L'arithmétique, pourquoi ?

Marc Guinot

L'arithmétique ou théorie des nombres (des nombres entiers s'entend : c'est le sens premier du mot « arithmos » en grec) occupe une place singulière en mathématiques. Aussi ancienne que les mathématiques elles-mêmes, elle n'eut, pendant longtemps, que des adeptes prestigieux mais peu nombreux, qui étaient soit des mathématiciens confirmés, comme Fermat, Euler, Lagrange, Legendre, Gauss ou Dirichlet, soit des amateurs brillants tel Christian Goldbach, correspondant d'Euler. Tout en œuvrant dans bien d'autres domaines des mathématiques, tous manifestèrent avec ardeur leur intérêt pour cette discipline particulière que Gauss qualifiait de « reine des mathématiques » et dont Euler célébrait dans sa correspondance la magnificence (« Herrlichkeit »).

Un des attraits de l'arithmétique réside dans le décalage qui existe entre la simplicité de beaucoup de ses énoncés et la subtilité qui préside à leurs démonstrations. On s'est beaucoup moqué d'Euclide qui crut bon de développer toute une argumentation pour démontrer, en géométrie, l'inégalité du triangle $AC \leq AB + BC$ (Euclide, *Les Éléments*, Livre I, proposition 20) alors que l'âne sait bien que pour aller à la botte de foin qu'il convoite, il n'est pas utile de faire un détour par un autre endroit. L'arithmétique elle-même n'est pas exempte de ce genre de travers et c'est avec le plus grand sérieux que le même Euclide démontre que la somme de plusieurs nombres pairs est encore un nombre pair (*Les Éléments*, Livre IX, proposition 21).

Mais prenons le célèbre théorème selon lequel *il existe une infinité de nombres premiers*. On ne sait pas justifier cette affirmation en donnant directement un exemple de famille infinie de nombres tous premiers comme on le ferait, par exemple, pour expliquer qu'il y a une infinité de nombres impairs. La démonstration astucieuse et évidemment irréfutable, se trouve dans les *Éléments* (Livre IX, proposition 20) et consiste à considérer *a priori* une quantité limitée de nombres premiers $p_1, \ldots, p_r$ et à montrer que cette liste, aussi étendue soit-elle, n'épuise pas tous les nombres premiers possibles, et cela, en donnant un exemple de nombre premier p qui ne figure pas dans la liste. Pour y parvenir, on forme le produit de tous les nombres premiers considérés et on lui ajoute 1. Comme tout nombre entier > 1, le nombre obtenu est soit lui-même un nombre premier (ce qui arrive), soit un multiple d'un nombre premier p (il suffit de considérer le plus petit diviseur $p > 1$ de N). Le nombre premier p mis ainsi en évidence ne figure pas parmi les nombres premiers $p_1, \ldots, p_r$ car s'il y figurait, ce serait un diviseur du produit $p_1 \ldots p_r$ et comme c'est aussi un diviseur de N, ce serait un diviseur de la différence $N - p_1 \ldots p_r$, c'est-à-dire de 1, ce qui est impossible pour un nombre premier normalement constitué. Le résultat est donc démontré.

Les entiers qui sont somme de deux carrés : quelques résultats préliminaires

Dans beaucoup de cas, on ne parvient à démontrer un résultat d'arithmétique qu'au prix de subtilités et de détours qui ont demandé à leurs auteurs des années d'efforts et de méditations. Prenons l'exemple des nombres entiers qui sont susceptibles de se décomposer d'une manière ou d'une autre en somme de deux carrés parfaits, comme 13 (qui est égal à $3^2 + 2^2$) ou 4 294 967 297 (qui est égal à $62\,264^2 + 20\,449^2$). Pour des raisons qui tiennent à l'identité algébrique

$$(x^2 + y^2)\,(x'^2 + y'^2) = (xx' - yy')^2 + (xy' + x'y)^2,$$

un premier point à élucider est de trouver tous les nombres premiers p qui admettent déjà ce genre de décompositions en deux carrés. Une observation sommaire dans la liste des premiers nombres premiers amène à proposer la conjecture suivante : sont décomposables en sommes de deux carrés le nombre premier 2 et les nombres premiers qui, comme le dit Fermat lui-même, « surpassent de l'unité un multiple de 4 ». Ces derniers sont donc les nombres premiers de la forme $4n + 1$ ou, pour employer le langage si commode des congruences (inventé par Gauss plus d'un siècle après Fermat), les nombres premiers p congrus à 1 modulo 4, ce qui peut encore s'écrire $p \equiv 1 \pmod 4$.

Dans sa correspondance, Fermat prétend qu'il est parvenu non sans mal à démontrer le résultat invoqué grâce à une variante subtile de sa méthode (déjà fort astucieuse) de descente infinie. Mais il n'en dit pas plus et Euler qui, sur l'incitation de Goldbach, reprit la question plus de 60 ans après la mort de Fermat, mit de nombreuses années à en venir à bout. Il vaut la peine de voir comment il s'y est pris.

Un premier préalable est de connaître les propriétés particulières des nombres premiers. En dehors du fait qu'un nombre de ce genre n'a pas de diviseurs hormis 1 et lui-même, l'essentiel réside dans la propriété (qu'on admettra ici sans discuter et qu'on trouve déjà dans Euclide : *Éléments*, Livre VII, proposition 30) que *tout nombre premier p qui divise un produit de facteurs divise nécessairement l'un de ces facteurs* (on notera que ce n'est pas le cas d'un nombre aussi simple que 4 qui divise le produit 6×2 sans diviser ni 6 ni 2). Dans le langage des congruences, cette propriété veut dire que la relation $xy \equiv 0 \pmod p$ implique $x \equiv 0 \pmod p$ ou $y \equiv 0 \pmod p$, ce qui est très satisfaisant (et faux quand on remplace le module premier p par un nombre composé tel que 4…).

Une autre propriété cruciale est le « petit théorème de Fermat » selon lequel *si x est un entier quelconque, $x^p - x$ est toujours divisible par p* si p est un nombre premier. Dans le langage des congruences, on a donc $x^p \equiv x \pmod p$. On notera que $2^{341} \not\equiv 2 \pmod{341}$, ce qui prouve au passage que 341 n'est pas premier…

Lorsque x est divisible par p, la congruence en question est évidente. Lorsque x n'est pas divisible par p, la décomposition $x^p - x = x\,(x^{p-1} - 1)$ et la propriété du paragraphe précédent fournissent une variante du théorème de Fermat: *si p est un nombre premier et si x n'est pas divisible par p, alors $x^{p-1} - 1$ est toujours divisible par p*. En d'autres termes, avec ces hypothèses, on a toujours $x^{p-1} \equiv 1 \pmod p$.

Pour démontrer la première des affirmations précédentes (et donc aussi la deuxième), on peut procéder comme l'avait fait Leibniz en son temps en remarquant que lorsqu'on développe $(a + b)^p$ par la formule du binôme, on obtient une somme qui s'écrit $a^p + pa^{p-1}b + \frac{p(p-1)}{2}a^{p-2}b^2 + \ldots + b^p$, avec des termes intermédiaires (extrêmes exclus) affectés de coefficients qui sont tous divisibles par p. En effet, comme ces coefficients, tous entiers, sont $C_p^k = \frac{p!}{k!(p-k)!}$ avec $1 \leq k \leq p-1$, on a la relation $p! = k!(p-k)!C_p^k$ qui montre que p divise le produit $k!(p-k)!C_p^k$. Comme, par hypothèse, p est plus grand que k et que $p - k$, il ne peut diviser ni $k!$ ni $(p - k)!$ comme on l'a vu précédemment, donc, pour la même raison, il divise C_p^k.

On déduit de là que $(a + b)^p \equiv a^p + b^p \pmod{p}$ quels que soient les entiers a et b. Par récurrence, on a aussitôt plus généralement

$$(a + b + \ldots + c)^p \equiv a^p + b^p + \ldots + c^p \pmod{p}$$

si a, b,..., c sont des entiers quelconques. Si, en particulier, on prend $a = b = \ldots = c = 1$, on voit alors que $x^p \equiv x \pmod{p}$, x étant le nombre de termes de la somme $a + b + \ldots + c$. D'où la relation cherchée dans le cas où x est un entier positif. Le cas négatif s'en déduit sans peine.

Les entiers qui sont somme de deux carrés : le noeud de l'affaire

Venons-en aux sommes de deux carrés, en étudiant celles d'entre elles qui s'écrivent $a^2 + b^2$ avec des entiers a et b premiers entre eux, c'est-à-dire sans diviseurs communs autres que 1. Comme l'a montré Euler, cette hypothèse n'est pas sans conséquences sur les facteurs premiers du nombre correspondant car *si un nombre* N *est de la forme* $a^2 + b^2$ *avec des entiers a et b premiers entre eux, alors tous les diviseurs premiers impairs de* N *sont de la forme* $4n + 1$. En parlant des « diviseurs premiers impairs » de N, on exclut expressément le diviseur 2, seul nombre premier pair possible. Raisonnons par l'absurde en supposant que le nombre N précédant admette un diviseur premier impair p qui ne soit pas de la forme voulue. Comme p est impair, il est de la forme $4n + 3$. Posons alors $m = 2n + 1$. De la relation $a^2 + b^2 \equiv 0 \pmod{p}$ qui résulte des hypothèses, on déduit successivement $a^2 \equiv -b^2 \pmod{p}$, $(a^2)^m \equiv (-b^2)^m \pmod{p}$ et $a^{2m} \equiv -b^{2m} \pmod{p}$ puisque m est impair. Comme $2m = 4n + 2 = p - 1$, cette dernière congruence s'écrit $a^{p-1} \equiv -b^{p-1} \pmod{p}$. Comme a n'est pas divisible par p (sinon b le serait aussi à cause de la relation $b^2 = N - a^2$) et que b n'est pas divisible par p (sinon a le serait aussi pour une raison analogue) – le tout en contradiction avec l'hypothèse que a et b sont premiers entre eux –, la congruence obtenue se réduit, grâce au petit théorème de Fermat, à $1 \equiv -1 \pmod{p}$, ce qui est absurde car cela voudrait dire que p divise 2. D'où le résultat.

Et réciproquement ?

Ce résultat admet une sorte de réciproque puisque *si p est un nombre premier de la forme* $4n + 1$, *alors il existe deux entiers a et b premiers entre eux tels que p divise* $a^2 + b^2$.

On s'approche ainsi du résultat final cherché qui veut que p soit en fait un nombre de cette forme (on notera que la relation $p = a^2 + b^2$ implique que a et b sont premiers entre eux).

Le nombre premier p étant de la forme $4n + 1$, considérons deux entiers x et y premiers entre eux et supposés non divisibles par p. Les deux entiers $a = x^n$ et $b = y^n$ sont encore deux entiers premiers entre eux et non divisibles par p. Comme $a^4 = x^{4n} = x^{p-1}$ et $b^4 = y^{4n} = y^{p-1}$, il résulte du théorème de Fermat que $a^4 \equiv b^4 \pmod{p}$, autrement dit que p divise $a^4 - b^4$, c'est-à-dire le produit $(a^2 - b^2)(a^2 + b^2)$ et donc $a^2 - b^2$ ou $a^2 + b^2$. Si p divise $a^2 + b^2$ pour au moins une valeur de x et une valeur de y, le résultat cherché est démontré. Si ce n'est pas le cas, p divise $a^2 - b^2 = x^{2n} - y^{2n}$ quelles que soient les valeurs de x et de y (soumises aux conditions que l'on sait). En particulier, en prenant $y = 1$, p doit diviser $x^{2n} - 1$ quel que soit x non multiple de p.

Pour montrer que cela est absurde (donc que p divise bien $a^2 + b^2$ pour des valeurs convenables de x et de y), Euler fait un petit détour par l'*opérateur de différence* D qui associe à toute fonction f (définie disons sur $\mathbf{R}$) la fonction $g = Df$ (définie sur le même ensemble) telle que $g(x) = f(x + 1) - f(x)$ pour tout x. Bref, on a

$$Df(x) = f(x + 1) - f(x) \text{ quel que soit } x.$$

On vérifie aussitôt que D est un opérateur linéaire, c'est-à-dire que

$$D(f + g) = Df + Dg \text{ et } D(af) = a\,Df \text{ si } a \in \mathbf{R}.$$

En particulier, avec des notations incorrectes, mais compréhensibles, on a

$$D(a_0 + a_1x + \ldots + a_mx^m) = a_0D(1) + a_1D(x) + \ldots + a_mD(x^m)$$

pour n'importe quel polynôme $a_0 + a_1x + \ldots + a_mx^m$. Comme

$$D(1) = 0,\ D(x) = x + 1 - x = 1,\ D(x^2) = (x + 1)^2 - x^2 = 2x + 1,\ \ldots,$$

$$D(x^m) = (x+1)^m - x^m = mx^{m-1} + \frac{m(m-1)}{2}x^{m-2} + \ldots + 1,$$

on voit aussitôt que si $a_m \neq 0$, $D(a_0 + a_1x + \ldots + a_mx^m)$ est un polynôme en x de degré $m - 1$, dont le coefficient dominant (celui du terme de degré $m - 1$) est ma_m.

On peut bien sûr répéter cette opération plusieurs fois, ce qui donne des opérateurs D^2, D^3, ... pour lesquels

$$D^2f(x) = f(x + 2) - f(x + 1) - f(x + 1) + f(x)) = f(x + 2) - 2\ f(x + 1) + f(x),$$

$$D^3f(x) = f(x + 3) - f(x + 2) - 2f(x + 2) + 2f(x + 1) + f(x + 1) - f(x)$$

$$= f(x + 3) - 3f(x + 2) + 3f(x + 1) - f(x)\ldots$$

Un raisonnement par récurrence, laissé au lecteur, montre que $D^mf(x)$ est une combinaison linéaire à coefficients entiers de $f(x), f(x + 1), \ldots, f(x + m)$.

D'autre part, si on applique D^2 à un polynôme de degré m, on trouve 0 si $m \leq 1$ et sinon un polynôme de degré $m - 2$, dont le coefficient dominant est $m\,(m - 1)\,a_m$.

On a des résultats analogues pour D^3, D^4, etc. En particulier, si m est un entier ≥ 1, $D^m(x^m)$ et $D^m(x^m - 1)$ se réduisent à la constante $m!$.

Il résulte de tout cela qu'en posant $f(x) = x^m - 1$, on a

$$m! = c_0 f(x) + c_1 f(x+1) + \ldots + c_m f(x+m)$$

où $c_0, c_1, \ldots, c_m$ sont des coefficients entiers. En particulier, avec $x = 1$, on a

$$m! = c_0 f(1) + c_1 f(2) + \ldots + c_m f(m+1).$$

Si on prend pour m le nombre $2n$ introduit plus haut il résulte de l'hypothèse faite sur $x^{2n} - 1$ que p divise $f(1)$, $f(2)$, ... , $f(4n)$, donc a fortiori $f(1)$, $f(2)$,..., $f(2n+1) = f(m)$ (car $2n + 1 < 4n$), et donc p divise $m!$ Mais cela n'est pas possible puisque $m = 2n < 4n - 1 = p$. CQFD.

Le Saint-Graal

Pour accéder au Saint-Graal, il reste un dernier résultat à démontrer : *si* N *est un entier de la forme* $a^2 + b^2$ *et* q *un diviseur premier de* N *de la forme* $x^2 + y^2$, *alors* N/q *peut être mis sous la forme* $u^2 + v^2$.

Comme $N = a^2 + b^2$ et $q = x^2 + y^2$, on a

$$Nq = (a^2 + b^2)(x^2 + y^2) = (ax - by)^2 + (ay + bx)^2$$

d'après une identité algébrique signalée plus haut. En fait, comme on peut remplacer sans dommages y par $-y$, on a

$$Nq = (ax \pm by)^2 + (ay \mp bx)^2 \qquad (*)$$

Comme q divise N, q divise

$$Ny^2 - b^2 q = (a^2 + b^2) y^2 - b^2 (x^2 + y^2) = a^2 y^2 - b^2 x^2 = (ay - bx)(ay + bx).$$

Comme q est premier, il divise alors $ay - bx$ ou $ay + bx$. S'il divise $ay - bx$, il divise aussi $Nq - (ay - bx)^2$, c'est à dire $(ax + by)^2$ d'après (*) et donc finalement $ax + by$. S'il divise $ay + bx$, il divise $Nq - (ay + bx)^2$, c'est à dire $(ax - by)^2$ et donc aussi $ax - by$. Bref, il existe un nombre ε égal à ± 1 tel que q divise à la fois $ay + \varepsilon bx$ et $ax - \varepsilon by$. On peut donc écrire $ax - \varepsilon by = qu$ et $ay + \varepsilon bx = qv$ où u et v sont deux entiers. Donc

$$Nq = (ax - \varepsilon by)^2 + (ay + \varepsilon bx)^2 = q^2 u^2 + q^2 v^2,$$

d'où la relation

$$\frac{N}{q} = u^2 + v^2$$

qu'il nous fallait démontrer.

Nous sommes maintenant à même de prouver entièrement le théorème attendu dont l'énoncé revient à Fermat et la démonstration à Euler : *tout nombre premier p de la forme* $4n + 1$ *est une somme de deux carrés.*

Le raisonnement pourrait être présenté comme une récurrence portant sur p (c'est peut-être comme cela qu'il faut comprendre Fermat quand il disait pouvoir appliquer à ce problème sa méthode de descente). Mais il est plus simple de raisonner par

l'absurde en supposant qu'il existe un nombre premier p de la forme $4n + 1$ qui ne soit pas somme de deux carrés et en choisissant p le plus petit possible. Comme p est de la forme $4n + 1$, il existe deux nombres a et b premiers entre eux, tels que p divise $a^2 + b^2$. Comme on l'a déjà noté, a et b ne peuvent pas être divisibles par p. Effectuons alors la division euclidienne de a et de b par p : cela permet d'écrire

$$a = kp + r \text{ avec } 0 < r < p, \text{ soit } r \equiv a \pmod{p},$$

$$b = lp + s \text{ avec } 0 < s < p \text{ soit } s \equiv b \pmod{p}.$$

Si $r < \frac{p}{2}$, posons $r' = r$. Si $r > \frac{p}{2}$ posons $r' = p - r$ (on notera que $\frac{p}{2}$ n'est pas un entier). On obtient ainsi un entier r' congru à r, donc à a, modulo p tel que $0 < r' < \frac{p}{2}$. Définissons de même, à partir de s, un entier s' tel que $s' \equiv s \equiv b$ avec $0 < s' < \frac{p}{2}$. Comme $r'^2 + s'^2 \equiv r^2 + s^2 \equiv a^2 + b^2 \equiv 0 \pmod{p}$, on voit que p divise $r'^2 + s'^2$. Soit g le PGCD de r' et de s'. Alors $r' = gr''$ et $s' = gs''$ où r'' et s'' sont des entiers premiers entre eux, vérifiant encore $0 < r'' < \frac{p}{2}$ et $0 < s'' < \frac{p}{2}$. Il est clair que p ne divise pas g, mais comme il divise $r'^2 + s'^2 = g^2 (r''^2 + s''^2)$, il divise nécessairement $r''^2 + s''^2$. Bref, en changeant hardiment de notations, on peut conclure à l'existence de deux entiers a et b premiers entre eux qui vérifient $0 < a < \frac{p}{2}$ et $0 < b < \frac{p}{2}$ et tels que p divise $a^2 + b^2$.

Soit $N = a^2 + b^2$. On peut alors écrire $N = a^2 + b^2 < \frac{p^2}{4} + \frac{p^2}{4} < p^2$, c'est à dire $0 < \frac{N}{p} < p$. Il en résulte que les facteurs premiers $q_1, \ldots, q_n$ (avec des répétitions éventuelles) de $\frac{N}{p}$ sont tous inférieurs à p et on a $N = p\, q_1 \ldots q_n$.

On peut alors affirmer que q_i est toujours une somme de deux carrés. C'est évident si $q_i = 2$ car $2 = 1^2 + 1^2$ et cela résulte de ce que nous avons vu plus haut et du choix de p, si q_i est impair. Nous pouvons maintenant appliquer le résultat ci-dessus qui ouvre la voie du Saint Graal, autant de fois que nécessaire : N/q_1 est une somme de deux carrés $u_1^2 + v_1^2$, $N/(q_1q_2)$ est une somme de deux carrés $u_2^2 + v_2^2$, et ainsi de suite jusqu'à $N/(q_1q_2 \ldots q_n)$ qu'on peut donc écrire $u_n^2 + v_n^2$. Mais cela est absurde car $N/(q_1q_2 \ldots q_n) = p$. Cette contradiction clôt la démonstration.

Wilson au secours d'Euler

Le décalage entre l'énoncé d'un théorème et sa démonstration illustré par ce qui précède, a pour conséquence qu'un même théorème a souvent d'autres démonstrations, de nature tout à fait différente. Ainsi, le laborieux raisonnement

d'Euler, passant par l'opérateur de différence, pour démontrer que tout nombre premier p de la forme $4n + 1$ divise une somme de deux carrés premiers entre eux peut être obtenu très rapidement et d'une toute autre manière grâce au *théorème de Wilson* selon lequel *pour tout nombre premier p, on a* $(p - 1)! \equiv -1$ (modulo p). Admettons en effet provisoirement ce résultat, supposons p de la forme $4n + 1$ et posons $m = 2n$ (ce qui fait $p = 2m + 1$). Ecrivons le produit $(p - 1)!$ sous la forme :

$$(p - 1)! = 1 \cdot 2 \ldots m\,(m + 1) \ldots (m + m)$$

ou mieux

$$(p - 1)! = (m + m)\,(m + m - 1) \ldots (m + 1)\,1 \cdot 2 \ldots m.$$

Puisque $p = 2m + 1$, il est clair que l'on a

$$m + m \equiv -1,\ m + m - 1 \equiv -2,\ \ldots,\ m + 1 \equiv -m \pmod{p}.$$

Par suite, le produit précédent, c'est à dire $(p - 1)!$, est congru, modulo p, à

$$(-1^2)\,(-2^2) \ldots (-m^2) = (-1)^m\,(m!)^2,$$

c'est à dire $(m!)^2$ puisque m est pair.
Comme p divise $(p - 1)! + 1$ d'après le théorème de Wilson, p divise aussi $(m!)^2 + 1$, c'est-à-dire une somme de carrés premiers entre eux…

Reste évidemment à démontrer le théorème de Wilson lui-même. Il y a, là encore, plusieurs façons de faire. L'une des plus simples consiste à vérifier que si a est un des nombres $1, 2, \ldots, p - 1$, il existe un nombre a' et un seul, figurant lui aussi dans la liste $1, 2, \ldots, p - 1$, tel que

$$aa' \equiv 1 \pmod{p}$$

On peut appeler a' l'inverse de a modulo p.

L'unicité de a' découle des propriétés usuelles des nombres premiers, car si a' et a'' sont deux entiers tels que $aa' \equiv 1 \pmod{p}$ et $aa'' \equiv 1 \pmod{p}$, on a aussi $aa' \equiv aa'' \pmod{p}$, ce qui veut dire que p divise $aa' - aa'' = a\,(a' - a'')$.
Comme p ne peut diviser a (car on a supposé $1 \leq a \leq p - 1$), p divise $a' - a''$. Mais si on suppose en outre $1 \leq a' \leq p - 1$ et $1 \leq a'' \leq p - 1$, on a

$$-(p - 2) \leq a' - a'' \leq p - 2$$

Comme le seul nombre compris entre $-(p - 2)$ et $p - 2$ qui soit divisible par p est 0, on a $a' - a'' = 0$, donc $a' = a''$.

L'existence de a' est une conséquence évidente du théorème de Bézout (que nous admettrons ici et qui en réalité était déjà connu de Bachet de Méziriac)) qui veut que *pour deux entiers a et b premiers entre eux, il existe des entiers x et y tels que* $ax + by = 1$. Appliqué au nombre a ci-dessus et au nombre premier p (qui sont premiers entre eux), il donne une relation du type $ax + py = 1$, qui implique $ax \equiv 1 \pmod{p}$, c'est-à-dire le résultat, avec $a' = x$.

Il est clair que a joue vis-à-vis de a' le rôle que a' joue vis-à-vis de a. Cela veut dire que $(a')' = a$. En d'autres termes, l'application $a \mapsto a'$ de l'ensemble $E = \{1, \ldots, p - 1\}$ dans lui-même est une transformation involutive (ou une involution).Bien entendu, il se peut que cette transformation ait des « points fixes », c'est-à-dire qu'il existe des éléments a de E tels que $a' = a$. Cela veut dire en fait que

$a^2 \equiv 1 \pmod{p}$. Comme cette congruence signifie que p divise $a^2 - 1 = (a-1)(a+1)$, elle veut dire aussi que p divise $a-1$ ou p divise $a+1$. Comme $1 \leq a \leq p-1$, le premier cas ne peut avoir lieu que pour $a = 1$ et le second pour $a = p-1$. Ce sont donc les deux seuls points fixes. Si on suppose $p \geq 5$ (les cas $p = 2$, $p = 3$ peuvent être traités à part), il est intéressant de restreindre l'application $a \to a'$ à l'ensemble $F = \{2, \ldots, p-2\}$. On vérifie aisément que si $a \in F$, alors $a' \in F$, de sorte qu'on a une nouvelle involution, cette fois de F, qui est sans point fixe.

Si on regroupe chaque élément a de F avec son inverse a∉, on partage F en $\frac{p-3}{2}$ classes différentes, ayant chacune deux éléments distincts. Si on procède à ce regroupement avec les facteurs du produit 2 . 3 … (p - 2), on obtient alors $\frac{p-3}{2}$ facteurs, tous congrus à 1 modulo p (puisque aa∉ ∫ 1). D'où

$$(p-1)! = 1 \cdot 2 \cdot 3 \ldots (p-2)\,(p-1) \equiv 1\,(p-1) = p-1 \equiv -1 \pmod{p}.$$

Le théorème de Wilson est démontré.

De mieux en mieux

On pourrait donner d'autres exemples de démonstrations multiples dans le domaine des sommes de deux carrés. La plus spectaculaire, découverte récemment, consiste à démontrer directement le principal résultat qui a fait l'objet des développements précédents sans passer par aucune des étapes intermédiaires décrites ci-dessus. Curieusement ce nouveau raisonnement fait intervenir les propriétés des transformations involutives vues dans la partie précédente et, en particulier, le résultat auquel on était parvenu en substance selon lequel *si on enlève d'un ensemble fini* E *les points fixes d'une involution f de* E, *le nombre d'éléments restants est pair.* Corollairement, *s'il existe dans* E *une involution ayant un nombre impair de points fixes, alors* le nombre d'éléments de E est lui aussi impair et surtout *toute involution de* E *admet au moins un point fixe.*

Considérons dans ces conditions un nombre premier p de la forme $4n + 1$ et appelons E l'ensemble des triplets $(x,y,z) \in \mathbf{N}^3$ tels que $x^2 + 4yz = p$. Il est clair que l'application $(x,y,z) \mapsto (z,y,x)$ est une involution de E et que tout point fixe de cette involution fournit une relation du type $x^2 + 4y^2$, donc une décomposition de p en sommes de deux carrés. Pour démontrer alors l'existence de ce genre de décomposition, il suffit, d'après ce qui précède, de trouver une seconde involution de E, ayant un nombre impair de points fixes.

Cette involution est l'application f définie par

$$(x,y,z) \mapsto \begin{cases} (x+2z, z, y-x-z) & \text{si } x < y-z, \\ (2y-x, y, x-y+z) & \text{si } y-z < x < 2y, \\ (x-2y, x-y+z, y) & \text{si } x > 2y. \end{cases}$$

Sans entrer dans tous les détails de la vérification, notons d'abord que dans tous les cas $f(x,y,z) \in \mathbf{N}^3$. Ensuite, si on appelle A, B, C les ensembles de triplets $(x,y,z) \in E$

tels que l'on ait $x < y - z, y - z < x < 2y, x > 2y$ respectivement, alors A, B et C forment une « partition » de E (ce sont trois ensembles deux à deux disjoints dont la réunion est E) ; on notera que les relations $x = y - z$ et $x = 2y$ sont impossibles, de même d'ailleurs que $xyz = 0$.

Puis pour vérifier que f est bien une involution, on montrera que f applique A dans C, B dans B et C dans A. Quant aux points fixes éventuels, ils ne peuvent être que dans B. Comme ce sont alors ceux pour lesquels $2y - x = x$ et $x - y + z = z$ on doit avoir $x = y$, donc $x^2 + 4xz = p$. Comme p est premier, on a nécessairement $x = 1$ et donc $y = 1$, et $z = \frac{p-1}{4}$. Ces relations fournissent effectivement un point fixe, le seul possible. CQFD.

Une petite pause du coté de l'infinité des nombres premiers

Pour ceux que les sommes de carrés commencent à fatiguer, revenons au problème tout différent de l'infinité des nombres premiers, déjà abordé dans l'introduction. Beaucoup d'autres méthodes, d'autres démonstrations ont été proposées depuis Euclide. En voici une, trouvée dans un numéro de l'*American Mathematical Monthly* (mars 1971, p. 272). Elle consiste à s'intéresser à la « série » $\frac{1}{p}$ où p parcourt l'ensemble des nombres premiers (attention, compte tenu de notre propos, il peut s'agir d'un ensemble fini…) et parallèlement à la « série » $\frac{1}{q}$ où q parcourt l'ensemble (fini ou non !) des nombres sans facteurs carrés, c'est-à-dire des nombres (entiers positifs) qui ne sont divisibles par aucun carré > 1.

En fait, il est facile de montrer que cette dernière série comporte nécessairement une infinité de termes et, même mieux, que c'est une série divergente. Pour cela, on commence par remarquer que tout entier $n \geq 1$ peut s'écrire sous la forme d'un produit qc^2 où q est un nombre sans facteurs carrés et c un entier ≥ 1. Cette écriture est d'ailleurs unique, mais on n'a pas besoin de le savoir pour la suite. Si on développe alors un produit de la forme $\left(\sum_{q \leq N} \frac{1}{q}\right)\left(\sum_{c \leq N} \frac{1}{c^2}\right)$ où N est un entier ≥ 1 fixé (et où q est sans facteurs carrés !), on obtient, entre autres choses, au moins une fois, tous les inverses $\frac{1}{n}$ où $1 \leq n \leq N$. On a donc l'inégalité

$$\left(\sum_{q \leq N} \frac{1}{q}\right)\left(\sum_{c \leq N} \frac{1}{c^2}\right) \geq \sum_{n \leq N} \frac{1}{n}.$$

La somme du second membre peut être rendue aussi grande que l'on veut (divergence de la série harmonique) et au contraire la somme $\sum_{n \leq N} \frac{1}{n^2}$ est bornée car,

à cause de la relation $\frac{1}{n^2} < \frac{1}{n(n-1)} = \frac{1}{n-1} - \frac{1}{n}$, on peut écrire

$$\sum_{n\leq N} \frac{1}{n^2} < 1+1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\ldots+\frac{1}{N-1}-\frac{1}{N} = 2-\frac{1}{N} < 2.$$

La série $\sum \frac{1}{q}$ ne peut donc être que divergente. On peut alors en déduire que la série $\sum \frac{1}{p}$ est aussi divergente. Pour cela on fait intervenir l'inégalité élémentaire $e^x > 1 + x$ valable pour tout $x > 0$. Si la série des inverses des nombres premiers convergeait vers un nombre a, on aurait (avec un entier $N \geq 1$ quelconque)

$$e^a > e^{\sum_{p\leq N} \frac{1}{p}} = \prod_{p\leq N} e^{\frac{1}{p}} > \prod_{p\leq N}\left(1+\frac{1}{p}\right)$$

Mais si on développe le produit final, on obtient entre autres tous les nombres $\frac{1}{q}$ avec $q \leq N$ et q sans facteurs carrés. D'où

$$e^a > \prod_{p\leq N}\left(1+\frac{1}{p}\right) \geq \sum_{q\leq N} \frac{1}{q},$$

ce qui est absurde car la somme obtenue dépasse, comme on l'a vu, n'importe quel nombre donné à l'avance.

Ainsi, non seulement il y a une infinité de nombres premiers p mais en plus on a $\sum \frac{1}{p} = +\infty$, résultat assez étonnant qui fut démontré pour la première fois par Euler en 1737.

Comme disait Gauss, « la diversité des méthodes aide beaucoup à jeter du jour sur les points les plus obscurs ». Gauss savait de quoi il parlait, lui qui a donné quatre démonstrations différentes du théorème fondamental de l'algèbre et huit de la loi de réciprocité quadratique !

Ultime retour sur les entiers somme de deux carrés

Tout ce foisonnement fait évidemment le bonheur de l'amateur, éclairé ou non. S'il est un peu au-dessus de la moyenne, il peut lui-même intervenir en démontrant des résultats ou en simplifiant des raisonnements qui ont pu échapper à l'attention de ses prédécesseurs plus ou moins illustres. L'histoire de la théorie des nombres en offre de multiples exemples.

Citons seulement le cas d'un certain L. Aubry qui fit paraître en 1912 dans une revue bizarrement appelée *Sphinx-Œdipe* une démonstration élémentaire sur les

sommes de carrés qu'auraient bien aimé connaître Fermat et Euler lorsqu'ils faisaient leurs propres recherches. C'est une autre façon de démontrer un résultat déjà étudié précédemment sans passer par l'astuce, pas facile à trouver, du signe à utiliser à bon escient dans la formule :

$$(x^2 + y^2)\,(x'^2 + y'^2) = (xx' \pm yy')^2 + (xy' \mp x'y)^2.$$

Rappelons qu'il s'agit de démontrer que, si q divise une somme N de deux carrés $a^2 + b^2$ et si q lui-même est une somme de deux carrés $x^2 + y^2$, alors $\dfrac{N}{q}$ est à son tour une somme de deux carrés $u^2 + v^2$. À cause de la formule rappelée ci-dessus (et avec un seul signe), on a

$$\frac{N}{q} = \frac{a^2+b^2}{x^2+y^2} = \frac{(a^2+b^2)(x^2+y^2)}{(x^2+y^2)^2} = \left(\frac{ax-by}{x^2+y^2}\right)^2 + \left(\frac{ay+bx}{x^2+y^2}\right)^2,$$

ce qui veut dire que N/q est une somme de deux carrés « in fractis » comme disait Euler. Peut-on en déduire que c'est une somme de deux carrés « in integris » ? C'est ce qu'affirme le théorème d'Aubry, qu'on peut présenter sous une forme agréablement géométrique en disant que *si un cercle du plan* $\mathbf{R}^2$, *d'équation* $x^2 + y^2 = N$ *(avec* N *entier* ≥ 1*), passe par un point rationnel* (i.e. un point à coordonnées rationnelles), *alors il passe aussi par un point entier* (i.e. un point à coordonnées entières).

On partira d'un résultat préliminaire simple selon lequel *pour tout point* $A = (x,y)$ *du plan* $\mathbf{R}^2$, *il existe un point entier* $A' = (x',y')$ *dont la distance à* A *est plus petite que* 1. Il suffit en fait d'appeler x' l'entier le plus proche de x (égal à la partie entière de $x + 1/2$) et y' l'entier le plus proche de y car on a alors $|x - x'| \leq 1/2$ et $|y - y'| \leq 1/2$, d'où

$$AA' = \sqrt{(x-x')^2 + (y-y')^2} \leq \sqrt{\frac{1}{4} + \frac{1}{4}} = \frac{\sqrt{2}}{2} < 1.$$

Cela étant, considérons un cercle d'équation $x^2 + y^2 = N$ (N entier ≥ 1) passant par un point rationnel A_0. Si A_0 est un point entier, il n'y a rien à démontrer. Sinon, considérons un point entier A'_0 tel que $A_0 A'_0 < 1$. Comme $A_0 \neq A'_0$, on peut considérer la droite $A_0 A'_0$, qui coupe le cercle en A_0, et en un « autre » point A_1 (qu'on prendra égal à A_0 si d'aventure la droite est tangente au cercle). Si A_1 est un point entier, on arrête. Sinon, on recommence en considérant un point entier A'_1 du plan tel que $A_1 A'_1 < 1$ et la droite $A_1 A'_1$ qui recoupe le cercle en un point A_2. Si A_2 est un point entier, on arrête sinon on recommence.

Tout le problème est de montrer que le processus ainsi décrit ne peut pas se poursuivre indéfiniment. Pour le voir, on commence par vérifier que tous les points A_i précédents sont rationnels. Il suffit de procéder par récurrence en vérifiant que lorsque A_i est un point rationnel non entier, alors A_{i+1} est encore un point rationnel. Sinon, posons $A_i = (x,y) = (m/d,n/d)$ où m, n, d sont des entiers et où $d \geq 2$. Comme

A'_i est supposé entier, on peut poser simplement $A'_i = (x',y') = (m',n')$ avec m' et n' entiers. Dans ces conditions, les points de la droite $A_i A'_i$ sont les points du plan de la forme $\left(m' + t\left(\frac{m}{d} - m'\right), n' + t\left(\frac{n}{d} - n'\right)\right)$ où t est un paramètre variable. La recherche des points de cette droite qui appartiennent au cercle $x^2 + y^2 = N$ conduit immédiatement à une équation du second degré en t qui s'écrit

$$\left[\left(\frac{m}{d} - m'\right)^2 + \left(\frac{n}{d} - n'\right)^2\right] t^2 + 2\left[m'\left(\frac{m}{d} - m'\right) + n'\left(\frac{n}{d} - n'\right)\right] t + m'^2 + n'^2 - N = 0.$$

Cette équation a une solution simple $t = 1$ qui correspond en fait à $A_i = (m/d, n/d)$, alors que l'autre solution

$$t = \frac{m'^2 + n'^2 - N}{\left(\frac{m}{d} - m'\right)^2 + \left(\frac{n}{d} - n'\right)^2}$$

(qu'on écrira $t = \frac{M}{D^2}$ où $M \in \mathbf{Z}$ et où D est la distance $A_i A'_i$) correspond au point A_{i+1}. Cette expression prouve bien que les coordonnées de A_{i+1} sont deux nombres rationnels, à savoir $m' + \frac{M}{D^2}\left(\frac{m}{d} - m'\right)$ et $n' + \frac{M}{D^2}\left(\frac{n}{d} - n'\right)$. Mais le carré D^2 vaut en fait $N + m'^2 + n'^2 - 2\frac{mm' + nn'}{d}$ comme on le vérifie aussitôt. C'est donc une fraction qu'on peut écrire $\frac{d_1}{d}$ avec d_1 entier > 0. Si on reporte cette fraction dans les expressions donnant les coordonnées de A_{i+1}, celles-ci se simplifient pour donner deux nombres s'écrivant $\frac{m_1}{d_1}$ et $\frac{n_1}{d_1}$ avec m_1 et n_1 entiers. Enfin comme $D^2 < 1$, on a nécessairement $d_1 < d$.

On voit alors que si le processus décrit ci-dessus se poursuivait indéfiniment, on aurait une suite illimitée d'entiers > 0, d_0, d_1, d_2, d_3, ... telle que $d_0 > d_1 > d_2 > d_3 > \ldots$, ce qui est absurde (impossibilité d'une descente infinie portant sur des entiers positifs). D'où le résultat.

Pour notre plaisir à tous

Si on désespère de rivaliser avec L. Aubry, on peut se transformer en amateur d'histoire des mathématiques, spécialité arithmétique, car, comme on vient de le voir, on peut faire de l'arithmétique en se référant aux grands noms du passé. C'est

d'autant plus facile que de 1636 à 1837 la théorie des nombres a une histoire assez simplissime qui se résume aux noms de Fermat, Euler, Lagrange, Legendre, Gauss et Dirichlet. Mais l'amateur peut remonter au Moyen Age, en se procurant par exemple le *Liber quadratorum* de Fibonacci ou en s'initiant aux mathématiques arabes par l'intermédiaire de Roshdi Rashid, ou même à l'Antiquité grecque puisque Fermat lui-même puisait son inspiration dans les *Arithmétiques* de Diophante. S'il a du courage, il peut s'attaquer à l'imposant XIX^{e} siècle où les découvertes en en arithmétique abondent (de Gauss, avec ses *Disquisitiones arithmeticae* de 1801 à Hilbert avec son célèbre *Zahlbericht* de 1897) et, pourquoi pas, au siècle tout juste défunt… même si on peut trouver plutôt déprimante la lecture du *Journal of Number Theory* ou des *Acta Arithmetica*.

Mais l'amateur lambda peut aussi, comme Lagrange à la fin de sa vie, se contenter « de jouir sur cette matière, comme sur plusieurs autres, du fruit des veilles d'autrui ». Sans aller jusqu'à éprouver ce plaisir si singulier qui accompagne, selon André Weil, la découverte chez les chercheurs (Weil va jusqu'à le comparer au plaisir érotique, en plus durable !), on peut faire son miel en arithmétique de beaucoup de bonnes choses. Outre les sommes de carrés ou la théorie des nombres premiers, bien d'autres domaines d'étude sont accessibles pour l'amateur de bonne volonté. On peut ainsi lire des merveilles sur les fractions continues, les nombres transcendants, le grand théorème de Fermat ou la fonction zêta. Même de nos jours, il se publie des choses propres à satisfaire l'amateur comme le montrent les exemples cités plus haut. Mais il faut faire le tri et ce n'est pas si simple, même si on dispose des « *Reviews in Number Theory* » de l'American Mathematical Society. En fait, ce qui manque, il faut bien le dire, c'est un Journal d'arithmétique pour amateurs (JAPA). Il faudrait des bonnes volontés et un éditeur. Qu'est-ce qu'on attend ?

Éléments bibliographiques

Peter BUNDSCHUH, *Einführung in die Zahlentheorie*, Springer-Verlag, Berlin Heidelberg, 1988

Marc GUINOT, *Les « resveries » de Fermat (Arithmétique pour amateur, Livre II)*, Aléas Editeur, Lyon, 1993

Michel DEMAZURE, *Cours d'algèbre*. Cassini, 1997.

HARDY, WRIGHT, *An Introduction to the Theory of Numbers*, 5^{e} éd., Oxford University Press, 1975.

H.E. ROSE (quelle manie, ces prénoms réduits à des initiales !), *A Course in Number Theory*, Oxford University Press, Oxford, 1988

André WEIL, *Number Theory, An approach through history, From Hammurapi to Legendre*, Birkhäuser, Bâle, 1983