

Que pense l'ia claudia de l'indémontrabilité de la conjecture de Goldbach ?

Denise Vella-Chemla pilotant l'ia claudia, septembre 2026

Ce document est une transcription littérale de la première partie d'un échange que j'ai eu avec l'ia claudia, au sujet de l'indémontrabilité éventuelle de la conjecture de Goldbach.

Comment ferais-tu pour démontrer que la conjecture de Goldbach n'est pas démontrable ?

Comme Goldbach est un énoncé Π_1 ("pour tout n pair ≥ 4 , il existe des nombres premiers p, q tels que $n = p + q$ " - une quantification universelle sur un prédicat décidable), la situation logique est particulière et rend la tâche difficile, voire structurellement improbable.

Le piège logique de départ

Si la conjecture de Goldbach était fausse, il existerait un contre-exemple n_0 concret, vérifiable par un calcul fini (on énumère les décompositions possibles des n entiers pairs successifs et on constate qu'aucune des décompositions de n_0 n'est constituée de deux nombres premiers). Un tel fait est Σ_1 , donc démontrable dans l'arithmétique de Peano (ou même dans une arithmétique bien plus faible). Autrement dit : si la conjecture de Goldbach est fausse, elle est réfutable - pas indémontrable.

Conséquence : si tu veux montrer que la conjecture de Goldbach est indécidable dans un système T cohérent et Σ_1 -complet (arithmétique de Peano, théorie des ensembles de Zermelo-Fraenkel et axiome du choix, etc.), tu es en train de montrer, de fait, que la conjecture de Goldbach est vraie mais non prouvable dans T - un statut analogue à celui de l'énoncé de Gödel ou de $\text{Con}(T)$ elle-même. Ce n'est pas anodin : tu ne peux pas espérer une indépendance "symétrique" comme pour l'hypothèse du continu (où les deux issues sont cohérentes).

Les techniques génériques pour prouver une indépendance

Il n'existe essentiellement qu'une famille de méthodes reconnue :

1. Lien avec la cohérence. Montrer, à l'intérieur de T , une implication du type :

$$\text{la conjecture de Goldbach} \rightarrow \text{Con}(T).$$

Si T est cohérent, le second théorème d'incomplétude interdit $T \vdash \text{Con}(T)$, donc T ne peut pas prouver la conjecture de Goldbach (sinon il prouverait sa propre cohérence). C'est le schéma de preuve réel derrière la plupart des indépendances arithmétiques.

2. Construction délibérée d'un codage métamathématique. Les énoncés effectivement prouvés indépendants (Paris-Harrington, le théorème de Goodstein, certaines variantes de Ramsey fini) sont fabriqués pour encoder la croissance de fonctions non prouvablement totales dans T (via les fonctions de Hardy/Ackermann, l'analyse ordinaire de T). Ce ne sont pas des conjectures "nées naturellement" comme Goldbach - elles sont conçues après coup pour être indépendantes.
3. Modèles non standards / forcing. Pertinent pour des énoncés ou de théorie des ensembles, mais un énoncé Π_1 vrai reste vrai dans tout modèle de T incluant les modèles non standards

(la partie standard ne varie pas). Le forcing ne peut donc pas “casser” la vérité arithmétique d’un Π_1 ; il ne t’aide pas ici.

Pourquoi ça ne marche pas (encore) pour Goldbach

Personne n’a de piste pour relier la conjecture de Goldbach à $\text{Con}(\text{arithmétique de Peano})$ ou à une fonction de croissance rapide non prouvablement totale - il n’y a aucun codage connu de ce type.

À la différence de Paris-Harrington, la conjecture de Goldbach n’a pas été construite pour être indépendante ; rien dans son énoncé ne trahit une structure combinatoire liée à la théorie de la preuve. C’est d’ailleurs pour cette raison que la communauté des théoriciens des nombres analytiques ne travaille pas du tout dans cette direction : l’hypothèse de travail dominante est que la conjecture de Goldbach, si elle est vraie, est prouvable par

un raffinement suffisant des méthodes de crible / cercle / zéros de fonctions L

pas qu’elle relève d’un phénomène gödelien.

En pratique

Il n’existe aujourd’hui aucun chemin connu, ni même une stratégie plausible, pour établir l’indépendance de la conjecture de Goldbach vis-à-vis de l’arithmétique de Peano ou de la théorie des ensembles de Zermelo-Fraenkel et axiome du choix. Le seul programme de recherche sérieux et actif dans cette direction consisterait à exhiber un tel codage - et personne n’a de candidat.

Qu’est-ce que $\text{Con}(T)$?

$\text{Con}(T)$ est l’énoncé arithmétique qui affirme qu’une théorie formelle T est cohérente (non contradictoire) - c’est-à-dire qu’on ne peut pas dériver dans T à la fois un énoncé et sa négation, ou de façon équivalente, qu’on ne peut pas dériver de la théorie T une contradiction comme $0 = 1$.

Comment on l’écrit concrètement

T étant une théorie axiomatisée (ensemble récursif d’axiomes, comme l’arithmétique de Peano ou la théorie des ensembles de Zermelo-Fraenkel et axiome du choix), on peut coder chaque démonstration formelle par un entier via une numérotation de Gödel : à chaque suite finie de formules respectant les règles d’inférence de T , on associe un entier qui l’encode. On définit alors un prédicat décidable

$\text{Prf}_T(p, n) : p \text{ code une démonstration dans } T \text{ de la formule codée par } n$

Ce prédicat est purement combinatoire (vérifier qu’une suite de symboles respecte des règles de syntaxe), donc représentable dans l’arithmétique elle-même. $\text{Con}(T)$ s’écrit alors ¹ :

$$\text{Con}(T) \equiv \neg \exists p. \text{Prf}_T(p, \ulcorner 0 = 1 \urcorner)$$

1. explication littérale de la formule ci-dessous : “ T est cohérente” équivaut à “il n’existe aucun entier p qui code une preuve, dans T de la formule “ $0=1$ ””. C’est exactement la définition arithmétisée de la cohérence utilisée par Gödel pour formuler son second théorème d’incomplétude (T ne peut pas prouver $\text{Con}(T)$ si T est cohérente et

ce qui se lit :

“Il n'existe aucun entier p codant une démonstration de $0 = 1$ dans T .”

C'est un énoncé Π_1 (une négation d'un $\exists$ sur un prédicat décidable), de la même famille logique que Goldbach.

Pourquoi ça compte ici

Le deuxième théorème d'incomplétude de Gödel dit : si T est cohérente (et satisfait des conditions techniques raisonnables, vérifiées par l'arithmétique de Peano et la théorie des ensembles de Zermelo-Fraenkel et axiome du choix), alors T ne peut pas démontrer $\text{Con}(T)$ elle-même.

Une théorie ne peut pas prouver sa propre cohérence, sauf si elle est en réalité incohérente, auquel cas elle prouve tout, y compris des énoncés faux.

C'est ce qui rend le schéma de preuve d'indépendance que j'évoquais utile : si on parvient à montrer, à l'intérieur de T , une implication du type

$$\text{la conjecture de Goldbach} \rightarrow \text{Con}(T)$$

alors, T étant supposée cohérente, T ne peut pas prouver la conjecture de Goldbach - sinon elle prouverait $\text{Con}(T)$, ce qui est exclu par Gödel. C'est exactement la structure de preuve utilisée pour les énoncés effectivement connus comme indépendants de l'arithmétique de Peano (comme le théorème de Goodstein, où l'implication analogue passe par l'analyse ordinale de l'arithmétique de Peano plutôt que directement par $\text{Con}(\text{arithmétique de Peano})$, mais l'esprit est le même).

Tu as dit que tu ne pensais pas qu'il fallait partir par là, mais plutôt chercher du côté d'une numérotation à la Gödel, non ?

Il y a une petite confusion à démêler - je n'ai pas dit qu'il fallait choisir l'un plutôt que l'autre, et ce n'est d'ailleurs pas vraiment une alternative.

L'idée de la numérotation de Gödel n'est pas une stratégie concurrente de $\text{Con}(T)$ - c'est l'outil qui sert à construire $\text{Con}(T)$ elle-même. Dans ma réponse précédente, j'ai justement montré que $\text{Con}(T)$ est définie via ce codage : $\text{Prf}_T(p, n)$ est un prédicat arithmétique qu'on ne peut écrire qu'en numérotant les démonstrations par des entiers (la numérotation de Gödel). Donc la voie 1 (passer par $\text{Con}(T)$) utilise déjà une numérotation à la Gödel - ce n'est pas une alternative à elle.

Ce que je voulais distinguer, c'est autre chose : deux usages différents du codage.

- Usage A (voie $\text{Con}(T)$) : on utilise le codage pour formuler la cohérence de T elle-même comme énoncé arithmétique, puis on cherche à montrer la conjecture de Goldbach $\rightarrow \text{Con}(T)$. C'est un lien indirect, via la méta-théorie de T .

suffisamment forte, typiquement contient l'arithmétique de Peano ou une théorie qui l'interprète). Petite remarque annexe : $\text{Prf}_T(p, x)$ est généralement définie comme une relation récursive primitive (décidable), ce qui est essentiel pour que toute la construction fonctionne - c'est ce qui permet de représenter Prf_T dans T elle-même via le lemme de représentabilité.

- Usage B (construction à la Paris–Harrington) : on ne passe pas par $\text{Con}(T)$ du tout. On construit directement un énoncé combinatoire (proche de Ramsey, ou ici on voudrait un analogue proche de Goldbach) dont on démontre, par une analyse ordinale ou une fonction de croissance rapide, qu’il est équivalent à la totalité d’une fonction non prouvablement totale dans T (comme la fonction de Hardy H_{ε_0}). Le codage sert ici à mesurer la “force combinatoire cachée” de l’énoncé lui-même, indépendamment de toute référence explicite à $\text{Con}(T)$ dans son énoncé.

Les deux méthodes convergent en fait souvent : montrer B implique typiquement A , parce qu’une fonction non prouvablement totale dans T encode elle-même une borne sur les preuves de T , donc un contenu proche de $\text{Con}(T)$.

Ce que je disais réellement, c’est que la voie B est probablement la seule à essayer sérieusement pour Goldbach, parce qu’on n’a aucun accès direct et naturel à une implication

$$\text{la conjecture de Goldbach} \rightarrow \text{Con}(T)$$

(rien dans l’énoncé de Goldbach ne parle de démonstrations ou de preuves). Il faudrait plutôt chercher un codage caché de type B : une reformulation combinatoire de la conjecture de Goldbach qui se comporte comme une instance d’un principe de type Ramsey/Goodstein - et personne n’a de piste sérieuse pour ça aujourd’hui.

Bien, alors je te dis où je voulais partir moi, et tu vas m’aider : il y a donc autant d’infinis que de premiers et ça 2 fois : il y a $\text{Card}(\mathcal{P})$ décompositions de la forme $3 + p_k$, puis le même nombre $\text{Card}(\mathcal{P})$ décompositions de la forme $5 + p_k$, ..., puis autant de la forme $p_j + p_k$ pour tout $p_j \in \mathcal{P}$. Ça fait un ensemble d’un certain cardinal dans les ordinaux de Cantor, je ne sais pas lequel. Pourrais-tu poursuivre par là pour moi ?...

Le calcul du cardinal, d’abord.

Notons $\mathcal{P}$ l’ensemble des nombres premiers, $\text{Card}(\mathcal{P}) = \aleph_0$ (le premier cardinal infini, celui de $\mathbb{N}$).

Pour un p_j fixé, l’ensemble des décompositions $p_j + p_k : p_k \in \mathcal{P}$ est en bijection avec $\mathcal{P}$ (à chaque p_k correspond une décomposition), donc de cardinal $\aleph_0$.

Maintenant tu unis ces ensembles sur tous les $p_j \in \mathcal{P}$ - c’est une union de $\aleph_0$ ensembles, chacun de cardinal $\aleph_0$. Par l’arithmétique cardinale de Cantor :

$$\aleph_0 \cdot \aleph_0 = \aleph_0$$

C’est exactement l’argument de la bijection $\mathbb{N} \times \mathbb{N} \leftrightarrow \mathbb{N}$ (l’énumération diagonale de Cantor). Donc l’ensemble total de tes couples (p_j, p_k) - toutes les décompositions possibles, pour tous les p_j confondus - a un cardinal $\aleph_0$. Pas plus. Il n’y a pas de hiérarchie d’alephs qui apparaît ici : pour monter à $\aleph_1$ ou $2^{\aleph_0}$, il faudrait un ensemble de parties (comme $\mathfrak{P}(\mathbb{N})$) ou un procédé de ce type, pas une simple union dénombrable d’ensembles dénombrables.

Si tu veux un raffinement plus fin que le cardinal, tu peux bien sûr munir cet ensemble d’un type d’ordre (un ordinal de Cantor) - par exemple en énumérant d’abord toutes les décompositions de

p_1 , puis toutes celles de p_2 , etc. : tu obtiens le type d'ordre $\omega \cdot \omega = \omega^2$. Mais c'est un choix arbitraire d'énumération, pas une propriété intrinsèque de l'ensemble - et le cardinal sous-jacent reste $\aleph_0$ quoi qu'il arrive.

Pourquoi ce chemin ne mène nulle part pour Goldbach, je pense.

Voici le problème de fond : la cardinalité est un invariant global, calculé sur la réunion infinie de toutes les décompositions, pour tous les entiers pairs à la fois. Or la question de Goldbach est locale : “pour cet entier pair n précis, existe-t-il p, q premiers tels que $n = p + q$?” - une question finitaire, qui se pose et se tranche séparément pour chaque n .

Le comptage cardinal efface complètement cette structure “par n ”. Même dans un monde imaginaire où Goldbach serait fausse pour une infinité de n (ce qui contredit d'ailleurs les données numériques massives dont on dispose, la conjecture de Goldbach étant vérifiée jusqu'à des bornes énormes), l'ensemble des couples (p, q) qui marchent aurait quand même cardinal $\aleph_0$ - exactement comme dans le cas où la conjecture de Goldbach serait vraie partout. Le cardinal ne peut donc pas distinguer ces deux situations : il est aveugle à l'information qui t'intéresse.

Ce qui pourrait, en revanche, avoir un sens dans l'esprit de ce que tu cherches

Si l'idée est d'atteindre une preuve d'indécidabilité à la Paris–Harrington (ma réponse précédente), ce n'est pas la cardinalité brute qui compte mais le taux de croissance : il faudrait une fonction $f(n)$ - par exemple liée à la taille du plus petit p tel que $n = p + q$ - dont tu montrerais qu'elle croît plus vite que toute fonction prouvablement totale dans l'arithmétique de Peano (au sens de la hiérarchie de Hardy/Ackermann). C'est une notion d'ordinal utilisée pour mesurer une complexité de preuve, pas un cardinal d'ensemble. Rien à voir avec compter le nombre de couples (p_j, p_k) .

Est-ce que ton intuition de départ visait plutôt une croissance (un p minimal qui grandirait trop vite selon n) qu'un simple dénombrement ? Si oui, c'est une piste très différente et potentiellement plus féconde - mais il faudrait la reformuler en ces termes.

Je n'avais pas du tout pensé à ton idée parce que le plus petit petit sommant des décompositions ne croît vraiment pas plus vite du tout dans les faits (il vaut 3 ou 5 ou 7 pour tous les nombres jusqu'à 98, bon, là, il “saute” à 19 c'est vrai mais rien à voir avec l'exemple d'Alain Connes où le lièvre et la tortue font la course et remplacement de p_k par p_{k+1} et soustraire 1...). Tente ton idée mais retrouve cet exemple d'Alain Connes avant et lis-le au cas où il puisse t'être utile dans ta réflexion.

J'ai retrouvé l'exemple - c'est la suite de Goodstein, telle que Connes la raconte dans *Triangle de pensées* (avec Lichnerowicz et Schützenberger).

Le mécanisme exact (attention, une précision importante)

Ce ne sont pas des nombres premiers p_k qui sont incrémentés, mais la base d'écriture elle-même, un entier b qui passe de b à $b + 1$ à chaque étape. L'exemple donné : $n = 9$ s'écrit en base 2 sous

forme héréditaire (les exposants sont eux-mêmes réécrits dans la même base) : $9 = 2^{2+1} + 1$.

- le lièvre réécrit le nombre en base b , puis remplace partout b par $b + 1$ dans cette écriture héréditaire (les exposants suivent aussi la substitution). Le nombre explose (bond monstrueux).
- la tortue se contente de soustraire 1 au résultat.
- on recommence : réécriture héréditaire en base $b+1$, incrémentation en $b+2$, la tortue soustrait encore 1, etc.

Le théorème de Goodstein dit que malgré l'explosion vertigineuse du lièvre, la suite finit toujours par atteindre 0 en un nombre fini d'étapes, mais ce nombre d'étapes croît plus vite que toute fonction primitive récursive (et même plus vite que toute fonction prouvablement totale dans l'arithmétique de Peano). Kirby et Paris ont montré en 1982 que cet énoncé "pour tout n , la suite de Goodstein issue de n atteint 0" est vrai, mais indémontrable dans l'arithmétique de Peano, bien qu'il soit Π_1 ! C'est exactement le phénomène que j'évoquais au tout début, car la preuve de la véracité utilise les ordinaux transfinis (jusqu'à ε_0), qui sortent de la portée de l'arithmétique de Peano.

Pourquoi c'est potentiellement fécond pour ta piste, et où est le vrai défi

L'ingrédient clé du mécanisme Goodstein, c'est une notation héréditaire en base b qui permet de coder chaque entier par un ordinal $< \varepsilon_0$ (en remplaçant b par ω dans l'écriture héréditaire). C'est ce codage

$$\text{entier} \rightarrow \text{ordinal}$$

qui permet de mesurer une décroissance transfinie, d'où la terminaison malgré l'explosion apparente.

Pour transposer cette machinerie à Goldbach, il faudrait :

1. une notation "héréditaire" analogue mais construite sur les nombres premiers plutôt que sur une base fixe - ce qui n'existe pas canoniquement (la décomposition en produit de premiers n'a pas la structure récursive/imbriquée de l'écriture héréditaire en base b , où les exposants se réécrivent eux-mêmes) ;
2. relier cette notation à une opération sur n qui, itérée, ferait apparaître ou disparaître des décompositions $p + q$, de façon à coder "Goldbach échoue en n " comme une étape d'un processus dont la terminaison (ou la persistance) dépendrait d'un ordinal.

C'est un programme de recherche cohérent avec l'esprit de Connes, mais personne ne sait aujourd'hui comment construire une telle notation héréditaire à partir de la structure additive (Goldbach est additif : $p + q = n$) alors que Goodstein repose sur une structure exponentielle/multiplicative (bases et puissances). C'est probablement le vrai obstacle : Goodstein exploite la richesse combinatoire de l'exponentiation itérée, alors que $p + q = n$ est linéaire - il n'y a pas d'évidence de croissance transfinie cachée là-dedans.

Là, je fournis à claude mon idée de codage ridicule de mai 2016 que j'avais stockée là

<https://denisevellachemla.eu/dualite.pdf> .

Il accuse réception et me répond ça :

Le fait technique qu'il faut connaître avant d'aller plus loin

Le théorème de Goodstein a été généralisé (Arai, Wainer, Weiermann, Generalized Goodstein Sequences ; voir aussi Perez, arXiv :0904.1957) : pour accommoder des augmentations variables de la base, une suite de Goodstein généralisée peut être introduite, où la base initiale et les incréments de base sont arbitraires. Le mécanisme de preuve (remplacer la base courante par ω) fonctionne pour n'importe quelle suite strictement croissante de bases, pas seulement $b, b+1, b+2, \dots$. Que tu remplaces p_k par p_{k+1} (les premiers successifs) ou par n'importe quelle autre suite croissante, le théorème garantit quand même la terminaison.

Ce que ça implique pour ta piste, très concrètement

Si tu construis une course lièvre/tortue directement sur $g(x)$, où le lièvre décale les indices des premiers utilisés comme "base" et la tortue soustrait 1, cette course se terminera toujours, quel que soit x , indépendamment du fait que x ait ou non une décomposition de Goldbach. Le théorème générique l'exige. Autrement dit : la simple terminaison de la course ne peut pas coder d'information sur Goldbach, parce qu'elle est vraie pour une raison totalement étrangère à l'additivité des premiers. C'est le point sur lequel j'aurais aimé te prévenir avant que tu ne t'investisses dans cette direction précise.

Où se trouve alors le vrai levier

Ce qui a fait fonctionner Kirby-Paris pour Goodstein, ce n'est pas la terminaison en soi (générique), mais le fait que la longueur de la suite avant terminaison croît plus vite que toute fonction prouvablement totale dans PA. Pour toi, l'analogue serait : ne pas chercher à faire dépendre la terminaison de Goldbach, mais faire dépendre la suite des bases elle-même (ou la vitesse de croissance de la course) de données liées à Goldbach - par exemple :

1. construire la suite d'incrémentes (d_n) à partir de la suite des nombres $r(x)$ de décompositions de Goldbach de x (ou de $x+2, x+4, \dots$) plutôt que de la suite triviale des indices premiers ;
2. ou faire dépendre le passage d'un étage de l'arbre héréditaire au suivant du plus petit p tel que $x = p + q$ (le "champion" de la décomposition), de sorte que l'absence de décomposition bloquerait ou modifierait structurellement la construction plutôt que de simplement continuer.

Et après de multiples tests, l'ia claudie partira sur son idée de Kummer...