

Représentations hyperbinaires, arbre de Calkin-Wilf et énumération des rationnels

Marc Lorenzi

9 juin 2026

Résumé

Des dénombrements, des arbres, une suite récurrente, quel est le lien entre tous ces objets ? Nous allons voir qu'en fin de compte, tout revient au même. L'arbre de Calkin-Wilf est un arbre binaire dont les étiquettes des nœuds sont en bijection avec les nombres rationnels positifs. Le nombre de décompositions *hyperbinaires* d'un entier est relié à l'étiquette des nœuds de cet arbre. Nous terminerons cet article par la description d'une suite récurrente très simple dont l'ensemble des valeurs prises est $\mathbb{Q}_+^*$, l'ensemble des rationnels strictement positifs.

1 Décompositions hyperbinaires d'un entier

Tout entier naturel s'écrit de façon unique comme une somme de puissances de 2 distinctes. Une *décomposition binaire* de l'entier n est une suite *strictement croissante* $(a_0, \dots, a_p)$ d'entiers naturels telle que

$$n = \sum_{k=0}^p 2^{a_k}$$

C'est le principe de l'écriture des entiers en base 2. Qu'en est-il si on autorise deux puissances de 2 égales (mais pas plus de deux) ?

Définition 1.0.1. Soit $n \in \mathbb{N}$. Une *décomposition hyperbinaire* de n est une suite croissante $(a_0, \dots, a_p)$ d'entiers naturels telle que au plus deux termes successifs de la suite sont égaux et

$$n = \sum_{k=0}^p 2^{a_k}$$

Notons $H(n)$ l'ensemble des décompositions hyperbinaires de n et $h(n)$ le cardinal de $H(n)$. Le tableau ci-dessous donne les premières valeurs de $H(n)$ et $h(n)$. La décomposition binaire usuelle de n est à chaque fois le dernier élément de la liste.

n	$H(n)$	$h(n)$
0	$\{()\}$	1
1	$\{(0)\}$	1
2	$\{(0, 0), (1)\}$	2
3	$\{(0, 1)\}$	1
4	$\{(0, 0, 1), (1, 1), (2)\}$	3
5	$\{(0, 1, 1), (0, 2)\}$	2
6	$\{(0, 0, 1, 1), (0, 0, 2), (1, 2)\}$	3
7	$\{(0, 1, 2)\}$	1
8	$\{(0, 0, 1, 2), (1, 1, 2), (2, 2), (3)\}$	4

FIGURE 1 – Les premiers ensembles $H(n)$ et leur cardinal

Pour les curieux, $h(100) = 19$ et

$$\begin{aligned}
H(100) = & \{(0, 0, 1, 2, 2, 3, 4, 5, 5), (0, 0, 1, 2, 2, 3, 4, 6), (0, 0, 1, 3, 3, 4, 5, 5), (0, 0, 1, 3, 3, 4, 6), \\
& (0, 0, 1, 4, 4, 5, 5), (0, 0, 1, 4, 4, 6), (0, 0, 1, 5, 6), (1, 1, 2, 2, 3, 4, 5, 5), (1, 1, 2, 2, 3, 4, 6), \\
& (1, 1, 3, 3, 4, 5, 5), (1, 1, 3, 3, 4, 6), (1, 1, 4, 4, 5, 5), (1, 1, 4, 4, 6), (1, 1, 5, 6), \\
& (2, 3, 3, 4, 5, 5), (2, 3, 3, 4, 6), (2, 4, 4, 5, 5), (2, 4, 4, 6), (2, 5, 6)\}
\end{aligned}$$

Proposition 1.0.1. *Pour tout $n \in \mathbb{N}$,*

$$h(2n+1) = h(n) \quad \text{et} \quad h(2n+2) = h(n) + h(n+1)$$

Démonstration. Soit $n \in \mathbb{N}$. Soit $(b_0, \dots, b_q) \in H(2n+1)$. Comme $2n+1$ est impair, $b_0 = 0$. De plus, $b_1 \neq 0$ car au plus deux des b_i sont égaux à 0. Ainsi,

$$(b_0, \dots, b_q) = (0, a_0 + 1, \dots, a_{q-1} + 1)$$

Où $(a_0, \dots, a_q)$ est une suite croissante d'entiers naturels telle que au plus deux a_i successifs sont égaux. De plus,

$$2n+1 = \sum_{k=0}^q 2^{b_k} = 1 + 2 \sum_{k=0}^{q-1} 2^{a_k}$$

d'où

$$n = \sum_{k=0}^{q-1} 2^{a_k}$$

Ainsi, $(a_0, \dots, a_{q-1}) \in H(n)$. Inversement, si $(a_0, \dots, a_p) \in H(n)$, alors $(0, 1 + a_0, \dots, 1 + a_p) \in H(2n+1)$. Les ensembles $H(n)$ et $H(2n+1)$ sont donc en bijection, d'où $h(2n+1) = h(n)$.

Soit $(b_0, \dots, b_q) \in H(2n+2)$. Deux cas se présentent.

- Cas 1, $b_0 = 0$. Comme $2n+2$ est pair, on a aussi $b_1 = 0$. Ainsi,

$$(b_0, \dots, b_q) = (0, 0, a_0 + 1, \dots, a_{q-2} + 1)$$

où $(a_0, \dots, a_{q-2})$ est une suite croissante d'entiers naturels telle que au plus deux a_i successifs sont égaux. De plus,

$$2n + 2 = \sum_{k=0}^q 2^{b_k} = 2 + 2 \sum_{k=0}^{q-2} 2^{a_k}$$

d'où

$$n = \sum_{k=0}^{q-2} 2^{a_k}$$

Ainsi, $(a_0, \dots, a_{q-2}) \in H(n)$. Inversement, si $(a_0, \dots, a_p) \in H(n)$, alors $(0, 0, 1+a_0, \dots, 1+a_p) \in H(2n+2)$.

- Cas 2, $b_0 \geq 1$. Écrivons

$$(b_0, \dots, b_q) = (a_0 + 1, \dots, a_q + 1)$$

où $(a_0, \dots, a_q)$ est une suite croissante d'entiers naturels telle que au plus deux a_i successifs sont égaux. De plus,

$$2n + 2 = \sum_{k=0}^q 2^{b_k} = 2 \sum_{k=0}^q 2^{a_k}$$

d'où

$$n + 1 = \sum_{k=0}^q 2^{a_k}$$

Ainsi, $(a_0, \dots, a_q) \in H(n+1)$. Inversement, si $(a_0, \dots, a_p) \in H(n+1)$, alors $(1+a_0, \dots, 1+a_p) \in H(2n+2)$.

On a ainsi une bijection entre $H(2n+2)$ et l'union (disjointe) $H(n) \cup H(n+1)$, d'où $h(2n+2) = h(n) + h(n+1)$. $\square$

Remarque. Pour résumer notre démonstration, on a pour tout $n \in \mathbb{N}$,

$$\begin{aligned} H(2n+1) &= \{(0, 1+a_0, \dots, 1+a_p) : (a_0, \dots, a_p) \in H_n\} \\ H(2n+2) &= \{(0, 0, 1+a_0, \dots, 1+a_p) : (a_0, \dots, a_p) \in H_n\} \\ &\quad \cup \{(1+a_0, \dots, 1+a_p) : (a_0, \dots, a_p) \in H_{n+1}\} \end{aligned}$$

La suite $(h(n))_{n \in \mathbb{N}}$ est appelée la *suite de Stern* ou aussi la *suite de Stern-Brocot*. Elle apparaît (avec un tout petit décalage dû au terme initial) dans l'incontournable On-Line Encyclopedia of Integer Sequences sous le numéro A002487.

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$h(n)$	1	1	2	1	3	2	3	1	4	3	5	2	5	3	4	1	5	4	7	3	8

FIGURE 2 – Les premières valeurs de $h(n)$

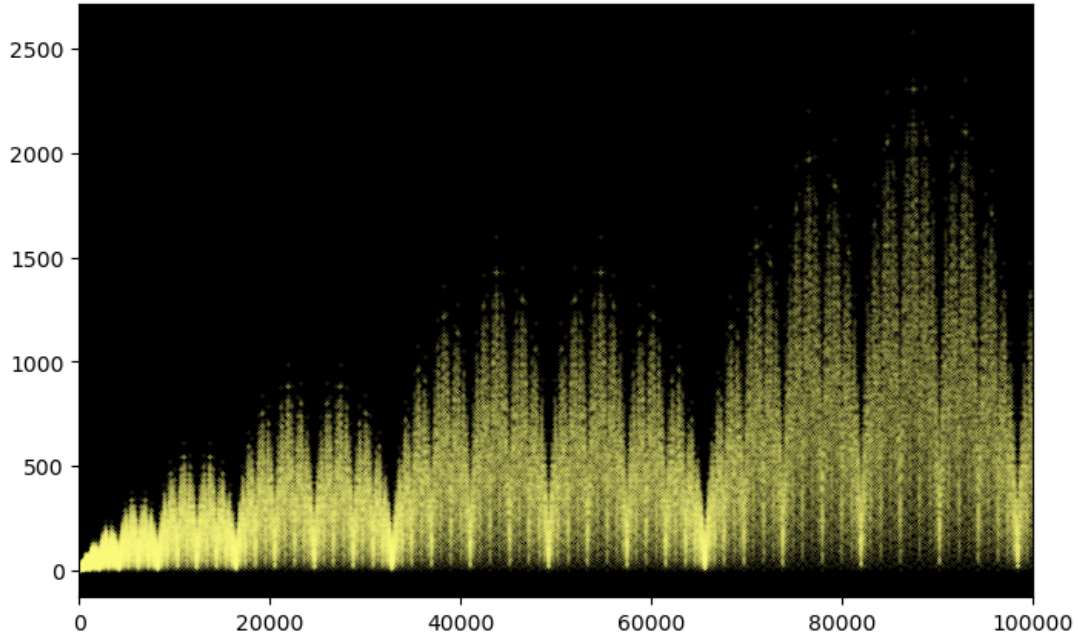


FIGURE 3 – La suite de Stern

2 L'arbre de Calkin-Wilf

L'arbre de Calkin-Wilf est l'arbre binaire $\mathcal{T}$ défini comme suit. L'ensemble des nœuds de $\mathcal{T}$ est $\mathcal{N} = \{0, 1\}^*$, l'ensemble des mots sur l'alphabet $\{0, 1\}$. La racine de $\mathcal{T}$ est le mot vide. Si $\nu \in \mathcal{N}$, alors le fils gauche de ν est le mot $\nu 0$ et son fils droit est le mot $\nu 1$. La *profondeur* d'un nœud ν de $\mathcal{T}$ est la longueur du mot ν .

L'*étiquette* $\varepsilon(\nu)$ du mot ν est définie par induction sur les mots : si ν est le mot vide, alors l'étiquette de ν est $(1, 1)$. Pour tout mot $\nu \in \{0, 1\}^*$, si $\varepsilon(\nu) = (a, b)$, alors $\varepsilon(\nu 0) = (a, a + b)$ et $\varepsilon(\nu 1) = (a + b, b)$. On vérifie facilement par récurrence sur la profondeur des nœuds que si $(a, b) \in \mathcal{E}$, alors $a, b \in \mathbb{N}^*$ ¹. Nous noterons $\mathcal{E} \subseteq \mathbb{N}^* \times \mathbb{N}^*$ l'ensemble des étiquettes des nœuds de $\mathcal{T}$.

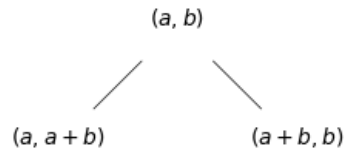


FIGURE 4 – La règle de formation des étiquettes

1. Toutes les présentations de l'arbre de Calkin-Wilf donnent pour étiquettes des nœuds les quotients a/b . Une fraction ne possède pas une écriture unique, alors qu'un couple si. Dit autrement, si je connais (a, b) alors je connais a/b . En revanche, si je connais a/b alors je ne peux pas retrouver a et b sans informations supplémentaires.

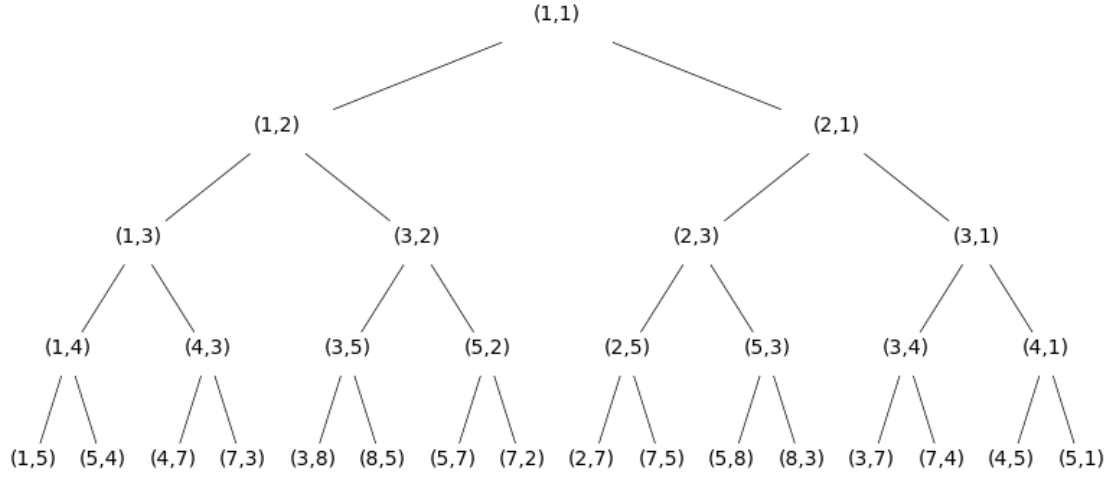


FIGURE 5 – Les premiers niveaux de l'arbre de Calkin-Wilf

Proposition 2.0.2. *Pour tout $(a, b) \in \mathcal{E}$, $\text{pgcd}(a, b) = 1$.*

Démonstration. Prouvons le résultat par récurrence sur la profondeur des nœuds. La racine de $\mathcal{T}$ a pour étiquette $(1, 1)$, et $\text{pgcd}(1, 1) = 1$. Soit $n \in \mathbb{N}$. Supposons la propriété vérifiée pour tous les nœuds de profondeur n . Soit (c, d) l'étiquette d'un nœud ν de profondeur $n + 1$. Soit (a, b) l'étiquette du père μ de ν .

- Si ν est le fils gauche de μ , alors $c = a$ et $d = a + b$. Comme $\text{pgcd}(a, b) = 1$, on a aussi $\text{pgcd}(c, d) = 1$.
- Si ν est le fils droit de μ , alors $c = a + b$ et $d = b$. Comme $\text{pgcd}(a, b) = 1$, on a aussi $\text{pgcd}(c, d) = 1$.

□

Définissons sur $\mathbb{N}^* \times \mathbb{N}^*$ une relation binaire $\preceq$ en posant $(a, b) \preceq (c, d)$ si et seulement si $b < d$ ou $(b = d \text{ et } a \leq c)$. La relation $\preceq$ est l'ordre *lexicographique*.

Propriété facile 2.0.3. *La relation $\preceq$ est un bon ordre sur $\mathbb{N}^* \times \mathbb{N}^*$.*

Démonstration. On vérifie facilement que $\preceq$ est une relation d'ordre. Montrons qu'il s'agit d'un bon ordre. Soit E une partie non vide de $\mathbb{N}^* \times \mathbb{N}^*$. L'ensemble

$$B = \{b \in \mathbb{N}^* : \exists a \in \mathbb{N}^*, (a, b) \in E\}$$

est une partie non vide de $\mathbb{N}^*$, donc B possède un plus petit élément b_0 . L'ensemble

$$A = \{a \in \mathbb{N}^* : (a, b_0) \in E\}$$

est une partie non vide de $\mathbb{N}^*$, donc A possède un plus petit élément a_0 . On a clairement $(a_0, b_0) \in E$. Soit $(a, b) \in E$. Comme $b \in B$, on a $b_0 \leq b$. Si $b_0 > b$, alors $(a_0, b_0) \preceq (a, b)$. Si $b_0 = b$, alors $a \in A$, donc $a \geq a_0$, d'où $(a_0, b_0) \preceq (a, b)$. Ainsi, (a_0, b_0) minore E . □

Proposition 2.0.4. $\mathcal{E} = \{(a, b) \in \mathbb{N}^* \times \mathbb{N}^* : \text{pgcd}(a, b) = 1\}$.

Démonstration. Nous avons déjà montré l'inclusion directe. Soit E l'ensemble des couples $(a, b) \in \mathbb{N}^* \times \mathbb{N}^*$ tels que $\text{pgcd}(a, b) = 1$ et $(a, b) \notin \mathcal{E}$. Supposons $E \neq \emptyset$. Tout d'abord, $(1, 1) \notin E$, puisque la racine de $\mathcal{T}$ a pour étiquette $(1, 1)$. Soit (a, b) le plus petit élément de E pour l'ordre $\preceq$.

- Cas 1, $a > b$. Soient $c = a - b$ et $d = b$. Par la minimalité de a , comme $c < a$, le couple (c, d) est l'étiquette d'un nœud de $\mathcal{T}$. Mais alors, l'étiquette du fils droit de ce nœud est $(c + d, d) = (a, b)$, contradiction.
- Cas 2, $a < b$. Soient $c = a$ et $d = b - a$. Comme $d < b$, le couple (c, d) est l'étiquette d'un nœud de $\mathcal{T}$. Mais alors, l'étiquette du fils gauche de ce nœud est $(c, c + d) = (a, b)$, contradiction.

Ainsi, $E = \emptyset$. $\square$

Proposition 2.0.5. Deux nœuds distincts de $\mathcal{T}$ ont des étiquettes distinctes.

Démonstration. Pour tout $(a, b) \in \mathbb{N}^* \times \mathbb{N}^*$, $a \neq a + b$ et $a + b \neq b$, donc $(1, 1)$ n'est l'étiquette du fils d'aucun nœud de $\mathcal{T}$. Ainsi, $(1, 1)$ est uniquement l'étiquette de la racine de $\mathcal{T}$.

Soit E l'ensemble des éléments de $\mathcal{E}$ qui sont l'étiquette d'au moins deux nœuds de $\mathcal{T}$. Supposons $E \neq \emptyset$. Soit (c, d) le plus petit élément de E pour l'ordre $\preceq$. Soient ν, ν' deux nœuds distincts d'étiquette (c, d) . Soient μ, μ' les pères de ν et ν' . Soient (a, b) et (a', b') les étiquettes de μ et μ' .

- Cas 1, $c > d$. Le nœud ν est alors le fils droit de μ et $(a, b) = (c - d, d)$. De même, ν' est le fils droit de μ' et $(a', b') = (c - d, d)$ donc μ et μ' ont la même étiquette. Or, $(c - d, d) \preceq (c, d)$ et $(c - d, d) \neq (c, d)$, donc $(c - d, d) \notin E$. De là, $\mu = \mu'$. Les nœuds ν et ν' sont les fils droits du même nœud, donc $\nu = \nu'$, contradiction.
- Cas 2, $c < d$. Le nœud ν est alors le fils gauche de μ et de μ' et $(a, b) = (a', b') = (c, d - c)$. En procédant comme dans le cas 1, on obtient $\nu = \nu'$, contradiction.

$\square$

3 Une énumération des nœuds

Pour tout nœud ν de l'arbre $\mathcal{T}$, notons $\varepsilon(\nu)$ l'étiquette de ν .

Énumérons les nœuds de $\mathcal{T}$ par profondeurs croissantes, puis de la gauche vers la droite. En appelant ν_n le nœud numéro n , $\varepsilon(\nu_0) = (1, 1)$, $\varepsilon(\nu_1) = (1, 2)$, $\varepsilon(\nu_2) = (2, 1)$, $\varepsilon(\nu_3) = (1, 3)$, etc. Avec cette numérotation, les fils gauche et droit de ν_n sont ν_{2n+1} et ν_{2n+2} .

Proposition 3.0.6. Pour tout $n \in \mathbb{N}$, ν_{n+1} est la décomposition binaire de $n + 1$.

Démonstration. Montrons ce résultat par récurrence forte sur n . Le résultat est vérifié par 0. Soit $n \in \mathbb{N}^*$. Supposons la propriété vérifiée pour tout $m < n$. Si $n = 2m + 1$ est impair, alors ν_n est le fils gauche de ν_m . Par l'hypothèse de récurrence, $\nu_m 1 = a_0 \dots a_{p-1} 1$ est la décomposition binaire de $m + 1$. Dit autrement,

$$m + 1 = \sum_{k=0}^{p-1} 2^{a_k} + 2^p$$

De là,

$$2m + 1 = \sum_{k=0}^{p-1} 2^{1+a_k} + 2^{p+1} - 1$$

TODO $\square$

Posons $\varepsilon(\nu_n) = (f(n), g(n))$.

On a les valeurs initiales $f(0) = g(0) = 1$ et $f(1) = 1, g(1) = 2$. De plus, pour tout $n \in \mathbb{N}$, les fils gauche et droit de ν_n sont ν_{2n+1} et ν_{2n+2} . On a donc les relations

$$(f(2n + 1), g(2n + 1)) = (f(n), f(n) + g(n))$$

$$(f(2n + 2), g(2n + 2)) = (f(n) + g(n), g(n))$$

Ainsi,

$$\begin{cases} f(2n + 1) &= f(n) \\ f(2n + 2) &= f(n) + g(n) \\ g(2n + 1) &= f(n) + g(n) \\ g(2n + 2) &= g(n) \end{cases}$$

Proposition 3.0.7. *Pour tout $n \in \mathbb{N}$, $g(n) = f(n + 1)$.*

Démonstration. Montrons ce résultat par récurrence forte sur n . Tout d'abord, $g(0) = 1 = f(1)$. Soit $n \in \mathbb{N}^*$. Supposons le résultat vrai pour tous les entiers strictement inférieurs à n .

- Si $n = 2m + 1$ est impair, alors on n'a pas besoin de l'hypothèse de récurrence.

$$g(n) = g(2m + 1) = f(m) + g(m) = f(2m + 2) = f(n + 1)$$

- Si $n = 2m + 2$ est pair, alors

$$g(n) = g(2m + 2) = g(m)$$

Par l'hypothèse de récurrence,

$$f(n + 1) = f(2m + 3) = f(m + 1) = g(m)$$

Ainsi, $g(n) = f(n + 1)$.

□

Nous avons donc pour tout $n \in \mathbb{N}$, $\nu_n = (f(n), f(n+1))$ ainsi que les relations de récurrence

$$\begin{cases} f(2n+1) &= f(n) \\ f(2n+2) &= f(n) + f(n+1) \end{cases}$$

Ces relations devraient nous rappeler quelque chose ...

Proposition 3.0.8. *Pour tout $n \in \mathbb{N}$, $f(n) = h(n)$.*

Démonstration. Les suites $(f(n))_{n \in \mathbb{N}}$ et $(h(n))_{n \in \mathbb{N}}$ ont les mêmes valeurs initiales et vérifient les mêmes relations de récurrence. □

4 Quotients

Définition 4.0.2. Soit ν un nœud de $\mathcal{T}$ d'étiquette (a, b) . La *valeur* de ν est le rationnel strictement positif $\text{val}(\nu) = a/b$.

Notons, pour tout $n \in \mathbb{N}$,

$$r_n = \text{val}(\nu_n) = \frac{f(n)}{f(n+1)}$$

Par l'étude qui précède, $\{r_n : n \in \mathbb{N}\} = \mathbb{Q}$. De plus, tous les termes de la suite $(r_n)_{n \in \mathbb{N}}$ sont distincts. Nous allons montrer que les r_n s'expriment de façon simple comme les termes d'une suite récurrente.

Lemme 4.0.9. *Pour tout $n \in \mathbb{N}$, $f(2n+1) < f(2n+2)$ et $f(2n+2) > f(2n+3)$.*

Démonstration. Soit $n \in \mathbb{N}$. On a

$$f(2n+2) = f(n) + f(n+1) = f(2n+1) + f(n) > f(2n+1)$$

De même,

$$f(2n+2) = f(n) + f(n+1) = f(n) + f(2n+3) > f(2n+3)$$

□

Soit $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}$ définie par

$$\varphi(x) = \frac{1}{[x] + 1 - \{x\}}$$

où $[x] \in \mathbb{Z}$ est la partie entière de x et $\{x\} = x - [x] \in [0, 1[$ est sa partie fractionnaire.

Remarque. Dans la suite, nous nous intéresserons uniquement à la valeur de φ sur les rationnels. Supposons que $x \in \mathbb{Q}_+$. Écrivons $x = a/b$ où $a \in \mathbb{N}$, $b \in \mathbb{N}^*$, et $\text{pgcd}(a, b) = 1$. La division

euclidienne de a par b s'écrit $a = bq + r$ où $q, r \in \mathbb{N}$ et $0 \leq r < b$. On a $\lfloor x \rfloor = q$ et $\{x\} = r/b$. De là,

$$\varphi(x) = \frac{1}{q + 1 - \frac{r}{b}} = \frac{b}{bq + b - r} = \frac{b}{a + b - 2r}$$

Proposition 4.0.10. *Pour tout $n \in \mathbb{N}$, $r_{n+1} = \varphi(r_n)$.*

Démonstration. Montrons le résultat par récurrence forte sur n . Pour $n = 0$, le résultat est immédiat. Soit $n \geq 1$. Supposons la propriété vérifiée par tous les entiers strictement inférieurs à n .

- Cas 1, $n = 2m + 1$ est impair. Ici, pas besoin de l'hypothèse de récurrence. On a

$$r_n = \frac{f(2m+1)}{f(2m+2)} = \frac{f(m)}{f(m) + f(m+1)}$$

Comme $f(2m+1) < f(2m+2)$, $\lfloor r_n \rfloor = 0$ et $\{r_n\} = r_n$. De là,

$$\varphi(r_n) = \frac{1}{1 - r_n} = \frac{f(m) + f(m+1)}{f(m+1)}$$

Par ailleurs,

$$r_{n+1} = \frac{f(2m+2)}{f(2m+3)} = \frac{f(m) + f(m+1)}{f(m+1)}$$

- Cas 2, $n = 2m + 2$ est pair. On a

$$r_n = \frac{f(2m+2)}{f(2m+3)} = \frac{f(m) + f(m+1)}{f(m+1)} = 1 + r_m$$

et

$$r_{n+1} = \frac{f(2m+3)}{f(2m+4)} = \frac{f(m+1)}{f(m+1) + f(m+2)} = \frac{r_{m+1}}{1 + r_{m+1}}$$

Soit $x \in \mathbb{R}_+$. On a $\lfloor 1 + x \rfloor = 1 + \lfloor x \rfloor$ et $\{1 + x\} = \{x\}$. De là,

$$\varphi(1 + x) = \frac{1}{2 + \lfloor x \rfloor - \{x\}} = \frac{\varphi(x)}{1 + \varphi(x)}$$

Par l'hypothèse de récurrence, $r_{m+1} = \varphi(r_m)$, et donc $r_{n+1} = \varphi(1 + r_m) = \varphi(r_n)$.

□

Pour conclure, la suite $(r_n)_{n \in \mathbb{N}}$ est une suite « magique ». Les termes de cette suite prennent comme valeurs, une et une seule fois, tous les rationnels strictement positifs. Il est remarquable que la relation de récurrence soit aussi simple ...

Bibliographie. Horreur, une bibliographie avec une seule référence? Rien de tel que le retour aux sources. Ces sources sont souvent plutôt complexes à décrypter, mais l'article original de Calkin et Wilf est limpide. Je me suis contenté de réorganiser les choses et de m'obliger à un peu plus de rigueur et de détails que ce qu'il est courant de rencontrer dans des articles de recherche, qui laissent souvent au lecteur le soin de compléter les détails. L'article en question est incontournable, on le trouve en quelques secondes sur Internet.

- Neil Calkin and Herbert S. Wilf, Recounting the rationals, The American Mathematical Monthly, Mathematical Association of America, Vol. 107, No 4, pp. 360-363, April 2000.

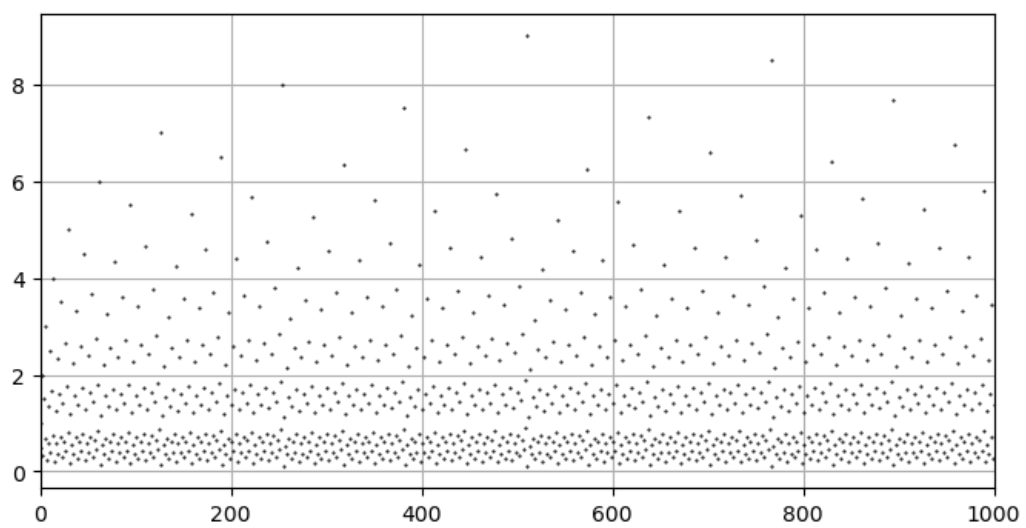


FIGURE 6 – Les mille premiers termes de la suite $(r_n)_{n \in \mathbb{N}}$