

N° SPÉCIAL

■ POUR LA

SCIENCE

Décembre 2000

édition française de

SCIENTIFIC
AMERICAN

www.pourlascience.com

Les
infinis

TRIBUNE DE LECTEURS

BLOC-NOTES

de Didier Noiret

POINT DE VUE

La pilule abortive aux États-Unis

par André Ulman

SCIENCE ET ÉCONOMIE

La stratégie du pénalty

par Ivar Ekeland



Tout au chocolat

par Hervé This

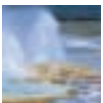


PERSPECTIVES SCIENTIFIQUES

Un nœud... peu gordien

de la structure des entiers

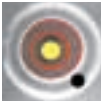
- Culte vaudou à Mayence
- Le chant d'amour des drosophiles
- Éclaircie sur Titan
- Virus cancéreux des herbes
- Porcelaines vietnamiennes
- La danse des liquides
- La cartographie de l'espace
- La fonction du prion
- Bateau romain... coulé
- Poisson fumé à la néandertalienne
- Dunes en liberté
- Un rat artificiel



IDÉES DE PHIL

Les geysers

par Philippe Soussol



ART ET SCIENCE

Un nœud... peu gordien

par Jean Baudry



ANALYSES DE LIVRES

Les sociétés animales. Évolution de la coopération

et organisation sociale, de Serge Aron et Luc Passera

- Chimie, l'art de se poser les bonnes questions, de J. Garratt, T. Overton et T. Threlfall
- Le défi paysan, de Luc Guyau

Deux encarts d'abonnement entre les pages 18 et 19, un encart broché service lecteurs et une carte d'abonnement entre les pages 146 et 147. Un encart parrainage pour les exemplaires destinés aux abonnés. Un dépliant trois volets Pour la Science



Chaque mois, retrouvez le sommaire complet de la revue **en ligne** avec pour chaque article une bibliographie et un complément d'information.

www.pourlascience.com

6

7

8

9

10

12

152

154

156

Les mathématiques, science de l'infini

24

par Javier Lorenzen et Lorenzo Martínez

L'infini, potentiel et actuel, semble assurer une fonction catalytique dans la formalisation mathématique.

L'infini est-il paradoxal en mathématiques?

28

par Jean-Paul Delahaye

Pour résoudre le paradoxe du tout et des parties et affronter l'hypothèse du continu, notre idée de l'infini actuel doit évoluer.

L'infini en Chine

44

par Jean-Claude Martzloff

Le calcul numérique a joué un rôle clé dans la perception que la Chine a eu de l'infini.

Thabit ibn Qurra et l'infini numérique

48

par Tony Levy

Face aux thèses d'Aristote, longtemps dominantes en philosophie, un savant arabe du IX^e siècle eut sur l'infini un point de vue aussi original qu'audacieux.

L'univers infini de Giordano Bruno

54

par Giulio Giorello

Giordano Bruno, mort sur le bûcher le 17 février 1600, fut plus audacieux que Copernic : il envisageait un univers infini, sans horizon ni limite.

La science du mouvement au XVII^e siècle

59

par Michel Blay

À la fin du XVII^e siècle et au début du XVIII^e, les mathématiciens décrivent le mouvement et introduisent la notion de vitesse, en évitant les paradoxes qui résultent de l'usage impropre et intuitif de l'infini.

Histoire d'infini

64

par Maria Reményi

Pendant 2 000 ans, le concept d'infini gêna les mathématiciens. John Wallis posa le symbole ∞ , les libérant des discussions philosophiques.

De la perspective à l'infini géométrique

66

par Jean-Pierre Le Goff

Au XVII^e siècle, grâce à la géométrie projective, née de l'union de la théorie de la perspective et de la théorie des coniques, l'infini géométrique acquiert une « dimension » humaine, une évidence perceptible.

L'ensemble triadique de Cantor

73

par André Lecomte

Les préparés au nettoyage de la planète Ter découvrent un ensemble de Cantor qui remplit un segment de largeur 1, mais son complémentaire aussi...

L'infini et l'univers des algorithmes

108

par Gilles Dowek

On ne peut pas toujours prévoir le nombre d'opérations nécessaires à un calcul. Aussi, pour permettre l'expression de tous les algorithmes, il faut introduire la possibilité que les calculs durent infiniment longtemps.

L'infini, pierre de touche du constructivisme

74

par Allan C. Calder

Selon les mathématiciens constructivistes, un objet mathématique n'existe que si on peut le construire. Cette définition de l'existence mathématique oppose, depuis plus d'un siècle, constructivistes et formalistes.

L'infiniment petit en physique

112

par Harald Fritzsch

En descendant l'échelle des dimensions, les physiciens ont découvert des particules si petites qu'ils les ont supposées ponctuelles. Le «modèle standard», la théorie construite sur cette hypothèse, accumule les succès depuis 30 ans, tout en étant pectée d'infinis...

L'infini en géométrie

84

par Marcel Berger

Lors d'un séminaire, un géomètre recourt aux éléments du décor pour illustrer le rôle important de l'infini en géométrie. L'analyste et l'algébriste seront convaincus...

L'infiniment grand

120

par Gerhard Eberhard

L'univers semble être brusquement sorti d'un état initial infiniment dense, que les physiciens ont de difficultés à imaginer et à décrire par une théorie.

L'analyse non standard

92

par Jean-Michel Salanskis

En introduisant l'analyse non standard, l'analyse non standard a conféré aux infiniment petits un statut honorable et justifié des opérations qui, bien que pratiquement satisfaisantes, posaient des problèmes de principe.

L'infiniment vide n'existe pas

128

par Timothy Boyer

Le vide classique est pur vide. Mais quand on a supprimé toute la matière et le rayonnement thermique d'une partie de l'espace, le vide de la physique contient encore des champs électromagnétiques, des particules fugaces, et possède une certaine énergie.

Les géométries non commutatives

98

par Daniel Biniatti

Les espaces d'opérateurs de dimensions infinies étudiés à l'aide des géométries non commutatives servent à élaborer de nouvelles théories en physique.

L'infini est un révélateur

138

par Patrick Dehornoy

L'infini est source d'inspiration. L'exploration d'ensembles hyperinfinis, impossibles à construire, mais pas inconcevables, a fait découvrir des objets mathématiques comme les tables de Laver et des propriétés des tresses.

L'infini est-il nécessaire?

102

par Patrick Dehornoy

Oui. Les suites de Goodstein gonflent et grossissent jusqu'à des tailles gigantesques... Et diminuent finalement pour atteindre zéro. Pour démontrer cette propriété paradoxale, il est inévitable de faire appel à l'infini.

L'infini littéraire

148

par Paul Braffort

Dans le sillage des philosophes, romanciers, poètes et essayistes ont été inspirés par l'infini.

TRIBUNE DES LECTEURS

Les promesses du WAP

Je réagis à l'article «Les promesses et les risques du WAP» (voir *Pour la Science*, novembre 2000).

Les faiblesses du WAP qui sont avancées comme étant à l'origine de son insuccès sont techniques pour la plupart. Pourtant, l'article reconnaît plus loin l'extraordinaire réussite d'*i-Mode* au Japon, alors que le protocole et les terminaux sont à peine plus riches, et sont plus ou moins équivalents au WAP sur accès GPRS.

C'est qu', et en tant que scientifique de formation et spectateur assidu de *Pour la Science* j'en suis le premier attristé, l'excellence technique n'est pas toujours synonyme de pertinence commerciale. Il manque au WAP un ingrédient essentiel : un outil simple et direct qui permette au fournisseur de service d'être rémunéré. Les services *i-Mode* qui font la majeure partie de l'audience nipponne sont tous payants (100, 200 ou 300 yens par mois). Les meilleurs obtiennent donc rapidement les moyens de développer leur activité, de communiquer et donc d'amener de nouvelles parts de marché à l'opérateur : cercle vertueux impossible à amorcer sur un modèle *Internet* «classique», c'est à dire reposant sur des hypothétiques modèles économiques maladroits.

Cependant, la vertu du WAP étant son ouverture, l'adoption d'un tel outil est tout à fait similaire à celle de *i-Mode*, mais également ouverte et donc multi-opérateur donnerait l'opportunité à tous les fournisseurs de service. Les éditeurs de service n'aiment pas plus que les opérateurs ne dépendre que d'un seul fournisseur.

Yves CHRISTOL, Paris

Fée à l'huile

J'ai lu avec intérêt l'article «La Fée électricité» (voir *Pour la Science*, septembre 2000). Son sujet est d'autant plus intéressant qu'un des collaborateurs à la réalisation du tableau écrivait que Raoul Dufy «exécuta cette œuvre immense en quatre mois, après avoir minutieusement étudié l'histoire de l'électricité...»

Je crois cependant utile de préciser que Jacques Maroger n'était pas chimiste, mais peintre, élève de Jacques-Émile Blanche, puis de Louis Anquetin, et restaurateur de tableaux (ce qui ne l'empêcha pas de devenir directeur technique des laboratoires du Musée du Louvre). Maroger allait rencontrer un chimiste des peintures au début des années 1930 en la personne de Marc Havel, chimiste chez Joseph Bourgeois (établissement qui fusionnera 40 ans plus tard avec Lefranc pour fonder la Société π). C'est d'ailleurs à la maison Bourgeois que sera confiée la préparation des peintures destinées à *La Fée électricité* (à signaler que ce

tableau, composé de 250 panneaux de contre-plaqué, mesure 60 mètres de longueur et 10 mètres de hauteur!). Cette peinture est composée de pigments dispersés dans une huile siccative diluée de 10 pour cent de résine dammar. Mais l'originalité de la formule de Maroger est d'ajouter un médium particulier (le médium est un liant ajouté à une peinture, en cours de travail ou peu avant, pour lui conférer les qualités techniques recherchées). En effet, il s'agit d'une émulsion, une «mayonnaise», d'huile siccative, cuite en présence d'oxyde de plomb, dans une solution de colle de peaux. Les peintures à l'huile étaient mélangées en début de séance à ce médium. Donc, cette peinture à l'huile se travaillait à l'eau!

Effectivement, le médium, comme la peinture prête à l'emploi, que Dufy a dû avoir sous des émulsions d'huile dans de l'eau. Or, dans les émulsions, c'est la phase externe qui détermine l'affinité. Dans le cas présent, la phase externe est une solution de colle, et de ce fait la peinture se diluait à l'eau. En séchant, la proportion de colle étant relativement faible, l'émulsion «claquait» pour ne laisser qu'un film de peinture à l'huile (incluant plus ou moins de colle de peaux).

L'originalité de cette technique est de permettre des œuvres à l'huile tout en terminant sur une peinture à l'eau mate, qui autorise de nouvelles possibilités picturales. On peut citer, entre autres comment Dufy, Maroger et Pierre Paulet, a pu «couvrir» 60 mètres carrés en si peu de temps.

François PÉRÉGO, Bécherel

Muscles et psychologie

J'ai lu avec intérêt l'article «Muscle, gènes et performances» par Jesper Andersen, Peter Sejerling et Bengt Saltin (voir *Pour la Science*, octobre 2000) et il me laisse une impression mitigée. En effet, il semble au premier abord être écrit par des biologistes et des généticiens, mais certains encadrés font part d'hypothèses plus psychologiques ou éthiques qui entraînent une certaine confusion.

En ce qui concerne les troubles de l'image corporelle, on devrait parler plus précisément de dysmorphophobie – peur d'être mal formé – et non de dysmorphie, terme qui dénote une malformation réelle et constatée par autrui. L'expression «dysmorphie musculaire» ne veut rien dire sur le plan psychologique.

Par ailleurs, les trois neuromédiateurs cités sont ceux que l'on étudie régulièrement dans les problèmes de dépressions, d'attaques de panique, de troubles du comportement alimentaire, de schizophrénie... Ils n'ont rien de spécifique dans le sujet en question – les

sports pour lesquels ils sont étudiés sont ceux de sensations fortes – et devraient être traités par les neurobiologistes plutôt que par des spécialistes du muscle.

Bernard GONZALEZ, Neaumont

Un problème de dimension

L'article «Les dimensions cachées de l'univers» (voir *Pour la Science*, octobre 2000) essaie de élucider l'existence de dimensions supérieures en comparant ce qui se passe dans des dimensions inférieures. Cet exercice ne m'a pas toujours convaincu.

Ainsi, dans le livre *Flatland*, cité dans le texte, quand on parle de figures à deux dimensions, on pense plutôt aux formes découpées dans du papier, ayant une troisième dimension infime mais réelle. Un vrai monde bidimensionnel serait celui d'ombres, qui peuvent se chevaucher sans jamais se heurter pour produire des ombres qui se chevaucheraient dans une troisième dimension.

Ensuite, le fil de coton avec son acarien, le coussin de tuyau avec une fourmi, gardera ses deux dimensions tant qu'on peut le discerner. Une ligne dans une dimension sera vu comme une surface sur un fond contrastant. Tant qu'il reste visible, le fil conserve ses deux côtés, quoiqu'il soit leur propre collection de points, rapprochés mais distincts.

Neuf dimensions ne sont pas une ligne comme une collection infinie de points comme parfois on a l'habitude de le faire : les dimensions zéro d'un point ne donneront que zéro, même multipliés à l'infini.

Une meilleure manière de passer d'une dimension n à une dimension $n+1$ est de voir une ligne comme la trajectoire d'un point, une surface comme la trajectoire d'une ligne, un volume comme celle d'une surface...

Le temps est une quatrième dimension donne la trajectoire des objets. En continuant sur cette voie vers des dimensions plus élevées, je ne vois pas bien comment arriver aux petites boucles à l'échelle de Planck, si ce n'est que puisse être l'hypothèse des cordes, qui nécessite aussi un nombre élevé de dimensions, pour expliquer la double nature corpusculaire des particules élémentaires et des forces qui les font interagir.

A. VAN DER SLUIS, La Régence de Mazenod

Écrivez à *Pour la Science* vos réactions, vos commentaires et vos réflexions sur l'actualité scientifique et vos relations avec la science. L'adresse de *Pour la Science* sur Internet est : tribune@poullascience.fr. Découvrez d'autres contributions de lecteurs sur le site Web www.poullascience.com.



BLOG-NOTES

DIDIER NORDON



Les épris de prix prisés primés

Les frères Goncourt seraient moins fameux si Edmond n'avait instauré le prix Goncourt. Fields, dont le nom inspire tant de respect aux mathématiciens, doit cela non à ses travaux, mais à la médaille instituée par lui. Enfin, rares seraient les érudits qui connaîtraient le nom de l'inventeur de la dynamite ou l'inventeur n'était à l'origine de la plus convoitée de toutes les récompenses, le prix Nobel.

À côté de cela, des centaines de donateurs ont financé des prix censés, année après année, honorer un vivant talentueux tout en perpétuant la mémoire du donateur. Hélas, même bien dotés, la plupart de ces prix végètent et ne sont connus de personne, sinon des heureux lauréats et des académies qui les décernent.

Immortaliser son nom en créant un prix n'est donc pas facile. Le prix Nordon ne mérite-t-il d'être, à leur tour, récompensés. J'institue donc un prix qui sera attribué, à titre posthume, à celui qui, pendant le siècle écoulé, aura le mieux su s'immortaliser grâce au prix qui porte son nom. Au titre du XX^e siècle, le prix Nordon est attribué à Nobel.

Néoloschisme

Inutile de connaître un mot pour le comprendre immédiatement lorsqu'on l'emploie devant vous. Inutile, même, qu'il existe.

Écoutant une conférence de Jean-Pierre Bourguignon, j'ai entendu ce dernier affirmer que la thèse de Riemann est un des textes les

plus «surloquants» de l'histoire des mathématiques. L'étymologie de ce mot m'est apparue en même temps que je l'entendais pour la première fois : latin *loquor*, je parle, précédé d'un préfixe créant un superlatif. Expressif, bien formé, «surloquant» sert à qualifier un propos extraordinairement dense et riche qui, poussant à son paroxysme l'efficacité de la parole, en dit plus qu'on n'en peut normalement dire et donne en peu de lettres de quoi méditer longuement. Le mot va sûrement se répandre.

Toutefois, poursuivant mon enquête, j'ai réalisé que j'avais mal compris. Bourguignon, par ailleurs, a dit que le texte de Riemann est «surécoulant». Ce mot, qui n'est pas non plus dans le dictionnaire, est l'adjectif dérivé d'une étymologie évidente grâce à laquelle il se comprend d'emblée. La décence, hélas, interdit de préciser. Disons que le mot relève de l'arsenal péjoratif, et tenons-nous, en là.

D'où naissent les idées?

Dans un cerveau, certes, mais cette réponse est vague, et des techniques perfectionnées de focalisation cérébrale sont mises à contribution pour déterminer quelles zones sont activées lors des périodes de créativité.

Seulement, quand sera posée la question «Où naissent les idées?», une deuxième se posera : «Où est née l'idée de se demander où naissent les idées?» Et ce sera le début de la fin. Qui dit deux, dit l'infini. La récurrence s'est amorcée, rien ne pourra plus l'arrêter. La deuxième question posée, une troisième viendra, inévitable : «Où est née l'idée de se demander où est née l'idée de se demander où naissent les idées?» Etc., jusqu'à l'infini.

Pour une fois, ce formalisme rigide ne se réduit pas à un pur jeu de l'esprit. À sa façon, il dit que, quels que soient les progrès de l'imagerie cérébrale, nous ne saurons jamais ni ce qui se produit dans un cerveau quand y jaillit une étincelle ni pourquoi elle y jaillit.

Cette impossibilité de savoir d'où viennent les idées rend décevantes les réponses des chercheurs et artistes censés sur la genèse de leur inspiration. Au lieu de nous enseigner dans quel état nous devons mettre notre cerveau pour qu'il soit siège d'idées fulgurantes, ils répondent par des anecdotes à côté du sujet. «Je ne puis travailler sans musique de fond», ou «Je fais de grandes marches dans la nature en ne pensant qu'au problème qui me préoccupe», ou encore «J'ai été soudain frappé par la ressemblance

inattendue entre un phénomène et un autre... Ces réponses, qui ne révèlent pas le secret grâce auquel, nous aussi, pourrions avoir des idées géniales. C'est que, le secret, nul ne sait le percer. Pas même celui qui en a, des idées géniales.

La loi de toutou rien

L'enseignement étant trop dogmatique, de nombreuses personnes quittent l'école en ayant l'impression que les mathématiques procèdent soit tout ou rien : vrai/faux, démontré/non démontré, rigoureux/non rigoureux. Du coup, lorsque des gens entendent dire que telle conjecture fameuse — le théorème de Fermat jusqu'en 1995 ou l'hypothèse de Goldbach — n'est pas démontrée, ils en déduisent que rien n'est fait dans le domaine. Même ceux qui ont gardé du goût pour les mathématiques se dispensent de s'informer des nombreux résultats partiels. Lorsqu'ils font une remarque, ils croient donc être les premiers à la faire. Mieux encore : parce qu'ils donnent un résultat, et qu'en mathématique c'est tout ou rien, ils pensent qu'elle donne le résultat.

Ils proposent alors leur démonstration aux professionnels, soupçonnant ces derniers de manifester un esprit de corps étroit. Ils ne précipitent pas, toutes affaires cessantes, vers leur démonstration comme vers une révélation. Vouloir se faire lire par les mathématiciens est légitime. Mais vouloir se faire lire d'eux sans avoir pris la peine de les lire au préalable ne l'est pas.



Jean-Michel Thiriet



La pilule abortive aux États-Unis

ANDRÉ ULMANN

Le RU 486 est désormais disponible aux États-Unis, mais ses inventeurs n'en tireront aucun bénéfice.

L'administration sanitaire des États-Unis, la *Food and drug administration* (FDA) vient d'autoriser la commercialisation, aux États-Unis, du RU 486, une pilule abortive. Cette décision constitue, pour ceux qui, comme moi, ont contribué au développement de ce produit et à sa mise sur les marchés européens, une immense satisfaction, entachée, toutefois, d'un certain dépit.

Une immense satisfaction de voir enfin reconnue, sur le premier marché pharmaceutique mondial une invention très originale due à une équipe de chercheurs français dirigée, au sein de *Roussel Uclaf*, par Georges Teutsch et Daniel Philibert, et remarquablement conseillée par Étienne Émile Baulieu. La communauté scientifique internationale s'accorde pour considérer qu'il s'agit sans doute de la seule vraie découverte dans le domaine de la contraception depuis l'introduction de la pilule contraceptive. Le produit a tenu ses promesses notamment comme moyen médical d'interruption de grossesse. Quand une femme souhaite avorter, on peut l'administrer dès que le test de grossesse se révèle positif et jusqu'à environ cinq semaines de grossesse.

Contrairement à d'autres domaines, les États-Unis sont en retard sur l'Europe en matière de contraception et d'avortement, tant sur le plan médical que juridique. Le débat sur l'avortement y est récurrent et d'une extrême violence (des médecins ont été accusés pour avoir pratiqué des avortements) du fait d'une minorité très active qui ne reflète pas l'opinion de la majorité de la population. L'arrivée du RU 486 devrait contribuer à combler ce retard et à assainir ce débat en permettant notamment aux gynécologues de pratiquer des interruptions de grossesse en toute sécurité tant pour leurs patientes que pour eux-mêmes.

De surcroît, l'homologation du produit a reposé exclusivement sur les études effectuées sous la direction de *Roussel Uclaf*, l'administration américaine n'ayant pas jugé nécessaire de demander des évaluations complémentaires. Les études

supplémentaires faites aux États-Unis n'ont fait que confirmer les résultats de *Roussel* quant à l'efficacité du produit et au fait qu'il est bien toléré.

Cette satisfaction se double d'une grande amertume devant la démission de *Roussel Uclaf*, poussé par sa maison mère, *Hoechst* : en 1993, suite à la demande de l'administration Clinton d'octroyer le produit aux États-Unis, le laboratoire a fait don, sans aucune contrepartie financière, des droits d'exploitation du RU 486 à un organisme à but non lucratif, le *Population Council*, donnant l'impression qu'il voulait s'en débarrasser. À son tour, celui-ci a vendu ses droits nouvellement acquis à une petite société américaine qui en a péniblement assuré le développement (il lui a fallu sept ans pour obtenir l'approbation de la FDA) et qui va maintenant le commercialiser.

Il est vrai que *Roussel Uclaf* avait déjà manqué le marché chinois du RU 486 en refusant d'en vendre en Chine, alors que l'autorisation de mise sur le marché y avait été accordée l'année précédente. *Hoechst* a refusé la paternité du produit. De même, en 1997, *Roussel Uclaf* a fait don du produit – ici encore sans aucune contrepartie financière – à l'un des sept anciens dirigeants afin de s'en déposséder définitivement. Ce dernier, Édouard Sakiz, a fondé la Société *Exel-gyn*, qui fabrique et commercialise le RU 486 en France.

Amertume également de constater que *Roussel Uclaf* a, à son tour, disparu dans le tourbillon des fusions répétées, et, avec ce Laboratoire, un savoir-faire unique dans le monde en matière de synthèse des hormones stéroïdes, définitivement perdu pour la France. Si cette disparition programmée a récemment ému un certain nombre de dirigeants ou de syndicalistes, ceux-ci n'ont rien fait, alors qu'il était peut-être encore temps, pour empêcher que l'on ne brade l'expérience des chercheurs de *Roussel Uclaf* et que l'on n'abandonne les recherches sur le RU 486, en un mot que l'on ne ferme

les laboratoires de Romainville. Aujourd'hui, comme en témoigne la presse outre-Atlantique, ce produit est devenu américain : son origine européenne a été définitivement oubliée.

À tout cela s'ajoute une certaine inquiétude dans la mesure où la société américaine chargée de commercialiser le produit n'est pas expérimentée ; elle risque d'avoir des difficultés à faire face aux procès qui surviendront inévitablement car ils restent la seule arme à la disposition des organisations anti-avortement. Inquiétude également de la part des femmes qui craignent de ne plus trouver facilement de médecins acceptant de leur prescrire la pilule qu'elles espèrent.

Toutefois, le RU 486 n'est pas la seule pilule abortive à laquelle les Américains peuvent avoir accès. Depuis quelques années, on prescrit une association de deux molécules : le méthotrexate (une substance anticancéreuse) et le misoprostol (mis sur le marché pour lutter contre les ulcères), utilisés seuls ou en association.

Malgré tout, l'autorisation de commercialisation de cette pilule abortive aux États-Unis suscite un grand espoir. Le RU 486 appartient à la famille des antiprogéstérone, mais l'étude des autres composants de la famille n'avait pas été possible en raison des blocages politiques constants qui avaient entravé le projet. L'homologation américaine élargit la classe des antiprogéstérone et il va désormais être possible de développer de nouveaux produits ayant éventuellement d'autres indications. Le RU 486 lui-même semble avoir des propriétés autres que celles pour lesquelles il est commercialisé : il serait efficace contre certaines tumeurs du cerveau et dans des maladies gynécologiques, tels les fibromes de l'utérus.

André ULMANN est le Directeur général du Laboratoire HRA Pharma.



La stratégie du penalty

IVAR EKELAND

Lors du tir des penaltys, le tireur et le gardien de but utilisent, consciemment ou inconsciemment, des tactiques «économiques». Celles-ci, pour être imprévisibles, sont fondées sur le hasard.

Le penalty est le jeu le plus simple qui soit. Le tireur doit frapper la balle sans aide de feinte, au terme d'une course rectiligne, et le gardien ne doit pas quitter sa ligne de but avant que la balle ait été frappée. Comme celle-ci va à plus de 150 km/h, elle met moins de 2/10 de seconde pour franchir les 11 mètres qui la séparent du but, et le gardien n'a pas le temps de réagir à la direction qu'elle prend. De son point de vue, le jeu consiste à deviner avant la frappe de quel côté la balle va partir et à s'élancer dans cette direction dès que le tireur touche la balle. S'il part du mauvais côté, le penalty est marqué, sans que le sort du cadre ; s'il part du bon côté, il a une chance d'arrêter le tir.

Les enjeux sont importants : la dernière Coupe de France a été remportée sur penalty. Lors de la dernière Coupe d'Europe des Nations, c'est sur penalty que se sont décidées l'élimination des Pays-Bas face à l'Italie et la qualification de la France face au Portugal.

Aussi, lors d'un penalty, le tireur et le gardien font attention à ce qu'ils font. Ce sont des professionnels, c'est un des actes les plus importants de leur métier. Ils s'y entraînent plusieurs fois par semaine, et ils connaissent parfaitement bien les qualités et les défauts de leur adversaire, soit qu'ils l'aient déjà eu en face d'eux dans des circonstances semblables, soit qu'ils l'aient étudié avec leur entraîneur sur des enregistrements vidéo. Bref, le penalty est un jeu simple, joué par des professionnels pour des enjeux importants. Peut-on l'analyser du point de vue théorique ?

Il existe une théorie des jeux, capable de modéliser mathématiquement des genres de situation. On lui fait le même reproche qu'à la théorie économique, c'est-à-dire de supposer que les joueurs sont de purs mathématiciens, infiniment calculateurs et parfaitement rationnels, bref de représenter un *homo economicus* qui serait une espèce distincte (et peut-être plus sympathique) que l'*homo sapiens* qui a envahi notre planète.

Jouons le jeu, regardons d'abord ce que font nos mathématiciens, et, après, nous les comparerons à Zidane et à Barthez.

Nos footballeurs idéaux ont des qualités personnelles : certains tirent mieux, ou arrêtent mieux, que les autres. Ils sont aussi dissymétriques : 85 % des tireurs sont droitiers et 15 % gauchers. D'ailleurs, plutôt que du côté droit ou du côté gauche, il vaut mieux mentionner le côté naturel du tireur, celui vers lequel il est le plus efficace. Le tireur a donc le choix entre son côté naturel (pour un droitier, le côté gauche, le côté opposé, et tout simplement le centre, sous la barre transversale. De même, le gardien a le choix entre le côté naturel du tireur (qu'il connaît), le côté opposé, ou rester sur place. Toutes les combinaisons de choix sont possibles et aboutissent à des probabilités de marquer différentes (qui sont propres au couple tireur/gardien considéré). Ainsi, si le tireur sait que le gardien va sauter à droite, il a toujours intérêt à tirer à gauche même si le côté droit est son côté naturel ; en d'autres termes, l'avantage de la surprise est toujours plus important que la puissance du tir.

Que dit alors la théorie ? Elle dit que le jeu possède un équilibre en «stratégies mixtes». Traduit en français, cela signifie que le gardien doit tirer au sort et l'autre calculer des probabilités, p_n , p_o et p_c pour le gardien, q_n , q_o et q_c pour le tireur, et qu'ils vont tirer au sort, chacun de son côté, en arrivant sur le terrain : le gardien choisira le côté naturel du tireur avec la probabilité p_n , le côté opposé avec la probabilité p_o , et le centre avec la probabilité p_c . De même pour le tireur, en remplaçant les p par des q .

Il peut être étonnant, à première vue, que des personnages supposés infiniment intelligents et calculateurs s'en remettent au sort pour décider ce qu'ils vont faire, quel un vulgaire dieu qui jouerait avec des dés, mais un instant de réflexion montre que c'est ce qu'ils ont de mieux à faire, par toute décision plus précise serait autodestructrice : si, par exemple, le tireur disposait d'un raisonnement ou d'un algorithme infallible qui lui dise qu'à coup sûr, compte tenu de tous les paramètres du problème, aujourd'hui c'est du côté naturel qu'il faut tirer, le gardien, par hypothèse tout aussi intelligent et cal-

culateur que le tireur, et en possession des mêmes informations, reproduirait ce raisonnement ou cet algorithme et sauterait de ce côté-là le moment venu, réduisant donc à néant l'efficacité de la prédiction.

Pour trancher cette circularité, il faut que, jusqu'au dernier moment, le tireur ne sache pas de quel côté il va shooter, ce que l'on exprime mathématiquement en disant qu'il tire au sort. En revanche, il peut parfaitement calculer les probabilités qu'il aura en moyenne le meilleur résultat : le gardien pourra reproduire son calcul, et donc connaître ses probabilités, mais il ne pourra pas prédire le résultat, car il a décidé de manière certaine du côté où la balle va partir aujourd'hui. Le calcul des probabilités a été fait par P. A. Chiappori, professeur d'économie à l'Université de Chicago, et S. Levitt, et ils ont montré que $p_c < q_c$ et $p_n > q_n$.

Ainsi un tireur parfaitement rationnel choisirait le centre plus souvent que le gardien et un gardien parfaitement rationnel choisirait le côté naturel plus souvent que le tireur. Remarquons que cette conclusion est vérifiable : il suffit de voir tous les penalty tirés dans les championnats français (1997-1999) et italien (1997-2000) de première division, soit 459 penaltys tirés par 162 tireurs contre 88 gardiens. L'expérience confirme toutes les prédictions théoriques. Pourquoi ? Aurait-on enfin localisé l'*homo economicus* dans les buts de l'équipe de France ? Barthez nous cacherait-il ses talents de mathématicien ?

Pas nécessairement. Il y a d'autres possibilités. La première, c'est qu'il est tombé dans la marmite quand il était petit, que les bonnes probabilités sont inscrites dans son code génétique, et que c'est justement pour cela qu'il est un bon gardien. La deuxième, c'est qu'il ne calcule pas, mais qu'il est capable d'apprendre : c'est l'expérience qui lui a enseigné les bonnes probabilités, et qui fait que ce sont celles-là qu'il pratique. Bref, selon une formule que j'emprunte ici à mon ami J.M. Lasry, on fait une théorie pour des gens superintelligents, et elle marche très bien avec des gens ordinaires.



Tout au chocolat

HERVÉ THIS

Comment introduire du chocolat dans toutes les pâtisseries.

Noël, le réveillon du Nouvel An... Le chocolat s'impose, mais où l'introduire? Une pâte feuilletée au chocolat, par exemple? Le passionné de la brune fève sait que le chocolat contient du beurre de cacao, et il voudrait substituer ce dernier au beurre normal, pour confectionner la pâte feuilletée. Toutefois, il sait aussi que le chocolat est d'une dureté qui interdit la réalisation souhaitée. Quelques idées simples sur les changements d'état donnent la solution du problème et conduisent à une transposition de la majorité des recettes classiques de pâtisserie.

Pour réaliser de la pâte feuilletée, le pâtissier forme d'abord une «détrempe», en travaillant de la farine avec un peu d'eau, parfois du beurre. Puis il étale son bâton, et place dessus une couche de beurre qu'il a aminci en le travaillant un peu. Il replie les bords de la détrempe sur le beurre afin de l'envelopper complètement, puis il effectue des «tours» : six fois de suite, il allonge cette enveloppe beurrée au rouleau et la replie sur elle-même, de sorte que la pâte qu'il a finalement est composée d'une alternance de couches de pâte et de beurre.

Comment y mettre du chocolat? Le chocolat à croquer, malgré son contenu en beurre de cacao, ne peut être utilisé à la place du beurre, parce qu'il est trop dur. Le beurre de cacao, qui est la matière grasse du chocolat (on n'évoquera pas ici les débats à propos des législateurs qui autorisent l'introduction d'autres matières grasses dans le chocolat), est composé à 80 pour cent de trois composés : trois «triglycérides», dont la molécule a un squelette de glycérol (ce que l'on nomme couramment la glycérine) auquel sont liés des «acides gras» (l'acide palmitique, l'acide stéarique et l'acide oléique).

Cette composition explique les remarquables propriétés physiques du beurre de cacao. Si le beurre de cacao n'était composé que d'une seule sorte de molécules, il fondrait à une température fixe, tout comme l'eau fond à 0 °C sous pression normale. Toutefois, comme il est formé de plusieurs types de molécules, sa température de fusion

s'étale sur une gamme de températures, qui commence à -7 °C (pour certains triglycérides) et s'achève à 34 °C. Malgré tout, le beurre de cacao est une matière grasse remarquable, car 75 pour cent de ses composés fondent entre 20 et 34 °C et près de 50 pour cent entre 30 et 34 °C. Autrement dit, le chocolat n'est pas un corps pur, mais il n'en est pas très éloigné.

Cette caractéristique, qui est un avantage pour la consommation du chocolat en tablette gêne le cuisinier, qui fait généralement ses préparations à des températures de l'ordre de 20 °C seulement. D'où les difficultés, notamment, pour la préparation de la pâte feuilletée. Pour la résoudre, comparons le beurre de cacao au beurre, plus hétérogène : le chocolat est une plus mauvaise émulsion, plus «impur». L'idée n'est pas neuve : les cuisiniers ont l'habitude de fondre du chocolat avec du beurre, afin d'obtenir des préparations chocolatées plus molles. Cependant, on peut poursuivre l'idée : à de l'huile neutre que l'on chauffe, on ajoute du chocolat et se mêle parfaitement et modifie à volonté les propriétés de l'émulsion.

Cette idée conduit à la réalisation attendue, d'une pâte feuilletée au chocolat (notons qu'une autre solution aurait été l'ajout de poudre de cacao à la farine ou au beurre laitière), mais la systématique entamée conduit aussi à une transposition générale de tous les desserts.

Tous les desserts au chocolat! De la pâte feuilletée, on arrive simplement à de la pâte à fonder au chocolat, à de la pâte briée ou de la pâte sablée, toujours au chocolat : il suffit de remplacer le beurre par du chocolat dont la composition lipidique a été changée. De même, on fera de la pâte à savarin au chocolat, de la brioche au chocolat, de la pâte à choux au chocolat, de la pâte à biscuit à l'amande et au chocolat sans farine... Parfois, on parviendra même à mettre le chocolat où on ne l'attend pas, comme dans une crème amandine au chocolat : à du chocolat fondu, on ajoute de la poudre d'amande, un alcool, puis on mêle de la crème pâtissière (obtenue par cuisson de jaunes d'œufs, de sucre de lait et de farine).

Dans d'autres cas, où la matière grasse utilisée est la crème, plutôt que le beurre, la stratégie devra être différente : c'est la crème qui sera remplacée par une préparation au chocolat. Cette fois la manipulation du point de fusion du chocolat ne suffit pas, car la crème est avant tout une émulsion, c'est-à-dire une dispersion de gouttelettes de matière grasse dans de l'eau (l'eau du lait) : les gouttelettes restent durablement (mais pas indéfiniment) séparées, parce qu'elles sont entourées de molécules «tensioactives», avec une partie qui est hydrophile (elle se place dans l'eau) et une partie hydrophobe (qui vient au contact de la matière grasse).

Pour remplacer la crème, le cuisinier devra confectionner une émulsion de chocolat... ce qui n'est pas difficile : dans une casserole, placez une solution aqueuse (eau, café, thé, Cognac...), puis chauffez en ajoutant des carrés de chocolat ; une émulsion de chocolat se forme. Cette «béarnaise au chocolat» est l'équivalent de la crème.

Enfin, pour confectionner une bavaroise au chocolat, une nouvelle opération s'avère nécessaire. Partons de la recette classique : on confectionne une crème anglaise, en fouettant du sucre en poudre avec des jaunes d'œufs, puis on ajoute du lait et on cuit la préparation ; on y dissout de la gélatine, et l'on incorpore enfin de la crème fouettée.

Où mettre le chocolat? Les jaunes d'œufs ne peuvent être remplacés, car c'est leur coagulation qui donne à la crème anglaise sa texture : les agrégats d'œuf cuit sont en suspension dans l'eau du lait. La gélatine, non placée, ne peut être remplacée par du chocolat. La crème, alors? Il s'agit, dans cette recette, de crème fouettée. Qu'à cela ne tienne : nous fouetterons une émulsion de chocolat et obtiendrons une mousse nommée chocolat Chantilly (voir *Pour la Science*, n° 230, décembre 1996). Et c'est cette mousse de chocolat qui sera ajoutée à la préparation. A vos tablettes!

Prochain rendez-vous France Info et Pour la Science le 27 décembre 2000, avec la chronique Info Sciences de Marie-Odile Monchicourt.

■ POUR LA SCIENCE

VOTRE ABONNEMENT 100 % NUMÉRIQUE

1 an au prix de 48 €

~~soit 4 € seulement par numéro au lieu de 5,20 €~~

Pour consulter et télécharger
en toute liberté sur www.pourlascience.fr
les 12 prochains numéros
~~dès leur parution.~~



Oui, je m'abonne
à Pour la Science 100% numérique
1 an (12 numéros) pour 48 €

~~Je préfère découvrir toutes~~
~~les offres d'abonnement disponibles~~
en cliquant ici.

Un nœud un peu gordien

Nous avons reconstitué un nœud en accord avec le mythe du nœud gordien tranché par Alexandre le Grand.

Jusqu'en 334 avant notre ère, en Phrygie, un temple érigé par le roi Midas et dédié à Zeus abrite un nœud noué par Gordius, le premier roi de Phrygie. Ce nœud est célèbre à travers le monde : une prophétie promet à celui qui le dénoue l'empire d'Asie ; ce fût Alexandre le Grand qui le trancha d'un coup d'épée.

Doit-on voir dans cette solution expéditive le tempérament radical d'un soldat ou plutôt, le fruit de la réflexion d'un ancien élève d'Aristote ? En d'autres termes, pouvait-on dénouer le nœud gordien sans le trancher ? Nous proposons ici quelques hypothèses et un exemple de nœud qui s'accorde avec la légende.

Que nous dit l'histoire ? Le nœud gordien, à l'abri dans son temple, a été une attraction très popu-

laire. Les nombreux prétendants, alléchés par la prophétie, n'ont pas réussi à dénouer le nœud.

Tous les nœuds fabriqués à partir d'une corde non fermée peuvent être dénoués, c'est-à-dire que, quelle que soit leur complexité, ils peuvent être transformés en une corde libre. Il s'agit d'une question de temps et de patience. Or la légende du nœud gordien n'indique aucune limitation dans le temps : on en déduit qu'il était fermé ou encore que les extrémités de la corde étaient jointes. En outre, l'historien grec Arrien (95-175), dans son ouvrage sur l'expédition d'Alexandre, indique qu'aucune extrémité du nœud n'était visible.

La configuration la plus simple d'un nœud fermé est une boucle. Dénouer un nœud revient alors à le transformer en une boucle simple. Cependant, certains nœuds fermés, dits non triviaux, ne peuvent être dénoués sans être coupés.

En mathématiques, cette observation est à la base de la classification des nœuds :

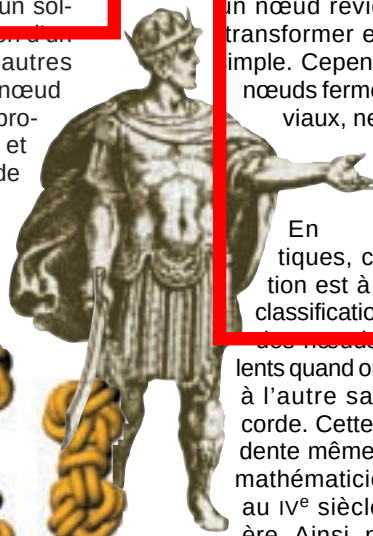
des nœuds sont équivalents quand on passe de l'un à l'autre sans rompre la corde. Cette propriété évidente même pour un non-mathématicien l'était déjà au IV^e siècle avant notre ère. Ainsi, nous sommes à la recherche d'un nœud équivalent à une boucle simple. Pourtant, le nœud

imaginé par Gordius ne devait pas être facile à dénouer. Nous pensons que c'était une boucle simple, enchevêtrée de telle façon qu'elle ne pouvait être dénouée par de simples manipulations.

Ces différentes propriétés nous conduisent à proposer un exemple de ce que pourrait être le nœud gordien (voir la figure) : une conformation équivalant à une boucle simple, mais que l'on ne peut pas transformer en une boucle simple sans étirer la corde dont est fait le nœud. Une corde fermée est pliée et nouée de telle façon qu'elle forme une extrémité de la boucle un domaine volumineux par lequel on fait passer la boucle terminale restante. Puis, la corde est rétrécie par un traitement, telle une désynchronisation, qui réduit sa longueur et son élasticité. Au total, les deux boucles terminales ne sont plus assez longues pour repasser au-dessus des domaines terminaux : l'épée s'impose !

À l'inverse, lorsque le nœud n'est pas assez serré ou si la corde est élastique, les boucles terminales peuvent repasser par dessus les domaines terminaux. Par ailleurs, des simulations sur ordinateur et des manipulations avec des cordes réelles ont montré que des domaines composés chacun d'un unique nœud simple ne retiennent pas les boucles. Par leurs configurations, ces domaines sont quasi sphériques et les boucles terminales n'ont pas besoin d'être étirées pour repasser par dessus (voir la figure). Dans le cas d'un nœud lâche, l'épée n'est plus nécessaire et l'Asie aurait peut-être été gouvernée par un homme plus pacifique.

Piotr PIERANSKI, École polytechnique de Poznan, Pologne
Andrzej STASIAK, Université de Lausanne



à la recherche d'un nœud équivalent à une boucle simple. Pourtant, le nœud



UN NœUD qui pourrait être le célèbre nœud gordien est composé de deux domaines volumineux constitués chacun de l'extrémité d'une corde fermée nouée plusieurs fois ; les boucles terminales sont passées par dessus ces domaines avant que la corde ne soit rétrécie (en haut). Cette opération empêche les boucles de repasser en sens inverse : seule une épée peut dénouer un tel nœud. En revanche, des domaines composés d'un unique nœud simple sont insuffisants pour retenir les boucles terminales (en bas). Si tel avait été le nœud gordien, Alexandre aurait pu le défaire sans le trancher (ci-contre).



AKG, Paris

La tortue des steppes

Pour retrouver son domaine, elle se repère à l'odorat.

Les mécanismes d'orientation des animaux révèlent d'ingéniosité : la boussole solaire des abeilles, l'orientation stellaire ou magnétique des oiseaux migrateurs... Se repérer au mieux pour optimiser l'exploitation des ressources est une question de survie. Les tortues des steppes (*Testudo Horsfieldii*) manifestent ainsi des capacités d'orientation étonnantes, sans doute grâce à l'odorat.

Ces petites tortues (20 centimètres de longueur et 1 kilogramme en moyenne) très répandues en Asie centrale, sont soumises à des températures extrêmes : 35 °C en été et -15 °C en hiver. Dès juin, quand les chaleurs de l'été arrivent, les tortues des steppes creusent le sol sablonneux et s'enterrent à 50 centimètres de profondeur. Elles ne sortiront de leur torpeur qu'en mars, quand les températures remontent. Ce sont des animaux ectothermes (leur température varie selon l'environnement, contrairement aux mammifères, endothermes, dont le métabolisme maintient la température corporelle), qui perdent moins de dix pour cent de leur masse pendant les neuf mois passés sans boire ni manger. À leur sortie, les tortues ne disposent que de trois mois pour se nourrir afin d'emmagasiner des réserves et pour se reproduire. D'autres espèces, par exemple les vipères péliades, qui vivent au-delà du cercle polaire, en Suède, ont une période d'activité aussi restreinte.



La tortue des steppes, qui vit en Asie centrale, passe neuf mois de l'année sous terre, en état d'hibernation, puis d'estivation. Pendant ses trois mois d'activité, elle optimise ses déplacements et se repère grâce à un mécanisme d'orientation très efficace, fondé sur l'odorat.

La période des accouplements des tortues des steppes dure seulement 15 jours, et les femelles pondent jusqu'à trois fois par saison. Pressées par le temps, les tortues font preuve d'une excellente connaissance de leur terrain. Les études menées depuis 1996 en Ouzbékistan ont montré que les tortues possèdent en effet un domaine vital vaste pour un animal de cette taille : une vingtaine d'hectares pour les mâles et une quarantaine pour les femelles. D'une année sur l'autre, elles restent fidèles à leur espace. Disposent-elles d'une capacité à se repérer dans l'espace, qui leur permettrait d'optimiser l'efficacité de leurs déplacements ?

Pour répondre à cette question, deux expériences ont été poursuivies durant l'été 2000. D'abord, 19 tortues ont été équipées d'un émetteur de radiofréquence, puis transportées dans des sacs hermétiques à 1 000 mètres hors de leur domaine vital. La position des tortues a alors été suivie jour après jour et leur trajet reconstitué. Trois jours plus tard, 18 tortues étaient déjà revenues sur leur lieu de capture, avec une vitesse moyenne de 85 mètres par heure, tandis que leur vitesse de déplacement habituelle ne dépasse pas 20 mètres par heure. Le jour croît, le retour a été précis : les tortues n'ont parcouru que 1 383 mètres en moyenne au lieu des 1 000 mètres du trajet direct.

Comment ces tortues relâchées en terrain inconnu ont-elles retrouvé le chemin de leur domaine ? L'orientation magnétique semble le mécanisme d'orientation principal chez les tortues marines bien que des possibilités d'orientation olfactive aient été invoquées pour certaines tortues terrestres. Pour trancher, nous avons répété l'expérience, mais cette fois après avoir collé de petits aimants sur la surface de 10 nouveaux animaux, afin de créer un « brouillage magnétique ». Les aimants perturbent localement le champ magnétique terrestre et désorientent les minuscules cristaux de magnétite présents dans le cerveau, comme on l'a montré chez les animaux capables de percevoir le champ magnétique terrestre. Ces animaux sont alors « déboussolés ».

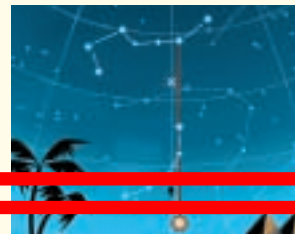
Le résultat fut sans appel : 12 des 13 tortues sont rentrées dans leur domaine sans difficulté et aussi rapidement qu'en l'absence de brouillage magnétique. L'hypothèse de l'orientation magnétique est ainsi éliminée. Les tortues des steppes retrouvent-elles leur chemin grâce à l'odorat ? Il reste désormais à tester cette hypothèse, après avoir mis au point des méthodes pour masquer l'olfaction, sans perturber les réactions naturelles des tortues.

A. LEGRAND et D. CHARFALLOT,
Centre d'études biologiques de Chizé

BRÈVES

L'âge des pyramides

Comment les grandes pyramides de Gizeh ont-elles été orientées vers le Nord avec une telle précision ? Selon l'égyptologue britannique Kate Spence, les Égyptiens se fondent sur le passage simultané à la verticale de deux étoiles : Kéchéb de la petite Ourse et Mizar de la grande Ourse. La déviation observée dans l'orientation des pyramides successive serait due à l'évolution de la carte céleste au fil des siècles, résultant du déplacement de l'axe de rotation de la terre. Si l'hypothèse est vérifiée, les astronomes pourraient calculer l'âge de ces pyramides à cinq ans près.



Aiguilles sur mesure

Pour qu'un vaccin intramusculaire soit efficace, le liquide injecté doit atteindre, comme son nom l'indique, un muscle. Or, une équipe américaine vient de constater que ce n'est pas le cas chez les personnes trop grosses : sur les 220 personnes vaccinées qui ont été suivies, 17 pour cent des hommes et près de 50 pour cent des femmes n'ont pas été correctement vaccinés, car l'aiguille standard (16 centimètres de longueur) n'a pu atteindre le muscle : la solution vaccinale avait été injectée dans la couche de graisse qui recouvre le muscle. Cette dernière est peu vascularisée, de sorte que l'antigène vaccinal y stagne sans activer le système immunitaire. Des aiguilles plus longues devraient être utilisées pour les personnes « enrobées ».

Les turbots respirent par la peau

Les turbots obtiendraient 30 pour cent de leur oxygène par l'intermédiaire de leur peau d'après une étude sur les poissons marins d'élevage de l'IFREMER. En effet, la réfraction de l'oxygène dans l'eau, l'apport d'oxygène par les branchies et la fréquence des mouvements qui facilitent le passage de l'eau dans les branchies. L'appareil respiratoire cutané prend alors le relais : la peau du turbot, mince et dépourvue d'écaillés, laisse diffuser l'oxygène. La peau de tous les poissons plats, tels la sole ou le flet, jouerait aussi le rôle d'organe respiratoire annexe.

L'origine de la cocaïne

Afin de lutter contre le trafic de la cocaïne, une équipe de l'Université de l'Utah et du département américain de la justice a mis au point une méthode pour retrouver le pays où a été cultivée la drogue ; il leur suffit de quelques grammes. Pour ce faire, ils ont analysé les isotopes du carbone et de l'azote dans des feuilles de coca provenant de cinq régions productrices d'Amérique du Sud. Les quantités de carbone 13 et d'azote 15 diffèrent selon les régions. Pour l'azote, les rapports isotopiques varient en fonction de la température, de la longueur de la longueurs de la saison des pluies. Les traces d'isotopes renseignent aussi sur l'origine des feuilles. On peut ainsi décrire les « cartes d'identité » des feuilles provenant des principaux sites de culture. De cette façon, les chercheurs ont retrouvé les régions d'origine de 192 des 200 échantillons analysés.

L'orbite de Néandertal

L'individu dont un fragment de crâne a été retrouvé en 1856 à Néandertal, près de Düsseldorf, a donné son nom à toute une espèce d'hominidés : les Néandertaliens. Le lieu de la découverte n'avait jamais été fouillé de façon systématique, jusqu'à ce



J. Thissen et R. Smitz

des archéologues allemands reprennent les fouilles. Sensation ! Un morceau de l'orbite de l'individu de 1856 se trouvait parmi les nombreux restes rassemblés. Le visage du premier Néandertalien va enfin être reconstitué.

Le dîner du dinosaure

Les dinosaures herbivores, notamment les diplodocus, ne dévoraient pas la cime des arbres. L'australien Roger Seymour a montré que ces grands mangeurs de feuilles ne pouvaient pas dresser leur cou pour atteindre les plus hautes branches : leur cœur ne leur permettait pas ; il était trop petit pour pomper le sang jusqu'à la tête, si les animaux avaient redressé leur cou. En comparant la taille du cœur de différentes espèces en fonction de la longueur du cou, il en a déduit que, pour relever la tête, il aurait fallu que le ventricule droit de leur cœur pèse... deux tonnes. Ils devaient se contenter des feuilles à la hauteur de leur tête, quand leur cou était à l'horizontale.

Culte celtique à Mayence

Des fouilles de sauvetage ont révélé des pratiques magiques dans un lieu de culte domestique romain.

Au II^e siècle, une femme trompée par son mari se rend au lieu de culte domestique de Mogontiacum, une ville rhénane de la frontière nord de l'empire romain, devenue Mayence aujourd'hui. Après s'être procurée une statuette chez un marchand, elle s'enferme dans l'une des maisonnettes de prière, allume des lampes à huile, brûle de l'encens sur l'autel en terre, puis y dispose des nourritures précieuses : figues et dattes séchées. Péniblement, car devant le nom d'une divinité celte, elle brise la statuette par le milieu, puis place la partie inférieure dans un trou, phallus vers le haut pour bien montrer aux dieux sur quoi s'abaisse sa vengeance. Pour renforcer encore sa malédiction, elle couche ensuite la partie supérieure de la statuette le nez dans la boue, comme on enterrait alors les criminels condamnés.

Les archéologues ignorent qui fut ainsi maudit, mais ils ont retrouvé chacun des éléments de cette scène. Au printemps 2000, une fouille de sauvetage a été menée dans le vieux Mayence. La mise au jour d'un lieu de culte domestique a révélé comment priait le peuple de cette ville de garnison romaine. L'endroit se trouve non loin de l'ancienne rue qui reliait l'un des camps militaires au pont sur le Rhin. Pour accéder aux vestiges romains, les archéologues ont évacué six mètres de remblais situés sous les restes d'un monastère moyenâgeux.

Mayence est l'une de ces cités, où l'on ne peut faire un trou sans découvrir des vestiges. La ville a été fondée sous l'empereur Auguste au cours de l'invasion romaine (-27 à 14). Afin de se garder des attaques sporadiques des germains qui vivaient au Nord du fleuve, les Romains y érigèrent progressivement un mur fortifié de 550 kilomètres de long : le limes. Il allait des troupes pour le garder et des bases arrière. En deçà du fleuve, le site de Mayence était stratégique, puisque placé face au confluent du Main et du Rhin. Jusqu'à 20 000 combattants, c'est-à-dire quatre légions y furent cantonnées. Comme le nouveau camp accueillait presque autant de troupes de soutien, il concentrait un énorme pouvoir d'achat, qui bientôt attira, marchands, artisans, tenanciers d'au-

bourg constituées... Mayence était née. La ville eut bientôt son théâtre, son aqueduc, son lieu de culte...

Le grouillant de vie, ce dernier se composait d'une série de fosses maçonnées, où les adorateurs célébraient leurs dieux. Des restes de pieux en bois attestent que ces fosses étaient probablement recouvertes d'une maisonnette à colombage. Une guilde de petits marchands s'y était fixée, qui servait les citoyens venus prier, leur vendait des ustensiles de culte, les approvisionnait... Aucune trace de temple n'a été retrouvée à proximité du lieu de culte domestique.

Nul ne sait quels dieux y étaient honorés. Trouvée exceptionnelle, une statuette masculine (*ci-dessous*) a été découverte, laquelle porte des traces manifestes de torturer. Une marque que les habitants pratiquaient des malédictions. S'en était-il là d'un reste de magie celte ? Nul ne le sait. Les archéologues ont dégagé une dizaine de tablettes de plomb. Les adorateurs romains y inscrivaient leur suppliques avant de les rouler et de les cacher dans un lieu sacré. En voie de restauration, elles n'ont pas encore été déroulées.

Une chose intriguait les Mayençais : pourquoi les anciens habitants de Mayence ont-ils consacré à leurs rites un endroit aussi central de la ville ? Un mystère que les fouilles ont éclairci en révélant la présence sur le lieu de culte de deux tumulus (« tertres celtiques »). Une sensation, car personne n'avait encore la preuve que les Celtes occupaient le site avant les Romains. L'hypothèse semblait pourtant bonne puisque le nom Mogontiacum dérive de celui du dieu celte de la lumière : Mogan. La majorité des habitants de la ville était probablement des celtes romanisés, qui respectaient les cimetières de leurs ancêtres.

Andrew COLE et Rainer FAHR



Lors des fouilles récentes d'un lieu de culte de la ville romaine de Mogontiacum (Mayence), la trouvaille la plus intéressante fut cette statuette ; cette dernière avait été brisée indiquant peut-être la pratique d'une sorte de malédictions, de magie celte.

Le chant d'amour des drosophiles

Pour séduire leur partenaire, les drosophiles mâles chantent avec leurs ailes.

Les drosophiles mâles séduisent les femelles... en chantant. Le chant de cour grésillement tenu inaudible pour l'oreille humaine, ne résulte pas d'une improvisation romantique mais suit, au contraire, un rythme précis et propre à chaque espèce. Depuis les années 1960, de nombreuses équipes, notamment celle de H. Bennet-Clark, en Angleterre, s'intéressent à la communication acoustique chez ces mouches. Pour «chuchoter des mots doux», le mâle déploie une aile – parfois les deux – et la fait vibrer. La femelle «écoute» par l'intermédiaire d'un segment d'antenne sensible aux variations de pression de l'air. Quel est le rôle du signal acoustique dans la parade nuptiale ? Fanny Rybak, Bruno Moulin et Thierry Aubin, du laboratoire de neurobiologie de l'apprentissage et de la mémoire du CNRS, à la Faculté des sciences d'Orsay, ont mis au point une technique d'enregistrement du chant de cour, et montré que la communication acoustique joue un rôle essentiel dans la reconnaissance des partenaires.

Le rituel de la parade nuptiale de la drosophile qui dure de une à dix minutes, est très précis et comporte divers signaux, visuels, tactiles, chimiques et acoustiques. Le mâle commence par s'orienter vers la femelle, puis, avec une patte antérieure pourvue de récepteurs gustatifs, se renseigne sur son identité, en touchant l'abdomen qui porte des substances chimiques, les phéromones. Il écarte alors une aile et la fait vibrer. Après quelques refus, la femelle finit par écarter ses plaques génitales et accepter l'accouplement.

Le chant de cour du mâle est constitué de deux types de signaux distincts, le chant pulsé et le chant sinusoïdal. Le chant pulsé est une succession d'unités sonores simples émises en rafale. C'est toujours la même unité qui est répétée, une douzaine de fois au plus. L'intervalle qui sépare deux unités est caractéristique de l'espèce : il est égal à

soixante millisecondes chez *Drosophila melanogaster*, et à 56 millisecondes chez *Drosophila simulans*, par exemple. Le chant sinusoïdal est une onde sinusoïdale régulière, une sorte de bourdonnement. Les deux signaux peuvent être émis isolément, successivement ou, parfois même, simultanément.

Le chant de cour ne porte qu'à très courte distance, tout juste à quelques millimètres. Au laboratoire, les deux partenaires sont placés dans une petite chambre acoustique bien isolée : l'équipe d'Orsay enregistre, d'une part, les bruits émis par le couple de drosophiles et le bruit de fond et, d'autre part, uniquement ce bruit de fond. La soustraction des signaux des deux enregistrements simultanés permet à F. Rybak et à T. Aubin d'isoler les signaux spécifiquement émis par le couple. Ils ont ainsi confirmé que les femelles sont presque insensibles aux variations de fréquence des deux types de chants, mais qu'elles sont très sensibles à l'intervalle de temps qui sépare les unités sonores à l'intérieur d'une même rafale du chant pulsé et à la durée de chaque rafale, c'est-à-dire au nombre d'unités sonores par rafale.

Toutefois, une question restait ouverte : lors de la parade nuptiale, l'attitude des femelles est-elle dictée plutôt par les signaux acoustiques (le chant) ou par les signaux chimiques (les phéromones) ? Pour dissocier odeurs et chant, l'équipe d'Orsay a réussi à obtenir des mâles muets et des mâles sans odeurs. S'il est aisé d'obtenir des mâles muets – il suffit de leur couper les ailes –, l'élimination des substances responsables des odeurs était plus difficile.

On a montré que le gène «transformeur», qui participe à la différenciation sexuelle et à la production des phéromones. Son expression peut être inhibée à volonté lorsqu'il est associé à un gène codant une protéine de choc thermique. Ce dernier n'existe pas naturellement dans le génome des drosophiles. F. Rybak, Gilles Sureau et T. Aubin l'ont inséré dans le génome des parents des mâles utilisés pour l'expérience : en faisant subir aux drosophiles un choc thermique à 37 °C, on déclenche l'expression de la protéine de choc thermique, laquelle inhibe la synthèse des phéromones. Ils ont alors obtenu des mâles sans odeurs, mais qui ont une parade normale, chantent et s'accouplent.

Les expériences ont révélé que les femelles sont plus sensibles aux signaux acoustiques qu'aux signaux chimiques. La communication acoustique semble être un mécanisme clé de reconnaissance de l'espèce, qui évite les accouplements entre espèces différentes.

Marie-Thérèse LANDOUSY

Éclaircies sur Titan

Bulletin météo sur Titan : importants passages de nuages de méthane.

Une partie du voile vient de se lever sur Titan, la plus grosse lune du système solaire jusqu'à présent masquée par d'épaisses brumes orange, opaques en lumière visible. Alain Griffith, de l'Université d'Arizona, a observé son atmosphère grâce au télescope infrarouge Keck, à Hawaii : des gros nuages s'y forment, y circulent, et se dissipent en quelques heures. Il pleut sur Titan, mais les pluies qui s'y abattent sont des poies de méthane !

L'atmosphère de Titan se compose d'azote (95 pour cent), d'argon, d'hydrogène et d'hydrocarbures, ce qui la rend similaire à l'atmosphère de la Terre primitive. Elle contient en outre deux à quatre pour cent de méthane, ce qui intriguait les astronomes. En effet, Titan est bombardée de photons, dont la densité est cent fois supérieure à celle qui atteint la Terre. Ils sont toutefois assez nombreux pour décomposer le méthane en molécules et atomes, des aérosols qui retombent sur Titan. Le méthane aurait dû disparaître peu à peu de l'atmosphère, et les réserves devraient être épuisées depuis 4,5 milliards d'années d'après les estimations des astronomes ! Or ce n'est pas le cas.

La partie basse de l'atmosphère (la troposphère) contient dix pour cent de méthane et les couches supérieures environ deux pour cent. Par conséquent, les réserves de méthane atmosphérique se renouvellent. Comment ? La découverte de nuages montrerait l'existence d'un cycle du méthane, similaire au cycle de l'eau sur Terre. Le méthane s'évaporerait du sol, serait de lacs ou de réserves dans un sol poreux. Il se condenserait dans la troposphère et formerait des nuages à une altitude de 15 kilomètres. Puis, il pleuvrait.

Toutefois, les nuages se condensent sur la Terre grâce à des germes (pollens, sels...) bien sûr inexistant sur Titan. Les astronomes expliquent alors le cycle du méthane grâce à une autre hypothèse, proposée par Régis Courtin, à l'Observatoire de Meudon. Selon lui, les couches basses de l'atmosphère seraient saturées en méthane, et lorsque la tension de vapeur saturante serait atteinte, le méthane passerait de l'état gazeux à l'état liquide et de grosses gouttes retomberaient sur le sol. Enfin, selon d'autres astronomes, l'atmosphère de Titan serait tellement instable

CHANT SINUSOÏDAL

CHANT PULSÉ

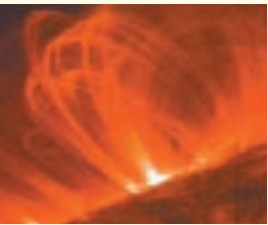
La structure du chant est un signe de reconnaissance chez les drosophiles.

Adam et Ève ne se sont pas croisés

Notre «père» ancestral aurait 80 000 ans de moins que notre «mère», d'après une étude génétique de l'équipe de Peter Oefner, de l'Université Stanford. L'ADN contenu dans de petits organites cellulaires nommés mitochondries est uniquement transmis par la mère, tandis que le chromosome sexuel Y est transmis de père en fils. En analysant ces deux marqueurs, les scientifiques ont retracé notre arbre généalogique. Ainsi, tous les êtres humains ont un ADN mitochondrial issu de l'Ève mitochondriale qui vivait il y a 140 000 ans. En revanche, l'Adam porteur du premier chromosome Y, dont est issu le chromosome sexuel de l'homme moderne, a vécu bien plus tard, il y a 50 000 à 70 000 ans.

La température de la couronne

Les astronomes savaient que l'atmosphère externe du Soleil, la couronne, est des centaines de fois plus chaude que la surface solaire, mais s'interrogeaient sur le mécanisme de chauffage. Des images



de cette couronne enregistrées par la sonde TRACE ont apporté des indices. Elles révèlent que chaque boucle

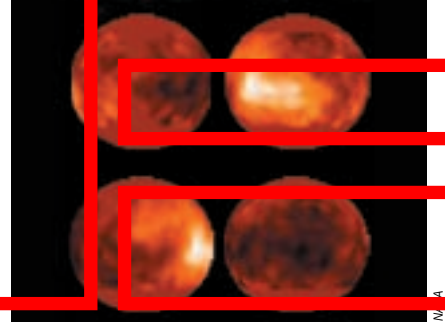
jaillissent et s'élèvent à des centaines de milliers de kilomètres) est constituée de nombreuses boucles fines. Contrairement à ce que l'on pensait, ces boucles seraient chauffées par leur base, près de la surface solaire, passant de 5 800 °C à plusieurs millions de degrés.

Synthèse vocale et accents

Un ordinateur saura-t-il prononcer Debreuil quand il lira *De Broglie*? La synthèse vocale est mise en défaut aussitôt qu'elle traite des noms propres, par exemple des noms de rues lors de la navigation assistée en voiture. Afin de tenir compte des particularités de prononciation, une Société belge la *Phonetic Topographics* a introduit une étape supplémentaire dans le procédé de synthèse vocale : elle fait prononcer directement les noms propres écrits en codes lus, puis énoncés par l'ordinateur, chacun des noms est transcrit en phonétique par des linguistes. On a demandé à des personnes de reconnaître des noms propres énoncés par l'ordinateur : avec le nouveau traitement, la reconnaissance des noms propres est notablement améliorée.

qu'elles s'effondreraient périodiquement avant de se recréer.

Le vaisseau *Cassini* devrait apporter des informations sur ce phénomène. Il a quitté la Terre le 15 octobre 1997 en direction de Saturne et transporte le module *Huygens* destiné à être parachuté à la surface de Titan en 2004. Outre la question de la présence de méthane dans l'atmosphère de Titan, c'est aussi celle de l'origine de la vie qui est en suspens. Même si les basses températures et l'absence d'eau liquide sur Titan ne semblent pas permettre l'apparition de la vie telle que nous la connaissons, Titan constitue néanmoins un



Vues de la surface de Titan photographié en infrarouge par Hubble. Les zones les plus brillantes, situées dans des régions de terrain élevé.

laboratoire naturel où s'élaborent et évoluent les molécules organiques complexes, qui, sur la Terre ont été le préambule à l'apparition de la vie.

Virus canin tueur de phoques

Un virus canin a été responsable de la mort de 20 000 phoques, l'été dernier.

Un virus canin, responsable de la maladie du jeune chien, la maladie de Carré, a causé la mort de 20 000 phoques de la mer Caspienne, l'été dernier. Comment ce virus s'est-il transmis? Comment a-t-il franchi la barrière qui sépare des espèces si éloignées? Est-il le seul facteur de l'épidémie? Autant de questions qui agitent la communauté des biologistes.

Les phoques de la mer Caspienne, *Phoca caspica*, étaient environ un million, au début du xx^e siècle; ils ne seraient plus aujourd'hui qu'environ 100 000. Répertoriés par l'organisation mondiale de la protection de la vie sauvage comme une espèce menacée, les phoques ont commencé à mourir par centaines dès le mois d'avril,

après leur hivernage au Nord de la mer Caspienne. Cette mer, la plus grande mer intérieure du monde, est aussi l'une des plus polluées : de fortes concentrations de substances toxiques telles que des métaux lourds, des pesticides (DDT, PCB) et des toxines issues de diverses activités industrielles et agricoles, y ont été enregistrées. À quel point que les biologistes ont d'abord accusé les usines qui bordent les côtes de la mer Caspienne. Cependant, les experts sont formels : la pollution ne peut être tenue pour seule responsable de la mortalité observée.

Une équipe conduite par le biologiste irlandais, Seamus Kennedy a étudié des échantillons de tissus prélevés sur 16 phoques morts : les examens ont révélé la présence d'un morbillivirus, dont le plus connu est celui de la rage. Une étude approfondie a révélé que le génome du virus ravageur est identique à celui du virus de la maladie de Carré, appartenant au genre morbillivirus. Les morbillivirus connus infectent les cellules épithéliales, les cellules nerveuses et les cellules du système immunitaire des animaux.

De précédentes contaminations de phoques par des morbillivirus provenant du chien domestique ont déjà été observées. Ces précédents, d'une moins grande ampleur, semblent indiquer la présence du virus à l'état latent dans la population des phoques de la mer Caspienne, ou des transmissions répétées à partir d'un «réservoir».

La sensibilité des phoques à ce virus a sans doute été favorisée par la pollution qui règne dans les eaux de la mer Caspienne et qui entraîne une baisse des défenses immunitaires augmentant leur vulnérabilité. Cette infection est un nouvel exemple non pas de l'émergence de nouvelles maladies, mais de la transmission d'anciennes maladies à de nouvelles espèces, qui ne sont pas préparées pour combattre ces micro-organismes.



Porcelaines vietnamiennes

La spectroscopie Raman révèle les techniques de fabrication des céramiques anciennes.

Les porcelaines vietnamiennes anciennes sont réputées pour leur finesse et leur fraîcheur, mais leurs techniques de fabrication restaient méconnues. Pour les retrouver, Philippe Colomban et ses collègues, du Laboratoire de dynamique, interactions et réactivité de Thiais en ont analysé des tessons ramassés au Nord du Vietnam et datés des XIII^e, XV^e et XVI^e siècles, d'après les données stratigraphiques des fouilles.

Après la fin de la domination chinoise, qui dure tout le premier millénaire de notre ère, un style original apparaît sur les porcelaines vietnamiennes, sous les dynasties Ly (1009-1225) et Tran (1225-1400). Les décors des jarres, plats, coupes et autres bols sont moins figés que ceux des pièces chinoises de la même époque et dégagent une impression d'harmonie. Comment ces porcelaines étaient-elles fabriquées ?

P. Colomban et ses collègues vietnamiens ont analysé des céramiques provenant de trois fours en activité au Nord du Vietnam, au cours des dynasties Ly et Tran : le four de Ha Lan, le four de Chu Dau et le four de Hop Lê. Les physico-chimistes ont utilisé la spectroscopie Raman. Cette méthode d'analyse est fondée sur la détection des modifications de la lumière diffusée par un objet éclairé par un laser monochromatique. Les lon-

gueurs d'ondes des raies de lumière diffusées sont caractéristiques des constituants de l'objet analysé et de leur état — amorphe ou cristallisé — imposé par les conditions de cuisson.

Les porcelaines analysées contiennent deux phases formées dans le four pendant la cuisson : la mullite (un aluminosilicate qui a cristallisé à partir du kaolin, l'argile réfractaire utilisée comme matière première) et du verre (le produit de la réaction entre du feldspath, du silicate et des cendres). Elles contiennent aussi du sable qui n'a pas fondu pendant la cuisson et qui était ajouté à l'argile (avant la cuisson). Le sable joue le rôle structural de l'eau en maintenant les feuillets de l'argile écartés, une fois que les molécules d'eau se sont évaporées, de sorte qu'il diminue la fissuration des pièces lors du séchage.

Parmi les échantillons de Ha Lan et de Hop Lê, certains sont des protoporcelaines, antérieures car leur cycle de cuisson ou la préparation du kaolin étaient imparfaits. Ce manque de maîtrise des techniques de fabrication se traduit par le fait que les protoporcelaines contiennent peu de mullite (qui se forment vers 1300 °C) et qu'elles contiennent encore des minéraux présents dans l'argile (probablement apportés par les alluvions), qui n'ont pas été complètement éliminés par la décantation. Progressivement, ces deux défauts ont été corrigés. À Ha Lan, la porosité augmente : plutôt que vers des porcelaines, les céramiques évoluent vers des faïences fines, plus légères et d'un blanc plus pur.

La plupart des porcelaines analysées sont recouvertes d'un émail vert, qui imite le jade. Cette coloration est due à la présence d'ions ferreux (bleus) et ferriques (jaunes), dont les proportions variaient en fonction du caractère oxydant ou réducteur de l'atmosphère du four, c'est-à-dire du réglage de l'alimentation en bois. La cuisson était lente (elle durait parfois une semaine), de sorte que des bulles microscopiques libérées par la réaction de l'émail sur le substrat restaient piégées à l'interface, diffusant la lumière et donnant une impression de profondeur.

Les physico-chimistes ont comparé les spectres Raman de l'émail d'une porcelaine ancienne et d'une copie moderne. Leur aspect diffère, car le premier indique que l'émail est entièrement vitreux, tandis que le second révèle la présence de petites cristaux que forment les pigments utilisés pour la coloration. La spectroscopie Raman est non destructive, rapide et depuis peu, réalisable avec un instrument portatif. Ainsi, elle autorise l'authentification des porcelaines anciennes et intéresse les archéologues, notamment pour étudier les pièces récemment mises au jour lors de fouilles en Mer de Chine.

La danse des liquides

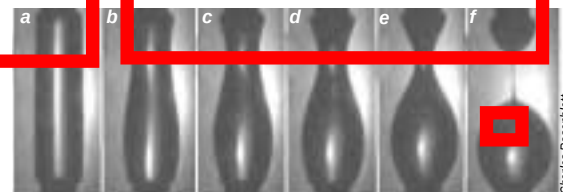
Des liquides magnétiques lèvent.

Comment étudier les phénomènes de microgravité à peu de frais ? En utilisant des liquides magnétiques. À l'Université Case Western, le groupe de Charles Rosenblatt fait varier la force gravitationnelle qui s'exerce sur ces milieux en les soumettant à une force magnétique qui s'oppose à la force de gravitation. Ainsi, on simule des conditions de gravité différentes, rencontrées lors des missions spatiales ou sur d'autres planètes.

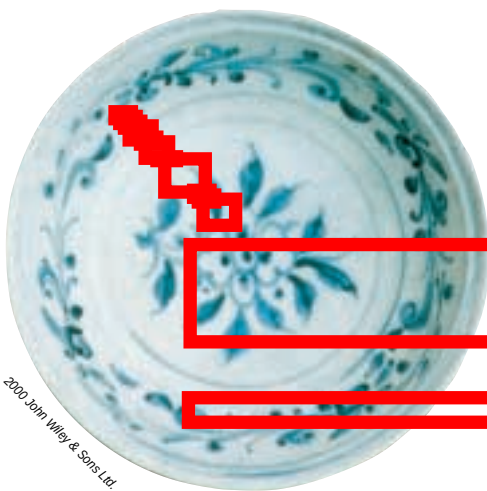
Quel soit sur la navette spatiale ou dans des avions en vol parabolique, les expériences en microgravité sont onéreuses. Peut-on obtenir les mêmes conditions, ou des conditions voisines, sur Terre, en laboratoire ? Sans la condition d'utiliser des liquides qui réagissent à un champ magnétique : on utilise des mélanges de glycérol et d'un liquide paramagnétique (le tétrahydure de chlorure de magnésium). Lorsque le mélange est soumis à un champ magnétique, les moments magnétiques élémentaires s'orientent parallèlement et confèrent à l'ensemble un champ magnétique intrinsèque.

Quand le champ magnétique appliqué est judicieusement choisi, il engendre un force sur le liquide qui, s'opposant à la force de gravitation, le place en position d'équilibre artificielle : le liquide lève. En faisant varier le champ magnétique appliqué, on simule des variations de la gravité avec le temps. Or, des variations temporelles de la gravité correspondent à des accélérations. Ainsi, au lieu de se contenter de l'étude statique des liquides magnétiques, les physiciens peuvent étudier leur dynamique.

On obtient une colonne de liquide — encore nommée pont liquide — suspendue entre deux points, comme si aucune force ne s'exerçait sur lui. En modulant le champ magnétique, on arrive même à faire danser le liquide avec un rythme dicté par la gravité et par les propriétés du



Un liquide magnétique est maintenu entre les pôles d'un électroaimant (a). Lorsque l'on coupe le champ magnétique, le pont liquide (la zone centrale étroite, s'effondre (de b à f).



La méthode de fabrication de cette porcelaine vietnamienne daté du XV^e siècle a été reconstituée grâce à son analyse par spectroscopie Raman.

fluide : le liquide entre en résonance. De ce mouvement, les physiciens déduisent les propriétés physiques du pont liquide, notamment sa tension superficielle, lorsque le liquide est en mouvement. Par cette méthode, on peut mesurer la tension de surface dynamique de solutions aqueuses contenant différentes molécules tensioactives.

Le contrôle doit être fin : si l'on fait trop varier la force magnétique, le pont s'effondre. En répétant cette expérience avec

diverses longueurs de pont et avec des forces magnétiques différentes, les physiciens ont montré que le temps que met un pont pour s'effondrer ne dépend pas directement de sa longueur.

Les conditions d'impesanteur obtenues ne sont ni moins ni plus que dans l'environnement spatial – il reste toujours une faible gravité résiduelle – mais ces méthodes sont bien moins onéreuses que les expériences spatiales et elles permettent aux chercheurs de simuler des

changements brusques d'accélération sans déplacer le fluide. Cette technique peut être utilisée pour simuler les effets des vibrations dues aux tremblements de Terre, le comportement des liquides à bord d'un vaisseau qui se pose sur une autre planète ou encore les fluides biologiques : la tension de surface des fluides qui tapissent les alvéoles pulmonaires varie au cours de la respiration et la méthode permet d'étudier les conditions d'une éventuelle rupture de ce film liquide.

La cartographie de l'espace

Les amas de galaxies se donnent rendez-vous aux carrefours de l'Univers.

Le satellite XMM-Newton de l'Agence spatiale européenne, qui scrute le ciel dans la gamme des rayons X, est en orbite depuis bientôt un an. Mis en place pour observer les phénomènes de haute énergie (où les températures sont élevées), tels que la physique des trous noirs, les restes de supernovae ou encore la nature du fond d'amas (l'émission continue de rayons X observé dans l'ensemble de l'espace interstellaire), il a récemment fourni ses premières données. Notamment, le satellite a fait des observations inédites sur les amas de galaxies, que l'équipe de Monique Arnaud et Philippe Ferrando, du Commissariat à l'énergie atomique de Saclay, a analysées (l'explication de l'ensemble

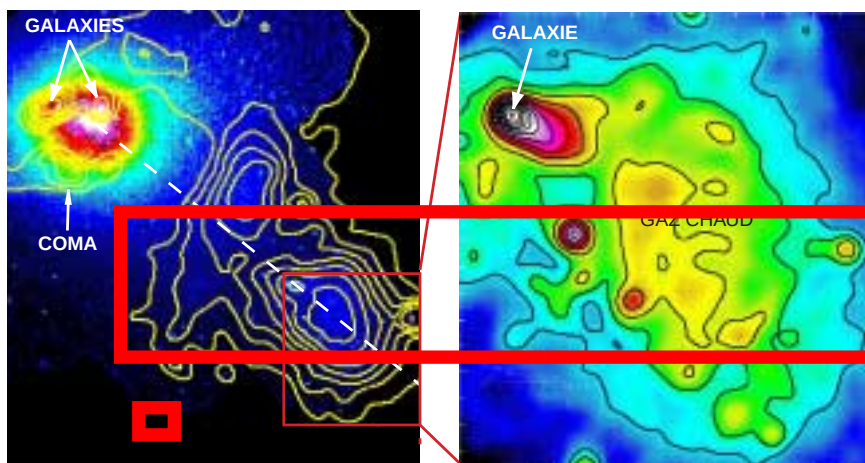
des données transmises par XMM-Newton occupe plus de 150 équipes de recherche dans le monde).

Les amas de galaxies sont les plus grandes structures de l'Univers. Ils sont constitués de quelques dizaines à plusieurs milliers de galaxies liées par leur propre attraction gravitationnelle. Ils se trouvent aux croisements de filaments, le long de quels la matière est répartie dans l'Univers, et qui forment une gigantesque toile d'araignée. Au cours du temps, la matière s'écoule le long des filaments et converge vers les intersections où se forment les amas de galaxies. Les galaxies émettent des rayons X, ainsi que le gaz chaud qui les entoure (dont la température atteint 10 à 100 millions de degrés).

Le satellite XMM-Newton a observé l'amas de galaxies Coma, situé dans notre «voisinage», à 400 millions d'années-lumière. Coma est un grand amas, dont la masse est 20 000 fois supérieure à celle de la Voie Lactée. Les rayons X, qui renseignent sur la température dans l'amas, ont révélé de grandes différences thermiques en son sein. Ces

ci indiquent que l'amas est le siège de vastes mouvements de matière. Pour les comprendre, les astrophysiciens ont reconstruit la carte d'émission X de Coma telle qu'elle serait s'il était au repos et l'ont soustraite de la carte obtenue avec XMM-Newton : ils disposent ainsi d'une carte des régions où la matière est en mouvement. Elle révèle plusieurs zones d'émission X intense. Le cœur de Coma est principalement composé de deux galaxies. Une seconde zone d'émission intense correspond à une structure filamentaire (un groupe de galaxies) attirée par Coma. C'est la première fois que ce filament est détecté avec une telle précision dans le domaine X. Les deux extrémités du filament sont orientées vers des amas de galaxies, ce qui confirme que les amas se forment au croisement des filaments. Dans l'équipe de M. Arnaud et Ph. Ferrando, Doris Neumann a étudié en détail le voisinage d'une galaxie qui se trouve à la frontière entre Coma et le filament. Elle est plus proche de l'amas que le gaz chaud du groupe de galaxies auquel elle appartient, car elle tombe plus vite que ce dernier, ralentit par le gaz chaud de l'amas.

Ces observations font partie de la phase de vérification de XMM-Newton, qui vient de s'achever. Elles montrent le bon fonctionnement du satellite, dont les résolutions spatiale et spectrale sont conformes aux prévisions. En outre, le satellite a bien résisté aux deux éruptions solaires qui se sont produites depuis son lancement (grâce à un système de filtres qui le protège des particules émises), laissant présager que sa durée de fonctionnement sera effectivement de dix ans, comme prévu. Cela suffira aux astronomes pour observer l'état dynamique d'autres amas de galaxies, ce qui nous éclairera sur la formation des grandes structures de l'Univers. Les prochains résultats sont d'autant plus attendus que XMM-Newton, sorti de sa phase de test, sera pointé vers des objets moins bien connus des astrophysiciens.



Cartes de la matière en mouvement (l'intensité de l'émission X augmente du bleu au rouge) dans l'amas de galaxies Coma. Son cœur est formé de deux galaxies (image de gauche). Cet amas se trouve à l'extrémité d'un filament (sa direction est celle des pointillés blancs) le long duquel un groupe de galaxies tombe vers lui. La galaxie de l'image de droite (en rouge) tombe vers l'amas Coma, mais moins vite que le gaz chaud qui l'entoure (en jaune et en vert).

La fonction du prion

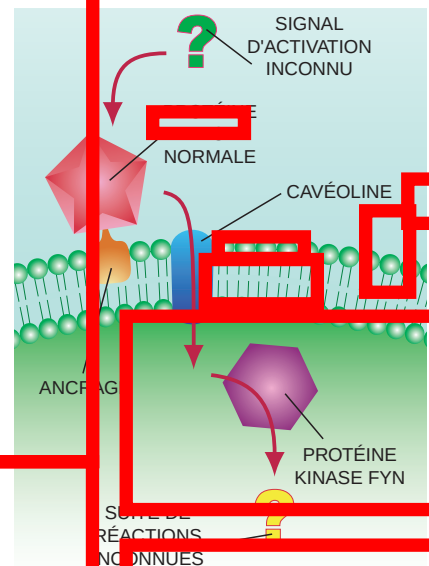
Le rôle de cette protéine se précise

Maladie de la vache folle, tremblante du mouton, maladie de Creutzfeldt-Jakob, ces maladies qui sont toutes des encéphalopathies spongiformes subaiguës transmissibles, c'est-à-dire des maladies qui, au bout de plusieurs années, finissent par transformer le cerveau en «éponge» sont très préoccupantes depuis l'épidémie britannique de la vache folle du début des années 1990. Elles auraient en commun le prion. Il s'agirait d'une protéine pathogène dont seul le repliement la distinguerait de la protéine prion normale. Or, on ignorait le rôle de cette protéine normale, rôle que l'équipe CNRS-Institut Pasteur d'Odile Kellermann vient de préciser.

Lors d'une contamination, la protéine prion anormale pénètre dans le système nerveux central et perturbe le repliement des protéines prions normales nouvellement synthétisées par les neurones. Les protéines mal repliées qui s'accumulent entraînent la dégénérescence et la mort des neurones, donnant ainsi au système nerveux cet aspect d'éponge. Des expériences antérieures avaient mon-

tré que cette protéine n'est pas vitale : des souris qui en sont dépourvues ont des anomalies mineures. Pour élucider la fonction de la protéine prion normale, les biologistes ont donc utilisé des cellules souches embryonnaires qui se différencient *in vivo* en neurones fonctionnels. La protéine prion est présente à la surface de ces cellules, et particulièrement dans les neurones des extensions des neurones). Les biologistes ont observé que lorsque la protéine prion reçoit un signal d'activation (dont on ignore encore la nature), l'activité d'une enzyme intracellulaire, une tyrosine kinase nommée FYN, augmente. Cette enzyme ajoute ou retire des groupes phosphates à la tyrosine, un acide aminé des protéines. La protéine kinase FYN est connue pour son rôle dans les cascades de signalisation, c'est-à-dire dans les successions de réactions enzymatiques qui transmettent les signaux extérieurs, tel celui qui déclenche les hormones ou les neurotransmetteurs, jusqu'au cœur des cellules, où les ordres véhiculés par ces signaux sont accomplis.

Comment le signal passe-t-il de l'extérieur à l'intérieur de la cellule ? L'équipe de O. Kellermann a identifié le partenaire de la protéine prion normale : il s'agit d'une autre protéine, la cavéoline, enchâssée dans la membrane cellulaire. Ainsi, la protéine prion intervient dans une cascade de signalisation dont on a élucidé trois étapes. Toutefois, le signal qui l'a



Ancrée dans la membrane des cellules grâce à une molécule (en orange), la protéine prion normale (en rouge) est activée par un signal encore inconnu. Ce signal (flèches rouges) active une protéine, la cavéoline (en orange), puis une enzyme (violet) nommée FYN (en violet). Seuls trois maillons de la chaîne ont été découverts.

déclenchée en amont, et la suite de réactions en aval, sont encore inconnus.

Ces résultats constituent, d'ores et déjà, un pas vers l'élucidation du processus à l'origine des encéphalopathies spongiformes subaiguës transmissibles.

Bateau romain... coulé

Depuis 1 500 ans, un bateau romain repose sur le fond de la mer Noire : il est en parfait état de conservation.

Le robot sous-marin *Argus* descend dans l'obscurité à 300 mètres sous la surface. Le fond n'est plus qu'à quelques mètres quand le spot lumineux accroché au robot dans la salle de contrôle, tous sont fascinés. Un mât de bois se dresse dans les eaux. Long de 11 mètres, ce mât est celui d'un navire de 15 mètres de longueur posé sur le fond de la mer Noire depuis 1 500 ans. Il vient d'être découvert par des archéologues et des océanographes américains, embarqués sur le navire d'exploration *Northern Horizon*.

Outre son intérêt archéologique, la trouvaille confirmerait le scénario d'une brusque invasion de la mer Noire par les eaux méditerranéennes. Selon les géologues américains William Ryan et Walter Pitmann, cette mer fut un lac d'eau

douce, il y a plus de 7 500 ans. Pour ces chercheurs de l'Université Columbia, la fonte des glaciers a fait monter les eaux de la mer Méditerranée, qui ont envahi ce lac. Certains archéologues suspectent que la région de la mer Noire était habitée par l'homme néolithique. Le déluge évoqué dans la bible correspondrait à la naissance de la mer Noire.

C'est aussi lors de cet événement que la couche d'eau sans oxygène qui garnit le fond de cette mer se serait constituée : les eaux salines provenant de la mer Méditerranée auraient coulé sous les eaux douces. La zone située à plus de 100 mètres de profondeur est pauvre en oxygène, dépourvue de vie et de micro-organismes. Stable depuis la naissance de la mer Noire, cette «zone morte» serait un réservoir de trésors archéologiques bien conservés, puisque les matières organiques, le bois notamment, n'y auraient pas été dégradées.

C'est parce qu'il était à l'abri de la décomposition, dans cette zone morte, que le bateau romain semble intact. Les datations au carbone 14 situent la construction du navire entre 410 et 520, peu de temps après le déplacement de



Le mât du bateau romain découvert au fond de la mer Noire apparaît dans la lumière du robot sous-marin *Argus*.

la capitale de l'empire romain de Rome à Constantinople, en 330. Les archéologues, qui ignorent encore ce que transportait le bâtiment, espèrent y retrouver des restes de vivres ou de tissus.

Trois bateaux marchands romains dans un état de conservation plus naturel ont également été découverts lors de cette expédition. Les vaisseaux reposent à 100 mètres de profondeur, où les eaux pauvres en oxygène se mêlent aux eaux de la surface. Les amphores qu'ils contiennent proviennent de la ville de Sinope, un port de la mer Noire, qui avait une grande importance dans l'économie locale, il y a 1 500 ans.

Poisson fumé à la néandertalienne

Il y a 50 000 ans, on faisait déjà fumer les saumons en Dordogne.

Les hommes ont occupé la vallée de la Dordogne de façon continue de 30 000 à 12 000 ans avant notre ère. Ils y ont laissé de nombreux vestiges, notamment dans des grottes de la falaise de Conte. La grotte 16 (la falaise en contient 23 au total), qui renferme à la fois des restes laissés par les Néandertaliens et par les hommes modernes (*Homo Sapiens Sapiens*) offre l'occasion rare de comparer leur mode de vie dans un même environnement. L'équipe de Jean-Philippe Rigaud, de l'Université Bordeaux I, et celle de Jan Simek de l'Université du Tennessee, ont découvert que ces deux espèces d'hominidés appréciaient le poisson et avaient des comportements voisins.

La grotte 16 mesure environ 25 mètres de largeur, 9 de hauteur et 20 de profondeur. Les sédiments qui recouvraient le sol ont livré des restes d'animaux (dans les foyers et aux alentours) et des outils de pierre, notamment des bifaces. Ils indiquent que des groupes de chasseurs (Néandertaliens ou hommes modernes) auraient occupé la grotte pendant de courtes périodes.

Ces hominidés semblent avoir pratiqué la pêche de façon intensive, si l'on en juge par l'abondance des vertèbres de saumon, de truite, de brochet... Cette découverte confirme le fait que les Néandertaliens exploitaient les ressources halieutiques de leur environnement. L'entrée de la grotte domine la confluence de la Dordogne et d'une petite rivière, le Céou, où se trouve un

gros ruisseau. Les conditions de pêche y sont favorables.

On ne sait que les Néandertaliens utilisaient le feu de façon rudimentaire, car, contrairement aux hommes modernes, leurs foyers ne comportaient pas de cercle de pierres, tels ceux que les archéologues ont retrouvés dans les niveaux du paléolithique supérieur de la grotte 16. Cependant, dans les foyers datant de 65 000 à 54 000 ans (les mieux préservés que l'on connaisse de cette époque), des restes de lichen, de mousse et d'herbes indiquent que les Néandertaliens utilisaient ces combustibles, avec du bois. Les herbes brûlent mal, mais produisent beaucoup de fumée : elles ont probablement servi à fumer le poisson ou la viande, ce qui assurait une meilleure conservation. Ce comportement semble plus évolué que celui que l'on attribue généralement aux Néandertaliens. Peut-être tiraient-ils parti des périodes de migration des animaux (surtout des poissons et des rennes), comme semblent l'avoir fait l'homme moderne : ils en auraient alors pris beaucoup sur une courte période,

ce qui paraît justifié l'emploi de techniques de conservation, tel le fumage, pour les denrées qu'ils ne consommaient pas immédiatement. Les archéologues reconstituent aujourd'hui les stratégies de chasse et de pêche des deux types d'hominidés, afin de confirmer cette hypothèse.

Dans la grotte 16, les vestiges laissés par les Néandertaliens datent de 75 000 à 36 000 ans, tandis que les vestiges les plus anciens laissés par les hommes modernes datent de 30 000 ans environ. Ils ne s'y seraient donc pas rencontrés et un remplacement rapide des uns par les autres, dû à leur affrontement, est exclu. Selon J.-P. Rigaud, les hommes modernes auraient progressivement remplacé les Néandertaliens car leur croissance démographique aurait été plus forte ou leur structure sociale plus solide.

Dunes en liberté

Quand on laisse la mer regagner le terrain conquis sur la terre.

Sur la côte atlantique de la France existent des vastes champs de dunes. Certaines, comme en Bretagne, sont fixées depuis longtemps par une végétation naturelle de pelouses ou de taillis. Ailleurs, surtout en Picardie et en Aquitaine, elles avaient conservé leur mobilité jusqu'à ce que l'homme décide de les stabiliser parce que, dans leur déplacement (parfois plusieurs kilomètres à l'intérieur des terres), elles menaçaient d'ensevelir habitations, champs, chemins et routes. Au XVII^e siècle déjà, des intendants recommandaient de planter des graminées, nommées oyats, pour freiner l'avancée des sables sur les rivages du Nord de la France. Au cours du XIX^e siècle, à la suite du décret impérial du 14 décembre 1810, les dunes vives de la côte atlantique firent l'objet de travaux de fixation de grande ampleur. Ainsi, la vaste forêt de pins maritimes qui s'étend sur plus de 80 000 hectares, à l'arrière de la côte d'Aquitaine, est un environnement artificiel que l'homme a substitué à un écosystème naturel de grandes dunes mobiles. Les étendues de sable blanc étaient le symbole de la misère des sociétés rurales et, dans les Landes, les ressources fournies par l'exploitation du bois et de la résine, ont favorisé le développement économique et social de la région.

Aujourd'hui, en Europe, la tendance est au contraire à la conservation des dunes mobiles là où elles existent encore, comme en Espagne, au Danemark ou en Pologne. Sur des côtes où la pression démographique n'a cessé de s'amplifier, elles constituent des espaces vides qu'il convient de préserver. Il faudrait-il pas aussi, là où il n'y a pas de risques, redonner une liberté de mouvement à des dunes artificiellement entravées ? Un premier pas vient d'être fait, en France, dans ce sens : dans le département du Pas-de-Calais, sur une réserve biologique domaniale de 450 hectares, on ne procède plus, depuis 1996, à aucuns travaux de fixation des dunes.

Déjà des langues de sable apparaissent et des dunes de forme parabolique évoluent. Les photographies aériennes du site montrent une remise en mouvement du sable. Ce nouveau type de gestion d'un espace naturel a été validé par la branche française de l'Union européenne pour la conservation des côtes (EUCC), une organisation qui s'attache à la sauvegarde de la richesse paysagère,



Cette coupe des sédiments de la grotte de la falaise de Conte, en Dordogne, montre des bandes noires, rouges, jaunes et blanches. Dans les foyers des Néandertaliens qui ont occupé la grotte, les matières ferrugineuses initialement présentes dans le sable recouvrant le sol ainsi que les cendres ont été oxydées par les feux ou réduites par le charbon de bois. Les produits de ces réactions sont responsables des différentes couleurs.

écologique et culturelle des littoraux du vieux continent. Cette approche innovante est assumée par l'Office national des forêts, qui, jusqu'à présent, considérait strictement le déplacement des sables sur les rivages maritimes. Cette perception renouvelée de la valeur des espaces littoraux conduit à restituer à la mer des terrains autrefois conquis par l'homme. Ainsi, deux polders, l'un dans l'iber de Crozon, en Bretagne, l'autre dans la baie des Veys, en Normandie ont déjà été « rendus » à la mer par le Conservatoire du littoral.

Roland PASKOFF,
Université Lumière - Lyon

Près de Melimont, sur le littoral du Pas-de-Calais, on laisse les dunes se déplacer à leur gré, à condition que cela ne menace pas les aménagements côtiers. Des langues de sable vif s'avancent aujourd'hui sur des dunes que l'homme avait naguère stabilisées par des plantations.

Roland Paskoff

Un rat artificiel

En prenant le rat comme modèle, biologistes, et «roboticiens» créent des animaux artificiels.

Les robots des premières générations étaient limités à des tâches répétitives, par exemple, peindre une portière de voiture ou visser un boulon. Ces machines n'avaient rien de commun avec les êtres vivants. Aujourd'hui, les biologistes, les informaticiens et les roboticiens s'associent pour créer des «animats», c'est-à-dire des animaux artificiels simulés sur ordinateur ou des robots réels, qui ont des comportements adaptatifs. Il s'agit en somme de construire des machines dont le comportement est proche de celui d'animaux vivants. Différents travaux semblent prouver que ce n'est pas de la science fiction.

Une nouvelle espèce est née : le *Ratus computatrix* issu des connaissances acquises sur la neurophysiologie et le comportement du rat. Cela ne signifie pas que la fabrication d'un robot, dont le système de contrôle et le comportement seraient identiques à ceux de l'animal, sera possible dans un avenir proche. En revanche, la combinaison de mécanismes et de capacités souvent étudiés isolément contribue à une meilleure compréhension des synergies et des compromis que l'évolution a inventés et mis en œuvre pour faire du rat l'un des animaux les mieux adaptés.

Angelo Arleo et Wulffram Gerstner, de l'Institut fédéral de technologie de Zurich, ont élaboré un système de navigation pour un robot autonome fondé sur l'utilisation de «cellules de lieu» et de «cellules de direction de la tête». De telles cellules s'activent respectivement selon le lieu et la direction de la tête. Le robot miniature *Khepera*, doté d'une caméra, de capteurs infrarouges et de deux roues, est commandé par deux

modules : l'un permet la représentation spatiale et permet au robot de rejoindre un lieu où il a déjà été «récompensé» (l'expérience lui a octroyé une «valeur» élevée parce que c'était le «bon chemin»), tout en évitant les obstacles et les endroits où il a été «punni». Mouvements propres du robot et informations visuelles externes se complètent.

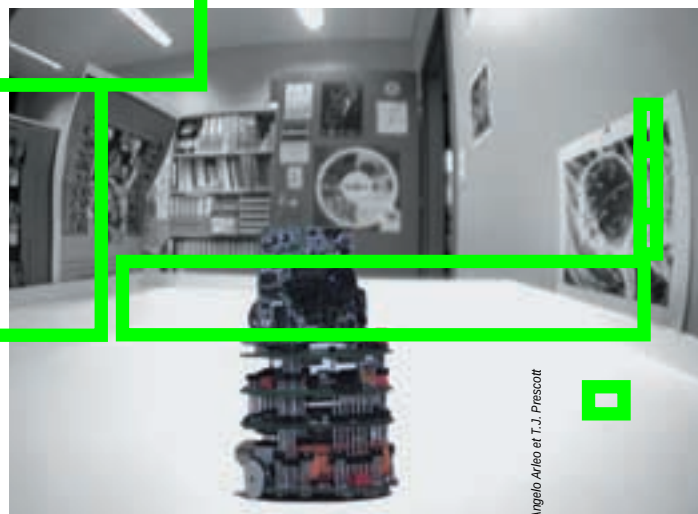
Ces résultats confortent celui qui a été obtenu par l'équipe d'Alain Berthoz, du Laboratoire de physiologie de la perception et de l'action, du Collège de France. L'équipe a montré que chez le rat les cellules de direction de la tête sont principalement sensibles aux informations visuelles, mais qu'elles dépendent également d'informations internes lorsque les informations visuelles ne sont pas fiables.

Jean-Academy Meyer et David Filliat, de l'Animatix au Laboratoire d'Informatique de l'Université Paris 6, ont étudié les stratégies de déplacement utilisées par des rats. Leur robot explore son environnement et en élabore une carte cognitive, constituée des informations topologiques et des informations métriques. À partir d'une telle carte, lorsqu'un robot est transporté d'un endroit à un autre, il se trompe temporairement sur sa localisation, mais se repositionne vite. Pour ce faire, il utilise un système de vision active et oriente sa caméra, soit vers les endroits où il a enregistré la présence de certains éléments caractéristiques qui lui servent de repères, soit vers les endroits qu'il n'a pas encore explorés visuellement.

Enfin, l'organisation des neurones que le rat utilise lorsqu'il doit choisir une action parmi plusieurs a inspiré l'équipe de Fernando Montes Gonzales, de l'Université de Sheffield, dans la conception d'un système de contrôle de robot. Muni d'un tel système, un robot *Khepera*

équipé d'un bras préhensile choisit correctement la quelle action effectuer : à l'image d'un rat placé dans un environnement inconnu et capable d'alterner de façon pertinente des déplacements exploratoires, des déplacements destinés à consommer de la nourriture et des retours vers une zone sombre sécurisante. Le robot, doté de motivations codées, tels la «faim» et le «stress», décide d'aller chercher de la nourriture (saisir un cylindre), de consommer cette nourriture (déposer le cylindre), et d'aller explorer l'environnement (déplacement aléatoire).

Ainsi, *Ratus computatrix* combine plusieurs sources d'informations sensorielles pour élaborer une représentation interne de son environnement. Il exploite cette information pour rejoindre des lieux où il peut satisfaire de façon cohérente des motivations simples, et choisir l'action la plus pertinente. Dans une prochaine étape, les relations sociales des rats seront appliquées à une collectivité de robots. Nous sommes, d'ores et déjà, dans l'ère de l'informatique murine.



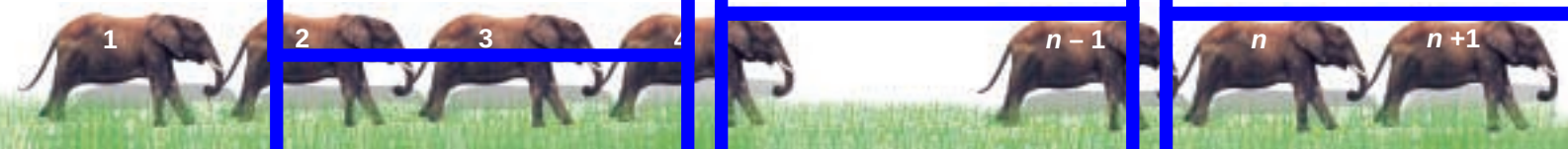
Angelo Arleo et T.J. Prescott

Grâce à une caméra, un robot *Khepera* explore son environnement. Il en établit progressivement une représentation spatiale interne.

Les mathématiques, science de l'infini

JAVIER DE LORENZO MARTÍNEZ

L'infini, sous ses aspects potentiel et actuel, semble assurer une fonction régulatrice dans la formalisation mathématique.



Faites l'expérience. Déclenchez la trotteuse de votre montre au début d'un exposé de mathématiques : il ne s'écoulera pas cinq minutes avant que l'orateur prononce le mot infini. Cette omniprésence de l'infini en mathématiques est étonnante car l'homme est un être fini, limité, embarqué sur un vaisseau, la Terre, limité et fini. Pourtant cet être fini examine l'infini et (s')en joue, au point que l'infini lui est indispensable pour appréhender le fini.

L'étude de l'infini est l'apanage du mathématicien, dont le travail consiste, pour l'essentiel, à répéter, comparer, ordonner et classer. L'itération et la comparaison conduisent à deux infinis différents, l'infini potentiel et l'infini actuel.

L'infini potentiel

Le marcheur, dans la succession de ses pas – un pas, puis un autre et encore un autre – sait que sa (dé)marche peut se répéter indéfiniment. En principe, il peut toujours faire «un pas de plus» dès lors qu'il vient d'en faire un. Cette répétition sans limitation conduit à l'intuition première d'un indéfini sans fin, l'infini potentiel ; il y a toujours un (dé)pas (sement) possible. L'infini potentiel, la faculté d'aller «un peu plus loin», est ainsi lié à la notion de successeur d'un entier naturel. À un nombre succède

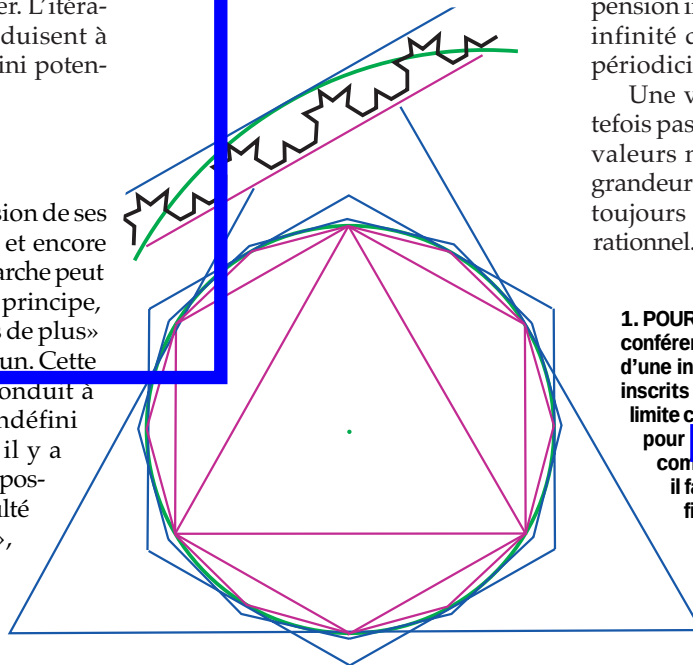
toujours un autre nombre et il n'en existe pas de dernier, car ce dernier nombre... a un successeur.

Les entiers permettent d'énumérer, c'est-à-dire de savoir combien de pas ont été faits. En comptant ou en mesurant, on associe une valeur numérique à une grandeur spatiale ou à une quantité. On compte les brebis d'un troupeau, les grains de sable du désert ou la longueur d'un bâton. Souvent ce décompte utilise implicitement un principe dit archimédien qui, comme nous le verrons, autorise les comparaisons : quelles que soient les valeurs

numériques x, y telles que x est inférieur à y , il existe un entier naturel n tel que nx est plus grand que y .

Pour comparer deux grandeurs de même nature, l'une étant prise comme unité de référence, il est pratique de leur associer un nombre. Le rapport de ces deux valeurs numériques associées pourra être entier, mais sera plus souvent rationnel, c'est-à-dire une fraction comme un huitième, un cinquième, etc. Dans cette comparaison, on opère une division où tout nombre est exprimable sous forme décimale. Ainsi, on écrit $9/10 = 0,9$; $1/3 = 0,333...$, où les points de suspension indiquent que l'écriture a une infinité de chiffres ou apparaît une périodicité de certains blocs.

Une valeur numérique n'est toutefois pas toujours le quotient de deux valeurs numériques associées à des grandeurs, c'est-à-dire qu'elle n'est pas toujours représentée par un nombre rationnel. Un nombre irrationnel s'écrit



1. POUR CALCULER π , la longueur de la circonférence, Archimède considère la limite d'une infinité de polygones circonscrits et inscrits dans le cercle et calcule la valeur limite comprise entre ces polygones. Ainsi, pour calculer une longueur infinie, mais comportant une infinité de décimales, il faut utiliser un processus infini. L'infini est cependant plein de pièges : même dans ce cas simple, on peut construire des fonctions de type fractal, de longueur infinie, qui restent comprises entre les polygones.

lui aussi, en base dix, avec une infinité de chiffres après la virgule, mais, à la différence du rationnel, ne présente pas de périodicité. Ainsi, dans le nombre $p = 3,14159\dots$ il n'y a pas de groupe de chiffres qui se répètent, et p n'est donc pas rationnel. Tout nombre réel, par exemple les distances à l'origine des points sur une courbe, s'exprime au moyen d'un nombre infini de chiffres. Cette infinité est caractéristique du nombre réel, mais pose certains problèmes quant à la définition constructive de celui-ci (voir *L'infini, pierre de touche du constructivisme*, pages 74 à 81).

L'infini émerge aussi de l'idée d'approximation du limité, de la valeur numérique associée à une grandeur commensurable ou incommensurable. Par exemple, $2/3$ ou $\sqrt{2}$ sont des nombres, mais leurs développements décimaux ne peuvent être donnés dans leur totalité ; néanmoins, les 10, 100 ou 1 000 premiers chiffres en constituent une approximation. Les écritures $2/3$ et $\sqrt{2}$ expriment des rapports de proportionnalité : $2/3$ est une quantité commensurable et le nombre est rationnel ; dans un carré, la diagonale est incommensurable avec le côté (on ne peut exprimer cette longueur comme un rapport, une fraction, de deux longueurs), et le nombre $\sqrt{2}$ est irrationnel. Dans ces deux cas, et de manière essentielle dans le second, l'infini apparaît lié à la convergence de fractions distinctes vers une limite.

En géométrie, l'infini apparaît aussi comme un processus d'approximation de limite. L'aire du cercle est la limite vers laquelle tendent les aires des polygones inscrits ou circonscrits lorsqu'on fait croître le nombre de côtés des polygones considérés. On dit alors que le nombre de côtés des polygones inscrits ou circonscrits tend vers l'infini ; la circonférence du cercle est alors un polygone ayant un nombre infiniment grand de côtés infiniment petits. Ici, l'infini apparaît plus comme un procédé de calcul, mais plutôt

Raisonnement par l'absurde

Le raisonnement par l'absurde est initialement lié à la preuve de l'irrationalité de nombre comme $\sqrt{2}, \sqrt{3}, \sqrt{5}, \dots$. Platon rapporte dans le *Théétète* que les mathématiciens ont démontré l'irrationalité de ces nombres jusqu'à $\sqrt{17}$. Cependant, il n'indique pas comment ils y parviennent ni pourquoi il s'arrête à $\sqrt{17}$.

Aristote, dans la *Métaphysique*, propose une démonstration de l'irrationalité de $\sqrt{2}$, désormais classique : supposons que $\sqrt{2}$ soit rationnel, c'est-à-dire exprimable comme rapport de deux entiers naturels, disons a et b , de telle manière que a et b soient premiers entre eux. On a $\sqrt{2} = a/b$. En multipliant par b de chaque côté, on obtient $\sqrt{2}b = a$. En élevant au carré, il vient $a^2 = 2b^2$, ce qui implique que a^2 est pair, et donc que a l'est aussi. On pose $a = 2p$. De là, il vient $2b^2 = (2p)^2 = 4p^2$, puis $b^2 = 2p^2$, ce qui signifie que b est pair lui aussi. On a ainsi montré que a et b sont pairs ; ils ne sont donc pas premiers entre eux, contrairement à ce qui a été supposé. Par conséquent l'hypothèse de la rationalité de $\sqrt{2}$ conduit à une absurdité et doit être rejetée.

comme le résultat du processus de calcul. Il est le nombre qui termine la succession des nombres entiers, la limite des valeurs numériques successives d'une même variable qui croît (ou décroît) à chaque étape. Similairement, pour atteindre la droite réelle, on a besoin de deux infinis, $+\infty$ et $-\infty$. Le plan euclidien, lui, se complète par la droite à l'infini, utile en perspective (voir *De la perspective, l'infini géométrique*, pages 66 à 72), où chacun des points est l'intersection des droites parallèles à une direction donnée.

Les points du plan à l'infini permettent l'addition d'admettre les principes de continuité et de dualité, de passer d'un espace métrique à un espace projectif. Par exemple, dans un plan euclidien, deux points déterminent une droite, mais deux droites ne déterminent pas toujours un point (quand elles sont parallèles) ; il n'y a pas de symétrie dans l'énoncé. Or, en admettant que deux droites parallèles se coupent à l'infini en un point unique, l'énoncé réciproque devient vrai.

Une symphonie sur le thème de l'infini

L'infini géométrique se distingue de l'illimité : on conçoit aisément un espace illimité (c'est-à-dire sans frontières), mais fini, comme la sphère. Il y a identité entre infinitude et illimité dans le seul contexte euclidien métrique. Pour un calcul numérique, cela signifie que la possibilité d'ajou-

ter un, qui caractérise l'infini potentiel, se présente comme une propriété métrique, quantitative ; l'illimité est ici encore identifié à l'infini potentiel.

Retournons à l'infini vu comme approximation d'une valeur donnée. En pratique, si l'on mesure la diagonale d'un carré, la précision des appareils limite la mesure à quelques décimales. Le résultat d'une mesure est toujours un nombre rationnel ; l'irrationnel n'a pas de place dans le champ expérimental.

Le mathématicien n'est pas handicapé par cette exclusion. Il comprend que $\sqrt{2}$ a un développement décimal infini non périodique, il affirme que $\sqrt{2}$ est irrationnel ; quelle que soit la mesure effectuée, elle restera toujours inexacte. Dans la possibilité de dépasser la cohérence d'un mode opératoire et de son résultat. Il invente des méthodes de démonstration qui assoient la cohérence des définitions des objets idéaux qu'il a introduits, l'infini identifié avec l'illimité, l'irrationalité de nombres comme $\sqrt{2}$.

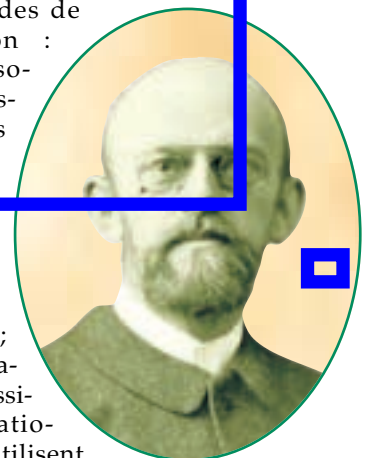
Les mathématiques s'enrichissent continuellement de nouvelles méthodes de

démonstration : méthodes associées au dépassement sans limitation méthodes liées à

l'itération et à la convergence, comme la démonstration par récurrence ; méthodes rattachées à l'impossibilité de l'irrationalité et qui utilisent



2. POINCARÉ (à gauche) pensait que l'infini potentiel suffisait aux mathématiciens et se méfiait des paradoxes induits par la hiérarchie des ensembles infinis de Cantor. **HILBERT (à droite)** était plus optimiste et ne voulait pas se priver du paradis que Cantor avait ouvert aux mathématiciens.



une démonstration par l'absurde, méthodes liées aux ensembles hyper-infinis (voir *L'infini est-il nécessaire?*, pages 102 à 106, et *L'infini est un révélateur*, pages 138 à 142).

Revenons au principe de répétition sans fin utilisé dans la démonstration de l'existence d'une infinité de nombres premiers. La proposition 20 du livre IX des *Éléments* d'Euclide établit que «les nombres premiers sont plus nombreux que toute multitude de nombres premiers proposée». Dans la preuve d'Euclide, on construit un nombre premier qui n'est pas dans la liste des premiers préalablement considérée. On l'y ajoute. On a alors un nouvel ensemble de nombres premiers auquel on peut appliquer à nouveau le procédé de construction ; aujourd'hui, nous disons qu'«il existe une infinité de nombres premiers». Cette formulation, qui identifie l'illimité et l'infini actuel, trahit peut-être la pensée d'Euclide.

Hilbert voyait dans l'analyse infinitésimale le lieu d'expression privilégiée de l'infini potentiel. Dans un certain sens, l'analyse infinitésimale n'est rien moins qu'une symphonie sur le thème de l'infini. L'analyse infinitésimale naît dans un contexte géométrique en relation avec le calcul de longueurs, d'aires et de volumes déter-

minés par une courbe, donnée comme graphe d'une fonction. D'un point de vue cinématique, cette courbe représente la trajectoire d'un point en mouvement en fonction du temps. L'étude du comportement d'une fonction se fait au voisinage d'un point, sur un intervalle élémentaire.

L'analyse infinitésimale est un cadre idéal pour comprendre les phénomènes physiques, mais elle a longtemps reposé sur des bases énigmatiques (voir *La science du mouvement au XVII^e siècle*, pages 38 à 63). Les physiciens établissent d'abord une équation différentielle qui régit le phénomène, et ils intègrent cette équation sur l'intervalle du domaine d'étude. Dans cette démarche apparaissent les concepts clés de l'analyse infinitésimale (dérivée, différentielle, intégrale, continuité, etc.), qui utilisent les notions de limite et d'approximation, c'est-à-dire d'infini potentiel.

Continuité et dérivabilité

La possibilité de tracer une courbe sur une feuille de papier sans lever le crayon est déjà formalisée par Augustin Cauchy qui définit qu'une fonction $f(x)$ est continue par rapport à x si un

accroissement infiniment petit de la variable produit toujours un accroissement infiniment petit de la fonction elle-même. Autrement dit, pour un accroissement h petit par rapport à x , la différence $f(x+h) - f(x)$ décroît indéfiniment avec h .

En termes de limite, une fonction f , définie pour tous les éléments x d'un segment fermé $[a, b]$, est dite continue en un élément x_0 du segment $[a, b]$ si la limite de $f(x)$, quand x tend vers x_0 , est égale à $f(x_0)$.

Cauchy précise qu'«une quantité variable devient infiniment petite quand sa valeur numérique décroît indéfiniment en convergeant vers zéro». Dans ce langage dynamique, les quantités sont des grandeurs qui augmentent ou diminuent, leurs valeurs numériques associées convergeant respectivement vers l'infini ou vers 0.

En termes statiques, les points de l'intervalle sont donnés «en acte». La formulation de Cauchy n'a alors pas véritablement de sens et peut être éliminée au profit de concepts de nature plus méthodique comme ceux de majoration, de minoration ou d'approximation. Dans ce cas, l'énoncé de la continuité d'une fonction en un point x_0 prend la forme suivante : la fonction f est continue en x_0 si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que si $|x - x_0| < \delta$ alors $|f(x) - f(x_0)| < \varepsilon$.

La nouvelle formulation traduit un glissement conceptuel. Les intervalles cessent d'être considérés comme des grandeurs étendues, qui diminuent ou augmentent, pour être vus comme des ensembles de points donnés en compréhension. Les intervalles pourront devenir plus grands ou plus petits, sans qu'ils puissent augmenter ou diminuer dynamiquement et, bien évidemment, sans que leurs éléments puissent changer. L'infini cesse d'être un processus. Nous sommes passés d'un infini potentiel à un infini actuel, comme nous étions passés du plan euclidien au plan projectif.

La notion de dérivabilité prolonge la notion de continuité. La dérivée d'une fonction en un point x_0 est donnée par :

$$f'(x_0) = \lim_{h \rightarrow 0} (f(x_0 + h) - f(x_0)) / h,$$

chaque fois que cette limite a un sens. Une fonction est dérivable en un point si elle admet une dérivée en ce point. Elle y sera alors aussi continue (en revanche, la réciproque n'est pas toujours vraie, comme nous le verrons, une courbe continue n'est pas toujours

Le raisonnement par récurrence ou induction complète

Essayons de comprendre le principe du raisonnement par récurrence à partir de l'inégalité dite de Bernoulli, qui s'énonce comme suit : $(1+x)^n > 1+nx$ pour tout $n \geq 2$.

Une récurrence se fait en trois étapes :

1. On commence par vérifier l'inégalité pour le premier entier naturel qui satisfait les conditions de l'énoncé, c'est-à-dire ici pour $n = 2$:

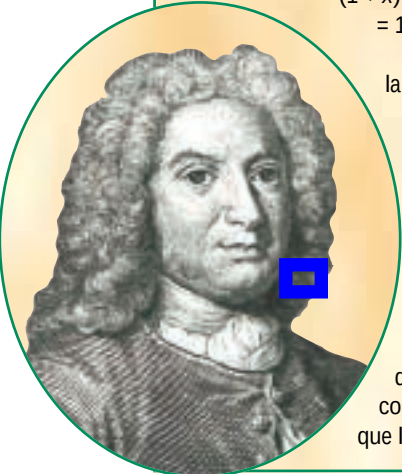
$$(1+x)^2 = 1+2x+x^2 > 1+2x \text{ car } x^2 \text{ est toujours positif}$$

2. On formule l'hypothèse de récurrence – dans notre exemple, on suppose que l'inégalité est vraie pour un certain entier k – et l'on montre qu'elle reste encore vraie pour l'entier $k+1$.

$$(1+x)^{k+1} = (1+x)^k (1+x) > (1+kx)(1+x) = 1+x+kx+kx^2 = 1+(1+k)x+kx^2 > 1+(1+k)x$$

3. (Clause de fermeture.) Comme k est un entier naturel quelconque, la proposition est démontrée pour tous les entiers naturels.

Il convient de distinguer la récurrence de l'induction ordinaire. Cette dernière consiste à vérifier la propriété entier après entier et n'est valable que pour les seuls entiers effectivement considérés. Il n'y a pas de clause de fermeture. Il peut donc arriver que la propriété ne soit pas satisfaite par un certain entier. En particulier, même si, dans les cas examinés, en nombre nécessairement fini, aucun contre-exemple n'a été trouvé, cela ne signifie pas qu'il n'en existe pas. Hume avait repéré cette faiblesse essentielle de l'induction ordinaire. La fermeture dans l'induction complète exclut la possibilité de l'existence du contre-exemple ; l'affirmation vaut pour tous les entiers. Ce qui exige que la propriété à vérifier soit définie pour tous les entiers.



dérivable. La courbe associée à la fonction f possède en ce point une tangente, la valeur de la dérivée étant égale à la pente de cette tangente. Il est alors possible d'étudier le comportement de la fonction – de savoir si elle est croissante ou décroissante, si elle admet un maximum ou un point d'inflexion – en examinant les dérivées successives de la fonction au point considéré.

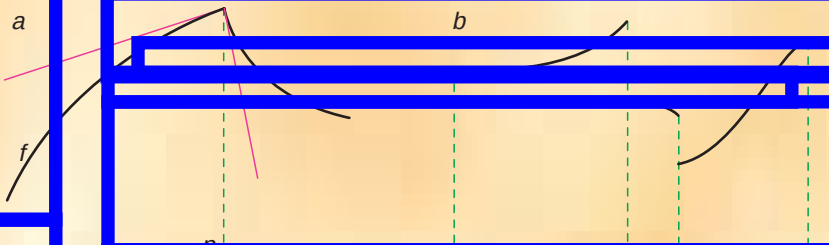
Lorsque la variable d'une fonction parcourt l'ensemble des entiers naturels, l'étude de son comportement se ramène à l'étude d'une suite $u_1, u_2, \dots, u_n, \dots$. De là, on passe au concept de série infinie formée par les sommes partielles successives des n premiers termes de la suite.

Les fonctions continues classiques sont en général très régulières ; indéfiniment dérivables, elles possèdent, en chaque point, une tangente. Mais il existe des fonctions continues qui ne sont dérivables en aucun de leurs points (en conséquence de quoi, elles ne peuvent avoir une représentation géométrique au moyen d'un graphe) ou qui remplissent le plan tout entier, comme la courbe construite par Peano en 1890 (voir l'encadré de la page 28). Pour comprendre ces courbes tératologiques, il faut avoir recours au concept d'infini potentiel, pris ici comme iteration d'un processus d'approximation. On peut approcher uniformément une fonction continue par des fonctions indéfiniment dérivables sur un intervalle ; on remplace la fonction en chacun de ses points par la moyenne des valeurs qu'elle prend sur un petit intervalle autour de ce point.

Les méthodes d'approximation sont très souvent utilisées en analyse infinitésimale. Par exemple, approximer une fonction consiste à remplacer la fonction considérée par une fonction polynomiale ; en d'autres termes, la remplacer par sa série de Taylor associée. Cette approximation est justifiée par le théorème de Karl Weierstrass, qui énonce que « toute fonction complexe continue sur un intervalle fermé $[a, b]$ peut être approchée uniformément sur $[a, b]$ par des polynômes ».

Une fonction se développe aussi en série entière sous la forme $f(z) = \sum c_n (z - z_0)^n$, (n variant de 0 à l'infini), le second membre convergeant sur un disque du plan complexe. Il y a donc une correspondance entre comportement de série et comportement de fonction. L'utilisation

Dérivabilité et continuité d'une fonction



La fonction f de la figure a admet, au point p , une dérivée à gauche et une dérivée à droite, de pentes différentes : la courbe a deux demi-tangentes au point considéré. La fonction n'est donc pas dérivable en ce point, bien qu'elle y soit continue.

Les notions de dérivabilité et de continuité n'ont pas vraiment été distinguées avant le milieu du x^e siècle, la nécessité ne s'en faisant pas sentir. Pendant très longtemps, les fonctions étudiées étaient suffisamment régulières ; les courbes représentatives admettaient une tangente en tout point ou avaient un nombre fini de discontinuités, faciles à comprendre, comme dans la figure b qui représente une fonction en escalier.

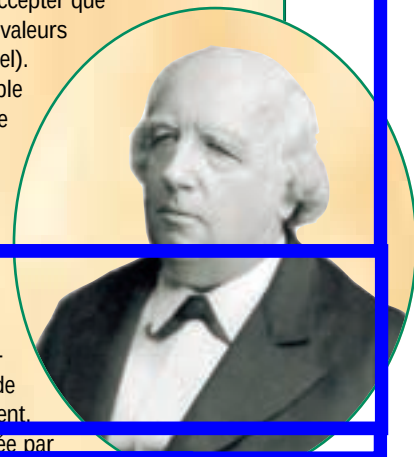
Avec l'apparition des fonctions pathologiques, comme celles qui recouvrent entièrement un carré – la courbe de Peano – ou les courbes ayant une infinité de points de discontinuité – comme celle donnée par Weierstrass –, la situation change et exige de préciser le sens des deux notions. Il faut pour cela commencer par accepter que l'ensemble des définitions d'une fonction et l'ensemble des valeurs qu'elle prend puissent contenir, comme élément, l'infini (actuel).

Le flocon de neige ou courbe de Koch est un autre exemple de courbe tératologique. C'est une courbe de longueur infinie qui délimite une surface finie. Elle s'obtient en partant d'un triangle équilatéral – les trois sommets sont donc des points anguleux en lesquels il n'y a pas de dérivées –, puis en construisant sur chaque côté du triangle un autre triangle équilatéral dont la longueur d'un côté est le tiers de la longueur d'un côté du triangle initial. Puis on réitère la construction une infinité de fois. La courbe obtenue est bien continue. Le flocon de neige est une courbe continue, mais n'est dérivable en aucun de ses points. À la différence des courbes de Weierstrass et de Peano, qui peuvent s'exprimer analytiquement,

le flocon de neige n'est pas le graphe d'une fonction. Publiée par Koch en 1916, le flocon de neige est un des premiers exemples de courbes fractales.

L'ensemble de Mandelbrot est peut-être encore plus pathologique. Il est défini comme la frontière des points du plan complexe obtenus en itérant indéfiniment la fonction de la variable complexe z définie par $z^2 + c$, où c désigne aussi un paramètre complexe. La définition se prête bien au dessin assisté par ordinateur ; la méthode mise au point par John Hubbard et utilisée par Peitgen et Richter consiste à colorier des points en fonction de leur position par rapport à une frontière dont ils s'approchent de plus en plus.

L'existence des courbes fractales généralise la notion de dimension en géométrie ; il existe des objets géométriques dont la dimension n'est pas entière.

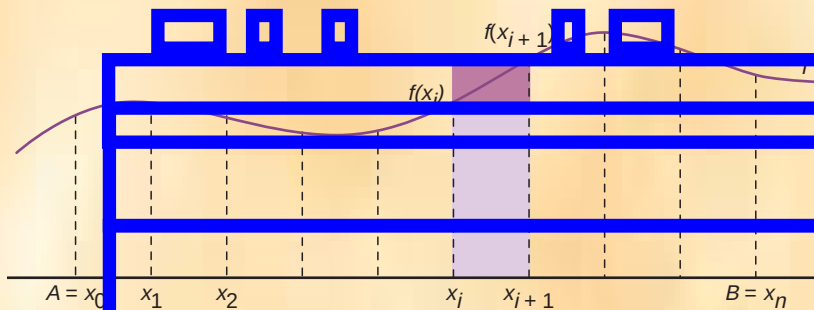


en analyse infinitésimale des méthodes d'approximation, des infinitésimaux ou des découpes peuvent résoudre le problème de l'homogénéité du découpage de tels infinitésimaux par un procédé itératif infini. L'arithmétisation de l'analyse par K. Weierstrass et son école n'y a pas apporté de solu-

tion satisfaisante. L'infini potentiel ne peut y être totalement éliminé : il se situe au cœur de l'analyse infinitésimale, comme nous l'avons vu dans la formalisation de la continuité d'une fonction en un point. Potentiel et chassé par une porte, l'infini réapparaît actuel à une autre.

Né dans un contexte géométrique, celui de la perspective et de l'inclusion du ou de points à l'infini, l'infini actuel autorise la quantification et la résolution de problèmes du monde réel. Il est l'essence des éléments idéaux (nomme infini, point à l'infini et séries avec une infinité d'éléments). La démonstration par récurrence se justifie par le passage de l'infini potentiel à l'infini actuel ; Pascal se réfère à la récurrence quand il affirme que nous connaissons qu'il y a un infini et que nous ignorons sa nature.

L'homogénéité dimensionnelle



Intuitivement, la droite est un objet de dimension 1 et le plan est un objet de dimension 2. Ainsi, la réunion quelconque de droites ne peut donner un plan : c'est le principe d'homogénéité dimensionnelle. Réciproquement, suivant le même principe, si l'on divise une droite en segments, puis ces segments en segments plus petits, et ainsi de suite, on obtiendra toujours des segments, certes infinitésimaux.

Cette idée d'homogénéité se trouve exprimée, entre autres, chez John Wallis en 1671 : «Une quantité finie (ici le segment AB) peut être supposée divisible (par dichotomie) en un nombre de parties infiniment grand (c'est-à-dire plus grand que n'importe quel nombre fini donné) : il n'y a pas de raison de croire que cette division s'arrête (puisque, même à la dernière étape effectuée, un segment, aussi petit qu'il soit, aura encore deux moitiés).»

Comment expliquer qu'une grandeur étendue quelconque, lorsqu'on la partage, reste de même dimension ? Cette question est centrale dans les œuvres de Leibniz et de Cauchy. Kant l'aborde explicitement dans son approche du continu. Pour lui, une droite n'est pas un ensemble de points défini en compréhension, mais plutôt une grandeur étendue ; quant aux points eux-mêmes, ils marquent les extrémités des segments. L'analyse non standard de Robinson propose une réponse qui ne revendique pas de filiation avec l'analyse leibnizienne. En outre, Leibniz refusait l'infini actuel, alors que précisément l'analyse non standard ne prend son sens qu'en l'admettant.

Pour calculer l'aire délimitée par la courbe représentative de f , les ordonnées des points A et B et l'axe des abscisses, Pascal remplace l'aire S sous la courbe par la somme de tous les rectangles de base «infinitésimale». Il obtient ainsi l'encadrement de S : $(x_{i+1} - x_i)f(x_i) < S < (x_{i+1} - x_i)f(x_{i+1})$, les intervalles $[x_i, x_{i+1}]$ devenant, à chaque itération, plus petits. Selon le principe d'homogénéité de Wallis, la répétition infinie du procédé ne fait pas apparaître de segments non divisibles. Les infinitésimaux fabriqués sont de même nature dimensionnelle que la grandeur initiale, le segment AB. L'infinitésimal n'est pas un indivisible, comme le point, de dimension 0, qui marque ou renferme les extrémités de l'intervalle.

Même si la différence $x_{i+1} - x_i$ tend vers 0, $[x_i, x_{i+1}]$ reste un segment de dimension 1. Ainsi, la somme des aires S_i des rectangles ne sera pas égale, même à la limite, à l'aire S cherchée, à moins de commettre systématiquement une erreur. En identifiant $x_{i+1} - x_i$ à l'indivisible – autrement dit, à un point – on pourrait obtenir une égalité, mais cela impliquerait de rompre avec le principe d'homogénéité dimensionnelle des figures de l'espace, les rectangles infinitésimaux devant, quant à eux, être identifiés, à la limite, à des droites.

Infini et bijections

L'ensemble des entiers naturels est un avatar de l'infini actuel comme ensemble de cardinal transfini, appelé $\aleph_0$. L'approche ensembliste de l'infini repose sur la comparaison bijective, une méthode qui supplée l'itération.

On peut comparer des ensembles sans avoir compté. Un berger analphabète se contente de comparer une à une les brebis de deux troupeaux entrant dans une bergerie pour savoir quel troupeau est le plus gros. Il saura si les troupeaux ont le même nombre de bêtes ou non. Pour pouvoir opérer ce genre de bijection, les troupeaux doivent être effectivement présents et dans leur totalité. Le point de départ est donc la donnée *a priori* de ces ensembles, : il importe peu qu'on puisse les énumérer indépendamment.

Existe-t-il un unique infini actuel ? Y en a-t-il plusieurs et combien ? Un ensemble est infini s'il peut être mis en correspondance bijective avec une de ses parties. Ainsi, l'ensemble des entiers naturels est en bijection avec l'ensemble des nombres pairs ou, plus étourdissant encore, avec l'ensemble des nombres premiers. Ces deux ensembles ont donc le même nombre d'éléments, le même cardinal.

D'un point de vue purement comptable ou, si l'on reprend l'argument euclidien déjà signalé, on a envie de dire que les nombres premiers sont en plus petit nombre que les entiers. Les nombres premiers constituent une sous-partie des entiers, et les entiers s'écrivent à partir des premiers (tout nombre se décompose, de façon unique, en produit de nombres premiers). En conséquence, dans la perspective ensembliste, «le tout n'est pas plus grand que la partie». Mieux, cette caractéristique fonde la différence entre fini et infini : dans le champ du fini elle est valable, dans le champ de l'infini elle cesse de l'être.



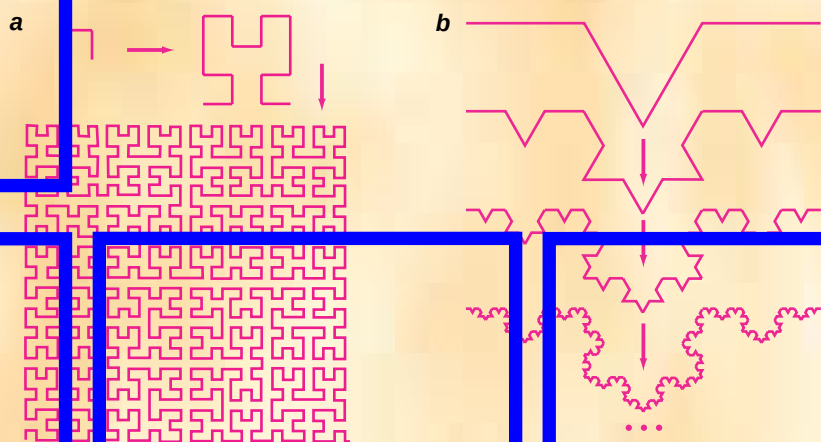
Les courbes continues sans dérivées

Afin de distinguer continuité et dérivabilité, le mathématicien allemand Karl Weierstrass construit, en 1872, une fonction définie par la série convergente :

$$f(x) = \sum a^n \cos(\pi b^n x),$$

fonction qui est continue en tout point et n'est dérivable en aucun point.

Cette fonction est difficilement visualisable. Peano construit une courbe (a) définie par une itération infinie, continue, non dérivable en aucun de ses points et qui, «à l'infini», remplit tout le plan de dimension 2. Von Koch élabore une courbe (b) de dimension non entière (4/3), la courbe en flocon de neige.



Y a-t-il un seul type d'infini actuel, celui associé aux entiers naturels pris dans leur ensemble? En d'autres termes, y a-t-il un unique cardinal transfini? Après avoir dépassé la notion d'illimité et accepté l'existence d'un infini actuel, l'unicité de l'infini paraît plutôt vraisemblable. Cantor anéantit cette intuition en établissant l'existence d'une échelle d'ensembles transfinis. Il indique même le chemin, la méthode dite de la diagonale, qui permet de passer d'un ensemble donné quelconque à un ensemble de cardinalité plus grande (voir *L'infini est-il paradoxal en mathématiques?*, pages 30 à 38).

Cantor se résout à comparer des ensembles donnés dans leur totalité actuelle. En définissant un ensemble, on se donne en même temps ses parties, ou plutôt l'ensemble de ses sous-ensembles, dont le cardinal est plus grand que celui de l'ensemble de départ. On dispose ainsi, outre la méthode de la diagonale, d'un procédé pour construire des ensembles transfinis de cardinalités croissantes.

L'infini actuel joue un rôle essentiel dans la création mathématique. Comme élément régulateur, il permet d'établir des propositions et des théorèmes qui sans lui, n'auraient pas de sens. Et cela, même si, en pratique, le mathématicien se contente de deux infinis, le dénombrable et le continu.

Les démonstrations mathématiques qui recourent à l'infini actuel perdent le caractère constructif que confère l'usage du seul infini potentiel. Elles prennent simplement un caractère existentiel sans qu'on puisse construire explicitement l'objet dont elles démontrent l'existence. C'est le cas lorsqu'une

preuve utilise l'«axiome du choix», qui affirme l'existence d'une fonction de choix mais ne la caractérise pas, ou l'«existence d'une relation de bon ordre» qui en est une formulation équivalente. Un raisonnement par l'absurde est souvent le point de départ de ces preuves existentielles. Il ne faut toutefois pas perdre de vue que, l'existence d'un objet étant acquise, il existe peut-être une voie constructive, inaccessible temporairement au chercheur, et qui mène à cet objet. Il arrive au contraire, que l'on ait démontré qu'il est impossible de construire de tels ensembles, comme par exemple les ensembles hyperinfinis, qui n'en sont pas moins utiles, voire indispensables.

Prétention formelle ou affectif?

L'infini actuel ou potentiel, n'existe pas dans la nature. L'homme, immergé dans cette nature, ne peut connaître que cette seule nature et ce qui la compose. À la rigueur lui concéderait-on l'infini potentiel, ou faux infini, perçu comme répétition sans limitation d'un processus de génération, par exemple des points sur une droite. Toutefois, il ne pourra prétendre donner ces points dans leur totalité actuelle.

Comment se représenter l'infini actuel? L'essentiel est moins de l'objectiver que de saisir ce qui lui est constitutif : les notions premières d'ensemble et d'application bijective entre ensembles.

Il importe de comprendre l'usage de ces notions fondamentales. Refuser le rôle conceptuel et régulateur des infinis – particulièrement celui d'infini

actuel – c'est prendre le risque de s'enfermer dans un constructivisme fini, voire borné, quelquefois justifié à tort par la omniprésence de l'ordinateur (voir *L'infini et l'univers des algorithmes*, pages 108 à 110). C'est aussi oublier que la théorie des fonctions récursives et de la calculabilité ont pour base conceptuelle la théorie des ensembles. Il en va de même pour le théorème d'arrêt d'Alan Turing et de ses conséquences sur le problème de la calculabilité.

L'infini, dans ses aspects potentiel et actuel, règle et structure l'activité mathématique. Faire des mathématiques, pensait Poincaré, c'est raconter quelque chose sur l'infini, alors même qu'aucun mathématicien n'a fait ou ne fera une démonstration ou un calcul infini...

Javier de LORENZO MARTÍNEZ est professeur de philosophie des sciences à l'Université de Valladolid.

Stan CUBILLICO, *Reaching for infinity*, Tab Books, 1994.

Infini des mathématiciens, infini des philosophes, sous la direction de F. Monnoyeur, Belin-Pour la Science, 1992.

D. HILBERT, *On the infinite*, sous la direction de P. Benacerraf et H. Putnam. *Philosophy of Mathematics : selected Readings*, Cambridge University Press, 1983.

A.-L. CAUCHY, *Analyse algébrique*, 1821, Éditions J. Gabay, 1989.

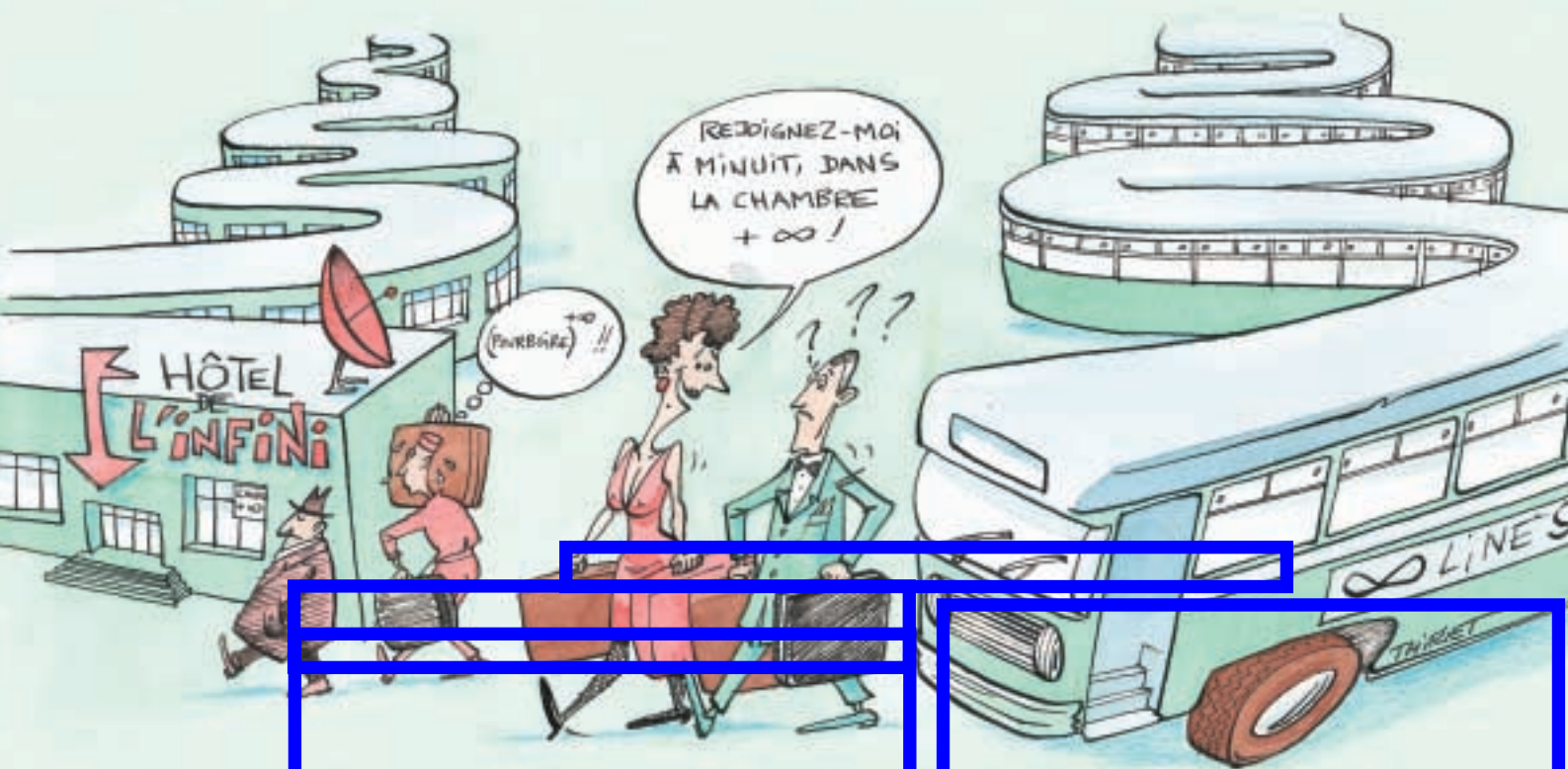
Ian STEWART, *Les mathématiques*, Belin-Pour la Science, 1989.

Paolo ZELLINI, *Breve Storia dell'Infinito*, Adelphi, Milan, 1980.

L'infini est-il paradoxal en mathématiques?

JEAN-PAUL DELAHAYE

Pour résoudre le paradoxe du tout et des parties et affronter l'hypothèse du continu, notre idée de l'infini actuel doit évoluer ; aujourd'hui encore, nous découvrons de nouveaux infinis.



1. LE PARADOXE DE L'HÔTEL DE HILBERT

L'histoire suivante illustre pourquoi des ensembles infinis actuels ont longtemps paru absurdes. Un hôtel infini dont les chambres sont numérotées par les entiers $0, 1, 2, 3, \dots$ est complet pour la nuit (un client occupe chaque chambre). Arrive un client : « Pas de problème », lui répond le responsable de l'accueil. Installez-vous dans la chambre 0. Je demande au client de la chambre 0 de passer dans la chambre 1, à celui de la chambre 1 de passer dans la chambre 2, etc. » L'accueil dispose bien sûr d'un téléphone spécial qui permet de téléphoner simultanément à toutes les chambres en demandant au client de la chambre n de passer en $n + 1$. Le nouveau client a pu être reçu. Dix minutes plus tard arrive un car (infini, bien sûr) de nouveaux clients qui demandent à passer la nuit dans l'hôtel. « Pas de problème », répond le responsable de l'accueil au chauffeur du car, et il utilise son téléphone

pour demander au client de la chambre n de passer dans la chambre $2n$. Il indique alors au chauffeur du car que le voyageur numéro i de son autocar peut disposer de la chambre $2i + 1$ (qui est effectivement libre, puisque toutes les chambres de numéro impair ont été libérées). Une demi-heure plus tard arrive un groupe plus important constitué d'une infinité de cars, chacun ayant à leur bord une infinité de passagers. « Pas de problème, je vous arrange ça », répond le responsable de l'accueil. Il demande au client de la chambre i de passer dans la chambre $2^i + 1$ (ce qui libère toutes les chambres ayant un numéro pair) et donne la consigne suivante au responsable du groupe d'autocars : le passager numéro i du bus j doit occuper la chambre $2^{i+1}(2j + 1)$. Tout se passe bien, et jamais deux voyageurs différents ne se sont vus attribuer la même chambre.

L'infini, c'est long, surtout vers la fin. (Alphonse Allais)

L'infini mathématique peut-il être maîtrisé? Autrement dit, peut-on faire une théorie de l'infini qui évite tout paradoxe et toute incohérence? Pour répondre à ces questions nous distinguerons paradoxes et situations logiquement peu satisfaisantes.

Un paradoxe au sein d'une théorie est la possibilité de démontrer une chose et son contraire. Dans cette situation, tout en n'utilisant que des raisonnements fondés sur les axiomes de la théorie nous pouvons en tirer une affirmation *A* et l'affirmation contraire *Non A*. En mathématiques, les paradoxes (on les dénomme aussi contradictions, inconsistances, antinomies) sont inacceptables, et les logiciens font tout pour les éviter.

Aujourd'hui, les mathématiciens connaissent toutes sortes de moyens pour contourner les paradoxes au sein des théories mathématiques de l'infini. La question subsiste : ces moyens atteignent-ils leurs fins?

Une situation logiquement insatisfaisante apparaît lorsqu'une théorie nous permet d'énoncer des propriétés étonnantes, parfois opposées à notre attente, sans toutefois qu'une véritable contradiction apparaisse. Nous pouvons alors continuer à développer la théorie en espérant que la difficulté sera résolue plus tard. Parfois nous finissons par accepter la situation jugée gênante, et ce qui était perçu comme douteux devient alors banal, comme si une métamorphose de nos conceptions profondes s'était produite.

Quelle est aujourd'hui la situation entre paradoxes infinis, situations insatisfaisantes tolérées et digérées?

Pour le savoir, nous allons passer en revue quelques paradoxes de l'infini mathématique, paradoxes qui décident de son histoire. À chaque fois, nous évoquerons les solutions mathématiques avancées et nous nous interrogerons pour savoir si ces solutions sont satisfaisantes en tout point ou si, au contraire, elles laissent des zones d'ombre qui marquent la limite de cette prétendue maîtrise de l'infini par le formalisme des mathématiques modernes.

L'infini n'a pas été accepté facilement, et on a longtemps espéré pou-

voir s'en passer. Aristote refusait l'infini actuel (ou infini actuel), c'est-à-dire pris d'un seul tenant. Il déniait toute existence physique à

l'infini, mais lui reconnaissait une certaine existence mathématique. Il lui semblait nécessaire d'envisager des grandeurs de plus en plus grandes, jusqu'à ce qu'un entier soit suivi d'un autre; aucun point n'est le dernier point d'une droite. Les mathématiciens ont tenté de se contenter de

cet infini potentiel ou en tout cas de s'y ramener, en évitant autant que possible l'infini actuel.

L'impassable infini actuel

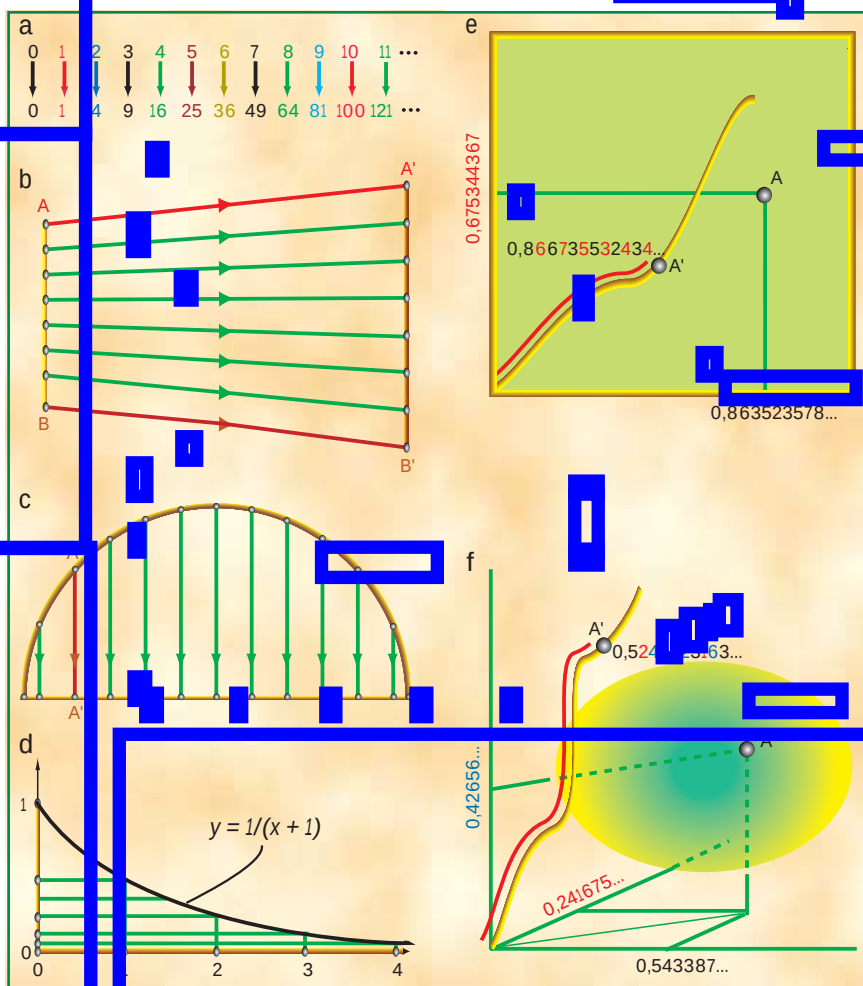
Euclide, par exemple, n'admettait pas qu'il existe une infinité de nombres premiers, mais que les nombres premiers sont en plus grande quantité que toute quantité de nombres premiers proposée», ce qu'il démontre en établissant que, si



Bernard Bolzano



Carl Friedrich Gauss



2. LES PARADOXES DE LA RÉFLEXIVITÉ : si les ensembles infinis existent en acte et non pas potentiellement, comme les ensembles que l'on complète petit à petit, alors l'ensemble des carrés est aussi grand que l'ensemble des nombres entiers : on peut associer chaque élément de l'un à un élément de l'autre (a). Similairement (b), le segment AB est aussi «grand» que le segment A'B', et le demi-périmètre d'un cercle est «égal» à son diamètre (c). On peut aussi mettre en correspondance (d) les points réels d'un segment et les points d'une demi-droite infinie par la relation $y = 1/(x + 1)$. Similairement, Cantor (il fut très surpris de cette possibilité) a mis en relation l'ensemble des points d'une surface plane avec les points d'une courbe quelconque. Pour cela, on associe aux deux coordonnées d'un point une valeur obtenue en intercalant les décimales des deux nombres représentant les coordonnées, et l'on associe à cette valeur un point de la courbe dont la distance à l'origine (prise le long de la courbe) est égale à la valeur obtenue (e). De la même façon, on associe les points d'un volume aux points d'une courbe (f).



Richard Dedekind

des nombres premiers sont donnés par avance $p_1, p_2, \dots, p_k$, on peut en construire un (le plus petit diviseur premier de $p_1 p_2 \dots p_k + 1$) qui sera différent de tous ceux qui étaient donnés.

La raison profonde de cette méfiance vis-à-vis de l'infini actuel est le paradoxe de la réflexivité : si un ensemble est infini, il est possible de le mettre en correspondance un à un – on dit aussi bijective, ou biunivoque – avec une de ses parties propres (c'est-à-dire différente de lui-même). La relation qui associe n^2 au nombre n , par exemple, établit une correspondance bijective entre les nombres entiers $0, 1, 2, 3, \dots$ et les nombres pairs $0, 1, 4, 9, \dots$ lesquels sont pourtant moins nombreux.

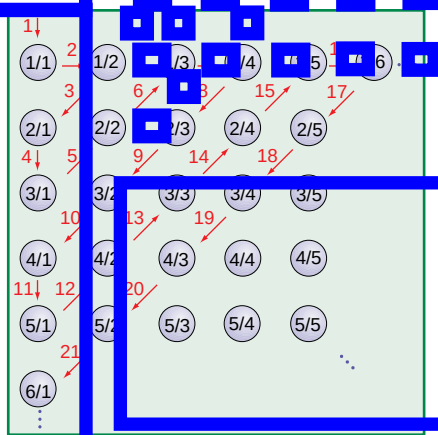
Le même problème surgit à propos du segment des nombres réels compris entre 0 et 1 (noté aujourd'hui $[0,1]$) et celui des nombres réels compris entre 0 et 2 (la correspondance bijective est celle qui, à x , associe $2x$) ; généralisant, on met sans aucune peine en correspondance bijective deux segments de droite AB et $A'B'$ quelconques. Plus gênante encore la correspondance qui à x associe $1/(x+1)$, car elle met en correspondance bijective l'intervalle des nombres réels compris entre 0 et 1, noté $[0, +1]$, et l'ensemble des nombres réels positifs, noté $[0, +\infty]$.

En quoi le paradoxe de la réflexivité est-il un paradoxe, nous demandons-nous aujourd'hui ? C'est un paradoxe, car le principe du tout et de la partie qui indique que «le tout est plus gros que la partie» y est mis en défaut. On n'imaginerait pas de renoncer à une vérité aussi claire. On craint



Gottfried Leibniz

que notre raison ne puisse résister à la remise en cause d'un tel principe immédiat. Reconsidérer ce principe du tout et de la partie apparaît d'une audace insensée, et c'est pourquoi on a bien souvent préféré conclure que seul un être infini lui-même, Dieu par exemple, peut penser l'infini. L'Église s'opposait d'ailleurs



3. L'ENSEMBLE INFINI des nombres rationnels (c'est-à-dire l'ensemble des nombres qui sont le quotient de deux nombres entiers) semble supérieur à l'ensemble des nombres entiers. Cantor démontra que l'on peut apparier les nombres rationnels dans un réseau carré et associer un nombre entier à chaque nombre rationnel en traçant le chemin coloré représenté sur la figure. L'ensemble des nombres rationnels est dénombrable !

à toute tentative des hommes de penser l'infini actuel. Saint Thomas d'Aquin considérait que quiconque pensait concevoir l'infini actuel entraînait en concurrence avec la nature unique et absolue de Dieu.

Il fallut attendre deux mille ans pour franchir cet obstacle. Le principe du tout et de la partie qui, à vrai dire, n'est guère utile en mathématiques, devait être reconsidéré. Ce principe éminemment paradoxal empêchait tout progrès dans la compréhension de l'infini actuel.

Cette audace sera le fait du philosophe mathématicien tchèque Bernhard Bolzano (1781-1848), dont l'ouvrage *Des paradoxes de l'infini*, publié après sa mort en 1851, envisage des



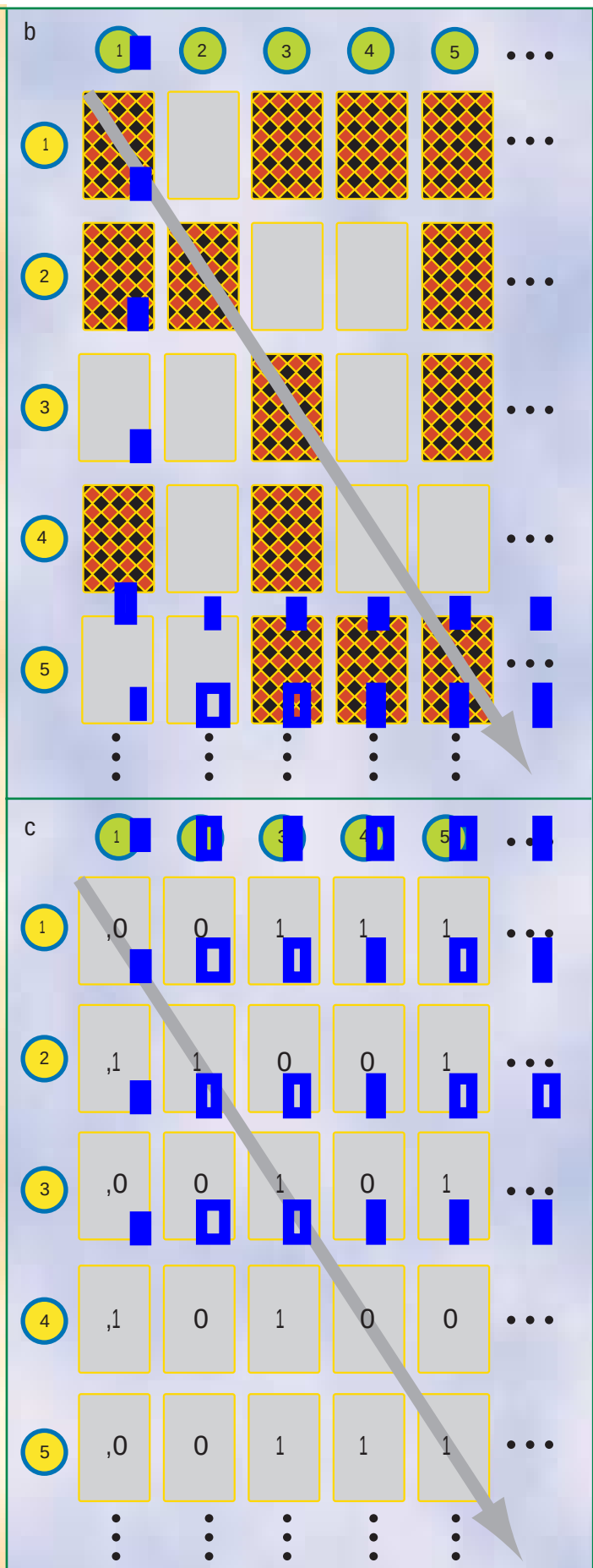
Georg Cantor

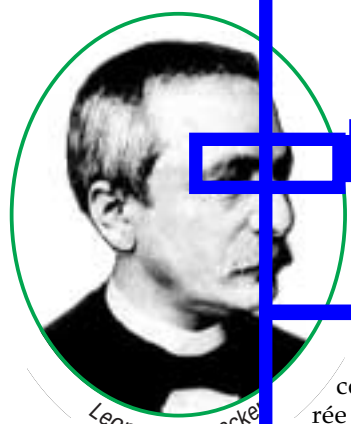
correspondances bijectives entre une totalité et l'une de ses parties propres, sans s'en émouvoir. Au contraire Bolzano propose de voir dans ces correspondances la caractéristique des totalités infinies, ce qui revient à abandonner, pour les totalités infinies, le principe du tout et de la partie. Plus tard, le mathématicien allemand Richard Dedekind (1831-1916) définira qu'un ensemble est infini s'il peut être mis en bijection avec une de ses parties propres : aujourd'hui, on adopte souvent cette définition en théorie des ensembles pour définir un ensemble infini.

Avant Bolzano, G. Leibniz avait défendu l'idée de l'infini actuel : «Je suis tellement pour l'infini actuel qu'au lieu d'admettre que la nature l'abhorre, comme l'on dit vulgairement, je tiens qu'elle l'affecte partout, pour mieux marquer la perfection de son Auteur.» (Leibniz, *Opera omnia studio Ludov. Duten...*, tome II, partie X.) Bolzano place cette citation en exergue de son ouvrage révolutionnaire.

Cependant l'infini actuel que défend Leibniz est différent de celui des totalités infinies considérées actuel dans le problème du paradoxe de la réflexivité. Leibniz défend plutôt un infini philosophique, celui du monde physique pris comme un tout. L'infiniment petit (qui est confronté

4. CANTOR a prouvé qu'il existe des infinis plus grands que l'infini dénombrable (ensemble dont on peut numéroter tous les éléments). Examinons d'abord un ensemble fini comportant trois éléments : nous dénombrons huit sous-ensembles (2^3) distincts (a). Pour un ensemble de N éléments, il existe 2^N sous-ensembles différents (chaque élément peut être, ou non, partie de l'ensemble). Construisons les sous-ensembles d'un ensemble infini dénombrable d'éléments (b). Nous démontrons par l'absurde qu'il est impossible de numéroter chaque élément de cet ensemble, donc que l'ensemble infini des sous-ensembles d'un ensemble dénombrable n'est pas dénombrable. Supposons que cela soit possible. Chaque sous-ensemble est représenté par un ensemble de cartes, chaque sous-ensemble s'étendant à l'infini vers la droite. Nous numérotions chaque sous-ensemble de haut en bas. Les numérotions nous ainsi tous ? Non. Examinons le sous-ensemble constitué par les cartes de la diagonale et supposons que nous avons retourné toutes les cartes : il n'est pas identique au premier sous-ensemble, car la première carte est différente ; il n'est pas identique au second, car la seconde carte, retournée, est différente, et ainsi de suite. Le n -ième sous-ensemble diffère du nouveau sous-ensemble, car la carte numéro n est retournée. Nous avons ainsi corrigé un sous-ensemble qui ne fait pas partie de la liste. Donc l'ensemble des sous-ensembles n'est pas numérotable... Cette preuve diagonale de Cantor sert à prouver que l'ensemble des nombres réels (les nombres rationnels et les nombres irrationnels) n'est pas dénombrable. Examinons (c) les nombres réels du segment de longueur unité : la distance de chaque point à l'origine est écrite sous forme binaire 0,011010010111... Le point dont la distance à l'origine est mesurée par les chiffres inversés de la diagonale ne figure pas dans l'énumération, et il est donc impossible de numéroter tous les nombres réels du segment unité... L'ensemble des nombres réels à la puissance $2^{\aleph_0}$ est un infini plus grand que l'infini dénombrable.





Leopold Kronecker



Kurt Gödel

d'autres paradoxes qu'on utilise parfois pour attaquer les nouvelles méthodes de calcul que Leibniz défend) ne saura vraiment trouver une place confortable et assurée en mathématiques qu'en se débarrassant de son objectivité, c'est-à-dire quand on ne parlera plus des infiniment petits comme d'objets mathématiques, mais comme de limites.

Les mathématiciens du XX^e siècle sauront reconnaître aux infiniment petits un statut d'objet authentique. La méthode utilisée au XIX^e siècle pour rendre rigoureux le calcul infinitésimal est un renoncement à l'infini actuel, auquel on substitue un infini potentiel, celui de quantités qui s'approchent de plus en plus de leur limite.

Le prince des mathématiciens, Gauss (1777-1855), exprimant le sentiment partagé par la communauté mathématique de son époque, écrivait par exemple : «Je conteste qu'on utilise un objet infini comme un tout complet ; en mathématiques, cette opération est interdite. L'infini n'est qu'une façon de parler.»

On peut donc affirmer que, malgré quelques tentatives diverses pour fonder une science mathématique de l'infini, c'est Bolzano et nul autre qui, en affrontant le paradoxe de la réflexivité, a le premier ouvert la voie à ce qui aujourd'hui est notre conception de l'infini mathématique pris comme un tout.

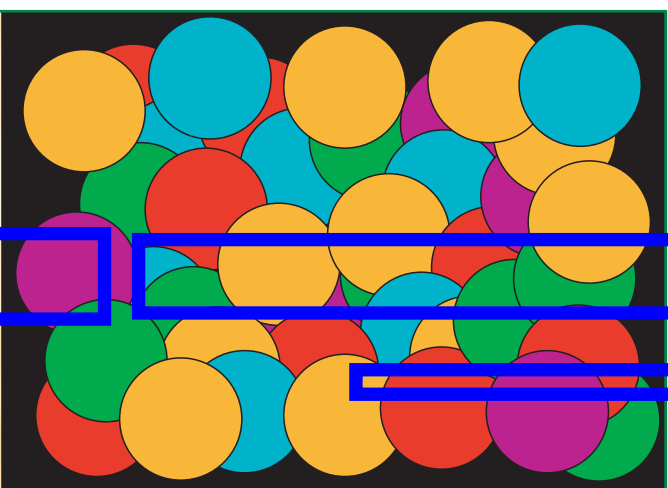
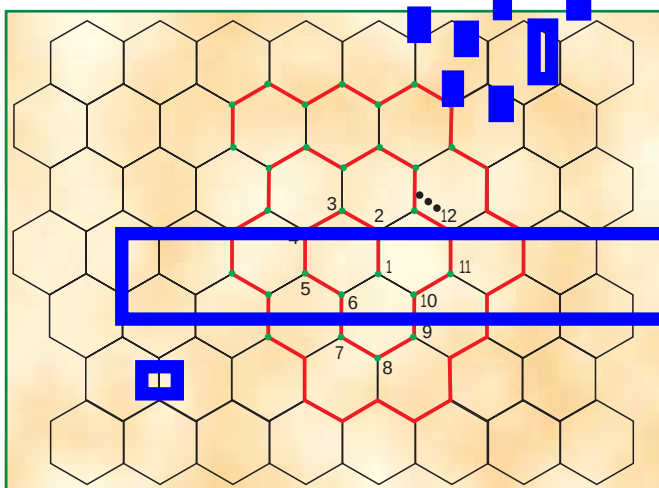
La solution actuelle du paradoxe de la réflexivité

La solution du paradoxe de la réflexivité proposée par Bolzano et rendue parfaitement claire par le développement ultérieur de la théorie des ensembles en termes modernes s'énonce ainsi : la relation «est contenu dans» entre ensembles ne doit pas être confondue avec la relation «avoir une taille plus petite que» : les nombres carrés sont contenus dans les nombres entiers, mais comme totalité ils ont même taille. Il est bien vrai que, si l'ensemble A est contenu dans l'ensemble B, alors la taille de A est inférieure ou égale à celle de B, mais elle peut être égale.

Au lieu de «taille», on emploie souvent le terme technique de «cardinal», mais c'est sans importance, l'idée unique de la solution du paradoxe est d'admettre qu'un ensemble A strictement contenu dans un ensemble B a parfois la même taille. Il faut accepter ce qui paraissait paradoxal et décréter que ça ne l'est plus par une opération de dédoublement de concept : «être contenu dans» n'est pas «avoir une taille plus petite que».

Ce pas franchi, n'allons-nous pas tomber sur des contradictions qui conduiraient à une mathématique inconsistante, et donc intenable ? Le jeu en vaut certainement la chandelle, mais il n'est pas facile à mener et ce sont d'autres que Bolzano qui le joueront, car, il faut bien le dire, Bolzano, comme essoufflé par l'audace de sa proposition initiale, ne mène pas le développement mathématique qu'elle appelle pourtant.

L'exploration mathématique de la réflexivité des bijections sera longue et tortueuse, et Bolzano inaugure une période d'instabilité grâce et de controverse parfois acerbe. Le travail principal sera fait par le mathématicien allemand Georg Cantor (1845-1918), qui découvrira de nombreuses propriétés des tailles d'ensembles infinis qui lui sembleront souvent à la limite du paradoxe. Cantor distingue les tailles des ensembles. Si le contraire avait été vrai – c'est-à-dire si chaque ensemble infini avait pu être mis en bijection avec chaque autre –, la théorie de la taille des ensembles infinis aurait singulièrement manqué d'intérêt ! Cantor s'aperçoit, en 1874, que l'ensemble des nombres réels ne peut pas être mis en bijection avec l'ensemble des nombres entiers : il est de taille strictement plus grande. Le raisonnement diagonal que Cantor découvre en simplifiant son raisonnement de 1874 sera le prototype même du raisonnement permettant de démontrer des résultats d'impossibilité en mathématiques. Dans ce théorème, on suppose que ce qu'on veut établir est faux (ici, qu'il existe une bijection entre l'ensemble $\mathbb{N}$ des entiers et l'ensemble $\mathbb{R}$ des réels) puis on joue sur la diagonale du tableau qu'elle définit (le i-ème chiffre du nombre réel f(i)), et l'on déduit une contradiction.



5. La différence entre $\aleph_0$ et $2^{\aleph_0}$ est importante en géométrie, quand apparaissent des ensembles infinis. Ainsi le nombre de sommets d'un pavage hexagonal du plan est $\aleph_0$, car on peut tracer une spirale (en

rouge), où l'on numérote tous les sommets. En revanche, le nombre de cercles que l'on peut disposer sur une feuille de taille finie est un infini $2^{\aleph_0}$, car le centre de chaque cercle est un point du plan.

Mieux (ou pire!), Cantor montre, toujours par un raisonnement diagonal, qu'il existe une infinité de tailles possibles différentes pour les ensembles infinis. Plus précisément, un ensemble E ne peut jamais être mis en bijection avec l'ensemble de ses parties, noté $P(E)$. L'ensemble $P(\mathbb{N})$ des parties de $\mathbb{N}$ (qui est composé de l'ensemble des nombres pairs, de l'ensemble des nombres premiers, etc.) ne peut pas être mis en bijection avec $\mathbb{N}$. De même, l'ensemble $P(P(\mathbb{N}))$ (ensemble des parties de l'ensemble $P(\mathbb{N})$) ne peut pas être mis en bijection avec $P(\mathbb{N})$; etc.

Ces résultats constituent une première avancée dans la compréhension de l'infini actuel et la preuve que ce qu'on découvre est digne d'intérêt : les résultats construisent une hiérarchie des totalités infinies où l'esprit trouve ses repères. Ces premiers résultats ne sont pas sans soulever reproches et critiques. Le célèbre mathématicien Kronecker est particulièrement sévère et reprochera un manquement de Cantor en retardant la parution dans le *Journal de Crelle*, l'un des plus prestigieux journaux de mathématiques, auquel Cantor refusera par la suite de proposer ses travaux.

Cantor découvre un résultat étonnant

L'article dont Kronecker a retardé la publication contient un résultat de Cantor tout à fait étonnant qui, bien que ne constituant pas un paradoxe, fut à l'époque considéré comme créant une situation logiquement peu satisfaisante. Cantor, toujours occupé à classer les infinis, a en effet découvert avec stupeur que l'ensemble des points d'une surface (un carré, par exemple) possède la même taille que l'ensemble des points d'un segment de droite. Du point de vue de leur taille d'ensemble infini, une droite et un plan (ou même un espace de dimension n) sont identiques. Il écrira à Dedekind à ce propos : «Je le vois, mais je ne le crois pas.»

Cantor adopte l'attitude téméraire qui consiste à accepter cette nouvelle vérité sans la déclarer paradoxale, malgré la gêne qu'elle engendre. Comme aucune contradiction véritable ne découle des raisonnements et des calculs que Cantor élabore avec soin, il faut avoir l'audace d'accepter ce que nos facultés de raisonnement produisent et, aussi déconcertante que soit

L'HYPOTHÈSE DU CONTINU ET LA NOTATION DES ALEPH

Cantor proposait de classer les infinis de la façon suivante :

- L'infini des entiers est noté $\aleph_0$ (ce qui se lit aleph zéro)
- L'infini suivant est noté $\aleph_1$, et

Par ailleurs, l'infini des nombres réels (le continu) est noté $2^{\aleph_0}$.

La justification de cette notation est que, d'une part, on démontre que l'ensemble des nombres réels a la même taille que l'ensemble des parties de $\aleph_0$, et que d'autre part, dans le cas fini, on montre que le nombre de parties d'un ensemble à n éléments est 2^n (par analogie, on note donc 2^a la taille de l'ensemble des parties d'un ensemble de taille a).

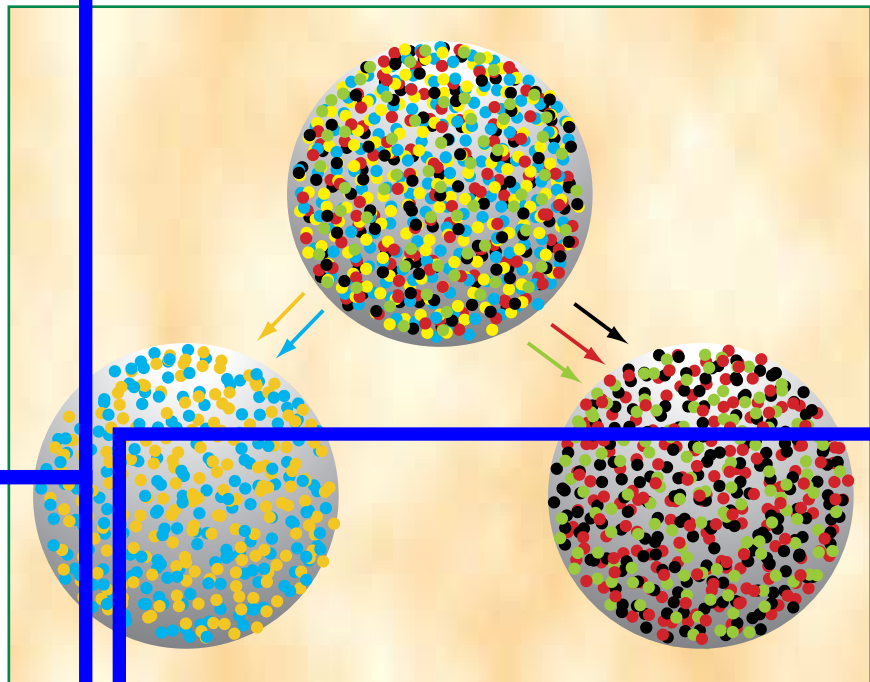
L'échelle des infinis donnée par le théorème fondamental de Cantor est notée : $\aleph_0, 2^{\aleph_0}, 2^{2^{\aleph_0}}$, etc.

La question de savoir s'il n'y a rien entre $\aleph_0$ et $2^{\aleph_0}$ s'écrit : $\aleph_1$ est-il égal à $2^{\aleph_0}$?

L'affirmation $\aleph_1 = 2^{\aleph_0}$ est appelée hypothèse du continu.

L'affirmation que, pour tout i : $\aleph_{i+1} = 2^{\aleph_i}$ est appelée hypothèse généralisée du continu.

Ces affirmations ont été démontrées indépendantes des axiomes classiquement acceptés pour la théorie des ensembles : jamais il ne sera possible à partir d'eux de savoir ce qu'il en est de la vérité de ces affirmations. Cela ne signifie cependant pas que tout espoir soit perdu de savoir si ces affirmations sont vraies ou fausses : il est possible que certains axiomes concernant les ensembles aient été oubliés et que, lorsque nous les découvrirons et les ajouterons aux axiomes connus, l'hypothèse du continu devienne démontrable.



6. LE PARADOXE DE BANACH-TARSKI : on établit qu'une sphère peut être décomposée en un nombre fini de morceaux, qui, une fois déplacés (sans déformation), se recomposent en deux sphères identiques à la sphère de départ. Plus généralement, une sphère peut être transformée en un cube sans aucune condition de volume sur la sphère initiale et sur le cube final. Et plus généralement encore, si deux parties A et B de l'espace sont de taille bornée et contiennent chacune au moins une sphère de rayon positif, alors on peut décomposer A en un nombre fini de morceaux qui, après déplacement, reconstitueront B . Ces découpages paradoxaux ont été découverts en 1924 par les mathématiciens polonais Stephan Banach et Alfred Tarski, qui travaillaient sur la théorie de la mesure (théorie abstraite des aires et des volumes). Bien que ces découpages choquent profondément le sens commun, qui s'attend à ce qu'en déplaçant des morceaux d'un objet on n'en modifie pas le volume, il n'y a pas «paradoxe» au sens strict : aucune contradiction n'est introduite dans la théorie des ensembles. Ces découpages montrent seulement que le monde du continu mathématique et les conséquences de l'axiome du choix ne correspondent pas à nos attentes intuitives.

PARADOXE DE RUSSELL

Soit l'ensemble des x qui ne sont pas éléments d'eux-mêmes : $E = \{x ; x \notin x\}$. Si $E \in E$, alors, par définition, $E \notin E$. Si $E \notin E$, alors, par définition, $E \in E$: c'est contradictoire !

Le paradoxe de l'ensemble des ensembles qui ne se contiennent pas eux-mêmes (paradoxe de Russell) est aujourd'hui correctement résolu en théorie des ensembles de la façon suivante.

On définit des méthodes qui permettent de créer de nouveaux ensembles à partir d'ensembles existant déjà (réunion, ensemble des parties, etc.), et il ne suffit pas de considérer une propriété (par exemple, $x \notin x$) pour pouvoir affirmer qu'il existe un ensemble regroupant tous les objets qui satisfont cette propriété. La notation $E = \{x ; x \notin x\}$ ne désigne pas nécessairement un objet légitime de la théorie des ensembles sur lequel on peut raisonner comme le fait le paradoxe ; si on veut l'utiliser, il faut établir que c'est un ensemble par utilisation des règles que fixe la théorie. Pour $E = \{x ; x \notin x\}$, on n'y arrive pas. Dans la théorie Zermelo-Fraenkel (désignée par ZF), la présentation correcte du raisonnement ci-dessus est donc :

Supposons qu'il existe un ensemble défini par : $E = \{x ; x \notin x\}$.

Si $E \in E$, alors, par définition, $E \notin E$. Si $E \notin E$, alors, par définition, $E \in E$, c'est contradictoire, donc il n'existe pas d'ensemble défini par l'égalité : $E = \{x ; x \notin x\}$.

Pour des raisons analogues, il n'existe pas d'ensemble de tous les ensembles, ni d'ensemble de tous les cardinaux, ou d'ensemble de tous les ordinaux. Cette solution classique aux paradoxes de la théorie des ensembles, découverts au début du XX^e siècle, réussit bien. Aucun nouveau paradoxe n'a été découvert dans la théorie ZF jusqu'à présent.

cette nouvelle réalité de l'infini actuel, il faut en poursuivre l'exploration.

À vrai dire, au tournant du XX^e siècle, certaines contradictions (ce qu'on appellera les antinomies) sembleront rendre la construction cantorienne intenable. Toutefois, en quelques années, ces difficultés seront circonscrites (plusieurs méthodes seront même proposées) et ce qui apparut à certains comme la faillite de la théorie mathématique de l'infini actuel sera en définitive l'occasion pour cette théorie de se forger une armure qu'aujourd'hui rien n'est venu sérieusement érafler.

L'hypothèse du continu

Revenons à Cantor qui, en théoricien consciencieux, développe une arithmétique de l'infini, c'est-à-dire une extension aux nombres qui lui servent à mesurer l'infini, des règles de calcul qu'on applique aux nombres entiers, servant à mesurer le fini. À cette occasion, il est contraint de distinguer deux types de nombres infinis, les cardinaux et les ordinaux. Les cardinaux sont les tailles des ensembles quand on les considère de manière brute, sans tenir compte d'un possible ordre entre leurs éléments. Cantor introduit en 1893 la notation $\aleph_0$ pour le cardinal de l'ensemble des nombres entiers, $2^{\aleph_0}$ pour le cardinal du continu (l'ensemble

des nombres réels), qui est le même que celui de $P(\mathbb{N})$, l'ensemble des parties de $\mathbb{N}$. Il utilise $2^{2^{\aleph_0}}$ pour le cardinal de $P(P(\mathbb{N}))$, etc.

L'autre type de nombres infinis de Cantor, les ordinaux, sert à mesurer la taille des ensembles lorsque leurs éléments sont ordonnés selon un bon ordre (un ordre tel que toute partie de l'ensemble possède un plus petit élément). Au-delà des ordinaux finis, qu'on assimile aux nombres entiers, il y a ω , le premier ordinal transfini, puis $\omega+1$, $\omega+2$, etc., qu'on obtient en ajoutant des unités une à une à ω . Plus loin encore, on trouve 2ω , $2\omega+1$,... puis ω^2 , ω^ω , etc.

Cantor découvre alors un problème qui finalement se révélera d'une difficulté qu'aujourd'hui encore nous n'avons pas totalement maîtrisée : l'hypothèse du continu.

Le cardinal de l'ensemble des nombres réels $\mathbb{R}$, ce qu'on appelle le continu et qui est mesuré par $2^{\aleph_0}$, est strictement plus grand que le cardinal des entiers (appelé infini dénombrable et mesuré, nous l'avons dit, par $\aleph_0$). Mais y a-t-il un autre cardinal entre les deux ? L'hypothèse du continu est l'affirmation qu'entre ces deux tailles d'ensembles infinis, $\aleph_0$ et $2^{\aleph_0}$, il n'y en a pas d'autres. Comme Cantor désigne par $\aleph_1$ le plus petit des cardinaux plus grands que $\aleph_0$,

l'hypothèse du continu est simplement l'affirmation $2^{\aleph_0} = \aleph_1$. (L'hypothèse généralisée du continu est l'affirmation que $2^{\aleph_s} = \aleph_{s+1}$ pour tout s .)

Deux situations sont envisageables : soit (i) il n'y a pas d'autres tailles infinies entre $\aleph_0$ et $\aleph_1$, ce qui est équivalent à dire que « tout sous-ensemble infini de $\mathbb{R}$ peut être mis en bijection avec $\mathbb{N}$ ou avec $\mathbb{R}$ » (c'est là un énoncé simple et, en apparence, très concret) ; soit (ii) il y en a d'autres et alors il faudrait savoir combien et comment on les obtient. Nul ne peut prétendre disposer d'une telle liste des nombres réels, il n'y a donc pas de réponse précise au problème de l'hypothèse du continu. Cantor s'épuiera à le résoudre et n'y réussira pas, bien qu'il ait parfois pensé, à tort, avoir résolu la question.

L'indécidabilité ne résout rien

Plus tard, la théorie des ensembles formalisée, deux résultats importants furent démontrés concernant l'hypothèse du continu. En 1938, le mathématicien autrichien Kurt Gödel prouva que, si la théorie habituelle des ensembles est consistante (ne conduit pas à une contradiction) alors la même théorie complétée par l'axiome qui affirme que l'hypothèse du continu est vraie ne conduit pas plus à une contradiction. Admettre l'hypothèse du continu n'est donc pas susceptible de faire s'effondrer la théorie usuelle des ensembles. En 1963, Paul Cohen ira plus loin : si la théorie des ensembles est consistante, alors la même théorie complétée en ajoutant l'affirmation que l'hypothèse du continu est fausse est, elle aussi, sans contradiction.

Ces deux résultats regroupés permettent de dire que celui qui admet la théorie usuelle des ensembles peut, sans risque supplémentaire de contradiction, prendre l'hypothèse du continu ou prendre sa négation. La théorie des ensembles laisse le choix aux mathématiciens de considérer l'hypothèse du continu comme vraie ou comme fausse.

Cette indécidabilité de l'hypothèse du continu n'est pas un paradoxe : la théorie ne se contredit pas, au contraire elle ne dit rien. Toutefois cette sous-détermination est logiquement très insatisfaisante et profondément troublante.

En effet, considérons le monde des nombres entiers : il n'est pas arbitraire, ce n'est pas nous qui décidons

si un nombre est pair ou impair, premier ou composé, etc. Pourquoi donc le monde des nombres réels serait-il arbitraire, et nous permettrait de décider en fonction de nos goûts qu'il existe de partie infinie de $\mathbb{R}$ impossible à mettre en bijection avec $\mathbb{N}$ ou $\mathbb{R}$, ou qu'au contraire il n'en existe pas ?

Comme le pensait Cantor lui-même, ainsi que Gödel, l'hypothèse du continu doit être vraie ou fausse, et la théorie des ensembles doit nous permettre de le savoir, à moins qu'elle ne soit qu'une illusion, ce qui signifierait que la compréhension qu'elle nous donne de l'infini actuel est factice. La situation créée par les résultats de Gödel et de Cohen sur l'hypothèse du continu est logiquement insatisfaisante, mais c'est toute la signification de la théorie de l'infini actuel proposée par Cantor qui est en cause, et, finalement, c'est l'idée même d'une théorie de l'infini actuel qui se joue.

Voies pour aller plus loin

L'indécidabilité de l'hypothèse du continu ne règle donc pas la question de l'infini, mais au contraire l'exacerbe : si vraiment la théorie des ensembles fournit une compréhension de l'infini actuel et n'est pas qu'un jeu gratuit entre symboles mathématiques, sans contrepartie dans le monde réel, il doit être possible d'avancer encore.

Une première voie consiste à remettre en cause la théorie usuelle des ensembles. Cette théorie, notée ZF (en l'honneur des deux mathématiciens E. Zermelo et A. Fraenkel qui la définirent au début du XX^e siècle), peut être jugée insatisfaisante pour diverses raisons. En particulier, parce qu'elle engendre des situations contraires à l'intuition concernant l'existence d'objets mathématiques qu'on ne peut pas construire (voir la figure 7). Diverses éventualités, théorie des types, théorie constructiviste des ensembles, théorie des ensembles avec ensemble universel, etc., ont été proposées sans toutefois attirer vraiment l'attention des mathématiciens, qui restent attachés à la simplicité de Zermelo-Fraenkel et ne semblent pas trouver très grave la situation créée par l'hypothèse du continu. Il est vrai que ces théories

alternatives obligent à réévaluer tout le travail mathématique déjà fait, alors qu'aucune véritable contradiction n'est apparue. Seule une insatisfaction philosophique pousse certains à vouloir abandonner Zermelo-Fraenkel.

PARADOXE DE L'AUTORÉFÉRENCE ET DE L'INFINI

Chacun connaît le paradoxe du menteur. Un personnage dit «je mens». S'il dit vrai, alors il ment, donc il ne dit pas vrai : c'est absurde. S'il ment, alors ce qu'il dit est vrai, et donc il ne ment pas : c'est absurde. On résout ce paradoxe en disant qu'il faut s'interdire les phrases autoréférentes. Les choses ne sont peut-être pas si simples, car on s'est aperçu récemment que l'introduction de l'infini permet d'éviter l'autoréférence tout en maintenant une situation paradoxale.

Soit une infinité de personnes placées en rang : $s(0)$, $s(1)$, ..., $s(n)$, ...

Chacune dit : «Au moins une personne derrière moi ment.»

Qui dit vrai ? Qui ment ?

D'après le sens des phrases prononcées :

- derrière toute personne qui dit vrai, il y a au moins une personne qui ment ;
 - si une personne ment, alors toutes les personnes derrière elle disent vrai.
- Si on désigne par M les personnes qui mentent et par H celles qui sont honnêtes, et ne mentent pas, les deux règles précédentes se traduisent :
- derrière tout H, il y a au moins un M ;
 - derrière un M, il n'y a que des H.

Il est impossible de concevoir une suite de M et de H qui vérifie ces deux règles (car tout M doit être suivi uniquement de H, ce qui ne se peut pas, puisque tout H doit être suivi d'au moins un M). Comme dans le cas du menteur mais cette fois à cause de l'infinité des personnes, la situation est paradoxale. Ces paradoxes, dits sémantiques, sont encore examinés par les logiciens, et aucune solution totalement satisfaisante ne semble avoir été trouvée.

La seconde voie consiste à admettre que Zermelo-Fraenkel est satisfaisante, c'est-à-dire qu'elle n'énonce que des choses acceptables pour les ensembles, mais qu'elle est incomplète. Certains axiomes manqueraient, et le travail du logicien et du mathématicien serait de les rechercher pour les ajouter. Ces axiomes découverts (ou au moins certains d'entre eux), l'hypothèse du continu ou sa négation deviendrait prouvable.

La seconde voie n'est pas du tout absurde, en insistant sur l'incomplétude actuelle de la théorie des ensembles, on ne fait que prendre au sérieux les théorèmes généraux d'incomplétude démontrés par Kurt Gödel en 1931, dont l'indécidabilité de l'hypothèse du continu ne serait qu'une manifestation en théorie des ensembles. Cette seconde voie conduit au programme de recherche de trouver de nouveaux axiomes à ajouter à la théorie des ensembles, programme que Kurt Gödel lui-même a défendu.

Les grands cardinaux

Les travaux faits en logique mathématique sur ce que l'on dénomme les grands cardinaux constituent le domaine de recherche d'où pourrait provenir la solution de l'énigme de l'hypothèse du continu (une partie infinie

de $\mathbb{R}$ est en bijection avec $\mathbb{N}$ ou $\mathbb{R}$), tout en étant le domaine où les mathématiciens rencontrent de nouveaux infinis actuels dont la taille vertigineuse aurait fait tourner la tête à Cantor lui-même.

Les axiomes de grands cardinaux sont des affirmations portant sur des ensembles monstrueusement grands. Délicats à définir, ils procèdent de principes analogues à celui qui consiste à affirmer que non seulement $\mathbb{N}$, $P(\mathbb{N})$, $P(P(\mathbb{N}))$, ... sont des infinis actuels légitimes, mais qu'il y en a de plus grands encore dont certains qui contiennent à la fois $\mathbb{N}$, $P(\mathbb{N})$, $P(P(\mathbb{N}))$, ...

Les axiomes de grands cardinaux sont vis-à-vis de la théorie Zermelo-Fraenkel comme l'est l'hypothèse du continu : on peut choisir, sans risquer d'introduction de contradiction, de les ajouter ou d'ajouter leur négation (en fait, les ajouter pourrait, en théorie, introduire une contradiction, mais personne en pratique n'en trouve jamais. Toutefois, contrairement à ce qui se passe pour l'hypothèse du continu, il est naturel de les tenir pour vrais. En effet, adopter leur négation serait limiter la taille des ensembles possibles, ce qui n'est pas naturel. De plus, cette même négation contredit un principe d'ouverture et de tolérance qui suggère d'accepter tous les objets qui n'introduisent pas d'incohérence. On dit que les axiomes des

grands cardinaux sont vrais *a priori*. Une autre raison conduit à accepter les axiomes des grands cardinaux. D'une façon qui a étonné les mathématiciens, ces axiomes de grands cardinaux se classent linéairement les uns par rapport aux autres comme s'ils désignaient véritablement une hiérarchie infinie (au-delà de celle mise en lumière par Cantor et la complétant). S'ils n'étaient qu'un

jeu formel et gratuit de logiciens, rien n'expliquerait qu'ils s'ordonnent ainsi les uns relativement aux autres. Pour les logiciens K. Kanamori et M. Magidor, «cet aspect hiérarchique de la théorie des grands cardinaux est quelque peu mystérieux, mais c'est aussi un argument fort en faveur de l'adoption des axiomes de grands cardinaux et du fait qu'ils fournissent des exten-

sions naturelles de Zermelo-Fraenkel».

La possibilité de trouver des liens entre les axiomes de grands cardinaux et l'hypothèse du continu est réelle, et des résultats partiels ont déjà été obtenus : certaines parties infinies de $\mathbb{R}$ dont on ne sait pas montrer dans Zermelo-Fraenkel qu'elles peuvent être mises en bijection avec $\mathbb{N}$ ou $\mathbb{R}$, peuvent être mises en bijection avec $\mathbb{N}$ ou $\mathbb{R}$ grâce aux axiomes de grands cardinaux.

La théorie que Cantor a créée semblerait donc tenir debout et, mieux que cela, progresser par ajouts de nouveaux axiomes naturels. L'infini actuel n'est ni paradoxal, ni logiquement insatisfaisant comme l'indécidabilité de l'hypothèse du continu nous a conduit à le craindre (un moment), mais au contraire cohérent et vraisemblable. Comme en physique, où notre conception du temps et de l'espace a dû être revue à la lumière de la théorie de la relativité (laquelle, en dépit des premières apparences, n'est nullement paradoxale), en mathématiques la conception de l'infini actuel force à reconstruire notre idée des objets et à revoir notre conception de la réalité mathématique. À celui qui accepte cette réforme profonde des conceptions, on ne peut opposer aucun paradoxe. On peut même dire que les questions logiquement insatisfaisantes qu'on croit encore percevoir se dissiperont à mesure que notre esprit acceptera pleinement le nouvel univers conceptuel proposé par ces mathématiques de l'infini actuel, qu'aujourd'hui encore les mathématiciens affinent et perfectionnent.

Jean-Paul DELAHAYE est professeur d'informatique théorique à l'Université de Lille.

A. KANAMORI, *The Higher Infinite*, Springer-Verlag, 1994.

J.W. DAUBIN, *Georg Cantor : His Mathematics and Philosophy of the Infinite*, Harvard University Press, 1979.

J.-L. TRIVIE, *Théorie des ensembles*, Cassini, Paris, 1998.

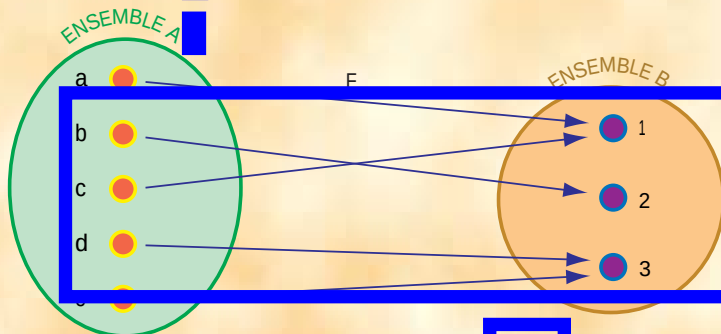
Bernard BOLZANO, *Les paradoxes de l'infini*, introduction, traduction de l'allemand et notes par Hourya Sinaceur, Seuil, 1993.

Jean-Paul DELAHAYE, *Jeux mathématiques et mathématiques des jeux*, chapitre 7 : Des jeux infinis aux grands ensembles, Belin/Pour la science, 1998.

Infini des mathématiciens, infini des philosophes, ouvrage collectif sous la direction de F. Monnoyeur, Belin, 1992.

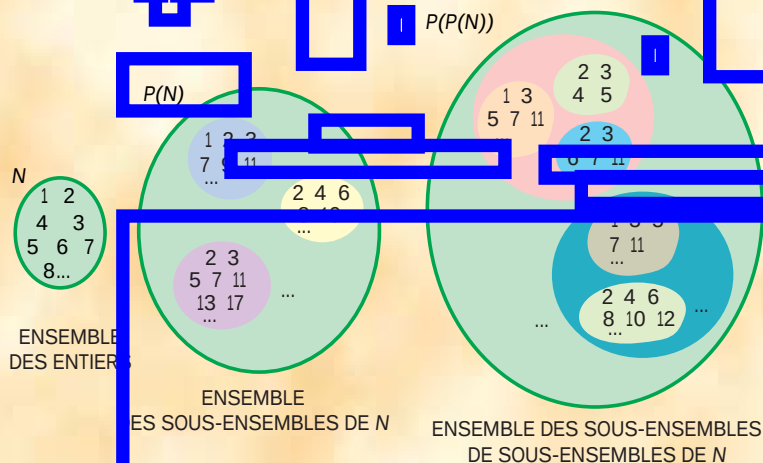
UN AXIOME DÉFINISSANT UN GRAND CARDINAL

Une surjection de A vers B est une application F qui atteint tout point de B : pour tout élément β dans l'ensemble B , il existe un élément α de l'ensemble A tel que $F(\alpha) = \beta$.



S'il existe une surjection de l'ensemble A vers l'ensemble B , on dit que A est plus gros que B . Si, de plus, il n'existe pas de surjection de B vers A , on dit que A est strictement plus gros que B .

Cantor a montré que l'ensemble $P(A)$ des parties de l'ensemble A est toujours strictement plus gros que l'ensemble A . Il en résulte qu'il existe une infinité d'ensembles de plus en plus gros, N , $P(N)$, $P(P(N))$, et, accessoirement, qu'il n'existe pas d'ensemble de tous les ensembles.

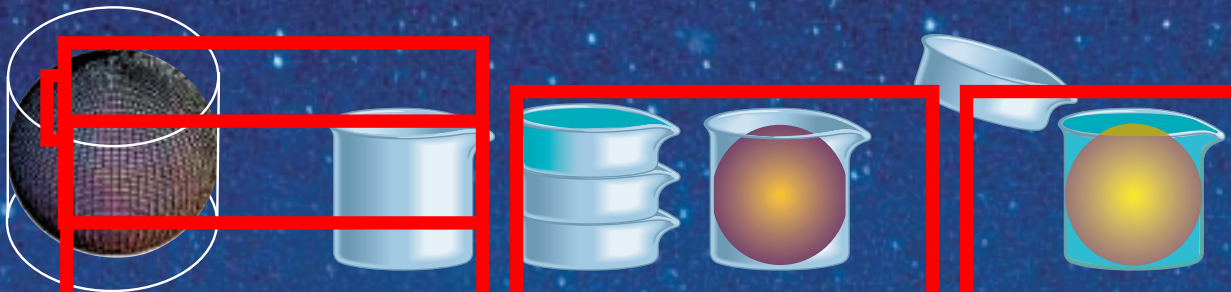


Voici un exemple d'axiome de grands cardinaux : il existe un ensemble non dénombrable A tel que, si A est strictement plus gros que B , alors A est aussi strictement plus gros que l'ensemble $P(B)$ des parties de B , plus gros que l'ensemble $P(P(B))$ des parties de $P(B)$, etc. et plus généralement, plus gros que tout ensemble définissable à partir de l'ensemble B . Le cardinal de A est un grand cardinal. Cependant ces axiomes sont, dans la grande majorité des cas, des indécidables de Gödel : avec les axiomes habituels de la théorie des ensembles, on ne peut prouver ni qu'ils sont vrais, ni qu'ils sont faux. Il faut donc les accepter tels quels, sans que l'on puisse espérer les démontrer à partir des axiomes de la théorie des ensembles. Ils sont pourtant utiles.

Archimède face à l'innombrable

ILAN VARDI

Archimède a inventé une notation pour compter de très grands nombres, mais une question reste ouverte : pourquoi s'est-il arrêté à son «plus grand nombre»?



«Certains estiment, roi Gélon, que le nombre de grains de sable est infiniment grand, et j'entends non seulement le sable des environs de Syracuse et du reste de la Sicile, mais encore celui qui est répandu dans toute la Terre, habitée ou inculte. D'autres, tout en admettant que ce nombre n'est pas infiniment grand, pensent qu'il n'existe pas de nombre exprimable assez grand pour dépasser la quantité de sable. Cependant, si ceux qui sont de cet avis s'imaginaient un volume de sable égal à celui de la Terre, ce volume comprenant toutes les mers et les vallées de la Terre remplies de sable jusqu'aux plus hauts montages, il est évident qu'ils reconnaîtraient encore beaucoup moins qu'on puisse énoncer un nombre dépassant le nombre de ces grains de sable. Or je tâcherai de te montrer, par des démonstrations géométriques que tu pourras suivre, que, parmi les nombres que j'ai énoncés et exposés dans mes écrits adressés à Zeuxippe, il y en a qui dépassent non seulement le nombre de grains de sable dont le volume serait égal à celui de la Terre remplie de la matière que nous avons indiquée, mais même ayant un volume égal à celui du monde.»

C'est ainsi qu'Archimède (287–212 avant notre ère) commence son ouvrage intitulé *L'Arénaire*. C'est le premier texte de vulgarisation scientifique et, selon moi, c'est aussi l'un des meilleurs jamais écrits. Dans cet article adressé à son roi, Archimède contredit la notion, fréquente à son époque, d'un nombre de grains de sable infini ou trop grand pour être évalué. Cette notion est présente dans la littérature grecque ; par exemple, au V^e siècle avant notre ère, Pindare écrit dans son *Deuxième chant olympique* : «Mais le sable échappe au calcul : les joies aussi que cet homme a données aux autres, qui pourrait en dire le nombre ?» L'idée est également présente dans la Bible, où il y a 21 références à la non-énumérabilité du sable.

Toutefois, si l'humanité a bien accepté les résultats d'Archimède (chacun sait que π est égal à 3,14...), son article le plus accessible, *L'Arénaire*, n'a eu aucun écho à son époque. Par exemple, dans le Nouveau Testament, écrit quelques centaines d'années après la mort d'Archimède, on lit encore que le sable est infini : «Naquirent des descendants comparables par leur nombre aux étoiles du ciel et aux grains de sable sur le rivage de la mer, innombrables.» (*Hébreux 11:12.*) Le Nouveau Testa-

ment ignore les progrès des sciences : on estime facilement que le nombre d'étoiles visibles à l'œil nu est inférieur à 10 000, ce qui était connu des astronomes de l'époque (comme en témoignent le catalogue d'étoiles d'Ératosthène, qui date du II^e siècle avant notre ère, et celui de Ptolémée, au II^e siècle de notre ère).

Dans son article, Archimède démontre que le nombre de grains de sable est fini, en donnant un nombre fini supérieur au nombre de grains de sable dans un univers rempli de sable. D'abord, il se demande ce qu'est «l'Univers» : pour être certain de ne pas être refuté, Archimède choisit le plus grand modèle connu alors, le système héliocentrique d'Aristarque. Cette partie de *L'Arénaire* est physique et expérimentale. Puis Archimède propose un système de numération capable de produire des nombres très grands : c'est la partie arithmétique de l'article.

L'Arénaire montre également que les Grecs n'ont pas toujours été émus d'abstraction, contrairement à ce qu'ont prétendu des historiens : l'ouvrage décrit une expérience astronomique stipulant que la Terre est ronde et tourne autour du Soleil, et Archimède y aborde la question des

nombre très grand, mais fini. Nous examinons ici le système de numération mis au point par Archimède et son application au problème du dénombrement du sable. Nous chercherons aussi pourquoi Archimède s'est arrêté à son «plus grand nombre».

Les nombres d'Archimède

Le dénombrement du sable était très difficile à l'époque d'Archimède, parce que le système de numération grec ne désignait pas les nombres supérieurs à 100 millions et que les Grecs ne disposaient pas d'un système formel, telle l'algèbre, pour écrire et manipuler les grands nombres. Les Grecs comptaient à l'aide de la myriade, égale à 10 000, et leur plus grand nombre était la myriade de myriades. Ce nombre apparaît deux fois dans la Bible.

Pour compter, Archimède commence donc par produire des nombres plus grands que la myriade de myriades. Il nomme «premiers nombres» les nombres compris entre 1 et la myriade de myriades (ce qui crée, même en grec, une confusion avec les nombres premiers 2, 3, 5, ...). Puis il définit les «deuxièmes nombres», dont l'unité est la myriade de myriades, et dont le plus grand est une myriade de myriade d'unités.

Ensuite viennent les «troisièmes nombres» et ainsi de suite jusqu'aux myriades de myriadièmes nombres.

Pour le calcul des grains de sable, ces nombres suffisent largement, mais Archimède veut démontrer qu'il peut aller plus loin et, surtout, plus loin qu'il n'est nécessaire pour les applications physiques : il nomme «nombres de la première période» la suite des nombres construits jusqu'à la myriade de myriadièmes nombres.

Puis il continue, en nommant «unité de la seconde période» le plus grand nombre de la première période, et ainsi jusqu'à la myriade de myriadième période. Il s'arrête au plus grand nombre : une myriade de myriades d'unités des myriades de myriadièmes nombres de la myriade de myriadième période. Comment calcule-t-on le nombre de grains de sable dans un univers rempli de sable avec les hypothèses d'Archimède et son système de numération ?

Pour trouver la réponse, on doit utiliser une méthode de multiplication des «nombres d'Archimède», qu'Archimède donnait ainsi : «Si des nombres sont en proportion à partir de l'unité et que certains de ceux qui sont dans la même proportion sont multipliés entre eux, le produit sera dans la même proportion éloigné du plus grand des

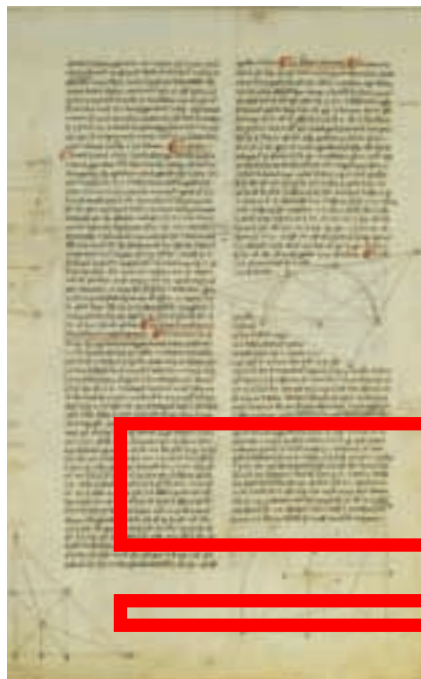
facteurs de autant de nombres dont le plus petit facteur est éloigné, en proportion de l'unité, et il sera éloigné de l'unité de la somme moins un des nombres dont les facteurs sont éloignés à l'unité.»

Un tel énoncé paraît confus ; il montre surtout combien on exprime difficilement les concepts mathématiques quand on ne dispose pas de la notation symbolique de l'algèbre. Traduisons cet énoncé en termes modernes (voir la figure 1).

Soit Ω la myriade de myriades, égale à 100 000 000, ou 10^8 en notation scientifique. Les premiers nombres sont l'ensemble $\{1, \dots, \Omega\}$. Les deuxièmes nombres sont l'ensemble $\{\Omega, \dots, \Omega^2\}$, les troisièmes nombres sont $\{\Omega^2, \Omega^2 + 1, \dots, \Omega^3\}$, et ainsi de suite. Les n -ièmes nombres, par exemple, sont l'ensemble $\{\Omega^{n-1}, \Omega^{n-1} + 1, \dots, \Omega^n\}$. Aussi les myriades de myriadièmes nombres sont l'ensemble $\{\Omega^{\Omega-1}, \Omega^{\Omega-1} + 1, \dots, \Omega^{\Omega}\}$.

Le nombre le plus grand ainsi construit est Ω^{Ω} . Le résultat est général : étant donné un nouveau symbole de numération A , on devrait pouvoir compter jusqu'à A^A : on forme la suite $A_1=A, A_2=A^2, A_3=A^3, \dots$ et le plus grand nombre avec un seul indice est $A_A=A^A$.

Ainsi, notre système de numération moderne pour les grands nombres



1. L'ŒUVRE D'ARCHIMÈDE a été retranscrite en 1270 par William de Moerbeke (à gauche). La traduction, en latin, de cette œuvre en a favorisé la diffusion, en Europe, auprès des lettrés du XIII^e siècle. D'autres traductions (au centre) ont suivi, par exemple celle de Jacobus Cremoniensis, qui date de 1458. Un manuscrit (à droite) a été découvert

en 1899, à Jérusalem. Il s'agit d'un parchemin où était écrit un texte d'Archimède intitulé «La méthode», que l'on avait cru disparu. Le texte d'origine avait été partiellement effacé et un texte religieux réécrit par dessus. Ce palimpseste, qui date du X^e siècle, a été retrouvé et vendu aux enchères, en 1998, à un inconnu, pour deux millions de dollars.

Première période	Deuxième période	...	Ω ⁱ ème période
Premiers nombres 1, ..., Ω	Premiers nombres Π, ..., ΩΠ	Premiers nombres ΠΩ-1, ..., ΩΠΩ-1	
Deuxièmes nombres Ω, ..., Ω ²	Deuxièmes nombres ΩΠ, ..., Ω ² Π	Deuxièmes nombres ΩΠΩ-1, ..., Ω ² ΠΩ-1	
Troisièmes nombres Ω ² , ..., Ω ³	Troisièmes nombres Ω ² Π, ..., Ω ³ Π	Troisièmes nombres Ω ² ΠΩ-1, ..., Ω ³ ΠΩ-1	
...	...	...	
n ⁱ èmes nombres Ω ⁿ⁻¹ , ..., Ω ⁿ	n ⁱ èmes nombres Ω ⁿ⁻¹ Π, ..., Ω ⁿ Π	n ⁱ èmes nombres Ω ⁿ⁻¹ ΠΩ-1, ..., Ω ⁿ ΠΩ-1	
...	...	...	
Ω ⁱ èmes nombres Ω ^{Ω-1} , ..., Ω ^Ω = Π	Ω ⁱ èmes nombres Ω ^{Ω-1} Π, ..., Ω ^Ω Π = Π ²	Ω ⁱ èmes nombres Ω ^{Ω-1} ΠΩ-1, ..., Ω ^Ω ΠΩ-1 = Π ^Ω	

2. LE SYSTÈME DE NUMÉRATION d'Archimède était fondé sur la myriade de myriades, égale à 10⁸ 000 000 et symbolisée ici par Ω.

est fondé sur celui qui a été proposé par Nicolas Chuquet en 1484 et qui utilise le nouveau symbole *illion*. Ainsi $M=10^6$ est un million, M^2 est un billion, M^3 un trillion, et ainsi de suite : un n -illion est M^n , et le plus grand nombre simple de cette forme est un *million-illion*, soit M^M .

Dans la deuxième étape, Archimède introduit les périodes, dont la base est $\Pi = \Omega^2$, le plus grand nombre construit jusque-là. D'après la remarque précédente, il devrait, en principe, continuer jusqu'à Π^2 . Archimède définit les premiers nombres de la seconde période comme l'ensemble $\{\Pi, \dots, \Omega\Pi\}$, les seconds nombres de la seconde période $\{\Omega\Pi, \dots, \Omega^2\Pi\}$, et ainsi de suite. C'est-à-dire que les n -ièmes nombres de la seconde période seront $\{\Omega^{n-1}\Pi, \dots, \Omega^n\Pi\}$. Et les Ω -ièmes nombres de la seconde période seront l'ensemble $\{\Omega^{\Omega-1}\Pi, \dots, \Omega^{\Omega}\Pi = \Pi^2\}$.

Maintenant, Π^2 est pris comme l'unité des nombres de la troisième période, qui seront donc l'ensemble $\{\Pi^2, \dots, \Pi^3\}$. De la même façon, les nombres de la n -ième période seront l'ensemble $\{\Pi^{n-1}, \dots, \Pi^n\}$, et Archimède continue jusqu'à Π^2 , le plus grand nombre nommé dans son système.

En notation moderne, ce nombre est égal à : $\Pi^2 = (\Omega^2)^{\Omega} = \Omega^{2\Omega} = (10^8)^{10^8} = 10^{8 \cdot 10^8}$, soit 10 à la puissance 80 millions de milliards.

Dans ce système, la loi des exposants est un peu compliquée par les choix d'Archimède : elle revient à dire que le produit de l'unité des m -ièmes nombres et de l'unité des n -ièmes

nombres est l'unité des $m+n-1$ -ièmes nombres. Archimède en donne une démonstration grecque, est un exemple simple qui contient toutes les idées de la preuve.

À la recherche de la période perdue

Le nombre Π^2 est souvent cité comme le plus grand nombre considéré dans l'Antiquité, mais personne n'a expliqué pourquoi Archimède s'est arrêté à lui. Vraisemblablement, il aurait dû continuer jusqu'à $\Pi^3 = (\Omega^2)^{\Omega^2} = \Omega^{2\Omega^2}$ qui est bien supérieur à 10 à la puissance 10 à la puissance 10 à la puissance 8, tandis que le «plus grand nombre» d'Archimède est inférieur à 10^{10} . Le mystère me semble découler des particularités du grec ancien et du manque de flexibilité du système d'Archimède : d'abord, soulignons qu'il existe une différence subtile, mais importante, entre les nombres cardinaux, qui indiquent la quantité, et les nombres ordinaux, qui indiquent l'ordre.

Ces concepts de la théorie des ensembles existent depuis longtemps dans toutes les langues. En français, le passage du cardinal à l'ordinal consiste à ajouter *ième* au nom du nombre cardinal. En revanche, l'examen des noms de cardinaux et d'ordinaux grecs ne révèle pas que les Grecs disposaient d'un tel algorithme.

Pour comprendre l'importance de ce concept, revenons au système de numération de Chuquet. Il ne s'agit

plus de la suite «un-illion, deux-illion, ..., n -illion», mais de la suite «premier-illion, deuxième-illion, ..., n -ième-illion» : un nouveau symbole A définit une suite $A_1, A_2, \dots$, où les indices sont ordonnés. Pour définir le dernier nombre de cette suite, on doit pouvoir dire *premier-illionième-illion* (par exemple, premier millionième de million). L'algorithme transforme n'importe quel cardinal en ordinal.

Donc, si l'on admet que les Grecs n'avaient pas d'algorithme pour transformer les cardinaux en ordinaux, on conclut qu'Archimède se limitait au nombre ordinal le plus élevé dans sa langue soit la myriade de myriadième. Autrement dit, le plus grand nombre qu'Archimède aurait pu nommer en utilisant sa langue courante aurait été la myriade de myriadième période, ou Π^2 .

Pour soutenir cette hypothèse, on remarque qu'Archimède n'avait pas conçu son système pour une utilisation générale, ce qui est encore plus intéressant puisque de très grands nombres apparaissent aussi dans son «problème des boeufs du Soleil» qui, selon l'*Odyssee* d'Homère, étaient 350. Il s'agit de classes de taureaux blancs, noirs, jaunes et tachetés, avec des vaches des mêmes variétés, en vérifiant des conditions telles que : «Le nombre total de taureaux blancs et noirs doit être un nombre carré», ou : «Le nombre total de taureaux jaunes et tachetés doit être un nombre triangulaire, de la forme $1+2+3+\dots+m$, où m est un entier positif.» Ce problème se ramène à un système de sept équations à huit variables, et une équation dite de Pell.

À la fin de l'*Arénaire*, Archimède écrit : «Un volume myriade de myriade de myriade de fois multiple du volume du monde.» Pour être cohérent, il aurait dû écrire : «Un volume une unité des troisièmes nombres de fois multiple du volume du monde.» Cet exemple illustre un autre problème linguistique dans l'invention de nouveaux nombres : il faut aussi inventer un adjectif pour chaque nombre (une fois, deux fois...).

On pourrait objecter que les ordinaux se disent aussi «numéro un, numéro deux, etc.», et qu'on lit la suite $A_1, A_2, \dots$ comme «A un, A deux...». De surcroît, la table des cardinaux et des ordinaux grecs montre qu'Archimède aurait pu facilement inventer une façon de transformer un cardinal en ordinal. Cependant, c'est ici qu'intervient la

principale faiblesse de son système : Archimède donne des noms à des ensembles de nombres, et non à des nombres eux-mêmes. Ainsi le nombre $\Pi = \Omega^2$ est l'«unité de la deuxième période», et, même en français moderne, il est difficile d'écrire l'ordinal correspondant.

Un thé au harem d'Archimède

Archimède était très isolé, à Syracuse, ce qui explique son originalité : son invention de grands nombres n'a rien en commun avec les mathématiques de son temps. Son isolement est probablement responsable des petites incohérences de son texte : s'il avait pu discuter de son système avec un autre mathématicien, Archimède aurait probablement amélioré ses résultats. Cherchons donc les améliorations qu'Archimède – sans trop le vexer – aurait pu apporter à son «théorème».

Tout d'abord, il aurait dû donner des noms à ses grands nombres. Par exemple, s'il avait nommé Ω^2 la période, alors Π^2 aurait été la deuxième période, qu'Archimède aurait pu nommer, en extrapolant les ordinaux grecs existants. Ainsi le grand nombre Π^2 aurait été atteint.

Cependant, Archimède aurait dû commencer par donner des noms directement à tous ses nouveaux nombres. Une méthode directe aurait consisté à nommer *la deuxième myriade* la myriade de myriades, *la troisième myriade* aurait été la myriade de deuxièmes myriades ; et ainsi de suite jusqu'à *la myriadième myriade*. Cette notation pour les petits cas a été utilisée par Diophante. Elle simplifie la lecture des exposants : le produit d'une m -ième myriade et d'une n -ième myriade est une $m+n$ -ième myriade.

Cependant, si le but est seulement de nommer de très grands nombres, et non de donner un système complet de numération, on peut aller beaucoup plus loin. On commence par nommer la myriadième myriade une *grande myriade*. On continue avec la grande myriadième grande myriade, qu'on nomme la *très grande myriade*. La très grande myriadième très grande myriade s'appelle *l'énorme myriade*.

On utilise de nouveau la nomenclature d'Archimède en nommant la grande myriade la *période*, la très grande myriade la *deuxième période*, l'énorme période la *troisième période*. Le procédé conduit à la *périodième*

Le nombre de grains de sable dans l'Univers

Les hypothèses d'Archimède

- Au plus 10 000 grains de sable dans un grain de pavot (sphérique).
- Au plus 40 diamètres de grains de pavot sont contenus dans le diamètre d'un doigt.
- Au plus 10 000 diamètres d'un doigt – au plus – sont dans un stade (une mesure ancienne de distance de l'ordre de 200 mètres).
- La Terre est une sphère dont le périmètre est inférieur à trois millions de stades.
- La distance entre le centre de la Terre et le centre du Soleil est inférieure à 10 000 diamètres de la Terre.
- L'Univers est une sphère, le Soleil est au centre, et la Terre tourne autour du Soleil selon un cercle, le rapport du diamètre de l'Univers au diamètre de l'orbite de la Terre autour du Soleil est inférieur au rapport entre le diamètre de l'orbite de la Terre et son diamètre.
- Le périmètre d'un cercle est supérieur au triple de son diamètre.
- Le volume d'une sphère est proportionnel au cube de son diamètre.

D'après les hypothèses d'Archimède, pourriez-vous trouver une borne supérieure pour le nombre de grains de sable dans un Univers rempli de sable ?

La réponse

1. Soit D_u le diamètre de l'Univers, D_o le diamètre de l'orbite de la Terre autour du Soleil, D_t le diamètre de la Terre.

La dernière hypothèse physique s'écrit : $D_u/D_o < D_o/D_t$.

L'avant-dernière hypothèse affirme que D_o/D_t est inférieur à 10 000, donc $D_u < 10^4 D_o$.

En outre, une fois, l'hypothèse $D_o/D_t < 10^4$ donne : $D_u < 10^4 \cdot 10^4 \cdot D_t = 10^8 D_t$.

L'hypothèse sur le périmètre de la Terre donne :

$D_t < 10^6$ stades, en raison de l'avant-dernière hypothèse

Donc : $D_u < 10^8 \cdot 10^6$ stades = 10^{14} stades.

Puisqu'il y a moins de 10 000 diamètres de doigts dans un stade, on a :

$D_u < 10^4 \cdot 10^{14} D_p = 10^{18} D_p$, où D_p est le diamètre d'un doigt.

Puisqu'il y a moins de 40 diamètres de grains de pavot dans un doigt :

$D_u < 40 \cdot 10^{18} D_g = 4 \cdot 10^{19} D_g$, où D_g est le diamètre d'un grain de pavot.

D'après la dernière hypothèse, le volume V_u de l'Univers satisfait l'équation :

$V_u < (4 \cdot 10^{19})^3 V_p = 64 \cdot 10^{57} V_p < 10^{59} V_p$, où V_p est le volume d'un grain de pavot..

D'après la première hypothèse, un grain de pavot contient au plus 10^4 grains de sable, donc le nombre de grains de sable dans l'Univers est inférieur à $10^{59} \cdot 10^4$, soit 10^{63} .

Notons que le diamètre maximal de l'Univers, d'après Archimède, est égal à 10^{14} stades.

Sachant qu'un stade est environ égal à 200 mètres, on obtient ainsi un rayon de 10^{12} kilomètres environ, soit une année-lumière.

période. Ces périodes croissent beaucoup plus vite que les périodes définies par Archimède, mais ces nouvelles périodes ne sont plus capables d'exprimer tous les nombres.

En notation moderne, une myriade correspond à $\mu = 10^4$, et une grande myriade à $\mu^2 = 10^{4 \cdot 10^4}$: c'est la période π_1 . Soit π_2 la deuxième période. En utilisant la formule $(10^a)^b = 10^{ab}$, on trouve : $\pi_2 = \pi_1 + 1 = 10^{10^{4 \cdot 10^4} + 1}$, valeur supé-

rieure à 10 à la puissance 10 à la puissance $10\,000$. De même, la troisième période serait supérieure à 10 à la puissance 10 à la puissance 10 à la puissance 10 .

En général, la n -ième période est d'un ordre de grandeur bien supérieur aux nombres conçus par Archimède. Ainsi, on peut calculer que la n -ième période a une valeur supérieure à celle d'une tour de puissances de 10 à $40\,000$ étages.

Thierry VARDI est mathématicien à l'Institut des hautes études scientifiques, IHES, à Bures-sur-Yvette.

<http://www.ihes.fr/~ilan/archimedes.html>

Archimède, *Ceuvres*, texte établi et traduit par C. Mugler, Les Belles Lettres,

Paris, 1970-1971.

I. VARDI, *Archimedes' Cattle Problem*, in *American Mathematical Monthly*, vol. 105, pp. 305-319, 1998.

Le thé au harem d'Archimède, film de Mehdi Charef, 1985.

L'infini en Chine

JEAN-CLAUDE MARTZLOFF

Le calcul numérique a joué un rôle clé dans la perception que la Chine a eu de l'infini.



Wang Anjie, *Jiezi yuan* (Manuel de peinture du jardin de grains de moutarde)

En Chine, l'infini et l'indéfiniment grand ou l'indéfiniment petit ont surtout été abordés à propos de problèmes de dialectique, de mathématiques, de cosmologie ou de synchronisation des cycles planétaires. À chaque fois, ces problèmes ont été traités pour eux-mêmes. Bien qu'il existe ailleurs, ce mode de perception de l'infini domine en Chine. Il s'oppose à celui des autres grandes civilisations de l'Antiquité et du Moyen Âge, qui ont plutôt abordé l'infini au travers de la métaphysique.

Au III^e siècle avant notre ère, les logiciens chinois élaborent des paradoxes qui rappellent les idées des sophistes grecs de la même époque. En Chine, ils sont les premiers à s'intéresser à la notion d'infini. Puis, au premier millénaire de notre ère, les traductions chinoises de centaines de textes indiens de théologie bouddhique propagent l'idée d'un monde infini, sans origine ni fin, qui évolue au rythme du cycle des réincarnations, répétées à l'infini. Les Chinois inventent des systèmes de notation pour représenter des nombres très grands ou très petits, correspondant aux intervalles de temps qui interviennent dans ce cycle.

Parallèlement, les mathématiciens chinois calculent la valeur approchée de π , établissent les formules de l'aire du cercle, du volume de la sphère et de la pyramide. Ils utilisent des raisonnements similaires à ceux, antérieurs, des mathématiciens grecs. On ignore si ces raisonnements sont issus exclusivement de la pensée chinoise ou empruntés à la civilisation hellénistique, dont les idées auraient atteint la Chine par la route de la soie, soit directement, soit indirectement, en passant d'abord par l'Inde.

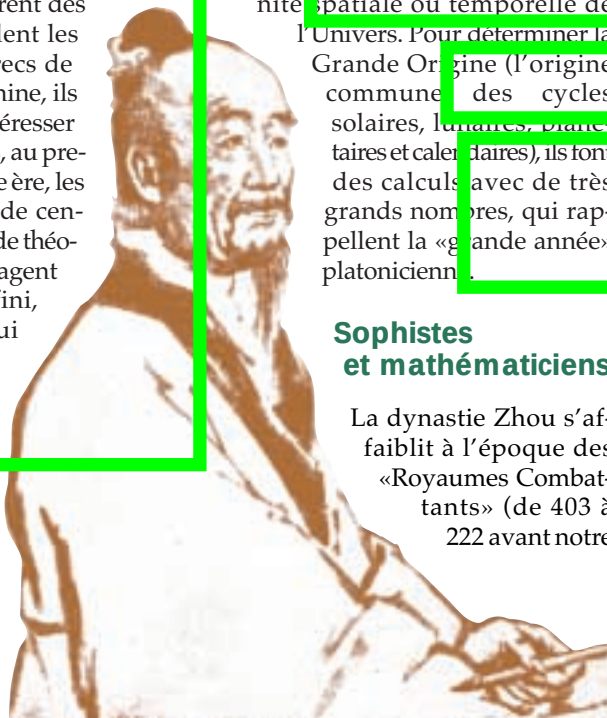
Les cosmologistes, les astrologues et les astronomes s'interrogent sur l'infinité spatiale ou temporelle de l'Univers. Pour déterminer la Grande Origine (l'origine commune des cycles solaires, lunaires, planétaires et calendaires), ils font des calculs avec de très grands nombres, qui rappellent la « grande année » platonicienne.

Sophistes et mathématiciens

La dynastie Zhou s'affaiblit à l'époque des « Royaumes Combattants » (de 403 à 222 avant notre

ère), marquée par des guerres incessantes. Pendant cette période troublée, des penseurs s'intéressent à la logique et à la dialectique. Des rhétoriciens, conseillers des princes, essaient de trouver les limites du langage à partir de paradoxes dont quelques-uns mettent en jeu l'infini, ou plus exactement l'illimité, le non-borné. À la fin du IV^e siècle avant notre ère, Hui Shi, Premier ministre du roi Hui de Wei, rhétoricien et auteur d'un code de lois aujourd'hui perdu, illustre l'infinie divisibilité spatiale d'un segment de droite par dichotomie : « Il restera toujours quelque chose d'un bâton d'un pied [de longueur] dont on prend chaque jour la moitié, même au bout de dix mille générations. » Son raisonnement est similaire à celui du sophiste grec Zénon d'Elée, selon lequel une flèche ne parvient jamais à son but, car elle doit d'abord parcourir la moitié de la distance qui en sépare, puis la moitié de la moitié et ainsi de suite. Hui Shi affirme même qu'il est à la fois impossible et possible d'accumuler l'infiniment petit : « Ce qui n'a pas de dimension ne peut pas être accumulé,

1. PORTRAIT IMAGINAIRE du mathématicien Liu Hui, peint par Jiang Zhaohe, à Pékin, en 1980. Même s'il ne s'appuie sur aucun document historique, ce portrait donne un visage à un « héros national » : le père des mathématiques chinoises. Les Chinois se sont intéressés à l'infini par l'intermédiaire des grands nombres, tel le nombre de configurations possibles d'un échiquier de *weiqi*, ce dernier mot signifiant échecs par encerclement (un jeu plus connu sous son nom japonais de *go*, ci-dessus).



mais finit pourtant par atteindre un millier de li [le li est une unité de mesure chinoise]. » Toutefois, il n'explique pas ce paradoxe. Les sinologues ont envisagé plusieurs contextes auxquels il pourrait se rapporter : un débat d'atomistes et de non-atomistes sur l'infinie divisibilité des objets du monde matériel, un débat de géomètres sur la dimension du point ou l'épaisseur du plan, une démonstration des limites du langage au moyen de phrases absurdes. Hui Shi en reprend peut-être une sorte de « déconstruction » du langage en juxtaposant des assertions dépourvues de sens, sans rapport les unes avec les autres. Ses phrases à propos de l'infini, rassemblées par Zhuangzi, le maître chinois de l'irrationnel, appuient cette dernière hypothèse : certaines sont effectivement absurdes et, de surcroît, sans rapport ni avec l'infini ni avec l'atomisme (par exemple, « la tortue est plus longue que le serpent » ou encore « un chien blanc est noir »).

À la fin du III^e siècle, l'infini apparaît pour la première fois en Chine, dans sa dimension mathématique, dans un texte de Liu Hui, qui commente les *Neuf Chapitres de procédés calculatoires*. Ce dernier ouvrage contient la synthèse des connaissances mathématiques jusqu'à l'époque des Han antérieurs (de 220 à 206 avant notre ère), la dynastie sous laquelle ils auraient été compilés pour la première fois.

Le mathématicien Liu Hui remarque que la racine carrée de certains nombres ne peut pas se calculer exactement et explique que l'opération se poursuit parfois indéfiniment. Il calcule cette racine chiffre par chiffre et parle des décimales du résultat en ces termes : « Les petits chiffres qui n'ont pas de nom [qui ne sont pas associés à un nom d'unité de mesure] correspondent à des numérateurs de fractions dont les dénominateurs valent successivement 10, 100, 1 000... » Ainsi, il aborde la notion d'infini par l'intermédiaire du calcul et du développement décimal illimité, qu'il a intuitivement découvert.

Liu Hui s'intéresse aussi à des processus algorithmiques susceptibles de se prolonger sans que l'on sache par avance s'ils vont s'arrêter. Sa méthode de calcul de l'aire du cercle et de la valeur de π en est un exemple. Elle consiste à encadrer le cercle entre deux suites illimitées de polygones et de roues dentées à $2n$ côtés, à partir de $n = 3$ (voir la figure 2). Liu Hui affirme qu'à la longue « le polygone se confond



2. L'AIRe DU CERCLE. Au III^e siècle, le mathématicien chinois Liu Hui, dans son commentaire des *Neuf Chapitres de procédés calculatoires* (reproduit ci-dessus), calcule l'aire du cercle en l'encadrant entre deux séries indéfinies de polygones réguliers inscrits (tel le polygone en bleu) et de « roues dentées » (telle la roue en vert). Si P_n = aire du polygone régulier inscrit à n côtés et R_n = aire de la roue dentée à n côtés, alors $P_n <$ aire du cercle $< R_n = P_n + 2(P_{2n} - P_n)$. Quand n devient assez grand, étant donné les approximations que fait Liu Hui, R_n devient numériquement égal à P_n .

avec le cercle», sans préciser d'avantage sa pensée. Il veut probablement dire qu'en menant les calculs assez loin il obtient une valeur numérique de l'aire du polygone égale à celle de la roue dentée correspondante. Selon le nombre de chiffres avec lesquels les opérations intermédiaires sont effectuées, la valeur de π obtenue est plus ou moins précise : elle est de 157/50 (3,14) d'après les calculs de Liu Hui et de 355/113 (dont le développement décimal contient les six premiers chiffres de π) d'après ceux de Zu Xuan, fils du mathématicien Zu Chongzhi (429-500), qui utilise la même technique de calcul. Liu Hui met le doigt sur un processus infini, ou plutôt indéfini (dont les limites ne sont pas connues à l'avance).

Nombres immenses et nombres infimes

Pour calculer l'aire du cercle, les démarches chinoise (celle de Liu Hui) et grecque (celle d'Archimède, voir *Les mathématiques, science de l'infini*, par Javier de Lorenzo Martínez, page 29) sont toutes les deux fondées sur des polygones dont le nombre de côtés double sans cesse, mais leur esprit s'oppose : Archimède s'appuie sur des pos-

tulats, les axiomes, des théorèmes et un raisonnement par l'absurde, tandis que Liu Hui se fonde exclusivement sur des calculs algorithmiques, justifiés de façon intuitive et visuelle.

Pour calculer des volumes, l'opposition est moins flagrante. Liu Hui établit le volume de la pyramide en la disséquant en une infinité d'éléments géométriques de plus en plus petits (voir la figure 3), à la manière d'Euclide dans le cinquième théorème du livre XII des *Éléments*, sans que sa technique dépende de calculs réitérés à précision variable, comme dans le cas de l'extraction de la racine carrée et du calcul de π . De même, dans le cas de la sphère, Liu Hui parvient au résultat un peu comme Archimède dans la *Méthode* d'Érathosthène, en s'appuyant sur un cas particulier du principe de Cavalieri pour déduire les volumes à partir des surfaces, la version chinoise de ce principe affirme que, si deux volumes de la même hauteur sont tels

que le rapport des surfaces de leurs sections reste constant, quel que soit le niveau auquel on effectue la coupe, alors ils sont égaux. Le volume est envisagé comme l'accumulation d'une infinité de tranches infiniment minces, ce qui rappelle le paradoxe de Hui Shi sur l'impossibilité d'accumuler ce qui est dépourvu de dimension.

Pour appréhender les nombres indéfiniment grands ou indéfiniment petits, les Chinois inventent des séries de termes inspirés des textes sanscrits indiens. Analogues à nos millions, milliards et billions, ces séries servent à nommer des puissances de dix de plus en plus grandes, mais elles se prolongent beaucoup plus loin. Les plus courantes commencent par 11 termes qui n'ont pas de signification particulière, mais se prolongent par *henghesha*, *buke-siyi*, *wuliangshu*, des expressions imagées qui signifient respectivement le «nombre de grains de sable du Gange», l'«impensable» et le «non-mesurable». Ces termes sont investis d'une connotation théologique éloignée des courants de pensée majoritaires dans le monde chinois (surtout du confucianisme), davantage préoccupés par des problèmes d'organisation politique de la société que de métaphysique.

Selon la façon d'attribuer une valeur à un terme donné de la série en fonction du terme précédent, les Chinois obtiennent trois séries dites «du degré inférieur», «du degré moyen» et «du degré supérieur» (les termes supérieurs à la myriade valent respectivement dix fois le terme précédent, 10^4 fois ce terme ou son carré).

Les Chinois ont aussi considéré des puissances de dix décroissantes jusqu'au *chun*, un mot d'origine sanscrite qui signifie «un instant», c'est-à-dire la plus petite unité de temps envisageable.

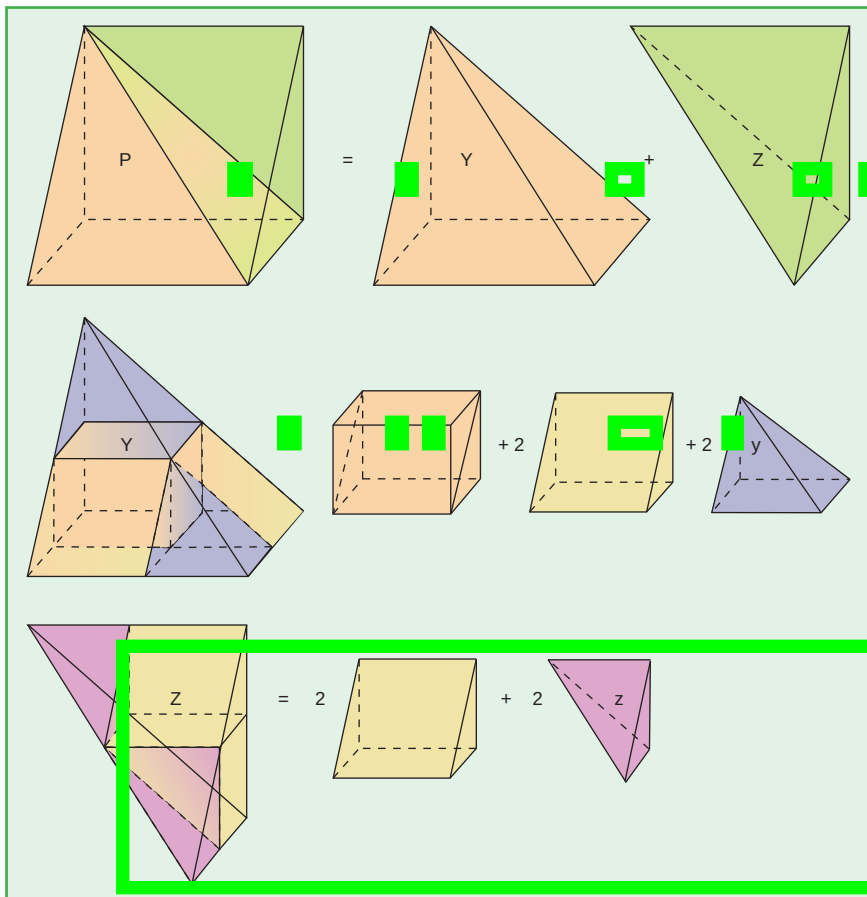
L'intérêt de ces représentations est purement théorique et spéculatif : aucune situation pratique ne nécessite de recourir à des nombres d'une taille aussi inhabituelle. Aussi l'on ne s'étonnera pas de constater que la représentation d'un nombre à l'aide des séries précédentes n'est pas immédiate, dès lors que la règle qui régit la succession

des puissances les unes à partir des autres est plus complexe qu'une simple multiplication par dix : cela n'a pas d'importance si l'on ne fait que réfléchir abstraitement sur l'existence de très grands ou très petits nombres. En revanche, les écrire explicitement est difficile. Dans un recueil de notes, le polygraphe Shen Gua (131-1095) essaie d'exprimer par écrit la solution du problème combinatoire suivant : quel est le nombre de configurations possibles d'un échiquier de *weiqi* ? Le jeu, plus connu sous son nom japonais, le *go*, se joue sur les intersections d'un damier de 361 cases, chaque intersection pouvant recevoir soit un pion noir, soit un pion blanc, soit rien. Shen Gua ne parvient pas à effectuer correctement les calculs et à trouver la réponse exacte, soit $3^{(19 \times 19)} = 174\,089\,172\,603$, qui se compose de 173 chiffres. En revanche, il imagine à cette occasion un système de représentation approchée des grands nombres, fondé sur la multiplication répétée du nombre 10 000 par lui-même.

Le problème de Shen Gua en rappelle un autre, célèbre en Europe depuis le Moyen Âge, où l'on cherche à déterminer le nombre total de grains de blé posés sur les cases d'un échiquier, le nombre de grains posés sur chaque case doublant à chaque nouvelle case remplie. La solution est égale à $2^{64} - 1$, un nombre très grand, mais encore petit comparé à celui de Shen Gua. Dans ces problèmes, c'est encore le calcul qui conduit à envisager l'infiniment grand et l'infiniment petit, plutôt que la métaphysique.

Les cycles astronomiques et la Grande Origine

Sous la dynastie des Han postérieurs (de 25 à 220), l'astronome Qi Meng est le premier à postuler que l'Univers est illimité dans le temps et dans l'espace. Pourtant, on ne trouve aucune réflexion substantielle sur l'infini dans la littérature scientifique ou philosophique chinoise postérieure aux Han. Au contraire, les Chinois semblent indifférents à la cosmologie. Pourtant, ils s'intéressent à l'astronomie pour déterminer la date de la conjonction du Soleil, de la Lune et des cinq planètes (Mercure, Vénus, Mars, Jupiter et Saturne). Cette conjonction bénéfique, en vertu des principes de l'astrologie chinoise, servait d'origine au décompte du temps dans le calendrier chinois. On la



3. LE VOLUME DE LA PYRAMIDE. Pour faire apparaître la pyramide Y, Liu Hui décompose le prisme P de sorte que $P = Y + Z$. Comme il connaît le volume du prisme $V_p = 1/2 \text{ largeur} \times \text{longueur} \times \text{hauteur}$, il lui suffit de montrer que $Y = 2Z$ pour prouver que le volume de la pyramide Y vaut $1/3 V_p$. Pour y parvenir, il dissèque Y et Z. Il constate ainsi que les divers éléments de la dissection sont dans la proportion du simple au double, mais avec un petit «reliquat», y et z, dont les éléments sont semblables à Y et Z, avec des dimensions divisées par deux, et pour lequel il faut vérifier à nouveau $y = 2z$. Pour y parvenir, il réitère la même décomposition indéfiniment, de façon à obtenir à chaque fois des reliquats de plus en plus petits. À la limite, $Y = 2Z$. La technique est intuitive, mais elle peut être présentée sous forme rigoureuse.

nomme la «Grande Origine». Elle était censée survenir à minuit, le jour du solstice d'hiver. À partir de cette date, les jours et les années étaient numérotés à l'aide du cycle de 60 (un nouveau cycle commence tous les 60 jours ou tous les 60 ans). Puis les Chinois ont voulu que la Grande Origine coïncide aussi avec le début du mois anomalistique (l'intervalle de temps qui sépare deux passages de la Lune à son périgée) et draconitique (l'intervalle de temps qui sépare deux passages de la Lune à son nœud ascendant, l'une des intersections de l'orbite lunaire avec l'écliptique terrestre). Dans le cas le plus complet, les Chinois ont calculé la date de la Grande Origine comme étant le point de départ de 11 cycles (cinq pour les planètes, un pour le Soleil, trois pour la Lune et deux pour le calendrier).

Ces préoccupations ont à nouveau fourni aux Chinois des raisons de s'intéresser aux très grands nombres : pour déterminer la date de la Grande Origine, ils ont calculé le plus petit multiple commun de certains ou de la totalité des cycles précédents. Les périodes correspondantes étant mutuellement incommensurables, ils ont reculé la Grande Origine de plus en plus loin dans le passé, il y a des milliers, puis des millions et enfin des centaines de millions d'années, dans le cas le plus complet. Avec la réforme de l'astronomie chinoise, promulguée en 1281, l'idée des cycles a été abandonnée et, dès lors, l'astronomie chinoise a cessé de taquiner les très grands nombres.

En dehors des exemples précédents, les quelques autres infinis de la culture chinoise ont presque tous émergé de processus calculatoires. Le plus frappant d'entre eux concerne les séries infinies de fonctions trigonométriques, introduites en Chine par les ouvrages mathématiques européens que les missionnaires jésuites ont apporté aux XVII^e et XVIII^e siècles. Les mathématiciens chinois du XVIII^e et du XIX^e siècle ont abordé le sujet à partir de calculs algébriques, sans lésiner sur la taille des nombres, qui atteignait parfois plus de 50 chiffres.

Si les sophistes et les logiciens de l'Antiquité avaient bénéficié d'une plus grande audience, si leurs théories n'avaient pas été associées aux charmes supposés fallacieux de la rhétorique, la notion d'infini en Chine s'en serait trouvée modifiée. L'importance croissante au cours du temps du confucianisme



4. LES MILLE BOUDDHAS de l'une des grottes du site de Dunhuang (au Nord-Ouest de la Chine) illustrent l'idée d'un monde infini, véhiculée par le bouddhisme en Chine à partir du premier millénaire de notre ère. Cette infinitude est symboliquement représentée par les très grands nombres (ici, par les mille bouddhas).

dans le monde chinois, ont été particulièrement influents au XI^e siècle, a probablement joué un rôle majeur à cet égard : dans ses *Entretiens familiaux*, Confucius explique qu'il serait vain de s'intéresser aux dieux, aux esprits et plus généralement à la métaphysique et à tout ce qui paraît fonciè-

rement irréalisable. Bien que la réflexion chinoise ne manque pas d'envolées spéculatives, ceux qui s'y sont essayés n'ont guère été suivis, et la question de l'infini n'a jamais retenu très longtemps l'attention des penseurs chinois, sauf des bouddhistes.

Jean-Claude MARTZLOFF est sinologue au Centre de recherche en civilisation chinoise du CNRS, à Paris.

Jacques GERNET, *Le monde chinois*, Armand Colin, Paris, 1999.

Jean-Claude MARTZLOFF, *A History of Chinese Mathematics*, Springer Verlag, Berlin, Heidelberg, New York, 1997.

Jean-Claude MARTZLOFF, *Chinese Mathematical Astronomy*, in *Mathematics Across Cultures, The History of Non-Western Mathematics*, sous la direction d'He-

laine S. pp. 373-407, Kluwer Academic Publishers, Dordrecht, 2000.

Joël P. N. *Notation et optimisation du calcul des grands nombres en Chine*, in *Nombres, astres, plantes et viscères*, sous la direction d'Isabelle Ang et Pierre-Étienne Will, Collège de France, Institut des hautes études chinoises, pp. 89-111, Paris 1994.

Kiyosi YABUUTI, *Une histoire des mathématiques chinoises*, Belin-Pour La Science, 2000.

Thābit ibn Qurra et l'infini numérique

TONY LÉVY

Face aux thèses d'Aristote, longtemps dominantes en philosophie, un savant arabe du IX^e siècle eut sur l'infini un point de vue aussi original qu'audacieux.



Jusqu'à la Renaissance, la conception d'Aristote a dominé tous les débats concernant l'infini. Commentée, précisée, expliquée, amendée par les uns, elle fut, plus rarement, récusée par d'autres. Sa portée est tout à la fois physique, cosmologique, mathématique, philosophique et théologique. Cet aspect de l'histoire de la notion d'infini est primordial : c'est dans l'unité de ces divers registres du savoir, dans les tensions qui les relient ou les opposent, que l'on saisit la dynamique du questionnement sur l'infini. Il nous faut donc retracer les grandes lignes de la théorie aristotélicienne de l'infini et indiquer la nature des premières oppositions qu'elle a suscitées : on mesurera ainsi la nouveauté et la hardiesse du point de vue développé sur cette question par Thābit ibn Qurra de Harrān (836-901).

L'infini est inachèvement

La pensée d'Aristote s'enracine dans un questionnement sur la structure du monde. La science de la nature – la « physique » pour Aristote – et l'analyse du mouvement qui en est le socle conduisent à affronter la question de l'infini : l'infini existe-t-il ? Si oui, en quel sens ? Le mouvement, une grandeur physique (une longueur, par exemple) ou le temps sont autant de notions soumises à l'alternative : être limité ou échapper à la limite. Puisque « la nature est principe de mouvement »,

que le mouvement relève du continu et que la divisibilité du continu est « sans terme » (c'est-à-dire sans fin), la question de l'infini est posée. Retenons ce point essentiel, qui oppose Aristote aux écoles de pensée atomistes : la division d'une grandeur physique est sans terme ; il faut donc rendre compte de cette propriété qui caractérise la divisibilité d'une grandeur telle qu'une ligne, une surface ou un solide.

Dans une première description, qui n'est pas encore une définition, on dira avec Aristote : « est infini (*apeiron*) » ce qu'on peut « par nature parcourir, mais qui ne se laisse pas parcourir et n'a pas de fin (*peras*) ». L'idée de « parcourir » de « traversée », et la métaphore spatiale qu'elle charrie, nous pousse à camper l'analyse sur le territoire du quantifiable. C'est bien dans la catégorie de la quantité, du « combien », qu'Aristote entend inscrire sa définition de l'infini. Qui dit quantité dit grandeur ou nombre et, pour mesurer ou compter, il faut pouvoir distinguer le tout de ses parties, et poser ainsi, d'emblée, que le tout en question est « partageable », c'est-à-dire « divisible ». L'infini concerne trois notions. La grandeur, qui est « divisible à l'infini » ; le nombre, qui est « composable à l'infini » (aussi grand que soit un nombre, il en est un qui est plus grand) ; le temps, qui est « composable et divisible ».

La première conclusion d'Aristote est négative : l'infini n'existe pas comme totalité, comme « être actuel »,

comme forme achevée. Cette impossibilité vaut aussi bien pour une totalité corporelle que pour une totalité incorporelle, substance ou principe. En particulier, le monde en tant que grandeur physique est fini, clos par la « dernière » des sphères célestes.

Toutefois, pour Aristote, l'infini existe selon un autre mode : en puissance. Il s'agit d'un mode d'existence inférieur, subordonné à l'être actuel, mais qui n'en est pas moins réel. L'infini en puissance, qui est possibilité, virtualité s'exprime dans cette propriété de la grandeur continue : aussi loin que je la divise, je peux encore poursuivre la division. Aristote écrit ainsi dans sa *Physique* : « Mais l'infini se trouve être le contraire de ce que l'on dit. En effet, il n'est pas ce qui n'a rien à l'extérieur de soi, mais c'est ce dont quelque chose est toujours à l'extérieur de lui : c'est cela, l'infini [...]. Est donc infini ce dont, quand on le prend selon la quantité, il est toujours possible de prendre quelque chose à l'extérieur. »

L'infini n'est qu'un « en puissance »

Cette divisibilité sans terme du continu se double d'une autre propriété. On peut la formuler ainsi : divisez un segment de droite en deux, puis un des deux sous-segments ainsi obtenus (dont la longueur vaut alors la moitié de celle du segment initial) à son tour en deux, et ainsi de suite ; cette division ou dichotomie

tomie n'a pas de terme, comme on l'a dit. Si, de plus, vous réunissez les sous-segments successivement produits par cette dichotomie, vous vous «rapprochez» du segment initial ; et plus le nombre des sous-segments est grand, plus vous vous rapprochez du segment fini initial. En s'écartant quelque peu de la lettre du texte d'Aristote, mais non de son esprit on pourrait dire que l'infini ainsi conçu est un détour nous permettant de retrouver le fini (le segment initial), sans jamais toutefois épuiser le fini. Cela fait bien sûr songer à ce que nous nommons aujourd'hui, la «convergence» d'une série géométrique : plus l'entier n est choisi grand dans la somme $S_n = 1/2 + 1/4 + 1/8 + \dots + 1/2^n$, plus le résultat se rapproche de 1 ; on dit que la somme S_n converge vers 1, ou que la limite de S_n est 1. On se gardera toutefois d'attribuer à Aristote l'idée moderne de «limite», qui n'a rien à voir avec son propos.

S'agissant du nombre, il sera dit «augmentable à l'infini en puissance». Il suffit de penser au nombre de sous-segments produits par la division répétée d'un segment de droite : quel que soit le nombre (de sous-segments) considéré on peut encore l'augmenter. Et de même que la division du segment ne s'achève jamais, il ne saurait y avoir un nombre ultime, un nombre «infini actuel».

Le troisième volet de l'analyse aristotélicienne concerne le temps, défini par le mouvement, et par ce biais, associé à la grandeur. Pour Aristote, le temps est à la fois divisible et augmentable. Le mouvement des sphères célestes est «sans commencement et sans fin», ce qui l'amène à faire apparaître l'infini «dans le temps et la génération des hommes». Ce thème, communément désigné comme celui de l'«éternité du monde» (le monde a toujours été là!), suscitera de nombreuses objections, surtout chez les partisans des religions révélées, soucieux de fonder philosophiquement l'idée de la création du monde par un acte divin. Soulignons que les premiers débats suscités par la théorie d'Aristote eurent un point de départ théologique.

Jean Philopon, un penseur chrétien d'Alexandrie (VII^e siècle), est à l'origine de plusieurs arguments dirigés contre l'idée de l'éternité du monde. Ces arguments, qui utilisent les définitions mêmes d'Aristote relatives à l'infini, visaient à ruiner les conclusions de ce dernier concernant l'infinité du temps,

pour en arriver à justifier l'idée d'un monde créé par la volonté souveraine de Dieu. C'est ainsi que Philopon met en défaut la théorie aristotélicienne de l'infini, suscitant ainsi chez les partisans d'Aristote, surtout au Moyen Âge, amendements et développements nouveaux.

L'infini aristotélicien confronté à ses paradoxes

La principale critique de Philopon, au demeurant pertinente, exploitait le paradoxe suivant : l'éternité temporelle affirmée par Aristote ne réintroduit-elle pas par la fenêtre l'infini actuel qu'il prétend chasser par la porte ? À cette occasion, Philopon recourait fréquemment à l'un des arguments développés par Aristote pour récuser la possibilité d'un infini actuel. Précisons pourquoi il revêt

une portée considérable pour l'histoire qui nous concerne ici.

Supposons, le temps de la réflexion, qu'existe réellement une substance (voire un principe, comme le pensaient par exemple les pythagoriciens) dont la seule essence est d'être un infini. Pour Aristote, la définition d'un être épuise la réalité ; autrement dit, la définition et la réalité d'une telle substance sont une seule et même chose. Si l'infini est une substance divisible, comment définir ses parties ? On répondra : de même que toute partie d'une substance comme l'air ne peut être définie que comme air, de même une partie de l'infini-substance ne peut être définie, à son tour, que comme de l'infini. Il en résulte que l'infini serait divisible en infinis. Arrivé à ce stade, Aristote conclut laconiquement : «Mais il est impossible que la même chose soit plusieurs infinis.»



1. DES SAGES AU TRAVAIL. Cette miniature arabe, qui date du XIII^e siècle, est l'œuvre du miniaturiste Al-Wasīṭī et fait partie des illustrations d'un classique de la littérature islamique, *Les séances d'Al-Hariri*.

Comment réfuter l'argument «un infini ne peut être plus grand qu'un infini»

«Nous avons aussi questionné au sujet de la proposition invoquée par nombre de commentateurs respectables, à savoir qu'un infini ne peut être plus grand qu'un infini. Il [Thābit] nous a montré la fausseté de cette proposition, encore une fois sur l'exemple du nombre.

En effet, le nombre en soi est infini, et dans le nombre, les nombres pairs, pris séparément, sont finis ; et il en est de même pour les nombres impairs ; or ces deux catégories sont égales, et chacune d'elles est la moitié du nombre dans sa totalité. Leur égalité [la catégorie des nombres pairs et celle des impairs] se montre à partir de ce que, de deux nombres successifs, l'un est pair et l'autre est impair. Que le nombre [dans sa totalité] soit le double de chacune d'elles [la catégorie des nombres pairs et celle des impairs], cela vient de ce qu'elles sont égales et qu'elles l'épuisent [le nombre] et qu'il ne contient aucune partie en dehors d'elles. Chacune d'elles est donc bien la moitié du nombre [dans sa totalité].

Il est clair aussi qu'un infini peut être le tiers d'un infini, ou le quart ou le cinquième ou toute autre partie du nombre considéré en lui-même. En effet, les nombres qui ont un tiers [les entiers multiples de 3] sont infinis et ils représentent le tiers du nombre dans sa totalité. Et les nombres qui ont un quart [les entiers multiples de 4] représentent le quart du nombre dans sa totalité ; et les nombres qui ont un cinquième représentent le cinquième du nombre dans sa totalité. Et l'on procède de même pour les autres parties du nombre. Cela vient de ce que nous trouvons dans tout triplet de nombres consécutifs, l'un qui a un tiers [un entier multiple de 3] ; dans tout quadruplet de nombres consécutifs, l'un qui a un quart ; dans tout quintuplet, l'un qui a un cinquième ; et dans toute suite de nombres consécutifs, quel qu'en soit le nombre, on trouve que l'un d'entre eux a une partie désignée d'après le nombrement [dans toute suite de n nombres consécutifs, il y a un nombre qui a une n -ième partie, c'est-à-dire un entier multiple de n]...»

Bien des auteurs ont compris ainsi cet argument d'impossibilité : si un tout est infini, et qu'une partie de ce tout est aussi infinie, alors, le tout étant nécessairement plus grand que sa partie, nous aurions un infini plus grand qu'un infini ; voilà qui est absurde. Rapportée au cadre conceptuel fixé par Aristote, cette «absurdité» signifie que l'infini n'est pas une entité, que c'est un processus d'inachèvement ; à ce titre, il n'y a aucun sens à le traiter comme une quantité, à distinguer différents infinis, à comparer des infinis, bref à le manipuler comme une forme.

Quoi qu'il en soit, tous ceux qui se plaçaient sur le terrain de la théorie aristotélicienne, pour la défendre ou pour en refuser certaines conclusions, ont abondamment invoqué ce type d'argument. Si le monde est sans commencement, écrira ainsi Philopon, alors on doit dire que le nombre des hommes engendrés jusqu'au temps de Socrate est infini ; en y ajoutant le nombre des hommes engendrés depuis Socrate jusqu'à aujourd'hui, on obtiendrait «quelque chose plus grand que l'infini». Autre paradoxe énoncé par Philopon : quand le Soleil accomplit une révolution, la Lune en effectue 12 ; en supposant le monde sans commencement, on doit bien admettre, d'une

manière ou d'une autre, que les révolutions du Soleil sont infiniment nombreuses, et que, par conséquent, les révolutions de la Lune représentent un infini deux fois plus grand qu'un autre infini - autre «absurdité». On devine la conclusion de Philopon, à savoir que le point de départ de ce raisonnement doit être récusé, et que les aristotéliens eux-mêmes doivent convenir que le monde n'a bien un commencement.

Les critiques de Philopon rencontrèrent un écho considérable dans les traditions grecque, arabe, hébraïque et latine ; elles donneront naissance à une foule d'arguments dérivés. Par exemple, certains théologiens musulmans opposeront aux philosophes arabes aristotéliens un argument qui fera fortune dans les débats de la scolastique médiévale. Supposons le monde sans commencement. Alors le nombre des âmes séparées des corps après la mort, et qui survivent éternellement à ces corps, constitue bien un nombre infini, et actuel ; n'est-ce pas là une contradiction dans le raisonnement des philosophes (aristotéliens), qui devrait les contraindre à reconnaître que la thèse de l'éternité du monde est (philosophiquement) intenable ? Maniant les outils logiques et conceptuels du discours philosophique avec

une efficacité souvent redoutable, ces théologiens concluaient, du reste, à l'impossibilité de toute espèce d'infini, potentiel comme actuel.

On s'en doute, les répliques des philosophes, soucieux de légitimer la cohérence de la pensée d'inspiration aristotélicienne, n'ont pas manqué ; et les arguments déployés à cette occasion ont nourri le débat sur l'infini, qui se prolongera dans les traditions théologiques et philosophiques d'expression latine et hébraïque. La critique du point de vue aristotélicien s'est aussi exercée dans d'autres directions. Beaucoup plus rarement, il est vrai, mais avec au moins autant de rigueur, certains ont adopté un infinitisme de principe, qui les a conduits à remettre en cause l'ensemble du dispositif conceptuel aristotélicien. C'est le cas de Thābit ibn Qurra.

Abū l-Ḥassan Thābit ibn Qurra al-Ḥarrānī, le Sabéen (mort en 901) est l'un des plus prestigieux représentants de la science arabe du IX^e siècle. Originaire de Harrān, en Anatolie, il fut remarqué pour ses dons exceptionnels par les trois frères Banū Mūsā, savants et mécènes de Bagdad, qui l'intégrèrent aux cercles de savants encouragés et financés par le calife. Traducteur du grec en arabe, d'œuvres d'Euclide, d'Archimède, d'Apollonius, de Ptolémée, de Nicomaque de Gérase, Thābit est lui-même l'auteur de traités novateurs, soit fondateurs, dans divers domaines des mathématiques (géométrie infinitésimale, algèbre, arithmétique et théorie des nombres) et de l'astronomie. On lui attribue aussi des écrits médicaux. Il semble avoir supervisé des traductions du syriaque en arabe - la langue syriaque fut un très important relais de la culture grecque avant l'avènement de l'islam ; à Harrān, centre syriaque important, dominait la religion syncrétiste des Sabéens, dont Thābit fut peut-être un dignitaire.

Comment Thābit ibn Qurra envisage l'infini actuel

Le texte de Thābit que nous allons commenter témoigne que ce dernier était, à côté de ses activités proprement scientifiques, profondément engagé dans les débats philosophiques et théologiques de son temps, à propos desquels il affichait avec autorité un point de vue original. Ce point de vue nous est parvenu sous la forme de réponses aux questions posées par un de ses disciples,

le chrétien Abū Mūsā 'Isā ibn Usayyid. De ce texte important, soulignons trois aspects qui fixent le cadre du débat : l'importance de la controverse théologique relative au dénombrement des âmes et à la connaissance par Dieu de l'infinité des êtres singuliers ; l'expression nette d'un infinitisme de principe qui se réclame des propriétés mathématiques du nombre ; la défense d'une conception philosophique du nombre et des figures géométriques qui rompt radicalement avec les principes aristotéliens.

Le texte s'ouvre d'emblée sur la controverse dont nous avons souligné les enjeux : est-il contradictoire d'envisager que les âmes constituent une multiplicité infinie ? Bien que le texte ne le mentionne pas, il est certain que cette question était associée à celle de l'éternité du monde. Toutefois, l'angle sous lequel elle est présentée ici est celui de la nature de la connaissance divine : si celle-ci envisage chaque être dans sa singularité (chaque âme) et dans son devenir, ne porte-t-on pas atteinte à la simplicité absolue de cette connaissance ? N'est-il pas plus juste de postuler que Dieu ne connaît que les « universaux » (les principes généraux : « le » nombre, « la » grandeur, l'immortalité de l'âme) et non les « singuliers » (chaque nombre, chaque grandeur, chaque âme séparée) ?

Intervenant à la fois sur le plan théologique et philosophique, et commentant un exemple astronomique, celui des éclipses, Thābit entend souligner que, de toute manière, on est conduit à admettre la possibilité d'un infini numérique actuel : au cours de sa réponse à la question relative au problème des âmes (à savoir sont-elles en nombre fini ou non ?), Thābit a dit : « Je m'étonne de celui qui dit que Dieu – qu'Il soit béni et exalte – ne connaît pas les particuliers, mais connaît les universaux, et ignore ce qu'il dirait si on l'interrogeait sur Sa connaissance, par exemple, de l'éclipse : est-ce qu'Il en connaît la nature et l'existence seulement, ou bien Il connaît les éclipses qui ont

eu lieu au cours du passé, et celles qui auront lieu, une à une, et le moment déterminé où elles ont eu lieu et où elles auront lieu. S'Il ne connaît rien d'autre que la nature générale de l'éclipse, il en résulterait que Son savoir est moindre que le savoir des astronomes... »

On l'a compris, Thābit n'accepte pas les arguments de ceux qui jugent impossible pour la connaissance divine d'appréhender les « particuliers ». Il rappelle que, même en acceptant cette thèse, on n'en est pas moins conduit à admettre l'idée d'un infini actuel. « De plus, si l'on acceptait, par esprit de conciliation, qu'Il ne connaît pas les particuliers, il n'en reste pas moins qu'il est des choses universelles qui sont des espèces infinies, et qu'Il connaît pour elles-mêmes, simultanément, en acte ; par exemple, les espèces des figures géométriques, dont Il connaît chacune, ses accidents, ses singularités, et tous ses autres

aspects ; et aussi les espèces des nombres. Ces deux espèces sont infinies. Il est donc possible qu'une chose existe en acte, et infinie en nombre ; si cela est possible, alors il n'est pas impossible qu'il en soit de même pour les âmes... »

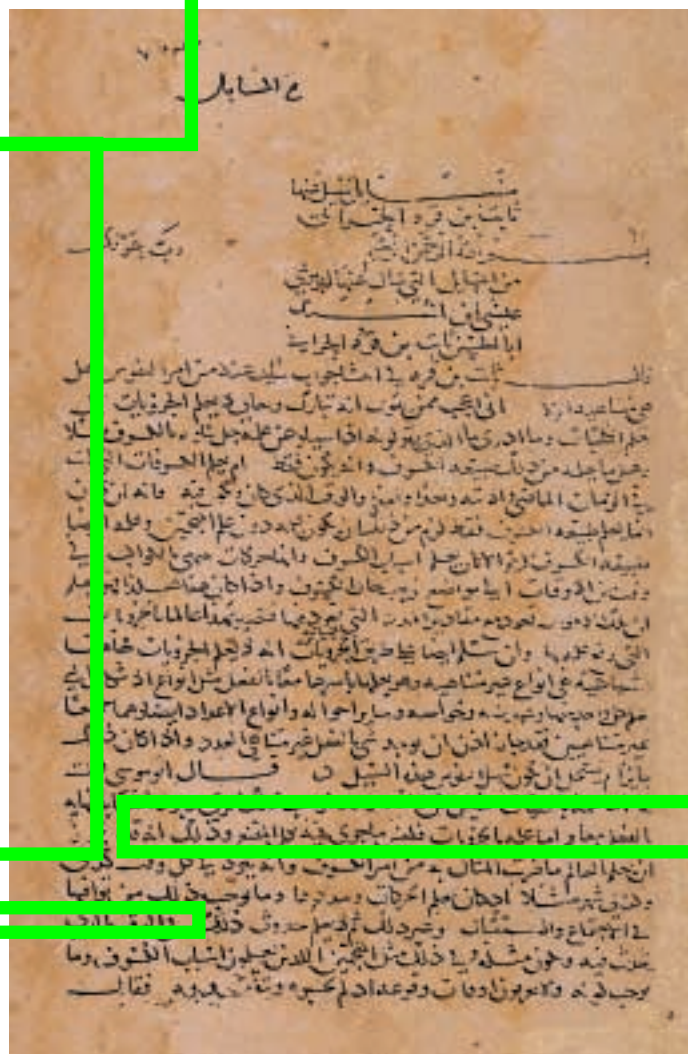
La comparaison des infinis numériques

Au terme de longs développements consacrés à ce problème de la connaissance divine, Thābit est invité par son questionneur à commenter l'argument si souvent invoqué dans les débats sur l'infini, selon lequel « un infini ne peut être plus grand qu'un infini ».

Au-delà de quelques difficultés de lecture, l'argument exposé dans ce texte est aussi clair que novateur. C'est la première fois, à notre connaissance, qu'une véritable arithmétique de l'infini est ébauchée à partir des propriétés arithmétiques du nombre entier. Du reste,

une formulation mathématiquement cohérente n'interviendra que 1 000 ans après Thābit, avec les travaux de Georg Cantor et sa théorie du transfini.

Précisons le contenu de l'argument, pour en mesurer l'audace. Pour Thābit, de toute évidence « il y a un nombre infini », et il s'agit bien d'un infini actuel. Aucune définition de cette notion n'est clairement énoncée, mais on comprend toutefois que l'infinie multiplicité des espèces numériques (on dirait aujourd'hui l'infinitude de l'ensemble des entiers) se identifie à la nature même du « nombre entier » (on dirait aujourd'hui l'infinitude du dénombrable). Répondant ici en mathématicien à l'argument classique : « un infini ne peut être plus grand qu'un infini », autrement dit, « il est absurde d'énoncer des relations d'ordre à propos de l'infini », Thābit réplique qu'il est



2. MANUSCRIT (premier folio) de Thābit ibn Qurra où il réfute qu'un infini ne peut être plus grand qu'un infini.

parfaitement possible de donner un sens mathématique à une telle relation d'ordre. Le geste opératoire décisif de Thābit consiste, pour parler en termes modernes, à définir l'«égalité» de deux ensembles infinis (les nombres pairs et les nombres impairs) par leur équipotence : l'ensemble infini des pairs «est égal» à l'ensemble des impairs, car à tout entier pair on peut associer un entier impair, et réciproquement. À partir de là, il peut conclure que les entiers pairs représentent la moitié de tous les entiers : il y a autant d'entiers pairs que d'entiers impairs, et tous les entiers sont épuisés par cette partition.

En affirmant que l'infini des entiers est «double» de l'infini des entiers pairs, Thābit invalide le fameux argument discuté (il y a bien un infini plus grand qu'un autre infini). Il va même plus loin : en éendant la remarque relative aux pairs et impairs, il exhibe un infini «égal» au tiers de l'infini des entiers (les entiers de la forme $3p$, qui ont la même infinité que les entiers de la forme $3p + 1$ et que celle des entiers de la forme $3p + 2$, ce qui permet de diviser l'infini des entiers en trois infinis «égaux»), un infini «égal» au quart de l'infini des entiers, ..., un infini égal au n -ième de l'infini des entiers.

Opérations sur l'infini

L'égalité naïve, celle qui passe par l'énumération, inapplicable à l'infini, a été remplacée par un autre type de comparaison permettant d'évaluer les rapports de proportion entre «le nombre dans sa totalité» et certaines de ses parties : les nombres pairs et impairs, les multiples de 3, les multiples de 4, les multiples de n .

Ce faisant, Thābit ne met pas en cause l'axiome selon lequel «le tout est plus grand que sa partie». L'infini des entiers, comme un tout, reste plus grand que l'infini des entiers pairs, puisqu'il en est le «double». D'autres

infinis, après lui, sauront orienter leur réflexion en réévaluant le rapport du tout à sa partie, lorsque des quantités infinies sont en jeu. Au XIX^e siècle, Richard Dedekind, à la même époque que Cantor, définira ainsi précisément l'ensemble infini comme étant un ensemble «égal» à l'une de ses parties.

Le lecteur averti de la définition du transfini cantorien aura relevé que Thābit ne dégage pas, comme Cantor, l'idée d'une cardinalité infinie, propre à tout ensemble dénombrable ; dans cette perspective, tous les ensembles infinis que nous avons évoqués seraient «équipotents», c'est-à-dire égaux au sens de Cantor, ce qui n'est pas le cas pour Thābit. Ainsi, Thābit ne dit pas qu'il y a «autant» de nombres entiers que de nombres entiers pairs, dans la mesure où l'on peut faire correspondre à tout entier un entier pair et réciproquement. Pour autant, il n'est pas nécessaire de voir dans Thābit une «anticipation» de Cantor pour mesurer l'intérêt de la pensée du savant arabe.

En revanche, malgré le millénaire d'histoire qui les sépare, on pourrait rapprocher ces deux savants par le sentiment qu'ils habitent tous deux quant au nécessaire élargissement de l'idée de nombre, telle qu'elle dominait les cultures respectives de leurs époques. Pour fonder cette remarque, il conviendrait d'analyser minutieusement les développements consacrés par Thābit au nombre, des catégories de pensée, au nombre des espèces de la quantité, à la quantification de la qualité, au rapport du genre à l'espèce, au statut des accidents. Voilà qui dépasse largement notre propos. Contentons-nous de résumer l'originalité et l'audace théorique de Thābit : non seulement il opère sur des infinis, leur conférant ainsi un statut que leur refusait la tradition aristotélésienne, mais, ce faisant, il élargit la notion même de nombre et le cadre conceptuel constitutif de cette tradition.

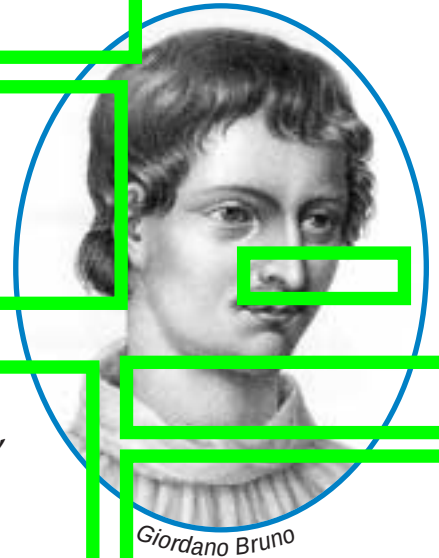
Tony LÉVY, Institut Warburg-CNRS, Centre d'histoire des sciences et des philosophies arabes et médiévales (Paris-Villejuif).

Shlomo PINES, *Thābit ibn Qurra's Conception of Number and Theory of the Mathematical Infinite*, in Pines, *Studies in Arabic Versions of Greek Texts and in Mediaeval Science (The Collected Works of Shlomo Pines, vol. II)*, Édition Magnes Press, Jérusalem, 1986, pp. 422-429.

Tony LÉVY, *Figures de l'infini. Les mathématiques au miroir des cultures*, Éditions du Seuil, Paris, 1987, pp. 101-109.

Abdelhamid I. SABRA, *Thabit ibn Qurra on the Infinite and Other Puzzles. Edition and Translation of his Discussions with Ibn Usayyid*, in *Zeitschrift für Geschichte der Arabisch-Islamischen Wissenschaften*, vol. 11, pp. 1-33, 1997.

L'univers infini de Bruno



GIULIO GIORELLO

Giordano Bruno, mort sur le bûcher le 17 février 1600, fut plus audacieux que Copernic : il envisageait un univers infini, sans horizon ni limite.

Philosophe ou astronome ? Poète ou visionnaire ? Écrivain ou homme de sciences ? Qui était Giordano Bruno ? Il se présente lui-même dans *Le souper des Cendres*, écrit en 1584, confiant à Teofilo le soin de prononcer son propre éloge : «Voici apparaître l'homme qui a franchi les airs, traversé le ciel, parcouru les étoiles, dépassé les limites du monde, dissipé les mirages imaginaires des premières, huitième, neuvième, dixième et autres sphères qui auraient pu leur être ajoutées selon les vains calculs des mathématiciens et suite à l'aveuglement des philosophes vulgaires.»

Quelques lignes auparavant, il comparait les «découvreurs de la «véritable nature» du ciel aux pionniers des explorations géographiques. Cette comparaison sera reprise par Galilée en 1623 dans les premières pages de *L'essayeur*. Galilée lisait Bruno, le paraphrasait, mais ne pouvait, en ces temps d'Inquisition, le citer explicitement.

Bruno est l'égal de navigateurs, tels Tiphys orientant Jason et les Argonautes vers la Colchide, pour conquérir la toison d'or, ou Christophe Colomb, le découvreur du Nouveau Monde. La comparaison s'arrête là : si les victimes des barbaries perpétrées par les Argonautes étaient des hommes, Bruno ne combattit que les croyances non fondées transmises par tradition et jamais remis en cause. Le «philosophe de la nature», tel que Bruno se décrit dans *Le souper des Cendres*, découvre le moyen de monter au ciel, de parcourir la circonférence des étoiles et de laisser derrière lui la surface convexe du firmament.

Selon Bruno, notre raison n'est plus entravée par des sphères imaginées par les mathématiciens pour représenter l'Univers : il n'y a qu'un ciel, une immense région éthérée où brillent de magnifiques foyers lumineux. Dans *Le souper des Cendres*, les étoiles sont des corps enflammés représentant, aux yeux de Bruno, les ambassadeurs de l'excellence de Dieu, les «héralds de sa gloire et de sa majesté». D'après Bruno, nous sommes conduits à découvrir l'effet infini de la cause infinie, la trace vivante et véritable de la vigueur infinie. Dieu aux pouvoirs infinis ne pouvait créer qu'un Univers infini. La puissance divine place Dieu à l'intérieur de l'homme, indiquant ainsi que ce n'est pas hors de nous qu'il faut chercher la divinité, puisqu'elle est en nous, en notre for intérieur.

Même pour d'hypothétiques habitants d'autres mondes, tels les Sélénites, habitants de la Lune, la présence divine n'est pas à rechercher auprès de nous, mais en nous : Bruno affirme que la Lune n'est pas plus pour nous un ciel que nous ne sommes un ciel pour elle. Cette symétrie, évoquée dans *Le souper des Cendres*, est reprise avec force par Galilée en 1632 dans la Première Journée des *Dialogues des systèmes majeurs* ; elle est l'essence même de ce que nous nommons aujourd'hui le principe cosmologique copernicien : la Terre ne peut être au centre de l'Univers et occuper une situation privilégiée. En 1543, Copernic place non plus la Terre au centre de l'Univers, mais le Soleil ; les astres, y compris la Terre, décrivent des sphères autour du Soleil.

Le dépassement du principe copernicien

Au Teofilo du *Souper* correspond le Filoteo de *L'infini, de l'univers et des Mondes*. Ensemble, ils reconnaissent à Copernic le mérite d'avoir détruit cette «erreur fausse supposition de la fixité de la Terre», les mouvements des astres de la voûte céleste ne sont pour Bruno que des apparences liées aux mouvements de l'astre où nous sommes et dont le mouvement ne nous est pas sensible. Cependant, Copernicien appelait aux «sphères ayant une surface concave et convexe» décrites par Bruno et aux «orbes déférents», autrement dit aux espaces circonscrits par l'orbite des astres. Dépassant Copernic, Bruno attribue un caractère d'infini à l'Univers et s'affranchit des sphères coperniciennes. Ainsi, la paternité du principe copernicien, tel qu'on l'énonce aujourd'hui, ne peut être attribuée au seul astronome polonais.

La principale objection faite à la doctrine copernicienne sur le mouvement terrestre reposait sur l'impossibilité qu'une pierre jetée verticalement en l'air retombe suivant la même trajectoire ; le très rapide mouvement de la Terre devait nécessairement la laisser loin derrière, «du côté de l'Occident». Ce problème avait fait discourir nombre de philosophes du Moyen Âge. Dans le troisième dialogue du *Souper*, Teofilo balaie cette objection, affirmant que se meut avec la Terre tout ce qui se trouve sur Terre. Bruno décrit une situation similaire : une pierre lancée en l'air, verticalement, par une personne à bord

d'un bateau retombera en suivant la même ligne droite, quelle que soit la vitesse du navire, et à condition qu'il ne gête pas. De cet exemple, Giordano Bruno explique que le mouvement régulier de la Terre rend impossible, à tout homme, l'observation de son mouvement, à moins qu'on ne l'observe de loin, du Soleil par exemple; ce qui semble difficilement réalisable!

La relativité des mouvements

L'idée de la relativité des mouvements affirmée par Bruno et reprise par Galilée dans la Seconde Journée du *Dialogue* ne présente rien d'original en soi. Il en existe de nombreux précédents. Bruno innove en postulant un mouvement relatif et en défendant la doctrine copernicienne, au sein d'un Univers sans limite. *De l'infini, de l'univers et des Mondes* nous décrit un ciel unique : un espace immense, un «contenant universel», une «région éthérée» où tout se déplace et se meut. Bruno évoque les sens qui nous montrent d'innombrables étoiles, astres, globes, soleils et terres, et la raison qui nous en fait admettre une infinité. Il affirme

que l'Univers immense et infini est composé de cet espace et de tous les corps qui le constitue.

Notre vision cosmologique est perturbée par nos sensations : spectateur immobile du théâtre de l'Univers, nous ne percevons pas les mouvements de la Terre. Le philosophe, tel le peintre, décrit l'Univers et en change notre connaissance. Cependant, nous lisons dans *Le scepter* que le peintre peint un paysage où, pour conformer son art à la nature, il montre ici le palais d'un roi, là une forêt, ailleurs un morceau de ciel. Dans sa comparaison entre le peintre et le philosophe, Giordano Bruno explique que le peintre est toujours susceptible d'arrêter son travail, d'en examiner les défauts avec le recul et la distance que prennent ordinairement les maîtres de l'art. Le philosophe, quant à lui, fait partie intégrante du monde qu'il prétend décrire. Il ne peut en sortir, à moins de tomber dans les abîmes du néant. Il fait partie de cette réalité qu'il essaie de comprendre. Dans l'Univers infini de Bruno, cette vision du philosophe sur le monde qui l'entoure intègre Dieu : créateur étranger à ses propres créations, Dieu finit par se retrouver nulle part. Le Seigneur ne contemple pas son

œuvre. Au plus haut des cieux, il y habite, il habite en nous, et c'est peut-être l'unique raison pour laquelle nous pouvons dire je suis».

La situation du philosophe peintre qui ne peut sortir de la scène qu'il dépeint ne constitue pas un frein à la connaissance, au contraire, elle est la condition même de sa possibilité. Bruno illustre par ses critiques de la séparation aristotélicienne entre la physique céleste et la physique terrestre : la Lune, qui dans les cosmologies traditionnelles est le premier des corps célestes parfait et immuable, est en réalité aussi imparfaite que la Terre. Seule la distance entre la Terre et la Lune interdit à nos sens de percevoir les aspérités lunaires. Bruno, dans *De l'infini, de l'univers et des Mondes*, souligne que, s'il y avait sur la Lune des habitants observant leurs cieux, ils verraient notre globe comme immuable et parfait.

L'ascension vers l'infini

En 1591, dans *De l'immensité*, Bruno imagine une ascension dans les cieux. Aux yeux du voyageur s'éloignant de la Terre, les détails des paysages terrestres s'estompent progressivement :



1. GIORDANO BRUNO a exploré le cosmos et découvert des mondes encore inconnus (gravure extraite de l'ouvrage de Camille Flammarion, *L'atmosphère : météorologie populaire*, publié en 1888).

les rivières, les forêts, les champs, les villes, tout finit par se confondre. Les seules différences perceptibles résultent de la réflexion de la lumière solaire sur les montagnes et sur les plaines. L'observateur lunaire voit la Terre tachetée, tout comme le terrien perçoit la surface lunaire. Bruno explique ainsi les taches lunaires comme des phénomènes optiques, en accord avec les critiques de la tradition héritée d'Aristote. Une position sur laquelle s'aligneront, au début du XVII^e siècle, Kepler et Galilée.

C'est parce que nous sommes éloignés des autres corps célestes que nous

les percevons immuables. À Galilée d'ajouter qu'il n'y a aucune qualité dans les cieux qui ne puisse se retrouver sur Terre. Quant à Bruno, il affirmait déjà qu'il n'y avait aucun point de vue privilégié.

Aujourd'hui, nous admettons aisément qu'il suffirait d'une très légère modification des constantes physiques ou des lois de la nature pour que l'Univers ne soit habité d'aucun «peintre», tel que le définissait Bruno, qui puisse le représenter. Bruno mentionne ainsi, dans *De l'infini*, que la plus petite erreur au commencement cause à la fin de

grandes différences ; un simple inconvenient se multiplie progressivement et se ramifie en une infinité d'autres, comme d'une racine minuscule naissent de grandes structures et des branches innombrables. Ces phénomènes évoqués par Bruno dans *De l'infini* sont aujourd'hui bien connus des spécialistes sous le qualificatif de système dépendant des conditions initiales.

Pour Giordano Bruno, c'est Dieu, ou peut-être le Diable, qui se cache derrière ces infimes détails, lourds de conséquences. L'infime perturbation des conditions initiales modifie-t-elle notre vision du monde ou le monde en lui-même ? Bruno montre que, si l'on admet la pluralité des mondes et des univers, alors l'homme apparaît comme un observateur doué d'intelligence, appartenant à un univers à un moment précis de son évolution, sans possibilité de revendiquer un quelconque point de vue privilégié.

Le philosophe se retrouve alors face à la question de penser cette «identité». Bruno le savait bien : son univers sans limite ni horizon, sans les «murailles imaginaires» des sphères célestes, apparaissait chaotique et confus à celui qui croyait que la Providence divine avait placé toute chose au bon endroit. Toutefois, aux yeux de Bruno, ce bel ordre et cette échelle de la nature ne sont qu'un songe, un conte pour enfants. Peut-être est-ce aussi pour cela que Bruno finit sur le bûcher le 17 février 1600. On raconte que, durant son incarcération à Rome, il n'a eu cesse d'enseigner à ses compagnons de cellule que tous ces points lumineux que nous appelons les étoiles et qu'ils devaient à travers l'étroite fenêtre de leur cachot former des mondes tout aussi complexes que le nôtre.

Giulio GIOIELLO est professeur d'épistémologie à l'Université de Milan.

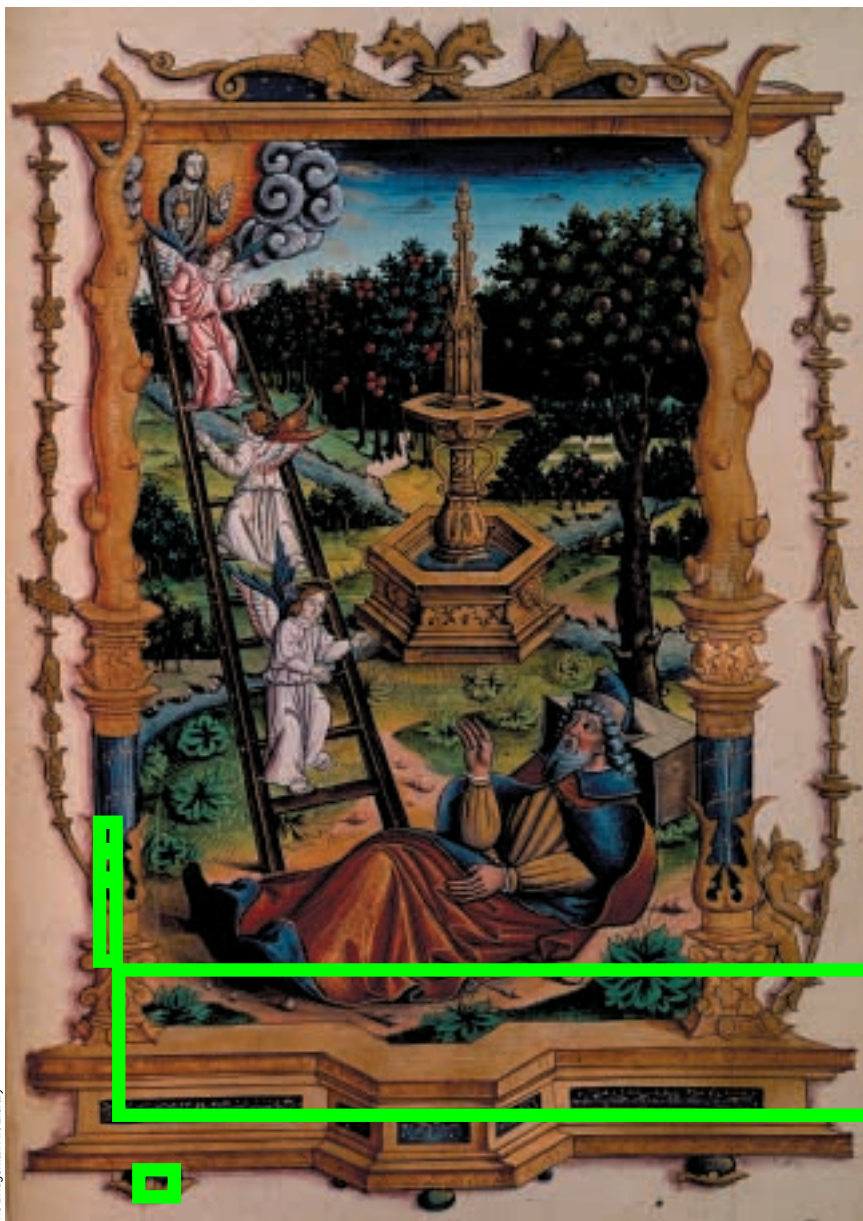
Giordano BRUNO, *Cœuvres Complètes*, Les Belles Lettres, 1994, tome II, *Le souper des Cieux*, traduction de Y. Hersant, tome III, *De l'infini, de l'univers et des mondes*, traduction J.-F. Cavanié.

GALILÉE, *L'Essayer*, Les Belles Lettres, Paris, 1990.

Galilée, *Dialogue sur les deux grands systèmes du monde*, Éditions du Seuil, 1992.

G. AQUILECCHIA, *Giordano Bruno*, Les Belles Lettres, Paris 2000.

P.-H. MICHEL, *La cosmologie de Giordano Bruno*, Hermann, Paris, 1962.



2. L'ÉCHELLE DE JACOB extraite des «Chants royaux sur la Conception couronnée du Puy de Rouen» (début du XVI^e siècle). Ce manuscrit montre Jacob endormi qui rêve de l'échelle permettant d'accéder aux cieux. «Alors il vit en songe une échelle, dont le pied était appuyé sur la terre, et le haut touchait le ciel, et des anges de Dieu montaient et descendaient cette échelle.» À l'instar de Jacob, Giordano Bruno a gravi quelques marches d'une échelle : non pas celle qui menait au ciel, mais celle qui conduisait à la connaissance des cieux.

La science du mouvement au XVII^e siècle

MICHEL BEAUF

À la fin du XVII^e siècle et au début du XVIII^e, les mathématiciens décrivent le mouvement et introduisent la notion de vitesse, en évitant les paradoxes qui résultaient de l'usage impropre et intuitif de l'infini.



Au V^e siècle avant notre ère, dans la ville grecque d'Élée, le philosophe Zénon bute sur un grave paradoxe quand il analyse le mouvement : pour atteindre un point donné, un mobile doit d'abord parcourir la moitié de la distance qui le sépare de ce but, puis la moitié de la distance restante, puis encore la moitié de la distance restante, et ainsi de suite à l'infini. Comment parcourir une infinité de moitiés en un temps fini ? Cet infini qui s'introduit dans l'analyse du mouvement troublera physiciens et mathématiciens jusqu'à la fin du XVII^e siècle. À cette époque, et au début du siècle suivant, la science du mouvement se clarifie enfin. Une partie notable des progrès est due au physicien français Pierre Varignon (1654-1722), qui publie plusieurs articles dans les *Mémoires de l'Académie royale des sciences*, autour des années 1700. Varignon se fonde sur les résultats géométriques obtenus entre 1684 et 1686 par Gottfried Wilhelm Leibniz (1646-1716), puis par Isaac Newton (1642-1727) : ce « calcul différentiel et intégral » permet de déterminer mathématiquement la pente de la tangente à une courbe, ou l'aire entre

une courbe et un axe. L'utilisant, Varignon trouve une formule très simple, qui décrit le mouvement d'un corps sur lequel s'exerce une force, notamment dans le cas des forces centrales, c'est-à-dire des forces qui s'exercent entre un point, attracteur ou répulsif, et un mobile qui se déplace sur une droite.

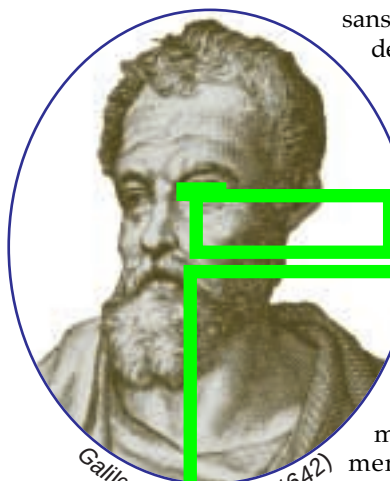
L'impossible géométrie du mouvement

Comment le calcul différentiel a-t-il permis de résoudre les difficultés antérieures de la science du mouvement ? Pour le comprendre, revenons d'abord aux enjeux de la géométrisation de la science du mouvement au XVII^e siècle. Les physiciens voulaient alors comprendre les mouvements par des interprétations géométriques déductives, sur le modèle des *Éléments* d'Euclide. Les principales questions portaient sur la continuité des mouvements, sur leur début et sur leur fin. Comment expliquer la variété des mouvements accélérés ? Étaient-ils, comme l'avaient proposé certains philosophes, un mélange de mouvements et de repos ? Ces questions avaient troublé Galilée (1564-1642), Bonaventura

Cavalieri (1598-1647), Blaise Pascal (1623-1662) et d'autres savants du XVII^e siècle.

Dans une lettre adressée à Galilée en date du 21 mars 1626, Cavalieri souligne l'importance et la difficulté des problèmes posés, dans le cadre de la géométrisation, par le commencement et l'évolution continue du mouvement : « Lorsque l'on en arrive à devoir prouver que le mobile qui, du repos, doit passer à un degré quelconque de vitesse, [...] par des degrés intermédiaires, je ne trouve aucune raison qui me tranquillise bien qu'il me semble que généralement il en soit ainsi. » La quête d'une raison qui tranquillise est à la fois un programme de travail et une attitude d'esprit : c'est la volonté de comprendre le commencement et l'évolution continue du mouvement ; c'est aussi la volonté d'expliquer le mouvement par les mathématiques.

Cette étude est très difficile, car l'infini y apparaît immédiatement. Pascal le mentionne explicitement dans son petit traité intitulé *De l'esprit géométrique* : « Quelque prompt que soit un mouvement, on peut en concevoir un qui le soit davantage, et hâter encore ce dernier ; et ainsi toujours à l'infini,



Galileo Galilei (1564-1642)

sans jamais arriver à un qui le soit de telle sorte qu'on ne puisse plus y ajouter. Et au contraire, que que lent que soit un mouvement, on peut le retarder davantage, et encore ce dernier ; et ainsi à l'infini, sans jamais arriver à un tel degré de lenteur qu'on ne puisse encore en descendre à une infinité d'autres, sans tomber dans le repos». En langage moderne, les questions sont : comment traiter mathématiquement de vitesses infiniment petites ? Comment traiter mathématiquement de vitesses infiniment grandes ?

Galilée rencontre les mêmes difficultés. Dans ses *Discours et démonstrations mathématiques concernant deux sciences nouvelles*, publiés en 1638, il fait dire à son ami et porte-parole Salviati : « Vous ne refuserez pas, je crois, de m'accorder qu'une pierre, tombant de l'état

de repos, acquiert ses degrés successifs de vitesse selon l'ordre dans lequel ces mêmes degrés diminueraient et se perdraient, si une force motrice la reconduisait à la même hauteur ; et le refuseriez-vous que je ne vois pas comment la pierre, dont la vitesse diminue et se consume en totalité au cours de son ascension, pourrait atteindre l'état de repos sans être passée par tous les degrés successifs de lenteur. »

Galilée souligne la continuité caractérisant, selon lui, la croissance ou la décroissance de la

vitesse dans un mouvement naturellement accéléré, comme la chute d'un corps pesant (un « grave »). Dans un tel mouvement, « un grave ne demeure en aucun de ces degrés de

vitesse pendant un temps fini ». Ce qui revient encore à dire que, dans un mouvement accéléré ou retardé, un corps qui sort du repos ou qui y retourne passe par une infinité de degrés de vitesses dans un intervalle de temps qui, si petit qu'il soit, contient une infinité d'instants.

Ces problèmes soulevés par cette analyse sont d'une extrême difficulté : « il n'y a une infinité de degrés de lenteur pour atteindre le repos, ne faut-il pas un temps infini pour que cet arrêt du mouvement puisse s'accomplir, pour que le mobile animé d'un tel mou-

vement s'arrête en passant successivement par tous les degrés de lenteur ? À l'inverse, pour qu'un mouvement commence en passant successivement par tous les degrés croissants de vitesse, ne faut-il pas, là aussi, un temps infini pour atteindre la moindre vitesse ? Dans un cas le repos est impossible à atteindre ; dans l'autre c'est le mouvement qui ne peut commencer. Or, bien évidemment, les mouvements commencent et finissent.

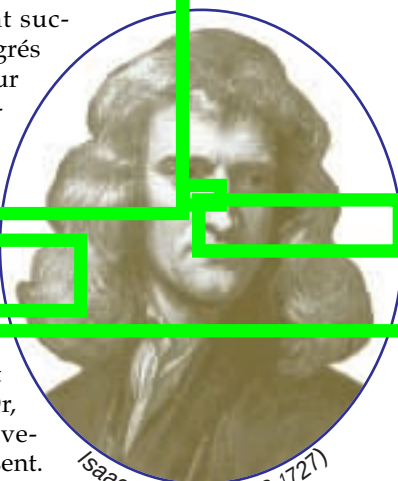
De l'état de repos à une vitesse finie

Une réponse consiste à rejeter l'idée d'un commencement non fini du mouvement ou, plutôt, à considérer que le mouvement de chute des graves commence non pas à partir d'une vitesse nulle par accroissements successifs continus, mais à partir d'une vitesse finie. C'est ainsi qu'Edme Mariotte (1620-1680), dans son *Traité de la percussion ou choc des corps*, publié à Paris en 1673, refuse l'idée qu'un mouvement accéléré le soit dès le premier instant. Son argumentation repose sur un rappel des difficultés engendrées

par les paradoxes de Zénon sur l'infini et aussi sur diverses expériences en rapport avec l'écoulement des fluides. Il propose d'éliminer l'infini au début du mouvement en considérant qu'il n'y a pas, au sens strict, de début du mouvement. Dès le premier instant, le corps est animé d'une vitesse très petite, mais finie.

De même, le mathématicien allemand Hartsoeker propose que le repos soit atteint sans que le mobile passe par tous les degrés de lenteur. Dans une lettre adressée à Leibniz le 6 janvier 1712, Hartsoeker écrit : « Il y a une loi dans la Nature, dites-vous, Monsieur, qui porte qu'il n'y a aucun passage *per saltum*

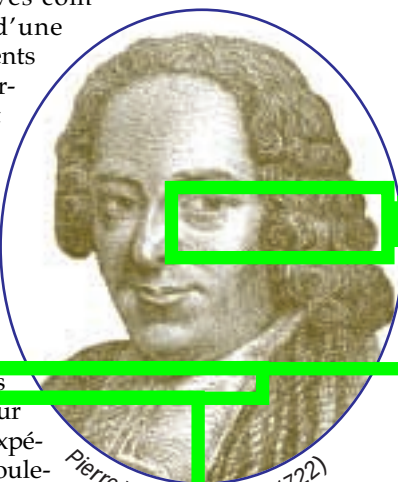
[par saut]. Je vous l'accorde dans un certain sens ; mais quand vous dites que cette loi ne permet pas qu'il n'y ait point de milieu entre le dur et le fluide, je n'y vois aucune nécessité. Si vous ne saviez pas par l'expérience, Monsieur, qu'un corps qui se meut avec



Isaac Newton (1642-1727)



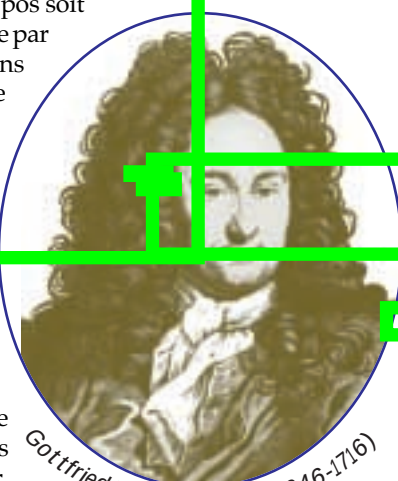
Blaise Pascal (1623-1662)



Pierre Varignon (1654-1722)



Bonaventura Cavalieri (1598-1647)

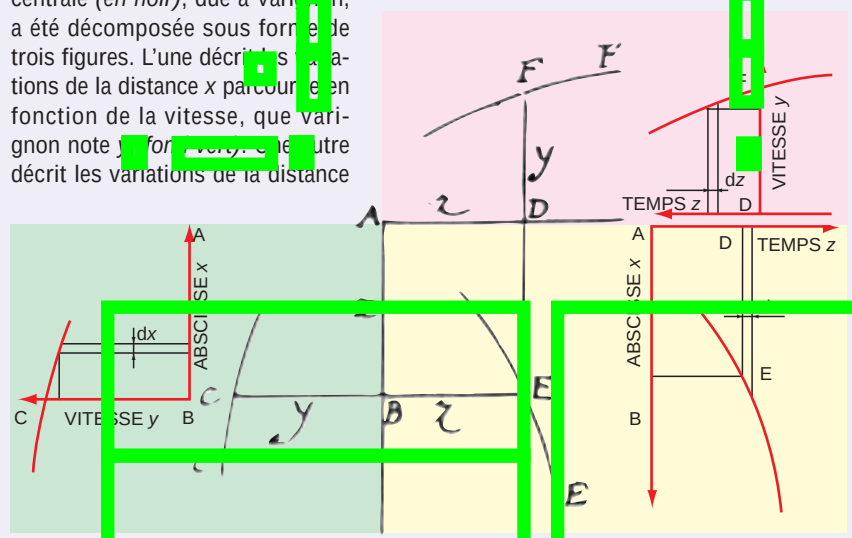


Gottfried Wilhelm Leibniz (1646-1716)

Le mémoire du 5 juillet 1698 de Varignon

Dans ce mémoire, Pierre Varignon explicite algébriquement les quantités finies et infiniment petites qui s'imposent pour la description des mouvements. La figure centrale (en noir), due à Varignon, a été décomposée sous forme de trois figures. L'une décrit les variations de la distance x parcourue en fonction de la vitesse, que Varignon note y (fond vert) ; une autre décrit les variations de la distance

x parcourue en fonction du temps, que Varignon note z (fond jaune). La troisième décrit les variations de la vitesse en fonction du temps z (fond rose).



tout autant de vitesse qu'il vous plaira peut demeurer en repos dès l'instant du choc, sans perdre peu à peu et par degré son mouvement, ne diriez-vous pas par votre loi que cela est impossible?»

À ces approches, on peut associer celles de Pierre Gassendi (1592-1655) ou, plus tard, de son commentateur François Bernier, pour lesquels le mouvement varié est un mélange de repos et de mouvements, la continuité n'étant qu'apparence, en raison d'une certaine persistance des images rétiniennes, pour parler en termes modernes. Toutefois ces approches n'offrent aucune prise réelle à la géométrisation. La pseudo-évidence expérimentale, de Mariotte en particulier, n'apprend rien.

Un premier effort de Leibniz

Alors qu'il ne développera son calcul différentiel et intégral qu'en 1676, Leibniz rédige en 1670 une *Théorie du mouvement abstrait*, qui a essentiellement pour objet d'édifier une théorie purement rationnelle du mouvement. Ce travail s'inspire de résultats mathématiques que Cavalieri avait exposés dans sa *Géométrie au moyen des indivisibles*, publiée à Bologne en 1635.

Leibniz considère que le mouvement est continu, c'est-à-dire qu'il n'est «nullement entrecoupé de petits repos», comme cela était parfois envisagé par les «atomistes», c'est-à-dire ceux qui acceptaient l'idée du discontinu dans

la matière, l'espace ou le temps. Continu, le mouvement est, non seulement divisible à l'infini, mais effectivement divisé, en ce sens qu'il y a des parties données en acte dans le continu et que ces-ci sont infinies en acte». Leibniz introduit donc un infini «en acte» ou «actuel», c'est-à-dire qu'il pose effectivement un infini. Cependant, pour Leibniz, «il n'y a pas de minimum dans l'espace ou le temps», car un tel minimum «implique contradiction» : s'il existait un minimum, il y aurait autant de minimums dans le tout que dans la partie, puisque toute partie de même espèce que le tout est encore infiniment divisible. Il s'agit là, toujours, des paradoxes de Zénon : un segment, aussi petit que l'on veut, contient une infinité de points ; c'est un infini en acte.

Leibniz échappe à cette contradiction en adaptant la méthode de Cavalieri, méthode qui vise, en restant à l'intérieur des mathématiques euclidiennes, à traiter de questions qui engageraient la considération de sommes infinies. Pour cela, Cavalieri introduit la notion d'agrégat (on parlera ainsi, par exemple, de l'agrégat de toutes les lignes ou de tous les degrés de vitesse), et établit les rapports entre l'agrégat de toutes les lignes de deux figures et celui des aires de ces deux figures. Cette approche précise et rigoureuse sur le plan mathématique sera, le plus souvent, interprétée comme la mise en œuvre d'une simple somme

de points, de lignes, etc., ce qui, bien évidemment, repose avec acuité la question du continu et de l'infini.

Leibniz est plus préoccupé par l'analyse du mouvement et des trajectoires que par les pures questions de géométrie, qui reposent le problème de la composition du continu. Il introduit son concept d'«indivisible» afin de résoudre le problème du mouvement : «Des indivisibles ou inétendus sont donnés, sans quoi ni le commencement ni la fin du mouvement et du corps ne sont concevables.»

Pour Leibniz, le commencement appartient à l'espace, au temps, au mouvement, sans pour autant être divisible. Comme on l'a vu, la notion d'un commencement divisible est contradictoire. Il y a donc des indivisibles, constitutifs de l'espace, du temps et du mouvement et cependant, hétérogènes à ce qu'ils constituent. Ce faisant, Leibniz introduit alors des êtres mathématiques pour le moins surprenants, que sont «le commencement du corps, de l'espace, du mouvement, du temps (à savoir le point, l'effort, l'instant)». C'est à Thomas Hobbes (1588-1679) que Leibniz emprunte, en le transformant, le concept d'effort pour en faire son indivisible de mouvement : «L'effort est au mouvement ce que le point est à l'espace, soit comme l'unité à l'infini, il est en effet le commencement et la fin du mouvement.» Il existe donc un indivisible de mouvement, l'«effort», qui, d'un certain point de vue, bloque la régression à l'infini, qui interdisait de penser le commencement ou la fin du mouvement.

Varignon et l'algorithme de la cinématique

Leibniz publie le texte fondateur du nouveau calcul différentiel et intégral en octobre 1684. Le titre de ce mémoire indique bien son objet : *Nouvelle méthode pour chercher les maxima et les minima, ainsi que les tangentes, méthode que n'embarassent ni les expressions fractionnaires ni les expressions irrationnelles*. Leibniz écrit notamment : «Lorsqu'on connaît l'algorithme, si je peux dire, de ce calcul que j'appelle différentiel, on peut trouver par le calcul ordinaire toutes les autres équations différentielles, les maxima et minima, ainsi que les tangentes.»

La diffusion du nouveau calcul est lente et difficile. À partir des années 1690, les deux frères Jean (1667-1748)

et Jacques Bernoulli (1654-1705), en contact étroit (principalement épistolaire) avec Leibniz, appliqueront les nouvelles méthodes de ce dernier aux sujets les plus divers, l'occasion leur en étant souvent fournie par les défis mathématiques qu'ont l'habitude de se lancer les savants d'Europe.

Au cours de son séjour parisien, pendant l'hiver 1691-1692, Jean Bernoulli initie, par ses *Leçons particulières*, le marquis Guillaume de L'Hospital (1661-1704) au calcul leibnizien. C'est lui qui saura lever les indéterminations du type $0/0$. Ces leçons, portant à la fois sur le calcul différentiel et sur le calcul intégral, constitueront la base à partir de laquelle L'Hospital rédigera le premier manuel de calcul différentiel. Celui-ci sera publié à la fin du mois de juin 1696 sous le titre *Analyse des infiniment petits, pour l'intelligence des lignes courbes*. L'introduction des nouvelles méthodes se trouve ainsi grandement facilitée.

Malgré les vives polémiques que suscitent les méthodes à l'Académie royale des sciences de Paris jusqu'à la fin des années 1710, Varignon assimile rapidement les principaux éléments du nouveau calcul, et il s'attache, entre les années 1695 et 1705 à reprendre l'étude du mouvement. Le travail de Varignon s'effectue en deux phases : d'abord, il introduit le concept de «vitesse dans chaque instant» ; puis il définit la «force accélératrice dans chaque instant». «Ouvrant les formules qu'il vient de construire à l'issue de cette double construction, Varignon en déduit ce que l'on nomme aujourd'hui l'«algorithme de la cinématique».

Dans deux mémoires lus aux séances de l'Académie royale des sciences, respectivement le samedi 5 juillet et le samedi 6 septembre 1698, Varignon donne un nouveau départ à la science du mouvement,

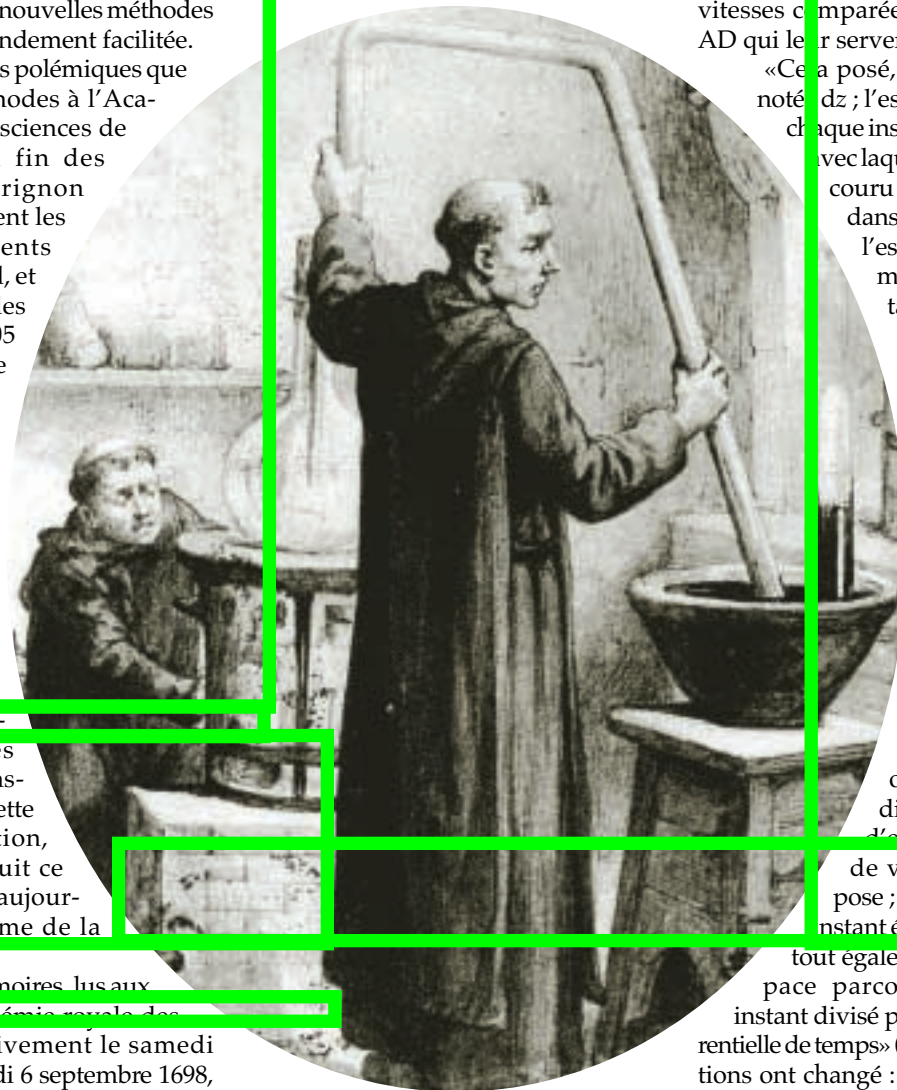
en élargissant d'abord le concept de «vitesse dans chaque instant». Le premier mémoire donne une expression générale de la vitesse susceptible de traiter tous les mouvements dans le cas de trajectoires rectilignes, quel que soit le mode de variation de la vitesse ; il est intitulé *Règle générale pour toutes sortes de mouvements de vitesses quelconques variées à discrétion*.

Varignon considère comme uniforme la vitesse d'un corps pendant chaque instant de son mouvement. Il définit tout d'abord un certain nombre de variables à partir desquelles la mise en place conceptuelle peut être réalisée : «Soient $AB = x$ (voir l'encadré page ci-contre) les espaces parcourus en quelque sens qu'on voudra, $BE = z$ les temps employés à les parcourir, et $BC = y = DF$ les vitesses à chaque point

B de ces espaces.» Les variables, espace, temps et vitesse, étant maintenant représentés par les «axes», ou segments de droite, AB , BE et BC (ou DF), il associe ces variables deux à deux, afin d'engendrer ce qu'on nommerait, en termes modernes, un diagramme des espaces EE , et deux diagrammes des vitesses CC et FF : «De manière que la courbe EE (quelle qu'elle soit) formée par tous les points E , exprime les espaces par ses abscisses AB , et les temps employés à les parcourir, par ses ordonnées correspondantes BE ; la courbe CC (quelle qu'elle soit aussi) formée par tous les points C exprime de même les espaces par les abscisses AB , et les vitesses à chaque point B de ces espaces, par ses ordonnées correspondantes BC ; enfin la courbe FF (quelle qu'elle soit encore) exprime aussi par ses ordonnées DF ces mêmes vitesses comparées aux temps BE ou AD qui leur servent d'abscisses.»

«Cela posé, les instants seront notés dz ; l'espace parcouru dans chaque instant, dx ; et la vitesse avec laquelle dx aura été parcouru sera noté y .» Ainsi, dans un instant égal à dz , l'espace parcouru par le mobile sera égal à dx , tandis que la vitesse de ce mobile sera égale à y . Varignon affirme alors que cette vitesse, dans chaque instant, peut «être regardée comme uniforme : de sorte que $y \pm dy = y$, la notion seule des vitesses uniformes donnera $y = dx/dz$ pour la règle de tous les mouvements variés comme on voudra, c'est-à-dire quelque rapport d'espace de temps ou de vitesse, qu'on suppose ; la vitesse de chaque instant étant toujours et par-

tout égale au quotient de l'espace parcouru dans chaque instant divisé par cette même différentielle de temps» (aujourd'hui les notations ont changé : le temps s'exprime par t et non plus par z , et la vitesse par v et non plus par y).



EDME MARIOTTE (1620-1684) a établi avec Robert Boyle la loi de compression des gaz, qu'il étudie ici expérimentalement. Il publia également un *Traité de la percussion ou choc des corps*, où il refusait l'idée qu'un mouvement accéléré le soit dès le premier instant. Il proposait que, dès le début d'un mouvement, le corps est animé d'une vitesse très petite, mais finie.

La règle générale de Varignon

Il est important de souligner, comme Varignon le fera explicitement dans un mémoire en date du 6 juillet 1707, intitulé *Des mouvements variés à volonté, comparés entre eux et avec les uniformes*, que «l'espace et le temps étant des grandeurs hétérogènes, ce n'est point proprement elles qu'on compare ensemble dans le rapport qu'on appelle vitesse, mais seulement les grandeurs homogènes qui les expriment, lesquelles sont ici, et seront toujours dans la suite ou deux lignes, ou deux nombres, ou deux autres grandeurs homogènes qu'on voudra». Ainsi, Varignon fait un pas décisif en montrant que l'on peut dépasser le débat sur les grandeurs homogènes en comparant non pas directement les grandeurs, mais les nombres qui les expriment en fonction des unités préalablement définies.

Le raisonnement qui conduit Varignon au concept de vitesse dans chaque instant repose directement sur les procédures de calcul des différences mises en œuvre en particulier dans l'*Analyse des infiniment petits, pour l'intelligence des lignes courbes*, publiée à Paris en 1696 par L'Hôpital. «On demande qu'on puisse prendre indifféremment l'une pour l'autre deux quantités qui ne diffèrent entre elles que d'une quantité infiniment petite ; ou (ce qui est la même chose) qu'une quantité qui n'est augmentée ou diminuée que d'une autre quantité infiniment moindre qu'elle puisse être considérée comme demeurant la même.» Ce qui conduit Varignon à confondre $v \pm dv$ avec v . Cette vitesse v peut alors être «considérée comme demeurant la même» pendant l'intervalle de temps dt au cours duquel la longueur dx a été parcourue ; en conséquence, «la notion seule des vitesses uniforme» donne bien, immédiatement, l'expression de la vitesse v dans chaque instant : $v = dx/dt$.

Ayant caractérisé la vitesse dans chaque instant, Varignon donne, dans son mémoire du 5 juillet 1698, une «Règle générale». Celle-ci n'est qu'une reprise de sa caractérisation, sous trois aspects différents, privilégiant successivement, à la faveur de manipulations algébriques, les vitesses, les temps et les espaces :

«Règle générale Des vitesses, Des temps, Des espaces ($v = dx/dt$ ou $dt = dx/v$ ou $dx = vdt$).»

De ces trois formules, il suit bien évidemment que, «quels que soient présentement la vitesse d'un corps

(accéléérée, retardée, en un mot variée comme on voudra), l'espace parcouru et le temps employé à le parcourir, deux de ces trois choses étant données à discrétion, il sera toujours facile de trouver la troisième par le moyen de cette règle, même dans les variations de vitesses les plus bizarres qui se puissent imaginer».

La construction du concept de vitesse dans chaque instant résulte donc finalement d'un double travail conceptuel : l'un autorisant, par le dépassement du débat classique sur les grandeurs homogènes, d'exprimer le concept de vitesse sous la forme d'un quotient ; l'autre permettant de faire intervenir aisément les concepts du calcul leibnizien, par l'introduction de l'idée selon laquelle la vitesse peut être considérée comme constante dans un instant de temps. L'énoncé de la Règle générale clôt la première partie du mémoire du 5 juillet 1698. La deuxième partie s'attache, à travers cinq exemples, à la mise en œuvre de cette règle générale.

Deux mois plus tard, le 6 septembre 1698, Varignon complète cette première étude par un deuxième mémoire consacré aux mouvements qui suivent des trajectoires incurvées : «Application de la Règle générale des vitesses variées, comme on voudra, aux mouvements par toutes sortes de courbes, tant mécaniques que géométriques. D'où l'on déduit encore une nouvelle manière de démontrer les chutes isochrones dans la cycloïde renversée.»

En possession du concept de vitesse dans chaque instant, Varignon utilise le calcul leibnizien, valable tant pour les mouvements rectilignes que pour les mouvements curvilignes. Puis, en 1700, s'appuyant sur la conceptualisation de la force accélératrice proposée par Isaac Newton dans le Lemme X de la section II du Livre I des *Principes mathématiques de la philosophie naturelle*, Varignon obtient l'expression de la «force accélératrice dans chaque instant», à savoir $y = ddx/dt^2$ ou $y = dv/dt$. Il énonce alors ce qu'il nomme les Règles générales des mouvements en lignes droites (il généralisera au cas des trajectoires curvilignes peu après) : $v = dx/dt$ et $y = dv/dt = ddx/dt^2$.

Ces règles nous révèlent que les expressions des concepts de vitesse dans chaque instant et de force accélératrice dans chaque instant peuvent être déduites l'une de l'autre par un simple calcul qui met en œuvre les algorithmes

leibniziens. Pour reprendre Auguste Comte (1798-1857) : «Toutes les questions relatives à cette théorie première du mouvement varié se réduiront immédiatement à de simples recherches analytiques, qui consisteront ou dans des différentiations ou, le plus souvent, dans des intégrations.»

Les enjeux de la physique mathématique

Au début de cet article, nous nous demandions comment la conceptualisation différentielle de la science du mouvement avait favorisé la réorganisation et le développement de cette science et permis de surmonter les difficultés antérieures. Varignon apporte des réponses ou du moins des éléments de réponses.

Il apparaît clairement que la conceptualisation différentielle de la science du mouvement au début du XVIII^e siècle n'est pas le résultat d'une simple transposition, en termes différentiels, des concepts et des démarches de la science du mouvement tels qu'ils s'exprimaient auparavant ; c'est plutôt le résultat d'une reconstruction conceptuelle qui a conduit à une réorganisation du champ de la science du mouvement autour des concepts de vitesse dans chaque instant et de force accélératrice dans chaque instant.

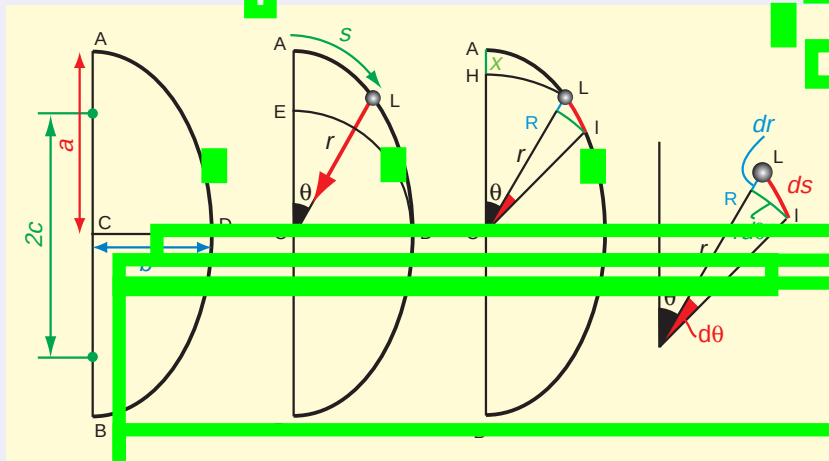
Cette reconstruction conceptuelle est riche d'enseignements : en utilisant le nouveau calcul leibnizien pour donner au concept de vitesse dans chaque instant ou en modernisant légèrement, de «vitesse instantanée», une expression explicite et opérante, Varignon montre *contrario* qu'on ne pouvait pas, avant, penser efficacement cette vitesse instantanée. En outre, et d'un point de vue plus général, le travail conceptuel de Varignon témoigne également que les mathématiques ont avec la physique un rapport de constitution, car c'est de leur couplage que se trouvent engendrés les nouveaux concepts, en l'occurrence ceux de vitesse dans chaque instant et de force accélératrice dans chaque instant.

Cette reconstruction permet alors de comprendre l'extraordinaire développement de la science du mouvement au XVIII^e siècle ; celui-ci résulte, d'une part, de la très grande généralité des deux nouveaux concepts mis en place par Varignon et, d'autre part, du fait que les expressions de ces concepts sont susceptibles de se

l'algorithme de la cinématique

Varignon calcule la force centrale y dans le cas où un corps décrit une ellipse, le centre de force étant confondu avec le centre de l'ellipse, cas qu'Isaac Newton examinait dans la proposition 1, problème 5, livre I de ses *Principes mathématiques de la philosophie naturelle*.

Varignon détermine d'abord la longueur ds d'un petit arc. Nous transcrivons ici son raisonnement et sa figure en termes modernes. Il considère une ellipse ABD de centre C, de grand axe $AB=2a$. Il définit le petit axe $CD=b$ et la distance $2c$ qui sépare les foyers. Puis il nomme p le paramètre de l'ellipse et lui donne implicitement la valeur $2b^2/a$.



La première étape du calcul est la détermination de $ds=LI$. L'abscisse curviligne s est égale à AL . Soit H le point de l'axe AC, à la distance $r=CL$ de C.

L'équation de l'ellipse est : $r^2 = b^2/(1 - e^2 \cos^2 \theta)$ où e est l'excentricité de l'ellipse égale à c/a et θ l'angle ACL. Cette équation s'écrit également : $b^2/r^2 = 1 - e^2 \cos^2 \theta$ qui donne, par différentiation : $-(b^2/r^3)dr = e^2 \cos \theta \sin \theta d\theta$. Or, d'après l'équation de l'ellipse $\cos^2 \theta = (1 - b^2/r^2)/e^2$. On en déduit, d'une part, $\cos \theta$ et, d'autre part, la valeur de $\sin^2 \theta = 1 - (r^2 - b^2)/e^2 r^2$.

Nous obtenons par substitution en prenant les valeurs positives de $\cos \theta$ et de $\sin \theta$: $-(b^2/r^3)dr = \sqrt{(r^2 - b^2)(e^2 r^2 - r^2 + b^2)}d\theta$. Par ailleurs : $dz = r d\theta = RI$.

En conséquence : $dz = b^2 dr / (r^2 - b^2) \sqrt{(e^2 r^2 - r^2 + b^2)}$.

Il ne reste plus qu'à exprimer ce résultat en fonction du paramètre $p = 2b^2/a$ et non plus de b et a : $dz = p^2 dr / (r^2 - p^2) \sqrt{(e^2 r^2 - r^2 + p^2)}$.

Nous pouvons maintenant obtenir une expression différentielle caractéristique d'un petit arc de trajectoire : $dz^2 = dr^2 + r d\theta^2$:

$$dz^2 = [(p^2 r^2 - 2p^2 r^2 + p^2 r^2) / (r^2 - p^2) \sqrt{(e^2 r^2 - r^2 + p^2)}]^2 dr^2 + r^2 d\theta^2$$

Calcul de la force centrale y

Il s'agit maintenant, par l'intermédiaire de simples manipulations algébriques-différentielles, de faire apparaître l'expression générale de la force centrale y dans la mesure où le mouvement est assujéti à la loi des aires.

Nous avons : $ds^2/dz^2 = 2a^2r^2 - 2r^4 + ap^2/a^3p$.

Soit, également : $2a^2 - 2r^2 + ap/a^3p = ds^2/r^2 dz^2$.

Or, d'après la loi des aires, «on aura partout ici suivant M. Newton (Phil. Nat. Princ. Math. lib. I, sect. 2, th. 1) les dt comme les rdz , ou $dt = rdz$ ».

En conséquence, l'expression précédente peut s'écrire :

$$ds^2/dt^2 = 2a^2 - 2r^2 + ap/a^3p$$

Varignon différencie alors les deux expressions : en considérant dt comme une constante : $2dsd(ds)/dt^2 = -4rdr/a^3p$, mais $dr = -dx$, d'où $4r/a^3p = 2dsd(ds)/dx dt^2$. Il fait ainsi apparaître, à l'issue de son calcul, et seulement d'un calcul, l'expression générale, trouvée précédemment, de la force centrale y , d'où $y = 4r/a^3p$. Celle-ci est donc bien comme la distance $CL(r)$ du centre C de l'ellipse «au corps L qui la décrit».

Varignon déduit alors de ce résultat ceux qui concernent la parabole et l'hyperbole.

déduire l'une de l'autre par le jeu de simples procédures bien réglées de différentiation et d'intégration. La science du mouvement devient algorithmique.

L'œuvre de Varignon est trop souvent oubliée ou, au mieux, limitée à ses contributions à la statique (le polygone de Varignon permet, notamment, de décrire les forces qui s'exercent dans des systèmes mécaniques à l'équilibre). Ce manque d'intérêt pour la plus grande partie de l'œuvre de Varignon s'explique simplement par le fait que la plupart des historiens de la mécanique classique ne voyaient, dans les mathématiques, qu'un instrument ou qu'un outil au service de la science, et privilégiaient l'analyse historique des concepts ou des principes de la mécanique ; ce faisant, ils négligèrent de penser la science du mouvement en tant que physique mathématique, c'est-à-dire qu'ils négligèrent le plus souvent de reconnaître son plein effet à l'action proprement mathématique de la physique mathématique dans la coordination, la cohérence et la genèse conceptuelle.

C'est précisément sur ce terrain que se situe l'effort de Varignon. À cet égard, n'est-il pas remarquable de constater que, si ses formules et règles générales émettent et rassemblent bien divers résultats, elles sont aussi et surtout des points de départ pour de nouvelles investigations ? Avec Varignon, la physique mathématique, qui avait reçu son impulsion de Galilée, de Christian Huygens (1629-1695) et de Newton, prend plus nettement conscience de sa valeur inductive spécifique.

En ces sens, on peut dire que l'œuvre de Varignon, caractérisée par son nouveau style analytique orienté vers l'obtention d'ensembles de lois ou de règles régies par des procédures algorithmiques, style qui culminera, en 1788, avec la mécanique analytique de Joseph Louis Lagrange, aura finalement ouvert la voie à ce qui sera la physique mathématique des XVIII^e et XIX^e siècles.

Michel BLAY est directeur de recherche au CNRS et directeur adjoint chargé de la recherche à l'École normale supérieure Lettres et Sciences humaines.

M. BLAY, *La Naissance de la mécanique analytique*, Presses universitaires de France, 1992.

M. BLAY, *Les Principia de Newton*, Presses universitaires de France, 1995.

Histoire d'infini

MARIA REMÉNYI

Pendant 2 000 ans, le concept d'infini gêna les mathématiciens. John Wallis posa le symbole ∞ , les libérant ainsi des discussions philosophiques.



En Europe, au XVII^e siècle, est né le calcul différentiel et son astucieux formalisme autorisant la manipulation des quantités infinitésimales. L'inverse de l'infiniment petit étant l'infiniment grand, il fallait que le même siècle voie l'apparition du symbole ∞ . C'est le mathématicien britannique John Wallis (1616–1703) qui, le premier, abrégea le concept «infini» par ce symbole. John Wallis a largement contribué au développement des mathématiques de son époque, tant dans leur contenu que dans leur forme.

Les sciences en avaient bien besoin de ce progrès conceptuel pour progresser. Dans le signe inventé par Wallis, la polémique sur le sens à donner au mot «infini» durerait peut-être encore. En fait, mathématiciens, philosophes et théologiens polémiquaient depuis deux mille ans. Une situation qui n'avait pas que des inconvénients. Ainsi, la littérature mathématique des siècles passés était, du moins de prime abord, plus accessible que de nos jours : ses auteurs se servaient de la langue écrite usuelle pour la description des objets et des relations mathématiques. Le langage moderne des formules, qui, pour le nombreux profanes, représente un obstacle insurmontable et contribue de ce fait à l'impopularité générale de la discipline mathématique, n'est pas dû à quelque nécessité alchimique de conservation des secrets mathématiques. L'histoire des sciences montre

que le désir de résoudre des problèmes concrets favorise l'introduction de symboles et d'abréviations, afin de mieux formuler les hypothèses et les démonstrations et d'obtenir des conclusions précises.

L'une des difficultés consistait à découvrir tout ce que cache le mot infini. Aristote (384 à –322) fit les premiers progrès. Le mot *apeiron* qu'utilisait le philosophe grec Aristote se caractérisait déjà, à l'époque présocratique, par des sens multiples. Il signifiait sans limite, indéfini, inconcevablement grand, mais désignait également, avec une connotation négative, le chaos duquel est né le monde. Ce n'est qu'au début de notre ère que «l'infini» fut identifié à l'«Un» divin. Pour Aristote, le mot infini était associé à l'expression de l'imperfection.

La précision qu'Aristote apporta à l'emploi de ce mot distinguant, entre infini actuel et infini potentiel, fut d'une importance décisive pour toutes les discussions mathématiques futures. Un processus qui poursuit sans fin, comme par exemple la production de nombres naturels par la consigne «toujours pour suivre la numération», n'est que potentiellement infini, car à chaque stade du processus il n'existe qu'un nombre fini d'objets. À l'inverse, un ensemble de taille infinie qui existe une fois pour toutes, correspond à un infini actuel.

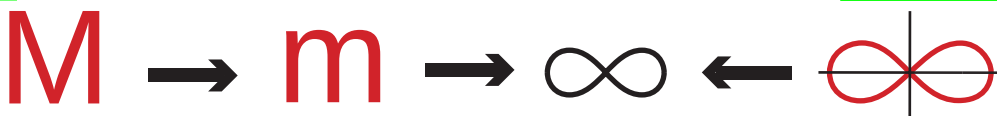
Notons que la mathématique grecque utilisait des méthodes infinitésimales cachées (méthode d'exhaus-

sion) pour le calcul des longueurs des courbes ou des surfaces délimitées par des lignes courbes, mais l'infini n'apparaissait pas en tant que terme mathématique. Euclide, par exemple, ne dit pas : «il existe une quantité infinie de nombres premiers», mais «il existe une quantité de nombres premiers supérieure à toute quantité donnée».

Réflexion métaphysique

Au Moyen Âge, ère obscure pour les sciences, les réflexions sur la nature de l'infini étaient essentiellement métaphysiques. Certaines peuvent toutefois être interprétées comme la préparation des considérations mathématiques du calcul infinitésimal au XVII^e siècle. C'est le cas des réflexions sur l'infini actuel et potentiel du théologien Jean de Nicolas de Cues (1401–1464). Johannes Kepler (1571–1630) s'en inspira, par exemple, lorsqu'il abandonna les méthodes antiques pour le calcul des surfaces.

Francesco Bonaventura Cavalieri (1598–1640), au XVII^e siècle, fut le premier à adopter une position considérée comme moderne sur l'infini, à savoir que des lois différentes gouvernent les grandeurs infinies et les grandeurs finies. Dans sa méthode des indivisibles, il se préoccupa moins de spéculations philosophiques que de trouver une façon de résoudre son problème. Pour y arriver, il contourna astucieusement les difficultés qui se



1. Le signe infini aurait été créé au XVII^e siècle par le mathématicien britannique John Wallis à partir d'une version cursive du m latin (ou M),

c'est-à-dire de 1 000 en chiffres romains. Peut-être a-t-il aussi pensé à la courbe de même forme (lemniscate), qui se parcourt sans fin?

présentaient lors du calcul de la somme d'un nombre infini de grandeurs.

Entre autres contributions, John Wallis poursuivit le développement des méthodes de Cavalieri. Après ses études supérieures, qui ne furent nullement marquées par une intense activité mathématique, il fut ordonné prêtre à Londres. Pendant cette période, il participa aux préparatifs de la fondation de la *Royal Society*. En 1643, il s'était illustré durant la guerre civile qui opposait royalistes et parlementaires, en décryptant très brillamment des informations secrètes. Lorsque Wallis fut nommé en 1649 professeur de géométrie à Oxford, ce n'était pas en raison de ses performances mathématiques, c'était probablement avant tout le résultat de ses activités politiques.

Cependant, l'estime dans laquelle Wallis était tenu dans le monde scientifique était méritée : il est aujourd'hui considéré comme l'un de ceux qui ont ouvert la voie au calcul infinitésimal et comme le précurseur le plus éminent de Newton, qui fut influencé par l'ouvrage majeur de Wallis, *Arithmetica Infinitorum* (1656). Au paravant, ses travaux avaient porté sur les sections coniques (*De sectionibus conicis*, 1655), où il délaissa l'approche purement géométrique de la mathématique grecque pour considérer les sections coniques comme des courbes lisses, dont il établit les équations algébriques. Il déduisit de ces équations les propriétés des sections coniques sans faire appel à des arguments géométriques, mais uniquement par le calcul. Il participa ainsi à un développement essentiel de l'histoire des mathématiques : l'algébrisation de la géométrie.

Dès le début de ces travaux, Wallis proposa une modification de la méthode des indivisibles de Cavalieri. Alors que Cavalieri divisait les surfaces en un nombre infini de segments, Wallis les décomposait en un nombre infini de parallélogrammes de même taille, cette taille étant «une fraction infiniment petite, égale à $1/\infty$ de la surface totale et le symbole ∞ désignant l'infini».

De M au signe infini

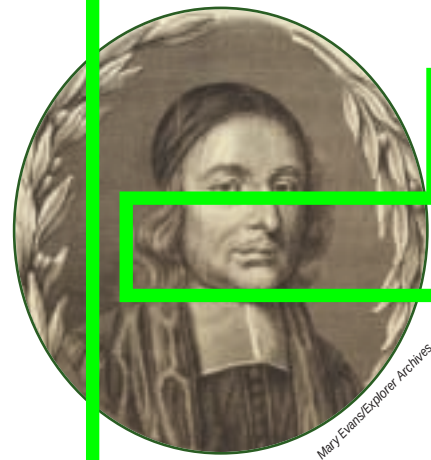
Pourquoi choisit-il cette représentation ? On ne peut émettre une opinion sur cette question qu'une réponse spéculative. Bien avant le début de ses activités mathématiques, Wallis s'était fait un nom en tant que philologue ; il savait probablement que le symbole ∞ désignait une ligature latine de m apparue

au VII^e siècle, c'est-à-dire une écriture cursive utilisée pour signifier 1 000 en chiffres romains, donc un «très grand» nombre. Un contemporain de Wallis, le mathématicien et philosophe hollandais Bernhard Nieuwentijt, utilise pour sa part le signe «m» dans son *Analysis infinitorum* de 1695. Le signe m, un symbole proche du futur ∞ , y désigne l'infini. Sans doute Wallis a-t-il aussi pensé que la boucle que représente le symbole ∞ faisait penser à l'infini, puisqu'elle peut être parcourue sans fin.

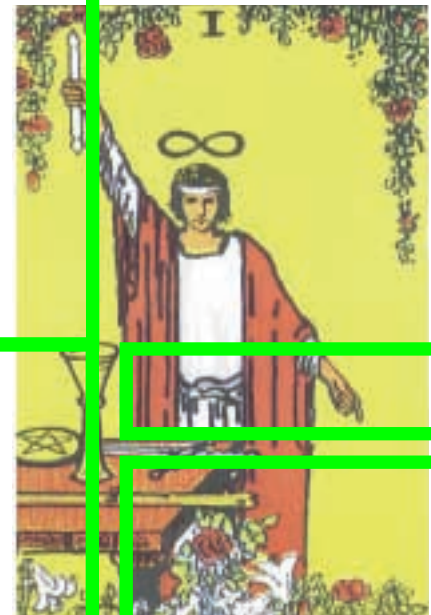
L'apparition du symbole ∞ contribua en tout cas fortement à la modernisation en marche des mathématiques. Au tournant du XVIII^e siècle, Isaac Newton (1643-1727) et Gottfried Wilhelm Leibniz (1646-1716) organisèrent les connaissances éparses sur le calcul différentiel et intégral en une véritable théorie. Tout était prêt pour que le symbole soit de plus en plus utilisé. Au début du XVII^e siècle, il apparut de plus en plus fréquemment dans la littérature mathématique et philosophique ; il était souvent lié au concept de l'infiniment petit, car la signification et l'admissibilité de cette notion dans le calcul infinitésimal alors en cours de développement posaient un problème tant mathématique que philosophique. Avec les travaux de Léonard Euler (1707-1783), le signe ∞ devint, dans la seconde moitié du XVIII^e siècle, un élément permanent du langage symbolique mathématique. Euler, contrairement à Newton et à Leibniz, rejetait complètement les légitimations métaphysiques des grandeurs infiniment petites et adoptait enfin un point de vue strictement formel.

La théorie des grandeurs infinitésimales, dont les bases restaient vacillantes, fut largement améliorée au cours du XIX^e siècle grâce à l'introduction du processus de passage à la limite, et les notions de continuité et de convergence. Une théorie du calcul différentiel et infinitésimal rigoureuse et qui répondait aux exigences formelles et logiques des mathématiques apparaissait enfin.

Au XIX^e siècle, le symbole ∞ servait surtout comme aujourd'hui d'ailleurs, à décrire les processus de passage à la limite, notamment pour les séries et les limites. Il décrivait, au sens d'Aristote, un infini potentiel. Toutefois, Georg Cantor (1845-1918), le fondateur de la théorie des ensembles, et par conséquent de la théorie moderne de l'infini, préféra différencier par des symboles différents les deux types d'infini. Il dési-



2. Le Britannique John Wallis contribua à l'algébrisation de la géométrie. Il écrivit, le premier, le signe ∞ pour décrire des parallélogrammes infiniment petits.



3. Le signe ∞ fut vite employé en dehors des mathématiques pour symboliser l'infini. Ici une carte d'un ancien jeu de tarots.

gna le premier nombre transfini (infini actuel) par la lettre hébraïque *aleph*. L'emploi de ce nouveau signe pour signifier l'infini ne paraît arbitraire qu'au premier abord, car, dès les années 1700, le symbole ∞ fut aussi employé en dehors des mathématiques et de la philosophie, afin de symboliser l'infinité ou l'éternité. Il fit, par exemple, son apparition sur la carte de tarot, appelée le Menteur ou le Mage. Or, le symbole kabbalistique qui fut associé à cette carte est la lettre hébraïque *aleph*.

Maria REMÉNYI, journaliste, est historienne des sciences.

De la perspective à l'infini géométrique

JEAN-PIERRE LE GOFF

Au XVII^e siècle, grâce à la géométrie projective, née de l'union de la théorie de la perspective et de la théorie des coniques, l'infini géométrique acquiert une «dimension» humaine, une évidence perceptible. L'infini potentiel des philosophes devient l'infini actuel des géomètres.



The Bridgman Art Library

Pour les philosophes et pour les géomètres de la Grèce antique, l'infini n'a pas de véritable existence. Il est considéré comme «potentiel», accessible en puissance. Ainsi, Aristote impose une vision géocentrique et «finitiste» du cosmos, limité par la sphère des fixes, sorte de voûte céleste qui porte les étoiles et marque les confins du monde avec le néant extérieur. Dans ce contexte, la géométrie euclidienne reconnaît et intègre l'infini potentiel, mais s'arrête aux rives de l'infini «actuel». L'infini inspire alors aux géomètres grecs méfiance et circonspection, car son usage immodéré ou inattentif conduit à des contradictions.

Un objet élémentaire tiré de la géométrie d'Euclide, une «ligne droite», illustre ces contradictions. Euclide la définit comme l'extension la plus «économique» d'un point à un autre. Aussi, une telle ligne s'apparente plutôt à notre moderne «segment» qu'à ce qu'un mathématicien d'aujourd'hui nomme une droite – à ceci près que, pour Euclide et ses successeurs, jusqu'au XVII^e siècle, cette «ligne» est prolongeable autant que nécessaire pour les besoins d'une construction ou d'une démonstration (dans une figure où il est nécessaire de

déterminer le point de concours de deux droites non parallèles, par exemple). De même, une «surface plane» est en fait une figure plane limitée par un contour fermé ou un segment de parabole, par exemple, et est limitée par la courbe et par l'une de ses cordes. Il en est de même d'un corps ou d'un solide, volume fini limité par une ou par plusieurs surfaces.

Ainsi la ligne droite, le plan ou le solide des géomètres strictement euclidiens sont fins, mais prolongeables à volonté, et non des supports rectilignes, plans ou volumiques d'un espace conçu d'emblée comme infiniment étendu.

Du fini prolongeable à l'infini

Depuis cette époque, on débat du statut de ce qu'on nomme le cinquième postulat d'Euclide, au premier Livre de ses *Éléments*. Ce postulat – une proposition non démontrée – énonce en substance que, si une sécante rencontre deux autres droites en faisant des angles internes situés d'un même côté de la sécante tels que leur somme soit inférieure à deux angles droits, les deux droites, prolongées indéfiniment (c'est-à-dire autant que nécessaire) se

rencontreront du côté où se trouvent les angles de somme inférieure à deux angles droits. Un équivalent moderne de ce postulat consiste à dire que, par un point hors d'une droite, il passe une parallèle à cette droite et une seule. Les géomètres ont soulevé la question suivante : ne pouvait-on pas démontrer le cinquième postulat à partir des autres postulats et propositions des *Éléments*?

Les premières tentatives de démonstration reposaient sur le prolongement indéfini des droites, c'est-à-dire leur prolongement potentiellement infini avec conservation de l'alignement.

On sait aujourd'hui que ces tentatives étaient vaines : le cinquième postulat d'Euclide ne peut être démontré. On sait aussi que cette aventure des parallèles conduisit à la création de trois types de géométries. L'une, dite euclidienne, admet ce postulat comme un axiome de la théorie. Les deux autres, dites non euclidiennes, ont été conçues à la fin du XVIII^e siècle et posent comme

axiome l'une des deux formes que peut prendre la négation du postulat des parallèles. La géométrie hyperbolique prend pour principe que par un point il passe plus d'une parallèle à une droite donnée (et, par voie de conséquence, on

démontre qu'il en passe une infinité). La géométrie elliptique, quant à elle, postule que par un tel point il ne passe aucune parallèle à une droite donnée.

Ce qu'il on sait moins, c'est que l'émergence des géométries non euclidiennes n'est devenue possible qu'après celle de la géométrie projective, qui devait servir de cadre général aux trois modèles, et que cette géométrie projective est née de la conjonction, au XVII^e siècle, de la science de la perspective, issue de la géométrie pratique, et de la théorie des coniques d'Euclide et surtout d'Apolonios de Perga (262?-180?). En somme, l'invention de la perspective centrale et ses prolongements géométriques ont rendues possibles l'irruption et la considération serene des éléments à l'infini dans la géométrie euclidienne (voir la figure 2), puis l'invention d'un cadre géométrique où l'on a pu envisager d'autres systèmes d'axiome jusque dans leurs conséquences ultimes.

L'évolution des idées sur l'infini géométrique est passée par la question de la perspective, c'est-à-dire la question de la représentation de la troisième dimension dans un plan. Un moment à la fois crucial et fondateur de cette question fut l'invention, à la Renaissance, de la «perspective linéaire».

La perspective centrale, invention du Quattrocento

Au travers de celle-ci, l'infini géométrique actuel, l'espace actuellement infiniment étendu, a trouvé une première représentation dans le dessin, pictural ou technique, et ce avant même d'avoir été pensé : il s'y trouve pointé, voire désigné sur l'horizon, mais aussi, très souvent, masqué par un mur ou par une porte.

La perspective linéaire consiste à projeter sur le plan du tableau les rayons rectilignes et imaginaires allant de l'objet à représenter jusqu'à l'œil, supposé ponctuel, du peintre ou du spectateur (voir l'encadré de la page 70). Elle est aussi nommée perspective centrale ou projection conique par les mathématiciens : l'œil constitue le centre O de cette projection, et tous les rayons passant par O et issus des points situés sur le contour de l'objet forment alors un cône de sommet O.

Dans la perspective linéaire, un ensemble de droites parallèles de l'espace a pour projection, sur le plan du tableau, un ensemble de droites parfois parallèles (lorsqu'elles sont parallèles au tableau), mais en général concourantes. Prenons, par

exemple, un faisceau de droites parallèles et perpendiculaires au plan du tableau, tracées sur le sol : les images de ces droites forment, sur le tableau, un ensemble de droites qui concourent en un point, le «point de fuite central» défini comme tel vers 1435 par Leon Battista Alberti (1404-1472). Ce sont les peintres ou les architectes de la Renaissance italienne du XV^e siècle – le Quattrocento –, tels que Filippo Brunelleschi (1377-1446), Alberti et Piero della Francesca (1416?-1492), qui ont codifié les règles de la construction légitime du dessin en perspective. Ils ont ainsi inauguré une pratique conventionnelle et géométrisée de la perspective, mais dont la grande adéquation avec la vision humaine, d'une part, et les conséquences pratiques, d'autre part, devaient assurer la postérité.

C'est ainsi, par la définition et l'usage dans le dessin d'un point de concours de droites en réalité parallèles, les artistes italiens de la Renaissance ont donné le premier exemple de représentation visuelle d'un infini actuel : le point de fuite figure sur le tableau un point en réalité situé à l'infini, là où les droites parallèles «se rencontrent». Il ne s'agit pas de dire ici que ces artistes



1. CETTE ANNONCIATION, conservée au Musée des Beaux-Arts de Caen, est due au peintre maniériste Pâris Bordone (1500-1571). Une analyse géométrique de ce tableau montre qu'il illustre le procédé singulier des «vues de l'en deçà» données par cet artiste : il ne peint pas ce qu'il voit dans le cadre fictif dont il entoure la scène représentée et dont on peut supposer que la base sert de ligne de terre,

mais aussi ce qu'il voit «en deçà» du cadre, devant le cadre, de son propre côté. Ainsi, tandis que l'ange fait l'annonce à Marie, les trajectoires visuelles des deux protagonistes ne peuvent se croiser. L'ange est au-delà de la fenêtre, tandis que la Vierge est située dans l'en deçà, du côté du peintre, sur un plan plus élevé que celui de l'observateur.

avaient pour projet de mettre en évidence la possibilité géométrique de penser l'infini actuel, ou qu'ils aient eu une quelconque conscience d'avoir participé à son invention. Les conceptions géométriques des théoriciens ou des praticiens de la Renaissance étaient encore profondément marquées par les idées finitistes des autorités antiques en la matière. Étaient encore à venir des notions telles que la droite infinie, le plan infini, ou l'espace antérieur à la matière qui l'habite.

Néanmoins, il est incontestable que les œuvres d'art, bientôt relayées par la gravure et par l'imprimerie, donneront à voir une multitude d'occurrences de l'infini représenté. En effet, le point de fuite central sera bientôt flanqué de points de distance, dont on s'avisa ensuite que ce sont aussi des points de fuite, donc des représentations de l'infini (voir la figure 4). C'est encore l'horizon (la ligne horizontale sur le tableau, située à même hauteur que l'œil du spectateur) qui même qui apparaîtra comme lieu d'une multitude – d'une infinité – de points de fuite; puis, vers 1600, viendra enfin l'idée que tout point du tableau est, en dernière instance, un point de fuite.

Desargues, le parallélisme et le concours à l'infini

Cette émergence d'une multiplicité de représentations visuelles de l'infini actuel devait conduire, au début du XVII^e siècle, en France, à une révolution dans le champ conceptuel de la géométrie elle-même. Girard Desargues (1591-1661), un géomètre lyonnais, va en effet théoriser ce que tout pratiquant de la perspective linéaire a vu sans toujours en croire ses yeux : il identifie le concours de droites ou de plans au parallélisme de ces objets. En d'autres termes, le parallélisme n'est rien d'autre qu'un concours à l'infini, des droites parallèles ne sont rien d'autre que des droites qui se rencontrent en un point situé à l'infini. Cette idée, Desargues la décline dans son *Brouillon projet d'une atteinte aux événements des rencontres du cône avec un plan*, c'est-à-dire un «essai pour atteindre à la compréhension synthétique des diverses occurrences de courbes coniques obtenues par intersection plane d'un cône à

base circulaire», imprimé à 50 exemplaires en 1639. Cette idée rompt avec la conception finitiste, héritée des Grecs, que l'œil se faisait de la ligne droite, de la surface ou figure plane, ou du solide.

Desargues a derrière lui deux siècles de pratique de la perspective, qui ont contribué à éduquer l'œil du spectateur occidental. En particulier, il est deux points qu'il convient de préciser pour comprendre les idées du géomètre. Le premier, c'est que le cadre dans lequel l'artiste cherche à figurer une scène – ce qu'Alberti nommait une «fenêtre» – était conçu, dans un premier temps, comme une ouverture sur un au-delà à représenter. Le second, c'est que cette fenêtre était en principe disposée verticalement, sur un géométral horizontal. Or les artistes ont très vite transgressé ces deux contraintes initiales.

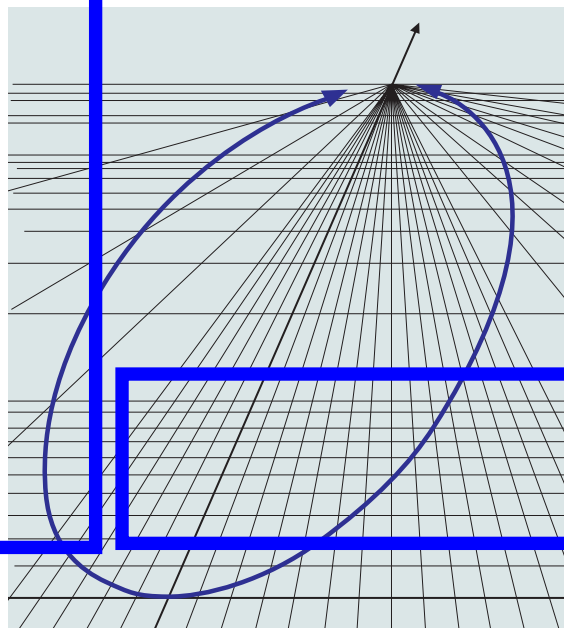
D'une part, ils se sont autorisés à représenter quelques éléments de l'en-deçà de la fenêtre : timidement, d'abord, en trompe-l'œil, ils ont dessiné le cadre du tableau ou l'encadrement de ladite fenêtre, prise alors au sens propre, ou tel objet posé sur le rebord d'une ouverture feinte (synonyme alors de peinture). Ainsi, dans le *Saint-Jérôme en son cabinet* d'Antonioello de Messine, un oiseau est perché sur le bord d'une fenêtre au-delà de laquelle on aperçoit le saint à son bureau. Un autre exemple, plus tardif

et contemporain de Desargues, est un portrait de Philippe de Champaigne (1602-1674) qui montre le personnage, la main négligemment posée sur le bord du cadre. Des exemples tout à fait singuliers de ce que nous appellerons ici des «vues de l'en-deçà» ont été créés au milieu du XVI^e siècle par un peintre vénitien maniériste, Paris Bordone (voir la figure 1).

D'autre part, les peintres se sont abstraits de la contrainte du fil à plomb et du niveau, comme en témoignent les décors plafonnants (tel *L'Oculus de la chambre des époux* d'Andrea Mantegna, voir la figure 7), ou plus tard les décors baroques de la Contre-Réforme jésuite (ceux d'Andrea Pozzo, par exemple) et enfin les trompe-l'œil sur voûtes ou toutes sortes de surfaces irrégulières (théorisés au début du XVII^e siècle en France par le jésuite Jean Dubreuil ou par le graveur Abraham Bosse, maître de Desargues). Ces œuvres indiquent qu'un tableau, voire une fresque, n'est pas nécessairement situé dans un plan orthogonal au sol.

De telles images ne pouvaient qu'imprimer avec force l'idée que la fenêtre ne constitue pas une borne et que cette fenêtre n'est pas nécessairement d'aplomb sur le géométral. Toutefois, il restait un pas à franchir. Desargues le fera, non dans le domaine de la représentation, mais dans celui de la conception : quand un rayon visuel théorique – conçu comme une demi-droite potentielle depuis Euclide et son *Optique* – s'étend à l'infini comme une ligne droite, pourquoi s'arrêterait-il à l'œil? Pour Desargues, un cône proprement dit est un cône à deux «nappes», obtenu en prolongeant en droites entières les demi-droites issues de son sommet. De façon analogue, on peut prolonger les rayons visuels en arrière de l'œil; celui-ci est aveugle dans ces directions, mais cela n'empêche pas d'imaginer cet espace arrière. Desargues repensa alors les

«coniques», courbes obtenues en coupant un cône à base circulaire par un plan. Ces courbes – l'ellipse, la parabole et l'hyperbole – suscitaient un intérêt renouvelé, de Johannes Kepler (1571-1630) et ses trajectoires planétaires elliptiques à Athanasius Kircher (1601-1680) et ses



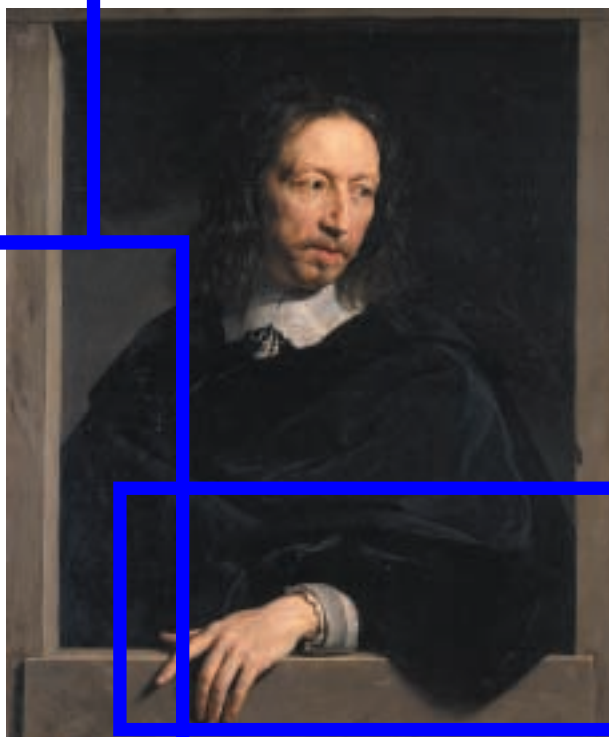
2. LES ÉLÉMENTS À L'INFINI ont longtemps été considérés avec circonspection. Aujourd'hui, on sait qu'une parabole, courbe ouverte, peut devenir, en perspective, une ellipse, courbe fermée, les branches infinies se rejoignant en un même point de l'horizon qui représente, à distance finie, la droite à l'infini commune aux plans horizontaux; ce point à l'infini se retrouve donc à distance finie de l'observateur, dans le tableau.

miroirs paraboliques, en passant par Descartes et sa classification algébrique des courbes du second degré.

Considérons les rayons visuels qui joignent un œil ponctuel aux points d'un cercle réel. Ces rayons forment un cône à base circulaire dont le sommet est l'œil. Toute coupe du cône par un plan produit une courbe qui est une perspective du cercle. Autrement dit, Desargues a pris conscience qu'une conique n'est jamais qu'un cercle, vu d'un bon point de vue.

Cela n'avait jamais été énoncé en ces termes auparavant. On concevait bien qu'une ellipse, courbe fermée, pouvait être l'image d'un cercle en perspective. Par exemple, Albrecht Dürer (1471-1528) a construit point par point des coupes d'un cône droit limité, dans les trois situations de coupe possibles (perpendiculaire à l'axe vertical, ce qui donne un cercle, parallèle à cet axe, ce qui donne une hyperbole, parallèle à une génératrice, et l'on obtient une parabole, et quelconque, ce qui donne une ellipse). Notons que son dessin d'ellipse a une forme ovoïde, comme s'il n'en croyait pas ses yeux, s'attendant plutôt à une courbe effilée près du sommet du cône plutôt qu'à une véritable ellipse. Pourtant, il paraissait inconcevable que l'on puisse assimiler au cercle des courbes non fermées, ayant des branches infinies, comme la parabole, ou présentant des ruptures de continuité, comme l'hyperbole à deux parties séparées, telle que nous la définissons aujourd'hui.

Kepler, dans son *Optique*, avait certes fait remarquer qu'une parabole n'est jamais qu'une ellipse dont l'un des foyers se serait éloigné à l'infini, la courbe ouvrant ses bras en quelque sorte, puis devenant une hyperbole lorsque cette ouverture se trouve limitée par deux droites asymptotes, dans le mouvement tournant que ferait un plan coupant un cône (illimité, mais à une nappe). Cette conception novatrice, réintégrant le mouvement dans la définition des courbes, a sans doute inspiré Desargues, mais le géomètre lyonnais est allé plus loin : il a dépassé l'infini potentiel de Kepler et a posé



3. UNE SCÈNE est délimitée, au XVI^e siècle par une «fenêtre» par laquelle on l'observe. Progressivement, les artistes s'en affranchissent et en jouent : ici la fenêtre apparaît sous forme d'un cadre : Philippe de Champaigne peint Robert Arnauld d'Andilly avec une main négligemment posée sur le bord du cadre. Notons que Philippe de Champaigne demanda à Desargues de dessiner pour lui le schéma d'un décor plafonnant pour la voûte de l'ancienne église des carmélites : le Christ en croix, flanqué de la Vierge et de saint Jean, paraissait être sur un plan vertical (en contre-plongée), alors qu'il était peint sur un plan horizontal.

d'emboîter le cône comme surface infinie à deux nappes.

Plus encore, en assimilant faisceaux de droites parallèles et faisceaux de droites concourantes, Desargues fut amené à concevoir qu'une droite n'est jamais qu'un cercle de rayon infini, c'est-à-dire dont le centre, une fois fixé l'un des points du cercle, s'en est éloigné indéfiniment. Et qu'un cylindre à base circulaire est un cône dont le sommet se trouve à l'infini – ce qui explique que les coupes d'un cylindre par un plan produisent aussi des coniques, en l'occurrence des ellipses ou des cercles.

Bassin circulaire, ellipse, parabole ou hyperbole

Pour mieux comprendre comment, dans le contexte de la perspective, les diverses coniques apparaissent comme les projections du cercle, imaginons une situation picturale banale : un peintre qui observe un bassin circulaire au travers de la «fenêtre» fictive où il souhaite le représenter, le tableau étant orthogonal au sol (voir la figure 6). On pourrait faire appa-

raître les diverses coniques en faisant varier l'inclinaison du tableau par rapport au sol. Modifions plutôt les positions relatives de l'œil, du tableau et du bassin.

Quand le bassin se situe au-delà de la fenêtre, son image est une ellipse ou un cercle, dans ce que l'on nomme, depuis Apollonios, le cas d'antiparallélisme. Rapprochons le bassin de l'observateur, de façon qu'une partie au moins du bassin se retrouve entre l'œil et le tableau : la projection du bassin sur le plan du tableau – un plan supposé illimité – reste une courbe fermée (une ellipse) dont une partie apparaît pas dans le tableau, puisqu'elle est située sous la ligne du terre. Rapprochons

le bassin de l'observateur, jusqu'au moment où celui-ci se retrouve debout sur son contour ; dans ce cas, le rayon visuel partant de l'œil et rejoignant son point d'aplomb est parallèle au tableau et ne le rencontrera jamais, ou plutôt à l'infini. Ce point du cercle ayant une image à l'infini, l'image du bassin circulaire devient une parabole (géométriquement, cette situation correspond à un plan de coupe parallèle à une génératrice du cône).

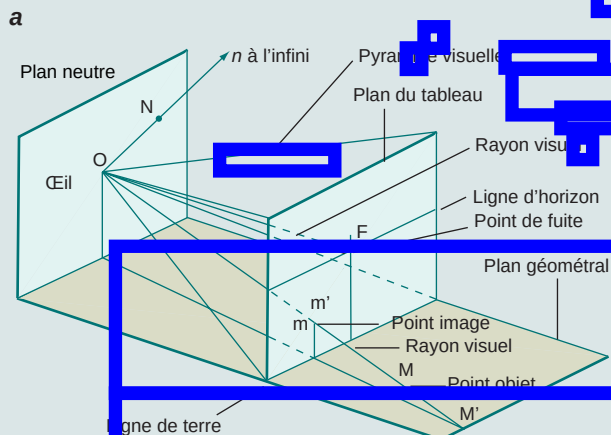
Que se passe-t-il enfin quand l'observateur met les pieds dans l'eau, de sorte que le bassin s'étend aussi derrière lui ? Le «plan neutre» – le plan parallèle au tableau et passant par l'observateur – coupe alors le cercle en deux points A et B. L'image sur le plan du tableau de chacun de ces points étant à l'infini, la partie du bassin circulaire qui s'étend devant l'observateur produit une courbe-image possédant deux branches infinies, c'est-à-dire une hyperbole. Plus précisément, il s'agit d'une hyperbole telle que l'avait définie Apollonios, c'est-à-dire une seule des deux courbes séparées qui composent une hyperbole dans sa définition moderne. Quant à l'arc du bassin qui se trouve derrière l'observateur, celui-ci ne le voit pas, mais cela n'empêche pas le géomètre de concevoir sa projection sur le plan du tableau. Celle-ci est obtenue en prolongeant jusqu'au plan

Le principe de la perspective centrale

Une perspective centrale, dite aussi linéaire, est, géométriquement parlant, une projection de l'espace dans un plan (P), menée à partir d'un point fixe O extérieur au plan (P). Chaque point M de l'espace a pour image le point m, lorsqu'il existe, intersection de la droite (OM) et du plan (P).

Le point O (voir la figure a) représente le centre de projection, c'est-à-dire, dans la pratique du dessin, un œil ponctuel, situé au-dessus du plan au sol, nommé géométral. Toutes les demi-droites issues de O – les rayons visuels – forment une pyramide ou un cône dits visuels, s'appuyant sur le cadre du tableau rectangulaire ou en forme de tondo (rond), qui est une partie du plan de projection (P).

Le tableau et le géométral se coupent selon la ligne dite de terre. Le rayon visuel principal est perpendiculaire au tableau et y détermine un point F, le point de fuite principal : ce point F détermine en fait le niveau d'une ligne parallèle à la ligne de terre passant par lui, que l'on nomme la ligne d'horizon. Cette ligne est aussi l'intersection du tableau avec le plan horizontal principal, parallèle au géométral et passant par l'œil. Enfin, le plan parallèle au tableau passant par l'œil O, qui contient tous les rayons visuels parallèles au tableau, est nommé plan neutre.

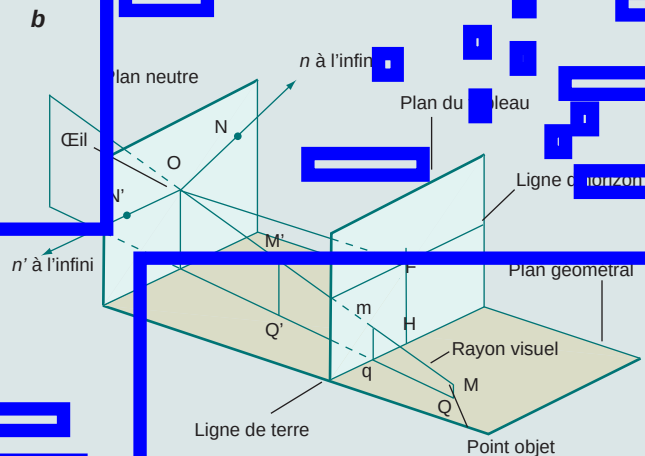


Un point M situé au-delà de la fenêtre aura pour image perspective ou pour apparence un point m du tableau par intersection du rayon visuel (OM) avec le tableau (voir la figure b) ; un point M' ou M'' aligné avec O et M aura la même image que M dans le tableau ($m' = m$).

La perspective centrale primitive ne considère que l'au-delà de la fenêtre, tandis que la projection conique moderne considère tous les points de l'espace, situés sur toutes les droites passant par O. Dans ce cas, les rayons visuels traversent l'œil et «cherchent» aussi tous les points situés en deçà du tableau : ceux-ci peuvent alors être situés entre l'œil et le tableau (un cadre en trompe-l'œil, par exemple, ou un objet hors cadre comme on en verra apparaître au début du XVI^e siècle) ; ou bien l'œil O, devenu purement géométrique, peut être interposé entre l'objet visé et le tableau ; dans ce dernier cas, le perspecteur aurait alors des idées derrière la tête, en quelque sorte, ce qui n'a, à notre connaissance, jamais été exploité par les peintres, du moins en termes géométriques.

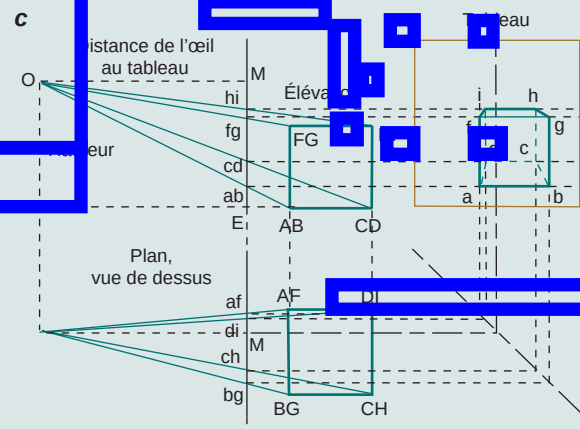
Notons enfin qu'un point N ou N' du plan neutre n'a pas d'image n ou n' à distance finie dans le tableau ; en géométrie projective,

on dira que l'horizon de la projection des plans parallèles au tableau est la droite à l'infini du tableau.



Le problème du dessin en perspective consiste, pour chaque point M de la scène à représenter, à déterminer le point m correspondant dans le tableau. Pour obtenir cette image, il existe plusieurs procédés, dont celui, dit aujourd'hui de la double projection, qui permet de repérer, au sol, l'écart qH de la projection q de m par rapport à la ligne PH, et, de profil, la hauteur mq du point m par rapport au sol (voir la figure c).

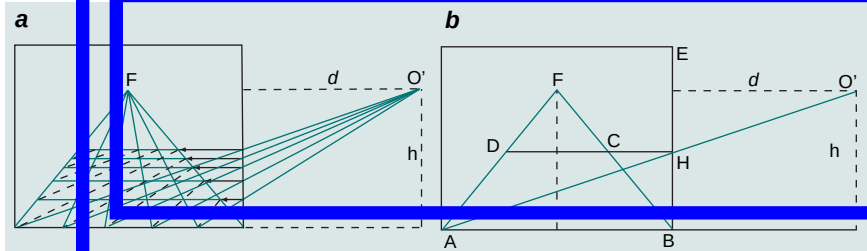
Cette méthode, probablement celle que Brunelleschi utilisa pour dessiner les premières architectures réputées être tracées en perspective linéaire, consiste à croiser les informations données par deux projections orthogonales. L'une sur le plan horizontal, l'autre sur un plan vertical, et qui peuvent être une vue de face et une vue de profil. Ces représentations conventionnelles étaient déjà connues de l'Antiquité, mais c'est Piero della Francesca qui, pour la première fois, exposa géométriquement la méthode. Le croisement des informations s'opère en juxtaposant plan et profil, où le tableau est réduit à une simple ligne et peut se lire par les points de la ligne du tableau de l'un sur la ligne du tableau de l'autre. Le principe est ici illustré avec le dessin d'un cube.



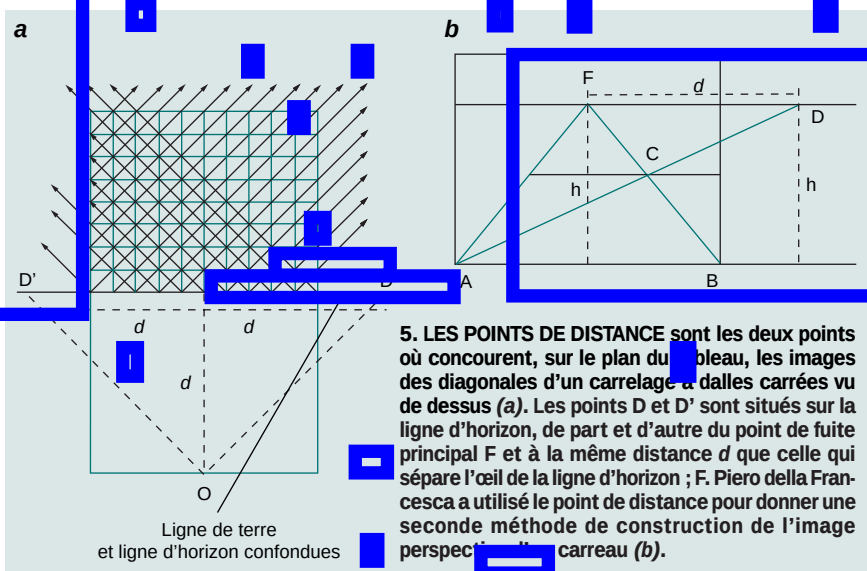
du tableau, les rayons visuels issus des points de l'arc et passant par l'œil. L'image ainsi produite est l'«hyperbole opposée» d'Apollonios, c'est-à-dire la courbe que Desargues adjoint à la première pour définir une seule et même hyperbole.

Dans la situation décrite (voir la figure 6c), le regard de l'observateur atteint les deux points A et B du cercle par des rayons parallèles au tableau, qui ne l'atteignent qu'à l'infini ; ces rayons définissent la direction des asymptotes de la courbe. Desargues conçoit ces droites asymptotes comme les tangentes aux points à l'infini de l'hyperbole, puisque ce sont les images perspectives des tangentes aux points A et B du cercle.

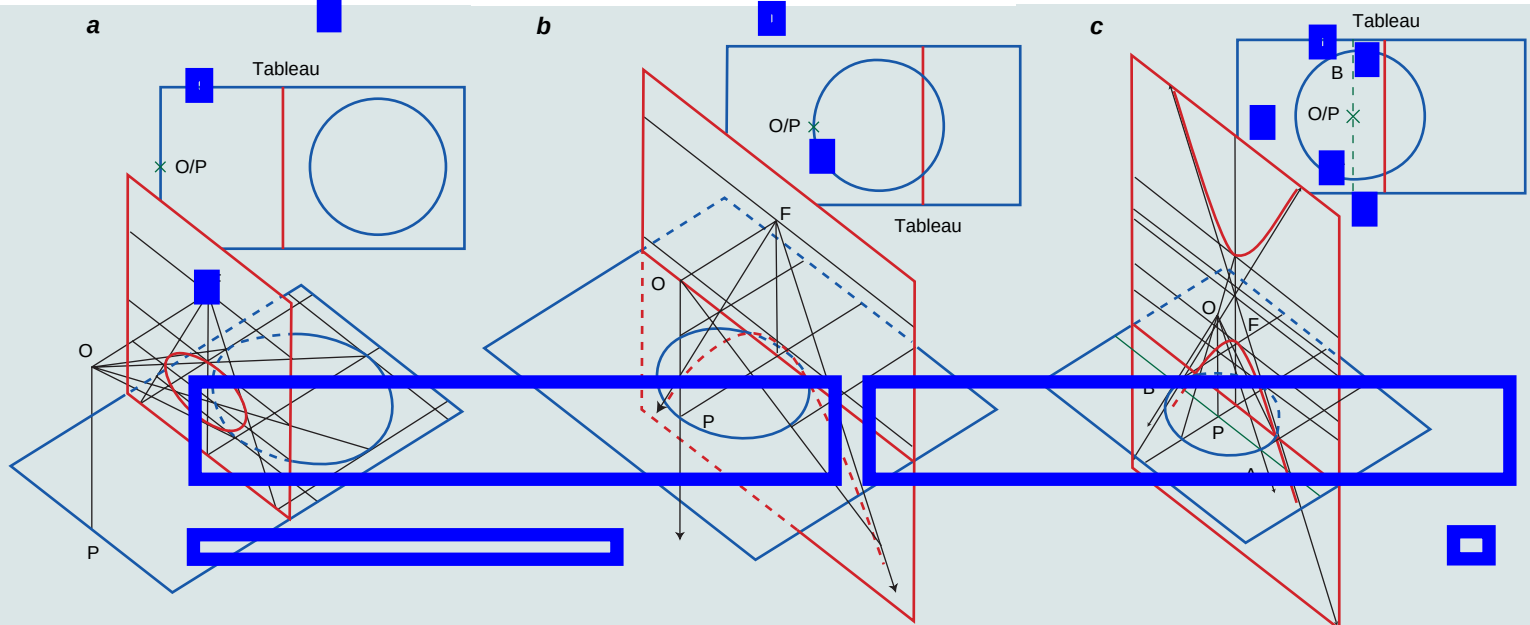
Ainsi, notre géomètre ne se borne pas à montrer qu'une perspective conserve l'alignement des points, l'intersection de droites en un point ou l'intersection de plans en une droite (étant entendu que ce point ou cette droite peuvent être à l'infini dans le cas de faisceaux parallèles). Desargues met en évidence d'autres propriétés qui restent inchangées lors des projections. Par exemple, la perspective conserve les tangentes aux courbes (si T est la tangente à une courbe C en un point P, l'image perspective de T est la tangente de la courbe-image de C au point-image de P, y compris dans le cas limite des asymptotes). De même, Desargues montre qu'un certain rapport numérique (qu'il



4. UN CARREPAGE VU DE DESSUS représenté par l'architecte et théoricien italien Leon Battista Alberti : il trace d'abord le faisceau de droites qui concourent au point de fuite principal F. Il place ensuite un point O' de côté, à une distance d du tableau. d est la distance entre l'œil du peintre O et le point de fuite F) et à la même hauteur que F. De O', il trace les segments de droite allant jusqu'aux points de départ des droites concourantes en F. L'intersection de ces segments avec le bord vertical droit du tableau fournit les hauteurs auxquelles doivent être placées les lignes transversales du carrelage. Piero della Francesca a simplifié la construction d'Alberti et l'a démontrée géométriquement avec l'image perspective d'un seul carreau (b).



5. LES POINTS DE DISTANCE sont les deux points où concourent, sur le plan du tableau, les images des diagonales d'un carrelage à dalles carrées vu de dessus (a). Les points D et D' sont situés sur la ligne d'horizon, de part et d'autre du point de fuite principal F et à la même distance d que celle qui sépare l'œil de la ligne d'horizon ; F. Piero della Francesca a utilisé le point de distance pour donner une seconde méthode de construction de l'image perspective d'un carreau (b).



6. SELON LA POSITION DU PEINTRE par rapport à l'objet à dessiner, la représentation change. En perspective centrale, où l'œil du peintre O constitue le centre de la projection, la représentation d'un bassin circulaire (en bleu) sur le plan du tableau (en rouge) peut être une ellipse (a),

une parabole (b) ou une hyperbole (c). On obtient une ellipse lorsque le bassin se trouve au-delà du peintre, une parabole lorsque les pieds P du peintre sont sur le contour du bassin, et une hyperbole lorsque le peintre est debout à l'intérieur du bassin.



7. LES DÉCORS PLAFONNANTS, tel l'Oculus de la chapelle des époux d'Andrea Mantegna (ci-dessus) ou les trompe-l'œil sur voûtes sont des représentations en perspective où l'artiste adopte un «point de vue» ou n'est ancré de la verticalité.

appelle une involution et auquel on préfère aujourd'hui le terme de birapport de quatre points) entre les segments définis par six points d'une droite reste inchangé par projection centrale ou par projection parallèlement à une direction donnée.

Les découvertes de Desargues ont rendue possible une théorie générale des projections, à laquelle s'emploieront les géomètres de la première moitié du XIX^e siècle, Gaspard Monge, Jean-Victor Poncelet et Michel Chasles, pour ne citer qu'eux. À son tour, la géométrie projective construite par ces mathématiciens conduira à entrevoir la possibilité des géométries non euclidiennes et à en concevoir des modèles euclidiens. On y rencontre aussi des représentations de l'infini. Par exemple, dans les modèles d'Henri Poincaré ou d'Eugenio Beltrami pour le «plan hyperbolique», l'infini est ramené à distance finie en assimilant les points à l'infini à un cercle ; l'intérieur de ce cercle constitue le «plan», et les «droites» de ce plan sont par définition des arcs circulaires orthogonaux au contour. Il s'agit d'un modèle euclidien d'une géométrie non euclidienne, puisque, par un point extérieur à une «droite», passent une infinité de «droites» parallèles à la droite en question. On notera que les extrémités des droites sont situées sur le contour, c'est-à-dire à l'infini : avec un tel modèle, le monde clos des Anciens fait en quelque sorte son retour...

La perspective cavalière

Les projections parallèles, projections où les rayons qui partent de l'objet à représenter sont parallèles à une direction donnée, semblent connues de longue date, au moins depuis l'architecte romain Vitruve (I^{er} siècle de notre ère) : beaucoup de dessins, qu'ils soient à vocation artistique ou technique, semblent relever de ce que l'on nomme aujourd'hui la perspective cavalière ou d'une projection orthogonale (projection dans une direction orthogonale au plan du dessin). En fait, il ne s'agit que de tentatives empiriques de représenter localement la profondeur pour des objets isolés. Ils ne peuvent en aucun cas relever d'une volonté de «représenter l'espace» de façon globale, cohérente et homogène. On ne peut véritablement parler de pers-

pective cavalière qu'à partir du moment où ses règles ont été codifiées géométriquement. Or celles-ci, bien que nettement plus simples que celles de la perspective centrale, n'ont été édictées qu'après celles de la perspective centrale du Quattrocento.

Les premiers dessins en perspective parallèle systématique apparaissent dans les théâtres de machines (des plans de machines) ou dans les ouvrages d'ingénieurs du XVI^e siècle, et les premiers traités de perspective cavalière ou militaire sont le fait de théoriciens français du début du XVII^e siècle. Quant aux architectes, qui ont donné la perspective centrale aux peintres, ils ne se sont saisis que tardivement, à partir de la fin du XIX^e siècle, de la perspective cavalière et de l'axonométrie (la représentation des trois dimensions par projection sur trois axes placés dans le plan du dessin).

Dans les perspectives parallèles, on considère les rayons visuels parallèles. Comme l'a montré Desargues, on peut aussi dire que ces rayons visuels se rencontrent à l'infini. En d'autres termes, la perspective cavalière équivaut à une perspective centrale où l'œil du dessinateur est «rejeté à l'infini». En 1921, le peintre russe Lissitzky s'exclame : «Le suprématisme a fait reculer l'extrémité de la pointe de la pyramide visuelle à l'infini... [Ce faisant] il a inversé la dernière illusion : l'extensibilité infinie vers l'arrière-plan ou l'avant-plan.» Il lui semblait qu'on en avait fini avec la peinture du sujet renaissant, que le point de vue n'existait plus.

Le peintre s'assimile alors au Créateur par le regard venu de l'infini. Il nous paraît pourtant que, dans une perspective parallèle, le point de vue est toujours présent, tout rejeté à l'infini qu'il soit et justement du fait qu'il est rejeté : une exclusion n'est pas une absence.

Jean-Pierre LE GOFF est professeur de mathématiques et d'histoire des mathématiques à l'IUFM (Institut universitaire de formation des maîtres) de Caen et à l'IREM (Institut de recherches sur l'enseignement des mathématiques) de Basse-Normandie (Université de Caen).

René TATON, *L'œuvre mathématique de Girard Desargues*, PUF, 1951 (réédition Vrin, 1988).

Philippe COMAR, *La perspective en jeu : Les dessous de l'image*, Éditions Gallimard/La Découverte, n° 138, 1992.

Jean-Pierre LE GOFF, *Desargues et la naissance de la géométrie projective*, in

Desargues et son temps, sous la direction de Jean Dombrowski et Joël Sakarovich, Blanchard, 1994.

Jean-Pierre LE GOFF, *Piero della Francesca. De la perspective en peinture*, préface de Hubert Damisch et postface de Daniel Arasse, in *Medias Res*, 1998.

Les Cahiers de la perspective, sous la direction de Daniel BESSOT et Jean-Pierre LE GOFF, IREM de Basse-Normandie, Université de Caen, 1981-1996.

D. BESSOT et J.-P. LE GOFF, *Mais où est donc passée la troisième dimension ?* in *Histoires de problèmes, histoire des mathématiques*, Éditions Ellipses, 1993.

73

L'infini, pierre de touche du constructivisme

ALLAN CALDER

Selon les mathématiciens constructivistes, un objet mathématique n'existe que si on peut le construire. Cette définition de l'existence mathématique oppose, depuis plus d'un siècle, constructivistes et formalistes.



La plus grande perfection de Dieu est la possibilité de créer un ensemble infini, et son immense bonté le conduit à le créer.

Georg Cantor

Si des hommes communiquaient avec une autre forme de vie dans l'Univers, les deux civilisations auraient, pensent les mathématiciens réalistes, en commun une forme élémentaire de mathématiques qu'elles utiliseraient pour dialoguer. Depuis l'époque de Platon, de nombreux mathématiciens et philosophes pensent que les mathématiques existent extrinsèquement, en dehors de la connaissance que nous en avons, ce qui leur confère une sorte de vérité absolue. La tâche du mathématicien consiste à découvrir cette vérité.

Cette croyance en une mathématique «pré-existante» n'est pas partagée par tous : au XIX^e siècle, le mathématicien allemand Leopold Kronecker était persuadé que Dieu avait créé les entiers et que «tout le reste est l'œuvre de l'homme». Selon ce point de vue, le mathématicien ne découvre pas les mathématiques, il les invente.

Mathématiques découvertes ou mathématiques inventées, la question n'est pas futile : selon leur option philosophique, les mathématiciens orientent différemment leurs travaux. Ainsi, l'idée des mathématiques constructives, selon laquelle l'existence

d'un objet mathématique n'est établie que lorsqu'on a montré comment le construire, est fondée sur la conviction que les mathématiques sont inventées. Les conséquences de ce choix sur des sujets aussi importants que la validité de démonstrations mathématiques ou la signification de l'existence mathématique sont mises en lumière par les méthodes et les idées des mathématiques grecques, qui ont guidé l'évolution des mathématiques pendant les 25 derniers siècles.

Pour les Grecs, les mathématiques étaient presque réduites à la géométrie, et les *éléments* d'Euclide en étaient le fleuron. Les méthodes déductives de l'école éréatique de philosophie (Parménide, Zénon, etc.), la logique aristotélicienne et l'idéalisme platonicien s'y fondaient en un système axiomatique. Euclide partit d'objets non définis, tels que le point, la droite et le plan, ainsi que d'un certain nombre d'axiomes, c'est-à-dire d'intuitions concernant ces objets ; il leur appliqua les règles de la déduction aristotélicienne pour obtenir de nouvelles propositions, ou théorèmes.

Bien que des objets comme le point (qui n'a pas de dimension) et la droite (qui a une longueur, mais pas de largeur) n'existent pas dans le monde réel, nous avons généralement le sentiment de savoir ce qu'ils sont. De surcroît, les théorèmes d'Euclide étaient confirmés

empiriquement : à l'aide d'un papier et d'un crayon, on vérifie, par exemple, que deux triangles dont les côtés sont égaux ont des angles égaux.

À l'époque moderne, les mathématiques ont continué à dépendre de la méthode axiomatique ; dans nombre de cas, cette méthode a été appliquée à des objets auxquels manquaient les vertus intuitives ou «constructibles» de la géométrie euclidienne. Pour les constructivistes, de telles applications ont faussé le sens des mathématiques. C'est pourquoi un certain nombre de propositions ont été avancées, ces 100 dernières années, afin de renouer avec la tradition intuitive. Nous examinerons ici cette longue controverse, dont les aspects les plus importants apparaissent au cours de l'évolution mathématique du concept d'infini.

Les constructions de l'infini

Les paradoxes de Zénon sur la divisibilité à l'infini et la continuité échauffèrent les mathématiciens grecs qui devinrent extrêmement prudents dans l'utilisation des infinis. L'objet fondamental de la mathématique grecque était la longueur, et Euclide, par exemple, se référait, non pas à des lignes sans fin, mais à des segments de droite extensibles à des longueurs arbitraires. Cette notion d'infini «potentiel» – on considère des lignes finies en

se réservant la possibilité de les étendre à l'infini - est une extrapolation raisonnable de l'expérience. Le concept opposé, l'infini «actuel», selon lequel il existe réellement des lignes de longueur infinie, est métaphysiquement différent ; il demande un effort d'imagination vertigineux. Aussi les mathématiciens ont-ils évité les infinis actuels pour se limiter aux infinis potentiels.

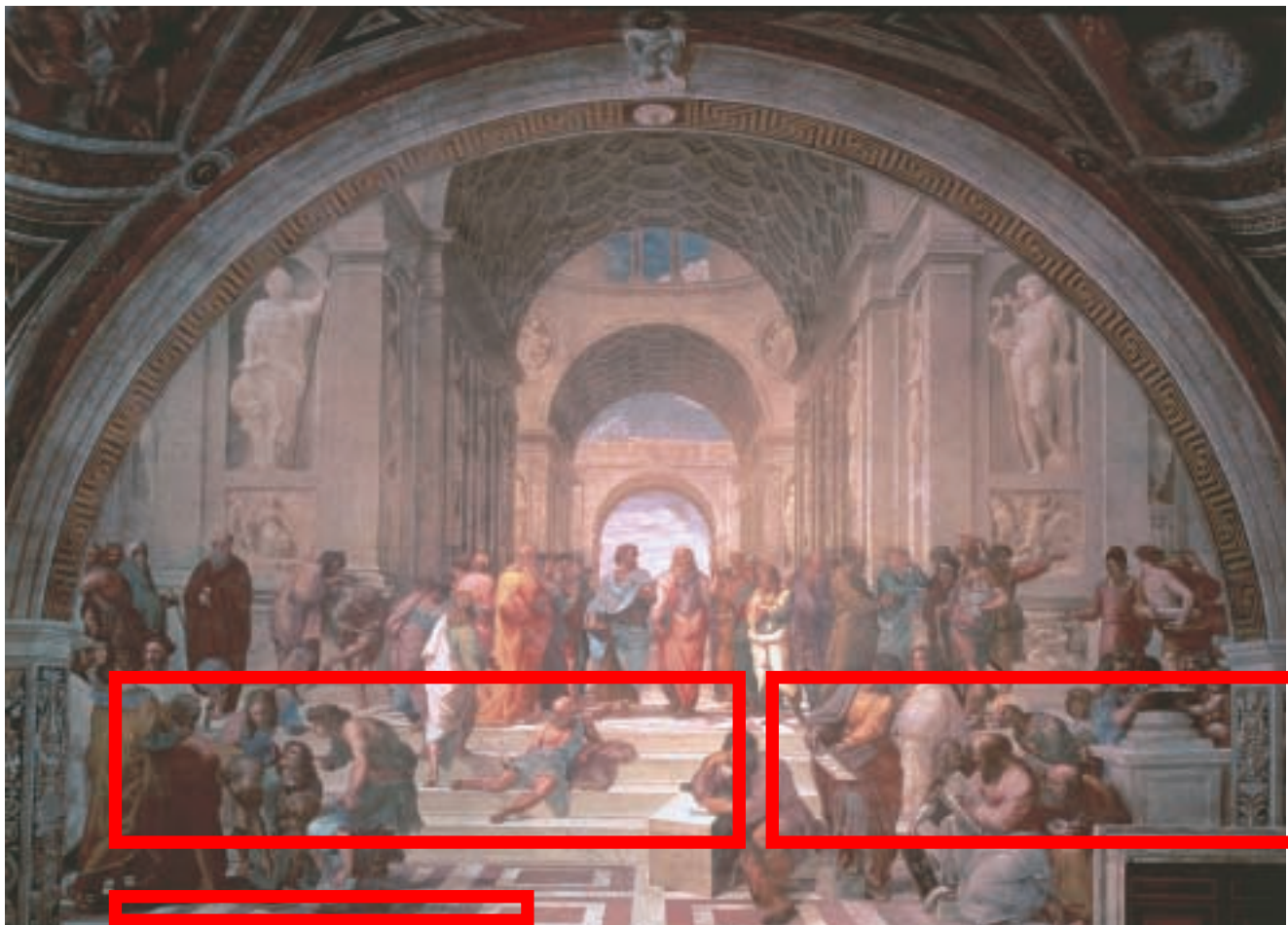
Au XVI^e siècle, René Descartes et Pierre de Fermat, en introduisant les coordonnées cartésiennes, transformèrent les problèmes sur les formes géométriques en problèmes sur les nombres. Cette numérisation de la géométrie conféra au nombre, plutôt qu'à la longueur, la place prépondérante dans les mathématiques. Les bases des travaux de Newton et de Leibniz sur le calcul infinitésimal étaient posées sous la forme d'un ensemble de lois pour manipuler des nombres qui sont infiniment petits sans être nuls.

Ce concept de quantité infiniment petite était troublant, à une époque où les objets mathématiques étaient porteurs de la même réalité que, par exemple, les électrons aujourd'hui. Il fallut un certain temps avant que l'indéniable utilité du calcul infinitésimal oblige les mathématiciens à accepter de plus ou moins bon gré l'idée des infinis actuels qui en résultait. Pendant les deux siècles qui suivirent, les mathématiciens forgèrent avec succès des techniques de calcul tout en contestant leurs bases philosophiques. À la fin du XIX^e siècle, cet irritant grain de sable au fond de l'histoire mathématique produisit une perle : la théorie des ensembles.

Depuis Georg Cantor, à qui l'on doit les premières ébauches de la théorie des ensembles, un ensemble est défini comme une collection d'objets, réels ou abstraits, munie d'une loi qui détermine leur appartenance à l'ensemble. Cette idée simple eut de profondes

conséquences sur le concept d'infini. Des ensembles comme les nombres réels sont, par leur nature même, des infinis actuels. L'étude de tels objets conduisit Cantor à tirer la conclusion déconcertante que, tout comme la cardinalité des ensembles finis (ou leur nombre d'éléments), la cardinalité des ensembles infinis pouvait être mesurée : certains ensembles étaient plus infinis que d'autres !

Cantor définit alors toute une hiérarchie d'ensembles infinis : l'ensemble des nombres rationnels, des nombres irrationnels (des nombres réels qui ne sont pas des fractions), des nombres transcendants (des nombres réels qui ne sont racine d'aucune équation de type $a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0$, où les nombres a_n sont des entiers relatifs), etc. Selon Cantor, deux ensembles ont la même taille, ou cardinalité, quand tous les éléments de l'un peuvent être, un à un, associés à tous les



1. DANS L'ÉCOLE D'ATHÈNES, qu'il exécuta au début du XVI^e siècle, Raphaël dépeint les grands philosophes et mathématiciens de la Grèce antique : au centre se tiennent Platon (la main levée) et Aristote ; en bas à gauche, Pythagore médite sur son système de proportions, alors qu'Euclide trace un cercle sur une ardoise, en bas à droite. Pour les Grecs, les mathématiques se limitaient presque exclusivement à la géométrie, et c'est dans son traité de géométrie, les *Éléments*, qu'Eu-

clide réunit pour la première fois les bases d'un système axiomatique. Dans ce système, certaines affirmations évidentes, initiatives ou axiomes, permettent d'obtenir, grâce aux règles aristotéliennes de la déduction, de nouveaux résultats nommés théorèmes. Les mathématiciens constructivistes proposent de reconstruire les mathématiques sans l'aide des concepts non intuitifs, tels que l'infini actuel, introduits en mathématiques depuis quelques siècles.

éléments de l'autre. Cette manière de faire correspondre deux ensembles est une bijection.

Ainsi un ensemble infini est dénombrable s'il a la même cardinalité que l'ensemble des naturels positifs $1, 2, 3, \dots$ (voir *L'infini est-il paradoxal en mathématiques?*, pages 30 à 38). Beaucoup de mathématiciens contemporains de Cantor trouvèrent que l'idée d'ensemble infini était absurde. Cantor lui-même admit qu'il avait été assailli «presque contre son gré» à considérer des ensembles infinis comme des entités uniques. Cette suppression de la distinction entre infini potentiel et infini actuel était «en contradiction avec des traditions appréciées». Et, plus choquant encore que la notion d'ensemble,

était l'usage que Cantor en faisait ; les mathématiciens furent scandalisés par sa démonstration que l'ensemble des nombres transcendants est infini non dénombrable.

Preuves non constructives

Un nombre transcendant est un nombre qui n'est pas algébrique, et un nombre x est algébrique s'il est solution d'une équation de la forme indiquée précédemment. Par exemple, les nombres rationnels sont des nombres réels qui peuvent se mettre sous la forme p/q , avec p et q entiers, q non nul ; ainsi, tout nombre rationnel x est algébrique, car il vérifie l'équation $qx - p = 0$. Cantor démontra qu'il

existe des nombres, les nombres transcendants, qui ne remplissent pas ces conditions ; il montra surtout que les nombres transcendants sont en nombre infini, en commençant par préciser que l'ensemble des nombres algébriques est infini dénombrable. Autrement dit, il mit en évidence une bijection entre l'ensemble des entiers naturels et l'ensemble des nombres algébriques.

L'ensemble des nombres réels n'étant pas dénombrable, il doit donc exister des nombres réels «laissés pour compte», c'est-à-dire qui ne sont pas algébriques. Supposons maintenant que l'ensemble de ces nombres transcendants soit dénombrable. Il n'est pas difficile de démontrer que l'ensemble formé de l'union de deux ensembles dénom-

2. Les ordinaux, des infinis toujours plus grands

Les premiers nombres connus furent les nombres entiers naturels $0, 1, 2, 3, \dots$ qui permettent de dénombrer les ensembles d'objets discontinus : des moutons, des sacs de blé, etc. On représente géométriquement ces nombres entiers naturels par des points régulièrement répartis sur une demi-droite, la distance entre l'origine et un point de la demi-droite représentant un nombre entier naturel. Les Égyptiens, les Grecs, les Arabes et les Européens de la Renaissance ont généralisé le concept de nombre entier naturel et défini les nombres rationnels, ou fractions, grâce aux opérations «algébriques» que sont la somme, la différence, le produit et le quotient de deux nombres entiers naturels.

Après les nombres rationnels, les Grecs introduisirent les nombres réels pour des raisons géométriques : $\sqrt{2}$, par exemple, est la longueur de la diagonale d'un carré de côté 1. Avec ces généralisations, tous les points d'une droite correspondent à un nombre réel, et inversement. Les opérations algébriques appliquées à ces nombres conduisirent à diverses généralisations, tels les nombres complexes.

Vers la fin du XIX^e siècle, le mathématicien allemand Georg Cantor voulut généraliser le concept de nombre, non pas de façon algébrique, comme l'avaient fait ses prédécesseurs, mais du point de vue de leur ordre : autrement dit, il voulut, à partir du principe de dénombrement utilisé dans la construction des nombres entiers, généraliser la relation «est supérieur à» (par exemple, 7 est supérieur à 4) et conserver l'acte de compter par lequel on place un nouveau nombre après des nombres déjà définis. Les nombres infinis, ou «transfinités» inventés par Cantor permettent de compter plus loin qu'avec les nombres entiers naturels. Ainsi les ordinaux sont, soit les nombres entiers naturels, soit les nombres transfinités.

Comment Cantor a-t-il proposé de dépasser l'infini des nombres entiers naturels ? On l'imagine en reprenant l'interprétation géométrique des nombres par des points d'une demi-droite : pour dépasser l'infini que représente cet ensemble de points, on juxtapose seulement une deuxième demi-droite à côté de la première, puis une troisième, ainsi de suite (voir la figure ci-contre). On juxtapose ainsi 10, ou 100 demi-droites, et comme le nombre de demi-droites n'est pas limité, on obtient finalement un quart de plan tout entier plein de demi-droites. Pourquoi s'arrêter là et ne pas chercher ensuite à empiler de tels plans ? Les possibilités de juxtaposition sont illimitées, et les objets obtenus sont les ordinaux.

Revenons à l'image intuitive des demi-droites et examinons les ordinaux successifs que nous créons par le procédé précédent. Après avoir compté tous les nombres entiers naturels (en nombre infini), on introduit le nombre transfini ω , que l'on représente comme l'origine de la deuxième demi-droite. En se déplaçant ensuite sur cette demi-droite on compte $\omega + 1, \omega + 2, \omega + 3, \dots$. Puis, quand on a ainsi compté tous les nombres de la forme $\omega + k$, où k est un nombre entier naturel, on introduit le nombre $\omega + \omega$, que l'on note aussi $\omega \times 2$, qui constitue l'origine de la troisième demi-droite. Enfin, quand on a compté tous les nombres de la forme $(\omega \times p) + k$ (où p et k sont des nombres entiers naturels quelconques), on introduit le nombre transfini ω^2 ; à ce stade, on a formé le quart de plan évoqué précédemment. Ainsi le nombre ordinal ω^2 qui vient après tous les nombres de ce réseau régulier plan est placé au-dessus de 0 dans un plan superposé au premier. On emploiera alors d'autres plans pour faire $\omega^2 \times 2, \omega^2 \times 3, \dots$

Dans un espace de dimension n supérieure à 3, l'ordinal ω^n est encore visualisable, mais la représentation des ordinaux a des limites : l'ordinal ω^n qui vient après tous les ω^n n'est pas représentable aussi simplement. Le leitmotiv de la construction des nombres transfinités est «après» : étant donné un ensemble X d'ordinaux, on définit un nouvel ordinal x par la condition que x soit plus grand (ou «après») que tous les éléments de l'ensemble X .

À ce stade nous n'avons pas encore défini rigoureusement les ordinaux. La définition la plus simple est fondée sur un concept de «bon ordre», c'est-à-dire une façon de classer (d'ordonner) les nombres entiers naturels. Le bon ordre le plus simple correspond à celui que nous utilisons quand nous énumérons les nombres entiers naturels et que nous avons mis en œuvre précédemment : son principe tient dans «est après». Toutefois d'autres bons ordres sont possibles. Par exemple, choisissons d'ordonner deux à deux les nombres entiers non nuls en utilisant leur décomposition en facteurs premiers : tout nombre entier naturel non nul est le produit des nombres premiers 2, 3, 5, 7, 11, 13... avec des exposants convenables. Par exemple, 84 est égal au produit $2^2 \times 3^1 \times 5^0 \times 7^1$. On obtient un bon ordre si, pour comparer deux nombres entiers naturels, on considère leur décomposition en facteurs premiers et l'on cherche le plus grand facteur qui n'a pas le même exposant dans les deux nombres ; on décide de placer

brables est dénombrable (cela ne pose aucun problème si un ensemble, ou les deux, sont finis ; si tous les deux sont infinis, pour montrer que leur union est dénombrable, il suffit de mettre l'un d'eux en bijection avec l'ensemble des nombres impairs, et l'autre avec les nombres pairs). Alors, si l'hypothèse était vraie, l'ensemble des nombres réels (qui est constitué de l'union des nombres algébriques et des nombres transcendants) serait dénombrable, ce qui est une contradiction. Par conséquent, l'hypothèse selon laquelle l'ensemble des nombres transcendants est dénombrable est fautive : cet ensemble est donc infini non dénombrable.

La démonstration de Cantor utilise la méthode classique du raisonnement par l'absurde, où l'on démontre une proposition en montrant que sa négation conduit à une contradiction. Des mathématiciens comme Kronecker furent scandalisés par une telle démonstration, parce que, dans la tradition mathématique, il n'y a pas d'existence sans construction : pour établir l'existence d'un objet mathématique, on doit spécifier par quelle méthode on peut, au moins en principe, construire l'objet. Que l'existence soit établie en établissant la négation de sa non-existence, ou même en un temps « raisonnablement » court, n'est pas une nécessité, du moment que la construction comporte un nombre fini d'étapes et que le passage d'une étape à la suivante est explicite. Le démon-

stration de Cantor ne remplissait pas ces conditions : elle ne permettait pas de construire des nombres transcendants et certainement pas une infinité d'entre eux!

La démonstration de Cantor fut un des premiers exemples d'une preuve d'existence non pure. Elle n'indique pas comment obtenir un nombre transcendant, mais établit l'existence de tels nombres en montrant la contradiction qu'entraînerait leur non-existence.

La preuve d'existence pure

Le spectre de l'infini se dessine. Un raisonnement par l'absurde qui établit l'existence d'un objet dans un ensemble fini est accepté par tous les

après le nombre pour lequel cet exposant est le plus grand. Par exemple pour les nombres 11 191 et 32 984, on cherche d'abord les décompositions $19^2 \times 31$ et $5 \times 7 \times 13 \times 31$. Le facteur 31, ayant le même exposant dans les deux décompositions, on compare les exposants de 19 : il est supérieur dans 11 191, de sorte qu'on classe ce nombre après 32 984. De la même façon, ce bon ordre place 2^{100} avant 5.

Ce bon ordre permet d'ordonner tous les nombres entiers non nuls de la façon suivante : $1, 2, 2^2, 2^3, 2^4, \dots, 3 \times 2, 3 \times 2^2, 3 \times 2^3, \dots, 3^2, 3^2 \times 2, \dots$ et l'on peut associer un ordinal à chaque nombre de cette suite : on associe l'ordinal 0 à 1, l'ordinal 1 à 2, l'ordinal 2 à 2^2 , l'ordinal 3 à 2^3 , ..., l'ordinal ω à 3 , l'ordinal $\omega + 1$ à 3×2 , l'ordinal $\omega \times 2$ à 3^2 , etc. Cette relation de bon ordre permet de représenter tous les nombres ordinaux jusqu'à ω^ω (non compris).

Les ordinaux sont à l'origine de la grande peur qu'ont ressentie les mathématiciens au début du siècle : le mathématicien italien Cesare Burali Forti (1861-1931) mit en évidence le paradoxe suivant : par définition, tout élément d'un ordinal est un ordinal, et la relation d'appartenance entre ordinaux est un bon ordre. L'ensemble des ordinaux est donc un ordinal x , et x est un élément de x (parce que x est formé de tous les ordinaux). Or le point x de l'ensemble x des ordinaux vérifie la proposition « x appartient à x », ce qui contredit l'irréflexivité de la relation d'appartenance sur x . Pour lever ce paradoxe, le logicien britannique Bertrand Russell (1872-1970) et d'autres mathématiciens imposèrent une définition *ad hoc* : une « classe » d'éléments définie par une propriété commune n'est un ensemble que dans certaines conditions, et, notamment, la classe de tous les ordinaux n'est pas un ensemble. De ce fait, on ne peut plus conclure « x appartient à x » dans la démonstration du paradoxe.

Nous avons vu que les bons ordres et les ordinaux sont deux manières de décrire la même réalité : à tout bon ordre correspond

un ordinal unique. Le mathématicien Ernst Zermelo (1871-1953) a montré (en utilisant l'axiome du choix, controversé) que, pour tout ensemble, on peut trouver une relation de bon ordre qui ordonne les éléments de l'ensemble. Si l'on oublie les noms des éléments et que l'on retient seulement la façon de

classer les éléments, on associe donc au moins un ordinal à un ensemble.

Ainsi, pour l'ensemble des nombres entiers naturels, nous connaissons au moins deux bons ordres : celui que nous utilisons quotidiennement pour classer les nombres ($0, 1, 2, 3, \dots$) et celui que nous avons construit à partir de la décomposition en facteurs premiers ($0, 1, 2, 2^2, 2^3, \dots, 3, 3 \times 2, 3 \times 2^2, \dots$). Le premier bon ordre est associé à l'ordinal ω , et le second à l'ordinal ω^ω .

Le plus petit des ordinaux associés à un ensemble est le « cardinal » (ou nombre d'éléments) de l'ensemble. Cette définition généralise aux ensembles infinis une notion familière dans le cas des ensembles finis. On montre facilement que les nombres entiers naturels sont des cardinaux, et l'on énumère les cardinaux infinis dans ce qu'on appelle la suite des alephs. Ainsi aleph zéro est le cardinal de l'ensemble des nombres entiers naturels ; c'est aussi le plus petit des cardinaux infinis, et il est confondu avec l'ordinal ω . Cantor montra d'autre part que le cardinal de l'ensemble des nombres réels est supérieur ou égal à aleph un (c'est ce que l'on appelle le paradoxe de Cantor). Il ruina sa santé à essayer de démontrer que le cardinal de l'ensemble des nombres réels était exactement égal à aleph un (hypothèse du continu). En fait, l'hypothèse du continu n'est pas décidable sur la seule base des axiomes de la théorie des ensembles.

Jean-Yves GIRARD
Institut de mathématiques (CNRS), Marseille-Luminy



Jan Brouwer (1881-1966)

mathématiciens, car on ne peut toujours (en principe) exhiber l'objet cherché en dressant la liste exhaustive des éléments de l'ensemble. Ce raisonnement ne s'applique pas aux nombres transcendants, eux-mêmes éléments de l'ensemble infini des nombres réels :

pour cette raison, de nombreux mathématiciens rejetèrent la démonstration de Cantor, arguant qu'une contradiction ne remplaçait pas un exemple tangible.

Cette controverse était loin d'être apaisée quand David Hilbert publia, en 1889, une preuve d'existence pure encore plus litigieuse. La démonstration concernait un problème posé par le mathématicien allemand Paul Gordon. Ce problème, l'une des questions mathématiques les plus brûlantes du temps, conjecturait l'existence d'ensembles (finis) d'objets présentant des propriétés particulières dans un certain nombre de cas. Hilbert, en raisonnant par l'absurde, montra que les ensembles en question existaient toujours. Cantor avait prouvé l'existence d'un ensemble d'objets dont on connaissait au moins quelques repré-

sentants : on avait découvert certains nombres transcendants quelque 30 ans auparavant. Hilbert soutenait à présent qu'il avait établi l'existence d'objets qu'on n'avait jamais vus et que personne ne savait construire. «Ce ne sont pas des mathématiques, s'écria Gordon, c'est de la théologie!»

Néanmoins, la preuve d'existence pure faisait merveille dans les mains de Hilbert, et comme la théorie des ensembles commençait à apporter d'importants résultats mathématiques, les objections s'estompèrent. Gordon en vint même à reconnaître que «la théologie, elle aussi, avait ses mérites». Après la mort de Kronecker, en 1891, l'opposition aux méthodes non constructives diminua d'intensité, jusqu'à ce que le mathématicien hollandais Luitzen Brouwer lui insufflât un sang neuf, en 1907. Selon Brouwer, les mathématiques perdraient tout sens si on ne faisait pas une distinction nette entre l'existence réelle, ou constructive, et l'existence pure. Ces deux notions avaient été confondues dans le prolongement de la logique classique du fini à l'infini, notamment à propos de la loi du tiers exclu.

Le tiers exclu... exclu

La loi aristotélicienne du tiers exclu postule que toute proposition est soit vraie, soit fausse. Sans cette loi, il n'est évidemment pas de preuve par l'absurde. Par exemple, dans la démonstration de l'existence des nombres transcendants par Cantor, l'hypothèse que tout nombre est algébrique mène à une contradiction ; ainsi, selon le principe du tiers exclu, il doit exister des nombres transcendants. La logique aristotélicienne fait si bien partie de la vie quotidienne que la remettre en cause paraît extravagant. La majorité des gens sont persuadés du bien-fondé de cette loi : si un sac contient un nombre fini de pommes, personne ne doute qu'il y a, soit une pomme rouge, soit pas. Personne n'a besoin de sortir les pommes du sac pour s'assurer que l'une de ces deux assertions est exacte (si l'on veut savoir laquelle est vraie, on répondra à la question après un temps fini).

La situation est différente quand on applique la loi du tiers exclu à un ensemble infini, par exemple un sac contenant une infinité de pommes, car on ne pourra jamais les examiner

L'INFINI ACTUEL ET LEIBNIZ
«Je suis tellement pour l'infini actuel qu'au lieu d'admettre que la nature l'abhorre je dirais qu'elle l'affecte partout, pour mieux marquer la perfection de son auteur. Ainsi, je crois qu'il n'y a aucune partie de la matière qui ne soit, je ne dis pas divisible, mais actuellement divisée, et, par conséquent, la moindre particule doit être considérée comme un monde plein d'une infinité de créatures différentes.»
Leibniz

toutes. Si on peut affirmer que le sac contient au moins une pomme rouge quand on en a tiré une, on ne peut pas conclure que le sac ne contient pas de pomme rouge tant qu'on n'a pas retiré toutes les pommes, ce qui est impossible puisqu'elles sont en nombre infini.

L'infini actuel n'étant pas présent dans la nature, les mathématiciens n'en ont donc pas d'idée intuitive. Pour appliquer la loi du tiers exclu à des ensembles infinis, ils doivent extrapoler à partir de leur connaissance des ensembles finis. Cette extrapolation semble d'autant plus délicate à justifier que les Grecs, à qui l'on doit cette loi du tiers exclu, avaient la plus grande méfiance envers les infinis actuels. Brouwer a même observé que les résultats de telles extrapolations pourraient ne pas être «constructivement» justes, car il serait possible de prouver l'existence d'objets «idéaux» impossibles à construire. Brouwer avançait qu'en passant à de tels objets idéaux la même «réalité» qu'aux objets constructibles on portait atteinte à la certitude mathématique. Les objets idéaux en appellent d'autres et, bientôt, on se voit plus si les mathématiques ont un fondement réel ou non.

Les critiques de Brouwer arrivèrent à un moment où la théorie des ensembles était vulnérable. Après 1900, la réorganisation ensembliste des mathématiques avait abouti à la formulation de paradoxes. Par exemple, comme pour tout ensemble, on peut trouver un ensemble de cardinalité supérieure. Cantor avait indiqué qu'il devait exister un ensemble V de cardinalité supérieure à celle de l'ensemble universel U , qui contient tous les ensembles ; d'autre part, comme U contient tout, tous les éléments de V doivent appartenir à U , et la cardinalité de V doit être inférieure ou égale



3. L'UTILISATION DU TIERS EXCLU est délicate dans le cas d'ensembles infinis. Ainsi comment décider si un ensemble dénombrable de pommes contient ou non une pomme rouge sans les examiner toutes, ce qui est impossible? Les démonstrations par contradiction reposent sur la loi du tiers exclu, qui interdit que toute proposition soit autre que vraie ou fausse. Les constructivistes rejettent l'application d'une telle loi à des ensembles infinis, parce qu'il peut être impossible de déterminer laquelle des deux propositions est vraie. En médaillon, le père du constructionnisme, Luitzen Egbertus Jan Brouwer (1881-1966).

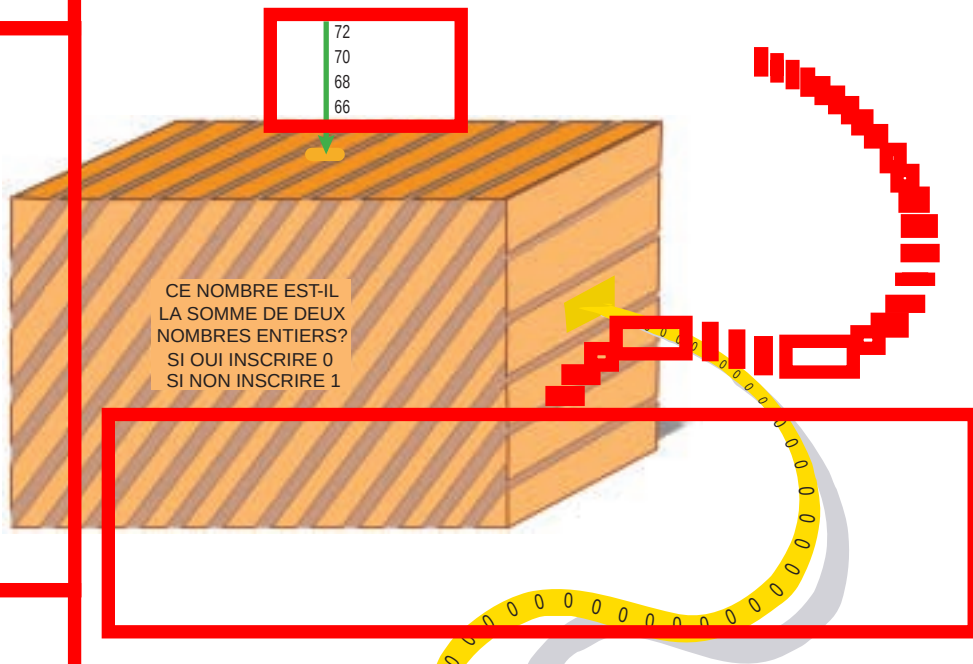
à celle de Cantor. Des paradoxes de ce genre sapèrent les fondements des mathématiques ensemblistes et, selon Brouwer, les erreurs provenaient de l'introduction des objets idéaux en mathématiques.

Brouwer était convaincu que l'on résoudrait la crise des paradoxes si l'on rejetait les objets et arguments idéaux, c'est-à-dire si l'on rebâtissait les mathématiques avec des méthodes exclusivement constructives. Comme ce programme, nommé intuitionnisme, aurait fait s'écrouler la presque totalité de l'édifice, il ne fut pas bien reçu, même de la part des mathématiciens ralliés à la cause de Brouwer. Ils étaient bien plus disposés à suivre le plan de Hilbert, qui ne cherchait pas à purifier les mathématiques du dedans, mais à les justifier du dehors. Ainsi les mathématiciens ne seraient pas, selon les propres mots de Hilbert, chassés du paradis que Cantor leur avait ouvert.

Hilbert croyait que, même si les objets idéaux n'avaient pas de signification propre, on pourrait en déduire des objets et des théorèmes importants et utiles. Selon lui, on pouvait espérer redonner un sens à la notion de certitude en transformant les mathématiques en un système axiomatique purement formel : une sorte de jeu dont l'intérêt résiderait non pas dans les objets qu'il manipule, mais dans les règles et dans les relations que ces dernières créent entre ces objets. Hilbert se hasardait même à dire que les mathématiques auraient un sens le jour où le système se révélerait cohérent, c'est-à-dire dépourvu de contradiction.

La théorie de la preuve

Le programme de Hilbert - le formalisme - avait pour premier objectif la représentation des mathématiques sous forme d'un système axiomatique formel. Il espérait qu'en transformant ainsi les données mathématiques en chaînes de symboles abstraits, combinées selon les lois de la logique, tout raisonnement porteur de germe de paradoxe serait démasqué. Mieux encore, une fois les mathématiques formalisées, on pourrait entreprendre de démontrer leur cohérence. À cette fin, Hilbert introduisit une nouvelle théorie, la théorie de la preuve, où il était possible de formuler des énoncés signifiants à partir des signes et



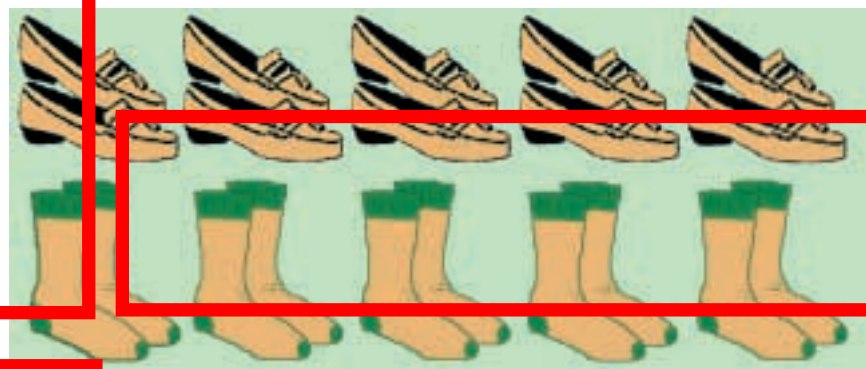
4. LA CONJECTURE DE GOLDBACH énonce que tout entier naturel pair est la somme de deux nombres premiers. Bien que cette proposition reste non démontrée, il existe un procédé fiable pour savoir si un nombre pair donné est la somme de deux nombres premiers. On peut donc programmer un ordinateur pour qu'il examine chaque nombre pair séparément et qu'il inscrive un 0 si ce nombre est la somme de deux nombres premiers, et un 1 dans le cas contraire. L'ensemble des nombres pairs étant infini, la chaîne des 0 (avec peut-être des 1) l'est aussi. Est-il correct d'appliquer la loi aristotélicienne du tiers exclu à cette chaîne infinie, et de dire qu'elle renferme un 1 ou qu'elle n'en renferme pas? Le problème est qu'on ne saura peut-être jamais, faute de démonstration, laquelle des deux hypothèses est la bonne.

des configurations dénués de sens du système axiomatique. Cette métamathématique devait être entièrement constructive, voire «finitiste», par essence. Il devenait, par conséquent, impermissible de mettre en doute la preuve métamathématique de la cohérence des mathématiques.

Brouwer refusa de croire que montrer la cohérence des mathématiques suffirait à leur redonner un sens et passa la majeure partie de sa vie à essayer de montrer que l'on pouvait «construire» les mathématiques. Son programme ne

fut guère populaire : les tentatives intuitionnistes étaient lourdes, artificielles et difficiles à utiliser. Quand Brouwer entreprit de débarrasser les mathématiques des notions idéales, tout en conservant certains aspects éprouvés de longue date, il fit appel à des notions qui parurent tout aussi idéalistes que celles qu'il éliminait.

Au total, ce n'est pas à cause d'une foi intense dans les mathématiques que les mathématiciens rejetèrent l'intuitionnisme, mais plutôt à cause de l'absence d'une telle foi. Beaucoup



5. L'AXIOME DU CHOIX, l'un des axiomes le plus utilisés et le plus controversés de la théorie des ensembles, admet que l'on peut toujours, partant d'une collection quelconque d'ensembles non vides, choisir un représentant dans chaque ensemble. Autrement dit, pour chacune des deux collections représentées ci-dessus, il devrait être possible d'indiquer une règle pour choisir des éléments, de sorte que deux personnes différentes opèrent les mêmes choix. Par exemple, pour la collection infinie de paires de chaussures, la loi «Choisissez la chaussure droite de chaque paire» est suffisante, car les deux chaussures sont dissemblables. Il n'existe pas de telle loi pour une collection infinie de paires de chaussettes toutes semblables.

6. LA DÉFINITION DES NOMBRES RÉELS CONSTRUCTIFS.

La question est : pourquoi ne définit-on pas les nombres constructifs à partir des développements décimaux ?

L'idée la plus naturelle pour définir un nombre réel constructif consiste à dire que c'est un nombre réel pour lequel on connaît un moyen numérique (c'est-à-dire algorithmique) d'en calculer les décimales aussi loin qu'on le veut. Appelons un tel nombre «décimalement constructif».

En voici deux exemples x et y :

x est égal à 1, défini comme le nombre dont la suite des décimales est 1,0000 ...

y est égal à 0 si tout nombre pair est somme de deux nombres premiers (conjecture de Goldbach, non démontrée aujourd'hui) et est égal à $1/10^n$ sinon, en prenant pour n le plus petit entier tel que $2n$ ne s'écrit pas comme somme de deux nombres premiers.

Le nombre y est «décimalement constructif», car, si vous voulez connaître ses dix premières décimales, il suffit de vérifier que 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 s'écrivent chacun comme somme de deux nombres premiers (cinq minutes suffisent pour s'en assurer). Donc x est inférieur à $1/10^{10}$ et s'écrit 0,0000000000... : on connaît avec certitude les dix premiers chiffres décimaux de y .

Considérons maintenant $z = x - y$. Puis-je connaître ses 10 premières décimales ? Non. Sans savoir si la conjecture de Goldbach est vraie ou fausse, on ne peut même pas connaître la première décimale de z . En effet, si la conjecture de Goldbach est vraie, alors $z = 1$, alors que, si elle est fausse, $z = 0,99999...999990000...$, les 9 se prolongeant jusqu'au chiffre n correspondant à un contre-exemple, $2n$, de la conjecture.

Cela montre que la différence de deux nombres «décimalement constructifs» n'est pas nécessairement un nombre «décimalement constructif». C'est pourquoi on adopte, pour définir les nombres constructifs, une méthode ne se référant pas aux chiffres décimaux. Cette définition consiste à dire que x est un nombre réel constructif si l'on connaît un moyen numérique de l'approcher d'autant près qu'on le veut par des nombres rationnels. Avec cette définition, les nombres x et y proposés ci-dessus sont des nombres réels constructifs, ainsi que $z = x - y$ (en vérifiant la conjecture de Goldbach jusqu'à 20, on en déduit que $0 \leq 1 - z < 1/10^{10}$). Plus généralement, avec la définition adoptée, la somme de deux nombres constructifs l'est, ainsi que leur différence et leur produit, etc.

Il est amusant de noter qu'Alan Turing, dans son article fondateur de l'informatique théorique de 1936, introduisait la notion de nombre réel calculable en se référant aux décimales (comme nous l'avons fait pour les nombres «décimalement constructifs»). Quelques mois après, prenant conscience de ce que sa notion était malcommode (la différence de deux nombres calculables ne l'est pas nécessairement), il suggérait une modification de sa définition.

étaient inquiets de voir leur discipline perdre son sens véritable ; la seule manière de garder le bateau à flots leur parut être d'accepter certaines notions idéales. Citons encore Hilbert : «À côté de l'immense essor des mathématiques modernes, quel sens auraient ces pitoyables vestiges, maigres résultats isolés, incomplets et sans lien, obtenus par les intuitionnistes ? » Le formalisme de Hilbert offrait à beaucoup un compromis attirant. Si l'on pouvait montrer que les mathématiques ne contenaient aucune contradiction, on serait sûr que, si deux personnes (ou la même en deux occasions différentes) jouaient aux mathématiques, elles n'aboutiraient pas à deux réponses exactes opposées.

En 1931, Kurt Gödel montra que le programme formaliste de Hilbert était

voué à l'échec. Il prouva qu'aucune métamathématique ne permet de démontrer qu'un système formel d'axiomes est cohérent si ce système est suffisamment riche pour inclure une arithmétique élémentaire ; dans le meilleur des cas, on peut déterminer quand le système est incohérent. Entretemps, la théorie des ensembles et le formalisme s'étaient tant imposés en mathématiques qu'il n'était plus question de les rejeter sous prétexte que la quête fondamentale du formalisme était une cause perdue.

Aussi le formalisme est-il devenu l'un des moteurs de la pensée mathématique, et l'intuitionnisme, une curiosité mathématique désuète. Les mathématiciens optent pour l'élégance des démonstrations d'existence pure (même quand ils ont à leur disposition des démonstrations constructives des

mêmes résultats), et le bien-fondé de telles mathématiques est rarement remis en cause. Quelques querelles subsistent au sujet de la théorie des ensembles (notamment l'emploi de l'axiome du choix, dont il sera question plus loin), mais la controverse sur les méthodes de la théorie ensembliste de Cantor et Hilbert est terminée.

Le constructivisme revisité

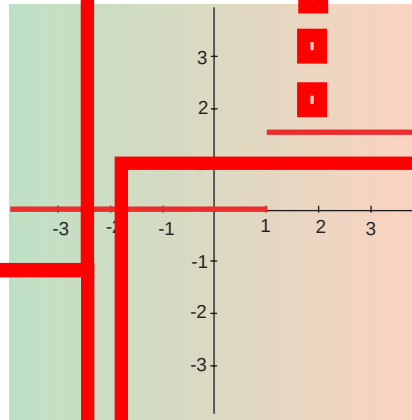
Le mathématicien Eric Bishop, en publiant en 1967 son livre, *Les fondements des mathématiques constructives*, raviva l'intérêt pour les mathématiques constructives. Contrairement à ses prédécesseurs, Bishop s'était appliqué à montrer que les résultats obtenus par ces méthodes constructives peuvent être aussi élégants et puissants que ceux qui sont obtenus par des méthodes formalistes. Alors que Brouwer voulait construire les certitudes en mathématiques en partant d'un système d'axiomes intuitifs pour l'arithmétique et travailler en remontant, Bishop aborde la «constructivisation» des mathématiques selon une optique toute différente. Son but n'est pas de reconstruire les mathématiques, mais de les pratiquer de façon constructive ; ainsi son point de départ est l'ensemble des nombres rationnels, muni des opérations classiques, et il progresse de façon «réaliste» à partir de là. Selon Bishop, on peut atteindre la certitude en mathématique en rejetant toute notion «idéaliste» et en vérifiant en permanence que les objets ou théorèmes obtenus conservent un sens intuitif. Un grand nombre de résultats importants peuvent être obtenus par cette méthode directe.

Dans les mathématiques constructives de Bishop, la définition des nombres réels ou celle des fonctions doivent être repensées radicalement. Comme dans les mathématiques non constructives, un nombre réel est défini s'il peut être approché par des nombres rationnels arbitrairement proches. D'un point de vue constructif, on doit, pour que la définition de notre nombre réel soit considérée comme valide, donner un moyen effectif (fini) pour l'approcher par des nombres rationnels à n'importe quel degré de précision voulu. Autrement dit, il doit être possible, après un nombre fini d'étapes, d'obtenir un nombre rationnel qui approche le nombre réel en question avec la précision demandée. Cela n'est pas tou-

jours aussi facile qu'il peut le paraître : par exemple, se référer simplement aux développements décimaux poserait des problèmes insurmontables.

Une autre conséquence de la définition constructive que donne Bishop d'un ensemble est surprenante: il est impossible de décomposer l'ensemble des nombres réels en deux ensembles non vides n'ayant aucun élément en commun. Cette restriction sur le recouvrement de l'ensemble des nombres réels se répercute sur le type de fonctions à valeurs réelles que l'on peut définir constructivement. La construction proposée par Bishop repose sur l'utilisation des fonctions en escalier (voir la figure 7).

Les critères d'acceptabilité des mathématiques constructives étant plus rigoureux que ceux des mathématiques non constructives, tout théorème vérifié constructivement l'est aussi non constructivement. Néanmoins, deux théorèmes distincts constructivement peuvent avoir la même interprétation non constructive. D'un autre côté, on peut transformer n'importe quel théorème non constructif en un théorème constructif, à la condition d'ajouter «en supposant le principe du tiers exclu». De plus, des démonstrations non constructives peuvent être réinterprétées constructivement. Bishop affirmait qu'une fois les enseignements positifs de son programme tirés, les mathématiques actuelles, telles qu'elles sont pratiquées, sortiraient de leur isolement pour devenir une branche des mathématiques constructives. Néanmoins, la plupart des mathématiciens n'ont pas encore trouvé de raison suffisante pour changer radicalement la conception de leur recherche. Les mathématiques ont tant progressé que personne ne peut affirmer que le manque de réalité que leur reprochent les constructivistes soit bien fondé. Après tout, l'utilité des mathématiques dans le monde réel n'est plus à démontrer. Cette «bataille entre chat et chiens mathématiciens», comme disait Albert Einstein, apparaît dénuée de tout intérêt à nombre de scientifiques. En outre, les problèmes mathématiques qui font évoluer les mathématiques sont d'ordre presque purement intuitif, interdisant ainsi de trop grands écarts par rapport à la réalité : la conviction que les mathématiques ne sont pas un jeu dénué de sens semble donc se confirmer.



7. LES FONCTIONS EN ESCALIER sont des fonctions très souvent appliquées aux nombres réels, mais qui ne peuvent être définies constructivement. En mathématiques constructives, une fonction sur les nombres réels est une loi qui associe à tout nombre réel r un autre nombre réel $f(r)$. Par exemple, la loi qui définit la fonction en escalier ci-dessus est : si r est inférieur à 1, $f(r)$ vaut 0, et si r est supérieur ou égal à 1, $f(r)$ vaut 1. Dans certains cas, il est possible de définir un nombre réel, sans pour autant pouvoir dire, au point d'un nombre fini d'étapes, s'il est inférieur, supérieur ou égal à 1. Par conséquent, il est impossible d'attribuer la valeur 0 ou 1 à l'image de ce nombre.

Même si la révolution constructive n'est pas vraiment une révolution, il n'y a pas de doute possible sur le rôle déterminant que joueront à l'avenir les mathématiques constructives dans l'évolution de la discipline. Le récent engouement pour la «constructivisation» de certains résultats montre que, dans certains cas, il y a de gros avantages à prendre les problèmes sous cet angle nouveau. On y gagne non seulement une meilleure analyse, mais aussi des théorèmes plus puissants : ainsi, lorsque l'existence constructive d'un nombre a été mise en évidence, il peut être calculé, en pratique comme en théorie.

La majorité des mathématiciens d'aujourd'hui, formés au moule du formalisme depuis plusieurs générations, ne doutent du bien-fondé des mathématiques que lorsque l'on met à jour d'insurmontables contradictions inhérentes au système axiomatique formel des mathématiques. Même si les mathématiciens pensent que de telles contradictions peuvent éventuellement apparaître, ils restent certains de pouvoir remettre les choses au point par quelques modifications mineures des axiomes. Dans ces conditions, la question de savoir si les mathématiques abstraites ont un sens réel ou non ne peut même pas être formulée.

Aux mathématiciens qui en viendraient à douter de la validité de leur science, on répète le conseil que le philosophe français Jean d'Alembert, plaçant Pascal, donnait à ceux qui doutaient du calcul infinitésimal : «Continuez, et la foi viendra.»

Toutefois, Jean-Paul Delahaye a remarqué qu'aujourd'hui la physique utilise sans aucune retenue les mathématiques non constructives et en particulier les nombres réels qui sont sans réalité physique : en physique, aucune mesure n'est infiniment précise!

Les mathématiques du continu sont donc un instrument formel mal ajusté aux besoins véritables du physicien. Les mathématiques constructives lui conviendraient mieux, puisqu'elles ne manipulent que des objets finiment définis et que l'infini y est, en quelque sorte, toujours potentiel, comme dans le monde physique. Il est donc à prévoir que, dans un avenir plus ou moins lointain, les physiciens en viendront à utiliser les mathématiques constructives (ou certaines de leurs variantes) pour éviter d'avoir à utiliser les formalismes surdéterminants (qui en disent plus qu'on ne le souhaiterait) des mathématiques usuelles non constructives.

Préoccupés par bien d'autres choses, et effrayés par la difficulté des mathématiques constructives, très peu de physiciens ont étudié à ce jour cette possibilité d'affinement de leurs outils théoriques.

Allan CALDER est professeur à l'Université d'État du Nouveau-Mexique.

E. BISHOP, *The Foundations of Constructive Analysis*, McGraw-Hill, 1967.

E. BISHOP et D. BRIDGES, *Constructive Analysis*, Springer-Verlag, 1985.

A.S. TROELSTRA et D. VAN DALEN, *Constructivism in Mathematics: an Introduction*, North-Holland, 1988.

M.J. BEESON, *Foundations of Constructive Mathematics: Metamathematical Studies*, Springer-Verlag Berlin/Heidelberg/New York, 1985.

M.B. DOUBIEL et J.I. RICHARDS, *Computability in Analysis and Physics*, Springer-Verlag, 1989.

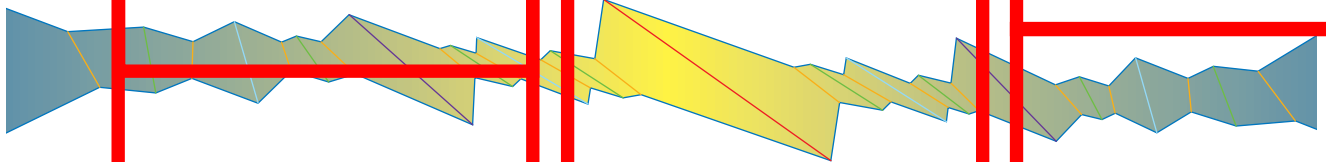
John MYHILL, *What is a Real Number?*, in *The American Mathematical Monthly*, vol. 79, n° 7, pp. 748-754.

I. LAKATOS, *A Renaissance of Empiricism in the Recent Philosophy of Mathematics*, in *Mathematics, Science and Epistemology*, Cambridge University Press, 1978.

L'infini en géométrie

MARCEL BERGER

Lors d'un dîner, un géomètre, recourt aux éléments du décor pour illustrer le rôle important de l'infini en géométrie. L'analyste et l'algébriste écoutent attentivement...



Àu sortir de la conférence, l'algébriste, l'analyste et le géomètre, dignes représentants de la trinité des mathématiques, méditent la phrase d'Hermann Weyl : «Les mathématiques sont la science de l'infini.» L'algébriste, sûr de lui, dit qu'en algèbre l'infini est omniprésent. L'analyste répond qu'il en est de même dans sa discipline. Puis tous deux se tournent, perplexes, vers le géomètre, silencieux. Observé avec insistance, celui-ci affirme également : «Oh! c'est également vrai en géométrie.» Devant les regards interrogateurs de ses deux amis, il pour-

suit : «Bien sûr, la géométrie semble être bien installée à distance finie, les cas d'égalité des triangles ou la longueur d'un cercle égale à 2π fois le rayon en sont une illustration... Où donc se loge l'infini en géométrie? Partout! Puis-je vous proposer de vous le montrer devant un bon dîner?» Les deux autres, intrigués, acceptent. Sur le chemin, le géomètre énumère en silence les différents domaines de sa discipline dans lesquels l'infini joue un rôle important. Dans chacun d'eux, l'infini est présent sous trois aspects : lorsque l'on considère des objets en nombre infini ; quand on s'intéresse

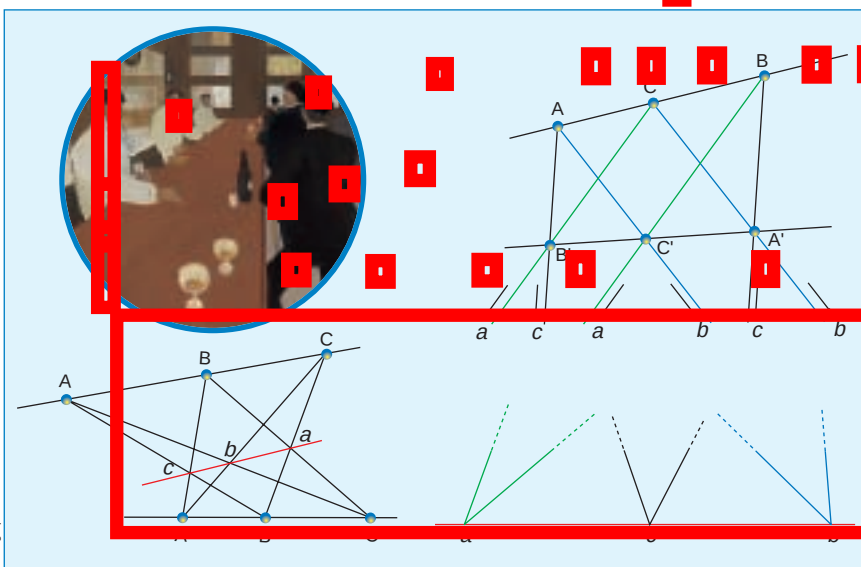
à des espaces de dimension infinie ; enfin, quand une opération est répétée un nombre infini de fois.

Nous trois mathématiciens franchissons le seuil de l'hôtel des Trois Faisans et patientons, alignés dans le corridor, en attendant l'hotesse des lieux. Là, au bar, trois rotaires passent le temps. Le géomètre ne peut manquer une telle occasion. «En guise d'introduction, vous voyez, messieurs, nous trois alignés ici, et ces trois hommes là-bas au bar, m'évoquent le théorème de Pappus (290 - 350), dont une démonstration aisée recourt à l'infini.»

L'infini selon Pappus

L'algébriste fouille sa mémoire : «Ce théorème n'affirme-t-il pas que, pour deux triplets de points alignés (A, B et C) et (A', B' et C') sur deux droites distinctes, les points d'intersection a de (BC') et $(B'C)$, b de (AC') et $(A'C)$ et c de (AB') et $(A'B)$ sont alignés (voir la figure 1)?»

L'analyste confirme : «Exact, et la démonstration requiert de voir ce résultat dans un plan projectif, c'est-à-dire un plan à l'infini (dans lequel les distances ne sont pas prises en compte), auquel on a ajouté une droite, dite à l'infini. Les points de cette droite sont les directions de toutes les droites du plan (voir *De la perspective à l'infini géométrique*, pages 66 à 72). Il n'y a plus de droites parallèles, celles qui l'étaient dans le plan seul sont maintenant sécantes en un point de la droite de l'infini. On décide, pour démontrer Pappus, que cette droite est la droite (ab) , ce que l'on peut toujours faire par une projection convenable. Ainsi, dans sa nouvelle



1. LE THÉORÈME DE PAPPUS affirme que les points d'intersection (a, b, c) des droites joignant convenablement les points de deux triplets alignés, (A, B, C) et (A', B', C') sont eux aussi alignés (à gauche). Ce théorème se démontre facilement avec l'infini : dans un plan dit projectif, on décide que les points de la droite (ab) , dite droite de l'infini, sont les directions de toutes les droites du plan ; deux droites parallèles sont alors sécantes en un point de cette droite. Puisque a et b sont sur cette droite, les droites (CB') et $(C'B)$ d'une part, et (AC') et $(A'C)$ d'autre part, sont parallèles. Démontrer le théorème de Pappus revient à montrer que le point c est aussi sur la droite de l'infini, donc que les droites (AB') et $(A'B)$ sont parallèles (à droite) ; le théorème de Thalès le montre aisément.

configuration, les droites (AC') et $(A'C)$ d'une part, et (BC') et $(B'C)$ d'autre part sont parallèles. Montrer que les points a, b, c sont alignés revient à montrer que le point c est aussi sur la droite de l'infini, c'est-à-dire que les droites de la troisième paire, (AB') et $(A'B)$, sont parallèles : le théorème de Thalès le montre aisément».

Après une petite réflexion, l'algébriste dit : «Oui, évidemment, cette démonstration est beaucoup plus simple que de se placer dans un repère et de calculer les coordonnées de tous les points».

Enfin, tout trois assis devant leur table, la rondeur des assiettes et la circularité des verres inspirent le géomètre.

Des surfaces et des volumes

«Vous connaissez tous les deux cette méthode qui requiert l'infini pour calculer la longueur d'un cercle. On y inscrit un carré, puis, bâti sur ce carré, on construit un octogone régulier, dont on divise ensuite encore en deux chacun des huit côtés, afin d'avoir un polygone régulier à 16 côtés, etc. On calcule les longueurs successives de ces polygones et, à la limite, quand le nombre de côtés tend vers l'infini, on obtient la longueur du cercle. Grâce à une telle itération, on calcule aussi la longueur d'une courbe quelconque : on y inscrit des

polygones de côtes de plus en plus petits, et la longueur de la courbe est la limite à l'infini, quand elle existe, de la suite des longueurs de ces polygones.»

Je dois pourtant me souvenir, interrompt l'algébriste, que l'on sait fabriquer des courbes de longueur infinie, même entre deux points quelconques de cette courbe.

Le géomètre reprend : «Oui, en revanche, on sait qu'il y a toujours une longueur lorsque la courbe est donnée par une équation suffisamment douce, c'est-à-dire que la courbe n'a pas de pic, ou en termes plus mathématiques, qu'elle est dérivable en chaque point.

Dans un plan, pour calculer l'aire d'un triangle, on ajoute un second triangle égal, donc de même aire : du parallélogramme obtenu, on ôte d'un côté, et on ajoute de l'autre un triangle égal afin de former un rectangle dont l'aire est égale au produit de la longueur par la largeur. Ainsi, l'aire d'un triangle est égale au demi produit de la base par la hauteur.

Toutefois, le calcul du volume d'un polyèdre est plus subtil. Dans l'espace, on peut toujours décomposer un parallélépipède en tétraèdres ; cependant, on ne sait pas de façon évidente s'ils ont même volume, même s'ils ont la même hauteur et la même aire de base. Ils ne sont pas déduits l'un de l'autre par un déplacement. Pour montrer

qu'ils ont même volume, on recourt à l'infini : on découpe les tétraèdres en tranches horizontales de plus en plus fines, chaque tranche contient un prisme et un morceau supplémentaire. Dans les deux tétraèdres, les prismes de deux tranches de même niveau ont le même volume, puisque l'aire de leur base est identique (voir la figure 5). Lorsque le nombre de prismes tend vers l'infini la somme des volumes des morceaux restants devient négligeable : les tétraèdres ont donc même volume, car ils sont tous construits d'un nombre infini de prismes de même volume.

Ainsi, on ne procède pas dans l'espace de la même façon que dans le plan, par des raisonnements de découpage en un nombre fini de morceaux de volume égaux déduits les uns des autres par des transformations de l'espace. Il est utile de rêver qu'un jour un génie trouvera un tel découpage, en 1903, Max Dehn (1878 - 1952) a démontré que ce rêve est impossible. Le cas du plan reste donc exceptionnel, l'espace nécessite l'infini.»

Et les cercles ?

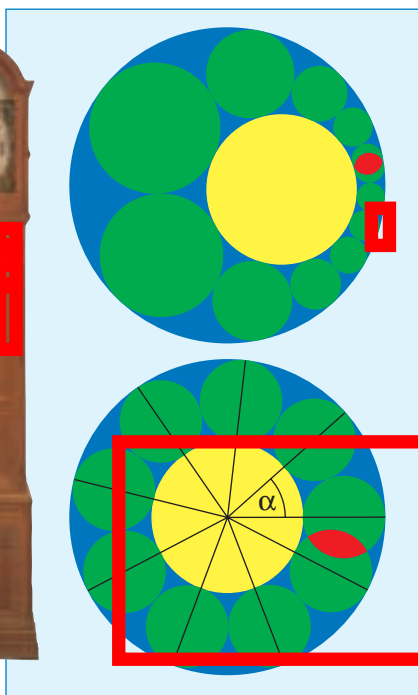
La serveuse apporte alors les apéritifs et intervertit deux verres. Encouragé par l'abondance des cercles des verres des assiettes, des soucoupes... disposées sur la table, le géomètre saisit l'occasion



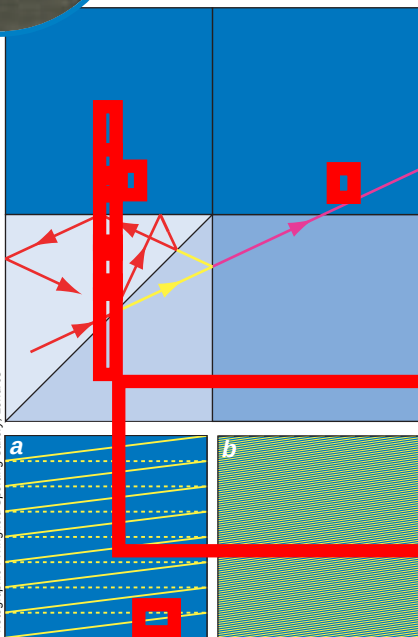
2. À L'HÔTEL DES TROIS FAISANS, on passe le temps, mais on parle aussi de l'infini en géométrie.



Cheltenham Art Gallery



3. L'ÉTUDE DES CERCLES (en vert) construits tangents à deux autres donnés (en haut) dont l'un est inclus dans l'autre (en bleu et en jaune) est évidente après une inversion : cette transformation géométrique rend concentriques les deux premiers cercles. Ainsi, les cercles construits sont égaux et il suffit d'étudier les rotations d'un angle α .



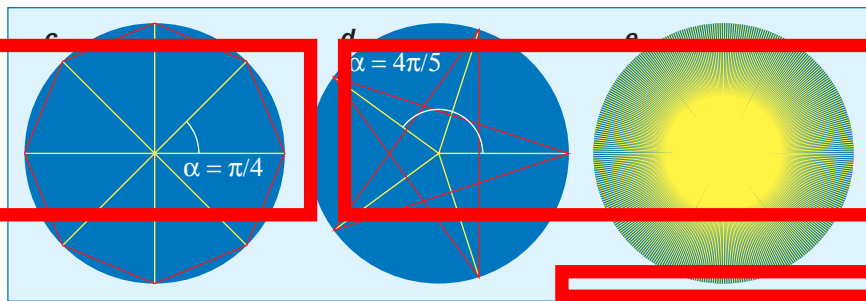
Photographie - Wingfield Sporting Gallery, Londres

4. LE COMPORTEMENT de deux particules qui rebondissent l'une contre l'autre sur un segment (à gauche, en haut) ou les rotations d'un angle dans un cercle (à droite) sont équivalents à des itérations d'une même opération. Dans le premier cas, la position des deux particules est représentée par un point dans un repère (un axe par particule). Puisqu'une particule ne peut pas passer par-dessus l'autre, la trajectoire de ce point est confinée dans un

et poursuit son exposé : « Avec la démonstration du théorème de Pappus, je vous ai montré que l'adjonction d'une droite à l'infini est parfaitement adaptée pour des problèmes de la géométrie des points et des droites. En revanche, elle n'est pas appropriée quand on s'intéresse à une géométrie de l'ensemble des cercles du plan euclidien, c'est-à-dire que l'on prend en compte la distance. La transformation du plan nommée inversion se révèle parfois pertinente. Elle consiste à se donner un point du plan, le centre d'inversion O . Tout point m autre que le centre est envoyé sur le point m' aligné avec O et m , et tel que le produit des longueurs $Om.Om'$ est constant : cette inversion transforme les cercles en cercles, sauf ceux passant par le centre O , qui sont transformés en droites.

A nouveau des itérations

Grâce à cette inversion, l'étude de problèmes sur les cercles devient plus facile : par exemple, Jacob Steiner (1796-1863) a étudié le comportement de cercles tangents à deux autres donnés, le premier étant inclus dans le second (voir la figure 3). Un cercle est construit tangent aux deux premiers, le suivant est tangent aux trois précédents, et ainsi jusqu'à avoir parcouru le périmètre du cercle intérieur : le dernier cercle construit est-il tangent au premier construit, ou le chevauche-t-il ? Une judicieuse inversion rend concentriques les deux cercles donnés au départ. Les cercles ajoutés sont donc égaux, et il s'agit maintenant d'étudier le comportement de la droite joignant le centre des cercles concentriques (voir la figure 3) à celui de ceux construits : à chaque nouveau cercle, la droite pivote d'un angle constant.»



triangle rectangle isocèle. On développe ce triangle dans un carré, puis dans le plan entier : la trajectoire est alors une droite infinie que l'on ramène dans un seul carré. Dans le second, on s'intéresse à la trajectoire d'un rayon du cercle. Certaines trajectoires sont périodiques (a, c et d), elles repassent par leur point de départ, d'autres, aperiodiques (b et e), approchent d'autant plus que l'on veut toutes les positions possibles.

Le son d'un carillon ponctue le discours du géomètre, dont le regard est attiré par les aiguilles de l'horloge : « De la même façon que l'aiguille de cette horloge tourne à chaque minute...

... nous pouvons poursuivre avec l'étude de telles itérations. Quel est l'effet sur un point du cercle de l'itération d'une rotation d'angle donné α ? Quand $\alpha = 2\pi/k$ (k est un entier), les images successives d'un point de départ donné constituent les sommets d'un polygone régulier à k côtés inscrit dans le cercle, c'est le cas de la pendule pour laquelle $k = 60$: lorsque $\alpha = 2\pi p/q$, où p et q sont deux entiers premiers entre eux, le polygone obtenu est étoilé, à q côtés et tournant p fois (voir la figure 4). Mais que se passe-t-il quand k est un nombre irrationnel ? La trajectoire du point ne se referme jamais, quel que soit le nombre d'itérations, même infini.

Toutefois, on peut en savoir plus. On montre que la trajectoire est partout et également dense dans le cercle : il y a une égale répartition des points « visités », aucune zone du cercle n'est négligée. En termes mathématiques, pour un ensemble infini $\{C_i\}_{i=1,2,3,\dots}$ de points du cercle, on dit qu'il y a égale répartition sur un arc m de longueur L du cercle lorsque k tend vers l'infini, la limite du nombre de points C_i ($i=1,2,3,\dots,k$) qui sont sur l'arc m est toujours égale à $L/2\pi$ (si le cercle est de rayon unité), cela quel que soit l'arc m ».

L'analyste coupe ici le géomètre et lui demande, entre deux bouchées : « C'est joli, les itérations, mais elles restent des opérations discrètes. Or, dans la nature, un grand nombre de systèmes sont continus dans le temps. Les géomètres étudient-ils aussi ces phénomènes lorsque le temps continue très longtemps, c'est-à-dire tend vers l'infini ? »

«Bien sûr! Le billard que j'aperçois là-bas m'inspire. En mécanique, on étudie le mouvement de deux particules de même masse sur un intervalle. Ces particules repartent en sens inverse avec la même vitesse quand elles se rencontrent. Quand on représente la position des deux particules par un point du plan, (les deux axes représentent chacun les positions d'une particule), la trajectoire de celui-ci, confinée dans un triangle rectangle isocèle, est celle d'une boule de billard qui se réfléchit sur les côtés avec un angle d'incidence égal à l'angle de réflexion par rapport à la normale, la droite perpendiculaire au côté. Nous simplifions le problème en développant la trajectoire par symétrie : le billard devient alors un carré (voir la figure 4). Puis, quand on développe la symétrie dans le plan tout entier, la trajectoire devient une droite infinie.

Vous remarquerez qu'il suffit de savoir ce qui se passe là où la trajectoire rencontre un bord du carré : le problème est alors de nouveau celui de l'itération d'un même procédé. Ici donc, deux cas se distinguent selon que la pente, qui est le rapport des vitesses initiales des deux particules, est en rapport rationnel ou irrationnel avec la longueur du carré (la longueur de l'intervalle).

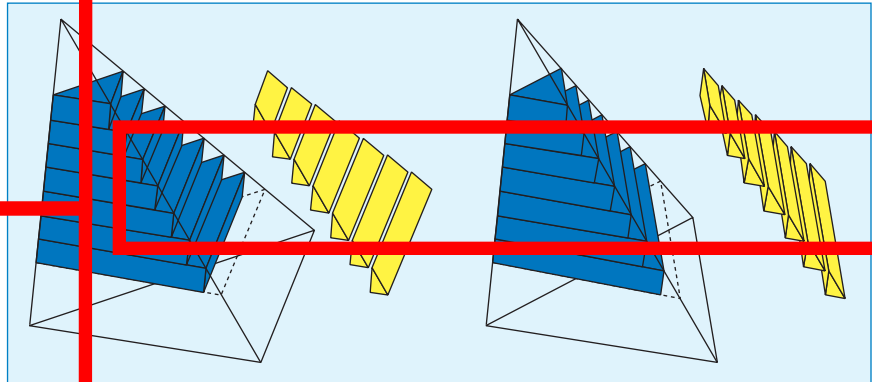
Dans le cas irrationnel, la trajectoire est partout dense et également répartie, c'est-à-dire que les deux particules occuperont toutes les positions possibles, et de façon équitablement répartie dans le temps. Sur un ordinateur, la trajectoire noircira l'écran de plus en plus, et de façon partout identique. Seule varie la vitesse de remplissage de l'écran. Dans le cas rationnel, la trajectoire est périodique ; après un temps, le mouvement reprend comme au départ : l'écran est seulement biffé de quelques lignes.»

L'algébriste et l'analyste, songeurs, demanderont alors : «Que se passe-t-il pour deux particules de masses différentes?»

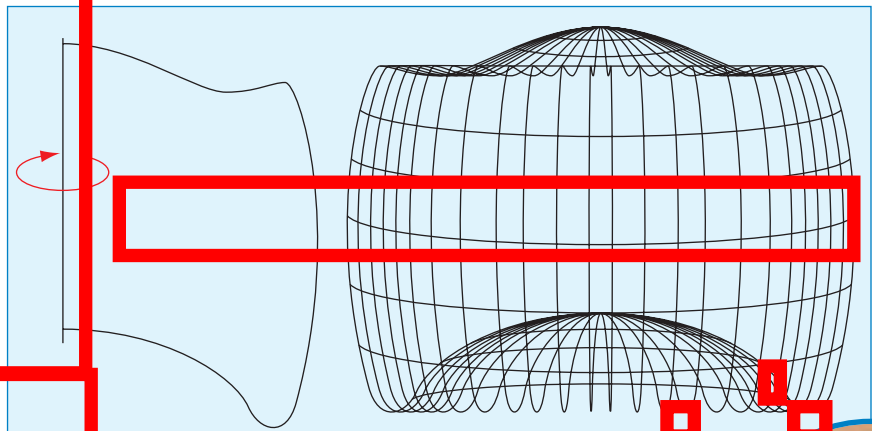
La réponse est immédiate : «Un calcul du choc élastique de particules de masses différentes, en tenant compte de la conservation de l'énergie et de la quantité de mouvement, montre que le problème devient alors celui du billard dans un triangle rectangle, mais non isocèle. Ici, les géométriciens ont un peu de résultats, car ils ne peuvent plus développer le triangle en un pavage de carrés. Le résultat dépend essentiellement du rapport des masses ; plus précisément, il s'agit de savoir quand l'angle α (un des angles du triangle), dont la tangente

est égale au produit des racines carrées des deux masses, est en rapport rationnel ou non avec π . Dans le cas rationnel, les trajectoires sont de trois types : dans des cas rares, elles sont périodiques ;

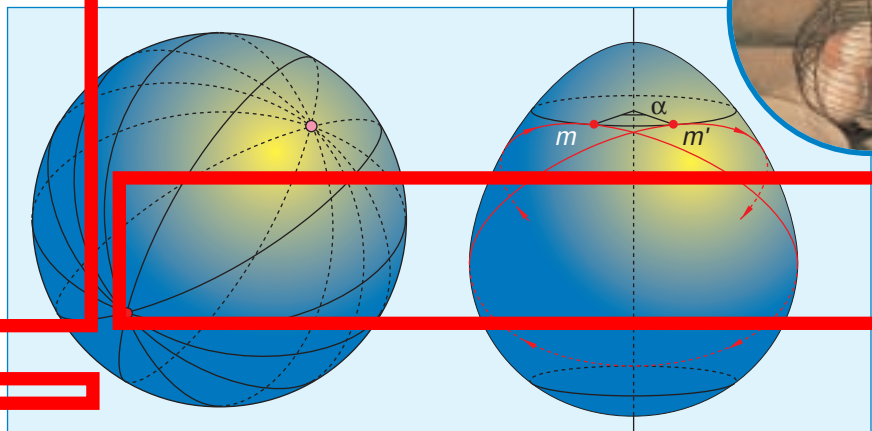
d'autres sont partout denses et bien réparties, c'est-à-dire qu'une zone du billard ne sera pas remplie avant une autre ; enfin, le plus souvent, elles sont partout denses, mais mal réparties. Dans



5. EN DÉCOUPANT DEUX TÉTRAÈDRES de même hauteur et de même aire de base en tranches, on montre qu'ils ont le même volume. Les tranches sont composées d'un prisme (en bleu) et d'un bout restant (en jaune). Dans chaque tétraèdre, les prismes du même niveau, ont un volume égal. Lorsque le nombre de prismes est infini, le volume des bouts restants est égal à zéro : les tétraèdres ont donc même volume.



6. LES SURFACES DE ZOLL sont des surfaces de révolution dont, à l'instar de la sphère, toutes les trajectoires d'une particule lancée à la surface sont des géodésiques périodiques.



7. SUR UNE SPHÈRE (à gauche), toutes les trajectoires d'une particule lancée sur la surface sans frottements et sans gravité sont des géodésiques périodiques : ce sont les grands cercles. Sur une surface de révolution (à droite), telle celle d'un œuf, les trajectoires (en rouge) sont des géodésiques périodiques pourvu qu'à partir d'un point m d'un parallèle quelconque (en noir), elles repassent sur ce même parallèle par un point m' décalé de m d'un angle α en rapport rationnel avec 2π .

tous les cas, les directions prises par une trajectoire en chaque point du billard sont en nombre fini : le système n'est donc pas ergodique en phase. Cependant, il est ergodique en position, car la trajectoire passe par tous les points. Par ailleurs, on sait que l'ordre de grandeur du nombre de trajectoires périodiques de longueur inférieure ou égale à une distance donnée croît avec le carré de la longueur de l'intervalle.»

L'analyste coupe le géomètre : «Et dans le cas irrationnel?»

«J'y venais. À partir de l'étude du cas rationnel et en approximant les triangles à angle irrationnel par de nombreux triangles à angle rationnel, on montre que l'ergodicité de phase est obtenue à l'aide de la seule ergodicité d'espace du cas rationnel. Cependant, c'est une existence complètement abstraite! On ne connaît aucune valeur de m et de m' qui permette de vérifier cette assertion. D'autre part, on ne connaît pas l'ordre de grandeur du nombre de trajectoires périodiques de longueur inférieure ou égale à une distance donnée.»

Le géomètre, intarissable, se remémorait les différents domaines de sa discipline où les itérations sont indispensables : «Puisque nous parlons d'itérations, une autre apparaît également dans la démonstration de l'inégalité isopérimétrique qui affirme que, parmi toutes les courbes fermées de longueur donnée, le cercle est celle qui entoure la plus grande aire. La démonstration de Wilhelm Blaschke (1885 - 1952), dans les années 1920, utilise une opération appelée symétrisation de Steiner (voir la figure 8). À partir d'un domaine dont le périmètre est une courbe quelconque, cette opération consiste, pour une direction donnée du plan, à découper le domaine en tranches infiniment fines, puis à entiler les tranches symétriquement autour d'un axe perpendiculaire à la direction donnée. Le domaine obtenu est de même aire, mais de périmètre plus petit que celui du domaine initial. Cette réduction de périmètre est due à une propriété des trapèzes, auxquels les tranches sont assimilées : le périmètre

d'un trapèze est minimal quand les deux côtés opposés non parallèles ont même longueur.»

L'algébriste confirme qu'un rapide calcul le montre.

Encore et toujours des itérations

Le géomètre poursuit, imperturbable : «L'itération de ces symétrisations avec infiniment de directions différentes a une limite : c'est un cercle, car lui seul est symétrique dans toutes les directions. De surcroît, la symétrisation fonctionne en toute dimension.»

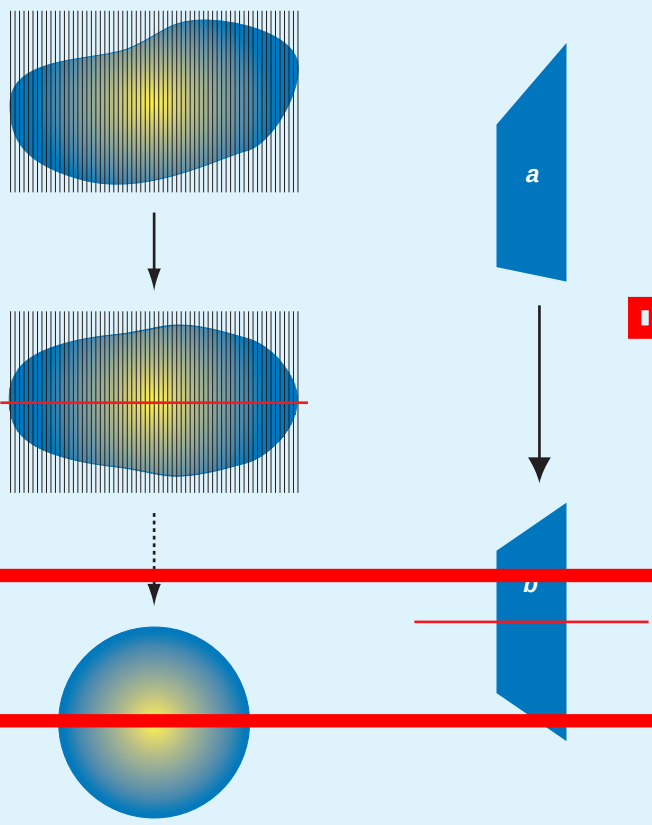
L'analyste et l'algébriste demandent alors : «Existerait-il une méthode plus naturelle pour minimiser le rapport de la longueur de la courbe frontière sur l'aire intérieure d'un domaine D ?»

Le géomètre répond : «Dans la réalité, on étudie plutôt le rapport du carré de la longueur sur l'aire pour manipuler des nombres sans dimensions. Il faut d'abord prouver l'existence d'une limite – une courbe – à ce rapport. Dans le plan, cette existence est démontrée. Toutefois, ce problème d'existence est à considérer ici dans l'espace infini de toutes les courbes : jusque dans les années 1960, dans des espaces de dimensions infinies, aucun théorème d'existence d'objets limite au sein de familles d'objets géométriques, telles les courbes, et de limites de variables elles aussi géométriques, telle l'aire d'un domaine, n'était disponible. Ce n'est qu'à la suite des travaux d'Herbert Federer, alors à l'Université Brown, qui constituent la théorie géométrique de la mesure, que l'existence de telles limites a été prouvée. On conclut alors, et ce en dimension quelconque, que la limite est, en deux dimensions, un disque, et au-delà, une boule.

Puisque nous évoquons la démonstration de l'inégalité isopérimétrique, nous pouvons aussi penser l'augmentation du rapport du carré de la longueur sur l'aire, non plus par l'opération de symétrisation, mais par une transformation continue, c'est-à-dire la déformation d'une courbe C en une courbe de plus en plus proche d'un cercle. Pour nous faire convaincre que les cercles sont des courbes dont la courbure est constante.»

L'algébriste interrompt ici le géomètre : «Tu peux nous rappeler ce qu'un géomètre entend par courbure?»

«Bien sûr! Pour déterminer la cour-



Photographie : Musée du Louvre

8. LA SYMÉTRISATION d'un domaine réduit son périmètre en maintenant son aire constante. On découpe le domaine initial en tranches infiniment fines (à gauche), puis on rassemble ces tranches symétriquement selon un axe perpendiculaire à celui des tranches. Lorsque cette opération est itérée un nombre infini de fois, le domaine tend vers un cercle qui minimise le périmètre. Celui-ci diminue en vertu d'une propriété des trapèzes de même aire (à droite) : le périmètre du trapèze b , symétrique, est inférieur à celui du trapèze a .

bure, le géomètre cherche le cercle osculateur à la courbe en un point. Pour ce faire, on considère le cercle tangent à la courbe en ce point et passant par un point voisin. Quand ce point voisin se rapproche du point étudié, ce cercle a une limite, c'est le cercle osculateur en ce point et son rayon est le rayon de courbure de la courbe. La courbure, elle, est par définition l'inverse de ce rayon. Ensuite, en chacun de ses points, la courbe se déplace le long de sa normale (la droite perpendiculaire à la tangente en ce point) d'une quantité inversement proportionnelle à la courbure en ce point (voir la figure 9). De manière plus précise, on cherche une famille de courbes $C(t)$, t est un paramètre continu, par exemple le temps, tel que la dérivée le long des normales soit inversement proportionnelle à la courbure. L'équation aux dérivées partielles qui décrit cette famille de courbes admet des solutions pour des temps assez petits. En revanche, l'existence pour un temps infini est une étude difficile.»

L'analyste interrompt de nouveau le géomètre : «Lorsque t tend vers l'infini, les courbes $C(t)$ ne convergent-elles pas vers un point, plutôt que vers un cercle?»

«Oui, cependant, convenablement renormalisées, c'est-à-dire que l'échelle de chaque courbe est modifiée de façon à conserver l'aire du domaine, les courbes $C(t)$ convergent vraiment vers un cercle. Enfin, en 1986, Gage a montré que le rapport isopérimétrique décroît lorsque t augmente. Cette technique d'évolution géométrique peut être utilisée dans des contextes plus généraux, par exemple pour obtenir des géodésiques périodiques sur les surfaces. Sujet que j'aimerais maintenant aborder avec vous, si vous le permettez.»

Les géodésiques du dessert

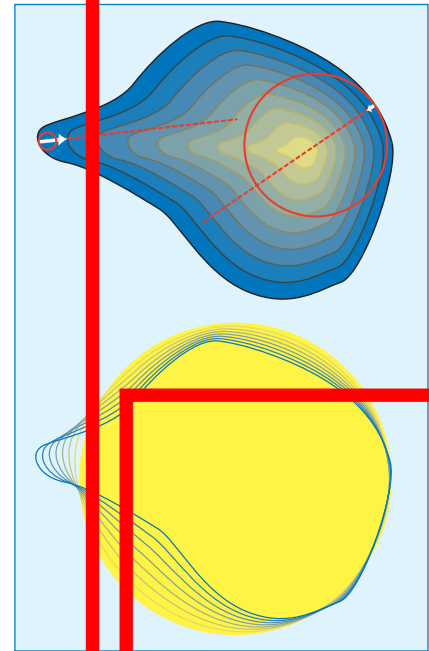
«Nous étudions un système dynamique continu dans lequel on décrit le mouvement d'une particule qui se meut sur une surface S donnée de l'espace, sans gravité. Depuis Johann Bernouilli (1667 - 1748), qui l'a enseigné à son non moins illustre élève Leonhard Euler (1707 - 1782), on sait qu'un tel mouvement suit une géodésique de la surface, c'est-à-dire une courbe dont l'accélération dans l'espace est normale à la surface, ou encore des courbes qui minimisent le chemin parcouru entre deux points voisins. Que devient cette

trajectoire, cette géodésique, lorsqu'on la prolonge indéfiniment? Est-elle périodique? À l'inverse, recouvre-t-elle toute la surface? Plus encore, prend-elle toutes les directions en chaque point? Et s'il existe des trajectoires périodiques, peut-on évaluer leur nombre en fonction de leur longueur? En 1905, Henri Poincaré a été le premier à s'intéresser à ce problème, car il s'agit du modèle mathématique le plus simple se rapprochant de celui qui décrit le Système solaire.

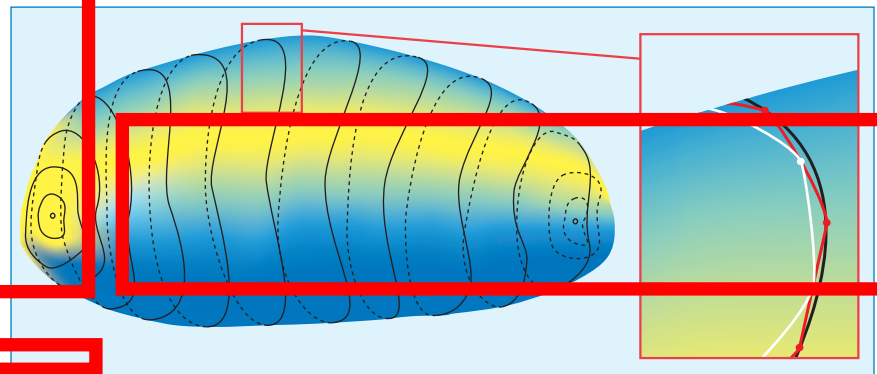
Dans le cas de la sphère, toutes les trajectoires, en nombre infini, sont périodiques et de même période : ce sont les grands cercles (les équateurs) de la sphère. Par ailleurs, d'autres surfaces, telles les surfaces de Zoll, qui ne sont pas des sphères, ont cette propriété (voir la figure 6). Pour toute surface de révolution, il y a une infinité de géodésiques périodiques. Sur une telle surface, la trajectoire d'une particule oscille entre deux parallèles. De par la révolution de la surface, une géodésique qui part d'une parallèle y reviendra après un certain temps, mais en un point qui a tourné d'un angle α par rapport au point initial. Nous sommes alors de nouveau dans le cas du billard que nous avons déjà évoqué : les géodésiques périodiques correspondent aux angles α proportionnels à 2π ; elles sont en nombre infini.

Pour les surfaces générales, on cherche d'une part à montrer l'existence de géodésiques périodiques, et à les évaluer asymptotiquement ; d'autre part, à déterminer l'ergodicité (en espace ou en phase) des trajectoires non périodiques. Posé simplement, le problème n'en demeure pas moins ardu : Poincaré écroula à prouver l'existence d'une

seule géodésique périodique. La raison essentielle de cette difficulté résultait, comme pour l'inégalité isopérimétrique, de l'absence d'un théorème d'existence de limites dans un espace géométrique de dimension infinie, ici celui des courbes fermées d'une surface. Ce n'est qu'en 1917 que George Birkhoff (1884-1944) a démontré, sans recours à un théorème d'existence d'une forme



9. LE PÉRIMÈTRE LE PLUS COURT pour une surface d'aire donnée (en bleu) est un cercle. L'une des méthodes pour le montrer consiste à déplacer les points du périmètre initial le long de leur normale d'une distance (en blanc) inversement proportionnelle au rayon du cercle osculateur (en rouge). En haut, le domaine tend vers un point, cependant, lorsque les domaines sont renormalisées, c'est-à-dire maintenues à aire constante, le domaine tend vers un cercle (en bas).



10. UNE TAPISSERIE est une famille de courbes qui «recouvrent» entièrement une surface. À partir d'une tapisserie quelconque (à gauche), on identifie une géodésique périodique en découpant les courbes en fragments d'une longueur inférieure à une distance donnée. Puis, entre deux points consécutifs, on prend le chemin le plus court (en rouge, à droite) ; sur une sphère, on prendrait les morceaux des grands cercles entre chaque paire de points. Enfin, des milieux des côtés du polygone géodésique, on trace un autre polygone strictement plus court (en blanc). La plus longue courbe d'une tapisserie économique est une géodésique périodique.

limite, l'existence d'une géodésique périodique pour une surface quelconque. Tout d'abord, il considère toutes les tapisseries de la surface, c'est-à-dire les familles de courbes allant d'une courbe réduite à un point m à une autre courbe réduite à un point m' : cette

famille de courbes recouvre entièrement la surface étudiée (voir la figure 10). Dans chacune des tapisseries, il y a une courbe de plus grande longueur, et l'on recherche une tapisserie pour laquelle la courbe de plus grande longueur est la plus petite possible.»

L'algébriste se permet une remarque : « En quelque sorte, il s'agit de rechercher une tapisserie économique. »

« À quelques subtilités près, c'est exact ! La plus grande courbe de cette tapisserie optimale est une géodésique périodique. Cela, à cause de la propriété de minimalité des géodésiques. »

L'analyste ponctue : « Reste à prouver l'existence d'une tapisserie économique. »

« C'est à qu'intervient la deuxième idée de Birkhoff : passer d'un espace de dimension infinie à un espace de dimension finie. Il a cherché à fabriquer, à partir d'une tapisserie quelconque, une deuxième tapisserie dont toutes les courbes sont strictement plus petites, à l'exception des géodésiques périodiques. Une tapisserie minimale contient toujours une géodésique périodique ; sinon, la transformation de nouveau appliquée fournirait une tapisserie dont toutes les courbes seraient plus courtes, et la tapisserie précédente ne serait donc pas minimale. Le procédé de Birkhoff

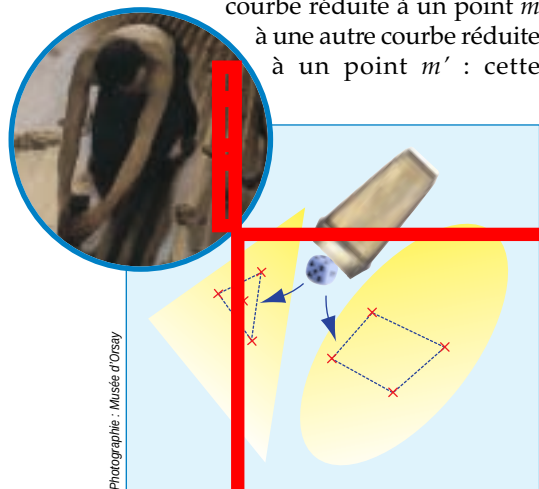
consiste à découper les courbes des tapisseries de la surface en N morceaux de longueurs inférieures à une distance donnée. Puis on remplace chaque morceau par le chemin le plus court, toujours sur la surface, joignant les deux extrémités du morceau. La deuxième opération consiste à prendre ensuite les milieux des côtés des polygones géodésiques obtenus et de les joindre en un nouveau polygone strictement plus petit, car, sur n'importe quelle surface, la longueur d'un côté d'un triangle est toujours inférieure à la somme des deux autres. Le nouveau polygone aura une longueur strictement inférieure à celle de la courbe initiale, à moins que tous les angles aux sommets soient égaux à π ; or cela arrive exactement pour une géodésique périodique. Ainsi, le raccourcissement simultané de toutes les courbes de la tapisserie entraîne la convergence de ces courbes en un point, sauf pour une d'entre elles : une géodésique périodique. Voilà comment le passage d'une dimension infinie à une dimension finie résout le problème.

Ce procédé ne fournit qu'une seule géodésique périodique, et ce n'est qu'en 1992 que J. Franks, de l'Université de Fribourg, en Allemagne, et Victor Bangert ont montré qu'il en existe une infinité. »

L'analyste ne put qu'approuver : « Admirable ! »

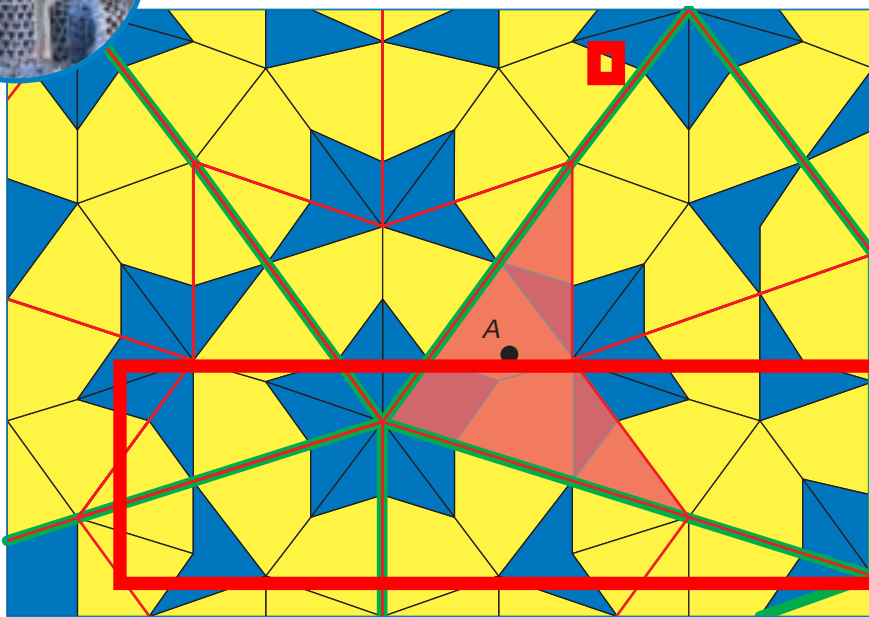
L'algébriste, après avoir fini son dessert, reprend la parole : « Face à un espace géométrique de dimension infinie, Birkhoff se tire d'affaire, certes de façon élégante, en approximant cet espace – l'espace de toutes les courbes planes d'une surface – avec un espace de dimension finie. Mais je subodore que ce n'est pas toujours possible. »

« Vrai, examinons par exemple celui des pavages de Penrose, c'est-à-dire ceux qui, construits à partir de deux types de pavés, recouvrent le plan à l'infini mais de façon nécessairement non périodique, grâce aux règles de construction (voir la figure 12). En 1975, Robinson a montré que l'on peut complètement les classer. Il a remarqué qu'à partir d'un pavage donné P , on peut grouper les pavés afin qu'ils forment des pavés plus grands de deux types seulement. Le nouveau pavage obtenu P^1 est également non périodique. De P^1 , par le même procédé, on obtient P^2 , puis P^3 , jusqu'à P^n , n tendant vers l'infini. Il s'agit ensuite de prendre un point quelconque à l'intérieur du pavage initial et de regarder dans quel

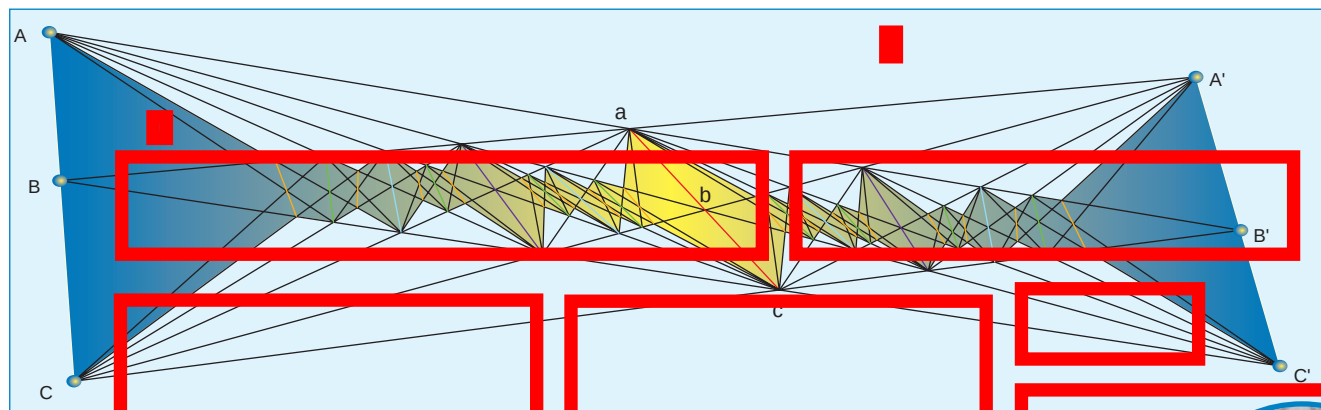


Photographie : Musée d'Orsay

11. QUATRE POINTS jetés dans un plan forment soit un quadrilatère, soit un triangle contenant le quatrième point. La probabilité d'obtenir le premier dépend de la géométrie du domaine : elle est maximale dans un domaine en ellipse, et minimale dans un domaine en triangle. La premier exemple de probabilités géométrique est le fameux problème de Buffon qui étudiait la probabilité d'une aiguille lâchée d'être à cheval sur deux lattes d'un plancher (en haut, à droite).



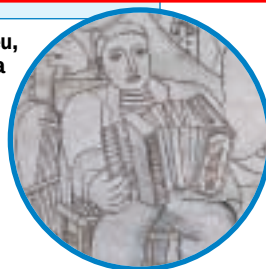
12. UN PAVAGE DE PENROSE est un pavage du plan non périodique avec deux types de pavés (en bleu et jaune). Groupés, les pavés forment un second pavage non périodique constitué de deux types de pavés plus grands (en rouge). À partir de ce pavage, on en construit un nouveau d'ordre supérieur (en vert). Dans chaque pavage, on associe à chaque type de pavés un chiffre (par exemple, 1 pour les pavés convexes et 0 pour les pavés concaves). On montre que l'espace des pavages de Penrose est similaire à celui des suites de 0 ou 1 qui coïncident à partir d'un certain rang et dans lesquelles un 1 est toujours suivi d'un 0. Ces suites sont construites en identifiant pour chaque ordre de pavage, le type de pavés dans lequel un point initial quelconque se trouve. Par exemple, le point A donne la suite 101...



Photographie : Musée Fernand Léger

13. LE THÉORÈME DE PAPPUS associe à deux triplets de points alignés (A, B, C) et (A', B', C') un troisième (a, b, c) : à partir des deux paires de triplets ainsi formées, le même procédé crée deux autres triplets. L'ensemble obtenu après l'itération à l'infini se compose de cinq

sont représentées ici : dans l'ordre, rouge, violet, bleu, vert et orange) de cette opération, est fractal et a la structure de groupe modulaire que l'on retrouve dans de nombreux domaines des mathématiques.



type de pavé il est dans P^1 , dans P^2 , jusqu'à P^n : quand, à chaque type de pavé, à un niveau d'itération donné, correspond un chiffre 0 ou 1, on obtient une série de 0 et de 1. Ainsi, les pavages de Penrose forment un espace infini qui est similaire – en mathématiques, on dit isomorphe – à celui de l'ensemble de toutes les suites infinies de 0 ou de 1 dans lesquelles tout 1 est nécessairement suivi d'un 0. Des pavages sont identiques quand leurs suites coïncident à partir d'un certain rang, et ce jusqu'à l'infini : cette caractéristique signifie que, dans un même pavage, deux points initiaux différents sont, à partir d'un certain ordre, dans des pavés de même type.»

Le dîner touchant à sa fin. Au moment du café, l'algébriste et l'analyste sont presque convaincus. Dans leur élan, un cure-dents négligé pendant l'appétif, roule et tombe sur plancher... à cheval sur deux lattes.

Des probabilités géométriques

Le géomètre reprend derechef : «Vous connaissez sans doute le problème des aiguilles de Buffon ? Il s'agit de calculer la probabilité qu'une aiguille a de tomber à cheval sur deux lattes d'un plancher. Ce problème est le premier exemple connu de probabilités géométriques. Nous ne sommes pas éloignés des itérations, puisqu'une probabilité dans la pratique signifie être une opération un nombre infini de fois. Je vous soumetts un autre exemple. Il est intéressant, car enfantin d'énoncé, et cependant en partie incompris aujourd'hui.

Lorsqu'on jette quatre points au hasard dans un domaine du plan, deux

cas sont possibles : ou les quatre points forment un vrai quadrilatère, ou bien l'un est à l'intérieur du triangle formé par les trois autres (voir la figure 12). Quelle est la probabilité d'avoir un vrai quadrilatère ?»

L'algébriste et l'analyste réfléchissent un instant, puis le premier répond : «De nouveau intuitivement, je pense que la réponse dépend de la forme du domaine ; si c'est un triangle, il sera plus difficile d'avoir un quadrilatère près des sommets que dans une autre région. À l'inverse, un quadrilatère est plus probable près du bord d'un domaine dont la frontière est une courbe lisse.»

Le géomètre poursuit : «Exact, en 1930, Blaschke a montré que la probabilité cherchée est maximale quand le domaine est une ellipse, et minimale pour un triangle. En revanche, dès que l'on s'intéresse à des dimensions plus grandes, par exemple au jet de cinq points dans l'espace tridimensionnel, et plus généralement de $n+2$ points dans un espace à n dimensions, les géomètres sont perdus. Pour la dimension trois, on montre, avec difficulté, que les ellipsoïdes sont les domaines les plus favorables à la formation d'un pentaèdre. En revanche, pour la plus mauvaise probabilité, on ne sait pas si c'est celle d'un tétraèdre. Plus grave encore : on ne connaît pas la valeur exacte de la probabilité pour le tétraèdre, c'est-à-dire que l'on ne sait pas calculer le volume moyen de l'infinité de tétraèdres contenu dans un tétraèdre donné. Les géomètres ont encore du pain sur la planche.»

Les mathématiciens s'apprêtent à partir, quand dehors, un accordéoniste passe avec son instrument.

De nouveau Pappus

«Messieurs, avant de nous quitter et puisqu'il a beaucoup été question de cercles dans mon exposé, je terminerai par où j'ai commencé, à savoir avec le théorème de Pappus. En 1993, K. Schwarz a remarqué que Pappus associe, à toute paire de triplets de points alignés sur une droite, un troisième triplet du même type. Il est donc naturel d'itérer l'opération et, à partir de chacune des deux paires créées, de construire de nouveaux triplets (voir la figure 13). Que se passe-t-il quand cette opération est répétée à l'infini ? L'ensemble des droites obtenus forme une fractale qui, de plus, a une structure de groupe modulaire. Cette structure que l'on retrouve dans tant de branches des mathématiques : la théorie des nombres, la théorie des groupes, la géométrie des surfaces, la théorie des nombres complexes, etc.

Vous voyez, messieurs, conclut le géomètre, alors que trois taxis se garent devant l'hôtel des Trois Faisans, à l'inverse de ce que vous pouviez penser, l'infini est omniprésent en géométrie et il permet même de relier nos trois disciplines.»

Et les trois de se quitter... maintenant tous convaincus que les mathématiques, toutes les mathématiques, sont la science de l'infini.

Marcel BERGER est l'ancien directeur de l'Institut des hautes études scientifiques.

Marcel BERGER, *Géométrie vivante ou l'échelle de Jacob de la géométrie*, Cassini, à paraître.

L'analyse non standard

JEAN-MICHEL SALANSKIS



En introduisant la notion d'objet insaisissable, l'analyse non standard a conféré aux infiniment petits un statut honorable et justifié des opérations qui, bien que pratiquement satisfaisantes, posaient des problèmes de principe.

Qui n'a pas été amusé à calculer sur sa calculatrice e^x pour une valeur de x positive et assez grande ? Le résultat affiché est alors zéro. De même, lorsque nous appliquons la fonction racine carrée à un nombre positif très grand, nous constatons qu'après un certain nombre d'applications successives, nous obtenons 1. En gros, la calculatrice a calculé $e^{-\infty} = 0$ et, en fin de compte, $\sqrt{1+\epsilon} = 1$. Étranges égalités qui nous paraissent être fausses ou exprimer de manière inadéquate une propriété de limite. Comme la calculatrice calcule, néanmoins, et comme nous avons spontanément traduit son calcul par des écritures reflétant ce qu'elle fait, il semble bien qu'elle traite certains nombres comme infiniment grands ou infiniment petits, et en tire des conséquences pour le calcul.

Ces conséquences évoquent une algèbre des calculs sur les limites. Par exemple, nous avons appris que l'inverse d'une fonction tendant vers l'infini tend vers zéro. Bien sûr, nous l'avons appris de façon correcte et nous avons été mis en garde contre l'équation illicite : $1/\infty = 0$. Toutefois, en pratique, quand nous résolvons des exercices,

nous passons parfois mentalement par l'usage de ces équations illicites.

L'analyse non standard est une méthode mathématique rigoureuse qui justifie ces équations illicites : elle autorise le raisonnement mathématique à faire usage des nombres infiniment petits ou grands réputés inexistant. Georg Cantor nous a appris à poser, comme des objets respectables de la théorie des ensembles, des cardinaux infinis associés à des ensembles infinis ; de même, l'analyse non standard a défini des nombres réels infiniment grands ou infiniment petits.

En examinant l'histoire de l'infini et de l'infiniment petit ou grand, nous ferons ressortir cinq étapes : le calcul infinitésimal des pionniers où l'on néglige les grandeurs « très petites » ; la conception de l'infini issue de la théorie des ensembles ; l'analyse non standard du mathématicien américain Abraham Robinson, où l'infiniment grand et l'infiniment petit apparaissent par l'intermédiaire de « l'axiome du choix » ; la théorie des ensembles internes, une extension de la théorie des ensembles qui permet d'apercevoir des infiniment petits ou grands au

1. LES MEASUREURS, tableau du XVII^e siècle attribué à Hendrick Van Balen (ci-dessus), illustre l'apologue du poète romain Horace : « Il y a une mesure pour toute chose. » Quelle que soit la précision des mesures, les quantités infiniment petites définies par l'analyse non standard resteront à jamais hors de portée.

milieu des nombres usuels ; enfin l'analyse non standard d'inspiration « constructiviste » suggérée par le mathématicien suisse bourgeois récemment disparu, Georges Reeb, qui conduit à envisager des entiers infiniment grands, au-delà des entiers saisissables.

La contradiction infinitésimale

Revenons à l'origine du calcul infinitésimal. Étant donné la relation entre variables $y = x^2$, nous voulons calculer l'accroissement dy de la variable y lorsque x est modifié de la quantité infiniment petite dx . Nous calculons alors $y + dy = (x+dx)^2$ soit, $dy = 2x dx + dx^2$; de là, nous passons à $dy = 2x dx$, en négligeant dx^2 comme carré d'un infiniment petit. Ainsi, dans le calcul infinitésimal, nous considérons comme licite d'affirmer que la différentielle de x^2 est $2x dx$.

Cela paraît une erreur, une «équation imparfaite», mais au fond, le concept de quantité infiniment petite n'étant pas clairement défini, nous ne savons pas très bien jusqu'à quel point il y a erreur. «Infiniment petit», pour les mathématiciens qui élaborent ce calcul, cela signifie «plus petit que toute quantité finie donnée». Chacun s'accorde, au fond, à considérer que la notion même de quantité réelle inclut la propriété que l'on dénomme «archimédianité» : tant qu'une quantité positive n'est pas nulle, on peut toujours trouver une autre quantité, du type $1/n$ (où n est un entier naturel non nul), plus petite qu'elle. Une quantité ne peut alors pas être «plus petite que toute quantité finie donnée», les nombres du type $1/n$ étant de (petites) quantités données. Autrement dit, un réel strictement positif ne peut être infiniment petit au sens strict : on en produit facilement un plus petit.

Pourtant, les pionniers ont développé un calcul où l'on utilise de telles quantités et où l'on en néglige certaines. Toutefois, en dépit de problème de principe, le calcul s'applique à la fois au sein des mathématiques, puisqu'il permet de caractériser localement les courbes, notamment la pente de la tangente en un point, de calculer les longueurs, des aires, etc. – et hors des mathématiques – puisqu'il autorise la modélisation des phénomènes, par intégration des «équations différentielles» établies sur des données élémentaires de la situation. Il doit donc y avoir une raisonnable dernière usage des infiniment petits et des «équations imparfaits» et la «pensée infinitésimale» des pionniers était une démarche mal fondée, mais qui allait dans la bonne voie.

La reformulation ensembliste

La pensée ensembliste contemporaine, c'est-à-dire la conception de l'infini issue de la théorie des ensembles, supprime les contradictions en traduisant le discours des pionniers dans son nouveau langage. Ainsi, l'équation précédemment citée, $dy = 2x dx + dx^2$, devient $f(x+dx) - f(x) = 2x dx + dx^2$. « dx » étant le nom de la fonction qui à x associe son carré x^2 . On la considère comme une équation exacte où dx est une quantité ordinaire, finie. On divise alors les deux membres par dx et l'on aboutit à : $dy/dx = (f(x+dx) - f(x))/dx = 2x + dx$. On constate maintenant que, «lorsque

dx tend vers 0, le taux d'accroissement $(f(x+dx) - f(x))/dx$ tend vers $2x$ ». Ce n'est pas $2x$ qui est assimilable à dy , c'est dy/dx qui «tend» vers $2x$. Cela suppose que l'on définit la notion de «tendre vers», ce qu'ont fait Karl Weierstrass, Georg Cantor et Richard Dedekind.

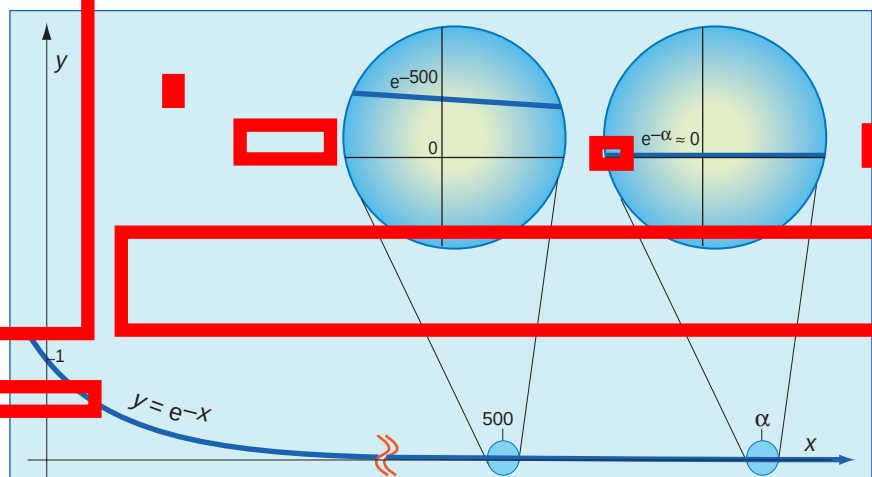
En substance, dire que le taux d'accroissement tend vers le nombre 1, c'est dire que l'on peut rendre sa différence avec 1 inférieure en valeur absolue à une valeur strictement positive arbitraire, quelle qu'elle soit, à condition de choisir un dx inférieur en valeur absolue à une valeur frontière strictement positive adaptée à la première. Cette définition en termes de contrôle satisfait les mathématiciens. Toutefois, le théoricien des fondements remarque, qu'avec cette définition, une propriété de convergence vers une limite apparaît comme établie à l'issue de la résolution d'une infinité de petits problèmes posés à chaque degré de petitesse que l'on veut réaliser. Plus sobrement, elle est la propriété d'un objet supposé bien défini qu'est une fonction, c'est-à-dire une application de l'ensemble des réels dans lui-même, ou encore, selon la conception ensembliste, un ensemble fabriqué à partir d'ensembles infinis.

Ainsi, l'interprétation «moderne» du discours des pionniers de l'analyse infinitésimale le sauve de l'équation imparfaite, qui était l'un des problèmes, mais pas de la référence à une quantité douteuse. Seulement celle-ci, au lieu d'être infiniment petite, est infinie «tout court». À vrai dire, une difficulté de l'ordre de celle de l'équation imparfaite surgit aussi, mais elle

ne le fait que pour un spécialiste soucieux des problèmes de fondement. En acceptant les ensembles infinis actuels de points, le mathématicien est obligé d'accepter une «logique de l'infini», de prolonger l'usage de ses règles logiques de discours de manière systématique à l'infini.

Notamment, il décide que la phrase exprimant que telle fonction tend vers l en zéro, phrase qui évoque une infinité de problèmes paramétrés par un nombre réel strictement positif (pour tout réel strictement positif ε , il existe un réel strictement positif δ tel que, quel que soit x , le fait que la valeur absolue de x soit inférieure à δ implique que la valeur absolue de $f(x) - l$ est inférieure à ε) est ou bien vraie ou bien fausse, selon le principe du tiers exclu. (Le principe du tiers exclu, voir *L'infini, pierre de touche du constructivisme*, pages 74 à 81, énonce que si l'on considère une affirmation P portant sur des objets mathématiques, alors soit P soit non- P est vraie, et l'on exclut toute autre possibilité.)

Or cela ne va pas sans difficulté théorique, ce serait-ce que parce que l'on est amené à formaliser les mathématiques et, par suite, à adopter un concept de déductibilité plus essentiel que le concept de vérité, concept pour lequel, on le sait depuis la démonstration par Gödel de son fameux théorème d'incomplétude, le tiers exclu ne vaut présentement pas (en 1930, Gödel a montré que dans tout système formel assez riche, on peut construire une proposition P telle que ni P ni non- P ne sont déductibles dans le système formel : une proposition «indécidable»).



2. LA FONCTION e^{-x} décroît très rapidement lorsque x augmente. Toutefois, si l'on prend un x fini, par exemple 500, en grossissant la courbe on notera une différence avec 0. En revanche, si l'on choisit un α infiniment grand, même en grossissant la courbe, la différence restera insaisissable. Selon cette intuition infinitésimale, on écrira $e^{-\alpha} \approx 0$.



La théorie des ensembles

C'est principalement Georg Cantor (ci-contre) qui a eu l'idée d'introduire en mathématiques la notion d'ensemble quelconque d'objets absolument quelconques et de démontrer des résultats généraux sur ces d'ensembles. La conception ensembliste stipule que tous les objets et toutes les notions des mathématiques peuvent être exposés en termes d'ensembles, de sorte que la «théorie des ensembles» devient la théorie sur laquelle se fondent toutes les mathématiques.

Les raisonnements sur les ensembles que l'on fait naturellement en partant de l'intuition de ce qu'est une collection et de ce qu'est, pour un objet, que d'appartenir à une telle collection, conduisent à des paradoxes (antinomie de Russell, paradoxe de Cantor, etc.). Les mathématiciens du début du xx^e siècle ont alors été conduits à formaliser la théorie des ensembles, à la présenter comme une théorie logique abstraite gouvernant l'emploi de la relation d'appartenance notée $\in$, symbole qui, désormais, n'a plus un sens intuitif, mais qui est soumis à des règles d'emploi consignées dans des axiomes. La théorie formelle ainsi obtenue, et qui constitue la référence implicite des mathématiques depuis près d'un siècle, est la théorie ZFC (Zermelo-Fraenkel avec axiome du choix).

Dans cette théorie, on définit la notion d'ensemble infini et d'ensemble fini comme Cantor en avait déjà eu l'idée : est infini un ensemble qui peut être mis en bijection avec une partie de lui différente de lui (ainsi, l'ensemble N des entiers naturels est mis en correspondance bijective avec l'ensemble P des nombres pairs par l'application $n \mapsto 2n$) : est fini un ensemble qui n'est pas infini. À tout ensemble, même infini, est associé un «cardinal», qui «compte» le nombre de ses éléments : ainsi N a «aleph 0» éléments, comme $\{0,1\}$ en a deux. Une application $f : E \rightarrow F$, de source l'ensemble E et de but l'ensemble F , est définie comme un triplet (E, G, F) où $G \subset E \times F$ est constitué de couples du type (x, y) avec $x \in E$ et $y \in F$, le graphe G ayant été choisi tel qu'à tout $x \in E$ peut être associé un unique $y \in F$ tel que $(x, y) \in G$ (le y en question s'appelle alors image de x). Si E est infini, le graphe G est alors, lui aussi, infini.

La reconquête de l'indéfini

Cette situation laissait soupçonner que l'on aurait pu employer le formalisme issu de la théorie des ensembles et les formalismes associés pour justifier le mieux possible le discours des pionniers plutôt que pour le disqualifier, comme le fait dans une certaine mesure la conception ensembliste que nous venons d'évoquer. Or, dans les années 1960, une telle réhabilitation a été proposée par le mathématicien américain Abraham Robinson, spécialiste de la branche alors triomphante de la logique mathématique nommée théorie des modèles.

La réinterprétation de Robinson du calcul et du discours des pionniers est en substance la suivante : elle sauve l'idée d'infiniment petit, en demandant aux mathématiciens de penser qu'il y a plus de nombres réels qu'ils ne le croyaient. Ou plutôt, les réels connus sont plongés dans un système plus vaste, qu'on peut nommer système des «hyperréels», et contenant des éléments strictement positifs plus petits que tout élément strictement positif du système de base, qu'on dénommera désormais système des

réels standard. Toute la subtilité de l'opération réside alors dans ceci que le nouveau système de nombres est une «extension élémentaire» du premier, c'est-à-dire qu'il en est jusqu'à un certain point un double, un élargissement non essentiel.

En d'autres termes, les nouveaux objets se peuvent être caractérisés en fonction des réels standard et des moyens discursifs qui les accompagnent. Ils sont des entités supplémentaires, indéterminées, simplement choisies de façon à procurer, à toute une classe de relations définies sur le système de base, des objets-limite dominant tous les objets standard suivant ces relations. Pour donner un exemple, disons que la relation μ a la propriété suivante : pour tout ensemble fini $\{x_1, \dots, x_n\}$ de réels standard, on peut trouver un réel standard y majorant chacun des x_i . Le processus d'élargissement de Robinson nous garantit alors dans le système $^*\mathbb{R}$ des hyperréels l'existence d'un μ tel que $x < \mu$ pour tout x standard, c'est-à-dire d'un μ «infiniment grand».

Pour construire les hyperréels, Robinson se fonde sur l'axiome du choix. Cet axiome – introduit explicitement par Zermelo en 1908 – pose,

pour résumer de manière commode, qu'il «existe» une fonction choisissant exactement un élément dans tout ensemble d'une famille infinie d'ensembles non vides donnés. Autrement dit, si l'on considère une famille d'ensembles non vides $(X_i) - i$ variant dans l'ensemble I des indices –, il y a une fonction qui «prend» exactement un élément dans chaque ensemble X_i , pour chaque indice i de l'ensemble d'indices I : à chaque indice i on associe un élément choisi dans X_i . Toutefois, le «choix» ainsi exprimé par la fonction censée exister n'est ni déterminé ni déterminable au-delà de l'axiome qui pose son existence. On ne peut pas décrire cette fonction autrement que par la sélection qu'elle opère : on ne peut pas l'exprimer par rapport aux objets déjà connus de la situation où elle intervient.

Néanmoins, la simple existence de la fonction est intéressante pour le mathématicien : elle lui garantit que certaines constructions ensemblistes sont possibles, que certains objets commodes peuvent être admis. Par exemple, elle permet l'introduction d'une base à laquelle rapporter tous les vecteurs de n'importe quel espace vectoriel.

Malgré les services rendus par cet axiome, plusieurs mathématiciens l'ont contesté au début du XX^e siècle. Ils estimaient que l'axiome de choix affirmait une existence trop idéale et trop peu contrôlée. Robinson justifie *a posteriori* cet axiome en montrant que le caractère peu contrôlé et indéterminé des objets introduits par l'axiome du choix peut être intéressant en lui-même.

L'axiome du choix amène la mathématique à jouer avec des objets dont elle n'a pas suivi le protocole de création, des objets qui s'ajoutent à son paysage, qui lui sont uniquement rattachés par une propriété miraculeuse. Les infiniment petits de Robinson sont justement de tels objets : leur statut se définit en termes de l'infinité des objets standard, mais ils sont pour le reste indéterminés. Robinson justifie le discours des pionniers en donnant un crédit ensembliste irréfutable aux infiniment petits et du coup, aux infiniment grands), mais en ajoutant à la figure de l'infini de totalité celle d'un infini-indéfini, d'un infiniement indéfini, ou quelque sorte (ou plutôt, cette figure, il ne l'ajoute pas, il la met en relief). C'est d'ailleurs en termes de cette figure que l'équation imparfaite est aussi interprétée : on écrira avec Robinson $dy/dx \approx 2x$, le signe $\approx$ exprimant que

la différence est infiniment petite, c'est-à-dire qu'elle est inférieure en valeur absolue à tout nombre standard. De manière philosophique, on peut dire que dy/dx ne diffère de $2x$ que d'un «indéfiniment petit».

La syntaxe de l'indéfini

Dans les années 1970, le mathématicien américain Edward Nelson a proposé une nouvelle théorie des ensembles qu'il nomme théorie interne des ensembles, ou IST (pour *Internal Set Theory*). La théorie IST étend la théorie formelle des ensembles usuels, la théorie Zermelo-Fraenkel avec choix, ou ZFC. Cette nouvelle théorie des ensembles fait intervenir en plus des symboles fondamentaux de l'égalité (=) et de l'appartenance ($\in$), un nouveau prédicat nommé st (comme standard). À l'inverse de $=$ ou de $\in$, qui sont des symboles à deux places ($x = y$, $x \in y$), le prédicat st est un prédicat à une place, c'est-à-dire qu'il affirme une propriété sur un objet (on écrit $st(x)$ pour x est standard). Cela signifie qu'on distinguera désormais, parmi tous les ensembles de l'univers des ensembles, ceux qui sont standard et ceux qui ne le sont pas (c'est un peu comme si la télévision ensembliste était passée du noir et blanc à la couleur, nous révélant des distinctions nouvelles).

Le prédicat st ne peut être réduit aux notions ensemblistes classiques. Sa signification est régie par trois nouveaux schémas d'axiome qui ont pour effet de lui faire exprimer plus ou moins la notion d'objet assignable, saisissable, logiquement connu dans le paysage ensembliste. Ce prédicat permet de distinguer les entités qui tiennent le rôle des infiniment petits chez Robinson. Ainsi, dans la mathématique nelsonienne, les réels infiniment petits non nuls sont des réels ne satisfaisant pas le prédicat st , des réels non standard.

Cette approche de Nelson met en évidence un rapport entre le fini et le

standard. Elle laisse entendre que les objets non standard sont là «à cause» de l'infiniment petit, et que, en un certain sens, tout ce qui est assignable tient dans le fini. Bien que la théorie IST qu'il propose soit une extension de la théorie classique, et puisse donc être considérée comme porteuse du même message logique et ontologique, elle nous engage à voir autrement dans l'univers des ensembles la répartition du fini et de l'infini, et leur rapport.

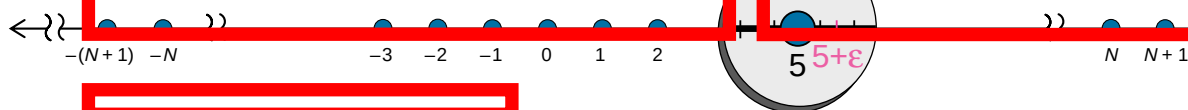
Les deux enseignements de la théorie IST sur le fini et l'infini sont les suivants : premièrement, tout ensemble infini contient un élément non standard, deuxièmement il y a un ensemble fini qui contient tous les objets standard de l'univers. Ces deux assertions sont des théorèmes qui découlent à peu près directement des axiomes spécifiques imposés au prédicat st par lesquels E. Nelson a complété la théorie des ensembles.

La première assertion se comprend au fond très bien : dans un ensemble que nous prenons pour infini, comment pourrait-il ne pas apparaître des objets échappant à notre contrôle logique, c'est-à-dire des objets non assignables, non standard ? La conclusion que tire le théorème semble être une simple conclusion raisonnable, dictée par la connaissance de notre finitude (bien que, naturellement, une telle considération métaphysique ne figure nullement parmi les justifications du dispositif de Nelson). Les justifications, à vrai dire, il n'en donne pas, c'est en cela notamment que consiste la démarche axiomatique).

La seconde assertion est plus surprenante : parmi les objets standard figurent, Nelson prend soin de nous le faire remarquer, tous les objets définissables dans la théorie à base des ensembles (classique) au moyen d'une formule qui les caractérise, c'est-à-dire tous les objets familiers du mathématicien, par exemple $\mathbb{R}$ l'ensemble des nombres réels ou $L^2(\mathbb{R}, \mathbb{C})$ l'espace des fonc-

tions complexes de carré sommable sur $\mathbb{R}$ (les fonctions de $\mathbb{R}$ sur $\mathbb{C}$ dont l'intégrale sur $\mathbb{R}$ du module au carré est finie) ; mais aussi bien sûr, les entiers «connus», 0, 1, 2, et ainsi de suite. La collection des objets standard contient donc toute une faune, apparemment infinie, d'objets qui sont eux-mêmes susceptibles d'être idéaux, c'est-à-dire que l'on peut penser, mais non construire, parce qu'ils se situent par principe au-delà de toute saisie immédiate, comme $\mathbb{N}$ ou $\mathbb{R}$ et du coup infinis. Néanmoins, ce que dit le théorème, c'est que la collection totale des objets de cette sorte peut être intégrée à un ensemble fini, et donc, en un sens, demeure justiciable d'un traitement analogue à celui des objets finis. Par exemple, on pourra former la somme finie d'une classe de nombres réels contenant tous les réels standard.

Cette affirmation contredit l'intuition qui résulte de la construction de Robinson : pour lui, les nombres réels standard sont les nombres usuels, et l'on n'obtient un système avec des nombres infiniment petits et grands qu'en ajoutant de nouveaux nombres à ces nombres usuels ; il y a donc une infinité continue de réels standard. Le point de vue de Nelson est autre : nous avons des nombres infiniment petits uniquement parce que nous reconnaissons que, parmi les nombres réels, figurent forcément des éléments insaisissables, non assignables, parce que nous entrons dans la logique de cette distinction entre assignables et non assignables. C'est cette distinction qui différencie fonctionnellement les nombres réels entre eux et qui dégage la possibilité «grammaticale» de l'infiniment petit ou grand. Elle fait aussi apparaître, à l'intérieur des réels, une collection de réels standard qui se laisse inclure dans une partie de $\mathbb{R}$ que l'on peut traiter comme une partie finie.



3. LES NOMBRES RÉELS INCLUENT les nombres entiers, positifs et négatifs, les nombres rationnels et les nombres irrationnels. Ces nombres réels peuvent être représentés par des points sur une droite nommée la droite réelle. L'analyse non standard nous conduit à envisager, autour de 0, tout un «halo» de nombres non standard infiniment petits. Plus généralement, chaque nombre réel usuel, comme $\sqrt{2}$ ou 5 sur le schéma, est environné d'un tel halo, consti-

tué de nombres de type $5 + \varepsilon$ où ε est un infiniment petit. Mais l'analyse non standard nous fait voir aussi, sur la droite réelle des nombres non limités, comme N , $-N$, $N + 1$, $-(N + 1)$ sur le schéma, qui sont supérieurs à tout nombre standard (infiniment grands positifs) ou inférieurs à tout nombre standard (infiniment grands négatifs) ; néanmoins aucun de ces nombres ne peut être dit «compter» l'infinité des points de la droite.

Dans le contexte de la théorie dénommée IST, gardons à l'esprit que fini ne signifie pas intuitivement fini, absolument fini, fini au sens des ordinateurs ou même de l'intuition ordinaire des nombres entiers. Fini a le sens que lui donne la théorie des ensembles ZFC, c'est-à-dire le sens formel de ce qui peut être mis en bijection avec une partie propre de soi. Un ensemble fini en ce sens peut ne pas paraître fini à notre intuition. Ainsi, lorsque μ est un entier infiniment grand, le segment $[0, \mu]$ est infini à nos yeux, mais fini du point de vue des manipulations mathématiques. C'est toute la difficulté qu'avaient déjà explorée les pionniers de la théorie des modèles et que Robinson avait exploitée pour faire émerger des réels infiniment petits. On peut, si l'on désire se raccrocher à une signification intuitive absolue du fini, nommer «hyperfini» ce fini qui n'est fini que formellement, au sens de la théorie des ensembles. Les objets standard restent dans l'orbe de l'hyperfini.

Assignable et constructif

La manière dont la théorie IST relie l'intention justificatrice de l'analyse non standard à l'égard du discours des pionniers du calcul infinitésimal et la question générale du fini et de l'infini prépare et clarifie l'étape suivante. Celle-ci consiste à mettre directement en rapport l'analyse non standard et les idées «constructivistes». Le constructivisme accorde une importance décisive à la distinction entre ce qui peut être effectivement construit, ou devrait pourvoir l'être en principe, et ce qui se situe au-delà, ce qui est «idéal». En effet, nous commençons à le comprendre, ce qui permet d'avoir des nombres infiniment petits dans la mathématique du continu sans engager les calculs sur les nombres réels dans des paradoxes, c'est de disposer d'un statut d'entité non assignable pour certains objets, certains nombres.

Ce statut, Robinson l'obtient en faisant jouer ensemble deux modèles de la même théorie dont l'un est une extension élémentaire de l'autre, et E. Nelson l'obtient en installant dans le langage une notion d'entité assignable. Le résultat de cette démarche est de lier l'assignable au fini, le non-assignable à l'infini.

Pourquoi, dans ces conditions, ne pas tenter de se fonder sur ce qui semble être la figure contemporaine de la notion d'entité assignable, et

RÉCURRENCE INTERNE ET EXTERNE

Le principe de récurrence nous dit que si une propriété $P(n)$ est vraie pour $n = 0$ et si elle est héréditaire (la vérité de $P(n)$ entraîne celle de $P(n + 1)$, alors elle est vraie pour tout n . Ce principe reste valable dans l'analyse non standard «nelsonienne», mais seulement pour les propriétés dites internes, qui ne font pas intervenir la notion d'objet standard (le prédicat σ). Pour une propriété externe, par exemple sur la nouvelle notion, sachant que $P(0)$ est vraie et que, pour le standard, $P(n)$ entraîne $P(n + 1)$, on pourra seulement conclure que $P(n)$ est vraie pour tout entier standard, la valeur de vérité de la propriété P pour les entiers infiniment grands restant parfaitement indéterminée.

poser qu'est assignable ce qui est représentable par l'informatique? Parce que cette notion est trop technique et contingente : elle semble dépendre d'un état actuel de la puissance des machines et de la taille des mémoires, par exemple. Les notions plus théoriques d'objets constructibles au sens de Brouwer (c'est-à-dire d'objets dont l'élaboration apparaît comme réalisable à l'intuition qui accompagne la succession indéfinie des nombres entiers) ou de calculabilité (dans les années 1940, des mathématiciens ont introduit cette notion en cherchant à préciser la notion intuitive de calcul) fournissent exactement ce qu'il est naturel de rechercher : des concepts d'entité assignable assez généraux pour dépasser les limites contingentes de la technique, mais qui n'entrent pourtant pas dans un type d'idéalité comparable à celui des ensembles infinis de nombres $\mathbb{N}$ ou $\mathbb{R}$. D'où l'idée de fonder un discours d'analyse non standard dont l'équation de base serait $\text{standard} = \text{constructif}$.

Il revient à Georges Reeb d'avoir eu l'idée générale et systématique d'une telle stratégie, et également d'avoir compris que c'était toujours de cela qu'il s'agit dans l'analyse non standard historique.

Il s'agit de comprendre que tout ce qui est nécessaire au développement de l'analyse non standard est un entier infiniment grand. On doit s'être également muni, bien entendu, d'une règle du jeu instruisant sur la façon de raisonner et de calculer avec lui, sur le type d'entités que l'on peut construire à partir des nombres entiers divisés en finis et infinis dont on dispose ; enfin, on doit se munir d'un bon motif «phi-

losophico fondationnel» pour ajouter crédit à tout ce montage intellectuel faisant intervenir un nombre entier infiniment grand. Dans ces conditions, on peut introduire un système de nombres comportant des infiniment petits et grands qui permet de faire de l'analyse. Le texte qui formule le programme d'une mathématique non standard fondée sur l'entier infiniment grand tout en interprétant ce dernier comme entier non constructif est l'article *La mathématique non standard, vieille de soixante ans?* de Georges Reeb, publié en 1979.

Pour Reeb, nous devons reconnaître que la posture rassemblant l'ensemble des entiers du formaliste, le $\mathbb{N}$ obéissant au principe du tiers exclu, que l'on définit comme le plus petit ensemble infini en théorie des ensembles, n'a aucune raison de contenir seulement les entiers intuitifs, les entiers qui viennent à l'esprit au fil de l'énumération 0, 1, 2, etc. C'est ce que les inlassables objections de Brouwer au formalisme naissant, au début de XX^e siècle, auraient montré sans ambiguïté. Brouwer a critiqué la mathématique ensembliste et formelle qui s'affirmait sous ses yeux, et a proposé une autre orientation, qu'il nomma «intuitionniste». Selon Reeb, Brouwer avait raison de considérer qu'une preuve formelle de la fausseté de la proposition «quel que soit l'entier n , la propriété $P(n)$ est vraie» – le plus probablement obtenue au moyen d'un raisonnement par l'absurde – n'apportait pas la garantie que l'on sache indiquer la construction d'un entier N tel que non $P(N)$ soit vraie. Toutefois, dans ces conditions, l'entier n tel que non $P(n)$ est vraie, dont la théorie affirme l'existence via la preuve formelle, semble avoir le statut d'un entier «au-delà» des entiers effectifs : un tel entier est réputé «exister» parce qu'une phrase de la théorie affirmant son existence est démontrée, mais on ne peut l'identifier avec aucun entier que nous sachions construire. Il est donc «plus grand» que tous ceux-ci.

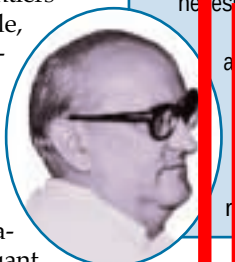
Dans le raisonnement de Reeb, l'entier infiniment grand est introduit par un étrange raisonnement, qui accepte l'existence de la théorie formelle des entiers et envisage même la mythique collection de nombres satisfaisant cette théorie, mais qui, en même temps, débat sur ce que cette collection peut et doit contenir comme éléments, en comparant la notion formelle d'entier à sa notion intuitive. Le résultat

auquel on arrive (que le putatif $\mathbb{N}$ doit contenir un entier infiniment grand) a beau être nécessaire au gré du raisonnement tenu, il ne peut pas être assumé dans le formalisme : si le prédicat « infiniment grand » était disponible dans le formalisme, on obtiendrait un paradoxe en appliquant le schéma d'induction à la classe des entiers finis ou à la classe des entiers infiniment grands. Par exemple, on dirait de substance que l'ensemble des entiers infiniment grands a un plus petit élément α , donc $\alpha - 1$ est fini, mais alors $\alpha - 1$ est aussi, d'où contradiction.

Tous les montages de l'analyse non standard nouvelle, jouant sur un entier infiniment grand sans passer sous les fourches Caudines de $\mathbb{R}$ ou de la théorie des ensembles, consistent à transformer le formalisme de la théorie des entiers pour que disparaissent ces paradoxes. L'usage du principe de récurrence devra, par exemple, être limité à des prédicats non bâtis au moyen de la notion d'entier « fini », « standard » ou « limité ». Il résulte alors de tels dispositifs que les entiers infiniment grands apparaissent comme une classe « insituable », intervenant sur la droite entière « après » ceux qui sont finis. Ils se distinguent des entiers standard, que l'on interprète désormais comme les « entiers constructifs », disons les entiers que l'on peut construire avec des ordinateurs, en imaginant ceux-ci suffisamment puissants.

L'infiniment grand

L'intention de ces théories non standard est ainsi de saisir l'infiniment grand comme l'excédent insituable au-delà du constructif. Ces théories idéalisent l'intuition des rapports entre entiers constructifs et entiers énormes, inaccessibles, rapports qui sont « concrètement » assez imprévisibles. Comme l'explique Jacques Harthong, les entiers constructibles ou accessibles sont nécessairement définis relativement à un langage de programmation choisi comme logiciel de production ou d'écriture de références. Ainsi le nombre $10^{\wedge} (10^{\wedge} (10^{\wedge} (10^{\wedge} 10))$, où $a^{\wedge} b$ désigne a^b , peut être calculé par un programme très simple, long d'une vingtaine de pas sur une calculatrice programmable : ce nombre est néanmoins gigantesque.



Des infinis aux visages différents

La notion d'infini a évolué avec les conceptions de l'analyse non standard. Abraham Robinson (à gauche) se fonde sur les ensembles infinis de la théorie des ensembles et sur l'axiome du choix pour produire un ensemble des hyperréels dans lequel on trouve, en sus des nombres réels connus, des nombres infiniment grands et infiniment petits. En reformalisant la théorie des ensembles, Edward Nelson (à droite) introduit la notion d'entité standard : les infiniment petits ou grands existent non pas comme nouveaux objets « ajoutés », mais parce que nous reconnaissons que, parmi les réels, existent nécessairement des éléments insaisissables.

Si Nelson relie l'infini à l'insaisissable et le fini à l'assignable, il revient au mathématicien Georges Reeb (à gauche) de fonder l'analyse non standard sur une sensibilité constructiviste. Autrement dit, à la suite de Reeb, on essaie de rapprocher le concept d'infini du concept d'incalculable. Les entiers standard sont alors les entiers « constructifs », ceux qu'on peut calculer à l'aide d'ordinateurs ou qui apparaissent au fil de l'énumération naïve.



Un nombre de même taille, mais n'ayant pas la même structure simple n'aurait aucune chance d'être programmable dans le même langage (à défaut de la structure, il ne resterait que la ressource d'écrire ses chiffres, mais il faudrait un nombre de symboles supérieur à celui des particules de l'Univers, et un temps dépassant la capacité de survie de tout vivant). La distribution des nombres « accessibles à la calculatrice programmable » est non régulière, tributaire des particularités du matériel symbolique et des opérations de base de la calculatrice considérée. La distribution des nombres qui apparaissent comme constructibles au sens de telle calculatrice programmable est donc caissée.

Dans le nouveau point de vue, l'analyse non standard est une systématisation théorique, indépendante de toute machine et de tout langage particuliers, des phénomènes que nous citons en exemple au début de l'article, lorsque nous remarquons que, pour les calculatrices, certains entiers finis jouent le rôle de l'infini. Si l'on comprend bien le concept de nombre fini accessible, constructif, on remarque que la pensée arithmétique contemporaine rend plausible une théorisation qualitative des entiers énormes qui leur donne un statut d'infini. Toutefois, dès que l'on dispose d'un contexte théorique d'arithmétique avec des entiers infiniment grands, on peut introduire un calcul sur des rationnels à dénominateur infiniment grand qui permet de retrouver, dans une certaine mesure, la mathématique du continu à laquelle nous sommes habitués. Alors, nous pourrions

faire de l'analyse non standard en justifiant les infiniment petits et les équations imparfaites des pionniers.

Pour y parvenir, il faut, encore une fois, faire valoir certains objets comme non assignables ou relativement mal assignés, ces entiers énormes cette fois. Selon les stratégies intellectuelles, la justification de ce caractère non assignable peut être plus ou moins subtile et utiliser soit la philosophie intuitionniste de Brouwer, soit le théorème d'incomplétude de Gödel, soit un examen précis des possibilités des langages de programmation.

Les conceptions successives de l'analyse non standard ont en commun d'avoir tenté de faire jouer en plus, à côté ou à la place de la notion cantorienne d'infini de totalité, une notion d'indéfini, d'entité non assignable. Cette notion tend jusqu'à un certain point des services techniques comparables à ceux que rendaient les notions infinitésimales des pionniers. En tout cas, elle justifie *a posteriori* un usage intuitif de l'infinitésimal auquel les mathématiciens et les physiciens sont restés attachés.

Jean-Michel SALANSKIS est professeur de philosophie des sciences à l'Université Paris X (Nanterre).

J.-M. SALANSKIS, *Le constructivisme non standard*, Presses du Septentrion, 1999.

M. PANZA et J.-M. SALANSKIS, *L'objectivité mathématique*, Dunod-Masson, 1995.

R. LALEMENT, *Logique, réduction, résolution*, Hermès, 1990.

F. DIENER et G. REEB, *Analyse non standard*, Hermann, 1989.

Les géométries non commutatives

DANIELA BIGATTI

Les espaces d'opérateurs de dimensions infinies étudiés à l'aide des géométries non commutatives servent à élaborer de nouvelles théories en physique.



CERN PHOTOS

Les espaces qui nous sont le plus familiers, ceux qui correspondent à notre expérience quotidienne, sont les espaces euclidiens. Au III^e siècle avant notre ère, Euclide définit les notions de point, de droite, de distance et d'angle. Il énonce les «règles», ou postulats, auxquels les objets présents dans ces espaces obéissent. Cette géométrie correspond si bien à notre monde que, pendant plus de 2 000 ans, personne ne pense à la remettre en cause.

Cependant, à partir du XVIII^e siècle, des mathématiciens tels que Hugo Riemann ou Nikolai Lobatchevski s'aperçoivent qu'ils peuvent faire de la géométrie dans des espaces où l'un au moins des postulats d'Euclide n'est plus vrai. Par exemple, Lobatchevski propose de considérer des espaces où, par un point donné, passe une infinité de droites parallèles à une droite donnée (dans les espaces euclidiens, il n'y en a qu'une) : il fonde ainsi une «extension» de la géométrie euclidienne, nommée géométrie hyperbolique.

Certaines extensions de la géométrie euclidienne, reposant sur des postulats qui semblent au départ arbitraires, trouvent des applications utiles en physique. Par exemple, Hermann Minkowski (qui fut le professeur d'Albert Einstein) ajoute une dimension supplémentaire aux espaces euclidiens, le temps. Ce nouveau concept, appelé espace-temps, est à la base de la relativité restreinte, dont les applications pratiques sont

nombreuses, notamment en électromagnétisme et en astronomie.

Pour élaborer la théorie de la relativité générale, Einstein utilise une autre de ces extensions : il considère la gravitation, la force universelle qui s'applique sur toute chose. Le seul moyen de s'y soustraire est d'y opposer une force d'inertie. Cette dernière se manifeste soit lors des accélérations et des décélérations sur les trajectoires rectilignes, mais qui apparaissent courbées dans l'espace-temps (sur une fusée qui décolle la force de gravitation est compensée par l'accélération), soit lors d'un mouvement sur une trajectoire courbe (la rotation des stations spatiales sur elles-mêmes crée une force d'inertie, qui joue le rôle de force de gravitation).

La courbure de l'espace-temps

Einstein a ainsi l'idée de modéliser la gravitation comme une courbure de l'espace. Grâce à cette théorie, il prévoit des effets qui n'existent pas dans les espaces euclidiens mais que l'on observe dans notre monde. Par exemple, un rayon lumineux passant à proximité du Soleil est dévié par la force de gravitation, tandis que, dans un espace euclidien, il se propage en ligne droite. La force de gravitation induit ainsi une courbure de l'espace (voir la figure 1). On le vérifie lors des éclipses de Soleil : les étoiles dont les coordonnées sont connues et dont les rayons passent près du Soleil

(celles qui sont à quelques minutes d'angle du Soleil) semblent légèrement décalées, par rapport à la position qu'elles devraient occuper. Cette observation montre que la lumière est effectivement déviée par le champ de gravitation du Soleil.

Puis les physiciens s'intéressent à la matière aux échelles microscopiques et élaborent la théorie de la mécanique quantique. Dans les espaces euclidiens, il y a une infinité de points sur la droite qui relie deux points donnés, si proches soient ils l'un de l'autre. Ce concept a-t-il un sens en physique ? Pour le savoir, les physiciens essaient de mesurer les positions des particules, indivisibles, qui sont l'équivalent des points dans les espaces euclidiens. Pour mesurer les dimensions d'une pièce, on la compare aux objets qu'elle contient, par exemple à un mètre. Plus la dimension que l'on veut mesurer est grande par rapport à l'objet qui sert de référence, plus la précision est grande (en admettant que l'objet de référence soit indivisible). Ainsi, pour mesurer la position d'une particule, on utilise l'objet de référence le plus petit possible, soit une autre particule. La précision est médiocre. En outre, la mesure, c'est-à-dire l'impact de la particule ou du faisceau de particules concé-

Des coordonnées non commutatives

La mécanique quantique, notamment la théorie quantique des champs, qui inclut les interactions quantiques et relativistes, ne tient pas compte de la gravitation. Aujourd'hui, on essaie d'élaborer des théories où la gravitation devrait obéir aux lois de la mécanique quantique. Dans ces théories, le principe de Heisenberg s'applique à la gravitation : la courbure de l'espace ne peut pas être complètement déterminée. Dès lors que l'on mesure l'une des coordonnées, on modifie cette courbure, donc les autres coordonnées. Ainsi, si l'on mesure les coordonnées d'un objet dans un certain ordre, par exemple x , y puis z , et qu'ensuite on les mesure dans un ordre différent, par exemple z , x , puis y , on n'obtient pas le même résultat : la mesure des coordonnées ne commute pas. Pour décrire cet état de fait, on utilise les géométries non commutatives.

Les espaces non commutatifs sont construits en découpant un espace de départ en régions, ou classes d'équiva-

lence (dont les points ont une propriété en commun) et en identifiant chacune d'elles à un point du nouvel espace (voir la figure 2). L'opération qui transforme une classe d'équivalence en un point du nouvel espace est dite opération de passage au quotient. Considérons notre monde comme le quotient d'un espace de départ (que l'on peut choisir parmi une multitude d'espaces distincts) du flou existant sur les coordonnées en physique quantique (correspondant aux différents états possibles d'un système à un instant donné) laisse supposer que l'espace de départ a plus de dimensions que celles que nous connaissons. Par exemple, il pourrait avoir sept dimensions spatiales et trois dimensions temporelles.

En mathématique, on utilise couramment des fonctions, qui assignent à chaque point un nombre, par exemple l'une de ses coordonnées. De même, en physique, la fonction température associe à chaque point de l'atmosphère un nombre qui exprime, en degrés, la température en ce point. La fonction que l'on peut nommer « fonction inter-

action » associe à chaque point un vecteur force à trois composantes, auquel serait soumise une particule supposée ponctuellement placée en ce point. Par définition, ces fonctions à valeurs numériques sont commutatives : l'attribution d'un nombre ou d'un vecteur à chaque point est indépendante des autres attributions effectuées. On ne peut donc pas décrire notre monde (en englobant la mécanique quantique et la gravitation, c'est-à-dire en considérant qu'il est non commutatif, comme nous l'avons vu précédemment sur les coordonnées d'une particule) avec des fonctions commutatives.

Puisque l'on ne peut pas utiliser de fonctions, on se sert d'opérateurs, bien adaptés, car ils ne commutent pas. Un opérateur est une procédure qui associe à chaque point ou entité de l'espace de départ une autre entité appartenant à l'espace d'arrivée (éventuellement, le même que l'espace de départ). Une rotation de 30 degrés du plan cartésien est un exemple d'opérateur d'un espace dans lui-même. Si un espace est invariant par rotation autour d'un point, on



1. CINQ IMAGES d'une même galaxie (indiquées par les flèches rouges) sont visibles sur cette photographie prise par le Télescope spatial Hubble. Cette observation illustre la courbure de l'espace-temps prévue par Einstein dans sa théorie de la relativité générale : les rayons lumineux sont déviés par les amas de galaxies (tel celui qui apparaît en jaune, au centre). Si l'on intègre cette propriété à la mécanique quantique, alors l'ordre dans lequel on mesure les coordonnées d'une particule n'est plus indifférent : la

mesure de chaque coordonnée modifie la courbure de l'espace-temps, donc la valeur des autres coordonnées. Pour modéliser ce phénomène, on a élaboré des géométries non commutatives. Pour choisir celle qui correspond le mieux à notre monde, on confronte les prévisions de chacune d'elles à nos observations. De même qu'Einstein a prévu la courbure de l'espace-temps avant de l'avoir observée, les géométries non commutatives révéleront peut-être des phénomènes insoupçonnés.

exprime ses propriétés de symétrie en affirmant simplement que l'opérateur de rotation transforme l'espace en lui-même. L'intérêt de l'approche opérationnelle réside dans son caractère synthétique.

Même dans un espace euclidien, une telle approche se justifie, car elle évite de considérer les points un à un et permet pourtant de retrouver les mêmes informations que par une approche classique. Prenons l'exemple d'une biopsie : elle consiste à prélever un échantillon en chaque point ou chaque endroit sur lequel on veut obtenir une information. Avec une radiographie, au contraire, on obtient une « synthèse » des données que l'on aurait récoltées en accumulant les biopsies, mais certaines informations sont perdues (la radiographie donne une information ou un niveau de gris pour un segment entier de points alignés dans la direction des rayons X, tandis que les biopsies donnent une information en chaque point). En mathématique, une radiographie peut être considérée comme un opérateur, défini sur le volume étudié (la par-

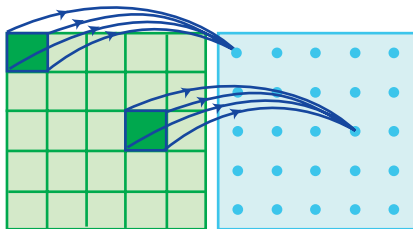
tie du corps radiographié) qui constitue l'espace de départ. L'espace d'arrivée est la plaque photographique. L'opérateur radiographie associe à chaque ensemble de points alignés dans la direction des rayons X, un point de la plaque photographique dont la nuance de gris est définie comme l'intégrale de l'intensité du rayon X qui a traversé le volume étudié (voir la figure 3). Dans ce cas, tous les points alignés dans la direction des rayons X sont « écrasés » en un seul point de la plaque photographique.

Des espaces d'opérateurs de dimension infinie

La mécanique quantique interdit toute description du monde en termes de fonctions d'onde. On préfère passer à des opérateurs. Aujourd'hui, pour élaborer de nouvelles théories, les physiciens envisagent notre monde comme le quotient d'un espace de départ, à imaginer. Ils construisent des géométries non commutatives pour étudier des espaces d'opérateurs qui feraient

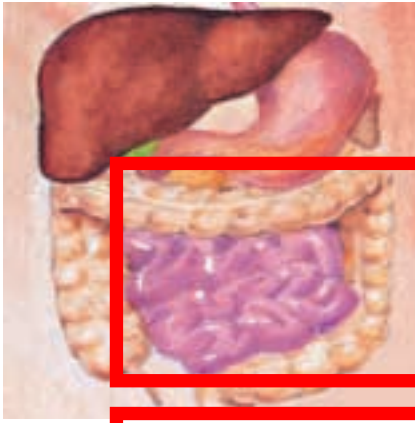
passer les espaces de départ possibles à notre monde. Ils ont déjà démontré que la partie non commutative de cet espace de départ (qui coexiste avec une partie commutative correspondant aux phénomènes ordinaires que nous connaissons) ne peut pas être décrite avec un espace d'opérateur de dimension infinie. Cette restriction est imposée par la théorie de la relativité restreinte. Cependant, ils n'ont que quelques indices sur l'espace de départ et procèdent un peu au hasard pour en choisir un parmi tous les espaces possibles, c'est-à-dire pour choisir la « jauge » : si l'on considère l'opération de passage au quotient comme une projection, pendant laquelle certaines informations présentes dans l'espace de départ sont perdues, la jauge serait la projection qui nous ramène à notre monde à l'espace de départ.

Comment accéder aux informations perdues, alors que l'on ne connaît même pas leur nature ? Les physiciens sont un peu dans la même situation que quelqu'un qui regarde les photographies d'une maison (une représentation à deux



2. DANS LES NOUVELLES THÉORIES DE PHYSIQUE, notre monde est considéré comme une projection d'un espace de départ sur un espace d'arrivée. De même, une pipe (à trois dimensions) est représentée par une projection sur un plan (la représentation est à deux dimensions ; ci-dessous, *La trahison des images*, de Magritte, 1929). Les géométries non commutatives étudient l'opération dite de passage au quotient, qui transformeraient un espace de départ (*ci-contre en vert*) en notre monde physique (*en bleu*). Chaque classe d'équivalence (où les points ont une propriété commune et schématisée ici par un carré vert) est identifiée à un point de notre monde, où les propriétés de l'espace de départ sont « agglutinées » en un point.





3. L'«OPÉRATEUR RADIOGRAPHIE» associe à chaque ensemble de points alignés dans la direction des rayons X qui traversent l'espace de départ (le volume du corps, à gauche), un point de la plaque photographique (à droite) ; la nuance de gris est définie comme l'intégrale de l'intensité du rayon X qui a traversé l'espace de départ. Les géométries non commutatives étudient les opérateurs qui font passer d'un espace de départ à notre monde. Lors de la transformation d'un espace de départ vers un espace d'arrivée, des informations sont perdues, de même qu'une partie de l'information sur les organes est perdue lorsque l'on effectue une radiographie : non seulement on ne voit pas que les os, mais on passe de structures tridimensionnelles à une photographie bidimensionnelle.

dimension) sans savoir à quoi ressemble effectivement une maison (c'est-à-dire sans savoir qu'elle existe dans un espace à trois dimensions). Pour décrire la maison, on pourrait croire qu'il faut accumuler le plus de photographies possible, car elles sont toutes différentes. En revanche, si l'on sait que la maison est un objet à trois dimensions, on se rend compte que trois photographies, prises dans les trois directions de l'espace, suffisent pour la décrire.

Les physiciens considèrent que la façon dont ils décrivent le monde avec la théorie quantique des champs est redondante, de même que la description de la maison avec un grand nombre de photographies. En quoi est-elle redondante ? Ils ne le sauront que lorsqu'ils auront fait un choix de jauge qui donnera un espace quotient dont les propriétés prédiront des phénomènes existant dans notre Univers, de même que la théorie d'Einstein prédisait la déviation des rayons lumineux par la force de gravitation, hypothèse confirmée ultérieurement.

Les choix de jauges

Pour l'instant, cette approche est une impasse. Les choix de jauge ont abouti à deux types de situations. Dans le premier, les descriptions des espaces de départ envisagés reproduisent les descriptions classiques de notre monde (notamment des interactions électromagnétiques), à quelques variantes près, mais ne prédisent que des phénomènes déjà connus. Ce résultat est aussi peu

surprenant que l'habileté d'un joueur de fléchettes qui vise toujours dans le mille, si ce joueur dessine la cible autour de la fléchette après l'avoir lancée. Dans le deuxième type de situation, les physiciens ont décrit l'espace de départ, dont notre monde serait le quotient, en s'affranchissant de certaines contraintes qui y existent. Leurs théories prédisent alors une foule de phénomènes qui n'existent pas dans notre monde.

Ils ne vont affiner le choix des contraintes à abandonner, mais ne pourront pas le faire à l'aide des seules géométries non commutatives. Les accélérateurs leur donneront peut-être des indices à mesure que les énergies employées augmentent, elles révéleront probablement des phénomènes insoupçonnés. Ces phénomènes désigneront alors la géométrie non commutative qui les prévoit comme celle qui fera progresser la physique.

Daniela BIGATTI est mathématicienne à l'Institut Weizmann, en Israël. Elle a récemment travaillé à l'Université catholique de Louvain sur les géométries non commutatives et leurs applications à la théorie des cordes.

Daniela BIGATTI, *Noncommutative spaces in physics and mathematics*, in *Class. Quant. Grav.*, n° 17, pp. 3403-3428, 2000.

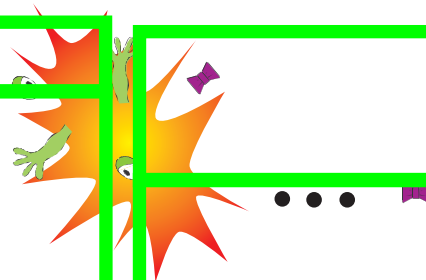
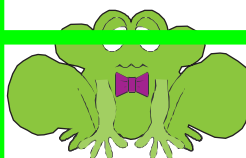
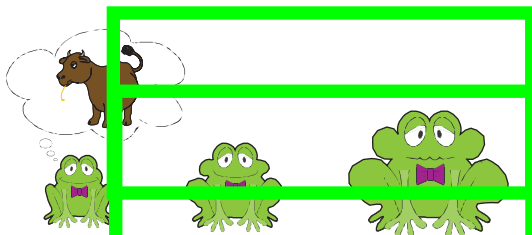
A. CONNES, *Noncommutative geometry*, Academic Press, 1995.

A. CONNES, *Geometrie non commutative*, InterÉditions, Paris, 1990.

L'infini est-il nécessaire?

PATRICK DEHORNOY

Oui. Les suites de Goodstein gonflent et grossissent jusqu'à des tailles gigantesques... Et diminuent finalement pour atteindre zéro. Pour démontrer cette propriété paradoxale, il est inévitable de faire appel à l'infini.



Les nombres entiers sont des objets finis : on pourrait donc espérer que toutes leurs propriétés se démontreraient sans recourir à l'infini. Erreur : seul l'infini permet de démontrer le théorème de Goodstein, qui s'applique aux suites inventées par le logicien anglais Reuben Louis Goodstein et dont les éléments sont pourtant des nombres entiers.

Pour comprendre ce paradoxe, voyons comment construire les suites de Goodstein et définissons pour cela le développement d'un entier en base p , puis en base p itérée.

On écrit habituellement les nombres entiers dans le système décimal, où l'on décompose un entier en une somme de puissances de 10, multipliées par des entiers compris entre 0 et 9. Ainsi l'entier noté 266 en base 10 correspond à la décomposition : $266 = 10^2 \times 2 + 10^1 \times 6 + 10^0 \times 6$, (par définition, a^1 est a et a^0 est 1). D'autres bases sont aussi employées. Les Babyloniens utilisaient la base 60, pratique grâce à ses nombreux diviseurs. Les ordinateurs effectuent les calculs en base 2. L'entier qui s'écrit 266 en base 10, se décompose en puissances de 2 : $266 = 256 + 8 + 2$, soit $2^8 + 2^3 + 2^1$.

En faisant figurer toutes les puissances de 2 : $2^8 \times 1 + 2^7 \times 0 + 2^6 \times 0 + 2^5 \times 0 + 2^4 \times 0 + 2^3 \times 1 + 2^2 \times 0 + 2^1 \times 1 + 2^0 \times 0$. D'où l'écriture de 266 en base 2 : 100001010.

Le développement en base p itérée est de même nature. Dans ce cas, on écrit aussi les exposants en base p , puis les exposants d'exposants, etc. Par exemple, le développement de 266 en base 2 est $2^8 + 2^3 + 2^1$; pour obtenir son développement en base 2 itérée, commençons par écrire les exposants en base 2 : $2^{2^3} + 2^{2^{1+1}} + 2^1$, puis l'entier 3, exposant d'exposant, est à son tour exprimé en base 2 : $266 = 2^{2^{2^1}+1} + 2^{2^{2^1+1}} + 2^1$. Cette expression est le développement en base 2 itérée de 266. Pour tout entier p plus grand que 2, le développement en base p itérée est construit de même : c'est une expression où apparaissent l'addition, la multiplication par des entiers plus petits que p et l'opération « p puissance».

La dilatation des entiers

Pour construire maintenant les suites de Goodstein, nous définissons une opération appelée dilatation, et désignée par d_p (p étant un entier au moins égal à 2), qui s'effectue, à partir d'un entier n quelconque, en deux étapes : on écrit n en base p itérée, puis on remplace partout p par $p+1$, (par exemple, 2^2 est remplacé par 3^3). Dilatons l'entier 266. Il s'écrit en base 2 : $2^{2^{2^1}+1} + 2^{2^{2^1+1}} + 2$, (nous avons omis les exposants 1), le dilaté $d_2(266)$ s'obtient en remplaçant les 2 par des 3, soit : $d_2(266) = 3^{3^{3^1+1}} + 3^{3^{3^1+1}} + 3 = 3^{81} + 84$,

dont l'écriture en base 10 a 38 chiffres. Remplacer « p puissance» par « $p+1$ puissance» engendre un nombre plus grand, voire beaucoup plus grand, que l'entier initial.

Forts de ces outils, nous définissons maintenant les suites de Goodstein $g_p(n)$. Une telle suite $g_1(n)$, $g_2(n)$, etc., se construit à partir d'un entier quelconque n que nous appellerons la «graine». La règle est la suivante : nous posons $g_1(n) = n$, puis nous calculons $g_2(n)$ en dilatant $g_1(n)$ à l'aide de l'opération d_2 et en retranchant 1 du résultat, et ainsi de suite en utilisant la relation de récurrence $g_p(n) = d_p(g_{p-1}(n)) - 1$.

Par exemple, pour la graine $n = 266$, nous trouvons $g_1(266) = 266$, puis $g_2(266) = d_2(266) - 1 = 3^{81} + 84 - 1 = 3^{81} + 83$. Ce nombre a , nous l'avons vu, 38 chiffres en base 10.

Le terme suivant $g_3(266)$ est égal à $d_3(g_2(266)) - 1$, soit $d_3(3^{3^{3^1+1}} + 3^{3^{3^1+1}} + 3) - 1 = 4^{4^{4^1+1}} + 4^{4^1+1} + 1$, un nombre de 616 chiffres en base 10.

Pourquoi retrancher 1 à chaque fois ? Les dilatations fabriquent des nombres tellement grands que le facteur — qui ne semble pas *a priori* influencer beaucoup sur le résultat ! — lui qui joue un rôle majeur dans les suites de Goodstein et engendre des phénomènes étranges.

Reprenons la suite de Goodstein de graine 266. Nous avons calculé $g_3(266) = 4^{4^{4^1+1}} + 4^{4^1+1} + 1$. Poursuivons : $g_4(266) = d_4(g_3(266)) - 1 = 5^{5^{5^{4^1+1}+1}} + 5^{5^{4^1+1}+1} + 1 - 1 = 5^{5^{5^5+1}} + 5^{5^5+1}$,



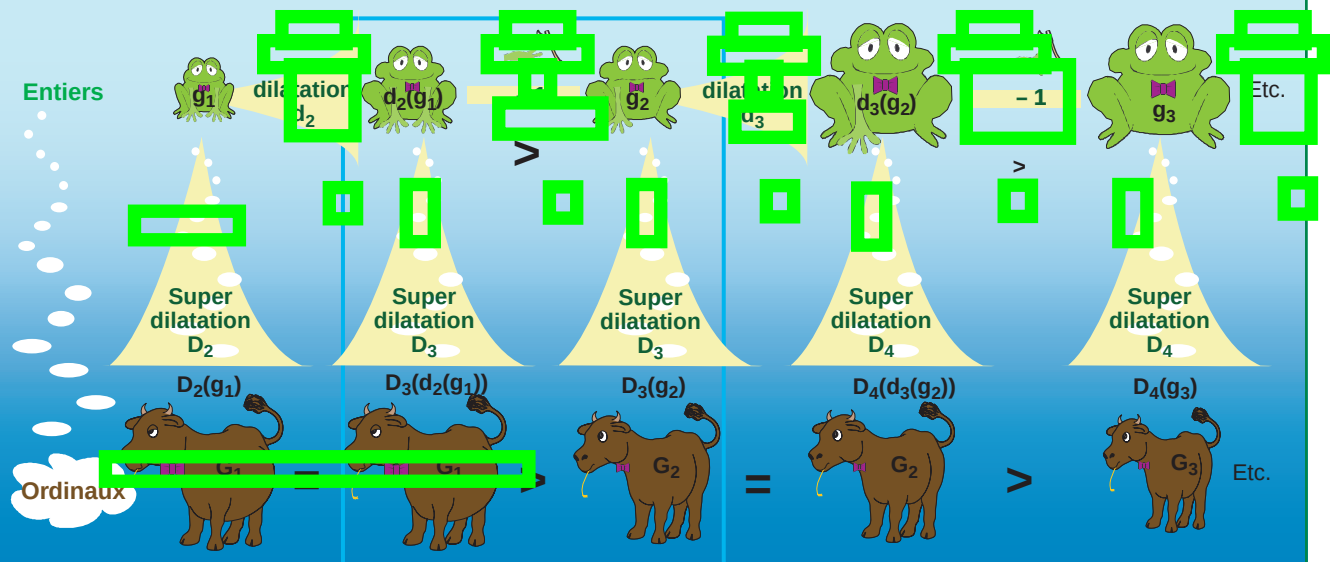
Reuben Goodstein

Difficile à croire? Nous allons pourtant démontrer ce résultat. Calculons g_p pour les premières graines. Dans le cas de la graine 2, nous trouvons $g_1(2) = 2$, puis $g_2(2) = d_2(2) - 1 = 3 - 1 = 2$, puis $g_3(2) = \frac{1+3+2}{2} - 1 = 1$, et $g_4(2) = 1 - 1 = 0$. Le résultat est là, mais n'est guère convaincant, car la graine est si petite qu'aucune opération «puissance» ne figure dans le développement. Le cas de 3 n'est pas plus palpitant : pour la même raison, les dilatations n'entrent

Ensuite viennent $\omega \times 2 + 1$, $\omega \times 2 + 2$, puis $\omega \times 3$, ... et $\omega \times n$; puis, après ce $\omega \times n$, qu'on note aussi $\omega \cdot n$, $\omega \cdot \omega$, et au delà, ω^ω , et beaucoup plus loin, ω^{ω^ω} , etc., la construction continue sans fin. Pour poursuivre, il faudrait introduire de nouvelles opérations, l'exponentiation:

[illegible]

dilatée de (g_1) : la grenouille g_2 est donc plus petite que la grenouille indienne $(g_1(g_1))$. Comme les suppositions concernent l'ordre, le bœuf G_2 est donc plus petit que le bœuf G_1 . Comme dans la fable, la grenouille a beau grossir, il n'en reste plus rien la fin...



pas vraiment en action, et nous aboutissons à $g_1(3) = 0$.

Le cas de la graine 4 gagne en revanche un intérêt... et devient plus compliqué. La suite atteint en effet la valeur 0 au bout d'un nombre d'étapes gigantesque (voir la figure 3). Dans le cas de la graine 5, et, *a fortiori*, des suivantes, les démonstrations paraissent hors de portée, car le nombre d'étapes augmente considérablement avec la valeur de la graine. Le théorème de Goodstein serait-il donc impossible à démontrer ?

Non. Il existe bien une démonstration, valable pour tout n , mais elle nécessite de sortir du cadre de l'arithmétique et de recourir à une autre méthode : les ordinaux. Les ordinaux

sont une prolongation de la suite des entiers, in-

roduite à la fin du siècle dernier par le mathématicien allemand Georg Cantor (voir la figure 1). Ils comprennent donc les entiers et, plus grands que tous les entiers, les ordinaux infinis, aussi appelés «transfins», c'est-à-dire «au-delà du fini». Deux points importants pour la démonstration du théorème de Goodstein : a) Les ordinaux sont munis d'une arithmétique similaire à celle des entiers (les quatre opérations de base) ; b) Toute suite décroissante d'ordinaux parvient à la valeur 0 en un nombre fini d'étapes (voir la figure 4).

Le théorème de Goodstein se démontre comme suit. Nous introduisons, pour chaque entier p , une «superdilatation» D_p définie comme d_p , à cela près qu'au lieu de remplacer p par $p+1$ dans le développement en base p itérée, p est remplacé par l'or-

dinal fini ω . Ainsi, la relation $266 = 2^{2^2} + 2^{2^1} + 2$ entraîne $D_2(266) = \omega^{\omega^{2^2}} + \omega^{\omega^{2^1}} + \omega$, à comparer avec $d_3(266) = 3^{3^{3^1+1}} + 3^{3^1+1} + 3$.

Examinons une propriété de la superdilatation : appliquer D_p revient à appliquer d_p puis D_{p+1} quel que soit le nombre de départ. En effet, dans le cas de la superdilatation directe D_p , on développe le nombre initial en base p itérée, puis on remplace les p par des ω . Dans le cas de la superdilatation indirecte (d_p suivie de D_{p+1}), on dilate d'abord le nombre, c'est-à-dire qu'on le développe en base p itérée et l'on remplace les p par des $p+1$, puis seulement intervient la superdilatation, où l'on remplace les $p+1$ par des ω , et le résultat final est le même. Vérifions avec 4 comme nombre de départ et avec $p = 2$. La superdilatation calculée directement est $D_2(4) = \omega^0$, car on remplace 2^2 par ω^0 .

La superdilatation indirecte mène à $d_2(4) = 3^3$, puis $D_3(d_2(4)) = D_3(3^3) = \omega^0$ aussi.

Une autre propriété de la superdilatation, déterminante pour démontrer le théorème de Goodstein, est que D_p est une fonction croissante pour chaque p . Par exemple, $D_2(4)$ est supérieure à $D_2(2)$.

Les supersuites de Goodstein

Après les superdilatations, définissons maintenant des suites d'ordinaux G , sorte de super-suites de Goodstein : l'élément $G_p(n)$ est l'ordinal $D_{p+1}(g_p(n))$. Voyons des exemples, et d'abord pour la graine 2 :

$G_1(2) = D_2(g_1(2)) = D_2(2) = \omega$,
 $G_2(2) = D_3(g_2(2)) = D_3(d_2(g_1(2)) - 1) = D_3(d_2(2) - 1) = D_3(3 - 1) = 2$.

Pour la graine 4, on trouve :

$G_1(4) = D_2(g_1(4)) = D_2(4) = \omega^0$,
 $G_2(4) = D_3(g_2(4)) = D_3(d_2(g_1(4)) - 1) = D_3(d_2(2^2) - 1) = D_3(3^3 - 1) = D_3(26) = D_3(3^2 \times 2 + 3 \times 2 + 2) = \omega^2 \times 2 + \omega \times 2 + 2$.

Dans ces exemples, on a $G_1(2) > G_2(2)$ et $G_1(4) > G_2(4)$. Le schéma de la

figure 2 généralise ce résultat et montre que les ordinaux $G_p(n)$ forment une suite strictement décroissante avec p pour chaque graine n .

Alors, une telle suite atteint 0 en un nombre fini d'étapes. Comme $G_p(n)$ ne peut valoir 0 que si $g_p(n)$ lui-même vaut 0, $g_p(n)$ atteint 0 en un nombre fini d'étapes, ce qui termine la démonstration du théorème de Goodstein.

3. SUITE DE GOODSTEIN DE GRAINE 4

Calculons la suite de Goodstein de graine 4. On trouve $g_1(4) = 4$, $g_2(4) = d_2(4) - 1 = d_2(2^2) - 1 = 3^3 - 1 = 26$.

$g_3(4)$ est égal à $d_3(26) - 1$, soit $d_3(3^2 \times 2 + 3 \times 2 + 2) - 1 = (4^2 \times 2 + 4 \times 2 + 2) - 1 = 41$.

Continuons : $g_4(4) = d_4(41) - 1 = d_4(4^2 \times 2 + 4 \times 2 + 1) - 1 = (5^2 \times 2 + 5 \times 2 + 1) - 1 = 60$.

$g_5(4) = d_5(60) - 1 = d_5(5^2 \times 2 + 5 \times 2) - 1 = (6^2 \times 2 + 6 \times 2) - 1 = 82$.

Quel que soit n , $g_n(4)$ a toujours un développement en base $n+1$ de la forme $(p+1)^2 \times a + (p+1) \times b_p + c_p$, avec $a_p \leq 2$, $b_p \leq p$ et $c_p \leq p$. Démontrons cette formule très utile pour prouver que la suite de graine 4 atteint 0. La formule est vraie pour $g_2(4) = 26 = 3^2 \times 2 + 3 \times 2 + 2$. Ici, on a $a_2 = 2$, $b_2 = 2$ et $c_2 = 2$, que nous notons $(2, 2, 2)$. La formule est conservée quand nous passons de $g_p(4)$ à $g_{p+1}(4)$, comme le montre l'étude des différents cas. L'évolution de a , b et c est similaire au compte rebours d'une horloge où a représenterait les heures, b les minutes et c les secondes ; cependant, l'horloge, détraquée, ne fonctionne plus en base 12, mais dans une base p qui augmente !

Premier cas : $c_p \neq 0$. La présence du facteur -1 diminue c et laisse a et b inchangés. Par exemple, on trouve $(a_3, b_3, c_3) = (2, 2, 2)$, $(a_4, b_4, c_4) = (2, 2, 2)$.

On a alors $a_{p+1} = a_p$, $b_{p+1} = b_p$, $c_{p+1} = c_p - 1$.

Deuxième cas : $c_p = 0$ et $b_p \neq 0$, comme dans $(a_4, b_4, c_4) = (2, 2, 0)$. Alors, à l'itération suivante, b diminue d'une unité et c réapparaît avec une valeur égale à la base diminuée de 1, c'est-à-dire p . On obtient $(a_5, b_5, c_5) = (2, 1, 5)$. Dans ce cas, $a_{p+1} = a_p$, $b_{p+1} = b_p - 1$ et $c_{p+1} = p$. Le coefficient b diminue seulement quand c égale 0, ce qui prend de plus en plus de temps car c réapparaît avec la valeur p , de plus en plus grande.

Troisième cas : $b_p = 0$ et $c_p \neq 0$, comme dans $(a_5, b_5, c_5) = (2, 0, 5)$. A l'itération suivante, on a $a_{p+1} = a_p - 1$, $b_{p+1} = p$, $c_{p+1} = p$, comme dans $(a_{23}, b_{23}, c_{23}) = (1, 23, 23)$. Ensuite, on obtient $(1, 0, 2)$ pour $p = 3 \times 2^{27} - 2$, à la suite de quoi a devient nul et les expressions prennent la forme $(p+1) \times b_p + c_p$. Ensuite, on obtient $(0, 1, 0)$, soit $g_p(4) = p+1$, pour $p = 3 \times 2^{27} \times 2^{3 \times 2^{27} - 2} - 1$; finalement, on atteint $(0, 0, 0)$, soit $g_p(4) = 0$ pour $p = 3 \times 2^{27+3 \times 2^{27} - 1} - 2$, soit $p = 3 \times 2^{402\,653\,211} - 2$, un entier dont l'écriture en base 10 a environ 130 millions de chiffres.

p	a, b, c
2	(2, 2, 2)
3	(2, 2, 1)
4	(2, 2, 0)
5	(2, 1, 5)
6	(2, 1, 4)
7	(2, 1, 3)
...	...
10	(2, 1, 0)
11	(2, 0, 11)
12	(2, 0, 10)
...	...
22	(2, 0, 0)
23	(1, 23, 23)
...	...
46	(1, 23, 0)
47	(1, 22, 47)
48	(1, 22, 46)
...	...

L'infini indispensable ?

Le théorème de Goodstein est une propriété d'arithmétique, au sens où il ne met en jeu que les nombres entiers et leurs quatre opérations élémentaires (addition, soustraction, multiplication, division), mais la démonstration que nous venons de donner repose sur l'utilisation d'objets infinis. En existe-t-il une autre qui ne passerait pas par l'infini ?

La démonstration d'un théorème n'utilise pas toujours que les outils avec lesquels il est formulé. De nombreux résultats d'arithmétique sont démontrés grâce à l'analyse, c'est-à-dire aux nombres réels et aux fonctions. Ceci a donné naissance à la théorie analytique des nombres (par opposition à la théorie algébrique des nombres, qui n'utilise que les entiers). Or, un nombre réel vraiment quelconque n'est exprimé ni par des fractions ni par des racines n -ièmes : c'est un objet dont le développement possède une suite infinie de chiffres après la virgule et, plus généralement, dont toutes les écritures sont infinies. Dès qu'une démonstration d'arithmétique utilise des nombres réels, une certaine forme d'infini y est donc présente. Notons cependant qu'il peut exister une autre démonstration alternative à l'utilisation de l'analyse et fondée seulement sur les entiers et leurs quatre opérations, mais elle est souvent très difficile à trouver. Ainsi, le célèbre théorème des nombres premiers, démontré à l'aide des nombres complexes par J. Hadamard et C. de La Vallée-Poussin en 1896, n'a été démontré par l'arithmétique que plus de cinquante ans plus tard, grâce à Selberg et Erdős.

Leur démonstration est par ailleurs beaucoup plus compliquée que l'originale. Précisons ce qu'on appelle une démonstration d'arithmétique. Les propriétés de base de l'arithmétique sont décrites par le système d'axiomes proposé à la fin du XIX^e siècle par Giuseppe Peano, mathématicien italien mort en 1932. Toutes les propriétés de base de l'arithmétique peuvent être démontrées dans le système de Peano, donc par récurrence : bien sûr, toutes les démonstrations d'un livre d'arithmétique ne sont pas des récurrences, mais on pourrait toujours s'y ramener, au moins de façon théorique.



Giuseppe Peano

Leur démonstration est par ailleurs beaucoup plus compliquée que l'originale. Précisons ce qu'on appelle une démonstration d'arithmétique. Les propriétés de base de l'arithmétique sont décrites par le système d'axiomes proposé à la fin du XIX^e siècle par Giuseppe Peano, mathématicien italien mort en 1932. Toutes les propriétés de base de l'arithmétique peuvent être démontrées dans le système de Peano, donc par récurrence : bien sûr, toutes les démonstrations d'un livre d'arithmétique ne sont pas des récurrences, mais on pourrait toujours s'y ramener, au moins de façon théorique.

4. SUITES DÉCROISSANTES D'ORDINAUX

Construisons une suite strictement décroissante à partir d'un entier, même très grand, disons 10^9 . Nous arriverons alors nécessairement à 0 après au plus 10^9 étapes. En revanche, puisqu'il y a une infinité d'entiers sous l'ordinal ω , nous pourrions penser que, partant de ω , il est possible de descendre une infinité de fois. Il n'en est rien, car, si nous construisons une suite strictement décroissante partant de $\alpha_1 = \omega$, il faut choisir pour α_2 un ordinal strictement plus petit que ω , c'est-à-dire un entier : si nous choisissons $\alpha_2 = 12$, nous atteignons 0 au plus 12 étapes plus tard. Si nous choisissons $\alpha_2 = 10^9$, nous pourrions descendre 10^9 marches avant d'arriver à 0, ce qui est plus, mais est toujours un nombre fini. Ainsi, à la différence du cas où nous partons d'un entier, il est impossible de donner de borne sur la longueur d'une suite descendant de ω à 0, il existe des suites aussi longues qu'on veut, mais chacune d'entre elles est néanmoins finie.

L'argument est le même lorsqu'on part d'ordinaux plus grands. Ainsi, si nous partons de $\alpha_1 = \omega^2$, nous devons choisir α_2 strictement plus petit, ce qui nous oblige à un saut infini : un tel α_2 est de la forme $\omega \times p + q$, $q \in \mathbb{N}$ et p est un entier. De là, au plus q étapes de descentes mènent à $\omega \times p$ (nous envisageons le pire cas), et, à l'étape suivante, nous arrivons à un ordinal de la forme $\omega \times (p-1) + q$, et ainsi de suite. Après au plus p sauts du même type, on arrive aux entiers, et donc à 0. Toute suite décroissante d'ordinaux partant de la valeur zéro en un nombre fini d'étapes.



Théorème de Kirby et Paris

Alors ? Pourrions-nous démontrer aussi, avec un peu d'astuce, le théorème de Goodstein par récurrence sans recourir à l'infini ? Non : un théorème, démontré en 1981 par Laurence Kirby et Jeffrey Paris, et basé sur une méthode que ce dernier avait mise au point en 1978 avec Leo Harrington, affirme que, quelle que soit notre imagination, nous n'arriverons jamais à montrer le théorème de Goodstein par récurrence, c'est-à-dire en ne nous servant que des entiers et des quatre opérations élémentaires. Ce résultat est remarquable (et sa démonstration très difficile !) et fait la spécificité du théorème de Goodstein. De façon plus précise, il affirme que la fonction qui, à chaque graine n , associe l'entier p pour lequel $g_p(n)$ s'annule est une fonction qui prend des valeurs tellement grandes qu'elle finit par dépasser n'importe quelle fonction dont on peut montrer l'existence par récurrence.

Comme elle utilise l'ordinal ω , objet infini, la démonstration du théorème de Goodstein n'est pas une preuve par

récurrence. Notons qu'il suffit d'ajouter l'hypothèse qu'il existe un objet infini au système de Peano pour obtenir, sans avoir à ajouter de nouveau postulats, toute l'arithmétique des ordinaux (utilisée dans la preuve de Goodstein). Du point de vue des hypothèses logiques sous-jacentes, ceci place le théorème de Goodstein juste au-dessus de l'arithmétique de Peano. Cet ajout représente exactement l'écart entre l'arithmétique (présent dans l'arithmétique, où le principe de récurrence postule l'existence d'une suite d'entiers sans fin) et l'infini actuel (présent dans les ordinaux, où existe un objet infini).

... Et Gödel

Il existe un rapport étroit entre ce qui précède et les célèbres théorèmes d'incomplétude démontrés par Kurt Gödel au début des années 1930.

Le premier théorème d'incomplétude énonce qu'il existe une propriété d'arithmétique qui est vraie, mais qui ne peut pas être démontrée par



Kurt Gödel

5. HERCULE CONTRE L'HYDRE

Le combat d'Hercule contre l'Hydre représente, comme le théorème de Goodstein, un exemple de propriété qui met en jeu des entiers et ne peut se démontrer sans utiliser l'infini. Hercule doit vaincre l'Hydre, mais, chaque fois qu'il lui coupe une tête, il en repousse de nouvelles ! Les règles du combat sont les suivantes : Hercule coupe une seule tête à chaque coup. S'il coupe une tête directement reliée au corps de l'Hydre, la tête ne repousse pas ; en revanche, si la tête coupée n'est pas directement reliée au corps, alors l'ensemble des têtes situées à côté de la tête coupée, c'est-à-dire provenant de la bifurcation commune au-dessous, se reproduit en n nouveaux exemplaires, si Hercule porte son n -ième coup du combat. Le nombre de tête de l'Hydre semble tendre vers l'infini irrémédiablement. Pourtant, un théorème énonce que, quelle que soit la forme initiale de l'Hydre et quelle que soit la stratégie d'Hercule, il finira toujours par tuer l'Hydre en coupant toutes les têtes. Comme le théorème de Goodstein, le théorème de l'Hydre se démontre en utilisant des ordinaux et, à nouveau, Kirby et Paris ont prouvé qu'il ne peut pas être démontré dans le système de Peano, c'est-à-dire par récurrence.



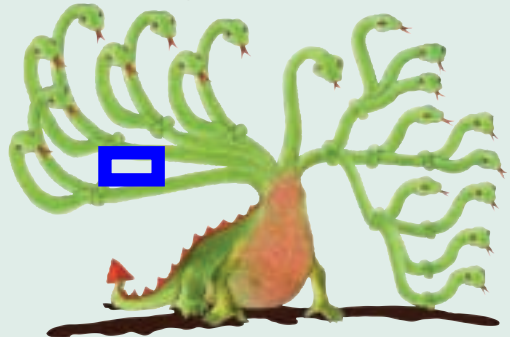
PREMIÈRE COUPURE
En brun, la partie excitée
par la coupe.
Cette partie est dupliquée.



DEUXIÈME COUPURE
En brun, la partie excitée
par la deuxième coupe.
Cette partie est ajoutée
deux fois.



TROISIÈME COUPURE
En brun, la partie excitée
par la troisième coupe.
La partie excitée est ajoutée
trois fois. Si l'on coupe la tête
bleue reliée directement
au corps de l'hydre,
elle ne repousse pas.



**Coupera-t-on
toutes les têtes de l'hydre ?**
Oui, et ce résultat
se démontre
avec les ordinaux.

récurrence. Il ne donne toutefois aucun exemple explicite d'une telle propriété.

Le second théorème d'incomplétude comble cette lacune, en livrant un exemple : c'est une propriété qui code, dans le langage de l'arithmétique, le fait que le principe de récurrence n'est pas contradictoire avec lui-même. Ce résultat, fort remarquable, continue de fasciner après bien des années. Pourtant, du point de vue de l'arithmétique, il nous laisse un peu sur notre faim, car la propriété qui code dans le langage de l'arithmétique le fait que le principe de récurrence n'est pas auto-contradictoire manque de clarté et ne correspond guère à ce qu'un arithméticien a l'habitude d'appeler une propriété d'arithmétique...

Il a fallu attendre cinquante ans avant de trouver des propriétés d'arithmétique plus communes qui ne soient pas démontrables par récurrence. Le premier exemple a été une propriété de type combinatoire isolée par Paris et Harrington en 1978, suivi du théorème de Goodstein, établi quelques années plus tard. C'est à ce jour l'exemple le plus simple d'une propriété d'arithmétique, qui se démontre à l'aide de l'infini, et dont il est prouvé qu'elle ne peut être démontrée par récurrence.

En un sens, un tel résultat constitue un argument puissant en faveur de l'utilisation de l'infini en mathématiques puisqu'il montre que certaines propriétés ne mettant en jeu que des objets finis ne peuvent être démontrées qu'à l'aide de l'infini : on aurait donc tort de se priver d'une telle possibilité. D'un autre côté, l'exemple de la graine 4 montre que le théorème de Goodstein met en jeu des entiers gigantesques, beaucoup plus grands par exemple que le nombre d'atomes dans l'univers, et nous pouvons aussi nous interroger sur l'usage de tels entiers et leurs rapports avec les entiers «de tous les jours»...

Patrick DEHORNOY est professeur de mathématiques à l'université de Caen.

G. GOODSTEIN, *On the Restricted Ordinal Theorem*, in *J. Symb. Logic*, n° 9, pp. 33-41, 1944.

L. KIRBY et J. PARIS, *Accessible Independence Results for Peano Arithmetic*, in *Bull. London Math. Soc.*, n° 14, pp. 285-293, 1982.

A. LEVY, *Basic Set Theory*, Springer, 1979.

L'infini et l'univers des algorithmes

GILLES DOWEK

On ne peut pas toujours prévoir le nombre d'opérations nécessaires à un calcul. Aussi, pour permettre l'expression de tous les algorithmes, il faut introduire la possibilité que les calculs durent infiniment longtemps.



Un algorithme décompose un calcul en des suites d'opérations élémentaires. Quand un algorithme prescrit un nombre infini d'opérations élémentaires, on n'obtient jamais le résultat. Aussi, faut-il tenter de maîtriser ou de chasser cet infini.

Pour multiplier 35 par 12, nous devons multiplier 5 par 2, poser le 0, retenir le 1, multiplier 3 par 2, ajouter la retenue... au bout de nos efforts, nous obtiendrons le résultat : 420. Cette recette qui décrit l'enchaînement des opérations élémentaires d'une multiplication est dénommée l'algorithme de la multiplication, ou plutôt un algorithme de la multiplication, car il existe d'autres méthodes de calcul possibles. Ainsi, $35 + 35 + 35 + 35 + 35 + 35 + 35 + 35 + 35 + 35 + 35 + 35$ est une méthode, un algorithme envisageable pour multiplier 35 par 12, à l'aide d'additions successives. Il en existe bien d'autres.

Ces algorithmes sont des méthodes de calcul qui donnent un résultat après un nombre fini d'opérations et dans un temps fini. L'infini ne semble ainsi jouer aucun rôle dans l'univers des algorithmes, et pourtant la notion d'algorithme ne

peut être définie sans utiliser l'infini. L'infini est indispensable, mais il faut le maîtriser.

La définition d'un algorithme

Si nous utilisons des algorithmes depuis l'Antiquité et que le mot algorithme vient du nom du mathématicien Al-Khwarizmi qui vécut à Bagdad au IX^e siècle, ce n'est qu'au XX^e siècle que cette notion a été définie avec précision.

Dans les années 1920, Thoralf Skolem (1887-1963) a proposé une première définition de la notion d'algorithme liée à la récurrence. Ce procédé consiste à définir une fonction en donnant sa valeur en 0 et en indiquant comment passer de sa valeur en n à sa valeur en $n + 1$. Par exemple, la fonction e , qui à chaque nombre entier n associe le nombre 2^n : $e(0) = 1$, $e(1) = 2$, $e(2) = 4$, $e(3) = 8$, $e(4) = 16$, ... est définie par récurrence, en indiquant que sa valeur en 0 est 1 et qu'on passe d'une valeur à la suivante en la doublant.

Les fonctions qui sont définies par leur valeur en 0, ou en une autre valeur

initiale et par une opération permettant de passer directement d'une valeur à la suivante sont appelées, selon un terme de J. Koszár Péter, récursives primitives. La fonction $e(n) = 2^n$ est récursive primitive. Elle est définie par $e(0) = 1$ et $e(n + 1) = 2e(n)$. Richard Dedekind, qui avait déjà étudié les définitions par récurrence en 1888, avait montré que l'addition, la multiplication, l'élevation à la puissance étaient des fonctions récursives primitives.

La définition d'une fonction par récurrence donne un algorithme pour le calcul des valeurs de cette fonction. Ainsi, pour calculer la valeur de la fonction e en 10, il suffit de calculer sa valeur en 0, puis en 1, en 2, ..., en 9 et enfin en 10. On obtient ainsi $e(0) = 1$, $e(1) = 2$, $e(2) = 4$, ..., $e(9) = 512$, et $e(10) = 1024$.

On programme aisément de telles fonctions récursives primitives en utilisant des boucles *For*. Chaque itération de la boucle exécute une opération élémentaire. On peut déterminer le nombre d'itérations nécessaires au calcul en fonction de la valeur que l'on souhaite calculer. Si l'on demande au programme de calculer la fonction e pour n égal à 10, après 10 itérations, 1024 s'affichera sur l'écran de l'ordinateur.

Avec cette notion de fonction récursive primitive, l'infini n'était pas convoqué, et les mathématiciens pensaient que ces fonctions suffisaient pour définir l'ensemble des algorithmes. Pourtant, en 1928, Wilhelm Ackermann a trouvé une fonction de deux variables m et n pour laquelle il y avait une recette de calcul, mais qui ne répondait pas au souhait de prévisibilité du nombre de pas de calcul. Cette fonction est définie par : $A(0, n) = n + 1$; $A(m + 1, 0) = A(m, 1)$ et $A(m + 1, n + 1) = A(m, A(m + 1, n))$.

Cette fonction n'est pas récursive primitive : le nombre de pas de calcul nécessaire pour passer d'une valeur de m ou de n à la suivante, $m + 1$ ou $n + 1$ n'est pas connu. Cette impossibilité n'est pas due à une mauvaise formulation de la fonction, mais à une impossibilité théorique démontrée par Ackermann : même si la fonction d'Ackermann peut toujours se calculer en un nombre fini d'étapes, le nombre d'itérations nécessaires au calcul est imprévisible. En informatique, le calcul d'une telle fonction ne peut être réalisé en utilisant des boucles *For*.

Les fonctions récursives primitives peuvent toujours se calculer par un algorithme ; en revanche, les fonctions calculables par un algorithme ne sont pas toujours récursives primitives, comme le prouve la fonction d'Ackermann. La définition de Skolem était donc trop restrictive, et le défi lancé aux mathématiciens, au début des années 1930, était de dépasser cette définition pour en proposer une qui englobe l'ensemble des algorithmes.

Le schéma *mu*

Plusieurs solutions ont été proposées par Jacques Herbrand, Kurt Gödel, Alonzo Church, Stephen Kleene, Alan Turing, Emil Post... Dans la définition la plus simple, avancée par Kleene en 1936, on ajoute aux définitions par récurrence un second procédé de calcul, portant le joli nom de schéma *mu*. Ce schéma nous oblige à continuer le calcul jusqu'à ce qu'une condition soit atteinte. Ainsi, à partir d'une fonction g , Kleene définit un nombre a qui est le plus petit nombre x pour lequel $g(x)$ est égal à 0 (condition recherchée ici). Par exemple, si g est la fonction qui à x associe le nombre $|x^2 - 9|$, 3 est le plus petit nombre x tel que $|x^2 - 9|$ soit égal à 0.

On sait calculer une fonction définie par un schéma *mu* au moyen d'un

La fonction d'Ackermann

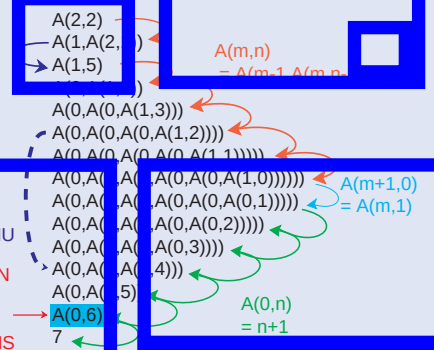
La fonction d'Ackermann est une fonction parfaitement définie et calculable, mais non récursive primitive. Prenons l'exemple du calcul de la fonction $A(2,2)$ en décomposant le calcul. Si cette fonction était récursive primitive, il nous suffirait de partir de la valeur de $A(0,0)$ puis de calculer $A(1,0)$, $A(2,0)$, $A(2,1)$ et enfin $A(2,2)$. Quatre étapes seraient suffisantes. Le problème principal de la fonction d'Ackermann est que les résultats obtenus pour des valeurs de m ou de n inférieures à celles que l'on souhaite calculer ne sont pas suffisants. La décomposition du calcul de la fonction $A(2,2)$, nécessite de calculer des fonctions où les valeurs de m et n sont supérieures à la valeur initiale. Seul un mécanisme de recherche par schéma *mu*, où la condition d'arrêt des itérations est l'obtention d'une seule valeur de m qui soit égale à 0, donne le résultat.

D'après la définition d'Ackermann, $A(2,2)$ se décompose en $A(1, A(2,1))$. D'après l'hypothèse, $A(2,1)$ est connu et égal à 5, nous obtenons alors la fonction $A(1,5)$, dont la valeur est inconnue. On décompose cette fonction jusqu'aux $A(0, n)$ dont les valeurs sont égales à $n + 1$. Ainsi $A(2,2)$ est égal à 7.

RÉSULTAT
CONNU
 $A(2,1) = 5$

SI ON
SAIT QUE
 $A(1,2) = 4$
EST CONNU

CONDITION
D'ARRÊT
DES
ITÉRATIONS



algorithme, c'est-à-dire que l'on peut calculer la valeur de a , bien que ce calcul soit moins simple que celui par récurrence : pour calculer a , nous devons rechercher le plus petit nombre x tel que $|x^2 - 9|$ soit égal à 0, c'est-à-dire calculer successivement $g(0)$, $g(1)$, $g(2)$, $g(3)$, ... jusqu'à une valeur d'annulation. Contrairement au calcul par récurrence classique des fonctions récursives primitives, le nombre d'itérations est imprévisible. Nous pouvons trouver la solution après quelques itérations, ou alors tester un grand nombre de valeurs avant de trouver celle qui convient.

Dans les langages de programmation, cette recherche de la solution s'effectue à l'aide des boucles *While*, qui de la même façon que les boucles *For* exécutent une opération plusieurs fois. Cependant, la boucle *For* n'est adaptée qu'aux fonctions récursives primitives, elle contient l'instruction d'arrêter le calcul lorsque le nombre d'itérations souhaité est atteint. La boucle *While* contient l'instruction d'arrêter le calcul lorsque la condition recherchée est atteinte, ici $g(x) = 0$, et ce quel que soit le nombre d'itérations. Le programme recherche la solution en itérant tant qu'elle (*While*) n'est pas trouvée.

Avec le schéma *mu*, on calcule toutes les fonctions où le nombre d'étapes n'est pas connu *a priori* et où l'on découvre au fur et à mesure la

structure du calcul. Le schéma *mu* permet d'élaborer progressivement la structure du calcul de la fonction d'Ackermann (voir l'encadré). Cette notion de découverte progressive était absente de la définition de Skolem.

Les fonctions définies avec ces deux procédés – définitions par récurrence et schéma *mu* – sont calculables par algorithme, et l'on a de bonnes raisons de penser que, réciproquement, toutes les fonctions qui peuvent se calculer par un algorithme peuvent se définir ainsi. Le défi lancé aux mathématiciens par l'introduction de la fonction d'Ackermann était ainsi relevé. Les fonctions récursives primitives et le schéma *mu* englobent l'ensemble des algorithmes.

L'arrivée de l'infini

Avec le schéma *mu*, contrairement aux définitions par récurrence, il n'est pas possible de prévoir le nombre d'itérations nécessaires pour trouver le résultat. Il se peut même que l'on ne trouve pas de solution, et, dans ce cas, le calcul se poursuivra éternellement.

Remplaçons dans l'exemple précédent le chiffre 9 par 10, et tentons de calculer le nombre a , qui soit le plus petit nombre x tel que $|x^2 - 10|$ soit égal à 0. Comme aucun nombre ne convient, on essaiera 0, 1, 2, 3, 4, 5... sans jamais trouver la valeur d'annulation. Techniquement, on dit que ce

nombre n'est pas défini. Concrètement, le calcul se poursuit à l'infini. C'est ce que Kleene appelle le mouvement perpétuel. Ce mécanisme de recherche à l'infini offert par le schéma *mu* se retrouve dans les boucles *While* des programmes informatiques ; le programme «boucle» tant que $|x^2 - 10|$ est différent de 0.

Traditionnellement, un algorithme était une recette de calcul qui fournissait toujours un résultat après un nombre fini d'opérations ; avec le schéma *mu*, l'univers des algorithmes s'élargit et, dans certains cas, le calcul se prolonge à l'infini et ne donne jamais de résultat. On dénomme parfois de tels algorithmes, des semi-algorithmes.

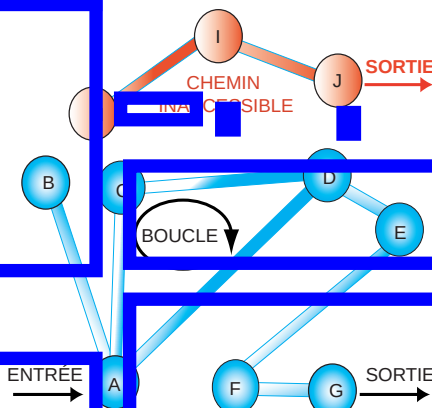
Maîtriser l'infini

Le schéma *mu* permet de définir des semi-algorithmes qui recherchent un certain objet dans un ensemble infini, sans même savoir si un tel objet existe. Un programme informatique qui recherche la sortie dans un labyrinthe est une illustration du schéma *mu*. Par exemple, on peut chercher un chemin qui va de A à G dans le labyrinthe de l'illustration ci-contre, et l'on trouvera peut-être le chemin A, D, E, F, G. Si l'on cherche un chemin allant de A à J, on n'en trouvera pas.

Comment procède-t-on pour trouver le chemin qui va de A à G ?

Le nombre de chemins dans un labyrinthe est infini : on peut très bien se perdre en parcourant une boucle, en passant de A en D, de D en C et de C en A, sans jamais prendre le couloir qui va de D vers E. Pour trouver la sortie d'un labyrinthe, il ne suffit pas de faire la liste des chemins en allant toujours droit devant soi ; il faut énumérer les chemins de manière systématique afin de n'en oublier aucun, par exemple en prenant tous les chemins de longueur 1 (qui sont en nombre fini), puis tous ceux de longueur 2 (qui sont également en nombre fini), puis ceux de longueur 3... L'énumération de tous les chemins de tailles finies du labyrinthe permet de trouver la sortie, si celle-ci existe et correspond à un chemin de taille finie, mais énumérera les chemins éternellement s'il n'en existe pas. Ce risque de recherche à l'infini, dans le cas où le problème n'a pas de solution, est inévitable dans un ensemble infini.

Dans le cas du labyrinthe, il est possible d'éviter cette recherche dans un



LA RECHERCHE DE LA SORTIE dans un labyrinthe peut être infinie. Il faut donc borner la recherche en excluant, par exemple, les boucles telles que ACDA.

ensemble infini. Pour traverser un labyrinthe, il n'est jamais nécessaire de passer deux fois par le même carrefour ; s'il y a un chemin qui comporte une boucle, on la supprime et l'on obtient un autre chemin qui mène directement au même endroit.

On se restreint ainsi à chercher une solution parmi les chemins sans boucle. Ces chemins sont alors en nombre fini, et s'il n'y a pas de solution (comme quand on essaie d'aller de A à J), après avoir énuméré tous les chemins et constaté qu'aucun d'eux ne convient, on saura que les deux points ne sont pas reliés. Qu'il y ait une solution ou non, la recherche arrivera à son terme.

Restreindre ainsi l'espace de recherche de manière à le rendre fini, sans pour autant perdre la solution, est le problème principal quand on cherche à concevoir certains algorithmes, par exemple des algorithmes de sortie de labyrinthe : cela s'appelle borner le schéma *mu*. Parfois – comme dans ce cas – cela est possible, mais dans d'autres cas, en particulier quand le problème qu'on cherche à résoudre est indécidable, cela ne l'est pas. Lorsque le problème est indécidable, on ne peut pas toujours, après avoir effectué un nombre fini d'étapes, prédire si la solution existe ou non. Le problème démarde une recherche dans un ensemble infini, et donc une recherche à l'infini quand il n'y a pas de solution.

Dans certaines situations, on ne peut éviter la recherche à l'infini, et le schéma *mu* est le bon outil. Mais dans d'autres, par exemple quand on cherche à définir la fonction d'Ackermann, dont les étapes sont certes imprévisibles, mais sont toujours finies, le schéma *mu* semble un outil mal

adapté, car trop puissant, puisqu'il convoque l'infini dans sa recherche. En introduisant le schéma *mu*, Kleene a fini par entrer tout dans la bergerie. L'infini dans la théorie des algorithmes qui était resté à l'extérieur du cadre du fini.

Cela est-il dû à une maladresse de Kleene ou cette intrusion de l'infini dans l'univers des algorithmes était-elle nécessaire ?

Chasse à l'infini ?

Church et Kleene ont montré qu'il ne s'agissait pas d'une maladresse et que cette intrusion était inévitable. En effet, tous les procédés de définition d'algorithmes qui se limitent à définir des fonctions qui donnent toujours un résultat après un nombre fini d'opérations (c'est-à-dire qui excluent le mécanisme de recherche à l'infini) sont incomplets : on pourra toujours construire une fonction, calculable par un algorithme en un nombre fini d'étapes, mais qui n'est pas définissable par ce procédé. Cette fonction se construit en utilisant la méthode de la diagonale introduite par Georg Cantor pour étudier les cardinaux infinis (voir l'article de Jean-Paul Delahaye, *L'infini est-il paradoxal en mathématiques ?*, page 30).

Il n'y a donc pas de définition qui englobe toutes les fonctions donnant un résultat après un nombre fini d'étapes et rien qu'elles. Certains procédés, comme le schéma *mu*, définissent toutes ces fonctions, mais aussi certains fonctions qui ne donnent pas de résultat après un nombre fini d'étapes. D'autres procédés comme ceux – plus complexes – proposés par Kurt Gödel en 1931 ou Jean-Yves Girard en 1971 ne permettent de définir que des fonctions qui donnent un résultat en un nombre fini d'étapes. Cependant, ils ne les définissent pas toutes.

Chasse à l'infini, il reviendra au galop.

Gilles DOWEK est chargé de recherches à l'Institut national de recherche en informatique et en automatique.

Jean-Luc CHABERT et al., *Histoire des algorithmes*, Boin, 1993.

Richard LASSAIGNE et Michel DE ROUGEMONT, *Logique et fondement de l'informatique*, Hermès, 1993.

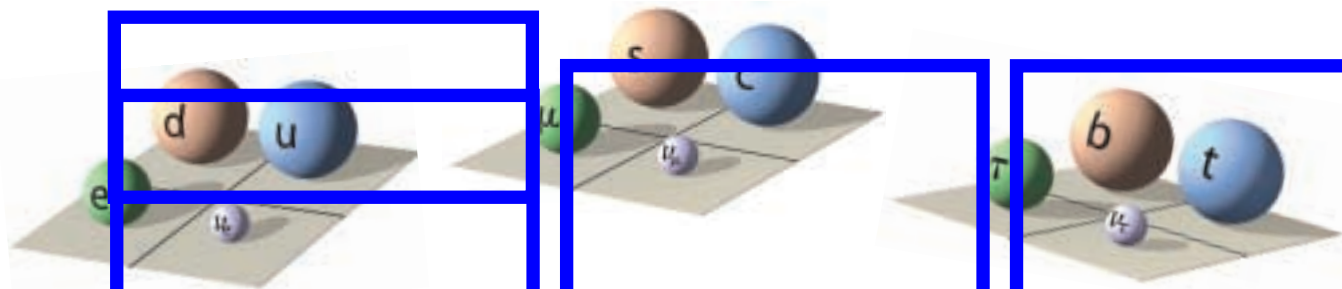
René CORI, *Logique mathématique*, Masson, 1993.

L'infiniment petit en physique

HARALD FRITZSCH

En descendant l'échelle des dimensions, les physiciens ont découvert des particules si petites qu'ils les ont supposées ponctuelles.

Le «modèle standard», la théorie construite sur cette hypothèse, accumule les succès depuis 30 ans, tout en étant pétrie d'infinis...



L'infiniment petit est omniprésent en physique des particules. En principe, aucune idée physique ne devrait faire appel à des notions infinies. Pourtant, les théories utilisées aujourd'hui par les physiciens qui étudient l'Univers à petite échelle supposent des constituants élémentaires ponctuels, donc infiniment petits. Les chercheurs n'ont pas adopté cette idée d'emblée. Partant de notre échelle macroscopique, régie par les lois de la physique classique, ils sont descendus d'étape en étape vers les petites dimensions, dans le domaine où règnent les lois quantiques. L'exploration des échelles atomique, puis nucléaire, puis subnucléaire les a confrontés à des comportements de plus en plus surprenants. Au-dessous du millionième de milliardième de mètre, tout se passe comme si les particules qui interagissent étaient infiniment petites. Les physiciens ont donc construit des théories fondées sur l'hypothèse que ces particules – les électrons et les quarks – sont ponctuelles, et qu'elles interagissent de façon ponctuelle. Extrême, cette thèse a d'abord

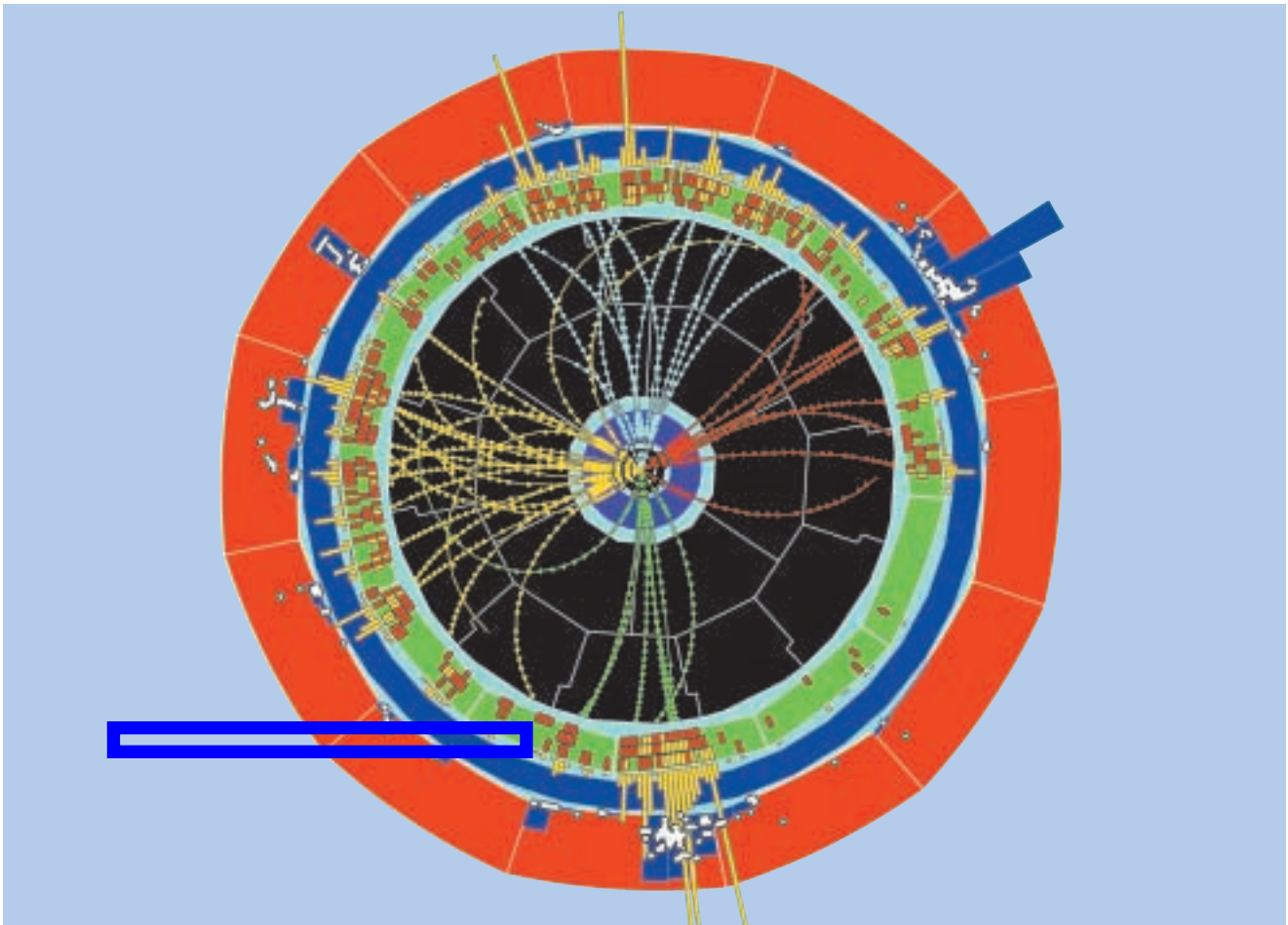
engendré d'immenses difficultés techniques. Toutefois, la première des théories fondée sur elle – l'électrodynamique quantique, qui décrit les interactions entre électrons et photons – devint rapidement la plus précise jamais élaborée. Encouragés par ce succès, les physiciens forgèrent des théories sur le même modèle pour décrire les autres interactions. Ils obtinrent ainsi le «modèle standard», qui rend compte avec succès des interactions de toutes les particules. Depuis 30 ans, cet énorme édifice théorique résiste à toutes les tentatives expérimentales pour le mettre en défaut. Les constituants élémentaires sont-ils vraiment ponctuels?

Les physiciens n'ont pas inventé l'infini

Les physiciens n'ont pas inventé l'infini : ils ont commencé à utiliser par nécessité ce concept de mathématicien (voir *Histoire d'infini*, page 64). En mathématique, la notion d'infini apparaît d'une façon «naturelle». On conçoit immédiatement que la suite des nombres naturels est sans fin, donc

qu'elle est infinie. Lorsque les mathématiciens disent qu'une quantité est infinie, ils veulent dire qu'elle (sa valeur absolue) est sans limite supérieure ou, ce qui revient au même, que supposer qu'une telle limite existe conduit à une contradiction logique : n ne peut être le plus grand entier, car $n + 1$ lui est supérieur. À l'inverse, quand ils parlent d'une quantité infiniment petite, cela signifie qu'elle (sa valeur absolue) n'a pas de limite inférieure, sinon zéro. La suite des inverses des nombres naturels, $1, 1/2, 1/3, 1/4, \dots$, ne contient

1. PLUS UN OBJET EST PETIT, plus les instruments nécessaires pour le mettre en évidence sont grands. En bas, le détecteur Aleph, à l'aide duquel les physiciens du CERN auraient récemment découvert une nouvelle particule : le boson de Higgs. Si elle se confirme, la découverte est importante, car le boson de Higgs est le dernier élément non découvert prévu par le modèle standard, la théorie la mieux acceptée par les physiciens des particules. Les expérimentateurs du CERN ont constaté que les trajectoires rassemblées dans les faisceaux vert et bleu proviennent de la désintégration d'une grosse particule. Sa masse semble correspondre à celle du boson de Higgs.



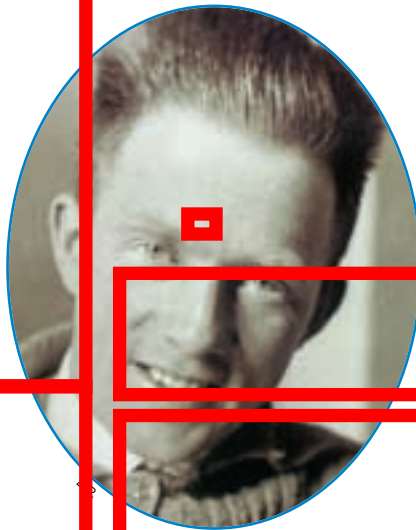
aucune limite inférieure. Pour les mathématiciens, les infinis, tant grands que petits, sont définis grâce à un processus de passage à la limite.

En physique, les situations qui imposent un tel passage ne manquent pas. Au début du XX^e siècle, les astronomes se représentaient l'Univers comme un espace plein d'étoiles, dont le Soleil. Puis ils ont découvert que les étoiles étaient les constituants élémentaires de galaxies, et que des milliards de galaxies existent. L'Univers s'était brusquement agrandi de plusieurs ordres de grandeur. Cette évolution invite à se demander quelles structures encore plus grandes pourraient exister. Et finalement, on s'interroge : le Cosmos est-il infini ? Ces questions restent ouvertes, même si plusieurs indices militent effectivement pour un Univers sans limites.

Il en va de même lorsqu'on essaie d'explorer l'infiniment petit. L'histoire des sciences révèle que l'homme a progressivement appris à diviser de plus en plus la matière. L'un des premiers à y penser fut Démocrite (460-370 avant notre ère). Le philosophe grec a supputé que la matière peut être divisée jusqu'à une échelle ultime, celle de petits constituants élémentaires invisibles, qu'il nomme *Atomos*. Pour Démocrite, comme pour les physiciens d'aujourd'hui, le processus de division de la matière est donc limité. Bien que très petit, un atome n'est pas infiniment petit. Son diamètre est d'environ un dixième de milliardième de mètre, taille qu'une expérience simple met en évidence : un film huileux répandu sur l'eau s'étend jusqu'à atteindre une hauteur qui correspond au diamètre d'une molécule, voire d'un atome. Or, grâce au microscope à effet tunnel, les physiciens sont capables d'observer des atomes. Que voient-ils ? De petites balles de tennis, dont il faudrait 100 millions accolées pour faire un centimètre.

Atomes signifie indivisible

Le mot *Atomos* signifie indivisible en grec ancien. Pourtant le caractère indivisible des atomes a été remis en cause à la fin du XIX^e siècle, quand les physiciens ont montré que les atomes peuvent être dissociés en constituants plus petits : le noyau et le nuage électronique. La taille de l'enveloppe impose celle des atomes, tandis que sa masse représente moins d'un millième de leur masse. Le plus simple des atomes est



2. WERNER HEISENBERG est l'un des pères de la mécanique quantique. Les physiciens du début du XX^e siècle ont inventé cette théorie, notamment pour mieux expliquer la structure des atomes.

l'hydrogène. Son enveloppe ne contient qu'un seul électron, chargé négativement, en rotation autour du noyau, qui porte une charge électrique positive. Sa rotation permanente l'empêche d'être attiré par l'attraction électrique qu'exerce le noyau. L'électron se comporte en fait comme une minuscule planète, dont l'orbite est stable parce que la force centrifuge due à la rotation équilibre l'attraction électrique du noyau.

Lorsqu'ils eurent compris cela, les physiciens s'étonnèrent des particularités des systèmes atomiques : tels une infinité de frères jumeaux, tous les atomes d'hydrogène ont la même structure. Or, dans le modèle planétaire, si l'électron tournait plus vite, il serait plus éloigné du noyau. Pourquoi le noyau de l'atome d'hydrogène ne fixe-t-il qu'un seul et unique électron, toujours à la même distance ? Autre mystère : les physiciens savent que les particules chargées accélérées rayonnent de l'énergie sous forme d'ondes électromagnétiques. Or, l'électron est soumis en permanence à une accélération et les particules chargées et accélérées perdent leur énergie en émettant un « rayonnement de freinage ». Pourquoi l'électron de l'atome d'hydrogène n'émet-il aucune onde électromagnétique ? Quel mécanisme le stabilise sur son orbite ? Les physiciens durent se rendre à l'évidence : la physique classique ne suffisait plus à expliquer la stabilité des atomes.

Pour répondre à ces questions, les physiciens du début du XX^e siècle ont créé la mécanique quantique. Le

concept central de la théorie quantique est celui de la probabilité d'événements. Ainsi, on calcule la probabilité qu'un électron se trouve en un endroit précis, à un instant donné ; qui plus est, la position et la vitesse d'un électron ne peuvent être déterminées simultanément. Le physicien allemand Werner Heisenberg fut le premier à formuler cette propriété quantique. Ces « relations d'incertitude de Heisenberg » s'appliquent à tous les systèmes qui ont un comportement quantique. Le calcul montre toutefois que l'approximation classique est excellente pour les systèmes macroscopiques. Si la probabilité quantique qu'une voiture traverse brusquement le mur du garage où elle était garée n'est pas nulle, elle est si faible qu'en pratique l'événement peut être considéré comme impossible !

Atomes hypothétiques

C'est justement cette incertitude sur la position de l'électron qui détermine la taille d'un atome. Dans le cas de l'atome d'hydrogène, cette taille est égale à 10^{-10} mètre. Supposons que nous observions un atome d'hydrogène, dont l'enveloppe serait beaucoup plus petite, disons cent fois plus petite. L'électron serait alors localisé avec une précision plus grande que dans l'atome d'hydrogène normal. Selon la relation d'incertitude, la précision de la vitesse de l'électron serait dans ce cas 100 fois inférieure, de sorte qu'il devrait se déplacer beaucoup plus vite que dans l'atome normal. Une vitesse plus élevée signifierait aussi une énergie supérieure. Un atome d'hydrogène hypothétiquement plus petit posséderait donc une énergie supérieure à celle d'un atome normal. Or, que ce soit en physique classique ou en mécanique quantique, les systèmes évoluent toujours vers leur état de moindre énergie. Un atome d'hydrogène plus petit que la normale ne serait pas stable, mais perdrait de l'énergie en rayonnant des ondes électromagnétiques : l'électron ralentirait, son enveloppe augmenterait et se stabiliserait lorsque l'atome aurait atteint sa taille « normale ». De même, un atome d'hydrogène hypothétique 100 fois plus grand qu'un atome normal ne serait pas stable. Pour le fabriquer, il faudrait éloigner l'électron du noyau, et donc lui apporter de l'énergie, de façon à contrecarrer l'attraction électrostatique. Pour évoluer vers un état de moindre énergie,

ce gros atome d'hydrogène perdrait très vite de l'énergie et reprendrait rapidement sa taille normale. En bref, c'est la relation d'incertitude entre la position et la vitesse des électrons qui détermine la taille des atomes.

Plus précisément, ce n'est pas la vitesse, mais la quantité de mouvement qui intervient dans la relation d'incertitude de Heisenberg. La quantité de mouvement est égale au produit de la vitesse par la masse. Pour cette raison, la taille de l'enveloppe électronique dépend directement de la masse de l'électron. La taille des atomes est donc déterminée par la masse de l'infime grain d'électricité élémentaire. Si cette masse était 100 fois plus petite que celle que l'on observe dans la nature, les atomes seraient tous 100 fois plus grands! Si elle était 1 000 fois plus grande, le diamètre de l'enveloppe de l'atome serait alors si petit qu'il serait égal à quelque dix fois le diamètre du noyau.

Ainsi, le comportement quantique des électrons impose la taille universelle des atomes. Impose-t-il aussi la taille des noyaux? Celle des électrons? Dès qu'ils ont commencé à mieux comprendre la constitution des atomes, les physiciens ont voulu répondre à ces questions. Ils sont alors entrés dans le domaine de la physique des particules. Malgré l'énorme travail d'exploration accompli, ces questions restent ouvertes.

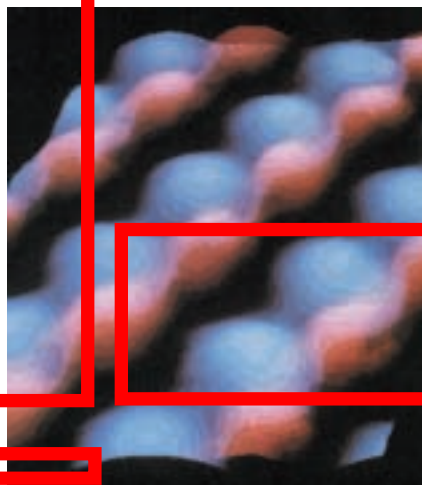
Pour comprendre pourquoi, examinons ce que les physiciens ont découvert du monde subnucléaire en commençant par les noyaux atomiques. Ces derniers représentent la plus grande partie de la masse atomique. Dans le cas de l'hydrogène, par exemple, le noyau est 1 840 fois plus lourd que l'électron. Le noyau de l'hydrogène est constitué d'une unique particule, le proton. Ceux des autres atomes contiennent tous des protons et des neutrons, des particules neutres dont la masse est comparable à celle des protons. Protons et neutrons sont apparentés, car ils sont soumis à l'interaction forte, force qui diffère de l'interaction électromagnétique (attraction électrique) laquelle s'exerce entre protons et électrons. Beaucoup plus intense, elle ne porte qu'à très courte distance : 10^{-15} mètre seulement! Pour évaluer cette portée, les physiciens ont mesuré le diamètre des protons et des neutrons du noyau, qu'ils nomment nucléons. Pour cela, ils bombardent

des noyaux avec des électrons, dont ils étudient la déviation. Ils ont ainsi montré que la charge électrique positive du proton n'est pas concentrée en un point, mais qu'elle se répartit sur une sphère d'environ 10^{-15} mètre de diamètre.

À la fin du XIX^e siècle, les physiciens constatèrent que certains éléments naturels sont instables et que leur noyau émet un rayonnement ou des particules : ils sont radioactifs. Bombardés avec des nucléons, les noyaux réagissent et se transforment en nouveaux noyaux. En 1938, Otto Hahn (1879-1968) découvrit le premier noyau (l'uranium 235) capable de se fragmenter. L'étude de ce phénomène de fission nucléaire et celle de nombreuses réactions nucléaires révélèrent de nouvelles particules.

Un bestiaire de dizaines de nouvelles particules

Les découvertes furent si nombreuses qu'à la fin des années 1950 un bestiaire de dizaines de particules différentes issues de réactions nucléaires encombrait l'esprit des physiciens. Il fallait mettre de l'ordre. Comment faire, sinon en regardant à une échelle encore plus petite? Les physiciens ont constaté qu'ils pouvaient décrire simplement les nucléons ainsi que les particules des réactions nucléaires, s'ils admettaient qu'ils sont formés de trois constituants : les quarks. Deux types de quarks différents, désignés par les symboles *u* (up) et *d* (down), sont nécessaires pour constituer les protons et les neutrons.



3. LES ATOMES D'UN ÉCHANTILLON d'arsénure de gallium aperçus au microscope à effet tunnel. Cette image est l'addition de deux prises de vue, car les propriétés électriques des atomes de gallium (en bleu) et ceux d'arsenic (en rouge) empêchent de les visualiser en même temps.

Un proton est composé de deux quarks *u* et d'un quark *d*, ce que les physiciens notent en écrivant $p = (uud)$. Dans le neutron, les rôles de *u* et de *d* sont inversés : $n = (udd)$.

Les charges électriques des quarks présentent un intérêt particulier, car elles ne sont pas des multiples entiers de la charge $+e$ du proton (opposée à la charge élémentaire $-e$ de l'électron). Ainsi, le quark *u* porte une charge égale à $2/3$ de la charge du positon, le quark *d* à un tiers de la charge de l'électron. Les charges des nucléons étant la somme de celles des quarks, la charge du proton est égale à $2e/3 + 2e/3 - e/3$, c'est-à-dire e , et celle du neutron : $2e/3 - e/3 - e/3$, c'est-à-dire 0. En fait, les charges des quarks du neutron s'équilibrent électriquement.

Au début, les physiciens supposèrent que les quarks n'étaient pas de véritables éléments constitutifs des nucléons, mais plutôt une sorte d'astuce d'écriture. Jusqu'à l'hypothèse des quarks, toutes les charges électriques observées étaient des multiples entiers de la charge e . Les physiciens considéraient que la charge de l'électron était élémentaire (la plus petite possible), et ils ne s'attendaient donc d'abord dans les quarks et leurs charges fractionnaires ($2e/3$, $e/3$, etc.) que des concepts pratiques pour décrire efficacement les particules et leur organisation. Toutefois, des expériences montrèrent bientôt que des «objets» se trouvent effectivement à l'intérieur des nucléons. Au SLAC, le grand accélérateur de particules de l'université de Stanford, on bombardait des noyaux atomiques avec des électrons préalablement accélérés jusqu'à une vitesse proche de celle de la lumière. On observa que la plupart des électrons n'étaient guère déviés de leur trajectoire, mais quelques-uns l'étaient très fortement. Ces constatations faites, les physiciens durent se rendre à l'évidence : toute substance comme si les électrons subissaient parfois, quand ils traversent de la matière nucléaire, un choc frontal avec un objet ponctuel électriquement chargé. L'analyse fine de ces expériences révéla que ces «objets» portent des charges électriques fractionnaires égales à $2e/3$ et de $-e/3$. Il s'agissait bien de quarks, des particules apparemment «infiniment» petites!

Les collisions entre les électrons et les quarks devaient être extrêmement violentes. Malgré cela, on ne détecta aucun quark pas parmi les «débris», les particules issues de la col-

lision. Il en est ainsi dans tous les chocs qui mettent en jeu des quarks. Manifestement, les forces de liaison entre quarks sont telles qu'ils sont inséparables! Quelle est la nature de forces capables d'un tel prodige?

Les physiciens des particules ne parlent pas de forces, mais plutôt d'interactions. Pour unir les concepts de la physique quantique et ceux de la relativité, ils ont dû créer une nouvelle théorie quantique : la théorie des champs. Dans ce cadre, toutes les forces qui existent dans la nature sont décrites par des échanges de particules dites médiatrices. L'existence d'une force entre quarks découle de l'échange permanent de particules médiatrices. L'importance de cette force résulte de l'intensité de l'interaction entre la particule intermédiaire et la particule massive ou chargée avec laquelle elle interagit. La première des théories des champs fut l'électrodynamique quantique, qui décrit les phénomènes électromagnétiques. Les forces qui résultent de l'interaction électromagnétique correspondent à des échanges de photons. Les forces nucléaires qui lient les protons aux neutrons du noyau, correspondent à l'échange de particules médiatrices entre quarks : les gluons (ainsi nommés parce qu'ils collent fortement les quarks les uns aux autres). Il en est de même pour les interactions faibles (bosons W et Z) et la gravitation (gravitons).

Or dans le cadre de la théorie des champs, les interactions entre particules médiatrices et particules de matière ont une particularité : elles sont « locales », c'est-à-dire qu'elles restent confinées

dans un volume infiniment petit (un point!). Ainsi, puisque la théorie des champs est fondée sur la notion d'interactions locales, l'infiniment petit y est omniprésent. L'électrodynamique quantique est devenue la plus précise des théories jamais élaborées. Elle décrit les interactions électromagnétiques avec une précision extrême. Avant d'arriver à ce succès sans précédent, les physiciens ont dû parcourir un long chemin.

L'antiparticule de l'électron

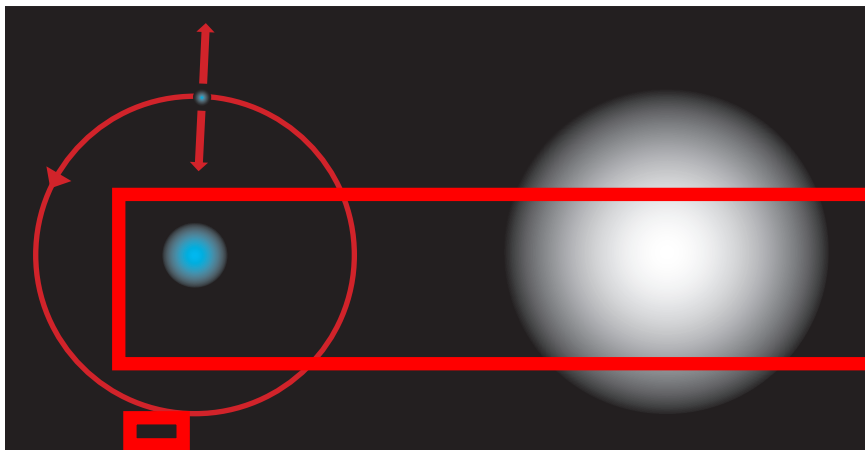
Le succès de l'électrodynamique quantique n'aurait pas été possible si l'on n'avait découvert (et découvert) une particule supplémentaire : l'antiparticule de l'électron ou positon. Les positons ont la même masse que l'électron et une charge opposée. Ils n'existent pas dans la nature, mais sont faciles à obtenir au sein des accélérateurs lors de collisions de particules. Ainsi, la collision de deux photons produit une paire électron-positon, quand l'énergie mise en jeu est assez grande. Le fait que cette collision entre particules sans masse (photons) produise des particules massives est une illustration de plus de l'équivalence entre l'énergie et la masse découverte par Albert Einstein au début du XX^e siècle. Comme en chimie, les réactions entre particules peuvent se produire dans les deux sens. Ainsi, lorsqu'un positon et un électron se rencontrent, ils s'annihilent en produisant deux photons.

L'électrodynamique quantique suppose que tous les électrons, positons et autres photons n'ont aucune structure

interne. Dans cette théorie, un électron, par exemple, est une particule ponctuelle, massive et chargée électriquement capable de participer aux interactions électromagnétiques. L'électron réel, c'est-à-dire celui qui est observable dans les expériences, diffère toutefois de cet « électron théorique », car un nuage de « particules virtuelles » l'entoure. Conséquence de la relation d'incertitude de Heisenberg, le « vide quantique » (celui que décrivent les théories quantiques) n'est pas vraiment... vide. L'énergie peut y fluctuer, ce qui rend possible la création erratique de particules, ou plus exactement de paires particule-antiparticule. Qualifiées de virtuelles par les physiciens parce qu'elles ne peuvent pas être observées dans les accélérateurs, ces paires (par exemple électron-positon) ont une « durée de vie » très brève et n'existent que dans une zone réduite. Sans influence sur les événements macroscopiques, elles altèrent toutefois l'environnement des électrons : supposons que nous amenions un électron réel en un point donné de l'espace. Chargé négativement, il repousse tous les électrons virtuels présents dans son environnement et attire les positons virtuels. Il finit par s'annihiler avec l'un d'eux. Resté seul, l'électron virtuel initialement apparié au positon (virtuel) tout juste annihilé devient réel et prend la place de l'électron d'origine (voir la figure 5). Pour un observateur extérieur (supposé suffisamment petit pour examiner une particule ponctuelle!), la trajectoire de l'électron prend une apparence étrange. Il semble disparaître sans cesse pour réapparaître instantanément ailleurs (en fait, c'est l'électron virtuel devenu réel qui réapparaît). La trajectoire de l'électron semble floue à l'observateur : un nouvel exemple concret d'incertitude sur la position au sens de Heisenberg.

Les physiciens nomment polarisation du vide cette création permanente de paires électrons-positons. Le fait que les positons virtuels prédominent au voisinage de l'électron crée une sorte de nuage de charges positives autour de la charge négative de l'électron. Ainsi cette dernière est quelque peu masquée par le nuage de particules virtuelles qui l'entoure, ce qui l'affaiblit. On ne peut observer des électrons ponctuels, et l'on dit

qu'ils sont « habillés » d'un nuage de particules positives. Les physiciens ont mesuré que la polarisation du vide commence à se manifester autour d'un électron à une distance de l'ordre du dixième de la taille d'un atome. Les électrons



4. L'ATOME D'HYDROGÈNE ne comporte qu'un seul électron. Celui-ci est attiré vers le noyau par l'attraction électrostatique et repoussé vers l'extérieur par la force centrifuge due à sa rotation autour du noyau. L'équilibre entre ces deux forces le maintient sur son orbite. Toutefois, cet « équilibre » s'entend au sens quantique. L'électron ne se manifeste pas sous la forme d'une petite masse ponctuelle en rotation sur son orbite, mais sous celle d'une présence probable au sein d'une zone appelée orbitale. L'orbitale de l'unique électron de l'hydrogène est représentée à droite. Sa taille détermine celle de l'atome.

«habillés» sont les seuls qui soient observables en pratique. Ainsi, l'électrodynamique quantique suppose des électrons infiniment petits, mais elle prédit aussi qu'ils ne peuvent être observés!

Quand les calculs mènent à un résultat infini

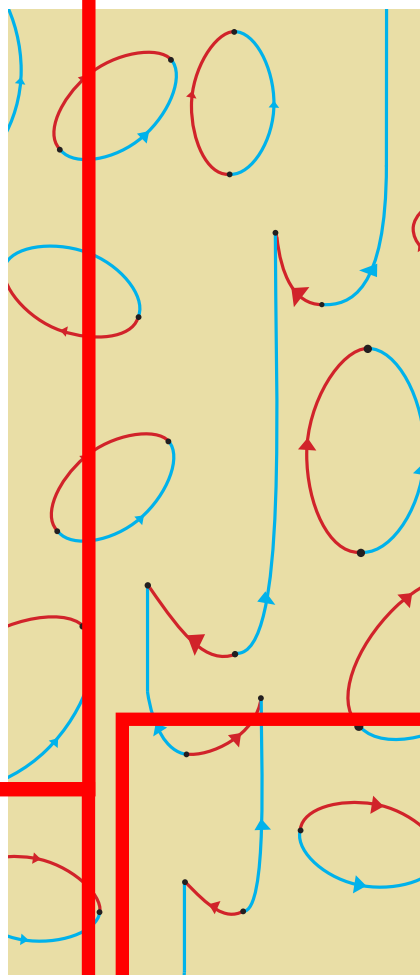
À la fin des années 1940, les physiciens ont calculé la charge de l'électron nu en prenant en compte les effets de la polarisation du vide. Ils ont eu une mauvaise surprise : les calculs mènent à un résultat infini. Ce n'est pas tout : par cette méthode, la masse de l'électron nu est également infinie! Manifestement ces résultats absurdes sont engendrés par le caractère ponctuel, donc infiniment petit, de l'électron nu. Rappelons que l'équivalence entre l'énergie et la masse nécessite que le champ électrique d'un électron contribue à la masse, puisque la présence d'un champ électrique dans l'espace implique celle d'une densité d'énergie. Puisque, dans le cadre de l'électrodynamique quantique, l'électron est ponctuel, le champ électrique devient très intense aux très courtes distances (il varie comme l'inverse du carré de la distance). À tel point que les calculs le montrent, sa contribution à la masse est infinie.

L'électrodynamique quantique fondée sur l'hypothèse d'électrons nus ponctuels et en interaction locale aboutit à un paradoxe. Cette hypothèse est-elle erronée? L'électron nu présente-t-il une structure interne, laquelle ne devient perceptible que pour de très petites distances, 10^{-20} mètre par exemple. S'il en est ainsi, cette particule n'est pas infiniment petite, mais a un rayon fini. Les calculs montrent alors que les valeurs infinies disparaissent. Toutefois, malgré d'intenses efforts, les physiciens n'ont jamais décelé de preuve expérimentale de l'existence d'une sous-structure électronique. Seul résultat obtenu : si le rayon de l'électron n'est pas nul, il est nécessairement inférieur au centième de la taille des nucléons.

Dans les années 1950, la situation restait paradoxale. L'hypothèse de l'électron nu ponctuel faisait vaciller une théorie par ailleurs prometteuse. Certains des résultats obtenus étaient précis, d'autres infinis. La description globale des interactions électromagnétiques donnée par l'électrodynamique quantique était correcte. Les physiciens ont alors tenté de séparer

le bon grain de l'ivraie, c'est-à-dire d'éliminer les résultats infinis des calculs. Pour ce faire, ils ont ajusté la charge de l'électron habillé au résultat trouvé expérimentalement, en ignorant tout simplement la charge de l'électron nu. Ils firent de même pour la masse de l'électron habillé, imposant ainsi la masse mesurée. Nommée renormalisation, cette méthode est d'une étonnante efficacité : les quantités infinies sont éliminées et les résultats obtenus pour les grandeurs physiques mesurables sont les plus précises jamais obtenues.

L'un des pères de la renormalisation fut le physicien américain Richard Feyn-



5. LE VIDE aussi a une nature quantique, ce qui se manifeste notamment par la création permanente de paires particule-antiparticule. Constituées d'un électron et d'un positron, elles sont dites virtuelles parce qu'on ne peut les observer directement. Conséquence : un électron qui se propage à travers le vide s'annihile souvent avec un positon issu d'une paire virtuelle. Il semble alors réapparaître là où l'électron issu de la paire virtuelle se trouve et y poursuivre sa trajectoire. Son parcours apparaît donc flou à un observateur extérieur supposé assez petit pour l'examiner de près. Il s'agit là de l'une des manifestations de l'incertitude quantique sur la position.

mann. Lors d'un congrès aux États-Unis, il monta pour la première fois comment la méthode élimine les infinis qui apparaissent dans les calculs de la théorie de l'électrodynamique quantique. Ce progrès fit de l'électrodynamique quantique la plus précise des théories.

Les problèmes d'infinis dans les théories locales (interactions ponctuelles mettant en jeu des particules ponctuelles) se rencontrent dans d'autres théories que l'électrodynamique quantique. Ainsi, les interactions entre quarks sont décrites dans le cadre d'une théorie similaire à l'électrodynamique : la chromodynamique quantique. Cette théorie décrit l'interaction forte, laquelle provoque, notamment, l'assemblage des quarks par groupes de trois. Elle tire son nom du grec *chromos* qui signifie couleur, car on attribue aux quarks, par exemple au quark *u*, un nouveau type de charge, dénommée couleur, qui ne peut prendre que trois valeurs : rouge, vert ou bleu. Il y a ainsi un quark *u* rouge, un vert et un bleu, comme il y a des quarks *d* rouges, verts ou bleus. Le fait qu'il y ait trois couleurs signifie que les nucléons sont composés de trois quarks. Nous avons déjà souligné que les interactions entre quarks sont véhiculées par des particules d'énergie nommées gluons, les «médiateurs» de l'interaction forte (si intense qu'il est impossible de séparer les quarks). Pour observer les quarks et les gluons, les physiciens ont été obligés de recourir à des méthodes indirectes : ils les ont bombardés avec des particules insensibles aux interactions fortes, des électrons, des neutrinos ou des photons. L'existence des gluons a été prouvée pour la première fois au synchrotron allemand de Hambourg (DESY) à la fin des années 1970.

Pas de structure interne dans les quarks

Dans la théorie actuelle, les quarks et les électrons sont des particules élémentaires, c'est-à-dire qu'ils n'ont pas de structures internes. Comme dans le cas des électrons, les physiciens ont essayé de trouver une sous-structure pour les quarks, mais sans succès jusqu'à aujourd'hui. Une limite supérieure pour le rayon d'une éventuelle sous-structure de quarks a été obtenue au grand accélérateur de particules LEP du CERN, à Genève, mais aussi avec le synchrotron DESY et le

Tevatron, l'accélérateur du laboratoire américain Fermilab, à Chicago. Elle est du même ordre de grandeur que celle des électrons : 10^{-18} mètre.

L'électrodynamique et la chromodynamique constituent deux des trois piliers de la théorie générale des particules acceptée aujourd'hui : le modèle standard. Cette théorie a fourni un cadre théorique cohérent, où l'on parvient à décrire toutes les interactions des particules élémentaires. Outre l'électron, il existe dans la nature cinq autres particules apparemment sans structure interne et qui ne sont pas soumises aux interactions fortes : le muon, le tau, le neutrino électronique, le neutrino muonique et le neutrino tau. Ce sont toutes des leptons ou des particules proches des leptons. L'électron est le plus léger des leptons électriquement chargés. Un autre lepton chargé est le muon, une sorte de gros électron dont la masse est environ 207 fois supérieure à celle de l'électron. Le tau est un lepton chargé dont la masse est 3 536 fois celle de l'électron. Le muon et le tau sont instables et se décomposent rapidement après leur production lors de collisions entre particules. Nous les avons déjà citées : trois particules électriquement neutres sont associées aux leptons chargés électriquement : les neutrinos (neutrino électronique, neutrino muon, neutrino tau). Dans le modèle standard, les neutrinos sont supposés sans masse (mais des expériences japonaises récentes leur en attribuent une). Il existe trois paires de leptons, chacune constituée d'une particule chargée et d'un neutrino.

Il existe six quarks différents

Si la chromodynamique quantique décrit la matière nucléaire (assemblage de nucléons) à partir des quarks *u* et *d*, il existe en fait quatre autres types de quarks. Ils sont désignés par les lettres *s* (quark étrange, *strange* en anglais), *c* (quark charme), *b* (pour *bottom*, qui signifie bas en anglais) et *t* (pour *top*, haut en anglais). Les physiciens les notent en général par paires de masses comparables. Les quarks (*c,s*) et (*t,b*) sont créés lors de la production de particules plus lourdes que les nucléons. À l'instar du muon ou du tau, ils sont instables, ce qui provoque la désintégration rapide des particules qu'ils constituent. La valeur de leurs masses laisse les physiciens perplexes : tandis que celle des quarks *u* et *d* sont quasi nulles, celle

du quark *t* est proche de la masse du noyau atomique de l'or. Un noyau qui contient pas moins de 197 nucléons ! On peut regrouper en trois paires les six leptons et quarks : si nous regroupons chaque paire de leptons avec une paire de quarks, nous obtenons trois familles de leptons-quarks. La première, la famille constituée par les quarks *u* et *d*, l'électron et son neutrino, contient tous les constituants de la matière atomique. Les quarks de la troisième famille *t* et *b*, et surtout le tau, sont les objets les plus lourds du modèle standard.

En fait, l'existence de l'interaction faible impose à elle seule le regroupe-

ment des leptons et des quarks par familles. Celle-ci est responsable de la désintégration radioactive (dite bêta) des noyaux atomiques. Comme les autres forces, l'interaction faible est due aussi à l'échange de particules médiatrices, les particules *W* et *Z*. À la différence des photons et des gluons de l'interaction forte, les particules médiatrices de l'interaction faible ont une masse élevée : près de 100 000 gigaélectronvolts (un gigaélectronvolt est égal à un milliard d'électronvolts). La force faible, véhiculée par les particules *W* agit à l'intérieur des paires de leptons et de quarks décrites. Lors de la désintégration radioactive d'un noyau atomique, par exemple, un quark *d* est changé en quark *u* ; un électron et un neutrino sont produits simultanément. Les interactions électromagnétiques faibles peuvent être réunies dans le cadre d'une théorie unifiée, appelée théorie des interactions électro-faibles. Le modèle standard actuel des particules élémentaires est donc la synthèse des théories des interactions électro-faibles et des interactions fortes.

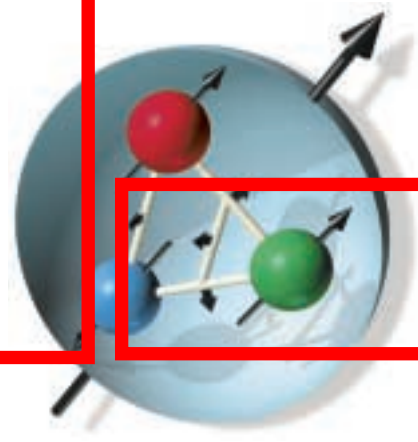
Cette théorie décrit bien les particules de notre monde et leurs interactions (l'interaction gravitationnelle reste à part). Par cette théorie, on évalue bien les masses des principaux constituants de la matière, les nucléons d'après celles de leurs constituants. Ainsi, la masse du neutron et celle, identique du proton correspondent à la somme des masses de leurs quarks internes et de l'énergie cinétique des gluons qu'ils échangent (un nouvel exemple d'équivalence masse-énergie au sens d'Einstein).

Reste une question ouverte : quelle est l'origine des masses des leptons et des quarks ? Malheureusement le modèle standard n'apporte pas de réponse. Pour les introduire dans la théorie en tant que paramètres, les physiciens ont fait appel à un subterfuge : ils ont introduit une particule hypothétique, dont l'unique fonction est de conférer leurs masses à ces particules. Il s'agit du boson de Higgs, une particule ainsi nommée d'après Peter Higgs, le théoricien qui la proposa. Le boson de Higgs a une masse si grande que les collisions hypothétiques qui pourraient le produire n'ont pu être réalisées.

Toutefois, les énergies nécessaires seront disponibles dès que le grand accélérateur LHC du CERN, actuellement en construction, fonctionnera, dès 2005. On devrait alors savoir si le boson de Higgs existe ou si la masse des



6. RICHARD FEYNMAN, l'un des pères de l'électrodynamique quantique. Il contribua à résoudre le problème des infinis qui apparaissent dans les calculs faits dans le cadre de cette théorie.



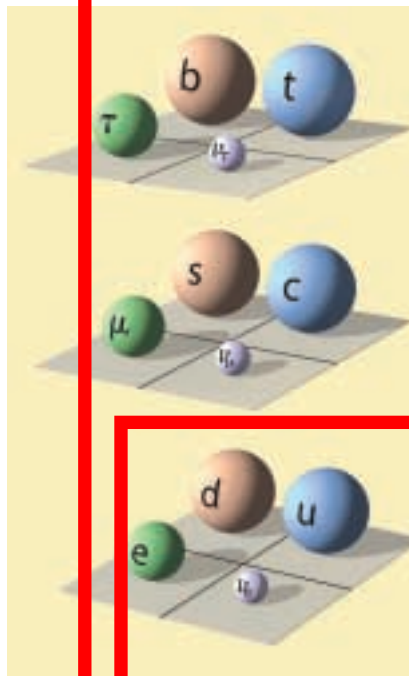
7. LE PROTON est constitué de trois quarks, lesquels peuvent prendre chacun trois «couleurs» : rouge, vert, bleu. Les flèches représentent les spins, les barres blanches, les échanges de gluons, les «particules collantes» qui maintiennent la cohésion des trois quarks du proton.

constituants élémentaires a une autre origine.

Ainsi, c'est la physique de l'infiniment petit. Au début du XX^e siècle, on a observé le spectre des états d'énergie de l'électron de l'hydrogène. Bien qu'il ait une structure simple, on ne sut l'expliquer avant l'avènement de la théorie quantique. Aujourd'hui, c'est le spectre des leptons et des quarks que les physiciens cherchent à expliquer. Dans cette science de l'infiniment petit, les progrès naissent le plus souvent d'expériences qui nous livrent de nouvelles structures simples, mais inattendues. Rien de ce genre ne s'est produit au cours des dernières années. Les arguments en faveur des leptons et des quarks du modèle standard sont-ils satisfaisants? Ces particules sont ponctuelles et représentent des singularités dans l'espace. Des particules infiniment petites peuvent-elles avoir une masse? Si oui, pourquoi la masse du muon est-elle 207 fois supérieure à celle de l'électron, bien que par ailleurs ces deux types de particules soient indiscernables. Peut-être serons-nous, un jour, contraints de renoncer au concept de particules ponctuelles. On peut supposer que les masses des leptons et des quarks résultent d'une sous-structure de ces particules, comme c'est le cas pour le proton. Peut-être obtiendra-t-on de nouveaux indices sur les constituants des leptons et des quarks grâce aux expériences futures réalisées dans les grands accélérateurs.

Perfectonner le modèle standard

Une autre raison incite les physiciens à perfectionner le modèle standard : ils voudraient ajouter la gravitation aux trois interactions – électromagnétique, faible et forte – qu'il décrit déjà. En fait, deux conceptions s'opposent : d'un côté, le modèle standard avec ses théories quantiques et ses interactions locales entre particules ponctuelles et, de l'autre, la relativité générale, où la gravitation résulte d'une déformation de l'espace-temps. Que faire? Considérer la relativité générale telle que nous la connaissons comme la théorie classique d'une théorie quantique plus générale, la gravitation quantique (qui restera à élaborer)? Une théorie quantique de la gravitation serait en même temps une théorie quantique de l'espace et du temps. Or, le modèle standard et la plupart des théories physiques sont bâties sur l'espace et sur le temps, et toute



8. LE MODÈLE STANDARD est la théorie la mieux acceptée par les physiciens des particules, les seuls le testent par l'expérience depuis 50 ans. Il est fondé sur trois familles au sein desquelles deux quarks sont associés à deux leptons. La matière et ses manifestations les plus courantes sont entièrement dues à la famille (u, d, e, ν_e), laquelle comprend les quarks u et d , l'électron et son neutrino associé. Les autres familles contiennent des particules beaucoup plus massives, qui se manifestent dans les désintégrations de particules lourdes, observables seulement dans les accélérateurs ou dans le Cosmos.

modification des fondations risquerait de les faire vaciller. Néanmoins, aux petites distances (inférieures à 10^{-35} mètre), les incertitudes dues à la théorie quantique perturbent vraisemblablement la structure de l'espace-temps. Si un électron est réellement une singularité de l'espace-temps avec une masse (une particule infiniment petite), cette singularité devrait disparaître à cette échelle. À quoi ressemble alors un électron? On l'ignore.

Selon certains théoriciens, les leptons et les quarks sont des manifestations de minuscules objets unidimensionnels, les «supercordes». Sortes de petites structures en forme de fils, les supercordes seraient moins singulières qu'un point. Or, la formulation de la gravitation quantique d'une théorie des supercordes. On devrait alors supposer que l'espace-temps décrit à l'aide de quatre dimensions (trois d'espace et une de temps) acquiert dix dimensions supplémentaires. Cela semble *a priori* contredire

nos observations, mais seules quatre des dix dimensions auraient des effets macroscopiques ; les autres ne seraient perceptibles qu'à petite distance, et y rendraient la gravitation plus intense. Prenons une image : si nous enroulons une feuille de papier sous forme de cylindre, elle ressemble, vue de loin, à une droite à une dimension. La deuxième dimension, celle qui permet de tourner autour du cylindre, n'est perçue qu'à courte distance. Ainsi en va-t-il des dimensions cachées de l'Univers : elles sont peut-être là, mais si petites que nous ne les voyons pas.

Si notre espace possède effectivement d'autres dimensions à petite échelle, elles pourraient être mises en relation avec certains phénomènes observés par les physiciens des particules. Ainsi, on ignore encore pourquoi la nature semble apprécier le chiffre trois : il y a trois dimensions d'espaces, mais aussi trois familles de leptons, trois familles de quarks, ces derniers ont trois couleurs... Le rôle particulier du chiffre 3 découle peut-être des dimensions cachées et leur nombre. En d'autres termes, peut-être existe-t-il une explication géométrique de ces structures fondamentales du modèle standard. Peut-être réaliserons-nous que les leptons et les quarks ne sont pas ponctuels, mais qu'ils ont une sous-structure. Nous pourrions aussi découvrir pourquoi leurs masses ont certaines valeurs? Pour le moment, le mystère demeure, mais Bertolt Brecht n'a-t-il pas écrit, en 1921, dans son Journal : «Là où il n'y a pas de mystère il n'y a pas de vérité?»

Harald FRITZSCH est professeur de physique théorique à l'Université de Munich.

Harald FRITZSCH et Zhi-zhong XING, *Neutrino Mixing and Maximal CP Violation*, Acta Phys. Polon. B31:1349-1364, 2000.

Harald FRITZSCH et Zhi-zhong XING, *Mass and Flavor Mixing Schemes of Quarks and Leptons*, in Prog. Part. Nucl. Phys., 45:1-31, 2000.

Harald FRITZSCH, *An Equation that Changed the World : Newton, Einstein, and the Theory of Relativity*, University of Chicago Press, 1994.

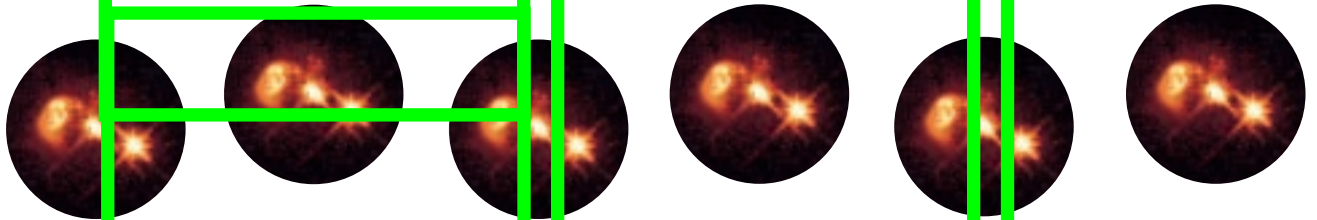
Harald FRITZSCH, *Quarks. Urstoff unserer Welt*, Piper, 1992.

Harald FRITZSCH, *Vom Urknall zum Zerfall. Die Welt zwischen Anfang und Ende*, Piper, 1999.

L'infiniment grand

GERHARD BÖRNER

L'Univers semble être brusquement sorti d'un état initial infiniment dense, que les physiciens ont des difficultés à imaginer et à décrire par une théorie.



L'infiniment grand, l'Univers, est-il sorti de l'infiniment petit, c'est-à-dire d'un point matériel? Nul ne le sait. Toutefois, de nombreuses théories prédisent un comportement extrême de la matière : la contraction sans fin des grandes masses sous l'effet de leur propre poids, jusqu'à la formation d'un point infiniment dense. Pour les physiciens, ces théories sont imparfaites ; elles annoncent des états aberrants, qu'ils ne sauraient admettre. Cependant, ils sont persuadés que ces points singuliers indiquent des tendances réelles des systèmes cosmiques, dont, au premier chef, l'Univers lui-même. L'observation de plus en plus probable de trous noirs, les traces laissées par une lointaine explosion originelle d'une nature mystérieuse sont autant d'indices qui les confortent dans leur conviction. Ils étudient donc ardemment les «maladies» de leurs édifices théoriques, au premier rang desquelles la relativité générale d'Albert Einstein et le modèle cosmologique standard, et inventent de nouvelles théories plus générales. Dans le même temps, ils observent de près les catastrophes cosmiques, telles les supernovae, les trous noirs, les quasars et les tourbillons ultra-rapides de nuages stellaires. Leur but : tout comprendre avant que l'Univers ne... disparaisse.

Intuitivement, il nous semble évident que les grandes masses soumises à la force de gravitation auront une destinée catastrophique. La gravitation agit à l'identique sur toutes les particules massives : elle les attire. De plus, elle a

une portée infinie, même quand elle est faible. Par ailleurs, si nous ajoutons particule sur particule à une masse quelconque, la force de gravitation croît avec le nombre de particules. Elle finit toujours par l'emporter sur toutes les forces qui s'opposent à son effet, que ce soit la pression thermonucléaire au cœur d'une étoile, la force répulsive interne aux noyaux atomiques ou la pression d'un gaz de particules. À cela s'ajoute le fait que non seulement la matière, mais aussi l'antimatière, ou même toute forme d'énergie, a un effet gravitationnel. Réciproquement, toute forme d'énergie subit la pesanteur. C'est pourquoi, si une gigantesque pression interne peut, jusqu'à un certain point, maintenir une masse en équilibre, les formes d'énergie qui créent cette pression contribuent à son champ de pesanteur et participent finalement à son effondrement gravitationnel.

Les singularités de l'espace-temps

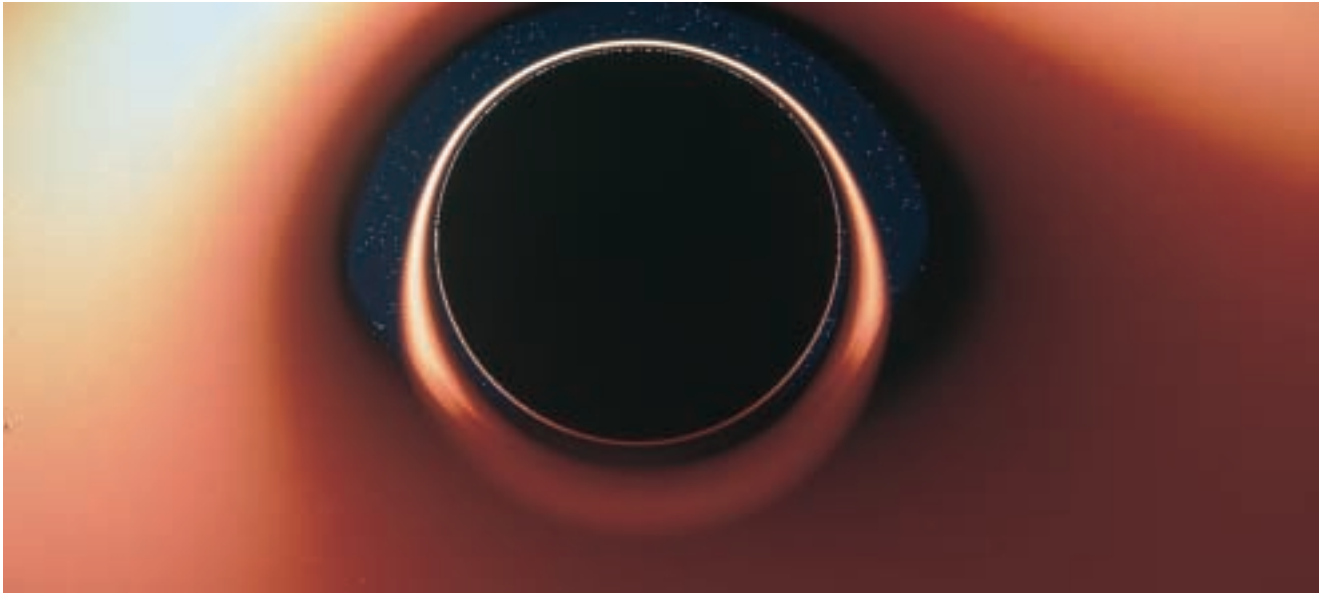
Les solutions des équations compliquées de la relativité générale présentent des singularités, dont on a longtemps pensé qu'elles étaient des conséquences des hypothèses de symétrie, introduites dans les calculs pour en faciliter la résolution. Les physiciens espéraient qu'elles disparaîtraient si l'on perturbait légèrement ces solutions, pour les rendre un peu moins symétriques. Toutefois, dans les années 1965 à 1970, les mathématiciens et les physiciens britanniques Roger Penrose et

Stephen Hawking montrèrent que des singularités de l'espace-temps apparaissent aussi dans le cas général, non symétrique, et qu'elles sont, en général, stables vis-à-vis de petites perturbations. En d'autres termes, ces physiciens découvraient que les singularités sont des cas limites incontournables en relativité générale, des «maladies» dont la nature ne pourra être mieux comprise sans une extension quantique de la théorie d'Einstein. Toutefois, de nombreux physiciens pensent que les théories physiques actuelles suffisent à décrire la situation à proximité de ces singularités, nommées «lieux d'infinis», quelle que soit leur nature réelle.

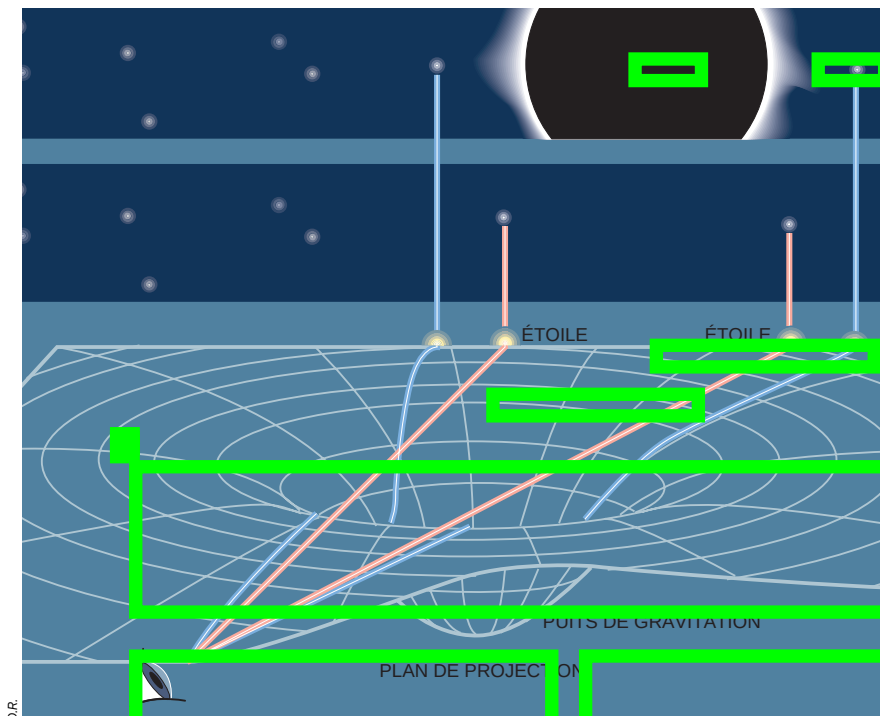
Les trous noirs et le Big Bang sont-ils des lieux d'infinis? Avant d'aborder cette question, rappelons quelques aspects fondamentaux des théories de la gravitation.

Isaac Newton, le savant anglais du XVII^e siècle, fondateur de la physique théorique classique, admettait l'existence d'un espace absolu. Il lui attribuait implicitement les propriétés géométriques que le Grec Euclide avait déduites d'axiomes fondamentaux «intuitifs». Newton supposait aussi qu'un «temps universel» mesurable sans ambiguïté, caractérise tous les événements que l'on peut étudier. Ainsi, un espace absolu muni de cadres solides et inaltérables fixe le cadre où évoluent les corps

1. DIVERSES SIMULATIONS de trous noirs. Au cours de leur aspiration par un trou noir, les gaz interstellaires s'échauffent au point de former un plasma ultrachaud : le disque d'accrétion (les zones blanc-jaune).



J.A. Marck/Sigma



D.R.

2. SELON LA THÉORIE de la relativité générale, les grandes masses courbent l'espace-temps dans leur voisinage, ce qui provoque une déviation apparente de la lumière. Ce phénomène a été prouvé par des observations pendant une éclipse totale du Soleil. Si la lumière n'était pas déviée par la masse du Soleil, elle suivrait les trajectoires directes (en rose). La présence du Soleil dévie la lumière, de sorte qu'un observateur voit les étoiles proches du Soleil (en bleu) plus à l'extérieur qu'elles ne le sont en réalité.

matériels au cours d'un temps lui aussi parfaitement défini. On constate que tous les objets tendent à se déplacer uniformément en ligne droite, sauf quand des forces extérieures perturbent ce déplacement. L'une de ces forces est la force de gravitation, laquelle se manifeste, par exemple, par la pesanteur. Cette force, attraction mutuelle entre deux corps massifs, est proportionnelle au produit de leurs masses et inversement proportionnelle au carré de la distance qui les sépare.

Albert Einstein reformule les conceptions de Newton

Albert Einstein a reformulé les conceptions de Newton sur l'espace et sur le temps. À la fin du XIX^e siècle, après les expériences de Michael Faraday et les nouvelles théories de James Clerk Maxwell, les physiciens ont commencé à considérer les champs électromagnétiques comme des objets physiques. On admettait donc que les ondes électromagnétiques se propagent à la vitesse de la lumière même dans le vide, et ce indépendamment des mouvements de leur source. Comment réconcilier cette observation avec la physique classique de Newton? Dans cette théo-

rie, la vitesse apparente de propagation de la lumière ne doit-elle pas augmenter ou diminuer, suivant que la source s'approche ou s'éloigne de l'observateur? Ces difficultés conduisirent Einstein à admettre, en 1905, la défaillance des concepts de temps et d'espace absolus, du moins pour la description de mouvements. Il postula qu'on ne peut déterminer si deux événements sont simultanés ou non, car la réponse dépend du référentiel de l'observateur. Einstein découvrit qu'il fallait abandonner les notions de distance spatiale entre deux points et d'écart temporel entre deux instants, pour utiliser plutôt les intervalles entre événements dans l'espace-temps à quatre dimensions (voir la figure 4).

Cette première extension de la physique classique, la «relativité restreinte», a fourni des bases théoriques fiables à toutes les branches de la physique où la force de gravitation peut être négligée. Elle est en particulier devenue incontournable en physique des particules.

En 1915, grâce à une autre amélioration de la description de l'espace-temps, Einstein parvint à prendre en compte la gravitation. L'idée sur laquelle repose sa théorie de la gravitation, nommée «relativité générale»,

est celle d'un espace-temps dont la structure n'est pas imposée *a priori*, mais qui découle de la répartition des masses et des énergies présentes. Chaque corps déforme – à proximité – l'espace-temps où il évolue. Réciproquement, la géométrie locale de l'espace-temps détermine la dynamique des corps. L'ensemble des déformations de l'espace-temps dues à tous les corps définit le champ de gravitation.

La gravitation agit aussi sur la lumière

La gravitation agit aussi sur la lumière, dont les rayons sont déviés lorsqu'ils approchent de corps massifs (voir la figure 2) : la propagation de rayonnements dépend de la répartition des

Ce lien intime entre espace, temps et matière ne simplifie pas les calculs ; cependant, la relativité générale représente un grand progrès en direction d'une physique unifiée. La gravitation fait partie intégrante de l'espace-temps et ne doit plus être traitée en tant qu'interaction physique indépendante, comme le faisait Newton.

Si l'on admet que la gravitation est inscrite dans la structure liant l'espace et le temps, comment détermine-t-on l'attraction mutuelle des objets massifs et comment en déduit-on les mouvements des corps célestes? Tentons de nous représenter l'espace-temps comme une toile de matériau élastique. Dans cette image, les objets de l'Univers pourraient être des boules de masses diverses posées sur la toile. Elles créent des déformations plus ou moins profondes. Une masse importante, telle celle du Soleil, se retrouve au fond d'un «entonnoir», profond et évasé ; une masse faible, comme celle de la Terre, crée à peine une légère dépression. Des corps qui s'aventureraient trop profondément dans l'entonnoir du Soleil tomberaient vers lui, entraînés par la pente, l'équivalent de l'attraction solaire. D'autres corps passant à distance ne seraient pratiquement pas perturbés par la déformation de la toile créée par le Soleil ; leurs trajectoires seraient quasi rectilignes (de tels corps ne seraient pas liés gravitationnellement au Soleil).

La Terre peut se maintenir sur un circuit fermé (l'orbite terrestre) le long des pentes de l'entonnoir solaire, parce que sa vitesse est très exactement ajustée à sa distance à l'astre. C'est bien la déformation de l'espace-temps dans

le voisinage du Soleil qui détermine les trajectoires.

Ainsi, l'existence d'une force gravitationnelle en un point de la toile accélère les astres. La Terre sur son orbite elliptique est, par exemple, soumise à une accélération centrifuge. Nous avons là une illustration du «principe d'équivalence» entre gravitation et accélération, l'un des principes fondamentaux de la théorie d'Einstein suivant lequel les effets de la gravitation se confondent avec ceux d'une accélération. Le plus surprenant est peut-être que, malgré sa formulation apparemment différente, la théorie d'Einstein rejoint celle de Newton lorsque les champs de force de gravitation et les vitesses en jeu restent faibles. Notons que la théorie de la gravitation de Newton inclut implicitement le principe d'équivalence, puisqu'elle pose l'égalité entre la masse inertielle (liée à l'accélération centrifuge) et la masse pesante (liée à la force de gravitation).

La relativité générale a pour le moment résisté à tous les tests expérimentaux. Ses succès sont impressionnants. Elle explique les apparentes anomalies de la précession de l'orbite de Mercure, connues depuis 1859. La déviation des rayons lumineux dans le champ de force de gravitation du Soleil, prédite par Einstein, fut confirmée par l'observation de rayons lumineux passant à proximité du Soleil lors d'une éclipse solaire en 1919 (et qui, normalement, ne sont pas observables en raison de la trop forte luminosité du Soleil). On a aussi vérifié, à plusieurs reprises, que la gravitation et le mouvement combinés avaient une action sur le temps que mesurent des horloges.

En 1916, deux mois après la publication par Einstein de la théorie de la relativité générale, Karl Schwarzschild obtint une solution des nouvelles équations : la «métrique de Schwarzschild». Devenue le prototype même du «trou noir», cette solution décrit d'une façon générale la déformation de l'espace-temps au tour d'une répartition de masses à symétrie sphérique.

Imaginons que nous puissions concentrer une boule de matière, telle que le Soleil, dans une sphère dont le rayon sera de plus en plus petit. Remarquons que c'est ce que faisait déjà Newton, quand il évoquait des «masses ponctuelles». Au voisinage de cette boule de matière, la solution de Schwarzschild est valable. Dépasse-t-on, toutefois, dans ce processus, un cer-

tain rayon limite, que les coordonnées trouvées par Schwarzschild, dans le cas statique (c'est-à-dire indépendant du temps), perdent leur légitimité. Le passage à la limite ne permet donc pas de retrouver la masse ponctuelle de Newton, mais montre seulement qu'un rayon limite existe, le «rayon de Schwarzschild». On a longtemps ignoré ce qui se produit à l'intérieur de ce rayon.

Puis, les théoriciens ont trouvé des solutions des équations de la relativité générale qui décrivent le domaine intérieur du rayon de Schwarzschild. La structure de l'espace-temps intérieur peut être analysée en étudiant comment la lumière se propage. Loin à l'extérieur du rayon de Schwarzschild, tout se passe comme dans l'espace à trois dimensions habituel. En revanche, quand on se rapproche du rayon de Schwarzschild, les rayons émis sont incurvés. Quelle que soit sa direction initiale, un rayon proche du rayon de Schwarzschild est si fortement dévié qu'il ne peut s'échapper. Il reste comme à la surface qui a pour rayon le rayon critique de Schwarzschild. Émise à l'intérieur du rayon de Schwarzschild, la lumière est soumise à la courbure de l'espace-temps ; elle aboutit toujours au point singulier de rayon nul.

L'Univers séparé en deux

Si la lumière extérieure pénètre sans problème à l'intérieur du rayon de Schwarzschild, aucun rayonnement lumineux ne peut en sortir. Le rayon de Schwarzschild définit une structure de l'espace-temps particulière en ce sens qu'un horizon sépare l'Univers en deux parties. Comme aucune information ne peut être transmise à une vitesse supérieure à celle de la lumière, les photons et aussi les particules restent confinés à l'intérieur. L'horizon équivaut à une mem-

brane imperméable à l'énergie et aux diverses informations, mais seulement de l'extérieur vers l'intérieur. Tout signal transmis vers l'intérieur aboutit *in fine* à la singularité centrale. En fait, la théorie de la relativité générale stipule que toute masse intérieure à l'horizon évolue pour devenir un point massif singulier. Au-delà de cet effondrement, espace et temps échangent leurs fonctions : en deçà de l'horizon l'espace «s'écoule et passe» comme le temps dans notre réalité. Les distances se consomment irrémédiablement jusqu'à la singularité, et tout se perd en un point unique.

L'espace-temps décrit par la solution de Schwarzschild est donc vide –privé de matière tant à l'intérieur qu'à l'extérieur de l'horizon–, à l'exception du centre où est concentrée toute la masse. Un observateur du monde extérieur perçoit les effets gravitationnels de ce point massif. Le phénomène que les astrophysiciens nomment «trou noir» n'est ni un objet matériel ni la manifestation d'un rayonnement, c'est un trou dans l'espace-temps.

L'existence d'horizons (que les physiciens nomment frontières de causalité) est un résultat remarquable de la relativité générale. La singularité se situe d'un côté de l'horizon et n'a aucun lien avec d'éventuels observateurs extérieurs, c'est-à-dire qu'elle ne leur envoie jamais aucune information. Elle est comme masquée par une espèce de «censure cosmique»! Les étoiles habituelles, les planètes et autres objets cosmiques sont bien sûr plus grands que leurs rayons de Schwarzschild respectifs. Celui-ci serait de trois kilomètres pour le Soleil, et seulement de 0,9 centimètre pour la Terre. La différence énorme entre les rayons de Schwarzschild et les tailles réelles des astres signifie que, dans leur voisinage, l'espace-temps est pratiquement plat.



3. UN DISQUE DE GAZ et de poussière de 3 700 années-lumière de diamètre a été découvert au cœur de la galaxie elliptique NGC 7052 par le télescope spatial Hubble. Ci-dessus, NGC 7052 telle qu'elle apparaît dans le télescope. Les astronomes ont déduit de la vitesse avec laquelle la matière tourne autour du centre qu'une masse égale à 300 millions de masses solaires doit s'y trouver. La seule explication plausible de l'existence d'un objet à la fois si massif et si petit est qu'il s'agit d'un trou noir.

Les particules subatomiques, les protons par exemple, ont 10^{39} fois la taille de leur rayon de Schwarzschild. Décidément non, la gravitation ne joue aucun rôle dans le monde des particules élémentaires!

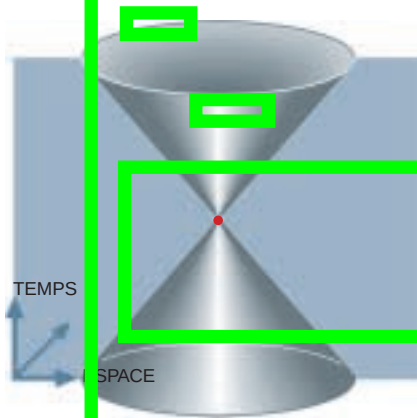
Il en va tout autrement pour une étoile à neutrons, dont la masse se compare à celle du Soleil. Son rayon égal à une dizaine de kilomètres seulement n'est que trois fois supérieur à la valeur de son rayon de Schwarzschild. C'est pourquoi la relativité générale est essentielle dans la description de la structure d'une telle étoile. Lorsqu'elle a épuisé son carburant nucléaire, une étoile à neutrons de masse suffisamment grande s'effondre sous son propre poids et disparaît finalement derrière son horizon de Schwarzschild, dans la singularité qu'elle crée. Pour un observateur situé à la surface, l'ensemble ne dure que le temps d'une chute libre, c'est-à-dire quelques secondes. En revanche, un observateur éloigné voit la surface stellaire se contracter de plus en plus lentement : elle n'atteint pas l'horizon en un temps fini. Pour lui, l'image se fige au rayon de Schwarzschild. Cette évolution s'explique par la propagation des ondes lumineuses près des horizons.

Si un cosmonaute suicidaire décidait de se laisser tomber dans un trou noir en émettant des éclairs lumineux à intervalles réguliers, un observateur détecterait des impulsions de plus en plus espacées dans le temps ; le dernier flash émis de l'horizon, ne parviendrait jamais à l'observateur (voir la figure 5). Ce décalage temporel est accompagné d'un autre phénomène : la longueur d'onde de la lumière émise est fortement décalée vers le rouge (c'est-à-dire qu'elle est décalée vers de plus grandes longueurs d'onde). Cet étirement de longueur d'onde croît exponentiellement à mesure que la source se rapproche de l'horizon de Schwarzschild. La luminosité reçue s'atténue rapidement, et l'étoile devient vite invisible.

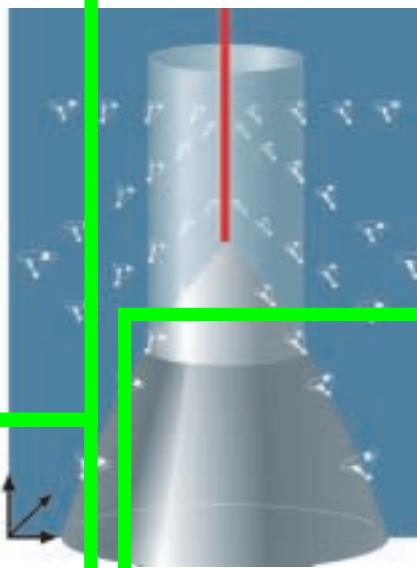
La solution de Schwarzschild ne dépend que de la masse du trou noir considéré. Elle n'est pas valable lorsque l'objet effondre tourne sur lui-même. Or, la plupart des étoiles tournent sur elles-mêmes. Un phénomène qui ne disparaît pas lors de l'effondrement. Dans ce cas de figure, les configurations possibles d'espace-temps et d'horizons se multiplient par rapport au cas simple de Schwarzschild. La situation devient encore plus complexe quand le trou

noir porte une charge électrique. Pour le monde extérieur, un trou noir est complètement défini par sa masse, son moment cinétique et sa charge.

Quoi qu'il en soit, par définition, les trous noirs n'émettent aucun rayonnement,



4. À PARTIR DU POINT de l'espace-temps marqué en rouge, seules les régions contenues à l'intérieur du cône de lumière sont accessibles. Le cône lui-même représente la zone de l'espace-temps que peut atteindre un signal lumineux (et l'information qu'il transporte) depuis le même point.



5. REPRÉSENTATION de l'effondrement gravitationnel d'un astre dans l'espace-temps à trois dimensions (deux d'espace et une de temps). Le temps s'écoule vers le haut, de sorte que les états successifs de l'astre sont représentés par chacune des coupes horizontales. L'astre s'effondre régulièrement (surface grise cylindrique), jusqu'à ce que son champ de gravité superficiel empêche tout rayon lumineux d'échapper et qu'apparaisse ainsi un horizon des événements (cylindre grisé). L'effondrement se termine en singularité (ligne rouge), c'est-à-dire par un état infiniment dense et infiniment petit. Les petits cônes de lumière illustrent la propagation de signaux lumineux en différents points de l'espace-temps. Plus un point est proche de la singularité, plus la lumière est déviée vers elle et plus son cône de lumière est resserré.

et qu'ils ne soient pas observables directement, ils peuvent, dans certaines circonstances, révéler leur présence : quand ils englobent la matière qui passe à proximité. Les masses ainsi aspirées ne se précipitent pas sur le trou noir en ligne droite, mais en spirale. Elles s'accumulent d'abord dans un anneau aplati, le «disque d'accrétion». Ce faisant, elles s'échauffent tant qu'elles deviennent des sources intenses de rayons X, que les astronomes enregistrent grâce à des instruments placés à bord de satellites.

Un certain nombre de sources X compactes, susceptibles d'abriter des trous noirs ont déjà été localisées. Certaines font partie de systèmes stellaires binaires (deux étoiles qui tournent l'une autour de l'autre) et gravitent autour d'un astre visible, directement observable (voir la figure 6). La mesure des paramètres dynamiques des orbites du système indique la masse de l'émetteur des rayons X. Quand cette dernière est supérieure à la masse maximale compatible avec une étoile à neutrons, il peut s'agir d'un trou noir. La source Cygnus X-1, dans la constellation du Cygne, est le trou noir probable le plus connu.

Un trou noir déniché au cœur de notre Galaxie

Les astronomes ont aussi découvert, à l'intérieur de galaxies actives (en évolution rapide) et de quasars (les noyaux très lumineux de galaxies lointaines), des concentrations surprenantes de matière et d'énergie, qui évoquent la présence de trous noirs gigantesques. Ainsi, au centre de la galaxie Messier 87, une masse de gaz tourne à une vitesse d'environ 750 kilomètres par seconde ; elle serait piégée par un trou noir qui contiendrait près d'un milliard de fois la masse solaire ! Un trou noir a même pu être déniché au cœur de notre propre Galaxie, la Voie lactée. Des mesures des longueurs d'onde infrarouges, qui traversent les écrans de poussière interstellaire masquant le centre galactique, ont montré que les mouvements des étoiles d'une région de 0,3 année-lumière de rayon résultent de la présence d'une masse d'environ trois millions de masses solaires. Cette masse ne saurait être qu'un amas d'étoiles ni une unique étoile immense, car tous les modèles le prouvent, de tels objets se seraient transformés en trou noir en moins d'un million d'années.

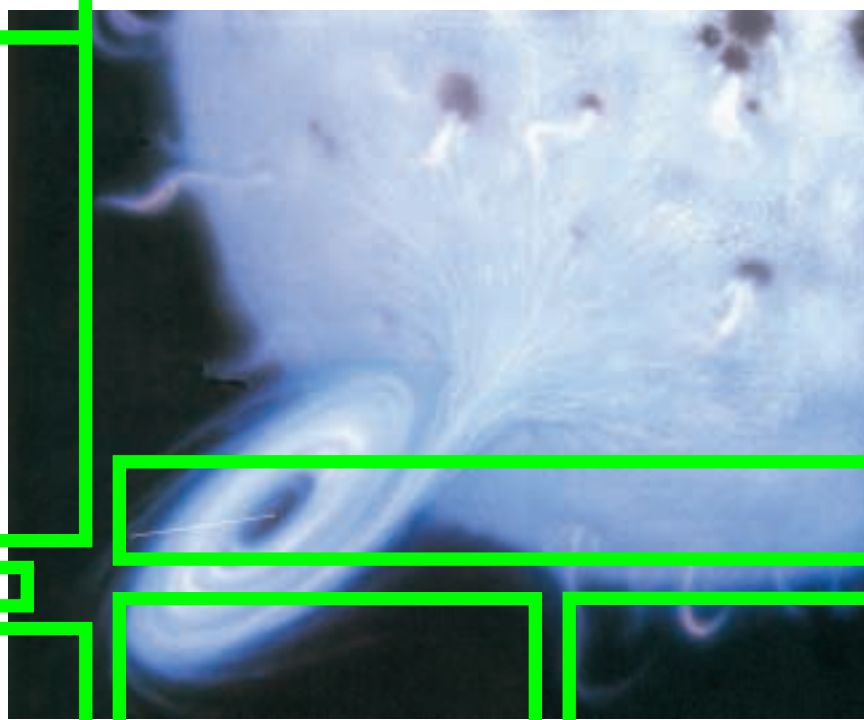
Quand une étoile s'effondre pour donner un trou noir, c'est-à-dire une

singularité, ni la physique classique ni la relativité générale ne suffisent à la décrire. Il faudrait faire appel à une théorie plus générale englobant la physique quantique et la relativité générale. Cet édifice théorique unifié porte déjà un nom, la «gravitation quantique». Toutefois, malgré l'activité intense des théoriciens, elle reste à construire.

Pour certains théoriciens, la gravitation quantique naîtra de la «théorie des cordes». Les briques élémentaires de cette théorie sont des «cordes» ou des membranes aux dimensions subatomiques dont les vibrations auraient engendré le monde à partir du vide quantique. Les vibrations de ces cordes ne se conçoivent que dans des espaces à plus de dix dimensions, dont six se trouveraient enroulées, enroulées dans d'infimes domaines. Si une corde s'approchait d'une singularité – lors d'un effondrement gravitationnel, par exemple – les dimensions enroulées réapparaîtraient, et se déploieraient à l'endroit de la singularité.

Trous noirs, pas si noirs

Plutôt que d'attendre l'avènement d'une théorie unifiée, certains physiciens ont tenu une démarche plus modeste. Ils cherchent les relations entre gravitation et comportement quantique, en adoptant un point de vue pratique : ils étudient les champs quantiques dans le cadre d'un espace-temps fixé *a priori*. En 1974, Stephen Hawking fit de cette façon une découverte théorique étonnante : les trous noirs ne seraient pas si noirs que cela ! Ils émettent un rayonnement, comme si une température bien définie régnait au rayon de Schwarzschild. S. Hawking étudiait le comportement du vide dans l'espace-temps d'un trou noir, c'est-à-dire les manifestations quantiques d'un état exempt de matière, quand il fit sa découverte. Dans ce vide naissent et disparaissent perpétuellement des paires de particules et d'antiparticules virtuelles. Près de l'horizon, il arrive parfois qu'une antiparticule ou qu'une particule tombe dans le trou noir, disparaissant ainsi sans retour. Puisqu'elle n'a plus d'existence réelle, elle s'annihile, la particule appariée acquiert alors une existence réelle. Il arrive qu'elle puisse s'éloigner du trou noir. Pour un observateur extérieur, le flux des particules qui s'échappent ainsi équivaut à un rayonnement thermique.



6. À LA FIN DE LEUR EXISTENCE, certaines étoiles peuvent donner naissance à des trous noirs. Si ces objets compacts font partie d'un système double, leur attraction gravitationnelle immense leur permet d'arracher du gaz à la surface de leur compagnon. Le gaz s'accumule en tournant dans un disque d'accrétion avant de disparaître finalement dans le trou noir.

Selon S. Hawking, le rayonnement extirpe énormément de l'énergie, donc de la masse, du trou noir, lequel s'amenuise. Toutefois, le temps nécessaire pour annihiler ainsi un trou noir de la masse du Soleil dépasserait l'âge de l'Univers de plusieurs ordres de grandeur. Qu'en resterait-il au final ? Du rayonnement ou une sorte de cicatrice dans la structure de l'espace-temps ? Ces questions restent ouvertes, car nous ignorons comment l'énergie émise réagit sur le trou noir.

Notre conception moderne du cosmos repose fondamentalement sur deux observations importantes : l'astronome américain Edwin Hubble découvrit à la fin des années 1920, que la distance entre la Voie lactée et pratiquement toutes les galaxies augmente et ce d'autant plus vite que ces galaxies sont éloignées. Ses deux compatriotes Arno Penzias et Robert Wilson détectèrent en 1964 un rayonnement cosmique dans le domaine spectral des micro-ondes, qui correspond à une température de 2,7 kelvins. Un observateur situé n'importe où dans l'Univers observerait que les galaxies s'éloignent les unes des autres. Il ne s'agit pas là de simples mouvements d'éloignement dans un espace statique, mais plutôt d'une expansion de l'espace-temps, lequel entraîne la

matière avec lui. Si l'on remonte le temps par le calcul pour retracer l'expansion de l'Univers, on constate qu'il y a 15 milliards d'années toute la matière était extrêmement concentrée et dense. L'Univers a traversé dans ses tout premiers instants une phase durant laquelle toute la matière qui nous entoure aujourd'hui sous forme d'étoiles et de galaxies était dissoute dans un mélange chaud, dense et sous forme de rayonnement et de matière. Le rayonnement cosmique omniprésent existe depuis le moment, où lumière et matière se sont découplées, quelque 380 000 ans après le Big Bang. Au fur et à mesure de l'expansion ultérieure, il s'est refroidi pour atteindre, aujourd'hui, la valeur de 2,7 kelvins.

Modèles de Big Bang

L'Univers en expansion est bien décrit par des modèles simples : des solutions particulières des équations de la gravitation d'Einstein. Il existe, en fait, toute une série de modèles compatibles avec les données astronomiques observées. Les astrophysiciens résument cette famille de théories par l'expression «modèle de Big Bang standard». Ce modèle standard est caractérisé par plusieurs propriétés qui sont communes à toutes les théories déjà évoquées. Après une phase d'expansion violente,

comparable à une explosion, l'Univers s'est développé jusqu'à son état actuel en passant par un stade précoce, dense et chaud. Les galaxies et les amas de galaxies sont nés de fluctuations initialement infimes de la densité de la matière, sous l'action de la gravitation. Les germes de ces fluctuations ont laissé leur empreinte dans les irrégularités du bruit de fond cosmologique diffus.

Des mesures récentes sur les supernovae des galaxies distantes nous ont fourni une information supplémentaire : l'expansion de l'Univers s'accélère. Le phénomène résulterait d'une répulsion cosmique due à une mystérieuse densité d'énergie de l'espace vide. Cette dernière intervient sous la forme de la « constante cosmologique », laquelle joue un rôle dans les équations de la relativité générale. Cette densité d'énergie confère son caractère irréversible à l'expansion de l'Univers, laquelle se poursuivra à jamais.

Tandis que nos concepts et nos lois physiques restent impuissants à

décrire le Big Bang lui-même, ils deviennent applicables dès la première seconde de l'Univers, quand le modèle standard décrit l'évolution de l'Univers.

La hiérarchie des interactions

Que s'est-il passé avant que le modèle standard s'applique? À des températures et à des densités aussi élevées, les différences entre les forces fondamentales commencent à s'évanouir. Ces forces se fondent alors en une seule et même force originelle. Ainsi, au cœur des stades primordiaux du développement du cosmos, il y a eu une transition brusque entre une « époque de haute symétrie » et les ères ultérieures de moindre symétrie. Tandis que, pendant la première époque, tout était gouverné par une unique force originelle, les ères suivantes furent marquées par la naissance de la hiérarchie des interactions gravitationnelles et électromagnétiques

(faible et forte). L'énergie qui fut disponible lors de cette transition détermine l'expansion de l'Univers aujourd'hui. Pendant de brefs instants, cette transition eut la violence d'une explosion caractérisée par une accélération exponentielle. Puis l'énergie du vide fut transférée au rayonnement, et l'évolution prévue dans le cadre du modèle standard prit la relève.

Le modèle de l'origine de l'Univers que nous venons d'ébaucher se nomme « modèle d'inflation ». Il justifie plusieurs aspects du modèle standard que, sinon, il aurait fallu accepter sans démonstration. Le premier est l'âge de l'Univers : 15 milliards d'années. Autre aspect : le modèle d'inflation implique une évolution vers un espace physique plan à trois dimensions en expansion perpétuelle, mais ralentie ; pour les cosmologistes, un tel univers est proche de celui que nous observons. L'inflation aurait donc gonflé l'Univers, ter un ballon, si rapidement qu'il serait devenu quasi plat. La partie du cosmos acces-

L'infiniment grand : une réflexion critique

Le modèle cosmologique standard, modèle d'évolution de l'Univers le plus communément diffusé auprès du public par la communauté scientifique, et décrit par Gerhard Hörter, a plusieurs lacunes. Il est tout d'abord bâti sur une solution particulière des équations de la relativité générale, que l'on obtient en supposant que la matière est uniformément répartie dans l'Univers. Or, cette hypothèse, outre le fait qu'elle présente des faiblesses théoriques, n'a jamais été vérifiée par l'observation. De plus, les singularités qu'elle prévoit – Big Bang, éventuellement Big Crunch – procèdent d'extensions abusives de la relativité générale à des échelles auxquelles les physiciens savent que cette théorie est difficilement applicable. Que pouvons nous, toutefois, retenir de ce modèle standard?

L'expansion de l'Univers est l'explication la plus convaincante du décalage vers le rouge du spectre électromagnétique des objets cosmiques lointains. De plus, cette hypothèse est compatible avec l'existence du rayonnement cosmologique, détecté pour la première fois en 1964, et mesuré depuis par plus d'une trentaine d'expériences.

L'hypothèse que l'Univers soit actuellement dans une phase d'expansion accélérée est une interprétation de données récentes (1998-1999) publiées à partir de l'observation de moins d'une centaine de supernovae éloignées. Elle est à envisager avec prudence car d'autres interprétations sont tout aussi plausibles.

L'inflation, si souvent invoquée depuis 20 ans, tente de résoudre certains problèmes du modèle standard, à l'aide d'un scénario décrivant les « premiers instants » de

l'Univers. Cette hypothèse engendre toutefois des problèmes nouveaux, qui ne sont toujours pas résolus. Pour tenter de pallier ces inconvénients, certains cosmologistes, sous l'impulsion du Russe Andreï Linde, ont multiplié les modèles d'inflation, de manière... inflationniste, compliquant encore la situation.

Quant au « pré-Big Bang » si Big Bang il y a eu, aucun moyen n'existe actuellement pour l'appréhender.

Quid de enfin des trous noirs? Les plus « classiques », c'est-à-dire ceux qui seraient issus de l'effondrement gravitationnel d'étoiles très massives, se formeraient au sein de systèmes, d'une taille comparable à celle du Système solaire, où l'on sait que la relativité générale s'applique. Toute la dynamique de leur formation est parfaitement décrite par cette théorie et par ce que les astrophysiciens savent de l'évolution stellaire. De plus, les observations dans le domaine des rayons X de candidats trous noirs indiquent un comportement qui semble plutôt conforme aux prédictions théoriques. L'hypothèse que de très grosses étoiles se transforment en fin de vie en trous noirs est donc un scénario plausible. Ce dernier n'évite pas les singularités – points de l'espace-temps où certaines grandeurs, *a priori* observables, deviennent infinies. Or, une théorie qui produit des infinis a toujours été considérée comme suspecte. La physique fondamentale en science est en marche et les singularités seront peut-être maîtrisées un jour.

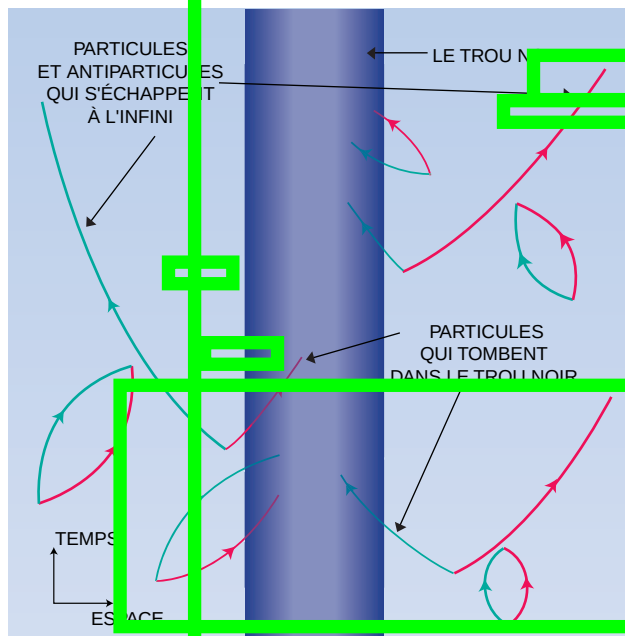
Marie-Noëlle CÉLÉRIER
Département d'astrophysique relativiste
de l'Observatoire de Paris-Meudon

sible aujourd'hui à nos observations se soit née d'un germe minuscule, dans lequel les liens entre événements étaient déterminés par les règles de la relativité (causalité).

Quelle taille devrait avoir eu ce domaine originel pour avoir englobé l'ensemble de l'Univers que l'on observe aujourd'hui? Grâce au modèle standard, on déduit de l'état actuel du Cosmos – c'est-à-dire de sa température (2,7 kelvins) et sa taille (égale à la distance des objets les plus éloignés connus) – ses caractéristiques à la fin de l'ère d'inflation : il mesurait alors quelque dix centimètres. Comme l'inflation a multiplié toutes les longueurs par un facteur 10^{29} , le germe de l'Univers actuel était confiné dans un domaine de l'ordre de 10^{-28} centimètre. Le physicien russe Andreï Linde a proposé des généralisations de cette image. Selon lui, le «Grand Univers» se composerait de bulles d'univers individuelles séparées les unes des autres, l'une se formant ici, l'autre s'évanouissant ailleurs, selon des évolutions simultanées et indépendantes. Nous nous trouverions dans l'une de ces bulles, qui aurait la particularité de nous avoir fourni des conditions de vie acceptables. Curieusement, le principe anthropique suivant lequel l'Univers a été créé pour que l'homme puisse exister trouverait dans la théorie de Linde une sorte d'explication : parmi la multitude des bulles cosmiques, il s'en serait trouvé une par hasard pour réunir les conditions nécessaires à la vie.

Le début de l'Univers

Qui ne se satisfait pas de ce genre d'explication de l'origine des choses doit se pencher plus avant sur la question des conditions initiales de l'Univers et s'interroger : «Qu'y avait-il avant le Big Bang?» Les physiciens ne savent comment répondre à cette question métaphysique. La gravitation quantique le permet-elle? C'est à une théorie définitive de toutes les interactions physiques qu'il faudrait se référer pour analyser la singularité initiale. Y a-t-il eu au début de l'Univers une «corde» à 10 ou 11 dimensions? Dans ses *Confessions* (livre 11) : saint Augus-



7. LE VIDE QUANTIQUE est en permanence le siège de la création de paires composées d'une particule (en vert) et d'une antiparticule (en rouge). Ces paires sont dites virtuelles parce qu'elles ne sont pas observables. À peine sont-elles créées qu'elles s'annihilent déjà. Lorsque, dans le cas à proximité d'un trou noir, il arrive que l'un de ses membres (particule ou antiparticule) y disparaisse. Ainsi libéré de sa compagne virtuelle, l'autre devient une particule réelle et échappe (parfois) à l'attraction du trou noir. Pour l'observateur extérieur, tout semble alors comme si le trou noir émettait un rayonnement thermique, le rayonnement de Hawking.

tin répondait à la question de la naissance de l'Univers par une pirouette : «À la question : "Que faisait Dieu, avant qu'il ne crée le Monde?", certains seraient tentés de répondre : "Il préparait l'Enfer, pour les personnes qui poseraient cette question."»

Même si nous ne disposons pas encore de la théorie de la gravitation quantique, nous pouvons spéculer sur la nature de l'Univers quantique. S. Hawking s'est beaucoup intéressé à cette question. Il propose de se limiter, pour le cosmos quantique primitif, à des espaces-temps d'une structure très simple, lisse, de courbure régulière. Le temps n'existe pas encore dans ce cosmos quantique constitué d'une série d'espaces à quatre dimensions, c'est-à-dire des surfaces d'espaces sphériques à cinq dimensions.

Notre Univers, caractérisé par une évolution temporelle, aurait émergé à partir d'une partie finie de cet univers quantique intemporel. La question du passé de ce «petit pois originel» n'a pas de sens dans le cadre normal de l'espace et du temps ; tout comme cela n'aurait pas de sens de demander la latitude et la longitude d'un point qui ne serait pas à la surface de la Terre.

Penrose lui aussi est convaincu que

l'Univers primordial était plat et uniforme. Il s'en est persuadé en quantifiant le désordre de l'Univers, c'est-à-dire son entropie. Il prend en compte non seulement le désordre associé au rayonnement et à la matière, mais aussi les sources d'entropie plus rares, dissimulées au sein du champ de gravitation, dans les vides et dans les courbures de l'espace-temps, ainsi qu'il en est dans les trous noirs. Les valeurs qu'il obtient sont gigantesques : les conditions initiales nécessaires à la naissance de notre Univers n'avaient qu'une chance parmi quelque 10 milliards à la puissance 120 de se produire. Une probabilité aussi faible a-t-elle un sens ?

Toutes ces réflexions sur les premières fractions de seconde de l'Univers relèvent du domaine de la spéculation. Les physiciens sont encore loin d'expliquer le Big Bang. Spéculer engendre un tel plaisir chez les cosmologistes pen-

chés sur ces questions qu'ils succombent parfois à la tentation de prendre leurs constructions théoriques pour la réalité. Pour reprendre les mots d'Einstein : «Ceux qui élaborent des théories dans ce domaine trouvent les fruits de leur imagination si naturels qu'ils ne les voient plus comme des édifices intellectuels, mais comme des réalités, qu'ils aimeraient voir reconnues.»

Gerhard BERNER est théoricien à l'Institut Max Planck d'astrophysique à Garching, près de Munich.

Jonathan MALLINELL, *La cosmologie quantique et la création de l'Univers*, Pour la Science, février 1992.

Stephen HAWKING, *Black Holes and Baby Universes and other Essays*, Bantam Books 1993.

Wendy FRIEDMAN, *L'expansion de l'Univers*, Pour la Science, janvier 1993.

Andreï LINDE, *L'univers inflatoire et autoreproducteur*, Pour la Science, janvier 1995.

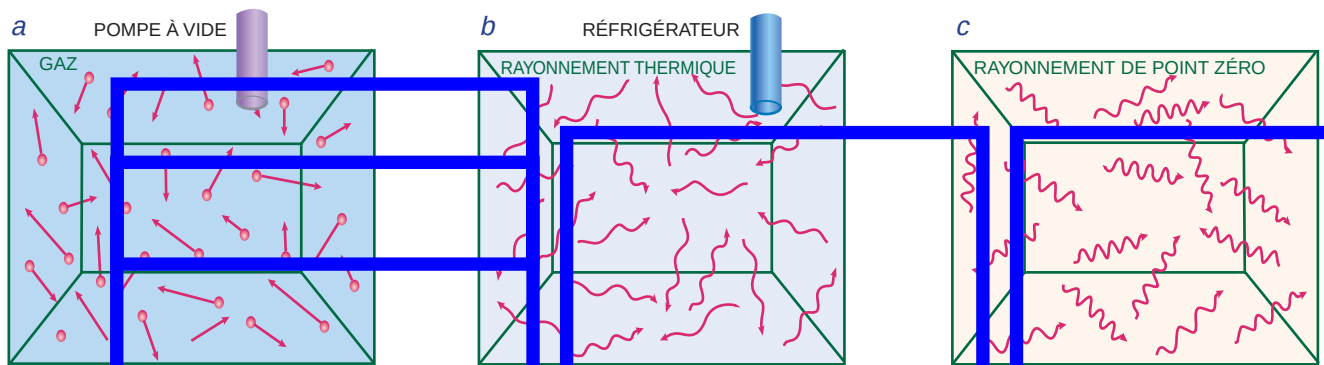
Marc LACHIÈZE-RAY et Jean-Pierre LUMINET, *La topologie de l'Univers*, Pour la Science, décembre 1996.

La révolution cosmologique, Pour la Science, mars 1999.

L'infiniment vide n'existe pas

TIMOTHY BOYER

Le vide classique n'est pas vide. Même quand on a supprimé toute la matière et le rayonnement thermique d'une partie de l'espace, le vide de la physique contient encore des champs électromagnétiques, des particules fugaces, et possède une certaine énergie.



Aristote et ses élèves croyaient qu'aucune région de l'espace ne pourrait être entièrement vide : le vide (le «rien»), leur semblait aussi hors d'atteinte que l'infini (le tout). Les physiciens du XVII^e siècle s'inscrivaient en faux contre l'affirmation aristotelicienne que «la nature a horreur du vide» et Torricelli démontra que le vide existait. Or, selon la physique moderne, l'idée d'Aristote est ressuscitée : si l'on peut, au moins en principe, retirer toute la matière ordinaire d'une partie de l'espace, le vide ainsi obtenu n'est pas synonyme de «rien».

La terminologie est alors ambiguë : si le vide n'est pas vide, mérite-t-il son nom ? Les physiciens modernes définissent le vide comme ce qui reste quand on a éliminé tout ce qu'il est expérimentalement possible d'enlever d'une partie de l'espace : ainsi le vide est défini par une limite expérimentale.

Pour faire le vide, on prélève d'une enceinte toute matière visible, solide, liquide ou gazeuse (voir la figure ci-dessus). Quand on a prélevé toute la matière, l'espace n'est toujours pas vide : il contient encore un rayonnement électromagnétique. Une partie de ce rayonnement est d'origine thermique, issue du rayonnement des corps matériels et peut donc être supprimée par refroidissement de l'enceinte contenant le vide, mais l'autre partie du rayonnement est d'origine différente. Quand bien même on réduirait (par la persée) la température du vide jusqu'au zéro absolu, il subsisterait un ensemble fluctuant d'ondes électromagnétiques. Ce rayonnement résiduel, le rayonnement de point zéro, qu'il est impossible de supprimer, est une caractéristique du vide.

En 1644, l'ancien secrétaire de Galilée, Évangélisme Toricelli, inventa le baromètre. Cette invention allait démontrer

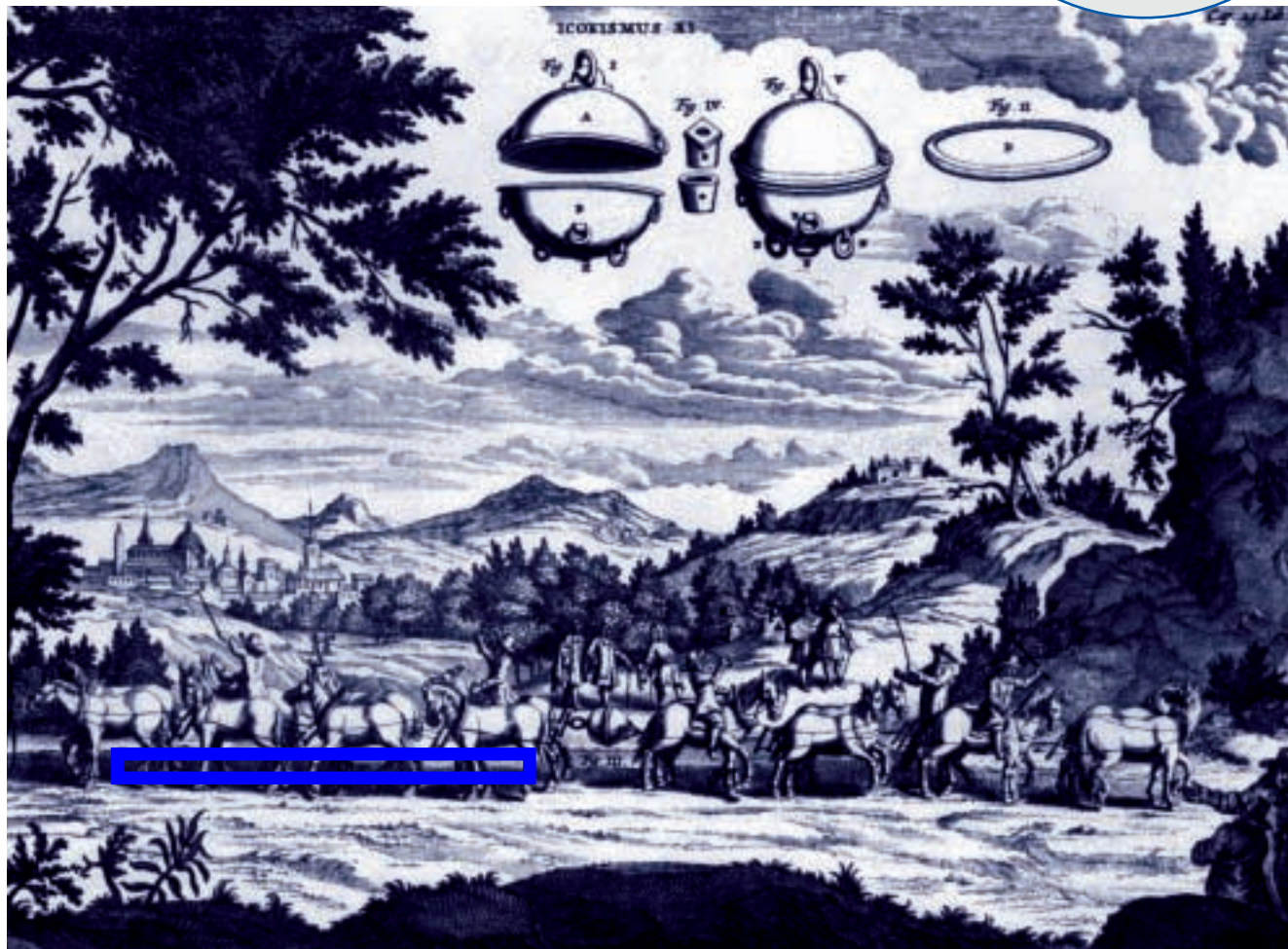
l'erreur de la doctrine d'Aristote, selon laquelle le vide était physiquement impossible. Toricelli avait placé du mercure dans un tube en verre fermé à une extrémité, puis il avait retourné le tube sur une cuve remplie de mercure. Le mercure était alors descendu dans le tube jusqu'à ce que sa surface libre se trouve à une hauteur de 76 centimètres au-dessus du niveau du mercure dans la cuve. L'espace sans mercure s'était constitué dans la partie supérieure du tube et cet espace était manifestement dépourvu de toute matière visible. Pour Toricelli, il ne contenait pas non plus de gaz et était donc vide. Une vive polémique s'engagea entre les partisans d'Aristote et ceux de Toricelli. Quelques années plus tard, Blaise Pascal imagina une série d'expériences qui confirmèrent toutes l'hypothèse de Toricelli.

Durant les décennies suivantes, de nombreux physiciens s'intéressèrent aux expériences sur le vide. La plus connue,

celle d'un gaz : la pression et la température du rayonnement augmentent du fait de sa compression. Ce rayonnement comprimé exerce alors une force qui s'oppose au retour du piston et celui-ci ne revient au fond du cylindre que

Dès le XIX^e siècle, les physiciens expliquèrent ce phénomène ; pendant que l'on maintient le piston écarté, les parois du cylindre émettent un rayonnement thermique dans le vide créé ; quand on repousse le piston, le rayonnement est comprimé. Or la compression d'un rayonnement thermique est analogue à

1. L'EXPÉRIENCE DES SPHÈRES DE MADGEBOURG, réalisée en 1654 par Otto von Guericke, mit en évidence «le vide». Après avoir appliqué les hémisphères l'une contre l'autre, von Guericke pompa l'air à l'intérieur. Deux attelages de huit chevaux de trait ne purent les séparer.



quand le rayonnement à haute température a été réabsorbé par les parois.

Le rayonnement thermique est constitué de champs électromagnétiques qui fluctuent de la façon la plus aléatoire possible. Paradoxalement, ce «hasard maximal» confère au rayonnement thermique une grande régularité statistique. Dans ces conditions d'équilibre thermique, c'est-à-dire quand la température est partout égale, le rayonnement est à la fois homogène et isotrope : ses propriétés sont identiques en tous les points de l'espace et dans n'importe quelle direction ; un instrument mesurant une propriété quelconque du rayonnement indiquerait le même résultat, quelles que soient sa position et son orientation.

Le spectre du rayonnement thermique

Un rayonnement thermique est caractérisé par sa température. En 1879, le physicien autrichien Josef Stefan étudia la densité d'énergie totale (c'est-à-

dire l'énergie par unité de volume) du rayonnement thermique. Se fondant sur certaines expériences préliminaires, il suggéra que cette densité d'énergie est proportionnelle à la puissance quatrième de la température absolue. Cinq ans plus tard, Ludwig Boltzmann, un élève de Stefan, démontra ce résultat par une analyse théorique subtile.

La température d'un rayonnement thermique caractérise non seulement sa densité d'énergie, mais aussi son spectre, c'est-à-dire la répartition de l'énergie du rayonnement en fonction de sa fréquence. La relation entre la température et le spectre thermique est une évidence quotidienne : quand on chauffe un objet métallique, par exemple, il commence par rougeoier, puis devient blanc ou même bleu, à mesure que le spectre contient une proportion croissante de rayonnements de fréquences supérieures. Le spectre thermique n'est pas monochromatique ; un ~~corps~~ métallique chauffée au rouge émet surtout du rouge, mais aussi, plus, faiblement, à toutes les autres fréquences.

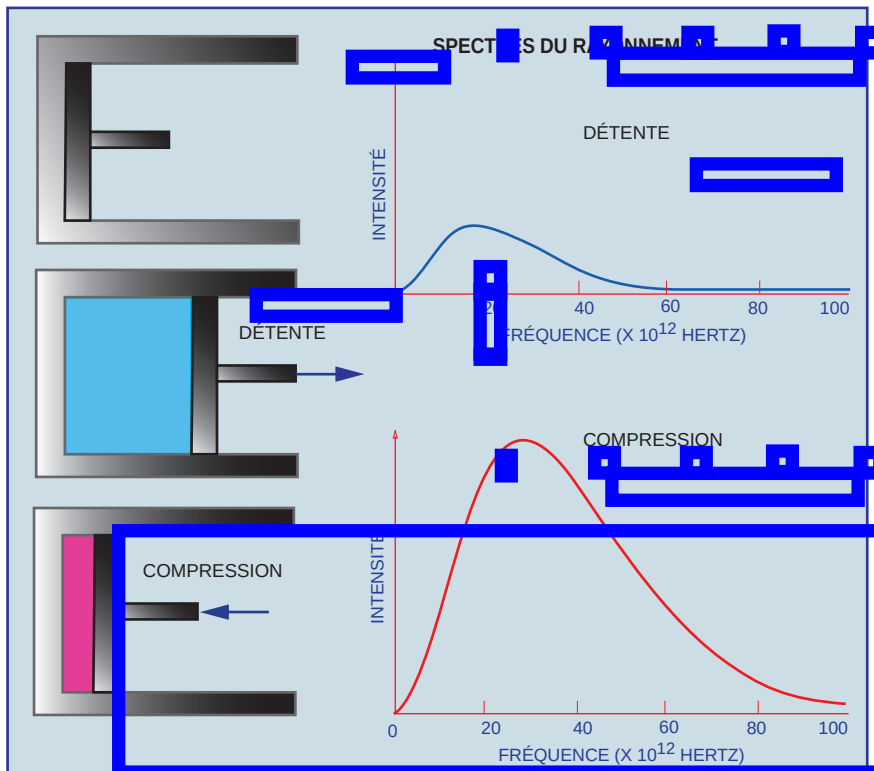
Vers la fin du XIX^e siècle, les physiciens déterminèrent expérimentalement la forme du spectre thermique et sa variation en fonction de la température, mais l'analyse théorique resta embryonnaire ; à l'époque on cherchait seulement une expression mathématique reliant l'intensité du rayonnement à la fréquence et à la température. En d'autres termes, pour une température donnée, on cherchait à prévoir l'intensité des divers rayonnements.

En 1896, le physicien allemand Wilhelm Wien analysa avec plus de pertinence le spectre thermique. Wien se fonda sur une expérience de pensée semblable à celle que nous avons décrite au début de cet article ; il faisait cependant une hypothèse supplémentaire : le cylindre devait être parfaitement isolé, afin que la quantité de chaleur emprisonnée soit constante. Wien calcula alors la modification du spectre due à une variation infinitésimale du volume interne du cylindre. Il en déduisit que l'expression mathématique décrivant le spectre devait comporter deux facteurs, dont le produit était égal à l'intensité du rayonnement, à une fréquence et à une température données. L'un des facteurs était le cube de la fréquence, l'autre une fonction du rapport entre la température absolue et la fréquence. Wien ne put toutefois déterminer la forme exacte de cette fonction ; il en proposa une... qui était fautive.

La théorie classique de l'électron

C'est Max Planck qui, en 1900, détermina la fonction mathématique nécessaire à la description du spectre. Planck découvrit que, pour décrire le rayonnement thermique, il fallait introduire une nouvelle constante fondamentale, égale à $6,626 \times 10^{-34}$ joule-seconde ; on désigne aujourd'hui h cette «constante de Planck» ; en cherchant à interpréter sa fonction, Planck élaborait la théorie des quanta.

Selon la formule de Stefan et Boltzmann la densité d'énergie du rayonnement thermique est proportionnelle à la puissance quatrième de la température, l'énergie doit être nulle au zéro absolu : à cette température, le rayonnement thermique disparaît complètement. Ainsi les résultats du XIX^e siècle, selon lesquels on pouvait éliminer tout le rayonnement thermique, rejoignaient une conception classique du vide, celle du XVII^e siècle : le vide parfait était



2. LE SYSTÈME IDÉAL constitué d'un piston glissant sans frottement dans un cylindre permet de constater l'existence d'un rayonnement thermique dans le vide. Initialement, le piston est au fond du cylindre, de sorte qu'aucun espace n'est compris entre le piston et le cylindre. On tire le piston et on le maintient tiré quelques temps dans cette position, à la température de la pièce. L'espace compris entre le piston et le fond du cylindre semble vide, mais, quand on relâche le piston, il ne revient pas à sa position initiale, même si on le pousse contre le fond du cylindre. Pendant que l'on maintient le piston écarté de sa position initiale, les parois du cylindre émettent un rayonnement qui remplit le «vide». Le retour du piston comprime le rayonnement dont la température augmente, ce qui modifie le spectre : le rayonnement, plus chaud, s'oppose à la compression.

toujours une région totalement vide de l'espace, mais pour atteindre cet état, il fallait non seulement retirer toute la matière visible et tous les gaz, mais aussi éliminer tout le rayonnement électromagnétique. On pouvait, théoriquement au moins, satisfaire cette dernière condition en refroidissant l'espace jusqu'à la température du zéro absolu.

Cette conception du vide classique était incluse dans la théorie physique fondamentale de l'époque, appelée aujourd'hui théorie classique de l'électron : on considérait les électrons comme des particules ponctuelles, dont les seules propriétés étaient la masse et la charge électrique ; on pouvait mettre les électrons en mouvement à l'aide de champs électriques et magnétiques et, inversement, leur mouvement produisait de tels champs (par exemple, un électron qui oscille régulièrement émet des ondes électromagnétiques à la fréquence de l'oscillation). En théorie classique de l'électron, on décrit les interactions entre les particules et les champs par les lois du mouvement établies par Newton et par les équations de l'électromagnétisme dues à James Clerk Maxwell. Pour obtenir des prévisions théoriques précises, il faut de plus spécifier certaines « conditions aux limites ». Les équations de Maxwell indiquent les variations d'un champ électroma-

gnétique dans l'espace ou dans le temps, mais pour calculer la valeur réelle du champ, il faut en connaître les valeurs initiales (appelées aussi valeurs limites), qui fournissent une référence par rapport à laquelle on calcule les modifications ultérieures.

C'est le choix des conditions aux limites qui, en théorie classique de l'électron, détermine la nature du vide : au XIX^e siècle, on pensait que le vide ne contenait aucun rayonnement et l'on supposait par conséquent que les champs électrique et magnétique y étaient initialement nuls. Autrement dit, selon la théorie classique de l'électron, au XIX^e siècle, on supposait que, dans un passé lointain, l'Univers contenait de la matière (les électrons), mais aucun rayonnement ; les rayonnements électromagnétiques résulteraient de l'accélération des charges électriques.

L'effet Casimir

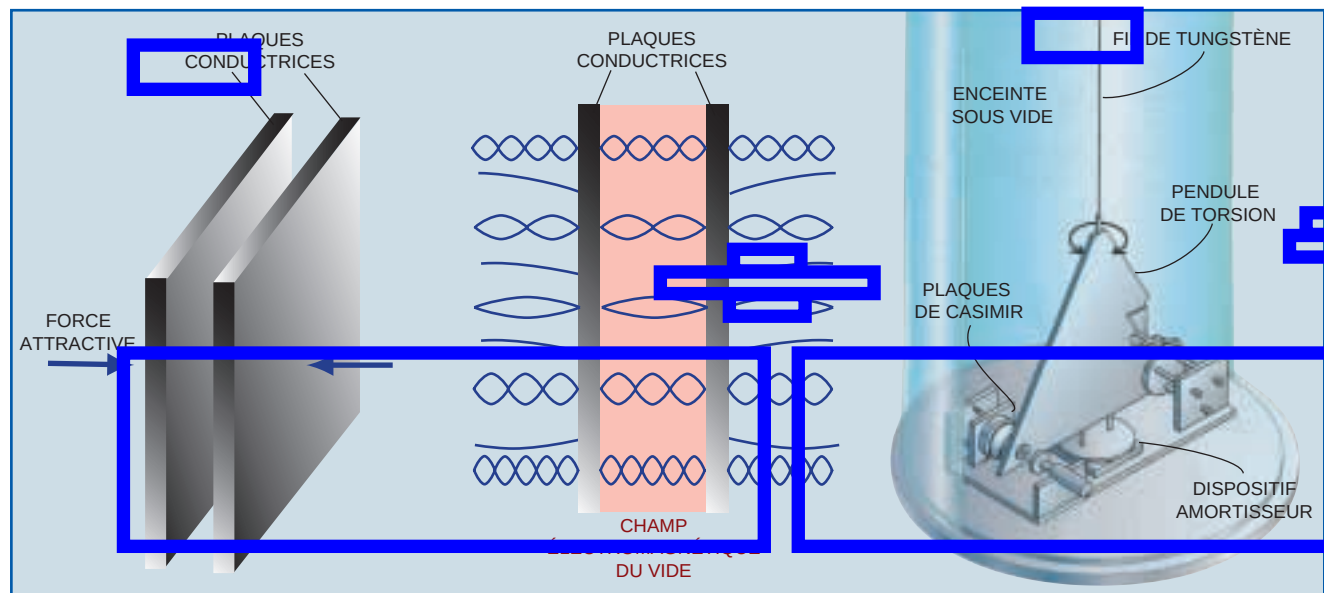
La théorie classique de l'électron reste aujourd'hui un édifice théorique intéressant, mais elle a pris une nouvelle tournure au XX^e siècle. En 1948, Hendrick Casimir, des Laboratoires de recherche Philips, en Hollande, imagina une expérience qui mettait en lumière la nécessité de revoir la théorie classique de l'électron : Casimir analysa les forces qui agissent sur deux plaques conduc-

trices parallèles et proches l'une de l'autre, dans le vide. Si les plaques sont chargées électriquement, elles doivent subir une attraction ou une répulsion selon les signes respectifs des charges qu'elles portent. Casimir considéra le cas où les plaques ne sont pas chargées : même alors, les plaques subissent une force en raison du rayonnement électromagnétique qui les entoure. Une analogie mécanique permet de comprendre l'origine de cette force.

Imaginons une corde lisse et souple que l'on fait passer par des trous percés au centre de deux blocs de bois. Si la corde n'est pas liée aux blocs, elle ne tend, au repos, ni à les séparer ni à les rapprocher. En revanche, si on fait vibrer transversalement la corde entre les blocs, ceux-ci sont soumis à des forces qui les éloignent l'un de l'autre le long de la corde. Cette force résulte du fait que la corde ne peut pas vibrer à l'endroit où elle traverse les blocs ; les ondes sont réfléchies à cet endroit. Lorsqu'une onde est réfléchie, une partie de sa quantité de mouvement est transmise au réflecteur, c'est-à-dire au bloc.

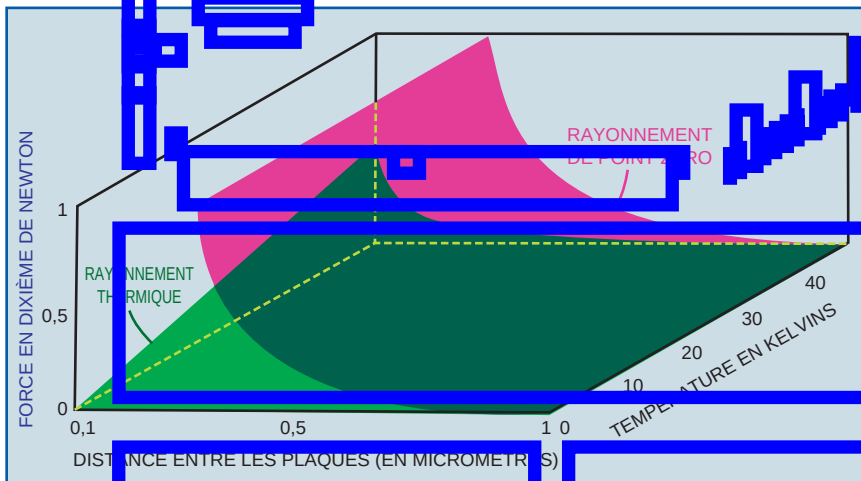


Hendrick Casimir



3. L'EFFET Casimir dénote l'existence de champs électromagnétiques dans le vide. Deux plaques métalliques placées dans le vide sont disposées parallèlement et à faible distance. Les plaques métalliques sont conductrices et elles réfléchissent les ondes électromagnétiques ; or la réflexion d'une onde s'accompagne de la présence d'un nœud – c'est-à-dire un point d'amplitude nulle – de champ électrique à la surface de la plaque. Le système d'ondes qui s'établit entre les plaques produit une force d'attraction car seules les ondes dont les longueurs

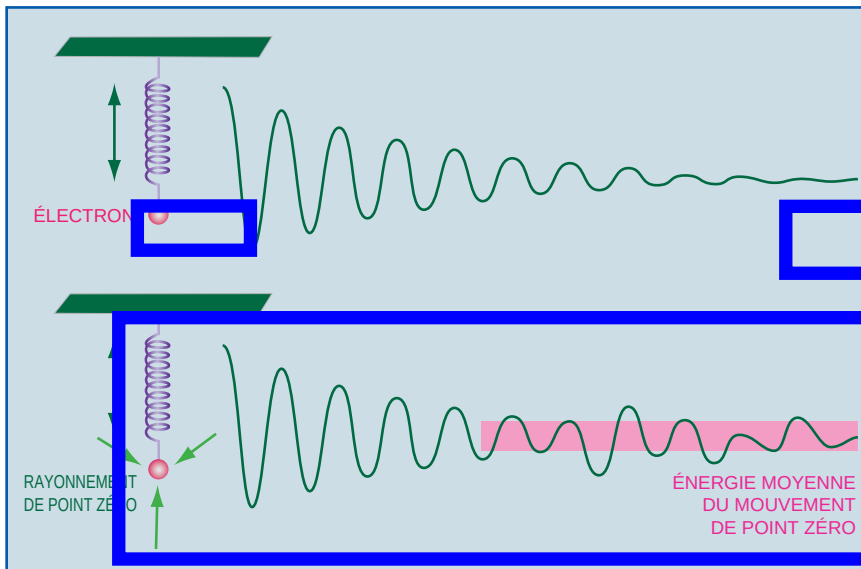
d'onde sont des sous-multiples de la distance entre les deux plaques existent dans l'espace entre les plaques. Les ondes dont les longueurs d'onde sont autres existent en dehors de l'espace entre les plaques et poussent celles-ci l'une contre l'autre, d'où la force attractive. À droite, le schéma de l'appareil utilisé par Steve Lamoreaux (Laboratoire de Los Alamos) pour mesurer l'effet Casimir. L'attraction des plaques tend à faire tourner le pendule de torsion, et l'on mesure la force à exercer pour ramener le pendule à sa position initiale.



4. LA FORCE D'ATTRACTION ENTRE LES DEUX PLAQUES. Dans l'expérience de Casimir, a deux composantes. À haute température, le rayonnement thermique (en vert) engendre une force directement proportionnelle à la température et inversement proportionnelle au cube de la distance entre les plaques. Cette force s'annule au zéro absolu, comme le rayonnement thermique lui-même. La force due au rayonnement de point zéro (en rose) est indépendante de la température et inversement proportionnelle à la puissance quatrième de la distance entre les plaques. Les forces correspondent ici à des plaques de un centimètre carré de surface.

Dans l'expérience de Casimir, la situation physique est analogue : les plaques métalliques correspondent aux blocs de bois et les champs électromagnétiques à la corde vibrante. Enfin, les trous par lesquels la corde traverse les blocs de bois représentent la conductivité des plaques métalliques. Les blocs réfléchissent les ondes électromagnétiques : comme

un rayonnement lèche les deux faces de chaque plaque les forces ont tendance à s'annuler. Mais, en fait, les forces ne sont pas exactement équilibrées : il reste une petite force résiduelle, proportionnelle à la surface des plaques et qui dépend aussi de leur écartement et du spectre du rayonnement électromagnétique aléatoire.



5. L'OSCILLATEUR HARMONIQUE est un système sur lequel on peut observer les effets du rayonnement de point zéro sur la matière. L'oscillateur est constitué d'un électron lié à un ressort idéal et sans frottement. Quand on écarte l'électron de sa position initiale, puis qu'on le relâche, il oscille autour de sa position d'équilibre et émet un rayonnement électromagnétique à la fréquence de l'oscillation. Ce rayonnement dissipe de l'énergie et, en l'absence du rayonnement de point zéro et à la température du zéro absolu, l'électron finit par s'immobiliser. En fait, le rayonnement de point zéro transmet sans arrêt des impulsions aléatoires à l'électron, de sorte qu'il ne s'arrête jamais totalement. Le rayonnement de point zéro confère à l'oscillateur une énergie moyenne, égale au produit de la fréquence de l'oscillation par la moitié de la constante de Planck.

Jusqu'ici, l'analyse est parfaitement conforme avec l'idée que les physiciens du XIX^e siècle se faisaient du vide. Dans ce cadre, la force agissant sur chaque plaque est due au rayonnement thermique aléatoire : quand on refroidit le système jusqu'à ce que sa température soit celle du zéro absolu, les forces doivent disparaître comme le rayonnement thermique. Or l'expérience contredit cette prédiction : en 1958, le physicien hollandais M. Sparnaay effectua une série d'expériences afin de tester les idées de Casimir et il observa que la force ne semblait pas s'annuler quand on réduisait fortement l'intensité du rayonnement thermique ; au contraire, on observait une force d'attraction résiduelle, qui doit persister même à la température du zéro absolu !

Cette force résiduelle est directement proportionnelle à la surface des plaques et inversement proportionnelle à la puissance quatrième de leur écartement ; la constante de proportionnalité est égale à $1,3 \cdot 10^{-27}$ newton-centimètre. La force résiduelle est faible, mais mesurable si les plaques sont suffisamment proches. Pour des plaques de un centimètre carré distantes de 0,5 micromètre (millième de millimètre), la force de Casimir est équivalente au poids d'une masse de 0,2 milligramme.

Quelle que soit l'importance exacte de l'effet Casimir, sa simple existence dénote une erreur fondamentale dans la théorie du XIX^e siècle sur le vide classique. Quand on compare la théorie et l'expérience, on constate que même à la température du zéro absolu, le vide classique n'est pas totalement vide ; il contient les champs électromagnétiques responsables de la force d'attraction mesurée par Sparnaay. On désigne aujourd'hui ces champs présents dans le vide sous le nom de rayonnement électromagnétique classique de point zéro.

Le spectre de point zéro

Quelles sont les caractéristiques de ce rayonnement de point zéro ? Le simple fait que ce rayonnement existe permet plusieurs constatations, car le rayonnement doit obéir aux postulats sur la nature du vide. Dans le vide, par exemple, aucun point n'est particulier, ni dans le temps, ni dans l'espace ; le vide doit paraître identique en tout lieu et dans toutes les directions. De ce fait, le rayonnement de point zéro, comme le rayonnement thermique, doit

être isotrope et homogène. D'autre part, aucune vitesse n'est particulière dans le vide : deux observateurs quelconques doivent en mesurer les mêmes propriétés physiques, quelle que soit leur vitesse l'un par rapport à l'autre, quand cette vitesse est constante. Autrement dit, le rayonnement de point zéro doit être invariant par la transformation de Lorentz (cette transformation, due au physicien hollandais H. Lorentz, permet de passer d'un référentiel à un autre, ces référentiels étant en mouvement uniforme les uns par rapport aux autres ; on tient compte du fait que la vitesse de la lumière est identique dans tous les référentiels ; la transformation de Lorentz joue un rôle important dans la théorie de la relativité restreinte, due à Albert Einstein).

L'invariance par la transformation de Lorentz est une condition très importante, comme nous allons le voir. Imaginons un voyageur qui hésite pour savoir si c'est son propre train qui se déplace, ou celui d'à côté : pour lever cette ambiguïté, il lui suffit de prendre pour repère un point qu'il sait fixe par rapport à la Terre. Or la condition d'invariance par la transformation de Lorentz stipule que de tels repères n'existent pas dans le vide et qu'aucune expérience ne permet à un observateur de connaître sa vitesse par rapport au rayonnement de point zéro. Pour vérifier cette condition, le spectre de rayonnement du vide doit, comme nous allons le voir, posséder des propriétés tout à fait particulières.

Supposons momentanément qu'un observateur observe le rayonnement de point zéro et qu'il constate que celui-ci n'est constitué que de lumière verte. Quelles que soient la position de l'observateur et la direction dans laquelle il effectue ses observations, le vide lui apparaît plein d'un rayonnement vert uniforme. Ce rayonnement est homogène et isotrope pour cet observateur particulier, mais l'est-il encore pour un autre observateur se dirigeant vers le premier à vitesse constante ? Non : en raison de l'effet Doppler, le second observateur en mouvement par rapport au premier observerait que le rayonnement est décalé en longueur d'onde, bleu devant lui et rouge derrière. D'autre part, en raison de la transformation de Lorentz, l'intensité du rayonnement est supérieure devant et inférieure derrière. Le rayonnement est donc différent selon l'observateur : il est isotrope pour l'un, mais pas pour l'autre.

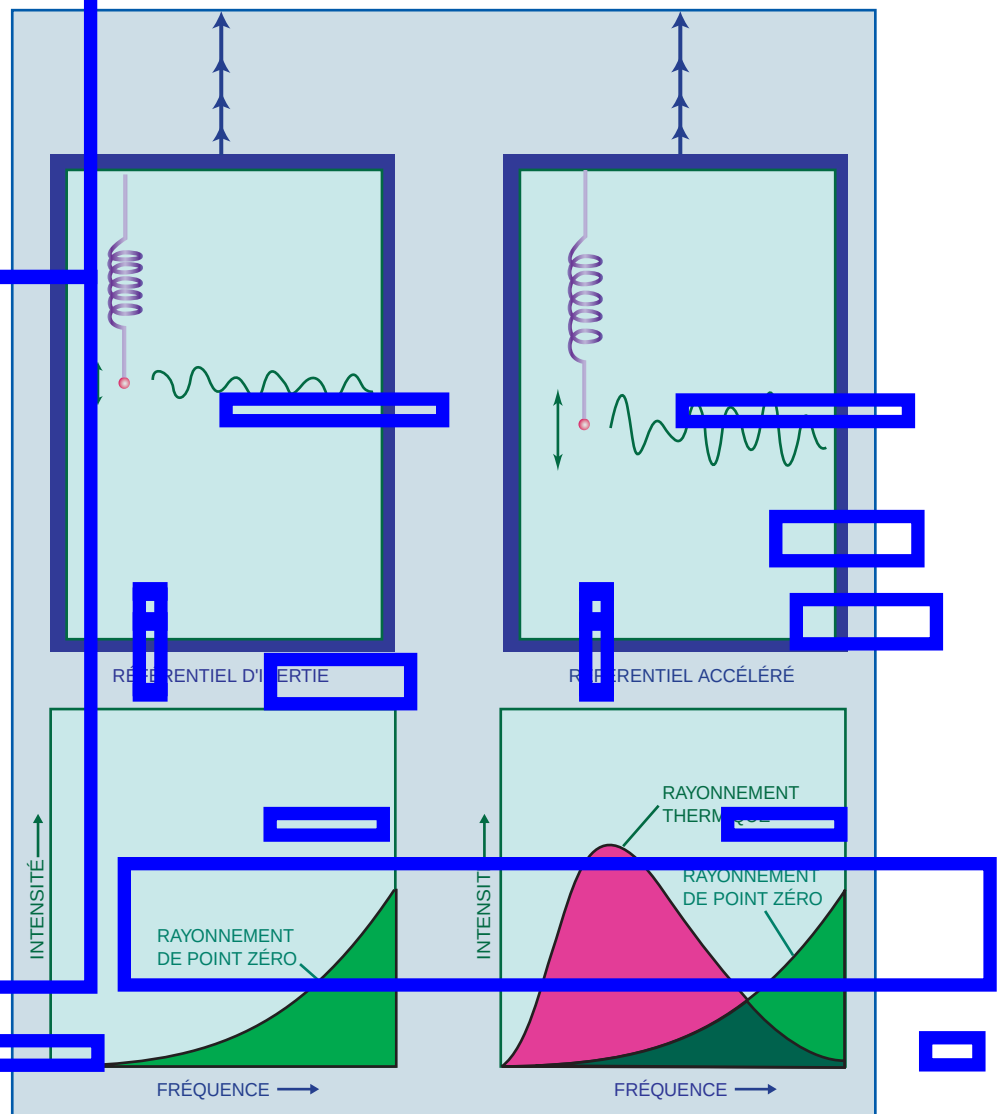
L'invariance par la transformation de Lorentz impose que le spectre de point zéro ait une structure bien particulière : l'intensité du rayonnement à une fréquence donnée doit être proportionnelle au cube de cette fréquence, car un spectre défini par une telle courbe apparaît identique à tous les observateurs, quelle que soit leur vitesse, à condition qu'elle soit constante. C'est en outre le seul spectre qui possède cette propriété.

L'énergie du vide

Le spectre de point zéro soulève une difficulté importante : comme l'intensité du rayonnement croît en fonction de la fréquence, la densité d'énergie du vide doit être infinie. Au XIX^e siècle, on aurait sans doute pensé que cette difficulté était

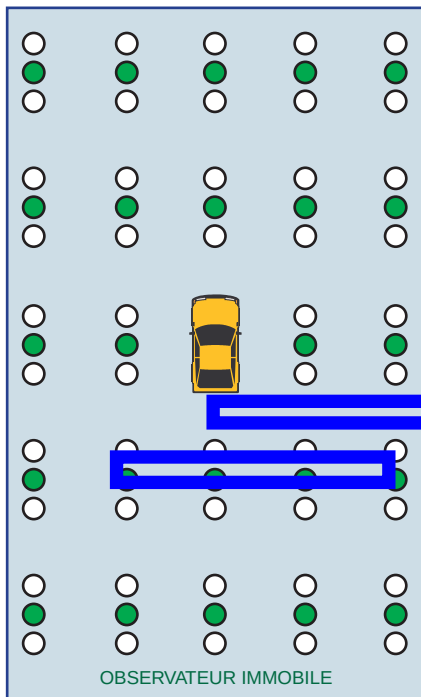
réductible, mais depuis les années 1940, on a rencontré des grandeurs infinies dans de nombreux domaines de la physique et on a mis au point des méthodes nouvelles pour en tenir compte. En général, on les a confrontées directement à des énergies infinies que pour les forces gravitationnelles ; tous les autres calculs font intervenir des variations ou des différences d'énergie qui, elles, sont toujours finies.

Si l'Univers baigne dans un rayonnement électromagnétique de point zéro, ce dernier devrait manifester sa présence par des phénomènes moins subtils que l'effet Casimir. On pourrait penser, par exemple, qu'il modifie le résultat de l'expérience du piston et du cylindre en s'opposant au retour du piston, même lors-

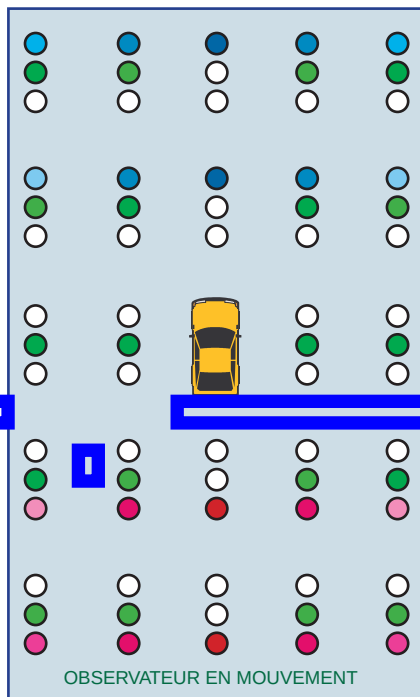


6. L'ACCÉLÉRATION d'un observateur, dans le vide, modifie le spectre du rayonnement qu'il observe. À la température du zéro absolu, un oscillateur harmonique placé dans un référentiel fixe ou se déplaçant à vitesse constante, n'est soumis qu'aux oscillations de point zéro. Dans un référentiel accéléré, la réponse de l'oscillateur est celle qu'il aurait si la température était supérieure à zéro.

SPECTRE "VERT"

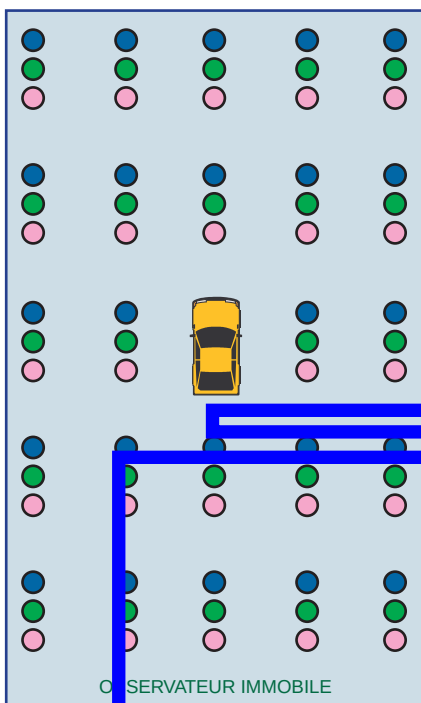


a

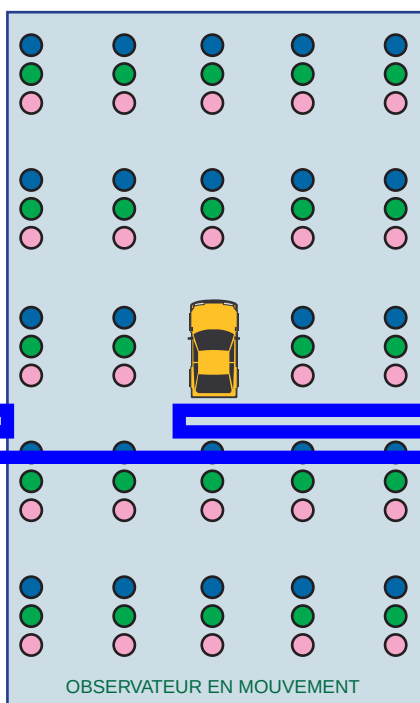


b

SPECTRE DE POINT ZÉRO



c



d

7. LE RAYONNEMENT DE POINT ZÉRO est invariant par la transformation de Lorentz, de sorte que deux observateurs en mouvement uniforme l'un par rapport à l'autre mesurent les mêmes propriétés du vide. La transformation de Lorentz relie les référentiels en mouvement uniforme relatifs. Pour qu'un rayonnement soit invariant par cette transformation, il faut que son spectre le soit. L'exemple considéré sur cette figure illustre la conséquence du mouvement sur le spectre. Un observateur dans une voiture est entouré de feux de signalisation d'un genre particulier : ces derniers indiquent l'intensité du rayonnement de point zéro à trois fréquences, correspondant au rouge, au vert et au bleu. Supposons par exemple qu'un observateur au repos par rapport aux signaux les voit tous verts (a), c'est-à-dire que le rayonnement soit concentré dans la partie verte du spectre électromagnétique. Si l'observateur se déplace (b), il voit les signaux bleus devant lui et rouges derrière lui, en raison de l'effet Doppler. La transformation de Lorentz indique en outre que les signaux qui se rapprochent sont plus brillants que ceux qui s'éloignent. Un seul spectre possède la propriété d'être invariant par cette transformation : celui tel que l'intensité d'un rayonnement est proportionnelle au cube de la fréquence. Lorsque les couleurs des feux de signalisation suivent cette règle, deux observateurs, dont l'un est au repos (c) et l'autre se déplace (d) observent les mêmes spectres.

qu'on a supprimé tout rayonnement thermique.

Les calculs indiquent que cette intuition est fautive : à l'équilibre, quand on n'applique aucune force extérieure au piston, le rayonnement baigne également l'intérieur et l'extérieur du cylindre, et les pressions qu'il exerce sur le piston s'équilibrent. Cet équilibre est ainsi assuré pour le rayonnement thermique comme pour le rayonnement de point zéro. Quand on repousse le piston à l'intérieur du cylindre, le rayonnement est comprimé. Wien avait, nous l'avons vu, étudié la modification du spectre en fonction d'une variation de volume ; j'avais calculé que le rayonnement thermique s'oppose à une compression, la température augmentant et exerçant une pression supérieure sur le piston. Le résultat est complètement différent pour le rayonnement de point zéro : son spectre n'est pas modifié lors d'une compression : seul un spectre décrit par une courbe cubique possède cette propriété remarquable.

Le spectre décrit par une courbe cubique permet-il d'interpréter l'effet Casimir lui-même ? Par chance, oui : quand on calcule les effets de ce spectre, on retrouve les résultats de Sparnaay : la force d'interaction entre les plaques doit être directement proportionnelle à leur surface et inversement proportionnelle à la puissance quatrième de leur écartement. Une fois encore, le spectre décrit par une courbe cubique est le seul qui permette d'obtenir une telle variation en fonction de la distance.

La nouvelle théorie classique de l'électron

Le spectre décrit par une courbe cubique n'est « unique » que par sa structure : il existe, bien sûr, une infinité de courbes de même forme, mais d'amplitudes différentes. Pour toutes ces courbes, l'intensité du rayonnement est proportionnelle au cube de la fréquence, le coefficient de proportionnalité étant différent pour chaque courbe. Il est impossible de calculer théoriquement la valeur de ce coefficient, mais grâce à la mesure de Sparnaay, on déduit cette valeur de celle de la force qui apparaît dans l'effet Casimir. Un calcul simple donne une valeur de $3,3 \times 10^{-34}$ joule-seconde, c'est-à-dire la moitié de la constante de Planck : la constante de Planck, présente en physique quantique, surgit dans un phénomène classique.

Pour décrire l'existence du rayonnement de point zéro à l'aide de la théorie classique de l'électron, il faut modifier beaucoup cette dernière : si l'on utilise encore les lois du mouvement de Newton, pour les électrons, et les équations de Maxwell, pour le champ électromagnétique, on impose de nouvelles conditions aux limites : le vide contient initialement des champs électromagnétiques aléatoires, dont le spectre est un spectre de point zéro. Cette théorie corrigée constitue la théorie classique de l'électron avec rayonnement électromagnétique classique de point zéro, que l'on appelle plus simplement l'électrodynamique stochastique.

Les nouvelles conditions aux limites modifient les prévisions théoriques. Considérons, par exemple, l'un des modèles favoris de la physique moderne : un oscillateur harmonique constitué d'un électron lié à un ressort parfaitement élastique et sans frottement. Plaçons ce système mécanique idéal dans le vide classique : quand on lâche l'électron après avoir tendu le ressort, l'électron oscille autour de sa position d'équilibre et il émet un rayonnement électromagnétique à une fréquence caractéristique.

L'oscillateur harmonique est un modèle pratique, car on peut facilement calculer le mouvement de l'électron.

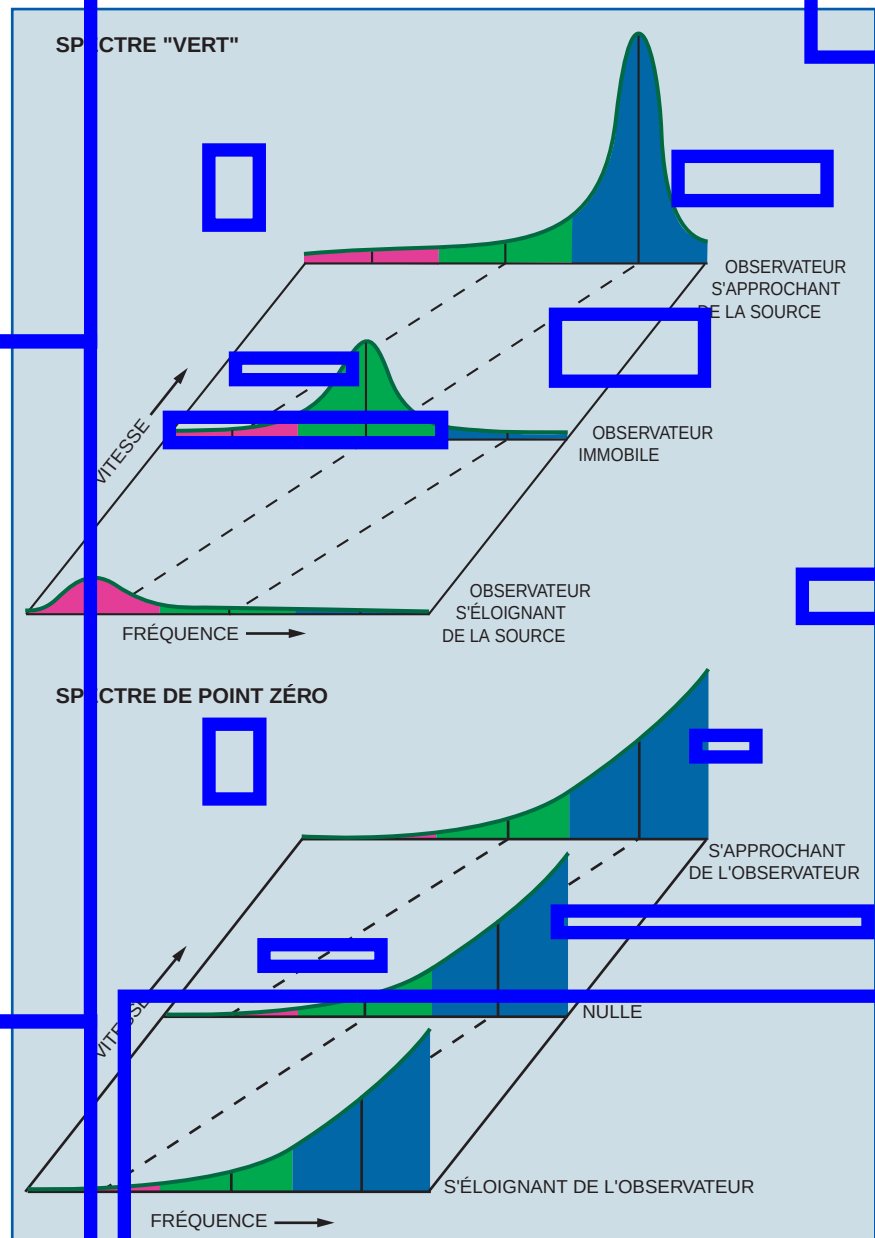
Selon la théorie classique de l'électron non modifiée, deux forces seulement agissent sur l'électron : la force de rappel du ressort et une force de réaction due à une émission de rayonnement ; comme la réaction s'oppose au mouvement de l'électron, les oscillations s'amortissent et l'électron finit par s'immobiliser. En revanche, selon l'électrodynamique stochastique, une force supplémentaire, due au rayonnement de point zéro, agit sur l'électron : l'électron interagit sans cesse avec les champs aléatoires de point zéro, de sorte qu'il ne s'immobilise jamais. L'oscillateur harmonique conserve une énergie moyenne, liée au spectre de point zéro et égale au produit de la moitié de la constante de Planck par la fréquence d'oscillation.

Jusqu'à ce point, nous avons décrit le vide classique tel que l'observe un observateur au repos ou en mouvement à vitesse constante. Les effets du rayonnement de point zéro sont cependant encore plus étonnants quand l'observateur est accéléré, c'est-à-dire lorsque sa vitesse change soit en intensité, soit en direction.

Les effets de l'accélération

Considérons un observateur dans une fusée qui accélère de façon constante par rapport à un référentiel supposé fixe, par exemple celui défini par des étoiles très éloignées. Comment cet observateur voit-il le vide classique ? Pour le savoir, il faut passer, par une transformation mathématique, du réfé-

rentiel fixe au référentiel accéléré. La transformation de Lorentz s'applique à des référentiels en mouvement uniforme les uns par rapport aux autres, mais ici, le problème est plus complexe, car la vitesse de l'observateur varie constamment. On peut cependant le résoudre en effectuant plusieurs transformations de Lorentz successives sur de petits intervalles de temps.



8. LE SPECTRE DE POINT ZÉRO est indépendant de la vitesse de l'observateur en raison des variations inverses de la fréquence et de l'intensité des rayonnements. Lorsqu'un observateur s'approche d'une source du rayonnement, toutes les fréquences sont décalées vers les hautes fréquences et toutes les intensités augmentent ; quand il s'en éloigne, c'est l'inverse qui se produit. Ainsi un spectre qui présente un pic dans le vert, pour un observateur immobile, présente un pic plus important dans le bleu pour un observateur qui s'approche de la source et un pic plus faible, dans le rouge, pour un observateur qui s'en éloigne. La courbe cubique qui décrit le spectre de point zéro est la seule où les décalages en fréquence et en intensité se compensent : la lumière qui apparaît verte dans le référentiel immobile devient bleue pour un observateur qui s'approche, mais son intensité devient égale à celle de la lumière bleue que voit l'observateur au repos ; de même, la lumière verte devient rouge pour un observateur qui s'éloigne, mais son intensité diminue en proportion.

L'énergie du vide est-elle récupérable?

Le principe d'incertitude d'Heisenberg limite nos connaissances... et détruit la possibilité que le vide ne soit «rien». Si nous supposons qu'une particule est dans une boîte nous limitons l'incertitude sur la position et notre incertitude sur la vitesse augmente. Le produit de l'incertitude sur la position par l'incertitude sur la vitesse est égale à une quantité faible, mais finie, la constante de Planck. Similairement l'incertitude sur l'énergie de tout objet dépend de l'incertitude sur le temps d'observation, ou plus simplement de la durée d'observation elle-même. Ainsi des particules peuvent apparaître dans le vide à cause de fluctuations d'énergie qui se transforment en masse pendant des durées extrêmement courtes.

Le vide n'est jamais vide! Les particules entrent et sortent venant de nulle part et cette mystique du vide est inséparable de la physique quantique.

Si l'énergie du vide existe, on devrait voir ses manifestations. Selon certains physiciens, cette énergie est peut-être l'élément stabilisateur de l'électron en orbite autour du noyau de l'atome. Dans la physique classique les électrons accélérés sur une orbite circulaire émettent un rayonnement électromagnétique et leur énergie diminue : les électrons devraient alors tomber, après très peu de temps, sur le noyau. Pour expliquer l'immobilité de leur trajectoire, le physicien américain controversé Harold Puthoff suppose que l'électron absorbe continuellement l'énergie du vide.

Les physiciens ont démontré, prolongeant ainsi l'expérience de Casimir, que la présence d'une cavité suffisamment petite empêche, dans la cavité, l'électron de tomber sur des niveaux d'énergie inférieure en émettant un rayonnement (voir Serge Haroche et Jean-Michel Raimond, *Électrodynamique quantique en cavité*, Pour la Science, juin 1993). Tout comme les grandes longueurs d'onde ne peuvent exister entre les plaques de l'expérience de Casimir, la cavité est si petite que les grandes longueurs d'onde du rayonnement du vide y sont absentes et les transitions correspondantes y sont interdites.

H. Puthoff pense que dans des configurations de confinement adéquat il serait possible de manipuler le spectre du vide de façon que le niveau inférieur de l'électron ait une énergie inférieure (l'atome serait plus petit). En tombant sur ce niveau, l'électron céderait une énergie récupérable. On ne peut encore imaginer de dispositif pratique pour récupérer l'énergie du vide fournie par des atomes de taille réduite. Les physiciens tentent aussi de concevoir un effet Casimir asymétrique permettant de propulser un vaisseau dans le vide intersidéral... sans résultat pour l'instant.

L'espoir de soutirer de l'énergie au vide est fondé sur l'hypothèse qu'une énergie notable est disponible. Certaines théories mathématiques avancent que l'énergie du vide est infinie : l'énergie qui nous baigne, argumentent les récupérateurs optimistes, est invisible car elle est uniforme et pénètre toutes les régions de l'espace, et ainsi la résultante des forces s'exerçant sur un objet est nulle.

Que les mathématiques indiquent une énergie du vide infinie n'est pas à prendre sans précaution, remarquent les pessimistes. Des infinis apparaissent dans certaines équations et les physiciens s'attachent à les faire disparaître, car ils pensent que ce sont des artefacts de calcul. Ainsi le calcul de la masse de l'électron donne un infini que les physiciens éliminent par un processus dit de renormalisation. Bienfait pour les mathématiciens, l'infini est abhorré par le physicien.

Sur la foi des résultats des expériences de type Casimir, beaucoup de physiciens pensent que l'énergie du vide est minime, les expériences élaborées n'ayant permis de n'extraire que 10^{-15} joule.

Pour extraire une quantité notable d'énergie il faudrait des appareils gigantesques.

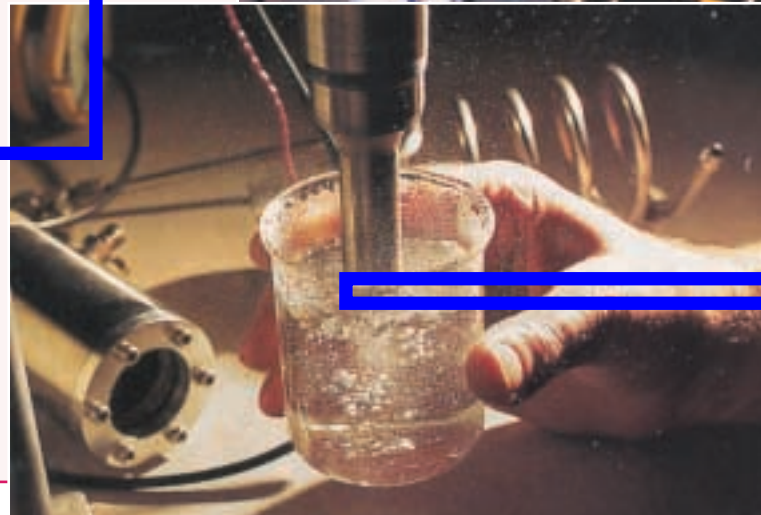
Un autre argument invalidant l'existence d'une énergie considérable dans le vide est la courbure de l'espace. L'énergie étant équivalente à la masse, l'énergie du vide devrait «courber» l'espace de façon notable, ce qui n'est pas le cas, la somme des angles d'un triangle étant égale à 180 degrés, aussi loin que nos mesures le permettent, sauf au voisinage de certaines étoiles massives. De plus, si cette masse équivalente existait dans tout l'Univers, elle déformerait l'espace dont la nature serait différente et la présence de cette masse serait observée. L'énergie du vide modifierait aussi la constante cosmologique, alors que les mesures astronomiques actuelles semblent montrer qu'elle est quasiment nulle.

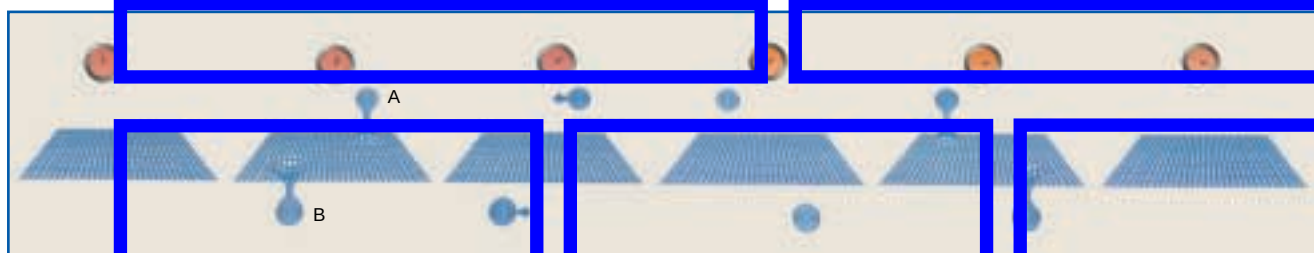
Si l'énergie du vide est importante, il doit exister une «énergie négative» qui l'annihile presque totalement. En résumé, l'énergie du vide ne serait importante que pour les pseudo scientifiques et les fervents mystiques du surnaturel : l'enthousiasme de Harold Puthoff est d'ailleurs critiqué.

Néanmoins certains physiciens invoquent l'énergie du zéro pour expliquer l'inertie, la résistance à l'accélération perçue par tous les corps. En réduisant l'énergie du zéro on pourrait peut-être accélérer plus facilement les masses. Et de rêver à de possibles voyages interstellaires... Il faudrait, pour étayer cette théorie, observer le pillage gravitationnel d'une masse en mouvement dans le vide.

L'exploitation de l'énergie du vide n'est plus un fantasme, mais elle reste irréaliste : trop bête pour être envisageable, voire possible.

DES PHYSICIENS, comme Harold Puthoff (ci-dessus, directeur de l'Institut des Énergies Avancées de Austin, au Texas, ont tenté d'extraire l'énergie du vide (ici au moyen d'un montage de production de bulles par ultrasons). Jusqu'à présent aucune tentative n'a donné de résultat probant.





9. LE VIDE EST SOUMIS AUX FLUCTUATIONS quantiques : en conséquence des particules virtuelles apparaissent spontanément dans le vide pendant des temps très courts, de l'ordre de la constante h de Planck. Ces particules sont émises en couple, une particule de matière, A et une particule d'antimatière, B. Ainsi, le vide n'est jamais tout à fait...

On pourrait penser que, pour un observateur accéléré, le spectre n'est plus isotrope et que l'«avant» et l'«arrière» paraissent différents ; on pourrait aussi imaginer que l'aspect du spectre varie tant que l'observateur est accéléré. En fait, ces intuitions sont erronées : le spectre est homogène et isotrope et ne change pas tant que l'accélération est constante. En revanche, le spectre observé par un observateur uniformément accéléré n'est pas celui que voit un observateur non accéléré. Pour chaque fréquence, le rayonnement est plus intense dans le référentiel accéléré que dans le référentiel au repos.

La forme du spectre électromagnétique classique pour un observateur accéléré n'est pas facile à comprendre, mais on peut l'interpréter en analysant le mouvement d'un oscillateur placé dans une fusée.

L'accélération, équivalente à un échauffement

L'équation du mouvement de cet oscillateur ressemble beaucoup à l'équation du mouvement dans un référentiel immobile. Il existe cependant deux différences notables : d'une part, la force de réaction au rayonnement comporte un nouveau terme, proportionnel au carré de l'accélération ; d'autre part, l'oscillateur est exposé à un nouveau rayonnement aléatoire lié à l'accélération. Au total, l'énergie moyenne de l'électron accéléré est supérieure à celle de l'électron oscillant dans un référentiel au repos. En d'autres termes, lorsqu'un oscillateur est accéléré, il vibre erratiquement plus fortement dans le vide que s'il était au repos.

Pour comprendre l'effet de l'accélération sur l'oscillateur harmonique, on peut chercher le spectre électromagnétique qui s'ajoute au rayonnement de point zéro pour engendrer ce mouvement supplémentaire. Il faut pour cela considérer le «principe

d'équivalence» sur lequel Einstein fonda sa théorie de la gravitation ; selon ce principe, un observateur, situé dans un petit laboratoire placé dans un champ gravitationnel, effectue exactement les mêmes mesures qu'un observateur placé dans une petite fusée qui accélère uniformément. La physique dans un référentiel accéléré est identique à celle dans un référentiel placé dans un champ gravitationnel uniforme. Comme les lois de la thermodynamique restent valables dans un champ gravitationnel, elles doivent l'être encore dans une fusée qui accélère uniformément : aussi, pour conserver l'équilibre, le spectre du rayonnement supplémentaire ne peut être qu'un spectre thermique. Dans tout autre cas, l'oscillateur ne serait pas en équilibre thermique avec son environnement et il pourrait servir à réaliser une machine à mouvement perpétuel.

La conclusion est directe et étonnante : un système physique uniformément accéléré dans le vide présente les mêmes propriétés à l'équilibre qu'un système non accéléré soumis à un rayonnement thermique à une température supérieure au zéro absolu.

William Unruh, de l'Université de Colombie britannique, et P. Davies, de l'Université de Newcastle-on-Tyne, ont trouvé vers 1976 la relation mathématique qui relie l'accélération à la température. Le spectre réel que peut enregistrer un observateur uniformément accéléré dans le vide est la somme de deux composantes : l'une est due au rayonnement de point zéro ; l'autre est le spectre du rayonnement thermique proposé par Planck en 1900. Planck ne put expliquer la forme de ce spectre qu'en introduisant, à contre-cœur, certaines idées de mécanique quantique ; on découvre aujourd'hui que l'on peut décrire cette courbe avec une analyse entièrement classique du rayonnement dans le vide !

Avec ce type de raisonnement, on obtient néanmoins un résultat surprenant : si l'on se fonde à nouveau sur le principe d'équivalence évoqué précédemment, il doit exister une température minimale au-dessous de laquelle on ne peut pas refroidir les systèmes placés dans les champs gravitationnels. Cette limite est une limite absolue, indépendante de toutes les difficultés pratiques qui s'opposent au refroidissement. À la surface de la Terre, cette limite est égale à 4×10^{-20} kelvins ; elle est bien inférieure aux températures que l'on sait réaliser, mais elle est supérieure à zéro.

Les relations entre le rayonnement thermique et la structure du vide classique révèlent une unité inattendue dans les lois de la physique, mais elle rend plus complexe notre conception du vide. Avec son ensemble de champs électromagnétiques aléatoires, le vide reste l'état le plus simple de la nature, aux étonnantes propriétés.

Timothy B. BOYER est professeur de physique à l'Université de New York.

H. PUTHOFF, a des sites sur internet, dont <http://www.newphys.se/elektromagnetism/physics/Puthoff/>

T.H. BOYER, *A Brief Survey of Stochastic Electrodynamics*, in *Foundation of Radiation Theory and Quantum Electrodynamics*, sous la direction de A.O. Larue, Plenum Press, 1980.

D.W. SCARMA, P. CANDELAS et D. DEBETS, *Quantum Field Theory, Horizons, and Thermodynamics in Physics*, vol. 30, n° 3, pp. 327-366, mai-juin 1981.

T.H. BOYER, *Derivation of the Blackbody Radiation Spectrum from the Equivalence Principle ...*, in *Physical Review D*, vol. 29, n° 6, 15 mars 1984.

S. LAMOREAUX, *Demonstration of the Casimir Force in Phys. Rev. Letters*, vol. 78, n° 1, pp. 5-8, janvier 1997.

L'infini est un révélateur

PATRICK DEHORNOY

L'infini est source d'inspiration. L'exploration d'ensembles hyperinfinis, impossibles à construire, mais pas inconcevables, a fait découvrir des objets mathématiques comme les tables de Laver et des propriétés des tresses.

Les tables de Laver sont des tableaux d'entiers dont les propriétés restent très mystérieuses, alors même qu'il est facile de les construire avec un ordinateur. Le plus étrange est que le peu que nous savions sur ces tables repose sur l'hypothèse indémontrable qu'il existe des ensembles "hyperinfinis". Cela illustre un nouveau type d'applications de l'infini en mathématiques, où celui-ci sert à fournir des intuitions plus que des démonstrations.

Opérations autodistributives

Les tables de Laver ont été découvertes dans les années 1980 par Richard Laver, théoricien des ensembles à l'Université de Boulder, au Colorado.

À côté d'opérations usuelles comme l'addition ou la multiplication, qui vérifient des lois telles que la com-

mutativité ($x * y = y * x$) ou l'associativité ($(x * y) * z = (x * (y * z))$), les mathématiciens étudient d'autres opérations qui vérifient des lois plus exotiques.

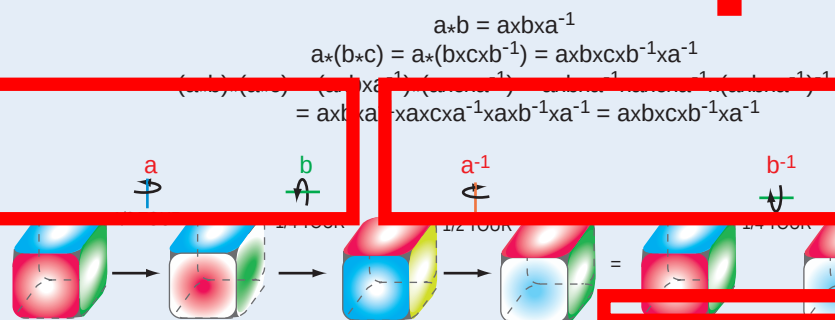
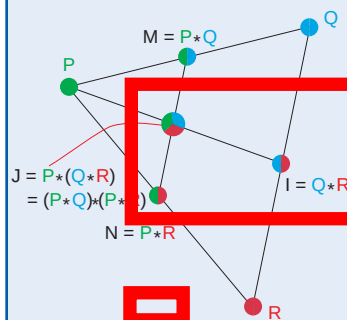
Parmi celles-ci figure l'autodistributivité, qui s'exprime par l'égalité $x * (y * z) = (x * y) * (x * z)$, et que nous noterons AD : le nom indique qu'une opération vérifiant la condition AD est distributive par rapport à elle-même, de la même façon qu'on dit la multiplication usuelle distributive par rapport à l'addition, parce qu'elle vérifie $x * (y + z) = (x * y) + (x * z)$. Nous appellerons donc opération autodistributive, ou simplement opération AD, une "multiplication" (sur des objets quelconques, pas nécessairement des nombres) qui vérifie la condition AD. Quels que soient les choix des éléments x, y, z ,

Il n'y aurait guère d'intérêt à choisir des conditions exotiques au hasard

et étudier pour le plaisir les opérations qui satisferaient de telles conditions (s'il en existe). L'intérêt pour les opérations AD provient de l'existence d'exemples non triviaux dans le monde mathématique, en liaison notamment avec d'autres structures importantes comme les groupes (voir l'encadré 1). Le rôle de l'infini dans le domaine tient à ce que la considération d'ensembles "hyperinfinis" a mené à la découverte d'exemples d'opérations AD d'un type complètement nouveau (voir l'encadré 4), lesquels ont donné un nouveau essor à toute la branche



1. EXEMPLES D'OPÉRATIONS AUTODISTRIBUTIVES



Dans le plan, on définit le «produit» de deux points P et Q comme étant le milieu du segment PQ. Par le théorème de Thalès, cette opération vérifie la condition d'autodistributivité : $P * (Q * R) = (P * Q) * (P * R)$. Un exemple d'opération autodistributive est la conjugaison dans un groupe, notamment

un groupe de rotations. Partant de la multiplication $\times$ du groupe, on définit une nouvelle opération $*$ en posant $a * b = a b c a^{-1}$. On vérifie que $a * (b * c)$ et $(a * b) * c$ sont égaux à $a b c b^{-1} a^{-1}$. Un exemple de conjugaison de rotations d'un cube est représenté à droite.

Ceci revient à chercher une sorte de table de Pythagore représentant une «multiplication». Si nous cherchons une table à N éléments, nous pouvons toujours supposer que les éléments de notre table sont les entiers $1, 2, \dots, N$, et notre problème est de construire un tableau à N lignes et N colonnes rempli de nombres entre 1 et N de sorte que, si nous dénommons $p * q$ le nombre qui est à l'intersection de la p -ième ligne et de la q -ième colonne, alors l'opération $*$ (la «multiplication») ainsi définie vérifie la condition AD. Si cela est le cas, nous dirons que nous avons obtenu une AD-table.

Tables de Live

Supposons que nous fixions les valeurs suivantes :

	1	2	3	N
1	2			
2	3			
3	4			
$\vdots$				
$\vdots$				
$\vdots$				
$\vdots$				
$N-1$	N			
N	1			

	1	3	4
1	2		
2	3		
3	4		
4	1		

Or $4 * 1$ est 1, donc on doit avoir $4 * 2 = 1 * 1 = 2$ et notre tableau devient :

Vérifions que les tables de Laver satisfont la loi d'autodistributivité (condition AD) : $p * (q * r) = (p * q) * (p * r)$. Sur le tableau à 4 éléments, choisissons $p = 4$, $q = 2$ et $r = 1$. Le premier membre est égal à $q * r = 2 * 2 = 8$ et $p * (q * r) = p * 8 = 4 * 8 = 8$. Calculons le second membre : $(p * r) = 4 * 2 = 6$; $(p * q) = (4 * 3) = 7$. $(p * q) * (p * r) = 7 * 6 = 42$. L'égalité n'est vérifiée. Il est facile de programmer la construction des tables de Laver sur un ordinateur, en suivant l'ordre de calcul des valeurs, pour éviter les débordements. En utilisant les propriétés de périodicité pour ne pas garder toute les valeurs en mémoire, on va jusqu'à la table à 1024 éléments sans problème, voire, en opérant avec modulo, jusqu'à 2^{20} éléments.

The diagram illustrates the construction of a 16x16 Latin square. It shows several 4x4 and 8x8 grids, some of which are highlighted with red boxes and labeled with 'p' and 'q'. These grids are arranged in a larger 16x16 grid, with some cells containing numbers 1 through 16. The red boxes and labels indicate the recursive construction of the Latin square.

Le lecteur pourra établir de même la valeur $4 * 4 = 4$. On passe alors à la ligne de 3. La première valeur à calculer est $3 * 2$, et il vient cette fois $3 * 2 = (1 * 1) * (3 * 1) = 4 * 4 = 4$. On remplit ainsi la ligne de 3, puis celle de 2 et de 1, parvenant finalement à :

	1	2	3	4
1	2	4	2	4
2	3	4	3	4
3	4	4	4	4
4	1	2	3	4

La période de la première ligne atteint 8 dans la table de Laver à 32 éléments, puis 16 dans la table à 512 éléments. Si le théorème de Laver est vrai, il existe des tables où la période de la première ligne vaut 32, et nous pouvons chercher quelle est la taille N de la première table où cela se produit, par exemple en programmant un ordinateur. Cette approche ne risque pas d'aboutir...

En effet, un théorème de Randall Dougherty, de l'université de Columbus, en Ohio, nous dit que le nombre N est au moins $A_9(A_8(A_8(255)))$, où A_k est la k -ième fonction d'Ackermann. Les fonctions A_k sont des fonctions sur les entiers, construites de proche en proche, et qui croissent de plus en plus vite : $A_1(n)$ est $n + 1$, $A_2(n)$ est $2n + 3$, $A_3(n)$ est de l'ordre de 2^n , $A_4(n)$ est approximativement une tour d'exponentielles de base 2 et de hauteur n , et, pour tout k , $A_{k+1}(n)$ s'obtient en appliquant A_k itérée n fois à 1. Le nombre $A_9(255)$ dépasse toute imagination, et que dire de $A_9(A_8(255))$, puis de $A_9(A_8(A_8(255)))$?

Sans risque d'être démenti, on peut affirmer que, quelle que soit la puissance de l'ordinateur utilisée ou l'astuce du programme développé, on trouvera toujours que la période de la première ligne dans une table de Laver est au plus 16. Pourtant le théorème de Laver nous dit que 32 est atteint, mais certainement bien plus loin que là où les ordinateurs savent encore compter...

La construction est la même pour toutes les tailles de tableau: notre règle du jeu donne à chaque fois une et une seule manière de remplir les cases de la dernière ligne, puis de l'avant-dernière, et ainsi de suite en remontant pour finir par la première ligne. Sont-ce des AD-tables? La construction effectuée garantit un certain nombre des égalités $p * (q * r) = (p * q) * (q * r)$, mais, en fait, seulement celles qui correspondent au cas particulier $r = 1$, car on se ramène toujours au cas précédemment calculé.

Pour les autres nombres du tableau, rien n'est garanti. La situation est alors la suivante: si la taille du tableau est une puissance de 2, alors le tableau que nous avons construit est bien une AD-table, sinon, ce n'en est pas une.

Ainsi, pour chaque entier n , nous avons obtenu une AD-table de taille 2^n , et c'est cette table qui est dénommée la n -ième table de Laver.

En dépit de la simplicité de leur construction, les tables de Laver sont des objets très compliqués et encore mal compris: en particulier, personne ne connaît de formule donnant explicitement la valeur du produit $p * q$ dans la n -ième table de Laver, et il est fort douteux qu'une telle formule existe!

Néanmoins, nous observons certaines propriétés sur les exemples de l'encadré 1 (et sur les exemples supplémentaires que le lecteur aura programmés sur son ordinateur...), qui ont été

démontrées en toute généralité par Richard Laver, et par Ales Drapal, un mathématicien de l'université Charles de Prague.

Nous citerons ici deux propriétés importantes. La première est que chaque table s'obtient simplement à partir de la table plus grandes: par exemple, nous obtiendrons la table à 8 éléments à partir de la table à 16 éléments en gardant le quart supérieur gauche et en «projetant modulo 8» les valeurs, c'est-à-dire en gardant les valeurs entre 1 et 8, et en retranchant 8 aux valeurs entre 9 et 16. La correspondance est la même à chaque niveau: la table à 6 éléments s'obtient à partir de la table à 32 éléments en projetant modulo 16, etc.

La seconde propriété concerne les lignes des tables. Dans la n -ième table (celle qui a 2^n éléments), par hypothèse, la p -ième ligne (en partant du haut) commence par la valeur $p * 1 = p + 1$. Ensuite, le long de la ligne, les valeurs augmentent jusqu'à atteindre la valeur maximale 2^n , à la suite de quoi on retrouve la valeur $p + 1$ et toutes les valeurs précédentes, de façon périodique. Par exemple, dans la table à 8 éléments, les valeurs successives sur la deuxième ligne sont 3, 4, 7, 8, 3, 4, 7, 8: il y a 4 valeurs différentes, et nous dirons que la période de la ligne est 4. Ainsi, dans chaque table de Laver, chaque ligne a une certaine période, dont nous savons qu'elle est elle-même nécessairement une puissance de 2: par exemple, dans la table à 8 éléments, les périodes des lignes 1 à 8 sont, respectivement, 4, 4, 2, 4, 2, 2, 1, et 8.

Le théorème de Laver

Intéressons-nous spécialement à la période de la première ligne. Nous lisons sur les exemples de l'encadré 2 que les valeurs pour les 5 premières tables sont 1, 1, 2, 4, et 4. Si le lecteur a programmé les tables suivantes, il trouvera ensuite les valeurs 8, 8, 8, 16, et 16 pour les tables à 32, 64, 128, 256, 512 et 1 024 éléments.

Qu'observons-nous? Que ces valeurs vont en croissant, plus exactement en ne décroissant pas. Cela était prévisible car nous savons que la n -ième table s'obtient en projetant modulo $2n$ la $n + 1$ -ième, et s'il y a m valeurs distinctes sur la première ligne de cette $n + 1$ -ième ligne, il y en a au plus m sur la première ligne de la n -ième table.

4. ENSEMBLES HYPERINFINIS ET RANGS AUTOSIMILAIRES

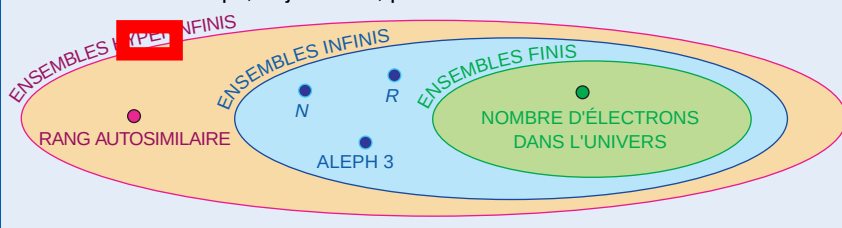
Un ensemble infini est, par définition, un ensemble tellement grand qu'il a des sous-ensembles aussi grands que lui. Par exemple, l'ensemble N des nombres entiers est infini, car la fonction qui envoie chaque entier n sur l'entier $n + 1$ établit une correspondance bijective entre N et le sous-ensemble de N constitué des entiers non nuls.

La théorie moderne des ensembles fait jouer un rôle fondamental à diverses extensions de la notion d'infini menant en fait ce qu'on peut appeler des ensembles hyperinfinis (on les dénomme aussi «grands cardinaux»). Un rang autosimilaire est un exemple de tel ensemble. L'idée est de repartir de la définition ci-dessus de l'infini, mais de la renforcer en demandant que la correspondance entre l'ensemble considéré X et l'un de ses sous-ensembles conserve toutes les structures additionnelles existant sur X . Reprenant l'exemple de la correspondance $n \rightarrow n + 1$

0	1	2	3	4	5	6	7	8	9	10	11	...	$p - 1$	p	...
1	2	3	4	5	6	7	8	9	10	11	...		p	$p + 1$	

entre N et un de ses sous-ensembles, nous voyons que cette correspondance conserve l'ordre des entiers ($p < q$ entraîne bien $p + 1 < q + 1$); en revanche, elle ne conserve pas l'addition ($p + q = r$ n'entraîne pas $(p + 1) + (q + 1) = r + 1$...). Les mathématiciens ont montré qu'il n'existe aucune correspondance sur N qui puisse convenir, et donc N , qui est infini, n'est pas hyperinfini au sens précédent. C'est cette notion forte d'infini qu'on dénomme autosimilarité et, donc, N n'est pas autosimilaire.

Pour élaborer de façon convenable une théorie des ensembles autosimilaires (s'il en existe), le contexte de la théorie des ensembles suggère d'utiliser des ensembles d'un type particulier appelé rang. L'existence d'un rang autosimilaire est l'un des axiomes d'un type infini que la théorie des ensembles étudie depuis les années 1970. Le lien avec les tables de Laver vient du fait qu'on peut construire à partir d'un rang autosimilaire une certaine opération AD (dont les tables de Laver sont des projections finies): c'est ainsi que Laver a, pour la première fois, introduit les tables qui, aujourd'hui, portent son nom.



Nous voyons aussi que les périodes doublent parfois. Sans qu'une loi apparaisse clairement sur un aussi petit nombre de valeurs, il apparaît plausible que la suite des périodes tende vers l'infini, c'est-à-dire atteigne successivement toutes les puissances de 2 (il est facile de voir que les périodes ne peuvent faire plus que doubler d'une table à la suivante, donc on ne peut sauter aucune puissance de 2). Qu'en est-il ?

En 1995, Richard Laver a prouvé le théorème suivant : «S'il existe un rang autosimilaire, alors la suite des périodes des premières lignes des tables de Laver tend vers l'infini.»

L'encadré 4 explique la notion de rang autosimilaire. Ce résultat est très étonnant : que la suite des périodes tende vers l'infini ne nous surprend guère, c'est ce que l'observation des exemples suggérait. Ce qui est surprenant ici est l'hypothèse avec laquelle Laver démontre son résultat, à savoir cette mystérieuse existence d'un rang autosimilaire. Cette hypothèse d'un rang autosimilaire exprime qu'il existe dans le monde mathématique des ensembles énormes, «hyperinfinis».

D'après le théorème d'incomplétude de Gödel, il n'existe aucun espoir de montrer que de tels ensembles existent, ni même de montrer que l'hypothèse qu'ils existent est non contradictoire. Ce qui est extrêmement surprenant alors, c'est l'apparition d'un lien entre les tables de Laver, qui sont des objets finis, calculables par ordinateur, et qui existent donc au sens le plus concret, et des ensembles hyperinfinis dont personne ne sait, ni ne saura jamais, s'ils existent.

Une alternative

La situation actuelle ne traduit que notre ignorance des propriétés des tables de Laver. Deux issues sont possibles :

- soit on finit par trouver une "vraie" démonstration du théorème de Laver, c'est-à-dire une démonstration qui n'utilise pas l'existence d'un ensemble hyperinfini,
- soit on finit par démontrer que toute preuve utilise comme hypothèse cette existence de façon incontournable.

Si la seconde issue est la bonne, la situation serait comparable à celle du

théorème de Goodstein, une propriété des entiers dont on sait qu'on ne peut pas la démontrer sans une certaine forme d'infini (voir *L'infini est-il nécessaire*, page 102) ; une distinction importante néanmoins est que l'hypothèse utilisée dans la preuve du théorème de Goodstein est l'existence d'un infini minimal, alors que celle utilisée dans la preuve du théorème de Laver est l'existence d'un infini extraordinairement plus grand.

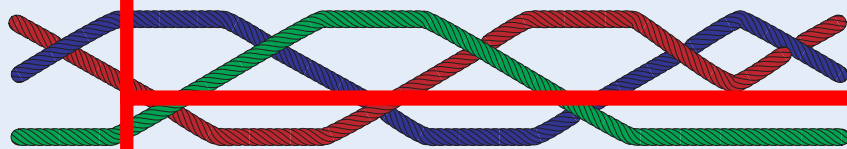
En théorie, rien n'empêche que cette issue soit la bonne (c'est même l'espoir de beaucoup de théoriciens des ensembles et de spécialistes de l'infini). Il semble cependant que la première issue ait bien davantage de chances d'être la bonne, même si les efforts pour démontrer «à la main» le théorème de Laver ont été jusqu'à présent infructueux ou, plus exactement, incomplets.

Applications de l'infini ?

Dans la situation actuelle, le théorème de Laver est une application de la notion d'infini, puisque la seule démonstration que nous en connaissons y fait

5. L'ORDRE DES TRESSSES

L'existence des tables de Laver et le théorème sur les périodes ne sont pas les seules applications de l'existence d'un rang autosimilaire. Une autre application a été la découverte de l'ordre des tresses. Les tresses mathématiques sont des diagrammes avec des brins qui se croisent, sur le modèle suivant :



Les tresses jouent un rôle important, car elles apparaissent dans de nombreuses branches des mathématiques et de la physique, en liaison notamment avec la théorie des nœuds et, de là, avec une grande variété de problèmes en physique, en chimie ou en biologie (voir *La science des nœuds*, Dossier hors série *Pour la Science*, avril 1997). On a également proposé récemment plusieurs applications à la cryptographie.

En 1992, un ordre remarquable sur les tresses a été découvert. Il a permis notamment de construire un nouvel algorithme de comparaison de tresses plus rapide que tout ce qui était connu auparavant.

Ce qui nous intéresse ici est la façon dont a été découvert cet ordre. Le point de départ est l'étude des rangs autosimilaires par Laver à la fin des années 1980 : comme mentionné dans l'encadré 4, cette étude amène naturellement à une certaine AD-table, et Laver, en 1989, a montré que cette table est sans cycle, ce qui signifie qu'on ne peut pas y trouver des éléments $x_1, \dots, x_n$ tels que chaque x_i apparaît dans la ligne de x_{i+1} et x_1 apparaît dans la ligne de x_n . Or on savait à l'époque que l'existence d'une seule opération AD sans cycle était importante pour l'étude de la condition AD en général, tout en n'ayant aucune raison de penser qu'un tel objet doive exister. Le théorème de Laver ne montrait pas vraiment cette existence, puisqu'il supposait lui-même l'existence indémontrable d'un ensemble hyperinfini, mais, au moins, il montrait qu'elle était plausible, et, du coup, suggérait qu'une construction directe devait être possible. Cette indication a été confirmée, et, de surcroît, elle a mené à des applications insoupçonnées : c'est la construction, sans l'aide d'un ensemble hyperinfini, d'une opération AD sans cycle, qui a naturellement amené à considérer des tresses et à les ordonner d'une façon que personne n'avait imaginée auparavant. Ainsi, il n'y a aucun lien direct entre les tresses et les ensembles hyperinfinis, si ce n'est que les premières ont été utilisées pour remplacer les seconds. Il n'empêche que, sans les ensembles hyperinfinis, nous en saurions aujourd'hui moins sur les tresses.

appel. S'il apparaît un jour une nouvelle démonstration qui ne l'utilise plus, le théorème cessera-t-il pour autant d'être une application ? En un sens strict, peut-être, puisque ni l'énoncé ni la nouvelle preuve n'utiliseront l'infini.

Mais, en fait, il me semble qu'il s'agira toujours d'applications : quoi qu'il arrive, c'est l'utilisation de l'infini qui aura permis de découvrir le théorème et il est plus que probable que personne n'aurait eu ni l'occasion ni la curiosité d'en chercher une démonstration si l'étude des rangs autosimilaires n'avait pas d'abord fourni un argument décisif, même fondé sur un principe indémontrable. C'est déjà ce qui s'est passé, il y a quelques années, avec l'ordre des tresses (voir l'encadré 5), et aussi avec

les tables de Laver elles-mêmes, d'abord déduites de l'existence d'un rang autosimilaire, puis, ensuite, reconstruites à la main en suivant la méthode que nous avons décrite ici.

Il est tentant de comparer ici la contribution de la théorie des ensembles hyperinfinis avec celle de la physique : se fondant sur des intuitions venues des modèles concrets, les physiciens proposent fréquemment des formules sans les démontrer rigoureusement ou, plutôt, en en donnant une première justification qui néglige des problèmes tels que divergence d'intégrales ou de séries (penser, par exemple, aux intégrales de Feynman). Il appartient alors aux mathématiciens de trouver d'autres démonstrations, complètes et rigoureuses. La situation est semblable lorsque des

ensembles hyperinfinis permettent de deviner et de justifier des propriétés comme l'existence des tables de Laver ou le théorème sur les périodes : comme précédemment, il reste ensuite à trouver d'autres démonstrations, libérées des hypothèses indémontrables.

Dans les deux cas, les contributions de la physique et de l'hyperinfini sont essentielles, car, sans elles, les propriétés n'auraient probablement pas été imaginées. Ces applications de l'infini sont d'un type nouveau en mathématiques : l'infini y sert de révélateur pour deviner des propriétés nouvelles, avant de disparaître ultérieurement lorsque sont trouvées les « bonnes » démonstrations.

L'existence de telles applications est un argument puissant en faveur de l'étude de l'infini (et de la théorie des ensembles) : elle montre que, même si l'on est tenté par une approche strictement constructiviste des mathématiques (voir l'article d'Allan Calder dans ce numéro) ou simplement si l'on ressent la plus grande méfiance et fort peu d'intérêt pour les ensembles infinis, et, *a fortiori*, hyperinfinis, on doit reconnaître à ces derniers la capacité de mener à la découverte de nouvelles propriétés des objets infinis.

Néanmoins, de ce point de vue, il importe peu que l'existence des ensembles infinis considérés soit plus ou moins plausible, puisque ceux-ci ont vocation à disparaître du paysage final : on peut même s'attendre à ce que les objets les plus puissants en termes d'applications soient les moins plausibles, comme ces rangs autosimilaires et autres ensembles hyperinfinis que nous avons rencontrés dans cet article et qui se sont déjà révélés fructueux.

Patrick DEHORNOY est professeur de mathématiques à l'université de Caen.

P. DEHORNOY, *Braids and self-distributivity*, Progress in Math., vol. 192, Birkhäuser, 2000.

J.-P. DELAHAÏE, *Les grands cardinaux*, in *Pour la science*, n° 224, pp. 60-67, juin 1996.

R. DUGHERTY, *Critical points in an algebra of elementary embeddings*, in *Ann. P. Appl. Logic*, n° 30, pp. 211-230, 1993.

C. KASSEL, *L'ordre de Denormoy sur les tresses*, Séminaire Bourbaki, exposé 865, novembre 1999.

R. LAVER, *On the algebra of elementary embeddings of a rank into itself*, in *Advances in Math.*, n° 110, pp. 334-346, 1995.

L'infini littéraire

PAUL BRAFFORT

Dans le sillage des philosophes, romanciers, poètes et essayistes ont été inspirés par l'infini

Voir un Monde dans un Grain de Sable
Et un ciel dans une Fleur Sauvage,
Tenir l'infini dans la paume de la main.
Et l'Éternité dans une heure.

William Blake
Augures d'innocence – 1803)

L'un, l'infini (la mathématique, la science), n'en finit pourtant pas : si la science veut décrire – voire expliquer – le monde, elle n'en saisit jamais qu'un infime fragment, et tout le reste... est littérature !

L'idée même de l'«infini» apparaît, dans le monde grec, au sein d'une culture qui ne s'est pas encore fragmentée. L'infini est donc littéraire et philosophique – voire linguistique – avant d'être mathématique (Platon et Aristote l'utilisent d'ailleurs pas le même mot pour le désigner). Et comme Paul Valéry, un écrivain qui, plus que tout autre, s'efforcera d'abolir les frontières au sein de la culture, l'observe en 1910 : «Il y a en nous une sensation finie de l'«infini.» Et ce n'est qu'un effet – une conséquence. Ce n'est pas une preuve de quoi que ce soit.» Avec Victor Hugo, Valéry sera ici notre guide.

Dès l'Antiquité, une dualité se dessine dans la problématique de l'infini : celle de l'arithmétique et du géométrique (du discret – dénombrable – et du continu). Une autre s'y ajoute au sein même du continu : celle de l'espace et du temps. Toutes intéressent et inspirent Valéry, comme elles ont intéressé Lucrèce, Bruno, Pascal, Goethe, Novalis, Hugo et intéresseront Aragon, Borges, Sinisgalli, Nabokov, entre autres. Ces dualités nous permettent de dessiner, dans le foisonnement littéraire, une certaine organisation de l'infini qui reprend la distinction traditionnelle entre les domaines de l'«extension» et de la «compréhension», mais doit tenir compte explicitement du contenu «privatif» qu'exprime le *in* de *infini*.

L'angoisse de Pascal – sur laquelle Valéry et Borges reviendront longuement – c'est aussi l'*Horror Vacui* (l'horreur du vide) de Sinisgalli (1945). Et cette angoisse se retrouve, comme le complément d'un ensemble, dans cette sensation de trop-plein qui parfois nous assaille, submergés souvent dans le déluge des signes, étouffés dans l'imbriication des réseaux, dans l'incendie des artifices. Il nous faudrait sans doute un nouvel Edgar Poe et un nouvel Eureka (ou peut-être suffit-il d'observer, avec Irène et Louis Sandomir, que «l'«*Pataphysique* contient tous les infinis»)...

La défense de l'infini

Quelqu'un va-t-il prendre enfin la défense de l'infini ?

Aragon
Paris-Journal, 13 avril 1923)

«L'infini se présente également, en arithmétique et en algèbre, comme un symbole d'impossibilité, comme une solution absurde et fautive...», observe le logicien Louis Couturat dans *De l'infini mathématique* (1896), et si, la même année, Valéry déclare, avec Teste : «J'ai mon *Infini*», il notera, dans ses *Cahiers*, en 1944 : «Les infinis sont des abus de langage et d'écriture, des développements fiduciaires sans garantie.» Borges avait dit aussi, en 1936, dans *Les avatars de la tortue* : «Il est un concept qui corrompt et dérègle tous les autres. Je ne parle pas du Mal, dont l'empire est circonscrit à l'éthique ; je parle de l'infini.»

Victor Hugo avait donné de ce sentiment de malaise une analyse quasi lexicale dans ses notes de travail de *William Shakespeare* (1864) : «Il semble que l'irrégulier, c'est l'inachevé, et que, dans l'inachevé, il y a de l'infini.» Il s'agit là d'une des nombreuses attestations de l'intérêt que porte Hugo à

un débat qui n'a cessé, depuis les éléates et les pères de l'Église, et que Louis Couturat met en scène, dans le livre III (*La critique de l'infini*) de son traité, sous la forme d'un dialogue «à la Galilée» entre «Le Finitiste» et «L'Infiniste». Et si, en 1923, Aragon se propose de prendre *La défense de l'infini*, il préférera détruire, en 1927, le manuscrit de l'ouvrage qu'il projetait : le logicien et le mathématicien sont désormais à l'œuvre, aux côtés du philosophe et du poète.

Dès le prologue de son grand traité en cours de publication, *La totalité* (la section III-2 du volume I, livre IV, s'intitule d'ailleurs *Totalité et Infini*), le philosophe Christian Godin s'exclame : «Tout conspirait chez Leibniz. Tout expirait chez la plupart des contemporains. Dédouction, dissémination, déterritorialisation, différence, tout un courant de pensée contemporain travaille sur ces syllabes *dé-, dis-,* qui sont devenues, en remplacement des *con- et syn-* classiques, les préfixes fétiches de la modernité.»

Le préfixe *in* est celui de *infini*, mais aussi de *incertitude*, d'*impossibilité*. L'expression littéraire la plus remarquable en est due à l'écrivain, peintre et philosophe polonais S. I. Witkiewicz, l'auteur d'*l'Inassouvissement* (1930), roman où est évoquée la théorie des types de Russell (remède contre les paradoxes), dont une version simplifiée fut imaginée par son ami le logicien, peintre et écrivain Léon Chwistek. Le thème de l'inassouvissement est ainsi directement lié à celui de l'infini !

Et c'est un mathématicien devenu peintre, poète et publiciste, Leonardo Sinisgalli, qui nous propose une version lyrique de *l'Infini* et de ses éternels veillements, dans *Horror Vacui* où l'on peut lire, dans le treizième fragment : «DÉCIMALES INFINIES – Qui a écrit que les nombres décimaux infinis donnent le sentiment d'une prolifération, d'un monstrueux travail de la nature, de

quelque chose qui rappelle les dynasties millénaires des Pontifes, des Empereurs, des Rois, des Princes ? Avec cette répétition périodique et cette succession chaotique des mêmes neuf chiffres, qui ne pense aux Ramsès, aux Innocents, aux Sixtes, aux Alexandre, aux Georges ? Quelque chose qui rappelle les vers de l'intestin, les bandelettes, les chaînes, des anneaux attachés à un billot, à une tête, à un nombre entier.»

Dans *Une vague de rêves* (1924), Aragon, lui, avait choisi d'accueillir courtoisement le monstre : «Qui est là ? Ah très bien : faites entrer l'infini.»

Escapes dans l'infini

La hiérarchie absolue de l'infini nous est inconnue. Il n'y a pas d'autre hiérarchie que la gravitation pour les corps et le rayonnement pour les esprits. Les deux phénomènes, qui se corrigent l'un par l'autre, sont en équilibre dans l'homme.

Victor Hugo

(William Shakespeare, *Notes de travail : l'infini dans l'art*, 1864)

Le premier infini qu'écrivains et mathématiciens s'efforcent de décrire est l'infini dénombrable, le plus proche de notre intuition, semble-t-il. Mais le logicien L. E. J. Brouwer, dans *Art, Life and Mysticism*, dès 1908, est réticent : «Dans la science tout ce qui est perçu est placé en dehors du Soi, dans un univers de perception indépendant du Soi ; son lien avec le Soi, son unique source et guide, est perdu. Cela construit alors un substratum logico-mathématique qui est complètement étranger à la vie, une illusion, et qui agit dans la vie comme une Tour de Babel avec sa confusion des langues.»

Tour de Babel, paradoxe de Zénon, autant de thèmes qui inspireront Valéry, Borges et Calvino : on n'a pas oublié la vingt-et-unième strophe du *Cimetière marin* : «Zénon ! Cruel Zénon ! Zénon d'Élée ! / Mas-tu percé de cette flèche ailée / ...». John Irwin consacre un chapitre entier du livre qu'il publie en 1994 au fameux paradoxe vu par Borges, dans *La mort et la boussole* (1942) et *L'aleph* (1945). Calvino en fera le thème de *Temps zéro* (1966), la nouvelle la plus réussie de son *Cycle des années* suivante, dans *Cybernétique et fantasmes*, il précise : «Je pourrais ajouter que la numérabilité, la finitude sont en train de triompher de l'indétermination des concepts qui ne peuvent pas être soumis aux mesures et aux délimita-

tions ; mais cette formulation risquerait de donner une idée un peu simpliste des choses alors que la situation réelle est exactement contraire : tout processus analytique, toute division en parties tend à donner du monde une image qui va se complexifiant, juste comme, en refusant un espace continu, Zénon d'Élée finissait par concevoir entre Achille et la tortue une subdivision infinie de points intermédiaires.»

Au paradoxe ancien de Zénon s'étaient ajoutées, au début du siècle, d'autres antinomies qui mettaient en cause la notion d'ensemble – et même le fonctionnement de notre langage : celles de Russell, de Richard, de Grelling, etc. Henri Poincaré s'exclamait alors, dans *La logique de l'infini* (1909) : «Est-il possible de raisonner sur des objets qui ne peuvent être définis en un nombre fini de mots ? Est-il possible même d'en parler en sachant de quoi l'on parle, et en prononçant autre chose

que des paroles vides ? Ou au contraire doit-on les regarder comme impensables ? Quand à moi je n'hésite pas à répondre que ce sont de purs neants.»

Valéry ne l'exprimait pas autrement : «Le résultat des arguments de Zénon, c'est la démonstration d'une confusion dans le langage» (1917), tout comme Sinigaglia qui, dans *Le Devin* (1944), fait dialoguer ainsi ses personnages :

Le Roi : *Ce sont des signes ou des choses ?*

Le Devin : *Ce sont des mots, Sire.*

La Clôture

Leur pouvoir suggestif ne dépend pas tant d'un signifiant caché ou latent que la lecture doit mettre au jour, il utilise plutôt toutes les ressources du signifiant et de la structure afin de poser un droit au sens multiple et, ainsi, de libérer la lecture à l'infini.

David Banonk
(*La lecture infinie*)



1. LE PROMENEUR SUR LA MER DE NUAGES, de Caspar David Friedrich (1818) évoque cet infini qui a inspiré de nombreux écrivains et poètes.

Le désir de la connaissance – qu'elle soit scientifique ou littéraire – a vite fait de devenir passion... ou mystique. L'expression qu'en donne Pascal, «la nature est une sphère infinie dont le centre est partout et la circonférence nulle part», a fait l'objet de nombreux commentaires de Valéry et de Borges. Le thème du cercle ou de la sphère est rassurant, il permet d'assigner à ce qui semble une expansion sans limite une forme de clôture», de compléter l'infini dénombrable par l'infini continu, et c'est encore le poète Sinisgalli qui, dans *Horror Vacui*, en célèbre l'invention (certains diraient la découverte) dans ce fragment :

«C'est le mérite de Georg Cantor de nous avoir fait sentir l'épaisseur, la densité, la puissance du continu, la mesure de l'infini, l'ordre de l'ensemble des nombres. La route qui du rien mène à l'unité, de l'unité au multiple, jusqu'à Dieu, nous pouvons aujourd'hui la parcourir pour la première fois sans danger d'y rencontrer des sauts ou des lacunes. Aucun nombre ne peut plus nous échapper. Cantor a trouvé la place qui convient à chacun. Il a ordonné les points d'un segment, d'une ligne, d'une surface, d'un espace, et il a trouvé par voie de correspondance la mesure, le terme de comparaison de ces infinis. Il a commencé par considérer l'ensemble des nombres entiers, l'ensemble des nombres pairs, l'ensemble des nombres premiers, il a trouvé que ces trois ensembles ont la même puissance, la puissance du dénombrable, le même nombre cardinal qu'il a appelé Aleph zéro. [...] Cantor a démontré que même l'ensemble des nombres rationnels et celui des nombres algébriques ont la capacité d'être dénombrés; mais l'infinité des nombres algébriques est infime comparée à l'infinité des nombres transcendants. Les nombres algébriques et les nombres transcendants forment l'ensemble des nombres réels, et Cantor a démontré que cet ensemble a la puissance du continu, exprimée par un deuxième nombre cardinal transfini. [...] Cantor a trouvé une loi d'engendrement de la multitude des nombres ordinaux finis et transfinis, il a

trouvé une dynastie, celle des Aleph, et cela, à l'aide de deux principes seulement, l'un immanent (additif), l'autre transcendant (passage à la limite): Cantor, législateur de l'infini.»

Pourtant toute écriture (littéraire ou mathématique) n'est jamais qu'engendrement d'un nombre fini de signes, et sa capacité à représenter le continu demeure problématique. Paul Valéry rappelle, dans ses *Cahiers*, en 1943, que «l'infini est le type de cette illusion, laquelle consiste à ne pas voir le retour invincible au point le plus proche de l'état de pouvoir recommencer un écart». Dès 1927, il évoque «l'idée d'un monde fermé contenant un ou plusieurs infinis» et précise, en 1935 : «L'opération de l'artiste consiste à tenter d'enfermer un infini. Un infini potentiel dans un fini actuel.» Comme en mathématique, les techniques «effectives» ont un caractère essentiellement combinatoire : labyrinthe, graphes aux subtiles bifurcations, arrangements tabulaires complexes, involutions spéculaires aux reflets illimités. L'OuLiPo (l'OUvoir de LIttérature POtentielle, groupe d'écrivains créé par le mathé-

maticien François Le Lionnais et par Raymond Queneau en 1960, qui a notamment accueilli I. Calvino et G. Perec) en a éprouvé les pouvoirs.

Comme Borges, Calvino lance un défi aux labyrintes : «Et qui croit pouvoir vaincre les labyrintes en fuyant leur complexité passe à côté de la question. Donc poser l'existence d'un labyrinthe puis demander à la littérature de fournir une clé pour en trouver l'issue apparaît comme une demande dépourvue de toute pertinence. Ce que peut faire la littérature est de définir la meilleure attitude possible pour trouver une voie de secours, même si cette issue ne consiste que dans le fait de passer d'un labyrinthe à un autre. C'est le défi au labyrinthe que nous voulons sauver, c'est la littérature du défi au labyrinthe que nous voulons mettre en évidence en la distinguant d'une littérature de reddition au labyrinthe.» Et dans *Cosmopolitique et fantasmes* (1967) Calvino ajoute : «Nous avons dit que la littérature est, tout entière, dans le langage, qu'elle n'est que la permutation d'un ensemble fini d'éléments et de fonctions. Mais la tension de la littérature ne viserait-elle pas sans cesse à échapper à ce nombre fini? Ne chercherait-elle pas à dire sans cesse quelque chose qu'elle ne sait pas dire, quelque chose qu'on ne peut pas dire, quelque chose qu'elle ne sait pas, quelque chose qu'on ne peut pas savoir? Telle chose ne peut pas être dite tant que les mots et les concepts pour l'exprimer et la pensée n'ont pas été employés dans cette position, n'ont pas été disposés dans cet ordre, dans ce sens. Le combat de la littérature est précisément un effort pour dépasser les frontières du langage; c'est du bord extrême du dicible que la littérature se projette; c'est l'attitude de ce qui est hors du vocabulaire qui meut la littérature.»

Le «défi au labyrinthe», c'est l'effort pour associer aux situations d'engendrement combinatoire des filtres «anti-combinatoires» (telle est la «contrainte» de l'OuLiPo). Dans le *Conte à votre façon* de Queneau, *Les villes invisibles* de Calvino, *La vie mode d'emploi* de Perec, *Feu pâle* de Nabokov, l'infini potentiel est



2. LE SCULPTEUR français Auguste Préault (1809-1879) a dessiné un portrait d'Eugène Delacroix, à qui il prête cette réflexion sur l'infini.

maîtrisé, mais fait naître un désir d'achèvement actuel (on songe au théorème de complétion d'un espace topologique uniforme). L'interactivité offerte par les systèmes du type *Conte à votre façon*, en particulier, n'est pas sans évoquer les «suites de libre choix» de Brouwer, et donc une forme élégante – et constructive – d'appréhension du continu.

Les sirènes de Titan

*C'est donc vers l'air que je déploie mes ailes
confiantes
Ne craignant nul obstacle, ni de cristal ni
de verre,
Je fends les cieux et m'érige à l'infini.
Et tandis que de ce globe je m'élève vers
d'autres globes
Et pénétre au-delà par le champ éthéré,
Je laisse derrière moi ce que d'autres voient
de loin.*

Giordano Bruno
(*De l'infini, de l'univers et des Mondes*)
Le titre de cette section, comme des précédentes, est celui d'une œuvre littéraire, ici un roman de Kurt Vonnegut, où l'auteur introduit des entités spatio-temporelles qu'il nomme «chrono-synclastic infundibula», événements semblables à ces «moments de synchronisation cosmiques» décrits par Nabokov. Ce sont des moments d'exploration simultanée de l'espace et du temps, de circulation le long de cette dimension d'«échelle» à laquelle les physiciens s'intéressent aujourd'hui, mais que Victor Hugo évoquait déjà dans *Les misérables* (La maison de la rue Plumet) : «Qui donc peut calculer le trajet d'une molécule ? que savons-nous si des créations de monde ne sont point déterminées par des chutes de grains de sable ? Qui donc connaît les flux et les reflux reciproques de l'infiniment grand et de l'infiniment petit. le retentissement des causes dans les précipices de l'ére et les avalanches de la création ?»

L'infini atteint ainsi la dimension d'échelle qui s'ajoute à celles de l'espace et du temps, et l'on peut imaginer que les nouvelles techniques, qui nous permettent de visualiser ces dimensions nouvelles, au service de la physique ou de la biologie, des formes littéraires et artistiques aussi bien que mathématiques de la représentation nous offriront des accomplissements, des assouvissements inédits...

Terminons par un véritable moment de synchronisation cosmique : on sait que les *Nouvelles Impressions d'Afrique* de Raymond Roussel (1932), dont la composition et les illustrations par H.-A. Zo ont été préparées avec un soin maniaque par l'auteur, se terminent par un dessin dont la spécification (qui lui avait été transmise par une agence de détectives privés) était : «Un pan de ciel étoilé sans paysage terrestre semblant vu d'un point de l'espace sidéral donnant l'impression de l'infini.» Ce poème et ses illustrations sont suivis, dans l'édition Lemerre, par un poème : *Mon Âme*, écrit à l'âge de 17 ans, réédité sous le titre *L'Âme de Victor Hugo*, et dont voici les deux derniers quatrains (le premier suggère évidemment l'audacieuse substitution : «Victor Hugo» → «Raymond Roussel») :

*À cette explosion voisine
De mon génie universel,
Je vois le monde qui s'incline
Devant ce nom : Victor Hugo.*

*Sur la terre, que je domine,
Je vois ce feu continu
Qui, seul et sans frère, illumine
Partout l'univers actuel.*

Et c'est dans *Abîme*, le chant XLI et dernier de *La légende des siècles*, que Victor Hugo, le 26 novembre 1853, conclut en donnant la parole à notre héros :
L'INFINI
L'être multiple vit dans mon unité sombre.

Paul BRAFFORT est responsable du programme scientifique de l'ALAMO (Atelier de littérature assistée par la mathématique et les ordinateurs)

H. POINCARÉ, *Dernières pensées*, Flammarion, 1913.

P. VALÉRY, *Cahiers*, Gallimard, 1974.

T. LÉVY, *Figures de l'infini*, Éditions du Seuil, 1987.

J.-L. BOERGE, *Avatars de la tortue, Œuvres complètes I*, Gallimard, 1993.

P. ZELLINI, *Breve storia dell'infinito*, Adelphi, 1992.

J.-F. LUMINET, *Les poètes et l'univers* (anthologie), Le Cherche Midi, 1996.

C. GODIN, *La Totalité*, Champ Vallon, 1998.

Paul BRAFFORT, *Science et Littérature*, Diderot, diffusion Éditions du Seuil, 1998.

Les geysers

PHILIPPE LEBLANC

Dans la cheminée d'un geyser, l'eau est sous pression et à température élevée. Périodiquement, un brusque changement d'état s'amorce et aboutit à une violente éruption : la colonne d'eau surchauffée passe de l'état liquide à l'état vapeur.

Alcalay/J. Bios

La Terre a parfois les allures d'une baleine expulsant l'eau par son évent. Les geysers ne sont cependant pas le signe d'une vie interne du globe, mais l'une des manifestations les plus étranges du cycle de l'eau. À intervalles réguliers ou irréguliers, un mélange d'eau chaude et de vapeur en provenance des entrailles de la Terre est brutalement projeté à plusieurs dizaines de mètres de hauteur. D'où provient cette vapeur ? Comment les projections sont-elles régulées ?

On dénombre quelques 30 champs de geysers dans le monde, dont les deux tiers seulement contiennent plus de cinq geysers actifs. Le parc national de Yellowstone (ci-dessus, Geyser Castle), aux États-Unis, en contient davantage, le plus connu étant *Old Faithful* qui, depuis 100 ans, laisse échapper toutes les 15 minutes, un jet d'eau et de vapeur de 50 mètres de hauteur. Certains geysers sont plus rapides : le *Strokkur*, en Islande, explose toutes les 10 à 15 minutes. D'autres encore sont beaucoup plus lents, les éruptions se produisant uniquement après plusieurs heures de repos.

Les geysers sont assez rares, car ils résultent de la conjonction de divers phénomènes, notamment d'une activité volcanique et d'une circulation d'eau à relativement faible profondeur. Un geyser évolue au rythme des modifications de la croûte terrestre, sur des centaines, voire des milliers d'années. L'activité d'un geyser change souvent quand un séisme se produit à proximité, car la structure du sous-sol, dont nous expliquerons l'importance, change. Dans les régions de volcanisme actif, les évolutions de la croûte terrestre expliquent aussi l'apparition de nouveaux geysers. Les solfatares sont des « geysers calmes », qui relarguent de la vapeur d'eau chargée de soufre.

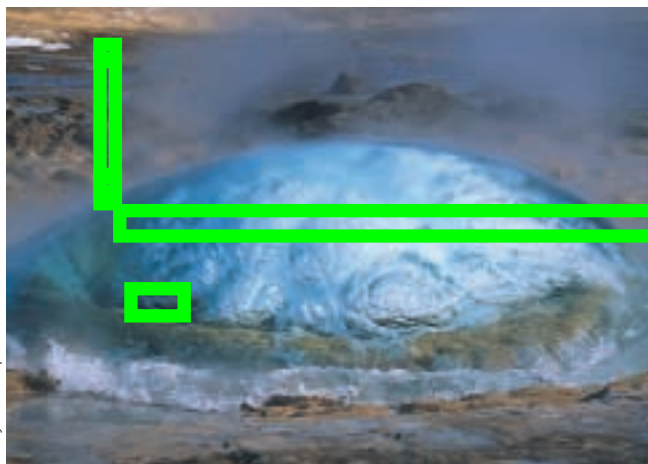
L'eau expulsée par les geysers (d'après le mot islandais *geysir* qui signifie *jaillir*) provient essentiellement de l'eau de

pluie qui s'est infiltrée très lentement dans le sol, à travers des fractures ou des roches poreuses sur plusieurs kilomètres. La température des roches s'élève progressivement avec la profondeur : elle augmente de 30 à 40 °C par kilomètre. Les premiers kilomètres, toutefois, la température est notablement supérieure à proximité des zones d'activité volcanique, car elle est proche du magma en fusion ou en cours de solidification. L'eau qui s'infiltré dans un sous-sol poreux se réchauffe progressivement et, parfois, s'accumule. Pour qu'un geyser se forme, il faut qu'une telle réserve d'eau soit connectée à l'air libre par un réseau de conduits sous-terrains.

RETARD À L'ÉBULLITION

Pourquoi un geyser donne-t-il l'impression d'exploser ? En raison d'une sorte de réaction en chaîne qui prend naissance au sein de la masse d'eau surchauffée et sous pression. En profondeur, l'eau supporte tout le poids de la colonne d'eau située au-dessus d'elle, qui est proportionnel à la profondeur considérée. Ainsi la pression qui s'exerce sous une colonne de 10 mètres de hauteur est égale à deux fois la pression atmosphérique (une colonne d'eau de 10 mètres de hauteur équivaut à une atmosphère à laquelle s'ajoute la pression atmosphérique qui s'exerce sur le haut de la colonne). Or la température d'ébullition de l'eau augmente avec la pression. L'eau bout à 100 °C sous pression atmosphérique normale, à 120 °C lorsque la pression est égale à deux fois la pression atmosphérique, à 158 °C sous 50 mètres (6 fois la pression atmosphérique). Dans les geysers les plus profonds la température de l'eau atteint 300 °C sans qu'il n'y ait d'ébullition. Considérons maintenant une colonne d'eau progressivement

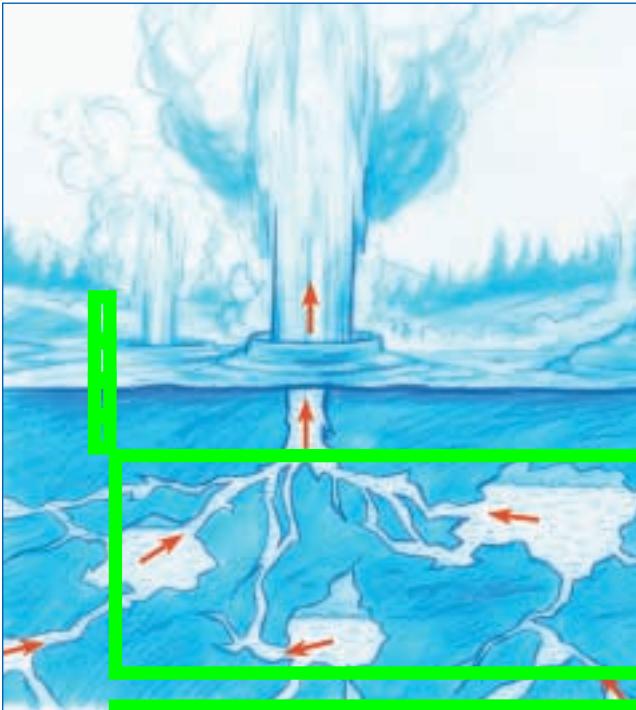
Alcalay/Jean-Jacques Bios



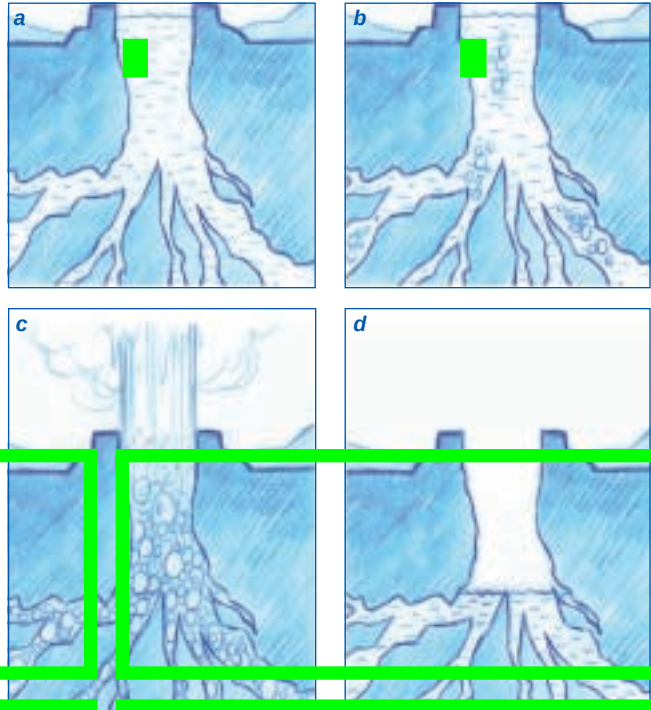
Klein/Hubert Bios



À gauche, le geyser *Strokkur*, en Islande, juste avant l'éruption et, à droite, le geyser *White Dome*, à Yellowstone, aux États-Unis.



Dessins de Bruno Vacaro



Les geysers sont alimentés par un réseau souterrain ramifié (à gauche), où circule de l'eau d'infiltration. Pour qu'il y ait éruption, il faut que la colonne qui est en contact avec l'air libre (à droite, a), soit pleine d'eau. Parfois, localement, l'eau surchauffe, c'est-à-dire que sa température dépasse la température d'ébullition pour la pression subie ; elle est alors dans un état métastable et la moindre perturbation locale rompt cet état : de petites bulles de vapeur d'eau se forment (b). Dès que de telles bulles apparaissent, elles remontent vers la surface, repoussant l'eau vers le haut. Dès lors, à divers endroits, la pression diminue

et de l'eau surchauffée passe de l'état liquide à l'état vapeur. Les bulles de vapeur qui se forment sont de plus en plus nombreuses et de plus en plus grosses, jusqu'à ce qu'elles emplissent toute la colonne (c). La brusque augmentation de volume fait jaillir le geyser. Lorsque la pression retombe, l'eau se stabilise dans la colonne, laquelle est partiellement vide (d). Lorsque la colonne sera à nouveau pleine d'eau et que la température de l'eau aura localement dépassé la température d'ébullition normale pour la pression subie, l'eau sera dans un état instable et la « réaction en chaîne » recommencera.

chauffée en son milieu. Lorsque la température d'ébullition est atteinte dans cette région, l'eau passe ponctuellement de l'état liquide à l'état vapeur, et des bulles de vapeur d'eau se forment et commencent à remonter vers la surface. Cette pression interne repoussée vers le haut de l'eau liquide, de sorte que la pression à laquelle est soumis le liquide dans la partie basse de la colonne diminue, jusqu'à ce que la température d'ébullition du liquide dans cette région diminue et devienne inférieure à la température ambiante. De nouvelles régions se mettent alors à bouillir : la vapeur d'eau qui se forme au sein du liquide tend à faire remonter davantage de liquide, ce qui abaisse encore localement la pression. Cette réaction en chaîne, très rapide, se traduit par la brutale projection d'un grand volume de vapeur et d'eau vers la surface.

vent saurée en silice, qui précipite juste à la sortie du geyser : la température baisse brusquement et avec elle la solubilité des électrolytes dissous, et ces derniers précipitent, formant un petit dôme de silice. Dans le cas des solfatares, la vapeur d'eau qui s'échappe est enrichie en soufre provenant du magma. À l'air libre, la solubilité du soufre dans l'eau diminue et le soufre pur cristallise. De l'acide sulfurique se forme également par réaction du sulfure d'hydrogène qui s'échappe de la Terre et de l'eau. À proximité de certains volcans s'étendent même des lacs d'acide sulfurique plus ou moins concentré. Dans d'autres cas, l'eau chaude dissout divers minéraux pour former des lacs de boue, desquels émergent de grosses bulles gaisantes qui éclatent en faisant un bruit étrange.

UNE SOURCE D'ÉNERGIE RENOUVELABLE

Ce phénomène s'arrête lorsque la température du liquide restant n'est plus assez élevée ou lorsque le volume d'eau présent dans le réseau de conduits est trop faible. Ainsi, les explosions sont intermittentes, séparées par des périodes de calme. Il faut attendre que le réseau des conduits directement connectés à la surface soit de nouveau plein d'eau chaude pour que l'ébullition et la propulsion recommencent. La longueur de la conduite principale, l'existence d'un bassin de stockage en profondeur, ou encore la configuration du réseau de fractures alimentant la conduite conditionnent la nature du geyser. Les vitesses de remplissage et de réchauffement imposent le rythme des pulsations du geyser. Les données sur la structure réelle des geysers sont rares, surtout parce qu'il est difficile de les explorer.

L'énergie emmagasinée dans le sol et éventuellement libérée sous forme de geyser est exploitée par l'homme. L'eau réchauffée au contact de roches chaudes en profondeur est utilisée pour les centrales géothermiques ou directement par diverses industries ou pour le chauffage des habitations. Ce sont principalement les États-Unis, les Philippines, l'Amérique du Sud, la Nouvelle-Zélande, mais aussi l'Islande, qui utilisent cette source d'énergie. Les sources chaudes riches en minéraux sont réputées bonnes pour la santé ; elles proviennent généralement des geysers plus doux, éjectant l'eau presque en continu.

La vapeur d'eau à haute température dissout la plupart des espèces chimiques à la surface des roches. Ainsi, elle est sou-

Philippe COUSSOT travaille au laboratoire des matériaux et des structures du Génie civil, LMSGC / UMR 113 LCPC-CNRS, à Champs-sur-Mame.
J.S. RHINEHART, *Geysers and geothermal energy*, Springer-Verlag, Berlin, 1980.
J. ELDER, *Geothermal systems*, Academic Press, New York, 1981.

Illusion ou réalité?

JEAN BULLIER

Un tableau dynamique défie les neurophysiologistes de la perception.

Regardez le tableau *Enigma* d'Isia Levant (voir la figure 2) en fixant le centre. Après quelques secondes, des nuées colorées apparaissent et se déplacent le long des anneaux. Parfois, à la faveur d'une saccade oculaire ou d'un clignement des yeux, le sens de rotation des nuées s'inverse. Elles repartent de plus belle dans la direction opposée. Cette sensation de mouvement alors que tout est stable dans le tableau d'I. Levant est une illusion visuelle dont les neurophysiologistes de la perception sont friands.

Nous sommes victimes d'une illusion lorsque nous percevons une chose qui n'existe pas ou qui ne peut être perçue par le système visuel. Les illusions visuelles ne sont pas que des curiosités de laboratoire : nous vivons dans un monde visuel reconstruit par notre cerveau et la différence entre illusion et réalité n'est pas aussi tranchée. Par exemple, la nuit, sur une autoroute éclairée par les lampadaires au néon, les boues des voitures que vous dépassez semblent tourner à l'envers : vous vous souvenez alors des roues des diligences au cinéma qui tournent également à l'envers, et vous en déduisez que cette illusion est due à l'éclairage discontinu des lampadaires au néon. Pourtant, l'impression est là : la roue tourne lentement dans le sens inverse de son mouvement réel. Illusion ou réalité ?

À l'horizon, une Lune pleine paraît deux fois plus grosse que d'habitude. Que se passe-t-il ? La distance de la Terre à la Lune est constante, et la Lune n'est pas grossie par un phénomène optique. Seulement, notre système visuel juge la Lune à l'horizon plus lointaine que lorsqu'elle est au zénith. Là encore, nous avons beau nous persuader que la lune ne change pas de taille, nous voyons deux fois plus grosse. Illusion ou réalité ?

De ces exemples tirés de la vie courante, nous concluons qu'il s'agit d'illusions parce qu'il y a contradiction entre ce que nous observons et ce que nous savons. Toutefois, lorsqu'il n'y a pas de contradiction, nous nous accommodons

de cette perception illusoire. Par exemple, à l'aide d'un ordinateur ou de deux projecteurs de diapositives, présentez une tache de lumière de n'importe quelle forme qui s'éteint alors qu'une même tache de lumière s'allume quelques centimètres plus loin et quelques millièmes de seconde plus tard. Vous voyez un mouvement rapide de la tache de lumière entre les deux positions, rien ne bouge, il s'agit seulement de taches de lumière qui s'allument et qui s'éteignent ! Quand les deux taches de lumière sont de couleurs différentes, la première semble toujours se déplacer, mais en plus, elle change de couleur à mi-chemin. Cette illusion, qui passe inaperçue tant qu'il n'y a pas contradiction avec d'autres perceptions ou d'autres connaissances, est nommée mouvement apparent ou mouvement ϕ . Notre cerveau reconstruit un phénomène qui n'a rien à voir avec la réalité.

Ainsi, nous vivons dans un monde visuel qui n'est qu'une interprétation de la réalité construite par notre cerveau,

et particulièrement par le cortex visuel. Ce dernier compense les limites de la première étape du traitement des informations visuelles : l'œil. Bien que célébré par de nombreux philosophes et par Darwin pour l'ingéniosité de sa « conception », l'œil est un bien pauvre capteur : la vision des couleurs est précise que dans la partie centrale du champ visuel ; au-delà, c'est le flou le plus total.

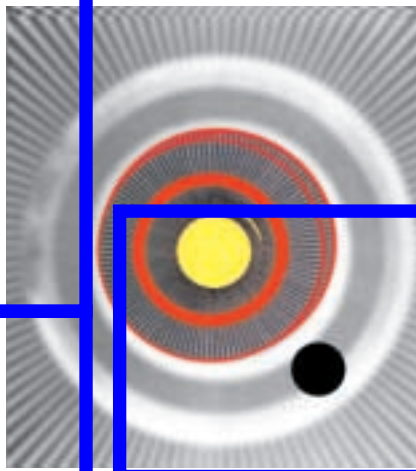
Par ailleurs, dans chaque œil, la rétine a une région aveugle, sans cellule sensible, qui correspond au point de convergence de toutes les cellules nerveuses de la rétine, et à l'origine du nerf optique.

Seule la région centrale de la rétine autorise la vision des couleurs (un disque dont le rayon est environ la moitié du rayon de la rétine), car, à la périphérie, les informations perçues par les cellules sont « mélangées » par les relais qui les véhiculent jusqu'au cerveau.

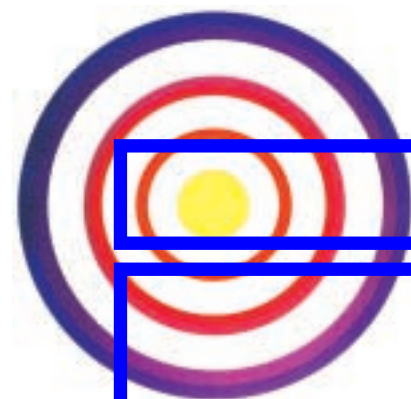
Enfin, l'œil n'est jamais stable : lorsque vous fixez une lettre sur cette page, l'image de cette lettre se déplace constamment sur la rétine à cause des micro-mouvements de l'œil ; l'amplitude de ces mouvements est de l'ordre de la taille de la lettre. De ces caractéristiques oculaires, nous pouvons reconstruire le tableau d'I. Levant tel que le voit notre œil sans reconstruction par le cerveau (voir la figure 1).

Toutefois, malgré les mouvements, la zone aveugle, l'acuité visuelle médiocre et la faible perception des couleurs, nous avons l'impression d'un monde visuel stable où tout est perçu avec la même acuité et avec un rendu parfait des couleurs.

L'illusion n'est donc pas l'exception, mais bien la norme : nous vivons dans une perception visuelle illusoire qui résulte de la reconstruction de la scène visuelle par le cortex visuel. Ce dernier est divisé en de nombreuses aires : la première, V1, reçoit ensemble des informations issues de la rétine que V2 répartit ensuite dans des aires secondaires spécialisées ; par exemple, V4 est dédiée aux formes et aux couleurs, tandis que V5 traite les informations relatives au mouvement et à la profondeur.



1. RECONSTITUTION D'ENIGMA tel que notre œil droit le voit (quand il est près du tableau) sans la reconstruction par notre cerveau : on y distingue la zone centrale où la vision est nette et colorée ; la zone où l'image est encore colorée, mais floue ; la zone où l'image perd sa couleur, la naissance du nerf optique (la tache noire, dite aveugle), et enfin, les mouvements saccadés de l'œil (la superposition d'images décalées, représentée par le double trait rouge et la portion centrale jaune).



2. L'ILLUSION D'ENIGMA (à droite) est due aux rayons noirs et blancs. Sans eux, les nuées colorées qui se déplacent le long des anneaux disparaissent (ci-dessus).

Quels mécanismes sous-tendent la reconstruction du monde visuel par notre cerveau ? Un exemple célèbre d'illusion utilisée par les neurophysiologistes pour expliquer ces mécanismes est le triangle de Kanisza (voir la figure à côté du titre) où trois disques noirs découpés créent l'impression illusoire de trois disques noirs entiers recouverts d'un triangle blanc : le système visuel interprète la complémentarité parfaite des coins découpés comme le résultat de la présence d'un triangle blanc posé sur trois disques noirs. On sait qu'en dehors de l'aire visuelle primaire, de nombreux neurones du cortex visuel sont activés par les bordures illusoires créées par ces stimuli. Il s'agit notamment de certaines sous-structures dans l'aire visuelle V1.

Nous avons récemment combiné l'illusion du triangle de Kanisza et le mouvement apparent et montré par des méthodes d'imagerie cérébrale chez l'homme que l'image active l'aire visuelle V5, mais aussi, de façon surprenante, l'aire V1, que l'on pensait dédiée aux informations «brutes». Or, le mouvement apparent et les contours illusoires sont des informations complexes : les informations traitées dans V5 retournent peut-être vers V1 pour y servir de référence.

L'intérêt du tableau d'I. Leviant réside dans la force de l'illusion de mouvement circulaire dans les anneaux alors que le tableau est parfaitement statique. Quel est la source de l'impression de mouvement ? Cette question a suscité un intéressant débat entre deux spécialistes britanniques de la vision, Semir Zeki et Richard Gregory. Ce dernier soutient que l'illusion est créée par le mouvement de l'image sur la rétine, tandis que S. Zeki pense que l'illusion résulte uniquement du traitement de l'information par le cortex visuel.

Tous deux ont partiellement raison. Nous l'avons rappelé, même lorsqu'on nous fixons intensément un objet, son image sur la rétine est en mouvement permanent, or quand on supprime ce mouvement de l'œil, par exemple, en solidifiant l'image à la rétine par des procédés optiques, toute perception visuelle disparaît. En d'autres termes, le mouvement est à la vision ce que l'oxygène est à la vie : sans mouvement, la vision n'existe pas. Ainsi, pour *Enigma* et pour les autres objets visuels, les micromouvements de l'image sur la rétine sont indispensables pour activer notre perception.

Cependant, le phénomène ne peut être purement rétinien : les micromouvements de l'œil ont une direction aléatoire ; ils n'expliquent donc pas à eux seuls que les nuées mouvantes perçues dans *Enigma* sont restreintes aux anneaux colorés et se déplacent de façon circulaire. Il s'agit sans doute d'un phénomène cortical, comme le suggèrent les résultats d'imagerie de S. Zeki. On sait que les rayons noirs et blancs du tableau activent fortement les neurones du cortex visuel, en particulier ceux des aires V1 et V2. Par ailleurs, certains neurones de V2 sont sensibles aux seules extrémités, ici, celles des rayons noirs et blancs

près des anneaux de couleur. Ces neurones participent à l'analyse d'objets d'une scène observée. Cette intense activité aléatoire est vraisemblablement réorganisée au niveau des aires corticales spécialisées dans le traitement de l'information du mouvement (V5) pour donner lieu à cette impression de mouvement continu dans les anneaux colorés : ainsi, un mouvement perçu au niveau des rayons est reporté aux anneaux.

Enfin, une fois, le cortex visuel recrée une impression cohérente à partir d'informations incomplètes ou incohérentes. Toutefois, tout n'est qu'une hypothèse et la compréhension des mécanismes sous-jacents pour expliquer ce terme d'activité neuronale cette perception cohérente de mouvement circulaire reste un défi pour les neurophysiologistes de la vision. Il y a encore un mystère derrière *Enigma*.

Jean BULLIER, Centre de recherche Cerveau et Cognition, Toulouse

Muriel BOUCART, Marie-Anne HENAFF et Catherine BELIN, *Vision : aspects perceptifs et cognitifs*, SOLAL, 1998.

Jacques NINIO, *La science des illusions*, Odile Jacob, 1998.

Paléontologie

Le paléontologue et l'évolution

Pascal Tassy. Le Pommier, 2000
(168 pages, 85 francs).

Pascal Tassy écrit justement qu'il existe entre paléontologie et évolution «une liaison passionnelle». Cependant, tous les paléontologues ne vivent pas cette liaison de la même façon : la lecture des nombreux livres que des spécialistes éminents, au fil des années, ont consacrés à cette question lui montre. Selon P. Tassy, le paléontologue volturait dans l'évolution, la reconstitution phylogénétique. Une bonne partie de son livre, son cours pourrait-on dire, est donc consacrée à un exposé des méthodes employées par les tenants de la systématique phylogénétique, plus connus sous le nom de cladistes, pour reconstituer les liens de parenté entre les espèces et les autres groupes taxinomiques.

Rassurons-nous : contrairement à ce qu'annonce l'étrange introduction générale à la collection dont fait partie cet ouvrage, il n'est nullement nécessaire de lire «crayon et cahier de notes à l'appui» : P. Tassy écrit dans un style clair, et le glossaire final définit les termes qui paraîtraient abscons à certains lecteurs peu familiers des sciences naturelles.

L'auteur ne prétend nullement broser ici un tableau complet de la paléontologie, et il fait même un inventaire rapide de diverses approches intéressantes qu'il n'a pas eu la place d'aborder, de la paléoclimatologie à l'anatomie fonctionnelle. Il rappelle aussi que «rien ne serait possible sans la découverte de fossiles» et, donc, dans le travail de recherche sur le terrain, qui est la clé des progrès futurs de la paléontologie. Cependant, le propos essentiel du livre est bien celui qu'indique son titre : l'évolution. Sans doute serait-il plus exact de dire «la phylogénie», car les deux termes ne sont pas synonymes. P. Tassy est un brillant défenseur de la méthode cladistique de reconstitution phylogénétique, et les phylogénies ainsi obtenues (comme celles des divers groupes de mammifères qu'il prend pour exemples) sont, pour lui, le but ultime du paléontologue préoccupé d'évolution. Toutefois, se contenter de reconstituer des phylogénies pour comprendre l'évolution des organismes, n'est-ce pas un peu comme étudier l'histoire humaine en se limitant aux généalogies des familles princières, en ne se souciant

ni de l'évolution des conditions socio-économiques, ni des conflits et des guerres ?

L'évolution telle que la présente P. Tassy est une suite de transformations anatomiques qui se produisent dans une sorte de vide. Tout au plus fait-il allusion aux déplacements continentaux et à leur influence sur la phylogénèse. Or, il ne fait aucun doute que l'histoire des êtres vivants est intimement liée à celle des milieux où ils vivent. La cladistique ne nous renseigne pas à ce sujet, mais est-ce une raison pour l'ignorer ? La gêne qu'éprouve P. Tassy à l'aborder est particulièrement claire lorsqu'il traite des extinctions en masse. On le sent assez peu familier de cette question, qu'il s'efforce de minimiser. Il suffit pourtant de parcourir les colonnes des grandes revues scientifiques pour comprendre l'importance qu'ont pris les débats sur les extinctions dans une communauté scientifique qui dépasse de loin le milieu paléontologique. Pour P. Tassy, cependant, «pour les paléontologues, les explications invoquant des événements instantanés — astéroïdes, comètes — fleurissent le catastrophisme dont on a vu qu'il était, au siècle dernier, le corollaire d'une vision fixiste de la vie et lié aux balbutiements d'une paléontologie en train de se constituer».

Autrement dit, l'étude des extinctions en masse serait un retour en arrière, à des conceptions périmées et pré- ou anti-évolutionnistes. L'argument est faible, car les chercheurs modernes qui ne craignent pas de se réclamer du catastrophisme n'en renient pas pour autant la théorie moderne de l'évolution. On peut admirer les intuitions de Cuvier en la matière sans partager son fixisme ; on peut voir en Jean-Baptiste Lamarck (1744-1829) un précurseur en matière d'évolution sans pour autant croire à l'hérédité des caractères acquis. D'ailleurs, il est assez surprenant de trouver, sous la plume de P. Tassy, une présentation erronée de certaines conceptions de Georges Cuvier (1769-1832), lorsqu'il écrit que celui-ci «émit l'hypothèse d'une succession de déluges anéantisant la vie sur terre, faune et flore étant recrées après chaque bouleversement». Il suffit de relire le *Discours sur les révolutions de la surface du globe* pour voir que Cuvier, contrairement à d'autres catastrophistes comme Alcide d'Orbigny (1802-1857), n'envisageait pas ces anéantisements complets et ces créations successives.

Pourquoi cette répugnance à prendre au sérieux le catastrophisme moderne, qui bouleverse certes notre vision de l'histoire du monde vivant, mais ne s'oppose en rien au cladisme cher à

P. Tassy ? Ce dernier l'explique peut-être lorsqu'il évoque les «relations ambiguës, sinon conflictuelles» entre la systématique cladistique et la paléontologie. Pour lui, «les raisons en sont d'abord d'ordre sociologique. La rupture avec les méthodes traditionnelles en paléontologie ne découle pas de ses propres évolutions épistémologiques, mais d'une autre science. Pour un paléontologue, un tel constat est toujours difficile à admettre». On peut aisément transposer cela aux relations tout aussi ambiguës et conflictuelles qu'entretiennent certains paléontologues avec les approches catastrophistes d'aujourd'hui. Celles-ci ont eu pour point de départ d'autres disciplines scientifiques, dont la géochimie, la physique et l'astrophysique. Elles n'en transforment pas moins la façon dont les paléontologues conçoivent aujourd'hui l'évolution. P. Tassy, et c'est bien compréhensible, pense comme certains sur l'aspect de la paléontologie évolutive qui l'intéresse le plus, la phylogénie ; il le fait avec talent, mais je regrette qu'il livre par la même occasion un combat d'arrière-garde contre des approches nouvelles qui ont replacé la paléontologie au cœur d'un débat scientifique impliquant nombre d'autres disciplines.

Eric BUFFETAUT

Éthologie

Les sociétés animales Évolution de la coopération et organisation sociale

Serge Aron et Luc Passera. Editions De Boeck, 2000 (336 pages, 320 francs).

En 1973, trois biologistes ont partagé le prix Nobel de médecine : Konrad Lorenz (1903-1989) étudiait le comportement social de la célèbre photographie où il était suivi par une ribambelle d'oisons ; Karl von Frisch (1886-1982) a commencé à décrypter la mystérieuse danse des abeilles ; et Nikolaas Tinbergen (1907-1988) s'est intéressé aux comportements instinctifs des animaux. Ils ont développé l'éthologie, ou science du comportement animal.

Où en est aujourd'hui la discipline ? Le behaviorisme (observation du comportement de l'extérieur sans recours aux explications physiologiques) est passé

de mode, l'éthologie s'est diversifiée, est devenue multidisciplinaire, les avis se nuancent. Les certitudes sont abandonnées, et les remises en question s'enchaînent, mais on se rend mal compte de cette évolution en France, car l'éthologie est peu présente dans les cursus universitaires, voire rognée dans les nouvelles maquettes de biologie. Les livres d'éthologie écrites sont rares, on s'en convainc en soumettant le mot clé «éthologie» aux moteurs de recherche sur Internet (de 10 à 30 références, avec peu d'auteurs : le choix est limité, et certains ouvrages sont dépassés). Quelle différence avec le résultat que l'on obtient quand on cherche les sites indexés par le mot *ethology* (plusieurs centaines de résultats)!

De ce fait, la publication du livre de Serge Aron et Luc Passera est une vraie bonne nouvelle. Les auteurs sont des spécialistes reconnus des insectes sociaux et dans *Les sociétés animales*, ils traitent de l'évolution sociale dans le cadre de la sélection naturelle et intègrent dans leur démarche l'évolution, l'écologie et l'éthologie.

Ils démontrent magistralement que les concepts sont faits pour être ébranlés, que les idées reçues n'existent que pour être balayées. La vie sociale des animaux n'est pas synonyme de vie en groupe, et le nombre de stades séparant l'animal solitaire de l'animal eusocial (mode de vie sociale le plus élaboré) : il y a l'animal grégaire (interattraction entre congénères), subsocial (apparition des comportements parentaux), colonial (existence d'un site d'élevage commun des jeunes, mais chaque femelle travaille pour elle-même), communal (les femelles coopèrent aux soins des jeunes).

L'eusocialité, beaucoup étudiée par le sociobiologiste Edward Wilson, semblait un concept simple. Les insectes eusociaux se trouvent dans l'ordre des Isoptères (termites) et dans celui des Hyménoptères (fourmis, abeilles et guêpes), mais il existe aussi quelques exemples de vertébrés eusociaux, dont les rats-taupes (*Heterocephalus glaber*, famille des Bathyergidés). Les critères de l'eusocialité semblaient particulièrement stricts : une coopération dans les soins aux jeunes et le chevauchement des générations (les descendants aident leurs parents), et des individus spécialisés dans la reproduction, tandis que d'autres, plus ou moins stériles (de façon définitive ou temporaire), se chargent des tâches communautaires en se les répartissant (polyéthisme). Le secret de l'altruisme de certains membres de la société qui sacrifient leur sexualité à un petit nombre d'individus chargés de les repré-

senter est leur apparemment, c'est la sélection de la parentèle.

Pourtant certaines fourmis (*Pristomyrmex pungens*) n'ont pas de reine et se reproduisent toutes par parthénogenèse thélytoque (ne donnant que des femelles). Malgré leurs particularités, il semble assez ridicule d'écarter ces fourmis japonaises des insectes eusociaux. D'autres exemples comme les chiens de prairie (*Cynomys ludovicianus*) pratiquent une vie sociale évoluée, mais tous les individus d'une coterie (8 à 12 individus) se reproduisent : la coopération dans l'élevage des jeunes et l'existence d'un nid collectif semblent ici des caractères essentiels de leur socialité. On admettrait aujourd'hui la distinction entre des espèces primitivement eusociales (celles où toutes les femelles se reproduiraient) et les espèces hautement eusociales (avec spécialisation dans la reproduction). Le bloc de l'eusocialité s'est scindé en deux.

On s'est beaucoup émerveillé du comportement altruiste des animaux qui se sacrifient pour leurs parents, telles les mangoustes (*Suricata suricatta*) qui vivent dans des régions désertiques du Sud de l'Afrique. On connaît leur silhouette dressée de sentinelles perchées sur un monticule, prêtes à alerter leurs congénères (occupés à la recherche de nourriture) dès qu'un prédateur menace. Ces vaillants gardiens affamés et vigilants semblaient être les symboles du développement. En fait, Clutton-Brock et ses collègues ont récemment dévoilé, dans la revue *Science* (1999), après plus de 2 000 heures d'observation, que les sentinelles ne montaient la garde que rassasiées, qu'elles étaient les premières à regagner le terrier en cas de danger et qu'elles échappaient au danger ! Beaucoup d'animaux prétendument altruistes sont de parfaits égoïstes !

Même si la vie sociale est, dans l'ensemble, avantageuse, ses inconvénients sont indéniables. Les bénéfices sont un meilleur repérage des prédateurs, une défense collective, la recherche de nourriture et l'élevage des jeunes en commun. Les inconvénients sont la concurrence alimentaire ou reproductive, les risques de maladies, le parasitisme social (certains profitant des autres sans restituer une contrepartie équivalente). Autrement dit, les sociétés animales ne sont pas idéales ; dans les deux cas, de multiples conflits sont possibles.

L'ouvrage est remarquablement construit, didactique dans le bon sens du terme, toujours passionnant, jamais ennuyeux, et la lecture en est aussi agréable pour le spécialiste que pour l'amateur.

Michèle FEBVRE

Chimie

Chimie l'art de se poser les bonnes questions

J. Garattot, T. Overton et T. Threlfall.
Éditions De Boeck, 2000
(182 pages, 130 francs).

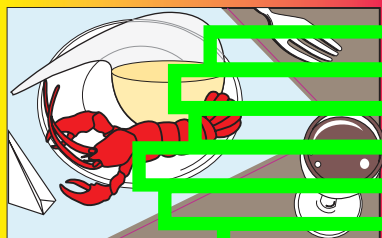
De nombreux livres d'enseignement ont bien faits... et très convenus : on salue alors une présentation remodelée de la discipline, une organisation interne nouvelle, un graphisme perfectionné. Toutefois, les livres d'enseignement véritablement originaux sont rares, en physique, par exemple, le cours de Richard Feynman reste isolé, et le livre de Lubert Stryer, en biologie moléculaire, a fait époque, tout comme la *Chimie physique* de Peter Atkins, en chimie physique, ou encore les traités de Sakurai, en mathématiques.

Posons aujourd'hui une pierre blanche, car *Chimie l'art de se poser les bonnes questions*, avec son titre ambigu (la chimie a-t-elle l'apanage d'être la discipline où l'on se pose de bonnes questions, ou bien est-ce ce livre qui fait poser les bonnes questions en chimie ?), est une remarquable innovation pédagogique.

Que l'on en juge : le sous-titre, qui décrit le contenu de l'ouvrage, est «Décryptage critique d'un texte scientifique», et les chapitres sont successivement «Comprendre une argumentation», «Construire une argumentation», «Lire de manière critique», «Se faire une opinion personnelle». L'objectif constant : enseigner aux étudiants comment réfléchir et non à quoi réfléchir. On l'a compris, les auteurs luttent contre le cliché selon lequel la chimie serait «une accumulation de constatations expérimentales disparates». Ils luttent à l'aide d'exercices (une trentaine par chapitres) dont la solution n'est donnée qu'une fois sur deux, en fin de document.

Examinons-en un, tiré du chapitre deuxième. L'énoncé donne trois propositions : (A) La synthèse d'un acide phosphorique de qualité engrais (c'est-à-dire impur) est moins onéreuse que celle de l'acide phosphorique pur ; (B) La synthèse de l'acide phosphorique pur implique plusieurs étapes incluant la calcination du phosphore blanc ; (C) L'acide phosphorique de qualité engrais peut être synthétisé en une seule étape, étant donné qu'il ne doit pas être d'une grande pureté. Le livre demande aux étudiants de former une «argumentation», c'est-à-dire un agencement séquentiel d'idées qui aboutissent à une conclusion logique, proposant des structures du type «X et Y, et dès lors Z»,

FRANCE INFO

et
POUR LA SCIENCEvous invitent
à écouter la chronique
de Marie-Odile MonchicourtInfo
SciencesTous les jours
sur **France Info**
à 15 heures 41,
17 heures 10,
20 heures 12,
22 heures 12,
et 23 heures 42Les prochains
résultats présentés
dans la rubrique
Science et Gastronomie
seront annoncés sur
France Info
le 27 décembre 2000.ou «X, Y, ce qui illustre ou démontre Z»
ou «X, Y, ce qui explique pourquoi Z».

L'exemple précédent est donné sans solution, car la réflexion pure conduit finalement à la solution, mais, pour d'autres cas, une bonne culture générale en chimie s'impose. Par exemple, le problème dont l'énoncé est «Plusieurs fois chaque année, des fidèles accourent nombreux vers la cathédrale de Naples pour être les «témoins de la liquéfaction miraculeuse du sang de saint Janvier». Un flacon à moitié rempli d'une substance rouge ressemblant à du sang leur est présenté. Au cours de la cérémonie, ledit flacon est retourné plusieurs fois de bas en haut, et l'officiant examine si le sang s'est liquéfié. Cette liquéfaction peut se produire au bout de quelques minutes, ou après des heures, voire des jours, et ce, quelle que soit la période de l'année, été ou hiver. Laquelle, parmi les assertions suivantes, fournit la meilleure hypothèse permettant d'expliquer ce phénomène? A : Il s'agit d'un véritable miracle. B : Le contenu du flacon est photosensible et se liquéfie lors de son exposition à la lumière. C : La liquéfaction et la resolidification du contenu du flacon résultent du développement périodique d'un micro-organisme. D : La substance est mixotrope, ce qui signifie qu'elle passe de l'état solide à l'état liquide sous l'effet d'une certaine agitation. E : La substance est un solide hygroscopique, déliquescent, qui devient liquide lorsqu'il absorbe l'humidité de l'air.

Cette fois, la discussion domine la solution (nommée «commentaire») apporte plus que la seule résolution de l'exercice.

Enfin, il y a aussi les cas beaucoup plus techniques, telle l'analyse de la réaction de Pictet-Spengler (qui fait intervenir une molécule organique contenant un cycle à cinq atomes, dont quatre de carbone et un d'azote), où une connaissance ciblée s'impose.

Nous commençons l'évocation de ce livre en mentionnant son originalité : elle est patente. Malgré le nombre de pages réduit, les questions du document sont grandes (la valeur n'attend pas le nombre des pages) et la traduction est généralement bonne. Un regret, à ce propos, toutefois : quelques belgicismes subsistent. Quel étudiant français comprendra la phrase «laquelle, parmi les assertions suivantes, détermine le plus le mécanisme proposé»? Déforce? Et une couche étherée est-elle une couche d'éther?

Malgré ces défauts véniels, le livre me semble hautement recommandable, car, quel que soit l'exercice, qu'il impose une connaissance chimique préalable ou qu'il ne demande qu'un peu de réflexion, il conduit à une connaissance solide, parce que raisonnée, de la chimie.

Hervé THIS

Agronomie

Le défi paysan

Luc Guyau. Éditions Le Cherche-Midi
(152 pages, 78 francs), 2000.

L'agriculture et l'alimentation sont aujourd'hui au cœur de nombreux débats de société qui, de toute évidence, concernent la recherche, agricole d'abord, mais aussi la recherche tout court : que l'on songe aux interrogations concernant les maladies à prions, les organismes génétiquement modifiés (OGM) et, même, sur un plan économique, les discussions de l'Organisation mondiale du commerce.

Il se trouve, et c'est une chance, que les agriculteurs passent progressivement d'une tradition marquée par la pratique à la réflexion. Luc Guyau, président de la puissante Fédération nationale des syndicats d'exploitants agricoles (FNSEA), paie, à cet égard, de sa personne. Il propose aujourd'hui, aux éditions Le Cherche-Midi, un ouvrage intitulé *Le défi paysan*. Je ne saurais trop conseiller la lecture de cet ouvrage à toute personne soucieuse de saisir, sur le fond, la question agricole.

L. Guyau s'adresse, en fait, de l'intérieur, aux agriculteurs. Dans un style direct et plaisant, il n'évite aucune question gênante. On est rapidement séduit par la prise en compte qui est la sienne des vrais problèmes des consommateurs et des citoyens : sécurité sanitaire, maîtrise de l'environnement, équité économique.

L'apport ainsi formulé à l'adresse des agriculteurs est responsable et tonique. L'idée que se fait L. Guyau du progrès technique est également très équilibrée et – pourquoi ne pas le dire? – saine. Il refuse tout retour en arrière, et milite pour une précaution responsable plutôt que culpabilisatrice.

Honnête qui pourrait être un peu accroché aux avis des militants de base de la FNSEA et les développements un peu techniques sur le devenir de la politique agricole commune et l'Organisation mondiale du commerce, *Le défi paysan* est, à coup sûr, un livre de référence pour tous ceux qui s'intéressent à l'agriculture, à l'alimentation, aux relations de l'homme «moderne» avec la nature. La sincérité et la simplicité des propos de L. Guyau donnent, de surcroît, à son ouvrage beaucoup de charme et d'authenticité.

Guy PAILLOTIN

La liste de tous les livres reçus en service de presse par la rédaction de *Pour la Science* est donnée sur notre site Internet, à l'adresse : <http://www.pourlascience.com>

index 2000

159

INFINI EST UN RÉVÉLATEUR (*L'*), par Patrick Dehornoy, n° 278, p. 138.

INFINI ET L'UNIVERS DES ALGORITHMES (*L'*), par Gilles Dowek, n° 278, p. 108.

INFINI, PIERRE DE TOUCHE DU CONSTRUCTIVISME (*L'*), par Allan Calder, n° 278, p. 74.

INTELLIGENCE EN ESSAIM (*L'*), par Éric Bonabeau et Guy Théraulaz, n° 271, p. 66.

LOGIQUE ET CALCULS DE LA TÉLÉPORTATION, par Jean-P. Delahaye, n° 272, p. 28.

MATÉRIAUX NUMÉRIQUES ET DÉGRADATION VIRTUELLE, par J. Dorsey et P. Hanrahan, n° 270, p. 76.

MATHÉMATIQUES, S'IL VOUS PLAIT, par Javier de Lorenz, n° 278, p. 24.

MÉMOIRE DES DISQUES DURS (*La*), par Jon Teigo, n° 273, p. 66.

ORDINATEURS MOLÉCULAIRES (*Les*), par Mark Reed et James Tour, n° 274, p. 78.

PROMESSES ET LES RISQUES DU WAP (*Les*), par Karen Bannan, n° 277, p. 34.

PHYSIQUE – CHIMIE

À LA RECHERCHE DE L'ÎLOT SUPERLOURD, par I. Oganessian, V. Utyonkov et K. Moody, n° 269, p. 60.

ÉNERGIE NÉGATIVE (*L'*), par Lawrence Ford et Thomas Roman, n° 271, p. 74.

GRAND COLLISIONNEUR DE HADRONS (*Le*), par Chris Llewellyn Smith, n° 275, p. 74.

HYDROGÈNE MÉTALLIQUE (*L'*), par William Nellis, n° 273, p. 84.

INFINIMENT GRAND (*L'*), par Gerhard Börner, n° 278, p. 120.

INFINIMENT PETIT EN PHYSIQUE (*L'*), par Harald Fritzsch, n° 278, p. 112.

INFINIMENT VIDE N'EXISTE PAS (*L'*), par Timothy Boyer, n° 278, p. 128.

LASERS À IMPULSIONS ULTRACOURTES, par John-Mark Hopkins et Wilson Sibbett, n° 277, p. 86.

LUMIÈRE EN CAGE (*La*), par Jean-Michel Lourtios et Daniel Maystre, n° 275, p. 66.

MATÉRIAUX NUMÉRIQUES ET DÉGRADATION VIRTUELLE, par J. Dorsey et P. Hanrahan, n° 270, p. 76.

MIROIRS À RETOURNEMENT TEMPOREL (*Les*), par Mathias Finelli, n° 268, p. 66.

PHYSIQUE DES AVALANCHES (*La*), par M. Naaim, F. Bouvet-Naaim et L. Vidal, n° 269, p. 52.

PLASTIQUES EXTRAITS DES PLANTES (*Des*), par Tillman Gerwig et Steven Slater, n° 276, p. 66.

QUAND L'EAU REFUSE DE GELER, par John Wettlaufer et Greg Dash, n° 270, p. 90.

SCIENCE AU XXI^e SIÈCLE (*La*), par John Maddox, n° 267, p. 26.

SCIENCE DU MOUVEMENT AU XVII^e SIÈCLE (*La*), par Michel Blay, n° 278, p. 58.

SOLIDES COULANTS (*Des*), par Ph. Coussot, H. van Damme et Ch. Ancey, n° 273, p. 84.

STOCKAGE DES GAZ À EFFET DE SERRE (*Le*), par Howard Herzog, Bengt Eliasson et Olav Kaarstad, n° 271, p. 88.

SYNTHÈSE PAR COMBUSTION (*La*), par Arvind Varma, n° 276, p. 90.

TÉLÉPORTATION QUANTIQUE (*La*), par Anton Zeilinger, n° 272, p. 36.

VERS L'UNIFICATION DE LA PHYSIQUE, par Steven Weinberg, n° 267, p. 32.

SCIENCE HUMAINES

BANQUE DE L'ESPOIR (*La*), par Muhammad Yunus, n° 278, p. 72.

COMMENT RATIONALISER LES DÉCISIONS, par J. Swets, R. Dawes et J. Monahan, n° 277, p. 52.

CULTURE ADAPTE LA SURPOPULATION (*La*), par F. de Waal, E. Aureli et P. Judge, n° 273, p. 60.

EXPANSION ET RECU DES FORÊTS ÉQUATORIALES, par Dominique Schwartz, n° 271, p. 58.

GRAPHOLOGIE, UN MAL FRANÇAIS (*La*), par Mari-Lou Bruchon-Schweitzer, n° 268, p. 60.

INFINI LITTÉRAIRE (*L'*), par P. Braffort, n° 278, p. 148.

L'ARQUE (*Les*), par Mario Rossi et Évelyne Peter-Defare, n° 272, p. 62.

MODIFICATIONS DU CLIMAT, par Thomas Karl et Kevin Trenberth, n° 267, p. 64.

MORALE JUSQU'AU CIEL (*Une*), par Sarah Simonson, n° 271, p. 48.

OUI, LE RÉCHAUFFEMENT DE LA PLANÈTE EST DANGEREUX, par Paul Epstein, n° 276, p. 80.

SÉLECTION NATURELLE DE LA CULTURE (*Une*), par Susan Blackmore, n° 277, p. 66.

SIX MILLIARDS D'HOMMES SUR TERRE, par Malcolm Potts, n° 269, p. 44.

TECHNIQUE

BALLON QUI A FAIT LE TOUR DU MONDE (*Le*), par Phil Scott, n° 269, p. 40.

BALLON DANS L'ESPACE (*Des*), par Steve Smith et James Cunniff, n° 269, p. 26.

CES SI DÉLICATS INSTRUMENTS DE VERRE, par Jean-Claude Chapuis, n° 272, p. 68.

PONTS LÉGERS DE ROBERT MAJELART (*Les*), par David Bellington, n° 275, p. 58.

PROMESSES ET LES RISQUES DU WAP (*Les*), par Karen Bannan, n° 277, p. 34.

SAUT DE LA TROISIÈME GÉNÉRATION (*Le*), par Leander Kahney, n° 277, p. 40.

STOCKAGE DES GAZ À EFFET DE SERRE (*Le*), par H. Herzog, B. Eliasson et O. Kaarstad, n° 271, p. 88.

ZEPPÉLINS DU XXI^e SIÈCLE (*Les*), par Klaus Hagenlocher, n° 268, p. 34.

ZOOLOGIE – BOTANIQUE

INTELLIGENCE EN ESSAIM (*L'*), par Éric Bonabeau et Guy Théraulaz, n° 271, p. 66.

MYXINE, CAROIGNARDS OCÉANIQUES (*Les*), par Frédéric Martini, n° 273, p. 76.

REINES ET CLAVAGISTES (*Les*), par Howard Topoff, n° 268, p. 74.

LOGIQUE ET PHILOSOPHIE

LA CRYPTOGRAPHIE RSA VINGT ANS APRÈS, par Jean-Paul Delahaye, n° 267, p. 104.

RACCOURCIS DANS LES DÉMONSTRATIONS, par J.-P. Delahaye, n° 268, p. 96.

LE DILEMME DU RENVOI D'ASCENSEUR, par J.-P. Delahaye et Philippe Mathieu, n° 269, p. 102.

LE TATOUAGE DES IMAGES NUMÉRIQUES, par Caroline Fontaine, n° 270, p. 102.

DES NOMS BRÉSIL À LA LETTRE..., par J.-P. Delahaye, n° 271, p. 104.

LA MUSIQUE EN IMAGES, par Christoph Oppé, n° 273, p. 98.

LA MISE EN PIÈCE DE LA MATHÉMATIQUE, par J.-P. Delahaye, n° 274, p. 96.

NUMÉROLOGIE ET COINCIDENCES, par J.-P. Delahaye, n° 275, p. 100.

L'ŒUF ET LA PHÈRE, par François Apéry, n° 276, p. 100.

MATHÉMATIQUES ET PHILOSOPHIE, par J.-P. Delahaye, n° 277, p. 100.

VISIONS MATHÉMATIQUES PAR IAN STEWART

L'EMPIRE DE LA CONSTANTE UN, n° 267, p. 100.

LES THÉORÈMES DE L'IMPOSSIBLE, n° 268, p. 92.

SCULPTURES RÉELLES ET VIRTUELLES, n° 269, p. 100.

LA STRATÉGIE DESSOUS-ENSEMBLES, n° 270, p. 100.

LE TROU NOIR DU SOLEIL, n° 271, p. 100.

SIMPLES D'AVÈS, n° 272, p. 106.

LA LOGIQUE DES PIRATES FÉROCES, n° 273, p. 96.

PARADOXES PERDUS ET RECHERCHÉS, n° 274, p. 96.

LA TOPOLOGIE MARINE, n° 275, p. 106.

UN GUIDE FRACTAL DU JEU DE MORPION, n° 276, p. 104.

AUTEURS

A – B

ASKEW Neville, *La conservation de la tombe de Majartari*, n° 270, p. 42.

ATKINS Puzos, *Un autobus entre les planètes*, n° 271, p. 44.

ALPERT L., *Comment aller sur Mars*, n° 271, p. 30.

ANDERSON Jasper, *Muscle, gènes et performances*, n° 276, p. 32.

ARKANI-HARMEY Nima, *Les dimensions cachées de l'univers*, n° 276, p. 56.

BACON Bruce, *Les défis de l'hépatite C*, n° 268, p. 38.

BANNAN Karen, *Les promesses et les risques du WAP*, n° 277, p. 34.

BECK Sarah, *Flambées d'étoiles dans les galaxies naines*, n° 274, p. 64.

BERGER Daniel, *L'infini en géométrie*, n° 278, p. 84.

BIGARD André-Xavier, *L'adaptation du muscle à l'entraînement*, n° 276, p. 42.

BIGATTI Daniela, *Les géométries non commutatives*, n° 278, p. 98.

BELLINGTON David, *Les ponts légers de Robert Majelart*, n° 275, p. 58.

BLACKMORE Susan, *La science du mouvement au XVIII^e siècle*, n° 278, p. 58.

BLAY Michel, *La science du mouvement au XVIII^e siècle*, n° 278, p. 58.

BONABEAU Éric, *L'intelligence en essaim*, n° 271, p. 66.

BÖRNER G., *L'infiniment grand*, n° 278, p. 120.

BORST Cornelius, *À cœur battant*, n° 277, p. 46.

BOUVET-NAIM Florence, *La physique des avalanches*, n° 278, p. 53.

BOYER Timothy, *L'infiniment vide n'existe pas*, n° 278, p. 128.

BRAFFORT Paul, *L'infini littéraire*, n° 278, p. 148.

BROWN Kathryn, *La lecture du génome humain*, n° 275, p. 40.

BRUCHON-SCHWEITZER Marilou, *La graphologie, un mal français*, n° 268, p. 60.

BUFFETAUT Eric, *Que nous apprendront les fossiles?*, n° 267, p. 52.

C – D

CALDER Allan, *L'infini, pierre de touche du constructivisme*, n° 278, p. 74.

CHAUSSIDON Marc, *La route des émeraudes anciennes*, n° 277, p. 58.

CHYBA Ch., *Une vie extraterrestre?*, n° 267, p. 82.

COLLINS F., *Le code de la vie déchiffré*, n° 267, p. 46.

CONNAN Jacques, *Les momies*, n° 270, p. 48.

CRAWFORD Ian, *Où sont-ils?*, n° 275, p. 24.

CUTTS J., *Des ballons dans l'espace*, n° 269, p. 26.

DAMASIO Antonio, *Le cerveau et l'esprit*, n° 267, p. 76.

DASH G., *Quand l'eau refuse de geler*, n° 270, p. 90.

DAWES Robyn, *Comment rationaliser les décisions*, n° 277, p. 52.

DEBENAY Jean-Louis, *Les foraminifères*, n° 276, p. 72.

DEHORNOY Patrick, *L'infini est-il nécessaire?*, n° 278, p. 100.

DEHORNOY Patrick, *L'infini est un révélateur*, n° 278, p. 138.

DELAHAYE Jean-Paul, *L'infini est-il paradoxal en mathématiques?*, n° 278, p. 30.

DELEDICQ André, *L'ensemble triadique de Cantor*, n° 278, p. 73.

DE WAAL Frans, *Inné contre acquis : la fin de l'opposition*, n° 277, p. 58.

DI BISCEGLIO Adrian, *Les défis de l'hépatite C*, n° 268, p. 38.

DIMOPOULOS Savas, *Les dimensions cachées de l'Univers*, n° 276, p. 56.

DOOLITTLE Ford, *L'arbre du vivant : un écosystème foisonnant*, n° 270, p. 84.

DORSEY Julie, *Matériaux numériques et dégradation virtuelle*, n° 270, p. 76.

DOWEK Gilles, *L'infini et l'univers des algorithmes*, n° 278, p. 100.

DVALI George, *Les dimensions cachées de l'Univers*, n° 276, p. 56.

E – F

ELIASSEN Bårdur, *Le stockage des gaz à effet de serre*, n° 271, p. 88.

ÉMILIE Dominique, *Cytokines et sensibilité aux maladies*, n° 270, p. 56.

EPSTEIN Paul, *Les défis de l'hépatite C*, n° 268, p. 38.

EZZELL Carol, *Au-delà du génome*, n° 275, p. 52.

FINK Mathias, *Les miroirs à retournement temporel*, n° 268, p. 66.

FORD Lawrence, *L'énergie négative*, n° 271, p. 74.

FRITZSCH Harald, *L'infiniment petit en physique*, n° 278, p. 112.

G – H

GALANAUD Pierre, *Cytokines et sensibilité aux maladies*, n° 270, p. 56.

GARDNER David, *Des plastiques extraits des plantes*, n° 276, p. 66.

GESLIN Jean-Louis, *Les foraminifères*, n° 276, p. 72.

GIORELLI Giorgio, *L'univers infini de Bruno*, n° 278, p. 54.

GIULIANI Gaston, *La route des émeraudes anciennes*, n° 277, p. 58.

GREELEY Robert, *L'océan caché d'Europe*, n° 268, p. 44.

HAGENBUCHER Klaus, *Les zeppelins du XXI^e siècle*, n° 269, p. 34.

HANRAHAN Pat, *Matériaux numériques et dégradation virtuelle*, n° 270, p. 76.

HARTENBERGER Jean-Louis, *La mort des mamifères des îles*, n° 276, p. 48.

HARVEY Gioia, *Votre futur compagnon numérique*, n° 277, p. 28.

HEAD James, *L'océan caché d'Europe*, n° 268, p. 44.

HERBER Jean-Marc, *Les anticoagulants de synthèse*, n° 274, p. 28.

HERZOG Howard, *Le stockage des gaz à effet de serre*, n° 271, p. 88.

HEUZÉ Michèle, *La route des émeraudes anciennes*, n° 277, p. 58.

HOFFMAN P., *Quand la Terre était glacée*, n° 267, p. 30.

HOLLOMAN M., *Des lacs maudits*, n° 275, p. 90.

HOPKINS Jean-Mark, *Lasers à impulsions ultracourtes*, n° 277, p. 86.

HOWARD Ken, *La bio-informatique*, n° 275, p. 46.

HUMBERT Marc, *Cytokines et sensibilité aux maladies*, n° 270, p. 56.

J – K

JEGALIAN K., *Le code de la vie déchiffré*, n° 267, p. 46.

JOHNSON Torrence, *La mission humaine sur Jupiter*, n° 270, p. 64.

KAARSTAD Glen, *Le stockage des gaz à effet de serre*, n° 271, p. 88.

KAHNEY Leander, *Le saut de la troisième génération*, n° 277, p. 40.

KARL Thomas, *Modifions-nous le climat?*, n° 267, p. 64.

KRAUSS John, *Le destin ultime de la vie*, n° 269, p. 80.

L – M

LARSEN Clark Spencer, *Le déclin des Indiens*, n° 274, p. 42.

LE GOFF Jean-Pierre, *De la perspective à l'infini géométrique*, n° 278, p. 66.

LÉVY Tony, *Thabit ibn Qurra et l'infini numérique*, n° 278, p. 48.

LOCOTHEIS Nikos, *La vision : une fenêtre sur la conscience*, n° 268, p. 80.

LORENZO MARTÍNEZ Javier de, *Les mathématiques, sciences de l'infini*, n° 278, p. 24.

LOURTIOZ J.-M., *La lumière en cage*, n° 275, p. 66.

MACKE André, *Les momies*, n° 270, p. 48.

MACKE-RIBET Ch., *Les momies*, n° 270, p. 48.

MADDOX J., *La science au XXI^e siècle*, n° 267, p. 26.

POUR LA SCIENCE

8, rue Férou 75278 PARIS CEDEX 06

Tél. : 01-55-42-84-00

<http://www.pourlascience.com>

Commande de dossier ou de magazine :

01-55-42-40-16 ou 01-55-42-40-17

POUR LA SCIENCE Directeur de la rédaction : Philippe Boulanger. Françoise Cinotti (Rédactrice en chef), Philippe Pajot, Loïc Mangin, Claire Le Poulennec, François Savatier, Emmanuel Leclerc, Laurence Plévert, Marie-Neige Cordonnier (Rédacteurs). Secrétaire de rédaction/Maquette : Annie Tacquenet, Pascale Thiollier, Céline Lapert, Sylvie Sobelman. Site Internet : Cyril Lamotte. Marketing et Publicité : Stéphane Montouchet, assisté de Séverine Merviel et Christophe Vitiello. Direction financière : Anne Gusdorf. Direction du personnel : Jean-Benoît Dutry. Fabrication : Jérôme Jalabert, assisté de Géraldine Le Coz. Presse et communication : Sylvie Gillet, assistée de Nathalie Choplain. Directeur de la publication et Gérant : Olivier Brossollet. Conseiller scientifique : Hervé This.

Ont également collaboré à ce numéro : Marie-Bernadette Causse, Marie-Noëlle Celerier, Évelyne Host-Platret, Richard Kerner, Marc Lachièze-Rey, Ariane Lançon, Alexis Michel, François Moutou, Claude Olivier.

SCIENTIFIC AMERICAN Editor : John Rennie. Board of editors : Michelle Press, Mark Alpert, Carol Ezzell, Wayt Gibbs, Kristin Leutwyler, Madhusree Mukerjee, George Musser, Sasha Nemecek, Ricki Rusting, Sarah Simpson, Gary Stix, Philip Yam, Glenn Zorpette. Chairman Emeritus : John Hanley. Chairman : Rolf Gruebach. President and Chief Executive Officer : Gretchen Teichgraber. Vice-President : Chuck McCullough and Frances Newburg.

PUBLICITÉ France

Chef de Publicité : Jean-François Guillotin (jf.guillotin@pourlascience.fr), assisté d'Irène Limon 8 rue Férou 75278 Paris Cedex 06

Tél. : 01 55 42 84 28 ou 01 55 42 84 97

Télécopieur : 01 43 25 18 29

Étranger : 415 Madison Avenue, New York.

N.Y. 10017 - Tél. (212) 754.02.62

SERVICE ABONNEMENTS

Anne-Claire Thoinis : 01 55 42 84 04

SERVICE VENTE RÉSEAU NMPP

Christophe Vitiello : 01 55 42 84 81.

DIFFUSION DE LA BIBLIOTHÈQUE

POUR LA SCIENCE

Canada : Edipresse : 945, avenue Beaumont, Montréal, Québec, H3N 1W3 Canada.

Suisse : Servis : Chemin des châlets, 1979 Chavannes - 2 - Bâle

Belgique : La Caravelle : 303, rue du Pré-aux-oyes - 1130 Bruxelles.

Autres pays : Éditions Belin : 8, rue Férou - 75278 Paris Cedex 06.

Toutes demandes d'autorisation de reproduire, pour le public français ou francophone, les textes, les photos, les dessins ou les documents contenus dans la revue «Pour la Science», dans la revue «Scientific American», dans les livres édités par «Pour la Science» doivent être adressées par écrit à «Pour la Science S.A.R.L.», 8, rue Férou, 75278 Paris Cedex 06.

© Pour la Science S.A.R.L.

Tous droits de reproduction, de traduction, d'adaptation et de représentation réservés pour tous les pays. La marque et le nom commercial «Scientific American» sont la propriété de Scientific American Inc. Licence accordée à «Pour la Science S.A.R.L.» En application de la loi du 11 mars 1957, il est interdit de reproduire intégralement ou partiellement la présente revue sans autorisation de l'éditeur ou du Centre français de l'exploitation du droit de copie (20, rue des Grands-Augustins - 75006 Paris).

Nos lecteurs trouveront des bulletins d'abonnement en pages 18A, 18B, 146A et 147B.

MAEKWA Shin, *La conservation de la tombe de Néfertari*, n° 270, p. 42.

MARTZOFF C., *L'infini en Chine*, n° 278, p. 44.

MATIC Ivan, *SOS génome : réparation et évolution*, n° 269, p. 66.

MATTEI NESAY, *Autrefois, nous n'étions pas seuls*, n° 269, p. 74.

MAYR Ernst, *L'influence de Darwin sur la pensée moderne*, n° 275, p. 82.

MAYSTED Daniel, *La lumière en cage*, n° 275, p. 66.

MIDANÉ-REYNES Béatrix, *La naissance de l'État égyptien*, n° 270, p. 28.

MONAHAN John, *Comment rationaliser les décisions*, n° 277, p. 52.

MOOD Kerion, *À la recherche de l'îlot superlourd*, n° 269, p. 60.

MORAVEC Hans, *L'ère des robots*, n° 267, p. 88.

MUSSE G., *Comment aller sur Mars*, n° 271, p. 30.

MYSLIVECKI, *La découverte d'un vizir*, n° 270, p. 34.

N - O

NAAIM Mohamed, *La physique des avalanches*, n° 269, p. 52.

NEMECK S., *Les premiers Américains*, n° 277, p. 76.

OBERG, *Un autobus entre les planètes*, n° 271, p. 44.

OGANESIAN Iouri, *À la recherche de l'îlot superlourd*, n° 269, p. 60.

ÖHMAN Arr, *La biologie des phobies*, n° 271, p. 50.

P - R

PAPPAARD Robert, *L'océan caché d'Europe*, n° 268, p. 44.

PAWSON Tony, *La communication dans les cellules*, n° 274, p. 50.

PENA S., *L'origine des Amérindiens*, n° 277, p. 82.

PETIT J Maurice, *Les anticoagulants de synthèse*, n° 274, p. 28.

POTTS Malcolm, *Six milliards d'hommes sur Terre*, n° 269, p. 44.

RADMAN Miroslav, *SOS génome : réparation et évolution*, n° 269, p. 66.

REED M., *Les ordinateurs moléculaires*, n° 271, p. 78.

REES M., *L'exploration de l'Univers*, n° 267, p. 40.

REMÉNI Maria, *Histoire d'infini*, n° 278, p. 64.

RODIER P., *Les origines de l'autisme*, n° 269, p. 88.

ROMAN Thomas, *L'énergie négative*, n° 271, p. 74.

ROSE Michel, *Le vieillissement est-il inéluctable?*, n° 269, p. 70.

S - T

SALAMINI Francesco, *La première céréale cultivée*, n° 278, p. 92.

SALTIN Bengt, *Muscle, gènes et performances*, n° 276, p. 32.

SANTOS F., *L'origine des Amérindiens*, n° 277, p. 82.

SCHJERLING Peter, *Muscle, gènes et performances*, n° 276, p. 32.

SCHRAG D., *Quand la Terre était gelée*, n° 268, p. 30.

SCHWARTZ Dominique, *Expansion et recul des forêts équatoriales*, n° 271, p. 58.

SCOTT John, *La communication dans les cellules*, n° 274, p. 50.

SCOTT Phil, *Le ballon qui a fait le tour du monde*, n° 269, p. 40.

SIBBETT Wilson, *Lasers à impulsions ultracourtes*, n° 277, p. 86.

SIEGEL Jerome, *La narcolepsie*, n° 268, p. 54.

SIMPSON S., *In moral jusqu'au ciel*, n° 271, p. 48.

SINGER Fred, *Mars via ses lunes*, n° 271, p. 42.

SLATER Kevin, *Des plastiques extraits des plantes*, n° 276, p. 66.

SMITH Charles Llewellyn, *Le Grand Collisionneur de hadrons*, n° 275, p. 74.

SMITH S. D., *Ballons dans l'espace*, n° 269, p. 26.

STARKMAN C., *Le destin ultime de la vie*, n° 269, p. 80.

STREMER, *Les foraminifères*, n° 276, p. 72.

SWENSON George, *Les communications galactiques*, n° 275, p. 30.

SWETS John, *Comment rationaliser les décisions*, n° 277, p. 52.

TADDEI François, *SOS génome : réparation et évolution*, n° 269, p. 66.

TARTER Bill, *Une vie extraterrestre?*, n° 267, p. 82.

TATTERSALL Ian, *Autrefois, nous n'étions pas seuls*, n° 269, p. 74.

THÉRAU AZ, *Intelligence et conscience*, n° 277, p. 86.

TOPOFF L., *Les reines esclavagistes*, n° 268, p. 74.

TOUR J., *Les ordinateurs moléculaires*, n° 274, p. 78.

TRENBERTH Kevin, *Modifions-nous le climat?*, n° 267, p. 64.

U - V

UTYONKOV Vladimir, *À la recherche de l'îlot superlourd*, n° 269, p. 60.

VARDI Ilan, *Archimède face à l'infini*, n° 278, p. 40.

VARMA A., *La synthèse par combustion*, n° 276, p. 90.

VERGER Bernard, *Entre terre et mer : le Mont-Saint-Michel*, n° 271, p. 22.

VIDAL L., *La physique des avalanches*, n° 269, p. 52.

W - Y - Z

WEINBERG Steven, *Vers l'unification de la physique*, n° 267, p. 32.

WETTLAUER John, *Quand l'eau refuse de geler*, n° 278, p. 92.

WETTLAUER John, *Quand l'eau refuse de geler*, n° 278, p. 92.

WETTLAUER John, *Quand l'eau refuse de geler*, n° 278, p. 92.

WETTLAUER John, *Quand l'eau refuse de geler*, n° 278, p. 92.

WETTLAUER John, *Quand l'eau refuse de geler*, n° 278, p. 92.

WETTLAUER John, *Quand l'eau refuse de geler*, n° 278, p. 92.