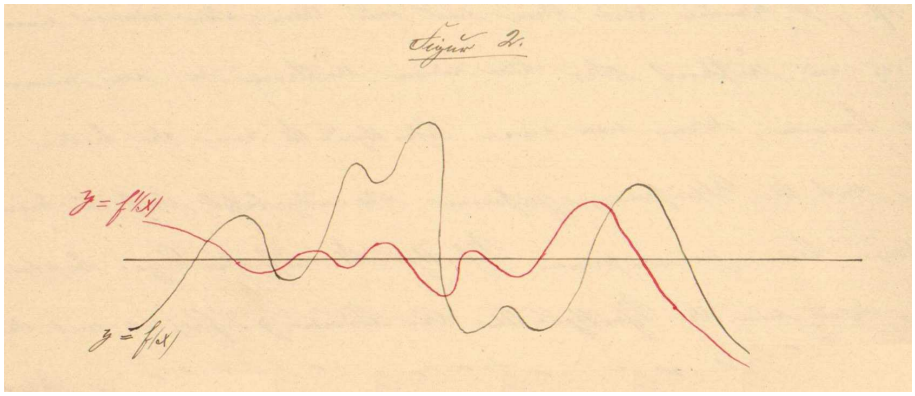
| Chaque courbe partage tous les points du plan en trois parties, celle qui 208|209
se trouve sur la courbe, celle qui se trouve, lors d'un déplacement sur la courbe dans un sens déterminé, à main droite et celle qui se trouve à main gauche. Si l'on exprime arithmétiquement cette répartition, alors on a la signification : nous pouvons séparer tous les systèmes de valeurs réelles et finies $(x, y)$ par rapport à une fonction quelconque $f(x, y)$ en trois classes, selon qu'il fait disparaître la fonction $f(x, y)$, qu'il la rend positive ou qu'il la rend négative. En même temps la fonction $f(x, y)$ doit bien sûr être donnée en une forme invariable, tout autant que nous devons toujours garder le même sens lors de notre cheminement sur la courbe. Si nous considérons par exemple tous les points d'un plan par rapport à un cercle dont l'équation a la forme $x^2 + y^2 - 1 = f(x, y) = 0$, alors $f(x, y)$ sera négative pour les valeurs de $(x, y)$ qui représentent des points à l'intérieur du cercle, donc, d'après la consigne choisie, qui se trouve sur la gauche pour le sens de rotation. Nous voulons généraliser cette façon de nous exprimer et nous désignerons comme se trouvant à l'intérieur de la courbe les points pour lesquels la fonction correspondante garde un signe négatif.



Si nous considérons sous ce point de vue nos courbes $y - f(x) = 0$ et $y - f_1(x) = 0$ ou plutôt l'ensemble des deux courbes, c'est-à-dire le produit $(y - f(x))(y - f_1(x))$,
209|210 alors | les points $(x, y)$ qui se trouvent à l'intérieur de la courbe $(y-f(x)).(y-f_1(x)) = \varphi(x, y) = 0$, sont ceux pour lesquels le produit a un signe négatif. Si l'on se déplace sur l'axe des abscisses (c'est-à-dire sur une troisième courbe $y = 0$) et que l'on considère donc seulement les systèmes de valeurs $(x, y)$ pour lesquels $y$ est constamment égal à zéro, nous irons alors et seulement alors de l'intérieur d'un domaine à l'extérieur ou de l'extérieur à l'intérieur d'un domaine, si nous passons par un point d'intersection de l'une des deux courbes avec l'axe des abscisses, en supposant que les deux fonctions n'ont pas de point d'intersection sur l'axe. Si l'on envisage maintenant le point d'intersection de la courbe $y - f(x) = 0$ avec l'axe, alors $sgn\ d(y - f(x))(y - f_1(x))$ devient, pour une valeur $x = \xi$ telle que $f(\xi) = 0$, égal à $+1$ ou à $-1$, selon que nous arrivons sur ce point de l'axe correspondant à cette valeur de l'intérieur vers l'extérieur (point de sortie), ou de l'extérieur vers l'intérieur (point d'entrée). Mais maintenant on a

$$[sgn\ d(y - f(x))(y - f_1(x))]_{x=\xi} = sgn\ f'(\xi)f_1(\xi),$$

d'où il ressort que la somme des signes $\sum_{x'<\xi<x''} sgn\ f'(\xi).f_1(\xi)$ correspond en bref au nombre de points de sortie qui se trouvent dans l'intervalle $x'...x''$, diminué du nombre de points d'entrée se trouvant dans le même intervalle.



La figure 1 au verso montre un exemple du cas précédent. La fonction $y = f(x)$ est

ici de degré impair ; nous | arrivons donc à de très grandes valeurs négatives pour les variables indépendantes $x$ et $y$, tandis que $f_1(x)$ est de degré pair. Le premier domaine, désigné par I, correspond à l'intérieur relativement au produit $(y - f(x))(y - f_1(x))$. Lors d'un trajet sur l'axe, nous quittons ce domaine au point $\xi_1$ et nous nous trouvons alors dans un domaine extérieur ; $\xi_1$ est donc un point de sortie et en tant que tel est désigné [29] par $A_1$. Pour le cas particulier du dessin, on a 210|211

$$\sum sgn\ f'(\xi).f_1(\xi) = 6 - 3 = 3.$$

Dans le cas particulier où $f'(x) = f_1(x)$, on obtient seulement des points de sortie, car entre deux zéros de la fonction se trouve toujours un nombre impair de zéro de la dérivée (Figure 2).

Ces remarques peuvent être généralisées en prenant à la place de l'axe des $x$ une troisième courbe en plus des deux autres ; il est cependant nécessaire, pour poursuivre arithmétiquement ces recherches, d'introduire une nouvelle notion : la notion de caractéristique d'un système de fonctions.

## II.) La caractéristique d'un système de deux fonctions d'une variable

15 Jan.

Soient $V(z)$ et $W(z)$ deux fonctions entières de degrés quelconques [30], et le déterminant

$$\begin{vmatrix} V(z) & W(z) \\ V'(z) & W'(z) \end{vmatrix} = \Delta(z)$$

de sorte que $\Delta(z)$ soit une fonction en $z$ de degré $(2n-1)$.

| Si l'on désigne alors par $\xi$ les valeurs pour lesquelles la fonction $V(z)$ s'annule, alors on appelle le nombre ou demi-nombre $-\frac{1}{2}\sum sign\ \Delta(\xi)$ la caractéristique du système de fonctions $(V(z), W(z))$ ou en termes arithmétiques : 211|212

1. $\chi(V(z), W(z)) = -\frac{1}{2}\sum_{(\xi)} sign\ \Delta(\xi)$

où la somme est à étendre à tous les $\xi$ pour lesquels la fonction $V(z)$ s'annule.

Il importe tout d'abord de montrer que lorsque $V(z)$ et $W(z)$ sont du même degré, on peut mettre à la place des racines de $V(z) = 0$ celles de $W(z) = 0$ dans l'équation de la définition 1). Dans ce but nous appliquons la formule que l'on a prouvée plus tôt (p. 186) :

$$\sum_{z_1<\xi<z_2} sign\ \Phi'(\xi) = -\frac{1}{2}\ sgn\ \Phi(z_1) + \frac{1}{2}\ sgn\ \Phi(z_2)$$

Ici $\Phi(z) = \Phi_0 z^n - \Phi_1 z^{n-1} + ...$ est une fonction entière de degré $n$. La somme sera étendue à toutes les valeurs $\xi$ qui se trouvent entre $z_1$ et $z_2$ et pour lesquelles $\Phi(\xi)$ s'annule. En particulier, lorsque $z_1$ est plus petit que la plus petite racine réelle et $z_2$ plus grand que la plus grande racine, on a [31] :

29. *N.d.T.* Le $A$ correspond au terme allemand de « sortie » (*Ausstritt*).
30. *N.d.T.* Correction sur la page de gauche : $n$
31. *N.d.T.* Correction dans la page de gauche :

$$\sum_{z_1<\xi<z_2} sgn\ \Phi'(\xi) = -\frac{1}{2}\ sgn\ z_1^n\Phi_0 + \frac{1}{2}\ sgn\ z_2^n\Phi_0$$
$$= 0\ \text{, si } n \text{ est pair}$$
$$= sgn\ \Phi_0\ \text{, si } n \text{ est impair}$$

$$\sum_{z_1<\xi<z_2} sgn\ \Phi'(\xi) = -\frac{1}{2}\ sgn\ z_1^n\Phi_0 + \frac{1}{2}\ sgn\ z_2^n\Phi_0$$
$$= 0\ ,\ \text{si } n \text{ est pair}$$
$$= 1\ ,\ \text{si } n \text{ est impair}$$

Comme $V(z)$ et $W(z)$ sont supposées du même degré, alors leur produit
212|213 $V(z).W(z)$ est de degré pair, la somme des signes de la dérivée est donc égal à zéro, c'est-à-dire que :

2. $\sum_{(\xi)}\ sgn\ V'(\xi).W(\xi) + \sum_{(\eta)}\ sgn\ V(\eta).W'(\eta) = 0$, si $\eta$ sont les valeurs pour lesquelles $W(\eta) = 0$, donc :

$$\frac{1}{2}\sum_{(\xi)}\ sgn\ V'(\xi).W(\xi) = -\frac{1}{2}\sum_{(\eta)}\ sgn\ V(\eta).W'(\eta)$$

Par conséquent, comme

$$\Delta(\xi) = -V'(\xi).W(\xi) \qquad \Delta(\eta) = V(\eta).W'(\eta),\ \text{on a :}$$

3. $\chi(V,W) = -\frac{1}{2}\sum_{(\xi)}\ sgn\ \Delta(\xi) = \frac{1}{2}\sum_{(\xi)}\ sgn\ V'(\xi).W(\xi)$
$= -\frac{1}{2}\sum_{(\eta)}\ sgn\ V(\eta).W'(\eta) = -\frac{1}{2}\sum_{(\eta)}\ sgn\ \Delta(\eta).$

Il ressort de plus de la forme de déterminant de $\Delta(z)$ que

4. $\chi(V,W) = -\chi(W,V)$ et que
5. $\chi(V, aV+W) = -\frac{1}{2}\sum_{(\xi)}\ sgn\ \begin{vmatrix} V & aV+W \\ V' & aV'+W' \end{vmatrix} = \chi(V,W)$

$$= -\frac{1}{2}\sum_{(\eta)}\ sgn\ \begin{vmatrix} V+aW & W \\ V'+aW' & W' \end{vmatrix} = \chi(V+aW, W).$$

Nous pouvons exprimer la somme de signes connue $\sum_{x'<\xi<x''}\ sgn\ f'(\xi).f_1(\xi)$ par la différence de deux caractéristiques. On a en effet :

$$\chi(f(z),(x''-z).f_1(z)) = \frac{1}{2}\sum_{(\xi)}\ sgn\ (x''-\xi).f_1(\xi).f'(\xi) = -\frac{1}{2}\sum_{(\xi)}\ sgn\ \Delta(\xi)$$

$$\chi(f(z),(x'-z).f_1(z)) = \frac{1}{2}\sum_{(\xi)}\ sgn\ (x'-\xi).f_1(\xi).f'(\xi) = -\frac{1}{2}\sum_{(\xi)}\ sgn\ \overline{\Delta}(\xi),$$

où la somme s'étend à toutes les racines $\xi$ de l'équation $f(z) = 0$. Comme maintenant $sgn\ (x''-\xi) = \pm sgn\ (x'-\xi)$, selon que $\xi$ se trouve à l'extérieur ou à l'intérieur de l'intervalle délimité par $x'$ et $x''$, alors on a :

$$\chi(f(z),(x''-z).f_1(z)) - \chi(f(z),(x'-y).f_1(z)) = \sum_{x'<\xi<x''}\ sgn\ f'(\xi).f_1(\xi)$$
$$= -\frac{1}{2}\sum_{(\xi)}\ sgn\ \Delta(\xi)$$
$$+\frac{1}{2}\sum_{(\xi)}\ sgn\ \overline{\Delta}(\xi)$$

213|214 La somme des signes est donc l'accroissement de la caractéristique | du système de fonctions $(f(z),(x-z)f_1(z))$ lors du passage de $x'$ à $x''$.

Ce dernier système de fonctions, dans lequel les coefficients de la deuxième fonction dépendent d'une grandeur variable $x$, peut être conçu comme un cas

particulier du système de fonctions $(V(z), W(z))$, où $V(z) = v_0 + v_1 z + ... + v_n z^n$, $W(z) = w_0 + w_1 z + ...w_n z^n$ et les $2n$ coefficients[32] $v_0, v_1 ... v_n; w_0, w_1 ... w_n$ sont des variables. Nous nous demandons à quel moment la caractéristique $\chi(V, W)$ change lors des variations des coefficients, et quelle est la valeur du changement lorsque l'on passe d'un système de coefficients déterminé $(v'_0..v'_n; w'_0...w'_n)$ à un autre $(v''_0..v''_n; w''_0...w''_n)$.

La caractéristique du système de fonctions peut alors manifestement seulement changer, si le signe du déterminant $\Delta(z)$ change pour l'une des racines $\xi$ de $V(z) = 0$ ou des racines $\eta$ de $W(z) = 0$, donc s'il passe par zéro. Une condition nécessaire pour que la caractéristique change est donc qu'on ait une racine commune aux deux équations

$$\begin{aligned} &\text{I.)}\ V(z).W'(z) - V'(z).W(z) = 0 \text{ et } V(z) = 0 \text{ ou soit} \\ &\text{II.)}\ V(z').W'(z') - V'(z').W(z') = 0 \text{ et } W(z') = 0 \end{aligned}$$

lors du changement continu des coefficients.

Cela est encore le cas pour I.) lorsque l'on a simultanément soit 1.) $V(z) = 0$, $V'(z) = 0$ soit 2.) $V(z) = 0$, $W(z) = 0$ ; et pour II.), lorsque l'on a simultanément soit 1a) $W(z') = 0$, $W'(z') = 0$ soit 2.) $W'(z') = 0$, $V(z') = 0$.

1.) ou 1a) (la disparition du discriminant | de l'une des deux fonctions) ne 214|215
sont pourtant pas des conditions suffisantes pour le changement de la caractéristique et seulement l'existence de l'équation 2.) (la disparition du résultant) a une telle conséquence nécessaire. En effet $\chi(V(z'), W(z'))$ s'annule si et seulement si $\chi(V(z'), W(z')) - x.V(z'))$ est aussi égal à zéro (éq. 5.)) donc lorsque par exemple d'après l'éq. II : 6) $W(z') - xV(z') = 0$ et 7.) $V(z')(W'(z') - x.V'(z') = 0$ alors 8.) $V(z').W(z') - W(z').V'(z') = 0$. Cette dernière équation 8.) détermine nos valeurs de $z'$, lesquelles doivent satisfaire l'équation 6.), et cela pour chaque valeur de la grandeur variable $x$ ; on doit donc avoir $V(z') = 0$ et $W(z') = 0$. C'est la seule condition qui vient de II.), tandis que de I.) il résultera encore $V(z) = 0$, $V'(z) = 0$. Mais du fait que cette condition ne soit pas suffisante aussi pour une modification de la caractéristique, il s'ensuit encore qu'elle ne peut avoir lieu que simultanément avec un changement de $\chi(V(z) - xW(z), W(z))$, changement pour lequel pourtant seulement $V(z) = 0$, $W(z) = 0$ ou $W(z) = 0$, $W'(z) = 0$ sont des conditions nécessaires. Ainsi il est donc démontré :

La valeur de la caractéristique est modifiée lors d'une variation continue du système des coefficients de ses fonctions si et seulement si le résultant des deux fonctions est nul.

Comme $\chi(V(z), W(z)) = -\frac{1}{2} \sum\ sgn\ \Delta(\xi)$, alors la caractéristique ne peut diminuer ou augmenter que d'une unité, selon que le $\Delta(\xi)$ en question, lequel se modifie, passe de $+1$ à $-1$ ou de $-1$ à $+1$. Si nous passons d'un système de coefficients $(v_k^{(1)}, w_k^{(1)})$ $(k = 0, 1, 2, ...n)$ à un autre $(v_k^{(2)}, w_k^{(2)})$, alors le changement de la caractéristique est :

$$-\frac{1}{2}\ sgn\ \Delta(v_k^{(1)}, w_k^{(1)}) + \frac{1}{2}\ sgn\ \Delta(v_k^{(2)}, w_k^{(2)}),$$

donc la modification | totale 215|216

9. $\chi(v'_0...; w'_0...) - \chi(v''_0...; w''_0...)$
$= \sum_{(h)} \left(-\frac{1}{2}\ sgn\ \Delta(v_k^{(h)}, w_k^{(h)}) + \frac{1}{2}\ sgn\ \Delta(v_k^{(h+1)}, w_k^{(h+1)})\right)$

où la sommation dans la première somme se rapporte à toutes les valeurs $h$ qui sont telles qu'entre les systèmes $(v_0^{(h)}...v_n^{(h)}; w_0^{(h)}...w_n^{(h)})$ et $(v_0^{(h+1)}...v_n^{(h+1)}; w_0^{(h+1)}...w_n^{(h+1)})$ se trouve un système $(v_0^{(h_1)}...v_n^{(h_1)}; w_0^{(h_1)}...w_n^{(h_1)})$ pour lequel $R = 0$ et en conséquence la sommation de la deuxième somme se fait sur toutes les valeurs $\xi_h$ pour lesquelles

32. *N.d.T.* Il y a en fait $2n + 2$ coefficients.

$V(\xi_h) = 0$ et $W(\xi_h) = 0$.

---

Pour maintenant vraiment calculer le changement de la caractéristique, nous pouvons très bien utiliser la même déduction dont nous nous sommes déjà servi lors de la démonstration du théorème généralisé de *Sturm* ; nous voulons récapituler rapidement le procédé avec la modification des notations correspondantes.

Soit

$$V(z) = v_0 + v_1 z + v_2 z^2 + ... + v_n z^n = v_n \prod (z - \xi)$$
$$W(z) = w_0 + w_1 z + w_2 z^2 + ... + w_n z^n = w_n \prod (z - \eta).$$

$R(V(z), W(z))$ n'est plus une fonction entière des coefficients, mais seulement entière en les quotients $\frac{v_0}{v_n}, ... \frac{v_{n-1}}{v_n}$. De plus, nous définissons :

$R_{n-m} = (-1)^{\frac{n(n-1)}{2}} T_{m,m}$ $(m = 0, 1...n)$ et nous obtenons ainsi la suite : $R_0, R_1 ... R_n$ dans laquelle on doit poser $R_0 = R$, $R_n = 1$.

Comme maintenant :

$$(-1)^{\frac{n(n-1)}{2}} T_{nn} = R(V, W) \text{ (p. 202) et}$$
$$(-1)^{\frac{n(n-1)}{2}} T_{n-1,n-1} = R(V, W) \sum_{i=1}^{n} \frac{1}{W(\xi_i).V'(\xi_i)} \text{ (p. 205),}$$

alors on obtient :

216|217 10. $\frac{R_1}{R} = \frac{T_{n-1,n-1}}{T_{n,n}} = \sum_{i=1}^{n} \frac{1}{W(\xi_i).V'(\xi_i)}$ ou : |

$$R_1 = \frac{R}{W(\xi_i).V'(\xi_1)} + \frac{R}{W(\xi_2).V'(\xi_2)} + ...$$

Il peut désormais être obtenu pour $V(z)$ et $W(z)$ par une variation continue des coefficients une racine commune, par exemple $\xi_1 = \eta_1$, de sorte que, comme on l'a vu dans le dernier développement, dans le voisinage de $\xi_1$ tous les termes du côté droit deviennent aussi petits que l'on veut, à l'exception du premier, dans lequel le numérateur et dénominateur deviennent aussi petits que l'on veut. De là il résulte que

$$sgn\ R_1 = \ sgn\ \frac{\delta R}{\delta(W(\xi_1).V'(\xi_1))}$$

ou

$$sgn\ R_1.\delta R = \ sgn\ \delta(W(\xi_1).V'(\xi_1)).$$

Maintenant de 9) il suit :

11. $\chi(v_0'...; w_0'...) - \chi(v_0''...; w_0''...) = -\frac{1}{2} \sum_{(h)} \ sgn\ \delta[W(x).V'(x)]_{x=\xi_h}$
$= -\sum_{(h_1)} \ sgn\ R_1(v_k^{(h_1)}.w_k^{(h_1)}).\delta R(v_k^{(h_1)}.w_k^{(h_1)})$ ou :

12. $\chi(v_0'...; w_0'...) - \chi(v_0''...; w_0''...) = \sum_{(h_1)} \ sgn\ R_1(v_k^{(h_1)}.w_k^{(h_1)}).\delta R(v_k^{(h_1)}.w_k^{(h_1)})$

en désignant $\chi(v_0''...; w_0''...)$ la caractéristique du système de fonctions

$$V^{(2)}(x) = v_0'' + v_1'' x + ... + v_n'' x^n \ , \ W^{(2)}(x) = w_0'' + w_1'' x + ... + w_n'' x^n$$

et $\chi(v_0'...; w_0'...)$ celle du système de fonctions

$$V^{(1)}(x) = v_0' + v_1' x + ... + v_n' x^n \ , \ W^{(1)}(x) = w_0' + w_1' x + ... + w_n' x^n.$$

On a maintenant (p. 180) : $T_{m+1,m+1}.T_{m-1,m-1} \equiv -(\ )^2 (mod\ T_{m,m})$ donc

13. $R_{h-1}.R_{h+1} \equiv -(\ )^2 (mod\ R_h)$.

La même chose était valable pour la suite

14. $T_0(x), T_1(x), ...T_n(x)$, c'est-à-dire que l'on avait aussi

$$T_{m+1}(x).T_{m-1}(x) \equiv -(\ )^2 (mod\ T_m(x)).$$

De là il a suivi que le nombre de changements de signes de | la suite 14) ne change 217|
pas si lors de la variation continue des $x$ une des fonctions, sauf $T_n(x)$, passe par zéro 218
et que le nombre de changements de signes perdus lorsque l'on laisse $x$ faire le chemin de $x'$ à $x''$ peut être exprimé par la somme (p. 187)

$$\begin{aligned} w' - w'' &= \sum_{x'<\xi<x''} sgn\ T'_n(\xi).T_{n-1}(\xi) \\ &= -\frac{1}{2}\sum_{h=1}^{n} sgn\ T_h(x').T_{h-1}(x') + \frac{1}{2}\sum_{h=1}^{n} sgn\ T_h(x'').T_{h-1}(x''), \end{aligned}$$

où la première somme s'étend à toutes les valeurs de $\xi$ pour lesquelles $T_n(\xi) = 0$ lors du passage continu de $x'$ à $x''$.

Nous pouvons immédiatement appliquer ce corollaire à notre suite

$$R_0, R_1, ...R_n$$

puisque la seule différence à noter est que l $R_0, R_1...R_n$ ne dépendent pas d'<u>une</u> grandeur variable $x$, mais dépendent des $2(n+1)$ coefficients de $V(x)$ et $W(x)$. C'est pourquoi on doit mettre aussi lors de la démonstration à la place de la dérivée $T'_n(x)$ la différentielle totale $\delta R(v_k, w_k)$. D'après cette réflexion, nous sommes autorisé à compléter l'équation 12) de la manière suivante :

15. $\chi(v''_0, w''_0) - \chi(v'_0, w'_0) = \sum_{(h_1)} sgn\ R_1(v_k^{(h_1)}, w_k^{(h_1)}).\delta R(v_k^{(h_1)}, w_k^{(h_1)}) = \frac{1}{2}\sum_{h=0}^{n-1} sgn\ R_h(v''_0, w''_0)R_{h+1}(v''_0; w''_0) - \frac{1}{2}\sum_{h=0}^{n-1} sgn\ R_h(v'_0, w'_0)R_{h+1}(v'_0; w'_0)$.

Si nous connaissons, pour un système de coefficients déterminé, la valeur de la caractéristique, alors nous la connaissons pour tous les systèmes de coefficients.

Supposons maintenant que les $2(n+1)$ coefficients $v_0...v_n; w_0....w_n$ sont dépendants des $\nu$ variables $x_1, x_2, ...x_\nu$, lesquelles sont mutuellement indépendantes et 218|219
forment ainsi une variété de dimension $v$. Alors il correspond à chaque point un système de fonctions $(V(z), W(z))$ et sa caractéristique, et inversement. L'équation $R = 0$ fixe une variété de dimension $(\nu - 1)$, laquelle est issue d'une variété de dimension $v$. Si l'on se déplace d'un point $(x'_1, x'_2, ...x'_\nu)$ à un autre $(x''_1, x''_2, ...x''_\nu)$, alors la valeur de la caractéristique changera si et seulement si on passe par la variété $R = 0$ de dimension $(\nu - 1)$. Considérons maintenant la variété $R.R_1 = 0$ de dimension $(v - 1)$. Celle-ci sépare notre variété de dimension $\nu$ en un domaine positif et un domaine négatif, et nous pouvons avec l'utilisation de la notion de point de sortie et d'entrée pour la variété $R.R_1 = 0$ exprimer la formule 15.), laquelle se laisse aussi écrire sous la forme

$$\chi(v''_0, w''_0) - \chi(v'_0, w'_0) = \sum_{(h_1)} sgn\ [\delta R(v_k^{(h_1)}, w_k^{(h_1)}).\delta R_1(v_k^{(h_1)}, w_k^{(h_1)})],$$

de la manière suivante :

La caractéristique ajoute sur le trajet d'un point à un autre de la variété de dimension $\nu$ à un autre autant d'unités que le nombre de points de sortie surpasse celui des points d'entrée.

Ainsi le contenu de la généralisation par *Sylvester* du théorème de *Sturm* (p. 213) s'exprime dans la forme suivante.

La généralisation par *Sylvester* du théorème de *Sturm* donne la modification de la caractéristique d'un système de deux fonctions, lorsque l'on conçoit l'une des fonctions comme une constante et la seconde comme une fonction linéaire d'une seule variable $x$. Considérons désormais le cas particulier de la caractéristique du système de 219|220

fonctions $(V(z), V(z)+V'(z))$. Pour cela, on a

$$\Delta(z) = \begin{vmatrix} V & V+V' \\ V' & V'+V'' \end{vmatrix} = V.V'' - V'^2.$$

Le résultant de $V(z)$ et $V'(z)$ est le discriminant de $V(z)$, et leur caractéristique est $-\frac{1}{2}\sum_{(\xi)} \; sgn\; \Delta(\xi) = +\frac{1}{2}\sum_{(\xi)} \; sgn\; (V'(\xi))^2$, donc elle est égale à la moitié du nombre de racines réelles de $V(x) = 0$.

De 11) il suit de plus la modification de la caractéristique :

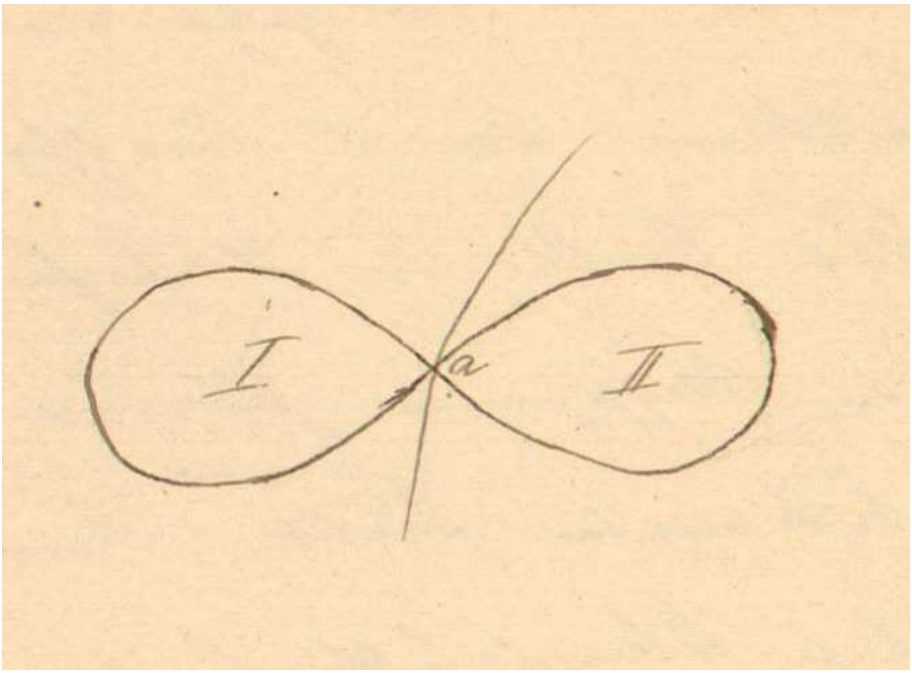
$$\chi(v_0''..., w_0''...) - \chi(v_0'..., w_0'...) = \frac{1}{2}\sum_{(h)} \; sgn\; [V''(x).\delta V(x)]_{x=\xi_h},$$

où on désigne par $\xi_h$ les valeurs pour lesquelles $V(x)$ et $V'(x)$ s'annulent en même temps.

Si nous nous représentons maintenant les coefficients de l'équation $V(x) = 0$ comme des fonctions de $\nu$ variables indépendantes $x_1, x_2...x_\nu$, alors nous pouvons attribuer à chaque point de la variété de dimension $\nu$ une équation avec des coefficients déterminés, et inversement fixer à chaque équation un point déterminé. Ensuite, d'après ce qui précède, la variété d'ordre $(\nu-1)$ dont l'équation est $D = 0$ sépare les domaines dans lesquels se trouvent les équations avec un nombre différent de racines réelles, et ce de sorte que l'on doive traverser inévitablement le discriminant pour aller d'une équation à une autre qui n'a pas le même nombre de racines réelles que la première.

Il ressort de ces recherches qu'en fait déjà le discriminant nous rend seul le même service que nous pouvions jusqu'à présent attendre seulement des suites de fonctions
220|221 de *Sturm* : de déterminer le nombre de racines réelles d'une | équation.

Cela nécessiterait en effet seulement une discussion exhaustive de la situation de




la variété-discriminante et des parties dans lesquelles il sépare les variétés supérieures pour pouvoir assigner une place à chaque équation dans ce dernier et afin de pouvoir déterminer le nombre de ses racines réelles. Nous appliquerons cela particulièrement au cas de l'équation du quatrième degré et nous verrons qu'il importe avant tout de trouver toutes les singularités de la variété.

Les fonctions de *Sturm* donnent des variétés contenant ces variétés singulières

d'ordre inférieur qui se trouvent sur la variété-discriminante, ou dans le cas général dans le résultant-variété, et par leurs signes elles déterminent précisément le domaine dans lequel se trouve l'équation concernée (à propos du système de fonctions en jeu.).

Par exemple, pour illustrer ce que nous venons de dire, supposons que le discriminant a la forme d'une lemniscate. Il est alors nécessaire pour pouvoir séparer les points des deux domaines I et II les uns des autres par des indications arithmétiques, de placer au point double $a$ une courbe. On fixe alors la position d'un point du plan par rapport à chacune des deux courbes par le signe qui donne ses coordonnées aux deux fonctions. Il est ainsi arithmétiquement précisément déterminé, duquel des trois domaines séparés par la lemniscate le point fait partie.

On observe de plus que les fonctions de *Sturm* ne sont pas absolument déterminées pour une fonction fixée et qu'elles ne sont aussi limitées concernant leur nombre seulement par une limite inférieure.

| Nous avons démontré que les domaines auxquels appartiennent les équations 221|222
ayant le même nombre de racines réelles sont séparés les uns des autres par la variété-discriminante ; dans la suite nous allons montrer que l'on peut aller continûment d'une équation à une autre qui a le même nombre de racines réelles sans passer par la variété-discriminante ; c'est-à-dire que donc ces domaines sont aussi connexes. La même chose a lieu pour le cas général de la caractéristique d'un système de fonctions et du résultant-variété associé.

Nous voulons montrer que nous pouvons « voyager » de l'équation 1.) $\prod_{k=1}^{n}(w - w'_k) = 0$ à une autre 2.) $\prod_{k=1}^{n}(w - w''_k) = 0$ sans dépassement du déterminant, sous l'hypothèse que les deux équations ont le même nombre de racines réelles et que toutes les racines de chacune des deux équations sont différentes les unes des autres. Pour atteindre ce but, nous supposons les racines $w_k = u_k + v_k i$ ordonnées selon les grandeurs de leur partie imaginaire $v_k$ (non pas ordonnées selon la valeur absolue, mais en considérant le signe de $i$) et les racines avec la même partie imaginaire d'après la grandeur de leur partie réelle $u_k$. Nous formons ainsi :

$$3.)\ \prod_{k=1}^{n}(w - (1-t)w'_k - t.w''_k) = 0$$

pour $0 \leq t \leq 1$, de sorte que nous passons de manière continue de la première équation à la seconde lorsque $t$ parcourt toutes les valeurs de 0 à 1. Il est inutile de montrer que pour aucune des valeurs de $t$ le discriminant de cette équation :

$$\prod_{h<k}\left((w'_k - w'_h)(1-t) + (w''_k - w''_h)t\right)^2$$

ne peut être nul.

| Si tel était le cas, un des facteurs, c'est-à-dire la partie réelle et imaginaire, 222|223
devrait disparaître séparément. Mais si l'on pose $w'_k - w'_h = u'_k + v'_k i - (u'_h + v'_h i) = w'_k - w_h + i(v'_k - v'_h)$, alors d'après notre classement on a soit $v'_k - v'_h > 0$, soit si $v'_k = v'_h$ alors $u'_k - u'_h > 0$, et de même pour le cas de $w''_k - w''_h$. Comme maintenant $t$ et $(1-t)$ sont aussi des grandeurs positives, alors la partie imaginaire du facteur est positive, ou tout au moins si elle doit s'annuler, la partie réelle est plus grande que zéro. Les deux parties peuvent seulement alors s'annuler si $w'_k = w'_h$ et $w''_k = w''_h$ ; mais ces égalités contredisent nos deux hypothèses.

Il faut cependant encore éliminer l'objection qui peut être facilement soulever contre cette argumentation : la preuve a été apportée que l'on peut arriver à partir de chaque équation à n'importe quelle autre sans franchir la variété-discriminante, puisque l'exigence d'un nombre égal de racines réelles, ce qui a été fait pour les

équations présentées, avait été loin d'être épuisé dans la preuve. À cela il faut répondre que seulement sous cette hypothèse l'équation 3) a des coefficients réels. Car lorsque par exemple l'équation 1) possède plus de racines complexes que la 2), alors il y aura assurément dans le produit du côté droit de 3) deux facteurs dans lesquels il apparaîtra un $w'_k$ complexe et un $w''_k$ réel et dans lesquelles les deux racines complexes $w'_k$ sont conjuguées l'une à l'autre. Mais le produit de ces deux facteurs linéaires en $w$ peut seulement ensuite avoir des coefficients réels, lorsque les deux racines réelles $w''_k$ sont égales l'une à l'autre, - et nous avions exclu ce cas.

223|224

## | III.) Application aux équations du quatrième degré.

Nous voulons désormais utiliser nos discussions générales au cas particulier où $V(z)$ est une équation du quatrième degré et pour cela déterminer la position et les limites du domaine dans lesquels notre variété sera découpée. Nous utilisons dans ce but la formule déduite précédemment (p. 176) :

$$1.)\ \sum_{(\xi_i)} sgn\ \Theta(\xi_i) = \sum_{m=0}^{n-1} sgn\ T_{m,m}.T_{m+1,m+1},$$

où la première somme s'étend à toutes les racines réelles de l'équation $V(y) = 0$ et $\Theta$ désigne une fonction quelconque. Pour $\Theta = 1$ le $T$ peut devenir $\overline{T}$ et le côté gauche représente alors le nombre $t$ de racines réelles de l'équation $V(y) = 0$. Si $w$ représente le nombre de changements de signes, $\varphi$ celui de la chaîne de signes dans la suite

$$2.)\ T_{00}, T_{11}, ...T_{nn},$$

il résulte de 1.) :

$$A = \sum_{m=0}^{n-1} sgn\ \overline{T}_{m,m}.\overline{T}_{m+1,m+1} = \varphi - w = n - 2w,$$

d'où il résulte que $w$ représente le nombre de paires de racines complexes de l'équation $V(y) = 0$. De plus on a :

$$\overline{T}_{m,m} = |s_{p+q}|\ (p, q = 0, 1, 2, ...m-1)$$

où $s_\alpha$ représente comme toujours la somme des puissances $\alpha-$ième de toutes les racines. Nous avons donc une suite de $(n+1)$ tels déterminants, puisque $m$ peut avoir les valeurs $0, 1, ...n$, bien que le nombre de changements de signes ne puisse pas être plus grand que $\frac{1}{2}n$, resp. $\frac{1}{2}(n-1)$, selon que $n$ est pair ou impair. Il s'agit encore de savoir s'il est possible de représenter linéairement par les $\frac{1}{2}n$, resp. $\frac{1}{2}(n-1)$ premiers
224|225 déterminants tous ceux qui restent | de sorte que l'on puisse y discerner clairement cette limitation du nombre de changements de signes. Pour les cas $n = 3, 4, 5$ ce type de représentation est possible, à savoir avec des coefficients qui sont des sommes de carrés tour à tour positives et négatives, de sorte que lorsque les premiers $\frac{n}{2}$ resp. $\frac{n-1}{2}$ termes de la suite présentent seulement des changements de signes, le nombre de changements de signes ne peut plus être augmenté par les autres.

Pour notre cas particulier nous prenons maintenant comme équation :

$$3.)\ F(w) = 3x + 4y.w \star -6z.v^2 + w^4 = 0,$$

dès lors que nous avons débarrassé celui-ci du terme qui contenait la puissance trois. Si l'on part de cette forme d'équation, alors on peut interpréter les trois coefficients

variables $x, y, z$ comme les coordonnées rectangulaires de l'espace se rapportant à un système d'axes $X, Y, Z$. Ainsi toutes les équations du quatrième degré remplissent un espace de dimension trois. Il s'agit pour nous de séparer les uns des autres les domaines dans lesquels se trouvent les équations avec 0, 1 et 2 paires de racines imaginaires. Nous les désignerons resp. par $G_0, G_1$ et $G_2$. Le discriminant de l'équation, laquelle est posée égale à zéro, représente une surface, qui est

$$\begin{aligned} 4.a)\ D &= |s_{p+q}|_{p,q=0,1,2,3} \\ &= (x+z^2)^3 - (3xz+y^2-z^3)^2, \text{ (à un facteur numérique près)} \\ 4.b)\ &= x(x-3z^2)^2 - 2y^2z(3x-z^2) - y^4, \end{aligned}$$

puisque $s_0 = 4$, $s_1 = 0$, $s_2 = 12z$, $s_3 = -12y$, $s_4 = 72z^2 - 12x$. Avec le discriminant $D$ les | 225|226

$$D_1 = \frac{1}{4.12^2}\begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix} = -xz - y^2 + 3z^3,$$

$$D_2 = \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix} = z$$

$$\frac{1}{4}s_0 = 1$$

forment les fonctions de notre suite de Sturm. Le nombre de racines réelles de l'équation est par conséquent :

$$5.)\ A = \ sgn\ D.D_1 + \ sgn\ D_1.z + \ sgn\ z + \ sgn\ 1.$$

L'équation $F(w) = 0$ avec le paramètre arbitraire $w$ représente pour le système de coordonnées $X, Y, Z$ une famille de plan, et comme, à une constante numérique multiplicative près

$$D = R\left(F(w), \frac{\partial F(w)}{\partial w}\right),$$

alors $D = 0$ est l'enveloppe de la famille de plan et par conséquent une surface développable.

La surface discriminante contient toutes les équations avec une paire de racines égales. Sa courbe en hélice se laisse représenter par l'intersection de deux surfaces dont les équations proviennent de $F(w) = 0$, $F'(w) = 0$, $F''(w) = 0$ par l'élimination de $w$, et contient les équations avec une racine triple. Elle semble être ainsi aussi la courbe commune des deux surfaces $D = 0$ et $D_1 = 0$. Le plus facile est de déduire ces équations de la représentation de $D$ dans $4.a)$ :

$$6.)\ x + z^2 = 0 \quad y^2 - 4z^3 = 0.$$

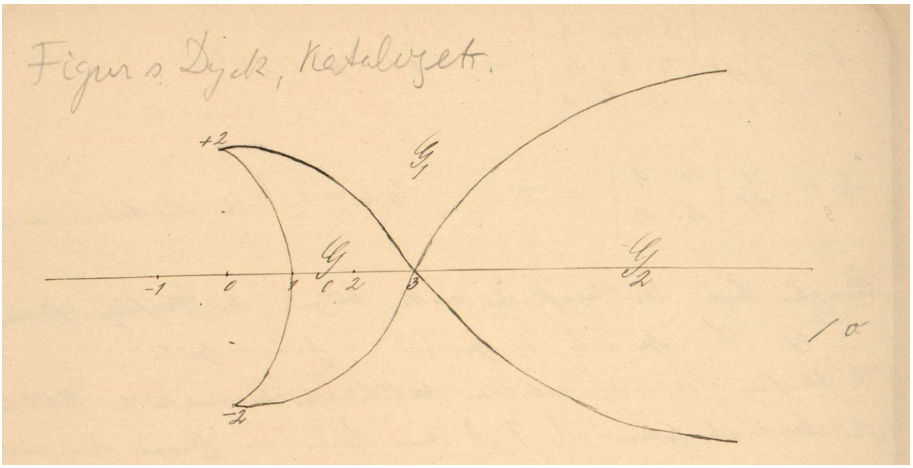
La surface discriminante possède cependant encore une deuxième courbe singulière, une courbe double, dont les équations proviennent de $4.b)$ sous la forme :

$$7.)\ x - 3z^2 = 0 \quad y = 0.$$

Celle-ci est une parabole | 226|227

dans le plan $XZ$ et le lieu de toutes les équations avec deux paires de racines égales.

Pour se faire une idée de la forme de la surface $D = 0$, nous voulons en montrer




l'intersection avec le plan $z = 1$. L'équation de la coupe est :

$$(x+1)^3 - (3x + y^2 - 1)^2 = 0$$

et elle est satisfaite pour :

$$x = -1\ ,\ y = \pm 2\,;\ x = +3\ ,\ y = 0\,;\ x = +1\ ,\ y = 0.$$

Le domaine $G_0$ dans lequel se trouvent les points qui représentent les équations avec seulement des racines réelles est le plus petit. Il se taille en forme de sac selon le plan $XY$ et possède dans celui-ci seulement l'origine des coordonnées. Sur le côté négatif du plan $XY$ le domaine, puisqu'il doit être connexe, ne peut posséder aucun prolongement. Par contre le domaine $G_2$ s'étend autant que $G_1$ par le plan $XY$, le premier le long de l'axe $Z$ et le dernier contient la surface discriminante dans toute sa longueur. Cependant la surface $D = 0$ se trouve sur le domaine des coordonnées où $Z$ est négatif exclusivement du côté droit (positif) du plan $YZ$. Nous obtenons en effet par exemple pour $z = -1$ à partir de 4.$a$) :[33]

$$8.)\ (x+1)^3 - (-3x + y^2 + 1)^2 = 0 \text{ ou } x^3 - 6y^2 + 9x - y^4 + 6xy^2 - 2y^2 = 0$$

une équation qui fournit pour des valeurs négatives de $x$ seulement des valeurs imaginaires de $y$. Les domaines $G_0$ et $G_2$ ont comme frontière commune la parabole 7.) dont le sommet est au point origine des coordonnées. Si l'on fait bouger de façon continue un point du domaine $G_0$ jusqu'à un point de $G_1$, alors dans les équations
227|228 attribuées à chaque point, au | moment du dépassement de la surface discriminante, deux des racines sont égales l'une à l'autre et sont ensuite complexes. Lors d'un autre déplacement jusqu'à un point de $G_2$ les deux autres racines sont égales l'une à l'autre au point d'intersection des surfaces et ensuite complexes. Si par contre on passe directement au travers de la parabole d'un point de $G_0$ à un de $G_2$, alors au point de la parabole les quatre racines sont égales par paire et ensuite complexes. Mais maintenant il se trouve <u>à l'intérieur</u> du domaine $G_2$ des équations avec deux paires identiques de racines complexes, et cela semble être comme si nous pouvions passer ici avec les racines doubles sans traverser la surface discriminante. Cependant l'autre branche de notre parabole s'étend à ce domaine, et sur celle-ci se trouvent les équations avec deux paires identiques de racines complexes. Cette branche de parabole apparaît par exemple dans la coupe caractérisée par l'équation 8.), en donnant cette

33. *N.d.T.* Correction dans la page de gauche :

$$8.)\ (x+1)^3 - (-3x + y^2 + 1)^2 = 0 \text{ ou } x^3 - 6x^2 + 9x - y^4 + 6xy^2 - 2y^2 = 0$$

et commentaire : « seulement une moitié »

équation pour $x = 3$ sauf pour les valeurs $y = \pm 4$ et aussi encore $y = 0$. Ce dernier point $(x = 3, y = 0, z = -1)$ est un point de la parabole, tout comme le point $x = 3, y = 0, z = +1$ dans la coupe annotée. Les quatre racines sont seulement égales à l'origine des coordonnées, où courbe double et courbe en hélice coïncident, et l'équation correspondant au point est $w^4 = 0$.

De cette discussion il résulte que les trois domaines $G_0, G_1, G_2$ sont caractérisés par les déterminations de signes suivantes :

| | | | | | |
|---|---|---|---|---|---|
| $G_0$ | $D > 0$ | $z > 0$ | $3z^2 > x$ | | |
| $G_1$ | $D < 0$ | | | | |
| $G_2$ | $D > 0$ | $z > 0$ | $3z^2 < x$ | $D > 0$ | $z < 0$ |

228|229

|

Le nombre de racines complexes sera déterminé par le nombre de changements de signes de la suite

$D, D_1, z, 1$ donc par exemple pour le domaine $G_1$ et
$z > 0$ - $+sgn\ D_1$ + +
$z < 0$ - $+sgn\ D_1$ - +
Comme il ne peut y avoir qu'un changement de signe, alors dans le premier cas on doit avoir $D_1 > 0$, et au contraire $D_1 < 0$ dans le second. - Pour $z < 0$, seul le signe de $D$ permet de décider dans quel domaine (dans $G_1$ ou dans $G_2$) nous nous trouvons ; au contraire pour $z > 0$ il faut encore décider si $3z^2 > x$ ou $3z^2 < x$.

Outre $D = 0$, les surfaces qu'une séparation des domaines devrait produire sont selon les indications de notre suite de fonctions :

$$D_1 = 0 \text{ et } z = 0.$$

La surface $D_1 = 0$ contient, comme indiqué précédemment, la courbe double. À la place de cette surface nous pouvons plus facilement prendre le cylindre qui possède comme courbe génératrice la parabole. Mais en aucun cas les deux égalités $D_1 = 0$ et $z = 0$ peuvent être remplacées par une seule, comme *Clebsch* en a donné une preuve fausse dans son livre « Theorie der linären algebraischen Formen ».

## IV.) Caractéristique d'un système de 3 fonctions de deux variables.

Nous avons eu à considérer, pour poser les fondements de la notion de caractéristique d'un système de fonctions, les trois fonctions :

$$y - f(x)\ ,\ y - f_1(x)\ ,\ y$$

et | à étudier les positions relatives de leurs courbes, que l'on obtient en posant ces 229|230
fonctions égales à zéro. À l'issue de ce préliminaire (p. 211) nous avions déjà montré que la question se généralise si l'on étudie trois fonctions quelconques de $x$ et $y$ :

$$f(x, y)\ ,\ \varphi(x, y)\ ,\ \psi(x, y)$$

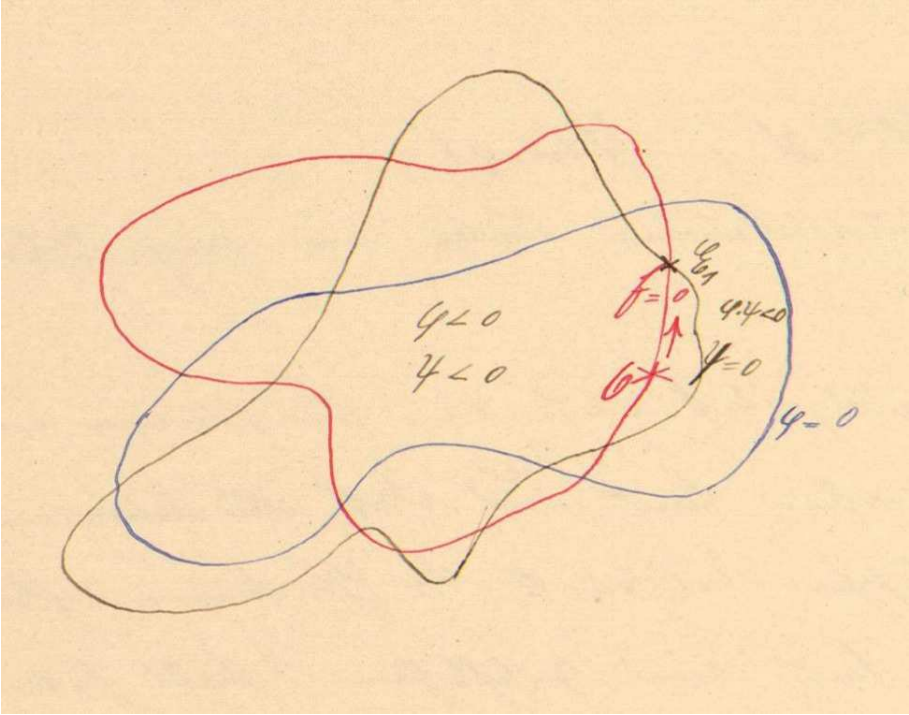
et les courbes correspondantes sous le point de vue mentionné.

Nous supposons que ces trois fonctions sont univalentes, continues et différentiables, et que les courbes dont les expressions analytiques sont les fonctions $f(x, y) = 0$, $\varphi(x, y) = 0$, $\psi(x, y) = 0$ sont fermées. Cette dernière propriété peut être caractérisée arithmétiquement de la façon suivante : il est possible de donner une grandeur finie fixée $R$ telle que pour toutes les valeurs $x, y$ qui satisfont une des trois équations $f = 0$, $\varphi = 0$, $\psi = 0$ on ait l'inéquation $x^2 + y^2 < R^2$.

Nous formons maintenant les déterminants :

1. $\begin{vmatrix} f & \varphi & \phi \\ f_1 & \varphi_1 & \phi_1 \\ f_2 & \varphi_2 & \phi_2 \end{vmatrix} = \Delta$, où l'indice 1 désigne la dérivée selon $x$ et l'indice 2 celle selon $y$. Si de plus $\Delta_1, \Delta_2, \Delta_3$ représentent les déterminants fonctionnels de resp. $\varphi, \psi; \psi, f; f, \varphi$, alors on peut écrire :
2. $\Delta = f\Delta_1 + \varphi\Delta_2 + \psi\Delta_3$, et $\Delta$ est lui-même le déterminant fonctionnel de $(z.f, z.\varphi, z.\psi)$ pour $z = 1$.

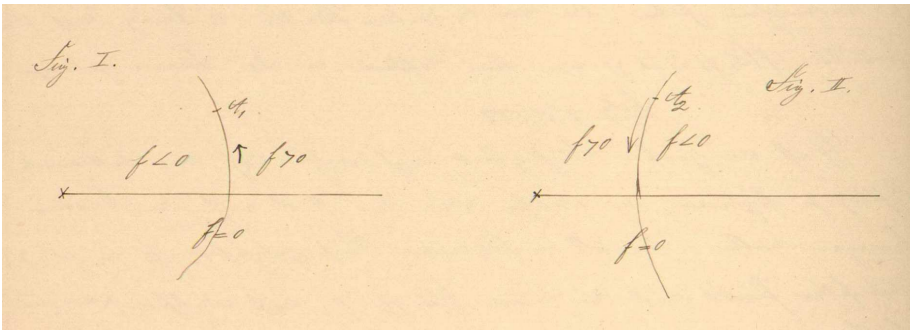
Si nous étudions maintenant la configuration générale formée par nos trois courbes
230|231 | de la même façon que nous l'avons fait dans le cas particulier, alors il s'avère qu'il y a ici six façons de les considérer, contre deux dans le cas précédent. Tandis que nous avions en effet parcouru à l'endroit cité soit le cours de la courbe $f = 0$ soit celui de la courbe $f_1 = 0$ et que nous avons considéré ses points d'intersection avec l'axe des abscisses, nous pouvons ici nous déplacer soit sur la courbe $f = 0$ soit sur $\varphi = 0$ ou soit enfin sur la courbe $\psi = 0$ et nous pouvons étudier les points d'intersection de la courbe en question avec l'une des deux autres. Si nous parcourons par exemple dans la figure ci-contre la courbe $f(x, y) = 0$ à partir de $O$, et que nous observons les points d'intersection avec la courbe $\psi(x, y) = 0$, nous devrons consigner en $\xi_1$ un point d'entrée, car en ce point, pour ce qui concerne le produit $\varphi(x, y).\psi(x, y)$, nous passons de l'extérieur à l'intérieur.




Principe de parcours

Il est d'une grande importance de fixer arithmétiquement une direction déterminée dans laquelle nous voulons toujours nous déplacer. Cela se fera de la manière suivante. Soit $(x', y')$ un point fixé sur la courbe $f(x, y) = 0$ à partir duquel nous voulons nous déplacer et par lequel passe une courbe quelconque $\Theta(x, y) = 0$. Ensuite pour le cas où on considère les points d'intersection avec la courbe $\varphi = 0$, on doit avoir $\frac{\partial|f,\Theta,\psi|}{\partial\psi} d\Theta > 0$, c'est-à-dire :

$$3.)\ sgn\ d\Theta = sgn\ \frac{\partial|f, \Theta, \psi|}{\partial\psi} = sgn\ (f_1\Theta_2 - f_2\Theta_1).$$




Par l'équation 3) on détermine si au point $(x', y')$ on passe, par rapport à la courbe $\Theta(x, y) = 0$, de l'intérieur à l'extérieur ou de l'extérieur à l'intérieur. Si l'on considère les points d'intersection avec $\psi = 0$, alors la direction est | l'opposée, car on doit bien 231|232
avoir

$$sgn\ d\Theta = sgn\ \frac{\partial|f, \varphi, \Theta|}{\partial\varphi}.$$

Le sens ainsi fixé reste le même quelle que soit la courbe $\Theta = 0$ que nous aurons prise. Ainsi, si par exemple $\Theta = y$ , alors on a $d\Theta = dy$, $f_1\Theta_2 - f_2\Theta_1 = f_1$, et donc notre équation 3) devient :

$$4.)\ f_1.dy > 0.$$

Maintenant, comme nous nous déplaçons sur la courbe $f = 0$, pour chaque point on a :

$f_1dx + f_2dy = 0$, $dx = -\dfrac{f_2dy}{f_1}$ et par l'introduction de ces valeurs dans $d\Theta = \Theta_1dx + \Theta_2dy$ on a $d\Theta = (f_2\Theta_2 - f_2\Theta_1)\frac{dy}{f_1}$, c'est-à-dire

$$sgn\ d\Theta = \ sgn\ (f_1.\Theta_2 - f_2.\Theta_1).sgn\ \frac{dy}{f_1},$$

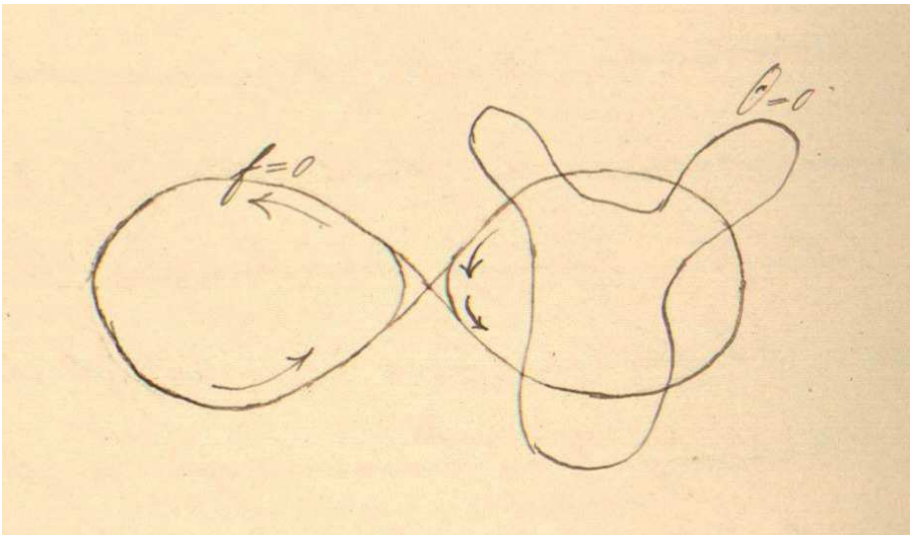
de sorte que sous l'hypothèse $dy.f_1 > 0$ on a aussi

$$sgn\ d\Theta = \ sgn\ (f_1.\Theta_2 - f_2.\Theta_1)$$

pour n'importe quel $\Theta$. Cette détermination de la direction - nous l'appellerons le principe de parcours - est dans le cas du plan équivalente à la règle fixant le sens de parcours de la courbe de sorte que l'intérieur reste à gauche. Ainsi par exemple dans la figure I pour les coordonnées $x_1, y_1$ du point $\mathcal{A}_1$ on a $f_1 > 0$, car lorsque $x$ croît et $y$ est constante les valeurs de la fonction $f(x, y)$ augmentent, donc d'après l'équation 4) $dy > 0$, les coordonnées en $y$ doivent augmenter et la courbe monter. Dans la figure II par contre les valeurs de $f(x, y)$ décroissent, alors nous devons parcourir la courbe en descendant. Mais dans les deux cas l'intérieur de la courbe se trouve à gauche. Cependant cette dernière détermination n'est déjà plus applicable pour un déplacement dans l'espace, car ici une « droite » ou une « gauche » ne pourrait pas être différencié, et pour entrer dans l'étude de variétés de plus grande dimension, fixer *arithmétiquement* un principe de parcours est bien une condition | nécessaire 232|233
préalable.

Imaginons maintenant dans la fonction $f(x, y)$ les variables $x$ et $y$ comme étant des fonctions d'une variable $z$, $x = V(z)$, $y = W(z)$ et demandons-nous comment l'équation 3) $sgn \begin{vmatrix} f_1 & f_2 \\ \Theta_1 & \Theta_2 \end{vmatrix} = sgn\ d\Theta$, laquelle définit le principe de parcours, se modifie. De $f(V(z), W(z))$ il suit :

$f_1V' + f_2W' = 0$ , $d\Theta = (V'.\Theta_1 + W'\Theta_2)dz$, de sorte que 3) se change en




$(f_1\Theta_2 - f_2\Theta_1)(V'.\Theta_1 + W'.\Theta_2)dz > 0$. Si l'on y pose $f_2 = -f_1.\frac{V'}{W'}$, alors on obtient :

$$(f_1\Theta_2 + f_1.\frac{V'}{W'}.\Theta_1)(V'\Theta_1 + W'\Theta_2)dz = \frac{f_1}{W'}(V'\Theta_1 + W'\Theta_2)^2dz > 0,$$

d'où résulte la détermination voulue

$$sgn\ f_1.W' = sgn\ dz = -sgn\ f_2.V'$$

tandis que pour la fonction initiale on doit avoir $sgn\ f_1 = sgn\ dz$, ou $-sgn\ f_2 = sgn\ dx$. De là il résulte que ce principe de parcours n'est pas applicable aux points pour lesquels $f_1 = f_2 = 0$, si bien que les singularités, comme par exemple les points doubles, ne peuvent pas être traversées, et on doit les parcourir séparément comme sur les deux parties d'une lemniscate.

Valeur et signification de la caractéristique.

Si $f(x,y) = 0$ et $\Theta(x,y) = 0$ sont deux courbes fermées, alors le nombre de points d'intersection des deux courbes où la courbe $\Theta = 0$ croît, est tout aussi grand que le nombre où elle décroît. On a donc ensuite l'égalité :

6.) $\sum_{f=0,\Theta=0} sgn\ d\Theta = \sum_{f=0,\Theta=0} sgn \begin{vmatrix} f_1 & f_2 \\ \Theta_1 & \Theta_2 \end{vmatrix} = 0$. Cette proposition facile (233)|(234) nous rend des services essentiels pour la théorie des caractéristiques. Tout d'abord nous l'utilisons pour le cas $\Theta = \varphi.\psi$, en supposant que $\varphi(x,y) = 0$ et $\psi(x,y) = 0$ représentent deux courbes fermées et qu'ainsi la courbe représentée par le produit $\varphi.\psi = 0$ a cette propriété.

Donc d'après l'égalité 6.), on doit avoir :

$$\sum_{f=0,\phi.\psi} sgn\ [f_1(\varphi_2.\psi + \varphi.\psi_2) - f_2(\varphi_1.\psi + \psi_1.\varphi)] = \sum_{f=0,\phi.\psi} sgn\ (\psi.\Delta_3 - \varphi.\Delta_2) = 0,$$

égalité qui peut aussi s'écrire sous la forme :

7.) $\sum_{f=0,\varphi=0} sgn\ \psi.\Delta_3 - \sum_{f=0,\psi=0} sgn\ \varphi.\Delta_2 = 0$. Il résulte aussi de 6.) :

8.) $\sum_{f=0,\varphi=0} sgn\ d(\varphi.\psi) = \sum_{f=0,\psi=0} sgn\ d(\varphi\psi)$, où dans la dernière somme le sens dans lequel on parcourt la courbe $f = 0$ doit être l'inverse de celui de la première

somme.

On a de plus avec l'utilisation de la condition $\frac{dy}{f_1} > 0$ :

$$\sum_{f=0,\varphi=0} sgn\ \psi.\Delta_3 = \sum_{f=0,\varphi=0} sgn\ \psi(f_1.\varphi_2 - f_2\varphi_1)$$

$$= \sum_{f=0,\varphi=0} sgn\ \psi.(f_1.\varphi_2 - f_2\varphi_1)\frac{dy}{f_1} = \sum_{f=0,\varphi=0} sgn\ \psi(\varphi_1 dx - \varphi_2 dy)$$

$$= \sum_{f=0,\varphi=0} sgn\ d(\varphi.\psi).$$

Donc :

$$9.) \sum_{f=0,\varphi=0} sgn\ d(\varphi.\psi) = \sum_{f=0,\psi=0} sgn\ d(\varphi.\psi) = \sum_{f=0,\psi=0} sgn\ \varphi.\Delta_2 = \sum_{f=0,\varphi=0} sgn\ \psi.\Delta_3$$

$$= \sum_{f=0,\varphi=0} sgn\ \Delta = \sum_{f=0,\psi=0} sgn\ \Delta \text{ (éq. 2))}.$$

On obtiendrait de manière analogue pour $\Theta = f.\varphi$ : 234|235

|

$$10) \sum_{\varphi=0,\psi=0} sgn\ d(\psi.f) = \sum_{\varphi=0,f=0} sgn\ d(\psi.f) = \sum_{\varphi=0,f=0} sgn\ \psi.\Delta_3 = \sum_{\varphi=0\psi=0} sgn\ f.\Delta_1$$

$$= \sum_{\varphi=0\psi=0} sgn\ \Delta = \sum_{\varphi=0,f=0} sgn\ \Delta$$

et comme le dernier membre de l'équation 10.) apparaît aussi dans la 9.), alors les termes de la suite 9.) sont égaux à ceux de la suite 10.). Nous obtiendrions une troisième série pour $\Theta = f.\varphi$. En particulier on a :

$$11.) \sum_{f=0,\varphi=0} sgn\ \Delta = \sum_{\varphi=0,\psi=0} sgn\ \Delta = \sum_{\psi=0,f=0} sgn\ \Delta.$$

Nous pouvons faire une deuxième utilisation de la formule 6.) lorsque $\Theta = \varphi$. On obtient alors de 6.) :

$$12.) \sum_{f=\varphi=0} sgn\ \Delta_3 = 0 = \sum_{f=\varphi=0,\psi>0} sgn\ \Delta_3 + \sum_{f=\varphi=0,\psi<0} sgn\ \Delta_3$$

et comme d'après 9.) :

$$\sum_{f=\varphi=0} sgn\ \Delta = \sum_{f=\varphi=0,\psi>0} sgn\ \Delta_3 - \sum_{f=\varphi=0,\psi<0} sgn\ \Delta_3,$$

alors on obtient par addition et soustraction de ces deux équations :

$$\sum_{f=\varphi=0} sgn\ \Delta = 2 \sum_{f=\varphi=0,\psi>0} sgn\ \Delta_3 = -2 \sum_{f=\varphi=0,\psi<0} sgn\ \Delta_3,$$

et donc :

$$13.) \ -\frac{1}{2} \sum_{f=\varphi=0} sgn\ \Delta = \sum_{f=\varphi=0,\psi<0} sgn\ \Delta_3$$

$$= \sum_{f=\varphi=0,\psi<0} sgn\ (f_1\varphi_2 - f_2\varphi_1) = -\frac{1}{2}\sum_{f=\varphi=0} sgn\ d(\varphi.\psi)\ \text{(éq. 9.)).}$$

$$= \chi(f,\varphi,\psi).$$

Les membres contenus dans la suite 13.) sont des expressions différentes pour la caractéristique de trois fonctions $f, \varphi$ et $\psi$. Pour être complet, nous voulons établir
235|236 toutes les formes possibles que peut prendre celle-ci :
|

14.) $\chi(f,\varphi,\psi) =$

$$\begin{cases} -\frac{1}{2}\sum_{f=\varphi=0} sgn\ \Delta & -\frac{1}{2}\sum_{\varphi=\psi=0} sgn\ \Delta & -\frac{1}{2}\sum_{\psi=f=0} sgn\ \Delta \\ \sum_{\substack{\psi<0\\ f=\varphi=0}} sgn\ (f_1\varphi_2 - f_2\varphi_1) & \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1) & \sum_{\substack{\varphi<0\\ \psi=f=0}} sgn\ (\psi_1 f_2 - \psi_2 f_1) \\ -\frac{1}{2}\sum_{\varphi=\psi=0} sgn\ d(f.\varphi) & -\frac{1}{2}\sum_{\psi=f=0} sgn\ d(\varphi.\psi) & -\frac{1}{2}\sum_{f=\varphi=0} sgn\ d(\psi.f) \\ -\frac{1}{2}\sum_{f=\psi=0} sgn\ d(f.\varphi) & -\frac{1}{2}\sum_{\varphi=f=0} sgn\ d(\varphi.\psi) & -\frac{1}{2}\sum_{\psi=\varphi=0} sgn\ d(\psi.f) \end{cases}$$

Pour les six dernières expressions, il faut tenir compte du fait que, comme il résulte de l'équation 8), on doit toujours se déplacer sur la courbe qui n'est pas sous le signe différentiel et que nous avons écrite sous le signe somme à la deuxième place. Si l'on prend par exemple $\chi(f,\varphi,\psi) = -\frac{1}{2}\sum_{\psi=f=0} sgn\ d(\varphi.\psi)$, alors on doit parcourir la courbe $f = 0$ et on doit chercher ses points d'intersection avec la courbe $\psi = 0$. La signification géométrique de cette caractéristique résulte alors facilement des considérations de la p. 231. Pour le point d'entrée noté $\xi_1$ dans la figure on aura en effet $d(\varphi.\psi) < 0$, par contre pour un point de sortie on aura $d(\varphi.\psi) > 0$, de sorte que si $\mathcal{E}$ désigne le nombre de points d'entrée et $\mathcal{A}$ celui de points de sortie, on a :

$$\chi(f,\varphi,\psi) = \frac{1}{2}(\mathcal{E} - \mathcal{A})$$

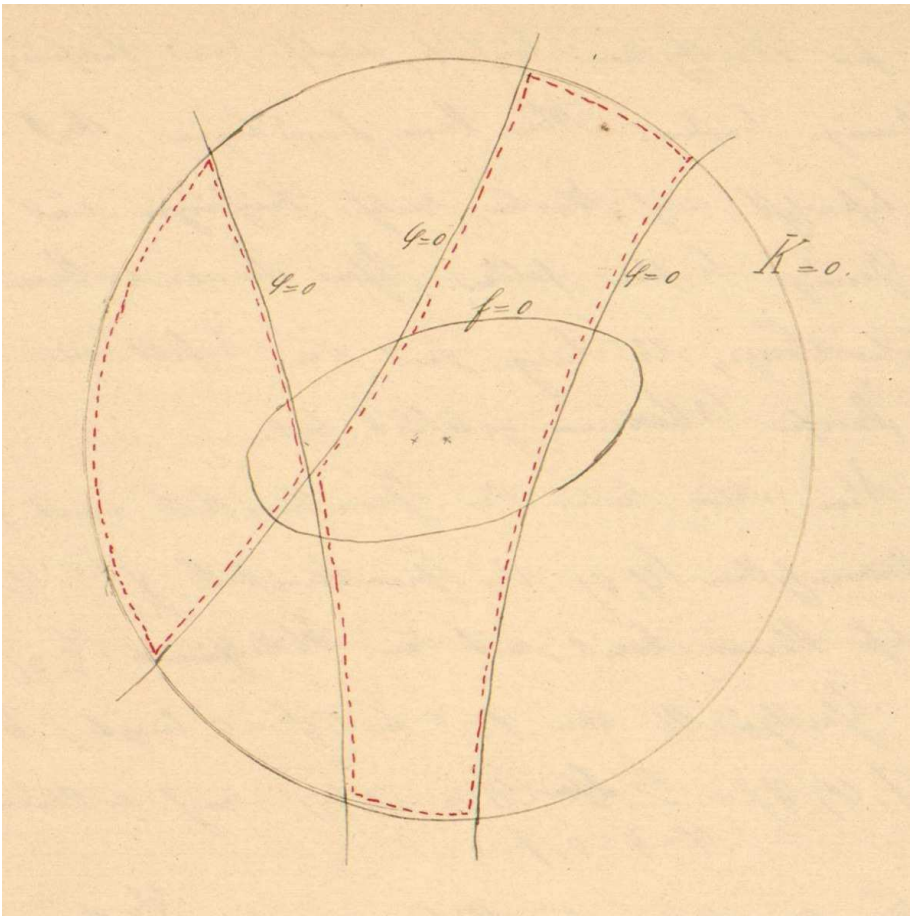
On peut de là, avec l'aide des expressions figurant dans la deuxième ligne de 14.), conclure que la différence $(\mathcal{E} - \mathcal{A})$ est toujours un nombre pair et que de plus elle reste la même quelle que soit celle des six combinaisons déjà mentionnées que l'on prenne comme fondement. Si l'on considère plus précisément les expressions de la deuxième ligne de 14.), alors on s'aperçoit que la caractéristique fournit aussi une
(236)|(237) somme algébrique de signes | pour un système de valeurs qui satisfait deux équations et une inéquation. On peut ajouter que la caractéristique gouverne toutes les questions sur les valeurs qui doivent satisfaire des équations et simultanément des inéquations. Déjà *Fourier* reconnaissait la grande portée de cette question, qui a longtemps été le contenu d'un grand sujet de concours de l'Académie de Paris.

Si l'on entend par caractéristique d'un point $(\xi,\eta)$ pour un système de fonctions $(\varphi,\psi)$ la caractéristique $\chi(\theta,\varphi,\psi)$, où $\theta$ est un très petit cercle de centre $(\varphi,\psi)$ dans lequel se trouve au plus un point d'intersection de $\varphi = 0$ et $\psi = 0$, alors nous pouvons aussi interpréter

$$\chi(f,\varphi,\psi) = \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1)$$

comme la somme des caractéristiques de tous les points isolés à l'intérieur de la courbe $f = 0$.

Extension des résultats au cas des courbes qui ne sont pas fermées.




Les recherches ci-dessus ne permettraient guère d'emploi notable si elles n'étaient pas transférables aux fonctions qui ne sont pas représentables par des courbes finies fermées. Il est ainsi nécessaire de montrer que la notion de caractéristique est aussi valable pour un système de fonctions à qui on impose exclusivement le respect de certaines conditions de continuité. Pour ce faire nous allons procéder de la manière suivante.

Imaginons le cercle $K = R^2 - x^2 - y^2 = 0$ d'un périmètre tel qu'il enferme tous les points d'intersection des deux courbes et considérons à la place des courbes non fermées $f = 0$, $\varphi = 0$ et $\psi = 0$ les courbes $F = f.K = 0$, $\Phi = \varphi.K = 0$, $\Psi = \psi.K = 0$. | Si nous les parcourons dans le sens donné par le principe de parcours, alors chacune 237|238
d'entre elle se divise en un certain nombre de courbes fermées. Ainsi par exemple dans la figure ci-contre les deux courbes fermées dessinées en pointillés représentent l'image géométrique qui entre uniquement en ligne de compte de $\varphi.K = 0$ et qui se substituent aux trois branches de $\varphi = 0$. Maintenant on définit :

$$\chi(f, \varphi, \psi) = \chi(F, \Phi, \Psi)$$

et comme :

$$\chi(F, \Phi, \Psi) = \sum_{\substack{F<0 \\ \Phi=\Psi=0}} sgn\ (\Phi_1\Psi_2 - \Phi_2\Psi_1) = \sum_{\substack{f<0 \\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1).K^2$$

(car les points d'intersection des courbes $\Phi = 0$ et $\Psi = 0$ sont à l'intérieur du cercle $K = 0$ exactement comme ceux de $\varphi = 0$ et $\psi$ et l'exigence $F < 0$ se réduit, comme

pour tous les points à considérer on a $K > 0$, à $f < 0$), alors on a :

$$15.)\ \chi(f,\varphi,\psi) = \sum_{\substack{f<0\\ \varphi=\psi=0}} sgn\ (\varphi_1\psi_2 - \varphi_2\psi_1)$$
$$= \sum_{\substack{\varphi<0\\ \psi=f=0}} sgn\ (\psi_1 f_2 - \psi_2 f_1)$$
$$= \sum_{\substack{\psi<0\\ f=\varphi=0}} sgn\ (f_1\varphi_2 - f_2\varphi_1),$$

car on peut permuter les fonctions dans l'expression $\chi(F,\Phi,\Psi)$. Pour la caractéristique des systèmes de fonctions généraux $\chi(f,\varphi,\psi)$, trois des douze énoncés du 14.) restent en vigueur.

Nous pouvons aussi maintenant expliquer comment la somme des signes a été interprétée dans l'introduction de ces considérations comme la caractéristique d'un système de trois fonctions. Nous construisons en effet à partir des trois fonctions $f(x) - y, f_1(x) - y, y$ pour lesquelles nous avons considéré comme configuration les
238|23) courbes correspondantes à la situation précédemment citée : |

$$\chi(f(x)-y, f_1(x)-y, y) = -\frac{1}{2} \sum_{\substack{f(x)-y=0\\ y=0}} sgn \begin{vmatrix} f(x)-y & f_1(x)-y & y \\ f'(x) & f_1'(x) & 0 \\ -1 & -1 & 1 \end{vmatrix}$$

$$= -\frac{1}{2} \sum_{\substack{f(x)-y=0\\ y=0}} sgn\ (f(x).f_1'(x) - f'(x).f_1(x)) = +\frac{1}{2} \sum_{(\xi)} sgn\ f'(\xi).f_1(\xi).$$

La dernière expression est la somme de signes que nous nous étions engagé à interpréter.

Modification de la caractéristique lors d'une variation de son système de fonctions.

D'après l'expression de la caractéristique valable pour toutes les fonctions continues on a [34] :

$$16.)\ \chi(f_1, \varphi - u.f, \psi - v.f) = \sum_{\varphi-u.f=\psi-v.f=0, f<0} sgn \begin{vmatrix} \varphi_1 - u.f_1 & \varphi_2 - u.f_2 \\ \psi_1 - v.f_1 & \psi_2 - v.f_2 \end{vmatrix}$$

$$\sum_{f=\psi-v.f=0,\varphi-u.f<0} sgn \begin{vmatrix} f_1 & f_2 \\ \psi_1 - v.f_1 & \psi_2 - v.f_2 \end{vmatrix} = \sum_{f=\psi=0,\varphi<0} \begin{vmatrix} f_1 & f_2 \\ \psi_1 & \psi_2 \end{vmatrix} = \chi(f,\varphi,\psi)$$

Si nous supposons maintenant que les coefficients des fonctions $f, \varphi$ et $\psi$ dépendent d'un paramètre variable $t$, alors on s'aperçoit d'après le deuxième membre de l'équation 16.) que la caractéristique $\chi(f,\varphi,\psi)$ peut seulement se modifier, si pour au moins

34. *N.d.T.* Version corrigée sur la page de gauche :

$$\sum_{f=\psi-v.f=0,\varphi-u.f<0} sgn \begin{vmatrix} \psi_1 - v.f_1 & \psi_2 - v.f_2 \\ f_1 & f_2 \end{vmatrix} = \sum_{f=\psi=0,\varphi<0} \begin{vmatrix} \psi_1 & \psi_2 \\ f_1 & f_2 \end{vmatrix} = \chi(f,\varphi,\psi)$$

Note sur la page de gauche : « Il sera montré que ce cas sans $f \neq 0$ ne peut pas se produire. Le cas $f = 0$ sera considéré plus tard. »

un couple $(x, y)$ on a simultanément :

$$\varphi - u.f = 0\ ,\ \psi - v.f = 0\ ,\ \begin{vmatrix} \varphi_1 - u.f_1 & \varphi_2 - u.f_2 \\ \psi_1 - v.f_1 & \psi_2 - v.f_2 \end{vmatrix} = \Delta_1 + u.\Delta_2 + v.\Delta_3 = 0.$$

Si l'on introduit dans la dernière expression $u = \frac{\varphi}{f}$, $v = \frac{\psi}{f}$, il résulte alors les trois conditions :

$$17.)\ \varphi - u.f = 0\ ,\ \psi - v.f = 0\ ,\ \Delta = 0.$$

De $\Delta = 0$ on obtient pour $x$ et $y$ seulement des valeurs qui ne dépendent pas de $u$ et $v$ | ; on doit avoir, comme $u$ et $v$ sont des indéterminées, pour de tels couples : 239|240

$$f = 0\ ,\ \varphi = 0\ ,\ \psi = 0.$$

Si l'on suppose maintenant que nous pouvons, de la même façon que plus tôt pour deux fonctions d'une variable, trouver une fonction symétrique des racines de $f = 0\ ,\ \varphi = 0\ ,\ \psi = 0$ dont la disparition est une condition nécessaire et suffisante pour l'existence d'un facteur commun entre les trois fonctions, et que nous avons aussi joint à cette fonction symétrique des racines le nom de « résultant », alors on peut dire :

La caractéristique d'un système de trois fonctions modifie sa valeur lors d'une variation continue du système de coefficients de ses fonctions si et seulement si le résultant des trois fonctions s'annule.

Si l'on suppose de même comme plus tôt les coefficients de $f$, $\varphi$ et $\psi$ comme des fonctions indépendantes de $v$ variables, alors à chaque point des variétés de dimension $v - 1$ fixées par ces dernières correspond un système de fonctions avec sa caractéristique. La variété $R = 0$ de dimension $(v - 1)$, lorsque $R$ désigne encore le résultant, sépare les uns des autres les domaines dont le système de fonctions a des caractéristiques différentes, de sorte que l'on ne peut sûrement jamais, sans traverser la variété $R = 0$, passer d'un système de fonctions à un autre dont la caractéristique est différente de celle du premier. 240|241

|

## V.) Application des résultats à la démonstration du théorème fondamental de l'algèbre de *Gauss* et des intégrales de *Cauchy*.

Nos recherches sur la caractéristique d'un système de fonctions nous permettent sans grande peine de démontrer la propriété des fonctions entières dont le contenu est désigné par *Gauss* comme le « Théorème fondamental de l'algèbre ». Celui-ci s'énonce ainsi :

Soit $\Theta(x + yi)$ une fonction entière de degré $n$ de la variable complexe $(x + yi)$ avec des coefficients complexes et, après le rassemblement de tous les termes réels et complexes :

$$\Theta(x + iy) = \varphi(x, y) + i.\psi(x, y);$$

alors les courbes définies par $\varphi = 0$ et $\psi = 0$ ont toujours $n$ points d'intersection.

Soit d'abord $\Theta(x + yi)$ une fonction quelconque continue et monogène, alors, comme on le sait, on a :

$$\varphi_1 = \psi_2\ ,\ \varphi_2 = -\psi_1.$$

Si maintenant $f(x, y) = 0$ représente une courbe fermée, alors la caractéristique est :

$$1.)\ \chi(f, \varphi, \psi) = \sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn\ \begin{vmatrix} \varphi_1 & \varphi_2 \\ \psi_1 & \psi_2 \end{vmatrix} = \sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn\ (\varphi_1^2 + \varphi_2^2) = \sum_{\substack{\varphi=\psi=0 \\ f<0}} sgn\ 1;$$

c'est-à-dire que $\chi(f,\varphi,\psi)$ nous donne le nombre de points d'intersection des courbes $\varphi=0$ et $\psi=0$ qui se trouvent dans le domaine $f<0$.

On obtient le contenu du théorème énoncé dans le cas particulier où $\Theta(x+yi)$ est une fonction entière de la forme

$$2.)\ \Theta(x+yi)=\varphi(x,y)+i.\psi(x,y)=(x+yi)^n-c_1(x+yi)^{n-1}+...\pm c_n$$

241|242 et | $f(x,y)=0$ est un cercle de rayon suffisamment grand pour que tous les points d'intersection des courbes $\varphi=0$ et $\psi=0$ se trouvent à l'intérieur.
Nous voulons d'abord montrer la possibilité de trouver un tel cercle. Si l'on pose $x+yi=R.e^{vi}$, alors on a :

$$\Theta(x+yi)=R^n.e^{nvi}\left\{1-c_1.\frac{e^{-vi}}{R}+c_2.\frac{e^{-2vi}}{R^2}-....\right\},$$

donc :

$$\left|c_1.\frac{e^{-vi}}{R}+c_2.\frac{e^{-2vi}}{R^2}-....\right|\leq\frac{|c_1|}{R}+\frac{|c_2|}{R^2}+...$$

et si $c_k<u$ (pour $k=1,2,...n$) : $<u.\frac{1}{R}.\frac{1-\frac{1}{R^n}}{1-\frac{1}{R}}<\frac{u}{R-1}<1$, lorsque $R$ est déterminé de sorte que : $u<R-1$, c'est-à-dire que $R>u+1$ ou $R>|c_k|+1$ $(k=1,2,...n)$. On a donc : $\Theta(R.e^{vi})=\varphi+i.\psi=R^n.e^{nvi}\alpha$, où $0<|\alpha|<1$.

Le cercle de rayon $R$ est donc tel que seulement en son intérieur (pas à l'extérieur ni sur la périphérie) les courbes $\varphi=0$ et $\psi=0$ peuvent se couper.

On sait maintenant que la caractéristique $\chi(f,\varphi,\psi)$ change seulement lorsque on a simultanément $f=0$, $\varphi=0$ et $\psi=0$. Dans notre cas, une modification de celle-ci ne peut donc en aucun cas se produire si les valeurs des coefficients $c_1,c_2...c_n$ diminuent, car le cercle $f=0$ garde encore *a fortiori* la propriété prescrite, et il ne peut jamais arriver jusqu'à une intersection commune des trois courbes. Si l'on pose donc :

$$3.)\ \varphi_0+i.\psi_0=(x+yi)^n-1=0,$$

alors on a aussi :

$$\chi(f,\varphi,\psi)=\chi(f,\varphi_0,\psi_0),$$

et comme la caractéristique $\chi(f,\varphi_0,\psi_0)$ conformément à sa signification donnée dans 1.) est manifestement égale à $n$, alors on a aussi pour chaque fonction entière 2.) :

$$\sum_{\substack{\varphi=\psi=0\\ f<0}} sgn\ 1=n,$$

242|243 c'est-à-dire que le nombre de points | d'intersection des courbes $\varphi=0$ et $\psi=0$ est aussi grand que le degré de cette fonction entière.

Le cœur de la première et la quatrième preuve de ce théorème de *Gauss* (1799 resp. 1849) n'est au fond rien de plus que la détermination de la caractéristique. Concernant la signification que *Gauss* joint au théorème, et à son équivalent algébrique, nous renvoyons aux discussions qui suivront plus tard.

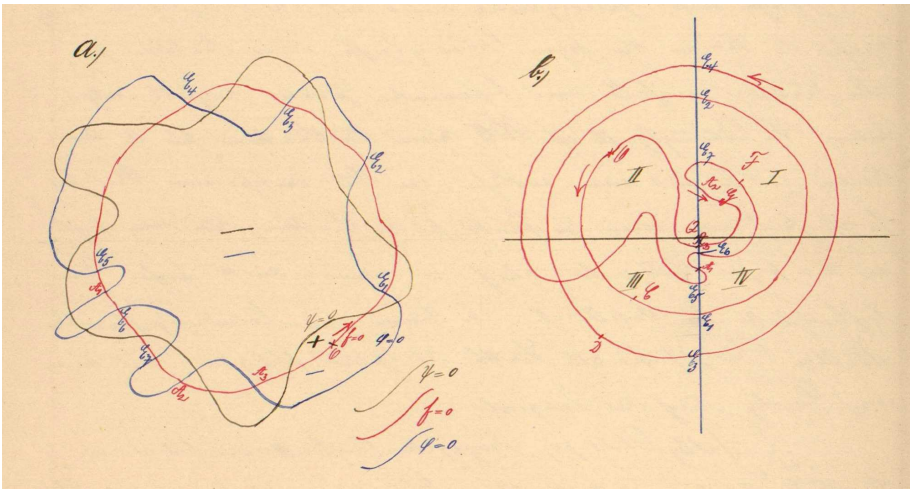
Il n'est pas plus difficile d'en dériver l'intégrale de Cauchy, laquelle donne la somme des nombres ordinaux d'une fonction pour tous les points à l'intérieur d'un domaine déterminé. Le nombre ordinal est nul lorsque la fonction a, pour les coordonnées du point, une valeur finie différente de zéro, il est égal à $+a$ lorsque le point est pour la fonction, comme on dit, un zéro d'ordre $a$, et il est égal à $-a$ lorsque le point est infiniment loin d'ordre $a$ pour celle-ci. La somme de ces nombres ordinaux de la fonction $\Theta(x+yi)$ pour tous les points du domaine $f(x,y)<0$ sera ensuite donnée

d'après *Cauchy* par l'intégrale [35]

$$\int d\log\ \Theta(x+yi)$$

(étendue à $f(x,y)=0$). Il faut tenir compte du fait que nous ne pouvons prendre ne considération les fonctions $\Theta(x+yi)$ qui prendraient des valeurs infinies pour des valeurs finies des arguments, car une telle propriété serait en contradiction avec les nécessaires conditions de continuité.

Considérons désormais la caractéristique $\chi(f,\varphi,\psi) = -\frac{1}{2}\sum_{f=\varphi=0} sgn\ d(\varphi,\psi)$




dans le cas particulier de la figure désignée par a.). (v. au verso).

Nous cheminons sur $f=0$ et nous obtenons, en ce qui concerne le produit $\varphi.\psi$, les points d'entrée $E_1, E_2, ...E_7$ et les | points de sortie $A_1, A_2, A_3$, de sorte que l'on ait 243|244

$$\chi = \frac{1}{2}(E-A) = \frac{7-3}{2} = 2.$$

On peut donner une interprétation très simple de ce nombre. Si nous représentons en effet la figure a) conformément aux équations

$$4.)\ \xi = \varphi(x,y)\ ,\ \eta = \psi(x,y),$$

alors la courbe $\varphi=0$ se transforme en l'axe $y$, $\psi=0$ en l'axe $x$ et $\varphi.\psi=0$ a dans la figure b.) seulement deux domaines intérieurs, à savoir les quadrants II et IV, et deux extérieurs, les quadrants I et III. Si l'on détermine ensuite à partir de 4.) $x$ et $y$ comme des fonctions de $\xi$ et $\eta$ et que l'on transforme $f(x,y)$ en une fonction de $\xi$ et $\eta$, alors la courbe $f(x,y)=0$ se représente dans la figure a.) d'une autre façon que dans la figure b.). Le point $O$ de la première se trouve dans b.) dans le deuxième quadrant et lorsque l'on continue à partir de ce point dans la direction de la flèche, la courbe doit franchir dans b.), conformément à son parcours dans la figure a.), d'abord l'axe des $x$, ensuite l'axe des $y$, puis de nouveau l'axe des $x$ et ainsi de suite. Les points $E_1...E_7; A_1, A_2, A_3$ se trouvent naturellement tous sur l'axe des $y$. De plus il est évident que dans la figure b.) pour tous les points d'entrée (c'est-à-dire lors du passage de I à II ou de III à IV) la direction est celle du sens contraire des aiguilles d'une montre, et pour les points de sortie (c'est-à-dire lors du passage de II à I ou de IV à III) la direction correspond au contraire à celle du sens des aiguilles d'une montre. Si l'on veut compter les tours complets de la courbe $f(x,y)=0$ autour du point zéro dans la fig. b.), alors on doit, lorsque l'on passe deux fois l'une après

35. *N.d.T.* Correction sur la page de gauche : $\frac{1}{2\pi i}\int d\log\ \Theta(x+yi)$

l'autre par un point d'entrée, compter un tour et tenir compte du fait que chaque point de sortie annule le point d'entrée précédent. En effet, lorsque par exemple je
244|245 me déplace (Fig b.) | de $\mathcal{C}$ à $E_1$ et $E_2$, alors je reviens en $\mathcal{D}$ dans le quadrant III et j'ai tourné autour du point d'origine des coordonnées une fois ; si par contre je me déplace de $\mathcal{F}$ à $E_7$ et $A_2$ après $\mathcal{G}$, alors je retourne certes aussi dans le quadrant I, mais sans avoir tourné autour du point zéro. Le nombre $\mathcal{A}$ de tours est donc :

$$5.)\ \mathcal{A} = \frac{1}{2}(E - A) = \chi(f, \varphi, \psi).$$

Nous pouvons illustrer la définition du nombre de tours de la manière suivante. Soit un clou vertical au plan fixé en $Q$ et soit posé un fil dans la position de la courbe $f(x, y) = 0$ ; Si nous attrapons alors ce dernier à un endroit quelconque et que nous le tendons, alors il fera deux fois le tour du clou.

Si nous posons, pour utiliser analytiquement cette signification de la caractéristique :

$$\xi = r.\cos v\ ,\ y = r.\sin v,$$

alors on a :

$$\chi(f, \varphi, \psi) = \frac{1}{2\pi}\int dv,$$

l'intégrale s'étendant à la courbe $f(x, y) = 0$, et comme

$$\frac{y}{\xi} = \text{tg } v\ ,\ v = \text{arctg } \frac{y}{\xi},$$

alors on a

$$6.)\ \chi(f, \varphi, \psi) = \frac{1}{2\pi}\int_{f(x,y)=0} d \text{ arctg } \frac{\psi}{\varphi} = \frac{1}{2\pi}\int_{f(x,y)=0} \frac{\varphi.d\psi - \psi.d\varphi}{\varphi^2 + \psi^2}.$$

Si l'on introduit aussi pour $x$ et $y$ des coordonnées polaires en posant :

$$x = \rho.\cos\omega\ ,\ y = \rho.\sin\omega,$$

de sorte que $\varphi(x, y) = \Phi(\omega)$, $\psi(x, y) = \Psi(\omega)$, car elle ne dépend que de la position des courbes les unes par rapport aux autres, alors on a aussi :

$$\chi(f, \varphi, \psi) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\Phi\Psi' - \Psi\Phi'}{\Phi^2 + \Psi^2} dw.$$

245|246 | Comme la caractéristique est un nombre entier, alors il est seulement nécessaire de calculer l'intégrale dans la dernière formule exactement jusqu'à la première décimale pour obtenir la valeur de la caractéristique.

L'intégrale de *Cauchy* apparaît pour le cas particulier de $\varphi$ et $\psi$ où

$$\varphi(x, y) + i.\psi(x, y) = \Theta(x + yi)$$

est une fonction de la variable complexe $x + yi$. Comme en effet :

$$-i.d\log(\varphi + \psi i) = -i\frac{d\varphi + id\psi}{\varphi + i\psi} = -i\frac{(d\varphi + id\psi)(\varphi - \psi i)}{\varphi^2 + \psi^2}$$

$$= -i.\frac{\varphi d\varphi + \psi d\psi}{\varphi^2 + \psi^2} + \frac{\varphi d\psi - \psi d\varphi}{\varphi^2 + \psi^2},$$

alors on a :

$$\frac{\varphi d\psi - \psi d\phi}{\varphi^2 + \psi^2} = -i.d\log(\varphi + \psi i) + \frac{i}{2}d\log(\varphi^2 + \psi^2)$$

et d'après 6.) :

$$\chi(f,\varphi,\psi) = \frac{1}{2\pi}\int_{f(x,y)=0} d\log\Theta(x+yi) + \frac{i}{4\pi}\int_{f(x,y)=0} d\log(\varphi^2+\psi^2)$$

Comme $d\log(\varphi^2+\psi^2)$ est une fonction univoque, alors la deuxième intégrale de la dernière équation disparaît d'après une proposition connue de *Cauchy* et il reste :

$$7.)\ \chi(f,\varphi,\psi) = \frac{1}{2\pi}\int_{f(x,y)=0} d\log\Theta(x+yi),$$

un résultat qui, d'après la signification du côté gauche de 7.), correspond à la proposition de *Cauchy*.

Si nous particularisons encore plus la formule 7.) en supposant que $\Theta(x+yi) = \varphi + i\psi$ est une fonction entière de degré $n$ et $f(x,y) = 0$ est un cercle qui contient tous les points d'intersection des courbes $\varphi = 0$ et $\psi = 0$, alors on obtient comme corollaire le théorème de *Gauss*. Soit en effet :

246|247

|

$$\Theta(x+yi) = (x+yi)^n - c_1(x+yi)^{n-1} + ... \pm c_n$$

et de là, si l'on pose [36] :

$$x + yi = R.e^{\omega i}$$

$$8.)\ \chi(f,\varphi,\psi) = \frac{1}{2\pi i}\int_0^{2\pi} \frac{n.R^{n-1}.e^{(n-1)\omega i} - (n-1)R^{n-2}.e^{(n-2)\omega i} + ...}{R^n.e^{n\omega i} - c_1R^{n-1}.e^{(n-1)\omega i} + ...}.d(R.e^{\omega i})$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \frac{n.R^n.e^{n\omega i} - (n-1)R^{n-1}.e^{(n-1)\omega i} + ...}{R^n.e^{n\omega i} - c_1R^{n-1}.e^{(n-1)\omega i} + ...}.d\omega = \frac{n}{2\pi}\int_0^{2\pi} d\omega.\frac{1 - \frac{n-1}{n.R.e^{\omega i}} + ...}{1 - \frac{c_1}{R.e^{\omega i}} + ...}$$

Si le rayon $R$ devient de plus en plus grand, alors l'expression du côté droit se rapproche de plus en plus du nombre $n$, $\chi(f,\varphi,\psi)$ reste inchangée, et comme elle doit être un nombre entier, alors nous avons démontré ici d'une seconde manière indépendante de la précédente et en toute rigueur le contenu du théorème de *Gauss*.

La notion de caractéristique et les recherches qui s'y rattachent peuvent être utilisées pour un système de $(n+1)$ fonctions à $n$ variables, sur une variété de dimension $n$. Le nombre de tours d'une spirale a déjà été exprimé par *Gauss* par une intégrale, *Tait* a abordé en détail la théorie du nombre de tours et de son calcul dans un livre.

## VI.) Clarification de la notion de caractéristique, des racines réelles et de la décomposition d'une fonction entière en fonctions réelles.

### Sur l'histoire du théorème fondamental de l'algèbre

L'histoire du théorème fondamental de l'algèbre est extrêmement instructif pour la genèse des mathématiques. On ne progresse pas en ligne droite : la connaissance

36. *N.d.T.* Correction sur la page de gauche :

$$8.)\ \chi(f,\varphi,\psi) = \frac{1}{2\pi i}\int_0^{2\pi} \frac{n.R^{n-1}.e^{(n-1)\omega i} - c_1(n-1)R^{n-2}.e^{(n-2)\omega i} + ...}{R^n.e^{n\omega i} - c_1R^{n-1}.e^{(n-1)\omega i} + ...}.d(R.e^{\omega i})$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \frac{n.R^n.e^{n\omega i} - c_1(n-1)R^{n-1}.e^{(n-1)\omega i} + ...}{R^n.e^{n\omega i} - c_1R^{n-1}.e^{(n-1)\omega i} + ...}.d\omega = \frac{n}{2\pi}\int_0^{2\pi} d\omega.\frac{1 - c_1\frac{n-1}{n.R.e^{\omega i}} + ...}{1 - \frac{c_1}{R.e^{\omega i}} + ...}$$

247|248 mathématique ramène constamment par toute sorte de tours | et de détours le chercheur à sa première source et la laisse ni plus ni moins comme une connaissance donnée de ses peu de vérités premières. L'expansion d'une toute nouvelle théorie et les nouvelles formations de concepts se rattachent à tellement de problèmes, les plus grandes espérances y ont été placées, jusqu'à ce que l'on parvienne à la conclusion que le but est atteignable aussi par les moyens simples précédents et ce maintenant avec la plus haute compréhension que le passage à la nouvelle théorie a permis. Notre problème doit avoir connu aussi un tel développement. Tandis que les mathématiciens grecs n'autorisaient comme racine d'une équation que les nombres rationnels, les Indiens ont introduit les irrationnels comme racine. Cardan considérait aussi les grandeurs imaginaires et les a introduites comme racines « impossibles » des équations.

La méthode pour résoudre les équations de deuxième, troisième et quatrième degré en représentant *explicitement* les « inconnues » comme des fonctions irrationnelles des coefficients a incité à croire que toute équation de degré quelconque d'un tel genre peut être résoluble par radicaux. On croyait être autorisé à supposer que toute fonction rationnelle entière se laisse décomposer en un produit d'autant de facteurs linéaires que ce que vaut son degré, et en effet seulement cette double sorte de représentation a pu conduire à la découverte des propriétés des racines des équations algébriques et à leurs preuves. Les grands mathématiciens, comme *Euler* et *D'Alembert* se sont bien
248|249 rendu compte que la proposition disant que toute équation algébrique de degré $n$ | a au plus $n$ racines de la forme $a+bi$ nécessite absolument une preuve. Mais ils s'étaient fait de telles fictions fantastiques concernant les propriétés des racines d'une équation algébrique que leurs preuves étaient totalement sans fondement, et la croyance que chaque équation est résoluble par radicaux était si solidement implantée qu'*Euler* a entrepris de démontrer avec l'aide de cette proposition fausse la proposition vraie précédente.

Ici, c'est encore *Gauss* qui a fourni la solution. Dans sa fameuse dissertation (*Demonstratio nova theorematis 1799*) il a découvert les défauts des anciennes preuves, bousculé les croyances à propos de la résolution d'une équation générale par radicaux et donné une nouvelle preuve pour ce théorème, preuve qui a encore été suivie plus tard par trois autres (1815, 1816, 1849). Ces preuves ont certes été d'une grande importance historique, et en particulier la seconde qui aura été d'une signification fondamentale pour la théorie des équations algébriques. Mais une chose reste non élucidée : la nature de la racine réelle. *Gauss* s'en tient à illustrer géométriquement sa signification et il avait la ferme conviction d'avoir donné ainsi une définition suffisante. Ce qu'il voulait prouver et ce qu'il a rigoureusement démontré a été donné dans une partie précédente avec précision. On aurait ainsi pu s'en contenter comme preuve pour le contenu purement analytique du théorème fondamental - que l'on énonce communément ainsi : toute équation de degré $n$ a $n$ racines - si l'on avait voulu « appliquer d'emblée la continuité géométrique sur des grandeurs algébriques, et donc,
249|250 pour l'esquisser | rapidement, faire correspondre à chaque distance entre deux points un nombre arithmétique » [37]. Tous ces mathématiciens utiliseront cela avec d'autant peu de scrupules qu'ils considéreront aussi les grandeurs irrationnelles et imaginaires comme « définies » par des quantités géométriques. Que Gauss ait aussi été de ce dernier avis ressort le plus clairement d'une remarque qu'il a faite dans une leçon à la société de Göttingen. (imprimé dans le Göttinger gelehrten Anzeigen de mars 1831, 64. Werke V) [38]. Il cherche, à la fin de ses essais, à donner aux grandeurs imaginaires un fondement solide : « Que l'on puisse la preuve de la possibilité de donner une si-

37. Traduction en allemand de Netto des preuves de Gauss.

On doit comprendre par « nombres arithmétiques » des grandeurs sur lesquelles on peut appliquer les opérations arithmétiques de l'addition, la multiplication, la division, etc. et pour lesquelles la proposition suivante reste valable : si un produit est nul, alors l'un des facteurs doit toujours être nul.

38. *N.d.T.* Il s'agit en fait des *Göttingische gelehrte Anzeigen* du 23 avril 1831

gnification intuitive à $\sqrt{-1}$ est donc ici complètement justifié, et il n'en faut pas plus pour autoriser ces grandeurs dans le domaine de l'arithmétique. » Les mathématiciens qui font leur ce point de vue qui nous parait douteux ne tiennent aucun compte du fait que certes les rapports arithmétiques sont transférables et applicables aux concepts géométriques intuitifs et que certes la géométrie a contribué et contribuera toujours aux progrès des sciences arithmétiques, mais que jamais un pur concept « arithmétique » ne pourra et ne devra être défini par une illustration géométrique.

Ainsi nombreux sont ceux qui pensent avoir par exemple suffisamment fixé la notion de racine carrée de 2 comme un nombre véritablement arithmétique lorsqu'ils disent qu'elle correspond à la longueur de la diagonale d'un carré dont le côté est égal à l'unité. Ils oublient à cette occasion qu'ici aussi ils ont déjà supposé qu'à tout point de la ligne graduée avec une unité arbitrairement choisie correspond à un nombre arithmétique. Et pour montrer la justesse de cette dernière supposition, cette définition géométrique des irrationnels leur est à nouveaux absolument nécessaire. Mais alors comment au juste un segment de droite géométrique peut « définir » un concept arithmétique ? Car la raison se conforme-t-elle aux sens, ou n'est-ce pas plutôt la raison qui tend à travers le monde sensible des fils qui permettent de le mettre en ordre ?

Non, $\sqrt{2}$ sera uniquement défini par l'équation $x^2 - 2 = 0$, et lorsque l'on dit que $\sqrt{2}$ est un nombre arithmétique - nombre qui ne peut certes pas être déterminé totalement précisément par des nombres rationnels, mais qui peut être approché aussi près que l'on veut - alors cela ne signifie rien d'autre que le fait qu'il y a un nombre $\delta$ arbitrairement petit pour lequel l'équation

$$x^2 - 2 = \pm\delta$$

est satisfaite par des nombres rationnels. Cela nous ramène à nos remarques sur la preuve de l'existence des racines. Si nous voulons « appréhender et fonder de façon strictement arithmétique le soi-disant théorème fondamental de l'algèbre » sans aucun soutien géométrique, alors nous pouvons prétendre que le véritable théorème fondamental, à savoir que chaque fonction entière rationnelle de degré $n$ est décomposable en un produit de $n$ facteurs linéaires réels ou imaginaires, n'est pas du tout correct. Ce que nous pouvons démontrer | et ce que nous démontrerons, et ce qu'en substance 251|252
Gauß a seulement démontré, c'est que à partir de chaque fonction rationnelle entière de degré $n$ on peut en fabriquer une autre aussi petite que l'on veut par une variation du système de coefficient de la première, cette dernière pouvant être découpée en un produit de $m$ facteurs linéaires et $\frac{n-m}{2}$ facteurs quadratiques et ce de sorte que les coefficients de tous les facteurs soient des nombres rationnels. Ce nombre $m$ peut être donné pour chaque fonction avec des coefficients déterminés [39].

On s'aperçoit pourtant qu'une fonction encore si voisine, pour ce qui concerne les propriétés algébriques, sera en général différente *toto genere* de celle présentée initialement, et que donc une telle décomposition ne promet vraiment aucun avantage pour des recherches de théorie algébrique. Mais ce n'est qu'un aspect de la version arithmétique du théorème fondamental. Pour les recherches algébriques nous allons prouver la possibilité d'une deuxième décomposition fondamentalement différente de celle donnée, qui n'est pas valable au sens de l'égalité mais seulement en tant que congruence pour un Primmodulsystem [40]. Nous allons en effet montrer [41] la validité de la congruence suivante :

$$F(x) = x^n - c_1x^{n-1} + c_2x^{n-2} - ... \pm c_n \equiv (x - x_1)(x - x_2)...(x - x_n)$$

39. Il est pourtant aussi encore permis lors des recherches suivantes d'employer par souci de brièveté l'expression habituelle concernant les racines réelles.

40. *N.d.T.* Nous avons choisi de laisser ici le terme allemand, qui n'a pas d'équivalent usuel en français.

41. V. 333 ss.

$$(modd.\ f_1 - c_1, f_2 - c_2, ..., f_n - c_n, y - c),$$

lorsque

$$(x - x_1)(x - x_2)...(x - x_n) = x^n - f_1 x^{n-1} + f_2 x^{n-2} - ... \pm f_n;$$

$x_1...x_n$ sont des variables indéterminées, $y$ une certaine variable dépendant de la
252|253 qualité algébrique de la fonction $F(x)$, qui n'est pas | une fonction symétrique de $x_1, x_2...x_n$, et $c$ est un nombre rationnel qui est pour $y$ dans le même rapport que $f_1, f_2, ...f_n$ pour resp. $c_1, c_2...c_n$. C'est seulement lorsque l'on sépare ainsi en toute rigueur la partie arithmétique et algébrique du théorème fondamental, qu'il est possible d'appréhender véritablement précisément son contenu.

26. Jan. La notion de caractéristique.

Nous voulons tout d'abord rendre la notion de caractéristique indépendante de l'hypothèse des racines réelles des fonctions entières.

On a, d'après notre explication précédente :

$$1.)\ \chi(V(x), W(x)) = -\frac{1}{2}\sum\ sgn\ \lim_{h=0}\begin{vmatrix} V(x) & W(x) \\ \dfrac{V(x+h)-V(x)}{h} & \dfrac{W(x+h)-W(x)}{h} \end{vmatrix}$$

$$= \frac{1}{2}\sum_{\xi}\ sgn\ V'(\xi)W(\xi),$$

les sommes s'étendent à toutes les valeurs $\xi$ pour lesquelles $V(\xi) = 0$, ou aussi la première somme à toutes les valeurs $\eta$ pour lesquelles $W(x)$ est nulle. Supposons que $V(x)$ et $W(x)$ ont des fonctions de même degré, et que ni $V(x)$ et $W(x)$, ni $V(x)$ et $V'(x)$, ni $W(x)$ et $W'(x)$ ne possèdent un facteur commun. De l'équation 1.) on obtient :

$$2.)\ \chi(V(x), W(x)) = -\frac{1}{2}\sum_{(\xi)}\ sgn\ \lim_{h=0}\frac{1}{h}\begin{vmatrix} V(x) & W(x) \\ V(x+h) & W(x+h) \end{vmatrix}$$

$$2^a.)\ = \frac{1}{4}\sum_{(\xi),(\eta)}\ sgn\ \lim_{h=0}\frac{1}{h}\left(W(x).V(x+h) - V(x).W(x+h)\right),$$

lorsque la somme dans 2.) s'étend à toutes les valeurs $\xi$, celle dans $2^a$.) non seulement aux valeurs $\xi$ mais aussi aux valeurs $\eta$.

253|254 | Nous pouvons maintenant supprimer l'utilisation de la notion de racine réelle dans cette dernière expression de la caractéristique en n'étendant pas la somme dans 2.) aux valeurs $\xi$ pour lesquelles $V(\xi) = 0$, mais à toutes les valeurs rationnelles de $x$ pour lesquelles $V(x)$ et $V(x+h)$ ont des signes opposés, et la somme dans $2^a$) à toutes les valeurs rationnelles de $x$ pour lesquelles non seulement [42] $V(x).V(x+h) < 0$ mais aussi [43] $W(x).W(x+h) < 0$. On a donc par exemple pour toutes les valeurs $x$ sur lesquelles on doit sommer dans la formule 2.) :

$sgn\ V(x+h) = sgn\ V'(\xi)$, et comme $sgn\ W(x) = sgn\ \lim W(x+h)$, alors on a aussi :

$$sgn\ \lim\begin{vmatrix} V(x) & W(x) \\ V(x+h) & W(x+h) \end{vmatrix} = sgn\ \lim\left(V(x).W(x+h) - V(x+h).W(x)\right)$$

42. *N.d.T.* Corrigé dans la page de gauche par *soit*
43. *N.d.T.* Corrigé dans la page de gauche par *soit*

$$= \ sgn \ \lim V(x+h).W(x) = \ sgn \ V'(\xi).W(\xi).$$

Nous supprimons le passage à la limite dans les expressions 2) et $2^a$) de la manière suivante.

Nous rendons l'intervalle $h$ si petit que jamais on ne peut trouver ensemble deux valeurs quelconques de $x$ pour lesquelles $V(x).V(x+h) < 0$ ou $W(x).W(x+h) < 0$ dans l'intervalle. Ensuite on a :

$$\frac{-\ sgn \ V(x).W(x+h) + \ sgn \ W(x).V(x+h)}{2}$$

$$= \ sgn \ (-V(x).W(x+h) + W(x).V(x+h))$$

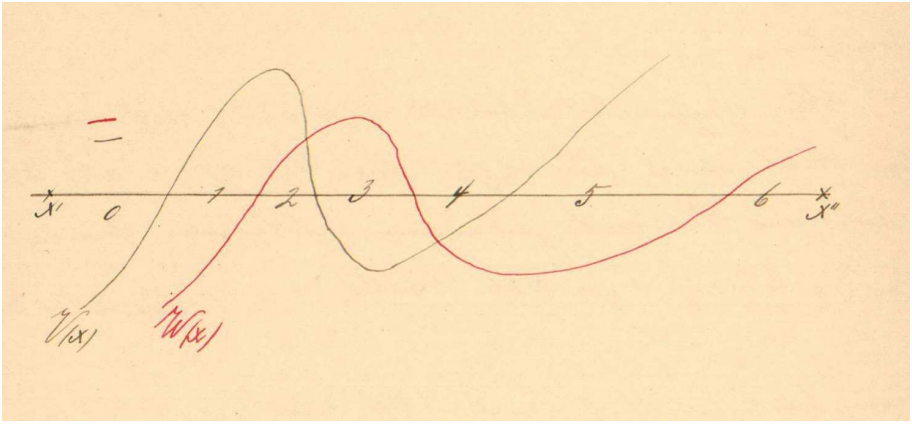
pour chaque valeur quelconque de $x$ [44]. On aura donc :

$$3.)\ \chi(V(x), W(x)) = -\frac{1}{8}\sum [\ sgn \ V(x).W(x+h) - \ sgn \ W(x).V(x+h)]\,,$$

où $h$ a une valeur donnée et où la somme s'étendra non seulement à toutes les valeurs de $x$ pour lesquelles $V(x).V(x+h) < 0$ mais aussi pour toutes celles pour lesquelles $W(x).W(x+h) < 0$. Nous pourrons aussi écrire cette équation 3.), | lorsque l'on met 254|255
$x_k$ à la place de $x$ et $x_{k+1}$ à la place de $x+h$, sous la forme :

$$4.)\ -8\chi(V(x), W(x)) = \sum_{(k)} \begin{vmatrix} sgn \ V(x_k) & sgn \ V(x_{k+1}) \\ sgn \ W(x_k) & sgn \ W(x_{k+1}) \end{vmatrix}$$

en étendant la somme comme auparavant non seulement à toutes les valeurs de $k$ pour lesquelles $V(x_k).V(x_{k+1}) < 0$ mais aussi à toutes celles pour lesquelles $W(x_k).W(x_{k+1}) < 0$. Mais, comme nous le savons, on peut délimiter un intervalle fini $x'...x''$ dans lequel se trouvent toutes les valeurs $x_k$ et $x_{k+1}$ (c'est-à-dire dans le langage habituel : dans lequel se trouvent toutes les racines de $V(x) = 0$ et $W(x) = 0$), de sorte que nous pouvons aussi étendre la somme dans 4.) à toutes les valeurs $x_k$ qui se forment lorsque nous subdivisons l'intervalle entier $x'...x''$ en des intervalles plus petits de la grandeur donnée $h$. Les déterminants qui s'ajoutent alors dans cette somme sont tous nuls à cause de l'égalité de leurs colonnes.




Ainsi par exemple la caractéristique des deux fonctions de degré trois qui correspondent aux courbes construites ci-contre est égale à $\frac{-3}{2}$, comme le montre le tableau ci-joint :

44. *N.d.T.* Ajout dans la page de gauche : « pour qui $V(x).V(x+h) < 0$ ou $W(x)W(x+h) < 0$ et $= 0$ pour toutes les autres valeurs de $x$. »

$$\begin{array}{lccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ V(x) & - & + & + & - & - & + & + \\ W(x) & - & - & + & + & - & - & + \end{array} ;$$

$$-8\chi(V,W) = \begin{vmatrix} -1 & +1 \\ -1 & -1 \end{vmatrix} + \begin{vmatrix} +1 & +1 \\ -1 & +1 \end{vmatrix} + \ldots$$
$$= +2+2+2+2+2+2$$
$$=12$$

Nous avons maintenant à montrer que la grandeur $h$ de l'intervalle se laisse déterminer selon nos exigences de sorte que jamais deux valeurs $x_k$ ne puissent se trouver
255|256 dans un seul et même intervalle, | ou en d'autres termes, que deux racines de l'équation $V(x) = 0$, deux racines de l'équation $W(x) = 0$, ou bien qu'une racine de l'équation $V(x) = 0$ et une de l'équation $W(x) = 0$ se trouvent dans le même intervalle. Nous pouvons cependant réduire ces conditions sur la grandeur $h$ à une seule disant que les deux équations $V(x) = 0$ et $W(x) = 0$ ne possèdent pas ensemble une racine dans le même intervalle. Ainsi, si par exemple $V(x) = 0$, qui ne contient pas de racine de l'équation $W(x) = 0$ dans un intervalle $a...b$, contient plus d'une racine, alors les déterminants qui correspondent à deux racines successives s'annuleront mutuellement toujours dans l'équation 4.). Si par exemple dans un intervalle $W(x)$ est constante et positive et si $V(x)$ passe en un point des positifs aux négatifs, puis de nouveau des négatifs aux positifs, alors les déterminants correspondant s'annulent l'un l'autre :

$$\begin{vmatrix} + & - \\ + & + \end{vmatrix} + \begin{vmatrix} - & + \\ + & + \end{vmatrix} = 0$$

Si le nombre de racines de $V(x) = 0$ est impair, alors il nous reste un seul déterminant qui, comme il est facile de le comprendre, a la même valeur que le déterminant :

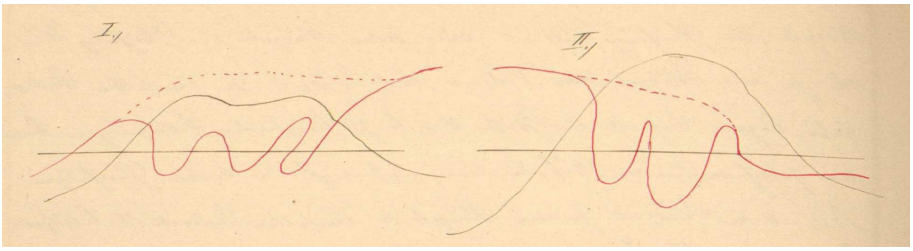
$$5.) \quad \begin{vmatrix} sgn\ V(a) & sgn\ V(b) \\ sgn\ W(a) & sgn\ W(b) \end{vmatrix} ;$$

Si le nombre est pair, alors la somme de tous les déterminants est nul et le déterminant 5.) a la même valeur, car ainsi

$$sgn\ V(a) = \ sgn\ V(b)\ ;\ sgn\ W(a) = \ sgn\ W(b).$$

On peut de là dans chacun des deux cas remplacer le nombre entier d'intervalles
256|257 par un intervalle unique $a...b$, sans faire entrer en ligne de compte | la justesse de la formule 4.) et par exemple dans les figures $I)$ et $II)$ on peut remplacer le parcours de la courbe initiale par la courbe en pointillés.

Nous voulons désormais déterminer la grandeur requise des intervalles. Il est impératif que dans notre démonstration nous épuisions l'hypothèse disant que les deux courbes $V(x) = 0$ et $W(x) = 0$ ne coupent pas l'axe en un même point, car sans cette hypothèse il est impossible de remplir les conditions que nous avons associées à la grandeur des intervalles.



D'après l'équation 3.) p. 204 le résultant est :

$$6.)\ R(V,W) = P(x).W(x) + Q(x).V(x)$$

à savoir dans notre cas que $R(V,W) \neq 0$. Si $V(x)$ et $W(x)$ sont respectivement de degrés $n$ et $n'$ (car pour la conclusion suivante il n'est pas nécessaire de se restreindre au cas où $V(x)$ et $W(x)$ sont de même degré), alors nous pouvons écrire $P(x)$ et $Q(x)$ sous la forme :

$$P(x) = p_0 + p_1 x + p_2 x^2 + ... + p_n x^{n-1}$$
$$Q(x) = q_0 + q_1 x + q_2 x^2 + ... + q_{n'} x^{n'-1}$$

Si maintenant les extrémités de l'intervalle sont données par les nombres rationnels $-r$ et $+r$, à l'intérieur duquel se trouvent toutes les racines de $V(x) = 0$ et $W(x) = 0$, alors pour chaque valeur de $x$ appartenant à cet intervalle [45] :

$$\begin{aligned} |P(x)| &< |p_0| + |p_1|.r + |p_2|.r^2 + ... + |p_{n-1}|.r^{n-1} \\ &< |p_k|.(1 + r + r^2 + ... + r^{n-1}) \\ &< p_k.\frac{r^n - 1}{r - 1}, \end{aligned}$$

lorsque $|p_k| > |p_i|$ $(i = 0, 1, 2, ...k-1; k+1...n-1)$.

| Et on a aussi : 257|258

$$\frac{P(x)}{|R|} = (s_1 - 1)\varphi(x),$$

où $s_1$ est un nombre entier déterminé et $\varphi(x)$ une fonction qui est plus petite que 1 pour chacune de ces valeurs de $x$ ; alors le quotient du côté gauche reste également au dessous d'une limite fixée.

De plus nous pouvons poser :

$$Q(x) = (s_2 - 1)|R|.\psi(x),$$

où $s_2$ est également un nombre entier déterminé et $\psi(x)$ est plus petite que 1 pour chacune de ces valeurs de $x$, et troisièmement :

$$P(x).\frac{W(x+\sigma) - W(x)}{\sigma} = (s-1)|R|.\chi(x),$$

étant donné que l'expression du côté gauche est une fonction entière car $x$ et $\sigma$ le sont, et donc on doit aussi avoir une limite supérieure fixée pour les valeurs de $\sigma$ pour lesquelles $|\sigma| < 1$.

Lorsque l'on pose pour $s$ un nombre entier fini qui est un nombre entier plus grand que $s_1, s_2$ et $s_3$ alors on peut poser :

$$7.)\ P(x) = (s-1)|R|.\varphi(x)$$
$$8.)\ Q(x) = (s-1)|R|.\psi(x)$$

45. *N.d.T.* Correction dans la page de gauche :

$$\begin{aligned} |P(x)| &< |p_0| + |p_1|.r + |p_2|.r^2 + ... + |p_{n-1}|.r^{n-1} \\ &\leq |p_k|.(1 + r + r^2 + ... + r^{n-1}) \\ &= |p_k|.\frac{r^n - 1}{r - 1}, \end{aligned}$$

lorsque $|p_k| \geq |p_i|$ $(i = 0, 1, 2, ...k-1; k+1...n-1)$.

$$7.)\ P(x)\frac{W(x+\sigma)-W(x)}{\sigma} = (s-1)|R|.\chi(x),$$

et les fonctions $\varphi(x), \psi(x), \chi(x)$ ont pour propriété d'avoir pour toutes les valeurs de $x$ qui sont dans l'intervalle $-r.... + r$ une valeur plus petite que 1. De l'égalité 6.) il résulte [46] :

$$10.)\ R(V,W) = P(x_1).(W(x_1) - W(x_2)) + P(x_1).W(x_2) + Q(x_1)V(x_1)$$

258|259 et sous l'hypothèse que $|x_2 - x_1| < 1$, de 9.) on obtient : |

$$11.)\ P(x_1).\frac{W(x_2)-W(x_1)}{x_2-x_1} = (s-1).|R|.\chi(x_1).$$

Avec l'utilisation des équations 7.), 8.) et 11.) il découle de 10.) :

$$12.)\ |R| = |(s-1).|R|.(x_1-x_2).\chi(x_1)+(s-1).|R|.\varphi(x_1).W(x_2)+(s-1).|R|.\psi(x_1).V(x_1)|$$

et après la division des deux côtés par $(s-1).|R|$ :

$$\frac{1}{s-1} = |(x_1-x_2)\chi(x_1) + \varphi(x_1).W(x_2) + \psi(x_1).V(x_1)|$$

donc en raison de la propriété des fonctions $\varphi(x), \psi(x)$ et $\chi(x)$ :

$$\frac{1}{s-1} < |x_1-x_2| + |V(x_1)| + |W(x_2)|$$

On a de plus $|x_2 - x_1| \leq \frac{1}{s}$, il suit alors par l'introduction de ces valeurs dans la dernière équation :

$$13.)\ -\frac{1}{s} + \frac{1}{s-1} = \frac{1}{s(s-1)} < |V(x_1)| + |W(x_2)|$$

Nous avons ainsi résolu notre problème. Nous avons en effet rendu $h \leq \frac{1}{s}$, de sorte que sûrement jamais une racine de $V(x) = 0$ et une racine de $W(x) = 0$ ne pourront se trouver dans un même intervalle, étant donné que si $V(x_1)$ et $W(x_2)$ sont les valeurs de fonction pour deux points $x_1$ et $x_2$ de l'intervalle, la somme de leurs valeurs absolues doit toujours excéder une grandeur finie déterminée (Eq. 13.)).

Après l'application de cette détermination nous pouvons dire que nous avons obtenu dans la formule 4.) une expression de la caractéristique qui se laisse déterminer de façon strictement arithmétique, sans l'utilisation de racines réelles et que nous évitons le calcul de la caractéristique par la proposition de *Sturm* généralisée par *Sylvester*.

En particulier si nous découpons l'intervalle pris en considération $-r...+r$ en $2s.r$
259|260 parties égales de taille $\frac{1}{s}$ et | que nous désignons les points frontières par :

$$x_0 = -r\ ,\ x_1 = -r + \frac{1}{s}\ ,\ x_2 = -r + \frac{2}{s}\ ,\ ...\ ,\ x_{2s.r-1} = -r + \frac{2s.r-1}{s}\ ,\ x_{2s.r} = +r,$$

alors on a :

$$14.)\ -8\chi(V(x),W(x)) = \sum_{k=0}^{k=2s.r} \begin{vmatrix} sgn\ V(-r+\frac{k}{s}) & sgn\ V(-r+\frac{k+1}{s}) \\ sgn\ W(-r+\frac{k}{s}) & sgn\ W(-r+\frac{k+1}{s}) \end{vmatrix}$$

Toutefois, il convient de noter ce qui suit pour cette représentation. Il peut se produire

46. Ajout dans la page de gauche : « comme le côté gauche est indépendant de $x$, »

lors de cette partition des intervalles ainsi induite que nous tombions par hasard en un point frontière sur une racine rationnelle de l'équation $V(x) = 0$ ou $W(x) = 0$, de sorte que par exemple $V(-r + \frac{k}{s}) = 0$ : dans un tel cas, il semble que la formule 14.) perde sa validité, comme d'après la dérivation elle-même (v. Formule 4.)), on doit avoir pour toute valeur de $k$ :

$$sgn\ V(-r + \frac{k}{s}).V(-r + \frac{k+1}{s} = \pm 1.$$

Mais en réalité dans ce cas aussi la validité de la formule est conservée, car lorsque $V(-r + \frac{h}{s}) = 0$, alors on a seulement besoin de se représenter les deux intervalles $-r+\frac{h-1}{s}....-r+\frac{h}{s}$ et $-r+\frac{h}{s}....-r+\frac{h+1}{s}$ réunis en un seul $-r+\frac{h-1}{s}....-r+\frac{h+1}{s}$, pour faire concorder de nouveau la formule 14.) avec l'équation 4.). D'après les conditions de l'équation 13.) la distance d'une racine de $V(x) = 0$ à une racine de $W(x) = 0$ doit en effet être plus grande que $\frac{1}{s}$ et ainsi dans notre intervalle double, à part la racine $(-r + \frac{h}{s})$ de $V(x) = 0$, on ne peut en trouver aucune de $W(x) = 0$, pas même sur les bords de l'intervalle.

Nous voulons désormais utiliser ces recherches pour déterminer, ce qui a déjà été calculé plus haut p. 213, la différence de deux caractéristiques : | 260|261

$$15.)\ \chi(f(x), (x'' - x)f_1(x)) - \chi(f(x), (x' - x).f_1(x)) = \sum_{x'<\xi<x''} sgn\ f_1(\xi).f'(\xi),$$

mais maintenant sans l'utilisation de racines réelles, en supposant dans les recherches suivantes, ce que l'on peut toujours obtenir, que $f_1(x)$ soit une fonction entière dont le degré soit plus petit de 1 par rapport à celui de $f(x)$.

Nous avons tout d'abord à déduire l'égalité :

$$\begin{aligned} 16.)\ &\frac{1}{2}\sum_{(\xi)} sgn\ f_1(\xi).f'(\xi) + \frac{1}{2}\sum_{(\eta)} sgn\ f(\eta).f_1'(\eta) \\ &- \frac{1}{4}\ sgn\ f(x'').f_1(x'') + \frac{1}{4}\ sgn\ f(x').f_1(x') = 0 \end{aligned}$$

où $\xi$ désigne les racines de $f(x) = 0$ et $y$ celles de $f_1(x) = 0$ [47]. Posons :

$$\begin{vmatrix} f(x) & (x'' - x).f_1(x) \\ f'(x) & (x'' - x).f_1'(x) - f_1(x) \end{vmatrix} = \Delta(x)$$

alors on a :

$$17.)\ -\frac{1}{2}\sum\ sgn\ \Delta(\xi) = \frac{1}{2}\ sgn\ (x'' - \xi).f_1(\xi).f'(\xi)$$

et

$$18.)\ -\frac{1}{2}\sum\ sgn\ \Delta(\eta) = -\frac{1}{2}\ sgn\ (x'' - \eta).f(\eta).f_1'(\xi) + \frac{1}{2}\ sgn\ f(x'').f_1(x'')$$

et soit de plus :

$$\begin{vmatrix} f(x) & (x' - x).f_1(x) \\ f'(x) & (x' - x).f_1'(x) - f_1(x) \end{vmatrix} = \Delta_1(x)$$

alors on aura :

47. Ajout dans la page de gauche : « entre x' et x" »

$$19.)\ -\frac{1}{2}\sum\ sgn\ \Delta_1(\xi) = \frac{1}{2}\ sgn\ (x'-\xi).f_1(\xi).f'(\xi)$$

et

$$20.)\ -\frac{1}{2}\sum\ sgn\ \Delta_1(\eta) = -\frac{1}{2}\ sgn\ (x'-\eta).f(\eta).f_1'(\xi) + \frac{1}{2}\ sgn\ f(x').f_1(x')$$

comme maintenant $x'-\xi \neq -(x''-\xi) < 0$,[48] alors on a :

$$\begin{aligned}\frac{1}{2}(17.) - 19.) - 18.) + 20.)) &= \frac{1}{2}\sum\ sgn\ f_1(\xi).f'(\xi) + \frac{1}{2}\sum\ sgn\ f(\eta).f_1'(\eta)\\ &= -\frac{1}{4}\ sgn\ f(x'').f_1(x'') + \frac{1}{4}\ sgn\ f(x').f_1(x') = 0\end{aligned}$$

Si nous supposons maintenant calculé le nombre entier $s$ pour les fonctions $f(x)$
261|262 et $f_1(x)$ et ensuite fixés les points frontières $x_0, x_1...$ pour l'intervalle $x'..x''$ comme | dans la formule 14.) pour l'intervalle $-r...+r$, de sorte que la somme algébrique des signes forme

$$\begin{aligned}21.)\ &-\frac{1}{4}\sum_{(k)}\begin{vmatrix} sgn\ f(x_k) & sgn\ f(x_{k+1}) \\ sgn\ f_1(x_k) & sgn\ f_1(x_{k+1})\end{vmatrix}\\ &+\frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}\ sgn\ f(x').f_1(x')\end{aligned}$$

alors nous pouvons aussi écrire ces valeurs sous la forme

$$\begin{aligned}22.)\ &\frac{1}{2}\sum_{x'<\xi<x''}\ sgn\ f_1(\xi).f'(\xi) + \frac{1}{2}\sum_{x'<\eta<x''}\ sgn\ f(\eta).f_1'(\eta)\\ &+\frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}\ sgn\ f(x').f_1(x')\end{aligned}$$

comme cela découle facilement de notre précédente discussion (V. 253 4)). Cette dernière expression n'est rien d'autre que $\sum_{x'<\xi<x''}\ sgn\ f_1(\xi)f'(\xi)$, comme on le voit tout de suite si l'on additionne à 22.) l'identité 16.). De plus 21.) est la valeur voulue de la différence de caractéristique 15.)

La somme de signes[49]

$$23.)\ -\frac{1}{2}\sum\begin{vmatrix} sgn\ f(x_k) & sgn\ f(x_{k+1}) \\ sgn\ f_1(x_k) & sgn\ f_1(x_{k+1})\end{vmatrix} + \frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}f(x').f_1(x'),$$

lorsque la somme s'étend seulement aux intervalles $x_k...x_{k+1}$ pour lesquels $f(x_k).f(x_{k+1}) < 0$, pendant que le signe de $f_1(x)$ reste constant dans l'intervalle et ses extrémités, est également égale à l'expression 22.), comme dans celle-ci :

$$\frac{1}{2}\sum_{x'<\xi<x''}\ sgn\ f_1(\xi).f'(\xi) = -\frac{1}{2}\sum_{x'<\eta<x''}\ sgn\ f_1(\eta).f'(\eta).$$

48. Ajout dans la page de gauche : « et également $x'-\eta$ et $-(x'-x) < 0$ »
49. *N.d.T.* Correction dans la page de gauche :

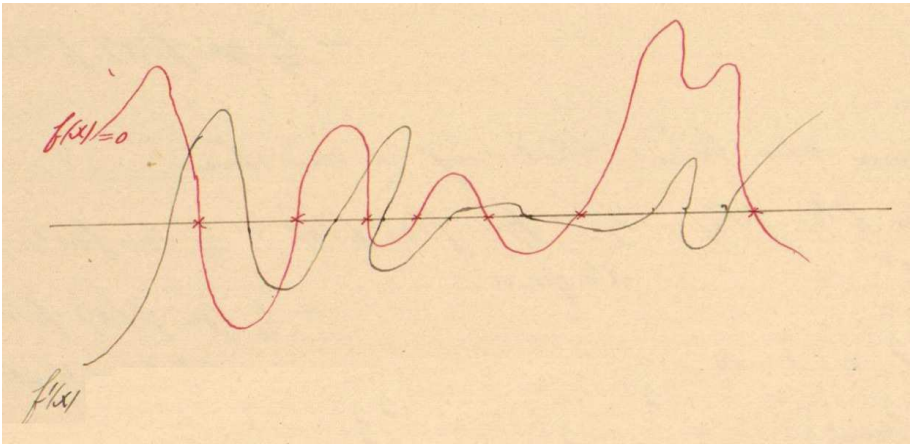
$$23.)\ -\frac{1}{4}\sum\begin{vmatrix} sgn\ f(x_k) & sgn\ f(x_{k+1}) \\ sgn\ f_1(x_k) & sgn\ f_1(x_{k+1})\end{vmatrix} + \frac{1}{4}\ sgn\ f(x'').f_1(x'') - \frac{1}{4}f(x').f_1(x')$$

Nous pouvons exprimer cette égalité sous la forme [50] :

$$24.)\ \chi(f(x),(x''-x).f_1(x))-\chi(f(x),(x'-x)f_1(x))$$
$$=\frac{1}{2}\sum_{f(x_k).f(x_{k+1})<0} sgn\ f_1(x_k).(\ sgn\ f(x_{k+1})-\ sgn\ f(x_k))$$
$$+\frac{1}{4}\ sgn\ f(x'').f_1(x'')-\frac{1}{4}\ sgn\ f(x')f_1(x').$$

262|263

| Pour le cas particulier où $f_1(x) = f'(x)$ nous obtenons dans 23.) et 24.) le nombre qui exprime combien il y a de « racines réelles » entre $x'$ et $x''$.




On observe dans le dessin de la figure que
1) $f_1$ est aussi une courbe parabolique
2) les zéros de $f'$ se trouvent entre le Max et le Min de $f$
3) Le Max et le Min de $f'$ se trouvent entre les points d'inflexion de $f$.

Mais ce nombre se laisse calculer d'une façon encore plus simple. Si l'on détermine en effet pour le système de fonctions $(f(x), f'(x))$ le nombre $s$ et que l'on découpe ensuite l'intervalle $x'...x''$ de la manière connue en intervalles plus petits, de sorte que pour ce système de courbe précis $f(x) = 0, f'(x) = 0$ la circonstance particulière suivante survient : la courbe $f(x) = 0$ coupe plus d'une fois l'axe des abscisses dans un petit intervalle, étant donné qu'entre deux tels points d'intersection il doit toujours se trouver au moins un point d'intersection de la courbe $f'(x) = 0$, si l'on suppose que le discriminant des deux équations $f(x) = 0$ et $f'(x) = 0$ ne s'annule pas. De là nous pouvons facilement identifier le nombre de racines réelles qui se trouvent dans l'intervalle $x'...x''$ à la somme de signes :

$$\frac{1}{2}\sum_{(k)}|\ sgn\ f(x_k)-\ sgn\ f(x_{k+1}|$$

50. *N.d.T.* Correction dans la page de gauche :

$$24.)\ \chi(f(x),(x''-x).f_1(x))-\chi(f(x),(x'-x)f_1(x))$$
$$=-\frac{1}{4}\sum_{f(x_k).f(x_{k+1})<0} sgn\ f_1(x_k).(\ sgn\ f(x_{k+1})-\ sgn\ f(x_k))$$
$$+\frac{1}{4}\ sgn\ f(x'').f_1(x'')-\frac{1}{4}\ sgn\ f(x')f_1(x').$$

où la somme est à étendre à toutes les valeurs de $x_k$ qui sont les points frontières de cet intervalle.

29. Jan. La nature des racines réelles et la factorisation en fonctions entières.

Nous voulons maintenant encore imposer à notre nombre entier $s$, en plus des conditions auxquelles nous l'avons déjà soumis (p. 258 éq. 7.) 8.) 9.) ) dans le cas
263|264 particulier du système de fonctions $(f(x), f'(x))$ à coefficients entiers, que : |

$$\left|\frac{f(x_2)-f(x_1)}{x_2-x_1}\right| < (s-1).|D|,$$

si $x_2 - x_1 < \frac{1}{s}$ ; $x_2 > x_1$, ou :

$$1.)\ |f(x_2)-f(x_1)| < (x_2-x_1)(s-1).|D|.$$

Si nous découpons maintenant chacun des intervalles de grandeur $\frac{1}{s}$ à nouveau en $r.|D|$ parties égales, ($D$ est un nombre entier), alors, lorsque $x'$ et $x''$ sont des valeurs rationnelles qui appartiennent à un seul et même de ces sous-intervalles, on a :

$$x''-x' \leq \frac{1}{r.s.|D|},$$

sous l'hypothèse que $x'' > x'$ ; et d'après l'égalité 1.) :

$$2.)\ |f(x'')-f(x')| < \frac{1}{r},$$

où $r$ est un nombre entier complètement quelconque.

Si l'on choisit l'un des intervalles initiaux de grandeur $\frac{1}{s}$ où la fonction a des signes opposés aux deux extrémités de celui-ci, alors il y aura aussi l'un de ses sous-intervalles qui aura la même propriété. Si $x'$ et $x''$ sont les extrémités rationnelles de cet intervalle, alors on doit avoir

$$|f(x')| < \frac{1}{r} \text{ et } |f(x'')| < \frac{1}{r},$$

étant donné que d'après 2.) même la somme de ces valeurs absolues est plus petite que $\frac{1}{r}$.

La même inégalité est valide pour toutes les valeurs rationnelles qui se trouvent à l'intérieur de l'intervalle. De cette façon nous obtenons donc à la place d'une racine réelle $\xi$ de $f(x) = 0$ un intervalle $x'...x''$ tel que

$$a.)\ f(x').f(x'') < 0 \text{ et } b.)\ |f(x'+(x''-x')\delta)| < \frac{1}{r},$$

où $0 \leq \delta \leq 1$ et $x', x''$ sont des nombres rationnels et déterminables.

Combien y a-t-il de tels intervalles ? Nous pouvons le calculer d'après les indications de la partie précédente. Si leur nombre est $v$, alors nous désignerons les intervalles
264|265 par |

$$\mathfrak{x}'_1\mathfrak{x}''_1, \mathfrak{x}'_2\mathfrak{x}''_2 ... \mathfrak{x}'_v\mathfrak{x}''_v$$

et par

$$\mathfrak{x}_1, \mathfrak{x}_2, ...\mathfrak{x}_v$$

les valeurs rationnelles qui se trouvent isolément dans chacun des intervalles. Ensuite

nous pouvons poser :

$$3.)\ f(x) = (x - \mathfrak{x}_1)(x - \mathfrak{x}_2)...(x - \mathfrak{x}_v)Q(x) + R(x),$$

où le reste $R(x)$ est en général différent de zéro. Cependant il peut être montré que ses coefficients peuvent être rendus aussi petits que l'on veut par la détermination appropriée du nombre $r$ resté indéterminé.

Nous appliquons dans ce but la formule d'interpolation de Lagrange, d'après laquelle :

$$4.)\ R(x) = \sum_{k=1}^{v} \frac{f(\mathfrak{x}_k)}{\varphi'(\mathfrak{x}_k)} \frac{\varphi x}{x - \mathfrak{x}_k},$$

étant donné que $f(\mathfrak{x}_k) = R(\mathfrak{x}_k)$ pour $k = 1, 2...v$. $\varphi'(\mathfrak{x}_k)$ est le produit des $(v-1)$ différences $(\mathfrak{x}_k - \mathfrak{x}_i)$. Mais deux $\mathfrak{x}_i, \mathfrak{x}_k$ différents sont au moins séparés l'un de l'autre de $\frac{1}{s}$. Ensuite, comme une racine de $f(x) = 0$ se trouve toujours éloignée d'au plus $\frac{1}{s}$ de la racine de $f'(x) = 0$ la plus proche, (p.259 éq. 13.)), mais entre deux racines de $f(x) = 0$ se trouve toujours une racine [51] de $f'(x) = 0$, alors l'écart entre deux racines de $f(x) = 0$ doit de toute façon s'élever à plus de $\frac{2}{s}$. La distance entre deux de ces valeurs rationnelles $\mathfrak{x}_i, \mathfrak{x}_k$ se trouvant à côté de ces racines ne peut pourtant être plus grande que $\frac{2}{r.s.|D|}$, étant donné que chaque $\mathfrak{x}_h$ ne peut pas être éloigné de plus de $\frac{1}{r.s.|D|}$ de la racine de $f(x) = 0$ correspondante. Toujours est-il que l'on a :

$$\begin{aligned} |\mathfrak{x}_i - \mathfrak{x}_k| &> \frac{2}{s} - \frac{2}{r.s.|D|} \\ &> \frac{1}{s}, \end{aligned}$$

si l'on a seulement $r.\,|D| \geq 2$ et [52] :

$$|\prod(\mathfrak{x}_i - \mathfrak{x}_k)| > \frac{1}{s^{v-1}}\ ,\ \frac{1}{|\prod(\mathfrak{x}_i - \mathfrak{x}_k)|} > s^{v-1}.$$

Comme de plus $|\ |f(\mathfrak{x}_k)| < \frac{1}{r}$, alors le coefficient de chaque terme de la somme est 265|266
dans l'équation 4.) :

$$5.)\ \left|\frac{f(\mathfrak{x}_k)}{\varphi'(\mathfrak{x}_k)}\right| < \frac{s^{v-1}}{r}.$$

Si l'on considère en outre les fonctions entières $\frac{\varphi(x)}{x - \mathfrak{x}_k}$ de degré $(v-1)$ qui apparaissent en plus, alors, puisque $|\mathfrak{x}_k| < r$ $(k = 1, 2, ...v)$ lorsque $r$ a la signification connue, la suite de leurs coefficients devient plus petite que :

$$(v-1) \cdot r\ ,\ \frac{(v-1)(v-2)}{2} \cdot r^2 ...\ ,\ r^{v-1}.$$

Si $t$ est un nombre qui est plus grand que chacune de ces valeurs, alors finalement chaque coefficient de la fonction entière $R(x)$ de degré $(v-1)$ sera en valeur absolue plus petit que $\frac{v.s^{v-1}.t}{r}$. On a donc :

51. *N.d.T.* Commentaire dans la page de gauche : « Il faut observer qu'ici le théorème de Rolle sera utilisé ».

52. *N.d.T.* Correction dans la page de gauche :

$$\left|\prod(\mathfrak{x}_i - \mathfrak{x}_k)\right| > \frac{1}{s^{v-1}}\ ,\ \frac{1}{|\prod(\mathfrak{x}_i - \mathfrak{x}_k)|} < s^{v-1}.$$

$$6.)\ R(x) = \sum_{k=1}^{v} \frac{f(\mathfrak{x}_k)}{\varphi'(\mathfrak{x}_k)} \cdot \frac{\varphi(x)}{x - \mathfrak{x}_k} = \frac{v \cdot s^{v-1} \cdot t}{r} \cdot \Theta(x)$$

et

$$7.)\ \prod(x - \mathfrak{x}_k) \cdot Q(x) + \frac{v \cdot s^{v-1} \cdot t}{r} \cdot \Theta(x),$$

où $\Theta(x)$ est une fonction entière de degré au plus $(v-1)$ dont les coefficients sont tous plus petits que 1. $s, v, t$ sont des nombres entiers que l'on peut déterminer précisément à partir des coefficients des fonctions $f(x)$ fournies, $r$ par contre est quelconque et peut être rendu suffisamment grand de sorte que les coefficients de la fonction de reste, que nous désignerons désormais à cause de cette dépendance à $r$ par $f_r(x)$, soient strictement plus petits que n'importe quelle valeur donnée aussi petite que l'on veut. Le résultat que nous avons ainsi obtenu est le suivant :

Il n'est en général pas possible de factoriser une fonction entière donnée $f(x)$ de degré $n$ en un produit de facteurs linéaires dont on pourrait donner des coefficients
266|267 une | véritable interprétation arithmétique exacte et exhaustive. Nous sommes par contre en mesure, lorsque nous désignons la valeur de $2\chi(f, f')$ par $v$, de factoriser la différence des deux fonctions $f(x) - f_r(x)$, où $f_r(x)$ est bien une fonction entière déterminable de degré au plus $(v-1)$ avec des coefficients aussi petits que l'on veut, en un produit qui se compose d'une fonction entière $Q(x)$ de degré $(n-v)$ avec des coefficients rationnels et $v$ facteurs linéaires rationnels, de sorte que l'on ait :
$8.)\ f(x) - f_r(x) = \prod_{k=1}^{v}(x - \mathfrak{x}_k)Q(x)$.

Nous pouvons représenter de la manière suivante la factorisation de $f(x) - f_r(x)$, lorsque nous supposons $f(x)$ exempte de toute racine rationnelle, en utilisant exclusivement les extrémités $x_k, x_{k+1}$ d'un intervalle de longueur $\frac{1}{s}$, dans lequel nous avons découpé l'intervalle $-r ... + r$ :

$$f(x) - f_r(x) = \prod_{(k)} \left\{ (x - x_k - 1).\frac{1}{2}\ sgn\ f'(x_k).[\ sgn\ f(x_{k+1} -\ sgn\ f(x_k)] + 1 \right\},$$

où le produit s'étend à toutes les valeurs de $k$. Car, si on a $x_{k+1} > x_k$ et que $f(x) = y$ change de signe dans l'intervalle $(x_k, x_{k+1})$, alors on peut avoir $sgn\ f'(x_k) = sgn\ f(x_{k+1})$, donc l'expression $\frac{1}{2}\ sgn\ f'(x_k)[\ sgn\ f(x_{k+1)} -\ sgn\ f(x_k)]$ est égale à $+1$ ou 0, selon que l'intervalle $(x_{k+1}, x_k)$ contient ou pas une « racine réelle ». Une autre sorte de représentation est :

$$f'(x) - f_r(x) = \prod_{(k)} (x - \delta_k.x_k - (1 - \delta_k)x_{k+1})^{\frac{1}{2}\ sgn\ f'(x_k).[\ sgn\ f(x_{k+1}) - sgn\ f(x_k)]}.Q(x),$$

où les $\delta_k$ (pour $k = 1, 2, ...$) sont n'importe quelles vraies fractions[53] positives, après la fixation desquelles $f_r(x)$ est aussi une fonction déterminée.

267|268 Nous pouvons exprimer aussi le contenu de l'équation 8.) de | la manière suivante : pour n'importe quelle fonction entière $f(x)$ pour laquelle $2\chi(f, f') = v$, on peut établir une fonction quelconque voisine qui a $v$ racines rationnelles.

Pour le cas particulier où $v = n$, nous avons atteint notre but, car alors $Q(x) = 1$. Mais si $v < n$, nous pouvons alors modifier la fonction de reste $f_r(x)$ de sorte que la différence $f(x) - f_r(x)$ puisse être représentée par le produit de $v$ facteurs linéaires et $\frac{n-v}{2}$ facteurs quadratiques avec des coefficients rationnels. Si nous avons montré cela et fourni aussi une méthode pour déterminer les coefficients des facteurs quadratiques, alors le contenu du théorème fondamental est entièrement représenté dans sa version et sa justification purement arithmétique et par conséquent pour le plus important de cette première partie, on pourrait presque dire que pour le seul théorème de la théorie des équations algébriques on a obtenu le « le dernier mot ».

53. *N.d.T.* Fraction dont la valeur absolue est inférieure à 1

Il faut prendre en considération le fait que, pour tout autre que les fonctions rationnelles entières, cette façon de factoriser n'est seulement encore possible que si l'on peut établir pour celles-ci les intervalles-isolants à partir de grandeurs déterminables. Que pour ce faire la condition de continuité ne soit pas suffisante, nous l'apprenons des fonctions données par *Riemann* et *Weierstrass*, par exemple $\sum_{n=1}^{\infty} \frac{\sin n^2 x\pi}{n^2}$, qui certes sont continues, mais permettent tout aussi peu la détermination des intervalles-isolants que ceux d'un quotient différentiel.

## VII.) Sur une propriété du domaine dans lequel les fonctions ont la même caractéristique.

Si nous supposons affecté d'une façon répétée donnée à chaque point d'une variété de dimension $v$ un | système de fonctions déterminé $(f(x), f_1(x))$ , alors 268|269
nous avons vu que, pour se déplacer d'un système à un autre dont la caractéristique est différente de celle du premier, nous devions impérativement franchir la variété-résultante. On peut cependant aussi donner pour cela la preuve que l'on peut toujours, sans passage par la variété-résultante, passer d'un système de fonctions à un autre dont la caractéristique est la même que celle du premier. Nous avons certes déjà donné cette preuve p. 222 pour le cas spécial où $f_1(x) = f'(x)$, mais à ce moment-là sous l'hypothèse de la factorisation d'une fonction entière de degré $n$ en $n$ facteurs linéaires réels ou complexes, de sorte que nous ne pouvons plus, d'après nos dernières recherches, lui attribuer complètement la même force arithmétique. À la place nous voulons confirmer sa justesse dans le cas général sans utiliser cette fiction.

Pour cela nous allons prendre le chemin suivant. Nous montrons que nous pouvons toujours passer d'un système de fonctions quelconque [54]

$$1.)\ (f(x), f_1(x))$$

au système :

$$2.)\ \begin{cases} \varphi(x) = (x-1)(x-3)(x-5)...(x-2\mu+1)(x^2+1^2)(x^2+3^2)...(x^2+(n-\mu-1)^2) \\ \psi(x) = (x-2)(x-4)(x-6)...(x-2\mu)(x^2+2^2)(x^2+4^2)...(x^2+(n-\mu)^2) \end{cases},$$

(dans lequel seulement et uniquement $\mu$ dépend de la caractéristique du système 1.)), sans traverser la variété-résultante. Car alors nous sommes aussi en mesure d'accéder à partir de chaque système de fonctions à un autre dont la caractéristique a la même valeur, en prenant toujours notre chemin à travers | un point fixe du domaine. 269|270

Nous effectuons le passage du système de fonctions 1.) au système 2.) pas à pas. Comme on a $\chi(f, f_1) = \chi(f, t.f + f_1)$ pour toutes les valeurs de $t$, alors nous pouvons passer par une variation continue du paramètre $t$ dans 0...1 à un système dont les fonctions sont de même degré. Nous transformons ensuite le système de fonctions, que nous renommons $(f, f_1)$ , de sorte que l'on ait

$$\chi(f, f') = \chi(f_1, f_1')$$

c'est-à-dire que le nombre de racines réelles des deux fonctions est le même. Cela peut aussi se voir de la façon suivante. Nous allons du système $(f, f_1)$ au système voisin $(\overline{f_1}, f_1)$ où on a $\overline{f_1} = (1-t)f + f.f_1$ et $t = \tau$, avec $\tau$ suffisamment petit pour que sur le chemin de $t = 0$ à $t = \tau$ la variété-discriminante $R(f, f') = 0$ ne soit pas traversée, c'est-à-dire qu'on ait donc $\chi(f, f') = \chi(f_1, f_1')$ [55]. C'est toujours possible,

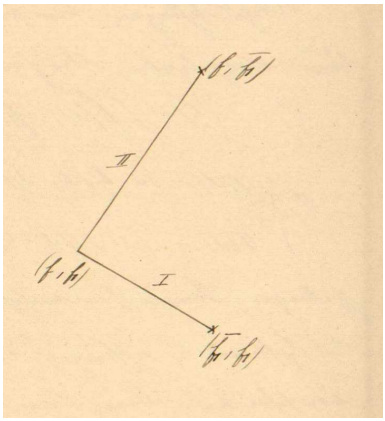
54. La seule chose à supposer est que $R(f, f_1) \neq 0$, $R(f, f') \neq 0$ et $R(f_1, f_1') \neq 0$.

55. *N.d.T.* Correction dans la page de gauche : $\chi(f, f') = \chi(\overline{f_1}, f_1')$

car d'après l'hypothèse le point $(f, f_1)$ doit se trouver à une distance finie de cette variété-discriminante. Mais on a maintenant :

$$\chi(f, \overline{f_1}) = \chi(f, (1-\tau).f + \tau.f_1) = \chi(f, f_1),$$

donc nous avons obtenu, en prenant comme point de départ $(f, f_1)$, le système de fonctions $(f, \overline{f_1})$ avec la propriété voulue.



Pour l'explication, on utilise la figure ci-contre. Nous pourrions faire le chemin $I$ sans traverser la variété $R(f, f') = 0$, et le chemin $II$ sans passage par la variété $R(f, f_1) = 0$.

Si maintenant on a $\chi(f, f') = \chi(\overline{f_1}, \overline{f_1'}) = v$, alors nous pouvons aller du système de fonctions $(f, \overline{f_1})$ sans passer par la variété $R(f, f_1) = 0$, - cette exigence ne doit
270|271 surtout plus être considérée comme évidente lors de notre prochaine étape - | à un système voisin quelconque qui sera désigné par $(F, F_1)$ et qui est tel que les deux fonctions ont chacune $v$ racines rationnelles :

$$3.) \quad \begin{cases} F(x) = (x-a_1)...(x-a_v)\Phi(x) \\ F_1 = (x-b_1)...(x-b_v)\Psi(x). \end{cases}$$

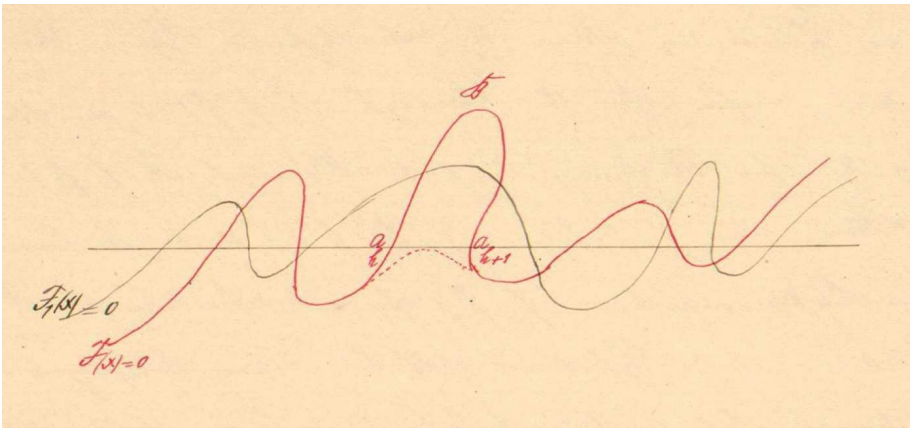
Le résultant de $F(x)$ et $F_1(x)$ est, comme en général

$$R(\varphi, \Theta.\psi) = R(\varphi, \psi).R(\varphi, \Theta) :$$

$$\begin{aligned} R(F, F_1) &= R(\Phi, \Psi).R(\Phi, \prod(x-b)).R(\Psi, \prod(x-a)).\prod(a_i - b_k) \\ &= R(\Phi, \Psi).\prod\Phi(b).\prod\Psi(a).\prod(a_i - b_k). \end{aligned}$$

La valeur de ce résultant peut, lors d'une variation du système de fonctions, passer par zéro seulement si $\prod(a_i - b_k)$ passe par zéro ; car $\Phi(x)$ et $\Psi(x)$ ne peuvent s'annuler pour aucune valeur réelle de $x$. Nous pouvons de plus supposer qu'entre deux racines rationnelles de $F(x) = 0$ il se trouve une et une seule racine de $F_1(x) = 0$. Car si par exemple les deux racines $a_h$ et $a_{h+1}$ de $F(x) = 0$ ne sont séparées par aucune racine de $F_1(x) = 0$, alors nous introduisons dans $F(x)$ à la place de $(x-a_h)(x-a_{h+1})$ le produit [56] $[(x - \frac{1}{2}(2-t)a_h + t.a_{h+1}][x - \frac{1}{2}(2-t)a_{h+1} + t.a_h]$ et nous laissons $t$ varier sur 0...1, moyennant quoi nous passons de $(x-a_h)(x-a_{h+1})$ à $(x - \frac{1}{2}(a_h + a_{h+1}))^2$. Si nous développons ce carré et que nous modifions d'aussi peu que l'on veut ses

56. *N.d.T.* Correction dans la page de gauche : $[(x - \frac{1}{2}((2-t)a_h + t.a_{h+1})][x - \frac{1}{2}((2-t)a_{h+1} + t.a_h)]$



coefficients dans une direction donnée, alors la fonction quadratique, supposée égale à zéro, possède deux racines complexes (v. la discussion sur l'équation du quatrième degré). Puisque deux telles racines se laissent donc toujours éliminer (lorsque, pour parler clairement, l'on tire vers le bas dans la figure ci-contre l'arc $\mathfrak{K}$ suffisamment pour qu'il traverse l'axe et qu'ensuite il se trouve sous l'axe), alors nous pouvons aussi supposer que 3.) se met sous la forme : | 271|272

$$4.)\quad \begin{cases} (x-\alpha_1)...(x-\alpha_\mu).\Phi(x) \\ (x-s_1)...(x-s_v)\Psi(x). \end{cases},$$

où $\alpha_1 < s_1 < \alpha_2 < s_2 < ... < \alpha_\mu < s_\mu$, et ainsi le nombre de racines rationnelles des deux équations doit aussi être encore le même. Ce système de courbes que représentent les fonctions 4.) a donc les mêmes propriétés que le système de courbes dans le cas particulier de la caractéristique $\chi(f, f')$, avec pour seule différence que pour 4.) il n'existe que des points de sortie, tandis que pour le système $(f, f')$ il n'y a que des points d'entrée. Comme maintenant $\chi(f, f') = \frac{1}{2}(E - A)$, alors on a dans notre cas [57]

$$\chi(F, F') = -\frac{1}{2}\mu.$$

Maintenant nous partons du système de fonctions 4.) de la manière suivante : avec le système 2. Fevr.

$$\prod_{h=1}^{\mu}[x-(\alpha_h(1-t)+(2h-1).t)].\Phi(x);\ \prod_{h=1}^{\mu}[x-(s_h(1-t)+2h\cdot t)].\Psi(x)$$

nous allons de $t = 0$ à $t = 1$, durant lequel la variété-résultante $R(F, F_1) = 0$ ne sera pas traversée, puisque l'inéquation pour les racines :

$$\alpha_1(1-t)+t < s_1(1-t)+2t < \alpha_2(1-t)+3t < ... < \alpha_\mu(1-t)+(2\mu-1)t < s_\mu(1-t)+2\mu\cdot t.$$

reste préservée pour toutes les valeurs de $t$ considérées. De cette manière nous obtenons pour le système de fonctions :

$$5).\ \prod_{h=1}^{\mu}(x-(2h-1))\Phi(x)\,;\ \prod_{h=1}^{\mu}(x-2h).\Psi(x).$$

Il n'est pas nécessaire maintenant de transférer le système de fonctions 5.) dans le

57. *N.d.T.* Correction dans la page de gauche : $\chi(F, F_1) = -\frac{1}{2}\mu$

système de fonctions 2.), que nous voulons écrire de façon analogue au système 5.) :

$$\varphi(x) = \prod_{h=1}^{\mu}(x-(2h-1))\Phi_0(x)\ ,\ \psi(x) = \prod_{h=1}^{\mu}(x-2h).\psi_0(x)$$

où donc la signification de $\Phi_0(x)$ et $\Psi_0(x)$ est claire.
Comme le résultant de[58] 5.) $\prod\Phi(2h)\prod\Psi(2h-1).R(\Phi,\Psi)$ disparaît seulement lorsque $R(\Phi,\Psi)$ est égal à zéro (en supposant que lors de la variation on conserve
272|273 toujours $\chi(\Phi,\Phi')=0$ et $\chi(\Psi,\Psi')=0$), alors nous pouvons | préciser les derniers problèmes restant :

On doit trouver un chemin qui relie entre eux les systèmes de fonctions $(\Phi,\Psi)$ et $(\Phi_0,\Psi_0)$, sans passer par la variété $R(\Phi,\Psi)=0$.

Si maintenant nous supposons comme toujours que les coefficients des termes de plus haut degré de la fonction 1.) sont positifs, alors les valeurs des deux fonctions $\Phi(x)$ et $\Psi(x)$, dans lesquelles les coefficients de la plus haute puissance sont aussi positifs pour des valeurs très grandes de $x$, sont ainsi pour toutes les valeurs réelles de l'argument positives, tout comme les valeurs des fonctions $\Phi_0$ et $\Psi_0$. À la place du système $(\Phi,\Psi)$ nous pouvons aussi prendre le système voisin très proche $(\Phi_1,\Psi_1)$, si nous veillons seulement à ce que sur le chemin de $(\Phi,\Psi)$ à $(\Phi_1,\Psi_1)$ nous ne traversions ni les variétés-discriminantes $R(\Phi,\Phi')=0$, $R(\Psi,\Psi')=0$ ni la variété-résultante $R(\Phi,\Psi)=0$. Le chemin voulu du point $(\Phi,\Psi)$ au point $(\Phi_0,\Psi_0)$ est alors le suivant :

Nous nous rendons de $(\Phi,\Psi)$ à $(\Phi_1,\Psi_1)$ (ce système sera déterminé plus précisément) et de $(\Phi_1,\Psi_1)$ à $(\Phi_0,\Psi_0)$ sur une ligne droite

$$6.)\quad \begin{array}{l}\Phi_t(x) = (1-t).\Phi_1(x) + t.\Phi_0(x)\\ \Psi_t(x) = (1-t).\Psi_1(x) + t.\Psi_0(x)\ \text{pour } t \text{ dans } 0...1\end{array}$$

Ces fonctions sont positives pour toutes les valeurs de $t$ par lesquelles nous passons, et toutes les valeurs de $x$, et toutes deux peuvent ainsi posséder seulement des facteurs linéaires complexes. Si nous traversons donc en cours de route la variété-résultante $R(\Phi,\Psi)=0$, alors le point d'entrée doit être un point d'une variété d'un ordre plus
273|274 bas qui se trouve | sur $R(\Phi,\Psi)=0$, car, si la fonction 6.) possède un diviseur commun, celui-ci doit être au moins du second degré. Si les fonctions $\Phi_1,\Psi_1$, donc aussi $\Phi_0,\Psi_0$, sont par exemple de degré quatre alors, avec notre conception et nos désignations précédentes (p. 227), cela signifie : une ligne droite entre deux points du domaine $G_2$ peut traverser la surface discriminante au plus en un point de la parabole faisant partie de la surface inclue dans $G_2$. Les fonctions appartenant à cette parabole possèdent un diviseur commun du second degré.

Pour montrer maintenant qu'une telle variété, qui a un degré inférieur de plus d'une unité à celle dans laquelle nous nous déplaçons, peut toujours être contournée, on procède de la manière la plus appropriée en s'approchant jusqu'à la variété en ligne droite et ensuite, lorsque cela est nécessaire, en l'évitant - ici il est plus facile de fixer le point de départ $(\Phi_1,\Psi_1)$, qui possède bien une certaine variabilité, de sorte que la ligne droite qui relie $(\Phi_1,\Psi_1)$ avec $(\Phi_0,\Psi_0)$ ne coupe pas la variété.

Nous avons d'abord à déterminer la variété, ou avec d'autres valeurs les deux équations que l'on peut écrire avec des coefficients indéterminés sous la forme

$$V(x) = v_0 + v_1x + v_2x^2 + ... + v_nx^n\ ;\ W(x) = w_0 + w_1x + w_2x^2 + ... + w_nx^n,$$

de sorte que celles-ci possèdent un facteur commun de degré deux. La condition pour que les deux fonctions $V$ et $W$ aient en général un diviseur commun est :

$$R(V,W) = R(v_0,v_1,...v_n;w_0,w_1...w_n) = 0.$$

58. *N.d.T.* Correction dans la page de gauche : 5.) $\prod\Phi(2h)\prod\Psi(2h-1).R(\Phi,\Psi)\prod(2h-1-2k)$

| Ainsi nous avons : 274|275

$$R(v_0 - V(u), v_1 ... v_n; w_0 - W(u), w_1 ... w_n) = 0,$$

puisque les fonctions $V(x) - V(u)$ et $W(x) - W(u)$ possèdent le facteur commun $x - u$. Comme on a $v_0 - V(u) = -v_1 u - v_2 u^2 - ... - v_n u^n$, alors la différenciation de cette égalité [59] donne :

$$-u.\frac{\partial R(v_0 - V(u), v_1, v_2 ...)}{\partial (v_0 - V(u))} + \frac{\partial R(v_0, v_1, v_2 ...)}{\partial v_1} = 0$$

et si on développe d'après les puissances de $V(u)$ et $W(u)$ on a :

$$-u.\frac{\partial R(v_0, v_1, ...; w_0 ...)}{\partial v_0} + \frac{\partial R(v_0, v_1 ...; w_0 ...)}{\partial v_1} \equiv 0 \text{ (modd. } V(u), W(u)).$$

Si l'on met à la place de l'inconnu $u$ la variable $x$, alors on a :

$$7) \quad -x.\frac{\partial R(V, W)}{\partial v_0} + \frac{\partial R(V, W)}{\partial v_1} \equiv 0 \text{ (modd. } V(u), W(u)).$$

Si maintenant $V(x)$ et $W(x)$ ont un facteur commun, alors le côté gauche de 7.) doit aussi être aussi divisible par celui-ci. Si ce facteur commun est une fonction linéaire, alors il n'est rien d'autre que

$$-x.\frac{\partial R(V, W)}{\partial v_0} + \frac{\partial R(V, W)}{\partial v_1};$$

Si par contre il est quadratique, alors

$$8.) \ \frac{\partial R(V, W)}{\partial v_0} = 0, \ \frac{\partial R(V, W)}{\partial v_1} = 0$$

reste valable.

Ces deux équations sont des conditions nécessaires pour l'existence d'un facteur quadratique commun aux fonctions $V(x)$ et $W(x)$ ; il est nécessaire et suffisant que l'on ait simultanément

$$9.) \ R = 0 \text{ et } \frac{\partial R(V, W)}{\partial v_0} = 0.$$

L'équation 9.) considère la seconde des équations 8.) comme une de ses conséquences nécessaires, car puisque nous pouvons tirer de 7.) et 9.) la conclusion que $\frac{\partial R(V, W)}{\partial v_1}$ est divisible par une fonction de $x$ du premier degré au moins, alors on doit avoir $\frac{\partial R(V, W)}{\partial v_1} = 0$. | 275|276

Pour notre cas particulier nous avons donc à montrer que les équations

$$10.) \ R(\Phi_t(x), \Psi_t(x)) = 0 \ , \ \frac{\partial R(\Phi_t(x), \Psi_t(x))}{\partial \varphi_{t,0}} = 0,$$

où $\varphi_{t,0}$ est le terme absolu [60] de la fonction $\Phi_t(x)$, ne peuvent être valables ensemble pour aucune valeur de $t$. Si l'on considère ces deux expressions qui dépendent de $t$ comme fonctions de cette grandeur, alors le résultant doit donc être

$$11.) \ R\left(R(\Phi_t(x), \Psi_t(x)), \frac{\partial R(\Phi_t(x), \Psi_t(x))}{\partial \varphi_{t,0}}\right) \neq 0.$$

59. *N.d.T.* Ajout dans la page de gauche : selon $v_1$

60. *N.d.T.* Le « terme absolu » d'un polynôme en $x$ est le terme ne contenant aucun $x$.

Ce résultant ne dépend plus de $x$ ni de $t$. Au contraire il change lors des variations du système de fonctions $\Phi_1(x), \Psi_1(x))$ et accepte d'autres valeurs. Si en particulier nous nous déplaçons dans une variété d'ordre $v$ et nous comprenons les coefficients de ces systèmes de fonctions comme des fonctions entières de $v$ paramètres $x_1, x_2, ...x_v$ que nous pouvons faire varier à l'intérieur de limites fixées, alors il y a pour 11.) deux cas possibles. Soit l'expression du côté gauche de 11.), qui est donc aussi une fonction entière de $x_1...x_v$, est identiquement nulle, soit seulement pour certaines valeurs de ces variables. Dans le deuxième cas notre souhait est réalisé : nous pouvons trouver un système $(\Phi_1, \Psi_1)$ pour lequel l'inéquation 11.) reste valable et qui ainsi peut servir comme point de sortie pour notre chemin (Éq. 6.)). Si l'on suppose possible le premier cas on est conduit en revanche à une contradiction. Car si l'expression dans 11.) s'annule pour toutes les valeurs de $x_1, x_2, ...x_v$, alors elle doit aussi être égale à
276|277 zéro pour les valeurs particulières de $x_1, x_2...x_v$ par lesquelles le système | $(\Phi_0, \Psi_0)$ résulte de $(\Phi_1, \Psi_1)$ ; nous pouvons en effet arriver par des modifications illimitées des $v$ paramètres indépendants les uns des autres à chaque système de fonctions d'une variété d'ordre $v$. Mais comme maintenant on a $R(\Phi_0, \Psi_0) \neq 0$, alors nous sommes devant la contradiction suivante : deux grandeurs indépendantes de $t$, $R(\Phi_0(x), \Psi_0(x))$ et $\partial R(\Phi_0(x), \Psi_0(x))\partial\varphi_0$, dont la première est de toute façon différente de zéro, doivent avoir un facteur commun dépendant de $t$.

Ainsi nous avons démontré que les domaines isolés par la variété $R(f, f_1) = 0$ qui contiennent les systèmes de fonctions avec la même caractéristique, sont connexes au sens usuel.

Pour le cas particulier $f_1 = f'$ ce résultat signifie que nous pouvons arriver de chaque équation $f(x) = 0$ avec $\mu$ racines réelles par une variation de leurs coefficients à l'équation

$$(x-1)(x-2)...(x-\mu)(x^2+1^2)(x^2+2^2)...(x^2+\left(\left(\frac{n-\mu}{2}\right)^2\right) = 0,$$

sans traverser la variété-discriminante et ainsi à n'importe quelle autre équation qui possède également $\mu$ racines réelles (c'est-à-dire pour laquelle $\chi(g(x), g'(x)) = \frac{\mu}{2}$).

# Annexe C

# Traduction des troisième et quatrième démonstrations de Gauss du théorème fondamental de l'algèbre

Nous allons donner les traductions des deux preuves du théorème fondamental de l'algèbre que Gauss a écrites en 1816 et en 1849. Ces traductions ont été réalisées à partir des textes suivants :

1. Pour la troisième : Carl Friedrich Gauss, *Theorematis de resolubilitate functionum algebraicarum integrarum in factores reales Demonstratio tertia supplentum commentationis praecedentis* (Gottingae), 1816. In Werke III, pp 57–64.
2. Pour la dernière : Carl Friedrich Gauss, *Beiträge zur Theorie der algebraischen Gleichungen*, 1849. In Werke III, pp 73–85.

Nous nous sommes par ailleurs aidé des traductions en allemand de Netto dans [Netto, 1876].

## C.1 Traduction de la troisième démonstration de Gauss (1816)

*Théorème sur la factorisation*
*des fonctions algébriques entière*
*en facteurs réels*

*Troisième démonstration*

*Supplément au commentaire précédent.*

Après que mon commentaire précédent fut déjà imprimé, encore une fois des réflexions sur le même sujet me conduisirent à une nouvelle démonstration du théorème, démonstration qui, ainsi que la précédente, est purement analytique, mais repose sur des principes totalement différents et au regard de sa simplicité, lui est de loin préférable. C'est pourquoi les pages qui suivent seront consacrées à cette troisième démonstration.

1.

Que soit représentée la fonction suivante d'indéterminée $x$ :

$$X = x^m + Ax^{c-1} + Bx^{m-2} + Cx^{m-3} + etc. + Lx + M \in \mathbb{R}$$

dans laquelle les coefficients A, B, C etc. sont des quantités réelles fixées. Soient $r$, $\varphi$ d'autres indéterminées, et posons :

$$\begin{aligned}
&r^m \cos m\varphi + Ar^{m-1}\cos(m-1)\varphi + Br^{m-2}\cos(m-2)\varphi + Cr^{m-3}\cos(m-3)\varphi \\
&+ etc. + Lr\cos\varphi + M = t \\
&r^m \sin m\varphi + Ar^{m-1}\sin(n-1)\varphi + Br^{m-2}\sin(m-2)\varphi + Cr^{m-3}\sin(m-3)\varphi \\
&+ etc. + Lr\sin\varphi = u \\
&mr^m \cos m\varphi + (m-1)Ar^{m-1}\cos(m-1)\varphi + (m-2)Br^{m-2}\cos(m-2)\varphi \\
&+ (m-3)Cr^{m-3}\cos(m-3)\varphi + etc. + Lr\cos\varphi + M = t' \\
&mr^m \sin m\varphi + (m-1)Ar^{m-1}\sin(n-1)\varphi + (m-2)Br^{m-2}\sin(m-2)\varphi \\
&+ (m-3)Cr^{m-3}\sin(m-3)\varphi + etc. + Lr\sin\varphi = u' \\
&m^2 r^m \cos m\varphi + (m-1)^2 Ar^{m-1}\cos(m-1)\varphi + (m-2)^2 Br^{m-2}\cos(m-2)\varphi \\
&+ (m-3)Cr^{m-3}\cos(m-3)^2\varphi + etc. + Lr\cos\varphi + M = t'' \\
&m^2 r^m \sin m\varphi + (m-1)^2 Ar^{m-1}\sin(n-1)\varphi + (m-2)^2 Br^{m-2}\sin(m-2)\varphi \\
&+ (m-3)^2 Cr^{m-3}\sin(m-3)\varphi + etc. + Lr\sin\varphi = u'' \\
&y = \frac{(t^2+u^2)(tt''+uu'') + (tu'-ut')^2 - (tt'+uu')^2}{r(t^2+u^2)^2}
\end{aligned}$$

Il est clairement permis de supprimer le facteur $r$ de la dernière formule puisqu'il divise $t', u', t''$ et $u''$. Ensuite, soit $R$ une quantité positive fixée, certes arbitraire, mais cependant plus grande que la plus grande des quantités

$$mA\sqrt{2},\ \sqrt{nB\sqrt{2}},\ \sqrt[3]{mC\sqrt{2}},\ \sqrt[4]{mD\sqrt{2}}, etc.$$

en faisant abstraction des signes des quantités $A, B, C$ etc., c'est à dire après avoir changé les négatifs, s'il y en a, en positifs. Cela étant posé, je dis que $tt' + uu'$ prend sans nul doute une valeur positive, s'il est établi que $r = R$, quelle que soit la valeur réelle attribuée à $\varphi$.

*Démonstration.* Posons

$$
\begin{aligned}
&R^m \cos 45 + AR^{m-1}\cos(45+\varphi) + BR^{m-2}\cos(45+2\varphi) + CR^{m-3}\cos(45+3\varphi)\\
&+ etc. + LR\cos(45+(m-1)\varphi) + M\cos(45+m\varphi) = T\\
&R^m \sin 45 + AR^{m-1}\sin(45+\varphi) + BR^{m-2}\sin(45+2\varphi) + CR^{m-3}\sin(45+3\varphi)\\
&+ etc. + LR\sin(45+(m-1)\varphi) + M\sin(45+m\varphi) = U\\
&mR^m \cos 45 + (m-1)AR^{m-1}\cos(45+\varphi) + (m-2)BR^{m-2}\cos(45+2\varphi)\\
&+ (m-3)CR^{m-3}\cos(45+3\varphi) + etc. + LR\cos(45+(m-1)\varphi) = T'\\
&mR^m \sin 45 + (m-1)AR^{m-1}\sin(45+\varphi) + (m-2)BR^{m-2}\sin(45+2\varphi)\\
&+ (m-3)CR^{m-3}\sin(45+3\varphi) + etc. + LR\sin(45+(m-1)\varphi) = U'
\end{aligned}
$$

et il est évident que :

I. $T$ est composée des parties

$$
\begin{aligned}
&\frac{R^{m-1}}{m\sqrt{2}}(R + mA\sqrt{2}\cos(45+\varphi))\\
+&\frac{R^{m-2}}{m\sqrt{2}}(R^2 + mB\sqrt{2}\cos(45+2\varphi))\\
+&\frac{R^{m-3}}{m\sqrt{2}}(R^3 + mC\sqrt{2}\cos(45+3\varphi))\\
+&\frac{R^{m-4}}{m\sqrt{2}}(R^4 + mD\sqrt{2}\cos(45+4\varphi))
\end{aligned}
$$

dont on perçoit facilement que chacune d'entre elles finit par être positive, pour n'importe quelle valeur déterminée réelle de $\varphi$ : de là $T$ devient nécessairement une valeur positive. De la même manière, on prouve que $U, T', U'$ deviennent aussi positives, d'où également $TT' + UU'$ devient nécessairement une quantité positive.

II. Pour $r = R$ les fonctions $t, u, t', u'$ se changent respectivement en

$$
\begin{aligned}
&T\cos(45+m\varphi) + U\sin(45+m\varphi)\\
&T\sin(45+m\varphi) - U\cos(45+m\varphi)\\
&T'\cos(45+m\varphi) + U'\sin(45+m\varphi)\\
&T'\sin(45+m\varphi) - U'\cos(45+m\varphi)
\end{aligned}
$$

On le prouve facilement en faisant les développements. De là la valeur de la fonction $tt' + uu'$, pour $r = R$, se dérive alors en $= TT' + UU'$, et est de plus une quantité positive. C.Q.F.D.

Par ailleurs, d'après les mêmes formules, nous pouvons conclure que la valeur de la fonction $tt + uu$, pour $r = R$, est $TT + UU$, et est en plus positive, d'où nous concluons que, pour aucune valeur de $r$ plus grande respectivement que $mA\sqrt{2}$, $\sqrt{mB\sqrt{2}}$, $\sqrt[3]{mC\sqrt{2}}$, *etc.*, on ne peut avoir en même temps $t = 0$ et $u = 0$.

2.

*THÉORÈME.* Entre les limites $r = 0$ et $r = R$, et $\varphi = 0$ et $\varphi = 360$, il existe toujours des valeurs des indéterminées $r, \varphi$, pour lesquelles on a en même temps $t = 0$ et $u = 0$.

*Démonstration.* Supposons que le théorème n'est pas vrai ; il est alors évident que la valeur de $tt + uu$ doit être une quantité positive pour toutes les valeurs des indéterminées comprises entre les limites fixées, et que par conséquent la valeur de $y$ est toujours définie. Considérons l'intégrale double

$$\iint ydrd\varphi$$

de $r = 0$ jusqu'à $r = R$ et depuis $\varphi = 0$ jusqu'à $\varphi = 360$, intégrale qui prend donc une valeur finie pleinement déterminée. Cette valeur, que nous noterons $\Omega$, devra rester la même, soit que l'intégration soit d'abord faite en fonction de $\varphi$ puis en fonction de $r$, soit qu'elle soit faite dans l'ordre inverse. Mais nous avons la primitive, en considérant $r$ comme constant,

$$\int yd\varphi = \frac{tu' - ut'}{r(t^2 + u^2)}$$

comme on le confirme facilement par une différenciation selon $\varphi$. Une constante ne doit pas être ajoutée si du moins nous supposons que l'intégrale doit commencer depuis $\varphi = 0$, puisque pour $\varphi = 0$, on a $\frac{tu'-ut'}{r(t^2+u^2)} = 0$. C'est pourquoi, puisque manifestement $\frac{tu'-ut'}{r(t^2+u^2)}$ s'annule pour $\varphi = 360$, l'intégrale $\int yd\phi$ depuis $\varphi = 0$ jusqu'à $\varphi = 360$ devient $= 0$, $r$ demeurant indéterminé. Or de là il s'ensuit que $\Omega = 0$. Nous avons pareillement la primitive, en considérant $\varphi$ comme constant,

$$\int ydr = \frac{tt' + uu'}{t^2 + u^2}$$

comme on le confirme également facilement par la différenciation selon $r$ ; une constante ne doit pas non plus être ajoutée, l'intégrale commençant depuis $r = 0$. C'est pourquoi l'intégrale depuis $r = 0$ jusqu'à $r = R$, en vertu de ce qui a été démontré dans la partie précédente, $= \frac{TT'+UU'}{T^2+U^2}$, et de plus, en vertu du théorème de la partie précédente, devient toujours une quantité positive pour n'importe quelle valeur réelle de $\varphi$. De là, $\Omega$ lui-même, c'est-à-dire la valeur de l'intégrale

$$\int \frac{TT' + UU'}{T^2 + U^2} d\varphi$$

depuis $\varphi = 0$ jusqu'à $\varphi = 360$ devient nécessairement une quantité positive [1]. Ce qui est absurde puisque nous avons auparavant trouvé la même valeur $= 0$ : c'est pourquoi cette supposition ne peut tenir et, de là, la vérité du théorème est prouvée.

3.

La fonction $X$ par la substitution $x = r(\cos\varphi + sin\varphi.\sqrt{-1})$ devient $t + u\sqrt{-1}$ et de même elle devient $t - u\sqrt{-1}$ par la substitution $x = r(\cos\varphi - sin\varphi.\sqrt{-1})$. Donc, si pour des valeurs déterminées de $r$, $\varphi$, par exemple pour $r = g$ , $\varphi = G$, se produit à la fois $t = 0, u = 0$ (il a été démontré dans la partie précédente que de telles valeurs existent), $X$ par l'une et l'autre des substitutions

$$x = g(\cos G + \sin G.\sqrt{-1}), x = g(\cos G - \sin G.\sqrt{-1})$$

vaudra 0 et de là sera divisible par $x - g(\cos G + \sin G.\sqrt{-1})$ et aussi assurément par $x - g(\cos G - \sin G.\sqrt{-1})$. Chaque fois que l'on n'a ni $\sin G = 0$, ni $g = 0$, ces diviseurs sont différents, et par conséquent $X$ sera également divisible par leur produit

1. Ce qui est déjà évident en soi. Par ailleurs, l'intégrale indéfinie s'extrait facilement en $= m\varphi + 45 - \arctan\frac{U}{T}$ et il peut être démontré autre part ( en effet, il n'est pas encore évident en soi de savoir quelle valeur parmi les nombreuses de la fonction multiforme $\arctan\frac{U}{T}$ qui conviennent pour $\varphi = 360$ il faut adopter) que la valeur de celle-ci jusqu'à $\varphi = 360$ doit être établie en $= m \times 360$ ou bien $= 2m\pi$. Mais cela n'est pas nécessaire pour ce que nous voulons établir.

$$xx - 2g\cos Gx + gg$$

Mais chaque fois que, soit $\sin G = 0$ et donc $\cos G = \pm 1$, ou bien que $g = 0$, ces facteurs sont assurément identiques à $= x \pm g$. C'est pourquoi il est certain que la fonction $X$ contient un diviseur réel de second ou de premier ordre, et puisque la même conclusion vaut à nouveau pour le quotient, $X$ sera complètement résoluble en de tels facteurs. C.Q.F.D.

4.

Bien que dans les parties précédentes nous nous soyons acquitté entièrement de la tâche que nous nous étions proposée, il ne sera cependant pas tout à fait inutile d'ajouter encore certains points au raisonnement de la partie 2. À partir de la supposition que $t$ et $u$ ne s'annulent en même temps pour aucune des valeurs des indéterminées $r, \varphi$ entre les limites qui leur ont été assignées, nous avons abouti à une contradiction inévitable, d'où nous avons conclu à la fausseté de la supposition même. Cette contradiction doit cesser, si réellement il y a des valeurs de $r, \varphi$ pour lesquelles $t$ et $u$ deviennent ensemble $= 0$. Et pour illustrer cela davantage, nous observons que pour de telles valeurs on a $tt + uu = 0$, et en outre que $y$ lui-même est indéfini, d'où il ne sera pas davantage permis de traiter l'intégrale double $\iint ydrd\varphi$ comme une quantité définie. En effet, de manière générale, en notant $\xi, \eta, \zeta$ les coordonnées de points dans l'espace, l'intégrale $\iint ydrd\varphi$ représente le volume d'un solide qui est contenu entre cinq plans dont les équations sont

$$\xi = 0, \eta = 0, \zeta = 0, \xi = R, \eta = 360$$

et par une surface dont l'équation est $\zeta = y$, en considérant comme négatives les parties dans lesquelles les coordonnées $\zeta$ sont négatives. Mais on peut comprendre que la condition tacite de la continuité de la sixième surface cesse lorsque $y$ est indéfini, et qu'il peut advenir que cette représentation n'ait plus de signification. Dans un tel cas, il ne peut y avoir de discussion sur l'interprétation de l'intégrale $\iint ydrd\varphi$, et il ne faut pas en outre s'étonner que les opérations analytiques appliquées par un calcul aveugle à des choses vaines conduisent à des absurdités. L'intégration $\int yd\varphi = \int \frac{tu' - ut'}{r(t^2+u^2)}$ est une vraie intégration, c'est-à-dire une somme, seulement dans la mesure où, dans les limites dans lesquelles elle s'étend, $y$ est partout une quantité définie, mais elle n'a pas de sens, si entre ces limites $y$ devient en quelque endroit indéfinie. Si nous expliquons selon les règles habituelles une telle intégrale $\int \eta d\xi$ , qui, généralement, représente une aire entre la ligne des abscisses et une courbe, dont l'ordonnée $= \eta$ avec pour abscisse $\xi$, sans nous préoccuper de la continuité, nous obtenons très fréquemment des contradictions. Par exemple en établissant $\eta = \frac{1}{\xi\xi}$, l'analyse fournit l'intégrale $C - \frac{1}{\xi}$, par laquelle une aire est bien définie tant que la courbe conserve sa continuité ; mais celle-ci étant interrompue pour $\xi = 0$, si quelqu'un demande de manière maladroite la grandeur de l'aire depuis l'abscisse négative jusqu'à la positive, il obtiendra de la formule une réponse absurde, que celle-ci est négative. Or pour ceux qui souhaitent d'autres cas semblables et d'autres paradoxes de l'analyse, nous continuerons plus abondamment à une autre occasion.

Qu'il me soit permis d'ajouter ici une unique observation. Quand sont proposées sans restriction des questions qui peuvent s'avérer absurdes dans certains cas, très souvent l'analyse délibère de manière à donner une réponse en partie vague. Ainsi pour la valeur de l'intégrale $\iint ydrd\varphi$ qui s'étend de $r = e$ jusqu'à $r = f$, et depuis $\varphi = E$ jusqu'à $\varphi = F$, si la valeur de $\frac{u}{t}$

pour $r = e$, $\varphi = E$ est désignée par $\theta$
pour $r = e$, $\varphi = F$ est désignée par $\theta$'
pour $r = f$, $\varphi = E$ est désignée par $\theta$"
pour $r = f$, $\varphi = F$ est désignée par $\theta$"'

on obtient facilement par des opérations analytiques

$$\arctan\theta - \arctan\theta' - \arctan\theta'' + \arctan\theta'''$$

Certes l'intégrale peut alors avoir réellement une valeur déterminée, chaque fois que $y$ demeure toujours définie dans les limites indiquées : cette valeur est certainement contenue dans la formule donnée, mais elle n'est cependant pas encore déterminée en totalité, puisque arctan est une fonction multiforme, et qu'il faudra décider par d'autres considérations (certes peu difficiles) quelles valeurs de la fonction devront être utilisées de préférence dans un cas précis. En revanche, chaque fois que $y$ s'avère indéfini en quelque endroit entre les limites assignées, la question de la valeur de l'intégrale $\int\int ydrd\varphi$ est absurde : si l'on s'obstine néanmoins à vouloir arracher une réponse à l'analyse, on obtiendra en raison de la diversité des méthodes tantôt une réponse, tantôt une autre qui cependant seront chacune contenues dans la formule générale donnée plus haut.

## C.2 Traduction de la quatrième démonstration de Gauss (1849)

# *CONTRIBUTIONS A LA THEORIE DES ÉQUATIONS ALGEBRIQUES DE CARL FRIEDRICH GAUSS*

*Présentées à la Société Royale des Sciences le 16 juillet 1849.*

*Traités de la Société Royale des Sciences. Volume 4. Göttingen.*

Dans cet essai sont traités deux sujets différents concernant les équations algébriques. Tout d'abord je présente la preuve que j'ai donnée il y a cinquante ans du théorème fondamental de la théorie des équations algébriques, dans une forme nouvelle et avec de nombreux ajouts. La deuxième partie est consacrée aux équations algébriques à trois termes, et contient des méthodes destinées à déterminer facilement les racines non seulement réelles mais aussi complexes de telles équations.

## PREMIER TRAITÉ

L'essai présenté en 1799, *Demonstratio nova theorematis, omnem functionem algebraicam rationalem inteqram unius variabilis in factores reales primi vel secundi gradus resolvi posse*, avait un double objectif ; celui de montrer en effet premièrement l'insuffisance et l'aspect illusoire de toutes les preuves tentées jusqu'alors de cet important théorème de la théorie des équations algébriques, et deuxièmement d'en donner une nouvelle preuve parfaitement rigoureuse. Il est inutile de revenir une fois encore au premier objet. J'ai fait suivre cette nouvelle preuve de deux autres, et une quatrième a été établie par Cauchy. Ces quatre preuves sont fondées sur autant de bases différentes, mais toutes quatre s'accordent en commençant par prouver seulement la présence d'un facteur dans la fonction étudiée. Cela ne diminue pas la rigueur des preuves : en effet, il est clair que si l'on retire un facteur de la fonction donnée, il reste une fonction similaire d'ordre inférieur, à laquelle le théorème peut être de nouveau appliqué, et finalement on obtient ainsi par la répétition du procédé une décomposition complète en facteurs de la fonction de départ. Cependant chaque argumentation gagne sans aucun doute un achèvement plus élevé s'il est prouvé qu'elle est capable de rendre directement claire la présence de tous les facteurs. J'ai déjà indiqué dans le mémoire susmentionné (Art. 23) que c'était le cas pour la première preuve, sans l'achever complètement : cela doit être maintenant complété, et je profite en même temps de cette occasion pour répéter les moments principaux de la preuve entière dans une forme différente, et je l'espère, plus claire. En ce qui concerne l'habillage extérieur du théorème, celui qu'utilisait celle de 1799 était que la fonction $x^n + Ax^{n-1} + Bx^{n-2} + etc.$ peut être résolue en facteurs réels du premier ou du second degré, habillage choisi pour éviter toute intervention de grandeur imaginaire. A l'heure actuelle où le concept de grandeur complexe est courant, il semble plus approprié d'utiliser cette forme et d'exprimer ainsi le théorème : toute fonction peut être décomposée en $n$ facteurs simples, où la partie constante de ces facteurs n'est pas forcément des grandeurs réelles, mais peut aussi être une valeur complexe. Par cet habillage, le théorème gagne en généralité, car la restriction à des valeurs réelles pour les coefficients $A, B, ...$ n'a plus besoin d'être supposée, et pour toutes les valeur la même permission est accordée.

### 1.

Nous considérons donc la fonction de la grandeur indéterminée $x$ :

$$x^n + Ax^{n-1} + Bx^{n-2} + etc. + Mx + N = X$$

où $A, B, ...M, N$ désignent des coefficients réels ou imaginaires. Le lien entre les racines de l'équation $X = 0$ et les facteurs simples de $X$ est donné par l'algèbre élémentaire. C'est-à-dire que lorsque l'équation est satisfaite par la substitution $x = p$ , alors $x - p$ est un facteur de $X$, et que s'il y a $n$ différentes façons de satisfaire cette équation, par les substitutions $x = p', x = p'', etc.$, alors le produit $(x-p)(x-p')(x-p'')...$ sera identique à $X$. Cependant, sous certaines circonstances, une solution, par exemple $x = p$, peut apparaître dans $X$ sous le facteur $(x - p)^2$ ou $(x - p)^3$ ou une puissance quelconque plus élevée, circonstance dans laquelle on considère la racine $p$ comme doublement présente, triplement présente, etc. Si l'on exige seulement la preuve que la fonction $X$ admet un facteur simple, il est suffisant d' établir uniquement l'existence d'une quelconque racine à l'équation $X = 0$. S'il doit être prouvé en une fois la factorisation complète de la fonction en facteurs simples, alors on doit montrer que l'équation $X = 0$ peut être satisfaite par $n$ valeurs différentes de $x$, ou par un plus petit nombre de ces solutions, dont une partie contient un caractère de multiplicité, de telle manière qu'en additionnant toutes les racines égales ou différentes, le total soit égal à $n$.

2.

L'ensemble des grandeurs complexes dans lequel les valeurs de $x$ qui vérifient l'équation $X = 0$ doivent être cherchées est un ensemble infini à deux dimensions, où l'on pose pour une telle valeur $x = t+iu$ (où $i$ signifie toujours l'unité imaginaire $\sqrt{-1}$) avec $t$ et $u$ des valeurs réelles comprises entre $-\infty$ et $+\infty$. Nous avons maintenant tout d'abord à extraire de cet ensemble un ensemble fini en dehors duquel il est sûr qu'aucune racine d'une certaine équation $X = 0$ ne peut se trouver. Cela peut se faire de plus d'une manière ; pour notre but, c'est la suivante qui semble le mieux convenir. À la place de la forme $t + iu$ on utilise

$$x = r(\cos\rho + i\sin\rho)$$

où, pour embrasser la totalité de l'ensemble des grandeurs complexes, $r$ doit s'étendre sur toutes les valeurs positives de 0 à $+\infty$ et $\rho$ de 0 à 360°, ou bien, ce qui est la même chose, d'une valeur de départ quelconque à une valeur finale plus grande de 360°. Afin d'obtenir une limite pour $r$ au-dessus de laquelle plus aucune valeur ne peut correspondre à une racine de l'équation $X = 0$, je pose tout d'abord les coefficients de chaque monôme de $X$ dans la même forme que $x$, c'est-à-dire

$$\begin{aligned} A &= a(\cos\alpha + i\sin\alpha) \\ B &= b(\cos\beta + i\sin\beta) \\ C &= c(\cos\gamma + i\sin\gamma) \\ &... \end{aligned}$$

où $a, b, c, ...$ doivent donc indiquer des grandeurs positives déterminées, hormis l'une ou l'autre d'entre elles qui peut être égale à 0. Je considère ensuite l'équation

$$r^n - \sqrt{2}(ar^{n-1} + br^{n-2} + cr^{n-3} + ...) = 0$$

laquelle, comme on le voit facilement, a une racine positive, et en fait (d'après le théorème d'Harriot) n'en a qu'une seule. Soit $R$ cette racine, et il est alors clair que pour chaque valeur positive de $r$ plus grande que $R$, la valeur de

$$r^n - \sqrt{2}(ar^{n-1} + br^{n-2} + cr^{n-3} + ...)$$

est positive, et que la même chose sera vraie pour la fonction

$$nr^n - \sqrt{2}\left((n-1)ar^{n-1} + (n-2)br^{n-2} + (n-3)cr^{n-3} + ...\right)$$

3.

J'affirme à présent que la valeur $R$ est appropriée pour donner une telle limite aux valeurs de $r$, comme il est requis dans l'article précédent. La preuve de cette affirmation est menée de la manière suivante.
Je pose de manière générale que $X = T + iU$, où naturellement $T$ et $U$ indiquent des grandeurs réelles et deviennent en fait

$$\begin{aligned} T = r^n \cos n\rho \quad & +ar^{n-1}\cos((n-1)\rho+\alpha) \\ & +br^{n-2}\cos((n-2)\rho+\beta) \\ & +cr^{n-3}\cos((n-3)\rho+\gamma)+... \\ U = r^n \sin n\rho \quad & +ar^{n-1}\sin((n-1)\rho+\alpha) \\ & +br^{n-2}\sin((n-2)\rho+\beta) \\ & +cr^{n-3}\sin((n-3)\rho+\gamma)+... \end{aligned}$$

On voit facilement que pour n'importe quel $r$ positif choisi plus grand que $R$, le signe de $T$ sera nécessairement le même que celui de $\cos n\rho$, tant que ce cosinus ne sera pas en valeur absolue inférieur à $\sqrt{\frac{1}{2}}$. C'est-à-dire que l'on a seulement besoin de poser $T$ sous la forme suivante

$$\begin{aligned} \pm T = \quad & \sqrt{\frac{1}{2}}r^n - ar^{n-1} - br^{n-2} - cr^{n-3} - ... \\ & + \left[\pm\cos n\rho - \sqrt{\frac{1}{2}}\right] r^n \\ & + \left[1 \pm \cos((n-1)\rho+\alpha))\right] ar^{n-1} \\ & + \left[1 \pm \cos((n-2)\rho+\beta))\right] br^{n-2} \\ & + \left[1 \pm \cos((n-3)\rho+\gamma))\right] cr^{n-3} + ... \end{aligned}$$

où les signes précédents doivent être valables pour le cas d'un $\cos n\rho$ négatif, et où la première partie du membre de droite est positive, ou tout du moins pas négative, d'après le théorème donné dans l'article précédent. De façon analogue, on montre (en écrivant seulement dans la formule ci-dessus $U$ à la place de $T$ et sinus à la place de cosinus) que, sous les mêmes hypothèses sur $r$, $U$ a toujours le même signe que $\sin n\rho$, tant que la valeur absolue de ce sinus n'est pas inférieure à $\sqrt{\frac{1}{2}}$. En conséquence dans tous les cas au moins l'une des deux grandeurs $T$ ou $U$ a un signe positif ou négatif défini d'avance, et par conséquent la fonction $X$ ne peut être $= 0$ pour aucune valeur de $\rho$. C.Q.F.D.

4.

Pour éclairer encore plus le comportement et les changements du signe de $T$ et $U$ (pour une valeur définie de $R$ plus grande que $r$ ), qu'on laisse $\rho$ parcourir toutes les valeurs entre deux limites de valeurs différentes de 360°, qui ne sont pas pourtant 0 et 360, mais $-\omega$ et $(8n-1)\omega$ où l'on a posé

$$\frac{45}{n} = \omega$$

Je partage l'intervalle tout entier en $4n$ parties égales, de telle façon que la première s'étende de $-\omega$ à $\omega$, la seconde de $\omega$ à $3\omega$, la troisième de $3\omega$ à $5\omega$,... On doit aussi dans un premier temps envisager la valeur des différentielles $\frac{dT}{d\rho}$, $\frac{dU}{d\rho}$ pour lesquelles on a

$$\frac{dT}{d\rho} = -nr^n \sin n\rho \quad -(n-1)ar^{n-1}\sin((n-1)\rho+\alpha)$$

$$\frac{dU}{d\rho} = nr^n \cos n\rho \quad \begin{array}{l} -(n-2)br^{n-1}\sin((n-2)\rho+\beta) \\ -... \\ +(n-1)ar^{n-1}\cos((n-1)\rho+\alpha) \\ +(n-2)br^{n-1}\cos((n-2)\rho+\beta) \\ +... \end{array}$$

Par une conclusion semblable à celle de l'article précédent et par le théorème de la conclusion de l'article 2, on voit facilement que $\frac{dT}{d\rho}$ a toujours le signe opposé à celui de $\sin n\rho$, tant que ce sinus n'est pas pris en valeur absolue plus petit que $\sqrt{\frac{1}{2}}$, et qu'en revanche $\frac{dU}{d\rho}$ a toujours le même signe que $\cos n\rho$, tant que la valeur absolue de ce cosinus n'est pas inférieure à $\sqrt{\frac{1}{2}}$. De là on observe la conclusion suivante.

Dans le premier intervalle, c'est-à-dire de $\rho = -\omega$ à $\rho = +\omega$, T est constamment positive, et $U$ en revanche est négative pour ses premières valeurs et positive pour ses dernières. Par conséquent entre les deux elle est certainement une fois et une fois seulement $= 0$ car dans l'intervalle entier $\frac{dU}{d\rho}$ est positif.
Dans le deuxième intervalle $U$ est certainement positive, $T$ est d'abord positive puis à la fin négative et entre les deux $T = 0$ une fois et une fois seulement, car dans cet intervalle $\dfrac{dT}{d\rho}$ est négatif.
Dans le troisième intervalle $T$ est constamment négative, $U$ est soumis à un changement de signe, de telle manière que $U$ devient $= 0$ une fois.
Dans le quatrième intervalle $U$ est constamment négative et $T = 0$ une fois.
Dans les intervalles suivants cette situation se répète dans le même ordre, de telle manière que le cinquième est équivalent au premier, le sixième au second, etc.

5.

A partir de la succession des valeurs positives et négatives de $T$ et de $U$ discutée dans l'article précédent, succession qui a lieu pour chaque valeur de $r$ plus grande que $R$ [2], on conclut que pour des valeurs plus petites de $r$ des croisements certains doivent avoir lieu, qui contiennent en eux l'essence de notre théorème à prouver.

Je représenterai l'exposé des preuves sous l'habillage du point de vue de la géométrie de position car ainsi on y gagne une plus grande clarté et une plus grande simplicité. Mais au fond le véritable contenu de toute l'argumentation fait partie d'un domaine plus élevé, indépendant de l'espace, domaine de la théorie générale et abstraite des grandeurs, qui a pour objet les combinaisons de grandeurs connectées par continuité, domaine qui est encore peu construit pour le moment, et dans lequel on ne peut pas non plus se mouvoir sans un langage emprunté à un dessin des lieux.

6.

Le domaine entier des nombres complexes sera représenté par un plan sans limite dans lequel chaque point, dont les coordonnées relatives à deux axes se coupant perpendiculairement sont $t, u$, sera considéré comme correspondant à la grandeur complexe $t + iu$ : si on donne ces grandeurs complexes sous la forme $x = r(\cos\phi + i\sin\phi)$,

2. Il est facile de montrer que pour la valeur $r = R$ elle-même une même progression reste valide, sous réserve seulement qu'ensuite dans certains cas particuliers une valeur de transition de $\rho$ (c'est-à-dire une valeur pour laquelle $T$ ou $U = 0$) peut coïncider avec l'une des valeurs $-\omega, \omega, 3\omega, 5\omega, ...$, alors que pour toute valeur plus grande de $r$ chaque valeur de transition de $\rho$ doit se trouver entre deux de ces grandeurs. Je ne m'attarde pourtant pas ici, car pour nos besoins il suffit qu'ait été prouvée l'existence d'une valeur de transition à partir d'une valeur de $r$.

alors $r, \phi$ représentent les coordonnées polaires des points correspondants. La représentation de l'ensemble des valeurs complexes pour lesquelles $r$ a une valeur fixée sera en conséquence un cercle dont le rayon est cette valeur, et dont le centre est l'origine des coordonnées. Celui de ces cercles pour lequel $r$ est plus grand que $R$ d'une différence choisie à loisir, je le désignerai par $K$. Et je désignerai par $(1), (2), (3), ..., (2n)$ les points de ce cercle qui correspondent aux valeurs de $\rho$ respectivement comprises entre $\omega$ et $3\omega$, entre $5\omega$ et $7\omega$, entre $9\omega$ et $11\omega$, ... entre $\omega$ et $3\omega$, pour lesquelles $T = 0$ d'après l'article 4. On remarque en même temps que pour les points $(1), (3), (5), ...,$ $U$ sera positif et que pour les points $(2), (4), (6), ...$ il sera en revanche négatif.

7.

On s'aperçoit que l'ensemble de ces points du plan pour lesquels $T$ est positif forme un ensemble de surfaces cohérentes quand on remarque que par un passage continu d'un point à un autre, $T$ varie continument. De même, tous les points pour lesquels $T$ sera négatif forment un ensemble de surfaces cohérentes. Entre la partie du plan de la première sorte et celle de la seconde se trouvent des points pour lesquels $T = 0$, et d'après la nature de la fonction $T$, ces points ne peuvent pas former aussi des bouts de surface, mais seulement des lignes, lesquelles délimitent d'un côté et de l'autre des surfaces. L'espace se trouvant à l'extérieur de $K$ contient $n$ surfaces de la première sorte, qui alternent avec tout autant de surfaces de la seconde sorte, et dont chacune s'étend de façon cohérente d'un bout du cercle $K$ jusqu'à l'infini. Mais en même temps il est clair que chacune de ces surfaces s'étend au delà du cercle dans l'espace intérieur, et que les cas suivants peuvent avoir lieu selon la présentation suivante.

1. La surface en question commençant à partir d'une partie de $K$ se termine isolée dans la surface du cercle ; sa délimitation périphérique se compose alors seulement de deux parties connectées, dont l'une est un composant de $K$ et l'autre se trouve dans la surface intérieure du cercle. Dans la figure jointe, laquelle se rapporte à une équation du cinquième degré et où les signes de $T$ sont inscrits dans les différentes surfaces, trois des surfaces se trouvent dans ce cas avec un $T$ positif, la première ayant pour frontières 10.1 et 1.11.10 ; la seconde 4.5 et 5.12.4 ; la troisième 6.7 et 7.17.6. On trouve deux surfaces d'une espèce semblable avec $T$ négatif.

2. La surface traverse simplement la surface du cercle de manière à former une surface cohérente avec une autre qui entre à un autre endroit. La totalité de la ligne périphérique se composera alors de quatre parties, dont deux appartiennent au cercle $K$, et les deux autres à son espace intérieur. Dans notre figure, ce cas se trouve dans la surface délimitée par 2.3 ; 3.0.8 ; 8.9 ; 9.11.2.

3. La surface se divise à l'intérieur du cercle une fois ou plus, de manière à former une surface cohérente avec deux autres ou plus qui entrent à un autre endroit, surface cohérente dont la totalité de la ligne périphérique se composera alors de six, huit ou plus de parties en nombre pair, parties qui appartiennent tour à tour au cercle et à l'espace intérieur. Dans notre exemple cela se produit pour la surface dont la frontière sera fournie par les six morceaux 3.4 ; 4.12.5 ; 5.6 ; 6.13.7 ; 7.8 ; 8.0.3, et dans laquelle $T$ est négatif.

8.

Pour avoir une énumération complète de toutes les organisations concevables des surfaces entrant à l'intérieur du cercle, il faudrait ajouter encore des modifications supplémentaires. Quand par exemple une surface se divise certes en deux branches, mais que celles-ci se réunifient dans l'espace intérieur, alors ce cas pourrait être attribué, selon que la surface trouve sa clôture à l'intérieur après l'unification ou bien que (sans nouvelle division) elle se poursuive jusqu'à un autre endroit du cercle, au

premier ou au deuxième cas du chapitre précédent, dans lesquels l'organisation de la surface serait juste modifiée par l'inclusion d'une île ne lui appartenant pas. D'ailleurs il ne serait pas difficile de prouver rigoureusement qu'avec la forme particulière de la fonction $T$ des modifications de cette sorte ne sont pas possibles. C'est pourtant superflu pour notre but, qui dépend seulement des séquences de morceaux de la limite extérieure de chacune des surfaces qui sont dans notre discussion (c'est-à-dire de celles pour lesquelles $T$ est positif).

Nous avons déjà remarqué que le nombre de ces parties est à chaque fois pair (deux dans le premier cas du chapitre précédent, quatre dans le deuxième, 6 ou plus dans le troisième) avec à tour de rôle l'un qui appartient au cercle et l'autre à l'espace intérieur. De plus, il est clair que lorsque cette limite extérieure sera parcourue dans un même sens, sens que l'on devra choisir ici, et dans lequel la numérotation des points de $K$ s'accroît (donc par exemple dans notre figure, de telle sorte que la surface se trouve à droite de la limite extérieure), le point de départ et celui d'arrivée d'une partie appartenant au cercle seront désignés par respectivement un nombre impair et un nombre pair plus grand d'une unité, par conséquent le point de départ et celui d'arrivée de chaque partie parcourant l'espace intérieur seront désignés chaque fois par un nombre impair et un nombre pair.

Il est clair aussi que parmi les $n$ lignes pour lesquelles $T = 0$ qui entrent dans l'espace intérieur par un point de $K$ désigné par un nombre impair, toutes celles d'une espèce bien définie [3] parcourent cet espace continûment, jusqu'à ce qu'elles ressortent à un autre endroit marqué d'un nombre pair.

Alors maintenant, comme il a été déjà remarqué plus haut (fin de l'article 6), la valeur de $U$ est positive à son point de départ et elle est négative à son point final, et ainsi par continuité de la fonction $U$ deviendra nécessairement $= 0$ entre ces deux points. Ce point représentera alors une racine de l'équation $X = 0$. Et le nombre de telles lignes est ici égal à $n$, de telle sorte qu'il en résulte $n$ solutions à cette équation.

9.

Lorsque les lignes imaginaires traversent l'intérieur du cercle sans se couper les unes les autres, il est alors clair que les $n$ racines ainsi obtenues sont nécessairement distinctes. De telles traversée libres se trouvent dans notre figure avec les lignes allant de 3 à 8, de 5 à 4 et de 7 à 6, et les racines représentées par les points 0, 12, 13 font partie de celles-ci. Lorsqu'en revanche deux ou plusieurs lignes de cette sorte ont un point en commun, il est possible, mais pour autant pas nécessaire, que ce point soit en même temps celui pour lequel $U = 0$, auquel cas alors deux ou plusieurs racines sont confondues, ou bien, comme on l'exprime habituellement, seront égales entre elles. Dans notre figure, les lignes 1.10 et 9.2 se rencontrent au point 11, et en même temps $U = 0$ à ce même point ; l'équation a donc, à part les trois racines inégales déjà représentées, deux racines égales.

10.

3. Qu'elles aient toujours une course définie repose sur le fait qu'elles doivent fixer une partie de la clôture extérieure de la surface pour laquelle $T$ a un signe défini. J'ai choisi ce signe positif, ce qui est en soi totalement arbitraire. Ainsi compris, par exemple la ligne entrant en 1 s'étend de 11 jusqu'à 10 : pour une partie de la limite d'une surface sur laquelle $T$ est négative, la ligne 1.11 continuerait après 2. Si l'on parlait en revanche plutôt d'une ligne sur laquelle $T = 0$, sans la voir comme une partie de la clôture d'une surface définie, alors dans ce cas ce sera plutôt la ligne 11.9 qui sera la continuation naturelle de la 1.11. Le point de vue choisi ici distingue mon procédé actuel de celui de 1799, et contribue de manière essentielle à la simplification de l'exposé des preuves.

Il reste encore seulement à démontrer que si le point $P$ représentant la racine $= p$ se trouve à la fois sur deux ou plusieurs lignes $T = 0$, le carré de $x - p$ ou la puissance plus haute correspondante au nombre de ces lignes sera présent comme facteur dans $X$. La preuve de cela repose sur les affirmations suivantes.

Que l'on introduise à la place de l'indéterminée $x$ une autre indéterminée $z$, en posant $x = z + p$. Que l'on substitue $X$ à $Z$, où $Z$ sera alors une fonction de $z$ de même ordre que $X$ de $x$, mais dont le membre constant manquera. En ordonnant celle-ci selon les puissances croissantes de $z$, on pose $Kz^m$ le terme de plus petit degré non nul et $Z = Kz^m(1 + \xi)$ où $\xi$ sera de la forme $L.z + L'z.z + L''z^3 + ... + \frac{1}{K}z^{n-m}$ ; posons enfin

$$z = s(\cos\psi + i\sin\psi)$$

La partie réelle et imaginaire de $z$ exprime la position de chaque point quelconque du plan en tant que coordonnées rectangulaires, et les grandeurs $s, \psi$ expriment les coordonnées polaires relatives aussi au point $P$, comme les parties de $x$ et les grandeurs $r$ et $\phi$ désignent les positions au point de départ. La mise en relation d'une valeur définie de $s$ avec toutes les valeurs de $\psi$ sur une étendue de 360°représente donc les points d'un cercle qui a son centre en $P$ et dont le rayon est $= s$. On pose maintenant $K = k(\cos x + i\sin x)$ et par conséquent

$$Kz^m = ks^m(\cos(m\psi + x) + i\sin(m\psi + x))$$

ainsi pour un $s$ infiniment petit la grandeur $\xi$, qui est au moins du même ordre que $s$, devient négligeable devant 1, et par conséquent on peut poser

$$T = ks^m\cos(m\psi + x)$$

d'où l'on tire que, lorsque $\psi$ augmente de 360°, le signe de $T$ est positif sur $m$ morceaux du cercle et négatif en alternance sur tout autant de morceaux, ou bien que $T = 0$ sur $2m$ points, et ce pour $\psi = \frac{1}{m}(x - 90°)$, $\psi = \frac{1}{m}(x + 90°)$, $\psi = \frac{1}{m}(x + 270°)$, etc. Il passe ainsi $2m$ lignes par $P$ sur lesquelles $T = 0$. Ou encore si, après avoir ordonné les lignes entrantes selon $\psi$ croissante, on relie ces $2m$ lignes chacune avec la suivante, les frontières ainsi considérées seront celles des passages de $T$ positif à $T$ négatif, et ces $m$ frontières concourront en $P$.

D'un autre côté il est clair qu'aussitôt que l'indéterminée $Z$ est divisible par $z^m$ et par aucune plus haute puissance de $z$, $X$ contiendra le facteur $(x-p)^m$, mais aucune puissance supérieure de $x - p$. Lorsque $p$ symbolise l'une des racines de l'équation $X = 0$, on a aussi toujours l'exposant de la plus grande puissance de $(x - p)$ par laquelle $X$ est divisible qui est égale au nombre de frontières se rencontrant en $P$ pour les surfaces avec $T$ positif, ou ce qui est pareil, au nombre de telles surfaces se rencontrant en $P$. D'ailleurs il est facile de donner à l'exposé des preuves un habillage totalement indépendant de l'utilisation de valeurs infiniment petites, et ce de manière totalement analogue à la fin des articles 3 et 4. On peut en effet prouver l'existence d'une valeur de $s$ pour laquelle, ainsi que pour chaque valeur plus petite, le cycle entier de toutes les valeurs de $\psi$ présente la même suite de $m$ morceaux qui alternent successivement entre $T$ positif et $T$ négatif. Cette propriété est vérifiée par la racine positive de l'équation

$$0 = m\sqrt{\frac{1}{2}} - (m+1)ls - (m+2)l'ss_(m+3)l''s^3 - ...$$

où $l, l', l'', ...$ représentent les racines carrées positives des normes des grandeurs complexes $L, L', L'', ...$, ou bien où l'on pose

$$L \quad = \quad l(\cos\lambda + i\sin\lambda)$$

$$\begin{aligned} L' &= l'(\cos\lambda' + i\sin\lambda') \\ L'' &= l''(\cos\lambda'' + i\sin\lambda'') \\ &\ldots \end{aligned}$$

Je crois pourtant pouvoir ici me passer du développement très facile de cette affirmation.

En conclusion on aimerait aussi faire remarquer que, dans l'exposé des preuves de 1799, l'observation de deux systèmes de lignes était nécessaire, l'un contenant les lignes $T = 0$, l'autre celles où $U = 0$, alors que dans notre procédé d'aujourd'hui un seul système a suffi ; j'ai choisi ici le système de lignes de démarcation avec $T$ positif, on aurait pu tout aussi bien pour le même but se servir de l'étude des frontières des surfaces avec $U$ positif (ou négatif).

# Annexe D

# Réseau de citation

Nous présentons ici le tableau récapitulatif de l'ensemble des textes qui apparaissent dans le réseau de citation construit dans notre introduction [1].

| Titre | Auteurs | Année |
|---|---|---|
| Recherche de la quantité qui satisfait à la fois deux équations algébriques données | Abel | 1826 |
| Ueber ein fundamentale Begründung der Invariantentheorie. | Aronhold | 1863 |
| A new mathematical and philosophical dictionary | Barlow | 1814 |
| Sopra l'equazioni algebriche con incognite | Betti | 1858 |
| Théorie générale des équations algébriques | Bézout | 1779 |
| Remarques sur les différences qui distinguent l'interpolation de M. Cauchy de la méthode des moindres carrés, et qui assurent la supériorité de cette méthode. | Bienaymé | 1853 |
| On a Certain Multiple Definite Integral | Boole | 1846 |
| Sur la quadrature définie des surfaces courbes | Borchardt | 1854 |
| Über eine interpolationformel für eine Art symmetrischer Funktionen | Borchardt | 1860 |
| Développements sur l'équation à l'aide de laquelle on détermine les inégalités séculaires du mouvement des planètes. | Borchardt | 1847 |
| Neue Eigenschaft der Gleichung, mit deren Hülfe man die secularen Störungen der Planeten bestimmt. | Borchardt | 1846 |
| Sur les questions 241 et 141 | Brioschi | 1855 |
| Cours d'analyse algébrique | Cauchy | 1821 |
| Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs égales | Cauchy | 1815 |
| Mémoire présenté à l'académie de Turin | Cauchy | 1831 |
| Mémoire sur l'interpolation | Cauchy | 1835 |

1. Voir p. 26

| | | |
|---|---|---|
| Nouveaux exercices de mathématiques | Cauchy | 1835 |
| Mémoire sur l'évaluation d'inconnues déterminées par un grand nombre d'équations approximatives du premier degré | Cauchy | 1853 |
| Exercices de mathématiques | Cauchy | 1826 |
| Sur l'équation à l'aide de laquelle ... | Cauchy | 1829 |
| Calcul des indices des fonctions | Cauchy | 1837 |
| Note sur les fonctions de Sturm | Cayley | 1846 |
| Nouvelles recherches sur les fonctions de M. Sturm | Cayley | 1848 |
| Sur la généralisation d'un théorème de M. Jellett, qui se rapporte aux attractions. | Cayley | 1848 |
| Sur quelques formules du calcul intégral | Cayley | 1847 |
| Mémoire sur les fonctions doublement périodiques | Cayley | 1845 |
| Aperçu historique | Chasles | 1837 |
| Theorie der bilinearen Formen | Christoffel | 1866 |
| Theorie der Abelschen funktionen | Clebsch et Gordan | 1866 |
| Ueber die simultane Integration linearer partieller Differentialgleichungen | Clebsch | 1866 |
| Ueber das pfaffsche problem | Clebsch | 1863 |
| Sur un moyen général de vérifier l'expression du potentiel relatif à une masse quelconque | Dirichlet | 1846 |
| Ueber eine neue Methode zur Bestimmung vielfacher Integrale | Dirichlet | 1839 |
| Extrait d'une Lettre adressée M. Charles Hermite | Eisenstein | 1852 |
| Uber die Vergleichung von solchen ternären quadratischen Formen, welche verschiedene Determinanten haben | Eisenstein | 1852 |
| Problema algebraicum | Euler | 1770 |
| Analyse des équations déterminées | Fourier | 1831 |
| Allgemeine Lehrsätze | Gauss | 1840 |
| Theoria attractionis corpum | Gauss | 1813 |
| Grundlehrsatz der theorie der algebraischen gleichungen | Gauss | 1849 |
| Disquisitiones Arithmeticae | Gauss | 1801 |
| Recension der "Untersuchungen über die Eigenschaften der positiven ternären quadratischen Formen von Ludwig August Seeber" | Gauss | 1840 |
| Disquisitio de elementis ellipticis Palladis | Gauss | 1811 |
| Theoria combinationis observationum erroribus minimis obnoxiae | Gauss | 1823 |
| Intensitas vis magneticae terrestris ad mensuram absolutam revocata | Gauss | 1833 |
| Allgemeine Theorie des Erdmagnetismus | Gauss | 1839 |
| Theoriae transcendentium Abelianarum primi ordinis adumbratio levis | Göpel | 1847 |

| | | |
|---|---|---|
| Sur l'extension du théorème de Sturm à un système d'équations linéaires | Hermite | 1852 |
| Sur la théorie de la transformation des fonctions abeliennes | Hermite | 1855 |
| Remarques sur le théorème de Sturm | Hermite | 1853 |
| Sur la théorie des formes quadratiques ternaires | Hermite | 1850 |
| Sur l'introduction des variables continues dans la théorie des nombres | Hermite | 1851 |
| Sur l'interpolation | Hermite | 1859 |
| Note sur la réduction des fonctions homogènes à coefficients entiers et à deux indéterminées | Hermite | 1848 |
| Sur la théorie des formes quadratiques | Hermite | 1854 |
| Dato systemate $n$ aequationum linearium inter $n$ incognitas | Jacobi | 1834 |
| Theoremata nova algebraica | Jacobi | 1835 |
| De fonctionibus alternatibus | Jacobi | 1826 |
| Disquisitiones analyticae de fractionibus simplicibus | Jacobi | 1825 |
| De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares | Jacobi | 1834 |
| Sopra le funzioni di Laplace che risultano dallo sviluppo dell'espressione | Jacobi | 1844 |
| De formatione et proprietatibus determinantium | Jacobi | 1841 |
| Über eine elementare Transformation | Jacobi et borchardt | 1857 |
| Nova methodus, aequationes differentiales partiales primi ordinis inter numerum variabilium quemcunque propositas integrandi | Jacobi | 1862 |
| Fundamenta nova theoriae functionum ellipticarum | Jacobi | 1829 |
| Theoria novi multiplicatoris systemati aequationum differentialium vulgarium applicandi | Jacobi | 1845 |
| Extrait d'une Lettre adressée à M. Liouville | Jellett | 1847 |
| über Systeme von Functionen mehrer Variabeln | Kronecker | 1869 |
| über die Charakteristik von Functionen-System | Kronecker | 1878 |
| über einige Interpolationformeln für ganze Functionen mehrer Variabeln | Kronecker | 1865 |
| Ueber bilineare Formen | Kronecker | 1866 |
| Sur le théorème de Sturm | Kronecker | 1869 |
| Bemerkungen über die cubische Gleichung, durch welche die Haupt-Axen der Flächen zweiten Grades bestimmt werden | Kummer | 1843 |
| Quatre suppléments à la théorie analytique des probabilités | Laplace | 1820 |
| Nouvelles méthodes pour la détermination des orbites des comètes | Legendre | 1805 |
| Sur les équations algébriques à plusieurs inconnues | Liouville | 1847 |

| | | |
|---|---|---|
| Mémoire sur quelques propositions générales de géométrie et sur la théorie de l'élimination | Liouville | 1841 |
| Remarques sur un théorème de Jacobi | Liouville | 1841 |
| Ueber gewisse Beziehung zwischen räumlichen | Lipschitz | 1866 |
| Untersuchungen über die Anwendung eines Abbildungsprincips auf die Theorie der Gravitation | Lipschitz | 1863 |
| Solution de la question 241 | Loxhay | 1852 |
| Ueber die Bestimmung des Grades einer durch Elimination hervorgehenden Gleichung | Minding | 1841 |
| Mémoire sur l'élimination | Poisson | 1802 |
| Grundlagen für eine allgemeine Theorie der Functionen | Riemann | 1851 |
| Neue Darstellung der Resultante der Elimination | Rosenhain | 1846 |
| Mémoire sur les fonctions de deux variables et à quatre périodes | Rosenhain | 1851 |
| Untersuchungen über die Eigenschaften der positiven Ternaeren quadratischen Formen | Seeber | 1831 |
| Cours d'algèbre | Serret | 1849 |
| Elementary theorems relating to determinants | Spottiswoode | 1851 |
| Uber parallele Flächen | Steiner | 1840 |
| Systematische Entwickelung der Abhängigkeit geometrischer Gestalten von einander | Steiner | 1832 |
| Démonstration d'un théorème d'algèbre de M. Sylvester | Sturm | 1842 |
| Démonstration d'un théorème de M. Cauchy, relatif aux racines imaginaires des équations | Sturm et Liouville | 1836 |
| Autre démonstration du même théorème | Sturm | 1836 |
| Mémoire sur la résolution des équations numériques | Sturm | 1835 |
| Analyse d'un Mémoire sur la résolution des équations numériques | Sturm | 1829 |
| On a theory of the Syzygetic Relation | Sylvester | 1853 |
| On rational derivation from equations of coexistence | Sylvester | 1839 |
| On a remarkable modification of Sturm's theorem | Sylvester | 1853 |
| Note on a remarkable modification of Sturm's theorem | Sylvester | 1853 |
| A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares | Sylvester | 1852 |
| On the relation between the minor determinants of linearly equivalent quadratic functions | Sylvester | 1851 |
| On a certain fundamental theorem of determinants | Sylvester | 1851 |
| An enumeration of the contacts of lines and surfaces of the second order | Sylvester | 1851 |

| | | |
|---|---|---|
| On the intersections, contacts, and other correlations of two conics expressed by indeterminate coordinates | Sylvester | 1850 |
| Sketch of a memoir on elimination, transformation and canonical forms | Sylvester | 1851 |
| Additions to the articles “On a new class of theorems" and “on Pascal’s theorem" | Sylvester | 1851 |
| An inslantaneous demonstration of Pascal’s theorem by the method of indeterminate coordinates | Sylvester | 1850 |
| On a new class of theorems in elimination between quadratic functions | Sylvester | 1850 |
| Solution de la question 141 | Tardy | 1854 |
| Sur les fractions continues | Tchebichef | 1858 |
| Miscellanea analytica de aequationibus algebraicis | Waring | 1762 |
| Neuer Beweis des Fundamentalsatzes der Algebra | Weierstrass | 1868 |
| Zur theorie der abelschen functionen | Weierstrass | 1847 |

# Bibliographie

# Bibliographie


[Abel *et al.*, 2004] Abel, N., Laudal, O. et Piene, R. (2004). *The Legacy of Niels Henrik Abel : The Abel Bicentennial, Oslo, June 3-8, 2002*, volume 1. Springer Science & Business Media.

[Abel, 1826] Abel, N. H. (1826). Algèbre élémentaire. Recherche de la quantité qui satisfait à la fois à deux équations algébriques données. *Annales de Gergonne*, 17:204–213.

[Abel, 1839] Abel, N. H. (1839). Sur la résolution algébrique des équations. *Oeuvres Complètes de NH Abel*, pages 185–209.

[Académie de Berlin, 1751] Académie de Berlin (1751). *Histoire de l'Académie royale des sciences et des belles lettres de Berlin : avec les mémoires pour la même année, tirez des registres de cette Académie*. Numéro vol. 7. Ambroise Haude.

[Aronhold, 1863] Aronhold, S. (1863). Ueber ein fundamentale Begründung der Invariantentheorie. *Journal für die reine und angewandte Mathematik*, 62:281–345.

[Bachmacova, 1960] Bachmacova, I. (1960). le théorème fondamental de l'algèbre et la construction des corps algébriques. *Archives internationales d'histoire des sciences*, 13:211–222.

[Baker, 1912] Baker, H. F. (1912). Biographical notice. *The Collected Mathematical Papers of James Joseph Sylvester, ed. HF Baker*, 4:XV–XXXVII.

[Baltzer, 1857] Baltzer, R. (1857). *Theorie und anwendung der Determinanten*. S. Hirzel.

[Begehr, 1998] Begehr, H. (1998). *Mathematics in Berlin : On behalf of the Berliner Mathematische Gesellschaft.* Birkhäuser.

[Belhoste, 1998] Belhoste, B. (1998). Pour une réévaluation du rôle de l'enseignement dans l'histoire des mathématiques. *Revue d'histoire des mathématiques*, 4(2):289–304.

[Bell, 1945] Bell, E. (1945). *The development of mathematics*. Courier Corporation.

[Bell, 1953a] Bell, E. (1953a). *Men of mathematics. Vol. 1.* Penguin.

[Bell, 1953b] Bell, E. (1953b). *Men of mathematics. Vol. 2.* Penguin.

[Berge, 1959] Berge, C. (1959). *Espaces topologiques : Fonctions multivoques*. Dunod, Chapman & Hall.

[Betti, 1858] Betti, E. (1858). Sopra l'Equazioni algebriche con più incognite. *Annali di Matematica Pura ed Applicata (1858-1865)*, 1(1):1–8.

[Bézout, 1779] Bézout, E. (1779). *Théorie générale des équations algébriques.* de l'imprimerie de Ph.-D. Pierres, rue S. Jacques.

[Bienaymé, 1853] Bienaymé, J. (1853). Remarques sur les différences qui distinguent l'interpolation de M. Cauchy de la méthode des moindres carrés, et qui assurent la supériorité de cette méthode. *Journal de Mathématiques Pures et Appliquées*, pages 299–308.

[Biermann, 1959] Biermann, K. (1959). *Johann Peter Gustav Lejeunde Dirichlet : Dokumente für sein Leben und Wirken ;(zum 100. Todestag).* Akademie-Verlag.

[Biermann, 1973] Biermann, K. (1973). *Die Mathematik und ihre Dozenten an der Berliner Universität, 1810-1920.* Akademie-Verlag.

[Biermann, 1981] Biermann, K. (1981). Kronecker, Leopold. *Dictionary of Scientific Biography. Scribner, New York.*

[Bôcher, 1895] Bôcher, M. (1895). Gauss's third proof of the fundamental theorem of algebra. *Bulletin of the American Mathematical Society*, 1(8):205–209.

[Bolzano, 1816] Bolzano, B. (1816). *Der binomische Lehrsatz und als Folgerung aus ihm der polynomische, und die Reihen, die zur Berechnung der Logarithmen und Exponentialgrössen dienen : genauer als bisher erwiesen.* Enders.

[Bolzano, 1973] Bolzano, B. (1973). *Theory of Science, ed. by Jan Berg. 1973.* D. Reidel.

[Bolzano *et al.*, 1905] Bolzano, B., Hankel, H. et Jourdain, P. (1905). Rein analytischer beweis des lehrsatzes : dass zwischen je zwey werthen, die ein entgegengesetztes resultat gewähren, wenigstens eine reelle wurzel der gleichung liege, von Bernard Bolzano.– Untersuchungen über die unendlich oft oszillierenden und unstetigen funktionen, von Hermann Hankel. (153).

[Boniface, 1999] Boniface, J. (1999). Kronecker et le concept de nombre, traduction et présentation. *La Gazette des mathématiciens*, pages 49–70.

[Boniface, 2004] Boniface, J. (2004). *Hilbert et la notion d'existence en mathématiques.* Cambridge Univ Press.

[Boniface, 2006] Boniface, J. (2006). Les Constructions des nombres réels dans le mouvement d'arithmétisation de l'analyse (Paris : Ellipses, 2002). *Revue d'histoire des sciences*, 59(1):151–153.

[Boniface et Schappacher, 2001] Boniface, J. et Schappacher, N. (2001). 'Sur le concept de nombre en mathématique' - Cours inédit de Leopold Kronecker à Berlin (1891). *Revue d'Histoire des Mathématiques*, 7:207–275.

[Borchardt, 1861] Borchardt, C. (1861). *Über eine Interpolationsformel für eine Art symmetrischer Functionen und über deren Anwendung.* Mathematische Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin. Königlich Akademie der Wissenschaften zu Berlin.

[Borchardt, 1846] Borchardt, C. W. (1846). Neue Eigenschaft der Gleichung, mit deren Hülfe man die seculären Störungen der Planeten bestimmt. *Journal für die reine und angewandte Mathematik*, 30:38–45.

[Borchardt, 1847] Borchardt, C. W. (1847). Développements sur l'équation à l'aide de laquelle on détermine les inégalités séculaires du mouvement des planètes. *Journal de Mathématiques pures et appliquées*, 12:50–67.

[Borchardt, 1852] Borchardt, C. W. (1852). Sur le théorème de M. Sturm ; d'après M. Borchardt. *Nouvelles annales de mathématiques*, 11:403–405.

[Borchardt, 1854] Borchardt, C. W. (1854). Sur la quadrature définie des surfaces courbes. *Journal de Mathématiques Pures et Appliquées*, pages 369–394.

[Borchardt, 1857] Borchardt, C. W. (1857). Bemerkungen über die beiden vorstehenden Aufsätze. *Journal für die reine und angewandte Mathematik*, 53:281–283.

[Bottazzini et Gray, 2013] Bottazzini, U. et Gray, J. (2013). *Hidden Harmony : Geometric Fantasies*. Springer.

[Bourbaki, 2006a] Bourbaki, N. (2006a). *Éléments de mathématiques*, volume Éléments d'histoire des mathématiques. Springer.

[Bourbaki, 2006b] Bourbaki, N. (2006b). *Éléments de mathématiques*, volume Algèbre commutative, Chapitres 5 à 7. Springer.

[Brechenmacher, 2006] Brechenmacher, F. (2006). *Histoire du théorème de Jordan de la décomposition matricielle*. Thèse de doctorat, EHESS.

[Brechenmacher, 2008a] Brechenmacher, F. (2008a). La controverse de 1874 entre Camille Jordan et Leopold Kronecker. *Revue d'Histoire des Mathématiques*, 2(13).

[Brechenmacher, 2008b] Brechenmacher, F. (2008b). L'identité algébrique d'une pratique portée par la discussion sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des planètes (1766-1874). *Sciences et Techniques en Perspective*, IIe série, fasc. 1 (2007):5–85.

[Brechenmacher, 2010] Brechenmacher, F. (2010). Le «Journal de M. Liouville» sous la direction de Camille Jordan (1885-1922). *Bulletin de la Sabix. Société des amis de la Bibliothèque et de l'Histoire de l'École polytechnique*, (45):65–71.

[Brechenmacher, 2011] Brechenmacher, F. (2011). On Jordan's measurements. *arXiv preprint arXiv :1110.2351*.

[Brechenmacher, 2012] Brechenmacher, F. (2012). A History of Galois fields.

[Brechenmacher et Ehrhardt, 2010] Brechenmacher, F. et Ehrhardt, C. (2010). On the identities of algebra in the 19th century. *In On the Identities of Algebra in the 19th Century*, volume 7, pages 604–612.

[Bret, 2 13a] Bret, L. (1812-13a). Correpondance. Lettre de M. Bret, professeur à la faculté des sciences de l'académie de Grenoble, au rédacteur des Annales. *Annales de Mathématiques pures et appliquées*, 3:31–34.

[Bret, 2 13b] Bret, L. (1812-13b). Lettre de M. L. Bret, professeur à la faculté des sciences de l'académie de Grenoble. *Annales de Mathématiques pures et appliquées*, 3:369–371.

[Brioschi, 1854] Brioschi, F. (1854). Sur les fonctions de Sturm. *Nouvelles annales de mathématiques*, 13:71–80.

[Brioschi, 1856] Brioschi, F. (1856). Sur les séries qui donnent le nombre de racines réelles des équations algébriques à une ou à plusieurs inconnues. *Nouvelles annales de mathématiques*, 15:264–286.

[Brown, 1999] Brown, R. (1999). Fixed point theory. *In History of topology*, pages 271–299. Elsevier.

[Bérard, 2 13] Bérard (1812-13). Lettre de M. Bérard, principale et professeur de mathématiques au collège de Briançon. *Annales de Mathématiques pures et appliquées*, 3:97–98.

[Calais, 1998] Calais, J. (1998). *Éléments de théorie des groupes*. Presses Universitaires de France - PUF.

[Calais, 2006a] Calais, J. (2006a). *Eléments de la théorie des anneaux : Anneaux commutatifs*. Ellipses.

[Calais, 2006b] Calais, J. (2006b). *Extensions de corps : Théorie de Galois*. Ellipses.

[Calinger, 1996] Calinger, R. (1996). The Mathematics Seminar at the University of Berlin : Origins, Founding, and the Kummer-Weierstrass Years. *Vita Mathematica*, pages 153–176.

[Cantor *et al.*, 1991] Cantor, G., Meschkowski, H. et Nilson, W. (1991). *Georg Cantor Briefe*. Springer.

[Cauchy, 1815a] Cauchy, A. L. (1815a). Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suite des transpositions opérées entre les variables qu'elles renferment. *Journal de l'école polytechnique*, 17:29–112.

[Cauchy, 1815b] Cauchy, A. L. (1815b). Mémoires sur la détermination du nombre des racines réelles dans les équations algébriques. *In Oeuvres complètes*, volume 1 de *2*, pages 170–257. Gauthier-Villars.

[Cauchy, 1820] Cauchy, A. L. (1820). Mémoires sur la détermination du nombre des racines réelles dans les équations algébriques (Rapport par Legendre). *In Oeuvres complètes*, volume 2 de *2*, pages 187–193. Gauthier-Villars.

[Cauchy, 1821] Cauchy, A. L. (1821). *Cours d'analyse de l'Ecole royale polytechnique par m. Augustin-Louis Cauchy. 1ère partie. Analyse algébrique*. de l'Imprimerie royale.

[Cauchy, 1829] Cauchy, A. L. (1829). Mémoire sur la résolution des équations numériques et sur la théorie de l'élimination. *Oeuvres completes, 2ième s.*, IX:87–161.

[Cauchy, 1831a] Cauchy, A. L. (1831a). Mémoires de l'Académie des sciences de Turin. *Oeuvres completes, 2ième s.*, XV:182–261.

[Cauchy, 1831b] Cauchy, A. L. (1831b). Sur les rapports qui existent entre le calcul des résidus et le calcul des limites. *Bulletin des Sciences de Férussac*, 16:116–128.

[Cauchy, 1837a] Cauchy, A. L. (1837a). Calcul des indices des fonctions. *Journal de l'école polytechnique*, 15:176–229.

[Cauchy, 1837b] Cauchy, A. L. (1837b). Calcul des indices des fonctions. *Oeuvres completes, 2ième s.*, I:416–466.

[Cauchy, 1837c] Cauchy, A. L. (1837c). Extrait d'une lettre de M. Cauchy sur un mémoire publié par lui à Turin, le 16 juin 1833, et relatif aux racines des équations simultanées. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 4:672–675.

[Cauchy, 1837d] Cauchy, A. L. (1837d). Note sur un théorème de relatif aux racines des équations simultanées. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 5:6–8.

[Cauchy, 1855a] Cauchy, A. L. (1855a). Sur le dénombrement des racines qui, dans une équation algébrique ou transcendante, satisfont à des conditions données. *CR*, 40:1329–1335.

[Cauchy, 1855b] Cauchy, A. L. (1855b). Sur les compteurs logarithmiques. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 40:1009–1016.

[Cayley, 1845] Cayley, A. (1845). On the theory of linear transformations. *Cambridge Math. J*, 4(1845):1–16.

[Cayley, 1846a] Cayley, A. (1846a). Note sur les fonctions de M. Sturm. *Journal de Mathématiques Pures et Appliquées*, pages 297–299.

[Cayley, 1846b] Cayley, A. (1846b). On linear transformations. *Cambridge and Dublin Mathematical Journal*, 1(104–122):180.

[Cayley, 1848] Cayley, A. (1848). Nouvelles recherches sur les fonctions de M. Sturm. *Journal de Mathématiques Pures et Appliquées*, pages 269–274.

[Chemla *et al.*, 2016] Chemla, K., Chorlay, R. et Rabouin, D. (2016). *The Oxford Handbook of Generality in Mathematics and the Sciences*. Oxford University Press.

[Chorlay, 2003] Chorlay, R. (2003). Fonctions Implicites : de la Notion au Théorème. *Mnémosyne*, 18:15–58.

[Chorlay, 2007] Chorlay, R. (2007). La multiplicité des points de vue en Analyse élémentaire comme construit historique. *Histoire et enseignement des mathématiques : rigueurs, erreurs, raisonnements. Lyon : INRP*, pages 203–227.

[Crelle, 1835] Crelle, A. (1835). Die Sätze von Fourier und Sturm zur Theorie der algebraischen Gleichungen. *Journal für die reine und angewandte Mathematik*, 13:119–144.

[Davidoglou, 1901a] Davidoglou, A. (1901a). Sur le nombre des racines communes à plusieurs équations. *Compt. Rend. hebd*, 133:784–786.

[Davidoglou, 1901b] Davidoglou, A. (1901b). Sur le nombre des racines communes à plusieurs équations. *Compt. Rend. hebd*, 133:860–863.

[Davies, 1844] Davies, C. (1844). *Elements of algebra, including Sturm's theorem.* AS Barnes.

[de Foncenex, 1759] de Foncenex, F. (1759). Reflexions sur les quantités imaginaires. *Miscellanea Phil.-Math. Soc. Taurinensis, 1*, pages 113–146.

[de Fourcy, 1850] de Fourcy, L. (1850). *Leçons d'algèbre.* Monceaux-Hoyois.

[de Maizière, 1813] de Maizière (1812-1813). Analise élémentaire. Considérations propres à fournir, dans un grand nombre de cas, des limites extrêmes, très-approchées, des racines des équations numériques. *Annales de Gergonne*, 3:41–46.

[de Zeipel, 1860] de Zeipel, V. (1860). Démonstration d'un théorème de M. Cayley sur les relations entre les fonctions de Sturm. *Nouvelles annales de mathématiques*, 19:220–224.

[Dedekind, 1873] Dedekind, R. (1873). Anzeige von "P. Bachmann. *Die Lehre von der Kreistheilung und ihre Beziehungen zur Zahlentheorie". Literaturzeitung der Zeitschrift für Mathematik und Physik*, 18:14–24.

[der Technische Hochschule zu Berlin, 1910] der Technische Hochschule zu Berlin, P. (1910). Programm der Technische Hochschule zu Berlin für das Studienjahr 1910/11.

[Descartes, 1987] Descartes, R. (1987). *Descartes, Discours de la méthode plus la dioptrique les météores et la géométrie.* Tours, Fayard.

[Dhombres et Alvarez, 2011] Dhombres, J. et Alvarez, C. (2011). *Une histoire de l'imaginaire mathématique : vers le théorème fondamental de l'algèbre et sa démonstration par Laplace en 1795.* Hermann.

[Dhombres et Alvarez, 2013] Dhombres, J. et Alvarez, C. (2013). *Une histoire de l'invention mathématique : les démonstrations du théorème fondamental de l'algèbre dans le cadre de l'analyse réelle et de l'analyse complexe de Gauss à Liouville.* Hermann.

[Dieudonné, 1964] Dieudonné, J. (1964). *Calcul infinitesimal.* Hermann, The Macmillan Co.

[Dieudonné, 1976] Dieudonné, J. (1976). *Cours de géométrie algébrique. Vol. 1 : Aperçu historique sur le développement de la géométrie algébrique*, chapitre Cinquième époque, Développement et chaos (1866-1920). PUF.

[Dieudonné, 1978] Dieudonné, J. (1978). *Abrégé d'histoire des mathématiques.* Hermann.

[Dinca et Mawhin, 2009] Dinca, G. et Mawhin, J. (2009). Brouwer degree and applications. *Preprint. http ://www. ann. jussieu. fr/smets/ULM/Brouwer_ Degree_ and_ applications. pdf.*

[Dirichlet, 1889] Dirichlet, G. (1889). *Dirichlet Werke, Tome I.*

[Droniou, 2006] Droniou, J. (2006). Degrés topologiques et applications. *Université Montpellier II.*

[Dubourguet, 1 12] Dubourguet (1811-12). Analise élémentaire. Démonstration du principe qui sert de fondement à la théorie des équations. *Annales de Mathématiques pures et appliquées*, 2:338–340.

[Dubourguet, 2 13] Dubourguet (1812-13). Correspondance. *Annales de Mathématiques pures et appliquées*, 3:94–97.

[Dugac, 1973] Dugac, P. (1973). Eléments d'analyse de Karl Weierstrass. *Archive for History of Exact Sciences*, 10(1-2):41–174.

[Dugac, 1976] Dugac, P. (1976). Problèmes de l'histoire de l'analyse mathématique au XIXème siècle. Cas de karl weierstrass et de Richard Dedekind. *Historia Mathematica*, 3(1):5–19.

[Duminil et Jaulin, 2008] Duminil, M.-P. et Jaulin, A. (2008). *Aristote. Métaphysique.* Paris, Flammarion.

[Dyck, 1884] Dyck, W. (1884). On the "Analysis situs" of three-dimensional spaces. *In Report of the Fifty-fourth Meeting of the British Association for the Advancement of Science : Held at Montreal in August and September*, volume 1885.

[Dyck, 1888] Dyck, W. (1888). Beiträge zur Analysis situs. *Mathematische Annalen*, 32(4):457–512.

[Dyck, 1890] Dyck, W. (1890). Beiträge zur Analysis situs II. *Mathematische Annalen*, 37(2):273–316.

[Dyck, 1894] Dyck, W. (1894). Sur la détermination du nombre de racines communes à un système d'équations simultanées et sur le calcul de la somme des valeurs d'une fonction en ces points. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 113:1254–1257.

[Dyck, 1895a] Dyck, W. (1895a). Bemerkungen zu Kronecker's Theorie der Charakteristiken von Functionen-Systemen. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 4:94–95.

[Dyck, 1895b] Dyck, W. (1895b). Sur les racines communes à plusieurs équations. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, pages 34–36.

[Dyck, 1896] Dyck, W. (1896). Beiträge zur Potentialtheorie. *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Klasse der Bayerischen Akademie der Wissenschaften zu München*, 25:261–277.

[d'Alembert, 1746] d'Alembert, J. (1746). Recherches sur le calcul integral. *Hist. Acad. Sci. Berlin*, 2(1746):182–224.

[Eckes, 2011] Eckes, C. (2011). *Groupes, invariants et géométries dans l'œuvre de Weyl : Une étude des écrits de Hermann Weyl en mathématiques, physique mathématique et philosophie, 1910-1931.* Thèse de doctorat, Lyon 3.

[Edwards, 1978] Edwards, H. (1978). On the Kronecker Nachlass. *Historia Mathematica*, 5(4):419–426.

[Edwards, 1980] Edwards, H. (1980). The genesis of ideal theory. *Archive for History of Exact Sciences*, 23:321–378.

[Edwards, 1984] Edwards, H. (1984). *Galois Theory.* Springer.

[Edwards, 1987] Edwards, H. (1987). An Appreciation of Kronecker. *The Mathematical Intelligencer*, 9:28–35.

[Edwards, 1988] Edwards, H. (1988). Kronecker's place in history. *History and philosophy of modern mathematics, Aspray and Kitcher*, pages 139–144.

[Edwards, 1990] Edwards, H. (1990). *Divisor Theory*. Birkhauser Boston.

[Edwards, 1995] Edwards, H. (1995). Kronecker on the Foundations of Mathematics. *In From Dedekind to Gödel*, pages 45–52. Springer.

[Edwards, 2005] Edwards, H. (2005). *Essays In Constructive Mathematics*. Springer.

[Edwards *et al.*, 1982] Edwards, H., Neumann, O. et Purkert, W. (1982). Dedekinds "Bunte Bemerkungen" zu Kroneckers "Grundzüge". *Archive for History of Exact Sciences*, 27:49–85.

[Elias et Kamnitzer, 1973] Elias, N. et Kamnitzer, P. (1973). *La civilisation des mœurs*, volume 49. Calmann-Lévy Paris.

[Euler, 1751] Euler, L. (1751). Recherches sur les racines imaginaires des équations. *Mém. Acad. Roy. Sci. et belles-lettres*, 5:222–288.

[Euler, 1768] Euler, L. (1768). *Institutionum calculi integralis*.

[Ewald, 2005] Ewald, W. (2005). *From Kant to Hilbert Volume 1 : A Source Book in the Foundations of Mathematics*, volume 1. Oxford University Press on Demand.

[Fauvel et Gray, 1987] Fauvel, J. et Gray, J. (1987). *The History of mathematics : a reader*. Macmillan Education in association with The Open University (Basingstoke).

[Feuer, 1984] Feuer, L. S. (1984). *America's First Jewish Professor : James Joseph Sylvester at the University of Virginia*. American Jewish Archives.

[Fine et Rosenberger, 1997] Fine, B. et Rosenberger, G. (1997). *The fundamental theorem of algebra*. Springer.

[Fine, 1914] Fine, H. (1914). An unpublished theorem of Kronecker respecting numerical equations. *Bulletin of the American Mathematical Society*, 20(7):339–358.

[Fourier, 1830] Fourier, J. B. J. (1830). *Analyse des équations déterminées*, volume 1. Firmin Didot.

[Francoeur, 1838] Francoeur, L. B. (1838). *Algèbre supérieure*. Meline, Cans et compagnie.

[Frege, 1884] Frege, G. (1884). Die Grundlagen der Arithmetik.

[Freudenthal, 1971] Freudenthal, H. (1971). Did Cauchy plagiarize bolzano ? *Archive for History of exact Sciences*, 7(5):375–392.

[Frobenius, 1893] Frobenius, F. G. (1893). Gedächtnisrede auf Leopold Kronecker : aus den Abhandlungen der König. Preuss. Akademie der Wissenschaften zu Berlin vom Jahre 1893.

[Fürstenau, 1860] Fürstenau, E. (1860). *Darstellung der reellen Wurzeln algebraischer Gleichungen durch Determinanten der Coefficienten*. Elwert.

[G., 1935] G., F. (1935). Dyck, W. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 45:89–98.

[Gagneux, 1989] Gagneux, B. (1989). La règle des signes de Descartes : le long chemin d'une imprécision. *In Mathématiciens français du XVIIème siècle Descartes, Fermat, Pascal*, pages 129–165. Presse universitaire Blaise Pascal.

[Gauss, 1799] Gauss, C. F. (1799). Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse. PhD thesis, Universitat Helmstedt. *In Werke Tome III*, pages 1–30. K. Gesellschaft der Wissenschaften zu Göttingen.

[Gauss, 1800] Gauss, C. F. (1800). Zur metaphysik der mathematik. *In Werke*, volume 12, pages 57–61.

[Gauss, 1801] Gauss, C. F. (1801). Disquisitiones Arithmeticae. *In Werke Tome I.* K. Gesellschaft der Wissenschaften zu Göttingen.

[Gauss, 1816a] Gauss, C. F. (1816a). Demonstratio nova altera theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse. Comm. Recentiores. *In Werke Tome III*, pages 31–56. K. Gesellschaft der Wissenschaften zu Göttingen.

[Gauss, 1816b] Gauss, C. F. (1816b). Theorematis de resolubilitate functionum algebraicarum integrarum in factores reales Demonstratio tertia supplentum commentationis praecedentis. *In Werke Tome III*, pages 57–64. K. Gesellschaft der Wissenschaften zu Göttingen.

[Gauss, 1822] Gauss, C. F. (1822). Theoria Attractionis Corporum Sphaeroidicorum Ellipticorum Homogeneorum. *In Werke Tome V*, pages 279–286. Springer.

[Gauss, 1840] Gauss, C. F. (1840). Allgemeine Lehrsätze in Beziehung auf die im verkehrten Verhältnisse des Quadrats der Entfernung wirkenden Anziehungs-und Abstossungs-Kräfte. *In Werke Tome V*, pages 195–242. Springer.

[Gauss, 1849] Gauss, C. F. (1849). Beiträge zur Theorie der algebraischen Gleichungen. *In Werke Tome III*, pages 71–85. K. Gesellschaft der Wissenschaften zu Göttingen.

[Gauss, 1856] Gauss, C. F. (1856). Troisième démonstration de la possibilité de décomposer les fonctions algébriques entières en facteur réels ; d'après Gauss. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 15(1):134–139.

[Gauss, 1900] Gauss, C. F. (1900). Zur Geometrie der Lage für zwei Raumdimensionen. *Werke (Collected Works)*, 8:282–286.

[Gauthier, 2007] Gauthier, S. (2007). *La géométrie des nombres comme discipline (1890-1945).* Thèse de doctorat, Paris VI.

[Gauthier, 2002] Gauthier, Y. (2002). *Internal Logic : Foundations of Mathematics from Kronecker to Hilbert.* Springer.

[Gauthier, 2009] Gauthier, Y. (2009). Classical function theory and applied proof theory. *Int. J. Pure Appl. Math*, 56:223–233.

[Gauthier, 2013] Gauthier, Y. (2013). Kronecker in Contemporary Mathematics, General Arithmetic as a Foundational Programme. *Reports on mathematical logic*, (48):37–65.

[Gauthier, 2015] Gauthier, Y. (2015). Towards an Arithmetical Logic. *Studies in Universal Logic*.

[Gerstell, 1975] Gerstell, M. (1975). Prussian education and mathematics. *American Mathematical Monthly*, pages 240–245.

[Gilain, 1991] Gilain, C. (1991). Sur l'histoire du théorème fondamental de l'algèbre : théorie des équations et calcul intégral. *Archive for History of Exact Sciences*, 42:91–136.

[Girard, 1629] Girard, A. (1629). *Invention nouvelle en Algebre*. Amsterdam, Guillaume Ianffon Blaeuw.

[Gispert, 1991] Gispert, H. (1991). La Société Mathématique de France. *Cahiers d'histoire et de philosophie des sciences*, (34):13–180.

[Goldstein, 1999] Goldstein, C. (1999). Sur la question des méthodes quantitatives en histoire des mathématiques : le cas de la théorie des nombres en france (1870-1914). *Acta historiae rerum necnon technicarum*, 3:187–214.

[Goldstein *et al.*, 2007] Goldstein, C., Schappacher, N. et Schwermer, J. (2007). *The shaping of arithmetic after CF Gauss's Disquisitiones Arithmeticae*. Springer Science & Business Media.

[Gozard, 2009] Gozard, Y. (2009). *Théorie de Galois*. Ellipses.

[Grattan-Guinness, 1970] Grattan-Guinness, I. (1970). Bolzano, Cauchy and the "new analysis" of the early nineteenth century. *Archive for History of exact Sciences*, 6(5):372–400.

[Grattan-Guinness, 2005] Grattan-Guinness, I. (2005). *Landmark Writings in Western Mathematics 1640-1940*. Elsevier Science.

[Gray, 1997] Gray, J. (1997). Algebraic Geometry between Noether and Noether – a forgotten chapter in the history of Algebraic Geometry. *Revue d'Histoire des Mathématiques*, 3:1–48.

[Gray et Parshall, 2011] Gray, J. et Parshall, K. (2011). *Episodes in the history of modern algebra (1800-1950)*, volume 32. American Mathematical Soc.

[Hadamard, 1921] Hadamard, J. (1921). L'oeuvre mathématique de Poincaré. *Acta mathematica*, 38(1):203–287.

[Haffner, 2014a] Haffner, E. (2014a). D'un point de vue rigoureux et parfaitement général : pratique des mathématiques rigoureuses chez Richard Dedekind. *Philosophia Scientiae*, 18(1):131–156.

[Haffner, 2014b] Haffner, E. (2014b). *The " Science of Numbers " in action in Richard Dedekind's works : between mathematical explorations and foundational investigations*. Thèse de doctorat, Université Paris Diderot - Paris 7.

[Hancock, 1901] Hancock, H. (1901). *Mémoire sur les systèmes modulaires de Kronecker*. Gauthier-Villars.

[Harriot, 1631] Harriot, T. (1631). *Artis, analyticae praxis, algebraicis nova methodo resolvendas*. London.

[Hashagen, 2003] Hashagen, U. (2003). *Walther von Dyck (1856-1934) : Mathematik, Technik und Wissenschaftsorganisation an der TH München*, volume 47. Franz Steiner Verlag.

[Hattendorff, 1874] Hattendorff, K. (1874). *Die sturm'schen Functionen*. Carl Rümpler.

[Hawkins, 1977] Hawkins, T. (1977). Weierstrass and the theory of matrices. *Archive for history of exact sciences*, 17(2):119–163.

[Hawkins, 1980] Hawkins, T. (1980). The origins of modern theories of integration. *In From the Calculus to Set Theory 1630-1910*, pages 149–180. Princeton University Press.

[Hawkins, 2001] Hawkins, T. (2001). *Lebesgue's theory of integration : its origins and development*, volume 282. American Mathematical Soc.

[Heinz, 1959] Heinz, E. (1959). An elementary analytic theory of the degree of mapping in n-dimensional space. *J. Math. Mech*, 8(2):231–240.

[Hensel et Landsberg, 1902] Hensel, K. et Landsberg, G. (1902). *Theorie der algebraischen Funktionen einer Variablen : und ihre Anwendung auf algebraische Kurven und Abelsche Integrale*. B.G. Teubner.

[Hermite, 1852] Hermite, C. (1852). Sur l'extension du théorème de M. Sturm à un système d'équations simultanées équations simultanées. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 35:52–54.

[Hermite, 1853] Hermite, C. (1853). Remarques sur le théorème de M. Sturm. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 36:294–297.

[Hermite, 1856] Hermite, C. (1856). Extrait d'une lettre de Mr. Ch. Hermite de Paris à Mr. Borchardt de Berlin sur le nombre des racines d'une équation algébrique comprises entre des limites données. *Journal für die reine und angewandte Mathematik*, 52:39–51.

[Hermite, 1969] Hermite, C. (1969). Sur l'extension du theoreme de M. Sturm a un systeme d'equations simultanees. *Oeuvres de Charles Hermite*, 3:1–34.

[Hilbert, 1888] Hilbert, D. (1888). *Bericht über meine Reise vom 9ten März bis 7ten April 1888*. Numéro Cod. Ms. D. Hilbert 741 : 1. Manucrit inédit, Niedersächsische Staats- und Universitätsbibliothek Göttingen.

[Hirsch, 2012] Hirsch, M. (2012). *Differential topology*, volume 33. Springer Science & Business Media.

[Houzel, 1989] Houzel, C. (1989). D'Alembert et le théorème fondamental de l'algèbre. *In Jean D'Alembert, savant et philosophe : portrait à plusieurs voix*, pages 351–360. Paris, édition des archives contemporaines.

[Hudde, 1661] Hudde, J. (1661). *Epistola prima de reductione æquationum.* Elsevier.

[Jacobi, 1835] Jacobi, C. G. J. (1835). Theoremata nova algebraica circa systema duarum aequationum, inter duas variabiles propositarum. *Journal für die reine und angewandte Mathematik*, 14:281–288.

[Jacobi, 1846] Jacobi, C. G. J. (1846). Über die Darstellung einer Reihe gegebner Werthe durch eine gebrochne rationale Function. *Journal für die reine und angewandte Mathematik*, 30:127–156.

[Jacobi, 1857] Jacobi, C. G. J. (1857). Über einen algebraischen Fundamentalsatz und seine Anwendungen [On a Fundamental Theorem in Algebra and Its Applications]. *Journal für die reine und angewandte Mathematik*, 53:275–280.

[Jacobi, 1881] Jacobi, C. G. J. (1881). Gesammelte Werke, Band 1. *Königlich Preussischen Akademie der Wissenschaften zu Berlin.*

[James, 1999] James, I. (1999). *History of topology.* Elsevier.

[Jordan, 1892] Jordan, C. (1892). Remarques sur les intégrales définies. *Journal de mathématiques pures et appliquées*, pages 69–100.

[Jordan, 1893] Jordan, C. (1893). *Cours d'analyse de l'École polytechnique*, volume 1. Gauthier-Villars et fils.

[Kant, 1980] Kant, I. (1980). *Critique de la raison pure (par Alexandre J.-L. Delamarre et François Marty).* Gallimard.

[Kant, 1991] Kant, I. (1991). *Correspondance.* Gallimard.

[Karp et Schubring, 2014] Karp, A. et Schubring, G. (2014). *Handbook on the history of mathematics education.* Springer.

[Klein, 1926] Klein, F. (1926). *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, volume 24. Springer-Verlag.

[Kneser, 1881] Kneser, A. (1881). *Algebraische Gleichungen.* Numéro Cod. Ms. A. Kneser B2. Manucrit inédit, Niedersächsische Staats- und Universitätsbibliothek Göttingen.

[Kneser, 1888] Kneser, A. (1888). Arithmetische Begründung einiger algebraischer Fundamentalsätze. *Journal für die reine und angewandte Mathematik*, 102:20–55.

[Kneser, 1925] Kneser, A. (1925). Leopold Kronecker. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 33:210–228.

[Knoblauch, 1913] Knoblauch, J. (1913). *Grundlagen der differentialgeometrie.* BG Teubner.

[Kolmogorov *et al.*, 2001] Kolmogorov, A., Yushvevich, A. et Shenitzer, A. (2001). *Mathematics of the 19th century : mathematical logic, algebra, number theory, probability theory*, volume 1. Springer Science & Business Media.

[Koralek, 1852] Koralek, M. (1852). Formules des fonctions de M. Sturm pour les équations du deuxième, du troisième et du quatrième degré. *Nouvelles annales de mathématiques*, 11:333–335.

[Kronecker, ] Kronecker, L. Werke, édité par Kurt Hensel, volumes 1–5, BG Teubner, Leipzig et l'Académie de Berlin, 1895–1930. Réimpression par Chelsea Publishing Compagny, New York (1969) et par l'American Mathematical Society, Providence, Rhode Island (2010).

[Kronecker, 1853] Kronecker, L. (1853). Über die algebraisch auflösbaren Gleichungen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 365–374. Werke IV, 1–12.

[Kronecker, 1857] Kronecker, L. (1857). Zwei Sätze über Gleichungen mit ganzzahligen Coëfficienten. *Journal für die reine und angewandte Mathematik*, 53:173–175.

[Kronecker, 1861] Kronecker, L. (1861). Antrittsrede von L. Kronecker bei der Aufnahme in die Akademie der Wissenschaften am 4. Juli 1861. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 637–642. Werke V, 387–393.

[Kronecker, 1865] Kronecker, L. (1865). Über einige Interpolationsformeln für ganze Functionen mehrer Variabeln. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 686–691. Werke I, 133–141.

[Kronecker, 1866] Kronecker, L. (1866). Ueber bilineare formen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 597–612. *Journal für die reine und angewandte Mathematik* 68 : 273–285. Werke I, 143–141.

[Kronecker, 1868] Kronecker, L. (1868). Ueber Schaaren von quadratischer Formen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 339–346. Werke I, 163–173.

[Kronecker, 1869a] Kronecker, L. (1869a). Bemerkungen zur Determinant-Theorie. *Journal für die reine und angewandte Mathematik*, 72:152–175. Werke I, 235–269.

[Kronecker, 1869b] Kronecker, L. (1869b). Sur le théorème de Sturm. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 68:1078–1082. Werke I, 229–234.

[Kronecker, 1869c] Kronecker, L. (1869c). Über Systeme von Functionen mehrer Variabeln. Erste Abhandlung. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 159–193. Werke I, 177–212.

[Kronecker, 1869d] Kronecker, L. (1869d). Über Systeme von Functionen mehrer Variabeln. Zweite Abhandlung. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 688–698. Werke I, 215–226.

[Kronecker, 1872] Kronecker, L. (1872). Zur algebraischen Theorie der quadratischen Formen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 490–504. Werke I, 283–301.

[Kronecker, 1873a] Kronecker, L. (1873a). *Cours du semestre d'hiver 1872/73, Theorie der Algebraischen Gleichungen.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1873b] Kronecker, L. (1873b). Über die verschiedenen Sturm'schen Reihen und ihre gegenseitigen Beziehungen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 117–154. Werke I, 305–348.

[Kronecker, 1874a] Kronecker, L. (1874a). Sur les faisceaux de formes quadratiques et bilinéaires. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 78:1181–1182. Werke I, 415–419.

[Kronecker, 1874b] Kronecker, L. (1874b). Uber die congruenten Transformationen der bilinearen Formen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 397–447. Werke I, 423–483.

[Kronecker, 1874c] Kronecker, L. (1874c). Über Schaaren von quadratischen und bilinearen Formen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 59–76, 149–156, 206–232. Werke I, 373–413.

[Kronecker, 1876] Kronecker, L. (1876). *Cours du semestre d'hiver 1875/76, Anwendung der Analysis des Unendlichen auf die Zahlentheorie.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1878a] Kronecker, L. (1878a). Notiz über Potenzreihen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 197–201. Werke V, 195–201.

[Kronecker, 1878b] Kronecker, L. (1878b). Über die Charakteristik von Funktionen-Systemen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 95–121. Werke II, 71–82.

[Kronecker, 1878c] Kronecker, L. (1878c). Über Sturm'sche Functionen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 95–121. Werke II, 39–70.

[Kronecker, 1879] Kronecker, L. (1879). *Cours du semestre d'hiver 1878/79, Theorie der algebraischen Gleichungen.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1881a] Kronecker, L. (1881a). *Cours du semestre d'hiver 1880/81, Theorie der Algebraischen Gleichungen.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1881b] Kronecker, L. (1881b). Ueber die Discriminante algebraischer Functionen einer Variabeln. *Journal für die reine und angewandte Mathematik*, 91:301–334. Werke II, 193–236.

[Kronecker, 1882a] Kronecker, L. (1882a). *Cours du semestre d'hiver 1881/82, Höhere Zahlentheorie.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1882b] Kronecker, L. (1882b). Grundzüge einer arithmetischen Theorie der algebraischen Grössen. *Journal für die reine und angewandte Mathematik*, 92:1–122. Werke II, 237–388.

[Kronecker, 1883] Kronecker, L. (1883). *Cours du semestre d'hiver 1882/83, Vorlesungen über Zahlentheorie.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1884a] Kronecker, L. (1884a). Beweis des Reciprocitätsgesetzes für die quadratischen Reste. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 519–537. Werke II, 497–522.

[Kronecker, 1884b] Kronecker, L. (1884b). Die Periodensysteme von Functionen reeller Variabeln. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 1071–1080. Werke IIIa, 31–46.

[Kronecker, 1884c] Kronecker, L. (1884c). Näherungsweise ganzzahlige Auflösung linearer Gleichungen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 1179–1193, 1271–1299. Werke IIIa, 47–109.

[Kronecker, 1885a] Kronecker, L. (1885a). *Cours du semestre d'hiver 1884/85, Theorie der Algebraischen Gleichungen, don du Dr Thiel.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1885b] Kronecker, L. (1885b). *Cours du semestre d'hiver 1884/85, Theorie der Algebraischen Gleichungen, Teil I.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1885c] Kronecker, L. (1885c). *Cours du semestre d'été 1885, Theorie der Algebraischen Gleichungen, Teil II.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1886] Kronecker, L. (1886). *Cours du semestre d'hiver 1885/86, Über die Affecte der Gleichungen, welche in der Theorie der elliptischen Funktionen vorkommen.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1887a] Kronecker, L. (1887a). *Cours du semestre d'hiver 1886/87, Theorie der Algebraischen Gleichungen.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1887b] Kronecker, L. (1887b). *Cours du semestre d'hiver 1886/87, Theorie der Algebraischen Gleichungen.* Manucrit inédit, Staats- und Universitätsbibliothek Göttingen.

[Kronecker, 1887c] Kronecker, L. (1887c). Ein Fundamentalsatz der allgemeinen Arithmetik. *Journal für die reine und angewandte Mathematik*, 100:490–510. Werke IIIa, 209–240.

[Kronecker, 1887d] Kronecker, L. (1887d). Über den Zahlbegriff. *Journal für die reine und angewandte Mathematik*, 101:337–355. Werke IIIa, 249–274.

[Kronecker, 1889a] Kronecker, L. (1889a). *Cours du semestre d'été 1889, Theorie der einfachen und mehrfachen Integrale.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1889b] Kronecker, L. (1889b). *Cours du semestre d'hiver 1888/89, Theorie der Algebraischen Gleichungen.* Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1889c] Kronecker, L. (1889c). Über symmetrische Systeme. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 349–362. Werke IIIa, 292–314.

[Kronecker, 1891a] Kronecker, L. (1891a). *Cours du semestre d'hiver 1890/91, Theorie der Algebraischen Gleichungen*. Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Kronecker, 1891b] Kronecker, L. (1891b). Sur le nombre de racines communes à plusieurs équations simultanées. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 113:1006–1012. Werke IIIb, 207–215.

[Kronecker, 1894] Kronecker, L. (1894). *Vorlesungen über die Theorie der einfachen und der vielfachen Integrale*. Teubner, Leipzig.

[Kronecker, 1901] Kronecker, L. (1901). *Vorlesungen über Mathematik*, volume 1. Teubner, Leipzig.

[Kronecker, 1903] Kronecker, L. (1903). *Vorlesungen über Mathematik*, volume 2. Teubner, Leipzig.

[Lacroix, 1819] Lacroix, S. F. (1819). *Traité élémentaire de calcul différentiel et du calcul intégral*. Courcier, 2e éd. édition.

[Lacroix, 1825a] Lacroix, S. F. (1825a). *Complément des éléments d'algèbre à l'usage de l'Ecole centrale des Quatre Nations*. Bachelier, 5e éd. édition.

[Lacroix, 1825b] Lacroix, S. F. (1825b). *Eléments d'Algébre*. Bachelier, 14e éd. édition.

[Lagrange, 1772] Lagrange, J. (1772). *Sur la forme des racines imaginaires des équations*. Académie royale des sciences et belles lettres.

[Lagrange, 1771] Lagrange, J. L. (1771). Réflections sur la résolution algébrique des équations.

[Lagrange, 1826] Lagrange, J. L. (1826). *Traité de la résolution des équations numériques de tous les degrés, avec des notes sur plusieurs points de la théorie des équations algébriques. Poinsot*. Bachelier.

[Lagrange, 1867] Lagrange, J. L. (1867). *Leçons élémentaires sur les mathématiques données à l'École Normale en 1795*, volume 7. ID., Oeuvres completes, a cura di J.-A. Serret, Gauthier-Villars, Paris.

[Laguerre, 1883] Laguerre, E. (1883). Mémoire sur la théorie des équations numériques. *Journal de Mathématiques Pures et Appliquées*, 9(3):99–146.

[Laguerre, 1884] Laguerre, E. (1884). Sur quelques points de la théorie des équations numériques. *Acta Mathematica*, 4:97–120.

[Lamla, 1911] Lamla, E. (1911). *Geschichte des Mathematischen Vereins an der Universität Berlin : 1861-1911*.

[Landsberg, 1911] Landsberg, G. (1911). Beiträge zur Topologie geschlossener Kurven mit Knotenpunkten und zur Kroneckerschen Charakteristikentheorie. *Mathematische Annalen*, 70(4):563–579.

[Laplace, 1835] Laplace, P. S. (1835). *Exposition du système du monde*. Bachelier.

[Laz, 1993] Laz, J. (1993). *Bolzano critique de Kant*. Vrin.

[Lebesgue, 1924] Lebesgue, H. (1924). Remarques sur les deux premières démonstrations du théorème d'Euler relatif aux polyèdres. *Bulletin de la Société Mathématique de France*, 52:315–336.

[Lebesgue, 1955] Lebesgue, H. (1955). L'œuvre mathématique de Vandermonde. *L'Enseignement mathématique*, 1:203–223.

[Liouville, 1836] Liouville, J. (1836). Démonstration d'un théorème dû à M. Sturm, et relatif à une classe de fonctions transcendantes. *Journal de Mathématiques Pures et Appliquées*, 1(1):269–277.

[Liouville, 1839] Liouville, J. (1839). Sur le principe fondamental de la théorie des équations algébriques. *Journal de Mathématiques Pures et Appliquées*, 4(1):501–508.

[Liouville, 1840] Liouville, J. (1840). Addition à la Note sur le principe fondamental de la théorie des équations algébriques. *Journal de Mathématiques Pures et Appliquées*, 5(1):31–33.

[Liouville, 1841] Liouville, J. (1841). Mémoire sur quelques propositions générales de géométrie et sur la théorie de l'élimination dans les équations algébriques. *Journal de Mathématiques Pures et Appliquées*, pages 345–411.

[Liouville et Sturm, 1836] Liouville, J. et Sturm, C. (1836). Démonstration d'un théorème de M. Cauchy, relatif aux racines imaginaires des équations. *Journal de Mathématiques Pures et Appliquées*, 1(1):278–289.

[Liouville et Sturm, 1837] Liouville, J. et Sturm, C. (1837). Note sur un théorème de M. Cauchy relatif aux racines des équations simultanées. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 4:720–724.

[Lipschitz et Scharlau, 1986] Lipschitz, R. et Scharlau, W. (1986). *Briefwechsel mit Cantor, Dedekind, Helmholtz, Kronecker, Weierstrass und anderen*. Springer.

[Mawhin, 2000] Mawhin, J. (2000). Poincaré's early use of Analysis situs in nonlinear differential equations : Variations around the theme of Kronecker's integral. *Philosophia Scientiae*, 4:103–143.

[Mayer et Choquet, 1832] Mayer et Choquet (1832). *Traité élémentaire d'algèbre*. Bachelier, Paris.

[Mehrtens *et al.*, 1981] Mehrtens, H., Bos, H. et Schneider, I. (1981). *Social history of nineteenth century mathematics*. Springer.

[Meschkowski et Cantor, 1967] Meschkowski, H. et Cantor, G. (1967). *Probleme des Unendlichen*. Springer.

[Meyer, 1934] Meyer, W. F. (1898–1934). Encyklopädie der Mathematischen Wissenschaften. *Leipzig, Münich and Vienna*.

[Meyer, 1904] Meyer, W. F. (1904). *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*. Teubner.

[Michel, 2002] Michel, A. (2002). Thèses d'existence et travail mathématique. *In* Serfati, M., éditeur : *De la méthode - Recherches en histoire et philosophie des mathématiques*. PU de Franche-Comté.

[Mignosi, 1925] Mignosi, G. (1925). Teorema di Sturm e sue estensioni. *Rendiconti del Circolo Matematico di Palermo*, 49(1):1–159.

[Milnor, 1997] Milnor, J. (1997). *Topology from the differentiable viewpoint*. Princeton University Press.

[Moigno, 1840a] Moigno, F. (1840a). *Leçons de calcul différentiel et de calcul intégral*. Bachelier.

[Moigno, 1840b] Moigno, F. (1840b). Note sur la détermination du nombre des racines réelles ou imaginaires d'une équation numérique comprises entre les limites données. - Théorèmes de Rolle, de Budan ou de Fourier, de Descartes, de Sturm et de Cauchy. *Journal de Mathématiques Pures et Appliquées*, 5(1):75–94.

[Molk, 1884] Molk, J. (1884). Sur une notion qui comprend celle de la divisibilité et sur la théorie générale de l'élimination. *Acta mathematica*, 6:1–166.

[Molk, 1904] Molk, J. (1904). *Encyclopédie des sciences mathématiques, pure et appliquées : v. 1. Arithmétique*, volume 1. Gauthier-Villars.

[Molk, 1995] Molk, J. (1995). *Encyclopédie des sciences pures et appliquées, Tome I*. J. Gabay.

[Nabonnand, 1999] Nabonnand, P. (1999). *La Correspondance entre Henri Poincaré et Gösta Mittag-Leffler : Avec en annexes les lettres échangées par Poincaré avec Fredholm, Gyldén et Phragmén*. Springer Science & Business Media.

[Nabonnand, 2008] Nabonnand, P. (2008). La théorie des Würfe de von Staudt–Une irruption de l'algèbre dans la géométrie pure. *Archive for history of exact sciences*, 62(3):201–242.

[Nabonnand, 2011] Nabonnand, P. (2011). G.K.C. von Staudt Geometrie der Lage Traduction et présentation par Philippe Nabonnand.

[Nabonnand et Rollet, 2009] Nabonnand, P. et Rollet, L. (2009). 1842-1927 : les Nouvelles annales de mathématiques, journal des candidats aux écoles spéciales. *In Séminaire d'histoire des mathématiques de l'Institut Henri Poincaré*, Paris, France.

[Nagumo, 1951] Nagumo, M. (1951). A theory of degree of mapping based on infinitesimal analysis. *American Journal of Mathematics*, pages 485–496.

[Netto, 1876] Netto, E. (1876). *Die vier Gauss'schen Beweise für die Zerlegung ganzer algebraischer Functionen in reelle Factoren oder zweiten Grades*. Leipzig, Verlag von Wilhelm Engelmann.

[Netto, 1879] Netto, E. (1879). Beweis des Wurzelexistenz algebraischer Gleichungen. *Journal für die reine und angewandte Mathematik*, 88:16–21.

[Netto, 1882] Netto, E. (1882). Zur Theorie der Discriminanten. *Acta Mathematica*, 1:371–399.

[Netto, 1900] Netto, E. (1900). *Vorlesungen über algebra*. Numéro vol 1 et 2. Leipzig, B.G. Teubner.

[Neumann, 2007] Neumann, O. (2007). Divisibility Theories in the Early History of Commutative Algebra and the Foundations of Algebraic Geometry. *Episodes in the History of Modern Algebra (1800-1950)*, 32:73–105.

[Noether, 1898] Noether, M. (1898). Sylvester, J. J. *Mathematische Annalen*, 50(1):133–156.

[Nouvelles-Annales, 1842] Nouvelles-Annales (1842). Théorèmes à démontrer. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 1:166–167.

[Ostrowski, 1920] Ostrowski, A. (1920). *Nachrichten der Gesellschaft der Wissenschaften*, chapitre Über den ersten und vierten Gauss'schen Beweis des Fundamentalsatzes der Algebra. Göttinden.

[Ostrowski, 1976] Ostrowski, A. (1976). On a theorem by Kronecker. *Aequationes Mathematicae*, 14:159–166.

[Outerelo et Ruiz, 2009] Outerelo, E. et Ruiz, J. (2009). *Mapping degree theory*, volume 108. American Mathematical Society Washington.

[Padovani et Pont, 2009] Padovani, F. et Pont, J.-C. (2009). *Collected Works of Charles François Sturm*. Springer.

[Parshall, 1998] Parshall, K. H. (1998). *James Joseph Sylvester : life and work in letters*. Oxford University Press.

[Paulsen, 1902] Paulsen, F. (1902). *Die deutschen Universitäten und das Universitätsstudium*. A. Asher & Company.

[Paulsen, 1908] Paulsen, F. (1908). *German education : Past and present*. TF Unwin.

[Penchèvre, 2006] Penchèvre, E. (2006). *Histoire de la théorie de l'élimination*. Thèse de doctorat, Paris VII.

[Penchèvre, 2007] Penchèvre, E. (2007). La théorie arithmétique des grandeurs algébriques de Kronecker. *actes du colloque "Journée en l'honneur de Christian Houzel" à l'IHP (Paris) le 23 novembre 2007 (http ://www.penchevre.fr/kro.pdf)*. À paraître.

[Perfect, 1991] Perfect, H. (1991). Leopold Kronecker : A great Gentleman in science. *Mathematical Spectrum*, 24(1):5–6.

[Perott, 1886] Perott, J. (1886). Démonstration du théorème fondamental de l'algèbre. *Journal für die reine und angewandte Mathematik*, 99:141–146.

[Petri et Schappacher, 2004] Petri, B. et Schappacher, N. (2004). From Abel to Kronecker : episodes from 19th century algebra. *In The Legacy of Niels Henrik Abel*, pages 227–266. Springer.

[Petri et Schappacher, 2007] Petri, B. et Schappacher, N. (2007). On arithmetization. *In The Shaping of Arithmetic after CF Gauss's Disquisitiones Arithmeticae*, pages 343–374. Springer.

[Petrova, 1973] Petrova, S. (1973). From the history of the analytic proofs of the fundamental theorem of algebra. *History and methodology of the natural sciences*, XIV:167–172.

[Phragmén, 1892] Phragmén, E. (1892). Sur une extension du théorème de Sturm. *CR Paris*, 114:3.

[Picard, 1889] Picard, E. (1889). Sur le nombre des racines communes à plusieurs équations simultanées. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 8:5–13.

[Picard, 1891a] Picard, E. (1891a). Du nombre de racines communes à plusieurs équations simultanées. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 113:1012–1014.

[Picard, 1891b] Picard, E. (1891b). Sur la recherche du nombre de racines communes à plusieurs équations simultanées. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 113:669–672.

[Picard, 1891c] Picard, E. (1891c). Sur le nombre de racines communes à plusieurs équations simultanées. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 113:356–358.

[Picard, 1891d] Picard, E. (1891d). *Traité d'analyse*. Numéro vol. 1 de Cours de la Faculté des Sciences de Paris. Gauthier-Villars.

[Picard, 1892] Picard, E. (1892). Sur le nombre de racines communes à plusieurs équations simultanées. *Journal de Mathématiques Pures et Appliquées*, 8(4):5–24.

[Picard, 1893] Picard, E. (1893). *Traité d'analyse*. Numéro vol. 2 de Cours de la Faculté des Sciences de Paris. Gauthier-Villars.

[Picard, 1896] Picard, E. (1896). *Traité d'analyse*. Numéro vol. 3 de Cours de la Faculté des Sciences de Paris. Gauthier-Villars.

[Picard, 1901] Picard, E. (1901). *Traité d'analyse*. Numéro vol. 1. Gauthier-Villars, 2e éd. édition.

[Picard, 1904] Picard, E. (1904). *Traité d'analyse*. Numéro vol. 2. Gauthier-Villars, 2e éd. édition.

[Picard, 1916] Picard, E. (1916). *L'Histoire des sciences : et les prétentions de la science allemande*, volume 2. Perrin.

[Pierobon, 2003] Pierobon, F. (2003). Kant et les mathématiques.

[Pierre, 2002] Pierre, W. (2002). Les philosophes et la science. *Paris, Gallimard*.

[Poincaré, 1883] Poincaré, H. (1883). Sur les fonctions $\Theta$. *Bulletin de la Société Mathématique de France*, 11:129–134.

[Poincaré, 1884] Poincaré, H. (1884). Mémoires et observations. Sur certaines solutions particulières du probléme des trois corps. *Bulletin Astronomique, Serie I*, 1:65–74.

[Poincaré, 1885] Poincaré, H. (1885). Sur l'équilibre d'une masse fluide animée d'un mouvement de rotation. *Acta mathematica*, 7(1):259–380.

[Poincaré, 1886] Poincaré, H. (1886). Sur les courbes definies par les equations differentielles (IV partie). *J. Math. Pures Appl.(4)*, 2:151–217.

[Poincaré, 1895] Poincaré, H. (1895). *Analysis situs*.

[Poincaré, 1912a] Poincaré, H. (1912a). Pourquoi l'espace a trois dimensions. *Revue de Métaphysique et de Morale*, 20(4):483–504.

[Poincaré, 1912b] Poincaré, H. (1912b). Sur un théoreme de géométrie. *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 33(1):375–407.

[Poincaré, 1986] Poincaré, H. (1986). La correspondance d'Henri Poincaré avec des mathématiciens de A à H. *Cahiers du séminaire d'histoire des mathématiques*, 7:59–219.

[Poincaré, 1989] Poincaré, H. (1989). La correspondance avec des mathématiciens de J à Z. *Cahiers du séminaire d'histoire des mathématiques*, 10:83–229.

[Pont, 1974] Pont, J.-C. (1974). *La topologie algébrique des origines à Poincaré*. Presses Universitaires de France, Paris.

[Pyenson, 1983] Pyenson, L. (1983). Neohumanism and the persistence of pure mathematics in Wilhelmian Germany. *Memoirs of the American Philosophical Society Philadelphia, Pa*, 150:1–136.

[Remmert, 1991] Remmert, R. (1991). The Fundamental Theorem of Algebra. *In Numbers*, pages 97–122. Springer.

[Rollet et Nabonnand, 2002] Rollet, L. et Nabonnand, P. (2002). Une bibliographie mathématique idéale ? Le Répertoire Bibliographique des Sciences Mathématiques. *La Gazette des Mathématiciens*, 92:11–25.

[Rosenhain, 1846] Rosenhain, J. G. (1846). Neue Darstellung der Resultante der Elimination von z aus zwei algebraischen Gleichungen f (z)= 0 und...(z)= 0 vermittelst der Werthe welche die Functionen f (z) und...(z) für gegebne Werthe von z annehmen. *Journal für die reine und angewandte Mathematik*, 30:157–165.

[Rouché et Lévy, 1900] Rouché, E. et Lévy, L. (1900). *Analyse infinitésimale à l'usage des ingénieurs*, volume 1. Gauthier-Villars.

[Rowe, 1988] Rowe, D. (1988). Gauss, Dirichlet, and the law of biquadratic reciprocity. *The Mathematical Intelligencer*, 10(2):13–26.

[Rowe, 2003] Rowe, D. (2003). Mathematical Schools, Communities, and Networks. *Cambridge history of science*, 5:113–132.

[Rowe, 2005] Rowe, D. (2005). Hilbert's early career : Encounters with allies and rivals. *The Mathematical Intelligencer*, 27(1):72–82.

[Runge, 1881] Runge, C. (1881). *Algebr[aische] Gleich[ungen].* Numéro Cod. Ms. C. Runge 52. Manucrit inédit, Niedersächsische Staats- und Universitätsbibliothek Göttingen.

[Russ, 2004] Russ, S. (2004). *The mathematical works of Bernard Bolzano.* OUP Oxford.

[Salmon, 1885] Salmon, G. (1885). *Lessons introductory to the modern higher algebra.* Hodges, Figgis, and Co.

[Schappacher et Volkert, 1997] Schappacher, N. et Volkert, K. (1997). Heinrich Weber ; un mathématicien à Strasbourg, 1895-1913. *L'Ouvert*, pages 1–18.

[Scholz, 1999] Scholz, E. (1999). The concept of manifold, 1850–1950. *In History of topology*, pages 25–64. Elsevier.

[Schubring, 2013] Schubring, G. (2013). *Hermann Günther Grassmann (1809-1877) : visionary mathematician, scientist and neohumanist scholar*, volume 187. Springer Science & Business Media.

[Sebestik, 1992] Sebestik, J. (1992). *Logique et mathématique chez Bernard Bolzano.* Vrin.

[Sebestik et Bolzano, 1964] Sebestik, J. et Bolzano, B. (1964). Bernard Bolzano et son Mémoire sur le théorème fondamental de l'Analyse. *Revue d'histoire des sciences et de leurs applications*, pages 129–164.

[Serres, 1813] Serres (1812-1813). Correspondance à propos de l'article : Analise élémentaire. Considérations propres à fournir, dans un grand nombre de cas, des limites extrêmes, très-approchées, des racines des équations numériques. *Annales de Gergonne*, 3:291–292.

[Serret, 1849] Serret, J. A. (1849). *Cours d'algèbre supérieure.* Mallet-Bachelier.

[Shields, 1989] Shields, A. (1989). Lejeune Dirichlet and the Birth of Analytic Number Theory : 1837-1839. *The Mathematical Intelligencer*, 11(4):7–11.

[Siegberg, 1981] Siegberg, H. (1981). Some historical remarks concerning degree theory. *The American Mathematical Monthly*, 88(2):125–139.

[Sinaceur, 1988] Sinaceur, H. (1988). Deux moments dans l'histoire du théorème d'algèbre de CH. F. Sturm. *Revue d'histoire des sciences*, 41(2):99–132.

[Sinaceur, 1991] Sinaceur, H. (1991). *Corps et Modèles : Essai Sur L'histoire de L'algèbre Réelle.* Mathésis (Paris). Isd.

[Sinaceur, 1992] Sinaceur, H. (1992). Cauchy, Sturm et les racines des équations. *Revue d'histoire des sciences*, 45(1):51–68.

[Smadja, 2010] Smadja, I. (2010). How Discrete Patterns Emerge from Algorithmic Fine-Tuning : A Visual Plea for Kroneckerian Finitism. *Topoi*, 29(1):61–75.

[Smith, 1929] Smith, D. (1929). *A Source Book in Mathematics.* McGraw-Hill.

[Smithies, 1997] Smithies, F. (1997). *Cauchy and the creation of complex function theory.* Cambridge University Press.

[Society, 1908] Society, R. (1908). *Catalogue of Scientific Papers : subject index vol. 1. Pure Mathematics*. University Press.

[Stäckel, 1893] Stäckel, P. (1893). Ueber Systeme von Functionen reeller Variabeln. *Journal für die reine und angewandte Mathematik*, 112:141–146.

[Stewart, 2003] Stewart, I. (2003). *Galois Theory, third edition*. Chapman and Hall/CRC.

[Struik, 1969] Struik, D. (1969). *A Source Book in Mathematics 1200-1800*. Harvard University Press, Cambridge.

[Sturm, 1829a] Sturm, C. (1829a). Analyse d'un mémoire sur la résolution des équations numériques. *Bulletin de Férussac*, 11:419–422.

[Sturm, 1829b] Sturm, C. (1829b). Extrait d'un mémoire présenté à l'Académie par M. Ch. Sturm. *Bulletin de Férussac*, 11:422–425.

[Sturm, 1835] Sturm, C. (1835). Mémoire sur la résolution des équations numériques. *Mémoire présenté par divers Savants étrangers à l'Acad. roy. sc., section math.phys.*, 6:273–318.

[Sturm, 1836] Sturm, C. (1836). Autres démonstrations du même théorème. *Journal de Mathématiques Pures et Appliquées*, 1(1):290–308.

[Sturm, 1842] Sturm, C. (1842). Démonstration d'un théorème d'algèbre de M. Sylvester. *Journal de Mathématiques Pures et Appliquées*, 7:356–368.

[Sturm *et al.*, 1901] Sturm, C., Prouhet, E., Laurent, H. et de Saint-Germain, A. (1901). *Cours d'analyse de l'École polytechnique*, volume 2. Gauthier-Villars.

[Sylvester, 1839] Sylvester, J. J. (1839). On rational derivation from equations of coexistence, that is to say, a new and extended theory elimination. *The London and Edinburgh Philosophical Magazine and Journal of Science*, 15(98):428–435.

[Sylvester, 1841] Sylvester, J. J. (1841). On the relation of Sturm's auxiliary functions to the roots of an algebraic equation. *Plymouth British Association Report*, pages 23–24.

[Sylvester, 1844] Sylvester, J. J. (1844). On the Existence of absolute criteria for determining the roots of numerical equations. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 25(168):442–445.

[Sylvester, 1852] Sylvester, J. J. (1852). A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 4(23):138–142.

[Sylvester, 1853a] Sylvester, J. J. (1853a). Note on a remarkable modification of Sturm's theorem. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 6:14–20.

[Sylvester, 1853b] Sylvester, J. J. (1853b). Nouvelle méthode pour trouver une limite supérieure et une limite inférieure des racines réelles d'une équation algébrique quelconque. *Nouvelles annales de mathématiques*, 12:329–336.

[Sylvester, 1853c] Sylvester, J. J. (1853c). On a remarkable modification of Sturm's theorem. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 5:446–456.

[Sylvester, 1853d] Sylvester, J. J. (1853d). On a theory of the syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraical common measure. *Philosophical transactions of the Royal Society of London*, pages 407–548.

[Sylvester, 1853e] Sylvester, J. J. (1853e). On the explicit values of Sturm's quotients. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 6:293–296.

[Sylvester, 1853f] Sylvester, J. J. (1853f). On the expressions for the quotients which appear in the application of Sturm's method to the discovery of the real roots of an equation. *Hull British Association Report*, pages 1–3.

[Sylvester, 1853g] Sylvester, J. J. (1853g). Théorème sur les limites des racines réelles des équations algébriques. *Nouvelles annales de mathématiques*, 12:286–287.

[Sylvester, 1904] Sylvester, J. J. (1904). *The collected mathematical papers of James Joseph Sylvester*, volume 1. Cambridge University Press.

[Sylvester, 1908] Sylvester, J. J. (1908). *The collected mathematical papers of James Joseph Sylvester*, volume 2. Cambridge University Press.

[Sylvester, 1909] Sylvester, J. J. (1909). *The collected mathematical papers of James Joseph Sylvester*, volume 3. Cambridge University Press.

[Sylvester, 1912] Sylvester, J. J. (1912). *The collected mathematical papers of James Joseph Sylvester*, volume 4. Cambridge University Press.

[Tannery et Hadamard, 1910] Tannery, J. et Hadamard, J. (1910). *Introduction à la théorie des fonctions d'une variable*, volume 2. A. Hermann.

[Terquem, 1844a] Terquem, O. (1844a). Théorèmes de Descartes, de Rolle, de Budan et Fourier, de MM. Sturm et Cauchy, déduits d'un seul principe. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 3(1):188–194.

[Terquem, 1844b] Terquem, O. (1844b). Théorèmes de Descartes, de Rolle, de Budan et Fourier, de MM. Sturm et Cauchy, déduits d'un seul principe. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 3(1):209–213.

[Terquem, 1844c] Terquem, O. (1844c). Théorèmes de Descartes, de Rolle, de Budan et Fourier, de MM. Sturm et Cauchy, déduits d'un seul principe. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 3(1):555–565.

[Terquem, 1844d] Terquem, O. (1844d). Théorèmes de Descartes, de Rolle, de Budan et Fourier, de MM. Sturm et Cauchy, déduits d'un seul principe. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 3(1):577–580.

[Transon, 1868a] Transon, A. (1868a). De la séparation des racines. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 7(2):25–36.

[Transon, 1868b] Transon, A. (1868b). De la séparation des racines. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 7(2):57–67.

[Transon, 1868c] Transon, A. (1868c). Démonstration de deux théorèmes d'algèbre. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 7(2):97–110.

[Tricot, 1979] Tricot, J. (1979). *Aristote. Organon IV. Les seconds analytiques : Traduction nouvelle et notes.* Paris : J. Vrin.

[Turner, 1980] Turner, R. (1980). The Prussian universities and the concept of research. *Internationales Archiv für Sozialgeschichte der deutschen Literatur (IASL)*, 5(1):68–93.

[Tzitzéica, 1901] Tzitzéica, G. (1901). Sur le nombre des racines communes a plusieurs equations. *CR Acad. Sci. Paris*, 133:918–920.

[Uspensky, 1948] Uspensky, J. (1948). *Theory of equation.* McGraw-Hill.

[van Dalen, 1999] van Dalen, D. (1999). Mystic, Geometer, and Intuitionist. The Life of LEJ Brouwer. Volume 1. The Dawning Revolution.

[Van Der Waerden, 1975] Van Der Waerden, B. (1975). On the sources of my book Moderne algebra. *Historia Mathematica*, 2(1):31–40.

[Van der Waerden, 1985] Van der Waerden, B. (1985). *A History of Algebra : From Al-Khwarizmi to Emmy Noether.* Springer.

[Vandermonde, 1771] Vandermonde, A. T. (1771). Mémoire sur la résolution des équations. *Histoire de l'académie des sciences*, pages 365–416.

[Verdier, 2010] Verdier, N. (2010). Le Journal de Liouville et la presse de son temps : une entreprise d'édition et de circulation des mathématiques au XIXème siècle (1836_1885). *Bulletin de la Sabix. Société des amis de la Bibliothèque et de l'Histoire de l'École polytechnique*, (45):57–64.

[Vincent, 1836] Vincent (1836). Note sur la résolution des équations numériques. *Journal de Mathématiques Pures et Appliquées*, 1(1):341–372.

[von Staudt, 1847] von Staudt, K. G. C. (1847). *Geometrie der lage.* Bauer und Raspe.

[Walton, 1871] Walton, W. (1871). Démonstration du théorème de Cauchy : toute équation a une racine. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 10(2):509–514.

[Waring, 1762] Waring, E. (1762). *Miscellanea analytica de aequationibus algebraicis.* Cambridge.

[Weber, 1893a] Weber, H. (1893a). Die allgemeinen Grundlagen der Galois'schen Gleichungstheorie. *Mathematische Annalen*, 43(4):521–549.

[Weber, 1893b] Weber, H. (1893b). Leopold Kronecker. *Mathematische Annalen*, 43(1):1–25.

[Weber, 1895] Weber, H. (1895). *Lehrbuch Der Algebra, Band 1.* Braunschweig.

[Weber, 1896] Weber, H. (1896). *Lehrbuch Der Algebra, Band 2*. Braunschweig.

[Weber, 1898] Weber, H. (1898). *Traité d'algèbre supérieure*. Gauthier-Villars.

[Weierstrass, 1881] Weierstrass, K. (1881). *Cours du semestre d'hiver 1881/82, Theorie der elliptischen Functionen*. Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Weierstrass, 1882a] Weierstrass, K. (1882a). *Cours du semestre d'hiver 1881/82, Theorie der Abelschen Funktionen. Theil 1*. Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Weierstrass, 1882b] Weierstrass, K. (1882b). *Cours du semestre d'hiver 1881/82, Theorie der Abelschen Funktionen. Theil 2*. Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Weierstrass, 1882c] Weierstrass, K. (1882c). *Cours du semestre d'hiver 1881/82, Theorie der elliptischen Functionen. Theil 1*. Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Weierstrass, 1882d] Weierstrass, K. (1882d). *Cours du semestre d'hiver 1881/82, Theorie der elliptischen Functionen. Theil 2*. Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Weierstrass, 1883] Weierstrass, K. (1883). *Cours du semestre d'hiver 1880/81, Theorie der analytischen Functionen*. Manucrit inédit, Strasbourg, Bibl. Univ. L. Pasteur.

[Weierstrass *et al.*, 1966] Weierstrass, K., Behnke, H. et Kopfermann, K. (1966). *Festschrift zur Gedächtnisfeier für Karl Weierstrass, 1815-1965 : herausgegeben von Heinrich Behnke und Klaus Kopfermann*, volume 33. Opladen.

[Weyl, 1940] Weyl, H. (1940). Algebraic theory of numbers.

[Weyr, 1890] Weyr, E. (1890). Zur theorie der bilinearen formen. *Monatshefte für Mathematik*, 1(1):163–200.

[Yandell, 2002] Yandell, B. (2002). *The Honors Class : Hilbert's Problems and Their Solvers*. AK Peters Natck, MA.

[Zolotareff, 1880] Zolotareff, M. (1880). Sur la théorie des nombres complexes. *Journal de mathématiques pures et appliquées*, 6:51–84 et 129–166.