La Méthode du Cercle

Rapport de Stage de L3
Mai-Juin 2023

Simon Pitte,
ENS de Lyon

encadré par **Didier Lesesvre,**
Université de Lille

Table des matières

Notations

On notera $A \ll B$ pour $A = O(B)$.

On notera $e(x) = e^{2i\pi x}$.

1 Motivation et idée générale

1.1 Le problème

On se donne une partie $A \subseteq \mathbb{N}$, et on veut savoir s'il existe un entier s tel que pour tout entier N, ou au moins tout entier assez grand

$$\exists x_1, \dots, x_s \in A,\ N = x_1 + \cdots + x_s.$$

Exemples :

— $A = \{x^2, x \in \mathbb{N}\}$ (on écrit N comme somme de carrés.)

— $A = \{x^k, x \in \mathbb{N}\}$ (c'est le problème de Waring.)

— $A = \{p \text{ premier}\}$ (la conjecture de Goldbach est que $s = 2$ dans ce cas.)

On peut aussi se donner, de manière plus générale, $(A_n)_{n\in\mathbb{N}}$ suite de parties de $\mathbb{N}$, et chercher un entier s tel que pour tout entier N, ou au moins tout entier assez grand

$$\exists x_1 \in A_1, \dots, x_s \in A_s,\ N = x_1 + \cdots + x_s.$$

Exemples :

— $A_n = \{x^{n+1}, x \in \mathbb{N}\}$ (on écrit N comme somme de puissances croissantes.)

— $A_n = \{x^{k_n}, x \in \mathbb{N}\}$ pour une suite $(k_n)_{n\in\mathbb{N}}$ fixée.

Remarque : La difficulté du problème, et la nécessité de la méthode, vient de la nature additive de la question posée, en dépit de la structure souvent multiplicative de A.

1.2 Idée générale de la méthode

Posons, pour $\alpha \in \mathbb{R}$,

$$f(\alpha) := \sum_{n\in A,\ n\leqslant N} e(\alpha n)$$

la fonction génératrice (partielle) de A, et $R(N)$ le nombre de façons d'écrire N comme somme de s éléments de A. Constatons alors que

$$F(\alpha) := f(\alpha)^s = \sum_{n_1,\dots,n_s\in A} e(\alpha(n_1 + \cdots + n_s)) = \sum_{n\in\mathbb{N},\ n\leqslant sN} R(n)e(\alpha n)$$

en regroupant les termes selon la somme $n_1 + \cdots + n_s$. Ainsi $R(N)$ est le N-ième coefficient de Fourier de F, et on obtient la reformulation analytique de $R(N)$ suivante :

$$\int_0^1 F(\alpha)e(-\alpha N)d\alpha = R(N). \tag{1}$$

Toute la méthode du cercle consiste à estimer cette intégrale.

Remarque : Dans le cas ou plutôt qu'un unique A on a une suite $(A_n)_n$ de parties de $\mathbb{N}$, on prendrait alors f_n la fonction génératrice partielle de A_n, et on poserait :

$$F(\alpha) := \prod_{n=1}^{s} f_n(\alpha)$$

et ce qui suit est identique.

Notons que le changement de variable $z = e(\alpha)$ donne une intégrale sur $\mathbb{S}^1$, d'où le nom de la méthode. Hardy et Littlewood, qui ont créé cette méthode, prenaient directement une série entière sur le disque comme fonction génératrice, obtenant (1) comme conséquence du théorème des résidus. Le changement de formalisme de la méthode qui consiste à prendre des sommes finies d'exponentielles (et qui se révèle être un point de vue plus simple à exploiter) est dû à Vinogradov.

La somme d'exponentielle $F(\alpha)$ est maximale en 0, et de manière générale est plus grande en les rationnels à petits dénominateurs. L'heuristique dépend du problème étudié : par exemple dans le cas des sommes de carrés, Hardy et Ramanujan obtiennent, lorsque ρ tend vers 1 :

$$\sum_{n=1}^{\infty} \rho^{n^2} e\left(\frac{an^2}{q}\right) \sim C\frac{S(q,a)}{q}(1-\rho)^{-1/2}$$

où $S(q,a) = \sum_{m=1}^{q} e(am^2/q)$ est de l'ordre de $\sqrt{q}$. Ainsi la série génératrice est de « taille »$1/\sqrt{q}$ en $e(a/q)$. Ci-dessous on a représenté le module de la fonction d'Euler $\phi(z) = \prod_{k=1}^{\infty}(1-z^k)$ (la fonction génératrice des partitions d'un entier) à titre d'exemple, où l'on constate bien le phénomène annoncé.

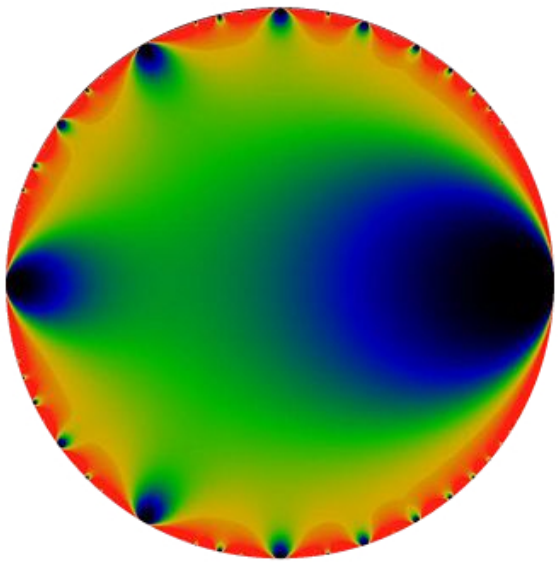

FIGURE 1 – Module de la fonction d'Euler sur le disque unité *(image tirée de Wikipédia)*

Pour maîtriser l'approximation des réels par les rationnels à « petits dénominateurs », on va employer le théorème suivant :

Théorème 1 (Dirichlet)**.** *Soit $\alpha \in \mathbb{R}$ et $X \geqslant 1$. Alors il existe un entier $q \leqslant X$ et $a \in \mathbb{N}$ tel que*

$$\left|\alpha - \frac{a}{q}\right| \leqslant \frac{1}{qX}.$$

Remarque : En particulier on a $\left|\alpha - \frac{a}{q}\right| \leqslant \frac{1}{q^2}$.

Démonstration. Quitte à changer X en $\lfloor X \rfloor$, on suppose X entier. Pour $1 \leqslant q \leqslant X$, soit $\alpha_q := \alpha q - \lfloor \alpha q \rfloor \in [0,1[$. On divise $[0,1[$ en les intervalles semi-ouverts $I_r := \left[\frac{r}{X}, \frac{r+1}{X}\right[$, pour $0 \leqslant r \leqslant X-1$. Si pour un certain q, $\alpha_q \in I_0$, alors $|\alpha_q| \leqslant 1/X$ donc avec $a = \lfloor \alpha q \rfloor$,

$$\left|\alpha - \frac{a}{q}\right| \leqslant \frac{1}{qX}.$$

Sinon par le principe des tiroirs, il existe $1 \leqslant u, v \leqslant X$ et $1 \leqslant r \leqslant X-1$ tels que $\alpha_u, \alpha_v \in I_r$. Donc $|\alpha_u - \alpha_v| \leqslant 1/X$, *i.e.* $|\alpha(u-v) - (\lfloor \alpha u \rfloor - \lfloor \alpha v \rfloor)| \leqslant 1/X$. Ainsi avec $q = |u-v| \leqslant X$ et $a = \lfloor \alpha u \rfloor - \lfloor \alpha v \rfloor$, on a bien

$$\left|\alpha - \frac{a}{q}\right| \leqslant \frac{1}{qX}.$$

□

Ainsi pour $Q > 1$ fixé, tout réel est au plus à $1/q^2$ d'un rationnel à dénominateur $q \leqslant Q$. Il s'agit donc de distinguer entre un réel proche d'un tel rationnel, et ceux qui en sont encore plus proches.

Pour ce faire, on va fixer $0 < \tau < 1$, $C > 2\tau$, et définir, pour $1 \leqslant a \leqslant q \leqslant N^\tau$, $a \wedge q = 1$

$$\mathfrak{M}(q,a) := \left[\frac{a}{q} - N^{-C}; \frac{a}{q} + N^{-C}\right]$$

qui sont deux-à-deux disjoints, puisque pour $1 \leqslant a \leqslant q \leqslant N^\tau$ et $1 \leqslant a' \leqslant q' \leqslant N^\tau$

$$\left|\frac{a}{q} - \frac{a'}{q'}\right| = \left|\frac{aq' - a'q}{qq'}\right| \geqslant \frac{1}{qq'} \geqslant N^{-2\tau} = N^{C-2\tau} N^{-C} \geqslant 2N^{-C}$$

pour N assez grand. On pose ensuite $\mathfrak{M} := \bigcup_{q,a} \mathfrak{M}(q,a)$. On les appelle les *arcs majeurs*.
On pose aussi $\mathfrak{m} := [N^{-C}; 1 + N^{-C}] \setminus \mathfrak{M}$ les *arcs mineurs*.

L'idée est que, avec une qualité d'approximation qui dépend de N, C, et τ,

$$\int_{\mathfrak{M}} F \approx \sum_{q,a} F\left(\frac{a}{q}\right).$$

Remarque : Notons que par le théorème de Dirichlet, un réel dans les arcs mineurs n'est jamais à plus de $1/q^2$ d'un rationnel à « petit dénominateur ».

Pour τ, C bien choisis pour le problème, on cherche à avoir

$$\int_{\mathfrak{m}} F(\alpha) e(-\alpha N) d\alpha = o\left(\int_{\mathfrak{M}} F(\alpha) e(-\alpha N) d\alpha\right),$$

de sorte que l'intégrale sur $\mathfrak{M}$ représente la contribution principale. Et si en plus on arrive à obtenir que

$$\int_{\mathfrak{M}} F(\alpha) e(-\alpha N) d\alpha \xrightarrow[N\infty]{} +\infty$$

on aura $R(N) \xrightarrow[N\infty]{} +\infty$. En particulier à partir d'un certain rang, $R(N) \geqslant 1$, *i.e.* il y a au moins une façon d'écrire N sous la forme voulue, ce qui conclut.

Dans un contexte concret, on se réferera à l'heuristique suivante : en majorant simplement la somme d'exponentielles f par $f(0)$ (*i.e.* le nombre de termes), on obtient la majoration triviale $F(\alpha) \ll F(0)$.
Très souvent, on obtient que $\int_{\mathfrak{M}} F(\alpha)e(-\alpha N)d\alpha \gg F(0)N^{-1}$ ou $\gg F(0)N^{-1-\varepsilon}$ pour tout $\varepsilon > 0$.
Ainsi la majoration à viser sur les arcs mineurs est de la forme

$$\int_{\mathfrak{m}} |F| \ll F(0)N^{-1-\delta}$$

pour un certain $\delta > 0$ fixé. On doit donc gagner $N^{1+\delta}$ sur la majoration triviale.

2 Estimations des arcs

2.1 Les arcs mineurs

De manière générale, on utilise deux types de lemmes pour les arcs mineurs. Tout d'abord, le lemme de Weyl.

Lemme 2 (Lemme de Weyl)**.** *Soit $\phi(x)$ polynôme réel de degré k et de coefficient dominant α, tel que l'on ait $|\alpha - a/q| \leqslant 1/q^2$. Alors*

$$\sum_{1 \leqslant x \leqslant Q} e(\phi(x)) \ll Q^{1+\varepsilon}\left(\frac{1}{q} + \frac{1}{Q} + \frac{q}{Q^k}\right)^{2^{1-k}}.$$

Remarque : La puissance 2^{1-k} le rend pratiquement inutile si $k > 3$, puisque l'estimée triviale est $\ll Q$. Notons que le terme $\frac{1}{q} + \frac{1}{Q} + \frac{q}{Q^k}$ impose que l'on choisisse q ni trop petit, ni trop grand par rapport à Q.

Idée de démonstration. Posons $T = \sum_{1 \leqslant x \leqslant Q} e(\phi(x))$ et notons que

$$|T|^2 = \sum_{1 \leqslant x,y \leqslant Q} e(\phi(x) - \phi(y)) = \sum_{h=1-Q}^{Q-1} \sum_{x \in I_1(h)} e(\phi(x+h) - \phi(x))$$

avec $I_1(h)$ intervalle entier dépendant de h. De plus, $\phi(x+h) - \phi(x)$ est un polynôme en x de coefficient dominant $k\alpha$, et de degré $k-1$.

Si on note $T_1(h) = \sum_{x \in I_1(h)} e(\phi(x+h) - \phi(x))$, on peut refaire la même chose, et par récurrence obtenir $|T|^K$ comme somme d'exponentielle de fonctions affines, dont l'estimation est plus facile.

Le reste de la preuve est purement technique et sans réelle difficulté (cf. [1]). Il est à noter cependant que la constante implicite ne dépend pas de α, mais seulement de $\alpha - a/q$. □

Pour les puissances plus grandes, le lemme de Weyl perd de son intérêt. Ainsi si on veut par exemple écrire les entiers comme somme de puissances croissantes, on aura :

$$F(\alpha) = \prod_{k=2}^{s} f_k(\alpha) \text{ avec } f_k(\alpha) = \sum_{x=1}^{N^{1/k}} e(\alpha x^k),$$

et on appliquera le lemme de Weyl uniquement à f_2, éventuellement f_3 :

$$\int_{\mathfrak{m}} |F(\alpha)|d\alpha \leqslant \sup_{\alpha \in \mathfrak{m}} |f_2(\alpha)| \sup_{\alpha \in \mathfrak{m}} |f_3(\alpha)| \int_0^1 \prod_{k=4}^{s} |f_k(\alpha)|d\alpha.$$

Pour le reste, on utilise Cauchy-Schwarz ou Hölder. Si on écrit $[\![4;s]\!] = J_1 \cup J_2$:

$$\int_0^1 \prod_{k=4}^{s} |f_k(\alpha)| d\alpha \leqslant \left(\int_0^1 \prod_{k \in J_1} |f_k(\alpha)|^2 d\alpha \right)^{1/2} \left(\int_0^1 \prod_{k \in J_2} |f_k(\alpha)|^2 d\alpha \right)^{1/2}$$

ou bien avec $\sum 1/2a_k = 1$, $a_k \in \mathbb{N}^*$:

$$\int_0^1 \prod_{k=4}^{s} |f_k(\alpha)| d\alpha \leqslant \prod_{k=4}^{s} \left(\int_0^1 |f_k(\alpha)|^{2a_k} d\alpha \right)^{1/2a_k}.$$

On peut estimer ces intégrales en prenant avantage du fait suivant, qui éclaire l'utilisation de Cauchy-Schwarz et Hölder. En développant une intégrale de la forme $\int_0^1 \prod_{n=1}^{m} |f_{k_n}(\alpha)|^2 d\alpha$, on obtient que cette intégrale compte exactement le nombre de solutions à l'équation diophantienne :

$$x_1^{k_1} + \cdots + x_m^{k_m} = y_1^{k_1} + \cdots + y_m^{k_m} \tag{2}$$

pour $x_n, y_n \leqslant N^{1/k_n}$. Pour majorer cela, on utilise ensuite des théorèmes de valeurs moyennes. On peut par exemple utiliser dans certains cas le lemme de Hua :

Lemme 3 (Lemme de Hua)**.** *Soit $k \geqslant 2$, $1 \leqslant j \leqslant k$. Alors :*

$$\int_0^1 |f_k(\alpha)|^{2^j} \ll N^{2^j - j + \epsilon}$$

Mais dans les cas où l'on a pas une puissance de 2 en exposant, ou si le lemme de Hua ne donne pas une majoration suffisante, on emploie des résultats qui permettent de compter les solutions de (2). Par exemple, sont d'usage courant plusieurs résultats itératifs dûs à Davenport, dont le lemme suivant (qui est une amélioration par Vaughan d'un résultat de Davenport).

Lemme 4 (Davenport, Vaughan)**.** *Soit $k \geqslant 3$, $P \geqslant 1$, $C > 0$, $1 - 1/k \leqslant \lambda \leqslant 1$, et $\nu = k\lambda - k + 1$.*
Notons S le nombre de solutions de

$$x_1^{k_1} + \cdots + x_m^{k_m} = y_1^{k_1} + \cdots + y_m^{k_m} \leqslant CP^{k\lambda}$$

avec conditions éventuelles sur les x_n, y_n.

Notons T le nombre de solutions de

$$x_1^{k_1} + \cdots + x_m^{k_m} + x^k = y_1^{k_1} + \cdots + y_m^{k_m} + y^k$$

avec $P \leqslant x, y \leqslant 2P$, $\sum x_n^{k_n}, \sum y_n^{k_n} \leqslant CP^{k\lambda}$, et les mêmes conditions éventuelles sur les x_n, y_n.

Notons enfin Σ le nombre de m-uplets $(x_1, \ldots, x_m)$ tels que, sous les mêmes conditions éventuelles, $\sum x_n^{k_n} \leqslant CP^{k\lambda}$.

Alors pour $1 \leqslant j \leqslant k-2$:

$$T \ll PS + P^{1+\nu-2^{1-j}} S + P^{1+\varepsilon+\nu(1-2^{-j})-(j+1)2^{-j}} S^{1-2^{-j}} \Sigma^{2^{1-j}}$$

Remarque : Constatons bien qu'on a enlevé deux inconnues à droite. T compte les solutions d'une équation à $2m+2$ inconnues, alors que S compte les solutions d'une équation à $2m$ inconnues, d'où le caractère itératif du lemme. La preuve, qu'on trouve dans [4], est peu instructive. On utilise ce lemme selon la méthode itérative suivante.

Supposons que $k_1 \geqslant \ldots \geqslant k_m$. On commence par résoudre la version simplifiée de (2) suivante :

$$x_1^{k_1} + x_2^{k_2} = y_1^{k_1} + y_2^{k_2} \tag{3}$$

Lemme 5. *Si on note S le nombre de solutions de l'équation (3), alors pour tout $\varepsilon > 0$,*

$$S \ll N^{1/k_1+1/k_2+\varepsilon}.$$

Démonstration. Si $x_1 \neq y_1$, et donc $x_2 \neq y_2$, on la réécrit $x_1^{k_1} - y_1^{k_1} = y_2^{k_2} - x_2^{k_2}$. Alors pour $N^{1/k_1}/2 \leqslant x_1 \neq y_1 \leqslant N^{1/k_1}$ fixés, et avec $p := x_1^{k_1} - y_1^{k_1}$, on a donc que (3) se réécrit

$$(y_2 - x_2)(y_2^{k_2-1} + y_2^{k_2-2}x_2 + \cdots + x_2^{k_2-1}) = p.$$

Ainsi pour toute solution, $y_2 - x_2$ est un diviseur de p. Et en outre, chaque diviseur de p ne peut correspondre qu'à au plus une solution, puisque si $y_2 - x_2 = d$, on a

$$y_2^{k_2-1} + y_2^{k_2-2}x_2 + \cdots + x_2^{k_2-1} = (x_2+d)^{k_2-1} + (x_2+d)^{k_2-2}x_2 + \cdots + x_2^{k_2-1}$$

qui est un polynôme strictement croissant en $x_2 \in \mathbb{N}$. Ainsi pour chaque p, on a au plus $d(p) \ll p^{\varepsilon} \ll N^{\varepsilon}$ solutions, pour tout $\varepsilon > 0$. Comme on a au plus N^{2/k_1} valeurs de p, ceci compte pour au plus $N^{2/k_1+\varepsilon}$ solutions de (3).

Si $x_1 = y_1$, alors $x_2 = y_2$, et ceci donne au plus $N^{1/k_1+1/k_2}$ solutions de (3). Ainsi, comme $1/k_1+1/k_2 \geqslant 2/k_1$ on a au plus $N^{1/k_1+1/k_2+\varepsilon}$ solutions de (3). □

On utilise ensuite le lemme 4 pour majorer le nombre de solutions de

$$x_1^{k_1} + x_2^{k_2} + x_3^{k_3} = y_1^{k_1} + y_2^{k_2} + y_3^{k_3}$$

en constatant que quitte à prendre N assez grand, la condition sur λ ne nous empêche pas de prendre ν arbitrairement petit. Ainsi, on aura toujours $PS \gg P^{1+\nu-2^{1-j}}S$ peu importe j. Reste alors à choisir j pour que le dernier terme soit le plus petit possible, voire plus petit que PS.

On continue ensuite itérativement en ajoutant de plus en plus de variables.

Remarque 1 : En pratique, il est toujours avantageux, tant que faire se peut, de commencer par les grandes puissances, puis d'affiner en utilisant ce lemme sur des puissances de plus en plus petites.

Remarque 2 : On trouvera un exemple de mise en oeuvre de cette méthode itérative dans la section 4.2.

2.2 Les arcs majeurs

Pour les arcs majeurs, il y a deux étapes dans la méthode. D'abord, on construit une fonction $G(\alpha; a, q)$ qui approche $F(\alpha)$ sur $\mathfrak{M}(q,a)$. Il reste ensuite à majorer

$$\Delta := \int_{\mathfrak{M}} |F(\alpha) - G(\alpha)| d\alpha,$$

de sorte que $\int_{\mathfrak{M}} F \approx \int_{\mathfrak{M}} G$, après quoi il suffit donc de minorer

$$I_N := \int_{\mathfrak{M}} G(\alpha) e(-\alpha N) d\alpha.$$

L'erreur Δ se majore avec une précision satisfaisante par des inégalités en général assez grossières, dès lors que les C et τ de la construction des arcs sont bien choisis.
En fait, on prend souvent τ presque arbitraire jusque là, et $C = C(\tau)$. C'est la borne obtenue sur Δ qui donnera une contrainte, parfois très forte, sur τ pour garantir qu'elle soit effective, et fixera donc une valeur de τ.

Une construction usuelle de G est la suivante. On se donne $\theta(m)$ une approximation plus « lisse » de $\mathbb{1}_A(m)$ (par exemple $\theta(m) = m^{1/k-1}/k$ la densité de probabilité que m soit une puissance k-ième, si on traite les puissances k-ièmes, ou encore $\theta = \Lambda$ fonction de Von Mangoldt si on traite les nombres premiers). Puis on pose :

— $v(\beta) := \sum_{n \leqslant N} \theta(n) e(\beta n)$
qui est une première approximation de f, puisqu'on la définit comme f mais en remplaçant $\mathbb{1}_A$ par θ.

— $S(q,a) := \sum_{n=1}^{q} t_n e(an/q)$
où les t_n sont des coefficients de somme q qui traduisent la proportion d'éléments de A congrus à n modulo q. Par exemple, si on traite le cas des sommes de puissances k-ièmes, on prendra $S(q,a) := \sum_{x=1}^{q} e(ax^k/q)$, de sorte que t_n est ici le nombre de solutions de $x^k \equiv n \pmod q$ pour $1 \leqslant x \leqslant q$.

Alors pour $\alpha \in \mathfrak{M}(q,a)$:

$$\begin{aligned} f(\alpha) &= \sum_{n \in A,\ n \leqslant N} e(\alpha n) = \sum_{n \in A,\ n \leqslant N} e\left(\left(\alpha - \frac{a}{q}\right) n\right) e\left(\frac{an}{q}\right) \\ &= \sum_{n \leqslant N} \mathbb{1}_A(n) e\left(\left(\alpha - \frac{a}{q}\right) n\right) e\left(\frac{an}{q}\right) \\ &\approx \sum_{n \leqslant N} \theta(n) e\left(\left(\alpha - \frac{a}{q}\right) n\right) q^{-1} S(q,a) \\ &= q^{-1} S(q,a) v\left(\alpha - \frac{a}{q}\right) =: g(\alpha; q, a). \end{aligned}$$

Puis on définit $G(\alpha; q, a)$ comme $F(\alpha)$, mais en remplaçant les $f(\alpha)$ par des $g(\alpha; q, a)$, c'est à dire

$$G(\alpha; q, a) = g(\alpha; q, a)^s$$

si on a une unique partie $A \subseteq \mathbb{N}$, ou

$$G(\alpha; q, a) = \prod_{n=1}^{s} g_n(\alpha; q, a)$$

si on a une suite $(A_n)_n$ de parties, où g_n est construite à partir de f_n fonction génératrice de A_n.
Si on traite à titre d'exemple le cas $F(\alpha) = f(\alpha)^s$, on a :

$$\begin{aligned} I_N &= \sum_{q \leqslant N^\tau} \sum_{a \wedge q = 1} \int_{\mathfrak{M}(q,a)} q^{-s} S(q,a)^s v\left(\alpha - \frac{a}{q}\right)^s e(-\alpha N) d\alpha \\ &= \sum_{q \leqslant N^\tau} \sum_{a \wedge q = 1} q^{-s} S(q,a)^s e\left(-\frac{aN}{q}\right) \int_{-N^{-C}}^{N^{-C}} v(\beta)^s e(-\beta N) d\beta \\ &= \mathfrak{S}^*(N) J^*(N) \end{aligned}$$

où $\mathfrak{S}^*(N)$ s'appelle la *série singulière* (partielle) et $J^*(N)$ l'*intégrale singulière* (partielle). On obtiendra toujours une factorisation de cette forme.

Remarque : Notons que par cette factorisation, on a séparé I_N en un objet purement arithmétique ($\mathfrak{S}$, qui ne contient aucune intégrable et aucun paramètre continue) d'un objet purement analytique (J, qui ne contient ni a, ni q).

Pour mieux estimer ces objets, on « complète » les deux facteurs singuliers, *i.e.* on approche $\mathfrak{S}^*(N)$ par

$$\mathfrak{S}(N) := \sum_{q=1}^{\infty} \sum_{a \wedge q = 1} q^{-s} S(q,a)^s e\left(-\frac{aN}{q}\right)$$

et on approche $J^*(N)$ par

$$J(N) := \int_{-1/2}^{1/2} v(\beta)^s e(-\beta N) d\beta.$$

Les restes sont bien contrôlés, par une convergence uniforme en N de la somme partielle $\mathfrak{S}^*$ vers la série $\mathfrak{S}$, et par décroissance assez rapide de $v(\beta)$ loin de $\beta = 0$.

Traitons d'abord $J(N)$. C'est, par définition, le N-ième coefficient de Fourier de $v(\beta)^s$. En développant $v(\beta)^s$, on obtient donc que

$$J(N) = \sum_{x_1 + \cdots + x_s = N} \prod_{i=1}^{s} \theta(x_i).$$

Les détails dépendent énormément de θ, et donc du problème, mais en général, avec une minoration plus ou moins fine des partitions de N, on peut obtenir $J(N) \gg F(0) N^{-1}$.

Traitons maintenant $\mathfrak{S}(N)$. Encore une fois, les détails dépendent beaucoup du problème, mais $\mathfrak{S}(N)$ a une structure multiplicative très forte, ce qui permet de la décomposer sous la forme d'un produit eulérien, dont il faut minorer chaque facteur. En effet, si on note

$$D(q) := \sum_{a \wedge q = 1} q^{-s} S(q,a)^s e\left(-\frac{aN}{q}\right),$$

on obtient bien souvent que D est une fonction multiplicative. Ainsi :

$$\mathfrak{S}(N) = \prod_{p} \sum_{\ell=0}^{\infty} D(p^\ell).$$

On se ramène donc à minorer chaque facteur du produit. Cette minoration se ramène souvent aux équations suivantes, pour chaque p premier et ℓ entier :

$$\sum_{i=1}^{s} x_i \equiv N \pmod{p^\ell}$$

pour $x_1, \ldots, x_s \in A$ et $\leqslant N$, dont il faut montrer qu'elle a suffisamment de solutions, en tout cas asymptotiquement. Il est à noter que ceci n'est pas du tout aussi dur que le problème original, puisque les équations dans $\mathbb{Z}/p^\ell\mathbb{Z}$ sont bien plus faciles que celles dans $\mathbb{Z}$.

Plus explicitement, le lien se fait par le lemme suivant dans le cas où l'on traite des sommes de puissances, pas nécessairement identiques :

Lemme 6. *Notons :*

$S_i(q,a) := \sum_{n=1}^{q} e(an^{k_i}/q)$,
$D(q) := \sum_{a\wedge q=1} q^{-s} \prod_{i=1}^{s} S_i(q,a) e\left(-\frac{au}{q}\right)$,
$M(N,q)$ *le nombre de solutions à l'équation* $\sum_{i=1}^{s} x_i^{k_i} \equiv N \pmod q$ *dans* $\mathbb{Z}/q\mathbb{Z}$.

Alors on a $\forall b \geqslant 1$

$$\sum_{q|b} D(q) = b^{-s} M(N,b).$$

Idée de démonstration. On écrit, en utilisant le fait que $\sum_{r=1}^{q} e(rh/q) = q\mathbb{1}_{h=0}$:

$$M(u,q) = \frac{1}{q} \sum_{r=1}^{q} \sum_{x_1=1}^{q} \cdots \sum_{x_s=1}^{q} e(r(x_1^{k_1} + \cdots + x_s^{k_s} - u)/q)$$

et le résultat suit en séparant la somme selon $r \wedge q$ parmi les diviseurs de q. □

Ainsi en particulier, comme $\sum_\ell D(p^\ell) = \lim_{\ell\infty} \sum_{q|p^\ell} D(q)$, on peut écrire, sous réserve de convergence, par l'écriture en produit eulérien précédente :

$$\mathfrak{S}(N) = \prod_p \lim_{\ell\infty} \sum_{q|p^\ell} D(q) = \prod_p \lim_{\ell\infty} p^{-\ell s} M(N,p^\ell)$$

et il faut donc bien s'assurer au moins que $M(N,p^\ell) \gg p^{\ell s}$.

On utilise, presque toujours, le lemme de combinatoire additive suivant, dû à Cauchy, Davenport et Chowla.

Lemme 7 (Cauchy (1813), Davenport et Chowla (1935))**.** *Soit A et B des parties de $\mathbb{Z}/q\mathbb{Z}$ de cardinaux respectifs r et s. On suppose que $0 \in B$, et que pour tout $b \in B$ non nul, b est inversible (soit $b \wedge q = 1$). Alors :*

$$\#(A+B) \geqslant \min(q, r+s-1).$$

Démonstration. La preuve se fait par récurrence sur s. L'initialisation est triviale (car si $s=1$, $B=\{0\}$ et $A+B=A$). L'idée de l'itération est que, si $r<q$ (le cas $r=q$ étant trivial) pour chaque $b \neq 0$, il existe $c \in A$ tel que $c+b \notin A$. Sinon, on aurait que $(a+b)_{a\in A}$ décrit A, et donc que

$$\sum_{a\in A} (a+b) = \sum_{a\in A} a$$

et par suite $rb=0$, absurde par inversibilité de b et non nullité de r. On peut donc fixer un $c \in A$ obtenu par ce moyen, artificiellement retirer de B les b tels que $c+b \notin A$, et ajouter à A les $c+b$ en question. Ayant diminué le cardinal de B, on conclut par hypothèse de récurrence. □

3 Une collection de résultats démontrés par cette méthode

Théorème 8 (Liu et Zhao (2021)). *Pour tout $N \in \mathbb{N}$ suffisamment grand, il existe $x_1, \dots, x_{13} \in \mathbb{N}$ tel que :*

$$\sum_{i=1}^{13} x_i^{i+1} = N.$$

Roth [5] avait obtenu 50, Thanigasalam [2] 35, Ford d'abord 15 dans [4], puis 14 l'année suivante. La conjecture est que 4 suffit. Roth [5] a d'ailleurs démontré que les entiers s'écrivant sous la forme $x_1^2+x_2^3+x_3^4+x_4^5$ sont de densité 1 dans $\mathbb{N}$.

Théorème 9 (Problème de Waring, Hua (1938)). *Pour $k \in \mathbb{N}$, $s = 2^k + 1$, et tout $N \in \mathbb{N}$ suffisamment grand, il existe $x_1, \dots, x_s \in \mathbb{N}$ tels que :*

$$\sum_{i=1}^{s} x_i^k = N.$$

De nombreuses améliorations existent pour ce résultat, dues notamment à Davenport, Vinogradov, et Vaughan. On ne connaît cependant la valeur minimale de s que pour $k = 2$ et 4 (respectivement $s_{\min}(2) = 4$ et $s_{\min}(4) = 16$). Pour $k = 3$, on sait que $4 \leqslant s_{\min}(3) \leqslant 7$ et on conjecture que $s_{\min}(3) = 4$. Si on s'intéresse au même problème, mais cette fois pour tout $N \in \mathbb{N}$, c'est à dire sans l'aspect asymptotique, on conjecture alors que

$$s_{\min}(k) = 2^k + \lfloor (3/2)^k \rfloor - 2.$$

Théorème 10 (Freiman-Scourfield (1960)). *Soit $(k_n)_n$ une suite d'entiers $\geqslant 2$. Alors on a :*

$$\forall n \in \mathbb{N}, \exists s = s(n) \in \mathbb{N} \text{ tel que } \forall N \text{ assez grand, } N = x_0^{k_n} + \cdots + x_s^{k_{n+s}}$$

si et seulement si $\sum \frac{1}{k_n}$ diverge.

Ce théorème, conjecturé et incomplètement prouvé par Freiman en 1949, est une généralisation du problème de Waring, mais sa preuve, détaillée dans Scourfield [6], ne donne malheureusement aucune information sur $s(n)$ en fonction de la suite $(k_n)_n$. En revanche, plus récemment, Brudern et Wooley [7], ont obtenus une version effective de ce théorème, en obtenant la condition suffisante suivante :

$$\forall m, \sum_{n \geqslant m} \frac{1}{k_n} \geqslant 2 \log k_m + 4,71$$

Théorème 11 (Conjecture de Goldbach ternaire, Helfgott (2013)). *Pour tout entier impair $N \geqslant 7$, il existe p_1, p_2, p_3 premiers tel que :*

$$p_1 + p_2 + p_3 = N.$$

La véritable conjecture de Goldbach est que tout entier pair $\geqslant 4$ s'écrit comme somme de 2 nombres premiers, ce qui implique la version ternaire en ajoutant 3. Il est à noter que le résultat ici n'est pas asymptotique. Pour arriver à ce genre de conclusion avec la méthode du cercle, il faut pousser la précision de la borne inférieure obtenue plus loin, afin de quantifier le « *suffisamment grand* » et surtout obtenir une borne assez petite pour pouvoir vérifier numériquement toutes les valeurs sous cette borne.

4 Un exemple explicite

On va montrer dans cette section en plus grands détails le résultat original suivant :

Théorème 12 (P. (2023))**.** *Pour tout $N \in \mathbb{N}$ suffisamment grand, il existe $x_1, x_2, y_1, y_2, y_3, z_1, z_2, z_3, z_4 \in \mathbb{N}$ tel que :*

$$x_1^2 + x_2^2 + y_1^3 + y_2^3 + y_3^3 + z_1^4 + z_2^4 + z_3^4 + z_4^4 = N$$

i.e. tout entier suffisamment grand est somme de 2 carrés, 3 cubes, et 4 puissances quatrièmes.

4.1 Définitions

On fixe $\tau = 1/22$ et on pose les arcs majeurs :

$$\mathfrak{M}(q,a) := \left\{\alpha, \left|\alpha - \frac{a}{q}\right| \leqslant N^{\tau-1}\right\} \text{ pour } 1 \leqslant a \leqslant q \leqslant N^\tau, a \wedge q = 1.$$

Puis on définit la fonction F :

$$\begin{aligned} f_k(\alpha) &:= \sum_{x=x_k}^{N^{1/k}} e(\alpha x^k) \text{ pour } k = 2,3,4, \\ \phi_k(\alpha) &:= f_k(\alpha)^k \text{ pour } k = 2,3,4, \\ F(\alpha) &:= \prod_{k=2}^{4} \phi_k(\alpha). \end{aligned}$$

Puis on définit son approximation, G :

$$\begin{aligned} S_k(q,a) &:= \sum_{x=1}^{q} e\left(\frac{ax^k}{q}\right) \text{ pour } k = 2,3,4 \text{ et } 1 \leqslant a \leqslant q \leqslant N^\tau, a \wedge q = 1, \\ v_k(\beta) &:= \sum_{x=x_k}^{N} \frac{x^{1/k-1}}{k} e(\beta x) \text{ pour } k = 2,3,4, \\ g_k(\alpha; q, a) &:= q^{-1} S_k(q,a) v_k\left(\alpha - \frac{a}{q}\right) \text{ pour } k = 2,3,4 \text{ et } \alpha \in \mathfrak{M}(q,a), \\ \psi_k(\alpha; q, a) &:= g_k(\alpha; q, a)^k \text{ pour } k = 2,3,4 \text{ et } \alpha \in \mathfrak{M}(q,a), \\ G(\alpha; q, a) &:= \prod_{k=2}^{4} \psi_k(\alpha) \text{ pour } \alpha \in \mathfrak{M}(q,a). \end{aligned}$$

Dans les définitions des f_k et des v_k, on a pris :

$$x_2 = 1,\ x_3 = \frac{N^{1/3}}{2},\ x_4 = \frac{N^{1/4}}{2}.$$

Tout ceci est l'incarnation dans le cadre du problème traité de ce qui a été présenté précédemment dans un cadre général.

Remarque : pour f_k quand $k = 3$ ou 4, la somme ne commence pas à 1 mais à $N^{1/k}/2$, et ce pour les besoins du lemme 4.

Ainsi le coefficient de Fourier $r(N)$ de F ne sera pas précisément égal à $R(N)$, le nombre de manière d'écrire N sous la forme voulue, mais on a tout de même $R(N) \geqslant r(N)$. Montrer que $r(N)$ tend vers l'infini, ce que nous allons faire, conclura donc quand même.

4.2 Les arcs mineurs

Comme annoncé dans l'exposé général de la méthode, comme l'estimation triviale sur F est ici $F(\alpha) \ll N^3$, l'intégrale sur les arcs majeurs sera d'ordre au moins $N^{2-\varepsilon}$ (ce qu'on montrera en plus grands détails dans la section suivante). On cherchera donc ici à gagner $N^{1+\delta}$, pour un certain $\delta > 0$, sur la majoration triviale.

On va découper $\int_{\mathfrak{m}} |F|$ comme suit, par inégalité triangulaire puis Cauchy-Schwarz :

$$\int_{\mathfrak{m}} |F(\alpha)| d\alpha \leqslant \sup_{\alpha \in \mathfrak{m}} |\phi_2(\alpha)| \left(\int_0^1 |\phi_3(\alpha)|^2 d\alpha \right)^{1/2} \left(\int_0^1 |\phi_4(\alpha)|^2 d\alpha \right)^{1/2}.$$

Lemme 13. *On a la majoration suivante :*

$$\sup_{\alpha \in \mathfrak{m}} |\phi_2(\alpha)| \ll N^{1-\tau+\varepsilon}.$$

Démonstration. Pour $\alpha \in \mathfrak{m}$, on peut par le théorème de Dirichlet trouver $1 \leqslant a \leqslant q \leqslant N^{1-\tau}$ tel que $|\alpha - a/q| \leqslant N^{\tau-1}$. Nécessairement $q > N^\tau$, sans quoi α serait dans un arc majeur.

Dès lors on peut appliquer le lemme de Weyl, pour obtenir :

$$\phi_2(\alpha) \ll \left((N^{1/2})^{1+\varepsilon} \left(\frac{1}{q} + \frac{1}{N^{1/2}} + \frac{q}{(N^{1/2})^2} \right)^{1/2} \right)^2 = N^{1+\varepsilon} \left(\frac{1}{q} + \frac{1}{N^{1/2}} + \frac{q}{N} \right)$$

Comme $N^{1-\tau} \geqslant q > N^\tau$, et $\tau \leqslant 1/2$, il suit :

$$\phi_2(\alpha) \ll N^{1-\tau+\varepsilon}$$

et puisque la constante implicite est indépendante de α comme précisé dans la preuve du lemme de Weyl, le résultat suit. □

Pour les deux autres morceaux, on va utiliser les lemmes 4 et 5 et la méthode itérative qui y est associée.

Lemme 14. *On a, pour $k = 3, 4$, la majoration suivante :*

$$\int_0^1 |\phi_k(\alpha)|^2 d\alpha \ll N^{1+\nu+\varepsilon}$$

pour $\nu > 0$ fixé, arbitrairement petit quitte à prendre N assez grand.

Démonstration. On notera $S_k(m)$ le nombre de solutions à $x_1^k + \cdots + x_m^k = y_1^k + \cdots + y_m^k$ pour $N^{1/k}/2 \leqslant x_i, y_i \leqslant N^{1/k}$, de sorte que

$$S_k(k) = \int_0^1 |\phi_k(\alpha)|^2 d\alpha \text{ pour } k = 3, 4.$$

On s'occupe d'abord par exemple des puissances quatrièmes.
Par le lemme 5, on a $S_4(2) \ll N^{1/2+\varepsilon}$.

Une première application du lemme 4, avec $P = N^{1/4}$, $S = S_4(2)$, $j = 1$, ν arbitrairement petit quitte à prendre N assez grand, donne puisqu'ici $T = S_4(3)$ et $\Sigma \ll N^{1/2}$:

$$S_4(3) \ll N^{1/4}N^{1/2+\varepsilon} + (N^{1/4})^{1+\varepsilon+\nu/2-1}(N^{1/2+\varepsilon})^{1/2}N^{1/2} = N^{3/4+\varepsilon} + N^{3/4+\varepsilon/4+\nu/8} \ll N^{3/4+\nu+\varepsilon}.$$

Puis une deuxième application, mais avec cette fois $j = 2$, donne :

$$\begin{aligned} S_4(4) &\ll N^{1/4}N^{3/4+\nu+\varepsilon} + (N^{1/4})^{1+\varepsilon-3\nu/4-3/4}(N^{3/4+\nu+\varepsilon})^{3/4}(N^{3/4})^{1/2} \\ &= N^{1+\nu+\varepsilon} + N^{1/16+3\nu/16+\varepsilon/4}N^{9/16+3\nu/4+3\varepsilon/4}N^{6/16} \ll N^{1+\nu+\varepsilon}. \end{aligned}$$

Notons que dans les deux cas ci-dessus, on n'a fait apparaître dans l'inégalité que deux des trois termes donnés par le lemme, celui du milieu étant à chaque fois négligeable devant les deux autres, par souci de concision.

Si on s'occupe à présent des puissances troisièmes, on obtient avec le lemme 5 et une seule application du lemme 4 que :

$$S_3(3) \ll N^{1+\nu+\varepsilon}.$$

□

Ainsi, on a par les lemmes 13 et 14 :

$$\int_{\mathfrak{m}} |F(\alpha)| d\alpha \ll N^{1-\tau+\varepsilon}(N^{1+\nu+\varepsilon})^{1/2}(N^{1+\nu+\varepsilon})^{1/2} = N^{2-(\tau-\nu)+\varepsilon}. \tag{4}$$

Quitte à fixer $\nu = \tau/2$ assez petit, on a donc gagné $N^{1+\tau/2} > N$ sur l'estimée N^3.

4.3 Les arcs majeurs

La factorisation de $\int_{\mathfrak{M}} G(\alpha)e(-\alpha N)d\alpha$ est la suivante :

$$\begin{aligned} \int_{\mathfrak{M}} G(\alpha)e(-\alpha N)d\alpha &= \left(\sum_{q \leqslant N^\tau} \sum_{a \wedge q = 1} \prod_{k=2}^{4} q^{-k} S_k(q,a)^k e\left(\frac{-aN}{q}\right)\right)\left(\int_{-N^{\tau-1}}^{N^{\tau-1}} \prod_{k=2}^{4} v_k(\beta)^k e(-\beta N) d\beta\right) \\ &= \mathfrak{S}^*(N)J^*(N). \end{aligned}$$

4.3.1 Estimation de l'intégrale singulière

On s'occupe d'abord de l'intégrale singulière $J^*(N)$. Posons donc

$$J(N) = \int_{-1/2}^{1/2} \prod_{k=2}^{4} v_k(\beta)^k e(-\beta N) d\beta.$$

L'écart entre J et J^* sera contrôlé par le lemme suivant, qui est le lemme 2.8 de Vaughan dans [1] :

Lemme 15. *Pour $|\beta| \leqslant 1/2$, $k \geqslant 2$ on a :*

$$v_k(\beta) \ll \min(N^{1/k}, |\beta|^{-1/k})$$

où la constante implicite est uniforme en β.

Il se démontre par sommation d'Abel, puisqu'on connaît les sommes partielles géométriques d'exponentielles.

On obtient alors que $v_k(\beta)^k \ll \min(N, |\beta|^{-1})$, puis que $\prod_{k=2}^4 v_k(\beta)^k \ll \min(N^3, |\beta|^{-3})$. Par suite :

$$|J(N) - J^*(N)| \leqslant 2\int_{N^{\tau-1}}^{1/2} \left|\prod_{k=2}^4 v_k(\beta)^k\right| d\beta \ll 2\int_{N^{\tau-1}}^{1/2} \min(N^3, |\beta|^{-3}) d\beta = \int_{N^{\tau-1}}^{1/2} 2\beta^{-3} d\beta$$

la dernière égalité étant vraie car sur le domaine d'intégration $\beta \geqslant N^{\tau-1} \geqslant N^{-1}$. En évaluant la dernière intégrale on obtient :

$$|J(N) - J^*(N)| \ll N^{2-2\tau} - 4 \ll N^{2-2\tau}. \tag{5}$$

On a donc gagné $N^{1+2\tau} > N$ sur l'estimée triviale.

Lemme 16. *On a $J(N) \gg N^2$.*

Démonstration. En développant la définition de v_k dans $J(N)$, on obtient :

$$J(N) = \sum_{x_1,\dots,x_9 \in P} \frac{(x_1x_2)^{-1/2}(x_3x_4x_5)^{-2/3}(x_6x_7x_8x_9)^{-3/4}}{2^2 3^3 4^4} \tag{6}$$

où on a noté P l'ensemble de 9-uplets $(x_1, \dots, x_9)$ qui vérifient

$$\begin{aligned} 1 \leqslant x_1, x_2 &\leqslant N, \\ N/8 \leqslant x_3, x_4, x_5 &\leqslant N, \\ N/16 \leqslant x_6, x_7, x_8, x_9 &\leqslant N, \\ x_1 + \cdots + x_9 &= N. \end{aligned}$$

Notons que le terme général de la somme (6), puisque pour tout i, $x_i \leqslant N$, se minore par :

$$\frac{(N^2)^{-1/2}(N^3)^{-2/3}(N^4)^{-3/4}}{2^2 3^3 4^4} = \frac{N^{-6}}{2^2 3^3 4^4}.$$

Et ainsi on a

$$J(N) \gg \#P \cdot N^{-6}. \tag{7}$$

Fixons $\alpha, \beta, \gamma \geqslant 1$ trois réels et posons s l'application injective suivante :

$$\begin{aligned} [\![1; N/\alpha]\!] \times [\![N/8; N/\beta]\!]^3 \times [\![N/16; N/\gamma]\!]^4 &\longrightarrow P \\ (x_i)_{1\leqslant i\leqslant 8} &\longmapsto \left(\left(N - \sum_{i=1}^8 x_i\right), x_1, \dots, x_8\right). \end{aligned}$$

s sera bien définie si pour tous choix de $(x_i)_{2\leqslant i\leqslant 9}$ dans l'ensemble de départ, $N - \sum_{i=2}^{9} x_i \geqslant 1$, *i.e.* $\sum_{i=2}^{9} x_i < N$. Ainsi il faut que $N/\alpha + 3N/\beta + 4N/\gamma < N$. L'ensemble de départ de s sera en outre non vide ssi $\alpha \leqslant N, \beta \leqslant 8, \gamma \leqslant 16$.

Or, constatons que :

$$\frac{N}{12} + \frac{3N}{6} + \frac{4N}{12} = \frac{11N}{12} < N$$

et que $12 \leqslant N$ (car N est arbitrairement grand), $6 \leqslant 8$, $12 \leqslant 16$.
Ainsi pour $(\alpha, \beta, \gamma) = (12, 6, 12)$, s est une injection bien définie, et son ensemble de départ est non vide. De fait

$$\#P \geqslant \#s^{-1}(P) = \left(\frac{N}{12}\right)\left(\frac{N}{6} - \frac{N}{8} + 1\right)^3 \left(\frac{N}{12} - \frac{N}{16} + 1\right)^4 \gg N^8. \tag{8}$$

Enfin, les deux inégalités (7) et (8) concluent la preuve. □

4.3.2 Estimation de la série singulière

Lemme 17 (Roth)**.** *Il existe $c > 0$ tel que*

$$\mathfrak{S}^*(N) \gg (\log\log N)^{-c}$$

et donc pour tout $\varepsilon > 0$,

$$\mathfrak{S}^*(N) \gg N^{-\varepsilon}.$$

Démonstration. Pour ce qui est de l'évaluation de $\mathfrak{S}^*(N)$, le développement est en fait presque identique à ce qui est présenté dans Roth [5] dans la section « *The Singular Series* » (lemmes 15 à 30). Dans cette section, il démontre le développement en produit eulérien annoncé, utilise à profit le lemme 6 et des résultats analogues au lemme 7, et introduit un argument supplémentaire qui s'appuie sur une réécriture de $D(q)$ en termes de caractères multiplicatifs.
Bien que son problème concerne un carré, un cube, une puissance quatrième et une puissance cinquième, on peut constater que dans ses arguments, rien de spécifique aux puissances cinquièmes n'est utilisé. Ainsi en substituant à $\{u \in \mathbb{Z}/q\mathbb{Z}, \exists x, u \equiv x^5\}$ l'ensemble

$$\{u \in \mathbb{Z}/q\mathbb{Z}, \exists x_1, y_1, y_2, z_1, z_2, z_3, u \equiv x_1^2 + y_1^3 + y_2^3 + z_1^4 + z_2^4 + z_3^4\}$$

aucune de ses conclusions n'est qualitativement altérée.

On peut donc utiliser son résultat final ([5] lemme 30), qui donne exactement

$$\mathfrak{S}^*(N) \gg (\log\log N)^{-c}.$$

Et puisque $\forall \varepsilon > 0,\ \log\log N \ll N^{\varepsilon}$, le résultat suit. □

4.3.3 Qualité de l'approximation de F par G

Lemme 18. *L'erreur d'approximation sur les arcs majeurs vérifie*

$$\int_{\mathfrak{M}} |F - G| \ll N^{7/4+5\tau}.$$

Démonstration. Le lemme 2.7 de Vaughan [1] donne exactement, pour $k = 2, 3, 4$, et $\alpha \in \mathfrak{M}(q, a)$:

$$f_k(\alpha) - g_k(\alpha; q, a) \ll N^{2\tau} \tag{9}$$

avec constante implicite ne dépendant encore pas de α mais seulement de $\alpha - a/q$.

De là notons que :

$$\phi_k(\alpha) - \psi_k(\alpha; q, a) = (f_k(\alpha) - g_k(\alpha; q, a)) \sum_{i=0}^{k-1} f_k(\alpha)^i g_k(\alpha; q, a)^{k-1-i}. \tag{10}$$

Comme $f_k(\alpha) \ll N^{1/k}$ trivialement, et que $2\tau < 1/k$ pour $k = 2, 3, 4$, on a par (9) :

$$g_k(\alpha; q, a) = f_k(\alpha) + (g_k(\alpha; q, a) - f_k(\alpha)) \ll N^{1/k} + N^{2\tau} \ll N^{1/k}.$$

Et donc pour tout $0 \leqslant i \leqslant k-1$, $f_k(\alpha)^i g_k(\alpha; q, a)^{k-1-i} \ll N^{(k-1)/k}$. D'où par (9) et (10) :

$$\phi_k(\alpha) - \psi_k(\alpha; q, a) \ll N^{2\tau} \cdot (k-1) N^{1-1/k} \ll N^{1-1/k+2\tau}. \tag{11}$$

On écrit alors :

$$\begin{aligned} F(\alpha) - G(\alpha; q, a) =& \phi_2(\alpha)\phi_3(\alpha)(\phi_4(\alpha) - \psi_4(\alpha; q, a)) \\ &+ \phi_2(\alpha)(\phi_3(\alpha) - \psi_3(\alpha; q, a))\psi_4(\alpha; q, a) \\ &+ (\phi_2(\alpha) - \psi_2(\alpha; q, a))\psi_3(\alpha; q, a)\psi_4(\alpha; q, a). \end{aligned}$$

Comme $\phi_k(\alpha) \ll N$ et $\psi_k(\alpha; q, a) \ll N$ (car $f_k \ll N^{1/k}$ et $g_k \ll N^{1/k}$), on en déduit, avec (11), que

$$F(\alpha) - G(\alpha; q, a) \ll N^{3-1/4+2\tau} + N^{3-1/3+2\tau} + N^{3-1/2+2\tau} \ll N^{3-1/4+2\tau}.$$

Ainsi, puisque chaque arc majeur est de largeur $2N^{\tau-1}$, et qu'il y en a au plus $N^{2\tau}$, il suit

$$\int_{\mathfrak{M}} |F - G| = \sum_{q,a} \int_{\mathfrak{M}(q,a)} |F(\alpha) - G(\alpha; q, a)| d\alpha \ll N^{2\tau} \cdot N^{\tau-1} \cdot N^{3-1/4+2\tau} = N^{7/4+5\tau},$$

ce qui est le résultat. □

Remarque : Pour que ce résultat soit exploitable, on veut absolument que l'exposant soit < 2, et c'est de là que vient la contrainte sur τ, comme annoncé dans l'exposé général de la méthode. Ici, on aura besoin de $\frac{7}{4} + 5\tau < 2$, donc il faut absolument

$$\tau < \frac{1}{20}.$$

La valeur précise $\tau = \frac{1}{22}$ n'a pas d'importance, mais simplifie légèrement les calculs dans ce qui va suivre. On pourrait ici être tenté de faire tendre τ vers 0 et de faire disparaître les arcs majeurs, ce qui rendrait futile la méthode du cercle.
Cependant, par (4), il est nécessaire de prendre $\tau > \nu$. Bien que ν soit arbitraire quitte à prendre N assez grand, les constantes implicites des inégalités asymptotiques dépendent *a priori* de ν. Ainsi, si ν est variable, N doit varier avec ν, et les constantes implicites dépendront de N, ce qui fait perdre tout leur sens aux estimations. Donc ν doit être fixe, et τ ne peut pas tendre vers 0.

4.4 Conclusion

On a donc, par (4), (5) et le lemme 18,

$$\begin{aligned}\int_0^1 F(\alpha)e(-\alpha N)d\alpha &= \int_{\mathfrak{M}} G(\alpha)e(-\alpha N)d\alpha + \int_{\mathfrak{M}} (F(\alpha)-G(\alpha))e(-\alpha N)d\alpha + \int_{\mathfrak{m}} F(\alpha)e(-\alpha N)d\alpha \\ &= \mathfrak{S}^*(N)J^*(N) + O(N^{7/4+5\tau}) + O(N^{2-\tau/2+\varepsilon}) \\ &= \mathfrak{S}^*(N)J(N) + O(\mathfrak{S}^*(N)N^{2-2\tau}) + O(N^{7/4+5\tau}) + O(N^{2-\tau/2+\varepsilon}).\end{aligned}$$

Comme $\tau = 1/22$, il suit $7/4+5\tau = 2-\tau/2 = 2-1/44$. Ainsi les deux derniers termes d'erreur sont $\ll N^{2-1/44}$.

De plus, par le lemme 16, $\mathfrak{S}^*(N)N^{2-2\tau} \ll \mathfrak{S}^*(N)J(N)N^{-2\tau} = o(\mathfrak{S}^*(N)J(N))$. Ainsi on a

$$\int_0^1 F(\alpha)e(-\alpha N)d\alpha = \mathfrak{S}^*(N)J(N) + o(\mathfrak{S}^*(N)J(N)) + O(N^{2-1/44}).$$

Par les lemmes 16 et 17, $\mathfrak{S}^*(N)J(N) \gg N^{2-\varepsilon}$. Donc de même, $\mathfrak{S}^*(N)J(N)+o(\mathfrak{S}^*(N)J(N)) \gg N^{2-\varepsilon}$. Ainsi

$$\int_0^1 F(\alpha)e(-\alpha N)d\alpha \gg N^{2-\varepsilon} + O(N^{2-1/44}) \gg N^{2-\varepsilon}$$

pour tout $\varepsilon < 1/44$. De là on déduit le Théorème 12. □

Références

[1] R. C. Vaughan, *The Hardy-Littlewood Method (Second Edition)*, Cambridge University Press, 1997.

[2] K. Thanigasalam, *On additive number theory*, Acta Arithmeticae, 1968.

[3] H. Davenport, *On sums of positive integral k-th powers*, Proceedings of the Royal Society, 1939.

[4] K. B. Ford, *The representation of numbers as sums of unlike powers*, Journal of the American Mathematical Society, 1995.

[5] K. F. Roth, *A problem in additive number theory*, Proceedings of the London Mathematical Society, 1948.

[6] E. J. Scourfield, *A generalization of Waring's problem*, Journal of the London Mathematical Society, 1960.

[7] J. Brüdern and T. D. Wooley, *On Waring's problem : beyond Freiman's theorem*, arXiv :2302.12920 [math.NT], 2023.