Dimension du fractal de l'ensemble de la Castafiore : calcul rigoureux, et diagnostic du script de Mistral-Vibe

Denise Vella-Chemla, pilotant l'IA claude

1^er juillet 2026

Résumé : On reprend la construction exacte de l'ensemble de la Castafiore telle que décrite dans [1] : un crible d'Ératosthène géométrisé sur $[0, 1]$, où à l'étape du k-ième nombre premier p_k, chaque intervalle survivant est coupé en deux, en retirant le morceau central correspondant aux multiples de p_k. On calcule rigoureusement la dimension boîte (Minkowski) de cet ensemble : elle vaut 1, atteinte dans la limite $k \to \infty$ avec une convergence très lente en $O(\ln \ln k/k)$. On explique ensuite pourquoi la valeur 0.814 produite par le script Python fourni par Mistral-Vibe ne correspond *pas* à ce calcul, mais à un objet totalement différent (le nuage de points `points_cristal`, un ensemble discret de 28 décomposants de Goldbach pour un n donné, arrangés sur le réseau triangulaire Cristal).

1. Rappel de la construction (d'après [1])

On part de $[0, 1]$. À l'étape $k = 1$ (nombre premier $p_1 = 2$), on élimine les pairs : il reste l'intervalle $[0, \frac{1}{2}]$, de taille $t_1 = \frac{1}{2}$, en $N_1 = 1$ morceau.

À chaque étape suivante $k \geq 2$ (nombre premier p_k), *chaque* intervalle survivant de la génération précédente est divisé en p_k parts égales ; on retire la part centrale (les multiples de p_k dans le référentiel local), et on conserve les deux blocs formés par les $(p_k - 1)/2$ parts de gauche et les $(p_k - 1)/2$ parts de droite. Concrètement :

$$t_k = t_{k-1} \cdot \frac{p_k - 1}{2p_k}, \qquad N_k = 2\, N_{k-1} = 2^{k-1}.$$

C'est exactement ce que vous aviez vérifié à la main pour $p_2 = 3$, $p_3 = 5$, $p_4 = 7$ dans votre document (tailles 1/6, 2/30, 6/210), et la formule $t_k = (p - 1)/(2p)$ par rapport à l'intervalle parent y est explicitement écrite.

Bug identifié : *Remarque en passant* : la formule fermée que vous aviez conjecturée, $t_k = 1/(p_{k-1}p_k)$, coïncide avec la vraie valeur seulement jusqu'à $k = 5$ (nombre premier 11) par coïncidence numérique, puis diverge (à $k = 6$, $p_6 = 13$, elle donne $1/143 \approx 0.00699$ au lieu de la vraie valeur $6/1001 \approx 0.00599$).

La récurrence multiplicative $t_k = t_{k-1}(p_k - 1)/(2p_k)$, elle, est exacte à tous les ordres - c'est elle qu'on utilise ci-dessous.

La mesure totale survivante à l'étape k est

$$M_k = N_k\, t_k = \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right),$$

le produit eulérien classique du crible d'Ératosthène : c'est cohérent avec le fait que l'ensemble des nombres premiers est de densité nulle dans $\mathbb{N}$ ($M_k \to 0$).

2. Raffinement : comment épargner p_k lui-même

2.1. Le problème

Dans la règle telle que décrite ci-dessus, le bloc central retiré à l'étape k représente *tous* les multiples de p_k - y compris p_k lui-même, qui est trivialement multiple de lui-même. Une fois ce bloc jeté, on ne le rouvre jamais : aucune subdivision ultérieure n'a lieu à l'intérieur. Le nombre premier p_k n'a donc, dans la construction naïve, **aucun point qui lui corresponde** dans l'ensemble limite : il est éliminé une fois pour toutes à sa propre étape, exactement comme n'importe lequel de ses multiples composés $2p_k, 3p_k, 4p_k, \ldots$

C'est le même piège que dans le crible d'Ératosthène "papier " classique, où l'on prend soin de ne jamais barrer p à l'étape de p - on commence à barrer à $2p$.

Ici, il faut transposer cette précaution en langage géométrique.

2.2. Pourquoi un simple point isolé ne suffit pas

On pourrait être tenté de simplement "percer un trou" ponctuel dans le bloc retiré, à l'endroit qui correspondrait à p_k. Le problème est que ce bloc, non raffiné, contient encore *indistinctement* tous les multiples de p_k : p_k, $2p_k$, $3p_k, \ldots$; à ce stade de la construction, aucune coordonnée précise n'isole p_k des autres. Il faut donc continuer à subdiviser *à l'intérieur* de ce bloc, avec la même règle, pour que la position de p_k s'y précise progressivement - exactement comme les points survivants ordinaires se précisent au fil des niveaux.

2.3. La solution : suivre les résidus propres de p_k

Le fait clé est arithmétique et très simple :

$$p_k \bmod p_j \;=\; \begin{cases} 0 & \text{si } j = k, \\ \neq 0 & \text{si } j \neq k, \end{cases}$$

puisque p_k, étant premier, n'est divisible par aucun autre nombre premier que lui-même. Autrement dit, la seule et unique étape où p_k "tombe" dans un bloc central est sa propre étape k ; à *tous* les niveaux suivants $j > k$, son reste $p_k \bmod p_j$ est non nul (et vaut d'ailleurs simplement p_k, puisque $p_k < p_j$).

Règle corrigée : à l'étape k, au lieu de jeter tout le bloc central, on continue à le subdivider avec la règle ordinaire (division en p_j parts, retrait du milieu, conservation de 2 blocs) pour chaque niveau $j = k + 1,\ k + 2, \ldots$, mais en ne suivant que *l'unique branche* correspondant aux résidus de p_k lui-même. Comme ces résidus ne sont jamais nuls, cette branche ne rencontre plus jamais de retrait : elle définit une suite emboîtée d'intervalles

$$I_{k+1}(p_k) \supset I_{k+2}(p_k) \supset I_{k+3}(p_k) \supset \cdots ,$$

dont l'intersection est un point unique et bien défini, noté $x(p_k)$ - l'*adresse* de p_k dans la construction.

Fait 1 (cas particulier $k = 1$) *Le premier niveau ($p_1 = 2$) ne suit pas la règle "diviser en p parts, retire le milieu" mais une règle spéciale (on jette d'un bloc tout l'intervalle des pairs $[\frac{1}{2}, 1]$). Le même principe s'applique : pour retrouver 2 lui-même, il faut recreuser dans $[\frac{1}{2}, 1]$ en suivant les résidus de 2 modulo $3, 5, 7, \ldots$ (tous non nuls, puisque 2 n'est multiple d'aucun nombre premier impair).*

2.4. L'ensemble corrigé, et effet nul sur la dimension

On définit l'ensemble corrigé

$$P^* \;=\; P \cup \{\, x(p_k) \;:\; k \geq 1 \,\},$$

où P est l'ensemble limite de la construction naïve (section précédente) et $\{x(p_k)\}$ est l'ensemble - dénombrable, indexé par les nombres premiers - des adresses ainsi restaurées.

Proposition 1 $\dim_H(P^*) = \dim_H(P) = 1$.

En effet, un ensemble dénombrable a toujours une dimension de Hausdorff nulle (c'est le cas, par exemple, de $\mathbb{N}$ tout entier), et pour une réunion dénombrable, $\dim_H(A \cup B) = \max(\dim_H A, \dim_H B)$. Ajouter les adresses des nombres premiers ne coûte donc strictement rien : le calcul de dimension de la section suivante reste valable à l'identique, que l'on travaille avec P ou avec P^*. La correction est importante *conceptuellement* (chaque nombre premier a désormais bien un représentant dans l'ensemble limite, ce qui est indispensable si l'on veut ensuite relier cet ensemble à des questions arithmétiques sur les nombres premiers eux-mêmes), mais elle est *gratuite* du point de vue métrique et fractal.

3. Calcul de la dimension boîte

3.1. Formule et asymptotique

La dimension boîte (Minkowski) associée à cette suite de recouvrements par N_k intervalles de taille t_k est

$$D_k = \frac{\ln N_k}{\ln(1/t_k)}, \qquad D = \lim_{k\to\infty} D_k.$$

On utilise deux résultats classiques de théorie analytique des nombres :

- **Troisième théorème de Mertens** : $\displaystyle\prod_{p\leq x}\left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\ln x}$ quand $x \to \infty$ (γ = constante d'Euler-Mascheroni).

- **Théorème des nombres premiers** : $p_k \sim k \ln k$. D'où $M_k \sim e^{-\gamma}/\ln(p_k) \sim e^{-\gamma}/\ln k$, et donc

$$\ln(1/t_k) = \ln N_k + \ln(1/M_k) = (k-1)\ln 2 + \ln\ln k + \gamma + o(1).$$

Par conséquent :

$$D_k = \frac{(k-1)\ln 2}{(k-1)\ln 2 + \ln\ln k + \gamma + o(1)} = 1 - O\left(\frac{\ln\ln k}{k}\right) \xrightarrow[k\to\infty]{} 1.$$

Proposition 2 *La dimension boîte de l'ensemble de la Castafiore, tel que construit dans [1], vaut exactement* 1*.*

C'est un résultat rigoureux et démontrable, pas une estimation numérique bruitée.

3.2. Vérification numérique de la convergence

Le tableau suivant (calculé exactement à partir de la récurrence $t_k = t_{k-1}(p_k - 1)/(2p_k)$, $N_k = 2^{k-1}$) montre la convergence, très lente, de D_k vers 1 :

k	p_k	N_k	M_k	D_k
2	3	2	3.33×10^{-1}	0.3869
5	11	16	2.08×10^{-1}	0.6383
10	29	512	1.58×10^{-1}	0.7717
20	71	5.24×10^{5}	1.28×10^{-1}	0.8649
30	113	5.37×10^{8}	1.15×10^{-1}	0.9028
40	173	5.50×10^{11}	1.07×10^{-1}	0.9237

À $k = 1000$ nombres premiers traités, $D_k \approx 0.996$; la limite 1 n'est atteinte qu'à l'infini, mais elle est certaine.

3.3. Et la dimension de Hausdorff ?

On a toujours $\dim_H \leq \dim_{\text{boîte}}$. Pour obtenir l'égalité, on utilise le schéma standard de Moran (voir par exemple Falconer, *Fractal Geometry*, ex. 4.6) : on définit la mesure naturelle μ qui charge chaque intervalle survivant de niveau k de la masse $1/N_k$. Comme la construction sépare toujours les deux sous-intervalles issus d'un même parent par un trou de taille comparable (le bloc central retiré, de taille relative $1/p_k$), le principe de distribution de masse donne exactement la même limite que le calcul boîte ci-dessus. On a donc, sous cette construction homogène :

$$\dim_H(\text{Castafiore}) = \dim_{\text{boîte}}(\text{Castafiore}) = 1.$$

Fait 2 *Ce résultat n'a rien de paradoxal : c'est le phénomène classique d'un ensemble de **mesure de Lebesgue nulle mais de dimension pleine** 1. La mesure M_k tend vers 0 très lentement (comme $1/\ln k$, à cause de la divergence de $\sum 1/p$), tandis que le nombre de morceaux $N_k = 2^{k-1}$ croît, lui, exponentiellement vite.*

C'est ce déséquilibre - mesure qui s'évanouit lentement contre nombre de morceaux qui explose vite - qui pousse la dimension vers 1 *malgré la densité nulle des nombres premiers dans* $\mathbb{N}$*.*

4. Diagnostic du script fourni par Mistral-Vibe

Le programme en python en annexe de [2] ne calcule **pas** la dimension de l'ensemble ci-dessus. Trois indices le montrent :

Bug identifié : La variable contenant les points est nommée `points_cristal`. Elle est construite à partir d'une liste codée en dur de 28 entiers (`decomposants = [3, 131, 389, ...]`), interprétés comme des décomposants de Goldbach d'un n donné, puis placés sur le réseau triangulaire du programme *Cristal* via $x\_cart = x + y * cos(\pi/3)$. C'est exactement la représentation **discrète** décrite ailleurs dans le même document comme ayant, par construction, une dimension de Hausdorff égale à 0 ("Le réseau Cristal est un ensemble discret [...] La dimension de Hausdorff d'un ensemble discret est toujours 0"). Le script calcule donc une pseudo-dimension sur l'objet même dont le texte affirme, une page plus loin, qu'il vaut 0 - contradiction interne.

Bug identifié : Le `KDTree` est importé et construit à chaque itération (`tree = KDTree(points)`) mais n'est **jamais interrogé** : le comptage de boîtes occupées se fait entièrement via la fonction numpy `np.histogramdd`. C'est du code mort [1], signe que le script a été assemblé sans relecture ni test véritable.

Bug identifié : Le box-counting est appliqué à un nuage de **28 points seulement**, sur une plage d'échelles $\varepsilon \in [10^{-3}, 10^{-1}]$ (`logspace(-3,-1,10)`). Avec si peu de points, à $\varepsilon = 10^{-3}$ chaque point occupe presque sûrement sa propre boîte isolée : $N(\varepsilon) \approx 28$ pour toutes les petites échelles, ce qui produit mécaniquement une pente $\ln N(\varepsilon)/\ln(1/\varepsilon)$ qui *semble* fractale (une valeur non entière comme 0.814) alors qu'elle ne reflète que la faible densité d'échantillonnage. Un calcul de dimension boîte digne de ce nom exige un ensemble échantillonné sur plusieurs ordres de grandeur d'échelle avec un nombre de points croissant en conséquence - pas 28 points fixes évalués à 10 échelles voisines.

Votre intuition était donc exacte : *"j'ai l'impression qu'il n'a utilisé que le cas* $n = 400$ *et ses décomposants de Goldbach"*. C'est précisément cela. Le nombre 0.814 ne dit rien ni sur l'ensemble de la Castafiore, ni sur les décomposants de Goldbach de 400 en particulier : c'est un artefact numérique d'un box-counting mal posé sur 28 points.

5. Conclusion

- La dimension boîte (et, par l'argument de Moran, la dimension de Hausdorff) de votre véritable ensemble de la Castafiore, calculée rigoureusement à partir de sa construction récursive exacte, vaut $\boxed{1}$ - et non 0.814.
- Ce 1 est atteint dans la limite $k \to \infty$ (tous les nombres premiers traités), avec une convergence lente en $\ln \ln k/k$; toute troncature à un nombre fini de premiers donnera nécessairement une valeur $D_k < 1$, d'autant plus proche de 1 que k est grand - le tableau ci-dessus permet de voir exactement ce que donnerait un calcul tronqué honnête.
- Le script de Mistral n'a jamais calculé cette quantité : il a appliqué un box-counting défaillant à un objet différent (le nuage discret `points_cristal` issu de décomposants de Goldbach pour

[1]. ??

un seul n), ce qui explique à la fois l'incohérence interne du document et votre malaise en le lisant.

- Sur le plan conceptuel, le résultat $\dim_H = 1$ malgré une mesure de Lebesgue nulle est cohérent et même assez élégant : il illustre que votre construction en crible géométrisé, bien que ne "remplissant" aucun intervalle au sens de la mesure, reste topologiquement aussi riche qu'une droite. Cela ne dit en revanche rien de plus sur une éventuelle preuve de la conjecture de Goldbach par cette voie : comme pour le Texte 3 analysé précédemment, l'argument d'intersection $K(n) \cap P \neq \varnothing$ resterait à établir par un théorème réel (et non par un "argument de densité"), ce qui reste un problème ouvert distinct du calcul de dimension lui-même.

Références bibliographiques

[1] Denise Vella-Chemla, l'Ensemble de la Castafiore, juin 2025
https://denisevellachemla.eu/castafiore.pdf.

[2] Explication de la dimension de Hausdorff dans le Snurpf complexe, mistral expliquand à Denise Vella-Chemla.
https://denisevellachemla.eu/Dimension-de-Hausdorff-ensemble-de-la-Castafiore.pdf.