pas vraiment en action, et nous aboutissons à $g_6(3) = 0$.

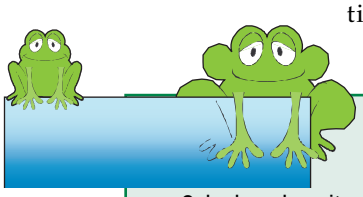
Le cas de la graine 4 gagne en revanche en intérêt… et devient plus compliqué. La suite atteint en effet la valeur 0 au bout d'un nombre d'étapes gigantesque *(voir la figure 3)*. Dans le cas de la graine 5, et, *a fortiori*, des suivantes, les démonstrations paraissent hors de portée, car le nombre d'étapes augmente considérablement avec la valeur de la graine. Le théorème de Goodstein serait-il donc impossible à démontrer?

Non. Il existe bien une démonstration, valable pour tout n, mais elle nécessite de sortir du cadre de l'arithmétique et de recourir à une autre méthode : les ordinaux. Les ordinaux sont une prolongation de la suite des entiers, introduite à la fin du siècle dernier par le mathématicien allemand Georg Cantor *(voir la figure 1)*. Ils comprennent donc les entiers et, plus grands que tous les entiers, les ordinaux infinis, aussi appelés «transfinis», c'est-à-dire «au-delà du fini». Deux points importent pour la démonstration du théorème de Goodstein : a) Les ordinaux sont munis d'une arithmétique similaire à celle des entiers (les quatre opérations de base) ; b) Toute suite décroissante d'ordinaux parvient à la valeur 0 en un nombre fini d'étapes *(voir la figure 4)*.

Le théorème de Goodstein se démontre comme suit. Nous introduisons, pour chaque entier p, une «superdilatation» D_p définie comme d_p, à cela près qu'au lieu de remplacer p par $p+1$ dans le développement en base p itérée, p est remplacé par l'ordinal infini ω. Ainsi, la relation $266 = 2^{2^{2+1}} + 2^{2+1} + 2$ entraîne $D_2(266) = \omega^{\omega^{\omega+1}} + \omega^{\omega+1} + \omega$, à comparer avec $d_2(266) = 3^{3^{3+1}} + 3^{3+1} + 3$.

Examinons une propriété de la superdilatation : appliquer D_p revient à appliquer d_p puis D_{p+1} quel que soit le nombre de départ. En effet, dans le cas de la superdilatation directe D_p, on développe le nombre initial en base p itérée, puis on remplace les p par des ω. Dans le cas de la superdilatation indirecte (d_p suivie de D_{p+1}), on dilate d'abord le nombre, c'est-à-dire qu'on le développe en base p itérée et l'on remplace les p par des $p+1$, puis seulement intervient la superdilatation, où l'on remplace les $p+1$ par des ω, et le résultat final est le même. Vérifions-le avec 4 comme nombre de départ et avec $p = 2$. La superdilatation calculée directement est $D_2(4) = \omega^\omega$, car on remplace 2^2 par ω^ω. La superdilatation indirecte mène à $d_2(4) = 3^3$, puis $D_3(d_2(4)) = D_3(3^3) = \omega^\omega$ aussi.

Une autre propriété de la superdilatation, déterminante pour démontrer le théorème de Goodstein, est que D_p est une fonction croissante pour chaque p. Par exemple, $D_2(4)$ est supérieur à $D_2(2)$.



3. SUITE DE GOODSTEIN DE GRAINE 4

Calculons la suite de Goodstein de graine 4. On trouve $g_1(4) = 4$, $g_2(4) = d_2(4) - 1 = d_2(2^2) - 1 = 3^3 - 1 = 26$.
$g_3(4)$ est égal à $d_3(26) - 1$, soit $d_3(3^2 \times 2 + 3 \times 2 + 2) - 1 = (4^2 \times 2 + 4 \times 2 + 2) - 1 = 41$.

Continuons : $g_4(4) = d_4(41) - 1 = d_4(4^2 \times 2 + 4 \times 2 + 1) - 1 = (5^2 \times 2 + 5 \times 2 + 1) - 1 = 60$.
$g_5(4) = d_5(60) - 1 = d_5(5^2 \times 2 + 5 \times 2) - 1 = (6^2 \times 2 + 6 \times 2) - 1 = 83$.

Quel que soit p, $g_p(4)$ a toujours un développement en base p +1 de la forme $(p+1)^2 \times a_p + (p+1) \times b_p + c_p$, avec $a_p \leq 2$, $b_p \leq p$ et $c_p \leq p$. Démontrons cette formule, très utile pour prouver que la suite de graine 4 atteint 0. La formule est vraie pour $g_2(4) = 26 = 3^2 \times 2 + 3 \times 2 + 2$. Ici, on a $a_2 = 2$, $b_2 = 2$ et $c_2 = 2$, que nous notons (2, 2, 2). La formule est conservée quand nous passons de $g_p(4)$ à $g_{p+1}(4)$, comme le montre l'étude des différents cas. L'évolution de a, b et c est similaire au compte à rebours d'une horloge où a représenterait les heures, b les minutes et c les secondes ; cependant, l'horloge, détraquée, ne fonctionne plus en base 12, mais dans une base p qui augmente !

Premier cas : $c_p \neq 0$. La présence du facteur – 1 diminue c et laisse a et b inchangés. Par exemple, on trouve $(a_2, b_2, c_2) = (2, 2, 2)$, $(a_3, b_3, c_3) = (2, 2, 1)$.
On a alors $a_{p+1} = a_p$, $b_{p+1} = b_p$, $c_{p+1} = c_p - 1$.

Deuxième cas : $c_p = 0$ et $b_p = 0$, comme dans $(a_4, b_4, c_4) = (2, 2, 0)$. Alors, à l'itération suivante, b diminue d'une unité et c réapparaît avec une valeur égale à la base diminuée de 1, c'est-à-dire p. On obtient $(a_5, b_5, c_5) = (2, 1, 5)$. Dans ce cas, $a_{p+1} = a_p$, $b_{p+1} = b_p - 1$ et $c_{p+1} = p$. Le coefficient b diminue seulement quand c égale 0, ce qui prend de plus en plus de temps car c réapparaît avec la valeur p, de plus en plus grande.

Troisième cas, $b_p = 0$ et $c_p \neq 0$, comme dans $(a_{22}, b_{22}, c_{22}) = (2, 0, 0)$. À l'itération suivante, on a $a_{p+1} = a_p - 1$, $b_{p+1} = p$, $c_{p+1} = p$, comme dans $(a_{23}, b_{23}, c_{23}) = (1, 23, 23)$. Ensuite, on obtient (1, 0, 0) pour $p = 3 \times 2^{27} - 2$, à la suite de quoi a devient nul et les expressions prennent la forme $(p+1) \times b_p + c_p$. Ensuite, on obtient (0, 1, 0), soit $g_p(4) = p + 1$, pour $p = 3 \times 2^{27} \times 2^{3 \times 2^{27}-2} - 1$; finalement, on atteint (0, 0, 0), soit $g_p(4) = 0$ pour $p = 3 \times 2^{27+3\times 2^{27}-1} - 2$, soit $p = 3 \times 2^{402\,653\,211} - 2$, un entier dont l'écriture en base 10 a environ 130 millions de chiffres.

p	(a, b, c)
2	(2, 2, 2)
3	(2, 2, 1)
4	(2, 2, 0)
5	(2, 1, 5)
6	(2, 1, 4)
7	(2, 1, 3)
…	…
10	(2, 1, 0)
11	(2, 0, 11)
12	(2, 0, 10)
…	…
22	(2, 0, 0)
23	(1, 23, 23)
…	…
46	(1, 23, 0)
47	(1, 22, 47)
48	(1, 22, 46)
…	…

Les supersuites de Goodstein

Après les superdilatations, définissons maintenant des suites d'ordinaux G, sorte de super-suites de Goodstein : l'élément $G_p(n)$ est l'ordinal $D_{p+1}(g_p(n))$. Voyons des exemples, et d'abord pour la graine 2 :
$G_1(2) = D_2(g_1(2)) = D_2(2) = \omega$,
$G_2(2) = D_3(g_2(2)) = D_3(d_2(g_1(2)) - 1) = D_3(d_2(2) - 1) = D_3(3 - 1) = 2$.
Pour la graine 4, on trouve :
$G_1(4) = D_2(g_1(4)) = D_2(4) = \omega^\omega$,
$G_2(4) = D_3(g_2(4)) = D_3(d_2(g_1(4)) - 1) = D_3(d_2(2^2) - 1) = D_3(3^3 - 1) = D_3(26) = D_3(3^2 \times 2 + 3 \times 2 + 2) = \omega^2 \times 2 + \omega \times 2 + 2$.
Dans ces exemples, on a $G_1(2) > G_2(2)$ et $G_1(4) > G_2(4)$. Le schéma de la figure 2 généralise ce résultat et montre que les ordinaux $G_p(n)$ forment une suite strictement décroissante avec p, pour chaque graine n. Alors, une telle suite atteint 0 en un nombre fini d'étapes. Comme $G_p(n)$ ne peut valoir 0 que si $g_p(n)$ lui-même vaut 0, $g_p(n)$ atteint 0 en un nombre fini d'étapes, ce qui termine la démonstration du théorème de Goodstein.

L'infini indispensable ?

Le théorème de Goodstein est une propriété d'arithmétique, au sens où il ne met en jeu que les nombres entiers et leurs quatre opérations élémentaires (addition, soustraction, multiplication, division), mais la démonstration que nous venons de donner repose sur l'utilisation d'objets infinis. En existe-t-il une autre qui ne passerait pas par l'infini ?

La démonstration d'un théorème n'utilise pas toujours que les outils avec lesquels il est formulé. De nombreux résultats d'arithmétique sont démontrés grâce à l'analyse, c'est-à-dire aux nombres réels et aux fonctions. Ceci a donné naissance à la théorie analytique des nombres (par opposition à la théorie algébrique des nombres, qui n'utilise que les entiers). Or, un nombre réel vraiment quelconque n'est exprimé ni par des fractions ni par des racines n-ièmes : c'est un objet dont le développement possède une suite infinie de chiffres après la virgule et, plus généralement, dont toutes les écritures sont infinies. Dès qu'une démonstration d'arithmétique utilise des nombres réels, une certaine forme d'infini y est donc présente. Notons cependant qu'il peut exister une autre démonstration alternative à l'utilisation de l'analyse et fondée seulement sur les entiers et leurs quatre opérations, mais elle est souvent très difficile à trouver. Ainsi, le célèbre théorème des nombres premiers, démontré à l'aide des nombres complexes par J. Hadamard et C. de La Vallée-Poussin en 1896, n'a été démontré par l'arithmétique que plus de cinquante ans plus tard, grâce à Selberg et Erdös. Leur démonstration est par ailleurs beaucoup plus compliquée que l'originale.

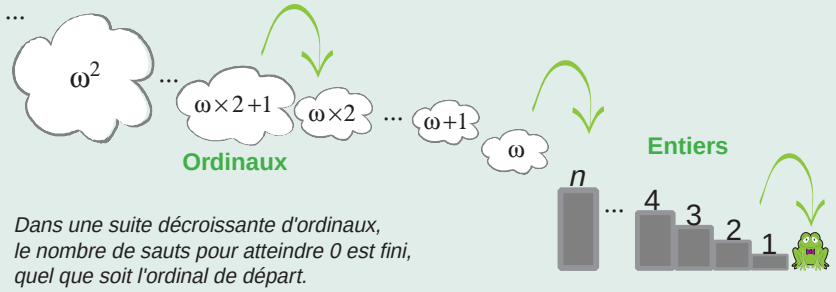

Guiseppe Peano

Précisons ce qu'on appelle une démonstration d'arithmétique. Les propriétés de base de l'arithmétique sont décrites par le système d'axiomes proposé à la fin du XIX^e^ siècle par Giuseppe Peano, mathématicien italien mort en 1932. Toutes les propriétés de base des entiers peuvent se démontrer dans le système de Peano, donc par récurrence : bien sûr, toutes les démonstrations d'un livre d'arithmétique ne sont pas des récurrences, mais on pourrait toujours s'y ramener, au moins de façon théorique.

4. SUITES DÉCROISSANTES D'ORDINAUX

Construisons une suite strictement décroissante à partir d'un entier, même très grand, disons 10^9. Nous arriverons alors nécessairement à 0 après au plus 10^9 étapes. En revanche, puisqu'il y a une infinité d'entiers sous l'ordinal ω, nous pourrions penser que, partant de ω, il est possible de descendre une infinité de fois. Il n'en est rien, car, si nous construisons une suite strictement décroissante partant de $\alpha_1 = \omega$, il faut choisir pour α_2 un ordinal strictement plus petit que ω, c'est-à-dire un entier : si nous choisissons $\alpha_2 = 12$, nous atteindrons 0 au plus 12 étapes plus tard. Si nous choisissons $\alpha_2 = 10^9$, nous pourrons descendre 10^9 marches avant d'arriver à 0, ce qui est plus, mais est toujours un nombre fini. Ainsi, à la différence du cas où nous partons d'un entier, il est impossible de donner de borne sur la longueur d'une suite descendant de ω à 0, il existe des suites aussi longues qu'on veut, *mais* chacune d'entre elles est néanmoins finie.

L'argument est le même lorsqu'on part d'ordinaux plus grands. Ainsi, si nous partons de $\alpha_1 = \omega^2$, nous devons choisir α_2 strictement plus petit, ce qui nous oblige à un saut infini : un tel α_2 est de la forme $\omega \times p + q$, où p et q sont des entiers. De là, au plus q étapes de descentes mènent à $\omega \times p$ (nous envisageons ici le pire cas), et, à l'étape suivante, nous arrivons à un ordinal de la forme $\omega \times (p - 1) + q$, et ainsi de suite. Après au plus p sauts du même type, on arrive aux entiers, et donc à 0. Toute suite décroissante d'ordinaux parvient à la valeur zéro en un nombre fini d'étapes.



Dans une suite décroissante d'ordinaux, le nombre de sauts pour atteindre 0 est fini, quel que soit l'ordinal de départ.

Théorème de Kirby et Paris

Alors ? Pourrions-nous démontrer aussi, avec un peu d'astuce, le théorème de Goodstein par récurrence sans recourir à l'infini ? Non : un théorème, démontré en 1981 par Laurence Kirby et Jeffrey Paris, et basé sur une méthode que ce dernier avait mise au point en 1978 avec Leo Harrington, affirme que, quelle que soit notre imagination, nous n'arriverons jamais à montrer le théorème de Goodstein par récurrence, c'est-à-dire en ne nous servant que des entiers et des quatre opérations élémentaires. Ce résultat est remarquable (et sa démonstration très difficile !) et fait la spécificité du théorème de Goodstein. De façon plus précise, il affirme que la fonction qui, à chaque graine n, associe l'entier p pour lequel $g_p(n)$ s'annule est une fonction qui prend des valeurs tellement grandes qu'elle finit par dépasser n'importe quelle fonction dont on peut montrer l'existence par récurrence.

Comme elle utilise l'ordinal ω, objet infini, la démonstration du théorème de Goodstein n'est pas une preuve par récurrence. Notons qu'il suffit d'ajouter l'hypothèse qu'il existe un objet infini au système de Peano pour obtenir, sans avoir à ajouter de nouveau postulat, toute l'arithmétique des ordinaux (utilisée dans la preuve de Goodstein). Du point de vue des hypothèses logiques sous-jacentes, ceci place le théorème de Goodstein juste au-dessus de l'arithmétique de Peano. Cet ajout représente exactement l'écart entre l'infini potentiel (présent dans l'arithmétique, où le principe de récurrence postule l'existence d'une suite d'entiers sans fin) et l'infini actuel (présent dans les ordinaux, où existe un objet infini).


Kurt Gödel

... Et Gödel

Il existe un rapport étroit entre ce qui précède et les célèbres théorèmes d'incomplétude démontrés par Kurt Gödel au début des années 1930.

Le premier théorème d'incomplétude énonce qu'il existe une propriété d'arithmétique qui est vraie, mais qui ne peut pas être démontrée par

5. HERCULE CONTRE L'HYDRE

Le combat d'Hercule contre l'Hydre représente, comme le théorème de Goodstein, un exemple de propriété qui met en jeu des entiers et ne peut se démontrer sans utiliser l'infini. Hercule doit vaincre l'Hydre, mais, chaque fois qu'il lui coupe une tête, il en repousse de nouvelles ! Les règles du combat sont les suivantes : Hercule coupe une seule tête à chaque coup. S'il coupe une tête directement reliée au corps de l'Hydre, la tête ne repousse pas ; en revanche, si la tête coupée n'est pas directement reliée au corps, alors l'ensemble des têtes situées à côté de la tête coupée, c'est-à-dire provenant de la bifurcation commune au-dessous, se reproduit en *n* nouveaux exemplaires, si Hercule porte son *n*-ième coup du combat. Le nombre de tête de l'Hydre semble tendre vers l'infini irrémédiablement. Pourtant, un théorème énonce que, quelle que soit la forme initiale de l'Hydre et quelle que soit la stratégie d'Hercule, il finira toujours par tuer l'Hydre en coupant toutes les têtes. Comme le théorème de Goodstein, le théorème de l'Hydre se démontre en utilisant des ordinaux et, à nouveau, Kirby et Paris ont prouvé qu'il ne peut pas être démontré dans le système de Peano, c'est-à-dire par récurrence.

PREMIÈRE COUPURE
En brun, la partie excitée par la coupe. Cette partie est dupliquée.

DEUXIÈME COUPURE
En brun, la partie excitée par la deuxième coupe. Cette partie est ajoutée deux fois.

TROISIÈME COUPURE
En brun, la partie excitée par la troisième coupe. La partie excitée est ajoutée trois fois. S'il coupe la tête bleue reliée directement au corps de l'hydre, elle ne repousse pas.

Coupera-t-on toutes les têtes de l'hydre? Oui, et ce résultat se démontre avec les ordinaux.

récurrence. Il ne donne toutefois aucun exemple explicite d'une telle propriété.

Le second théorème d'incomplétude comble cette lacune, en livrant un exemple : c'est une propriété qui code, dans le langage de l'arithmétique, le fait que le principe de récurrence n'est pas contradictoire avec lui-même. Ce résultat, fort remarquable, continue de fasciner après bien des années. Pourtant, du point de vue de l'arithmétique, il nous laisse un peu sur notre faim, car la «propriété qui code dans le langage de l'arithmétique le fait que le principe de récurrence n'est pas autocontradictoire» manque de clarté et ne correspond guère à ce qu'un arithméticien a l'habitude d'appeler une propriété d'arithmétique...

Il a fallu attendre cinquante ans avant de trouver des propriétés d'arithmétique plus communes qui ne soient pas démontrables par récurrence. Le premier exemple a été une propriété de type combinatoire isolée par Paris et Harrington en 1978, suivi du théorème de Goodstein, établi quelques années plus tard. C'est à ce jour l'exemple le plus simple d'une propriété d'arithmétique, qui se démontre à l'aide de l'infini, et dont il est prouvé qu'elle ne peut être démontrée par récurrence.

En un sens, un tel résultat constitue un argument puissant en faveur de l'utilisation de l'infini en mathématiques puisqu'il montre que certaines propriétés ne mettant en jeu que des objets finis ne peuvent être démontrées qu'à partir de l'infini : on aurait donc tort de se priver d'une telle possibilité. D'un autre côté, l'exemple de la graine 4 montre que le théorème de Goodstein met en jeu des entiers gigantesques, beaucoup plus grands par exemple que le nombre d'atomes dans l'univers, et nous pouvons aussi nous interroger sur l'usage de tels entiers et leurs rapports avec les entiers «de tous les jours»...

Patrick DEHORNOY est professeur de mathématiques à l'université de Caen.

G. GOODSTEIN, *On the Restricted Ordinal Theorem*, in *J. Symb. Logic*, n° 9, pp. 33-41, 1944.

L. KIRBY et J. PARIS, *Accessible Independence Results for Peano Arithmetic*, in *Bull. London Math. Soc.*, n° 14, pp. 285-293, 1982.

A. LEVY, *Basic Set Theory*, Springer, 1979.

L'infini et l'univers des algorithmes

GILLES DOWEK

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

Un algorithme décompose un calcul en des suites d'opérations élémentaires. Quand un algorithme prescrit un nombre infini d'opérations élémentaires, on n'obtient jamais le résultat. Aussi, faut-il tenter de maîtriser ou de chasser cet infini.

Pour multiplier 35 par 12, nous devons multiplier 5 par 2, poser le 0, retenir le 1, multiplier 3 par 2, ajouter la retenue… au bout de nos efforts, nous obtiendrons le résultat : 420. Cette recette qui décrit l'enchaînement des opérations élémentaires d'une multiplication est dénommée l'algorithme de la multiplication, ou plutôt un algorithme de la multiplication, car il existe d'autres méthodes de calcul possibles. Ainsi, 35 + 35 + 35 + 35 + 35 + 35 + 35 + 35 + 35 + 35 + 35 + 35 est une méthode, un algorithme envisageable pour multiplier 35 par 12, à l'aide d'additions successives. Il en existe bien d'autres.

Ces algorithmes sont des méthodes de calcul qui donnent un résultat après un nombre fini d'opérations et dans un temps fini. L'infini ne semble ainsi jouer aucun rôle dans l'univers des algorithmes, et pourtant la notion d'algorithme ne peut être définie sans utiliser l'infini. L'infini est indispensable, mais il faut le maîtriser.

La définition d'un algorithme

Si nous utilisons des algorithmes depuis l'Antiquité et que le mot algorithme vient du nom du mathématicien Al Khwarizmi qui vécut à Bagdad au IXe siècle, ce n'est qu'au XXe siècle que cette notion a été définie avec précision.

Dans les années 1920, Thoralf Skolem (1887-1963) a proposé une première définition de la notion d'algorithme liée à la récurrence. Ce procédé consiste à définir une fonction en donnant sa valeur en 0 et en indiquant comment passer de sa valeur en n à sa valeur en $n + 1$. Par exemple, la fonction e, qui à chaque nombre entier n associe le nombre 2^n : $e(0) = 1$, $e(1) = 2$, $e(2) = 4$, $e(3) = 8$, $e(4) = 16$, … est définie par récurrence, en indiquant que sa valeur en 0 est 1 et qu'on passe d'une valeur à la suivante en la doublant.

Les fonctions qui sont définies par leur valeur en 0, ou en une autre valeur initiale, et par une opération permettant de passer directement d'une valeur à la suivante sont appelées, selon un terme dû à Rosza Péter, récursives primitives. La fonction $e(n) = 2^n$ est récursive primitive. Elle est définie par $e(0) = 1$ et $e(n + 1) = 2e(n)$. Richard Dedekind, qui avait déjà étudié les définitions par récurrence en 1888, avait montré que l'addition, la multiplication, l'élévation à la puissance étaient des fonctions récursives primitives.

La définition d'une fonction par récurrence donne un algorithme pour le calcul des valeurs de cette fonction. Ainsi, pour calculer la valeur de la fonction e en 10, il suffit de calculer sa valeur en 0, puis en 1, en 2,…, en 9 et enfin en 10. On obtient ainsi $e(0) = 1$, $e(1) = 2$, $e(2) = 4$, …, $e(9) = 512$, et $e(10) = 1\,024$.

On programme aisément de telles fonctions récursives primitives en utilisant des boucles *For*. Chaque itération de la boucle exécute une opération élémentaire ; on peut déterminer le nombre d'itérations nécessaires au calcul en fonction de la valeur que l'on souhaite calculer. Si l'on demande au programme de calculer la fonction e pour n égal à 10, après 10 itérations, 1 024 s'affichera sur l'écran de l'ordinateur.

Avec cette notion de fonction récursive primitive, l'infini n'était pas convoqué, et les mathématiciens pensaient que ces fonctions suffisaient pour définir l'ensemble des algorithmes. Pourtant, en 1928, Wilhelm Ackermann a trouvé une fonction de deux variables m et n pour laquelle il y avait une recette de calcul, mais qui ne répondait pas au souhait de prévisibilité du nombre de pas de calcul. Cette fonction est définie par : $A(0, n) = n + 1$; $A(m + 1, 0) = A(m, 1)$ et $A(m + 1, n + 1) = A(m, A(m + 1, n))$.

Cette fonction n'est pas récursive primitive : le nombre de pas de calcul nécessaire pour passer d'une valeur de m ou de n à la suivante, $m + 1$ ou $n + 1$ n'est pas connu. Cette impossibilité n'est pas due à une mauvaise formulation de la fonction, mais à une impossibilité théorique démontrée par Ackermann : même si la fonction d'Ackermann peut toujours se calculer en un nombre fini d'étapes, le nombre d'itérations nécessaires au calcul est imprévisible. En informatique, le calcul d'une telle fonction ne peut être réalisé en utilisant des boucles *For*.

Les fonctions récursives primitives peuvent toujours se calculer par un algorithme ; en revanche, les fonctions calculables par un algorithme ne sont pas toujours récursives primitives, comme le prouve la fonction d'Ackermann. La définition de Skolem était donc trop restrictive, et le défi lancé aux mathématiciens, au début des années 1930, était de dépasser cette définition pour en proposer une qui englobe l'ensemble des algorithmes.

Le schéma *mu*

Plusieurs solutions ont été proposées par Jacques Herbrand, Kurt Gödel, Alonzo Church, Stephen Kleene, Alan Turing, Emil Post… Dans la définition la plus simple, avancée par Kleene en 1936, on ajoute aux définitions par récurrence un second procédé de calcul, portant le joli nom de schéma *mu*. Ce schéma nous oblige à continuer le calcul jusqu'à ce qu'une condition soit atteinte. Ainsi, à partir d'une fonction g, Kleene définit un nombre a qui est le plus petit nombre x pour lequel $g(x)$ est égal à 0 (condition recherchée ici). Par exemple, si g est la fonction qui à x associe le nombre $|x^2 - 9|$, 3 est le plus petit nombre x tel que $|x^2 - 9|$ soit égal à 0.

On sait calculer une fonction définie par un schéma *mu* au moyen d'un algorithme, c'est-à-dire que l'on peut calculer la valeur de a, bien que ce calcul soit moins simple que celui par récurrence : pour calculer a, nous devons rechercher le plus petit nombre x tel que $|x^2 - 9|$ soit égal à 0, c'est-à-dire calculer successivement $g(0)$, $g(1)$, $g(2)$, $g(3)$, … jusqu'à une valeur d'annulation. Contrairement au calcul par récurrence classique des fonctions récursives primitives, le nombre d'itérations est imprévisible. Nous pouvons trouver la solution après quelques itérations, ou alors tester un grand nombre de valeurs avant de trouver celle qui convient.

Dans les langages de programmation, cette recherche de la solution s'effectue à l'aide des boucles *While*, qui de la même façon que les boucles *For* exécutent une opération plusieurs fois. Cependant, la boucle *For* n'est adaptée qu'aux fonctions récursives primitives, elle contient l'instruction d'arrêter le calcul lorsque le nombre d'itérations souhaité est atteint. La boucle *While* contient l'instruction d'arrêter le calcul lorsque la condition recherchée est atteinte, ici $g(x) = 0$, et ce quel que soit le nombre d'itérations. Le programme recherche la solution en itérant tant qu'elle (*While*) n'est pas trouvée.

Avec le schéma *mu*, on calcule toutes les fonctions où le nombre d'étapes n'est pas connu *a priori* et où l'on découvre au fur et à mesure la structure du calcul. Le schéma *mu* permet d'élaborer progressivement la structure du calcul de la fonction d'Ackermann *(voir l'encadré)*. Cette notion de découverte progressive était absente de la définition de Skolem.

Les fonctions définies avec ces deux procédés – définitions par récurrence et schéma *mu* – sont calculables par algorithme, et l'on a de bonnes raisons de penser que, réciproquement, toutes les fonctions qui peuvent se calculer par un algorithme peuvent se définir ainsi. Le défi lancé aux mathématiciens par l'introduction de la fonction d'Ackermann était ainsi relevé. Les fonctions récursives primitives et le schéma *mu* englobent l'ensemble des algorithmes.

L'arrivée de l'infini

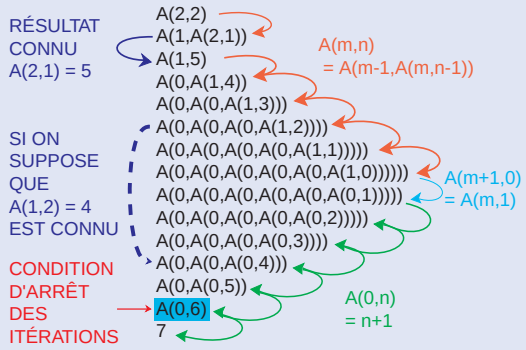
Avec le schéma *mu*, contrairement aux définitions par récurrence, il n'est pas possible de prévoir le nombre d'itérations nécessaires pour trouver le résultat. Il se peut même que l'on ne trouve pas de solution, et, dans ce cas, le calcul se poursuivra éternellement.

Remplaçons dans l'exemple précédent le chiffre 9 par 10, et tentons de calculer le nombre a, qui soit le plus petit nombre x tel que $|x^2 - 10|$ soit égal à 0. Comme aucun nombre ne convient, on essaiera 0, 1, 2, 3, 4, 5… sans jamais trouver la valeur d'annulation. Techniquement, on dit que ce

La fonction d'Ackermann

La fonction d'Ackermann est une fonction parfaitement définie et calculable, mais non récursive primitive. Prenons l'exemple du calcul de la fonction A(2,2) en décomposant le calcul. Si cette fonction était récursive primitive, il nous suffirait de partir de la valeur de A(0,0), puis de calculer A(1,0), A(2,0), A(2,1) et enfin A(2,2). Quatre étapes seraient suffisantes. Le problème principal de la fonction d'Ackermann est que les résultats obtenus pour des valeurs de m et de n inférieures à celles que l'on souhaite calculer ne sont pas suffisants. La décomposition du calcul de la fonction A(2,2), nécessite de calculer des fonctions où les valeurs de n sont supérieures à la valeur initiale. Seul un mécanisme de recherche par schéma mu, où la condition d'arrêt des itérations est l'obtention d'une seule valeur de m qui soit égale à 0, donne le résultat.

D'après la définition d'Ackermann, A(2,2) se décompose en A(1, A(2,1)). D'après l'hypothèse, A(2,1) est connu et égal à 5, nous obtenons alors la fonction A(1,5), dont la valeur est inconnue. On décompose cette fonction jusqu'aux A(0, n) dont les valeurs sont égales à n + 1. Ainsi A(2,2) est égal à 7.



nombre n'est pas défini. Concrètement, le calcul se poursuit à l'infini. C'est ce que Kleene appelle le mouvement perpétuel. Ce mécanisme de recherche à l'infini offert par le schéma *mu* se retrouve dans les boucles *While* des programmes informatiques ; le programme «boucle» tant que $|x^2 - 10|$ est différent de 0.

Traditionnellement, un algorithme était une recette de calcul qui fournissait toujours un résultat après un nombre fini d'opérations ; avec le schéma *mu*, l'univers des algorithmes s'élargit et, dans certains cas, le calcul se prolonge à l'infini et ne donne jamais de résultat. On dénomme parfois de tels algorithmes, des semi-algorithmes.

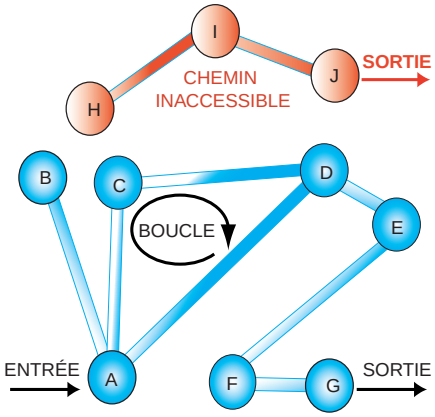
Maîtriser l'infini

Le schéma *mu* permet de définir des semi-algorithmes qui recherchent un certain objet dans un ensemble infini, sans même savoir si un tel objet existe. Un programme informatique qui recherche la sortie dans un labyrinthe est une illustration du schéma *mu*. Par exemple, on peut chercher un chemin qui va de A à G dans le labyrinthe de l'illustration ci-contre, et l'on trouvera peut-être le chemin A, D, E, F, G. Si l'on cherche un chemin allant de A à J, on n'en trouvera pas.

Comment procède-t-on pour trouver le chemin qui va de A à G?

Le nombre de chemins dans un labyrinthe est infini : on peut très bien se perdre en parcourant une boucle, en passant de A en D, de D en C et de C en A, sans jamais prendre le couloir qui va de D vers E. Pour trouver la sortie d'un labyrinthe, il ne suffit pas de faire la liste des chemins en allant toujours droit devant soi ; il faut énumérer les chemins de manière systématique afin de n'en oublier aucun, par exemple en prenant tous les chemins de longueur 1 (qui sont en nombre fini), puis tous ceux de longueur 2 (qui sont également en nombre fini), puis ceux de longueur 3… L'énumération de tous les chemins de tailles finies du labyrinthe permet de trouver la sortie, si celle-ci existe et correspond à un chemin de taille finie, mais énumérera les chemins éternellement s'il n'en existe pas. Ce risque de recherche à l'infini, dans le cas où le problème n'a pas de solution, est inévitable dans un ensemble infini.

Dans le cas du labyrinthe, il est possible d'éviter cette recherche dans un ensemble infini. Pour traverser un labyrinthe, il n'est jamais nécessaire de passer deux fois par le même carrefour ; s'il y a un chemin qui comporte une boucle, on la supprime et l'on obtient un autre chemin qui mène directement au même endroit.



LA RECHERCHE DE LA SORTIE dans un labyrinthe peut être infinie. Il faut donc borner la recherche en excluant, par exemple, les boucles telle que ACDA.

On se restreint ainsi à chercher une solution parmi les chemins sans boucle. Ces chemins sont alors en nombre fini, et s'il n'y a pas de solution (comme quand on essaie d'aller de A à J), après avoir énuméré tous les chemins et constaté qu'aucun d'eux ne convient, on saura que les deux points ne sont pas reliés. Qu'il y ait une solution ou non, la recherche arrivera à son terme.

Restreindre ainsi l'espace de recherche de manière à le rendre fini, sans pour autant perdre la solution, est le problème principal quand on cherche à concevoir certains algorithmes, par exemple des algorithmes de sortie de labyrinthe : cela s'appelle borner le schéma *mu*. Parfois – comme dans ce cas – cela est possible, mais dans d'autres cas, en particulier quand le problème qu'on cherche à résoudre est indécidable, cela ne l'est pas. Lorsque le problème est indécidable, on ne peut pas toujours, après avoir effectué un nombre fini d'étapes, prédire si la solution existe ou non. Le problème demande une recherche dans un ensemble infini, et donc une recherche à l'infini quand il n'y a pas de solution.

Dans certaines situations, on ne peut éviter la recherche à l'infini, et le schéma *mu* est le bon outil. Mais dans d'autres, par exemple quand on cherche à définir la fonction d'Ackermann, dont les étapes sont certes imprévisibles, mais sont toujours finies, le schéma *mu* semble un outil mal adapté, car trop puissant, puisqu'il convoque l'infini dans sa recherche. En introduisant le schéma *mu*, Kleene a fait entrer le loup dans la bergerie : l'infini dans la théorie des algorithmes qui était *a priori* le cadre du fini.

Cela est-il dû à une maladresse de Kleene ou cette intrusion de l'infini dans l'univers des algorithmes était-elle nécessaire?

Chasser l'infini?

Church et Kleene ont montré qu'il ne s'agissait pas d'une maladresse et que cette intrusion était inévitable. En effet, tous les procédés de définition d'algorithmes qui se limitent à définir des fonctions qui donnent toujours un résultat après un nombre fini d'opérations (c'est-à-dire qui excluent ce mécanisme de recherche à l'infini) sont incomplets : on pourra toujours construire une fonction, calculable par un algorithme en un nombre fini d'étapes, mais qui n'est pas définissable par ce procédé. Cette fonction se construit en utilisant la méthode de la diagonale introduite par Georg Cantor pour étudier les cardinaux infinis (voir l'article de Jean-Paul Delahaye, *L'infini est-il paradoxal en mathématiques?*, page 30).

Il n'y a donc pas de définition qui englobe toutes les fonctions donnant un résultat après un nombre fini d'étapes et rien qu'elles. Certains procédés, comme le schéma *mu*, définissent toutes ces fonctions, mais aussi certaines fonctions qui ne donnent pas de résultat après un nombre fini d'étapes. D'autres procédés comme ceux – plus complexes – proposés par Kurt Gödel en 1941 ou Jean-Yves Girard en 1970 ne permettent de définir que des fonctions qui donnent un résultat en un nombre fini d'étapes. Cependant, ils ne les définissent pas toutes.

Chassez l'infini, il reviendra au galop...

Gilles DOWEK est chargé de recherches à l'Institut national de recherche en informatique et en automatique.

Jean-Luc CHABERT et al., *Histoire des algorithmes*, Belin, 1995.

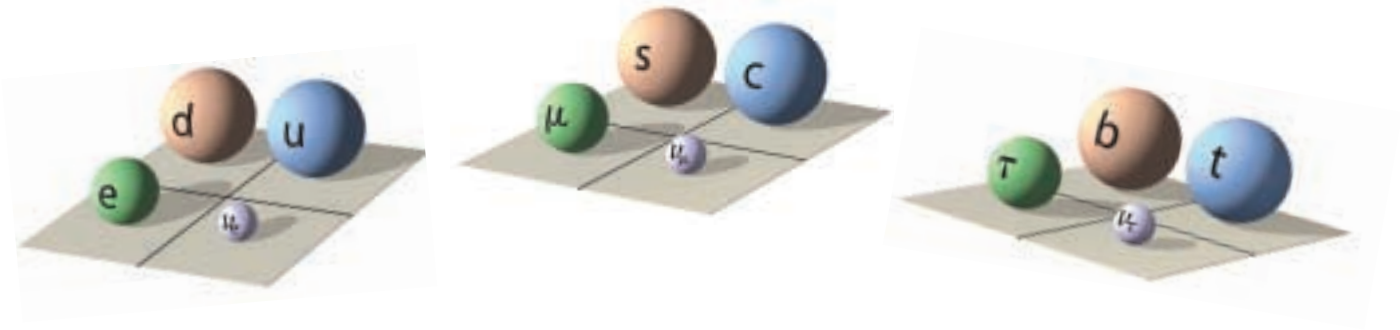
Richard LASSAIGNE et Michel DE ROUGEMONT, *Logique et fondement de l'informatique*, Hermès, 1993.

René CORI, *Logique mathématique*, Masson, 1993.

L'infiniment petit en physique

HARALD FRITZSCH

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

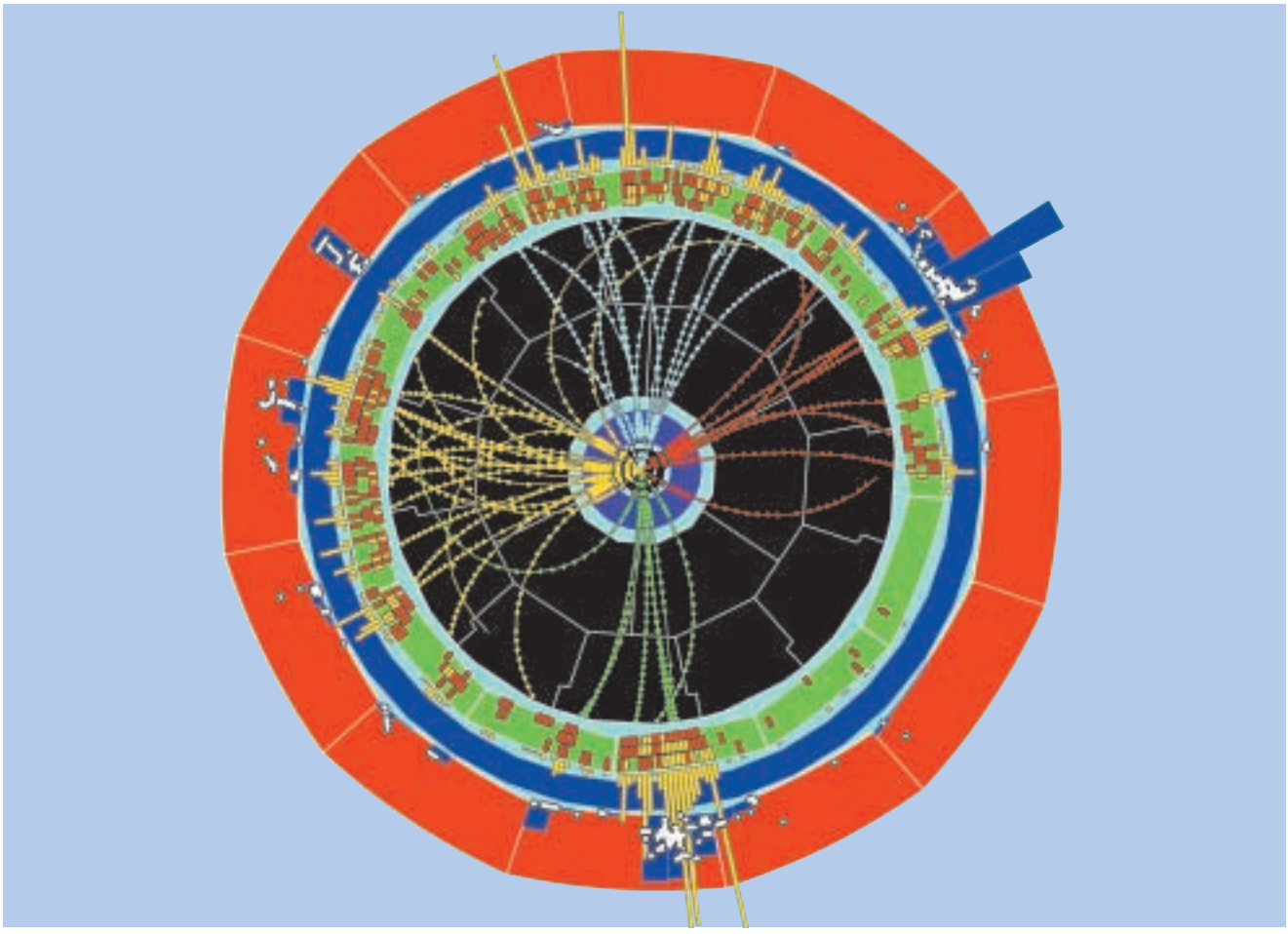


L'infiniment petit est omniprésent en physique des particules. En principe, aucune idée physique ne devrait faire appel à des notions infinies. Pourtant, les théories utilisées aujourd'hui par les physiciens qui étudient l'Univers à petite échelle supposent des constituants élémentaires ponctuels, donc infiniment petits. Les chercheurs n'ont pas adopté cette idée d'emblée. Partant de notre échelle macroscopique, régie par les lois de la physique classique, ils sont descendus d'étage en étage vers les petites dimensions, dans le domaine où règnent les lois quantiques. L'exploration des échelles atomique, puis nucléaire, puis subnucléaire les a confrontés à des comportements de plus en plus surprenants. Au-dessous du millionième de milliardième de mètre, tout se passe comme si les particules qui interagissent étaient infiniment petites. Les physiciens ont donc construit des théories fondées sur l'hypothèse que ces particules – les électrons et les quarks – sont ponctuelles, et qu'elles interagissent de façon ponctuelle. Extrême, cette thèse a d'abord engendré d'immenses difficultés techniques. Toutefois, la première des théories fondée sur elle – l'électrodynamique quantique, qui décrit les interactions entre électrons et photons – devint finalement la plus précise jamais élaborée! Encouragés par ce succès, les physiciens forgèrent des théories sur le même modèle pour décrire les autres interactions. Ils obtinrent ainsi le «modèle standard», qui rend compte avec succès des interactions de toutes les particules. Depuis 30 ans, cet énorme édifice théorique résiste à toutes les tentatives expérimentales pour le mettre en défaut. Les constituants élémentaires sont-ils vraiment ponctuels?

Les physiciens n'ont pas inventé l'infini

Les physiciens n'ont pas inventé l'infini : ils ont commencé à utiliser par nécessité ce concept de mathématicien (voir *Histoire d'infini*, page 64). En mathématique, la notion d'infini apparaît d'une façon «naturelle». On conçoit immédiatement que la suite des nombres naturels est sans fin, donc qu'elle est infinie. Lorsque les mathématiciens disent qu'une quantité est infinie, ils veulent dire qu'elle (sa valeur absolue) est sans limite supérieure ou, ce qui revient au même, que supposer qu'une telle limite existe conduit à une contradiction logique : n ne peut être le plus grand entier, car $n + 1$ lui est supérieur. À l'inverse, quand ils parlent d'une quantité infiniment petite, cela signifie qu'elle (sa valeur absolue) n'a pas de limite inférieure, sinon zéro. La suite des inverses des nombres naturels, 1, 1/2, 1/3, 1/4, ..., ne contient

1. PLUS UN OBJET EST PETIT, plus les instruments nécessaires pour le mettre en évidence sont grands. En bas, le détecteur *Aleph*, à l'aide duquel les physiciens du CERN auraient récemment découvert une nouvelle particule : le boson de Higgs. Si elle se confirme, la découverte est importante, car le boson de Higgs est le dernier élément non découvert prévu par le modèle standard, la théorie la mieux acceptée par les physiciens des particules. Les expérimentateurs du CERN ont constaté que les trajectoires rassemblées dans les faisceaux vert et bleu proviennent de la désintégration d'une grosse particule. Sa masse semble correspondre à celle du boson de Higgs.



CERN

aucune limite inférieure. Pour les mathématiciens, les infinis, tant grands que petits, sont définis grâce à un processus de passage à la limite.

En physique, les situations qui imposent un tel passage ne manquent pas. Au début du XX^{e} siècle, les astronomes se représentaient l'Univers comme un espace plein d'étoiles, dont le Soleil. Puis ils ont découvert que les étoiles étaient les constituants élémentaires de galaxies, et que des milliards de galaxies existent. L'Univers s'était brusquement agrandi de plusieurs ordres de grandeur. Cette évolution invite à se demander quelles structures encore plus grandes pourraient exister. Et finalement, on s'interroge : le Cosmos est-il infini? Ces questions restent ouvertes, même si plusieurs indices militent effectivement pour un Univers sans limites.

Il en va de même lorsqu'on essaie d'explorer l'infiniment petit. L'histoire des sciences révèle que l'homme a progressivement appris à diviser de plus en plus la matière. L'un des premiers à y penser fut Démocrite (460-370 avant notre ère). Le philosophe grec a supputé que la matière peut être divisée jusqu'à une échelle ultime, celle de petits constituants élémentaires invisibles, qu'il nomme *Atomos*. Pour Démocrite, comme pour les physiciens d'aujourd'hui, le processus de division de la matière est donc limité. Bien que très petit, un atome n'est pas infiniment petit. Son diamètre est d'environ un dixième de milliardième de mètre, taille qu'une expérience simple met en évidence : un film huileux répandu sur l'eau s'étire jusqu'à atteindre une hauteur qui correspond au diamètre d'une molécule, voire d'un atome. Or, grâce au microscope à effet tunnel, les physiciens sont capables d'observer des atomes. Que voient-ils? De petites balles de tennis, dont il faudrait 100 millions accolées pour faire un centimètre.

Atomos signifie indivisible

Le mot *Atomos* signifie indivisible en grec ancien. Pourtant le caractère indivisible des atomes a été remis en cause à la fin du XX^{e} siècle, quand les physiciens ont montré que les atomes peuvent être dissociés en constituants plus petits : le noyau et le nuage électronique. La taille de l'enveloppe impose celle des atomes, tandis que sa masse représente moins d'un millième de leur masse. Le plus simple des atomes est

2. WERNER HEISENBERG est l'un des pères de la mécanique quantique. Les physiciens du début du XX^{e} siècle ont inventé cette théorie, notamment pour mieux expliquer la structure des atomes.

l'hydrogène. Son enveloppe ne contient qu'un seul électron, chargé négativement, en rotation autour du noyau, qui porte une charge électrique positive. Sa rotation permanente l'empêche d'être attiré par l'attraction électrique qu'exerce le noyau. L'électron se comporte en fait comme une minuscule planète, dont l'orbite est stable parce que la force centrifuge due à la rotation équilibre l'attraction électrique du noyau.

Lorsqu'ils eurent compris cela, les physiciens s'étonnèrent des particularités des systèmes atomiques : tels une infinité de frères jumeaux, tous les atomes d'hydrogène ont la même structure. Or, dans le modèle planétaire, si l'électron tournait plus vite, il serait plus éloigné du noyau. Pourquoi le noyau de l'atome d'hydrogène ne fixe-t-il qu'un seul et unique électron, toujours à la même distance? Autre mystère : les physiciens savent que les particules chargées accélérées rayonnent de l'énergie sous forme d'ondes électromagnétiques. Or, l'électron est soumis en permanence à une accélération et les particules chargées et accélérées perdent leur énergie en émettant un «rayonnement de freinage». Pourquoi l'électron de l'atome d'hydrogène n'émet-il aucune onde électromagnétique? Quel mécanisme le stabilise sur son orbite? Les physiciens durent se rendre à l'évidence : la physique classique ne suffisait plus à expliquer la stabilité des atomes.

Pour répondre à ces questions, les physiciens du début du XX^{e} siècle ont créé la mécanique quantique. Le concept central de la théorie quantique est celui de la probabilité d'événements. Ainsi, on calcule la probabilité qu'un électron se trouve en un endroit précis, à un instant donné ; qui plus est, la position et la vitesse d'un électron ne peuvent être déterminées simultanément. Le physicien allemand Werner Heisenberg fut le premier à formuler cette propriété quantique. Ces «relations d'incertitude de Heisenberg» s'appliquent à tous les systèmes qui ont un comportement quantique. Le calcul montre toutefois que l'approximation classique est excellente pour les systèmes macroscopiques. Si la probabilité quantique qu'une voiture traverse brusquement le mur du garage où elle était garée n'est pas nulle, elle est si faible qu'en pratique l'événement peut être considéré comme impossible!

Atomes hypothétiques

C'est justement cette incertitude sur la position de l'électron qui détermine la taille d'un atome. Dans le cas de l'atome d'hydrogène, cette taille est égale à 10^{-10} mètre. Supposons que nous observions un atome d'hydrogène, dont l'enveloppe serait beaucoup plus petite, disons cent fois plus petite. L'électron serait alors localisé avec une précision plus grande que dans l'atome d'hydrogène normal. Selon la relation d'incertitude, la précision de la vitesse de l'électron serait dans ce cas 100 fois inférieure, de sorte qu'il devrait se déplacer beaucoup plus vite que dans l'atome normal. Une vitesse plus élevée signifierait aussi une énergie supérieure. Un atome d'hydrogène hypothétiquement plus petit posséderait donc une énergie supérieure à celle d'un atome normal. Or, que ce soit en physique classique ou en mécanique quantique, les systèmes évoluent toujours vers leur état de moindre énergie. Un atome d'hydrogène plus petit que la normale ne serait pas stable, mais perdrait de l'énergie en rayonnant des ondes électromagnétiques : l'électron ralentirait, son enveloppe augmenterait et se stabiliserait lorsque l'atome aurait atteint sa taille «normale». De même, un atome d'hydrogène hypothétique 100 fois plus grand qu'un atome normal ne serait pas stable. Pour le fabriquer, il faudrait éloigner l'électron du noyau, et donc lui apporter de l'énergie, de façon à contrecarrer l'attraction électrostatique. Pour évoluer vers un état de moindre énergie,