Une équation fonctionnelle

Daniel PERRIN

1 La question

Existe-t-il une fonction f de $\mathbf{N}$ dans $\mathbf{N}$ telle que $f \circ f(n) = n^2$ pour tout $n \in \mathbf{N}$?

2 Solution

2.1 Théorème. *Il existe une infinité d'applications $f : \mathbf{N} \to \mathbf{N}$ vérifiant $f \circ f(n) = n^2$ pour tout $n \in \mathbf{N}$.*

Démonstration. La preuve repose sur le lemme [1] suivant :

2.2 Lemme. *Il existe une suite d'entiers $a_0 < a_1 < \cdots < a_n < \cdots$ tels que $\mathbf{N}$ soit la réunion disjointe des ensembles $A_n = \{a_n, a_n^2, a_n^4, \ldots, a_n^{2^k}, \ldots\}$.*

Démonstration. (du lemme) On prend $a_0 = 0$ (donc $A_0 = \{0\}$), $a_1 = 1$ (donc $A_1 = \{1\}$), puis $a_2 = 2$, de sorte que A_2 est l'ensemble des puissances de 2 dont l'exposant est lui-même une puissance [2] de 2 : $2, 4, 16, 256, 65536, \ldots$. La construction se fait ensuite par récurrence sur $n \geq 2$ en prenant pour a_{n+1} le plus petit entier non contenu dans la réunion $A_0 \cup A_1 \cup \cdots \cup A_n$. (Il existe de tels entiers, par exemple un nombre premier distinct de $a_2, \ldots, a_n$).

Notons d'abord que l'on a $a_n < a_{n+1}$. En effet, a_{n+1} est différent des a_i, $i \leq n$ (sinon il serait dans A_i). S'il est $\leq a_n$ il est donc strictement compris entre a_i et a_{i+1}, mais c'est une contradiction avec la construction de a_{i+1} comme le plus petit entier non contenu dans $A_0, A_1, \ldots, A_i$.

Montrons ensuite que $\mathbf{N}$ est réunion disjointe des A_n pour $n \in \mathbf{N}$.

D'abord il est réunion. Sinon il existe $a \in \mathbf{N}$ qui n'est pas dans la réunion. On choisit le plus petit a ainsi, tous les $b < a$ sont dans la réunion donc sont dans $A_{n_0}, \ldots, A_{n_k}$ avec $n_0 < n_1 < \ldots < n_k$. Mais si a n'est pas dans la réunion des A_n pour $n \leq n_k$, c'est le plus petit entier ainsi et la construction précédente montre qu'on a $a = a_{n_{k+1}}$.

La réunion est disjointe. En effet, si l'on a deux ensembles de puissances $A_k = \{a_k, a_k^2, \ldots\}$ et $A_l = \{a_l, a_l^2, \ldots\}$ avec $k < l$ (donc $a_k < a_l$), qui ont un point commun : $a_k^{2^m} = a_l^{2^n}$, en extrayant les racines carrées successives de ces nombres on voit qu'on a $m \geq n$, puis $a_l = a_k^{2^{m-n}}$, ce qui montre que a_l est dans A_k. Mais cela contredit la construction de a_l.

1. Pour une interprétation en termes d'orbites, voir le paragraphe suivant.
2. Les nombres de Fermat diminués de un.

On peut alors prouver le théorème. On pose $f(0) = 0$ et $f(1) = 1$ (on peut aussi prendre $f(0) = 1$ et $f(1) = 0$). On regroupe ensuite les a_n, pour $n \geq 2$, deux par deux, par exemple[3] a_{2p} et a_{2p+1}, et l'on définit $f(a_{2p}) = a_{2p+1}$, $f(a_{2p}^{2^k}) = a_{2p+1}^{2^k}$, puis $f(a_{2p+1}) = a_{2p}^2$, $f(a_{2p+1}^{2^k}) = a_{2p}^{2^{k+1}}$. On a $f \circ f(a_{2p}^{2^k}) = f(a_{2p+1}^{2^k}) = a_{2p}^{2^{k+1}} = \left(a_{2p}^{2^k}\right)^2$ et de même avec les a_{2p+1}, de sorte que f satisfait bien la condition de l'énoncé.

2.3 *Remarque.* Si l'on prend, par exemple, $a_{2p} = 2$ et $a_{2p+1} = 3$, le schéma de l'application f sur $A_{2p} \cup A_{2p+1}$ est le suivant :

$$\begin{array}{ccccccccccc} 2 & & 4 & & 16 & & 256 & & ... \\ \downarrow & \nearrow & \downarrow & \nearrow & \downarrow & \nearrow & \downarrow & \nearrow & \\ 3 & & 9 & & 81 & & 6561 & & ... \end{array}$$

et on voit bien que $f \circ f$ consiste, sur chaque ligne, à se déplacer d'un pas vers la droite, donc est l'élévation au carré.

3 Généralisation

Sur ce sujet, la question fondamentale est de déterminer les applications $g : \mathbf{N} \to \mathbf{N}$ qui sont de la forme $f \circ f$. Je me contente de traiter ici le cas où g est injective.

3.1 Orbites associées à g

3.1 Proposition-Définition. *Soit $g : \mathbf{N} \to \mathbf{N}$ une application injective. La relation $\mathscr{R}$ définie sur $\mathbf{N}$ par $x\mathscr{R}y \iff \exists u \in \mathbf{N}, \exists m, n \in \mathbf{N}$, $x = g^m(u)$ et $y = g^n(u)$ est une relation d'équivalence. Les classes d'équivalence associées sont appelées* **orbites** *de g.*

Démonstration. La relation est réflexive, comme on le voit en prenant $x = y = u$ et $m = n = 0$. Elle est clairement symétrique. Montrons qu'elle est transitive. Si l'on a $x\mathscr{R}y$ et $y\mathscr{R}z$, il existe u, m, n comme ci-dessus avec x, y et v, p, q tels que $g^p(v) = y$ et $g^q(v) = z$. Supposons par exemple $p \geq n$. On a $y = g^n(u) = g^n(g^{p-n}(v))$. Comme g^n est injective, on en déduit $u = g^{p-n}(v)$ donc $x = g^{m+p-n}(v)$ et on a $x\mathscr{R}z$.

3.2 *Remarques.* 1) On note que si x est dans une orbite Ω tous ses transformés par les g^n y sont aussi.

2) Si g est bijective, les orbites au sens de 3.1 ne sont rien d'autre que les orbites sous le sous-groupe engendré par g.

3. Il y a une infinité de façons de faire cela, donc une infinité de f convenables.

3.2 Classification des orbites

3.3 Proposition. *Soit $g : \mathbf{N} \to \mathbf{N}$ une application injective. Les orbites associées à g sont de l'un des types suivants :*

1) Des orbites finies, de la forme $\Omega(a) = \{a, g(a), \ldots, g^{n-1}(a)\}$ avec $n \in \mathbf{N}^*$. Sur une telle orbite g agit par translation de 1 sur les exposants, vus comme éléments de $\mathbf{Z}/n\mathbf{Z}$.

2) Des orbites "de type $\mathbf{N}$", $\Omega(a) = \{a, g(a), g^2(a), \ldots, g^n(a), \ldots\}$ où a n'est pas dans l'image de g. Sur une telle orbite g agit par translation de 1 sur les exposants.

3) Des orbites "de type $\mathbf{Z}$", $\Omega(a) = \{g^n(a) \mid n \in \mathbf{Z}\}$, contenues dans l'image de g. Sur une telle orbite g agit par translation de 1 sur les exposants.

Démonstration. 1) Si l'orbite de a est finie, elle contient tous les $g^n(a)$ pour $a \in \mathbf{N}$. Il existe donc p et q avec $p < q$ tels que $g^p(a) = g^q(a)$. Comme g est injective, on a $g^{q-p}(a) = a$. Je dis qu'alors l'orbite est formée de $a, g(a), \ldots, g^{q-p-1}(a)$. En effet, si b est dans l'orbite, il existe u, m, n avec $a = g^m(u)$ et $b = g^n(u)$. Mais le même argument que ci-dessus montre qu'il existe k tel que $g^k(u) = u$ et on peut supposer k arbitrairement grand. On en déduit $g^k(u) = g^{k-m+m}(u) = g^{k-m}(a)$, ce qui montre que u est de la forme $g^n(a)$ et on a le résultat. Le complément sur l'action de g est alors évident.

2) Supposons que l'orbite Ω contient a qui n'est pas dans l'image de g. Je dis qu'alors elle est formée des $g^n(a)$ pour $n \in \mathbf{N}$. Il est clair que ces éléments sont dans Ω. Inversement, si b est dans Ω, il existe u, m, n avec $a = g^m(u)$ et $b = g^n(u)$. Mais comme a n'est pas dans l'image de g, l'unique possibilité est $m = 0$, donc $a = u$ et on a le résultat.

3) Enfin, supposons que Ω est dans l'image de g. On choisit $a \in \Omega$. Il est clair que les $g^n(a)$ avec $n \in \mathbf{N}$ sont dans Ω, mais aussi les $g^{-n}(a)$ avec $n > 0$ à cause de la relation $g^n(g^{-n}(a)) = a$ qui montre qu'on a $a \mathscr{R} g^{-n}(a)$. Le même argument montre que l'orbite est formée exactement de ces éléments.

3.4 *Remarques.* 1) Voici un exemple d'application admettant une (unique) orbite de type $\mathbf{Z}$: on définit $g(2n) = 2n+2$ pour tout $n \in \mathbf{N}$, $g(2n+1) = 2n-1$ pour $n \geq 1$ et $g(1) = 0$.

2) Il est très facile d'obtenir des exemples qui cumulent des orbites de différents types, en nombre infini si l'on veut. On peut en effet trouver dans $\mathbf{N}$ une infinité d'ensembles disjoints avec une relation d'ordre du type de $\mathbf{N}$ (par exemple les puissances des différents nombres premiers) et il suffit de bâtir une application du type voulu sur chacun de ces ensembles.

Le point fondamental, s'agissant des orbites, est le suivant :

3.5 Proposition. *Soit $g : \mathbf{N} \to \mathbf{N}$ une application injective. On suppose que g s'écrit $f \circ f$. Pour $a, n \in \mathbf{N}$ on a $f(g^n(a)) = g^n(f(a))$. Si $\Omega(a)$ est une orbite de g on a $f(\Omega(a)) \subset \Omega(f(a))$, précisément :*

1) Si $\Omega(a)$ est finie de cardinal n, $\Omega(f(a))$ aussi et on a $f(\Omega(a)) = \Omega(f(a))$.

2) Si $\Omega(a)$ est de type $\mathbf{Z}$, $\Omega(f(a))$ aussi et on a $f(\Omega(a)) = \Omega(f(a))$.

3) Si $\Omega(a)$ est de type $\mathbf{N}$, $\Omega(f(a))$ aussi et on a $f(\Omega(a)) \subset \Omega(f(a))$ (mais l'inclusion peut être stricte).

Démonstration. Comme $g = f \circ f$, la première assertion est évidente, les deux expressions étant égales à $f^{2n+1}(a)$. Comme f et g commutent, si a est dans l'image de g, $f(a)$ aussi car, si $a = g(u)$, on a $f(a) = f(g(u)) = g(f(u))$, et de même avec g^n.

Soit $\Omega(a)$ une orbite de g. Elle est de l'une des formes vues en 3.3.

1) Si $\Omega(a)$ est finie de cardinal n, tout élément de $\Omega(a)$ est de la forme $g^k(a)$ avec $k \in \mathbf{N}$ et, comme on a $f(g^k(a)) = g^k(f(a))$, il en résulte qu'on a $f(\Omega(a)) \subset \Omega(f(a))$. De plus, on a $f(g^n(a)) = f(a) = g^n(f(a))$, de sorte que $\Omega(f(a))$ est finie, de cardinal $\leq n$ et comme f est injective, les cardinaux sont égaux.

2) Si $\Omega(a)$ est de type $\mathbf{Z}$, et si $b \in \Omega(a)$ est de la forme $b = g^n(a)$ avec $n \in \mathbf{N}$ la formule montre que $f(g^k(a))$ est dans $\Omega(f(a))$. Si b est de la forme $b = g^{-n}(a)$ avec $n \in \mathbf{N}$, on a $g^n(b) = a$ donc $f(a) = f(g^n(b)) = g^n(f(b))$, ce qui assure que $f(b)$ est bien dans l'orbite de $f(a)$. On a donc $f(\Omega(a)) \subset \Omega(f(a))$ ce qui montre que $\Omega(f(a))$ est infinie.

Comme a est dans l'image de g^n pour tout $n \in \mathbf{N}$ il en est de même de $f(a)$ comme on l'a vu. Cela montre que l'orbite de $f(a)$ est de type $\mathbf{Z}$. De plus, si b est dans l'orbite de $f(a)$, il existe $n \in \mathbf{Z}$ tels que $b = g^n(f(a)) = f(g^n(a))$. Cela montre que l'inclusion est une égalité.

3) Si $\Omega(a)$ est de type $\mathbf{N}$, on peut supposer que a n'est pas dans l'image de g. Le même argument que ci-dessus montre qu'on a $f(\Omega(a)) \subset \Omega(f(a))$, de sorte que $\Omega(f(a))$ est infinie. Si elle est de type $\mathbf{Z}$, pour tout $n \geq 2$, $f(a)$ s'écrit $g^n(u)$. En appliquant f on a $g(a) = f(g^n(u)) = g^n(f(u))$ donc $a = g^{n-1}(f(u))$ est dans l'image de g ce qui est absurde. L'orbite $\Omega(f(a))$ est donc de type $\mathbf{N}$. L'inclusion $f(\Omega(a)) \subset \Omega(f(a))$ peut être stricte comme le montre l'exemple de l'orbite $\Omega(3)$ dans 2.3.

3.3 Étude des translations

3.6 Lemme. *1) Une permutation circulaire $g = (a_1, \ldots, a_n)$ est le carré d'une permutation de $a_1, \ldots, a_n$ si et seulement si n est impair.*

2) L'application $g : \mathbf{N} \to \mathbf{N}$ donnée par $g(n) = n + 1$ n'est pas le carré d'une application de $\mathbf{N}$ dans $\mathbf{N}$.

3) L'application $g : \mathbf{Z} \to \mathbf{Z}$ donnée par $g(n) = n + 1$ n'est pas le carré d'une application de $\mathbf{Z}$ dans $\mathbf{Z}$.

Démonstration. 1) Posons $X = \{a_1, \ldots, a_n\}$. Si $n = 2k + 1$ est impair on a $g^{2k+1} = \mathrm{Id}_X$ donc $g = g^{-2k} = (g^{-k})^2$. Si n est pair, la signature de g comme permutation de X est négative et g ne peut être le carré d'une permutation de X.

2) Supposons qu'on a $g = f \circ f$ et posons $a = f(0)$. En vertu de 3.5, on a, pour tout $n \in \mathbf{N}$, $g^n(f(0)) = f(g^n(0))$, donc $g^n(a) = a + n = f(n)$, donc $f^2(n) = n + 2a$, et en particulier $f^2(0) = g(0) = 1 = 2a$ ce qui est absurde.

3) Le calcul est identique à celui du point 2).

3.4 Le théorème

3.7 Théorème. *Soit $g : \mathbf{N} \to \mathbf{N}$ une application injective. Il existe $f : \mathbf{N} \to \mathbf{N}$ telle que $g = f \circ f$ si et seulement si les orbites de g qui sont finies de cardinal pair ou de type $\mathbf{N}$ ou de type $\mathbf{Z}$ sont, soit en nombre fini pair, soit en nombre infini.*

Démonstration. Supposons d'abord que la condition de parité du nombre d'orbites est remplie. On construit f en la définissant sur les orbites de g. Sur les orbites finies de cardinal impair on trouve f grâce à 3.6. Pour les orbites de type fini pair ou $\mathbf{N}$ ou $\mathbf{Z}$, on associe deux par deux les orbites de même type[4] et il reste à montrer le lemme suivant :

3.8 Lemme. *Soient $\Omega(a)$ et $\Omega(b)$ deux orbites de même type sur lesquelles g agit par translation des exposants. Il existe f, défini sur $\Omega(a) \cup \Omega(b)$, telle que l'on ait $f \circ f = g$.*

Démonstration. (du lemme) Si $\Omega(a)$ est formée des $g^n(a)$ et $\Omega(b)$ des $g^n(b)$ on définit f par $f(g^n(a)) = g^n(b)$ et $f(g^n(b)) = g^{n+1}(a)$ et on a bien $f \circ f = g$.

Si la condition n'est pas réalisée, il y a un nombre impair d'orbites de l'un des types. Si $g = f \circ f$ on a vu en 3.5 que f permute les orbites de g en conservant le type. Comme elles sont en nombre (fini) impair, l'une au moins est stable par f (et évidemment par g). Sur cette orbite on devrait donc avoir $g = f \circ f$, mais cela contredit 3.6.

4. Si le nombre d'orbites est infini (dénombrable) on les met en bijection avec $\mathbf{N}$ et on associe les orbites de rangs $2n$ et $2n + 1$.