Héritage des décomposants de Goldbach : théorème de synchronisation modulaire, vérification numérique et modèle probabiliste

Denise Vella-Chemla, pilotant l'IA claude

29 juin 2026

1. Introduction

Cette note formalise une observation faite par l'auteure il y a de nombreuses années, en lisant l'annexe d'un livre de Davis et Hersh [1] listant les décompositions de Goldbach des entiers pairs autour de 20902 et 20962 : les décomposants apparaissant dans les deux séries, espacées de 60, sont remarquablement les mêmes, dans le même ordre. Cette observation est consignée dans deux notes personnelles de l'auteure datant d'octobre 2007 [2, 3], qui y proposait déjà l'idée d'un ordre lexicographique sur les représentations par restes chinois pour l'expliquer, ainsi qu'un argument de descente infinie de Fermat. La présente note reprend cette idée, vieille de près de vingt ans, sous le nom *d'héritage des décomposants de Goldbach*, l'énonce comme théorème, la teste numériquement sur un cas particulier (l'écart de 6), puis explore un modèle probabiliste pour quantifier le phénomène - modèle dans lequel une observation de l'auteure (un lemme de divisibilité élémentaire) permet d'isoler une part **rigoureusement démontrée**, non heuristique, de la corrélation observée. Toute la note utilise la convention suivante : n **désigne directement un entier pair** (et non $2n$).

2. Le critère de Vella-Chemla et le crible pivot

Définition 1 (Décomposant de Goldbach)**.** *Pour n pair, on appelle décomposant de Goldbach de n tout nombre premier p, $2 \leq p \leq n-2$, tel que $n-p$ est également premier.*

Définition 2 (Critère de Vella-Chemla)**.** *Soit n pair et q un nombre premier impair, $q \leq \sqrt{n}$. Si p est un décomposant de Goldbach de n avec $p > \sqrt{n}$, alors*

$$p \not\equiv 0 \pmod{q} \qquad \textit{et} \qquad p \not\equiv n \pmod{q}.$$

Ce critère a été conjecturé par Denise Vella-Chemla et démontré par Leila Schneps.

Définition 3 (Crible pivot)**.** *Pour n pair et $B \geq 2$, on appelle crible pivot de n en base B la famille, indexée par les premiers impairs $q \leq B$, des paires de classes interdites $\{\dot{0}_q, \dot{n}_q\}$ (réduites à un singleton si $q \mid n$). On note $\mathcal{A}(n, B)$ l'ensemble des entiers pivot-autorisés, c'est-à-dire évitant ces classes pour tout $q \leq B$.*

Par construction, tout décomposant de Goldbach $p > \sqrt{n}$ de n appartient à $\mathcal{A}(n, \sqrt{n})$.

3. Théorème d'héritage

Théorème 1 (Invariance du crible pivot par congruence partagée). *Soient n, n' pairs et $B \leq \min(\sqrt{n}, \sqrt{n'})$. Si*
$$n \equiv n' \pmod q \quad \textit{pour tout premier impair } q \leq B,$$
alors $\mathcal{A}(n, B) = \mathcal{A}(n', B)$.

Démonstration. La paire de classes interdites en base q ne dépend que de $n \bmod q$ (via $\dot{n}_q$) et de 0. Si $n \equiv n' \pmod q$, ces deux paires coïncident pour ce q ; ceci étant vrai pour tout $q \leq B$ par hypothèse, les ensembles pivot-autorisés coïncident. □

Corollaire 1 (Condition suffisante par primorielle). *Si $n' \equiv n \pmod{\#B}$, où $\#B = \prod_{q \leq B} q$ est la primorielle de B, alors l'hypothèse du théorème 1 est automatiquement vérifiée.*

Remarque 1. *C'est exactement le phénomène observé par l'auteure en 2007 entre* 20902 *et* 20962 *(écart $60 = 2^2 \times 3 \times 5$) : l'héritage n'est garanti que pour $q = 3$ et $q = 5$ (pas pour $q = 7, 11, 13, \ldots$), ce qui correspond très précisément à l'observation de l'auteure que les représentations par restes des deux séries "se correspondent terme à terme seulement selon les trois premières coordonnées" (c'est-à-dire $q = 2, 3, 5$).*

4. Vérification numérique : la conjecture de l'écart 6

Conjecture 1 (Vella-Chemla). *Pour tout n pair, $n > 12$, n partage au moins un décomposant de Goldbach avec $n - 6$.*

Remarque 2. *Puisque $6 = 2 \times 3$ ne contient que le facteur premier impair 3, l'hypothèse du théorème 1 n'est garantie que pour $q = 3$: cette conjecture n'est donc* ***pas*** *une consequence directe du théorème d'héritage général, mais un phénomène empirique plus fort, à vérifier à part.*

Méthode. Crible d'Ératosthène jusqu'à $N = 10^7$, puis pour chaque n pair de 14 à N, recherche d'un nombre premier p tel que p, $n - p$ et $(n - 6) - p$ soient simultanément premiers (donnant un décomposant commun de n et $n - 6$).

Résultat. Aucun contre-exemple trouvé pour les 4 999 994 valeurs paires testées entre 14 et 10 000 000 (calcul effectué en 51 secondes). La conjecture résiste largement à cette borne.

5. Identification avec la constante de Hardy-Littlewood

On définit la *densité de Vella-Chemla* tronquée à B :

$$\delta(n, B) = \prod_{q \mid n,\, q \leq B} \frac{q-1}{q} \times \prod_{q \nmid n,\, q \leq B} \frac{q-2}{q}$$

(produits sur les premiers impairs), qui mesure la proportion réelle d'entiers pivot-autorisés. On démontre, par une identité algébrique exacte ($\frac{q-2}{q-1} = (1 - \frac{1}{(q-1)^2}) \cdot \frac{q-1}{q}$), que cette densité coïncide, terme à terme, avec la constante singulière $\mathfrak{S}(n)$ de Hardy-Littlewood (1923), à un facteur de renormalisation universel $\prod_{q\le B} \frac{q}{q-1}$ près (facteur de Mertens, indépendant de n). Le détail complet de cette identification fait l'objet d'une note séparée [4]. Cette identification confirme que la densité de l'auteure, obtenue par un raisonnement élémentaire (identité de Poincaré sur le partage de coordonnées modulo les petits premiers), est exactement la heuristique standard de la littérature - sans rien y perdre, et sans que cela invalide l'intérêt propre du théorème d'héritage (§ 3), qui porte sur une relation *entre* deux entiers pairs, absente de la formulation classique de $\mathfrak{S}(n)$.

6. Modèle probabiliste d'héritage

6.1. Position du problème

Soient n, n' partageant le même crible pivot $\mathcal{A} = \mathcal{A}(n, B) = \mathcal{A}(n', B)$ (théorème 1). On note $\mathrm{Dec}(n)$ l'ensemble des décomposants effectifs de n supérieurs à $\sqrt{n}$. On souhaite estimer

$$\Pr\Big[\mathrm{Dec}(n) \cap \mathrm{Dec}(n') = \varnothing\Big].$$

6.2. Trois ingrédients, de nature très différente

Pour chaque $p \in \mathcal{A}$, on a

$$p \in \mathrm{Dec}(n) \cap \mathrm{Dec}(n') \iff p \text{ premier } \wedge \ (n-p) \text{ premier } \wedge \ (n'-p) \text{ premier.}$$

Ingrédient 1 (heuristique de Cramér). On modélise la primalité d'un entier m comme un évènement de Bernoulli indépendant, de paramètre $1/\ln m$. Sous ce modèle, comme "p premier" est un évènement **partagé** entre les deux probabilités ci-dessus (il n'est compté qu'une fois), on obtient, pour $p \sim n/2$,

$$\Pr[p \in \mathrm{Dec}(n) \cap \mathrm{Dec}(n')] \approx \frac{1}{(\ln n)^3},$$

alors qu'une indépendance totale entre les deux évènements "$p \in \mathrm{Dec}(n)$" et "$p \in \mathrm{Dec}(n')$" donnerait $\frac{1}{(\ln n)^4}$ - soit un facteur correctif $\ln n$ dû au partage du candidat p.

Ingrédient 2 (hypothèse de Poisson). On suppose que le nombre de coïncidences suit une loi de Poisson de paramètre égal à l'espérance, ce qui permet d'écrire $\Pr[\# = 0] \approx e^{-\mathbb{E}[\#]}$.

Ingrédient 3 (approximation $p \sim n/2$ uniforme). On utilise une densité moyenne $1/\ln n$ plutôt que la vraie distribution $1/\ln p$, $1/\ln(n-p)$ variable sur $\mathcal{A}$.

6.3. Ce qui bloque, et pourquoi

L'ingrédient 1 ne peut pas être affaibli pour donner une preuve. Ce n'est pas une hypothèse quantitativement trop forte qu'on pourrait assouplir : c'est une hypothèse de nature *probabiliste*, simulant un phénomène qui est en réalité *déterministe* (la primalité dépend de divisibilités exactes, corrélées par construction entre entiers voisins). Aucune version affaiblie de cette hypothèse ne reste dans la catégorie des énoncés démontrables sur les entiers : c'est précisément l'écart, ouvert depuis Hardy-Littlewood (1923), entre heuristique convaincante et preuve.

L'ingrédient 2 pourrait en principe être remplacé par une borne de second moment (Chebyshev), strictement plus faible que Poisson. Mais le contrôle de la variance nécessaire demanderait de contrôler les corrélations entre candidats p distincts, ce qui requiert en l'état de l'art des méthodes de crible (Selberg, grand crible) qui ne donnent que des *majorations* de $\#\mathrm{Dec}(n)$, jamais de minoration non triviale pour tout n - minorer $\#\mathrm{Dec}(n) > 0$ pour tout n est exactement Goldbach lui-même. On ne peut donc pas se défaire de l'ingrédient 2 sans disposer déjà du résultat visé.

L'ingrédient 3, lui, est un simple raffinement technique, sans obstruction conceptuelle : il pourrait être remplacé par une intégration sur la vraie distribution de $\mathcal{A}$, ce qui affinerait la constante sans rien changer à l'ordre de grandeur ni au statut (heuristique vs. prouvé) du résultat.

6.4. Une partie de l'ingrédient 1 peut être rendue déterministe : le lemme de Denise

L'observation suivante, due à Denise Vella-Chemla, permet d'extraire de l'ingrédient 1 une composante qui n'est **pas** heuristique mais strictement démontrée.

Lemme 1 (Synchronisation déterministe des cribles par différence connue)**.** *Soient n, n' pairs, $\delta = n' - n$, et $d \geq 2$ un entier divisant δ. Pour tout entier p,*

$$d \mid (n - p) \iff d \mid (n' - p).$$

Démonstration. $(n' - p) - (n - p) = n' - n = \delta$. Si $d \mid (n - p)$, alors $n - p \equiv 0 \pmod d$, et puisque $d \mid \delta$, $n' - p = (n - p) + \delta \equiv 0 \pmod d$. La réciproque est symétrique. □

Corollaire 2. *Pour q premier impair, $q \mid (n' - n)$: la divisibilité par q de $n - p$ et celle de $n' - p$ sont **le même évènement**, pour un même p - non pas des évènements de même probabilité sous un modèle, mais littéralement la même proposition arithmétique.*

Ce lemme permet de restreindre, sans aucune perte d'information ni la moindre hypothèse probabiliste, l'univers de recherche d'un décomposant commun à $\mathcal{A}(n, B)$ (au lieu de $\{1, \ldots, n\}$ tout entier) : c'est exactement le théorème d'héritage (§ 3), relu au niveau de chaque entier p individuellement plutôt qu'au niveau global des classes de congruence.

6.5. Vérification numérique de la part déterministe

On a testé, pour plusieurs couples (n, δ), le rapport entre le nombre réel de décomposants communs $|\mathrm{Dec}(n) \cap \mathrm{Dec}(n')|$ et l'espérance "naïve" calculée soit sur l'univers brut $\{1, \ldots, n\}$, soit sur l'univers pivot-autorisé $\mathcal{A}(n, B)$ (en utilisant exclusivement le lemme 1, donc sans aucune hypothèse probabiliste supplémentaire). Le détail complet des calculs figure en annexe A ; on résume ici les résultats.

n	δ	$\lvert\mathrm{Dec}(n)\rvert$	$\lvert\mathrm{Dec}(n')\rvert$	$\lvert\cap\rvert$	ratio brut	ratio pivot
10^6	30	10804	10842	2764	23.60	6.29
10^6	210	10804	11316	3581	29.29	5.58
2×10^6	2310	19440	19346	6622	35.22	5.49
1000006	30	9742	8590	2739	32.73	6.55

Ratio observé/espéré, selon que l'espérance naïve est calculée
sur l'univers brut ou sur l'univers pivot-autorisé via le lemme 1.
$\ln(10^6) \approx 13.82$, $\ln(2 \times 10^6) \approx 14.51$.

Première observation (trompeuse). Sur les quatre couples du tableau, tous resserrés autour de $n \sim 10^6$–2×10^6 (soit $\ln n \in [13.8, 14.5]$, un intervalle très étroit), le *ratio brut* dérive nettement avec δ (23.6 à 35.2), tandis que le *ratio pivot* (c'est-à-dire le résiduel *après* application du lemme déterministe de Denise - semble rester à peu près stable, dans la fourchette 5.5–6.6. Cette stabilité apparente, sur une plage de n aussi resserrée, avait initialement suggéré que le résidu pourrait être *borné* par une constante absolue. L'analyse à plus grande échelle qui suit montre que ce n'est pas le cas.

6.6. Croissance réelle du résidu : analyse à grande échelle

On reprend le même protocole ($\delta = 30$, univers pivot sur $\{3, 5\}$), mais sur six échelles de n couvrant un facteur 500 (n de 2×10^5 à 10^8), en ayant soin d'éviter un piège méthodologique : si n (ou $n + \delta$) possède par coïncidence un petit facteur premier impair supplémentaire (typiquement 3 et 5 simultanément), la constante de Hardy-Littlewood locale $\mathfrak{S}(n)$ est elle-même anormalement élevée (facteur $\prod_{q|n} \frac{q-1}{q-2}$, § 5), ce qui gonfle artificiellement $|\mathrm{Dec}(n)|$ et fausse le ratio sans rapport avec le phénomène étudié. On a vérifié que c'est précisément ce qui se produit pour $n = 15\,000\,000 = 2^6 \times 3 \times 5^7$ (qui contient à la fois 3 et 5, et montre un ratio brut anormalement élevé) ; on choisit donc systématiquement, pour chaque échelle, un n pair voisin tel que $n/2$ soit premier, ce qui garantit l'absence de tout petit facteur impair parasite.

n	$\ln n$	$\lvert\mathrm{Dec}(n)\rvert$	$\lvert\cap\rvert$	ratio pivot
200 006	12.21	2 141	673	5.79
1 000 018	13.82	8 121	2 650	6.63
5 000 018	15.43	31 881	7 664	7.19
15 000 026	16.52	82 809	17 534	7.69
40 000 006	17.50	195 037	40 598	8.13
100 000 034	18.42	437 301	82 080	8.57

Ratio pivot pour $\delta = 30$, sur six échelles de n choisies sans petit facteur premier impair parasite
Le ratio croît de façon monotone et régulière avec $\ln n$.

Constat corrigé. Sur cette plage élargie, le ratio pivot croît clairement et régulièrement avec n : ce n'est **pas** une constante bornée. La première observation (§6.5, plage étroite) était un artefact d'échelle insuffisante : $\ln n$ ne variant que de 13.8 à 14.5 sur les quatre premiers essais, une croissance lente passait pour une stabilité.

Identification de la loi de croissance. On a comparé plusieurs lois candidates par régression sur les six points de la table 2 : une constante seule échoue nettement (résidu maximal 1.55, contre moins de 0.1 pour les lois suivantes) ; $\ln \ln n$ donne un ajustement honorable (coefficient directeur ≈ 6.6) ; mais c'est la loi

$$\text{ratio pivot}(n) \;\approx\; \tfrac{1}{2}\ln n + b$$

qui s'ajuste le mieux *à coefficient directeur fixé a priori à* $\frac{1}{2}$ (écart-type des résidus 0.143, contre 0.658 pour $\sqrt{\ln n}$ ajusté de la même façon avec un coefficient libre, qui dérive systématiquement). Le coefficient directeur obtenu par régression libre sur la table 2 est 0.437, sensiblement plus proche de $\frac{1}{2}$ que de 1 (la valeur prédite par l'ingrédient 1 seul, sans le lemme de Denise).

Piste d'explication du facteur $\frac{1}{2}$. L'événement $p \in \mathrm{Dec}(n) \cap \mathrm{Dec}(n')$ couple trois primalités deux à deux : $(p,\ n{-}p)$, $(p,\ n'{-}p)$, et $(n{-}p,\ n'{-}p)$. Le lemme 1 rend la troisième paire **entièrement déterministe** dès que $q \mid \delta$ (c'est ce qui produit le passage de l'univers brut à l'univers pivot). Les deux premières paires, en revanche, restent de nature heuristique (ingrédient 1) : chacune contribuerait, prise isolément, un facteur correctif de l'ordre de $\ln n$ par rapport à une indépendance totale (partage du candidat p). Si l'on ne "paie" cette corrélation heuristique qu'une fois sur les deux paires restantes plutôt que deux fois indépendamment - une hypothèse à vérifier plutôt qu'un fait établi à ce stade - on obtiendrait précisément un facteur $\sqrt{\ln n}$ ou $\frac{1}{2}\ln n$ selon la façon exacte dont les deux corrélations résiduelles se combinent. C'est une piste de nature qualitative, pas une démonstration : c'est précisément le point sur lequel l'avis d'une probabiliste serait précieux.

7. Synthèse et question ouverte

On a établi :

- un théorème démontré (synchronisation du crible pivot par congruence partagée, § 3) ;
- une identification démontrée avec la constante de Hardy-Littlewood (§ 5, détaillée dans [4]) ;
- une vérification numérique exhaustive jusqu'à 10^7 d'une conjecture spécifique (écart 6, § 4) ;
- un lemme démontré (§ 6.4) isolant une composante **non heuristique** de la corrélation entre décomposants de deux entiers pairs liés par une différence connue, et une vérification numérique de sa contribution quantitative, à deux échelles différentes (§ 6.5 et § 6.6) ;
- la mise en évidence, après correction d'un premier artefact d'échelle trop étroite, d'une croissance réelle et régulière du résidu en $\frac{1}{2}\ln n$ (table 2), sensiblement plus lente que le facteur $\ln n$ brut prédit par l'heuristique de Cramér seule.

Reste ouverte la question précise suivante : le facteur $\frac{1}{2}$ observé empiriquement provient-il d'un second mécanisme partiellement déterministe (de même nature que le lemme 1, mais portant sur les paires $(p, n-p)$ et $(p, n'-p)$ plutôt que sur $(n-p, n'-p)$), ou résulte-t-il d'une simple propriété de l'heuristique de Cramér standard qu'une analyse plus soigneuse expliquerait sans rien ajouter de nouveau ? C'est cette question précise que l'auteure souhaite soumettre à une lectrice spécialiste des probabilités.

A. Détail des calculs numériques

Tous les calculs ont été effectués par crible d'Ératosthène (`numpy`), en Python. Le code et les sorties complètes sont disponibles auprès de l'auteure. On donne ici le détail méthodologique pour chaque ligne de la table 1.

A.1. Cas $\delta = 30 = 2 \times 3 \times 5$ ($n = 10^6$)

Univers pivot défini par les facteurs premiers impairs de δ, soit $\{3, 5\}$: $p \in \mathcal{A}(n, 5)$ si $p \not\equiv 0, n \pmod 3$ et $p \not\equiv 0, n \pmod 5$. Taille de cet univers mesurée : 266 666 (densité 0.2667, conforme à la valeur théorique $\delta(n, 5)$).

A.2. Cas $\delta = 210 = 2 \times 3 \times 5 \times 7$ ($n = 10^6$)

Univers pivot sur $\{3, 5, 7\}$. Taille mesurée : 190 476 (densité 0.1905).

A.3. Cas $\delta = 2310 = 2 \times 3 \times 5 \times 7 \times 11$ ($n = 2 \times 10^6$)

Univers pivot sur $\{3, 5, 7, 11\}$.

A.4. Cas de contrôle ($n = 1\,000\,006$, $\delta = 30$)

Repris pour vérifier la stabilité du ratio pivot à δ fixé mais n légèrement différent (non multiple de 3 ni 5 contrairement au cas précédent, modifiant les facteurs $q \mid n$ dans $\delta(n, B)$).

A.5. Calculs à grande échelle (table 2, § 6.6)

Crible d'Ératosthène jusqu'à $N = 10^8 + 50$. Pour chaque échelle cible, recherche du premier n pair voisin tel que $n/2$ soit premier (ce qui garantit l'absence de tout petit facteur premier impair parasite dans n), et tel que ni n ni $n + 30$ ne soient multiples de 3 ou 5. Calcul exhaustif (non échantillonné) de $\mathrm{Dec}(n)$ et $\mathrm{Dec}(n + 30)$ par crible vectorisé (`numpy`), puis de leur intersection. Détection de l'artefact à $n = 15\,000\,000 = 2^6 \times 3 \times 5^7$ par comparaison de la densité observée $|\mathrm{Dec}(n)|/(n/\ln^2 n)$ à la valeur attendue ≈ 2 (rapport anormal 4.01 contre ≈ 2 pour les valeurs voisines), confirmée par factorisation explicite ($n = 15\,000\,000$ et $n - 30 = 14\,999\,970$ partagent les deux facteurs 3 et 5, contrairement à leurs voisins immédiats).

Références

[1] P.J. Davis, R. Hersh, *L'Univers mathématique*, Gauthier-Villars, 1985.

[2] D. Vella, *Les deux partages de décomposants les plus chouettes qu'il m'ait été donné de voir*, note personnelle, ca. 2007 (contient l'observation initiale sur 122, 1802, et les décompositions de Davis-Hersh autour de 20902 et 20962 qui sont à l'origine de la présente note).

[3] D. Vella, *Changer l'ordre sur les entiers naturels pour comprendre le partage des décomposants Goldbach*, note personnelle, octobre 2007 (version développée, avec la notion d'ordre lexicographique sur les représentations par restes chinois et l'argument de descente infinie de Fermat).

[4] D. Vella-Chemla, *Identification de la densité de Vella-Chemla avec la constante singulière de Hardy-Littlewood*, note personnelle, 2026.