Formule exacte de Legendre pour $|R_N|$ et obstruction combinatoire
rédigé par l'ia claude pour Denise Vella-Chemla

Résumé

On reprend la réduction combinatoire de Goldbach au non-recouvrement de $V_N \setminus \bigcup_r (A_r \cup B_r)$ et on met en œuvre le point 1 du programme proposé [1] : une formule *exacte* par inclusion-exclusion pour $|R_N|$, dans l'esprit du crible de Legendre. On montre que la formule est correcte et vérifiable numériquement, mais que son coût de calcul explose en $2^{\pi(\sqrt{N})}$ termes, ce qui la rend inexploitable dès que N dépasse quelques centaines - c'est très exactement l'obstruction classique qui a historiquement poussé Brun puis Selberg à remplacer le crible de Legendre par des cribles tronqués. Ce document ne prouve donc pas Goldbach ; il rend précis, calculable et vérifié le verrou identifié à la section 10 du document source [2].

1. Rappel des notations

Pour N pair, $N > 2$, et r premier, $r \leq \sqrt{N}$, on pose

$$c_r = \begin{cases} 1 & \text{si } r \mid N, \\ 2 & \text{si } r \nmid N, \end{cases}$$

le nombre de classes de résidus interdites modulo r (voir section 2 du document source). On note $P = P(\sqrt{N}) = \{r \text{ premier} : r \leq \sqrt{N}\}$, $M = \prod_{r \in P} r$, et

$$R_N = \left\{ p \in \left[3, \tfrac{N}{2}\right] \ : \ \operatorname{pgcd}(p, M) = 1, \ \operatorname{pgcd}(N - p, M) = 1 \right\}.$$

Le Lemme 1 du document source donne : $R_N \neq \varnothing \Rightarrow N$ est somme de deux premiers.

2. Formule exacte par inclusion-exclusion

Pour $r \in P$, notons $B_r \subset [3, N/2]$ l'ensemble des p tels que $p \equiv 0$ ou $p \equiv N \pmod r$; c'est une union de c_r classes de résidus modulo r. On a $R_N = [3, N/2] \setminus \bigcup_{r \in P} B_r$.

Pour un diviseur sans facteur carré $d \mid M$, $d = r_1 \cdots r_k$, notons T_d l'ensemble des résidus modulo d obtenus par théorème des restes chinois en choisissant, pour chaque r_i, l'une des c_{r_i} classes interdites modulo r_i ; ainsi $|T_d| = \prod_{r \mid d} c_r$. Pour $t \in T_d$, le nombre exact d'entiers de $[a, b]$ congrus à t modulo d est

$$\operatorname{cnt}(a, b, d, t) = \left\lfloor \frac{b - t}{d} \right\rfloor - \left\lfloor \frac{a - 1 - t}{d} \right\rfloor.$$

Posons $S_d = \sum_{t \in T_d} \operatorname{cnt}\Big(3, \lfloor N/2 \rfloor, d, t\Big)$, avec $S_1 = \lfloor N/2 \rfloor - 2$.

1. Je ne sais pas de quel programme il est question.
2. Je ne sais pas de quel document source il s'agit, ce que c'est que d'être désorganisée...

Proposition 1 (Formule de Legendre pour $|R_N|$)**.**

$$|R_N| \;=\; \sum_{\substack{d \mid M \\ d \text{ sans facteur carré}}} \mu(d)\, S_d.$$

Démonstration. C'est l'inclusion-exclusion usuelle appliquée aux événements $(p \in B_r)_{r \in P}$: $\left|[3, N/2] \setminus \bigcup_r B_r\right| = \sum_{S \subseteq P} (-1)^{|S|} \left|\bigcap_{r \in S} B_r\right|$, et $\bigcap_{r \in S} B_r$ est exactement, par CRT, l'union des $\prod_{r \in S} c_r$ classes de résidus modulo $d = \prod_{r \in S} r$ décrite ci-dessus, d'où $\left|\bigcap_{r \in S} B_r \cap [3, N/2]\right| = S_d$, et $\mu(d) = (-1)^{|S|}$ puisque d est sans facteur carré.

□

Cette proposition est *exacte*, pas asymptotique : c'est bien une "borne exacte par inclusion-exclusion" au sens du point 1 du programme. Elle règle en principe le problème (i) énoncé à la section 5 du document source (non-recouvrement local) de façon totalement explicite et calculable.

3. Pourquoi cela ne suffit pas : explosion combinatoire

Le nombre de diviseurs sans facteur carré de M est $2^{|P|} = 2^{\pi(\sqrt{N})}$. La formule de la Proposition 1 demande donc, en toute généralité, le calcul de $2^{\pi(\sqrt{N})}$ termes S_d.

Remarque : Par le théorème des nombres premiers, $\pi(\sqrt{N}) \sim \dfrac{2\sqrt{N}}{\ln N}$. Le nombre de termes croît donc en $2^{2\sqrt{N}/\ln N}$, une fonction sur-polynomiale en N. Dès $N \approx 4000$ ($\sqrt{N} \approx 63$, $\pi(\sqrt{N}) = 18$), on a déjà $2^{18} \approx 2{,}6 \times 10^5$ termes ; dès $N = 10^6$ ($\pi(1000) = 168$), on a 2^{168} termes, un nombre totalement hors de portée. Le calcul reste correct - c'est une identité, pas une approximation - mais il devient inexploitable bien avant que N ne soit grand.

C'est très exactement, formulé de façon quantitative et vérifiable, le verrou décrit informellement section 10 du document source : "non éliminé par les petits premiers" n'est pas la même chose qu'"exhibable explicitement", et ici on voit précisément *pourquoi* l'approche la plus directe (Legendre) ne peut pas être menée jusqu'au bout : son coût explose exponentiellement en $\pi(\sqrt{N})$, bien avant tout obstacle propre à la théorie des nombres analytique. C'est historiquement le constat qui a conduit Brun (1920) puis Selberg à remplacer la somme complète sur tous les diviseurs sans facteur carré de M par une somme tronquée aux diviseurs $d \le z$ pour un paramètre $z \ll M$ bien choisi, au prix d'un terme d'erreur à contrôler - exactement la piste évoquée à la section 12 du document source ("terme principal – erreur").

4. Vérification numérique

Le script joint `goldbach_legendre_exact.py` calcule $|R_N|$ de deux façons indépendantes pour de petites valeurs de N (là où $2^{\pi(\sqrt{N})}$ reste raisonnable) :

1. par un crible direct (dénombrement brut, en $O(N)$) ;
2. par la formule exacte de la Proposition 1.

Les deux méthodes coïncident sur toute la plage testée, ce qui constitue une vérification de la Proposition 1. Le script affiche également, pour chaque N testé, le nombre de termes $2^{\pi(\sqrt{N})}$ requis par la formule exacte, afin de rendre visible l'explosion combinatoire annoncée ci-dessus.

5. Conclusion et suite

Le point 1 du programme ("borne exacte par inclusion-exclusion pour $|R_N|$") est donc réalisé : la Proposition 1 est une identité exacte, prouvée et vérifiée numériquement. Mais, comme prévu à la section 12 du document source, une identité exacte n'est utile que si elle peut être évaluée ou minorée effectivement ; ici son évaluation directe est hors de portée dès que N cesse d'être petit.

La suite naturelle - qui rejoint les points 3 et 4 du programme[3] - est donc de tronquer la somme de la Proposition 1 à $\omega(d) \leq K$ ou $d \leq z$, d'estimer le terme principal $\sum_{d \leq z} \mu(d) \prod_{r \mid d} c_r \cdot \frac{N/2}{d}$ (qui se factorise en un produit eulérien fini, apparenté à une série singulière de type Hardy-Littlewood), et de majorer rigoureusement le reste. C'est précisément le programme d'un crible de Selberg pour ce problème à deux contraintes de résidu par premier - un travail substantiel, non entrepris ici, mais dont la présente note délimite l'entrée exacte.

Résultat de l'exécution du programme, car un programme a été fourni.

```
C:\Users\Denise_Vella\Desktop>python goldbach_legendre_exact.py
Verification de la formule exacte de Legendre pour |R_N|
================================================================
     N |  |R_N| direct |  |R_N| Legendre | accord |  2^pi(sqrt N)
----------------------------------------------------------------
    10 |             1 |               1 |    oui |             4
    20 |             1 |               1 |    oui |             4
    30 |             3 |               3 |    oui |             8
    50 |             2 |               2 |    oui |            16
   100 |             5 |               5 |    oui |            16
   200 |             6 |               6 |    oui |            64
   400 |            11 |              11 |    oui |           256
   700 |            22 |              22 |    oui |           512
  1000 |            24 |              24 |    oui |          2048
  2000 |            34 |              34 |    oui |         16384

Illustration de l'explosion combinatoire 2^pi(sqrt(N)) (formule non evaluee) :
  N =         1,000 -> pi(sqrt(N)) =   11 -> 2^pi(sqrt(N)) ~ 10^3.3 termes
  N =        10,000 -> pi(sqrt(N)) =   25 -> 2^pi(sqrt(N)) ~ 10^7.5 termes
  N =       100,000 -> pi(sqrt(N)) =   65 -> 2^pi(sqrt(N)) ~ 10^19.6 termes
  N =     1,000,000 -> pi(sqrt(N)) =  168 -> 2^pi(sqrt(N)) ~ 10^50.6 termes
  N = 1,000,000,000 -> pi(sqrt(N)) = 3401 -> 2^pi(sqrt(N)) ~ 10^1023.8 termes
```

3. qu'on ne connaît pas.