

Géométrie algorithmique : des données géométriques à la géométrie des données

Jean-Daniel Boissonnat

▶ **To cite this version:**

Jean-Daniel Boissonnat. Géométrie algorithmique : des données géométriques à la géométrie des données. Collège de France / Fayard, 2017. hal-01975080

HAL Id: hal-01975080
https://inria.hal.science/hal-01975080

Submitted on 9 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Géométrie algorithmique : des données géométriques à la géométrie des données

Jean-Daniel Boissonnat
Leçon inaugurale 23 mars 2017

18 juillet 2017

L'un des plus grands géomètres du XXème siècle, Harold Coxeter, écrivait en 1967 :

> *La géométrie possède encore l'attrait esthétique qu'elle a toujours eu, et la beauté de ses résultats ne s'est pas estompée. En fait, elle est plus utile et même plus nécessaire aux savants et aux mathématiciens qu'elle ne le fut jamais.*
> *Au cours des siècles, la géométrie s'est développée. De nouveaux concepts, de nouvelles méthodes d'action furent forgés. Par les moyens qui nous conviendront le mieux, revenons donc à Euclide ; et découvrons quelques-uns des plus récents résultats. Peut-être pourrons-nous, ainsi, retrouver un peu de l'intimidation émerveillée que suscita en nous le premier contact avec la géométrie...*

Quel plus bel endroit que le Collège de France pour répondre à l'invitation d'Harold Coxeter et présenter un nouveau développement de la géométrie ?

Le sujet de ce cours, que les professeurs du Collège me font l'honneur de me confier, est la *géométrie algorithmique*[1]. La géométrie, dont les origines se confondent avec celles de l'humanité[2], ne pouvait pas ignorer les évolutions spectaculaires de l'informatique et des sciences du numérique décrites ici même par Gérard Berry et mes prédécesseurs à cette chaire. Cette leçon inaugurale est une invitation à découvrir le rôle que la géométrie joue dans notre monde devenu numérique et quelques unes des nouvelles questions nées du rapprochement de la géométrie et de l'informatique.

1 Les origines de la géométrie algorithmique

La conception assistée par ordinateur

L'histoire de la géométrie algorithmique commence dans les années 1970 avec la conception assistée par ordinateur (CAO). Il s'agit de résoudre un problème indus-

1. Computational Geometry en anglais.
2. M. Serres. *Les origines de la géométrie.* Flammarion, 1993.

triel : créer un outil informatique simple qui permette aux dessinateurs de modéliser des surfaces en trois dimensions, comme des carrosseries de voitures, et qui facilite la programmation des machines à commande numérique. Deux ingénieurs français, Pierre Bézier qui travaillait chez Renault et Paul de Casteljeau qui travaillait chez Citröen, vont fonder le sujet. Leurs contributions seront majeures : les courbes et les surfaces de Bézier et les algorithmes de de Casteljeau qui les construisent sont universellement utilisés aujourd'hui.

En utilisant l'informatique, science de l'immatériel, les ingénieurs ont franchi un pas décisif : la CAO s'affranchit des maquettes physiques, lourdes et encombrantes, et des représentations 2d réductrices, plans, dessins industriels, coupes, ou projections. Elle leur substitue une véritable représentation 3d, virtuelle mais effective. Ces nouveaux modèles objectivent la représentation de formes tridimensionnelles et en simplifient radicalement le traitement, la reproduction et la transmission. Ils ouvrent de nouveaux champs d'applications aux créateurs de formes tridimensionnelles, ingénieurs, médecins [3] ou artistes [4].

L'invention de la numérisation 3d

Une deuxième étape décisive dans le développement de la modélisation 3d va être l'invention de capteurs permettant de numériser des formes tridimensionnelles. Permettez moi d'évoquer un souvenir personnel. C'était en 1980 à l'IRIA, l'institut de recherche en informatique et en automatique dirigé par Jacques-Louis Lions. Je venais juste d'être recruté comme chercheur dans le groupe de François Germain. Avec George Kryzé, François Germain avait inventé un capteur 3d novateur pour l'époque. Ce capteur comportait un faisceau laser dont la direction était contrôlé à l'aide de deux miroirs, et trois caméras linéaires qui repéraient le point éclairé par le laser, ce qui permettait de calculer par triangulation les coordonnées dans l'espace du point éclairé. Plus besoin de palpeurs et de systèmes mécaniques complexes ; pour la première fois, un système simple, rapide et précis permettait de mesurer des formes tridimensionnelles.

Les progrès de la numérisation 3d ont ensuite été spectaculaires. Des descendants du prototype de F. Germain et G. Kryzé se trouvent aujourd'hui sur les étagères de grandes surfaces et seront bientôt intégrés dans nos téléphones portables. De nombreux autres systèmes de numérisation 3d ont été inventés, révolutionnant de nombreuses applications : la tomographie, très utilisée en imagerie médicale, ainsi qu'en géophysique, en astrophysique et en mécanique des matériaux, la télédétection par laser (LIDAR), la microscopie confocale couramment utilisée en biologie et en science des matériaux, la cryo-microscopie électronique qui permet d'étudier la structure de complexes macromoléculaires. On mesure aujourd'hui de façon routinière des formes tridimensionnelles, de l'échelle atomique à l'échelle astronomique, donnant ainsi accès à une dimension inaccessible à l'œil humain. Cette fabrique du regard 3d

3. N. Ayache. *Des images médicales au patient numérique.* Collège de France/Fayard 2015.

4. M-P. Cani. *Façonner l'imaginaire : de la création numérique 3d aux mondes virtuels animés.* Collège de France/Fayard 2015.

(pour reprendre l'expression de Monique Sicard [5]) ouvre des applications inédites. Le monde numérique n'est maintenant plus limité au texte, au son et aux images, et les représentations numériques de formes tridimensionnelles jouent un rôle central dans de très nombreux domaines. Citons, parmi bien d'autres, l'ingénierie, la cartographie, le cinéma et les jeux vidéos, l'architecture, la préservation du patrimoine culturel, l'exploration pétrolière, la médecine, la conception de médicaments.

La question de la complexité algorithmique

L'approche des ingénieurs de la CAO est avant tout destinée à la modélisation de formes relativement simples. Elle est bien résumée par la jolie expression "Formes et formules". Mais elle ne permet pas de modéliser les formes complexes de la nature ou les statues de Michel Ange. L'apparition de données 3d va changer la perspective. Les formes sont très variées, les données massives : plusieurs millions de points peuvent être nécessaires pour représenter précisément un objet complexe. La taille des données et la complexité des algorithmes deviennent alors critiques.

Ces questions sont au centre de la thèse que M. Shamos soutient à Yale en 1978 et qu'il intitule "Computational Geometry". M. Shamos est un théoricien de l'algorithmique qui veut étudier les interactions entre géométrie et calcul. Son objectif est de fonder un nouveau sujet en étudiant de manière systématique les problèmes géométriques du point de vue algorithmique. Voici un extrait de la préface de sa thèse :

> *Cette thèse examine les questions qui se posent quand on cherche à résoudre des problèmes géométriques avec un ordinateur, ce qui nous oblige à considérer des aspects du calcul géométrique qui ne sont simplement pas abordés par les mathématiques classiques. De nouvelles méthodes sont requises.*

Un problème résolu par M. Shamos va permettre de comprendre sa démarche. Il s'agit de trouver dans un ensemble de n points du plan les deux points qui sont les plus proches. C'est une question d'apparence simple. Il suffit de considérer tous les segments reliant deux points et de retenir le plus court. Comme il y a $n \times (n-1)/2$ paires de points, cet algorithme naïf a une complexité quadratique. L'algorithme de Shamos évite de calculer toutes les paires, et effectue un nombre d'opérations élémentaires proportionnel à $n \log n$. On dit que sa complexité est en $O(n \log n)$ [6]. Le temps de calcul se trouve divisé par un facteur considérable quand n est grand : par exemple, si $n = 10.000.000$ (le nombre de pixels dans un appareil photo ordinaire), le temps de calcul passe de la semaine à la seconde.

5. M. Sicard. *La Fabrique Du Regard.* Images de science et appareils de vision (15ème-20ème siècle). Editions Odile Jacob 1998.

6. La notation ignore la constante de proportionnalité pour mettre en évidence l'ordre de grandeur de la complexité pour n grand.

2 Structures de données géométriques

À la différence des images numériques, les données 3d sont peu ou pas structurées car les informations qu'on peut recueillir sur une forme 3d ne sont en général que partielles et il est nécessaire de combiner les données acquises de plusieurs points de vue : Flatland, le chef d'œuvre d'Edwin Abbott, est une démonstration pédagogique brillante de cette limitation.. Si la thèse de M. Shamos ne répond pas à la question de la modélisation des formes géométriques 3d à partir de mesures, elle va fournir quelques clés essentielles à travers la définition et la construction de *structures de données géométriques*. En informatique, une structure de données est une manière d'organiser les données en un modèle interne qui permet de les traiter plus facilement. Cette notion n'est pas très éloignée de celle de représentation en psychologie et de modèle interne en neurosciences. Notre cerveau est prodigieusement doué pour cartographier notre environnement et en construire des représentations internes qui vont lui permettre de planifier des actions. C'est ce que fait un enfant qui apprend à marcher. C'est aussi ce que fait un chirurgien qui regarde des images tomographiques et doit, à partir de ces images bidimensionnelles, construire une représentation mentale tridimensionnelle de l'organe qu'il va opérer.

Peut-on imaginer construire in silico des représentations qui intégrent les mesures géométriques aujourd'hui disponibles et permettent d'effectuer des calculs efficacement, reprenant le schéma représentation-plan-programme que décrit Marc Jeannerod dans son livre *Le cerveau volontaire*[7] ?

Un des succès de la géométrie algorithmique, est d'avoir identifié et étudié quelques structures géométriques discrètes fondamentales, qui vont permettre de structurer les données.

Polyèdres convexes, mosaïques et triangulations

Un premier exemple d'une telle structure géométrique discrète est celui des polyèdres convexes qui sont connus depuis la haute antiquité et n'ont pas cessé depuis de fasciner les plus brillants mathématiciens comme l'attestent les gravures de Léonard de Vinci et de Képler. Représenter le monde qui nous entoure avec des polyèdres convexes peut sembler outrageusement réducteur et réservé aux amateurs de beauté formelle. Pourtant, comme on le verra plus loin, les polyèdres convexes jouent un rôle central en modélisation géométrique. On peut y voir un nouvel exemple de la déraisonnable efficacité des mathématiques dans les sciences.

Un deuxième exemple de modèle géométrique discret est celui des diagrammes (ou mosaïques) de Voronoï, du nom du mathématicien ukrainien George Voronoï. En mathématiques, un diagramme de Voronoï est un découpage en cellules à partir d'un ensemble fini de points qu'on va appeler « germes », les points rouges sur la figure 1. Chaque cellule enferme un seul germe, et forme l'ensemble des points du plan plus proches de ce germe que des autres. Les diagrammes de Voronoï représentent des relations de proximité et permettent de modéliser des phénomènes de croissance en

7. M. Jeannerod. *Le cerveau volontaire.* Editions Odile Jacob, 2009.

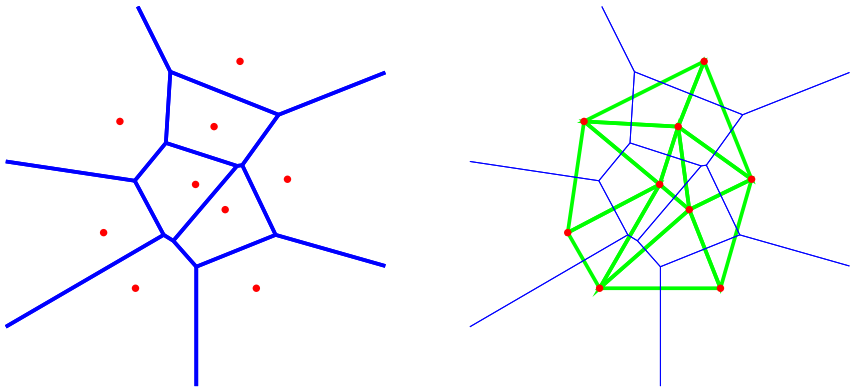


FIGURE 1 – Le diagramme de Voronoï d'un ensemble fini de germes P(à gauche) et la triangulation de Delaunay duale à droite.

toutes dimensions. Bien avant Voronoï, Descartes les avait dessinés pour décrire le système solaire. Les diagrammes de Voronoï ont été utilisés (sous différents noms) pour modéliser des territoires en écologie, en épidémiologie et en sciences humaines. Les géographes peuvent les observer sur le sol plat du salar d'Uyuni, les zoologistes sur le dos des girafes réticulées et les carapaces des tortues, les botanistes sur les méristèmes, les chimistes dans les mousses.

Les *triangulations* sont un troisième exemple de modèle géométrique discret universel, au cœur de l'interaction entre géométrie et calcul. En approchant la géométrie des objets par des assemblages de simplexes – triangles, tétraèdres, ou leurs analogues en dimensions supérieures – , les triangulations permettent de visualiser les objets, de faire des calculs sur ces objets, de résoudre des équations aux dérivées partielles, de simuler numériquement des phénomènes physiques. Pour des besoins très différents et avec des contraintes qui leur sont propres, l'analyse numérique, l'informatique graphique, la robotique triangulent leurs objets d'étude, espaces fonctionnels, mondes virtuels ou espaces de configuration.

Un ensemble de points admet de très nombreuses triangulations. L'une d'elles se détache par la richesse de ses propriétés mathématiques et de ses applications. C'est la triangulation de Delaunay [8], définie en 1934 par Boris Delaunay, un mathématicien russe qui a dédié son article fondateur à G. Voronoï. La triangulation de Delaunay est duale du diagramme de Voronoï dont elle se déduit en reliant les germes dont les cellules de Voronoï partagent une arête (Figure 1).

Une correspondance très utile peut être établie entre diagrammes de Voronoï, triangulations de Delaunay et polyèdres convexes. Il faut, pour cela, quitter l'espace

8. B. Delaunay. *Sur la sphère vide.* Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk, 7 (1934).

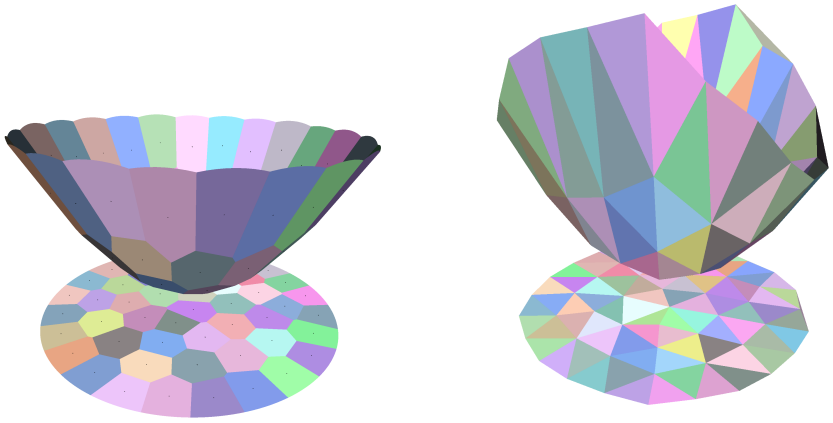


FIGURE 2 – Correspondance entre diagrammes de Voronoï, triangulations de Delaunay et polyèdres convexes.

de départ qui contient les points de données et se placer dans un espace qui a une dimension de plus. Le diagramme de Voronoï de n points du plan apparait alors comme la projection d'un polyèdre convexe, intersection de n demi-espaces de l'espace à trois dimensions (figure 2). Et la triangulation de Delaunay de ces mêmes points apparait comme la projection d'un polyèdre convexe dual. Ce polyèdre est défini en relevant les points de départ dans cet espace de dimension 3, et en prenant leur enveloppe convexe, c'est-à-dire le plus petit convexe qui les contient. Ainsi les diagrammes de Voronoï et les triangulations de Delaunay du plan ne sont que des polyèdres convexes de dimension 3 déguisés. De même, les diagrammes de Voronoï et les triangulations de l'espace de dimension 3 sont la projection de polyèdres de dimension 4. D'une manière générale, la construction d'une triangulation de Delaunay de l'espace de dimension d ou, par dualité, d'un diagramme de Voronoï se ramène à la construction d'une enveloppe convexe en dimension $d + 1$. Il est donc important de pouvoir construire efficacement des enveloppes convexes de points, et de pouvoir le faire au delà de la dimension trois.

Algorithmes

Une enveloppe convexe de n points de l'espace à 2 dimensions est une suite d'au plus n arêtes reliées par des points de données. Pour la calculer, le mathématicien Ronald Graham a proposé un algorithme optimal très simple qui pourrait être étudié dans les collèges. On commence par ordonner les points autour d'un point central, ce qui fournit un polygone qui passe par tous les points. On parcourt ce polygone dans le sens trigonométrique à partir de son point le plus à gauche. On avance tant qu'on tourne à gauche. Si on tourne à droite, on recule en éliminant le sommet concave. Il est facile de voir que la complexité de cet algorithme est $O(n \log n)$. En effet, il est bien connu que la complexité du tri initial est $O(n \log n)$. Le coût du parcours qui suit est linéaire puisque soit on avance le long du polygone soit on recule et dans ce cas on supprime un point qu'on n'examinera plus.

La situation se complique un peu en dimension 3. Une enveloppe convexe de points

est un assemblage de facettes (en général triangulaires) qui sont attachées par des arêtes qui elles-mêmes ont pour extrémités des points de données. Combien y a-t-il d'arêtes et de facettes ? La réponse à cette question est donnée par la fameuse formule d'Euler

$$s - a + f = 2$$

qui relie les nombres de sommets s, d'arêtes a et de facettes f de tout polyèdre convexe. On peut en déduire facilement que les nombres d'arêtes et de faces dépendent linéairement du nombre de sommets. On dira que la *complexité combinatoire* de l'enveloppe convexe est linéaire. On déduit de la correspondance vue plus haut que la complexité combinatoire du diagramme de Voronoï de points du plan est linéaire, et il en va de même par dualité pour la triangulation de Delaunay. Il faut faire un saut dans le temps de plus de 200 ans pour que ces résultats soient étendus en toutes dimensions et borner la complexité combinatoire d'un polyèdre convexe de dimension d qui a n sommets. Comme le montre Peter Mc Mullen en 1971 [9], les polyèdres convexes voient leurs nombres de faces (de toutes dimensions, de 0 jusqu'à $d-1$) croître de manière exponentielle avec la dimension d de l'espace ambiant. On en déduit que la complexité des diagrammes de Voronoï et des triangulations est quadratique en dimension 3 dans le cas le pire, et qu'elle croît de manière exponentielle avec la dimension d de l'espace, ce qui est une difficulté majeure si on veut utiliser les diagrammes de Voronoï en grandes dimensions. J'y reviendrai.

Maintenant qu'on connait la complexité combinatoire d'une enveloppe convexe, il reste à la calculer. L'algorithme de Graham résoud le problème dans le plan de façon optimale. Qu'en est-il en dimension 3 ? On a vu que l'enveloppe convexe de n points en dimension 3 a un nombre linéaire d'arêtes et de faces mais peut-on la calculer en temps $O(n \log n)$, ce qui serait optimal ? La réponse est positive mais les algorithmes connus qui ont cette complexité sont difficiles à programmer. En revanche, il existe un algorithme incrémental très simple qui insère les points un à un et met à jour l'enveloppe convexe après chaque insertion. Cet algorithme est quasiment aussi simple que l'algorithme de Graham et pourrait lui aussi être enseigné au collège. Mais cet algorithme n'est pas optimal et sa complexité en dimension 3 est quadratique dans le cas le pire. Comment concilier le point de vue du programmeur qui recherche un algorithme simple et efficace en pratique, et celui du théoricien qui recherche un algorithme optimal dans tous les cas, ce qui peut conduire à des constructions complexes peu efficaces pour les valeurs de n utilisées en pratique.

C'est l'introduction d'un part de hasard qui va permettre de concilier les points de vue et constituer une des avancées les plus fécondes en géométrie algorithmique et, au delà, dans toute l'algorithmique [10]. Le premier algorithme randomisé [11] est un algorithme géométrique dû à Michael Rabin, prix Turing 1976, mais ce sont Ken Clarkson et Peter Shor [12] qui, dans les années 1990, vont introduire un cadre général pour concevoir et analyser des algorithmes géométriques randomisés. Ils montrent

9. P. Mc Mullen. *The maximum number of faces a convex polytope.* Mathematika (1971).

10. B. Chazelle. *L'algorithmique et les sciences.* Collège de France/Fayard 2013.

11. Un mot français ancien moins usité que sa version anglaise, qui a la même racine que randonner.

12. K. Clarkson, P. Shor. *Applications of Random Sampling in Computational Geometry II.* Discrete and Computational Geometry, 4 (1989).

en particulier que si les points sont insérés dans un ordre aléatoire, l'algorithme incrémental de construction d'une enveloppe convexe est optimal en toutes dimensions. L'algorithme effectue des choix aléatoires mais il fournit toujours la solution exacte, et aucune hypothèse n'est faite sur les données d'entrée. C'est uniquement le déroulement de l'algorithme et non son résultat qui comporte une part de hasard. L'introduction de hasard a donc permis d'adapter un algorithme très simple et, tout en lui gardant sa simplicité, de le rendre optimal en moyenne et très efficace en pratique (on peut calculer la triangulation de Delaunay d'un million de points en dimension 3 en quelques secondes). De très nombreux travaux auxquels nous avons participé activement [13], ont permis d'étendre le travail fondateur de Clarkson et Shor. Plusieurs livres ont été écrits sur le sujet et les algorithmes randomisés occupent aujourd'hui une place essentielle en géométrie algorithmique.

Le calcul géométrique

Calculer vite est important mais calculer de manière fiable l'est encore plus. Cette question a été ignorée des premiers travaux de géométrie algorithmique ; la thèse de M. Shamos utilise pour analyser la complexité des algorithmes un modèle d'ordinateur capable d'effectuer toutes les opérations élémentaires sur les nombres *réels* exactement et en un coût unitaire, ce qui n'est pas réaliste puisque les ordinateurs ne savent représenter que des nombres de précision finie. Le modèle simpliste de M. Shamos, en évacuant les questions arithmétiques, permet d'utiliser la puissance du continu et de s'appuyer aveuglément sur un immense corpus de résultats mathématiques. Ceci a conduit à un essor rapide des aspects théoriques de la géométrie algorithmique, et ce modèle reste le modèle standard dans le domaine. Mais un algorithme correct pour ce modèle de calcul ne l'est plus si on utilise pour représenter les nombres réels leur approximation par des nombres machine. On pourrait penser que la très grande précision fournie par les ordinateurs est suffisante en pratique et qu'il suffit d'ajuster quelques ε. En fait il n'en est rien.

La difficulté vient du fait que les structures géométriques ont une double nature, combinatoire et géométrique. Pour bien comprendre cette question, revenons au calcul de l'enveloppe convexe de points du plan. Le déroulement de l'algorithme est contrôlé par l'évaluation d'une seule opération numérique, qu'on appelle un prédicat, qui décide si un triangle est orienté positivement (tourne à gauche) ou négativement (tourne à droite). L'évaluation d'un prédicat aussi simple que le prédicat d'orientation peut poser des problèmes lorsque les points sont presque alignés et qu'on utilise l'arithmétique flottante des ordinateurs. Plus grave, d'infimes erreurs d'arrondi peuvent se propager et avoir des conséquences fatales, parfois bien plus tard dans le déroulement de l'algorithme. Il est donc indispensable d'évaluer exactement le prédicat d'orientation si on veut construire une enveloppe convexe de manière fiable.

Cette question des liens entre géométrie et arithmétique n'est pas nouvelle et nous

13. J-D. Boissonnat, O. Devillers, R. Schott, M. Teillaud and M. Yvinec. *Application of Random Sampling to On-line Algorithms in Computational Geometry*, Discrete and Computational Geometry, Vol. 8, pp. 51–71 (1992).

sommes un peu dans la même situation que les Grecs qui savaient démontrer par l'absurde que $\sqrt{2}$, la longueur de la diagonale d'un carré de côté unité, n'est pas un nombre rationnel : $\sqrt{2}$ n'était donc pas représentable dans leur système arithmétique (qui ne connaissait que les nombres entiers et les fractions) alors que des approximations étonnament précises étaient connues depuis longtemps comme en témoignent des tablettes babyloniennes. Le programmeur est comme l'esclave Ménon : il ne dispose que des nombres machine représentés dans son ordinateur, typiquement les nombres entiers et les nombres flottants, et il ne peut pas représenter exactement certaines quantités qui interviennent dans les algorithmes géométriques.

Evaluer un prédicat géométrique revient, en général à évaluer le signe $+, -$, ou 0 d'un polynôme. Pour éliminer les problèmes numériques, il suffirait d'évaluer ce polynôme de manière exacte en utilisant du calcul formel. Mais les polynômes en jeu peuvent être de degré élevé (le degré du prédicat d'orientation est d en dimension d) et cette solution serait très coûteuse et ralentirait beaucoup les programmes. Il est alors important de calculer le signe du polynôme sans évaluer le polynôme lui-même. Ceci suffit pour construire la structure combinatoire cherchée et assurer la robustesse des programmes, et on peut tirer parti du fait que le signe du polynôme est facile à évaluer quand on est loin des situations dégénérées, comme trois points alignés dans le cas du prédicat d'orientation dans le plan. Les situations dégénérées requièrent des méthodes plus coûteuses pour être traitées exactement mais sont souvent rares en pratique. On dispose aujourd'hui de méthodes très performantes qui permettent de faire du calcul géométrique rigoureusement fiable sans pénaliser de façon significative les temps de calcul en pratique. Ces avancées ont fait le succès de la bibliothèque d'algorithmes géométriques CGAL [14] qui apporte à ses utilisateurs la sécurité et l'efficacité des algorithmes théoriques.

CGAL

C'est en 1995 qu'est né le projet d'une bibliothèque logicielle implémentant sous une forme cohérente les principaux algorithmes de géométrie algorithmique. L'objectif était double. D'une part, confronter les modèles théoriques à la réalité et les faire évoluer dans une dialectique vertueuse. D'autre part, diffuser un savoir hautement spécialisé sous une forme facilement accessible et utile aux autres domaines scientifiques et à l'industrie.

Le développement de CGAL a été une longue aventure scientifique et humaine qui a engagé plusieurs universités et instituts de recherche européens pendant 20 ans. Les principaux contributeurs sont l'ETH de Zurich, le Max Plank Institute à Saarbrucken, l'université de Tel Aviv et mon équipe à l'INRIA.

Le développement de logiciels de grande envergure comme CGAL est un travail de nature fondamentale et de longue haleine, qui se distingue de la programmation d'applications. Il est mené dans le cadre de la recherche académique, au moins dans une (longue) phase initiale. Il a un rôle structurant pour la discipline avant d'être un produit exportable vers les autres sciences et le monde industriel. La mise au point de CGAL n'a été possible qu'en réunissant des compétences scientifiques

14. *The Computational Geometry Algorithms Library :* www.cgal.org

variées en algorithmique, programmation, arithmétique des ordinateurs, et aussi des compétences touchant à l'organisation du travail de développement (ce à quoi les chercheurs ne sont nécessairement bien préparés).

CGAL comporte aujourd'hui 700 000 lignes de code, une documentation de 3 000 pages et a plus de 1 000 utilisateurs. CGAL est devenu le standard de fait en calcul géométrique. La bibliothèque est distribuée sous une license open source, ce qui a permis sa large diffusion dans le milieu de la recherche et de l'enseignement. Elle est également commercialisée par une startup d'Inria, GeometryFactory, et est intégrée à de nombreux produits commerciaux dans des domaines variés à travers le monde.

3 Maillage de surfaces

Passage continu-discret

Maintenant que les bases de la Géométrie Algorithmique ont été posées, revenons aux questions du début de cette leçon. Comment représenter les objets, organes, monuments, molécules que nous pouvons maintenant numériser ? Comment construire des modèles informatiques de ces formes complexes ? Ces modèles doivent avoir une complexité finie pour au moins deux raisons. Tout d'abord, un algorithme ne peut traiter qu'une quantité finie d'informations. D'autre part, l'objet d'étude n'est souvent connu qu'à travers un nombre fini de mesures.

Comment représenter les formes géométriques continues par des modèles discrets ? Cette question du passage du continu au discret, a suscité de nombreuses recherches picturales et occupe une place centrale en traitement du signal et des images, après les travaux fondateurs de Claude Shannon dans les années 1950. En comparaison, l'étude du lien continu-discret en géométrie est beaucoup plus récente. Les données géométriques étant par nature des ensembles plutôt que des fonctions, les méthodes de l'analyse numérique se révèlent dans une large mesure inadaptées, et une théorie de l'échantillonnage géométrique est requise.

Maillages de surfaces

Commençons par examiner le cas des surfaces qui interviennent dans les multiples applications de la visualisation d'objets tridimensionnels. Mailler une surface consiste à échantillonner la surface et à relier ces points de façon à former une surface triangulée dans l'espace de dimension 3.

On comprend que pour offrir des garanties, l'échantillon devra être suffisamment dense et l'échantillon devra être d'autant plus dense que la forme qu'on veut mailler est complexe. Pour caractériser la complexité d'un objet géométrique, Herbert Federer, le fondateur de la théorie de la mesure géométrique, a introduit une quantité, baptisée la *portée* (reach en anglais), qui résume en une seule valeur la façon dont l'objet est géométriquement plongé dans l'espace. La portée se définit comme le plus petit nombre r tel que tout point à distance au plus r de l'objet a un unique plus

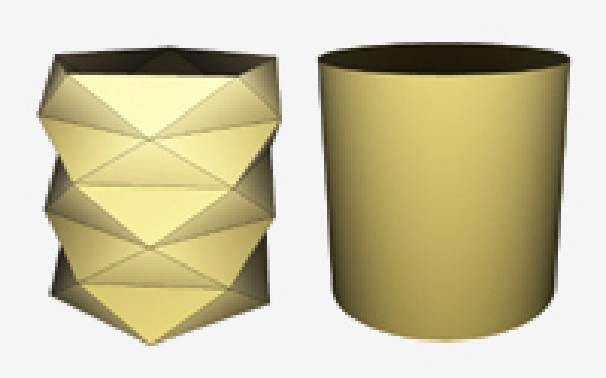


FIGURE 3 – Le lampion de Schwartz.

proche point sur l'objet. Le concept de portée prend en compte la courbure de l'objet (une notion locale) mais aussi son étranglement (une notion globale).

Quelles sont les qualités attendues d'un maillage ? On veut tout d'abord contrôler la distance entre la surface et le maillage. La distance habituelle pour comparer deux ensembles est la distance de Hausdorff : M et M' sont à distance de Hausdorff d si tout de point de M est à distance au plus d d'un point de M' et vice versa. On voudrait également que les facettes du maillage soient proches des plans tangents à la surface. Ceci n'est pas nécessairement vrai même si les facettes sont petites comme l'illustre le cas du lampion de Schwartz (Figure 3). Le lampion de Schwartz est un polyèdre dont les sommets sont placés sur un cylindre et dont les faces sont aussi petites et proches que l'on veut du cylindre, mais disposées en accordéon. Une telle approximation serait catastrophique si on voulait rendre de manière réaliste la façon dont la lumière est réfléchie par le cylindre ou si on voulait estimer son aire en calculant celle du polyèdre. Enfin, il faut que l'approximation respecte la connectivité de la surface, son nombre de composantes, le nombre de trous. On veut que le maillage ait le même *type topologique* que la surface, c'est-à-dire qu'on puisse déformer continûment l'un en l'autre, sans coupure ni déchirure.

Pour répondre à ces exigences, les géomètres algorithmiciens ont adapté la notion de triangulation de Delaunay au cas d'une surface, à travers le concept de *triangulation de Delaunay restreinte*. La figure 4 montre un ensemble de points pris sur une courbe du plan et son diagramme de Voronoï en rouge. On indique en traits pleins les arêtes du diagramme qui coupent la courbe, et on indique en traits pleins bleus leurs arêtes duales dans la triangulation de Delaunay des points. L'ensemble des arêtes en traits pleins bleus constitue la triangulation de Delaunay des points restreinte à la courbe. Comme on le voit, c'est une bonne approximation de la courbe. La notion de triangulation de Delaunay restreinte s'étend au cas des surfaces de l'espace de dimension 3. On obtient dans ce cas un ensemble de triangles, chaque triangle étant dual d'une arête de Voronoï qui traverse la surface.

Une série de travaux initiés par Amenta et Bern [15] ont montré que, si P est un

15. N. Amenta et Marshall Bern. *Surface Reconstruction by Voronoi Filtering.* Discrete Comput. Geom. 22-4 (1999).

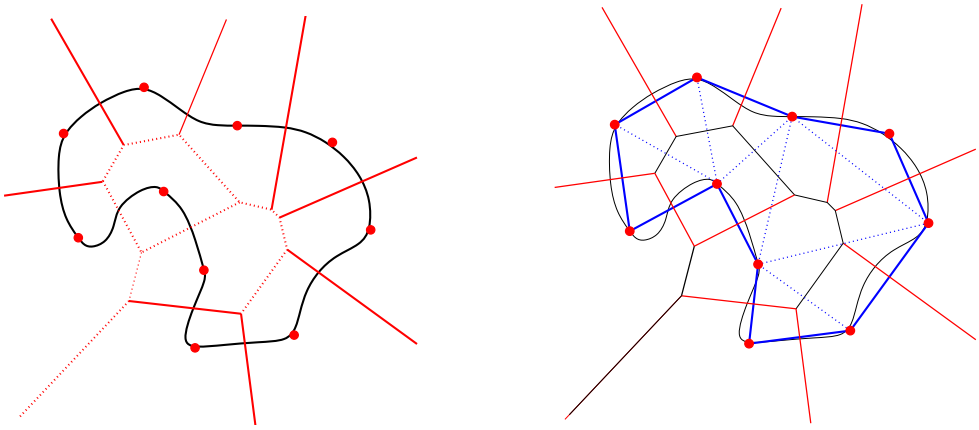


FIGURE 4 – La triangulation de Delaunay restreinte à une courbe.

échantillon suffisamment dense d'une surface, la triangulation restreinte à cette surface est une excellente approximation qui satisfait tous les critères évoqués plus haut.

Avec Steve Oudot [16], nous avons prolongé ce résultat et proposé un algorithme qui échantillonne et maille des surfaces. Cet algorithme utilise un oracle qui sonde la surface en lançant des rayons le long de certaines arêtes de Voronoï, un peu comme un aveugle découvre à tâtons son environnement, et raffine progressivement l'approximation jusqu'à satisfaire les conditions du théorème. Cet algorithme a été utilisé avec succès dans de nombreuses applications. Il peut, par exemple, être utilisé pour mailler des surfaces de niveau dans des images médicales tridimensionnelles ou des horizons sismiques dans des images géologiques, ou pour reconstruire des surfaces à partir de nuages de points (Figure 5).

Les maillages bidimensionnels suffisent pour visualiser des surfaces mais, pour planifier des opérations chirurgicales robotisées ou simuler des phénomènes physiques, il est nécessaire de mailler non seulement les surfaces bords des objets mais également les volumes qu'elles délimitent. Il faut alors utiliser des maillages tétraédriques. La méthodologie utilisée pour les surfaces reste valide et le raffinement de la triangulation de Delaunay utilisé pour les surfaces peut être étendu au cas de volumes tridimensionnels limités par des surfaces. Bien que la génération de maillages de domaines tridimensionnels complexes soit beaucoup plus délicate que celle de maillages plans ou surfaciques, les progrès réalisés au cours des dix dernières années ouvrent la voie à de nouvelles applications. Dans le domaine médical, des maillages construits à partir d'images 3d, permettent de visualiser les organes, simuler des phénomènes physiologiques ou planifier des interventions chirurgicales. Un autre exemple d'application est la reconstruction de modèles 3d à partir de vues du sol et de vues aériennes. On peut reconstruire ainsi des bâtiments, des villes ou des usines et créer des modèles 3d de l'environnement dans lequel doivent s'intégrer de nouvelles infrastructures comme des routes ou des voies ferrées.

16. J-D. Boissonnat, S. Oudot. *Provably good sampling and meshing of surfaces.* Graphical Models, 67 (2005).

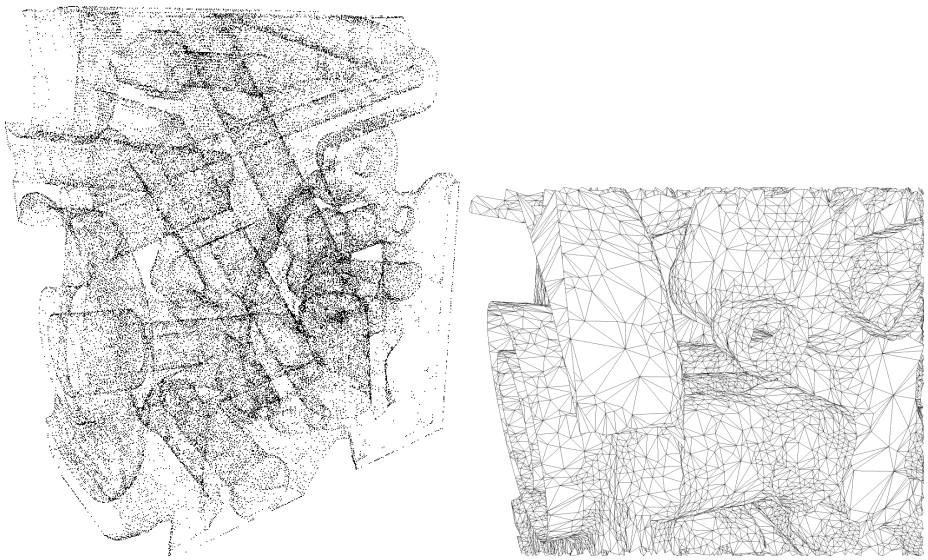


FIGURE 5 – Reconstruction d'une pièce mécanique.

Au delà de la dimension 3

On n'a considéré jusqu'ici que les surfaces de l'espace de dimension 3 mais certains problèmes se posent en dimensions supérieures à 3, et il est utile de trianguler des surfaces plongées dans des espaces de plus grandes dimensions. Celles-ci se rencontrent naturellement quand on s'intéresse aux systèmes dynamiques et à la modélisation du mouvement de systèmes articulés. Un exemple est donné par l'étude des paysages énergétiques des molécules. Une molécule est un ensemble d'atomes animés de mouvements les uns par rapport aux autres. Les différentes positions prises par la molécule lorsque ces déformations se produisent s'appellent les conformations de la molécule. On peut produire un grand nombre de conformations d'une molécule en simulant la dynamique moléculaire. Si la molécule est formée de n atomes, une conformation sera caractérisée par la donnée de $3n$ coordonnées qui repèrent la position dans l'espace de chaque atome, et on peut représenter chaque conformation comme un point dans un espace à $3n$ dimensions. L'ensemble des conformations obtenues par dynamique moléculaire sera un nuage de points dans ce grand espace. Récemment, des mathématiciens et des chimistes ont pu décrire précisément l'espace des conformations du cyclo-octane C_8H_{16}, un alcane cyclique utilisé dans la fabrication de plastiques et étudié depuis plus de quarante ans [17]. En analysant un ensemble d'un million de points de l'espace de dimension 72, chacun décrivant une conformation du cyclo-octane, les auteurs ont établi que l'espace des conformations du cyclo-octane est une surface de dimension 2 composée d'une sphère et d'une bouteille de Klein (une surface non orientable qu'on ne peut pas plonger en dimension 3). Au delà de son intérêt intrinsèque, cette découverte ouvre de nouvelles pistes pour comprendre le paysage énergétique du cyclo-octane. Etendre ce type d'analyse aux plus grandes molécules, en particulier aux protéines, aurait des implications

17. S. Martin, A. Thompson, E. A. Coutsias, J-P. Watson. *Topology of cyclo-octane energy landscape.* The Journal of Chemical Physics 132 (234115), 2010.

considérables.

Lorsque l'espace ambiant est de grande dimension, on doit faire face à ce qu'on appelle, à la suite du mathématicien Richard Bellman, le fléau de la dimension : l'explosion combinatoire du coût de résolution d'un problème lorsque la dimension augmente. En géométrie algorithmique, ce phénomène se traduit notamment par le fait que les structures géométriques utilisées en petites dimensions, même les plus simples comme les polyèdres, les triangulations ou les grilles, voient leur taille grossir de façon exponentielle avec la dimension de l'espace. Si on choisit une résolution d'un millième, une image 2d aura 1 million d'éléments (pixels), une image 3d en aura 1 milliard, et une image en dimension 6 en aura 10^{18}. On ne peut donc pas trianguler un espace de grande dimension comme l'espace de dimension 72 dans lequel sont données les conformations du cyclo-octane. On peut en revanche contourner la difficulté lorsque la structure qu'on cherche à trianguler est une *variété* de petite dimension comme c'est le cas du cyclo-octane dont l'espace des conformations est une surface de dimension 2.

La notion de *variété topologique* généralise celle de courbe et de surface qui sont les variétés de dimensions 1 et 2. D'une manière générale, une variété de dimension m est un objet non linéaire tel que le voisinage de tout point peut être déformé en un morceau de plan de dimension m topologiquement équivalent. La dimension m de la variété ne doit pas être confondue avec la dimension de l'espace ambiant dans laquelle cette variété est plongée. Comme on l'a vu, les conformations du cyclo-octane forment une surface de dimension 2 qui est plongée dans un espace de dimension 72. Lorsque la dimension de la variété est petite, il est utile de développer des algorithmes dont la complexité dépend essentiellement de la *dimension intrinsèque* m et pas de la dimension de l'espace ambiant. On peut ainsi contourner le fléau de la dimension.

La question de la *triangulation des variétés* a traversé tout le XXème siècle mathématique. Beaucoup d'illustres mathématiciens s'y sont penchés et en particulier Hassler Whitney, qui a démontré en 1957 que toutes les variétés différentiables sont triangulables. Il est intéressant de noter que sa démonstration est constructive bien que d'une complexité algorithmique inaccessible.

Avec mes collaborateurs R. Dyer et A. Ghosh [18], nous avons élaboré les éléments d'une théorie de la triangulation de Delaunay des variétés. Contrairement au cas des surfaces, il ne suffit pas de raffiner l'échantillon pour assurer l'existence d'une triangulation de Delaunay sur une variété de dimension supérieure à 2. En revanche, sous des conditions plus fortes, on peut prouver l'existence de telles triangulations et les construire efficacement. Pour cela, on calcule des triangulations de Delaunay dans des cartes locales autour de chaque point de donnée, ce qu'on peut faire quasiment aussi rapidement que si on calculait une triangulation dans un espace linéaire de la même dimension que celle de la variété. Il faut ensuite recoller ces triangulations locales en une triangulation globale. Le fait que la triangulation de Delaunay soit canonique va le permettre si l'échantillon est assez dense. En fait, il faut assurer que les triangulations locales restent stables quand on passe d'un point à un point

18. J-D. Boissonnat, R. Dyer, A. Ghosh. *Delaunay triangulation of manifolds.* J. of Foundations of Computational Mathematics (2017).

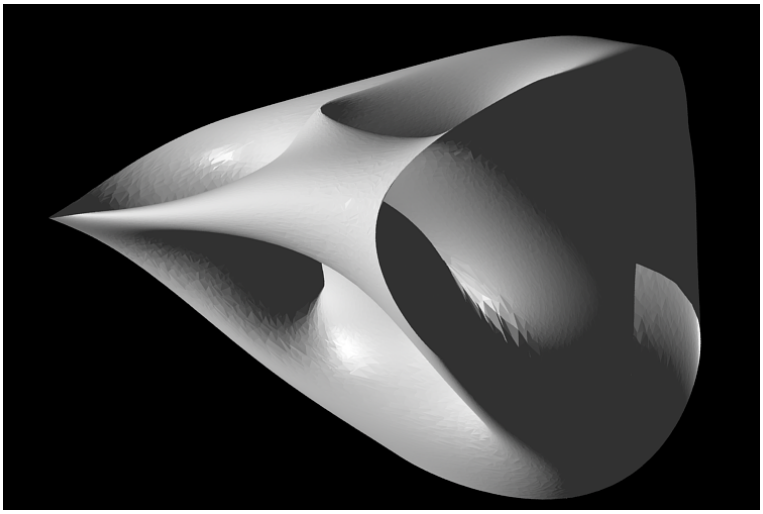


FIGURE 6 – Triangulation d'une surface de Riemann.

voisin sur la variété. Il est intéressant d'observer que la question de la stabilité de la triangulation de Delaunay qui est, comme on l'a vu, critique pour la robustesse des algorithmes, l'est également pour contourner le fléau de la dimension dans les problèmes de reconstruction des variétés. On voit sur la figure 6, le maillage d'une surface de Riemann plongée en dimension 8. Trianguler cette surface a permis d'en calculer la caractéristique d'Euler et de conforter une conjecture mathématique.

4 Analyse géométrique et topologique des données

Si les données géométriques ont révolutionné notre perception et notre interaction avec le monde tridimensionnel, d'une manière plus générale, les données, géométriques ou non, ont pris une place essentielle dans la science moderne et, au delà, dans la société toute entière [19]. De façon peut-être un peu inattendue, la géométrie algorithmique a un rôle à jouer dans ce contexte. Développer une approche géométrique et topologique de l'analyse des données, en construire les fondements mathématiques et algorithmiques, mettre au point et diffuser des programmes informatiques rapides et fiables constituent quelques uns des nouveaux objectifs de la géométrie algorithmique.

Géométrie des données

Les données géométriques fournissent les coordonnées de points dans l'espace. Plus généralement, une donnée est un ensemble de valeurs associées à certains paramètres comme la pression, la température ou le temps, et peut être représentée par un point dans un espace de paramètres qu'on appellera *l'espace des observations*. Cet espace

19. S. Abiteboul. *Sciences des données : de la logique du premier ordre à la toile.* Collège de France/Fayard (2012).

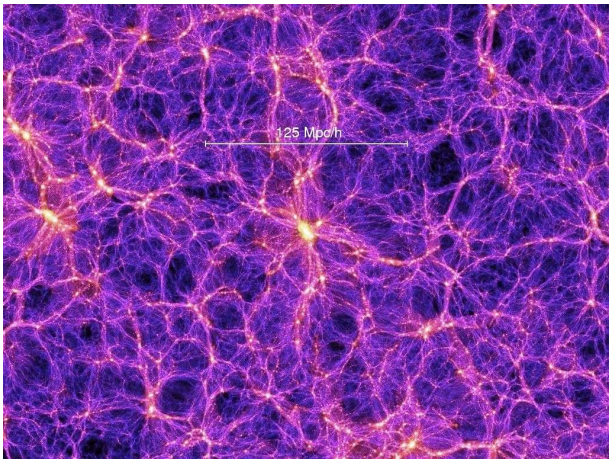


FIGURE 7 – La toile cosmique.

des observations ne ressemble pas nécessairement à notre espace euclidien mais, si on sait définir une notion de proximité ou de ressemblance entre les données, on va pouvoir aborder l'analyse des données d'un point de vue géométrique.

Le nombre de paramètres qui caractérisent une donnée est souvent grand et l'espace des observations est alors de grande dimension. Mais cette grande dimension n'est qu'apparente car les données sont en général corrélées et ne sont pas uniformément distribuées dans l'espace des observations. Comprendre la forme des données est alors utile pour mieux comprendre le système physique sous-jacent, comme on l'a vu dans le cas du cyclo-octane.

Un autre exemple est celui des structures à grandes échelles de l'univers (figure 7). Plusieurs grands relevés de galaxies, comme Le Redshift Survey, ont montré que la matière n'est pas distribuée de façon uniforme dans l'univers : les galaxies semblent plutôt distribuées selon des filaments et des feuillets entourant des zones gigantesques quasiment vides de galaxies. L'univers apparait aujourd'hui comme une structure évoquant celle d'une gigantesque éponge ou d'une toile d'araignée. Au delà de l'observation toujours plus fine de l'univers, un des buts de la cosmologie moderne est de comprendre ces structures à grandes échelles et de pouvoir les reproduire par simulation numérique. Il s'agit d'un problème qui ressemble au problème de reconstruction discuté plus haut mais la démarche n'est pas la même : il ne s'agit pas de reconstruire un objet dont l'existence est avérée, comme un organe ou une pièce mécanique, mais d'inférer une structure plausible à partir de données. C'est moins la précision de l'approximation qui nous intéressera que la recherche de nouvelles formes de symétrie et d'informations qualitatives robustes.

Homologie persistante

Rassembler les données en clusters, c'est-à-dire en groupes de données voisines, est une première question importante pour comprendre l'organisation spatiale des données. Quand les clusters ont une forme complexe, une méthode couramment uti-

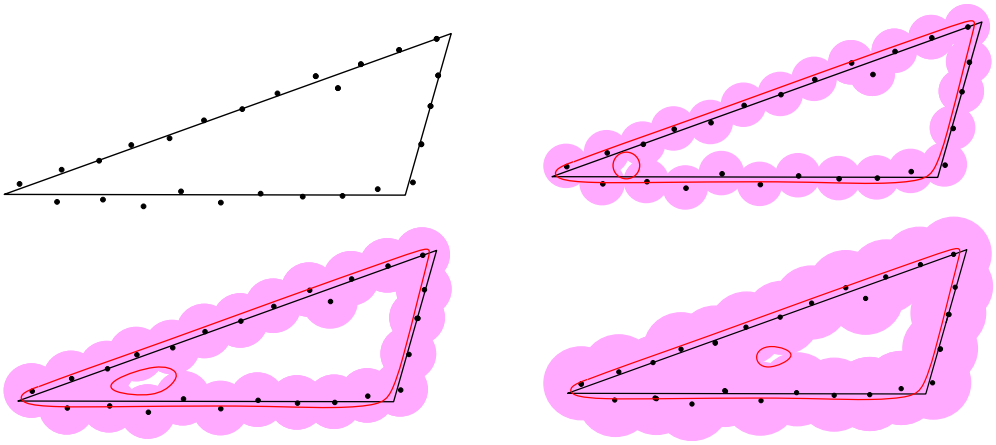


FIGURE 8 – Evolution de l'union des boules centrées aux points de données.

lisée consiste à estimer localement la densité des données et rechercher des pics de cette densité. En pratique, l'estimation de la densité est bruitée, et il s'agit d'en détecter les modes significatifs et de les distinguer de ceux qui sont peu importants ou dûs au bruit. Cette question n'est pas sans rappeler la notion de hauteur de culminance des alpinistes qui est la différence d'altitude entre un sommet et le col le plus élevé permettant d'atteindre une cime encore plus haute. On peut ainsi distinguer l'aiguille du midi qui est un sommet majeur des sommets voisins qui ont presque la même hauteur. Une question de même nature se pose en traitement du signal et en analyse d'images. Elle se pose également en modélisation géométrique : comment distinguer les détails (qui peuvent n'être que du bruit) des structures importantes ?

L'*homologie persistante* apporte des réponses à cette question. Cette théorie prend ses racines dans la théorie de Morse, « cet énigmatique monolithe qui domine un bonne part du paysage mathématique contemporain » comme disait R. Thom. Remplaçons les points de données par des boules dont le rayon va jouer le rôle d'un facteur d'échelle, et faisons croître la valeur de ce rayon. L'union des boules va grossir, des composantes peuvent fusionner, des trous se créer puis disparaître comme on le voit sur la figure 8. En dimensions supérieures, peuvent aussi apparaître et disparaître des caractéristiques topologiques de plus grandes dimensions.

Ces caractéristiques topologiques peuvent être calculées grâce à l'*homologie*. On peut voir la formule d'Euler pour les polyèdres comme un premier résultat de la théorie de l'homologie mais c'est H. Poincaré qui en a vraiment bâti les fondations. L'homologie peut être décrite par des nombres qu'on appelle les nombres de Betti. Le premier nombre de Betti, β_0, compte le nombre de composantes, β_1 le nombre de cycles indépendants, β_2 le nombre de cavités. Les nombres de Betti suivants comptent les « trous » de plus grandes dimensions. En analyse de données, β_0 pourra compter les clusters, β_1 les cycles qui apparaissent notamment quand des phénomènes périodiques sont en jeu, β_2 les cavités qui sont en particulier importantes dans les structures poreuses.

L'homologie est une notion plus grossière que l'équivalence topologique mais elle conduit à des algorithmes efficaces d'algèbre linéaire quand on représente les espaces

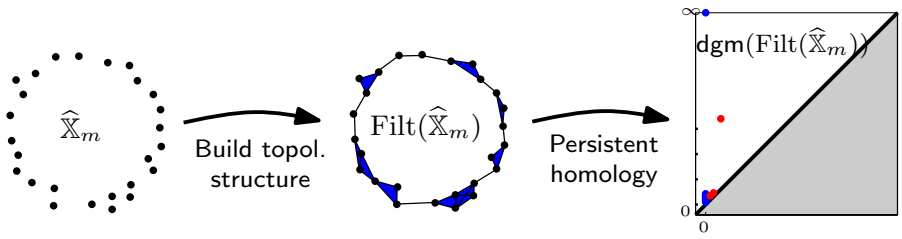


FIGURE 9 – La chaîne de traitement de l'homologie persistante.

par des *complexes simpliciaux*. La notion de complexe simplicial généralise celle de graphe et de surface triangulée. Un complexe simplicial est défini comme un assemblage de simplexes (arêtes, triangles, tétraèdres, etc.) qui se recollent le long de faces communes de dimensions plus petites.

Un complexe simplicial utile est le complexe de Čech qui se définit de la manière suivante. Si deux boules ont une intersection commune, on crée une arête joignant leurs centres, si trois boules ont une intersection commune on crée un triangle, et, d'une manière générale, si k boules ont une intersection commune, on crée le simplexe de dimension $k-1$ dont les sommets sont les k centres associés. Il faut considérer ce complexe comme une structure combinatoire abstraite qui n'est pas nécessairement plongée dans l'espace contenant les points.

Le théorème du nerf de Jean Leray, éminent mathématicien qui a enseigné ici même au Collège de France entre 1947 et 1978 va établir un pont entre la topologie des objets géométriques et les complexes simpliciaux [20]. Il découle du théorème du nerf que le complexe de Čech a la même *homologie* que l'union des boules.

Nous avons les principaux ingrédients pour construire l'algorithmique de l'homologie persistante. C'est à partir des années 2000, sous l'impulsion d'Herbert Edelsbrunner et de Gunnar Carlsson, que l'homologie persistante s'est vraiment développée et que les questions algorithmiques ont été abordées. L'homologie persistante est aujourd'hui un outil central en analyse topologique des données, et l'objet de recherches très actives de nature mathématique et algorithmique [21].

La chaîne de traitement de l'homologie persistante est la suivante. Partant d'un nuage de points, on construit une suite de complexes simpliciaux emboîtés qu'on appelle une *filtration*. On peut prendre par exemple le complexe de Čech que nous venons de définir, et considérer la suite de complexes emboîtés obtenus quand on fait croître le rayon des boules. L'algorithme de persistance parcourt la filtration, repère les changements topologiques du complexe et construit un *diagramme de persistance* qui associe à une caractéristique topologique (composante connexe, cycle, cavité et leurs analogues en plus grandes dimensions) un point du plan ayant pour abscisse la valeur de la filtration à laquelle elle est apparue et en ordonnée la valeur de la filtration à laquelle elle a disparu. Les points qui sont loin de la diagonale représentent

20. J. Leray. *Sur la forme des espaces topologiques et sur Ies points fixes des représentations*. J. Math. Pure Appl. 24 (1945).

21. S. Oudot. *Persistence Theory : from Quiver Representations to Data Analysis*. American Mathematical Society (2015).

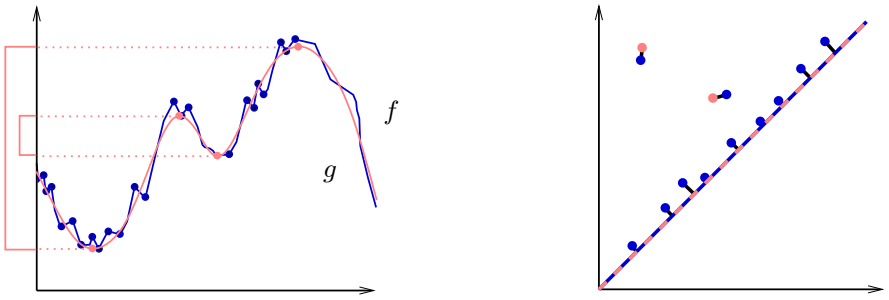


FIGURE 10 – Stabilité des diagrammes de persistance.

des éléments topologiques qui perdurent au cours de la filtration. La figure 9 montre l'exemple d'un ensemble de points échantillonnant approximativement un cercle. Sur le diagramme à droite de la figure, les points bleus représentent les composantes. Au départ il y a autant de composantes que de points. Les composantes fusionnent rapidement quand le rayon grandit et on obtient une seule composante qui ne disparaît jamais. Les cycles sont représentés par les points rouges. De nombreux cycles sont créés au début. Ils disparaissent rapidement puis apparait un cycle qui a une longue durée de vie : il correspond au cercle échantillonné. Ce cycle finit par disparaître quand le rayon des boules est suffisamment grand.

On peut se demander quel crédit accorder à une représentation aussi succinte d'un jeu de données. Une partie de la réponse vient d'un résultat de stabilité. Considérons les deux fonctions de la figure 10. Comme attendu, la fonction rouge qui est lisse a deux modes persistants indiqués par deux points rouges sur le diagramme de persistance. La fonction bleue bruitée a de nombreux maxima et minima, ce qui conduit à de nombreux points dans le diagramme de persistance, mais la plupart sont très proches de la diagonale. Pour une certaine distance appelée la distance d'étranglement (la longueur du plus grand segment bleu sur la figure de droite), les deux diagrammes de persistance (en incluant la diagonale) sont proches. Cette observation est en fait un théorème important de D. Cohen-Steiner, H. Edelsbrunner et J. Harer [22] qui ont montré que des fonctions proches ont des diagrammes de persistance qui sont proches pour la distance d'étranglement. Les composantes persistantes correspondent donc à des modes qui restent stables si on perturbe la fonction. Ce sont des modes significatifs qui se distinguent du bruit.

L'analyse topologique que fournit l'homologie persistante est intéressante à plus d'un titre. Elle est peu sensible au choix de la métrique, indépendante du choix des coordonnées, et elle fournit des informations intrinsèques globales sur les données sous la forme de diagrammes de persistance robustes au bruit. L'homologie persistante offre de nouveaux outils pour l'analyse de données. Une application est la détection de clusters [23] illustrée sur la figure 11 : le diagramme de persistance présente deux

22. H. Edelsbrunner, D. Letscher, A. Zomorodian. Topological Persistence and Simplification. Discrete Comput. Geom. 28 (2002).

23. F. Chazal, L. J. Guibas, S. Y. Oudot, P. Skraba. *Persistence-Based Clustering in Riemannian Manifolds*. Journal of the ACM, 60 (2011).

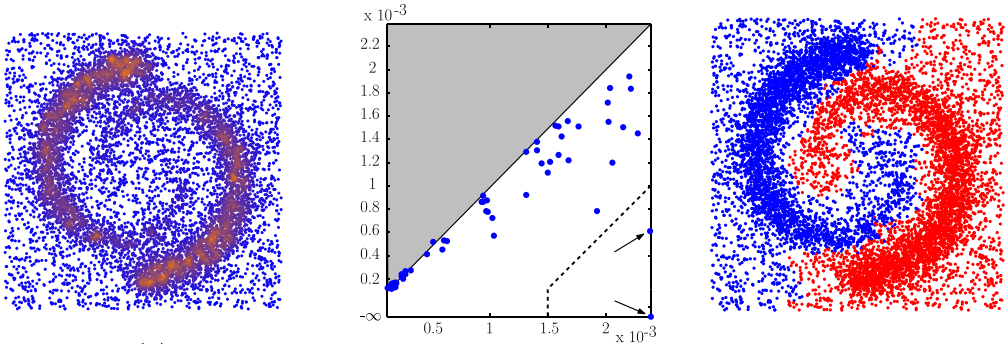


FIGURE 11 – Détection de clusters.

points éloignés de la diagonale, ce qui permet de classer les points en deux clusters, bleu et rouge.

On a jusque là considéré le diagramme de persistance d'un seul nuage de points dont on voulait analyser la structure géométrique. Une question tout aussi importante est de comparer des jeux de données pour pouvoir les comparer ou les classer. On peut, pour cela, considérer les diagrammes de persistance comme des signatures des nuages de points et utiliser la distance d'étranglement. On construit ainsi une géométrie dans l'espace des diagrammes de persistance. C'est un nouveau champ de recherche passionnant mêlant topologie algorithmique et statistiques.

La mise en œuvre de cette théorie dans des applications réelles demande une algorithmique spécifique pour l'ensemble de la chaîne de l'homologie persistante. C'est ce que nous développons dans le cadre du projet Gudhi financé par le Conseil Européen de la Recherche [24]. Le calcul du complexe de Čech est impossible en pratique et il faut concevoir d'autres types de complexes simpliciaux qui aient de bonnes propriétés topologiques et soient facilement calculables. Il faut aussi savoir coder de façon optimale ces complexes simpliciaux et leurs filtrations de façon qu'ils occupent le moins de place mémoire possible mais puissent néanmoins être utilisés comme structures de données efficaces pour le calcul de la persistance. Les algorithmes de persistance eux-mêmes sont l'objet de recherche très actives et on ne connaît pas encore précisément la complexité algorithmique du calcul de la persistance. D'autres questions essentielles portent sur le développement d'algorithmes approchés. Enfin, pour que ces nouvelles techniques soient réellement efficaces et accessibles aux autres domaines scientifiques, il faut également développer des outils logiciels robustes, efficaces et faciles d'utilisation.

5 Conclusion

J'ai essayé de vous montrer comment la démarche algorithmique s'est imposée en géométrie sous la pression combinée des besoins en modélisation tridimensionnelle et de la recherche en informatique. Le traitement des formes géométriques tridimen-

24. https ://project.inria.fr/gudhi/

sionnelles est devenu une nouvelle forme du traitement de l'information qui trouve des retombées visibles dans tous les domaines. Ceci a été rendu possible non seulement par les développements technologiques spectaculaires des capteurs, des moyens de calcul et des systèmes de visualisation, mais aussi par l'invention de nouvelles représentations des formes et de nouveaux paradigmes de calcul.

On a vu que la modélisation géométrique ne se limite pas aux formes tridimensionnelles. En voulant modéliser le mouvement, on est amené à modéliser des espaces de configurations de dimensions plus grandes que 3. On a vu également que la géométrie algorithmique a des modèles et des programmes à proposer pour explorer le monde vertigineux des données. C'est un sujet en plein développement que je n'ai qu'effleuré mais qui sera développé dans le cadre du cours et des colloques qui suivront, et qui constitue un champ de recherches pour les années à venir.

Je voudrais, pour terminer, remercier mes collaborateurs, collègues et étudiants, qui ont contribué aux résultats que j'ai évoqués dans cette leçon et que je ne peux pas tous citer ici. Je voudrais également remercier ceux qui m'ont encouragé et aidé avant que le mot de géométrie algorithmique n'existe. Claude Puech qui dirigeait à l'époque le département d'informatique de l'Ecole Normale Supérieure, Philippe Flajolet, un maître de l'informatique mathématique, disparu trop tôt mais qui reste un modèle scientifique et humain, Olivier Faugeras qui m'a confié un jour la thèse de M. Shamos et invité à pénétrer dans le nouveau monde de la géométrie algorithmique.

Je vous remercie pour votre attention et laisse le mot de la fin à Henri Poincaré :

> *Une géométrie ne peut être plus vraie qu'une autre, elle peut simplement être plus commode.*

Références

[1] F. Aurenhammer, R. Klein, D-T. Lee. *Voronoi diagrams and Delaunay triangulations.* World Scientific (2013).

[2] J-D. Boissonnat, F. Chazal, M. Yvinec. *Geometric and Topological Inference.* Cambridge University Press (2017).

[3] J-D. Boissonnat, M. Teillaud ed. *Effective Computational Geometry for Curves and Surfaces.* Springer (2007).

[4] J-D. Boissonnat, M. Yvinec. *Algorithmic Geometry.* Cambridge University Press (1998).

[5] S-W. Cheng, T. K. Dey, J. R. Shewchuk. *Delaunay Mesh Generation.* CRC Press (2013).

[6] T. K. Dey. *Curve and Surface Reconstruction : Algorithms with Mathematical Analysis.* Cambridge University Press (2007).

[7] H. Edelsbrunner, J. L. Harer. *Computational topology : an introduction.* American Mathematical Soc. (2010).

[8] R. Ghrist. *Elementary applied topology.* (2014).

[9] S. Y. Oudot. *Persistence Theory : From Quiver Representations to Data Analysis*. AMS Mathematical Surveys and Monographs, Vol. 209 (2015).