Petits décomposants de Goldbach : une heuristique de comptage, son erreur de sommation, et sa correction par le produit de Mertens

d'après la note de Denise Vella-Chemla (18 avril 2023)

Résumé

On formalise une note consacrée au phénomène empirique suivant : la quasi-totalité des nombres pairs n possèdent un décomposant de Goldbach *petit*, c'est-à-dire inférieur à $\sqrt{n}$ - phénomène vérifié exhaustivement jusqu'à 10^8, avec seulement 73 exceptions sous 10^5 et aucune entre 10^5 et 10^8. La note propose deux heuristiques de comptage pour estimer la probabilité qu'un nombre premier $p \leq \sqrt{n}$ donné soit un décomposant de Goldbach. La première, fondée sur une *sommation* des probabilités d'évitement modulo chaque petit premier, contient une erreur de principe : elle confond la probabilité d'une conjonction d'événements indépendants (qui s'obtient par produit) avec celle d'une union (liée à une somme via inclusion-exclusion), ce qui produit des valeurs numériques dépourvues de sens probabiliste (largement supérieures à 1). La seconde heuristique, antérieure (« été 2019 »), évite cet écueil en utilisant le produit de Mertens : c'est la forme correcte, et elle rejoint l'heuristique de Hardy-Littlewood déjà rencontrée dans d'autres notes de la même auteure.

1 Le phénomène empirique

Définition 1 *Pour n pair, $n \geq 6$, on appelle petit décomposant de Goldbach de n tout nombre premier p avec $3 \leq p \leq \sqrt{n}$ tel que $n - p$ soit également premier.*

Constat 1 (Constat numérique) *Jusqu'à 10^5, environ 92,7 % des nombres pairs possèdent un petit décomposant de Goldbach. Le nombre pair*

$$81\,099\,776 = 139 + 81\,099\,637,$$

avec

$$139 < \sqrt{81\,099\,776} \approx 9005{,}54,$$

est cité comme exemple de petit décomposant proche de la borne $\sqrt{n}$. Aucun contre-exemple (nombre pair sans petit décomposant) n'a été trouvé par recherche exhaustive entre 10^5 et 10^8, contre 73 exceptions sous 10^5 (la liste est donnée dans la note, allant de 6 à 63 274).

Remarque 1 *Ce phénomène est cohérent avec, et apparenté à, des résultats connus de la littérature (notamment les travaux de Granville et Pomerance sur les petits décomposants de Goldbach, et plus généralement l'heuristique de Hardy-Littlewood, qui prédit non seulement l'existence mais l'abondance de décomposants pour tout n pair suffisamment grand). C'est un fait empirique solide et un point de départ légitime pour une heuristique probabiliste.*

2 Première heuristique : la formule par sommation, et son erreur

Constat 2 (Caractérisation modulaire d'un petit décomposant) *Un nombre premier p, $3 \le p \le \sqrt{n}$, est un décomposant de Goldbach de n si et seulement si*

$$\forall q \text{ premier},\ 3 \le q \le \sqrt{n} : \quad p \not\equiv n \pmod{q}.$$

La note propose d'estimer la probabilité de cet événement conjoint en attribuant à chaque congruence individuelle « $p \not\equiv n \pmod{q}$ » la probabilité $1 - \dfrac{1}{q}$, puis en *additionnant* ces probabilités sur tous les premiers $q \le \sqrt{n}$:

$$S := \sum_{\substack{3 \le q_k \le \sqrt{n} \\ q_k \text{ premier}}} \left(1 - \frac{1}{q_k}\right) = \Big(\pi(\sqrt{n}) - 1\Big) - \sum_{\substack{3 \le q_k \le \sqrt{n} \\ q_k \text{ premier}}} \frac{1}{q_k},$$

en utilisant l'estimation de Mertens $\sum_{q \le x} \dfrac{1}{q} = \ln\ln x + M + o(1)$ (la note retient le terme principal $\ln\ln\sqrt{n}$), ce qui conduit à

$$S \approx \Big(\pi(\sqrt{n}) - 1\Big) - \ln\ln\sqrt{n}.$$

Constat 3 (L'erreur de principe) *Pour des événements indépendants $A_1, \dots, A_k$, la probabilité de leur conjonction $A_1 \cap A_2 \cap \cdots \cap A_k$ est le produit $\prod_i P(A_i)$, non leur somme. La somme des probabilités intervient dans le calcul de la probabilité d'une union $A_1 \cup \cdots \cup A_k$ (par la formule du crible / inclusion-exclusion), pas d'une conjonction. La quantité S telle que définie additionne les probabilités d'évitement individuelles ; elle estime (très approximativement, et de façon erronée même pour cela) une notion sans rapport avec la probabilité recherchée, qui est celle de la conjonction de tous les évitements.*

Contre-exemple 1 (Confirmation numérique) *Le tableau suivant compare S (somme) au produit correct $B := \prod_{q \le \sqrt{n}} \left(1 - \dfrac{1}{q}\right)$ (défini à la Section 3) :*

n	$\pi(\sqrt{n})$ *(premiers ≥ 3)*	S *(somme)*	B *(produit)*
50	3	2,324	0,457
1 000	10	8,934	0,306
10 000	24	22,697	0,241
100 000	64	62,478	0,193

La quantité S ***dépasse*** *1* ***dès*** *$\pi(\sqrt{n}) \ge 2$ et croît sans borne avec n (chaque terme ajouté est positif et proche de 1), alors qu'une probabilité authentique doit demeurer dans $[0, 1]$. Le produit B, lui, décroît doucement vers 0, comme attendu d'une probabilité de conjonction d'événements de plus en plus nombreux. Le fait que la note elle-même observe « S est supérieur à 1 dès $n = 50$ » aurait dû, en toute rigueur, signaler l'incohérence : une probabilité ne peut excéder 1.*

Remarque 2 *L'erreur n'invalide pas l'intuition de départ (Constat 2, qui est correct), mais uniquement la façon de combiner les probabilités individuelles. Le bon outil de combinaison, pour une conjonction d'événements supposés indépendants, est le produit - c'est précisément ce que fait la seconde heuristique, antérieure, de la même note.*

3 Seconde heuristique : le produit de Mertens (forme correcte)

Théorème 1 (Mertens)

$$B(\sqrt{n}) := \prod_{\substack{p \text{ premier} \\ p \le \sqrt{n}}} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\ln p_{\max}} \left(1 + \mathrm{O}\left(\frac{1}{\ln p_{\max}}\right)\right),$$

où $p_{\max}$ est le plus grand nombre premier $\le \sqrt{n}$, et γ la constante d'Euler-Mascheroni.

Proposition 1 (Forme corrigée de l'heuristique) *En notant $A := \dfrac{n/2}{\ln(n/2)}$ une minoration de $\pi(n/2)$, l'heuristique antérieure (« été 2019 ») propose*

$$S' := A \cdot B(\sqrt{n}) = \frac{n/2}{\ln(n/2)} \cdot \frac{e^{-\gamma}}{\ln p_{\max}} \left(1 + \mathrm{O}\left(\frac{1}{\ln n}\right)\right)$$

comme minoration heuristique du nombre de décomposants de Goldbach de n (petits ou non). La note rapporte que $S' > 1$ pour $n = 6$ et $n = 8$, $S' < 1$ de façon quasi continue jusqu'à $n = 2422$, puis $S' > 1$ de façon apparemment définitive au-delà.

Remarque 3 (Identité avec les heuristiques précédentes) *Cette formule $S' = A \cdot B$ est, à la définition de A près (ici $\pi(n/2)$ minoré simplement par $\dfrac{n/2}{\ln(n/2)}$, plutôt que par la forme effective de Rosser-Schoenfeld), **structurellement identique** à la minoration corrigée obtenue dans la note Annexe à la proposition denitac.pdf (23 juillet 2019) déjà formalisée précédemment, et c'est bien la même heuristique de Hardy-Littlewood qui ressort, par une troisième voie de calcul indépendante, dans cette note d'avril 2023.*

4 Portée et limites de l'heuristique, même corrigée

Remarque 4 ***(ce que l'heuristique corrigée explique, et ce qu'elle ne peut pas prouver)*** *La forme corrigée $S' = A \cdot B$ rend compte, de façon qualitativement satisfaisante, du phénomène empirique de la Section 1 : le produit de Mertens $B(\sqrt{n})$ décroît seulement comme $1/\ln\ln n$ environ (puisque $p_{\max} \sim \sqrt{n}$, donc $\ln p_{\max} \sim \dfrac{1}{2}\ln n$), bien plus lentement que A ne croît, ce qui explique que S' devienne rapidement, et reste, largement supérieur à 1 : c'est cohérent avec la raréfaction empirique des exceptions au-delà d'un seuil modeste (2422 selon le calcul de la note).*

Cette heuristique souffre cependant des mêmes limites de principe que toutes les approches apparentées déjà rencontrées : elle repose sur une hypothèse d'indépendance, entre les événements de congruence aux différents petits modules premiers, qui n'est nulle part démontrée. C'est exactement le contenu du problème de parité de Selberg : aucun crible ne peut, par lui seul, certifier une borne inférieure non triviale sur le nombre d'entiers premiers dans un ensemble criblé - l'estimation heuristique de Hardy-Littlewood, retrouvée ici sous une troisième forme, demeure une conjecture corroborée numériquement, non un théorème.

Remarque 5 (Sur l'objet « petit décomposant« spécifiquement) *Il faut noter, enfin, que ni S ni S' n'estiment directement la probabilité qu'existe un décomposant petit ($\leq \sqrt{n}$) - elles estiment (de façon heuristique, pour S', et erronée, pour S) le nombre total de décomposants de Goldbach de n, sans distinguer les petits des grands. Le phénomène spécifique observé en Section 1 (rareté décroissante des nombres pairs sans petit décomposant) est un fait plus fin que ne le capture cette heuristique globale ; il relève de résultats plus spécifiques sur la distribution des plus petits décomposants de Goldbach, dans l'esprit des travaux de Granville et Pomerance, et reste, à ce stade, un fait observé plutôt qu'expliqué par le calcul proposé.*

Conclusion

Le phénomène empirique de départ - la rareté décroissante des nombres pairs sans petit décomposant de Goldbach - est solide et bien documenté dans la note. La première tentative de quantification heuristique, par sommation des probabilités d'évitement modulo les petits premiers, contient une erreur de principe (somme au lieu de produit pour une conjonction d'événements), confirmée numériquement par l'explosion de S bien au-delà de 1. La seconde heuristique, antérieure et fondée sur le produit de Mertens, est la forme correcte ; elle s'avère, une fois reconnue, identique en substance à l'heuristique de Hardy-Littlewood déjà rencontrée dans d'autres notes de la même auteure, avec les mêmes vertus (cohérence numérique qualitative) et la même limite de principe (hypothèse d'indépendance non démontrée, problème de parité de Selberg).

```
import math

def est_premier(k):
    if k < 2: return False
    if k < 4: return True
    if k % 2 == 0: return False
    i = 3
    while i*i <= k:
        if k % i == 0: return False
        i += 2
    return True

def premiers_jusqu_a(x):
    return [p for p in range(2, int(x)+1) if est_premier(p)]

print(f"{'n':>6} {'pi(sqrt n)':>10} {'S (somme)':>12} {'B (produit)':>12}")
for n in [50, 98, 1000, 10000, 100000]:
    racine = math.isqrt(n)
    premiers = [p for p in premiers_jusqu_a(racine) if p >= 3]
    S = sum(1 - 1/q for q in premiers)
    B = 1.0
    for q in premiers:
        B *= (1 - 1/q)
        print(f"{n:>6} {len(premiers):>10} {S:>12.4f} {B:>12.6f}")
```

Notes bibliographiques

- D. Vella-Chemla, *Petits décomposants de Goldbach*, 18 avril 2023.
  https://denisevellachemla.eu/petits-premiers.pdf.
- D. Vella-Chemla, *Annexe à la proposition denitac.pdf*, 23 juillet 2019.
  https://denisevellachemla.eu/annexdenita.pdf.
- A. Granville, C. Pomerance, travaux sur les petits décomposants de la conjecture de Goldbach.