



Université de Strasbourg

École Doctorale Mathématiques, Sciences de l'Information et de
l'Ingénieur

THÈSE

présentée pour obtenir le grade de

Docteur en sciences

Mention Informatique

par

Rémi Imbach

**Résolution de contraintes géométriques
en guidant une méthode homotopique
par la géométrie**

Soutenue publiquement le **8/10/2013** devant la commission d'examen composée de :

Pascal Schreck *Directeur de thèse*
Professeur à l'Université de Strasbourg

Dominique Michelucci *Rapporteur*
Professeur à l'Université de Bourgogne

Bernard Mourrain *Rapporteur*
Directeur de recherches à l'INRIA-Sophia Antipolis

Mohamed Tajine *Président du jury*
Professeur à l'Université de Strasbourg

Philippe Serré *Examinateur*
Maître de conférences à l'institut SUPMECA

Pascal Mathis *Examinateur*
Maître de conférences à l'Université de Strasbourg

Résumé

Suivant le domaine où on les sollicite, les solutions d'un système de contraintes géométriques peuvent être :

1. formelles et exactes : elles prennent alors la forme d'un plan de construction, qui est une suite d'instructions primitives (par exemple des constructions à la règle et au compas) produisant toutes les solutions ;
2. numériques et approchées : le système de contraintes est ré-écrit en un système d'équations ; on peut chercher une ou toutes ses solutions, réelles ou complexes.

Ces deux attentes sont distinctes mais pas incompatibles. La première est souvent atteinte en appliquant de manière systématique des règles dérivées de lemmes de géométrie. Cependant beaucoup de problèmes, la plupart quand on se place dans un contexte 3D, résistent à cette simplification du raisonnement humain. Les traduire en systèmes d'équations grâce à la géométrie analytique permet alors de les résoudre grâce à des méthodes numériques générales pour répondre au second objectif. Elles sont efficaces quand elles ne recherchent qu'une solution. Par contre, de par la nature des problèmes traités, trouver toutes les solutions en une fois conduit à une complexité exponentielle.

Les méthodes par continuation, ou homotopie, permettent d'obtenir toutes les solutions d'un système d'équations polynomiales. Très étudiées des points de vue de la géométrie algébrique et de la topologie différentielle depuis au moins 30 ans, elles bénéficient d'une assise théorique permettant de les rendre très fiables. Leur application à des systèmes d'équations issus de problèmes de géométrie reste cependant coûteuse et difficilement sujette aux raisonnements permis par l'origine géométrique du problème, en dehors de ceux réalisés en amont sur le système lui-même, car elles opèrent hors de l'espace des figures géométriques.

Notre travail a pour objet la spécialisation d'une méthode par continuation à des systèmes de contraintes géométriques. Leur origine permet de simplifier et de justifier sa mise en œuvre dans l'espace des figures, où des raisonnements géométriques sont possibles. On aborde également les cas où l'ensemble de solutions d'un problème contient des éléments isolés et des continus. Des solutions proches d'une esquisse fournie par un utilisateur sont d'abord trouvées. La recherche d'autres solutions, malgré sa complexité exponentielle, est rendue envisageable par une approche itérative. Une méthode de décomposition, qui simplifie des problèmes traditionnellement considérés comme indécomposables, est proposée pour maîtriser le coût de la résolution.

Abstract

Depending on the required application field, the solutions of a geometric constraints system are either :

1. symbolic and exact such as construction plans, sequences of simple instructions (for instance ruler and compass constructions) providing all the solutions ;
2. or numerical and approximated : the system of constraints is translated into a system of equations ; one or all the solutions, real or complex, can be sought.

These two expectations are separated but not incompatibles. The first one can be satisfied by systematically applying geometric rules. Unfortunately many problems, mostly in a 3D context, resist to this approach, seen as an emulation of human reasoning. Solving the equations built from the constraints by generic numerical methods makes possible to achieve the second goal, but is efficient when only one solution is wanted. However, searching all the solutions leads to an exponential computation cost, due to the nature of problems.

Continuation methods, also called homotopic methods, find all the solutions of a polynomial system. Researches in algebraic geometry and differential topology that have been carried out for at least 30 years make these methods robust and reliable. Using them to solve systems of equations associated to systems of constraints is nevertheless costly. Moreover, combining them with geometric reasoning is a challenge, except in order to simplify the system of polynomial to be solved, because they act in a projective complex space and not in the realizations space.

The aim of this work is to specialize a continuation method to geometric constraints satisfaction problems. Geometry is exploited to simplify and justify its adaptation in the space of realizations, so allowing geometric reasoning. Cases where the connected components of the solution space of a problem have heterogeneous dimensions are addressed. The method discussed here provides in a first step solutions that are similar to a sketch drawn by the user in a computer aided design context. Then a procedure is proposed to search new solutions. Its iterative nature seems to make the exponential complexity of this task bearable. A decomposition method is proposed, that restrains the resolution cost and achieves to simplify problems that usually are not decomposed by existing approaches.

Table des matières

| | | |
|----------|--|------------|
| 1 | Introduction | 9 |
| 1.1 | Résolution de contraintes géométriques | 10 |
| 1.2 | Problématique et organisation | 15 |
| 2 | État de l'art | 19 |
| 2.1 | Systèmes de contraintes géométriques | 20 |
| 2.2 | Méthodes par continuation | 31 |
| 3 | Homotopie | 43 |
| 3.1 | Définitions et notations | 44 |
| 3.2 | Résolution d'un SCG par homotopie | 47 |
| 3.3 | Chemins d'homotopie | 55 |
| 3.4 | Conclusion | 61 |
| 4 | Utiliser un plan de construction | 65 |
| 4.1 | Définitions et notations | 66 |
| 4.2 | Guider la résolution par homotopie | 70 |
| 4.3 | Résolution par reparamétrisation | 78 |
| 4.4 | Conclusion | 85 |
| 5 | Points critiques et incidences | 87 |
| 5.1 | Étude empirique | 88 |
| 5.2 | Éviter les points critiques | 94 |
| 5.3 | Détecter les continuums de solutions avec la géométrie d'incidence | 98 |
| 5.4 | Conclusion | 105 |
| 6 | Décomposition | 107 |
| 6.1 | Graphes de contraintes et décomposition | 108 |
| 6.2 | Une décomposition non constructive | 112 |
| 6.3 | Perspectives : résolution et assemblage | 117 |
| 7 | Conclusion et perspectives | 121 |
| 7.1 | Bilan | 121 |
| 7.2 | Apports | 122 |
| 7.3 | Perspectives | 123 |
| A | Rappels de topologie différentielle | 133 |
| B | Univers géométriques | 137 |

Chapitre 1

Introduction

En informatique graphique, la modélisation par contraintes peut être vue comme une facette de la modélisation déclarative. Elle permet de spécifier un objet en énonçant des propriétés, ou contraintes, qu'il vérifie. Ces propriétés peuvent être mécaniques, topologiques, géométriques... Dans le dernier cas, elles portent par exemple sur sa forme ou sur les positions relatives de primitives qui composent l'objet. Le but d'un système implantant un processus de modélisation déclarative est de produire un, plusieurs, ou tous les objets satisfaisant des contraintes imposées.

Quand les propriétés concernent uniquement sa géométrie, l'objet à construire est une figure et les contraintes portent sur des primitives géométriques : il s'agit par exemple de points, de droites, de surfaces. L'ensemble des contraintes, ainsi que les primitives sur lesquels elles portent, est appelé Système de Contraintes Géométriques (SCG). Les SCG interviennent en Conception Assistée par Ordinateur (CAO), en Enseignement Assisté par Ordinateur (EAO), où ils correspondent à des problèmes de constructions géométriques, en robotique, où un système de pilotage calcule les positions d'un manipulateur, ou encore en chimie moléculaire, pour connaître les positions des atomes d'une molécule.

Ces différents domaines d'application imposent des cahiers des charges différents, et les solutions d'un SCG prennent des formes différentes selon le contexte dans lequel elles sont sollicitées. Un concepteur dans l'industrie dessine une esquisse en spécifiant certaines de ses dimensions dans un logiciel dédié et attend une figure à l'échelle qui ressemble à son esquisse. En EAO, les solutions importent moins que la façon dont elles sont construites ; un plan de construction des solutions, c'est à dire une suite d'instructions pour construire toutes les solutions du SCG, est alors fourni quand il existe. Un chimiste souhaite en général connaître toutes les formes que peut prendre une molécule, alors qu'un roboticien désire aussi une analyse de l'espace des solutions pour savoir si son manipulateur peut atteindre une configuration critique (un continuum de solutions possibles pour la même configuration des mécanismes d'articulation).

Ces différentes attentes ne sont pas incompatibles, et les méthodes de résolution de chacun de ces domaines sont confrontées à des problématiques communes ; on en énonce quelques-unes. Certains SCG résistent à toutes stratégies "diviser pour régner", et peuvent impliquer des dizaines, voire des centaines de contraintes et d'inconnues. Sans le résoudre, décider si un SCG est bien déterminé (s'il possède des solutions et si elles sont en nombre fini) est difficile. Enfin, un SCG bien déterminé admet en général un nombre de solutions exponentiel en le nombre de contraintes.

On introduit dans la section 1.1 quelques notions de résolution de contraintes géométriques pour préciser la problématique de nos travaux et l'organisation de ce mémoire dans la section 1.2.

1.1 Résolution de contraintes géométriques

On présente ici quelques notions fondamentales en résolution de contraintes géométriques en les illustrant sur des exemples. Certaines ne constituent pas le cœur du travail présenté dans ce document et sont évoquées ici informellement. On y fera cependant référence par la suite. D'autres seront précisées quand cela sera nécessaire.

Voici deux exemples simples de problèmes de construction géométrique :

Problème 1 Construire dans l'espace un tétraèdre de sommets P_0, P_1, P_2, P_3 en connaissant la longueur de ses six arêtes.

Problème 2 Construire dans l'espace un solide de sommets P_0, P_1, \dots, P_5 , dont les cinq faces sont (P_0, P_1, P_2, P_3) , (P_0, P_3, P_4, P_5) , (P_1, P_2, P_4, P_5) , (P_0, P_1, P_5) , (P_2, P_3, P_4) , en connaissant la longueur de ses neuf arêtes.

1.1.1 D'un problème de construction à un système de contraintes géométriques

Dans les exemples donnés ci-dessus, on cherche à construire des solides dans l'espace décrits par leurs sommets, leurs faces, et la longueur de leurs arêtes. On peut distinguer les inconnues, les contraintes qu'une solution doit respecter, et les paramètres du problème. Dans le problème 1, les inconnues sont les sommets du tétraèdre (des points) et les contraintes les longueurs des arêtes ; elles dépendent implicitement de paramètres, qui sont les longueurs données dans une unité quelconque. Dans le problème 2, l'énumération des faces sous-entend des contraintes de coplanarité.

On appellera Système de Contraintes Géométriques (SCG) un énoncé énumérant des ensembles d'inconnues (ainsi que leurs types), de paramètres, et de contraintes les impliquant. Parmi les contraintes d'un SCG, certaines dépendent d'un paramètre (ici les contraintes de distance entre deux points), et seront dites *métriques*. Celles n'impliquant aucun paramètre (ici celles de coplanarité) seront dites *booléennes*.

Pour un problème dans l'espace, il est naturel de représenter chaque inconnue par un ensemble de coordonnées cartésiennes et d'assimiler chaque objet (par exemple un point) à un vecteur de coordonnées (pour un point, un vecteur à trois composantes). En confondant un vecteur de coordonnées avec sa représentation graphique dans un espace affine réel, la donnée des coordonnées de chaque inconnue d'un problème est une *figure géométrique*. La figure 1.1 présente une figure géométrique pour chacun des deux problèmes ci-dessus.

Souvent, et particulièrement en CAO, un système de contraintes géométriques est saisi interactivement par un utilisateur dans une interface dédiée par l'intermédiaire d'une esquisse. C'est une figure, au sens donné plus haut, de l'ensemble des inconnues du SCG. Les contraintes métriques sont parfois représentées sur une esquisse sous forme de cotes ; on dira alors que l'esquisse est *cotée*, ou *dimensionnée*. La figure 1.1 donne des esquisses cotées pour les problèmes 1 et 2 où les contraintes de distance sont matérialisées par des arêtes reliant les points, accompagnées des paramètres. En général ces paramètres, représentés sur la figure 1.1 par des noms les désignant, ont des valeurs numériques qui sont fournies par l'utilisateur. On a choisi ici de ne pas représenter les contraintes booléennes sur les esquisses pour garder des figures simples à appréhender, mais de les donner à part. Si l'esquisse d'un SCG satisfait les contraintes booléennes de l'énoncé (ce qui, en 3D, peut relever du défi), elle est une solution du SCG pour certaines valeurs des paramètres des contraintes métriques ; ces valeurs sont celles qui peuvent être mesurées sur l'esquisse.

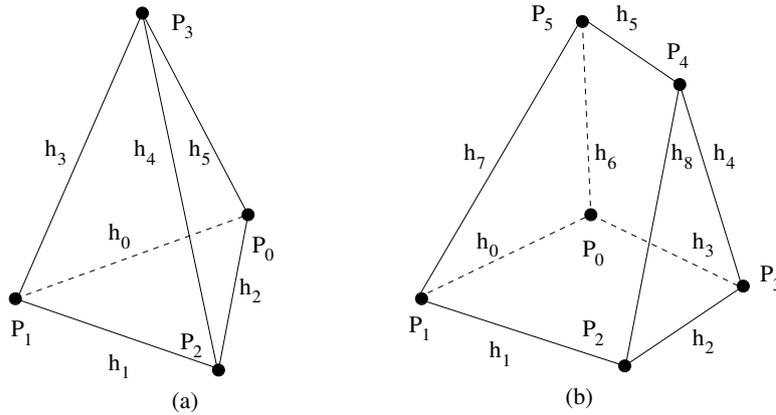


FIGURE 1.1 – Deux esquisses cotées pour les problèmes respectifs 1 et 2. Les arêtes représentent des contraintes de distance de paramètres h_i . À droite, les contraintes de coplanarité sur les faces (P_0, P_1, P_2, P_3) , (P_0, P_3, P_4, P_5) et (P_1, P_2, P_4, P_5) ne sont pas représentées.

Pour écrire formellement un SCG, on peut par exemple utiliser le cadre de la logique du premier ordre. Une *signature* décrit des sortes d’objets et des symboles, fonctionnels et prédicatifs. Chaque inconnue ou paramètre est alors vu comme une variable, d’une sorte spécifiant sa nature géométrique (un point, une droite, une longueur). Ces variables interviennent dans des termes traduisant les contraintes. Pour simplifier, les contraintes métriques sont traduites par des termes de symboles fonctionnels et les contraintes booléennes par des termes de symboles prédicatifs. Une signature permet donc de définir une syntaxe précise dans laquelle un problème est exprimé. Une *interprétation* est l’association d’une sémantique à une signature. Elle spécifie, entre autres, dans quels domaines sont cherchées les inconnues et donne une signification aux contraintes. Un *univers géométrique* consiste en la donnée d’une signature et d’une interprétation, et un SCG s’entend dans un univers géométrique précis.

L’association de coordonnées cartésiennes réelles à chaque inconnue et de réels à chaque inconnue décrite plus haut est appelée dans ce mémoire interprétation *numérique*. Elle associe aussi des fonctions numériques à chaque symbole fonctionnel, et une équation à chaque symbole prédicatif ; elle permet de traduire les contraintes, booléennes et métriques, par des équations.

Une autre interprétation, qui permettra d’étudier la structure combinatoire d’un SCG est celle des degrés de liberté. Habituellement, le *degré de liberté* (abrégé en DoF pour “Degree of Freedom”) d’un objet est le nombre de coordonnées nécessaires pour le déterminer. Par exemple, un point dans l’espace euclidien est décrit par trois coordonnées cartésiennes (ou polaires) et a un degré de liberté de 3. Cette définition est ambiguë lorsqu’on considère entre autres un plan dans l’espace. Son équation cartésienne est déterminée par quatre paramètres, mais son degré de liberté est de 3 : en effet, pour un plan donné, on peut trouver une infinité de vecteurs de paramètres définissant ce plan. Par contre, trois coordonnées (une distance à l’origine et deux angles) polaires suffisent à le déterminer précisément. Le *degré de restriction* d’une contrainte (abrégé en DoR pour “Degree of Restriction”) peut être défini par le nombre d’équations nécessaires pour la décrire ou par le nombre de DoF qu’un objet perd quand il la satisfait. Une contrainte de distance peut s’exprimer par la norme euclidienne d’un vecteur de l’espace et a un DoR de 1. Une contrainte de coplanarité peut s’exprimer comme un déterminant nul et a aussi un DoR de 1. Par contre, une incidence point-droite peut être traduite par l’incidence du point à d’intersection de deux plans, et a un DoR de 2.

1.1.2 Résolution d'un système de contraintes géométriques

Résoudre un SCG peut signifier produire des figures, c'est-à-dire des valeurs pour les coordonnées de ses inconnues qui satisfont ses contraintes. Ces solutions ne sont valables que pour un jeu donné de valeurs de paramètres et sont appelées solutions numériques. Elles sont obtenues en traduisant les contraintes sous forme d'équations et en utilisant une méthode numérique pour les résoudre. On peut aussi chercher des solutions symboliques de ce système d'équations avec des méthodes de calcul formel, qui sont alors des expressions fonctions des paramètres. On parlera dans ce cas de résolution algébrique. La résolution symbolique consiste à trouver une suite d'instructions permettant de construire les figures solutions, appelée solution formelle. Pour un certain jeu de valeurs des paramètres, elles peuvent être appliquées pour obtenir des figures satisfaisant les contraintes.

Résolution symbolique

Considérons le problème 1 spécifié par l'esquisse de la figure 1.1(a). Pour construire dans l'espace une figure qui satisfasse les contraintes de ce problème, on peut appliquer la suite d'instructions suivante :

- (I1) Choisir un point P_0 , un plan Pl et une droite L de Pl passant par P_0 .
- (I2) Construire P_1 comme une des intersections de la sphère centrée en P_0 de rayon h_0 et de L .
- (I3) Construire P_2 comme une des intersections de la sphère centrée en P_0 de rayon h_2 , de celle centrée en P_1 de rayon h_1 et de Pl .
- (I4) Construire P_3 sur une des intersections des sphères centrées respectivement en P_0, P_1, P_2 et de rayons respectifs h_5, h_3, h_4 .

On appelle une telle liste d'instructions un *Plan de Construction* (PC) des solutions du problème 1, car il permet d'en obtenir toutes les solutions. Ce plan de construction est valide quelles que soient les valeurs des paramètres h_i ; quiconque souhaite, pour une valeur donnée des paramètres, construire une solution de ce problème peut exécuter ces instructions en remplaçant les h_i par des valeurs numériques, et en choisissant à chaque étape une intersection où placer le point. Bien sûr, pour certaines valeurs de paramètres et ou certains choix, il peut être impossible d'appliquer ces étapes, par exemple quand trois sphères n'ont aucune intersection. Ce dernier cas se produit quand une inégalité triangulaire n'est pas respectée ; notons qu'il est possible d'appliquer cette étape si l'on cherche des solutions complexes (*c.à.d* dont les coordonnées des points sont complexes). L'étape (I1) de cette liste consiste à choisir un repère affine de l'espace euclidien. On appellera P_0, Pl, L un *repère*.

Ce plan de construction a été obtenu en remarquant que chaque objet se trouve à l'intersection de lieux géométriques définis par les contraintes qui l'impliquent ; cette méthode est appelée LIM, pour "Locus Intersection Method". On la décrit en quelques mots : quand un objet est impliqué par autant de contraintes (chacune de degré de restriction 1) que son degré de liberté, il est bien défini comme intersection de lieux géométriques. Il est considéré comme déterminé, et les contraintes qui ont servi à le placer sont mises de côté, pour appliquer à nouveau le raisonnement aux objets et contraintes restantes. On appelle ce processus une propagation des degrés de liberté. Quand cette propagation ne peut plus être poursuivie, alors si l'ensemble des objets non déterminés peut l'être en choisissant un repère (comme P_0, P_1, P_2 dans l'exemple ci-dessus), un plan de construction des solutions peut être déduit.

Un tel plan de construction permet aussi de remarquer qu'un SCG admet, en général, un nombre exponentiel de solutions en le nombre de ses inconnues. En effet, si chaque instruction du PC nécessite de choisir parmi deux possibilités pour placer un objet (par exemple quand il

est déterminé par l'intersection de trois sphères), il peut construire de l'ordre de 2^m solutions différentes, où m est le nombre d'inconnues, donc d'objets à placer.

La méthode des lieux ne parvient à résoudre que peu de problèmes en 2D, très peu en 3D. Par exemple, dans le SCG correspondant au problème 2, chaque point P_i est impliqué par trois contraintes de distance et deux de coplanarité : aucune inconnue ne peut être construite, et LIM est inefficace.

Résolution algébrique

En utilisant le point de vue de la géométrie analytique, on peut associer à chaque objet un vecteur d'inconnues, et exprimer les contraintes comme des équations en ses inconnues. Ainsi, le problème 1 peut être écrit comme un système de six équations (une par contrainte de distance) en douze inconnues (trois par points). Pour que ce problème admette un nombre fini de solutions, on doit fixer six inconnues, ce qui correspond au choix du repère dans la méthode LIM. Le système d'équations P obtenu peut alors être résolu formellement pour obtenir des solutions symboliques.

Dans le cas du problème 1, on pourrait attribuer comme suit les inconnues x_0, \dots, x_5 aux points P_0, \dots, P_3 : $P_0(0, 0, 0)$, $P_1(x_0, 0, 0)$, $P_2(x_1, x_2, 0)$, $P_3(x_3, x_4, x_5)$, pour poser le système d'équations :

$$(P) \begin{cases} x_0^2 - h_0^2 = 0 \\ x_1^2 + x_2^2 - h_2^2 = 0 \\ (x_1 - x_0)^2 + x_2^2 - h_1^2 = 0 \\ x_3^2 + x_4^2 + x_5^2 - h_5^2 = 0 \\ (x_3 - x_0)^2 + x_4^2 + x_5^2 - h_3^2 = 0 \\ (x_3 - x_1)^2 + (x_4 - x_2)^2 + x_5^2 - h_4^2 = 0 \end{cases}$$

En résolvant P grâce à une méthode symbolique, on trouve comme solutions :

$$[x_0 = h_0] \text{ ou } [x_0 = -h_0],$$

$$\left[x_1 = \frac{x_0^2 + h_2^2 - h_1^2}{2x_0}, x_2 = -\frac{\sqrt{-x_0^4 + 2h_2^2x_0^2 + 2h_1^2x_0^2 - h_2^4 + 2h_1^2h_2^2 - h_1^4}}{2x_0} \right]$$

ou

$$\left[x_1 = \frac{x_0^2 + h_2^2 - h_1^2}{2x_0}, x_2 = \frac{\sqrt{-x_0^4 + 2h_2^2x_0^2 + 2h_1^2x_0^2 - h_2^4 + 2h_1^2h_2^2 - h_1^4}}{2x_0} \right]$$

...

Cette méthode n'est cependant pas applicable à des problèmes impliquant beaucoup d'inconnues et de contraintes, car d'une part les méthodes symboliques utilisées sont très coûteuses et d'autre part le nombre de termes des expressions des solutions explose. Ainsi, dans l'exemple donné, l'expression de x_5 en fonction de $x_0, x_1, x_2, h_3, h_4, h_5$ est une racine d'une fraction rationnelle dont le numérateur est composé d'environ 30 monômes.

Résolution numérique

Lorsqu'on remplace tous les paramètres par leurs valeurs, le système d'équations P peut en revanche être résolu plus efficacement avec des méthodes numériques. Les solutions trouvées sont alors valables uniquement pour les valeurs des paramètres utilisées, et sont approchées. Ces méthodes

sont efficaces quand elles ne cherchent qu'une solution, mais à nouveau très coûteuses quand elles sont utilisées pour les chercher toutes, ne serait-ce qu'à cause du nombre potentiellement exponentiel de solutions.

Quelle que soit la méthode de résolution choisie, appliquer en amont une méthode de décomposition de SCG permet de dominer son coût. Ces méthodes peuvent agir directement sur le SCG ou sur le système d'équations associé.

1.1.3 Constriction et ensembles de solutions

Si une figure est solution d'un SCG toutes les figures obtenues en lui appliquant des déplacements (rotations et translations) sont également solutions du SCG; on parle d'invariance des contraintes par isométrie. On peut ainsi définir une relation d'équivalence sur l'ensemble des solutions d'un SCG pour laquelle deux solutions sont dans la même classe si et seulement si l'une est l'image de l'autre par une isométrie. Quand on parle de l'ensemble des solutions d'un SCG, on désigne en fait l'ensemble des solutions *modulo* cette relation. Le fait de choisir un repère comme montré plus haut permet de chercher les solutions modulo les isométries.

Quand, pour des valeurs *génériques* des paramètres, le nombre de solutions modulo les isométries, *c.à.d* le nombre de classes pour la relation introduite ci-dessus, d'un SCG est fini et non nul (respectivement nul, infini), le problème est dit *bien* (resp. *sur*, *sous*) *contraint*. Le terme générique est ici utilisé comme raccourci pour "en dehors d'un ensemble de mesure de Lebesgue nulle", sous-entendu dans l'espace de toutes les valeurs possibles des paramètres.

Cette définition est ambiguë si l'univers géométrique dans lequel les solutions sont cherchées n'est pas précisé. Si son interprétation est réelle, le problème 1, par exemple, est à la fois bien et sur contraint car les valeurs des paramètres pour lesquelles toutes les inégalités triangulaires sont respectées (respectivement ne sont pas respectées) forment un ensemble de mesure de Lebesgue non-nulle dans l'espace des valeurs des paramètres. En se plaçant dans le cadre où les contraintes d'un SCG peuvent s'écrire sous la forme de fonctions polynomiales, considérer le nombre de solutions projectives complexes permet de caractériser sa constriction de manière cohérente grâce à cette définition¹.

On propose d'illustrer cette notion sur le problème 2, dont une esquisse est donnée sur la figure 1.1(b). Remarquons tout d'abord que chaque paramètre h_i de ce problème prend sa valeur dans les réels positifs ou nuls. L'espace des valeurs possibles des paramètres est donc $(\mathbb{R}_+)^9$. Quand les valeurs de h_4 et de h_6 sont toutes deux nulles, les points P_3 et P_4 , ainsi que P_0 et P_5 , sont confondus. Si de plus la valeur de h_2 (la distance de P_2 à P_3) est différente de celle de h_8 (la distance de P_2 à P_4), le problème n'admet aucune solution. Sinon, si la valeur de h_0 est égale à celle de h_7 , le problème admet une infinité de solutions, car quatre distances ne suffisent pas à déterminer 4 points dans un plan. Cependant, ces situations se produisent sur un hyperplan de dimension 7 (celui où $h_4 = h_6 = 0$), qui est de mesure nulle dans $(\mathbb{R}_+)^9$. On ne peut donc rien en déduire de la constriction générique de ce problème.

Les ensembles de solutions d'un SCG, pour une valeur générique des paramètres, ne sont pas nécessairement homogènes en dimension. On illustre ce phénomène par deux nouveaux exemples dont la figure 1.2 donne des esquisses.

Problème 3 *Construire dans l'espace un solide de sommets P_0, P_1, \dots, P_7 , dont les six faces sont (P_0, P_1, P_2, P_3) , (P_1, P_2, P_4, P_7) , (P_2, P_3, P_5, P_4) , (P_3, P_0, P_6, P_5) , (P_0, P_1, P_7, P_6) , (P_4, P_5, P_6, P_7) , en connaissant la longueur de ses 12 arêtes.*

1. Comme cela sera vu plus bas, d'autres notions, comme celle de la rigidité des solutions, permettent de caractériser la bonne constriction d'un SCG.

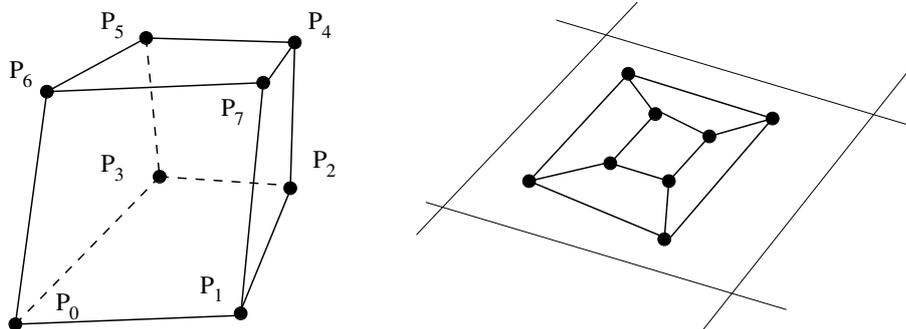


FIGURE 1.2 – Deux esquisses cotées pour les problèmes respectifs 3 et 4. Les arêtes représentent des contraintes de distance. À gauche, les quatre sommets de chacune des six faces sont impliqués dans une contrainte de coplanarité.

Problème 4 Construire dans l'espace huit points P_0, P_1, \dots, P_7 tous coplanaires en connaissant 12 distances entre deux points.

Le problème 4 est sous contraint ; en effet, si on choisit de placer tous les points sur le plan d'élévation nulle en fixant leur dernière coordonnée à 0, la contrainte de coplanarité impliquant tous les points est automatiquement satisfaite. En fixant 3 autres coordonnées (ce qui correspond à fixer un repère dans un plan), on peut associer à ce problème un système de 12 équations en 13 inconnues qui admet, en général, un ensemble de solutions de dimension 1.

Le problème 3, en revanche, admet pour des valeurs génériques des paramètres des solutions isolées, correspondant chacune à une disposition où les plans définis par les contraintes de colinéarité sont tous différents. Cependant, en remarquant que chaque solution du problème 4 satisfait ses contraintes (huit points coplanaires sont aussi coplanaires quatre à quatre) il admet, en plus de ses solutions isolées, des ensembles de solutions de dimension 1, et ce pour des valeurs génériques des paramètres. En particulier, ce problème n'est pas génériquement bien contraint.

On peut cependant exprimer des propriétés satisfaites par les hexaèdres non aplatis qui ne le sont pas par les solutions du problème 4, notamment en voyant un hexaèdre comme un tétraèdre dont deux coins sont tronqués par des plans. Ainsi, comme les droites (P_0P_1) et (P_2P_3) sont coplanaires, elles sont sécantes ou parallèles, et on peut noter H leur intersection, éventuellement un point à l'infini. De même, on note K l'intersection de (P_0P_6) et (P_3P_5) ; on montre alors que les droites (HK) , (P_4P_5) , (P_6P_7) , de même que les droites (HK) , (P_1P_7) , (P_2P_4) sont concourantes, et on en déduit des contraintes de colinéarité que ne vérifient pas, en général, les solutions aplatis.

1.2 Problématique et organisation

On attend d'une méthode de résolution de SCG qu'elle présente une ou plusieurs des qualités suivantes :

- 1 Trouver toutes les solutions d'un problème.
- 2 Trouver des solutions exactes (formelles ou symboliques), réutilisables pour d'autres valeurs de paramètres.
- 3 Trouver des solutions en un temps "raisonnable" (permettant une interaction).

- 4 Trouver des solutions proches d'une esquisse fournie par un utilisateur.
- 5 Trouver des solutions à tous les problèmes.
- 6 Analyser les cas de sous ou sur constriction.

Initialement, l'approche strasbourgeoise de la satisfaction de contraintes géométriques a conduit à étudier des méthodes utilisant des raisonnements géométriques, et répondant aux *desiderata* 1 et 2 notamment grâce à des systèmes à base de connaissances [Schreck 93]. L'exigence 3 est abordée grâce aux stratégies de décomposition de systèmes [Dufourd 98, Schreck 06, Mathis 10]. De telles méthodes produisent des plans de construction et d'assemblage des solutions; leur évaluation, c'est-à-dire le remplacement des paramètres par des valeurs numériques, conduit en général à un arbre d'interprétation produisant un nombre exponentiel de figures. [Essert-Villard 00] propose une méthode pour sélectionner certaines solutions parmi celles-ci pour satisfaire le point 4 et des outils pour naviguer dans l'espace des solutions structuré par l'arbre d'évaluation. Beaucoup de problèmes en 2D et encore plus en 3D (par exemple les problèmes 2 et 3) résistent aux approches géométriques et [Fabre 07] utilise une méthode numérique pour pallier les échecs de ces méthodes dans une approche dite par reparamétrisation qui tente de répondre au point 5. Finalement, les travaux exposés dans [Thierry 07, Thierry 11] prennent en compte la nature sous contrainte d'un problème pour le traiter de manière adéquate (demande 6).

Le fil conducteur de ces travaux est la prise en compte *a priori* de la nature géométrique des problèmes à résoudre, qui fait contraste avec les approches algébriques provenant de la résolution de systèmes d'équations. Ces dernières, quand les contraintes admettent une expression sous la forme d'un polynôme, répondent aux demandes 1, 2, 5 et 6, en appliquant des méthodes de calcul symbolique, mais leur coût explose avec l'accroissement du nombre de contraintes d'un problème. Résoudre le système d'équations grâce à une méthode numérique cherchant toutes les solutions revient à sacrifier les qualités 2 et 6 pour obtenir le bénéfice du point 3, partiellement atteint. En effet, ces méthodes restent lentes de par le nombre exponentiel de solutions cherchées. Souvent, en CAO, une seule solution est cherchée, pour satisfaire au point 3, et au 4 si possible.

Les travaux présentés dans ce mémoire portent sur l'utilisation de la nature géométrique des problèmes considérés pour guider une méthode numérique, l'homotopie. Cette dernière, très étudiée depuis au moins 30 ans, bénéficie d'une assise en géométrie algébrique et en topologie différentielle qui la rend très fiable et opère de la manière suivante. En définissant un système d'équations dit *initial* dont les solutions sont connues et en l'interpolant continûment en le système à résoudre, dit système *cible*, ses solutions sont déformées en les solutions isolées de ce dernier. Pendant sa déformation, une solution décrit une courbe, appelée *chemin d'homotopie* qui peut être suivie par une méthode numérique. Toutes les solutions du système cible sont trouvées en suivant ces chemins à la condition suffisante que le système initial en possède au moins autant. Un système initial adéquat, dont toutes les solutions sont connues, est alors construit. Le nombre de solutions du système cible, a priori inconnu avant sa résolution, est majoré aussi finement que possible, car un excès de solutions du système initial ralentit inutilement la résolution.

Dans la littérature consacrée à la résolution de contraintes géométriques, on trouve une alternative à la construction d'un système initial consistant en l'utilisation de l'esquisse : en remarquant qu'elle vérifie les contraintes métriques de l'énoncé pour certaines valeurs de leur paramètres, et en supposant qu'elle répond aux contraintes booléennes, elle permet de définir une homotopie qui interpole ces valeurs en celles pour lesquelles les solutions sont souhaitées. Les points du chemin suivi pendant cette interpolation correspondent alors tous à des figures. Cependant, comme une seule solution du système initial (l'esquisse) est connue, une seule solution est trouvée. Elle a néanmoins l'avantage d'être similaire à l'esquisse dans un sens défini dans [Essert-Villard 00].

En choisissant de réaliser l'homotopie dans l'espace des figures, les propriétés assurant le succès de la phase de suivi de chemin et donc de l'obtention de solutions sont perdues. Les deux

phénomènes suivants peuvent alors se produire et mettre en échec la résolution, et on propose dans nos travaux de les aborder en analysant leur origine dans l'espace des figures :

- (i) un chemin a une longueur infinie sans aboutir à une solution,
- (ii) un chemin en croise d'autres, ou change de topologie.

Le chapitre 3 montre comment embarquer l'interpolation des paramètres dans une fonction d'homotopie dont les chemins portent plusieurs solutions du problème considéré. Une méthode générale de résolution de SCG en est déduite. Une caractérisation rigoureuse des chemins d'homotopie, utilisant le cadre théorique de la géométrie différentielle, est proposée.

Le chapitre 4 porte sur l'utilisation de Plans de Constructions (PC) pour guider cette méthode. On considère des problèmes résistant aux méthodes géométriques, et pour lesquels on ne sait pas obtenir de PC fournissant toutes les solutions. Il est cependant possible, en remplaçant quelques contraintes du SCG, d'en obtenir un; les figures qu'il construit ne respectent pas toutes les contraintes. On l'utilise d'abord comme observateur pour détecter un chemin de longueur infinie (problème (i)), puis pour générer de nouvelles esquisses dans le but d'obtenir de nouvelles solutions. Quand les contraintes qui ont été ajoutées pour produire le plan de construction sont métriques, on peut voir les figures qui en résultent comme dépendant de leur paramètres, et chercher pour quelles valeurs ces figures satisfont les contraintes ôtées du problème initial. Cette méthode est appelée reparamétrisation dans la littérature. On expose brièvement une adaptation de notre méthode par homotopie à la recherche de ces valeurs.

Le problème (ii) apparaît quand le chemin suivi contient un *point critique*, en lequel la fonction d'homotopie n'est pas suffisamment régulière. Les méthodes d'homotopie classiques règlent ce problème par un argument probabiliste, qui garantit qu'un tel point n'est pas sur un chemin suivi, même quand il est solution du système d'équation à résoudre. Dans ce dernier cas, il correspond à une solution non-isolée, c'est à dire à un élément d'un continuum de solutions du système. Le chapitre 5 traite ces problèmes : les points critiques sont dus à la satisfaction de théorèmes de géométrie amenant un problème à être sous-contraint. On montre comment modifier la fonction d'homotopie afin d'appliquer l'argument probabiliste utilisé par les méthodes classiques pour éviter les points critiques quand ils ne sont pas solutions. On y propose également l'esquisse d'une méthode permettant, le cas échéant, d'analyser la figure correspondant au point critique et de détecter la cause de la sous-constriction quand elle est due à la géométrie d'incidences.

La recherche de toutes les solutions d'un SCG impliquant un grand nombre de contraintes est trop coûteuse si une stratégie de décomposition ne lui est pas appliquée en amont de la résolution. Une telle stratégie consiste à chercher dans un SCG des sous-systèmes bien déterminés, et cette recherche se base souvent sur des critères combinatoires. Le chapitre 6 présente une méthode de décomposition, utilisant la même méthode de remplacement de contraintes que celle utilisée dans la reparamétrisation.

Le chapitre 2 est dédié à la présentation de la littérature portant sur les systèmes de contraintes géométriques et les méthodes par homotopie.

Terminons cette introduction en précisant le sens des mots "homotopie" et "continuation". L'homotopie désigne la déformation continue d'un système initial en un système cible. La continuation désigne la déformation des solutions du système initial en les solutions du système cible. Dans ce mémoire, on utilisera les termes "méthode par homotopie", ou "méthode homotopique" et "méthode par continuation" de façons équivalentes.

Chapitre 2

État de l'art

Dans les débuts de la Conception Assistée par Ordinateur (CAO), les contraintes sont des équations sur les coordonnées des objets géométriques considérés. [Sutherland 64] est à notre connaissance la première communication scientifique au sujet d'un système de résolution de contraintes ; les auteurs y décrivent un système muni d'une interface dédiée à la saisie et à la visualisation, plus tard appelée "sketcher". Des contraintes sur les objets géométriques peuvent être spécifiées par l'utilisateur pour l'aider dans sa démarche de conception. Les équations correspondant aux contraintes sont résolues par une méthode de relaxation. [Lin 81] décrit un système de dessin technique spécialisé dans la résolution de contraintes géométriques, où une fonction numérique multivariée en les coordonnées des objets représente le système de contraintes. L'esquisse est utilisée comme solution initiale, et améliorée par des itérations de Newton-Raphson. Deux problèmes majeurs liés à cette démarche sont soulevés. D'abord, appliquer la méthode de Newton-Raphson requiert un nombre égal d'équations et d'inconnues ; son succès nécessite également que les équations soient bien réparties, et non redondantes. Ensuite, la solution obtenue pour un jeu de paramètres n'est pas réutilisable pour d'autres valeurs de ces paramètres.

Ces deux problèmes révèlent deux aspects de la résolution de contraintes géométriques. Le premier est l'étude des Systèmes de Contraintes Géométriques (SCG) et particulièrement de leur constriction. Elle peut être caractérisée par la nature de son espace de solutions, qui est difficile à déterminer *a priori*. Une approche combinatoire, justifiée par la théorie de la rigidité ou le théorème de König-Hall, en fournit une approximation, qui est la constriction structurelle d'un SCG. Le second aspect est la résolution d'un SCG par des méthodes répondant à certaines demandes.

Des solutions formelles sont souvent trouvées par des méthodes dites *géométriques* car elles utilisent un raisonnement géométrique pour résoudre un problème. Elles fournissent toutes les solutions d'un SCG. Beaucoup de SCG résistent cependant à ces méthodes, auquel cas ils peuvent être résolus numériquement. Utilisée en amont d'un processus de résolution, une stratégie de décomposition permet de diminuer son coût et d'offrir à l'utilisateur une vision plus intelligible des solutions fournies, ou une analyse structurelle de la constriction d'un SCG le cas échéant. La section 2.1 propose un aperçu de la résolution de contraintes géométriques. Le niveau de détails avec lequel sont décrits les différents aspects de ce domaine est fonction de leur proximité avec les travaux présentés dans ce mémoire.

Parmi les méthodes numériques, l'homotopie (ou continuation) permet de trouver toutes les solutions isolées d'un système d'équations. Le but des travaux présentés ici est d'exhiber les problèmes liés à son utilisation pour résoudre des SCG et d'apporter à certains quelques réponses, c'est pourquoi la section 2.2 en présente le principe et la littérature qui s'y rapporte.

2.1 Systèmes de contraintes géométriques

Issues du domaine de l'Enseignement Assisté par Ordinateur (EAO) et de la résolution automatique de problèmes de géométrie, les méthodes géométriques sont souvent implantées par des systèmes à bases de règles [Scandura 74, Buthion 79]. Elles nécessitent une formalisation et une axiomatisation précise d'un univers géométrique, mais fournissent des solutions formelles, valables pour des valeurs de paramètres quelconques. L'énoncé est traduit sous forme de formules logiques, et un moteur d'inférence applique des lemmes de géométrie pour les simplifier. Le résultat prend la forme d'un programme de construction qui est une généralisation des plans de construction incluant des structures de contrôle du type de celle des langages de programmation [Schreck 93], souvent des constructions à la règle et au compas. [Verroust 92, Sunde 87] propose un ensemble de règles de constructions pour la CAO, et caractérise l'ensemble des problèmes que son système est capable de résoudre. [Brüderlin 87, Aldefeld 88, Sunde 87] détectent et rendent compte à l'utilisateur des parties sous-déterminées d'un système de contraintes. [Kramer 90] utilise le cadre formel de l'analyse cinématique, où les systèmes de contraintes sont des pièces de mécanismes (qui sont les objets géométriques dont la position est inconnue) reliés par des joints qui sont les contraintes. À chaque objet (respectivement contrainte) est associé un degré de liberté (resp. degré de restriction). Les contraintes sont résolues itérativement par des intersections de lieux géométriques dictées par une base de règles.

Un graphe de contraintes est une représentation combinatoire d'un système de contraintes géométriques obtenue en associant un sommet à chaque inconnue et une arête à chaque contrainte reliant les sommets associés aux objets sur lesquels elle porte. On suppose ici que les contraintes d'un SCG agissent toutes sur exactement deux objets géométriques, pour éviter le formalisme des hyper-graphes. Un graphe de contraintes facilite l'étude de la structure combinatoire du SCG associé. Cette structure donne lieu à des approches par décomposition de systèmes (par exemple [Owen 91, Bouma 95]) qui seront décrites plus bas.

Le domaine de la preuve automatique de théorème de géométrie élémentaire (donc sans considération différentielle) a développé des méthodes symboliques coûteuses mais complètes dans certaines géométries. Par exemple, le principe de la méthode de Ritt-Wu [Wen-Tsun 86] dont est dérivé le prouveur de Chou [Chou 88] est le suivant : les hypothèses d'un théorème forment un système d'équations polynomiales, et sa conclusion un polynôme. Le système est transformé par pseudo-divisions successives, jusqu'à obtenir un reste et des conditions subsidiaires ; si le reste est nul, la conclusion est une conséquence des hypothèses sous réserve de la réalisation des conditions subsidiaires qui peuvent s'interpréter géométriquement. Dans le cas de la résolution de contraintes géométriques, le système d'équations associé aux contraintes peut être résolu de manière partiellement symbolique en le triangularisant à l'aide de la méthode de Wu ou d'une base de Gröbner (voir par exemple [Elkadi 07]). Cette approche, qui doit être complétée par une résolution numérique de chaque polynôme, est coûteuse et mal adaptée aux systèmes provenant de problèmes de CAO.

Le système d'équations associé à un SCG peut être résolu de manière numérique. Ces méthodes seront présentées plus en détail dans la sous-section 2.1.2. Elles sont rapides quand elle ne doivent fournir qu'une solution et agissent en général sur des SCG bien déterminés. La sous-section 2.1.1 aborde cette problématique de constriction d'un SCG. Dans la sous-section 2.1.3, on présentera les méthodes dites "hybrides", parce qu'elles mêlent des concepts issus de plusieurs domaines. La sous-section 2.1.4 présente des méthodes de décomposition. Celles-ci implémentent une stratégie "diviser pour régner" pour résoudre des SCG : ils sont réduits en sous-systèmes qui peuvent être résolus séparément en respectant un ordonnancement. Les solutions de ces sous-systèmes sont ensuite ré-assemblées.

2.1.1 Ensembles de solutions

La définition de la constriction d'un SCG introduite dans le chapitre 1 fait intervenir le cardinal de l'ensemble de ses solutions, pour des valeurs *génériques* des paramètres. Deux difficultés majeures proviennent de sa nature probabiliste (l'ensemble des valeurs génériques est l'ensemble des valeurs privé d'un ensemble de mesure de Lebesgue nulle) et non constructive (elle permet de qualifier un problème de contraintes géométriques de sous, bien, sur contraint si on connaît l'ensemble de ses solutions). Cependant, résoudre un SCG nécessite de connaître sa constriction, ou au moins de supposer qu'il soit bien contraint. Notons que les contraintes considérées ici sont invariantes par changement de repère et ont une expression polynomiale homogène.

Cette notion de constriction est le pendant dans le domaine de la résolution de contraintes géométriques de l'étude des variétés algébriques définies par un système d'équations polynomiales en géométrie algébrique, ou de celle des variétés différentielles qui sont les zéros d'une fonction analytique en topologie différentielle, que nous présentons brièvement ici.

Considérons une fonction $C^\infty F : \mathbb{C}^m \rightarrow \mathbb{C}^n$, où $m \geq n$. On dit que 0 est une valeur régulière de F si en tous ses antécédents, la Jacobienne de F est de rang maximal (c'est à dire n) ; si 0 est une valeur régulière de F , les composantes connexes de $F^{-1}(0)$ sont des variétés différentielles de dimension $m-n$; une introduction à la topologie différentielle peut être trouvée dans [Milnor 65], ou [Gramain 71] en français¹. Si F est une fonction numérique associée à un SCG, qu'elle comporte autant de composantes que de variables ($m = n$), et qu'elle admet 0 comme valeur régulière, les solutions de $F = 0$ sont des points isolés. On dira alors que ces solutions sont *rigides*, ce qu'on explique intuitivement par : soumises à n'importe quel déformation continue (autre que des déplacements), une solution du SCG n'en est plus une ; ou encore : la propriété "être une solution" n'est invariante que par déplacement. La question de la rigidité des solutions d'un système est traduite par le calcul du rang de la jacobienne de la fonction numérique F , plus précisément le calcul des points en lesquels la jacobienne de F n'est pas de rang plein. Ce calcul n'est pas réalisable symboliquement pour des SCG issus de la CAO. De plus, si F est considérée analytique, cette notion de rigidité ne dit rien du caractère fini ou non du nombre de solutions.

Supposons à présent que les composantes de F soient polynomiales, et que $m \geq n$. L'ensemble des solutions du système $F = 0$ dans \mathbb{C}^m est appelé ensemble algébrique. Ses composantes connexes sont aussi des ensembles algébriques. En considérant une de ces composantes connexes \mathcal{Z} , on définit l'ensemble $\mathcal{Z}_{reg} \subseteq \mathcal{Z}$ formé de ses points réguliers (pour simplifier, l'ensemble des points de \mathcal{Z} en lesquels la jacobienne de F est de rang n), et on appelle variétés algébriques les composantes connexes de \mathcal{Z}_{reg} . On définit leur dimension en les voyant comme des variétés différentielles. Posons à présent $m = n$, et supposons que F soit la fonction numérique dont les composantes correspondent aux contraintes d'un SCG. Ce SCG est bien contraint si les variétés algébriques définies par $F = 0$ sont toutes de dimension 0, et sont en nombre fini. Le théorème de Bézout ([van der Waerden 49]) établit que le nombre de solutions isolées de F dans \mathbb{C}^m est au plus égal au produit des degrés des composantes de F ; les solutions rigides du SCG sont donc en nombre fini. Cependant, rien n'est dit des variétés de dimension strictement positives. Le calcul symbolique permet de trouver ces ensembles, en cherchant une décomposition de l'idéal engendré par les polynômes de F . [Sommese 05] propose une introduction à ces notions (voir [Elkadi 07] pour un ouvrage plus complet en français).

La constriction structurelle d'un système d'équation, plus faible mais ayant l'avantage d'être combinatoire et plus facilement calculable, et son pendant dans le domaine des systèmes de contraintes, remplace souvent celle de la constriction générique.

1. Un rappel des notions de bases, tirées de ces ouvrages, est donné en annexe A.

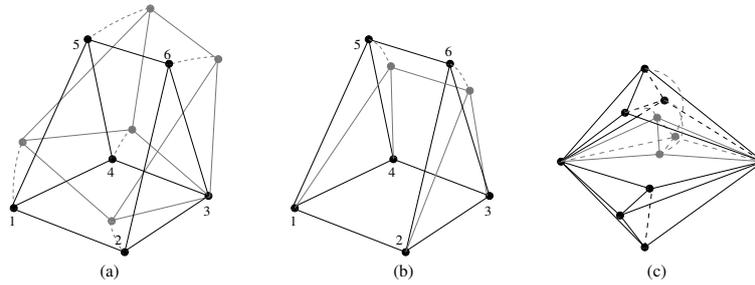


FIGURE 2.1 – Un mouvement rigide (a) et un non rigide (b) d’un assemblage de tiges métalliques en 2D. À droite (c), un mouvement non rigide de l’exemple en 3D de la double banane. Les points sont les joints, les arêtes les tiges.

Constriction structurelle et théorie de la rigidité.

La notion de constriction structurelle d’un système F de c équations en m inconnues répond à la question suivante : F possède-t-il assez d’équations, et les inconnues sont-elles assez bien réparties dans ces équations pour être déterminées par F . La question est combinatoire, et la réponse apportée l’est aussi ; elle utilise la notion de graphe biparti associé à un système d’équations. Ce graphe est construit en associant à chaque inconnue et chaque équation un sommet. Les arêtes de ce graphe relient un sommet associé à une équation à un sommet associé à une inconnue si et seulement si l’inconnue intervient dans l’équation (et inversement). Le théorème de König-Hall établit que F est structurellement bien contraint si et seulement si le graphe biparti qui lui est associé admet un couplage parfait (voir [Plummer 86]). On dira par la suite d’un SCG qu’il est *bien contraint au sens de König-Hall* si le système d’équations qui lui est associé est structurellement bien contraint. Cette notion ne tient pas compte de la “qualité” des équations de F ; par exemple, deux équations peuvent être redondantes, et le système est génériquement sous-contraint, ou contradictoires, entraînant une sur-constriction. D’autre part, les contraintes d’un SCG peuvent être les hypothèses et la conclusion d’un théorème de géométrie qui le rend génériquement sous-contraint.

Un algorithme dû à Dulmage et Mendelsohn ([Dulmage 63]) agissant sur ce graphe biparti permet de décomposer un système d’équations en une partie (structurellement) bien contrainte, une sous-contrainte, et une sur-contrainte.

La *théorie de la rigidité* apporte une réponse au problème de la constriction d’un SCG quand celui-ci implique uniquement des points et des distances. Elle étudie la *rigidité* (cette notion a bien sûr un rapport avec celle déjà introduite) d’assemblages de tiges métalliques reliées par des joints flexibles dans le plan ou dans l’espace. On appelle *mouvement* d’un tel assemblage un mouvement des joints tel que la distance entre chaque paire de joints reliés par une tige soit préservée (la tige ne plie pas, et n’a pas d’élasticité). Un mouvement est dit *rigide* si la distance entre chaque paire de joints, non nécessairement reliés par une tige, est préservée par ce mouvement ; ces mouvements sont des déplacements du plan ou de l’espace. Enfin, un assemblage est dit *rigide* si tous ses mouvements sont rigides. La figure 2.1(a) présente un assemblage de tiges métalliques (correspondant à une figure du problème 2) et un de ses mouvements rigides. Cet assemblage admet un mouvement non rigide (voir figure 2.1(b)), et n’est donc pas rigide.

Jusque là il a été fait abstraction des longueurs des tiges de l’assemblage. Or un assemblage peut être non rigide pour certaines longueurs de ses tiges mais rigide pour d’autres valeurs. Ainsi, le même assemblage que celui de la figure 2.1(a) avec des longueurs de tiges pour lesquelles les

droites (12), (34) et (56) ne sont pas parallèles est rigide. On dira qu'un jeu de longueurs pour les tiges d'un assemblage est *générique* quand les longueurs sont algébriquement indépendantes. En admettant que l'ensemble des jeux de longueurs non algébriquement indépendantes forme en ensemble de mesure de Lebesgue nulle dans l'ensemble de tous les jeux de longueurs possibles pour les tiges d'un assemblage, on a alors (voir [Hendrickson 92]) :

- (i) Si pour un jeu de longueurs générique un assemblage est rigide, il l'est pour tous les jeux de longueurs génériques.
- (ii) Si pour un jeu de longueurs générique un assemblage est rigide, il l'est pour presque tous les jeux de longueurs.
- (iii) Si pour un jeu de longueurs un assemblage est rigide, il l'est pour presque tous les jeux de longueurs.

La proposition (i) est parfois appelée le théorème de Gluck, et (iii) l'hypothèse de généricité. En pratique, les valeurs des longueurs sont représentés par des flottants, donc des rationnels, et l'ensemble des jeux de paramètres représentable est de mesure de Lebesgue nulle dans l'espace des jeux de longueurs.

Dans le cas des assemblages dans le plan (en 2D), une caractérisation combinatoire de ceux qui sont *génériquement rigides* (c.à.d rigides pour tous les jeux de longueurs génériques) existe, et est appelée critère de Laman ([Laman 70]). En appelant *sous-assemblage* d'un assemblage un assemblage formé d'un sous-ensemble de ses joints, et dont les tiges sont celles de l'assemblage original ayant leurs deux extrémités dans ce sous-ensemble, un assemblage vérifie la caractérisation de Laman si, pour c tiges et n joints :

$$(C1) \quad c = 2n - 3,$$

$$(C2) \quad \text{dans tout sous-assemblage de } c' \leq c \text{ tiges et } 2 \leq n' \leq n \text{ joints, } c' \leq 2n' - 3.$$

Si un assemblage du plan vérifie la caractérisation de Laman, il est rigide pour presque tous les jeux de longueurs des tiges. La condition (C2) est a priori coûteuse à vérifier, du fait de la nécessité de prendre en compte tous les sous-assemblages; [Lovasz 82], entre autres, décrit un algorithme polynomial pour tester ce critère efficacement.

Si on considère à présent un assemblage, pour un jeu de longueurs de tiges donné, comme une solution d'un SCG dont les inconnues sont des points (les joints de l'assemblage) et les contraintes des distances entre chaque paire de points reliés dans l'assemblage par une tige, le critère de Laman permet de caractériser ce SCG en 2D génériquement bien contraint. Si il vérifie ce critère alors pour presque toutes les valeurs des paramètres des distances, toutes ses solutions sont rigides. En nommant F la fonction numérique qui lui est associée, cela signifie que les points en lesquels F s'annule sont isolés, et les deux notions de rigidité, celle exposée précédemment et celle des assemblages se rejoignent. De plus, on peut facilement écrire F de façon à ce que ses composantes soient des polynômes, et les variétés algébriques du système $F = 0$ sont toutes de dimension 0. Comme le nombre de solutions isolées de ce système est fini (plus petit que la borne de Bézout), le SCG correspondant est génériquement bien contraint.

Cependant, l'extension du critère de Laman en 3D ($c = 3n - 6$ et $c' \leq 3n - 6$) ou en dimension supérieure n'implique pas la rigidité des solutions d'un SCG comme le montre l'exemple classique de la double banane, qui admet pour presque tous jeux de longueurs des tiges tels qu'il existe un assemblage (les sommets communs aux deux bananes coïncident) le mouvement non rigide représenté par la figure 2.1(c). Son extension à des SCG contenant des contraintes d'incidence, par exemple, n'implique pas non plus la rigidité.

Détecter le niveau de constriction

Aucune caractérisation combinatoire ou structurelle (comme le critère de Laman) n'est connue pour décider de la rigidité d'assemblages en 3D, ou de la constriction générique de SCG, (impliquant d'autres contraintes que des contraintes métriques). Pourtant détecter si un système est bien contraint, ou au moins si ses solutions sont rigides pourrait permettre d'analyser l'échec de la résolution d'un problème dans le cadre d'un système de CAO.

Un premier résultat, appelé théorème de Cauchy ([Cauchy 13]), établit que tout assemblage correspondant à un polyèdre strictement convexe de \mathbb{R}^3 est rigide si ses faces sont rigides. Utilisant l'hypothèse de généricité (la proposition (iii) ci-dessus), la méthode probabiliste numérique propose de choisir au "hasard" un jeu de longueurs pour les tiges d'un assemblage et de tester si il est rigide en calculant tous ses mouvements (ce qui peut se faire en calculant le noyau d'une application linéaire associée à une matrice jacobienne).

L'extension de cette méthode au cadre des SCG n'impliquant pas seulement des contraintes de distances est dite méthode du témoin ([Michelucci 06a, Michelucci 07]). Un témoin est une figure (par exemple une esquisse) satisfaisant les contraintes booléennes d'un SCG (les contraintes métriques sont satisfaites pour certaines valeurs de paramètres). À nouveau, l'idée est que si le témoin est choisi au "hasard" la jacobienne de la fonction numérique associée en ce témoin partage ses propriétés, en particulier son rang, avec les jacobiennes en les solutions. Cette méthode permet de détecter dans un système les dépendances entre contraintes, par exemple quand certaines sont les conclusions d'un théorème de géométrie.

On a évoqué, en introduction de cette sous-section, la géométrie algébrique, dont le sujet d'étude est l'ensemble de solutions d'un système polynomial, et on a mis en évidence qu'un SCG pouvant être écrit sous la forme d'un système d'équations polynomiales $F = 0$ était génériquement bien contraint si les variétés algébriques définies par ce système étaient toutes de dimension 0. Le calcul symbolique de ces variétés est hors d'atteinte dans le cas de SCG issus de la CAO. La dernière décennie a vu l'émergence d'une littérature au sujet du calcul par des méthodes numériques, pour des valeurs de paramètres de F fixées, des variétés algébriques définies par le système $F = 0$; ce sont les méthodes de la *numerical algebraic geometry* (voir [Sommese 05]). Elles s'appuient essentiellement sur l'homotopie, aussi seront-elles détaillées dans la section 2.2. Elles sont évoquées ici car en choisissant des valeurs pour les paramètres de F au "hasard" et en calculant grâce à de telles méthodes les variétés algébriques définies par $F = 0$ et leurs dimensions, on peut conclure sur la bonne constriction ou non d'un SCG. Bien que n'effectuant que du calcul numérique, elles restent cependant très coûteuses pour des systèmes d'équations issus de SCG.

2.1.2 Résolution numérique

Les méthodes numériques permettent, pour la plupart, de résoudre des problèmes structurellement bien contraints. En supposant que toutes les contraintes considérées ont un degré de restriction de 1 et donc s'interprètent par une équation, la fonction numérique associée à un tel SCG dépend de 6 variables de plus que de composantes (en 3D). La constante 6 correspond aux degrés de liberté d'un objet rigide dans l'espace. Pour chercher les solutions modulo les déplacements, 6 variables sont fixées dans un repère affine de l'espace. La fonction numérique F obtenue après le choix d'un tel repère a donc autant de composantes que de variables, et on notera par la suite m le nombre de composantes de F . On peut alors appliquer à cette fonction des méthodes numériques, en général itératives, pour trouver des vecteurs v de \mathbb{K}^m , où $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , tels que $F(v) = 0$. Les solutions attendues sont en général des vecteurs de \mathbb{R}^m (qui correspondent à des solutions réelles).

Les premiers travaux publiés ([Sutherland 64, Lin 81]) décrivant des systèmes de CAO utilisent des méthodes itératives, qui améliorent une solution initiale. Une seule solution est alors

trouvée, et sa “proximité” avec l’esquisse est utilisée comme argument en faveur de ces méthodes. Cependant la notion de proximité de deux figures relève du jugement de l’utilisateur. Quand la méthode numérique ne converge pas, ou converge vers une solution qui ne lui convient pas, ce dernier peut redessiner une esquisse différente pour la soumettre au solveur. Plus récemment, des méthodes recherchant toutes les solutions d’un problème de contraintes géométriques ont été proposées [Durand 00, Elber 01, Tahari 10]. Elles semblent plus adaptées à des domaines comme la chimie moléculaire ([Porta 07]) et sont directement parallélisables ; leur coût est exponentiel. On distinguera les méthodes par bisection (qui découpent l’espace des figures en régions dans lesquelles chercher les solutions) et les méthodes par homotopie, qui déforment les zéros d’une fonction initiale connus en les zéros recherchés. Ce lien entre solutions connues et solutions cherchées est appelé continuation.

Méthodes itératives

Un zéro de la fonction numérique associé à un SCG peut être trouvé par la méthode de Newton-Raphson. Au voisinage d’un vecteur v_0 initial, F est assimilée à sa dérivée d’ordre 1 en v , qui est une application linéaire définie par la matrice jacobienne de F . Si cette matrice est inversible, le noyau de l’application linéaire associée est un vecteur v_1 qui approxime un zéro de F . Cette approximation est utilisée comme nouveau vecteur initial, et l’étape décrite plus haut est répétée pour construire une suite d’approximations $(v_n)_{n \in \mathbb{N}}$, qui converge de manière quadratique vers un zéro de F si v_0 est dans une zone de convergence autour de ce zéro. Dans le cadre de la CAO, [Lin 81, Serrano 91, Pérez 93], entre autres, utilisent l’esquisse fournie par l’utilisateur comme solution initiale. Cette méthode est néanmoins sensible aux optimums locaux de F , et numériquement instable dans le sens où une solution initiale ne converge pas forcément vers la solution dont elle est la plus proche. Pour décrire ce phénomène de manière plus rigoureuse, on utilise la notion de bassins d’attraction. Pour un zéro v_* de F , on appelle bassin d’attraction de v_* le sous-ensemble de \mathbb{R}^m constitué des points v pour lesquels la méthode de Newton-Raphson avec comme vecteur initial v construit une suite d’approximations qui convergent vers v_* . Les frontières des bassins d’attraction des solutions d’une fonction F sont très irrégulières ; l’ensemble de Julia associé à la méthode de Newton a une structure fractale et est parfois appelé fractale de Newton. Les conséquences dans le cadre de l’application de cette méthode à la CAO sont qu’une solution trouvée peut parfois être très éloignée de l’esquisse fournie par l’utilisateur ².

[Ge 99] suggère d’utiliser une méthode d’optimisation, appelée BFGS, où méthode de quasi-Newton, pour minimiser la fonction $S(X) = \sum F_i(X)^2$, où les F_i sont les composantes de F . Cette méthode permet donc de résoudre de manière homogène des SCG sous, bien et sur contraints. BFGS est une méthode itérative, et c’est à nouveau l’esquisse qui est utilisée comme solution initiale. Les auteurs introduisent également une mesure de la proximité d’une solution avec l’esquisse ; les solutions trouvées par leur méthode sont “souvent proches de la meilleure solution”, au regard de cette mesure. Cette méthode a été peu utilisée.

Méthodes par bisections

Le principe des méthodes par bisection, issues de l’arithmétique par intervalles (voir [Moore 66, Neumaier 90]), est de diviser l’espace \mathbb{K}^m en boîtes (extensions en dimension m d’un intervalle), et de tester pour chaque boîte :

1. si un unique zéro de F s’y trouve, dans ce cas la boîte est considérée comme une solution (précisée par une méthode itérative),
2. si aucun zéro de F ne s’y trouve, et la boîte est disqualifiée,

2. voir [Lamure 95] pour plus de précisions.

3. si aucun ou plusieurs zéros de F s’y trouvent, la boîte est subdivisée, et chaque boîte obtenue est à nouveau soumise à cette procédure de test.

Dans [Ait-Aoudia 02], la procédure de décision décrite ci-dessus est réalisée par le test de Krawczyk ([Krawczyk 69]) qui utilise une expression de F et de sa jacobienne par intervalle ([Moore 66]). La subdivision de la boîte, dans le cas 3, se fait selon la longueur de son plus long côté.

Les solveurs à base de Bernstein ([Mourrain 09, Tahari 10]) utilisent une représentation des composantes de F , quand elles sont polynomiales, dans la base tensorielle de Bernstein. La propriété fondamentale de cette base est que les valeurs de F sont contenues dans l’enveloppe convexe de ses coefficients dans cette base. Deux autres propriétés sont que les coefficients dépendent de l’espace sur lequel on considère F , et que la représentation d’un polynôme multivarié dans cette base nécessite un nombre de termes exponentiel en le nombre de ses variables. Une conversion initiale de F dans la base de Bernstein pour la boîte où sont initialement recherchées les solutions est d’abord effectuée. Pour savoir si F s’annule sur cette boîte, les signes des coefficients dans cette base sont inspectés. Calculer l’enveloppe convexe de ces coefficients permet aussi de réduire significativement la boîte considérée, et de tester si la boîte contient un seul zéro, ou plusieurs. Dans ce dernier cas, la boîte est subdivisée selon son plus long côté, et les coefficients de F dans chaque nouvelle boîte peuvent être recalculés grâce à l’algorithme de de Casteljaou. Quand une boîte ne contenant qu’une solution est trouvée, quelques itérations de Newton sont utilisées avec comme solution initiale le centre de la boîte pour obtenir une solution précise. [Elber 01] utilise une représentation en splines de F .

Ces solveurs permettent donc d’obtenir tous les zéros de F (en un temps nécessairement exponentiel si le nombre de solutions est exponentiel). Leur principale limitation est due à la taille de la représentation d’un polynôme dans la base de Bernstein, exponentielle en le nombre de ses variables. Ceci la rend en pratique inapplicable aux systèmes de contraintes qui dépendent en général de plusieurs dizaines, voir centaines, de variables. [Fünzig 09] spécialise un tel solveur aux systèmes d’équations quadratiques en exploitant la programmation linéaire. Cette méthode permet de résoudre des systèmes en beaucoup d’inconnues, car elle considère un polytope défini par un nombre polynomial (en m) de faces plutôt que l’enveloppe convexe des 2^m coefficients de F dans la base de Bernstein. Ces faces permettent d’encadrer F et de réduire la boîte considérée. Le coût de cette méthode augmente de façon drastique quand les équations ont un degré plus grand que 2 car de nombreuses variables et équations intermédiaires sont ajoutées.

Méthodes par continuation

Les méthodes par continuation permettent de trouver toutes les solutions isolées d’un système polynomial. Elles interviennent pour une part importante dans nos travaux et la littérature les concernant est présentée dans la section 2.2 de ce chapitre. On en présente ici un bref résumé, pour montrer comment elles ont été utilisées pour résoudre des systèmes de contraintes.

Étant donné un système polynomial $F = 0$ appelé *système cible*, une borne supérieure du nombre de solutions isolées qu’il admet consiste en la borne de Bézout. Un système admettant autant de solutions, toutes connues, que cette borne peut alors être construit. Il est appelé *système initial*. En interpolant de manière continue le système initial en le système final, les solutions de ce dernier sont déformées en les solutions cherchées. Pour suivre ces solutions pendant leur déformation, on utilise une méthode appelée méthode de suivi de chemin, car elles appartiennent à des courbes appelées chemins d’homotopie, qui sont des variétés différentielles de dimension 1. Sous certaines conditions sur F et sur le système initial, toutes les solutions cherchées sont trouvées en suivant les chemins depuis toutes les solutions.

Le principal inconvénient de la méthode décrite ici est que beaucoup de chemins ne mènent pas à des solutions, ou à des solutions trouvées par des chemins précédents, car la borne de Bézout est en général bien supérieure au nombre de solutions cherchées. Beaucoup de travaux ont été menés afin de définir de nouvelles bornes plus fines que la borne de Bézout, menant à des méthodes efficaces. Cependant, dans le cas des systèmes issus de contraintes géométriques, où les équations sont en général quadratiques et “creuses” (elles font intervenir peu de termes), ces bornes n’améliorent pas la borne de Bézout. [Durand 00] propose de modifier le système $F = 0$ par des méthodes symboliques avant de le soumettre à une telle méthode, de façon à ce que ces bornes soient plus petites que celle de Bézout.

[Lamure 95] propose d’utiliser l’esquisse donnée par l’utilisateur pour construire le système initial. Comme une seule solution de ce système est connue (l’esquisse), une seule solution est trouvée. Elle présente cependant l’avantage d’être “proche” de l’esquisse, donc de la solution attendue par l’utilisateur. Le coût de cette méthode est alors comparable à celui d’une méthode itérative quand le chemin suivi présente de bonnes propriétés.

2.1.3 Méthodes hybrides

On qualifie d’hybrides des méthodes de résolution de systèmes de contraintes géométriques alliant plusieurs types d’approches. Par exemple les systèmes de CAO appliquent en général une méthode de décomposition au SCG à résoudre. Les sous-systèmes indécomposables sont ensuite résolus indépendamment par exemple par une méthode numérique. Les solutions des sous-systèmes sont ensuite ré-assemblées. Cette stratégie est classique, et il est sous-entendu en général qu’une méthode de résolution s’applique à un système indécomposable, résultant d’une phase de décomposition.

[Joan-Arinyo 99] s’intéresse à des problèmes de contraintes où certains paramètres de contraintes métriques ne sont pas constants, mais sont donnés par une relation fonctionnelle impliquant d’autres paramètres. Un graphe biparti inconnues-équations de ces relations fonctionnelles est construit, qui permet de donner un ordre pour attribuer des valeurs aux paramètres non constants en fonction de valeurs de paramètres déjà déterminés. Cette décomposition équationnelle est embarquée sous forme d’une règle dans une méthode géométrique.

Reparamétrisation

La méthode par *reparamétrisation* utilise successivement résolution géométrique ou combinatoire et résolution numérique. Elle fut initialement introduite par [Gao 02] qui l’applique à des systèmes de contraintes impliquant moins de 6 inconnues, et procède en deux phases :

Lors d’une première phase symbolique, un système est résolu par une méthode géométrique. Si la résolution est un succès, le plan de construction est évalué pour produire les solutions. Le cas échéant, certaines contraintes du système à résoudre sont remplacées par des contraintes métriques pour obtenir un système de contraintes *reparamétré* que la méthode géométrique parvient à résoudre. Un plan de construction des solutions du système reparamétré est alors obtenu, et son évaluation dépend des paramètres des contraintes qui ont été ajoutées, qui sont appelés paramètres guides. Quand une solution produite par le plan de construction satisfait les contraintes qui ont été supprimées du problème initial, elle est aussi une solution du problème initial.

Le plan de construction permet d’obtenir toutes les solutions du système reparamétré pour une valeur donnée des paramètres guides en appliquant ses instructions l’une après l’autre. Souvent, une instruction requiert de faire un choix parmi deux possibilités pour placer un objet géométrique. Appliquer les instructions pour tous les choix possibles construit toutes les solutions. Appliquer les

instructions pour une série de choix donnée construit une solution. En associant chaque choix à un nœud d'un arbre représentant toutes les combinaisons de choix possibles, on dira qu'une série de choix est une *branche* du plan de construction.

Trouver une solution du problème initial revient alors à trouver un couple formé d'une branche et d'un jeu de valeurs pour les paramètres guides tels que la figure construite en évaluant le plan de construction sur cette branche satisfasse les contraintes supprimées. La seconde phase, numérique, de la méthode par reparamétrisation, consiste à trouver un, plusieurs ou tous ces couples, afin de trouver toutes les solutions du système de contraintes initial.

Dans [Gao 02], c'est la méthode LIM qui est utilisée lors de la première phase. La phase numérique est réalisée en échantillonnant l'espace des paramètres guides. Le plan de construction est évalué pour chaque échantillon sur chacune de ses branches. Cette méthode est efficace pour de petits problèmes (moins de 6 inconnues) car d'une part la dimension de l'espace des paramètres guide est en général un ou deux³ et d'autre part, le plan de construction a peu de branches (moins de 2^5).

Quand on applique cette méthode à un SCG portant sur beaucoup d'inconnues, le nombre de branches explose et plus de paramètres guides sont nécessaires. [Fabre 07, Fabre 08] utilise une méthode à base de règles pour réaliser la phase symbolique, puis construit une fonction numérique dont les variables sont les paramètres guides et dont les composantes s'annulent quand toutes les contraintes supprimées sont satisfaites par la figure résultant de l'évaluation du plan de construction. L'esquisse est alors utilisée comme solution initiale de la méthode itérative de Newton-Raphson appliquée à cette fonction. Cette méthode est peu robuste car le plan de construction n'est pas toujours défini pour toutes valeurs des paramètres guides. Une itération fait souvent "sortir" le vecteur itéré de son domaine de définition.

En fait, la réalisation de la phase numérique est rendu difficile par le caractère multi-fonctionnel du plan de construction et la difficulté à caractériser le domaine de définition d'un plan de construction. Les problèmes liés à la phase symbolique sont dépendants de la méthode utilisée pour la phase numérique. Si c'est une méthode par échantillonnage, son coût augmente de manière exponentielle avec le nombre de paramètres guides. Minimiser le nombre de ces paramètres devient crucial, or c'est un problème difficile, non résolu à notre connaissance. De même, le SCG pourrait être reparamétré en minimisant le nombre de branches du plan de construction, ou de façon à faciliter la caractérisation de son domaine de définition.

2.1.4 Approche par décomposition

Les considérations structurelles d'un graphe de contraintes, ou d'un système d'équations, associés à un SCG permettent de le décomposer, afin d'une part de maîtriser le coût de sa résolution, et d'autre part de proposer à un utilisateur un parcours de l'espace des solutions guidé par les sous-systèmes. On considère en général que le coût de la décomposition d'un SCG en sous-systèmes est négligeable devant celui de sa résolution en entier.

On a déjà évoqué plus haut la décomposition due à Dulmage et Mendelsohn qui agit sur le graphe biparti d'un système d'équations, en le séparant de manière unique en parties structurellement sous, bien et sur-contraintes. Un autre résultat dû également à König, Dulmage et Mendelsohn et présenté dans [Ait-Aoudia 93] permet de décomposer la partie structurellement bien contrainte en sous-systèmes irréductibles, qui peuvent être résolus l'un après l'autre.

[Owen 91] propose de décomposer le graphe de contraintes d'un SCG en 2D selon ses paires d'articulation. La figure 2.2 représente un SCG en 2D, son graphe de contraintes, et une décomposition

3. On peut montrer, dans le cas des SCG en 2D à moins de six inconnues, qu'un paramètre guide suffit.

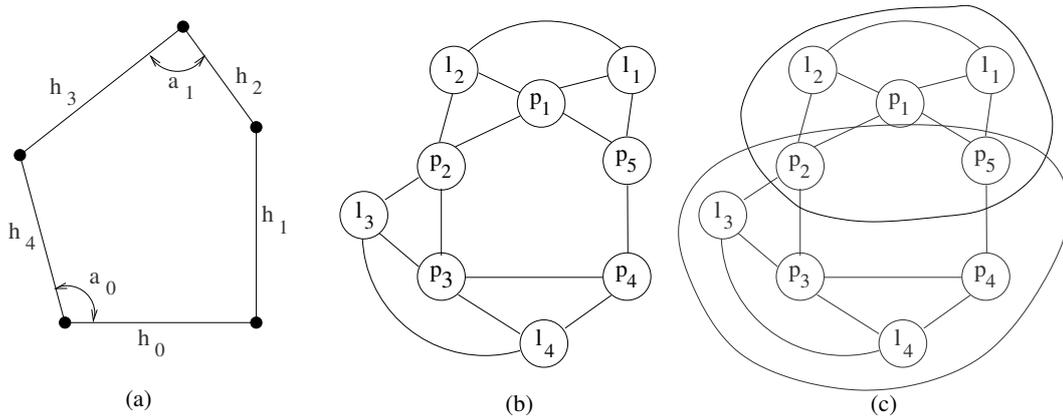


FIGURE 2.2 – (a) un système de contraintes en 2D; (b) le graphe de contrainte associé : les droites l_i sont introduites pour être les objets des contraintes d'angle, et ainsi éviter le formalisme des hyper-graphes; (c) une décomposition selon une paire d'articulation.

selon la paire d'articulation formée des sommets du graphe P_2 et P_5 . Si le SCG original est bien contraint (au sens de Laman), deux sous-graphes sont obtenus, et un est bien contraint (dans l'exemple de la figure 2.2 celui restreint à l'ensemble de sommets $\{P_2, P_5, P_1, l_2, l_1\}$), et l'autre est sous-contraint. Des contraintes virtuelles, impliquant les objets communs aux deux sous-systèmes (les sommets de la paire d'articulation selon laquelle le graphe a été décomposé), sont ajoutés à la partie sous-contrainte (ici par exemple une distance entre P_2 et P_5) pour la rendre bien contraint. Dans la méthode originale, les sous-systèmes bien contraints sont résolus en les décomposant encore suivant leurs paires d'articulation jusqu'à ce que chaque sous système (contenant éventuellement des contraintes virtuelles) soit réduit à un système rigide facilement résoluble, par exemple un graphe correspondant à un triangle rigide. Au moins un de ceux-ci ne contient pas de contrainte virtuelle, et peut être résolu; les valeurs des paramètres des contraintes virtuelles, que l'on suppose métriques, sont lues sur les solutions des sous-systèmes résolus pour résoudre les autres, et assembler les solutions.

Cette méthode peut aussi être utilisée uniquement pour la décomposition : une fois des sous-systèmes non-décomposables obtenus, ceux qui sont difficiles à résoudre sont traités par une méthode quelconque (en général numérique dans un système de CAO). D'autre part, elle s'étend facilement en 3D en considérant des triplets d'articulation plutôt que des paires.

Plutôt que de procéder de manière descendante pour décomposer le graphe de contraintes, la méthode de [Bouma 95] construit des *clusters* correspondant à des ensembles de sommets du graphe. Un repère, consistant en deux objets reliés par une contrainte, est d'abord choisi pour former un premier cluster; par exemple, sur l'exemple de la figure 2.2, l'ensemble $\{P_1, l_1\}$. Ce cluster est ensuite augmenté en y ajoutant itérativement des sommets reliés par deux contraintes à des sommets déjà présents dans le cluster. Dans l'exemple considéré, les sommets P_5, l_2 puis P_2 sont ajoutés au cluster initial. Quand plus aucun objet ne peut être ajouté au cluster, soit celui-ci contient tous les sommets du graphe, auquel cas le SCG est résoluble par la méthode des lieux LIM, soit ce n'est pas le cas, et un nouveau repère est choisi parmi les sommets n'appartenant pas aux clusters déjà trouvés. Dans l'exemple, supposons que le repère choisi soit $\{P_4, P_5\}$; le cluster ne peut pas être augmenté. Finalement, en choisissant $\{P_4, l_4\}$ comme repère, un dernier cluster $\{P_4, l_4, P_3, l_3, P_2\}$ est obtenu. La figure 2.3 présente les trois clusters obtenus pour l'exemple considéré.

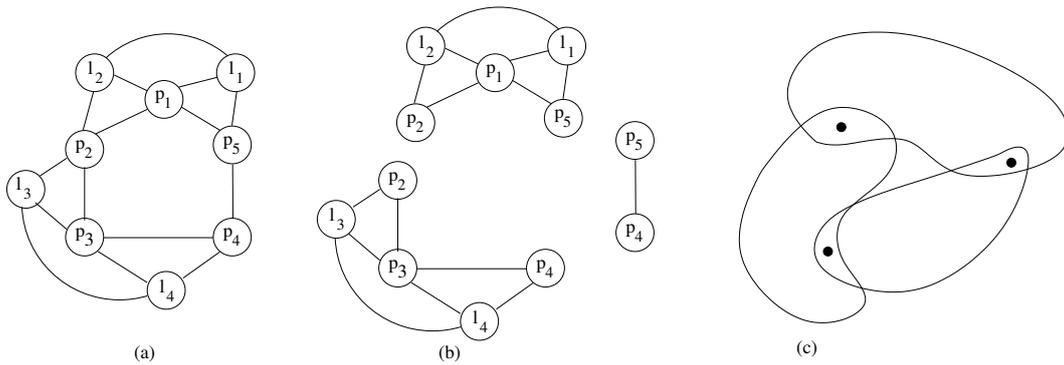


FIGURE 2.3 – (a) le graphe de contrainte d’un système en 2D ; (b) trois clusters obtenus par la méthode décrite dans [Bouma 95] ; (c) une des règles d’assemblage de la méthode des clusters.

Quand tous les sommets du graphe sont répartis dans des clusters, les objets rigides correspondant à chaque cluster sont assemblés grâce à trois règles d’assemblage, dont une est représentée sur la figure 2.3(c), qui peut s’énoncer par “si trois clusters partagent deux à deux un unique objet, l’assemblage des objets rigides solutions de chaque cluster faisant coïncider les points partagés est rigide”. Les solutions des trois clusters du graphe de contraintes représentés dans la figure 2.3 peuvent ainsi être assemblés. Cependant, si aucune des règles ne permet d’assembler les clusters, la résolution est un échec. Remarquons que cette méthode ne parvient à décomposer un SCG que si elle parvient également à le résoudre.

[Ait-Aoudia 01] allie décomposition géométrique du système de contraintes, décomposition équationnelle du système d’équations et résolution numérique. La première étape de la méthode décrite consiste en le calcul des clusters d’un graphe de contraintes. Les sommets appartenant à plusieurs clusters sont utilisés pour construire un nouveau graphe de contraintes, en y ajoutant si besoin des contraintes virtuelles. Ce nouveau graphe est appelé un squelette, et la méthode est appliquée itérativement à ce squelette, jusqu’à en obtenir un qui corresponde à un problème basique, par exemple trois points reliés deux à deux par une contrainte de distance. Quand pour un squelette (au pire le graphe de contraintes initial), aucune simplification n’est possible, le problème est traduit sous forme d’un système d’équations, décomposé par la méthode de [Ait-Aoudia 93] (présentée au début de cette sous-section), et chaque sous-système est résolu grâce à une méthode de bissection.

Les décompositions géométriques présentées ici trouvent leur justification dans l’invariance sous le groupe des transformations rigides des contraintes considérées. [Schreck 06] propose une généralisation de la décomposition en considérant d’autres groupes de transformations, par exemple celui des homothéties. [Mathis 10] propose une formalisation de la décomposition de SCG en sous-systèmes.

Citons enfin la méthode probabiliste du témoin déjà évoquée dans la sous-section 2.1.1 qui permet de détecter dans un SCG des sous-systèmes rigides maximaux. L’idée est de choisir dans un système une contrainte à supprimer. Le graphe de contraintes associé reste connexe, et les sous-systèmes rigides maximaux du système de contraintes correspondant sont trouvés grâce au test probabiliste du témoin. Ce dernier permet également de vérifier qu’une décomposition donnée par une autre méthode est correcte, notamment en ce qui concerne les dépendances dues à des théorèmes de géométrie. En effet, les méthodes de décomposition supposent que les contraintes structurellement indépendantes sont aussi algébriquement indépendantes.

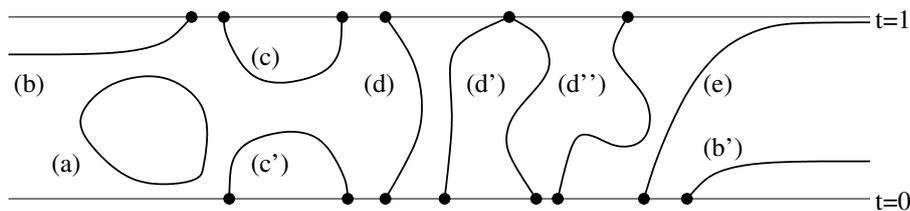


FIGURE 2.4 – Différentes formes de chemins d’homotopie. Les points noirs sur les hyperplans $t = 0$ et $t = 1$ sont les intersections des chemins avec ces hyperplans.

2.2 Méthodes par continuation

Le but d’une méthode par continuation, ou méthode par homotopie, est de trouver les zéros isolés d’une fonction $F : \mathbb{K}^m \rightarrow \mathbb{K}^m$, où $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . À partir de F , une fonction $H : \mathbb{K}^m \times [0, 1] \rightarrow \mathbb{K}^m$, appelée fonction d’homotopie, est définie telle que $\forall x \in \mathbb{K}^m, H(x, 1) = F(x)$. Si H satisfait à certaines conditions différentielles, $H^{-1}(0)$, l’ensemble des zéros de H , consiste en plusieurs composantes connexes, qui sont des variétés différentielles de dimension 1 (*c.à.d.* des courbes plongées dans $\mathbb{K}^m \times [0, 1]$). En posant $\mathcal{P}_0 = \mathbb{K}^m \times \{0\}$ et $\mathcal{P}_1 = \mathbb{K}^m \times \{1\}$, les bords de ces composantes connexes sont exactement leurs intersections avec $\mathcal{P}_0 \cup \mathcal{P}_1$. Pour chaque x tel que $F(x) = 0$, il existe une composante connexe \mathcal{S}_i de $H^{-1}(0)$ qui “arrive” en $(x, 1)$ (*c.à.d.* avec $(x, 1) \in \mathcal{S}_i$), et pour chaque x tel que $H(x, 0) = 0$, il en existe une, \mathcal{S}_j , qui “part” de $(x, 0)$ (*c.à.d.* $(x, 0) \in \mathcal{S}_j$). Ces composantes connexes sont appelées des chemins d’homotopie de H .

En supposant qu’un zéro x_j de $H(x, 0)$ soit connu, le chemin d’homotopie \mathcal{S}_j peut-être suivi depuis $(x_j, 0)$ par une méthode de suivi de chemins jusqu’à ce qu’il croise l’hyperplan \mathcal{P}_1 . Plusieurs cas de figure peuvent se présenter, selon la topologie de \mathcal{S}_j . Un résultat de topologie différentielle, rappelé en annexe A, établit que le chemin \mathcal{S}_j est difféomorphe (*c.à.d.* qu’il est en bijection dérivable, d’inverse dérivable) à un cercle de \mathbb{R}^2 , ou à un intervalle de \mathbb{R} . Comme $\mathcal{S}_j \subset (\mathbb{K}^m \times [0, 1])$, il ne peut être difféomorphe ni à un cercle, ni à un intervalle ouvert, car ces deux ensembles n’ont pas de bord, donc aucune intersection avec \mathcal{P}_0 . \mathcal{S}_j est alors difféomorphe à un intervalle fermé ou semi-ouvert de \mathbb{R} . Dans le premier cas, le chemin d’homotopie intersecte soit \mathcal{P}_1 , et un zéro de $H(x, 1)$ donc de $F(x)$ est trouvé, soit une seconde fois \mathcal{P}_0 . Dans le second cas, \mathcal{S}_j n’intersecte ni \mathcal{P}_0 ni \mathcal{P}_1 en dehors de x_j , et a une longueur infinie. Quand la dernière composante de \mathcal{S}_j est croissante⁴, celle-ci tend vers une limite finie t_∞ car elle est majorée. On peut alors en extraire une suite $(x_n, t_n)_{n \in \mathbb{N}}$ avec $\lim_{n \rightarrow \infty} t_n = t_\infty \in]0, 1[$, et on dira que x_n tend vers une solution à l’infini du système $H(x, t_\infty) = 0$. Le suivi de \mathcal{S}_j prend alors un temps infini si une telle situation n’est pas détectée.

La figure 2.4 présente différentes formes possibles de chemins d’homotopie. Les chemins de forme (a), difféomorphes à des cercles, n’ont pas d’intersection avec l’hyperplan \mathcal{P}_0 et ne sont donc jamais suivis. Les chemins de forme (b), (b’) et (e) sont difféomorphes à des intervalles semi-ouverts, et ont une longueur infinie; (e) tend vers l’infini quand sa composante en t tend vers 1, c’est à dire que sa limite est une solution à l’infini de F . Un chemin de forme (c) (respectivement (c’)) part d’une solution de $H(x, 0)$ (resp. F) pour arriver à une solution du même système. Les chemins de la forme (d) et (d’’) partent d’une solution de $H(x, 0)$ pour arriver à une solution de F . Si il existe $0 \leq t \leq 1$ tel que deux chemins se croisent en un point (x, t) , alors la composante connexe formée par ces deux chemins n’est pas une variété de dimension 1. La composante (d’) illustre ce cas de figure, avec $t = 1$; cependant, l’intersection de cette composante connexe avec l’ensemble

4. Comme \mathcal{S}_j est difféomorphe à un intervalle $[a, b[\subset \mathbb{R}$, il existe une paramétrisation $\varphi_j : [a, b[\rightarrow \mathbb{K}^m \times [0, 1]$ avec $\varphi_j([a, b[) = \mathcal{S}_j$ et $\varphi_j(a) = (x_j, 0)$. La croissance de la dernière composante de \mathcal{S}_j s’entend par rapport à la variable de φ_j .

$\mathbb{K}^m \times [0, 1[$ est composée de deux chemins qui sont des variétés de dimension 1. Enfin, les chemins de la forme (c), (c') et (d'') ont leur dérivée par rapport à t qui s'annule en certains points.

Les conditions sous lesquelles les composantes connexes de l'ensemble $H^{-1}(0)$ sont des variétés différentielles de dimension 1 portent d'une part sur la dérivabilité de H (on considérera en général que H est de classe C^∞), et d'autre part sur le rang des applications linéaires $\partial H_{(x,t)} : \mathbb{K}^{m+1} \rightarrow \mathbb{K}^m$ qui sont les dérivées de H en les points $(x,t) \in H^{-1}(0)$. Si en un point (x,t) , $\partial H_{(x,t)}$ n'est pas de rang plein, on dit que 0 est une valeur critique de H . Dans ce cas, la caractérisation de l'ensemble $H^{-1}(0)$ en variétés différentielles de dimension 1 n'est plus vraie (et le suivi de ses composantes connexes par une méthode de suivi de courbe n'est plus possible). Si $H^{-1}(0)$ ne contient pas de point critique, on dira que 0 est une valeur régulière de H .

Quand $\mathbb{K} = \mathbb{C}$, les points critiques peuvent être évités en appliquant une petite perturbation complexe à la fonction d'homotopie. Un argument probabiliste assure alors que 0 est une valeur régulière de H , et les chemins d'homotopie de H ont leur dernière composante strictement croissante, ce qui facilite la phase numérique de suivi de chemin. Quand $\mathbb{K} = \mathbb{R}$, cet argument n'est plus valable, et la dérivée partielle par rapport à la dernière composante des chemins d'homotopie peut s'annuler.

Dans le cas où les composantes de F sont des polynômes, il est possible de construire H de façon à ce que tous les zéros de $H(x, 0)$ soient connus, et soient en nombre supérieur au nombre de zéros de F . Si chaque chemin partant d'un zéro de $H(x, 0)$ arrive à un zéro de F (c.à.d chaque chemin est de la forme (d), (d') ou (d'') sur la figure 2.4), autant de zéros (non nécessairement différents) de F que de zéros de $H(x, 0)$ sont trouvés.

Cette introduction a présenté les principaux points à considérer pour appliquer une méthode par continuation cherchant les zéros d'une fonction $F : \mathbb{K}^m \rightarrow \mathbb{K}^m$. La littérature les aborde dans le cas où F est polynomiale, ce qui permet d'estimer le nombre de ses zéros pour poser un système initial. Elle décrit des méthodes permettant de trouver toutes les solutions isolées du système d'équations $F(\cdot) = 0$ dont, pour certaines, des implémentations libres sont disponibles. Notons finalement qu'on a choisi ici le point de vue de la topologie différentielle pour décrire ces méthodes. Beaucoup des articles cités ici utilisent le cadre de la géométrie algébrique. Cette approche, plus à même de traduire les structures induites par la nature polynomiale des systèmes considérés est cependant moins accessible. Elle sera utilisée dans la sous-section 2.2.6 qui présente des méthodes, utilisant l'homotopie, permettant de trouver en plus des solutions isolées d'un système, ses ensembles de solutions de dimension strictement positives. Ces méthodes sont issues d'un domaine récent nommé "numerical algebraic geometry", pour lequel nous garderons l'appellation anglophone.

2.2.1 Poser la fonction d'homotopie

On considère un système P de m équations polynomiales P_i en les m inconnues $x = (x_1, \dots, x_m)$ et on appelle *degré* d_i de P_i le degré de son monôme dominant. Le nombre de Bézout, égal au produit des d_i , est une borne supérieure du nombre de solutions de P (voir [van der Waerden 49]). Une idée utilisée dans [Drexler 78, Garcia 79, Chow 79, Morgan 86a] est alors de construire un système canonique Q de m équations en m inconnues, appelé système initial, ayant au moins autant de solutions que le nombre de Bézout de P . La fonction d'homotopie H est alors construite en interpolant Q en P grâce au paramètre d'interpolation $t \in [0, 1]$ par $H(x, t) = (1-t)Q(x) + tP(x)$. De chacune des d' solutions de Q part un chemin d'homotopie, et si 0 est une valeur régulière de H , ils sont tous difféomorphes à un intervalle fermé ou semi-ouvert de \mathbb{R} . Un choix simple pour le système initial est le système Q dont chaque composante est le polynôme $Q_i = x_i^{d'_i} - a_i$, avec $d'_i \geq d_i$ pour chaque i .

En choisissant $d'_i = d_i + 1$ pour tout i et en considérant H comme une fonction de variables complexes à valeurs complexes, [Garcia 79] montre que la dernière composante (celle correspondant au paramètre d'interpolation) de la dérivée de H en un point (x, t) de $H^{-1}(0)$ ne s'annule jamais (*c.à.d* que les chemins d'homotopie ne sont pas de la forme (c), (c') ou (d'')), et qu'il existe pour chaque $0 \leq t < 1$ un compact $K_t \subset \mathbb{C}^m \times [0, 1]$ tel que les zéros de H soient dans K_t (donc les chemins ne sont pas de la forme (b) ou (b')). Si 0 est une valeur régulière de H , ses chemins d'homotopie arrivent soit à une solution de P , soit tendent vers une solution à l'infini quand t tend vers 1 (les chemins sont strictement croissants en leur dernière composante, car la dérivée en t ne s'annule jamais). Comme plusieurs chemins ne peuvent arriver à la même solution de P (sinon ces chemins ne seraient pas des variétés), ils sont tous de la forme (d) ou (e). En notant d le nombre réel de solutions de P dans \mathbb{C}^m , cette construction de H implique que $d' - d$ chemins aient une longueur infinie (chemins de la forme (e)). Cependant, en suivant tous les chemins d'homotopie depuis les solutions de Q , toutes les solutions de P sont obtenues.

[Chow 79] utilise un système initial Q avec $d' = \prod d_i$ solutions, et réalise l'homotopie en posant $H(x, t) = (1 - t)Q(x) + tP(x) + t(t - 1)R(z)$. Le troisième terme $t(t - 1)R(z)$, qui s'annule quand $t = 0$ ou $t = 1$, est construit de telle sorte que la dernière composante de la dérivée de H ne s'annule pas, et ne tende vers l'infini que quand t tend vers 1. 0 est considéré comme une valeur régulière seulement pour la restriction de H à l'ensemble $\mathbb{C}^m \times [0, 1[$, ce qui signifie que pour certains $(x, 1)$ tels que $H(x, 1) = 0$, la dérivée de H en $(x, 1)$ n'est pas nécessairement de rang plein, et que plusieurs chemins d'homotopie peuvent arriver à la même solution.

Contrairement aux travaux déjà présentés, [Morgan 89] n'utilise pas de système canonique, mais se place dans un cadre où les coefficients des monômes des P_i dépendent de paramètres. Si pour certaines valeurs de ces paramètres les solutions de P sont connues, ces valeurs sont interpolées en un jeu de valeurs pour lesquelles on cherche les solutions de P . Toutes les solutions cherchées peuvent être trouvées à partir de toutes les solutions complexes du système initial (*c.à.d* le système avec les valeurs des paramètres pour lesquelles les solutions sont connues). En général, le système initial a (beaucoup) moins de solutions qu'un système canonique, et il découle que moins de chemins doivent être suivis. Les résultats sont prouvés dans un cadre suffisamment large pour permettre à \mathbb{K}^m d'être toute compactification de \mathbb{C}^m , ce qui rend cette méthode compatible avec celles présentées plus bas, qui permettent d'une part (dans la sous-section 2.2.3) de détecter les solutions à l'infini, et d'autre part (dans la sous-section 2.2.4) de diminuer le nombre de solutions du système initial, et ainsi le nombre de chemins à suivre.

2.2.2 Régularité de la valeur 0

Quand $H^{-1}(0)$ contient un point (x, t) tel que l'application linéaire $\partial H_{(x,t)}$ n'est pas de rang plein, la composante connexe de $H^{-1}(0)$ qui contient ce point n'est pas nécessairement une variété différentielle de dimension 1. Il peut s'agir, par exemple, de la réunion d'une variété de dimension 1 et d'une variété de dimension supérieure. La figure 2.5 illustre le cas où une composante connexe est la réunion de deux ensembles de dimension 1 et d'un de dimension 2. Un théorème prouvé par Sard (voir [Milnor 65]) établit que l'ensemble des valeurs régulières de H est dense dans \mathbb{K}^m .

[Chow 79] utilise un argument probabiliste pour s'assurer que 0 soit une valeur régulière de H . Si les composantes Q_i du système initial Q sont $Q_i = x_i^{d'_i} - a_i$ (ou d'une autre forme impliquant un paramètre a_i), H peut être vue comme une fonction $H(x, t, a) : \mathbb{C}^m \times \mathbb{R} \times \mathbb{C}^m$. Grâce au théorème de transversalité, appelé aussi théorème de Sard (voir [Chow 79], ou [Allgower 97], page 77), on peut montrer que si 0 est une valeur régulière de F , alors pour presque toutes les valeurs (au sens de la mesure de Lebesgue) des coefficients a_i de Q , 0 est une valeur régulière de H . En conséquence,

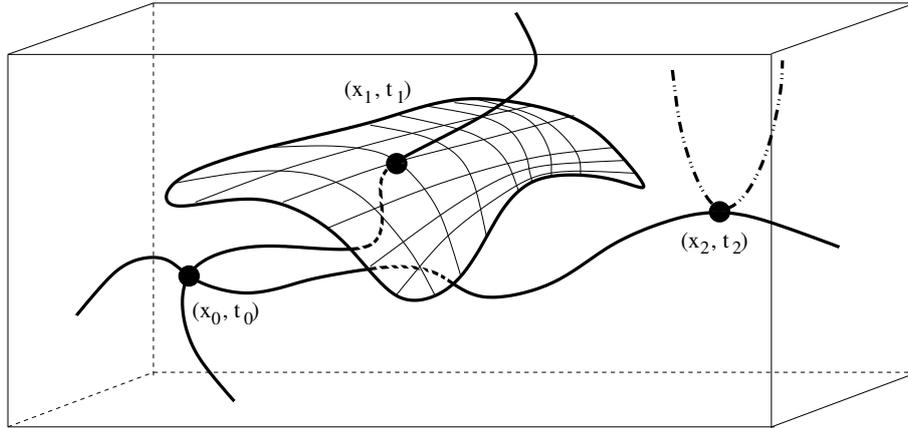


FIGURE 2.5 – Une composante connexe de $H^{-1}(0)$ quand 0 n'est pas une valeur régulière de H . Les (x_i, t_i) sont des points critiques de H . La courbe en pointillés (un long, deux courts) est une composante imaginaire pure de H .

une petite perturbation de $a = (a_0, \dots, a_m) \in \mathbb{C}^m$ assure, avec une probabilité égale à 1, que 0 est une valeur régulière de H .

[Li 93] montre que quand $\mathbb{K} = \mathbb{R}$, les points critiques de H constituent un ensemble fini de points en lesquels la dernière composante de la tangente au chemin s'annule. Ces points sont critiques si on considère H comme une fonction complexe, car ils sont l'intersection d'une composante réelle pure et d'une imaginaire pure de $H^{-1}(0)$, mais n'affectent pas le suivi de la courbe si il est réalisé dans $\mathbb{R}^m \times \mathbb{R}$. Sur la figure 2.5, le point (x_2, t_2) est un de ces points critiques, il est l'intersection d'une composante réelle pure (en trait plein), et d'une composante imaginaire pure (en traits pointillés) de la composante connexe.

2.2.3 Détecter les solutions à l'infini

Dans les contributions présentées ci-dessus, les chemins suivis sont de la forme (d), (d') et (e) (voir figure 2.4). Les chemins de la forme (d) et (d') mènent à des solutions du système P à résoudre, et les chemins de la forme (e) tendent vers des solutions à l'infini de P . Cependant, le suivi de ces derniers nécessite un temps de calcul infini. La procédure de suivi doit donc être stoppée quand l'une des composantes x_j du chemin devient plus grande qu'un certain seuil. Quand ce seuil est grand, le suivi des chemins dont une des composantes dépasse ce seuil est pénalisant en terme de temps de calcul, et quand il est petit, le risque qu'un chemin dont le suivi a été stoppé arrive à une solution de P de très grande norme augmente.

Pour résoudre ce problème, [Wright 85] réalise l'homotopie dans l'espace projectif complexe \mathbb{P}^m , qui est l'ensemble des droites de \mathbb{C}^{m+1} passant par 0. Une fonction polynôme $P_i(x_1, \dots, x_m)$ définie sur \mathbb{C}^m de degré d_i est transformée en une fonction $P'_i(x_0, x_1, \dots, x_m)$ définie sur \mathbb{P}^m en multipliant chaque monôme de P_i par $x_0^{d_i - d''}$, où d'' est le degré de ce monôme. On a alors $P'_i(1, x_1, \dots, x_m) = P_i(x_1, \dots, x_m)$, et chaque racine (x_1^*, \dots, x_m^*) de P_i correspond à une unique racine $(1, x_1^*, \dots, x_m^*)$ de P'_i et à une droite paramétrée par x_0 dont les points sont $(x_0, \frac{x_1^*}{x_0}, \dots, \frac{x_m^*}{x_0})$. Les x_l , pour $0 \leq l \leq m$, jouent un rôle symétrique dans P'_i et cette droite peut être paramétrée par tout x_l . On dit que P'_i est une homogénéisation de P_i , et que $P' = (P'_1, \dots, P'_m)$ est une homogénéisation de $P = (P_1, \dots, P_m)$. [Wright 85] cherche les zéros d'une homogénéisation de P et construit le système

initial Q dans l'espace projectif \mathbb{P}^m . Les chemins sont suivis depuis les solutions de Q de la forme $(1, x_1, \dots, x_m)$. Lors du suivi d'un chemin dont la norme d'une composante x_l tend vers l'infini, son équivalent paramétré par x_l est calculé, et le chemin passant par ce point, qui a une longueur finie, peut être suivi jusqu'à l'hyperplan \mathcal{P}_1 . Si son intersection avec cet hyperplan est un point dont la première composante est nulle et la l -ème est 1, c'est une solution à l'infini de P .

[Morgan 86b] réalise également l'homotopie dans \mathbb{P}^m , et résout le système $P'' = (P'_0, P'_1, \dots, P'_m)$ où P'_0 est une nouvelle équation polynomiale construite de sorte que P'' n'ait pas de solutions à l'infini. La fonction d'homotopie associée n'a pas de chemin de longueur infinie.

2.2.4 Réduire le nombre de chemins à suivre

Un frein à l'efficacité des méthodes par continuation est la sur-estimation du nombre de solutions de P , le système à résoudre (ou de son homogénéisation P'). En effet, si P a d solutions et le système initial $d' \geq d$ solutions, le suivi de $d' - d$ chemins est superflu (il n'amène pas de nouvelles solutions). En choisissant un système initial ayant autant de solutions que la borne de Bézout de P , on a en général $d' \gg d$.

Une idée présentée dans [Morgan 87] est d'exploiter une structure h -homogène du système à résoudre pour réduire la borne supérieure du nombre de solutions. L'homogénéisation de P est réalisée de telle sorte que l'ensemble des variables puisse être partitionnée en h ensembles. La borne de Bézout de cette homogénéisation est plus petite que la borne de Bézout de P . [Wampler 93] présente une méthode efficace pour construire un système initial adaptée à cette structure h -homogène, et [Li 01] un algorithme permettant de trouver une partition des variables diminuant la borne de Bézout, sans toutefois assurer qu'il soit minimum. [Malajovich 07] montre que trouver ce minimum est un problème de la classe NP. Finalement, [Verschelde 94a] exploite une symétrie dans la distribution des variables dans les monômes de P , et montre que cette symétrie permet de réduire encore sa borne de Bézout.

Le théorème de Bernshtein [Bernshtein 75] fournit une borne supérieure au nombre de solutions d'un système de polynômes P plus précise que la borne de Bézout et ses adaptations à un système h -homogène, quand les composantes de P sont creuses (*c.à.d* impliquent peu de monômes). Cette borne utilise le concept de polytope de Newton, construit à partir des exposants des monômes des P_i . Une borne au nombre de solutions de P , appelée borne BKK⁵, peut être calculée à partir de ces polytopes, et [Verschelde 94b] l'utilise pour définir un système initial canonique Q . [Huber 95] propose un algorithme pour calculer la borne BKK utilisant des subdivisions des polytopes, et utilise autant de systèmes initiaux que de subdivisions pour définir autant de fonctions d'homotopie, dont tous les chemins d'homotopie sont suivis. Cette méthode est appelée *homotopie polyédrale*.

Quand les polynômes de P dépendent d'un vecteur de coefficients $c \in \mathbb{C}^l$, et que d est une borne supérieure du nombre de solutions de P atteinte pour un $c_0 \in \mathbb{C}^l$, alors l'ensemble des vecteurs de paramètres tels que P admet d solutions est dense dans \mathbb{C}^l . En particulier, si c est choisi au hasard, P admet d solutions pour les paramètres c , et pour tout autre choix de c , P admet au plus d solutions. La méthode appelée *cheater's homotopy* ([Li 89]) utilise cette propriété dans le cas où P doit être résolu plusieurs fois avec différentes valeurs de paramètres. Le système est d'abord résolu par une des méthodes ci-dessus avec un vecteur de paramètres c_0 choisi au hasard. d solutions sont alors obtenues (c'est la phase de triche). Lors des résolutions suivantes de P avec comme paramètres ceux demandés par l'utilisateur, les d solutions trouvées lors de la phase de triche sont utilisées comme solutions initiales dans une homotopie où le système initial est P , avec comme paramètres c_0 . Cette homotopie est réalisée dans l'espace des paramètres, comme justifié dans [Morgan 89] (en général, d est très inférieur à la norme de Bézout ou à la borne BKK).

5. BKK pour Bernshtein, Kushnirenko et Khovanskii

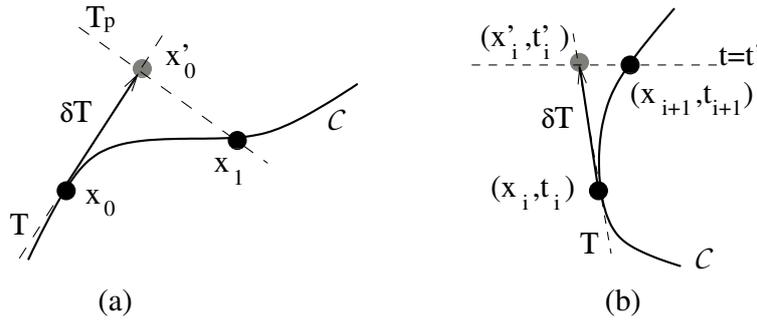


FIGURE 2.6 – (a) : la première étape d’une méthode de prédiction-corréction classique depuis la solution initiale x_0 . (b) : une étape de prédiction correction d’une *climbing homotopy*.

2.2.5 Méthodes de suivi de courbe

Quand la fonction d’homotopie $H(x, t) : \mathbb{K}^m \times [0, 1] \rightarrow \mathbb{K}^m$ a été définie, certains zéros de $H(x, 0)$ (tous dans le cas des méthodes décrites plus haut) sont connus ; chacune de ces solutions appartient à un chemin d’homotopie de H qui est supposé être une variété différentielle de dimension 1. Une méthode numérique de suivi de courbe est alors appliquée. Les chemins sont des courbes plongées dans $\mathbb{K}^m \times [0, 1]$. On présente ici des méthodes permettant de suivre des courbes de \mathbb{R}^{m+1} qui sont l’ensemble des zéros d’une fonction $H : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$. Dans le cadre des méthodes décrites plus haut, les fonctions d’homotopie ont en général des variables complexes et sont à valeurs complexes. Cependant, on se ramène facilement au cas réel en voyant \mathbb{C} comme \mathbb{R}^2 .

Deux familles de méthodes permettent de calculer les points d’une telle courbe. La première, appelée prédiction-corréction, calcule une suite $(x_n)_{n \in \mathbb{N}}$ de points de la courbe à partir d’un point initial x_0 , où x_{i+1} est construit à partir de x_i en une étape de prédiction, pour obtenir x'_i , et une étape de correction pour obtenir x_{i+1} (la figure 2.6(a) présente la première itération depuis x_0). La prédiction est réalisée le long de la tangente T à la courbe en x_i , avec un pas de prédiction δ . x'_i est corrigé en x_{i+1} avec une méthode numérique de type quasi-Newton. La seconde méthode pave \mathbb{R}^{m+1} par des simplexes de dimension $m + 1$, et calcule l’ensemble des simplexes contenant la courbe suivie, en approximant H sur chaque simplexe par une application affine $H^\vee : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$. Cette méthode est appelée approximation linéaire par morceaux, abrégé en PLI, pour Piecewise Linear Approximation.

Prédiction-corréction

Dans sa version la plus classique, une itération de prédiction-corréction se résume à :

- (E1) Calculer la tangente T de la courbe \mathcal{C} en x_i .
- (E2) Choisir un pas de prédiction δ , et calculer $x'_i = x_i + \delta T$
- (E3) Corriger x'_i en $x_{i+1} \in \mathcal{C}$ grâce à une méthode numérique.

L’étape (E1) est effectuée en calculant la matrice jacobienne J_{x_i} de H en x_i ; J_{x_i} est la matrice de application linéaire $\partial H_{x_i} : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$. La tangente à \mathcal{C} en x_i est alors le noyau de cette application linéaire, calculée avec les outils classiques de l’algèbre linéaire. La jacobienne de H en x_0 peut être obtenue soit de manière symbolique si une expression de H est connue, soit numériquement par la méthode des différences finies, soit encore en utilisant une méthode de différentiation de code (voir par exemple [Bischof 08]). Un vecteur directeur T (en général de norme unitaire) de la tangente est choisi. Dans le cas des méthodes présentées plus haut où $x = (u, t)$ et où la dérivée partielle

$(\frac{\partial H}{\partial t})_x$ ne s'annule pour aucun $x \in \mathcal{C}$, la dernière composante de T n'est jamais nulle ; il est alors possible de choisir T tel que sa dernière composante soit égale à 1. Ces homotopies (pour lesquelles les chemins sont monotones en t) sont parfois appelées *climbing homotopies*. Ici la prédiction est calculée le long de la tangente, et un tel prédicteur est appelé dans la littérature prédicteur d'Euler. Les dérivées d'ordre supérieur peuvent également être utilisées ; le lecteur intéressé se référera à l'état de l'art de [Allgower 97].

Lors de l'étape (E3) x'_i est corrigé en un point x_{i+1} tel que $H(x_{i+1}) = 0$. Mais cette équation est sous-déterminée (elle admet une infinité de solutions, qui sont les points de \mathcal{C}). Pour y remédier, x_{i+1} peut être cherché comme la solution du problème de minimisation $\min_x \{\|x'_i - x\|_2 | H(x) = 0\}$ grâce à une méthode dite de Gauss-Newton (voir par exemple [Ortega 87]). Une seconde approche concatène à H une équation additionnelle mesurant l'avancement sur \mathcal{C} , d'où son nom de *pseudo longueur d'arc*. Un choix simple pour cette équation est $x.T = 0$, où $.$ est le produit scalaire usuel. L'utiliser revient à chercher x_{i+1} sur l'hyperplan T_p de vecteur normal T , comme illustré dans la figure 2.6(a). Dans le cas des *climbing homotopies*, cette équation peut être $t = t'$ où $t' \in \mathbb{R}$ est la dernière composante de x'_i , ce qui revient à chercher x_{i+1} tel que $H(x_{i+1}, t') = 0$; ce principe est illustré sur la figure 2.6(b). Le système $(H(x), f_{add}(x)) = 0$ où f_{add} est l'équation additionnelle peut alors être résolu par une méthode itérative comme la méthode de Newton.

L'étape (E2) est réalisée en fonction de la méthode de correction utilisée dans l'étape (E3). Dans une implémentation où aucune stratégie d'adaptation du pas δ n'est mise en œuvre, ce pas est fixé à l'avance et l'étape (E2) est triviale. C'est néanmoins dans cette étape que réside la plus grande possibilité d'optimisation, en termes de robustesse et d'efficacité, pour une méthode de prédiction-corrrection : si le pas choisi est trop grand le point prédit peut se trouver en dehors de la zone de convergence de la méthode itérative de correction, et s'il est trop petit, le suivi du chemin est inefficace. [Durand 98] propose de doubler ce pas (jusqu'à atteindre une valeur maximale) tant que la méthode de correction converge vers une solution, et de le diviser par deux tant qu'elle échoue. Si ce pas atteint une valeur minimale, le suivi est stoppé. Des approches plus analytiques utilisent la théorie de Newton-Kantorovich pour calculer le rayon de convergence de la méthode de correction (Newton ou Gauss-Newton) afin de déterminer le pas (voir par exemple [Yakoubsohn 95] ou [Den Heijer 81]).

Si deux chemins d'homotopie, ou deux portions d'un même chemin, sont proches l'un de l'autre, un pas inadéquat peut amener le correcteur à "changer" de chemin, ou à rater une portion d'un chemin. Pour éviter ces erreurs, des approches plus récentes utilisent l'arithmétique par intervalles pour garantir le suivi de courbe. La correction est alors effectuée par une adaptation d'une méthode itérative. [Kearfott 94] présente un algorithme pour choisir un pas optimum, au sens où il s'agit du plus grand pas tel que le correcteur converge et ne change pas de chemin. [Faudot 07] propose une méthode garantissant que le correcteur ne change pas de courbe, compatible avec différentes arithmétiques par intervalles, capable de détecter que la courbe passe près d'un point critique. Il est suggéré d'utiliser alors la méthode PLI pour suivre la courbe au voisinage de ce point.

[Bates 08] remarque que beaucoup d'échecs lors du suivi d'un chemin sont dus à l'imprécision du calcul flottant, et qu'à l'inverse une précision trop élevée est coûteuse et souvent superflue. Les auteurs présentent une méthode permettant d'adapter cette précision pour le calcul de la jacobienne, de son inverse et des itérations du correcteur.

Concluons en remarquant que dans le cadre d'une méthode par continuation classique, le processus de suivi de courbe est stoppé quand la prédiction dépasse l'hyperplan \mathcal{P}_1 . La solution du système initial est alors obtenu en effectuant une prédiction sur cet hyperplan (en choisissant un pas δ approprié), et une correction avec l'équation ajoutée $t = 1$.

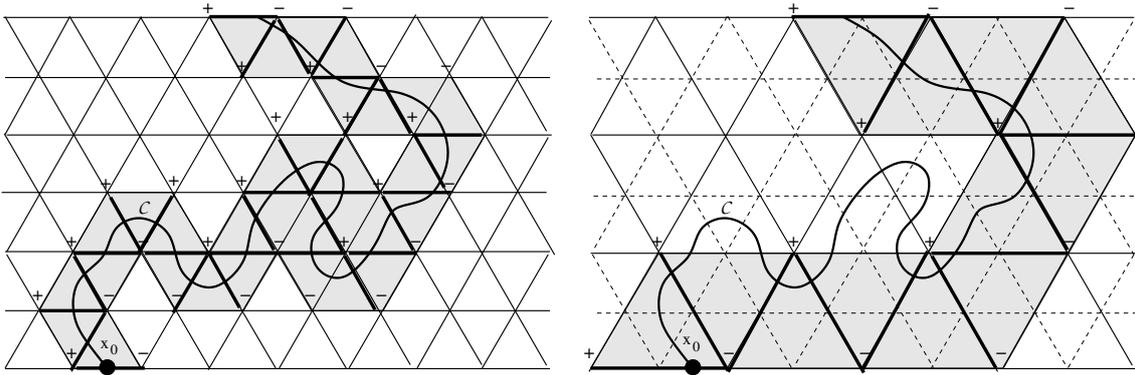


FIGURE 2.7 – La méthode de suivi de courbe PLI quand $m = 1$, pour deux pavages de \mathbb{R}^2 par des simplexes de tailles différentes. En gris, l'ensemble de simplexes résultant du suivi de la courbe \mathcal{C} par PLI.

Approximation Linéaire par Morceaux

Le principe de la méthode de suivi de courbe PLI suit un principe très intuitif comme on peut le voir en étudiant le cas où $m = 1$. La courbe \mathcal{C} à suivre est donc plongée dans \mathbb{R}^2 et définie par $H^{-1}(0)$, où $H : \mathbb{R}^2 \rightarrow \mathbb{R}$. \mathbb{R}^2 peut être pavé par des simplexes de dimension 2, notés ${}^2\mathcal{S}^i$, qui sont des triangles. Un tel simplexe est vu comme l'enveloppe convexe d'un ensemble de simplexes de dimension 1, notés ${}^1\mathcal{S}^i$ qui sont ses faces (ici les côtés du triangle), eux-mêmes enveloppe convexe de 2 simplexes de dimension 0 (ici les sommets du triangle), qui sont des points de \mathbb{R}^2 . La figure 2.7, qui illustre la méthode PLI, présente deux pavages de \mathbb{R}^2 par des triangles ; sur sa partie gauche, les triangles ont des côtés deux fois plus petits que sur sa partie droite.

Supposons qu'un point $x_0 \in \mathcal{C}$ appartenant à une unique face ${}^1\mathcal{S}^0$ (simplexe de dimension 1) du pavage soit connu. Cette face est commune à exactement deux simplexes de dimension 2 ; on en choisit un que l'on note ${}^2\mathcal{S}^0$ et ce choix correspond au sens dans lequel on va suivre \mathcal{C} . Si x_0 est la seule intersection de \mathcal{C} avec ${}^1\mathcal{S}^0$ (H s'annule une seule fois sur l'ensemble des points de ${}^1\mathcal{S}^0$), alors si H est continue, elle prend des valeurs de signes opposés en les deux sommets de ${}^1\mathcal{S}^0$. Sur la figure 2.7, les signes à côté des sommets des triangles indiquent les signes de la valeur de H . Si H est non nulle en le troisième sommet de ${}^2\mathcal{S}^0$, son signe est opposé à celui que prend H en un des deux sommets de ${}^1\mathcal{S}^0$. Si H est continue, elle s'annule au moins une fois sur la face ${}^1\mathcal{S}^1$, enveloppe convexe des deux sommets de signes opposés par H . ${}^1\mathcal{S}^1$ appartient à un unique simplexe ${}^2\mathcal{S}^1$ différent de ${}^2\mathcal{S}^0$ qui contient, par le même raisonnement, une unique autre face ${}^1\mathcal{S}^2$ dont les sommets ont des signes opposés par H . En itérant ce processus, on obtient deux ensembles $({}^1\mathcal{S}^n)_{n \in I_1}$ et $({}^2\mathcal{S}^n)_{n \in I_2}$ qui approximent \mathcal{C} . Dans la figure 2.7, $({}^2\mathcal{S}^n)_{n \in I_2}$ est l'ensemble des triangles grisés, et $({}^1\mathcal{S}^n)_{n \in I_1}$ l'ensemble des arêtes en traits épais.

Quand les deux sommets d'une face ${}^1\mathcal{S}^j$ prennent par H des valeurs de signes opposés, \mathcal{C} ne coupe pas nécessairement ${}^1\mathcal{S}^j$ une unique fois, mais un nombre impair de fois ; de même, si les deux sommets prennent des valeurs de même signe, \mathcal{C} coupe ${}^1\mathcal{S}^j$ un nombre pair de fois (qui peut être nul). Dans sa partie droite, la figure 2.7 présente un cas où des faces du pavage ont des sommets qui prennent par H des valeurs de signes opposés, et où \mathcal{C} a un nombre pair non nul d'intersections avec ces faces. En conséquence, \mathcal{C} traverse des simplexes et coupe des faces qui ne sont pas dans $({}^1\mathcal{S}^n)_{n \in I_1}$ et $({}^2\mathcal{S}^n)_{n \in I_2}$.

Examinons à présent la généralisation de PLI en dimension quelconque ($m \geq 1$). Un simplexe ${}^{m+1}\mathcal{S}$ en dimension $m + 1$ est l'enveloppe convexe de $m + 2$ points v_0, \dots, v_{m+1} de \mathbb{R}^{m+1} ,

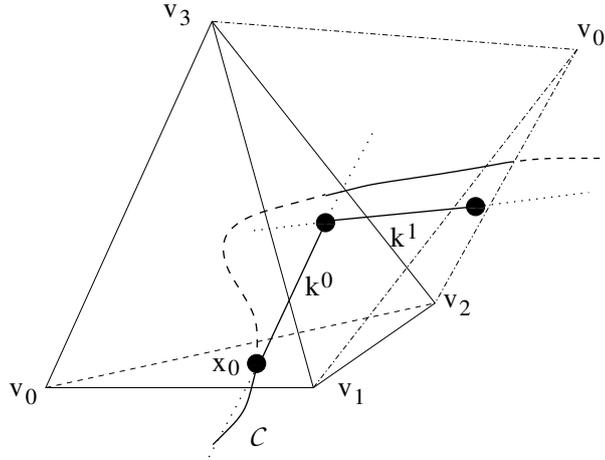


FIGURE 2.8 – La méthode PLI quand $m = 2$.

tels que pour tout $0 \leq i \leq m + 1$, v_i n'appartienne pas à l'hyperplan \mathcal{H}_i passant par les points $v_0, \dots, v_{i-1}, v_{i+1}, \dots, v_{m+1}$. ${}^{m+1}\mathcal{S}$ comporte $m + 2$ faces ${}^m\mathcal{S}_i$ (des simplexes de dimension m), enveloppes convexes des $m + 1$ points $v_0, \dots, v_{i-1}, v_{i+1}, \dots, v_{m+1}$. On appellera ${}^m\mathcal{S}_i$ la i -ème face de ${}^{m+1}\mathcal{S}$.

Par construction, $\langle v_0, \dots, v_{m+1} \rangle$ forme un repère affine de \mathbb{R}^{m+1} (c.à.d la famille composée des vecteurs $(v_i - v_0)$ forme une base de \mathbb{R}^{m+1}) et l'évaluation de H en chacun des points de ce repère détermine une unique application affine, dont la matrice est :

$$\left(\begin{array}{cccc} (H(v_1 - v_0)) & (H(v_2 - v_0)) & \dots & (H(v_{m+1} - v_0)) \end{array} \right).$$

Supposons à présent que x_0 (vérifiant $H(x_0) = 0$) appartienne à la face i de ${}^{m+1}\mathcal{S}^0 = \langle v_0, \dots, v_{m+1} \rangle$. Dans ce simplexe, \mathcal{C} est assimilée au noyau de l'approximation affine H^0 de H , définie comme ci-dessus. Le noyau de H^0 est au moins de dimension 1 d'après le théorème du rang. Si il est de dimension 1 c'est une droite et on considère le segment $[a, b]$ défini comme l'intersection de $\text{Ker}(H^0)$ avec ${}^{m+1}\mathcal{S}^0$. Comme $a \in {}^m\mathcal{S}_i^0$, il existe $j \neq i$ tel que $b \in {}^m\mathcal{S}_j^0$, et si j est unique (c'est à dire que b n'appartient pas à un simplexe de dimension inférieure à m), ${}^m\mathcal{S}_j^0$ est la face du simplexe par laquelle sort l'approximation de la courbe. Comme dans le cas bi-dimensionnel, cette face appartient à un unique simplexe ${}^{m+1}\mathcal{S}^1$ différent de ${}^{m+1}\mathcal{S}^0$. Une nouvelle approximation linéaire de H sur ${}^{m+1}\mathcal{S}^1$ est calculée et, en itérant le processus, deux ensembles $({}^m\mathcal{S}_n)_{n \in I_1}$ et $({}^{m+1}\mathcal{S}^n)_{n \in I_2}$ sont obtenus ; ils approximent \mathcal{C} .

La figure 2.8 présente deux itérations de cette méthode pour suivre une courbe de \mathbb{R}^3 . La courbe entre dans le simplexe $\langle v_0, v_1, v_2, v_3 \rangle$ par sa face $\langle v_0, v_1, v_2 \rangle$, et le noyau k^0 de l'application affine approximant H dans ce simplexe est calculé. Ce noyau intersecte la face $\langle v_1, v_2, v_3 \rangle$ du simplexe. Dans le nouveau simplexe $\langle v_0', v_1, v_2, v_3 \rangle$ partageant sa face $\langle v_1, v_2, v_3 \rangle$ avec le simplexe initial, un nouveau noyau k^1 est calculé et ainsi de suite.

En pratique, il n'est nul besoin de construire de pavage de tout \mathbb{R}^{m+1} . Étant donné un point initial x_0 de \mathcal{C} , un premier simplexe peut être construit tel que x_0 appartienne à une de ses faces. [Dobkin 90] propose de réaliser les calculs dans le repère barycentrique induit par les sommets du simplexe. Le calcul de la face du simplexe qu'intersecte le noyau est alors un simple problème d'optimisation, et il permet d'obtenir un nouveau point x_1 d'une nouvelle face. Un nouveau simplexe auquel appartient cette face peut alors être obtenu par symétrie par rapport à cette face. Les deux

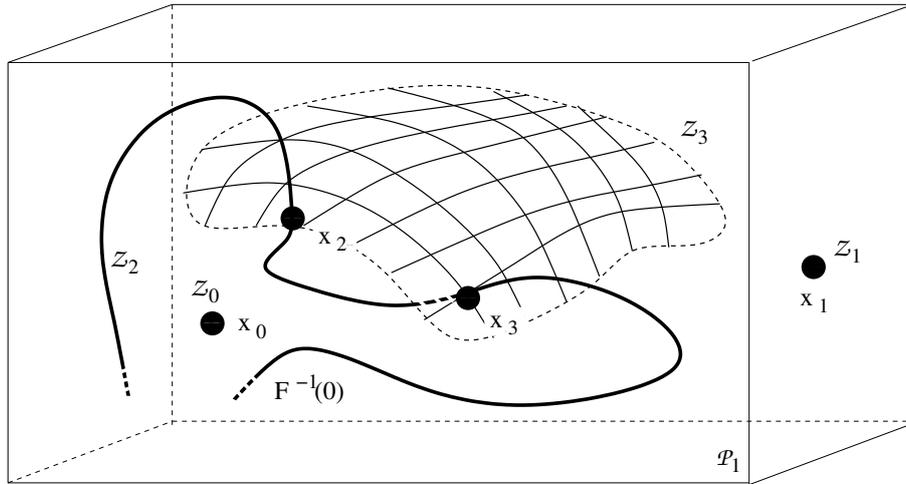


FIGURE 2.9 – Une projection sur \mathbb{R}^3 de l'ensemble $F^{-1}(0)$. Si F est “embarquée” dans une résolution par homotopie, on peut voir cet ensemble comme l'intersection de $H^{-1}(0)$ avec l'hyperplan \mathcal{P}_1 .

simplexes ne diffèrent alors que par un de leurs sommets. Notons que le calcul du noyau dans un simplexe revient à inverser une matrice de taille $m \times m$ ($m + 1 \times m + 1$ si les calculs sont effectués dans un repère barycentrique de $m + 2$ points). Dans le simplexe suivant, la matrice à inverser est la même matrice, à une colonne près. Il est alors bénéfique d'utiliser une méthode d'actualisation de l'inverse de cette matrice, par exemple celle décrite dans [Hager 89].

2.2.6 Numerical algebraic geometry

La “numerical algebraic geometry” a pour objectif de calculer numériquement les ensembles de solutions de dimensions positives d'un système polynomial, c'est-à-dire ses solutions non isolées. Ses moyens principaux sont les méthodes par homotopie présentées plus haut. On rappelle pour commencer quelques notions déjà introduites dans la section 2.1.

On considère une fonction multivariée $F : \mathbb{C}^m \rightarrow \mathbb{C}^n$ dont les composantes sont des polynômes, et on suppose que $m \geq n$. L'ensemble des solutions du système $F = 0$ dans \mathbb{C}^m est appelé ensemble algébrique. Ses composantes connexes sont aussi des ensembles algébriques. En considérant une de ces composantes connexes \mathcal{Z} , on définit l'ensemble $\mathcal{Z}_{reg} \subseteq \mathcal{Z}$ formé de ses points réguliers (pour simplifier, l'ensemble des points de \mathcal{Z} en lesquels la jacobienne de F est de rang n), et on appelle variétés algébriques les composantes connexes de \mathcal{Z}_{reg} . On définit leur dimension en les voyant comme des variétés différentielles. [Elkadi 07] est un ouvrage dédiée à l'étude des solutions des systèmes polynomiaux.

La figure 2.9 représente, pour une fonction F avec $m \geq 3$, des composantes connexes de $F^{-1}(0)$ (les ensembles de solutions de $F = 0$). La composante \mathcal{Z}_0 (respectivement \mathcal{Z}_1) est formée d'un unique point x_0 (resp. x_1) et l'ensemble de ses points réguliers est \mathcal{Z}_0 (resp. \mathcal{Z}_1), qui est de dimension 0. Les ensembles \mathcal{Z}_2 et \mathcal{Z}_3 ont pour intersection $\{x_2, x_3\}$, formé de points critiques de F . L'ensemble des points réguliers de \mathcal{Z}_2 (respectivement \mathcal{Z}_3) est donc $\mathcal{Z}_2 \setminus \{x_2, x_3\}$ (resp. $\mathcal{Z}_3 \setminus \{x_2, x_3\}$) qui est un ensemble de dimension 1 (resp. 2).

Si $m = n$ et que les solutions de F sont cherchées par homotopie grâce à la fonction $H : \mathbb{C}^m \times [0, 1] \rightarrow \mathbb{C}^m$ seules les solutions isolées, c'est à dire les variétés algébriques de dimension 0,

sont trouvées, grâce aux constructions présentées dans 2.2.2. En effet, si un chemin d'homotopie de H aboutissait à une solution d'un ensemble de dimension strictement positive, cette solution serait un point critique de H ; or ces points critiques sont presque sûrement évités.

Le principe de la recherche des composantes de $F^{-1}(0)$ de dimension strictement positive par homotopie est introduit dans [Sommese 96]. Notons que si F est définie sur \mathbb{C}^m et n'est pas identiquement nulle, les variétés algébriques définies par $F = 0$ sont de dimensions au plus $m - 1$. Par ailleurs, l'intersection d'une variété de dimension $0 \leq k < m$ avec presque tout hyperplan de dimension $m - k$ consiste en $d > 0$ points isolés (et d est en général appelé le degré de la variété). Les variétés de dimension $m - i$ définies par $F = 0$ peuvent alors être cherchées en ajoutant à F $m - i$ équations linéaires définissant un hyperplan de dimension i , et $m - i$ variables qui s'annulent en les points de ces hyperplans.

[Sommese 00] propose d'embarquer F dans une cascade d'homotopies, permettant de réutiliser les solutions trouvées lors de la recherche de variétés de dimension $m - i$ par la méthode précédente pour chercher les variétés de dimension $m - i - 1$. Un test probabiliste y est également présenté pour tester si une solution appartient à une variété.

Finalement, [Sommese 07] utilise les méthodes présentées ici pour résoudre une système polynomial équation par équation.

2.2.7 Implantations

Terminons cet état de l'art des méthodes par continuation en présentant quelques logiciels libres qui en implantent certaines.

HOM4PS-2.0 ([Lee 08]) qui est l'abréviation de HOMotopy FOR Polynomial Systems, résout des systèmes d'équations polynomiales par homotopie, en réduisant le nombre de chemins à suivre grâce à l'homotopie polyédrale (voir [Huber 95]). Il est développé et maintenu en FORTRAN-90 à l'université du Michigan (USA). L'utilisateur peut également choisir d'utiliser comme borne maximale du nombre de solutions la borne de Bézout. Il est disponible uniquement pour les machines 64 bits, sur les systèmes Linux, MacOs et Windows XP, à l'adresse : http://www.math.nsysu.edu.tw/~leetsung/works/HOM4PS_soft.htm.

Une interface pour Matlab est également disponible. La version HOM4PS-3.0, actuellement en développement, devrait proposer entre autres de résoudre le système par la "cheater's homotopy" (voir [Li 89]), et de gérer les ensembles de solutions de dimension strictement positive.

Bertini ([Bates 06]) est développé et maintenu à l'université de Notre-Dame (USA). Il propose une méthode générale de résolution par homotopie pour des systèmes polynomiaux. Il implémente la détection des solutions à l'infini, et réduit le nombre de chemins à suivre en utilisant une multi-homogénéisation du système (voir [Morgan 86b],[Morgan 87]). Il adapte la précision du suivi du chemin grâce à [Bates 08] (il permet aussi à l'utilisateur de suivre un chemin avec la précision qu'il souhaite). Il gère également les systèmes qui admettent des ensembles de solutions de dimension positive en implémentant les méthodes présentées dans la sous-section 2.2.6. Bertini doit son nom à un théorème de géométrie algébrique du mathématicien éponyme. Il est disponible pour les machines 32 et 64 bits, pour les systèmes Linux, MacOs et Windows XP, à l'adresse : <http://www3.nd.edu/~sommese/bertini/>.

PHCpack ([Vershelde 99]) comme Polynomial Homotopy Continuation, implémente, en plus des fonctionnalités de Bertini, l'homotopie polyédrale pour réduire le nombre de chemins à suivre. Il est développé en langage Ada. Les sources et des versions exécutables pour la plupart des systèmes/architecture peuvent être obtenus à l'adresse : <http://homepages.math.uic.edu/~jan/PHCpack/phcpack.html>.

Chapitre 3

Résolution par homotopie

Un Système de Contraintes Géométriques (SCG) est défini et résolu dans un univers géométrique donné (2D ou 3D), selon une interprétation des variables et des contraintes. En associant des coordonnées cartésiennes aux objets géométriques, des constantes réelles aux paramètres et des équations aux contraintes, on associe au système de contraintes un système d'équations dont les inconnues sont les coordonnées des objets géométriques à déterminer. Quand le SCG est bien contraint, le système comporte autant d'équations que d'inconnues, et admet, pour des valeurs génériques des paramètres, un nombre fini de solutions complexes. Celles qui sont réelles fournissent les coordonnées des objets géométriques des figures qui sont solutions du problème. Leur nombre est en général exponentiel en le nombre des inconnues.

Le système d'équations peut être résolu par une méthode de continuation classique : un système initial dont les solutions sont connues est déformé continûment en le système à résoudre. Les chemins d'homotopie de longueur infinie sont détectés et les autres sont suivis jusqu'à aboutir à une solution, complexe ou réelle, du système cible. Elles sont toutes obtenues quand le système initial en admet au moins autant. Un nombre exponentiel de chemins est donc suivi, menant le plus souvent à des solutions complexes, ce qui rend cette méthode très coûteuse pour des problèmes impliquant beaucoup de contraintes et d'objets géométriques.

Dans le cadre de la Conception Assistée par Ordinateurs (CAO), le problème a été défini par l'intermédiaire d'une esquisse : c'est une figure composée des objets géométriques du système, respectant en général ses contraintes d'incidence (colinéarités, coplanarités,...). Les contraintes métriques (distances, angles,...) sont considérées comme satisfaites pour d'autres valeurs des paramètres que celles demandées et ces valeurs peuvent être mesurées sur la figure. Une esquisse permet donc de définir un système d'équations utilisé comme système initial dans un contexte de résolution par continuation. Une telle méthode, opérant dans un espace réel, déforme l'esquisse en une solution du système à résoudre, en interpolant les valeurs des paramètres lues sur l'esquisse en les valeurs pour lesquelles les solutions sont cherchées. Une seule solution est trouvée, car un seul chemin est suivi. Elle est cependant proche de l'esquisse (et de la solution attendue) dans le sens où elle résulte d'une déformation continue de cette dernière.

On présente dans ce chapitre un cadre général de résolution de SCG par homotopie utilisant un système initial défini par l'esquisse. La fonction d'homotopie réalise l'interpolation du système initial en le système cible dans l'espace des paramètres, et on considère le chemin auquel appartient l'esquisse dans son ensemble : plusieurs solutions sont obtenues. La section 3.1 pose un cadre formel et définit des notations utilisées par la suite.

La section 3.2 montre sur un exemple comment une méthode par homotopie est appliquée à un

SCG, puis effectue un survol de notre méthode de résolution pour mettre en avant les problèmes rencontrés. En effectuant la continuation dans l'espace réel des paramètres du SCG, le succès de la phase numérique de suivi d'un chemin d'homotopie depuis l'esquisse n'est plus assuré : les chemins de longueur infinie ne peuvent être détectés à l'avance, et des chemins peuvent se croiser ou croiser des ensembles de solutions de dimension supérieure à 1, compromettant ainsi la terminaison et la correction du suivi du chemin. Les chapitres qui suivent aborderont ces problèmes d'un point de vue géométrique.

La topologie des chemins d'homotopie suivis est caractérisée dans la section 3.3, et la section 3.4 présente quelques résultats quantitatifs pour conclure.

3.1 Définitions et notations

On formalise quelques notions et on introduit quelques notations utiles par la suite en les illustrant avec le problème suivant :

Problème 5 *Construire dans l'espace un octaèdre (8 faces, 12 arêtes, 6 points) en connaissant les longueurs de ses douze arêtes.*

3.1.1 Systèmes de contraintes géométriques

La signature suivante :

Sortes : *point, longueur*
Symboles : *distance(point, point) → longueur,*

permet d'écrire de façon formelle les contraintes du problème 5. En notant $X = \{P_0, \dots, P_5\}$ où les P_i sont des inconnues de sorte *point*, $A = \{h_0, \dots, h_{11}\}$ où les h_i sont des paramètres de sorte *longueur*, chaque contrainte c_i s'écrit $distance(P_i, P_j) = h_k$. Pour expliciter la dépendance de c_i en les inconnues P_i, P_j et en le paramètre h_k , on notera $c_i[\{P_i, P_j\}, \{h_k\}]$ cette contrainte.

Définition 1 *Un système de contraintes géométriques G sur une signature consiste en la donnée d'un ensemble de variables sortées $X \cup A$ se décomposant en un ensemble d'inconnues X et un ensemble de paramètres A , et d'un ensemble C de contraintes $c_i[X_i, A_i]$ où $X_i \subset X$ et $A_i \subset A$. On le notera $G = C[X, A]$.*

Le SCG correspondant au problème 5 dans la signature ci-dessus est donné à la figure 3.1. Dans le cadre de la résolution numérique invoqué ici, on utilise l'interprétation numérique, qui fait correspondre à chaque sorte un ensemble de coordonnées cartésiennes réelles permettant de la décrire, et à chaque symbole une fonction numérique.

En 3D, la sorte *point* est donc interprétée par un triplet de réels, et la sorte *longueur* par un réel. Le symbole *distance* est interprété par la norme euclidienne du vecteur reliant les interprétations des objets de type *point* impliqués. Pour une variable sortée x , on appellera *son espace d'interprétation* l'espace dans lequel son interprétation prend ses valeurs. Par exemple, l'espace d'interprétation d'une variable de sorte *point* en 3D est \mathbb{R}^3 .

Un univers géométrique est la donnée d'une signature, fournissant une syntaxe, et d'une interprétation qui pose un cadre. Tous les problèmes donnés par la suite le sont dans l'univers géométrique 3D donné dans l'annexe B.

Inconnues :
point P_0, \dots, P_5
Paramètres :
longueur h_0, \dots, h_{11}
Contraintes :
distance(P_0, P_1) = h_0
...
distance(P_3, P_5) = h_{11}

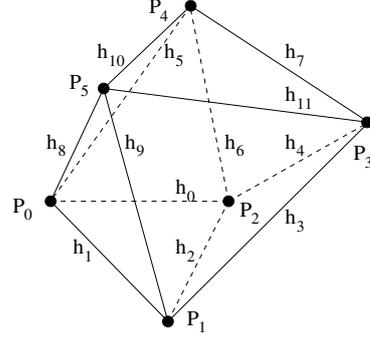


FIGURE 3.1 – À gauche, un énoncé formel du problème 5. À droite, une esquisse cotée : les arêtes représentent des contraintes de distance de paramètres h_i .

Définition 2 Pour un ensemble X de variables dans un univers géométrique donné, une valuation est une fonction σ définie sur l'ensemble $\mathcal{P}(X)$ des parties de X , qui à un ensemble $\{x_1, \dots, x_n\}$ associe un élément $y \in \mathbb{K}_1 \times \dots \times \mathbb{K}_n$, où \mathbb{K}_i est l'espace d'interprétation de x_i . Si $x \in X$, on notera $\sigma(x)$ pour $\sigma(\{x\})$.

Pour une valuation $\sigma_*(A)$ des paramètres et une valuation $\sigma(X)$, la contrainte $distance(P_0, P_1) = h_0$ du SCG associé au problème 5 est écrite :

$$f_d(\sigma(P_0), \sigma(P_1)) = \sigma_*(h_0), \text{ où } f_d(\sigma(P_0), \sigma(P_1)) = \|\overrightarrow{\sigma(P_0)\sigma(P_1)}\|_2.$$

Ici, f_d est la fonction numérique associée au symbole distance. Pour alléger les notations quand il n'y a pas d'ambiguïté sur la valuation de X considérée, on notera $\overrightarrow{P_0P_1}$ le vecteur reliant $\sigma(P_0)$ à $\sigma(P_1)$. On dira que $\sigma(X)$ satisfait cette contrainte pour la valuation $\sigma_*(A)$ quand l'égalité est vérifiée.

Une valuation $\sigma(X)$ d'un univers 3D, dont l'interprétation est numérique, peut être dessinée dans \mathbb{R}^3 ; on appellera par la suite *figure* une valuation pour insister sur son caractère géométrique, ou valuation une figure pour insister sur la valeur des coordonnées attribuée à chaque objet de X . L'esquisse fournie par un utilisateur pour un problème est notamment une valuation de X , que l'on notera en général $\sigma_{sk}(X)$. Pour le problème 5, on donne une esquisse dans la figure 3.1.

3.1.2 Solutions

Étant donné une interprétation et une valuation des paramètres $\sigma_{so}(A)$, une valuation $\sigma(X)$ est une *solution* de $G = C[X, A]$ si $\sigma(X)$ satisfait toutes les interprétations des contraintes c_i de G pour $\sigma_{so}(A)$.

L'esquisse $\sigma_{sk}(X)$ donné dans la figure 3.1 est une solution du SCG $G = C[X, A]$ associée au problème 5 pour une valuation particulière $\sigma_{sk}(A)$ des paramètres, qui peut être lue sur la figure. Par exemple, la longueur $\sigma_{sk}(h_0)$ peut être mesurée sur l'esquisse en interprétant le terme $distance(P_0, P_1)$, c'est à dire en calculant $\|\overrightarrow{P_0P_1}\|_2$. Comme cette notion interviendra souvent, on pose une définition pour la formaliser.

Définition 3 Soit $G = C[X, A]$ un SCG tel que chaque paramètre h_i intervienne dans une unique contrainte métrique $c_i[\{x_1, \dots, x_n\}, h_i]$, et $\sigma(X)$ une valuation de X . On appelle valuation de A lue sur $\sigma(X)$ la valuation σ_* telle que pour tout $h_i \in A$, on ait $\sigma_*(h_i) = f_i(\sigma(x_1), \dots, \sigma(x_n))$ où f_i est l'interprétation du symbole de $c_i[\{x_1, \dots, x_n\}, h_i]$.

3.1.3 Fonction numérique

À chaque coordonnée de chaque objet de l'ensemble X d'un SCG $G = C[X, A]$, on peut associer une variable réelle χ . Si un repère a été choisi pour chercher les solutions de G modulo les transformations rigides, certaines de ces variables sont fixées. On dira que les coordonnées non fixées, et les variables associées, sont *libres*¹.

Dans le cas du problème 5, on munit l'espace du repère $(O, \vec{i}, \vec{j}, \vec{k})$ et on cherche des solutions $\sigma(X)$ telles que $\sigma(P_0) = O$, $\sigma(P_1)$ est sur la droite passant par O et dirigée par \vec{i} , et $\sigma(P_2)$ dans le plan passant par O de vecteur normal \vec{k} , ce qui revient à fixer les trois coordonnées de P_0 , deux coordonnées de P_1 et une coordonnée de P_2 . Par la suite, on dira qu'on cherche les solutions dans le repère $P_0P_1P_2$. On associe donc à X 18 variables, dont 12, que l'on notera χ_0, \dots, χ_{11} , sont libres.

Pour une valuation $\sigma_*(A)$ des paramètres, on peut définir une fonction F des variables libres de X , dont chaque composante $F_i(\chi_0, \dots, \chi_{m-1})$ correspond à l'interprétation du symbole d'une contrainte c_i de G , et s'annule si et seulement si c_i est vérifiée pour $\sigma(X)$ quand chaque coordonnée libre prend par σ la valeur de la variable χ_i qui lui est associée. Pour alléger les notations, on assimilera X à l'ensemble de ses coordonnées libres, et $\sigma(X)$ à l'attribution d'une valeur pour chacune des variables libres χ_i . $\sigma(X)$ est alors vu comme un vecteur de \mathbb{R}^m . On construit donc la fonction :

$$F : \mathbb{R}^m \rightarrow \mathbb{R}^n$$

$$\sigma(X) \mapsto (F_0(\sigma(X)), \dots, F_n(\sigma(X)))^t$$

où n est le nombre de contraintes de G .

Pour une contrainte métrique $c_i[\{x_1, \dots, x_n\}, h_i]$, la composante F_i de F correspondant est $F_i(\sigma(X)) = f_i(\sigma(x_1), \dots, \sigma(x_n)) - \sigma_*(h_i)$, où f_i est l'interprétation du symbole de c_i . Pour une contrainte $c_i[\{x_1, \dots, x_n\}, \emptyset]$, la composante correspondant est $F_i(\sigma(X)) = f_i(\sigma(x_1), \dots, \sigma(x_n))$.

On considère ici que F dépend aussi de la valeur des paramètres des contraintes métriques.

Définition 4 Soit $G = C[X, A]$ un SCG, dont X contient m coordonnées libres, A contient m' paramètres et $C = \{c_1, \dots, c_n\}$ n contraintes. On appellera fonction numérique associée à G la fonction $F : \mathbb{R}^m \times \mathbb{R}^{m'} \rightarrow \mathbb{R}^n$ dont la i -ème composante s'écrit $F_i(\sigma(X), \sigma_*(A)) = f_i(\sigma(x_1), \dots, \sigma(x_n)) - \sigma_*(h_i)$ si $c_i[\{x_1, \dots, x_n\}, h_i]$ est métrique, $F_i(\sigma(X), \sigma_*(A)) = f_i(\sigma(x_1), \dots, \sigma(x_n))$ si $c_i[\{x_1, \dots, x_n\}, \emptyset]$ est booléenne, et où f_i est l'interprétation du symbole de c_i .

On désignera par $F(\sigma(X), \sigma(A))$ une valeur prise par F pour des valuations précises, mais on écrira $F(X, A)$ pour marquer la dépendance de F des variables libres de X et des paramètres de A . De même, on écrira $F_i(X, A) = f_i(x_1, \dots, x_n) - h_i$ pour s'abstraire des valuations, mais $F_i(\sigma(X), \sigma(A))$ la valeur prise par F_i pour les valuations $\sigma(X)$ et $\sigma(A)$. On écrira aussi $F(X, \sigma(A))$ la fonction numérique spécialisée par les valeurs des paramètres $\sigma(A)$.

Pour le problème 5, on a $m' = 12$, et la fonction numérique associée est $F : \mathbb{R}^{12} \times \mathbb{R}^{12} \rightarrow \mathbb{R}^{12}$. Sa composante $F_0(X, A)$, par exemple, s'écrit $\|\vec{P_0P_1}\|_2 - h_0$.

1. En géométrie dynamique, on appelle variable libre un paramètre. Ici il s'agit simplement d'une coordonnée non fixée par le choix d'un repère.

3.2 Résolution d'un SCG par homotopie

Le but de cette section est de présenter un cadre de résolution de SCG par une méthode homotopique dans lequel s'articulent nos contributions. On commence par le résumer en quelques phrases. On explicite ensuite la structure du reste de cette section.

L'esquisse est utilisée pour définir un système initial dont elle est solution. Le chemin d'homotopie partant de l'esquisse est suivi; pendant ce suivi, l'esquisse est transformée continûment en une solution. Ce processus permet de n'obtenir qu'une solution si le suivi est stoppé dès que l'hyperplan \mathcal{P}_1 est rencontré. En revanche, on peut considérer que le chemin d'homotopie est défini au delà de cet hyperplan et si le chemin le recoupe, d'autres solutions sont trouvées. Cette approche est appelée *global homotopy* dans [Allgower 97]; elle nécessite de caractériser la topologie des chemins d'homotopie (ce qui sera fait dans la section 3.3) pour d'une part assurer que ce chemin soit une variété de dimension 1 et garantir la réussite de la phase de suivi, et d'autre part déterminer une condition d'arrêt au processus de suivi. Quand un chemin d'homotopie est difféomorphe à un cercle, stopper le suivi quand le chemin "repassse" par l'esquisse permet d'assurer que toutes les solutions du chemin (*c.à.d* toutes ses intersections avec l'hyperplan \mathcal{P}_1) sont trouvées (si la phase numérique de suivi n'a pas "raté" de portions du chemin). Les solutions trouvées après le suivi du premier chemin sont toutes des déformation continues de l'esquisse. Dans un second temps, on génère de nouvelles esquisses à partir des solutions trouvées pour explorer de nouveaux chemins et ainsi découvrir de nouvelles solutions.

Ce cadre de résolution s'inscrit entre deux approches de résolution de contraintes par homotopie. Rappelons avant de les résumer qu'appliquer naïvement une méthode par continuation à un SCG issu de la CAO est très coûteux, car un nombre de chemins exponentiel en le nombre de ses inconnues est suivi et la plupart aboutissent à des solutions complexes. Connaissant une valuation $\sigma_{so}(A)$ des paramètres d'un SCG, [Durand 00] modifie la fonction numérique $F(X, \sigma_{so}(A))$ lors d'une phase systématique de simplification par calcul symbolique en une fonction $F' : \mathbb{R}^l \rightarrow \mathbb{R}^l$. Le degré du système $F'(\cdot) = 0$ est alors plus petit que celui de $F(X, \sigma_{so}(A)) = 0$, et moins de chemins d'homotopie doivent être suivis. Cette approche permet d'obtenir toutes les solutions du SCG. Dans [Lamure 95], l'esquisse $\sigma_{sk}(X)$ est utilisée pour définir la fonction d'homotopie qui interpole les valeurs $\sigma_{sk}(A)$ lues sur l'esquisse en les valeurs $\sigma_{so}(A)$ pour lesquelles on cherche des solutions (cette construction est justifiée dans [Morgan 89]). Le chemin passant par l'esquisse est suivi jusqu'à ce qu'il croise l'hyperplan \mathcal{P}_1 . L'unique solution trouvée est proche de l'esquisse.

Dans la sous-section 3.2.1, on montre comment résoudre le SCG associé au problème 5 (l'octaèdre) par homotopie d'abord de manière naïve, puis avec les méthodes de [Durand 00] et de [Lamure 95] et enfin avec l'approche présentée dans ce mémoire.

La sous-section 3.2.2 décrit en détail le cadre de résolution dont il est question dans cette section en exhibant les problèmes auxquels il faut faire face pour appliquer une homotopie agissant dans l'espace des figures, comme c'est le cas ici. Ces problèmes seront abordés dans les chapitres suivants. Notons que la section 3.4 présente quelques résultats numériques obtenus par notre méthode.

3.2.1 Un exemple : l'octaèdre

On considère le problème 5, dont le SCG associé $G = C[X, A]$ et une esquisse cotée sont donnés dans la figure 3.1. L'espace est muni du repère $(O, \vec{i}, \vec{j}, \vec{k})$, et les solutions sont cherchées modulo les déplacements dans le repère $P_0P_1P_2$ comme décrit dans 3.1.3. Pour la valuation $\sigma_{so}(A)$ des paramètres donnée dans la table 3.1, G admet 32 solutions modulo les déplacements. À partir de chacune de ces solutions, on peut en obtenir 7 autres par compositions des symétries par rapport aux plans de vecteurs normaux \vec{i} , \vec{j} , et \vec{k} . On regroupe ces 8 solutions dans une orbite

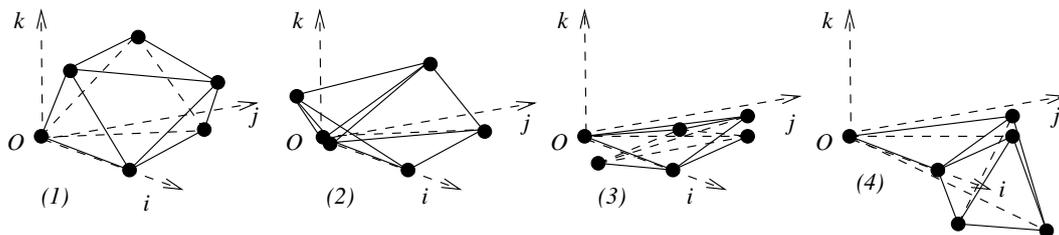


FIGURE 3.2 – Les 4 solutions, modulo les déplacements et les symétries, du problème 5.

sous l'action des compositions de ces symétries. L'ensemble des 32 solutions de G peut alors être partitionné en 4 classes d'équivalence. La figure 3.2 présente un élément de chacune de ces classes.

| | | | |
|-----------------|-----------------|--------------------|--------------------|
| $h_0 = 1.11983$ | $h_1 = 1.17389$ | $h_2 = 1.18117$ | $h_3 = 1.06129$ |
| $h_4 = 1.12482$ | $h_5 = 1.18592$ | $h_6 = 1.00796$ | $h_7 = 1.15857$ |
| $h_8 = 1.17417$ | $h_9 = 1.07569$ | $h_{10} = 1.19071$ | $h_{11} = 1.16643$ |

TABLE 3.1 – Une valuation $\sigma_{so}(A)$ des paramètres du problème 5.

Application naïve d'une méthode classique

Si F est la fonction numérique associée à G , on construit une fonction polynomiale Fp : pour chaque composante $F_i(X, A) = \|\overrightarrow{P_j P_k}\|_2 - h_i$ de F , on définit la i -ème composante de Fp par $Fp_i(X, A) = (\|\overrightarrow{P_j P_k}\|_2)^2 - h_i^2$. Le système $Fp(X, \sigma_{so}(A)) = 0$ formé des 12 équations $Fp_i(X, \sigma_{so}(A)) = 0$ a alors un degré de $2^{12} = 4096$, et le résoudre par une méthode d'homotopie classique, par exemple celle décrite dans [Morgan 86b] conduit à suivre 4096 chemins. Notons qu'utiliser une homotopie polyédrale ne réduit pas ce nombre de chemins. Le logiciel libre HOM4PS-2.0 effectue cette résolution en environ 126 secondes sur un Intel(R) Core(TM) i5 CPU 750 @ 2.67GHz. 64 solutions sont trouvées, parmi lesquelles 32 sont réelles et correspondent aux 32 solutions modulo les déplacements.

Simplification symbolique pour réduire le nombre de chemins à suivre

La phase de simplification de $Fp(X, A) = 0$ décrite dans [Durand 00] revient, pour simplifier, à résoudre symboliquement les équations impliquant les points P_1 et P_2 du repère en choisissant pour chacun un demi-plan où ils se trouvent, et à réduire le nombre d'inconnues et d'équations en remplaçant les coordonnées (x, y, z) d'un point contraint par une distance, donc sur une sphère, par un couple d'angles (ϕ, ψ) . Le résultat est un système $F'(\cdot) = 0$ de trois équations de degré 4 en trois inconnues ; le choix du placement de P_1 et P_2 disqualifie les trois quarts des solutions du système original, qui peuvent être obtenues par symétries par rapport aux plans de vecteurs normaux \vec{i} et \vec{j} . Ce système, qui a pour degré $4^3 = 64$ est résolu par HOM4PS-2.0 en environ 0.05 secondes avec une méthode d'homotopie classique, et en moins de 0.01 secondes avec l'homotopie polyédrale. Dans ce dernier cas, seulement 16 chemins sont suivis. 8 solutions réelles sont obtenues, qu'on peut répartir en deux orbites sous l'action des symétries par rapport au plan dirigé par \vec{k} .

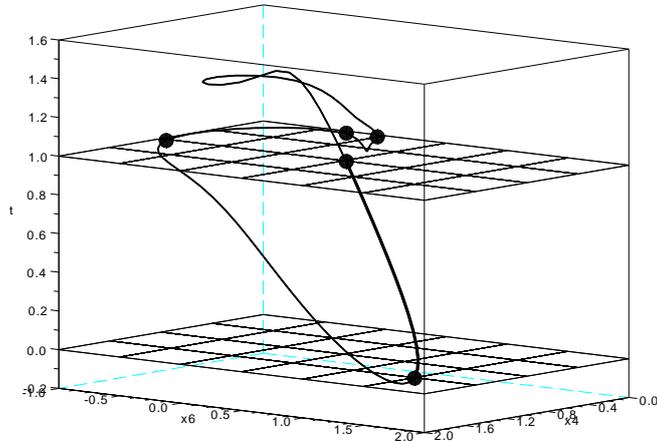


FIGURE 3.3 – En gras, le chemin d’homotopie suivi par [Lamure 95] à partir de l’esquisse, pour le problème 5. En trait mince, ce chemin en entier. Il croise 4 fois l’hyperplan \mathcal{P}_1 , en les solutions données à la figure 3.2.

Utilisation de l’esquisse pour construire un système initial

En utilisant l’esquisse donnée dans la partie droite de la figure 3.1 et les valeurs $\sigma_{sk}(A)$ des paramètres lues sur cette esquisse pour poser le système initial comme dans [Lamure 95], les 12 composantes de la fonction d’homotopie $H : \mathbb{R}^{12} \times [0, 1] \rightarrow \mathbb{R}^{12}$ sont $H_i(X, t) = \|\overrightarrow{P_j P_k}\|_2 - (1 - t)\sigma_{sk}(h_i) - t\sigma_{so}(h_i)$. Le chemin suivi depuis l’esquisse est représenté en gras sur la figure 3.3, la solution obtenue à son intersection avec \mathcal{P}_1 est celle de la figure 3.2(1). Ce chemin est exploré en environ 0.03 secondes avec la méthode de prédiction-corrrection décrite dans la sous-section 3.2.2.

Suivi du chemin d’homotopie défini par l’esquisse dans son ensemble

En considérant que H est définie sur $\mathbb{R}^{12} \times \mathbb{R}$ et non sur $\mathbb{R}^{12} \times [0, 1]$, le chemin d’homotopie de H est une courbe dans $\mathbb{R}^{12} \times \mathbb{R}$, et non plus dans $\mathbb{R}^{12} \times [0, 1]$. En l’explorant en entier, on obtient le chemin représenté sur la figure 3.3, qui intersecte 4 fois \mathcal{P}_1 , en les solutions (1), (2), (3) et (4) de la figure 3.2, et revient sur lui-même ; ce chemin est suivi dans son ensemble (jusqu’à ce qu’il passe à nouveau par l’esquisse) en environ 0.1 secondes par la même méthode de prédiction-corrrection.

3.2.2 Guider la résolution par homotopie avec la géométrie

On présente le cadre de résolution par homotopie dans lequel s’articulent nos contributions, en faisant apparaître les problèmes qu’il soulève.

On considère un SCG $G = C[X, A]$ que l’on suppose bien contraint, avec m contraintes, dont m' sont métriques, dont les inconnues sont déterminées par m coordonnées libres (après le choix d’un repère). On lui associe la fonction numérique $F(X, A)$ des coordonnées libres et des paramètres définie comme dans la section 3.1.

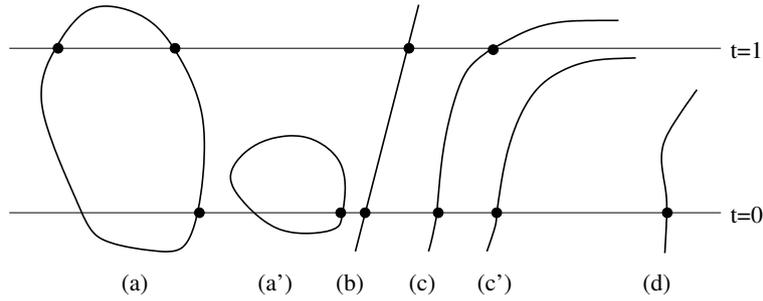


FIGURE 3.4 – Différents types de chemins d'homotopie. Les points noirs sur l'hyperplan $t = 0$ correspondent à des esquisses, ceux sur l'hyperplan $t = 1$ les solutions appartenant aux chemins.

Construire la fonction d'homotopie

Pour une esquisse $\sigma_{sk}(X)$ de G , $\sigma_{sk}(A)$ la valuation des paramètres lue sur l'esquisse, et une valuation $\sigma_{so}(A)$ pour laquelle on cherche les solutions de G , on considère la fonction d'homotopie associée à G :

$$H : \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m \quad (3.1)$$

$$(X, t) \mapsto F(X, d(t))$$

où $d(t)$ est une fonction C^∞ d'une variable réelle, à valeurs dans $\mathbb{R}^{m'}$, appelée *fonction d'interpolation* des paramètres, telle que $d(0) = \sigma_{sk}(A)$ et $d(1) = \sigma_{so}(A)$.

Caractériser ses chemins sous condition de sa régularité

Sous une condition de régularité de H , la proposition 7 de la section 3.3 permet de caractériser les chemins d'homotopie de H . Ils sont difféomorphes soit à un cercle de \mathbb{R}^2 , soit à un intervalle ouvert de \mathbb{R} . La figure 3.4 présente différentes formes de chemins d'homotopie.

Le cas où un chemin est difféomorphe à un cercle (chemins (a) et (a') sur la figure 3.4) est le plus avantageux, de deux points de vue. D'abord, le chemin peut contenir plusieurs solutions (chemin (a)), ensuite le suivi du chemin peut être stoppé quand un cycle est constaté. Ce critère d'arrêt est testé ici en calculant une norme ; un problème intéressant qui ne sera pas abordé est de proposer un test prouvable et exact pour ce critère, en établissant par exemple l'unicité du chemin suivi dans une boîte contenant le point courant. Quand un chemin est difféomorphe à un intervalle, il peut avoir une longueur infinie, et son suivi peut ne jamais terminer. Il peut être borné en sa dernière composante (celle correspondant au paramètre d'interpolation t), ou non.

- Si non, pour tout $M \in \mathbb{R}$, il contient une solution de $H(X, t) = 0$ pour tout t tel que $|t| > M$ (chemin (b) sur la figure 3.4).
- S'il l'est, il existe $t_\infty \in \mathbb{R}$ tel que le chemin converge vers une solution à l'infini du système $H(X, t_\infty) = 0$ (chemins (c) et (c') sur la figure 3.4).

Un chemin difféomorphe à un intervalle ouvert de \mathbb{R} peut aussi être contenu dans un compact (chemin (d) sur la figure 3.4). Dans ce cas, on montre qu'il converge vers un point en lequel H n'est pas définie, ou pas dérivable. Une telle situation est facilement détectée.

La section 3.3 montre comment construire la fonction d'homotopie en utilisant une fonction d'interpolation d des paramètres non linéaire permettant de modifier la nature des chemins d'homotopie, afin que ceux de longueur infinie soient uniquement des chemins tendant vers une solution à l'infini ; pour parler de manière informelle, les chemins non bornés en t sont transformés en chemins difféomorphes à des cercles. Le chemin d'homotopie auquel appartient l'esquisse peut alors être du

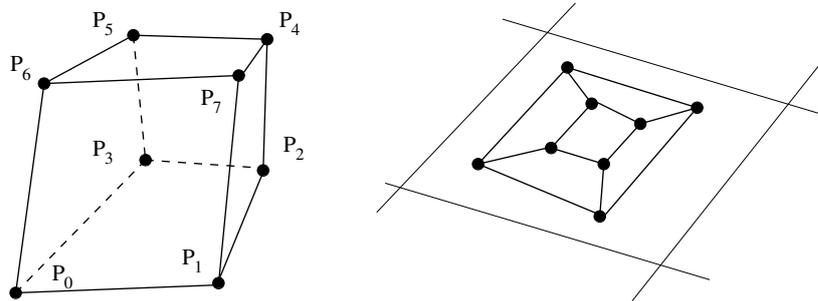


FIGURE 3.5 – Deux esquisses cotées pour les problèmes 3 (à gauche) et 4 (à droite). Les arêtes représentent des contraintes de distance. À gauche, les 4 points de chacune des 8 faces sont impliqués dans une contrainte de coplanarité, et à droite les 8 points sont tous coplanaires.

type d'un des chemins (a), (a'), (c), (c') et (d) de la figure 3.4. Bien sûr, un chemin difféomorphe à un intervalle ouvert peut avoir une de ses extrémités qui converge vers une solution à l'infini, et l'autre qui converge vers un point où H n'est pas dérivable ou pas définie.

On propose dans le chapitre 4 d'utiliser un plan de construction d'un problème "proche" du SCG G à résoudre (on suppose que G résiste à une méthode géométrique permettant d'obtenir un plan de construction de ses solutions) comme observateur pour détecter qu'un chemin tend vers une solution à l'infini.

Gérer les cas où la condition de régularité n'est pas satisfaite

Cette caractérisation est valable quand H satisfait à la condition de régularité suivante : sa matrice jacobienne doit être de rang plein en toutes les solutions de $H(X, t) = 0$. Quand cette condition n'a pas lieu, on dit que 0 n'est pas une valeur régulière de H et on appelle *point critique* un point $(\sigma(X), t)$ en lequel la jacobienne de H n'est pas de rang plein.

Quand 0 n'est pas une valeur régulière de H , les composantes connexes de $H^{-1}(0)$ ont une topologie plus compliquée que celle d'un cercle ou d'une droite. Elles peuvent être composées de deux variétés de dimension 1 s'intersectant, ou d'une variété de dimension 1 croisant une variété de dimension supérieure. Dans ces cas, le suivi de ces ensembles peut soit échouer, soit ne jamais terminer quand l'ensemble présente des embranchements. Dans le cadre de la résolution par des méthodes classiques, la probabilité pour qu'un chemin d'homotopie contienne un point critique est nulle. Avec notre méthode, quand un système de contraintes géométriques implique des contraintes booléennes, comme des alignements ou des coplanarités, ce phénomène se rencontre cependant.

C'est le cas notamment quand on résout le problème 3 (dont une esquisse cotée est donnée dans la figure 3.5), déjà introduit dans le chapitre 1, qui consiste à construire un hexaèdre en connaissant douze distances et six coplanarités entre quatre points. Ce problème, et le système d'équations $F(X, A) = 0$ où F est la fonction numérique associée, est structurellement bien contraint (au sens de König-Hall) et admet des solutions isolées. Pourtant, il n'est pas génériquement bien contraint, car toutes les solutions du problème 4 (dont une esquisse cotée est donnée dans la figure 3.5), qui existent en un nombre infini, satisfont ses contraintes. On peut en fait montrer que si $F'(X, A)$ est la fonction numérique associée au problème 4 et $F'_{\sigma(A)}(X)$ la fonction définie par $F'_{\sigma(A)}(X) = F'(X, \sigma(A))$, les composantes connexes de $(F'_{\sigma(A)})^{-1}(0)$ sont des variétés différentielles de dimension 1, et ce pour presque toutes les valuations $\sigma(A)$.

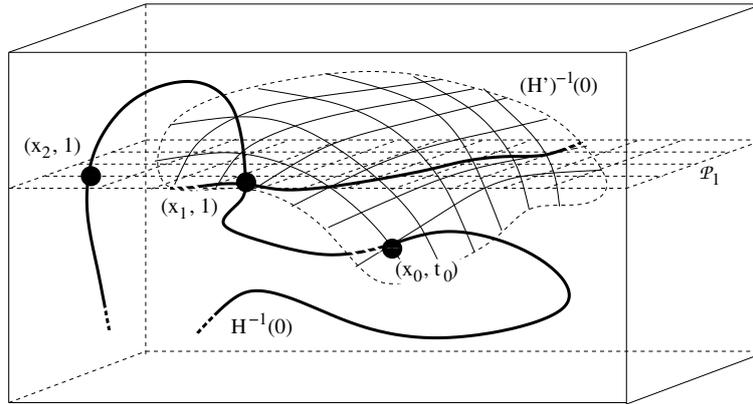


FIGURE 3.6 – Une représentation schématisée d’une projection sur \mathbb{R}^3 d’une partie d’une composante connexe de $H^{-1}(0)$ pour le problème 3 : ici, H ne satisfait pas la condition de régularité.

En appelant H (respectivement H') la fonction d’homotopie associée au problème 3 (resp. au problème 4), les composantes connexes de $(H')^{-1}(0)$ sont des variétés de dimension 2, et il est clair que $(H')^{-1}(0) \subset (H)^{-1}(0)$. On dira que les solutions du problème 4 et, d’une manière générale, les figures dont tous les points sont coplanaires, sont *aplaties*. En suivant la composante connexe de $(H)^{-1}(0)$ à laquelle appartient l’esquisse (non aplatie) donnée dans la figure 3.5, si une figure de cette composante connexe est aplatie, elle appartient à une composante connexe de $(H')^{-1}(0)$ qui est une variété de dimension 2. Cette figure correspond à un point critique, et le suivi du chemin qui la contient est un échec. Comme on le verra, cette situation (le chemin passe par une figure aplatie) n’est pas un hasard dans le sens où l’ensemble des figures aplaties solutions de $H(\cdot) = 0$ sépare l’ensemble des solutions de $H(\cdot) = 0$. Sur la figure 3.6, on a représenté de manière schématisée les ensembles $(H)^{-1}(0)$ et $(H')^{-1}(0)$. Tous les points de ce dernier ensemble correspondent à des figures aplaties. L’intersection de ces deux ensembles avec l’hyperplan \mathcal{P}_1 forme l’ensemble de solutions du problème 3, dont les composantes connexes sont hétérogènes en dimension.

Dans le chapitre 5, on montre comment modifier la fonction d’homotopie de façon à ce que ses chemins ne contiennent d’autres points critiques que ceux qui sont solutions du SCG G à résoudre, pour les valuations $\sigma_{so}(A)$ ou $\sigma_{sk}(A)$, avec une probabilité de 1. Cette construction permet d’utiliser pour un problème des esquisses qui ne satisfont pas ses contraintes booléennes. Les points critiques solutions de G qui peuvent être croisés appartiennent tous à l’hyperplan \mathcal{P}_1 (ou \mathcal{P}_0) ($(x_1, 1)$ et $(x_2, 1)$ sur la figure 3.6), et on propose à la fin de cette sous-section une adaptation de la méthode de prédiction-corrrection pour gérer ces cas.

Ces points critiques trouvés par notre méthode sont des éléments d’un ensemble de solutions de G de dimension strictement positive (par exemple les solutions aplaties du problème 3). Ces phénomènes apparaissent lorsque des théorèmes de géométrie amènent des contraintes à être automatiquement satisfaites, pour certains ensembles de figures. On présente dans le chapitre 5 l’idée d’une méthode pour détecter de telles situations quand elles sont liées à la géométrie d’incidence, et apporter une analyse géométrique au problème. Dans le cas de l’exemple décrit ici, cette méthode détecte qu’une solution aplatie du problème 3 est solution du problème 4 et qu’il est structurellement sous-contraint.

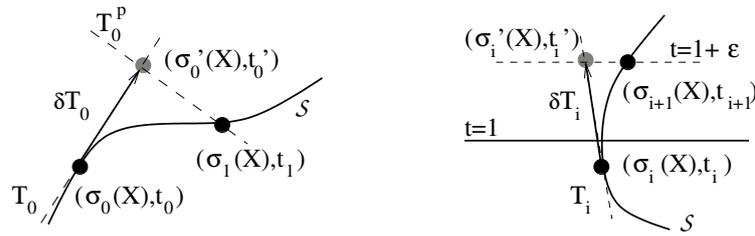


FIGURE 3.7 – Une étape de prédiction-corréction à gauche, et son adaptation dans le cas où un point critique, sur l’hyperplan \mathcal{P}_1 , peut être rencontré, à droite.

Obtenir d’autres solutions

Suivre le chemin d’homotopie de H auquel appartient l’esquisse permet d’obtenir les solutions qui appartiennent à cette composante connexe de $H^{-1}(0)$. Une question importante à laquelle on a tenté de répondre est comment trouver les solutions appartenant aux autres composantes connexes. Une approche qui a été envisagée est d’utiliser un SCG sous-contraint construit à partir du SCG G : en supprimant une contrainte de degré de restriction 1 de G , on obtient un SCG G' et une fonction numérique F' dont l’ensemble des zéros est une variété de dimension 1 ; les solutions de G sont les points de $F'^{-1}(0)$ en lesquels la contrainte qui a été supprimée de G est satisfaite ; elles peuvent donc être trouvées en explorant l’ensemble $F'^{-1}(0)$. De plus, chaque solution trouvée sur le chemin d’homotopie initialement suivi appartient à une composante connexe de cet ensemble. Cependant, sans interpoler les paramètres de G , on ne peut pas transférer les propriétés des composantes connexes de $H^{-1}(0)$ à celles de $F'^{-1}(0)$ et on ne peut pas garantir qu’elles soient difféomorphes à des cercles. Notons que si cette méthode ne permet de n’explorer que les composantes connexes de $F'^{-1}(0)$ dans lesquels on a déjà trouvé une solution, on peut espérer en atteindre d’autres en considérant un SCG sous-contraint G' défini par la suppression d’une autre contrainte. On présente dans ce mémoire une approche utilisant la géométrie constructive pour trouver des solutions appartenant à d’autres composantes connexes.

Quand le SCG G peut être résolu par une méthode géométrique qui fournit un plan de construction, toutes les solutions de G peuvent être obtenues en évaluant ce plan de construction ; nous considérons ici des SCG qui ne peuvent pas être résolus par de telles méthodes. Il est cependant possible de produire un plan de construction d’un problème G' proche de G , partageant ses inconnues et certaines de ses contraintes. Évaluer ce plan de construction en utilisant des valeurs de paramètres lues sur les solutions contenues par le chemin d’homotopie passant par l’esquisse initiale permet de construire de nouvelles esquisses pour G . Ces nouvelles esquisses permettent de définir autant de nouvelles fonctions d’homotopie, et de suivre autant de nouveaux chemins pour obtenir de nouvelles solutions. Ce processus est détaillé dans le chapitre 4. Il peut être itéré tant que de nouvelles solutions sont trouvées sur ces chemins. Intégré dans un logiciel de CAO, il peut aussi permettre de proposer de nouvelles esquisses à un utilisateur ; celui-ci peut alors sélectionner celles qui lui conviennent, afin de trouver des solutions proches de celles-ci.

Suivre les chemins d’homotopie

On utilise une méthode classique de prédiction-corréction pour suivre un chemin \mathcal{S} depuis l’esquisse $(\sigma_{sk}(X), 0)$, qu’on adapte pour gérer l’existence de points critiques dans les hyperplans \mathcal{P}_0 ou \mathcal{P}_1 . \mathcal{S} est définie par $H^{-1}(0)$, où $H : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$. Aucune stratégie d’adaptation du pas de prédiction n’est utilisée.

Soit $(\sigma_0(X), t_0) \in \mathcal{S}$. La matrice jacobienne $J_{(\sigma_0(X), t_0)}$ de H en $(\sigma_0(X), t_0)$ est calculée numériquement grâce à la méthode des différences finies, et la tangente T_0 à \mathcal{S} en $(\sigma_0(X), t_0)$ est calculée en résolvant l'équation linéaire :

$$J_{(\sigma_0(X), t_0)}v = 0, \quad (3.2)$$

où $v \in \mathbb{R}^{m+1}$, et $J_{(\sigma_0(X), t_0)}$ est une matrice de taille $m \times (m+1)$. Si le rang de $J_{(\sigma_0(X), t_0)}$ est m (c.à.d. $(\sigma_0(X), t_0)$ n'est pas un point critique de H), alors en appliquant le théorème du rang, l'ensemble des solutions v de (3.2) est un espace vectoriel de dimension 1. On appelle tangente T_0 à \mathcal{S} en $(\sigma_0(X), t_0)$ une des deux solutions v de (3.2) telle que $\|v\|_2 = 1$ (choisir une de ces deux solutions revient à choisir un sens pour le suivi de \mathcal{S}).

Une prédiction $(\sigma'_0(X), t'_0)$ est alors calculée dans la direction de T_0 , avec le pas δ :

$$(\sigma'_0(X), t'_0) = (\sigma_0(X), t_0) + \delta T_0$$

Cette prédiction est alors corrigée en une solution $(\sigma_1(X), t_1) \in \mathcal{S}$ telle que $(\sigma_1(X), t_1) \in T_0^p$, où T_0^p est l'hyperplan perpendiculaire à T_0 . On définit la fonction $H_{m+1}^0 : \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ qui à v associe $v \cdot T_0$, où \cdot est le produit scalaire usuel, valant 0 quand v est dans l'hyperplan de vecteur normal T_0 . On applique alors la méthode de Newton-Raphson à la fonction

$$\begin{aligned} \mathbb{R}^{m+1} &\rightarrow \mathbb{R}^{m+1} \\ v &\mapsto (H(v), H_{m+1}^0(v))^t \end{aligned} \quad (3.3)$$

avec la solution initiale $(\sigma'_0(X), t'_0)$.

La partie gauche de la figure 3.7 présente une étape de prédiction correction. En itérant ces étapes, on construit une suite $((\sigma_i(X), t_i))_{i \in \mathbb{N}}$ de points de \mathcal{S} . À chaque étape, la tangente T_i est choisie entre les deux vecteurs T_i^1 et T_i^2 de façon à minimiser $\|T_i^j - T_{i-1}\|_2$.

Dans le cas où le SCG résolu contient des contraintes booléennes, si $\mathcal{S} \cap (\mathcal{P}_1)$ contient des points critiques (le cas $\mathcal{S} \cap (\mathcal{P}_0)$ est similaire), une de ces étapes peut échouer aux abords de \mathcal{P}_1 :

1. si $t_i = 1$, la matrice jacobienne n'est pas de rang plein et l'ensemble des solutions de (3.2) n'est pas un espace vectoriel de dimension 1 ; la prédiction n'a plus de sens,
2. si $t'_i = 1$ la matrice jacobienne n'est pas de rang plein et la correction par la méthode de Newton-Raphson échoue,
3. si $(t_i - 1) \cdot (t'_i - 1) < 0$ (c.à.d. t_i et t'_i sont de part et d'autre de \mathcal{P}_1), $t_{i+1} = 1$ et $(\sigma_{i+1}(X), t_{i+1})$ est un point critique, l'étape suivante de prédiction échoue.

On prévient ces cas d'échec en appliquant les règles suivantes :

- (i) si les cas 3 ou 2 se présentent, le pas de prédiction δ est choisi tel que $t'_i = 1 + \epsilon$ si $t'_i < 1$, $t'_i = 1 - \epsilon$ sinon, avec $\epsilon > 0$, et $(\sigma'_i(X), t'_i)$ est corrigé en $(\sigma_{i+1}(X), t_{i+1})$ tel que $t_{i+1} = t'_i$ (la dernière composante de H' dans (3.3) est remplacée par $H_{m+1}^i(X, t) = t - t'_i$);
- (ii) si le cas 1 se présente, $(\sigma_i(X), t_i)$ est remplacé par le dernier point connu sur \mathcal{S} , et la règle (i) est appliquée.

La partie droite de la figure 3.7 illustre la règle (i).

Le processus de prédiction-correction est itéré jusqu'à ce qu'une des situations ci-dessous soit détectée :

- \mathcal{P}_0 est croisé, et la figure courante est similaire à l'esquisse (c.à.d. $\|\sigma_i(X) - \sigma_{sk}(X)\| < \epsilon$, avec $\epsilon > 0$),
- la figure $\sigma_i(X)$ est proche d'une configuration où H n'est pas C^∞ ,
- la figure $\sigma_i(X)$ présente la configuration d'une solution à l'infini.

Dans les deux derniers cas, le chemin est suivi à nouveau depuis l'esquisse, mais dans la direction opposée. La proposition 9 de la section 3.3 justifie que ce processus termine, et que toutes les solutions de G sur le chemin \mathcal{S} sont trouvées.

3.3 Chemins d'homotopie

Dans les méthodes par homotopie classiques, au moins autant de solutions du système initial que le système cible en admet sont connues ; caractériser les chemins d'homotopie permet alors d'assurer la continuation des solutions du premier système (initial) en solutions du second (cible). Certains chemins de cet ensemble ont cependant une longueur infinie, et tendent vers une solution à l'infini du système cible. Poser la fonction d'homotopie dans un espace projectif permet de détecter un tel chemin avant son suivi.

La méthode proposée ici utilise une esquisse du problème à résoudre pour poser le système initial ; de ce fait une seule solution en est connue. Cette esquisse est déformée continûment en des solutions du système à résoudre, et chaque point d'un chemin d'homotopie correspond à une figure, permettant d'y appliquer des raisonnements géométriques.

La fonction d'homotopie est construite de façon à ce que ses chemins d'homotopie soient difféomorphes à des cercles. Un tel chemin coupe un hyperplan de l'espace un nombre pair de fois. Ainsi, l'explorer dans son ensemble peut fournir plusieurs solutions. Les chemins de la fonction d'homotopie qui ne sont pas difféomorphes à un cercle convergent :

1. vers un point en dehors du domaine de définition ou de dérivabilité de la fonction d'homotopie,
2. vers une solution à l'infini de la fonction d'homotopie.

Le cas 1. peut être facilement détecté car le domaine de dérivabilité de la fonction d'homotopie est l'intersection des domaines de dérivabilité de ses composantes (donc des contraintes du SCG considéré). Il sera montré dans le chapitre 4, en utilisant un plan de construction, que les solutions à l'infini de la fonction d'homotopie correspondent à des configurations géométriques particulières, qui peuvent être lues sur les figures d'un chemin convergeant vers une solution à l'infini (cas 2.).

Les notions basiques de topologie différentielle utilisées dans ce chapitre sont rappelées en annexe A.

3.3.1 Fonction d'homotopie

On considère un système de contraintes géométriques $G = C[X, A]$, où $C = \{c_0, \dots, c_{m-1}\}$ est un ensemble de m contraintes, chacune dépendant d'un paramètre $h_i \in A$. Les contraintes de C qui sont booléennes sont considérées ici comme impliquant un paramètre de valeur nulle. À chaque contrainte c_i est associée son interprétation $F_i(X, h_i)$, qui est une fonction numérique à valeurs réelles. On suppose ici que G est bien contraint, et qu'après avoir défini un repère, X contient m variables libres. La fonction numérique $F : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ associée à G est alors définie comme dans la section 3.1.

On suppose qu'on cherche à résoudre G pour la valuation des paramètres $\sigma_{so}(A)$, et qu'on dispose d'une esquisse $\sigma_{sk}(X)$ de G . Sur cette esquisse, on peut lire la valuation $\sigma_{sk}(A)$ telle que $F(\sigma_{sk}(X), \sigma_{sk}(A)) = 0$. (Rappelons qu'on cherche l'ensemble des figures $\sigma(X)$ telles que $F(\sigma(X), \sigma_{so}(A)) = 0$).

Notre homotopie déforme les paramètres de l'esquisse en les paramètres pour lesquels on cherche les solutions. On définit de manière formelle la notion de fonction d'interpolation des paramètres, déjà introduite dans la section 3.2 :

Définition 5 On appelle fonction d'interpolation de $\alpha \in \mathbb{R}$ à $\beta \in \mathbb{R}$ une application $C^\infty d : \mathbb{R} \rightarrow \mathbb{R}$ telle que $d(0) = \alpha$ et $d(1) = \beta$. Soit $A = \{a_0, \dots, a_{p-1}\}$ un ensemble de paramètres. On appelle $d : \mathbb{R} \rightarrow \mathbb{R}^p$ une fonction d'interpolation de $\sigma_1(A)$ à $\sigma_2(A)$ si chacune de ses composantes d_j , pour $0 \leq j \leq p-1$, est une fonction d'interpolation de $\sigma_1(a_j)$ à $\sigma_2(a_j)$.

Inconnues :
point P_0, \dots, P_5
Paramètres :
longueur $h_0, h_1, h_5, \dots, h_9$
angle a_1, a_2, a_3
Contraintes :
distance(P_0, P_1) = h_0
...
distance(P_2, P_0) = h_9
angle(P_2, P_3, P_4) = a_1
angle(P_2, P_3, P_0) = a_2
angle(P_1, P_0, P_3) = a_3
coplanaires(P_0, P_1, P_2, P_3)
coplanaires(P_1, P_2, P_4, P_5)

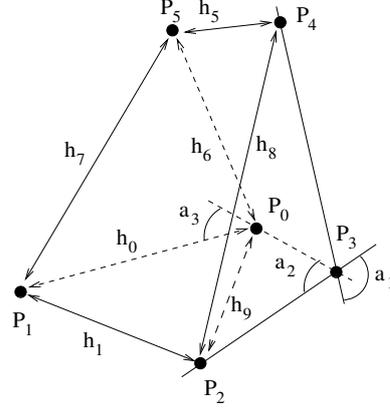


FIGURE 3.8 – L'énoncé formel et une esquisse cotée du problème 6p7d3a2c. Les doubles flèches droites représentent les contraintes de distance de paramètres h_i et les arcs de cercles entre deux droites joignant deux points sont des contraintes d'angle de paramètres a_i .

Si d est une fonction d'interpolation de $\sigma_{sk}(A)$ à $\sigma_{so}(A)$, on définit la fonction d'homotopie H comme dans l'expression (3.1) que l'on rappelle ici pour la commodité de la lecture :

$$H : \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m \quad (3.4)$$

$$(X, t) \mapsto F(X, d(t))$$

et on suppose que H est C^∞ sur une variété $\mathcal{D} \subseteq \mathbb{R}^m \times \mathbb{R}$ de dimension $m + 1$ sans bord.

On introduit ici un nouveau problème afin d'illustrer notre propos, dont les solutions sont des solides à 6 sommets respectant 7 contraintes de distances sur ses arêtes, 3 contraintes d'angles orientés entre trois points, et 2 coplanarités de 4 points. On appellera ce problème 6p7d3a2c. Le SCG G associé à ce problème, ainsi qu'une esquisse cotée, sont donnés dans la figure 3.8. On cherche des solutions $\sigma(X)$ telles que $\sigma(P_0) = (0, 0, 0)$, $\sigma(P_1) = (\chi_0, 0, 0)$, et $\sigma(P_2) = (\chi_1, \chi_2, 0)$ où χ_0, χ_1, χ_2 sont des variables libres. En fixant à 0 la dernière coordonnée de P_3 , la contrainte *coplanaires*(P_0, P_1, P_2, P_3) est automatiquement satisfaite, et la fonction numérique associée à G est $F : \mathbb{R}^{11} \times \mathbb{R}^{11} \rightarrow \mathbb{R}^{11}$.

Les composantes F_i de F associées aux contraintes de distances *distance*(P_j, P_k) = h_i s'écrivent

$$F_i(X, h_i) = \|\overrightarrow{P_j P_k}\|_2 - h_i.$$

Elles sont définies pour tous $\sigma(P_j), \sigma(P_k), \sigma(h_i)$, mais dérivables seulement quand $\sigma(P_j) \neq \sigma(P_k)$. Dans \mathbb{R}^{11} , l'espace de toutes les valuations de X , les valuations telles que $\sigma(P_j) = \sigma(P_k)$ forment un hyperplan \mathcal{H}_i de dimension $11 - 3 = 8$.

Les composantes F_i de F associées aux contraintes d'angles *angle*(P_j, P_k, P_l) = a_i s'écrivent

$$F_i(X, a_i) = \frac{\overrightarrow{P_k P_j} \cdot \overrightarrow{P_k P_l}}{\|\overrightarrow{P_k P_j}\|_2 \|\overrightarrow{P_k P_l}\|_2} - \cos(a_i).$$

Elles sont définies et infiniment dérivables quand $\sigma(P_j) \neq \sigma(P_k)$ et $\sigma(P_l) \neq \sigma(P_k)$, donc dans l'ensemble \mathcal{H}_i qui est l'union de deux hyperplans de dimension 8 dans l'ensemble de toutes les

valuations de X . Remarquons que pour obtenir une expression sans division d'une contrainte d'angle, on pourrait utiliser la fonction

$$F'_i(X, a_i) = \|\overrightarrow{P_k P_j}\|_2 \|\overrightarrow{P_k P_l}\|_2 F_i(X, a_i),$$

où F_i est la fonction numérique associée à une contrainte d'angle donnée ci-dessus, définie pour tous $\sigma(P_j), \sigma(P_k), \sigma(P_l), \sigma(a_i)$, mais dérivable seulement quand $\sigma(P_j) \neq \sigma(P_k)$ et $\sigma(P_l) \neq \sigma(P_k)$. Quand $\sigma(P_j) \neq \sigma(P_k)$ et $\sigma(P_l) \neq \sigma(P_k)$, les zéros de F'_i sont exactement ceux de F_i ; cependant, F'_i et F_i sont C^∞ sur le même domaine.

Enfin, les composantes F_i de F associées aux coplanarités *coplanaires* $(P_{k_1}, P_{k_2}, P_{k_3}, P_{k_4})$ s'écrivent

$$F_i(X, h_i) = \det(\overrightarrow{P_{k_1} P_{k_2}}, \overrightarrow{P_{k_1} P_{k_3}}, \overrightarrow{P_{k_1} P_{k_4}}) - h_i,$$

et sont C^∞ sur l'ensemble de toutes les valuations de X .

Considérons la fonction d'homotopie $H(X, t) = F(X, d(t))$, où $H : \mathbb{R}^{11} \times \mathbb{R} \rightarrow \mathbb{R}^{11}$ et d est une fonction d'interpolation de $\sigma_{sk}(A)$ vers $\sigma_{so}(A)$. Si d est C^∞ sur \mathbb{R} , il est clair que H est définie et infiniment dérivable sur un ensemble $\mathcal{D} \subset \mathbb{R}^{11} \times \mathbb{R}$, où \mathcal{D} s'écrit

$$\mathcal{D} = (\mathbb{R}^{11} \times \mathbb{R}) \setminus \cup_i (\mathcal{H}'_i),$$

où les \mathcal{H}'_i sont des hyperplans ou des unions d'hyperplans de dimensions 9 dans l'espace $\mathbb{R}^{11} \times \mathbb{R}$. Pour chacun des points u de \mathcal{D} , on peut construire un voisinage U tel que $id|_U : \mathbb{R}^{12} \rightarrow \mathbb{R}^{12}$ (la fonction identité) soit un difféomorphisme, et l'image V de U par ce difféomorphisme est un ouvert de \mathbb{R}^{12} . \mathcal{D} est donc bien une variété de dimension 12 sans bord.

3.3.2 Caractérisation des chemins d'homotopie

On introduit ici la notion de régularité d'une valeur prise par une fonction, portant sur le rang de sa matrice jacobienne en les antécédents de cette valeur :

Définition 6 Soit $f : \mathcal{C} \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ une fonction C^1 avec $p \geq q$, $x \in \mathcal{C}$ et $Jf(x)$ sa matrice jacobienne en x . Si $\text{rang}(Jf(x)) < q$ (c.à.d. le rang de $Jf(x)$ n'est pas maximal), on dit que x est un point critique de f , et que $f(x)$ est une valeur critique de f .

Soit $H : \mathcal{D} \subseteq \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m$ une fonction d'homotopie C^∞ sur la variété \mathcal{D} de dimension $m+1$ sans bord. Comme conséquence directe de deux lemmes tirés de [Milnor 65] et rappelés en annexe A (voir les lemmes 23 et 24), on établit la proposition suivante :

Proposition 7 Si 0 est une valeur régulière de H , les composantes connexes de $H^{-1}(0)$ sont des sous-variétés fermées de \mathcal{D} de dimension 1, sans bord.

On rappelle que les composantes connexes de $H^{-1}(0)$ sont appelées chemins d'homotopie de H . D'après le théorème de classification des variétés de dimension 1 (théorème 25 en annexe A), une variété de dimension 1 est difféomorphe soit au cercle $\mathcal{S}^1 = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 1\}$, soit à un intervalle ouvert $]0, 1[$, fermé $[0, 1]$, ou semi-ouvert $[0, 1[$ (ou $]0, 1]$) de \mathbb{R} . Ici, comme les chemins d'homotopie de H sont sans bord, ils sont difféomorphes soit à \mathcal{S}^1 , soit à l'intervalle ouvert $]0, 1[$.

Un chemin d'homotopie difféomorphe à l'intervalle $]0, 1[$ peut avoir une longueur infinie. Dans ce cas, si on note $\pi_t : \mathbb{R}^{11} \times \mathbb{R} \rightarrow \mathbb{R}$ la projection canonique d'un vecteur de $\mathbb{R}^{11} \times \mathbb{R}$ sur sa dernière composante (celle correspondant au paramètre d'interpolation), et \mathcal{S} un tel chemin d'homotopie, on peut distinguer deux cas de figure :

1. $\pi_t(\mathcal{S})$ est borné,

2. $\pi_t(\mathcal{S})$ n'est pas borné.

Dans le second cas, G admet une solution pour toutes les valuations $d(t)$ où $t \in \mathbb{R}$. La proposition 9 et sa preuve à la section 3.3.4 montrent que sous une condition simple sur la fonction d'interpolation d , un tel chemin \mathcal{S} difféomorphe à $]0, 1[$ et de longueur infinie est du type 1, et qu'il existe $t_l \in \mathbb{R}$ tel que \mathcal{S} converge vers une solution à l'infini du système $H(\cdot, t_l) = 0$. On y montre également que si \mathcal{S} n'est pas de longueur infinie, il converge vers un point où H n'est pas définie ou pas C^∞ .

Définition 8 Soit d une fonction d'interpolation de $\alpha \in \mathbb{R}$ vers $\beta \in \mathbb{R}$. On appelle support positif de d , et on le note $\text{Supp}(d)$, l'ensemble $\{x \in \mathbb{R} \mid d(x) \geq 0\}$. Si $d = (d_0, \dots, d_{p-1})$ est une fonction d'interpolation de $a \in \mathbb{R}^p$ vers $b \in \mathbb{R}^p$, on appelle support positif de d , et on le note $\text{Supp}(d)$, l'intersection des supports positifs des composantes d_j de d .

Remarquons avant d'établir la proposition 9 que si H_i est une composante de H correspondant à une contrainte de distance, et d_i sa fonction d'interpolation, alors $\forall (x, t) \in H^{-1}(0)$, $t \in \text{Supp}(d_i)$. En effet, une distance ne peut pas être négative.

Proposition 9 Soit \mathcal{S} une composante connexe de $H^{-1}(0)$; on suppose que H_0 correspond à une contrainte de distance, et que $\text{Supp}(d_0)$, le support positif de la fonction d'interpolation d_0 pour le paramètre de cette contrainte, est compact. Si $0 \in \mathbb{R}^m$ est une valeur régulière de H , \mathcal{S} est difféomorphe soit à un cercle, soit à un intervalle ouvert de \mathbb{R} . Dans le dernier cas, au moins une des deux assertions suivantes a lieu :

- (i) $\exists t_l \in \text{Supp}(d_0)$ tel que les éléments de \mathcal{S} convergent vers une solution à l'infini du système $H(X, t_l) = 0$,
- (ii) $\exists (\sigma_l(X), t_l) \in (\mathbb{R}^m \times \mathbb{R}) \setminus \mathcal{D}$ tel que $(\sigma_l(X), t_l)$ est un point de la fermeture de \mathcal{S} .

Cette proposition est prouvée dans la sous-section 3.3.4. On montre dans le chapitre 4 que toute solution à l'infini de $H(X, t_l) = 0$ correspond à une configuration géométrique des objets d'une figure $\sigma(X)$, et que cette configuration peut-être détectée pendant le suivi de \mathcal{S} en utilisant un plan de construction. Dans le cas où un chemin \mathcal{S} vérifie l'assertion (ii), il existe une suite de ses éléments qui converge vers un point de $(\mathbb{R}^m \times \mathbb{R}) \setminus \mathcal{D}$, c'est à dire vers un point où H n'est pas définie ou pas C^∞ . Or de tels points correspondent à des configurations sur $\sigma(X)$ faciles à énumérer en observant les contraintes de G . On propose d'illustrer les conclusions de cette proposition sur l'exemple introduit plus haut.

3.3.3 Un exemple : les chemins d'homotopie de 6p7d3a2c

On résout le problème 6p7d3a2c pour trois valuations $\sigma_1(A)$, $\sigma_2(A)$ et $\sigma_3(A)$ des paramètres différentes. Ces valuations un peu artificielles ont été construites dans l'objectif de faire apparaître les trois types de chemins d'homotopie qui peuvent être rencontrés d'après la proposition 9. On rappelle que $\sigma_{sk}(A)$ est la valuation des paramètres lue sur l'esquisse.

À l'exception des trois mesures d'angles, les valeurs de $\sigma_i(A)$ pour $i = 1, 2, 3$ sont les mêmes que $\sigma_{sk}(A)$. On définit d^i la fonction d'interpolation de $\sigma_{sk}(A)$ vers $\sigma_i(A)$ par

- $d_0^i(t) = -t^2 + (\sigma_i(h_0) + \sigma_{sk}(h_0) + 1)t + \sigma_{sk}(h_0)$ (H_0 correspond à une contrainte de distance),
- $d_j^i(t) = (1-t)\sigma_{sk}(a_j) + t\sigma_i(a_j)$ pour $1 \leq j \leq m$ si H_i correspond à une contrainte d'angle,
- $d_j^i(t) = (1-t)\sigma_{sk}(h_j) + t\sigma_i(h_j)$ pour $1 \leq j \leq m$ sinon.

Les fonctions d_0^i ont un support positif compact et, à l'exception des composantes correspondant à une contrainte d'angle, les d_j^i pour $j \neq 0$ sont constantes. Pour simplifier, par la suite, on notera d_1^i , d_2^i et d_3^i les composantes de d^i correspondant aux contraintes d'angles de paramètres respectifs a_1 , a_2 et a_3 .

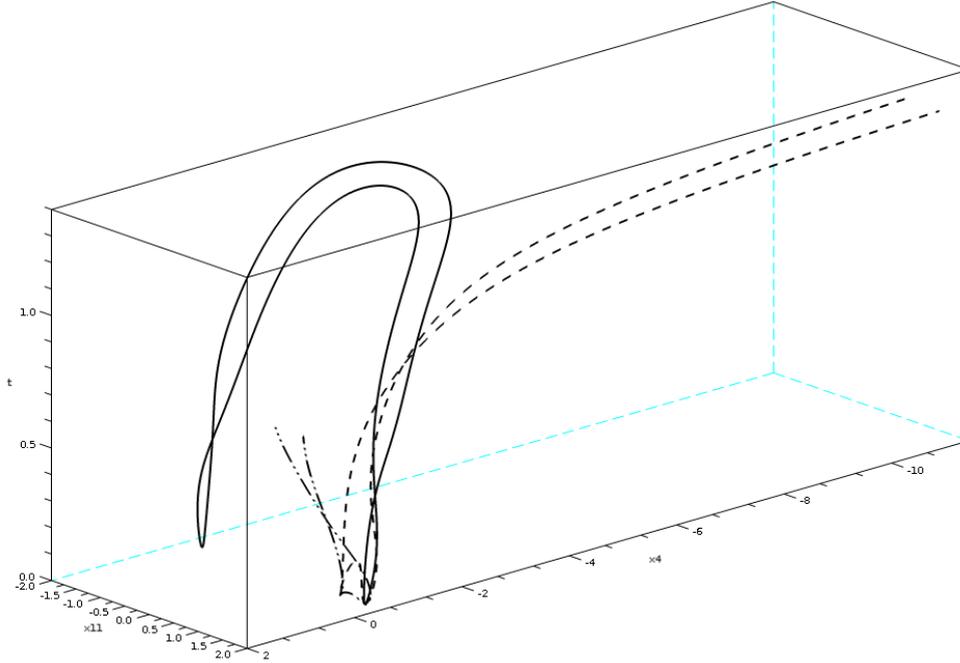


FIGURE 3.9 – Une projection sur \mathbb{R}^3 de trois chemins d’homotopie pour le problème 6p7d3a2c avec les valuations des paramètres à résoudre σ_1 (en trait plein), σ_2 (en pointillés) et σ_3 (en pointillés, avec un trait, deux points).

On considère la résolution de G pour la valuation $\sigma_2(A)$ des paramètres, où $\sigma_2(a_3) = \sigma_{sk}(a_3)$ (d_3^2 est constante). Supposons qu’il existe $t_2 \in \text{Supp}(d)$ tel que $d_1^2(t_2) = d_2^2(t_2) = 0$. Si une figure $\sigma(X)$ est solution de G pour la valuation $d^2(t_2)$, elle vérifie $\widehat{P_2P_3P_4} = \widehat{P_2P_3P_0} = 0$. Sur une telle figure, l’intersection des trois droites (P_0P_3) , (P_2P_3) , (P_4P_3) , qui est le lieu géométrique de P_3 , se situe infiniment loin de P_0 , et P_3 a une norme infinie, et $\sigma(X)$ est une solution à l’infini du système $H(\cdot, t_2) = 0$. La courbe en pointillés de la figure 3.9 est une projection sur \mathbb{R}^3 d’un chemin d’homotopie de H construit avec la fonction d’interpolation d^2 convergeant vers une solution à l’infini de $H(\cdot, t_2) = 0$, avec $t_2 \simeq 1.25$.

Dans le cas de la valuation σ_3 des paramètres, il existe t_3 tel que $d_3^3(t_3)$ (la mesure de l’angle $\widehat{P_1P_0P_3}$) est égale à la mesure de l’angle $\widehat{P_1P_0P_2}$, et $d_2^3(t_3)$ (la mesure de l’angle $\widehat{P_2P_3P_0}$) est une constante appartenant à l’intervalle $]0, \pi[$. Sur une figure solution de G pour la valuation $d^3(t_3)$, P_3 coïncide soit avec P_0 , soit avec P_2 (car $\widehat{P_2P_3P_0}$ n’est pas un angle plat), et mesurer l’angle $\widehat{P_1P_0P_3}$ n’a plus de sens. H n’est pas définie en une telle figure, et le chemin d’homotopie de H converge vers cette figure. La courbe en pointillé (un trait, deux points) de la figure 3.9 est une projection d’un tel chemin d’homotopie, où $t_3 \simeq 0.7$.

Quand on résout G pour la valuation σ_1 , aucun des phénomènes ci-dessus ne se produit, et la composante connexe de $H^{-1}(0)$ à laquelle appartient $\sigma_{sk}(X)$ est difféomorphe à un cercle ; sa projection est représentée sur la figure 3.9 en trait plein. Remarquons qu’elle coupe quatre fois l’hyperplan \mathcal{P}_1 et que pour la valuation $\sigma_1(A)$, quatre solutions de G sont trouvées quand le chemin d’homotopie est suivi entièrement.

3.3.4 Preuve de la proposition 9

Soit \mathcal{S} un chemin d'homotopie de la fonction H , tel qu'il existe un difféomorphisme $\varphi :]0, 1[\rightarrow \mathcal{S}$. On notera x le vecteur dont les composantes sont les coordonnées de X à déterminer. On montre tout d'abord qu'il existe t_∞ tel que $\lim_{z \rightarrow 1} \pi_t(\varphi(z)) = t_\infty$.

Lemme 10 *Soit $t \in \mathbb{R}$, et $H_t : \mathbb{R}^m \rightarrow \mathbb{R}^m$ la fonction définie par $H_t(x) = H(x, t)$. Si 0 est une valeur régulière de H , alors les vecteurs $x \in \mathbb{R}^m$ tels que $H_t(x) = 0$ sont isolés, et en nombre fini.*

D'après le lemme de la page 11 de [Milnor 65], $H_t^{-1}(0)$ est une variété différentielle de dimension 0, donc consiste en des points isolés. D'autre part, les vecteurs $x \in \mathbb{R}^m$ tels que $H_t(x) = 0$ sont les solutions d'un système polynomial², qui admet seulement un nombre fini de solutions isolées. \square

Comme $Supp(d_0)$ est compact, il existe $t_0 \in \mathbb{R}_+$ vérifiant

$$\forall z \in]0, 1[, -t_0 < \pi_t(\varphi(z)) < t_0.$$

Soit $t_1 \in]-t_0, t_0[$. D'après le lemme 10, il existe $M_1 \in]0, 1[$ tel que

$$\forall z > M_1, \pi_t(\varphi(z)) < t_1 \text{ ou } \pi_t(\varphi(z)) > t_1,$$

sinon il existerait un nombre infini de $x \in \mathcal{S}$ tels que $\pi_t(x) = t_1$ et $H_{t_1}(x) = 0$. Donc pour $z \in]M_1, 1[$,

(i) $\pi_t(\varphi(z)) \in]-t_0, t_1[$, ou

(ii) $\pi_t(\varphi(z)) \in]t_1, t_0[$.

On notera $\mathcal{I}_1 =]-t_0, t_1[$ si (i) est vrai, $\mathcal{I}_1 =]t_1, t_0[$ sinon. Soit t_2 le milieu de l'intervalle \mathcal{I}_1 . D'après le lemme 10, il existe $M_2 \in]M_1, 1[$ tel que

$$\forall z > M_2, \pi_t(\varphi(z)) < t_2 \text{ ou } \pi_t(\varphi(z)) > t_2.$$

Soit $\mathcal{I}_2 =]\inf(\mathcal{I}_1), t_2[$ si $\pi_t(\varphi(z)) \in]\inf(\mathcal{I}_1), t_2[$, $\mathcal{I}_2 =]t_2, \sup(\mathcal{I}_1)[$ sinon, où $\inf(\mathcal{I})$ (respectivement $\sup(\mathcal{I})$) est la borne inférieure (respectivement supérieure) d'un intervalle \mathcal{I} . En appliquant itérativement ce raisonnement, on construit une suite d'intervalles $(\mathcal{I}_n)_{n \in \mathbb{N}}$ et en notant $\text{midd}(\mathcal{I})$ le milieu d'un intervalle \mathcal{I} , trois suites de réels $(\inf(\mathcal{I}_n))_{n \in \mathbb{N}}$, $(\text{midd}(\mathcal{I}_n))_{n \in \mathbb{N}}$ et $(\sup(\mathcal{I}_n))_{n \in \mathbb{N}}$. On montre alors facilement que ces trois suites convergent vers la même limite $t_\infty \in Supp(d_0)$. On a alors bien $\lim_{z \rightarrow 1} \pi_t(\varphi(z)) = t_\infty$.

Pour terminer la preuve de la proposition, on peut distinguer deux cas : soit \mathcal{S} a son adhérence $\overline{\mathcal{S}}$ incluse dans \mathcal{D} , soit il existe $(\sigma_\infty(X), t_\infty) \in (\mathbb{R}^m \times \mathbb{R}) \setminus \mathcal{D}$ et une suite de vecteurs de \mathcal{S} qui converge vers $(\sigma_\infty(X), t_\infty)$.

Dans le premier cas ($\overline{\mathcal{S}} \subset \mathcal{D}$), \mathcal{S} est fermé dans $\mathbb{R}^m \times \mathbb{R}$. Si \mathcal{S} était borné, il serait un compact de $\mathbb{R}^m \times \mathbb{R}$, donc difféomorphe à un intervalle fermé de \mathbb{R} et aurait un bord non vide. Donc \mathcal{S} n'est pas borné, d'où $\forall M \in \mathbb{R}, \exists s_1, s_2 \in \mathcal{S}$ t.q. $\|s_2 - s_1\| > M$. Alors $\lim_{z \rightarrow 1} \|\varphi(z)\| = +\infty$, et les éléments de \mathcal{S} sont solution à l'infini du système $H_{t_i}(x) = 0$, et on a montré (i). Dans le deuxième cas, on a exactement (ii). \square

2. Quand une (ou plusieurs) composante de H correspond à une contrainte d'angle, on peut utiliser son expression sans division donnée plus haut ; les ensembles de zéros coïncident sur le domaine où H est C^∞ . Une expression polynomiale de H s'obtient aisément en élevant ses composantes au carré.

3.4 Conclusion

On a introduit dans ce chapitre un cadre général de résolution de systèmes de contraintes géométriques par homotopie. À la différence des méthodes par continuation classiques, le système initial utilisé n'est pas un système générique dont toutes les solutions sont connues. C'est ici l'esquisse qui est utilisée pour définir le système initial, dont elle est la seule solution connue. Les paramètres du système initial sont alors déformés en les paramètres du système cible, et l'esquisse est déformée en des solutions.

Implantation

Les travaux présentés ici ont pour but d'être intégrés à la plate-forme de résolution de contraintes géométriques `Constraint` développée et maintenue par Pascal Mathis au sein de l'équipe IGG du laboratoire ICube. Cependant, les résultats présentés plus bas sont le fruit d'une implantation ad hoc, en C++. Un SCG y est représenté par sa fonction numérique associée et la méthode de suivi de courbe par prédiction-corrrection présentée plus haut (voir sous-section 3.2.2) effectuée uniquement du calcul numérique (aucune représentation symbolique, notamment pour le calcul de la matrice jacobienne, n'est utilisée). Les sorties graphiques présentées dans tout le manuscrit ont été réalisées grâce au logiciel `Scilab`³.

Quelques résultats

Voici quelques résultats quantitatifs d'application de la démarche présentée dans ce chapitre pour quatre problèmes en 3D, tous structurellement bien contraints. Deux d'entre eux n'impliquent que des contraintes de distance, les deux autres comportent en plus des coplanarités entre quatre points. Ils résistent tous aux stratégies de décomposition ; ils peuvent aussi être vus comme des sous-problèmes résultant d'une méthode de décomposition. Aucun d'entre eux n'est résoluble par LIM.

Pour chacun de ces problèmes, la table 3.2 donne le nombre de solutions trouvées sur le chemin de l'esquisse et le temps en secondes pris pour suivre ce chemin dans la ligne "chemin de l'esquisse". Quand c'est possible (*c.à.d* quand le problème n'a pas trop de contraintes) on compare ces résultats avec ceux obtenus par le logiciel HOM4PS-2.0 (voir [Lee 08]) en posant naïvement le système d'équations. Pour les problèmes faisant intervenir des contraintes booléennes, et dont certaines solutions sont des points critiques, le nombre de telles solutions trouvées par notre méthode est donné entre parenthèses. Les tests ont été réalisés sur une machine munie d'un processeur Intel(R) Core(TM) i5 CPU 750 @ 2.67GHz.

TABLE 3.2 – Temps d'exécution sur un Intel(R) Core(TM) i5 CPU 750 @ 2.67GHz.

| | Disulfide | Pentaèdre | Hexaèdre | Icosaèdre |
|------------------------|-----------|-----------|----------|-----------|
| HOM4PS-2.0 : | | | | |
| nb solutions | 18 | 9 (6) | 16 (0) | - |
| temps | 6129s | 23s | 12800s | - |
| chemin de l'esquisse : | | | | |
| nb solutions | 8 | 9 (6) | 7 (3) | 28 |
| temps | 3s | 0.4s | 1.6s | 9s |

3. www.scilab.org/fr

La molécule de disulfide

Le premier problème vient du domaine de la chimie moléculaire et est emprunté à [Porta 07]. Il consiste en la construction de 8 points dans l'espace en connaissant 18 distances entre deux points, et correspond au placement des atomes d'une molécule de disulfide. La partie gauche de la figure 3.10 exhibe une esquisse pour ce problème. Pour la valuation des paramètres des distances donnée dans [Porta 07], ce problème admet 18 solutions, toutes trouvées par une méthode numérique par bisection en plus de 10 minutes.

Pour résoudre ce problème par continuation, on lui associe une fonction $H : \mathbb{R}^{19} \rightarrow \mathbb{R}^{18}$. Le chemin contenant l'esquisse intersecte 8 fois \mathcal{P}_1 (donnant 8 solutions) et est suivi entièrement en 3 secondes.

Le pentaèdre, ou problème 2

Le problème 2, pour la valuation :

$$\begin{array}{lll} h_0 = 1.3, & h_1 = 2.9, & h_2 = 2.6, \\ h_3 = 1.35, & h_4 = 1.9, & h_5 = 1.5, \\ h_6 = 0.7, & h_7 = 1.1, & h_8 = 2.6, \end{array}$$

admet 9 solutions, trouvées en environ 23 secondes par HOM4PS-2.0. Parmi celles-ci, 6 sont aplaties. Ces 9 solutions se trouvent sur le chemin de l'esquisse, suivi entièrement en 0.4 secondes.

L'hexaèdre, ou problème 3

Le problème 3 consiste en la construction d'un solide à 8 points, 12 arêtes et 8 faces (contenant chacune 4 points). Il est illustré par la figure 3.5. Pour le résoudre par notre méthode, on lui associe une fonction d'homotopie $H : \mathbb{R}^{18} \rightarrow \mathbb{R}^{17}$. Pour la valuation des paramètres :

$$\begin{array}{llll} h_0 = 1.3, & h_1 = 1.3, & h_2 = 1.3, & h_3 = 0.5, \\ h_4 = 0.7, & h_5 = 0.9, & h_6 = 0.8, & h_7 = 0.8, \\ h_8 = 0.5, & h_9 = 1.9, & h_{10} = 1.8, & h_{11} = 1.5, \end{array}$$

il admet 16 solutions non aplaties, trouvées en environ 3 heures 30 minutes par HOM4PS-2.0, et une infinité de solutions aplaties qui forment une variété de dimension 1. Aucune de ces dernières n'est obtenue par HOM4PS-2.0 qui assure par une perturbation des paramètres que les chemins suivis ne contiennent aucun point critique. Notre méthode *traverse* ces points critiques, et fournit les figures correspondantes comme solutions (le continuum auquel elles appartiennent n'est pas exploré). Lors du suivi du chemin contenant l'esquisse donnée à la figure 3.5, 4 solutions non aplaties et 3 aplaties sont trouvées.

L'icosaèdre

Un icosaèdre est un solide à 12 sommets, 30 arêtes et 20 faces toutes triangulaires. Le SCG correspondant au problème :

Problème 6 *Construire un icosaèdre en connaissant la longueur de ses 30 arêtes,*

contient donc en tout 30 contraintes de distances, et la fonction d'homotopie associée est $H : \mathbb{R}^{31} \rightarrow \mathbb{R}^{30}$. La partie droite de la figure 3.10 présente une esquisse de ce problème. Pour presque toute valuation des paramètres, ses solutions sont toutes des points isolés. Ce problème a trop de contraintes pour être résolu par HOM4PS-2.0 (le processus ne termine pas).

Avec notre méthode, le chemin contenant l'esquisse de la figure 3.10 contient 28 solutions, trouvées en environ 9 secondes.

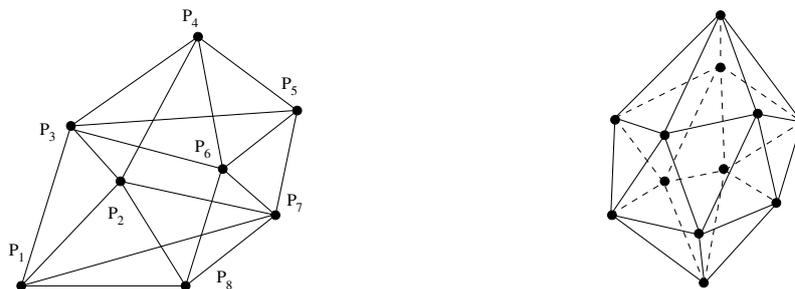


FIGURE 3.10 – Deux esquisses cotées : pour le problème de la molécule de disulfide (à gauche), et pour le problème de l’icosaèdre (à droite). Les arêtes représentent les contraintes de distance.

Contributions

La contribution de ce chapitre réside dans l’interpolation non linéaire de ces paramètres qui influe sur la topologie des chemins d’homotopie : ils sont, dans les cas favorables, des cercles et permettent de trouver plusieurs solutions. Les cas défavorables sont les suivants :

1. le chemin auquel appartient l’esquisse ne mène à aucune solution (il ne croise pas l’hyperplan \mathcal{P}_1);
2. le chemin auquel appartient l’esquisse a une longueur infinie et tend vers une solution à l’infini;
3. le chemin auquel appartient l’esquisse contient un point critique.

Notons que les cas 2 et 3 ne signifient pas qu’aucune solution n’appartient au chemin. Le cas 3 est abordé dans le chapitre 5, et le cas 2 dans le chapitre 4, qui présente aussi une méthode pour générer de nouvelles esquisses à partir des solutions obtenues.

En appliquant cette méthode pour générer de nouvelles esquisses, on trouve pour la plupart des “petits” problèmes sur lesquels on a appliqué notre méthode toutes les solutions, comme le montre la table 3.3 dans la ligne “processus complet”. Dans le cas de la molécule de disulfide seulement 13 solutions sur les 18 sont obtenues ; on discute de ce résultat dans le chapitre suivant, en proposant une alternative de notre méthode avec laquelle les 18 solutions sont trouvées. L’icosaèdre est un problème trop conséquent pour mener la recherche de nouvelles solutions à son terme ; de nouvelles solutions sont toutefois trouvées.

TABLE 3.3 – Temps d’exécution sur un Intel(R) Core(TM) i5 CPU 750 @ 2.67GHz.

| | Disulfide | Pentaèdre | Hexaèdre | Icosaèdre |
|---------------------|-----------|-----------|-----------|-----------|
| HOM4PS-2.0 : | | | | |
| nb solutions | 18 | 9 (6) | 16 (0) | - |
| temps | 6129s | 23s | 12800s | - |
| processus complet : | | | | |
| nb solutions | 13 | 9 (6) | 251 (235) | - |
| temps | 108s | 1.8s | 136s | - |

On propose pour terminer une réflexion sur la façon de remédier au cas 1. Il a été mis en évidence que la fonction d’interpolation utilisée pour définir la fonction d’homotopie modifiait la topologie des chemins d’homotopie. Elle a évidemment aussi une incidence sur la géométrie de ceux-ci, et

sur le fait qu'elle croise ou non l'hyperplan \mathcal{P}_1 . Cette influence de la fonction d'interpolation sur la géométrie des chemins d'homotopie paraît plus difficile à caractériser, mais pourrait sans doute être l'objet d'heuristiques. En tous cas, modifier cette fonction change le nombre de solutions obtenues, et quand le cas 1 se produit, changer la fonction d'interpolation peut résoudre le problème. Dans les résultats présentés au début de cette section, la fonction d'interpolation choisie est toujours la même. Sa composante correspondant à une contrainte de distance est une hyperbole de coefficient dominant -1 (qui a donc un support positif compact), ses autres composantes sont linéaires. Modifier le coefficient dominant de la composante hyperbolique peut changer le nombre de solutions obtenues. Une autre possibilité pour faire face au problème 1 est d'utiliser la méthode présentée dans le chapitre 4 permettant de générer d'autres esquisses. On met en avant son utilisation sur les solutions obtenues, mais elle peut être appliquée sur l'esquisse.

Chapitre 4

Utiliser un plan de construction

Le cadre de résolution présenté dans le chapitre précédent consiste en une homotopie dans l'espace des paramètres d'un système de contraintes géométriques, et déforme une esquisse en certaines de ses solutions. Les chemins d'homotopie sont suivis dans un espace réel (l'espace des figures) et non dans un espace projectif complexe. En conséquence, certaines propriétés garantissant le succès de la continuation (*c.à.d* la possibilité de trouver toutes les solutions du système cible à partir de celles du système initial) et la terminaison des méthodes générales sont perdues, notamment :

1. toutes les solutions du système cible ne sont pas trouvées,
2. un chemin d'homotopie peut avoir une longueur infinie.

On propose d'utiliser des raisonnements géométriques, rendus possibles par le fait que les points d'un chemin d'homotopie correspondent toutes à des figures, pour pallier ces défauts. Dans ce chapitre, ce raisonnement prendra la forme d'un Plan de Construction (PC).

Souvent, un PC est obtenu par une méthode géométrique ou combinatoire et son interprétation numérique fournit toutes les solutions d'un problème de contraintes géométriques, auquel cas un processus de résolution numérique comme celui décrit dans ce mémoire est superflu. Ces méthodes ont toutefois une efficacité limitée en 3D, et les SCG considérés ici leur résistent. Dans ce chapitre, les plans de constructions considérés ne sont pas utilisés pour construire des solutions : en fait les figures produites ne répondent pas à toutes les contraintes de l'énoncé. Ils permettent d'observer les figures le long des chemins d'homotopie pour détecter qu'un chemin tend vers l'infini. Ils permettent aussi de construire de nouvelles esquisses et ainsi définir de nouveaux chemins à suivre et espérer découvrir de nouvelles solutions.

La section 4.2 présente ces contributions, après avoir montré comment obtenir de tels plans de construction pour accomplir ces tâches. Une première approche consiste à produire un PC dit *canonique* ne tenant aucunement compte des contraintes du système à résoudre. Une seconde fournit un PC produisant des figures qui satisfont autant de contraintes du SCG que possible, en utilisant la méthode par reparamétrisation. Si les deux types de plans de construction sont valides, celui obtenu par la deuxième méthode présente des avantages qualitatifs, mais difficilement évaluable, pour la détection des solutions à l'infini. Précisons pour éviter les confusions que nous n'utilisons jusqu'ici de la reparamétrisation que sa phase symbolique, c'est à dire sa capacité à produire un PC dont les figures satisfont la plupart des contraintes du problème ; il est obtenu en remplaçant certaines contraintes du SCG initial par des contraintes métriques.

C'est la section 4.3 qui aborde le problème de la phase numérique de la reparamétrisation, à savoir trouver pour quelles valeurs des paramètres des contraintes ajoutées au SCG initial les figures produites par le plan de construction satisfont les contraintes qui en ont été ôtées. Le PC

y est considéré comme une fonction, construisant une figure pour une valuation des paramètres guides. Évaluer les contraintes non prises en compte sur cette figure permet de construire une fonction numérique dont les zéros sont les solutions du SCG initial. Cette fonction numérique dépend de beaucoup moins de variables que celle associée dans le chapitre précédent au SCG, et y appliquer une méthode numérique peut être beaucoup moins coûteux. Cependant, cette application est rendue difficile par le caractère multifonctionnel du PC ; une adaptation de la méthode de suivi de chemins par approximation linéaire par morceaux (PLI) à la multifonction correspondant à l'évaluation du PC est proposée. Les résultats en termes d'efficacité et de robustesse seront discutés dans cette section.

La section 4.1 formalise la notion de plan de construction.

4.1 Définitions et notations

Un plan de construction (PC) pour un ensemble d'objets géométriques X est une liste (ordonnée) d'instructions dépendant d'un ensemble de paramètres A , qui construisent les objets de X . Chacune de ces instructions correspond à la construction d'un élément de X à partir de ceux déjà définis et pour certains comme intersection de lieux géométriques. Quand un ensemble de lieux a plusieurs intersections, un choix est fait entre celles-ci. Ces instructions seront formalisées par des termes de l'univers géométrique. Évaluer un plan de construction (*c.à.d* appliquer dans l'ordre les instructions de la liste avec une série de choix) conduit à construire une figure. En utilisant le paradigme de l'arbre, on appelle une suite de choix une *branche* du plan de construction, et on dit que les figures résultant de son évaluation sont ses feuilles.

4.1.1 Plan de construction

Voici un plan de construction simple pour un ensemble $X = \{S_0, S_1, S_2, P_3\}$ dépendant des paramètres $A = \{P_0, P_1, P_2, h_0, h_1, h_2\}$ dans un "formalisme" type collège :

- (I0) construire la sphère S_0 centrée en P_0 de rayon h_0 ;
- (I1) construire la sphère S_1 centrée en P_1 de rayon h_1 ;
- (I2) construire la sphère S_2 centrée en P_2 de rayon h_2 ;
- (I3) construire P_3 à l'intersection de S_0, S_1, S_2 .

Pour l'écrire formellement, la signature de l'univers géométrique considéré est augmenté des éléments suivants :

Sortes : $\dots, sphere$
Symboles : $\dots,$
 $sphere(point, longueur) \rightarrow sphere,$
 $interSSS(sphere, sphere, sphere) \rightarrow point.$

Le symbole *sphere* définit une variable de sorte *sphere* à partir d'une variable de sorte *point* (son centre) et d'une de sorte *longueur* (son rayon), et *interSSS* définit une variable de sorte *point* à partir de trois variables de sorte *sphere*. Chaque instruction peut alors être écrite formellement comme un terme d'un symbole et de variable sortées de cette signature. Ainsi, l'instruction (I0) du PC ci-dessus peut s'écrire $S_0 = sphere(P_0, h_0)$, et l'instruction (I3) peut s'écrire $P_3 = interSSS(S_0, S_1, S_2)$. Comme dans le cas des contraintes d'un SCG, on écrit $i_0[S_0, \{P_0, h_0\}]$ le terme correspondant à l'instruction (I0), et $i_3[P_3, \{S_0, S_1, S_2\}]$ celui correspondant à (I3).

Les instructions d'un PC des inconnues X pour les paramètres A sont organisées en une séquence triangulaire : chacune définit une variable en fonction de variables définies précédemment, ou de

| | |
|-------------------------------------|---------------------------------|
| Inconnues : | $S_2 = sphere(P_1, k_{10})$ |
| point P_3, P_4, P_5 | $P_3 = interSSS(S_0, S_1, S_2)$ |
| sphere S_0, \dots, S_8 | $S_3 = sphere(P_1, k_9)$ |
| Paramètres : | $S_4 = sphere(P_2, k_8)$ |
| point P_0, P_1, P_2 | $S_5 = sphere(P_3, k_4)$ |
| longueur $k_2, k_3, k_4, k_5, k_6,$ | $P_4 = interSSS(S_3, S_4, S_5)$ |
| k_7, k_8, k_9, k_{10} | $S_6 = sphere(P_0, k_6)$ |
| Instructions : | $S_7 = sphere(P_1, k_7)$ |
| $S_0 = sphere(P_0, k_3)$ | $S_8 = sphere(P_4, k_5)$ |
| $S_1 = sphere(P_2, k_2)$ | $P_5 = interSSS(S_6, S_7, S_8)$ |

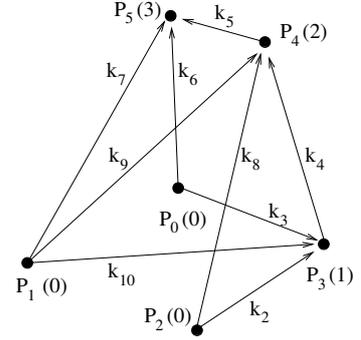


FIGURE 4.1 – À gauche un plan de construction de l'ensemble $\{P_3, P_4, P_5\}$. À droite une représentation graphique de ce plan de construction : une flèche représente une sphère utilisée pour construire le point vers lequel elle pointe, les nombres entre parenthèses indiquent l'ordre de construction des points.

paramètres. D'autre part, un élément de X apparaissant comme inconnue d'une instruction i_j ne peut pas apparaître comme inconnue d'une instruction suivante $i_{j'}$, et un paramètre de i_j est soit un élément de A , soit un élément de X apparaissant comme inconnue d'une instruction précédente.

Définition 11 Un plan de construction P d'un ensemble d'inconnues X et de paramètres A est un ensemble ordonné I de termes $\{i_0, \dots, i_{m-1}\}$ tel que :

- (i) pour tout $x \in X$, il existe un unique $0 \leq j \leq m-1$ tel que $i_j[x, A_j] \in I$,
- (ii) pour tout terme $i_j[x, A_j] \in I$, on a $x \in X$ et pour tout $a \in A_j$, soit $a \in A$, soit il existe $i_{j'}[a, A_{j'}] \in I$ où $0 \leq j' < j$.

On le notera $P = I[X, A]$.

À ces nouveaux éléments de la signature, on associe la sémantique de l'interprétation numérique. Une variable de sorte *sphere* est interprétée par un élément de $\mathbb{R}^3 \times \mathbb{R}_+$, qui est l'adjonction du rayon (un réel positif) au vecteur des coordonnées du centre. L'interprétation d'un terme de symbole *interSSS* est l'intersection de trois lieux géométriques. Ici, chaque lieu est une sphère ; un point P appartenant à une sphère de centre P_j et de rayon h_j vérifie $\|\overrightarrow{PP_j}\|_2 - h_j = 0$. L'interprétation du terme i_3 du plan de construction ci-dessus est donc le système d'équations :

$$\begin{cases} \|\overrightarrow{P_3 P_0}\|_2 - h_0 = 0 \\ \|\overrightarrow{P_3 P_1}\|_2 - h_1 = 0 \\ \|\overrightarrow{P_3 P_2}\|_2 - h_2 = 0 \end{cases}$$

Dans le plan de construction de l'ensemble $\{P_3, P_4, P_5, S_0, \dots, S_8\}$ donné dans la partie gauche de la figure 4.1, on peut voir les sphères S_0, \dots, S_8 comme des objets intermédiaires, servant uniquement à construire les points P_3, P_4 , et P_5 et on le notera par abus $P = I[\{P_3, P_4, P_5\}, A]$.

Parmi les instructions d'un plan de construction, on distingue celles qui construisent directement un objet, par exemple celles de symbole *sphere*(...), de celles qui définissent un objet comme l'intersection de lieux géométriques, par exemple celles de symboles *interSSS*(...), qu'on appellera *instruction d'intersection*.

4.1.2 Système d'équations associé

On a vu qu'une instruction d'intersection était interprétée par un système d'équations. En concaténant les systèmes interprétant chaque instruction d'intersection d'un plan de construction P , on lui associe un système d'équations :

Définition 12 Soit $P = I[X, A]$ un plan de construction, et $\mathcal{J} = \{j_0, \dots, j_{n-1}\}$ l'ensemble ordonné des indices des instructions d'intersection de P . On appelle système d'équations associé à P le système \mathcal{P} résultant de la concaténation, dans l'ordre donné par les indices de \mathcal{J} , des interprétations \mathcal{P}_k des instructions d'intersection i_{j_k} . On appelle \mathcal{P}_k un bloc de \mathcal{P} .

À titre d'exemple, on rappelle ci-dessous le plan de construction donné dans la figure 4.1, et on donne le système d'équations \mathcal{P} qui lui est associé.

| | | | | |
|--------------------------------|--|---|---------------|-----------------|
| Inconnues : | $S_0 = \text{sphere}(P_0, k_3)$ | } | \mathcal{P} | |
| point P_3, P_4, P_5 | $S_1 = \text{sphere}(P_2, k_2)$ | | | \mathcal{P}_0 |
| sphere S_0, \dots, S_8 | $S_2 = \text{sphere}(P_1, k_{10})$ | | | |
| Paramètres : | $P_3 = \text{interSSS}(S_0, S_1, S_2)$ | | | |
| point P_0, P_1, P_2 | $S_3 = \text{sphere}(P_1, k_9)$ | | | \mathcal{P}_1 |
| longueur $k_2, k_3, k_4, k_5,$ | $S_4 = \text{sphere}(P_2, k_8)$ | | | |
| $k_6, k_7, k_8, k_9, k_{10}$ | $S_5 = \text{sphere}(P_3, k_4)$ | | | |
| Instructions : | $P_4 = \text{interSSS}(S_3, S_4, S_5)$ | | | \mathcal{P}_2 |
| | $S_6 = \text{sphere}(P_0, k_6)$ | | | |
| | $S_7 = \text{sphere}(P_1, k_7)$ | | | |
| | $S_8 = \text{sphere}(P_4, k_5)$ | | | |
| | $P_5 = \text{interSSS}(S_6, S_7, S_8)$ | | | |

On dira que le système associé à un plan de construction est *triangulaire par bloc*, car chacun de ses blocs \mathcal{P}_k fait intervenir une seule inconnue n'intervenant pas dans les blocs précédents, et peut donc être vu comme un système de m_k équations en m_k inconnues, où m_k est le nombre de coordonnées de l'interprétation de cette inconnue. Chaque bloc peut être résolu l'un après l'autre, en utilisant les inconnues déterminées par les blocs précédents.

En fait, ce système d'équations peut être vu comme le système d'équations $F = 0$, où F est la fonction numérique associée au SCG dont P fournit toutes les solutions.

4.1.3 Évaluation d'un plan de construction

Pour une valuation $\sigma_*(A)$ des paramètres donnée, évaluer un plan de construction signifie :

1. dans le système d'équations associé, remplacer chaque paramètre par sa valeur par σ_* ,
2. résoudre l'un après l'autre les blocs du système d'équations.

Chaque bloc peut admettre aucune, une ou plusieurs solutions, ou une infinité, correspondant aux intersections des objets intervenant dans l'instruction qui lui est associée. Quand un bloc admet plusieurs solutions, une de ces solutions est choisie pour être utilisée dans la résolution des blocs suivants et les solutions obtenues pour les blocs suivants dépendent de ce choix, des choix précédents et des suivants. Si le nombre de choix à chaque bloc est fini, le résultat de cette évaluation est un ensemble fini, éventuellement vide, de figures, c'est à dire des valuations de X .

Une série de choix (un par bloc) peut être représentée par un vecteur d'entiers, dont la composante k indique la solution choisie pour le bloc \mathcal{P}_k . En supposant que l'évaluation de chaque

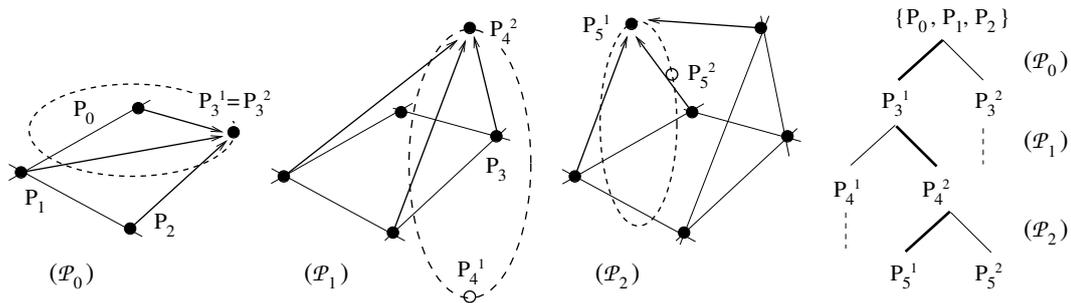


FIGURE 4.2 – L'évaluation des blocs \mathcal{P}_k du plan de construction de la figure 4.1, sur sa branche $b = (1, 2, 1)$. À droite, l'arbre associé; la branche choisie est celle dont les arêtes sont en traits épais.

instruction mène à intersecter des lieux qui ont un nombre fini d'intersections, l'ensemble de toutes les séries possibles de choix est fini, et on indexe les éléments de cet ensemble sur les entiers. Un élément de cet ensemble (*c.à.d* une série de choix) est une branche du plan de construction. Si b est une telle branche, on dira qu'on évalue un PC sur la branche b si on l'évalue en choisissant pour le bloc k la solution indiquée par la k -ème composante de la branche b . La figure 4.2 illustre l'évaluation du plan de construction ci-dessus, sur sa branche $b = (1, 2, 1)$. Cette évaluation est la résolution successive des trois blocs $\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2$, chacun laissant le choix entre deux solutions P_j^1 et P_j^2 . (voir la figure 4.2).

Quand pour une valuation des paramètres et une branche b donnée, un bloc \mathcal{P}_k du système associé à un PC n'admet aucune solution ou une infinité dans l'espace considéré, on dira que l'évaluation n'est pas possible. Dans l'exemple ci-dessus, si on résout \mathcal{P}_0 dans \mathbb{R}^3 (3 est le nombre de coordonnées de P_3 à déterminer), et que les trois sphères S_0, S_1, S_2 n'ont aucune intersection, \mathcal{P}_0 n'admet aucune solution. Dans cet exemple, l'évaluation du plan de construction n'est possible sur aucune de ses branches. Dans toute la suite, on évalue les plans de construction considérés (*c.à.d* on résout le système d'équation associé) dans l'ensemble \mathbb{R}^m , où m est le nombre de coordonnées des inconnues à déterminer.

Enfin, on considérera par la suite que résoudre un bloc \mathcal{P}_k dans un corps réel nécessite de choisir une solution parmi n_k , où n_k est le nombre maximum de solutions isolées que peut admettre \mathcal{P}_k . Dans ce mémoire, $n_k \leq 2$. Si un bloc correspond à l'intersection de trois sphères, ses solutions peuvent être : une sphère ou un cercle (et l'évaluation du PC n'est pas possible), un ou deux points, ou l'ensemble vide; ainsi $n_k = 2$. Quand $n_k = 2$ et que \mathcal{P}_k n'admet qu'une solution, par exemple quand deux sphères intersectées sont tangentes, on considérera que les deux choix possibles mènent à choisir la même solution. Il est donc possible, en particulier dans ce cas, que l'évaluation du plan de construction sur deux branches différentes mène à construire la même figure. Par exemple, dans la figure 4.2, les deux solutions possibles pour \mathcal{P}_0 sont égales car deux des sphères sont tangentes. Dans ce cas, l'évaluation du même plan de construction pour les mêmes paramètres sur sa branche $b' = (2, 2, 1)$ amène à construire la même figure.

4.1.4 Multifonction associée

On peut voir l'évaluation d'un plan de construction $P = I[X, A]$ comme une multifonction, qui associe à une valuation $\sigma(A)$ un ensemble de figures, qui sont le résultat de l'évaluation de P sur toutes ses branches $\{b_0, \dots, b_{n-1}\}$.

Définition 13 Soit $P = I[X, A]$ un plan de construction, \mathcal{P} le système d'équations associé, et b une branche. On appelle fonction associée à l'évaluation de P sur sa branche b , la fonction P_b qui à une valuation $\sigma(A)$ des paramètres associe la valuation $\sigma(X)$ des inconnues obtenue en évaluant P sur sa branche b (c.à.d $\sigma(X)$ est la solution obtenue en résolvant \mathcal{P} avec la série de choix b).

Comme il a déjà été remarqué, P_b n'est pas définie pour toute valuation; on notera \mathcal{D}_b le domaine de définition de P_b , c'est à dire l'ensemble des valuations pour lesquelles l'évaluation de P sur sa branche b est possible. Remarquons qu'il est facile de construire un exemple où \mathcal{D}_b n'est ni ouvert ni fermé dans l'ensemble de toutes les valuations de A . On notera ${}^\circ(\mathcal{D}_b)$ le plus grand ouvert contenu dans \mathcal{D}_b .

On dira alors que l'évaluation d'un plan de construction P est continue si pour toutes ses branches b , la fonction P_b est continue sur ${}^\circ(\mathcal{D}_b)$, et on supposera par la suite que les plans de construction considérés sont tels que leur évaluation est continue.

Définition 14 Soit $P = I[X, A]$ un plan de construction. On appelle multifonction associée à l'évaluation de P , et on la notera P lorsqu'il n'y a pas d'ambiguïté, la multi-fonction qui à une valuation $\sigma(A)$ des paramètres associe l'ensemble $\{P_{b_0}(\sigma(A)), \dots, P_{b_{n-1}}(\sigma(A))\}$ où $\{b_0, \dots, b_{n-1}\}$ est l'ensemble des branches de P .

4.2 Guider la résolution par homotopie

On suppose que le SCG G résiste aux méthodes géométriques de résolution et en particulier à la méthode des lieux LIM qui est en général utilisée pour fournir un plan de construction de X , dont l'évaluation sur toutes ses branches fournit toutes les solutions de G .

Si H est la fonction d'homotopie associée à G , tout point $(\sigma(X), t)$ d'un chemin d'homotopie de H correspond à une figure (qui est solution de G pour une certaine valuation de A). On utilisera ici un plan de construction P de X pour lequel pour chaque $\sigma(X)$ il existe une branche et une valuation des paramètres tels que son évaluation résulte en la construction de $\sigma(X)$.

On montre dans la sous-section 4.2.2 qu'une solution à l'infini de H correspond à une solution à l'infini d'un bloc du système d'équations associé à P et donc à une configuration particulière de la figure correspondant à la solution à l'infini. Analyser P avant le suivi du chemin permet de déterminer ces configurations possibles, pour détecter lors du suivi si une figure tend vers une telle configuration.

D'autre part, quand des solutions ont été obtenues pour G , on peut évaluer P sur toutes ses branches avec des valuations de paramètres lues sur chacune de ces solutions pour obtenir de nouvelles esquisses et suivre de nouveaux chemins d'homotopie; cet usage du plan de construction est présenté dans la sous-section 4.2.3.

On commence par proposer deux méthodes pour produire un plan de construction pour X . On discutera dans la sous-section 4.2.4 de la pertinence de chacune d'elles, c'est à dire de la qualité des plans de construction obtenus.

4.2.1 Produire un plan de construction

Soit $G = C[X, A]$ un SCG sur un univers géométrique 3D, et $\sigma_*(X)$ une valuation de X qui vérifie les contraintes booléennes de G . Sur la figure $\sigma_*(X)$, il est possible de lire les paramètres $\sigma_*(A)$. On suppose, pour simplifier, que X ne contient que des points.

Soit un plan de construction $P = I[X', A']$ tel que $X \subset (X' \cup A')$; une figure $\sigma(X')$ résultant de l'évaluation de P peut naturellement être vue comme une valuation de X . Si, de plus, les éléments de A' sont soit des éléments de X , soit des variables scalaires correspondant à des grandeurs géométriques (distances, angles,...) qui peuvent être lues sur une figure de X , une valuation $\sigma(A')$ des paramètres de P est uniquement déterminée par une valuation $\sigma(X)$, et il existe une branche b de P telle que l'évaluation de P sur cette branche pour $\sigma(A')$ résulte en $\sigma(X)$.

Prenons comme exemple le SCG $G = C[X, A]$ associé au problème de construction d'un pentaèdre (problème 2) où $X = \{P_0, \dots, P_5\}$, et considérons le plan de construction $P = I[X', A']$ donné à la figure 4.1, où $X' = \{P_3, P_4, P_5\}$ (on omet volontairement les sphères, qui sont des objets intermédiaires). On a bien $X \subset (X' \cup A')$, et en connaissant une valuation $\sigma(X)$, on peut facilement exhiber une valuation $\sigma_*(A')$ pour laquelle il existe une branche b_* telle qu'évaluer P pour $\sigma_*(A')$ sur la branche b_* conduise à construire $\sigma(X)$.

En effet, on a directement $\sigma_*(\{P_0, P_1, P_2\}) = \sigma(\{P_0, P_1, P_2\})$. Ensuite, comme P_3 est défini dans P comme l'intersection de trois sphères, P_3 appartient à chacune de ces sphères, et notamment $\sigma_*(k_3)$ (le rayon de S_0) est égal à $\|\overrightarrow{P_0 P_3}\|_2$. Dans ce cas particulier, en notant $\sigma(A)$ la valuation de A lue sur $\sigma(X)$, on aurait pu utiliser de manière équivalente $\sigma_*(k_3) = \sigma(h_3)$. Une valuation des autres paramètres de A' peut se déduire de la même façon de $\sigma(X)$.

Pour obtenir une branche b_* telle que $\sigma(X)$ résulte de l'évaluation de P sur b_* , il suffit alors d'évaluer une à une les instructions de P . À chaque fois qu'un choix entre deux solutions doit être fait, par exemple lors de la résolution du bloc correspondant à $P_3 = \text{interSSS}(S_0, S_1, S_2)$, il suffit d'envisager les deux solutions possibles pour P_3 et de choisir la plus proche¹ (en norme) de $\sigma(P_3)$.

Même si nous avons posé de manière informelle la condition, pour le PC $P = I[X', A']$, sous laquelle une figure $\sigma(X)$ peut être obtenue en évaluant P pour une valuation de A' sur une branche b lues sur l'esquisse, il est clair qu'il est toujours possible de construire un tel plan de construction. Deux méthodes sont proposées pour cela.

Utiliser des constructions d'Henneberg

En théorie de la rigidité, on représente un assemblage de tiges métalliques par un graphe, dont les sommets sont les joints et les arêtes les tiges métalliques. On dit d'un graphe qu'il est rigide si l'assemblage correspondant est rigide. Les constructions d'Henneberg sont des opérations d'ajout d'arêtes et de sommets à un graphe qui préservent sa rigidité. De plus, tout graphe rigide du plan peut être construit par une suite de telles constructions. On les exprime ici dans le contexte des SCG sur un univers géométrique 2D, dont les inconnues sont des points et les contraintes des distances.

- (C1) Ajouter à X une variable P de sorte *point*, à A deux paramètres h_1 et h_2 de sorte *longueur*, et à C deux contraintes $h_1 = \text{distance}(P, P_1)$ et $h_2 = \text{distance}(P, P_2)$, où $P_1 \in X$ et $P_2 \in X$.
- (C2) Ajouter à X une variable P de sorte *point*, à A trois paramètres h_1, h_2 et h_3 de sorte *longueur*, et à C trois contraintes $h_1 = \text{distance}(P, P_1)$ et $h_2 = \text{distance}(P, P_2)$, $h_3 = \text{distance}(P, P_3)$ où $P_1 \in X$, $P_2 \in X$ et $P_3 \in X$ sont tels qu'il existe une contrainte de distance $h_0 = \text{distance}(P_1, P_2)$ dans C . Retirer cette distance de C .

Tout SCG d'un univers 2D obtenu à partir de deux points impliqués dans une contrainte de distance par une suite de telles opérations est structurellement et génériquement bien contraint

1. Si les calculs étaient symboliques ou exacts, une des deux intersections coïnciderait avec $\sigma(P_3)$. En pratique, les calculs sont numériques et réalisés sur les flottants.

car respecte le critère de Laman. La réciproque est vraie aussi : tout SCG impliquant uniquement des points et des distances et qui est génériquement bien contraint peut être obtenu par une suite de telles constructions. Il est clair aussi que tout SCG obtenu par une suite de construction du premier type est soluble par LIM ; la suite d'opérations est directement transposable en un plan de construction.

La plupart de ces propriétés ne sont plus vraies quand on étend ces constructions en 3D. La propriété suivante reste vraie : tout SCG obtenu à partir de trois points impliqués deux à deux dans trois contraintes de distances par une suite d'opération du type :

(C1) ajouter à X une variable $P \notin X$ de sorte *point*, à A trois paramètres $h_1 \notin A$, $h_2 \notin A$, $h_3 \notin A$ de sortes *longueur*, et à C trois contraintes $h_1 = \text{distance}(P, P_1)$, $h_2 = \text{distance}(P, P_2)$ et $h_3 = \text{distance}(P, P_3)$ où $P_1 \in X$, $P_2 \in X$ et $P_3 \in X$;

est résoluble par LIM². Le triplet de points et de contraintes de distances de base peut être vu comme un repère. Un plan de construction du SCG obtenu par une suite de construction (C1) est déduit directement à partir des constructions.

Par exemple, le PC de la figure 4.1 a été obtenu par une telle suite de constructions à partir des points P_0, P_1, P_2 comme repère. Chacun des points de X encore à construire l'est l'un après l'autre en choisissant à chaque fois trois points déjà construits et en le définissant comme l'intersection de trois sphères centrées en ces points.

Utiliser un problème reparamétré

Le principe de la reparamétrisation est de transformer un problème $G = C[X, A]$ non résoluble par une méthode géométrique en un problème $G' = C'[X, A']$ résoluble par cette méthode en remplaçant certaines de ses contraintes par des contraintes métriques. [Gao 02] propose un algorithme pour obtenir un tel SCG G' pour la méthode géométrique LIM, rappelé ici dans nos notations.

Définissons dans le but d'exposer cet algorithme le degré de restriction d'un objet géométrique x pour un ensemble de contraintes C , que l'on note $DoC(C, x)$, comme la somme des degrés de restrictions des contraintes $c_i[X_i, A_i]$ de C telles que $x \in X_i$. On se donne la possibilité de marquer les objets et les contraintes ; on considère également que toutes les contraintes ont un degré de restriction de 1.

L'algorithme 1 (voir ci-contre) est la transcription dans le formalisme utilisé ici de l'algorithme de [Gao 02]. Il termine quand, avec les notations de l'algorithme, $d < 0$ et qu'il n'est pas possible de construire d contraintes métriques non redondantes entre elles et avec celles qui ne sont pas marquées. Dans cette situation, les éléments non marqués de X , c'est à dire les éléments de X_m forment un repère. Alors, avec les notations de l'algorithme 1, soit $|C^+| = |C^-|$ et le repère est bien déterminé (en fixant 6 coordonnées, le nombre de coordonnées libres de X_m est égal au nombre de contraintes de C_m), soit $|C^-| - |C^+| > 0$, et il faut ajouter $|C^-| - |C^+|$ contraintes au repère pour qu'il soit bien déterminé. On ajoute également ces contraintes à $|C^+|$.

En notant $G^+ = C^+[X, A^+]$ et $G^- = C^-[X, A^-]$, le SCG $G' = C'[X, (A \setminus A^-) \cup A^+]$ où $C' = (C \setminus C^-) \cup C^+$ est résoluble par LIM, et la suite des étapes de l'algorithme 1 permet d'obtenir un plan de construction des solutions de G' , qui est un plan de construction des inconnues X dépendant des paramètres $(A \setminus A^-) \cup A^+$.

2. L'extension en 3D de ces constructions donne lieu à des problèmes de combinatoires intéressants hors de l'utilisation qu'on en fait ici. Par exemple, l'icosaèdre, illustré dans la figure 3.10 est génériquement bien contraint, mais ne peut pas être obtenu par une suite de construction d'Henneberg.

Algorithme 1 Reparamétrisation

Entrées: $G = C[X, A]$ un SCG structurellement bien contraint, C^- et C^+ des ensembles de contraintes initialement vides. Aucune contrainte, ni objet, ne sont marqués.

1: Poser C_m l'ensemble des contraintes non marquées de C et X_m l'ensemble des objets non marqués de X_m .

2: Choisir $x \in X_m$ tel que $d = DoC(C_m, x) - DoF(x)$ est minimum.

3: **Si** $d > 0$,

choisir un ensemble $C_{supp} \subset C_m$ de d contraintes impliquant x , concaténer C_{supp} à C^- .

Si $d < 0$,

Si il est possible de construire un ensemble C_{add} de d contraintes métriques entre des objets de $X_m \setminus \{x\}$ et x , non redondantes avec les contraintes de C_m , concaténer C_{add} à C^+ .

Sinon aller à 5.

Sinon ($d = 0$) ne rien faire.

4: Marquer x et les contraintes de C_m impliquant x , retourner à 1.

5: X_m forme un repère.

Retourner C^- et C^+ .

4.2.2 Détecter les solutions à l'infini

Dans cette sous-section, on montre que les solutions à l'infini d'un SCG correspondent à des configurations géométriques particulières d'une figure, et qu'un plan de construction permet de les détecter.

Si $G = C[X, A]$ est le SCG que l'on cherche à résoudre, on considère un PC $P = I[X, A']$ comme décrit dans la sous-section précédente, et \mathcal{P} le système d'équations qui lui est associé. Soient encore H la fonction d'homotopie pour G construite avec la fonction d'interpolation d à support compact et \mathcal{S} le chemin d'homotopie auquel appartient l'esquisse. Si $(\sigma_0(X), t_0) \in \mathcal{S}$, $\sigma_0(X)$ est solution de \mathcal{P} , pour la valuation des paramètres $\sigma_0(A')$ lue sur $\sigma_0(X)$ (c.à.d. il existe une branche b telle que $P_b(\sigma_0(A')) = \sigma_0(X)$).

Supposons à présent que \mathcal{S} soit un chemin difféomorphe à un intervalle ouvert de \mathbb{R} de longueur infinie. D'après la proposition 9 (voir chapitre 3, section 3.3.1), on peut extraire de \mathcal{S} une suite $(\sigma_n(X), t_n)_{n \in \mathbb{N}}$ telle que :

(i) $\lim_{n \rightarrow +\infty} \|(\sigma_n(X), t_n)\| = +\infty$, (c.à.d le chemin a une longueur infinie)

(ii) $\exists t_\infty \in \text{supp}(d)$ tel que $\lim_{n \rightarrow +\infty} t_n = t_\infty$ (c.à.d sa projection en t converge).

En écrivant $X = \{x_0, \dots, x_m\}$, où les x_i sont les objets de X indexés selon leur ordre d'apparition dans \mathcal{P} , il existe un plus petit indice $i \in \{0, \dots, m\}$ tel que :

(iii) $\lim_{n \rightarrow +\infty} \|\sigma_n(x_i)\| = +\infty$,

(iv) $\exists M \in \mathbb{R}$ tel que $\forall 0 \leq j < i, \forall n \in \mathbb{N}, 0 \leq \|\sigma_n(x_j)\| \leq M$.

Si P respecte l'hypothèse de continuité donnée dans la sous-section 4.1.4, on peut facilement établir :

(v) la suite $(\sigma_n(\{x_0, \dots, x_{i-1}\}))_{n \in \mathbb{N}}$ converge.

En effet, cette suite peut être vue comme le résultat de l'évaluation d'un plan de construction P'' contenant seulement les premières instructions de P (celles déterminant les objets $\{x_0, \dots, x_{i-1}\}$) pour une suite de valuations $\sigma_n^*(A'')$ où $A'' \subset A'$ est l'ensemble des paramètres de P'' . On montre que $\sigma_n^*(A'')$ converge. En effet, on peut répartir les éléments de A'' en deux ensembles distincts : $A'' \cap A$ et $A'' \setminus A$. La suite $\sigma_n^*(A'' \cap A)$ converge, car ses éléments sont des projections de $d(t_n)$, qui converge

Inconnues :
point P_0, \dots, P_5
Paramètres :
longueur $h_0, h_1, h_5, \dots, h_9$
angle a_1, a_2, a_3
Contraintes :
distance(P_0, P_1) = h_0
...
distance(P_2, P_0) = h_9
angle(P_2, P_3, P_4) = a_1
angle(P_2, P_3, P_0) = a_2
angle(P_1, P_0, P_3) = a_3
coplanaires(P_0, P_1, P_2, P_3)
coplanaires(P_1, P_2, P_4, P_5)

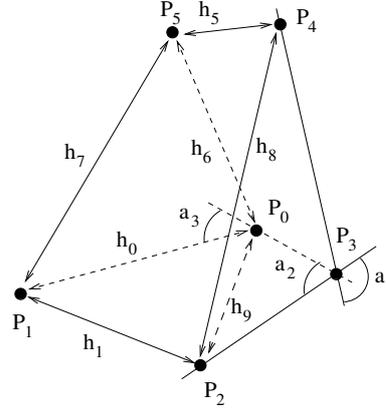


FIGURE 4.3 – L'énoncé formel et une esquisse cotée du problème 6p7d3a2c. Les doubles flèches droites représentent les contraintes de distance de paramètres h_i et les arcs de cercles entre deux droites joignant deux points sont des contraintes d'angle de paramètres a_i .

car d est continue. Les éléments de la suite $\sigma_n^*(A'' \setminus A)$ sont lus sur la figure $\sigma_n(\{x_0, \dots, x_{i-1}\})$, qui varie continûment en fonction de t par hypothèse. Comme t_n converge, il en va de même de la suite $\sigma_n^*(A'' \setminus A)$. \square

On a alors :

(vi) la limite de la suite $(\sigma_n(\{x_0, \dots, x_{i-1}\}))_{n \in \mathbb{N}}$ est solution du système formé des i premiers blocs de \mathcal{P} , pour la valuation des paramètres lue sur cette limite.

On déduit de (iii) et (vi) :

(vii) la suite $(\sigma_n(x_i))_{n \in \mathbb{N}}$ converge vers une solution à l'infini du $(i + 1)$ -ème bloc de \mathcal{P} , pour les valeurs $\lim_{n \rightarrow +\infty} \sigma_n(\{x_0, \dots, x_{j-1}\})$ des objets géométriques $\{x_0, \dots, x_{i-1}\}$ intervenant dans ce bloc.

Donc une solution à l'infini de H est nécessairement une solution de norme infinie d'un bloc de \mathcal{P} . De plus, il est clair que si une suite d'éléments d'un chemin d'homotopie de H converge vers une solution de norme infinie de \mathcal{P} , d'une part elle converge vers une solution à l'infini de H , et d'autre part il existe une projection de cette suite sur un sous-ensemble de ses composantes qui converge vers une solution de norme infinie d'un bloc de \mathcal{P} . On propose alors d'énumérer les cas où un bloc \mathcal{P}_i de \mathcal{P} admet des solutions de normes infinies, et plus précisément les valeurs $\sigma(A')$, $\sigma(\{x_0, \dots, x_{i-1}\})$ qui conduisent \mathcal{P}_i à admettre une solution de norme infinie, pour détecter sur la figure si une telle situation se produit lors du suivi de \mathcal{S} .

Illustrons cela sur le problème 6p7d3a2c déjà introduit dans le chapitre 3, dont le SCG $G = C[X, A]$ et une esquisse cotée $\sigma_{sk}(X)$ sont rappelés sur la figure 4.3. On a exhibé une valuation $\sigma_{so}(A)$ (appelée $\sigma_2(A)$ dans le chapitre 3) pour laquelle il existait t_∞ telle que dans $d(t_\infty)$, les mesures des angles $\widehat{P_2P_3P_4}$ et $\widehat{P_2P_3P_0}$ étaient nulles. On a également montré un chemin duquel on pouvait extraire une suite $(\sigma_n(X), t_n)_{n \in \mathbb{N}}$ avec $\lim_{n \rightarrow +\infty} \|(\sigma_n(X), t_n)\| = +\infty$ et $\lim_{n \rightarrow +\infty} t_n = t_\infty$. On montre comment détecter cette situation avec un plan de construction obtenu par des constructions d'Henneberg.

Le plan de construction $P = I[X', A']$ donné dans la figure 4.1 est un plan de construction des objets de X . Soit \mathcal{P} le système qui lui est associé (donné dans la sous-section 4.1.2). Tous ses blocs

correspondent à l'intersection de trois sphères, et son bloc \mathcal{P}_i est de la forme :

$$\mathcal{P}_i \begin{cases} \|\overrightarrow{P_{i'}P_{j_0}}\|_2 - k_{j_0} = 0 \\ \|\overrightarrow{P_{i'}P_{j_1}}\|_2 - k_{j_1} = 0 \\ \|\overrightarrow{P_{i'}P_{j_2}}\|_2 - k_{j_2} = 0 \end{cases}$$

où $P_{j_0}, P_{j_1}, P_{j_2}, k_{j_0}, k_{j_1}, k_{j_2}$ sont considérés comme des paramètres déterminés (soit ils appartiennent à A' , soit ils sont déterminés par des blocs précédents), et les points sont indexés dans leur ordre d'apparition dans \mathcal{P} .

Supposons qu'il existe une suite $(\sigma_n(X), t_n)_{n \in \mathbb{N}}$ avec $\lim_{n \rightarrow +\infty} \|(\sigma_n(X), t_n)\| = +\infty$, où i est le plus petit indice tel que $\lim_{n \rightarrow +\infty} \|(\sigma_n(\{P_0, \dots, P_i\}), t_n)\| = +\infty$.

Comme $(\sigma_n(\{P_{j_0}, P_{j_1}, P_{j_2}\}))_{n \in \mathbb{N}}$ converge, les normes de $\sigma_n(P_{j_0}), \sigma_n(P_{j_1}), \sigma_n(P_{j_2})$ sont finies, et \mathcal{P}_i admet une solution $\lim_{n \rightarrow +\infty} \sigma_n(P_{i'})$ de norme infinie si et seulement si $\lim_{n \rightarrow +\infty} \sigma_n(k_{j_0}), \lim_{n \rightarrow +\infty} \sigma_n(k_{j_1})$ et $\lim_{n \rightarrow +\infty} \sigma_n(k_{j_2})$ sont de norme infinie.

On peut alors distinguer deux cas :

- si une des contraintes $\text{distance}(P_{i'}, P_{j_0}) = a, \text{distance}(P_{i'}, P_{j_1}) = b$ ou $\text{distance}(P_{i'}, P_{j_2}) = c$ est dans C , alors $\sigma_n(a)$ ou $\sigma_n(b)$ ou $\sigma_n(c)$ est nécessairement bornée (car t_n est borné), et $\lim_{n \rightarrow +\infty} \sigma_n(P_{i'})$ ne peut pas être de norme infinie (ce qui entre en contradiction avec l'hypothèse $\lim_{n \rightarrow +\infty} \|(\sigma_n(X), t_n)\| = +\infty$),
- sinon $\lim_{n \rightarrow +\infty} \sigma_n(k_{j_0}), \lim_{n \rightarrow +\infty} \sigma_n(k_{j_1})$ et $\lim_{n \rightarrow +\infty} \sigma_n(k_{j_2})$ sont de norme infinie.

Dans le système \mathcal{P} associé au PC P des objets du problème 6p7d3a2c, seul le bloc \mathcal{P}_0 peut admettre des solutions de norme infinie, car $\text{distance}(P_2, P_4) = h_8$ et $\text{distance}(P_0, P_5) = h_6$ appartiennent à l'ensemble de contraintes C du SCG G , donc $\sigma_n(k_8)$ et $\sigma_n(k_6)$ sont bornés et $\mathcal{P}_1, \mathcal{P}_2$ ne peuvent admettre de solution de norme infinie qui sont aussi solutions de G pour la valuation $d(t_n)$.

Cette analyse, qui dépend fortement du PC utilisé, si elle est effectuée avant le suivi du chemin permet soit d'écarter la possibilité de solution à l'infini, soit d'énumérer les configurations de la figure correspondant à une solution à l'infini : ici, les distances k_3, k_2 et k_{10} deviennent infiniment grandes, et P_3 est infiniment loin des autres points de la figure. Un plan de construction plus "proche" du SCG initial (obtenu par reparamétrisation), aurait pu permettre de détecter que les angles de paramètres a_1 et a_2 , s'ils sont plats pour la même valeur de t , impliquent cette solution à l'infini, car trois droites formant des angles plats et non confondues se coupent "à l'infini".

4.2.3 Générer de nouvelles esquisses

Dans cette sous-section, un plan de construction est utilisé pour générer de nouvelles esquisses permettant de définir de nouvelles fonctions d'homotopie et donc de nouveaux chemins, qui peuvent amener de nouvelles solutions.

On considère un SCG $G = C[X, A]$ dont on cherche les solutions pour la valuation $\sigma_{s_0}(A)$. En appliquant la méthode par continuation décrite dans le chapitre 3, on a obtenu un ensemble de solutions $\{\sigma_0(X), \dots, \sigma_{n-1}(X)\}$. Soit $P = I[X', A']$ un plan de construction pour X .

Sur chacune des solutions $\sigma_i(X)$, on peut lire la valuation $\sigma_i(A')$, et la branche b_i , telles que $P_{b_i}(\sigma_i(A')) = \sigma_i(X)$. L'évaluation de P sur toutes ses branches, pour la valuation $\sigma_i(A')$, fournit un ensemble de figures $\{\sigma_{i,0}(X), \dots, \sigma_{i,m-1}(X)\}$ (où, pour au moins un $0 \leq j \leq m-1$, $\sigma_{i,j}(X) = \sigma_i(X)$). Les figures $\sigma_{i,j}(X)$ avec $0 \leq i \leq n-1$, $0 \leq j \leq m-1$ et $\sigma_{i,j}(X) \neq \sigma_i(X)$ peuvent être utilisées comme esquisses ; pour chaque $\sigma_{i,j}(X)$, on peut définir une fonction d'homotopie comme

Inconnues :
point P_0, \dots, P_5
Paramètres :
longueur h_0, \dots, h_8
Contraintes :
distance $(P_0, P_1) = h_0$
...
distance $(P_2, P_4) = h_8$
coplanaires (P_0, P_1, P_2, P_3)
coplanaires (P_0, P_3, P_4, P_5)
coplanaires (P_1, P_2, P_4, P_5)

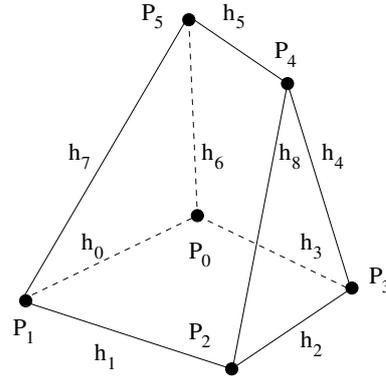


FIGURE 4.4 – À gauche, un énoncé formel du problème 2. À droite, une esquisse cotée : les arêtes représentent des contraintes de distance de paramètres h_i .

décrit dans le chapitre 3, et suivre son chemin d’homotopie contenant cette esquisse, pour obtenir de nouvelles solutions.

Bien sûr, selon le PC P considéré, les esquisses $\sigma_{i,j}(X)$ ne vérifient pas nécessairement les contraintes booléennes de G . Cependant, la méthode décrite dans le chapitre 5 pour éviter les points critiques permet d’utiliser des esquisses qui ne satisfont pas les contraintes booléennes d’un SCG. On se permet donc d’ignorer ce problème.

En évaluant P sur toutes ses branches pour $\sigma_i(A')$, on obtient un nombre exponentiel en le nombre d’objets géométriques de X d’esquisses, ce qui peut rendre cette phase très coûteuse. Cependant, comme en général un SCG admet un nombre exponentiel de solutions, produire toutes ces solutions a nécessairement un coût exponentiel (si le SCG est résoluble par LIM, un plan de construction de toutes les solutions est obtenu en un temps quadratique, mais ce plan de construction possède un nombre exponentiel de branches). On peut néanmoins ne considérer que des esquisses différentes à symétries près en fixant certains choix lors de l’évaluation de P .

Dans le cadre de la CAO, des esquisses à utiliser en priorité pourraient être choisies par l’utilisateur de manière interactive, par exemple en ayant recourt au gel de branche ([Essert-Villard 00]), ce qui rendrait cette méthode itérative plus supportable (en terme de temps de traitement) que la recherche d’emblée de toutes les solutions.

Les nouvelles solutions obtenues en suivant les chemins d’homotopie déterminés par les nouvelles esquisses sont utilisées pour ré-itérer ce processus, tant que de nouvelles solutions sont trouvées.

On illustre cette recherche de solutions sur le problème de la construction d’un pentaèdre (problème 2), dont on donne le SCG ainsi qu’une esquisse cotée dans la figure 4.4. Soit la valuation $\sigma_{so}(A)$ donnée dans la table 4.1. En suivant le chemin déterminé par l’esquisse donnée dans la partie droite de la figure 4.4, on obtient les solutions (1), (2), (3) et (4) de la figure 4.5. En évaluant le plan de construction P donné sur la figure 4.1 pour les paramètres lus sur la solution (1) sur toutes ses branches, on obtient les esquisses données dans la figure 4.6. (a) et (b) sont différentes à symétries près et seule (b) est différente de la solution (1). En suivant le chemin d’homotopie déterminé par cette dernière esquisse, on retrouve les mêmes solutions que celles initialement trouvées. Par contre, en évaluant le PC pour les paramètres lus sur la solution (4) et en suivant les chemins d’homotopie déterminés par les esquisses construites, on obtient les trois nouvelles solutions (5), (6), (7).

En réitérant le processus avec les esquisses obtenues à partir de (4), (5), (6), on ne trouve pas de nouvelles solutions. En fait, pour la valuation des paramètres considérée ici, le problème 2 n’admet

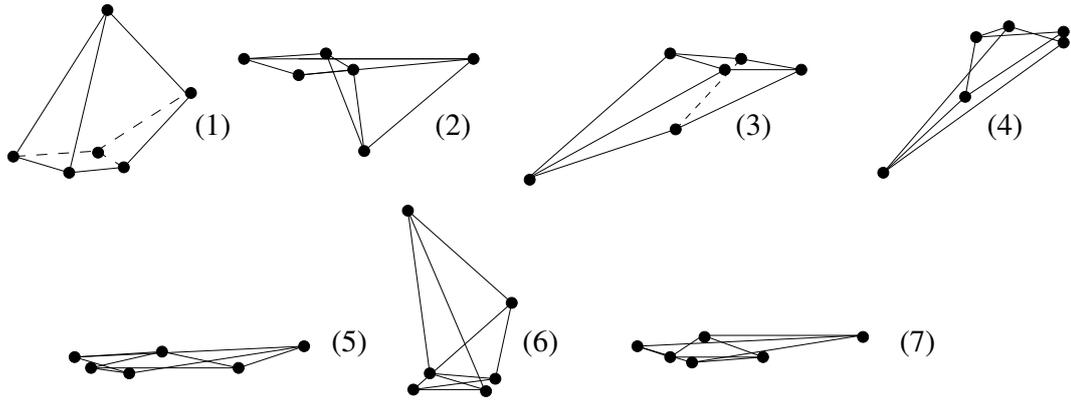


FIGURE 4.5 – Les 7 solutions différentes à symétries près du problème 2 pour la valuation des paramètres donnée à la table 4.1. Les points des solutions (5) et (7) appartiennent tous à un même plan.

| | | |
|-------------|-------------|-------------|
| $h_0 = 1.3$ | $h_1 = 0.7$ | $h_2 = 0.9$ |
| $h_3 = 0.8$ | $h_4 = 1.2$ | $h_5 = 1.5$ |
| $h_6 = 2.2$ | $h_7 = 2.3$ | $h_8 = 1.4$ |

TABLE 4.1 – Une valuation $\sigma_{so}(A)$ des paramètres du problème 2.

que ces 7 solutions qui sont donc toutes trouvées ici. Cependant, tout ce que l'on peut garantir ici, c'est qu'aucune autre solution ne peut être obtenue par cette méthode.

4.2.4 Conclusion

Implantation

Il a été choisi d'implanter des plans de constructions de haut niveau, utilisant des fonctions d'intersection de lieux des bibliothèques géométriques de la plate-forme de résolution de contraintes géométriques **Constraint** (développée en C++ dans l'équipe IGG du laboratoire ICube).

L'algorithme de reparamétrisation (voir algorithme 1) utilise une représentation d'un SCG par son graphe de contraintes et fournit comme résultat le graphe de contraintes du SCG reparamétré, ainsi que les listes de contraintes ajoutées et supprimées. La liste ordonnée des termes du plan de

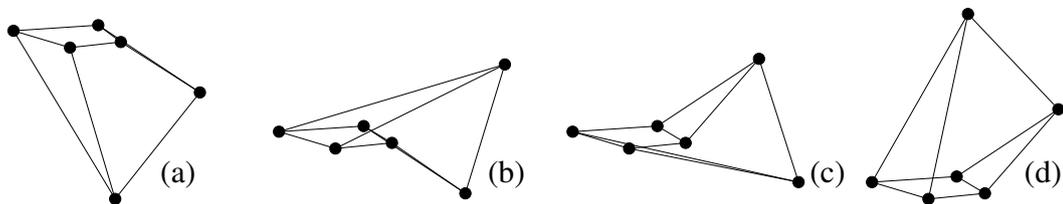


FIGURE 4.6 – Les 4 esquisses obtenues à partir de la solution (1) de la figure 4.5 en évaluant le plan de construction de la figure 4.1 sur toutes ses branches.

construction est également générée pendant le déroulement de l’algorithme. La génération de PC grâce à des constructions d’Henneberg n’a pas été implantée.

Pour leur évaluation, les termes sont interprétés numériquement l’un après l’autre avec les paramètres voulus, et les branches sont représentés par des vecteurs d’entiers.

De la qualité du plan de construction

On a présenté dans la sous-section 4.2.1 deux façons d’obtenir un plan de constructions également valides pour les deux utilisations qu’on en fait, à savoir obtenir de nouvelles esquisses et détecter des configurations de solutions à l’infini. Voici pour conclure quelques réflexions sur les aspects “qualitatifs” de ces plans de construction.

Tout d’abord pour ce qui est de l’obtention de nouvelles esquisses, on peut juger de la qualité d’un plan de construction par les nouvelles solutions obtenues en suivant les chemins d’homotopie auxquels appartiennent les esquisses qu’il a fournies. Notre intuition était qu’en utilisant un PC issu de la reparamétrisation plutôt qu’obtenu avec des constructions d’Henneberg, les esquisses obtenues étaient plus proches de solutions du SCG à résoudre, dans le sens où le nombre de contraintes non satisfaites par ces esquisses étaient peu nombreuses (le nombre de contraintes d’un SCG remplacées pour qu’il soit résoluble par LIM est en général bien plus petit que le nombre de contraintes totales).

Les expériences que nous avons effectuées ont démenti cette intuition. Dans l’exemple donné à la sous-section précédente, utiliser un PC issu de constructions d’Henneberg ou de reparamétrisation donne le même résultat en termes de nouvelles solutions obtenues. Pour le cas de la molécule de disulfide, un des exemples sur lesquels on présente les résultats dans le chapitre 3, un PC obtenu par constructions d’Henneberg est avantageux. Dans la sous-section 3.4 on évoque une valuation des paramètres pour laquelle le problème admet 18 solutions. En utilisant un PC obtenu par reparamétrisation, seulement 13 de ces solutions sont trouvées. Par contre, un PC issu de constructions d’Henneberg permet d’obtenir les 18 solutions en suivant les chemins définis par les esquisses qu’il construit.

Pour ce qui est de détecter les solutions à l’infini, un plan de construction issu de reparamétrisation semble fournir une analyse plus proche du SCG à résoudre des situations menant à des solutions à l’infini puisque la plupart de ses contraintes interviennent dans ce plan de construction et qu’il dépend de moins de paramètres ajoutés, qui ne sont pas nécessairement bornés.

4.3 Résolution par reparamétrisation

La résolution d’un système de contraintes géométriques $G = C[X, A]$ par reparamétrisation (voir [Gao 02, Fabre 07, Fabre 08]) se décompose en deux phases.

Lors de la première, dite *phase symbolique*, $G' = C'[X, A']$ résoluble par une méthode géométrique donnée est obtenu à partir de G en lui ajoutant m_+ contraintes métriques, ce qui est compensé par la suppression de m_+ contraintes du système initial. Un algorithme pour obtenir G' a été présenté dans la section 4.2 quand la méthode géométrique considérée est LIM. À l’issue de cette phase, on distingue les ensembles suivants :

- A^+ , l’ensemble des m_+ paramètres des contraintes métriques ajoutées, appelés par la suite paramètres guides.
- $C^-[X, A^-]$ l’ensemble des *contraintes supprimées*, que l’on notera G^- .
- $A' = (A \setminus A^-) \cup A^+$ l’ensemble des paramètres de G' .

On note P le plan de construction des solutions de G' , et $P_b(\sigma(A'))$ la figure résultant de l'évaluation de P sur sa branche b pour la valuation $\sigma(A')$

La seconde phase, dite *phase numérique* a pour but de trouver des valeurs pour les paramètres guides afin que les solutions fournies par le plan de construction de G' satisfassent les contraintes supprimées. Plus précisément, il s'agit de trouver des couples $(b, \sigma(A^+))$ tels que $P_b((\sigma(A^+), \sigma_{so}(A \setminus A^-)))$ soit solution de G , c.à.d satisfasse les contraintes de G^- .

Cette dernière phase est difficile à réaliser. [Gao 02] utilise une méthode d'échantillonnage de l'espace des paramètres guides pour chaque branche du plan de construction qui devient trop coûteuse quand le nombre de branches et la dimension de l'espace échantillonné augmente. [Fabre 08] applique la méthode itérative de Newton-Raphson à une fonction numérique évaluant le plan de construction pour une valeur des paramètres guides, et s'annulant quand la figure obtenue satisfait les contraintes de G^- . Cette application est difficile car la fonction P_b associée au plan de construction n'est pas définie partout : une itération fait souvent "sortir" le vecteur itéré de son domaine de définition. Une adaptation de la méthode de Newton-Raphson est proposée par [Fabre 08] pour limiter ce phénomène.

On propose ici d'adapter la méthode d'homotopie présentée dans le chapitre 3 au caractère multifonctionnel du plan de construction. La sous-section 4.3.1 présente l'idée d'une telle adaptation sur un exemple. On reste volontairement imprécis pour deux raisons. D'une part, la formalisation de cette idée est plus ardue à comprendre que ce dont il retourne. D'autre part, l'application pratique de cette adaptation n'est pas probante. La sous-section 4.3.2 présente un algorithme de suivi de chemin par approximation linéaire par morceaux permettant de passer d'une branche du PC à l'autre pendant le suivi de chemin. La sous-section 4.3.3 conclut. On termine cette partie en rappelant ([Fabre 06]) que la méthode par reparamétrisation est correcte, et complète si la phase numérique l'est.

Soit $\sigma(X)$ une figure résultant de l'évaluation du plan de construction associé à G' . $\sigma(X)$ satisfait toutes les contraintes de $C \setminus C^-$. Si $\sigma(X)$ vérifie par ailleurs les contraintes de C^- , elle satisfait toutes les contraintes de G . Supposons à présent que $\sigma(X)$ soit une solution de G . Comme les contraintes ajoutées sont métriques, on peut lire sur $\sigma(X)$ une valuation $\sigma(A^+)$ telle que $\sigma(X)$ soit une solution de G' , pour la valuation des paramètres $(\sigma(A^+), \sigma_{so}(A \setminus A^-))$. On peut en déduire qu'il existe une branche b telle que $\sigma(X) = P_b((\sigma(A^+), \sigma_{so}(A \setminus A^-)))$.

4.3.1 Résolution par continuation

On utilise comme exemple le problème 5 rappelé dans la figure 4.7(a). Pour obtenir un SCG G' résoluble par LIM, on peut par exemple remplacer sa contrainte $distance(P_3, P_5) = h_{11}$ par la contrainte $distance(P_0, P_3) = k_0$. On a alors $A^+ = \{k_0\}$, et $G^- = \{distance(P_3, P_5) = h_{11}\}$. Le plan de construction $P = I[X, A']$ des solutions de G' est représenté sur la figure 4.7(b). Les figures obtenues lors de l'évaluation de P dépendent des paramètres de $A \setminus A^-$, (ici $A \setminus \{h_{11}\}$) donnés par l'énoncé, et des paramètres guides A^+ , ici $\{k_0\}$. Pour faire la distinction entre ces deux types de paramètres, on notera $P(A^+, A \setminus A^-)$ l'évaluation de P , ou plus simplement $P(A^+, A)$. On associe à G^- la fonction numérique $F^-(X, A^-)$, ou $F^-(X, A)$. Dans l'exemple, l'unique composante de F^- est $\|\overrightarrow{P_3P_5}\|_2 - h_{11}$.

Pour une branche b de P , on définit la fonction numérique :

$$F_b : \begin{array}{ccc} \mathbb{R}^{m+} \times \mathbb{R}^m & \rightarrow & \mathbb{R}^{m+} \\ (A^+, A) & \mapsto & F^-(P_b(A^+, A), A) \end{array}$$

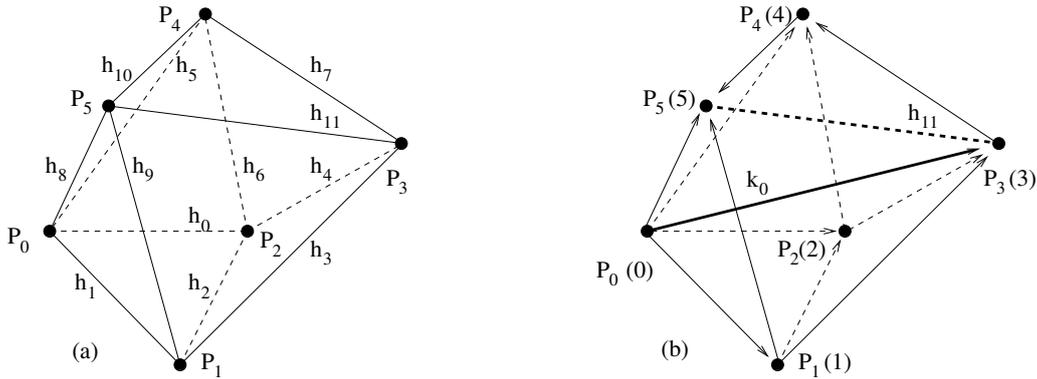


FIGURE 4.7 – (a) une esquisse cotée du problème 5. (b) une représentation du plan de construction du problème reparamétré. k_0 est le paramètre de la distance ajoutée, et h_{11} celui de la distance supprimée.

Soit $\sigma_{so}(A)$ la valuation des paramètres pour laquelle on veut résoudre G , $\sigma_{sk}(X)$ une esquisse, et d une fonction d'interpolation de $\sigma_{sk}(X)$ vers $\sigma_{so}(A)$. On définit alors les n fonctions d'homotopie $H_b(A^+, t) = F_b(A^+, d(t))$, où n est le nombre de branches de P .

Les fonctions d'homotopie $H_b(A^+, t)$ sont définies sur une variété de dimension $m_+ + 1$, dont le bord, pour simplifier, est le bord de \mathcal{D}_b , le domaine de définition de P_b ³. En utilisant le lemme 24 donné en annexe A, on peut affirmer que si 0 est une valeur régulière de H_b , ses chemins d'homotopie sont des variétés de dimension 1, dont le bord coïncide avec celui de \mathcal{D}_b .

La figure 4.7(a) donne une esquisse $\sigma_{sk}(X)$ pour le problème 5. Sur cette esquisse peuvent se lire les deux valuations $\sigma_{sk}(A)$ et $\sigma_{sk}(A^+)$ ($\sigma_{sk}(A^+)$ correspondant à l'ensemble formé par l'unique réel qui est la distance entre $\sigma_{sk}(P_0)$ et $\sigma_{sk}(P_3)$). Il existe alors une branche, disons b_0 telle que $(\sigma_{sk}(X), 0)$ appartienne à un chemin d'homotopie de H_{b_0} . Sur la figure 4.8, la courbe en gras représente ce chemin, et les deux points noirs à ses extrémités son bord. En le suivant depuis l'esquisse, on rencontre alors un point de son bord, qui est aussi un point du bord du domaine de définition de H_{b_0} , et donc de P_{b_0} .

On peut caractériser en terme de configuration géométrique un point de ce bord. Dans l'exemple considéré, les instructions du plan de construction sont toutes des intersections de trois sphères. Si un point $(\sigma(X), t)$ appartient au bord de P_{b_0} , c'est que pour une instruction c_i du PC, les trois sphères intersectées n'ont qu'une intersection (au lieu de deux). En effet, si elles en avaient deux, on pourrait trouver un ouvert U autour de $(\sigma(X), t)$ tel qu'en tout point de cet ouvert, ces trois sphères auraient encore deux intersections, donc P_{b_0} serait défini sur tout l'ouvert U (d'où la contradiction). Si les trois sphères n'avaient aucune intersection ou une infinité, P_{b_0} ne serait pas défini, et H_{b_0} non plus.

En un point $(\sigma(X), t)$ du bord d'un chemin \mathcal{S}_{b_0} , on peut trouver une nouvelle branche $b_1 \neq b_0$ telle qu'il existe un chemin de H_{b_1} , disons \mathcal{S}_{b_1} , qui "prolonge" de manière continue \mathcal{S}_{b_0} , c'est à dire tel que la figure géométrique correspondant aux points de $\mathcal{S}_{b_0} \cup \mathcal{S}_{b_1}$ soit déformée continûment le long du chemin, en inversant, dans la branche b_0 , le choix correspondant à l'instruction c_i . Sur la figure 4.8, \mathcal{S}_{b_1} est une des courbes jouxtant la courbe en gras (\mathcal{S}_{b_0}), selon le sens de suivi.

3. Dans la section 4.1, on a déjà introduit cette notion de domaine de définition. Ici \mathcal{D}_b est le domaine de définition de la fonction $(A^+, t) \mapsto P_b(A^+, t)$.

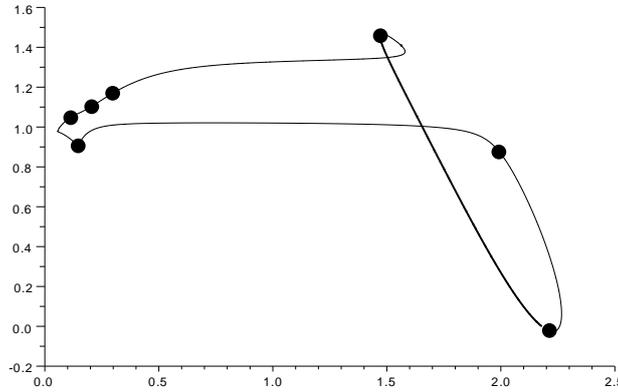


FIGURE 4.8 – Le chemin d’homotopie suivi lors de la phase numérique de la reparamétrisation du problème 5 par notre méthode de continuation. Les points noirs représentent les changements de branche. Sur l’axe des abscisses, la valeur du paramètre guide k_0 , en ordonnée la valeur du paramètre d’interpolation t .

Sur la partie gauche de la figure 4.9, on a représenté en gras le chemin $\mathcal{S}_{b_0} \cup \mathcal{S}_{b_1}$. Le point B représente son intersection avec le bord $d\mathcal{D}_0$ du domaine de définition \mathcal{D}_0 de H_{b_1} . À droite, la situation (b) représente les trois sphères intersectées dans l’instruction c_i . Le point A est un point de \mathcal{S}_{b_0} , et correspond à la situation (a). La nouvelle branche b_1 est choisie telle que le point C corresponde à la situation (c) pour les trois sphères de l’instruction c_i : c’est l’autre intersection de ces trois sphères qui est choisie dans b_1 .

En fait, en suivant le chemin d’homotopie depuis l’esquisse, on est amené à suivre la suite de chemins $(\mathcal{S}_j)_{j \in \mathbb{N}}$, en changeant de branche à chaque fois qu’on arrive sur un bord d’un chemin. On peut montrer qu’il existe un indice j_0 dans cette suite tel que $\mathcal{S}_{j_0} = \mathcal{S}_0$. La figure 4.8 présente cette suite de chemins $(\mathcal{S}_j)_{j=0}^{j_0}$, délimités par les points noirs.

L’idée de la preuve est la suivante : on considère la fonction d’homotopie définie sur $\mathbb{R}^m \times \mathbb{R}^{m_+} \times \mathbb{R}$, dont les m premières composantes correspondent à la fonction d’homotopie associée à G' (le problème reparamétré), et les m_+ dernières correspondent aux contraintes de G^- . On peut alors appliquer la proposition 9 pour établir que ses chemins d’homotopie sont difféomorphes à des cercles, ou à des intervalles ouverts. La réunion des chemins de la suite $(\mathcal{S}_j)_{j \in \mathbb{N}}$ peut alors être vue comme la projection sur $\mathbb{R}^{m_+} \times \mathbb{R}$ (où \mathbb{R}^{m_+} est l’espace des paramètres guides) des chemins d’homotopie de cette fonction.

4.3.2 Suivi de chemin

Le principal problème posé par la mise en œuvre de cette méthode réside dans le suivi de la suite de chemins $(\mathcal{S}_j)_{j=0}^{j_0}$, et plus précisément dans la gestion du changement de branche. À proximité du bord d’un chemin \mathcal{S}_j , une prédiction, ou une correction par la méthode de Newton-Raphson peut résulter en l’obtention d’un point hors du domaine de définition de la fonction d’homotopie H_j , et le suivi échoue. On propose ici d’adapter une méthode de suivi de chemin par approximation linéaire par morceau (PLI), qui construit une suite de simplexe de $\mathbb{R}^{m_+} \times \mathbb{R}$ contenant⁴ la courbe

4. Pour être plus précis, la suite de simplexes contient une approximation linéaire de la courbe suivie.

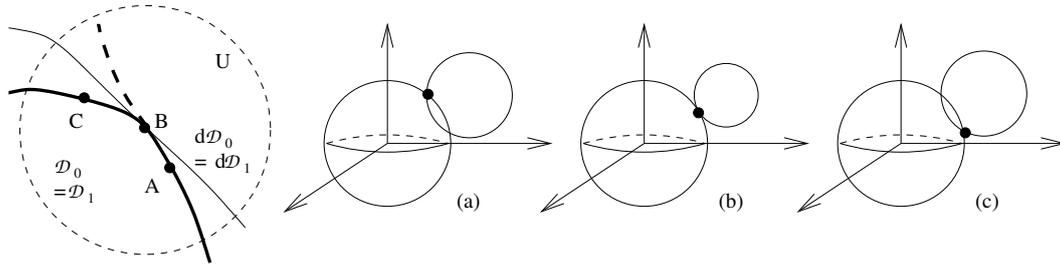


FIGURE 4.9 – À droite, la configuration, dans l'espace des paramètres, d'un changement de branche. À gauche, les situations correspondantes dans l'espace des figures ; en trait gras, le chemin suivi.

suivie. Pour résumer simplement cette adaptation, quand un simplexe a un de ses sommets hors du domaine de définition de H_j , le bord de H_j dans ce simplexe est approximé par un hyperplan de $\mathbb{R}^{m_+} \times \mathbb{R}$. La nouvelle branche $j + 1$ pour laquelle le chemin d'homotopie doit être suivi est alors déterminée comme indiqué dans la sous-section précédente. Une heuristique, qu'on présente plus bas, amène à déplacer le sommet de ce simplexe dans le domaine de définition de H_{j+1} en prenant son symétrique par rapport à l'hyperplan approximant le bord du domaine de définition de H_j . Cette approximation est actualisée à chaque nouveau simplexe calculé jusqu'à ce que tous les sommets du simplexe soient dans le domaine de définition de H_{j+1} .

L'heuristique évoquée plus haut consiste à considérer que pour chaque point x du bord du domaine de définition \mathcal{D}_j de H_j , on peut trouver un ouvert U tel que :

- $U \cap \mathcal{D}_j = U \cap \mathcal{D}_{j+1}$ donc $U \cap d\mathcal{D}_j = U \cap d\mathcal{D}_{j+1}$ (\mathcal{D}_{j+1} est le domaine de définition de H_{j+1}) ;
- dans U , le bord $d\mathcal{D}_j$ consiste en un hyperplan de $\mathbb{R}^{m_+} \times \mathbb{R}$.

La partie gauche de la figure 4.9 présente, autour du point B , un tel ouvert U . Cette heuristique est justifiée de la manière suivante : quand trois sphères sont "presque" tangentes, les figures construites en choisissant l'une ou l'autre des intersections de ces sphères sont "presque" symétriques. Dans la figure 4.9, la branche b_0 correspond au choix du point A , et la branche b_1 correspond au choix du point C .

Intuitivement, l'algorithme qu'on présente ci-dessous "déplie" les domaines de définition \mathcal{D}_j et \mathcal{D}_{j+1} (qui d'après l'heuristique se superposent) autour de l'axe de symétrie $d\mathcal{D}_j$: quand un sommet d'un simplexe est hors de \mathcal{D}_j , on considère qu'il est dans le symétrique, par rapport à $d\mathcal{D}_j$, de \mathcal{D}_{j+1} .

Adaptation de PLI

On rappelle brièvement le principe du suivi de chemin par approximation linéaire par morceau. Un simplexe en dimension $m_+ + 1$ est l'enveloppe convexe de $m_+ + 2$ vecteurs $v_0, \dots, v_{m_+ + 1}$, qui forment un repère affine de $\mathbb{R}^{m_+ + 1}$ (c.à.d. la famille $\langle (v_1 - v_0), \dots, (v_{m_+ + 1} - v_0) \rangle$ est une base de $\mathbb{R}^{m_+ + 1}$). H_b est approximé sur ce simplexe par l'unique application affine H'_b telle que $H'_b(v_i) = H_b(v_i)$, pour $0 \leq i \leq m_+ + 1$. Si $H'_b(v_i)$ est de rang plein (donc m_+)⁵, son noyau est un espace affine de dimension 1 (donc une droite) dans $\mathbb{R}^{m_+ + 1}$. Supposons à présent un point x tel que $H'_b(x) = 0$ appartenant à une face du simplexe soit connu. Le noyau de H'_b contient entre autre x , et croise une autre face du simplexe. On peut alors redéfinir un nouveau simplexe, partageant cette face avec le simplexe précédent, et construire une nouvelle approximation affine de H_b sur ce simplexe.

Lors des itérations de cette méthode, on détecte que \mathcal{S}_{b_0} , le chemin suivi, s'approche du bord du domaine de définition \mathcal{D}_{b_0} quand H_{b_0} n'est pas définie en un sommet, disons v_0 , du simplexe

5. Cette condition est une conséquence du fait que 0 soit une valeur régulière de H_b .

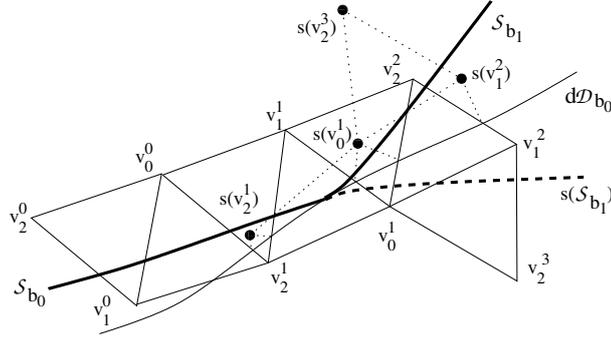


FIGURE 4.10 – Le suivi par PLI d’une courbe $\mathcal{S}_{b_0} \cup \mathcal{S}_{b_1}$ passant d’une branche du plan de construction à une autre.

courant. La nouvelle branche b_1 permettant de prolonger \mathcal{S}_{b_0} par \mathcal{S}_{b_1} est déterminée comme décrit plus haut. Il est cependant impossible de calculer H'_{b_0} sur ce simplexe, puisque $H_{b_0}(v_0)$ n’existe pas. On fait alors intervenir notre heuristique en considérant que la fonction H_{b_0} est prolongée au delà de $d\mathcal{D}_0$ par la fonction $H(x) = H_{b_1}(s(x))$ définie sur $s(\mathcal{D}_0)$, où s désigne le symétrique d’un point (ou d’un ensemble), par rapport à l’hyperplan approximant $d\mathcal{D}_0$. On utilise alors, pour calculer H'_{b_0} , $H(v_0)$ à la place de $H_{b_0}(v_0)$, ce qui revient à suivre $s(\mathcal{S}_{b_1})$ à la place de \mathcal{S}_{b_1} . Sur la partie gauche de la figure 4.9 $s(\mathcal{S}_{b_1})$ est présenté en trait gras pointillé.

Lorsque tous les sommets du simplexe courant sont en dehors de \mathcal{D}_0 , on calcule les symétriques de ces sommets par rapport au bord pour construire un nouveau simplexe en lequel H_{b_1} est complètement définie et continuer à suivre la courbe \mathcal{S}_{b_1} définie par $H_{b_1} = 0$. Le processus de changement de branche est alors un succès.

À chaque étape, l’hyperplan \mathcal{P} approximant $d\mathcal{D}_0$ est calculé en considérant les arêtes $v_i v_j$ du simplexe telles que $v_i \in \mathcal{D}_0$ et $v_j \notin \mathcal{D}_0$. Il y en a au moins $m_+ + 1$ si au moins un des sommets est en dehors de \mathcal{D}_0 . On recherche alors par dichotomie pour $m_+ + 1$ de ces arêtes un point aussi proche que souhaité de $d\mathcal{D}_0$. Les $m_+ + 1$ points ainsi obtenus définissent l’hyperplan \mathcal{P} approximant $d\mathcal{D}_0$.

La figure 4.10 illustre cette procédure dans le cas d’une courbe \mathcal{S}_{b_0} de \mathbb{R}^2 ($m_+ = 1$). \mathcal{S}_{b_0} sort du simplexe de sommets v_0^0, v_1^0, v_2^0 par la face opposée au sommet v_2^0 , et le nouveau sommet v_2^1 calculé est hors du domaine de définition \mathcal{D}_0 de H_{b_0} . L’hyperplan \mathcal{P} approximant $d\mathcal{D}_0$ est alors calculé grâce aux arêtes $v_0^0 v_2^1$ et $v_1^0 v_2^1$, et H_{b_0} est approximé par l’application affine valant $H_{b_0}(v_1^0)$ en v_1^0 , $H_{b_0}(v_0^0)$ en v_0^0 et $H_{b_1}(s(v_2^1))$ en v_2^1 , où $s(v_2^1)$ est le symétrique de v_2^1 par rapport à \mathcal{P} . Le noyau de cette application sort du simplexe v_0^0, v_1^0, v_2^1 par la face opposée au sommet v_1^0 . À l’itération suivante, le noyau approximant la courbe suivie sort du simplexe v_0^0, v_1^1, v_2^1 opposée à v_0^0 , et le nouveau sommet v_0^1 n’est pas dans \mathcal{D}_0 . On continue ainsi jusqu’à ce qu’un simplexe ait tous ses sommets en dehors de \mathcal{D}_0 . Alors le nouveau simplexe est $s(v_0^1), s(v_1^1), s(v_2^1)$ et la fonction considérée est H_{b_1} .

4.3.3 Résultats et conclusion

La méthode de suivi de courbe PLI présente l’avantage d’être facile à implanter. Un premier simplexe est calculé de façon à ce que le point initialement connu sur la courbe suivie se trouve sur une de ses faces. L’approximation affine de la fonction d’homotopie H est représentée par la matrice dont chaque vecteur colonne est l’image par H d’un sommet du simplexe, et calculer son noyau revient à inverser cette matrice. Le calcul des intersections de ce noyau avec les faces du

simplexe est effectué en coordonnées barycentriques, et comme à chaque étape une seule colonne de la matrice représentant l'approximation affine de H change, l'inverse de la matrice est actualisée en un temps quadratique (en $m_+ + 1$) grâce, par exemple, à la méthode décrite dans [Hager 89]. En supposant que l'évaluation du plan de construction sur une de ses branches ait un coût linéaire en m (où m est le nombre des objets géométriques qu'il construit), chaque itération a un coût de l'ordre de $\max(m, (m_+ + 1)^2)$.

Le principal inconvénient de la méthode PLI est que le nombre de simplexes traversés par une courbe en dimension d peut augmenter de façon factorielle en d ; en effet, le volume d'un simplexe régulier en dimension d est $\frac{a^d}{d!} \sqrt{\frac{d+1}{2^d}}$ où a est la longueur d'une arête du simplexe. Il faut donc $d! \sqrt{\frac{2^d}{d+1}}$ simplexes d'arêtes de longueurs 1 pour remplir un pavé de volume 1. Une courbe qui "traverse" un pavé de volume 1 peut alors potentiellement traverser $d! \sqrt{\frac{2^d}{d+1}}$ pavés; son suivi peut donc nécessiter $d! \sqrt{\frac{2^d}{d+1}}$ itérations de la méthode PLI, pour des simplexes d'arêtes de longueurs 1. On suppose cependant ici que le nombre de paramètres guides nécessaires pour reparamétriser un problème est bien inférieur au nombre d'équations et d'inconnues de ce problème. Le problème 5 utilisé tout au long de cette section le montre : il a 12 contraintes, et sa reparamétrisation nécessite l'ajout d'une seule contrainte; le chemin d'homotopie est suivi dans un espace de dimension 2, contre 13 dans le cas de la méthode présentée dans le chapitre 3. Dans le cas du problème de l'icosaèdre, (voir 3.4), qui se formalise en un SCG avec 30 contraintes, seulement 3 paramètres guides sont utilisés, et le chemin d'homotopie est suivi dans un espace de dimension 4 (contre 31 sans reparamétriser le problème).

Malgré les arguments présentés ici en faveur de la résolution par homotopie de la phase numérique de la reparamétrisation, les résultats des tests menés, en terme d'efficacité et de robustesse, ne sont pas probants :

1. Pour un même problème, le suivi du chemin d'homotopie est en général plus long quand il est défini par la fonction d'homotopie de sa reparamétrisation que quand il est défini par la fonction d'homotopie du problème lui-même.
2. Dans le cas de certains problèmes, le suivi de la courbe définie par la fonction d'homotopie de leur reparamétrisation échoue quelle que soit la taille des simplexes choisis.

On cite en exemple, pour illustrer le point 1, le cas du problème 5. Le chemin d'homotopie défini par sa reparamétrisation est donné dans la figure 4.8. Son suivi nécessite environ 24 secondes, et 70000 itérations de la méthode PLI. Le suivi du chemin d'homotopie du même problème non reparamétrisé nécessite environ 0.1 secondes sur la même machine, et 115 itérations de la méthode de prédiction-corrrection proposée dans le chapitre 3. On peut avancer les arguments suivants pour justifier ces constatations.

Tout d'abord, l'évaluation d'un plan de construction entraîne des erreurs numériques, d'autant plus importantes que ce plan de construction contient d'instructions d'intersections. En effet, celles-ci sont évaluées numériquement l'une après l'autre, en utilisant les résultats des précédentes. Or le processus de changement de branche nécessite une bonne précision lors de l'évaluation du plan de construction pour être robuste; ceci amène à devoir choisir des simplexes de très petite taille pour obtenir la précision nécessaire. Par exemple, pour le suivi du chemin d'homotopie dans le cas du problème 5, on est amené à choisir des simplexes dont l'arête a une longueur d'environ 10^{-4} , quand le suivi du chemin sans reparamétrisation par une méthode de prédiction-corrrection réussit avec un pas de prédiction de 0.1.

D'autre part, notre heuristique assimilant le bord du domaine de définition de H_b dans un ouvert centré en un point de ce bord à un hyperplan peut ne pas refléter la réalité, ou la refléter

dans un ouvert si petit que la taille d'un simplexe dans lequel elle est valable est en deçà de la précision machine.

Comme cela a déjà été dit plus haut, l'utilisation d'une méthode de suivi de courbe par prédiction-correction est rendue difficile par le fait que le plan de construction n'est défini que pour certaines valeurs des paramètres : les points prédits ou le résultat d'une itération de la méthode de Newton peuvent être en dehors du domaine de définition. Dans nos essais de résolution de la reparamétrisation avec cette méthode, un mécanisme de surveillance des objets géométriques pendant le suivi de la courbe a été mis en place, pour détecter quand celle-ci s'approche du bord du domaine de définition du PC et ainsi adapter le pas de prédiction à la situation géométrique. Le problème qui n'a pas été surmonté est de trouver un point de la courbe suffisamment proche de son intersection avec le bord du domaine de définition pour être en mesure de suivre la courbe sur la nouvelle branche. C'est ce qui a mené à la mise en place de l'algorithme décrit plus haut.

Finalement, une piste que nous n'avons pas explorée est l'utilisation d'un plan de construction de très bas niveau représenté par un graphe orienté sans cycle (un "DAG") dont les sommets correspondent aux opérations élémentaires $+$, $-$, $*$, \div et $\sqrt{\quad}$ et dont le parcours de la racine à une feuille sur une branche fournit une expression symbolique des contraintes supprimées en fonction des paramètres guides et du paramètre d'interpolation. En supposant que les changements de branches se produisent exactement quand, dans cette expression, un radicande s'annule, cette représentation pourrait permettre de fournir une caractérisation symbolique du bord du domaine de définition ainsi que de l'intersection de la courbe suivie avec ce bord, dans le but de "guider" la résolution numérique par prédiction-correction. Les deux limitations que nous entrevoyons à cette approche sont d'une part la stabilité numérique lors de l'évaluation d'un empilement de racines et d'autre part la taille des expressions symboliques des contraintes supprimées⁶.

4.4 Conclusion

Ce chapitre a abordé trois des problématiques posées par l'application d'une méthode d'homotopie à la résolution de contraintes géométriques qui déforme l'esquisse en des solutions et on a proposé d'y répondre grâce à des plans de constructions. Quand un SCG n'est pas résoluble par une méthode géométrique comme LIM, un plan de construction de ses solutions n'est pas disponible. Deux façons d'obtenir un plan de construction produisant des figures d'un SCG qui satisfont certaines de ses contraintes ont été proposées.

Les figures obtenues par ce plan de construction peuvent être utilisées pour construire de nouvelles esquisses. Ces nouvelles esquisses permettent de construire de nouvelles fonctions d'homotopie dont suivre les chemins. Bien sûr, ces chemins ne sont pas d'autres chemins de la fonction d'homotopie construite à partir de l'esquisse, et il n'est pas possible d'assurer qu'ils permettent d'obtenir de nouvelles solutions. Cependant, dans les exemples présentés, de nouvelles solutions sont trouvées par cette méthode, et parfois elles le sont toutes. Cette phase a un coût très élevé, exponentiel en le nombre d'inconnues du SCG. Il est à mettre en regard avec le coût des méthodes d'homotopie classiques qui les rend inapplicables à de gros problèmes, et avec le nombre exponentiel de solutions que peut admettre un SCG. La nature itérative de cette méthode, et le fait qu'on puisse imaginer la faire guider par un utilisateur (par exemple lui faire sélectionner des esquisses qui lui conviennent) nous semble être un atout dans le cadre d'un système de CAO.

Le plan de construction peut également être utilisé pour détecter qu'un chemin tend vers une solution à l'infini. On a montré qu'une telle solution à l'infini correspondait à une configuration

6. Par exemple, l'expression de la cote de l'intersection de trois sphères, dont une est centrée en l'origine, une sur l'axe des abscisses et la dernière sur le plan de cote nulle, en fonction des 3 rayons de ces sphères et des 3 réels nécessaires pour placer les trois centres, est un polynôme de 31 termes.

géométrique, et que l'analyse du plan de construction permettait d'énumérer ces configurations avant le suivi du chemin. Dans les méthodes par homotopie classiques, des chemins menant à des solutions à l'infini ne peuvent pas contenir de solution et peuvent être détectés avant leur suivi. Cette détection a priori ne semble pas possible ici, et les chemins de longueur infinie sont suivis jusqu'à ce qu'une configuration particulière soit détectée. Leur suivi n'est cependant pas vain, puisqu'ils peuvent contenir des solutions du SCG.

On a enfin proposé d'utiliser un plan de construction obtenu par reparamétrisation d'un SCG pour diminuer la dimension de l'espace dans lequel le chemin d'homotopie doit être suivi, et réduire la complexité de ce suivi. Cette méthode n'est cependant pas satisfaisante de par la nature multifonctionnelle du plan de construction, qui rend le suivi de chemin complexe à mettre en œuvre, peu robuste aux imprécisions, et en somme très coûteux. Une méthode de décomposition sera proposée, qui semble plus à même de maîtriser le coût de la résolution par homotopie d'un SCG.

Chapitre 5

Points critiques et incidences

La méthode par homotopie présentée dans le chapitre 3 déforme continûment l'esquisse d'un Système de Contraintes Géométriques (SCG) en ses solutions. Chacune des figures obtenues pendant cette déformation est une solution pour une certaine valuation des paramètres. Dans l'espace des coordonnées des objets géométriques du problème, ces figures correspondent toutes à un point d'un chemin d'homotopie.

Lors de sa déformation, en suivant le chemin d'homotopie, une figure peut présenter des configurations spéciales, souvent appelées *dégénérées*. En Conception Assistée par Ordinateur (CAO), elle sont considérées comme assez rares pour être ignorées. Certaines d'entre elles correspondent à des solutions à l'infini, ou à des figures en lesquelles les contraintes ne sont pas définies, et ces cas ont été traités dans le chapitre précédent.

D'autres correspondent à des points en lesquels les contraintes du problème ne satisfont pas certaines conditions de régularité. Ces points sont appelés points critiques si l'application linéaire dérivée de la fonction d'homotopie en ce point n'est pas de rang plein. Ils peuvent correspondre à des "croisements" de chemins, ou à l'intersection du chemin suivi avec des ensembles de solutions de dimensions positives (des nappes, des volumes,...).

Ce chapitre montre qu'ils apparaissent dans le cadre de la méthode exposée quand le système de contraintes contient des contraintes booléennes ; ils ne peuvent pas être évités si chaque figure du chemin d'homotopie satisfait ces contraintes.

On présente dans la section 5.1 deux exemples de SCG contenant des contraintes booléennes, pour lesquels des points critiques sont sur le chemin d'homotopie. Ils correspondent à des configurations géométriques particulières.

La section 5.2 est le fruit d'une analyse très pragmatique de ce problème : puisque satisfaire les contraintes booléennes d'un problème implique de passer par des points critiques, celles-ci sont transformées en contraintes métriques. Les solutions cherchées doivent toutefois satisfaire ces contraintes booléennes, et des points critiques solutions du SCG à résoudre sont par conséquent trouvés. Notons qu'en utilisant des méthodes d'homotopie classiques, ces points critiques solutions sont ignorés.

Les points critiques solutions sont des éléments de continus de solutions. Rencontrer un tel continuum est problématique pour des méthodes destinées à approximer des points isolés. La section 5.3 expose une méthode pour détecter ces continus de solutions, quand ils sont conséquences de théorèmes de géométrie d'incidence. Une structure de matroïde est utilisée pour effectuer cette détection. Les travaux présentés dans cette section reflètent nos recherches actuelles, et ne sont pas complètement aboutis. Ils apparaissent ici surtout comme des perspectives.

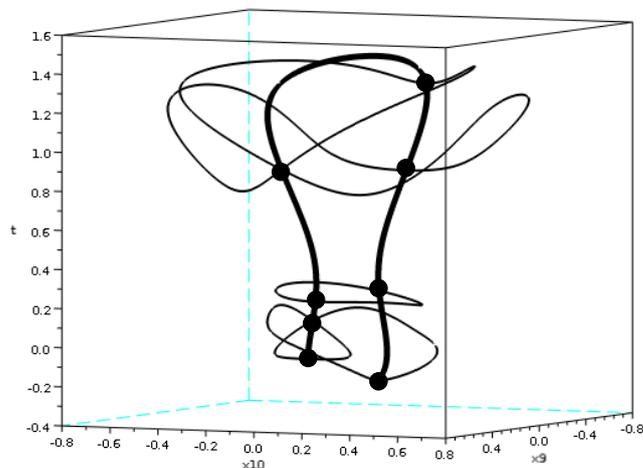


FIGURE 5.1 – Une projection sur \mathbb{R}^3 d’une composante connexe de $H^{-1}(0)$ pour le problème du pentaèdre (problème 2). Les cercles noirs représentent les points critiques de cette composante.

5.1 Étude empirique

On propose dans cette section d’étudier la résolution par la méthode homotopique proposée dans le chapitre 3 de deux problèmes impliquant des contraintes booléennes. Dans les deux cas, les chemins suivis depuis l’esquisse contiennent des points critiques, et ce pour presque toutes les valeurs des paramètres pour lesquels on cherche des solutions, et presque toutes les esquisses. La figure 5.1 présente un tel chemin : son exploration par une méthode classique de suivi de chemins est rendue difficile par le fait que la courbe présente des embranchements, compromettant ainsi la terminaison de la méthode. On verra que ces points critiques sont dus aux contraintes booléennes. Le second exemple abordé, structurellement bien contraint au sens de König-Hall, admet néanmoins, en plus de solutions isolées, des ensembles de solutions de dimensions positives, et certains des points critiques sont des éléments de ces ensembles de solutions. Quand le problème est embarqué dans une résolution par homotopie, ces ensembles de solutions deviennent des ensembles de dimension deux ou plus, et la matrice jacobienne de la fonction d’homotopie en les points de ces ensembles n’est pas inversible ; suivre un tel chemin traversant de tels ensembles échoue dès qu’un des points prédits ou corrigés appartient à cet ensemble. On rappelle brièvement quelques-unes des notations utilisées par la suite.

Les contraintes considérées ici (distances, coplanarités) étant invariantes par les déplacements, on cherche les solutions d’un SCG $G = C[X, A]$ en fixant un repère, c’est à dire en fixant 6 coordonnées d’objets de X . Si déterminer X revient à déterminer m coordonnées libres (non fixées par le repère choisi), et si le problème considéré est bien contraint au sens de König-Hall, C contient m contraintes. Si, parmi ces m contraintes, m' sont métriques, on construit une fonction numérique $F : \mathbb{R}^m \times \mathbb{R}^{m'} \rightarrow \mathbb{R}^m$, dépendant des m coordonnées libres de X et des m' paramètres de A . On note $F(X, A)$ cette fonction, ou $F(X, \sigma(A))$ quand F est “spécialisée” à une valuation de A , ou $F(\sigma(X), \sigma(A))$ la valeur qu’elle prend quand elle est évaluée en $(\sigma(X), \sigma(A))$.

Si $\sigma_{sk}(A)$ est la valuation de A lue sur une esquisse, $\sigma_{so}(A)$ celle pour laquelle on cherche les

Inconnues :
point P_0, \dots, P_5
Paramètres :
longueur h_0, \dots, h_8
Contraintes :
distance(P_0, P_1) = h_0
...
distance(P_2, P_4) = h_8
coplanaires(P_0, P_1, P_2, P_3)
coplanaires(P_0, P_3, P_4, P_5)
coplanaires(P_1, P_2, P_4, P_5)

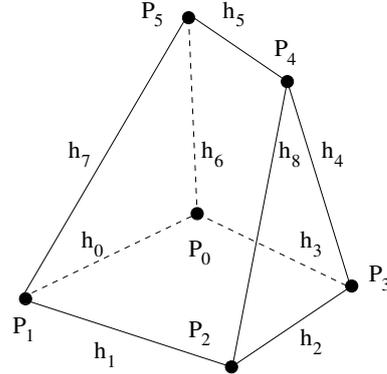


FIGURE 5.2 – À gauche, un énoncé formel du problème 2. À droite, une esquisse cotée : les arêtes représentent des contraintes de distance de paramètres h_i .

solutions de G , on définit la fonction d'interpolation $d : \mathbb{R} \rightarrow \mathbb{R}^{m'}$ telle que $d(0) = \sigma_{sk}(A)$ et $d(1) = \sigma_{so}(A)$, et d est C^∞ .

Si d est une fonction d'interpolation de $\sigma_{sk}(A)$ vers $\sigma_{so}(A)$, on définit la fonction d'homotopie $H : \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m$ par $H(X, t) = F(X, d(t))$, et on suppose que H est C^∞ sur un ouvert \mathcal{D} de $\mathbb{R}^m \times \mathbb{R}$. Les points critiques de H sont les points $(\sigma_0(X), t_0) \in \mathcal{D}$ tels que $H(\sigma_0(X), t_0) = 0$ et la jacobienne de H en $(\sigma_0(X), t_0)$ n'est pas de rang m . Si H admet un ou plusieurs points critiques, on dira que 0 est une valeur critique de H , sinon que c'est une valeur régulière de H . Rappelons que si 0 est une valeur régulière de H , les composantes connexes de $H^{-1}(0)$ sont des variétés différentielles de dimension 1, et sont par conséquent difféomorphes à un cercle ou un intervalle ouvert de \mathbb{R} .

5.1.1 Pentaèdre

On considère dans cette sous-section le problème de la construction d'un pentaèdre, appelé problème 2 dans ce mémoire; le SCG $G = C[X, A]$ qui lui est associé, ainsi qu'une esquisse cotée $\sigma_{sk}(X)$ sont donnés à la figure 5.2. Nous allons constater que pour presque tous les couples $(\sigma_{so}(A), \sigma_{sk}(X))$ de valuations des paramètres, la composante connexe \mathcal{S} de $H^{-1}(0)$ à laquelle appartient l'esquisse contient des points critiques.

Pour chercher les solutions de G modulo les déplacements, on les cherche dans le repère P_0, P_1, P_2 , c'est-à-dire telles que $\sigma(P_0) = (0, 0, 0)$, $\sigma(P_1) = (., 0, 0)$, $\sigma(P_2) = (., ., 0)$. Dans ce repère, la contrainte *coplanaires*(P_0, P_1, P_2, P_3) peut être automatiquement satisfaite en fixant la dernière coordonnée de P_3 à 0. On associe donc à G une fonction numérique $F : \mathbb{R}^{11} \times \mathbb{R}^9 \rightarrow \mathbb{R}^{11}$, et on construit la fonction d'homotopie $H : \mathbb{R}^{11} \times \mathbb{R} \rightarrow \mathbb{R}^{11}$.

Soit $G_2 = C_2[X_2, A_2]$ le SCG donné à la figure 5.3, dont les solutions $\sigma(\{P_0, \dots, P_5\})$ sont 6 points appartenant à un même plan de l'espace. On a $X_2 = X$ et $A_2 = A$ et on remarque les faits suivants :

- (1) L'ensemble des solutions de G_2 pour la valuation $\sigma(A_2)$ est inclus dans l'ensemble des solutions de G pour $\sigma(A)$. En effet, si $\sigma(\{P_0, \dots, P_5\})$ sont coplanaires, les trois contraintes de coplanarité de G sont respectées. Si de plus, ces points satisfont les contraintes de distance de G_2 , ils satisfont celles de G (qui sont les mêmes).
- (2) G_2 peut être vu comme un SCG G'_2 d'un univers 2D, n'impliquant que des points et des distances. G'_2 est bien contraint au sens de Laman, donc génériquement bien contraint : pour

Inconnues :
point P_0, \dots, P_5
Paramètres :
longueur h_0, \dots, h_8
Contraintes :
distance $(P_0, P_1) = h_0$
...
distance $(P_2, P_4) = h_8$
coplanaires (P_0, P_1, P_2, P_3)
coplanaires (P_0, P_1, P_2, P_4)
coplanaires (P_0, P_1, P_2, P_5)

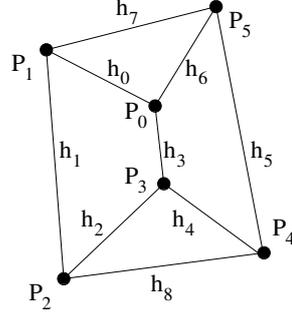


FIGURE 5.3 – Un énoncé et une esquisse cotée d’un SCG en 3D, dont tous les points sont coplanaires.

presque toute valuation $\sigma(A)$, G'_2 admet un nombre fini non nul de solutions isolées. Comme on peut établir une bijection entre les solutions de G'_2 et celles de G_2 , G_2 est génériquement bien contraint et, pour presque toute valuation $\sigma(A)$, admet un nombre fini non nul de solutions isolées.

- (3) En posant F_2 la fonction numérique associée à G_2 , et $H_2(X, t) = F_2(X, d(t))$ où d est la même fonction d’interpolation que celle de H , on a $H_2^{-1}(0) \subseteq H^{-1}(0)$.
- (4) Si on suppose que la fonction d’homotopie d’un SCG génériquement bien contraint n’impliquant que des contraintes métriques admet 0 comme valeur régulière pour presque toute esquisse et presque toutes valuations des paramètres cibles, les composantes connexes de $H_2^{-1}(0)$ sont difféomorphes à des cercles ou à des intervalles ouverts. On utilise le même raisonnement que dans le point (2). G'_2 est génériquement bien contraint et n’implique que des contraintes de distances. En utilisant la bijection entre les solutions de G_2 et celles de G'_2 déjà évoquée, on peut facilement plonger les chemins d’homotopie de G'_2 dans \mathbb{R}^{12} , l’espace des chemins d’homotopie de G_2 .

On suppose à présent que $\sigma_{sk}(X)$, l’esquisse de G , n’est pas *aplatie* (c.à.d n’est pas solution de G_2). On note \mathcal{S} le chemin d’homotopie de H auquel appartient $\sigma_{sk}(X)$. Si il existe un point $(\sigma_2(X), t_2) \in \mathcal{S}$ tel que la figure $\sigma_2(X)$ soit aplatie (c.à.d $\sigma_2(X)$ est solution de G_2), alors $H_2(\sigma_2(X), t_2) = 0$, et $(\sigma_2(X), t_2)$ appartient à une composante connexe \mathcal{S}_2 de $H_2^{-1}(0)$. D’après (3), on a $\mathcal{S}_2 \subseteq \mathcal{S}$, et comme $\sigma_{sk}(X)$ n’est pas solution de G_2 , on a $(\sigma_{sk}(X), 0) \notin H_2^{-1}(0)$ et $\mathcal{S}_2 \subset \mathcal{S}$ (l’inclusion est stricte). Il s’en suit, d’après (4), que \mathcal{S} n’est pas une variété de dimension 1. La conclusion de la proposition rappelée en préambule (voir proposition 7 du chapitre 3) est donc fautive, et l’hypothèse de régularité de 0 pour H n’est pas satisfaite. $(\sigma_2(X), t_2)$ est un point critique pour la valeur 0.

La figure 5.1 présente une projection sur \mathbb{R}^3 de la composante connexe $\mathcal{S} \subset H^{-1}(0)$; cette projection a été choisie de sorte à faire apparaître les points critiques. Les axes des abscisses et des ordonnées correspondent à x_9 et x_{10} , les deux dernières coordonnées du point P_5 . Le troisième axe correspond à t , le paramètre d’interpolation. La courbe en gras représente \mathcal{S}_2 . Pour chacun de ses points $(\sigma(X), t)$, $\sigma(X)$ est une solution de G_2 pour la valuation $d(t)$ des paramètres (donc vérifie $\sigma(P_5) = (., ., 0)$). Les points noirs sur \mathcal{S}_2 sont les points critiques de H .

Terminons la discussion sur cet exemple en remarquant que le fait d’avoir tous les points d’une figure d’un chemin d’homotopie de H coplanaires n’est pas une coïncidence : l’ensemble de telles figures sépare l’ensemble de tous les zéros de H en deux demi-espaces.

Inconnues :
point P_0, \dots, P_7
Paramètres :
longueur h_0, \dots, h_{11}
Contraintes :
 $\text{distance}(P_0, P_1) = h_0$
...
 $\text{distance}(P_6, P_7) = h_{11}$
 $\text{coplanaires}(P_0, P_1, P_2, P_3)$
 $\text{coplanaires}(P_1, P_2, P_4, P_7)$
 $\text{coplanaires}(P_2, P_3, P_5, P_4)$
 $\text{coplanaires}(P_3, P_0, P_6, P_5)$
 $\text{coplanaires}(P_0, P_1, P_7, P_6)$
 $\text{coplanaires}(P_4, P_5, P_6, P_7)$

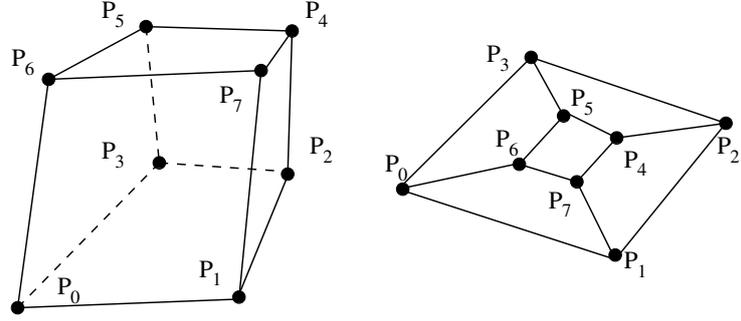


FIGURE 5.4 – L'énoncé (à gauche) et une esquisse cotée (au centre) du problème 3. À droite, une esquisse cotée du problème 4.

D'abord, sur une solution de G où P_0, P_3, P_4 et P_1, P_2, P_4 ne sont pas alignés, P_4 est dans le plan passant par P_0, \dots, P_3 si et seulement si P_5 est aussi dans ce plan (sinon cette solution ne respecte pas les contraintes $\text{coplanaires}(P_0, P_3, P_4, P_5)$ et $\text{coplanaires}(P_1, P_2, P_4, P_5)$). En se souvenant que dans le repère choisi le plan passant par P_0, \dots, P_3 est celui dont tous les points ont leur dernière coordonnée nulle, on en déduit que tous les points d'une solution $\sigma(X)$ de G sont coplanaires si et seulement si la dernière coordonnée de $\sigma(P_4)$ est nulle. Or l'ensemble des figures $\sigma(X)$ dont la dernière coordonnée de $\sigma(P_4)$ est nulle est un hyperplan, qui sépare l'espace de toutes les figures (c.à.d l'ensemble de toutes les valuations de X).

Ensuite, l'ensemble des figures sur lesquelles P_0, P_3, P_4 et P_1, P_2, P_4 sont alignés correspond à l'intersection de deux hyperplans de dimension 9 dans l'espace de toutes les figures \mathbb{R}^{11} (car une colinéarité induit deux degrés de restriction sur un point), qui n'existe pas en général.

5.1.2 Hexaèdre

On a déjà introduit le problème 3, dont les solutions sont des hexaèdres, dans le chapitre 1 ; on donne dans la figure 5.4 le SCG $G = C[X, A]$ qui lui est associé, ainsi qu'une esquisse cotée. Ce problème est bien contraint au sens de König-Hall et admet, pour presque toutes valeurs des paramètres, un nombre fini de solutions isolées. Cependant, il n'est pas génériquement bien contraint, car pour presque toutes valeurs de paramètres, il admet une infinité de solutions non isolées (qui forment un ensemble de dimension 1). Cet ensemble de dimension 1 parmi les solutions de G implique l'existence de points critiques dans les composantes connexes de $H^{-1}(0)$, où H est la fonction d'homotopie associée à G .

Considérons le SCG $G_2 = C_2[X, A]$, avec les mêmes inconnues et paramètres que G , obtenu en substituant dans C les 6 contraintes de coplanarité par les 5 contraintes :

$$\begin{aligned} &\text{coplanaires}(P_0, P_1, P_2, \mathbf{P_3}), \\ &\text{coplanaires}(P_0, P_1, P_2, \mathbf{P_4}), \\ &\text{coplanaires}(P_0, P_1, P_2, \mathbf{P_5}), \\ &\text{coplanaires}(P_0, P_1, P_2, \mathbf{P_6}), \\ &\text{coplanaires}(P_0, P_1, P_2, \mathbf{P_7}). \end{aligned}$$

La partie droite de la figure 5.4 présente une esquisse cotée de G_2 . Ce SCG est celui du problème 4 introduit dans le chapitre 1. Comme dans la sous-section précédente, les solutions de G_2 sont *aplaties*, dans le sens où tous les points de ces solutions sont dans un même plan.

Les solutions de G_2 sont de manière évidente des solutions de G . De plus, 18 coordonnées doivent être déterminées pour fixer les éléments de X à mouvement rigides près¹, et C_2 comporte 17 contraintes. La fonction numérique associée à G_2 est alors $F_2 : \mathbb{R}^{18} \times \mathbb{R}^{12} \rightarrow \mathbb{R}^{17}$ et, en fixant une valuation de A , $F_2(\cdot, \sigma(A)) : \mathbb{R}^{18} \rightarrow \mathbb{R}^{17}$. En supposant que 0 soit une valeur régulière de $F_2(X, \sigma(A))$, les composantes connexes de $(F_2(\cdot, \sigma(A)))^{-1}(0)$ sont des variétés différentielles de dimension 1.

Alors pour presque toutes valuations $\sigma_{sk}(A)$ et $\sigma_{so}(A)$ des paramètres, la fonction d'homotopie $H_2 : \mathbb{R}^{19} \rightarrow \mathbb{R}^{17}$ associée au SCG G_2 définit des composantes connexes de $H_2^{-1}(0)$ qui sont des variétés de dimension 2.

En supposant que l'esquisse $\sigma_{sk}(X)$ n'est pas aplatie et que \mathcal{S} est la composante connexe de $H^{-1}(0)$ à laquelle appartient l'esquisse, si \mathcal{S} contient un point $(\sigma_2(X), t_2)$ tel que $\sigma_2(X)$ soit aplatie (donc solution de G_2 pour la valuation $d(t_2)$), $(\sigma_2(X), t_2)$ appartient à une composante connexe de $H_2^{-1}(0)$ qui est une variété de dimension 2 strictement incluse dans \mathcal{S} ; \mathcal{S} n'est donc pas une variété différentielle de dimension 1, et contient des points critiques, dont $(\sigma_2(X), t_2)$. À l'instar du problème 2, l'ensemble des valuations de X pour lesquelles tous les points sont coplanaires forment un hyperplan de dimension 18 qui sépare l'espace des valuations.

5.1.3 Points critiques et théorèmes de géométrie

On a vu deux exemples de SCG impliquant des contraintes d'incidence pour lesquels la fonction d'homotopie n'admet pas 0 comme valeur régulière. Dans ces deux cas, $(\sigma(X), t) \in H^{-1}(0)$ est un point critique quand $\sigma(X)$ est solution d'un SCG G_2 dont les solutions sont aussi solutions de G , et que G' est bien contraint ou sous contraint.

Cela conduit à poser la conjecture suivante : si

- (i) $G = C[X, A]$ est bien contraint au sens de König-Hall,
 - (ii) il existe un SCG $G_2 = C_2[X, A]$ tel que pour toute valuation des paramètres, l'ensemble des solutions de G_2 est strictement inclus dans celui de G ,
 - (iii) G_2 est bien contraint ou sous-contraint au sens de König-Hall,
 - (iv) l'esquisse $\sigma_{sk}(X)$ n'est pas solution de G_2 ,
- alors si $(\sigma(X), t) \in H^{-1}(0)$ et $\sigma(X)$ est solution de G_2 , $(\sigma(X), t)$ est un point critique de H .

Dans les deux exemples précédents, ces situations sont dues à la satisfaction d'hypothèses d'un théorème de la géométrie d'incidence qu'on pourrait énoncer par : "si $n \geq 2$ points sont coplanaires, alors tout sous ensemble de $m \leq n$ de ces points est formé de points coplanaires".

La section 5.3 présente des pistes de réflexions à propos de la détection :

- avant le suivi s'il existe, pour un problème G donné, d'un problème G' bien ou sous-contraint dont les solutions sont solutions de G ;
- pendant le suivi, aux abords d'un point potentiellement critique, du problème G' s'il est dû à la géométrie d'incidence.

On propose, dans la section 5.2, une méthode pour construire la fonction d'homotopie de façon à ce qu'elle n'admette pas de points critiques $(\sigma(X), t)$ avec $t \notin \{0, 1\}$, qu'on justifie par un argument

1. En voyant G_2 comme un SCG en 3D, les coordonnées libres de ses points sont au nombre de 18.

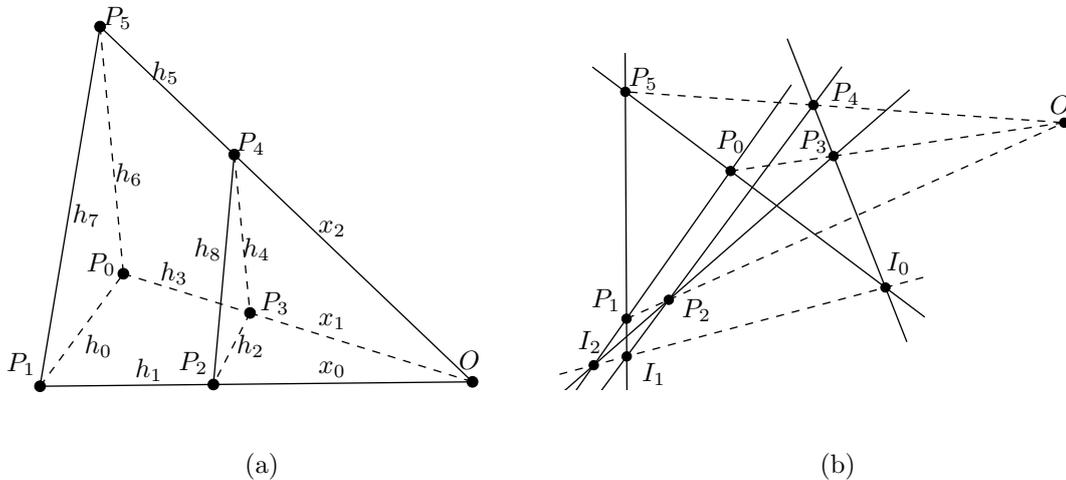


FIGURE 5.5 – (a) un pentaèdre inscrit dans un tétraèdre de sommets $OP_0P_1P_5$; (b) l'illustration du théorème de Desargues.

probabiliste. $(\sigma(X), t)$ est alors un point critique de H pour la valeur 0 si et seulement si $t = 0$ ou $t = 1$, et $\sigma(X)$ est un point critique de $F(\cdot, \sigma_{so}(A))$ ou de $F(\cdot, \sigma_{sk}(A))$.

La fin de cette section montre comment, dans le cas des deux exemples présentés ici, écrire un système d'équations qui n'admette que des solutions non aplaties.

5.1.4 Réécriture du système de contraintes

Dans les deux exemples utilisés pour illustrer ce chapitre, le pentaèdre et l'hexaèdre, l'existence d'ensembles de solutions de dimension strictement positive et de points critiques sont dus aux solutions aplaties qui satisfont les contraintes de coplanarité. On propose, pour chacun de ces deux exemples, de poser le système d'équations de façon à ce qu'il n'admette que des solutions non aplaties, en considérant les configurations de Desargues que ces dernières satisfont. La généralisation des constructions présentées ici à d'autres polyèdres semble difficile, mais donne lieu à un problème intéressant.

La figure 5.5(a) montre comment construire un pentaèdre $P_0 \dots P_5$ dans un tétraèdre $OP_1P_0P_5$. Réciproquement, toute solution $P_0 \dots P_5$ non aplatie du problème 2 peut s'inscrire dans un tétraèdre, dont un sommet est éventuellement à l'infini. En effet, les trois droites (P_0P_3) , (P_1P_2) , (P_4P_5) sont concourantes ou parallèles. Cette propriété est en général fautive quand les points $P_0 \dots P_5$ sont tous coplanaires; d'après le théorème de Desargues, (P_0P_3) , (P_1P_2) et (P_4P_5) sont concourantes ou parallèles si et seulement si une des conditions suivantes est réalisée :

- (C1) $(P_0P_1) \parallel (P_2P_3)$, $(P_1P_5) \parallel (P_2P_4)$, $(P_0P_5) \parallel (P_3P_4)$,
- (C2) les trois couples de droites de la condition (C1) sont formés de droites sécantes dont les intersections respectives sont alignées,
- (C3) deux des trois couples de droites de la condition (C1) sont formés de droites sécantes et la droite passant par leurs intersections respectives est parallèle aux deux droites du troisième couple.

La figure 5.5(b) illustre une configuration où $P_0 \dots P_5$ sont coplanaires et où la condition (C2) est réalisée.

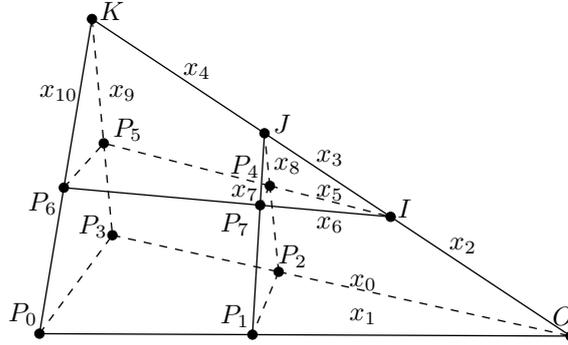


FIGURE 5.6 – Un hexaèdre inscrit dans un tétraèdre de sommets OP_0P_3K tronqué.

En posant x_0 la distance de P_2 à O , x_1 celle de P_3 à O et x_2 celle de P_4 à O (voir figure 5.5(a)), on peut utiliser les formules d’Al Kashi dans les trois triangles formant les faces du tétraèdre $OP_0P_1P_5$, pour obtenir un système de trois polynômes de degrés trois en les inconnues x_0 , x_1 et x_2 . Par exemple, on peut écrire en considérant successivement les triangles OP_2P_4 et OP_1P_5 les relations :

$$\cos(OP_2P_4) = \frac{x_0^2 + x_2^2 - h_8^2}{2x_0x_2} \text{ et } \cos(OP_1P_5) = \frac{(x_0 + h_1)^2 + (x_2 + h_5)^2 - h_7^2}{2(x_0 + h_1)(x_2 + h_5)}$$

et en remarquant que $\cos(OP_2P_4) = \cos(OP_1P_5)$, on obtient un polynôme en x_0 et x_2 .

Les solutions du système d’équations ainsi obtenu sont des triplets (x_0, x_1, x_2) . Ceux qui sont réels et permettent de construire un tétraèdre $OP_2P_3P_4$ correspondent chacun à une solution du problème 2. De la même façon, il est possible de poser un système de 11 équations de degrés 3 ou 4 en les 11 inconnues x_0, \dots, x_{10} pour trouver les solutions non-aplaties du problème de la construction de l’hexaèdre, comme illustré sur la figure 5.6.

Dans le cas du pentaèdre, le système d’équations obtenu a un degré beaucoup plus petit (27 contre 4608) que le système construit à partir des contraintes; sa résolution par une méthode d’homotopie nécessite donc de suivre moins de chemins. Par contre, dans le cas de l’hexaèdre, le degré du système décrit ici est de 1769472 contre 995328 pour le système construit à partir des contraintes. De plus la méthode par homotopie décrite dans le chapitre 3 est peu adaptée pour le résoudre : lors de la déformation de l’esquisse, dès qu’une intersection de deux droites coplanaires part vers l’infini (les deux droites deviennent parallèles), le chemin suivi tend vers une solution à l’infini. Même si elle peut être détectée, cette situation peut conduire à ne trouver aucune solution. C’est ce qui a été observé sur les expériences menées.

On présente dans la suite une approche pour éviter les points critiques ne nécessitant pas de modifier le système d’équations et applicable à tous les SCG.

5.2 Éviter les points critiques

Dans la section précédente ont été présentés deux SCG impliquant des contraintes booléennes dont la fonction d’homotopie n’admettait pas 0 comme valeur régulière. Il en résulte que les chemins d’homotopie ne sont pas des variétés de dimension 1, et ne peuvent pas être suivis par une

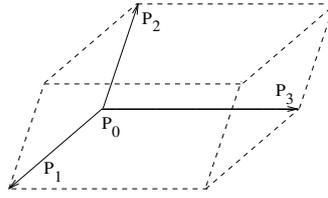


FIGURE 5.7 – Le parallélépipède engendré par les trois vecteurs $\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \overrightarrow{P_0P_3}$.

méthode de suivi de chemins classique. On suppose dans cette partie que ces points critiques sont dus aux contraintes booléennes. Une façon pragmatique d'éviter ces points critiques est alors de remplacer ces contraintes booléennes par des contraintes métriques. Par exemple, une contrainte de coplanarité de quatre points peut être remplacée par une contrainte de volume, dont le paramètre (le volume à satisfaire) est nul. L'heuristique que l'on utilisera alors est qu'en construisant la fonction d'homotopie de façon à ce que pour $t \neq 1$, ces volumes ne soient pas nuls, cette fonction n'admet comme points critiques que ceux qui sont solution du SCG à résoudre. Cette heuristique sera justifiée par le théorème de Sard².

Dans cette section, on fera l'hypothèse que souvent, et dans la plupart des problèmes issus de la CAO, la fonction numérique $F : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ associée à un SCG n'impliquant que des contraintes métriques admet 0 comme valeur régulière.

Comme on le verra dans la section suivante, cette hypothèse implique qu'un SCG bien contraint au sens de König-Hall et ne contenant que des contraintes métriques est génériquement bien contraint, ce qui est faux en général, et notamment quand le SCG contient des contraintes redondantes ou contradictoires. Ces cas peuvent être détectés grâce à la méthode du témoin (voir [Michelucci 06a, Michelucci 07]) introduite dans le chapitre 2. Cependant, dans la plupart des exemples rencontrés et entre autres ceux utilisés dans ce mémoire, la bonne constriction structurelle implique la bonne constriction générique.

Afin d'utiliser cette hypothèse, on considère que toute contrainte booléenne d'un SCG peut-être écrite comme une contrainte métrique de paramètre nul. Par exemple, une contrainte de coplanarité $coplanaires(P_0, P_1, P_2, P_3)$ peut être vue comme la contrainte de volume $volume(P_0, P_1, P_2, P_3) = 0$, interprétée par la fonction numérique :

$$det(\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \overrightarrow{P_0P_3}) = 0,$$

mesurant le volume du parallélépipède engendré par les trois vecteurs $\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \overrightarrow{P_0P_3}$ (voir figure 5.7). Indépendamment de leurs interprétations, ces deux contraintes sont équivalentes, c'est-à-dire que si $\sigma(\{P_0, P_1, P_2, P_3\})$ satisfait l'une, elle satisfait l'autre.

D'un point de vue formel, cela suppose d'étendre l'univers géométrique considéré en ajoutant à sa signature, pour chaque symbole de contrainte booléenne, le symbole de contrainte métrique adéquat et une sorte correspondante. Par exemple, pour le symbole $colineaires(point, point, point)$, on ajoute le symbole $aire(point, point, point) \rightarrow surface$ et la sorte $surface$.

En remplaçant chaque contrainte booléenne par une contrainte métrique dans G de cette façon, on peut alors voir la fonction numérique qui lui est associée comme $F : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, et lui appliquer l'hypothèse précédente assurant la régularité. Cela n'entre pas en contradiction avec les

2. On considère ici des contraintes booléennes de degré de restriction 1. Celles de degré de restriction supérieurs (par exemple une incidence point droite en 3D) nécessitent d'introduire autant de paramètres que leur degré de restriction, sous peine de rendre le problème sous-contraint.

exemples précédents, car on y considèrerait la restriction de F aux valuations des paramètres pour lesquels les valeurs des paramètres des contraintes booléennes étaient nulles.

5.2.1 Théorème de Sard

Le théorème de Sard est un outil central en topologie différentielle. Il permet de transférer presque sûrement une propriété d’une régularité d’une fonction dépendant de paramètres sur sa restriction à certaines valeurs de paramètres. On adapte ici la version de [Allgower 97] à nos notations :

Théorème 15 (Théorème de Sard) *Soient A, B, C des variétés différentielles de dimension finie, avec $\dim(A) \geq \dim(B)$, et soit $\mathcal{F} : A \times B \rightarrow C$ une fonction C^∞ . Si 0 est une valeur régulière de \mathcal{F} , alors pour presque tout $b \in B$ (au sens de la mesure de Lebesgue sur B), la restriction $\mathcal{F}(\cdot, b) : A \rightarrow C$ admet 0 comme valeur régulière.*

Supposons que $F(X, A)$, la fonction associée à un SCG $G = C[X, A]$ dont toutes les contraintes sont métriques, admette 0 comme valeur régulière. Alors en appliquant ce théorème à F , on en déduit que pour presque toute valuation $\sigma(A)$ des paramètres, $F(X, \sigma(A))$ admet 0 comme valeur régulière ; donc l’ensemble des 0 de $F(X, \sigma(A))$ (c.à.d l’ensemble de solutions de G pour la valuation $\sigma(A)$) est formé de points isolés. Si de plus F est polynomiale, ces points isolés sont en nombre fini, et G est génériquement bien contraint, d’où la remarque de la section précédente.

Dans la sous-section 5.1.2, on a présenté un exemple (l’hexaèdre) justifiant le “pour presque tout $b \in B$ ”. En effet, on avait montré que pour presque toutes les valuations $\sigma(A)$ telles que $\sigma(h_i) = 0$ si h_i est le paramètre d’une contrainte de coplanarité, $F(X, \sigma(A))$ n’admettait pas 0 comme valeur régulière. Or dans \mathbb{R}^m , l’ensemble des valuations nulles pour un ensemble non vide de ses paramètres est un ensemble de mesure de Lebesgue nulle.

On peut cependant remarquer que, sous l’hypothèse de régularité de la valeur 0 pour $F(X, A)$, si 0 n’est pas une valeur régulière de $F(X, \sigma(A))$, alors pour presque toute “perturbation” $\sigma'(A)$, $F(X, \sigma(A) + \sigma'(A))$ admet 0 comme valeur régulière, où $\sigma(A) + \sigma'(A)$ est défini comme l’ensemble $\{\sigma(h_0) + \sigma'(h_0), \dots, \sigma(h_{m-1}) + \sigma'(h_{m-1})\}$ pour $A = \{h_0, \dots, h_{m-1}\}$. On propose dans la sous-section suivante d’utiliser la fonction d’interpolation pour créer une petite perturbation des paramètres.

5.2.2 Perturbation des paramètres

Soient $c_0, \dots, c_{m''-1}$ les contraintes booléennes d’un SCG (ici considérées comme métriques), où $m'' = m - m'$ avec m' le nombre de contraintes métriques originales du SCG. Soit encore $\{h_0, \dots, h_{m''-1}\} \subset A$ l’ensemble des paramètres ajoutés pour transformer les contraintes booléennes en contraintes métriques. Soit enfin d une fonction d’interpolation de $\sigma_{sk}(A)$ à $\sigma_{so}(A)$, où pour $0 \leq i \leq m'' - 1$, d_i interpole le paramètre h_i .

Remarquons que si les composantes d_i , pour $0 \leq i \leq m'' - 1$, sont linéaires, et que $\sigma_{sk}(A) = \sigma_{so}(A)$, alors pour tout $t \in \mathbb{R}$, $d_i(t) = d_i(0) = d_i(1)$, et si 0 n’est pas une valeur régulière de $F(X, \sigma_{so}(A))$, il n’y a aucune raison qu’elle le soit pour $H(X, t) = F(X, d(t))$. On construit alors d pour qu’elle interpole de manière non linéaire les paramètres $h_0, \dots, h_{m''-1}$, c’est à dire les paramètres correspondant aux contraintes métriques.

Soit d telle que :

- (1) $(d_0(0), \dots, d_{m''-1}(0)) = (\sigma_{sk}(h_0), \dots, \sigma_{sk}(h_{m''-1}))$,
- (2) $(d_0(1), \dots, d_{m''-1}(1)) = (\sigma_{so}(h_0), \dots, \sigma_{so}(h_{m''-1}))$,
- (3) $t \in \mathbb{R} \setminus \{0, 1\} \Rightarrow \forall 0 \leq i \leq m'' - 1, d_i(t) \neq \sigma_{so}(h_i)$,

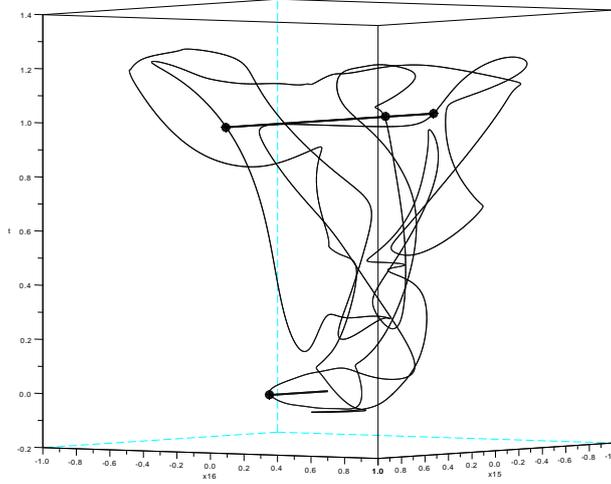


FIGURE 5.8 – Une projection sur \mathbb{R}^3 d’une composante connexe de $(H')^{-1}(0)$ pour le problème 3. x_{15}, x_{16} , les dimensions des axes horizontaux, sont les deux dernières coordonnées de P_7 . L’axe vertical a pour dimension t , le paramètre d’interpolation.

on suppose alors que la fonction $H_{|\mathbb{R} \setminus \{0,1\}} : \mathbb{R}^m \times (\mathbb{R} \setminus \{0,1\}) \rightarrow \mathbb{R}^m$ définie par $H_{|\mathbb{R} \setminus \{0,1\}}(X, t) = F(X, d(t))$ admet “presque sûrement” 0 comme valeur régulière.

Pour montrer cette affirmation, il faudrait introduire une mesure sur l’ensemble de toutes les fonctions d’interpolation d de $\sigma_{sk}(A)$ à $\sigma_{so}(A)$ vérifiant (1), (2) et (3). (On pourrait aussi se restreindre à des fonctions d’interpolations quadratiques, déterminées uniquement par les coefficients des monômes de ses composantes). On pourrait alors montrer que pour presque toute fonction d’interpolation d de $\sigma_{sk}(A)$ à $\sigma_{so}(A)$ vérifiant (1),(2) et (3), et presque tout $t \in \mathbb{R} \setminus \{0,1\}$, $F(X, d(t))$ admet 0 comme valeur régulière. Le transfert de cette propriété à $H_{|\mathbb{R} \setminus \{0,1\}}$, où t intervient comme une variable, est alors montré dans [Li 93].

On se bornera à montrer, sur les deux exemples considérés précédemment, que pour un choix simple de la fonction d’interpolation d , les points critiques $(\sigma(X), t)$ avec $t \notin \{0,1\}$ sont évités. Ceux qui sont solutions de G pour la valuation $\sigma_{so}(A)$ (c.à.d tels que $t \in \{0,1\}$) appartiennent toujours aux composantes connexes de $H^{-1}(0)$.

Remarquons enfin que cette construction de la fonction d’homotopie permet d’utiliser une esquisse qui ne satisfait pas les contraintes booléennes de l’énoncé, puisqu’on peut lire les valeurs $\sigma_{sk}(h_i)$, pour $0 \leq i \leq m'' - 1$, sur cette esquisse, comme dans le cas des paramètres des contraintes métriques.

5.2.3 Résultats

Pour le problème 2 étudié dans la sous-section 5.1.1 (le pentaèdre), on construit le SCG G' en remplaçant les trois coplanarités notées c_9, c_{10}, c_{11} par les contraintes de volume c'_9, c'_{10}, c'_{11} de paramètres respectifs h_9, h_{10}, h_{11} . Si pour au moins une coplanarité c'_i , la fonction d’interpolation d_i de la contrainte de volume associée est telle que $t \notin \{0,1\} \Rightarrow d_i(t) \neq 0$ (la condition (3) est

alors satisfaite), alors si $(\sigma(X), t)$ est un point d'un chemin d'homotopie \mathcal{S} de H avec $t \notin \{0, 1\}$, les points de la figure $\sigma(X)$ ne sont pas tous coplanaires, sinon la contrainte de volume c'_i ne serait pas respectée. Dans le cas particulier de ce problème, comme G_2 (le SCG dont les solutions sont aplaties) est bien contraint, il admet seulement un nombre fini de solutions pour la valuation $\sigma_{so}(A)$, donc si H_2 est la fonction d'homotopie associée à G_2 les composantes connexes de $H_2^{-1}(0)$ intersectent \mathcal{P}_1 et \mathcal{P}_0 seulement en un nombre fini de points.

Dans le cas du problème 3, étudié en sous-section 5.1.2 (l'hexaèdre), si la fonction d'interpolation satisfait les conditions (1),(2) et (3), les points d'un chemin d'homotopie \mathcal{S}' ne sont solutions de G_2 que quand $t = 0$ ou $t = 1$. Par contre, ici G_2 n'est pas bien contraint et l'intersection d'une composante connexe de $H_2^{-1}(0)$ avec \mathcal{P}_1 (ou \mathcal{P}_0) est une variété de dimension 1. Si pour un point $(\sigma(X), t) \in \mathcal{S}'$, avec $t \in \{0, 1\}$, les points de $\sigma(X)$ sont tous coplanaires, $(\sigma(X), t)$ appartient à une composante connexe \mathcal{S}_2 de $H_2^{-1}(0)$, et $\mathcal{S}_2 \cap \mathcal{P}_1$ (ou $\mathcal{S}_2 \cap \mathcal{P}_0$) est inclus dans \mathcal{S}' . $(\sigma(X), t)$ est donc toujours un point critique de H . La figure 5.8 présente une projection dans \mathbb{R}^3 de \mathcal{S}' pour le problème 3 : les lignes en gras sont des parties de $\mathcal{S}_2 \cap (\mathcal{P}_0 \cup \mathcal{P}_1)$.

Pour un SCG quelconque dont les contraintes booléennes ont été transformées en contraintes métriques, et d vérifiant les conditions (1), (2) et (3) ci-dessus, on notera $\mathcal{S}'' = \mathcal{S}' \setminus (\mathcal{P}_0 \cup \mathcal{P}_1)$, et on considérera $\overline{\mathcal{S}''}$ l'ensemble des points de $\mathcal{P}_0 \cup \mathcal{P}_1$ qui sont des points d'adhérence de \mathcal{S}'' (les points noirs sur la figure 5.8 représentent cet ensemble dans le cas du problème 3).

On supposera alors :

- Les composantes connexes de \mathcal{S}'' sont des variétés différentielles de dimension 1.
- L'ensemble $\mathcal{S}'' \cup \overline{\mathcal{S}''}$ est une variété différentielle de dimension 1 (difféomorphe à un cercle ou à un intervalle ouvert).

Ces deux hypothèses justifient l'utilisation de la méthode de suivi de chemin donnée dans le chapitre 3 pour explorer $\mathcal{S}'' \cup \overline{\mathcal{S}''}$. En appliquant cette méthode de suivi pour explorer le chemin d'homotopie, les points critiques solutions (par exemple les points noirs de la figure 5.8) sont "traversés", et considérés comme des solutions.

5.3 Détecter les continus de solutions avec la géométrie d'incidence

On a présenté dans le chapitre 3 une adaptation d'une méthode de suivi de chemin par prédiction-corrrection permettant d'explorer un chemin d'homotopie contenant des points critiques uniquement sur les hyperplans \mathcal{P}_0 et \mathcal{P}_1 . Cette adaptation s'appuie sur l'heuristique expliquée dans la section précédente. Une condition suffisante pour que H admette des points critiques solutions de G est que G admette des continus de solutions en plus de solutions isolées. Ces points sont alors dus à un "changement de dimension" de l'ensemble suivi.

Le problème de la construction d'un hexaèdre admet, pour presque toute valuation des paramètres, des solutions isolées et des ensembles de solutions qui sont des variétés différentielles de dimension 1, qui peuvent être "traversées" par un chemin d'homotopie suivi depuis une esquisse. On peut voir ce phénomène comme conséquence de la satisfaction de l'hypothèse du théorème de géométrie d'incidence : si 8 points sont coplanaires, ils sont coplanaires 4 à 4. Les 6 contraintes de coplanarités du SCG associées au problème de la construction d'un hexaèdre sont donc automatiquement satisfaites par une figure sur laquelle les 8 points sont coplanaires. D'autre part, 5 contraintes de coplanarité de 4 points suffisent à contraindre 8 points à se trouver sur un même plan, donnant lieu à un SCG sous-contraint.

On propose dans cette section une piste, qui a fait l’objet d’expérimentations, pour détecter soit à l’avance, par une analyse *a priori* du SCG à résoudre, soit pendant le suivi du chemin, quand il croise l’hyperplan \mathcal{P}_1 , si une telle situation se produit. Cette détection permet de répondre à deux questions :

- (Q1) Les chemins d’homotopie contiennent-ils des points critiques liés à un changement de dimension de l’ensemble de solutions ?
- (Q2) Le problème à résoudre admet-il des continuums de solutions ?

La question (Q1) est liée à la méthode par continuation présentée ici. Les points critiques solutions de G sont trouvés, et proposés à l’utilisateur comme solution. Ces solutions n’étant pas isolées, elles ne sont pas trouvées par des méthodes d’homotopie classiques. La question (Q2) est moins prosaïque, et y apporter une réponse renseigne sur la structure de l’espace des solutions. On propose d’utiliser pour répondre à ces questions un *matroïde géométrique* ([Michelucci 04, Michelucci 06b]), qui permet de résoudre des problèmes de géométrie d’incidence. La méthode envisagée n’est donc capable de traiter que des questions relatives à l’incidence.

On présente en sous-section 5.3.1 la notion de matroïde géométrique pour un ensemble de points d’un espace vectoriel de dimension d . Il peut être représenté par un graphe orienté contenant un sommet par sous-ensemble de ces points, et dont les arêtes traduisent la notion d’inclusion directe. Ainsi, pour n points, le graphe contient 2^n sommets. Cet ensemble de points est muni d’une application de l’ensemble de ses parties dans les entiers naturels, qui associe à chaque sous-ensemble de points la dimension du sous-espace affine qu’il engendre ; aussi cette application est-elle appelée fonction *rang* associée au matroïde. Cette application traduit les coïncidences, colinéarités, coplanarités et les extensions en dimensions supérieures, de tout sous-ensemble de points. Compléter un matroïde signifie déterminer l’application rang associée.

La sous-section 5.3.2 montrera comment le matroïde peut être utilisé pour répondre aux questions (Q1) et (Q2). Pour répondre à la question (Q2), il permet de déterminer un ensemble minimal de contraintes d’incidences que satisfait une figure. Pour répondre à (Q1), on extrait un ensemble minimal de contraintes dont les solutions satisfont nécessairement les contraintes d’un SCG.

On donnera enfin une méthode naïve pour extraire cet ensemble minimal de contraintes en sous-section 5.3.3.

Remarquons par ailleurs que des méthodes symboliques de résolution de systèmes d’équations polynomiales permettent de trouver tous les ensembles de solutions d’un système, sous forme d’idéaux de polynômes [Wen-Tsun 86, Chou 90]. Une analyse de ces idéaux permet de déterminer la nature géométrique d’un ensemble de solutions. Cependant ces méthodes ont un coût doublement exponentiel, ce qui compromet leur utilisation pour des problèmes issus de la CAO.

Les méthodes de la “numerical algebraic geometry” présentées dans le chapitre 2 permettent aussi de trouver tous ces ensembles de solutions. Par exemple, on peut résoudre avec le logiciel **Bertini**, qui les implémente, le problème de l’hexaèdre. Avec la valuation des paramètres donnée dans le chapitre 3, en environ 48 heures, les 16 solutions réelles isolées sont trouvées, et des ensembles de dimension 1, dont toutes les solutions sont aplaties, sont trouvés. Ce problème peut être posé comme dépendant de 17 variables, et **Bertini** cherche les ensembles de solutions de dimensions 16, puis 15, ..., pour finir par chercher les solutions isolées. C’est ce qui explique ce temps d’exécution.

5.3.1 Matroïde géométrique

Les matroïdes permettent d’inférer des propriétés de dépendance ou d’indépendance entre les sous-ensembles d’un ensemble E . Plusieurs formalisations des matroïdes sont possibles et on propose ici d’utiliser celle traduisant les relations de dépendances affines de points dans un espace

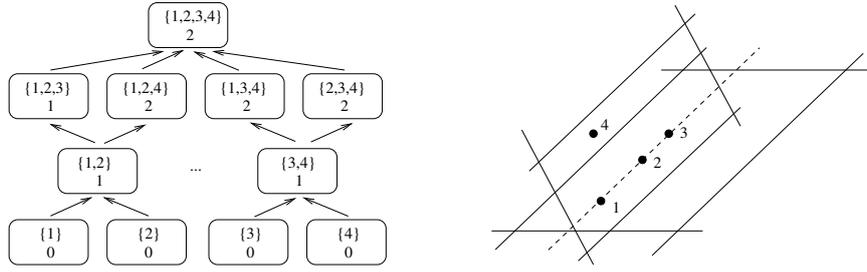


FIGURE 5.9 – Un matroïde pour 4 points dont 3 sont alignés dans \mathbb{R}^d où $d = 3$. On n’a pas représenté tous les sommets associés aux ensembles de 2 points, ni toutes les arêtes.

affine de dimension d (les cas qui nous intéressent ici sont ceux où $d = 2$ ou $d = 3$). Ces relations peuvent se traduire en termes de rangs.

On considère un ensemble E de points dans un espace affine de dimension d . Pour tout sous-ensemble $F = \{e_0, \dots, e_{n-1}\}$ de E , on appelle *rang* de F , qu’on notera $rg(F)$ la dimension de l’espace vectoriel engendré par la famille $\langle \overrightarrow{e_0 e_1}, \dots, \overrightarrow{e_0 e_{n-1}} \rangle$ (c.à.d la taille de sa plus grande sous-famille libre). Il est alors clair que :

- $rg(F) \leq d$,
- $rg(F) \leq |F| - 1$ (où $|F|$ est le cardinal de F).

Parfois la fonction rang est définie comme attribuant à un ensemble de points le cardinal d’un plus grand sous-ensemble de points linéairement indépendants. Cette définition permet d’attribuer un rang égal à 0 à l’ensemble vide, ce qui n’a pas de sens avec notre définition. Les rangs d’un ensemble pour ces deux définitions sont les mêmes à la constante 1 près ; on décide que la fonction rang vaut -1 pour l’ensemble vide, ce qui garantit sa définition pour les opérations ensemblistes. [Michelucci 04] présente trois axiomes satisfaits par la fonction rang, qu’on adapte ici à nos notations :

- (A1) $|F| = 1 \Rightarrow rg(F) = 0$,
- (A2) $F \subseteq G \Rightarrow rg(F) \leq rg(G)$ (monotonie),
- (A3) $rg(F \cup G) + rg(F \cap G) \leq rg(F) + rg(G)$ (sous-modularité).

On considère aussi la propriété suivante, qui est une conséquence de la sous-modularité :

- (P4) $e \in E \Rightarrow rg(F \cup \{e\}) \leq rg(F) + 1$.

On représente le *matroïde* associé à un ensemble fini de points E en dimension d par un graphe orienté $\mathcal{G} = \langle S, A \rangle$ muni d’une fonction d’étiquetage de ses sommets $f : S \rightarrow \mathbb{N}$, tels que

- S soit en bijection avec $\mathcal{P}(E) \setminus \{\emptyset\}$ (chaque sommet de \mathcal{G} correspond à un sous-ensemble non-vide de E),
- $(s_1, s_2) \in A$ si et seulement si F_1 et F_2 , les sous-ensembles associés respectivement à s_1 et s_2 , sont tels que $F_1 \subset F_2$ et $|F_2 \setminus F_1| = 1$,
- si $s \in S$, et F est le sous-ensemble de E qui lui est associé, $f(s) = rg(F)$.

Notons que la fonction d’étiquetage n’est pas nécessairement connue ; on dira qu’un matroïde associé à E est *compatible* si la fonction d’étiquetage satisfait les axiomes (A1),(A2) et (A3).

Par soucis de simplicité, on confondra un sommet s de \mathcal{G} avec le sous-ensemble F de E qui lui est associé. En utilisant la terminologie des DAG³, si $(E_1, E_2) \in A$, on dira que E_1 est un *fil*s de E_2 , et que E_2 est un *père* de E_1 .

3. DAG pour graphes orientés sans cycles.

Remarquons que la définition du matroïde comme un graphe orienté est superflue, car la relation, au sens de la théorie des graphes, induite par ses arêtes est celle de l'inclusion ensembliste. La donnée d'un ensemble E et de la fonction rang suffit pour formaliser la notion de matroïde. On utilise ici la notion de graphe orienté (en fait, c'est un treillis) pour insister sur sa nature combinatoire. La figure 5.9 représente un matroïde associé à l'ensemble de points $\{1, 2, 3, 4\}$ dans \mathbb{R}^3 , où 1, 2, 3 sont alignés. On y a étiqueté chaque sommet, correspondant à un sous-ensemble $F \subset \{1, 2, 3, 4\}$, par son rang.

Soit un ensemble E de points placés dans l'espace \mathbb{R}^d , c'est à dire, avec les notations de ce mémoire, qu'on connaît une valuation $\sigma(E)$ associant à chaque point de E un élément de \mathbb{R}^d . Avec du calcul exact dans \mathbb{R}^d , il est trivial mais très couteux de construire un matroïde compatible associé à E , en calculant le rang de chaque sous-ensemble de E ; ce matroïde est unique. Avec du calcul flottant, trouver un matroïde compatible n'est pas aussi aisé.

Si aucune valuation de E n'est connue (par exemple quand les éléments de E sont les inconnues d'un SCG), mais que les rangs $rg(F_1) = r_1, \dots, rg(F_n) = r_n$ de certains sous-ensembles de E sont connus (par exemple par des contraintes de colinéarité ou de coplanarité), on appelle *compléter* un matroïde le fait de chercher une fonction d'étiquetage f (c.à.d étiqueter les sommets du matroïde) telle que $f(F_1) = r_1, \dots, f(F_n) = r_n$, et f satisfasse les axiomes (A1),(A2) et (A3). Si une telle fonction existe (le matroïde est compatible), il est possible de trouver une valuation de E qui satisfasse ces contraintes; sinon l'ensemble de contraintes n'admet aucune solution. Cette fonction d'étiquetage n'est pas nécessairement unique et en complétant un matroïde à partir d'un ensemble de rangs connus, on obtient une liste de matroïdes compatibles, éventuellement vide.

Dans [Michelucci 04, Michelucci 06b], un matroïde est utilisé pour prouver des théorèmes de géométrie portant sur un ensemble fini E de points de \mathbb{R}^d , dont les hypothèses et les conséquences peuvent s'exprimer sous forme de rangs. S'il est possible, en partant des rangs donnés par les hypothèses, de compléter le matroïde associé à E et que toute fonction d'étiquetage d'un matroïde compatible vérifie les conséquences du théorème, le théorème est prouvé.

5.3.2 Calcul des ensembles de solutions de dimension positive

On propose d'utiliser ici un matroïde pour calculer les relations de dépendances entre des contraintes satisfaites par une figure, et ainsi calculer le degré de restriction d'un ensemble de contraintes. En effet, on peut voir un matroïde comme un ensemble de contraintes. Dans l'exemple de la figure 5.9, le sommet $\{1, 2, 3\}$, avec $f(\{1, 2, 3\}) = 1$, peut se traduire par *colineaire*(1, 2, 3). Cette contrainte, dans un univers 3D, a un degré de restriction de 2, car un point appartenant à une droite en 3D a un seul degré de liberté. On peut ainsi étendre la notion de degré de restriction aux sommets d'un matroïde dont la fonction d'étiquetage est connue. Aussi, le sommet $\{1, 2, 3, 4\}$, avec $f(\{1, 2, 3, 4\}) = 2$ a un degré de restriction de 1, mais cette contrainte est une conséquence directe de *colineaire*(1, 2, 3), et le degré de restriction de l'ensemble de ces deux contraintes est 2, et non 3. Des cas plus compliqués peuvent avoir lieu.

On introduit la notion d'ensemble de contraintes indépendantes d'un matroïde dont la fonction d'étiquetage f est connue. C'est un ensemble minimal, au sens de l'inclusion, de sommets F_1, \dots, F_m , tel qu'en complétant un matroïde à partir des rangs connus $f(F_1), \dots, f(F_m)$, la fonction d'étiquetage f' de tout matroïde compatible obtenu vérifie $F \subseteq E \Rightarrow f'(F) \leq f(F)$. On suppose dans cette partie qu'on est capable d'extraire d'un matroïde un ensemble de contraintes indépendantes, et qu'on est capable de calculer le degré de restriction de cet ensemble.

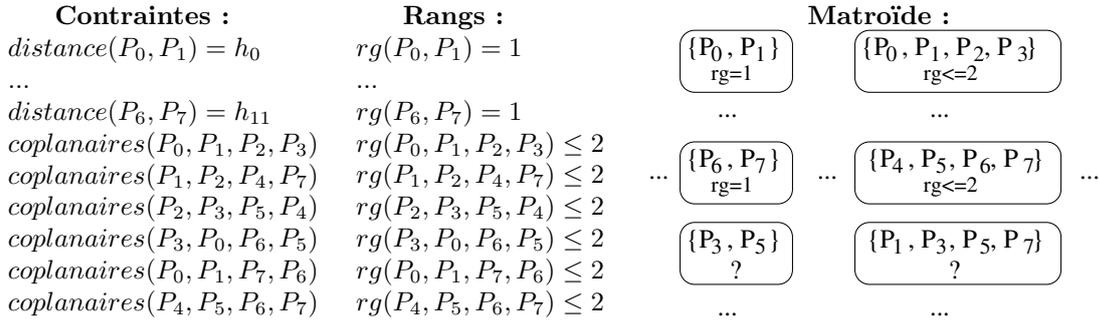


FIGURE 5.10 – Construire un matroïde à partir de l’ensemble de contraintes d’un SCG. À droite, dans le matroïde obtenu, on a étiqueté les sommets dont les rangs sont à déterminer par ‘?’.

Analyser a priori un SCG

La première utilisation d’un matroïde consiste à trouver, pour un SCG G faisant intervenir un ensemble C de contraintes d’incidences, des ensembles de contraintes indépendantes impliquant les contraintes d’incidence du SCG. Pour un tel ensemble de contraintes C' , si son degré de restriction est strictement plus petit que celui de C , alors le SCG G' obtenu à partir de G en remplaçant C par C' est sous-contraint. Ses solutions, qui forment un ensemble de dimension strictement positive, sont alors solutions de G . Voici l’esquisse d’un algorithme réalisant cette tâche :

Soit E un ensemble de points, C un ensemble de contraintes portant sur E .

- (1.) Construire le matroïde associé à E , avec les contraintes de C .
- (2.) Compléter ce matroïde pour obtenir un ensemble de matroïdes compatibles.
- (3.) Pour chaque matroïde compatible, extraire un ensemble C' de contraintes indépendantes.
- (4.) Calculer le degré de restriction de C' .

L’étape (1.) est illustrée sur la figure 5.10, pour les contraintes du problème de construction d’un hexaèdre. L’étape (2.) est très coûteuse. Elle peut être réalisée en propageant les rangs connus dans les sommets pères et fils grâce à des règles tirées de (A1),(A2),(A3) et (P4). En général, le nombre de matroïdes obtenus est très élevé, car le problème est sous-déterminé. Pour diminuer ce nombre de matroïdes, on peut par exemple utiliser les contraintes de distance pour déterminer le rang des ensembles de points sur lesquelles elles agissent (comme suggéré sur la figure 5.10). [Michelucci 06b] montre que le problème de la satisfaction d’une formule logique (SAT) se ramène à la recherche de matroïdes compatibles, et l’étape (2.) correspond par conséquent à un problème NP-complet. La question des étapes (3.) et (4.) sera abordée dans la section 5.3.3.

En appliquant cet algorithme au problème 3, on trouve deux matroïdes desquels on peut extraire plusieurs ensembles de contraintes indépendantes ayant un degré de restriction égal à 6, notamment l’ensemble des 6 contraintes de coplanarités du problème, et un seul ensemble de contraintes indépendantes ayant un degré de restriction inférieur à 6, à savoir l’ensemble formé de l’unique contrainte $coplanaires(P_0, P_1, P_2, P_3, P_4, P_5, P_6, P_7)$, qui a un degré de restriction de 5. En effet, trois des points servent à déterminer un plan, et chacun des 5 points restants voit son degré de liberté diminuer de 1. On en déduit que le problème de l’hexaèdre admet un ensemble de solutions, de dimension $6 - 5 = 1$, dont tous les éléments correspondent à des figures où les 8 points sont coplanaires.

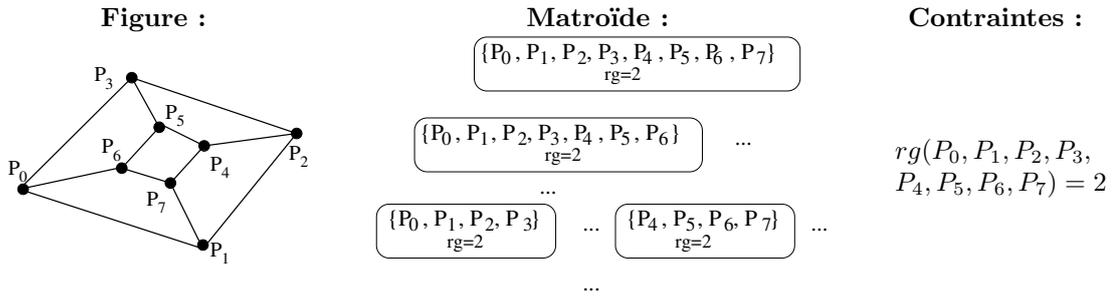


FIGURE 5.11 – À gauche, une solution du problème de l’hexaèdre. Au centre, quelques sommets du matroïde construit à partir de cette valuation. À droite, un ensemble de contraintes indépendantes.

Détecter un continuum de solutions pendant le suivi

Un matroïde peut aussi être utilisé pendant le suivi d’un chemin d’homotopie, à chaque fois que l’hyperplan \mathcal{P}_1 est croisé. Un matroïde peut être construit à partir de la figure correspondant à cette intersection, en lisant les rangs de chaque sous-ensemble de points. En extrayant une liste de contraintes indépendantes, et en calculant le degré de restriction de cet ensemble de contraintes, on peut déterminer si la figure, qui est une solution du SCG G à résoudre, est aussi solution d’un SCG G' sous-contraint, dont l’ensemble de solutions est inclus dans celui de G . Voici les étapes d’une telle détection. Soit $\sigma(E)$ une valuation de E , solution d’un SCG G :

- (1.) Construire le matroïde en lisant les rangs sur $\sigma(E)$.
- (2.) Extraire de ce matroïde un ensemble C' de contraintes indépendantes.
- (3.) Calculer le degré de restriction de C' .

Ces trois étapes sont illustrées sur la figure 5.11, pour une solution du problème de l’hexaèdre. Le rang de chaque sous-ensemble de $\{P_0, \dots, P_7\}$ est lu sur la figure. L’ensemble de contraintes indépendantes est réduite à $\{coplanaires(P_0, P_1, P_2, P_3, P_4, P_5, P_6, P_7)\}$ (ou son énoncé en termes de rangs), qui a un degré de restriction de 5 (au lieu des 6 des contraintes de coplanarités originales). On peut alors affirmer que la solution trouvée est un point critique, et qu’il existe un ensemble de solutions de dimension 1 auquel appartient cette solution.

Cette détection a été implémentée, et testée sur le problème de l’hexaèdre. En l’appliquant à chaque solution trouvée à l’intersection du chemin d’homotopie avec l’hyperplan \mathcal{P}_1 , la courbe représentée dans la figure 5.8 (présentée dans la section précédente) est suivie en environ 5 secondes. Pour chaque solution aplatie, un continuum de solutions est trouvé.

Limites de l’approche

La principale difficulté rencontrée, mise à part l’extraction d’un ensemble de contraintes indépendantes, est la prise en compte de l’inexactitude numérique. En effet, les solutions trouvées sont approchées, de là il n’est pas facile de savoir si des points sont “presque coplanaires”, ou s’ils le sont vraiment. De plus, la taille de la structure de matroïde ($2^{|E|}$ sommets) fait que le coût de cette approche explose. Il n’est cependant pas certain qu’il soit nécessaire de calculer les rangs de tous les sous-ensembles de E pour extraire l’ensemble de contraintes indépendantes.

Remarquons qu’on pourrait aussi étudier le rang de la matrice jacobienne en chaque solution trouvée (*c.à.d.* utiliser chaque solution comme un témoin) pour déterminer sa rigidité. On est alors confronté au même problème d’inexactitude numérique.

Mentionnons enfin le fait que cette approche n'est certainement pas à même de détecter toutes les dépendances dues à des théorèmes de géométrie d'incidence. Par exemple, il n'est pas possible de démontrer le théorème de Pappus avec un matroïde en deux dimensions, mais un matroïde en 3D permet de prouver le théorème de Desargues [Michelucci 04].

5.3.3 Contraintes et constriction

L'utilisation qu'on a proposée de faire des matroïdes nécessite d'être capable d'extraire un ensemble de contraintes indépendantes d'un matroïde, en voyant ses sommets comme des contraintes. On suppose dans cette partie qu'on dispose d'un matroïde dont on connaît la fonction d'étiquetage. Précisons à quelle condition un sommet d'un matroïde est une contrainte (et est donc susceptible de faire partie d'un ensemble de contraintes indépendantes).

Un sommet F d'un matroïde en dimension d ayant pour fonction d'étiquetage f est appelé une *contrainte* si $f(F) < \min(d, |F| - 1)$. En effet, si $f(F) \geq d$ ou $f(F) = |F| - 1$, les points de F sont affinement indépendants. L'ensemble des contraintes indépendantes d'un matroïde est inclus dans l'ensemble de ses contraintes.

On propose alors une caractérisation constructive "heuristique" de l'ensemble des contraintes indépendantes d'un matroïde, qui passe par la notion suivante. On dira qu'une contrainte F telle que $f(F) = r$ est *autonome* si :

- pour tous ses fils F_i , on a $f(F_i) = r$,
- pour tous ses pères F_j , on a $f(F_j) = r + 1$.

Cette définition est justifiée de la manière suivante :

- Soit F_i un fils de F . On a par (P4), $r - 1 \leq f(F_i) \leq r$. Si $f(F_i) = r - 1$, alors encore par (P4), $f(F) \leq r$, donc la contrainte du sommet F est une conséquence de la contrainte du sommet F_i .
- Pour F_j un père de F , si $f(F_j)$ était égale à r , F serait une conséquence de F_j .

On affirme⁴ que l'ensemble des contraintes autonomes permet de déduire à partir de (A1),(A2),(A3) et (P4) toutes les étiquettes du matroïde, *c.à.d* qu'en complétant un matroïde à partir de ces contraintes indépendantes, toutes les fonctions d'étiquetage f' des matroïdes obtenus vérifient $f'(F) \leq f(F)$ pour tous $F \subseteq E$. Si cette affirmation est vraie, l'ensemble des contraintes indépendantes est inclus dans l'ensemble des contraintes autonomes.

Pour calculer le degré de restriction d'un ensemble de contraintes, on étend la notion de degré de restriction à une contrainte d'un matroïde en dimension d par la formule :

$$DoR(F) = \max(0, (|F| - 1 - rg(F))(d - rg(F)))$$

que l'on justifie de la manière suivante :

- $|F|d$ est le nombre de degrés de liberté de $|F|$ points dans un espace de dimension d ,
- $(rg(F) + 1)d$ est le nombre de degrés de liberté de $(rg(F) + 1)$ points engendrant un espace de dimension $rg(F)$,
- $rg(F)$ est le degré de liberté d'un point appartenant à un espace de dimension $rg(F)$.

Parmi les points de F :

- $(rg(F) + 1)$ points sont utilisés pour former un repère affine de l'espace de dimension $(rg(F))$ et ont d degrés de liberté chacun,
- $|F| - (rg(F) + 1)$ points appartiennent à cet espace et ont $rg(F)$ degrés de liberté chacun.

4. On pense être en mesure de le montrer.

Donc les $|F|$ points de F ont un degré de liberté de $(rg(F) + 1)d + (|F| - (rg(F) + 1))rg(F)$. En simplifiant $|F|d - (rg(F) + 1)d + (|F| - (rg(F) + 1))rg(F)$, c'est à dire le nombre de degrés de liberté de $|F|$ points en position générique moins le nombre de degrés de liberté de $|F|$ points engendrant un espace de dimension $rg(F)$, on obtient l'expression $(|F| - 1 - rg(F))(d - rg(F))$.

Dans l'ensemble $\{F_1, \dots, F_m\}$ de contraintes autonomes obtenues, il peut se produire que $F_i \subset F_j$. Calculer le degré de restriction de cet ensemble comme la somme des degrés de restriction de ses éléments amènerait à compter plusieurs fois les degrés de restriction de F_i . L'idée est alors de tenir compte de ces inclusions et de réaliser le calcul des degrés de restriction en partant des contraintes portant sur des ensembles minimaux au sens de l'inclusion, et de soustraire leurs degrés de restriction aux contraintes qui les contiennent.

Terminons par illustrer cette idée sur un exemple. Considérons l'ensemble de contraintes autonomes en dimension 3 :

$$(C1) \quad f(1, 2, 3) = 1$$

$$(C2) \quad f(1, 2, 3, 4, 5, 6) = 2$$

La contrainte (C1) a un degré de restriction de 2. Dans la contrainte (C2), on peut remplacer pour le calcul les points 1, 2, 3 par les points 1, 2, car 1, 2, 3 ne détermine qu'une droite. Si cet ensemble de contraintes comprend toutes les contraintes autonomes d'un matroïde, on sait que $f(1, 2) = 1$, sinon $f(1, 2) = 0$ appartiendrait à cet ensemble. Alors le degré de restriction de la contrainte (C2), sachant que la contrainte (C1) est satisfaite, est le degré de restriction de la contrainte $f(1, 2, 4, 5, 6) = 2$, c'est à dire 2. Le degré de restriction de l'ensemble des deux contraintes est donc 4.

5.4 Conclusion

Ce chapitre a mis en évidence le fait que les contraintes d'incidence d'un SCG amenaient le chemin d'homotopie suivi à passer par des points critiques, et, dans certains cas, à admettre des ensembles de solutions hétérogènes en dimensions. Les points critiques sont aisément évités quand ils ne sont pas solutions du SCG à résoudre, en transformant les contraintes booléennes en contraintes métriques. Les autres sont traversés par le chemin suivi.

On a également présenté une idée pour apporter une analyse géométrique, en utilisant un matroïde. Cette analyse permet de déterminer si une solution trouvée est isolée, et la configuration géométrique qu'elle présente le cas échéant. De manière moins prosaïque, elle permet aussi d'aborder le problème de la constriction générique d'un SCG en appliquant un raisonnement qui semble plus fort que les approches combinatoires de Laman, ou de constriction structurelle. En revanche, cette méthode n'est capable de détecter des dépendances qu'entre les contraintes d'incidence, de par la nature des matroïdes. Une sérieuse limitation à son utilisation pour analyser *a priori* un problème grâce à cette méthode est son coût, qui est augmenté dans les exemples étudiés ici, par leur symétrie.

De plus, on s'est borné à considérer des contraintes d'incidence comme représentants des contraintes booléennes, ignorant les contraintes d'orthogonalité, et celles de parallélisme, qui sont l'objet de nombreux théorèmes de géométrie, et pourraient tout autant amener un SCG à admettre des ensembles de solutions hétérogènes en dimensions.

Insistons sur le fait que des méthodes symboliques et numériques existent pour déterminer de tels ensembles de solutions de systèmes d'équations, et donc décider de la bonne constriction générique d'un problème; elles sont trop coûteuses pour être applicables aux exemples qu'on a présentés ici. La nôtre est applicable, mais nous pensons que dans l'état, sans trouver des stratégies

pour compléter plus efficacement un matroïde par exemple, son coût est asymptotiquement plus élevé que les méthodes purement numériques.

Chapitre 6

Décomposition

Les systèmes de contraintes géométriques provenant de la CAO impliquent souvent des dizaines de primitives et des centaines de contraintes. En 3D, ils résistent pour la plupart à une approche géométrique, et leur résolution par une méthode numérique, si toutes les solutions sont cherchées, est très coûteuse.

Utiliser une approche par décomposition en amont de la résolution ouvre la possibilité, si les sous-systèmes obtenus n'impliquent pas trop de primitives, de les résoudre tous de manière exhaustive ; bien sûr, le nombre de solutions reste exponentiel. Le principe de la décomposition d'un SCG est de découvrir ses sous-systèmes bien contraints, c'est-à-dire dont les solutions sont rigides. Cette recherche est souvent réalisée en étudiant la structure combinatoire d'un SCG et en utilisant l'approximation structurelle de la bonne constriction donnée par le critère de Laman. Chercher dans un graphe de contraintes tous les sous-systèmes structurellement bien contraints est un problème de classe NP.

Deux principales stratégies de décomposition en temps polynomial existent. La première est dite descendante (ou “top-down”) et le graphe est considéré dans son ensemble. Des paires, ou triplets en 3D, d'articulation peuvent par exemple être cherchés dans ce graphe. La seconde, dite ascendante (ou “bottom-up”), forme des clusters pour les assembler ensuite. Une des méthodes qui l'applique part d'un repère, qui est un système minimal structurellement bien contraint, et augmente ce système en y ajoutant les inconnues qui peuvent être construites par la méthode des lieux à partir des inconnues déjà construites. En répétant ce processus, des “clusters” sont formés, puis assemblés selon des motifs définis.

On propose dans ce chapitre une méthode ascendante formant des clusters qui ne sont pas nécessairement résolubles par la méthode des lieux, en adaptant un algorithme de reparamétrisation à une stratégie de décomposition. Les sous-systèmes obtenus sont structurellement bien contraints, et résolus indépendamment de la façon dont ils ont été obtenus.

La section 6.1 formalise les notions de graphes de contraintes et de leur décomposition, puis propose une adaptation de la méthode des clusters. Cette adaptation introduit la méthode de décomposition non constructive, car les sous-systèmes obtenus ne sont pas résolus, que l'on présente dans la section 6.2. Enfin, la section 6.3 aborde brièvement les problèmes de non-déterminisme de l'algorithme, de la résolution des sous-systèmes qu'elle produit et de leur assemblage.

6.1 Graphes de contraintes et décomposition

On s'intéresse dans ce chapitre à la structure combinatoire d'un SCG, c'est-à-dire à la façon dont les contraintes sont réparties entre les inconnues. Pour étudier cette structure, on considérera les degrés de liberté des inconnues et les degrés de restriction des contraintes. On rappelle que le degré de liberté d'une inconnue $x \in X$ est le nombre de coordonnées nécessaires pour la déterminer. Par exemple, une inconnue de sorte *point* a deux degrés de libertés si le SCG est défini dans un univers 2D, trois s'il l'est dans un univers 3D. On note $DoF(x)$ le degré de liberté de $x \in X$. Le degré de restriction d'une contrainte $c[\{x_i, x_j\}, \{h\}] \in C$ est le nombre de degrés de liberté qu'elle enlève à x_j (respectivement x_i) si x_i (resp. x_j) est déterminée et que c est satisfaite. Par exemple, en 3D, une contrainte de distance entre deux points $distance(P_i, P_j) = h$ où P_i est connu et $h > 0$, laisse P_j sur une sphère si elle est satisfaite. P_j a alors deux degrés de liberté au lieu de trois ; une contrainte de distance entre deux points a donc un degré de restriction de 1. On note $DoR(c)$ le degré de restriction d'une contrainte $c \in C$.

Si $\sigma(X)$ est une solution de G pour une valuation des paramètres $\sigma(A)$, toutes les figures obtenues en appliquant un déplacement à $\sigma(X)$ sont également solutions de G (pour la valuation $\sigma(A)$). Si les seules transformations par lesquelles l'image de $\sigma(X)$ est une solution de G sont les isométries, on dit que $\sigma(X)$ est rigide. Si toutes les solutions de G sont rigides, et sont en nombre fini modulo les isométries, pour une valuation générique des paramètres, on dit que G est génériquement bien contraint. Dans toute la suite, on fera abstraction de la valeur des paramètres, et on dira dans ce cas que G est bien contraint.

On considère en général que les SCG bien contraints dans un univers géométrique de dimension d vérifient les relations :

$$\sum_{x \in X} DoF(x) - \sum_{c \in C} DoR(c) = \binom{d+1}{2}$$

et, pour tout sous-ensemble de contraintes $C'[X', A'] \subseteq C[X, A]$, où X' est l'ensemble des inconnues impliquées par C' ,

$$\sum_{x \in X'} DoF(x) - \sum_{c \in C} DoR(c) \geq \binom{d+1}{2},$$

où la constante $\binom{d+1}{2}$ correspond aux degrés de liberté en translations et en rotations d'une figure. C'est aussi le nombre minimum de degrés de liberté qui doit être fixé par un repère.

Ces relations sont appelées critère de Laman. Quand $d = 2$ et que X ne contient que des points et C des distances, elles permettent de caractériser les SCG génériquement bien contraints. Dans les autres cas, notamment quand $d = 3$, elles n'expriment qu'une condition nécessaire à la bonne constricton d'un SCG, mise en défaut pour certains SCG¹. Elles ne s'étendent pas non plus aux cas où le système implique d'autres objets et contraintes que des points et des distances. On dira d'un SCG qu'il est bien contraint au sens de Laman s'il satisfait ces relations.

Les exemples donnés dans ce chapitre le sont dans un univers géométrique 2D, détaillé dans l'annexe B, ainsi que son interprétation par des degrés de liberté.

6.1.1 Graphe de contraintes

Un SCG $G = C[X, A]$ peut être représenté par un graphe de contraintes $\mathcal{G} = \langle X, C \rangle$ dans lequel les sommets sont les éléments de X , et les arêtes, non orientées, sont les contraintes de C . Afin

1. Voir [Jermann 02], page 62 pour plus de détails.

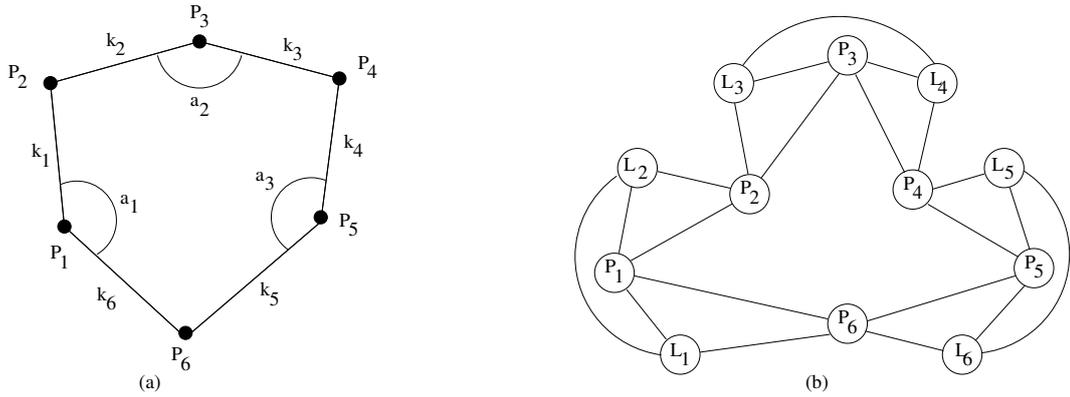


FIGURE 6.1 – (a) l’esquisse cotée d’un SCG en 2D. (b) le graphe de contraintes associé.

d’éviter le formalisme des hyper-graphes pour simplifier le discours, on supposera que toutes les contraintes de C portent uniquement sur deux éléments de X , et ont un degré de restriction de 1.

On dira également qu’un graphe de contraintes \mathcal{G} est bien contraint si le SCG G qu’il représente est bien contraint au sens de Laman. De plus si $\sigma(X)$ est une solution de G (pour une valuation quelconque des paramètres), on dira que c’est une solution de \mathcal{G} .

Soit $x \in X$, on notera le degré du sommet x dans le graphe. En considérant uniquement des contraintes de degré de restriction 1, $D(x)$ est donc le nombre de contraintes impliquant x dans C . On considérera ici que toutes les inconnues d’un SCG ont un degré de liberté de d , où $d = 2$ ou 3 est la dimension de l’univers géométrique dans lequel est défini le SCG.

La figure 6.1(a) représente l’esquisse cotée d’un SCG en 2D avec six points, six distances et trois contraintes d’angle. Les contraintes d’angles étaient jusque là exprimées comme impliquant trois points, par exemple $angle(P_6, P_1, P_2)$ pour l’angle $\widehat{P_6 P_1 P_2}$. Pour rentrer dans le cadre posé ici et éviter d’utiliser des hyper-arêtes, elles sont exprimées comme des angles entre deux droites, ajoutées à l’ensemble des inconnues, ainsi que des contraintes d’incidences point-droite permettant de les déterminer. La contrainte $angle(P_6, P_1, P_2)$ est ainsi traduite par les quatre contraintes d’incidence points-droites $incidents(P_1, L_1), \dots, incidents(P_2, L_2)$, et la contrainte d’angle entre deux droites $angle(L_1, L_2)$. Chacune de ces contraintes correspond alors à une arête dans le graphe de contraintes, représenté dans la figure 6.1(b). De la même façon, dans un univers 3D, une contrainte de coplanarité entre quatre points est traduite par 4 contraintes d’incidences point-plan, et un objet de sorte *plan* est ajouté à l’ensemble des inconnues.

Soit $\mathcal{G} = \langle X, C \rangle$ un graphe de contraintes. On notera $\mathcal{E}[\mathcal{G}]$ l’ensemble de ses sommets (donc l’ensemble de ses inconnues) et $\mathcal{V}[\mathcal{G}]$ l’ensemble de ses arêtes (donc l’ensemble de ses contraintes). Si $X_1 \subseteq X$, on note $\mathcal{G}[X_1]$ le graphe induit dont les arêtes sont celles de \mathcal{G} ayant leurs deux extrémités dans X_1 , et dont l’ensemble des sommets est X_1 . On définit de manière similaire le sous-graphe induit par un ensemble d’arêtes $C_1 \subseteq C$ dont les sommets sont ceux de \mathcal{G} qui sont incident à au moins une arête de C_1 et dont l’ensemble des arêtes est C_1 . Le sous-graphe induit par l’ensemble d’arêtes C_1 est noté $\mathcal{G}[C_1]$.

Dans un graphe de contraintes $\mathcal{G} = \langle X, C \rangle$ d’un SCG où $d = 2$ où 3 est la dimension de l’univers géométrique, on appellera repère un ensemble de d sommets reliés entre eux par d arêtes (en supposant que dans \mathcal{G} , chaque paire de sommets est relié par au plus une arête). Dans le graphe de la figure 6.1, par exemple, $\mathcal{G}[\{P_1, L_1\}]$ est un repère. Dans un graphe d’un SCG en 3D,

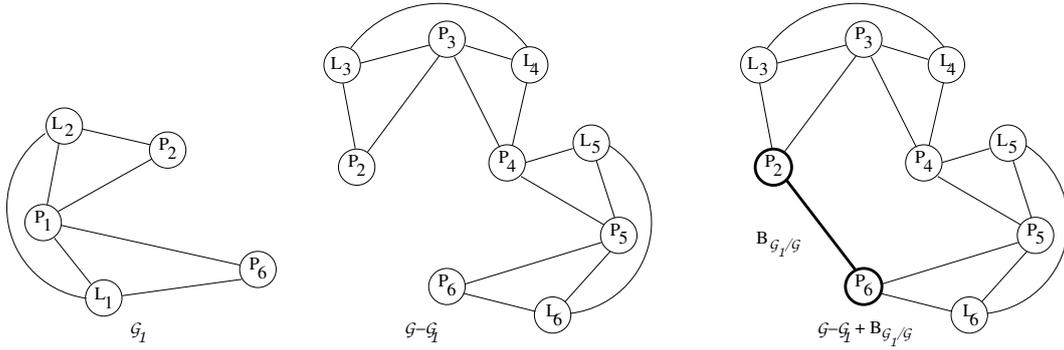


FIGURE 6.2 – À gauche un sous-graphe \mathcal{G}_1 bien contraint du graphe \mathcal{G} de la figure 6.1. Au centre le graphe $\mathcal{G} - \mathcal{G}_1$, qui est sous-contraint. À droite, $\mathcal{G} - \mathcal{G}_1$ augmenté du bord de \mathcal{G}_1 .

un repère est par exemple un point incident à une droite, tous deux incidents à plan. Remarquons que si X_1 est l'ensemble des sommets d'un repère, le SCG correspondant au graphe de contraintes $\mathcal{G}[X_1]$ est rigide modulo les déplacements, et peut être résolu par la méthode d'intersection des lieux géométriques.

6.1.2 Bords et décomposition

Soient $\mathcal{G} = \langle X, C \rangle$ un graphe de contraintes, et $\mathcal{G}_1 = \langle X_1, C_1 \rangle$ un sous-graphe de \mathcal{G} .

La différence $\mathcal{G} - \mathcal{G}_1$ est le graphe $\mathcal{G}[C_2]$ où $C_2 = \mathcal{V}[\mathcal{G}[C \setminus C_1]]$. Pour deux sous-graphes $\mathcal{G}_1 \langle X_1, C_1 \rangle$ et $\mathcal{G}_2 \langle X_2, C_2 \rangle$ de \mathcal{G} , on définit l'intersection $\mathcal{G}_{1 \cap 2} = \mathcal{G}[X_1 \cap X_2]$.

On sait ([Mathis 10]) que si \mathcal{G} et \mathcal{G}_1 sont bien-contraints, il en est de même pour $\mathcal{G}_2 = \mathcal{G} - \mathcal{G}_1$ si et seulement si $\mathcal{G}_{1 \cap 2}$ est bien contraint et non réduit à un point en 2D et à deux points en 3D. La figure 6.2 représente un sous-graphe \mathcal{G}_1 du graphe \mathcal{G} donné dans la figure 6.1(b) et, au centre, $\mathcal{G}_2 = \mathcal{G} - \mathcal{G}_1$. Dans cet exemple, le graphe $\mathcal{G}_{1 \cap 2}$ est le graphe sans arête formé des deux sommets P_2 et P_6 ; il n'est pas bien contraint, il en découle que \mathcal{G}_2 ne l'est pas non plus.

Avec ces notations, on appelle sommet intérieur d'un sous-graphe \mathcal{G}_1 un sommet de $\mathcal{G}_1 - \mathcal{G}_{1 \cap 2}$.

Si \mathcal{G} et \mathcal{G}_1 sont bien-contraints et que $\mathcal{G}_{1 \cap 2}$ ne l'est pas directement, on peut le compléter par des informations tirées de \mathcal{G}_1 . Par exemple, dans la figure 6.2, supposons qu'on connaisse une solution $\sigma(X_1)$ de \mathcal{G}_1 ; on peut alors lire sur $\sigma(X_1)$ la valuation $\sigma(h)$ du paramètre h d'une contrainte métrique $c[\{P_2, P_6\}, \{h\}]$, par exemple $distance(P_2, P_6) = h$, et ajouter une arête entre P_2 et P_6 à $\mathcal{G} - \mathcal{G}_1$ pour qu'il soit bien contraint. Cette arête ajoutée correspond dans ce cas à l'"arête virtuelle" dans la méthode de [Owen 91].

Le graphe de contrainte $\mathcal{G}_{1 \cap 2}$, augmenté des arêtes qui font de lui un graphe bien contraint, est appelé le *bord* de \mathcal{G}_1 dans \mathcal{G} , et noté $B_{\mathcal{G}_1/\mathcal{G}}$. On notera alors $\mathcal{G} - \mathcal{G}_1 + B_{\mathcal{G}_1/\mathcal{G}}$ le graphe dont les sommets sont $\mathcal{V}[\mathcal{G} - \mathcal{G}_1]$ et les arêtes sont $\mathcal{E}[\mathcal{G} - \mathcal{G}_1] \cup \mathcal{E}[B_{\mathcal{G}_1/\mathcal{G}}]$. La figure 6.2 représente $\mathcal{G} - \mathcal{G}_1 + B_{\mathcal{G}_1/\mathcal{G}}$; on y a marqué $B_{\mathcal{G}_1/\mathcal{G}}$ en traits gras.

On appellera *décomposition* d'un graphe de contraintes \mathcal{G} et du SCG correspondant une suite finie $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n$ où :

- \mathcal{G}_1 est un sous-graphe de \mathcal{G} bien contraint,
- $\mathcal{G}_2, \dots, \mathcal{G}_n$ est une décomposition de $\mathcal{G} - \mathcal{G}_1 + B_{\mathcal{G}_1/\mathcal{G}}$.

Sur la figure 6.2, $\mathcal{G}_1, \mathcal{G}_2 = \mathcal{G} - \mathcal{G}_1$ est une décomposition du graphe donné dans la figure 6.2(b). Il faut noter qu'une telle décomposition n'est en général pas unique. On dit qu'elle est maximale lorsque la longueur de la suite de sous-graphes est maximale.

6.1.3 Une décomposition constructive

On présente ici, dans le but de faciliter l'explication d'une méthode plus efficiente en section 6.2, une méthode de décomposition constructive, au sens où elle résout un SCG en même temps qu'elle le décompose. Elle peut être vue comme une extension de la méthode de [Bouma 95], puisqu'elle procède aussi en trouvant des clusters. Ces clusters ne sont cependant pas résolus indépendamment puis ré-assemblés selon des motifs prédéfinis, mais sont résolus séquentiellement et assemblés en faisant coïncider leurs bords.

Cette approche utilise la méthode des lieux LIM, déjà présentée dans le chapitre 1, qui construit itérativement les objets des solutions d'un SCG en intersectant des lieux géométriques. Si toutes les solutions d'un SCG peuvent être construites par une telle suite d'intersections de lieux géométriques, une propagation des degrés de liberté dans le graphe de contraintes $\mathcal{G} = \langle X, C \rangle$ correspondant fournit cette suite de constructions.

La démarche décrite ici est toutefois compatible avec une résolution géométrique plus sophistiquée.

Propagation des degrés de liberté

La propagation des degrés de liberté peut être réalisée selon deux stratégies. La première, appelée rétro-propagation procède par déconstruction du graphe : un des sommets x satisfaisant $D(x) = DoF(x)$ est retiré du graphe avec les arêtes ayant ce sommet comme extrémité. Si après avoir répété cette étape autant de fois qu'il est possible, le graphe est réduit à un repère, la résolution est un succès.

La seconde stratégie, ou propagation avant, consiste à choisir dans \mathcal{G} un repère, et à marquer les sommets correspondants comme construits. Puis, si un sommet x est adjacent à $DoF(x)$ sommets déjà construits, il est marqué lui aussi. Cette opération est appliquée jusqu'à ce qu'aucun nouveau sommet ne puisse être marqué. Lorsque tous les sommets sont marqués, le SCG est résolu.

Ces deux stratégies permettent de résoudre les mêmes problèmes et traduisent une constructibilité symbolique par LIM. L'algorithme 2, où $DM(x)$ représente le nombre de sommets marqués adjacents à un sommet x , implante la propagation avant des degrés de liberté.

Algorithme 2 Propagation avant des degrés de liberté

Entrées: $\mathcal{G} = \langle X, C \rangle$: graphe de contraintes non-vide

- 1: Choisir un repère et marquer ses sommets
 - 2: Choisir un sommet non-marqué x tel que $DM(x) = DoF(x)$, marquer x
 - 3: Retourner à 2 tant qu'il est possible de choisir un sommet
-

En appliquant l'algorithme 2 au graphe \mathcal{G} de la figure 6.1, le repère choisi est par exemple la paire de sommets $\{P_1, L_1\}$. P_6, L_2 , puis P_2 sont marqués, après quoi aucun sommet non marqué x ne satisfait $DM(x) = DoF(x)$, et l'algorithme s'arrête. L'ensemble des sommets marqués est alors $\mathcal{V}[\mathcal{G}_1]$, où \mathcal{G}_1 est le graphe représenté dans la partie gauche de la figure 6.2. Il n'est aucun choix de repère qui permette à l'algorithme 2 de marquer tous les sommets de \mathcal{G} .

Notons que la réussite de l’algorithme 2 dépend du repère choisi dans l’étape 1 : pour un même graphe, choisir un repère peut amener à marquer tous ses sommets, alors que l’algorithme s’arrêtera sans que tous les sommets ne soient marqués avec un autre repère.

Décomposition en sous-systèmes résolubles par LIM

Comme dans [Bouma 95], on peut utiliser la propagation avant des degrés de libertés pour obtenir une décomposition d’un graphe de contrainte $\mathcal{G} = \langle X, C \rangle$, en appliquant la méthode suivante, qu’on explique sur l’exemple de la figure 6.1.

La première étape consiste à choisir un repère, et à appliquer à \mathcal{G} l’algorithme 2. Quand il s’arrête, on dispose d’un ensemble $X_1 \subseteq X$ de sommets marqués. Si $X_1 = X$, le SCG correspondant à \mathcal{G} est résolu par LIM. Sinon, si $\mathcal{G}_1 = \mathcal{G}[X_1]$ contient au moins un sommet interne, $\mathcal{G} - \mathcal{G}_1 + B_{\mathcal{G}_1/\mathcal{G}}$ contient moins de sommets que \mathcal{G} . On peut alors recommencer, en appliquant l’algorithme 2 à \mathcal{G}_1 . Enfin, si \mathcal{G}_1 ne contient aucun sommet interne, $\mathcal{G} - \mathcal{G}_1 + B_{\mathcal{G}_1/\mathcal{G}}$ contient autant de sommets que \mathcal{G} . On choisit alors un autre repère pour appliquer l’algorithme 2, jusqu’à obtenir un ensemble X_1 tel que $\mathcal{G}[X_1]$ contienne au moins un sommet interne, où jusqu’à avoir essayé tous les repères possibles.

En appliquant cette méthode au graphe de la figure 6.1(b), le repère choisi initialement pour appliquer l’algorithme 2 est par exemple formé de la paire $\{P_1, L_1\}$. Le graphe \mathcal{G}_1 obtenu est donné dans la figure 6.2. Il contient les deux sommets internes L_1 et L_2 . En le remplaçant par son bord $B_{\mathcal{G}_1/\mathcal{G}}$ dans \mathcal{G} , on obtient le graphe $\mathcal{G} - \mathcal{G}_1 + B_{\mathcal{G}_1/\mathcal{G}}$ donné dans la partie droite de la figure 6.2. On peut alors recommencer en appliquant l’algorithme 2 à ce graphe.

Cette méthode permet d’obtenir une décomposition $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n$ de \mathcal{G} avec, éventuellement, $\mathcal{G}_n = \mathcal{G}$ si \mathcal{G} est résolvable par LIM, ou si aucun choix de repère ne permet d’obtenir par l’algorithme 2 un graphe \mathcal{G}_1 contenant un sommet interne.

On considère dans un premier temps le cas où \mathcal{G}_n est un sous-graphe au sens strict de \mathcal{G} . Alors chaque graphe \mathcal{G}_i pour $i < n$ est résolvable par LIM, et si \mathcal{G}_n l’est aussi, il est facile d’obtenir un plan de construction pour les solutions de chaque graphe \mathcal{G}_i de cette décomposition, et un plan d’assemblage découle naturellement de la décomposition. Si \mathcal{G}_n n’est pas résolvable par LIM, on considère que la méthode échoue ; remarquons que \mathcal{G}_n peut être résolu à part, par exemple par une méthode numérique. Si $\mathcal{G}_n = \mathcal{G}$ et que \mathcal{G}_n n’est pas résolvable par LIM, \mathcal{G} n’a pas été décomposé.

6.2 Une décomposition non constructive

La figure 6.3, à gauche, présente un graphe de contraintes qui n’est pas décomposable par la méthode de décomposition donnée dans la section 6.1, car quel que soit le repère choisi, l’ensemble des sommets marqués par l’algorithme 2 ne contient aucun sommet interne. On présente dans cette section une méthode qui trouve une décomposition de ce graphe en un temps polynomial. Elle utilise la propagation avant des degrés de liberté, mais l’adapte pour découvrir des sous-graphes bien contraints non nécessairement résolubles par LIM. Quand pour aucun repère, l’algorithme 2 ne parvient à marquer de sommets internes, des contraintes sont “virtuellement” ajoutées pour parvenir à marquer des sommets ; le graphe induit par les sommets marqués (en ignorant les contraintes ajoutées) est alors sous-contraint. Des sommets adjacents à des sommets marqués peuvent ensuite être ajoutés même s’ils sont reliés à ces derniers par plus d’arêtes qu’ils n’ont de degrés de liberté. Les arêtes en surplus sont virtuellement supprimées. Le graphe correspondant au graphe initial, dans lequel on remplace les contraintes virtuellement supprimées par celles virtuellement ajoutées est résolvable par l’algorithme 2. En ceci, cette méthode adapte la phase symbolique de la méthode

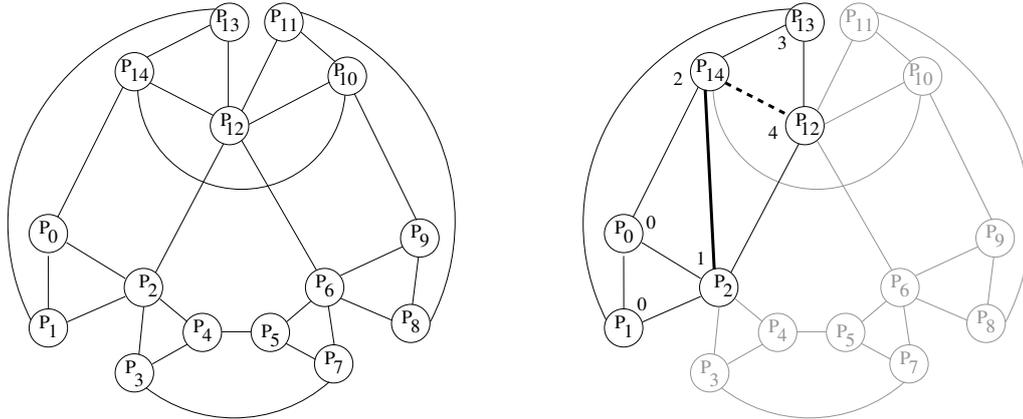


FIGURE 6.3 – À gauche le graphe de contraintes d'un SCG en 2D. À droite le résultat de l'application de la propagation avant des degrés de liberté avec remplacement virtuel de contraintes. Les nombres à côté des sommets indiquent l'ordre dans lesquels ils sont marqués, l'arête en trait gras est virtuellement ajoutée, celle en trait gras pointillé est virtuellement supprimée.

par reparamétrisation; cependant, la résolution des sous-systèmes est indépendante de la façon dont ils ont été obtenus. On peut en trouver des solutions par toute méthode numérique ou symbolique, comme la méthode de Newton-Raphson, ou la méthode d'homotopie présentée dans le chapitre 3.

6.2.1 Reparamétrisation en propagation avant

L'algorithme 3 implémente la propagation avant des degrés de libertés avec remplacement virtuel d'arêtes. Comme la propagation avant des degrés de libertés, il augmente un ensemble de sommets dont les primitives géométriques correspondantes peuvent être construites par la méthode LIM, à partir d'un repère initialement choisi. Si à une étape, il ne reste aucun sommet non marqué dont le degré dans le graphe est égal au degré de liberté, des contraintes sont virtuellement ajoutées ou supprimées pour marquer un sommet.

Algorithme 3 Propagation avant avec remplacement virtuel de contraintes

Entrées: $\mathcal{G} = \langle X, C \rangle$: graphe de contraintes non-vide

- 1: Choisir un repère et marquer ses sommets
 - 2: Choisir un sommet non encore marqué x tel que $DM(x) > 0$ et $DM(x) - DoF(x)$ est maximum
 - 3: **Si** $DM(x) < DoF(x)$
 ajouter virtuellement $DoF(x) - DM(x)$ contraintes
 Si $DM(x) > DoF(x)$
 enlever virtuellement $DM(x) - DoF(x)$ contraintes
 - 4: Marquer x
 - 5: Soit X_m l'ensemble des sommets marqués
 Arrêter si $\mathcal{G}[X_m]$ est bien contraint et contient un sommet interne
 sinon retourner à 2
-

On va montrer que dans le point 5, la condition $\mathcal{G}[X_m]$ est bien contraint équivaut au fait que le nombre d'arêtes virtuellement ajoutées est égal au nombre d'arêtes virtuellement supprimées si \mathcal{G} est bien contraint. Cette remarque simplifie le test de la condition d'arrêt.

Soit donc un graphe bien contraint $\mathcal{G} = \langle X, C \rangle$; l'étape 1 marque les sommets d'un repère X_m . On a alors nécessairement $DM(x) \leq DoF(x)$ pour tous les sommets x non marqués adjacents à X_m . En effet, si x non marqué vérifie $DM(x) > DoF(x)$ alors le graphe $\mathcal{G}[X_m \cup \{x\}]$ est sur-contraint ce qui contredit l'hypothèse de bonne constriction de \mathcal{G} . Si il existe un sommet x tel que $DM(x) = DoF(x)$, il est marqué.

Considérons à présent la première étape où il n'y a aucun sommet non marqué x satisfaisant $DM(x) = DoF(x)$. En utilisant le même raisonnement que précédemment, on a nécessairement $DM(x) < DoF(x)$ pour tous les sommets x non marqués. Soit alors x un sommet connecté à X_m , et $a_m = DoF(x) - DM(x)$ le nombre de contraintes virtuellement ajoutées pour marquer x . En choisissant x qui maximise $DM(x) - DoF(x)$ (donc qui minimise a_m), on a choisi un des sommets les plus connectés.

Plaçons-nous maintenant dans le cas où il existe un sommet non marqué x adjacent aux sommets marqués avec $DM(x) > DoF(x)$ pour la première fois. Soit $a_{m'}$ le nombre total de contraintes virtuellement ajoutées à toutes les étapes précédentes, et $X_{m'}$ l'ensemble des sommets marqués. Alors on a $DM(x) - DoF(x) \leq a_{m'}$. En effet, $\mathcal{G}[X_m]$ augmenté des arêtes virtuellement ajoutées est bien contraint. Si on avait $DM(x) - DoF(x) > a_{m'}$ alors $\mathcal{G}[X_m \cup \{x\}]$ serait sur-contraint, ce qui contredirait l'hypothèse de bonne constriction de \mathcal{G} .

Enfin, supposons qu'à une étape les nombres d'arêtes virtuellement ajoutées et supprimées soient égaux. Soit X_m l'ensemble des sommets marqués. On appelle $\mathcal{G}[X_m]'$ le graphe résultant des suppressions et ajouts virtuels d'arêtes; il est bien contraint (car obtenu par une propagation des degrés de liberté). Alors $\mathcal{G}[X_m]$ est bien contraint car il contient autant d'arêtes que $\mathcal{G}[X_m]'$ et que, par hypothèse de bonne constriction de \mathcal{G} , aucun de ses sous-graphes n'est sur-contraint.

Réciproquement, supposons que le graphe $\mathcal{G}[X_m]$ induit par l'ensemble des sommets marqués soit bien contraint, et contienne c arêtes. Comme $\mathcal{G}[X_m]'$ est également bien contraint (car obtenu par une propagation des degrés de liberté), il contient aussi c arêtes, et le nombre d'arêtes ajoutées virtuellement est égal à celui des arêtes supprimées virtuellement.

En appliquant l'algorithme 3 au graphe $\mathcal{G} = \langle X, C \rangle$ donné dans la partie gauche de la figure 6.3 avec comme repère initial $\{P_0, P_1\}$, le point P_2 est construit. $\mathcal{G}[\{P_0, P_1, P_2\}]$ est bien contraint, mais ne contient aucun sommet interne. Tous les sommets x adjacents à $\{P_0, P_1, P_2\}$ sont tels que $DM(x) < DoF(x)$, et on choisit par exemple P_{14} , qu'on pourrait construire si l'arête P_2P_{14} appartenait à C . En ajoutant virtuellement cette arête, P_{14} est marqué. P_{13} est un des deux sommets adjacents aux sommets marqués maximisant $DM(x) - DoF(x)$, et il peut être marqué car $DM(P_{13}) - DoF(P_{13}) = 0$. Enfin, P_{12} maximise $DM(x) - DoF(x)$, et est construit en supprimant virtuellement, par exemple, l'arête $P_{12}P_{14}$. Comme $\mathcal{G}[\{P_0, P_1, P_2, P_{12}, P_{13}, P_{14}\}]$ est bien contraint et contient un sommet interne, l'algorithme 3 s'arrête. La partie droite de la figure 6.3 présente le sous-graphe obtenu.

L'algorithme 3 n'est pas déterministe : différents choix de sommets peuvent aboutir à l'obtention de différents sous-systèmes bien contraints.

Notons enfin que si l'on y remplace la condition d'arrêt du point 5 par :

5 : retourner à 2

cet algorithme est équivalent à l'algorithme de reparamétrisation en propagation arrière de [Gao 02], dans le sens où le graphe résultant des remplacements virtuels d'arêtes est résoluble par LIM.

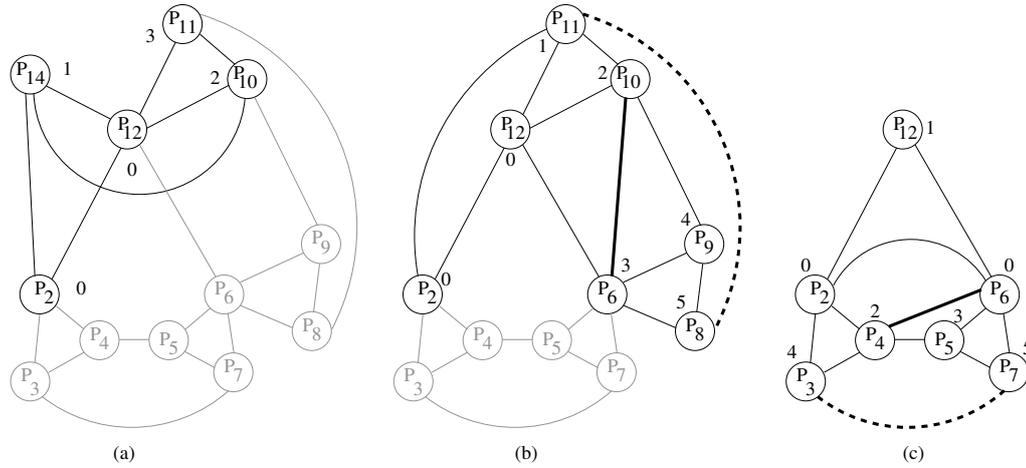


FIGURE 6.4 – Trois étapes de l’application de la méthode de décomposition non constructive. Pour chaque étape, les sommets marqués sont représentés en noir, et ceux non marqués en gris. L’ordre dans lequel les sommets sont marqués est indiqué à côté de chaque sommet.

6.2.2 Algorithme de décomposition

On donne à présent notre méthode de décomposition, dont le principe est d’appliquer au maximum la propagation avant des degrés de liberté, puis d’utiliser le remplacement virtuel de contraintes quand cette dernière échoue. Elle se traduit plus précisément en l’algorithme 4.

Algorithme 4 Décomposition

Entrées: $\mathcal{G} = \langle X, C \rangle$: graphe de contraintes non-vide

Tant que tous les sommets ne sont pas marqués **faire**

- Appliquer l’algorithme 2 sur \mathcal{G} autant que possible.

Pour chaque sous-système \mathcal{G}_1 obtenu contenant des sommets internes, remplacer \mathcal{G}_1 par son bord dans \mathcal{G} .

- Appliquer l’algorithme 3 sur \mathcal{G} .

Remplacer le graphe \mathcal{G}_1 obtenu par son bord dans \mathcal{G} .

fin Tant que

Son application à l’exemple de la figure 6.3 se déroule comme suit. L’algorithme 2 ne permet, pour aucun repère, de marquer un ensemble de sommets avec un sommet interne. Quand l’algorithme 3 utilise comme repère $\{P_0, P_1\}$, le graphe \mathcal{G}_1 obtenu est celui donné dans la partie droite de la figure 6.3, qui contient trois sommets internes. Son bord est formé des sommets P_2, P_{12} , et P_{14} . L’arête P_2P_{14} est ajoutée à $\mathcal{G}[\{P_2, P_{12}, P_{14}\}]$ pour obtenir $B_{\mathcal{G}_1/\mathcal{G}}$ bien contraint, et $\mathcal{G} - \mathcal{G}_1 + B_{\mathcal{G}_1/\mathcal{G}}$ est donné dans la figure 6.4(a) (graphe dans son ensemble).

L’algorithme 2, avec $\{P_2, P_{12}\}$ comme repère, marque les sommets P_{14}, P_{10} , puis P_{11} (voir le graphe en noir dans la figure 6.4(a)). P_{14} étant un sommet interne, le graphe obtenu peut être remplacé par son bord, et l’application de l’algorithme 2 ne permet plus de marquer un ensemble de sommets contenant un sommet interne. Le graphe à l’issue de cette étape est représenté dans la figure 6.4(b) (graphe dans son ensemble).

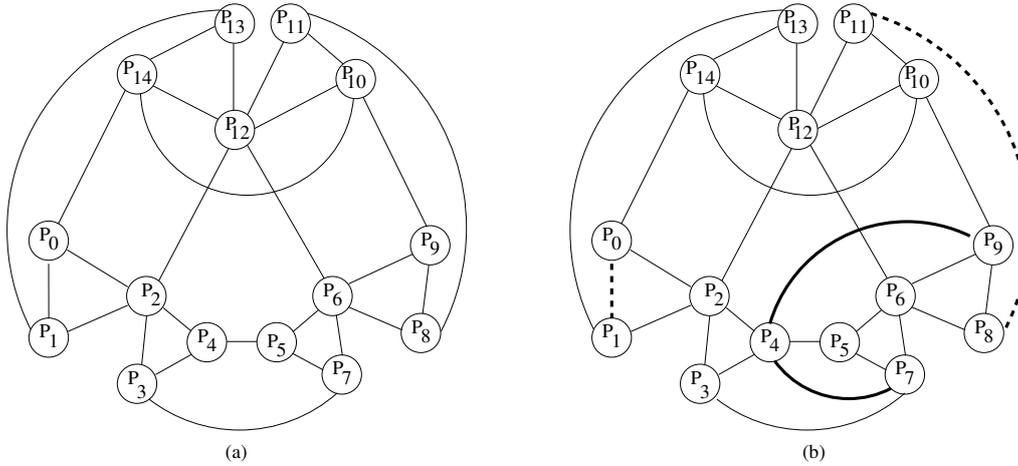


FIGURE 6.5 – (a) un graphe de contrainte. En appliquant l’algorithme 3 avec comme repère P_4P_5 , le graphe bien contraint (b) est obtenu en ajoutant virtuellement les arêtes en traits gras, et en retirant celles en pointillés.

Appliqué deux fois de suite avec pour repères $\{P_2, P_{12}\}$ puis $\{P_2, P_6\}$, l’algorithme 3 permet de marquer tous les sommets et de terminer la décomposition, comme illustré dans les figures 6.4(b) et 6.4(c).

6.2.3 Améliorer la décomposition

Cette méthode n’est pas déterministe, et la décomposition obtenue dépend fortement des repères choisis à chaque étape. Par exemple, en choisissant dans l’exemple de la figure 6.3(a) (rappelé sur la figure 6.5(a)) le repère formé des deux sommets P_4 et P_5 , le premier graphe bien contraint marqué par l’algorithme 3 est le graphe en entier. La figure 6.5(b) présente en traits gras les arêtes virtuellement supprimées (en pointillés) et ajoutées (en traits pleins). L’algorithme 4 ne parvient donc pas à décomposer le graphe.

On présente ici une méthode heuristique pour améliorer cette décomposition. Supposons que \mathcal{G} soit un système bien contraint avec une décomposition maximum en deux sous-graphes \mathcal{G}_1 et \mathcal{G}_2 , chacun contenant des sommets internes et chacun pouvant être obtenu par l’ajout virtuel d’une seule contrainte. Le raisonnement suivant se généralise à davantage de sous-graphes. L’objectif est ici de trouver la répartition des contraintes virtuellement ajoutées dans les deux sous-systèmes.

Nous savons que \mathcal{G}_1 est bien-contraint. Par ailleurs, \mathcal{G}_2 est bien contraint si $\mathcal{G}_{1 \cap 2}$ est bien contraint ; il est sous-contraint sinon. Bien entendu, l’algorithme 4 démarre sans connaître cette décomposition. Idéalement, si un repère est fixé dans \mathcal{G}_1 , on peut espérer isoler \mathcal{G}_1 . Mais supposons qu’un repère soit fixé dans \mathcal{G}_2 et soit $\mathcal{G}f$ le premier sous-graphe déterminé par 4, nous avons alors deux cas :

- soit $\mathcal{G}_{1 \cap 2}$ est bien contraint, dans ce cas $\mathcal{G}f = \mathcal{G}_2$,
- soit au contraire $\mathcal{G}_{1 \cap 2}$ n’est pas bien-contraint, ce qui est le cas le plus courant en CAO, et alors $\mathcal{G}f = \mathcal{G}$

En effet, dans ce dernier cas, comme \mathcal{G}_2 est sous-contraint, toute propagation démarrante à l’intérieur de \mathcal{G}_2 ne peut s’arrêter qu’après avoir marqué tous les sommets de \mathcal{G}_1 pour produire le sous-graphe bien contraint $\mathcal{G}f$. Ainsi, si des contraintes ont été ajoutées dans \mathcal{G}_2 , la dernière

contrainte retirée appartenait à \mathcal{G}_1 . Cette remarque conduit à l’heuristique suivante : si deux (ou plus) contraintes sont virtuellement ajoutées dans un sous-graphe \mathcal{G}_f , l’algorithme 4 est relancé sur \mathcal{G}_f mais en partant d’un nouveau repère impliquant la dernière contrainte virtuellement retirée. Alors, le marquage des sommets a de meilleures chances d’isoler \mathcal{G}_1 comme composante rigide.

Dans l’exemple de la figure 6.5, cette méthode amène à reconsidérer la décomposition car deux contraintes ont été virtuellement ajoutées. Comme l’arête P_0P_1 a été virtuellement supprimée, l’algorithme 3 est relancé avec comme repère les sommets P_0 et P_1 , et la décomposition présentée à la sous-section précédente est obtenue.

6.2.4 Terminaison et complexité

À chaque itération, l’algorithme extrait un sous-graphe et le remplace par son bord. Soient n et e le nombre de sommets et d’arêtes du graphe de contraintes. L’algorithme 2 met en œuvre $O(n^2 + e)$ opérations s’il résout le problème. Dans chaque boucle, les sommets restants sont parcourus, un sommet x est ôté et ses voisins sont mis à jour. La boucle est itérée n fois et la mise à jour des voisins nécessite de visiter toutes les arêtes adjacentes. L’algorithme 2 échoue quand tous les repères possibles ont été essayés. Dans le plan, le nombre de repères est en $O(e)$ et pour chacun d’eux tous les sommets sont parcourus. Dans ce cas, le nombre d’opérations effectuées est en $O(ne)$. Or dans un graphe de contraintes e est proportionnel à n , on a $e = O(n)$ et l’ensemble du processus donne tous les sous-graphes possibles avec un coût de $O(n^2)$. Dans l’espace, un repère est formé par une combinaison de trois contraintes donc le nombre de repères est au plus en $O(e^3)$, donc l’algorithme a un coût en $O(n^3)$. Remarquons qu’en 3D, un repère peut contenir une contrainte ajoutée quand aucun repère ne peut être extrait du graphe de contrainte original.

Le bord est calculé en ajoutant des contraintes entre les primitives du bord. En 2D, un bord structurellement bien contraint peut être produit par des constructions d’Henneberg (voir [Gao 09]). Dans l’espace, [Thierry 11] présente un algorithme glouton pour déterminer ce bord : la méthode du témoin permet de détecter dans un système un sous-système non sur-contraint. En construisant à partir du bord un système saturé (par exemple en liant tous les couples d’objets par une distance), un bord bien contraint est déterminé en utilisant la méthode du témoin pour enlever les contraintes sur-numéraires.

L’algorithme 3 diffère d’un algorithme de propagation avant des degrés de liberté uniquement par le fait que si aucun sommet ne peut être construit, des contraintes virtuelles sont ajoutées, et les coûts de ces deux algorithmes sont les mêmes. En appliquant l’heuristique de la sous-section précédente, l’algorithme 3 est appelé $k < n$ fois où k est le nombre de contraintes virtuellement supprimées. Donc, dans le pire des cas, l’algorithme général a un coût en $O(n^4)$ en 3D, et $O(n^3)$ en 2D.

6.3 Perspectives : résolution et assemblage

On termine ce court chapitre en soulevant deux problèmes liés à la méthode de décomposition proposée.

Comme il a déjà été dit, elle n’est pas déterministe, et la décomposition obtenue dépend fortement des repères choisis à chaque étape, et des sommets marqués. En l’appliquant au SCG d’un univers 3D donné dans la figure 6.6, elle peut fournir, entre autres, les deux décompositions données dans la figure 6.7. Celle de gauche peut sembler préférable du point de vue de la complexité de la résolution des sous-systèmes. Notons également qu’un “mauvais choix” de repère initial, par exemple celui formé des points P_0, P_1 et du plan passant par P_0, P_1, P_2, P_3 ne permet pas de

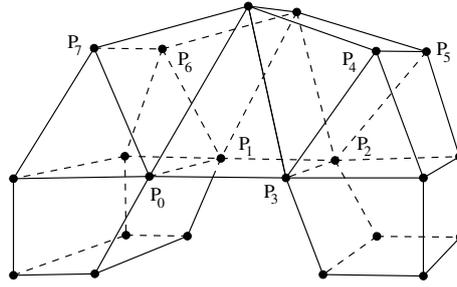


FIGURE 6.6 – Un SCG en 3D. Les arêtes représentent des contraintes de distance. 15 contraintes de coplanarité ne sont pas indiquées : une pour chaque face des deux hexaèdres, et les trois suivantes : $\text{coplanaires}(P_0, P_1, P_2, P_3)$, $\text{coplanaires}(P_0, P_1, P_6, P_7)$ et $\text{coplanaires}(P_2, P_3, P_4, P_5)$.

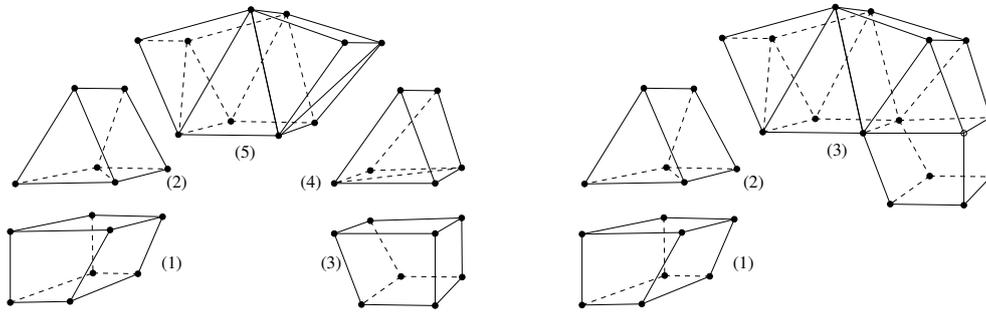


FIGURE 6.7 – Deux décompositions du SCG de la figure 6.6. Les nombres entre parenthèses indiquent l'ordre dans lequel ont été trouvés les sous-systèmes, et l'ordre dans lequel ils sont résolus.

décomposer le SCG : le premier graphe bien contraint trouvé par l'algorithme 3 est le graphe dans son ensemble. L'heuristique qu'on a proposée permet alors d'obtenir la décomposition présentée dans la partie gauche de la figure 6.7. On ne peut cependant pas garantir qu'une décomposition maximale est obtenue.

Le second problème concerne la résolution des sous-systèmes et leur assemblage. Si chaque système peut être résolu indépendamment, dans l'ordre imposé par la décomposition, leur assemblage nécessite que les bords des solutions de deux systèmes qui doivent coïncider soient les mêmes. Soit par exemple la décomposition donnée dans la partie droite de la figure 6.7. L'hexaèdre étiqueté (1) doit être résolu en premier. Supposons qu'on soit capable de trouver l'ensemble de ses s_1 solutions $\{\sigma_1(X_1), \dots, \sigma_{s_1}(X_1)\}$. Pour chaque $\sigma_i(X_1)$, avec $1 \leq i \leq s_1$, on peut lire les informations de son bord nécessaires pour résoudre le sous-système étiqueté (2), on obtient alors l'ensemble des s_2^i solutions $\{\sigma_{i,1}(X_2), \dots, \sigma_{i,s_2^i}(X_2)\}$. Pour que les bords coïncident automatiquement, on peut choisir de les fixer (avec les valeurs lues sur $\sigma_i(X_1)$) lors de la résolution de (2) ; on a alors éventuellement $s_2^i = 0$. Bien sûr, si le problème global admet des solutions, il existe au moins un i tel que $s_2^i \neq 0$, si on néglige l'imprécision numérique de la résolution. Dans ce cas, pour garantir la réussite de la phase d'assemblage, il faut être capable de trouver toutes les solutions des sous-systèmes. En associant à la phase d'assemblage un arbre, où chaque branche correspond à un choix consécutif de solutions des sous-systèmes, cela signifie qu'il faut parcourir en profondeur cet arbre pour trouver des solutions.

La phase d'assemblage semble encore plus difficile quand la décomposition obtenue est celle de la partie gauche de la figure 6.4, car le sous-système étiqueté (5) doit coïncider, à ses bords, avec

deux figures. Il paraît alors pertinent d'étudier plus en détail cette phase d'assemblage pour fournir des décompositions qui la facilitent.

Chapitre 7

Conclusion et perspectives

L'étude et la résolution de systèmes de contraintes géométriques en trois dimensions ne peuvent être considérées comme un " $n + 1$ " par rapport au cas bidimensionnel. Le passage de la 2D à la 3D représente un saut qualitatif pour les méthodes géométriques : mise à part la méthode des lieux qui résout très peu de problèmes, elles n'y sont pas adaptées, ne serait-ce qu'à cause de l'explosion combinatoire du nombre de cas à considérer. Pour les méthodes de résolution numériques, le saut est quantitatif, puisque les systèmes d'un univers 3D impliquent plus de contraintes, donc plus d'équations. Leur sensibilité aux problèmes de constriction est liée à leur nature, et non à la dimension de l'univers géométrique dans lequel elles opèrent. Le but de la décomposition est de faire face à l'augmentation du coût de la résolution, mais les méthodes existantes ne parviennent à simplifier guère plus que des problèmes relativement simples en 3D et sont basées sur une approche combinatoire de la rigidité qui est inexacte.

Les travaux présentés dans ce mémoire concernent une méthode numérique, guidée par la géométrie constructive ; bien que prenant en compte la nature géométrique des problèmes traités, les raisonnements utilisés sont valides dans le plan comme dans l'espace.

7.1 Bilan

L'homotopie est une méthode robuste pour obtenir toutes les solutions isolées complexes d'un système d'équations. Son utilisation pour résoudre des systèmes de contraintes géométriques est cependant trop coûteuse, car un nombre exponentiel de chemins d'homotopie, la plupart aboutissant à des solutions complexes qui ont peu d'intérêt en résolution de contraintes géométriques, sont suivis. En CAO, où une solution proche d'une esquisse fournie par un utilisateur est demandée, cette méthode est modifiée en construisant le système initial à partir de l'esquisse. L'unique solution obtenue est proche de l'esquisse.

On a présenté dans ce mémoire une nouvelle méthode suivant le chemin d'homotopie auquel appartient l'esquisse dans son ensemble ; plusieurs solutions sont obtenues, qui résultent toutes d'une déformation continue de l'esquisse. Cette méthode a donc l'avantage de proposer plusieurs solutions, proches de l'esquisse, en un temps raisonnable, c'est-à-dire permettant une interaction. Elle est justifiée par une caractérisation des chemins d'homotopie : ils sont difféomorphes dans la plupart des cas à des cercles. Dans un second temps, un processus de recherche de nouvelles solutions peut-être engagé et, dans beaucoup de cas, il permet de trouver toutes les solutions d'un problème. Bien que coûteux, ce processus itératif semble adapté à la CAO puisqu'il pourrait être guidé par un utilisateur.

Quand ils ne sont pas difféomorphes à des cercles, les chemins d’homotopie peuvent avoir une longueur infinie ; il est alors nécessaire de détecter ces situations pour que le processus termine. Notre homotopie opère dans l’espace des figures, et les “limites” de ces chemins correspondent à des configurations géométriques particulières. On utilise des plans de construction pour détecter avant le suivi d’un chemin si une figure peut présenter une telle configuration ; cette analyse permet, pendant le suivi du chemin, de stopper ce processus si elle est détectée. Contrairement aux cas similaires rencontrés par les méthodes homotopiques classiques, ces chemins de longueur infinie peuvent mener à des solutions.

Quand un système de contraintes géométriques implique des contraintes d’incidence, il peut avoir des ensembles de solutions hétérogènes en dimensions, et cette situation amène le chemin suivi par notre méthode à changer de topologie. Suivre un tel chemin est impossible ou infini, notamment si il présente des embranchements, ou si il coupe des ensembles de solutions de dimensions plus grandes que deux. En perturbant l’interpolation des paramètres de la fonction d’homotopie, on garantit presque sûrement que ces situations se produisent uniquement en des solutions du problème à résoudre ; la méthode de suivi de chemin utilisée est adaptée pour prendre en compte ces cas. On a également esquissé une méthode permettant de détecter à l’avance que de telles situations peuvent se produire, ou d’analyser lors de l’obtention d’une solution si elle est isolée ou non.

Enfin, le problème de la décomposition de système de contraintes géométriques a été abordé. En 3D, ce problème est très difficile et de nombreux systèmes courants en CAO, comme les exemples utilisés dans ce mémoire, sont indécomposables au sens où ils ne contiennent aucun sous-système rigide. L’approche proposée s’appuie sur la reparamétrisation et parvient à décomposer des systèmes qui sont des compositions de plus petits systèmes rigides. Si cette partie des travaux présentés ici peut paraître orthogonale aux autres par son aspect combinatoire, elle s’y inscrit naturellement comme une tentative de diminution du coût de la résolution d’un système de contraintes.

7.2 Apports

Pour résumer, voici les principaux apports présentés dans ce mémoire :

- on présente une méthode d’homotopie pour résoudre des SCG utilisant l’esquisse comme système initial, produisant plusieurs solutions proches de l’esquisse ;
- cette méthode peut être guidée par la géométrie constructive pour obtenir de nouvelles solutions ; lors d’une recherche exhaustive, elles sont souvent toutes trouvées. Dans le cas du problème de la construction d’un hexaèdre, toutes les solutions sont fournies en un temps 100 fois plus petit qu’avec le logiciel `HOM4PS-2.0` ;
- on donne des exemples simples et non-artificiels de SCG dont les ensembles de solutions ne sont pas des variétés algébriques : ils sont réunions d’ensembles hétérogènes en dimensions ;
- on montre comment adapter notre méthode pour résoudre de tels SCG ;
- on propose une méthode utilisant le formalisme des matroïdes permettant de détecter la cause, quand elle est liée à la satisfaction de théorèmes de géométrie d’incidence, de ce phénomène, pour déterminer la dimension des variétés solutions. Sur l’exemple de l’hexaèdre, cette méthode est environ 10000 fois plus rapide que le logiciel `Bertini` mais ne garantit pas que tous les ensembles de solutions soient trouvés (sur l’exemple en question ils le sont) ;
- à cette fin on extrait, de l’ensemble des incidences respectées par les points d’une figure, un ensemble de contraintes indépendantes impliquant ces incidences. Ceci pourrait sans doute être utilisé pour corriger des dessins en perspective (par exemple des dessins d’architecture) ¹.

Les travaux sur la caractérisation des chemins d’homotopie, l’utilisation d’un plan de construction pour détecter les chemins de longueur infinie et générer de nouvelles esquisses ainsi que la

1. voir par exemple [Sosnov 02, Macé 04]

modification des contraintes booléennes en contraintes métriques pour éviter les points critiques ont donné lieu à la rédaction d'un article pour la conférence internationale GD/SPM 2013. Il paraîtra dans la revue CAD (Computer Aided Design) de Elsevier ([Imbach 13]).

La méthode de décomposition de SCG avec remplacement de contraintes a donné lieu à une publication et une communication lors de la conférence internationale SAC (Symposium On Applied Computing) en 2012 ([Mathis 12]), et aux journées du GTMG (Groupe de Travail en Modélisation Géométrique) en 2012 ([Imbach 12]).

Enfin, la résolution de la reparamétrisation par homotopie a été le sujet de [Imbach 11], présenté à la conférence internationale SYNASC (International Symposium on Symbolic and Numeric Algorithms for Scientific Computing) en 2011.

7.3 Perspectives

Les travaux exploratoires exposés dans ce mémoire visent à cerner une problématique, celle de l'utilisation de la géométrie pour guider la méthode par homotopie appliquée à la résolution de contraintes en géométrie. Les contributions décrites ici ont permis d'une part de mettre en lumière les difficultés posées par l'application d'une méthode par continuation, et d'autre part d'apporter quelques débuts de réponses à cette problématique dans le cadre des domaines où apparaissent des contraintes géométriques et notamment en CAO. Elles débouchent sur un large champ de perspectives.

On a évoqué, dans la conclusion du chapitre 3, l'influence que les fonctions d'interpolations pouvaient avoir sur la géométrie des chemins d'homotopie (en plus de celle qu'elles ont sur leur topologie). Modifier les fonctions d'interpolation peut changer la forme d'une composante connexe et surtout le nombre de solutions qu'elle contient. Si la caractérisation de cette influence semble difficile, elle peut néanmoins faire l'objet d'heuristiques, par exemple pour "transformer" les chemins convergeant en dehors du domaine de définition de la fonction d'homotopie, ou de longueur infinie en chemins difféomorphes à des cercles.

L'obtention du plan de construction utilisé pourrait être menée de façon à optimiser la détection de ces cas ou le nombre de nouvelles solutions obtenues en l'utilisant pour générer de nouvelles esquisses. On a vu l'influence que ce plan de construction avait sur la phase de recherches de nouvelles solutions.

En ce qui concerne la décomposition de SCG, il semble intéressant d'étudier la phase d'assemblage des solutions obtenues et notamment d'aborder la question suivante : quelle est la précision requise dans la résolution des sous-systèmes pour assurer que les bords des solutions coïncident ?

Une problématique a depuis le début accompagné les travaux présentés ici mais elle ne transparaît pas dans ce mémoire ; c'est celle de l'obtention de toutes les solutions réelles d'un système de contraintes. Quand l'esquisse est utilisée pour construire le système initial, trouver toutes les solutions réelles et complexes de ce système initial pour assurer la continuation des solutions est aussi difficile que de résoudre le problème lui-même. La méthode proposée pour découvrir de nouvelles solutions en générant de nouvelles esquisses présente l'avantage d'être itérative et de pouvoir être guidée par un utilisateur ; elle ne peut cependant garantir l'obtention de toutes les solutions, puisqu'elle procède en construisant de nouvelles fonctions d'homotopie.

Enfin, la détection des continuums de solutions conséquences de la satisfaction de théorèmes de géométrie d'incidence semble prometteuse d'abord parce qu'elle touche au problème de la constriction générique des systèmes de contraintes géométriques, ensuite parce qu'elle permet de

prédire l'existence de configurations critiques susceptibles d'intéresser un utilisateur d'un système de modélisation par contraintes, qu'il soit architecte ou roboticien.

Bibliographie

- [Ait-Aoudia 93] Samy Ait-Aoudia, Roland Jegou & Dominique Michelucci. *Reduction of constraint systems*. 1993.
- [Ait-Aoudia 01] Samy Ait-Aoudia, Hakim Badis & Messaoud Kara. *Solving geometric constraints by a hybrid method*. Dans Information Visualisation, 2001. Proceedings. Fifth International Conference on, pages 749–753. IEEE, 2001.
- [Ait-Aoudia 02] Samy Ait-Aoudia. *Numerical solving of geometric constraints*. Dans Information Visualisation, 2002. Proceedings. Sixth International Conference on, pages 125–129. IEEE, 2002.
- [Aldefeld 88] Bernd Aldefeld. *Variation of geometries based on a geometric-reasoning method*. Computer-Aided Design, vol. 20, no. 3, pages 117–126, 1988.
- [Allgower 97] E.L. Allgower & K. Georg. *Numerical path following*. Handbook of Numerical Analysis, vol. 5, no. 3, page 207, 1997.
- [Bates 06] Daniel J Bates, Jonathan D Hauenstein, Andrew J Sommese & Charles W Wampler. *Bertini : Software for numerical algebraic geometry*, 2006.
- [Bates 08] Daniel J Bates, Jonathan D Hauenstein, Andrew J Sommese & Charles W Wampler. *Adaptive multiprecision path tracking*. SIAM Journal on Numerical Analysis, vol. 46, no. 2, pages 722–746, 2008.
- [Bernshtein 75] David N Bernshtein. *The number of roots of a system of equations*. Functional Analysis and its Applications, vol. 9, no. 3, pages 183–185, 1975.
- [Bischof 08] Christian Bischof. Advances in automatic differentiation, volume 64. Springer, 2008.
- [Bouma 95] William Bouma, Ioannis Fudos, Christoph Hoffmann, Jiazhen Cai & Robert Paige. *Geometric constraint solver*. Computer-Aided Design, vol. 27, no. 6, pages 487–501, 1995.
- [Brüderlin 87] Beat Brüderlin. *Constructing three-dimensional geometric objects defined by constraints*. Dans Proceedings of the 1986 workshop on Interactive 3D graphics, pages 111–129. ACM, 1987.
- [Buthion 79] Michel Buthion. *Un Programme qui résout formellement des problèmes de constructions géométriques*. RAIRO Informatique, vol. 3, no. 4, pages 353–387, 1979.
- [Cauchy 13] Augustin Louis Cauchy. *Sur les polygones et polyèdres, Second mémoire*. J. Ecole Polytechnique, vol. 9, page 87, 1813.
- [Chou 88] Shang-Ching Chou. *An introduction to Wu’s method for mechanical theorem proving in geometry*. Journal of Automated Reasoning, vol. 4, no. 3, pages 237–267, 1988.

- [Chou 90] Shang-Ching Chou & Xiao-Shan Gao. *Ritt-Wu's decomposition algorithm and geometry theorem proving*. Dans 10th International Conference on Automated Deduction, pages 207–220. Springer, 1990.
- [Chow 79] Shui-Nee Chow, John Mallet-Paret & James A Yorke. *A homotopy method for locating all zeros of a system of polynomials*. Dans Functional differential equations and approximation of fixed points, pages 77–88. Springer, 1979.
- [Den Heijer 81] C Den Heijer & WC Rheinboldt. *On steplength algorithms for a class of continuation methods*. SIAM Journal on Numerical Analysis, vol. 18, no. 5, pages 925–948, 1981.
- [Dobkin 90] David P Dobkin, Allan R Wilks, Silvio VF Levy & William P Thurston. *Contour tracing by piecewise linear approximations*. ACM Transactions on Graphics (TOG), vol. 9, no. 4, pages 389–423, 1990.
- [Drexler 78] FJ Drexler. *A homotopy method for the calculation of all zeros of zero-dimensional polynomial ideals*, 1978.
- [Dufourd 98] Jean-François Dufourd, Pascal Mathis & Pascal Schreck. *Geometric construction by assembling solved subfigures*. Artificial Intelligence, vol. 99, no. 1, pages 73–119, 1998.
- [Dulmage 63] AL Dulmage & NS Mendelsohn. *Two algorithms for bipartite graphs*. Journal of the Society for Industrial & Applied Mathematics, vol. 11, no. 1, pages 183–194, 1963.
- [Durand 98] C. Durand & CM Hoffmann. *Continuum : A Homotopy Continuation Solver for Systems of Algebraic Equations*. Rapport technique, TR 98-028, Department of Computer Sciences, Purdue University, 1998.
- [Durand 00] C. Durand & C.M. Hoffmann. *A systematic framework for solving geometric constraints analytically*. Journal of Symbolic Computation, vol. 30, no. 5, pages 493–519, 2000.
- [Elber 01] Gershon Elber & Myung-Soo Kim. *Geometric constraint solver using multivariate rational spline functions*. Dans Proceedings of the sixth ACM symposium on Solid modeling and applications, pages 1–10. ACM, 2001.
- [Elkadi 07] Mohamed Elkadi & Bernard Mourrain. *Introduction à la résolution des systèmes polynomiaux*, volume 59. Springer, 2007.
- [Essert-Villard 00] Caroline Essert-Villard, Pascal Schreck & J-F Dufourd. *Sketch-based pruning of a solution space within a formal geometric constraint solver*. Artificial Intelligence, vol. 124, no. 1, pages 139–159, 2000.
- [Essert-Villard 01] Caroline Essert-Villard. *Sélection dans l'espace des solutions engendrées par un plan de construction géométrique*. Thèse de doctorat, 2001.
- [Fabre 06] Arnaud Fabre. *Contraintes géométriques en dimension 3*. Thèse de doctorat, 2006.
- [Fabre 07] Arnaud Fabre & Pascal Schreck. *A formal-numerical approach to solve 3D geometric constraints systems*. Dans Geometric Modeling and Imaging, 2007. GMAI'07, pages 54–59. IEEE, 2007.
- [Fabre 08] Arnaud Fabre & Pascal Schreck. *Combining symbolic and numerical solvers to simplify indecomposable systems solving*. Dans Proceedings of the 2008 ACM symposium on Applied computing, pages 1838–1842. ACM, 2008.
- [Faudot 07] D. Faudot & D. Michelucci. *A new robust algorithm to trace curves*. Reliable computing, vol. 13, no. 4, pages 309–324, 2007.

- [Fünfzig 09] Christoph Fünfzig, Dominique Michelucci & Sebti Foufou. *Nonlinear systems solver in floating-point arithmetic using lp reduction*. Dans 2009 SIAM/ACM Joint Conference on Geometric and Physical Modeling, pages 123–134. ACM, 2009.
- [Gao 02] Xiao-Shan Gao, Christoph M. Hoffmann & Wei-Qiang Yang. *Solving spatial basic geometric constraint configurations with locus intersection*. Dans Proceedings of the seventh ACM symposium on Solid modeling and applications, SMA '02, pages 95–104, New York, NY, USA, 2002. ACM.
- [Gao 09] Heping Gao & Meera Sitharam. *Characterizing 1-Dof Henneberg-I graphs with efficient configuration spaces*. Dans Proceedings of the 2009 ACM symposium on Applied Computing, pages 1122–1126. ACM, 2009.
- [Garcia 79] CB Garcia & Willard I Zangwill. *Finding all solutions to polynomial systems and other systems of equations*. Mathematical Programming, vol. 16, no. 1, pages 159–176, 1979.
- [Ge 99] Jian-Xin Ge, Shang-Ching Chou & Xiao-Shan Gao. *Geometric constraint satisfaction using optimization methods*. Computer-Aided Design, vol. 31, no. 14, pages 867–879, 1999.
- [Gramain 71] André Gramain. Topologie des surfaces, volume 7. Presses universitaires de France, 1971.
- [Hager 89] William W Hager. *Updating the inverse of a matrix*. SIAM review, vol. 31, no. 2, pages 221–239, 1989.
- [Hendrickson 92] Bruce Hendrickson. *Conditions for unique graph realizations*. SIAM Journal on Computing, vol. 21, no. 1, pages 65–84, 1992.
- [Huber 95] B. Huber & B. Sturmfels. *A polyhedral method for solving sparse polynomial systems*. Mathematics of computation, vol. 64, no. 212, pages 1541–1556, 1995.
- [Imbach 11] R. Imbach, P. Mathis & P. Schreck. *Tracking method for reparametrized geometrical constraint systems*. Dans 2011 13th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, pages 31–38. IEEE, 2011.
- [Imbach 12] R. Imbach, P. Mathis & P. Schreck. *Une approche par décomposition et reparamétrisation de systèmes de contraintes géométriques*. Dans Journées du Groupe de Travail en Modélisation Géométrique, 2012.
- [Imbach 13] R. Imbach, P. Schreck & P. Mathis. *Leading a continuation method by geometry for solving geometric constraints*. Computer-Aided Design, 2013.
- [Jermann 02] Christophe Jermann. *Résolution de contraintes géométriques par rigidifications récursive et propagation d'intervalles*. Thèse de doctorat, Université de Nice Sophia-Antipolis, 2002.
- [Joan-Arinyo 99] R Joan-Arinyo & A Soto-Riera. *Combining constructive and equational geometric constraint-solving techniques*. ACM Transactions on Graphics (TOG), vol. 18, no. 1, pages 35–55, 1999.
- [Kearfott 94] R Baker Kearfott & Zhaoyun Xing. *An interval step control for continuation methods*. SIAM Journal on Numerical Analysis, vol. 31, no. 3, pages 892–914, 1994.
- [Kramer 90] Glenn A Kramer. *Solving geometric constraint systems*. Dans Proceedings of the eighth National conference on Artificial intelligence-Volume 1, pages 708–714. AAAI Press, 1990.

- [Krawczyk 69] Rudolf Krawczyk. *Newton-algorithmen zur bestimmung von nullstellen mit fehlerschranken*. Computing, vol. 4, no. 3, pages 187–201, 1969.
- [Laman 70] G. Laman. *On graphs and rigidity of plane skeletal structures*. Journal of Engineering Mathematics, vol. 4, pages 331–340, 1970.
- [Lamure 95] Hervé Lamure & Dominique Michelucci. *Solving geometric constraints by homotopy*. pages 263–269, 1995.
- [Lee 08] Tsung-Lin Lee, Tien-Yien Li & Chih-Hsiung Tsai. *HOM4PS-2.0 : a software package for solving polynomial systems by the polyhedral homotopy continuation method*. Computing, vol. 83, no. 2-3, pages 109–133, 2008.
- [Li 89] TY Li, Tim Sauer & JA Yorke. *The cheater’s homotopy : an efficient procedure for solving systems of polynomial equations*. SIAM Journal on Numerical Analysis, vol. 26, no. 5, pages 1241–1251, 1989.
- [Li 93] TY Li & Xiao Shen Wang. *Solving real polynomial systems with real homotopies*. mathematics of computation, vol. 60, no. 202, pages 669–680, 1993.
- [Li 01] Tiejun Li & Fengshan Bai. *Minimizing multi-homogeneous Bézout numbers by a local search method*. Mathematics of computation, vol. 70, no. 234, pages 767–787, 2001.
- [Lin 81] Vincent C Lin, David C Gossard & Robert A Light. *Variational geometry in computer-aided design*. Dans ACM SIGGRAPH Computer Graphics, volume 15, pages 171–177. ACM, 1981.
- [Lovasz 82] Laszlo Lovasz & Yechiam Yemini. *On generic rigidity in the plane*. SIAM Journal on Algebraic Discrete Methods, vol. 3, no. 1, pages 91–98, 1982.
- [Macé 04] Pierre Macé, Dominique Michelucci & Louis-Paul Untersteller. *Solving Space Orthogonality Problems in a 2D Drawing*. Dans 8th International Workshop on Automated Deduction in Geometry, 2004.
- [Malajovich 07] Gregorio Malajovich & Klaus Meer. *Computing minimal multi-homogeneous Bézout numbers is hard*. Theory of Computing Systems, vol. 40, no. 4, pages 553–570, 2007.
- [Mathis 97] Pascal Mathis. *Constructions géométriques sous contraintes en modélisation à base topologique*. Thèse de doctorat, 1997.
- [Mathis 10] Pascal Mathis & Simon EB Thierry. *A formalization of geometric constraint systems and their decomposition*. Formal Aspects of Computing, vol. 22, no. 2, pages 129–151, 2010.
- [Mathis 12] P. Mathis, P. Schreck & R. Imbach. *Decomposition of geometrical constraint systems with reparameterization*. Dans Proceedings of the 27th Annual ACM Symposium on Applied Computing, pages 102–108. ACM, 2012.
- [Michelucci 04] D Michelucci & P Schreck. *Detecting induced incidences in the projective plane*. Dans isiCAD Workshop. Citeseer, 2004.
- [Michelucci 06a] Dominique Michelucci & Sebti Foufou. *Geometric constraint solving : the witness configuration method*. Computer-Aided Design, vol. 38, no. 4, pages 284–299, 2006.
- [Michelucci 06b] Dominique Michelucci & Pascal Schreck. *Incidence constraints : a combinatorial approach*. International Journal of Computational Geometry & Applications, vol. 16, no. 05n06, pages 443–460, 2006.

- [Michelucci 07] Dominique Michelucci & Sebti Foufou. *Detecting all dependences in systems of geometric constraints using the witness method*. Dans Automated Deduction in Geometry, pages 98–112. Springer, 2007.
- [Milnor 65] J. Milnor. *Topology from the differentiable viewpoint*. Univ. Press Virginia, 1965.
- [Moore 66] Ramon E Moore. Interval analysis, volume 2. Prentice-Hall Englewood Cliffs, 1966.
- [Morgan 86a] Alexander P Morgan. *A homotopy for solving polynomial systems*. Applied Mathematics and Computation, vol. 18, no. 1, pages 87–92, 1986.
- [Morgan 86b] A.P. Morgan. *A transformation to avoid solutions at infinity for polynomial systems*. Applied mathematics and computation, vol. 18, no. 1, pages 77–86, 1986.
- [Morgan 87] Alexander Morgan & Andrew Sommese. *A homotopy for solving general polynomial systems that respects m-homogeneous structures*. Applied Mathematics and Computation, vol. 24, no. 2, pages 101–113, 1987.
- [Morgan 89] Alexander P Morgan & Andrew J Sommese. *Coefficient-parameter polynomial continuation*. Applied Mathematics and Computation, vol. 29, no. 2, pages 123–160, 1989.
- [Mourrain 09] Bernard Mourrain & Jean-Pascal Pavone. *Subdivision methods for solving polynomial equations*. Journal of Symbolic Computation, vol. 44, no. 3, pages 292–306, 2009.
- [Neumaier 90] Arnold Neumaier. Interval methods for systems of equations, volume 37. Cambridge university press, 1990.
- [Ortega 87] James M Ortega & Werner C Rheinboldt. Iterative solution of nonlinear equations in several variables, volume 30. Society for Industrial and Applied Mathematics, 1987.
- [Owen 91] John C Owen. *Algebraic solution for geometry from dimensional constraints*. Dans Proceedings of the first ACM symposium on Solid modeling foundations and CAD/CAM applications, pages 397–407. ACM, 1991.
- [Pérez 93] Alfredo Pérez & David Serrano. *Constraint based analysis tools for design*. Dans Proceedings on the second ACM symposium on Solid modeling and applications, pages 281–290. ACM, 1993.
- [Plummer 86] MD Plummer & L Lovász. Matching theory, volume 121. North Holland, 1986.
- [Porta 07] Josep M Porta, Lluís Ros, Federico Thomas, Francesc Corcho, Josep Cantó & Juan Jesús Pérez. *Complete maps of molecular-loop conformational spaces*. Journal of computational chemistry, vol. 28, no. 13, pages 2170–2189, 2007.
- [Scandura 74] Joseph M Scandura, John H Durnin & Wallace H Wulfeck II. *Higher order rule characterization of heuristics for compass and straight edge constructions in geometry*. Artificial Intelligence, vol. 5, no. 2, pages 149–183, 1974.
- [Schreck 93] Pascal Schreck. *Automatisation des constructions géométriques à la règle et au compas*. Thèse de doctorat, 1993.
- [Schreck 06] Pascal Schreck & Pascal Mathis. *Geometrical constraint system decomposition : a multi-group approach*. International Journal of Computational Geometry & Applications, vol. 16, no. 05n06, pages 431–442, 2006.

- [Serrano 91] David Serrano. *Automatic dimensioning in design for manufacturing*. Dans Proceedings of the first ACM symposium on Solid modeling foundations and CAD/CAM applications, pages 379–386. ACM, 1991.
- [Sommese 96] Andrew J Sommese, Jan Verschelde & Charles W Wampler. *Numerical algebraic geometry*. Dans The Mathematics of Numerical Analysis, volume 32 of Lectures in Applied Mathematics. Citeseer, 1996.
- [Sommese 00] Andrew J Sommese & Jan Verschelde. *Numerical homotopies to compute generic points on positive dimensional algebraic sets*. Journal of Complexity, vol. 16, no. 3, pages 572–602, 2000.
- [Sommese 05] Andrew John Sommese & Charles Weldon Wampler. The numerical solution of systems of polynomials : Arising in engineering and science. World Scientific Publishing Company, 2005.
- [Sommese 07] Andrew J Sommese, Jan Verschelde & Charles W Wampler. *Solving polynomial systems equation by equation*. Dans IMA Volume 146 : Algorithms in Algebraic Geometry. Citeseer, 2007.
- [Sosnov 02] Alex Sosnov, Pierre Macé & Gérard Hégron. *Semi-metric formal 3D reconstruction from perspective sketches*. Dans Computational Science—ICCS 2002, pages 285–294. Springer, 2002.
- [Sunde 87] Geir Sunde. *A CAD system with declarative specification of shape*. Dans Intelligent CAD Systems I, pages 90–104. Springer, 1987.
- [Sutherland 64] Ivan E Sutherland. *Sketch pad a man-machine graphical communication system*. Dans Proceedings of the SHARE design automation workshop, pages 6–329. ACM, 1964.
- [Tahari 10] Abdelkarim Tahari, Christoph Fünfzig, Dominique Michelucci, Sebti Foufou & Sami Ait-Aoudia. *Bases tensorielles de Bernstein et solveurs*. Technique et Science Informatiques, vol. 29, no. 6, page 629, 2010.
- [Thierry 07] Simon EB Thierry, Pascal Mathis & Pascal Schreck. *Towards an homogeneous handling of under-constrained and well-constrained systems of geometric constraints*. Dans Proceedings of the 2007 ACM symposium on Applied computing, pages 773–777. ACM, 2007.
- [Thierry 11] Simon EB Thierry, Pascal Schreck, Dominique Michelucci, Christoph Fünfzig & Jean-David Génevaux. *Extensions of the witness method to characterize under-, over-and well-constrained geometric constraint systems*. Computer-Aided Design, vol. 43, no. 10, pages 1234–1249, 2011.
- [van der Waerden 49] BL van der Waerden. *Le théorème de Bézout pour les hypersurfaces*. Annali di Matematica Pura ed Applicata, vol. 30, no. 1, pages 73–74, 1949.
- [Verroust 92] A Verroust, F Schonek & Dieter Roller. *Rule-oriented method for parameterized computer-aided design*. Computer-Aided Design, vol. 24, no. 10, pages 531–540, 1992.
- [Verschelde 94a] Jan Verschelde & Ronald Cools. *Symmetric homotopy construction*. Journal of Computational and Applied Mathematics, vol. 50, no. 1, pages 575–592, 1994.
- [Verschelde 94b] Jan Verschelde, Pierre Verlinden & Ronald Cools. *Homotopies exploiting Newton polytopes for solving sparse polynomial systems*. SIAM Journal on Numerical Analysis, vol. 31, no. 3, pages 915–930, 1994.

- [Vershelde 99] Jan Vershelde. *Algorithm 795 : PHCpack : A general-purpose solver for polynomial systems by homotopy continuation*. ACM Transactions on Mathematical Software (TOMS), vol. 25, no. 2, pages 251–276, 1999.
- [Wampler 93] Charles W Wampler. *An efficient start system for multi-homogeneous polynomial continuation*. Numerische Mathematik, vol. 66, no. 1, pages 517–523, 1993.
- [Wen-Tsun 86] Wu Wen-Tsun. *Basic principles of mechanical theorem proving in elementary geometries*. Journal of Automated Reasoning, vol. 2, no. 3, pages 221–252, 1986.
- [Wright 85] Alden H Wright. *Finding all solutions to a system of polynomial equations*. mathematics of computation, vol. 44, no. 169, pages 125–133, 1985.
- [Yakoubsohn 95] Jean-Claude Yakoubsohn. *A universal constant for the convergence of Newton's method and an application to the classical homotopy method*. Numerical Algorithms, vol. 9, no. 2, pages 223–244, 1995.

Annexe A

Rappels de topologie différentielle

Les définitions, lemmes et théorèmes donnés ici sont tirés de [Milnor 65]. [Gramain 71] constitue aussi une introduction à la topologie différentielle, en français. La terminologie française en est issue.

A.1 Variétés différentielles

Définition 16 Soit $F : U \rightarrow V$ une fonction, où $U \subset \mathbb{R}^k$ et $V \subset \mathbb{R}^l$ sont des ouverts. F est C^∞ si toutes ses dérivées partielles $\frac{\partial^n F}{\partial x_{i_1} \dots \partial x_{i_n}}$, où $n \in \mathbb{N} \setminus \{0\}$ et $\forall 1 \leq j \leq n, i_j \leq k$, existent et sont continues. On appellera dérivée de F en $x \in U$, et on la notera ∂F_x l'application linéaire définie par sa matrice jacobienne JF_x , dont les colonnes sont les vecteurs $\frac{\partial F}{\partial x_i}(x)$ pour $1 \leq i \leq k$.

On introduit alors la notion de difféomorphismes entre deux sous-ensembles d'espaces vectoriels réels, centrale en topologie différentielle et dans notre caractérisation des chemins d'homotopie.

Définition 17 Soient $X \subset \mathbb{R}^k$ et $Y \subset \mathbb{R}^l$. On appelle $f : X \rightarrow Y$ un difféomorphisme, et on dit que X est difféomorphe à Y , si pour tout $u \in X$, il existe un ouvert $U \subset \mathbb{R}^k$ avec $u \in U$ et une fonction $F : U \rightarrow \mathbb{R}^l$ telle que :

- (i) $\forall x \in U \cap X, F(x) = f(x)$,
- (ii) F est un homéomorphisme (F est bijective et continue, F^{-1} est continue),
- (iii) F et F^{-1} sont C^∞ .

Cette notion de difféomorphisme permet alors de définir ce qu'est une variété différentielle.

Définition 18 Un ensemble $\mathcal{M} \subset \mathbb{R}^k$ est une variété différentielle de dimension m si chacun de ses points $x \in \mathcal{M}$ admet un voisinage $W \cap \mathcal{M}$ difféomorphe à un ouvert U de \mathbb{R}^m . On appellera paramétrisation de la région $W \cap \mathcal{M}$ tout difféomorphisme $g : U \rightarrow W \cap \mathcal{M}$.

Par définition, un ensemble $\mathcal{M} \subset \mathbb{R}^k$ est une variété différentielle de dimension 0 si chacun de ses points $x \in \mathcal{M}$ admet un voisinage $W \cap \mathcal{M}$ réduit à lui-même. Dans le présent mémoire, quand il n'y a pas de confusion possible avec les variétés algébriques, les variétés différentielles sont aussi appelées variétés.

A.2 Fonctions entre variétés, dérivation

On s'intéresse à des fonctions entre variétés $f : \mathcal{M} \rightarrow \mathcal{N}$, où $\mathcal{M} \subset \mathbb{R}^k$ et $\mathcal{N} \subset \mathbb{R}^l$, et notamment à leur propriétés différentielles.

Définition 19 Soit $f : \mathcal{M} \rightarrow \mathcal{N}$, où $\mathcal{M} \subset \mathbb{R}^k$ et $\mathcal{N} \subset \mathbb{R}^l$ sont des variétés de dimensions respectives m et n . f est C^∞ si pour chaque $x \in \mathcal{M}$, il existe un ouvert $U \subset \mathbb{R}^k$ avec $x \in U$, et une fonction $F : U \rightarrow \mathbb{R}^l$ telle que F soit C^∞ et coïncide avec f sur $U \cap \mathcal{M}$ (c.à.d. $\forall u \in U \cap \mathcal{M}$, $F(u) = f(u)$).

Afin de définir la dérivée d'une telle fonction, on introduit la notion d'espace tangent à une variété.

Définition 20 Soit $\mathcal{M} \subset \mathbb{R}^k$ une variété de dimension m , $x \in \mathcal{M}$, et $g : U \rightarrow \mathcal{M} \subset \mathbb{R}^k$ une paramétrisation d'un voisinage $g(U)$ de x où U est un ouvert de \mathbb{R}^m . Soit encore $u \in U$ tel que $g(u) = x$, et $\partial g_u : \mathbb{R}^m \rightarrow \mathbb{R}^k$ la dérivée de g en u . On appelle espace tangent de \mathcal{M} en x , et on le note $T\mathcal{M}_x$ l'espace vectoriel de dimension m égal à l'image par ∂g_u de \mathbb{R}^m .

On propose d'illustrer cette notion avec une variété $\mathcal{M} \subset \mathbb{R}^k$ de dimension k , qui admet, au voisinage de chacun de ses points x , une paramétrisation par l'identité $id : U \rightarrow \mathbb{R}^k$, où $U = id(U)$ est un voisinage de x . L'application linéaire ∂id_x est donc l'application identité de \mathbb{R}^k (représenté par la matrice identité). Pour tout $x \in \mathcal{M}$, l'espace tangent $T\mathcal{M}_x$ est alors l'espace vectoriel \mathbb{R}^k de dimension k .

On définit alors la dérivée d'une fonction entre deux variétés :

Définition 21 Soit $f : \mathcal{M} \rightarrow \mathcal{N}$ une fonction C^∞ , où $\mathcal{M} \subset \mathbb{R}^k$ et $\mathcal{N} \subset \mathbb{R}^l$ sont des variétés de dimensions respectives m et n . Soit $x \in \mathcal{M}$ et $f(x) = y$. Il existe alors un voisinage ouvert W de x dans \mathbb{R}^k et une fonction C^∞ $F : W \rightarrow \mathbb{R}^l$ qui coïncide avec f sur $W \cap \mathcal{M}$.

On appelle dérivée de f en x , et on la note ∂f_x , l'application linéaire $\partial f_x : T\mathcal{M}_x \rightarrow T\mathcal{N}_y$ définie par $\partial f_x(v) = \partial F_x(v)$, pour tout $v \in T\mathcal{M}_x$.

On peut montrer, et le lecteur est renvoyé à [Milnor 65] pour cette preuve, que les définitions d'espace tangent et de dérivée d'une fonction entre deux variétés ne dépendent pas des paramétrisations de ces variétés.

Considérons une telle fonction C^∞ $f : \mathcal{M} \rightarrow \mathbb{R}^l$, où $\mathcal{M} \subset \mathbb{R}^k$ est une variété de dimension k , qui admet en chaque point une paramétrisation par l'identité. En remarquant que \mathbb{R}^l est une variété de dimension l qui admet en chacun de ses points une paramétrisation par l'identité, on peut affirmer que pour tout $x \in \mathcal{M}$, la dérivée de f en x est une application linéaire $\partial f_x : \mathbb{R}^k \rightarrow \mathbb{R}^l$ (voir ci-dessus). De plus, pour tout $v \in \mathbb{R}^k$, $\partial f_x(v) = \partial F_x(v)$, où F est une fonction C^∞ d'un voisinage de x dans \mathbb{R}^k . On en déduit que pour tout $x \in \mathcal{M}$, ∂f_x est l'application linéaire $\partial f_x : \mathbb{R}^k \rightarrow \mathbb{R}^l$ dont la matrice est la matrice jacobienne JF_x . Cette spécialisation, bien que triviale, est suffisante dans le cadre de ce mémoire.

A.3 Valeurs critiques, valeurs régulières

Les notions de valeurs critiques et régulière présentées maintenant portent sur le rang de l'application linéaire dérivée d'une fonction, et jouent un rôle majeur dans la suite.

Définition 22 Soit $f : \mathcal{M} \rightarrow \mathcal{N}$ une fonction C^∞ d'une variété de dimension m dans une variété de dimension n . Soit C l'ensemble des points $x \in \mathcal{M}$ tels que l'application linéaire $\partial f_x(v) : T\mathcal{M}_x \rightarrow T\mathcal{N}_{f(x)}$ a un rang strictement plus petit que n . C est appelé l'ensemble des points critiques de f , $f(C)$ l'ensemble des valeurs critiques de f et le complément $\mathcal{N} \setminus f(C)$ est l'ensemble des valeurs régulières de f . On appellera $x \in C$ un point critique de f , $y \in f(C)$ une valeur critique, et $y \in \mathcal{N} \setminus f(C)$ une valeur régulière de f .

On peut alors, pour une valeur régulière, caractériser l'ensemble de ses antécédants par une fonction entre variétés.

Lemme 23 *Soit $f : \mathcal{M} \rightarrow \mathcal{N}$ une fonction C^∞ entre variétés de dimensions respectives m et n , et $m \geq n$. Si y est une valeur régulière de f , alors l'ensemble $f^{-1}(y) \subset \mathcal{M}$ est une variété de dimension $m - n$.*

On introduit ici la notion de variété avec bord. Posons $H^m = \{(x_1, \dots, x_m) \in \mathbb{R}^m \mid x_m \geq 0\}$. On appelle bord de H^m l'hyperplan $\mathbb{R}^{m-1} \times \{0\}$. Un ensemble $\mathcal{M} \subset \mathbb{R}^k$ est une variété avec bord de dimension m si chaque $x \in \mathcal{M}$ a un voisinage difféomorphe à un ouvert de H^m . On appelle bord $\overline{\mathcal{M}}$ de \mathcal{M} l'ensemble des antécédents du bord de H^m par un tel difféomorphisme.

L'intérieur de \mathcal{M} , égal à $\mathcal{M} \setminus \overline{\mathcal{M}}$ est une variété de dimension m , et le bord $\overline{\mathcal{M}}$ une variété de dimension $m - 1$.

Lemme 24 *Soit $f : \mathcal{M} \rightarrow \mathcal{N}$ une fonction C^∞ entre variétés de dimensions respectives m et n , et $m \geq n$. Supposons que \mathcal{M} ait un bord $\overline{\mathcal{M}}$. Si y est une valeur régulière de f et de la restriction de f à $\overline{\mathcal{M}}$, alors l'ensemble $f^{-1}(y) \subset \mathcal{M}$ est une variété de dimension $m - n$, dont le bord est exactement égal à l'intersection de $f^{-1}(y)$ avec $\overline{\mathcal{M}}$.*

A.4 Classification des variétés de dimension 1

Théorème 25 *Toute variété connexe de dimension 1 est difféomorphe soit au cercle $\mathcal{S}^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$, soit à un intervalle de \mathbb{R} .*

Un intervalle de \mathbb{R} peut être ouvert ($]0, 1[$), semi-ouvert ($[0, 1[$ ou $]0, 1]$), ou fermé ($[0, 1]$). Une variété connexe sans bord de dimension 1 est donc difféomorphe à \mathcal{S}^1 , ou à l'intervalle ouvert $]0, 1[$.

Annexe B

Univers géométriques

On présente ici brièvement les univers géométriques utilisés dans ce mémoire. Dans les chapitres 3 à 5, les exemples considérés sont définis dans un univers 3D. L'interprétation utilisée est numérique : aux variables sont associées des coordonnées cartésiennes réelles, et aux symboles des fonctions numériques ou des équations.

Le chapitre 6 s'appuie principalement sur des exemples d'un univers 2D, et c'est l'interprétation des degrés de liberté et de restriction qui est considérée. L'extension de cette sémantique à un univers 3D est sous-jacente dans la section 6.3 et on la précise à la fin de cette annexe.

Dans le texte, on évoque les sortes *longueur*, *angle*, *surface* et *volume*. Ici on ne considère que la sorte *mesures*, interprétée par des scalaires.

Univers 3D

Signature

Sortes :

mesure
point
sphere

Symboles :

$distance(point, point) \rightarrow mesure$
 $angle(point, point, point) \rightarrow mesure$
 $aire(point, point, point) \rightarrow mesure$
 $volume(point, point, point, point) \rightarrow mesure$
 $sphere(point, mesure) \rightarrow sphere$
 $interSSS(sphere, sphere, sphere) \rightarrow point$
 $colineaires(point, point, point)$
 $coplanaires(point, point, point, point)$

Interprétation numérique

Les variables de sorte *mesure* sont interprétées par des réels, celles de sorte *point* par un triplet de coordonnées cartésiennes réelles $x \in \mathbb{R}^3$, et celles de sorte *sphere* par un quadruplet $x \in \mathbb{R}^4$ dont les trois premiers éléments sont les coordonnées du centre, et le dernier son rayon.

Le symbole *distance* est interprété par la fonction numérique :

$$f_d : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}_+ \\ (x_1, x_2) \mapsto \|x_1 - x_2\|_2$$

qui est définie sur $\mathbb{R}^3 \times \mathbb{R}^3$ et est C^∞ pour $x_1 \neq x_2$. Au symbole *angle*, on associe la fonction :

$$f_a : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow [0, \pi] \\ (x_1, x_2, x_3) \mapsto \text{acos}\left(\frac{(x_1 - x_2) \cdot (x_3 - x_2)}{f_d(x_1, x_2)f_d(x_3, x_2)}\right)$$

où \cdot est le produit scalaire de \mathbb{R}^3 , et *acos* la fonction arc cosinus. Elle est définie et C^∞ dès que $x_1 \neq x_2$ et $x_3 \neq x_2$. Le symbole *aire* associe à trois points x_1, x_2, x_3 l'aire du parallélogramme de côtés $x_1 - x_2, x_3 - x_2$. Son interprétation n'est pas utilisée dans ce mémoire mais on peut toutefois utiliser un produit vectoriel pour le traduire en fonction numérique. Le symbole *volume* est interprété par la fonction :

$$f_v : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}_+ \\ (x_1, x_2, x_3, x_4) \mapsto \det(x_2 - x_1, x_3 - x_1, x_4 - x_1)$$

où $\det(\dots)$ est le déterminant de la matrice 3×3 dont les colonnes sont les vecteurs $x_i - x_j$. f_v est définie et C^∞ sur $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$. Au symbole *sphere* on associe la concaténation d'un vecteur de \mathbb{R}^3 (l'interprétation de la variable de sorte *point*) et d'un réel (l'interprétation de la variable de sorte *mesure*). Le symbole *interSSS* est interprété par la multifonction qui à trois points $(x_i, r_i) \in \mathbb{R}^3 \times \mathbb{R}$ associe l'ensemble des solutions du système d'équations en l'inconnue $x \in \mathbb{R}^3$:

$$\begin{cases} f_d(x, x_1) - r_1 = 0 \\ f_d(x, x_2) - r_2 = 0 \\ f_d(x, x_3) - r_3 = 0 \end{cases}$$

On choisit comme convention que cette fonction n'est pas définie quand ce système n'a aucune solution réelle, ou quand il en admet une infinité.

Les symboles *colineaires* et *coplanaires* sont dits *prédicatifs*, alors que les autres sont *fonctionnels*, car ils correspondent à la vérification de prédicats, en l'occurrence la colinéarité de trois points, où la coplanarité de trois points. Ils sont alors interprétés par des équations, qui sont satisfaites quand le prédicat l'est. Ainsi, au prédicat *coplanaires*(P_0, P_1, P_2, P_3), où P_i est interprété par x_i , on associe l'équation :

$$f_v(x_1, x_2, x_3, x_4) = 0$$

qui est vérifiée si et seulement si les points x_1, x_2, x_3, x_4 sont tous dans un même plan de \mathbb{R}^3 .

Univers 2D

Signature

Sortes :

mesure
point
droite

Symboles :

distance(point, point) → mesure
angle(droite, droite) → mesure
incidents(point, droite)

Interprétation des degrés de liberté

Les variables de sorte *point* et *droite* ont 2 degrés de liberté. Les variables de sorte *mesure* interviennent dans notre cadre comme des paramètres, et ne sont pas interprétées. Des contraintes de symbole *distance*, *angle* et *incidents* ont toutes un degré de restriction de 1.

L'extension de cette sémantique à un univers 3D est en général directe : les variables de sorte *point* et *plan* ont un degré de liberté de 3. Une variable de sorte *droite* a par contre 4 degrés de liberté. En effet, elle peut être représentée de manière unique comme les coordonnées de son intersection avec le plan de cote nulle, et deux angles.

Des contraintes de distance, d'incidence entre point et plan et d'angles entre deux droites ont toutes un seul degré de restriction, alors qu'une contrainte d'incidence entre un point et une droite a deux degrés de restriction.

Résolution de contraintes géométriques en guidant une méthode homotopique par la géométrie

Résumé

Suivant le domaine où on les sollicite, les solutions d'un système de contraintes géométriques (SCG) peuvent être :

- formelles et exactes : elles prennent par exemple la forme d'un plan de construction produisant toutes les solutions, obtenu en appliquant des règles dérivées de lemmes de géométrie. Beaucoup de SCG, surtout en 3D, résistent à cette approche ;
- numériques et approchées : elles sont les solutions d'un système d'équations construit à partir des contraintes et trouvées grâce à des méthodes numériques efficaces quand elles ne recherchent qu'une solution.

De par la nature des problèmes traités, chercher toutes les solutions conduit à une complexité exponentielle.

Les méthodes par continuation, ou homotopie, permettent d'obtenir toutes les solutions d'un système d'équations polynomiales. Leur application à des SCG est coûteuse et difficilement sujette aux raisonnements permis par l'origine géométrique du problème car elles opèrent hors de l'espace des figures géométriques.

Notre travail a pour objet la spécialisation d'une méthode par continuation à des SCG. La géométrie simplifie et justifie sa mise en œuvre dans l'espace des figures, où des raisonnements géométriques sont possibles. On aborde également les cas où l'ensemble de solutions d'un problème contient des éléments isolés et des continus. Des solutions proches d'une esquisse fournie par un utilisateur sont d'abord trouvées. La recherche d'autres solutions, malgré sa complexité exponentielle, est rendue envisageable par une approche itérative. Une nouvelle méthode de décomposition est proposée pour maîtriser le coût de la résolution.

Mots Clés

Résolution de contraintes géométriques, Méthodes par continuation, Homotopie, Méthodes Hybrides, Modélisation géométrique

Abstract

Depending on the required application field, the solutions of a geometric constraints system (GCS) are either :

- symbolic and exact such as construction plans, providing all the solutions, obtained by applying geometric rules. Many problems, mostly in a 3D context, resist to this approach ;
- or numerical and approximated : they are the solutions of a system of equations built from the constraints, provided by generic numerical methods that are efficient when only one solution is sought.

However, searching all the solutions leads to an exponential computation cost, due to the nature of problems.

Continuation methods, also called homotopic methods, find all the solutions of a polynomial system. Using them to solve systems of equations associated to systems of constraints is nevertheless costly. Moreover, combining them with geometric reasoning is a challenge, because they act in a projective complex space and not in the realizations space.

The aim of this work is to specialize a continuation method to GCS. Geometry is exploited to simplify and justify its adaptation in the space of realizations, so allowing geometric reasoning. Cases where the connected components of the solution space of a problem have heterogeneous dimensions are addressed. The method discussed here provides in a first step solutions that are similar to a sketch drawn by the user. Then a procedure is proposed to search new solutions. Its iterative nature seems to make the exponential complexity of this task bearable. A new decomposition method is proposed, that restrains the resolution cost.

Keywords

Geometric constraints solving, Continuation methods, Homotopy, Hybrid methods, Geometric modeling