

Algorithmes pour le calcul de cartes routières d'ensembles semi-algébriques

*Encadrement par Mohab Safey El Din, Professeur à Sorbonne Université
Membre honoraire de l'Institut Universitaire de France*

Préambule: ensembles semi-algébriques, applications et algorithmes

Un sous-ensemble S de \mathbb{R}^n est dit *semi-algébrique* lorsqu'il est défini par une disjonction *finie* de conjonctions de contraintes *polynomiales* à coefficients réels. De tels ensembles jouissent de propriétés de *finitude* et *stabilité* importantes : le nombre de composantes connexes de tels ensembles est fini ; aussi, la classe des ensembles semi-algébriques est stable par projection [4].

Ces propriétés constituent les pierres angulaires de nombreux développements fondamentaux pour la conception d'algorithmes permettant la résolution de systèmes polynomiaux à variables réelles.

Ceux-ci sont essentiels : les ensembles semi-algébriques apparaissent naturellement dans de nombreuses applications. Par exemple, en biologie, l'évolution du vivant est modélisée par des systèmes dynamiques à coefficients polynomiaux. Obtenir les points d'équilibres (stables ou instables) de tels systèmes revient à calculer les solutions réelles d'un système polynomial.

Un autre exemple d'applications qui va occuper une place prépondérante dans cette thèse est la robotique et la conception de mécanismes articulés. Mathématiquement, de tels mécanismes sont vus comme des applications (polynomiales) qui à un point de l'espace articulaire associe un point de l'espace de travail. Cette application est appelée application cinématique. Contrôler de tels mécanismes est évidemment crucial mais difficile. Une condition nécessaire est d'éviter les singularités de l'application cinématique. Ainsi, un problème important est de *compter* le nombre de composantes connexes du complémentaire de l'ensemble de ces singularités (qui est un ensemble semi-algébrique). Mieux, lorsque deux points d'un semi-algébrique vivent dans une même composante connexe de cet ensemble, on souhaite pouvoir construire (donc calculer) un chemin les reliant.

Du fait des propriétés de finitude mentionnées plus haut, ces problèmes algorithmiques sont bien fondés. Pour l'heure, ces derniers sont traités en pratique par des méthodes purement numériques qui présentent le défaut d'être *incomplètes*: n'étant pas certifiées, elles ne permettent pas de garantir que deux points sont connectés lorsqu'elles n'ont pas convergé. Ainsi, des *postures* peuvent être «manquées», la compréhension géométrique et l'exploitation du potentiel de ces mécanismes peuvent être sous-optimaux.

Dans cette thèse, on vise le développements d'algorithmes certifiés (dont les réponses sont complètes) sans sacrifier à la nécessité de passer à l'échelle pour traiter les applications mentionnées ci-dessus. Deux leviers importants seront développés:

- le développement de résultats mathématiques permettant de mieux comprendre la géométrie des ensembles semi-algébriques et leurs composantes connexes ;
- l'utilisation de méthodes de *calcul formel* (calcul algébrique) qui sont particulièrement bien adaptées pour la résolution *globale* et *garantie* des systèmes polynomiaux.

Le calcul formel en France, positionnement de l'équipe d'accueil

Les méthodes algébriques (i.e. de calcul formel) pour la résolution de systèmes polynomiaux constituent le levier qui va nous permettre de garantir l'exactitude des algorithmes de cette thèse.

Le calcul formel est surtout connu pour les systèmes de calcul formel MAPLE, MATHEMATICA ou SAGE-MATH qui fournissent une large panoplie d'algorithmes et implantations pour le calcul mathématique.

La communauté française développe surtout des bibliothèques spécialisées, afin de se concentrer sur l'optimisation algorithmes et des calculs en pratique. Parmi les grands acteurs du calcul formel en France, mentionnons l'équipe de J. van der Hoeven (fondée par M. Giusti) à l'école polytechnique et qui développe le logiciel MATHMAGIX ; l'équipe ARIC de l'ENS de Lyon, portée par B. Salvy qui développe GFUN (pour la manipulation de séries et fonctions D -finies) et, conjointement aux équipes de Grenoble et Montpellier, la bibliothèque LINBOX et enfin l'équipe AROMATH à Inria Sophia-Antipolis, portée par B. Mourrain et qui développe des bibliothèques (AXEL) pour la modélisation géométrique.

L'équipe de Sorbonne Université s'inscrit dans ce paysage ; sa singularité réside surtout dans le développement de plusieurs bibliothèques pour la résolution des systèmes polynomiaux (pour le calcul de bases de Gröbner – FGB – et la résolution sur les réels – RAGLIB). Elle est reconnue pour ses contributions qui ont permis des progrès pratiques et théoriques spectaculaires ces dernières années et son investissement dans le traitement des applications de la résolution des systèmes polynomiaux.

Elle a développé récemment une forte collaboration avec l'équipe de calcul formel au RICAM (Research Institute of Computational Mathematics) de l'Univ. J. Kepler à Linz en Autriche autour des outils mathématiques et algorithmiques pour la conception de mécanismes articulés. Des roboticiens autrichiens et français sont impliqués et le doctorant pourra ainsi confronter ses idées et contributions directement aux applications. Les aspects algorithmiques et mathématiques directement reliés à cette thèse ont fait l'objet de collaborations nombreuses avec É. Schost, de l'Université de Waterloo au Canada. Le doctorant effectuera des visites de quelques semaines dans cette université pour y développer une collaboration. Elle portera sur l'étude d'un des objets fondamentaux sur lesquels les développements algorithmiques de cette thèse s'appuient : il s'agit de la notion de *carte routière* d'un semi-algébrique.

État de l'art

Soit $S \subset \mathbb{R}^n$ un ensemble semi-algébrique et $\mathcal{P} \subset S$ un ensemble fini de points. Une *carte routière* pour (S, \mathcal{P}) est un ensemble semi-algébrique \mathcal{C} de dimension au plus un, qui contient \mathcal{P} et dont l'intersection avec chaque composante connexe de S est non vide et connexe. Ainsi, étant donné x et y dans \mathcal{P} , on peut décider si il existe un chemin dans S qui les relie en décidant si il existe une composante connexe C de \mathcal{C} qui les contient tous les deux. On a ainsi ramené un problème de connexité en dimension arbitraire à un problème de connexité en dimension au plus un. Dans ce dernier cas, on dispose déjà d'algorithmes convaincants. Le nœud de l'affaire consiste donc à calculer ces cartes routières. La notion de carte routière est introduite par J. Canny dans [5]. Dans la suite, pour simplifier, on suppose que $\mathcal{P} = \emptyset$.

Lorsque S est défini par des équations polynomiales de degré D , Canny a proposé un algorithme (probabiliste) permettant de calculer une carte routière. La méthode proposée s'appuie sur un «résultat de connexité» qu'on énonce comme suit. Notons π_i la projection canonique $(x_1, \dots, x_n) \rightarrow (x_1, \dots, x_i)$. Sous l'hypothèse que S est compacte et lisse, et qu'on soit en coordonnées génériques, on considère le lieu critique de la restriction de la projection π_2 à S ; on le note $W(\pi_2, S)$. Il s'agit d'une courbe (critique). Comme illustré sur l'exemple du tore (ici π_2 est la projection sur un plan perpendiculaire à l'axe de révolution), cette courbe intersecte chaque composante connexe de S (grâce à la compacité) mais ces intersections ne sont pas nécessairement connexes.

Il faut donc réparer ces «défauts de connexité» en rajoutant des courbes à notre lieu critique. Ces nouvelles courbes sont obtenues en considérant des tranches bien choisies de S . On considère ainsi les valeurs critiques de la restriction de π_1 à \mathcal{C} . Une version algébrique du théorème de Sard assure qu'elles sont en nombre fini (voir par exemple [9]). On les note v_1, \dots, v_ℓ et on considère les tranches $\cup_{i=1}^\ell (S \cap \pi_1^{-1}(v_i))$. Canny prouve alors que l'union de ces tranches avec $W(\pi_2, S)$ a une intersection non vide et connexe avec chaque composante connexe de S . La dimension de ces intersections est inférieure à celle de S . Ainsi, si S était de dimension 2, on obtient une carte routière de S . Le procédé est illustré par la figure ci-jointe. Lorsque la dimension de S est supérieure à 2 strictement, on profite de la chute de

dimension pour faire un appel récursif à l'algorithme. Ainsi, si on part d'un ensemble de dimension d , on se retrouve à effectuer un appel récursif en dimension $d - 1$, puis $d - 2$, et ainsi de suite jusqu'à la dimension 1. La profondeur de l'algorithme récursif ainsi obtenu est de l'ordre de n .

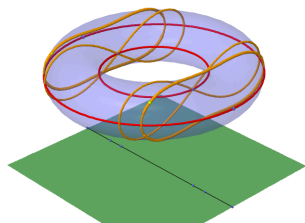


Figure 1. calcul de carte routière

À ce stade, il convient de « mesurer » la taille des objets calculés. Il faut d'abord noter qu'ils sont tous de nature algébrique. En effet, les courbes critiques sont définies par des systèmes algébriques ; leur « complexité » est mesurée par le degré de la courbe algébrique définie par ce système dans \mathbb{C}^n , c'est à dire le nombre maximal de points obtenus en intersectant ces courbes avec un hyperplan génériquement choisi. Le théorème de Bézout assure qu'en partant d'équations polynomiales de degré D , une courbe critique a un degré borné par $D^{O(n)}$. Les valeurs critiques de la restriction de π_1 à \mathcal{C} sont elles de degré en $D^{O(n)}$. Ainsi une analyse de complexité facile montre que les objets calculés par l'algorithme de Canny sont de degré borné par $D^{O(n^2)}$. Ainsi, tous les travaux (voir entre autres [7, 2]) qui ont suivi les premiers résultats de Canny qui continuaient de s'appuyer sur son résultat de connectivité, ne pouvait espérer obtenir mieux qu'une complexité en $D^{O(n^2)}$ (des versions déterministes ont néanmoins été obtenues).

Une nouvelle famille d'algorithmes pour le calcul de cartes routières

Si les résultats obtenus par Canny et la famille d'algorithmes qui lui sont associés ont rencontré un certain écho, le caractère exponentiel en n^2 de leur complexité n'est pas satisfaisant. En effet, il est connu que le nombre de composantes connexes d'un semi-algébrique est borné par $(sD)^{O(n)}$ (où s est le nombre d'inégalités). Cette borne pouvant être atteinte dans les pires situations, on ne peut espérer mieux qu'une complexité simplement exponentielle en n . Néanmoins, un coût exponentiel en n^2 n'est pas justifié.

Il faut attendre [8], presque 25 ans après publication des travaux de Canny, pour obtenir une percée significative. L'algorithme obtenu dans [8] permet en effet de calculer des cartes routières d'hypersurfaces lisses et compactes en temps $D^{O(n\sqrt{n})}$. Le levier qui a permis cette avancée est un nouveau résultat de connectivité pour les *variétés algébriques réelles* (des ensembles semi-algébriques définis uniquement par des équations) lisses et compactes. Le point de départ réside dans l'observation que pour obtenir une meilleure complexité il faudrait pouvoir réduire la profondeur de la récursion. Le schéma de résolution dû à Canny est trop contraint pour cela : on calcule une courbe critique (donc de dimension 1) et répare ses défauts de connectivité en y ajoutant des tranches de co-dimension 1 dans le semi-algébrique étudié.

Il faut donc se donner plus de liberté. Sous l'hypothèse que notre semi-algébrique S est une variété lisse, équi-dimensionnel et compacte, le point de départ consiste à considérer le lieu critique de la restriction de π_i à S ; on le note $W(\pi_i, S)$. De manière analogue à Canny, on peut alors considérer le lieu critique \mathcal{H} de la restriction de π_1 à $W(\pi_i, S)$ mais au lieu de projeter ces points sur l'axe de coordonnées x_1 , on le projette sur le sous-espace de coordonnées (x_1, \dots, x_{i-1}) ; on note $F_{i-1} \subset \mathbb{R}^{i-1}$ l'ensemble des points obtenus. Le nouveau résultat de connexité établit que $\mathcal{H} = W(\pi_i, S) \cup (S \cap \pi_{i-1}^{-1}(F_{i-1}))$ a une intersection connexe avec chaque composante connexe de S , sous réserve que \mathcal{H} soit un ensemble fini de points.

Il est établi dans [8] qu'en coordonnées génériques, \mathcal{H} est fini, et que si $i \leq \frac{\dim(S)}{2}$, $W(\pi_i, S)$ est de dimension au plus $i - 1$, lisse et équi-dimensionnel.

Aussi, en jouant sur les valeurs de i choisies, on peut réduire la profondeur de la récursion. Si on choisit $i \simeq \lfloor \sqrt{n} \rfloor$, on obtient un schéma de résolution du type « pas de bébé – pas de géant », de profondeur \sqrt{n} qui aboutit à une complexité en $D^{O(n\sqrt{n})}$. Le cas des hypersurfaces lisses et compactes est le cas de base auquel on peut ramener des situations plus générales. C'est ce que nous avons fait dans [1] mais le prix à payer est l'introduction d'un nombre constant d'infinitésimaux pour rendre effectives des techniques de déformation. Ceci est bon du point de vue de la complexité théorique mais l'alourdissement du coût des

opérations arithmétiques de base rend ces stratégies impraticables en vue d'un déploiement applicatif.

Évidemment, si on choisit $i \simeq \lfloor \log_2(n) \rfloor + 1$, on obtient une profondeur en $\simeq \log_2(n)$. La mise en œuvre de ce schéma «diviser pour régner» induit des complications techniques importantes qu'on ne détaillera pas ici. Elles sont surmontées une première fois dans [3] mais le prix à payer est aussi prohibitif en pratique : le nombre d'infinitésimaux introduit est de l'ordre de $O(\log(n))$ (!). Sous l'hypothèse que S est une variété lisse équi-dimensionnelle et compacte, on obtient enfin dans [9] une complexité en $(nD)^{O(n \log(n))}$, ce qui est proche de l'optimal et sans introduire d'infinitésimaux. Le but de cette thèse est de développer les méthodologies nécessaires pour généraliser ces derniers résultats aux situations non compactes et semi-algébriques, tout en visant un niveau de performance pratique qui permet de traiter les applications en robotique précédemment mentionnées.

Lever l'hypothèse de compacité

On considère ici un ensemble $S \subset \mathbb{R}^n$ qu'on suppose être une variété algébrique réelle lisse et équadimensionnelle mais on ne suppose pas que S soit compact. L'ensemble S est donné implicitement comme solution d'un système d'équations polynomiales de degré borné par D . Nous élaborons ci-dessous les pistes de recherche qui permettraient d'obtenir un algorithme pour le calcul de cartes routières et de complexité $(nD)^{O(n \log(n))}$.

Une première étape est de se doter d'un résultat de connectivité plus général que celui élaboré dans [8] et qui suppose que S soit compact. Cette hypothèse de compacité intervient à plusieurs endroits de la preuve qui, elle-même s'appuie sur deux résultats intermédiaires qu'on énonce ci-dessous.

Lemme 1. *Soit Z une variété algébrique réelle lisse et compacte, $x \in \mathbb{R}$ tel que $(Z \cap \pi_1^{-1}(x)) \cap W(\pi_1, Z) = \emptyset$ et une composante connexe C de $Z \cap \pi_1^{-1}(]-\infty, x])$. Alors, si $C \cap \pi_1^{-1}(]-\infty, x[)$ est non vide, il est connexe.*

Lemme 2. *Soit Z une variété algébrique réelle lisse et compacte. Soit $v < w$ dans \mathbb{R} tel que $(Z \cap \pi_1^{-1}(]v, w[)) \cap W(\pi_1, Z) = \emptyset$ et C une composante connexe de $Z \cap \pi_1^{-1}(]-\infty, w[)$. Alors $C \cap \pi_1^{-1}(]-\infty, v[)$ est une composante connexe de $Z \cap \pi_1^{-1}(]-\infty, v[)$.*

Les points les plus sensibles de la preuve s'appuient sur des versions semi-algébriques du théorème de fibration d'Ehresmann. Mais ce dernier théorème est valide lorsque la restriction de π_1 est propre (c'est-à-dire lorsque pour tout $a \in \mathbb{R}$, il existe $\varepsilon > 0$ tel que $Z \cap \pi_1^{-1}([a - \varepsilon, a + \varepsilon])$ soit compact).

Ces deux lemmes sont ensuite utilisés pour montrer que les intersections des composantes connexes de $S \cap \pi_1^{-1}(]-\infty, a])$ avec celles de $\mathcal{R} \cap \pi_1^{-1}(]-\infty, a])$ sont non vides et connexes lorsque a parcourt \mathbb{R} en allant de $-\infty$ à $+\infty$. Là encore, la compacité de S est utilisée d'une part pour montrer que l'intersection est non vide et d'autre part pour assurer la connexité lorsqu'on traverse une valeur critique de π_1 . Mais, pour cette dernière étape, il suffit là encore d'assurer la propriété de π_1 .

Ainsi, généraliser notre résultat de connexité aux situations non compactes est un pari raisonnable si on sait remplacer π_1 par une fonction polynomiale dont on peut assurer la propriété a priori. Une voie facile est alors de substituer π_1 par la fonction $x \rightarrow \|x\|^2$ qui est, par construction, propre.

L'obtention d'un algorithme de calcul de cartes routières s'appuyant sur un tel résultat de connectivité généralisé n'est néanmoins pas encore immédiat. Il faut en effet contrôler les dimensions des objets qui sont les analogues de $W(\pi_i, S)$ and $S \cap \pi_1^{-1}(F_{i-1})$. Dans le cas compact, où l'on considèrerait la projection π_1 , on pouvait pour cela s'appuyer sur des théorèmes de transversalité (qui sont des analogues algébriques des résultats de Mather) pour les variétés dites polaires. Cela garantit que nos besoins sont satisfaits en coordonnées génériques. Le remplacement de π_1 par «une fonction distance» rend cette étape assurément plus délicate. Il faut néanmoins remarquer qu'on retrouve une certaine marge de manœuvre pour appliquer des résultats de transversalité si on s'autorise à mettre S en coordonnées générique avant de considérer tout lieu critique d'une part et si on s'autorise ensuite à translater génériquement l'origine de notre espace. Nous avons donc bon espoir d'aboutir à un algorithme de complexité $(nD)^{O(n \log(n))}$ sans hypothèse de compacité pour les variétés algébriques lisses et équi-dimensionnelles.

Passage au cas semi-algébrique

Cela ne sera néanmoins pas suffisant pour le traitement d'applications en robotique : celles-ci font naturellement intervenir des semi-algébriques. On considère donc le cas d'un semi-algébrique défini par

$$f_1 = \dots = f_p = 0, \quad g_1 \geq 0, \dots, g_s \geq 0$$

avec les f_i et les g_j dans $\mathbb{Q}[x_1, \dots, x_n]$. On fera l'hypothèse de régularité (raisonnable pour les applications visées) suivante : pour tout $\{i_1, \dots, i_\ell\} \subset \{1, \dots, s\}$, la suite $(f_1, \dots, f_p, g_{i_1}, \dots, g_{i_\ell})$ est régulière et réduite. Grâce aux résultats intermédiaires attendus pour la levée de l'hypothèse de compacité, on devrait pouvoir obtenir encore une généralisation du théorème de connectivité au cas semi-algébrique (mais sous les hypothèses de régularité qu'on vient de décrire). En effet, on peut s'appuyer sur l'extension de la notion de points et valeurs critiques aux semi-algébriques et qui consiste à considérer les points et valeurs critiques des fonctions considérées sur les bords définis par les suites régulières réduites $(f_1, \dots, f_p, g_{i_1}, \dots, g_{i_\ell})$. L'hypothèse de régularité assure de bonnes propriétés topologiques (on peut obtenir «facilement» une stratification de Whitney de S) qui permettent alors de s'appuyer sur les versions semi-algébriques du théorème d'isotopie de Thom pour mimer la preuve du résultat de connectivité qu'on obtiendra pour les variétés algébriques lisses et équi-dimensionnelles.

Ici encore, le passage du résultat de connectivité à l'algorithme nécessitera de nouveaux résultats. Le fait de devoir considérer le bord de S induira facteur combinatoire dans la complexité qui devrait alors être en $s^n (nD)^{O(n \log(n))}$. Il faut noter qu'à ce stade, nous n'avons toujours pas introduit d'infinitésimaux.

Applications et réalisations logicielles

Cela nous conforte dans l'espoir que les algorithmes obtenus pourront passer à l'échelle. Cet espoir est renforcé par [6] où en exploitant la structure du problème nous parvenons à calculer des cartes routières pour l'analyse des singularités cinématiques d'une familles de robots sériels vendus pour l'industrie par l'entreprise danoise Universal Robots. Il s'agit là d'une démonstration de la faisabilité du calcul de cartes routières dans un contexte applicatif qui a un impact industriel avéré. L'objectif ultime de cette thèse est ainsi d'implanter les algorithmes obtenus avec un niveau d'efficacité suffisant pour étudier d'autres familles de robots et c'est ici que la collaboration avec nos collègues roboticiens est un atout.

References

- [1] S. Basu, M.-F. Roy, M. Safey El Din, and É. Schost. "A Baby Step-Giant Step Roadmap Algorithm for General Algebraic Sets". In: *Foundations of Computational Mathematics* 14.6 (2014), pp. 1117–1172.
- [2] S. Basu, R. Pollack, and M.-F. Roy. "Computing roadmaps of semi-algebraic sets on a variety". In: *Journal of the American Mathematical Society* 13.1 (2000), pp. 55–82.
- [3] S. Basu and M.-F. Roy. "Divide and conquer roadmap for algebraic sets". In: *Discrete & Computational Geometry* 52.2 (2014), pp. 278–343.
- [4] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*. Springer-Verlag, 1998.
- [5] J. Canny. *On the complexity of robot motion planning*. MIT Press, 1988.
- [6] J. Capco, M. Safey El Din, and J. Schicho. "Robots, computer algebra and eight connected components". In: *submitted to the Proceedings of the International Symposium on Symbolic and Algebraic Computation*. 2020.
- [7] L. Gournay and J.-J. Risler. "Construction of roadmaps in semi-algebraic sets". In: *Applicable Algebra in Engineering, Communication and Computing* 4.4 (1993), pp. 239–252.
- [8] M. Safey El Din and É. Schost. "A baby steps/giant steps probabilistic algorithm for computing roadmaps in smooth bounded real hypersurface". In: *Discrete & Computational Geometry* 45.1 (2011), pp. 181–220.
- [9] M. Safey El Din and É. Schost. "A nearly optimal algorithm for deciding connectivity queries in smooth and bounded real algebraic sets". In: *Journal of the ACM (JACM)* 63.6 (2017).