

formulée, elle n'en donne pas un moyen de calcul explicite. Bien que plus techniques et moins élégantes que la preuve par le biais de la théorie de la complexité (plus tardive), les preuves de Lebesgue et Sierpinski semblent être de la même nature : elles désignent de façon théorique des nombres dont il est prouvé qu'ils sont absolument normaux, mais dont le calcul des décimales n'est pas effectué et n'est peut-être pas possible. Nous allons voir cependant que ces démonstrations techniques possèdent un avantage sur celle de la théorie de la complexité.

Dans les motivations que Lebesgue donne au début de son travail, apparaît la question de l'axiome du choix, qui préoccupait les mathématiciens du début du XX<sup>e</sup> siècle. L'axiome du choix affirme qu'à toute famille  $\{A_i\}$  d'ensembles non vides est associé au moins un ensemble  $C$  ayant exactement un élément commun avec chaque  $A_i$ . Lebesgue exprime sa conviction que pour les définitions « qui ne sont pas illusoire, il est possible d'apporter des précisions qui conduisent à la définition d'un des ensembles [ou nombres] dont on démontre l'existence ». Or on sait aujourd'hui que certaines démonstrations d'existence ne peuvent pas être rendues constructives et seraient donc considérées illusoire par Lebesgue.

Paradoxalement, tel est justement le cas pour les ensembles non mesurables « au sens de Lebesgue » ! Les ensembles mesurables sont ceux constitués de nombres réels auxquels on peut attribuer une longueur, et dans le cas du plan, ceux auxquels on peut attribuer une aire. L'axiome du choix permet de prouver qu'il existe des ensembles non mesurables (n'ayant donc pas de longueur, ou d'aire), mais ne permet d'en définir aucun. Or en 1970, Robert Solovay a démontré qu'on pouvait, sans introduire de contradiction, ajouter à la théorie des ensembles habituelle l'axiome « tout ensemble est mesurable au sens de Lebesgue ». Cela signifie qu'aucune définition précise d'un ensemble non mesurable n'est possible dans le cadre de la théorie des ensembles sans l'axiome du choix. La démonstration classique de l'existence d'ensembles non mesurables

## 3. Les nombres normaux sont majoritaires

Borel a démontré que les nombres qui ne sont pas normaux en base 10 sont en quantité négligeable : en prenant un nombre réel au hasard entre 0 et 1, on tombera, sauf cas exceptionnellement rare, sur un nombre normal en base 10 (et même normal en toute base de numération).

C'est assez choquant, car on n'arrive même pas à démontrer que les nombres usuels comme  $\pi$ ,  $e$ ,  $\log(2)$ ,  $\sqrt{2}$ , etc., sont normaux en base 10. Voici une façon de présenter le problème qui n'est pas une preuve rigoureuse, mais qui aide à saisir l'étonnante affirmation de Borel.

Tirer au hasard un nombre réel entre 0 et 1 revient à tirer au hasard une suite infinie de 0, 1, 2, ..., 9 avec un dé équitable, par exemple un icosaèdre où chaque chiffre apparaît sur

deux faces. Or quand on mène une telle suite de tirages équitables, chaque séquence de longueur  $k$ , par exemple la séquence 1789 de longueur 4, a la probabilité  $1/10^k$  d'être obtenue à



chaque instant (1789 apparaît en moyenne une fois sur 10 000).

C'est une conséquence de la loi des grands nombres, car les événements sont indépendants et chaque chiffre se présente avec une probabilité  $1/10$ . Pour une séquence fixée de longueur  $k$ ,

la probabilité pour que la fréquence limite soit  $1/10^k$  est donc de 100 %. Comme une réunion dénombrable d'ensembles négligeables est aussi négligeable, la probabilité qu'une suite infinie de tirages ne donne pas la fréquence limite  $1/10^k$  pour une séquence de longueur  $k$  est nulle. Autrement dit, un nombre réel pris au hasard entre 0 et 1 a toutes les chances d'être normal en base 10. La conclusion s'étend à la propriété d'être absolument normal (normal dans toute base), car il y a une infinité dénombrable de bases de numération possibles (l'ensemble  $NN(b)$  des nombres non normaux en base  $b$  est négligeable, donc la réunion de tous les  $NN(b)$  aussi ; son complémentaire est l'ensemble des nombres absolument normaux).

est irrémédiablement non constructive : elle s'appuie sur l'axiome du choix, mais pire, sans lui, elle est impossible !

## En quête d'un algorithme

Arrivés à ce point et conscients, grâce à la logique mathématique, de l'inconfort de la situation concernant les nombres absolument normaux, notre problème se résume alors ainsi :

– La preuve d'existence de Borel est non constructive, mais ne l'est pas aussi gravement que celle des ensembles non mesurables, car les reformulations de Lebesgue et de Sierpinski ainsi que les approches par la théorie de la complexité donnent des définitions précises de nombres absolument normaux.

– Cependant, la logique nous a aussi appris qu'être bien défini n'implique pas être calculable. Reste donc la question, passée sous silence par Lebesgue et Sierpinski faute d'idées assez claires sur le sujet, et devenue parfaitement mathématique grâce à Turing qui introduisit le concept de fonction

calculable et de nombre réel calculable : existe-t-il des nombres absolument normaux et calculables ?

Celui grâce à qui la question de l'existence a pris une forme algorithmique claire et mathématique est aussi celui qui la résout positivement pour les nombres absolument normaux dans son article non publié de 1934. Turing refait en quelque sorte le travail de Sierpinski et Lebesgue, mais il cherche à aller plus loin et veille à définir un nombre absolument normal accompagné d'un algorithme de calcul. Malheureusement, ce qu'on a retrouvé de l'article de Turing ne va pas tout à fait au bout et reste incomplet sur un point qui ne sera définitivement éclairci qu'en 2002 par Verónica Becher et Santiago Figueira. Enfin rendue parfaitement constructive et calculable, la démonstration d'existence des nombres absolument normaux de Borel est-elle pour autant définitive ?

Non, la satisfaction éprouvée par ce succès récent est atténuée par deux remarques qui suggèrent qu'il reste encore du travail à faire et que les mathématiciens, les logiciens

## 4. La complexité des irrationnels algébriques

Le magnifique et puissant théorème de Adamczewski et Bugeaud permet de démontrer la conjecture de Loxton et van den Porten que tout nombre dont les chiffres en base  $b$  sont produits par un automate fini est soit rationnel, soit transcendant (non solution d'une équation polynomiale à coefficients entiers). Cette conjecture signifie que les nombres algébriques (solutions d'équations polynomiales à coefficients entiers) sont ou bien rationnels (et leurs chiffres sont ultimement périodiques, donc très simples), ou bien complexes (leurs chiffres ne peuvent être écrits par un automate fini).

Le théorème donne des informations sur la complexité des chiffres d'un nombre réel mesurée par la fonction  $p_b(m)$ . Si  $x$  est un nombre réel écrit en base  $b$ , on note  $p_b(m)$  (et on dénomme fonction complexi-

té de  $x$ ) le nombre de séquences différentes de longueur  $m$  rencontrées dans son développement en base  $b$ .

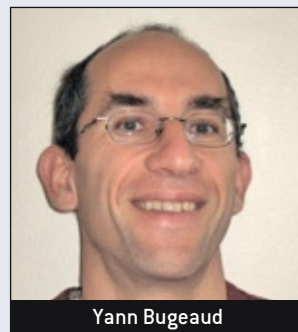
Par exemple,  $1/5$  s'écrit, en base 2, 0,0011001100110011... Le nombre de mots de longueur 3 rencontrés dans ce développement est 4, puisque les seules séquences de longueur 3 qu'on trouve sont 001, 011, 110 et 100. Ils en suit que  $p_2(3) = 4$  pour le nombre  $1/5$ . Notons que  $p_2(m)$  est inférieur ou égal à  $2^m$ , et que  $p_b(m)$  est inférieur à  $b^m$ .

THÉORÈME DE ADAMCZEWSKI ET BUGEAUD. Si  $x$  est un nombre irrationnel algébrique écrit en base  $b$ , alors  $p_b(m)/m$  croît indéfiniment (pour  $m$  assez grand,  $p_b(m)/m$  dépasse 10, et ne revient jamais en dessous, plus loin il dépasse 100, etc.).

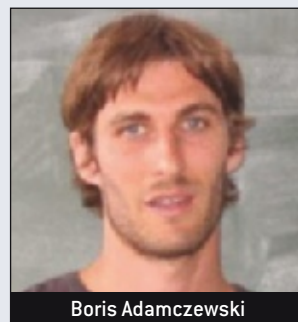
Dit autrement, la variété des séquences de longueur  $m$  qu'on trouve dans le développement d'un nombre algébrique irrationnel est grande :

elle augmente plus vite que n'importe quelle fonction linéaire. Ce théorème ne permet pas encore de prouver que les nombres irrationnels algébriques sont normaux, ce que l'on pense vrai, mais on tient là une affirmation forte sur les décimales des nombres algébriques irrationnels qui nous approche un peu du but.

Le théorème est aussi un outil puissant pour démontrer que des nombres sont transcendants et il permet de retrouver le résultat que le nombre de Prouhet-Thue-Morse 0,0110100110010110..., souvent rencontré en arithmétique, est transcendant (ce nombre est obtenu en écrivant, après la virgule, 0 puis le complément 1, ce qui donne 01 ; on ajoute alors le complément 10, ce qui donne 0110 ; on ajoute encore le complément 1001, ce qui donne 01101001 ; et ainsi de suite).



Yann Bugeaud



Boris Adamczewski

et les informaticiens n'en ont pas fini avec les nombres absolument normaux.

### Quatre niveaux d'existence mathématique

D'une part, l'algorithme explicité est affreusement inefficace : il est « doublement exponentiel », ce qui signifie que pour connaître  $n$  décimales du nombre absolument normal qu'il calcule, il faut le faire fonctionner durant un temps en gros proportionnel à  $\exp(\exp(n))$  (ou 2 à la puissance  $2^n$ ), qui est une fonction de la variable  $n$  à croissance extrêmement rapide. À ma connaissance, personne n'a essayé de produire des décimales du nombre calculable que l'algorithme de Turing-Becher-Figueira égraine indéfiniment... en théorie. Cette double exponentielle illustre un quatrième niveau dans les degrés de l'existence mathématique : après l'existence pure (preuve de Borel), l'existence avec caractérisation (théorie

de la complexité ou preuves de Lebesgue et Sierpinski), l'existence avec algorithme de calcul (Turing, Becher, Figueira), on aimerait bien atteindre l'existence avec un algorithme de calcul praticable et donnant donc « pour de vrai » les décimales d'un nombre absolument normal qu'on pourrait enfin voir écrit sur une page imprimée !

La deuxième raison d'insatisfaction est que la question posée à propos des nombres absolument normaux n'est pas tant celle d'en définir un et de le calculer, mais plus simplement :  $\sqrt{2}$ ,  $e$  ou  $\pi$ , ou les nombres irrationnels que l'on connaît, et dont l'étude des chiffres suggère qu'ils sont absolument normaux, le sont-ils vraiment ?

Sur ces dernières questions, des progrès ont été faits assez récemment. Bien sûr le calcul des chiffres de  $\pi$ , qui a maintenant été mené jusqu'à la décimale en position  $10^{13} = 10\,000\,000\,000\,000$ , permet d'étudier empiriquement la question. Un travail portant sur les quatre premiers milliards de ces chiffres et utilisant une modélisation à l'aide d'un processus de Poisson a conduit David Bailey, Jonathan Borwein, C. Calude,

Michael Dinneen, Monica Dimiteescu et Alex Yee à la conclusion suivante : « Il est extraordinairement improbable que  $\pi$  ne soit pas normal en base 16, étant donné la normalité de son segment initial. » Malheureusement, le nombre décimal qui aurait le même développement pendant quatre milliards de décimales et qui se poursuivrait par des 0 aurait lui aussi conduit à la même conclusion, qui n'est donc en rien l'esquisse d'une preuve.

### Vers la normalité de $\log(2)$ ou de $\pi$

D. Bailey et Richard Crandall ont proposé une condition simple assurant la normalité du nombre  $\log(2)$  (qui, comme  $\sqrt{2}$ ,  $e$  ou  $\pi$ , intéresse particulièrement les mathématiciens). Si la suite définie par  $x_0 = 0$  et  $x_n = (2x_{n-1} + 1/n) \bmod 1$  est équidistribuée [ $x_n$  se retrouve dans l'intervalle  $[a, b]$  à la fréquence  $b - a$ , quels que soient les nombres réels  $a$  et  $b$  tels que  $0 < a \leq b < 1$ ], alors  $\log(2)$  est normal en base 10. Même si la condition semble plus simple à établir que