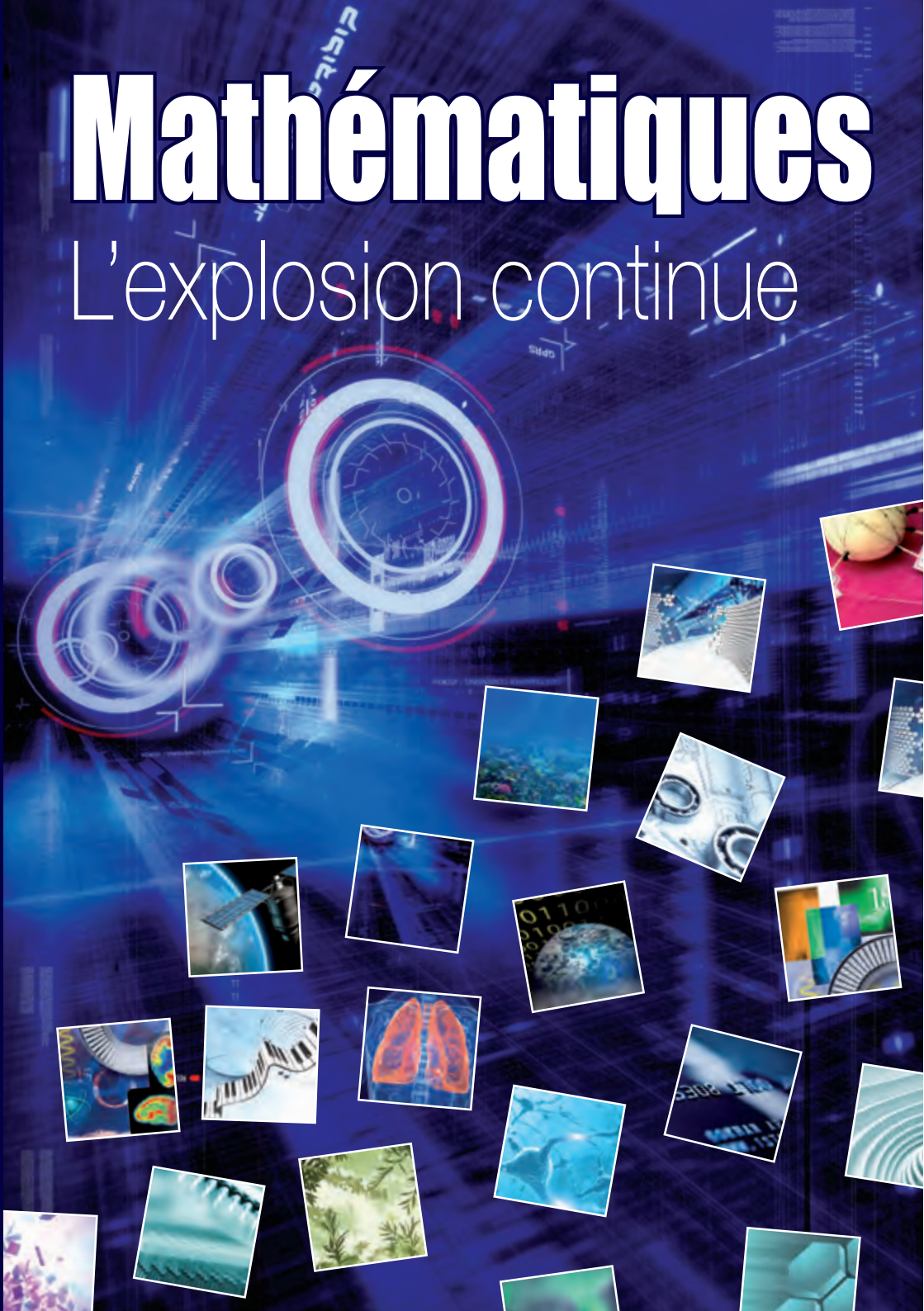
Mathématiques

L'explosion continue



Mathématiques, l'explosion continue

Photos : *© thinkstock (2013).*
© Collections École polytechnique (page 162).
© O. Boulanger, J. Cotera, FSMP.

Création Centre Poly-*media*

Achevé d'imprimer en septembre 2013. Dépôt légal : 3e trimestre 2013
ISBN 978 - 2 - 85629 - 375 - 1. *Imprimé en France*

La brochure « Mathématiques, l'explosion continue »,
conçue par la Fondation Sciences Mathématiques de Paris (FSMP),
la Société Française de Statistiques (SFDS),
la Société de Mathématiques Appliquées et Industrielles (SMAI)
et la Société Mathématique de France (SMF),
a été réalisée grâce au soutien financier de Cap'Maths.

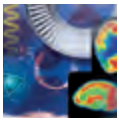
FSMP Institut Henri Poincaré 11 rue Pierre et Marie Curie 75231 Paris CEDEX 05, France
Tel. 01 44 27 68 03 http://www.sciencesmaths-paris.fr

SFDS Institut Henri Poincaré 11 rue Pierre et Marie Curie 75231 Paris CEDEX 05, France
Tél. 01 44 27 66 60 http://www.sfds.asso.fr

SMAI Institut Henri Poincaré 11 rue Pierre et Marie Curie 75231 Paris CEDEX 05, France
Tél. 01 44 27 66 61 http://www.smai.emath.fr

SMF Institut Henri Poincaré 11 rue Pierre et Marie Curie 75231 Paris CEDEX 05, France
Tél. 01 44 27 67 96 http://smf.emath.fr

Sommaire



Avant-propos

Maria J. Esteban, *présidente de la Société de Mathématiques Appliquées et Industrielles 2009-2012*
Bernard Helffer, *président de la Société Mathématique de France 2010-2012*
Jean-Michel Poggi, *président de la Société Française de Statistique 2011-2013*

Depuis une dizaine d'années, de nombreuses initiatives ont vu le jour en France pour mieux appréhender le rôle des mathématiques dans notre société. Les mathématiciens sont ainsi devenus plus conscients qu'ils se devaient de mieux faire connaître les spécificités et l'utilité de leur discipline. Une des premières initiatives fut la publication de l'Explosion des Mathématiques *en 2002. Ce recueil a été largement diffusé et traduit en plusieurs langues. Il est maintenant épuisé même s'il reste accessible sur le web. Les quelque dix ans qui nous séparent de cette première édition ont vu une évolution très rapide de toutes les branches des mathématiques et leur développement croissant dans tous les domaines de la société: l'explosion continue! Le congrès Maths à Venir, dont le but était de témoigner de cette place grandissante des mathématiques, a été en 2009 un véritable succès, auquel se sont associés de nombreux industriels. Les fonds recueillis à cette occasion nous permettent une réédition. C'est pourquoi nous avons décidé de publier une nouvelle brochure reprenant trois articles mis à jour de l'ancienne brochure et vingt-et-un nouveaux textes tenant compte des évolutions récentes et de la diversité croissante des interactions et applications des mathématiques.*

Avant de laisser le lecteur découvrir ces textes, replaçons ce volume dans son contexte. Nous vivons en effet aujourd'hui encore une situation paradoxale. Les mathématiques sont un instrument irremplaçable de formation à la rigueur et au raisonnement; elles développent l'intuition, l'imagination, l'esprit critique; elles sont aussi un lan-

gage universel, et un élément fort de la culture. Elles jouent en outre, par leurs interactions avec les autres sciences et par leur capacité à décrire et expliquer des phénomènes complexes mis en évidence dans la nature et dans le monde technologique, un rôle grandissant dans notre vie quotidienne. Cet état de fait est bien souvent ignoré ou du moins minoré par la majorité de nos concitoyens, pour qui les mathématiques sont une discipline abstraite, qui n'évolue plus, figée dans une perspective de formation et qui a peu à voir avec le monde réel.

On peut trouver bien sûr à ce paradoxe des explications qui tiennent à la spécificité des mathématiques. C'est une discipline qui se nourrit de ses liens avec les autres sciences, avec la société et le monde industriel, mais qui également s'enrichit elle-même : les nouvelles théories ne détruisent pas les précédentes mais les utilisent, les améliorent ou les dépassent. Réciproquement, même si bon nombre de chercheurs en mathématiques sont intéressés avant tout par le côté intellectuel voire même esthétique de leur discipline, nombreux sont ceux qui s'investissent dans des directions de recherche liées à des applications réelles de celle-ci. Ainsi, les applications enrichissent la recherche en mathématiques, mais ne peuvent seules la piloter. Cet équilibre subtil entre les facteurs de développement interne et externe doit absolument être préservé. Vouloir définir ou mesurer l'activité ou la recherche en mathématiques par ses applications existantes ou potentielles reviendrait à les faire disparaître. À l'opposé, privilégier l'axiomatisation, l'étude des structures et la dynamique interne de la discipline comme l'ont fait, certes avec de beaux succès, les mathématiques françaises dans la période 1940-1970 a conduit à retarder le développement en France des mathématiques dites appliquées, contrairement à ce qui se passait au même moment aux États-Unis et en Union Soviétique. Les facteurs de progrès sont très souvent aux frontières de la discipline. Aujourd'hui, et nous nous en réjouissons, les mathématiques ont rétabli, et parfois créé, des liens forts avec de nombreux secteurs économiques et avec les autres sciences. La frontière entre mathématiques pures et mathématiques appliquées est devenue perméable : les mathématiques les plus fondamentales servent à résoudre des problèmes concrets de plus en plus difficiles alors que de nouveaux problèmes théoriques sont posés par des questions appliquées.

Le but du présent document est de faire découvrir tous les attraits et atouts du monde des mathématiques et en particulier la remarquable efficacité des mathématiques dans la résolution de problèmes sociétaux et technologiques ; de montrer les mathématiques sous leurs aspects les plus divers – scientifiques, techniques, culturels, sociologiques ; de souligner la diversité et l'universalité

d'une discipline qui entretient des liens aussi bien avec la physique, la chimie, la mécanique, l'informatique, l'économie et la biologie qu'avec les sciences humaines ou sociales, en passant par toutes les facettes de la technologie. Les mathématiques sont partout. Sans elles, pas d'ordinateurs, pas de systèmes d'information, pas de téléphonie mobile ; pas d'ateliers de conception pour les constructeurs automobiles et aéronautiques ; pas de systèmes de localisation par satellite, de traitement du signal, de décryptage du génome, de prévisions météo, de cryptographie, de cartes à puce, de robots, de moteurs de recherche ou de traitements de grandes masses de données…

Au-delà de leur rôle de science académique et de formation de base à l'école, les mathématiques sont omniprésentes dans la société d'aujourd'hui. Elles suivent, accompagnent et quelquefois précèdent les développements scientifiques et technologiques actuels, qui font aussi bien appel aux résultats de la recherche fondamentale contemporaine la plus récente qu'aux découvertes accumulées dans le passé. Enfin, les besoins en mathématiques croissent avec l'accélération des mutations et créations technologiques. On ne peut s'en passer, alors que l'on est confronté à la nécessité d'élaborer, de maîtriser, d'analyser ou d'optimiser des systèmes de complexité croissante. Les États-Unis l'ont bien compris, puisque la NSF (National Science Foundation, l'organisme fédéral nord-américain chargé de distribuer les crédits pour la recherche universitaire) a augmenté considérablement son soutien financier aux mathématiques à partir des années 2000. Plus récemment des pays émergents comme la Chine ont investi massivement dans ce domaine.

C'est dans ce contexte qu'il nous a semblé précieux de compléter et mettre à jour cet outil de diffusion et de popularisation des mathématiques. Nous souhaitons au lecteur de belles et stimulantes découvertes.

Couper, attendrir, trancher, réduire :

un conte culinaire sur la résolution informatique des problèmes difficiles

Nicolas Schabanel, *directeur de recherches CNRS à l'Université Paris Diderot*
Pierre Pansu, *professeur à l'Université Paris-Sud*

Comment répartir ses invités sur deux tablées sans placer ensemble ceux qui se détestent ? Résoudre ce casse-tête quel que soit le nombre d'invités est un problème extrêmement difficile à resoudre quand le nombre d'invités augmente. À l'aide d'un peu de programmation géométrique, on peut trouver une réponse approchée de garantie présumée optimale.

Ce soir, je convie mes amis à dîner. Cependant, je ne dispose pas d'une table assez grande et je vais devoir les répartir sur deux tables. Quels groupes former ?

Fermement décidé à résoudre ce problème de façon rationnelle, j'attrape un crayon et mon bloc-note et appelle Maelys, toujours la première informée des derniers potins, pour savoir qui s'est fâché avec qui dernièrement. Tout en tenant le téléphone d'une main, je me retrouve à représenter chaque invité par un petit rond (on dit un *sommet* en théorie des graphes) et à tracer un trait (une *arête*, c'est plus chic) entre chaque paire d'invités qu'il vaudrait mieux séparer en ce moment. Je me rends rapidement compte que la situation est assez compliquée : il y a beaucoup plus d'animosités entre eux que je ne le croyais ; par exemple, aucun groupe ne semble se détacher clairement sur mon dessin (on dit un *graphe*).

Figure 1. Le graphe des inimitiés entre mes amis

Je vais donc essayer de trouver la répartition de mes amis en deux groupes qui minimise le nombre d'inimitiés à l'intérieur de chaque table, ce qui revient à maximiser le nombre d'arêtes allant d'une table à l'autre. Mathématiquement, je cherche à *partitionner* les sommets de mon graphe en deux parties de sorte à maximiser le nombre d'arêtes cassées par cette partition, c'est-à-dire à cheval entre les deux tables. (Pour ne pas compliquer inutilement, nous ne chercherons pas à obtenir deux parties de même taille. Les techniques de résolutions avec ou sans cette contrainte ne diffèrent d'ailleurs pas significativement.)

Quand est-ce qu'on mange ?

Si j'étais physicien, j'aurais certainement monté le dispositif suivant : pour chaque invité je prends une bille, et pour chaque arête je place un gros ressort comprimé simulant l'inimitié entre les deux billes ; puis je lâche le tout dans un couloir, les ressorts écartent les billes fâchées qui s'étalent sur toute la longueur du couloir, et j'obtiens mon plan de table en coupant le couloir au milieu : je répartis les invités sur les deux tables en fonction du bout du couloir où a atterri leur bille. Le problème est que je ne disposerais alors d'aucune garantie sur la qualité de la solution que j'aurais ainsi trouvée.

Un mathématicien prouverait très facilement que l'on peut trouver une solution optimale. En effet, j'ai deux choix possibles pour chaque invité : le placer à la première ou à la seconde table. Il n'y a donc que $2 \times 2 \times \ldots \times 2 = 2^n$ façons possibles de répartir les n invités sur les deux tables. Il me suffit donc de toutes les énumérer et de garder celle qui coupe le maximum d'arêtes. Cependant, cette méthode vaut-elle mieux que la précédente ?

Certes, elle trouvera la solution optimale, mais au bout d'un temps *exponentiellement* long : 2^n essais. Par exemple, avec 60 invités, il faudrait énumérer et tester 2^{60} soit environ 10^{18} partitions, ce qui prendrait approximativement 10^9 secondes soit à peu près 100 ans sur un hypothétique processeur surpuissant cadencé à 100 GHz. Et si l'on invitait 61 personnes, ce serait 200 ans qu'il faudrait attendre, 400 ans pour 62, etc. Autant dire que cette solution n'en est pas une dès lors que le nombre d'invités augmente un peu : ajouter un seul invité double le temps de calcul (et ce *indépendamment* de la vitesse du processeur) ! Inversement, doubler la vitesse du processeur ne permet de résoudre une situation qu'avec un invité de plus seulement !

Avec 60 invités, il faudrait énumérer et tester 2^{60} soit environ 10^{18} partitions, ce qui prendrait approximativement 10^9 secondes soit à peu près 100 ans sur un hypothétique processeur surpuissant cadencé à 100 GHz.

Réduire et tomber sur un os

Peut-on trouver la meilleure solution à notre problème dans un temps raisonnable ? En tant qu'informaticien, je vous répondrai que nous avons de bonnes raisons de penser que

Méthodes efficaces et inefficaces

En informatique, le consensus traditionnel est de considérer qu'une méthode de résolution est *efficace* si le nombre d'opérations à effectuer est *au plus polynomial* en fonction de la taille du problème (au plus n^{α} pour un problème de taille n, avec α quelconque). Pour une méthode polynomiale, doubler la taille du problème augmente le temps de résolution d'un facteur constant $2^{\frac{1}{\alpha}}$: les variations du temps de calcul « suivent » celles de la taille des données dans des proportions bornées.

À l'opposé, l'effet de la taille sur un algorithme de résolution en temps exponentiel (par exemple 2^n) est catastrophique :

le temps de calcul d'un algorithme en temps exponentiel double chaque fois que la taille du problème augmente de un seulement ! ($2^{n+1} = 2 \times 2^n$). Par exemple, si la puissance de calcul des ordinateurs double tous les 18 mois, alors, avec un algorithme polynomial, on peut espérer résoudre des problèmes $2^{\frac{1}{\alpha}}$ fois plus grands tous les 18 mois, alors qu'avec un algorithme exponentiel, on ne peut espérer résoudre que des problèmes de taille +1 tous les 18 mois…

Le nombre d'opérations en fonction de la taille n de l'entrée du problème à résoudre est une mesure *intrinsèque* de la complexité d'un problème : *indépendamment de la machine utilisée*, plus cette fonction croît rapidement, plus il sera difficile de résoudre ce problème en grande taille.

ce n'est pas gagné car ce problème (dont le nom scientifique est *Max-Cut*) est « un des plus durs ». Voici ce que je veux dire par là.

Pour résoudre un problème avec un ordinateur, il faut lui donner une recette (le *programme*) qui le résolve et ce quelles que soient les valeurs des paramètres (ici : le nombre d'invités et la liste des arêtes d'inimitié entre eux). Pour écrire cette recette (je veux dire, ce *programme* ; on parle aussi d'*algorithme*), on peut l'écrire de toutes pièces, ou bien s'appuyer sur d'autres recettes afin d'éviter de perdre trop de temps (ou risquer d'introduire de nouveaux bugs). Par exemple, on peut imaginer que la recette qui résoudrait le problème hypothétique *A* = « sortir ce robot de cette pièce » pourrait s'appuyer sur la recette *B* = « trouver la porte de cette pièce » et qu'une fois la porte trouvée il est relativement simple d'y aller pour sortir. Puisqu'on sait comment résoudre *A* facilement si on sait résoudre *B*, cela démontre que le problème *A* est *moins difficile* que *B*. C'est ainsi que les informaticiens démontrent qu'un problème *A* n'est pas plus dur qu'un autre *B* : en supposant connue une recette pour *B* et en écrivant alors une recette efficace pour *A* qui utilise celle de *B*. On dit alors que la résolution de *A se réduit* à celle de *B*.

Or, il se trouve que notre problème de placement d'invité (*Max-Cut*) fait parti du club des problèmes auxquels on peut réduire n'importe quel problème « *raisonnable* » : si vous me fournissez une recette efficace à notre problème Max-Cut, je peux vous écrire une recette qui résout n'importe quel problème « raisonnable ». Cela veut dire que parmi les problèmes que l'on pourra résoudre en un temps « raisonnable », Max-Cut est l'un des plus durs ! On dit que Max-Cut est *NP-complet* (voir l'encadré *Les problèmes les plus durs : les problèmes NP-complets*).

La grande question ouverte en informatique est : est-ce que l'on pourra un jour trouver

P =? NP : *la mécanisation de l'intuition est-elle possible ?*

Une question fondamentale soulevée par l'informatique est exprimée par la formule

P =? NP

qui demande s'il existe une méthode de résolution efficace pour tous les problèmes pour lesquels on peut *vérifier* efficacement la solution une fois qu'on la connaît.

P est l'ensemble des problèmes pour lesquels il existe une construction efficace de la solution (voir l'encadré *Méthodes efficaces et inefficaces*). Pour de nombreux problèmes importants (tout particulièrement pour l'industrie), aucun algorithme efficace n'est connu actuellement. Par exemple, sur les ordinateurs actuels, aucune méthode connue ne trouve efficacement le chemin le plus court passant par toutes les villes d'un pays et retournant à son point de départ avec plus de 1500 villes.

Pourtant, pour la plupart de ces problèmes importants, il est très facile de vérifier efficacement qu'une solution apportée par quelqu'un est effectivement correcte ou non (ces problèmes sont dits NP). Par exemple, on ne connaît aucune méthode efficace pour trouver un coloriage d'une carte avec trois couleurs de sorte que deux pays voisins aient des couleurs différentes, mais cette propriété est très facile à vérifier sur une proposition de coloriage. Cela signifie que pour les problèmes NP, si l'on a d'abord la bonne *intuition* sur ce que devrait être la solution, on peut ensuite vérifier efficacement que la solution est effectivement correcte.

La question P =? NP demande donc si tous les problèmes NP admettent une solution efficace. Elle résume ainsi la question philosophique « l'intuition est-elle mécanisable » ? Cette question figure parmi la liste des sept problèmes mathématiques pour le prochain millénaire sélectionnés par l'institut Clay en 2000 et dont la résolution serait récompensée d'un million de dollars. La réponse, aujourd'hui inconnue, aurait un retentissement bien au-delà du seul cadre de l'informatique.

Les problèmes les plus durs : les problèmes NP-complets

La plupart des problèmes de calcul à résoudre dans la vie réelle appartiennent à la classe NP, ceux où l'intuition peut fonctionner (voir l'encadré *P =? NP: la mécanisation de l'intuition est-elle possible ?*).

Certains problèmes NP sont-ils plus difficiles ou plus importants que d'autres, au sens où ils concentreraient toute la difficulté des problèmes NP ? En 1971, de part et d'autre du mur de Berlin, Stephen A. Cook et Leonid A. Levine ont démontré qu'il existait un problème dans NP qui était effectivement plus difficile que tous les autres, au sens que tous les autres problèmes de NP s'y ramènent : s'il existe une méthode de résolution efficace pour celui-ci, on peut la transformer en une méthode de résolution efficace pour tous les problèmes de NP !

Un tel problème est dit NP-*complet*. Richard M. Karp a démontré l'année suivante qu'en fait de très nombreux problèmes de NP qui n'avaient a priori rien à voir (coloration d'une carte, voyageur de commerce…) sont tous NP-complets ! La résolution d'un seul d'entre eux entraîne la résolution de tous les autres ! C'est par exemple le cas du problème Max-Cut étudié ici.

La plupart des nombreux problèmes étudiés en informatique sont soit dans P, soit NP-complets, et rares sont ceux qui échappent à cette dichotomie – parmi eux cependant un problème emblématique : la factorisation des grands entiers, clé de voûte de la sécurité sur Internet actuellement avec le protocole RSA (voir l'article p. 83).

Malgré des efforts importants, aucune avancée déterminante n'a été encore accomplie indiquant l'existence ou l'impossibilité d'une résolution efficace d'aucun de ces problèmes NP-complets. Vu la diversité de ces problèmes, la plupart des chercheurs pensent qu'il n'existe pas de moyen efficace pour résoudre les problèmes NP-complets.

une solution efficace pour le club des problèmes les plus durs ? La plupart des chercheurs pensent que non, mais nous n'avons pas de réponse claire pour l'instant (voir l'encadré *P =? NP: la mécanisation de l'intuition est-elle possible ?*).

Attendrissons notre problème…

En attendant que cette question trouve sa réponse, il faudrait tout de même que l'on arrive à placer nos invités. Si le problème est trop dur, peut-être qu'en assouplissant quelques contraintes, nous pourrions le

rendre plus sympathique ! Par exemple, nous pourrions nous satisfaire d'une recette qui garantisse de couper au moins 99 % du nombre d'arêtes coupées par la solution optimale. Malheureusement, on peut démontrer qu'approcher à 16/17 soit 94,1 % le nombre d'arêtes coupées optimal fait lui aussi partie du « club des problèmes les plus durs » et n'a donc probablement pas de recette efficace.

Si le problème est trop dur, peut-être qu'en assouplissant quelques contraintes, nous pourrions le rendre plus sympathique !

Essayons donc d'assouplir encore les contraintes. Pour cela, nous allons utiliser une forme originale de programmation : la programmation géométrique. Un des problèmes principaux avec l'approche physicienne évoquée au départ est que les billes et les ressorts ont beaucoup de mal à circuler dans le couloir trop étroit et se gênent les unes les autres, ce qui complique énormément la convergence vers un état d'équilibre optimal. Si l'on se donnait un peu plus d'espace, peut-être que l'on pourrait assurer une convergence efficace du dispositif vers un état d'équilibre optimal et même certifier la qualité de la solution produite.

Un gros oignon et quelques clous de girofle

Voici comment nous allons procéder. Au lieu de les représenter par des billes, nous allons représenter les n invités par autant de vecteurs de taille 1 (on dit *unitaires*) attachés à une origine commune, ceci dans un espace à n dimensions (cela veut juste dire que nous choisissons de donner n composantes à chaque vecteur, afin qu'il ait de la place pour bouger – notre bille d'avant n'avait qu'une seule dimension pour se déplacer : sa position dans le couloir, c'était trop contraignant). Comme pour les billes, nous souhaitons éloigner deux vecteurs correspondant à des invités reliés par une arête d'inimitié. Au lieu de relier ces vecteurs par des ressorts comprimés, nous allons chercher à les faire pointer dans des directions opposées. Pour cela, nous allons utiliser le fait que le *produit scalaire* entre deux vecteurs est positif s'ils pointent dans la même direction et négatif dans le cas contraire, et même d'autant plus négatif qu'ils pointent dans des directions opposées. Nous recherchons alors à disposer les vecteurs de sorte de *minimiser la somme des produits scalaires* entre les vecteurs correspondant à des invités liés par une arête d'inimitié. Or,

Figure 2. Une projection sur un oignon tridimensionnel de nos n vecteurs n-dimensionnels dans la configuration optimale déterminée par nos contraintes sur les produits scalaires.

il se trouve que l'on dispose d'une recette (un programme) pour résoudre *efficacement* ce genre de contraintes portant sur des produits scalaires entre n vecteurs n-dimensionnels, c'est la « programmation semi-définie » généralisant la « programmation linéaire ».

Tranchons dans le tas

Grâce à notre recette, nous avons maintenant obtenu une configuration de nos n vecteurs unitaires n-dimensionnels qui minimise la somme des produits scalaires des vecteurs correspondant aux invités fâchés. Pour les répartir en deux groupes, il me suffit alors de trancher par un plan (en fait on dit un *hyperplan* en dimension n) passant par l'origine des vecteurs, dans une direction choisie *au hasard*: les invités seront sur l'une ou l'autre des tables selon que leurs vecteurs tombent de l'un ou l'autre côté du plan.

Le choix d'une direction aléatoire pour trancher est un ingrédient essentiel de la recette. On peut aussi procéder sans aléatoire, mais c'est un peu plus compliqué et repose paradoxalement de manière cruciale sur l'analyse de la méthode aléatoire. Michel Goemans et David Williamson ont démontré en 1994 que cette solution coupe toujours au moins 87,8 % du nombre optimal d'arêtes. J'ai donc rempli ma mission avec une marge d'erreur relativement faible, garantie inférieure à 12,2 %.

Coupe toujours

De nombreuses variantes de ce problème existent : on peut chercher à couper en deux parties égales, en k morceaux, ou bien ajouter des poids sur les arêtes pour pondérer les inimitiés entre individus : la solution proposée ci-dessus peut être adaptée.

Figure 3. Une fois tranchée suivant une direction aléatoire, nous obtenons la partition des invités ci-dessus, qui coupe effectivement un grand nombre d'inimitiés (en noir), certifié en moyenne supérieur à 87,8 % de l'optimum !

Les problèmes de coupes dans les graphes ont de très nombreuses applications dans la vie pratique. On peut citer par exemple la détection de *communautés* dans les graphes dont les applications sont multiples : en sociologie, où on pourra isoler des groupes d'amis (ayant peu d'inimitié entre eux) ; en marketing, où on pourra suggérer à une personne un achat via son appartenance à tel profil de consommation ; dans la conception de micro-processeurs, où on pourra regrouper les composants en minimisant les câbles qui se croisent…

Pour conclure, on peut voir ici un très joli lien entre une géométrie très *continue* (je peux déplacer mes vecteurs petit à petit dans l'espace) et un problème purement discret (les solutions possibles sont clairement séparées les unes des autres, il n'y a pas d'intermédiaire possible entre placer un invité sur une table ou sur l'autre). Bien que l'informatique soit le royaume du discret (tout y est codé sous la forme d'entiers représentés en mémoire par des suites finies de bits, 0 ou 1), il semble que les interactions entre continu et discret ne se limitent pas à l'obtention de bonnes approximations. En effet, des résultats récents établissent des connexions intrigantes entre des conjectures sur la difficulté de résoudre informatiquement certains problèmes et certaines conjectures purement géométriques.

Ces problèmes ont de très nombreuses applications dans la vie pratique : en marketing, où on pourra suggérer à une personne un achat via son appartenance à tel profil de consommation ; dans la conception de micro-processeurs, où on pourra regrouper les composants en minimisant les câbles qui se croisent…

Mais je me rends compte que mes invités arrivent et qu'il est temps de passer à table avant de risquer une indigestion mathématico-informatique !

Garder le contrôle… … à l'aide des mathématiques

Karine Beauchard, *chargée de recherche CNRS à l'École polytechnique*
Jean-Michel Coron, *professeur à l'Université Pierre et Marie Curie*
Pierre Rouchon, *professeur à Mines-ParisTech*

Contraindre la trajectoire d'un satellite, réguler la température de sa maison, stabiliser le niveau d'eau d'un canal… les situations nécessitant que l'on contrôle une donnée, une quantité, une position, sont omniprésentes. Ces problèmes sont l'objet d'une théorie mathématique très riche : la théorie du contrôle.

Posons un balai verticalement, à l'envers, sur une surface plane (le bout du balai est rond). La position où le balai est vertical, immobile, est une position d'équilibre : en théorie, lorsque le balai se trouve dans cette configuration, il y reste. En pratique, ce n'est pas ce que l'on constate (le balai s'éloigne de la verticale et finit par tomber) parce que le balai n'est pas exactement sur la position d'équilibre (il n'est pas parfaitement vertical ni parfaitement immobile). Ainsi, de petites perturbations produisent de grands effets : même si le balai est près de la position d'équilibre au début de l'expérience, il finit par s'en éloigner énormément. On dit alors que l'équilibre est instable.

Maintenant, posons le balai sur notre index. En bougeant notre doigt habilement, on peut éviter que le balai ne tombe, même si cela requiert un certain entraînement. Dans le jargon de la théorie du contrôle, le mouvement de notre doigt est une *commande* qui agit sur le *système* (le balai). En fait, notre doigt exerce sur le balai une rétroaction, aussi appelée *feedback*, de façon à rendre stable un équilibre instable. Ce feedback repose sur le fait que nos yeux voient le balai tomber et donc si le balai commence par pencher à droite, on bougera notre index vers la droite, et ce d'autant plus vite que le balai penchera vite. Ainsi la position de notre index, la commande, dépend à chaque instant d'une combinaison de la position et de la vitesse du balai : c'est un feedback stabilisant le balai autour de son équilibre vertical instable.

Notre doigt exerce sur le balai une rétroaction, aussi appelée « feedback », de façon à rendre stable un équilibre instable.

Ce problème de stabilisation du balai est typiquement un problème dit de *contrôle*. La théorie du contrôle étudie des systèmes sur lesquels on peut agir au moyen d'une commande. La commande est ici utilisée pour amener le système d'un état initial donné à un état final souhaité malgré des perturbations. Le problème de la stabilisation est une question centrale de la théorie du contrôle.

Les horloges à eau d'Alexandrie

Historiquement, l'un des plus anciens feedbacks construits par l'homme est celui inventé par Ctésibios, ingénieur grec d'Alexandrie au IIIe siècle avant J.-C., pour les horloges à eau ou clepsydres, dont le but est de mesurer le temps qui passe, problème qui remonte aux premières civilisations. Le principe de la clepsydre est le suivant : un premier récipient, percé en son fond d'un trou, est rempli d'eau. L'eau s'écoule par le trou dans un autre récipient. On peut alors mesurer le temps en mesurant la hauteur de l'eau dans le second récipient. Cette méthode a été utilisée par les égyptiens dès le milieu du deuxième millénaire avant J.-C. Le problème de cet appareil est qu'il est difficile d'assurer un débit constant du premier récipient vers le second : plus il y a d'eau dans le premier récipient, plus le débit est important, comme illustré sur la figure 1. Ainsi, le débit de l'eau décroît au cours du temps et le niveau d'eau dans le deuxième récipient n'est pas proportionnel au temps écoulé.

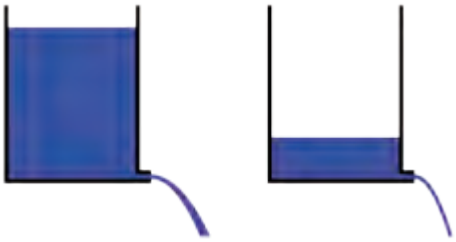


Figure 1. Quand le niveau baisse, le débit diminue.

Pour pallier ce problème, Ctésibios a introduit au IIIe siècle avant J.-C. le dispositif représenté sur la figure 2. Il y a maintenant trois récipients ; l'eau s'écoule du premier récipient vers le second, puis du second vers le troisième. C'est dans le deuxième récipient que se trouve le feedback. Quand le niveau de l'eau dans le récipient 2 est trop élevé, le flotteur rouge vient toucher la coiffe verte. Cela empêche l'eau de passer du récipient 1 vers le récipient 2. Par contre, si le niveau de l'eau dans le récipient 2 est bas, le flotteur rouge ne restreint pas le débit de l'eau du récipient 1 vers le récipient 2. Par construction, le débit du récipient 1 vers le récipient 2 est alors supérieur au débit du récipient 2 vers le récipient 3, donc le niveau de l'eau va monter dans le récipient 2, jusqu'à ce que le flotteur rouge soit de nouveau trop haut. On garde ainsi un niveau d'eau presque constant dans le récipient 2, ce qui assure un débit constant du récipient 2 vers le récipient 3. Le niveau d'eau dans le récipient 3 donne donc une mesure relativement bonne du temps écoulé.

Le premier régulateur industriel

Le premier régulateur utilisé au niveau industriel est le régulateur de James Watt (1736-1819), qui sert à réguler la vitesse de rotation des machines à vapeur. On le voit sur l'image 3 et le schéma 4. Son fonctionnement est le suivant. Lorsque la machine tourne, elle transmet (par un système non représenté) son mouvement à l'axe du régulateur, qui tourne alors autour de la verticale à une vitesse proportionnelle à la vitesse de rotation de la machine. Le cadre du régulateur est solidaire de son axe et tourne donc à la même vitesse. Quand la vitesse de rotation augmente, les deux boules, sous l'action de la force centrifuge, s'écartent de l'axe vertical et le manchon rouge monte. Ce manchon rouge, par un jeu de tringles non représenté, agit sur la vanne d'admission de la vapeur dans la machine. Ainsi, si la machine tourne plus vite que souhaité, le manchon monte : cela provoque une réduction de l'arrivée de vapeur, et la machine finit par tourner moins vite. Par contre, si la machine tourne moins vite que souhaité, le manchon descend : cela provoque une ouverture plus importante de la vanne d'admission, rendant plus importante l'arrivée de vapeur, et la machine se met à tourner plus vite. Le régulateur de Watt exerce ainsi un feedback sur la machine à vapeur. Cette action dépend, à chaque instant, de l'état de la machine à vapeur.

Ce n'est que quatre-vingts ans plus tard que la première analyse mathématique de ce type de régulation sera faite, par James Clerk Maxwell (1831-1879). Le régulateur de Watt y est modélisé par une équation différentielle.

Dans notre monde actuel, les régulateurs sont omniprésents. Par exemple, le thermostat d'une maison stabilise la tempé-

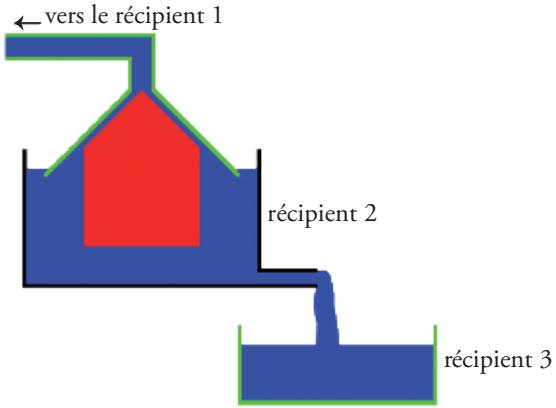


Figure 2. Grâce au flotteur rouge, le débit est maintenant constant.

Figure 3. Ce régulateur est installé sur la célèbre machine à vapeur « Lap » de la société Boulton & Watt. Elle date de 1788 et se trouve actuellement au Science Museum à Londres.

rature de la maison autour de la valeur de consigne imposée par l'usager : le thermostat ajuste en temps réel la position de la vanne, qui détermine si l'eau du circuit des radiateurs entre dans la chaudière (vanne ouverte) ou repart dans le circuit sans être réchauffée (vanne fermée). Les suspensions actives des automobiles réalisent une stabilisation de l'habitacle par feedback. Les drones volants sont munis de feedbacks stabilisants qui les rendent plus faciles à piloter…

Même si James Watt a réussi à réguler sa machine à vapeur sans analyse mathématique, les deux exemples récents de régulateurs que nous allons maintenant découvrir reposent de façon cruciale sur des mathématiques : le premier stabilise des voies navigables, le second des états quantiques fragiles.

Des équations de Saint-Venant dans les écluses

Les canaux navigables sont constitués de biefs, c'est-à-dire d'une succession de petits canaux séparés d'une part par des vannes mobiles, qui permettent d'agir en temps réel sur le niveau de l'eau, et d'autre part par des écluses pour le passage des bateaux (voir figure 5). Sur ces canaux navigables, il est important de réguler le niveau de l'eau ainsi que son débit. Pourquoi ? Tout d'abord, il faut permettre aux transporteurs de conteneurs de calculer la charge maximale qu'ils peuvent embarquer sans risque

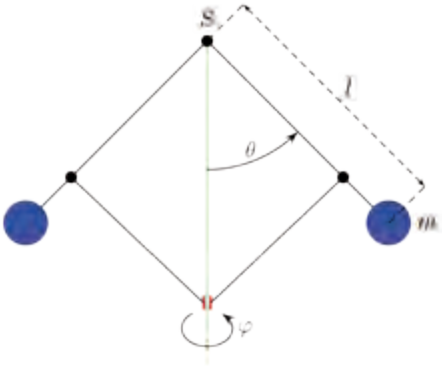


Figure 4. Schéma du régulateur de Watt.

de toucher le fond, ce qui impose de réguler le niveau de l'eau avec précision car les canaux sont souvent peu profonds. Ensuite, il faut garantir l'approvisionnement des centrales électriques et des industries consommatrices d'eau situées sur les berges, ce qui nécessite une régulation du débit.

La commande des vannes d'une écluse se fait de façon automatique. Le processus est géré par des algorithmes qui reposent sur un modèle de contrôle traduisant le comportement de l'eau dans les biefs.

La commande des vannes se fait aujourd'hui de façon automatique. Le processus est géré par des algorithmes qui reposent sur un modèle de contrôle traduisant le comportement de l'eau dans les biefs. Le modèle le plus utilisé est celui des équations de Saint Venant. Son origine remonte au XIX^{e} siècle. Adhémar-Jean-Claude Barré de Saint Venant fut élève de l'École polytechnique, puis enseigna les mathématiques à l'École nationale des ponts et chaussées. C'est en étudiant les rivières en crue et le mouvement des marées dans les estuaires qu'il publia son modèle en 1871. Il s'agit d'équations aux dérivées partielles, traduisant la conservation de la masse et de la quantité de mouvement dans le fluide.

Pour concevoir des lois de commande régissant le comportement des vannes, on utilise une quantité représentative du comportement global du bief : une *fonction de Lyapunov*, c'est-à-dire une fonction qui atteint son minimum lorsque le système est dans l'état d'équilibre souhaité. Cette méthode de stabilisation par l'utilisation de fonctions de Lyapunov est très courante en théorie du contrôle. Elle possède l'avantage

Figure 5. Vannes mobiles pour la régulation des canaux. A droite, la vanne a été sortie de l'eau pour une opération de maintenance.

de conduire à des comportements robustes vis-à-vis des incertitudes de modélisation et des perturbations extérieures. De plus, ces lois de commande sont, dans notre exemple, faciles à mettre en œuvre concrètement.

Cette méthode d'analyse et de conception du contrôle des voies navigables a été utilisée notamment pour le réglage des vannes de la Sambre par l'administration des voies hydrauliques du ministère de l'Équipement et des Transports de la région wallonne en Belgique. On a pu constater des gains de performance significatifs : l'amplitude des variations de niveau est divisée par deux et les perturbations de débit sont amorties deux fois plus vite.

Stabiliser des états quantiques de la lumière grâce aux probabilités

L'expérience qui va suivre constitue la première réalisation expérimentale de feedback sur un système quantique. Il s'agit d'une étape significative vers l'ordinateur quantique, pour lequel la protection d'états quantiques par feedback serait alors une méthode efficace pour lutter contre les perturbations liées à l'environnement (la décohérence). L'objectif de cette expérience est de stabiliser certains états quantiques de la lumière. Ces états comportent un nombre entier et bien déterminé de photons. Ils sont fragiles et difficiles à observer, car très différents de ceux de la lumière qui nous entoure. L'objectif est de maintenir ce nombre de photons. Comme illustré de façon allégorique sur la figure 6, il faut surmonter une difficulté supplémentaire, car à l'échelle quantique, la mesure modifie le système.

Dans cette expérience, les photons sont confinés entre deux miroirs supraconducteurs qui se font face, formant ainsi une cavité ouverte sur les cotés, comme l'illustre la figure 7. La mesure repose sur des atomes qui traversent l'un après l'autre la cavité, interagissent avec les photons, et sont mesurés en sortie. Pour la commande, on dispose d'une source de lumière appropriée, utilisée entre le passage des atomes. (Une description bien plus détaillée de cette expérience se trouve notamment sur le site http://www.lkb.ens.fr/Un-asservissement-quantique où des animations basées sur des données expérimentales sont également présentées.)

Les principes de la mécanique quantique conduisent à une description probabiliste très précise du modèle reliant les signaux de commande et les signaux de mesure. Cette description mathématique, bien que relativement complexe, permet de déterminer un feedback stabilisant qui là encore repose sur une fonction de Lyapunov.

Bien que les modèles utilisés pour les canaux de navigation et les états quantiques de la lumière soient de natures très différentes, leurs contrôles s'appuient sur des méthodes mathématiques proches, dont l'origine remonte aux travaux fondamentaux d'Alexandre Lyapunov sur la stabilité.

Ainsi, bien que les modèles utilisés pour les canaux de navigation et les états quantiques de la lumière soient de natures très différentes, le premier étant déterministe en *temps continu* et le second aléatoire et en *temps discret*, leurs contrôles s'appuient sur des méthodes mathématiques proches dont l'origine remonte aux travaux fondamentaux d'Alexandre Lyapunov sur la stabilité.

Dans les problèmes de contrôle, le progrès ne vient donc pas uniquement d'inventions purement techniques ou expérimentales. Il naît aussi des recherches abstraites de la théorie mathématique du contrôle. Les outils mathématiques qui y sont utilisés sont très variés (équations différentielles, équations aux dérivées partielles, processus stochastiques…) et les recherches s'y font souvent à l'interface avec d'autres disciplines (mécanique des fluides, ingénierie hydraulique, mécanique quantique, optique quantique…).

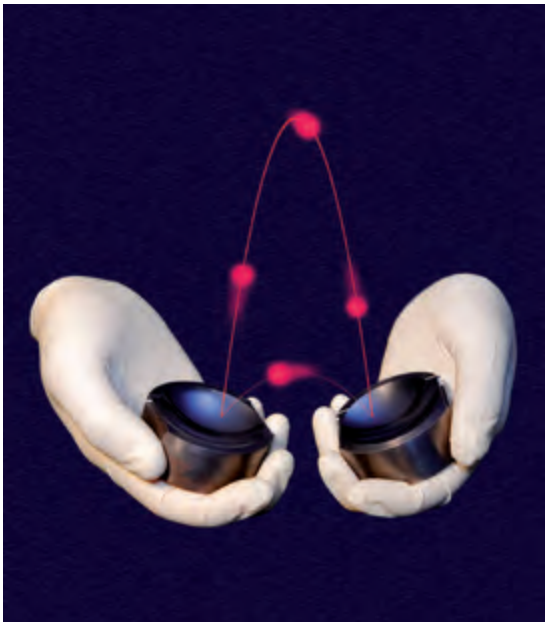


Figure 6: Les états quantiques à stabiliser sont des états avec un nombre entier et bien déterminé de photons. Ces derniers rebondissent entre les deux miroirs. La difficulté pour le feedback, ici représenté de façon allégorique par les deux mains du jongleur, vient du fait que l'observation de ces photons par le jongleur les perturbe nécessairement. Cette action en retour supplémentaire due à la mesure doit être prise en compte dès la conception d'un feedback stabilisant.

Remerciements
Les auteurs remercient Michel Brune pour les figures 6 et 7 illustrant le contrôle d'états quantiques.

Bibliographie

Coron J.-M., d'Andréa-Novel B., Bastin G., (mars 2008). *Penser globalement, agir localement*. La Recherche, n° 417, p. 82-83.

Mayr O., (1970). *The origins of feedback control*. The M.I.T. Press (Cambridge Massachusetts, and London, England), p. vii +151.

Sayrin C., Dotsenko I., Zhou X., Peaudecerf B., Rybarczyk T., Gleyzes S., Rouchon P., Mirrahimi M., Amini H., Brune M., Raimond J.-M., Haroche S., (septembre 2011). *Real time quantum feedback prepares and stabilizes photon number states*. Nature, vol. 477, p. 73-77.

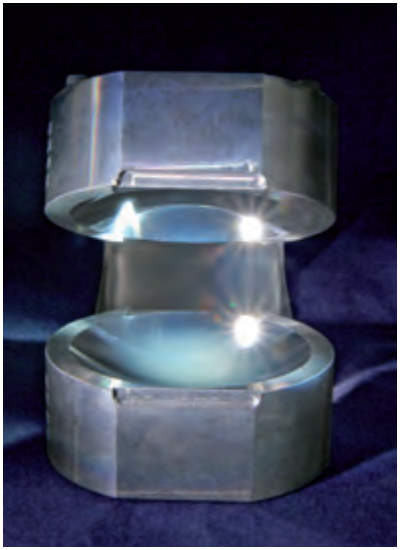


Figure 7. la lumière à contrôler est un champ électromagnétique d'une fréquence voisine de 10 GHz confiné entre les deux miroirs supraconducteurs qui se font face comme sur la figure. Ils forment ainsi une cavité ouverte sur les cotés. La mesure repose sur des atomes (non représentés sur la figure) qui traversent horizontalement la cavité à mi-chemin entre les deux miroirs. Ces atomes interagissent avec le champ électromagnétique et sont mesurés ensuite à leur sortie. Pour la commande, on dispose d'une source électromagnétique classique réglable (non représentée sur la figure) qui, après chaque atome, éclaire à la façon d'un flash la cavité.

Le théorème de Green-Tao et autres secrets des nombres premiers

Michel Waldschmidt, *professeur émérite à l'Université Pierre et Marie Curie*

Bien que les mathématiciens s'y intéressent depuis l'Antiquité, les nombres premiers continuent de fasciner. En les additionnant ou en les soustrayant entre eux, on trouve une mine de problèmes dont certains sont longtemps demeurés ouverts ou restent encore irrésolus.

Obtenir tous les nombres entiers positifs en utilisant les seuls symboles + et 1 est facile : il suffit d'écrire chaque entier $n > 0$ sous la forme 1 + 1 + … + 1 avec n fois le symbole 1 et (n-1) fois le symbole +.

Remplaçons l'addition par la multiplication et le symbole + par le symbole x (ou •). Pour écrire tous les nombres entiers positifs comme des produits, on a donc besoin de *briques* élémentaires qui jouent le rôle tenu par 1 pour l'addition. Ces briques sont les *nombres premiers*, ceux qui ne peuvent pas s'écrire comme produits de nombres entiers plus petits. Chaque entier supérieur ou égal à 2 est produit, de manière unique (à permutations près), de nombres premiers : c'est ce qu'on appelle le *théorème fondamental de l'arithmétique*. Par exemple 4200 = 2 x 2 x 2 x 3 x 5 x 5 x 7.

Des nombres aussi grands que l'on veut

La liste des nombres premiers commence par 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37…

On trouve le début de cette liste sur la toile, dans l'encyclopédie des suites de nombres entiers, qui est une mine d'informations pour ce genre de questions. Les nombres supérieurs à 1 qui ne sont pas premiers sont appelés nombres *composés*. Leur liste commence par 4, 6, 8, 9, 10, 12, 14, 15, 16…

Cette deuxième liste comporte entre autres tous les nombres pairs à partir de 4. Il est donc facile d'écrire un nombre composé aussi grand que l'on veut. En revanche, écrire (par exemple sous forme décimale) un nombre premier aussi grand que l'on voudrait est une opération que l'on ne sait pas réaliser.

Le plus grand nombre premier explicitement connu (depuis le 2 mai 2013) est $2^{57\,885\,161} - 1$ qui a 17 425 170 chiffres décimaux. Si on écrivait tous ces chiffres avec disons deux chiffres par centimètre, l'écriture de ce nombre s'étendrait sur plus de 87 kilomètres !

Pourtant on sait, depuis Euclide, que la liste des nombres premiers ne s'arrête jamais : *il existe une infinité de nombres premiers*. Les nombres premiers sont donc un exemple de suite dont on sait qu'elle est infinie, mais dont on ne sait pas expliciter des éléments aussi grands que l'on voudrait.

Des propriétés additives étonnantes

Les nombres premiers ont été introduits en arithmétique pour la multiplication, comme nous venons de le voir. Étudier leurs propriétés additives peut sembler une idée bizarre. Elle mène cependant à des problèmes profonds, dont l'étude a conduit à des théories très élaborées d'une grande richesse, nécessitant l'utilisation d'outils mathématiques sophistiqués. Le plus fascinant dans cette théorie est qu'elle continue à se développer. Elle a permis de résoudre un grand nombre de questions qui sont restées ouvertes fort longtemps, mais elle n'est pas encore suffisamment puissante pour répondre à certaines questions qui sont pourtant faciles à formuler.

Additionnons deux nombres premiers ; on obtient les valeurs suivantes :

p	2,	3,	5,	7,	11,	13,	17,	19,	23,	29,	31,	37…
$p+2$	4,	5,	7,	9,	13,	15,	19,	21,	25,	31,	33,	39…
$p+3$	5,	6,	8,	10,	14,	16,	20,	22,	26,	32,	34,	40…
$p+5$	7,	8,	10,	12,	16,	18,	22,	24,	28,	34,	36,	42…
$p+7$	9,	10,	12,	14,	18,	20,	24,	26,	30,	36,	38,	44…
$p+11$	13,	14,	16,	18,	22,	24,	28,	30,	34,	40,	42,	48…
…	…											

Les seuls nombres impairs que l'on puisse trouver dans ce tableau figurent sur la première ligne (celle qui donne les $p + 2$) et sur la première colonne (celle qui donne les $2 + p$, puisque le tableau est symétrique). En effet, la somme de deux nombres impairs est forcément un nombre pair. Comme 2 est le seul nombre premier pair, les seules sommes impaires de deux nombres pre-

miers sont celles qui consistent à ajouter 2 à un nombre premier impair ; toutes les autres sommes sont des nombres pairs.

Ce qui est remarquable est qu'il semble qu'on obtienne avec ce tableau tous les nombres pairs à partir de 4. Beaucoup de nombres pairs sont obtenus de plusieurs façons différentes, par exemple,
10 = 5 + 5 = 7 + 3 ; 14 = 7 + 7 = 3 + 11…

Que tous les nombres pairs supérieurs ou égaux à 4 puissent s'écrire comme somme de deux nombres premiers est un problème ouvert.

Que tous les nombres pairs supérieurs ou égaux à 4 puissent s'écrire comme somme de deux nombres premiers est appelé problème binaire de Goldbach ; il a été proposé par Euler en réponse à une lettre de Goldbach datant de 1742, dans laquelle Goldbach suggérait que tout nombre entier supérieur ou égal à 7 était somme de trois nombres premiers. Ce dernier résultat, moins fort que le problème binaire de Goldbach, a été annoncé par Harald Helfgott en mai 2013. Mais le problème binaire de Goldbach est toujours ouvert et c'est la source de nombreuses recherches depuis bientôt trois siècles. C'est un problème difficile, qui a donné naissance à des théories élaborées ayant d'autres applications. On peut citer par exemple la théorie additive des nombres, qui étudie les sommes de nombres entiers appartenant à un ensemble donné (ici les nombres premiers).

La différence entre deux nombres premiers : une source de questions passionnantes

Une autre façon de marier les nombres premiers avec l'addition consiste à étudier la différence entre deux nombres premiers consécutifs.

La liste des nombres premiers est $(p_1, p_2, p_3\ldots)$ = (2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37…) et la liste $(p_n - p_{n-1})$ des différences entre deux nombres premiers consécutifs commence par 1, 2, 2, 4, 2, 4, 2, 4, 6, 2, 6…

Il est facile de voir qu'à part la première valeur, 1, toutes les autres sont des nombres pairs : cela provient une fois de plus du fait que 2 est le seul nombre premier pair. Y a-t-il une infinité de 2 dans cette liste ? On ne le sait pas encore ; c'est la conjecture des nombres premiers jumeaux, qui s'énonce : *il existe une infinité de nombres premiers p ayant la propriété que $p + 2$ est aussi premier.*

On sait que la suite des différences entre deux nombres consécutifs de la suite des nombres premiers n'est pas bornée : il existe des « trous » aussi grands qu'on veut (l'étude des « trous » dans la suite des nombres premiers remonte à Legendre et Gauss). On sait aussi estimer l'ordre de grandeur moyen de tels trous, grâce au Théorème des Nombres Premiers, démontré en 1896 par Hadamard et de la Vallée Poussin. La légende prétendait que le mathématicien qui démontrerait cette conjecture serait immortel. D'une certaine manière Jacques

Hadamard (1865-1963) et Charles de la Vallée Poussin (1866-1962) laissent leur nom dans l'histoire par leur théorème, mais ils ont aussi réalisé la prédiction de façon aussi complète que possible, puisqu'ils sont décédés presque centenaires.

Récemment, de nouveaux progrès ont été accomplis sur la question des trous entre deux nombres premiers. Le postulat de Bertrand, démontré par Tschebycheff vers 1850, affirme qu'entre un nombre entier et son double, il y a toujours un nombre premier.

Au lieu du coefficient 2 qui intervient dans le double, on peut prendre n'importe quel facteur $c > 1$: pour n suffisamment grand, il y a toujours un nombre premier p entre n et cn. Cela donne déjà une information sur les écarts entre deux éléments consécutifs de la suite des nombres premiers. Le théorème des nombres premiers implique plus précisément que la différence entre le n-ième nombre premier p_n et le suivant p_{n+1} est de l'ordre de grandeur de $\log p_n$, le logarithme népérien de p_n.

En 2013, Yitang Xang a démontré qu'il existe une infinité de couples de nombres premiers dont l'écart est inférieur à 70 millions.

C'est loin de ce que l'on attend, si on pense à la conjecture des nombres premiers jumeaux ci-dessus, selon laquelle la différence entre deux nombres consécutifs de la liste des nombres premiers devrait être 2 une infinité de fois. Mais c'est un progrès considérable par rapport à ce que l'on savait avant.

Progressions arithmétiques

L'étude des différences entre les nombres premiers peut être généralisée d'une autre manière, en étudiant les progressions arithmétiques composées de nombres premiers. Une suite de nombres forme une *progression arithmétique* si la différence entre deux termes consécutifs de la suite est toujours la même : cette valeur est alors la *raison* de la progression.

Par exemple la suite 5, 11, 17, 23, 29 est une progression arithmétique de 5 termes de raison 6.

La suite 7, 37, 67, 97, 127, 157 est une progression arithmétique de 6 termes de raison 30 et la suite 199, 409, 619, 829, 1039, 1249, 1459, 1669, 1879, 2089 est une progression arithmétique de 10 termes de raison 2100. Dans ces exemples, les termes de ces suites sont tous des nombres premiers.

Y a-t-il des progressions arithmétiques aussi longues que l'on veut composées uniquement de nombres premiers ?

La plus longue progression arithmétique explicitement connue composée uniquement de nombres premiers comporte 26 termes, la raison en est le produit de 23 681 770 par 223 092 870 ; c'est la suite $43\,142\,746\,595\,714\,191 + 23\,681\,770 \cdot 223\,092\,870\,n$ pour n prenant les valeurs de 0 à 25. Elle a été trouvée par Benoît Perichon.

Ben Green

Terence Tao

Y a-t-il des progressions arithmétiques aussi longues que l'on veut composées uniquement de nombres premiers ? Pendant longtemps cette question a été ouverte, c'était encore un défi de plus pour les mathématiciens. C'est un travail mené en collaboration par les mathématiciens Ben Green et Terence Tao qui a conduit finalement à la solution en 2004. Leur théorème dit ainsi que oui, il existe des progressions arithmétiques, de longueur aussi grande que l'on veut, composées uniquement de nombres premiers. Leur démonstration n'est pas constructive, c'est-à-dire qu'elle ne dit pas comment construire une telle suite !

Pour aller plus loin

Tenenbaum G., Mendès-France M., (1997). *Les Nombres premiers* ; Presses Universitaires de France (PUF), coll. Que Sais-je ?.

Ribenboim P., (2000). *Nombres premiers : mystères et records* ; Presses Universitaires de France (PUF), coll. Mathématiques.

Hardy G.H., Wright E.M., (2008). *An Introduction to the Theory of Numbers* ; Oxford Science Publications, sixth ed. Oxford : Oxford University Press. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

Sur Internet

http://www.mersenne.org/
http://primes.utm.edu/largest.html
https://oeis.org/

Le blog de Terence Tao :
http://terrytao.wordpress.com/

La supraconductivité

Sylvia Serfaty, *professeur à l'Université Pierre et Marie Curie*

La supraconductivité, capacité d'un métal à laisser passer le courant électrique sans perte d'énergie, peut avoir des applications étonnantes. L'étude de ce phénomène fait intervenir divers domaines des mathématiques, comme le calcul des variations, les équations aux dérivées partielles, l'analyse asymptotique. Plusieurs questions ouvertes y sont associées.

La supraconductivité vient de fêter ses 100 ans. Plus exactement, c'est en 1911 que le physicien hollandais Heike Kamerlingh Onnes a découvert qu'une fois refroidi à une température très proche du zéro absolu, le mercure perdait complètement sa résistance électrique : autrement dit, on pouvait y voir circuler des courants électriques indéfiniment, sans dissipation d'énergie ni perte de chaleur.

Lévitation et tourbillons

La supraconductivité, qui concerne de nombreux métaux et alliages, entraîne d'autres phénomènes surprenants, notamment en réponse à un champ magnétique. Lorsqu'un supraconducteur, refroidi en-dessous de sa *température critique*, est soumis à un champ magnétique extérieur, il l'expulse littéralement. Ceci s'appelle l'*effet Meissner*.

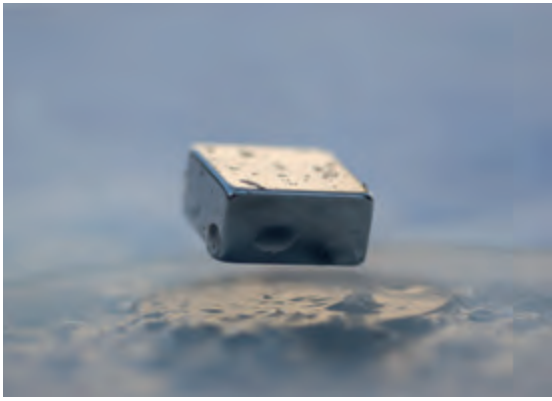


Figure 1. Supraconducteur en lévitation au dessus d'un aimant : l'effet Meissner. Copyright : J. Bobrov

L'une des conséquences les plus spectaculaires de cet effet est qu'un aimant posé sur un supraconducteur est soumis à une force magnétique qui le repousse : on voit l'aimant léviter au-dessus du supraconducteur !

Cependant, lorsque certains matériaux supraconducteurs sont soumis à un champ magnétique plus intense, la situation change : au-delà d'un certain seuil de champ, appelé *premier champ critique*, le champ magnétique commence à pénétrer un peu, et aux endroits où il pénètre, on voit se former des tourbillons de vorticité ou *vortex*, qui sont comme les tourbillons d'un mini-cyclone, le cœur étant dans l'état normal (non supraconducteur) entouré de phase supraconductrice, autour desquels tournent des boucles de *courant supraconducteur*. Plus le champ appliqué est élevé, et plus on voit de vortex. Une fois devenus nombreux, ceux-ci s'organisent en réseaux hexagonaux parfaits, comme en nid d'abeille (cf. photo). Mais au-delà d'une certaine intensité, appelée le *second champ critique*, les vortex sont tellement serrés les uns contre les autres qu'il n'y a plus de place pour la phase supraconductrice, et le matériau retrouve l'état normal.

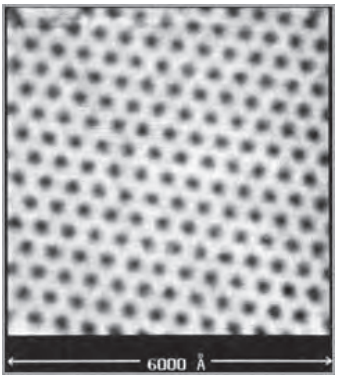


Figure 2. Vortex organisés en réseau d'Abrikosov, d'après H. F. Hess et al., Bell Labs, Phys. Rev. Lett. 62, 214 (1989).

Un aimant posé sur un supraconducteur est soumis à une force magnétique qui le repousse : on voit l'aimant léviter au-dessus du supraconducteur !

Un potentiel énorme en termes d'applications

Les applications de la supraconductivité sont potentiellement énormes. En effet, ce phénomène permet en principe de transporter des courants très forts sans perte d'énergie. Déjà, les systèmes de trains à sustentation magnétique, comme le japonais Maglev, reposent sur la supraconductivité. Ces trains sont équipés de bobines supraconductrices leur permettant de glisser sans frottement en lévitation au-dessus de rails aimantés. Les supraconducteurs sont également utilisés pour engendrer de très forts champs magnétiques, en y faisant circuler des courants de forte intensité, comme par exemple dans le nouvel accélérateur de particules LHC de Genève. Ils servent aussi à détecter de très faibles champs magnétiques, ce qui est utilisé entre autres dans les sous-marins, en imagerie médicale, pour la détection d'objets cachés et dans les nano-circuits. Pour l'instant, la principale limitation est que les supraconducteurs doivent être équipés de systèmes de refroidissement pour des-

cendre à très basse température, ce qui est coûteux et contraignant. Les physiciens, en partenariat avec les chimistes, continuent donc d'explorer la possibilité de fabriquer des composés chimiques qui soient supraconducteurs à des températures plus élevées.

Un peu de physique quantique

Que se passe-t-il dans un matériau qui fait qu'il devient supraconducteur à basse température ? Seule la physique quantique peut nous l'expliquer. Dans un métal normal, les électrons se comportent comme des ondes étalées sur chaque atome, indépendantes les unes des autres. Quand le métal devient supraconducteur, à basse température, les électrons s'associent par paires, appelées *paires de Cooper*. Grossièrement, toutes les paires d'électrons se mettent alors dans le même état quantique pour former comme une seule onde, un peu comme des poissons dans un banc. Ce comportement collectif des électrons leur permet de traverser le matériau sans être sensibles aux obstacles et ainsi la résistance électrique disparaît. La *condensation* dans un seul état quantique est d'ailleurs exactement ce qui se produit aussi pour les atomes dans les *superfluides* (tel l'hélium liquide à basse température) ou les condensats de Bose-Einstein, prédits par Einstein et récemment mis en évidence expérimentalement. Les superfluides et les condensats sont deux types de liquides qui perdent toute viscosité à très basse température, d'où leur nom de *liquides superfluides*. Et lorsqu'on met un liquide superfluide en rotation rapide, on y observe, comme dans les supraconducteurs, des tourbillons de vorticité.

Dans les années 30, le physicien London a fourni le premier modèle descriptif de la supraconductivité, mais ce n'est que dans les années 50 qu'est née la théorie de Ginzburg et Landau, la plus universellement admise de nos jours, qui décrit l'apparition de la supraconductivité et le comportement des supraconducteurs en présence de champ magnétique. Cette théorie décrit un supraconducteur par deux « fonctions » indiquant son état local : une *fonction d'onde* ψ, et un champ magnétique A, chacun dépendant du point dans l'échantillon où l'on se place.

A ces deux fonctions on associe alors, par une formule explicite, un nombre unique, dépendant de l'intensité du champ appliqué, et qui donne ce qui s'appelle *l'énergie de Ginzburg-Landau*. Les états stables du système, ceux que l'on observe lorsqu'il est au repos, sont ceux pour lesquels l'énergie de Ginzburg-Landau est la plus basse. En 1956, la théorie BCS de Bardeen, Cooper et Schrieffer est venue compléter la théorie de Ginzburg et Landau, en expliquant la supraconductivité à partir du niveau quantique, par la formation des paires de Cooper.

Que viennent faire les mathématiques dans tout cela ?

L'énergie de Ginzburg-Landau, qui est une fonction de l'état du système, constitue une sorte de fonction mathématique. En outre,

de manière pas si surprenante, c'est quasiment la même fonction mathématique que celle qui décrit l'énergie d'un superfluide ou d'un condensat de Bose-Einstein en rotation. On peut donc faire semblant d'oublier le modèle physique dont cette énergie provient et, la regardant avec un œil de mathématicien, l'étudier comme fonction mathématique, à l'aide de démonstrations, obtenant ainsi du même coup des résultats sur ces trois problèmes physiques. On se demande alors : quelle est la valeur du minimum de cette fonction ? À quoi ressemblent les états qui atteignent ce minimum ? Comment varient-ils en fonction de la valeur du champ appliqué ? Peut-on prédire l'apparition des vortex (caractérisés comme les points où la fonction ψ vaut 0) et leur emplacement ?

Il se trouve que le même genre de questions se pose pour toutes sortes d'énergies qui proviennent de la physique, et même de l'ingénierie, de l'économie… où l'on cherche à minimiser un certain coût (ou, dans le cas de l'économie, à maximiser un gain ou une *fonction d'utilité*) et où l'on veut comprendre quels sont les états ou configurations optimaux. Mathématiquement, cela relève de la théorie du *calcul des variations*, qui remonte à Bernoulli, Euler et Lagrange, et qui nous dit quand on peut garantir l'existence d'un optimum, et comment le caractériser. Le plus souvent, celui-ci est caractérisé par une *équation aux dérivées partielles*, qui n'est autre qu'une relation exacte entre la fonction d'état (qui dépend du lieu dans l'échantillon) et ses dérivées. Or, les mathématiciens ont également développé depuis un siècle une gigantesque théorie et classification de ces équations aux dérivées partielles. Même si de nombreuses et difficiles questions subsistent, on sait assez bien dire lesquelles ont des solutions, comment celles-ci se comportent qualitativement dans l'espace et dans le temps, etc.

Les mathématiciens savent dire, par exemple, comment les vortex vont se déplacer sous l'effet d'un courant appliqué, question importante puisque les mouvements des vortex entraînent de la déperdition d'énergie.

Mais revenons à cette fameuse énergie de Ginzburg-Landau : comment savoir si cette énergie, proposée à l'origine sans démonstration, est un bon modèle pour décrire la supraconductivité ? Précisément, en regardant si, mathématiquement, elle prédit le même comportement que celui qui est observé dans les expériences. Ce problème se pose d'ailleurs pour tout modèle physique : l'étude mathématique permet d'abord de valider le modèle, puis au-delà, de prédire le comportement du système physique avec plus de détails, et dans des situations qui ne sont par exemple pas accessibles à l'expérience ! Ainsi, le physicien Abrikosov, peu après l'introduction du modèle de Ginzburg et Landau, a fait des calculs, quelque peu formels, qui l'ont amené à dire que le modèle devrait avoir des solutions avec des vortex disposés périodiquement, et ainsi à prédire que l'on devrait les voir dans les expériences. Les physiciens ont ensuite essayé d'observer de telles situations et ils ont effectivement découvert les arrangements de vortex en réseau hexagonal tels que sur la photo de la

page 38. Ceux-ci ont été baptisés *réseaux d'Abrikosov*.

Par rapport à la physique et au calcul formel, les mathématiques apportent toute la rigueur et la précision de leur méthode hypothético-déductive. De ce point de vue, le modèle de Ginzburg-Landau est déjà très largement validé mathématiquement. On sait en effet démontrer pour quelles valeurs (dites critiques) du champ magnétique appliqué les vortex apparaissent dans les minimiseurs de l'énergie, combien il y en a, et comment ils se disposent en moyenne. On retrouve et précise ainsi le comportement observé ou calculé par les physiciens. Les mathématiciens savent aussi dire, par exemple, comment les vortex vont se déplacer sous l'effet d'un courant appliqué, question importante puisque les mouvements des vortex entraînent de la déperdition d'énergie.

Analyse asymptotique

L'analyse mathématique de la supraconductivité fait donc intervenir le calcul des variations, la théorie des équations aux dérivées partielles, la *théorie spectrale*, mais aussi de manière importante, un autre ingrédient qui s'appelle l'*analyse asymptotique*. Dans les supraconducteurs, le comportement des vortex dépend d'une constante K, appelée le paramètre de Ginzburg-Landau, qui est propre à chaque matériau et qui apparaît naturellement dans le calcul de l'énergie de Ginzburg-Landau. En fait, ce n'est que dans les matériaux pour lesquels cette constante est assez grande que les vortex apparaissent. L'idée de l'analyse asymptotique est alors de « forcer le trait » : si l'on s'intéresse aux matériaux pour lesquels les vortex apparaissent – donc pour lesquels cette constante est grande – pourquoi ne pas regarder ce qui se passe mathématiquement lorsqu'on la rend *infiniment grande* ? Cette idée s'avère en fait très fructueuse dans toute l'analyse d'équations ou d'énergies provenant de modèles physiques. La philosophie, qui est très générale, est donc d'arriver à réduire un problème de minimisation compliqué (sur un espace de dimension infinie) à un problème plus simple, en dimension plus petite (c'est-à-dire avec moins de paramètres possibles à examiner). La difficulté est d'arriver à justifier rigoureusement cette simplification (en quelque sorte, à montrer qu'elle est « légale ») et à donner la formule explicite de l'énergie réduite.

La meilleure disposition possible

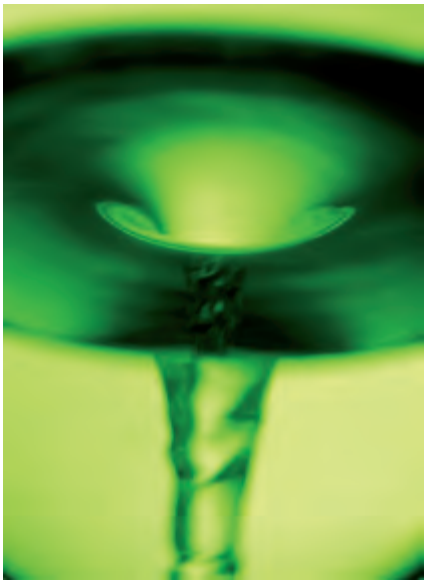
Ce n'est que très récemment que nous sommes parvenus à expliciter, à partir de Ginzburg-Landau, la formule de l'énergie simplifiée, qui régit la distribution et la répartition des vortex. Cette énergie n'est pas inconnue des physiciens ni des mathématiciens : elle correspond à une interaction de type électrostatique entre les vortex. Ceux-ci se comportent donc comme des charges électriques ponctuelles qui se repoussent entre elles, mais sont en même temps confinées par une force extérieure (la force du champ magnétique appliqué). C'est la compétition entre ces deux effets

qui dicte la répartition. Par analogie, imaginez des personnes nombreuses qui se détestent mais qui sont forcées de rester toutes ensemble dans la même pièce : quelle est la meilleure manière que toutes ont de se positionner pour être le plus loin possible les unes des autres ?

Ce dernier problème ne se limite pas à la supraconductivité, c'est en fait un type de question courant dans la nature. Pourquoi les atomes dans un cristal se disposent-ils exactement selon un réseau cristallin (en général cubique), c'est-à-dire de manière périodique ? De nouveau, c'est la somme de toutes les interactions électrostatiques entre les atomes, et entre les atomes et les électrons, qui doit dicter la disposition. On peut mesurer le coût de chaque répartition via une énergie, dont on pense que le minimum est justement atteint par un tel arrangement en réseau. Hélas, on ne sait pas démontrer que ces structures cristallines sont bien celles qui minimisent cette énergie, même si on les observe clairement dans la nature… Le seul cas où l'on sache démontrer quelque chose d'analogue est celui où l'on cherche la meilleure distribution de boules rigides toutes identiques dans un plan, telles des boules de billard, lorsqu'on veut en mettre le maximum possible par unité de volume. Dans ce cas, il a été démontré assez récemment que la meilleure configuration est celle en réseau hexagonal (en dimension 3, le résultat analogue donne la meilleure manière d'empiler des oranges sur l'étal d'un marchand !)

Ces questions de périodicité sont donc générales, fondamentales dans la nature, et constituent un réel défi mathématique. Dans le cas de la répulsion électrostatique qui gouverne l'interaction entre les vortex dans Ginzburg-Landau, on pense que le meilleur arrangement est de nouveau le réseau hexagonal, toujours le fameux réseau d'Abrikosov ! Mais, pour le moment, on ne sait pas le démontrer rigoureusement. Le modèle de Ginzburg-Landau n'a donc pas encore livré tous ses secrets…

La meilleure distribution de boules rigides identiques dans un plan, telles des boules de billard, est celle en réseau hexagonal (en dimension 3, le résultat analogue donne la meilleure manière d'empiler des oranges sur l'étal d'un marchand !)



Inspiration mathématique : la modélisation du poumon

Céline Grandmont, *directrice de recherche à Inria*

La complexité de notre système respiratoire en fait un joli sujet d'application des mathématiques. Le fonctionnement de l'appareil respiratoire est décrit par des équations qui servent à effectuer des simulations venant compléter l'expérience et permettant de mieux comprendre ou prévoir les phénomènes qui se produisent lorsque nous respirons.

Quels sont les phénomènes physiologiques intervenant au cours d'une crise d'asthme ? Où se déposent les aérosols thérapeutiques ou les particules polluantes émises par les gaz d'échappement dans le poumon ? Comment ventiler au mieux un patient ayant besoin d'une assistance respiratoire ? Comment l'emphysème ou la fibrose, maladies touchant les tissus pulmonaires, affectent la ventilation et la diffusion de l'oxygène dans le sang ?

Toutes ces questions intéressent des médecins – pneumologues, allergologues – des biologistes, des mécaniciens des fluides mais aussi des mathématiciens, qui cherchent à y amener des éléments de réponses.

L'arbre bronchique : une géométrie complexe

La respiration fait intervenir des phénomènes physiques nombreux, à des échelles très différentes. La fonction principale de l'appareil respiratoire est d'assurer les échanges gazeux entre l'air et le sang. L'air, inspiré par le nez ou la bouche, est ensuite transporté dans les voies aériennes : pharynx, larynx, trachée, bronches… L'arbre bronchique humain présente une géométrie d'une grande complexité : il se présente sous la forme d'un arbre *dyadique* (dont chaque branche se divise en deux, et ainsi de suite) à environ 23 niveaux de bifurcations, chaque génération étant constituée de conduits plus petits que la précédente,

les longueurs diminuant de 10 cm (trachée) à 1 mm (dernières bronchioles respiratoires). Cet arbre est entouré d'un tissu viscoélastique, le *parenchyme pulmonaire*, constitué entre autres d'élastane, de fibres de collagène et d'un réseau de vaisseaux sanguins.

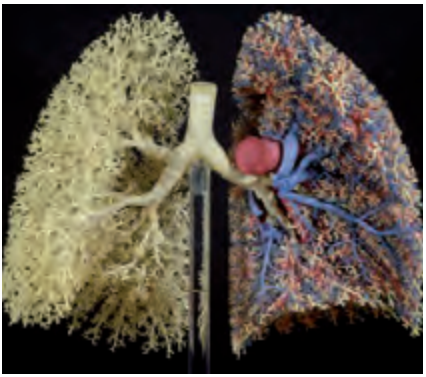


Figure 1. Moulage de poumon humain. Institut d'Anatomie, Université de Berne, Prof. Ewald R. Weibel.

La première partie de l'arbre bronchique, jusque vers la quinzième génération, est purement conductrice et a pour fonction principale le transport de l'air. Au-delà, des alvéoles viennent se greffer sur les bronchioles et permettent un échange de gaz (oxygène et dioxyde de carbone) avec les capillaires sanguins qui tapissent l'extérieur de quelque 300 millions d'alvéoles. Le mouvement de l'air dans ce réseau de tubes est assuré par le déplacement du diaphragme et de la cage thoracique, véritables moteurs de la respiration. Cela entraîne, à l'inspiration, l'augmentation du volume du poumon et induit une différence de pression entre la bouche (ou le nez) et les alvéoles pulmonaires, créant alors un flux d'air dans l'arbre (de ce point de vue, le poumon ressemble à un soufflet).

On ne sait pas forcément mesurer expérimentalement toutes les quantités souhaitées, et on ne peut pas non plus répéter les expériences à l'infini pour infirmer ou confirmer certaines hypothèses. C'est pourquoi il peut être utile de modéliser.

Décrire, comprendre le transport de l'air ou le dépôt de particules dans cette géométrie complexe, et dans toute sa généralité, peut s'avérer difficile. On ne sait pas forcément mesurer expérimentalement toutes les quantités souhaitées, et on ne peut pas non plus répéter les expériences à l'infini pour infirmer ou confirmer certaines hypothèses. C'est pourquoi il peut être utile de modéliser – c'est-à-dire ici de mettre en équations et simuler sur ordinateur – de manière simple mais représentative, certains phénomènes particuliers tels que les variations du volume global du poumon au cours des cycles respiratoires. On peut aussi chercher à décrire plus précisément une zone particulière, par exemple la partie supérieure (dite *proximale*) de l'arbre, afin d'obtenir une cartographie de l'écoulement de l'air dans cette région.

La modélisation dépendra bien sûr des besoins qui la motivent. Il pourra s'agir de comprendre un mécanisme global et l'effet de pathologies sur ce mécanisme, ou encore d'obtenir une description tridimensionnelle de l'écoulement pour, notamment, prédire dans quelles zones de l'arbre se déposeront des particules curatives ou polluantes.

Un modèle simplifié…

Reprenons l'image du soufflet. Imaginons que l'on cherche à reproduire les résultats d'une expérience de *spirométrie* (du latin *spiro*, respirer, et du grec *metron*, mesure : mesure de la respiration). Au cours de cette expérience, on demande au patient d'inspirer et d'expirer le plus intensément possible et on mesure le volume respiré, que l'on note V, et le débit associé (qui n'est autre que la vitesse de variation du volume au cours du temps), noté $\dot{V}$.

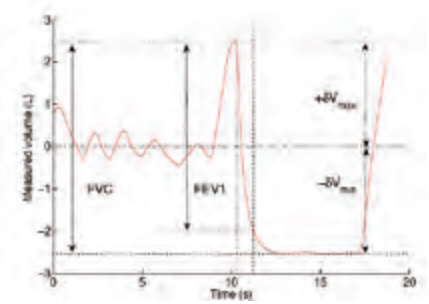


Figure 2. Evolution du volume respiré au cours du temps (d'après S. Martin, T. Similowski, C. Straus, B. Maury, ESAIM Proc., vol. 23, 2008, p. 30-47)

Imaginons que l'on veuille décrire le comportement global de la dynamique du volume respiré (autrement dit la manière dont ce volume respiré évolue au cours du temps). Le modèle mathématique le plus simple, très utilisé, auquel l'on puisse penser est ce que l'on appelle le modèle *monocompartiment*. Dans ce modèle, connu et utilisé depuis longtemps, le flux d'air est proportionnel à la différence entre la pression alvéolaire P_a et la pression à la bouche P_{atm}. La constante de proportionnalité correspond à la résistance R de l'arbre bronchique à l'écoulement (la résistance dépend des dimensions du tube et de la viscosité du fluide, en particulier, plus le tube est long et fin, plus il est difficile de faire passer l'air dedans). On peut exprimer cela par une équation, appelée Loi de Poiseuille du nom du médecin français qui l'a établie en 1844, qui dit que :

$$P_{atm} - P_a = R\dot{V}$$

Par ailleurs le volume du « ballon » dépend de l'élastance des tissus (notée E) et des forces qui lui sont appliquées (ici la pression alvéolaire P_a et la pression P exercée par le diaphragme). Cela s'exprime par une autre équation traduisant l'équilibre élastique du ballon :

$$P_a - P = EV$$

Finalement, on combinant nos deux équations, on obtient :

$$P_{atm} - P = R\dot{V} + EV$$

qui décrit l'évolution au cours du temps du volume du ballon.

Il s'agit d'une équation différentielle puisqu'elle fait intervenir à la fois le volume V et sa variation $\dot{V}$, linéaire, et on sait en calculer explicitement la solution.

Pour exploiter un tel modèle, si simple soit-il, il faut trouver les valeurs des paramètres

que sont la résistance des voies aériennes, l'élastance des tissus et la pression exercée.

Cette recherche des bons paramètres peut être faite « à la main » ou, si l'on sait mesurer la pression exercée, de manière systématique à l'aide de méthodes mathématiques afin que la solution soit le plus proche possible du volume mesuré.

Caler « à la main » consiste à essayer différentes valeurs des paramètres afin de reproduire au mieux les mesures de volume et débit de l'expérience. Si la pression exercée est mesurable, il est également possible de résoudre un problème de minimisation où l'on va chercher R et E minimisant l'écart entre la pression mesurée P_m et la pression P donnée par l'équation :

$$P = P_{atm} - R\dot{V} - EV$$

où V et $\dot{V}$ nous sont donnés par l'expérience.

Ainsi, une fois les paramètres « calés », en résolvant notre équation différentielle, on dispose d'une formule mathématique exprimant le volume en fonction du temps et de P_{atm}, P_a, P et E :

$$V(t) = V(0)e^{-\frac{E}{R}t} + \int_0^t \frac{P_{atm} - P(s)}{R} e^{\frac{E}{R}(s-t)} ds$$

… qui peut être enrichi !

Même si on ne s'intéresse qu'à une dynamique globale du volume, ce modèle simplifié ne rend pas compte de tous les phénomènes pouvant avoir un impact sur ce que l'on cherche à mesurer. En particulier, l'expansion du poumon est limitée par la cage thoracique, l'arbre bronchique se déforme au cours des cycles respiratoires, la résistance à l'écoulement augmente avec le débit… Il est donc nécessaire d'enrichir le modèle.

On peut imaginer que la résistance R et l'élastance E dépendent du volume V et du débit $\dot{V}$. Mais en tenant compte de cette dépendance, on obtient finalement une équation dont on ne sait plus calculer explicitement la solution. Toutefois, on sait dire si cette solution existe et il existe des méthodes permettant d'en calculer une approximation. Le travail du mathématicien appliqué est donc de comprendre le comportement des solutions et de calculer au cours du temps et avec l'aide de l'ordinateur, le volume respiré.

Si la solution du modèle mathématique et numérique est proche des expériences spirométriques, alors c'est un « bon » modèle et l'on peut ensuite « jouer » avec les paramètres et les faire varier afin de reproduire ou de comprendre les phénomènes en jeu. La résistance R augmente lors d'une crise d'asthme, l'élastance E des tissus diminue en cas d'emphysème. On peut alors apporter un éclairage différent de celui des mesures expérimentales et contribuer à répondre aux questions telles que :

Comment l'emphysème ou une crise d'asthme affectent globalement la ventilation ?

Bien sûr, une telle modélisation ne peut pas rendre compte de l'écoulement de l'air dans toute sa complexité et ne permet pas de connaître les vitesses de l'air dans la trachée ou dans des petites bronches. Néanmoins, si l'on cherche à décrire le comportement de l'air en chaque point de l'arbre bronchique, la démarche du mathématicien appliqué sera la même : déterminer un système d'équations capables de rendre compte le plus fidèlement possible de la réalité, comprendre ces équations et les propriétés des solutions à l'aide d'outils mathématiques, trouver des stratégies – puisqu'on ne sait pas calculer les solutions exactement – pour en déterminer des approximations numériques, exploiter ces calculs et les comparer aux mesures expérimentales disponibles pour valider ou invalider le modèle…

Avec plusieurs niveaux de description

Supposons donc que l'on cherche à décrire l'écoulement de l'air dans l'arbre bronchique, non plus uniquement grâce au volume respiré, mais de manière plus précise. Par exemple, si on veut déterminer la vitesse et la pression du fluide en chaque point de cet arbre au cours d'un cycle respiratoire. Compte tenu de la complexité de la géométrie, les calculs ne sont, à l'heure actuelle, pas envisageables dans la totalité de l'arbre. On peut donc dans un premier temps se restreindre à la partie supérieure de l'arbre respiratoire (jusqu'à la dixième génération). Appelons cette zone Ω.

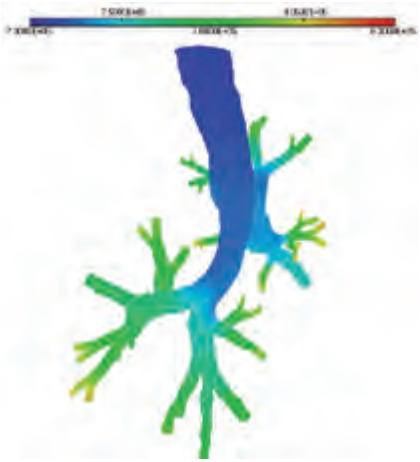


Figure 3. Zone Ω

Dans Ω, il est raisonnable de supposer que la vitesse et la pression de l'air vérifient des équations bien connues en mécanique des fluides : les équations de Navier-Stokes (voir encadré p. 49). Cependant, cette vitesse et cette pression dépendent fortement de ce qui se passe au delà de la dixième génération, et en particulier du mouvement du diaphragme et de l'élasticité des tissus pulmonaires. Il faut donc trouver un moyen de décrire mathématiquement cette partie de l'appareil respiratoire et de la coupler avec les équations de Navier-Stokes. On peut s'inspirer du modèle simple précédent. On obtient alors un système d'équations avec deux niveaux de description : une des-

cription fine de Ω et une plus grossière au delà. Il s'agit ensuite d'étudier le système obtenu : existe-t-il une solution et quelles sont ses propriétés mathématiques ? Comment calculer efficacement des vitesses et des pressions approchées ? Ce système représente-il en partie la réalité ?

Les réponses à la première question permettent de donner un cadre, de mieux comprendre le domaine de validité du modèle. Les difficultés ici sont que l'on est en présence d'équations *non-linéaires*, couplant différentes échelles de description. De plus, on regarde l'écoulement dans une région donnée et on « coupe » artificiellement ce qui ne nous intéresse pas. Il s'agit donc d'un système ouvert dans lequel de la matière et de l'énergie (cinétique) peuvent rentrer. Un des problèmes mathématiques importants auquel on est alors confronté est d'estimer cette énergie.

Toutes ces étapes posent de passionnantes questions qui ont leur intérêt mathématique propre.

Cette problématique est également cruciale quand on cherche à répondre à la deuxième question, sur le calcul numérique des vitesses et des pressions. Il s'agit d'un calcul approché car, ici encore, on ne sait pas exhiber de solutions explicites (c'est-à-dire écrire la vitesse en chaque point et en tout temps comme pour le volume dans la formule p. 46). Pour cela, on va tâcher de décrire le domaine de calcul Ω, qui est obtenu à partir de l'imagerie médicale et est propre à chaque patient, à l'aide d'un nombre fini de points : des *nœuds* reliés par des *arêtes*. C'est ce que l'on appelle un *maillage*. On va prendre l'image issue d'une IRM ou de scanner et on va la « segmenter » et la « mailler », ce qui nécessite ici encore des développements mathématiques importants, tout particulièrement pour ces géométries complexes, courbes avec plusieurs embranchements telles celle de l'arbre bronchique.

Figure 4. Maillage de la zone Ω

C'est en ces nœuds que l'on va ensuite déterminer la vitesse et la pression du fluide.

Enfin, calculer numériquement les valeurs approchées des vitesses et pressions en ces points de l'espace et en un nombre finis d'instants demande la conception de méthodes spécifiques et adaptées au modèle, et qui par exemple traitent efficacement le fait que le système soit « ouvert ».

Toutes ces étapes posent de passionnantes questions qui ont leur intérêt mathématique propre et qui font aujourd'hui l'objet de recherches actives. Malgré tout, ces modélisations ne peuvent se faire sans comparer les résultats des simulations à des expériences. Les mathématiciens interagissent avec les spécialistes de la respiration, qui leur disent si les modèles élaborés sont pertinents, et comment ils peuvent être améliorés. L'objectif à terme est d'obtenir des outils de prédiction fiables et d'apporter un éclairage complémentaire à ce que l'on obtient via les expériences *in vivo* ou *in vitro*.

Les équations de Navier-Stokes

Les équations de Navier-Stokes utilisées ici décrivent l'écoulement d'un fluide visqueux incompressible, et relient donc la vitesse u de ce fluide à sa densité ρ et à la pression p. Elles s'écrivent

$$\begin{cases} \rho\partial_t u + \rho(u\cdot\nabla)u - \nu\Delta u + \nabla p = 0 \\ \nabla\cdot u = 0 \end{cases}$$

La première équation exprime le principe fondamental de la dynamique de Newton en l'absence de force extérieure, tandis que la seconde traduit l'incompressibilité du fluide, dont la viscosité est notée ν. Il faut, pour pouvoir résoudre ces équations numériquement, se donner la vitesse sur le bord du domaine où on les applique, ainsi que dans tout le domaine à l'instant initial.

Le temps qu'il fera

Claude Basdevant, *professeur à l'Université Paris 13 et à l'École polytechnique*

La prévision météorologique ou climatique n'est pas une mince affaire. Elle implique la modélisation de nombreux phénomènes de natures différentes et l'intervention de plusieurs sciences, des mathématiques à la biologie, en passant par l'informatique, la physique ou la chimie.

Derrière les petits nuages gris ou soleils radieux qui parsèment la carte de France au bulletin météo du soir, il y a longtemps qu'il n'y a plus de grenouille et de thermomètre, mais des ordinateurs super puissants, auxquels on a fait absorber un grand nombre de mesures (obtenues principalement par satellite), beaucoup de lois de la mécanique et de la physique, mais aussi beaucoup de mathématiques, parfois très récentes.

Pour que les ordinateurs fournissent des prévisions, il faut élaborer au préalable ce qu'on appelle un *modèle numérique de prévision du temps*. Schématiquement, un tel modèle de prévision à l'échéance de huit à dix jours représente l'état de l'atmosphère à chaque instant par les valeurs des paramètres météorologiques (vitesse du vent, température, humidité, pression, nuages, aérosols, etc.) aux centres de « boîtes » d'environ deux à vingt kilomètres de côté et de quelques dizaines à quelques centaines de mètres de hauteur. Ce découpage imaginaire de toute l'atmosphère en boîtes est inévitable, car il est impossible de spécifier les paramètres météorologiques en tous les points de l'atmosphère (ces points sont en nombre infini !). En principe, plus les boîtes sont petites – et donc nombreuses –, plus la description de l'état atmosphérique est précise, et plus les prévisions le seront aussi. Mais en pratique, les boîtes ne font pas moins de quelques kilomètres pour une prévision régionale à courte échéance, et quelques dizaines de kilomètres pour une prévision à dix jours qui est nécessairement sur toute la planète ; en deçà, la puissance

des plus gros ordinateurs ne suffirait pas : il faut bien que le calcul s'achève en temps utile, c'est-à-dire en nettement moins de 24 heures !

Partant de l'état de l'atmosphère supposé connu au début de la période à prévoir, le modèle fait calculer par l'ordinateur son évolution future en utilisant les lois de la dynamique et de la physique. L'évolution dans le temps est calculée pas à pas, par intervalles de quelques minutes. Tel est le principe de la prévision numérique du temps, un principe connu depuis le début du XX^e siècle mais qui a attendu les années 1940-1950 et les premiers ordinateurs avant d'être mis en œuvre.

Des mesures inexploitables directement

Premier problème dans le schéma idéal de prévision qui vient d'être décrit : savoir construire l'*état initial de l'atmosphère*. Les observations sont loin d'être bien adaptées à cet exercice. Les stations météo au sol sont fort mal réparties sur le globe et fournissent très peu de mesures en altitude. Quant aux satellites, ils sont pour la plupart à défilement, c'est-à-dire qu'ils balayent continûment la Terre. Leurs mesures ne sont donc pas obtenues au même instant en tous points. De plus, les satellites mesurent des quantités intégrées sur toute l'épaisseur de l'atmosphère (il s'agit en général des flux d'énergie reçus dans une certaine gamme de longueurs d'onde) et non pas les grandeurs météorologiques (vent, température, humidité, etc.) qui entrent en jeu dans les équations des modèles.

On dispose donc d'une masse de données disparates, mal distribuées à la surface du globe, étalées sur 24 heures, connues avec des incertitudes variables, avec lesquelles il faut *initialiser* une prévision, c'est-à-dire construire un état météorologique initial dont le modèle simulera l'évolution. Or grâce aux travaux sur l'*optimisation dynamique*, domaine auquel ont beaucoup contribué le chercheur russe Lev Pontriaguine (1908-1988) et l'école mathématique française, on a pu mettre au point, dans les années 1980, des méthodes dites d'*assimilation variationnelle* qui permettent de reconstruire de façon optimale l'état initial. L'idée sous-jacente à ces méthodes, opérationnelles depuis l'année 2000 à Météo-France, est d'obliger en quelque sorte la trajectoire du modèle numérique à passer « près » des données observées pendant les 24 heures précédentes tout en préservant l'équilibre du modèle. L'assimilation variationnelle n'est d'ailleurs pas la seule technique mathématique moderne qui a bouleversé le traitement des observations : l'utilisation des réseaux neuromimétiques ou des ondelettes, inventés dans les années 1980, a donné lieu à des gains spectaculaires en efficacité, précision et rapidité dans le traitement des données fournies par les satellites.

On dispose donc d'une masse de données disparates, mal distribuées à la surface du globe, étalées sur 24 heures, connues avec des incertitudes variables.

Quand l'analyse numérique entre en action

Une fois connu l'état atmosphérique initial dont a besoin le modèle numérique de prévision, reste à écrire le programme informatique capable de calculer le temps futur à partir de cet état initial et des lois de la physique. Ces lois reposent sur une description continue de l'espace et du temps. Or notre modèle numérique, lui, ne connaît qu'un nombre, certes grand, mais fini, de boîtes. De plus, les intervalles de temps entre deux états calculés sont de plusieurs minutes – on dit que le problème a été *discrétisé*. Passer des équations continues à des schémas numériques pour le modèle discrétisé, tout en gardant la meilleure précision possible, tel est le domaine de l'analyse numérique, une branche des mathématiques qui a explosé depuis l'arrivée des ordinateurs. L'analyse numérique a pour but de savoir résoudre des équations et mener les calculs jusqu'au bout, c'est-à-dire jusqu'à l'obtention de valeurs numériques précises, en investissant le moins de temps et d'efforts possible. Elle est indispensable pour que simulation ne soit pas synonyme de simulacre et pour évaluer l'incertitude des prévisions. Par exemple, des progrès très importants ont été obtenus récemment concernant les méthodes permettant de simuler le déplacement des espèces chimiques ou des particules dans la turbulence atmosphérique. Ces avancées ont significativement amélioré l'étude et la prévision de la pollution de l'air.

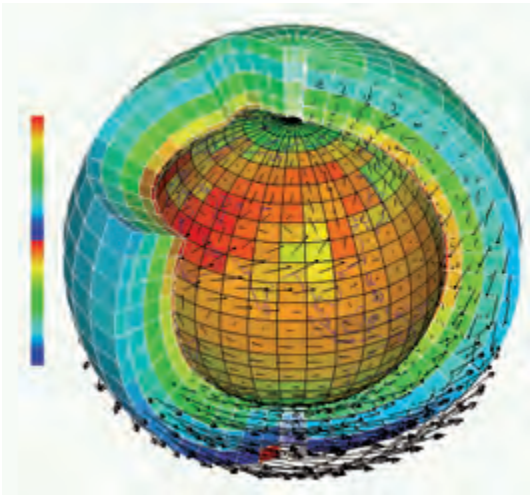


Figure 1. Vue d'artiste des boîtes de calcul d'un modèle de prévision du temps ou du climat. (L. Fairhead LMD/CNRS)

Les prévisions météo sont-elles fiables ?

Un des grands défis de la prévision météorologique est de pouvoir estimer la qualité de ses prévisions ; sont-elles fiables à l'échéance de trois, quatre, voire dix jours ? Au-delà de l'intérêt pour tout un chacun voulant organiser ses loisirs ou son activité, cette question a des impacts humains ou économiques forts, que ce soit pour la prévision d'évènements extrêmes qui peuvent être dévastateurs ou pour la gestion de productions d'énergie ou de crèmes glacées ! Si les progrès constatés a posteriori sont très importants, il reste difficile d'estimer a priori la fiabilité des prévisions. Depuis quelques années, des outils issus du croisement entre la théorie des systèmes dynamiques et celle des probabilités ont permis de développer ce qui s'appelle la *prévision d'ensemble* : celle-ci consiste à ne plus faire une prévision à partir d'un seul état initial mais à faire en parallèle un grand nombre de prévisions (50 en général) à partir d'état initiaux perturbés. La théorie des systèmes dynamiques permet de choisir les perturbations initiales les plus sensibles, alors que la théorie des probabilités permet d'extraire de la divergence des simulations une information sur la fiabilité de celles-ci et de calculer la probabilité qu'un événement précis se produise.

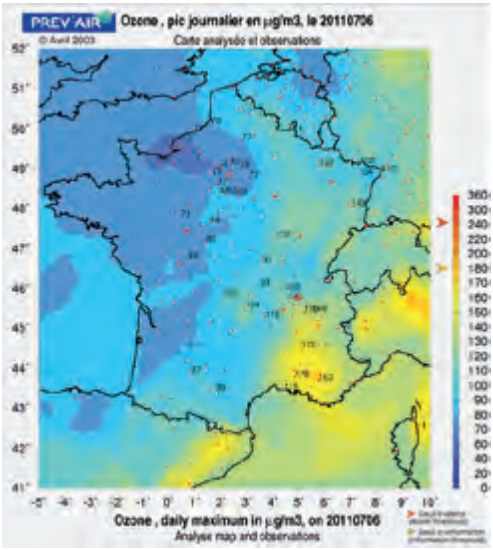


Figure 2. Concentrations d'ozone en surface le 6 juillet 2011 données par le système de prévision national de la qualité de l'air PREVAIR (INERIS/IPSL/Météo-France). En couleurs, les concentrations prévues par le modèle de chimie-transport CHIMERE ; les valeurs numériques sont les pics de concentration observés dans la journée par le réseau de surveillance.

Peut-on prédire le temps longtemps à l'avance ? Non, indique la théorie des systèmes dynamiques.

On a évoqué jusqu'ici la prévision du temps à courte échéance, de huit à dix jours. Mais pourquoi ne fait-on pas des prévisions à plus longue échéance ? Le météorologue américain Edward N. Lorenz, dans un célèbre article de 1963, a montré que c'était probablement sans espoir. L'atmosphère est un système chaotique, c'est-à-dire que toute erreur sur l'état météorologique initial, aussi petite soit-elle, s'amplifie rapidement au cours du temps ; si rapidement qu'une prévision à l'échéance d'une dizaine de jours perd toute sa pertinence. Néanmoins, cela ne veut pas dire que l'on ne peut pas prévoir le climat – c'est-à-dire faire une prévision de type statistique plutôt que déterministe, s'intéresser à la moyenne des températures ou des précipitations sur une période et une région, plutôt qu'au temps précis qu'il fera sur Quimperlé en Bretagne tel jour du mois de juillet. L'enjeu est d'importance : notre climat futur est menacé par les rejets de gaz dus aux activités humaines et il faut prévoir l'effet à long terme de ces perturbations, conséquences de l'effet de serre. C'est la théorie des systèmes dynamiques qui donne des outils pour justifier cette modélisation du climat. Ce domaine, pour lequel le mathématicien Henri Poincaré, au début du XX^e^ siècle, fut un grand précurseur, a connu des progrès très importants dans les dernières années du XX^e^ siècle. La théorie des systèmes dynamiques permet par exemple de dégager ce que les mathématiciens appellent des *attracteurs*, ou des *régimes de temps* pour les météorologues. Elle permet aussi de savoir quels sont les régimes de temps les plus prévisibles et ceux qui sont les plus instables. Dans les situations d'instabilité, un bon outil serait la modélisation probabiliste du climat, c'est-à-dire la conception de modèles prenant explicitement en compte le caractère aléatoire de la prévision. Encore peu développées, les modélisations de ce type doivent s'appuyer sur des outils très récents de la théorie des équations aux dérivées partielles stochastiques et des statistiques.

Des prévisions météorologiques aux prévisions climatiques

Les modèles numériques de prévision du climat ressemblent comme des frères aux modèles de prévision du temps, à deux différences essentielles près. Pour des raisons de temps de calcul, leurs « boîtes » sont nécessairement plus grandes (100 à 300 km de côté) ; les temps simulés allant de quelques mois à des centaines voire des milliers d'années, il est impossible d'être plus précis. Mais la différence importante tient au fait que les variations climatiques ont lieu à de longues échelles de temps, et qu'il n'est alors plus possible de négliger dans les interactions entre l'atmosphère, l'océan, les glaces de mer, la biosphère les évolutions de ces autres domaines. C'est pourquoi un modèle de climat doit combiner un modèle d'atmosphère, un modèle d'océan, un modèle de glaces de mer, un modèle de biosphère et tous les cycles d'échanges d'énergie, de quantité de mou-

vement, d'humidité, de gaz, ainsi que les couplages chimiques entre ces milieux. Au-delà de la complexité informatique d'une telle construction, se posent de délicats problèmes mathématiques sur la bonne manière de coupler ces domaines et sur la spécification des conditions aux interfaces atmosphère-océan, océan-glaces, etc. Et pour que le calcul dans les « grandes boîtes » reste significatif, il faut évaluer l'effet statistique, à l'échelle de ces boîtes, de processus qui se produisent à des échelles beaucoup plus petites (par exemple : quel est l'effet statistique, sur le bilan d'énergie d'une boîte de 300 km de côté, des petits cumulus de quelques km de diamètre qui s'y développent ?). Il reste, dans toutes ces questions, encore beaucoup de matière à développements interdisciplinaires dans lesquels les mathématiques tiennent une place importante.

Bibliographie

Jouzel J., Debroise A., (2004). *Le climat : jeu dangereux, quelques prévisions pour les siècles à venir*, éd. Dunod.

Le Treut H., Jancovici J.-M., (2004). *L'effet de serre : allons-nous changer le climat ?*, éd. Flammarion.

Sadourny R., (2002). *Le climat est-il devenu fou ?*, éd. Le Pommier.

–, (2003). *Peut-on croire à la météo ?*, éd. Le Pommier.

–, (2005). *D'où viennent les tempêtes ?*, éd. Le Pommier

Temam R., Wang S., (2000). *Mathematical Problems in Meteorology and Oceanography*, Bull. Amer. Meteor. Soc., 81, p. 319-321.

Autres références possibles :

Joussaume S., (2000). *Le Climat : d'hier à aujourd'hui*, CNRS Editions.

Hauglustaine D., Jouzel J., Le Treut H., (2004). *Climat : chronique d'un bouleversement annoncé*, éd. Le Pommier.

Internet, feux de forêt et porosité : trouvez le point commun

Marie Théret, *maître de conférences à l'Université Paris Diderot*

Echanges de données entre internautes, propagation d'un feu de forêt, infiltration de l'eau dans une roche : un modèle mathématique simple utilisant des graphes permet de mieux comprendre ces phénomènes.

Quel est le point commun entre une forêt, un réseau de communication et une roche poreuse ? Pour préserver les forêts, il est important de comprendre comment se propage un feu, ou une maladie, d'arbre en arbre. Pour améliorer les réseaux de communication, il faut savoir comment les données circulent entre les différents utilisateurs. Pour étudier la porosité d'une roche, il faut décrire la façon dont l'eau peut s'écouler à l'intérieur. Ces trois problèmes très concrets, bien qu'ayant des formulations différentes dans le langage courant, peuvent se représenter par un même modèle mathématique : le modèle de percolation.

Propagation d'une maladie dans une forêt

Concentrons-nous sur l'exemple de la forêt pour décrire ce modèle. On souhaite étudier la propagation d'une maladie : la contamination causée par un arbre malade va-t-elle rester circonscrite dans une petite région, ou risque-t-elle de s'étendre à des arbres situés très loin de la zone de contamination initiale ? L'approche que l'on adopte est celle de la physique statistique : pour comprendre un phénomène macroscopique, c'est-à-dire à grande échelle, on s'intéresse aux éléments microscopiques qui composent le système. Le système étudié est la forêt tout

entière, constituée d'un très grand nombre d'éléments « microscopiques » (c'est-à-dire ici de très petite taille par rapport à celle de la forêt), les arbres.

Il faut commencer par décrire comment se fait la contagion d'un arbre à l'autre. On suppose que la contamination directe ne peut avoir lieu qu'entre arbres assez proches, et qu'elle a un certain risque de se produire qui dépend uniquement de la virulence de la maladie (et non pas des caractéristiques individuelles des arbres, ni de la zone où ils se trouvent). Ces hypothèses sont à la fois crédibles du point de vue biologique, et suffisamment simplificatrices pour permettre un traitement mathématique efficace.

On forme un graphe dont chaque arbre est un sommet, et où l'on relie deux arbres par une arête s'ils sont suffisamment proches pour se contaminer l'un l'autre.

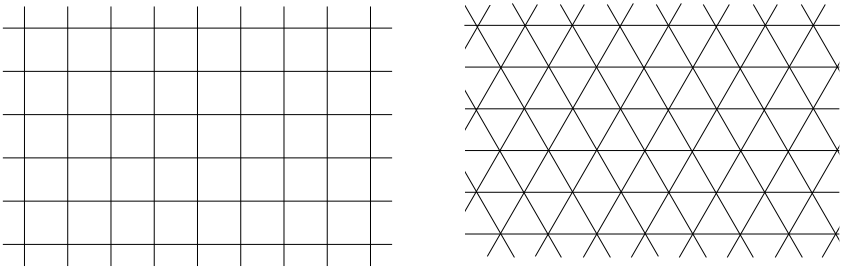


Figure 1. Deux graphes très utilisés en théorie de la percolation : le réseau carré et le réseau triangulaire

Le graphe obtenu décrit comment la maladie pourrait se propager dans le pire scénario, si chaque arbre infectait tous ses voisins. Mais on veut tenir compte du fait que la transmission de la maladie ne se produit que de manière aléatoire. Il y a donc une certaine probabilité de contamination entre voisins, et on colorie une arête en rouge quand la contamination a lieu.

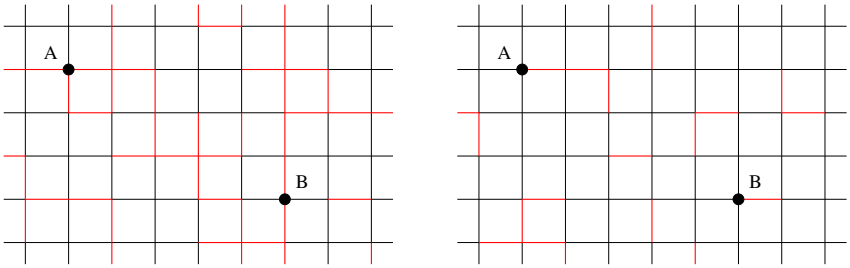


Figure 2. Les arêtes sont coloriées aléatoirement, indépendamment les unes des autres, lorsque la contamination entre deux arbres voisins a lieu. Cette contamination est supposée instantanée : en particulier, l'ordre dans lequel on colorie les arêtes n'a pas d'importance. À gauche : L'arbre A contamine l'arbre B, ainsi que d'autres arbres situés plus loin. Ceci se produit typiquement quand la probabilité de transmission est élevée. À droite : L'arbre A ne contamine que 3 autres arbres. Il ne contamine pas B. Ceci se produit typiquement quand la probabilité de transmission est faible.

Ce que l'on vient de décrire est le modèle de percolation par arête : partant d'un graphe donné, on colorie aléatoirement chaque arête avec une certaine probabilité, indépendamment les unes des autres.

Notons que ce modèle simplificateur ne rend pas compte de l'aspect temporel de la propagation de la maladie : toutes les contaminations entre voisins ont lieu simultanément, instantanément, et indépendamment les unes des autres. Les arêtes non coloriées correspondent aux couples d'arbres entre lesquels la maladie ne se transmet pas.

Ce que l'on vient de décrire est le modèle de percolation par arête. Pour traduire la différence d'échelle entre l'arbre et la forêt, on peut considérer qu'il y a une infinité d'arbres

dans la forêt, c'est-à-dire que le graphe à étudier est infini. Les questions que l'on se pose sur la transmission de la maladie se reformulent dans ce contexte : pour savoir si l'épidémie causée par un arbre malade va rester circonscrite, ou au contraire si elle risque de toucher des arbres très lointains, il faut regarder quels sommets du graphe sont reliés par un chemin de couleur rouge au site correspondant à l'arbre malade initialement. On se demande en particulier si ces arbres contaminés sont en nombre fini ou infini. Notre modèle dépend d'un unique paramètre : la probabilité de contamination entre voisins, et il s'agit d'étudier, selon les valeurs de ce paramètre, les propriétés du graphe aléatoire décrit ci-dessus.

Ce modèle est également pertinent pour décrire un réseau de communication ou une roche poreuse. Dans le premier cas, les arêtes du graphe sont les câbles du réseau, qui sont en bon état (arêtes coloriées) ou déficients (arêtes non coloriées), et on peut étudier à qui est transmise une information envoyée par un utilisateur. Dans le deuxième cas, les arêtes du graphe sont des tubes microscopiques dans la roche, qui peuvent laisser passer l'eau mais aussi se boucher. On s'intéresse alors à la zone de la roche qui est mouillée si l'eau s'infiltre à tel ou tel endroit. Le terme « percoler » signifie, pour un liquide, passer à travers des matériaux poreux : c'est donc cette interprétation du modèle en terme de roche poreuse qui a donné son nom au modèle mathématique de percolation, et c'est effectivement pour modéliser un milieu poreux que ce modèle a été introduit par Broadbent et Hammersley en 1957.

Un modèle simple mais riche

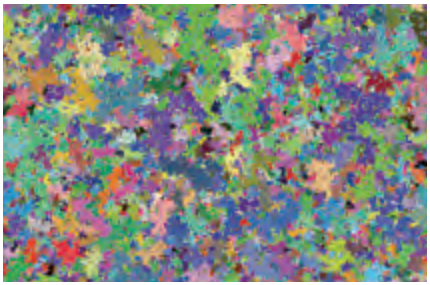
Le modèle de percolation est très simple mais aussi très riche.
D'une part, parce qu'il présente une transition de phase, c'est-à-dire un changement brutal de comportement quand le paramètre du système – la probabilité avec laquelle une arête du graphe est coloriée - dépasse une certaine valeur critique. En effet, si chaque arête est coloriée avec une probabilité suffisamment grande, la transmission des maladies se fait à l'échelle du système tout entier, c'est-à-dire qu'un arbre malade peut infecter d'autres arbres arbitrairement éloignés de lui. Mathématiquement, ceci s'exprime en disant qu'« il existe une composante connexe infinie dans le graphe aléatoire ». Mais dès que la probabilité de transmission passe au dessous d'un certain seuil, la transmission des maladies reste locale, c'est-à-dire que l'infection est limitée à de petits domaines de la forêt : « il n'existe pas de composante connexe infinie » (voir les figures de la page suivante). Ce théorème a été démontré par Broadbent et Hammersley.

D'autre part, parce que malgré sa simplicité, il soulève de nombreuses questions qui restent encore sans réponses, à commencer par la plus évidente : comment calculer cette valeur critique ? De grandes avancées dans l'étude de la percolation ont eu lieu ces dernières années, notamment grâce à l'utilisation d'outils issus de l'analyse complexe, autre branche des mathématiques (citons notamment les travaux de Lawler, Schramm et Werner dans les années 2000), et la recherche continue.

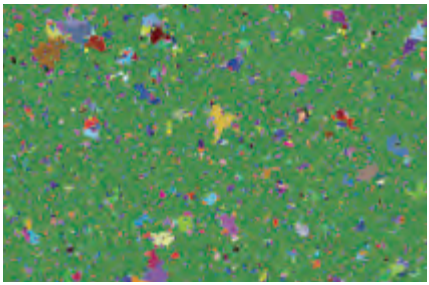
Malgré sa simplicité, ce modèle soulève de nombreuses questions qui restent encore sans réponses, à commencer par la plus évidente : comment calculer la valeur critique ?

Pour une étude plus précise de ces phénomènes, Hammersley et Welsh ont introduit en 1965 un modèle plus complexe, le modèle de percolation de premier passage. Au lieu de colorier aléatoirement les arêtes du graphe, ils leur associent un temps aléa-

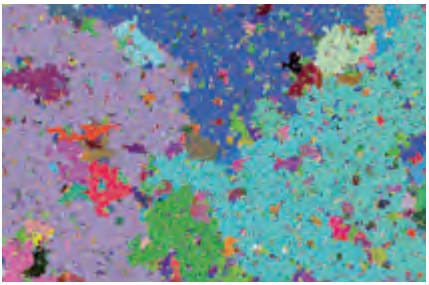
(a)



(b)



(c)



Figue 3. Phénomène de transition de phase (simulations informatiques) : on regarde le système à grande échelle, chaque pixel représente ici un arbre de la foret. Un arbre contamine ceux qui sont situés dans la même tache de couleur que lui. (a) : La probabilité de transmission vaut ici 0,4. Les taches de couleur sont de petite taille : un arbre malade ne contamine qu'une petite région autour de lui. (b) : p=0,506, légèrement au dessus de la valeur critique. La tache verte est de taille infinie : un arbre peut contaminer des arbres situés arbitrairement loin de lui. (c) : p=0,5, régime critique, c'est la transition entre les deux autres, le régime le moins bien connu. Les taches sont de tailles finies, mais très grosses. Simulations numériques par R. Cerf.

toire (le temps nécessaire au feu pour se propager entre deux arbres ou à l'eau pour traverser un tuyau) ou une capacité aléatoire (la quantité maximale d'information qui peut traverser un câble par seconde). Les questions soulevées sont alors encore plus nombreuses : combien de temps faut-il pour que de l'eau se propage à travers une couche de roche poreuse ? Quels sont les arbres touchés par un feu de forêt dans ses premières vingt-quatre heures ? Quelle est la quantité maximale d'information qu'un utilisateur d'un réseau peut envoyer à un destinataire par seconde ? De nombreux théorèmes ont déjà été obtenus (citons les résultats présentés dans les années 80 par Kesten, et plus récemment les travaux de Benjamini, Kalai et Schramm en 2003 et Chatterjee et Dey en 2009), mais il reste encore beaucoup de défis à relever !

Pour aller plus loin

Duminil-Copin H., *La Percolation, jeu de pavages aléatoires*, http://images.math.cnrs.fr/La-percolation-jeu-de-pavages.html

Grimmett G., (1989). *Percolation*. Springer-Verlag.

A la recherche de la forme idéale

Grégoire Allaire, *professeur à l'École polytechnique*
François Jouve, *professeur à l'Université Paris Diderot*

Les objets issus de la fabrication industrielle sont pensés de façon à optimiser un certain nombre de paramètres comme le poids ou la solidité. Pour éviter de chercher à tâtons la meilleure forme possible, on peut aujourd'hui compter sur plusieurs méthodes mathématiques d'optimisation.

Nos sociétés modernes sont éprises de *design*, ce mot anglais sans équivalent en français, qui traduit notre volonté d'allier le beau à l'utile. Le grand public connaît bien les designers célèbres et médiatiques comme Pininfarina ou Starck, mais beaucoup moins les scientifiques, ingénieurs ou chercheurs, qui font de la *conception optimale* (*optimal design* en anglais) : loin de toute préoccupation esthétique, ils améliorent les formes des objets industriels (structure mécanique, profil aérodynamique, composants électroniques, etc.) afin d'en augmenter les performances (solidité, efficacité) tout en tenant compte de contraintes, parfois contradictoires, comme leur poids ou leur coût. Il est clair par exemple que la solidité d'une structure varie à l'inverse de son poids (ce qui est lourd est plus solide que ce qui est léger). Ainsi, l'optimisation de la robustesse d'un avion est limitée par la contrainte d'une consommation minimale de carburant, qui est directement liée au poids. Un problème classique en mathématique consiste justement à chercher la *solution optimale* d'un problème d'optimisation d'une fonction (appelée *fonction objectif*) tout en respectant des contraintes.

Il est clair que la solidité d'une structure varie à l'inverse de son poids (ce qui est lourd est plus solide que ce qui est léger).

La méthode traditionnelle d'optimisation procède par essais et erreurs, suivant le savoir faire et l'intuition de l'ingénieur : on choisit une forme dont on calcule la perfor-

mance puis, en fonction de cette dernière, on la modifie pour essayer de l'améliorer et on recommence jusqu'à obtention d'une forme satisfaisante (à défaut d'être optimale). Cette façon de faire « manuelle » est très lente, coûteuse et peu précise. Grâce au formidable développement de la puissance de calcul des ordinateurs, ainsi qu'au progrès mathématique, cet empirisme est de plus en plus souvent remplacé par des logiciels numériques qui automatisent ce processus d'optimisation.

Optimiser la géométrie avec la méthode d'Hadamard

Tout algorithme d'optimisation est *itératif*: on construit une nouvelle forme à partir d'une variation de la précédente. Puis on calcule la performance de cette nouvelle forme, que l'on compare à celle de la première. Enfin, si la performance de la structure s'avère améliorée, on retravaille à partir de la nouvelle forme.

Il faut deviner la bonne topologie à imposer à notre forme de départ, or cela est impossible dans la plupart des cas.

En 1907, le mathématicien Jacques Hadamard a proposé une méthode de variation de forme qui porte désormais son nom et qui, bien que d'origine théorique, s'applique en pratique pour simuler certains problèmes sur ordinateur. La méthode consiste, en partant d'une forme initiale, à déplacer ses bords petit à petit, sans en créer de nouveaux. Cette méthode modifie donc la *géométrie* de la forme initiale mais elle préserve sa *topologie*: en effet les formes

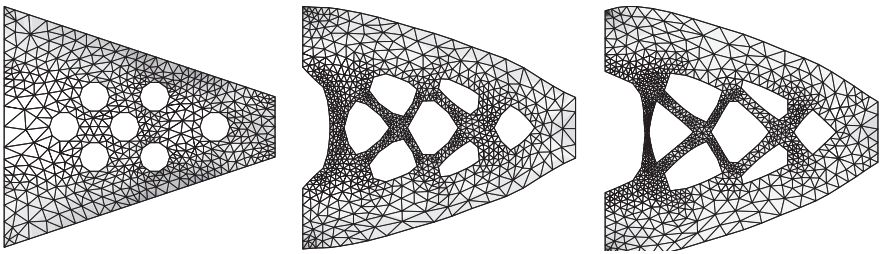


Figure 1. Initialisation (à gauche), itération intermédiaire (au centre) et forme optimale (à droite) d'une console, obtenues par la méthode d'Hadamard.

obtenues successivement conservent toujours le même nombre de trous, comme on peut le voir sur les illustrations représentant les résultats numériques de la méthode d'Hadamard.

Le fait que la topologie ne change pas constitue une limitation assez gênante. En effet, cela signifie qu'il faut deviner la bonne topologie à imposer à notre forme de départ, puisqu'on ne pourra pas la modifier pour améliorer la performance. Or, cela est impossible dans la plupart des cas. D'où la motivation des mathématiciens pour inventer d'autres méthodes capables d'optimiser également la topologie, c'est à dire le nombre de trous.

De plus, la méthode d'Hadamard présente aussi le désavantage d'être très coûteuse en temps de calcul.

De l'importance des matériaux composites

Dans les années 1990, les mathématiciens ont inventé une méthode d'optimisation topologique de forme, dite *méthode d'homogénéisation*, qui est désormais largement utilisée par les ingénieurs dans de nombreux logiciels industriels.

L'idée sous-jacente est de transformer un problème d'optimisation de formes en un problème d'optimisation d'une *densité de matière*.

En tout point de l'espace, la densité de matière est une valeur comprise entre 0 et 1. La valeur 0 correspond à un trou ou du vide (pas de matière), la valeur 1 correspond à du matériau plein et les valeurs intermédiaires (par exemple la valeur 0.5) correspondent à un matériau composite poreux, comme une éponge par exemple. Plus la valeur est proche de 0, et plus la proportion de trous dans le matériau est donc importante.

Ici, on remplace donc le problème original d'optimisation *discrète* du type 0 ou 1 (en chaque point de l'espace, on a soit du vide, soit de la matière), par un nouveau problème d'optimisation *continue*, où la variable à optimiser, la densité de matière, parcourt l'intervalle complet [0,1].

L'idée sous-jacente est de transformer un problème d'optimisation de formes en un problème d'optimisation d'une densité de matière.

Avec cette nouvelle approche, on n'est plus prisonnier de la paramétrisation des formes proposée par Hadamard : la topologie va pouvoir changer et des trous vont pouvoir apparaître ou disparaître au gré des variations de la densité.

Notons que la densité de matière ne suffit pas à caractériser complètement un matériau composite : pour une densité donnée, la forme des trous compte aussi pour évaluer les propriétés effectives du matériau. Par exemple, la rigidité équivalente d'un matériau composite ne sera pas la même pour une structure laminée que pour une structure en nid d'abeilles. La méthode d'homogénéisation se propose donc d'optimiser

non seulement la densité de matière mais aussi la microstructure (la forme des trous) du matériau composite.

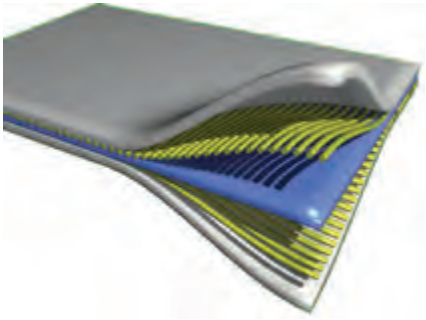
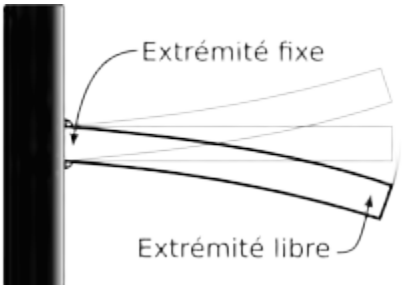


Figure 2. Exemple de matériau composite

Un autre avantage de la méthode d'homogénéisation est que sa résolution numérique est plus rapide : elle demande moins de calculs. L'une des raisons à cela est que pour changer la densité en un point donné, l'algorithme ne va utiliser que les valeurs de la densité autour de ce point, et non les valeurs des points plus éloignés, comme c'était le cas pour la méthode d'Hadamard.

La console optimale

Voici un problème d'optimisation de forme classique, celui de la *console optimale*. Ici, le bord de gauche de la console est fixé tandis qu'une force verticale est appliquée au milieu du bord de droite.



Nous traçons la densité de matière, représentée par un niveau de gris (le noir correspondant à du matériau plein, le blanc à du vide). La solution optimale présente de larges zones de gris, correspondant à du matériau composite, qu'il est difficile d'interpréter comme une forme. Pour retrouver une forme « classique » proche de la forme *composite* optimale, une solution est de « pénaliser » les matériaux composites à la fin de l'optimisation numérique, c'est-à-dire de rajouter des contraintes dans l'algorithme pour que la solution optimale soit plutôt composée de zones soit pleines, soit vides, que de zones composites.

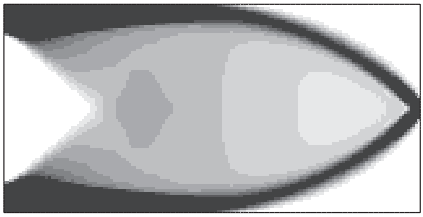


Figure 4. Console optimale composite (le niveau de gris indique la densité de matière).

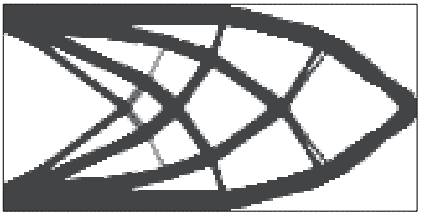


Figure 5. Console optimale après pénalisation et sans composites

Le résultat est impressionnant : la zone composite se transforme en un treillis de barres, qui rappelle de nombreuses structures en génie civil ou mécanique.

Ces méthodes d'optimisation géométrique et topologique de formes sont utilisées quotidiennement dans l'industrie automobile ou aéronautique, par exemple quand il s'agit de trouver la forme d'une structure qui soit à la fois rigide et légère. De nombreuses pièces mécaniques (triangles de suspension, longerons, etc.) dans les voitures ou les avions sont ainsi allégées par optimisation afin de réduire, in fine, la consommation de carburant.

Optimisation d'un pont.

Optimisation d'un pignon de roue.

Pour aller plus loin

Site web :
http://www.cmap.polytechnique.fr/~optopo

Allaire G., (2007). *Conception optimale de structures*, Collection Mathématiques et Applications, vol. 58, Springer Verlag.

Henrot A., Pierre M., (2005). *Variation et optimisation de formes*, Collection Mathématiques et Applications, vol. 48, Springer Verlag.

Hildebrandt S., Tromba A., (1986). *Mathématiques et formes optimales, Pour la Science*, Belin, Paris.

La biodiversité mise en équations… ou presque

Sylvie Méléard, *professeur à l'École polytechnique*

Prédire l'évolution d'une population animale sur une longue période, connaître le fonctionnement d'un écosystème, comprendre l'avantage de la reproduction sexuée pour la survie des espèces… les problèmes issus de la biodiversité sont complexes et leur résolution fait appel à des outils mathématiques sophistiqués.

"After years, I have deeply regretted that I did not proceed far enough at least to understand something of the great leading principles of mathematics: for men thus endowed seem to have an extra-sense".

(Les années passant, j'ai profondément regretté de ne pas être allé assez loin pour comprendre quelques aspects des principes essentiels des mathématiques: car les hommes dotés d'un tel entendement semblent posséder un sixième sens.)

Charles Darwin,
Autobiography

Il peut sembler étonnant d'associer les mathématiques, vues souvent comme une succession de formules figées, et la biodiversité, associée à la vie fourmillante des écosystèmes. Pourtant, la complexité extrême qui apparaît dès que l'on se penche sur la biodiversité justifie l'intérêt des mathématiques et des mathématiciens pour ce domaine. L'enjeu est de simplifier intelligemment les phénomènes pour pouvoir les mettre en équations et construire ainsi des modèles mathématiques novateurs. Le but est de pouvoir en déduire des quantités calculables telles que, par exemple, le temps d'extinction d'une espèce, l'estimation d'une abondance d'espèce, la vitesse d'invasion d'un parasite… ou encore de développer des algorithmes de simulation qui sont des outils d'expérimentation théo-

rique indispensables pour prédire certains comportements. C'est vrai en particulier pour les très grandes échelles de temps de l'évolution, pour lesquelles les observations n'existent pas. Les simulations sont aussi un support visuel qui facilite les échanges scientifiques avec les biologistes. La modélisation peut enfin servir à construire des outils statistiques permettant de prédire et quantifier différents scénarios de la biodiversité au vu des données observées.

Au XIXe et début du XXe siècle, de nombreux modèles, principalement déterministes, c'est-à-dire pour lesquels une même cause provoque toujours le même effet, ont été développés pour étudier la dynamique des populations (Malthus, Verhulst, Lotka, Volterra…). D'autres modèles plus probabilistes, c'est-à-dire faisant intervenir le hasard, ont également été proposés, pour étudier la génétique des populations (Fisher, Haldane, Wright…). Dans les dernières décennies, les biologistes se sont largement concentrés sur la biologie moléculaire et sur la formidable masse de données que les nouvelles technologies leur ont permis d'obtenir. Les méthodes de séquençage ont apporté beaucoup d'information qu'il a fallu trier, classifier, ordonner et analyser. Forts de ces données, les biologistes ressentent de nouveau la nécessité de créer des modèles pour leur donner une cohérence et en déduire des prédictions sur les comportements biologiques et écologiques.

La complexité des problèmes issus de la biodiversité nécessite donc des outils mathématiques sophistiqués. Réciproquement, les mathématiciens vont puiser dans ces problèmes une motivation et une inspiration pour créer de nouveaux modèles et théorèmes d'une grande richesse, qui permettront de comprendre certains phénomènes écologiques.

Quelques problématiques d'écologie

Voici quelques exemples de questions posées par les écologues et auxquelles les mathématiciens s'efforcent de donner des éléments de réponses :

– Quel est le comportement d'une population en temps long ? Va-t-elle se stabiliser ou s'éteindre ? Si elle s'éteint, quel sera le temps d'extinction ?

– Les espèces cohabitent et interagissent dans une compétition pour les ressources, dans un réseau hôte-parasite ou une chaîne alimentaire, ou au contraire coopèrent pour leur survie. Comment ces réseaux évoluent-ils ?

– Quel est l'impact de la migration et de la fragmentation de l'habitat sur la biodiversité ?

– Quel est l'avantage de la reproduction sexuée pour la biodiversité et la survie des espèces ? En effet, même si la reproduction sexuée apporte de la variation génétique par recombinaison, nourrir les mâles, rechercher des partenaires, s'exposer à la prédation pendant l'acte sexuel sont des arguments qui rendent plus favorable la reproduction asexuée.

– Quel est l'impact des variations de l'environnement sur la biodiversité ?

– Quel est le lien entre l'ADN des populations, leurs interactions et leur évolution ?

La notion d'espèce et les modèles probabilistes

Pour comprendre la biodiversité, il faut commencer par comprendre ce qu'est une espèce.

La notion d'espèce a évolué au cours du temps. Déjà, Maupertuis (1698-1759), mathématicien et naturaliste, soulignait l'apparition au hasard de mutants qui se développent et forment une nouvelle population : « Nous voyons paraître des races de chiens, de pigeons, de serins qui n'étaient point auparavant dans la nature. Ce n'ont été d'abord que des individus fortuits ; l'art et les générations répétées en ont fait des espèces ». Dans l'œuvre de Lamarck (1744-1829), l'image linéaire de la grande chaîne des êtres est remplacée par un « arbre buissonnant ». Finalement, dans son fameux ouvrage *On the origin of species by natural selection*, Darwin (1809-1882) introduit l'idée de sélection naturelle et d'arbre des espèces, avec disparition de certaines espèces par sélection des plus adaptées, et apparition de nouvelles, telles les différentes espèces de pinsons qu'il met en évidence sur les îles Galapagos.

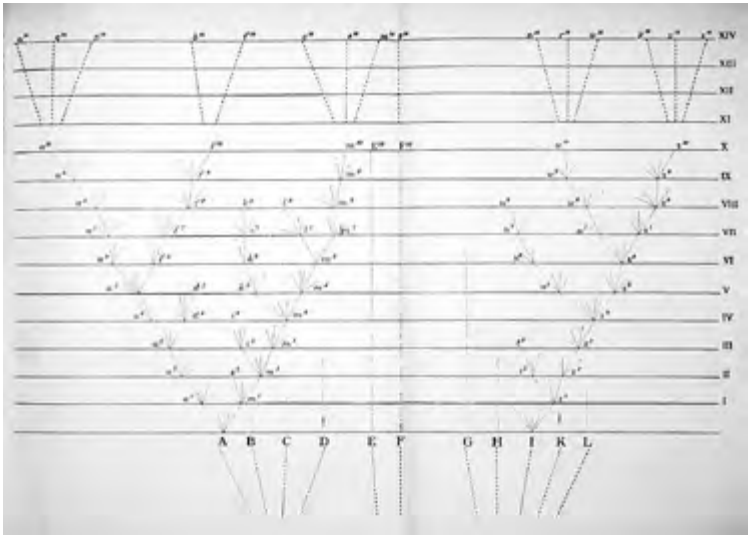


Figure 1. Arbre des espèces, avec disparition de certaines espèces et apparition de nouvelles, d'après C. Darwin, On the origin of species by means of natural sélection.

C'est cette dynamique de la biodiversité que l'on souhaite mieux appréhender, grâce à des modèles mathématiques et en particulier à des modèles probabilistes qui prennent en compte les naissances et les morts aléatoires des individus et que l'on appelle de manière générique *processus de branchement* ou *processus de naissance et mort*. Le prototype en est le processus de Galton-Watson (voir encadré). Il peut permettre de calculer la probabilité d'extinction d'une espèce.

Une démarche inverse, que le séquençage de l'ADN a beaucoup stimulée, consiste à reconstruire l'histoire de vie d'une population. Étant donné un groupe d'individus, à quelle époque se situait leur plus récent ancêtre commun ? Comment retrouver les lignées ancestrales des individus ? Cette démarche est celle des généticiens des populations et a conduit au modèle de Wright-Fisher (voir encadré).

Cette dualité entre la dynamique des populations au cours du temps et la reconstruction des généalogies permet ainsi d'estimer la biodiversité passée en fonction des relations de parenté des espèces actuelles.

Des modèles construits sur les comportements individuels et leurs interactions écologiques

Les modèles cités précédemment sont mathématiquement très simples. Même si les processus de Galton-Watson et les modèles de Wright-Fisher jouent un rôle fondamental – leur simplicité permettant de faire des calculs et de donner certaines réponses quantitatives – ils ne permettent pas de prendre en compte la diversité génétique et les interactions écologiques. Aussi, les recherches actuelles des biologistes théoriciens et des mathématiciens ont pour but d'introduire de la diversité génétique (ADN) et de l'interaction (compétition, coopération, prédation) dans les paramètres démographiques des modèles.

Les recherches actuelles ont pour but d'introduire de la diversité génétique et de l'interaction.

Les modèles les plus anciens introduisant de la compétition sont des modèles déterministes décrivant le comportement de grandes populations. Le plus célèbre est *l'équation logistique*, qui décrit la dynamique de la taille d'une population :

$$n'(t) = n(t)(b - cn(t))$$

Dans cette équation, le paramètre b décrit le taux de croissance de la population et c mesure la compétition entre deux individus pour le partage des ressources. Pour des temps t très grands, cette taille se stabilise sur la valeur limite b/c, appelée *capacité de charge*. Si la population comporte des individus de plusieurs types, elle est alors décrite par un système d'équations. Dans le modèle de proies-prédateurs par exemple, si $n_1(t)$ et $n_2(t)$ sont le nombre de proies et de prédateurs au temps t, alors on peut écrire les équations :

$$n_1'(t) = a\, n_1(t) - b\, n_1(t)\, n_2(t)$$

$$n_2'(t) = -c\, n_2(t) + d\, n_1(t)\, n_2(t)$$

On observe alors un comportement cyclique de ces dynamiques.

Dans le cas d'un continuum de types (la taille à la naissance, l'âge à la maturité…), la dynamique d'une grande population sera décrite par des équations plus compliquées (équations aux dérivées partielles ou intégro-différentielles) reliant les types des individus et le temps.

Ces modèles déterministes sont des modèles macroscopiques (c'est-à-dire qu'ils considèrent le comportement global d'un ensemble d'individus) et ne prennent pas en compte les comportements individuels, ni les naissances aléatoires d'individus mutants. Ils ne sont pas représentatifs des comportements des petites populations pour lesquels les fluctuations aléatoires sont essentielles. Ils ne peuvent pas non plus rendre compte de toutes les échelles de temps auxquelles l'on peut être confronté : par exemple le lien entre l'échelle écologique (les temps de naissance et de mort des individus) et celle de l'évolution (apparition et fixation des mutations). Pour pouvoir intégrer ces événements au modèle, il est important de revenir à une description du comportement aléatoire de chaque individu en prenant en compte ses propres caractéristiques. Les modèles probabilistes sont alors des généralisations du processus de Galton-Watson.

Exemple d'une population caractérisée par la taille des individus à la naissance

Dans cet exemple, les individus sont caractérisés par leur taille x à la naissance, qui varie par exemple entre 0 et 4. Le temps qu'un individu de taille x met à se reproduire est une variable aléatoire de moyenne $b(x)$. On peut supposer par exemple que $b(x) = 1/(4 - x)$, ce qui traduit le fait qu'un grand individu dépense beaucoup d'énergie à se nourrir et en a moins pour se reproduire.

A chaque reproduction, l'individu transmet héréditairement son type x, mais il arrive que son descendant mute en un individu de type y. Ce dernier va se reproduire et créer ensuite sa propre sous-population de type y. Les mutations introduisent ainsi de la variabilité génétique dans le modèle. Par ailleurs, la mort de l'individu dépend de ses propres paramètres génétiques, mais aussi de la compétition pour les ressources qu'il subit de la part de ses congénères. Par exemple, un individu de type $y > x$ consommera plus de ressources, au dépend de celui de trait x qui va ainsi voir ses possibilités de survie s'amoindrir.

La sélection naturelle, qui va permettre à la population de s'adapter et entrainer son évolution, est donc le résultat d'un compromis entre le fait de favoriser les petits individus pour accroître leurs capacités de reproduction et celui de favoriser les gros individus pour leur permettre d'être plus forts dans la compétition pour les ressources. C'est ce type d'optimisation évolutive que les mathématiciens cherchent à comprendre

ainsi que les mécanismes mathématiques qui peuvent expliquer l'apparition de nouvelles espèces.

Le modèle décrit ci-dessus peut expliquer la diversité des pinsons de Darwin à travers la taille de leur bec. La figure ci-dessous montre une simulation du modèle (voir figure 2) qui explique l'apparition de quatre espèces à partir d'une espèce unique, dans une très longue échelle de temps.

Il est bien sûr totalement illusoire de vouloir réduire la biodiversité à de simples équations, mais les modèles mathématiques peuvent fournir un nouvel angle d'étude et un point de vue dépassionné et objectif des écosystèmes.

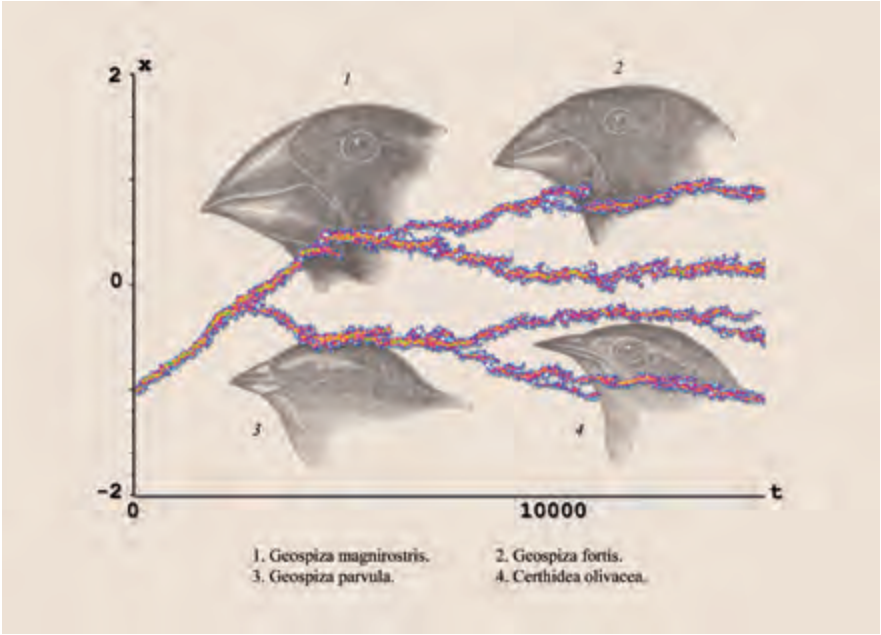


Figure 2. Simulation du modèle des pinsons de Darwin.

Le processus de Galton-Watson

Le processus de Galton-Watson est une suite de variables aléatoires (Z_n) décrivant l'évolution d'une population qui, à chaque génération, se reproduit aléatoirement. Si Z_n est la taille de la population à la $n^{\text{ième}}$ génération, alors $Z_{n+1} = X_1 + \ldots + X_{Zn}$, où X_i est le nombre d'enfants du $i^{\text{ème}}$ individu. Les variables X_i ont toutes la même loi X, et en particulier le nombre moyen m d'enfants par individu est constant. La probabilité d'extinction est alors obtenue comme limite d'une suite récurrente (dont le terme général donne la probabilité d'extinction à la génération n), et est solution de l'équation $g(p) = p$, où $g(s)$ est la moyenne de s^X. On peut montrer que si m est un nombre supérieur à 1, la population peut exploser exponentiellement, alors que si m est inférieur ou égal à 1, la population va s'éteindre. Par exemple, sachant que le nombre moyen des petits de baleines noires est estimé à 0,976 et que l'abondance des femelles était estimée à 150 en 1994, on peut montrer par ce modèle qu'il y a 99 % de chances pour que ces baleines aient totalement disparu en 2389.

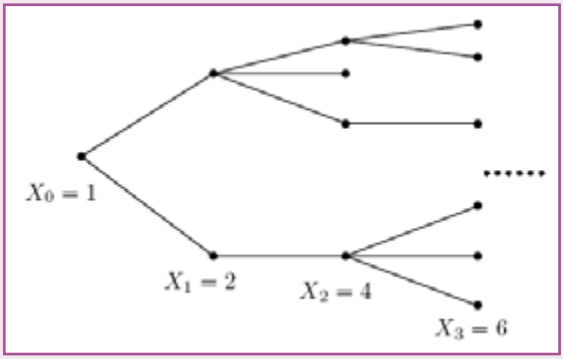


Figure 3. Une réalisation du processus de Galton-Watson.

Le modèle de Wright-Fisher

Dans ce modèle, la taille de la population N est constante et, à chaque génération, chaque individu choisit son ancêtre uniformément au hasard dans la génération précédente. Ainsi la probabilité que deux individus aient le même parent est $1/N$, qui tend vers 0 quand N tend vers l'infini. Il faut donc changer l'échelle de temps pour observer les généalogies pour une grande population et prendre des temps de génération proportionnels à N. Dans ce cas, on construit un objet limite appelé *coalescent de Kingman* qui est un objet géométrique probabiliste décrivant les lignées ancestrales.

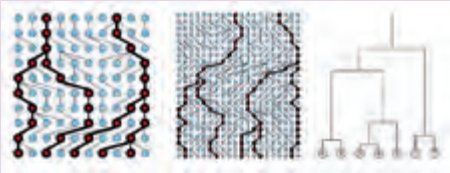


Figure 4. Le coalescent de Kingman comme limite d'échelle d'arbres généalogiques (d'après B. Mallein, Culture Math, ENS Ulm, 2011)

Pour aller plus loin

Bacaer N., (2008). *Histoires de mathématiques et de populations.* Le sel et le fer, Cassini.

Haccou P., Jagers P., Vatutin V.A., (2005). *Branching Processes : Variation, growth and extinction of populations*, Cambridge University Press.

Istas J., (2000). *Introduction aux modélisations mathématiques pour les sciences du vivant*. Springer.

Mallein B., (2011). *Généalogie de populations : le coalescent de Kingman*, Culture Math, ENS Ulm.
http://www.math.ens.fr/culturemath/articles-ens/mallein11/coalescent-de-kingman.html

Tangente Hors série n° 42 (2011). *Mathématiques et biologie – L'organisation du vivant.*

Darwin online
http://darwin-online.org.uk/content/frameset?pageseq=60&itemID=F1497&viewtype=side

La restauration de vieux films

Julie Delon, *chargée de recherche CNRS à Telecom ParisTech*
Agnès Desolneux, *directrice de recherche CNRS à l'École Normale Supérieure de Cachan*

Le papillonnage fait partie des défauts qui affectent couramment les bandes abîmées. À travers ce cas particulier, voyons comment les mathématiques aident à créer des algorithmes permettant de corriger automatiquement les imperfections des vieux films.

L'apparition des techniques de numérisation permet aujourd'hui d'accéder à une part importante de l'héritage cinématographique. Cependant, le processus de numérisation doit être accompagné d'une restauration des nombreux défauts qui altèrent les films et les vidéos. Parmi ces défauts, citons en particulier le *papillonnage* (appelé aussi *flicker* pour reprendre le terme anglais), qui se traduit par des variations artificielles de l'intensité de l'image, les rayures, la dérive des couleurs, les *blotches* (petites taches dues à la présence de poussière ou à la perte de morceaux de gélatine sur la pellicule), etc.

La variété des défauts observés dans les films rende leur restauration particulièrement difficile. Un film de 90 minutes contient environ 130 000 images. Pour le restaurer complètement en un temps raisonnable, et pour éviter un investissement humain trop important, des algorithmes de restauration rapides et les plus automatiques possibles deviennent indispensables. Le développement de tels algorithmes est d'autant plus critique que la tendance actuelle vers des résolutions d'images toujours plus hautes (TV HD ou support blu-ray par exemple) accentue la visibilité du moindre défaut dans les images.

Le papillonnage

Concentrons-nous sur un défaut très fréquent dans les films anciens : le papillonnage (ou *flicker*). Ce défaut consiste en

Figure 1. Trois images extraites du film « The Cure » (1917) de Charlie Chaplin. L'image centrale est nettement plus sombre que les deux autres. C'est ce défaut de luminosité, appelé papillonnage ou flicker, qui donne l'impression que le film « clignote », et que l'on va chercher à corriger de façon automatique.

des variations non naturelles de contraste d'une image à l'autre du film : les images deviennent artificiellement sombres ou claires, alternativement. Ces variations de contraste peuvent être dues à la fois à la dégradation chimique du support du film (qui crée alors des zones plus sombres ou plus claires lors du visionnage), mais aussi à une variation du temps d'exposition (temps pendant lequel la pellicule est exposée à la lumière) d'une image à l'autre. Ceci est en particulier vrai pour les films tournés à l'époque où la pellicule était entraînée manuellement.

Contrairement à d'autres défauts couramment observés dans les films (rayures, poussières, etc.), le papillonnage ne fait pas apparaître de nouvelles structures dans les images. Sa particularité est donc d'être transparent, voire quasiment invisible sur une image isolée. Seul le visionnage des images successives du film permet de se rendre compte de sa présence. Par conséquent, la restauration d'un tel défaut ne peut pas se faire sur chaque image indépendamment des autres, il faut impérativement utiliser plusieurs images successives du film et chercher à en « moyenner le contraste ».

Afin de corriger ce défaut, les films sont d'abord numérisés, ce qui veut dire que la pellicule est scannée, image par image, et que cette suite d'images numériques est stockée sur un ordinateur. En général, une seconde de film comporte 24 images. Un film d'une heure, une fois scanné, contient donc 86400 images numériques. Une image numérique en noir et blanc est modélisée mathématiquement comme une fonction définie sur une grille rectangulaire de carrés (appelés pixels pour la contraction de « picture elements ») et à valeurs dans l'ensemble des nombres positifs. La valeur de l'image en un pixel est appelée le niveau de gris de ce pixel.

Changement de contraste

Une fois le film numérisé, on peut appliquer à toutes ses images ce qu'on appelle des *changements de contraste* (nous verrons plus loin comment les construire pour éli-

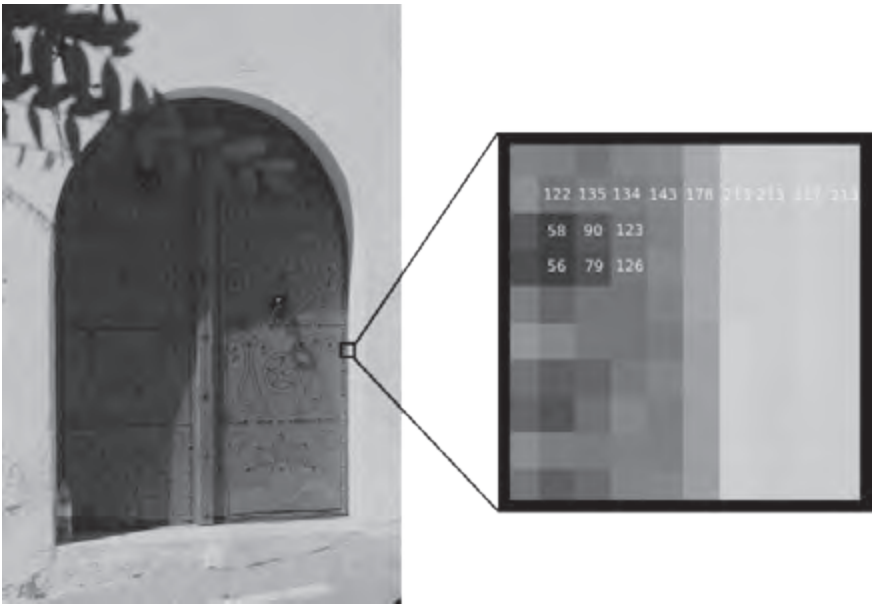


Figure 2. Une image numérique et un zoom sur un petit morceau de cette image, faisant apparaître quelques pixels et leurs niveaux de gris.

miner le papillonnage). Dire qu'une image I subit un changement de contraste veut dire qu'elle est transformée en $f(I)$ où f est une fonction croissante sur l'ensemble des nombres positifs : chaque pixel de coordonnées (x,y) voit son niveau de gris $I(x,y)$ devenir $f(I(x,y))$. Ainsi, deux pixels ayant des niveaux de gris identiques dans l'image I auront encore des niveaux de gris identiques dans la nouvelle image $f(I)$. L'intérêt et la nécessité d'utiliser une fonction f croissante est qu'elle conserve l'ordre des niveaux de gris : si un pixel est plus sombre qu'un autre pixel dans l'image I, cette propriété reste vraie dans l'image $f(I)$. En conséquence, un changement de contraste ne modifie pas le contenu géométrique d'une image, c'est-à-dire qu'on voit la même chose dans l'image avant et après un changement de contraste. Il n'y a pas d'apparitions de nouveaux objets dans l'image.

> *Un changement de contraste ne modifie pas le contenu géométrique d'une image, c'est-à-dire qu'on voit la même chose dans l'image avant et après un changement de contraste.*

Les changements de contraste utilisés pour supprimer le papillonnage d'un film sont construits de la manière suivante. Imaginons que l'on souhaite restaurer la $n^{\text{ième}}$ image du film, notée I_n. Pour chaque pixel

Figure 3. Changement de contraste. L'image de droite est obtenue en appliquant un changement de contraste à l'image de gauche. On peut observer que le contenu géométrique de l'image n'a pas changé.

(x,y) de I_n, on regarde quel est le rang R de ce pixel dans l'image I_n lorsque tous les pixels de cette image sont ordonnés de façon croissante suivant leur niveau de gris (on suppose que ce rang est le même pour des pixels de niveaux de gris égaux). Regardons maintenant, dans les 10 images qui précèdent I_n et dans les 10 images qui la suivent dans le film (voir l'encadré *Le nombre d'images*), quels sont les niveaux de gris des pixels ayant aussi le rang R (ou le rang le plus proche de R) dans ces images. Ces pixels ont des valeurs de gris qui peuvent être très différentes de celle de $I_n(x,y)$. On prend donc la moyenne de tous ces niveaux de gris (21 valeurs en tout), ce qui détermine la nouvelle valeur $f(I_n(x,y))$ au pixel (x,y), pour la $n^{\text{ième}}$ image du film. Cette correction s'appuie sur l'observation suivante : si un objet du film voit ses niveaux de gris changer d'une image à l'autre, le rang des pixels qu'il contient varie très peu d'une image à l'autre. Par exemple, les pixels de niveau median (tels qu'il y ait autant de pixels plus sombres qu'eux que de pixels plus clairs qu'eux dans l'image) correspondront vraisemblablement aux mêmes objets dans les différents images du film (voir l'encadré *L'histogramme cumulé des niveaux de gris*).

On a ainsi réalisé une *égalisation de contraste* à travers les images du film, ce qui fait disparaître l'impression de papillonnage.

Figure 4. Les trois images du film de Charlot « The Cure » après restauration automatique du contraste par la méthode décrite dans cet article. Ces trois images ont maintenant des luminosités similaires, et l'effet de papillonnage (qui se traduisait par une impression de clignotement) du film a disparu.

Bibliographie

Cet article est inspiré de celui que nous avions écrit pour le site Images des Mathématiques :

Delon J., Desolneux A., (2011) *Papillonnage et mathématiques des images – Images des Mathématiques*, CNRS.

En ligne, URL : http://images.math.cnrs.fr/Papillonnage-et-mathematiques-des.html

Handbook of Mathematical Models in Computer Vision, Edited by Nikos Paragios, Chen Yunmei, et Olivier Faugeras, Springer, 2005

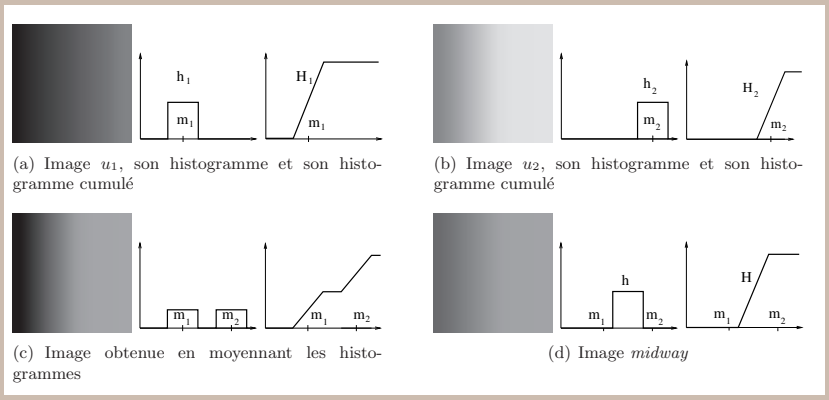
Handbook of Mathematical Methods in Imaging, Edited by Otmar Scherzer, Springer, 2011

Le nombre d'images

Le nombre d'images utilisées pour la restauration (on a pris ici les 10 images précédentes et les 10 images suivantes) est un choix heuristique que l'on peut faire varier. Plus ce nombre est grand, plus on prend en compte un grand nombre d'images dans le film pour restaurer l'image courante, et plus on élimine les fluctuations dues au papillonnage. En contrepartie, prendre un nombre trop grand d'images n'a pas forcément de sens si le film comporte beaucoup de mouvements de caméra ou d'objets qui se déplacent.

L'histogramme cumulé des niveaux de gris

Le changement de contraste appliqué à l'image I_n du film s'interprète de manière précise grâce à la notion d'histogramme cumulé (appelé aussi fonction de répartition) des niveaux de gris de cette image. L'histogramme cumulé des niveaux de gris d'une image est défini comme étant la fonction qui à chaque valeur de niveau de gris (c'est à dire généralement à chaque nombre entre 0 et 255) associe le nombre de pixels dans l'image ayant un niveau de gris inférieur à cette valeur. C'est donc une fonction croissante du niveau de gris. L'opération qui consiste à moyenner les niveaux de gris des pixels de même rang R dans les 10 images qui précèdent et qui suivent chronologiquement I_n dans le film revient en fait à changer l'histogramme cumulé de I_n en la moyenne harmonique (inverse de la moyenne arithmétique des fonctions inverses) des histogrammes cumulés des 21 images considérées. D'autres types de moyenne seraient éventuellement envisageables, en particulier la moyenne arithmétique, mais on aurait alors apparition d'artefacts, comme l'illustre la figure 5. En fait, on peut montrer que la moyenne harmonique est la seule qui garantit que, sur un film fait uniquement d'une image fixe, on moyenne correctement le contraste.



(a) Image u_1, son histogramme et son histogramme cumulé

(b) Image u_2, son histogramme et son histogramme cumulé

(c) Image obtenue en moyennant les histogrammes

(d) Image *midway*

Figure 5. (a) Image d'un dégradé sombre, avec son histogramme et son histogramme cumulé. (b) Image d'un dégradé clair, avec son histogramme et son histogramme cumulé. (c) Image obtenue si on avait utilisé la moyenne arithmétique pour changer le contraste : on a ici créé une discontinuité dans le dégradé. (d) Image obtenue avec la moyenne harmonique des histogrammes cumulés : le résultat est conforme à ce qu'on attend - c'est un dégradé de niveau de gris « moyen ».

Cryptage et décryptage : communiquer en toute sécurité

Jean-Louis Nicolas, *professeur à l'Université Claude Bernard Lyon 1*
Christophe Delaunay, *professeur à l'Université de Franche-Comté*

La sécurisation de nos cartes bleues, ainsi que d'autres procédés de cryptages utilisés couramment, se basent sur l'impossibilité, en pratique, de factoriser de très grands nombres. Ce type de cryptage pourrait cependant être détrôné par d'autres méthodes, sa fiabilité étant sans cesse remise en question par les progrès de l'informatique.

En mars 2000, un gros titre avait fait la une des journaux : « Alerte à la sécurité des cartes bancaires ». Que s'était-il passé ? En France, le secret des cartes à puce était protégé depuis 1985 grâce à une méthode de cryptage faisant intervenir un grand nombre N, constitué de 97 chiffres. Ce nombre N doit être le produit de deux grands nombres premiers, c'est-à-dire de nombres qui, comme 7 ou 19, ne sont divisibles que par 1 et par eux-mêmes. Le secret d'une carte bancaire est constitué précisément par ce couple de nombres premiers ; les calculer à partir de N était pratiquement impossible dans la décennie 1980. Mais avec l'augmentation de la puissance des ordinateurs et l'amélioration des méthodes mathématiques, la taille des nombres N dont on peut calculer les facteurs premiers en un temps raisonnable a dépassé la centaine de chiffres dans les années 1990 (le record actuel, obtenu en décembre 2009, est la factorisation d'un nombre de 232 chiffres). Ainsi, un informaticien astucieux, Serge Humpich, avait pu trouver les deux nombres premiers ultra-secrets dont le produit valait le nombre N d'alors. Pour garantir la sécurité de nos petits rectangles de plastique, l'organisme de gestion des cartes ban-

caires avait été obligé de construire aussitôt de nouveaux nombres N, nettement plus grands. La page web de Paul Zimmerman (voir les références en fin d'article) offre une mise à jour des différents records de factorisation des entiers.

La taille des nombres N dont on peut calculer les facteurs premiers en un temps raisonnable a dépassé la centaine de chiffres dans les années 1990.

La cryptographie moderne, au croisement des mathématiques et de l'informatique

Cette péripétie illustre l'importance considérable que revêt aujourd'hui la science du cryptage, c'est-à-dire du codage de messages en vue de les rendre illisibles par des personnes indiscrètes. Crypter et décrypter des messages secrets est une activité vieille de plusieurs siècles, voire millénaires. Et cette activité a largement débordé du cadre strictement diplomatique ou militaire pour investir des pans entiers de l'univers des communications civiles : procédures d'authentification, transactions bancaires, commerce électronique, protection de sites et fichiers informatiques, etc.

La cryptographie a connu beaucoup d'avancées au cours des dernières décennies. Ce faisant, elle est devenue une science complexe, où les progrès sont généralement le fait de spécialistes ayant reçu une formation poussée en mathématiques et en informatique.

Cette spécialisation s'est manifestée dès la Deuxième guerre mondiale. On le sait aujourd'hui, le déchiffrage par les Alliés des messages codés par les fameuses machines allemandes Enigma a joué un rôle déterminant dans ce conflit. Or c'est un éminent mathématicien britannique, Alan Turing, par ailleurs l'un des pères de l'informatique théorique, qui a apporté une contribution essentielle à ce décryptage.

Dans les années 1970, la cryptographie a connu une petite révolution : l'invention de la cryptographie à « clé publique », avec la méthode RSA. De quoi s'agit-il ? Jusque-là, les correspondants voulant échanger des messages secrets devaient partager une clé secrète, et le risque d'interception de cette clé par l'ennemi était grand. Le protocole RSA, nommé ainsi d'après ses trois inventeurs (Ronald Rivest, Adi Shamir et Leonard Adleman), résout ce problème. Cette méthode utilise deux clés : une clé de

cryptage publique – elle peut être connue de tous – et une clé de décryptage, qui reste secrète. Elle est fondée sur le principe (utilisé par la suite pour protéger les cartes bancaires, comme on l'a vu plus haut) qu'il est possible de construire de grands nombres premiers (de cent, mille chiffres, voire plus), mais qu'il est extrêmement difficile de retrouver les facteurs premiers p et q d'un grand nombre $N = p \times q$ lorsque l'on connaît seulement N. Schématiquement, la connaissance de N revient à celle de la clé publique de cryptage, tandis que la connaissance de p et q revient à celle de la clé secrète de décryptage.

Évidemment, si quelqu'un trouvait une méthode pour décomposer rapidement en leurs facteurs premiers de grands nombres, le protocole RSA deviendrait caduc. Mais il se pourrait aussi que les mathématiciens prouvent qu'une telle méthode n'existe pas, ce qui renforcerait la sécurité du protocole RSA. Ce sont là des sujets de recherche décisifs.

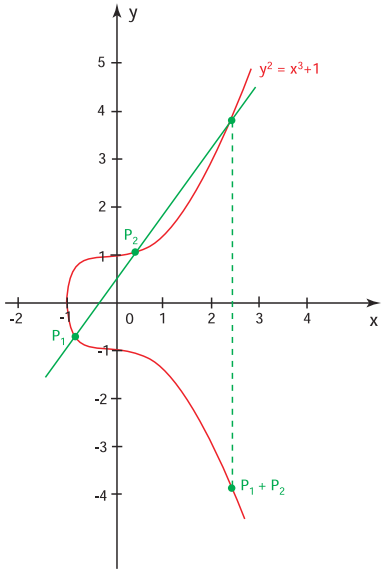
Les méthodes qui, comme le protocole RSA, font intervenir de la théorie des nombres élaborée, apportent une grande leçon : des recherches mathématiques (sur les nombres premiers notamment) tout à fait désintéressées peuvent se révéler, des années ou des décennies plus tard, cruciales pour telle ou telle application ; et ce de manière imprévisible.

Dans son livre *L'apologie d'un mathématicien*, le grand théoricien des nombres britannique G. H. Hardy (1877-1947), qui était un fervent pacifiste, se targuait de travailler dans un domaine parfaitement pur, l'arithmétique, et de n'avoir rien fait qui puisse être considéré comme « utile ». Ses travaux étaient peut-être « inutiles » à son époque. C'est faux aujourd'hui.

Courbes elliptiques : la géométrie algébrique au service des agents secrets

Et cela ne concerne pas uniquement la théorie des nombres. D'autres domaines des mathématiques, auparavant considérés comme dépourvus d'applications, contribuent à la science du cryptage. Des méthodes cryptographiques prometteuses et fondées sur des principes voisins de ceux du protocole RSA sont apparues au cours des dernières années. Il en est ainsi de la méthode dite du *logarithme discret*. Celle-ci a servi à son tour à concevoir des méthodes qui s'appuient sur les propriétés des *courbes elliptiques*. Il ne s'agit pas de courbes ayant la forme d'une ellipse, mais de courbes dont l'étude a débuté au XIX^e^ siècle pour résoudre

le problème difficile du calcul du périmètre d'une ellipse. Ces courbes, dont les coordonnées (x, y) de leurs points vérifient une équation de la forme $y^2 = x^3 + ax + b$, ont d'intéressantes propriétés – dont l'étude fait partie de la *géométrie algébrique*, très vaste domaine des mathématiques actuelles. Par exemple, à l'aide d'une construction géométrique appropriée, il est possible de définir une addition entre les points d'une courbe elliptique (voir figure ci-dessous).



Le graphe de la courbe elliptique d'équation $y^2 = x^3 + 1$. Les courbes elliptiques ont une propriété remarquable : on peut « additionner » leurs points selon le procédé représenté sur le dessin. L'« addition » ainsi définie respecte les lois arithmétiques attendues, telles que $(P_1 + P_2) + P_3 = P_1 + (P_2 + P_3)$. Certaines méthodes modernes de cryptographie font appel aux courbes elliptiques et à leurs propriétés algébriques.

Plus généralement, les objets géométriques que sont les courbes elliptiques possèdent des propriétés arithmétiques – que l'on continue d'explorer – susceptibles de rendre service à la cryptographie. C'est ainsi qu'a été développée une méthode cryptographique intitulée *logarithme discret sur les courbes elliptiques*. De façon anecdotique, les courbes elliptiques fournissent aussi une méthode originale pour factoriser les entiers (cependant, des méthodes bien plus techniques sont nécessaires pour obtenir les records actuels).

Les objets géométriques que sont les courbes elliptiques possèdent des propriétés arithmétiques susceptibles de rendre service à la cryptographie.

L'ordinateur quantique : l'outil de demain ?

Une autre direction, totalement différente, est apparue assez récemment. Il s'agit de la cryptographie quantique. Que signifie ce terme ? Il y a quelques années, des physiciens et des mathématiciens ont imaginé qu'il serait un jour possible de réaliser un ordinateur quantique, c'est-à-dire dont le fonctionnement exploiterait les lois bizarres de la physique quantique, celles qui règnent dans le monde infiniment petit. Or, on s'est rendu compte qu'un tel ordinateur, s'il était réalisable, serait capable de factoriser très vite de grands nombres et rendrait ainsi totalement inefficace la méthode RSA (dans ce contexte, lors du congrès international des mathématiciens à Berlin en 1998,

Peter Shor, des laboratoires AT & T, obtenait le prix Nevanlinna pour ses travaux sur la factorisation à l'aide des ordinateurs quantiques). Des recherches visant la réalisation concrète d'un ordinateur quantique ont d'ailleurs été publiées dans la revue britannique *Nature* (cf. référence ci-dessous). D'un autre côté, des chercheurs ont élaboré des protocoles de cryptographie quantique, c'est-à-dire des méthodes de cryptage utilisant des objets (photons, atomes…) obéissant aux lois de la physique quantique. Ces protocoles quantiques garantissent (du moins théoriquement) une sécurité infaillible. Tout cela est à l'étude et risque de devenir opérationnel dans un futur proche (par exemple, un câble de communication quantique reliant la ville de Genève à Lausanne fonctionne déjà depuis plusieurs années…).

Pour aller plus loin

La page web de Paul Zimmerman sur les records de factorisation des entiers : www.loria.fr/~zimmerma/records/factor.html

Kahn D., (1980). *La guerre des codes secrets* (Interéditions).

Stern J., (Jacob O. 1998). *La science du secret*.

Singh S., (Lattès J.-C., 1999) *Histoire des codes secrets.*

Delahaye J.-P., (2000). *Merveilleux nombres premiers* (Belin/Pour la Science).

Stinson D. (Vuibert, 2001). *Cryptographie, théorie et pratique*.

Vandersypen L. M. K., *et al.*, (2001) *Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance, Nature*, vol. 414, p. 883-887.

Pourquoi et comment nager dans le miel ?

François Alouges, Guilhem Blanchard, Sylvain Calisti, Simon Calvet, Paul Fourment, Christian Glusa, Romain Leblanc et Mario Quillas-Saavedra,
respectivement professeur et élèves de l'École polytechnique

La nage dans des milieux très visqueux comme le miel est un sujet de recherches actuel, qui touche des disciplines aussi diverses que la mécanique des fluides, les mathématiques appliquées ou la biologie. Mais pourquoi donc s'intéresser à la natation dans du miel ? Et quelles sont les différences entre la nage dans du miel et celle dans de l'eau ?

Qui donc pourrait vouloir nager dans du miel ? Personne. Si l'on s'intéresse à la nage dans le miel, ou dans les milieux très visqueux en général, c'est pour étudier des problèmes qui, bien qu'apparemment très différents, mettent en fait en jeu les mêmes mécanismes. En effet, la nage dans des milieux très visqueux à notre échelle (l'échelle macroscopique, de l'ordre du mètre) possède les mêmes caractéristiques que la nage dans de l'eau à très petite échelle (l'échelle microscopique, avec des tailles de l'ordre du centième de millimètre). Or, il est plus facile de réaliser des expériences de taille normale dans du miel que des expériences miniatures dans de l'eau !

Grâce à cet artifice, on peut étudier la nage dans des milieux aqueux (c'est-à-dire à base d'eau) à très petite échelle. Et, cette fois, les applications pratiques sont multiples. Par exemple, les bactéries sont des micro-organismes qui évoluent en nageant dans des fluides proches de l'eau comme le sang. Mieux comprendre leurs capacités de mouvement pourrait permettre soit de les empêcher de se déplacer afin de lutter contre certaines maladies soit, au contraire, de les aider à se développer (dans le cas par exemple des bactéries qui composent la flore intestinale et qui nous servent à digérer la nourriture que nous avalons). Un autre exemple, aux applications médicales encore plus évidentes, est celui des micro-robots nageurs. En effet, si nous arrivions à construire des minuscules robots capables de se mouvoir dans des milieux comme l'eau ou le sang, nous pourrions nous en

servir pour transporter des substances médicamenteuses au sein même des cellules malades, effectuer des opérations chirurgicales dans le corps humain sans avoir besoin d'y découper une ouverture pour le scalpel du chirurgien, ou encore réaliser des réparations de taille microscopique, inaccessibles aux outils, même très perfectionnés, maniés par les humains.

Inertie et viscosité

Quelles sont donc les différences entre la nage de Michael Phelps, 1m93, et celle de la bactérie Escherichia Coli (voir figure 1), qui mesure quelques micromètres (millièmes de millimètres) ? Pourquoi ne peut-on pas étudier le mouvement de ces deux êtres vivants de la même façon ?

D'une façon générale, les deux types d'effets qui interviennent lors du mouvement dans un fluide sont les effets d'inertie et les forces de viscosité. Les effets inertiels, ce sont par exemple ceux que l'on ressent à l'intérieur d'un avion qui décolle : on est alors « plaqué » contre son siège pendant que l'avion accélère brutalement. Ces effets sont liés au fait que les forces physiques (frottements, poids, etc.) n'agissent pas directement sur notre vitesse, mais sur notre accélération.

Il est plus facile de réaliser des expériences de taille normale dans du miel que des expériences miniatures dans de l'eau.

Les forces de viscosité, elles, représentent les interactions entre les différentes molécules qui constituent le fluide. On peut les associer à la résistance que le fluide oppose lorsqu'on essaye de le déformer.

Figure 1. Un champion de natation vs. *E. Coli. Une différence de… taille.*

Ces deux effets ont des intensités très différentes en fonction de l'échelle à laquelle on se place. Pour mesurer leur importance relative, on utilise un nombre sans dimension appelé nombre de Reynolds. Le nombre de Reynolds d'un écoulement est donné par la formule $Re = \rho UL/\nu$, dans laquelle ρ et ν représentent la densité et la viscosité du fluide, et L et U sont respectivement une taille et une vitesse caractéristiques de l'écoulement (celles du nageur par exemple) : un nombre de Reynolds très grand (par rapport à 1) signifiera que les effets inertiels sont beaucoup plus importants que les effets visqueux, que l'on négligera alors ; au contraire, si le nombre de Reynolds est très petit, on se contentera de prendre en compte la viscosité, et pas l'inertie. C'est donc là que se situe toute la différence entre Michael Phelps et E. Coli : le nombre de Reynolds caractérisant la nage du champion olympique vaut environ 10 millions, tandis que celui caractérisant la nage d'une bactérie est de l'ordre de 0,00001.

Deux écoulements possédant des nombres de Reynolds semblables ont des caractéristiques similaires même s'ils correspondent à des situations très différentes.

Deux écoulements possédant des nombres de Reynolds semblables ont des caractéristiques similaires même s'ils correspondent à des situations très différentes comme par exemple pour des microorganismes évoluant dans l'eau, des objets de taille centimétrique évoluant dans un fluide très visqueux (miel, silicone, peinture, etc.) ou bien des écoulements au sein des glaciers. Dans ces trois cas, on est à faible nombre de Reynolds.

Le théorème de la Coquille Saint-Jacques

La nage à faible nombre de Reynolds, avant tout liée aux effets de viscosité, a des propriétés intéressantes et parfois inattendues. L'écoulement du fluide est décrit par des équations appelées *équations de Stokes* qui sont en fait une version simplifiée des *équations de Navier-Stokes* dans lesquelles l'inertie a été négligée. Elles ont une particularité extrêmement importante du point de vue mathématique : elles sont linéaires, c'est-à-dire que la modification d'un paramètre (pression, vitesse, etc.) entraîne une modification proportionnelle de tous les autres paramètres.

Cette linéarité a des conséquences importantes pour la nage à faible nombre de Reynolds, comme la réversibilité : si un objet se déforme d'une certaine façon, puis reprend sa forme initiale à nouveau de façon exactement inverse – on dit alors que le mouvement est *réciproque* – il reviendra aussi à sa position initiale et ne se sera pas déplacé. C'est le « théorème de la Coquille Saint-Jacques », établi par E.M. Purcell en 1977 : ces coquillages nageant en ouvrant et fermant leur coquille ne peuvent pas se déplacer dans des milieux dominés par la viscosité.

Figure 2. La coquille Saint-Jacques ne pourrait pas se déplacer dans un milieu dominé par la viscosité.

Pour contourner cette difficulté, Purcell imagine dès 1977 un système à deux batteurs capable de nager à faible nombre de Reynolds, en les actionnant successivement comme schématisé sur la figure 3. On notera qu'ici le mouvement n'est pas réciproque.

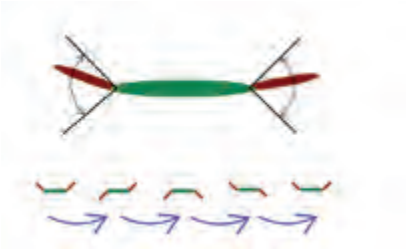


Figure 3. Le « Three-Link-Swimmer » de Purcell est un robot-nageur à deux batteurs. En actionnant successivement les deux batteurs selon le schéma du bas, le nageur progresse latéralement dans un écoulement dominé par la viscosité.

Optimiser sa natation

Le problème de la natation consiste à chercher des changements de forme périodiques, des brassées, qui produisent, par interaction avec le fluide, un déplacement du nageur. Le nageur (ou le robot-nageur) peut ensuite enchaîner les brassées de façon à avancer dans le fluide. Mathématiquement, ce problème revient à chercher un *chemin fermé* dans l'ensemble des formes que peut prendre l'objet qui conduise à un changement de sa position.

Le problème consistant à trouver des brassées minimisant l'énergie mécanique dépensée par le nageur afin de produire un déplacement donné est un problème de contrôle optimal.

C'est un problème de *contrôle*. En agissant sur des paramètres de contrôle – ici la forme du nageur – on souhaite réaliser un objectif : déplacer le nageur. De la même façon, conduire sa voiture est un problème de contrôle : en actionnant le volant et l'accélérateur, on cherche à déplacer le véhicule et l'amener à un endroit précis. La théorie du contrôle est l'ensemble des outils mathématiques qui sont utilisés pour démontrer rigoureusement qu'un dispositif est contrôlable, c'est-à-dire pour notre nageur, qu'il peut effectivement se déplacer en effectuant des brassées. Cette théorie utilise beaucoup d'arguments géométriques et d'équations différentielles ordinaires. En utilisant la théorie du contrôle, on peut ainsi démontrer rigoureusement que certains robots nageurs proposés par des physiciens

peuvent effectivement se déplacer dans un fluide à un régime dominé par la viscosité. Les figures 3 et 4 donnent des exemples de tels robots.

Souvent, savoir que l'on peut contrôler un système est insuffisant. On veut aussi le faire en optimisant un critère, comme par exemple, lorsqu'on se déplace en voiture et que l'on veut minimiser le temps de parcours ou le carburant consommé. Le domaine mathématique est alors celui du contrôle optimal. Ainsi, pour revenir à la natation, le problème consistant à trouver des brassées minimisant l'énergie mécanique dépensée par le nageur afin de produire un déplacement donné (ou celui d'aller le plus vite possible avec une quantité d'énergie fixée) est un problème de contrôle optimal.

Il est possible de construire des algorithmes permettant de calculer numériquement (à l'aide d'un ordinateur) les brassées optimales de certains nageurs. Après modélisation du problème à résoudre, il s'agit de trouver une solution d'une équation différentielle ordinaire qui part d'un point connu (la position et la forme du nageur initialement) et arrive à un autre point connu (la position et la forme finales). On utilise alors ce que l'on appelle des « méthodes de tir » (qui portent leur nom par analogie au problème de balistique qui consiste à envoyer un boulet de canon d'un point à un autre en réglant l'orientation du canon). Pour les équations différentielles que l'on doit résoudre pour la natation optimale, on connaît le point de départ et l'on veut viser le point d'arrivée en réglant la vitesse initiale.

Ces recherches en théorie du contrôle optimal sont très actuelles (voir *Pour aller plus loin*). Qui sait, peut-être permettront-elles de concevoir les micro-robots de demain !

Pour aller plus loin

Le lecteur intéressé par l'approfondissement de ce sujet est invité à lire le très bel article de Purcell :

Purcell E.M., (1977). *Life at low Reynolds number*, Am. J. Phys, 45(1):3–11.

Les travaux de recherche d'un des auteurs de cet article sont également disponibles :

Alouges F., DeSimone A., Lefebvre A., (2008). *Optimal strokes for low Reynolds number swimmers : an example*, *J. Nonlinear Sci.*, 18(3):277-302.

Alouges F., DeSimone A., Lefebvre A., (2009). *Biological fluid dynamics : swimming at low Reynolds numbers*, *Encyclopedia of Complexity and System Science*, Springer Verlag.

A l'adresse http://www.cmap.polytechnique.fr/~alouges/nage.php on peut voir des vidéos montrant les stratégies optimales de natation de certains mécanismes et leur comparaison avec d'autres modes moins performants. La figure 4 est tirée de l'une d'entre elles.

Très instructif aussi, le film réalisé par G. I. Taylor (**G. I. Taylor**, *Low Reynolds Number Flow*, *National Committee for Fluid Mechanics Films*, Education Development Center. Inc., Newton, MA.) – et accessoirement disponible sur YouTube – montre un éventail exhaustif des multiples particularités des écoulements à faible nombre de Reynolds.

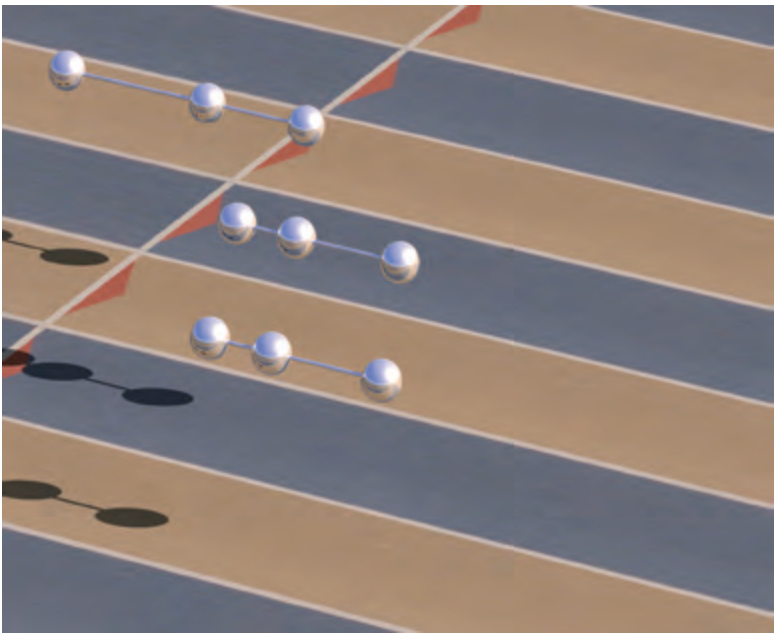


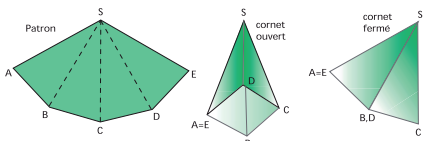
Figure 4. Simulation numérique d'une course entre trois nageurs composés de sphères alignées et reliées par des vérins extensibles. Le nageur au premier plan est calculé de façon à nager optimalement, et nage plus vite que les deux autres pour la même énergie dépensée. La course a lieu de la gauche vers la droite.

Le théorème du soufflet

Étienne Ghys, *directeur de recherche CNRS à l'École Normale Supérieure de Lyon*

Une règle, un crayon, du carton, des ciseaux et de la colle : il n'en faut guère plus pour procurer aux mathématiciens du plaisir et de jolis problèmes – dont l'étude se révèle souvent, après coup et de manière inattendue, utile dans d'autres métiers.

Construisons une pyramide en carton… Pour cela, on commence par découper un patron SABCDE dans une feuille de carton comme indiqué sur la figure, puis on plie le long des lignes pointillées et enfin, on colle les côtés AS et ES.



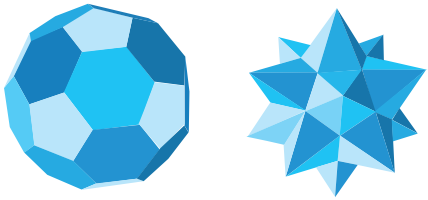
Le résultat est une espèce de cornet dont le sommet est le point S et dont le bord est un quadrilatère ABCD. Cet objet est *flexible*. Si on le tient dans la main, le quadrilatère ABCD peut se déformer et s'ouvrir plus ou moins : la construction n'est pas très solide. Pour compléter la pyramide, il faut encore découper un carré en carton et le coller sur le quadrilatère pour former la base. Après cette opération, la pyramide est solidifiée, rigidifiée. Si on la pose sur une table, elle ne s'écroule pas. Si on la prend dans la main et si on essaye de la déformer (avec douceur !), on n'y parvient pas, à moins de déformer les faces en carton. De même, un cube en carton est *rigide* comme tout le monde l'a souvent constaté. Qu'en est-il pour un *polyèdre* plus général, possédant peut-être des milliers de faces ? La géode de la Villette est-elle rigide ? Cette dernière question laisse entrevoir que le sujet de la rigidité et de la flexibilité n'est peut-être pas seulement théorique !

Le cas des polyèdres convexes

Le problème de la rigidité de ce type d'objets est très ancien. Euclide en avait probablement connaissance. Le grand mathématicien français André-Marie Legendre s'y intéresse vers la fin du XVIII[e] siècle et en parle à son collègue Joseph-Louis Lagrange qui suggère à son tour au jeune Augustin-Louis Cauchy d'étudier cette question en 1813. Ce sera le premier résultat marquant du baron A.-L. Cauchy, qui deviendra par la suite l'un des plus grands mathématiciens de son siècle.

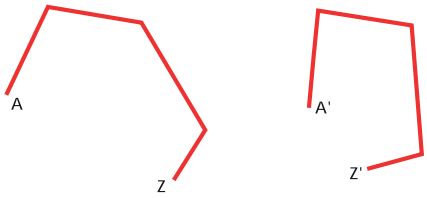
Augustin-Louis Cauchy (1789-1857), l'un des grands mathématiciens de son époque. (Cliché Archives de l'École polytechnique)

A.-L. Cauchy s'intéresse aux polyèdres *convexes*, c'est-à-dire aux polyèdres qui n'ont pas d'arêtes rentrantes. Par exemple, la pyramide que nous avons construite ou le ballon de football est convexe mais l'objet dessiné à droite de la figure ci-dessous ne l'est pas.



Le théorème établi par A.-L. Cauchy est le suivant : *tout polyèdre convexe est rigide*. Cela signifie que si on construit un polyèdre convexe avec des polygones indéformables (en métal par exemple) ajustés par des charnières le long d'arêtes, la géométrie globale de l'ensemble empêche les jointures de jouer. Le cornet que nous avons construit est flexible mais ceci n'invalide pas le théorème : il lui manque une face et c'est la dernière face qui rigidifie la pyramide !

Faire des mathématiques, c'est *démontrer* ce qu'on affirme ! La démonstration de A.-L. Cauchy est superbe (même si certains ont fait remarquer par la suite qu'elle était incomplète). Il n'est malheureusement pas question dans ce petit article de donner une idée de cette preuve, mais j'aimerais en extraire un « lemme », c'est-à-dire une étape dans la démonstration. Posons sur le sol une chaîne constituée de quelques barres métalliques assemblées bout à bout, comme sur la figure ci-après.



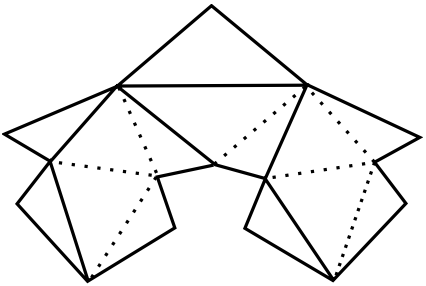
Si l'on diminue les angles de la chaîne, la distance entre les extrémités diminue : AZ est plus grand que A'Z'.

En chacun des angles de cette ligne polygonale, rapprochons les deux barres de façon à diminuer l'angle correspondant. Alors, les deux extrémités de la chaîne se rapprochent. Évident ? Essayez de le démontrer…

Construction d'un « flexidron »

Pendant longtemps, beaucoup de mathématiciens se sont demandés si les *polyèdres non convexes* étaient également rigides. Peut-on trouver une preuve de la rigidité qui n'utiliserait pas l'hypothèse de convexité ? Les mathématiciens aiment les énoncés dans lesquels *toutes* les hypothèses sont utiles pour obtenir la conclusion. Il a fallu attendre plus de 160 ans pour connaître la réponse dans ce cas particulier.

En 1977, le mathématicien canadien Robert Connelly crée la surprise. Il construit un polyèdre (assez compliqué) qui est flexible, bien sûr non convexe pour ne pas contrarier A.-L. Cauchy ! Depuis, sa construction a été quelque peu simplifiée, en particulier par Klaus Steffen. Je présente sur la figure un patron qui permettra au lecteur de construire le « flexidron » de K. Steffen. Découpez, pliez le long des lignes. Les lignes en continu sont des arêtes saillantes et les lignes en pointillé correspondent aux arêtes rentrantes. Collez les bords libres de la manière évidente. Vous obtiendrez une espèce de *Shadok* et vous verrez qu'il est effectivement flexible (un peu…).



Le flexidron de Steffen. Les côtés mesurent 5, 10, 11,12 et 17.

À l'époque, les mathématiciens furent enchantés par ce nouvel objet. Un modèle métallique fut construit et déposé dans la salle de thé de l'Institut des Hautes Études Scientifiques, à Bures-sur-Yvette, et on pouvait s'amuser à faire bouger cette chose (à vrai dire pas très jolie, et qui grince un peu). L'histoire raconte que Dennis Sullivan eut l'idée de souffler de la fumée de cigarette à l'intérieur du flexidron de R. Connelly et qu'il constata que lorsqu'on mettait en mouvement l'objet, aucune fumée ne sortait… Il eut donc l'intuition que lorsque le flexidron se déforme, son volume ne varie pas ! L'anecdote est-elle vraie ? Peu importe ! Quoi qu'il en soit, R. Connelly et D. Sullivan conjecturèrent que *lorsqu'un polyèdre se déforme, son volume est constant*. Évidem-

ment, il n'est pas difficile de vérifier cela dans l'exemple particulier du flexidron de R. Connelly ou encore pour celui de K. Steffen (au prix de calculs compliqués mais sans intérêt). La conjecture en question considère **tous** les polyèdres, y compris ceux qui n'ont jamais été construits en pratique ! Ils ont appelé cette question la « conjecture du soufflet » : le soufflet au coin du feu éjecte de l'air quand on le presse ; autrement dit, son volume diminue (et c'est d'ailleurs sa fonction). Évidemment un vrai soufflet ne répond pas au problème de R. Connelly et D. Sullivan : il est fabriqué en cuir et ses faces se déforment constamment, contrairement à nos polyèdres dont les faces sont rigides.

En 1997, R. Connelly, I. Sabitov et A. Walz ont finalement montré cette conjecture. Démonstration grandiose, montrant une fois de plus les interactions entre toutes les parties des mathématiques. Dans cette question évidemment géométrique, les auteurs utilisent des méthodes très fines d'algèbre abstraite moderne. Il ne s'agit pas d'une démonstration que A.-L. Cauchy « aurait pu trouver » : les techniques mathématiques dont il disposait n'étaient pas suffisantes.

Je voudrais « rappeler » une formule qu'on apprenait autrefois à l'école secondaire. Si les longueurs des côtés d'un triangle sont *a*, b et c, on peut calculer facilement la superficie du triangle. Pour cela, on calcule d'abord le *demi-périmètre* $p = (a+b+c)/2$ et ensuite on obtient la superficie S en extrayant la racine carrée de $p(p-a)(p-b)(p-c)$. C'est une « jolie formule » qui porte le nom du mathématicien grec Héron et qui nous vient de la nuit des temps. Peut-on calculer de même le volume d'un polyèdre si on connaît les longueurs de ses arêtes ? C'est ce qu'ont montré nos trois auteurs contemporains. Ils partent d'un polyèdre construit à partir d'un patron formé d'un certain nombre de triangles et ils appellent l_1, l_2, l_3 etc. les longueurs des côtés de ces triangles (qui peuvent être très nombreux). Ils trouvent alors une équation du n^{ieme} degré satisfaite

Le dôme de la Villette (1730 triangles et 36 mètres de diamètre), et un dôme plus petit.

par le volume V du polyèdre, de la forme $a_0 + a_1V + a_2V^2 + \ldots + a_nV^n = 0$.
Le degré n dépend du patron utilisé et les coefficients de l'équation a_0, a_1 etc. dépendent explicitement des longueurs des côtés l_1, l_2, l_3 etc. Autrement dit, si on connaît le patron et les longueurs des côtés, on connaît l'équation. Si le lecteur se souvient qu'une équation a en général une solution lorsqu'elle est du premier degré, deux solutions lorsqu'elle est du second degré, il pourra deviner qu'une équation de degré n n'a guère que n solutions. Conclusion : si on connaît le patron et les longueurs, on ne connaît pas nécessairement le volume, mais on sait au moins que ce volume ne peut prendre qu'un nombre fini de valeurs… Lorsque le flexidron se déforme, son volume ne peut donc pas varier continûment ; ce volume est « bloqué » et la conjecture du soufflet est établie…

L'ancien, le simple et l'appliqué

Ce problème est-il utile, intéressant ? Qu'est ce qu'un problème mathématique intéressant ? Question difficile à laquelle les mathématiciens réfléchissent depuis longtemps, bien sûr. Voici quelques éléments de réponse, quelques indices de « qualité ». L'ancienneté est un premier critère : les mathématiciens sont très sensibles à la tradition. Un bon problème doit également s'énoncer simplement, sa solution doit mener à des développements surprenants, si possible mettant en jeu des domaines très différents. De ces points de vue, le problème de la rigidité que nous venons d'aborder est intéressant. La question de savoir si un bon problème doit avoir des applications utiles dans la pratique est plus subtile : les mathématiciens y répondent de manière très variable. Incontestablement, les questions « pratiques », issues par exemple de la physique, servent bien souvent de motivations pour les mathématiques. Parfois, il s'agit de résoudre un problème bien concret, mais le lien est souvent plus flou : le mathématicien ne se sert alors de la question concrète que comme d'une *source d'inspiration* et la résolution effective du problème initial n'est plus la motivation véritable. Le problème de rigidité appartient à cette dernière catégorie. L'origine physique est assez claire : la stabilité et la rigidité de structures, par exemple métalliques. Il est douteux que les exemples de R. Connelly puissent un jour être d'une quelconque utilité pour l'ingénieur. Cependant, il paraît bien clair que ce genre de recherche ne manquera pas, dans un avenir indéterminé, de permettre une meilleure compréhension globale de la rigidité des vastes structures constituées d'un grand nombre d'éléments individuels (macromolécules ou bâtiments…). Il s'agit donc de recherches théoriques et « désintéressées », mais qui pourraient peut-être un jour se rendre utiles…

Ce genre de recherche ne manquera pas, dans un avenir indéterminé, de permettre une meilleure compréhension globale de la rigidité des vastes structures constituées d'un grand nombre d'éléments individuels.

Pour aller plus loin

Berger M., (1977). *Géométrie, v. 3. – Convexes et polytopes, polyèdres réguliers, aires et volumes* (CEDIC/Nathan Information).

Connelly R., Sabitov I., Walz A., (1997). *The bellows conjecture*, Beiträge Algebra Geom., 38, n° 1, p. 1-10.

Connelly R., (1977). *A counterexample to the rigidity conjecture for polyhedra*, Institut des Hautes Études Scientifiques, Publication Mathématique n° 47, p. 333-338.

Kuiper N. H., (1979). *Sphères polyédriques flexibles dans E^3*, d'après Robert Connelly, Séminaire Bourbaki, 30^{e} année (1977/78), exposé n° 514, pp. 147-168 (Lecture Notes in Math. 710, Springer).

La détection de spams : un jeu d'enfant ?

Tristan Mary-Huard, *chargé de recherche INRA à INRA-AgroParisTech*

Comment distinguer automatiquement un spam d'un message normal? Les filtres anti-spams analysent le texte des messages en utilisant des algorithmes de classification en forme d'arbres. Ceux-ci comportent un nombre optimal de nœuds correspondant à autant de questions pertinentes permettant de déterminer la nature d'un message.

Il y a quelques années, un célèbre fabriquant de jouet proposait le jeu « *Qui est-ce ?* ». Le principe du jeu était simple : chaque joueur devait retrouver, parmi une collection de personnages, celui choisi par son adversaire. Pour cela, les joueurs posaient tour à tour une question de la forme : « Le personnage a-t-il des moustaches ? » ou « Le personnage porte-t-il des lunettes ? ». Le premier joueur à trouver le personnage de l'adversaire gagnait la partie. La victoire appartenait donc au joueur posant successivement les questions les plus pertinentes pour identifier le plus rapidement possible le personnage mystère.

L'histoire ne dit pas si Leo Breiman, mathématicien et chercheur américain, et ses collaborateurs de l'Université de Berkeley étaient des joueurs passionnés de « *Qui est-ce ?* » Toujours est-il que la méthode qu'ils proposèrent en 1984 reprend scrupuleusement la logique de ce jeu et l'utilise pour la résolution de problèmes de classification. Cette méthode a été ensuite régulièrement reprise, et sert aujourd'hui - entre autres - à la détection de spam.

Parmi l'ensemble des messages reçus par un individu sur sa messagerie électronique, on distingue deux types : les messages « réguliers » (envoyés par des amis ou par des sites internet auquel l'individu est abonné) et les messages « spams » (communication électronique non sollicitée). Ces spams sont généralement envoyés à des milliers, voire des millions d'individus sur internet, et les conséquences de tels envois

en masse ne sont pas négligeables : en France, on estime que 95 % des courriers reçus en 2009 étaient des spams.

Une possibilité est de classer les messages à partir de l'étude des fréquences d'apparition de certains mots.

Afin d'éviter aux utilisateurs de voir leur messagerie surchargée, les fournisseurs d'accès à internet cherchent à élaborer des méthodes de filtrage anti-spam, capables de distinguer automatiquement un spam d'un message régulier. De telles méthodes sont appelées algorithmes de classification.

« Gagner », « loterie » : les mots typiques des spams

Sur quelles informations l'algorithme de classification peut-il se baser ? Tout message contient un texte. C'est donc à partir de l'analyse du contenu de ce texte qu'il faut décider de classer un message en spam ou régulier. Bien souvent, un spam capte l'attention du destinataire par la promesse d'un gain quelconque (généralement d'ordre financier), et propose de récupérer ce gain en se connectant sur un site internet qui sera ainsi fréquemment visité. Ces courriers se caractérisent donc par la présence de mots comme « gagner », « loterie », ou encore « cliquer » dans le corps du texte. Il existe donc des mots révélateurs de la nature du message. De ce fait, une possibilité est de classer les messages à partir de l'étude des fréquences d'apparition de certains mots.

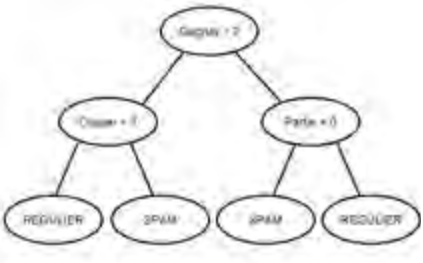


Figure 1. Exemple d'arbre de classification pour le classement des messages électroniques.

L'algorithme proposé par Breiman et ses collègues est illustré en figure 1. Il représente le parcours que doit accomplir un message pour être classé en spam/régulier. Le parcours commence par le haut, en lisant la question inscrite dans le nœud racine (les ellipses sont appelées *nœuds*, les traits entre deux ellipses *branches*, et le premier nœud en haut est appelé *nœud racine*). Si la réponse à la question est affirmative, c'est-à-dire si la fréquence d'apparition du mot « gagner » est strictement supérieure à 2, le message poursuit son chemin en empruntant la branche droite. Sinon, le message emprunte la branche gauche. Le message atteint alors un autre nœud, et comme précédemment poursuit son chemin à droite ou à gauche suivant la réponse faite à la nouvelle question. Ceci se répète jusqu'à ce que le message atteigne une feuille, c'est-à-dire un nœud d'où ne part aucune branche. Dans cette feuille se lit le classement qui doit être attribué au message. Remarquons que le vocabulaire employé (racine, nœuds, branches et feuilles) n'est pas dû au hasard : l'algorithme de Breiman porte en effet le nom d'arbre de classification.

Considérons les messages suivants :

Message 1

Bravo,
Vous venez de gagner à notre grand tirage internet.
Cliquez ici pour recevoir 1.000.000 de dollars !!!
Et pour gagner d'autres cadeaux, cliquez ici.

Message 2

Salut Benoît,
Demain je joue contre Bruno.
Si je gagne la partie je me qualifie directement.
Si je ne gagne pas demain mais que je gagne la suivante,
je monte en pool 3 l'année prochaine !
Bises, Sandra

Dans le premier message, le mot « gagner » n'apparaît que deux fois. La réponse à la question du nœud racine est donc négative, et le message suit la branche gauche. Arrivé au nœud suivant, on vérifie que le mot « cliquer » apparaît deux fois. Le message prend donc la branche droite, pour finalement atteindre une feuille SPAM. Le message sera donc classé en spam. Dans le deuxième message, le mot « gagner » apparaît trois fois, mais il est ici employé au sens de « gagner une partie » plutôt que « gagner de l'argent ». Cette utilisation alternative du mot « gagner » est prise en compte (par le nœud situé en dessous à droite du nœud racine), et le message sera bien classé en REGULIER par l'arbre de la figure 1.

Bien évidemment, l'arbre et les messages considérés ici ne sont qu'une illustration très simplifiée du problème général. Les arbres de classification utilisés en pratique sont souvent plus grands (en termes de nombre de nœuds) et les questions associées à chaque nœud portent sur les fréquences d'apparition de dizaines de mots différents.

Evaluer les différents arbres

Reste la question du choix de l'arbre de classification. Comment choisit-on le nombre de questions à poser ? Ou sur quels mots doivent porter les différentes questions ? Ces choix peuvent s'opérer de la manière suivante : chacun peut proposer l'arbre de classification qui lui semble plus pertinent, puis tous ces arbres sont évalués sur un échantillon test. Un échantillon test est une collection de messages dont la nature est connue parce qu'ils ont déjà été lus et classés en spam/régulier par un individu.

Chaque arbre de classification est tour à tour appliqué à cet échantillon test, et pour chaque message le classement par l'algorithme est comparé au vrai classement. Un arbre classant trop souvent un message régulier en spam ou l'inverse est disqualifié. Parmi les arbres restants, on choisira celui possédant le moins de nœuds, c'est-à-dire celui parvenant à prédire la nature du message en un minimum de questions. L'arbre gagnant est donc celui qui pose les questions les plus pertinentes pour déterminer le plus rapidement possible la nature des messages. On retrouve ici la stratégie du jeu *Qui est-ce ?* !

Améliorer le filtrage

La longévité des arbres de classification, utilisés depuis maintenant 25 ans, ainsi que la diversité des applications qui leur ont été trouvées (dans les domaines de la biologie, de l'écologie ou de la médecine par exemple, voir encadré) ne s'expliquent pas uniquement par le fait que les performances obtenues en pratique par cet algorithme sont satisfaisantes. En effet, du point de vue mathématique, l'algorithme pose bien des questions : peut-on l'utiliser pour tout problème de classification, ou existe-t-il des problèmes où l'algorithme se révélerait inefficace ? Peut-on démontrer que les arbres de classification sont meilleurs (en termes de performance de classement) que d'autres algorithmes couramment utilisés ? Gagnerait-on à utiliser plusieurs arbres de classification plutôt qu'un, comme lorsque l'on convoque un panel d'experts plutôt qu'un seul ? De nombreux articles ont été consacrés à ces questions, et plus généralement à déterminer les propriétés des arbres de classification, et ont permis de justifier théoriquement leur utilisation. Mais ces recherches mathématiques sur les arbres de classification, loin de ne servir qu'à valider la méthode initiale, ont aussi permis aux chercheurs de suggérer des pistes pour faire évoluer l'algorithme et en améliorer les performances. Ainsi, l'application des arbres de classification au filtrage de spams s'est considérablement raffinée au fil des ans. D'un unique arbre de classification, on est passé à une forêt (méthode des *random forests*, cf. encadré), et les arbres sont devenus dynamiques : au fur et à mesure que le propriétaire de la messagerie reçoit des messages, chaque arbre continue de collecter les informations sur les fréquences de mots pour améliorer ses performances de classification. Les arbres de classification, et plus généralement l'ensemble des algo-

rithmes de classification constituent donc un sujet de recherche très actif, auquel de nombreuses conférences mathématiques internationales et prestigieuses sont dédiées.

Au fur et à mesure que le propriétaire de la messagerie reçoit des messages, chaque arbre continue de collecter les informations sur les fréquences de mots pour améliorer ses performances de classification.

Du coté des émetteurs de spams, on n'est pas en reste, et les spams sont aujourd'hui rédigés de manière à contourner les filtrages. Les mots susceptibles de trahir la nature du message sont soigneusement évités, ou volontairement mal orthographiés pour rendre leur identification par l'algorithme de classification plus difficile. La course entre spammeurs et filtreurs ne fait donc que commencer, et continuera à l'avenir de poser aux futurs mathématiciens des problèmes passionnants… et critiques pour le développement informatique et économique de nos sociétés !

La méthode des random forests

Par Robin Genuer, *maître de conférences à l'Université Bordeaux Segalen*

La méthode des *random forests* (forêts aléatoires) peut s'avérer très utile dans des applications en médecine. Par exemple, lorsqu'on étudie les facteurs responsables d'une maladie et qu'on dispose d'information à propos des gènes de sujets atteints ou non de cette maladie. Le but est alors de rechercher quels sont les gènes qui permettent de distinguer au mieux les sujets malades des sujets sains.

Cependant, la plupart du temps, nous disposons d'informations sur des dizaines de milliers de gènes, et donc d'autant de questions potentielles pour découper un nœud de l'arbre. Il est alors très difficile de proposer un bon arbre « en un coup » (faire une partie de *Qui est-ce ?* avec 10000 questions possibles semble en effet assez ardu !).

L'idée des random forests est de construire une collection d'arbres où pour chaque arbre on se restreint à un petit paquet de questions. Après avoir mis en commun tous ces arbres, on peut évaluer le classement donné par la forêt (sur un échantillon test).

De nombreuses études montrent qu'une forêt semble mieux se comporter qu'un seul arbre pour ce type de problème (même si les raisons de ce phénomène restent encore assez méconnues).

Pour aller plus loin

Breiman L., Friedman J., Olshen R., Stone C., (1984) *Classification And Regression Trees*, Chapman & Hall.

Article Spam de Wikipédia, (http://fr.wikipedia.org/wiki/Spam).

European Network and information Security Agency (ENISA), *Spam Survey*, 16 décembre 2009

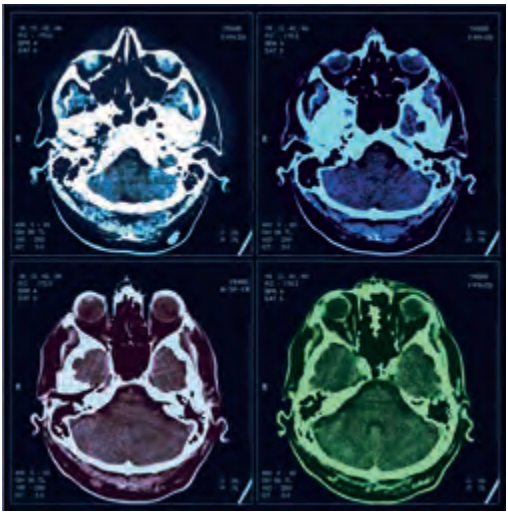
Surhone L.M., Tennoe M.T., Henssonow S.F., (2010) *Random Forest*, Verlag.

L'art de couper les têtes sans faire mal

Erwan Le Pennec, *chargé de recherche Inria à l'Université Paris-Sud*

Le principe du scanner implique de savoir retrouver un objet à partir d'une collection de radiographies de cet objet. Il s'agit de ce que l'on appelle, en mathématiques, un problème inverse. Sa résolution s'avère difficile et constitue toujours une source de questions pour les mathématiciens.

Il paraît difficile de « couper » une tête pour en voir l'intérieur sans faire mal à son propriétaire. C'est pourtant ce que les équipes médicales peuvent réaliser à l'aide des scanners, ou tomodensitomètres en français. Cet outil de diagnostic des dommages internes est disponible depuis les travaux de Godfrey Hounsfield et Allan McLeod Cormack, deux physiciens qui ont partagé le prix Nobel en 1979 pour cette invention. Le scanner n'aurait pu exister sans le mariage de la physique, des mathématiques et de l'informatique.



Problème direct, problème inverse

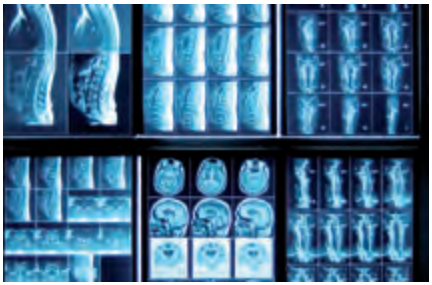
A l'inverse de la radiographie classique à rayons X, l'image d'un scanner n'est pas obtenue directement par une mesure physique mais est créée par un ordinateur à l'aide d'algorithmes mathématiques à partir d'une collection de radiographies de la tête du patient prise sous des angles différents. Pour réaliser une telle prouesse, il faut savoir résoudre deux problèmes : comprendre comment cette collection de radiographies dépend de l'intérieur de la tête du patient, on parle de *problème direct*, et comprendre comment remonter de la collection de radiographies à ce qui se passe à l'intérieur de la tête, on parle alors de *problème inverse*.

Le problème direct est ici le même que celui de la radiographie à rayons X. Son principe a été découvert par Wilhelm Conrad Roentgen en 1895 et a immédiatement été utilisé pour diagnostiquer des blessures internes. Il a été largement étudié d'un point de vue physique : on sait ainsi que la radiographie obtenue dépend essentiellement de la quantité de matière absorbante rencontrée par le rayonnement X entre la source et le capteur. C'est pour cela que les os, plus absorbants que les tissus, sont les structures naturelles les plus visibles. En prenant en compte les caractéristiques physiques des instruments de radiographie utilisés, on sait alors prévoir mathématiquement de manière très fine la radiographie d'un objet supposé entièrement connu. Il en est donc de même pour la collection de radiographies obtenue par le scanner.

Bien que très précise, cette modélisation n'est pas parfaite : la radiographie mesurée sera légèrement différente de la radiographie prédite du fait de simplifications physiques, de l'imprécision des différents capteurs et d'autres aléas.

Retrouver l'objet à partir des radiographies

Retrouver l'objet à partir de la collection de radiographies, le problème inverse, s'avère être un problème beaucoup plus difficile. Il faut réussir à inverser la procédure d'acquisition du scanner par un principe implémentable dans un ordinateur. De plus, il faut que de petites imprécisions sur le modèle n'induisent pas de grandes erreurs sur le résultat. C'est la mise en œuvre d'une solution à ce problème qui a permis l'existence de la première génération de scanners à la fin des années 60. Depuis lors, les évolutions technologiques des sources de rayons X, des capteurs ainsi que des autres mécanismes instrumentaux ont permis une amélioration de la qualité des images obtenues. Celle-ci aurait été bien moindre sans le travail des mathématiciens et le développement de nouveaux algorithmes d'inversion.



Un principe simple, souvent appelé principe du *rasoir d'Ockham*, est à la source de la plupart des méthodes modernes pour résoudre ce problème. Guillaume d'Ockham, un moine philosophe franciscain du XIVe siècle, expliquait que « les multiples ne doivent pas être utilisés sans nécessité ». Dans son incarnation moderne, le rasoir d'Ockham est un principe de parcimonie : une bonne solution est une solution *cohérente* avec les données observées tout en n'étant pas inutilement *complexe*.

Pour appliquer ce principe de bon sens, il faut bien sûr donner une définition précise (et mathématique) de la « cohérence » et de la « complexité », ainsi que proposer une manière d'obtenir une solution à la fois cohérente et peu complexe.

Une bonne solution est une solution cohérente avec les données observées tout en n'étant pas inutilement complexe.

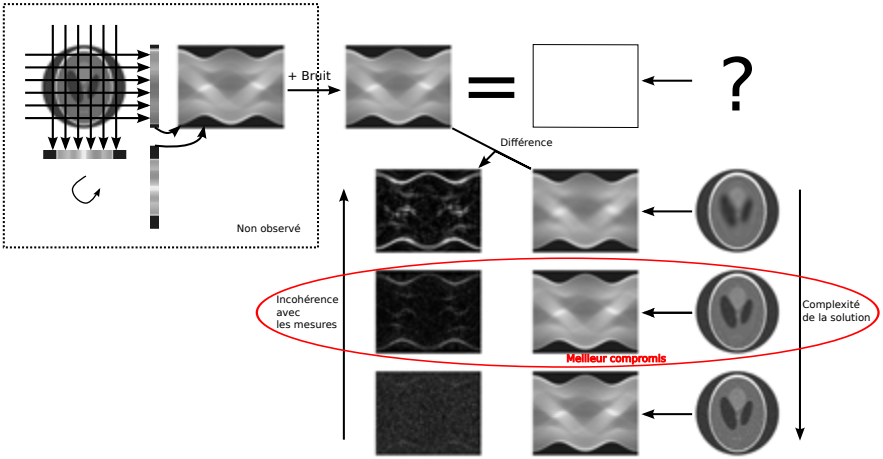
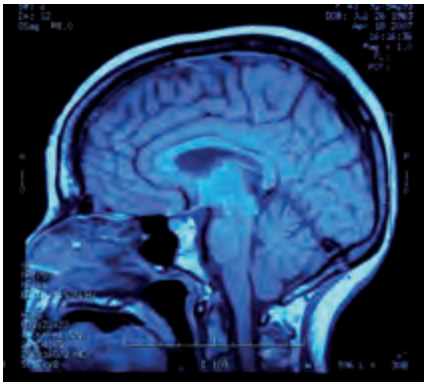


Figure 1. Complètement à gauche se trouve une image test appelée fantôme de Logan-Shepp, souvent utilisée pour évaluer les méthodes ; l'image immédiatement à sa droite correspond aux mesures d'un scanner parfait (cette image bidimensionnelle est obtenue en juxtaposant les « tranches » correspondant à différents plans de coupe). L'image suivante correspond à des mesures légèrement perturbées ; enfin on a représenté au dessous trois équilibres différents entre l'incohérence (l'image de la colonne de gauche est d'autant plus blanche que l'incohérence avec les mesures est grande) et la complexité (l'image de la colonne de droite correspond à des solutions de complexité croissante du haut vers le bas).

Aucune des méthodes de résolution du problème inverse ne faisant l'unanimité, les mathématiciens, ainsi que les fabricants de scanners, continuent d'explorer les possibilités. L'une d'entre elles, sur laquelle a travaillé l'auteur de cet article, est illustrée par la figure 1.

On observe que la *meilleure* solution semble bien être celle qui réalise le *meilleur* compromis entre cohérence et complexité.



IRM, imagerie sismique et autres problèmes similaires

De nombreux problèmes similaires existent en imagerie, on peut citer par exemple :

– l'IRM (Imagerie en Résonance Magnétique), où l'on cherche à déterminer les structures à l'intérieur d'un corps en mesurant leurs influences sur des champs magnétiques,

– la *déconvolution optique*, où l'on cherche à enlever le flou dû à la lentille dans une photographie,

– l'imagerie sismique, où l'on cherche à déterminer les structures souterraines à partir d'enregistrements de propagations d'ondes.

Dans chacun d'eux, on retrouve un problème direct, l'étude de la formation des mesures à partir d'un objet, et un problème inverse, dans lequel on cherche à revenir des mesures à l'objet. Pour chacun d'eux, le principe du rasoir d'Ockham a prouvé son efficacité pour obtenir des résultats théoriques et pratiques. Les questions autour de ces problèmes inverses sont cependant loin d'être closes : de nouveaux dispositifs d'acquisition sont régulièrement introduits, de nouveaux moyens de calcul toujours plus puissants permettent d'envisager des méthodes de plus en plus complexes, des nouveaux outils mathématiques permettent également d'explorer de nouvelles voies…

Pour aller plus loin…

La modélisation la plus commune du problème évoqué dans cet article est l'observation d'un objet O à travers sa collection de radiographies R obtenue par l'application d'une fonction Φ venant du modèle direct auquel est ajoutée une perturbation B inconnue. On observe ainsi

$$R = \Phi(O) + B$$

et l'on souhaite retrouver au mieux O à partir de la collection de radiographies R.

Une idée naturelle pour un mathématicien est de déterminer l'inverse Φ^{-1} de la fonction Φ et de l'appliquer à R pour retrouver O. Malheureusement, on vérifie que l'objet ainsi obtenu

$$\Phi^{-1}(R) = \Phi^{-1}(\Phi(O)+B))$$

peut être très différent de O même si B est petit. On dit que l'inversion est *instable*.

Pour appliquer le principe d'Ockham, il faut définir le critère de *cohérence* ainsi que de *complexité*.

Le mathématicien définit ainsi une fonction d'incohérence $F(R,O)$ qui quantifie la différence entre la radiographie observée R et celle prévue par le modèle direct pour un objet O : cette fonction est petite lorsque l'objet O est cohérent avec les observations et grande lorsqu'il est incohérent.

De même, il définit une mesure de complexité $C(O)$ grande quand l'objet est complexe et petite quand l'objet simple.

Il peut alors interpréter le principe d'Ockham comme la recherche d'un objet O réalisant le *meilleur* compromis entre ces deux quantités, un objet pour lequel la somme

$$F(R,O) + C(O)$$

est la plus petite possible.

Les expériences numériques de la figure ont été obtenues en choisissant pour critère de cohérence essentiellement la distance usuelle entre l'image de l'objet $\Phi(O)$ et R, et pour critère de complexité, la parcimonie, le nombre de coordonnées non nulles, dans une représentation adaptée. Ce choix permet d'obtenir des garanties théoriques sur la reconstruction ainsi qu'un algorithme efficace de détermination du meilleur compromis.

Climatologie et statistiques

Pascal Yiou, *directeur de recherche au CEA*
Philippe Naveau, *chargé de recherche CNRS à l'Institut Pierre Simon Laplace*

L'étude du climat et de ses variations est basée sur un grand nombre d'outils et concepts statistiques. Ceci se reflète dans le langage employé par les différents rapports du GIEC (Groupe International d'Experts sur le Climat), qui mettent largement l'accent sur les incertitudes et leur quantification.

L'estimation du changement climatique au cours des derniers siècles, et la mise en perspective du comportement des dernières décennies dans ce contexte pluri-millénaire, constituent un des grands défis scientifiques actuels. Quand on regarde des courbes de températures moyennes d'observations, sur le globe ou sur chaque hémisphère (ou sur des sous-régions), on constate une augmentation générale au cours du XX^{e} siècle, avec une accélération à partir des années 1970. Un grand nombre d'arguments physiques et thermodynamiques, ainsi que des mesures, montrent que cette augmentation est liée à l'activité humaine, responsable du rejet de gaz à effet de serre dans l'atmosphère.

Climat *vs.* météo

La première manifestation des statistiques en climatologie se situe dans la définition elle-même du climat, et la différence entre météorologie et climatologie. Pour reprendre un aphorisme célèbre, attribué à E.N. Lorenz, le météorologue à l'origine de l'expression « effet papillon » et de l'étude du chaos en sciences atmosphériques : « climate is what you expect, meteorology is what you get », soit « le climat est ce qui arrive en général, la météorologie est ce qui arrive vraiment ». Climatologie et météorologie utilisent et décrivent les mêmes variables physiques (température, pression, vitesse de vent, précipitations…), mais les météorologues s'intéressent à l'état des variables à un endroit donné et à un instant donné,

alors que les climatologues regardent des grandeurs moyennes, sur une saison ou une région. Ces grandeurs moyennes peuvent être les variables physiques habituelles (par exemple, la température moyenne en hiver sur l'Île-de-France), ou des grandeurs plus sophistiquées, comme la variance de la vitesse du vent ou l'ensoleillement moyen sur une région. En général, les variables climatiques sont définies sur des périodes de 30 ans (par exemple entre 1961 et 1990) pour des raisons heuristiques liées à l'inférence statistique des grandeurs moyennes. C'est pour cette raison qu'il n'est pas raisonnable de comparer des tendances de température sur une décennie ou moins, avec des tendances estimées sur l'ensemble du XX^e siècle.

La première manifestation des statistiques en climatologie se situe dans la définition elle-même du climat.

L'évolution des températures moyennes

Une question scientifique légitime est de savoir si les valeurs de température observées maintenant ont pu être vécues au cours du dernier millénaire, alors que la configuration des continents et celle de la position de la Terre autour du soleil étaient assez semblables. Pour accomplir cette tâche, alors qu'il n'y a pas d'enregistrements thermométriques avant le XVII^e siècle, on utilise des indicateurs climatiques à base de cernes d'arbres, de sédiments lacustres ou marins, de carottes de glace, ou encore des témoignages écrits (comme les dates de vendanges). La combinaison de ces enregistrements indirects se fait par des techniques statistiques multi-variées parfois sophistiquées (par exemple des techniques d'*ondelettes*, des *modèles bayésiens hiérarchiques*, des *composantes principales*…) qui permettent de reconstituer un signal interprété comme une variation de température au cours des derniers siècles. Le résultat le plus frappant de ces études, réalisées par une quinzaine d'équipes de recherche dans le monde, est la robustesse de la forme des variations de températures. Cette forme est dite en « crosse de hockey » : on observe une légère décroissance de la température entre l'an Mil et 1900, puis une augmentation rapide au cours du XX^e siècle (figure 1). Cette crosse de hockey est un peu tordue, avec un optimum médiéval climatique chaud (entre 1000 et 1400) et un petit âge de glace froid (entre 1500 et 1800). L'utilisation de méthodes statistiques permet d'obtenir des intervalles de confiance pertinents pour répondre à la question : « Quelle est la

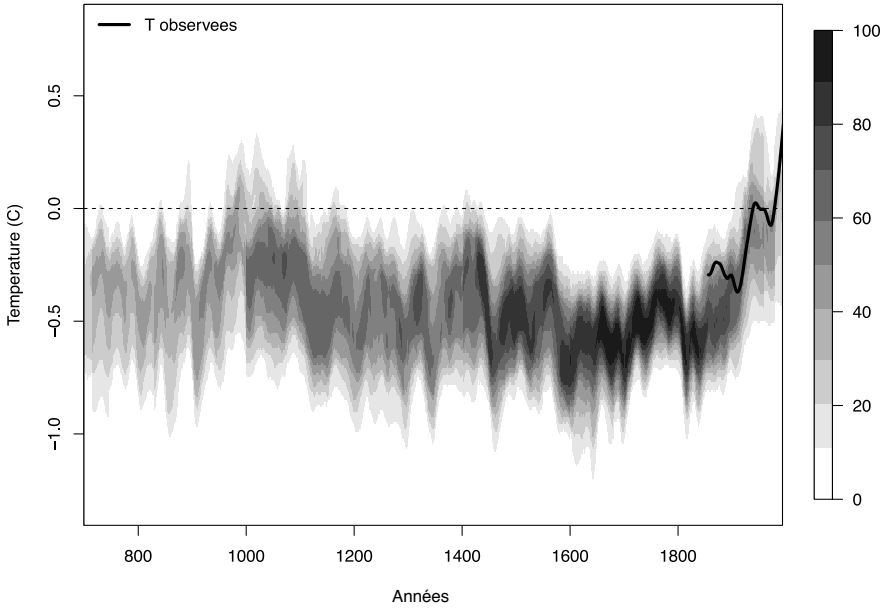


Figure 1. Recouvrement des reconstructions de températures de l'hémisphère nord depuis l'an 800. Les températures sont exprimées en différence par rapport à la moyenne de 1961 à 1990 (train pointillé). Les niveaux de gris indiquent les probabilités (en %) que les reconstructions de température pour chaque année se trouvent dans une gamme de température. La courbe en trait plein indique la moyenne des observations de température sur l'hémisphère nord (adapté du rapport de l'IPCC 2007).

probabilité pour que la température actuelle fasse partie de la gamme de températures reconstruites depuis l'an Mil ? » La réponse : une probabilité excessivement faible, quelle que soit l'approche statistique utilisée dans la reconstruction ! Autrement dit, il n'y a aucune chance que les changements climatiques actuels résultent d'une évolution normale du climat.

Quelle est la probabilité pour que la température actuelle fasse partie de la gamme de températures reconstruites depuis l'an Mil ?

Variations autour de l'état moyen

Le corollaire de l'étude de l'évolution des températures moyennes est celle des variations autour de l'état moyen : peut-on dire qu'on observe de plus en plus d'événements extrêmes (comme des canicules, précipitations intenses, etc.) vers la fin du XX^{e} siècle ?

Pour répondre à cette question, les climatologues peuvent utiliser des modèles statistiques construits pour la circonstance. Cette approche nécessite la disponibilité d'observations climatiques de qualité et de durée suffisante.

Si on s'intéresse aux vagues de chaleur, de froid ou aux précipitations intenses, le cadre statistique qui s'impose est celui de la Théorie des Valeurs Extrêmes, développée par Emile Gumbel, dans les années 1940. Pour une variable aléatoire continue X, l'objet de cette théorie est de décrire la répartition des grandes valeurs. L'enjeu principal est de décrire les extrêmes d'une loi que l'on ne connaît pas a priori (ce qui est le cas de la plupart des variables que l'on mesure).

Théorie des valeurs extremes

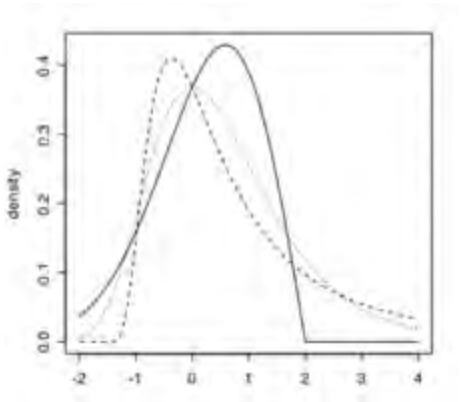


Figure 2. Densité de probabilité des lois de Fréchet (--), Weibull (-) et Gumbel (..).

Maximum de *n* variables aléatoires → *n* tend vers l'infini → **Loi de Fréchet, Weibull ou Gumbel**

M-R. Fréchet

Des mathématiciens des années 70 ont résolu cette question en énonçant que si la loi du maximum de X converge vers une loi, alors cette loi est une des trois lois d'extrêmes : les lois de Gumbel, Weibull et Fréchet (qui diffèrent essentiellement par la manière dont leur densité se comporte pour les très grandes valeurs). Pour en revenir au climat, on s'intéresse au type de loi de probabilité que suivent les températures maximales à chaque saison. Dans la plupart des régions, les lois du maximum de température suivent des lois de Weibull, c'est-à-dire que la loi du maximum décroît vers 0 plus vite qu'une loi exponentielle (ou une gaussienne). En revanche, les lois du maximum de précipitation suivent des lois de Gumbel ou de Fréchet : les valeurs peuvent être très grandes de manière relativement fréquente.

E-J. Gumbel

Le dernier rapport spécial du GIEC sur les extrêmes montre que le nombre de vagues de chaleur en Europe a crû au XX^{e} siècle. C'est aussi le cas des précipitations intenses dans le sud de l'Europe, même si cette région connaît également une augmentation du nombre de jours sans pluie.

W. Weibull

Les manières dont les propriétés de ces lois d'extrêmes évoluent dans le temps sont des sujets de recherche très actifs dans le domaine de la climatologie statistique. Une estimation de ces extrêmes, et des incertitudes associées, est indispensable pour faire des prévisions climatiques dont se servent des secteurs économiques aussi variés que l'énergie, l'agriculture, l'assurance et les transports.

Pour aller plus loin

Jeandel C., Mosséry R., (éd.) (2011). *Le Climat à Découvert : Outils et méthodes en recherche climatique*, CNRS Éditions, Paris.

Chercher sur le Web : juste un point fixe et quelques algorithmes

Serge Abiteboul, *directeur de recherche Inria à l'École Normale Supérieure de Cachan*

Le Web met à notre disposition une masse considérable d'information, plusieurs dizaines de milliards de documents. Sans les moteurs de recherche, ces systèmes de plus en plus sophistiqués qui nous aident à nous focaliser sur un petit nombre de pages, le Web ne serait qu'une poubelle à ciel ouvert, gigantesque et inutilisable. Le rôle de ces systèmes est de faire surgir de la masse des internautes une intelligence collective pour évaluer, classer, filtrer les informations. Comment les moteurs de recherche gèrent-ils ces volumes d'information véritablement phénoménaux ? Comment aident-ils les utilisateurs à trouver ce qu'ils cherchent dans cette masse ? Retour sur un des plus beaux succès du Web.

Ce Web introduit par Tim Berners-Lee vers 1990 auquel nous nous sommes si rapidement habitués est fait de documents hypermédia. L'information est en langue naturelle (voir l'encadré *Les langues du Web*), non pas en langage informatique, et les textes sont vaguement structurés avec les balises HTML pour, par exemple, des titres ou des énumérations. Des ancres sur lesquelles il peut cliquer permettent à l'internaute de découvrir des images, de la musique, des films. Elles lui permettent aussi de passer de page en page au gré de son humeur. Et pour trouver de l'information dans ce bazar, quoi de plus simple ? Il suffit d'utiliser cette petite merveille technologique qu'est un moteur de recherche du Web. L'internaute choisit quelques mots clés. Le moteur sait retrouver en un instant les pages qui hébergent ces mots. La magie, c'est qu'il sait aussi proposer parmi les dizaines voire centaines de millions de pages possibles, les quelques pages qui contiennent si souvent ce que l'internaute recherche.

Les langues du Web

En 2012, l'anglais représente environ 45 % des pages du Web, l'allemand 7 %, le français est la troisième langue avec 5 % ; 90 % des langues du monde ne sont pas représentées sur Internet.

Un index du Web

Tout débute par un index. Il s'agit d'une liste ordonnée dans le style d'un index en fin de livre, qui associe à chaque mot la liste des pages du Web qui contiennent ce mot. Par exemple, une entrée dans cet index serait :

Casablanca → http://www.imdb.com/title/tt0034583/

qui indique que le mot « Casablanca » est présent dans cette page Web du site imdb, une base de données en ligne sur le cinéma. Si vous donnez à l'index plusieurs mots comme « Casablanca Bogart Bergman », il retournera la liste des pages du Web qui contiennent l'un de ces trois mots, et notamment la page précédente. La difficulté essentielle ici est la taille de cet index, qui représente des téraoctets de données (un téraoctet = mille milliards d'octets = un million de méga-octets) pour quelques milliards de pages Web. Le serveur qui gère un tel index va connaître des problèmes de passage à l'échelle :

1. Pour indexer plus de pages, le serveur a besoin de plus en plus de stockage pour garder l'index, et chaque requête devient de plus en plus coûteuse à évaluer.
2. Si le nombre d'utilisateurs croît, le serveur reçoit de plus en plus de requêtes.

Dans les deux cas, le serveur est vite submergé. Pour résoudre ce problème nous allons utiliser une technique fondamentale de l'informatique, la technique du hachage. L'échelle du Web nécessite de paralléliser cette technique.

Pour répartir l'index sur, disons, $K = 10$ machines M_1, M_{10}… nous allons utiliser une fonction aléatoire H qui transforme un mot en un entier entre 1 et 10. (Cette fonction est appelée fonction de hachage.) La machine

numéro $H(w)$ sera celle chargée du mot w. Par exemple, les données du mot « France » sont stockées sur la machine H(« France »), disons M_7. Si H est bien aléatoire, les données seront donc partagées relativement équitablement entre les dix machines, ce qui résout le premier problème. Supposons que quelqu'un veuille les données correspondant à France, il interroge seulement la machine M_7. Les requêtes aussi sont donc partagées relativement équitablement entre les dix machines, ce qui résout le second problème.

La taille des données ou le nombre d'utilisateurs peuvent croître, il suffit alors d'adapter le système en augmentant le nombre de machines. Le parallélisme nous a tirés d'affaire et nous permet de passer à l'échelle supérieure. Par exemple Google utilise des milliers de machines dans des « fermes » et disperse ses dizaines de fermes aux quatre coins du monde.

Pourquoi est-ce que cela marche ? Grâce au parallélisme. De manière générale, peut-on prendre n'importe quel algorithme et l'accélérer à volonté en utilisant plus de machines ? La réponse est non ! La recherche a montré que tous les problèmes ne sont pas aussi parallélisables, ou bien pas parallélisables de manière aussi simple. Mais il se trouve que la gestion d'index est un problème très simple, très parallélisable : nous pouvons sans frémir envisager d'indexer de plus en plus de pages, des dizaines de milliards ou plus.

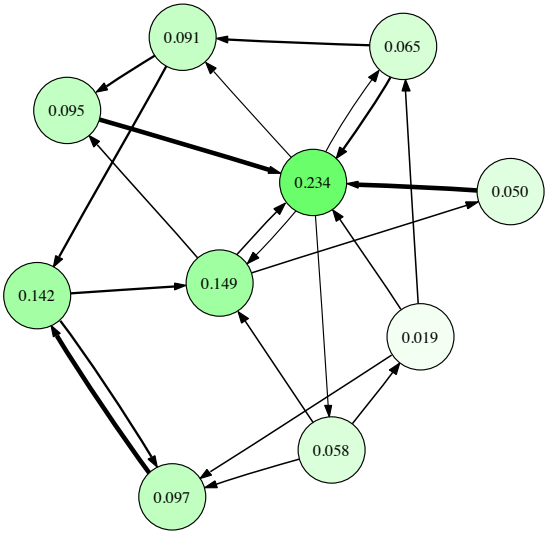


Figure 1. Pagerank : calcul de la popularité des images du web. Chaque cercle représente une image, les flèches et leur épaisseur représentent le nombre de liens d'une page vers l'autre. Les nombres dans chaque cercle, et la couleur, indiquent la popularité.

PageRank

Le cœur du problème reste maintenant de choisir, parmi les millions de pages qui contiennent les mots d'une requête, lesquelles sont les plus intéressantes. C'est essentiel car un utilisateur ira rarement au-delà des dix ou vingt premiers résultats affichés.

Au départ, les moteurs de recherche comme AltaVista utilisaient pour classer les pages uniquement des techniques statistiques de recherche d'information. Une page était jugée plus intéressante si le terme apparaissait dans un titre, ou en caractère gras. Plus un terme est répété dans le document plus il « pèse ». Et également, plus le terme est rare dans l'ensemble des documents, plus il pèse lorsqu'il apparaît. Ce genre de techniques qui marchent bien sur de petits corpus s'avéraient assez décevantes pour le Web.

Les jeunes créateurs de Google ont eu l'idée de baser également le choix des pages sur une technique connue en mathématiques, la marche aléatoire. C'est cette idée inspirée de travaux de Jon Kleinberg qui est à l'origine du succès de Google, un des succès industriels les plus étonnants de l'histoire de l'humanité.

La marche aléatoire

Imaginez un surfeur du Web. Il part d'une page du Web, disons la page www.inria.fr. Et puis il se balade sur le Web en choisissant à chaque étape de cliquer sur un des liens sortant de la page. Si la page n'a pas de lien sortant, il choisit aléatoirement une page n'importe où sur le Web. Et il continue encore et encore, pour toujours, un temps infini. Quel est à l'infini sa probabilité de se trouver sur une page précise ? C'est ce que nous définirons comme la popularité de cette page (cf. figure 1). sur un micro-Web. A priori, une définition abstraite, un joli concept de mathématiques totalement inutile ? Et pourtant ! En pratique cette popularité correspond assez bien à la « popularité » de la page et à ce que les internautes attendent le plus souvent.

Le point fixe

Reste à calculer cette popularité. Pour cela, il faut écrire une équation dont la popularité sera la solution, équation qui code le comportement du surfeur aléatoire. Au cours

des actions du surfeur, chaque page va transférer sa popularité vers les pages vers lesquelles elle pointe (si une page est un cul-de-sac qui ne conduit nulle part, elle partage sa popularité entre toutes les autres pages). En ignorant quelques détails, cela nous conduit à une matrice Q qui capture ces « échanges » de popularité. Le vecteur des popularités pop (la liste de toutes les popularités $pop[i]$ pour chaque page i) se trouve être la solution de l'équation de point fixe :

$$pop = Q \times pop,$$

une notation bien compacte pour un système de dix milliards d'équations à dix milliards d'inconnues (une inconnue pour chaque page Web)…

Et là, banco ! Une technique connue va nous permettre de calculer cette solution. Partons du vecteur pop_0 défini par

$$pop_0[i] = 1/N,$$

c'est-à-dire qu'au départ toutes les pages sont supposées aussi importantes. Et définissons :

$$pop_1 = Q \times pop_0; \quad pop_2 = Q \times pop_1;$$
$$pop_3 = Q \times pop_2 \ldots$$

En continuant ainsi, on arrive assez rapidement à un point fixe qui se trouve être la solution de notre équation. On a donc calculé le vecteur de popularité ! (En pratique six ou sept itérations suffisent pour arriver à une convergence suffisante.)

Pour réaliser ce calcul efficacement avec des volumes de données pareils, il faut des algorithmes très sophistiqués, un engineering de fou. Ce n'est peut-être plus des mathématiques mais c'est de l'informatique de toute beauté. J'ai pu crawler le Web et implémenter un tel algorithme de PageRank avec des étudiants. Cela a été une de mes plus fantastiques expériences de chercheur.

Des recherches à poursuivre

Nous avons présenté une version hyper simplifiée d'un moteur de recherche. Par exemple, celui de Google considère que les textes qui apparaissent près d'un pointeur vers une page font partie du contenu informatif de cette page. C'est le point exploité par le « bombing » (*cf.* encadré *Trouver Chuck Norris*). Les moteurs de recherche doivent se sophistiquer sans cesse, ne serait-ce que pour contrer les attaques des bombeurs qui cherchent à manipuler le Web ou des spamdexeurs qui trichent pour être mieux représentés. Le PageRank de Google actuel utiliserait des dizaines de critères combinés dans une formule gardée secrète. Mais ces moteurs posent encore des problèmes essentiels. Pour n'en citer que quelques-uns :

– Une mesure comme PageRank privilégie la popularité et donc a pour effet d'encourager l'uniformité, les pages populaires devenant de plus en plus populaires et les autres sombrant dans l'anonymat. Est-ce vraiment souhaitable ?

– Pour interroger les moteurs du Web, nous utilisons une série de mots-clés, une « langue » primitive quasiment sans gram-

maire. Ne pourrions-nous pas faire beaucoup mieux ?

– Faut-t-il exclure des pages, parce qu'elles sont racistes, vulgaires, fausses, ou pour favoriser un client, ou ne pas déplaire à un gouvernement ?

– Il y a quelque chose d'extrêmement embarrassant dans la puissance considérable que les moteurs de recherche ont de par leur contrôle de l'information. Devons-nous leur faire confiance sans comprendre leur classement ? Et pourquoi ce secret ?

Pour gérer les volumes d'information du Web, de nouvelles techniques utilisant plus à fond le parallélisme sont sans cesse proposées. De nombreuses techniques sophistiquées sont étudiées pour évaluer, classer, filtrer l'information. Nous pourrions citer par exemple les techniques exploitant les systèmes de notation (l'internaute est invité à noter des services comme dans eBay), les systèmes de recommandation automatisée (comme Netflix ou Meetic), les systèmes qui évaluent l'expertise d'internautes (comme Mechanikal Turk). Le domaine est extraordinairement actif.

Trouver Chuck Norris

À l'heure où ce texte est écrit, si vous tapez « trouver Chuck Norris » sur le moteur de recherche de Google, le premier résultat vous emmène à une page qui dit « Google ne recherchera pas Chuck Norris car il sait que personne ne peut trouver Chuck Norris, c'est lui qui vous trouvera. » Pour arriver à cela, de nombreux internautes ont créé des pages Web qui pointent vers cette page avec comme légende : « trouver Chuck Norris ». Ils sont arrivés à redéfinir collectivement l'expression « trouver Chuck Norris ».

Juste un point fixe et quelques algorithmes

Je me trouvais dans le groupe de recherche sur les systèmes d'information à Stanford en 1995 quand deux jeunes étudiants, Sergei Brin et Larry Page, y travaillaient sur le prototype du moteur de recherche Google. Ils ont développé l'algorithmique nécessaire, mais leurs propositions pouvaient passer pour farfelues. Elles auraient été irréalistes quelques années plus tôt, quand les tailles des mémoires et leur prix auraient demandé d'utiliser un nombre improbable de machines hyper coûteuses. Dans les années 95, cela devenait possible avec un nombre raisonnable de machines bon marché. Les utilisateurs allaient plébisciter leur moteur de recherche. Comme base à ce succès extraordinaire, nous pourrions mentionner un engineering exceptionnel pour faire fonctionner des milliers de machines 24 heures sur 24, des modèles commerciaux révolutionnaires, des techniques de management originales fondées sur un culte de la créativité. Mais en ce qui me concerne, je préfère me rappeler qu'au début, il y avait juste un point fixe et quelques algorithmes.

Pour aller plus loin

Brin S., Page L., (1998). *The Anatomy of a Large-Scale Hypertextual Web Search Engine*. WWW Conference.

Abiteboul S., Manolescu I., Rigaux P., Rousset M.-C., Senellart P., *Web data management*, Cambridge University Press. http://webdam.inria.fr/Jorge

Kleinberg J., (1999). *Authoritative sources in a hyperlinked environment. Journal of the ACM*.

Abiteboul S., Preda M., Cobena G., (2003). *Adaptive On-Line Page Importance Computation*. WWW Conference.

La brouette de Monge ou le transport optimal

Yann Brenier, *directeur de recherche CNRS à l'École polytechnique*

Né d'un problème concret – comment déplacer au mieux un tas de sable – le transport optimal est un outil qui trouve des applications aussi bien à l'intérieur des mathématiques (de la géométrie à l'analyse fonctionnelle) que dans d'autres domaines, comme la gestion de ressources, par exemple.

Dès le XVIII[e] siècle, le mathématicien français Gaspard Monge, dans son *Mémoire sur la théorie des déblais et des remblais*, étudiait un problème des plus concrets (déplacer au mieux un tas de sable !) en lui appliquant une méthode rigoureuse, « optimale » dirions-nous aujourd'hui. Ce problème est l'un des premiers exemples de *recherche opérationnelle*, la branche des mathématiques qui s'intéresse aux méthodes pour traiter de manière efficace des problèmes combinatoires (voir l'encadré *La combinatoire*) et qui est une théorie encore très vivace aujourd'hui.

En quoi ce problème (qu'on pourrait appeler « Brouette de Monge », le but étant de déplacer du sable) est-il un problème combinatoire ? Nous pouvons le lire de la façon suivante. Nous avons un tas de sable, composé de petits éléments (un élément étant un grain ou une poignée). Nous avons aussi un trou à combler, dont on connaît la forme et dont le volume est supposé coïncider avec celui du tas. On peut penser le trou comme composé de petites cases, chacune pouvant contenir un élément de sable. La question est : où envoyer chaque élément pour minimiser les chemins parcourus au total ? On peut essayer de répondre à partir d'une description combinatoire du problème, mais dans certains cas il peut être plus utile d'en donner une description différente, basée sur les formes du profil du tas et du trou, considérées comme des fonctions. Et nous allons voir aussi que ce

problème d'optimisation peut tout à fait être utilisé pour parler d'autres choses, et pas seulement de sable à déplacer.

Des applications diverses

La question du transport a été réexaminée dans les années 1940 par Leonid Kantorovich (un des quelques mathématiciens lauréats du prix Nobel d'économie) à propos d'allocation optimale des ressources. Dans son œuvre, il cherche à résoudre des problèmes de ce genre : comment utiliser les rares facteurs de production (travail, capital…) afin de réaliser une satisfaction maximale des besoins ? Par exemple, étant données des mines produisant du métal et des usines le travaillant, on peut essayer de déterminer quelle mine doit approvisionner quelle usine pour minimiser les dépenses dues au transport du matériel : cette question est finalement la même que celle du tas de sable.

Bien d'autres questions (files d'attente, gestion de stock, etc.), que nous ne détaillerons pas ici, font partie de ce domaine. Il s'agit toujours de trouver comment passer de la situation initiale à la situation finale de la meilleure manière possible, c'est-à-dire par le meilleur chemin dans l'espace compliqué de tous les choix possibles.

Déplacer un nuage de points

Le point clé de la théorie est probablement le concept de *géodésique*, mot qui désigne le plus court chemin entre deux points. Prenons deux points dans le plan : quel est le meilleur chemin qui les relie ? La réponse est bien connue : c'est la ligne droite. Et si on se place sur une sphère ? Ce sont les arcs de cercle maximal : l'équateur, les méridiens (mais pas les parallèles) ; on le voit d'ailleurs si l'on pense à la trajectoire d'un vol transatlantique.

Ce qui est vrai pour des points peut être vrai pour des « nuages de points » ou d'autres objets : les manipuler au mieux revient à trouver le meilleur chemin possible pour amener ces objets tous ensemble de leur état initial à l'état final désiré.

On va considérer tout d'abord un exemple : le cas d'un nuage fini constitué de quatre points dans le plan (voir la figure 1).

On se donne la configuration du nuage à deux instants donnés, et on cherche comment le transporter de façon optimale entre les deux configurations. À l'instant de départ, le nuage est donc défini par une liste de quatre points A_1, A_2, A_3, A_4, et à l'instant final par quatre autres points B_1, B_2, B_3, B_4.

Un point clé à comprendre est qu'on ne se soucie aucunement de l'individualité des « particules » et que leur numérotation est arbitraire. Ainsi, il est possible de transporter, par exemple, A_1 vers B_1, A_2 vers B_3, A_3 vers B_4 et A_4 vers B_2. (En revanche, on n'est pas autorisé à déplacer A_1 et A_2 vers B_1 en laissant B_4 vacant, par exemple.) Il faudra donc chercher une solution optimale parmi tous les arrangements possibles.

Un arrangement correspond à une manière d'associer à chaque point A un point B, de

manière à ce que chaque point de départ soit connecté à un et un seul point à l'arrivée. Le nombre d'arrangements possibles des N premiers entiers s'appelle la factorielle de N et s'écrit $N!$ Il vaut 24 pour $N = 4$ (et plus de 20 milliards pour $N = 15$).

L'association σ correspondant à notre exemple (figure du milieu) s'écrit : $\sigma(1) = 1$, $\sigma(2) = 3$, $\sigma(3) = 4$, $\sigma(4) = 2$, alors que l'association optimale serait $\sigma(1) = 4$, $\sigma(2) = 2$, $\sigma(3) = 3$, $\sigma(4) = 1$.

Trouver la solution optimale revient à minimiser un *coût*, qui est choisi au cas par cas d'après les applications envisagés. Le coût de Monge était donné par la somme de toutes les distances entre les points de départ et les points d'arrivée correspondants. D'autres coûts sont possibles. Par exemple, on pourrait considérer la somme des carrés de ces distances (c'est le coût *quadratique*).

Un problème d'optimisation combinatoire

Tel que nous l'avons formulé, notre problème de transport optimal appartient à une branche importante des mathématiques : l'optimisation combinatoire. Dans cette discipline, notre problème est considéré comme « facile ». En effet, il existe des algorithmes qui permettent de trouver l'arrangement optimal en un nombre d'opérations proportionnel au cube du nombre de points (cela veut dire que, sur un ordinateur où les opérations s'effectuent séquentiellement, sans parallélisme, le temps de calcul pour N points vaudra plus ou moins une constante fois N^3).

Rappelons que, comme le nombre d'arrangements est $N!= N(N-1)x...x3x2x1$, il est

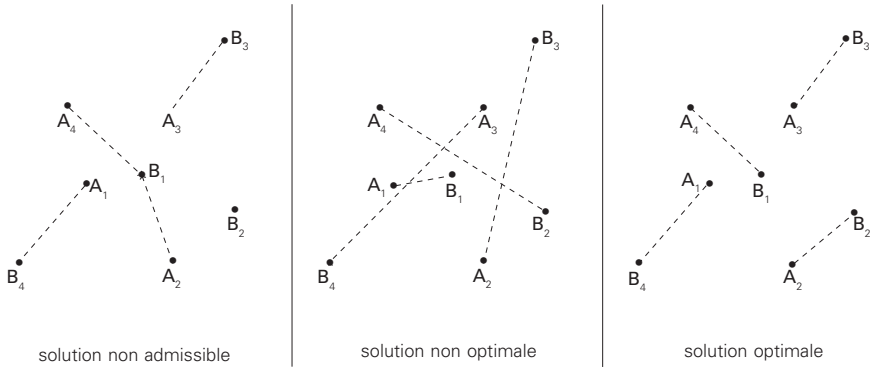


Figure 1. Transport d'un nuage de quatre points.

donc déjà remarquable de rendre le temps de calcul cubique par rapport à N (en effet, si pour N = 4 la factorielle vaut 24 et reste modérée, lorsque N est grand, elle augmente presque comme N^N, bien plus vite que N^3 en tous cas, et devient rapidement astronomique).

Cependant, même si on a remarqué que N^3 est beaucoup plus petit que $N!$, en pratique, les algorithmes connus sont peu performants lorsque N devient grand (au delà du millier, disons). On peut rêver d'un algorithme dont le coût de calcul croîtrait à peu près proportionnellement à N. Les applications en seraient spectaculaires. Mais pour approcher ce rêve, il faut changer de point de vue, et considérer non plus un nombre limité de points, mais des points en très grand nombre, voire une infinité de points. On passe alors d'un problème *discret* à un problème *continu*.

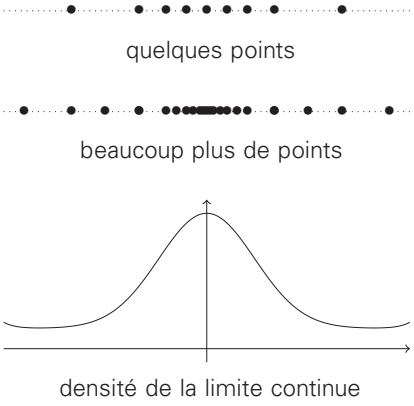


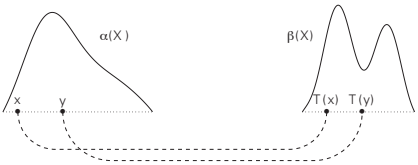
Figure 2. Cas d'un nuage continu de points

On peut rêver d'un algorithme dont le coût de calcul croîtrait à peu près proportionnellement à N. Les applications en seraient spectaculaires.

Transport optimal d'un nuage continu

Comme souvent en mathématiques et en physique, la limite continue du problème discret précédemment discuté permet de mettre en jeu toute la force du calcul différentiel et intégral (en particulier, les sommes que l'on effectue dans le cas discret sont remplacées par des intégrales dans le cas continu).

Dorénavant, notre nuage de points sera décrit, respectivement au départ et à l'arrivée, par deux fonctions (deux *densités de probabilité*) donnant, en tout point X de l'espace, la densité du nuage. On note $\alpha(X)$ la densité de départ et $\beta(X)$ celle d'arrivée. Ce sont deux fonctions réelles positives, dont l'intégrale sur tout l'espace vaut 1 et l'intégrale sur chaque ensemble A vaut la proportion de points du nuage se trouvant dans A. Dans ce cadre continu, on cherche donc à déterminer le transport optimal qui sera une fonction T envoyant les points de départ vers des points d'arrivée.



Dans le cas continu, en considérant le coût quadratique, c'est-à-dire l'intégrale du carré de la distance entre le nuage de départ et celui d'arrivée (au lieu de la somme des carrés des distances entre les points de départ et les points d'arrivée du cas discret), des résultats théoriques obtenus dans la seconde moitié des années 1980 permettent de trouver la solution.

Pour les comprendre, il est utile de voir ce qui se passe dans le cas de l'espace à une seule dimension, c'est-à-dire quand la variable X et les densités α et β vivent sur une droite. Dans ce cas, le résultat nous dit que le transport optimal T est le *transport monotone croissant*, c'est-à-dire celui qui envoie le point le plus à droite de la densité α sur le point le plus à droite de β, et qui respecte l'ordre ($X \leq Y$ donne $T(X) \leq T(Y)$; on peut démontrer qu'il n'y a qu'une seule fonction T avec cette propriété qui respecte les densités α et β).

En dimension supérieure à 1, un problème se pose cependant : qu'est-ce que cela signifie, qu'une fonction de deux variables est « croissante » ? A priori, cela n'a pas vraiment de sens… Pourtant, on peut utiliser une astuce : en dimension 1, T est croissante si et seulement si c'est la dérivée d'une fonction convexe (c'est-à-dire d'une fonction dont le graphe est courbé vers le haut, dont la pente augmente avec l'abscisse, et qui est caractérisée par le fait que la dérivée de sa dérivée est positive). La notion de fonction convexe a aussi un sens en dimension supérieure, et le langage qui a été choisi pour énoncer le théorème concernant le transport optimal en dimension supérieure est bien celui-ci.

Le résultat des années 80 évoqué auparavant fait bien intervenir une fonction convexe $\Phi(x,y)$, que l'on appelle « potentiel ». Il nous dit que le transport optimal T entre les deux fonctions de densité α et β est égal au *gradient* de cette fonction $\Phi(x,y)$, noté $D\Phi(x,y)$, qui est le vecteur dont les composantes sont, en chaque point, la dérivée de Φ par rapport à x et sa dérivée par rapport à y.

Reste cependant à trouver cette fonction « potentiel » Φ qui nous permettra de définir la trajectoire optimale de chacune des particules. En exprimant le fait que T envoie la densité α sur β, on obtient des équations différentielles (faisant intervenir des dérivées). En dimension 1 l'équation impose une égalité sur la dérivée de T, donc la dérivée seconde de Φ. En dimension supérieure, on obtient une formule exprimant la valeur du déterminant de la matrice donnée par les dérivées secondes de Φ (par rapport aux différentes variables), qui doit être égal au ratio entre les densités α et β. Cette équation admet une solution unique Φ et permet donc de trouver le potentiel et, ensuite, la fonction T.

Ainsi, paradoxalement, le fait d'avoir « compliqué » le problème (en passant du discret au continu) nous a permis de le résoudre.

Une application inattendue : modifier les images

Ce problème d'essence mathématique rencontre déjà des applications bien concrètes, comme le montre cet exemple étudié dans les années 2000 par l'équipe de Jean-Michel Morel à l'École Normale Supérieure de Cachan.

On dispose de l'image numérisée d'une peinture dont les couleurs sont défraîchies. Ces couleurs sont numériquement codées selon les couleurs élémentaires bleue, rouge et verte. À chaque pixel est donc associée une certaine proportion de bleu, de rouge ou de vert. En balayant les N pixels de l'image numérique, on obtient un nuage de N points dans l'espace tridimensionnel des couleurs. (Il s'agit d'un espace abstrait, pas de notre espace physique.) Par ailleurs, on a une idée de la palette de l'auteur, obtenue en considérant ses autres œuvres. Si l'on peut en faire une statistique raisonnable, on pourra ainsi définir un nuage « de référence », constitué de N points dans le même espace de couleurs. Une technique de rehaussement des couleurs consiste alors à effectuer le transport optimal, dans l'espace des couleurs, du nuage de points fourni par l'image défraîchie vers le nuage de points de référence. On déplacera alors en conséquence les valeurs des couleurs de chaque pixel.

Notons que dans cet exemple, c'est bien grâce à la formulation continue que le problème a pu être étudié, ceci indépendamment du fait que le calcul numérique est effectué sur un nombre fini de points.

« Jeunes filles au bord de la mer » de P. Puvis de Chavanne.

Une autre application… aux mathématiques elles-mêmes !

Dans les années 1990, le transport optimal devient un outil de démonstration puissant et élégant pour démontrer de nombreuses « inégalités » de la géométrie et de l'analyse fonctionnelle.

Le cas le plus simple à décrire est la fameuse « inégalité isopérimétrique », solution du problème de Didon déjà mentionné par Virgile, qui peut s'énoncer ainsi : de toutes les courbes fermées de longueur donnée, celle qui entoure l'aire la plus grande est le cercle. Si on veut s'en tenir à la légende, Didon avait obtenu du roi local des terres, mais « autant qu'il en pourrait tenir dans la peau d'un bœuf ». Elle prit alors une peau, la fit découper en ficelles et fit une longue corde pour contourner la plus grande surface possible, en résolvant donc un « problème isopérimétrique » et en fondant la ville de Carthage.

Or, si on prend une figure F dans le plan et on la compare au cercle C de même aire, en considérant le transport optimal entre deux densités uniformes sur F et sur C et en utilisant les propriétés de Φ, on peut démontrer que le périmètre de F est forcément plus grand que celui de C. De même, dans l'espace, le volume maximal entouré par une surface donnée est réalisé par la sphère.

Ces résultats géométriques étaient bien connus avant le transport optimal et cela en donne juste une preuve alternative ; une méthode similaire est néanmoins utilisée pour réaliser la démonstration d'inégalités non démontrées auparavant.

Mathématiques pures et appliquées se répondent.

La voici donc bouclée, cette boucle de deux siècles. Une boucle où mathématiques pures et appliquées se répondent, où l'in-

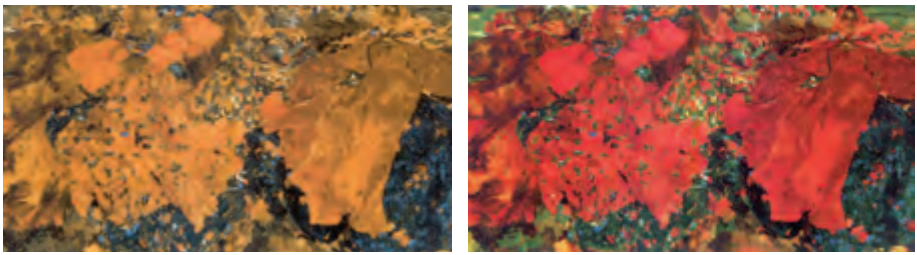


Figure 3. À droite : « Whirlwind » de F. Maliavin. À gauche : le tableau de F. Maliavin repeint avec les teintes de Puvis de Chavanne, par des méthodes de transport optimal. D'après Delon-Salomon-Sobolevski, SIAM J. Appl. Math.

formatique théorique permet de mieux comprendre ce qui peut se calculer efficacement. Et surtout, où c'est en « changeant de point de vue » (ici, en passant d'un problème discret à un problème continu) qu'un problème compliqué peut livrer une solution inattendue.

Pour aller plus loin

Cet article est inspiré de celui du même auteur paru dans *Interstices*.

La combinatoire

La combinatoire est le domaine des problèmes mathématiques où l'on énumère et classifie les manières de relier ou d'arranger un nombre fini, mais souvent élevé, d'objets. Par exemple :

Combien de plaques minéralogiques différentes du type LL CCC LL peuvent être émises ?

Et si l'on veut que deux plaques différentes se distinguent par deux caractères au moins ?

Combien de couleurs sont nécessaires au minimum pour colorier un planisphère de manière à ce que deux états ayant une frontière commune n'aient pas la même couleur ?

Des statistiques pour détecter les altérations chromosomiques

Emilie Lebarbier, *maître de conférences*
Stéphane Robin, *directeur de recherche INRA, à AgroParisTech*

Les altérations chromosomiques sont responsables de nombreuses maladies, parmi lesquelles certains cancers. La détection de petites altérations, essentielle pour le diagnostic du médecin, fait appel à un modèle classique en statistiques : la segmentation.

Dans chaque cellule du corps humain, chaque chromosome est présent en deux exemplaires ou *copies*. Un écart à cette règle (*perte* ou *gain*) entraîne un déséquilibre pouvant être à l'origine de certaines maladies. La *perte* correspond à l'absence ou à la présence en une seule copie du chromosome et le *gain* à la présence de trois, quatre copies, voire plus encore. Ce déséquilibre du nombre de copies est appelé *altération chromosomique*. La plus connue est la trisomie, pour laquelle un chromosome est présent en trois copies. Ainsi, le syndrome de Dawn (aussi appelé *mongolisme*) est dû à la présence de trois copies complètes du 21e chromosome. Il est possible de détecter ce type d'altérations à l'aide du traditionnel *caryotype*.

D'autres maladies sont liées à des altérations qui ne touchent pas des chromosomes entiers mais seulement des portions de chromosomes. Certaines portions sont « perdues » (absentes ou présentes en une seule copie) ou « amplifiées » (présentes en trois ou quatre copies, voire plus encore). C'est notamment le cas d'un grand nombre de cancers. Les cellules du tissu malade présentent typiquement des pertes de régions contenant des gènes « suppresseurs de tumeurs » ou, au contraire, des amplifications de régions contenant des *oncogènes* qui favorisent le développement tumoral.

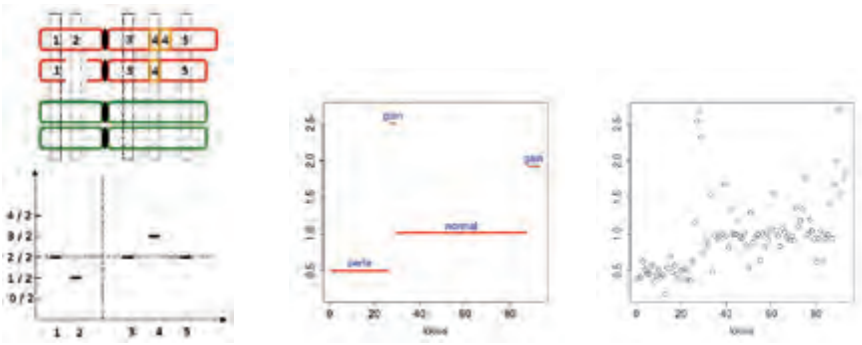
Quand elles sont trop petites, ces altérations peuvent malheureusement passer inaperçues dans l'étude du caryotype. Leur détection est donc un enjeu essentiel de la recherche médicale afin d'identifier les gènes affectés, d'essayer de comprendre leur implication dans le développement de la maladie et de concevoir ainsi des thérapies adaptées.

La détection des altérations chromosomiques est un enjeu essentiel de la recherche médicale.

Évaluer le nombre de copies d'une portion de chromosome

C'est dans les années 1990 que l'étude de ces petites altérations a été rendue possible grâce à l'arrivée de techniques de biologie moléculaire. Le principe consiste à compter le rapport du nombre de copies de différents gènes, dont la position (appelée *locus*) sur le génome est connue, entre un individu malade et un individu sain. Ce principe est illustré sur un exemple très simple en figure 1a, pour cinq locus.

Il existe deux altérations chez l'individu malade (rouge) par rapport à l'individu sain (vert), qui possède toujours deux copies de chaque gène : une portion perdue au locus 2 qui donne un rapport du nombre de copies égale à 1/2 et un gain au locus 4 qui donne 3/2. Si l'on s'intéresse à beaucoup plus de locus (figure 1b), plusieurs locus successifs peuvent avoir le même statut biologique – normal, gain(s) ou perte(s) – et forment ainsi des *régions*. En réalité, il n'est pas possible de calculer directement le vrai nombre de



(a) (b) (c)

Figure 1. (a) Rapports « théoriques » entre le nombre de copies d'un individu malade et celui d'un individu sain (2 copies) suivant le numéro du locus pour cinq locus fictifs. (b) Ces rapports théoriques découpés suivant des régions, obtenus après travail statistique pour 93 locus réels. (c) Rapports « empiriques » de copies mesurés expérimentalement suivant le locus pour ces 93 locus.

copies mais les expériences biologiques vont permettre d'obtenir ce nombre à une erreur près (erreur de mesure, variabilité naturelle), comme le montre la figure 1c. Le rapport du nombre de copies ainsi obtenu pour les différents locus ne sera donc pas exactement égal à l'ensemble de valeurs théoriquement possibles 1/2, 1, 3/2, 2… ce qui rend alors impossible une analyse manuelle.

L'objectif statistique est de retrouver les « vraies » valeurs à partir des données que l'on a observées. En particulier, comme on l'a vu précédemment, il s'agit de détecter et de localiser automatiquement les différentes régions formées par plusieurs locus successifs ou un seul. Une fois l'analyse effectuée, le médecin pourra fonder son diagnostic sur la figure 1b plutôt que sur la figure 1c.

Un modèle mathématique

Comme souvent en statistique, la résolution de ce problème passe par la définition d'un modèle, c'est-à-dire, d'une traduction aussi fidèle que possible du processus biologique en langage mathématique (et suffisamment simple pour permettre une résolution mathématique). Un modèle classique en statistique dédié à la détection de régions (ou segments) est le modèle de *segmentation*. Il consiste à supposer que la vraie valeur du nombre de copies est la même au sein de chaque région, que cette valeur change quand on change de région, et que la valeur observée à un locus donné est égale à la vraie valeur de la région auquel il appartient plus un terme aléatoire, l'erreur. L'objectif statistique consiste alors à répondre à trois questions : Combien il y a de régions ? Où se trouvent-elles (c'est-à-dire quelles sont les limites entre les régions) ? Quelles sont les vraies valeurs au sein de chaque région ?

La troisième question se résout simplement : de façon naturelle, la vraie valeur de chaque région sera estimée par la moyenne des observations dans cette région. La résolution statistique des deux autres questions se déroule en sens contraire. Pour un nombre de régions fixé – notons-le k - on cherche le positionnement de ces régions le mieux ajusté aux données. C'est un problème de nature algorithmique. Sa résolution va nous fournir un algorithme qui nous donnera la « meilleure segmentation » des données pour k régions.

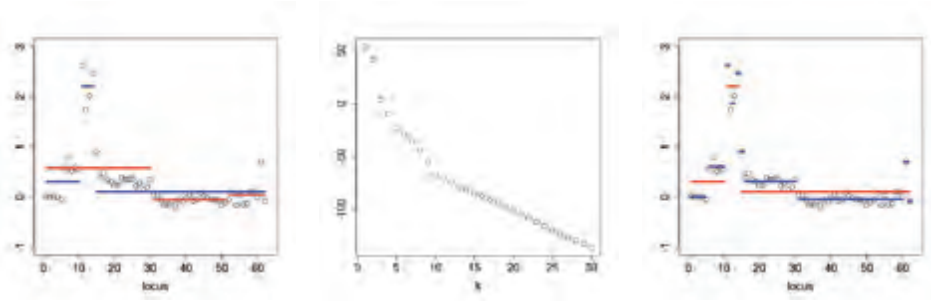
Bien sûr, en général, k n'est pas connu. Même si l'on dispose d'information a priori

sur les données, il est difficile de se le fixer à l'avance. Il faut donc le déterminer, ou plus exactement, choisir, encore une fois, le « meilleur » possible. Connaissant la « meilleure segmentation » des données pour différents k, il s'agira alors d'en prendre une parmi toutes. Cette fois, le problème est plutôt de nature statistique.

Comment obtenir la meilleure segmentation

Notons n le nombre de locus et fixons à k le nombre de régions. Il existe bien sûr plusieurs segmentations possibles des n locus en k régions et il s'agit de trouver la « meilleure ». La figure 2a présente deux exemples de segmentations, une bleue et une rouge, sur des données recueillies le long du chromosome 6 d'un patient atteint d'un cancer de la vessie. Entre ces deux segmentations, on préférera la segmentation bleue à la segmentation rouge. En effet, les données au sein des régions « bleues » (c'est-à-dire délimités par la segmentation bleue) sont moins dispersées autour de leur « vraie valeur » qu'au sein des régions « rouges ». La segmentation bleue s'ajuste donc mieux aux données que la segmentation rouge. On peut alors choisir ce critère d'ajustement, classique en statistique, comme mesure de qualité des segmentations. Reste à essayer toutes les segmentations possibles et à retenir la meilleure.

Un problème se pose alors : celui du nombre total de segmentations. On montre facilement que, si on observe n locus et qu'on cherche à segmenter les données en k régions, il existe autant de segmentations possibles que de combinaisons de $(k-1)$ objets parmi $(n-1)$, que l'on note $C(n-1, k-1)$. Par exemple, pour $n = 1000$



(a) (b) (c)

Figure 2. (a) Deux segmentations en trois régions. (b) Evolution de l'ajustement en fonction du nombre de régions. (c) Meilleures segmentations en trois et dix régions.

locus et k = 10 segments, il existe plus de 10^{138} segmentations possibles et même les ordinateurs les plus puissants ne peuvent pas explorer l'ensemble de ces combinaisons. Une solution existe cependant en reformulant légèrement le problème.

La meilleure segmentation pour un nombre de régions donné est celle qui s'ajuste le mieux aux données.

On définit le « coût » d'une région comme étant la dispersion des données autour de la moyenne de cette région. On appelle une « bonne » région celle qui a un faible coût. Sachant qu'une segmentation est un ensemble de k régions de la forme $[1 ; t_1]$, $[t_1+1 ; t_2]\ldots$, $[t_{k-2}+1 ; t_{k-1}]$, $[t_{k-1}+1 ; n]$, la « meilleure » segmentation correspond alors au chemin qui permet :

– d'aller du point 1 au point n,
– en faisant $(k - 1)$ étapes à des points t_1, t_2, … t_{k-2}, t_{k-1} à déterminer,
– pour le plus faible coût total.

Un tel problème est un problème de « plus court chemin », que l'on sait résoudre avec un algorithme qui nécessite $k \times n^2$ opérations. Pour n = 1000 locus et k = 10 segments, il faut donc faire 10^7 opérations, ce qui est bien moins que 10^{138}. Sur la figure 2a, la segmentation bleue est la meilleure de toutes les segmentations possibles en trois régions.

Choisir le nombre de régions

Une première idée serait de se fonder sur le coût des segmentations défini plus haut. Pour chaque nombre de régions k possible (k = 1, 2…), on détermine la segmentation de plus faible coût au moyen de l'algorithme du plus court chemin et on retient le coût de cette meilleure segmentation, notée $J(k)$. On choisit ensuite le nombre de régions k qui, parmi toutes ces solutions, mène au coût le plus faible. On peut montrer que ce raisonnement conduit toujours à choisir le plus grand nombre de régions possible. En effet, le coût $J(k)$, qui traduit l'ajustement de la segmentation aux données, diminue avec le nombre de régions : plus il y a de régions, plus la segmentation résultante sera proche des données (voir figure 2c). La segmentation alors choisie est la segmentation où chaque locus est une région à lui tout seul (l'ajustement est parfait). Mais ce choix n'a aucun intérêt puisqu'il revient à rendre au médecin la figure 1c en guise de figure 1b.

Afin d'obtenir un résultat interprétable, on préfère choisir une segmentation suffisamment bien ajustée aux données sans être trop complexe, c'est-à-dire avec un nombre raisonnable de régions. La figure 2b montre l'évolution du coût $J(k)$ en fonction du nombre de régions k et illustre bien cet objectif. En effet, $J(k)$ diminue fortement pour les premières valeurs de k, signifiant que l'on gagne fortement en ajustement, puis à partir d'un certain k, cette diminution devient moins importante (augmenter k n'améliore que marginalement l'ajustement alors que le modèle devient de plus en plus complexe). La valeur de k correspondant à cette cassure devrait donc être un bon choix. L'idée consiste donc à rendre plus coûteuses les segmentations avec plus de régions. D'un point de vue mathématique, cela se traduit par l'ajout au coût $J(k)$ d'un terme appelé *pénalité*, qui devra refléter la complexité de la segmentation et augmenter avec elle.

Il reste à définir ce que l'on entend par « complexité » d'une segmentation à k segments. Pour décrire une segmentation, il faut déterminer les k valeurs des moyennes mais aussi les limites t_k entre les régions. On a vu que cette dernière recherche nécessitait l'exploration d'un trop grand nombre de segmentations. Ainsi, une pénalité adaptée devra prendre en compte ces deux quantités, typiquement sous la forme

$$a \times k + b \times \log[C(n-1, k-1)].$$

Il serait illusoire de penser que le critère ainsi construit (et qui nécessite donc de choisir des constantes a et b) est infaillible, c'est-à-dire mène à la meilleure solution dans toutes les situations possibles. Disposer du maximum d'informations est donc important. En plus des informations que détient le biologiste (qui connaît ses données mieux que personne), une analyse du comportement du coût $J(k)$ peut permettre d'extraire plusieurs solutions pour k pouvant être biologiquement pertinentes. Par exemple, la figure 2b montre que $J(k)$ présente deux inflexions pour k = 3 et k = 10. Si le critère mène à la solution k = 3, il sera important d'étudier la segmentation pour k = 10, qui comme on va le voir dans la suite permet la détection d'une altération supplémentaire.

La meilleure segmentation permet la détection d'une altération supplémentaire.

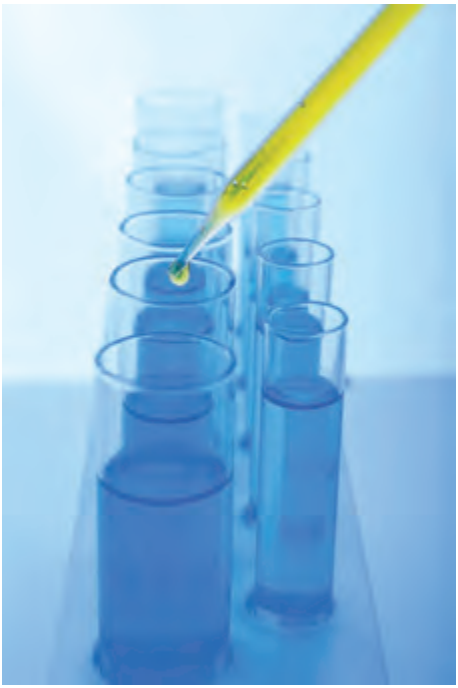
Une détection plus fine

Le critère décrit ci-dessus sélectionne $k = 10$ régions pour les données présentées en figure 2. La segmentation associée est donnée figure 2c en bleu. Une région amplifiée est détectée (des locus 11 à 14). Cette région du chromosome 6 contient le gène E2F3 dont l'amplification est connue pour être associée au développement du cancer de la vessie. Une amplification de cette région est également repérée chez d'autres patients atteints du même type de cancer.

Une variation au locus 61 est également détectée par cette segmentation. Ce locus est connu par les biologistes pour être lié à un polymorphisme (une variation génétique locale spécifique au patient). Il n'est donc pas surprenant d'observer une altération à ce locus chez certains patients.

Comme il a été évoqué ci-dessus, un deuxième point d'inflexion existe dans le coût $J(k)$ pour $k = 3$. La segmentation associée est représentée figure 2c en rouge. Cette segmentation, moins fine, permet déjà de détecter la région contenant le gène E2F3 mais pas le polymorphisme au locus 61.

Les auteurs remercient Théo Robin pour sa relecture attentive et ses commentaires.



Le piano rêvé des mathématiciens

Juliette Chabassier, *chargée de recherche Inria, équipe Magique-3D*

Comme de nombreux phénomènes physiques, le fonctionnement d'un piano peut être modélisé grâce aux mathématiques. Mais le modèle obtenu permet aussi d'aller plus loin, de rêver de pianos impossibles ou d'imaginer des sons nouveaux. La recherche offre ainsi au compositeur un formidable champ d'exploration et de création.

Le piano est un système acoustique et mécanique assez sophistiqué. On peut résumer très schématiquement son fonctionnement de la façon suivante (voir figure 1) : le doigt du pianiste frappe une touche du clavier, un mécanisme très précis démultiplie le mouvement de la touche et met en mouvement un marteau. Le marteau frappe entre une et trois cordes à la fois (selon la note choisie), qui se mettent en vibration. Le chevalet permet de transmettre l'énergie des cordes à la table d'harmonie, qui vibre elle aussi, mettant en mouvement les molécules d'air avoisinantes, et entraînant la propagation d'un son dans l'air.

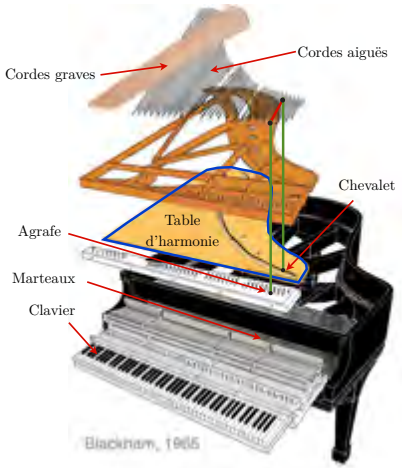


Figure 1. Vue éclatée d'un piano à queue.

Des EDP derrière les touches

Comme la plupart des phénomènes physiques qui nous entourent, le fonctionnement du piano peut être décrit mathématiquement à l'aide d'équations aux dérivées partielles (abrégées en EDP). Les EDP sont des équations qui mettent en relation les dérivées d'une ou plusieurs fonctions par rapport à leurs différentes variables : le temps, l'espace… Certaines EDP sont assez célèbres : *l'équation de la chaleur*, qui modélise l'évolution de la température dans un milieu ; les *équations de propagation d'ondes*, qui modélisent la vibration d'une corde et la propagation du son ; les *équations de Navier-Stokes*, qui interviennent en mécanique des fluides ; les *équations de Maxwell*, qui sont celles de l'électromagnétisme ; les *équation de l'élastodynamique*, qui modélisent les ondes sismiques mais aussi la déformation des matériaux ; ou encore l'*équation de Schrödinger*, qui modélise l'évolution d'une particule quantique.

Revenons à notre piano : chacune des étapes acoustiques citées plus haut peut donc être modélisée par une EDP (pour les cordes, ce sera une équation des ondes monodimensionnelle avec raideur, pour la table d'harmonie, une équation de vibration de plaque, pour l'air, une équation de propagation des ondes tri-dimensionnelle) ou par une équation d'une autre nature (pour le marteau on aura une équation différentielle ordinaire, au chevalet on exprimera l'égalité des vitesses, et pour le couplage table d'harmonie/air, l'équation traduira l'égalité aussi des vitesses mécanique et acoustique).

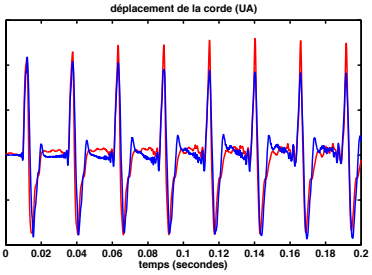
Figure 2. Maillage d'une table d'harmonie.

De manière générale, les EDP sont des équations complexes dont il est en général impossible de donner la solution au moyen d'une formule. Les problèmes physiques qu'elles traduisent, et qui combinent la plupart du temps plusieurs de ces équations, le sont aussi, à plus forte raison. Diverses méthodes permettent d'en calculer les solutions approchées décrivant le mieux possible la réalité (voir l'encadré *Analyse numérique des EDP*). Cela passe souvent par la *discrétisation* du problème et par le calcul sur ordinateur de solutions approchées.

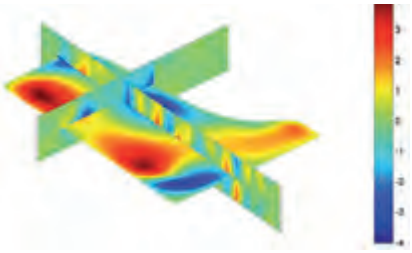
Un modèle général du piano est établi grâce aux EDP, puis discrétisé sous la forme d'un maillage, et enfin les solutions sont calculées numériquement.

C'est ce que l'on fait dans le cas du piano : un modèle général est établi grâce aux équations, puis discrétisé sous la forme d'un *maillage* (voir un maillage de table d'harmonie en figure 2), et enfin les solutions sont calculées numériquement grâce à des méthodes d'analyse numérique développées précisément pour l'occasion.

Grâce à ce modèle et aux simulations numériques qui en découlent, on parvient à reproduire fidèlement des formes d'ondes (voir figure 3), mais aussi, par exemple, à connaître le champ de pression de l'air sur un ensemble de points tout autour du piano, ce qui serait impossible à mesurer sans perturber le système : si on essayait de faire cette mesure en entourant le piano d'un réseau de capteurs, la présence même



(a) Comparaison simulation numérique (bleu)/expérience (rouge) : déplacement de la corde D#1 pendant les 200 premières millisecondes.



(b) Déplacement sur la table d'harmonie en couleurs arbitraires et pression dans l'air (deux plans de coupe) en Pa (échelle de couleur sur la droite), 16 millisecondes après la frappe du marteau sur la corde C2.

Figure 3. Résultats de simulation numérique : vérification et prédictions

de ce réseau provoquerait un phénomène de diffraction du son et fausserait donc la mesure effectuée.

Pianos virtuels

L'une des applications les plus enthousiasmantes de ce modèle et de sa discrétisation numérique est l'aide à la facture instrumentale. Aujourd'hui, la facture du piano est majoritairement basée sur un savoir empirique issu de siècles d'expérimentations, d'échecs, de succès… Les facteurs, c'est-à-dire ceux qui conçoivent et réalisent les instruments, ont ainsi acquis un ensemble de connaissances extrêmement précises. On trouve par exemple dans la littérature spécialisée des affirmations telles que : « les rouleaux des marteaux influent sur le toucher du piano à queue de façon décisive », ou encore « la table d'harmonie est en épicéa de Sitka, un bois d'une densité et d'une élasticité idéales pour la résonance ». Ces affirmations intriguent beaucoup les chercheurs en acoustique musicale qui s'attachent, en utilisant des méthodes scientifiques, à leur donner raison ou tort, et à aller plus loin dans la compréhension des phénomènes mis en jeu. Le modèle de piano peut alors permettre d'isoler certains phénomènes afin de comprendre leur influence sur le son, sur le rayonnement, ou encore sur la transmission de l'énergie… mais aussi de construire virtuellement des pianos qui n'existent pas (changer la forme ou la taille de la table, les matériaux utilisés, la façon de l'utiliser…) et de pouvoir écouter le son qu'ils auraient s'ils étaient construits.

Encore plus intriguant, le modèle permet de générer des sons d'objets qui ne peuvent pas exister pour des raisons pratiques (matériaux inventés, cordes de sept mètres

de long, piano flottant sans cadre ni pieds…) mais qui pourtant respectent les lois de la physique et dont le son paraît donc plausible à l'oreille (en ce sens que le cerveau parvient à les attribuer à un événement, une cause physique, une dynamique associée). On touche là une application très intéressante du point de vue de l'interaction entre science et musique : la recherche offre au compositeur des instruments nouveaux, un matériau sonore jusqu'ici inexistant et adaptable à souhait. A une époque où le timbre et la manipulation du son sont au cœur de la création musicale, elle lui permet d'aller toujours plus loin.

Le modèle permet de générer des sons d'objets qui ne peuvent pas exister pour des raisons pratiques.

Un outil pour le compositeur

Ces possibilités sont parfaitement illustrées par les travaux du compositeur Guillaume Loizillon, qui utilise un logiciel de synthèse sonore par modèles physiques développé à l'Institut de Recherche et Coordination Acoustique/Musique, Modalys, permettant de modéliser une multitude d'instruments existants ou chimériques. Il dit à ce sujet : « Ce que je trouve intéressant dans Modalys, plus que l'imitation d'un timbre, c'est que cela permet, au niveau d'un son unique, de donner, de manière assez évidente, l'idée de la matière, du métal, du bois. Deuxième chose, cela permet de créer des dynamiques, des événements et non pas des objets. […] C'est-à-dire que le son raconte un événement : cela roule, cela rebondit. »

Analyse numérique des EDP

Quand on s'intéresse à des applications concrètes des EDP (en météorologie, en sismologie, en aéronautique, en construction automobile…), il ne suffit pas de connaître l'existence et le comportement d'une solution dans son espace abstrait mathématique. Les physiciens ont besoin de pouvoir au moins estimer cette solution quantitativement. La science qui consiste à chercher une approximation de la solution d'une EDP de façon précise et performante grâce à l'utilisation d'ordinateurs est celle du calcul scientifique et de l'analyse numérique. L'idée est de ramener l'infini de l'espace mathématique de la solution de l'EDP au caractère foncièrement fini et discret (par opposition à continu) de l'ordinateur, mais en s'assurant que plus on fera d'efforts (en termes de temps de calcul ou de puissance du processeur), plus on s'approchera de la « vraie » solution. Une des méthodes les plus connues et les plus répandues dans l'industrie est probablement la *méthode des éléments finis*.

Le compositeur peut même dépasser le monde physique existant et en imaginer un autre, aux lois différentes : « Ensuite ce que j'aime c'est aller dans des modèles plus utopiques » poursuit Guillaume Loizillon. « Avec Jean Claude Risset [NDLR : compositeur également], on avait discuté du sujet *est-ce que l'on peut créer une physique non terrienne*. [...] On peut imaginer une physique où les objets rebondissent pendant des heures et des heures, où la pesanteur n'est plus la même. »

Ainsi, la modélisation instrumentale ouvre au musicien des champs d'exploration inouïs.

Pour aller plus loin

Site web : http://modelisation.piano.free.fr

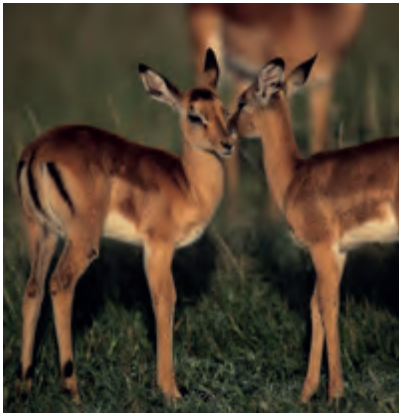
Les citations de Guillaume Loizillon sont extraites de *La synthèse par modèles physiques*, mémoire de maîtrise de Aline Hufschmitt, Université Paris Sorbonne (Paris IV), disponible à l'adresse alinehuf3.free.fr

Comment faire coopérer des individus égoïstes ?

Yannick Viossat, *maître de conférences à l'Université Paris-Dauphine*

La coopération est au cœur de nombreux comportements sociaux ou biologiques. La théorie des jeux permet d'expliquer le choix de stratégies coopératives et d'en comprendre les mécanismes dans des contextes où les individus se trouvent en concurrence, des situations de guerre à la régulation de la pêche.

Le monde vivant offre de nombreux exemples de coopération : les insectes sociaux, les animaux criant pour signaler la présence d'un prédateur, les impalas se nettoyant réciproquement… Darwin considérait ces comportements coopératifs comme une énigme, et un défi à sa théorie. La sélection naturelle ne devrait-elle pas favoriser les comportements égoïstes ? Ce paradoxe de la coopération est au cœur de l'analyse de transitions majeures dans l'évolution, comme l'apparition de la vie ou celles des êtres multicellulaires. Dans certaines de ces transitions, de petites entités coopèrent pour former une entité plus complexe. Ces constructions subsistent, alors qu'elles semblent à la merci de l'apparition d'entités parasites, tirant profit de la coopération des autres mais ne coopérant pas elles-mêmes. Comprendre l'origine et le maintien de la coopération dans le monde vivant est donc un problème central de la biologie de l'évolution.



Intérêt collectif, intérêts individuels

Des questions similaires se posent en sciences sociales. De nombreux conflits existent entre l'intérêt collectif et les intérêts individuels. La coopération peut-elle alors émerger spontanément ? Plus concrètement : pourquoi continuons-nous à trop pêcher, et à épuiser les stocks de poissons ? A l'inverse, pourquoi, dans certaines tranchées de la guerre de 14-18, les soldats faisaient-ils exprès de ne pas gêner le ravitaillement de leurs ennemis ?

Les mathématiques permettent d'éclairer ces questions en analysant des modèles de conflits entre intérêt collectif et intérêts individuels. Le plus simple de ces modèles, appelé dilemme du prisonnier pour des raisons anecdotiques, se présente ainsi : deux individus, appelés joueur 1 et joueur 2, doivent décider d'aider l'autre (coopérer) ou de ne pas l'aider (faire défection). Chaque joueur doit choisir sans connaître le choix de l'autre.

Coopérer rapporte un bénéfice de 4 à l'autre mais coûte 1 à celui qui coopère. Faire défection ne rapporte rien à l'autre et ne coûte rien. Ainsi, si les deux joueurs coopèrent, ils gagnent tous les deux 4 – 1 = 3. Le tableau suivant résume les actions possibles et les gains correspondants. Les gains du joueur 1 (J1) sont en rouge, ceux du joueur 2 (J2) en bleu. La valeur précise de ces gains importe peu tant que leur ordre (du plus petit au plus grand) reste le même. On pourrait donc supposer que faire défection nous rapporte quelque chose, ou autre variante.

J1 \ J2	Si J2 coopère (C)	Si J2 fait défection (D)
Si J1 coopère (C)	3,3	–1,4
Si J1 fait défection (D)	4, –1	0,0

De nombreuses situations de la vie courante présentent des analogies : les enfants qui peuvent prêter leurs jouets (C) ou ne pas les prêter (D) ; les grandes puissances, lors de la guerre froide, qui pouvaient limiter leur armement (C) ou s'armer toujours plus (D) ; le citoyen qui peut recycler ses déchets (C) ou ne pas le faire (D), etc.

L'analyse de ce type de situations relève de la théorie des jeux. Cette branche des mathématiques examine l'interaction d'acteurs stratégiques. Elle concerne toute situation où des acteurs doivent chacun prendre des décisions, qui déterminent l'issue de cette situation.

Assurer l'avenir

La théorie des jeux a permis d'identifier plusieurs mécanismes qui permettent de surmonter les conflits entre intérêt collectif et intérêts individuels. Voyons maintenant le

fonctionnement et les limites d'un de ces mécanismes : la réciprocité directe (aider ceux qui nous ont aidés).

D'autres mécanismes de coopération

La réciprocité directe n'est pas le seul mécanisme sur lequel peut se baser la coopération entre individus. Parmi les autres, on peut citer : la construction d'institutions permettant de faire appliquer des contrats, la réciprocité indirecte (aider ceux qui ont aidé les autres) ou encore la sélection de parentèle (aider les individus qui nous sont apparentés), fondamentale en biologie.

Revenons au dilemme du prisonnier. Quand le jeu n'est joué qu'une fois, jouer D rapporte toujours 1 unité de plus que de jouer C (4 au lieu de 3 si l'autre joue C, 0 au lieu de -1 s'il joue D). Les deux individus ont donc individuellement intérêt à faire défection. Le dilemme est qu'ils obtiennent alors 0, alors qu'ils auraient pu obtenir 3 s'ils avaient coopéré tous les deux.

Si l'interaction se répète la situation est différente : mes choix d'aujourd'hui influent sur les choix que fera mon partenaire demain. Coopérer peut être une forme d'investissement : cela me rapporte moins aujourd'hui que de faire défection, mais peut augmenter mes gains futurs, si cela rend l'autre plus coopératif. En revanche, si l'autre joueur fait systématiquement défection, j'ai intérêt à le faire aussi. Il n'y a donc pas de stratégie qui soit toujours la meilleure.

Mes choix d'aujourd'hui influent sur les choix que fera mon partenaire demain.

Toutefois, certaines stratégies obtiennent souvent de bons résultats. La stratégie Donnant-Donnant, notamment, a remporté plusieurs tournois informatiques, où chaque stratégie jouait contre chacune des autres et où le but était d'avoir le meilleur score total. Cette stratégie de simple réciprocité commence par coopérer, puis fait ce qu'a fait l'autre au tour précédent.

Ce n'est pas en exploitant les autres que Donnant-Donnant obtient de bons résultats. En fait, dans un duel, elle ne gagne jamais plus que son adversaire. Mais elle parvient à enclencher une logique de coopération, et en moyenne obtient de très bons scores. L'analyse de tournois de dilemme du prisonnier fait apparaître des caractéristiques des stratégies à succès : essayer de coopérer, ne pas chercher à gagner plus que l'autre, ne pas se laisser exploiter, pardonner une défection isolée, et être lisible.

Le système du « Vivre et laisser vivre » de la première guerre mondiale offre une illustration concrète de coopération dans un environnement pourtant défavorable. Ce système consistait à ne pas attaquer l'ennemi pourvu qu'il ne nous attaque pas non plus. Cette attitude était fortement découragée par le commandement. Un officier anglais racontait ainsi sa visite des

tranchées : « Je fus ébahi de voir les soldats allemands se promener à portée de fusil. Nos hommes semblaient ne pas remarquer (…). Ces choses-là ne devraient pas être permises. » Pourtant, ce système s'établit en de nombreux points du front, car dans cette guerre de tranchée, les mêmes unités se faisaient face suffisamment longtemps pour que la logique de la réciprocité fonctionne. Un soldat expliquait : « Ce serait un jeu d'enfant de bombarder la route derrière les tranchées, encombrée comme elle doit l'être de chariots de ravitaillement (…) mais dans l'ensemble tout est calme (…). Si vous empêchez l'ennemi de se ravitailler, son remède est simple : il vous empêchera de vous ravitailler. »

Une analyse précise du dilemme du prisonnier permet de mieux comprendre les facteurs favorisant les stratégies de réciprocité. L'avantage de la défection ne doit pas être trop fort, et l'importance de l'avenir doit être suffisamment grande. Pour cela, il faut que les joueurs soient patients et que l'interaction se répète suffisamment longtemps. Il faut aussi que le moment où l'interaction s'arrête ne soit pas connu à l'avance. Sinon, à la dernière étape, les deux joueurs ont intérêt à faire défection et, anticipant cela, ils sont amenés à faire défection à l'avant-dernière étape, et ainsi de suite.

Théorie des jeux d'évolution

Pour mieux comprendre les conditions qui permettent l'émergence de la coopération, il est utile de recourir à la théorie des jeux d'évolution. Celle-ci étudie des situations où un grand nombre d'individus interagissent et font évoluer leur comportement en fonction du résultat des interactions précédentes. On suppose que les stratégies qui obtiennent de bons résultats ont tendance à se répandre dans la population. Une formulation précise de ce principe donne lieu à un système dynamique, qui peut modéliser tant des phénomènes de sélection naturelle que d'apprentissage.

Dans un monde où des erreurs peuvent se produire, il est bon de réparer les dommages causés aux autres involontairement, et de ne pas se venger systématiquement des agressions subies.

Dans les modèles de mutation-sélection, on observe une alternance de phases de défection et de coopération. Partant d'un état où personne ne coopère, des mutations peuvent faire apparaître des stratégies du type Donnant-Donnant. Si celles-ci parviennent à une fréquence suffisamment élevée, elles obtiennent de meilleurs résultats que la stratégie de défection et envahissent

la population. On aboutit alors à un état où tous les individus jouent des stratégies coopératives. La stratégie qui consiste à coopérer de manière indiscriminée peut alors augmenter en fréquence, car elle n'est plus contre-sélectionnée. Mais si les coopérateurs indiscriminés sont trop fréquents, la stratégie de défection est favorisée et peut redevenir prédominante, jusqu'à un nouveau cycle.

Supposons maintenant qu'à chaque étape, avec une petite probabilité, les joueurs n'arrivent pas à implémenter l'action qu'ils ont choisie : ils coopèrent au lieu de faire défection, où l'inverse. Quand Donnant-Donnant joue contre elle-même, la première erreur déclenche une vendetta destructrice (voir figure 1). De ce fait, Donnant-Donnant ne peut pas s'établir de manière stable dans la population.

Joueur 1	C … C D C D C … D C C
Joueur 2	C … C C D C D … C C C

Figure 1. Donnant-Donnant contre un autre Donnant-Donnant, avec erreurs (C signifie « coopérer » et D « faire défection »).

Une défection involontaire (en rouge) entraîne une alternance de coopération et de défection. Il faut attendre une coopération involontaire (en bleu) pour que la coopération reprenne. D'autres stratégies coopératives prennent la relève : « Donnant-Donnant contrit », par exemple, prévoit de compenser une défection erronée en coopérant deux fois de suite quoi que fasse l'autre ; « Donnant-Donnant généreux » coopère avec probabilité positive même si l'autre vient de faire défection. Ces petites différences permettent de surmonter les crises engendrées par une défection malencontreuse. Dans un monde où des erreurs peuvent se produire, il est bon de réparer les dommages causés aux autres involontairement, et de ne pas se venger systématiquement des agressions subies.

La pêche ou la tempête de neige : d'autres types de modèles

Nous avons jusqu'ici supposé que l'interaction se répétait mais que les paiements au cours du jeu ne changeaient pas. Or, imaginons un modèle de régulation volontaire de la pêche, où un grand nombre d'acteurs doivent décider de la quantité de poisson qu'ils pêcheront chaque année. Pêcher peu permet de laisser les stocks de poissons se renouveler, et est analogue à coopérer. Pêcher le plus possible épuise les ressources et est analogue à faire défection. Chaque acteur ayant une influence faible sur l'évolution du stock de poissons, il a intérêt à pêcher le plus possible même si, collectivement, il est préférable de limiter la quantité pêchée. La situation est donc analogue à celle du dilemme du prisonnier.

Toutefois, une modélisation plus précise devrait tenir compte de l'évolution du stock de poisson. « Pêcher le plus possible », par exemple, rapporte de moins en moins au fur et à mesure que les stocks de poissons s'épuisent. Les gains des acteurs à un instant donné ne dépendent donc pas seule-

ment de leurs actions présentes mais aussi de l'état global du système, qui est ici le stock de poissons. Cet état évolue en fonction des actions des différents acteurs ainsi que de divers aléas. L'étude de ce type de situation demande des modèles plus évolués qui relèvent des *jeux stochastiques*.

Une autre hypothèse que nous avons employée implicitement est qu'un individu a autant de chances d'interagir avec n'importe lequel des autres individus. En réalité, nous interagissons surtout avec nos collègues, nos amis ou nos voisins. Ceci peut être pris en compte en supposant que les individus sont répartis dans l'espace et interagissent surtout avec leurs voisins. L'analyse de ces modèles a montré que, dans le dilemme du prisonnier, la coopération émerge plus facilement lorsque les interactions sont locales que lorsque les individus interagissent avec n'importe quel autre individu.

Il est également instructif d'étudier des modèles où le conflit entre intérêts individuels et intérêt collectif n'est pas aussi brutal. Dans le jeu de la tempête de neige, par exemple, deux conducteurs voient la route coupée par la neige. Chacun a une pelle dans sa voiture, et peut soit déblayer la route (coopérer) soit attendre dans sa voiture (faire défection) en espérant que l'autre déblaiera tout seul. L'idéal est que l'autre déblaie la route tout seul, puis de déblayer la route à deux, et il est encore préférable de déblayer la route tout seul plutôt que de passer la nuit dans sa voiture, ce qui se produira si personne ne déblaie. S'il y a toujours un conflit entre l'intérêt collectif (déblayer la route) et les intérêts individuels (laisser l'autre s'en charger), la situation n'est plus celle du dilemme du prisonnier. En effet, faire défection n'est pas toujours avantageux : cela ne l'est que si l'autre coopère. L'analyse de ce jeu fait apparaître des différences plus subtiles. Ainsi, nous avons vu que le fait que les individus interagissent surtout avec leurs voisins favorisait la coopération dans le dilemme du prisonnier. Cet effet des interactions locales n'est pas présent dans le jeu de la tempête de neige. Plus généralement, si le dilemme du prisonnier a été très étudié, ce n'est pas le cas d'autres modèles, tout aussi pertinents pour étudier les phénomènes de coopération. Une analyse poussée de ces modèles est nécessaire, et encore largement à faire, pour comprendre quels résultats sont valables en général. Si les mathématiques permettent de mieux appréhender les phénomènes de coopération, il reste beaucoup d'efforts à fournir pour parvenir à une compréhension plus complète.

Pour aller plus loin

Axelrod R., (1992). *Donnant-Donnant, théorie du comportement coopératif*, éditions O. Jacob.

Sigmund K., (2010). *The Calculus of Selfishness*, Princeton University Press.

Page web :
http ://www.univie.ac.at/virtuallabs : programmes accessibles librement, pour explorer soi-même la dynamique des dilemmes sociaux.

AVC : les mathématiques à la rescousse

Emmanuel Grenier, *professeur à l'École Normale Supérieure de Lyon*

L'AVC, qui touche des milliers de personnes chaque année, est une pathologie complexe dont le diagnostic et le traitement nécessitent encore d'être améliorés. C'est un des domaines où la modélisation mathématique peut venir en aide à la recherche médicale, en complétant notamment l'expérimentation sur les animaux.

D'après l'Association d'aide aux patients et aux familles de patients victimes d'AVC, la France compterait 150000 nouveaux cas d'accident vasculaire cérébral (AVC en abrégé) chaque année. Un AVC ischémique commence par l'obstruction d'une artère du cerveau, par exemple par un petit caillot sanguin. Le territoire cérébral irrigué normalement par cette artère est alors privé de son afflux sanguin habituel. Les neurones de cette zone n'ont plus de quoi fonctionner normalement et meurent. Mais ce n'est pas tout : la présence de ces cellules mortes (et en particulier nécrosées, c'est à dire qui ont explosé) entraînant des dysfonctionnements, il est nécessaire de les éliminer. L'inflammation survient dans ce but, mais cause elle-même des dommages en faisant gonfler le tissu et en libérant des substances toxiques. Ces dommages viennent s'ajouter à ceux déjà causés par l'AVC. De plus, des radicaux libres normalement contenus dans les cellules, mais qui, après leur éclatement suite à l'AVC, se retrouvent dans l'espace extracellulaire, sont oxydés, ce qui entraîne aussi la production de substances toxiques. Les dommages causés peuvent ainsi s'enchaîner et évoluer pendant plusieurs jours voire plusieurs semaines après le début de l'accident. L'AVC est donc une pathologie très complexe, impliquant de multiples acteurs. Selon les zones touchées, le patient souffre de paralysie, d'aphasie…

L'arsenal thérapeutique dont on dispose pour guérir les AVC est hélas limité. Les cliniciens cherchent à déboucher l'artère atteinte pour rétablir au plus vite un flux

sanguin normal, mais cela n'est pas sans risque. De plus, le diagnostic de l'AVC est lui-même délicat.

Le mathématicien va traduire les connaissances biologiques et médicales sous forme d'équations. Puis, à partir de ces équations, il va faire des simulations sur son ordinateur.

Pour faire avancer la recherche dans ce domaine, les médecins et les biologistes travaillent sur des « modèles animaux », en provoquant des AVC chez des rongeurs (rats, gerbilles) et en les étudiant par imagerie ou biopsie. Le problème survient lors du passage à l'essai clinique chez l'homme, où les résultats peuvent être différents de ceux observés sur les animaux, et contraires au but recherché, comme nous le verrons par la suite.

Modéliser et simuler les phénomènes

En quoi les mathématiques peuvent-elles aider la recherche clinique ? Le principal avantage des mathématiques est qu'elles permettent de quantifier des phénomènes. Le médecin et le biologiste pensent et analysent les systèmes vivants à travers des relations, influences, réseaux… Mais ils ne peuvent pas toujours quantifier précisément les effets, leur donner une valeur numérique.

Confronté à un système complexe, un biologiste ou un médecin aura l'intuition de son fonctionnement. Il va développer des scénarios. Toutefois, les systèmes vivants sont tellement complexes qu'il est difficile d'avoir une intuition correcte et complète de leur fonctionnement.

Le mathématicien va pour sa part tenter de modéliser mathématiquement les phénomènes. Autrement dit, il va traduire les connaissances biologiques et médicales sous forme d'équations. Puis, à partir de ces équations, il va faire des simulations sur son ordinateur. Au lieu d'avoir un modèle animal, il va donc construire un modèle abstrait (des équations). Au lieu de faire des expériences, il va faire des calculs. Bien sûr, modèle abstrait et modèle animal sont deux approximations de la réalité. Le modèle abstrait ne saurait rendre compte de toute la complexité du vivant. Il est par essence faux, approximatif, en progression constante et est seulement une mise en scène d'une partie des connaissances.

Un mathématicien ne va pas découvrir une nouvelle molécule. Mais en mettant en équation les intuitions des médecins, il va tester leur cohérence et leur validité. Parfois, les divers acteurs biologiques s'organisent comme l'avaient prévu les cliniciens, mais parfois ce n'est pas le cas. La simulation numérique peut alors montrer d'autres formes de réponses collectives des acteurs, des scénarios originaux, non initialement envisagés. Elle peut aussi expliquer des résultats paradoxaux. Les premiers résultats dans ce sens commencent à se développer, en cancérologie, en neurologie ou encore en cardiologie.

Un résultat inattendu… mais prévisible mathématiquement

Au tout début d'un AVC, les neurones n'ont plus suffisamment d'énergie pour maintenir les concentrations de divers ions à leurs valeurs physiologiques. De nombreux industriels ont donc tenté de développer de nouveaux médicaments pour contrer ces phénomènes ioniques. L'idée était de bloquer le passage de certains ions à travers la paroi neuronale, en bloquant les canaux qu'ils empruntent. De multiples molécules ont été développées et testées avec succès chez le rongeur. Mais lorsqu'il a été question de passer chez l'homme, vers les années 2000, la déception fut grande : des dizaines de molécules, représentant des années de travail et des centaines de millions d'euros se sont révélées totalement inefficaces, voire nocives ! Des tests cliniques ont même été interrompus avant leur fin car ils induisaient une surmortalité chez l'homme : la molécule avait exactement l'effet inverse de celui qui était prévu !

Cet échec total s'est produit alors que les canaux ioniques sont très étudiés et très connus depuis les années cinquante (Hodgkin et Huxley ont reçu le prix Nobel de médecine pour cette découverte). Est-ce qu'un tel gâchis était prévisible ? Pouvait-on anticiper certaines des difficultés rencontrées par une analyse mathématique ?

Le modèle mathématique des échanges ioniques montre que le rapport cellules gliales/neurones a un effet crucial sur l'efficacité des bloqueurs de canaux ioniques. Or, l'effet de ce rapport est difficile à trouver par intuition pure.

Ici, pour construire le modèle mathématique, la démarche consiste d'abord à faire la liste des ions importants, ainsi que des canaux ioniques, pompes, échangeurs… dans lesquels ces ions circulent. Puis, il

s'agit d'établir tous les paramètres physiologiques liés aux objets listés (concentrations, nombre de portes, potentiels d'ouverture…). Ensuite, le tout est mis en équation (on utilise les équations de Nernst, les potentiels de membrane…). Ce travail long et délicat est basé sur une démarche développée depuis quelques décennies, notamment pour le cœur, entre autres par l'équipe du Professeur Denis Noble, d'Oxford.

Le modèle construit indique bien des effets très positifs des différents médicaments chez le rongeur… mais aussi des effets très faibles, voire négatifs, chez l'homme ! Autrement dit, si dans les années 90 on avait mis en équation les échanges ioniques dans un AVC et simulé ces modèles, on aurait pu prédire ce qui s'est passé.

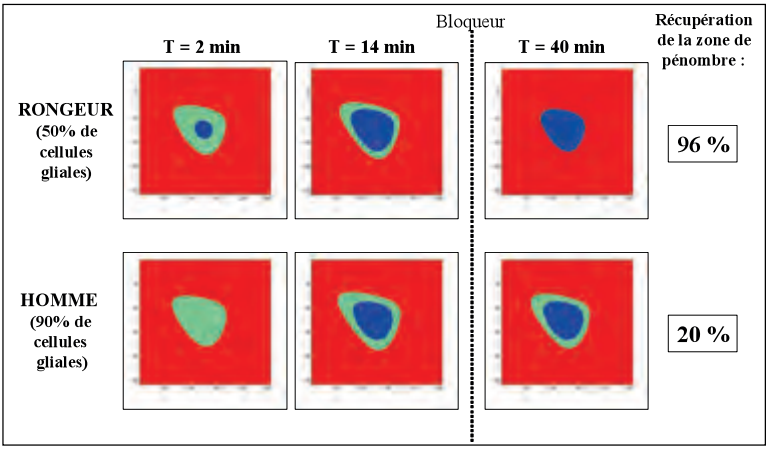


Figure 1 : Résultats de simulations obtenus avec le modèle de l'AVC permettant de voir l'évolution des différentes zones ischémiées sur 40 minutes suite à une obstruction artérielle (à t = 0 min) et suite à l'introduction d'un bloqueur d'un canal ionique sodique spécifique (à t = 20 min) pour le « modèle rongeur » et pour le « modèle homme » (différant par leur proportion de cellules gliales). La zone infarcie (où les cellules sont totalement mortes) est en bleu, la zone de pénombre (où les cellules sont endommagées ou affaiblies, et peuvent soit mourir, soit se réparer et survivre) en vert et la zone saine en rouge. L'enjeu de tout traitement est évidemment la zone de pénombre, le but étant qu'elle évolue au maximum vers l'état sain. Les résultats de ces simulations montrent que, suite à l'introduction de ce bloqueur, la récupération de la zone de pénombre est très importante dans le cas du rongeur alors qu'elle est assez faible dans le cas de l'homme. (d'après M.A. Dronne, E. Grenier, G. Chapuisat, M. Hommel, J.-P. Boissel, Progress in Biophysics and Molecular Biology, 2008, vol. 97).

Peut-on comprendre l'échec clinique à partir du modèle ? Ici, le point crucial réside dans la composition différente des tissus nerveux du cerveau de l'homme et du rongeur. Dans le cerveau, on distingue les neurones et les cellules gliales (cellules de support, qui nourrissent les neurones, les protègent, contribuent à l'homéostasie…). Chez le rat, il y a 2 cellules gliales pour un neurone, alors que chez l'homme il y a près de 9 cellules gliales pour un neurone. Le modèle mathématique des échanges ioniques montre que ce rapport cellules gliales/neurones a un effet crucial sur l'efficacité des bloqueurs de canaux ioniques : plus ce rapport augmente et moins le médicament est efficace. Les cellules gliales étant proportionnellement plus nombreuses chez l'homme, l'effet du médicament chute et devient faible, voire négatif dans certains cas. Or, l'effet de ce rapport cellules gliales/neurones est difficile à trouver par intuition pure. C'est pourquoi le recours à un modèle mathématique et à sa simulation était nécessaire pour tester l'influence de ce paramètre de façon plus précise et de dégager ce scénario inattendu.

Point de vue sur les mathématiques françaises depuis l'étranger

John Ball, *professeur à l'Université d'Oxford*

Les mathématiques françaises ont la réputation d'être parmi les meilleures au monde. Qu'est-ce qui explique leur exceptionnelle qualité ? Voici un décryptage de cette particularité hexagonale à travers le regard du mathématicien britannique John Ball.

Pierre-Louis Lions (copyright FSMP)

On pourrait difficilement contester que la France se situe actuellement parmi les meilleurs pays au monde en mathématiques. Elle occupe peut-être le second rang, devancée seulement par les États-Unis, voire le tout premier si l'on se rapporte à la taille de sa population. Bien entendu, de telles affirmations doivent être traitées avec beaucoup de prudence. Elles sont symptomatiques de cette tendance moderne à vouloir établir un ordre pour tout, que ce soit en classant les gens les plus grands, les plus riches ou les plus puissants du monde, les meilleurs footballeurs de tous les temps, les meilleures universités, et ainsi de suite. Les mathématiques ne sont pas un sport où s'affrontent divers pays du monde. Il n'y a pas de vainqueur de la Coupe du Monde de Mathématiques. Il est plus juste de considérer que les pays possèdent des communautés mathématiques distinctes, avec des styles mathématiques différents et des points forts dans des disciplines différentes. Quelques indicateurs, cependant, viennent confirmer l'idée d'une suprématie française. Par exemple, le palmarès des médailles Fields, distinctions les plus pres-

tigieuses historiquement, décernées par l'Union Mathématique Internationale lors du Congrès International des Mathématiciens qui se tient tous les quatre ans (la dernière fois en 2010, à Hyderabad, en Inde, et la prochaine, en 2014 à Séoul, en Corée du Sud). A ce jour, plus d'un cinquième des médailles Fields ont été attribuées à des Français. Ensuite, les conférenciers de ce même Congrès International. Ceux-ci sont choisis à la suite d'un processus rigoureux par des comités formés par plus d'une centaine de mathématiciens de premier plan. Être invité à parler au Congrès International est considéré comme un grand honneur. Or la France est toujours représentée par un contingent important de conférenciers invités, ce qui témoigne de la profondeur au plus haut niveau de ses mathématiques.

A ce jour, plus d'un cinquième des médailles Fields ont été attribuées à des Français.

Cédric Villani (copyright J. Cotera)

Héritage intellectuel et organisation académique

Comment expliquer ce caractère exceptionnel des mathématiques françaises ? Il y a en premier lieu l'héritage intellectuel d'un nombre extraordinaire de grands mathématiciens français, comme d'Alembert, Cauchy, Fermat, Fourier, Galois, Hermite, Laplace, Lagrange, Liouville, Monge, Poisson et, dans la première moitié du XX^e^ siècle, Poincaré, Hadamard, Lebesgue et Leray. La plupart d'entre eux ont donné leurs noms à des rues de Paris, un honneur rarement accordé à des mathématiciens dans d'autres pays.

Alain Connes (copyright FSMP)

Il y a eu aussi la création des grandes écoles à partir de la fin du XVIII^e^ siècle, où l'enseignement des mathématiques figurait en bonne place. Dans les écoles françaises, la maîtrise d'outils mathématiques était, et est encore, considérée comme un moyen d'évaluer l'intelligence. Plus récemment, la

disponibilité de nombreux postes de chercheurs à plein temps (donc sans charge d'enseignement) comme ceux du CNRS a offert aux chercheurs l'occasion de profiter de périodes de concentration prolongée précieuses pour la création et le développement de nouvelles mathématiques. Ce type de postes est rare, voire inexistant, dans les autres pays. Certains diront aussi que la qualité des mathématiques françaises a été maintenue par le pilotage au niveau national de certaines étapes des processus de recrutement et de promotion (par opposition au système actuellement mis en place en France permettant aux universités autonomes de prendre ces décisions purement localement).

Une tradition locale dans une discipline de plus en plus internationale

Comme les joueurs de football, les mathématiciens, quelle que soit leur origine, sont de plus en plus internationaux. Par exemple, mon propre département à Oxford compte parmi ses membres des citoyens allemands, américains, chinois, coréens, costariciens, danois, français, grecs, hongrois, roumains et russes. De nombreux départements de mathématiques aux États-Unis ont des compositions semblables. Cet effet est beaucoup moins prononcé en France, où les universitaires des meilleurs départements sont majoritairement français. Ceci est largement, on peut le supposer, dû à la barrière de la langue, et on voit des situations similaires dans d'autres pays européens. Cela a des aspects positifs et négatifs. Certes, il existe un socle culturel commun à la plupart des chercheurs, ce qui facilite les collaborations ; a contrario, différentes nationalités apportent des perspectives et des approches différentes. D'un autre point de vue, cependant, les mathématiques en France sont admirablement internationales dans leurs perspectives, grâce à la tradition de coopération avec des pays moins développés, en particulier en Afrique, en Amérique du Sud, en Inde et au Cambodge. Une partie de cette activité est organisée grâce à l'excellent Centre International de Mathématiques Pures et Appliquées basé à Nice. Créé en France et reconnu par l'UNESCO, le CIMPA est un organisme international œuvrant pour l'essor des mathématiques dans les pays en voie de développement.

La France n'est pas non plus à l'abri des pressions tendant à modifier les méthodes d'enseignement dans les écoles…

Sylvia Serfaty
(copyright Olivier Boulanger)

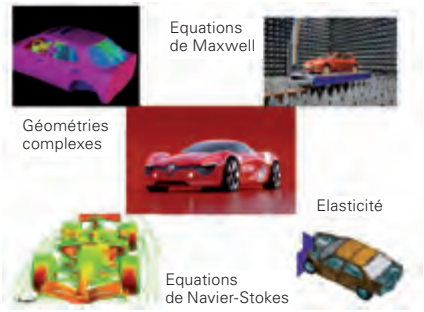
L'avenir des mathématiques françaises

Que réserve l'avenir aux mathématiques françaises ? L'histoire montre la grande importance de la tradition dans le maintien d'une recherche de haute qualité, et la France se trouve en position de force à cet égard. D'un autre côté, de nombreux changements ont lieu actuellement dans l'organisation de l'enseignement supérieur et de la recherche française, un exemple étant l'autonomie accrue des universités. Ces changements doivent être opérés avec soin afin de permettre une saine évolution tout en préservant les nombreuses qualités du système actuel. La France n'est pas non plus à l'abri des pressions tendant à modifier les méthodes d'enseignement dans les écoles, telles que l'élargissement du programme d'études menant à une réduction des heures allouées aux mathématiques, ce qui pourrait avoir des effets négatifs sur le niveau de préparation des étudiants en mathématiques entrant dans les grandes écoles et les universités. Comme toujours, un effort constant est nécessaire pour préserver la tradition d'excellence.

EDP à la française

John Ball, *professeur à l'Université d'Oxford*

Spécialiste des équations aux dérivées partielles, John Ball raconte ici comment ce domaine des mathématiques appliquées a connu un formidable essor en France, grâce en particulier à une figure emblématique : le professeur Jacques-Louis Lions.



La France est un acteur important sur la scène mathématique mondiale dans presque toutes les branches de la discipline, avec une représentation particulièrement forte en algèbre, en géométrie, en théorie des nombres, en probabilités, ou encore dans le domaine des équations aux dérivées partielles (EDP). C'est de cette dernière école dont je parlerai ici car je m'en sens particulièrement proche, du fait de mon parcours et de mes longs et heureux séjours à Paris, au Laboratoire Jacques-Louis Lions, l'un des principaux centres mondiaux pour les EDP et leur analyse numérique. A Paris vivent plus d'experts de ce domaine que dans toute autre ville du monde. En fait, il y en a plus que dans bien des pays !

Le rôle essentiel de Jacques-Louis Lions

La création de cette école a été en grande partie le travail d'un homme, Jacques-Louis Lions. Après des études à l'École normale

supérieure de 1947 à 1950, et l'obtention de son doctorat sous la direction de Laurent Schwartz, Lions a commencé à s'attaquer au problème de résoudre, tant d'un point de vue théorique que numérique, diverses équations aux dérivées partielles qui servaient à décrire certains phénomènes réels. Il a perçu la nécessité, pour les mathématiques appliquées, de relever les défis résultant de l'utilisation croissante des simulations sur ordinateur. Comment savoir que de telles simulations produisaient la bonne réponse ? S'appuyer sur l'intuition physique ne suffisait plus dans le monde non-linéaire. Une compréhension fondamentale du comportement des solutions était nécessaire, afin de fournir une trame pour des schémas numériques.

La théorie moderne des équations aux dérivées partielles se construisait également de l'autre côté de l'Atlantique, au Courant Institute, à New-York, conduite par des personnalités tels que Kurt Friedrichs, Peter Lax, Louis Nirenberg, Fritz John, Joe Keller et Jürgen Moser, ainsi qu'en Russie, entre autres par les grandes mathématiciennes Olga Ladyzhenskaya et Olga Oleinik. Mais l'approche adoptée par Lions était plus abstraite, basée sur l'analyse fonctionnelle et en particulier la *convergence faible*, c'est-à-dire la convergence des moyennes, motivée par les travaux de Jean Leray et Eberhard Hopf sur les équations de Navier-Stokes, qui décrivent le flux de fluides visqueux tels que l'huile et l'eau.

Lions a synthétisé et mis au point un éventail de techniques permettant l'utilisation de la convergence faible dans l'étude des équations aux dérivées partielles non linéaires. Le corpus théorique résultant de ces travaux s'est révélé suffisamment souple pour permettre de traiter une large gamme de problèmes importants en mécanique et en physique. Une fois maîtrisées, ces idées sont devenues très faciles à utiliser. Bien qu'elles ne permissent pas de répondre aux questions techniques les plus profondes, qui exigeaient des méthodes d'analyse spécifiques, leur caractère systématique signifiait qu'elles pouvaient être utilisées efficacement dans de nombreuses situations. L'essor de cette école, propulsé par de nombreux étudiants de Lions ayant eux-mêmes acquis une renommée internationale, a été extraordinaire. Et son influence sur l'industrie française a été profonde.

Jacques-Louis Lions au Centre national d'études spatiales

L'essor de cette école a été extraordinaire et son influence sur l'industrie française a été profonde.

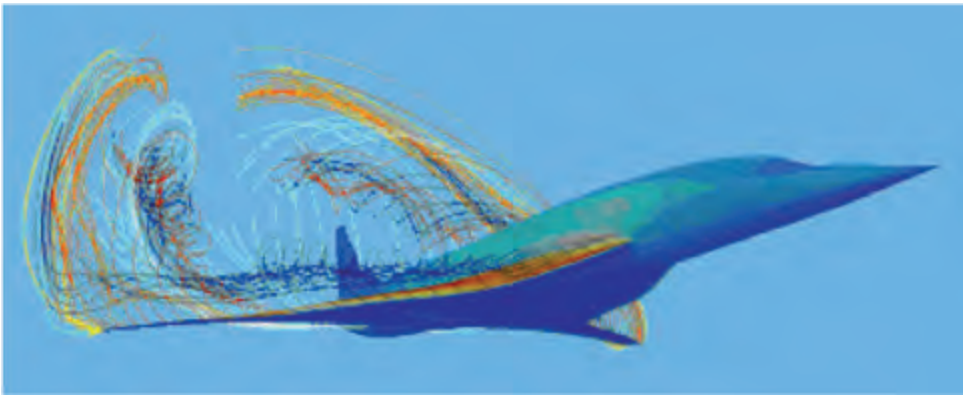
Cauchy et les modèles mathématiques

L'approche des équations aux dérivées partielles adoptée en France est une manière particulière, très influente et à l'impact profond, de faire des mathématiques appliquées. Mais ce n'est pas la seule façon de faire. Un contraste intéressant peut être relevé avec l'un des plus grands mathématiciens français, le baron Augustin-Louis Cauchy (1789-1857), dont le nom est connu de tous les étudiants en mathématiques, par exemple grâce à la formule intégrale de Cauchy en analyse complexe et aux suites de Cauchy. Il semble que Cauchy n'ait pas été un collègue très agréable, mais son approche des mathématiques était admirable et tout à fait moderne en ce qu'il ne faisait pas de distinction artificielle entre mathématiques pures et appliquées. En effet, Cauchy a fondé l'analyse complexe en même temps qu'il introduisait la notion de contrainte dans un milieu continu, d'une importance incalculable pour l'ingénierie moderne, et écrivait pour la première fois les équations de l'élasticité linéaire, en faisant un lien entre les théories des solides à l'échelle atomique et celle des milieux continus (même si les atomes n'avaient pas encore été observés). Ainsi, Cauchy assumait la responsabilité des modèles mathématiques sur lesquels il travaillait, ainsi que celle de leur analyse. Une telle attitude n'est pas tout à fait dans la tradition française moderne des mathématiques appliquées, dans laquelle la création de modèles mathématiques est largement laissée aux chimistes, physiciens, biologistes, mécaniciens et aux économistes. Pour être tout à fait juste, l'attitude moderne est aussi celle de nombreux mathématiciens dans le monde entier, mais il y aurait certainement beaucoup à gagner en associant la modélisation et l'analyse. Les travaux de Cauchy en offrent un bon exemple à travers sa découverte du théorème de décomposition polaire pour les matrices, qui fut une conséquence directe de son travail sur la mécanique des milieux continus.

Si les travaux de Lions et d'autres analystes appliqués ont été une source d'inspiration mathématique, ils m'ont aussi apporté le réconfort psychologique de savoir qu'il n'y avait pas lieu d'opposer bonnes mathématiques appliquées et raisonnements rigoureux !

Mathématiques pures ou appliquées ?

Je n'ai jamais bien compris ce qui a conduit à la polarisation des mathématiques pures d'un côté et appliquées de l'autre pendant la majeure partie du XX^e siècle, tant celle-ci me semble en contradiction avec l'esprit de l'unité de la science et des mathématiques représentées par plusieurs des plus grands mathématiciens universels, comme Archimède, Newton, Euler, Gauss, Riemann, Poincaré et Cauchy. En Grande-Bretagne, G.H. Hardy en fut peut-être en partie responsable, même si, ironiquement, ses mathématiques ont trouvé de nombreuses applications aux équations aux dérivées partielles résultant des sciences de la matière. En France, l'école Bourbaki a adopté dans ses traités (souvent précieux) une approche pure sans compromis, distanciant systématiquement les mathématiques fondamentales des applications. Quand j'étais étudiant en doctorat dans les années 1960 et au début des années 1970, j'ai certainement ressenti en Grande-Bretagne que la majorité des mathématiciens appliqués n'accordaient pas aux théorèmes une grande valeur (et inversement, les spécialistes de mathématiques dites pures ne s'intéressaient guère aux applications). C'est pourquoi, si les travaux de Lions et d'autres analystes appliqués ont été une source d'inspiration mathématique, ils m'ont aussi apporté le réconfort psychologique de savoir qu'il n'y avait pas lieu d'opposer bonnes mathématiques appliquées et raisonnements rigoureux !



F.A.Q. (et idées reçues) sur les mathématiciens

Quels sont les débouchés des cursus de mathématiques ? Comment devient-on mathématicien(ne) ? En quoi consiste ce métier ? Des professionnels, hommes et femmes, répondent à ces questions et tordent le cou aux idées reçues.

Quels métiers peut-on exercer avec un diplôme de mathématiques ?

Edwige Godlewski,
responsable de la spécialité Ingénierie mathématique du Master de l'UPMC :

Les secteurs où les jeunes formés en mathématiques trouvent une embauche couvrent presque tous les domaines : automobile, aéronautique, espace, énergie, transport, télécommunications, traitement du signal, de l'image, industrie pharmaceutique, secteur biomédical, génie civil, environnement, logistique, banque, assurance, prévision, media, etc. Le titre de l'emploi comporte souvent le terme ingénieur (si c'est au niveau master), ou analyste, chargé d'études, etc. Cela peut être statisticien, très rarement mathématicien. Le contenu du travail est lui aussi varié, l'utilisation des compétences en mathématiques pouvant rester très majoritaire (métiers nécessitant une expertise technique pointue, en méthodes statistiques ou numériques par exemple) ou se limiter, parfois dès le début, aux qualités transverses traditionnellement attendues de la discipline : abstraction, rigueur, esprit d'analyse, etc.

D'une façon générale, la demande de jeunes formés aux méthodes statistiques est très importante, et les applications sont plutôt bien connues (actuariat, biostatistique, statistique industrielle, études de marketing, prévision de consommation, sondages, épidémiologie, etc.). Mais il existe aussi de réels débouchés pour les

masters formant dans tous les domaines des mathématiques appliquées, à condition que la formation comporte un minimum de maîtrise des outils informatique et logiciel, et que les jeunes diplômés y apprennent à appréhender des questions du point de vue de l'entreprise, sans être enfermés dans l'idée de ne pouvoir appliquer que des théorèmes. Un stage en entreprise en fin de cursus représente enfin une étape quasi indispensable. Une fois ces conditions réunies, les responsables de formations de masters type Ingénierie mathématique constatent que leurs étudiants, s'ils manifestent un certain dynamisme, vont trouver assez rapidement un emploi intéressant et que leur formation leur permettra soit de rester dans un domaine d'expertise scientifique, soit d'évoluer vers la conduite de projets ou vers d'autres postes.

Faut-il aller jusqu'au doctorat pour faire un métier qui utilise des mathématiques ?

Adeline Samson,
responsable de la licence professionnelle «Statistique et Informatique décisionnelle pour la Santé», de l'IUT Paris Descartes :

Non ! C'est la richesse des débouchés en mathématiques. Des métiers qui utilisent des mathématiques sont accessibles avec un master, un diplôme d'ingénieur, mais aussi un DUT ou une licence professionnelle. La différence entre ces métiers et ceux accessibles après un doctorat est dans la façon d'utiliser les mathématiques. Un docteur sera amené à développer, inventer de nouveaux modèles mathématiques, de nouvelles méthodes d'analyse, de calcul, d'estimation statistique. Avec un diplôme à Bac+5 ou Bac+2, +3, on demandera de savoir utiliser les mathématiques au travers de logiciels ou de méthodes déjà développées par des chercheurs et surtout de savoir choisir la méthode ou le modèle le plus adapté à la problématique sur laquelle on travaille.

Faculté de Jussieu

En quoi consiste le métier de chercheur en maths ? Où travaillent les chercheurs ? Que font-ils ?

Cédric Villani,
professeur à l'Université Claude Bernard Lyon 1 et directeur de l'Institut Henri Poincaré (CNRS/UPMC):

Dans les universités, dans les instituts de recherche et les centres de rencontres scientifiques, les chercheurs mathématiciens travaillent à développer des théories mathématiques et à transmettre leur savoir. C'est ainsi que de nouveaux théorèmes continuent à

naître, plus maintenant que jamais. On estime le nombre de nouveaux théorèmes, chaque année, entre 100000 et un million ! Cette naissance n'est pas un processus simple : elle est faite d'essais et d'erreurs, de rencontres et de discussions, de travail patient et d'éclairs, d'inspiration glanée on ne sait où. Les outils de travail : papier et crayon bien sûr ; ordinateur, soit pour effectuer des calculs soit pour communiquer ; de la documentation, il en faut toujours ; des échanges, des échanges et des échanges. Les chercheurs ont une conscience aiguë de former une communauté pleinement internationale. Ils voyagent sans cesse à travers le monde pour donner des exposés dans lesquels ils évoquent leurs découvertes, leurs problèmes, leurs espoirs.

Y a-t-il encore des choses à trouver en maths ?

Filippo Santambrogio,
professeur à l'Université Paris-Sud :

Bien sûr ! Et les questions qu'on se pose sont de nature très variée. D'une part il y en a de très dures qui traînent depuis des siècles. Certaines sont d'ailleurs très simples à raconter : par exemple, la conjecture de Goldbach (prouver que tout nombre pair supérieur à 2 peut s'écrire comme la somme de deux nombres premiers : elle est vraie pour tous les nombres que les ordinateurs ont testés, mais personne ne sait prouver qu'elle l'est toujours !).

D'autre part, il y a les questions issues des applications. Quand on s'est mis, avec mon collègue Bertrand Maury, à étudier des modèles de mouvement de foules, on a vu qu'on avait besoin de prendre la limite d'une relation non linéaire entre pression et densité : ça se raconte beaucoup moins bien que Goldbach, et on ne l'aurait jamais trouvé intéressant sans en voir l'application… mais c'est comme ça que la plupart des mathématiques se font : on étudie ce dont on a besoin, sans savoir à l'avance qu'on le fera !

Faut-il être un génie pour devenir mathématicien ?

D'après Terence Tao, l'un des meilleurs mathématiciens actuels (lauréat de la médaille Fields en 2006) : « La réponse est non, absolument pas. Apporter des contributions belles et utiles aux mathématiques nécessite de beaucoup travailler, de se spécialiser dans un domaine, d'apprendre des choses dans d'autres domaines, de poser des questions, de parler aux autres mathématiciens, et de réfléchir aux grandes lignes du paysage mathématique considéré. Et oui bien sûr, une intelligence raisonnable, de la patience, de la maturité sont aussi nécessaires. Mais en aucun cas on aurait besoin de posséder une sorte de gène magique du génie mathématique ou d'autres superpouvoirs, qui inspireraient spontanément, et à partir de rien, des idées profondes ou des solutions totalement inattendues à des problèmes. »

Par ailleurs, rien qu'en France, environ 4000 mathématiciens travaillent dans les laboratoires de recherche des universités, sans compter tous les mathématiciens travaillant

dans l'industrie… tous ne sont certainement pas des génies !

En quoi consiste un doctorat en maths ?

Amélie Rambaud,
thèse à l'Institut Camille Jordan sous la direction de Francis Filbet :

Trois ans, ça paraît long, mais c'est ce qu'il faut ! Pour explorer, se familiariser avec le monde de la recherche (en maths j'entends !). Il s'agit de s'attaquer à un problème ouvert de mathématiques proposé par un directeur, mais qui peut conduire en cours de route à d'autres pistes, si par exemple le sujet initial soulève des questions semblant plus pertinentes. En fait, c'est assez subjectif, instinctif. Dans une thèse, il y a souvent des moments où l'on est perdu, bloqué dans des calculs par exemple… Ce brouillard peut durer ! Mais en persévérant, en travaillant, il arrive un moment de déblocage, et on « accouche » d'un résultat, c'est l'impression que j'ai du moins ! Puis vient une phase de rédaction, de défense de son résultat devant d'autres chercheurs. C'est un exercice difficile et en anglais, mais très intéressant, qui permet de resituer son travail dans son contexte, le mettre en perspective, bref, lui trouver sa place dans le foisonnement des publications de mathématiques. Lorsque l'on a terminé, on en retire une grande satisfaction : voir son travail publié est un accomplissement, motivant pour poursuivre ses recherches ! Ainsi, je dirais que faire une thèse, c'est mettre un pied dans l'univers fou de la recherche. Quel qu'en soit le résultat, c'est enrichissant, mais attention, on peut attraper le virus, et continuer la recherche !

Peut-on travailler dans l'industrie avec un diplôme de mathématiques ?

Marc Bernot,
ingénieur de recherche chez Thales Alenia Space, Cannes :

Absolument ! J'ai rejoint le service de recherche de Thales Alenia Space (une entreprise qui fabrique des satellites de toute sorte) il y a quatre ans. Je travaille sur la qualité optique des télescopes, ce qui me met en contact avec des collègues aux profils variés (opticiens, mécaniciens…). Cette diversité s'est révélée très stimulante après une formation 100% mathématiques (École normale supérieure, agrégation, M2, doctorat) sans passer par une école d'ingénieur.

Mais que font les mathématiques là-dedans, si la modélisation d'un télescope se base sur la physique ? Eh bien c'est que souvent, la seule connaissance du modèle ne suffit pas… Une étape importante est la modélisation numérique, et sa mise en œuvre rapide et précise ouvre un monde de questionnement. Les techniques ne sont pas toujours les plus complexes que j'ai vues lors de ma formation : c'est le cadre d'application qui est complexe et multi métier. Pourtant ma formation (même la théorie de Galois, dont je ne me sers plus du tout) est essentielle : c'est grâce à elle que je ne ressens pas d'effroi face à la technicité, mais un désir de clarifier et trouver l'essentiel.

Sébastien Marque,
directeur du département Biométrie de Danone Research :

Oui ! Pour ma part, je suis titulaire d'une maîtrise de mathématiques appliquées et d'un doctorat de biostatistique. Sur la vingtaine de personnes qui travaillent dans mon département, les deux tiers sont titulaires d'un master professionnel du domaine de la statistique ou biostatistique. La plupart des autres sont des data managers, métier que l'on peut exercer à des niveaux et avec des diplômes variés : après un diplôme à bac +2 (DUT STID, diplôme de mathématiques professionnalisant), bac +3 (licence professionnelle) et aussi avec des diplômes plus proches de l'informatique voire même de la chimiométrie.

En plus des compétences techniques en mathématique/statistique obtenues par leur diplôme, j'attends des personnes qui travaillent dans mon équipe une grande ouverture d'esprit. Ce sont elles qui font les ponts en traduisant la question scientifique (de nature biologique ou épidémiologique très souvent) en un problème statistique et qui, une fois celui-ci résolu, doivent à nouveau faire un pont en communiquant leurs résultats de manière didactique afin de s'assurer que les messages seront bien compris et utilisés. Et puis bien sûr, le monde évolue vite, elles doivent sans cesse se renouveler et s'adapter, tant sur l'organisation des entreprises que dans le domaine mathématique et statistique.

J'aime les mathématiques et j'apprécie vraiment le travail en groupe, est-ce compatible ?

Travail en solitaire ou en groupe, ceci dépend avant tout de la personnalité de chaque chercheur. On fait toujours énormément de rencontres dans les métiers de la recherche, et le brassage et les échanges d'idées ont une importance cruciale dans le développement de la recherche en mathématiques. Les meilleures idées peuvent aussi bien venir d'une discussion animée après le repas, que de l'isolement d'une nuit blanche. Le travail, même individuel, n'est de toute façon jamais isolé ne serait-ce qu'au travers des communications électroniques et des avancées qui se construisent les unes à partir des autres…

La recherche en mathématiques est-elle compatible avec une vie de famille ?

Fabienne Castell,
professeur d'Université à Aix-Marseille, trois enfants :

Cela doit être compatible puisque de fait, je suis à la fois mère de famille et mathémati-

cienne, mais il est difficile pour moi de juger de l'intérieur de la qualité de ma propre vie de famille ! Ce que je peux dire, c'est que je n'ai pas l'impression que les maths aient « brimé » ma vie familiale, en tous les cas ni plus ni moins que beaucoup d'autres métiers que j'aurais pu exercer. Que l'on soit homme ou femme, il y a toujours un équilibre à trouver entre vie de famille et vie professionnelle, et cette question n'est pas spécifique aux mathématicien(ne)s. Par exemple, je n'ai pas calculé ou organisé ma vie personnelle (je pense essentiellement à la naissance de mes enfants) par rapport aux grandes étapes de mon métier. Je me déplace sûrement moins que des personnes qui n'ont pas de famille, et si cela a certainement un impact sur ma carrière, il n'est pas si important. Mon travail est prenant et mes enfants le savent et sont habitués. Ce qui est spécifique dans ce métier est que c'est un travail de création et que parfois, même de retour à la maison, j'ai l'esprit préoccupé par mes recherches et j'ai du mal à penser à autre chose.

Je n'arrive pas à me décider entre les maths et la biologie : pourrai-je faire de la biologie si je suis un cursus en maths ?

Hélène Morlon,
chercheuse en biologie au Centre de Mathématiques Appliquées de l'École polytechnique :

Oui, tout à fait. De plus en plus de domaines de la biologie s'appuient sur des modèles, et les personnes ayant de fortes compétences en maths et un intérêt pour la biologie sont recherchées. Une bonne approche peut être de favoriser, en parallèle aux maths pures ou appliquées, les formations en statistique, en analyse numérique, en programmation… et en biologie bien entendu ! Passionnée de biologie mais aimant aussi beaucoup les maths, j'ai choisi un cursus en maths après le bac pour me laisser un maximum d'ouverture. Après une prépa en maths et des études à l'École normale supérieure de Cachan (dont le passage de l'agrégation de maths), j'ai pu intégrer un master d'écologie et commencer une reconversion vers la biologie. C'est définitivement un challenge, mais l'interdisciplinarité est de plus en plus valorisée. Le challenge en vaut la peine, puisqu'il permet de concilier plusieurs passions ! Plusieurs de mes collègues ont suivi un parcours similaire, et je ne pense pas qu'aucun regrette.

Peut-on créer son entreprise quand on est mathématicien ?

Stéphane Mallat,
professeur à l'École normale supérieure et créateur de l'entreprise Let it Wave :

Clairement oui, et de nombreux mathématiciens l'ont fait avec succès aussi bien en France qu'à l'étranger, et surtout aux États-Unis. Les mathématiques sont une source extraordinaire d'innovation technologique permettant de mettre au point de nouveaux produits et d'améliorer les processus industriels. C'est le cas en informatique, dans les télécommunications, l'aéronautique, la santé et dans la plupart des domaines industriels.

Le métier d'enseignant-chercheur est une excellente formation pour devenir entrepreneur, contrairement à beaucoup d'idées reçues. Créer une entreprise est aussi un projet de recherche. On part d'une première idée pour aborder la recherche du marché, de l'implémentation technologique et créer une équipe, avec beaucoup de surprises et de tournants en chemin. L'expérience d'enseignant aide à expliquer et convaincre de la valeur de ses idées.

Il n'est pas nécessaire de connaître la finance, le commerce, le marketing ou le management, mais simplement d'avoir envie d'apprendre, pour découvrir de nouveaux horizons. C'est ainsi que je l'ai vécu.

Comment gagne-t-on sa vie dans les métiers des mathématiques ?

Cela dépend évidemment beaucoup du domaine d'exercice… Les salaires d'embauche des jeunes diplômés de mathématiques se dirigeant vers la finance sont certainement attractifs, même si une part de mystère (voire de fantasme) est entretenu sur le niveau réel des rémunérations dans le domaine. Par comparaison, les carrières dans le public (enseignement à l'université, recherche publique…) sont bien sûr moins rémunératrices. D'un extrême à l'autre, les revenus semblent au final très variés, de l'enseignement à la banque en passant par la recherche industrielle ou l'environnement et la santé…


IMPRIM'VERT®


PEFC™
10-31-1238

« Mais à quoi ça sert ? » Cette question, les collégiens la posent souvent à leur professeur de mathématiques. Elle est bien sûr légitime pour des élèves, mais l'est tout autant pour des étudiants et des adultes, qu'ils exercent ou non une activité liée à la science, et encore plus pour des adultes investis de responsabilités collectives.

En effet, même si de tout temps, les mathématiques ont été liées à de multiples activités : administratives, techniques, scientifiques ou culturelles, on assiste depuis une quarantaine d'années à une explosion continue du nombre de domaines dans lesquels la recherche mathématique la plus avancée se révèle indispensable.

De la cryptographie au traitement d'images, de la compréhension du climat à celle de la biodiversité, de la lutte contre les spams au fonctionnement des moteurs de recherche, de la détection des maladies génétiques à la prévention des AVC, de l'univers académique à celui des entreprises, les applications des mathématiques ne se comptent plus et couvrent un large spectre de plus en plus étendu. Inversement, de façon concomitante, les questions posées par le développement de la technologie, de la biologie ou de la gestion des données massives – pour ne citer que ceux-là – suscitent la création et le développement de nouvelles théories mathématiques.

Parmi les vingt-quatre articles de cet ouvrage, trois proviennent de *l'Explosion des mathématiques* éditée en 2002 ; les autres sont des textes originaux. Tous illustrent l'ubiquité toujours croissante des mathématiques dans le monde d'aujourd'hui. Le point de vue du mathématicien anglais John Ball sur l'école mathématique française et quelques questions-réponses pour mettre à jour les idées reçues sur le métier de mathématicien complètent aimablement l'ouvrage.

Société Française de Statistique

Société de Mathématiques Appliquées et Industrielles

Société Mathématique de France

Fondation Sciences Mathématiques de Paris

CAP MATHS

ISBN 978-2-85629-375-1