



SCIENCE

édition française de

SCIENTIFIC
AMERICAN

La télévision numérique

Le chant des cigales

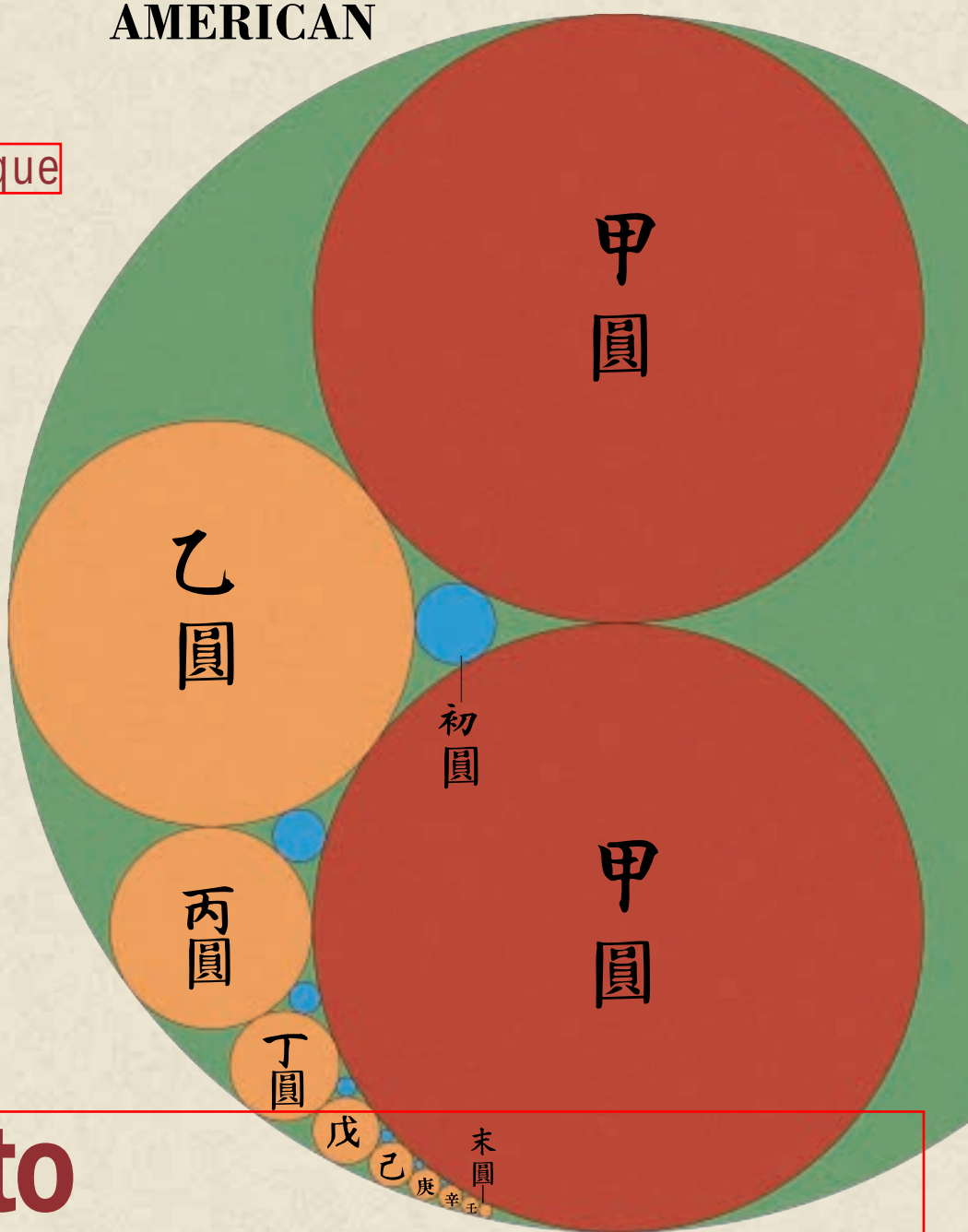
La contraception
masculine

Les écrans plats

La pollution
par les nitrates

Les ancêtres
des tortues

La physique
du mélange



Ex-voto mathématiques japonais



BLOC-NOTES

de Didier Nordon

5

JEU-CONCOURS

Quadrilatères

par Pierre Tougne

6

TRIBUNE DES LECTEURS



POINT DE VUE

Les Académies

par Gérard Toulouse

9



PRÉSENCE DE L'HISTOIRE

Premières images au microscope

par Brian Ford

10



SCIENCE ET ÉCONOMIE

La mondialisation

par Bernard Guerrien et Francisco Vergara

14

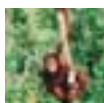


SCIENCE ET GASTRONOMIE

Jambons crus d'Espagne

par Hervé This

17

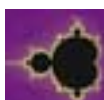


PERSPECTIVES SCIENTIFIQUES

- Lancements maritimes ■ Hello, le Soleil vibre...
- De Mars à Tatahouine ■ L'oreille absolue

20

- Kampala : la campagne à la ville ■ La masse du neutrino
- Né de père absent ■ Moussons inattendues
- Exploration virtuelle du côlon ■ Séisme en Afghanistan
- Crânes virtuels ■ Épiderme de culture
- Machiavel aurait-il pu être un singe?



LOGIQUE ET CALCUL

Certitudes sans démonstration

par Jean-Paul Delahaye

100



VISIONS MATHÉMATIQUES

Quelle coïncidence!

par Ian Stewart

106



L'IMAGE DU MOIS

Rater n'est pas fausser

109



ANALYSES DE LIVRES

- MAPLE V release 5
- Éloge du mixte, matériaux nouveaux et philosophie,

110

de Bernadette Bensaude-Vincent ■ Le guide du Paris savant, d'Anna Alter et Philippe Testard-Vaillant ■ Pour Darwin, sous la direction de Patrick Tort ■ Les avatars du gène : la théorie néo-darwinienne de l'évolution, de Pierre-Henri Gouyon, Jean-Pierre Henry et Jacques Arnould

La contraception masculine

34

par Sophie Christin-Maitre
et Philippe Bouchard

Le contraceptif masculin idéal devra être rapidement efficace, dépourvu d'effets secondaires et réversible. Il ne devra pas perturber l'activité sexuelle. Bien que ces exigences soient difficiles à concilier, plusieurs substances contraceptives sont prometteuses.



L'évolution des tortues

42

par G. Rougier,
M. de la Fuente et A. Arcucci

La récente découverte, en Argentine, d'étonnants spécimens fossiles précise l'origine des tortues et révèle une diversité écologique et systématique inattendue chez les Chéloniens.

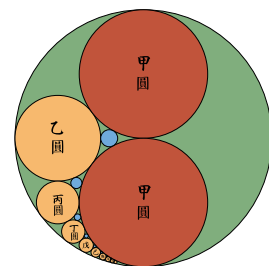


Géométrie et religion au Japon

52

par Tony Rothman
et Hidetoshi Fukagawa

Quand le Japon s'isola, entre 1639 et 1854, des pratiques mathématiques originales apparurent. Les mathématiciens amateurs dédiaient aux divinités des ex-voto géométriques.



Nitrate : un polluant de longue durée

60

par André Mariotti

La durée de résidence dans le sol de l'azote apporté par les engrais est très longue. Les nitrates qui polluent aujourd'hui les nappes aquifères résultent de nombreuses années de fertilisation.



Le chant des cigales

66

par Henry Bennet-Clark

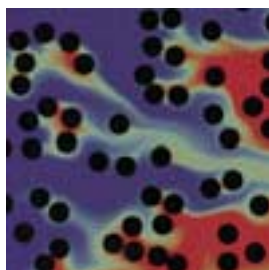
Les cigales mâles sont les plus bruyants des insectes. On a découvert comment fonctionne leur système acoustique.

**Mélange et désordre figé**

70

par Étienne Guyon
et Jean-Pierre Hulin

Deux liquides se mélangent quand ils traversent un solide poreux. Inversement, les solides poreux peuvent aussi séparer des mélanges. La physique explore les raisons de ces différences.

**La télévision de demain****Les écrans plats**

78

par Alan Sobel

Après des décennies de perfectionnements, les écrans à plasma sont devenus lumineux, de grande taille et si plats qu'on peut les accrocher au mur. Les ingénieurs cherchent maintenant à réduire leur prix.

La norme européenne de télévision numérique

86

par Jean-Pierre Hénot,
Vincent Michon et Bertrand Sueur

La nouvelle norme européenne de télévision numérique remplacera bientôt la norme de transmission analogique dans tous les équipements télévisuels.

La télévision numérique

92

par Jae Lim

Après des années de travail, les États-Unis ont déterminé la norme de télévision numérique qui sera imposée aux chaînes du pays.

É D I T O R I A L**Les propriétés des nombres...**

... apparaissent toujours paradoxales. Un segment est totalement «rempli» par les nombres fractionnaires : il existe un tel nombre aussi près que l'on veut d'un point du segment. Et pourtant, d'autres infinités de nombres «remplissent» encore mieux le segment : les nombres irrationnels (comme $\sqrt{2}$), les nombres transcendants (comme π)... Des infinités de nombres de divers types s'insèrent entre deux nombres infiniment proches. La jungle des nombres est diverse et extraordinairement dense.

Mais pas impénétrable. Des propriétés mathématiques adaptées guident la découverte de l'identité d'un nombre de la jungle. Ainsi, vous calculez par un moyen informatique un nombre et vous trouvez 7,8539... Hélas, une telle suite de décimales vous renseigne sur la valeur du nombre, pas sur sa nature. Pour identifier votre valeur numérique, (voir *Certitudes sans démonstration*, par Jean-Paul Delahaye, page 100) vous comparez votre nombre, par divers moyens mathématiques, aux 70 millions de nombres de la table élaborée par Simon Plouffe, et des indices vous persuadent que votre nombre est $\pi/4$. Victoire : le résultat de votre calcul est ce nombre transcendant et vous connaissez, par cette identification, sa valeur exacte. Peut-être, cela s'est vu, cette propriété amènera une autre découverte sur le processus physique qui a donné ce nombre.

En mathématiques, le passage de l'incompréhension à l'évidence est relativement facile. «N'importe quel imbécile peut comprendre les mathématiques, argumentait le grand logicien Alonzo Church ; il suffit d'être consciencieux, travailleur et patient. En revanche, découvrir une propriété est autrement difficile», soupirait-il, pensant certainement à ses présomptions de jeunesse. Les méthodes exposées dans l'article cité peuvent aider le processus de découverte.

Cela étant, certaines propriétés des nombres restent d'une mystérieuse simplicité, comme la loi de Benford. Comment est-il possible que les premiers chiffres des 70 millions de constantes de la liste de Simon Plouffe soient plus fréquemment des 1, moins souvent des 2, encore moins souvent des 3, etc.? Lisez la page 105.

Philippe BOULANGER

L'horreur du vide

«**S**a disparition laisse un grand vide parmi nous.» Par ces mots, l'orateur rend un dernier hommage au défunt.

Un grand vide? Est-ce là un compliment? Si j'ai bien compris la physique moderne, le vide est bourré d'animalcules quantiques qui, par myriades, fluctuent sans cesse, naissant, s'entretenant, disparaissant, renaissant... J'imagine que, dans un grand vide, il y a place pour plus d'animalcules que dans un petit. Autrement dit, un grand vide est moins vide qu'un petit. Ainsi, un petit vide est, en tant que vide, plus grand qu'un grand vide. CQFD.

Tout bien réfléchi, lorsque je prendrai possession de mon ultime logis, je serai heureux si, entre deux pelletées, j'entends l'orateur affirmer : «Sa mort laisse un vide minuscule parmi nous...»

Science et graphologie

Les graphologues prétendent déterminer si un scripteur est, ou non, intelligent, honnête, cultivé, original, etc. Les entreprises tirent parti des progrès de la graphologie et demandent que les lettres de candidature à un poste soient manus-

crites. Pourquoi les scientifiques ne profiteraient-ils pas, eux aussi, de ces progrès? Si les articles paraissaient sous forme manuscrite, le lecteur saurait, rien que par l'examen de l'écriture, quel crédit leur apporter. Plus facile.

«Et si l'écriture de l'auteur est illisible?» Eh bien, en ce cas, on ne lira pas son article. Cela ne changera pas grand-chose aux pratiques actuelles.

D'accord? Pas d'accord...

Une plaie pour l'orateur, c'est l'auditeur qui, à la fin de la conférence, déclare : «Je suis parfaitement d'accord avec tout ce que vous venez de dire.» Suite à quoi, en général, il aplatit les propos de l'orateur vers quelque lieu commun à la mode, ou bien les déforme jusqu'à ce qu'ils collent avec ses propres idées. Coincé par l'amabilité de l'auditeur, le malheureux orateur n'ose pas répondre en toute franchise, comme il le ferait face à un contradicteur déclaré.

Suis-je seul à ressentir ainsi les choses? Non. Un aphorisme de Louis Scutenaire exprime mon sentiment : «Si je vous donne raison, c'est dix fois sur dix que je n'ai pas compris ce que vous entendiez me dire.»

Chères mathématiques

Qu'une démonstration soit maladroite, excessivement longue, voire inutile – il n'en coûtait qu'au mathématicien, jusqu'à des temps récents. C'est lui qui avait perdu son temps, torturé pour rien sa cervelle – et voilà tout. Il payait seul. N'ayant de comptes à rendre à personne, il était en droit, royal, de mépriser les questions matérielles, il pouvait se permettre de ne «pas savoir compter». Tel n'est pas et n'a jamais été le cas des sciences de la nature. La cherté d'une expérience ou d'un procédé a toujours fait partie des paramètres que l'expérimentaliste doit prendre en considération.

Le mathématicien a perdu sa belle indifférence. Désormais, que ce soit pour démontrer des théorèmes ou pour étayer des conjectures, il recourt massivement (comme on dit) à l'ordinateur. Avec ce corollaire obligé : le devoir de limiter les dépenses. L'informatique a ramené les mathématiques dans le giron commun des soucis économiques. Cette nouvelle n'a rien pour réjouir.

Épistémologie

Quand un épistémologue étudie et interprète telle ou telle science, il est rare que les pratiquants de ladite science se reconnaissent, si par hasard ils lisent ses travaux. Ils sont, au mieux, sceptiques, et plus souvent fâchés : que de contresens sur eux, estiment-ils!

Qui voit juste? Où se trouve la vérité? Donner *a priori* raison au scientifique, qui agit au cœur du processus alors que l'épistémologue se contente d'observer et de commenter – c'est croire à la parfaite lucidité de l'homme sur ses actes. Une telle position n'est pas tenable. Donner *a priori* raison à l'épistémologue puisqu'il a du recul, c'est prendre le scientifique pour un idiot qui ne comprend rien à ce qu'il fait. Cette position n'est pas non plus tenable. Le plus probable, donc, est que les deux ont partiellement raison et partiellement tort.

Conclusion : pour «se faire une idée» sur une science, il faut à la fois l'étudier et lire ses commentateurs. Pourtant, si rude soit-elle, cette accumulation de tâches permet tout au plus d'élaborer une interprétation personnelle – laquelle, comme toute interprétation, est susceptible d'être à son tour interprétée. On n'en sort jamais.

Le niveau baisse

«**L**es élèves qui entrent dans les classes préparatoires aux Grandes Écoles n'y apportent que des connaissances superficielles et une formation technique à peu près nulle ; ils abordent cette période décisive de leurs études avec une absence totale de maturité d'esprit, une puissance de travail inexistante, une incompréhension générale de la vie intellectuelle qui n'a d'égale qu'une ambition souvent démesurée. (...) Malgré la diffusion de l'enseignement, et, peut-être, à cause de cette diffusion même, nous assistons à un recul général de la vie intellectuelle.» Propos excessifs? Non : ils sont confirmés par le rapport sur le concours d'admission à l'École polytechnique. «Très peu de copies portent la marque d'une personnalité. Toutes les autres dénotaient l'absence d'esprit critique et de jugement. Partout la même abondance décevante de calculs sans la moindre trace d'effort personnel. Souvent la rédaction présentait le même manque d'ordre et de soins.»

Attention! Le premier extrait provient d'un exposé fait en 1935, le second du rapport sur le concours 1933. Aujourd'hui, ces alarmes nous font sourire : les professeurs sévères qui, dans notre jeunesse, nous reprochaient notre nullité (il m'en souvient encore...) avaient été précisément les nullards stigmatisés par les professeurs de 1930! En disant «le niveau baisse», ils ne faisaient donc que répéter la leçon qu'ils avaient eux-mêmes subie...

La baisse de niveau serait-elle une loi de nature? Je n'en crois rien. S'effarer du niveau est l'immuable attitude de ceux qui entretiennent un rapport crispé avec leur savoir. Les jeunes ont à s'adapter à un monde différent de celui auquel les vieux ont eu à s'adapter. Ils y mettent ni plus ni moins de qualités, mais ils les utilisent autrement. Dire «le niveau baisse» est une facilité de pensée que doit s'interdire quiconque veut comprendre quelque chose à la marche du monde et des générations – et à son propre vieillissement.

DES CAILLOUX DANS LES CHOSSES SÛRES

DIDIER NORDON

Des cailloux
dans les choses sûres



Didier Nordon épingle les scientifiques et les politiques qui extraient de la science plus qu'elle ne dit. Nul doute que chacun reconnaîtra dans ces piques... le défaut de son condisciple.

144 pages - 65 F
Code 1897

**POUR LA
SCIENCE**
DIFFUSION BELIN

Bon de commande en p. 98

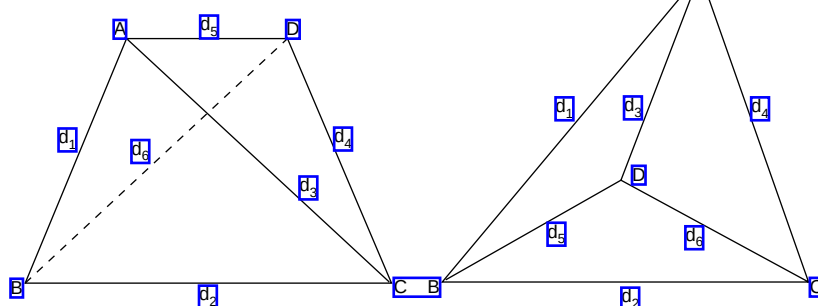
JEU-CONCOURS N° 49

PIERRE TOUGNE

Quadrilatères

Un quadrilatère ABCD dans le plan détermine six distances mutuelles entre ses sommets. Ces six distances ne peuvent toutes être égales (la justification est laissée au soin du lecteur). En revanche, si ces six distances sont différentes, il n'existe que deux types de quadrilatères, selon qu'un sommet, que nous appelons D, est intérieur ou

extérieur au triangle ABC formé par les trois autres sommets comme cela est montré sur les figures ci-dessous. Sauriez-vous déterminer (graphiquement) toutes les familles de quadrilatères essentiellement distinctes telles que les six distances mutuelles ne prennent que deux ou trois valeurs ?



Envoyez vos réponses aux questions sur carte postale à *Pour la Science*, 8, rue Férou, 75006 Paris. Parmi les réponses exactes reçues pendant le mois de juillet 1998, dix gagnants tirés au sort recevront un livre.

RÉPONSE AU JEU-CONCOURS N° 47

La solution de ce jeu-concours est un exemple typique de la récursivité. On désigne par $N(i, j)$, le nombre de chemins arrivant à la case (i, j) . Si les déplacements autorisés sont à droite ou en bas, on ne peut accéder à la case (i, j) qu'à partir des cases $(i-1, j)$ et $(i, j-1)$ d'où la relation de récurrence $N(i, j) = N(i-1, j) + N(i, j-1)$. De plus, on a $N(1, 1) = 1$ pour la case départ et $N(a, b) = 0$ pour les cases interdites. Il suffit alors d'appliquer cette règle comme indiqué sur l'échiquier 1 de la figure de gauche (un tableur est particulièrement bien adapté pour traiter ce genre de problème). Ainsi, pour les trois échiquiers proposés, les réponses étaient 982, 828, 712. Remarquons qu'en l'absence de

cases interdites, il existe une formule simple pour le nombre de chemins aboutissant à la case (n, p) , soit $C_{n,p} = n!p!/(n-p)!$. Si, de plus, on autorise un déplacement en diagonale Sud-Est, la formule de récurrence devient $N(i, j) = N(i-1, j) + N(i-1, j-1) + N(i, j-1)$. Un exemple d'application de l'algorithme sur l'échiquier 2 est représenté sur la figure de droite. Les réponses étaient donc 11 388, 13 760, 9 960.

En ce qui concerne le damier, la relation était $N(i, j) = N(i-1, j-1) + N(i-1, j+1)$, ce qui conduit aux valeurs 126, 334, 436, 409, 251 pour les cases blanches à partir de la gauche (et les mêmes valeurs pour les cases noires à partir de la droite).

1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8
1	3	6	10	15	21	28	36
1	4	10	0	0	21	49	85
1	5	15	0	0	21	70	156
1	6	21	21	21	42	112	267
1	7	28	49	70	112	224	491
1	8	36	85	155	267	491	982

1	1	1	1	1	1	1	1
1	3	5	7	9	11	13	15
1	5	0	12	28	48	0	28
1	7	12	24	64	140	188	216
1	9	28	64	152	356	684	1 088
1	11	0	92	308	816	0	1 772
1	13	24	116	516	1 640	2 456	4 228
1	15	52	192	824	2 980	7 076	13 760

Oiseaux et dinosaures

L'article de Kevin Padian et Luis Chiappe *L'origine des oiseaux et de leur vol* (voir *Pour la Science*, avril 1998) affirme que les oiseaux sont issus des dinosaures théropodes. Que l'*Archaeopteryx* soit un dinosaure, on peut le croire. Qu'il ait eu pour ancêtres de petits reptiles coureurs, c'est très possible. Mais qu'il soit un oiseau, c'est une autre affaire : seule la présence de plumes en fait un oiseau. On possède une demi-douzaine d'individus sur lesquels pas une seule plume n'a de barbes séparées, alors qu'un coup de vent peut suffire pour cela. À mon avis, les plumes de l'*Archaeopteryx* sont des lames cornées finement nervurées. Les barbes étaient soudées entre elles et cela représenterait un état primitif de la plume d'oiseau.

L'organisation des oiseaux est trop complexe pour que leur origine remonte seulement à 150 millions d'années. *Archaeopteryx*, que l'on n'a retrouvé qu'en Bavière, ne serait, au Jurassique supérieur, qu'un «fossile vivant», et il y aurait eu, à la même époque, de vrais oiseaux. Ceux-ci descendraient des Archosauriens qui vivaient au début de l'ère secondaire. Leurs ancêtres de petite taille auraient évolué dans des forêts tropicales où, en sautant de branche en branche et grâce à l'apparition de plumes, ils auraient appris à voler et seraient devenus des oiseaux. Leurs ancêtres n'ont jamais été coureurs. La course suppose un mouvement alterné des bras, alors que le saut demande leur synchronisation, propice au vol. Mais une augmentation de la taille a pu inciter certains oiseaux à prendre l'envol en courant. L'évolution a été très discrète, le milieu forestier étant fermé et peu favorable à la fossilisation.

Maurice POMARÈDE, Montpellier

Réponse d'Eric Buffetaut

Maurice Pomarède soulève deux points importants : la nature des plumes d'*Archaeopteryx* et les modalités d'acquisition du vol. Les plumes d'*Archaeopteryx*, plus ou moins bien conservées suivant les spécimens, ont fait l'objet de nombreuses études détaillées qui montrent bien qu'elles sont de véritables plumes. Elles diffèrent peu des plumes des oiseaux modernes, et ne sont pas des lames cornées.

Par ailleurs, *Archaeopteryx* est seulement le plus ancien oiseau connu. Qu'il ait vécu il y a 150 millions d'années n'implique pas que l'origine des oiseaux soit aussi récente : il a eu des ancêtres moins aviens que lui. Vivaient-ils dans les arbres ou couraient-ils sur le sol? Cette question est très débattue, et depuis long-

temps. Les deux hypothèses sont étayées par de bons arguments. On peut espérer que la découverte d'oiseaux ou de proto-oiseaux antérieurs à *Archaeopteryx* nous éclairera un jour.

Pucerons et miellat

La légende de l'illustration qui accompagne l'analyse de deux livres consacrés aux pucerons (voir *Pour la Science*, mai 1998) affirme que le miellat produit par les pucerons sort de deux «cornicules» placés à l'extrémité de l'abdomen. Est-ce vraiment le cas? Monique Paranthou, Auch

Réponse de Jacques d'Aguilar

Le puceron absorbant la sève des plantes riche en sucres, mais pauvre en acides aminés, il est contraint d'en ingérer de grandes quantités pour satisfaire ses besoins en protéines. Le produit de la digestion, amassé dans la partie dilatée du rectum, est rejeté à l'extérieur sous forme de fines gouttelettes. Ce miellat est donc expulsé par l'anus, non par les cornicules. Riche en sucres, il favorise le développement de champignons saprophytes qui entraînent des fumigines, maladies associées au dépôt d'un enduit noir qui entrave l'assimilation chlorophyllienne. Quant aux cornicules, dont le rôle est mal connu, ils permettraient l'émission d'hémocytes à inclusions cireuses (l'équivalent de cellules sanguines) et de phéromones.

Les carburants de demain

Le dossier sur l'avenir du pétrole (voir *Comment prévenir la prochaine crise pétrolière*, *Pour la Science*, mai 1998) est excellent, mais il me semble qu'on ne prend pas une hauteur suffisante pour bien centrer la recherche des solutions à l'épuisement du pétrole brut. Nous avons deux besoins principaux : les consommateurs fixes et les transports. Les premiers sont typiquement la génération d'électricité, le chauffage et la pétrochimie. Leur combustible idéal est le gaz. Il peut être économiquement attrayant, ici où là, de le liquéfier, mais cela restera une aberration sans avenir. Pour les transports, c'est entendu, il faut un hydrocarbure liquide. Le gaz liquéfié, sauf les petites quantités de propane et de butane, ne convient pas, car il faut une pression énorme (à l'échelle domestique) ou une température cryogénique. La source naturelle de liquides combustibles me semble plutôt être le charbon, aux réserves immenses et dont on ne fera rien d'autre.

Roger PÈRÈS, Le Taillan-Médoc

Systèmes ouverts ou fermés?

Max Dauchet soulève le problème du monopole de la Société *Microsoft* sur la micro-informatique, et celui de l'aptitude du pouvoir politique à orienter le marché et la recherche (voir *Informatique et société*, *Pour la Science*, juin 1998). L'alternative entre les systèmes ouverts ou fermés pose toutefois incomplètement le problème. Le choix tranche entre la recherche et les fournisseurs de logiciels, mais il ne crée en aucun cas les conditions favorables au développement d'une industrie logicielle crédible et prospère. Au mieux, les bonnes idées d'une recherche sur les systèmes ouverts seront récupérées par la Société *Microsoft*. Au pire, les recherches fondées sur des systèmes fermés seront étouffées. La clé pourrait être, au-delà des lois antitrust, d'obliger juridiquement tous les fournisseurs de logiciels à respecter les normes qu'ils établissent ensemble, et à organiser une politique de recherche européenne qui tienne aussi compte de ces normes.

Henri SINTUREL, Paris

Réponse de Max Dauchet

L'établissement et le respect de normes publiques sont en effet souhaitables. Les normes sont une solution qui assure les compatibilités techniques et garantit l'ouverture des marchés. Le caractère public des normes évite aussi la concentration incontrôlable de pouvoir. Je précisais bien que l'alternative entre systèmes ouverts et systèmes fermés n'est qu'un exemple des problèmes nouveaux que pose la révolution numérique à la société, et je conclus, comme H. Sinturel, que l'Europe doit être ambitieuse dans ses approches. Je ne pense pas que le pouvoir politique doive orienter le marché et la recherche, si par orienter il faut entendre encadrer et décréter, mais il doit aider l'émergence de perspectives et inciter les acteurs à construire des solutions nouvelles. Au niveau de la formation professionnelle initiale et continue des informaticiens, le ministère de l'Éducation nationale, de la Recherche et de la Technologie vient de me confier une mission dans cet esprit.

Écrivez à *Pour la Science* vos commentaires et vos réflexions sur l'actualité scientifique. Courriel électronique : 101641.3525@compuserve.com. Retrouvez la Tribune sur l'Internet, à l'adresse : <http://www.pourlascience.com>



Les Académies

GERARD TOULOUSE

A quoi servent-elles ?

Les académies, notamment les académies des sciences et les académies d'ingénierie, sont des institutions qui ont encore, et plus que jamais, des rôles importants à jouer : information et conseil (auprès des décideurs et du public), coopération internationale, soutien aux talents individuels, réflexion éthique sur les relations entre science et société. Au regard de pareilles missions, les compagnies académiques, émanations originales de la communauté scientifique, méritent donc une attention vigilante et critique.

Toutefois l'opinion commune, parmi mes compatriotes scientifiques, se situe à l'opposé. Envers les académies règne une attitude de «négligence indulgente» : ces institutions sont vues comme des clubs de vieux messieurs, des musées du troisième âge, des survivances désuètes et inoffensives (l'un allant de pair avec l'autre). On ne sait pas bien ce qui s'y fait, et on ne s'en soucie pas trop. La négligence incline à l'indulgence, et l'indulgence renforce la négligence. Nombre d'académiciens eux-mêmes, attachés avant tout à la conservation pérenne de leurs avantages et privilèges, s'octroient quêttement, en miroir, une semblable «négligence indulgente».

Comment organiser les académies pour qu'elles remplissent leurs fonctions malgré le petit nombre de leurs membres ? Dans la population active des pays industrialisés, la proportion de scientifiques est de l'ordre du pour cent ; et, parmi les scientifiques, la proportion des académiciens se situe entre le pour mille et le pour cent. Pourtant, à travers le monde, les académies ne sont pas toutes du même acabit. Notre Académie des sciences, à Paris, et la *Royal Society* de Londres se prêtent bien à comparaison, toutes deux étant nées il y a trois siècles environ.

Pendant la première moitié de leur existence, la compagnie parisienne a pu sembler mieux conçue. En 1734, Voltaire jugeait ainsi que l'académie londonienne «n'est pas si bien réglée que la nôtre, et cela par la seule raison peut-être qu'elle est plus ancienne ; car si elle avait été formée après l'Académie de Paris, elle en aurait adopté quelques sages lois et eût perfectionné

les autres. La Société royale de Londres manque des deux choses les plus nécessaires aux hommes, de récompenses et de règles. C'est une fortune sûre à Paris pour un géomètre, pour un chimiste, qu'une place à l'Académie ; au contraire, il en coûte à Londres pour être de la Société royale».

Toutefois, vers la moitié du XIX^e siècle, la conjoncture se renverse : l'Académie des sciences s'engage dans un long déclin, à cause de sa structure rigide, tandis que la *Royal Society* entreprend une réforme lucide et courageuse, qui lui permettra de rester dorénavant en prise sur le progrès des sciences.

En 1966, Jacques Monod (1910-1976) dressait ainsi une comparaison : «[Ma conviction est] que l'Académie [des sciences], dans sa structure actuelle, ne peut pas ou, du moins, ne peut plus remplir le rôle qui devrait être le sien et cela pour un motif très simple. L'Académie compte aujourd'hui, si je ne me trompe, environ quatre-vingts membres ; elle en comptait, je crois, soixante il y a cent cinquante ans. (...) Il n'était pas possible à l'Académie, sans une adaptation de ses structures et sans une augmentation considérable du nombre de ses membres, de continuer de représenter comme elle le doit l'ensemble de la science française et de compter en son sein la majorité sinon la totalité des meilleurs et des plus dignes dans toutes les disciplines. De là vient, à mon avis, l'effacement progressif de l'Académie (...). Cette évolution n'était nullement inéluctable. Vous n'ignorez pas, bien entendu, que la Société royale, sœur jumelle de l'Académie des sciences française, a subi une évolution toute différente. Elle compte aujourd'hui six cents membres et tous les hommes de science britanniques de renom et de valeur y figurent ou y figureront bientôt. Le prestige de la Société royale n'a nullement souffert, au contraire, de l'accroissement du nombre de ses fellows. Son rayonnement en Angleterre même comme à l'étranger reste considérable ; de même que demeurent très importantes les fonctions qu'elle remplit comme conseiller des pouvoirs publics dans tous les domaines de la science et de la technologie.»

En 1976, une cote d'alerte fut atteinte. Trop de lauréats du prix Nobel ou de la médaille Fields demeuraient à l'écart de l'Académie des sciences, menacée de perdre définitivement toute légitimité internationale. Une réforme fut mise en œuvre, mais, parant au plus pressé (afin d'améliorer la qualité), ce fut une demi-réforme, esquivant le problème de la quantité. Or il n'y a pas de qualité à long terme sans quantité suffisante.

Aujourd'hui, la *Royal Society* comporte environ 1150 fellows, soit dix fois plus que le nombre de membres à Paris (ou quatre fois plus que le total du nombre de membres et du nombre de correspondants, car il y a deux classes à Paris pour une seule à Londres). Autant la *Royal Society* est ouverte sur le Commonwealth et la planète, autant l'Académie des sciences est recroquevillée sur l'Hexagone et même, dans certaines disciplines, sur Paris. La préférence nationale (ou même capitale) y joue à plein, au détriment de l'espace francophone (et de la province).

Comment la situation évoluera-t-elle ? Au cours de ce siècle, marqué par l'essor et l'impact des techniques, des académies d'ingénierie sont nées dans l'ensemble des pays industrialisés, notamment nordiques et anglo-saxons (en Suède en 1919 ; au Danemark en 1937 ; en Norvège en 1955 ; aux États-Unis en 1964, etc.). Il est envisagé qu'en France, en l'an 2000, soit enfin créée une telle académie autonome, qui remplacerait l'actuel Conseil des applications de l'Académie des sciences, le CADAS, créé en 1982.

Choisira-t-on de bâtir cette nouvelle Académie sur le modèle continental étriqué, avec hiérarchies et privilèges ? Ou bien se tournera-t-on vers la modernité en adoptant le modèle nordique, ouvert et dynamique ? Pour toutes les académies, en ce pays et dans sa zone d'influence, ce sera une heure de vérité.

Gérard TOULOUSE est physicien à l'École normale supérieure. Il vient de publier *Regards sur l'éthique des sciences*, aux éditions Hachette Littératures, 1998.



Premières images au microscope

BRIAN FORD

En refaisant les toutes premières expériences de microscopie optique, on a observé ce que les biologistes voyaient avec leurs instruments à une seule lentille.

En 1674, le Hollandais Antonie Van Leeuwenhoek découvrit, grâce à l'un de ses microscopes, un nouvel univers. Il observait de la vase recueillie dans un lac lorsqu'il découvrit des organismes inconnus : il décrit le «spectacle fascinant» des animalcules qui virevoltaient si rapidement dans l'eau.

Leeuwenhoek (1632-1723) ouvrait ainsi, par inadvertance, la voie de la microbiologie. Il fabriqua lui-même plus de 500 microscopes, alors que la science de l'optique n'était pas encore à son apogée. À l'aide de ces instruments, il observa des animalcules (des microbes), mais aussi plusieurs types de cellules, tels des hématies et des spermatozoïdes. Il décrit aussi des bactéries, des protozoaires, des cellules végétales et des champignons. Néanmoins, Leeuwenhoek fut considéré par ses contemporains comme

un dilettante à l'imagination fertile, et l'on n'accorda guère de crédit à ses observations. Il fallut attendre le milieu des années 1800, et Louis Pasteur, pour que l'existence des micro-organismes soit reconnue. Aujourd'hui encore, quelques scientifiques doutent que Leeuwenhoek ait réellement observé des micro-organismes avec ses instruments rudimentaires à une seule lentille.

Des biologistes, ayant observé des préparations avec un microscope de Leeuwenhoek, ont conclu que les grossissements disponibles ne leur permettaient pas de distinguer les détails que Leeuwenhoek avait décrits ; dans l'une de ces expériences, les hématies ressemblaient plutôt à des taches. De telles accusations ont aussi été lancées contre Robert Brown. En 1827, ce jeune botaniste écossais observa à l'intérieur de cel-

lules le mouvement incessant de minuscules particules. Il venait de mettre en évidence le phénomène qui porte aujourd'hui son nom : le mouvement brownien. Quelques années plus tard, Brown observa à travers son microscope à une seule lentille un autre phénomène qui n'avait encore jamais été observé. Au cours d'un voyage en Australie, à la recherche de plantes exotiques, il avait été fasciné par les orchidées qu'il cueillait. En les examinant, il découvrit, dans chaque cellule, une structure circulaire, plus opaque que la membrane cellulaire. Brown montra que les cellules de nombreux organismes en sont pourvues, et nomma cette structure le «noyau cellulaire».

Comme Leeuwenhoek, Brown fut contesté. Quelques-uns de ses contemporains ont même affirmé qu'il avait simplement confondu le mouvement des

a

b

1. DES CELLULES de la fleur *Tradescantia virginiana* ont été observées au moyen des microscopes de Robert Brown. Elles étaient éclairées par la lumière du Soleil (en haut de la page de droite). Le microscope d'Antonie Van Leeuwenhoek, conservé à l'Université d'Utrecht, aux Pays-Bas, révèle des hématies (a), voire même le noyau de certaines cellules, des leucocytes par exemple. Un microscope actuel à lentille unique donne l'image de cellules de levure (b) ou des spores de champignons (c).

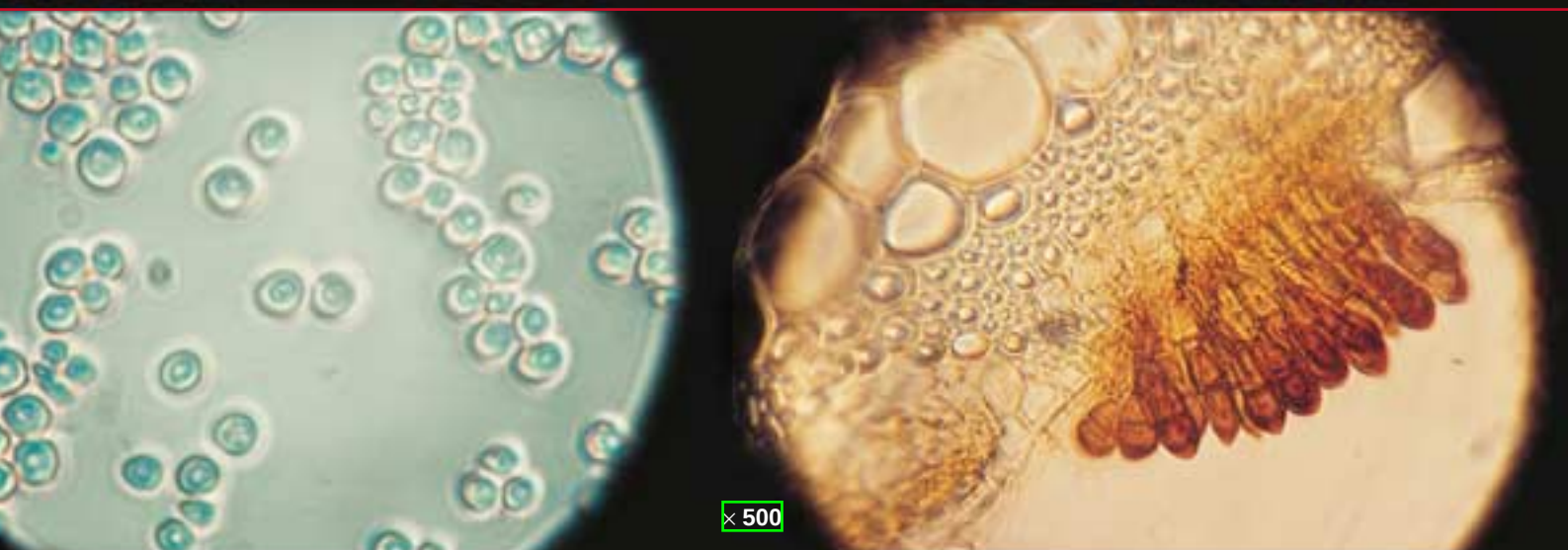
× 170

× 446



× 110

C



× 500



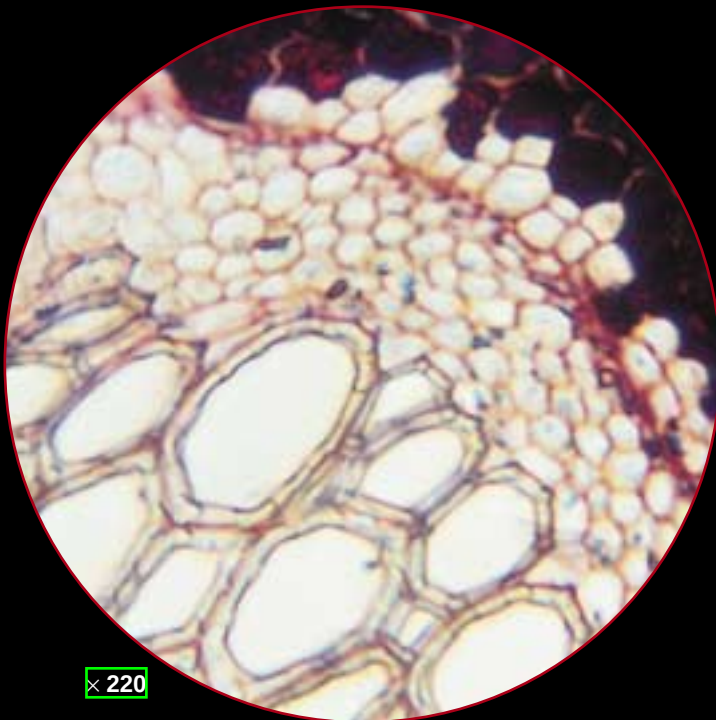
× 600

2. DES SPORES DE TRUFFE (*Tuber melanosporum*) représentent un excellent test pour le microscope de Leeuwenhoek. Les spores minuscules ont la taille d'une cellule (de l'ordre de 30 micromètres) et les épines celle d'une bactérie (de l'ordre de 3 micromètres). En dépit de sa simplicité, ce microscope vieux de 300 ans révèle spores et épines.



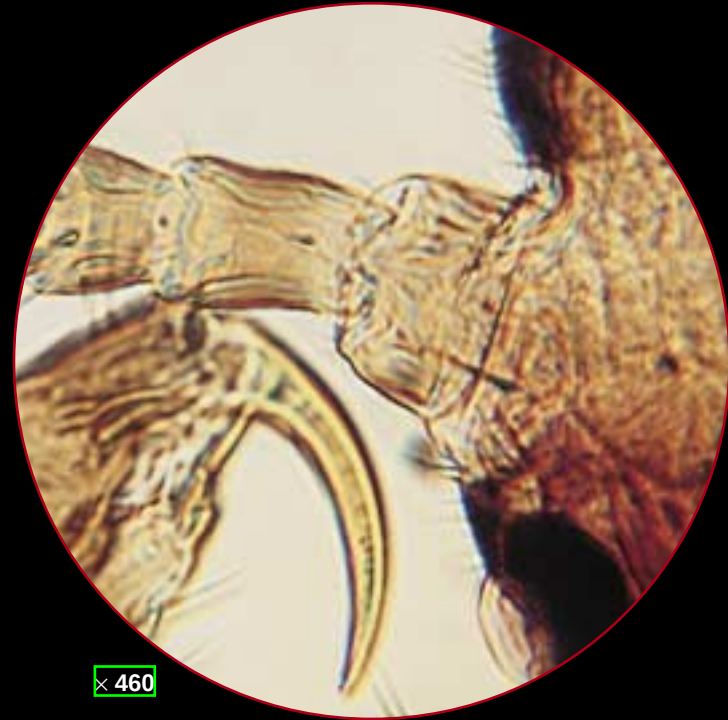
× 1 200

3. DES BACTERIES ont effectivement été observées par Leeuwenhoek. Cette image de *Spirillum*, l'une des espèces décrites par ce Hollandais ingénieux, prouve son talent. Pour observer ces bactéries, on doit illuminer la préparation correctement et faire une mise au point très minutieuse.



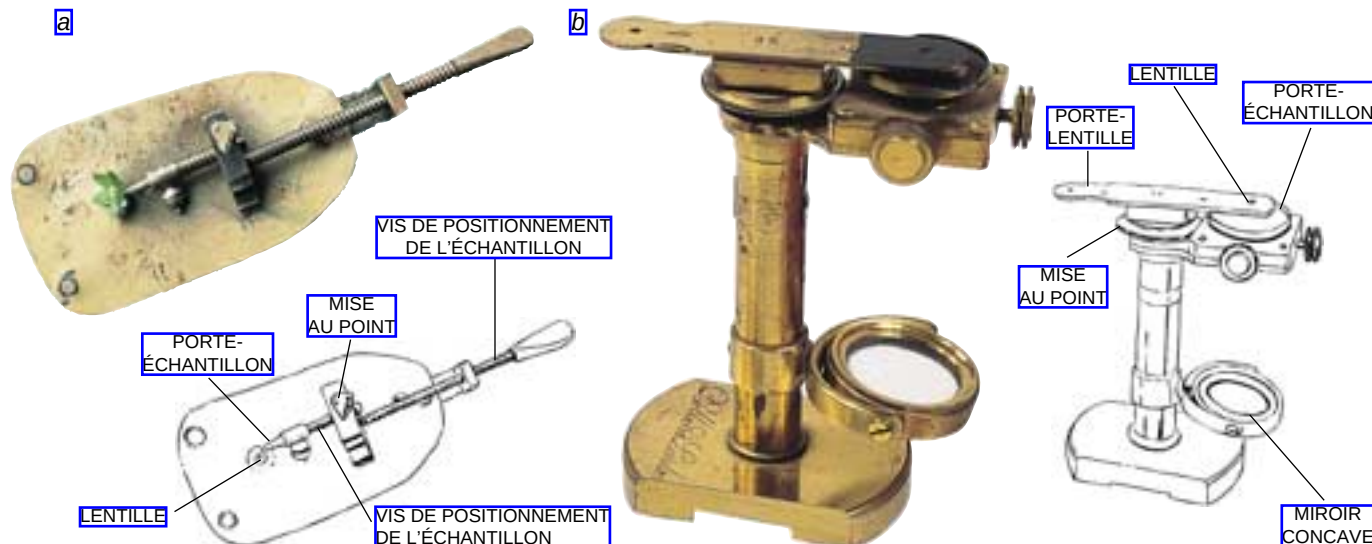
× 220

4. LES PAROIS DES CELLULES d'une racine d'*Osmunda regalis* ont été observées au moyen d'un microscope à une seule lentille datant des années 1700. La netteté de cette image est comparable à celle des images réalisées avec des microscopes optiques plus récents.



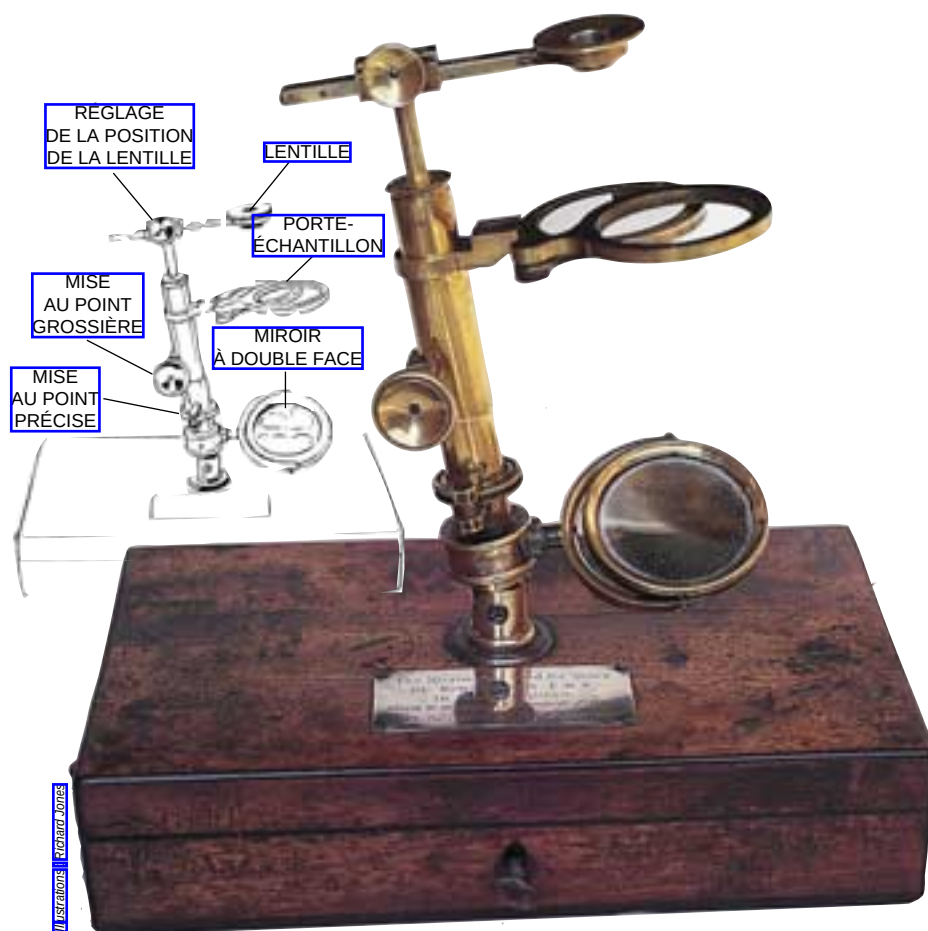
× 460

5. UNE TETE DE POU a été observée au microscope avec une lentille neuve : on distingue parfaitement l'antenne et les yeux brun foncé de cet insecte, que les biologistes des siècles passés ont étudié tout à loisir. Ces observations prouvent la qualité des images fournies par les microscopes du XVII^e siècle.



6. LES PREMIERS MICROSCOPES des années 1670 ressemblaient à ceux de Leeuwenhoek. Ce microscope en laiton a été fabriqué à la main (a). L'unique lentille, un peu plus grosse qu'une tête d'épingle, assurait un grossissement tel qu'on voyait des bactéries. On l'aperçoit sous la feuille ; elle était maintenue entre deux feuilles de laiton par deux écrous (en bas à gauche). Pour observer l'objet, on retournait l'instrument après avoir positionné l'objet sur le porte-objet (une pointe). Une tige filetée permettait de positionner l'échantillon devant la len-

tille, une autre de rapprocher ou d'éloigner l'échantillon de la lentille, c'est-à-dire de faire la mise au point. Ces instruments ont été progressivement perfectionnés pour atteindre leur apogée, 150 ans plus tard (b) avec les instruments fabriqués par la société londonienne Dollond. Robert Brown possédait l'un de ces instruments. Au sommet de la colonne centrale se trouve une vis de mise au point précise, et deux vis agissent sur un plateau qui soutient le porte-échantillon. L'un et l'autre microscopes mesurent moins de huit centimètres.



7. BROWN, le père du mouvement brownien, a fait ses observations au moyen de microscopes petits et robustes, fabriqués par la société londonienne *Bancks & Son*. Ce modèle mesure environ 15 centimètres de hauteur. Sur le pied du microscope sont adaptées une vis de mise au point grossière et une vis de mise au point précise, et un miroir à double face. Avec de tels microscopes, on obtient des grossissements compris entre 30 et 150.

grains de pollen observés avec celui, infiniment plus ténu, des structures microscopiques que contenaient ces grains. Des détracteurs de Brown ont même soutenu que son microscope avait pu lui être utile pour faire des dissections, mais qu'il n'avait, en aucun cas, pu y observer des noyaux cellulaires.

Comment le savoir ? J'ai repris les microscopes d'origine de Leeuwenhoek et de Brown, et j'ai moi-même fabriqué des lentilles. J'ai ensuite refait plusieurs de leurs expériences. Malheureusement, je ne disposais d'aucun compte rendu détaillé des méthodes utilisées par ces deux biologistes ; de plus, il faut un soin extrême pour mettre au point les microscopes et reproduire fidèlement les expériences. Contrairement à ce qu'affirmaient certains, j'ai alors constaté qu'avec un microscope rudimentaire on observe des bactéries vivantes, des noyaux cellulaires et le mouvement brownien. Avec le microscope de Brown, j'ai même distingué des mitochondries, qui sont des centaines de fois plus petites que le noyau cellulaire. Leeuwenhoek et Brown avaient honnêtement décrit ce qu'ils avaient observé.

Les images présentées ici n'ont pas été retouchées. Elles illustrent la qualité des images qu'offraient les instruments les plus simples – ici des microscopes à une seule lentille datant des XVII^e et XIX^e siècles.

Brian FORD est un biologiste anglais spécialiste des microscopes.



La mondialisation

BERNARD GUERRIEN • FRANCISCO VERGARA

La localisation des services et l'automatisation de l'industrie limiteront toujours la mondialisation.

La «mondialisation» effraie : nous serions «envahis» par des produits étrangers, ordinateurs, fruits et légumes, vêtements, etc. Parallèlement certains secteurs de l'économie souffriraient de la «délocalisation» de leurs activités vers les pays à bas salaires.

Aussi les alarmistes s'étonnent-ils que tant d'activités subsistent et se développent encore en France. Or, malgré un chômage important, l'emploi (le nombre d'actifs occupés) continue d'augmenter en France : en 1997, il y avait un million d'emplois de plus qu'il y a dix ans. En outre, la France dégage un excédent des exportations sur les importations, excédent qui augmente régulièrement depuis sept ans et qui a atteint 230 milliards de francs en 1997 (2,8 pour cent du PIB, mieux que le Japon). Si un pays en «envahit» d'autres avec ses produits, ne serait-ce pas la France ?

D'autre part, en France et dans nombre de pays, la mondialisation (c'est-à-dire le degré d'intégration dans l'économie mondiale) a atteint, vers 1974, un palier dont la valeur correspond... à celle de 1913 ! La mondialisation n'est qu'un retour à une situation d'échanges interrompue par les deux guerres mondiales et la grande récession des années 1930.

Comment expliquer cette stabilité ? En 1974 (juste avant la flambée du prix du pétrole), la France importait 75 pour cent de l'énergie qu'elle consommait, ce qui représentait 4,5 pour cent de son PIB. En 1996, la part importée était tombée à 46 pour cent, et ne correspondait plus qu'à 1,5 pour cent de son PIB. La production de maïs est un autre exemple. Le consommateur ne s'aperçoit pas de cette évolution – vers moins de «mondialisation» –, l'origine des produits n'étant pas toujours identifiable.

Même dans les branches où l'industrie nationale semblait avoir été irrémédiablement évincée (téléviseurs, magnétoscopes, habillement bon marché), on observe souvent un retour de la fabrication et de l'assemblage national, voire européen. Ainsi, les livraisons

de téléviseurs par l'industrie française, qui avaient stagné à partir de 1978 au niveau de deux millions d'unités par ans, ont repris leur croissance à partir de 1989 et atteignaient 3,5 millions d'unités en 1996, et cela dans un marché en stagnation absolue. Cette relocalisation résulte des avantages que procure la proximité des fournisseurs et des clients : les dépenses liées à leur mise en rapport sont souvent plus importantes que les coûts de fabrication.

Dans les industries qui «assemblent», le faible coût de la main-d'œuvre a expliqué les délocalisations vers l'Asie des dernières décennies. Mais, les robots commandés par ordinateur tendent à réduire l'importance du coût directement salarial dans ces branches. Sans ce facteur, on ne pourrait comprendre pourquoi les grands entreprises co-réennes et japonaises ouvrent des usines de production en Europe et aux États-Unis.

Pour les matières premières qui ne disparaissent pas avec la consommation du produit (fer, aluminium, plastique, papier, etc.), à moins d'imaginer des décharges publiques éternellement croissantes, il faut admettre que le recyclage (c'est-à-dire la production domestique) occupera une place croissante.

La part des services dans le PIB français est passée de 59 pour cent en 1974 à 72 pour cent en 1996 et continue à augmenter. Si on ajoute à ce chiffre la valeur ajoutée non industrielle de la production industrielle (secrétariat, comptabilité, gestion des ressources humaines, nettoyage, sécurité, etc.), les services approchent

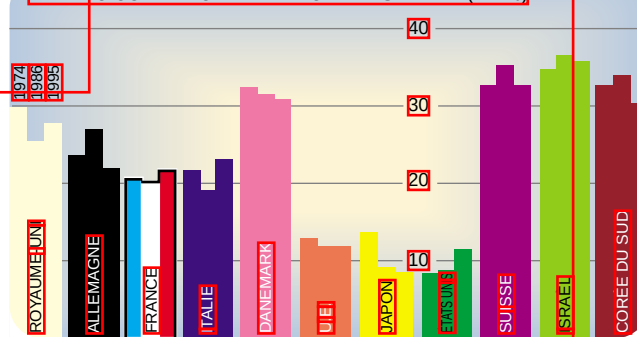
90 pour cent de la production nationale. Or ces chiffres sont d'une extraordinaire stabilité depuis 20 ans : la part importée dans la consommation nationale de services qui était de 7,7 pour cent en 1975 est passée à 7,1 pour cent en 1996.

Si certains services peuvent être délocalisés, notamment ceux liés au tourisme, il existe un noyau dur d'activités du tertiaire qui doivent demeurer à proximité des consommateurs : la construction (bâtiment et infrastructures), les coiffeurs, les restaurants, les services d'entretien et de réparation (soudage, plomberie et au garagiste), les services municipaux, une partie importante des transports, tout ce qui a trait à l'enseignement et à la santé. À cela s'ajoutent la fourniture de l'eau, du gaz et de l'électricité, ainsi que les différents «services» procurés par les communes et par l'État (armée, police, justice). Ces activités – marchandes ou non marchandes – occupent plus du tiers de la population des pays développés.

Ainsi, pour pratiquement tous les produits qui ne se caractérisent pas par une fourniture directe du service au client – tels les produits manufacturés et agricoles –, l'activité d'intermédiation entre le producteur et le consommateur est prépondérante. Remarquons que le café qui arrive au Havre à 4 francs le kilo est vendu chez notre détaillant autour de 60 francs : une grande partie de cette «valeur» ajoutée entre le port et l'épicier du coin, rémunère le transport, le stockage, l'emballage, le travail du grossiste et du détaillant, la publicité. Toutes ces activités ne peuvent être délocalisées. Comme elles sont rémunérées à des salaires «locaux» plus élevés que ceux qui prévalent dans les pays producteurs, la disproportion entre prix de production et prix de vente au détail est frappante.

Ainsi, l'existence d'activités qui ne peuvent être délocalisées impose une contrainte «objective» à la mondialisation. Elle explique pourquoi celle-ci est, et demeurera sans doute (au-delà des apparences), un phénomène limité.

PART DU COMMERCE EXTERIEUR DANS LE PIB (EN %)





Jambons crus d'Espagne

HERVÉ THIS

L'attribution d'appellations d'origine contrôlée impose une exploration rationnelle du goût.

Ces dernières années, la production des jambons crus espagnols a augmenté jusqu'à atteindre 30 millions de pièces par an. Certaines d'entre elles, obtenues à partir du porc ibérique local (plus d'un million de jambons par an) par des méthodes traditionnelles, dans le Sud-Ouest de l'Espagne, ont un goût si remarquable que l'Union européenne leur a accordé le privilège d'appellations d'origine contrôlée. Comment ce goût se forme-t-il ? Jésus Ventanas et ses collègues de l'École vétérinaire de Caceres ont analysé les diverses étapes de la longue préparation de ces jambons et identifié les raisons de leur originalité.

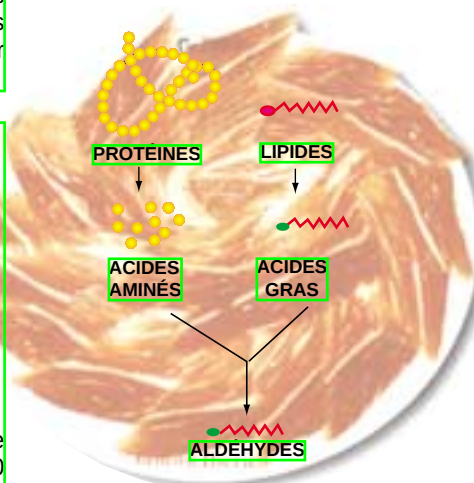
Le procédé traditionnel de fabri-

cation a ses particularités qui commencent dès l'élevage : les porcs, de race ibérique, sont engraisés en liberté, se nourrissant majoritairement de glands et d'herbes, jusqu'à ce que leur masse atteigne 160 kilogrammes. Après l'abattage, les jambons sont conservés pendant deux jours à une température de 0 °C ; puis ils sont frottés avec du sel contenant un pour cent de nitrate et placés dans un lit de ce sel pendant une semaine à basse température (entre 0 et 4 °C). Débarrassés du sel, ils sont ensuite conservés aux mêmes températures pendant deux à trois mois avant d'être séchés : le séchage se fait à une température croissante pendant un mois et demi jusqu'à 18 °C. Puis, pendant l'été, les jambons sont encore conservés un mois et demi à la température ambiante, avant la maturation finale, qui s'effectue en cave pendant 14 à 22 mois.

Ce procédé très long est une adaptation empirique aux particularités du climat de l'Estramadure, mais, aujourd'hui, les chambres froides et les thermostats évitent les aléas climatiques. Pour adapter les méthodes traditionnelles aux outils de production modernes, les fabricants doivent choisir des températures et des durées précises et, si possible, moins longues. Comment faire ces choix, afin

de conserver le goût particulier des jambons traditionnels ?

Dans les années 1970, des chimistes avaient montré que des produits formés lors de la dégradation des protéines (des molécules formées par l'enchaînement d'acides aminés) et des lipides contribuent à l'arôme de certains aliments, tels le fromage. Ces produits étaient-ils responsables des arômes des jambons espagnols ou étaient-ils seulement des précurseurs de ces arômes ? À partir de



Dans le jambon espagnol, les protéines et les lipides se dégradent en formant de nombreux aldéhydes aromatiques.

1990, les chercheurs de Caceres se sont attachés à reproduire des fabrications traditionnelles en analysant, à chaque étape, les modifications des protéines et des lipides. Ils ont ainsi observé que les acides aminés libérés par la décomposition des protéines, au cours du salage, étaient ensuite dégradés : ce sont des précurseurs de molécules aromatiques.

Plusieurs réactions étaient envisageables : par exemple, les réactions de Maillard entre les acides aminés et les sucres sont responsables des arômes de la viande grillée, de la croûte du pain ou du café torréfié, mais elles se produisent également lors du stockage prolongé

des produits alimentaires, qu'elles font brunir. D'autre part, les dégradations de Strecker sont des réactions des acides aminés avec des acides, tels les acides gras libérés au cours de la dégradation des lipides ; elles engendrent des aldéhydes, souvent aromatiques.

Dans les jambons ibériques, l'accumulation de produits de Maillard augmente proportionnellement à la durée de maturation : on ne saurait obtenir de bons jambons sans une longue maturation. De surcroît, les aldéhydes, qui sont formés après le salage, participent aux réactions de Maillard, formant des produits qui retardent le rancissement des graisses.

Une autre étude, plus récente, a pris le problème à l'envers : les chimistes ont identifié les molécules volatiles et cherché l'origine de ces molécules. Les plus abondantes sont les alcanes, molécules composées exclusivement d'atomes de carbone et d'atomes d'hydrogène. Dans le jambon cru espagnol, on en trouve deux types : d'une part, les alcanes linéaires proviennent vraisemblablement de la décomposition des lipides ; d'autre part, les alcanes ramifiés sont dus à l'alimentation à base de glands et ils sont caractéristiques des porcs ibériques. Ainsi la chimie justifie que l'appellation d'origine contrôlée ne soit attribuée qu'aux jambons provenant d'animaux élevés en liberté dans les chênaies.

Une autre classe aromatiquement importante est celle des aldéhydes linéaires, qui sont formés par les réactions de Strecker, mais aussi par les réactions de rancissement des acides gras insaturés. Là encore, la chimie justifie l'appellation : les porcs ibériques ont une chair marbrée d'une graisse qui contient beaucoup d'acides gras insaturés.

Prochain rendez-vous **France Info** et **Pour la Science** le 28 juillet 1998, avec la chronique **Info Science** de Marie-Odile Monchicourt.

Lancements maritimes

La plate-forme de lancement de fusées est opérationnelle.

La guerre des étoiles n'est plus militaire, mais économique. Après les nations, ce sont les grands groupes industriels qui s'engagent dans la course à l'espace. Ainsi un consortium baptisé *Sea Launch* se prépare à lancer des fusées à partir d'une plate-forme pétrolière reconvertie en pas de tir semi-submersible.

Des centaines de nouveaux satellites devront bientôt être mis en orbite autour de la Terre et, dans une optique industrielle de lancement, la fiabilité cède le pas à la rentabilité : des fusées de fiabilité de 80 pour cent seulement ont des coûts de lancement bien inférieurs à ceux des lancements actuels.

Le consortium (qui regroupe l'américain *Boeing*, l'armateur anglo-norvégien *Kvaerner* et des constructeurs de fusées russes et ukrainiens) réutilise du matériel ancien, moins fiable, mais meilleur marché.

Le pas de tir comporte deux bâtiments : la plate-forme mobile *Odyssey* et le navire de commandement *Sea-Launch-Commander*. À Vyborg et à Saint-Petersbourg, ces bâtiments ont reçu leurs derniers équipements, dont un système de guidage des fusées et les terminaux de commande. Dans le courant de l'été, les deux bâtiments quitteront la Russie pour le port de Long

Beach, en Californie, situé près des centres de production des satellites. Le lanceur se trouve déjà à bord du bateau et le premier satellite prendra place sous sa coiffe. La fusée montée sera ensuite transférée sur la plate-forme de tir.

La plate-forme et le navire de commandement se rendront alors vers les îles Gilbert (République de Kiribati), à 2 500 kilomètres d'Hawaii, d'où le lanceur sera tiré pour la première fois au mois d'octobre. Le lancement est automatique : la plate-forme est évacuée et le navire s'éloigne à cinq kilomètres. La fusée est alors érigée, les réservoirs sont remplis de kérosène et d'oxygène liquide (stockés à bord de la plate-forme) et le tir est effectué.

La zone de lancement est proche de l'équateur, où la vitesse de rotation de la Terre est maximale. L'économie est de taille : lancée de l'équateur, la fusée *Zénit* met en orbite une charge utile de cinq tonnes, tandis que de Baïkonour, située à 51 degrés de latitude, cette même fusée ne place en orbite que deux tonnes.

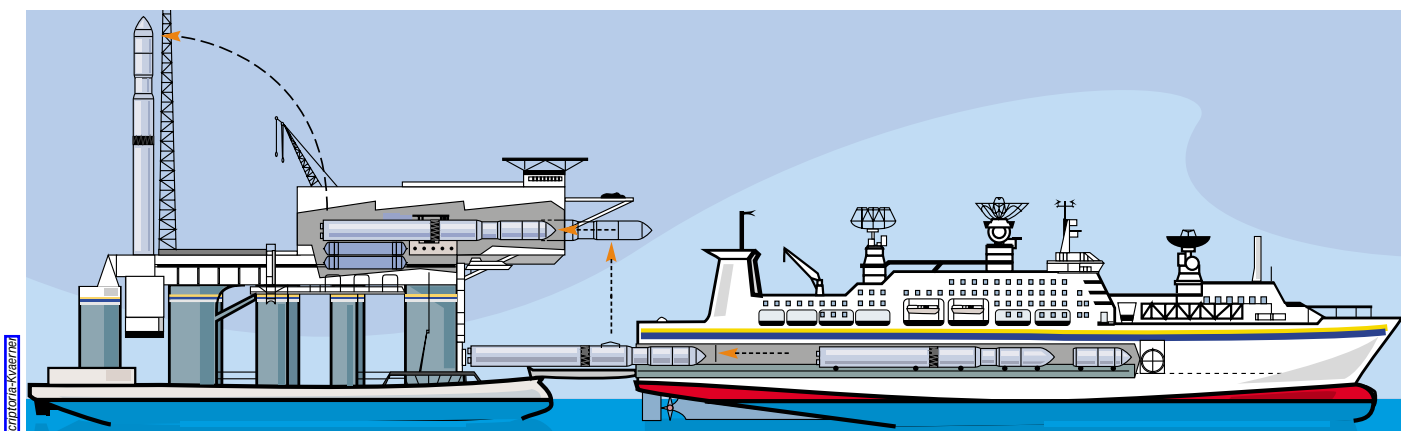
Les ingénieurs envisagent de transborder la fusée sur la plate-forme à quai. Dans un an ou deux seulement, lorsque le pas de tir sera opérationnel, le navire transportera deux fusées, voire trois, et le transfert se fera en mer. En effet, le transbordement en pleine mer d'une fusée entre deux bâtiments non solidaires est une opération délicate : le moindre choc endommage irrémédiablement une fusée.

Sur un pas de tir de la taille d'une plate-forme pétrolière, l'espace est limité : les réservoirs de stockage du kérosène et de l'oxygène liquide sont distants de moins d'une centaine de mètres. Le danger d'accident en est augmenté.

Avant le tir, l'alignement des centrales à inertie du lanceur nécessite une immobilité quasi parfaite. Sur *Ariane*, cette phase dure 45 minutes, où seuls les mouvements dus au remplissage des réservoirs de la fusée sont tolérés. Pour l'alignement des centrales en mer, les ingénieurs utiliseront le dispositif de positionnement global par satellite (GPS). Seuls des mouvements verticaux du lanceur inférieurs à cinq centimètres seront admissibles. Si la hauteur des vagues reste inférieure à trois mètres, la plate-forme, enfoncée de 17 mètres dans la mer, sera alors suffisamment stable. Selon les responsables de *Sea Launch*, les relevés météorologiques de la région sont « encourageants ». Lors du décollage, des mouvements de contrepois garantiraient la stabilité de la structure semi-submersible.

La fiabilité d'un lanceur constitué du premier et du deuxième étage de la fusée *Zénit*, et du moteur du troisième étage de la fusée *Proton*, utilisée à Baïkonour sera-t-elle suffisante ? Ce moteur a connu tant de difficultés d'allumage (dont l'une est à l'origine de la perte de la sonde russe *Mars 1996*) que, sur le lanceur *Proton*, les Russes ont décidé de le remplacer !

Premier tir en octobre.



Le lanceur *Sea Launch* sera assemblé à bord du bateau accompagnateur, dans le port de Long Beach, en Californie, puis transbordé sur la plate-forme de lancement. Les deux bâtiments autonomes se dirigeront alors vers l'équateur, près des îles Gilbert. Commandée du navire, la phase automatique de mise à feu

commencera : érection du lanceur, remplissage en kérosène et en oxygène liquide, alignement des centrales à inertie. Durant cette période, la plate-forme sera enfoncée de 17 mètres dans la mer, ce qui assurera la stabilité. Le bateau de commandement mesure 200 mètres de longueur.

Hello, le Soleil vibre...

Les éruptions solaires créent des ondes sismiques.

Comme une gigantesque cloche sphérique, le Soleil vibre, et les premières observations des mouvements de sa surface datent de 1960. Aujourd'hui, on étudie le déplacement de la surface au moyen d'instruments placés sur *SOHO*, l'observatoire spatial solaire. Alexander Kosovichev, de l'Université Stanford, et Valentina Zharkova, de l'Université de Glasgow, ont ainsi, pour la première fois, observé l'onde sismique déclenchée par une éruption solaire.

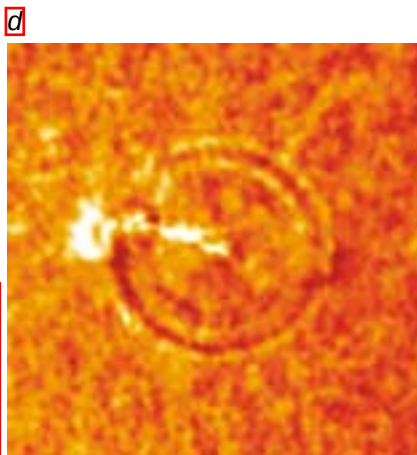
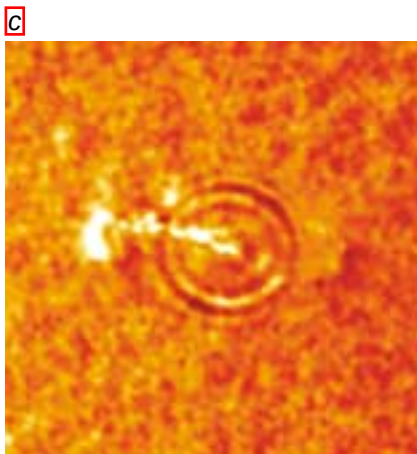
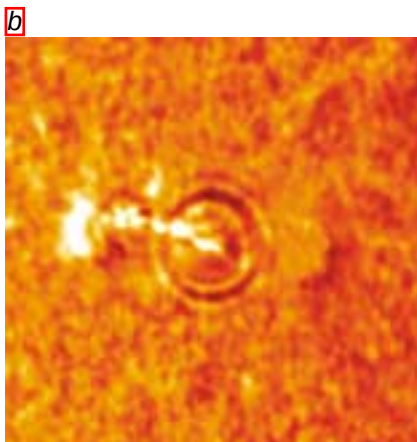
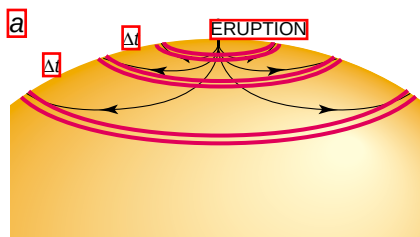
Dans le Soleil, les éruptions émettrices de rayons X résultent de la formation sporadique de tubes de champ magnétique où les électrons de haute énergie sont canalisés. À la surface du Soleil, ces électrons ont un double effet : ils entrent en collision avec les ions de l'atmosphère, ce qui déclenche une bouffée de rayons X ; ils chauffent le plasma ambiant qui engendre une onde de choc, l'onde sismique.

Le 9 juillet 1996, alors que l'activité magnétique du Soleil était faible, une éruption X s'est produite. Grâce à la caméra Doppler-Michelson de *SOHO*, les astronomes ont décelé un fort mouvement ascendant et descendant du gaz à proximité de l'éruption. Avec cet instrument, les astrophysiciens mesurent les variations de la longueur d'onde de la lumière émise en des millions de points du Soleil. Selon que ces points s'éloignent ou s'approchent radialement de la caméra, la longueur d'onde de la lumière augmente ou diminue en raison de l'effet Doppler. Après dépouillement des 35 minutes d'enregistrements, on a reconstruit l'évolution de l'onde sismique en fonction du temps.

L'éruption a créé, dans la surface solaire, une onde qui s'est éloignée de l'épicentre. Le front d'onde s'est d'abord éloigné à 30 kilomètres par seconde, puis il a atteint 100 kilomètres par seconde avant de se dissiper.

Pourquoi cette accélération ? Parce que les ondes sont la manifestation en surface d'ondes sismiques à trois dimensions qui s'enfoncent dans le Soleil. Ces ondes se propagent à une vitesse qui croît avec la température. Comme, dans le Soleil, la température augmente avec la profondeur, leur vitesse augmente et le front d'onde en surface accélère (voir le schéma).

Maintenant qu'ils savent que les éruptions solaires créent des ondes sismiques observables, les astronomes vont les rechercher plus systématiquement.



Les ondes sismiques, créées par l'éruption solaire du 9 juillet 1996, s'enfoncent dans le Soleil (a). À des intervalles de temps égaux (Δt), ces ondes parcourent des trajets plus longs, mais plus rapidement, de sorte que le front d'onde à la surface (en rouge) accélère. Sur les reconstitutions (b, c, d), l'écart entre deux images est de dix minutes et le diamètre final est égal à 250 000 kilomètres.

BRÈVES

Changement de sexe

Quand deux femelles de *Gobio histrio*, un petit poisson, se rencontrent au détour d'un massif de corail, l'une d'entre elles se transforme souvent en mâle. Dans des conditions symétriques, les mâles se transforment en femelle. Ce comportement est un avantage adaptatif : plutôt que courir la grande aventure, dans la mer dangereuse, à la recherche d'un partenaire, plutôt changer de sexe sur place !

Première lumière

Le premier des quatre télescopes du très grand télescope européen (VLT) est achevé. Le 24 mai, son miroir de 8,20 mètres de diamètre a capté pour la première fois la lumière d'un astre lointain. Un dispositif de 300 vérins hydrauliques soutient et ajuste la forme du miroir avec une précision de l'ordre du micromètre. Lorsqu'il sera achevé, en 2002, le télescope sera l'instrument le plus sensible au monde, capable de traquer les premières galaxies et les planètes extra-solaires.



Volcans réfrigérants

Une trentaine d'énormes éruptions volcaniques ont refroidi l'hémisphère Nord au cours des 600 dernières années. K. Briffa et ses collègues, de l'Université d'East Anglia, à Norwich, ont associé les températures estivales basses, enregistrées par les cernes des arbres, et les plus fortes éruptions volcaniques. L'amplitude du refroidissement indique la puissance de l'éruption (la quantité de cendres et de gaz qui interceptent le rayonnement solaire). L'éruption du Huaynaputina, au Pérou, en 1600, responsable d'un refroidissement de 0,8°C pendant l'été 1601, serait la plus forte éruption de cette période historique, supérieure à celles du Krakatoa en 1883 et du Pinatubo en 1991.

Dépistage précoce de la dyslexie

Le dépistage de la dyslexie, difficulté de lecture et d'écriture frappant dix pour cent des enfants scolarisés, est trop tardif. En systématisant le dépistage dès trois ans, les enfants dyslexiques suivront plus tôt une rééducation appropriée, avant l'apprentissage de la lecture au cours préparatoire. Plusieurs tests de dépistages existent pour les jeunes enfants, et ils sont bien acceptés parce que ludiques.

Suite page 22

Mots à gauche, images à droite

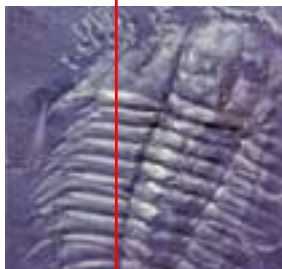
Comment mémorisons-nous les objets que nous voyons? Si ce sont des mots, plutôt avec l'hémisphère gauche de notre cerveau. Si ce sont des visages inconnus, que nous ne pouvons décrire simplement avec des mots, c'est l'hémisphère droit qui est le plus actif. Selon une équipe de l'Université Washington, les mémoires verbale et non verbale sont associées lorsque nous devons mémoriser des objets dont nous connaissons les noms, et nous nous souvenons mieux de ces derniers que de mots ou d'images seuls.

Cyclope et cholestérol

Les brebis qui mangent un lis sauvage mettent au monde des agneaux cyclopes. Or, cette plante contient une toxine dont deux équipes américaines viennent de montrer qu'elle perturbe la synthèse et le transport du cholestérol au cours de l'embryogenèse. Voici une nouvelle fonction pour cette molécule essentielle du métabolisme de l'adulte, et une nouvelle voie de recherche pour les anomalies de développement de l'embryon.

Une faune durable

Les plus anciens animaux connus, regroupés sous le nom de «faune d'Ediacara» sont des créatures au corps mou, dont les fossiles ont été préservés dans plusieurs formations sédimentaires de 600 à 544 millions d'années. Les paléontologues pensaient qu'ils s'étaient éteints ou transformés avant la diversification de la vie animale au début du Cambrien, il y a 540 millions d'années. On se trompait. Sören Jensen, de l'Université de Cambridge, et ses collègues, ont découvert, en Australie, des fossiles de la faune d'Ediacara dans les mêmes couches sédimentaires que des animaux apparus au Cambrien.

**Gauche, droite : à égalité**

Une équipe anglaise a montré que, contrairement à l'opinion commune, la durée de vie des gauchers n'est pas inférieure à celle des droitiers. L'enquête a porté, à neuf ans d'intervalle, sur six mille britanniques. Bien que les gauchers vivent dans un environnement conçu pour les droitiers, leur espérance de vie est analogue.

Suite page 24

De Mars à Tatahouine

Des structures minérales qui étayeraient l'hypothèse de vie sur Mars se forment aussi sur Terre.

En août 1996, une équipe de la NASA annonce la découverte de traces de vie dans une météorite d'origine martienne, nommée ALH84001 et ramassée en 1984 dans les glaces de l'Antarctique. Les indices de vie les plus convaincants sont des dépôts de carbonate et, dans ces dépôts, des structures allongées de quelques dizaines à quelques centaines de nanomètres de long. Ces structures pourraient être des fossiles de bactéries. Des extraterrestres prennent forme... microscopique.

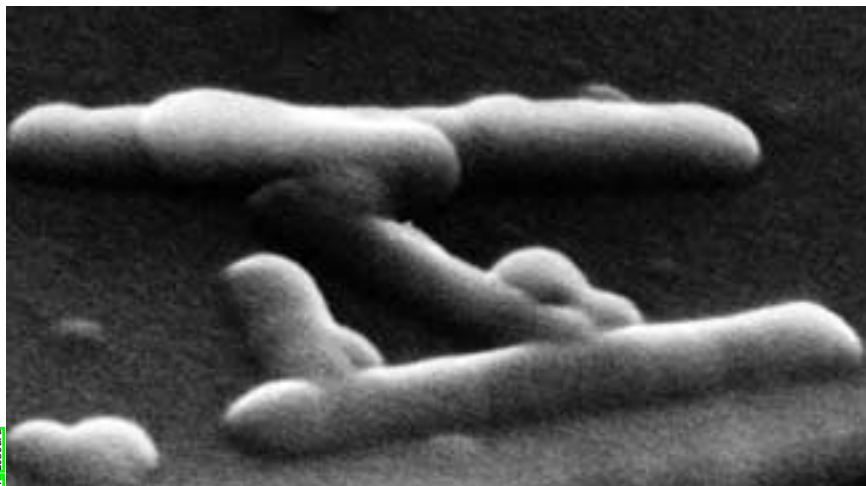
L'interprétation en termes de vie des structures allongées a suscité des polémiques. Bien que l'origine martienne des dépôts de carbonate ne fasse aucun doute, leurs conditions de formation restent âprement discutées : sont-elles compatibles avec la vie? De plus, on ne connaissait pas de forme de vie de si petite taille sur Terre. Or nous avons trouvé, dans les débris d'une météorite tombée il y a 67 ans dans le Sud de la Tunisie, des objets allongés tout à fait semblables à ceux de la météorite martienne, et qui pourrait correspondre à des restes d'organismes.

Revenons à l'histoire de la météorite de Tatahouine : dans la nuit du 27 juin 1931, les habitants de Tatahouine, en Tunisie, sont réveillés par un fort grondement, accompagné d'une lueur intense durant une demi-minute. Après quelques

recherches, les militaires français découvrent des fragments d'une roche curieuse dont les spécialistes reconnaissent l'origine extraterrestre. Ils en expédient quelques morceaux au Muséum national d'histoire naturelle de Paris, où ils sont toujours. En 1994, Alain Carion, un collectionneur, retourne sur le site d'impact et récolte de nouveaux fragments de la météorite.

La météorite de Tatahouine ressemble beaucoup à ALH84001. Elles sont toutes les deux composées presque exclusivement de cristaux de couleur vert olive, de l'orthopyroxène, un silicate de fer et de magnésium. En examinant au microscope électronique des échantillons collectés en 1994, nous avons eu la surprise de découvrir à l'intérieur de certaines fractures des agrégats de calcite, un carbonate de calcium. Ces agrégats, plus petits qu'une tête d'épingle, se sont formés sur la Terre : leurs compositions isotopiques en oxygène et en carbone sont quasi identiques à celles des carbonates du sol où la roche a séjourné, et on n'a observé aucun agrégat de calcite dans l'échantillon recueilli en 1931.

Ces carbonates sont particulièrement intéressants : les traces de vie supposées dans ALH84001 sont associées à de petits grains de carbonate, formés il y a plusieurs milliards d'années sur Mars, dans des conditions qui ne sont toujours pas connues. Ainsi, selon les hypothèses, leur température de formation s'échelonne de 0 à 700 °C. Jusqu'à présent, tous les modèles de formation des carbonates dans ALH84001 supposaient que la roche était saturée de fluides et que le dépôt avait eu lieu en une seule fois. Les observations faites sur la météorite de Tatahouine incitent à penser que les carbonates s'y sont formés lors de différents épisodes pluvieux,



Ces structures d'environ 500 nanomètres de long se sont formées sur Terre, dans les fragments d'une météorite qui ont séjourné dans le sol de la Tunisie pendant plus de 60 ans. Elles ressemblent aux «fossiles martiens» découverts dans une autre météorite.

au cours desquels le sol est partiellement détrempé. De tels dépôts sont instantanés à l'échelle des temps géologiques.

L'observation à fort grossissement révèle que les cristaux de carbonate ainsi que les surfaces des cristaux en contact avec le sol sont couverts d'objets cylindriques à terminaisons arrondies, de 100 à 600 nanomètres de longueur. Ces structures ressemblent à celles qui ont été trou-

vées dans ALH84001 et que l'équipe de la NASA a considérées comme des fossiles martiens. Jusqu'à présent, on n'en avait jamais observé ailleurs.

Dans le cas de Tatahouine, leur origine terrestre ne fait aucun doute. Il n'y en a pas dans les fragments prélevés en 1931, et les mêmes petits cylindres ont été observés sur d'autres objets récoltés dans le sol environnant, des gravillons et une vieille pièce de monnaie. S'agit-il

de précipités minéraux ou de restes d'organismes? Dans ce dernier cas, ils seraient deux à dix fois plus petits que les plus petites bactéries connues aujourd'hui. Des biologistes explorent aujourd'hui cette hypothèse. Si elle se confirme, une météorite aura, pour la première fois, permis la découverte de... nouveaux terriens.

J.A. BARRAT, Université d'Angers,
P. GILLET, ENS de Lyon

L'oreille absolue

Ce «don» a une composante génétique, mais ne s'exprime qu'avec une pratique précoce de la musique.

Mozart identifiait sans jamais se tromper n'importe quelle note émise par n'importe quel instrument, même isolée de tout contexte musical : il avait ce que l'on nomme l'oreille absolue. Ce don était-il inné? Était-ce une faculté qui résultait de sa pratique instrumentale précoce? Nous l'ignorons, mais quelques équipes ont précisé la question : elles ont montré que l'oreille absolue s'exprime préférentiellement chez les musiciens qui ont commencé à étudier la musique dans la petite enfance.

Siamak Baharloo et ses collègues de l'Université de San Francisco ont étudié 620 musiciens professionnels qui devaient reconnaître quarante notes par séries de dix, émises à trois secondes d'intervalle et pendant une seconde. Toutes les notes des sept premières octaves étaient indiquées sur un tableau, et les personnes testées

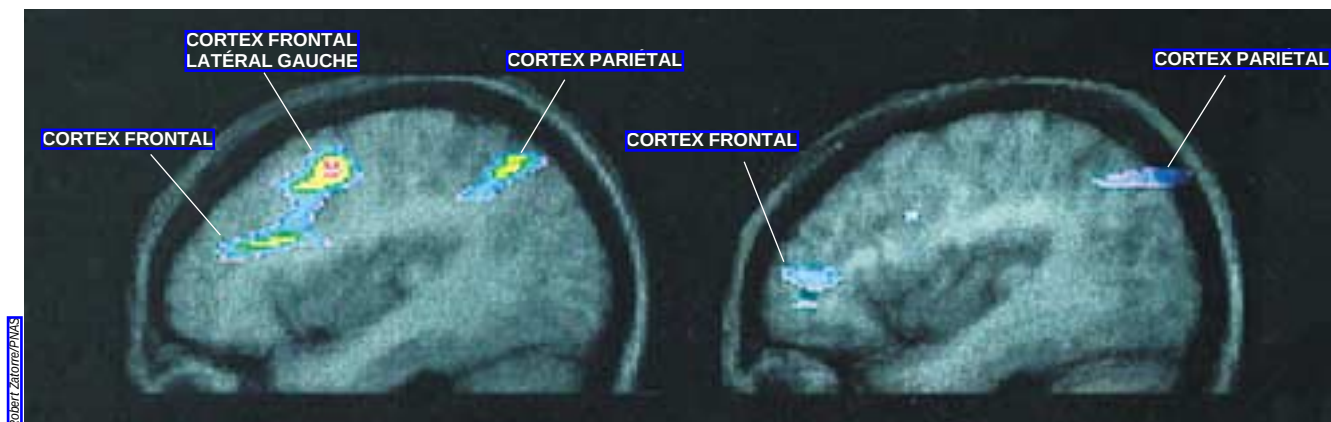
devaient identifier sur le tableau les notes qu'elles entendaient : la reconnaissance des notes est immédiate et infaillible chez les personnes qui ont l'oreille absolue.

Dans la population générale, la proportion des personnes ayant l'oreille absolue est inférieure à un pour cent, mais 15 pour cent des musiciens testés l'avaient. De surcroît, l'âge où les enfants commencent à étudier la musique influe sur leurs chances d'avoir l'oreille absolue quand ils seront adultes : la moitié de ceux qui ont commencé à apprendre la musique avant l'âge de quatre ans l'ont, le tiers de ceux qui ont commencé entre quatre et six ans et deux pour cent seulement de ceux qui ont commencé après 12 ans. Pour rechercher un éventuel gène de l'oreille absolue, l'équipe californienne va collecter des échantillons de sang de personnes appartenant à des familles où plusieurs membres ont l'oreille absolue.

Quels sont les mécanismes cérébraux spécifiques des personnes ayant l'oreille absolue? Robert Zatorre et ses collègues de l'Université McGill, à Montréal, ont utilisé la tomographie par émission de positons pour rechercher les bases neuronales de l'oreille absolue. Ils ont mesuré le flux sanguin cérébral

chez des personnes ayant l'oreille absolue à qui l'on faisait entendre des notes de musique et chez des musiciens qui n'avaient pas cette capacité. Comme prévu, l'audition se traduisait par une augmentation du débit sanguin dans les aires du cortex auditif des deux groupes. De surcroît, chez les personnes ayant l'oreille absolue, une région du cortex frontal latéral gauche est spécifiquement activée. Or cette région intervient dans l'apprentissage des associations réflexes. Ainsi, les personnes ayant l'oreille absolue attribueraient un nom à une note, sans faire appel aux zones cérébrales «de travail» : dès qu'ils entendent une note, ils la nomment spontanément, par une sorte d'association réflexe.

Simultanément, l'imagerie médicale fournit d'autres données : les personnes ayant l'oreille absolue ont une région du cerveau, le planum temporale, plus développée que la moyenne. Or, même si l'on ignore le rôle précis de cette zone, on sait qu'elle contient des neurones du cortex auditif dont les terminaisons atteignent la zone postérieure du cortex frontal latéral gauche, précisément la région activée lors de l'examen par tomographie par émission de positons, chez les personnes ayant l'oreille absolue.



Dans le cerveau des musiciens à qui l'on fait entendre des sons purs, les zones du cortex frontal et du cortex pariétal sont activées, comme le montre un examen par tomographie par émission de positons. Chez

les musiciens ayant l'oreille absolue (à gauche), une zone spécifique, localisée dans le cortex frontal latéral gauche, s'active. Ce n'est pas le cas chez les musiciens n'ayant pas l'oreille absolue (à droite).

Pontages hélicoïdaux

Les pontages coronariens seraient plus efficaces s'ils avaient une structure hélicoïdale : Colin Caro, biophysicien de l'Imperial College, à Londres, a montré que les vaisseaux présentant une telle structure facilitent l'écoulement du sang, tout comme une bouteille pleine d'eau se vide plus rapidement quand on fait tourner le liquide.

Bacille de la tuberculose

La tuberculose tue plus de trois millions de personnes dans le monde, continue à se répandre dans les pays en développement et réapparaît dans les pays industrialisés. Le génome du bacille de la tuberculose a été entièrement séquencé par l'équipe de Stewart Cole, de l'Institut Pasteur, et une équipe britannique, ce qui facilitera l'identification des gènes de résistance à certains antibiotiques, de latence (la bactérie se terre dans l'organisme infecté, mais n'est pas éliminée) et de virulence (la maladie est plus ou moins grave selon la souche contaminante).

Le déclin des mâles

La proportion de nouveau-nés garçons par rapport aux filles augmente après les guerres. Nul n'a pu expliquer ce phénomène en termes physiologiques, aussi a-t-on invoqué la providence divine qui pourvoyait ainsi au remplacement des soldats morts au champ d'honneur. Le surplus de garçons par rapport aux filles est en train de diminuer, faiblement, mais certainement, dans les pays développés. Pourquoi, nul ne le sait, mais l'absence de cause identifiée amène à accuser la pollution, l'anti-providence moderne.

Des parfums sur la toile

Un brevet (EP 831 384) a été déposé : l'objet de l'invention porte sur un dispositif distributeur de parfums, commandé par des instructions liées au message véhiculé sur l'Internet. Un texte traitant des forêts déclencherait ainsi la libération d'arômes forestiers préalablement encapsulés dans les ordinateurs, un article sur les océans serait accompagné d'odeurs iodées. Les déclarations passionnées des messageries roses auraient pour accompagnement des senteurs torrides de fleurs entêtantes.

Kampala : la campagne à la ville

Les crises politiques et économiques africaines modifient le paysage urbain.

Pour étudier les relations entre violence et urbanisation en Afrique, Bernard Calas, géographe à l'Université de Saint-Étienne, partit en 1988 à Kampala, capitale de l'Ouganda, où la guerre avait fait rage pendant 20 ans. Kampala avait été épargnée par les combats, mais la ville avait été occupée par l'armée qui pillait, violait et terrorisait la population. À ces maux s'ajoutait un effondrement économique dû à la guerre. Pourtant, la ville continua à croître, alimentée par un nombre croissant de ruraux déracinés, notamment ceux qui fuyaient les zones de combat.

Comment survivaient les citadins, anciens et nouveaux ? En cultivant d'abord leurs jardins. À l'exception d'un quartier de type européen qui rassemble les administrations, l'agglomération urbaine est parsemée de champs : 40 pour cent des ressources alimentaires kampalaises sont produites en ville. Les citadins ne pouvaient survivre à l'aide de leur seul salaire (le pouvoir d'achat associé au salaire minimal de 1984 équivalait à dix pour cent de celui de 1972), et recherchaient des revenus complémentaires : une classe de salariés-commerçants-cultivateurs est apparue. La culture assurait aussi l'approvisionnement urbain, l'acheminement des produits de la campagne étant handicapé par la détérioration du réseau

routier, la réduction du nombre de véhicules et les barrages militaires sur les routes.

Socialement, il n'y a pas non plus de rupture entre milieux ruraux et urbains : le réseau familial est étendu, et les familles sont réparties entre la ville et la campagne. Ces réseaux familiaux protègent citadins et ruraux : lorsque l'argent manque en ville, les ruraux envoient de la nourriture et accueillent les enfants des villes ; lorsque les combats dévastent les campagnes, ou en période de sécheresse, les citadins envoient matériel agricole et argent. Ces milieux sont complémentaires dans un réseau de « protection sociale » assuré par les familles (l'État ne dispense aucune aide).

Avec la reprise économique, commençant à Kampala en 1986, le retour de la confiance et des capitaux étrangers inversent-ils la situation, en redonnant à la ville un visage plus urbain au sens européen ? Non, au contraire : le système « agro-rurbain » mis en place pendant la guerre se perpétue, car les inégalités se creusent entre les classes sociales qui bénéficient de la reprise (employés qualifiés, commerçants) et les autres (petits fonctionnaires, ruraux déracinés, etc.). Le système agro-rurbain, qui produit des biens alimentaires bon marché, assure la survie des plus pauvres. Le tissu de solidarités familiales modère ainsi des effets négatifs du libéralisme.

B. Calas poursuit ses recherches dans d'autres villes d'Afrique orientale, telles Zanzibar et Dar-Es-Salaam, en Tanzanie, et Nairobi, au Kenya, qui souffrent de graves crises économiques, mais pas de violences. Les formes d'urbanisation de ces villes sont similaires à celles de Kampala : le développement des systèmes agro-rurbains n'est pas spécifique d'un contexte de violence.



Le paysage de la ville de Kampala s'est « ruralisé » lors de la crise militaire, politique et économique qui a sévi entre 1966 et 1986. L'auto-suffisance alimentaire a ainsi été assurée.

La masse du neutrino

L'expérience SuperKamiokande découvre les oscillations des neutrinos.

En 1931, Wolfgang Pauli postulait l'existence d'une particule dénommée neutrino dont l'énergie compensait le déficit énergétique observé dans la désintégration bêta des noyaux radioactifs. Sa masse, comme celle du photon, était alors considérée comme nulle, mais aujourd'hui, cette hypothèse est remise en question : une équipe américano-japonaise vient d'annoncer l'existence d'un neutrino massif.

Le neutrino est une particule neutre qui n'interagit avec la matière que par l'intermédiaire de la force nucléaire faible, l'une des quatre forces de la nature. Cette force n'agit qu'à très faible distance et l'étude du neutrino n'en est que plus délicate : la plupart des neutrinos traversent la Terre sans interagir.

La traque des neutrinos a commencé dans les années 1950 et l'on a découvert non pas un neutrino, mais trois, chacun associé à une famille de leptons, des particules insensibles à la force nucléaire forte. Les physiciens ont identifié trois familles de leptons : celle de l'électron, celle du muon et celle du tau. Chaque famille comprend des représentants chargés, comme l'électron, et des particules neutres, tel le neutrino.

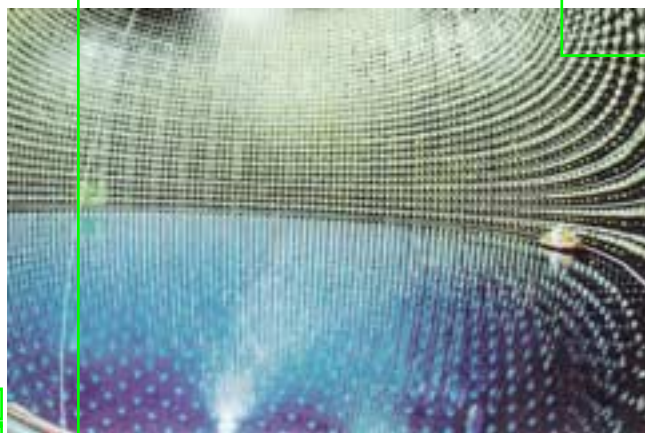
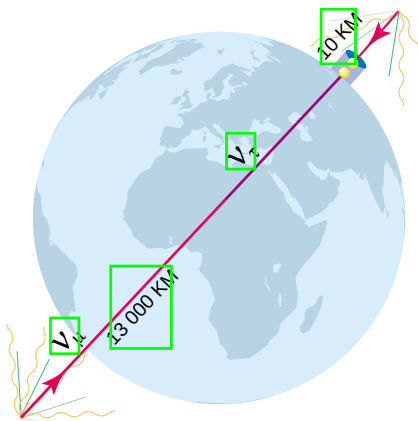
Lors d'une réaction nucléaire, un électron peut se transformer en un neutrino. Si les neutrinos ont une masse strictement nulle, la physique des particules stipule que ces échanges doivent se faire en famille : un neutrino électronique ne peut se changer en neutrino muonique. En revanche, si les neutrinos ont une masse, les échanges peuvent se faire « en dehors du clan ». Ainsi, un neutrino électronique peut se changer en neutrino muonique au cours du temps, et réciproquement. On nomme un tel changement une « oscillation des neutrinos ». La probabilité de changement dépend de deux facteurs : la différence des carrés des masses entre les neutrinos échangés et la durée entre l'instant de leur formation et l'instant de leur interaction (dénommée temps de vol).

Plus ce temps de vol est long, plus les neutrinos oscillent. Pour étudier des neutrinos dont le temps de vol varie, on

mesure la distance parcourue entre le point de formation du neutrino (sa source) et son point de disparition (l'interaction).

L'expérience *SuperKamiokande*, située au Japon, étudie les neutrinos atmosphériques. Lorsque les protons du rayonnement cosmique entrent en collision avec des atomes de la haute atmosphère, à dix kilomètres d'altitude, des gerbes de particules contenant des neutrinos sont émises. Le détecteur reçoit des neutrinos qui proviennent de toute l'atmosphère et qui ont parcouru des distances différentes : les neutrinos créés à la verticale du détecteur ont parcouru une dizaine de kilomètres, tandis que ceux créés aux antipodes du détecteur ont traversé toute la Terre, soit près de 13 000 kilomètres. Ainsi, la direction d'arrivée des neutrinos indique la distance parcourue, donc leur temps de vol, car on connaît leur vitesse.

Installé dans une mine, à 700 mètres de profondeur pour éviter la contamination par d'autres particules, le détecteur de *SuperKamiokande* est constitué d'une cuve cylindrique remplie de 50 000 tonnes d'eau très pure. Toutes les heures environ, un neutrino atmosphérique interagit



La cuve de 50 000 tonnes d'eau de *SuperKamiokande* (en bas), ici à moitié remplie, est recouverte de 13 000 photomultiplicateurs. Ces derniers enregistrent la lumière émise par l'interaction d'un neutrino venant de l'atmosphère avec un quark contenu dans un noyau d'oxygène de l'eau très pure de la cuve. Les neutrinos détectés sont créés dans l'atmosphère, vers dix kilomètres d'altitude (en haut). Plus la distance parcourue est grande, moins on détecte de neutrinos muoniques (ν_μ), car ils se sont transformés en neutrinos tauïques (ν_τ).

avec un quark contenu dans un noyau d'oxygène de l'eau. Une particule chargée de haute énergie, un électron ou un muon, est alors émise. Cette particule poursuit son chemin quelques mètres dans l'eau avant d'être absorbée.

Lors de ce déplacement, un cône de lumière est émis ; la lumière est alors enregistrée par quelques-uns des 13 000 photomultiplicateurs qui couvrent la paroi interne de la cuve. De l'analyse de cette lumière, on déduit la direction de la particule, son énergie et sa nature. Grâce à cette détermination, on sait quel type de neutrino a interagi dans la cuve, car un neutrino électronique produit un électron et un neutrino muonique, un muon.

Quels sont les résultats ? Connaissant, par une autre méthode, combien de neutrinos électroniques et de neutrinos muoniques sont produits dans l'atmosphère, les résultats de *SuperKamiokande* montrent que certains neutrinos muoniques produits disparaissent, tandis que les neutrinos électroniques sont stables (en accord avec les résultats de l'expérience de Chooz, publiés en 1997).

Les neutrinos muoniques disparaissent, parce qu'ils se transformeraient en neutrinos tauïques. Or, de tels échanges entre familles n'ont lieu que si les particules ont une masse. On en déduit que ces deux neutrinos ont une masse, mais aucune valeur ne peut être déduite de cette observation (au mieux connaît-on la différence du carré de leur masse).

Si les résultats de *SuperKamiokande* sont probants, ils supposent, toutefois, que l'on connaisse bien la source des neutrinos atmosphériques. Or, les incertitudes sur le flux de ces neutrinos sont voisines de 15 pour cent. Pour confirmer les oscillations, les physiciens projettent des expériences qui utiliseront des sources connues de neutrinos, tels les accélérateurs de particules.

La première étape aura encore lieu au Japon, où l'on dirigera un faisceau de neutrinos de l'accélérateur KEK, près de Tokyo, vers le détecteur de *SuperKamiokande* situé à 250 kilomètres. Le faisceau sera prêt au printemps 1999, et il faudra trois ans pour accumuler 600 événements. Un programme européen projette d'envoyer un faisceau entre le CERN et le tunnel du Gran Sasso, en Italie, à 730 kilomètres. Si les décisions sont prises rapidement, le faisceau et les détecteurs seront prêts en 2003.

Yves DECLAIS, Institut de physique nucléaire de Lyon

Né de père absent

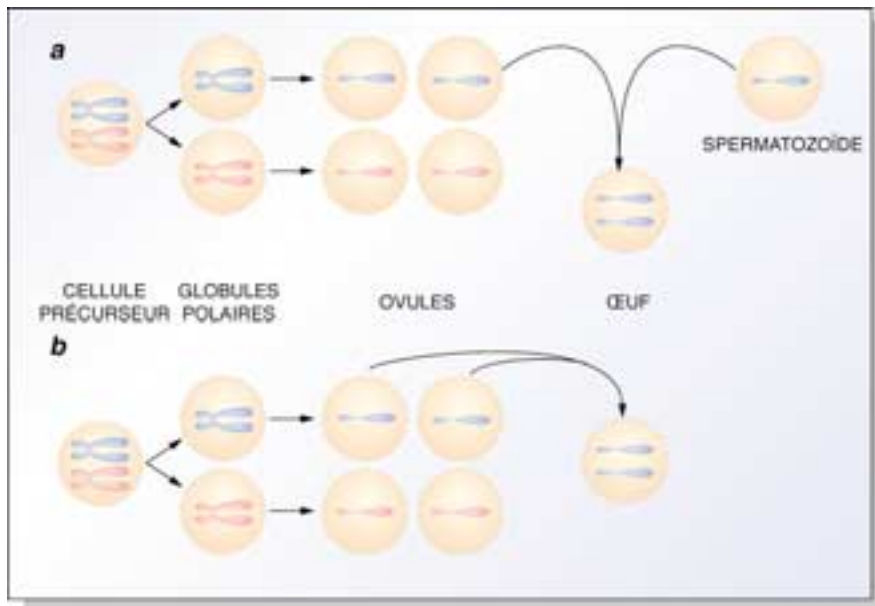
Des serpents femelles donnent naissance à des mâles sans s'accoupler.

Un animal naît généralement de l'accouplement de deux individus de sexes différents. Toutefois, dans certaines espèces, les femelles peuvent donner naissance à des jeunes sans le concours d'un mâle : c'est la parthénogenèse. On observe ce mode de reproduction chez des invertébrés (abeilles, pucerons...) et chez quelques vertébrés, dont plusieurs reptiles.

Ainsi le gecko, petit lézard du Pacifique Sud, est une espèce unisexuée qui se reproduit par clonage : tous les individus sont des femelles. Jusqu'à présent, la parthénogenèse n'était attestée que pour une seule des 2 800 espèces de serpents recensées. Des biologistes viennent de découvrir que cinq autres espèces peuvent aussi se reproduire par parthénogenèse, les femelles ne donnant alors naissance qu'à des mâles.

Des serpents femelles d'espèces bisexuées, maintenus en captivité en l'absence de mâles, produisent des jeunes. Fécondent-elles leurs ovules avec du sperme stocké dans leurs voies génitales, comme c'est souvent le cas chez les reptiles ? C'est peu probable, car les spermatozoïdes ainsi emmagasinés ne resteraient viables qu'une dizaine de mois, et l'une des femelles procréatrices n'a pas eu de contact avec un mâle depuis 14 ans. En outre, la comparaison de l'ADN des femelles et de leurs jeunes confirme l'absence d'éléments d'origine paternelle. Aucun doute possible : les jeunes sont issus de parthénogenèse.

Pourquoi, au contraire de ce qui se passe pour les lézards, tous ces jeunes serpents nés sans père sont-ils des mâles ? Dans le cas de la reproduction sexuée « normale », les cellules sexuelles des deux sexes, les gamètes, sont produites lors de la méiose à partir d'une cellule précurseur. Celle-ci, qui contient n paires de chromosomes (n dépend de l'espèce), se divise d'abord en deux globules polaires, qui contiennent chacun n chromosomes (un de chaque paire). L'un de ces globules est éliminé et, dans le second, les deux brins d'ADN qui composent chaque chromosome se séparent. Ce second globule se divise alors en deux nouveaux globules polaires, qui contiennent chacun n brins d'ADN. Chaque globule devient un gamète fonctionnel : un œuf résultera de la fusion de ce gamète avec un gamète de l'autre sexe.



Chez les reptiles, la femelle porte deux chromosomes sexuels différents (on n'a figuré que ces deux chromosomes, en bleu et en rouge, dans les cellules), et le mâle en porte deux identiques (ici, bleus). Lors de la production de cellules sexuelles par la femelle (a), une cellule précurseur (à gauche) se divise en deux « globules polaires », qui contiennent chacun un chromosome sexuel différent. Puis chaque globule polaire se divise à son tour, et forme un ovule fonctionnel. Dans plusieurs espèces de serpents, deux évolutions sont possibles. Si la femelle s'accouple avec un mâle, un œuf est formé par fusion de l'ovule avec un spermatozoïde (en haut à droite). En l'absence d'accouplement (b), les deux ovules issus du même globule polaire fusionnent et forment un œuf, qui contient ainsi deux chromosomes sexuels identiques : c'est un mâle.

Dans le cas des serpents parthénogénétiques, les deux ovules issus de la seconde division fusionnent, comme si l'un d'eux était un spermatozoïde, et forment un œuf. Chez les reptiles, la femelle porte deux chromosomes sexuels différents, Z et W, tandis que le mâle possède deux chromosomes sexuels identiques (ZZ) (chez les mammifères, en particulier chez l'homme, c'est le mâle qui porte deux chromosomes sexuels différents). Lors de la méiose, la division de la cellule précurseur du gamète produit deux globules polaires qui ne contiennent chacun que l'un des deux chromo-

somes sexuels : Z ou W. Les deux globules formés par la seconde division contiennent donc tous les deux le même chromosome sexuel. La fusion de ces deux globules produit un œuf de type ZZ ou WW. Le développement d'un œuf qui contient deux chromosomes sexuels Z produit un mâle. Les œufs qui contiennent deux chromosomes W, en revanche, ne correspondent à aucun des deux sexes et ne se développent donc pas : on observe effectivement une forte proportion d'œufs stériles chez ces femelles qui ne s'accouplent pas.

Catherine PERRIN

Moussons inattendues

Des moussons vieilles de 525 000 ans remettent en question les modèles d'évolution des climats.

Depuis un million d'années, les pluies tropicales d'été, nommées moussons, sont renforcées tous les 20 000 ans environ (quand la Terre passe au plus près du Soleil pendant l'été de l'hémisphère Nord), pendant quelques milliers d'années. Martine Rossignol-Strick (CNRS-Université Paris VI), Martine Paterne et Franck Bassinot (CEA)

ont découvert que les moussons africaines et indiennes ont été énormes il y a 528 000 - 525 500 ans, à un moment où elles auraient dû être à peine renforcées. Cette anomalie serait le signe d'un événement climatique exceptionnel à l'échelle planétaire.

La Terre a enregistré l'amplitude de ces moussons dans les sédiments marins. M. Rossignol-Strick et ses collègues ont étudié une carotte forée en Méditerranée, à l'Est de la Sicile. Dans cette carotte, les sédiments généralement clairs sont entrecoupés de quelques couches noires, très riches en matière organique, qui indiquent que des moussons exceptionnellement fortes ont frappé l'Afrique de l'Ouest.

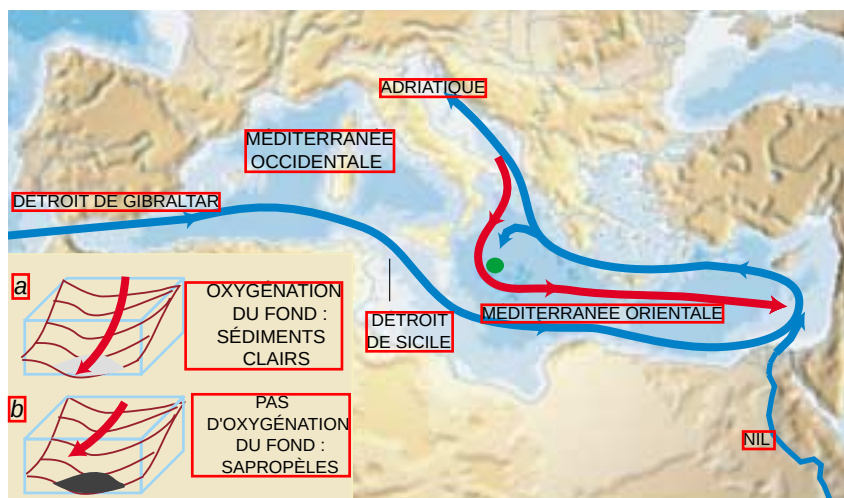
Dans ce bassin oriental de la Méditerranée, l'eau de surface vient de l'océan Atlantique. En revanche, la circulation d'eau profonde est bloquée par les détroits, et l'eau de fond résulte de l'enfoncement, en hiver, de l'eau dense (froide et salée) de la surface de l'Adriatique. Cette eau alimente le fond du bassin en oxygène et assure la survie des bactéries aérobies. Ces bactéries oxydent la quasi-totalité des déchets organiques, de sorte que les sédiments au fond de la mer sont de couleur claire.

Sur le Nord de l'Éthiopie, les pluies de la mousson d'Afrique de l'Ouest alimentent la crue du Nil, qui se déverse au Sud-Est de la Méditerranée. Quand la mousson est particulièrement forte, l'eau de ces crues, mélangée à l'eau de mer, est entraînée jusqu'à l'Adriatique. L'apport d'eau douce est alors tel que l'eau de surface n'est plus assez dense pour s'enfoncer beaucoup : le fond de la mer n'est pas assez oxygéné pour que les bactéries aérobies survivent. Les bactéries anaérobies, qui les remplacent, ne dégradent pas les débris organiques les plus résistants, qui s'accumulent et colorent en noir le sédiment, nommé alors sapropèle.

Afin de dater ces sédiments, on relie l'alternance des sédiments clairs et noirs à celle des périodes glaciaires et interglaciaires dont on connaît les dates. Pour cela, on prélève les coquilles de certains micro-organismes marins dans les carottes sédimentaires. Leur analyse renseigne sur la proportion d'oxygène lourd par rapport à l'oxygène normal, dans l'eau des océans, proportion qui dépend du volume total des glaces.

Ainsi, on a trouvé que le dépôt d'une des couches de sédiments noirs de Méditerranée a commencé il y a 528 000 ans, et s'est terminé il y a 525 000 ans. Pourtant, à cette époque, l'insolation d'été était relativement faible dans l'hémisphère Nord, et les pluies de mousson n'auraient pas dû être si abondantes sur l'Afrique de l'Ouest : habituellement, seule une forte insolation d'été dans l'hémisphère Nord, en accentuant les différences de pression entre l'hémisphère Sud froid et des zones très chaudes de l'hémisphère Nord, renforce les vents de mousson responsables des pluies.

En outre, l'étude de sédiments prélevés dans l'océan Indien a révélé de fortes moussons indiennes, exactement à la même période. Il reste aux paléoclimatologues à trouver le phénomène climatique responsable de ces énormes moussons africaines et indiennes, que n'explique pas l'intensité de l'insolation terrestre.



Circulations des eaux méditerranéennes : les flèches bleues représentent la circulation des eaux de surface, et les flèches rouges les eaux profondes. En hiver, les eaux froides et salées de l'Adriatique s'enfoncent : la matière organique est oxydée (a). En cas de moussons d'Afrique exceptionnellement fortes, les abondantes eaux de pluies diminuent la densité des eaux de surface : les eaux de l'Adriatique s'enfoncent moins, et n'apportent plus l'oxygène aux matières organiques qui ne sont plus complètement oxydées. La fraction non oxydée s'accumule dans les sédiments comme ceux découverts à l'Est du détroit de Sicile (point vert).

Exploration virtuelle du côlon

Une méthode non invasive dépiste les polypes, sources potentielles de cancer du côlon.

Le dépistage précoce est l'une des conditions d'un traitement efficace des cancers. Dès lors, comment détecter les signes avant-coureurs de la maladie ? Une nouvelle méthode permet une exploration virtuelle du côlon, reconstitué à partir d'images prises au scanner.

Le cancer du côlon est la deuxième cause de mortalité par cancer en France, et les polypes en sont un indice précurseur. Ces petites tumeurs, de quelques millimètres de diamètre, bénignes au début, dégénèrent souvent en cancer. Pour les dépister, on recherche du sang dans les selles, mais le test (*Hémocult*) est peu sensible et peu fiable.

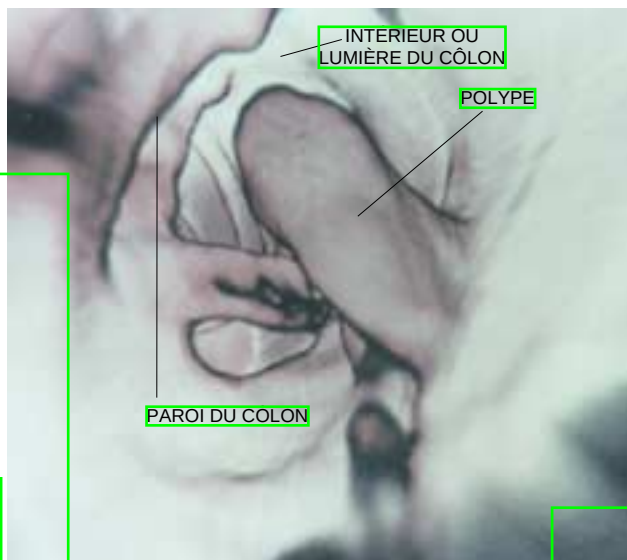
Les gastro-entérologues espèrent disposer bientôt de la coloscopie virtuelle, née de l'association du scanner à rayons X et de logiciels informatiques performants. L'endoscopie est une méthode fiable, mais qui présente quelques inconvénients : on introduit dans le côlon, par l'anus, un endoscope, une sorte de caméra miniaturisée équipée d'un minuscule scalpel qui sert à l'ablation des polypes visualisés grâce à la caméra.

Cette méthode n'est pas envisageable pour un dépistage de masse, car elle nécessite une anesthésie, qui augmente le coût de l'opération et n'est pas dénuée de risques. De surcroît, exceptionnellement, l'endoscope risque de perforer la paroi colique.

Au contraire, un examen par endoscopie virtuelle se déroule comme un scanner abdominal classique, sans injection d'un produit améliorant les contrastes. Comme lors d'une endoscopie, la personne subit une préparation pour nettoyer le côlon des selles ; au début de l'examen, le côlon est distendu par de l'air à l'aide d'une sonde introduite par l'anus, et le médecin administre un médicament qui relaxe les muscles du côlon. L'abdomen est ensuite exploré par scanner X : la table où est placé le patient avance lentement, pendant que le tube à rayons X tourne autour. Des détecteurs enregistrent les rayons X qui traversent l'organisme et fournissent les données nécessaires à la reconstruction des images des « coupes » de l'intestin. L'examen dure une vingtaine de minutes. Ensuite, la coloscopie virtuelle est réalisée sur les données stockées par l'ordinateur, tandis que la personne examinée est déjà repartie. Les programmes informatiques reconstituent l'abdomen et, notamment le côlon, en trois dimensions. Le médecin étudie à l'aide d'un logiciel adapté cet abdomen selon toutes les incidences qui l'intéressent. Il peut même observer l'inté-

rieur du côlon, grâce à une lumière virtuelle qu'il y déplace. Les images brutes sont en noir et blanc, comme des radiographies, mais le système les colore artificiellement.

Dans ces conditions, les polypes apparaissent comme des excroissances dans le côlon. On détecte aisément les polypes de plus de 10 millimètres de diamètre, mais la sensibilité est médiocre pour les polypes de moins de cinq millimètres (on n'en détecte qu'un sur trois environ). Une des principales difficultés tient aux risques de confusion entre un polype et un résidu fécal. Selon Gérard Schmutz, du Service de radiologie du Centre hospitalo-universitaire de Caen, qui évalue cette nouvelle méthode, la coloscopie virtuelle devrait être utilisée lors des premières explorations des malades à risques. En cas de doute, si l'on détecte un polype à l'endoscopie virtuelle, on doit pratiquer une endo-



Grâce à la coloscopie virtuelle, on observe l'intérieur du côlon sans y faire pénétrer aucun endoscope.

scopie classique, plus sensible, et qui, seule, permet l'ablation des polypes.

Simultanément, poursuivant leur quête de méthodes d'exploration ou de traitement peu invasives, les méde-

cins cherchent comment détruire les polypes par voie externe. Ils étudient, par exemple, des substances qui se fixeraient spécifiquement sur les polypes et qui, quand on les activerait par un faisceau laser, libéreraient un facteur nécrosant le polype sur lequel ils se sont fixés. D'autres études visent à étendre le domaine d'application de l'endoscopie virtuelle du côlon à toute structure tubulaire, par exemple la trachée, l'œsophage ou encore les vaisseaux sanguins.

Un examen par coloscopie virtuelle coûte environ trois fois moins cher qu'une coloscopie classique. Il est dénué de risques, mais l'irradiation par les rayons X équivaut à neuf radiographies pulmonaires. Malgré cet inconvénient, la coloscopie virtuelle devrait devenir une méthode de prévention du cancer du côlon.

Séisme en Afghanistan

Aura-t-il un successeur?

Il y a deux mois, un séisme de magnitude 6 avait ébranlé les montagnes de l'Hindu Kush, en Afghanistan. Le 30 mai 1998, un nouveau séisme de même intensité a frappé à nouveau, durant la nuit : 4 000 personnes ont péri, ensevelies dans leurs habitations qui avaient été sans doute fragilisées par le premier séisme.

Pourquoi cette répétition? Pendant une semaine, les seules informations dont ont disposé les géologues ont été les enregistrements sismologiques du réseau mondial, et les relevés tectoniques effectués lors de précédentes missions. La géologie de la région est déterminée par le mouvement de deux plaques tectoniques : l'Inde, qui comprime l'Asie à la vitesse de quatre à cinq centimètres par an depuis 40 millions d'années, s'enfonçant en partie sous elle, provoque la surrection de l'Himalaya et du Tibet ; la collision des deux plaques découpe dans l'Asie de gros blocs, limités par des failles. On avait étudié certains de

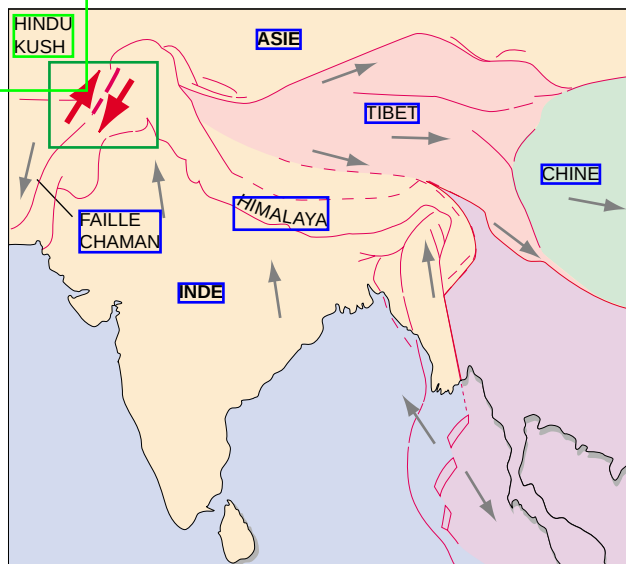
ces blocs qui se déplacent vers l'Est (tel le Tibet) ; le séisme du 30 mai 1998 concerne un bloc à l'Ouest de l'encoche produite par l'enfoncement de l'Inde dans le continent asiatique.

Comment ce séisme est-il né? Le réseau mondial de sismomètres, dont fait partie le réseau français Géoscope, a pu identifier la position et la profondeur du séisme, superficiel. En déterminant le sens des premiers mouvements du

sol et en mesurant l'amplitude des ondes sismiques guidées par la surface du Globe, on a déterminé l'orientation de la faille qui a coulissé : le plan de décrochement orienté Nord-Nord Est/Sud-Sud Ouest s'étend sur une vingtaine de kilomètres de longueur, sur dix kilomètres de profondeur, et la rupture de ce plan s'est répercutée en surface. Le décrochement a atteint environ un mètre, le long d'une cassure irrégulière.

A quoi servent de telles informations? Tout d'abord, les tectoniciciens ont été intéressés d'observer que la faille qui a coulissé est dans le prolongement d'un système de failles déjà connu, plus au Sud (faille de Hodarab), ainsi que dans le prolongement direct d'une très grande faille qui, encore plus au Sud (la faille de Chaman), forme la limite occidentale des plaques indienne et asiatique. D'autre part, comme la faille fautive n'est pas exactement au même endroit que celle qui a rompu il y a deux mois (80 kilomètres plus au Nord) et que son mécanisme est identique, on peut imaginer qu'une zone encore sous contrainte, entre les deux, risque de lâcher bientôt.

R. ARMIGO ET P. BERNARD, Institut de physique du Globe

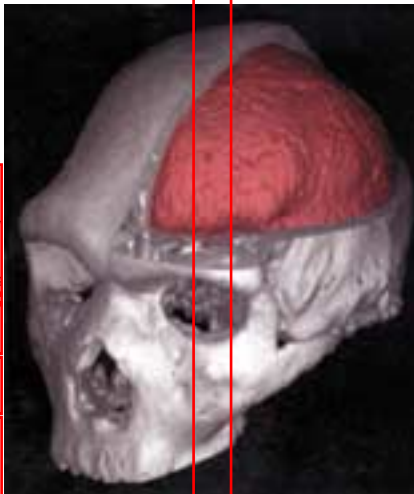


L'Inde, qui heurte l'Asie, crée des chaînes de montagnes et repousse de gros blocs (en couleurs) limités par des failles (en rouge). Le séisme du 30 mai a eu lieu à l'endroit encadré en vert ; les flèches rouges indiquent le décrochement détecté par les sismomètres.

Crânes virtuels

Les fossiles sont les seuls témoins directs de nos ancêtres. Toutefois, les conservateurs de musées prêtent difficilement ces Pierres de Rosette. Les crânes fossiles, en outre, renferment beaucoup d'informations inaccessibleles par des méthodes traditionnelles. Aujourd'hui, les images tridimensionnelles (obtenues à l'aide de scanners identiques à ceux utilisés par les médecins), la réalité virtuelle et les techniques de modélisation par ordinateur viennent à la rescousse des paléo-anthropologues, qui réalisent ainsi des analyses impossibles à pratiquer sur les fossiles originaux. Les parties manquantes d'un crâne peuvent être reconstruites à l'aide des parties conservées, des déformations postérieures à la mort de l'individu corrigées, et de minuscules structures cachées (par exemple de l'oreille interne) grossies. De plus, les anthropologues reconstituent, sur leur écran, des fossiles brisés.

On mesure aussi, virtuellement, tous les paramètres accessibles par les mesures classiques, et, souvent, avec la même précision. En créant un moule virtuel de l'intérieur du crâne, on détermine le volume du cerveau : Glenn Conroy, de l'Université de Washington, a ainsi montré que le volume du cerveau d'un australopithèque de 2,5 millions d'années avait été surestimé de près de 20 pour cent. Les estimations des capacités crâniennes d'autres fossiles pourraient également être incorrectes : selon les nouveaux résultats obtenus avec l'imagerie virtuelle, il faudra peut-être reconsidérer les modèles d'évolution du cerveau.



Ce crâne fossile virtuel d'un homme archaïque de Petralona, en Grèce, révèle les structures cachées comme les sinus. Un moule de l'intérieur du crâne permet l'étude du cerveau.

Après les images tridimensionnelles, les anthropologues vont découvrir la réalité virtuelle : un modèle en résine transparente est réalisé, par stéréolithographie, à partir d'une image informatique. La résine est polymérisée par un faisceau laser, couche après couche, de sorte que toutes les structures internes, notamment la boîte crânienne, les sinus et les canaux où passaient les nerfs, sont reproduits. Des crânes qui se ressemblent extérieurement peuvent cacher des différences internes, de sorte que l'étude des caractéristiques morphologiques internes des crânes fossiles apportera sans doute des éléments nouveaux au débat sur les origines de l'homme de Néandertal et de l'homme moderne.

D'autre part, selon une hypothèse récente, fondée sur la taille d'un canal qui innerve la langue, les Néandertaliens auraient eu une capacité de langage proche de la nôtre ; la réalité virtuelle aidera les anthropologues à tester cette hypothèse.

Les paléontologues envisagent de créer des archives numériques, contenant les images des principaux fossiles du monde entier. Grâce à l'Internet et à un logiciel approprié, ils pourront étudier n'importe quel fossile, faire leurs propres mesures, et, éventuellement, réaliser un modèle par stéréolithographie. Ces techniques d'imagerie ne remplaceront pas l'observation directe des fossiles ou de leurs moulages classiques, seul moyen d'obtenir une précision submillimétrique ; l'imagerie complétera ces méthodes traditionnelles, surtout quand son coût aura notablement diminué.

Épiderme de culture

Les grands brûlés peuvent être sauvés par des autogreffes de cellules d'épiderme multipliées en culture.

Une personne brûlée profondément sur plus des deux tiers de la peau n'a qu'une chance sur dix de survivre. Peut-être parce qu'on ne s'occupe pas assez tôt de sa peau. Les grands brûlés sont aussi de grands blessés et les réanimateurs tentent de sauver le blessé avant de traiter ses brûlures. Or, selon Hervé Carsin de l'Hôpital d'instruction des armées Percy, à Clamart, les chances de survie dépendent en tout premier chef du traitement de la peau brûlée : comme cette dernière libère des toxines dans l'organisme, on doit la retirer chirurgicalement le plus rapidement possible. L'équipe de l'Hôpital Percy a montré, au

terme d'un suivi de 30 personnes, que les chances de survie des grands brûlés traités par la technique de greffe d'épiderme de culture atteignent 90 pour cent.

Pour traiter les malades, les chirurgiens retirent la peau brûlée et prélèvent, quand ils le peuvent, des fragments de



Des fragments d'épiderme de culture, d'environ 70 centimètres carrés, ont été placés sur l'abdomen de cette personne gravement brûlée.

peau indemne, qu'ils étirent jusqu'à former une sorte de filet dont ils recouvrent les plaies. La multiplication des cellules épidermiques assure la reconstitution de l'épiderme. Que faire quand la surface de peau indemne est insuffisante pour recouvrir les plaies ? On ne peut pas greffer des cellules épidermiques d'un donneur, car, contrairement aux cellules du derme, les cellules de l'épiderme activent le système immunitaire, qui les rejette. Les cellules greffées doivent donc provenir du brûlé lui-même.

L'équipe de Clamart prélève deux biopsies de un à deux centimètres carrés (généralement, il reste quelques zones indemnes dans les plis cutanés). Les biopsies sont envoyées aux États-Unis, à la Société *Genzyme Tissue Repair*, spécialisée dans la culture d'épiderme pour les grands brûlés. Les cellules épidermiques, les kératinocytes, prélevées sur le brûlé sont mises en culture avec des facteurs de croissance. Chaque kératinocyte engendre un clone de cellules filles qui lui sont identiques.

On obtient ainsi des feuillets (deux à huit) de kératinocytes d'environ 10 centimètres sur sept ; il faut une vingtaine de jours pour que la culture ait la surface nécessaire. Pendant cette période, on recouvre le patient de peau prélevée sur des personnes décédées. Lorsque l'épiderme de culture arrive à Clamart, les chirurgiens retirent l'épiderme du donneur et le remplacent par l'épiderme de culture (le derme du donneur reste en place).

L'étude menée pour examiner l'efficacité de la nouvelle technique de greffage de l'épiderme cultivé a porté sur sept enfants et 23 adultes, de deux ans et

demi à 70 ans. Ces malades étaient brûlés sur 50 à 95 pour cent de la surface corporelle. Sur certains il a fallu utiliser 210 greffons d'épiderme (de 75 centimètres carrés environ). Dans le meilleur des cas, 95 pour cent de l'épiderme greffé a adhéré au derme ; dans le pire des cas, 25 pour cent seulement de la greffe a «pris». La proportion de greffes réussies a été supérieure chez les enfants. Selon H. Carsin, les greffes d'épiderme de culture sont indispensables lorsque les lésions profondes dépassent 70 pour cent de la surface corporelle.

Sur les 30 personnes greffées à l'Hôpital Percy, trois sont décédées au cours

des deux mois qui ont suivi l'accident, tandis que les estimations ne prévoyaient que trois survivants. Le handicap lié à ces greffes est lourd : le néoépiderme est fragile, adhère parfois insuffisamment au derme, de sorte que des bulles apparaissent, sources de plaies et d'infections ; les personnes restent hospitalisées pendant environ quatre mois.

Sur les 3 000 grands brûlés recensés chaque année en France dans les centres spécialisés, une cinquantaine, parmi les plus gravement atteints, pourraient bénéficier de la technique de l'Hôpital Percy, qui ne peut en traiter que sept par an.

Machiavel aurait-il pu être un singe ?

Mensonge et bluff chez les animaux.

Mentir est-il l'apanage des seuls êtres humains ? En France la tradition cartésienne place l'homme au-dessus du règne animal, et il semblait inconcevable qu'un «animal» puisse user de subterfuges. Pour mentir, il faut avoir conscience que des mots ou expressions peuvent modifier le comportement d'autrui, pouvoir distinguer sa connaissance de celle d'un autre, et comprendre les désirs, les intentions et les convictions d'un congénère. Seul l'être humain était doté de telles capacités, qui apparaissaient subitement, pensait-on, à l'âge de cinq ans. Or il semblerait que des enfants plus jeunes mentent, et de jeunes singes aussi.

Dans le test de Sally-Anne, fondement des études, le sujet du test voit un personnage, baptisé Sally, placer l'objet dans une première cache, puis disparaître. Un second personnage, Anne, apparaît, place l'objet dans une des autres caches et s'en va. Sally revient. On demande alors au sujet d'indiquer l'endroit où Sally pense que l'objet est placé. S'il désigne l'endroit où est effectivement abrité l'objet, c'est-à-dire où Anne l'a placé, il ne distingue pas sa connaissance de celle d'un autre (Sally) et échoue au test : mentir lui est impossible.

Récemment, V. Reddy et P. Newton, de l'Université de Portsmouth, ont montré que des

enfants de trois ans seulement réussissent le test et, de plus, étaient capables de comportements plus élaborés, comme de feindre l'indifférence envers un objet convoité pour ne pas attirer l'attention d'un frère ou d'une sœur.

En 1994, grâce à l'enregistrement vidéo des enfants, des pédiatres ont noté que la bonne réponse ne dépend pas des facultés intellectuelles, mais plutôt d'un problème de verbalisation : quatre-vingt dix pour cent des enfants de trois ans regardent vers le bon emplacement où Sally avait caché l'objet.

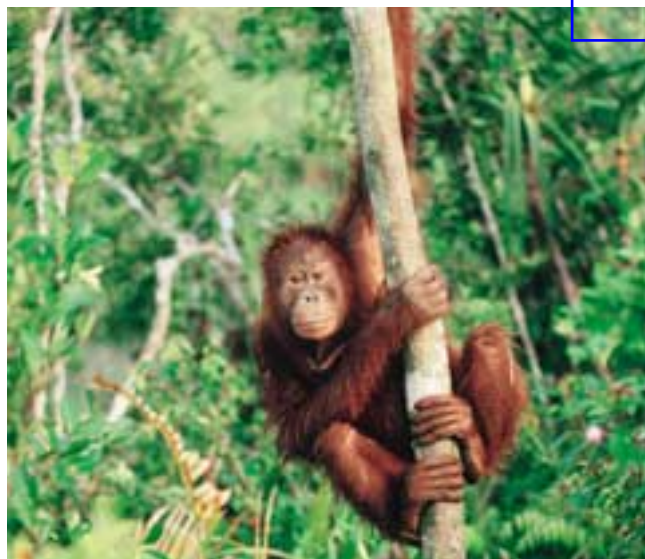
Pour les singes, le spécialiste de leur langage, D. Premack, a publié, en 1979, une étude montrant des chimpanzés capables, lorsqu'ils sont en concurrence avec un homme, d'essayer de le tromper, ce qu'ils ne font pas en situation de coopération. Toutefois, ces conditions artificielles ne parviennent pas à convaincre les spécialistes.

D'autres tromperies ont été observées chez les singes. Par exemple, un

singe, acculé par ses congénères en colère regarde l'horizon et émet un signal d'alarme. La troupe se disperse dans les arbres environnants, la tension retombe, mais aucun prédateur n'apparaît... Comportement issu d'une série d'essais-erreurs ou mensonge prémédité ? Est-il possible de distinguer une tromperie intentionnelle d'un comportement acquis ? La question reste en suspens.

M. Tomasello, utilisant une version non verbale du même test, compare les aptitudes au mensonge des chimpanzés, des orangs-outans et des enfants de quatre ou cinq ans. Il conclut, contrairement à V. Reddy, que seuls les humains, à partir de cinq ans, distinguent ce qu'ils savent de ce que les autres savent. La polémique rebondit, mais M. Tomasello concède que les singes anthropoïdes peuvent posséder les rudiments nécessaires pour une manipulation d'autrui sans pour autant qu'elle soit décelable par les tests. Tout le problème résidait peut-être dans la méthode employée : on ne peut pas traiter de la même façon humains et non-humains.

Le mensonge se serait-il développé progressivement, au cours de l'évolution, tout comme un organe spécialisé ? Des rudiments de mystification, communs à tous les primates, ont-ils permis à nos ancêtres, très vulnérables au sortir de la forêt, de survivre et ainsi de nous permettre d'être là aujourd'hui ? La communauté scientifique reste toutefois très sceptique envers ce nouvel assaut contre l'exception humaine. Peut-on réellement assimiler à l'animal sur la base du mensonge ? Le Prince est-il l'aboutissement de millions d'années d'évolution ?



Sait-il mentir ?

LOÏC MANGIN

La contraception masculine

SOPHIE CHRISTIN-MAITRE • PHILIPPE BOUCHARD

Le contraceptif masculin idéal devra être rapidement efficace, dépourvu d'effets secondaires et réversible. Il ne devra pas perturber l'activité sexuelle. Bien que ces exigences soient difficiles à concilier, plusieurs substances contraceptives sont prometteuses.

Un contraceptif efficace inhibe réversiblement la capacité de reproduction, sans perturber l'activité sexuelle. La contraception hormonale destinée aux femmes a notablement progressé au cours des années 1960, mais pourquoi la contraception masculine a-t-elle pris tant de retard ? Parce que la fonction des testicules est double : ils produisent les spermatozoïdes et sécrètent les hormones masculines. Toute la difficulté tient à cette dualité. Bien que les hommes soient encore peu nombreux à demander une contraception, plusieurs équipes recherchent le contraceptif masculin idéal, qui devra être efficace aussi vite que possible après son administration, réversible, facile à utiliser et bon marché. Examinons le contexte physiologique où doit agir un contraceptif masculin.

La physiologie testiculaire

Les testicules contiennent des cellules de Leydig et des cellules de Sertoli. Les premières représentent moins de cinq pour cent des différentes cellules. Elles sont situées dans l'espace interstitiel testiculaire, c'est-à-dire entre les tubules séminifères, où sont libérés les spermatozoïdes matures. Elles assurent surtout la production de la testostérone. Cette hormone est la principale des hormones masculines (ou androgènes). Elle est responsable de l'apparition, à la puberté, des caractères sexuels secondaires : la musculature de type masculin, la pilosité, la voix grave. Son rôle est essentiel dans le désir sexuel.

Les cellules de Sertoli constituent l'essentiel des cellules testiculaires. Elles tapissent les tubules séminifères, de petits tubes creux de 250 micro-

mètres de diamètre environ, où se développent les spermatozoïdes. Les cellules de Sertoli servent de cellules nourricières aux spermatozoïdes, mais aussi de barrière hémato-testiculaire, entre le sang et les cellules germinales. Cette dernière protège les cellules souches, qui produisent les spermatozoïdes, des substances sanguines toxiques.

Chez l'adulte, le développement des spermatozoïdes, ou spermatogénèse, se déroule dans les tubules séminifères en 74 jours environ chez l'homme, et selon trois phases : la phase proliférative, qui correspond à la maturation des cellules souches, ou spermatogonies, en spermatocytes ; une phase de multiplication et de maturation des spermatocytes ; la phase de spermatogénèse à proprement parler, où les spermatocytes se métamorphosent en spermatides, puis en spermatozoïdes. La quatrième phase correspond à la libération des spermatozoïdes dans les tubules séminifères.

La fabrication des spermatozoïdes est continue : un homme produit 30 millions de spermatozoïdes par jour. Quand ils quittent le testicule, les spermatozoïdes sont immobiles, incapables de féconder un ovocyte. Ils s'accumulent dans l'épididyme, où leur maturation leur confère mobilité et pouvoir fécondant. Lors de l'éjaculation, les spermatozoïdes empruntent le canal déférent. Le nombre de spermatozoïdes dans l'éjaculat normal est compris, en moyenne, entre 25 et 100 millions par millilitre, soit 50 à 300 millions par éjaculat.

La production des spermatozoïdes n'est pas autonome : la fonction testiculaire est commandée par deux hormones, d'une part la FSH, ou hormone

de stimulation des follicules et, d'autre part, la LH, ou hormone lutéinisante. Ces deux hormones, nommées gonadotrophines, sont synthétisées et sécrétées par l'hypophyse, une glande localisée à la base du cerveau. Pour être actives, l'hormone lutéinisante et l'hormone de stimulation des follicules doivent se lier à leurs récepteurs membranaires spécifiques situés sur les cellules-cibles : la LH agit surtout sur les cellules de Leydig, tandis que la FSH agit sur les cellules de Sertoli.

Quand elle se lie à ses récepteurs, sur les cellules de Sertoli, la FSH stimule la synthèse de diverses protéines qui commandent la différenciation et la production des spermatozoïdes. Quand la LH se lie à ses récepteurs membranaires, sur les cellules de Leydig, elle stimule la libération de testostérone intratesticulaire (une partie de l'hormone circule dans le sang, mais une proportion notable reste localisée dans les testicules).

Les deux systèmes FSH-cellules de Sertoli et LH-cellules de Leydig sont liés, puisque les deux hormones sont stimulées par la gonadolibérine GnRH (de l'anglais *Gonadotropin releasing hormone*, ou hormone qui libère les gonadotrophines). La GnRH, une protéine d'origine hypothalamique, est libérée par bouffées, toutes les 60 à 90 minutes, chez l'homme. Dans l'hypophyse, elle se lie à ses récepteurs, portés par les cellules gonadotropes, qui libèrent la FSH et la LH. Un excès de testo-

1. LA PRODUCTION DES SPERMATOZOÏDES, dans les testicules, est commandée par les hormones FSH et LH produites par l'hypophyse (flèche rouge), sous le contrôle de la GnRH, synthétisée par l'hypothalamus. La testostérone, produite dans les testicules, exerce une rétroaction sur la production de GnRH (flèche orange), donc sur sa propre production.

stérone inhibe la production de la GnRH, et celle de la FSH et de la LH, (et, par conséquent, la production de testostérone). Au contraire, une faible concentration en testostérone stimule la production de la LH. De surcroît, une deuxième connexion relie la LH et la FSH : la testostérone libérée dans les testicules sous l'influence de la LH agit aussi sur les cellules de Sertoli, donc sur la production des spermatozoïdes.

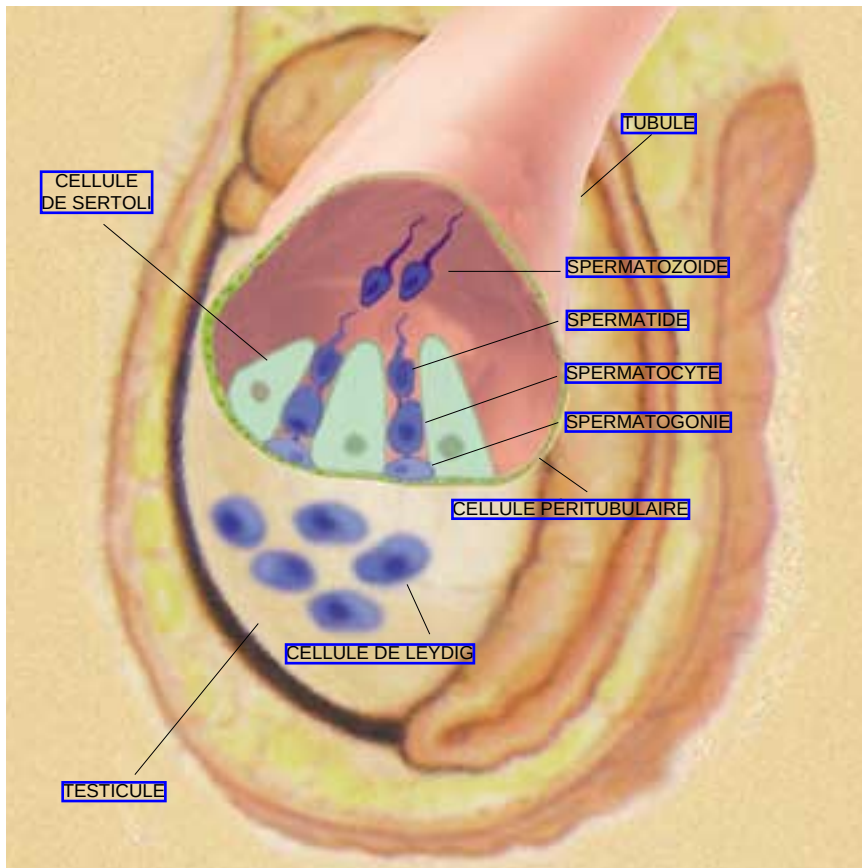
Chez le rat, la spermatogenèse nécessite une synergie entre la testostérone intratesticulaire et la FSH. On a montré qu'après ablation chirurgicale de l'hypophyse (les hormones FSH et LH ne sont plus produites), l'administration de testostérone seule déclenche la spermatogenèse. Toutefois, pour que la spermatogenèse soit normale, on doit aussi administrer de la FSH. Chez le singe, la testostérone peut, à elle seule, déclencher la spermatogenèse, la maintenir et, même, la restaurer après ablation de l'hypophyse.

La régulation de la spermatogenèse

Quels sont les rôles respectifs de la testostérone, de la LH et de la FSH chez l'homme ? L'étude de divers cas pathologiques permet de les préciser. Certains enfants ont, avant la puberté, une maladie rare, la testotoxicose. Les récepteurs de la LH ne fonctionnent pas correctement, car ils sont codés par un gène muté : ils fonctionnent en permanence, même quand ils ne sont pas activés par la LH. L'activation permanente des récepteurs de la LH déclenche une production anormalement élevée de testostérone : ces garçons produisent des spermatozoïdes. Cet exemple montre que la testostérone intratesticulaire, seule, peut déclencher la production de spermatozoïdes.

Au contraire, chez les garçons ayant une anomalie de la production de GnRH,





2. LES TUBULES SÉMINIFÈRES (au premier plan), sont localisés dans les testicules (à l'arrière plan). Ils contiennent les cellules de Sertoli, et sont le siège de la formation des spermatozoïdes. Avant de devenir des cellules fécondantes matures, les spermatozoïdes sont passés par plusieurs stades de développement : les spermatogonies, les spermatocytes et les spermatides.

L'administration de testostérone ne déclenche pas de spermatogenèse. Dans une de ces affections, le syndrome de Kallmann-De Morsier, les malades n'ont pas d'odorat et leurs testicules sont atrophiés ; les hormones FSH et LH ne sont pas (ou peu, dans les formes partielles de la maladie) synthétisées dans l'hypophyse. Les neurones qui produisent la GnRH et les neurones du système olfactif sont issus de la même zone embryonnaire (la placode olfactive) : quand leur migration est anormale, au cours de l'embryogenèse, ni le bulbe olfactif, ni les neurones à GnRH ne se développent correctement. Les patients sont dépourvus d'odorat, et comme leur hypothalamus ne contient pas les neurones qui auraient dû libérer l'hormone GnRH, la production de FSH et de LH par l'hypophyse n'est pas déclenchée : les gonades ne se développent pas. Dans ces conditions, l'administration de testostérone ne suffit pas à déclencher la spermatogenèse. Ainsi, la testostérone exogène ne joue pas le même rôle que la testostérone intratesticulaire.

Diverses études ont montré que la LH et la testostérone intratesticulaire sont nécessaires à la spermatogenèse dans l'espèce humaine. En résumé, l'association des deux gonadotrophines, LH et FSH, est nécessaire pour que la spermatogenèse soit qualitativement et quantitativement normale.

L'inhibine est une autre hormone qui régule la spermatogenèse. Elle est produite dans les testicules par les cellules de Sertoli sous l'action de la FSH, et inhibe la synthèse de la FSH par l'hypophyse. D'autres molécules participent à la régulation de la spermatogenèse, notamment la vitamine A, le facteur de croissance des cellules-souches SCF, et la protéine codée par le proto-oncogène *c-Fos*. Les spermatozoïdes des rats dépourvus de vitamine A se développent anormalement. Les souris transgéniques dont l'expression du gène codant le facteur de croissance des cellules-souches a été inhibée ont de petits testicules et une spermatogenèse anormale. Plusieurs systèmes hormonaux combinent leur action pour assurer une spermatogenèse correcte.

On déduit des mécanismes de la spermatogenèse les différents sites d'action possibles d'un contraceptif : on pourrait agir sur la GnRH hypothalamique, sur la sécrétion de la LH et de la FSH hypophysaires, sur le développement des spermatozoïdes dans les testicules, sur la maturation des spermatozoïdes dans l'épididyme, ou encore interrompre le trajet des spermatozoïdes.

La contraception hormonale

Les hommes qui ne sécrètent pas de LH ou de FSH, en particulier ceux qui ont un syndrome de Kallmann sont azoospermes, c'est-à-dire que l'éjaculat ne contient pas de spermatozoïdes. Quel est le seuil, exprimé en nombre de spermatozoïdes dans l'éjaculat, pour qu'un contraceptif soit efficace ? Selon une étude de l'Organisation mondiale de la santé en 1990, ce nombre doit être inférieur à trois millions par millilitre. Ainsi, un contraceptif efficace n'aurait pas à garantir une azoospermie totale, mais le nombre de spermatozoïdes en deçà duquel il n'y a pas fécondation reste controversé.

Une contraception hormonale doit concilier deux exigences difficilement conciliables : inhiber la sécrétion de LH et de FSH pour bloquer la production des spermatozoïdes, tout en conservant une concentration (ou imprégnation) en hormones masculines normale pour que les caractères sexuels secondaires, la libido et la vie sexuelle ne soient pas perturbés.

D'après la description précédente de la régulation hormonale de la spermatogenèse, on peut soit augmenter la concentration en testostérone en administrant l'hormone masculine (ce qui bloque la production de LH et de FSH), soit inhiber la production de la FSH et de la LH (en freinant directement l'activité de la GnRH). Ainsi, les traitements possibles sont la testostérone et ses dérivés, les progestatifs (un type d'hormones féminines), seuls ou en association avec des hormones masculines, les agonistes de la GnRH (qui par un effet de saturation, expliqué plus loin, inhibent l'action de cette hormone) et, enfin, les antagonistes de la GnRH (qui en bloquent l'action) combinés aux hormones masculines.

En théorie, un traitement par la testostérone et ses dérivés supprime la production des hormones LH et FSH et,

par conséquent, la spermatogenèse, tout en maintenant une concentration en hormones masculines normale. Dans une étude réalisée par l'Organisation mondiale de la santé, sur 271 hommes suivis dans sept pays différents, le traitement consistait en une injection intramusculaire de testostérone par semaine, pendant six mois.

Injection de testostérone

Ce traitement a entraîné une absence de production de spermatozoïdes chez 35 à 60 pour cent des hommes après 7 à 20 semaines de traitement. Une seule grossesse est survenue pendant la durée de l'essai. Cette étude a révélé une grande disparité ethnique : on a constaté que la proportion d'absence totale de production de spermatozoïdes (ou azoospermie) chez les Asiatiques (90 pour cent) était notablement plus élevée que chez les Européens et les Nord-Américains (25 à 75 pour cent).

Ce traitement par les hormones masculines a des effets secondaires qui résultent d'une concentration trop élevée en testostérone dans le plasma. L'excès de testostérone engendre une prise temporaire de poids, l'apparition d'acné, l'augmentation de taille de la

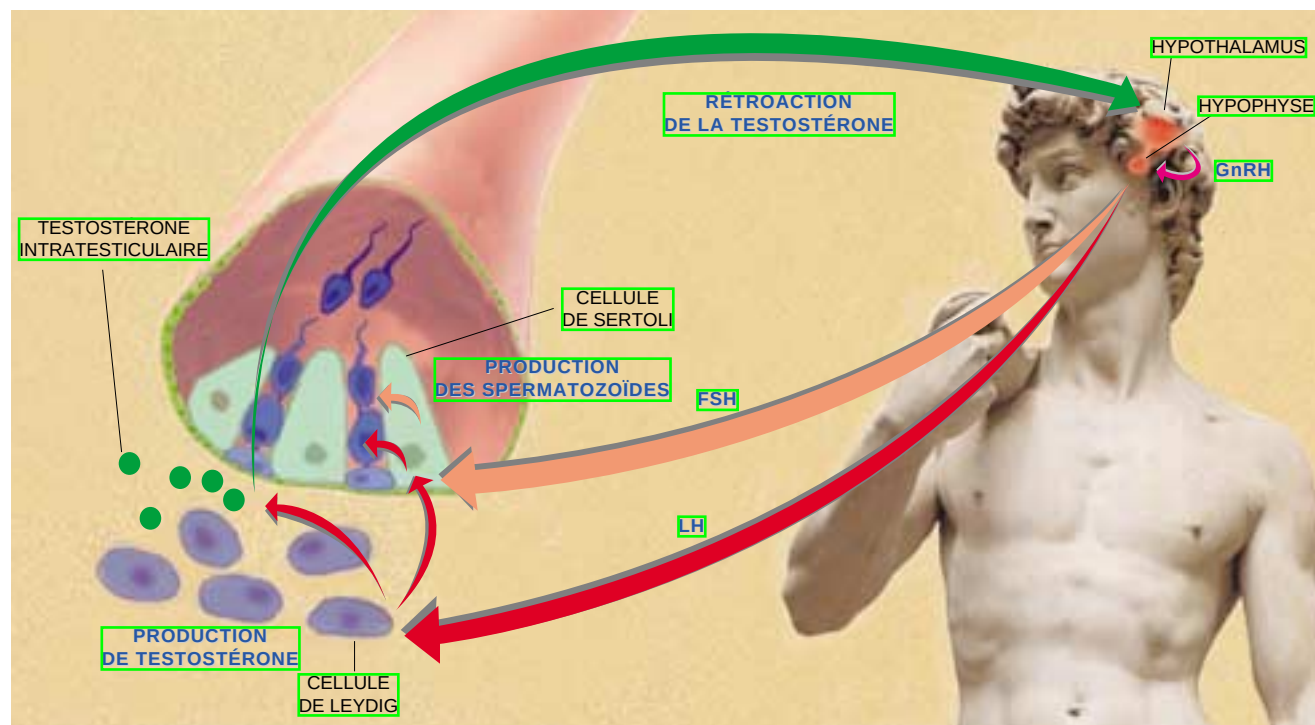
glande mammaire et une diminution de la concentration en cholestérol HDL, le «bon» cholestérol. Ce traitement diminue le volume des testicules. De surcroît, les androgènes risquent, en théorie, d'entraîner une augmentation de la taille de la prostate et de favoriser les carcinomes prostatiques.

Les hommes traités par des injections hebdomadaires de testostérone ne présentent pas de modification de l'humeur, ni d'agressivité. Toutefois, ce traitement pose la difficulté intrinsèque de tout contraceptif masculin hormonal : il lui faut trois mois pour agir, puisque la spermatogenèse dure plus de 70 jours. Le traitement est réversible, mais trois ou quatre mois seulement après l'arrêt du traitement. Un autre inconvénient du traitement par les androgènes reste le mode d'administration : aujourd'hui, la seule méthode efficace nécessite une injection hebdomadaire. La concentration en testostérone atteint des concentrations efficaces au cours des jours qui suivent l'injection.

En revanche, les androgènes, administrés par voie orale, trois fois par jour n'ont pas d'action contraceptive efficace, vraisemblablement parce que la concentration sanguine en testostérone reste insuffisante. Des recherches sont

en cours pour tourner cette difficulté : des équipes tentent de mettre au point des dispositifs à diffusion transcutanée, c'est-à-dire des timbres qui délivreraient la testostérone à travers la peau. D'autres étudient la possibilité de poser des implants sous la peau ou encore des dispositifs qui libéreraient les androgènes de façon prolongée et régulière, respectant la sécrétion physiologique.

L'équipe allemande d'Eberhardt Nieschlag a récemment administré du buciclate de testostérone (un ester de la testostérone) à 12 hommes. Une seule injection de 1,2 gramme a notablement abaissé le nombre des spermatozoïdes, pendant plusieurs semaines, chez la moitié des hommes traités. La proportion reste insuffisante, mais les médecins ont constaté que la concentration en testostérone était restée dans les limites de la normale pour tous les hommes traités (l'augmentation brutale de la concentration en testostérone, après l'injection, bloque la GnRH pendant plusieurs jours, mais, après ce pic, la concentration globale se normalise). Cette voie semble prometteuse, mais on devra évaluer quelle est la dose minimale qui assurera une contraception efficace chez tous les hommes traités.



3. L'HORMONE GnRH produite par l'hypothalamus commande la synthèse, par l'hypophyse, des hormones FSH et LH. La FSH agit sur les cellules de Sertoli, dans les testicules, où elle règle la production des spermatozoïdes (flèches orange). La LH a pour cible les cellules de Leydig qui produisent la testostérone (flèches rouges). D'une part, la testostérone qui a des récepteurs dans

tout l'organisme exerce une rétroaction sur la production de GnRH par l'hypothalamus (flèche verte) : quand la concentration en testostérone est trop élevée, la production de GnRH, et par conséquent de testostérone, est inhibée. D'autre part, la testostérone agit sur les cellules de Sertoli et, donc, sur la production des spermatozoïdes.

Les progestatifs

Les progestatifs agissent sur les cellules gonadotropes et inhibent aussi la libération de la LH et de la FSH et, par conséquent, la spermatogenèse. Ils sont utilisés seuls ou en association avec les androgènes. L'association d'un dérivé de la progestérone (l'acétate de médroxyprogestérone) et de la testostérone bloque la production des spermatozoïdes chez près de la moitié des hommes traités. Certains dérivés s'accumulent dans le tissu adipeux, ce qui augmente la durée d'action de ces progestatifs. Dans ce cas, le traitement est réversible, mais un an seulement après l'arrêt du traitement (en raison de l'imprégnation du tissu adipeux).

Certaines équipes ont utilisé la forme orale de l'acétate de médroxyprogestérone ou de l'acétate de noéthistérone avec des androgènes administrés par voie percutanée ou orale. La testostérone renforce l'action inhibitrice du progestatif sur l'hypophyse. Le groupe de Carlo Flamigni, à Bologne, a récemment confirmé ces résultats en associant de l'acétate de cyprotérone, ou Androcur (un progestatif), et de la testosté-

rone chez 15 hommes. Les endocrinologues utilisent généralement l'acétate de cyprotérone pour traiter certains effets néfastes de la testostérone : chez les hommes ayant un cancer de la prostate, chez les femmes dont la pilosité est trop importante ou chez les personnes ayant de l'acné.

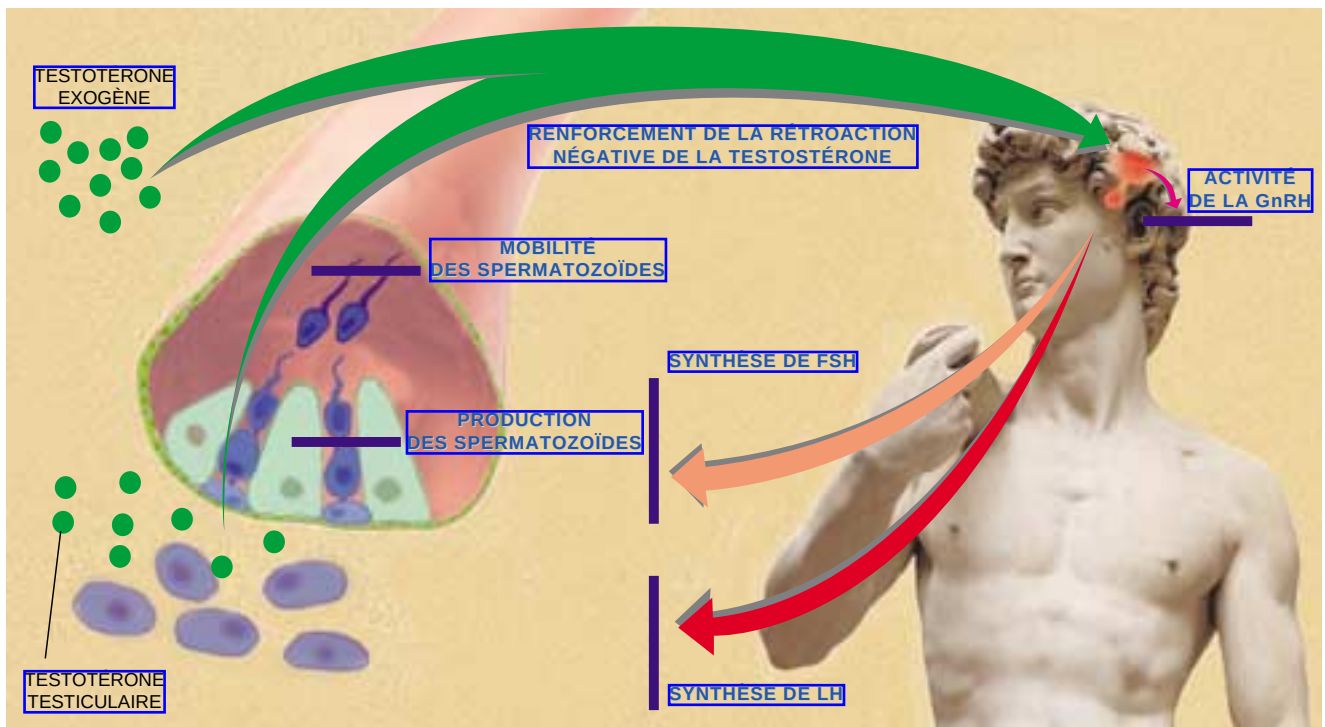
L'acétate de cyprotérone entre en compétition avec la testostérone, car elle se fixe sur les récepteurs des hormones masculines et les leurre. Elle inhibe la production de FSH et de LH et, ainsi, bloque la spermatogenèse, mais seulement chez une faible proportion des hommes traités. Au contraire, l'administration combinée de cet acétate et de l'énanthate de testostérone (une forme de testostérone à libération prolongée) est suivie d'une absence de production de spermatozoïdes, chez la totalité des hommes traités. La production des spermatozoïdes est bloquée au bout de sept semaines, environ deux fois plus rapidement qu'avec la testostérone seule.

D'après une étude réalisée sur huit volontaires, ce type de traitement, associant l'acétate de cyprotérone et la testostérone par voie orale, est bien toléré. Le nombre de spermatozoïdes a nota-

blement diminué chez les huit hommes traités pendant 16 semaines, mais l'absence totale de production des spermatozoïdes ou une production minimale n'ont été constatées que chez six des huit participants. L'efficacité d'un traitement combinant l'acétate de cyprotérone et la testostérone n'est pas suffisante. Plusieurs équipes recherchent de nouveaux androgènes qui pourraient être administrés par voie orale, en association avec l'acétate de cyprotérone. Ces traitements devraient diminuer la concentration en LH et en FSH et, par conséquent, inhiber la spermatogenèse tout en maintenant une imprégnation suffisante en testostérone.

Agonistes et antagonistes de la GnRH

La GnRH est une hormone responsable de la sécrétion de la LH et de la FSH. En modifiant certains acides aminés de cette molécule, les biologistes ont mis au point des substances qui se fixent sur les récepteurs hypophysaires de la GnRH. Les premières molécules étaient des agonistes, ayant une affinité plus élevée pour les récepteurs hypophysaires de la GnRH que la GnRH elle-même ; en



4. LES CONTRACEPTIFS MASCULINS agissent en bloquant l'une des voies de production des spermatozoïdes ou leur trajet après qu'ils ont été produits. L'injection de testostérone (testostérone exogène) renforce la rétroaction négative de l'hormone sur l'hypothalamus, ce qui inhibe la production de l'hormone GnRH. D'autres substances bloquent l'activité biologique de la GnRH : ce sont des agonistes de cette hormone qui finissent par «insensibiliser» (et donc par inactiver) les

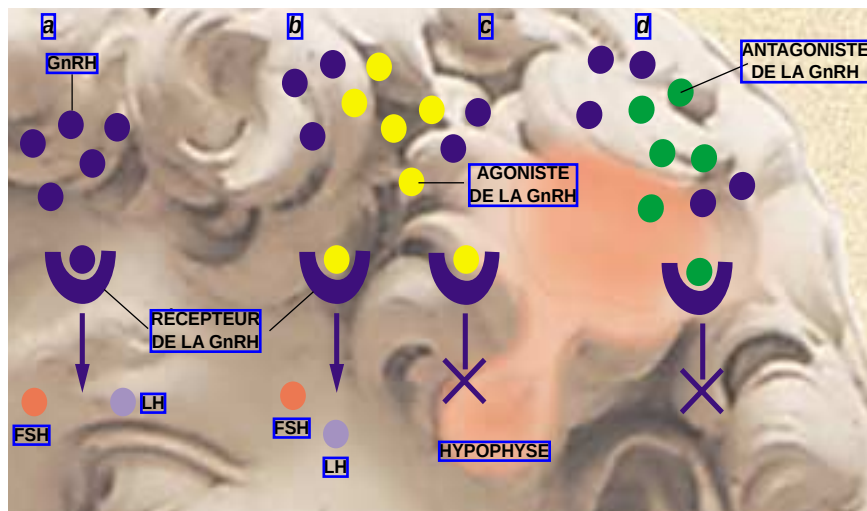
récepteurs de la GnRH, soit des antagonistes qui bloquent ces récepteurs. De même, d'autres substances, tels les progestatifs, agissent aussi dans la région de l'hypothalamus et de l'hypophyse, bloquant la synthèse de la FSH et de la LH. Ces quatre types de substances inhibent la production des spermatozoïdes. D'autres contraceptifs sont également envisageables : ils réduiraient la mobilité des spermatozoïdes déjà produits et, par conséquent, leur pouvoir fécondant.

outre, ces agonistes sont éliminés plus lentement de l'organisme que la GnRH.

Comment agissent les agonistes de la GnRH? Pendant environ deux semaines, ils commencent par stimuler la sécrétion de LH et de FSH, car ils imitent la GnRH et leurrent les récepteurs. Puis, dans une deuxième phase, survient un phénomène de désensibilisation : le seuil d'activation des récepteurs de la GnRH, stimulés en permanence, augmente, de sorte que les messages qui aboutissent à la production de la FSH et de la LH ne sont plus transmis. Douze études utilisant des agonistes de la GnRH ont été menées. Elles ont toutes montré que ce type de traitement ne diminuait pas suffisamment le nombre de spermatozoïdes pour réaliser une contraception efficace.

Au contraire, les antagonistes de la GnRH semblent plus prometteurs. Ils se fixent aussi sur les récepteurs de la GnRH endogène, mais ne transmettent aucun message. Ils agissent plus rapidement que les agonistes de la GnRH, car la phase initiale de stimulation, que l'on constate avec les agonistes et qui dure deux semaines, n'existe pas. Les études menées sur les rongeurs et les primates non humains ont montré que l'association d'un antagoniste de la GnRH et d'un androgène entraîne une azoospermie réversible, avec un comportement sexuel normal. Chez l'homme, quelques études ont été réalisées avec le Nal-Glu et des androgènes en injection une fois toutes les deux semaines : la spermatogenèse a été bloquée chez sept des huit hommes traités. Dans une autre étude, un traitement hebdomadaire par des concentrations assez élevées d'hormones masculines a entraîné une absence totale de production de spermatozoïdes chez tous les hommes traités. D'autres types d'antagonistes de la GnRH sont en cours de développement, notamment le Cétrorelix, des Laboratoires *Asta Medica*, à Bordeaux.

Le mode d'administration reste la principale contrainte de ce type de traitement. Les pharmacologues recherchent des formes galéniques qui autoriseraient une administration mensuelle, voire trimestrielle. De surcroît, les nouveaux antagonistes ne déclenchent plus les réactions allergiques que provoquaient les premiers antagonistes de la GnRH ; ils s'accompagnent seulement, dans quelques cas, de rougeurs cutanées, au point d'injection. Comme avec les traitements par la testostérone et ses dérivés, les antagonistes de la



5. LA GnRH EST INHIBÉE dans l'hypophyse de diverses façons. Normalement, quand elle se fixe sur ses récepteurs, elle déclenche la production de la LH et de la FSH (a). Quand on administre des agonistes, ils se fixent sur les récepteurs de la GnRH avec une affinité supérieure à celle de l'hormone elle-même. Dans un premier temps, ils agissent comme la GnRH, libérant de la LH et de la FSH (b), mais le seuil d'excitation des récepteurs augmente au bout de quelques jours, de sorte qu'ils ne sont plus activés et ne transmettent plus de messages : la production de LH et de FSH cesse (c). Quand on injecte des antagonistes, ils prennent la place de la GnRH sur ses récepteurs et bloquent leur action : à nouveau, la production de LH et de FSH est inhibée (d).

GnRH ne sont efficaces qu'au bout de trois mois et ils sont réversibles en trois à quatre mois.

D'autres méthodes contraceptives

Jusqu'à présent, nous avons énuméré des méthodes qui visent à bloquer la production des spermatozoïdes. D'autres traitements ont été envisagés, mais ils ont tous un inconvénient rédhibitoire : ils ne sont pas réversibles. Ainsi, dans les années 1950, les médecins chinois avaient constaté de nombreux cas de stérilité masculine chez les hommes qui cultivaient le coton. Cette stérilité résultait d'un contact avec le gossypol, une substance contenue dans le coton. Le gossypol inhibe la mobilité des spermatozoïdes et leur maturation. Malheureusement, cette substance est toxique : 30 pour cent des hommes qui ont été en contact avec elle n'ont jamais retrouvé une spermatogenèse normale. En Inde, quelques équipes étudient la toxicité du gossypol et la cause de son effet irréversible.

D'autres moyens sont envisageables. Existerait-il un vaccin efficace? On sait que des anticorps dirigés contre les spermatozoïdes sont responsables de stérilité chez cinq pour cent des couples infertiles. On a donc imaginé de vacciner les hommes avec certains antigènes de surface portés par les spermatozoïdes. Les anticorps produits

après la vaccination se lient à leurs antigènes spécifiques, inhibant l'activité des spermatozoïdes, notamment la fusion du spermatozoïde potentiellement fécondant avec la zone pellucide de l'ovocyte. Toutefois, les antigènes des spermatozoïdes sont encore mal identifiés et, même si certaines protéines de surface sont aujourd'hui à l'étude, aucune n'est encore un vaccin potentiel.

Des vaccins contenant des anticorps dirigés contre la zone pellucide sont à l'étude. Toutefois, la contraception immunologique pose une difficulté de principe : les réactions du système immunitaire, qui conditionnent l'efficacité du vaccin, fluctuent d'une personne à l'autre. Le délai d'action et la réversibilité de ces vaccins sont en cours d'étude.

Une autre voie de recherche repose sur l'inhibition de la spermatogenèse par des moyens physiques : la spermatogenèse est normale à condition que la température des testicules soit inférieure de deux degrés à celle du reste de l'organisme. Or, les testicules, en position anormale, c'est-à-dire inguinale ou abdominale, présentent des anomalies de la spermatogenèse. L'équipe de R. Mieusset et L. Bujan, à Toulouse, a mis au point une méthode qui déclenche une légère hyperthermie par compression des testicules à l'entrée des canaux inguinaux, durant les heures d'éveil, 15 heures par jour. Cette méthode réduit le nombre des spermatozoïdes et leur

Les contraceptifs masculins de demain

La testostérone et les hormones masculines :

- Elles agissent sur l'hypothalamus en inhibant la production de la GnRH ; elles ont un délai d'action supérieur à deux mois ; elles nécessitent des injections hebdomadaires ;
- Une des difficultés sera la détermination de la dose minimale efficace chez tous les hommes traités. Les effets secondaires sont à évaluer.

Les progestatifs :

- Ils bloquent la synthèse de la LH et de la FSH ;
- Un progestatif, l'acétate de cyprotérone, combiné à un dérivé de la testostérone, bloque la spermatogenèse chez la totalité des hommes traités par cette association.

Les antagonistes et les agonistes de la GnRH :

- Les agonistes se lient aux récepteurs de la GnRH qui, stimulés en permanence, finissent par ne plus réagir. Ils ne sont pas des contraceptifs efficaces.
- Les antagonistes bloquent directement les récepteurs de la GnRH. Combinés à des hormones masculines, ils représentent des contraceptifs potentiellement efficaces.

Les méthodes non-hormonales :

- Les vaccins posent deux difficultés : les antigènes portés par les spermatozoïdes, qui seraient des agents vaccinaux, sont mal connus ; les réactions du système immunitaire, lesquelles conditionnent l'efficacité d'un vaccin, fluctuent d'une personne à l'autre.
- Inhibition de la spermatogenèse par augmentation de la température des testicules.

mobilité. Plusieurs études cliniques ont montré que l'hyperthermie testiculaire est un contraceptif masculin efficace. La réversibilité est obtenue en six à douze mois. Toutefois, chez la souris, la proportion de malformations embryonnaires après traitement hyperthermique du mâle est notablement plus élevée que chez les témoins. Enfin, on ignore encore les effets à long terme de l'hyperthermie sur la spermatogenèse.

La vasectomie

La vasectomie est un moyen contraceptif utilisé en Chine, aux États-Unis et en Grande-Bretagne. Chaque année, plus de 500 000 hommes choisissent cette méthode de contraception aux États-Unis. Elle est illégale en France. Elle consiste à lier, au cours d'une intervention chirurgicale, les canaux déférents ou à les couper de quelques centimètres. Le canal déférent est interrompu par la ligature ou par la résection. La stérilité est obtenue deux à trois semaines après l'opération. *A priori*, elle est peu réversible. Le rétablissement de continuité des canaux déférents restaure un spermogramme normal chez 60 pour cent des hommes opérés, mais seulement 40 pour cent d'entre eux sont féconds. On ignore pourquoi, mais après une vasectomie, des anticorps anti-spermatozoïdes apparaissent systématiquement. Ils seraient responsables de la faible fertilité des hommes chez qui l'on rétablit la continuité des canaux déférents. Le suivi de 10 500 hommes ayant subi une vasectomie a confirmé

l'innocuité de la méthode, même si quelques effets secondaires mineurs sont signalés : saignement local, inflammation ou infection au voisinage des ligatures, ou encore quelques douleurs.

On peut réaliser une vasectomie non chirurgicale en injectant sous la peau, dans les canaux déférents, une substance qui polymérise dans les canaux. Toutefois cette méthode est irréversible, et le polyuréthane injecté est potentiellement cancérogène. Avant toute vasectomie, on pourrait recueillir du sperme et le conserver à -196°C , dans l'azote liquide ; il retrouverait son pouvoir fécondant après décongélation. Une insémination artificielle, au moment de l'ovulation de la partenaire, aboutirait à un taux de fécondation élevé.

Enfin, n'oublions pas le tout premier contraceptif masculin (du moins historiquement) : le préservatif. Ses premières utilisations datent du IV^e siècle, et il semble être apparu en Inde. Les premiers préservatifs étaient constitués de vessie ou d'intestins de porc. Depuis quelques années, le préservatif est très utilisé, car il protège contre les maladies sexuellement transmissibles. Toutefois, son efficacité est inférieure à celle des contraceptifs féminins. Le contraceptif masculin idéal devrait être aussi efficace que les contraceptifs féminins et protéger contre les maladies sexuellement transmissibles.

D'autres techniques seront sans doute développées, telles que la modification de la maturation des spermatozoïdes et la réduction de leur mobilité dans les voies excrétrices des testicules,

mais aussi dans les voies génitales féminines. Pourtant, ces techniques ne verront pas le jour avant que la physiologie testiculaire et les interactions des testicules et des autres acteurs du système de la reproduction ne soient totalement élucidées.

La recherche en contraception masculine progresse plus lentement que la recherche en contraception féminine, car on sépare difficilement, dans les testicules, la production de spermatozoïdes de celle des hormones masculines. Chez la femme, la contraception fonctionne comme un mécanisme de tout ou rien : l'émission mensuelle de l'ovule est inhibée chez toutes les femmes traitées. En revanche, une même concentration d'hormones ne bloque que partiellement la production des millions de spermatozoïdes : une des difficultés est d'inhiber efficacement la spermatogenèse chez tous les hommes traités. Outre les difficultés biologiques, deux conditions seront nécessaires au développement de la contraception masculine : la femme devra faire confiance à son partenaire, lequel devra se sentir concerné par la planification du nombre de ses enfants. Quoi qu'il en soit, sans moyen contraceptif largement diffusé, la population mondiale dépassera six milliards de personnes en l'an 2 000.

Sophie CHRISTIN-MAITRE, praticien hospitalier-universitaire est endocrinologue à l'Hôpital Saint-Antoine, dans le service que dirige Philippe BOUCHARD.

B. JEGOU, C. PINEAU et A. DUPAIX, *La cellule de Sertoli : actualisation du concept de cellule nourricière*, in *Médecine Sciences*, n° 4, vol. 11, pp. 519-527, 1995.

S. CHRISTIN-MAITRE, J.-F. GUÉRIN et P. BOUCHARD, *La contraception masculine*, in *Médecine Sciences*, n° 4, vol. 11, pp. 579-590, 1995.

R.A. ANDERSON et al., *Comparison between Testosterone Oenanthate-induced Azoospermia and Oligozoospermia in Male Contraceptive Study*, in *Human Reproduction*, vol. 12, n° 8, pp. 1657-1662, 1997.

M.C. MERIGGIOLA et W.J. BREMNER, *Progestin-androgen Combination Regimens for Male Contraception*, in *J. Androl.*, vol. 18, n° 3, pp. 240-244, mai-juin 1997.

S. CHRISTIN-MAITRE, *Gonadotrophines et physiologie ovarienne*, in *Médecine thérapeutique*, vol. 4, hors série, pp. 29-33, 1998.

L'évolution des tortues

GUILLERMO ROUGIER • MARCELO DE LA FUENTE • ANDREA ARCUCCI

La récente découverte, en Argentine, d'étonnants spécimens fossiles précise l'origine des tortues et révèle une diversité écologique et systématique inattendue chez les Chéloniens.

Si la carapace des tortues rend ces animaux facilement identifiables, son origine reste l'un des mystères de la paléontologie des vertébrés. La tortue peut rétracter son cou et enfouir sa tête sous la carapace, et cela de deux façons possibles qui correspondent aux deux types de tortues actuelles. La découverte de tortues fossiles en Argentine a précisé la date de l'apparition de cette bifurcation.

Les tortues sont amniotes, comme les lézards, les crocodiles, les oiseaux les mammifères actuels et un grand nombre de formes fossiles. Chez les amniotes, l'embryon est recouvert d'une membrane, l'amnios, qui diminue fortement les pertes d'eau. Cette structure a permis aux amniotes de conquérir tous les environnements terrestres, puisque, à la différence des amphibiens, ils n'avaient pas l'obligation de retourner dans l'eau pour pondre.

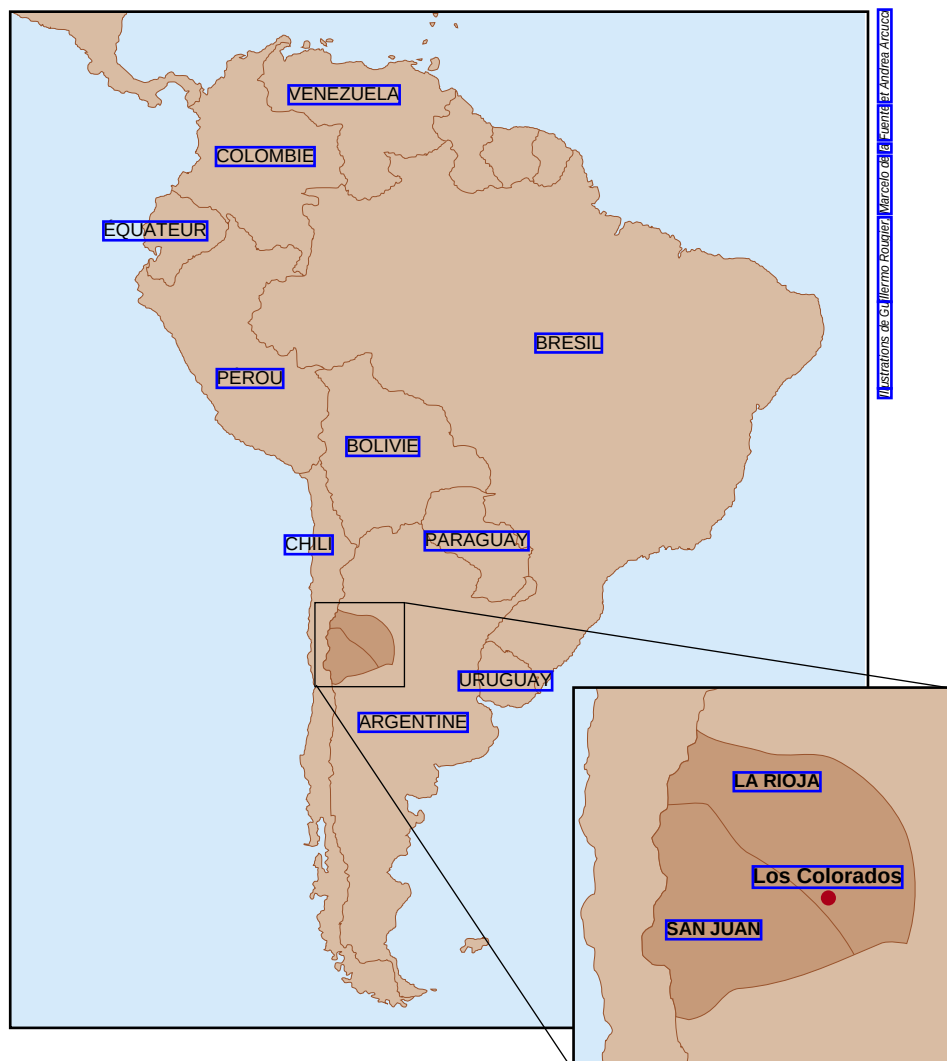
Cette évolution a débuté au Carbonifère, qui commença il y a 322 millions d'années et dura 30 millions d'années. L'acquisition de l'amnios entraîna, durant la période suivante, le Permien (entre 290 et 245 millions d'années), une explosion de nouveaux groupes. Bien que l'on n'ait pas retrouvé de fossiles de tortues datant du Permien, toutes les lignées que l'on pense liées aux tortues sont issues de cette période.

Les déductions de Watson

Après la découverte récente de nouveaux fossiles en Argentine, les paléontologues ont réexaminé et précisé les hypothèses sur l'origine des tortues et leurs relations avec les principaux groupes d'amniotes.

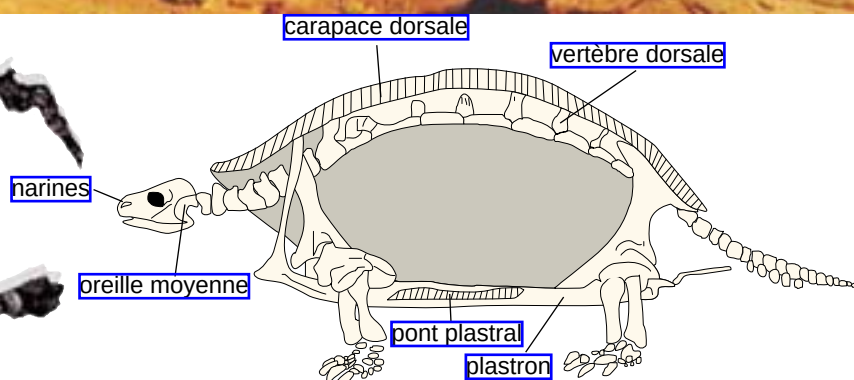
L'identification des ancêtres des tortues est une longue histoire. En 1914, D. Watson, paléontologue anglais, postula que *Eunotosaurus africanus*, un petit amniote d'Afrique du Sud qui vivait au Permien moyen, il y a environ 260 millions d'années, était un ancêtre des tortues. Selon Watson, la carapace des tortues résultait de l'élargissement et de l'épaississement des côtes

d'*Eunotosaurus*. Hélas pour la phylogénie, cette ébauche de carapace n'a pas les mêmes caractéristiques chez les tortues primitives du Trias, aussi l'hypothèse n'a pas semblé convaincante. De plus, ajoutaient les paléontologues, contrairement aux tortues actuelles et fossiles qui n'ont pas de dents mais une crête osseuse recouverte d'une couche cornée, *Eunotosaurus* a de véritables



Illustrations de et par Guillermo Rougier, Marcelo De La Fuente et Andrea Arcucci

1. LA FORMATION LOS COLORADOS présente de grands affleurements, de 100 kilomètres de long, 30 kilomètres de large et plusieurs centaines de mètres d'épaisseur (a). C'est dans le secteur supérieur (b) de cette formation que l'on a découvert des vertébrés fossiles du Trias (205-220 millions d'années), dont *Palaeochersis talampayensis*, la plus ancienne tortue fossile connue d'Amérique du Sud (c). Sur le squelette de la tortue, on distingue les narines externes allongées et le début d'une cavité de l'oreille moyenne. Dans le squelette, on observe les processus épiploastraux (clavicules), la ceinture scapulaire primitive, un court pont plastral et l'ischion soudé au plastron.



dents implantées dans les prémaxillaires et les maxillaires ; contrairement aux tortues, il ne possède pas une cuirasse de plaques osseuses ; enfin, les ceintures scapulaires (os de l'épaule) sont disposées au-dessus des côtes, comme chez la majorité des amniotes et non pas dessous, comme chez les tortues.

L'étude de spécimens anciens et nouveaux, dont un crâne complet d'*Eunotosaurus africanus*, poussèrent B. Cox en 1969, et A. Keyser et C. Gow en 1981, à considérer qu'*Eunotosaurus* était un élément d'un groupe d'amniotes primitifs, les captorhinidés, qui ressemblaient à des lézards, mais n'avaient pas de relation claire avec les tortues.

Comment était apparue la carapace ? Pas par épaissement des côtes et des épines neurales, comme le pensait Watson, mais par extension et fusion d'ossifications sur les vertèbres dorsales et les côtes, commencèrent à supputer les paléontologues.

En suivant cette hypothèse, ils ont apparenté divers groupes d'amniotes aux tortues. Après les avoir liés aux captorhinidés, Robert Reisz et Michel Laurin ont précisé un lien de parenté avec les procolophonidés, un sous-groupe des captorhinidés. Enfin, Michael Lee, en 1993, étudia à fond

les parentés entre les amniotes primitifs et identifia quatre groupes Pareiasaurida, *Sclerosaurus*, Procolophonidae et Nyctiphruretida qui, quand on remonte le temps, se distinguent de plus en plus du groupe des tortues. Ces quatre groupes ont principalement vécu au Permien : à l'exception du premier, nettement plus gros, ils étaient tous de taille petite ou moyenne, semblables aux lézards actuels.

Les paréiasauridés, ancêtres des tortues ?

Michael Lee prôna le regroupement des paréiasauridés et des Chéloniens sur la base de 16 caractéristiques communes. Quant à la carapace, elle résultait d'une fusion des plaques osseuses du dos et des flancs des paréiasauridés.

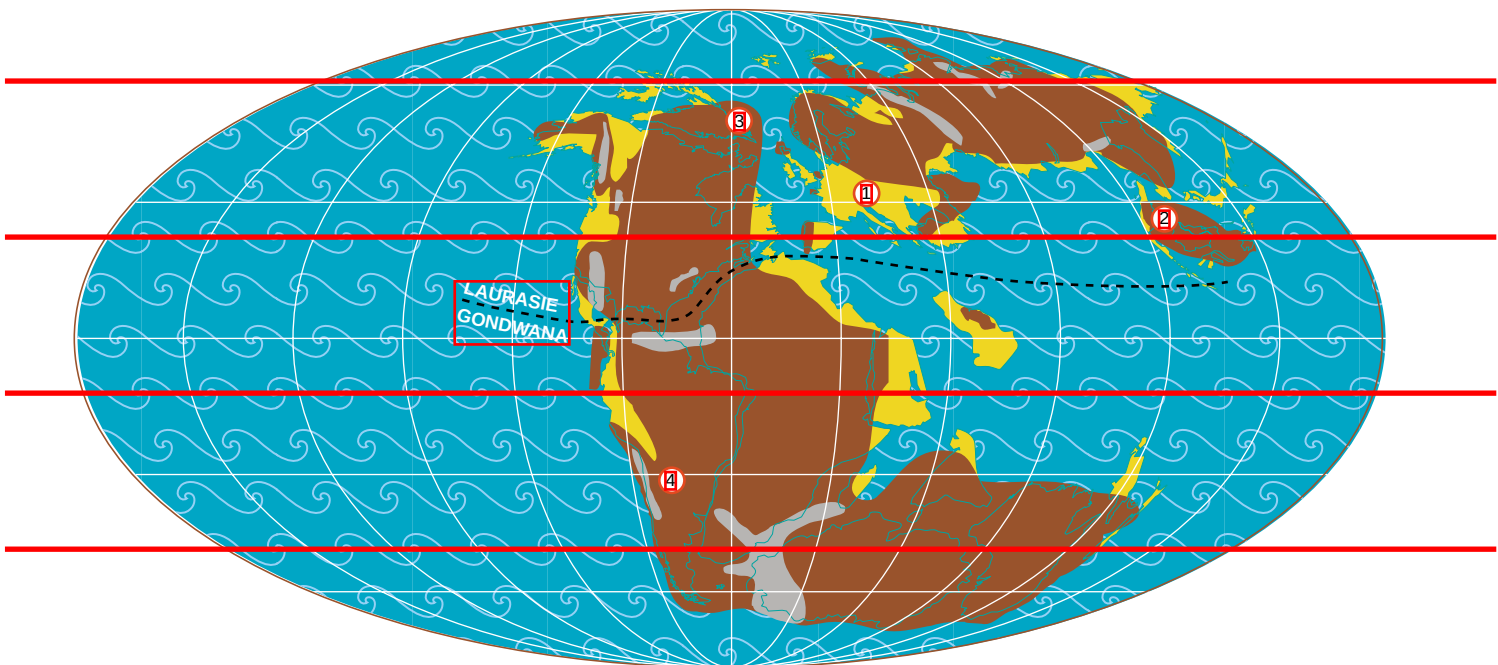
Le fait que les tortues du Trias possèdent un nombre d'éléments de carapace supérieur à celui des tortues modernes semble étayer l'hypothèse qu'au cours des âges, la carapace s'est formée par fusion de multiples ossifications indépendantes. La tendance au raccourcissement du squelette et la présence de seulement cinq vertèbres cervicales dans le cou des paréiasauridés confirment leur parenté étroite avec les tortues. Dans ce contexte, il semble que

les huit vertèbres cervicales et la disposition interne des ceintures des tortues actuelles résultent, chez les paréiasauridés, de l'incorporation de la première, deuxième et troisième vertèbres dorsales à leur cou par un déplacement vers l'arrière des os de l'épaule.

Dramatique absence : pas de tortues au Permien !

Le Norien est la période du Trias supérieur qui s'étend de 220 à 205 millions d'années. Comme les premières tortues fossiles connues datent de cette époque, 30 millions d'années au moins séparent les Chéloniens les plus anciens et les représentants des groupes qui leur sont apparentés. D'où l'incertitude : on n'a pas découvert de fossile montrant la transformation la plus draconienne de l'histoire des Chéloniens, l'apparition de la carapace.

On n'avait, avant la découverte argentine, retrouvé des restes de Chéloniens du Trias que dans trois régions, l'Allemagne, l'Asie du Sud-Est et le Groenland oriental. La dérive des continents explique cette disposition : au Norien, les emplacements des continents étaient radicalement différents de leurs emplacements actuels. Au Trias, un supercontinent, la Pangée, résultait de



2. CARTE PALÉOGÉOGRAPHIQUE où sont indiquées les lignes côtières de la fin du Trias, il y a 220 millions d'années. À l'époque, le supercontinent Pangée commençait à se séparer en Laurasia (Nord) et Gondwana (Sud). L'Amérique du Nord est encore liée au Gondwana. Les quatre régions où l'on a localisé des gisements fossilifères avec des tortues datant du Norien sont indiquées. Ainsi, on a retrouvé en

Allemagne six squelettes de *Proganochelys quenstedti* et une douzaine de carapaces de *Proterochersis robusta* (1). Des spécimens fort incomplets de *Proganochelys ruchae* proviennent de la formation de Huai Hin Lat, au Nord-Est de la Thaïlande (2). Les fossiles de *Proganochelys* sp. ont été découverts dans la formation Fleming Fjord, dans le Jameson Land, au Groenland oriental (3). Enfin, une quinzaine de *Palaeochersis talampayensis* ont été mis au jour dans la formation Los Colorados, en Argentine (4).

l'agrégation de tous les grands blocs continentaux ; les trois régions mentionnées se trouvaient dans la partie Nord du supercontinent. À la fin du Norien et au début du Jurassique, la Pangée se scinda, donnant la Laurasia au Nord et le Gondwana au Sud. Après cette scission, des faunes distinctes apparaissent dans chacune des masses continentales.

La première allusion aux tortues triasiques date de 1884. Edward Cope, paléontologue Nord-américain, désigna par ce terme un spécimen découvert à proximité de Tübingen en Allemagne. En 1887, le paléontologue allemand G. Baur baptisa le spécimen de Cope *Proganochelys quenstedti*. Au fil du temps, on découvrit dans les carrières allemandes de Trossingen-Aixheim et Halberstadt six autres spécimens de *Proganochelys quenstedti* datés du Trias supérieur.

À Halberstadt, l'une des tortues était dans des sédiments contenant divers spécimens de *Plateosaurus* (dinosaur prosauropode), de phytosaures (reptiles semblables aux crocodiles), d'amphibiens labyrinthodontes et de poissons. Ces sédiments correspondent à des environnements continentaux, mais proches de la mer. Au Norien, Halberstadt se trouvait effectivement à proximité de la côte.

À Murrhardt et Bentental bei Lorch, à l'Ouest de l'Allemagne, l'on découvrit une dizaine de carapaces d'une autre tortue triasique, légèrement plus ancienne que *Proganochelys*. En 1913, E. Fraas la baptisa *Proterochersis robusta*, et on la considéra comme la plus ancienne tortue pleurodire. Les tortues pleurodires, l'un des deux grands groupes de tortues actuelles, sont caractérisées par un cou qui se replie horizontalement pour faire rentrer la tête à l'intérieur de la carapace et par leur bassin fusionné avec la carapace. Chez l'autre groupe, les cryptodires, le cou se replie verticalement et le bassin n'est pas attaché à la carapace.

Toutefois, les quelques éléments connus des *Proterochersis* se limitent à la carapace et au bassin ; en l'absence des os du cou et du crâne, Fraas ne disposait pas, pas plus que nous ne disposons aujourd'hui, d'éléments suffisants pour relier *Proterochersis* de manière certaine aux pleurodires. Pendant près de 70 ans, les connaissances sur les premiers étapes de l'évolution des tortues se limitaient à ce qui pouvait être déduit des rares restes décou-

verts à la fin du siècle dernier et au début de ce siècle.

En 1982, France de Broin, du Muséum national d'histoire naturelle de Paris, fit part de la découverte de tortues triasiques en Thaïlande. Ces fossiles, fragments de carapaces trouvés dans les dépôts lacustres, servirent à définir une nouvelle espèce de *Proganochelys*, *Proganochelys ruchae*. Ces tortues proviennent d'affleurements situés dans trois localités distinctes du

Nord-Ouest de la Thaïlande, associées à une faune de vertébrés et d'invertébrés. En 1993, Farish Jenkins et ses collègues de l'Université Harvard mentionnèrent des restes fragmentaires de carapaces, attribués à une espèce indéterminée de *Proganochelys*, dans des affleurements continentaux, eux aussi du Trias supérieur du Groenland oriental. Cette faune comprend divers groupes présents dans les localités allemandes, tels que des dinosaures prosauropodes et aétosauriens (reptiles présentant une forte cuirasse de plaques), des ptérosaures, des mammifères et divers poissons, dont des requins.

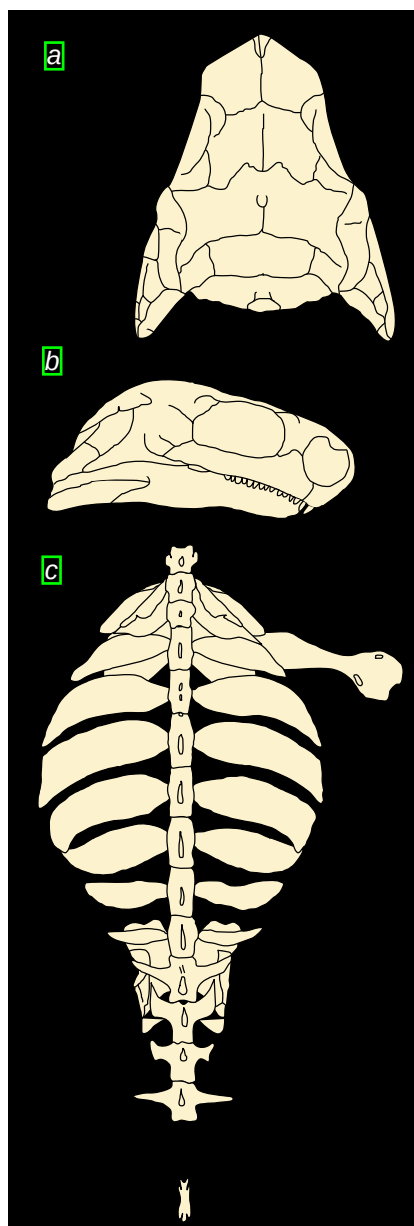
Les tortues argentine

Telle était la situation lorsque, entre 1992 et aujourd'hui, nous avons trouvé des fossiles apportant quelques lumières sur les tortues qui vivaient il y a plus de 200 millions d'années. Nous découvrîmes des spécimens exceptionnels d'une nouvelle espèce de tortue triasique provenant de la province de La Rioja, au Nord-Est de l'Argentine.

Cette découverte a plusieurs facettes. En premier lieu, l'abondance de spécimens et leur excellent état de conservation contrastent avec le maigre échantillonnage de Chéloniens triasiques connus ; en outre, il s'agit du premier gisement de tortues de cet âge situé dans la partie méridionale de la Pangée ; enfin, ces restes sont plus anciens de 60 millions d'années que la date généralement associée à l'apparition de tortues en Amérique du Sud.

Les prospections amenèrent la découverte de deux spécimens de tortues. L'un des squelettes récupérés lors de la première campagne, en 1992, est extrêmement bien conservé : l'on a retrouvé son crâne et ses mandibules, ainsi qu'une grande partie de la carapace et du squelette. Le rapprochement de ce spécimen avec d'autres, découverts par la suite, nous a incité à créer une nouvelles espèce appartenant à un nouveau genre : *Palaeochersis talampayensis*.

Le cadre majestueux des premiers chaînons de la précordillère andine sert de toile de fond au gisement de Chéloniens de La Rioja. Géologiquement, il fait partie du bassin de Ischigualasto-Villa Union, où s'est déposée une séquence sédimentaire qui s'étend de la fin du Trias inférieur à la fin du Trias



3. VUES DU CRÂNE (a-b) ET DU SQUELETTE (c) de *Eunotosaurus africanus*, petit amniote du Permien moyen d'Afrique du Sud (260 millions d'années). On a longtemps cru que c'était l'ancêtre des tortues, car ses côtes élargies couvraient presque complètement le tronc, formant un semblant de carapace. À la différence des tortues actuelles, il a des dents, dont les formes suggèrent un régime insectivore.

supérieur (240-205 millions d'années). Cette séquence triasique culmine avec la formation Los Colorados, dont les roches sont constituées de grès rougeâtres. Ces roches surmontent la formation Ischigualasto où l'on a trouvé les restes de *Eoraptor* et *Herrerasaurus*, parmi les plus anciens dinosaures.

La faune de Los Colorados est extrêmement variée ; on y a découvert des formes proches des premiers mammifères et archosaures (groupe qui comprend les crocodiles, les oiseaux, les dinosaures et les ptérosaures). Parmi les archosaures trouvés, citons des aétosauriens, des crocodiles primitifs, des thécodontes (ancêtres des dinosaures et des crocodiles) et des dinosaures prosauropodes. L'étude de ces vertébrés par José Bonaparte dans les années 1970, a établi qu'il s'agissait d'une association faunistique continentale, la partie haute de la formation datant du Norien. Pour parvenir à cette conclusion, on a comparé les taxons de Los Colorados avec ceux d'autres régions du monde ; la répartition de ces animaux dans les sédiments indiquent que la formation de Los Colorados date du Norien.

En septembre 1994, nous sommes retournés au gisement en compagnie de Gerardo Cladera, sédimentologue de l'Université de Buenos Aires. À notre grande surprise, dans les 140 derniers mètres d'épaisseur du gisement découvert deux ans auparavant, nous avons identifié deux niveaux fossilifères, situés à une centaine de mètres l'un de l'autre. Le niveau inférieur livra des restes d'un dinosaure prosauropode, ainsi que de nombreux restes isolés d'archosaures de moyenne et de grande taille. Dans le niveau fossilifère supérieur situé à 30 mètres du sommet de la formation, nous avons découvert au moins 15 squelettes articulés de *Palaeochersis*, à la base d'une strate argileuse, vestige d'une coulée de boue.

Les tortues étaient enfouies dans une coulée de boue

Les coulées de boues résultent de la mise en mouvement de sédiments par des pluies abondantes ou des tremblements de terre. Les tortues furent ensevelies et emprisonnées par une coulée ; un seul squelette, à la base de la coulée, était retourné. Les spécimens sont magnifiquement conservés.

Les roches des alentours sont en

majorité composées de grès, résultant de la compaction du sable des dunes. Entre ces dunes coulaient parfois des torrents, puis le cours de ces rivières temporaires étaient remplis par des dunes ou, de manière plus sporadique, par des coulées de boue, comme celle qui engloutit les tortues. Ces caractéristiques de la séquence rocheuse suggèrent que les organismes découverts vivaient dans une zone aride, ce qui correspond bien à la nature des reptiles et des mammifères de la formation Los Colorados et aux conditions climatiques arides du Trias supérieur.

Palaeochersis et Proganochelys

Proganochelys, la tortue allemande du Norien, présente un ensemble de caractères primitifs absents chez les casichélidés modernes (le groupe dont font partie les pleurodires et les cryptodires). Ces circonstances ont incité E. Gaffnet et L. Meeker, en 1983, à considérer *Proganochelys* comme un groupe frère des autres tortues. Chez *Proganochelys*, on reconnaît nombre de caractéristiques crâniennes primitives : articulation entre le palais et le basicrâne qui permet le mouvement des deux éléments ; narines externes séparées par une projection du prémaxillaire (chez les tortues actuelles, les fosses nasales sont séparées par des éléments cornés) ; dents sur le palais ; os qui se perdent ou fusionnent chez les groupes les plus avancés, comme les supratemporaux et les vomers. La partie inférieure de la carapace de *Proganochelys* a des caractéristiques primitives, telles que la persistance de projections épipastrales et un processus entoplastral qui s'étend vers l'arrière (vestiges des clavicules et interclavicules).

Bien que chez *Palaeochersis*, la tortue argentine, subsistent bon nombre des traits primitifs présents chez *Proganochelys*, d'autres sont déjà plus évolués. Citons, par exemple, la fusion du palais avec la base du crâne, l'apparition d'une cavité auditive et la fusion du bassin avec la carapace. Il en ressort que *Palaeochersis* est une tortue plus évoluée que *Proganochelys*. Ainsi, *Palaeochersis* est le plus ancien des rhabdochélidés, nouvelle catégorie créée par E. Gaffney et J. Kitching en 1994 pour regrouper *Australochelys africanus* et les tortues casichélidées. *Australochelys* est la tortue africaine la plus ancienne, dont nous sont parve-

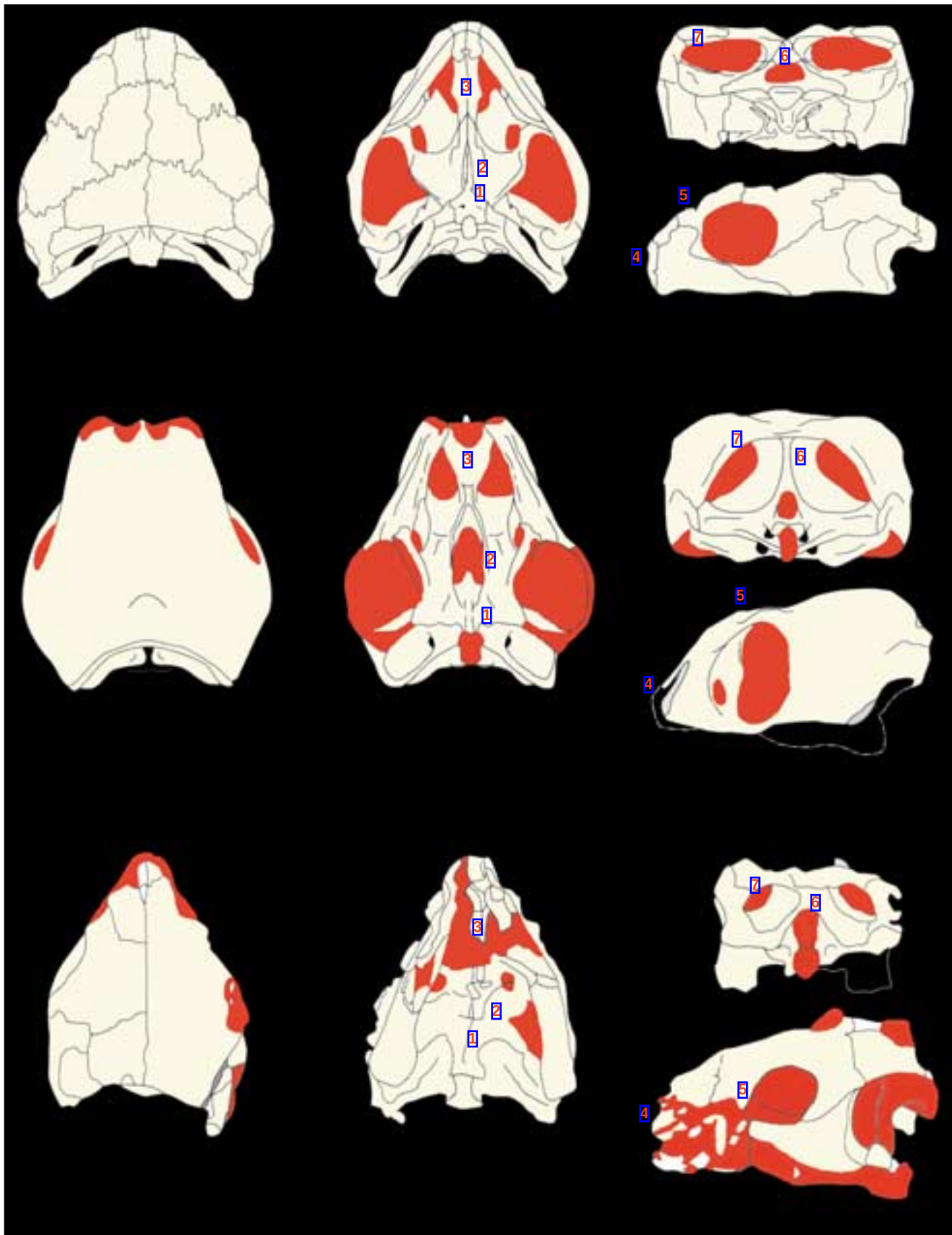
nus seulement un crâne et un fragment de carapace du Jurassique inférieur.

Australochelys, d'Afrique du Sud, et *Palaeochersis*, d'Argentine, sont regroupés au sein de la même famille, Australochelyidae, dans la mesure où ils partagent de nombreux traits dérivés communs : entre autres, des fosses nasales externes allongées, une plate-forme nasale, une plaque occipitale large avec des dépressions pour l'insertion de la musculature du cou et une fosse temporale partiellement refermée par un rebord du toit crânien.

La situation géographique des deux membres de la famille Australochelyidae, respectivement l'Afrique du Sud et l'Argentine, suggère qu'ils vivaient au Sud de la Pangée. Leur comparaison indique que la famille des Australochelyidae constitue le taxon frère de *Proterochersis* et des casichélidés. Cette position dans la séquence phylogénétique est intermédiaire entre *Proganochelys* et les autres tortues connues.

Palaeochersis et *Proganochelys* possèdent des vertèbres cervicales amphicoèles (corps cervicaux avec des faces articulaires concaves). Ils ne peuvent donc pas rentrer le cou à l'intérieur de la carapace. Cette tactique défensive, remarquable chez les tortues actuelles, s'effectue à présent, via la flexion du cou, dans un plan sagittal (Cryptodires) ou dans un plan latéral (Pleurodires). Chez *Palaeochersis*, en revanche, ce mécanisme est remplacé par une extension de la carapace qui couvre le cou. Chez *Proganochelys* le cou non rétracté est protégé par des ossifications dermiques en forme de pointes au-dessus des vertèbres cervicales. En résumé, au moins trois méthodes ont été utilisées pour protéger un cou vulnérable au cours de l'évolution du groupe : au moyen de pointes osseuses ou de la projection de la carapace chez les formes triasiques, par la flexion à l'intérieur de la carapace chez les groupes modernes.

D'autre part, *Palaeochersis* et *Proganochelys* partagent deux caractéristiques spécifiques : un hypo-ischion, ossification articulée au niveau du bassin dont sont dépourvus les autres tétrapodes, et la présence de seulement deux phalanges aux doigts de l'autopode antérieur (main) et postérieur (pied). Contrairement à *Proganochelys*, doté de cinq doigts comme la majorité des amniotes primitifs, la main et le pied de *Palaeochersis* ne possèdent



4. VUES DORSALES, ventrales, postérieures et latérales des crânes de (a) *Proganochelys quenstedti* (Trias supérieur allemand), (b) *Australochelys africanus* (Jurassique inférieur d'Afrique du Sud) et (c) *Palaeochersis talampayensis* (Trias supérieur argentin). Les crânes de *Palaeochersis* et d'*Australochelys* présentent des caractères plus évolués que *Proganochelys* et précisent la phylogénie des tortues :

la fusion du basiptérigoïde (1), la réduction des plaques dentaires à une seule rangée sur une crête (2), des vomers sans plaques dentaires (3), des fosses nasales externes allongées (4), une plateforme nasale (5), une large plate-forme occipitale (6) et des fosses temporales partiellement refermées par les rebords du toit crânien (7). Les zones rouges indiquent la présence de sédiment ou une ouverture.

que quatre doigts. La réduction du nombre de doigts et de phalanges, ajoutée à la morphologie particulière des membres antérieurs et postérieurs, suggèrent que *Palaeochersis* se déplaçait, du moins partiellement, en milieu terrestre, tout comme, pensent E. Gaffney et J. Kitching *Australochelys*.

Une vingtaine de nouveautés évolutives justifient la création du groupe Rhaptochelydia, qui comprend la famille Australochelyidae, les tortues modernes et les groupes apparentés. La majorité de ces caractères sont en relation avec la consolidation et la modification du palais et le développement de l'oreille. D'autres résultent de modifications de la musculature de la partie postérieure du crâne et du cou. Des 20 caractères communs qui définissent les rhaptochéliens, seuls deux portent sur l'arrière du crâne : le crâne des tortues a subi de nombreuses transformations, tandis que le squelette postcrânien, et particulièrement la carapace, n'a pas changé autant. La

carapace des tortues a gardé sa structure de base depuis le Trias.

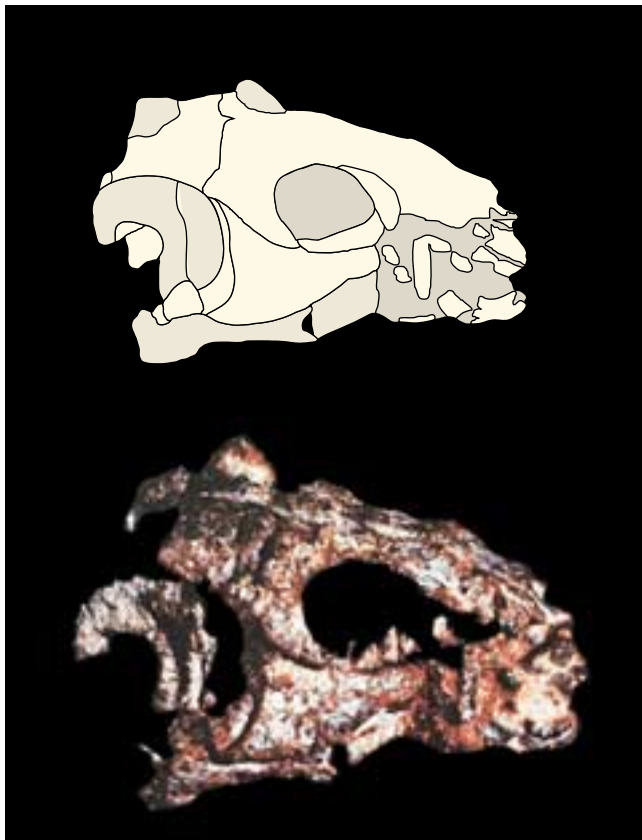
La séparation des deux groupes de tortues

Nous avons déjà mentionné que le bassin fusionné de *Proterochersis*, du Trias allemand, a été pris comme preuve de sa parenté avec les pleurodires. Toutefois, l'association d'un bassin fusionné avec la carapace et d'un crâne ne présentant pas les caractéristiques d'un pleurodire chez *Palaeochersis* nous amènent à exclure *Proterochersis* des pleurodires et à le considérer plus ancien que la séparation entre pleurodires et cryptodires. Dans un cadre plus général, notre hypothèse de relations des tortues primitives implique que la fusion pelvienne, considérée depuis toujours comme une caractéristique distinctive des pleurodires, constitue un trait primitif de la majorité des groupes de tortues, qui s'est par la suite perdu chez les cryptodires.

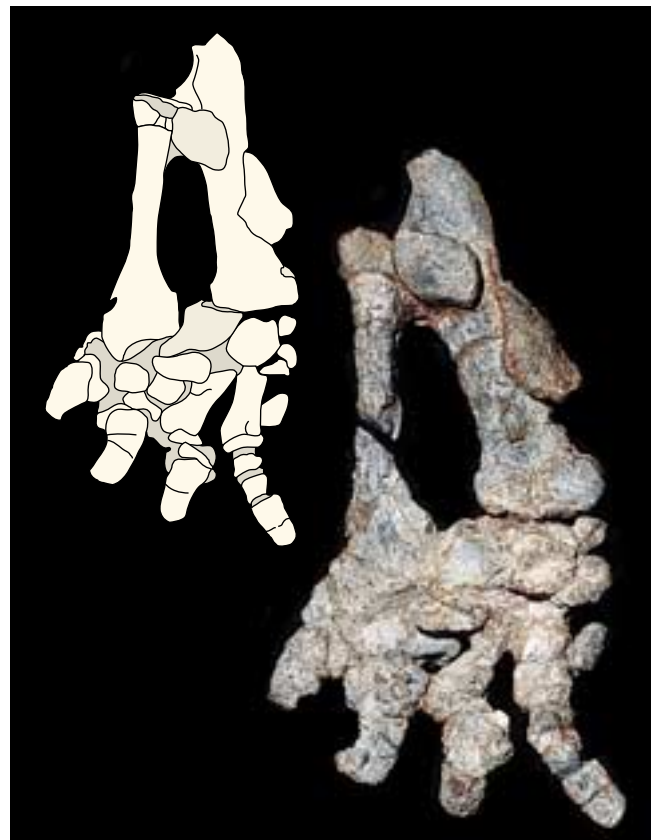
On peut déduire de la description

de *Proterochersis* que les tortues triasiques ne s'inscrivent dans aucun des groupes modernes, et donc que pleurodires et cryptodires (les deux groupes qui comprennent toutes les tortues vivantes) seraient apparus au Trias supérieur. Si les tortues casichéliées sont apparues à l'époque, leur apparition coïncide avec celle de la majorité des grands groupes d'animaux actuels, les anoures, les crocodiles ou les mammifères. Les spécimens les plus anciens de cryptodires sont ceux de *Kayentachelys* du Jurassique inférieur d'Arizona (185 millions d'années) et de pleurodires, ceux de *Platychelis* du Jurassique supérieur suisse, datant respectivement (156 millions d'années).

Les phylogénies des tortues sont fondées sur des caractères crâniens et, dans une moindre mesure, sur des caractères du squelette. Il n'est donc pas étonnant que les relations de parenté suggérées diffèrent de celles obtenues quand on se contente d'observer les traits postcrâniens. Il n'est pas facile d'harmoniser les deux critères, car nous



5. CRANE (a), AVANT-BRAS ET MAIN (b) de *Palaeochersis talampayensis*. Sur le crâne on reconnaît des traits primitifs tels que la persistance d'un supratemporal, disparu chez les tortues modernes, et le développement d'une cavité tympanique peu profonde, qui chez les quasichéliés, présente une forme d'embout caractéristique. La main de *Palaeochersis* ne dispose que de quatre doigts, comptant deux phalanges chacun. La réduction du nombre de phalanges



est présente chez des tortues terrestres de la sous-famille *Testudininae*, ainsi que le pélomédusidé aquatique *Pelomedusa* ; les tortues terrestres *Meiolaniidae* ont aussi perdu leur cinquième doigt. Cette morphologie suggère l'absence d'une membrane interdigitale chez *Palaeochersis*, présente chez les formes amphibiennes vivantes. Les spécialisations des pieds, la cambrure de la carapace et le paléoenvironnement indiquent que la tortue de Los Colorados vivait sur terre.

ne possédons parfois que les crânes de certains fossiles, tandis que nous n'avons que des carapaces pour d'autres.

Si nous avions découvert uniquement le crâne de *Palaeochersis*, nous serions parvenu à la conclusion phylogénique proposée ici. Toutefois, si nous n'avions récupéré qu'une carapace de Los Colorados, nous aurions pensé que certains pleurodires primitifs étaient présents dans le Trias d'Argentine (*Palaeochersis*) de même qu'en Allemagne (*Proterochersis*). Le bassin fusionné de *Palaeochersis* l'aurait inévitablement apparenté aux pleurodires, alors que la conclusion est différente lorsque l'on considère à la fois les caractères crâniens et postcrâniens.

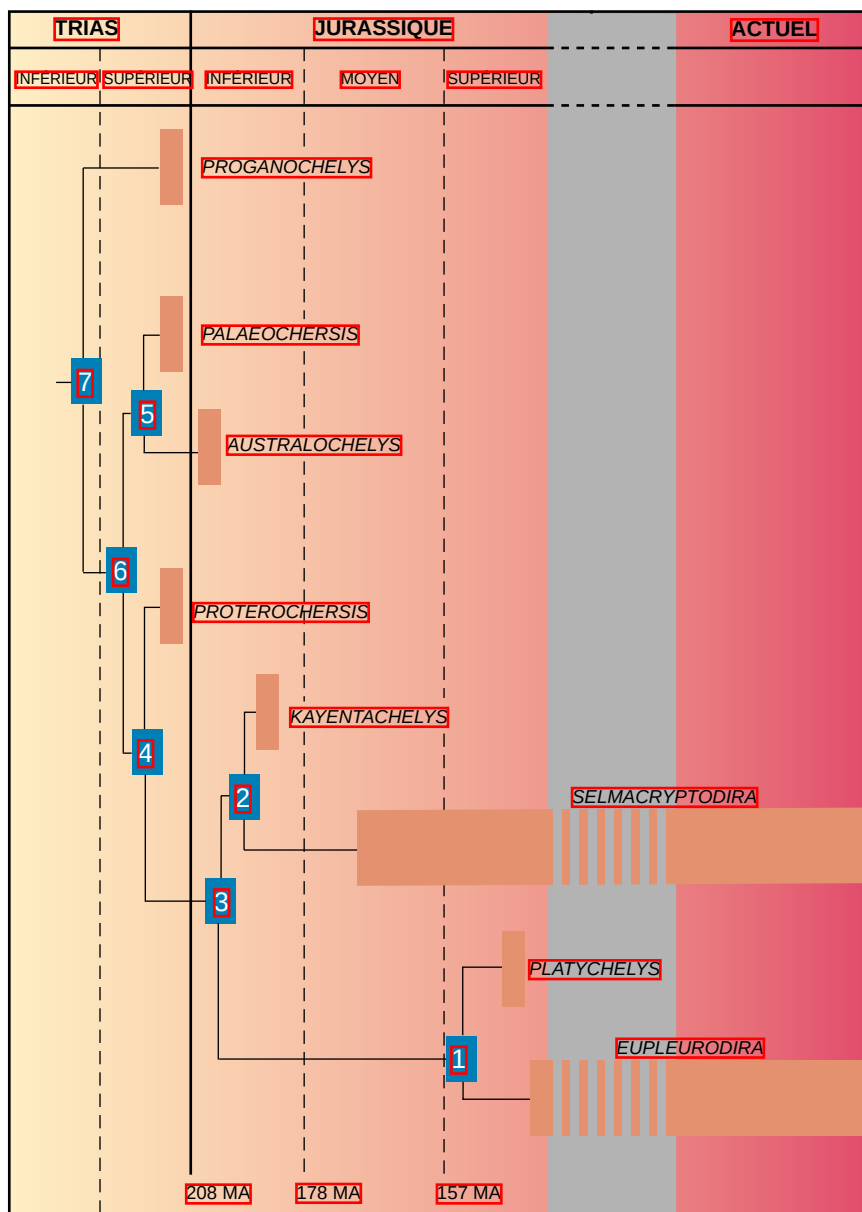
L'histoire des tortues enfin reconstituée

Nous constatons donc que, au cours de l'évolution des tortues, divers caractères sont successivement apparus, puis disparus. Par conséquent, la découverte de crânes et de squelettes associés est une condition indispensable pour élaborer des parentés entre les groupes de tortues.

Dès la fin du Trias, les tortues avaient montré des adaptations diverses dans des zones géographiques variées. Il existait trois groupes différents (*Proganochelys*, *Australochelyidae* et *Proterochersis*) qui couvraient les régions septentrionales et méridionales de la Pangée. Ce tableau du trio révèle une importante diversité écologique, avec des formes amphibies d'eau douce (*Proganochelys*) et des formes terrestres (*Proterochersis* et peut-être les *Australochelyidae*).

La faune du Trias semble très homogène, dans la mesure où des genres d'une même famille se retrouvent un peu partout. Sans remettre en cause cette affirmation, nous précisons que certaines familles sont mieux représentées dans la portion septentrionale de la Pangée triasique (tels les phytosaures) ou dans les continents du Sud (*Australochelyidae*). Cela suggère que, avant que ne se produise la scission du supercontinent Pangée, deux grandes sous-unités faunistiques s'étaient déjà différenciées.

Les caractéristiques principales des tortues se sont affirmées au Trias supérieur. Par comparaison avec les mammifères ou les dinosaures triasiques et les oiseaux, les modifications opérées depuis chez les Chéloniens



6. NOUVEL ARBRE PHYLOGENETIQUE DES TORTUES. Dès la fin du Trias, les tortues se sont diversifiées en au moins trois groupes habitant des régions septentrionales (*Proganochelys quenstedti*, *Proganochelys rucha*, *Proterochersis robusta*) et méridionales (*Palaeochersis talampayensis*) de la Pangée. Les *Australochelyidae* (*Australochelys* et *Palaeochersis*) constituent un groupe sans descendance intermédiaire dans la séquence phylogénétique entre *Proganochelys* et les autres tortues, qui n'ont pas laissé de descendants vivants. 1 = pleurodires ; 2 = cryptodires ; 3 = Casychelydia ; 4 = sans dénomination ; 5 = *Australochelyidae* ; 6 = *Rhaptochelydia* ; 7 = Testudines.

semblent constituer de légers ajustements d'un modèle ancien à l'efficacité prouvée. La diversité des groupes qui sont nés de ce plan commun, leur dispersion géographique et les divers environnements dans lesquels ils se trouvent prouvent leur faculté d'adaptation. De leur train de sénateur, les tortues ont sillonné une planète changeante pendant 200 millions d'années, alors que nombre de groupes, dont les dinosaures, ont disparu à jamais.

G. ROUGIER, Marcelo de la FUENTE et A. ARCUCCI sont spécialistes des tortues. G. ROUGIER est professeur à l'Université de Buenos Aires.

E.S. GAFFNEY et J.W. KITCHING, *The Most Ancient African Turtle*, in *Nature*, n° 369, pp. 55-58, 1994.

M.S.Y. LEE, *The Origin of the Turtle Body Plan*, in *Science*, vol. 261, pp. 1716-1720, 1993.

G.W. ROUGIER, et al, *Late Triassic Turtles from South America*, in *Science*, vol. 328, pp. 855-858, 1995.

Géométrie et religion au Japon

TONY ROTHMAN • HIDETOSHI FUKAGAWA

Quand le Japon s'isola, entre 1639 et 1854, des pratiques mathématiques originales apparurent. Notamment le peuple dédiait aux divinités des tablettes de bois ornées de problèmes géométriques.

De 1639 à 1854, quand le Japon ferma ses frontières, les idées scientifiques occidentales cessèrent de s'introduire, et les mathématiques se développèrent sans échange avec celles des autres pays du monde. C'est alors que fleurit une tradition d'une rare élégance : des samouraïs, des marchands et des paysans passionnés de mathématiques résolvaient des problèmes de géométrie variés, puis inscrivaient leurs résultats sur des tablettes de bois, qu'ils suspendaient sous les toits des édifices religieux. Selon les cas, ces *sangaku*, littéralement « tablettes mathématiques », étaient des offrandes aux divinités ou des défis aux autres visiteurs des lieux saints.

La plupart des *sangaku* traitaient de géométrie euclidienne, mais ils mentionnent des problèmes bien différents de ceux des cours de géométrie occidentaux : les cercles et les ellipses y jouent un rôle prépondérant. On considère ainsi des cercles dans des cercles, des cercles dans des ellipses, des ellipses dans des cercles... Certains de ces problèmes sont élémentaires ; d'autres sont des casse-tête dont la résolution nécessite des techniques avancées, notamment de calcul intégral ou de géométrie affine. La plupart de ces problèmes appartiennent aujourd'hui à la catégorie des mathématiques récréatives ou pédagogiques, mais certains résultats inscrits sur les tablettes votives japonaises anticipaient des résultats qui ne furent découverts qu'ultérieurement en Occident : le problème de Malfatti (inscrire dans un triangle trois

cercles tangents entre eux deux à deux et tangents chacun à deux côtés du triangle), le théorème de Casey (trouver la relation entre les rayons de quatre cercles tangents à un cinquième cercle ou à une droite), le théorème des six sphères de Soddy (voir la page 56) ou le théorème des cercles de Descartes (trouver la relation entre les rayons de quatre cercles tangents deux à deux) furent démontrés au Japon avant de l'être en Occident. Nombre de tablettes votives japonaises tiennent autant de l'œuvre d'art que de la publication mathématique.

Mathématiques divines

Comment cette pratique est-elle apparue ? Et quand ? L'accrochage de tablettes votives aux toits des sites religieux se pratiquait depuis plusieurs siècles quand apparurent les premiers *sangaku*. Le shintoïsme, religion officielle du Japon jusqu'en 1945, est un polythéisme qui admet « huit cents myriades de dieux », les *kamis*. Ces derniers, disait-on, aimaient les chevaux, et les fidèles trop pauvres pour leur offrir un cheval vivant pouvaient, à la place, en offrir une représentation dessinée sur bois. Aussi, jusqu'au XV^e siècle, les tablettes votives représentaient souvent des chevaux.

Le plus ancien *sangaku* connu, découvert dans la préfecture de Tochigi,

date de 1683. On a également trouvé, à Tokyo, une deuxième tablette datant de 1686, puis une troisième datant de 1691. Au XIX^e siècle, le mathématicien japonais Kazu Yamaguchi consigna dans son journal de voyage la description d'une tablette encore plus ancienne, qui aurait été réalisée en 1668. Les historiens estiment que la tradition des *sangaku* date de la seconde moitié du XVII^e siècle.

Le premier recueil de problèmes de *sangaku* fut publié en 1789. D'autres recueils parurent aux XVIII^e et XIX^e siècles. Écrits à la main ou imprimés à partir de matrices en bois, ces ouvrages sont très beaux. Plus de 880 tablettes qui subsistent aujourd'hui font elles-mêmes référence à des centaines d'autres tablettes qui figurent dans divers recueils. Un recensement des *sangaku* existants révèle que tous les milieux, urbains ou ruraux, pratiquaient la tradition de ces ex-voto mathématiques ; toutefois ces tablettes sont deux fois plus nombreuses dans les lieux de culte shintoïste que dans les temples bouddhistes.

La plupart des *sangaku* conservés contiennent plusieurs théorèmes et sont souvent ornés de couleurs brillantes. Les théorèmes sont généralement livrés sans démonstration. À côté du résultat figurent le nom de l'auteur et la date de l'offrande. Les problèmes ne traitent pas uni-

1. AU JAPON, de nombreux lieux de pèlerinages shintoïstes et bouddhistes abritent des *sangaku* : ces tablettes votives étaient traditionnellement suspendues sous les avant-toits des édifices religieux. Les problèmes consignés sur les *sangaku* traitent presque toujours de géométrie. Nombre de ces tablettes sont délicatement peintes ; certains problèmes ont été inscrits en idéogrammes dorés.



quement de géométrie. Certains, qui calculent le volume de solides, imposent l'usage du calcul intégral, de sorte qu'on peut se demander quelles techniques étaient connues et utilisées par les Japonais. Nous examinerons plus loin cette question. D'autres tablettes proposent des équations diophantiennes, c'est-à-dire des équations algébriques dont on cherche des solutions en nombres entiers.

Aujourd'hui, les *sangaku* n'intéressent plus que quelques passionnés des mathématiques japonaises traditionnelles, et, parmi eux, Hidetoshi Fukagawa, un enseignant de lycée dans la préfecture d'Aichi (à mi-chemin entre Tokyo et Osaka). Il y a une trentaine d'années, H. Fukagawa décida d'étudier l'histoire des mathématiques japonaises, afin d'enrichir son enseignement. La mention de ces tablettes mathématiques, dans un vieil ouvrage trouvé en bibliothèque, le surprit, car il ignorait l'existence de tels objets. Depuis, H. Fukagawa parcourt le Japon à la recherche de ces tablettes et il est devenu un expert en *sangaku*, mais aussi en histoire des mathématiques japonaises traditionnelles.

Pour ses recherches, H. Fukagawa a dû apprendre le *kambun*, du chinois classique augmenté de signes qui facilitent la lecture par les Japonais. Équivalent japonais du latin, le *kambun* était la langue scientifique de la période d'Edo (1603-1867). H. Fukagawa est mis à contribution chaque fois que l'on découvre de nouvelles tablettes.

Les origines d'une tradition

Si l'origine des *sangaku* ne peut être datée avec précision, elle peut être localisée. En japonais, le mot *wasan* désigne les mathématiques développées au Japon, par opposition au *yosan*, qui désigne les mathématiques occidentales. L'histoire des mathématiques japonaises ancienne explique comment le *wasan* et la curieuse coutume du *sangaku* sont apparus.

Au milieu du VI^e siècle, le bouddhisme s'introduisit au Japon, ainsi que les mathématiques chinoises. Dès le début du VIII^e siècle, le Japon avait importé les grands classiques chinois sur l'arithmétique, l'algèbre et la géométrie. Le plus ancien de ces ouvrages est le *Chou-pei Suan-ching*, qui contient le théorème de Pythagore et la figure habituellement associée à

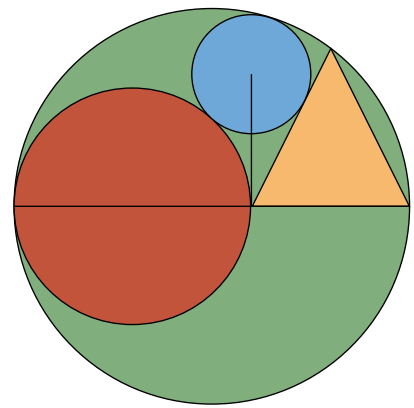
sa démonstration. Cette partie de l'ouvrage date au moins du VI^e siècle avant notre ère.

L'ouvrage intitulé *Chiu-chang Suan-shu*, qui fut le document chinois le plus important pour le développement des mathématiques japonaises, montre un état de connaissances plus avancé : il indique comment déterminer l'aire de triangles, de cercles et d'autres figures, et il résout des problèmes analogues à ceux des écoles d'aujourd'hui (Si cinq bœufs et deux moutons coûtent huit taels d'or, et si deux bœufs et huit moutons coûtent huit taels, combien coûte chaque animal?). La majeure partie de l'ouvrage fut certainement rédigée au III^e siècle avant notre ère : ainsi ce document serait le plus ancien à mentionner des nombres négatifs et à étudier une équation du second degré (certains historiens des mathématiques pensent toutefois que les Égyptiens avaient abordé les équations du second degré dès le troisième millénaire avant notre ère).

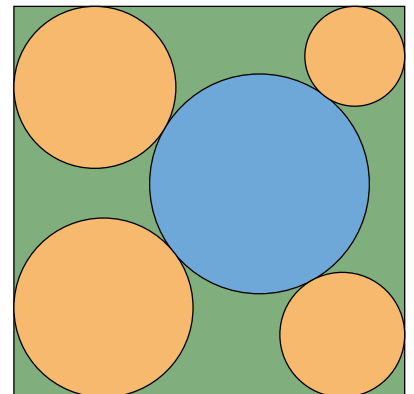
Malgré l'apport chinois, les mathématiques ne prirent pas racine au Japon. Le pays connut même une période d'obscurantisme quand l'Europe occidentale vivait son Moyen Âge. Comme les églises et les monastères européens, les temples bouddhistes devinrent les seuls centres culturels, même si les mathématiques ne semblent pas y avoir joué un rôle majeur. Durant le shogunat des Ashikaga (1338-1573), le Japon n'avait plus d'expert en division.

Les mathématiques japonaises ne naquirent véritablement qu'au début du XVII^e siècle. Un certain Shigeyoshi Mori fut particulièrement productif aux alentours de 1600. Bien qu'une seule de ses œuvres – un simple opuscule – nous soit parvenue, on sait qu'il développa le calcul sur le *soroban*, le boulier japonais, et fit connaître ce type de calcul dans le pays.

Un de ses élèves, Mitsuyoshi Yoshida (1598-1672), a laissé un ouvrage important. Ce document intitulé *Jinko-ki* (littéralement, «Des nombres grands et petits») fut publié en 1627. Il traite également du calcul sur le *soroban*, et son influence fut telle que son titre fut longtemps synonyme d'arithmétique : les mathématiciens japonais s'intéressèrent plus au calcul qu'à la logique. Mori et Yoshida furent ainsi les pères des mathématiques japonaises.



Dans ce problème, qui provient d'un *sangaku* datant de 1803 et trouvé dans la préfecture de Gumma, la base d'un triangle isocèle repose sur un diamètre du grand cercle vert. En outre, ce diamètre partage en deux le cercle rouge intérieurement tangent au cercle vert et en contact avec le sommet du triangle. Le cercle bleu est tangent au triangle et au cercle rouge, et intérieurement tangent au cercle vert. Un segment de droite relie le centre du cercle bleu au point de contact du cercle rouge et du triangle. Montrer que ce segment est perpendiculaire au diamètre du cercle vert.

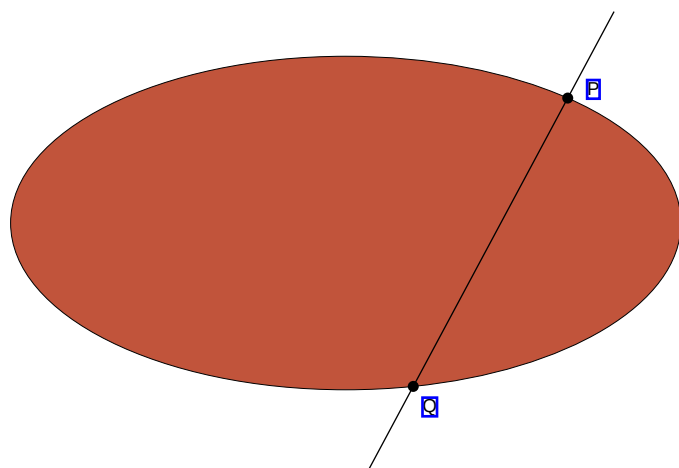
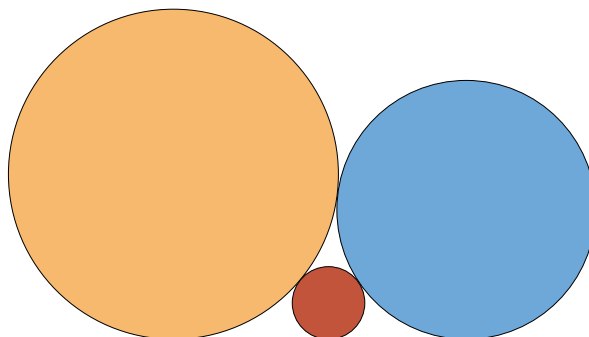


Ce problème provient d'une tablette de 1874 trouvée dans la préfecture de Gumma. Un grand cercle bleu se trouve dans un carré. Quatre petits cercles orange, de rayons différents, sont tangents d'une part au cercle bleu, d'autre part aux côtés du carré. Quelle relation unit les rayons des quatre cercles et le côté du carré (indication : ce problème peut se résoudre par le théorème de Casey, qui donne la relation entre quatre cercles tangents à un cinquième cercle ou à une droite)?

Quelques problèmes de *sangaku* dans le plan

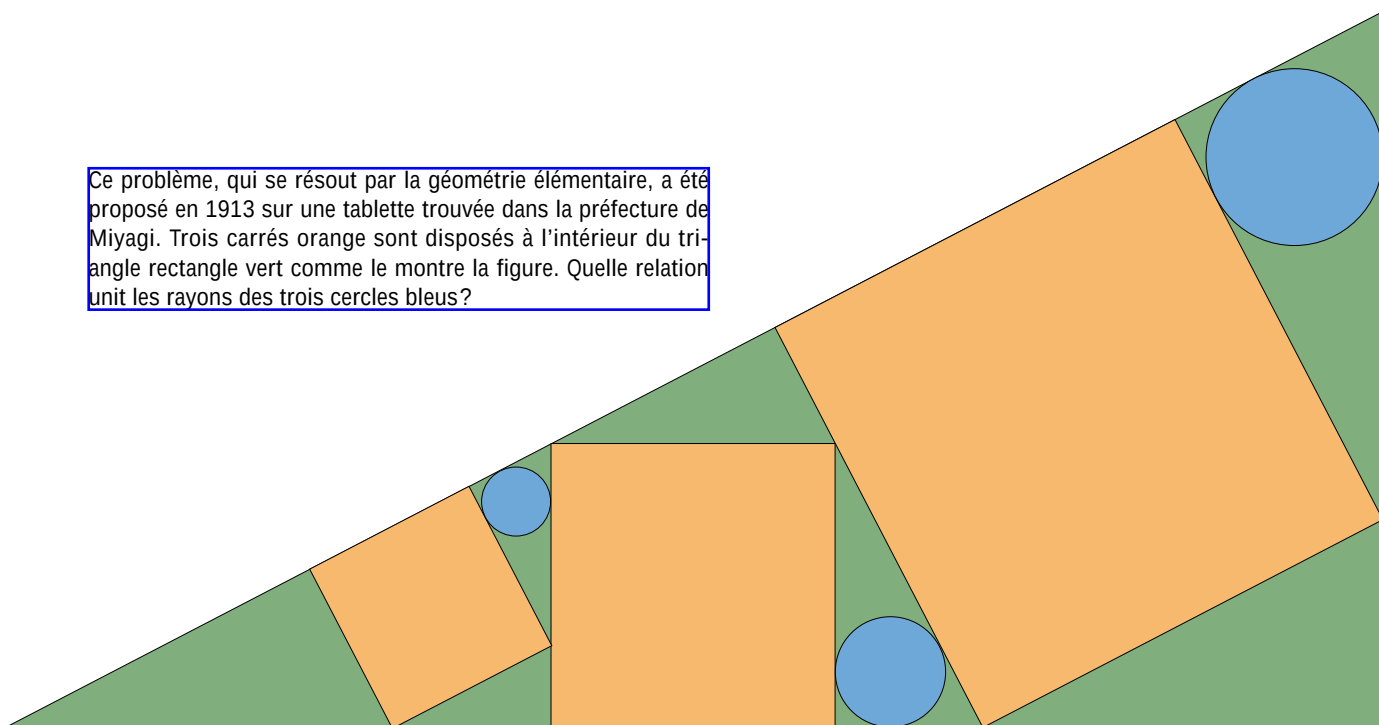
(les réponses sont données en page 58).

Ce problème simple figurait sur une tablette datant de 1824, retrouvée dans la préfecture de Gumma. Les cercles orange et bleu sont tangents et tous deux tangents à une droite. Le petit cercle rouge est tangent aux deux premiers cercles et à cette même droite. Quelle relation unit les rayons des trois cercles ?

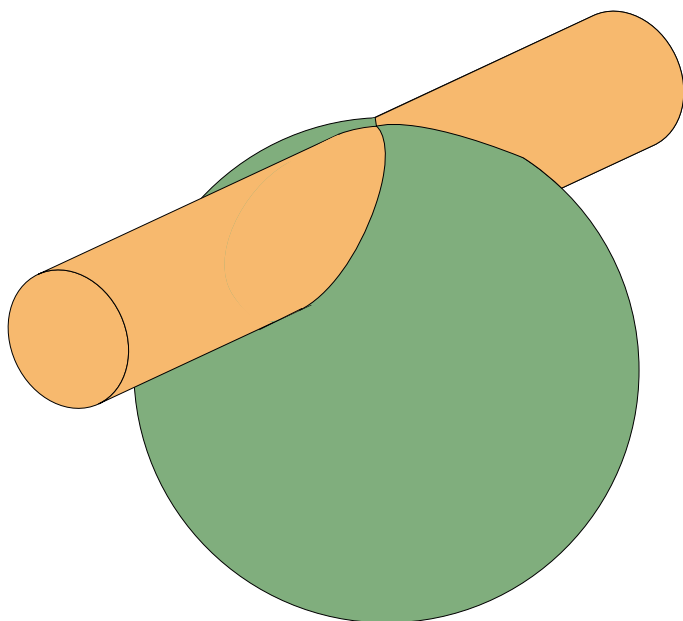


Ce problème fut consigné en 1912 sur une tablette conservée dans la préfecture de Miyagi. Par un point P d'une ellipse, on trace la normale à l'ellipse, et l'on nomme Q l'intersection de cette normale avec l'autre côté de l'ellipse. Quelle est la valeur minimale de PQ ? Ce problème semble simple : la valeur minimale de PQ est le petit axe de l'ellipse. Ce n'est toutefois la solution que si $b < a \leq \sqrt{2}b$, où a et b sont respectivement le grand axe et le petit axe de l'ellipse. En fait, la tablette ne donne pas cette solution, mais une autre, correspondant à $2b^2 < a^2$.

Ce problème, qui se résout par la géométrie élémentaire, a été proposé en 1913 sur une tablette trouvée dans la préfecture de Miyagi. Trois carrés orange sont disposés à l'intérieur du triangle rectangle vert comme le montre la figure. Quelle relation unit les rayons des trois cercles bleus ?

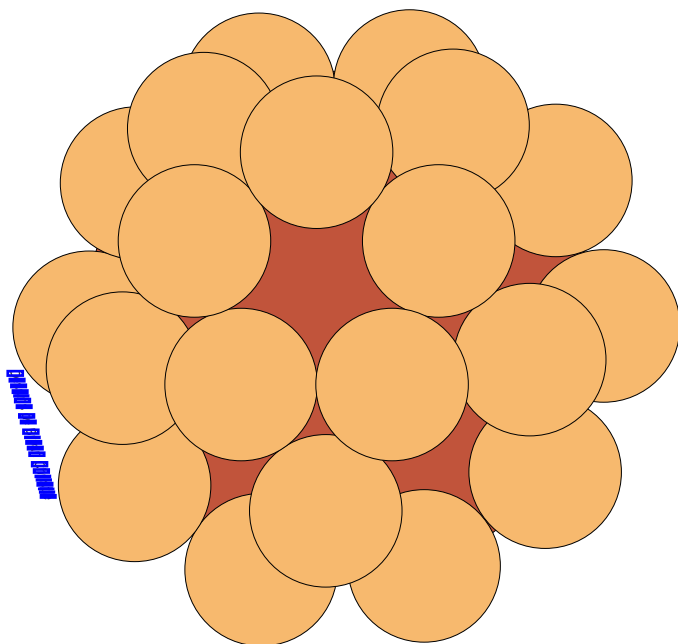
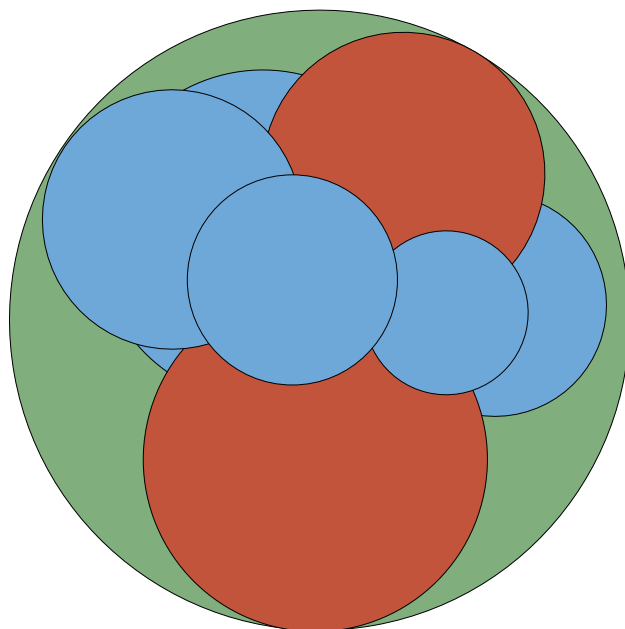


Quelques problèmes dans l'espace



Ce problème, provenant d'un *sangaku* de 1825, fut probablement résolu à l'aide de l'*enri*, ou principe du cercle dans les mathématiques japonaises. Un cylindre coupe une sphère tout en étant intérieurement tangent à cette sphère. Quelle est l'aire de la surface du cylindre contenue dans la sphère?

Ce problème était mentionné sur une tablette de 1822, retrouvée dans la préfecture de Kanagawa. Il anticipe de plus d'un siècle un théorème de Frederick Soddy, le chimiste britannique qui, avec Ernest Rutherford, découvrit la transmutation des éléments chimiques. Deux sphères rouges tangentes entre elles sont intérieurement tangentes à la grande sphère verte. Un «collier» de sphères bleues plus petites et de tailles différentes entoure le point de tangence des deux sphères rouges. Les sphères bleues du collier sont deux à deux tangentes, tangentes aux sphères rouges et tangentes à la sphère verte. Quel est le nombre des sphères bleues? Quelle relation unit les rayons des sphères bleues?



Hidetoshi Fukagawa fut si fasciné par ce problème, qui date de 1798, qu'il en a construit un modèle en bois. Soit une grande sphère entourée de 30 petites sphères identiques, chacune tangente à ses quatre voisines et à la grande sphère. Quelle relation unit le rayon de la grande sphère et celui des petites sphères?

Isolement et géométrie

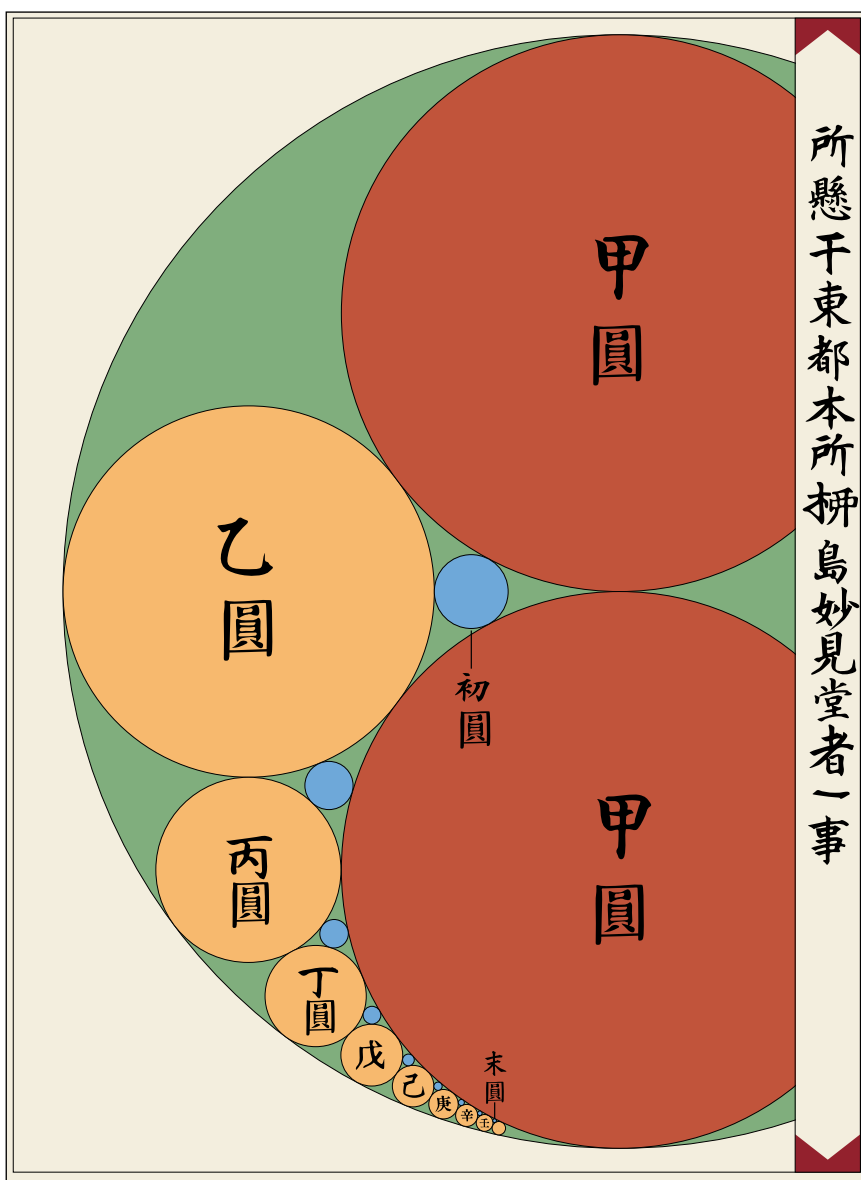
Les mathématiques qui s'affichèrent sur les *sangaku* furent toutefois l'œuvre de nombreux individus, et non celle d'une élite. En 1639, le shogunat des Tokugawa (durant la période d'Edo) décréta de fermer les frontières du Japon afin de renforcer son pouvoir. Durant cette période de *sakoku*, de «fermeture du pays», le gouvernement proscrivit les ouvrages étrangers, interdit les voyages à l'étranger, persécuta les chrétiens et interdit l'accostage des navires portugais et espagnols. L'isolement persista durant plus de deux siècles, jusqu'à ce que le capitaine Matthew Perry, à la tête d'une flotte de guerre américaine, oblige le Japon à mettre fin à son *sakoku*, en 1854.

L'isolement du pays eut quelques avantages : la fin du XVII^e siècle fut une renaissance, ou *Genroku*. Les arts du *haïku*, des théâtres *no* et *kabuki* atteignirent leur apogée ; l'*ukiyo-e*, l'art des peintures d'un «monde flottant», apparut ; la cérémonie du thé et l'art du bouquet se développèrent, tandis que les mathématiques prospéraient autour de Takakazu Seki (1642-1708).

Au Japon, Seki est aussi connu qu'Isaac Newton ou Antoine Laurent de Lavoisier en Europe. On lui a attribué un très grand nombre de manuscrits et l'on sait qu'il forma de nombreux élèves, qui développèrent les mathématiques japonaises.

Le premier accomplissement de Seki fut sa théorie des déterminants, qui est plus puissante que celle de Gottfried Wilhelm Leibniz et la précède d'au moins une décennie. Il trouva également des méthodes pour résoudre des équations de degré élevé (une grande partie des mathématiques traditionnelles japonaises de cette époque concerne des équations de degré atteignant plusieurs centaines ; l'une d'elles est de degré 1 458). Enfin Seki semble avoir été à l'origine de l'*enri*, ou «principe du cercle».

L'*enri* était un champ de recherches mathématiques qui s'est développé après Seki. À la base, il utilise des méthodes analogues à celles qu'employaient Eudoxe et Archimède pour calculer l'aire des cercles : à l'aide de polygones à n côtés, on approche la circonférence du cercle et l'on utilise de telles approximations pour calculer des aires et des volumes, par exemple. Ainsi l'*enri* fut une forme de calcul intégral, appliquée ensuite à d'autres figures,



2. LES PLAQUETTES VOTIVES accrochées dans les temples japonais (*en haut*) mentionnent souvent des problèmes mathématiques où figurent des cercles emboîtés ou des sphères inscrites à l'intérieur d'autres figures. Le problème reproduit en bas figurait sur une plaquette votive peinte en 1788, dans la préfecture de Tokyo. On cherche le rayon du n -ième cercle bleu en fonction du rayon r du cercle vert. Les deux cercles rouges sont identiques et leur rayon est égal à $r/2$ (indication : le rayon du cinquième cercle bleu est égal à $r/95$). La solution de ce problème fait intervenir l'équivalent japonais du théorème des cercles de Descartes, qui détermine la relation entre les rayons de quatre cercles tangents deux à deux ; elle est donnée page 58.

telles les sphères et les ellipses. Les techniques de l'*enri* ont été mises en œuvre dans la réalisation de *sangaku*.

Des sphères dans des ellipsoïdes

Les premiers livres utilisant l'*enri* et les premiers *sangaku* attestés sont contemporains de Seki. Il ne s'agit pas là d'une coïncidence : les disciples de Yoshida et de Seki contribuèrent probablement au développement du *wasan* en même temps qu'ils en subissaient l'influence.

Certains racontent que Seki a rencontré son premier *sangaku* alors qu'il se rendait au château du shogun, dont il était le mathématicien officiel ; les tablettes votives l'auraient incité à approfondir ses recherches. Est-ce une légende ?

On ne le saura sans doute jamais, mais, au siècle suivant apparurent des livres contenant des problèmes typiquement japonais, mettant en jeu des cercles à l'intérieur de triangles, des

sphères à l'intérieur de pyramides, des ellipsoïdes entourant des sphères... Les problèmes trouvés dans ces ouvrages diffèrent peu de ceux qui ont été retrouvés sur les tablettes, les historiens ont conclu que l'engouement pour tous les problèmes de *wasan* – et la vogue des *sangaku* – résulta de la politique d'autarcie du Japon.

Cette autarcie fut-elle totale ? À l'exception des Hollandais, autorisés à accoster dans le port de Nagasaki, sur l'île méridionale de Kyushu, tous les commerçants occidentaux étaient interdits de séjour. De surcroît, les déplacements des Japonais étaient sévèrement limités : un voyage à l'étranger était un acte de trahison, passible de la peine de mort. L'autarcie fut suffisamment complète pour minimiser l'influence des mathématiques étrangères sur les mathématiques japonaises.

En revanche, au XIX^e siècle, le *wasan* fut progressivement supplanté par le *yosan*. Apparurent alors des manus-

crits mêlant des textes en *kambun* et des notations mathématiques occidentales. Puis, après l'ouverture du Japon et l'effondrement du shogunat des Tokugawa, en 1867, le nouveau gouvernement adopta le *yosan* et délaissa les mathématiques locales. Certains praticiens du *wasan* continuèrent de suspendre des tablettes (quelques *sangaku* datent même de la dernière décennie), mais la quasi-totalité des problèmes proposés durant le XX^e siècle sont des plagiat.

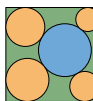
Qui étaient les auteurs des *sangaku* ? Les théorèmes merveilleusement dessinés sur les tablettes votives étaient-ils l'œuvre de mathématiciens professionnels ou de simples amateurs ? La question est difficile. Le traité d'histoire des mathématiques japonaises de David Smith et de Yoshio Mikami ne mentionne que quelques *sangaku*. Il cite le recueil *Shimpeki Sampo* («Problèmes mathématiques suspendus devant le temple»), publié en 1789 par le mathématicien profes-

Réponses aux problèmes de *sangaku*

Nous ne donnons ici que l'indication des solutions. Les lecteurs intéressés trouveront les solutions détaillées sur le site Internet de la revue *Pour la Science*, à l'adresse <http://www.pourlascience.com>.

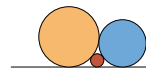


Réponse : Dans sa solution, l'auteur du problème trace un segment de droite qui passe par le centre du cercle bleu et qui est perpendiculaire au diamètre du cercle vert. Il suppose ensuite que ce segment est différent du segment décrit dans l'énoncé du problème. Ainsi, ces deux segments devraient couper le diamètre tracé en deux points différents. Il montre ensuite que la distance entre ces deux points est nécessairement nulle : autrement dit, les deux segments sont identiques et, donc, perpendiculaires au diamètre du cercle vert.



Réponse : Si a désigne le côté du carré et r_1, r_2, r_3 et r_4 sont respectivement les rayons des cercles supérieur droit, supérieur gauche, inférieur gauche et inférieur droit, alors :

$$a = \frac{2(r_1 r_3 - r_2 r_4) + \sqrt{2(r_1 - r_2)(r_1 - r_4)(r_3 - r_2)(r_3 - r_4)}}{r_1 - r_2 + r_3 - r_4}$$



Réponse : $\frac{1}{\sqrt{r_3}} = \frac{1}{\sqrt{r_1}} + \frac{1}{\sqrt{r_2}}$, où r_1, r_2 et r_3 sont respectivement les rayons des cercles orange, bleu et rouge. On peut résoudre le problème à l'aide du théorème de Pythagore.



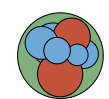
Réponse : $PQ = \sqrt{27 a^2 b^2 / (a^2 + b^2)^{3/2}}$. On peut résoudre ce problème à l'aide de la géométrie analytique : on cherche une équation pour PQ ; puis on prend la dérivée de cette équation et on l'égale à zéro pour obtenir la valeur minimale de PQ. On ignore si les premiers auteurs de ce problème l'ont résolu en recourant au calcul différentiel.



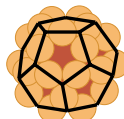
Réponse : $r_2^2 = r_1 r_3$, où r_1, r_2 et r_3 sont respectivement les rayons des grand, moyen et petit cercles bleus (autrement dit, r_2 est la moyenne géométrique de r_1 et r_3). On peut résoudre le problème en observant que les triangles verts délimités par les carrés orange sont tous semblables. La solution, telle qu'elle fut initialement proposée, consiste alors rechercher la relation unissant les trois carrés.



Réponse : $\frac{16t}{\sqrt{t(r-t)}}$, où r et t sont respectivement les rayons de la sphère et du cylindre.



Réponse : Six sphères. Le théorème des six sphères de Soddy stipule que le nombre de sphères est égale à six. Curieusement, les rayons $t_1, t_2, \dots, t_6$ des diverses sphères bleues qui composent le collier sont liés par les relations : $\frac{1}{t_1} + \frac{1}{t_4} = \frac{1}{t_2} + \frac{1}{t_5} = \frac{1}{t_3} + \frac{1}{t_6}$.



Réponse : $R = \sqrt{5r}$, où R et r sont respectivement les rayons de la grande sphère et des petites sphères. On peut résoudre le problème en observant que le centre de chaque petite sphère se trouve au milieu d'une arête de dodécaèdre régulier – un solide régulier à douze faces pentagonales.



Réponse : $r / [(2n - 1)^2 + 14]$. La solution originelle de ce problème fait intervenir à plusieurs reprises la version japonaise du théorème du cercle de Descartes. On peut toutefois utiliser l'inversion, qui était à l'époque inconnue des mathématiciens japonais.



3. LES EX-VOTO MATHÉMATIQUES DU JAPON furent parfois l'œuvre de mathématiciens professionnels, mais, parfois aussi, ils furent réalisés par des commerçants, des samourais, des enfants... Sur

ces plaques de bois, sont généralement mentionnés le problème, le nom de celui qui a réalisé l'ex-voto, la date de l'offrande ; certains *sangaku* sont des chefs-d'œuvre.

sionnel Kagen Fujita. D. Smith et Y. Mikami mentionnent une tablette sur laquelle la solution proposée est suivie de la référence suivante : « District féodal de Kakegawa, province d'Enshu, troisième mois de 1795, Sonobei Keichi Miyajima, élève de Sadasuke Fujita, de l'école de Seki ». Dans son ouvrage consacré au développement des mathématiques en Chine et au Japon, Y. Mikami mentionne le « problème du temple de Gion », qui fut suspendu au temple de Gion, à Kyoto, par Enkyu Tsuda, élève d'Enri Nishimura. En outre, ces tablettes furent rédigées en *kambun*, signe que leurs auteurs appartenaient à une classe sociale cultivée.

D'autres observations montrent toutefois que les tablettes votives ne furent pas l'apanage des mathématiciens professionnels. Notamment nombre de problèmes proposés sont élémentaires et ne sont pas de ceux qu'étudieraient des professionnels. Dans la préfecture de Mie, H. Fukagawa a trouvé une tablette signée par un marchand. Sur d'autres tablettes figurent des noms de femmes et d'enfants âgés de 12 à 14 ans. Selon H. Fukagawa, la majorité des problèmes furent inventés par les samourais, mais quelques uns furent proposés par des paysans : le mathématicien Sen Sakuma (1819-1896), par exemple,

enseignait le *wasan* aux fermiers de la préfecture de Fukushima et eut jusqu'à 2 000 élèves.

Ce type d'enseignement rappelle la période d'Edo, où le Japon n'avait ni lycées ni universités. Le peuple japonais venait alors apprendre à lire, à écrire et à calculer sur des bouliers dans des écoles privées ou dans des temples.

Des individus de toutes les classes sociales semblent donc avoir contribué aux *sangaku* : de nombreux Japonais s'adonnaient avec autant de plaisir aux mathématiques qu'à la poésie et à d'autres formes d'art, parce qu'ils étaient attirés par la beauté de la géométrie. Les maîtres, au terme de leur journée d'enseignement, ou les samourais qui avaient cessé de

combattre s'isolaient pour concevoir des problèmes mettant en jeu des sphères et des ellipsoïdes ; la sereine contemplation mathématique était alors un loisir. Quand la solution était trouvée, on l'écrivait sur une plaque de bois que l'on suspendait ensuite au temple local. Les visiteurs du temple remarquaient la tablette colorée et admiraient sa beauté. Nombre d'entre eux quittaient alors le temple en se demandant comment l'auteur du problème était parvenu à une telle solution. Certains décidaient de résoudre le problème ou d'étudier la géométrie afin de tenter de le résoudre. D'autres quittaient le temple en se demandant « Que se passerait-il si l'on modifiait légèrement le problème et que... »

Tony ROTHMAN est professeur de mathématiques à l'Université Harvard. Hidetoshi FUKAGAWA est professeur de mathématiques dans la préfecture d'Aichi.

Annick HORIUCHI, *Les mathématiques japonaises à l'époque d'Edo (1600-1868), une étude des travaux de Seki Takakazu et de Takebe Katahiro*, Librairie philosophique J. Vrin, Paris, 1994.

Yoshio MIKAMI, *The Development of Mathematics in China and Japan*, deuxième édition, Chelsea Publishing Company, New York, 1974.

H. FUKAGAWA et D. PEDOE, *Japanese Temple Geometry Problems*, Charles Babbage Research Foundation, Winnipeg, Canada, 1989.

David E. SMITH et Yoshio MIKAMI, *A History of Japanese Mathematics*, Open Court Publishing Company, Chicago, 1914 (également disponible sur microfilm).

H. FUKAGAWA et D. SOKOLOWSKY, *Traditional Japanese Mathematics Problems from the 18th and 19th Centuries*, Science Culture Technology Publishing, Singapour, à paraître.

Nitrate : un polluant de longue durée

ANDRE MARIOTTI

La durée de résidence dans le sol de l'azote apporté par les engrais est très longue. Les nitrates qui polluent aujourd'hui les nappes aquifères résultent de nombreuses années de fertilisation.

L'azote, constituant des protéines et de nos gènes, et nutriment essentiel des plantes, est, avec l'eau et le carbone, l'élément dont le cycle naturel est le plus perturbé par les activités humaines. Sur la frange superficielle de la Terre, l'azote est principalement présent dans la molécule diazote, formée de deux atomes d'azote (N_2), qui entre pour quatre cinquièmes dans la composition de l'atmosphère (soit 39×10^{20} grammes d'azote). La fixation, transformation de cette molécule en composés, organiques ou minéraux, où l'azote est combiné, était de l'ordre de 100 milliards de kilogrammes d'azote par an dans la période pré-industrielle. Elle était principalement l'œuvre de bactéries, terrestres ou aquatiques, et des algues bleues. Aujourd'hui, les productions d'engrais, combustions et cultures de légumineuses ont plus que doublé la quantité fixée annuellement.

Cette perturbation a des conséquences dans l'atmosphère, localement, par la formation de *smog* et la chute de pluies acides dus à des excès d'oxydes d'azote gazeux, mais aussi à l'échelle globale : les bulles d'air piégées dans les glaces polaires enregistrent depuis le XIX^e siècle une augmentation de la concentration d'oxyde nitreux dans l'atmosphère, qui est actuellement de 0,3 pour cent par an. Dans ces glaces, notamment au Groenland, une augmentation de la concentration en nitrate a aussi été mesurée, de 40 parties par milliard vers 1750 à plus de 120 parties par milliard aujourd'hui.

La composition chimique des eaux souterraines ou de surface est aussi modifiée : elle contient de plus en plus de nitrate. Parmi les plus grands réservoirs d'eaux souterraines, beaucoup présentent des qualités chimiques ou bactériologiques satisfaisantes pour des usages domestiques, à l'exception d'un critère, le nitrate, dont la concentration dépasse la norme de potabilité de 50 milligrammes par litre, établie par l'Organisation mondiale de la santé. L'élévation de ces concentrations a commencé dès la fin des années 1950 et, dans la plupart des régions du globe, elle se poursuit encore.

Le nitrate (ion négatif formé d'un atome d'azote et de trois atomes d'oxygène, NO_3^-) est au cœur du cycle de l'azote. Il est synthétisé dans le sol par l'action de bactéries qui transforment d'abord en ammonium (ion positif formé d'un atome d'azote et de quatre atomes d'hydrogène, NH_4^+) une partie de l'azote contenu dans les matières organiques mortes du sol (c'est la minéralisation), puis oxydent l'ammonium en nitrate (c'est la nitrification).

Le nitrate est stable en présence d'oxygène, dans les sols aérés que l'on peut cultiver. Il est chargé négativement et il n'est pas retenu par les colloïdes argileux ou organiques du sol, eux aussi chargés négativement : il est facilement dissous dans l'eau d'infiltration, qui l'entraîne dans les nappes aquifères. L'ion nitrate, fabriqué à partir de matières organiques, peut être retransformé en matières organiques (c'est la réorganisation) ; nutriment des plantes, il est aussi consommé par

la biomasse des sols, bactéries ou champignons, et contribue, via ces micro-organismes, à la synthèse de matières organiques du sol, puis d'humus. Ces réorganisations, minéralisations et oxydations successives constituent le cycle interne de l'azote dans le sol (voir la figure 1).

La perturbation du cycle de l'azote dans les agrosystèmes est exemplaire des défis que l'agriculture du XXI^e siècle devra affronter. Depuis le début du XX^e siècle, en effet, la productivité agricole a augmenté grâce aux engrais et aux produits phytosanitaires, générateurs de pollutions des sols, des eaux et de l'atmosphère. Dans les prochaines années, nous devons maintenir la productivité, voire l'accroître encore, tout en réduisant ces pollutions.

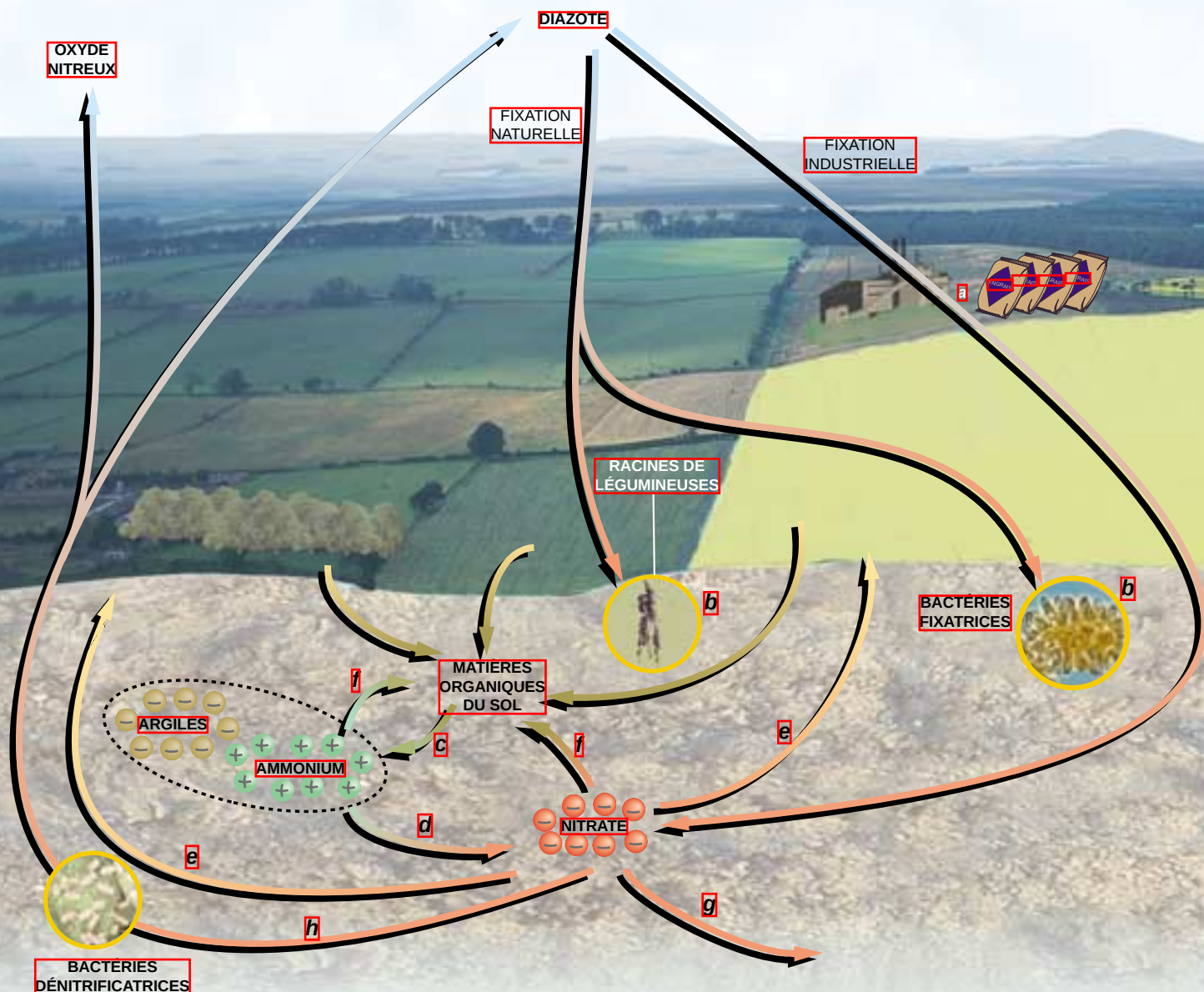
Le cas des nitrates illustre l'intérêt de l'approche géochimique des questions d'environnement. Au-delà de l'identification chimique d'un composé polluant, nous déterminons deux données essentielles : l'origine du composé et son temps moyen de résidence dans le milieu. Nous avons ainsi montré que l'azote contenu dans les nitrates lessivés a été introduit dans le sol plusieurs années auparavant et que les matières organiques du sol en sont la principale source. Ces matières organiques du sol productrices de nitrates peuvent aussi les éliminer mais, dans le cycle très complexe de l'azote, on court sans cesse le risque de transférer vers l'atmosphère la pollution du sol et de l'eau : aussi les bilans écologiques raisonnables et les tentatives de dépollution prennent-ils en compte l'ensemble de la biosphère.

Les agriculteurs ne fertilisent pas trop

L'agriculture est en grande partie responsable de l'accroissement des concentrations de nitrate dans les nappes aquifères. Toutefois, aujourd'hui, les agriculteurs savent épandre les engrais afin que les pertes restent faibles. Examinons, par exemple, le bilan azoté d'une culture de blé. Si l'objectif de production est de 85 quin-

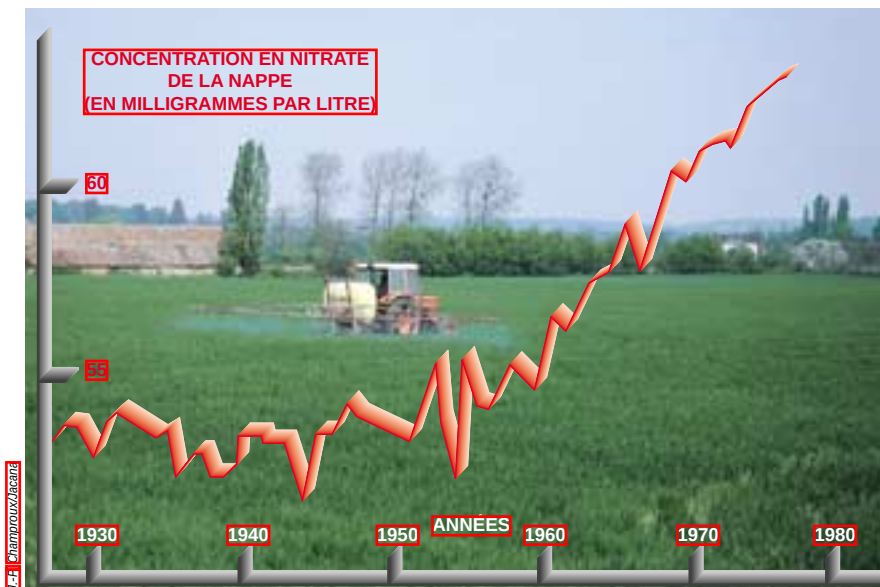
taux à l'hectare, il convient de fournir à la plante 255 kilogrammes d'azote par hectare. Les matières organiques du sol en apportent une partie par minéralisation et par nitrification, et le reste provient des engrais, principalement sous forme de nitrate et d'ammonium. Même lorsque la fertilisation azotée est bien gérée (analyse préalable des sols, mesure des reliquats d'azote avant la fertilisation, utilisation de modèles de fertilisation et de logiciels),

des pertes se produisent : les eaux d'infiltration emportent des nitrates. Ces pertes ne dépassent généralement pas 30 kilogrammes d'azote par hectare et par an, soit 11,5 pour cent seulement de la quantité totale mise en jeu pour la culture. Peu de systèmes techniques fonctionnent avec aussi peu de pertes. Malheureusement, ces faibles pertes relatives, dissoutes dans la lame d'eau qui s'infiltre en climat tempéré, sont à l'origine d'une concentration de



1. LE CYCLE DE L'AZOTE a été fortement perturbé par la fixation artificielle du diazote atmosphérique en engrais (a). La quantité d'azote ainsi injectée dans le sol est le double de celle fixée naturellement, surtout par les racines des plantes légumineuses et par des bactéries du sol (b). L'azote est stocké dans les matières organiques du sol. Celles-ci sont minéralisées en ammonium (c), stabilisé par son association électrostatique avec les colloïdes argileux. La nitrifica-

tion de l'ammonium (d), produit du nitrate, qui est assimilé par les plantes (e), réorganisé en matières organiques (f), lessivé dans la nappe aquifère (g), ou réduit par des bactéries dénitrificatrices (h). Le lessivage et la réduction de trop grandes quantités de nitrate sont à l'origine de pollutions, de l'eau mais aussi de l'atmosphère : la dénitrification produit du diazote, mais aussi de l'oxyde nitreux, gaz à effet de serre et destructeur de l'ozone stratosphérique.



2. LA CONCENTRATION EN NITRATE augmente dans la nappe aquifère de la Brie à cause de l'épandage d'engrais azotés. Bien que les agriculteurs maîtrisent mieux cet épandage depuis quelques années, la norme de potabilité de l'eau, de 50 milligrammes par litre, est dépassée, comme dans la plupart des régions de grandes cultures.

53 milligrammes de nitrate par litre en moyenne, déjà supérieure à la norme de potabilité.

Comment réduire ces concentrations? Les agriculteurs pourraient-ils réduire les pertes par une gestion encore meilleure de la fertilisation et un contrôle plus fin du cycle de l'azote dans le sol? C'est peu vraisemblable, pour deux raisons. D'abord, le lessivage des nitrates dépend beaucoup de paramètres incontrôlables, tels que la température ou la pluviométrie. Ensuite, les nitrates lessivés au cours d'une année donnée ne proviennent que pour une toute petite part du lessivage direct des engrais apportés cette année-là : leur origine principale est la matière organique du sol, qui stocke l'azote pendant très longtemps.

À la station INRA de Châlons en Champagne, avec Bernard Nicolardot, nous suivons les flux des minéraux qui sortent du sol dans les eaux d'infiltration, à l'aide de cases expérimentales. Il s'agit de cubes de sol non perturbé de deux mètres de côté, dont les flancs latéraux sont murés et dont le plancher est une plaque d'inox perforée : nous recueillons, dans un espace vide situé au-dessous, l'eau qui a traversé la case et nous l'analysons.

Chaque année, le sol est fertilisé en surface avec l'équivalent de 150 kilogrammes d'azote par hectare. En 1982, nous avons utilisé un engrais qui contenait un isotope stable de

l'azote, rare dans le milieu naturel, l'azote 15 (dont le noyau atomique contient 15 protons et neutrons, au lieu de 14 dans la forme d'azote la plus répandue). Depuis 16 ans, des nitrates contenant ces atomes d'azote 15 ressortent de la case expérimentale à deux mètres de profondeur, à raison, chaque année, de 1 à 1,5 pour cent de l'apport initial de 1982 (voir la figure 3). La même expérience montre qu'un sel non réactif (qui n'est pas, comme le nitrate, réorganisé dans le sol), tel le chlorure, appliqué dans les mêmes conditions, est éliminé en trois ou quatre ans.

L'azote apporté sous forme d'engrais et qui n'est pas consommé dans l'année par les plantes reste donc longtemps dans le sol. Il est réorganisé en matières organiques à renouvellement rapide qui participent au cycle interne de l'azote, que nous avons décrit au début de cet article. Les matières organiques se minéralisent progressivement, puis sont nitrifiées et elles libèrent des nitrates dans le milieu. Les nitrates lessivés proviennent donc en partie de fuites de ce cycle, qui surviennent à toute époque de l'année. Par exemple, l'hiver, si la température remonte et qu'il pleut, les bactéries qui assurent la minéralisation et la nitrification reprennent leur activité alors qu'aucune plante en période de forte croissance ne consomme le nitrate : celui-ci est emporté dans la nappe aquifère.

Des pertes faibles, mais durables

Sur 15 ans, alors qu'il sort en moyenne du sol 38 kilogramme d'azote par hectare et par an sous forme de nitrate, la perte moyenne annuelle correspondant à l'apport de 1982 est de seulement 0,7 kilogramme d'azote par hectare. Chaque apport annuel est donc responsable d'une petite perte annuelle, mais pendant très longtemps : ces pertes s'additionnent, expliquant les grandes quantités de nitrate perdues après une quarantaine d'années de fertilisation régulière.

Cet énorme tampon de matière azotée que sont les matières organiques mortes du sol et l'humus (environ 50 000 milliards de kilogrammes d'azote sur Terre) est donc une source importante des nitrates lessivés, et qui les produit lentement. Nous ne comprendrons complètement le cycle de l'azote que lorsque nous aurons défini et caractérisé les différents types de matières organiques impliqués dans le cycle interne court (organisation et minéralisation), et que nous connaîtrons leur comportement à long terme et leurs interactions avec les transferts d'eau.

Nous pouvons toutefois dès à présent proposer une modélisation simple de l'accroissement des concentrations de nitrate dans les nappes, en additionnant les pertes correspondant aux apports annuels successifs (voir la figure 4). On obtient d'abord une période d'accroissement de la concentration en nitrate, puis cette concentration se stabilise : un nouvel équilibre dynamique, qui inclut les apports par l'homme, s'établit. On comprend mieux ainsi pourquoi la pollution azotée, qui a débuté vers les années 1950, a mis tant de temps à s'installer.

Avons-nous atteint le nouvel équilibre? Nous ne le savons pas. Nous ne savons même pas où nous en sommes de la courbe d'accroissement de la concentration en nitrate : au milieu? en haut? près de l'équilibre dynamique? Cette modélisation démontre toutefois que même si l'on arrêta tout apport d'engrais, hypothèse absurde, le retour à l'état antérieur durerait plusieurs décennies.

On comprend ainsi la déception apportée par des expérimentations de diminution à grande échelle des doses d'engrais. La variabilité inter-annuelle des pertes en azote est très grande : dans le meilleur des cas, il ne pleut pas, et il

n'y a pas d'azote qui sort du sol, parce qu'il n'y a pas d'eau qui le quitte. Ainsi, pendant trois ou quatre années peu pluvieuses, les concentrations en nitrate n'augmentent plus dans des nappes à renouvellement rapide. Les promoteurs de ces expérimentations crient alors victoire, et... les concentrations remontent à la première pluie.

La lenteur de la production du nitrate est accompagnée de la lenteur de son transfert dans le sous-sol au-dessus de la nappe. Ainsi, en Champagne, les profils de concentration du nitrate dans la zone du sous-sol non saturée en eau, au-dessus de la nappe de la craie, montrent que la progression des nitrates vers la nappe est d'environ un mètre en une année et demie. Les concentrations dans la partie supérieure du sous-sol non saturé étant plus élevées que dans la nappe, la concentration dans celle-ci va certainement s'accroître dans les prochaines années.

Les matières organiques, précieux réservoir d'azote

Que peuvent alors faire les agriculteurs? D'abord, ne pas aggraver la situation, en préservant en priorité les matières organiques du sol. Pendant plusieurs décennies, cette préservation était un objectif des agronomes : ils croyaient que la concentration des sols en matières organiques était un paramètre essentiel de la fertilité. Puis cette pratique a été abandonnée : la fertilisation azotée a permis l'accroissement régulier des rendements tandis que les concentrations en matières organiques des sols cultivés chutaient. Ces matières organiques détruites sont parfois la source d'une bonne part des nitrates contenus aujourd'hui dans les nappes!

Ainsi, la teneur en matières organiques du sol de la prairie Nord américaine a diminué de 70 pour cent depuis son retournement en 1888 pour la culture de blé. L'incorporation de matières organiques venues du blé ne compense pas cette diminution. Notons d'ailleurs que le carbone de ces matières organiques détruites a été oxydé en dioxyde de carbone, gaz à effet de serre émis dans l'atmosphère : ce n'est que l'une des nombreuses relations qui couplent les cycles de l'azote et du carbone.

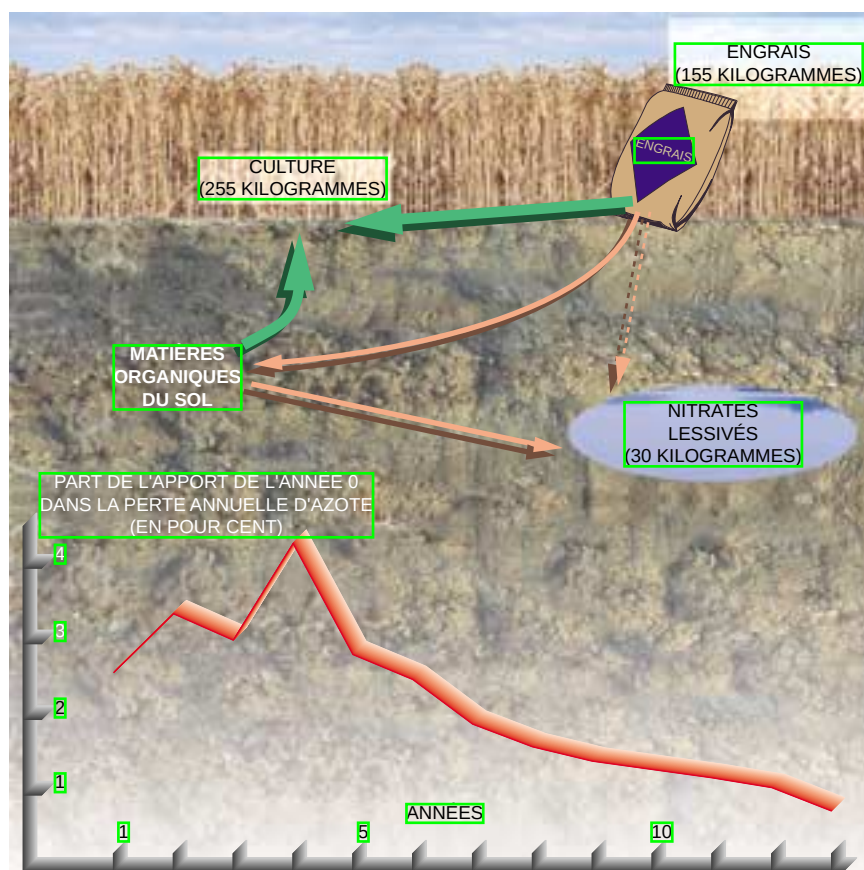
Un autre exemple de dégradation du sol, naturelle cette fois, illustre le lien entre la destruction des matières organiques du sol et la concentration du nitrate dans la nappe sous-jacente.

Ainsi la concentration du nitrate est-elle très élevée dans la nappe aquifère du Grand Erg, au Sahara. Dans les zones d'affleurement de la nappe, le nitrate précipite même en sels, encore utilisés il y a quelques décennies par des tribus touarègues pour la fabrication de poudre. Ces nitrates proviennent en grande partie de la minéralisation, puis de la nitrification, de l'azote contenu dans les matières organiques des sols qui s'étaient formés lors de la période dite pluviale, il y a 6000 ans : la végétation et la faune se sont développées dans le Sahara, au moins localement, puis ont disparu à la suite d'importants changements climatiques et de l'aridification qui a conduit à la situation actuelle.

Au-delà de leur intervention dans le cycle de l'azote, les matières organiques du sol ont d'autres rôles : elles transportent, quand elles sont solubles, des micropolluants (pesticides, métaux lourds) et contrôlent en partie le devenir de ces molécules et de ces éléments dans le milieu naturel. Elles inter-

viennent aussi sur la stabilité des sols et sur leur potentiel de résistance à l'érosion. Enfin, contenant beaucoup de carbone organique, à l'état réduit, elles sont des réducteurs, et réagissent chimiquement avec de nombreux composés dans le sol, notamment avec les nitrates. Ainsi, dans de nombreux milieux anaérobies, où il n'existe pas de dioxygène dissous, vivent des bactéries qui respirent l'oxygène de l'ion nitrate et transforment celui-ci en diazote : le carbone des matières organiques s'oxyde en gaz carbonique, tandis que le nitrate est réduit. Ce processus naturel de dépollution est la dénitrification, qui boucle le cycle externe de l'azote. C'est encore un exemple du couplage entre les cycles de l'azote et du carbone.

Cette dénitrification se produit naturellement dans des eaux souterraines profondes, sous les terres cultivées, dans de très nombreux sites en France, en Angleterre, en Afrique du Sud ou aux États-Unis. Ainsi, dans la nappe de la craie du Nord de la



3. LES NITRATES LESSIVÉS au cours d'une année ne proviennent que pour une faible part des engrais apportés cette année là. Ces derniers sont surtout utilisés par les plantes et réorganisés en matières organiques du sol (en haut, les chiffres correspondent à une année). La part, dans la perte annuelle, de l'azote apporté une année donnée décroît lentement au cours de la douzaine d'années suivantes (courbe en rouge).

France, la concentration du nitrate chute par endroits de plus de 100 milligrammes par litre à zéro, en quelques centaines de mètres, grâce à l'action de ces bactéries.

Dans la plaine de l'Allemagne du Nord, la concentration de nitrate dans les nappes est aussi souvent quasi nulle, en dépit d'une agriculture intensive. C'est, là encore, grâce à l'action de bactéries dénitrifiantes. Ces bactéries ont une action légèrement différente : elles réduisent les nitrates en oxydant des sulfures de fer (où le soufre est réduit), abondants dans le sol, et non pas le carbone réduit des matières organiques. La réaction libère ainsi des sulfates (où le soufre est oxydé). Depuis les années 1960, il n'y a pas d'augmentation de la concentration de nitrate dans les nappes, malgré une agriculture intensive, mais, en contrepartie, on y observe un accroissement régulier des concentrations en sulfate, qui se rapprochent dangereusement de la norme de potabilité pour ce composé.

La dénitrification se produit aussi en surface dans des zones humides, tels les sols engorgés des bordures boisées des grands fleuves. Là, sur quelques dizaines de centimètres de profondeur dans le sol et sur quelques dizaines de mètres latéralement, les eaux chargées de nitrate, lessivés latéralement, peuvent être dénitrifiées par des bactéries, qui oxydent simultanément du carbone en dioxyde de carbone. Cette dénitrification est efficace (trois kilogrammes d'azote par hec-

tare et par jour) : la pollution est ainsi réduite avant que ces eaux de surface ne rejoignent les rivières et les fleuves.

Divers travaux ont montré que l'on peut facilement promouvoir la dénitrification par des bactéries dans les eaux souterraines peu profondes, par exemple dans des nappes alluviales : il suffit d'injecter dans la nappe polluée par les nitrates de petites quantités de carbone organique soluble, de l'alcool par exemple. Dans toutes les expériences, la dénitrification démarre, sous l'action des bactéries naturellement présentes dans le milieu, et devient très efficace. Toutefois, très rapidement, les pores de l'aquifère sont colmatés par les corps bactériens de la biomasse dénitrifiante qui se reproduit trop rapidement : à terme, l'eau circule plus mal, introduisant un dysfonctionnement hydraulique de la nappe.

La pollution passe de l'eau à l'air

Par ailleurs, la dénitrification naturelle a une contrepartie : elle transfère vers l'atmosphère une partie de la pollution du sol et de l'eau. Le nitrate, forme la plus oxydée, est transformé en diazote, forme réduite, par une chaîne de réactions qui produisent de nombreux composés intermédiaires. Or, le rendement de la dernière de ces réactions, qui transforme l'oxyde nitreux en azote, est faible dans de nombreux milieux. La dénitrifica-

tion produit alors de l'oxyde nitreux, gaz stable dans la basse atmosphère (son temps moyen de résidence y est de 150 ans), responsable d'un fort effet de serre (six pour cent de l'effet additionnel depuis 1765) et qui contribue à la destruction de l'ozone dans la stratosphère, qu'il atteint à cause de sa longue durée de vie.

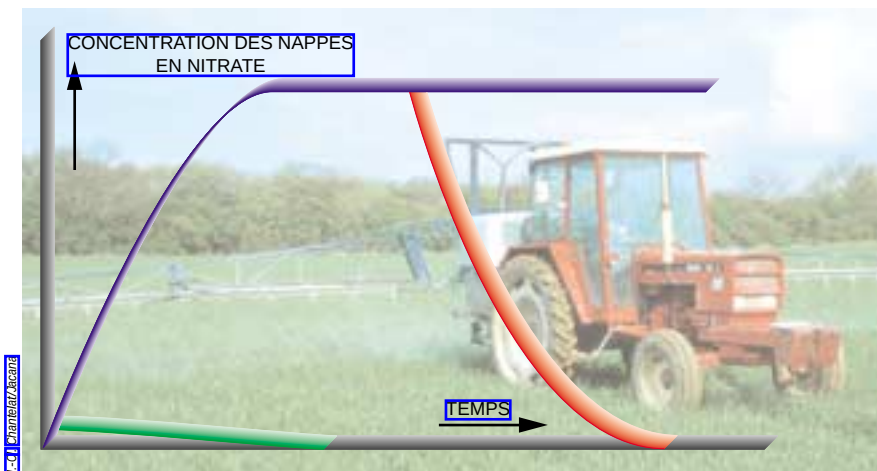
Ainsi, on cultive du colza pour produire du biocarburant, afin d'économiser du combustible fossile et de rejeter moins de dioxyde de carbone dans l'atmosphère. Mais cette culture nécessite des apports d'engrais, et elle est responsable, à certaines périodes de l'année, d'excédents de nitrate dans le sol. Des équipes de l'INRA de Grignon et de Dijon ont montré que la dénitrification naturelle peut alors se produire et émettre de l'oxyde nitreux, dont le fort effet de serre annihilerait le gain associé au dioxyde de carbone qui n'a pas été émis.

Quelles voies de recherche permettront-elles de relever le défi de l'agronomie écologique et économiquement rentable? L'exemple de l'oxyde nitreux montre que nous devons prendre en compte l'ensemble des cycles biogéochimiques de la surface (cycles de l'eau, de l'azote, du carbone), dans le sol, dans la plante et dans l'atmosphère, ainsi que leurs couplages.

L'échelle naturelle d'étude de ces cycles est celle du bassin versant, voire de la région. Nous avons vu en effet que le fonctionnement écologique des agrosystèmes est couplé à celui des écosystèmes naturels qui les jouxtent : la dénitrification n'est que l'un des processus de dépollution qui se produisent dans les zones engorgées d'eau des bords de rivière.

En outre, les durées caractéristiques des cycles sont supérieures à l'année, donc à la période de la production végétale, et souvent en décalage avec celle-ci. Contrairement aux expériences de type agronomique, de courte durée, les expérimentations environnementales requièrent donc de longues durées, dans le cadre d'observatoires de terrain.

La compréhension des écosystèmes naturels guiderait aussi les efforts vers une agronomie à faible niveau d'intrant, notamment d'engrais. Ainsi, avec le laboratoire Fonctionnement et évolution des systèmes écologiques, de l'Université Paris 6 et de l'École normale supérieure, nous avons étudié le



4. LA QUANTITE DE NITRATE qui arrive chaque année dans une nappe aquifère (courbe en bleu) est la somme des pertes dues aux apports annuels successifs d'engrais (dont un élément est montré en vert). Lorsque l'on commence à utiliser des engrais, cette quantité augmente pendant plusieurs années, puis elle se stabilise : les pertes de nitrate compensent les apports. Le retour à l'équilibre antérieur (courbe en rouge), par la suppression des engrais, ne serait obtenu qu'après une longue durée.

cycle de l'azote dans la savane : celui-ci y est presque totalement fermé, sans pertes par lessivage et sans dénitrification notable.

Bien que les sols soient peu fertiles (ils produisent peu d'azote minéral, notamment de nitrate), la productivité atteint presque celle d'un champ de blé. Pourquoi n'y a-t-il pas de nitrate dans le sol ? Les plantes herbacées de la savane sont pérennes. Après leur mort, leur azote organique est minéralisé sous forme d'ammonium. L'ammonium, chargé positivement contrairement au nitrate, est retenu par l'argile et par l'humus ; il n'est pas lessivé, pas dénitrifié. Il est réutilisé, presque sans aucune perte, par la nouvelle génération de plantes. Il semble que les plantes empêchent le sol de nitrifier : elles posséderaient un système inhibiteur de nitrification, tel qu'en ont recherché, sans beaucoup de succès, les agronomes.

Pour remédier aux pollutions associées à l'azote, il reste à inventer de nouvelles pratiques dont nous savons que les résultats ne seront apparents qu'après de nombreuses années. L'interdiction des chlorofluorocarbones, dont physiciens et chimistes de l'atmosphère ne nous prédisent les premiers effets bénéfiques sur la couche d'ozone que dans plusieurs décennies, sera un exemple à suivre.

André MARIOTTI est professeur à l'Université Pierre et Marie Curie, à Paris, où il dirige l'UMR 7618 INRA / CNRS.

A. P. KINZING et R. H. SOCOLOW, *Human Impacts on the Nitrogen Cycle*, in *Physics Today*, novembre 1994.

A. MARIOTTI, *Dénitrification in situ dans les eaux souterraines : processus naturels ou provoqués : une revue*, in *Hydrogéologie*, vol. 3, pp. 43-68, 1994.

D. DENYS, A. MARIOTTI, J.C. MULLER, M. GRABLY, B. MARY et B. NICOLARDOT, *Utilisation conjointe de la lysimétrie et des techniques de marquage isotopique pour étudier le bilan de l'azote des sols cultivés*, in *Comptes-rendus de l'Académie d'agriculture de France*, vol. 81, n°4, pp. 131-162, 1995.

C. HÉNAULT et J. C. GERMON, *Quantification de la dénitrification et des émissions de protoxyde d'azote (N_2O) par les sols*, in *Agronomie*, vol. 15, pp. 321-355, 1995.

Vaclav SMIL, *Engrais et démographie*, in *Pour la Science*, septembre 1997.

Le chant des cigales

HENRY BENNET-CLARK

Les cigales mâles sont les plus bruyants des insectes. On a découvert comment fonctionne leur système acoustique.

Dans de nombreuses régions du globe, la fin du printemps est sonore : chaque année, quelques-unes des milliers d'espèces de cigales sortent du sol, et les mâles commencent à entonner un chant rauque et assourdissant. Parmi ces virtuoses du monde des insectes, les mâles de la cigale australienne *Cyclochila australasiae* semblent être ceux dont le chant est le plus puissant. Cette cigale a un cri aussi strident

qu'une alarme domestique. De plus, un mâle ne chante pas tout seul : il est généralement accompagné de dizaines, voire de centaines, d'autres cigales. Imaginez une pièce où retentiraient simultanément des dizaines d'alarmes !

Comment ces petits animaux (ils sont longs de quelque six centimètres) produisent-ils un tel vacarme ? Bien que la sortie de terre de ces insectes soit parfois un événement médiatique, surtout

aux États-Unis, leur système acoustique était mal connu. En 1954, J. Pringle, spécialiste anglais des insectes, publia un article sur le chant des cigales qui fit date, mais son travail ne fit guère d'émules. Depuis sept ans, avec David Young, de l'Université de Melbourne, nous étudions la production des sons par ces insectes. Nous avons utilisé de minuscules microphones, que nous avons mis au point, afin d'étudier le dispositif acoustique de *Cyclochila australasiae*.



LE MALE DE LA CIGALE AUSTRALIENNE (*Cyclochila australasiae*) a le chant d'appel le plus puissant jamais enregistré.

Un chant assourdissant

Les cigales, apparentées aux pucerons et aux cicadelles, sucent la sève des plantes. Pour des insectes qui se nourrissent ainsi, elles sont grandes et bruyantes : les plus petites mesurent un centimètre de longueur, les plus grandes dix. Les femelles pondent leurs œufs dans des tiges de plantes ou dans des arbres. Après l'éclosion, les larves tombent sur le sol et s'y enfouissent, à la recherche de racines de plantes à percer. Elles restent sous terre pendant toute leur vie larvaire, qui dure parfois plusieurs années ; les cigales américaines vivent entre 13 et 15 ans, et ce cycle de vie est parmi les plus longs, chez les insectes. Quand les larves sortent à l'air libre, elles muent en adultes ailés et vivent quelques semaines. Alors, les mâles chantent à tue-tête pour attirer les femelles (les cigales peuvent également produire des chants de protestation et un paisible chant nuptial).

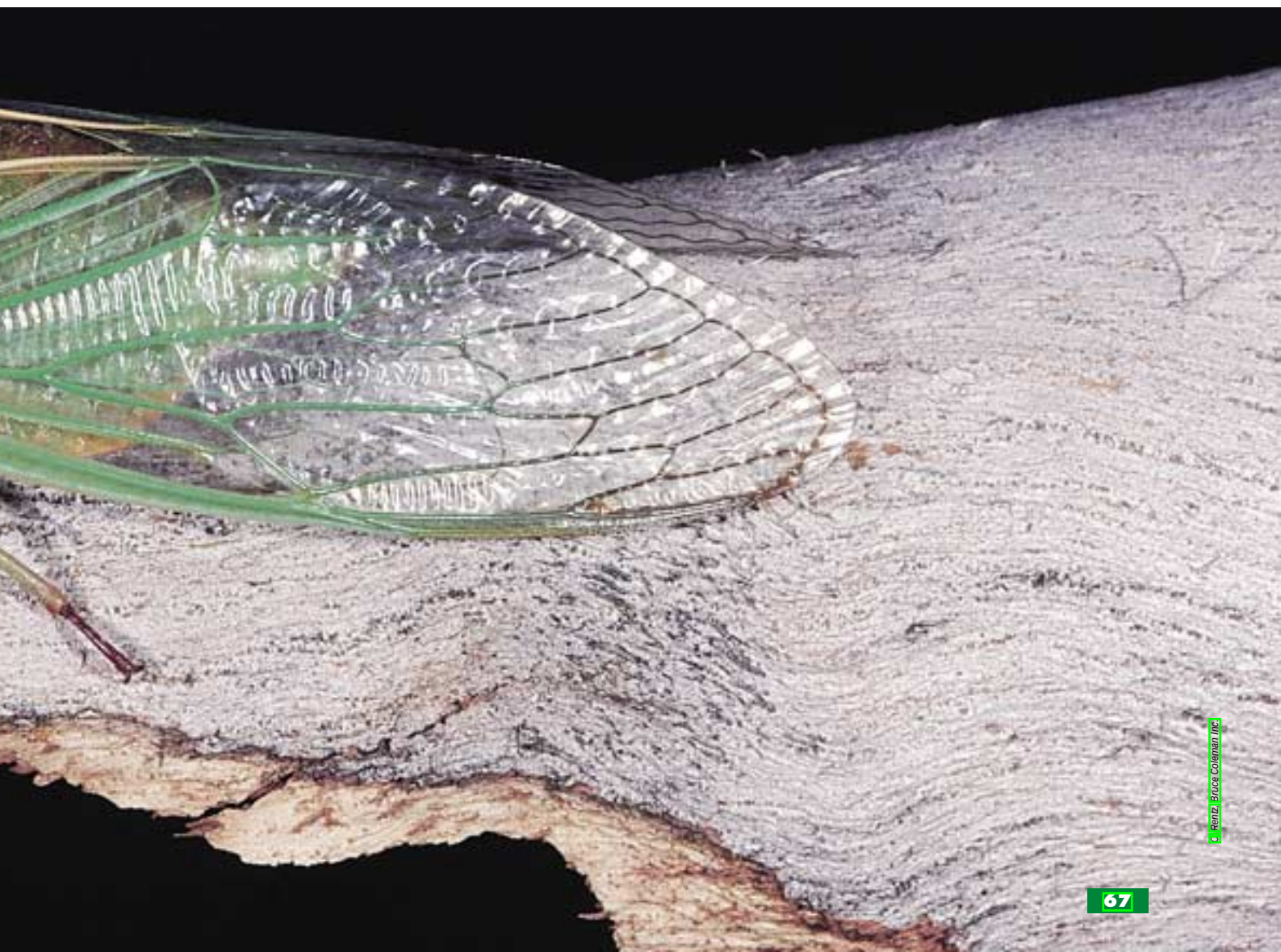
Pourquoi le chant sexuel de ces insectes est-il si bruyant ? Peut-être pour assourdir les prédateurs, qui ne peuvent plus localiser précisément les cigales. Peut-être les grandes cigales ont-elles un habitat très étendu, ce qui les oblige à chanter haut et fort pour attirer leurs partenaires. Par ailleurs, on sait que les femelles entendent bien les sons de plus de 30 décibels. Comme les grillons femelles, les cigales femelles reconnaissent vraisemblablement les mâles par l'intensité et la qualité de leur chant.

Des tambours incorporés

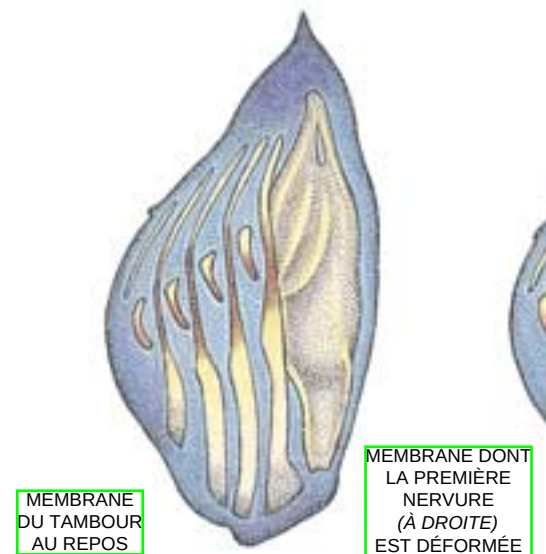
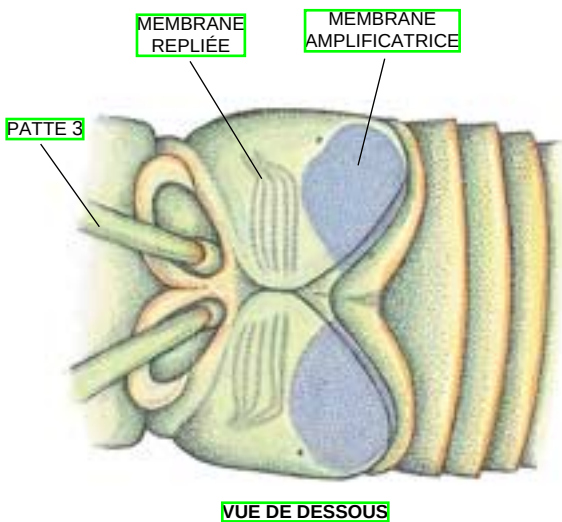
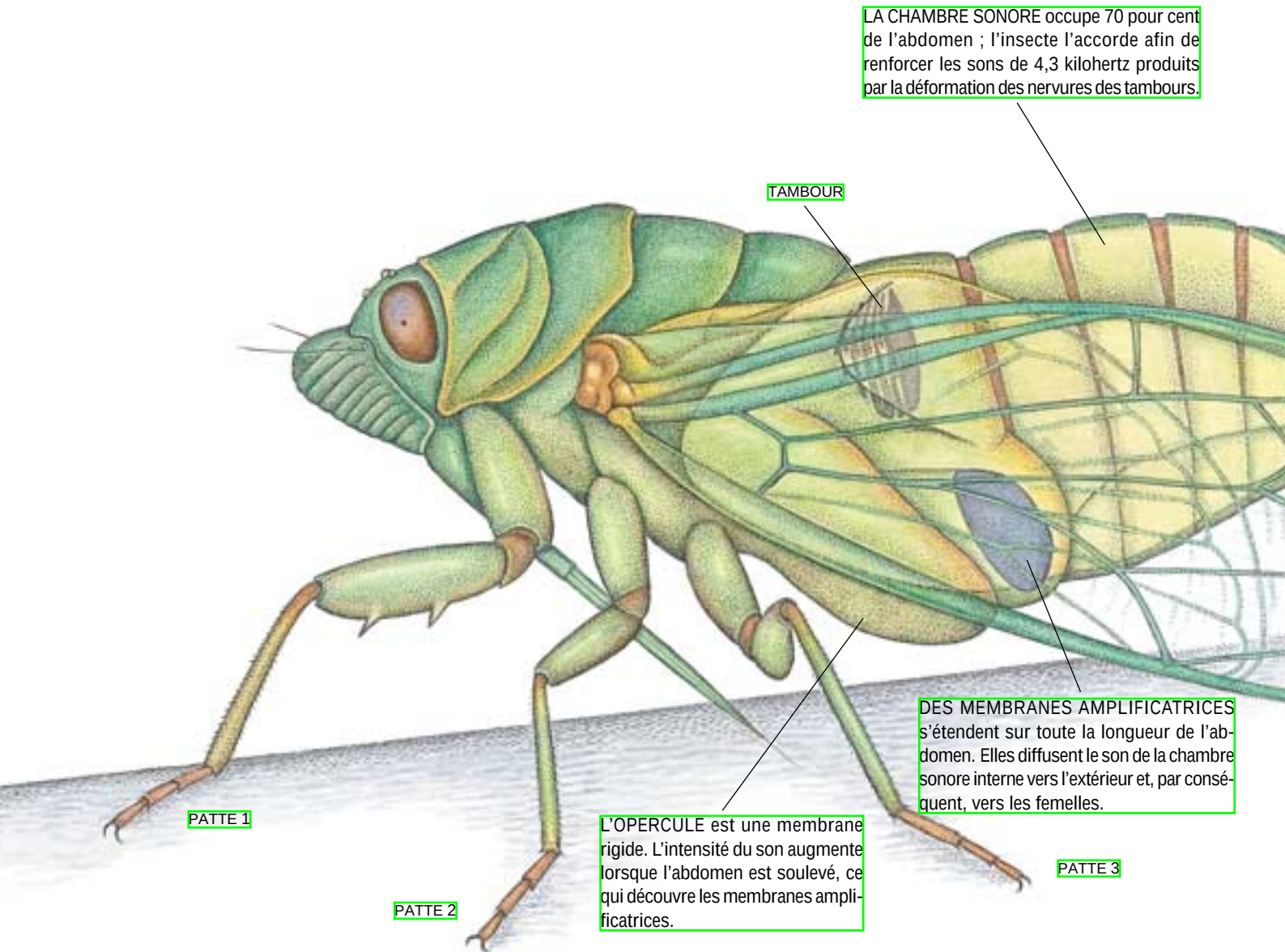
Pour comprendre comment les cigales mâles chantent aussi fort, on doit analyser les différents mécanismes de production de ces sons. Les mâles sont pourvus d'une paire de tambours, situés de chaque côté de l'abdomen. Chacun de ces tambours élastiques est une cavité résonante, renforcée par quatre nervures convexes qui courent

sur les deux faces des tambours. Ces nervures sont reliées les unes aux autres par une membrane, et les deux membranes sont réunies par un muscle puissant.

La contraction de ces muscles déforme les membranes, produisant un son. En fait, la meilleure comparaison serait celle d'une boîte de conserve dont le couvercle et le fond seraient légèrement bombés. Le son émis est comparable à celui que l'on produirait en appuyant sur le fond et sur le couvercle bombés qui, la pression relâchée, reprendraient leur position initiale. Ici le couvercle et le fond sont attirés l'un vers l'autre par le muscle qui les joint ; quand ils se déforment vers l'intérieur, ils produisent deux « clics ». Quand ils reprennent leurs positions initiales, ils émettent deux nouveaux clics. Comme ces muscles se contractent, alternativement, 120 fois par seconde chacun, le son produit par les tambours a une fréquence double, égale à 240 hertz (un hertz est l'unité de fréquence qui cor-

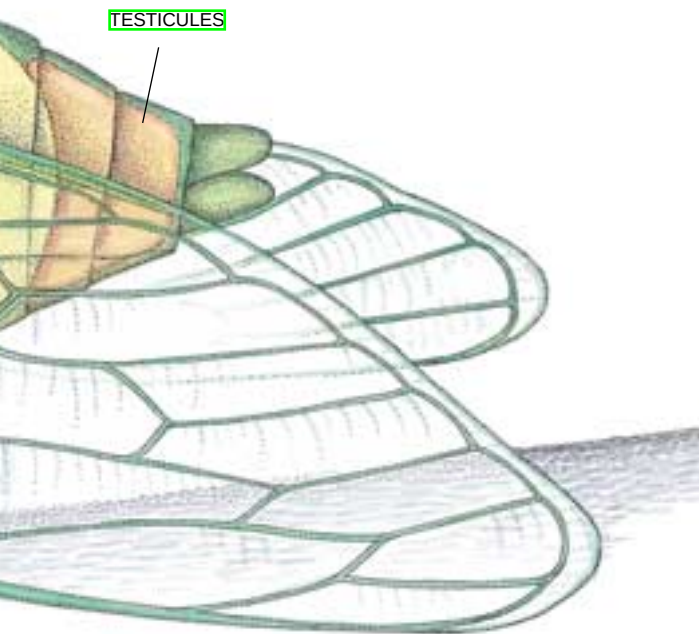


Anatomie du chant de la cigale

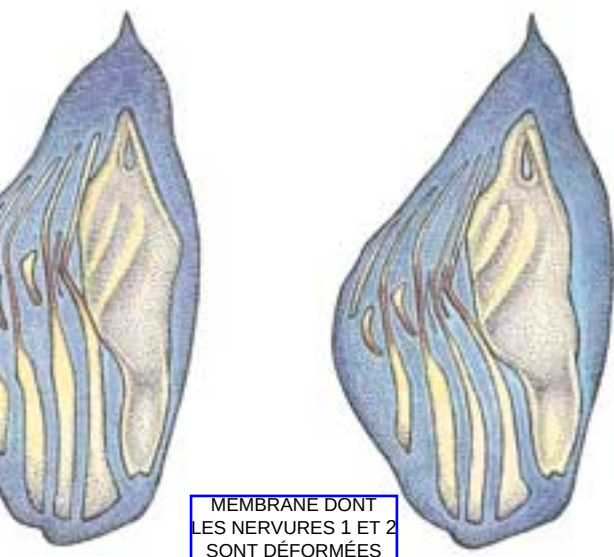


respond à un cycle par seconde). En fait, les deux membranes des tambours ne se déforment pas d'un bloc, mais nervure après nervure.

En se contractant rapidement, les muscles produisent suffisamment d'énergie, durant les trois premières milli-secondes de chaque contraction, pour déformer, les unes après les autres, deux ou trois nervures de chaque «cou-vercle». Ce mouvement déplace, en deux ou trois fois, les deux faces de chaque tambour vers l'intérieur. À chaque fois qu'une nouvelle nervure se déforme, un claquement



LES NERVURES sont déformées les unes après les autres par le puissant muscle de chaque tambour. Chaque déformation crée une pression sonore atteignant 158 décibels. Les nervures des membranes sont entourées d'un tissu élastique (en bleu).



bref et intense retentit. Comme, à chaque contraction des muscles, chaque nervure produit plusieurs clics, que les nervures d'une même membrane sont déformées les unes après les autres et que les nervures des deux membranes sont actionnées séquentiellement, l'ensemble aboutit à 4 300 clics par seconde : l'ensemble des clics produit un train de vibrations dont la fréquence résultante est celle du chant de la cigale, soit 4,3 kilohertz.

Le claquement d'une des nervures produit des pressions sonores élevées, atteignant 158 décibels, dans l'abdomen de la cigale, ce qui équivaut au bruit produit par l'explosion d'une grenade à un mètre de distance (un décibel est un dixième de bel, unité d'intensité sonore relative ; un son est un bel plus haut qu'un autre lorsque le rapport de leurs intensités est égal à dix). Chez la plupart des cigales, l'abdomen contient une grande poche d'air, ou chambre sonore ; chez la cigale australienne, cette poche, dont le volume atteint 1,8 millilitre, occupe 70 pour cent de l'abdomen de l'insecte. L'abdomen contient une paire de membranes amplificatrices qui s'étendent sur toute la largeur de la face ventrale de l'insecte et qui relie la chambre sonore à l'extérieur.

Les impulsions sonores produites par les tambours et par les nervures résonnent dans la chambre sonore abdominale. Les membranes amplificatrices transmettent efficacement le son vers l'extérieur et l'amplifient : le son est environ 20 fois plus puissant que s'il était émis par les tambours seuls. De cette façon, l'insecte est audible de beaucoup plus loin.

Pourquoi ce son de 158 décibels n'endommage-t-il pas le système auditif de la cigale ? Parce qu'il est transmis grâce aux membranes amplificatrices, tandis que la partie sensorielle du système auditif du mâle est localisée dans une capsule séparée et n'est reliée à ces membranes que par un petit canal. Cette séparation semble protéger les mâles de la surdité.

L'orchestration du chant

L'abdomen et les tympanes ne diffusent pas seulement les sons vers l'extérieur : ils garantissent la qualité du chant nuptial. En étirant l'abdomen et en ouvrant les opercules qui recouvrent les membranes amplificatrices, la cigale ajuste son résonateur abdominal à la fréquence de 4,3 kilohertz, celle des vibrations sonores produites par les tambours. Les structures abdominales améliorent à la fois la pureté du chant et son intensité : elles corrigent et équilibrent les tambours.

De nombreuses espèces de cigales semblent pourvues d'un mécanisme similaire de production des sons, et d'autres insectes sonores «chantent» quasiment de la même façon. Toutefois, pour autant que nous le sachions, les mâles des cigales australiennes restent les insectes les plus bruyants du monde... au grand dam des habitants de Melbourne et de Sydney.

Henry BENNET-CLARK est attaché au département de zoologie des invertébrés de l'Université d'Oxford.

H.C. BENNET-CLARK et D. YOUNG, *A Model of the Mechanism of Sound Production in Cicadas*, in *Journal of Experimental Biology*, vol. 173, pp. 123-153, 1992.

D. YOUNG et H.C. BENNET-CLARK, *The Role of the Tymbal in Cicada Sound Production*, in *Journal of Experimental Biology*, vol. 198, pp. 1001-1019, 1995.

Mélange et désordre figé

ETIENNE GUYON • JEAN-PIERRE HULIN

Deux liquides se mélangent quand ils traversent un solide poreux. Inversement, les solides poreux peuvent aussi séparer des mélanges. La physique explore les raisons de ces différences.

Devant la gare de Dijon, la grande place Henry Darcy rappelle la mémoire de l'ingénieur en chef du corps des Ponts qui, entre 1834 et 1842, met en place le réseau d'alimentation en eau de cette ville. Il en fait alors la ville la plus propre de France, avec une consommation en eau de 200 litres par jour rapportés à chacun des 30 000 habitants de l'époque.

Dijon est traversé par le Suzon, qui, lit-on dans un écrit du XVIII^e siècle, «se mouvait dans un lit de roc pourry, entremêlé de sable mouvant». En appendice d'un article sur «les fontaines publiques de la ville de Dijon», Darcy mentionne ce lit poreux du Suzon et il exprime la loi de l'écoulement d'un fluide dans un milieu poreux. Son nom est resté attaché à cette loi (voir la figure 2).

Bien des problèmes que Darcy se posait sont encore ouverts et d'actualité. Ils se présentent sous des formes variées : comment des polluants déchargés accidentellement se mélangent-ils aux eaux souterraines et à quelle vitesse se dispersent-ils ? Comment évaluer la sécurité d'un site de stockage de déchets chimiques ou radioactifs ? Comment récupérer efficacement un polluant (ou du pétrole) dans le sous-sol ? Nous verrons plus loin que d'autres questions, apparemment disjointes, sont pourtant très proches : comment mélanger les constituants visqueux d'un mélange ou d'une pâte ? Comment séparer les éléments d'un mélange chimique ?

La réponse à ces questions impose des études en laboratoire et des tests pratiques des théories : l'écoulement désordonné d'un liquide dans un réseau complexe de canaux, à l'intérieur du sol ou d'une roche poreuse, mélange souvent les espèces entraînées par le liquide. De même, certains

mélangeurs industriels, sans aucune pièce mobile, contiennent un ensemble de cloisons qui imposent au fluide un écoulement analogue à celui qui aurait lieu dans une roche poreuse ; les espèces à mélanger sont divisées en fines lanières qui s'entremêlent. Comment optimiser de tels systèmes ?

Nombre de ces questions simples restent ouvertes. Nous verrons toutefois que certains écoulements que l'on inverse renseignent sur la structure des milieux poreux. D'autre part, la compréhension des phénomènes d'hydrogéologie tire également parti d'expériences de laboratoires sur des milieux stratifiés ou de porosité particulière. On ne pourra vraiment interpréter les écoulements observés dans le sous-sol que lorsque les mécanismes de base auront été analysés finement par de telles expériences.

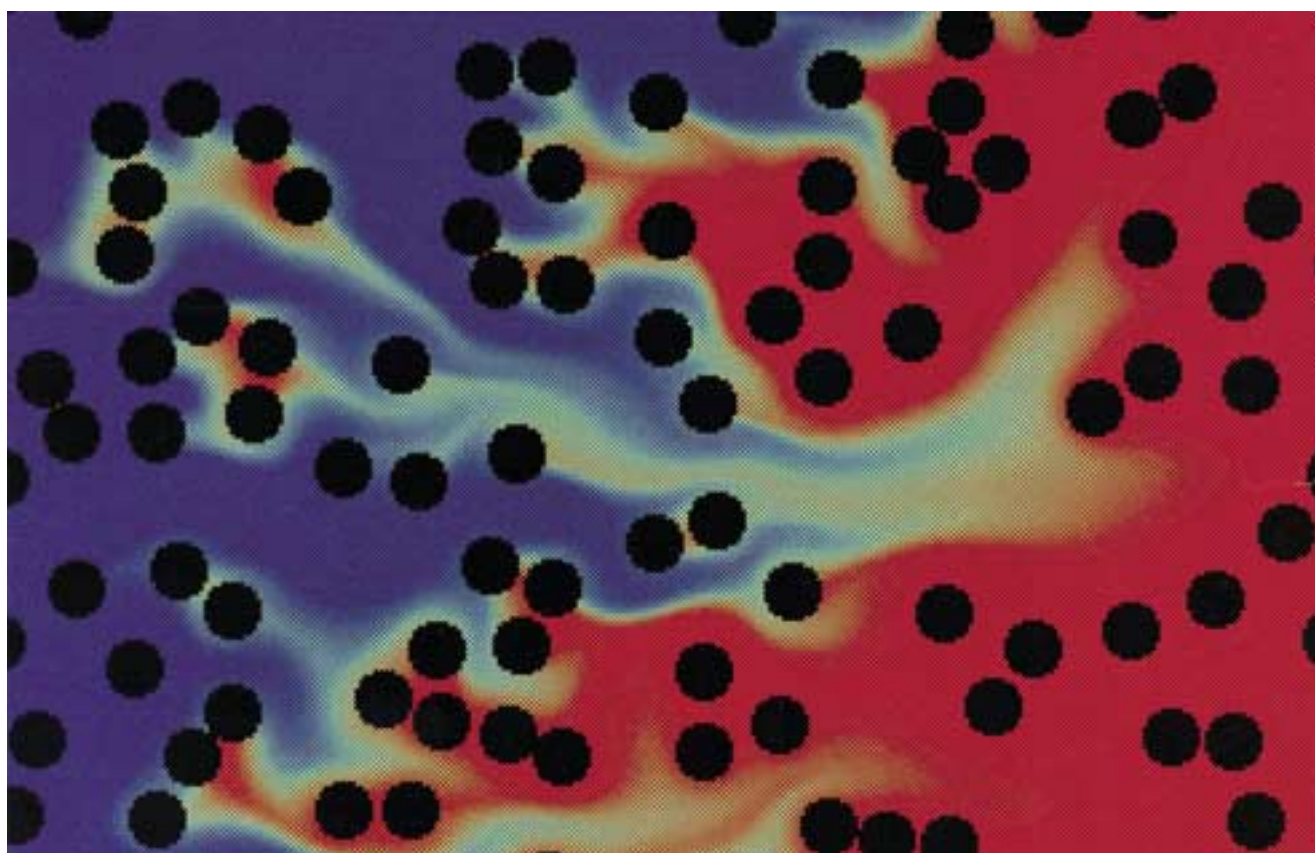
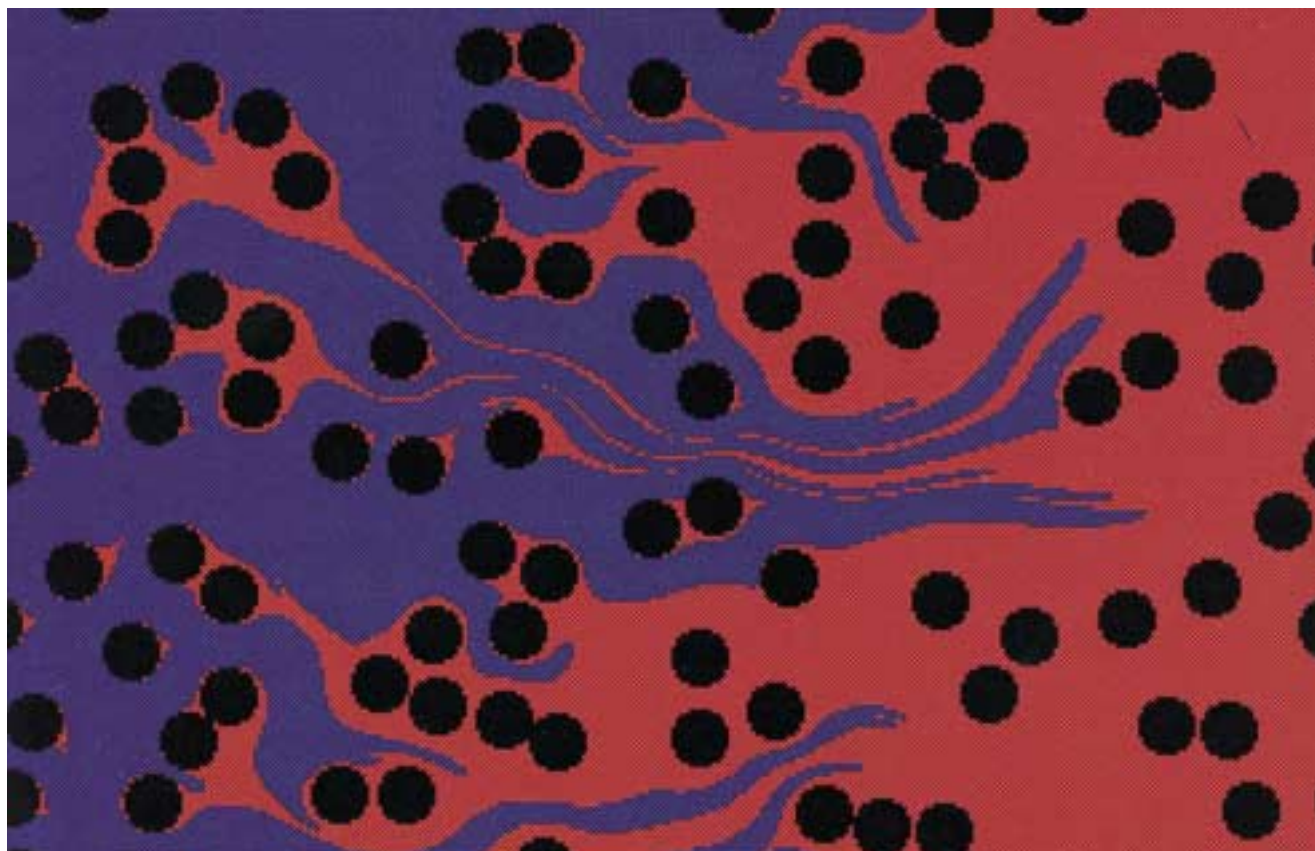
Enfin le mélange n'est pas le seul effet que produit la traversée d'un solide poreux : ce passage peut, au contraire, séparer des espèces et produire une eau pure. C'est ce que réalisent les bons filtres, qui ne sont que des matériaux poreux particuliers : ils arrêtent les particules trop grosses pour passer dans les pores, ou les molécules qui se lient aux parois. Les chimistes séparent ainsi les molécules dissoutes dans les solvants à l'aide de colonnes de chromatographie, où un empilement de grains aux propriétés chimiques particulières retient différemment les divers constituants : au lieu d'être arrêtés, comme dans un filtre, ces constituants sont plus ou moins retardés, selon l'intensité de leurs interactions avec les grains. L'examen des mécanismes de mélange et de séparation montre en quoi ces deux phénomènes sont bien des frères ennemis.

Diffusion et convection

Le mécanisme physique qui mélange le plus intimement deux matériaux est la diffusion moléculaire : dans un bol de café, ajoutons un peu de lait ; même si nous déposons celui-ci précautionneusement, dans un coin du bol, l'agitation naturelle des molécules assure le mélange. Cette diffusion moléculaire agit très lentement, car les trajectoires des molécules sont erratiques : les molécules se propagent en ligne droite jusqu'à ce qu'elles heurtent d'autres molécules ou la paroi du récipient qui les contient ; elles rebondissent et repartent en ligne droite, et ainsi de suite. En raison de cette progression en zigzag, les molécules s'éloignent très lentement du point où on les a placées. La distance atteinte augmente seulement comme la racine carrée du temps écoulé : une molécule met 100 fois plus de temps pour aller seulement dix fois plus loin ! Et une molécule d'eau, dans la tasse à café, ne franchit le rayon de la tasse qu'en plusieurs centaines d'heures, quand elle n'est mue que par l'agitation moléculaire.

Pourtant, ce mécanisme est essentiel, à la fin de tout mélange, car c'est le seul qui reste efficace jusqu'aux distances moléculaires. Quand on a mélangé macroscopiquement deux liquides au point de former des zones de petite taille composées des deux liquides, la diffusion parachève le mélange.

C'est ainsi que, pour obtenir rapidement un café au lait homogène, on agite le contenu du bol à l'aide d'une cuiller. Le mouvement désordonné du liquide, nommé convection, étire chaque goutte de lait, réduit ses dimensions transverses et favorise ensuite le mélange par diffusion moléculaire avec le café environnant. La séquence d'opérations ressemble beaucoup à celle qu'ef-



N. Martys / Institute of Statistics and Science, EPFL, Lausanne, Switzerland

1. DANS UN MILIEU POREUX, l'écoulement d'un liquide est fortement modifié par la diffusion moléculaire. N. Martys, du Bureau américain de normalisation, a simulé un écoulement d'un liquide bleu (circulant de gauche à droite) dans un liquide rouge qui emplissait initialement un solide poreux. Les plots noirs représentent les parties solides du

matériau poreux. Leur répartition aléatoire divise le liquide bleu en filets désordonnés. En l'absence de diffusion moléculaire (*en haut*), le liquide bleu se divise en lanières qui ne se mélangent pas parfaitement au liquide rouge. En présence de diffusion moléculaire (*en bas*), les divers filets se mélangent transversalement.

fectuent les boulangers quand ils pétrissent le pain : ils étirent et replient la pâte à plusieurs reprises, afin de disperser les levures responsables de la fermentation. Dans un liquide, le mélange est partout efficace si ces séquences s'opèrent de façon aussi désordonnée que possible : si l'on tournait lentement, toujours dans le même sens, sans induire d'échange entre le fond et la surface du bol, on risquerait de conserver de grandes zones non mélangées.

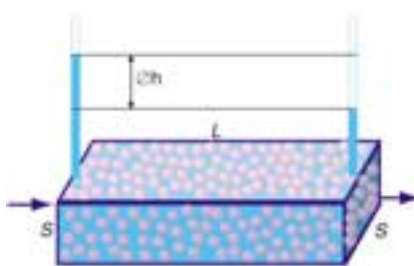
Étirement et repliement, mouvements désordonnés turbulents : voilà les ingrédients qui permettront un mélange intime final, par diffusion. Ces opérations sont efficaces quand le mélange a été optimisé selon la nature et les propriétés des constituants à mélanger : c'est le travail de l'ingénieur de génie des procédés ou du responsable de l'environnement dans l'industrie.

Du café aux poreux : le mélange figé

Toutefois, dans de nombreuses situations expérimentales, tels les mélanges de pâtes ou les écoulements lents dans les sols poreux, la turbulence est absente. Dans un matériau poreux, par exemple, la vitesse d'un liquide en chaque point ne varie pas, au cours du temps, car les frottements visqueux du liquide contre le solide régularisent les écoulements, amortissant quasi instantanément les fluctuations de vitesse qui engendrent la turbulence dans les écoulements libres.

Le mélange qui s'opère quand on injecte dans un tel milieu deux liquides ensemble, ou l'un après l'autre, est «figé». La répartition des vitesses n'évolue pas au cours du temps, mais cette vitesse du liquide diffère selon les points : les éléments de fluide subissent ainsi des déformations analogues à celles qui résultent des opérations d'étirement et de repliement dans les écoulements turbulents.

Caractériser et comprendre les mélanges dans les milieux poreux est nécessaire pour nombre d'applications. Par exemple, les hydrogéologues chargés du contrôle de la qualité de l'eau qui circule dans les nappes phréatiques doivent évaluer si des déchets chimiques ou radioactifs déposés au voisinage des nappes sont entraînés vers celles-ci. Ainsi, il y a une dizaine d'années, à la suite d'un accident de la route, un camion citerne avait perdu plusieurs mètres cubes de tétrachlorure de carbone dans la plaine d'Alsace. Quelques



2. HENRY DARCY, ingénieur en chef du corps des Ponts, alimenta en eau la ville de Dijon, au début du XIX^e siècle. À cette fin, il utilisa un cours d'eau qui traversait la ville dans un massif rocheux poreux, et il énonça la loi qui porte aujourd'hui son nom : le débit d'écoulement d'un liquide dans un solide poreux est proportionnel à la différence de pression entre l'entrée et la sortie du milieu, à la section S, et inversement proportionnel à la longueur L de matériau traversé et à la viscosité du liquide ; le coefficient de proportionnalité est la perméabilité du milieu. Elle se mesure... en darcys.

années plus tard, ce composé fut détecté dans des puits profonds d'une centaine de mètres, à plusieurs kilomètres en aval. Dans une polémique juridique qui dure encore, il fut argué que la pollution provenait non de cet accident, mais d'une usine située à proximité. La résolution du conflit nécessite que l'on modélise les conditions de transport des déchets, leur vitesse de migration et leur trajectoire en fonction de la nature du sol.

Une expérience classique de laboratoire reproduit de tels phénomènes dans de petits échantillons : on fait d'abord couler à vitesse moyenne constante V un liquide dans un cylindre poreux, tel un échantillon de roche ou de sol ; puis, pendant un court instant, on ajoute à l'entrée du solide une substance qui ne modifie pas l'écoulement et n'interagit pas avec le solide (colorant, composé fluorescent, traceur radioactif, etc.). On choisit cette substance afin d'en détecter la sortie à l'extrémité de l'échantillon poreux et de mesurer les variations de sa concentration au cours du temps.

Le traceur parcourt l'échantillon en un temps moyen de l'ordre de L/V , où L est la longueur du solide poreux. Toutefois, même si l'injection du traceur est brève, les molécules ne séjournent pas toutes pendant le même temps dans l'échantillon : à la sortie d'un matériau homogène, la courbe qui représente la concentration du traceur en fonction du temps a approximativement la forme d'une courbe en cloche, nommée courbe gaussienne. Cet étalement associé à un écoulement est nommé dispersion.

Pourquoi l'étudier alors qu'on s'intéresse aux mélanges ? Parce que le colorant qui se disperse dans le fluide environnant se mélange à lui : le volume où s'effectue ce mélange grossit au cours du temps, en raison du phénomène de dispersion.

Dans cette première expérience, nous avons supposé que la concentration en traceur était uniforme dans une section de l'échantillon poreux : nous n'avons examiné l'étalement que dans la direction de l'écoulement moyen, ce que l'on nomme la dispersion longitudinale. Si nous déposons maintenant une petite goutte de traceur au milieu de l'écoulement, nous voyons celle-ci s'étaler aussi bien dans la direction perpendiculaire à l'écoulement que parallèlement à celui-ci, mais de façon différente ; pour un fluide qui s'écoule dans un solide poreux, la

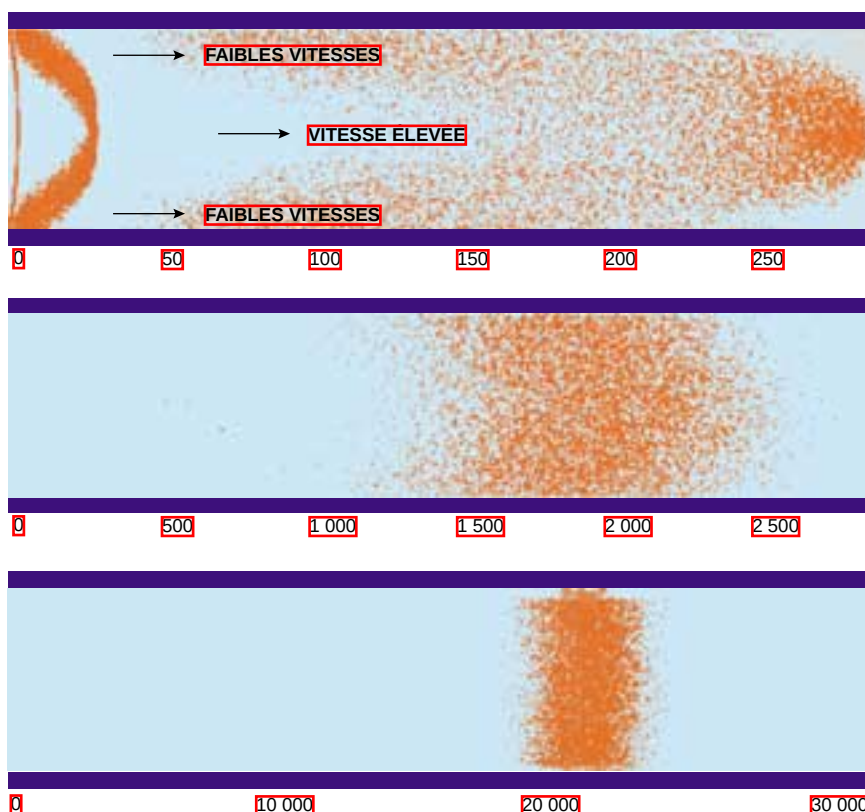
dispersion longitudinale diffère de la dispersion transverse, tandis que, dans les fluides simples, la diffusion moléculaire est identique dans toutes les directions (elle est «isotrope»). Dans le premier cas, une goutte sphérique de traceur s'allonge dans la direction de l'écoulement, tandis que, dans le second cas, elle reste sphérique.

La dispersion au microscope

Pourquoi cette différence ? Afin d'en identifier la cause, utilisons un milieu transparent et un microscope qui permette une observation à l'échelle des pores. En pratique, nous employons un solide poreux composé de grains de verre entassés (du verre pilé, par exemple), où nous faisons passer un liquide transparent de même indice de réfraction que le verre. Le milieu, alors transparent, permet de suivre une tache de colorant. Un peu comme un bouchon à la surface d'un torrent hérissé de nombreux obstacles rocheux, une molécule du traceur entraînée par le liquide accélère, ralentit, est déviée en tous sens...

La figure 1 montre une simulation numérique à deux dimensions de cette expérience : les grains solides du matériau poreux sont représentés sous la forme de plots circulaires. Un fluide coloré est déplacé par un autre fluide d'une couleur différente : le front initialement plan prend une forme de plus en plus complexe, à mesure que les fluides traversent les pores et contournent les grains. On voit notamment le fluide envahissant se diviser en filets étroits, qui sont souvent vite étalés latéralement par l'effet de la diffusion moléculaire, si l'écoulement n'est pas trop rapide.

Si nous nous déplaçons à la vitesse moyenne imposée au liquide, nous verrions que tout se passe comme si chaque particule de fluide effectuait une marche au hasard dans l'échantillon : elle va parfois vers l'avant (plus vite que la vitesse moyenne), vers l'arrière (moins vite que le reste du liquide), tout à fait comme des molécules ayant un mouvement d'agitation thermique. Cette marche au hasard qui se superpose à un déplacement d'ensemble explique la différence des temps de traversée du milieu (on a le même effet pour un ensemble d'ivrognes qui s'éloigneraient d'un café, en effectuant chacun une série de pas de directions imprévisibles et indépendantes).



3. UNE LIGNE VERTICALE DE COLORANT injectée dans un tube fin se disperse par un mécanisme hydrodynamique identifié par le physicien britannique Geoffrey Ingram Taylor. L'image supérieure montre trois états de la ligne, déformée par l'écoulement du liquide, à des temps successifs 1, 10 et 100 (mesurés en nombre de pas de la simulation numérique). Les autres images, à des échelles différentes, montrent l'étalement à des temps respectivement 1 000 et 10 000. La différence d'échelle horizontale entre les trois figures donne l'impression que l'étalement diminue, mais la goutte continue de s'étaler (comme la racine carrée de la distance parcourue).

Si l'on injecte initialement une goutte de colorant dans un solide poreux et si on laisse le liquide circuler suffisamment, la distribution des distances parcourues par le traceur prend une forme de courbe gaussienne, comme si la goutte s'étalait par diffusion moléculaire dans un volume de fluide immobile. Cependant cette répartition gaussienne n'apparaît qu'après un nombre suffisant de changements de direction, c'est-à-dire après un nombre suffisant de pas dans la marche au hasard. Au tout début du processus, quand la distance de déplacement est inférieure à la taille d'un pore, la répartition des distances résulte simplement de la répartition des vitesses dans l'écoulement.

L'existence de ces deux régimes est illustrée par une expérience récente de Jacques Leblond et de Luc Lebon, à l'École supérieure de physique et de chimie de Paris : à l'aide d'une technique de résonance magnétique nucléaire, ils suivent le déplacement des molécules d'un liquide sur une distance comprise entre une faible

fraction de la taille de pore et quelques tailles des pores. Quand le temps de mesure est plus court que le temps moyen de parcours d'un pore par le liquide, la mesure révèle la distribution des vitesses dans les canaux. Lorsque le trajet moyen devient de l'ordre d'une dizaine de diamètres de pore, la répartition des vitesses est presque gaussienne, et centrée sur la distance parcourue à la vitesse moyenne dans l'échantillon. Enfin, aux déplacements intermédiaires (deux pores environ), la courbe que représente le nombre de molécules en fonction de la vitesse présente deux bosses : les particules les plus rapides ont parcouru plusieurs diamètres de pores et ont eu des vitesses très différentes, mais les plus lentes sont restées piégées à l'intérieur du canal où elles étaient initialement.

Ces expériences expliquent ainsi l'existence des deux dispersions longitudinale et transverse. La différence quantitative entre ces deux dispersions, qui résulte de l'anisotropie du problème, reste toutefois un problème ouvert.

Des capillaires de Poiseuille à ceux de Taylor

Une autre forme de dispersion dans les solides poreux a une origine hydrodynamique. Les travaux du médecin Jean-Louis Poiseuille, au XIX^e siècle, ont montré que, dans un tube de diamètre suffisamment petit pour que la turbulence soit absente, un vaisseau sanguin de faible diamètre, par exemple, la vitesse est nulle aux bords du tube et maximale au centre.

Ce phénomène provoque un étalement longitudinal d'un segment de colorant injecté dans l'écoulement et initialement localisé suivant un diamètre du tube : en l'absence de diffusion moléculaire, la ligne de colorant dessine une parabole qui s'étire au cours du temps.

L'effet d'étiement longitudinal par les différences de vitesse est cependant contrebalancé en partie par la diffusion moléculaire dans la direction perpendiculaire à l'écoulement, qui tend à homogénéiser la distribution du colorant dans la section du tube. La figure 3 montre ainsi le résultat du jeu combiné de ces deux effets. Aux temps courts (*figure supérieure*), le profil parabolique s'étale de plus en plus transversalement sous l'effet de la diffusion. Il disparaît aux temps assez longs, et la concentration en colorant le long du tube devient une gaussienne. De plus, la largeur de cette gaussienne augmente proportionnellement à la racine

carrée du temps, comme pour les autres mécanismes que nous avons examinés. D'une image à l'autre, les distributions paraissent de plus en plus étroites, car la largeur augmente plus lentement que le temps de trajet.

Où est la marche au hasard, dans ce problème ? La durée d'un pas de marche au hasard est ici le temps que met une molécule de fluide (ou de traceur) pour « oublier » sa vitesse à un instant donné. Dans ce cas, c'est le temps qu'il faut pour que, par diffusion transverse à l'écoulement, les molécules aillent du centre de l'écoulement jusqu'au bord ou inversement.

À l'Université de Cambridge, le grand mécanicien des fluides Geoffrey Ingram Taylor, qui avait aussi réalisé les premières études du mélange dans des tubes capillaires, a montré que le coefficient de dispersion hydrodynamique qui caractérise cet étalement est inversement proportionnel à la diffusion moléculaire : la dispersion est d'autant plus forte que la diffusion est faible.

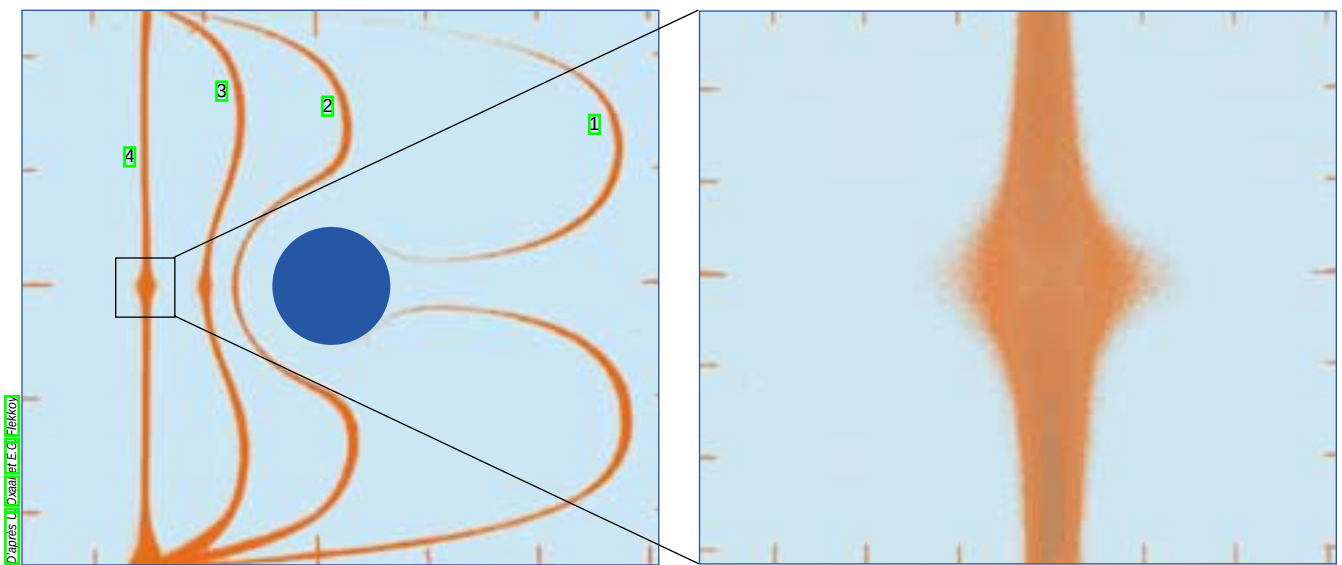
Dans les cas précédemment examinés, la diffusion, prenant le relais de la convection aux petites distances, favorise l'étalement. Dans le cas présent, si la diffusion moléculaire est très efficace, elle fera sans cesse passer les molécules de colorant de la proximité des parois vers la partie centrale du tube : l'effet sur l'étalement des vitesses créé par l'effet Poiseuille sera très limité. En revanche, si la diffusion moléculaire est faible, les particules de traceur

resteront longtemps à une même distance de l'axe. L'étiement produit par le profil parabolique de vitesse, et la dispersion, seront importants.

Cette dispersion de Taylor intervient dans de nombreux phénomènes naturels, telle l'homogénéisation du sang dans les artères. Elle intervient beaucoup moins dans les milieux poreux habituels, à trois dimensions, tels des empilements de grains : si l'on trouve encore des profils proches de celui de Taylor à l'intérieur des canaux d'écoulement, les trajectoires des particules fluides sont plus complexes quand on passe d'un canal à l'autre, et ne suivent plus nécessairement les parois. Les molécules de fluides qui arrivent à un embranchement peuvent choisir l'une ou l'autre des branches ; au cours de ce parcours, une molécule qui se déplace rapidement, parce qu'elle est au centre d'un écoulement, dans un pore, peut se retrouver contre la paroi, dans le pore suivant, où elle se déplace alors lentement (*voir la figure 5*). L'effet du profil Poiseuille est limité à la longueur d'un canal et s'en trouve donc très diminuée.

Dans les roches

Les mécanismes que nous avons considérés sont surtout valables pour de petits échantillons relativement homogènes. En analysant la dispersion, on obtient alors des informations sur la structure à petite échelle des milieux poreux.



4. UNE LIGNE DE COLORANT injectée dans un écoulement (de gauche à droite) a été déformée par un cylindre (en bleu foncé) d'axe perpendiculaire au plan de la figure (1). Quand on inverse le sens de l'écoulement, après quelque temps, la ligne (en rouge) reprend progressivement sa forme initiale (positions successives

2, 3, 4), à l'exception de la partie centrale, qui a été étirée davantage et sur laquelle l'influence de la diffusion moléculaire a été supérieure : les molécules de colorant s'étant mélangées aux molécules du fluide sur une distance supérieure, le mélange a été notable.

Toutefois l'observation d'une coupe de terrain, sur le bord d'une autoroute, montre que le sol est rarement homogène : il est souvent constitué de bandes qui résultent d'accumulations successives de sédiments. Les propriétés varient fortement d'une zone à l'autre, et des liquides s'écoulent dans les diverses couches à des vitesses différentes. La dispersion diffère alors beaucoup de celle dans un milieu poreux homogène.

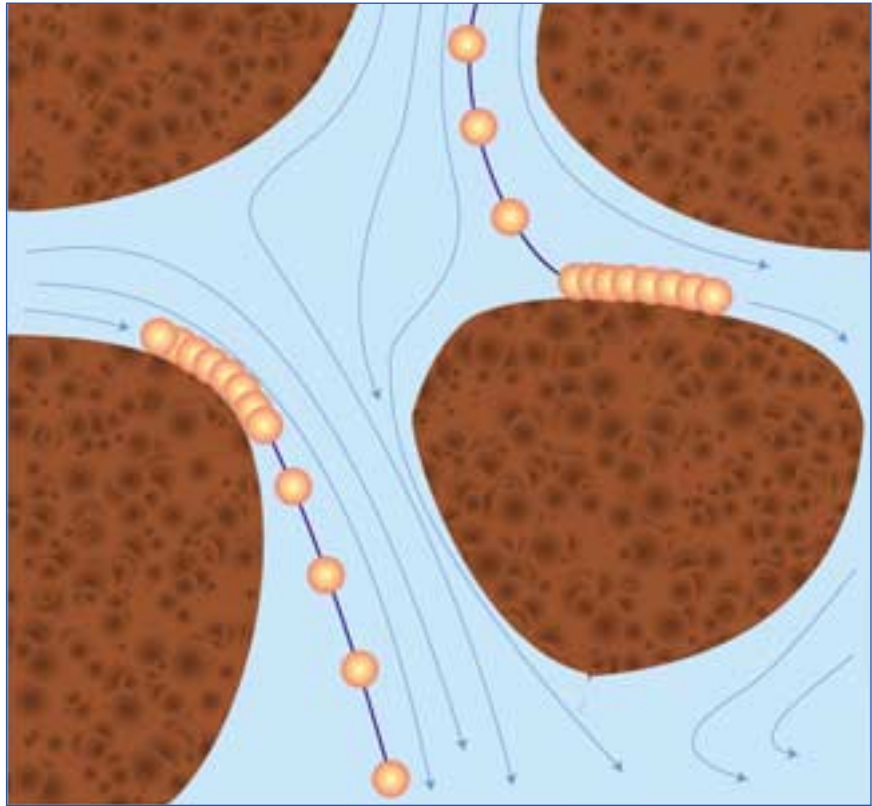
Notamment la loi de dispersion gaussienne s'applique rarement. Parfois un polluant se propage beaucoup plus vite qu'on ne le prévoyait. C'est le cas quand il existe des canaux de forte perméabilité, où l'écoulement est beaucoup plus rapide que l'écoulement moyen, comme dans les massifs calcaires où des dissolutions créent des zones où l'écoulement est direct. Parfois, au contraire, un polluant reste piégé très longtemps dans des poches contenant du liquide au repos, et il n'est relargué que bien après le temps moyen de transit. On parle dans ces deux cas de dispersion anormale.

Les descriptions précédentes ne s'appliquent donc plus que par morceaux dans ces situations pratiques de terrain, et les résultats globaux diffèrent généralement du modèle de dispersion gaussienne classique. Notamment, pour de longs trajets dans ce type de milieu, les coefficients de dispersion calculés (le rapport du carré de l'étalement au temps d'écoulement) augmentent avec la distance.

On comprend ces effets à partir du cas très simple d'un écoulement parallèle à des nappes poreuses, telles que la perméabilité varie de tranche en tranche. L'écoulement est le plus rapide là où la perméabilité est supérieure. Si l'on injecte ainsi un traceur dans un milieu formé de deux strates parallèles en contact, où l'eau s'écoule à des vitesses différentes V_1 et V_2 , la différence des distances parcourues après un temps donné (elle représente l'étalement du traceur) augmente proportionnellement au temps et non plus à la racine du temps comme dans un poreux homogène. Un tel processus résulte d'une convection à différentes vitesses d'une couche à l'autre.

Mélanges et instabilités

Il n'y a d'ailleurs pas besoin d'avoir des couches de terrain différentes pour créer des écoulements variables par



5. DES MOLECULES qui se déplacent rapidement, au centre de l'écoulement dans un pore (*en haut*), parce qu'elles se trouvent au centre de l'écoulement, peuvent être ralenties, dans le pore suivant, au contact d'un grain. Inversement, des molécules qui se déplaçaient lentement (*en bas à gauche*), peuvent accélérer. L'effet de dispersion de Taylor n'a donc pas le temps de s'établir.

tranche. Imaginons le cas d'un milieu poreux rempli d'un liquide A que l'on pousse en injectant à la suite un second liquide B miscible au premier. Les deux liquides se mélangent localement, et la frontière nette entre les deux liquides est remplacée par une zone intermédiaire dont l'épaisseur croît lentement au cours de l'écoulement. Cependant, si le liquide A est plus visqueux que le liquide B et si l'injection est rapide, alors le liquide B pénètre beaucoup plus rapidement sous la forme de doigts dans le milieu A qui le précède : de tels effets de digitation sont classiques en mécanique des fluides et résultent de l'instabilité du front entre les deux liquides. Une fois un doigt formé, la dispersion transverse au doigt favorise naturellement le mélange, comme dans la dispersion de Taylor. L'interpénétration pourra être inégale : lorsqu'un doigt prend de l'avance, il limite la croissance des autres.

Une autre instabilité, favorisant également le mélange, intervient quand deux couches liquides superposées s'écoulent horizontalement dans un matériau poreux. Normalement le front

de séparation s'élargit uniquement en raison de la dispersion transverse à l'écoulement, mais, si le liquide supérieur est plus dense que le liquide inférieur, la gravité tend à faire redescendre le premier liquide en doigts et à faire monter le second. L'effet de la dispersion transverse est notablement amplifié par un tel échange.

Notons cependant que de telles instabilités produisent une interpénétration des fluides sous la forme de doigts d'assez grandes dimensions, mais ne permettent pas, par leur seule action, d'obtenir un mélange efficace à toutes les échelles. L'interpénétration pourra être inégale : lorsqu'un doigt prend de l'avance, il limite la croissance des autres. D'autres mécanismes doivent intervenir pour fractionner encore le mélange jusqu'aux petites échelles, où la diffusion moléculaire devient efficace.

Renverser le mélange

Comment étudier les milieux poreux réels, alors que leur structure est inconnue. Une technique récemment imaginée donne des indications précieuses. Les écoulements lents de fluides vis-

queux sont généralement réversibles : si l'on filme un écoulement et qu'on passe ensuite le film à l'envers, on voit un autre écoulement compatible avec les lois de la mécanique des fluides (à condition de négliger la diffusion moléculaire qui, elle, provoque toujours un mélange irréversible ; toutefois celle-ci agit lentement, de sorte qu'elle ne détruit pas toujours la réversibilité).

C'est le cas dans une expérience modèle à deux dimensions, réalisée par Eric Flekkoy et ses collègues de l'Université d'Oslo : on déforme une ligne de colorant en l'entraînant par un écoulement vers un plot cylindrique. Une fois que la ligne de colorant s'est déformée, on inverse le sens de l'écoulement, et l'on retrouve des positions identiques à celles de l'aller (voir la figure 4). Finalement on obtient une ligne presque droite et presque aussi fine qu'au départ, sauf dans la partie centrale : c'est là que l'étirement de la ligne a été maximal. Lorsque la ligne retrouve sa forme et sa longueur initiales, l'étalement transversal dû à la diffusion moléculaire et acquis pendant la phase d'étirement est fortement amplifié.

Dans un milieu poreux réel, la présence de nombreux grains multiplie les étirements et les divisions du front d'interpénétration des fluides. L'effet de mélange de la diffusion moléculaire s'en trouve amplifié. Qu'en est-il alors du front de séparation entre deux fluides injectés successivement dans un milieu poreux : ce front de séparation deviendra-t-il de plus en plus étroit si l'on inverse le sens de l'écoulement ?

Injectons, par exemple, du colorant dans un écoulement et inversons le sens de ce dernier après un temps T . Comment le colorant sera-t-il réparti après le temps $2T$? Si les particules suivaient exactement les mêmes lignes de fluide, mais dans le sens inverse, on obtiendrait un «écho de dispersion» très peu étalé, où on retrouverait concentrées toutes les particules de traceur. C'est effectivement ce que l'on trouve si le temps T est court (quand le liquide parcourt en moyenne deux ou trois canaux). Toutefois, si le temps T est long par rapport au temps de parcours d'un pore, les phénomènes entre T et $2T$ sont indépendants des phénomènes entre 0 et T : le front de dispersion se déforme comme si l'on avait laissé couler le liquide continûment dans la même direction pendant le temps $2T$. Notre collègue Patrick Rigord a fait l'expérience et a observé que cette perte de mémoire d'une ligne de courant entre l'aller et le retour est presque complète dès que le liquide a parcouru une distance de l'ordre d'une dizaine de diamètres de grains, dans un solide poreux homogène.

On se représente mieux cet effet en pensant à ce qui se passerait dans un labyrinthe : à chaque bifurcation, une petite déviation latérale aléatoire peut nous emmener dans un couloir ou un autre, et séparer irrémédiablement des trajectoires en faisant perdre la mémoire du chemin déjà parcouru.

Il existe cependant des cas où les liquides gardent la mémoire plus

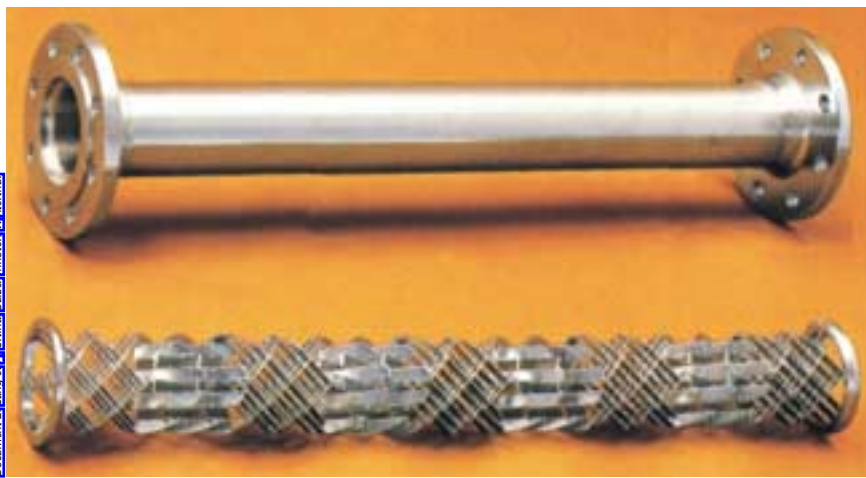
longtemps et où la réversibilité apparaît : ainsi pour l'écoulement parallèle dans des couches de perméabilités différentes, l'étalement du front de séparation entre les liquides est en grande partie réversible. En effet, la majeure partie de l'étalement vient de la différence de vitesse d'une couche à une autre : quand on inverse le sens de l'écoulement, les particules situées dans une couche d'écoulement rapide (celles qui sont allées le plus loin) reviennent vers leur point de départ à vitesse supérieure. Elles l'atteindront en même temps que les particules d'une couche «lente», qui parcourent moins de chemin, mais se déplacent plus lentement. La seule condition à cette réversibilité est qu'il n'y ait pas trop d'échange de particules de fluide entre couches par diffusion transverse : autrement dit, les particules doivent garder la mémoire de la couche de terrain où elles étaient initialement.

De telles expériences, d'«écho de dispersion», permettent de distinguer les différents types d'hétérogénéités : celles de grande taille, qui donnent des effets réversibles, et celles de petite taille et distribuées aléatoirement, qui donnent la même contribution à la dispersion quand on inverse le sens de l'écoulement (ce ne sont naturellement que les deux cas extrêmes).

Les échos de dispersion pourraient également permettre de sonder la distribution de perméabilités d'un massif poreux où l'on n'a accès que par une extrémité : c'est le cas, par exemple, dans un forage pétrolier. Les mesures que l'on y fait, les diagraphies, introduites par les frères Schlumberger dans les années 1920, ne sont-elles pas un des outils les plus utilisés pour la recherche pétrolière ? Des mesures de dispersion comme celles que nous venons d'examiner représenteraient une extension possible de telles techniques.

Tuyaux et mélange

Le déplacement conjoint de deux fluides miscibles initialement séparés dans un solide poreux conduit généralement, par dispersion, à l'apparition et à la croissance d'une zone de mélange. Cette idée simple sous-tend le fonctionnement de dispositifs pratiques et bien adaptés au mélange de fluides visqueux dans des écoulements qu'on ne peut rendre turbulents.



6. LES MELANGEURS STATIQUES de type «Sulzer» contiennent une âme constituée de nombreuses lames d'orientations variées qui découpent et dévient les liquides dans toutes les directions. Deux liquides injectés simultanément, mais complètement séparés, à l'entrée du mélangeur sont divisés en filets de plus en plus fins, qui sont mêlés, jusqu'à ce que la diffusion moléculaire vienne parfaire le mélange.

Les «mélangeurs statiques» sont bien adaptés aux problèmes de mélange des fluides complexes et visqueux : ils ne contiennent pas de pièces mobiles et ne sont en fait que des tuyaux à l'intérieur desquels sont disposés un ensemble de lamelles d'orientations variées (voir la figure 5). Deux écoulements de fluide différents, injectés côte à côte sont divisés et déviés par les lamelles, ce qui finit par les entremêler de plus en plus finement, jusqu'à ce que le mélange soit bon : ce processus ressemble à la dispersion transverse, où les filets de fluide sont divisés et entremêlés au fur et à mesure des passages d'un pore à un autre. Évidemment l'espacement entre les lamelles sera pris bien supérieur à la taille des pores dans une roche : sinon la dissipation d'énergie dans l'écoulement serait excessive, surtout dans le cas de fluides visqueux. On fait d'ailleurs une pierre deux coups, car on augmente simultanément la dispersion longitudinale et la dispersion transverse, et la largeur de la zone mélangée croît beaucoup plus rapidement.

Ces mélangeurs servent aussi à mélanger des suspensions de grains solides. En général, les méthodes de mélange des liquides s'appliquent mal aux mélanges granulaires secs et provoquent plutôt des ségrégations. En mettant les grains en suspension dans des liquides, on réalise des pâtes auxquelles on applique des techniques de mélange classique, quitte à sécher ensuite le mélange.

Séparation dans un poreux

Nous avons annoncé que les milieux poreux séparent parfois, au lieu de mélanger. C'est le cas en filtration, en microfiltration ou en ultrafiltration, où l'on bloque respectivement le transport de petites particules, de grosses molécules ou de virus. De façon générale, les milieux poreux sont utilisés chaque fois que l'on veut séparer des molécules différentes, en chimie et en biologie.

Ainsi, le chimiste qui effectue une séparation de molécules par chromatographie en phase liquide réalise en apparence une expérience de dispersion : il injecte un petit volume de mélange à séparer à l'entrée d'un milieu poreux (la colonne de chromatographie) où il établit un écoulement régulier d'un fluide porteur

(l'éluant). Comme pour la dispersion, un détecteur placé en fin de colonne analyse les variations des concentrations en divers composés dissous : on observe une succession de pics, souvent étroits, qui correspondent chacun à un composé du mélange.

Qu'est devenue la dispersion ? Elle est toujours là, mais c'est un parasite qu'on cherche à minimiser en réduisant la longueur des pas de la marche au hasard pour les composés du mélange dans le solide poreux. Les résultats précédents montrent que, à cette fin, le poreux doit être aussi homogène que possible, il doit présenter aucune strate, ni aucun chemin préférentiel, ce qui pose de difficiles problèmes techniques ; en outre, il doit être composé de grains très petits (quelques micromètres de diamètre). Des canaux aussi fins présentent une forte résistance à l'écoulement : on doit parfois appliquer des pressions de quelques centaines d'atmosphères pour obtenir une vitesse d'écoulement notable. Une vitesse d'écoulement trop faible finirait par rendre sensible l'influence de la diffusion moléculaire et conduirait à des temps de mesure prohibitifs. Une vitesse excessive augmenterait la dispersion.

Ainsi, dès qu'on regarde l'écoulement dans une colonne chromatographique à une échelle un peu plus grande que la taille des grains microscopiques qui emplissent la colonne, on constate qu'il est parfaitement uniforme sur toute la section de la colonne, et non parabolique, comme dans un tuyau simple.

La chromatographie hydrodynamique permet de séparer des particules de tailles variables en suspension dans un liquide que l'on fait circuler dans un milieu poreux dont les canaux sont plus larges que la taille des particules. Ce sont paradoxalement les plus grosses particules qui arrivent en tête, parce qu'elles doivent emprunter les canaux les plus larges où la vitesse est grande ; les plus petits objets «perdent du temps» dans les canaux étroits ou dans les zones de faible vitesse, près des parois.

Mélangeurs statiques industriels et colonnes de séparation chromatographique ne sont donc pas si différents, malgré leurs objectifs opposés : ils diffèrent surtout par la taille des pores, la façon dont les fluides sont injectés et l'interaction des composants en écoulement et des parois solides.

Le désordre «figé» d'un milieu poreux ou d'un mélangeur statique peut donc assurer le mélange tout comme le désordre perpétuellement changeant d'un écoulement turbulent. Pour mieux comprendre la différence de fonctionnement entre ces deux situations, revenons à l'exemple du mélange que crée la petite cuiller dans le bol de café au lait, et qui est bien adaptée aux fluides peu visqueux. Ce mélange turbulent fait intervenir de multiples échelles de tailles caractéristiques, allant de la taille globale de l'écoulement (celle du bol) à celle des plus petits tourbillons. Pour obtenir un bon mélange, on doit veiller à ce que cette interpénétration soit réalisée jusqu'aux plus petites échelles, où la diffusion assure le mélange.

C'est la combinaison de ces mouvements désordonnés qui conduit à un mélange efficace. Dans un milieu poreux homogène, le mélange est dominé par la seule taille caractéristique des grains et sera peu efficace, ce qui permettra d'obtenir de bonnes séparations en jouant sur l'interaction variable des espèces à séparer et des grains du milieu. En revanche, dans les solides poreux usuels, il existe des hétérogénéités à des échelles de tailles supérieures à celle des grains, qui réalisent une interpénétration des fluides sur une échelle supérieure, ce que font également les écoulements turbulents.

Bien que l'on soit encore loin d'un modèle universel du mélange, que nous avons nommé GUM (pour «grande unification du mélange»), dans un congrès international de notre discipline, certaines caractéristiques communes applicables à des systèmes très différents permettent de mieux comprendre et d'améliorer les opérations de mélange... et de séparation.

Étienne GUYON est directeur de l'École normale supérieure, à Paris. Jean-Pierre HULIN est directeur de recherche au CNRS ; il travaille au Laboratoire Fast de l'Université Paris Sud, à Orsay.

Étienne GUYON et Jean-Pierre HULIN, *Granites et fumées, un peu d'ordre dans le mélange*, Éditions Odile Jacob, 1998.

U. OXAAL, E.G. FLEKKOY et J. FEDER, *Irreversible Dispersion at a Stagnation Point : Experiments and Lattice Boltzmann Simulations*, in *Physical Review Letters*, vol. 75, p. 3514, 1994.

Les écrans plats

ALAN SOBEL

Après des décennies de perfectionnements, les écrans à plasma sont devenus lumineux, de grande taille et si plats qu'on peut les accrocher au mur. Les ingénieurs cherchent maintenant à réduire leur prix.

Le téléviseur est probablement l'appareil électronique le plus répandu au monde. Dans les pays industrialisés, presque tous les foyers ont un poste, et, dans beaucoup de pays, les téléviseurs sont plus nombreux que les téléphones.

Pourtant ces appareils restent très imparfaits : ils sont encombrants, lourds et fragiles. Pis encore, les écrans sont petits : les tubes cathodiques dont la diagonale mesure plus d'un mètre sont onéreux, très lourds (plusieurs centaines de kilogrammes) et si encombrants qu'ils ne passent plus par les portes. De ce fait, les plus grands écrans actuels utilisent des systèmes de projection moins lumineux et dont les angles de vue sont inférieurs à ceux des tubes cathodiques.

Avec l'arrivée imminente de la télévision en haute définition (TVHD), la mise au point de grands écrans devient une urgence industrielle. La nouvelle technique de diffusion, disponible au Japon et présente en Europe et aux États-Unis, permet la réception d'images plus précise que celle des systèmes classiques. Toutefois les images ne seront appréciées des téléspectateurs que si elles apparaissent sur des écrans dont la diagonale est au moins égale à un mètre : de grands écrans permettront d'apprécier tous les détails et captiveront le spectateur comme le fait un écran de cinéma.

Afin de fabriquer ces grands écrans plats, des sociétés de plusieurs pays améliorent des techniques inventées il y a quelques années déjà. De toutes ces techniques, la seule qui soit proche de la viabilité commerciale est celle des écrans à plasma, aussi nommés écrans à décharge gazeuse : les ingénieurs de la Société NEC viennent de présenter un écran de 1,27 mètre de

diagonale pour la TVHD, au format 16/9. Les Sociétés Fujitsu, Mitsubishi, Toshiba et Plasmaco ont également présenté des écrans allant jusqu'à 1,056 mètre (42 pouces) de diagonale, au format 4/3 de la télévision habituelle comme au format 16/9 de la TVHD (dans le format 4/3 des téléviseurs classiques, la longueur de l'image est égale à 4/3 fois sa largeur ; le format 16/9 est plus proche du format du cinéma).

Réduire les prix

En juin 1998, un moniteur (sans récepteur ni haut-parleur) Fujitsu de 1,056 mètre de diagonale coûtait 100 000 francs : ils n'étaient achetés que par des sociétés. La Bourse de New York, par exemple, a installé un millier de ces écrans pour l'affichage des cours. Ces écrans apparaissent également dans les manifestations commerciales où l'image de marque est importante. Ils ne se démocratiseront toutefois que si la production augmente : s'ils sont produits en grande série, ces écrans coûteront moins de 6 000 francs.

Les industriels évaluent le marché potentiel de ces écrans à plusieurs centaines de millions d'unités, soit plusieurs centaines de milliards de francs. Aussi les entreprises japonaises mentionnées précédemment ont-elles investi près de six milliards de francs dans l'adaptation de leurs usines et dans la construction de nouvelles unités de production, tandis qu'elles s'allaient à des sociétés américaines spécialisées dans la télévision. D'autres

1. LES ÉCRANS À PLASMA, minces et légers, ont une résolution suffisante pour afficher les images de haute définition qui seront produites selon les nouvelles normes.



© Chris Burke / Quasidance Photomani



sociétés élaborent aussi des panneaux à plasma de grande taille, mais elles n'ont pas encore construit d'écran de diagonale supérieure à un mètre : les Sociétés *Philips* et *Thomson Tubes Électroniques*, par exemple, portent leur effort sur des écrans plus petits, de haute résolution.

Non seulement les écrans à plasma utilisent à plein les performances des nouvelles normes de télévision, mais ils concrétisent un vieux rêve des ingénieurs : plats et légers, ils s'accrocheront au mur comme un tableau. Au cours des années, plusieurs dispositifs plats de visualisation sont apparus, et quelques-uns, notamment les écrans à cristaux liquides, se sont imposés : dans les ordinateurs portables, dans les tableaux de bord d'avions et dans divers matériels électroniques. Cependant, seule la technique des écrans à plasma autorise la production d'écrans suffisamment grands pour l'affichage d'images très précises et changeant rapidement.

Le fonctionnement

Par de nombreux aspects, les écrans plats sont supérieurs aux tubes cathodiques, ainsi qu'aux systèmes de projection grand format (voir les encadrés des pages suivantes). Beaucoup moins lourds et moins encombrants, ils ont une résolution et une luminosité uniformes sur toute la surface d'affichage, tandis que les tubes cathodiques et les projecteurs ont une résolution moins bonne et sont moins lumineux sur les bords et dans les coins. Les écrans plats sont plus efficaces (la quantité de lumière produite pour une puissance donnée est supérieure) que des dispositifs de projection de même taille et ils consomment moins d'énergie qu'eux, mais ils demeurent moins efficaces que les tubes cathodiques.

Le principe des écrans plats qui semblent le mieux fonctionner ressemble à celui des tubes fluorescents : un gaz traversé par un courant électrique est ionisé, c'est-à-dire que ses atomes et ses molécules perdent des électrons (un gaz ionisé est nommé plasma). Dans les écrans à plasma, le gaz est généralement un mélange d'hélium et de xénon. Le gaz ionisé émet un rayonnement ultraviolet qui excite un luminophore déposé sur la paroi interne du dispositif. En se désexcitant, ce dernier émet de la lumière visible.

Cette émission de lumière se produit dans chacun des millions d'éléments d'image (les pixels) qui constituent la surface de l'écran. Chaque pixel est composé de trois cellules, et chaque cellule contient des électrodes ainsi que des composés photo-émissifs peints sur les parois, sur le fond et, parfois, à l'avant de la cellule.

Chaque cellule fonctionne comme un tube fluorescent dont la couleur rouge, bleue ou verte dépend du type de luminophore. Ces luminophores sont le plus souvent des oxydes dopés. Pour le rouge, on utilise, par exemple, un oxyde d'yttrium dopé avec de l'euporium. Le vert peut être émis par un oxyde de zinc et de silicium dopé avec du manganèse. Enfin, le bleu s'obtient avec un oxyde de baryum, d'aluminium et de manganèse, dopé avec de l'euporium.

Cette brève description ne montre pas pourquoi la mise au point des écrans à plasma est si difficile. L'ionisation des gaz, tout d'abord, pose des problèmes. Lorsque l'on crée une différence de potentiel entre deux électrodes placées dans un gaz, aucun courant ne circule tant qu'il n'y pas de particules chargées (ion ou électron) pour le transporter. Dès sa création, la première particule chargée se déplace dans le gaz sous l'action de la force engendrée par la différence de potentiel ; elle rencontre alors d'autres atomes et molécules, qu'elle ionise et qui, à leur tour, ionisent d'autres atomes. Ainsi augmente le courant transporté entre les électrodes. Lorsque les ions positifs du gaz captent des électrons, ils émettent de la lumière ultraviolette, visible et infrarouge.

Pour obtenir des écrans efficaces, on doit à la fois faciliter la création d'ions primaires et limiter le courant qui circule entre les électrodes. Pour fournir les premières particules chargées, les systèmes à plasma utilisent une décharge auxiliaire, de sorte que la commutation puisse avoir lieu plusieurs fois par seconde.

D'autre part, les cellules doivent être conçues de manière que le courant qui circule entre leurs électrodes n'augmente pas excessivement. Sans une limitation de courant, le gaz s'ioniserait totalement, sa résistance électrique diminuerait et la décharge se transformerait en un arc électrique qui détruirait la cellule.

Les signaux organisés

Les industriels fabriquent deux types d'écrans à plasma : les écrans à courant alternatif et les écrans à courant continu. Dans les écrans à courant alternatif, les plus répandus, le courant est limité grâce à la conception même des cellules : les deux électrodes situées dans chaque cellule sont séparées du gaz par une mince couche de matériau électriquement isolant, qui, comme un condensateur, ne peut être traversée que par un courant alternatif. En outre, ce condensateur limite le débit de courant et empêche la décharge de se transformer en arc.

Un écran à plasma ne se réduit pas à une série d'éléments lumineux : les ingénieurs doivent organiser la manière dont les signaux électriques commandent les cellules pour former des images. Comme pour tous les écrans plats, les cellules d'un panneau à plasma sont disposées selon un quadrillage. Les conducteurs électriques menant aux électrodes sont les lignes et les colonnes de ce quadrillage, dont chaque intersection comporte une cellule photo-émissive. Chaque conducteur est commandé par un circuit électronique qui émet le courant de commande.

Tous les écrans actuels étant des descendants des tubes cathodiques, de nombreuses techniques utilisées aujourd'hui pour la commande des écrans à plasma sont issues de leur vénérable ancêtre. Dans un tube cathodique, un faisceau d'électrons balaie l'écran. Le faisceau commence en haut à gauche, balaie une ligne vers la droite, pixel après pixel, descend d'un cran, puis balaie une nouvelle ligne, et ainsi de suite jusqu'à la ligne du bas, après laquelle il repart en haut de l'écran. Une image complète forme une trame, et le tube affiche 25 trames par seconde (30 en Amérique du Nord). Les écrans d'ordinateur, eux, sont «rafraîchis» 50 fois par seconde (60 fois aux États-Unis), que l'image soit modifiée ou non.

Avec leur arrangement matriciel et leurs cellules qui réagissent lentement, les écrans plats ne peuvent utiliser la technique de balayage des tubes cathodiques. Le signal commande l'affichage simultané de toutes les cellules d'une ligne, et une trame est obtenue par l'affichage de lignes successives.

Dans un écran à plasma à courant alternatif, on applique des séries d'im-

pulsions de tension trop faibles pour déclencher une décharge. Cette tension est nommée «tension de maintien». Pour allumer une cellule particulière, on ajoute une tension (d'«excitation») à la ligne et à la colonne correspondante. Le courant qui circule charge alors le

condensateur de la cellule ; puis l'accumulation de charges produit une tension de sens contraire à la tension d'excitation, la décharge s'arrête ainsi que l'émission de lumière. Lors de l'impulsion suivante, de polarité opposée, l'effet de la charge résiduelle

s'ajoute à celui de la tension de maintien, et la cellule s'allume de nouveau. Elle reste allumée aussi longtemps que cette tension de maintien est appliquée. Chaque décharge ne dure qu'une microseconde environ et n'émet qu'une faible quantité de lumière, mais comme le nombre d'impulsions par seconde atteint 50 000, la quantité totale de lumière émise peut être notable.

On éteint une cellule en appliquant une tension qui réduit la charge électrique déposée sur les condensateurs. D'autre part, on règle la brillance des cellules en commandant le temps pendant lequel elles restent allumées. Ce système a un inconvénient, commun aux écrans plats : on doit disposer de très nombreux circuits de commande autour de l'écran. Ainsi, pour un écran classique d'ordinateur, avec ses 480 lignes et ses 640 colonnes de pixels, composés chacun de trois cellules, on doit utiliser 2 400 circuits ($480 + 640 \times 3$),



UN PIXEL



FACE AVANT EN VERRE

ISOLANT PROTECTEUR

COUCHE ISOLANTE

ISOLANT PROTECTEUR

ELECTRODES D'ADRESSAGE

RIB

LUMINOPHORE ROUGE

LUMINOPHORE VERT

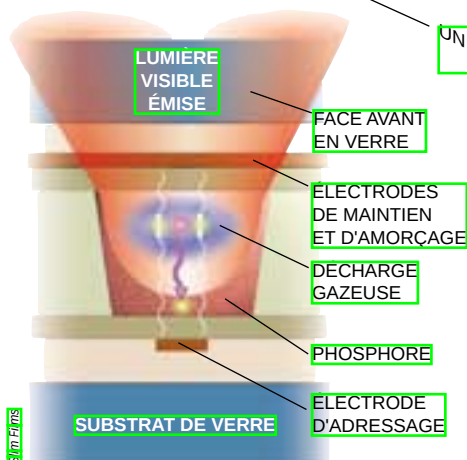
LUMINOPHORE BLEU

UN PIXEL

ELECTRODES D'AMORÇAGE

ELECTRODES DE MAINTIEN

SUBSTRAT DE VERRE



2. UN ELEMENT D'IMAGE (ou pixel) d'un écran à plasma est composé de trois minuscules cellules, dont chacune est intérieurement recouverte d'atomes de phosphore (ci-dessous). Quand le gaz d'une cellule est ionisé, la cellule émet une couleur primaire. L'ionisation a lieu lorsqu'une impulsion de tension est appliquée entre une électrode d'adressage et des électrodes de maintien. Le gaz ionisé, nommé plasma, émet un rayonnement ultraviolet qui excite le luminophore. À leur tour, ceux-ci émettent de la lumière visible (en bas, à gauche). Les électrodes d'amorçage, adjacentes aux électrodes de maintien, délivrent les impulsions de courant nécessaires au déclenchement de chaque décharge. Comme dans un écran cathodique, les niveaux de luminance de chaque couleur primaire émise par chacun des milliers de pixels permettent l'affichage de n'importe quelle couleur et de n'importe quelle image (à gauche).

quelle que soit la taille matérielle de l'écran (en comparaison, un tube cathodique n'a que deux circuits pour commander les déviations horizontales et verticales du faisceau, et trois circuits de modulation, un pour chaque couleur). Même avec la diminution du coût des circuits intégrés, ceux-ci sont si nombreux que les écrans à plasma restent chers.

Dans les écrans à plasma à courant continu, le courant est également appliqué par impulsion, mais les impulsions ont toutes la même polarité. Différents systèmes sont testés, mais, aujourd'hui, le seul grand écran à courant continu a été mis au point par les Laboratoires NHK (le centre public japonais de recherches en télé-

communication). Ces derniers ont présenté un écran de 1,056 mètre de diagonale, avec une bonne luminance et un bon contraste, mais ce système demeure lourd et moins efficace que les écrans à courant alternatif. De plus, il chauffe de manière excessive, et sa structure complexe le rend coûteux à fabriquer.

Un autre problème concerne la durée de vie. L'électrode négative (la cathode) s'use : soumis à un bombardement continu, les atomes sont éjectés de la cathode et se redéposent partout sur l'intérieur de la cellule. Le problème est plus gênant avec les écrans à courant continu, car le matériau pulvérisé est un métal et non un isolant, de sorte que des courts-

circuits peuvent se produire. À la longue, la cellule s'assombrit et l'écran devient moins lumineux.

Un groupe de Saint-Petersbourg a présenté un autre dispositif à courant continu, qui semble prometteur : cet écran est constitué d'un assemblage de morceaux de 20 centimètres carrés de côté, comportant chacun 64×64 pixels. Les raccords, très fins, sont quasi invisibles. Chaque morceau étant piloté de façon indépendante, la luminance de l'écran est indépendante de sa taille, ce qui n'est pas le cas pour l'écran japonais, où la luminance est inversement proportionnelle au nombre de lignes. Cette structure composite ouvre la voie aux écrans de très grande dimension.

Grands, lumineux et organiques

Parmi les nombreuses structures d'écrans plats, celle des écrans électroluminescents organiques semble particulièrement prometteuse : les molécules photoémissives sont ici des molécules organiques, c'est-à-dire composées d'atomes de carbone, d'hydrogène, d'oxygène et d'azote. Avec leurs couches actives 1 000 fois plus minces qu'un cheveu, ces dispositifs ont été présentés pour la première fois il y a juste dix ans. Les meilleurs d'entre eux, beaucoup plus brillants que les tubes cathodiques des téléviseurs actuels, fonctionnent sous une tension inférieure à dix volts durant des dizaines de milliers d'heures. Suffisamment rapides pour afficher des films vidéo, les matériaux organiques sont des composants de choix pour les écrans d'ordinateurs et de télévision de demain.

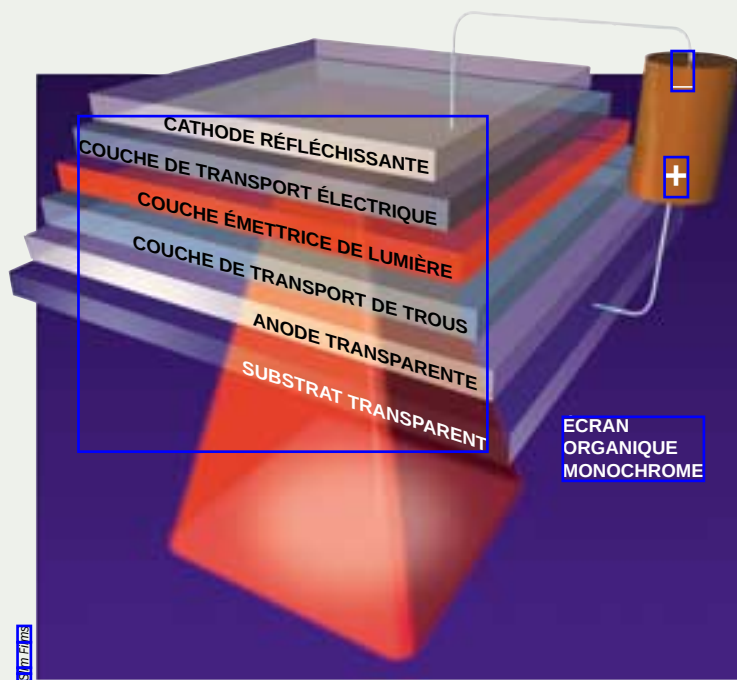
Ces appareils ont des caractéristiques étonnantes. Notamment les faibles liaisons qui s'exercent entre les molécules organiques permettent de les déposer sur des substrats plastiques minces et flexibles. Les écrans peuvent ainsi être roulés ou moulés sur des contours variés. Les couches actives et leurs supports plastiques sont si légers que le poids d'un écran d'ordinateur portable pourrait, en théorie, être réduit à une centaine de grammes.

Pour l'instant, l'écran le plus grand mesure 12,5 centimètres de diagonale, mais, en raison du faible coût des matériaux utilisés, on imagine un mur recouvert d'un écran organique géant. Les potentialités de ces matériaux intéressent des Sociétés comme *Eastman Kodak* ou *Motorola*, qui ont investi des milliards de francs dans des programmes de recherche appliquée, mais aussi de petites sociétés spécialisées.

De la lumière, naturellement

La plupart des écrans organiques comportent trois couches de semi-conducteurs organiques : la couche de transport électronique, la couche émettrice de lumière et la couche de transport de trous. Ces couches sont entou-

rées par une cathode et une anode (voir la figure ci-dessous). Les particules négativement chargées (les électrons) et les porteurs de charges positives (les trous, qui correspondent à l'absence d'un électron) sont émis respectivement par la cathode et par l'anode sous l'action du champ électrique. Ils passent alors dans la région émettrice de lumière, où ils se recombinent, formant ce que l'on nomme des excitons. Ces derniers, électriquement neutres, sautent d'une molécule à l'autre et se dés excitent en quelques milliardièmes de seconde, émettant un photon dont l'énergie détermine la couleur. En choisissant une molécule adéquate pour la couche émettrice, on obtient de la lumière rouge, verte ou bleue. Selon la configuration des



Encore des difficultés

Nous l'avons vu, l'obstacle principal à la réalisation des écrans matriciels est le grand nombre de circuits de commande nécessaire : un pour chaque ligne et pour chaque colonne, plus un autre pour activer les électrodes au bon moment. Pour les écrans à plasma, ce problème est aggravé par les tensions élevées (de l'ordre de 100 volts) que doivent supporter les circuits intégrés. De surcroît, ces circuits doivent résister à l'intense impulsion de courant qui circule au début de chaque décharge. Pour cette raison, les systèmes électroniques de ces écrans coûtent plus cher que ceux des systèmes de visualisation concurrents. Ce coût

diminuera cependant avec le volume de production.

Un autre problème récurrent des écrans à plasma est l'usure : le bombardement ionique détruit progressivement le revêtement luminescent, ce qui réduit son efficacité lumineuse. Naguère, ce phénomène limitait la durée de vie des appareils à 10 000 heures, mais aujourd'hui certains constructeurs annoncent des durées de vie de 30 000 heures, analogues à celles des tubes cathodiques ; après cette période, la luminance est inférieure à la moitié de la luminance initiale. Pour allonger la durée de vie des écrans, les constructeurs maintiennent la décharge éloignée des luminophores.

L'efficacité et la luminance des panneaux à plasma à courant alternatif s'amélioreront. Aujourd'hui, l'efficacité est de un lumen par watt environ, et elle augmente. La luminance est en train de dépasser 100 candelas par mètre carré, et un constructeur a annoncé 300 candelas par mètre carré, valeur peu inférieure aux 350 candelas par mètre carré habituels dans les écrans de télévision à tube cathodique. Les écrans de télévision doivent être plus lumineux que les écrans d'ordinateurs, parce qu'ils sont plus loin de l'utilisateur et qu'ils doivent apparaître lumineux, même dans des pièces éclairées.

La luminance des écrans à plasma, qu'ils utilisent du courant alternatif

couches, la cathode ou l'anode doit être transparente pour que la lumière atteigne l'observateur.

Avec les deux électrodes transparentes (en oxyde d'étain et d'indium, par exemple), on peut imaginer des applications nouvelles, comme des vitres écrans où l'image serait visible des deux côtés de l'écran.

Les composés organiques peuvent aussi être des polymères, c'est-à-dire de longues chaînes composées de nombreuses sous-unités identiques (parfois plus de 100 000). Les écrans à polymères mis au point par cette méthode sont presque aussi lumineux que leurs cousins qui utilisent de petites molécules déposées sous vide, mais leur durée de vie est dix fois inférieure, en partie parce qu'on les purifie moins bien.

Applications commerciales

De tous les écrans organiques, le plus avancé émet une lumière verte. Son efficacité et sa durée de vie sont

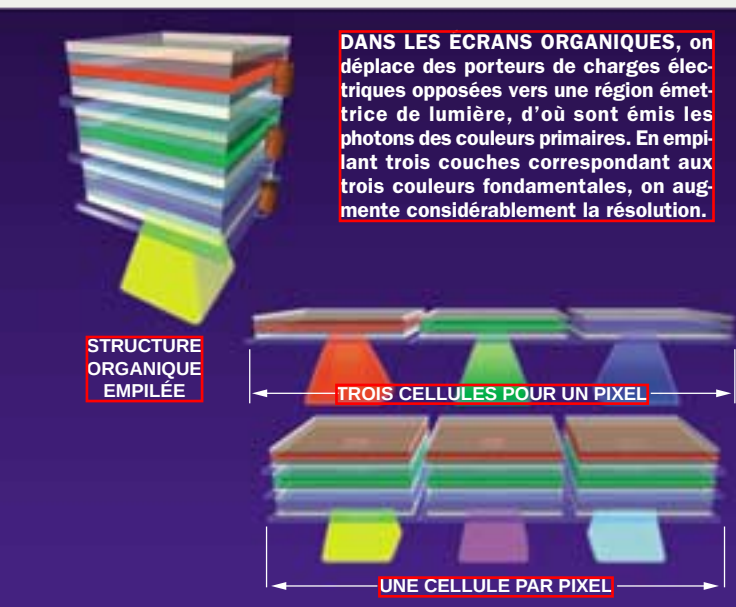
supérieurs à ceux des appareils qui émettent dans le rouge ou dans le bleu. Son efficacité lumineuse est de 18 lumens par watt, soit cinq à dix fois supérieure à celle d'un afficheur à cristaux liquides, mais quatre fois inférieure à celle d'un tube cathodique. Protégé hermétiquement de la vapeur d'eau, un écran organique conserve pendant des dizaines de milliers d'heures une luminance supérieure à 50 pour cent de la luminance initiale. La principale difficulté concerne la surchauffe, mais on devrait bientôt disposer de matériaux émettant efficacement dans le rouge et dans le bleu sans chauffer trop.

Les écrans organiques monochromes émettant dans le vert commencent à être commercialisés. La Société Pioneer vend aussi un écran organique monochrome de 2,5 centimètres de hauteur et de 10 centimètres de large, qui affichera des informations de navigation dans les automobiles.

Ces écrans monochromes préfigurent les écrans organiques du futur. En empilant trois couches organiques transparentes, chacune émettant une couleur particulière, rouge, bleue et verte, nous avons produit n'importe quelle combinaison de ces trois couleurs (voir le schéma, à gauche). Ces couches organiques empilées pourraient remplacer les architectures classiques d'écrans couleur, où les pixels (éléments d'image) rouges, verts et bleus sont juxtaposés sur une seule couche.

Dans la configuration empilée, la définition est bien meilleure qu'avec les tubes cathodiques ou même qu'avec les écrans à plasma, où trois pixels, un bleu, un rouge et un vert, sont nécessaires pour chaque point d'image. Dans la structure empilée, chaque pixel peut émettre n'importe quelle couleur : on triple ainsi la résolution du dispositif, tout en éliminant des zones sombres inutiles entre les pixels. De surcroît, cet empilement de couleurs donne un affichage en couleurs quelle que soit la distance à l'écran, parce que l'œil n'a pas à résoudre de problèmes de proximité de pixels.

Paul BURROWS et Stephen FORREST, Université Princeton, Mark THOMPSON, Université de Californie du Sud



Les concurrents

Aujourd'hui, les écrans à plasma dominent le marché des grands écrans, mais plusieurs autres techniques s'améliorent parallèlement.



Panasonic Consumer Electronics

Les projecteurs. Les systèmes de projection ne sont pas des écrans que l'on accroche au mur, mais ils affichent de très grandes images. L'image projetée provient soit d'une tête constituée de trois tubes cathodiques monochromes mesurant 180 à 230 millimètres de diamètre, soit d'une source qui émet une lumière intense à travers trois «valves optiques» (des projecteurs de diapositives électroniques).

Plus massifs que des panneaux à plasma, les projecteurs ont une résolution si médiocre qu'on n'envisage pas leur utilisation pour la télévision en haute définition.

Les tubes cathodiques. Gros, lourds et gourmands en énergie, les tubes cathodiques modernes résultent de 50 ans de perfectionnements. Leurs principaux défauts sont leur taille limitée (la diagonale est inférieure à 90 centimètres) et leur encombrement.

Les cristaux liquides. Ces écrans se sont imposés avec les ordinateurs portables. Pour satisfaire une demande croissante, l'industrie améliore sans cesse la résolution, les angles de vue, la couleur et l'efficacité de ces écrans. Les meilleurs d'entre eux, ceux à matrice active, sont composés de cellules commandées chacune par un commutateur électronique (un transistor en couches minces). Cette configuration augmente le coût, mais améliore les performances. De grands écrans sont à l'étude. Par exemple, la Société Sharp a présenté un écran couleur de 1,015 mètre de diagonale, mais le coût augmente considérablement avec la taille. En outre, les cristaux liquides sont peu appropriés pour la télévision, parce que leurs cellules réagissent lentement.

Les cristaux liquides adressés par plasma sont des hybrides intéressants. Inventés par la Société Tektronix, ils sont aujourd'hui perfectionnés par les Sociétés Sony, Sharp et Philips. On allume une

ligne en ionisant de minces canaux remplis de gaz (moins onéreux que des transistors). Comme dans la plupart des écrans à cristaux liquides classiques, on commande les cellules individuelles en commandant la tension des électrodes de colonnes. Cette approche est particulièrement intéressante pour les grands écrans. Des panneaux de 1,065 mètre ont déjà été présentés lors de salons. La luminance est excellente, mais, malheureusement, l'affichage d'images animées souffre des problèmes inhérents aux cristaux liquides et l'angle de vue est un peu inférieur à celui qu'on obtient sur certains autres écrans à cristaux liquides.



Sanjiv

Les diodes électroluminescentes. Ce sont des composants semi-conducteurs où les électrons et les trous (une absence d'électron) sont injectés dans une région où ils se recombinaient en émettant de la

lumière (et un peu de chaleur). Après des années de recherches, on sait construire des diodes bleues et l'on peut concevoir des écrans en couleur à diodes. Ces écrans sont constitués de plusieurs milliers de pixels, chacun comportant une diode rouge, une diode verte et une diode bleue. Ces écrans peuvent mesurer plusieurs mètres de large et sont suffisamment lumineux pour être vus en plein jour. Leur prix élevé restreint toutefois leur utilisation aux grandes manifestations.



Sharp

Les panneaux électroluminescents. Dans ces écrans, une mince couche d'atomes photo-émisifs est en sandwich entre deux couches isolantes. Une différence de potentiel entre les deux couches externes crée des électrons entre l'isolant et le phosphore ; ces électrons émettent de la lumière comme les électrons des tubes cathodiques. Les écrans à couches minces



USI Micro-Electronics

3. CET ECRAN A PLASMA MESURE 1,056 mètre de diagonale et coûte environ 100 000 francs. Avec une épaisseur inférieure à 15 centimètres et une masse de 40 kilogrammes, cet écran, affiche 16,7 millions de couleurs et préfigure la télévision de demain.

ou du courant continu, est suffisante pour beaucoup d'applications, mais elle demeure inférieure à celle des tubes cathodiques. Comment l'augmenter ? Si l'on augmente la puissance, on se heurte à des augmentations rédhibitoires du prix des circuits de commande et des alimentations. Aujourd'hui, les ingénieurs explorent deux directions : ils cherchent à augmenter le nombre de photons produits à chaque décharge et à améliorer l'efficacité de conversion des photons ultraviolets en photons visibles.

Comme pour tout produit technique, le succès commercial d'un écran tient autant à l'invention proprement dite qu'à l'amélioration des techniques de production. Un appareil révolutionnaire, mais trop cher pour être compétitif ne peut réussir : les bonnes idées périssent, sauf quand

électroluminescentes sont robustes, possèdent de grands angles de vue et une longue durée de vie. S'ils demeurent chers, petits et monochromes, ils se sont néanmoins imposés pour plusieurs applications.



Les écrans à émission de champ. Il y a 20 ans, des chercheurs de l'Université Stanford ont montré que de faibles tensions électriques appliquées à des pointes de molybdène ou d'autres métaux réfractaires créent des champs électriques suffisamment intenses pour extraire des électrons du métal à la température ambiante. Cette émission électronique «de cathode froide» est aujourd'hui le fondement d'un dispositif de visualisation sans tube qui utilise plusieurs cathodes par pixel.

A l'époque où l'émission de pointe fut découverte, on ignorait comment réaliser industriellement ces cathodes. Les progrès de la micro-électronique et de la physique des matériaux ont résolu le problème. La Société française *PixTech* commercialise les premiers écrans à émission de champ ; des sociétés américaines la suivent.

La Société *Canon* vient d'annoncer un autre type de cathode, qui faciliterait la fabrication de ces écrans. On évite de sculpter des pointes microscopiques sur du silicium monocristallin en créant une mince bande isolante entre deux zones métalliques (qui servent d'électrodes) déposées sur du verre. Si ce dispositif fonctionne comme prévu, des cathodes de n'importe quelle taille pourraient être réalisées par des systèmes bon marché analogues aux imprimantes à jets d'encre.

Si de nombreux problèmes demeurent, les écrans à émission de champ sont de sérieux concurrents des écrans à cristaux liquides, pour les écrans de petites dimensions. Pour ce qui concerne les grandes dimensions, ce seront peut-être même des concurrents des écrans à plasma.



Les écrans fluorescents. Ces écrans, qui s'inspirent d'une idée ancienne, sont des sortes de tubes cathodiques plats fonctionnant à basse tension. Les électrons émis par des filaments cathodiques froids sont dirigés vers des atomes qui émettent de la lumière avec un bon rendement malgré la basse tension. Les pixels peuvent être sélectionnés et allumés par un réseau de grilles métalliques situées entre les cathodes et les phosphores. Ces dispositifs fluorescents sous vide sont souvent utilisés pour les horloges de voitures, dans les affichages de fours et, aussi, pour les magnétoscopes. Par cette technique, les ingénieurs n'ont, pour l'instant, réalisé que de petits affichages.

elles sont soutenues par des entreprises qui disposent de fonds importants et qui jouissent d'un soutien d'État.

Les entreprises japonaises bénéficient aujourd'hui de leur volonté et des investissements effectués depuis 30 ans. Elles doivent aussi beaucoup à l'inventivité de leurs ingénieurs et de leurs ouvriers, qui ont assidûment perfectionné les techniques de production. Pour ces raisons, les usines d'écrans de visualisation sont pour la plupart en Asie.

L'amélioration de ces écrans est un travail motivant, source d'idées novatrices. De nouvelles techniques apparaissent, alors même que les anciennes s'améliorent. Ces écrans, qu'ils soient à plasma ou issus d'une autre technique, sont des éléments primordiaux de l'interaction entre les humains et les machines de demain.

Alan SOBEL, consultant à Evanston (Illinois), travaille depuis 35 ans dans le domaine de l'affichage électronique.

Flat-Panel Displays and CRTs, sous la direction de Lawrence E. Tannas, Jr., Van Nostrand Reinhold, 1985.

Electro-optical Displays, sous la direction de M. Karim, Marcel Dekker, 1992.

T. MAKINO et al., *Improvement of Video Image Quality in AC-Plasma Display Panels by Suppressing the Unfavorable Coloration Effects with Sufficient Gray Shades Capability*, in *Asia Display*, 1995.

Display Systems: Design and Applications, sous la direction de Lindsay W. MacDonald et Anthony C. Lowes, Wiley SID series in display technology, published in association with the Society for Information Display, John Wiley & Sons, 1997.

Display Systems: Design and Applications, sous la direction d'A. Lowe, John Wiley & Sons, 1997.

La norme européenne de télévision numérique

JEAN-PIERRE HENOT • VINCENT MICHON • BERTRAND SUEUR

La nouvelle norme européenne de télévision numérique remplacera bientôt la norme de transmission analogique dans tous les équipements télévisuels.

Depuis 20 ans, le monde de la télévision prépare son évolution vers les systèmes numériques, où les sons et les images sont codés par des nombres au lieu d'être transmis sous forme de signaux continus. Dans les années 1980, les effets spéciaux numériques se multiplient. Aujourd'hui, la numérisation s'est imposée dans l'ensemble du monde audiovisuel : elle déborde du studio professionnel et envahit tous les maillons de la chaîne de l'image, des caméras aux téléviseurs, en passant par les satellites ou les réseaux câblés.

Cet avènement du numérique est facilité par l'adoption d'une norme européenne issue des travaux du groupe DVB (*digital video broadcast*, soit «diffusion vidéo numérique»). Ainsi, les grandes sociétés d'informatique, de télécommunication et d'audiovisuel multiplient les produits entièrement numériques. L'utilisateur bénéficiera de programmes de télévision plus nombreux. Pour les professionnels de l'audiovisuel, la numérisation diminue le coût de diffusion des programmes.

Pourquoi la numérisation des images nécessite-t-elle une compression des données et comment cette compression est-elle faite ? Une fois numérisée, l'image est transmise aux téléspectateurs par des satellites, par des câbles ou par voie hertzienne terrestre. Quels sont les avantages et les inconvénients de ces modes de transmission ? Enfin, quelle que soit la méthode choisie, la transmission numérique nécessite des «codes correcteurs d'erreur», recettes mathématiques qui permettent de retrouver le message original, même lorsque celui-ci a été en partie altéré. Comment ces codes fonctionnent-ils et comment

mélange-t-on plusieurs flots de données dans un flot plus important, ce qu'on nomme le multiplexage ? L'examen de ces questions montrera combien le traitement numérique des sons et des images est supérieur à l'ancien système analogique de diffusion.

Nous verrons aussi comment la numérisation permet de crypter l'information pour contrôler l'accès aux programmes de télévision payants. Aujourd'hui, près d'une centaine de chaînes numériques sont diffusées par satellite ou par le câble. La qualité encore imparfaite de ces transmissions devrait être améliorée par l'augmentation des débits : la télévision en haute définition pourra alors s'imposer.

Réduire les débits

La norme numérique professionnelle d'échantillonnage des images se fonde sur les caractéristiques de la télévision analogique, qui affiche 25 images par seconde, avec 625 lignes par image, dont 576 sont effectivement visibles à l'écran (les autres lignes transmettent des informations complémentaires, comme le télétexte). Chaque ligne est composée de 720 points, codés par un nombre à 16 chiffres binaires (16 bits). Ainsi, la transmission des images numérisées nécessite un débit d'information égal à 165 mégabits par seconde.

Ce débit considérable est nécessaire dans les studios professionnels qui doivent manipuler les images et garantir des enregistrements parfaits, même après de multiples copies, mais il est excessif pour des utilisations domestiques : la capacité des canaux de transmission ou le débit des magnétoscopes grand public sont 10 à 40 fois inférieurs.

Grâce aux techniques de compression numérique, on diminue le débit au prix d'une dégradation maîtrisée de la qualité des images et des sons. Cette compression résulte de deux opérations principales : d'une part, on évite de transmettre la totalité de chaque image fixe ; d'autre part, on utilise les redondances entre images, de sorte que, dans la transmission d'une succession d'images, on ne transmet que les changements entre les images.

Une image fixe, telle la photographie d'un clocher gris sur un ciel bleu, se décompose en plusieurs zones. Un point bleu est le plus souvent entouré d'autres points bleus, tandis que la plupart des points gris sont cernés d'autres points gris. Cette règle s'applique à toute l'image, excepté sur les contours, où le gris de la pierre jouxte le bleu du ciel.

En pratique, pour analyser les détails de l'image, on la décompose en pavés jointifs de huit lignes par huit points. À l'intérieur de ces pavés, on effectue une opération mathématique qui transforme les 64 nombres qui codaient la couleur des 64 points en 64 autres nombres. On choisit cette opération réversible de manière à présenter l'information sur le bloc en fonction de la qualité subjective de l'image : la suite de 64 nombres obtenus est caractéristique de la quantité de détails de l'image, c'est-à-dire du nombre de contours. Chacun des 64 nombres correspond à un niveau de détail. Comment transmettre une image plus grossière que l'image initiale ? En choisissant de transmettre seulement une partie de ces nombres (par exemple les 16 premiers). Puis, dans le récepteur, l'opération mathématique inverse reconstruit l'image initiale, mais moins détaillée.

On affine cette compression des images fixes en utilisant les imperfections du système visuel humain, qui remarque plus facilement une tache sur un pantalon uni que sur un pull-over chiné. Ainsi, on réduit peu la qualité des images si l'on transmet plus grossièrement les zones chinées que les zones unies. La quantité d'informations sur les discontinuités ainsi sacrifiées lors de cette étape détermine le débit de l'image comprimée, et donc, sa qualité.

Dans les séquences animées, l'impression du mouvement résulte de l'affichage d'une nouvelle vue fixe tous les 1/25 de seconde. Or, deux images successives d'une série sont comparables à deux photographies d'une même scène, prises à une fraction de seconde d'intervalle : elles se ressemblent beaucoup. Les systèmes de codage exploitent cette redondance temporelle en distinguant trois types d'images. Premièrement, des images codées intrinsèquement, c'est-à-dire sans référence aux images voisines (images de type *I*, comme *Intra*), sont indispensables pour amorcer le décodage lorsque le téléspectateur change de chaîne ou après un changement de plan, et leur transmission nécessite la moitié du débit total. Deuxièmement, les images codées à partir des ressemblances avec une image précédente (image de type *P*, comme *Prédit*) : on transmet uniquement les différences entre ces images et des images

précédentes ; leur transmission nécessite quatre fois moins de débit que les précédentes. Enfin, pour les images interpolées de type *B*, on transmet les différences entre ces images et la moyenne des deux images *I* et *P* situées de part et d'autre (les images sont remises dans le bon ordre à l'arrivée) ; elles n'occupent que quelques pour-cent du débit total (voir la figure 3).

Ces principes ont été définis par l'Organisation internationale de normalisation (ISO) qui a publié des algorithmes nommés MPEG1 et MPEG2 vidéo (*Moving Picture Expert Group*, soit «groupe d'experts pour les images animées»). Le premier, prévu pour l'enregistrement de séquences d'images sur disque compact à des débits inférieurs à 1,5 mégabit par seconde, traite des images quatre fois plus petites que les images normales de télévision. La norme MPEG2 étend la norme MPEG1 au format de la télévision classique et peut même être utilisée pour de la télévision en haute définition.

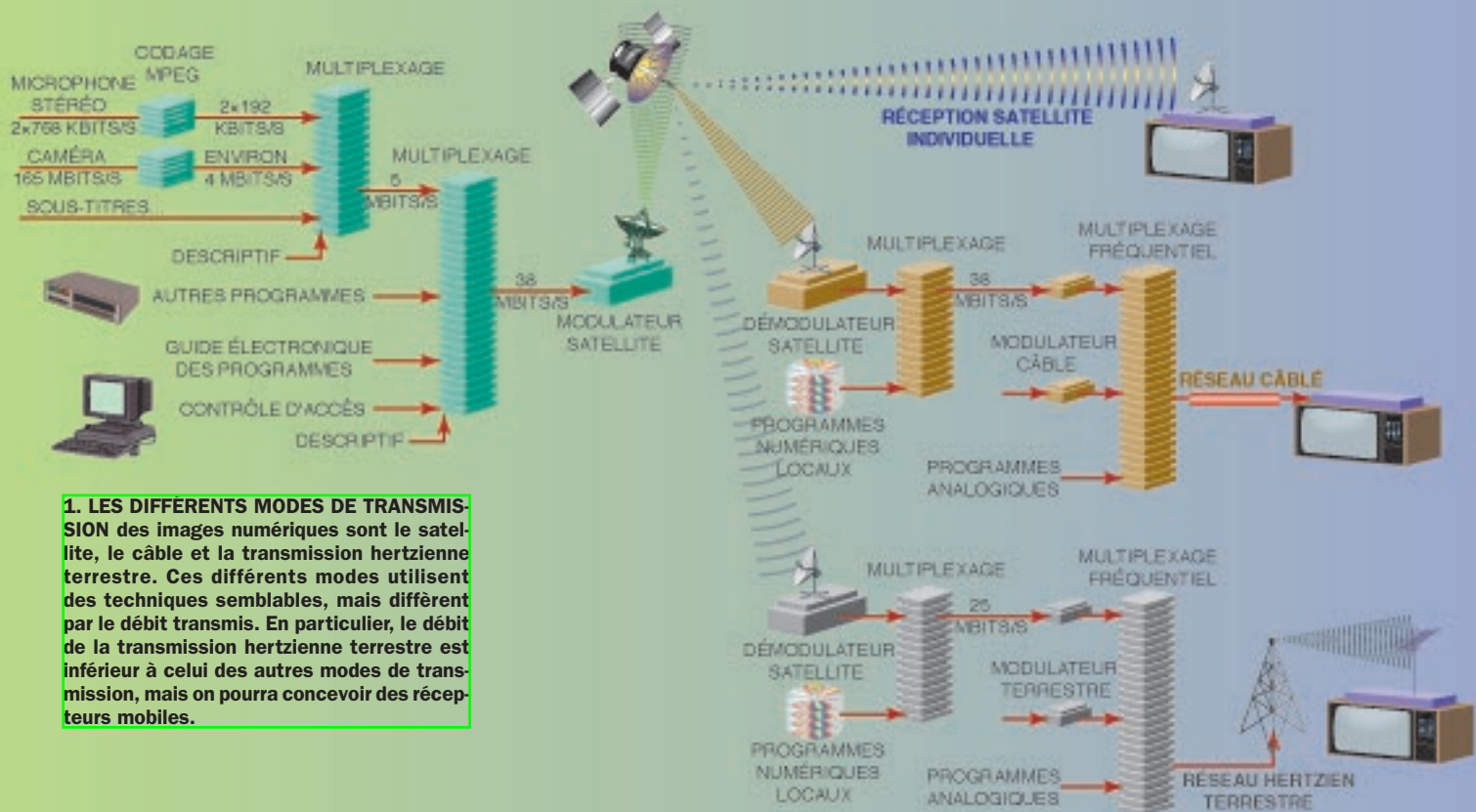
Le son, à l'appui de l'image

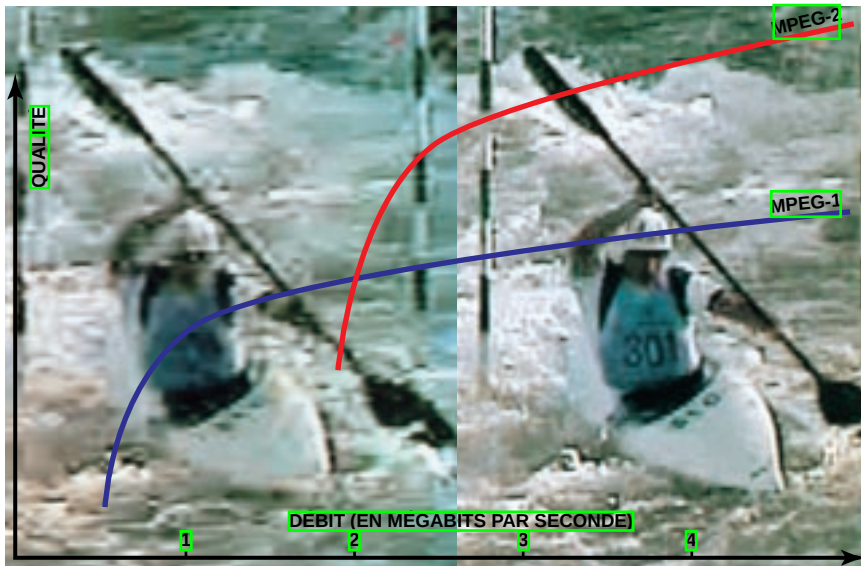
La télévision ne se réduit pas aux images : elle s'enrichit grâce au son. La numérisation des bandes sonores qui accompagnent les images vidéo est similaire à celle des images elles-mêmes. Dans la norme MPEG2, les sons sont échantillonnés à 48 kilohertz sur 16 bits

minimum : 48 000 fois par seconde, on code l'intensité du son selon 65 536 niveaux (2^{16}). Le débit occupé par le son est ainsi de 768 kilobits par seconde et, comme la bande passante est une ressource rare, les données audio sont également comprimées.

L'oreille étant le dernier élément de la chaîne de transmission des sons, les ingénieurs utilisent ses imperfections afin de réduire le débit sans dégrader la qualité perçue. On divise par exemple la gamme des fréquences audibles, des plus graves aux plus aiguës, en 32 bandes de fréquences distinctes. Toutes les 8 millisecondes, on mesure l'intensité des 32 bandes. Puis on utilise une particularité du système auditif humain, qui ne perçoit pas un son lorsqu'il est émis en même temps qu'un autre son plus puissant et de fréquence voisine : le démarrage d'un camion couvre le bruit sourd du moteur de la voiture qui le suit. Par exemple, lorsque la bande 25 est si forte que l'on n'entend pas beaucoup la bande 24 ni la bande 23, on transmettra une information minimale pour ces bandes, à condition que cette réduction d'information reste inaudible.

La norme MPEG1 code sans défaut audible un son stéréophonique à un débit de deux fois 256 kilobits par seconde, voire deux fois 192 kilobits par seconde si l'on élimine certaines redondances entre les deux voies. Cette norme est aujourd'hui utilisée pour la





2. LES CODAGES MPEG1 ET MPEG2 ont été mis au point pour transmettre des images animées avec des qualités différentes : le standard de compression MPEG1 vise la transmission des images de qualité moyenne à très bas débit. Au-delà de deux mégabits par seconde, le standard MPEG2 prend le relais. Au prix d'un débit élevé, la qualité d'image est alors améliorée, mais cette dernière dépend aussi de la nature de l'image : les séquences de sport sont plus difficiles à coder que les films. À gauche, l'image du kayak est codée à huit mégabits par seconde, tandis que la même image (à droite), codée à seulement deux mégabits par seconde, est plus grossière.

plupart des applications sonores pour le grand public, comme la radiodiffusion numérique sonore DAB (*digital audio broadcast*, soit «diffusion radio numérique»). La norme MPEG2 a étendu le codage MPEG1 aux sons à plusieurs canaux, tout en préservant la compatibilité : un son codé en MPEG2 sur cinq canaux, en «pentaphonie», peut être restitué en stéréophonie par des décodeurs MPEG1, et, inversement, les sons codés en MPEG1 sont interprétables par les décodeurs MPEG2. À l'avenir, les cinq canaux de la norme MPEG-audio pourront être utilisés, en particulier pour inclure des effets sonores comme l'effet «spatial» utilisé dans les salles de cinéma, où le spectateur est entouré de sons.

Comment choisir le débit des données vidéo et des données sonores pour satisfaire le téléspectateur ? En demandant à des groupes de personnes de comparer des séquences audio ou vidéo non comprimées à ces mêmes séquences comprimées à différents degrés. Ces testeurs, composés de spécialistes pour le son et de personnes ordinaires pour l'image, notent la qualité du son ou de l'image sur une échelle dont les extrêmes correspondent à «franchement gênante» et «non perceptible».

Au Centre commun d'études de télédiffusion et télécommunications (CCETT), à Rennes, nous avons amélioré ces méthodes : au lieu tester de courtes

séquences notées une à une, les jurys disposent d'un curseur allant de «bon» à «mauvais» qu'ils actionnent pendant toute la durée du test. Les séquences tests sont ainsi plus longues et plus réalistes. Connaissant la relation quantitative entre la compression et la qualité perçue, les diffuseurs ont les outils pour choisir le compromis entre le coût et la satisfaction du consommateur (plus on comprime, plus on économise de la bande passante et moins onéreuse est la diffusion).

Transmission multiple

Une fois les images et les sons convenablement comprimés, les diffuseurs se chargent de la transmission aux téléspectateurs. Si la plupart des Français reçoivent aujourd'hui la télévision par une antenne «râteau» installée sur leur toit, la transmission par câble et par satellite se développe. Quelles techniques permettent d'émettre plusieurs millions de bits par seconde vers tous les téléspectateurs ?

Qu'elle soit analogique ou numérique, la transmission des signaux utilise des ondes électromagnétiques dont on fait varier dans le temps l'amplitude et la phase, en fonction du signal à transmettre : c'est la modulation.

Lorsque la modulation est analogique, comme c'est le cas avec le codage SECAM aujourd'hui, l'amplitude et la

phase de l'onde modulée varient continuellement. En transmission numérique, en revanche, la suite binaire (de 0 et de 1) qui code les images, les sons, etc. est découpée en séquences successives de p bits (en général 2, 3, 4, 5 ou 6). Il existe 2^p séquences de p bits différentes. Par exemple, les séquences de deux bits sont au nombre de quatre : 00, 01, 10 et 11. À chaque séquence de p bits, on fait correspondre un couple phase/amplitude parmi 2^p valeurs. Sur un oscilloscope, les 2^p valeurs possibles du couple dessinent une «constellation» de points (voir la figure 4). Toutes les T secondes, on émet une onde de phase et d'amplitude données, ce qui correspond à un point de la constellation. La succession des valeurs de ces points contient les données transmises. Plus la période T est faible, plus grand est le débit transmis, mais plus grande est aussi la bande passante nécessaire pour acheminer le signal (on peut montrer que la bande passante est légèrement supérieure à $1/T$).

Les canaux de transmission par satellite bénéficient d'une bande passante relativement confortable, de l'ordre de 36 mégahertz. La principale contrainte technique vient des restrictions de la puissance disponible : à 36 000 kilomètres d'altitude, le satellite ne dispose que de l'énergie produite par ses panneaux solaires, dont la taille est limitée par le coût et par des contraintes techniques. La modulation choisie n'autorise que quatre valeurs du couple amplitude/phase. Toutes les 36 nanosecondes, on transmet ainsi deux bits différents ($4 = 2^2$). Nous verrons que ce choix d'une modulation à deux bits résulte des atténuations du signal lors de la propagation entre le satellite et le sol. Le débit résultant est de l'ordre de 55 mégabits par seconde, ce qui correspond à un débit utile, après correction des erreurs de transmission, de 38 mégabits par seconde.

Les contraintes techniques des réseaux câblés sont inverses : les fréquences disponibles sont rares (il est coûteux d'enterrer des câbles), et les canaux sont limités à huit mégahertz de largeur ; en revanche, la puissance disponible est assez confortable. Pour pallier le manque de bande passante, on adopte une constellation à 64 valeurs : toutes les 145 nanosecondes, on obtient une série de six bits ($64 = 2^6$). On transmet plus de bits par intervalle de temps, car l'atténuation est plus faible que pour un satellite. Le débit

atteint est d'environ 41 mégabits par seconde, ce qui correspond encore, après correction des erreurs, à un débit utile de 38 mégabits par seconde (les débits sont volontairement identiques sur câble et sur satellite pour que le câble puisse transmettre facilement les mêmes données numériques).

Lors de la transmission, les récepteurs captent un signal très atténué qui est amplifié. Plus le signal est amplifié, plus il est altéré. L'introduction de ce bruit dégrade inévitablement toutes les transmissions : c'est lui qui ajoute cette «neige» caractéristique aux images analogiques de la télévision. Lorsque les signaux sont numériques, les points qui représentent le signal dans la constellation deviennent des taches. Dès que le bruit est trop important, les taches se chevauchent, et il devient plus difficile pour le récepteur de décider à coup sûr quel point de la constellation avait été émis. Lorsque la constellation est plus dense, par exemple la constellation à 64 états de la transmission par câble, les points sont plus proches et le chevauchement fréquent engendre des erreurs de transmission. C'est pourquoi, dans les transmissions par satellite où le bruit est important, on a préféré une constellation à quatre états seulement (voir la figure 4).

Pour lutter contre les erreurs de transmission, on doit réduire leur fréquence d'apparition à une erreur pour un milliard de bits transmis, ce qui correspond à quelques erreurs par heure seulement. De puissants mécanismes de correction d'erreurs ont été mis en œuvre à cette

fin : on ajoute des informations supplémentaires aux bits à transmettre, afin qu'un bit mal reçu puisse être «deviné» par un décodeur adéquat.

Le principe de cette correction s'explique simplement sur des messages composés de lettres. Supposons que nous voulions transmettre un message composé de la suite de voyelles : AEIOEAEUEIUEIO. Avec beaucoup de chance, ce message peut parvenir à son destinataire intact, mais, souvent, les défauts de transmission le transformeront en quelque chose comme AEIOEOEIEUEIO, avec deux erreurs en sixième et en huitième position. On donne au destinataire un moyen de détecter les erreurs en y insérant des consonnes, de façon à créer des mots intelligibles. Par exemple, la suite précédente de voyelles se retrouve intégralement dans le message plus long suivant : ATTENTION MESSAGE NUMERIQUE IMPORTANT. Perturbé lors de la transmission, ce message parviendrait par exemple à son destinataire sous la forme ATTELTION MESSOGE NOMERIQUE IMPORTANT, où les voyelles de rangs 6 et 8 sont encore erronées. Toutefois, le destinataire rectifie alors facilement MESSOGE, qui n'existe pas, en le mot voisin MESSAGE, et NOMERIQUE en NUMERIQUE. Après élimination des consonnes, il récupère la suite des voyelles intacte.

La remarquable fiabilité des lecteurs optiques de codes à barres, utilisés par exemple aux caisses de supermarchés, est due à la puissance de ces méthodes de correction d'erreur. Les lecteurs de disques compacts recourent aussi à de

telles techniques de codage pour éviter les erreurs de lecture. Aujourd'hui, ces codes, inventés à l'Institut de technologie du Massachusetts par Irving Reed et Gustave Solomon en 1960, trouvent une nouvelle application dans la transmission de la télévision numérique par les réseaux câblés. Pour la transmission par satellite, intrinsèquement moins fiable, le code choisi pour le câble a été conservé, mais il est renforcé par un deuxième niveau de correction. Enfin, un troisième niveau de correction a même été adjoint pour fiabiliser les liaisons hertziennes terrestres, encore moins fiables.

La transmission hertzienne terrestre cumule en effet les difficultés techniques. D'une part, les fréquences disponibles sont rares, car la majorité d'entre elles sont déjà utilisées pour la télévision analogique. D'autre part, les puissances d'émission sont limitées, en raison du risque de brouillage possible des chaînes de télévision existantes. Enfin, les transmissions terrestres sont perturbées par la réflexion parasite des ondes électromagnétiques sur les collines ou sur les constructions qui entourent l'antenne du récepteur. Les échos sont décalés dans le temps, parce que la lumière se propage à vitesse finie. Les décalages de quelques microsecondes entre l'instant d'arrivée du signal principal et les instants d'arrivée des signaux parasites engendrent les images fantômes que l'on voit parfois sur les téléviseurs analogiques (voir la figure 5).

Pour fiabiliser les transmissions, la télévision numérique terrestre utilisera



3. LA COMPRESSION DES IMAGES ANIMEES par le codage MPEG2. Les images *I*, qui reviennent périodiquement, sont codées sans référence aux autres images. Les images *P* sont prédites à partir d'une image

précédente : on transmet une image fictive qui code la différence entre l'image *P* et l'image *I*. Les images *B* sont reconstruites à partir de la moyenne des deux images *I* ou *P* les plus proches.

une technique de transmission multiporteuse mise au point au CCETT : l'unique onde électromagnétique porteuse est remplacée par plusieurs milliers d'ondes porteuses, portées par des fréquences proches, simultanément transmises. Dans cette transmission parallèle de plusieurs milliers d'ondes porteuses, chaque porteuse transporte moins d'un millième du débit total et peut donc être modulée 1 000 fois plus lentement que si elle était seule. La durée pendant laquelle on transmet un point de la constellation étant égale à quelques centaines de microsecondes au lieu de quelques centaines de nanosecondes pour la modulation monoporteuse classique, on élimine les interférences entre deux émissions successives en les espaçant d'une courte durée destinée à absorber les échos.

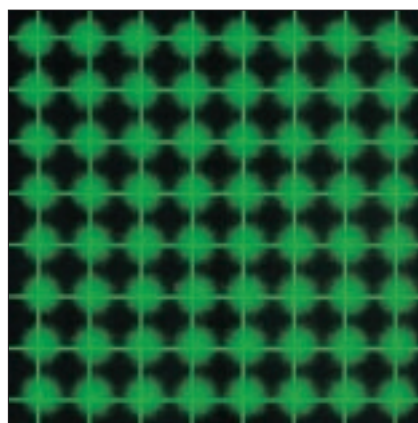
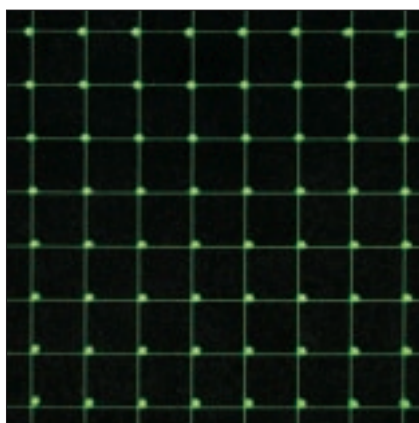
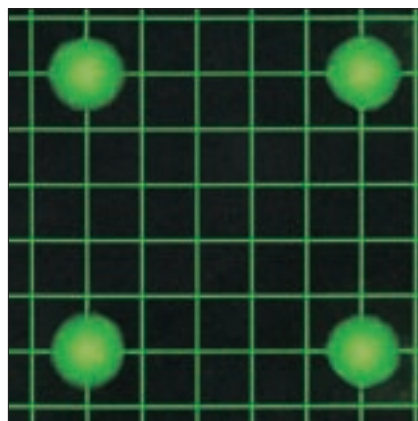
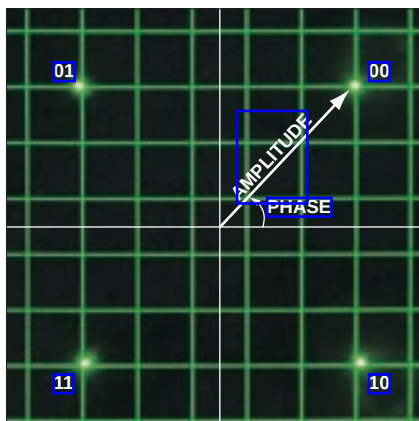
Cette transmission multiporteuse résout un autre problème de la diffu-

sion hertzienne terrestre. La portée d'un émetteur de télévision est de l'ordre de quelques dizaines de kilomètres. Lorsqu'une antenne de réception capte les signaux de deux émetteurs, elle reçoit d'abord le signal émis par l'émetteur le plus proche, et un peu après, le signal émis par l'émetteur le plus éloigné. Ce second signal est identique au premier, mais décalé dans le temps, si bien qu'il est équivalent à un écho du premier signal. Pour éviter que deux émetteurs ne se brouillent mutuellement, on fait émettre les émetteurs voisins à des fréquences porteuses différentes (c'est également le cas pour la transmission radio analogique : lorsqu'on change de région, la fréquence des stations FM change). Les fréquences, si précieuses, sont ainsi gaspillées.

Comme la transmission multipor-

teuse ne se brouillent plus. Un grand territoire peut être couvert avec autant d'émetteurs que nécessaire, tous fonctionnant sur la même fréquence. On gagne ainsi un facteur dix dans l'utilisation de la bande passante. La modulation multiporteuse résiste également aux brouillages involontaires dus à une télévision analogique qui émettrait à proximité : si quelques porteuses numériques sont brouillées, la proportion de l'information perdue reste assez faible pour que le système correcteur d'erreurs la recouvre intégralement.

Malgré ces efforts, chaque canal de télévision terrestre, avec une largeur de 8 mégahertz, ne peut acheminer qu'un débit maximum de 35 mégabits par seconde, ce qui correspond, après correction des erreurs de transmission, à un débit utile de 25 mégabits par seconde. Ce débit reste bien inférieur aux 38 mégabits par seconde du câble ou du satellite : on transmet moins de programmes sur un canal de télévision terrestre que sur un canal du câble ou du satellite. En revanche, contrairement aux téléviseurs qui reçoivent les programmes par câble ou par satellite, les récepteurs terrestres seront librement déplacés tels les téléphones sans fil. On expérimente même des récepteurs à écran plat, alimentés par batterie, et installés à l'arrière des voitures ou dans les transports en commun.



4. LES DONNÉES NUMÉRIQUES qui sont transmises par les diffuseurs sont décomposées en séquences que l'on peut représenter par des points dans un quadrillage. À intervalles de temps réguliers, on module l'amplitude et la phase d'une onde porteuse afin de coder l'un des points de cette « constellation ». Pour la transmission par satellite, la constellation est composée de quatre points (*en haut, à gauche*). Les données sont ainsi transmises par séquences de deux bits. Sur le câble, on transmet un point parmi 64 (*en bas, à gauche*) à chaque période, soit six bits ($2^6=64$). Le parcours du signal sur de grandes distances introduit du bruit : le signal reçu (*à droite*) n'est pas constitué de points, mais de taches, qui, lorsqu'elles se recouvrent introduisent des erreurs dans la transmission. Ainsi, pour les satellites, on choisit une constellation à quatre états : les points sont éloignés et se recouvrent moins facilement que dans la constellation à 64 états.

Tous dans le tuyau

Que ce soit par satellite, par câble ou par diffusion terrestre, les canaux de transmission numérique ont conservé leur largeur en fréquence des temps analogiques. Or, grâce à la compression, la place nécessaire à un programme numérique de qualité identique à celle des programmes analogiques a diminué. La quarantaine de mégabits par seconde d'un canal de satellite ou d'un canal de réseau câblé est partagée entre plusieurs programmes. Ce nombre est compris entre deux et vingt selon que l'on y place des programmes en haute définition ou des programmes de qualité réduite (comme les émissions de téléachat).

Les suites de bits sont regroupées en paquets de plusieurs sortes : des paquets vidéo, audio, de sous-titres, de télétexte, de messages de contrôle d'accès, etc. Ces paquets ne sont pas tous indépendants et l'on doit conserver la synchronisation entre le son et l'image. Avant la transmission, les paquets des

différents composants sont rassemblés en un seul flot : les paquets qui correspondent aux images de *France 2* sont mélangés aux paquets audio de *France 2* (son stéréo, voire plusieurs sons pour un programme multilingue) et aux paquets de sous-titre, etc. Comme ces paquets peuvent être transmis par des moyens asynchrones, voire, dans certains cas, par des itinéraires différents, ils sont munis d'une étiquette qui indique leur rang par rapport aux autres. À la réception, les paquets sont classés selon l'ordre indiqué par les étiquettes et présentés à l'écran dans le bon ordre et au rythme indiqué. Ensuite, les flots des différentes chaînes sont rassemblés en un seul flot global (TF1, *France 2*, *France 3*, etc.).

Finalement, plusieurs de ces flots numériques globaux sont mélangés, soit entre eux, soit avec des signaux analogiques, à des fréquences différentes. Ils sont alors prêts pour la transmission sur le même réseau. On entend ici par «réseau» un ensemble de satellites, un réseau câblé ou un ensemble d'émetteurs terrestres (voir la figure 1).

Accès protégé

Comment limiter l'accès aux programmes payants ? Ce contrôle d'accès est effectué grâce à des suites de nombres totalement imprévisibles à qui-conque n'en connaît pas le principe de construction. Pour cette raison, ces suites sont nommées pseudo-aléatoires. Comment deviner, par exemple, que la suite 2 5 7 2 9 1 0 1 1 2 3 5 8 se poursuit par 3, si l'on ignore que chaque chiffre est le reste de la division par dix de la somme des deux chiffres précédents ? Quand on additionne une telle suite au flot numérique des images, lors de l'émission, on transmet une série de nombres qui, utilisés sans décryptage, engendrent une image incohérente. Des variantes de cette technique permettent de créer des images cryptées qui conservent quelque lisibilité malgré une très mauvaise qualité, afin de susciter l'envie de les obtenir en clair. Toutefois, seule la soustraction, dans un décrypteur associé au téléviseur, de la suite secrète de nombres aux données reçues cryptées recrée les images initiales.

L'appareil de décryptage retrouve la suite secrète s'il connaît les deux valeurs initiales (2 et 5, dans notre exemple) qui constituent la clé instantanée de décryptage. Cette clé est la même pour tous les récepteurs, mais elle diffère pour



5. LES ECHOS apparaissent, en transmission analogique, lorsqu'un récepteur capte le signal de deux émetteurs terrestres proches. Comme ces deux émetteurs ne sont pas à la même distance, les deux signaux reçus sont décalés, ce qui provoque les échos. En télévision numérique, le problème sera résolu par l'utilisation de la transmission multiporteuse.

chaque émission cryptée. Cette clé, dont la durée de vie ne dépasse pas quelques secondes, est transmise sous forme chiffrée dans un message spécial nommé ECM (*Entitlement Checking Message*, soit «message de contrôle du cryptage»), qui contient aussi les références de l'émission (il s'agit, par exemple, de la retransmission d'un match de football). L'appareil de décryptage s'assure d'abord que le message ECM qu'il reçoit n'a pas été falsifié en vérifiant sa signature électronique ; les algorithmes d'authentification des ECM reçus sont secrets. Lorsque la carte à puce introduite dans le récepteur confirme que l'abonnement «football» a bien été payé, la clé instantanée est déchiffrée, et l'image reçue est décryptée. Sinon, le décrypteur refusera de calculer la clé, mais proposera l'achat du match si un crédit suffisant figure sur la carte à puce.

Le flot des ECM est permanent pour autoriser un décryptage rapide en cas de changement de chaîne. Comment les abonnements mémorisés par la carte à puce sont-ils mis à jour ? Grâce à un autre message, l'EMM (*Entitlement Management Message*, soit «message de gestion du cryptage»), également protégé du piratage par signature électronique. Cette fois, les EMM sont particuliers à chaque usager, c'est-à-dire à chaque carte à puce. Ils permettront de faire évoluer à distance les droits de cet usager. Les systèmes concurrents diffèrent par leurs

méthodes – jalousement gardées secrètes – de chiffrement des clés et d'authentification des messages EMM ou ECM reçus. En raison de ces différences, les appareils de décryptage des systèmes concurrents sont différents.

Voilà presque 60 ans, les premières émissions expérimentales diffusées de la tour Eiffel montraient la faisabilité d'une technique – la télévision – dont les répercussions sociales étaient alors inimaginables. Sommes-nous aujourd'hui, avec la télévision numérique, à la veille d'un bouleversement du même type ? Sans doute pas : nous n'avons franchi qu'un pas vers la révolution numérique, qui mariera de façon indissociable séquences vraies et images de synthèse. La révolution numérique se borne pour l'instant à élargir le choix du kiosque audiovisuel : limité hier à quelques quotidiens généralistes, il proposera désormais plusieurs dizaines ou centaines de revues spécialisées sur papier glacé.

Jean-Pierre HENOT, Vincent MICHON et Bertrand SUEUR ont travaillé, à Rennes, au Centre commun d'études de télédiffusion et télécommunication (CCETT) du groupe *France Télécom*.

Nicholas NEGROPONTE, *L'homme numérique*, Éditions Robert Laffont, 1995.
Hervé BENOÎT, *La télévision numérique*, Éditions Dunod, 1998.

La télévision numérique

JAE LIM

Après des années de travail, les États-Unis ont déterminé la norme de télévision numérique qui sera imposée aux chaînes du pays.

Dans quelques mois, la télévision américaine changera autant que lorsque la télévision en couleurs est apparue : en novembre 1998, dans les grandes villes américaines, les chaînes de télévision émettront leurs émissions sous forme numérique, afin de transmettre des images de meilleure définition, et avec un meilleur son, que par le codage analogique actuel. Ce nouveau système de télévision numérique émettra également des données récupérables par les ordinateurs.

De la norme américaine actuelle (NTSC, qui date des années 1950) à la nouvelle norme numérique, la transition a été difficile. Après des années de compétition et de coopération, les professionnels concernés, chaînes de télévision, constructeurs de matériel électronique et universitaires, ont abouti à une norme qui ne rend pas obsolètes les téléviseurs existants. Certains détails ne sont pas encore fixés, notamment les modalités d'introduction de services informatiques dans le domaine de la télévision.

Cependant, la situation internationale reste floue. Pour l'instant, seuls le Canada et la République de Corée ont adopté la norme numérique américaine.

La plupart des difficultés ont été résolues, et les progrès réalisés dans le domaine des communications, dans le traitement du signal et en micro-électronique ont dépoussiéré une technique vieille d'un demi-siècle. Le nouveau système utilisera des canaux situés essentiellement dans la bande UHF (Ultra High Frequency, soit «ultra haute fréquence»), entre 470 et 890 mégahertz (canaux 14 à 83). L'ancien et le nouveau système coexisteront jusqu'en 2006, c'est-à-dire jusqu'au moment où les chaînes de télévision cesseront vraisemblablement d'utiliser les signaux NTSC à la fois dans la bande VHF (Very High Frequency, soit «très haute fréquence»), entre 54 et 216 mégahertz

(canaux 2 à 13), et dans la bande UHF. La Commission américaine des télécommunications (FCC) pourra alors allouer ces canaux vacants à la télévision numérique ou à d'autres services, comme les radiocommunications.

Initialement, la Commission américaine des télécommunications n'imposait de normes techniques qu'aux chaînes de télévision hertziennes, mais son influence s'est ensuite étendue aux opérateurs du câble et du satellite, parce que ceux-ci ont voulu limiter les études et réduire les coûts. Une difficulté est apparue quand certains opérateurs du câble et du satellite ont fait cavalier seul en émettant des signaux numériques, mais pas en haute définition. L'évolution vers la haute définition a duré trois décennies.

À la fin des années 1960, la chaîne japonaise NHK collabora avec les fabricants japonais de composants électroniques pour mettre au point un codage analogique nommé MUSE (Multiple Sub-nyquist Encoding, soit «codage par sous-échantillonnage multiple»). Utilisant cinq fois plus de données qu'auparavant pour produire chaque image avec une meilleure définition, cette diffusion nécessitait également cinq fois plus de bande passante : alors que le codage NTSC utilise un canal de six mégahertz de largeur pour transmettre ses signaux audio et vidéo, le système MUSE nécessitait 30 mégahertz. Toutefois, à l'époque déjà, alors que le spectre électromagnétique était très peu utilisé, la norme MUSE semblait excessivement lourde.

La norme NTSC profitait de l'espace électromagnétique libre ; elle semble aujourd'hui très peu parcimonieuse. Dans la transmission d'images, les intensités des éléments adjacents constituant l'image (les pixels) sont souvent très similaires ou, du moins, dépendent de l'intensité des pixels contigus. Le système NTSC transmet des images

entières sans exploiter cette dépendance, et une grande partie de l'information transmise est redondante. Progressivement, des interférences entre les différents signaux NTSC apparaissent.

Pour y remédier, la Commission américaine des télécommunications décida de ne pas utiliser des bandes du spectre électromagnétique – les canaux tabous – entre les bandes attribuées aux chaînes. Généralement, dans les régions densément peuplées des États-Unis, seul un canal VHF sur deux et un canal UHF sur six sont attribués. Avec l'avènement des téléphones portables et des autres services de télécommunications, le spectre électromagnétique américain est devenu trop petit pour des normes telles que MUSE. En 1988, des industriels et des laboratoires de recherche publics et privés proposèrent des normes de diffusion numérique.

Des approches antagonistes

Alors que l'évaluation technique des différentes normes proposées commençait, la Commission américaine des télécommunications prit une décision déterminante : en mars 1990, elle opta pour une incompatibilité des nouvelles normes avec les récepteurs existants.

Une approche compatible aurait permis aux téléviseurs actuels de former des images de qualité acceptable à partir des signaux numériques de télévision en haute définition. Une telle compatibilité avait été conservée lors du passage à la télévision en couleurs : les téléviseurs noir et blanc pouvaient former une image noir et blanc avec le signal d'une image couleur. Toutefois, lors des études, la Commission américaine jugea que la conservation de la compatibilité des récepteurs nécessitait une augmentation excessive de la largeur de bande, et elle préféra une approche nom-



© 2011 KENNEDY KIMBLE LEBLANC

TELEVISION ANALOGIQUE



1. LES DIFFUSEURS SE PREPARENT au passage à la télévision numérique. Ainsi, la chaîne NBC, à Washington (*en haut*), s'est équipée de caméras de télévision numériques et de systèmes d'enregistrement de qualité supérieure. La télévision numérique peut

TELEVISION NUMERIQUE EN HAUTE DEFINITION



KESSELSHOEVE INSTITUT OF LEUWEN ST

offrir une image très nette, en haute définition, et un rapport de la largeur à la hauteur de l'écran supérieur à celui du système actuel (*comparaison simulée ci-dessus*). Le format 16/9 est proche du format utilisé au cinéma.

mée *Simulcast*. Un signal en *Simulcast* serait diffusé dans son propre canal de six mégahertz de large, indépendamment du signal NTSC. De cette façon, le système de transmission partirait sur des bases neuves.

Comment éviter de pénaliser les possesseurs de téléviseurs à la norme NTSC? La Commission américaine des télécommunications décida d'attribuer un nouveau canal à chaque chaîne qui en ferait la demande, à condition que les deux diffusions coexistent pen-

dant une période de transition. Initialement, la Commission voulait que le même programme soit diffusé simultanément ou avec un léger décalage sur les deux canaux, d'où le nom *Simulcast* ou «diffusion simultanée», mais cette exigence fut ensuite abandonnée. Une fois la télévision en haute définition (TVHD) suffisamment répandue, la diffusion à la norme NTSC cessera. L'espace hertzien dégagé sera alors attribué à des canaux TVHD supplémentaires ou à d'autres services.

L'approche *Simulcast* a de nombreux avantages : elle requiert moins de puissance et interfère moins avec d'autres signaux, notamment avec le signal NTSC, de sorte que les canaux tabous seront utilisables. De surcroît, après la période de transition, le procédé *Simulcast* éliminera les canaux NTSC utilisant le spectre avec un mauvais rendement. Enfin, la suppression des signaux NTSC, qui engendrent de fortes interférences, permet une meilleure utilisation des autres canaux.

L'Europe face aux États-Unis

Dans la lutte d'influence qui s'est engagée pour conquérir les marchés asiatique, australien et Sud-américain, la norme de télévision haute définition américaine a plusieurs handicaps dont l'Europe pourrait profiter.

L'Europe s'est dotée d'une structure technique et décisionnelle unique, le Consortium DVB (pour *Digital Video Broadcast*, soit «diffusion vidéo numérique»), qui rassemble plus de 200 acteurs européens du monde audiovisuel (centres de recherche, fabricants d'équipements professionnels ou de téléviseurs grand public, instances de régulation, sociétés de production de programmes, opérateurs de télécommunications, etc.). Ces experts ont su s'entendre sur un certain nombre de normes qui définissent progressivement le paysage audiovisuel européen du siècle prochain : on trouve ainsi une norme DVB pour le codage des images et des sons, une autre pour la diffusion par satellite, une pour la distribution par câbles, une pour la diffusion terrestre, une pour l'interface des équipements de studios, etc. Tous les aspects de la chaîne de l'image numérique ont ainsi fait l'objet d'une recommandation applicable à l'Europe entière.

Cette démarche globale garantit une cohérence technique pour l'ensemble du monde audiovisuel et évite les solutions nationales ou privées incompatibles. L'Europe ne veut pas réitérer les erreurs des années 1960, quand s'affrontèrent deux normes de qualité identique, mais incompatibles (PAL et SECAM), déclinées de surcroît en de multiples variantes nationales. N'ayant pas su canaliser la révolution numérique, les États-Unis se retrouvent avec différents systèmes incompatibles, tant sur satellites que sur réseaux câblés, et ils n'ont pas cette palette de normes qui donne au concept DVB sa cohérence.

Pour diffuser la télévision en haute définition par le réseau terrestre classique, les États-Unis ont retenu le concept *Simulcast*, après nombre de redéfinitions. Ce principe consiste à promouvoir la haute définition en attribuant gratuitement à chaque diffuseur d'un programme classique un autre canal, à la condition qu'il le remplisse avec un programme numérique en haute définition. Avec l'avènement des écrans plats de grande taille, la télévision en haute définition est techniquement envisageable à moyen terme. Toutefois, en raison du prix élevé des ces écrans, les publicitaires doutent de l'intérêt d'annoncer sur les nouveaux canaux à audience trop restreinte. Qui financera alors la produc-

tion de programmes en haute définition? Pour élargir l'audience de ces programmes, les chaînes ont réclamé des décodeurs numériques bon marché que les téléspectateurs adjoindraient à leur téléviseur classique : l'émission diffusée en haute définition serait alors convertie en une image ordinaire. N'eût-il pas été préférable de transmettre dans le canal alloué plusieurs programmes numériques classiques, plutôt qu'un unique programme en haute définition voué, dans l'immense majorité des cas, à être visualisé comme une image ordinaire?

L'approche européenne est plus pragmatique, car elle utilise toute la flexibilité des techniques numériques et laisse aux diffuseurs le libre choix d'utiliser comme ils l'entendent – et comme le marché les y encourage – la ressource de chaque canal. Avec la norme européenne, on peut transmettre des images en résolution ordinaire et des images en haute résolution selon le contenu esthétique des images. Le système américain diffusera un unique programme en haute définition à raison de 19 mégabits par seconde. En Europe, le système retenu est plus souple puisque le débit sera ajustable, entre 4 et 32 mégabits par seconde.

Comment les images et les sons codés seront-ils transmis? La télévision terrestre américaine a répondu à cette question cruciale (elle détermine la qualité des programmes effectivement reçus par les téléspectateurs) par des solutions techniques classiques, impropres à une réception portable et, encore plus, à une réception mobile. En revanche, la solution multiporteuse des Européens – initialement mise au point pour la réception de programmes audionumériques dans les voitures (DAB) – est d'une souplesse et d'une robustesse telles qu'elle élargit le champ des services futurs possibles. Seule technique connue autorisant la conception de réseaux monofréquences, elle utilise les précieuses ressources spectrales (ces fameuses bandes de fréquences si disputées...) bien plus efficacement que la norme américaine ne le permet.

L'Europe a récemment imposé sa norme GSM de téléphonie numérique mobile ; saura-t-elle rééditer l'exploit et transformer en suprématie mondiale sa supériorité technique dans le domaine de la télévision numérique?

Bertrand SUEUR, Laboratoire diffusion et distribution hertzienne, CCETT, Rennes

Ce choix du *Simulcast*, en 1990, fut un élément déterminant dans le processus de standardisation de la télévision numérique aux États-Unis. Peu après, cinq normes fondées sur *Simulcast* furent proposées. L'une d'elles préconisait un système analogique, tandis que les quatre autres étaient entièrement numériques.

En février 1993, un comité étudia les résultats et conclut que, si les quatre systèmes numériques étaient nettement plus performants que le système analogique, chacun d'entre eux avait ses points forts, si bien qu'il ne pouvait en recommander un en particulier. Les experts préconisèrent de tester à nouveau les quatre normes numériques quand elles auraient été améliorées par leurs concepteurs. Toutefois, les autorités recommandèrent aux auteurs des quatre projets de combiner les meilleures parties de leurs normes, afin de soumettre un seul système à évaluation.

Vers un standard unique

Les quatre concepteurs de systèmes numériques acceptèrent cette idée d'un projet commun. En mai 1993, ils constituèrent un consortium formé de sept organismes et sociétés dont la Société AT&T, la Société Zenith, l'Institut de technologie du Massachusetts, la Société Philips Electronics et la Société française Thomson Électronique. Entre 1993 et 1994, ce consortium choisit les meilleurs éléments techniques des quatre normes et les améliora. Une norme technique fut élaborée.

Pour diffuser des images en haute définition sans compression de données, on doit transmettre un milliard de bits par seconde. Comme le débit d'un canal à six mégahertz est seulement de 20 millions de bits par seconde, le consortium adopta le format de compression nommé MPEG-2 (l'acronyme de *Moving Pictures Experts Group*, soit «groupe d'experts en images animées») : les images ne sont pas transmises intégralement ; dans le format MPEG, seules les modifications des images sont communiquées. Par exemple, lors de la diffusion du journal télévisé, le système MPEG-2 ne transmet pas continuellement les images du plateau ou du fond, qui restent fixes, mais principalement les mouvements du présentateur.

Les signaux numériques comprimés provenant de différentes sources, vidéo, audio ou autres, sont d'abord multiplexés, c'est-à-dire transformés



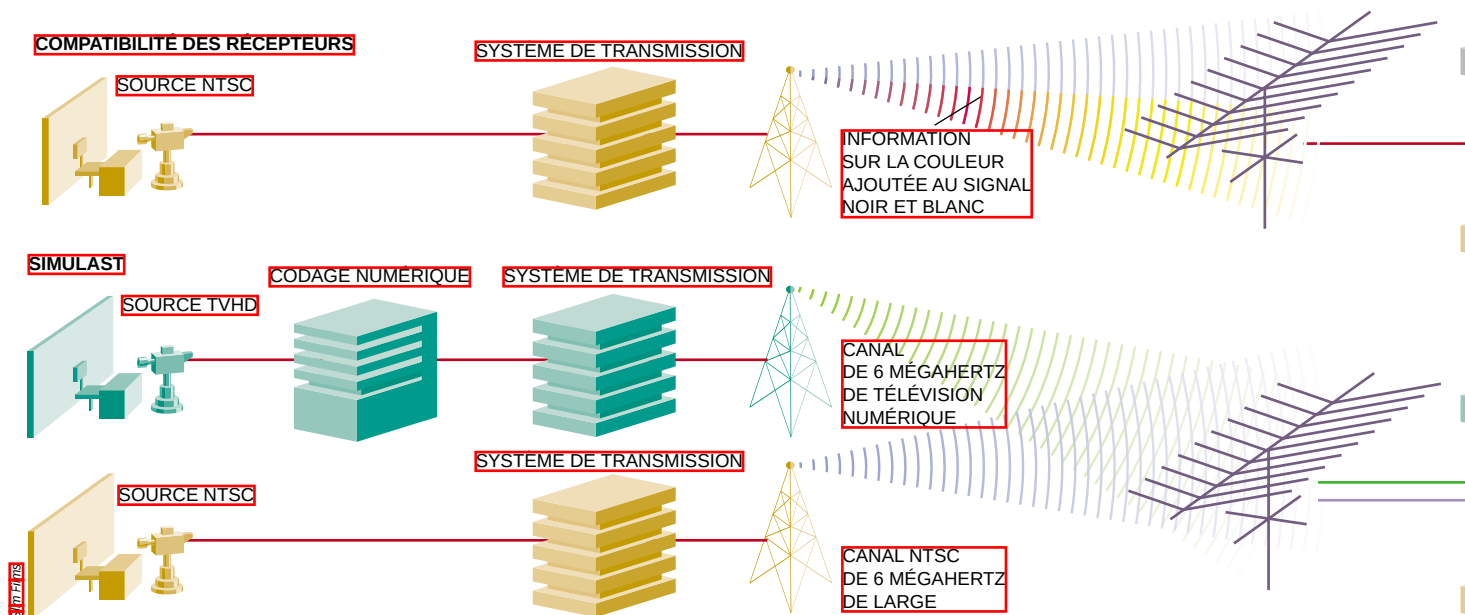
ONDES ET ONDELETTES

B. Burke Hubbard

L'élaboration des ondelettes est une saga moderne. Fourier a inventé la décomposition d'un signal en éléments simples ; pour les besoins de l'analyse du signal, physiciens et mathématiciens ont inventé une nouvelle décomposition en ondelettes. Celle-ci est utile dans de nombreuses applications et ses richesses ne sont pas toutes exploitées.

Barbara Burke Hubbard est écrivain scientifique : elle retrace avec simplicité et chaleur comment les scientifiques pensent et hésitent avant que n'écloie une idée, et comment ils coopèrent. Voir bon de commande p. 98

POUR LA
SCIENCE
DIFFUSION BELIN



2. LA COMPATIBILITÉ DES SYSTÈMES a été abandonnée par la Commission américaine des télécommunications : les téléviseurs actuels ne sauront pas utiliser les nouveaux signaux numériques (TVHD), pour former des images en haute définition. La compatibilité fut conservée lors de l'introduction de la télévision en couleurs.

dont les informations pouvaient être transmises dans un canal de six mégahertz. Comme un seul canal ne peut transmettre à la fois un signal en haute définition et un signal NTSC, la Commission américaine des télécommunications a adopté l'approche *Simulcast*, qui demande aux diffuseurs d'émettre les deux types de signaux.

en une unique séquence de bits. Puis on utilise cette séquence de bits pour modifier l'amplitude d'une onde électromagnétique, nommée porteuse, qui est alors transmise par voie hertzienne. Quand une antenne capte le signal, celui-ci est démodulé par un récepteur, qui soustrait l'intensité de la porteuse et récupère une séquence de chiffres binaires. Cette séquence est démultiplexée (on sépare les données audio, sonores, etc.), et les signaux sont finalement décomprimés.

En novembre 1995, la Commission des télécommunications a adopté la norme MPEG-2 (également retenue en Europe). Dès l'automne prochain, les principaux diffuseurs transmettront des signaux de télévision numérique dans les grandes zones urbaines, notamment à Boston, Los Angeles et Washington.

De surcroît, le nouveau standard est évolutif. Par exemple, la norme MPEG-2 spécifie uniquement la syntaxe des séquences binaires codées et le procédé de décodage. Cette liberté autorise une évolution du codage, sans que le standard soit modifié.

Que faire des téléviseurs actuels, qui ne pourront interpréter le nouveau signal numérique ? Pour bénéficier de l'image numérique en haute définition, on devra soit leur adjoindre un décodeur de signal numérique, soit le changer pour un nouvel appareil. Des téléviseurs numériques seront com-

mercialisés dès l'automne 1998. Un téléviseur numérique à tube cathodique grand écran qui affichera la haute résolution coûtera initialement près de 30 000 francs, mais les prix devraient baisser rapidement.

Réception de l'image numérique

L'adjonction d'un boîtier de décodage à un téléviseur classique coûtera moins de 3 000 francs. Ce boîtier transformera le signal numérique en un signal analogique NTSC. Le spectateur ne profitera alors pas de la haute résolution, mais la qualité de l'image sera légèrement supérieure à celle du même programme diffusé par le canal NTSC analogique.

L'amélioration de la qualité de réception sera également perceptible, même sans téléviseur numérique : dans les limites d'une zone de couverture donnée, le signal vidéo sera parfait, et notamment, exempt des parasites gênants, tels les images fantômes ou le bruit de fond (neige sur l'écran). Hors de la zone de couverture, on ne recevra aucune image. Cette situation du tout ou rien est analogue à ce qui se passe en musique numérique, où la lecture d'un disque compact sale ou endommagé est impossible.

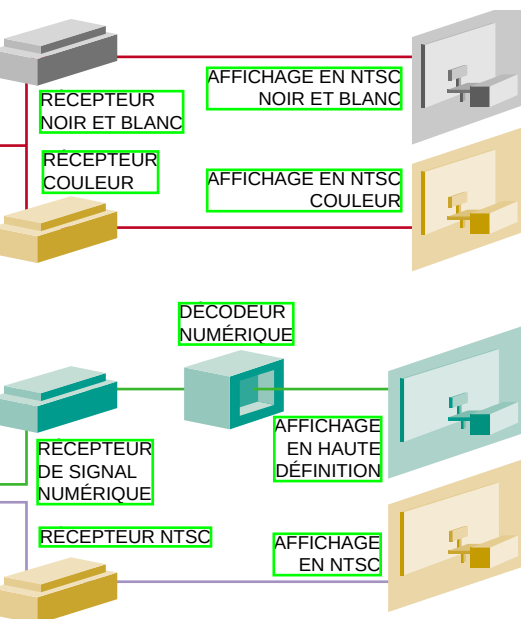
Pour ceux qui disposeront des téléviseurs à la nouvelle norme, l'amé-

lioration de la qualité sera considérable. Avec le système NTSC, on ne perçoit pas le balayage par lignes à partir d'une distance égale à sept fois la hauteur de l'écran. Ainsi, pour un écran de 66 centimètres de hauteur (ou un mètre de diagonale), l'éloignement recommandé est de 4,25 mètres, ce qui rend difficile l'installation d'un téléviseur grand écran dans de nombreux foyers et réduit l'angle de vision à dix degrés environ.

Pour la télévision numérique en haute définition, l'éloignement recommandé sera seulement trois fois la hauteur de l'image : pour le téléviseur décrit précédemment, on pourra se placer à 1,80 mètre. Avec un écran plus grand, on disposera d'un champ de vision plus large et plus réaliste, d'environ 30 degrés.

Avec la nouvelle norme, le rapport de la largeur à la hauteur de l'image est augmenté. Un téléviseur NTSC affiche des images avec un rapport largeur sur hauteur égal à 4/3. Pour se rapprocher du format large du cinéma, la télévision numérique proposera des images au format 16/9 (format déjà disponible en SECAM).

Le son sera de qualité égale à celle des disques compacts. Grâce à des canaux audio additionnels, les spectateurs pourront être «entourés de son», comme dans les salles de cinéma. Ces canaux additionnels pourront égale-



ment transmettre des doublages des films en plusieurs langues.

Le nouveau système numérique entraînera un changement, peut-être plus radical, dans l'utilisation du poste de télévision. Alors que la télévision est aujourd'hui un appareil de loisir, la transmission numérique apportera des données variées, le courrier électronique. Le téléviseur pourra même être utilisé comme visiophone, ou comme moniteur d'ordinateur.

Cette intégration des systèmes a poussé l'industrie informatique à pénétrer dans un domaine dominé jusqu'ici par les chaînes de télévision. Des deux côtés, il reste encore beaucoup de travail, mais la convergence semble inévitable. Cette intégration pourrait finalement modifier l'approche de la télévision bien plus radicalement que l'amélioration de l'image et du son numérique.

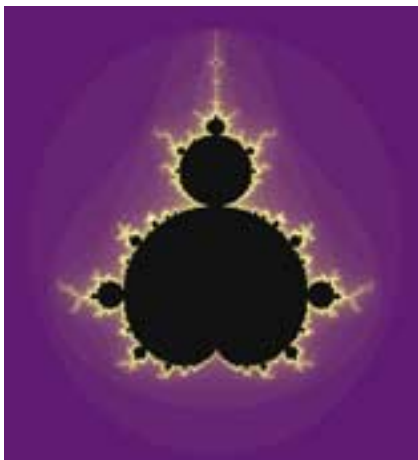
Jae LIM enseigne au département de génie électrique et d'informatique de l'Institut de technologie du Massachusetts.

Jae S. LIM, *Two-Dimensional Signal and Image Processing*, Prentice-Hall, 1990.

A.N. NETRAVALI et B.G. HASKELL, *Digital Pictures : Representation, Compression and Standards*, deuxième édition, Plenum Press, 1995.

HDTV and the New Digital Television, numéro spécial de IEEE Spectrum, vol. 32, n° 4, pp. 34-79, avril 1995.

ATSC Digital Television Standard, Advanced Television Systems Committee, 16 septembre 1995.



Certitudes sans démonstration?

JEAN-PAUL DELAHAYE

Identifier les constantes mathématiques nécessite des tables numériques, de bonnes idées et d'excellentes machines.

Vous devez calculer l'intégrale de $\sin(77x)\sin(5x)/x^2$ entre 0 et l'infini et, grâce à votre logiciel, vous obtenez la valeur M égale à 7,853981634. Cela ne vous semble pas très satisfaisant, et vous soupçonnez que cette intégrale a une expression simple, laquelle vous permettrait des simplifications quand vous la combinez avec d'autres expressions qui entrent dans votre calcul. Cette valeur est peut-être une fraction p/q avec p et q entiers ; ou une fraction multipliée par $\sqrt{2}$ (c'est-à-dire quelque chose de la forme $p\sqrt{2}/q$ avec p et q entiers) ; ou une fraction multipliée par π (quelque chose de la forme $\pi \times p/q$) ; ou une fraction multipliée par e (quelque chose de la forme $e \times p/q$). Comment faire

pour reconnaître M ? Nous verrons que l'inverseur de Simon Plouffe (*voir les encadrés 2 et 4*) répond à la question, et plus encore. Nous allons examiner différentes méthodes d'identification.

ENUMÉRATION ET COMPARAISON

La première idée consiste à calculer tous les nombres p/q , $\sqrt{2} \times p/q$, $\pi \times p/q$, $e \times p/q$, puis à rechercher celui qui se rapproche le plus de M . Il existe une infinité de tels nombres et parmi eux certains sont aussi proches qu'on le veut de M . Abordons donc le problème progressivement. Essayons en supposant d'abord que p et q sont tous deux compris entre 1 et 10.

Il faut 100 essais pour comparer M aux nombres de la forme p/q , 100 autres pour ceux de la forme $\sqrt{2} \times p/q$, etc., soit, au total, 400 essais à faire. Si on ne trouve rien, il faudra essayer en faisant l'hypothèse que p et q sont tous deux des entiers compris entre 1 et 100. Cela nécessitera 40 000 essais. Puis, si nécessaire, on passera à l'intervalle de 1 à 1 000, ce qui conduira à 4 millions de calculs, etc. Le nombre d'essais en fonction de l'intervalle de recherche de 1 à n augmente très rapidement, comme le carré de n .

Dans notre cas, dès la première tentative, avec p et q entre 1 et 10, on découvre que, parmi les 400 nombres essayés pour approcher M , le nombre $5\pi/2$ se trouve situé à un milliardième

1. RAISONNEMENTS HEURISTIQUES

Tous les nombres considérés sont entre 0 et 40. On dispose d'un nombre M qu'on soupçonne être parmi 400 nombres candidats. On compare M aux 400 nombres, on découvre qu'il s'approche de l'un d'eux, $5\pi/2$, de moins 10^{-9} (un milliardième). Ne doit-on pas considérer comme assuré que M vaut $5\pi/2$?

Raisonnement 1

Si M était choisi au hasard entre 0 et 40, alors la probabilité pour qu'il se trouve à un milliardième d'un des 400 nombres essayés serait inférieure à : $400.2.10^{-9} / 40 = 2.10^{-8}$, car la longueur totale de l'intervalle considéré est 40 (les cas possibles) et qu'il y a 400 petits intervalles chacun de longueur 2.10^{-9} correspondant aux cas favorables (les nombres situés à moins de 10^{-9} d'un des 400 nombres essayés), ce qui au total constitue un ensemble d'intervalles d'une longueur cumulée inférieure à $400.2.10^{-9}$. Si M était choisi au hasard, la probabilité de trouver M à moins d'un milliardième d'un des 400 nombres essayés serait donc inférieure à 2.10^{-8} . C'est vraiment très peu, ce n'est donc pas par hasard que cela s'est produit, c'est parce que M vaut $5\pi/2$.

Raisonnement 2 (utilisant la formule de Bayes)

Supposons que nous ayons quelques raisons de considérer que M est l'un des 400 nombres testés (événement noté A) et que plus précisément *a priori* nous évaluons la probabilité de A à $1/10$. La probabilité, *a priori*, que M ne soit pas parmi les 400 nombres (événement B) est de $9/10$. La formule de Bayes indique comment recalculer cette probabilité après l'expérience conduisant à la découverte que " M est à une distance de moins de 10^{-9} d'un des 400 nombres" (événement E).

La nouvelle probabilité pour que M soit l'un des 400 nombres (c'est-à-dire $5\pi/2$) est :

$P_A' = P_A Q_A / (P_A Q_A + P_B Q_B)$ où
 P_A = probabilité *a priori* de A = $1/10$
 P_B = probabilité *a priori* de B = $9/10$

Q_A = probabilité de E sachant A (probabilité que M soit situé à une distance de moins de 10^{-9} d'un des 400 nombres essayés sachant que A est vrai) = 1

Q_B = probabilité de E sachant B (probabilité que M soit situé à une distance de moins de 10^{-9} d'un des 400 nombres essayés sachant que A n'est pas vrai) = $400.2.10^{-9}/40 = 2.10^{-8}$

P_A' = probabilité de A sachant E (probabilité *a fortiori* que A soit l'un de 400 nombres, donc $5\pi/2$) = $P_A Q_A / (P_A Q_A + P_B Q_B) = P_A' = (1/10) / (1/10 + (9/10) \cdot (2.10^{-8})) = 0,99999982$

La probabilité est passée de 10% à 99,999982 %

En partant de $P_A = 1\%$, on arriverait à : $P_A' = 99,99980\%$.

En partant de $P_A = 0,1\%$, on arriverait à : $P_A' = 99,99802\%$.

Donc même si, au départ, on a très peu de raisons de croire que M est parmi les 400 nombres essayés, une fois le calcul fait on en est quasiment certain.

Notons bien qu'aucun des deux raisonnements n'est totalement justifié, car :

- M n'est pas choisi au hasard (raisonnement 1),
- on ne peut donner aucun sens rigoureux à l'évaluation *a priori* $P_A = 1/10$ (raisonnement 2)

Avec les raisonnements heuristiques, on aboutit à des quasi-certitudes, mais pas à des preuves mathématiques.

de M , alors que les 399 autres nombres sont tous situés à une distance de M plus grande que un dixième (le plus proche est $8/1$, situé à $0,15$ de M).

N'a-t-on pas une bonne raison de croire que M est exactement $5\pi/2$ et que ce sont les erreurs de calcul et d'arrondi qui expliquent qu'il n'y a pas égalité parfaite? La question est très délicate et constitue le cœur d'une certaine pratique expérimentale des mathématiques. Que $5\pi/2$ ressemble fort à M n'est bien sûr pas une preuve mathématique que les deux nombres sont égaux. Toutefois, on prend peu de risques en acceptant d'identifier les deux nombres, car on raisonne ainsi. Les raisonnements permettant de justifier cette identification sont présentés dans l'encadré 2.

Une variante de la méthode par énumération et comparaison est de prendre M , $M/\sqrt{2}$, M/π et M/e , et de multiplier ces quatre nombres par $q = 1, 2, 3, \dots, 10$, en regardant à chaque fois si le résultat est un entier (si, par exemple pour q égal à 2, on trouve que $qM/\pi \sim 5$, cela signifie que, vraisemblablement, $M \sim 5\pi/2$).

Si on recherche comme tout à l'heure le nombre M en imposant à p et q d'être entre 1 et 10, cette méthode conduit à faire 43 multiplications (trois au départ, puis 10 pour chaque nombre). Pour p et q entre 1 et 100, il y a 403 multiplications à faire, etc. L'augmentation est plus modérée que précédemment.

Cette seconde méthode est plus efficace : elle aboutit plus rapidement au résultat que vraisemblablement M vaut $5\pi/2$. Bien sûr, on n'obtient, là aussi, qu'une forte présomption.

Une troisième idée consiste à utiliser qu'un nombre est rationnel si et seulement si son développement décimal est périodique à partir d'un certain moment. On pourrait donc tenter de reconnaître la période de M , $M/\sqrt{2}$, M/π et M/e . On serait amené à faire quatre multiplications, puis quatre recherches de périodes, ce qui semble encore meilleur que la seconde idée. Malheureusement, la période du développement d'un nombre p/q peut être de longueur $q-1$. C'est le cas, par exemple, pour p/q avec $0 < p < q$ et $q = 7, 17, 19, 23, 29, 47, 59, 61, 97, 109, 113, 131, 149, 167, 179, 181, 193$ (il s'agit de ce qu'on appelle les entiers premiers longs). Cela signifie que, si $M = 13/109$, on ne reconnaîtra la période que si l'on dispose de plus de 109 chiffres (disons 200), ce qui est difficile : la méthode du repérage de période est mauvaise car elle demande trop de chiffres pour aboutir.

La quatrième idée (qui va nous conduire aux techniques les plus modernes de l'identification des constantes) provient d'une méthode ancienne de recherche d'approximations des nombres par des fractions disposées

sur de multiples étages, appelées fractions continues, dont les propriétés sont explicitées dans l'encadré 3.

LE DÉVELOPPEMENT EN FRACTION CONTINUE

Le développement en fractions continues d'un nombre est une représentation très intéressante de ce nombre (comme le sont les développements décimaux et binaires). Cette représentation fait apparaître sa nature mieux que le développement décimal, car on a les remarquables résultats (a) (b) et (c) indiqués plus bas. Ils ont été prouvés aux XVII^e et XVIII^e siècles par Huygens, Euler et Lagrange.

(a) Un nombre x est rationnel si et seulement si son développement en fractions continues est fini.

(b) Un nombre x est quadratique, c'est-à-dire de la forme $p/q + \sqrt{r/s}$ avec p, q, r et s entiers (racine d'une équation du second degré à coefficients ration-

nels), si et seulement si son développement en fractions continues est périodique à partir d'un certain point.

(c) Certains nombres transcendants (par définition, ils ne sont pas solution d'une équation polynomiale à coefficients entiers) ont un développement en fractions continues qui suit une règle clairement identifiable. Le nombre e qui est transcendant s'écrit ainsi : $e = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, 1, 1, 12, 1, 1, 14, \dots]$

Si on tombe sur une fraction continue infinie et non périodique ayant une règle bien claire (c'est le cas de l'exemple précédent), l'utilisation des résultats (a) et (b) nous donne une très sérieuse raison de croire que le nombre est irrationnel et non quadratique (irrationnel, car ayant une fraction continue infinie ; non quadratique, car ayant une fraction continue non périodique). Le mathématicien qui, par exemple, calculerait les 60 premiers coefficients de la fraction continue de e et verrait tout le long 2 1 2 1 1

2. QUELQUES CONSTANTES MATHÉMATIQUES PARMI LES 75 MILLIONS DE LA TABLE DE SIMON PLOUFFE

Le nombre d'Archimède

π : rapport de la circonférence d'un cercle à son diamètre.

$= 3,14159265358979323846264338327\dots$

Connu avec une précision de 51 539 600 000 de décimales.

Durée du calcul : 29 heures 3 minutes + 37 heures 8 minutes (pour la vérification).

Machine : Hitachi SR2201 avec 1 024 processeurs.

Auteur du calcul : Y. Kanada et D. Takahashi, de l'Université de Tokyo

(calculs menés du 6 au 8 juin 1997 et du 4 au 6 juillet 1997).

La racine carrée de deux

$\sqrt{2} = 1,41421356237309504880168872421\dots$

Connu avec une précision de 137 438 953 444 décimales.

Durée du calcul : 7,5 heures. Machine : Hitachi SR2201 avec 1 024 processeurs.

Auteur du calcul : Y. Kanada et D. Takahashi, de l'Université de Tokyo (3 août 1997).

La constante des logarithmes népériens

$e = 1 + 1/1! + 1/2! + 1/3! + \dots = 2,71828182845904523536028747135\dots$

Connue avec une précision de 50 000 817 décimales.

Durée du calcul : 714 heures. Machine : HP 9000/778 160 Mhz.

Auteur du calcul : Patrick Demichel (30 septembre 1997).

(un calcul de un milliard de décimales aurait été mené à bien par D. et G. Chudnovsky).

Logarithme de deux

$\log(2) = 1 - 1/2 + 1/3 - 1/4 + 1/5 - \dots = 0,693147180559945309417232121458\dots$

Connue avec une précision de 58 486 400 décimales.

Durée du calcul : 53 heures. Machine : SGI R10000.

Auteur du calcul : Xavier Gourdon (décembre 1997).

Le nombre d'or

$F = (1+\sqrt{5})/2 = 1,61803398874989484820458683436\dots$

Connu avec une précision de 10 000 000 de décimales.

Durée du calcul : 29 minutes. Machine : SGI R10000; 194 Mhz.

Auteur du calcul : Simon Plouffe (2 décembre 1996).

La constante d'Euler-Mascheroni

$\gamma = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \log(n) \right) = 0,577215664901\dots$

Connue avec une précision de 7 286 255 décimales.

Durée du calcul : 47 heures 36 minutes. Machine : SGI R10000.

Auteur du calcul : Xavier Gourdon (décembre 1997).

La constante d'Apéry

$\zeta(3) = 1 + 1/2^3 + 1/3^3 + 1/4^3 + 1/5^3 + \dots$

$= 1,20205690315959428539973816151\dots$

Connu avec une précision de 32 000 279 décimales.

Durée du calcul : 35 heures, 21 minutes. Machine : 9x MIPS R10000 180 Mhz.

Auteur du calcul : Sebastian Wedeniwski (17 mars 1998).

3. LES FRACTIONS CONTINUES

Tout nombre x peut s'écrire sous la forme d'une fraction à une multitude d'étages, appelée fraction continue :

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \ddots}}}}$$

$a_0, a_1, a_2, \dots$ sont des nombres entiers.

Voici comment obtenir à la main la fraction continue de $233 / 177$.

On fait une succession de divisions en reprenant pour dividende de la division n le diviseur de la division $n - 1$, et pour diviseur de la division n le reste de la division $n - 1$, ce qui conduit à la disposition suivante :

$$\begin{array}{r} 233 \overline{) 177} \quad 1 \text{ reste } 56 \\ 177 \overline{) 56} \quad 3 \text{ reste } 9 \\ 56 \overline{) 9} \quad 6 \text{ reste } 1 \\ 9 \overline{) 1} \quad 9 \text{ reste } 0 \end{array}$$

La première division s'interprète $233/177 = 1 + 56 / 177 = 1 + 1 / (177 / 56)$.

La seconde donne :

$233/177 = 1 + 1 / (177 / 56) = 1 + 1 / (3 + 9 / 56) = 1 + 1 / (3 + 1 / (56 / 9)) \dots$, etc.

Au total, la fraction continue est : $[1, 3, 6, 4, 2]$

$$\begin{array}{c} 233/177 = 1 + \cfrac{1}{3 + \cfrac{1}{6 + \cfrac{1}{4 + \cfrac{1}{2}}}} \end{array}$$

Un nombre est rationnel si et seulement si sa fraction continue est finie. En effet :

- si la fraction continue est finie, le nombre est rationnel (on réduit la fraction continue de proche en proche jusqu'à l'avoir sous la forme p/q),
- quand on applique la méthode indiquée ci-dessus, le diviseur de l'étape n est le reste de l'étape $n - 1$, donc le diviseur diminue strictement à chaque étape, et il n'y a donc qu'un nombre fini d'étapes.

Si un nombre a un développement en fraction continue périodique à partir d'un certain point, alors il s'écrit sous la forme $p/q + \sqrt{r/s}$ avec p, q, r , et s entiers.

Exemple. Prenons le nombre $x = [2, 1, 1, 1, 4, 1, 1, 1, 4, 1, 1, 1, 4, \dots]$

Considérons le nombre dont le développement est :

$y = [4, 1, 1, 1, 4, 1, 1, 1, 4, 1, 1, 1, 4, \dots]$

Le nombre y vérifie :

$$y = 4 + 1/(1 + 1/(1 + 1/(1 + 1/y))) = 4 + 1/(1 + 1/(1 + y/(y + 1))) = 4 + 1/(1 + (1 + y)/(2y + 1))$$

$$= 4 + (2y + 1)/(3y + 2) = (14y + 9)/(3y + 2)$$

$$3y^2 + 2y = 14y + 9$$

$$3y^2 - 12y - 9 = 0$$

$$y^2 - 4y - 3 = 0$$

$$y_1 = 2 + \sqrt{7}$$

$$y_2 = 2 - \sqrt{7} \text{ (à éliminer car } y > 4)$$

$$\text{d'où } x = y - 2 = \sqrt{7}$$

Pour tout nombre dont le développement en fraction continue est périodique, on peut faire un raisonnement analogue (à chaque fois, après réduction de la fraction, on aboutit à une équation du second degré). Donc tout nombre dont le développement en fraction continue est périodique est un nombre quadratique (de la forme $p/q + \sqrt{r/s}$).

Il n'est pas aussi simple de démontrer la réciproque : si un nombre est de la forme $p/q + \sqrt{r/s}$ avec p, q, r et s entiers, alors son développement est périodique à partir d'un certain point. C'est pourtant ce que Lagrange a établi.

Les fractions continues des racines carrées

$$\sqrt{2} = [1, 2, 2, 2, \dots]$$

$$\sqrt{3} = [1, 1, 2, 1, 2, 1, 2, \dots]$$

$$\sqrt{5} = [1, 4, 4, 4, 4, 4, \dots]$$

$$\sqrt{6} = [2, 2, 4, 2, 4, 2, 4, 2, 4, \dots]$$

$$\sqrt{7} = [2, 1, 1, 1, 4, 1, 1, 1, 4, 1, 1, \dots]$$

$$\sqrt{8} = [2, 1, 4, 1, 4, 1, 4, 1, 4, 1, 4, \dots]$$

$$\sqrt{9} = [3]$$

$$\sqrt{10} = [3, 6, 6, 6, 6, 6, 6, 6, 6, 6, \dots]$$

$$\sqrt{11} = [3, 3, 6, 3, 6, 3, 6, 3, 6, 3, 6, \dots]$$

$$\sqrt{12} = [3, 2, 6, 2, 6, 2, 6, 2, 6, 2, 6, \dots]$$

$$\sqrt{13} = [3, 1, 1, 1, 6, 1, 1, 1, 6, 1, 1, \dots]$$

$$\sqrt{14} = [3, 1, 2, 1, 6, 1, 2, 1, 6, 1, 2, \dots]$$

Le nombre d'or

C'est le nombre dont le développement en fraction continue est le plus simple.

$$\begin{array}{c} (1 + \sqrt{5})/2 = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\ddots}}}} \end{array}$$

$$(1 + \sqrt{5})/2 = [1, 1, 1, 1, 1, 1, 1, 1, \dots]$$

4 1 1 6 1 1 8 1 1 10 1 1 12 jusqu'à 1 1 40 ne pourrait pas croire que cela peut changer plus loin ; il serait donc convaincu que e n'est ni rationnel, ni quadratique... mais n'aurait pas de démonstration.

Cette situation illustre plus nettement encore que notre premier exemple qu'on peut avoir de fortes raisons de croire un certain résultat mathématique, sans en avoir la moindre preuve. En mathématiques, il existe des situations où le bon sens nous commande de croire une certaine vérité que nous sommes pourtant incapable de prouver d'une façon qui satisfasse un mathématicien classique.

L'objectif des mathématiciens expérimentalistes est justement de rechercher des vérités mathématiques par tous les moyens, y compris informatiques, de collecter le plus possible de raisons de croire aux vérités découvertes et, si c'est possible, de les démontrer au sens habituel, qui reste la seule façon d'avoir des certitudes parfaites (personne ne le conteste).

Leur philosophie est qu'il vaut mieux connaître une relation mathématique qu'on ne sait pas prouver, que l'ignorer. Lorsqu'en cryptographie on fabrique des nombres premiers par des méthodes probabilistes (qui ne garantissent pas entièrement que les nombres produits sont premiers) on parle de nombres premiers à usage commercial. Par analogie, on pourrait dire que les mathématiciens expérimentalistes s'intéressent aux vérités mathématiques à usage commercial.

Notons que, contrairement au développement de e ou de racine de 2, le développement en fractions continues du nombre π malheureusement ne présente pas de régularité apparente.

$$\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, 2, 2, 2, 1, 84, 2, 1, 1, 15, 3, 13, 1, 4, 2, \dots]$$

QUATRIÈME MÉTHODE D'IDENTIFICATION

Les fractions continues fournissent la meilleure méthode pour résoudre notre problème d'identification de M : prendre les quatre nombres $M, M/\sqrt{2}, M/\pi$ et M/e et calculer leur développement en fractions continues jusqu'à ce qu'un terme soit presque entier (ce qui signifiera qu'on est très proche d'un nombre rationnel).

Ici en s'arrêtant, soit au dixième terme du développement en fractions continues, soit quand on a un entier à 10^{-5} près, on obtient les résultats suivants.

Pour M : [7, 1, 5, 1, 5, 1, 1, 2, 28, 1,

26] rien n'est trouvé ;

Pour $M/\sqrt{2}$: [5, 1, 1, 4, 6, 9, 1, 2, 1,

2, 5] rien n'est trouvé ;

Pour M/π : [2, 1, 1+1/765855891] ~

$$[2, 1, 1] = 2 + 1/(1 + 1/1) = 5/2 ;$$

Pour M/e [2, 1, 8, 28, 1, 1, 1, 2, 1, 3, 48] rien n'est trouvé.

Le fait numérique que, pour M/π , on ait trouvé, à très peu près, la fraction $5/2$ nous montre à nouveau que vraisemblablement M/π vaut $5/2$, c'est-à-dire que M vaut $5\pi/2$.

Pour s'assurer de la validité de notre méthode (c'est-à-dire pour avoir de meilleures raisons encore de croire que M vaut $5\pi/2$), on calcule non plus 10, mais 20 termes des fractions continues (ce qui correspond à des millions d'essais de constantes) ; on trouve cette fois deux candidats pour M . Le premier est toujours $5\pi/2$, le second est le nombre rationnel : 3926990817/500000000 (simplification de 7853981634/1000000000 par 2).

Entre les deux candidats, le plus simple est sans hésitation possible $5\pi/2$, c'est lui qu'on retiendra donc si l'on est mis en demeure de choisir sans pouvoir faire de calculs ou de preuves supplémentaires. Remarquons que, concernant l'exemple de départ, les trois méthodes essayées aboutissent au même résultat (ce qui semble le rendre encore plus probable). Un calcul direct de l'intégrale à l'origine du problème est d'ailleurs possible et confirme que l'identification $5\pi/2$ est bonne.

Considérons un autre exemple pour illustrer l'idée des fractions continues : $M = 4,222012200165$.

Pour M : [4, 4, 1, 1, 58, 4, 2, 2, 37, 3, 1] ; pour $M/\sqrt{2}$: [2, 1, 67, 1, 1, 3, 1, 14, 1, 26, 5] ; pour M/π : [1, 2, 1, 9, 1, 5, 3, 1, 3, 4, 1] ; pour M/e : [1, 1, 1, 4, 4, 1, 1/829123] ~ [1, 1, 1, 4, 4, 1] = 73/47. Cela suggère fortement que $M = 73e/47$ (qui est effectivement l'expression utilisée pour définir M).

La méthode des fractions continues va droit au but. Aussi est-elle beaucoup plus puissante que les méthodes envisagées au début, car les nombres quadratiques ont des développements périodiques, et certains nombres transcendants ont des développements en fractions continues simples qu'un œil attentif peut identifier.

L'INVERSEUR DE SIMON PLOUFFE

Aujourd'hui il existe une méthode supplémentaire, bien plus puissante encore et plus commode que celles décrites ci-dessus. Cette

méthode est l'inverseur de Simon Plouffe, qui est à la fois un immense dictionnaire de constantes mathématiques et une machine à rechercher des expressions simples pour les constantes numériques qu'on lui soumet.

Alors que d'habitude on donne une expression à une machine pour qu'elle la calcule, l'inverseur, lui, retrouve l'expression à partir du résultat, d'où son nom.

Lorsque vous avez un nombre dont vous connaissez quelques chiffres, vous pouvez le soumettre à l'inverseur de Plouffe en ouvrant la page internet de son site : <http://www.lacim.uqam.ca/pi/>. Vous pouvez aussi lui soumettre une question par courrier électronique à : numbers@math.uqam.ca. Exemple de question : «lookup 1.4142». L'inverseur, après quelques secondes de calcul, vous renverra sa réponse par courrier électronique.

Si on lui propose une constante M , l'inverseur de Plouffe l'examine en mettant en œuvre progressivement des méthodes de recherche de plus en plus complexes et puissantes.

Dans un premier temps, il se contente de comparer votre M à une liste de

constantes mathématiques qui sont mémorisées dans une grande table. Ensuite, il utilise une série d'algorithmes (dont celui des fractions continues) pour rechercher des combinaisons de ces constantes ou des expressions remarquables vérifiées par le nombre étudié.

TABLEAU DE MENDELEÏEV

La grande table de constantes, noyau de l'inverseur de Simon Plouffe, peut être vue comme un tableau de Mendeleïev des constantes mathématiques. Elle tente de réunir tous les nombres réels simples rencontrés en mathématiques, quel qu'en soit le domaine. La table ne contient pas les constantes physiques qui souvent dépendent de choix arbitraires dans les unités de mesure (cela n'empêche pas que de nombreux physiciens utilisent l'inverseur pour mener leurs travaux théoriques).

Les nombres de la table sont stockés avec la plus grande précision possible (tous ne sont pas connus avec le même nombre de décimales, certains sont d'ailleurs très difficiles à calculer).

On trouvera bien sûr π , $\sqrt{2}$, e , mais aussi : $\log(3)$, $\exp(\pi)$, le nombre d'or, le nombre de Champernowne 0,123456789101112..., la constante d'Artin (qui indique la proportion de nombres premiers longs), la constante de Feigenbaum liée à la théorie du chaos, la somme des $1/n^n$, et mille autres choses. Non, pas mille, bien plus : aujourd'hui, la table contient 75 millions de constantes.

Cela semble beaucoup, pourtant cela résulte d'un choix : il serait trop facile de mettre $\log(2)$, $\log(3)$, $\log(4)$ jusqu'à $\log(1000000000)$. L'intérêt de la table provient du fait que tout ce qui y est placé est assez simple. On n'y met pas π^{435256} , pas plus que $(\log(463052)\sqrt{2543})^{12}$.

Les calculs des constantes de la table sont vérifiés et complétés en permanence. Des centres de recherche au Canada, aux États Unis ou en France prêtent leurs machines pour ces calculs. Simon Plouffe évalue que plus de deux ans de temps de calculs d'un processeur rapide sont ainsi matérialisés dans sa table. Le volume de la table est de

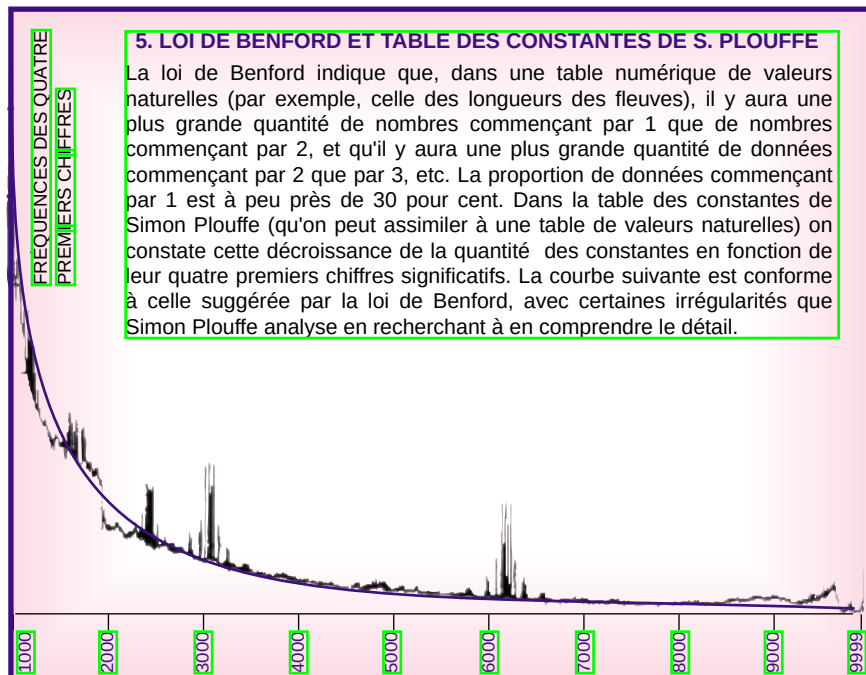
4. L'INVERSEUR DE SIMON PLOUFFE

Quand on pose à l'inverseur de Plouffe la question : "Quelle est la constante : 7,853981634 ?" on obtient un extrait de la table. La réponse correcte ($5\pi/2$) est une constante tellement basique qu'elle y est représentée plusieurs fois sous des codages différents. On reconnaît $\pi/4$ qui possède les mêmes chiffres que $5\pi/2$ puisque $5\pi/2 = 10\pi/4$

7853977541646791 = (p212) S10(1,2)(8,-8,8,2,-9)
7853977611656700 = (a261) $-3-2x+3x^2+3x^3-2x^4-6x^5$
7853978825673085 = (q399) 5/63662
7853979266757639 = (v131) $\sum(1/(8^n-8^{n+52}), n=1..inf)$
7853979538716315 = (m001) $(\exp(-1/2\pi i) \log(\gamma^n) \ln(1+sr(2))) / \ln(1+sr(2))$
7853980101855253 = (a311) $\cos(\pi/158) \sin(\pi/1759)$
7853980314963090 = (a215) Discriminant of a Real Cubic Field
7853980378630278 = (a334) $21^{1/2} (12^{1/4} - 19)$
7853980996271272 = (a215) Discriminant of a Real Cubic Field
7853981032645112 = (v131) $\sum(1/(5/2^n + 145/2^{n-44}), n=1..inf)$
7853981126778084 = (m001) $(1/2 - \text{Backhouse} * \text{GAM}(5/6)) / \text{Backhouse}$
7853981633974126 = (v131) $\sum(1/(4^n - 4^{n+101}), n=1..inf)$
7853981633974483 = (a007) $\pi/4$
7853981633974483 = (m001) $\pi \cos(\pi/12) \sin(\pi/12)$
7853981633974483 = (m405) $1/4 \cdot sr(\pi/4)$
7853981633974483 = (p199) $S(1,1)[1,1,1,0,-1,1,-1,0]$
7853981633974483 = (s207) A003881 Concatenated sequence from EIS
7853981689970672 = (v003) $\sum(1/(2^n n + (29/2^n n^2 + 31/2^n n - 12)), n=1..inf)$
7853981829182755 = (f259) $F(7/12, 9/10; 1/10, 2/9, 4/7; 1)$
7853982300884955 = (q399) 355/452
7853982572371080 = (m001) $\text{BesK}(0,1) + \cos(1) \cdot \text{GAM}(13/24)$
7853982658543988 = (f259) $F(7/12, 5/8; 1/7, 4/5, 2/3; 1)$
7853982819699420 = (p212) S10(1,5)(6,5,0,-3,-2)
7853983015535401 = (v003) $\sum(1/(3^n n + (17/2^n n^2 - 25/2^n n + 28)), n=1..inf)$
7853983588391700 = (m001) $\text{BesJ}(0,1) / \text{BesK}(0,1) \cdot \exp(-\pi)$
7853983588391700 = (m001) $\text{BesJ}(0,1) / \exp(\pi) / \text{BesK}(0,1)$
7853983685097973 = (m001) $\exp(1/2) \cdot \text{Zeta}(1,2) \cdot \text{GAM}(1/12)$
7853983963809479 = (p210) $\sqrt[3]{3} \cdot T_2(0,1,-1,-1,1,1,-1,0,0,-1,-1)$
7853984493158376 = (v015) $\sum((19/6^n n^3 - 19^n n^2 + 353/6^n n - 28) \cdot n! / n^n, n=1..inf)$
7853984748948144 = (q404) $-\text{Psi}(2/15) + \text{Psi}(9/14) - \text{Psi}(13/20)$
7853984752358357 = (m001) $(sr(2) + \text{GAM}(11/12)) \cdot \text{BesI}(0,2)$
7853984917319985 = (v131) $\sum(1/(113/2^n n^2 - 173/2^n n + 47), n=1..inf)$

5. LOI DE BENFORD ET TABLE DES CONSTANTES DE S. PLOUFFE

La loi de Benford indique que, dans une table numérique de valeurs naturelles (par exemple, celle des longueurs des fleuves), il y aura une plus grande quantité de nombres commençant par 1 que de nombres commençant par 2, et qu'il y aura une plus grande quantité de données commençant par 2 que par 3, etc. La proportion de données commençant par 1 est à peu près de 30 pour cent. Dans la table des constantes de Simon Plouffe (qu'on peut assimiler à une table de valeurs naturelles) on constate cette décroissance de la quantité des constantes en fonction de leur quatre premiers chiffres significatifs. La courbe suivante est conforme à celle suggérée par la loi de Benford, avec certaines irrégularités que Simon Plouffe analyse en recherchant à en comprendre le détail.



reconnut que c'était $\pi^2/6$, ce qui ensuite le guida pour trouver une démonstration du résultat qu'il finit effectivement par obtenir. Gauss, autre géant des mathématiques, précisait aussi que sa façon d'arriver à la vérité mathématique était l'expérimentation systématique.

Aujourd'hui, les moyens informatiques suggèrent de reprendre cette conception empirique des mathématiques. Le développement systématique de cette approche iconoclaste ne fait aucun doute, d'autant que les systèmes de calcul formel (comme *Mathematica* ou *Maple*, qui ne manipulent plus seulement des nombres, mais aussi des formules mathématiques) étendent le champ de l'expérimentation mathématique, et que les démonstrateurs automatiques de théorèmes mécanisent parfois une part du raisonnement, permettant à l'approche expérimentale d'aller encore plus loin que le calcul numérique.

Le monde est mathématique, et certains agencements de ses composants peuvent aussi faire des mathématiques et révéler aux humains des parties de la réalité abstraite qu'ils ne verraient pas sans leur aide.

Les cerveaux humains ne sont plus les seuls microscopes mathématiques. Les vérités mathématiques commerciales (presque certaines, mais non démontrées) produites par les nouveaux outils d'explorations mathématiques seront sans doute de plus en plus nombreuses.

Jean-Paul DELAHAYE est professeur d'informatique à l'Université de Lille. La société mathématique de France lui a décerné le prix d'Alembert pour son livre, *Le fascinant nombre π* .

M. BORWEIN et P. B. BORWEIN, *A Dictionary of Real Numbers*; Wadsworth, 1990. (Un dictionnaire de constantes, aujourd'hui largement dépassé par la base de nombres de l'inverseur de Simon Plouffe).

C. BREZINSKI, *History of Continued Fractions and Padé Approximants*. Springer-Verlag, 1991.

J.-P. DELAHAYE, *Le fascinant nombre π* , Édition Pour La Science/Belin, 1997. (Pour des détails sur la nouvelle formule de calcul de π découverte par Simon Plouffe.)

G. H. HARDY et E. M. WRIGHT, *An Introduction to the Theory of Numbers*, Oxford Science Publications, Clarendon Press, Oxford, First Edition 1938, Fifth Edition 1979. (Pour un exposé mathématique sur les fractions continues.)

F. LE LIONNAIS, *Les nombres remarquables*, Hermann, Paris, 1983. Un des tout premiers dictionnaires de nombres (mélangeant nombres réels et nombres entiers).

plusieurs milliards de chiffres, en augmentation constante. Sachant que, dans un livre moyen on met à peu près un million de chiffres on peut en déduire que, si on l'imprimait, l'encyclopédie de Simon Plouffe comporterait plusieurs milliers de tomes. Seule l'informatique permet la consultation et la mise à jour de telles bases de données! Simon Plouffe envisage que, dans l'avenir, sa table pourrait contenir jusqu'à un milliard de constantes de base! (ce qui ferait alors plusieurs centaines de milliards de chiffres).

Si la constante que vous avez indiquée à l'inverseur de Plouffe n'est pas dans sa table, une batterie d'algorithmes se met en marche pour reconnaître votre nombre. Dans un premier temps la recherche de combinaisons simples des constantes les plus fréquentes est tentée. Cette phase est l'équivalent de l'exploration d'une table virtuelle de plusieurs centaines de milliards de constantes. Puis des méthodes de décomposition en fractions continues, en fractions égyptiennes, en produits infinis, etc., sont lancées, ainsi que d'autres méthodes qui généralisent les fractions continues comme l'algorithme PLSQ, dont l'objectif est de trouver des relations complexes impliquant plusieurs constantes. Cet algorithme a déjà permis plusieurs découvertes mathématiques remarquables, dont celle, il y a trois ans, par Simon Plouffe d'une formule nouvelle pour π , qui a conduit à des méthodes de calcul des chiffres binaires indépendamment les uns des autres, et qui est à l'origine du record de plongée dans π par le fran-

çais Fabrice Bellard en octobre 1997. Ce record nous apprend que le mille milliardième chiffre binaire de π est un 1 suivi de 000011111.

En mars 1998, Robert Munafo, à Malden, dans le Massachusetts, a mené un calcul approché du centre de gravité de l'ensemble de Mandelbrot, représenté en vignette à côté du titre de cet article. L'inverseur de Simon Plouffe a été utilisé, et il a proposé la valeur $(\log(3) - 1/3)$ Feig1 où Feig1 est la première constante de Feigenbaum (4,669201609102...). Les constantes des Feigenbaum sont des constantes qu'on rencontre en théorie du chaos et dans l'étude des fractales. Il apparaît donc très vraisemblable que la formule trouvée pour le centre de gravité est juste. Pour en être plus sûr, il faudrait disposer de quelques chiffres supplémentaires dans l'évaluation du centre de gravité. Hélas, ce centre de gravité est très coûteux à évaluer et l'on considère qu'il faudra une dizaine d'années pour disposer de machines donnant un chiffre significatif de plus.

L'EXPÉRIMENTATION MATHÉMATIQUE

Les mathématiques expérimentales qui sont à la base de la conception de l'inverseur de Simon Plouffe ne sont pas totalement nouvelles. L'exemple historique le plus remarquable de cette approche est celui de la série :

$$1 + 1/4 + 1/9 + 1/16 + \dots$$

qui intriguait les mathématiciens au XVIII^e siècle. Euler (en compétition sur ce problème avec les Bernoulli), après avoir évalué grossièrement cette somme,



Quelle coïncidence !

IAN STEWART

Les coïncidences sont parfois plus probables qu'il n'y paraît.

Voilà quelques années, un de mes amis passait sa lune de miel dans un coin perdu d'Irlande. Sa femme et lui marchaient sur une plage déserte lorsqu'ils virent s'approcher deux personnes... qui n'étaient autres que le patron de mon ami et sa femme. Aucun des couples n'avait eu vent des projets de l'autre : leur rencontre était une pure coïncidence.

Si nous sommes si fascinés par les coïncidences, c'est parce que nous avons une piètre intuition des probabilités qui s'y rapportent. Examinons l'une des coïncidences les plus communes : celle des dates d'anniversaire. Combien de personnes doit-on réunir dans une pièce pour avoir une probabilité supérieure à un demi que deux d'entre elles au moins aient la même date de naissance ? Pour simplifier, supposons qu'il n'y ait pas d'années bissextiles, de sorte qu'il n'y ait que 365 dates d'anniversaire possibles. Supposons également chaque date équiprobable, ce qui n'est pas tout à fait exact : il naît plus d'enfants à certaines époques de l'année qu'à d'autres (on pourrait tenir compte de ces facteurs, mais cela compliquerait le calcul sans modifier les conclusions).

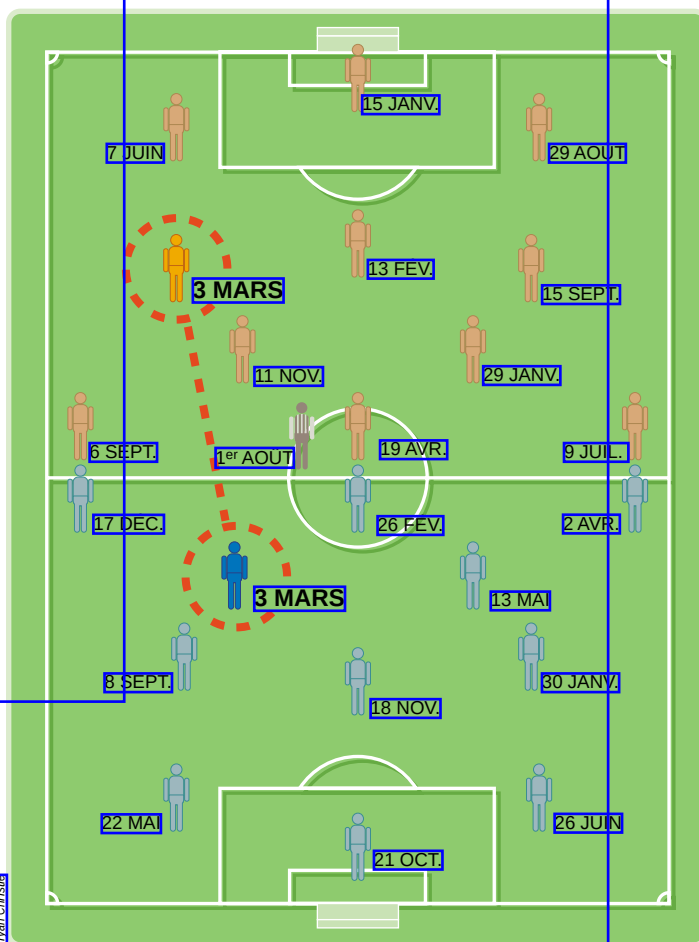
Alors, combien de personnes faut-il dans la pièce ? Cent ? Deux cents ? Mille ? Lorsque des chercheurs posèrent cette question à des étudiants d'Université, la moyenne de leurs estimations fut 385. Valeur excessive à l'évidence, puisque, dès qu'il y a 366 personnes,

une coïncidence au moins est garantie d'après le principe des tiroirs (le même, qui stipule que si l'on place trois chemises dans les tiroirs d'un meuble à deux tiroirs, un des tiroirs contient nécessairement plus d'une chemise ; ici les dates sont l'équivalent des tiroirs, et les personnes correspondent aux chemises). En fait, la réponse correcte est bien inférieure : seulement 23 personnes.

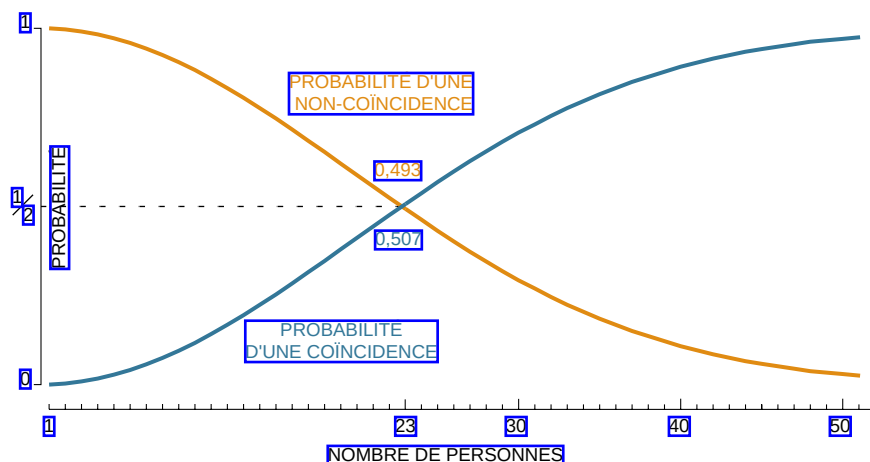
Pour ce calcul, il est plus aisé de déterminer la probabilité qu'un évé-

ment ne se produise pas. En soustrayant ce nombre à un, on obtient la probabilité que l'événement se produise. Par exemple, quand l'événement « Deux personnes au moins ont la même date d'anniversaire » n'est-il pas réalisé ? Lorsque toutes les dates sont différentes. Supposons qu'il y ait d'abord une personne dans la pièce et introduisons les autres personnes une à une. Nous pouvons calculer la probabilité qu'une nouvelle personne ait une date d'anniversaire différente de toutes celles des personnes déjà entrées. Ces probabilités décroissent constamment, et celles d'une coïncidence croissent en conséquence. Or, un événement est plus vraisemblable que son contraire si et seulement si sa probabilité est supérieure à 1/2. Dès que la probabilité de dates toutes différentes devient inférieure à 1/2, nous savons que l'existence d'une coïncidence au moins est plus probable que son contraire.

Avec une seule personne, que nous nommerons Arthur, pas de coïncidence possible : la probabilité cherchée est donc égale à un. Voici qu'entre Élise. Son anniversaire peut être l'un des 365 jours de l'année, mais, comme Arthur est né l'un de ces jours, la probabilité d'avoir des dates distinctes est égale à $364/365$ (le quotient du nombre de cas favorables et du nombre de cas possibles). Puis arrive Pierre. Il reste 363 dates non prises, d'où la probabilité d'avoir une date distincte des deux autres, qui est de $363/365$. La probabilité combinée d'avoir trois dates distinctes est égale au produit $(364/365)$



1. Lors d'un match de football hypothétique, la probabilité que deux des personnes présentes aient leur anniversaire le même jour est supérieure à 1/2.



2. La probabilité que deux dates de naissance soient identiques (au moins) est supérieure à 1/2 dès que 23 personnes sont présentes. Elle augmente rapidement et atteint 0,9 avec 40 personnes.

$\times (363/365)$ (de la même façon que la probabilité d'obtenir deux face de suite, à pile ou face, est égale à $1/2 \times 1/2$).

Nous voyons apparaître la formule générale. Lorsque Catherine entre dans la pièce, la probabilité d'avoir quatre dates distinctes est égale à :

$$(364/365) \times (363/365) \times (362/365).$$

Plus généralement, lorsqu'il y a n personnes dans la pièce, la probabilité d'avoir n dates différentes est :

$$(364/365) \times (363/365) \times \dots \times ((365 - n + 1)/365).$$

Il nous suffit donc de calculer les valeurs successives de cette expression pour savoir à partir de quel n elle devient inférieure à $1/2$. Vingt-deux personnes ont une probabilité de 0,524 d'avoir des dates de naissance toutes distinctes, et ce résultat descend à 0,493 pour 23 personnes. Donc, dès que 23 personnes sont rassemblées, la probabilité d'une coïncidence au moins est $1 - 0,493$, soit 0,507 : l'événement est légèrement plus probable que son contraire.

Testez ce résultat dans des rassemblements de plus de 23 personnes en prenant des paris : à la longue, vous gagnerez. Plus il y aura de monde, plus vite vous gagnerez. La plupart des participants jugeront qu'une coïncidence est improbable, parce qu'ils se focaliseront sur un aspect trompeur du problème : le nombre de personnes dans la pièce. Bien que 23 soit petit, il existe pour 23 personnes, 253 manières de les associer par paires (pour n , il y en a $n(n-1)/2$). Ce grand nombre est révélateur de la probabilité d'une coïncidence.

Robert Matthews, un journaliste britannique et mathématicien, et Fiona Stones ont vérifié ce résultat d'une manière expérimentale. Dans un match de football, il y a 23 personnes sur le

terrain : les deux équipes de 11 joueurs et l'arbitre. Donc, l'événement «deux joueurs au moins ont la même date d'anniversaire» est plus probable que son contraire. Ils se sont intéressés aux matchs de première division du Royaume-Uni joués le 19 avril 1997. Sur dix rencontres, six étaient avec coïncidences, et quatre sans.

En fait, dans deux des matchs, il y avait deux coïncidences ! Dans celui opposant Liverpool à Manchester United, deux joueurs étaient nés le 21 janvier et deux le 1^{er} août ; entre Chelsea et Leicester City les coïncidences étaient le 1^{er} novembre et le 22 décembre. Cette double coïncidence est également prédite par la théorie des probabilités. Elle évalue la probabilité d'avoir deux coïncidences parmi 23 personnes à 0,111... : l'événement se produit donc en moyenne tous les neuf matchs de football. La probabilité d'avoir trois coïncidences parmi 23 personnes est égale 0,018..., et celle d'avoir une coïncidence triple (trois personnes parmi les 23 ont la même date de naissance) est égale 0,007... : on l'observe une fois sur 143 rencontres en moyenne. Amusez-vous à vérifier ces valeurs avec les matchs de la Coupe du monde.

INTUITION, INTUITION...

Une question légèrement différente à présent : combien de personnes en plus de vous-même doivent être dans la pièce pour que l'événement «une personne au moins a votre date de naissance» soit plus probable que son contraire ? On pourrait croire que la réponse est $(364/2) + 1$, soit 183, puisqu'il y a 364 dates différentes de la vôtre et que, si plus de la moitié de ce nombre de personnes se trouve dans la salle, vous avez plus d'une chance sur deux

de partager votre date de naissance avec l'une d'elles. Mais la réponse correcte est 253.

Pour obtenir ce nombre, reprenons la technique déjà utilisée : on cherche la probabilité que les dates d'anniversaire soient différentes de la vôtre, puis l'on soustrait ce nombre à un. Vous êtes déjà dans la pièce et les autres personnes se présentent l'une après l'autre : Arthur, Élise, Pierre, Catherine, etc. La probabilité qu'Arthur ait une date de naissance différente de la vôtre est $364/365$. La probabilité qu'Élise ait une date de naissance différente de la vôtre est encore $364/365$. Il en est de même pour tous les autres. Nous nous désintéressons ici des coïncidences entre dates de naissance de personnes autres que vous, par exemple si Arthur et Élise sont nés le 3 avril. Seul compte si une date de naissance est la même que la vôtre, de sorte que, après l'entrée de n personnes, la probabilité pour qu'elles aient toutes des dates différentes de la vôtre est égale à $(364/365)^n$. La première valeur de n pour laquelle cette expression est inférieure à $1/2$ est 253 ($(364/365)^{253} = 0,499...$).

Incidemment, le fait que cette réponse soit égale au nombre de paires dans le premier problème (avec 23 personnes) semble n'avoir aucune signification mathématique. Une coïncidence donc !

Que nous enseignent ces calculs de probabilité ? D'abord ils nous engagent à ne pas être impressionné outre mesure par des événements qui semblent improbables : ils ne le sont peut-être pas. Une paire de footballeurs ayant la même date d'anniversaire seront probablement surpris de la coïncidence, mais tout match a plus d'une chance sur deux d'en posséder une. Les joueurs se souviendront probablement de cette coïncidence pendant des années, mais les 252 autres paires de joueurs ne s'étonneront pas que leurs dates de naissance ne coïncident pas, car nous remarquons les coïncidences et ignorons les non-coïncidences, et faisons de la sorte paraître les premières plus importantes qu'elles ne le sont en réalité.

La rencontre de mon ami durant sa lune de miel paraît moins surprenante si l'on songe au nombre de couples qu'il a rencontrés durant sa vie et qui n'étaient pas son patron et son épouse.

Robert MATTHEWS et Fiona STONES, *Teaching Statistics*, printemps 1998.

Martin GARDNER, *La magie des paradoxes*, Bibliothèque Pour la Science, 1997.

Henri BROCH, *Le paranormal*, Seuil (Poche), 1989.

Rater n'est pas fausser

Une pièce de monnaie mal frappée, anormalement lourde ou légère, ou encore mal sertie si elle a plusieurs éléments, est probablement fausse. Aujourd'hui, la valeur nominale d'une pièce est supérieure à celle du métal qu'elle contient : son cours est garanti par l'Etat. Ainsi, le coût de fabrication d'une pièce de dix francs est aujourd'hui de 0,996 franc seulement, celui d'une pièce d'un franc de 0,742 franc. Paradoxalement, celui d'une pièce de 50 centimes est de 0,551 franc. Comme pour les billets de banque, seule la qualité de fabrication d'une pièce en certifie la provenance, donc la valeur.

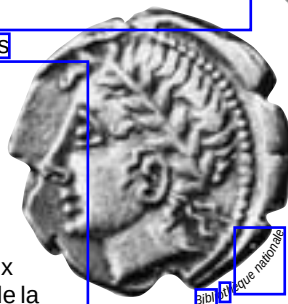
La pièce dont les deux faces sont présentées ici en couleur, un statère averne (mal) frappé en Gaule au II^e siècle avant notre ère, n'est pas fausse : elle a circulé comme celle, mieux fabriquée, qui est présentée en noir et blanc, et elle avait la même valeur. Les motifs portés par ces pièces, une tête ornée d'une couronne de laurier et un cheval accompagné d'une amphore, imitent ceux du statère macédonien frappé par Philippe II, père d'Alexandre le Grand, et utilisé dans tout le monde méditerranéen aux III^e et II^e siècles avant notre ère. Plus généralement, pendant toute l'Antiquité, les pièces mal frappées, très nombreuses, valaient autant que les autres : leur cours était celui du métal qu'elles contenaient.

Comment fabriquait-on les pièces ? On coulait d'abord des rondelles de métal de la taille souhaitée, nommées flans, dans des moules en terre. Puis l'on déposait chaque flan, chauffé au rouge, sur une matrice enchâssée dans une enclume et qui portait en creux l'empreinte de l'une des faces de la pièce. Le monnayeur posait sur le dessus un tampon portant, en creux, l'autre face, et frappait d'un grand coup de marteau.

Le flan n'était pas toujours posé exactement au centre de la matrice, ni le tampon au centre du flan : l'empreinte s'imprimait alors seulement en partie. Le flan pouvait aussi se fendre sous le coup de marteau, ou son diamètre varier légèrement.



Musée national suisse, Zurich



Musée national suisse, Zurich



Musée national suisse, Zurich



Bibliothèque nationale

Errare humanum est, exposition au Musée d'archéologie du Jura à Lons-le-Saunier, jusqu'au 13 septembre 1998.

Logiciels

Maple V release 5 pour Windows 95

Éditions Waterloo Maple Software.

Le logiciel Maple, dont le nom signifie «érable» en anglais, a pris racine en 1980 à l'Université de Waterloo, au Canada. Bénéficiant des fruits de la recherche en calcul formel, apprécié des milieux scientifiques, c'est aujourd'hui un produit commercial largement répandu. *Maple V*, la cinquième génération, porte un nouveau bougeon, la version 5. C'est l'occasion d'évaluer les capacités de cet outil d'aide aux mathématiques, qui est conçu pour résoudre des problèmes de nature scientifique et technique, à l'école, à l'université, dans les laboratoires de recherche et dans les entreprises.

Un logiciel de calcul formel maîtrise le calcul exact et le calcul symbolique. Le calcul est dit «exact», car le logiciel stocke chaque nombre, aussi grand ou aussi petit soit-il, avec toute la précision requise. L'arithmétique est programmée en conséquence. Cette compétence, qui paraît naturelle, échappe pourtant aux systèmes classiques, notamment les calculatrices. Leur mode de calcul, dit «numérique», représente les nombres avec une précision limitée, provoquant ainsi des erreurs d'ar-

rondi. Dans les systèmes formels, le calcul exact, d'abord programmé sur les nombres entiers, est étendu aux fractions grâce à des règles d'algèbre connues de tous les collégiens. L'enjeu de ce mode de calcul «symbolique» est immense : automatiser la partie technique de l'activité mathématique, de l'algèbre, dans sa signification étymologique de calcul, jusqu'aux nombreuses méthodes analytiques.

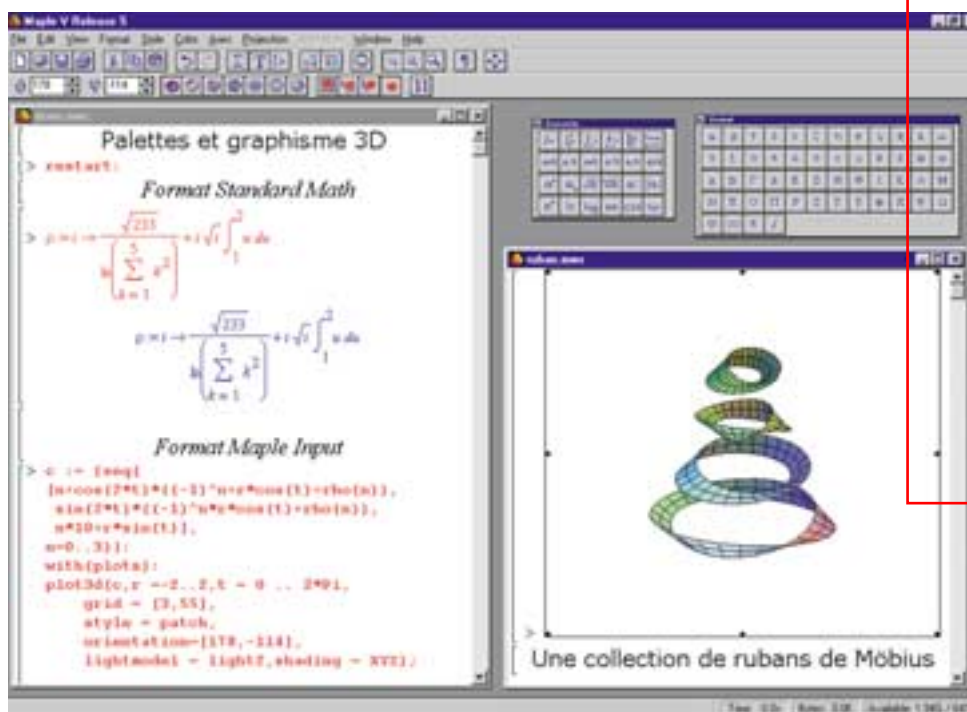
Si *Maple V* est avant tout un système de calcul formel, c'est aussi un excellent dessinateur de graphiques et un bon outil de départ pour le calcul numérique. Son interface utilisateur est conviviale, son langage de programmation est accessible, et le programme est doté d'une vaste bibliothèque de fonctions, en évolution permanente. Toutes ces particularités placent *Maple*, avec *Mathematica* et *Macsyma*, dans le peloton de tête du marché des logiciels commerciaux d'aide aux mathématiques. Cette concurrence a largement profité aux utilisateurs, de version en version, avec des progrès notables des produits, en puissance comme en qualité.

Certaines nouveautés viennent combler des lacunes des versions antérieures. Par exemple, pour désigner le résultat du dernier calcul effectué, la version 5 de *Maple* renonce à son guillemet ("), au profit du % de *Mathematica* ; le guillemet retrouve alors son rôle usuel, pour encadrer un mot ou une phrase. Toutefois, *Maple* continue de comprendre des fichiers écrits avec des versions antérieures et d'effectuer les corrections qui s'imposent.

La Société *Waterloo Maple Software* est particulièrement fière de son interface utilisateur, avec sa multitude de menus, boutons, formats, contextes et la richesse de ses pages d'aide en ligne. Avec la version V, un nouveau format d'entrée des expressions mathématiques, dit «Standard Math», permet enfin d'écrire une instruction de calcul conformément à la calligraphie mathématique, et non dans le formalisme *Maple*, comme c'était le cas jusqu'à présent. Un «éditeur d'expressions» permet de bâtir dynamiquement une formule mathématique, par simple clic de la souris sur les boutons d'une palette de symboles. Ce joli format d'entrée – réaction claire aux palettes de *Mathematica 3.0* – était attendu, puisqu'il existait en sortie (pour l'affichage de résultats) depuis plusieurs versions. L'enseignant l'appréciera pour rédiger cours, exercices et sujets d'examen. Autre nouveauté, les menus contextuels, accessibles par le bouton droit de la souris, optimisent les manipulations. On remarquera la commande «Smartplot», qui transforme une sortie en un graphisme 2D ou 3D, que l'on peut ensuite orienter sous toutes ses faces, au rythme des mouvements de la souris, grâce à de puissants algorithmes de calcul.

Fidèle à son origine universitaire, la firme de Waterloo collecte dans la «share library» de *Maple* les contributions des équipes de recherche en calcul formel du monde entier. Les codes développés par ces chercheurs, lorsqu'ils sont d'utilité générale, sont intégrés à la nouvelle version. C'est ainsi que *Maple V.5* propose une gestion plus systématique des équations différentielles ordinaires. Sa fonction «odeadvisor» permet au profane d'évaluer ses chances de résolution d'une équation différentielle, sans avoir recours à un ouvrage spécialisé ou aux conseils d'un mathématicien. Dans la version V.5, le calcul de bases de Gröbner, qui permet d'aborder le redoutable problème de la réduction des systèmes d'équations algébriques, a été amélioré et étendu.

Depuis 1995, le programme des classes préparatoires scientifiques (biologie exceptée) prévoit d'aborder l'informatique à travers un logiciel de calcul formel qui inclut un langage de programmation. Par ce biais, *Maple* ou *Mathematica* ont été introduits dans la



L'écran présenté à l'utilisateur du programme *Maple*.

plupart des lycées. Les enseignants du secondaire et du supérieur en ont perçu l'utilité, aussi bien pour un premier contact avec l'informatique que pour une meilleure assimilation des programmes de mathématiques. Si l'on ne peut pas découvrir les mathématiques grâce à *Maple*, on peut néanmoins progresser et consolider des savoirs préexistants par la répétition des expériences et le contrôle, notamment visuel, de certaines hypothèses. Par son interface graphique, ses pages d'aide et ses menus contextuels, *Maple* semble désormais assez convivial pour être manipulé au lycée, voire au collège. D'où un impact certain sur la représentation des connaissances et la nature des savoirs transmis.

Alain GIORGETTI

Épistémologie

Eloge du mixte, matériaux nouveaux et philosophie

Bernadette Bensaude-Vincent.
Éditions Hachette, 1998.

Un cheik yéménite visitait mon laboratoire. Après lui avoir fait tâter la flexibilité d'un tube de caoutchouc, nous plongeâmes le tube dans l'air liquide et le projetâmes par terre : il se brisa en menus éclats. Le chef de tribu, bouleversé, se prosterna et invoqua «Allah est grand» : notre perception sensible d'un matériau est un fait culturel, ce qui va dans le sens d'une prise en compte par la philosophie, objet de ce livre sympathique.

Le sujet, si bien trouvé, est superbe, car le mixte (cet aspect du livre est un hommage au livre sur *Le Mixte et la combinaison chimique*, publié en 1902 par Pierre Duhem) est à la racine de la chimie. Tout comme le jeune enfant combinant les couleurs de sa boîte de peinture, le chimiste met ensemble des réactifs «pour voir ce que ça donne». Les résultats de telles combinaisons, même seulement binaires (A + B), sont d'une grande diversité, selon les proportions relatives et selon les paramètres externes, telle la température. Ce livre fleurit l'enjouement d'un renouvellement de la thématique de son auteur : Bernadette Bensaude-Vincent saute ici de la science chimique du XVIII^e siècle à la chimie industrielle au seuil du XXI^e ; on sent son exaltation à traiter de nouveaux matériaux plutôt que de vieux grimoires.



Le carbonate de calcium forme la coquille des mollusques en s'appliquant sur des protéines. Ce principe a été utilisé dans l'industrie.

L'auteur ne s'empêtré pas (trop) dans le piège sociologique en vogue consistant à ne voir dans un laboratoire qu'un lieu de pouvoir où se fabriquent des opinions mandarinales qui plaquent sur notre perception du monde physique (entité elle aussi suspectée) des valeurs sociales idéologiques et datées. En revanche, elle tombe dans le substantialisme qui guette les non-chimistes : les molécules sont conçues comme autant d'objets de collection, figés, avec chacun ses propriétés spécifiques. Elle s'est aussi laissée contaminer par l'empirisme érigé en dogme dans certains milieux industriels.

La dynamique de la formation d'un matériau, ce fameux «mixte», interpelle davantage que le statisme des propriétés mécaniques du matériau fini. Tout solide s'édifie en fonction du partenaire présent, à preuve la cristallisation du carbonate de calcium, qui forme la coquille d'un mollusque en venant s'appliquer sur une protéine lui servant de plan directeur ; les structures locales sont complémentaires ; elles valent d'être envisagées pour les dynamiques responsables d'une telle complémentarité et, surtout, la garantissent.

Que de belles questions philosophiques soulevées (mais, dans l'ensemble, peu traitées ici) : la délimitation de ces objets de pensée que sont les nouveaux matériaux ; la dénomination de ces nouveaux objets ; qu'est-ce que la réalité, dans la «chimiosphère» que nous occupons aujourd'hui ? ; comment le chimiste choisit-il telle action, telle réaction plutôt qu'une autre (un «hyperchoix», selon B. Bensaude-Vincent) ? ; ou encore les dualismes du familier et de l'étrange, du rassurant et de l'angoissant (pensons au boyau plastique : acceptons-nous d'en

manger avec les saucisses ?). J'ai goûté que B. Bensaude-Vincent rapproche de manière *a priori* insolite et pourtant féconde la science des matériaux et la casuistique.

Citons pour finir ce vers philosophique de Jorge Luis Borges (et non José Luis, comme B. Bensaude-Vincent le nomme curieusement), extrait de *La rose profonde* (Éditions Gallimard) :

Le sable ne sait pas qu'il est le sable.

Aucune

Chose ne sent sa forme avec étrangeté.

Pierre LASZLO

Histoire des sciences

Le guide du Paris savant

Anna Alter et Philippe Testard-Vaillant. Éditions Belin, 1997.

Quelle excellente idée que d'avoir écrit ce livre ! On le lit avec beaucoup de plaisir, parce qu'il nous fait mieux connaître ou reconnaître plus de 500 savants, ainsi que les principales institutions scientifiques françaises et les musées scientifiques parisiens.

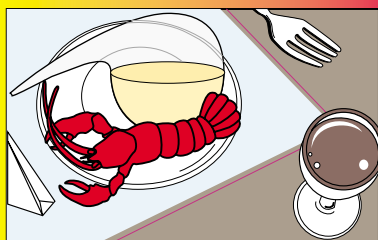
Comme il s'agit de la capitale de la France, une très grande majorité de savants français sont honorés, illustres ou moins illustres, à partir du nom qu'ils ont

FRANCE INFO et POUR LA SCIENCE

vous invitent
à écouter la chronique
de Marie-Odile Monchicourt :

Info Sciences

Tous les jours
sur France Info
à 15 heures 41,
17 heures 10,
20 heures 12,
22 heures 12,
et 23 heures 42



Les prochains
résultats présentés
dans la rubrique
Science et Gastronomie
seront annoncés sur
France Info
le 28 juillet 1998.

FRANCE

info

donné aux rues. Le temps a fait son œuvre. Les savants étrangers ne sont pas oubliés pour autant : Galilée, Fleming, Leibniz ou Newton sont présents, mais Einstein ?

Les articles sont intéressants, généralement bien documentés. Peu d'erreurs ou de lacunes subsistent après le travail d'enquête effectué par les deux auteurs principaux, journalistes scientifiques réputés. Le style est quelquefois un peu libre : «La pomme que Newton reçut sur la poire». Il faut peut-être y aller un peu plus doucement, mais tout cela est bien vivant.

Nous espérons qu'il y aura bientôt une seconde édition si, comme nous le souhaitons, celle-ci vient à être épuisée. Il faudra alors en profiter pour ajouter des plans de Paris, qui faciliteront la localisation des rues qui sont décrites et permettront de programmer des promenades scientifiques. Pour éviter la propagation de quelques erreurs qui subsistent dans le livre (et qui sont colportées souvent depuis des années), il serait bon de solliciter des historiens de toutes les disciplines scientifiques. Par exemple, Pierre Curie n'est pas le «fondateur du magnétisme», et les professeurs au Collège de France n'ont pas attendu 1979 pour décider de la «création d'un enseignement de chimie» : Marcelin Berthelot, qui a sa statue devant le Collège de France (d'ailleurs décrit dans

le livre), Alain Horeau, plus près de nous, et d'autres sont là pour montrer que la chimie fut célébrée il y a déjà plus d'un siècle au Collège de France. Seuls les littéraires sont cités : et les physiciens ? Et les biologistes ?

A l'inverse, à propos de l'Ecole Normale Supérieure, on aurait pu mentionner Raymond Aron, Jean-Paul Sartre et combien d'autres hommes de lettres, ainsi que des hommes politiques : Léon Blum, Georges Pompidou (qui savait écrire, disait le général de Gaulle, lorsqu'il s'attira ses services).

L'article sur le CNRS (rue Michel-Ange) est un peu léger, car combien de découvertes n'ont été possibles, dans notre pays, que grâce à l'initiative prise, juste avant la Seconde Guerre mondiale, de créer cet instrument irremplaçable de progrès scientifique ? Pour avoir plus d'informations, on pourra toutefois se reporter à l'article consacré à Jean Perrin, qui fut un artisan important de cette création.

Au total, c'est avec beaucoup d'intérêt et d'enthousiasme que nous avons lu le *Guide du Paris savant*. Excellente initiative, originale... En attendant la seconde édition, procurez-vous la première, vous ne serez pas déçus.

Michèle BEUGELMANS
et Pierre POTIER



Urbain Le Verrier découvre par le calcul une planète : Neptune. Puis il postule l'existence d'une nouvelle planète qu'il nomme Vulcain. Cette bourde lui vaut ce dessin ainsi légendé : «Ne retrouvant pas sa planète, M. Le Verrier se décide à coller à sa place dans le firmament un calcul qui prouve son existence.» Le Verrier est honoré, à Paris, par la rue Le-Verrier, dans le VI^e arrondissement, et par une statue à l'Observatoire de Paris.

Pour Darwin

Sous la direction de Patrick Tort.
Presses universitaires de France
(1 094 pages), 1997.

« Ah ! ils étaient moins nombreux voici 30 ans, les darwiniens dans les couloirs de la Sorbonne et autres hauts lieux universitaires », diront les esprits chagrins, surpris de la frénésie éditoriale qui paraît s'emparer de la communauté évolutionniste, intarissable, semble-t-il, sur le grand homme et ses écrits.

Ne croyons cependant pas que cette profusion d'ouvrages et de conférences centrés sur Darwin et son œuvre soit propre à notre pays et due à l'activisme de quelques-uns, avec Patrick Tort comme meneur. On doit, certes, à ce philosophe la monumentale *Encyclopédie du darwinisme et de l'évolution* (Presses universitaires de France, 1996) et quelques autres ouvrages qui replacent l'œuvre dans les courants de pensée de notre temps. Toutefois, hors de France, on est tout aussi actif dans la relecture des écrits de Charles Darwin, comme dans l'approfondissement de la réflexion autour du thème de la sélection naturelle. Dans les pays anglo-saxons, le combat que les biologistes doivent quotidiennement mener contre les créationnistes aiguillonne les auteurs, mais on constate aussi l'élargissement de l'approche darwinienne à de nouvelles disciplines : l'anthropologie, la sociologie, la psychologie, la médecine se sentent maintenant concernées par cette vision transformiste, et chacune dans son domaine se trouve en situation d'innover grâce à elle.

L'ouvrage *Pour Darwin* paraît dans la foulée d'un colloque qui s'est tenu à Romainville, en septembre 1997. Est-ce le délai très court, entre la manifestation et la parution des actes, qui a empêché la retranscription des discussions qui n'ont pas manqué de s'y déclencher ? Sans doute. Pourtant, même si l'on apprécie d'avoir si tôt les documents du colloque, on regrette que le document ait laissé de côté une science de marque : la médecine. C'est dommage pour un livre qui a l'ambition, selon la quatrième de couverture, d'illustrer « la vitalité d'une science ». On y reviendra.

Pour l'heure, on doit trouver normal que le premier quart de l'ouvrage



Ce fossile d'oursin figure dans le catalogue d'un musée de Vérone : découvert en 1622, il ressemble tant à une coquille moderne qu'il fut correctement identifié et contribua à l'évolution des idées en biologie.

rassemble des communications autour du concept de « l'effet réversif en évolution », que P. Tort prétend avoir trouvé dans l'œuvre de Darwin et qu'il a lui-même nommé. Les auteurs l'appliquent ici à l'anthropologie (langage, comportements sociaux, psychisme). Le deuxième quart est un ensemble de réflexions autour du concept de niveau d'intégration et de son pouvoir heuristique dans diverses disciplines (neurosciences, métabolisme mitochondrial, fondements biologiques de la théorie de la conscience). Ensuite viennent une succession d'articles qui passent en revue et critiquent les critiques qui ont été faites au darwinisme, de l'antitransformisme du XVIII^e siècle aux écrits de Derek Denton. Pour terminer le volume, la parole est donnée à des biologistes et à des paléontologistes qui exposent l'usage qu'ils font de Darwin en cette fin de siècle.

On ne lira pas d'une traite un livre aussi riche, mais on le consultera pour s'ouvrir sur les usages que font des collègues disciplinairement éloignés de *L'origine des espèces* et des autres travaux de Darwin. L'ouvrage aurait gagné à bénéficier des écrits d'autres spécialistes de Darwin, tels Jean Gayon ou Michael Ruse, penseurs parmi les plus originaux, mais on ne peut pas inviter tout le monde...

Je signalais que, curieusement, la médecine était absente de ces réflexions. C'est un regret sur lequel je veux revenir : il serait temps que les maladies ne soient plus considérées comme des accidents qui concernent des individus, mais que soit abordé le problème général de la vulnérabilité des populations humaines aux maladies. Quels processus sélectifs nous ont communiqué une sensibilité aux cancers ou aux virus ? Lesquels nous ont bâti un cerveau ? Questions trop naïves pour être ainsi posées.

De plus, les problèmes de « médecine darwinienne » ne relèvent pas tous de recherches s'adressant à des maux mortels, mais concernent aussi la vie de

tous les jours : est-il raisonnable de couper la fièvre d'un malade, d'empêcher une femme enceinte d'avoir des nausées, etc., alors que ces manifestations sont programmées pour éviter des maux plus graves, telle l'intoxication. Dans un autre domaine, on sait par exemple que l'alimentation, au cours de l'histoire des hommes, a débouché sur des « spécialisations » : certains peuples ne vivent que de laitages ; l'obésité est une qualité des populations insulaires qui leur permet de surmonter les périodes de disette.

Le courant de cette approche darwinienne de la médecine existe en France comme dans les autres pays, et il méritait d'être représenté à Romainville. Des spécialistes comme Jacques Ruffié ou Mirko Grmek, pour ne citer que ces deux noms, auraient certainement enrichi le débat : les adaptations des populations humaines aux climats extrêmes, à des régimes alimentaires étroits, la dispersion des épidémies, l'acquisition immunologique sont des problèmes d'évolution justiciables de la vision transformiste. Cette absence est d'autant plus regrettable que certaines recherches génétiques médicales très (trop ?) largement financées gagneraient à être passées au crible darwinien.

Faisons confiance au dynamisme de la communauté évolutionniste pour que ces thèmes trouvent preneurs dans de prochains congrès.

Jean-Louis HARTENBERGER

Biologie

Les avatars du gène : la théorie néodarwinienne de l'évolution

Pierre-Henri Gouyon, Jean-Pierre Henry et Jacques Arnould. Éditions Belin, 1997

A vātara est un mot sanscrit qui désigne les incarnations successives de Vishnu : le dieu hindouiste revêt des formes diverses pour se manifester dans le monde sensible. Il y a Matsya, le poisson ; Kurma, la tortue ; Vahara, le sanglier ; Narasimha, l'homme-lion ; Vamana, le nain ; Parasurama, le héros qui délivra les brahmanes ; Rama, le héros conquérant ; Krishna, le prince berger ; Bouddha, le Kshatriya ; et Kalki, le cavalier blanc.

Aussi variés sont les définitions et les rôles du gène : cette petite sous-unité

POUR LA SCIENCE

8, rue Férou 75278 PARIS CEDEX 06
Tél : 01-55-42-84-00
<http://www.pourlascience.com>

POUR LA SCIENCE Directeur de la rédaction : Philippe Boulanger. Hervé This (Rédacteur en chef), Françoise Cinotti (Rédactrice en chef adjointe), Bénédicte Leclercq (Rédactrice en chef adjointe), Luc Allemand, Yann Esnault, Hélène Maurey, Philippe Pajot (Rédacteurs). Secrétariat de rédaction : Annie Tacquenot, Pascale Thiollier, Céline Lapert. Direction Marketing et Publicité : Henri Gibelin, assisté de Séverine Merviel et Christelle Dumas. Direction financière : Pierre Lecomte, assisté de Anne Gusdorf. Fabrication : Jérôme Jalabert, assisté de Delphine Savin. Directeur de la publication et Gérant : Olivier Brossollet.

Ont également collaboré à ce numéro : Eric Buffetaut, Raymond Campan, Daniel Collobert, Paul Decaix, Philippe Delbonnel, Philippe Hers, Annick Horiuchi, Évelyne Host-Platret, Christian Jeanmougin, Anne Masé, Claude Olivier, Roseline Talbot.

SCIENTIFIC AMERICAN Editor : John Rennie. Board of editors : Michelle Press, Timothy Beardsley, Wayt Gibbs, Alden Hayashi, Kristin Leutwyler, Madhusree Mukerjee, Sacha Nemecek, Ricki Rusting, David Schneider, Gary Stix, Paul Wallich, Philip Yam, Glenn Zorpette. Publisher : Joachim Rosler. Chairman and Chief Executive Officer : John Hanley. Corporate Officers : Robert Biewen, Frances Newburg, Anthony Degutis.

PUBLICITÉ France

Chef de Publicité : Susan Mackie, assistée de Anne-Claire Ternois, 8 rue Férou 75278 Paris Cedex 06
Tél. : 01 55 42 84 28.

Télex : LIBELIN 202978F.

Télécopieur : 01 43 25 18 29

Étranger : John Moeling, 415 Madison Avenue, New York. N.Y. 10017 - Tél. (212) 754.02.62

SERVICE ABONNEMENTS

Ginette Grémillon : 01 55 42 84 04.

SERVICE DE VENTE RÉSEAU NMPP

Henri Gibelin.

DIFFUSION DE LA BIBLIOTHÈQUE POUR LA SCIENCE

Canada : Modulo - 233, avenue Dunbar Mont-Royal, Québec, H3P 2H2 Canada

Suisse : GM Diffusion - Rue d'Etraz, 2 - CH 1027 Lonay

Autres pays : Éditions Belin - B.P. 205, 75264 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 adressés 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 d'exploitation du droit de Copie (3, rue Hautefeuille - 75006 Paris).

Nos lecteurs trouveront des bulletins d'abonnement en pages 18A, 18B, 98A et 98B.

de la longue molécule d'ADN est composée de nucléotides (chacun est constitué d'un sucre, le désoxyribose, d'un groupe phosphate et d'une base azotée : adénine, guanine, cytosine ou thymine) ; elle est une unité de fonction, de mutation, de recombinaison, transmise de génération en génération, qui commande et contrôle le développement et les caractères d'un individu. Un gène peut s'associer à d'autres gènes sur un chromosome, se localiser dans le noyau de la cellule ou dans son cytoplasme, constituer un ensemble, ou génome, caractéristique d'une cellule (ou d'un virus), ou former des pools géniques à l'échelle d'une population, d'une espèce ou d'un écosystème selon des groupements hiérarchisés.

La place prééminente dans le titre de l'ouvrage de P.-H. Gouyon, J.-P. Henry et J. Arnould ne doit pas laisser croire que ces trois auteurs consacrent ici la suprématie du gène «égoïste», uniquement consacré à sa survie, le reste étant accessoire. Les auteurs ne font pas non plus l'apologie d'un réductionnisme radical : ils explorent tous les niveaux du vivant, du plus simple au plus complexe et *vice versa*. On en retire qu'il est difficile d'appréhender l'évolution biologique avec des certitudes et que le génome est un «lieu de conflits perpétuels».

Enfin, des chercheurs français se sont attaqués à un bilan complet, à fois épistémologique, philosophique et scientifique, de l'évolution, jusque dans ses aspects les plus actuels ! On avait presque oublié que le concept d'évolution des organismes vivants était issu des spéculations du siècle des Lumières et qu'il avait pris tournure au début du XIX^e siècle, avec le transformisme du bien malmené Jean-Baptiste Lamarck. Ce dernier avait osé avancer que les espèces n'étaient pas fixes, mais le puissant Cuvier avait réduit au silence (et de façon durable) les esprits novateurs. Venu d'Angleterre, le darwinisme n'a pas troublé la France, qui a perdu l'initiative des débats dans ce domaine.

Ce livre ambitieux veut relancer les esprits quelque peu engourdis. Il troublera ceux qui ont des certitudes sur l'origine de la vie et sur les mécanismes d'apparition des espèces. Il perpétuera peut-être l'impression fausse que, plus la science avance, et moins on comprend les mystères du vivant, alors à quoi bon réfléchir ? Il donnera sans doute aussi l'idée que tous ces biologistes qui modélisent, expérimentent, défendent une théorie, bientôt combattue avec apreté par les chercheurs concurrents, renvoient une image assez fidèle de la théorie darwinienne, de la survie du plus apte (ou du plus malin médiatiquement). Certes, mais

isons attentivement, et notre effort sera récompensé : l'analyse fine des théories de l'évolution est proposée avec un esprit critique et avec la volonté de simplifier le plus possible des notions plus accessibles aux spécialistes qu'au grand public, sans dogmatisme. Admirez en passant le tour de force de ces évolutionnistes néo-darwiniens, qui ont raison de se placer dans ce cadre conceptuel. P.-H. Gouyon a proposé une définition du néo-darwinisme : «On peut dire pour simplifier que c'est une théorie scientifique qui rend compte de l'évolution grâce à la combinaison d'une variation du matériel héréditaire et d'une sélection qui trie *a posteriori*.»

Tous ceux qui annoncent la fin du darwinisme avec une hargne digne de l'époque victorienne sont des égarés qui défendent une biologie dépassée ou des convictions religieuses figées. Un paradigme qui dure depuis plus d'un siècle et qui a engendré autant de programmes de recherche et de débats offre certaines garanties. Même les théories du neutralisme de Motoo Kimura et des équilibres ponctués de Niles Eldredge et Stephen Jay Gould, qui apparaissent comme des pensées dissidentes, sont aujourd'hui intégrées au néo-darwinisme : le neutraliste se fonde sur la très grande variabilité génétique révélée par la biologie moléculaire, au point que la plupart des variations ne se posent pas en termes de survie des organismes, leur persistance ou leur élimination étant aléatoire et sélectivement neutre ; le punctualisme suppose que la plupart des événements importants de l'évolution surviennent dans de très courts moments de spéciation, suivies de longues périodes de stabilité. En revanche, les débats sur la sociobiologie (l'étude systématique des fondements biologiques des comportements sociaux, d'abord proposée chez les animaux, puis étendue à l'espèce humaine) et l'eugénisme, science qui présente des moyens d'«améliorer» les populations humaines, n'ont en rien perdu de leur actualité, et les comités d'éthique sont les garants de l'utilisation des manipulations génétiques. Chercher à comprendre l'évolution biologique est une façon d'envisager le monde futur en connaissance de cause. Ce livre en fournit les clés.

Michèle FEBVRE

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>