Maths Société Express

Comité international des jeux mathématiques

Sommaire

Retrouvez les compléments d'enquête de la brochure sur www.cijm.org
(« Productions CIJM » puis « Maths Express ») :

Les mystérieux carnets de Ramanujan

Archéologie de l'impression 3D,
une nouvelle littéraire et mathématique

Introduction

Marie Ekeland

Co-fondatrice du fonds d'investissement Daphni
Membre du Conseil national du numérique

Il est loin le temps où je maniais habilement les intégrales, les nombres imaginaires, les dérivées premières ou secondes et tous ces concepts mathématiques qui m'ont accompagnée tout au long de mes études. Ils sont aujourd'hui comme de vieux amis que j'aurais perdus de vue depuis longtemps.

Il ne se passe en revanche pas un jour où je ne résous des problèmes complexes en appliquant un raisonnement mathématique. Ces réflexes de pensée et de modélisation, je les ai acquis de manière définitive. Tout d'abord, poser le problème : délimiter son contour et identifier ses données et hypothèses essentielles. Ensuite, assembler ces données et hypothèses afin de tracer un chemin de pensée, d'articuler un raisonnement. Enfin, déboucher sur une issue, une solution qui se traduit la plupart du temps par une prise de décision. Si l'application de cette décision apporte effectivement une solution au problème, le sujet est clos, sinon, on remet en question les hypothèses de départ et on repart dans une boucle de pensée.

Cette mécanique de pensée fondamentale, j'ai eu l'occasion de l'expérimenter dans des missions professionnelles très différentes avec toujours la même puissance de réalisation.

Dans mon activité de « codeuse », cela m'aidait à concevoir de nouveaux programmes informatiques en décomposant le comportement global souhaité du programme en une suite logique d'actions. Cela s'appelle de l'*algorithmique* !

Dans mon métier de « capital-risqueuse », cela me permet de mieux analyser les chances de succès d'une *startup* et de prendre une décision d'investissement en ayant une conscience plus claire des risques pris.

Enfin dans mes incursions au sein du monde politique, au travers de l'association France Digitale, cette approche pragmatique de résolution de problèmes, partant de l'analyse de données issues du terrain tout en cherchant une solution équitable et performante, nous a permis de pousser avec succès des réformes qui étaient essentielles au développement de l'entrepreneuriat numérique français.

La dimension collective de cette approche est essentielle, et les jeux mathématiques ont joué un rôle important dans son apprentissage à plusieurs, éduquant nos sens en groupe, en résonance avec d'autres logiques et avec un plaisir partagé.

Pour moi, c'est avant tout cela les mathématiques : un bagage essentiel pour construire sa vie, au gré de ses envies, de ses rencontres ou de sa curiosité, en s'appuyant avec confiance sur sa raison en plus de son intuition.

M.E.

Mathématiques et réseaux sociaux

Daniel Kious

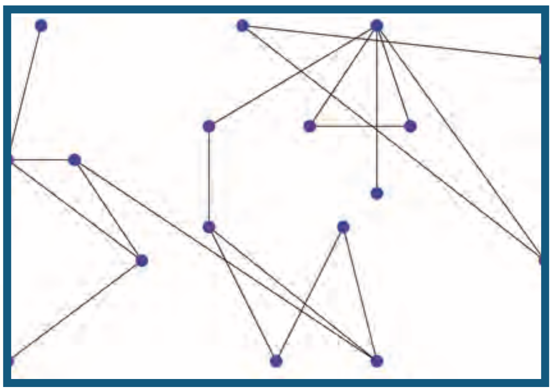
Enseignant-chercheur
à l'École polytechnique fédérale de Lausanne

Comment communique-t-on et comment créons-nous des liens avec les gens qui nous entourent ? Voici une question qui semble être réservée aux sociologues. Pourtant, on peut tenter d'y répondre à travers les mathématiques ! L'utilité de bien comprendre de tels mécanismes s'étend bien au-delà des sciences humaines.

Qu'est-ce qu'un *réseau social* ? Ces mots sont sans doute familiers pour la plupart des jeunes lecteurs. Cependant, la notion de réseau social est bien plus large que Facebook, Twitter… Observer un réseau social, c'est observer un groupe de personnes et les liens qui les unissent. Ces liens peuvent être de différentes natures : liens d'amitié virtuels sur un site Internet ; relations commerciales ; collaborations politiques, scientifiques, ou artistiques… Les possibilités ne sont limitées que par notre imagination.

Les graphes et le phénomène du « petit monde »

Un réseau peut être représenté comme un *graphe*, c'est-à-dire un ensemble de points (ou *sommets*) et d'*arêtes*. Les sommets symbolisent les personnes impliquées dans le réseau et les arêtes représentent les liens entre elles, liens dont la nature reste à clarifier.

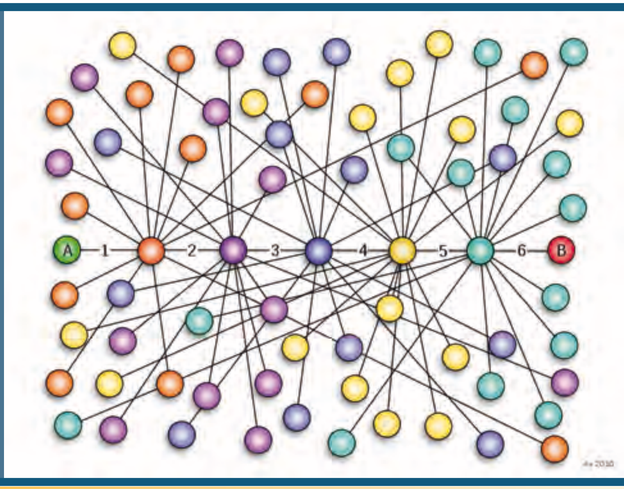


La structure de graphe permet de représenter tout type de réseaux.

Observons un tel réseau social dans le monde réel et tentons d'en dégager certaines caractéristiques importantes. Ensuite, de manière abstraite cette fois, considérons un ensemble de points (représentant les personnes du réseau réel) et lions-les entre eux en suivant certaines règles, que nous aurons préalablement choisies. Est-il alors possible de trouver quelques règles simples qui nous permettraient de recréer un réseau ayant les caractéristiques remarquées dans le réseau réel ?

Prenons l'une des principales caractéristiques que l'on observe dans un réseau social réel : le *phénomène du petit monde,* aussi appelé les *six degrés de séparation*. Choisissons deux personnes A et B au hasard dans le monde. Ces deux personnes ne se connaissent sans doute pas et peuvent être très éloignées géographiquement. En partant de A et en passant d'amis en amis (*« Je suis avec A, puis avec un ami de A, puis avec un ami de cet ami… »*), combien d'amis intermédiaires me suffit-il de rencontrer avant d'être avec B ? La *distance* entre A et B mesure précisément le nombre minimal d'étapes pour aller de A à B de cette manière.

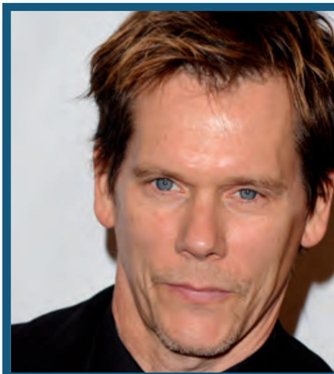
La réponse à notre question est étonnante : dans la très grande majorité des cas, il suffirait de cinq intermédiaires, et donc six étapes en tout ! La première expérience fut menée par le psychologue social américain Stanley Milgram (1933–1984) dans les années 1960, mais l'idée date des années 1930. Ce phénomène surprenant a été observé dans beaucoup de contextes, dont les réseaux de collaborations scientifiques ou artistiques, le réseau électrique californien, le réseau Internet ou encore les réseaux de neurones.



Souvent, dans un réseau, moins de cinq intermédiaires sont nécessaires pour aller d'un sommet à un autre sommet. *Source : Wikipedia*

Partout, des grands connecteurs… même à Hollywood !

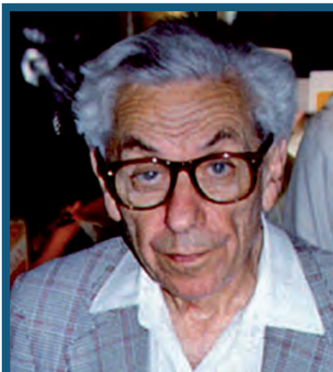
Le Kevin Bacon Game est devenu un exemple assez célèbre du phénomène du petit monde. En quoi consiste ce jeu nommé en référence à l'acteur de *Footloose* (Paramount Pictures, 1984) ? Un chercheur américain a étudié le graphe composé par tous les acteurs affiliés à Hollywood et dans lequel deux acteurs sont connectés s'ils ont joué dans un même film (les liens d'amitié sont donc ici remplacés par les collaborations cinématographiques). Il en a conclu que Kevin Bacon a une place particulièrement centrale dans l'univers hollywoodien, sa distance moyenne à tous les acteurs étant de 3.



L'acteur américain Kevin Bacon (né en 1958).
Source : biosstars.com

Vous pouvez le vérifier sur le site oracleofbacon.org avec tout acteur de votre choix !

Les mathématiciens ont un concept similaire, appelé *nombre d'Erdös*, qui consiste à mesurer leur distance, en termes de collaborations scientifiques, au mathématicien Paul Erdös, connu pour avoir été extrêmement prolifique.



Le mathématicien hongrois Paul Erdös (1913–1996).
Source : Wikipedia

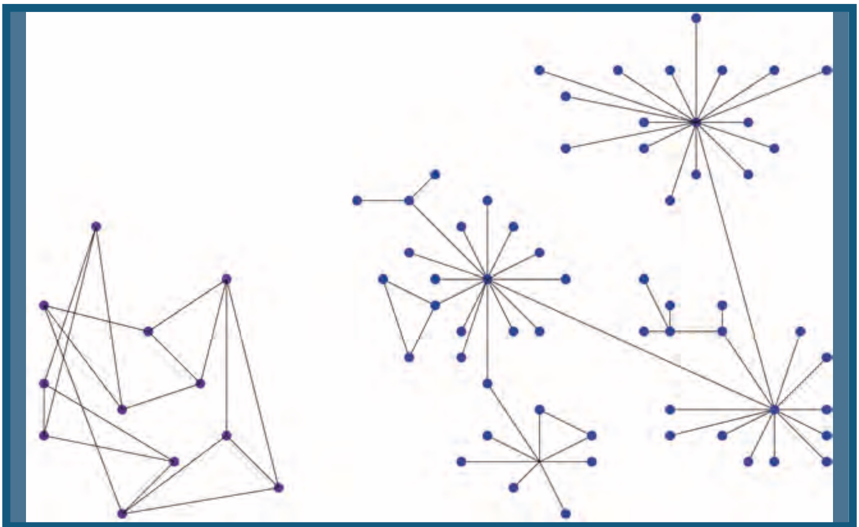
L'une des raisons du phénomène du petit monde pourrait être la présence de *grands connecteurs* dans le réseau, c'est-à-dire de quelques personnes particulièrement bien connectées, à l'image de Kevin Bacon ou de Paul Erdös. En effet, un acteur ayant joué avec Kevin Bacon est en moyenne à distance 4, au plus, d'un autre acteur !

Quand les riches s'enrichissent…

Ainsi émergent deux caractéristiques que l'on aimerait pouvoir recréer à l'aide de quelques règles simples : une distance typique très petite entre deux sommets, et la présence de grands connecteurs.

Tentons une première construction, simple, d'un graphe. Considérons les sommets d'un graphe et lions-les au hasard, de la manière suivante : chaque paire de sommets est connectée avec une certaine probabilité p strictement comprise entre 0 et 1 et déconnectée sinon (avec probabilité $1-p$, donc). Le graphe obtenu de cette manière s'appelle un *graphe d'Erdös–Rényi*. Son étude révèle que la distance typique entre deux sommets est très petite comparée à la taille du graphe, ce qui va dans le sens des six degrés de séparation ! En revanche, les sommets ont tous, à peu près, le même nombre de voisins (ou d'amis), ce qui indique malheureusement l'absence de grands connecteurs…

Essayons donc une seconde approche, dans l'idée de faire apparaître ces grands connecteurs. Commençons avec un petit nombre de sommets connectés de manière arbitraire puis, à chaque étape, ajoutons un nouveau sommet au graphe qui sera lié aléatoirement à un certain nombre d'anciens sommets : plus un ancien sommet a de voisins, plus il aura de chances que le nouveau sommet se connecte à lui. Cette procédure s'appelle l'*attachement préférentiel*, parfois illustrée par l'expression « le riche s'enrichit ». L'analyse mathématique de ce modèle dévoile une petite distance typique, ainsi que la présence de grands connecteurs !



Deux structures de graphes aux propriétés très différentes :

À gauche : Type Erdös–Rényi :
tous les sommets ont trois ou quatre arêtes.

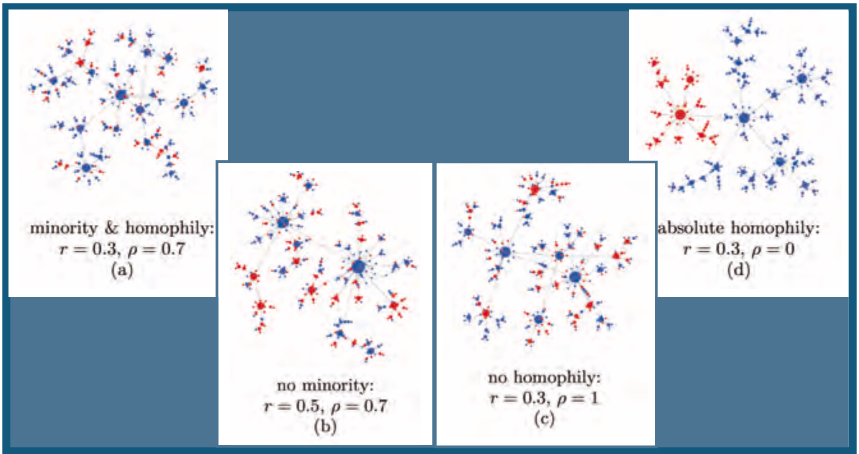
À droite : Type attachement préférentiel :
quatre sommets ont un nombre particulièrement grand d'arêtes.

Nous obtenons donc un exemple de modèle mathématique dont certaines caractéristiques typiques semblent proches de celles d'un réseau social réel. Mais les modèles et constructions de graphes disponibles aujourd'hui sont encore loin d'être entièrement satisfaisants : beaucoup de recherche est encore nécessaire pour la compréhension de tels phénomènes.

Tout ce travail est-il utile ? Il paraît ici opportun de citer Henri Poincaré dans *Science et Méthode* (Flammarion, 1908) : *« La mathématique est l'art de donner le même nom à des choses différentes. »* En effet, comprendre la structure d'un réseau social (et notamment comprendre pourquoi Kevin Bacon est central à Hollywood) permettrait de comprendre la structure d'une multitude d'autres réseaux partageant les mêmes caractéristiques globales. Ainsi, la recherche dans ce domaine pourrait avoir des applications dans les sciences politiques (pour l'étude de l'évolution des coopérations au sein d'une grande organisation), dans l'analyse de la propagation de maladies ou de virus informatiques, dans la mise au point de réseaux électriques, de communication ou de distribution, ainsi que dans la compréhension de la création du langage ou de la formation de réseaux de neurones dans notre cerveau. Rien que ça !

Cette dernière application semble particulièrement actuelle : elle aurait des répercussions en intelligence artificielle. En effet, les algorithmes d'apprentissage profond (*deep learning*) s'inspirent du fonctionnement du système nerveux humain, que l'on peut schématiser comme un graphe dont les sommets reçoivent des messages et établissent des connexions en conséquence, selon certaines règles. Ainsi, mieux comprendre la structure des réseaux de neurones (qui possèdent une petite distance typique et des grands connecteurs) permettrait peut-être d'améliorer nos compétences en intelligence artificielle. En mars 2016, AlphaGo (voir https://deepmind.com/alpha-go.html), un algorithme d'apprentissage profond de Google DeepMind, a gagné assez largement une partie de jeu de go contre l'un des meilleurs joueurs du monde.

Dans un domaine plus proche de la sociologie, en 2015, un groupe de chercheurs français, israéliens et suisses ont été offusqués par la grande proportion d'hommes en recherche informatique, d'autant plus prononcée parmi les chercheurs les plus influents. Ils ont alors proposé un modèle basé sur des règles simples, telles que l'attachement préférentiel et l'homophilie (on se sent plus proche des gens qui nous ressemblent), pour lequel on voit apparaître un *plafond de verre* : dans un réseau où les hommes sont plus nombreux que les femmes, il sera disproportionnellement plus difficile pour une femme d'atteindre un haut niveau d'influence !



Sur ces images, le nombre r est la proportion de femmes et ρ est un coefficient d'homophilie (qui vaut 0 pour l'intolérance au sexe opposé et qui vaut 1 pour une indifférence totale).
Plus un sommet est « gros », plus il est connecté.

Source : Homophily And The Glass Ceiling Effect In Social Networks. *Chen Avin, Barbara Keller, Zvi Lotker, Claire Mathieu, David Peleg et Yvonne-Anne Pignolet,* Proceedings Of The 2015 Conference On Innovations In Theoretical Computer*, 2015.*

Ce modèle, bien que simpliste, permettrait déjà d'envisager de possibles solutions au plafond de verre rencontré par certaines minorités sociales…

D.K.

Pour en savoir (un peu) plus :

Small Worlds: The Dynamics Of Networks Between Order And Randomness. Duncan Watts, Princeton University Press, 2003.

Emergence Of Scaling In Random Networks.
Albert-Laszlo Barabasi et Réka Albert, *Science* 286, 1999.

The Large-Scale Organization Of Metabolic Networks.
Hawoong Jeong, Balint Tombor, Réka Albert, Zoltan Oltvai et Albert-Laszlo Barabasi, *Nature* 405, 2000.

Classes Of Behavior Of Small-World Networks.
Luis Nunes Amaral, Antonio Scala, Marc Barthélémy et Eugene Stanley, *Proceedings Of The National Academy Of Sciences* 97, 2000.

Homophily And The Glass Ceiling Effect In Social Networks.
Chen Avin, Barbara Keller, Zvi Lotker, Claire Mathieu, David Peleg et Yvonne-Anne Pignolet, *Proceedings Of The 2015 Conference On Innovations In Theoretical Computer*, 2015.

Jeux vidéo : les univers créés par les algorithmes

FibreTigre

Game designer

En quoi consiste le travail d'un concepteur de jeux vidéo (ou *game designer*) ? Il a à charge d'écrire sur du papier des « règles du jeu », en espérant qu'elles seront amusantes. C'est à partir de ce document que naîtra, sous les efforts des artisans du jeu vidéo (codeurs, graphistes, ergonomes, musiciens…), le produit final.

C'est un travail très proche des mathématiques appliquées à la physique, dans le sens où le concepteur de jeux vidéo imagine un univers : il doit développer des fonctions, des conditions, élaborer des algorithmes pour atteindre son objectif, qui est principalement d'amuser, parfois de faire passer un message ou de raconter une histoire, et, dans quelques cas déontologiquement discutables, d'encourager le joueur à la dépense…

Un principe fondamental de création de modèles mathématiques pour le jeu vidéo est issu d'une maxime : *« Pensez bien votre univers, les principes de* gameplay *[c'est-à-dire les règles du jeu] émergeront d'eux-mêmes. »* La base de travail est donc de « constater » la nature d'un univers paradoxalement imaginaire.

Une hypothèse chère au monde mathématique…

Cette pratique reste très jeune (moins de 30 ans) et même une grammaire basique et unifiée pour expliquer les méthodes de base fait encore défaut. Ces méthodes restent empiriques, intuitives. Il n'est pas rare qu'en détaillant un projet de jeu on commence par expliquer des concepts nommés arbitrairement, car on avance dans l'inconnu. Il n'existe pas encore de classiques universels de référence, sauf ceux que chacun veut bien reconnaître comme tels.

Statistiques et polynômes au service du jeu

À haut niveau, les connaissances mathématiques requises pour exercer le métier de *game designer* concernent surtout les statistiques ; c'est le cas des *cash machines* de très grande envergure comme Candy Crush Saga (King, 2012) ou Clash Of Clans (Supercell, 2012, ce dernier dégageant plus d'un million d'euros de bénéfices par jour actuellement).

Pour les artisans du commun, plus modestes, le niveau de mathématiques requis ne dépasse pas la seconde : il faut connaître le théorème de Pythagore, qui reste une façon élégante de mesurer la distance entre deux points dans un jeu en deux dimensions, ou bien savoir manipuler des polynômes du second degré, car ils permettent, par exemple, de créer des courbes de progression (ou de difficulté) faciles à appréhender intuitivement pour un joueur.

L'un des aspects mathématiques du métier concerne la *génération procédurale*. Imaginons un monde médiéval fantastique avec des montagnes, des châteaux, des dragons… Il existe deux méthodes principales pour créer un tel univers de jeu. On peut d'abord le fabriquer « à la main », avec des professionnels qui placent chaque élément de jeu en le pensant, en suivant des guides… ou leur intuition. Mais on peut aussi avoir recours à la génération procédurale d'univers : on substitue des algorithmes au travail de placement d'objets ou de création de lieux. On obtient ainsi des mondes générés *à la volée* (créés à la demande en début de partie) ou *persistants* (générés informatiquement mais sélectionnés ensuite).

On associe à certains mondes créés à la volée des *graines* (*seeds* en anglais), à savoir de courtes chaînes de chiffres (comme 1–2–3–5–6–1–5–6–1–5), qui sont le petit battement d'ailes de papillon qui vont créer un monde unique sous le fracas primordial des algorithmes de génération.
Avec la « bonne » graine, on peut retrouver le monde correspondant : l'algorithme est une fonction qui, avec la graine X, produit l'univers Y.

Le procédural au service du débogage

Une des phases capitales de la création de jeu est le débogage. Durant plusieurs mois, des joueurs professionnels explorent et poussent le jeu dans ses limites pour supprimer les *bugs*. La faiblesse de densité de *bugs* est essentielle pour que le jeu soit bien noté et se vende bien. Dans un univers fabriqué à la main, si on s'aperçoit d'un *bug* (une planète de couleur violette, ce qui n'est pas prévu), on le corrige. Et si jamais on trouve une autre planète violette, on la corrige à nouveau.
Dans un univers généré de manière procédurale, corriger un *bug* signalé agit sur toutes les occurrences possibles ultérieures, qu'elles existent ou non.
C'est un gain de temps considérable !

Des univers, passionnants ou ennuyeux, ludiques ou artificiels

Face à ces options, une question reste prépondérante : quelle est la solution la plus ludique ? Un terme de robotique, *vallée dérangeante*, exprime que plus un robot s'approche de l'apparence humaine, plus ses infimes différences nous rappellent qu'il reste une création artificielle. Existe-t-il une vallée dérangeante de la création algorithmique d'univers, qui ferait que l'on peut pointer du doigt un monde généré algorithmiquement (par opposition à un monde généré à la main) ? Les expériences de jeu sont-elles identiques ?

Historiquement, les deux écoles ont avancé de front : d'un côté, les jeux de rôles, d'action ou de plateau proposent une aventure unique mais parfaitement circonscrite. De l'autre, des jeux générés de manière procédurale proposent une expérience émotionnelle parfois plus faible ; mais comme chaque univers est différent à chaque partie, ils offrent une « rejouabilité » sans limite.

Daggerfall (Bethesda, 1996), dont la suite est Morrowind (Bethesda, 2002), est un jeu procédural : il propose un monde de montagnes et de plaines, de cavernes labyrinthiques, de villages et de villes, grand comme environ l'Allemagne, avec autant d'habitants. Tout est généré avec des fonctions mathématiques. Morrowind, au contraire, est un jeu où chaque caillou, objet, personnage a été placé et défini à la main. Sa surface est de 24 km², soit un cinquième de Paris.

Daggerfall.
Cette carte montre l'une des nombreuses provinces de l'aire de jeu. Chaque point représente un village, une ville ou un donjon.

© *Bethesda*

Morrowind, qui propose une longueur de jeu presque aussi importante que son prédécesseur Daggerfall, a beaucoup plus marqué les esprits. Pourquoi ? Le monde généré de manière procédurale pour Daggerfall avait, dans son foisonnement, une forme de vacuité : les motifs étaient perceptibles, le jeu était immense comme une frise infinie qui jamais ne déborderait sur de la folie ou sur une forme d'humanité. C'est un exemple rare où une tâche accomplie par un robot est améliorée par un humain !

Morrowind.
Chaque emplacement de terrain (y compris les profondeurs sous-marines proches des côtes) a été modélisé à la main avec des éditeurs et des *level designers*.

© Bethesda

En dépit de cette démonstration, qui a orienté beaucoup de jeux dans le début des années 2000, la génération procédurale a fait des progrès et revient en force.

Minecraft (Mojang, 2011), dont l'univers est basé sur du procédural, a été racheté par Microsoft 2,3 milliards d'euros et reste un jeu dominant sur la scène mondiale.

Minecraft.
Le jeu est généré de manière procédurale à la volée au fil des explorations du joueur.

© *Mojang*

No Man's Sky (Hello Games), présenté durant l'été 2015 à l'E3, un salon majeur du monde de l'informatique et de l'électronique, propose un univers à l'échelle d'une galaxie où, selon le communiqué de l'éditeur, *« chaque atome »* est généré de manière procédurale.

No Man’s Sky.
Cette planète, les couleurs, la densité de faune et de flore, leur nature et tous types d’éléments physiques sont générés de manière procédurale… tout comme les milliards d’autres planètes proposées. © *Hello Games*

Star Citizen (Cloud Imperium Games, 2016), jeu financé de manière participative à hauteur de 100 millions d’euros (!), est généré de façon procédurale pour offrir un univers également à l’échelle interstellaire.

Star Citizen.
Les systèmes solaires sont générés de façon procédurale.
Pour le plus grand succès de financement participatif de l’histoire, la génération procédurale est au service de la création d’univers qualitatifs plus que quantitatifs. © *Cloud Imperium Games*

Excepté pour Minecraft, dont l'univers en lui-même est moins interressant que les interactions de ses mécaniques internes, la question initiale demeure, et elle vaut aujourd'hui cent millions d'euros : les joueurs trouveront-ils de l'intérêt à visiter une planète générée par une fonction mathématique, aussi sophistiquée soit-elle ?

Quand les algorithmes touchent à la sensibilité

Cette question entre en résonance avec les performances récentes d'AlphaGo, intelligence artificielle (développée par Google DeepMind), contre Lee Sedol, champion de go. Nous sommes désormais au-delà du constat simpliste de Deep Blue contre Garry Kasparov affirmant que la machine avait une faculté d'anticipation supérieure à celle de l'homme. Les performances informatiques dépassent celles de l'humain à un niveau tel que l'on ne peut pas dissocier, au moment de l'écriture de cet article, le travail d'algorithmes qui tâtonnent dans le noir du champ des possibles de la notion un peu métaphysique et rattachée jusqu'alors à l'humain de « créativité ».

Certes, l'œuvre la plus poignante (ou la plus amusante dans notre cas) peut émerger de l'infini des nombres, grâce à une astucieuse procédure de codage. Mais il faut un humain pour sélectionner la beauté dans cet infini. On peut trouver la Joconde dans une séquence soigneusement sélectionnée de décimales d'un nombre univers. Mais l'œuvre humaine doit préexister pour que cette sélection soit opérée. Quand Léonard de Vinci peint la Joconde, il est dans un travail de sélection : à tout moment il sélectionne pinceau, couleur, modèle, trait et de nombreux paramètres selon l'algorithme de sa sensibilité. On découvre dans son travail la familière altérité d'une production humaine plutôt que la reproduction d'une quelconque réalité.

Sommes-nous capables de concevoir des algorithmes qui génèrent des mondes circonscriptibles par notre intime et qui nous fassent voir la familière altérité de la sensibilité de leurs auteurs ?

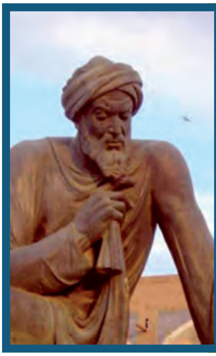
Ou bien les mathématiques resteront-elles ce terrain radicalement étranger, teinté de ses concepts intimidants (le zéro, l'infini…), qui ne pourra jamais se substituer à la création ? Cette question pèse aujourd'hui de façon remarquable dans les vingt milliards d'euros de chiffre d'affaire annuel de la première industrie de loisirs au monde !

F.T.

Informatique et mathématiques, un partenariat gagnant

Gérard Berry
Professeur au Collège de France,
membre de l'Académie des sciences
et de l'Académie des technologies,
médaille d'or 2014 du CNRS

Les mathématiques et l'informatique partagent deux ancêtres communs : le nombre et le calcul, vite devenus indispensables pour le comptage des troupeaux, la mesure du temps, celle des distances et des surfaces, l'architecture, l'observation des astres, le commerce, l'impôt… Les procédés de calcul sur les nombres datent de plusieurs milliers d'années pour la multiplication et la solution des équations simples. Devenus petit à petit systématiques et se faisant aider par des outils variés (abaques, bouliers…), ils permettaient de ne plus faire appel à l'intelligence pour réaliser les opérations : c'étaient déjà des *algorithmes* avant la lettre.



Le mot « algorithme » désigne un procédé de calcul systématique. Il vient du nom du savant persan al-Khuwarizmi (vers 783, 850), qui a importé les connaissances mathématiques des Indiens (dont les chiffres sont du coup devenus nos chiffres arabes).

Il a écrit le traité *Al Jabr* sur la solution des équations ; de là vient le mot « *algèbre* ».

Photo : Maurice Nivat

De la démonstration à l'ordinateur : une évolution « logique »

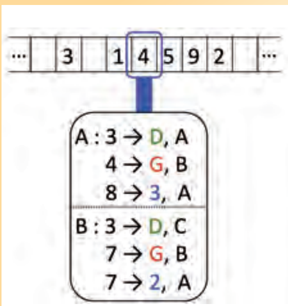
Les mathématiques telles que nous les connaissons et les pratiquons aujourd'hui sont arrivées plus tard. On les retrouve déjà chez les pythagoriciens pour justifier les algorithmes empiriques, par exemple pour montrer que $a^2 + b^2 = c^2$ est vrai quel que soit le triangle rectangle de côtés quelconques *a*, *b* et *c*, et pas seulement pour les nombreux exemples connus. Il a fallu introduire la nouvelle notion de démonstration, clef de la pensée mathématique. Cette notion a ensuite été elle-même mathématisée par les logiciens, qui l'ont transformée en calcul : si inventer une démonstration est intellectuellement difficile, la vérifier doit être une opération algorithmique conceptuellement réalisables par une machine. Cette approche calculatoire de la logique a culminé dans le fameux théorème de Gödel, qui montre les limites du procédé en exprimant que toutes les vérités mathématiques ne sont pas démontrables par calcul logique.

Alan Mathison Turing (1912–1954) a grandement participé à la victoire des Alliés en cassant le code secret allemand Enigma pendant la Seconde Guerre mondiale.

L'informatique n'est vraiment née qu'en 1936, quand Alan Turing a défini sa fameuse machine (voir ci-dessous). Appliquant les idées de Gödel à cette construction ultra-minimale, Turing a démontré des résultats extraordinaires : d'abord, sa machine permet de faire tous les calculs dont on a besoin en pratique ; ensuite, on peut construire une machine universelle programmable, où un programme enregistré à côté des données définit les calculs à faire, la machine restant constante ; enfin, la machine universelle n'est pas omnipotente, car elle ne peut pas évaluer par le calcul si un programme précis va s'arrêter sur une donnée précise.

La machine de Turing

La machine conceptuelle mise au point par Turing comporte une tête de lecture, qui permet de lire et d'écrire des lettres appartenant à un alphabet fini sur une bande infinie découpée en cases, et de décaler la bande d'une case à gauche ou à droite. Elle détermine ses actions seulement à partir de la lettre lue sous la tête et d'un état fini géré par un contrôleur qui dit simplement quelle action faire :
« écrire une autre lettre »
ou
« bouger la bande d'une case ».



Paradoxalement, la machine de Turing possède les mêmes capacités (mais bien sûr pas la même efficacité) que tous les ordinateurs connus, et, selon la thèse de Church, que tous les procédés de calcul à venir.

L'ENIAC
a été le premier ordinateur.
Il pesait plusieurs tonnes
et avait un temps de fonctionnement
sûr de seulement quelques minutes.
Son programme était câblé,
pas encore enregistré en mémoire.
© *US Army*
(Encadré) Le microprocesseur
Haswell d'Intel comporte plus
d'un milliard de transistors.
Crédit : AnandTech

Mais il faudra attendre la fin des années 1940 pour que l'informatique passe à l'acte avec les premiers ordinateurs électroniques construits par des ingénieurs et des mathématiciens (dont John von Neumann), puis les inventions des transistors et des circuits intégrés modernes pour passer à l'échelle actuelle : si le calculateur ENIAC avait seize mille tubes à vide, un processeur moderne possède plusieurs milliards de transistors.

Le développement d'une pensée algorithmique propre

Les ordinateurs ont d'abord servi d'outils de calcul scientifique pour mathématiciens. Mais ils ont vite pénétré d'autres secteurs, remplaçant les machines mécanographiques dans la gestion des entreprises. Dès la fin des années 1960, avec le progrès exponentiel des ordinateurs, l'explosion des domaines d'application et l'accroissement constant de sa communauté, l'informatique a commencé à se détacher des mathématiques classiques pour développer sa propre *pensée algorithmique* : invention des structures de données et développement des algorithmes associés, conception de langages permettant une programmation plus riche et non limitée aux nombres, *etc*. Cette évolution a demandé l'élaboration de nouvelles approches mathématiques, comme en cryptographie [voir notre article en page 21], pour résoudre des questions… typiquement informatiques.

La *complexité des algorithmes* cherche à estimer le coût en temps, mémoire et énergie des calculs pour un problème donné. Elle a conduit à une classification fine de la difficulté intrinsèque des problèmes algorithmiques, et aussi à la fameuse conjecture « la classe P est-elle égale à la classe NP ? » (énoncée par Stephen Cook en 1972), qui fait partie de la liste des problèmes de maths mis à prix à un million de dollars.

P *versus* NP et SAT

La classe P est celle des problèmes solubles en temps polynomial, donc « rapidement » (en pratique). La classe NP est celle des problèmes pour lesquels il est « facile » de vérifier une solution (proposée par exemple par un oracle), mais pas de proposer la bonne solution. Il existe des centaines de problèmes pratiques dans ce cas ! Par exemple, le problème de la *k*-coloration de graphe cherche s'il est possible de colorer un graphe non orienté avec *k* couleurs de manière à ce que deux sommets reliés par un arc ne soient pas de la même couleur. Il est très facile de vérifier si une coloration a bien la propriété cherchée, mais très difficile d'en trouver une bonne. Le problème NP le plus typique est SAT, le vieux problème de la satisfaction d'une formule booléenne (sur vrai / faux, ou 0 / 1) écrite avec des variables et les opérateurs et, ou, non. Une telle combinaison de variables booléennes étant donnée, savoir s'il existe des valeurs (0 ou 1) de ces variables qui rendent l'expression globale égale à 1 est un problème NP-complet, ce qui veut dire que tout autre problème NP peut être réduit à lui. La conjecture est que SAT ne peut pas être résolue en temps polynomial dans le pire cas. Mais cela ne veut pas dire qu'il ne peut pas être résolu dans les cas pratiques ! Avant 2000, on ne pensait pas pouvoir résoudre SAT au-delà de quelques centaines de variables. Avec de nouveaux algorithmes étonnants, on résout maintenant des problèmes à deux millions de variables, y compris en cadre industriel !

La *théorie de l'information* cherche quant à elle à comprendre comment transmettre efficacement des messages sur des lignes bruitées. Créée par Claude Shannon en 1947, elle a permis des gains de facteur 10 000 entre la transmission par modems des années 1980 et l'ADSL actuel sur les mêmes fils téléphoniques, et des gains tout aussi considérables pour la 5G en téléphonie sans fil. Information et probabilités sont liées. La transmission en télécoms, le routage des paquets sur Internet et la gestion des réseaux pair-à-pair font appel à des algorithmes probabilistes sophistiqués. En fait, le hasard et les probabilités jouent un rôle de plus en plus important en algorithmique.

Géométrie, théorie des nombres, analyse : toutes les maths mobilisées !

La *géométrie algorithmique* cherche à développer des algorithmes très efficaces pour les grands calculs géométriques nécessaires pour la CAO (conception assistée par ordinateur) des avions, voitures, ou tout autre objet, pour la synthèse d'images, pour la modélisation des molécules et de leurs interactions…

Le *traitement d'images et de sons* a pris une importance considérable, avec des résultats insoupçonnables auparavant. Par exemple, au lieu de regarder séparément des images par radio, scanner et IRM, le médecin va regarder désormais des *fusions algorithmiques* de ces images sous formes 3D ou 4D (films), qui lui fourniront des possibilités de diagnostic bien plus fines.

La *modélisation* et la *simulation numériques* de phénomènes complexes se développent dans de nombreuses branches : médecine, avionique, motorisation, astrophysique, agronomie, biologie. On commence même à faire réaliser des calculs destinés à agir sur des réseaux métaboliques par une *machine biochimique intra-cellulaire*, qui remplace les transistors par des brins d'ADN…

Le vieux rêve de Turing commence à se réaliser

N'oublions pas le rôle majeur de la logique mathématique dans l'informatique moderne ! Les premiers constructeurs d'ordinateurs ont constaté, à leur grande surprise, qu'il était très difficile d'écrire des programmes justes. Ils ont dû inventer le *débogage,* activité apparemment fort éloignée des mathématiques. Ici encore, Turing a été le grand précurseur, dans une note de 1949 exprimant que la meilleure façon de savoir si un programme fait ce que l'on veut serait de le vérifier mathématiquement. Il a proposé des solutions à base d'*assertions logiques* placées sur les lignes de programmes et d'*invariants mathématiques* permettant de prouver ces assertions. Après une lente évolution (qui a cependant donné lieu au prix Turing du Français Joseph Sifakis en 2006), la vérification de programmes est devenue mature. Elle se décompose en *vérification automatique* pour des sous-domaines dédiés (circuits électroniques…) et *vérification par assistants de preuves mathématiques.* Les deux ont fait des progrès extraordinaires ces quinze dernières années, avec des succès pratiques comme la vérification systématique de la conception des circuits électroniques, celle du compilateur C CompCert de Xavier Leroy ou celles de protocoles de sécurité pour la carte bleue. Le vieux souhait de Turing se réalise !

Si les mathématiciens utilisent depuis longtemps les systèmes informatiques de calcul formel pour gérer leurs gros calculs, l'utilisation d'assistants informatiques de preuve pour vérifier les théorèmes est plus récente. Les plus puissants sont fondés sur le λ-calcul (« lambda-calcul »), introduit en 1936 par Alonzo Church en même temps que la machine de Turing, et devenu la base des langages de programmation et des logiques modernes. Le système français Coq, leader actuel du domaine, résulte de trente années de recherches sur le λ-calcul et la théorie des types par Jean-Yves Girard, Gérard Huet, Thierry Coquand et bien d'autres.

En utilisant Coq, Georges Gonthier et Benjamin Werner ont démontré en machine le fameux théorème des quatre couleurs (voir en encadré). Puis Gonthier et son équipe ont réalisé un exploit exceptionnel, véritablement historique, en démontrant en Coq le célèbre théorème de Feit–Thompson (datant de 1963) sur la classification des groupes d'ordre impair. La preuve publiée de ce difficile théorème d'algèbre occupe pas moins de deux cent cinquante pages de maths lourdes, que l'on n'a jamais vraiment réussi à simplifier de manière significative. Enfin, en 2014, Thomas Hales a fourni la première démonstration de la fameuse conjecture de Kepler sur le rangement de sphères toutes identiques, qui dit simplement que les épiciers ont raison de mettre leurs oranges plan par plan sous forme d'hexagones. La preuve de Hales a demandé d'une part une utilisation massive du calcul formel sur les polynômes, et d'autre part

Le théorème des quatre couleurs

Le *théorème des quatre couleurs* exprime qu'il suffit de quatre couleurs pour colorier toute carte de géographie de façon à ce que deux États avec une frontière commune n'aient pas la même couleur. La question de savoir si toute carte peut être ainsi coloriée a été posée en 1852. En 1976, Appel et Haken rédigent un article montrant qu'il suffisait de vérifier 1 478 configurations de cartes, ce qu'ils font sur ordinateur. Leur article n'est pas contesté, mais la vérification sur ordinateur n'est pas vue comme une « vraie » preuve. En 2006, à l'aide du système Coq, Gonthier et Werner prouvent réellement le théorème, simplifient la preuve, et montrent que, si les vérifications sur ordinateur de 1976 étaient bien correctes, le texte mathématique original avait des preuves pas toujours complètes et devait être largement repris !

des vérifications assistées de preuves en utilisant le système anglais HOL (Higher Order Logic).

À la suite de ces exploits informatico-mathématiques, le fameux Institute For Advanced Study de Princeton (États-Unis) a dédié une année spéciale à la mécanisation des preuves mathématiques en Coq et poursuit ce travail dirigé notamment par le médaillé Fields Vladimir Voïevodsky. Aux États-Unis toujours, la National Science Foundation a créé un grand projet, The Science Of Deep Specification, sur la vérification

G.B.

Pour en savoir (un peu) plus :

Pourquoi et comment le monde devient numérique. Gérard Berry, Fayard–Collège de France, 2008.

Vidéos et supports des cours de 2008 à 2016 au Collège de France : ***college-de-france.fr/site/gerard-berry***.

Les métamorphoses du calcul. Gilles Dowek, Le Pommier, 2007.

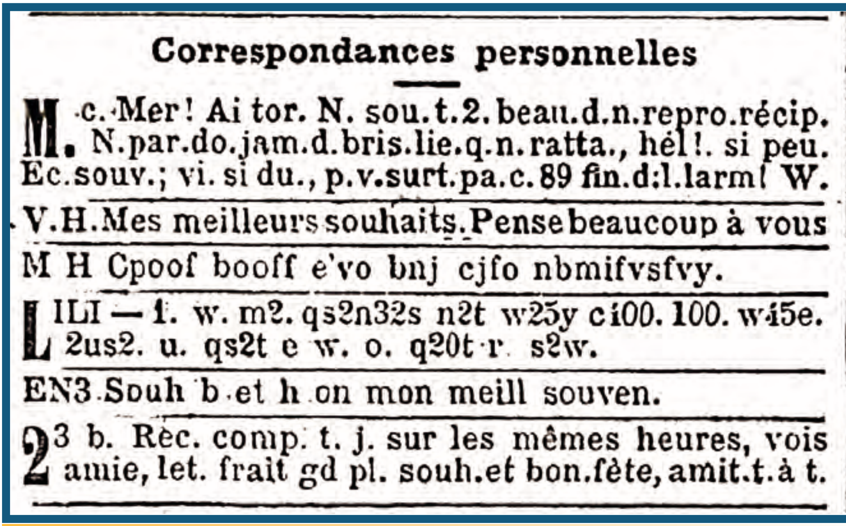
Une histoire illustrée de l'informatique. Pierre-Étienne Mounier-Kuhn et Emmanuel Lazard, EDP Sciences, 2016.

Le secret des correspondances

Hervé Lehning

Agrégé de mathématiques, journaliste et écrivain scientifique

À la fin du XIX^{e} siècle, les journaux étaient utilisés pour la communication entre particuliers. Ainsi, voici un extrait de la rubrique « correspondances personnelles » du *Figaro* du 1er janvier 1890 :



Correspondances personnelles

M. c. Mer! Ai tor. N. sou. t. 2. beau. d. n. repro. récip. N. par. do. jam. d. bris. lie. q. n. ratta., hél!. si peu. Ec. souv.; vi. si du., p. v. surt. pa. c. 89 fin. d:l. larm! W.

V. H. Mes meilleurs souhaits. Pense beaucoup à vous

M H Cpoof booff e'vo bnj cjfo nbmifvsfvy.

LILI — 1. w. m2. qs2n32s n2t w25y c100. 100. w15e. 2us2. u. qs2t e w. o. q20t r. s2w.

EN3 Souh b et h on mon meill souven.

23 b. Rec. comp. t. j. sur les mêmes heures, vois amie, let. frait gd pl. souh. et bon. fête, amit. t. à t.

La rubrique « correspondances personnelles » du *Figaro* du 1er janvier 1890.

Deux messages sont chiffrés dans cet extrait. Le premier, rédigé par M H, est limpide ; il correspond à un décalage d'une lettre et signifie : *« Bonne année d'un ami bien malheureux. »* Le second, proposé par LILI, semble correspondre à une substitution alphabétique. En collectant une cinquantaine de messages, dans différentes éditions du *Figaro*, portant le même indicatif LILI, il est possible de reconstituer la substitution.

Pour commencer, le symbole 2 étant le plus fréquent, il représente sans doute la lettre e, la lettre la plus fréquente en français. On cherche alors une faille pour casser LILI. Nous la trouvons dans le message du 12 janvier :

« **LILI. Vot pens ne me quitte pas est tout mon bonh. voud. vs v. 32. u. 13. n2.** »

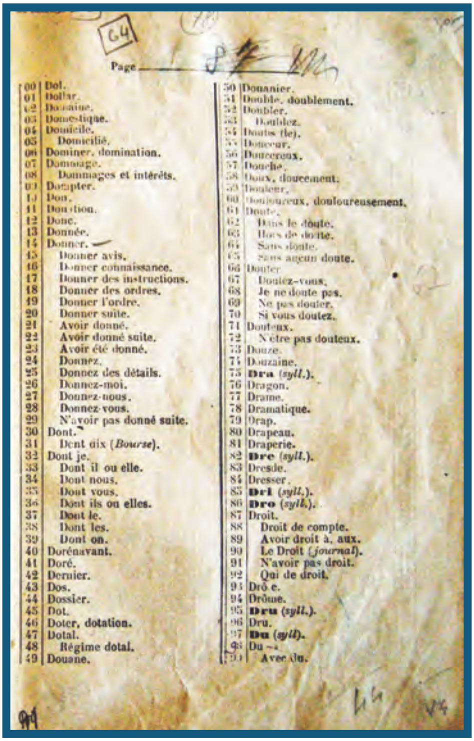
L'unique partie chiffrée, *« 32. u. 13. n2 »*, est un dernier mot que l'expéditeur a voulu cacher. La position des 2 fait penser à *« je t'aime »*, les lettres i et j étant confondues. Le chiffre s'écroule alors progressivement.

Aussi étrange que cela puisse paraître, ces décryptements ont eu un rôle dans l'histoire de la cryptologie. En effet, le commandant français Étienne Bazeries (1846–1931) s'amusait à les lire et, au mess des officiers de sa garnison, régalait ses collègues des histoires scabreuses qu'il lisait sans peine… jusqu'au jour où il annonça qu'il pouvait également lire les messages chiffrés de l'armée. Son général prit cette remarque au sérieux et lui demanda de décrypter quelques dépêches du ministère, ce que Bazeries fit aisément. C'est ainsi qu'il devint l'un des grands cryptologues de l'armée française, puis du ministère des Affaires étrangères de l'époque.

Des maths… plus sophistiquées… pour plus de sécurité !

Pour chiffrer, les entreprises de l'époque utilisaient plutôt des dictionnaires chiffrés ce qui, outre l'assurance d'un certain secret des correspondances, économisait le coût des messages.

L'expression « dommages et intérêts » pouvait ainsi se coder en 9108, soit quatre lettres au



64

Page 8 7 44

00 Dol.
01 Dollar.
02 Domaine.
03 Domestique.
04 Domicile.
05 Domicilié.
06 Dominer, domination.
07 Dommage.
08 Dommages et intérêts.
09 Dompter.
10 Don.
11 Donation.
12 Donc.
13 Donnée.
14 Donner.
15 Donner avis.
16 Donner connaissance.
17 Donner des instructions.
18 Donner des ordres.
19 Donner l'ordre.
20 Donner suite.
21 Avoir donné.
22 Avoir donné suite.
23 Avoir été donné.
24 Donnez.
25 Donnez des détails.
26 Donnez-moi.
27 Donnez-nous.
28 Donnez-vous.
29 N'avoir pas donné suite.
30 Dont.
31 Dont dix (*Bourse*).
32 Dont je.
33 Dont il ou elle.
34 Dont nous.
35 Dont vous.
36 Dont ils ou elles.
37 Dont le.
38 Dont les.
39 Dont on.
40 Dorénavant.
41 Doré.
42 Dernier.
43 Dos.
44 Dossier.
45 Dot.
46 Doter, dotation.
47 Dotal.
48 Régime dotal.
49 Douane.
50 Douanier.
51 Double, doublement.
52 Doubler.
53 Doublez.
54 Doubs (le).
55 Douceur.
56 Doucereux.
57 Douche.
58 Doux, doucement.
59 Douleur.
60 Douloureux, douloureusement.
61 Doute.
62 Dans le doute.
63 Hors de doute.
64 Sans doute.
65 Sans aucun doute.
66 Douter
67 Doutez-vous.
68 Je ne doute pas.
69 Ne pas douter.
70 Si vous doutez.
71 Douteux.
72 N'être pas douteux.
73 Douze.
74 Douzaine.
75 **Dra** (*syll.*).
76 Dragon.
77 Drame.
78 Dramatique.
79 Drap.
80 Drapeau.
81 Draperie.
82 **Dre** (*syll.*).
83 Dresde.
84 Dresser.
85 **Dri** (*syll.*).
86 **Dro** (*syll.*).
87 Droit.
88 Droit de compte.
89 Avoir droit à, aux.
90 Le Droit (*journal*).
91 N'avoir pas droit.
92 Qui de droit.
93 Drôle.
94 Drôme.
95 **Dru** (*syll.*).
96 Dru.
97 **Du** (*syll*).
98 Du —
99 Avec du.

44 V4 91

Une page du *Code télégraphique chiffré* de F.-J. Sittler (Lefebre, 1879). Le numéro de la page est écrit au crayon et plusieurs fois modifié car il fait partie du chiffre. Si le numéro de la page est ici 91, « dragon » se chiffre 9176 (ou 9716… selon la convention passée entre les correspondants).

© Hervé Lehning

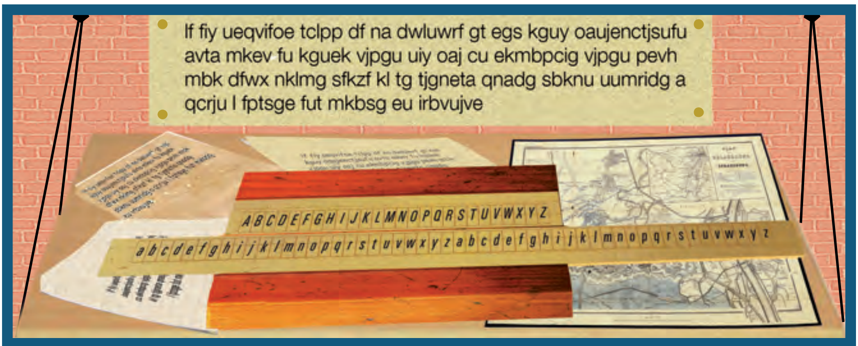
lieu de dix-huit, quand on utilisait le *best-seller* des dictionnaires chiffrés de l'époque, celui de Sittler, qui eut cours de 1890 à 1920 environ.

Même si la numérotation personnalisée des pages compliquait le décryptement, celui-ci restait relativement élémentaire. Tous les messages chiffrés arrivant à la poste étaient envoyés au ministère des Affaires étrangères, pour lequel Étienne Bazeries travaillait. Ils étaient ainsi décryptés.

L'autre grand nom de la cryptographie de la fin du XIX^{e} siècle fut le Néerlandais Auguste Kerckhoffs (1835–1903). Il fut le premier à énoncer le principe de base de la cryptographie moderne : un système de chiffrement ne doit pas reposer sur son secret, mais sur celui d'une clef que l'on change périodiquement. Tous les systèmes que nous venons de voir sont défaillants de ce point de vue.

Dès la Renaissance, le diplomate français Blaise de Vigenère (1523–1596) avait pourtant conçu un système répondant au principe de Kerckhoffs, même s'il fut peu utilisé à l'époque. Son principe est simple mais son application à la main est pénible, il consiste en un décalage des lettres variant selon une clef. Par exemple, la clef « abc » correspond à un décalage de 0 pour la première lettre, de 1 pour la seconde, de 2 pour la troisième, puis on recommence (0 pour la quatrième lettre, 1 pour la cinquième…). Ainsi, « mathématiques » se chiffre en « mbvhfoaukqvgs ».

Bien entendu, les clefs sont en général plus complexes, mais le principe reste le même. Comme l'utilisation manuelle d'un tel chiffre est pénible, il existe des instruments pour le faire, comme la *réglette de Saint-Cyr*, utilisée autrefois dans cette école d'officiers pour l'enseignement de la cryptographie.



La *réglette de Saint-Cyr* permet de réaliser des décalages selon une clef (en italiques ici). Ainsi, la clef *i* transforme A en I, B en J, *etc.*

© *Hervé Lehning*

Le décryptement est d'autant plus difficile que la clef est longue et choisie aléatoirement. L'idéal est qu'elle soit aussi longue que le message et qu'on ne l'utilise qu'une fois, on parle alors de *masque jetable*. C'est le chiffre utilisé dans le fameux téléphone rouge, qui relie Washington et Moscou depuis 1963. Ce chiffre est aussi l'un des chiffres des espions de la guerre froide, d'où les petits carnets étranges ci-dessous.

Carnet de nombres aléatoires avec loupe.

© Archives de la DGSE. Exposition Archives nationales : Le secret de l'État.

Le principe est simple. On chiffre d'abord le message en une suite de nombres par la méthode que l'on veut (par exemple, A = 01, B = 02…), puis on ajoute au message obtenu un nombre aléatoire de même longueur (sans effectuer de retenue). Ainsi, « mathématiques » devient d'abord :

13 01 20 08 05 13 01 20 09 17 21 05 19.

On lui ajoute ensuite les vingt-six premiers chiffres du carnet, le message chiffré est donc :

03691 91070 16694 49382 63998 8.

Clair	13012	00805	13012	00917	21051	9
+ clef	90689	91275	03682	49475	42947	9
= chiffré	03691	91070	16694	49382	63998	8

Si l'on connaît la clef, c'est-à-dire les nombres du carnet, on déchiffre ce message en faisant la différence entre le message chiffré et la clef :

Chiffré	03691	91070	16694	16694	63998	8
– clef	90689	91275	03682	49475	42947	9
= clair	13012	00805	13012	00917	21051	9

Ce système de chiffrement est solide mais repose sur la sécurité de transmission de la clef. Il est très probablement utilisé dans les *stations de nombres*.

Surveillance de masse : prise de conscience et solutions

De nos jours, la confidentialité des communications sur Internet est assurée par des systèmes de ce type, que l'on dit *symétriques* car chiffrement et déchiffrement sont symériques l'un de l'autre. Ces deux opérations s'opèrent à la vitesse de l'addition, c'est-à-dire instantanément. La transmission de la clef est assurée par un chiffrement *asymétrique* (où savoir chiffrer ne suffit pas pour déchiffrer). Le plus célèbre de ces chiffres est le code RSA, inventé par Ronald Rivest, Adi Shamir et Leonard Adleman en 1976 et fondé sur la difficulté de la factorisation des nombres. Son chiffrement repose sur une exponentiation, donc est très lent et gourmand en énergie, ce qui explique que l'on limite son utilisation à la transmission des clefs. Le système PGP (Pretty Good Privacy, en français Assez bonne confidentialité) repose sur ce principe.

A priori, nous pouvons disposer ainsi de chiffrements pratiquement indécryptables. C'est le cas du système de communication des hautes personnalités de l'État créé par Thalès et nommé TEOREM (*sic.*).

Le TEOREM de Thalès,
téléphone sécurisé des personnalités de l'État.
© *Hervé Lehning*

Bien entendu, ceci ne peut que déplaire aux services de renseignements. Une solution serait de demander à chacun de laisser ses clefs chez un tiers de confiance qui ne les donnerait qu'aux autorités autorisées, sur décision de justice.

Edward Snowden
(né en 1983) a révélé
la surveillance de masse
de la NSA.
© *Laura Poitras*

Selon Edward Snowden, la National Security Agency (NSA, l'agence de sécurité américaine) a préféré adopter une autre démarche et a demandé aux fournisseurs de logiciels américains de créer des portes dérobées permettant de contourner leurs algorithmes de chiffrement.

Le pire ennui de cette méthode est que, si une porte dérobée existe pour pénétrer dans une citadelle, tout le monde peut la découvrir et l'utiliser ! La NSA a ainsi créé une faiblesse dans tous les systèmes de chiffrements qu'elle contrôle. Aussi, pourquoi des *hackers*, des entreprises ou d'autres organismes se priveraient-ils de les utiliser ?

À l'inverse, ces portes dérobées pourraient disqualifier les entreprises américaines et créer une opportunité extraordinaire aux sociétés françaises. D'autant plus que le marché de la cryptologie est florissant. Pour satisfaire le désir de l'État de surveiller les escrocs et les terroristes, il suffirait alors de créer un tiers de confiance gardant les clefs de chacun, et d'instaurer une législation prévoyant d'y faire appel dans le cadre de certaines enquêtes.

H.L.

Pour en savoir (un peu) plus :

Les archives du Figaro (1890), disponibles en ligne *via* Gallica, Bibliothèque nationale de France.

L'univers des codes secrets, de l'Antiquité à Internet, Hervé Lehning, Ixelles, 2012.

Les stations de nombres, radio des espions, Philippe Baudouin, Arte radio, 2013.

Le phénomène Big Data et les mathématiques

Jean-Michel Loubes

Institut de mathématiques de Toulouse

Traditionnellement, un statisticien manipule des données certes en grand nombre, mais homogènes : elles présentent un même format, ou partagent une structure commune. Or, l'évolution rapide des systèmes d'information gérant des données de plus en plus volumineuses a causé de profonds changements de paradigme dans le travail du statisticien, devenant maintenant expert en données, ou *data scientist*. Caractériser précisément ce que cache l'appellation « Big Data » est une question sophistiquée : ce terme regroupe diverses réalités selon les différents champs d'application. La compétence d'un expert en données, rare encore sur le marché, se dispute aujourd'hui à prix d'or par les employeurs. Le rapport du McKinsey Global Institute (2011, disponible en ligne) a popularisé la caractérisation du Big Data par trois « V » : volume, variété, vélocité. De nombreux domaines des mathématiques ont un rôle à jouer à tous les niveaux : passons en revue quelles compétences sont mobilisées !

Apprentissage : des algorithmes qui doivent maintenant être revisités

Volume : le traitement de grandes masses de données impose une parallélisation des calculs pour obtenir des résultats en temps raisonnable. Cela est d'autant plus vrai lors du traitement en temps réel d'un flux (ou *streaming*) de données. Les acteurs majeurs que sont Google Inc., Amazon.com, Yahoo!, Twitter Inc., Netflix, Facebook, Microsoft Corporation, SAP, Oracle Corporation… développent dans des centres de données (ou *data centers*) des architectures spécifiques pour stocker, à moindre coût, de grandes masses de données (pages Web, listes de requêtes, informations sur les clients ou les articles à vendre, messages…) sous forme brute, sans unité de format.

L'infrastructure dominante, Hadoop (Apache Software Foundation, 2012), pour distribuer et archiver les données, est issue directement des besoins de Google (archiver des pages, compter des occurrences de mots…) et de son système de fichiers. Ce choix impose de repenser complètement la façon dont sont écrits les algorithmes : la parallélisation devient le maître mot ! Ceux qui ont entendu parler du modèle de programmation MapReduce (Google, 2004), qui domine actuellement, en savent quelque chose…

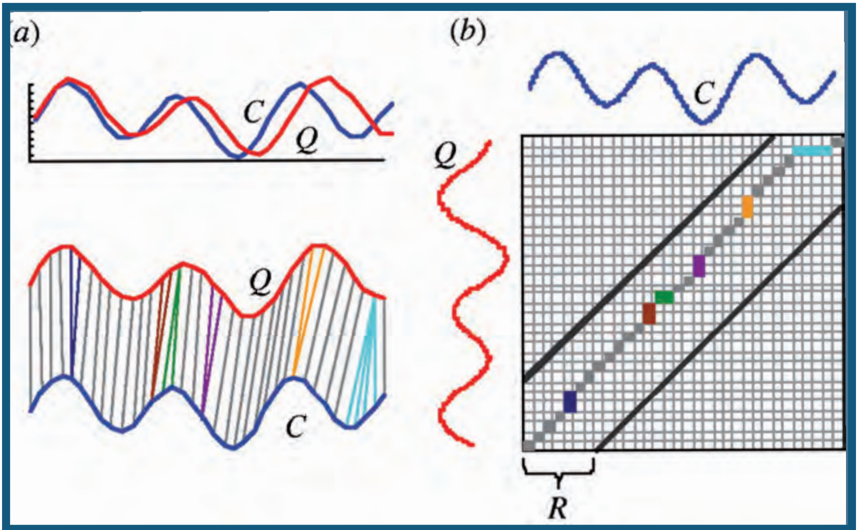
Tous les secteurs industriels sont touchés. Prenons ainsi la problématique, très générale, de la *régression logistique* : cette technique prédictive vise à construire un modèle permettant de prédire ou expliquer les valeurs prises par une variable cible qualitative (comme l'appréciation d'un client sur un produit) à partir d'un ensemble de variables explicatives qualitatives ou quantitatives (grandeur physique, âge, salaire, chiffre d'affaires…). La médecine, le secteur bancaire, l'économétrie, les assurances, le marketing, les sciences sociales, entre autres, y sont confrontés quotidiennement. Le problème de la régression logistique peut désormais être traité avec des outils stochastiques récents (comme les « classifieurs bayésiens naïfs » ou les « algorithmes du gradient stochastique »), ou des méthodes d'apprentissage directement interfacées avec Hadoop. La plus célèbre d'entre elles est le fameux algorithme des *k*-moyennes (ou *k-means algorithm*), utilisé pour réaliser un partitionnement d'une énorme collection de documents. De la même manière, de nombreux algorithmes et techniques classiques en mathématiques doivent être revisités, approfondis, améliorés.

On a besoin d'experts
en statistiques, en algorithmique,
en programmation, en apprentissage et en optimisation !

Classification : vers de nouvelles idées utilisant la géométrie

Variété : le grand nombre de données étudiées entraîne de fait une très grande hétérogénéité. Sous le terme de variété se cache donc de multiples problématiques. Tout d'abord, les données, notamment dans un environnement industriel, peuvent se présenter sous des formes aussi hétérogènes que des textes, des séquences, des fonctions, des courbes, des spectres, des images, des graphes, des tableaux, des sons, des chaînes de caractères… voire des combinaisons de l'ensemble. Au-delà des difficultés techniques inhérentes au traitement d'une telle diversité se pose le problème de l'utilisation de toute cette masse d'informations hétéroclites. Chaque structure de données soulève des questions originales spécifiques, et comparer ces différentes structures entre elles devient également compliqué à la fois d'un point de vue pratique mais également théorique. En effet, calculer ne serait-ce qu'une simple moyenne (ou des distances entre les objets concernés) requiert une analyse particulière pour intégrer leur géométrie.

On a besoin d'experts en géométrie !



Une utilisation en anthropologie pour quantifier les différences et les similarités entre deux crânes de gorilles :
(*a*) deux séries temporelles similaires mais déphasées ;
(*b*) pour aligner les deux séries, une matrice de recalage est construite pour rechercher le chemin de recalage optimal (en carrés colorés).

D'autre part, les données sont observées avec une grande variabilité, qui masque bien souvent la nature du phénomène étudié. Il importe dès lors de révéler la structure contenue dans ces observations afin d'en extraire l'information qu'elles contiennent. Ainsi, pour calculer la distance entre deux courbes, celle usuelle (la « distance dans les espaces L^2 ») est beaucoup trop sensible à de légères déformations ou décalages des courbes. Un recalage préalable (ou *time warping*) est nécessaire.

Une autre approche prometteuse (le *scattering*) en analyse d'images consiste à représenter ces dernières dans une base d'ondelettes possédant des propriétés d'invariance pour certains groupes de transformations (comme Stéphane Mallat et Laurent Sifre l'ont proposé en 2012). Cette représentation permet alors d'identifier ou de regrouper des formes ou des textures d'images particulières.

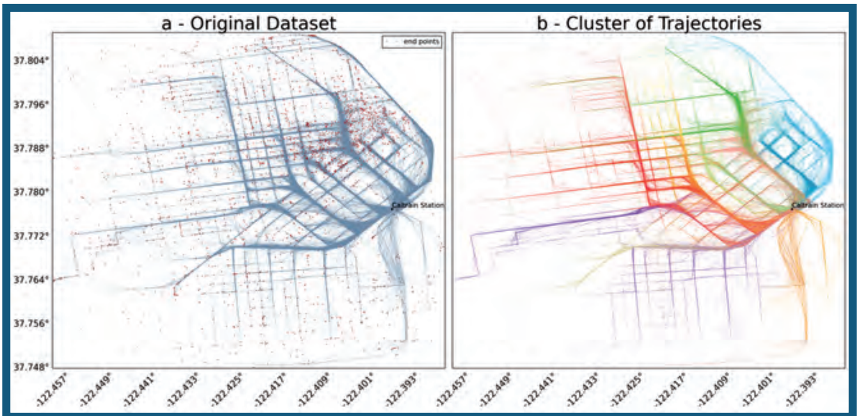
On a besoin d'experts en analyse !

Le dernier exemple concerne initialement l'analyse d'images, conjointement avec des principes de parcimonie. Emmanuel Candès, Justin Romberg et Terence Tao ont montré en 2004 que l'acquisition comprimée (ou *compressed sensing*) permet d'acquérir beaucoup plus rapidement, avec moins de mesures, une très bonne précision des images médicales à partir du moment où celles-ci sont structurées. La reconstruction de l'image fait ensuite appel à la résolution parcimonieuse d'un système linéaire. L'acquisition comprimée a bien d'autres applications, notamment pour l'analyse des très grandes matrices creuses produites en génétique, en génomique, en fouille de textes, en analyse d'incidents dans l'industrie…

On a besoin d'experts en théorie du signal, en analyse d'image et en algèbre linéaire !

Prise de décision : aller au-delà d'une vision séquentielle

Vélocité : la vélocité des données est également motrice de nouvelles recherches sur les algorithmes des méthodes de décision, qui deviennent nécessairement adaptatives ou séquentielles. Comment traiter au mieux



Estimation de flux de populations à partir de données GPS à San Francisco.
© Brendan Guillouet et Jean-Michel Loubes

le flux de données au fil de l'eau ? Le cadre classique de l'apprentissage en ligne ne s'y prête guère : il suppose un échantillon fixé, pour lequel le statisticien a le temps de faire tous les calculs nécessaires afin d'obtenir une règle de décision qui ne changerait pas de sitôt. Cette période est révolue !

Les approches intrinsèquement séquentielles, permettant de s'adapter progressivement aux nouvelles données qui arrivent, connaissent ainsi un vif regain d'intérêt. Pour l'étape d'optimisation conduisant aux règles de décision, c'est notamment le cas de la descente de gradient stochastique (comme l'ont montré Francis Bach et Éric Moulines en 2013). En outre, il convient de prendre en compte le vieillissement des données.

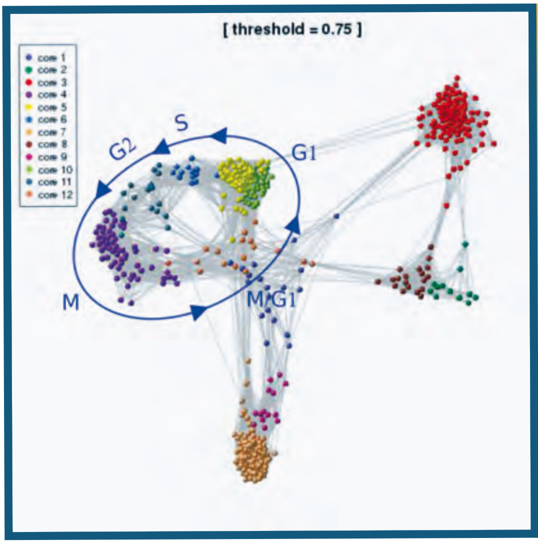
À titre d'exemple, on peut penser aux systèmes de recommandation, comme ceux qui interviennent dans la gestion du contenu des pages Web : le succès fulgurant de Criteo (www.criteo.com) souligne l'importance stratégique de cette problématique. Par ailleurs, l'organisation et le choix des prix de réserves dans les systèmes d'enchères au second prix à grande vitesse (étudiés par Nicolo Cesa-Bianchi, Ofer Dekel et Ohad Shamir en 2013) nécessitent également la mise en production ultra rapide et décentralisée d'algorithmes issus de modèles probabilistes avancés qui rappellent les problématiques du *trading* à haute fréquence.

On a besoin d'experts en aide à la décision,
en théorie des jeux et en analyse stochastique !

Ce dernier sujet, le *trading*, illustre en particulier le fait que, dans un cadre séquentiel, les décisions prises à un instant peuvent influencer les observations futures, qui du coup ne peuvent pas être traitées comme un échantillon. Ainsi, le système de recommandation ne peut se contenter de proposer les contenus dont l'appétence estimée est la plus grande car, s'il le fait, il se trouvera vite piégé, ne servant plus qu'un nombre très restreint de contenus pas assez diversifiés – les autres n'étant simplement pas assez proposés pour que l'intérêt que leur porte l'utilisateur puisse être perçu.

Toutes les branches des mathématiques sont mobilisées

À l'intérieur même des mathématiques, les contributions tendent à se diversifier. L'apprentissage profond (ou *machine learning*) s'est ainsi beaucoup appuyé sur la modélisation stochastique et statistique des données afin de construire des algorithmes fournissant des règles de décision pertinentes et afin d'obtenir des garanties théoriques d'efficacité ou d'optimalité de ces méthodes en se fondant principalement sur des outils probabilistes (parfois d'inspiration géométrique). Les gigantesques masses de données et de signaux suscitent un immense intérêt pour l'étude des matrices aléatoires de grande dimension, et en particulier de leur spectre asymptotique, alors que les premiers résultats ont été initiés dans ce domaine par des questions de… physique théorique. Le passage à l'échelle du Big Data rend indispensables et centraux de nouveaux travaux dans le domaine de l'optimisation convexe et, plus généralement, de l'analyse. La modélisation des systèmes complexes, à commencer par les grands réseaux, suscitent des rapprochements avec certains travaux de géométrie, d'algèbre ou de théorie des graphes. Comprendre la géométrie d'un espace euclidien de plusieurs dizaines de dimensions est ainsi devenu un enjeu majeur dans des domaines tels que l'assurance !



Résultat d'un partitionnement de données : estimation d'un graphe d'interaction de gènes pour des mécanismes liés au diabète.

© Anne-Claire Brunet et Jean-Michel Loubes

Le rôle du mathématicien est important pour apporter un changement de perspective, souvent contre-intuitif mais efficace, en élevant le niveau d'abstraction ou en intégrant une approche stochastique. Il y a forcément un domaine d'activité qui saura vous séduire par la façon dont il mobilisera les mathématiques !

J.-M. L.

Les statistiques pour aider au développement de médicaments

Marc Lavielle

Inria Saclay & CMAP, École polytechnique

Entre la découverte d'une molécule et la mise sur le marché d'un nouveau médicament, il peut se passer entre dix à quinze ans moyennant un investissement de plus d'un milliard d'euros… De cent mille molécules initialement passées aux cribles, seulement cent seront testées, dix seront candidates comme possibles médicaments, et seule une molécule sera éventuellement retenue comme médicament.

Développer un médicament : pas une mince affaire !

1. Une première phase de découverte permet l'identification de nouvelles molécules potentiellement utilisables comme médicaments.

2. Les études pré-cliniques permettent ensuite d'étudier le comportement des molécules candidates en interaction avec le vivant, sur des cellules en culture et sur des animaux. L'objectif est multiple : identifier les molécules qui présentent un potentiel pour traiter une maladie, mettre en évidence les possibles effets secondaires, déterminer les doses maximales tolérées.

3. Les études cliniques se font chez l'homme, avec différentes phases :

- Les essais de phase I sont menés chez un petit groupe de volontaires sains (vingt à quatre-vingts participants). L'objectif est d'évaluer son devenir dans l'organisme, son seuil de tolérance et les effets indésirables.

- La phase II consiste à tester l'efficacité du médicament et à déterminer la dose optimale sur un nombre limité de patients (quelques centaines).

• La phase III permet de comparer l'efficacité thérapeutique du traitement soit à un placebo, soit à un traitement de référence. Les groupes sont généralement consitués de plusieurs milliers de participants.

La statistique : un outil incontournable d'aide à la décision

Le développement d'un médicament est un travail d'équipe qui réunit non seulement des biologistes, des pharmacologues et des cliniciens, mais aussi des informaticiens et des mathématiciens. Le statisticien joue un rôle majeur d'aide à la décision dans toutes les étapes des développements pré-cliniques et cliniques. L'enjeu est grand : décider d'écarter une molécule ou la laisser dans la course avec des critères valables du point de vue statistique peut représenter un bénéfice considérable.

Le statisticien participe également à la conception et la planification des essais cliniques *i*) en déterminant le nombre de sujets nécessaires qui permettront de mettre en évidence avec une grande certitude un effet lorsqu'il existe, *ii*) en optimisant le protocole de l'étude (tailles des groupes, doses à administrer, instants de mesures…) afin d'obtenir le maximum d'informations à partir d'un minimum de données.

Il tente d'expliquer au mieux la variabilité des réponses à un même traitement au sein d'une population par des caractères physiologiques connus (sexe, poids, âge…). Il s'agit de mieux comprendre l'ensemble de la relation dose–réponse afin d'aider, notamment par simulation, à planifier les essais cliniques suivants en prenant mieux en compte les sources de variabilité et d'incertitude. Ces analyses reposent sur des modèles physiologiques de plus en plus complexes qui nécessitent eux aussi des outils statistiques de plus en plus complexes.

Tester des médicaments sur des patients virtuels

Comparer différents traitements sur de « vrais » patients est extrêmement long et couteux. L'utilisation de modèles mathématiques va permettre de simuler, au moyen d'un programme informatique, autant d'essais cliniques que l'on souhaite et donc comparer facilement et rapidement les effets attendus de différents traitements.

On a besoin d'équations différentielles…

Prenons l'exemple d'un médicament administré par voie orale et intéressons-nous à la quantité de ce médicament en circulation dans l'organisme. *Définir un modèle pharmacocinétique* signifie faire des hypothèses sur la façon dont le médicament est absorbé, c'est-à-dire comment il passe du système digestif au système sanguin, comment il est distribué dans le système sanguin puis éliminé. On peut ainsi supposer *i*) la vitesse d'absorption proportionnelle à la quantité de médicament dans le système digestif, *ii*) la concentration du médicament uniforme dans l'organisme à tout instant, *iii*) la vitesse d'élimination proportionnelle à la quantité de médicament en circulation.

Une fois le modèle biologique décrit, il faut le « traduire » sous forme d'équations mathématiques. Si l'on note ici $Q_s(t)$ la quantité de médicament en circulation dans le système sanguin au temps t et $Q_d(t)$ celle dans le système digestif, nos hypothèses concernant les vitesses d'absorption et d'élimination sont représentées au moyen d'équations différentielles ordinaires (EDO) :

$$\dot{Q}_d(t) = -k_a Q_d(t)$$

$$\dot{Q}_s(t) = k_a Q_d(t) - k_e Q_s(t)$$

où $\dot{Q}_d$ représente la dérivée de Q_d, k_a est la constante d'absorption du médicament et k_e la constante d'élimination. Si une dose unique D est administrée au temps 0, la solution de ce système est, pour $t \geq 0$, $Q_s(t) = Dk_a(e^{-k_e t} - e^{-k_a t}) / (k_a - k_e)$.

Prélever quelques millilitres de sang permet de mesurer la concentration plasmatique $C_s(t)$ définie par $C_s(t) = Q_s(t) / V$, où V est le volume de sang dans lequel circule le médicament. Le modèle est alors totalement défini par ces trois paramètres pharmacocinétiques, k_a, V et k_e.

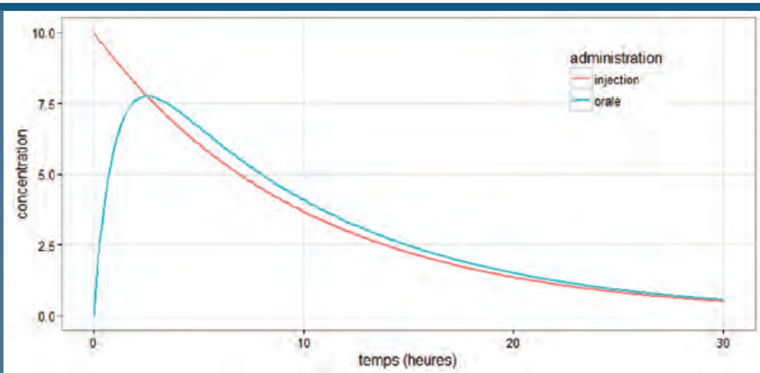


Figure 1 : Concentrations après une administration intraveineuse ou orale de 100 mg au temps 0.

La figure 1 représente les concentrations d'un médicament administré par voie orale ou par injection intraveineuse ($C_s(t) = De^{-k_e t} / V$) pour une dose de 100 mg, avec les valeurs suivantes de paramètres :
$k_a = 1\text{h}^{-1}$, V = 10 *l*, $k_e = 1\text{h}^{-1}$ (unités : milligramme mg, litre *l*, heure h).

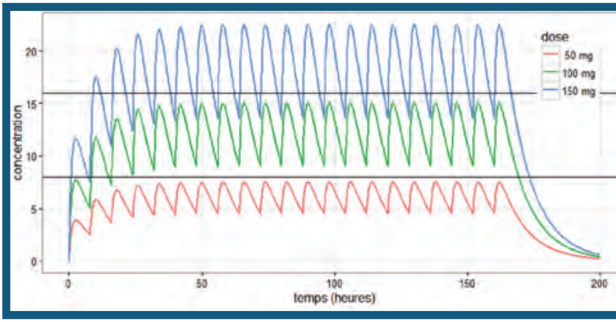


Figure 2 : Concentration d'un médicament administré toutes les huit heures pendant sept jours, avec différents dosages.

La figure 2 montre que le modèle utilisé pour une administration orale unique peut également être utilisé pour des administrations répétées toutes les huit heures, pendant sept jours. L'objectif consiste généralement à maintenir la concentration dans une fenêtre thérapeutique : cette concentration doit être suffisament importante pour que le traitement soit efficace (> 8 mg / *l* dans cet exemple), mais pas trop pour ne pas devenir toxique (< 16 mg / *l* ici). Dans cet exemple, des doses de 100 mg permettent d'atteindre ce double objectif.

… et on a besoin de probabilités

Ces paramètres sont des paramètres physiologiques qui varient d'un individu à l'autre. On les considère alors comme des variables aléatoires afin de prendre en compte cette variabilité inter-individuelle.

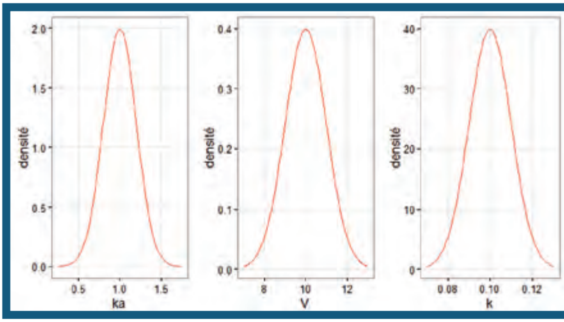


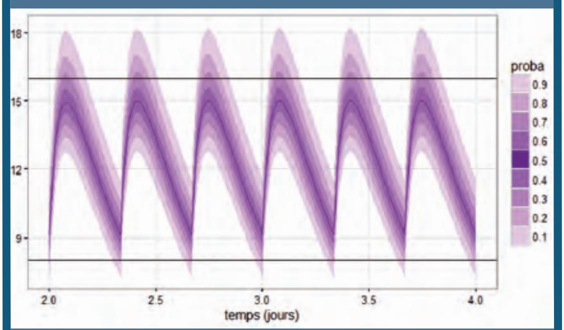
Figure 3 : Distributions de probabilité des paramètres pharmacocinétiques au sein de la population.

La figure 3 représente un exemple de distributions normales* (ou gaussiennes) – vous savez, ces fameuses « courbes en cloche » que l'on rencontre partout en probabilités et en statistiques –, pour les trois paramètres pharmacocinétiques : $k_a \sim N(1\ ;\ 0{,}2)$, $V \sim N(10\ ;\ 1)$, $k_e \sim N(0{,}1\ ;\ 0{,}01)$. Dans ce modèle, la constante d'élimination typique de la population vaut $k_e = 0{,}1\text{h}^{-1}$. Des individus éliminent le médicament plus vite ($k_e > 0{,}1\text{h}^{-1}$) et d'autres moins vite ($k_e < 0{,}1\text{h}^{-1}$).

On peut alors utiliser ces distributions de probabilité pour générer un très grand nombre de patients virtuels avec leurs propres paramètres individuels (k_a, V, k_e) et calculer pour chacun d'eux la concentration de médicament dans le sang au moyen du modèle et pour différentes doses. Les paramètres étant tous distincts, les courbes de concentration sont également toutes différentes et la fenêtre thérapeutique peut être atteinte pour certains individus mais pas pour d'autres. On peut donc, au moyen des données générées par simulation, évaluer, pour une dose donnée, la proportion d'individus pour qui le traitement n'est pas efficace ou est toxique. La figure 4 montre que le seuil de toxicité est dépassé pendant un certain intervalle de temps pour près de 20 % des patients lorsqu'une dose de 100 mg est administrée toutes les huit heures.

Figure 4 : Distribution de probabilité de la concentration du médicament les jours 2 et 3, après une administration de 100 mg toutes les huit heures.

Les deux lignes horizontales indiquent la *fenêtre thérapeutique* du médicament.



Ce modèle très simple de pharmacocinétique est couramment utilisé dans la pratique. On le couple généralement à un modèle *pharmacodynamique,* qui va décrire comment le médicament agit sur l'organisme. Ainsi, un modèle de dynamique virale du VIH va décrire comment des cellules (les lymphocytes CD4) sont infectées par le virus et produisent à leur tour des virus, et comment la thérapie antivirale inhibe la production de virus et limite l'infection des cellules cibles. En oncologie, l'utilisation de modèles permet de prédire à l'avance le développement d'une tumeur en fonction du traitement, ou la probabilité

* Une distribution normale $N(m, \sigma^2)$ est complètement caractérisée par deux paramètres, sa *moyenne m* et sa *variance* σ^2. La variance est le carré de l'écart type σ.

de décès en fonction de la taille de la tumeur. Ces modèles dépendent de paramètres physiologiques qui sont des paramètres individuels et qui expliquent pourquoi les patients ne répondent pas tous de la même façon à un même traitement. On peut là encore, au moyen de la simulation, générer des patients virtuels et comparer différents traitements pour optimiser la probabilité que le traitement soit efficace.

Du théorème à la création de *startup*, la statistique mène à tout !

Le groupe de travail Monolix réunissant statisticiens, biologistes et pharmacométriciens a été créé en 2004 dans le but de développer de nouvelles méthodes et de nouveaux algorithmes pour la modélisation mathématique en pharmacométrie. Parmi les travaux de ce groupe, certains ont donné lieu à des résultats mathématiques – des « théorèmes » – mais toujours dans le but d'une utilisation pratique, grâce en particulier au développement du logiciel Monolix.

Ces travaux ont très vite intéressé le secteur pharmaceutique. Inria a alors mis en place un projet de développement logiciel en 2009 afin de poursuivre le développement de Monolix. Le soutien financier de plusieurs industriels, dont Novartis, Roche, Johnson & Johnson, AstraZeneca et Sanofi, ont ainsi permis de convertir Monolix en un outil professionnel utilisable au quotidien dans l'industrie.

La *startup* Lixoft a été créée en 2011 dans le but d'assurer les développements futurs, la maintenance et le support de Monolix. Ce logiciel est aujourd'hui largement utilisé aussi bien dans le monde académique (pour qui il est gratuit) que dans l'industrie pharmaceutique.

Cet exemple de collaboration efficace avec le secteur industriel démontre que les mathématiques peuvent être utiles… et utilisables !

M.L.

Pour en savoir (un peu) plus :

• Introduction à la modélisation pharmacocinétique
http://model.webpopix.org/movies/PKmodelling.swf 4

• Les métiers de la statistique dans l'industrie pharmaceutique
http://gdr.statsante.fr/metiers/index.html

• Rôle et limites de la statistique dans l'évaluation des risques sanitaires liés aux OGM
http://publications-sfds.fr/index.php/stat_soc/article/view/146

• Société française de statistique (SFdS), groupe biopharmacie et santé
http://www.sfds.asso.fr/57-Presentation_des_objectifs_du_groupe

Maths, médecine, entreprises : des collaborations gagnantes

Jean-Frédéric Gerbeau
et Céline Grandmont

Équipe-projet Reo. Centre de recherche Inria
de Paris et Laboratoire Jacques-Louis Lions.

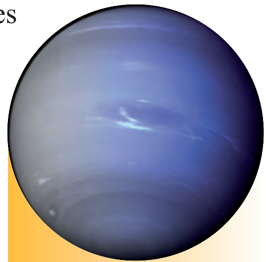
En 1954 paraît l'album de Tintin *On a marché sur la lune* (Casterman, 1954). Vous souvenez-vous de Dupond et Dupont marchant d'un pas léger malgré leur scaphandre ? Vous êtes-vous déjà demandé comment Hergé pouvait savoir que sur la Lune le poids serait six fois plus faible que sur Terre ? Quelqu'un y était-il déjà allé pour y faire des mesures ? Non, bien sûr !

C'est seulement quinze ans plus tard qu'un homme y marchera réellement. Si l'on savait depuis longtemps que le poids était différent sur la Terre et sur la Lune, c'est grâce à la théorie de la gravitation que Newton a publiée à la fin du XVIIe siècle. Plus précisément, c'est grâce au modèle de la gravitation, c'est-à-dire à une équation mathématique, que l'on peut prédire le poids sur la Lune. Sans même y aller !

Dupond et Dupont sur la Lune.

L'existence de Neptune découverte… sans lunette astronomique !

C'est aussi grâce au modèle de Newton que, en 1846, l'astronome français Urbain Le Verrier prédit mathématiquement l'existence et la position de Neptune. C'était la première fois qu'une planète était découverte grâce à la résolution d'équations mathématiques, et non grâce à l'observation. En janvier 2016, Michael Brown et Konstantin Batygin annonçaient l'existence possible d'une neuvième planète dans le système solaire, elle aussi découverte… grâce au calcul. Mais, à la différence de Le Verrier, les astronomes du XXI^{e} siècle utilisent des algorithmes programmés sur des ordinateurs pour faire leurs calculs. La résolution des équations d'un modèle à l'aide d'un ordinateur est une sous-discipline récente des mathématiques appliquées, la *simulation numérique*.



Survol de Neptune par la sonde Voyager2 en 1969.
© *NASA*

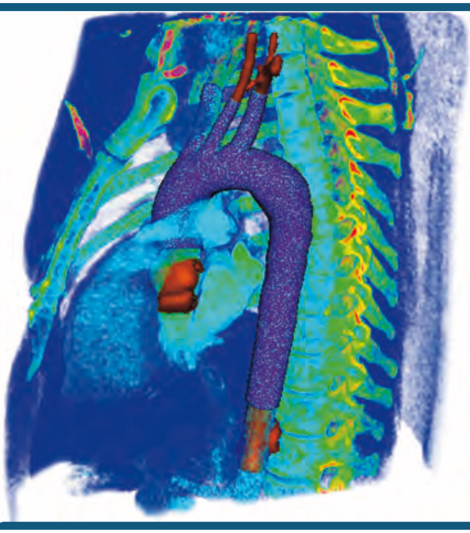
Depuis une cinquantaine d'années, la simulation numérique s'est développée dans tous les secteurs des sciences et de l'ingénierie. La combinaison des modèles physiques, des méthodes mathématiques et des ordinateurs permet, dans de nombreux domaines, la réalisation d'expériences numériques inédites. Grâce à elles, on peut prédire le temps qu'il va faire, l'échauffement d'une navette spatiale lors de son entrée dans l'atmosphère, l'évolution du climat, *etc*. Lors de la conception d'une voiture, avant de faire des *crash tests* réels, c'est d'abord par simulation numérique que l'on étudie le comportement du véhicule en cas de choc. Même lorsque la simulation ne remplace pas complètement l'expérience, elle permet d'envisager, à moindre coût et en toute sécurité, un grand nombre de scénarios, et donc de préparer au mieux les expériences réelles.

Une aide nouvelle aux médecins et aux entreprises bio-médicales

Depuis quelque temps, mathématiciens, ingénieurs, médecins et biologistes tentent d'étendre au domaine biomédical cette démarche de modélisation et de simulation. Il serait en effet formidable de pouvoir simuler un dispositif médical avant de le tester sur des animaux et sur des humains, de pouvoir envisager numériquement plusieurs scénarios chirurgicaux avant une opération, de pouvoir tester l'effet de molécules sur les cellules cardiaques avant qu'elles ne deviennent des médicaments ! Les expériences numériques dans ce domaine s'appellent parfois « expériences *in silico* », en référence au silicium qui constitue les processeurs des ordinateurs, et par analogie avec les classiques expériences *in vivo* et *in vitro* des sciences de la vie. Ces expériences *in silico* commencent à sortir des laboratoires de recherche et à apporter une aide nouvelle aux médecins et aux entreprises de bio-ingénierie.

Dans cette aorte, extraite d'une image médicale, sont résolues les équations de la mécanique des fluides pour simuler l'écoulement du sang, et de la mécanique du solide pour simuler le comportement de la paroi artérielle.

Source : équipes-projets Reo et Macs, Inria



Mais, plus précisément, en quoi consistent la modélisation et la simulation numérique dans le domaine biomédical ? Que viennent y faire les mathématiques ?

Modéliser signifie écrire des équations qui permettent de décrire au mieux le phénomène considéré : écoulement du sang dans les artères, de l'air dans l'appareil respiratoire, dépôt d'aérosols dans les bronches… Une des difficultés rencontrées dans les applications biomédicales vient souvent de l'impossibilité d'obtenir des mesures directes des paramètres intervenant dans les équations. On cherche alors à avoir un nombre restreint de paramètres représentatifs qui permettront, par exemple, de décrire certaines pathologies. Traiter les équations du modèle, pour des paramètres donnés, revient à s'intéresser au *problème direct*. L'estimation des paramètres, à l'aide de mesures ou d'examens médicaux, correspond, elle, à la résolution du *problème inverse*. Les problèmes directs et inverses sont abordés de manière approchée, grâce à la simulation numérique.

Jusqu'à la fin des années 1990, les méthodes numériques utilisées dans ce domaine se sont fortement inspirées de celles développées pour l'ingénierie classique. Comme les équations décrivant l'écoulement de l'air autour d'une aile d'avion sont, à la base, les mêmes que celles décrivant l'écoulement du sang dans les artères ou de l'air dans les poumons, les méthodes numériques développées pour l'aéronautique ont été un point de départ naturel. Cependant, de nouveaux algorithmes ont rapidement vu le jour afin de mieux prendre en compte les spécificités des applications biomédicales, conduisant ainsi également à des développements mathématiques originaux.

Les modèles mathématiques et la simulation numérique interviennent ensuite à plusieurs niveaux. Il peut s'agir, dans un premier temps, de mieux *comprendre* les phénomènes en jeu ou les observations médicales. Quand les modèles sont suffisamment élaborés, précis, adaptables à différentes pathologies, on peut alors tenter de *prédire* l'effet d'un traitement ou d'un dispositif médical, comme l'on prédit le temps qu'il fera demain. Enfin, lorsque l'on est capable de créer des modèles adaptés à chaque patient, on peut chercher à *optimiser* un traitement ou un geste chirurgical.

Comprendre :

Il n'est pas rare que l'efficacité d'un médicament ou d'un dispositif médical soit constaté par le praticien sans qu'il ne s'explique complètement pourquoi un bénéfice est constaté chez certains patients et pas chez d'autres. C'est le cas des mélanges hélium–oxygène donnés à des patients en détresse respiratoire pour améliorer la ventilation. Cependant, pour une même pathologie, le ressenti des patients peut grandement varier. Ainsi, des entreprises comme Air Liquide, qui fabriquent des gaz thérapeutiques, font appel à des mathématiciens et des physiciens pour les aider à comprendre, à l'aide de simulations numériques, les phénomènes en jeu, la disparité des réponses et ainsi essayer d'identifier pour quelles pathologies un réel bénéfice est attendu.

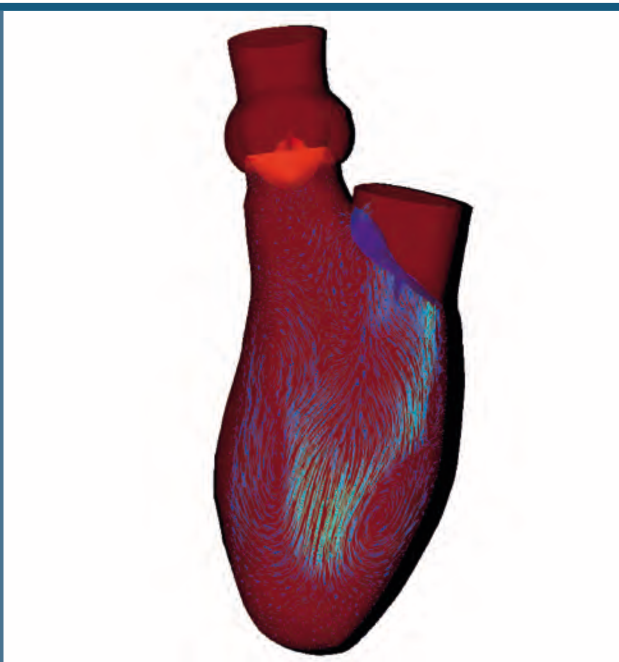
Prédire :

On crée de plus en plus de dispositifs destinés à être implantés dans le réseau artériel : des *stents*, sortes de petits grillages, que l'on dispose dans les artères pour rétablir leur diamètre, ou que l'on place devant des anévrismes pour éviter qu'ils ne se rompent ; des valves cardiaques artificielles qui ne nécessitent plus d'opération lourde ; des pompes d'assistance ventriculaire qui aident le cœur en cas d'insuffisance cardiaque. À l'avenir, les possibilités qu'offre l'impression 3D rendront ces dispositifs de mieux en mieux adaptés à chaque patient.

Les entreprises qui les développent sont souvent des *startups* (Cardiatis, Epygon, Kephalios, Corwave…) qui font appel à des mathématiciens et des biomécaniciens pour modéliser ces dispositifs avant de les tester sur des êtres vivants.

Cette démarche transforme les habitudes : au lieu de constater *a posteriori* les impacts sur le système cardiovasculaire, on cherche à les connaître *a priori*, à prédire leurs effets. Les conséquences sont évidentes en termes de coût, de sécurité et d'éthique.

Naturellement, la complexité des phénomènes et la variabilité des situations font qu'un modèle ne peut complètement remplacer le recours à l'expérience, et des allers-retours entre modélisation et expérimentation sont nécessaires. Mais pensez aux *crash tests* automobiles, aux navettes spatiales, *etc.* : ce sont aussi des situations terriblement complexes, et la simulation y joue un rôle décisif depuis une cinquantaine d'années. Il en sera peut-être de même en bio-ingénierie dans un avenir proche !



Champ de vitesse dans le ventricule gauche du cœur.
Source : équipe-projet Reo, Inria

Optimiser :

Récemment, grâce à une collaboration entre Inria (Paris), University Of California à San Diego (États-Unis) et Stanford University (États-Unis), la simulation numérique a également été utilisée pour planifier des opérations chirurgicales complexes chez des enfants souffrant de malformations congénitales du cœur. Chez cinq patients, plusieurs scénarios ont été réalisés numériquement pour aider le chirurgien à choisir la meilleure option, qui a été ensuite effectivement mise en œuvre chez de nouveaux patients.

Naturellement, l'usage de la simulation pour la planification en chirurgie vasculaire reste encore exceptionnel. Certaines solutions arrivent cependant à maturité, et pourraient se généraliser rapidement.

Ainsi, depuis quelques années, la société HeartFlow, fondée par un précurseur de la simulation numérique des écoulements sanguins, propose d'estimer par simulation les conséquences du rétrécissement d'une artère coronaire. Le gain pour le patient et le système de santé est appréciable puisque la méthode proposée évite de recourir à un examen par cathéter, coûteux et invasif.

Demain, l'explosion de la science des données

La simulation numérique dans le domaine biomédical est aujourd'hui suffisamment mature pour être utilisée non seulement dans les entreprises de bio-ingénierie, mais aussi au service des médecins. Notre époque est marquée par l'explosion de la « science des données » : des techniques d'apprentissage statistique permettent d'exploiter d'une manière redoutablement efficace les quantités de données accumulées par des capteurs, toujours plus nombreux, et souvent connectés.

La modélisation biomédicale devra aussi tirer parti de cette démarche. Il ne faut cependant pas perdre de vue la puissance qu'offrent des modèles basés sur la physique. Quand on veut prédire la gravité sur la Lune, on s'appuie sur les lois de Newton, pas sur des statistiques !

L'approche statistique est classique en médecine. La modélisation et la simulation numérique le sont beaucoup moins. Savoir tirer parti des deux est un des défis à relever pour l'avenir.

J.-F. G. & C.G.

Circulation de flux dans le cerveau et calcul scientifique

Stéphanie Salmon

Professeure au laboratoire de mathématiques de l’université de Reims

Si depuis longtemps on se pose la question de savoir comment coule un fluide, celle de savoir comment s’écoule l’air dans nos poumons ou le sang dans nos vaisseaux sanguins est plus récente. La particularité de ces fluides biologiques est qu’ils sont vivants ! Ils ne font pas que s’écouler, ils sont toujours en interaction avec leur milieu, ils transportent des nutriments, de l’oxygène, ils ont aussi une fonction biologique qui complique beaucoup leur comportement. Ainsi, les artères ne sont pas de simples tuyaux, ce sont des structures très élastiques qui se déforment au passage du sang et qui agissent sur l’écoulement de ce dernier, qui en retour déforme la structure…

Circulation de flux dans le cerveau et simulation

Notre cerveau ne représente que 2 % de notre masse corporelle mais demande pour son bon fonctionnement 16 % de notre capacité sanguine. La circulation sanguine est un circuit fermé : le sang est distribué de la tête aux pieds par les grandes artères, qui se divisent en petites artères et artérioles. Après avoir perfusé tous les organes, il est recueilli par les capillaires, conduit par les veinules jusqu’aux veines, qui le ramènent au niveau du cœur et des poumons. Les tissus artériels et veineux sont très différents en termes de diamètre et d’épaisseur, le sang s’y écoule dans un sens opposé et à des vitesses très différentes.

Notre cerveau est alimenté par un réseau tridimensionnel complexe de vaisseaux. Le sang est principalement apporté par les deux carotides (gauche et droite) et les deux artères vertébrales. Il est ensuite distribué par les artères cérébrales antérieures et postérieures. Les veines cérébrales, quant à elles, sont mal connues et peu étudiées car les principales anomalies vasculaires se développent sur les artères. Intéressons-nous donc à la simulation de l’écoulement du sang dans ces veines.

Notre cerveau contient un autre liquide : le liquide céphalo-rachidien (LCR), principalement composé d'eau. Il entoure notre cerveau et notre moelle épinière. Il les protège des chocs et a des propriétés immunologiques. La pression qui y règne, appelée pression intracrânienne, est régulée à chaque instant et assure le bon fonctionnement de notre cerveau. Quand cette pression augmente de façon délétère, les conséquences peuvent être dramatiques : on ne peut plus parler ou plus marcher car le liquide appuie sur les zones concernées de notre cerveau ! L'objectif de simuler les écoulements de liquide céphalo-rachidien est d'essayer de comprendre comment cette pression est régulée, et aussi comment on pourrait l'obtenir de manière non invasive. Car pour mesurer cette pression, seule une ponction dans le rachis ou le cerveau permet de l'évaluer ! Avec des simulations numériques de l'écoulement et des mesures de flux par IRM, l'espoir est d'obtenir une estimation de cette pression.

Comprendre et connaître les limites du travail de modélisation

Vu la complexité des systèmes étudiés, des hypothèses simplificatrices sont nécessaires. Le sang, par exemple, est un fluide *incompressible* : un volume donné ne peut tenir dans un contenant de plus petit volume, contrairement aux gaz. Mais il comporte des globules rouges, dont le diamètre est supérieur au diamètre des vaisseaux de petit calibre ! Dans ce cas, pour y circuler, les globules rouges se déforment et l'écoulement est rendu d'autant plus complexe à simuler. On supposera que le sang se comporte dans les grands vaisseaux comme l'eau qui le compose principalement. La deuxième hypothèse permettant de simplifier le problème de l'écoulement du sang dans le cerveau est de ne considérer qu'un écoulement de fluide seul et non une interaction fluide–structure.

Les interactions fluides–structures

Dans les années 1980, on a commencé à étudier et simuler sur ordinateur le couplage entre un fluide et une structure : comment réagit la structure d'un pont au vent ? Ne suffit-il pas alors d'appliquer cette procédure à l'écoulement du sang dans les artères ? Hélas, les premiers essais ne fonctionnaient pas du tout, et il a fallu des années avant que l'on comprenne pourquoi (pour simplifier, la masse volumique du sang et des artères, tous deux essentiellement composés d'eau, sont trop proches l'une de l'autre). Depuis, les mathématiciens ont amélioré leur technique de couplage. Il a donc fallu attendre pour les fluides biologiques non seulement le développement de la puissance des ordinateurs, mais aussi celle des méthodes mathématiques permettant de simuler de tels phénomènes.

Cette hypothèse est clairement fausse dans les veines jugulaires, qui se déforment bel et bien en fonction de la vitesse du sang qui s'y écoule. Le modèle ne sera donc adéquat que dans la partie intracrânienne. C'est tout l'enjeu de la modélisation : un modèle n'est pas juste ou faux, il est réaliste sous certaines hypothèses. Comprendre et bien connaître les limites du modèle fait partie du travail de modélisation.

L'écoulement des fluides dans le cerveau sera connu quand on connaîtra la vitesse et la pression du fluide en tout point des vaisseaux sanguins qui nous intéressent. Les équations à résoudre sont celles de la dynamique des fluides incompressibles, connues sous le nom d'*équations de Navier–Stokes*. D'un point de vue mathématique, on ne sait presque rien sur ces équations (c'est même un des problèmes du millénaire, mis à prix à un million de dollars par le Clay Mathematics Institute de Boston, aux États-Unis). On sait néanmoins en trouver des solutions approchées à l'aide d'ordinateurs.

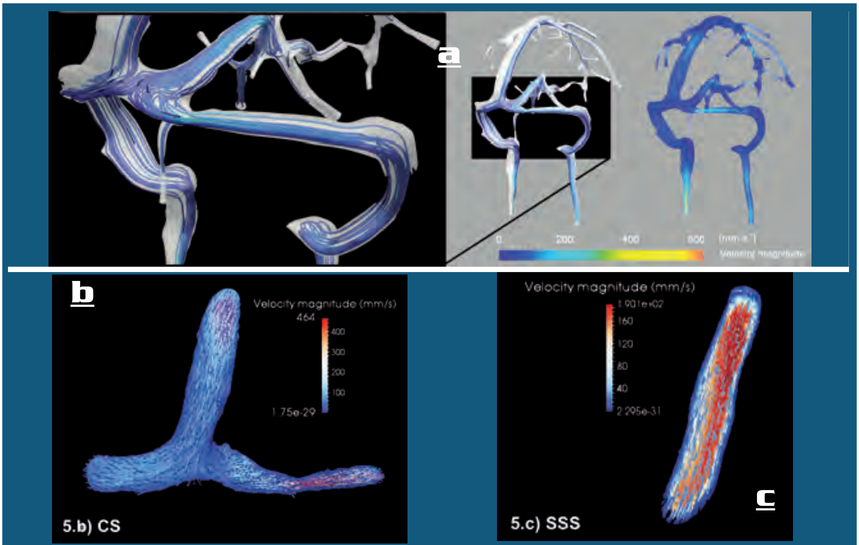
Dans ces équations interviennent la vitesse $u = (u_1, u_2, u_3)$ du fluide, sa pression, la viscosité dynamique qui le caractérise, les forces volumiques qui lui sont appliquées (la gravité, par exemple), le domaine d'intérêt Ω dans lequel on résout nos équations, et l'intervalle de temps considéré, $[0, T]$. La dernière difficulté à surmonter est le problème des conditions limites au bord de Ω (si on travaille avec des veines, des artères ou le LCR). En effet, on ne peut travailler sur le corps humain en entier, on « coupe » donc artificiellement les tuyaux pour ne regarder que ce qui nous intéresse. Mais on ajoute alors des frontières artificielles qui ne sont pas des murs et qui doivent être traitées différemment (si ce sont des entrées ou des sorties, si on travaille avec des veines ou des artères…).

En effet, en entrée des artères, le sang provient du cœur, c'est donc un flux pulsé (chaque seconde, le cœur envoie le sang aux organes). En sortie des artères, le fluide doit continuer son trajet dans le réseau et il faut donc faire en sorte qu'il y ait quelque chose à la sortie qui remplace ce réseau.

Dans les cas des veines, le sang provient de la micro-circulation, c'est donc à petite vitesse et relativement uniformément que le sang rentre dans les veines. Il en ressort, attiré par le cœur plutôt que poussé. En ce qui concerne le LCR, il oscille dans les différents compartiments qui le contiennent de façon synchronisée avec la circulation sanguine : il entre et sort de façon pulsée.

Des opportunités immenses pour des recherches futures

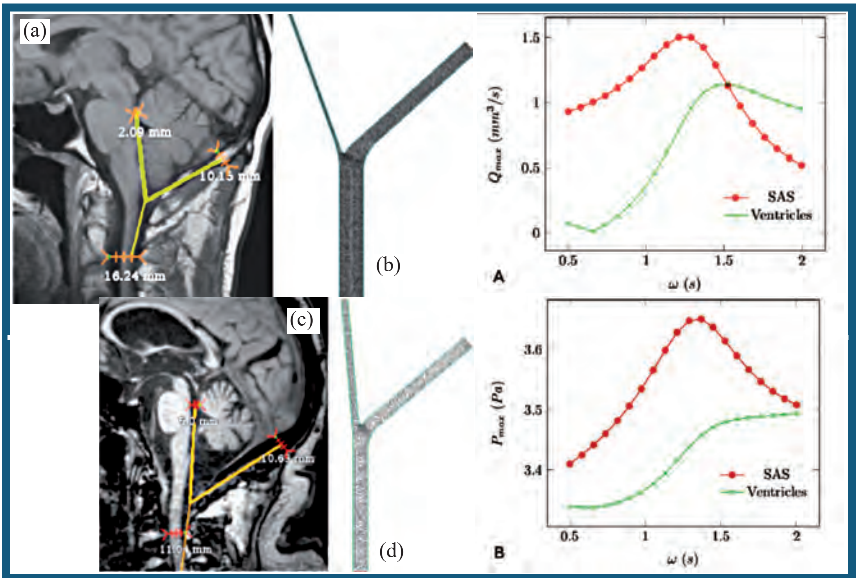
Essayons donc de comprendre la distribution du LCR dans les espaces sub-arachnoïdiens (SAS) et les ventricules. Dans certaines pathologies comme l'hydrocéphalie, le volume de LCR dans les ventricules est beaucoup trop important, entraînant une trop forte pression sur le cerveau. Ces pathologies sont souvent dues à une obstruction d'un petit canal, l'aqueduc de Sylvius. Les ventricules reçoivent normalement seulement 10 % du volume total de LCR contenu dans tout le système, mais quand il n'y a pas obstruction de l'aqueduc de Sylvius on ne sait pas toujours expliquer la raison du développement anormal du volume de LCR dans les ventricules.



Sur le réseau veineux cérébral, l'écoulement est assez simple dans le sinus sagittal supérieur, mais présente une recirculation au niveau de la confluence des sinus. On remarque aussi que l'écoulement n'est pas symétrique car la géométrie ne l'est pas non plus : la jugulaire droite draine le sang du sinus droit et des veines cérébrales profondes, la jugulaire gauche draine le sang provenant du sinus sagittal supérieur et des veines périphériques du cerveau. On lit ces informations avec ce type de simulation !

Simulation de l'écoulement du sang dans le réseau veineux cérébral :
(a) lignes de courant dans le réseau,
(b) champ de vitesse à la confluence des sinus,
(c) champ de vitesse dans le sinus sagittal supérieur.
Les vecteurs et les couleurs représentent la direction et la magnitude de la vitesse respectivement.

Quels sont alors les paramètres qui peuvent influencer la distribution du LCR ? Pour répondre à cette question, simplifions la géométrie. Considérons une bifurcation : le tube d'entrée représente la moelle épinière, le petit tube représente les ventricules, et le plus grand, les espaces sub-arachnoïdiens. Simulons alors l'écoulement du LCR en résolvant les équations de Navier–Stokes. Le flux, que l'on peut mesurer par IRM, est imposé en entrée de la moelle épinière. En sortie, mimons à l'aide de conditions limites particulières la résistance des structures.



En haut à gauche : (a) IRM morphologique, (b) maillage construit à partir des données dans un cas sain. En haut à droite : débit du LCR en fonction de la fréquence cardiaque (courbe rouge = dans les SAS, courbe verte = dans les ventricules). En bas à gauche : (c) IRM morphologique, (d) maillage construit à partir des données dans un cas pathologique. En bas à droite : pression du LCR en fonction de la fréquence cardiaque. Le débit peut être supérieur (et donc la pression) dans les ventricules avec certaines fréquences cardiaques de l'échelle physiologique.

Ainsi la fréquence cardiaque influe sur la distribution du LCR et le volume de LCR dans les ventricules peut augmenter de façon anormale et éventuellement être une cause d'hydrocéphalie.

Une des difficultés du calcul scientifique pour le vivant est la validation des résultats : on ne peut pas vraiment « voir » l'écoulement du sang dans nos artères car il est impossible d'ouvrir un crâne pour y faire des observations (qui seraient par ailleurs perturbées par l'opération). Alors, comment valider les simulations numériques ? La première piste consiste à réaliser des « fantômes physiques », des dispositifs en silicone représentatifs des artères qui nous intéressent. On y fait couler un fluide (de l'eau). On peut alors regarder l'écoulement grâce à des séquences spécifiques d'imagerie par résonnance magnétique (IRM). Une deuxième piste est de reconstruire des angiographies à l'aide des résultats de simulation. À partir d'angiographies réelles, on peut reconstituer la géométrie du réseau vasculaire cérébral des patients, puis simuler à l'intérieur de ce réseau l'écoulement du sang. À partir de ces résultats numériques, on simule l'image IRM qui aurait été acquise avec cet écoulement. Si ces images virtuelles ressemblent aux images de départ acquises sur des sujets, on aura validé toute la chaîne de travail !

Chaque étape de ce travail ouvre un champ de recherches. Celle des images médicales aux volumes reconstruits s'appelle la *segmentation* et concerne le traitement d'images. Une autre consiste à passer des volumes reconstruits à des maillages adéquats pour y faire la simulation. Puis arrive l'étape de modélisation, et enfin celle de simulation. Sans oublier la validation des simulations.

Ces simulations numériques ne sont que des premiers essais pour le cerveau : chaque étape demande en général de surmonter de nouvelles difficultés. Néanmoins, dans les prochaines années, ces simulations pourraient éventuellement expliquer une hydrocéphalie non obstructive.

S.S.

Pour en savoir (un peu) plus :

La page Web du projet Vivabrain porté par l'Agence nationale de la recherche, ***http://vivabrain.fr*** .

Biology And Mechanics Of Blood Flows. Marc Thiriet, Springer, deux volumes, 2008.

Basic Mathematical Models And Motivations. Luca Formaggia, Karl Perktold et Alfio Quarteroni, dans *Cardiovascular Mathematics. Modeling And Simulation Of The Circulatory System*, Springer, 2009.

Cell And Tissue Organization In The Circulatory And Ventilatory Systems. Marc Thiriet, Springer, 2011.

A Phase-Contrast MRI Study Of Physiologic Cerebral Venous Flow. Souraya Stoquart-ElSankari, Pierre Lehmann, Agnès Villette, Marek Czosnyka, Marc-Étienne Meyer, Hervé Deramond et Olivier Balédent, *Journal Of Cerebral Blood Flow And Metabolism* 29, 2009.

Blood Flow Simulations In The Cerebral Venous Network. Olivia Miraucourt, Stéphanie Salmon, Marcela Szopos et Marc Thiriet, *Computational And Mathematical Biomedical Engineering*, 2013.

A Windkessel Model For The Cerebrospinal Fluid. Simon Garnotel, Stéphanie Salmon et Olivier Balédent, *Computational And Mathematical Biomedical Engineering*, 2015.

Interstellar : cinéma, espace-temps et… géométrie !

Roland Lehoucq

Astrophysicien, CEA Saclay

En pleine catastrophe écologique majeure, la Terre se meurt, battue par les vents et la poussière. Les récoltes flétrissent, la nourriture manque. À cours de ressources, l’humanité est à l’agonie. Un groupe de scientifiques reclus (la NASA !) cherche une échappatoire et propulse une petite équipe dans l’espace en quête d’une nouvelle Terre. Se baladant de trous de ver en planètes exotiques, le héros finit sa course dans un trou noir pour nous apprendre que, finalement, seul l’amour transcende l’espace et le temps…

Une « théorie du tout » capable de changer notre destin ?

S’il est possible de s’accorder sur la faiblesse de l’accroche d’*Interstellar*, réalisé par Christopher Nolan et sorti dans les salles françaises en 2014, le film, héritier revendiqué du classique de Stanley Kubrick *2001, l’Odyssée de l’espace*, n’en possède pas moins un réel intérêt pédagogique tant la science y est présente et tant la volonté du réalisateur de « bien faire » est incontestable. Pour son long métrage, Christopher Nolan s’est d’ailleurs attaché les conseils du physicien américain Kip Stephen Thorne, spécialiste de la relativité générale et inventeur du concept de trou de ver.

Dans le film, le professeur Brand, sans doute le reflet de Thorne, tente de bâtir une nouvelle théorie physique capable de rendre compte en même temps de toutes les interactions fondamentales, gravitation et interactions à l’échelle microscopique. Le professeur explique que la mise au point de cette « théorie du tout » pourrait changer le destin de l’humanité en nous rendant capable d’envoyer dans l’espace les énormes vaisseaux à même de sauver l’humanité. Si l’on oublie la naïveté d’une telle prétention, il est intéressant de se demander si les équations fugitivement montrées sur les immenses tableaux noirs de son bureau ont la moindre signification pour un physicien.

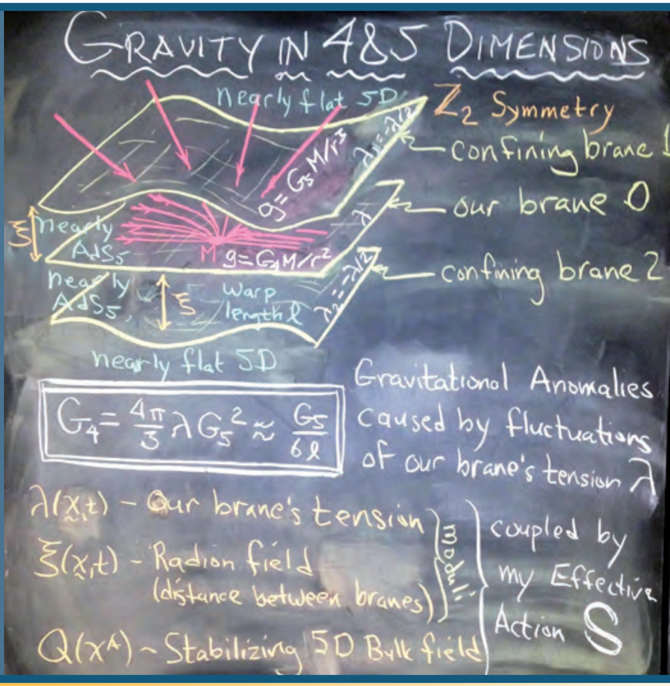


Figure 1 : © *Interstellar, Christopher Nolan*

Sur le premier tableau (voir ci-dessus), un dessin représente notre univers comme une surface (*« our brane 0 »*) prise en sandwich entre deux autres *(« confining brane 1 »*, *« confining brane 2 »*) situées dans une dimension « perpendiculaire » aux deux nôtres. Cela montre clairement que le professeur Brand bâtit sa théorie dans un super-espace – le *bulk* – ayant plus de dimensions que notre univers. En 1999, les physiciens américains Lisa Randall et Raman Sundrum proposèrent justement un modèle selon lequel notre univers n'est qu'un sous-ensemble d'une structure ayant plus de dimensions. Ce *bulk* y est décrit par un espace à cinq dimensions dit *anti de Sitter* (noté AdS_5 sur le tableau). La motivation de Randall et Sundrum était d'expliquer la faiblesse de la gravitation vis-à-vis des autres interactions : à l'échelle microscopique, l'attraction gravitationnelle entre le proton et l'électron d'un atome d'hydrogène est 10^{39} fois plus faible que la force électrique qui les lie. Imaginons alors que notre univers tridimensionnel soit plongé dans un espace comportant une dimension spatiale supplémentaire et que toutes les interactions, sauf la gravité, restent confinées dans les trois dimensions qui nous sont familières. Pouvant s'exprimer dans toutes les dimensions, la gravité semble être, pour nous, observateurs tridimensionnels, victime de « fuites » qui se traduisent par son apparente faiblesse quand nous l'expérimentons dans les trois dimensions ordinaires.

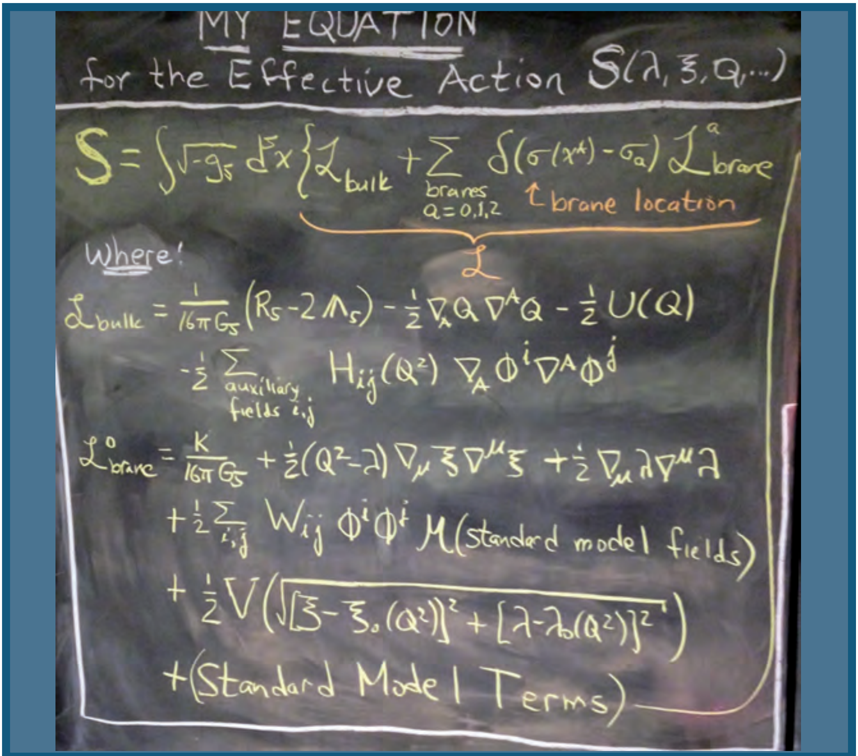


Figure 2 : © *Interstellar, Christopher Nolan*

L'intimidante formule du second tableau (ci-dessus) confirme cette interprétation. Elle exprime l'« action effective » de la théorie du professeur Brand. En physique, l'action est une grandeur physique scalaire ayant pour dimension le produit d'une énergie par un temps. Son importance est due à un principe très général appelé *principe de moindre action* : le trajet effectivement suivi par un objet entre deux points donnés est celui qui conduit à un extrémum de l'action. Ainsi, en mécanique, au lieu de penser la trajectoire comme résultant d'accélérations produites par des forces, on l'imagine comme la courbe qui optimise l'action. Ce principe de moindre action s'est avéré simple, puissant et général non seulement en mécanique classique, où il est strictement équivalent aux lois de Newton, mais aussi en électromagnétisme et en mécanique quantique ou relativiste où sa généralisation a été si fructueuse qu'elle fonde toute la physique théorique moderne. Les termes g_5 et d^5x présents dans l'équation de Brand indiquent que sa théorie est bien construite dans un espace à cinq dimensions : une de temps et quatre d'espace. Le problème est qu'il a été montré que le « sandwich AdS_5 » est instable car les deux branes qui confinent la nôtre subissent une compression ! Dans son équation, Brand tente donc d'écrire une action effective qui pourrait résoudre ce problème : elle contient des termes associés à chaque brane (numérotées de 0, 1 et 2) et un terme associé au *bulk* censé stabiliser l'ensemble.

Quand le temps devient une dimension physique

La quête du professeur, l'unification de la relativité générale et de la mécanique quantique, n'est toujours pas résolue aujourd'hui. Diverses approches du problème, comme la gravité quantique à boucles, la théorie des cordes, la géométrie non-commutative, et bien d'autres, font toujours l'objet d'intenses recherches, notamment en mathématiques.

Interstellar aborde d'une autre façon, plus visuelle, la question des dimensions spatiales supplémentaires. Vers la fin du film, Cooper, l'un des astronautes, pénètre dans le trou noir Gargantua qui est au cœur de l'intrigue. Il se retrouve alors dans un espace-temps étrange nommé Tesseract. En géométrie, le tesseract est un objet quadridimensionnel qui est au cube ce que le cube est au carré et le carré au segment de droite. De même que la surface d'un cube est constituée de six faces carrées, l'hypersurface délimitant un tesseract est constituée de huit cellules cubiques. Dans le film, cette structure géométrique est en fait un passage qui permet à Cooper de passer de la singularité centrale du trou noir au *bulk* à cinq dimensions. Il se retrouve alors hors de notre brane et, de là, il devient capable de percevoir le temps comme une dimension physique. C'est pour cela qu'une fois dans le tesseract Cooper voit la chambre de sa fille Murph, la future adjointe du professeur Brand, démultipliée à l'infini mais avec des temporalités différentes. Il est alors capable de communiquer avec elle à travers le temps au moyen de signaux gravitationnels, fournissant les données dont elle a besoin pour résoudre l'équation du professeur Brand.

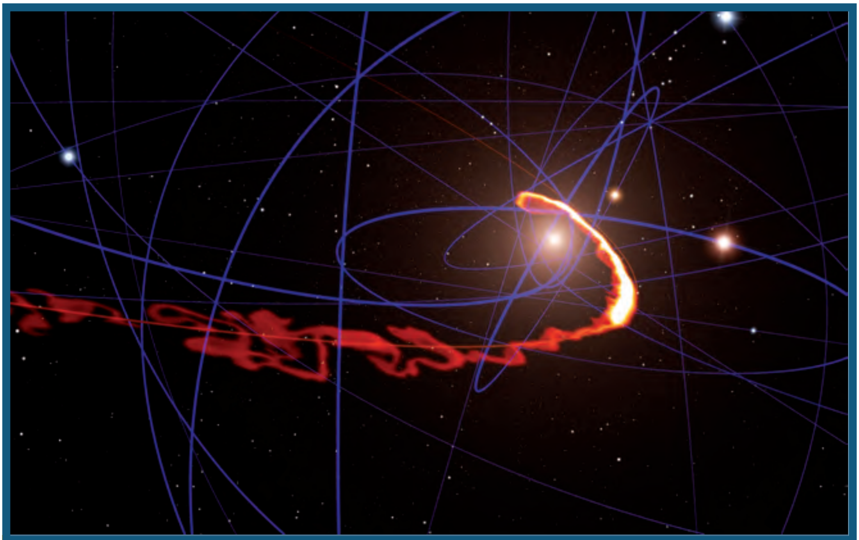
Un objet géométrique qui n'a pas fini de nous fasciner

Par ses propriétés troublantes, le tesseract, ou hypercube, apparaît dans bien d'autres œuvres. Ainsi, dans le film *Cube 2 : hypercube* (Andrzej Sekula, 2002), huit personnages sont piégés dans un tesseract qui apparaît comme un réseau de cubes connectés dont certains altèrent le déroulement du temps. Dès 1941 et sa nouvelle *la Maison biscornue* (*Astounding Stories*, 1941), Robert Heinlein décrivait une maison construite selon le patron tridimensionnel d'un tesseract. C'est ce même patron que Salvador Dalí utilisa pour sa peinture de 1954 *Crucifixion (Corpus Hypercubus)*, suggérant que, tout comme Dieu pourrait exister dans une dimension qui nous est incompréhensible, le tesseract existe dans un espace quadridimensionnel inaccessible à l'esprit humain.

Crucifixion (Corpus Hypercubus).
© Salvador Dalí, Fundació Gala-Salvador Dalí, Figueres, 2014
Photo © 1987 The Metropolitan Museum Of Art

La compréhension fine de la plupart des phénomènes décrits dans *Interstellar* requiert des connaissances en relativité générale, en mécanique quantique et une touche de théorie des cordes. Le résultat est un film à la fois intéressant et intrigant, qui tente de marier histoire forte et exactitude scientifique. Mais *Interstellar* reste une œuvre de science-fiction où liberté artistique et extrapolation scientifique font partie du jeu. En invitant les spectateurs à s'interroger sur la nature de l'espace et du temps et sur la façon dont nous les percevons, il a une vertu pédagogique et se situe en plein dans ce qu'en 1909 l'écrivain Maurice Renard nommait *« merveilleux scientifique »*, c'est-à-dire *« l'aventure d'une science poussée jusqu'à la merveille ou d'une merveille envisagée scientifiquement »*.

R.L.



Cette simulation d'un nuage de gaz passant à proximité du trou noir supermassif du centre de notre galaxie reproduit l'évènement tel qu'il était au milieu de l'année 2013. Les observations effectuées au moyen du VLT de l'ESO confirment que le nuage est à présent si étiré que sa partie avant a dépassé le point le plus proche de sa trajectoire et s'éloigne du trou noir à plus de 10 millions de km / h, tandis que la partie arrière continue de chuter sur lui.

© ESO/S. Gillessen / MPE / Marc Schartmann

Les mystérieux carnets de Ramanujan enfin décryptés !

Édouard Thomas

Mathématicien
et journaliste scientifique

« Les schémas du mathématicien, comme ceux du peintre ou du poète, doivent être beaux ; les idées, comme les couleurs ou les mots, doivent s'assembler de façon harmonieuse. La beauté est le premier test : il n'y a pas de place durable dans le monde pour des mathématiques laides. » (Godfrey Hardy)

La sortie annoncée en septembre 2016 sur nos écrans du long-métrage réalisé par Matt Brown *The Man Who Knew Infinity*, d'après l'ouvrage éponyme de Robert Kanigel (Washington Square Press, 1992), est l'occasion rêvée de parler de mathématiques en évoquant l'un de ses plus fascinants représentants, Srinivasa Ramanujan. Coïncidence inouïe : après des décennies d'acharnement, les dernières formules mystérieuses que le mathématicien indien a laissées à la postérité dans plusieurs carnets viennent d'être décryptées !

Pour la première fois au cinéma : les maths à l'écran !

Un long-métrage va nous faire revivre un épisode étonnant de l'histoire des mathématiques. Que sait-on du film ? Manjul Bhargava, médaille Fields 2014, et Ken Ono, spécialiste des mathématiques de Ramanujan, ont été associés de près à la réalisation depuis de nombreuses années. Les seules libertés artistiques prises semblent liées au physique de Ramanujan (plus corpulent que Dev Patel, sauf sur la fin de sa vie), à l'âge de sa femme Janaki (plus jeune que Devika Bhise) et de Hardy (lui aussi plus jeune que l'acteur qui joue son rôle, Jeremy Irons ; la ressemblance physique entre les deux hommes est d'ailleurs frappante). Et, cinéma oblige, l'histoire d'amour a peut-être été trop romancée.

Une décision sans retour qui va changer l'histoire

Srinivasa Ramanujan naît en 1887 dans l'État du Tamil Nadu, dans le Sud de l'Inde, alors sous domination britannique. Il appartient à une famille brahmane très modeste. Dès son plus jeune âge, il fait montre de talents scolaires certains… et d'une santé fragile. Au collège, ses dons pour les mathématiques font de lui une célébrité locale. Un enseignant lui met alors entre les mains un livre obscur, composé par un certain George Carr, répétiteur de mathématiques : il donne des leçons privées pour préparer les étudiants britanniques aux redoutables et redoutées Tripos de l'université de Cambridge, ces épreuves de mathématiques réputées les plus difficiles du monde. Pour s'aider dans sa tâche, Carr a compilé, sans aucune démonstration, tout le savoir nécessaire pour réussir les Tripos, sous la forme de deux intimidants tomes de formules et d'identités remarquables. Ces ouvrages (et leur auteur) seraient totalement tombés dans l'oubli sans Ramanujan, qui, à des milliers de kilomètres de Cambridge, découvre littéralement des milliers de formules qu'il ne connaît pas.

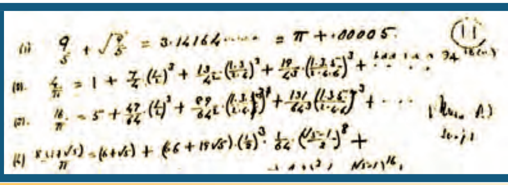
On ne saura sans doute jamais ce qu'il se passe dans la tête du farouche adolescent indien à la vue de ce flot ininterrompu de formules, données sans aucune explication, indication ni hypothèse. Y voit-il autant de défis à sa sagacité, à son intelligence ? Toujours est-il qu'à l'âge de 15 ans, semble-t-il, Ramanujan prend une décision extrême, radicale, qui va bouleverser sa vie… et le cours des mathématiques : il décide qu'il se consacrera désormais entièrement à cette science, corps et âme.

Au collège, chez lui, le jour, la nuit, Ramanujan travaille dur, avec la bénédiction de sa mère, qui voit dans la décision de son fils une quête spirituelle digne d'un brahmane observant, une quête qui honore sa famille. Ramanujan échoue à tous ses examens : il est désormais trop occupé à ses propres recherches pour s'intéresser aux matières qu'on veut lui enseigner.

Des carnets remplis de maths en guise de carte de visite

Les années passent. Ramanujan n'a toujours pas levé le nez de son ardoise, sur laquelle il réalise de complexes calculs que personne ne semble comprendre. Il a pris l'habitude de consigner ses découvertes sur des feuilles de papier récupérées et grossièrement reliées, qu'il appelle « carnets ». Le papier est une denrée que la famille ne peut s'offrir, alors Ramanujan est parcimonieux : il ne recense que les résultats qui lui semblent les plus élégants, les plus aboutis… sans la moindre explication, indication ni hypothèse.

Ramanujan est fier et conscient de ses exceptionnelles capacités en mathématiques ; il veut épater le monde par son habileté technique à manipuler les objets les plus délicats de l'analyse : ses trois carnets, qui ne le quittent pas, sont sa carte de visite. Avec eux, il compte bien en mettre plein la vue à ses interlocuteurs et trouver enfin un mécène ou une institution qui reconnaisse son génie et lui offre, en échange du fruit de ses recherches, de quoi se loger et se nourrir, et du temps pour sa longue quête solitaire.



Ramanujan a recensé cette jolie formule donnant une estimation de π comme somme de 9 / 5 et de sa racine carrée. La précision est étonnante ! Aucune trace d'une telle formule n'a apparemment jamais été trouvée auparavant : elle semble totalement originale.

Komalatammal, la mère de Ramanujan, décide finalement qu'il est temps que son fils de 21 ans s'assume dans la vie : elle lui arrange un mariage avec la jeune Janaki. Ramanujan redouble d'efforts pour faire reconnaître la valeur de son travail. Il monte à Madras, aujourd'hui Chennaï, et rencontre des mathématiciens, qui ne savent se prononcer sur la qualité des résultats patiemment produits par un génie obstiné… ou par un illuminé. Petit à petit, à force de persévérance et d'audace, il se constitue un réseau. Après bien des errances, il trouve un emploi de bureau.

Deux hommes que tout oppose réunis par une même passion

Ayant consulté, sans succès, tout ce que le Sud de l'Inde comptait comme mathématiciens professionnels, Ramanujan écrit au plus grand scientifique britannique depuis Newton, Godfrey Harold Hardy. Ce dernier saura reconnaître le génie du jeune Indien et réussira à le faire venir à Cambridge… quatre mois avant que n'éclate la Première Guerre mondiale. S'ensuivra l'une des collaborations les plus fructueuses de l'histoire des sciences, malgré des conditions épouvantables, entre guerre et maladie, entre privations et solitude. Les deux chercheurs ouvrent mille chantiers, réussissent à publier une vingtaine d'articles, dans un contexte particulièrement hostile. Un vrai choc culturel entre deux hommes que tout oppose (de la pratique des mathématiques au rapport à la religion en passant par les codes sociaux ou les habitudes alimentaires) et qu'une même passion a réunis.

Durant les années 1917 et 1918, Ramanujan est souffrant, sans que personne ne sache bien de quel mal il est atteint. Confiné dans des sanatoriums, aucun médecin ne sait le soulager. De son côté, Hardy s'active pour le faire élire en 1918 à la Royal Society, vénérable société savante s'il en est. En Inde, c'est l'euphorie : leur compatriote est entré dans l'histoire. Avec une telle reconnaissance, Ramanujan retrouve des forces pour ses recherches, sa santé s'améliore. En 1919, il est même en état de retourner au pays ! Il y retrouve sa femme, qui a maintenant 20 ans et peut prendre soin de lui. Hélas, sa santé se dégrade à nouveau, et en Inde non plus personne ne réussit à le soulager. Début 1920, Ramanujan se meurt. Du fond de son lit de douleur, il envoie à Hardy une lettre au contenu mathématique mystérieux… et prophétique. Quelques jours plus tard, Ramanujan s'éteint, le 26 avril 1920. Il avait 32 ans.

Les mathématiques au-delà de la mort : transmission et héritage

Dès la mort prématurée de Ramanujan, Hardy comprend l'importance de réaliser une « édition » des trois carnets, qui ont été compilés en Inde entre 1903 et 1914. Il inspecte les documents (originaux ou copies manuscrites) à sa disposition. Ils contiennent des milliers de formules patiemment compilées, des formules si alambiquées que *« si Ramanujan ne les avait pas écrites, personne ne les aurait trouvées, même dans cent ans, même dans deux cents ans »* (Bruce Berndt). Hardy précise très vite les « règles du jeu » de l'édition : (*i*) pour les résultats déjà connus, produire une référence précise (la plus ancienne possible) ; (*ii*) pour les résultats corrects originaux, produire une démonstration (si possible dans l'esprit de Ramanujan) ; (*iii*) pour les résultats faux, un résultat correct doit se nicher pas loin (!), il faut le chercher…

Mais il y a un problème : dans ses écrits, Ramanujan ne définit pas les objets qu'il manipule, il invente ses propres notations, il ne s'embête pas avec les contre-exemples « évidents » à ses énoncés, il n'explique rien. Les contenus ne sont pas structurés, pas organisés, sauf dans le deuxième carnet. Ces documents personnels n'ont en effet jamais été destinés à être lus ou publiés ! En outre, de son vivant, leur auteur n'a révélé ni les clés ni les sources de sa puissante intuition, ni ses mystérieuses méthodes, pas même à Hardy. Aussi le chantier éditorial est-il immense.

Hardy s'entoure de deux collègues, Bertram Wilson et George Watson, qui, dès 1929, vont se consacrer pleinement à la validation complète des formules de Ramanujan contenues dans les trois précieux documents*. Le deuxième carnet est privilégié. La tâche est estimée à pas moins de cinq années de travail. Ce sera autrement plus délicat : dix années plus tard, en 1940, près de quarante articles ont été publiés, correspondant peu ou prou à l'édition de… six à huit chapitres du deuxième carnet ! Hélas, Wilson est décédé prématurément en 1935, et Watson ne peut poursuivre seul très longtemps. À la mort de ce dernier en 1965, tous ses documents de travail sont collectés, puis envoyés à la bibliothèque du Trinity College. Ils dormiront là plusieurs années : leur importance ayant été largement sous-estimée, peu de personnes en connaissent l'existence.

On comprend sur cette page manuscrite toute la difficulté du travail de validation des milliers de formules laissées par Ramanujan : quantités non définies, énigmatiques points de suspension, version originale quasi illisible…

Un travail d'édition moderne, systématique, tenace

Nous sommes en 1974. Le mathématicien américain Bruce Berndt, spécialiste de théorie analytique des nombres, s'aperçoit que, avec les outils qu'il a développés plusieurs années plus tôt, il peut démontrer élégamment certaines formules dues à Ramanujan. C'est le déclic : il décide de poursuivre le travail de Watson et Wilson. Au même moment, en 1976, son collègue George Andrews retrouve un peu par hasard, dans les archives de la Trinity College Library, un quatrième carnet de la main de Ramanujan, qu'il s'empresse de baptiser *Lost Notebook* (« carnet perdu ») ! Ce document de 138 pages était en possession de Hardy, qui l'a transmis à Watson.

Finalement, entre 1977 et 1997, en s'aidant des travaux de ses prédécesseurs, et avec le soutien de collègues, de plus de vingt doctorants et de l'éditeur Springer, Berndt réalise intégralement, dans cinq ouvrages publiés entre 1981 et 1997, la colossale tâche d'édition des trois carnets selon les règles édictées par Hardy. Ce travail monumental (plus de deux

* Entre 2013 et 2016, une version digitalisée des trois carnets (351 pages, 356 pages et 33 pages respectivement), du « carnet perdu » (138 pages) et de bien d'autres documents a enfin été rendue disponible ; retrouvez-les en ligne sur http://ramanujan.sirinudi.org et sur le site du projet Janus (https://janus.lib.cam.ac.uk).

mille pages pour tout décortiquer avec méthode et rigueur) reste à ce jour l'un des plus importants chantiers éditoriaux jamais entrepris en sciences : les trois carnets recensent, d'après Berndt, plus de trois mille résultats.

Mieux : le « carnet perdu » a lui aussi cédé ! Entre 2005 et 2013, quatre volumes ont été publiés par Andrews et Berndt. Aujourd'hui, ce nouveau chantier s'achève : tous les écrits connus de Ramanujan ont été impitoyablement passés au crible des experts, et le cinquième et dernier volume va paraître prochainement (nous avons pu le consulter).

Le bilan laisse sans voix : sur les milliers de résultats laissés à la postérité, moins d'une centaine sont irrémédiablement faux, et les deux tiers sont totalement originaux. En particulier, le contenu de la mystérieuse dernière lettre de Ramanujan a enfin révélé une partie de ses secrets. On reste admiratifs devant le caractère prophétique de la vision mathématique qu'elle contient et qui occupe aujourd'hui de nombreux chercheurs.

Attention, ce n'est que le début du travail !

Le chantier est-il achevé ? Non, il commence seulement, car une grande partie des démonstrations maintenant publiées fait appel à l'outil informatique, ou à des théories modernes, ou que Ramanujan ne maîtrisait pas (comme la théorie des fonctions d'une variable complexe). Mais alors, comment faisait-il, lui, pour *prouver* ses formules, et surtout pour les *trouver*, en premier lieu ?

Berndt pense que Ramanujan fonctionnait comme tout mathématicien, mais avec plus de perspicacité, de façon plus pénétrante (*« with more insight »*). Nous sommes encore loin d'avoir retrouvé toutes ses méthodes ! Ce qui est à peu près unique dans l'histoire des mathématiques, surtout à l'échelle d'une telle masse de résultats. Retrouver l'intuition originelle, maintenant que l'on sait que les résultats recensés sont corrects : voilà le nouveau chantier…

É.T.

Pour en savoir (un peu) plus :

The Man Who Knew Infinity. Robert Kanigel, Washington Square Press, 1992.

Les carnets indiens de Srinivasa Ramanujan. Bernard Randé, Cassini, 2002.

Pour un aperçu grand public des mathématiques de Ramanujan, vous pouvez consulter le document PDF de présentation issu de la conférence de l'auteur sur le site du CIJM :
www.cijm.org/accueil/productions-cijm/90-maths-express
(« Complément d'enquête sur Ramanujan »).
Conférence publique donnée sur l'espace « rencontres » du Salon de la culture et des jeux mathématiques le samedi 28 mai 2016 à 11 h.

Des outils mathématiques pour votre GPS

Nicolas K. Blanchard
et Leila Gabasova
Institut de recherche en informatique fondamentale et Université Paris-Saclay.

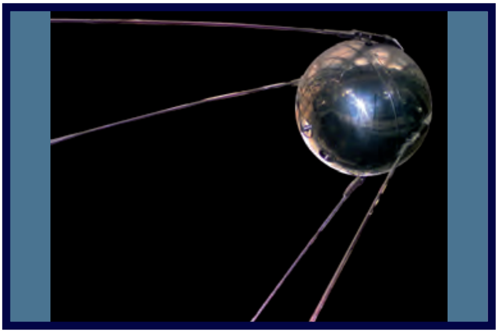
Développé durant la guerre froide, le GPS (pour Système de positionnement global) permet à une personne munie d'un récepteur radio sophistiqué de trouver avec grande précision sa propre position sur Terre. Ce système, fruit de prouesses tant scientifiques que technologiques, repose sur un ensemble de trente-deux satellites en orbite basse. Chacun de ces satellites est équipé d'une horloge atomique, c'est-à-dire d'un dispositif permettant de connaître le temps écoulé avec une très grande précision – les horloges atomiques modernes pouvant fonctionner pendant plusieurs centaines de millions d'années avant d'accumuler une seconde de décalage. Ces satellites diffusent en permanence l'heure ainsi que leur propre position. En regardant la différence entre l'heure actuelle et l'heure reçue, on peut trouver la distance au satellite ! C'est en utilisant ces distances avec plusieurs satellites différents et leur position que l'on peut calculer sa propre position, avec une technique nommée *trilatération*. Avant de voir cela en détail, remontons un peu dans le passé.

Se repérer grâce aux satellites : une histoire militaire

Le 4 octobre 1957, Spoutnik 1 devient le premier objet d'origine humaine à être envoyé en orbite, déclenchant par cela la course à l'espace. Ce satellite artificiel n'embarque qu'un émetteur radio, diffusant une série de « bips » sur une fréquence radio standard. Plusieurs scientifiques américains essayent immédiatement d'en savoir le plus possible sur ce satellite, mais ce sont ceux du laboratoire de physique de l'université Johns-Hopkins qui vont changer la donne.

Une réplique
du satellite Spoutnik
envoyé en orbite
par l'URSS
le 4 octobre 1957.

© *NASA*



Ces derniers, faute de moyen, n'ont accès qu'à une antenne radio standard, mais se rendent compte que le signal est très proche de 20 MHz, avec un petit décalage qu'ils estiment dû à l'effet Doppler – un signal provenant d'une source qui s'approche paraît plus aigu, au contraire d'une source qui s'éloigne, comme on peut le vérifier dans la rue lorsque passe une ambulance toutes sirènes hurlantes. Les scientifiques essayent de deviner l'orbite du satellite à partir de cette simple information et obtiennent, après bien des efforts, une précision correcte. Le directeur de leur institut leur demande alors s'il est possible de faire le contraire, c'est-à-dire de connaître la position d'une personne à l'aide de satellites (le positionnement précis des navires était un des problèmes majeurs de la marine états-unienne de l'époque). Deux ans plus tard, le 13 avril 1960, un premier satellite est lancé, inaugurant le système NAVSAT, ancêtre du GPS.

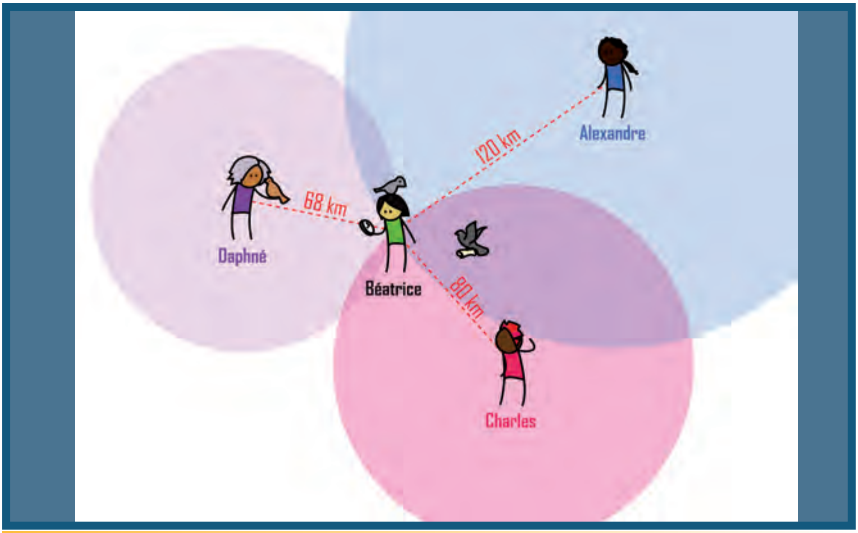
Au cours des décennies suivantes, les systèmes deviennent extrêmement perfectionnés, et massivement accessibles aux civils, avec une contrainte cependant : le système étant toujours contrôlé par la défense américaine, une erreur volontaire est introduite pour les appareils civils et étrangers, faisant chuter la précision (qui passe d'une dizaine de mètres à une centaine de mètres). Cette erreur volontaire est l'un des arguments utilisés pour la création du système Galileo, équivalent européen du GPS en cours de déploiement. D'autres systèmes concurrents, collectivement nommés GNSS (pour Système global de navigation par satellite), sont aussi développés par la Russie, la Chine et l'Inde, afin d'améliorer les performances et de ne pas dépendre du gouvernement des États-Unis. Comment ce système fonctionne-t-il en pratique ?

Une idée simple : le positionnement par pigeon voyageur

Pour simplifier temporairement, imaginons que nous sommes sur un plan, où Alexandre envoie un pigeon voyageur à Béatrice à 4 h 05, avec un message contenant l'heure d'envoi et la position d'Alexandre. Si

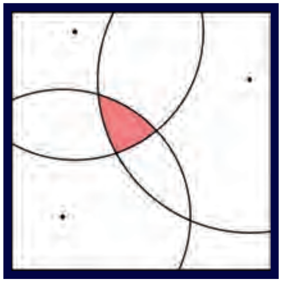
Béatrice reçoit le pigeon à 5 h 35 et que le pigeon voyageait en ligne droite à 80 km / h, Béatrice sait qu'elle est à 120 km d'Alexandre.

L'ensemble des points où Béatrice peut se trouver est donc le cercle de rayon 120 km autour d'Alexandre. Si Charles envoie un autre pigeon (allant à la même vitesse) à 4 h 30 et que Béatrice le reçoit à 5 h 30, elle sait qu'elle est à 80 km de Charles, et donc à l'un des deux points d'intersection des deux cercles (comme sur la figure ci-dessous). Avec un troisième pigeon provenant d'une autre personne, Béatrice pourra trouver où elle se trouve parmi les deux points restants. Un problème subsiste toutefois : que se passe-t-il si la montre de Béatrice est en retard ?

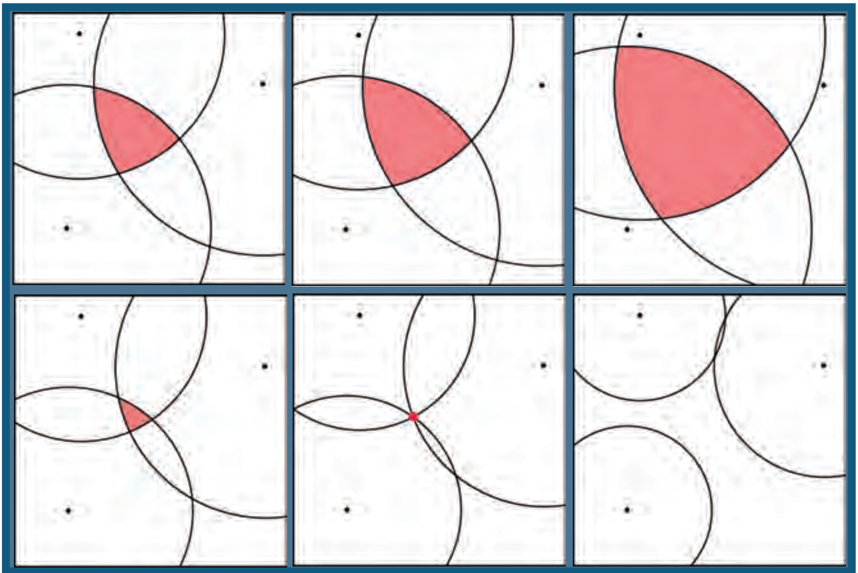


En sachant combien de temps il faut au pigeon pour arriver, Béatrice peut deviner qu'elle se trouve à l'intersection de plusieurs cercles, chacun correspondant à une distance donnée depuis Alexandre, Charles et Daphné.

Si la montre de Béatrice retarde de six minutes, elle va penser être à 8 km de plus de chacune des autres personnes, et les trois cercles ne se croisent plus en un point unique. Comment faire alors pour retrouver sa position réelle ? Il existe *a priori* toute une zone dans laquelle Béatrice pourrait se trouver…



La solution est pourtant plus simple qu'il n'y paraît : en fait, on n'a qu'une inconnue, le retard de la montre ! En effet, si la montre a x minutes de retard, chacune des distances va augmenter de x fois la vitesse. En retrouvant le retard, on peut donc obtenir la position exacte de Béatrice. En faisant des approximations, on peut retrouver x, par exemple en utilisant la méthode des moindres carrés : on essaye successivement plusieurs valeurs possibles pour x, en tentant de minimiser la taille de la zone d'incertitude.



En essayant des décalages positifs (en haut) ou négatifs (en bas) de une minute (à gauche), deux minutes (au centre) et cinq minutes (à droite), on peut trouver la solution la plus probable, ici deux minutes de retard.

Géométrie dans l'espace : quatre satellites suffisent… en théorie

Dans un cadre plus réaliste, comme la Terre, on peut appliquer la même méthode. Cette fois on n'a plus besoin de l'intersection de trois cercles, mais de quatre sphères – dans l'espace, l'intersection de deux sphères est (au plus) un cercle, celle de trois sphères correspond à deux points, et celle de quatre sphères donne donc un point unique. Quatre satellites suffisent pour retrouver la position de Béatrice. En cas de problème, on peut même se contenter de trois satellites, si l'on suppose que Béatrice est à la surface de la Terre, une des sphères étant alors déjà donnée (mais la précision sera alors très mauvaise). Les GNSS ont donc besoin de nombreux satellites, entre vingt et trente, pour garantir que chaque personne soit à tout instant visible par au moins trois satellites distincts. Pour avoir une meilleure précision, on essaye généralement d'en avoir plus (entre six et douze pour le GPS).

Dans ses travaux au début du XX^{e} siècle sur la relativité spéciale et générale, Albert Einstein montre que le temps s'écoule différemment pour des observateurs allant à des vitesses différentes, et que l'espace ainsi que le temps peuvent être déformés par la présence d'objets massifs comme une planète ou une étoile. Avec son collègue Paul Langevin, ils réalisent l'expérience de pensée suivante. Ils considèrent deux vrais jumeaux, l'un restant sur Terre, l'autre voyageant à haute vitesse dans l'espace avant de revenir. Leurs réflexions théoriques les conduisent à conclure que le jumeau voyageant dans l'espace devrait « vivre plus lentement », c'est-à-dire que le temps se dilaterait pour lui et qu'il devrait revenir en ayant moins vieilli que celui resté sur Terre.

Un autre effet contrebalance cette conclusion : la Terre, massive, déforme l'espace-temps autour d'elle. Ces deux effets seront testés pour la première fois en 1971 en embarquant des horloges atomiques dans des avions faisant le tour de la Terre : sans surprise, après un tour complet, les horloges avaient un décalage avec celle qui était restée au point de départ. À cause de ces effets, les horloges à bord des satellites accumulent un retard, correspondant à près de quarante nanosecondes par jour (une seconde dure un milliard de nanosecondes). Cet écart peut sembler faible mais c'est déjà suffisant pour avoir une erreur de plusieurs kilomètres sur la position de Béatrice !

Heureusement, des outils théoriques permettent de prédire de manière extrêmement précise ce décalage, et les ordinateurs à bord des satellites le corrigent automatiquement.

De l'utilité des imprécisions pour la météorologie

Béatrice allume donc sa radio, reçoit des signaux de plusieurs satellites indiquant avec une très haute précision leur heure d'envoi, et calcule sa position la plus probable en devinant le retard qu'a sa propre montre. Cela suffit pour deviner sa position, souvent à quelques mètres près. Comment donc peut-on expliquer l'imprécision qui subsiste ? Une partie vient bien sûr de petits décalages dans les orbites des satellites ou dans les horloges, ou de trajectoires d'ondes qui ne sont pas rectilignes (car on se trouve rarement au milieu d'une plaine). Cependant, ces erreurs sont trop faibles pour expliquer cette précision relativement mauvaise.

Il se trouve qu'en effet des obstacles subsistent : l'humidité de l'air dans les couches basses de l'atmosphère, et la présence d'ions dans les couches supérieures, perturbent et retardent les ondes. Quand l'onde envoyée par le satellite traverse l'atmosphère, elle est ralentie. Ce ralentissement n'est pas connu de manière exacte, et se trouve être à la source de cette relative imprécision. On peut l'éviter en envoyant plusieurs signaux à des fréquences différentes (qui vont être ralentis de manière différente, permettant d'estimer le délai supplémentaire), et cette technique est mise en pratique dans les appareils militaires ou scientifiques de haute précision, allant jusqu'au millimètre.

Cependant, on peut aussi l'exploiter : si l'on connaît avec précision notre position, on peut en déduire le ralentissement dû à l'humidité, et trouver, à partir de ce résultat, l'humidité, et aussi la température de l'air !

Cette technique très récente a déjà de nombreuses applications en météorologie, car elle nous permet d'étudier avec précision la température de l'air en tout point du globe, et donc le réchauffement climatique avec des données globales de haute qualité.

N.K. B. & L.G.

Explorer la surface du globe… sans sortir de chez soi

En combinant des photographies avec des données GPS (comme sur les Google Cars), des photos par satellite, des plans de ville accessibles publiquement et des corrections fournies par les utilisateurs, plusieurs services existent permettant d'explorer la surface de la Terre. Certains proposent même de reconstruire partiellement en 3D les villes, en utilisant des techniques mathématiques et algorithmiques avancées (et dont les détails sont secrets ; les enjeux financiers semblent en effet gigantesques) pour deviner la forme des bâtiments à partir de plusieurs images. Ces développements récents sont aussi dus à l'apparition de nouvelles techniques en reconnaissance d'image et en intelligence artificielle.

Pour en savoir (un peu) plus :

Genesis of Satellite Navigation. William Guier et George Weiffenbach, 1998, disponible en ligne.

Global Positioning System: The Mathematics Of GPS Receivers. Richard Thompson, *Mathematics Magazine* 71, 1998, disponible en ligne.

Scientists Use GPS Signals To Measure Earth's Atmosphere. Turner Brinton, 2007, disponible en ligne.

Des mathématiques pour le climat

Didier Bresch

Directeur de recherche CNRS
au laboratoire de mathématiques
de l'université de Savoie-Mont-Blanc

Quelles mathématiques peuvent être utiles pour aider à lutter contre le réchauffement climatique ? La question semble compliquée, mais la réponse sera simple : toutes les mathématiques, leurs développements actuels et futurs !

Savoir mobiliser toutes les branches des maths

Le mathématicien étant un observateur privilégié du monde, il peut créer des outils conceptuels efficaces. Lorsque l'on s'intéresse aux mathématiques pour le climat et la prévention des risques naturels, ou plus généralement aux mathématiques pour la planète Terre, il faut garder en tête le triptyque suivant :

- *Mathématiques en émergence* : développement de nouvelles mathématiques motivé dès le départ par des problématiques sociétales et environnementales (faire face aux hétérogénéités, aux phénomènes à très forts gradients, aux changements d'états brutaux, aux différentes échelles ; gérer l'aléa et l'environnement incertain, les évènements rares…) ;
- *Mathématiques du monde réel* : outils de résolution d'un problème environnemental en assurant le transfert avec les mathématiques les plus avancées (comprendre grâce à la théorie ; observer et simuler des phénomènes non reproductibles en laboratoire ; aider à la gouvernance des territoires et à la gestion durable des écosystèmes…) ;
- *Mathématiques numériques* : pont entre l'abstrait et le concret se frottant à des problèmes provenant des applications (couplage de processus ; validation de modèles ; paramétrages ; calcul haute performance ; rareté *versus* abondance des données ; quantification d'incertitudes…).

Le système climatique a été créé par l'évolution interactive de l'atmosphère, des océans, de la cryosphère (glace terrestre ou marine, manteau neigeux), des systèmes hydrologiques (océans, lacs, rivières, nappes souterraines), de la biosphère (tous les organismes vivants, que ce soit dans l'air, dans la mer ou sur la terre). Il est donc extrêmement complexe, les problématiques sont fortement hétérogènes (multi-échelles et multiprocessus).

Un travail par « nature » hautement pluridisciplinaire

En 2013, Michael Ghil, chercheur confirmé en dynamique du climat, a rédigé un article intitulé *Mathématiques et Étude du climat* pour la brochure *Maths de la Planète Terre Express* (disponible en ligne, comme toutes les brochures éditées par le Comité international des jeux mathématiques). Son dernier paragraphe est sans appel : *« Plusieurs grandes personnalités dans l'étude du climat, dont Jule Gregory Charney et Edward Lorenz, possédaient une importante formation en mathématiques. Actuellement de nombreuses thèses ne consistent qu'en l'analyse d'une simulation obtenue à l'aide d'un ordinateur ou d'un jeu de données satellitaires. J'aimerais bien être témoin d'une renaissance de l'application des méthodes mathématiques les plus avancées à l'étude du climat. L'étude rigoureuse de la robustesse des résultats de modèles très complexes est un champ fascinant avec des retombées importantes dans d'autres domaines, tels que les modèles biologiques ou économiques. »*

Il faut éviter de ne privilégier qu'une ou deux parties du triptyque Mathématiques en émergence / Mathématiques du monde réel / Mathématiques numériques. En outre, il est dangereux de déconnecter mathématiques et autres disciplines comme la physique. Le point clé est de tendre vers la science en lien avec la Terre, sans barrières entre les disciplines, tout en préservant la richesse d'une articulation : approche fondamentale et approche vers l'application, approche globale sur modèles complexes et approche locale sur modèles simplifiés.

Les questions provenant de problématiques concrètes doivent dicter la mise en place des mathématiques pour la Terre, que ce soient les mathématiques en émergence, les mathématiques du monde réel ou les mathématiques du numérique.

Le vivant n'est pas oublié : l'étude des déplacements collectifs a des répercussions sur la compréhension des phénomènes biologiques, mais aussi sociaux ! Ainsi, il convient d'élaborer de nouveaux outils et formalismes mathématiques permettant d'aborder les phénomènes d'auto-organisation et d'émergence.

Météorologie, modélisation du climat et prévision du temps

En termes de systèmes dynamiques, la météorologie est chaotique et n'est pas prévisible au-delà de quelques jours, mais le climat, lui, pourrait être prévisible. De nombreux aspects universels émergent, qui peuvent être compris grâce au couplage analyse–probabilités–statistiques. La compréhension de la statistique des écoulements turbulents, et plus généralement de la mécanique statistique des systèmes complexes, est un enjeu multidisciplinaire ; les mathématiques ont un rôle important à y jouer. Les mathématiques peuvent aider à structurer les discussions autour du climat ! Sous l'impulsion de von Neumann, le Meteorological Project est lancé, en 1946, à Princeton (États-Unis). En moins de dix ans, l'équipe dirigée par Jule Gregory Charney obtient une prévision du temps opérationnelle pour tout le territoire américain. Mais le calcul par ordinateur permet aussi d'effectuer facilement des expériences numériques, autrement dit de tester des hypothèses sur les mécanismes climatiques.

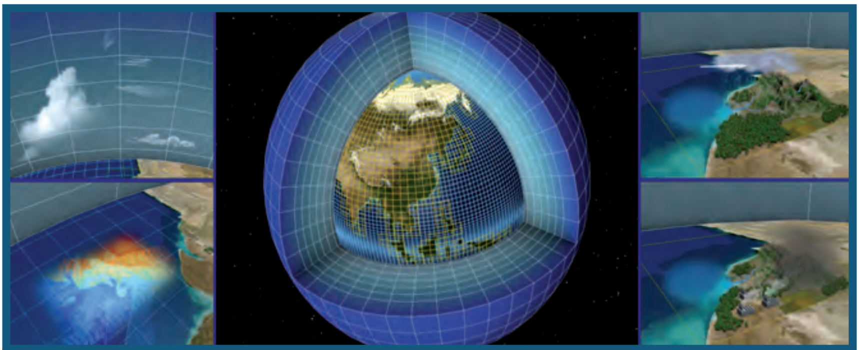
L'amélioration de la prévision du temps passe par une meilleure compréhension des mécanismes généraux du climat. Dès le début des années 1950, les scientifiques de Princeton développent, à côté des *modèles de prévision*, des *modèles de circulation générale*, dont l'objectif est de reproduire les propriétés moyennes des mouvements de l'atmosphère. Ce sont les ancêtres des modèles de climat actuels. Le premier modèle numérique de circulation générale de l'atmosphère est achevé par Norman Philipps en 1955. À partir de formulations simplifiées de la dynamique, il parvient à reproduire les grands traits de la circulation à l'échelle du globe (courants, jets, perturbations, transports d'énergie).

Projet Météorologie à l'Institut de recherches avancées, Princeton, 1952.
De gauche à droite : Jule Charney, le calculateur MANIAC I, Norman Phillips, Glenn Lewis, Norma Gilbarg, George Platzman.

Courtesy AIP Emilio Segrè Visual Archives.

Mouvements de l'atmosphère : la dynamique et la physique

Dans les années 1970 s'effectue la convergence entre les modélisations du climat et du temps : les modèles opérationnels de prévision cherchent à simuler des mouvements de l'atmosphère en se basant sur les lois de la physique. L'atmosphère est représentée par un maillage à trois dimensions, et l'ordinateur calcule pour chaque maille et à chaque pas de temps les variables caractérisant l'état de l'atmosphère (température, pression, vent, humidité…) à partir de leurs valeurs au pas de temps précédent. Le modèle comprend deux parties distinctes : une *partie dynamique* décrivant les masses d'air (les algorithmes utilisés sont tirés des équations de la mécanique des fluides) et une *partie physique* qui décrit les échanges verticaux entre l'atmosphère et l'espace, et entre l'atmosphère et les surfaces océanique et continentale (échange de rayonnements, de chaleur, de quantité de mouvement, d'eau…). Ces phénomènes se situent à une échelle bien inférieure à la taille de la maille ; ils sont donc traités indirectement, par des paramètres par lesquels est estimé statistiquement leur effet climatique à l'échelle de la maille. Sont ainsi paramétrés des processus physiques, chimiques ou biologiques comme la convection moyenne ou profonde, la turbulence près de la surface, les échanges radiatifs, l'évapotranspiration sur un couvert végétal…



Un maillage à trois dimensions de l'atmosphère.
Une compréhension plus fine des phénomènes mis en jeu nécessite des mailles plus denses dans les simulations.
Images issues d'un film présentant la modélisation du climat.© CEA

Le développement des deux parties du modèle relève donc de tâches assez différentes ! Dans la partie dynamique, il s'agit de formuler des algorithmes de sorte qu'ils puissent être intégrés par ordinateur et soient stables, exacts, maniables : le choix d'algorithmes dépend des contraintes imposées par la capacité de l'ordinateur. Pour la partie physique, les paramétrages relèvent plutôt d'un travail de physicien et varient suivant les processus à représenter. Si leur noyau dynamique est semblable, les modèles de prévision et ceux de climat diffèrent surtout par leurs parties physiques, et donc leurs paramétrages, car les processus d'échanges les plus importants ne sont pas les mêmes à court et à long terme.

L'évolution majeure de la modélisation dans les années 1990 réside dans l'intégration d'un nombre croissant de milieux, d'interaction dans les modèles. Après les modèles couplés océan–atmosphère, on introduit des couplages entre chimie, végétation, hydrologie… La dynamique propre de la modélisation du climat et la croissance exponentielle de la puissance des ordinateurs, en exigeant des projections à long terme, amplifient la tendance.

Redoutables équations de la dynamique des fluides !

Dans le système climatique, l'atmosphère joue un rôle déterminant en distribuant l'eau et la chaleur et donc la géographie des climats. Les équations qui déterminent le mouvement de l'atmosphère, les *équations de Navier–Stokes avec force de Coriolis*, ont été formulées au XIXe siècle par Henri Navier et George Stokes… entre autres, dont Adhémar Barré de Saint-Venant et Gaspard-Gustave Coriolis, qui ont joué un rôle central. Elles font écho aux travaux de Leonhard Euler de 1755 sur la mécanique des fluides. L'étude mathématique de ces équations (mise à prix à un million de dollars par le Clay Mathematics Institute) amène à s'interroger sur leur domaine de validité, sur les solutions à rechercher. Leur non-linéarité traduit que toutes les échelles spatiales et temporelles des écoulements atmosphériques et océaniques interagissent. Ceci oblige, comme le souligne le climatologue français Hervé Le Treut, à pondérer la force brute des ordinateurs par une bonne intelligence de ce que l'on cherche à simuler.

Les équations de Navier–Stokes, comme celle d'Euler, restent mal comprises. Un résultat mathématique étonnant établi par Vladimir Scheffer et Alexander Shnirelman dans les années 1990 énonce ainsi qu'un fluide idéal initialement au repos peut, tout à coup, s'agiter spontanément, sans qu'aucune force n'ait été exercée sur lui ! Après quoi il revient au repos, violant outrageusement le principe de conservation de l'énergie. Ce résultat incroyable pousse à se demander si le modèle physique est le bon, ou si la notion mathématique de solution n'a pas intérêt à être revue…

La question de la validité des équations de Navier–Stokes a également animé la communauté scientifique après plusieurs travaux de Howard Brenner au début des années 2000. Le point central de sa théorie est la notion de *vitesse barycentrique*, qui permettrait de différencier deux types de vitesses (une dite *de volume* et une dite *de travail*), généralisant ainsi les équations de Navier–Stokes compressibles dans les cas fortement hétérogènes. Les équations doivent-elles être changées dans ce cas ? Ce raisonnement est-il important pour la dynamique du climat ? À voir. Mais l'auteur de ce texte avec Benoît Desjardins et Ewelina Zatorska ont montré récemment l'importance du concept de vitesses barycentriques pour les équations de Navier–Stokes compressibles avec viscosité dépendant de la densité : une hydrodynamique à plusieurs vitesses semblerait encodée dans un modèle qui, pourtant, ne semble présenter qu'une unique vitesse !

Nous ne sommes donc qu'au début de l'histoire de notre compréhension des équations de la mécanique des fluides.

Un étonnant phénomène d'amortissement des perturbations

La stabilité des solutions, et donc des écoulements qu'elles modélisent, est un champ d'analyse mathématique incessant depuis les travaux de Lord Kelvin et Lord Rayleigh à la fin du XIX^{e} siècle. La stabilité du *flot de Couette* – écoulement dont le profil des vitesses croît linéairement avec la hauteur – est un problème majeur du domaine. Récemment, Jacob Bedrossian et Nader Masmoudi ont prouvé la stabilité au sens de Lyapunov du flot de Couette. Ils ont en outre mis en évidence un phénomène d'amortissement des perturbations similaire à l'amortissement Landau en physique des plasmas dû à Clément Mouhot et Cédric Villani (ce qui a en partie valu la médaille Fields à ce dernier).

Que se passe-t-il pour des équations plus complexes et pour des solutions de base plus générales ? Selon certains physiciens, il y a amortissement et stabilité dans le cas de cisaillements plus généraux. Un travail à l'interface des mathématiques et de la physique permettra d'aborder des configurations plus réalistes pour une meilleure compréhension en météorologie par exemple.

D.B.

Pour en savoir (un peu) plus :

Brèves de maths : mathématiques de la planète Terre. Collectif, Nouveau Monde, 2014.

Changement climatique : dynamiques scientifiques, expertise, enjeux géopolitiques. Amy Dahan-Dalmedico et Hélène Guillemot, *Sociologie du travail* 48, 2006.

Le site Internet « Un jour, une brève (mathématiques de la planète Terre) » : **breves-de-maths.fr**

Le site Internet de l'atelier réflexion prospective MathsInTerre : **mathsmonde.math.cnrs.fr**

« How Math Helps Structuring Climate Discussions. » Conférence en anglais de Rupert Klein à Berlin (Allemagne) le 17 décembre 2013, 45 minutes, disponible en ligne.

Les maths dans la société : quelles perspectives pour les jeunes, filles et garçons

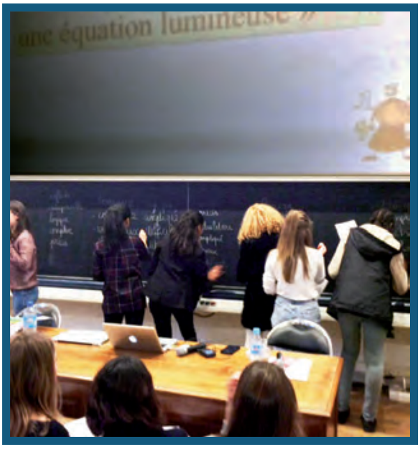
Véronique Slovacek-Chauveau
et Annick Boisseau

Vice-présidente et secrétaire
de l'association Femmes et Mathématiques

« Oh moi, j'ai toujours été nul en maths ! »

Cette réaction fréquente de la part d'adultes de tous milieux, en particulier politique et médiatique, est, de plus, souvent affirmée avec fierté. Les mathématiques sont la seule discipline à provoquer de tels propos, de façon aussi systématique et péremptoire. Pour beaucoup de parents, les mathématiques ne sont qu'un instrument de sélection, souvent associé à des souvenirs douloureux. Lorsque nous interrogeons des élèves de lycée sur les idées qu'ils et elles se font des mathématiques et des personnes qui en font, « calculs », « logique », « complexe », « raisonnement », « intelligent », « difficile » sont les mots revenant le plus souvent. Nous n'obtenons que très rarement les mots « intuition », « beauté », « plaisir », « utiles », « indispensables », « omniprésentes », pourtant volontiers utilisés par les enseignant.e.s et les chercheur.e.s.

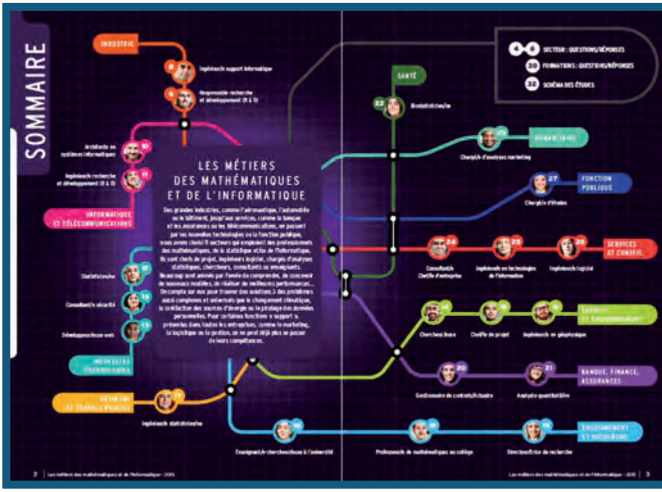
Pour le grand public et la majorité des élèves, les mathématiques sont mortes, tout a déjà été trouvé… et les études dans ce domaine ne débouchent que sur le métier de professeur ou de chercheur.



Pour les personnes proches du milieu scientifique et les élèves intéressé.e.s, les mathématiques n'ont jamais été aussi foisonnantes, ni autant en interaction avec les autres sciences. Les métiers auxquels elles conduisent se diversifient et les besoins augmentent. Communiquer sur les mathématiques est indispensable !

« Les maths, à quoi ça sert ? »

Éternelle question, un peu provocatrice mais légitime, posée par les élèves. Les sociétés savantes et l'ONISEP y ont répondu dans le *Zoom sur les métiers des mathématiques* de 2007 : les mathématiques constituent un instrument irremplaçable de formation à la rigueur et au raisonnement ; elles participent au développement de l'intuition, de l'imagination, de l'esprit critique, voire du rêve ! Elles sont aussi un langage universel, et un élément fort de la culture, indispensable aux futur.e.s citoyennes et citoyens.
En outre, par leurs interactions avec les autres sciences et par leur utilité à décrire, expliquer, prévoir des phénomènes complexes dans la nature et dans le monde technologique, elles jouent un rôle grandissant dans notre vie quotidienne.



Zoom sur les métiers des mathématiques et de l'informatique, ONISEP, 2015.

Les mathématiques sont partout… mais elles sont invisibles

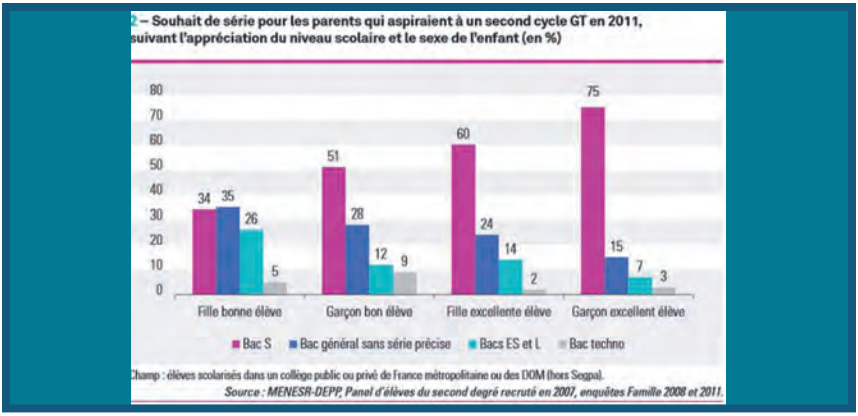
Dans son rapport publié en 2012, l'UNESCO déclare : *« Il est unanimement reconnu que les mathématiques sont omniprésentes dans le monde actuel, notamment dans les objets technologiques qui nous entourent ou dans les processus d'échange et de communication, mais elles le sont généralement de façon invisible. Cette invisibilité rend problématique la perception de l'intérêt de développer une culture mathématique, au-delà des apprentissages les plus basiques, concernant nombres, mesures et calcul. Il est important que la scolarité de base contribue à lever cette invisibilité, et ce d'autant plus que les besoins actuels de ce que l'on appelle la* littératie mathématique *vont bien au-delà des besoins traditionnellement associés au* savoir compter. *»*

Les mathématiques : une matière appréciée à l'école primaire

La plupart des jeunes enfants ont soif de connaissances et prennent plaisir à découvrir les nombres et les figures géométriques, à résoudre des énigmes, à se poser des questions et réfléchir sur des sujets de plus en plus complexes. Différentes études montrent que les mathématiques sont une matière appréciée par les élèves du primaire. Michèle Artigue (professeure émérite didacticienne des mathématiques) décrit en 2013 *« une érosion sensible de l'intérêt pour cette discipline au fil de la scolarité secondaire, mais pas forcément plus que pour d'autres disciplines, et peut-être plus encore (…) une érosion de la confiance des élèves dans leur capacité à continuer à réussir dans cette discipline ».*

C'est au moment des premiers choix d'orientation, à l'issue de la classe de troisième, que commencent à apparaître des différences de comportement et des blocages. Ces différences s'accentuent ensuite à chaque étape où des choix d'options sont à effectuer et s'accroissent à chaque palier d'orientation au lycée, puis dans les études supérieures.

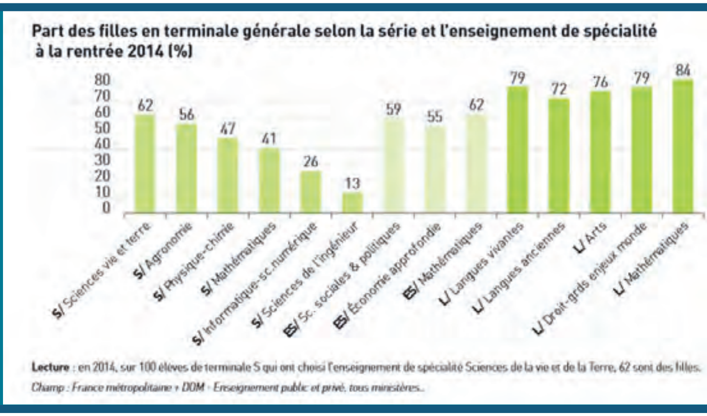
Érosion de la confiance en soi face à cette discipline, impact des idées reçues sur la discipline, pressions de l'entourage (parents, pairs…), influence du milieu scolaire, projection dans un futur rôle social, se conjuguent et conduisent à cette baisse de l'intérêt pour les mathématiques.



Deux facteurs interviennent de façon spécifique : l'origine sociale et géographique et le sexe. Selon Erick Roser, inspecteur général de l'Éducation nationale, dans son rapport de 2015 : *« la probabilité pour un élève de troisième d'accéder à une première scientifique ou technique peut varier de 12 % à 32 % selon les départements. (…) Les orientations en classes préparatoires vont du simple au double selon les territoires ».*

D'autre part, il met en évidence *« une surreprésentation marquée des catégories favorisées dans les classes terminales et les études supérieures scientifiques. (…) On constate que 24,4 % des élèves issus de PCS [professions et catégories socioprofessionnelles] défavorisées accèdent à une première S quand c'est le cas pour 47,1 % des élèves issus de PCS très favorisées »*.

Quant à la mixité, elle est rarement atteinte. Par exemple, en terminale S, la proportion de filles n'a encore jamais dépassé 46 % (mais 48 % des mentions Bien ou Très Bien au bac), avec une répartition très inégale selon les spécialités.



À souligner aussi, *« pour la série S, le pourcentage d'orientation des garçons est, dans tous les milieux, supérieur d'environ 9 % à celui des filles »*.
L'écart se creuse encore dans le supérieur avec 24 % de filles en classes préparatoires scientifiques, 25 % en sciences fondamentales à l'université et 30 % en écoles d'ingénieurs (contre 16 % il y a trente ans).

Pourquoi conseiller à des jeunes de se lancer dans des études de mathématiques ou scientifiques

Comme le sport ou la musique, la pratique des mathématiques peut procurer de belles satisfactions personnelles. La diversité des applications des mathématiques dans tous les domaines (industrie, médecine, arts, écologie, *etc.*) permet d'allier des compétences multiples et de se réaliser selon ses goûts.
Le développement des nouvelles technologies conduit à la création de nouveaux métiers et ouvre des perspectives d'emploi nombreuses et variées, avec une rémunération attractive.

La difficulté supposée de ce type d'études n'a rien d'exceptionnel si on la compare à d'autres formations, comme médecine par exemple, qui est très prisée, surtout par les filles, malgré un taux de réussite en PACES (première année commune aux études de santé) extrêmement faible.

Depuis une vingtaine d'années, du DUT (diplômes universitaires de technologie) au doctorat, les formations se sont largement diversifiées et adaptées au niveau des étudiant.e.s ainsi qu'aux enjeux de la société.

« Plus il y a de maths, moins il y a de filles ! »

C'est ainsi que Martin Andler (président de l'association Animath) résume la situation. La généralisation de la mixité à l'école en 1975 n'a pas conduit spontanément à l'égalité.

C'est pourquoi l'association Femmes et Mathématiques organise des actions en direction des collégiennes et des lycéennes, mais aussi des étudiantes et des doctorantes.

Les journées *« Filles et maths : une équation lumineuse »*, organisées dans des universités rassemblant chacune une centaine de filles et le *« Forum des jeunes mathématiciennes »* annuel, rencontrent un succès grandissant.

Mathématicienne : une espèce en voie de disparition

Un comble : non seulement la proportion de femmes parmi les chercheurs et enseignants-chercheurs n'augmente plus, mais elle diminue depuis quelques années.

En France, seules 18 % des maîtres de conférences et 6 % des professeurs des universités en mathématiques fondamentales sont des femmes. *« Le nombre de mathématiciennes est très faible, cela contribue à véhiculer l'idée fausse selon laquelle les mathématiques ne seraient pas une activité adaptée aux filles »* regrette Laurence Broze (présidente de Femmes et Mathématiques).

Et pourtant elles existent…

La France compte de brillantes mathématiciennes. Au dernier Congrès international des mathématiciens, à Séoul en 2014, la délégation française comptait sept mathématiciennes (sur quarante et un membres) : Viviane Baladi, Zoé Chatzidakis, Isabelle Gallagher, Monique Laurent, Sandrine Péché, Laure Saint-Raymond, Michela Varagnolo. C'est lors de ce congrès que la prestigieuse médaille Fields a été attribuée pour la première fois à une femme, l'Iranienne Maryam Mirzakhani. Plus près de nous, la mathématicienne Claire Voisin vient d'être nommée professeure au Collège de France ; sa leçon inaugurale aura lieu le jeudi 2 juin 2016 à 18 h au sein de la vénérable institution du Quartier latin.

Jeunes, filles et garçons, les mathématiques vous intéressent ? Elles ont besoin de vous, alors n'hésitez pas : foncez !

V. S.-C. & A.B.

De haut en bas et de gauche à droite :

Nalini Anantharaman, Valérie Berthé, Virginie Bonnaillie-Noël, Nadia Brauner Vettier, Laurence Broze, Irène Marcovici, Maryam Mirzakhani, Natacha Portier, Nathalie Revol, Marie-Françoise Roy, Sylvia Serfaty, Laure Saint-Raymond, Camille Ternynck.

Quand les maths débarquent dans les tribunaux

Leila Schneps

Institut de mathématiques de Jussieu

Les maths débarquent dans les tribunaux ! Grâce aux nouvelles techniques d'analyse scientifique de tous les types de traces laissées sur une scène de crime, elles constituent d'ores et déjà un élément incontournable des procès.

Les résultats des expertises sont de nature probabiliste !

Pourquoi les maths, alors que ce sont entre autres la chimie, la biologie, la génétique, les sciences de l'ingénieur qui permettent de réaliser les analyses des indices matériels ? Parce que la plupart du temps ces autres sciences fourniront des résultats de type probabiliste. Une empreinte pourra être identifiée comme provenant d'un certain type de chaussure ; au vu des traces d'usure, on pourra même effectuer une comparaison avec la botte du suspect, mais là aussi, à moins d'avoir une trace exceptionnellement détaillée, la comparaison ne donnera en fin de compte qu'une probabilité d'identité.

Même les analyses ADN, réputées pour produire des résultats d'identification au-delà de tout doute possible, peuvent donner lieu à de nombreuses difficultés dans les cas d'échantillons mixtes, dégradés ou disponibles en quantité très réduite. Les gendarmes affectés aux analyses en laboratoire des échantillons d'ADN provenant de scènes de crime l'affirment : le résultat est suffisamment clair pour indiquer un coupable au-delà de tout doute raisonnable dans moins d'un cas sur dix. Vous vous en souvenez peut-être : de l'attentat d'Omagh de 1998 en Irlande à l'accusation de meurtre lancée en 2007 contre la jeune Américaine Amanda Knox, les analyses de faibles traces d'ADN ont donné lieu à des débats acharnés entre experts incapables de se mettre d'accord.

Il est donc fréquent qu'un fait concret, scientifiquement établi, apporte un renseignement sur la culpabilité éventuelle d'un suspect, sans que ni les experts scientifiques, ni le juge, ni les avocats, ni les membres du jury, sans parler de l'accusé lui-même ou du public, sachent véritablement expliquer la portée dudit renseignement. Oui, le mélange d'ADN trouvé sur le vêtement du cadavre contient certains allèles en commun avec ceux du présumé meurtrier. Oui, l'écriture de l'accusé a des points communs avec celle de la lettre de menaces. Mais que peut-on en conclure ? Quel est le véritable poids de ces preuves, à charge ou à décharge de l'accusé ?

Bien comprendre la notion d'évènements indépendants

Le système judiciaire s'attend à ce que les membres du jury se fient à leur seul instinct. Ce système, en place depuis des siècles, vise à contrebalancer les effets d'ignorance et de préjugés des individus en les plaçant au sein d'un groupe hétérogène, afin de parvenir à un jugement le plus objectif possible. Mais ce système présente un problème dans les cas, de plus en plus fréquents, où une thèse (accusation ou défense) utilise des arguments basés sur des probabilités. Les probabilités sont souvent terriblement contre-intuitives, et les penser correctement est un exercice subtil qui nécessite un certain entraînement.

Voici un exemple frappant pour vous en convaincre. Pour déterminer la probabilité que deux évènements se produisent, sachant la probabilité de chacun d'entre eux, il ne faut multiplier les deux probabilités que lorsque les deux évènements sont totalement indépendants. Le sexe d'un enfant à naître est un évènement indépendant de toute autre naissance (excluons le cas de jumeaux identiques, ou toute autre particularité génétique qui peut influer sur cette question). Disposant de cette information, proposons l'énigme suivante : vous apprenez par hasard que votre interlocuteur a deux enfants, dont un garçon. Quelle est la probabilité que l'autre enfant soit une fille ?

Si vous répondez immédiatement *« Une chance sur deux ! »*, vous avez tort. En effet, les parents de deux enfants sont équitablement répartis en quatre groupes : GG, GF, FG et FF (G désigne un garçon, F une fille). Si votre interlocuteur a un garçon, c'est qu'il appartient aux 75 % de la population représentés par les groupes GG, GF et FG. Il y a donc deux chances sur trois (et non une chance sur deux) que l'autre enfant soit une fille !

Notre intuition nous fourvoie dès qu'il est question de probabilités. Dans le cadre d'un procès, souvent tout le monde se fourvoie dans le même sens… Que de condamnations parce que le jury, aidé seulement par son intuition, s'est laissé convaincre que deux bébés dans une même famille ne pouvaient pas être tous les deux victimes de mort subite du nourrisson ! ou que telle infirmière ne pouvait pas avoir été purement par hasard présente à toutes les morts au sein de l'hôpital où elle travaillait !

Inversement, que de personnes acquittées parce que le jury avait trop de mal à incarner, dans l'être humain assis devant eux, une culpabilité indiquée non pas par des faits établis, mais par une puissante improbabilité théorique d'innocence ! On pense à cet homme, accusé du crime d'avoir tiré depuis sa voiture sur deux jeunes, mais acquitté malgré la lourde preuve à charge que constituait un taux de poudre résiduelle de tir deux cents fois plus élevé que celui trouvé typiquement dans une voiture de chasseur, explication présentée par la défense pour justifier la présence de ces résidus. Ou alors à cette jeune mère américaine acquittée du meurtre de sa fille de 3 ans : elle avait omis de prévenir la police pendant un mois après la disparition de sa fille, abreuvant ses parents, ses amis et la police d'un flot de mensonges, facilement démontés ; en outre, son propre père, ex-policier, avait témoigné avoir senti une forte odeur de cadavre décomposé dans le coffre de la voiture de sa fille. La raison de l'acquittement ? La reconstitution du meurtre présentée par l'accusation contenait des éléments mineurs non démontrés par des faits avérés, mais justifiés par un raisonnement probabiliste, qui laissait aux yeux du jury la place à un doute raisonnable.

La science face au doute : le raisonnement bayésien

Une méthode existe pour incorporer des analyses scientifiques qui ne donnent qu'un renseignement probabiliste au sujet d'un crime particulier dans la pensée *a priori* intuitive d'un membre du jury. Elle s'appelle le *raisonnement bayésien*, et se réalise mathématiquement au moyen de réseaux bayésiens. On trouve des logiciels de réseaux bayésiens très performants, destinés à être utilisés plutôt pendant la phase de l'instruction ou au cours du travail d'un expert avec les avocats que devant le jury. Pour le moment, ils sont peu utilisés car ils soulèvent surtout de l'inquiétude et des soupçons de la part des juristes.

Comment ça marche ? Essentiellement, la méthode consiste à proposer une estimation de probabilité subjective de culpabilité de façon traditionnelle, intuitive, basée sur la présentation des preuves non scientifiques, puis d'utiliser la formule de Bayes pour incorporer une preuve de nature purement probabiliste dans cette estimation, de façon à la « mettre à jour ».

Thomas Bayes
(1702–1761).

La version de la *formule de Bayes* appliquée le plus souvent aux preuves scientifiques lors d'un procès est la dernière formule proposée en encadré. Pour comprendre son application dans des situations vraies, il faut considérer que l'évènement C correspond à « L'accusé est coupable », alors que l'évènement T correspond à « Cette trace précise a été trouvée sur la scène du crime ». T peut indiquer, par exemple, la présence sur la scène du crime d'une fibre de laine ou de coton, d'une trace de sang d'un groupe sanguin donné, d'une empreinte digitale avec certaines caractéristiques particulières, une trace de chaussure ayant telles particularités qui permettent d'en identifier la taille et la marque, ou alors d'un échantillon d'ADN.

En général, l'accusé correspondra à la trace laissée, car sinon – s'il possèdait un autre groupe sanguin ou un ADN différent ou une taille de chaussures plus petite… – il ne serait pas l'accusé. La probabilité P(T si C) est donc considérée comme ayant la valeur 1, puisqu'elle mesure la probabilité que si l'accusé est bien le coupable, c'est lui qui aurait laissé la trace. La formule de Bayes se simplifie alors :

$$P(C \text{ si } T) / P(\text{non } C \text{ si } T) = P(C) / [P(T \text{ si non } C) \times (1 - P(C))].$$

La probabilité P(T si non C) mesure la probabilité qu'une trace du type T peut être présente si l'accusé n'est pas le coupable, c'est-à-dire la probabilité de trouver la trace T dans la population générale. C'est un élément purement numérique, indépendant de l'accusé et de tout autre élément de preuve, qui sera apporté par un expert scientifique. La probabilité P(C) représente la probabilité *a priori* de culpabilité de l'accusé, c'est-à-dire une probabilité subjective formulée indépendamment par chaque juré après avoir entendu, lors du procès, toutes les preuves de nature non scientifique. Le fait de prendre cette estimation initiale P(C), subjective, de culpabilité et de la multiplier par le facteur $1 / [P(T \text{ si non } C) \times (1 - P(C))]$ est la méthode bayésienne de « mise à jour ».

Le résultat de cette opération est le quotient P(C si T) / P(non C si T), qui fournit une comparaison entre la probabilité que l'accusé soit coupable étant données (*i*) la présence de la trace T, (*ii*) l'estimation préalable de culpabilité P(C), (*iii*) l'information scientifique, et d'autre part la probabilité que l'accusé soit non coupable étant donnés les même trois éléments. Plus le quotient s'avère grand, plus l'information scientifique s'avérera avoir été à charge de l'accusé. Plus le quotient est petit (inférieur à 1), plus l'apport scientifique est à sa décharge.

Du bon usage de la formule de Bayes

La formule de Bayes est dérivée d'une identité simple concernant la probabilité que deux évènements, T et C, se produisent en même temps. Cette probabilité s'écrit P(C et T), et est égale au produit P(C si T) × P(T), où P(C si T) indique la probabilité que C survient étant donné que T est survenu, et P(T) est la probabilité de l'évènement T. Comme T et C jouent un rôle symétrique dans la probabilité P(C et T), nous avons également la formule suivante :

P(C et T) = P(T si C) × P(C).

On en déduit l'égalité P(C si T) × P(T) = P(T si C) × P(C), ou encore, en divisant les deux côtés par P(T) :

Formule de Bayes : P(C si T) = P(T si C) × P(C) / P(T).

Si on note « non C » l'évènement opposé à C, soit « C n'est pas survenu », la probabilité P(non C) est égale à 1 – P(C). Il peut être intéressant de comparer P(C) à P(non C) au moyen du quotient P(C) / P(non C). Si ce quotient est supérieur à 1, c'est que C est plus probable que non C, et *vice versa* si le quotient est inférieur à 1.

Si l'on applique maintenant la formule de Bayes aux évènements T et non C, on obtient : P(non C si T) = P(T si non C) × P(non C) / P(T), ce qui permet de calculer le quotient :

P(C si T) / P(non C si T) = P(T si C) × P(C) / [P(T si non C) × (1 – P(C))].

La place des probabilités dans l'appareil judiciaire

Cela peut paraître purement théorique, mais cette formule est d'ores et déjà utilisée dans de nombreux procès ! Voyons un cas réellement jugé. Rentrant d'une sortie nocturne avec des amis, une jeune femme, M^lle^ M., est abordée par un homme qui lui demande l'heure, puis l'attaque et la viole. La police obtient un échantillon de l'ADN de l'agresseur suffisamment précis pour n'appartenir qu'à un individu sur deux cents millions. Les enquêteurs obtiennent de plus les renseignements suivants de la part de M^lle^ M. : l'agresseur était de race blanche, il avait environ 25 ans et parlait avec l'accent local. Elle a assuré pouvoir le reconnaître facilement si jamais on le retrouvait.

Dans un premier temps, il fut impossible de mettre la main sur un suspect plausible. Au bout de deux ou trois ans, une correspondance fut trouvée entre l'échantillon d'ADN et un autre échantillon, ajouté récemment dans une base de données. L'individu fut donc arrêté, mais M^lle^ M. ne l'a pas identifié, soutenant qu'il était beaucoup trop vieux (il avait 37 ans) et ne ressemblait pas à l'homme dont elle se souvenait. De plus, il avait un alibi, fourni par son amie, qui soutenait avoir passé la nuit en question avec lui.

Le jury s'est donc trouvé confronté à un ensemble de preuves allant toutes dans le sens de la défense : la non-identification, l'erreur sur l'âge, et l'alibi. La défense a présenté un témoin expert, qui a expliqué le théorème de Bayes au jury de la manière suivante. On commence par assigner des probabilités subjectives à toutes les preuves non scientifiques, en calculant le quotient de probabilité à chaque fois. Ainsi,

P(ne pas reconnaître si coupable) / P(ne pas reconnaître si innocent) = 1 / 10,
P(avoir un alibi si coupable) / P(avoir un alibi si innocent) = 1 / 2,
P(avoir accent local si coupable) / P(avoir accent local si innocent) = 1 / 200 000.

L'estimation *a priori*, subjective, de culpabilité peut être calculée : P(C) / P(non C) = 1 / 4 000 000, une présomption très forte d'innocence. On incorpore maintenant le quotient suivant :
P(ADN trouvé si coupable) / P(ADN trouvé si innocent) = 200 000 000.
En multipliant l'*a priori* par ce nombre, on obtient pour P(C) / P(non C) la nouvelle estimation 2 000 / 40, qui est égale à 50. L'avocat de la défense a argué qu'une chance de culpabilité cinquante fois plus grande que la chance d'innocence revient à une chance d'innocence d'à peu près P(non C) = 2 %, qu'il interprétait comme un doute raisonnable. Au vu des 98 % de chance d'être coupable, le jury a tout de même condamné l'accusé, mais le procès a été annulé en appel… sous prétexte que l'on ne peut pas donner un cours de maths au jury pour leur dire comment penser !

Si la science apporte de précieux renseignements, les probabilités peuvent fournir une méthode fiable pour les incorporer dans un raisonnement par ailleurs non scientifique. Mais elles sont rarement bien comprises, jusqu'à être parfois interdites de séjour dans les prétoires. Il reste à trouver la juste place de la théorie des probabilités au sein de l'appareil judiciaire.

L.S.

Pour en savoir (un peu) plus :

Les maths au tribunal. Quand les erreurs de calcul font les erreurs judiciaires.
Coralie Colmez et Leila Schneps, Le Seuil, 2015.

Mathématiques électorales : vers plus de démocratie

Karel Janeček

Mathématicien et entrepreneur

Est-il possible de concevoir un système de vote qui garantirait l’égalité des suffrages exprimés et qui, dans le même temps, donnerait plus de poids aux votants informés et engagés ? Contre toute attente, c’est possible. Le système de vote Democracy 2.1, à l’initiative de l’auteur et qui repose sur les mathématiques, résout ce paradoxe apparent. En outre, il permet d’assainir le processus de vote.

« La démocratie est le pire système de gouvernement, à l’exception de tous les autres qui ont pu être expérimentés dans l’histoire » affirmait déjà Winston Churchill à Londres le 11 novembre 1947 devant la chambre des Communes. Sommes-nous condamnés à voter pour le candidat « le moins pire », à ne pas être entendus ?

Exprimer plus par son vote que sa préférence pour un parti

Sur le long terme, tout système autre que démocratique semble non viable. En effet, dès que le principe d’égalité des suffrages est transgressé, le risque est grand de voir le système social dégénérer en une forme de dictature, ou devenir discriminatoire.

Pour un État, disposer d’une représentation politique de qualité est un élément clé vers un bon fonctionnement de la société. Si des systèmes de vote démocratiques existent, le principe de l’égalité des suffrages à lui seul conduit aux dangers du populisme et des extrémismes. Comment se débarrasser de la manipulations de l’opinion, du contrôle des masses ?

L’engagement citoyen dans la cité, s’il demande du temps, de l’énergie, de l’investissement personnel, est à encourager. Imaginons que ces citoyens engagés et informés, qui orienteront leurs votes vers des candidats de qualité, soient mieux entendus par nos représentants politiques.

Imaginons que le citoyen puisse exprimer plus, par son vote, que sa préférence pour un parti ou un candidat. Chaque citoyen sera motivé pour s'informer, débattre de la question politique et sociale, se faire entendre. L'engagement politique retrouverait son sens, et une participation accrue de la population aux élections s'ensuivrait mécaniquement.

Quand la théorie des jeux réinvente la politique

De tels systèmes de vote sont étudiés depuis longtemps. L'un d'entre eux, le système de vote à la semi-proporionnelle Democracy 2.1, a été déployé dans différents contextes en Tunisie, en République tchèque, au Portugal, à New York, en Chine… Illustrons-en le fonctionnement sur un scrutin local fictif typique : deux mandats sont à pourvoir dans la circonscription, chaque parti en lice peut présenter un ou deux candidats, des candidats indépendants peuvent se présenter (s'ils réunissent suffisament de parrainages ou de signatures par exemple). Chaque électeur dispose pour sa part de quatre « votes d'adhésion » et de deux « votes sanctions », qu'il peut distribuer à sa guise à chacun des candidats. Il ne peut pas attribuer plus d'un vote à un candidat donné, comme il peut choisir de ne pas faire usage de toutes ses cartouches.

Dans un premier temps, regardons ce qu'il se passe dans le cas, simplifié, d'un seul mandat à pourvoir et deux votes d'adhésion disponibles par candidat (mais pas de vote sanction). Cela suffira à illustrer les mécanismes du système.

Les candidats en présence sont A, présenté par un parti populiste de droite, crédité de 20 % des intentions de vote ; B et C, deux candidats classés à droite, crédité chacun de 12 % des intentions de vote ; D et E, deux candidats classés à gauche, crédité chacun de 12 % des intentions de vote ; et F, présenté par un parti populiste de gauche, crédité de 20 % des intentions de vote.

Dans un scrutin uninominal majoritaire à un tour, si les données sont fiables, l'un des candidats extrémistes l'emporterait. Dans un scrutin à deux tours, ils seraient tous deux qualifiés pour le second tour. Avec un système de type Democracy 2.1, chaque votant dispose d'un second vote d'adhésion. Un électeur populiste votant A (et de même pour F) donnerait son second vote à B ou C (respectivement D ou E), ou n'en ferait pas usage du tout. Un partisan de B ou C attribuerait son second vote à l'autre candidat de droite, plus rarement à D ou E, encore plus rarement à A. De même, un partisan de D ou E attribuerait son second vote à l'autre candidat de gauche, plus rarement à B ou C, encore plus rarement à F.

Quelle serait l'issue du scrutin ? Le populiste de droite, A, obtiendrait un peu plus de 20 % de l'ensemble des suffrages exprimés. De même pour celui de gauche, F. Les deux candidats de droite B et C obtiendraient chacun plus de 24 % des voix. De même pour les deux candidats de gauche D et E. Les résultats des candidats populistes sont mécaniquement relativisés. En outre, les candidats de droite B et C, tout comme les candidats de gauche D et E, sont moins dans une rivalité qui favoriserait la montée des populismes. Ils peuvent se focaliser sur les détails qui caractérisent et distinguent leur programme politique.

Votes sanctions et votes multiples pour assainir la politique

Revenons à notre situation, plus réaliste, avec deux mandats à pourvoir, quatre votes d'adhésion et deux votes sanctions. Ces derniers permettent à l'électeur de manifester son désir de désavouer jusqu'à deux candidats. Ce simple aspect devrait permettre, à lui seul, de faire retrouver le chemin des urnes à nombre d'abstentionnistes, d'une part, et de modérer l'influence des partis populistes ou extrémistes, d'autre part. Les votes sanctions permettent de lutter contre la corruption et d'assainir l'appareil politique. Quant aux votes d'adhésion, s'ils sont plus nombreux que le nombre de sièges à pourvoir, ils favoriseront les candidats ouverts au consensus.

Pour l'illustrer, prenons un nouvel exemple illustratif avec un unique mandat à pourvoir, deux partis politiques en présence, quatre votes d'adhésion et deux votes sanctions. Les deux partis représentés sont la droite (D) et la gauche (G). Chacun propose deux candidats, un corrompu

Dynamiser les élections législatives

Une faiblesse des scrutins majoritaires est leur aptitude limitée à représenter, à travers un unique candidat, la diversité des suffrages exprimés. Une piste serait d'élire plus d'un candidat par scrutin ! Et si on pouvait élire deux candidats ?
On aurait un siège supplémentaire, et plus de voix prises en compte. Cela favoriserait les candidats indépendants, les « petits » partis : des idées politiques nouvelles émergeraient plus facilement.

En outre, les campagnes de dénigrement entre candidats, qui polluent le débat et font primer les jeux politiques sur les enjeux, auraient moins d'impact, et seraient peut-être monnaie moins courante. Un parlement ou une assemblée législative élue selon un tel système (que Democracy 2.1 autorise) comprendrait plus de diversité. Et un parti déjà très populaire serait sans doute encore plus fort que dans un système à la proportionnelle.

Quels partis seraient alors affaiblis par un tel système ? Ceux de taille moyenne, qui attirent peu d'électeurs. L'environnement électoral, propice à la collaboration, favoriserait la formation d'un gouvernement stable.

et un honnête (D^- et D^+ respectivement pour la droite, G^- et G^+ pour la gauche). Dans un monde idéal, l'honnêteté des candidats serait un critère décisif pour l'électeur, des votes sanctions seraient attribués aux candidats corrompus D^- et G^- et les suffrages iraient à D^+ et G^+.

Regardons maintenant comment se comporte notre modèle avec le pire scénario, celui où les partisans de chaque parti plébisciteraient leur candidat corrompu au détriment du candidat honnête de l'autre parti. Le vote tactique des électeurs consiste à attribuer un vote sanction au candidat honnête du parti opposé, et non au candidat corrompu de son propre parti. Eh bien, même dans ce scénario, le vote sanction assainit le système ! Les candidats honnêtes sont élus dès lors que le résultat est serré entre la droite et la gauche. Plus précisément, D^+ et G^+ emportent la mise dès que le rapport entre le nombre n_D d'électeurs de droite et le nombre n_G d'électeurs de gauche est supérieur à 3 / 4 et inférieur à 4 / 3.

Dans le pire scénario, si l'écart entre les deux partis est « faible » (le ratio entre les nombres d'électeurs des deux partis se situe dans l'intervalle [3 / 4, 4 / 3]), le vote sanction permet d'éviter l'élection des candidats corrompus. Dans tout autre scénario, l'effet assainissant du vote sanction apparaît dès que le ratio entre les nombres d'électeurs des deux partis se situe dans l'intervalle [1, 2]. Saurez-vous établir ces résultats ? Les raisonnements utilisés sont typiques de la *théorie des jeux*, dont les mathématiques électorales font partie.

Sans vote sanction disponible, les deux candidats du parti politique qui dispose du plus grand nombre de partisans sont élus. Avec prise en compte du vote sanction, les deux candidats honnêtes sont élus dès lors que $n_D / (n_D+n_G)$ appartient à la fourchette [3 / 7, 4 / 7] ou, ce qui revient strictement au même, dès lors que $n_G / (n_D + n_G)$ appartient à cette même fourchette. On mesure ainsi, quantitativement, l'effet assainissant du vote sanction !

Le théorème d'Arrow : choix social et inévitables paradoxes

Les élections dans la vie réelle sont plus sophistiquées que ces modèles simplifiés. Néanmoins, les mathématiques sont riches d'enseignements même dans les situations les plus complexes. Ainsi, le *théorème d'impossibilité d'Arrow* établit que, dès que quatre hypothèses « naturelles » sont satisfaites, aucun système de vote ne peut agréger les préférences individuelles dans un classement global à l'échelle de la collectivité. Les quatre hypothèses sont les suivantes :

- *Universalité* : pour chaque individu, tous les profils de préférences sont autorisés (tous les classements possibles des candidats doivent être autorisés pour chacun des électeurs) ;
- *Non-dictature* : aucun individu ne peut imposer, indépendamment du choix des autres électeurs, son propre classement des candidats ;
- *Unanimité* : lorsque tous les électeurs préfèrent un candidat donné à un autre, le classement global des candidats doit refléter cette préférence (autrement dit, il n'est pas possible de favoriser un candidat dans le classement global sans en défavoriser un autre) ;
- *Indépendance des options non pertinentes* : la préférence d'un candidat par rapport à un autre dans le classement global ne doit pas dépendre de l'existence d'autres candidats (ajouter un candidat fantôme pour lequel personne ne vote ne change pas le résultat du scrutin).

Quatre hypothèses « de bon sens », mais hélas dès que trois candidats sont en lice avec deux électeurs le théorème d'impossibilité d'Arrow s'applique, et nul système de vote ne peut satisfaire les quatre hypothèses. Il n'existe aucun processus de choix qui permette d'exprimer une hiérarchie des préférences qui soit valide pour une collectivité à partir de l'agrégation des préférences individuelles exprimées par chacun des membres de cette collectivité. Et il n'existera jamais de tel processus ; c'est l'objet du théorème démontré par Kenneth Arrow en 1951. Le résultat s'applique à tous les types d'élections et de scrutins ; il est typique des *mathématiques de la décision*. Il n'est donc pas possible de trouver un système de vote « parfait » : dès que les quatre hypothèses précédentes sont satisfaites, le théorème d'Arrow s'applique. On peut dès lors étudier d'autres critères de choix social. Un critère, très restrictif mais naturel, qui vient à l'esprit est celui du gagnant de Condorcet,

Le paradoxe de Condorcet

Marie Jean Antoine Nicolas de Caritat, marquis de Condorcet, est un célèbre représentant des Lumières. Il a mis en valeur le paradoxe qui porte son nom : prenez une assemblée ayant le choix entre trois candidats, A, B et C. Imaginons que 40 % des électeurs préfèrent A à B et B à C (écrivons A > B > C) ; que 35 % préfèrent B à C et C à A (B > C > A) ; et que 25 % préfèrent C à A et A à B (C > A > B). Alors, quel que soit le mode de scrutin utilisé pour désigner le vainqueur, une majorité des électeurs sera déçue ! Par exemple, si B l'emporte, alors le groupe de 40 % et le groupe de 25 % (soit 65 % de l'assemblée) préféreraient élire A. Et de même pour toutes les situations possibles : exercez-vous à vous en convaincre ! On retrouve cette situation paradoxale dans le jeu pierre–feuille–ciseaux : la feuille l'emporte sur la pierre, qui bat les ciseaux, qui gagnent sur la feuille… Le théorème d'impossibilité d'Arrow montre que le problème est en fait lié aux difficultés de l'agrégation des préférences.

selon lequel si un candidat, confronté en duel à tous les autres, sort toujours vainqueur, alors ce candidat (nommé « gagnant de Condorcet ») doit être élu. Le critère du gagnant de Condorcet met fin au paradoxe de Condorcet.

Deux problèmes subsistent cependant : d'une part, l'existence d'un gagnant de Condorcet n'est pas garantie. D'autre part, la plupart des systèmes de vote ne satisfont pas ce critère. C'est le cas en particulier des scrutins uninominaux majoritaires à un ou à deux tours (on a vu leur incapacité à juguler les candidats extrémistes). Democracy 2.1 non plus ne vérifie pas le critère du gagnant de Condorcet : considérons trois candidats A, B et C. Supposons que pour $n + 1$ électeurs, A est préféré à B, qui lui-même est préféré à C. Pour n autres électeurs, B est préféré à C, qui lui-même est préféré à A. Le gagnant de Condorcet est alors A dès que n est supérieur à 2 (assurez-vous-en !). Or, le système de vote Democracy 2.1 fera émerger B. Le candidat B est appelé un « gagnant de Borda » (voir en encadré). Ce choix est rationnel et légitime, surtout si n est « grand » (à l'échelle d'une collectivité). Ainsi, le critère du gagnant de Condorcet peut être discuté…

Méthode Borda et vote tactique

La *méthode Borda* est un système de vote pondéré : une fois que chaque électeur a classé les candidats par ordre de préférence, chaque candidat reçoit un nombre de points correspondant au nombre de candidats plus mal classés que lui ; est élu celui parmi les candidats qui obtient le plus grand nombre de points. La méthode Borda est un mode de scrutin basé sur le consensus, mais qui est sujet au *vote tactique* : les électeurs d'un candidat assuré de passer le premier tour peuvent avoir intérêt, tactiquement, à classer en premier non pas leur candidat, mais le « moins dangereux » des autres candidats, par exemple. Il est donc possible d'améliorer les chances d'obtenir l'élection d'un candidat en le reculant sur la liste de ses préférences. De même, on peut réduire les chances d'un candidat en l'avançant dans la liste de ses préférences ! Certains systèmes de vote ne sont pas sensibles au vote tactique.

De la même manière, les autres critères de choix social peuvent être analysés, discutés, critiqués. Vous seriez surpris de voir, même sans sortir de l'Union européenne, la diversité et l'originalité des modes de scrutin qui sont utilisés quotidiennement, et parfois depuis des siècles ! De son côté, Democracy 2.1 attend toujours de rencontrer un scrutin pour lequel le résultat n'irait pas dans le sens de la maximisation de l'utilité sociale (le « bon sens »). Ce système est fondé à la fois sur de solides bases théoriques et sur un nombre significatif d'expérimentations en situation réelle. Le XXI^{e} siècle devra composer avec de tels systèmes plus à l'écoute des citoyens !

K.J.

Un mathématicien aux prises avec la société

Édouard Thomas
Mathématicien
et journaliste scientifique

Au XX^{e} siècle émerge la question, jusque-là inédite, de l'éthique du scientifique dans son travail, que ce soit en physique nucléaire, en génie chimique, en biologie moléculaire, en génétique… On assiste en outre à une sophistication du langage mathématique (précision des notions, nouveaux concepts opératoires) permettant de quantifier les données de l'expérience. Que de bouleversements ! Répétition des observations et recours aux statistiques pour gommer les imperfections et aléas sur les phénomènes observés permettent de tendre vers une certaine forme, *faible*, d'objectivité au sein de la communauté scientifique. Plus précisément, le consensus sur une adéquation entre des énoncés et des faits réels tend alors à s'imposer, au détriment de l'objectivité *forte*, idéale, portant sur les choses elles-mêmes, qui a cependant toujours ses adeptes.

Supprimer les dessins : une fausse bonne idée !

Tournons-nous vers les mathématiques « pures », pour lesquelles limitations matérielles et protocoles expérimentaux ne sont généralement pas des entraves aux progrès. L'éthique du chercheur est-elle alors déconnectée des préoccupations sociétales dans lesquelles ces recherches sont conduites ? L'objectivité forte a-t-elle seulement cours ?

Un petit schéma explicatif dans un texte présente le risque d'être perçu comme un cas particulier, une instance d'un phénomène. La subjectivité n'est pas loin… On peut dès lors imaginer supprimer toute figure des ouvrages scientifiques, comme l'a d'ailleurs fait le groupe de mathématiciens Bourbaki, menant à son paroxysme une démarche initiée par David Hilbert en 1899 avec ses *Grundlagen der Geometrie*.

Aujourd'hui encore, il reste impensable de faire reposer une démonstration sur un schéma, un diagramme ou une figure : ces derniers ne sont là que comme supports de l'intuition et ne peuvent en aucun cas se substituer à un raisonnement en bonne et due forme. Ils aident néanmoins à la diffusion des idées et, à ce titre, participent à l'objectivité, faible, des résultats établis. Au cours du XXe siècle se développent ainsi une plus grande rigueur, une exigence logique perçue comme nécessaire et une méfiance accrue vis-à-vis des figures : un tracé, aussi précis soit-il, n'est pas une construction ! L'objectivité forte aurait-elle trouvé un asile ?

Une activité sociale et profondément humaine

L'objectivité forte ferait penser à la validité intangible d'un théorème parfaitement rigoureux, à un raisonnement implacable. Or, le processus de validation d'un résultat nouveau n'est pas définitif et se fait par consensus des experts du domaine, puis par un ensemble de plus en plus élargi de la communauté. On commence à voir poindre un lien entre le mathématicien et le monde qui l'environne.

Examinons en particulier deux questions d'actualité très concrètes. La première est : que faire face au risque qu'une erreur subtile échappe à tous, comme cela a pu être le cas avec des « démonstrations » de théorèmes invalidées, parfois des années après leur publication ?

La seconde question peut se formuler ainsi : comment gérer une proposition de preuve trop longue pour être embrassée par des pairs ou trop complexe pour être formalisée à l'aide d'un assistant de preuve ?

Au-delà des preuves formelles (comme le théorème des quatre couleurs) ou des démonstrations qui nécessitent de longs calculs informatiques (telle la conjecture de Kepler), l'interrogation est bel et bien légitime, comme on peut le voir aujourd'hui avec la preuve toujours en chantier depuis les années 1980 de la classification des groupes finis simples, ou l'intimidante proposition de démonstration de la conjecture ABC par le mathématicien japonais Shinichi Mochizuki en août 2012.

Des risques d'erreur et des preuves trop longues…

Ces deux interrogations suscitent depuis une trentaine d'années des débats enflammés, à tel point que le médaillé Fields Vladimir Voïevodsky a mis un terme à ses recherches en géométrie algébrique pour se lancer dans un projet de refonte des fondements des mathématiques.

Pour trouver une réponse à nos deux questions spécifiques (comment débusquer les erreurs subtiles et comment examiner les preuves complexes), la seule solution semble consister à donner du temps à des professionnels qui ont la déontologie bien chevillée au corps. Ainsi, les grandes lignes de la gigantesque théorie élaborée par Mochizuki ont été comprises par un groupe extrêmement réduit d'experts courageux, qui ont passé des années à en examiner la structure… au détriment de leurs propres recherches. L'heure est aujourd'hui à la consolidation des arguments, à l'inspection de la théorie dans ses moindres rouages, à la validation (on l'espère) et à l'élargissement de la communauté des experts capables et désireux de s'approprier ces nouveaux outils. Viendra ensuite, si tous les voyants sont au vert, la diffusion des idées apportées par le mathématicien japonais.

Une petite histoire de l'objectivité

Au Moyen Âge, une chose existe objectivement quand elle existe en tant que chose connue, en tant qu'objet de l'esprit. En 1641, Descartes écrit dans ses *Méditations métaphysiques* : *« Être objectivement, [c'est] être dans l'entendement en la manière que les objets ont coutume d'y être. »* On découvre en effet que les sens nous trompent, la perception elle-même n'est pas fiable. L'objectivité ne saurait donc être donnée par l'expérience. La pensée de Kant marque un tournant épistémologique. Il rétorque que les perceptions multiples fournies par l'expérience sont organisées et ordonnées par notre entendement. Ce qui est connu n'est pas l'objet par lui-même, mais l'objet tel qu'il est représenté par l'esprit : *« Des principes* a priori *fondent des jugements objectivement valables tout en étant empiriques »* (*Prolégomènes à toute métaphysique future qui pourra se présenter comme science,* 1783). La connaissance est le résultat de deux facteurs : l'objet et le sujet. Mais comment parler d'objectivité lorsque objet et sujet ne font qu'un ? C'est le cas des sciences humaines et sociales, la sociologie notamment, qui émergent à cette époque… Pour atteindre une certaine objectivité, la science du XIX^{e} siècle propose alors de dissocier, voire d'opposer, l'objet connu et le sujet connaisseur. Les scientifiques se voient eux-mêmes comme de potentiels obstacles à la connaissance. Le XX^{e} siècle voit arriver le doute vis-à-vis de la notion même d'objectivité : les lois scientifiques les plus sûres (comme la gravitation de Newton) sont remises en cause ! Des instruments de mesure de plus en plus précis et de puissantes technologies permises par l'avènement de l'informatique obligent à revoir certains modèles théoriques. Une connaissance scientifique pourra-t-elle jamais être totalement objective ? La science se révèle elle-même être irréductiblement construite sur des valeurs subjectives ! L'idée de l'objectivité forte, portant sur les choses elles-mêmes, correspondrait à une science idéale, indépendante de l'histoire et de la société. Or, la science est une activité sociale. Henri Poincaré l'exprimait ainsi dans *la Valeur de la science* en 1905 : *« Ce qui est objectif doit être commun à plusieurs esprits, et par conséquent pouvoir être transmis de l'un à l'autre. »*

L'objectivité se construit petit à petit, elle possède un caractère dynamique. Les notions subjectives de perceptions, de concepts et de valeurs sont en fait l'unique moyen de l'atteindre. L'observation, l'expérience, l'expérimentation, la méthode en sont les garants.

En pratique, le chercheur, qui d'autant plus est en charge d'une multitude de tâches et de missions administratives, est censé appliquer dans ses recherches une méthode scientifique, basée sur l'observation, l'expérience et le doute constructif, porteuse de garde-fous et garante d'une certaine valeur des résultats produits. Hélas, la mise en œuvre de telles méthodologies nécessite des moyens considérables en termes de temps et de ressources humaines, denrées dont nos politiques de recherche sont de plus en plus avares. Comment valoriser le travail d'expertise, idéalement anonyme, réalisé par un professionnel qui évalue minutieusement la qualité d'une étude originale soumise à la communauté ? Il y a là matière à réunir scientifiques et décideurs autour d'une même table pour y réfléchir.

La passion et la curiosité, le vrai moteur des maths !

Les progrès techniques de ces dernières décennies (saisie et traitement de texte, édition scientifique, communication instantanée, baisse des coûts d'impression, partage massif de documents…) marquent une rupture : ils facilitent le travail du scientifique et donnent dans le même temps naissance à des problématiques nouvelles et des défis passionnants (les précédents articles consacrés au calcul scientifique, à la simulation numérique et à la transdisciplinarité en attestent). Par exemple, le chercheur peut vite se trouver totalement submergé par la massive production de ses pairs ! Comment maintenir une veille bibliographique sérieuse, dans un domaine d'activité donné, quand des centaines de nouveaux articles sont rendus disponibles chaque jour ? La solution la plus immédiate pour le chercheur consiste à se spécialiser dans un secteur extrêmement pointu (où les experts, peu nombreux, sont parfaitement identifiés et où les progrès sont lents) ou dans une thématique « à la mode ». Ces choix stratégiques se font au détriment d'une exploration plus audacieuse, plus originale, plus personnelle, et sont une conséquence de la gestion et de l'évaluation de la recherche scientifique dans nos sociétés. Ne risque-t-on pas alors de voir à terme la science se scléroser ?

Heureusement, la curiosité, la passion, l'enthousiasme et l'abnégation de nombreux chercheurs dans tous les domaines du savoir rendent ce scénario peu probable. En mathématiques, la beauté d'un résultat, la finesse d'un raisonnement, l'esthétique d'une construction, l'émerveillement devant une équation ou une théorie nouvelle, la jouissance de « cracker » un problème ardu qui énerve les experts depuis des années, l'admiration d'un(e) collègue à la pensée féconde ou totalement investi(e) dans ses recherches, autant de notions subjectives, animent les défricheurs de nouvelles terres. Et ce, quelles que soient les difficultés : haute technicité du domaine, débouchés incertains, indifférence des médias, chute des effectifs dans les filières scientifiques, charges d'enseignement et missions administratives lourdes, racket éhonté des éditeurs scientifiques (le développement de journaux spécialisés de nouvelle génération, comme récemment *Discrete Analysis* sous l'impulsion de Tim Gowers, montre cependant que des solutions existent pour ce problème… et pour les autres).

La science a besoin de temps, pas de tampons !

Un autre sujet d'inquiétude majeure est apparu, lui aussi consécutif de l'ingérence du politique et des instances administratives dans le pilotage de la recherche fondamentale : le « *publish or perish* » est devenu une réalité dévastatrice ! Obtenir un poste relève de la mission impossible et tous les coups sont permis pour publier vite. Il faut opposer à ces pressions un retour à la déontologie : la science a besoin de temps, pas de tampons ! Les chercheurs parlent de passion, de partage, de soif de connaissance, de découvrir les lois subtiles de l'univers, quand les décideurs et politiques raisonnent en termes de projets à court terme (à ce titre, l'obsession de certains de figurer en bonne place dans le prochain « classement de Shanghai » des universités est significative). La tendance est au financement massif de « grands » projets d'envergure, aux objectifs bien balisés *a priori* (!), mobilisant des dizaines voire des centaines de chercheurs.

Ce modèle n'est pas bien adapté aux mathématiques, dans lesquelles les collaborations se font rarement à plus de quatre ou cinq chercheurs et où les besoins matériels restent relativement modestes. Une autre répartition des moyens financiers (moins conséquents mais à plus d'équipes) serait préférable. Qui peut sincèrement croire que l'électricité a été découverte en investissant massivement dans le perfectionnement de la bougie ? N'est-ce pas plutôt la liberté de pensée de quelques défricheurs courageux qui permet les innovations (les vraies) ? Et si décideurs et scientifiques se retrouvaient, là encore, pour en discuter ?

Revenons aux valeurs fondamentales et fondatrices pour continuer à donner envie à nos jeunes d'embrasser des carrières exaltantes, et pour que cessent les affaires de fraude scientifique. Ces dernières se sont multipliées ces dernières années, notamment (mais pas seulement) dans les sciences humaines et sociales : études bidonnées, résultats créés de toutes pièces ou non vérifiables, plagiat, données inventées, expériences biaisées, théories non testables, conflits d'intérêt… Publier le premier et se faire remarquer pour obtenir le financement nécessaire à une étude peuvent pousser le plus vertueux des scientifiques à franchir la ligne rouge, surtout si ses collègues n'ont plus le temps de vérifier ses travaux.

Des enjeux sociétaux majeurs, de la finance au Big Data

Les mathématiques ne sont pas épargnées par ces considérations ! On a ainsi vu en 2006, après la démonstration de la conjecture de Poincaré par le Russe Grisha Perelman, des tentatives d'attribution indue de ce glorieux résultat scientifique. La communauté s'est immédiatement indignée, et les deux chercheurs chinois concernés ont dû faire amende honorable.

Quelques années plus tard, les médias se sont déchaînés sur la responsabilité des mathématiques financières et des enseignants-chercheurs dans la crise actuelle. L'analyse stochastique appliquée a permis l'avènement de produits financiers douteux. On a créé des outils théoriques qui peuvent être détournés de leurs fins et s'avérer concrets et dangereux ! Avait-on déjà oublié le rôle de l'algèbre dans la fabrication de la bombe atomique ? Qui est responsable de l'exploitation illicite d'un théorème assorti de ses hypothèses, ou d'un concept sorti du cadre conceptuel dans lequel il a été défini : le découvreur ou l'utilisateur ? Qui est responsable du détournement d'un outil, livré avec sa notice d'utilisation, à des fins malveillantes ?

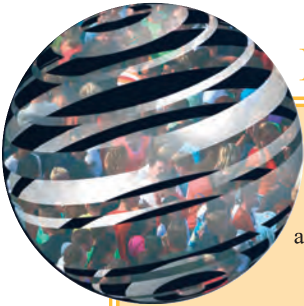
Plus proche de nous, l'affaire Snowden illustre certaines dérives liées aux données et au Big Data. Ainsi, la National Security Agency est l'une des principales sources de financement de la recherche américaine en mathématiques. Mais l'exploitation du fruit de ces recherches invite à s'interroger quant à la moralité voire à la légalité de plusieurs activités de l'agence, notamment concernant certaines collectes de données à des fins de surveillance électronique massive de la population. Rares sont les mathématiciens qui ont le courage de proposer de rompre tout lien, scientifique, financier ou autre, avec la NSA…

Plus encore que la cryptographie hier ou la finance aujourd'hui, demain c'est le Big Data et toutes les sciences des données qui seront concernées par ces questions déontologiques profondes. Oui, le mathématicien, par son travail, par son indépendance, a un impact sur le monde, sur la société : plus aucun domaine technologique, aujourd'hui, ne peut se passer des mathématiques. Notre communauté doit en prendre conscience et agir en conséquence. Même si le plaisir de faire des maths, la curiosité et la passion seront toujours nos moteurs !

É.T.

Pour en savoir (un peu) plus :

La science selon Henri Poincaré. Henri Poincaré, Dunod, 2013.
Les trois classiques de Poincaré, très lisibles et toujours d'actualité : la Science et l'Hypothèse, la Valeur de la science *et* Science et Méthode.



Bibliographie « maths et société »

Le Comité international des jeux mathématiques est une association indépendante. Nous avons sélectionné les ouvrages suivants pour les lecteurs de *Maths Société Express* qui souhaitent interroger plus avant la place des mathématiques dans nos sociétés.

Le dernier théorème de Fermat. Simon Singh, Fayard–Pluriel, 2011.
La quête historique d'un résultat mathématique qui a enflammé le monde pendant trois cent cinquante ans. L'aventure humaine par excellence.

Le chercheur fantôme. Robin Cousin, Flblb, 2013.
Et si un problème de mathématiques fondamentales (« P versus *NP ») pouvait révolutionner la conception des algorithmes, des ordinateurs, des robots, et bouleverser notre quotidien ?*

L'équation du millénaire. Fondation Sciences Mathématiques de Paris, 2013.
Disponible en ligne : http://www.sciencesmaths-paris.fr/fr/bd-462.htm
Une bande dessinée autour de la dynamique des fluides et des équations de Navier–Stokes.

Homo mathematicus. Jean-Pierre Boudine, Vuibert, 2001.
Le plus large éventail de réponses à la question : « Les mathématiques, à quoi ça sert ? »

Mathématiques en liberté. Pierre Cartier, Jean Dhombres, Gerhard Heinzmann et Cédric Villani, La Ville Brûle, 2012.
Deux mathématiciens, un philosophe et un historien des sciences débattent sur la place des maths aujourd'hui dans notre société, à partir de questions que vous vous posez peut-être.

« De la petite vérole au XVIIe siècle au cancer aujourd'hui : ce que peuvent apporter les mathématiques. » Conférence de Dominique Barbolosi à la Bibliothèque nationale de France (Paris) le mercredi 30 mars 2016, disponible en ligne. (vimeo.com/162821676).
Comment convaincre en une heure trente un adolescent de s'engager dans les sciences : oui, les mathématiques, ça peut aussi sauver des vies !

Un mathématicien aux prises avec le siècle. Laurent Schwartz, Odile Jacob, 1997.
L'un des plus brillants mathématiciens du XXe siècle fut également un intellectuel engagé. Il a rassemblé dans cet ouvrage majeur une présentation de ses différents travaux scientifiques et une synthèse de ses combats militants.

Des mathématiciens et des guerres. Collectif, CNRS Éditions, 2013.
L'une des premières études du genre : quel fut l'engagement des mathématiciens dans les grands conflits du passé ?

Mathématiques d'ailleurs. Marcia Ascher, Le Seuil, 1998.
Ou comment certaines peuplades ont dû développer d'autres types de mathématiques que celles que nous connaissons pour répondre à leurs besoins.

Nous recommandons également les ouvrages suivants, tous accessibles, aux visiteurs du Salon qui souhaiteraient acquérir une culture de base en mathématiques :

Logicomix. Apostolos Doxiadis et Christos Papadimitriou, Vuibert, 2010.
Une plongée dans l'histoire moderne des mathématiques, de la « crise des fondements » au « théorème d'incomplétude de Gödel », sous la forme d'un grand roman graphique.

Thalès, Pythagore, Euclide, Archimède. Stéphane Favre-Bulle, Ellipses, 2004.
Quatre portraits à l'aquarelle par un professeur de maths, dessinateur de talent.

Grigori Perelman face à la conjecture de Poincaré. Donal O'Shea, JC Lattès, 2007.
Le récit incroyable de la plus grande découverte scientifique des années 2000.

Alex au pays des chiffres. Alex Bellos, Robert Laffont, 2011.
Une voie privilégiée pour « entrer en mathématiques ».

Merveilleux nombres premiers. Jean-Paul Delahaye, Belin, 2013.
Pour tout savoir sur les nombres premiers, tout simplement.

Éloge des mathématiques. Alain Badiou et Gilles Haéri, Flammarion, 2015.
Le plaidoyer d'un philosophe pour les maths vues comme clé pour une vie heureuse.

L'éternité dans une heure. Daniel Tammet, Les Arènes, 2013.
De l'art et la manière de poser des questions aussi pertinentes que poétiques.

La symphonie des nombres premiers. Marcus Du Sautoy, Héloïse D'Ormesson, 2011.
La passion d'une quête, celle de la compréhension de la répartition des nombres premiers (à travers l'hypothèse de Riemann), par un vulgarisateur hors pair.

Théorème vivant. Cédric Villani, Grasset, 2013.
La quête de l'explication par les mathématiques d'un phénomène physique aussi insaisissable que fondamental : l'amortissement Landau.

Amour et maths. Edward Frenkel, Flammarion, 2015.
Une déclaration d'amour pour les mathématiques, par un spécialiste du très ambitieux programme de Langlands.

Le théorème du perroquet. Denis Guedj, Le Seuil, 2000.
L'un des premiers romans historiques sur l'histoire des mathématiques. Un classique !

Les maths à toutes les sauces. Bernadette Guérite-Hess, Isabelle Causse-Mergui et Marie-Céline Romier, Le Pommier, 2005.
Vous avez des difficultés avec les fondamentaux en mathématiques ? Mettez la main à la pâte avec ces activités minutieusement décortiquées.

L'affaire Olympia. Mickaël Launay, Le Pommier, 2013.
Retrouvez l'auteur sur le site Micmaths, bric-à-brac mathématique et ludique (www.micmaths.com).

Samadhi, le pouvoir oublié. Xavier Debarge et Frédéric Veber, Un Autre Reg'Art, 2015.
Une bande dessinée fantastique qui mêle intelligemment action et géométrie.

La recherche MATHÉMATIQUE se prend au JEU

LES MATHS
VECTEUR D'INNOVATION

AMIES est
l'*Agence pour les Mathématiques*
en Interaction avec
l'Entreprise et la Société.
Sa mission principale est de mettre en relation des mathématiciens et des entreprises françaises, au travers d'initiatives incitatrices telles que les PEPS, les SEME, les FEM, et bien d'autres.

C'est une agence nationale,
portée par le CNRS en partenariat avec Inria et UGA
(Université Grenoble Alpes)
De ce fait, ses services sont absolument gratuits.

- Les PEPS sont des projets proposés par une entreprise et subventionnés par AMIES.

- Chaque SEME est une semaine d'étude où des problèmes industriels sont planchés par groupes de 5-6 thèsards en maths, avec une restitution le vendredi même. Il y est demandé d'explorer des stratégies de réponse inédites pour l'industriel.

- Les FEM sont les Forums Emploi Maths, où des entreprises et des étudiants peuvent se rencontrer à l'occasion de présentations mutuelles et de conférences.

◊ Depuis la création d'AMIES, plus de 50 PEPS ont été alloués, avec des montants d'aide pouvant aller jusqu'à 50k€. La première rencontre s'effectue le plus souvent avec un facilitateur AMIES, qui est mathématicien lui-même et peut ainsi rediriger l'entreprise vers le bon interlocuteur.

◊ Les SEME ont lieu approximativement tous les trimestres : la 17ème se déroulera à Marseille en juin 2016. Elles ont réuni jusqu'à maintenant plus de 75 entreprises et plus de 370 étudiants.

◊ Le FEM2016 constituera le grand rendez-vous national le 15 décembre 2016 à Paris-La Villette.

www.agence-maths-entreprises.fr

Centre de recherche Inria de Paris en bref

Inria
INVENTEURS DU MONDE NUMÉRIQUE

600 personnes dont
520 scientifiques
35 équipes de recherche

20 bourses ERC depuis 2009

1 à **2** nouvelles start-up par an

En partenariat avec :

Les universités Paris Dauphine, Paris Diderot, Paris-Est Marne-la-Vallée, Pierre et Marie Curie, École des Ponts ParisTech, École Normale Supérieure de Paris, Mines ParisTech, Ircam, Cerema, Cnrs, Inserm.

Membre de :

PSL RESEARCH UNIVERSITY

SORBONNE UNIVERSITÉS

USPC Université Sorbonne Paris Cité

Inria.fr/Centre/Paris

@inria_paris

Centre de recherche Inria de Paris en bref

Inria
INVENTEURS DU MONDE NUMÉRIQUE

600
personnes dont
520 scientifiques
35 équipes
de recherche

20 bourses ERC depuis 2009

1 à **2** nouvelles start-up par an

En partenariat avec :

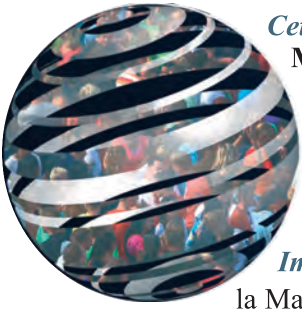
Les universités Paris Dauphine, Paris Diderot, Paris-Est Marne-la-Vallée, Pierre et Marie Curie, École des Ponts ParisTech, École Normale Supérieure de Paris, Mines ParisTech, Ircam, Cerema, Cnrs, Inserm.

Membre de :

PSL RESEARCH UNIVERSITY — SORBONNE UNIVERSITÉS — USPC Université Sorbonne Paris Cité

Inria.fr/Centre/Paris

@inria_paris



Cette brochure placée sous le parrainage de
Marie Ekeland

a été réalisée par le
Comité International des Jeux Mathématiques

sous la direction de
Marie José Pestel et Édouard Thomas

Imprimée grâce à
la Mairie de Paris et Sciences sur Seine,
la Région Île-de-France, l'AMIES, l'INRIA, Cap'Maths, le CNRS,
l'UPMC, le Crédit Mutuel Enseignant et les Éditions POLE

Elle réunit les signatures de

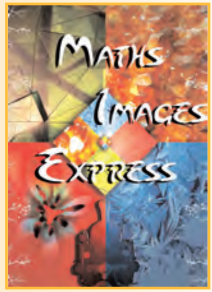
Marie Ekeland
Daniel Kious
FibreTigre
Gérard Berry
Hervé Lehning
Jean-Michel Loubes
Marc Lavielle
Céline Grandmont et Jean-Frédéric Gerbeau
Stéphanie Salmon
Roland Lehoucq
Édouard Thomas
Nicolas K. Blanchard et Leila Gabasova
Didier Bresch
Véronique Slovacek-Chauveau et Annick Boisseau
Leila Schneps
Karel Janeček
Christian Lavigne

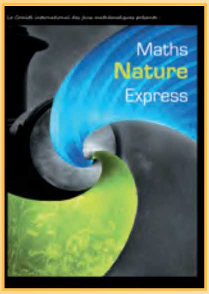
Que tous ces auteurs soient ici remerciés pour leur enthousiasme,
leur patience et leur gentillesse. Grâce à eux, nous espérons
que le lecteur prendra plaisir à découvrir que mathématiques et société
sont indissociables dans l'histoire des hommes, de leurs cultures
mais surtout de leur avenir.

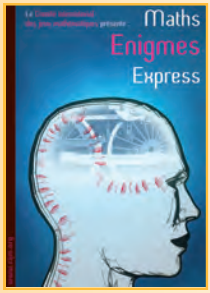
Réalisation : Patrick Arrivetz
Maquette de couverture et bandeau : Elsa Godet – www.sciencegraphique.com
Achevé d'imprimer sur les presses de CIA GRAPHIC – 03 86 90 96 10

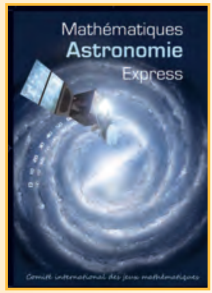
Maths Express

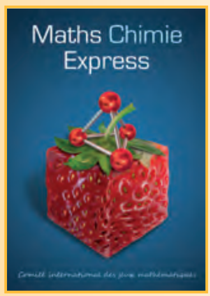
une collection CIJM - www.cijm.org

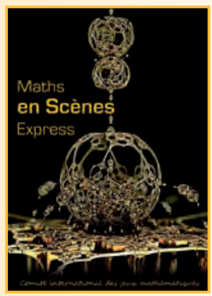


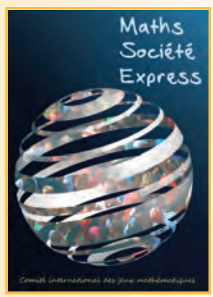












MINISTÈRE
DE L'ÉDUCATION NATIONALE
DE L'ENSEIGNEMENT SUPÉRIEUR
ET DE LA RECHERCHE

CIJM
Institut Henri Poincaré
11 rue Pierre et Marie Curie
75231 Paris Cedex 05

www.cijm.org