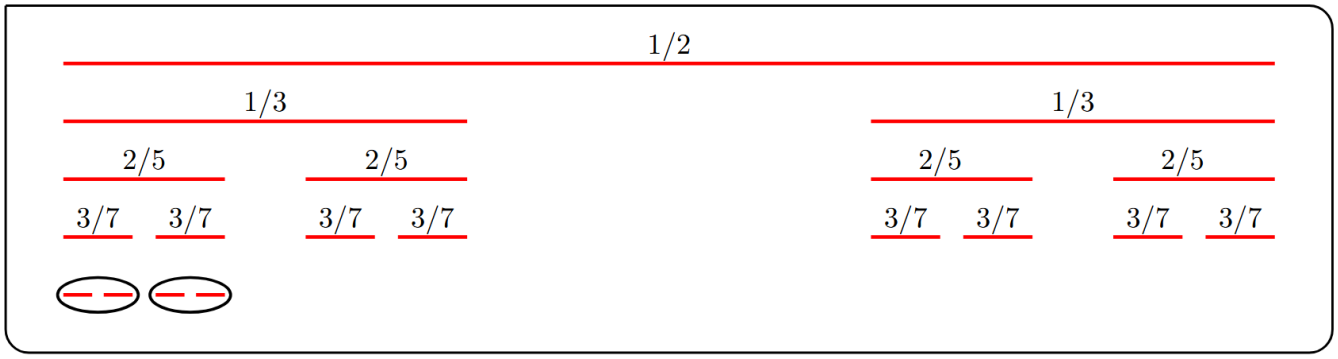
L'ensemble de la castafiore, Denise Vella-Chemla, juin 2025.

Au cours de ces recherches que je mène en free-lance autour de la conjecture de Goldbach, j'avais utilisé un système de numération par les restes modulaires (de Gauss), que j'avais appelé le Snurpf pour Système de numération par les restes dans les parties finies de $\mathbb{N}$. J'ai récemment retrouvé dans un des livres de Knuth une référence à de tels systèmes de numération.

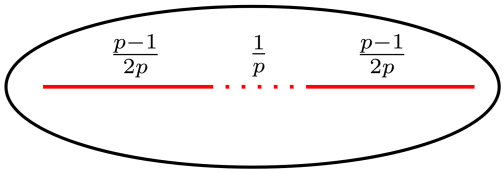
Un nombre premier supérieur à $\sqrt{n}$ est un décomposant de Goldbach de n s'il ne partage aucun reste modulaire avec n. Leila Schneps a démontré cette caractérisation.

Si on représente les nombres par leurs restes modulaires modulo l'ensemble infini des nombres premiers $\mathcal{P}$, les nombres premiers n'ont qu'un seul reste nul, c'est leur reste modulo eux-mêmes. Et cela amène naturellement au souhait de voir l'ensemble des nombres premiers comme un ensemble de Cantor, dit aussi "poussière de Cantor", selon le dessin suivant :



Il s'agit d'éliminer, selon chaque module premier p_k, les nombres qui ont pour reste modulaire 0 selon ce nombre premier, i.e. qui sont divisibles par lui. Comme on le voit, on a, comme pour l'ensemble triadique de Cantor, systématiquement positionné l'intervalle à éliminer au milieu des intervalles d'un niveau donné.

Si on zoome sur l'un des petits intervalles d'un niveau donné, on conserve à gauche et à droite de l'un des intervalles de son niveau associé, deux petits intervalles de même taille, de taille $\frac{p-1}{2p}$ chacun.



On consultera avec profit l'article de référence de Cantor qui s'intitule *De la puissance des ensembles parfaits de points* et qu'on a transcrit en LaTeX ici
https://denisevellachemla.eu/transc-Cantor-ensembles-parfaits.pdf.

Procédons comme Cantor nous l'explique.

Au premier niveau, on élimine tous les nombres pairs, car à part 2, aucun nombre pair ne peut être un nombre premier. On avait l'intervalle initial $[0, 1]$, on obtient l'intervalle $\left[0, \frac{1}{2}\right]$.

Au niveau suivant, on s'occupe du nombre premier 3, l'intervalle $\left[0, \frac{1}{2}\right]$ est coupé en 3. On élimine l'intervalle central qui représente les multiples de 3. On conserve deux intervalles de taille $\frac{1}{6}$, commençant respectivement en 0 et $\frac{2}{6}$; ce sont les intervalles $\left[0, \frac{1}{6}\right]$ et $\left[\frac{2}{6}, \frac{1}{2}\right]$.

Au niveau suivant, on s'occupe du nombre premier 5, on va éliminer l'intervalle central qui représente les multiples de 5. On se retrouve avec 4 intervalles, dont les débuts respectifs sont les nombres 0, $\frac{3}{30}, \frac{10}{30}$ et $\frac{13}{30}$. Les 4 intervalles conservés sont tous de la même taille, cette taille est $\frac{2}{30}$. Voici les 4 intervalles :

$$\left[0, \frac{2}{30}\right] \qquad \left[\frac{3}{30}, \frac{1}{6}\right] \qquad \left[\frac{2}{6}, \frac{6}{15}\right] \qquad \left[\frac{13}{30}, \frac{1}{2}\right]$$

Au niveau suivant, on s'occupe du nombre premier 7, on va éliminer l'intervalle central qui représente les multiples de 7. On se retrouve avec 8 intervalles, qui sont tous de la même taille, cette taille est $\frac{6}{210}$. Voici les 8 intervalles :

$$\left[0, \frac{6}{210}\right], \qquad \left[\frac{8}{210}, \frac{2}{30}\right], \qquad \left[\frac{3}{30}, \frac{27}{210}\right], \qquad \left[\frac{1}{6} - \frac{6}{210}, \frac{1}{6}\right],$$

$$\left[\frac{2}{6}, \frac{2}{6} + \frac{6}{210}\right], \qquad \left[\frac{12}{30} - \frac{6}{210}, \frac{12}{30}\right], \qquad \left[\frac{13}{30}, \frac{13}{30} + \frac{6}{210}\right], \qquad \left[\frac{1}{2} - \frac{6}{210}, \frac{1}{2}\right].$$

Quel est le schéma récursif derrière ces nombres ? À chaque niveau k, les intervalles sont tous de la même taille t_k, on les notera $[a_{k,i}, b_{k,i}]$ avec i allant de 1 à une puissance de 2 (on a vu qu'on avait 2 intervalles pour le nombre premier 3, 4 pour le nombre premier 5, 8 pour le nombre premier 7, 16 pour le nombre premier 11, etc.) et ils ont comme extrémités gauche pour moitié les extrémités gauche des intervalles du niveau précédent $a_{k-1,i}$, et pour moitié $a_{k-1,i} + t_{k-1} - t_k$. L'extrémité droite de chaque intervalle s'obtient en ajoutant à son extrémité gauche la taille du niveau courant.

Au niveau du nombre premier p_k, il y a 2^{k-1} intervalles, tous de taille

$$t_k = \frac{\prod\limits_{\substack{p_j \text{ premier} \\ 2 \leq p_j \leq p_{k-2}}} p_j}{\prod\limits_{\substack{p_j \text{ premier} \\ 2 \leq p_j \leq p_k}} p_j}.$$

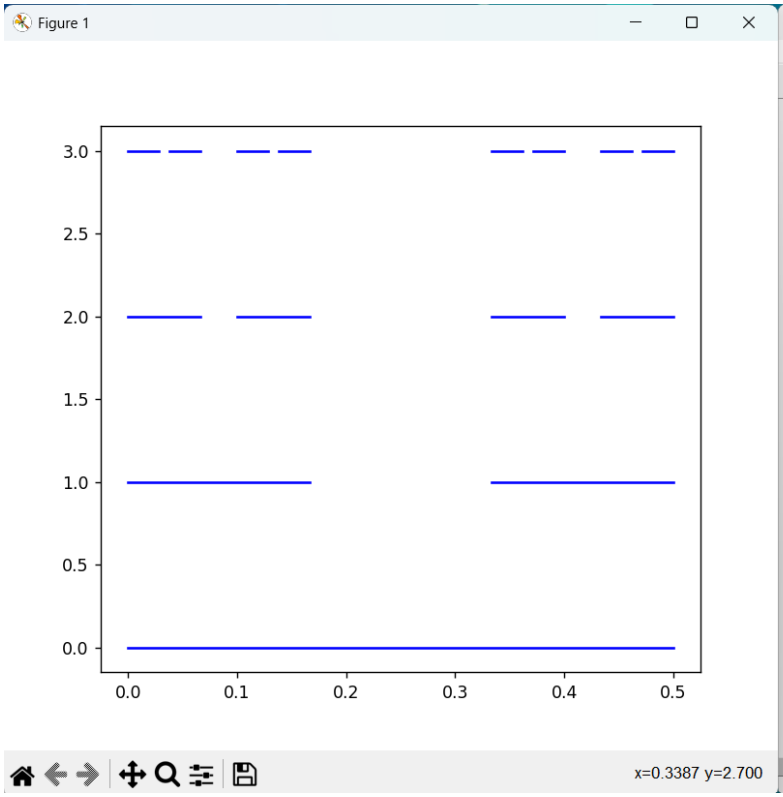
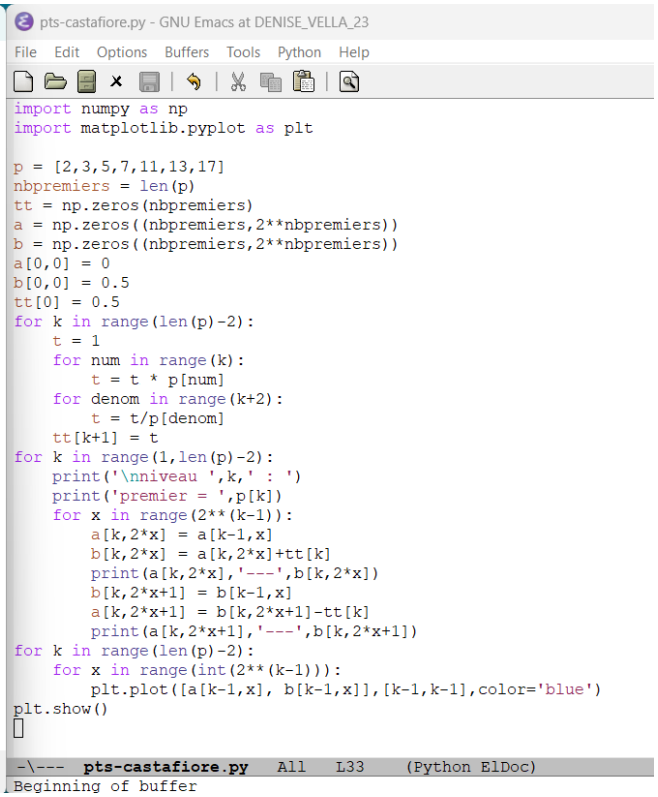
Les extrémités gauches des intervalles sont, quant à celles d'indices impairs, identiques aux extrémités gauches de l'intégralité des extrémités gauches du niveau précédent :

$$a_{2k-1,i} = a_{k,i-1}.$$

Et quant à celles d'indices pairs, on les calcule ainsi :

$$a_{2k,i} = a_{2k-1,i} + t_{i-1} - t_i.$$

Quant aux extrémités droites, on a :

$$b_{k,i} = a_{k,i} + t_i.$$

Ce qui est à souligner par rapport à un ensemble de Cantor, c'est que c'est un ensemble qui a la puissance du continu, c'est une infinité de la même taille que $\mathbb{R}$.