A propos d’une curieuse famille de fonctions récursives imbriquées dues à Hofstadter

Pierre Letouzey

9 octobre 2023

Avertissement : la préparation de cet exposé a subi…

La loi d'Hofstadter : il faut toujours plus longtemps que prévu, même en tenant compte de la loi d'Hofstadter.

Code Coq + cet exposé + ancien rapport technique

https://github.com/letouzey/hofstadter_g

(branche: `generalized`)

La fonction G d'Hofstadter (OEIS A5206)

Douglas Hofstadter, “Gödel,Escher,Bach”, chapitre 5 (p.135)

$$G(0) = 0$$
$$G(n) = n - G(G(n-1)) \qquad \text{pour tout entier } n > 0$$

Etude préliminaire de G

$G(n) = n - G(G(n-1))$

- Existence et premier encadrement: $0 \leq G(n) \leq n$
- $G(0) = 0$, $G(1) = 1$ puis $1 \leq G(n) < n$
- G “monte” par pas de $+0$ ou $+1$
- Jamais deux $+0$ de suite
- Jamais trois $+1$ de suite

Etude préliminaire de G

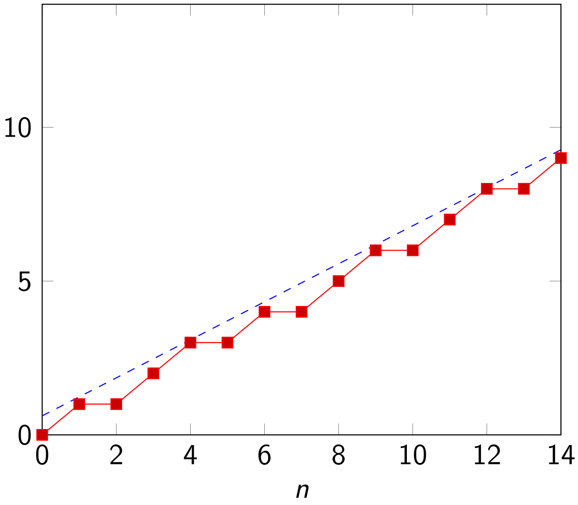
$G(n) = n - G(G(n-1))$

- Existence et premier encadrement: $0 \leq G(n) \leq n$
- $G(0) = 0$, $G(1) = 1$ puis $1 \leq G(n) < n$
- G "monte" par pas de $+0$ ou $+1$
- Jamais deux $+0$ de suite
- Jamais trois $+1$ de suite

Ici en fait : $G(n) = \lfloor (n+1)/\phi \rfloor$ où ϕ est le nombre d'or

Etude préliminaire de G

Graphes de $G(n)$ et $(n+1)/\phi$:



Généralisons : la fonction H (OEIS A5374)

Comme Hofstadter, varions le nombre d'appels imbriqués:

$$H(0) = 0$$
$$H(n) = n - H(H(H(n-1))) \qquad \text{pour tout entier } n > 0$$

Mêmes propriétés de base que G, sauf que:

- Au plus trois $+1$ successifs
- Pas d'équation simple et exacte à base de $\lfloor\ \rfloor$
- Par contre: $H(n) = \lfloor \tau n \rfloor + 0$ ou 1

  avec τ racine réelle de $X^3 + X - 1$ ($\tau = 0.6823$)

Genéralisons encore : une famille f_k de fonctions

Notons $k+1$ le nombre d'appels récursifs:

$$\begin{aligned} f_k(0) &= 0 \\ f_k(n) &= n - f_k^{(k+1)}(n-1)) \qquad \text{pour tout entier } n > 0 \end{aligned}$$

où $f_k^{(p)}$ note p itérations de f_k : $f_k^{(0)} = id$ et $f_k^{(p+1)} = f_k \circ f_k^{(p)}$

On retrouve en particulier $G = f_1$ et $H = f_2$

NB: ce k est choisi pour éviter le cas "0 appel récursif" (sans intérêt et non uniforme avec le reste)

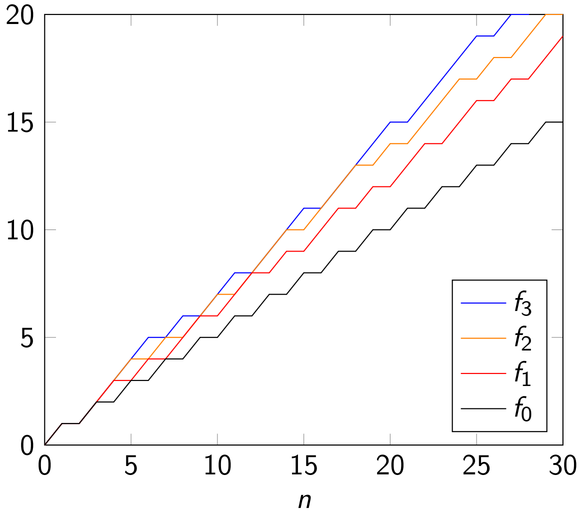
Le cas initial f_0 : un seul appel récursif

$f_0(n) = n - f_0(n-1)$

On alterne $+0$ et $+1$, c'est en fait une fonction moitié :

$f_0(n) = \lfloor (n+1)/2 \rfloor = \lceil n/2 \rceil$

Graphes



Premières propriétés de f_k

$$f_k(n) = n - f_k^{(k+1)}(n-1))$$

- Existence et premier encadrement: $0 \leq f_k(n) \leq n$
- $f_k(0) = 0$, $f_k(1) = 1$ puis $1 \leq f_k(n) < n$
- f_k “monte” par pas de $+0$ ou $+1$
- Jamais deux $+0$ de suite
- Au plus $k + 1$ pas de $+1$ de suite

NB: Pour k>1, $f_k(n)$ n'a pas d'expression simple via $\lfloor\ \rfloor$.

Deux équations intéressantes pour G puis f_k

Surjectivité "explicite"

- $G(n + G(n)) = n$
- $f_k(n + f_k^{(k)}(n)) = n$

Deux équations intéressantes pour G puis f_k

Surjectivité “explicite”

- $G(n + G(n)) = n$
- $f_k(n + f_k^{(k)}(n)) = n$

Equation “renversée”

- $G(n) + G(G(n+1) - 1) = n$
- $f_k(n) + f_k^{(k)}(f_k(n+1) - 1) = n$

Et en Coq ?

Cf `FunG.v` `FunG_prog.v` `GenG.v` :

- Décroissance non structurelle : pas de `Fixpoint` Coq ainsi
- Spécification via un prédicat inductif
- `recf` : une définition remaniée avec un compteur `p`
- Possibilité d'utiliser `Program Fixpoint` (mais lourd)
- Plus rapide : `fopt` fonctionnant par table

Conjecture: croissance des f_k point-à-point

Conjecture: $\forall k, \forall n, f_k(n) \leq f_{k+1}(n)$

Conjecture: croissance des f_k point-à-point

Conjecture: $\forall k, \forall n, f_k(n) \leq f_{k+1}(n)$

Ici, on comparera toujours les fonctions via l'ordre produit.

Donc formulation alternative : (f_k) est une suite croissante.

Conjecture: croissance des f_k point-à-point

Conjecture: $\forall k, \forall n, f_k(n) \leq f_{k+1}(n)$

Ici, on comparera toujours les fonctions via l'ordre produit.

Donc formulation alternative : (f_k) est une suite croissante.

Preuve générale ??

Conjecture: croissance des f_k point-à-point

Quelques éléments préliminaires:

- Facile: $\forall k, f_0 \leq f_k$

Conjecture: croissance des f_k point-à-point

Quelques éléments préliminaires:

- Facile: $\forall k, f_0 \leq f_k$
- Preuves ad-hoc (et dures) : pour $k \leq 9$, $f_k \leq f_{k+1}$.

  Utilise une forme de quasi-additivité (et des calculs!)

Conjecture: croissance des f_k point-à-point

Quelques éléments préliminaires:

- Facile: $\forall k, f_0 \leq f_k$
- Preuves ad-hoc (et dures) : pour $k \leq 9$, $f_k \leq f_{k+1}$.

  Utilise une forme de quasi-additivité (et des calculs!)
- "Petits" n : $\forall k, \forall n \leq (k+4)(k+5)/2 - 3, f_k(n) \leq f_{k+1}(n)$.

  Cf le "bas" des arbres à venir juste après

Conjecture: croissance des f_k point-à-point

Quelques éléments préliminaires:

- Facile: $\forall k, f_0 \leq f_k$
- Preuves ad-hoc (et dures) : pour $k \leq 9$, $f_k \leq f_{k+1}$.

  Utilise une forme de quasi-additivité (et des calculs!)
- "Petits" n : $\forall k, \forall n \leq (k+4)(k+5)/2 - 3, f_k(n) \leq f_{k+1}(n)$.

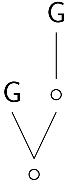
  Cf le "bas" des arbres à venir juste après
- "Grands" n : $\forall k, \exists N, \forall n \geq N, f_k(n) \leq f_{k+1}(n)$

  Lorsque $n \to \infty$ on a l'équivalent $f_k(n) \sim n.\tau_k$

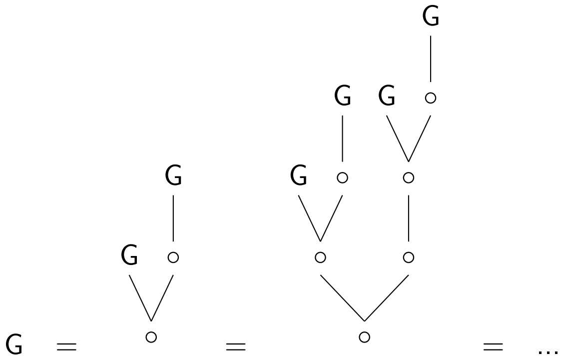
  où τ_k est la racine réelle positive de $X^{k+1} + X - 1$

Arbres rationnels

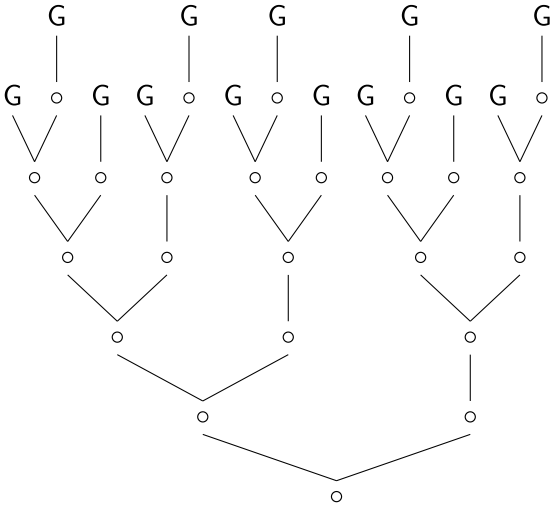
Un arbre infini rationnel

$G =$ 

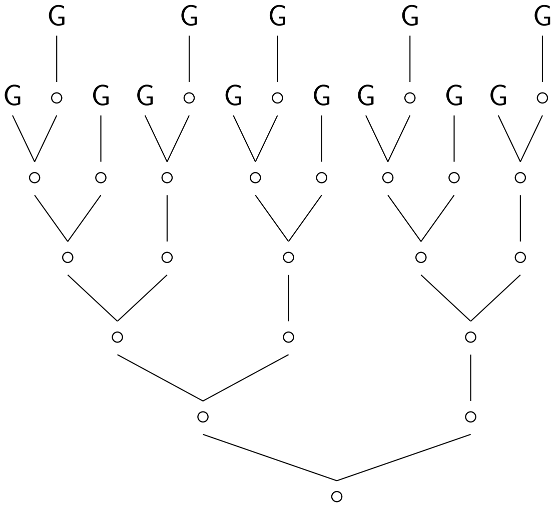
Un arbre infini rationnel



Un arbre infini rationnel



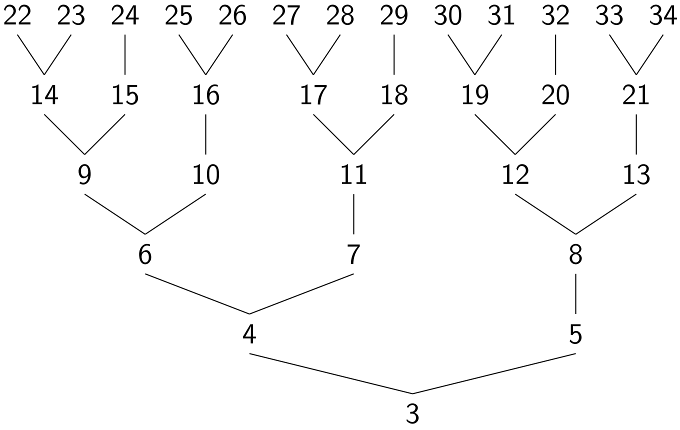
Un arbre infini rationnel



Combien de noeuds par niveau ?

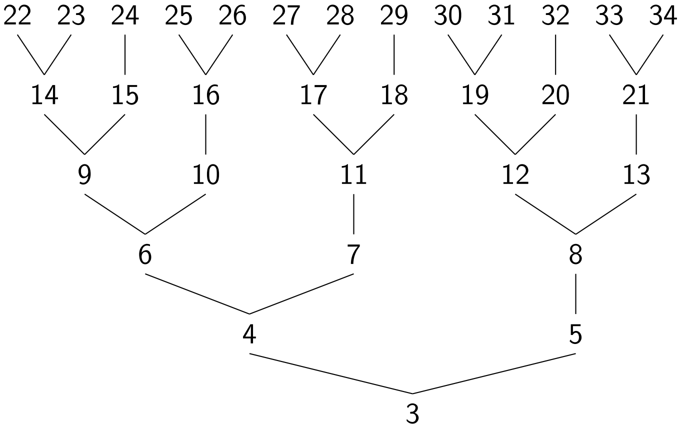
Numérotons !

Parcours en largeur, de gauche à droite



Numérotons !

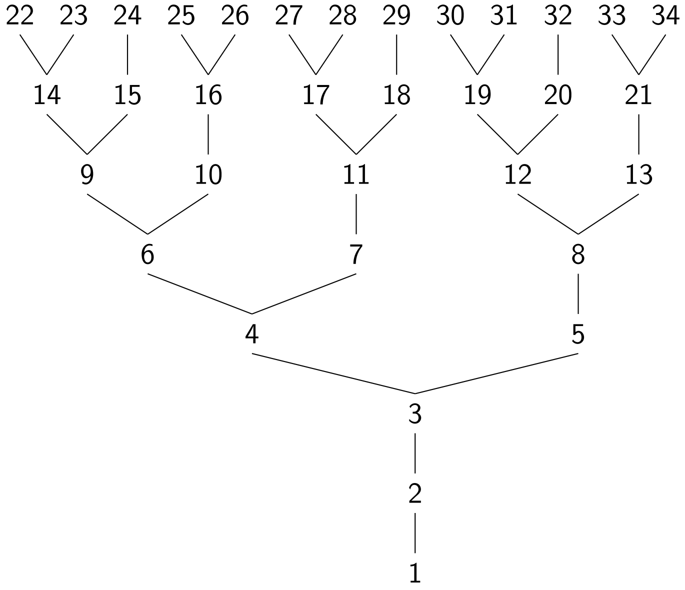
Parcours en largeur, de gauche à droite



Départ à 3 ? Pour expliciter les nombres de Fibonacci...

Et ainsi, le noeud n a $G(n)$ comme parent.

Ajout d'une racine ad-hoc : l'arbre de G



Aparté : arbre de fonction, fonction d'arbre

Soit un arbre:

- infini
- dont les noeuds ont des arités finies non nulles
- numéroté via un parcours en largeur

Que peut-on dire de sa fonction parent ?

Aparté : arbre de fonction, fonction d'arbre

Soit un arbre:

- infini
- dont les noeuds ont des arités finies non nulles
- numéroté via un parcours en largeur

Que peut-on dire de sa fonction parent ?

Recip. que faut-il sur une fonction $\mathbb{N} \to \mathbb{N}$ pour qu'elle soit la fonction parent d'un et un seul tel arbre ?

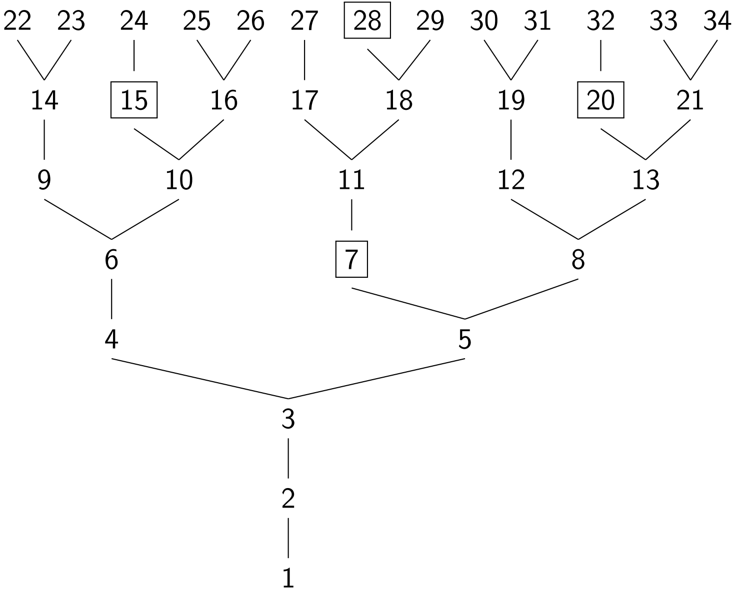
Aparté : arbre de fonction, fonction d'arbre

- f croissante
- f(n)$<$n hormis à la racine
- f surjective
- f ne stationne pas (i.e. tend vers $+\infty$)

Hofstadter: A problem for curious readers is. . .

Suppose you flip diagram G around as if in a mirror, and label the nodes of the new tree so that they increase from left to right. Can you find a recursive *algebraic* definition for this “flip-tree” ?

Arbre miroir $\overline{G}$



Solution ?

- Il y avait une conjecture sur https://oeis.org/A123070
- Mais pas de preuve...
- Hofstadter devait probablement avoir au moins cette formule

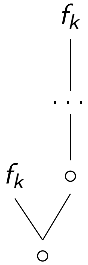
$$\begin{aligned}
\overline{G}(n) &= n + 1 - \overline{G}(\overline{G}(n-1)+1) && (n > 3)\\
\overline{G}(n) &= n && (n = 0, 1)\\
\overline{G}(n) &= n - 1 && (n = 2, 3)
\end{aligned}$$

- Preuve papier pénible, multiples cas (vive Coq! cf fichier `FlipG.v`)

Arbre généralisé

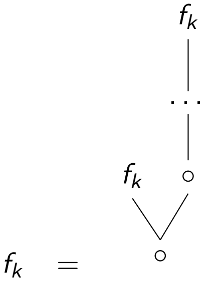
On allonge la branche de droite ($k+1$ segments)

$f_k \quad =$



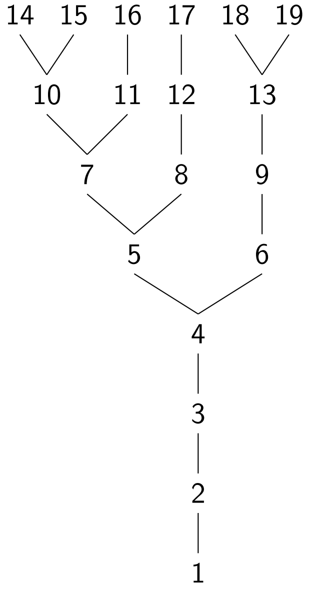
Arbre généralisé

On allonge la branche de droite ($k+1$ segments)

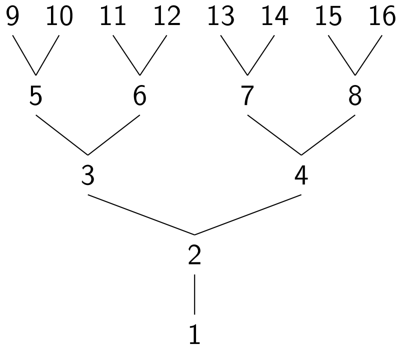


Et toujours une racine ad-hoc (1 puis $k+1$ segments)

Arbre pour f_2 (H de Hofstadter)



Arbre pour f_0



Equation de l'arbre miroir $\overline{f}_k$?

Quasiment comme pour $\overline{G}$:

$$
\begin{aligned}
\overline{f}_k(n) &= n + 1 - \overline{f}_k^{(k)}(\overline{f}_k(n-1) + 1) && (n > k+2) \\
\overline{f}_k(n) &= n && (n = 0, 1) \\
\overline{k}_k(n) &= n - 1 && (2 \leq n \leq k+2)
\end{aligned}
$$

Fibonacci généralisé et numération

Fibonacci

$$
\begin{aligned}
F_0 &= 1 \\
F_1 &= 2 \\
F_{n+2} &= F_n + F_{n+1}
\end{aligned}
$$

Fibonacci

$$F_0 = 1$$
$$F_1 = 2$$
$$F_{n+2} = F_n + F_{n+1}$$

NB: indices décalés pour éviter 0 et un double 1

Théorème de Zeckendorf

Une décomposition $n = \Sigma F_i$ est *canonique* si elle est :

(1) sans doublons
(2) sans termes consécutifs

Décomposition *relachée* : (1) mais pas forcément (2)

Théorème de Zeckendorf

Une décomposition $n = \Sigma F_i$ est *canonique* si elle est :

(1) sans doublons
(2) sans termes consécutifs

Décomposition *relachée* : (1) mais pas forcément (2)

Thm: tout entier naturel a une unique décomposition canonique.

Zeckendorf, variante

Def: le *rang* d'une décomposition est l'indice du plus petit terme.

Algo: canonisation d'une décomposition relachée de n

- le nombre de termes décroît ou stagne
- le rang augmente (par pas de 2) ou stagne

Tableau de Wythoff / Zeckendorf (k=1)

Colonne c: les nombres de rang c par ordre croissant

1	2	3	5	8	13	21	...
4	7	11	18	29	47	76	...
6	10	16	26	42	68	110	...
9	15	24	39	63	102	...	...
12	20	32	52	84	...	...	...
14	23	37	60	97	...	...	...
17	28	45	73	118	...	...	...
...	...	...	...	...	...	...	...

G et Fibonacci

- $G(F_i) = F_{i-1}$ (avec la convention $F_{0-1} = F_0 = 1$)

G et Fibonacci

- $G(F_i) = F_{i-1}$ (avec la convention $F_{0-1} = F_0 = 1$)
- Plus généralement: $G(\Sigma F_i) = \Sigma F_{i-1}$

G et Fibonacci

- $G(F_i) = F_{i-1}$ (avec la convention $F_{0-1} = F_0 = 1$)
- Plus généralement: $G(\Sigma F_i) = \Sigma F_{i-1}$
- Cela marche même pour des décompositions relachées
- Preuve selon le rang de la décomposition (0, pair>0, impair).
- Nombreuses conséquences concernant G et le rang.

Au passage, différences entre $\overline{G}$ et G

Def: n est de rang 1-impair si sa décomposition canonique commence par $F_1 + F_{2p+1} + \ldots$.

Thm: $\overline{G}(n) = 1 + G(n)$ si n est de rang 1-impair, sinon $\overline{G}(n) = G(n)$.

Au passage, différences entre $\overline{G}$ et G

Def: n est de rang 1-impair si sa décomposition canonique commence par $F_1 + F_{2p+1} + ...$

Thm: $\overline{G}(n) = 1 + G(n)$ si n est de rang 1-impair, sinon $\overline{G}(n) = G(n)$.

Preuve: encore pire que pour l'équation de $\overline{G}$, pléthore de cas.

Cor: $\overline{G}$ et G diffèrent pour $7 = F_1 + F_3$, puis tous les 5 ou 8 entiers.

Fibonacci généralisé

Soit k un entier naturel.

$$
\begin{aligned}
A_0^k &= 1 \\
A_1^k &= 2 \\
&\ldots \\
A_k^k &= k+1 \\
A_{n+1}^k &= A_n^k + A_{n-k}^k \qquad\qquad \text{pour } n \geq k
\end{aligned}
$$

Fibonacci généralisé

- A^0 : 1 2 4 8 16 32 64 128 256 512
- A^1 : 1 2 3 5 8 13 21 34 55 89
- A^2 : 1 2 3 4 6 9 13 19 28 41
- A^3 : 1 2 3 4 5 7 10 14 19 26

NB: A^2 est nommé Narayana's Cows, cf. OEIS A930

Zeckendorf généralisé

Soit k fixé.

k-décomposition $n = \Sigma A_i^k$ est *canonique* : indices distants $\geq (k+1)$

k-décomposition *relachée* : indices distants d'au moins k

Zeckendorf généralisé

Soit k fixé.

k-décomposition $n = \Sigma A_i^k$ est *canonique* : indices distants $\geq (k+1)$

k-décomposition *relachée* : indices distants d'au moins k

Thm: tout entier naturel a une unique k-décomposition canonique.

Algo: on peut "renormaliser" une k-décomposition relachée.

Un peu d'arithmétique avec ces décompositions

La décomposition de $n+1$ et $n-1$ peut s'obtenir raisonnablement bien à partir de celle de n.

Par contre pas d'addition, multiplication, etc.

f_k et Fibonacci généralisé

- $f_k(A_i^k) = A_{i-1}^k$ (avec la convention $A_{0-1}^k = A_0^k = 1$)

f_k et Fibonacci généralisé

- $f_k(A_i^k) = A_{i-1}^k$ (avec la convention $A_{0-1}^k = A_0^k = 1$)
- Plus généralement: $f_k(\Sigma A_i^k) = \Sigma A_{i-1}^k$

f_k et Fibonacci généralisé

- $f_k(A_i^k) = A_{i-1}^k$ (avec la convention $A_{0-1}^k = A_0^k = 1$)
- Plus généralement: $f_k(\Sigma A_i^k) = \Sigma A_{i-1}^k$
- Cela marche pour des décompositions canoniques ou relachées
- Important : f_k “stagne” en n lorsque le rang de n est 0 (i.e. lorsque n a 1 dans sa décomposition)

Quasi-additivité de f_k ?

Un exemple d'utilisation des décompositions:

```
Lemma additivity_bounded k p : k<>0 ->
 forall n, exists m,
   m < add_bound k p /\
   f k (p+n) - f k n = f k (p+m) - f k m.

Lemma decide_additivity k p a b : k<>0 ->
 calc_additivity k p (add_bound k p) = (a,b) ->
 forall n, a + f k n <= f k (p+n) <= b + f k n.
```

Quasi-additivité de f_k ?

Un exemple d'utilisation des décompositions:

```
Lemma additivity_bounded k p : k<>0 ->
 forall n, exists m,
   m < add_bound k p /\
   f k (p+n) - f k n = f k (p+m) - f k m.

Lemma decide_additivity k p a b : k<>0 ->
 calc_additivity k p (add_bound k p) = (a,b) ->
 forall n, a + f k n <= f k (p+n) <= b + f k n.
```

Ceci a permis de prouver $f_1 \leq f_2$ jusqu'à $f_9 \leq f_{10}$ (en Coq: seulement jusqu'à $f_5 \leq f_6$).

Lien avec des mots morphiques

Une substitution de lettres

Soit k un entier naturel. On utilise $\mathcal{A} = [0..k]$ comme alphabet.

$$\begin{aligned} \mathcal{A} &\to \mathcal{A}^* \\ \sigma_k(n) &= (n+1) && \text{pour } n < k \\ \sigma_k(k) &= k.0 \end{aligned}$$

Ceci engendre un mot infini m_k à partir de la lettre k (on parle de mot *morphique*)

Par exemple $m_2 = 2012202012012201220 2...$

Equation récursive

m_k est la limite de $\sigma_k^n(k)$ quand $n \to \infty$

Mais aussi la limite de préfixes finis $M_{k,n}$ définis ainsi:

- $M_{k,n} = k.0...(n-1)$ pour $n \leq k$
- $M_{k,n+1} = M_{k,n}.M_{k,n-k}$ pour $k \leq n$

Equation récursive

m_k est la limite de $\sigma_k^n(k)$ quand $n \to \infty$

Mais aussi la limite de préfixes finis $M_{k,n}$ définis ainsi:

- $M_{k,n} = k.0...(n-1)$ pour $n \leq k$
- $M_{k,n+1} = M_{k,n}.M_{k,n-k}$ pour $k \leq n$

Remarque : $|M_{k,n}| = A_n^k$

Lien avec f_k

La n-ième lettre $(m_k)_n$ du mot infini m_k est le rang de la k-decomposition de n (ou k si ce rang est plus de k).

En particulier cette lettre est 0 si $f_k(n) = f_k(n+1)$

En cumulant : le nombre de 0 dans m_k entre 0 et n est $n - f_k(n)$.

Plus généralement, compter les lettres au dessus de p donne $f_k^{(p)}$.
En particulier le nombre de k est $f_k^{(k)}$.

Fréquences ?

Quelle limite pour $f_k(n)/n$ lorsque $n \to \infty$?

- Si elle existe, facile à déterminer, racine positive de $X^{k+1} + X - 1$.
- Preuve d'existence non triviale

Cf. K. Saari, *On the Frequency of Letters in Morphic Sequences*.

En Coq, il fallait déjà parler de racines, et d'équivalent infini de suites linéaires comme A^k.

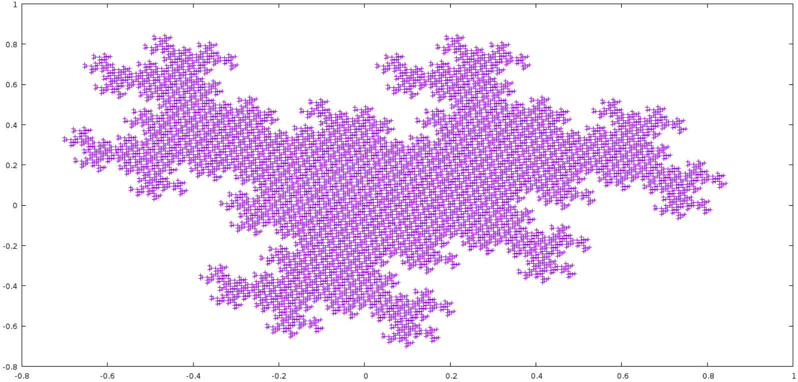
De fil en aiguille, preuve de la formule de Leibniz du determinant et determinant des matrices de Vandermonde. . .

Assure la croissance des f_k pour n suffisemment grand.

Cas k=2 (i.e. H)

Surprise il y a quelques années

Affichage des points $(\delta(i), \delta(H(i)))$ avec i=0..10000 et $\delta(n) = H(n) - n.\tau_2$



Fractale de Rauzy et variante

Apparemment, la factale précédente est nommée Jacobi-Perron, proche de la fractale de Rauzy.

G. Rauzy, *Nombres algébriques et substitutions*, 1982

- Dans son cas, suites de Tribonacci additionnant les trois derniers termes
- Ici on additionne dernier et avant-avant-dernier termes

L'étude est très similaire.

Application ici

On obtient finalement:

- $|H(n) - n.\tau_2| < 0.996 < 1$
- Et donc $H(n) = \lfloor n.\tau_2 \rfloor + 0$ ou 1
- Et quasi-additivité de H :
  $\forall nm, -2 \leq H(n+m) - H(n) - H(m) \leq 2$

Nombres de Pisot

Dixit Wikipédia: En mathématiques, un nombre de Pisot-Vijayaraghavan est un entier algébrique réel strictement supérieur à 1, dont tous les éléments conjugués ont un module strictement inférieur à 1.

Ici la limite τ_2 de $H(n)/n$ est la racine positive de $X^3 + X - 1$ mais aussi l'inverse de la racine positive de $X^3 - X^2 - 1$ qui est le nombre de Pisot P_3.

Cas k=3, Pisot sans jolie fractale...

Résultat principal pour k=3

En suivant le même cheminement (pas encore formalisé en Coq)

- $|f_3(n) - n.\tau_3| < 1.998$
- Et donc $-1 \leq f_3(n) - \lfloor n.\tau_3 \rfloor \leq 2$
- Et quasi-additivité de f_3 :
  $\forall nm, -5 \leq H(n+m) - H(n) - H(m) \leq 5$

Cas k>3, $f_k(n) - n.\tau_k$ diverge

Conclusions & Perspectives

- On trouve encore des conjectures “abordables” sur OEIS
- Et aussi parfois des petites choses fausses. . .

Conclusions & Perspectives

- On trouve encore des conjectures “abordables” sur OEIS
- Et aussi parfois des petites choses fausses...
- Des preuves étonnemment délicates pour de “simples” entiers.
- Merci Coq !
- Peut-on éviter ces “détours” via $\mathbb{R}$ et $\mathbb{C}$?
- Quid de la conjecture ?
- Des questions restantes concernant l'irréductibilité des polynômes rencontrés