ENS de Rennes
Université Rennes 1

Département de Mathématiques
Rapport de Stage

Corrélations entre fonctions multiplicatives

Élève :
Titouan Donnart

Encadrant :
Pierre-Yves Bienvenu

Introduction :

Une heuristique en théorie analytique des nombres stipule que les ensembles de nombres entiers positifs ne peuvent pas être simultanément structurés de manière additive et multiplicative. La vérification pratique de cette heuristique est à l'origine d'un grand nombre de problèmes difficiles. Les conjectures de ce type sont également au moins équivalentes sur le plan conceptuel à l'attente selon laquelle une fonction multiplicative quelconque se comporte de manière aléatoire sur des ensembles additifs.

Dans ce rapport, nous examinons l'un de ces problèmes qui est resté ouvert jusqu'en 2015 : *le problème de la discrépance d'Erdos*. Il s'agit de voir si oui ou non la borne supérieure prise (en module) par les sommes des éléments d'une suite $(f(n))_{n\in\mathbb{N}}$ à valeurs dans $\{-1,+1\}$ le long de progressions arithmétiques homogènes est infinie.

Après avoir rappelé divers résultats de base sur la théorie analytique des nombres, nous nous intéresserons au point de départ de la preuve du problème de la discrépance d'Erdos : les résultats obtenus par Matomäki et Radziwiłł en 2014 sur les moyennes de fonctions multiplicatives sur de petits intervalles. Ces résultats ont bouleversé la recherche en théorie analytique des nombres, et la direction de notre rapport emprunte seulement l'une des nombreuses conséquences importantes de ces résultats.

Cela nous amènera à discuter de la conjecture de Chowla et d'Elliott, ainsi que des versions dites "en moyenne" de ces conjectures, qui seront démontrées par les résultats obtenus par Matomäki et Radziwiłł. Enfin, nous nous intéresserons à la démonstration de Tao du problème de la discrépance d'Erdos.

Je tiens personnellement à remercier chaleureusement mon maître de stage, Pierre-Yves Bienvenu, de m'avoir fait découvrir ce domaine des mathématiques dont je n'avais presque aucune connaissance avant ce stage. Je tiens aussi à le remercier pour toute l'aide et toutes les connaissances qu'il m'a apportées pendant ces deux mois. Je tiens également à remercier l'ensemble du personnel du Trinity College de m'avoir permis de réaliser ce stage dans un bon cadre. Enfin, je remercie l'ENS de Rennes et le centre Henri Lebesgue de m'avoir octroyé une bourse pour que je puisse réaliser ce stage à Dublin malgré le coût très élevé du logement là-bas.

Je tiens enfin à signaler que ce présent n'est qu'un aperçu de mon travail réalisé pendant ces deux mois. Sur ma page web, on retrouvera mon rapport de stage complet comportant la plupart des démonstrations manquantes dans celui-ci.

Table des matières

Notations :

— Si A est un ensemble fini on notera $|A|$ son cardinal.

— Si A est une partie de $\mathbb{R}$ mesurable (pour la mesure de Lebesgue), on notera mes(A) sa mesure.

— Si A est un ensemble, on notera $\mathbb{1}_A$ la fonction indicatrice de A ($\mathbb{1}_A(x) = 1$ si $x \in A$ et est égal à 0 sinon). De plus si S est un énoncé mathématique on notera $\mathbb{1}_S$ la fonction qui vaut 1 quand S est vraie et 0 quand S est fausse. Ainsi on notera par exemple $\mathbb{1}_{2|n}$ l'indicatrice des nombres pairs.

— On notera log le logarithme népérien.

— Si $x \in \mathbb{R}$, on notera $\lfloor x \rfloor$ sa partie entière.

— Si $x \in \mathbb{R}_+$, on notera une somme $\sum_{n \leqslant x}$ la somme $\sum_{n=0}^{\lfloor x \rfloor}$. Une somme indicée sur p sera toujours une somme sur les nombres premiers.

— Pour tous $m, n \in \mathbb{Z}$, on notera (m, n) leur PGCD.

— On rappelle qu'une fonction $f : \mathbb{N} \to \mathbb{C}$ est dite multiplicative si pour tous m, n premiers entre eux $f(mn) = f(m)f(n)$, on dit qu'elle est complètement multiplicative si celle égalité est valable pour tous $m, n \in \mathbb{N}$.

— Si $f : \mathbb{N} \to \mathbb{C}$, on notera $\mathcal{D}f$ sa série de Dirichlet qui vérifie pour tous $s \in \mathbb{C}$ (lorsque cette quantité est définie) :

$$\mathcal{D}f(s) = \sum_{n=0}^{+\infty} \frac{f(n)}{n^s}.$$

— On notera φ l'indicatrice d'Euler définie par $\varphi(n) = |\{1 \leq k \leq n;\ k \text{ est premier avec } n\}|$.

— On note μ la fonction de Mobiüs définie par $\mu(1) = 1$,$\mu(n) = 0$ si n possède un facteur carré et $\mu(n) = (-1)^k$ si n est le produit de k nombres premiers distincts. Elle est multiplicative.

— On note λ la fonction de Liouville définie par $\lambda(n) = (-1)^{\Omega(n)}$ où $\Omega(n)$ désigne le nombre de facteurs premiers comptés avec multiplicité de l'entier n. C'est une fonction complètement multiplicative.

— On note π la fonction qui à $x \in \mathbb{R}$ compte le nombre de nombres premiers inférieurs ou égaux à x.

— On adoptera les notations asymptotiques dans deux contextes, un où il n'y a pas de paramètre asymptotique présent et un où il y en a. Dans le contexte sans paramètre asymptotique, on utilisera $X = \mathcal{O}(Y)$, $X \ll Y$ ou $Y \gg X$ pour noter une estimation de la forme $|X| \leqslant CY$ où C est une constante fixe.

— Dans certains cas où la constante C dépend de paramètres additionnels comme k on le notera en indice comme ceci $X = \mathcal{O}_k(Y)$, $X \ll_k Y$ ou $Y \gg_k X$.

— Quand il existe un paramètre asymptotique (par exemple x) tous les objets mathématiques pourront dépendre de ce paramètre (à part contre-indications). On utilisera $X = \mathcal{O}(Y)$, $X \ll Y$ ou $Y \gg X$ pour noter une estimation de la forme $|X| \leqslant CY$ où C est une constante qui peut dépendre d'autre paramètres tant que ceux-ci sont fixés. On notera aussi $X = o(Y)$ pour noter une estimation de la forme $|X| \leqslant cY$ où c est une quantité qui tend vers 0 quand le paramètre asymptotique tend vers $+\infty$.

Les autres notations apparaissant dans ce rapport seront précisées au fur et à mesure du document.

I Rappels et quelques éléments de théorie analytique des nombres

I.1 Quelques rappels

I.1.1 Formule de sommation par parties

L'objectif de cette partie est de présenter la formule de sommation par parties qui va être un résultat technique classique en théorie analytique des nombres.

Théorème I.1: Formule de sommation par parties

Soit $x \in \mathbb{R}_+, a \in \mathbb{N}, f : [a, x] \to \mathbb{C}$ et g une fonction de classe $\mathcal{C}^1([a, x])$. Alors :

$$\sum_{a \leqslant n \leqslant x} f(n)g(n) = \sum_{a \leqslant n \leqslant x} f(n)g(x) - \int_a^x g'(t) \left(\sum_{a \leqslant n \leqslant t} f(n) \right) \mathrm{d}t.$$

De plus, si $x \geqslant 2$, alors :

$$\sum_{p \leqslant x} f(p)g(p) = \sum_{p \leqslant x} f(p)g(x) - \int_2^x g'(t) \left(\sum_{p \leqslant t} f(p) \right) \mathrm{d}t.$$

Démonstration : Par le théorème fondamental de l'intégration, on peut écrire :

$$\sum_{a \leqslant n \leqslant x} f(n)g(n) = \sum_{a \leqslant n \leqslant x} f(n) \left(g(x) - \int_n^x g'(t)\,\mathrm{d}t \right) = \sum_{a \leqslant n \leqslant x} f(n) \left(g(x) - \int_{\mathbb{R}} g'(t) \mathbb{1}_{[n,x]}(t)\,\mathrm{d}t \right).$$

Donc on obtient :

$$\begin{aligned} \sum_{a \leqslant n \leqslant x} f(n)g(n) &= \sum_{a \leqslant n \leqslant x} f(n)g(x) - \int_{\mathbb{R}} g'(t) \sum_{a \leqslant n \leqslant x} f(n) \mathbb{1}_{[n,x]}(t)\,\mathrm{d}t \\ &= \sum_{a \leqslant n \leqslant x} f(n)g(x) - \int_{\mathbb{R}} g'(t) \mathbb{1}_{[a,x]}(t) \sum_{a \leqslant n \leqslant t} f(n)\,\mathrm{d}t \\ &= \sum_{a \leqslant n \leqslant x} f(n)g(x) - \int_a^x g'(t) \sum_{a \leqslant n \leqslant t} f(n)\,\mathrm{d}t \end{aligned}$$

Pour obtenir le deuxième point on applique le premier point à la fonction $\tilde{f}$ qui vaut $f(p)$ si p est premier et 0 sinon. □

I.1.2 Théorème d'approximation de Dirichlet

Dans deuxième courte sous-partie, on va démontrer le résultat suivant :

Théorème I.2: Théorème d'approximation de Dirichlet

Soient $\theta \in \mathbb{R}$ et Q un entier supérieur ou égal à 2. Alors il existe un rationnel $\frac{p}{q}$ tel que :

$$0 < q < Q \quad \text{et} \quad \left|\theta - \frac{p}{q}\right| \leqslant \frac{1}{qQ}.$$

Démonstration : La preuve est assez rapide et repose essentiellement sur "le principe des tiroirs". Considérons les $Q+1$ nombres $0, 1\{\theta\}, \{2\theta\}, \dots \{(Q-1)\theta\}$. Tous ces nombres appartiennent à $[0,1]$. Donc au moins un des Q intervalles $\left[\frac{k}{Q}, \frac{k+1}{Q}\right]$ pour $k \in \{0, \dots, Q-1\}$ contient deux de ces nombres. Ainsi il existe des entiers a_1, a_2, b_1, b_2 tels que :

$$0 \leqslant a_1 < a_2 \leqslant Q \quad \text{et} \ |(a_1\theta - b_1) - (a_2\theta - b_2)| \leqslant \frac{1}{Q}.$$

Donc en posant $q = a_2 - a_1 \in [1, Q-1]$ et $p = b_2 - b_1$, on obtient le résultat attendu. □

I.1.3 Caractères, transformée de Fourier sur un groupe fini et sommes Gaussiennes

Dans cette partie on va s'intéresser aux caractères de Dirichlet et à la transformée de Fourier sur un groupe fini. Nous n'allons pas démontrer les résultats ici car ce sont plutôt des résultats classiques de cours de L3/M1 sur les représentations de groupes. Pour les preuves on fera référence à [1] (sauf contre-indication).

Définition I.3

- *Soit G un groupe fini abélien. On appelle* ***caractère*** *sur le groupe G, tout morphisme $\chi : G \to \mathbb{C}$. On notera $\hat{G}$ l'ensemble des caractères de G.*
- *Soit $q > 0$. On appelle* ***caractère de Dirichlet modulo*** *q toute fonction $\chi : \mathbb{N} \to \mathbb{C}$ telle qu'il existe un caractère $\tilde{\chi} : (\mathbb{Z}/q\mathbb{Z})^* \to \mathbb{C}$ tel que :*

$$\forall n \in \mathbb{N} : \chi(n) = \mathbb{1}_{(n,q)=1}\tilde{\chi}(n \mod q).$$

Deux exemples de base de caractères de Dirichlet sont les suivants :

Définition I.4

- *Le caractère de Dirichlet valant 1 sur les entiers premiers avec n et 0 ailleurs est appelé* ***caractère principal*** *modulo n.*
- *Le caractère de Dirichlet principal modulo 1 (valant 1 sur tous les entiers) est dit* ***caractère trivial****.*

Citons pour commencer quelques propriétés de base sur les caractères :

Proposition I.5

Il y a exactement $\varphi(q)$ caractères de Dirichlet modulo q pour tout $q \geqslant 1$.

Intéressons-nous maintenant à la transformée de Fourier sur un groupe fini abélien G que l'on fixe pour le reste de cette partie. Sur l'ensemble $\mathbb{C}[G]$ des applications de G dans $\mathbb{C}$, on définit le produit scalaire hermitien par :

$$\forall f, g \in \mathbb{C}[G] = \frac{1}{|G|} \sum_{x \in G} f(x)\overline{g(x)}.$$

Définition I.6

— *Soient $f \in \mathbb{C}[G]$ et $\chi \in \hat{G}$. On définit le* ***coefficient de Fourier*** *$c_f(\chi)$ par : $c_f(\chi) = \langle f \mid \chi \rangle$.*

— *On appelle transformée de Fourier sur G, l'application $\mathcal{F}$ qui à $f \in \mathbb{C}[G]$ associe $\hat{f}$ définie par :*

$$\forall \chi \in \hat{G} : \hat{f}(\chi) = |G|\, c_f(\overline{\chi}) = \sum_{x \in G} f(x)\chi(x).$$

Par les relations d'orthogonalité entre les caractères on obtient les deux résultats suivants :

Proposition I.7

Soit $f \in \mathbb{C}[G]$, on a la formule d'inversion :

$$f = \sum_{\chi \in \hat{G}} c_f(\chi)\chi = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \hat{f}(\chi)\chi^{-1}.$$

Proposition I.8

Soient $f, g \in \mathbb{C}[G]$, on a la formule de Plancherel :

$$\sum_{s \in G} f(s)\overline{g(s)} = |G| \sum_{\chi \in \hat{G}} c_f(\chi)\overline{c_g(\chi)} = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \hat{f}(\chi)\overline{\hat{g}(\chi)}.$$

On termine cette partie en énonçant un résultat sur les sommes gaussiennes (voir Théorème 8.7, [2]).

Proposition I.9

Soit χ un caractère de Dirichlet modulo q. On note pour tout $n \geqslant 1$:

$$G(n, \chi) := \sum_{m=1}^{q} \chi(m) e\left(\frac{mn}{q}\right).$$

Alors on a pour tout $n \geqslant 1$:

$$G(n, \chi) = \overline{\chi(n)} G(1, \chi).$$

I.1.4 Entropie de Shannon

Dans cette partie, on va s'intéresser à un objet de la théorie de l'information appelé l'entropie de Shannon. Cette quantité prendra de l'importance dans notre rapport lorsque qu'on utilisera un argument dit de *décrément d'entropie*.

Définition I.10

Soient X, Y deux variables aléatoires discrètes.

— *On appelle* ***entropie*** *de X la quantité $H(X)$ définie par :*

$$H(X) = \sum_{x \text{ atome}} \mathbf{P}(X = x) \log\left(\frac{1}{\mathbf{P}(X = x)}\right).$$

— *On définit* ***l'entropie de X conditionnellement à Y*** *par :*

$$H(X|Y) = \sum_{y \text{ atome}} \mathbf{P}(Y = y) H(X|Y = y),$$

où :

$$H(X|Y = y) = \sum_{x \text{ atome}} \mathbf{P}(X = x|Y = y) \log\left(\frac{1}{\mathbf{P}(X = x|Y = y)}\right).$$

On notera aussi $H(X, Y)$ comme l'entropie de Shannon du couple (X, Y) de variables aléatoires discrètes.

En fait on appelle dans ce texte l'entropie ce qui est couramment appelé dans la littérature *entropie de Shannon* (en prenant un logarithme en base 2). Cette grandeur est surtout utilisée en théorie de l'information, c'est une mesure de l'incertitude dans une source de données. En effet, lorsque la variable aléatoire X est presque sûrement constante son entropie de Shannon est nulle, c'est-à-dire, si une source envoie toujours le même symbole alors son entropie est nulle (c'est-à-dire ici minimale car l'entropie de Shannon est positive).

Proposition I.11

Soient X, Y, Z trois variables aléatoires discrètes :

— $H(X, Y) = H(X|Y) + H(Y) = H(X) + H(Y|X)$,

— $H(X|Y) \leqslant H(X)$,

— $H(X, Y) \leqslant H(X) + H(Y)$,

— $H(X, Y|Z) \leqslant H(X|Z) + H(Y|Z)$.

Démonstration :

— Le premier point vient d'un calcul direct.

— On développe l'expression de $H(X|Y)$:

$$H(X|Y) = \sum_{x,y \text{ atomes}} \mathbf{P}(Y = y)\mathbf{P}(X = x|Y = y) \log\left(\frac{1}{\mathbf{P}(X = x|Y = y)}\right).$$

Donc d'après l'inégalité de Jensen appliquée à la fonction concave $x \in [0, 1] \mapsto x \log\left(\frac{1}{x}\right)$:

$$H(X|Y) \leqslant \sum_{x \text{ atome}} \left(\sum_{y \text{ atome}} \mathbf{P}(Y = y)\mathbf{P}(X = x|Y = y)\right) \log\left(\frac{1}{\sum_{y \text{ atome}} \mathbf{P}(X = x|Y = y)}\right).$$

On en déduit le deuxième point par la formule des probabilités totales.

— Le troisième point est une conséquence immédiate des deux autres.

— Le quatrième point se prouve de la même manière.

□

Définition I.12

Soient X, Y deux variables aléatoires discrètes, on définit leur ***information mutuelle*** *par :*

$$I(X,Y) = H(X) + H(Y) - H(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X).$$

Par la proposition précédente, on remarque que : $I(X,Y) = I(Y,X) \geqslant 0$. On peut voir $I(X,Y)$ comme la mesure de la non-indépendance de X et Y : plus elle est petite plus les variables aléatoires sont d'avantage indépendantes (en fait X et Y sont indépendantes si et seulement si $I(X,Y) = 0$) et plus elle est grande pour les variables aléatoires X et Y ne sont pas indépendantes. On termine cette introduction à l'entropie de Shannon par cette dernière proposition qui nous sera très utile en pratique.

Proposition I.13

Soit X une variable aléatoire discrète prenant au maximum N valeurs. Alors $H(X) \leqslant \log(N)$.

Démonstration : Notons A l'ensemble des valeurs prises par X. On a $|A| \leqslant N$ par hypothèse. D'après l'inégalité de Jensen appliqué à encore $x \in [0,1] \mapsto x\log\left(\frac{1}{x}\right)$ (en prenant les poids tous égaux à $\frac{1}{|A|}$) :

$$H(X) = \sum_{x\in A} \mathbf{P}(X=x)\log\left(\frac{1}{\mathbf{P}(X=x)}\right) \leqslant |A| \sum_{x\in A} \frac{1}{|A|}\mathbf{P}(X=x)\log\left(\frac{|A|}{\sum_{y\in A}\mathbf{P}(X=y)}\right) = \log(|A|).$$

□

I.2 Outils de base de la théorie analytique des nombres

I.2.1 Théorèmes fondamentaux de la théorie analytique des nombres

Le but de cette partie est de présenter sans preuve (on pourra, à l'exception du théorème des nombres premiers, les retrouver dans le rapport complet) quelques résultats fondamentaux de la théorie analytique des nombres. On commence par énoncer trois résultats dus à Mertens :

Théorème I.14: Premier théorème de Mertens

Pour tout $x \geqslant 1$:

$$\sum_{p\leqslant x} \frac{\log(p)}{p} = \log(x) + \mathcal{O}(1).$$

Théorème I.15: Deuxième théorème de Mertens

Il existe deux constantes $c_1, c_2 \in \mathbb{R}$ telles que pour tout $x \geqslant 2$:

$$\sum_{p \leqslant x} \frac{1}{p} = \log\log(x) + c_2 + \mathcal{O}\left(\frac{1}{\log(x)}\right).$$

Proposition I.16: Formule de Mertens

Il existe une constante $C \in \mathbb{R}$ telle que pour $x \to +\infty$:

$$\prod_{p \leqslant x} \frac{1}{1 - \frac{1}{p}} = C \log(x) + \mathcal{O}(1).$$

En fait on peut avoir un résultat plus fin et montrer que $C = \exp^{-\gamma}$ où γ est la constante d'Euler, mais ce n'est pas important pour la suite du rapport.

On présente maintenant le fameux théorème des nombres premiers ainsi que ses formes équivalentes qu'on utilisera dans la suite de ce rapport.

Théorème I.17: Théorème des nombres premiers

On a l'équivalent suivant :

$$\pi(x) \underset{x \to +\infty}{\sim} \frac{x}{\log(x)}.$$

De manière équivalente on a :

$$\lim_{x \to +\infty} \frac{1}{x} \sum_{n \leqslant x} \lambda(n) = \lim_{x \to +\infty} \frac{1}{x} \sum_{n \leqslant x} \mu(n) = 0.$$

On va maintenant présenter un dernier résultat classique de théorie analytique des nombres, le théorème d'Hardy-Ramanujan, il se base sur la propriété suivante dont on se servira dans la suite du rapport.

Proposition I.18: Inégalité de Turan-Kubilius

Soit $I \subset \mathbb{R}$ un intervalle de longueur au moins 1 et n un entier tiré au sort uniformément dans cet intervalle. Soit $\mathcal{P}$ une partie finie de $\mathbb{N}$ ne contenant que des nombres premiers de taille au plus $P \geqslant 2$. Alors la variable aléatoire $\omega = \sum_{p \in \mathcal{P}} \mathbb{1}_{p|n}$ vérifie :

$$\forall \lambda > 0 : \quad \mathbf{P}(|\omega - \ell(\mathcal{P})| \geqslant \lambda) \ll \frac{\ell(\mathcal{P}) + \frac{P^2}{\mathrm{mes}(I)}}{\lambda^2}.$$

Théorème I.19: Théorème d'Hardy-Ramanujan

Pour tout x suffisamment grand :

$$\frac{1}{x} \sum_{n \leqslant x} |\omega(n) - \log\log(x)|^2 \ll \log\log(x).$$

I.2.2 Quelques applications de la théorie des cribles

On présente pour finir deux résultats généraux que l'on peut obtenir grâce à la théorie des cribles (Voir rapport général pour plus de détails). Ils nous serviront plusieurs fois dans la suite dans nos estimations.

Proposition I.20

Soit $x \geqslant 1$ et soit I un intervalle de longueur x et soit $\mathcal{P}$ un ensemble de nombres premiers inférieurs ou égaux à x. Si l'on retire une classe de résidus modulo p de I pour tout $p \in \mathcal{P}$ alors il restera au maximum $\mathcal{O}\left(|I| \prod_{p\in\mathcal{P}} \left(1 - \frac{1}{p}\right)\right)$ éléments.

Proposition I.21

On note $\Xi(x, y, z)$ le nombre d'entiers $n \leqslant x$ qui n'ont pas de facteurs premiers dans $]y, z]$. On a :

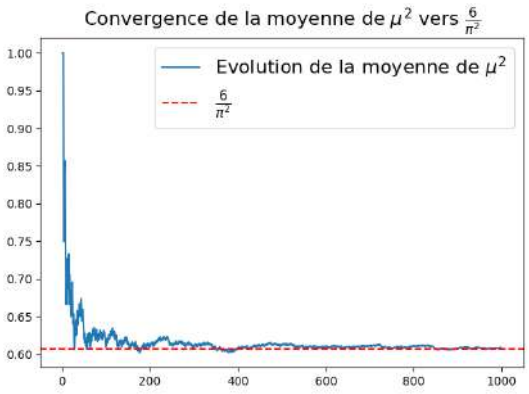
$$\Xi(x, y, z) \ll \frac{x \log(y)}{\log(z)}.$$

I.3 Moyennes de fonctions multiplicatives

I.3.1 Quelques exemples

Étudier directement le comportement d'une fonction arithmétique paraît difficile car le comportement d'une telle fonction est très souvent erratique.

Par contre si on trace ça moyenne sur $[1, x]$ par rapport à x, on obtient quelque chose de plus "classique" sous sa forme. Par exemple ci-dessous on a tracé l'évolution de la moyenne de μ^2 sur $[1, x]$ en fonction de $x \in [1, 1000]$. On conjecture que la moyenne de μ^2 sur $[1, x]$ converge lorsque $x \to +\infty$ vers $\frac{6}{\pi^2}$ (et on peut le montrer).



Cela nous amène à poser la définition suivante :

Définition I.22

On dit qu'une fonction arithmétique f ***admet une valeur moyenne*** *si la limite suivante existe.*

$$\lim_{x\to+\infty} \frac{1}{x} \sum_{n\leqslant x} f(n).$$

Dans ce cas, on appelle ***valeur moyenne d'une fonction arithmétique*** *f la valeur de cette limite.*

Se ramener à l'étude de la valeur moyenne d'une fonction arithmétique est un raisonnement qui se rejoint beaucoup à la théorie des probabilités, on construit des objets qui sont plus simples à étudier et

qui permettent de connaître quelques informations sur la loi étudiée (comme on le fait avec les moments ou la fonction caractéristique d'une variable aléatoire).

En fait, on n'étudie que les valeurs moyennes de fonctions multiplicatives car une fonction arithmétique est trop malléable pour obtenir un résultat général. En effet, on peut construire une fonction arithmétique non multiplicative qui n'admet pas de valeur moyenne. Pour cela on pose :

$$f(n) = \begin{cases} 1 & \text{s'il existe } k \in \mathbb{N} \text{ tel que } 2^{2k} < x \leqslant 2^{2k+1} \\ 0 & \text{s'il existe } k \in \mathbb{N} \text{ tel que } 2^{2k+1} < x \leqslant 2^{2k+2} \end{cases}.$$

On montre facilement que :

$$\liminf_{x\to+\infty} \frac{1}{x}\sum_{n\leqslant x} f(n) = \frac{1}{3} \quad \text{et} \quad \limsup_{x\to+\infty} \frac{1}{x}\sum_{n\leqslant x} f(n) = \frac{2}{3}.$$

Pourtant l'étude de la moyenne d'une fonction multiplicative est loin d'être simple, par exemple connaître le comportement de la valeur moyenne de manière suffisamment précise, on devra faire appel au théorème des nombres premiers (par exemple pour la fonction μ ou la fonction τ). On peut aussi calculer la valeur moyenne de certaines fonctions multiplicatives assez simples par des techniques encore plus élémentaires. Par exemple on peut montrer de manière élémentaire que (voir rapport complet pour quelques détails) :

$$\frac{1}{x}\sum_{n\leqslant x} \tau^2(n) \ll \log^3(x). \tag{I.1}$$

Il arrive même que certaines fonctions multiplicatives n'aient même pas de moyenne, par exemple $n \mapsto n^{it}$, on a :

$$x \mapsto \frac{1}{x}\sum_{n\leqslant x} n^{it} \text{ n'a pas de limite quand } x \to +\infty. \tag{I.2}$$

En effet d'après la formule de sommation par parties :

$$\sum_{n=1}^{x} n^{it} = \lfloor x \rfloor\, x^{it} - it\int_1^x \lfloor u \rfloor\, u^{it-1}\,\mathrm{d}u = x^{it+1}\left(1 + \frac{it}{it+1}\right) + \mathcal{O}(1).$$

Ainsi :

$$\frac{1}{x}\sum_{n\leqslant x} n^{it} \underset{x\to+\infty}{\sim} x^{it}\left(1 + \frac{it}{it+1}\right).$$

On en déduit le fait annoncé. En fait, on remarque qu'en module l'équivalent calculé est constant mais qu'il varie en argument en fonction de x, c'est parce que la fonction $n \mapsto n^{it}$ oscille trop. Cette fonction joue un rôle important dans le cadre général, une fonction multiplicative qui se situera assez loin de $n \mapsto n^{it}$ (dans un sens qu'on définira), alors on pourra contrôler la valeur moyenne et certaines autocorrélations de cette fonction multiplicative.

I.3.2 Distance prétentieuse

On va introduire une notion de distance entre deux fonctions multiplicatives. Nous allons nous cantonner aux fonctions multiplicatives dite *1-bornées* c'est à dire des fonctions $f : \mathbb{N} \to \mathbb{C}$ vérifiant pour tout $n \in \mathbb{N} : |f(n)| \leqslant 1$. Avant cela, on va avoir besoin du lemme suivant :

Lemme I.23

Soient u, v, w trois vecteurs d'un espace préhilbertien $(H, \langle . \mid . \rangle, \|.\|)$ de norme inférieure ou égale à 1. Alors, on a l'inégalité suivante :

$$\sqrt{1 - \langle u \mid w \rangle} \leqslant \sqrt{1 - \langle u \mid v \rangle} + \sqrt{1 - \langle v \mid w \rangle}.$$

Démonstration :

— On va d'abord démontrer ce lemme dans le cas où u, v, w sont de norme égale à 1. On a alors :

$$\sqrt{1 - \langle u \mid v \rangle} = \frac{1}{\sqrt{2}} \|u - v\|.$$

On a évidemment la même relation en remplaçant u ou v par w. On obtient donc l'inégalité demandée grâce à l'inégalité triangulaire pour $\|.\|$.

— Revenons au cas général. On va travailler dans l'espace préhilbertien $(\tilde{H}, \langle . \mid . \rangle_{\tilde{H}}, \|.\|_{\tilde{H}})$ où $\tilde{H} = H \times \mathbb{R}^3$. On pose :

$$\begin{aligned} \tilde{u} &= (u, \sqrt{1 - \|u\|}, 0, 0) \\ \tilde{v} &= (u, 0, \sqrt{1 - \|v\|}, 0) \\ \tilde{w} &= (u, 0, 0, \sqrt{1 - \|w\|}) \end{aligned}$$

Les vecteurs $\tilde{u}, \tilde{v}, \tilde{w}$ sont unitaires donc on peut appliquer le point précédent dans $\tilde{H}$ à ces trois vecteurs :

$$\sqrt{1 - \langle u \mid w \rangle_{\tilde{H}}} \leqslant \sqrt{1 - \langle u \mid v \rangle_{\tilde{H}}} + \sqrt{1 - \langle v \mid w \rangle_{\tilde{H}}}.$$

On en déduit le résultat annoncé en calculant les produits scalaires apparaissant ci-dessus.

□

On va maintenant introduire l'objet de ce paragraphe :

Définition I.24

Soient f, g deux fonctions multiplicatives 1-bornées et fixons un seuil $X \in \mathbb{R}_+^* \cup \{+\infty\}$. ***La distance prétentieuse*** *$\mathbb{D}(f, g; X)$ entre f et g à l'échelle X est :*

$$\mathbb{D}(f, g; X) = \sqrt{\sum_{p \leqslant X} \frac{1 - \mathcal{R}e\Big(f(p)\overline{g(p)}\Big)}{p}}.$$

On va maintenant prouver quelques propriétés de base sur cet objet :

Proposition I.25

Soient f, g, h des fonctions multiplicatives 1-bornées et soit $X \geqslant 2$. On a les propriétés suivantes :

— $\mathbb{D}(f, g; X) \geqslant 0$ et il y a égalité si et seulement si $f(p) = g(p)$ et $|f(p)| = 1$ pour tout $p \leqslant X$,

— $\mathbb{D}(f, g; X) = \mathbb{D}(g, f; X)$,

— $\mathbb{D}(f, h; X) \leqslant \mathbb{D}(f, g; X) + \mathbb{D}(g, h; X)$.

Avant la preuve de ces résultats, la première chose que l'on remarque est que la distance prétentieuse n'est pas une distance. Pourtant, elle possède quand même des propriétés appréciables que vérifie une distance comme l'inégalité triangulaire et la symétrie et c'est pour cela qu'elle est utile en pratique.

Démonstration :

— Vérifions d'abord que $\mathbb{D}(f, g; X)$ est positif. Pour cela, on va vérifier que chaque terme à l'intérieur de la somme est positif. Soit p premier, on a les majorations suivantes :

$$\begin{aligned}\mathcal{R}e\Big(f(p)\overline{g(p)}\Big) &= \mathcal{R}e(f(p))\mathcal{R}e(g(p)) + \mathcal{I}m(f(p))\mathcal{I}m(g(p)) \\ &\leqslant \frac{1}{2}\left(\mathcal{R}e(f(p))^2 + \mathcal{R}e(g(p))^2\right) + \frac{1}{2}\left(\mathcal{I}m(f(p))^2 + \mathcal{I}m(g(p))^2\right) \\ &\leqslant \frac{1}{2}\left(|f(p)|^2 + |g(p)|^2\right) \leqslant 1.\end{aligned}$$

On en déduit que chaque terme de la somme de $\mathbb{D}(f, g; X)$ est bien positif.

— On va maintenant montrer l'équivalence. Si $f = g$ et que $|f(p)| = 1$ pour tout p alors clairement $\mathbb{D}(f, g; X) = 0$. Réciproquement, supposons que $\mathbb{D}(f, g; X) = 0$, alors comme tous les termes de la somme sont positifs, on en déduit que :

$$\forall p : \quad \mathcal{R}e\Big(f(p)\overline{g(p)}\Big) = 1 \qquad (*).$$

Cela veut donc dire que les majorations du premier point sont en fait des égalités. On en déduit que :

$$\forall p : \quad \mathcal{R}e(f(p))\mathcal{R}e(g(p)) = \frac{1}{2}\left(\mathcal{R}e(f(p))^2 + \mathcal{R}e(g(p))^2\right).$$

Donc pour tout p : $\mathcal{R}e(f(p)) = \mathcal{R}e(g(p))$. De la même manière : $\mathcal{I}m(f(p)) = \mathcal{I}m(g(p))$ pour tout p. On obtient alors que pour tout p premier inférieur ou égal à x : $|f(p)| = |g(p)| = 1$ et $f(p) = g(p)$.

— On a pour tout p :

$$\mathcal{R}e\Big(f(p)\overline{g(p)}\Big) = \mathcal{R}e(f(p))\mathcal{R}e(g(p)) + \mathcal{I}m(f(p))\mathcal{I}m(g(p)),$$

qui est symétrique en f et en g. On en déduit que $\mathbb{D}(f, g; X) = \mathbb{D}(g, f; X)$.

— On applique le lemme I.23 à $H = \mathbb{D}$ muni du produit scalaire $\langle u \mid v\rangle = \mathcal{R}e(u\overline{v})$, pour tous $u, v \in \mathbb{D}$ (tous les termes sont positifs) :

$$\begin{aligned}1 - \mathcal{R}e\Big(f(p)\overline{h(p)}\Big) \leqslant &\left(1 - \mathcal{R}e\Big(f(p)\overline{g(p)}\Big)\right) + \left(1 - \mathcal{R}e\Big(g(p)\overline{h(p)}\Big)\right) \\ &+ 2\sqrt{\left(1 - \mathcal{R}e\Big(f(p)\overline{g(p)}\Big)\right)\left(1 - \mathcal{R}e\Big(g(p)\overline{h(p)}\Big)\right)}.\end{aligned}$$

On en déduit en sommant sur p puis par l'inégalité de Cauchy-Schwarz :

$$\mathbb{D}(f,h;X)^2 \leqslant \mathbb{D}(f,g;X)^2 + \mathbb{D}(g,h;X)^2 + 2\mathbb{D}(f,g;X)\mathbb{D}(g,h;X) = \left(\mathbb{D}(f,g;X) + \mathbb{D}(g,h;X)\right)^2.$$

On en déduit l'inégalité triangulaire car chaque distance prétentieuse est positive par le premier point.

□

On finit par introduire quelques définitions liées à la distance prétentieuse qui nous seront plutôt utiles dans les parties suivantes.

Définition I.26

Pour toute fonction multiplicative g 1-bornée, pour tout $Q \geqslant 1$ et pour tout $X \geqslant 2$, on introduit les quantités suivantes :

$$M(g;X) = \inf_{|t|\leqslant X} \mathbb{D}(g, n \mapsto n^{it}; X).$$

$$M(g;X,Q) = \inf_{\chi[q], q\leqslant Q} M(g\overline{\chi};X) \quad \text{et} \quad M(g;\infty,\infty) = \inf_{\chi,t} \mathbb{D}(g, n\mapsto \chi(n)n^{it}, \infty)^2.$$

Lorsque $M(g;X)$ est petit alors g "prétend" être $n \mapsto n^{it}$ et $M(g;X,Q)$ est petit lorsque g "prétend" être un caractère de Dirchlet modulo au plus Q tordu de facteur de torsion au plus X.

I.3.3 Inégalité faible de Halasz

Le théorème d'Halasz est un des théorèmes fondamentaux sur l'existence de moyennes de fonctions multiplicatives. Il est difficile à démontrer, c'est pour cela que l'on va se contenter d'un version faible du théorème d'Halasz qu'on appellera *inégalité faible d'Halasz* dont on trouvera la preuve dans le rapport complet (car même pour une version dite faible, c'est un résultat assez long à obtenir)

Théorème I.27: Inégalité faible de Halasz

Soit $f : \mathbb{N} \to \mathbb{C}$ une fonction multiplicative 1-bornée. Soient $T \geqslant 1$ et $M \geqslant 0$ et soit x que l'on suppose suffisamment grand. Supposons aussi que :

$$\forall\, |t| \leqslant T : \mathbb{D}(f, n \mapsto n^{it}; x) \geqslant M.$$

Alors :

$$\frac{1}{x}\sum_{n\geqslant x} f(n) \ll (1+M)e^{-\frac{M}{2}} + \frac{1}{T}.$$

II Théorèmes de Matomäki-Radziwiłł

L'objectif de cette partie est de s'intéresser à des moyennes sont pas sur des intervalles de la forme $[0,x]$ quand $x \to +\infty$ mais de la forme $[x,x+h]$ quand $x \to +\infty$ et $h \underset{x\to+\infty}{\to} +\infty$ aussi lentement que l'on souhaite. Par exemple, on peut prendre $h = \sqrt{x}$ ou $h = \log(x)$ ou $h = \log\log(x)$...

II.1 Quelques résultats préliminaires

Dans cette partie on va démontrer quelques lemmes utiles lorsqu'on s'intéressa aux résultats de Matomäki et Radziwiłł.

Lemme II.1

Soit $(a_n)_{n\in\mathbb{N}}$ une suite de nombres à valeurs complexes, nulle à partir d'un certain rang. On définit le polynôme de Dirichlet associé à cette suite par :

$$\forall y \in \mathbb{R}, \quad A(y) = \sum_{n\in\mathbb{N}} a_n n^{iy}.$$

Soit $T \geqslant 1$, alors :

$$\int_0^{+\infty} \left| \sum_{xe^{-\frac{1}{T}} < n \leqslant xe^{\frac{1}{T}}} a_n \right|^2 \frac{\mathrm{d}x}{x} = \frac{2}{\pi} \int_{-\infty}^{+\infty} |A(y)|^2 \left(\frac{\sin\left(\frac{y}{T}\right)}{y} \right) \mathrm{d}y.$$

Démonstration : On introduit la fonction $f : x \in \mathbb{R} \mapsto \sum_{e^{x-\frac{1}{T}} \leqslant n \leqslant e^{x+\frac{1}{T}}} a_n$. Sa transformée de Fourier est donnée par :

$$\forall \xi \in \mathbb{R} : \hat{f}(\xi) = \int_{-\infty}^{+\infty} f(x) e^{-ix\xi} \,\mathrm{d}x = \sum_{n\in\mathbb{N}} \int_{\log(n)-\frac{1}{T}}^{\log(n)+\frac{1}{T}} e^{-ix\xi} \,\mathrm{d}x = A(-\xi) \frac{2\sin\left(\frac{\xi}{T}\right)}{\xi}.$$

Or, par le théorème de Plancherel :

$$\int_{-\infty}^{+\infty} |f(x)|^2 \,\mathrm{d}x = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left|\hat{f}(\xi)\right|^2 \mathrm{d}\xi.$$

Et le changement de variable "$u = e^x$" donne :

$$\int_{-\infty}^{+\infty} |f(x)|^2 \,\mathrm{d}x = \int_0^{+\infty} \left| \sum_{ue^{-\frac{1}{T}} < n \leqslant ue^{\frac{1}{T}}} a_n \right|^2 \frac{\mathrm{d}u}{u}.$$

On en déduit le résultat annoncé. □

Lemme II.2

Soit X suffisamment grand et a_n, A définies comme précédemment. On suppose aussi qu'il existe deux constantes c_1, c_2 telles que $a_n = 0$ en dehors de $[c_1X, c_2X]$. Soit $h \in \left[1, \frac{c_1X}{10}\right]$. Alors :

$$\int_0^{+\infty} \left| \sum_{x<n\leqslant x+h} a_n \right|^2 \ll \frac{c_2^2}{c_1} X \int_{-\infty}^{+\infty} |A(y)|^2 \min\left(\frac{h^2}{c_1^2X^2}, \frac{1}{y^2}\right) \mathrm{d}y.$$

Démonstration : On définit la fonction $\mathcal{A} : x \in \mathbb{R} \mapsto \sum_{n\leqslant x} a_n$. Pour tout $\nu \in [2h, 3h]$, on a par l'inégalité triangulaire et par l'inégalité de Cauchy-Scwharz :

$$\int_0^{+\infty} |\mathcal{A}(x+h) - \mathcal{A}(x)|^2 \,\mathrm{d}x \leqslant 2 \int_0^{+\infty} \left(|\mathcal{A}(x+\nu) - \mathcal{A}(x)|^2 + |\mathcal{A}(x+h) - \mathcal{A}(x+\nu)|^2 \right) \mathrm{d}x.$$

Donc en intégrant sur $[2h, 3h]$, on obtient :

$$\begin{aligned}\int_0^{+\infty} |\mathcal{A}(x+h) - \mathcal{A}(x)|^2 \,\mathrm{d}x &\ll \int_{2h}^{3h} \int_0^{+\infty} |\mathcal{A}(x+\nu) - \mathcal{A}(x)|^2 \,\mathrm{d}x \,\mathrm{d}\nu + \int_{2h}^{3h} \int_0^{+\infty} |\mathcal{A}(x+h) - \mathcal{A}(x+\nu)|^2 \,\mathrm{d}x \,\mathrm{d}\nu \\ &\ll \int_{C_1\frac{X}{2}}^{C_2X} \int_h^{3h} |\mathcal{A}(x+\nu) - \mathcal{A}(x)|^2 \,\mathrm{d}\nu \,\mathrm{d}x \quad \text{en utilisant l'hypothèse sur la suite } a_n.\end{aligned}$$

Maintenant dans l'intégrale en ν, on fait le changement de variable "$\delta = x/\nu$. Ceci nous permet d'obtenir :

$$\int_{C_1\frac{X}{2}}^{C_2X} \int_h^{3h} |\mathcal{A}(x+\nu) - \mathcal{A}(x)|^2 \,\mathrm{d}\nu \,\mathrm{d}x \ll \int_{C_1\frac{X}{2}}^{C_2X} \int_{\frac{h}{C_2X}}^{\frac{6h}{C_1X}} |\mathcal{A}(x(1+\delta)) - \mathcal{A}(x)|^2 \, x \,\mathrm{d}\delta \,\mathrm{d}x.$$

Et avec la majoration $x \leqslant \dfrac{C_2^2X^2}{x}$ valable pour tout $x \in \left[\dfrac{C_1X}{2}, C_2X\right]$, on obtient finalement :

$$\int_0^{+\infty} |\mathcal{A}(x+h) - \mathcal{A}(x)|^2 \,\mathrm{d}x \ll \frac{c_2^2hX}{c_1} \max_{\frac{h}{c_2X}\leqslant\delta\leqslant\frac{6h}{C_1X}} \left(\int_{C_1\frac{X}{2}}^{C_2X} |\mathcal{A}(x(1+\delta)) - \mathcal{A}(x)|^2 \,\mathrm{d}\delta \frac{\mathrm{d}x}{x} \right).$$

Soit $\dfrac{h}{c_2X} \leqslant \delta \leqslant \dfrac{6h}{C_1X}$. D'après le lemme précédent appliqué avec $T = \dfrac{2}{\log(1+\delta)}$ après avoir effectué le changement de variable "$u = x\sqrt{1+\delta}$:

$$\frac{c_2^2hX}{c_1} \max_{\frac{h}{c_2X}\leqslant\delta\leqslant\frac{6h}{C_1X}} \left(\int_{C_1\frac{X}{2}}^{C_2X} |\mathcal{A}(x(1+\delta)) - \mathcal{A}(x)|^2 \,\mathrm{d}\delta \frac{\mathrm{d}x}{x} \right) \ll \frac{c_2^2hX}{c_1} \max_{\frac{h}{c_2X}\leqslant\delta\leqslant\frac{6h}{C_1X}} \left(\int_0^{+\infty} |A(y)|^2 \min\left(\frac{1}{T^2}, \frac{1}{y^2}\right) \right).$$

Or :

$$\frac{1}{T^2} \ll \frac{\delta^2}{4} \ll \frac{h^2}{c_1^2X^2}.$$

On en déduit que :

$$h \int_0^{+\infty} \left| \sum_{x<n\leqslant x+h} a_n \right|^2 \ll \frac{c_2^2}{c_1} hX \int_{-\infty}^{+\infty} |A(y)|^2 \min\left(\frac{h^2}{c_1^2X^2}, \frac{1}{y^2}\right) \mathrm{d}y.$$

Le fait annoncé arrive directement en divisant cette relation par h. □

Lemme II.3

Pour tout $\delta > 0$, on a uniformément en $t \in \mathbb{R}$:

$$\sum_{n \leqslant x} \lambda(n) n^{it} \ll x \exp\left(\frac{-\log(x)}{\log(x+|t|)^{\frac{2}{3}+\delta}}\right),$$

$$\sum_{p \leqslant x} p^{it} \ll \frac{\pi(x)}{1+|t|} + x \exp\left(\frac{-\log(x)}{\log(x+|t|)^{\frac{2}{3}+\delta}}\right).$$

On ne prouvera pas ce lemme, on pourra trouver une démonstration dans [3]. Elle se base sur des résultats d'analyse complexe qu'on a pas eu le temps d'aborder pendant le stage.

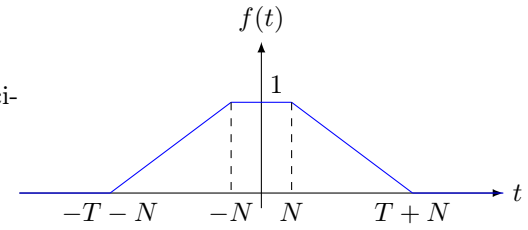
Lemme II.4

Pour toute suite de nombres complexes $(a_n)_{n \in \mathbb{N}}$, on a :

$$\int_{-T}^{T} \left| \sum_{n \leqslant N} a_n n^{it} \right|^2 \mathrm{d}t \ll (T+N) \sum_{n \leqslant N} |a_n|^2 .$$

Démonstration :

Soit f la fonction continue par morceaux définie ci-contre.



Ainsi on peut écrire :

$$\int_{-T}^{T} \left| \sum_{n \leqslant N} a_n n^{it} \right|^2 \mathrm{d}t \leqslant \int_{\mathbb{R}} f(t) \left| \sum_{n \leqslant N} a_n n^{it} \right|^2 \mathrm{d}t = \sum_{1 \leqslant m,n \leqslant N} a_m \overline{a_n} F\left(\frac{m}{n}\right),$$

où on note F la fonction $x \mapsto \displaystyle\int_{\mathbb{R}} f(t) x^{it} \, \mathrm{d}t$. On remarque que $F(1) = \displaystyle\int_{\mathbb{R}} f(t) \, \mathrm{d}t = 2T + N$. Un calcul un peu pénible donne que pour tout $x \neq 1$:

$$F(x) = -\frac{2\sin(T)}{N \log(x)^2} - \frac{2\cos(N+T)}{N \log(x)^2} \ll \frac{1}{N} \frac{1}{\log(x)^2}.$$

Donc pour $1 \leqslant m \neq n \leqslant N$, on a :

$$F\left(\frac{m}{n}\right) \ll \frac{1}{N} \frac{1}{\log\left(\frac{m}{n}\right)^2} \ll \frac{1}{n} \left(\frac{m+n}{m-n}\right)^2 \ll \frac{N}{(m-n)^2}.$$

On déduit de ces deux points :

$$\int_{-T}^{T}\left|\sum_{n\leqslant N}a_n n^{it}\right|^2 \mathrm{d}t \ll F(1)\sum_{1\leqslant n\leqslant N}|a_n|^2+\sum_{1\leqslant m\neq n\leqslant N}a_m\overline{a_n}F\left(\frac{m}{n}\right)$$
$$\ll (2T+N)\sum_{1\leqslant n\leqslant N}|a_n|^2+N\sum_{1\leqslant m\neq n\leqslant N}|a_n a_m|\,\frac{1}{(m-n)^2}.$$

Le dernier terme étant négligeable devant le premier (par exemple par l'inégalité $2|ab|\leqslant |a|^2+|b|^2$ pour tous $a,b\in\mathbb{C}$), on en déduit le résultat annoncé. □

II.2 Théorème de Matomäki et Radziwiłł

Le résultat fondamental obtenu par Matomäki et Radziwiłł s'énonce comme ceci.

Théorème II.5

Soit $f:\mathbb{N}\to[-1;1]$ une fonction multiplicative. Alors il existe des constantes absolues $C,C'>1$ telles que pour tous $2\leqslant h\leqslant X$ et $\delta>0$:

$$\left|\frac{1}{h}\sum_{x\leqslant n\leqslant x+h}f(n)-\frac{1}{X}\sum_{X\leqslant n\leqslant 2X}f(n)\right|\leqslant \delta+C'\frac{\log\log(h)}{\log(h)}$$

pour tous les entiers $x\in[X,2X]$, sauf au plus $CX\left(\frac{\log(h)^{\frac{1}{3}}}{\delta^2 h^{\frac{\delta}{25}}}+\frac{1}{\delta^2\log(X)^{\frac{1}{50}}}\right)$ entiers qu'on appellera *exceptions*.

Nous n'allons pas le démontrer en généralité mais dans le cas particulier de la fonction de Liouville λ et en obtenant une borne du nombre d'exceptions moins satisfaisante. Cependant, on retrouve les mêmes idées générales mais avec un peu moins d'estimations techniques. Le résultat qu'on va démontrer est le suivant.

Théorème II.6

Pour tout $\varepsilon>0$, il existe $H(\varepsilon)$ tel que pour tout $H(\varepsilon)<h\leqslant X$, on a :

$$\int_X^{2X}\left|\sum_{x<n\leqslant x+h}\lambda(n)\right|^2\leqslant \varepsilon X h^2.$$

En admettant d'abord ce résultat et d'après l'inégalité de Markov, on obtient le nombre d'entiers $x\in[X,2X]$ tels que $\left|\sum_{x<n\leqslant x+h}\lambda(n)\right|\leqslant \varepsilon^{\frac{1}{3}}h$ est majoré par $\varepsilon^{1/3}X$. Ainsi on sait que pour presque tous les entiers $x\in[1,2X]$ on a l'estimation : $\sum_{x<n\leqslant x+\psi(X)}\lambda(n)=o(\psi(X))$. pour $\psi(X)\to+\infty$ quand $X\to+\infty$. En appliquant directement le théorème II.5 à la fonction de Liouville λ on obtient le même résultat. C'est en ceci que le théorème II.6 constitue un résultat plus faible que le théorème II.5 obtenu par Matomäki et Radziwiłł. Passons désormais à la preuve du théorème II.6.

Démonstration : Nous allons seulement démontrer le résultat pour $h \geqslant \exp\left((\log(X))^{\frac{3}{4}}\right)$. On peut aussi supposer $h \leqslant \sqrt{X}$: si on prouve le résultat pour $h \leqslant \sqrt{X}$ alors on peut en déduire le résultat juste pour $h \geqslant \exp\left((\log(X))^{\frac{3}{4}}\right)$. En effet si $H := \exp\left((\log(X))^{\frac{3}{4}}\right)$, alors il existe $k \in \mathbb{N}$ tel que $H - kh \in [\sqrt{x}, \sqrt{x}+h]$. Alors on peut écrire $H = \sum_{i=0}^{k} h_i$ avec $h_i \leqslant \sqrt{X}$. Et ainsi :

$$\sum_{x \leqslant n \leqslant x+H} \lambda(n) = \sum_{i=0}^{k-1} \sum_{x+\sum_{j=0}^{i} h_j \leqslant n \leqslant x+\sum_{j=0}^{i+1} h_j} \lambda(n).$$

Ainsi si on montre le résultat pour tout $h \in [H, \sqrt{X}]$, alors on démontrera le résultat dans le cas $h \geqslant H$. On va maintenant adopter quelques notations :

- On note $\mathcal{P}$ l'ensemble des nombres premiers dans $[H, h]$.
- Pour tout $j \in \mathbb{N}$, on note $\mathcal{P}_j := \mathcal{P} \cap [P_j, P_{j+1}]$ où on définit pour tout $j \in \mathbb{N}$: $P_j := 2^j H$.
- On pose $J := \left\lfloor \frac{\log(h) - (\log(h))^{\frac{9}{10}}}{\log(2)} \right\rfloor$, ainsi on a $\mathcal{P} = \bigcup_{j=0}^{J} \mathcal{P}_j$.
- On note $\omega_{\mathcal{P}}$ l'application ω où on ne considère que les nombres premiers dans $\mathcal{P}$, c'est-à-dire :

$$\forall n \in \mathbb{N} : \omega_{\mathcal{P}}(n) = \sum_{\substack{p \in \mathcal{P} \\ p | n}} 1.$$

- On note, en référence à I.18 qu'on va utiliser la quantité $W(\mathcal{P}) = \sum_{p \in \mathcal{P}} \frac{1}{p}$. D'après I.15, on a : $W(\mathcal{P}) \sim \frac{1}{10} \log\log(h)$.
- Ce choix de $\mathcal{P}$ a été fait pour qu'il vérifie la condition suivante : tout élément $p \in \mathcal{P}$ vérifie $\exp\left(\log(h)^{\frac{9}{10}}\right) \leqslant p \leqslant h$ donc $\exp\left(\log(X)^{\frac{2}{3}}\right) \leqslant \exp\left(\log(X)^{\frac{27}{40}}\right) \leqslant p \leqslant h$ (en anticipant l'application du lemme II.3).

Soit $(a_n)_{n \in \mathbb{N}}$ une suite qui vérifie pour tout $y \in \mathbb{R}$:

$$A(y) := \sum_{n \in \mathbb{N}} a_n n^{iy} = \sum_{j=0}^{J} \sum_{p \in \mathcal{P}_j} \sum_{\frac{X}{P_{j+1}} \leqslant m \leqslant \frac{2X}{P_j}} \lambda(p) p^{iy} \lambda(m) m^{iy}.$$

Cette suite vérifie :

- $a_n = 0$ pour tout n en dehors de $\left[\frac{X}{2}, 4X\right]$,
- $a_n = \lambda(n)\omega_{\mathcal{P}}(n)$ pour tout $n \in \{X, \dots, 2X\}$.

Par inégalité triangulaire, on a :

$$\begin{aligned} W(\mathcal{P})^2 \int_X^{2X} \left(\sum_{x < n \leqslant x+h} \lambda(n) \right)^2 \mathrm{d}x & \\ \ll \int_X^{2X} \left(\sum_{x < n \leqslant x+h} \lambda(n)\omega_{\mathcal{P}}(n) \right)^2 \mathrm{d}x + \int_X^{2X} \left(\sum_{x < n \leqslant x+h} \lambda(n)(\omega_{\mathcal{P}} - W(\mathcal{P})) \right)^2 \mathrm{d}x, & \\ \ll X \int_{\mathbb{R}} |A(y)|^2 \min\left(\frac{h^2}{X^2}, \frac{1}{y^2} \right) \mathrm{d}y + XW(\mathcal{P})h^2. & \end{aligned}$$

On a obtenu la dernière estimation en appliquant le lemme II.2 pour le premier terme et l'inégalité de Turan-Kubilius I.18 pour le deuxième terme. Il nous reste donc à estimer l'intégrale. Pour cela on va d'abord gérer la contribution des $|y| \geqslant X$ à l'intégrale, pour cela on décomposer cet ensemble en sous-ensembles dyadiques et appliquer II.4 sur chacun de ces sous-ensembles :

$$\begin{aligned}\int_{|y|>X} |A(y)| \min\left(\frac{h^2}{X^2}, \frac{1}{y^2}\right) \mathrm{d}y &= \int_{|y|>X} |A(y)| \frac{1}{y^2} \,\mathrm{d}y, \\ &= \sum_{k=0}^{+\infty} \int_{2^k X \leqslant |y| \leqslant 2^{k+1} X} |A(y)|^2 \frac{\mathrm{d}y}{y^2} \\ &\ll \frac{1}{X^2} \sum_{k=0}^{+\infty} \frac{1}{2^{2k}} \int_{|y| \leqslant 2^{k+1} X} |A(y)|^2 \,\mathrm{d}y, \\ &\ll \frac{1}{X^2} \sum_{k=0}^{+\infty} \frac{2^{k+1}X + 4X}{2^{2k}} \sum_{n \leqslant 4X} |a_n|^2 .\end{aligned}$$

Or par définition de la suite (a_n) et par l'inégalité de Turan-Kubilius I.18, on a :

$$\sum_{n \leqslant 4X} |a_n|^2 \leqslant \sum_{n \leqslant 4X} \omega_{\mathcal{P}}(n)^2 \ll X W(\mathcal{P})^2.$$

On en déduit que :

$$\int_X^{2X} \left(\sum_{x < n \leqslant x+h} \lambda(n) \right)^2 \mathrm{d}x \ll X + \frac{X}{W(\mathcal{P})^2} \int_{|y| \leqslant X} |A(y)|^2 \min\left(\frac{h^2}{X^2}, \frac{1}{y^2}\right) \mathrm{d}y + \frac{Xh^2}{W(\mathcal{P})}.$$

Notons que le premier terme (que l'on vient d'obtenir) est négligeable devant $\dfrac{Xh^2}{W(\mathcal{P})}$. Il reste donc à regarder l'intégrale restante. Pour cela, on utilise l'inégalité de Cauchy-Schwarz dans $|A(y)|^2$ pour montrer que :

$$\int_{|y| \leqslant X} |A(y)|^2 \min\left(\frac{h^2}{X^2}, \frac{1}{y^2}\right) \mathrm{d}y \ll W(\mathcal{P}) \max_{0 \leqslant j \leqslant J} I_j,$$

où pour tout $j \in \{0, \dots, J\}$:

$$I_j = (\log(P_j))^2 \int_{-X}^{X} \left| \sum_{p \in \mathcal{P}_j} p^{iy} \right|^2 \left| \sum_{\frac{X}{P_{j+1}} \leqslant m \leqslant \frac{2X}{P_j}} \lambda(m) m^{iy} \right|^2 \min\left(\frac{h^2}{X^2}, \frac{1}{y^2}\right) \mathrm{d}y.$$

Il reste donc à estimer I_j pour tout j. On fixe un tel j et on utilise le lemme II.3 pour dire que :

$$\sum_{p \in \mathcal{P}_j} p^{iy} \ll \frac{P_j}{\log(P_j)} \frac{1}{1+|y|} + P_j \exp\left(-(\log(X))^{\frac{27}{40} - \frac{2}{3} - \delta}\right) \ll \frac{P_j}{\log(P_j)} \left(\frac{1}{1+|y|} + \frac{1}{\log(P_j)} \right). \tag{II.1}$$

En utilisant d'abord cette borne pour $X \geqslant |y| \geqslant \log(P_j)$, on obtient :

$$(\log(P_j))^2 \int_{\log(P_j)\leqslant|y|\leqslant X} \left|\sum_{p\in\mathcal{P}_j} p^{iy}\right|^2 \left|\sum_{\frac{X}{P_{j+1}}\leqslant m\leqslant\frac{2X}{P_j}} \lambda(m)m^{iy}\right|^2 \min\left(\frac{h^2}{X^2},\frac{1}{y^2}\right)\mathrm{d}y$$
$$\ll \frac{P_j^2}{(\log(P_j)^2}\int_{\log(P_j)\leqslant|y|\leqslant X}\left|\sum_{\frac{X}{P_{j+1}}\leqslant m\leqslant\frac{2X}{P_j}}\lambda(m)m^{iy}\right|^2 \min\left(\frac{h^2}{X^2},\frac{1}{y^2}\right)\mathrm{d}y.$$

Or, en faisant en découpant en sous-ensembles dyadiques (de la même manière qu'au dessus) et en appliquant le lemme II.4, on montre que :

$$\frac{P_j^2}{(\log(P_j)^2}\int_{\log(P_j)\leqslant|y|\leqslant X}\left|\sum_{\frac{X}{P_{j+1}}\leqslant m\leqslant\frac{2X}{P_j}}\lambda(m)m^{iy}\right|^2 \min\left(\frac{h^2}{X^2},\frac{1}{y^2}\right)\mathrm{d}y \ll \frac{h^2}{P_j^2}.$$

Pour gérer la contribution $|y| \leqslant \log(P_j)$ à I_j, on utilise le lemme II.3 qui nous dit que :

$$\sum_{\frac{X}{P_{j+1}}\leqslant m\leqslant\frac{2X}{P_j}}\lambda(m)m^{iy} \ll \frac{X}{P_j(\log(X)^{10}}.$$ [1]

Ainsi en utilisant l'estimation II.1, on arrive à :

$$(\log(P_j))^2\int_{|y|\leqslant\log(P_j)}\left|\sum_{p\in\mathcal{P}_j}p^{iy}\right|^2\left|\sum_{\frac{X}{P_{j+1}}\leqslant m\leqslant\frac{2X}{P_j}}\lambda(m)m^{iy}\right|^2\min\left(\frac{h^2}{X^2},\frac{1}{y^2}\right)\mathrm{d}y \ll \frac{h^2}{\log(X)^{-19}}.$$

Ainsi on obtient que $I_j \ll \dfrac{h^2}{(\log(P_j)^2}$. Donc en tout :

$$\int_X^{2X}\left(\sum_{x<n\leqslant x+h}\lambda(n)\right)^2\mathrm{d}x \ll \frac{Xh^2}{\log\log(h)}.$$

On obtient ainsi le théorème demandé pour $h \geqslant \exp\left((\log(X))^{\frac{3}{4}}\right)$.

□

La démonstration dans le cas général de Matomäki et Radziwiłł reprend le même cheminement que la notre mais en ajoutant quelques estimations plus fines pour ne pas avoir besoin d'une borne inférieure pour h. De plus ils ont restreint leur somme sur un ensemble S qu'ils appellent dense. En fait il comprend "beaucoup" d'entiers et permet d'enlever certains cas pathologiques qui sont les entiers qui ont une décomposition en facteurs premiers particulière. Pour plus de détails on renvoie vers le papier de Matomäki et Radziwiłł [**MatomÃď ki2016multiplicative**]

1. L'exposant 10 sur le log n'est pas obligatoire, on le choisi (et le lemme II.3 nous permet de le faire) suffisamment grand pour qu'à la fin on puisse négliger ce terme.

III Conjectures de Chowla et d'Elliott en moyenne logarithmique

III.1 Un petit raisonnement probabiliste vers la conjecture de Chowla

Connaître les corrélations à k points de fonctions multiplicatives est étroitement lié problèmes additifs dans les nombres premiers. On peut par exemple étudier le nombre de facteurs premiers d'un nombre n par rapport au nombre de facteurs premiers de $n+h$ pour $n, h \in \mathbb{N}$, s'intéresser aux corrélations à k points de la fonction de Liouville est donc une question naturelle.
En supposant la véracité de l'hypothèse de Riemann, la différence entre le nombre de valeurs de $n \leqslant x$ telles que $\lambda(n) = +1$ et le nombre de valeurs de $n \leqslant x$ telles que $\lambda(n) = -1$ est approximativement de l'ordre de $O\left(x^{1/2+\varepsilon}\right)$. Ceci est cohérent avec le comportement d'une suite de variables aléatoires indépendantes de Rademacher $\mathbf{X}_n$ (en supposant les valeurs de la fonction de Liouville comme indépendantes), prenant les valeurs $+1$ et -1 avec une probabilité égale par le théorème limite central.
Or, par indépendance des ces variables aléatoires, on sait que aussi que pour tout $k, n \in \mathbb{N}^*$, $h_1, \ldots, h_k$ des entiers naturels distincts et $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, +1\}$: $\mathbf{P}(\mathbf{X}_{n+h_1} = \varepsilon_1, \ldots, \mathbf{X}_{n+h_k} = \varepsilon_k) = 2^{-k}$. Si les valeurs de λ sont effectivement essentiellement indépendantes, il est donc raisonnable de penser que les composantes des vecteurs $(\lambda(n+h_1), \ldots, \lambda(n+h_k))$ changent indépendamment lorsque n varie, c'est à dire :

$$\frac{1}{x}\left|\{n \leqslant x \quad \text{tel que} \quad \lambda(n+h_1) = \varepsilon_1, \ldots, \lambda(n+h_k) = \varepsilon_k\}\right| \xrightarrow[x\to+\infty]{} 2^{-k}.$$

Ceci est lié aux corrélations à k points de la fonction de Liouville par la proposition suivante :

Proposition III.1

Soient $k \in \mathbb{N}^*$, $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, +1\}$, $h_1, \ldots, h_k \in \mathbb{N}^*$. Les deux propositions sont équivalentes :

(i) $\quad \frac{1}{x}\left|\{n \leqslant x \quad \text{tel que} \quad \lambda(n+h_1) = \varepsilon_1, \ldots, \lambda(n+h_k) = \varepsilon_k\}\right| \xrightarrow[x\to+\infty]{} 2^{-k}.$

(ii) $\quad \sum_{n=1}^{x} \lambda(n+h_1)\ldots\lambda(n+h_k) = o_{x\to+\infty}(x).$

Démonstration : Dans cette preuve on va adopter la notation suivante, pour $S \subset \{1, \ldots, k\}$, on posera $\varepsilon_S = \prod_{j\in S} \varepsilon_j$. On va procéder directement par équivalence : la proposition (i) est équivalente à

$$\lim_{x\to+\infty} \frac{1}{x} \sum_{n\leqslant x} \mathbb{1}_{\lambda(n+h_1)=\varepsilon_1} \ldots \mathbb{1}_{\lambda(n+h_k)=\varepsilon_k} = 2^{-k}.$$

Or comme on a pour tout $j \in \{1, \ldots, k\}$: $\dfrac{1+\varepsilon_j\lambda(n+h_j)}{2} = \mathbb{1}_{\lambda(n+h_j)=\varepsilon_1}$, ceci est équivalent à :

$$\lim_{x\to+\infty} \frac{1}{x} \sum_{n\leqslant x} \prod_{j=1}^{k} (1+\varepsilon_j\lambda(n+h_j)) = 1.$$

C'est-à-dire :

$$\lim_{x\to+\infty} \frac{1}{x} \sum_{n\leqslant x} \sum_{S\subset\{1,\ldots,k\}} \varepsilon_S \prod_{j\in S} \lambda(n+h_j) = 1,$$

ce que l'on peut encore réécrire :

$$\lim_{x\to+\infty} \sum_{S\subset\{1,\dots,k\}} \frac{1}{x} \sum_{n\leqslant x} \varepsilon_S \prod_{j\in S} \lambda(n+h_j) = 1.$$

Or, ici tous les termes de la somme sur $S \subset \{1,\dots,k\}$ sont positifs donc cette dernière proposition est équivalente à (ii) (en basculant de l'autre côté de l'égalité le terme associé à $S = \varnothing$). *De facto*, on vient de prouver l'équivalence entre (i) et (ii). □

C'est sous cette deuxième forme que Sarvadaman D. S. Chowla a émis cette conjecture dans [4].

Conjecture III.2: (Chowla)

Soient $k \in \mathbb{N}^*$, $\varepsilon_1,\dots,\varepsilon_k \in \{-1,+1\}$, $h_1,\dots,h_k \in \mathbb{N}^*$. On a :

$$\sum_{n=1}^{x} \lambda(n+h_1)\dots\lambda(n+h_k) = o_{x\to+\infty}(x).$$

C'est toujours à l'heure actuelle une conjecture mais les résultats de Matomäki et Radziwiłł permettent d'obtenir de nombreuses avancées. On va dans cette partie présenter deux versions "en moyenne" : une sur les paramètres $(h_1,\dots,h_k)$ qui sera admise (et qui est faite dans le rapport complet), la deuxième "en moyenne logarithmique" qu'on démontrera et qui sera un des ingrédients principaux à la démonstration de la conjecture d'Erd**H**os dans la dernière partie.

III.2 Conjecture d'Elliott et sa version en moyenne

Une généralisation de la conjecture de Chowla est la suivante, appelée conjecture d'Elliott :

Conjecture III.3: (Elliott)

Soient $g_1,\dots,g_k : \mathbb{N} \to \mathbb{R}$ des fonctions multiplicatives 1-bornées et soient $(a_1,\dots,a_k,b_1,\dots,b_k) \in \mathbb{N}^{2k}$ tels que les (a_i,b_i) soient $\mathbb{Q}$-libres deux à deux. On suppose de plus qu'il existe $j_0 \in \{1,\dots,k\}$ tel que $M(g_{j_0};\infty,\infty) = +\infty$. Alors :

$$\sum_{1\leqslant n\leqslant X} \prod_{j=1}^{k} g_j(a_j n + b_j) = o(X).$$

Cette conjecture affirme que pour tous (a_i,b_i) $\mathbb{Q}$-libres deux à deux et pour toutes fonctions multiplicatives 1-bornées $g_1,\dots,g_k$, on a l'asymptotique ci-dessus à part si tous les g_j prétendent être un caractère de Dirichlet tordu. Comme la fonction λ de Liouville vérifie l'hypothèse de la conjecture d'Elliott, cette dernière implique la conjecture de Chowla. Si on veut généraliser ce résultat pour des fonctions multiplicatives à valeurs complexes il faut remplacer l'hypothèse sur les g_j par la suivante [2] :

$$\text{Il existe } j_0 \in \{1,\dots,k\} \text{ tel que pour tout } Q \geqslant 0 : \lim_{X\to+\infty} M(g_{j_0};X,Q) = +\infty.$$

Une première idée pour arriver à un tel résultat est de rechercher une version en moyenne sur des paramètres de *shift*. En voici une version quantitative dont la preuve sera passée (on pourra retrouver une partie entière consacrée à sa démonstration dans le rapport complet).

2. Elles sont d'ailleurs équivalentes dans le cas des fonctions g_i à valeurs réelles.

Théorème III.4: Conjecture d'Elliott en moyenne

Soient $10 \leqslant H \leqslant X$ et $A \geqslant 1$. Soient $g_1, \dots, g_k : \mathbb{N} \to \mathbb{C}$ des fonctions 1-bornées et soient $a_1, \dots, a_k, b_1, \dots, b_k \in \mathbb{N}$ tels que $a_j \leqslant A$ et $b_j \leqslant AX, \forall j \in \{1, \dots, k\}$. Soient $1 \leqslant j_0 \leqslant k$ et supposons que g_{j_0} est multiplicative. Alors on a :

$$\sum_{1 \leqslant h_1, \dots, h_k \leqslant H} \left| \sum_{1 \leqslant n \leqslant X} \prod_{j=1}^{k} g_j(a_j n + b_j + h_j) \right| \ll A^2 k \left(e^{-\frac{M}{80}} + \frac{\log\log(H)}{\log(H)} + \frac{1}{\log(X)^{\frac{1}{3000}}} \right) H^k X,$$

où $M = M(g_{j_0}; 10AX, Q)$ où $Q = \min\left(\log(X)^{\frac{1}{125}}, \log(H)^{20}\right)$. En fait on a même légèrement mieux en enlevant un paramètre :

$$\sum_{1 \leqslant h_2, \dots, h_k \leqslant H} \left| \sum_{1 \leqslant n \leqslant X} g_1(a_1 n + b_1) \prod_{j=2}^{k} g_j(a_j n + b_j + h_j) \right| \ll A^2 k \left(e^{-\frac{M}{80}} + \frac{\log\log(H)}{\log(H)} + \frac{1}{\log(X)^{\frac{1}{3000}}} \right) H^{k-1} X.$$

On va quand même donner deux résultats que l'on démontre pendant la preuve de la conjecture d'Elliott en moyenne et que l'on utilisera dans la suite.

Lemme III.5

Soient $10 < P_1 < Q_1 \leqslant X$ et $\sqrt{X} \leqslant \mathbf{X}_0 \leqslant X$ tels que $Q_1 \leqslant \exp(\sqrt{\log(\mathbf{X}_0)})$. Alors, pour tout X suffisamment grand :

$$|\{1 \leqslant n \leqslant X : n \notin S_{P_1, Q_1, \mathbf{X}_0, X}\}| \ll \frac{\log(P_1)}{\log(Q_1)} X.$$

Proposition III.6

Soient $X, H, W \geqslant 10$ tels que $(\log(H))^5 \leqslant W \leqslant \min\left(H^{\frac{1}{250}}, (\log(X))^{\frac{1}{125}}\right)$ et soit g une fonction multiplicative 1-bornée telle que $W \leqslant \exp\left(\frac{M(g; X, Q)}{3}\right)$. Soit enfin $S = S_{P_1, Q_1, \sqrt{X}, X}$ où $P_1 = W^{200}$ et $Q_1 = \frac{H}{W^3}$. Alors pour tout $\alpha \in \mathbf{S}^1$, on a :

$$\int_{\mathbb{R}} \left| \sum_{x \leqslant n \leqslant x+H} \mathbb{1}_S g(n) e(\alpha n) \right| \mathrm{d}x \ll \frac{\log(H)^{\frac{1}{4}} \log\log(H)}{W^{\frac{1}{4}}} HX.$$

III.3 Moyenne logarithmique de fonctions multiplicatives

On va dans ce paragraphe introduire une nouvelle manière de considérer une moyenne d'une fonction multiplicative qui est moins contraignante et qui permettra d'établir des résultats "en moyenne". Pour cela on va introduire la grandeur suivante.

Définition III.7

Soit $f : \mathbb{N} \to \mathbb{C}$ une fonction arithmétique. On dit que f admet une ***moyenne logarithmique*** *si la limite suivante existe :*

$$\lim_{x\to+\infty} \frac{1}{\log(x)} \sum_{k\leqslant x} \frac{f(k)}{k}.$$

Le résultat principal de ce paragraphe est le suivant :

Proposition III.8

Soit f une fonction arithmétique. Si elle possède une moyenne alors elle possède une moyenne logarithmique et elles sont égales. La réciproque est fausse.

Démonstration : Le premier fait est une conséquence directe de la formule de sommation par parties. Pour le sens réciproque, on va reprendre l'exemple I.2. On a vu que la fonction multiplicative $n \mapsto n^{it}$ ne possédait pas de valeur moyenne pour $t \neq 0$. Par contre elle possède une moyenne logarithmique car pour tout $x \geqslant 1$:

$$\sum_{n\leqslant x} n^{it-1} = \int_1^x u^{it-1}\,\mathrm{d}u + \mathcal{O}\left(\int_1^x (it-1)u^{it-2}\,\mathrm{d}u\right).$$

□

En fait la moyenne logarithmique ne prend pas en compte les oscillations de n^{it} lorsque n est grand ce qui donne une plus grande perméabilité quant à l'admission de moyenne logarithmique.

III.4 Présentation du résultat principal et réduction du problème

L'objectif de cette partie est d'établir une conjecture d'Elliott non asymptotique en moyenne logarithmique :

Théorème III.9

Soient $a_1, a_2 \in \mathbb{N}$, $b_1, b_2 \in \mathbb{Z}$ tels que $a_1b_2 - a_2b_1 \neq 0$. Soit $\varepsilon > 0$ et soit A suffisamment grand devant a_1, a_2, b_1, b_2 et $\frac{1}{\varepsilon}$. Soit $x \geqslant \omega \geqslant A$ et soient $g_1, g_2 : \mathbb{N} \to \mathbb{C}$ deux fonctions multiplicatives 1-bornées. Supposons que g_1 est non-prétentieuse, c'est-à-dire pour tout caractère de Dirichlet χ de module au plus A et pour tout $t \in [-Ax, +Ax]$:

$$\mathbb{D}(g_1; n \mapsto \chi(n)n^{it}, x) \geqslant A.$$

Alors :

$$\left|\sum_{\frac{x}{\omega}<n\leqslant x} \frac{1}{n} g_1(a_1n+b_1)g_2(a_2n+b_2)\right| \leqslant \varepsilon \log(\omega).$$

On peut montrer que pour démontrer le théorème précédent il suffit de démontrer le théorème suivant. On pourra en retrouver une preuve dans le rapport complet.

Théorème III.10: (Réduction du problème)

Soient $a \in \mathbb{N}$, $b \in \mathbb{Z}$ et $h \neq 0$. Soit $\varepsilon > 0$ et soit A suffisamment grand devant a, b, h et $\frac{1}{\varepsilon}$. Soit $x \geqslant \frac{x}{\log(x)} \geqslant \omega \geqslant A$ et soient $g_1, g_2 : \mathbb{N} \to \mathbf{S}^1$ deux fonctions complètement multiplicatives. Supposons que g_1 est non-prétentieuse, c'est-à-dire pour tout caractère de Dirichlet χ de module au plus A et pour tout $t \in [-Ax, +Ax]$:

$$\mathbb{D}(g_1; n \mapsto \chi(n)n^{it}, x) \geqslant A.$$

Alors :

$$\left| \sum_{\frac{x}{\omega} < n \leqslant x} \frac{1}{n} g_1(an+b)g_2(an+b+h) \right| \leqslant \varepsilon \log(\omega).$$

III.5 Passage à un point de vue probabiliste

Soient a, b, h, ε comme dans le théorème III.10. Supposons que celui-ci est faux pour ce choix de paramètres. Quitte à diminuer ε, on peut supposer que ε est suffisamment faible devant a, b, h. On introduit H_- suffisamment grand $a, b, h, \frac{1}{\varepsilon}$, H_+ suffisamment grand $a, b, h, \frac{1}{\varepsilon}, H_-$ et $A > 0$ suffisamment grand $a, b, h, \frac{1}{\varepsilon}, H_-, H_+$. On gardera en tête les deux hiérarchies suivantes (avec certaines notations qui seront introduites plus tard) :

$$a, b, h \ll \frac{1}{\varepsilon^2} \ll H_- \ll p \ll H \ll H_+ \ll A \leqslant \omega \leqslant \frac{x}{\log(x)} \leqslant x.$$

$$x \geqslant n \geqslant \frac{x}{\omega} \geqslant \log(x) \leqslant \log(A) \ll H_+.$$

Par hypothèse, il existe $x \geqslant \omega \geqslant A$ et $g_1, g_2 : \mathbb{N} \to \mathbf{S}^1$ complètement multiplicatives tels que pour tout caractère de Dirichlet χ de module au plus A et pour tout $|t| \leqslant Ax$:

$$\mathbb{D}(g_1; n \mapsto \chi(n)n^{it}, x) \geqslant A,$$

et tels que :

$$\left| \sum_{\frac{x}{\omega} < n \leqslant x} \frac{1}{n} g_1(an+b)g_2(an+b+h) \right| > \varepsilon \log(\omega). \tag{III.1}$$

Proposition III.11

Pour tout $H_- \leqslant H \leqslant H_+$, on a :

$$\sup_{\alpha \in \mathbf{S}^1} \sum_{\frac{x}{\omega} < n \leqslant x} \frac{1}{Hn} \left| \sum_{j=1}^{H} g_1(n+j)e(j\alpha) \right| \ll \frac{\log\log(H)}{\log(H)} \log(\omega).$$

En particulier :

$$\sup_{\alpha \in \mathbf{S}^1} \sum_{\frac{x}{\omega} < n \leqslant x} \frac{1}{Hn} \left| \sum_{j=1}^{H} g_1(n+j)e(j\alpha) \right| = o_{H_- \to +\infty}(\log(\omega)).$$

Démonstration : Soient $\alpha \in \mathcal{S}^1$, et $X \in \left[\frac{x}{2\omega}, 2x\right]$. On définit S comme au théorème III.6 avec $P_1 = W^{200}$ et $Q_1 = \dfrac{H}{W^3}$ où $W = \log(H)^5$. On écrit :

$$\sum_{X<n\leqslant 2X} \frac{1}{HX}\left|\sum_{j=1}^{H} g_1(n+j)e(j\alpha)\right| \ll \sum_{X<n\leqslant 2X} \frac{1}{HX}\left|\sum_{j=1}^{H} g_1(n+j)e(j\alpha)\right| + |S^c|\,.$$

Or d'après le lemme III.5, on a :

$$|S^c| \ll \frac{\log(P_1)}{H\log(Q_1)} \ll \frac{\log\log(H)}{H\log(H)}.$$

Ainsi d'après la proposition III.6, on obtient que :

$$\sum_{X<n\leqslant 2X} \frac{1}{HX}\left|\sum_{j=1}^{H} g_1(n+j)e(j\alpha)\right| \ll \frac{\log\log(H)}{\log(H)}.$$

On obtient alors par sommation partielle que :

$$\sum_{X<n\leqslant 2X} \frac{1}{n}\left|\sum_{j=1}^{H} g_1(n+j)e(j\alpha)\right| \ll \frac{\log\log(H)}{\log(H)}\log(X).$$

On en déduit facilement les deux points de la proposition en moyennisant. □

Il sera plus adapté dans le contexte de la sous-partie suivante, d'utiliser un langage probabiliste. On introduit une variable aléatoire discrète $\mathbf{n}$ sur $\left\{n \in \mathbb{N} : \frac{x}{\omega} < n \leqslant x\right\}$ en posant :

$$\mathbf{P}(\mathbf{n} = n) = \frac{1/n}{\sum_{\frac{x}{\omega}<k\leqslant x}\frac{1}{k}}.$$

Montrons que uniformément $H \in [H_-, H_+]$:

$$\sup_{\alpha\in\mathbf{S}^1} \mathbf{E}\left[\left|\sum_{j=1}^{H} g_1(\mathbf{n}+j)e(j\alpha)\right|\right] = o_{H\to+\infty}(H). \tag{III.2}$$

Comme $x \geqslant \omega \geqslant A$ et $\omega \leqslant \dfrac{x}{\log(x)}$, on a :

$$\sum_{\substack{n\in\mathbb{N}\\ \frac{x}{\omega}\leqslant n\leqslant x}} \frac{1}{n} = (1 + o_{A\to+\infty})\log(\omega).$$

Donc par (III.1), pour A suffisamment grand.

$$\mathbf{E}\left[g_1(a\mathbf{n}+b)g_2(a\mathbf{n}+b+h)\right] \gg \varepsilon.$$

On obtient le fait annoncé par la proposition III.11.

Lemme III.12

Soit $q \in \mathbb{N}$, $q \leqslant H_+$ et $r \in \mathbb{Z}$ tel que $|r| \leqslant H_+$. Alors pour toute variable aléatoire $X(\mathbf{n})$ à valeurs complexes dépendant de $\mathbf{n}$ et bornée en module par un nombre indépendant des paramètres du problème :

$$\mathbf{E}\left[X(\mathbf{n})\mathbb{1}_{\mathbf{n}\equiv r[q]}\right] = \frac{1}{q}\mathbf{E}\left[X(q\mathbf{n}+r)\right] + o_{A\to+\infty}(1).$$

Démonstration : La preuve ne présente pas de difficulté, pour les détails on se référera à [5]. □

Proposition III.13

Notons $\mathcal{P}_H$ l'ensemble des nombres premiers dans $\left[\frac{\varepsilon^2 H}{2}, \varepsilon^2 H\right]$. Pour tout p premiers on note $c_p = \overline{g_1(p)g_2(p)} \in \mathbf{S}^1$. Alors on a :

$$\mathbf{E}\left[\sum_{p\in\mathcal{P}_H}\sum_{j:\ \substack{j\in[1,H]\\ j+pH\in[1,H]}} c_p\mathbb{1}_{a\mathbf{n}+j\equiv pb[ap]}g_1(a\mathbf{n}+j)g_2(a\mathbf{n}+j+ph)\right] \gg \frac{\varepsilon H}{\log(H)}.$$

Démonstration : On note $X = \mathbf{E}\left[\mathbb{1}_{\mathbf{n}\equiv b[a]}g_1(\mathbf{n})g_2(\mathbf{n}+h)\right]$. Par complète multiplicativité de g_1 et g_2 et par définition de c_p, on voit que pour tout $p \in \mathcal{P}_H$:

$$X = \mathbf{E}\left[c_p\mathbb{1}_{p\mathbf{n}\equiv pb[ap]}g_1(p\mathbf{n})g_2(p\mathbf{n}+ph)\right].$$

D'après le lemme III.12, on a pour tout $j \in \{1,\dots,H\}$ et pour tout $p \in \mathcal{P}_H$:

$$\mathbf{E}\left[c_p\mathbb{1}_{\mathbf{n}+j\equiv pb[ap]}g_1(\mathbf{n}+j)g_2(\mathbf{n}+j+ph)\right] = \frac{X}{p} + o_{A\to+\infty}(1).$$

En sommant ceci pour $j = 1,\dots,H$, on obtient pour tout $p \in \mathcal{P}_H$:

$$\mathbf{E}\left[c_p\sum_{j=1}^{H}\mathbb{1}_{\mathbf{n}+j\equiv pb[ap]}g_1(\mathbf{n}+j)g_2(\mathbf{n}+j+ph)\right] = \frac{X}{p} + o_{A\to+\infty}(1).$$

On fixe désormais $p \in \mathcal{P}_H$ et on introduit pour $s \in \{0,\dots,a-1\}$, la quantité :

$$Q(s) = \mathbf{E}\left[c_p\sum_{j=1}^{H}\mathbb{1}_{\mathbf{n}+j\equiv pb[ap]}g_1(\mathbf{n}+j)g_2(\mathbf{n}+j+ph)\mathbb{1}_{n\equiv s[a]}\right].$$

Par le calcul précédent, on sait que $\sum_s Q(s) = \frac{HX}{p} + o_{A\to+\infty}(1)$, l'idée est de comparer $Q(s)$ et $Q(s+1)$ pour tout s et d'utiliser cette relation pour obtenir une estimation de $Q(0)$ et de la sommer sur $p \in \mathcal{P}_H$ afin d'obtenir une quantité pas trop éloignée de ce que l'on cherche. Tout d'abord, d'après le lemme III.12, on a :

$$\begin{aligned}Q(s+1) = Q(s) &+ \mathbf{E}\left[c_p\mathbb{1}_{\mathbf{n}+1\equiv pb[ap]}g_1(\mathbf{n}+1)g_2(\mathbf{n}+1+ph)\mathbb{1}_{n\equiv s[a]}\right]\\ &- \mathbf{E}\left[c_p\mathbb{1}_{\mathbf{n}+1+H\equiv pb[ap]}g_1(\mathbf{n}+H+1)g_2(\mathbf{n}+1+H+ph)\mathbb{1}_{n+H\equiv s[a]}\right] + o_{A\to+\infty}.\end{aligned}$$

Or les deux quantités dans les espérances sont nulles avec probabilité $1-\mathcal{O}\left(\frac{1}{p}\right)$. En dehors de cet événement elles sont des $\mathcal{O}(1)$. Ainsi on obtient que pour tout $s\in\{0,\dots,a-1\}$:

$$Q(s+1)=Q(s)+\mathcal{O}\left(\frac{1}{p}\right).$$

Ainsi $Q(0)=\frac{HX}{ap}+\mathcal{O}\left(\frac{a}{p}\right)$. Donc en sommant sur $p\in\mathcal{P}_H$, on obtient :

$$\mathbf{E}\left[\sum_{p\in\mathcal{P}_H}\sum_{j=1}^{H}c_p\mathbb{1}_{\mathbf{n}+j\equiv pb[ap]}g_1(\mathbf{n}+j)g_2(\mathbf{n}+j+ph)\mathbb{1}_{n\equiv 0[a]}\right]=\left(\frac{HX}{a}+\mathcal{O}(a)\right)\sum_{p\in\mathcal{P}_H}\frac{1}{p}.$$

Donc d'après le deuxième théorème de Mertens (I.15) [3] :

$$\mathbf{E}\left[\sum_{p\in\mathcal{P}_H}\sum_{j=1}^{H}c_p\mathbb{1}_{\mathbf{n}+j\equiv pb[ap]}g_1(\mathbf{n}+j)g_2(\mathbf{n}+j+ph)\mathbb{1}_{n\equiv 0[a]}\right]\gg\frac{\varepsilon H}{a\log(H)}.$$

Ainsi d'après le lemme III.12 :

$$\mathbf{E}\left[\sum_{p\in\mathcal{P}_H}\sum_{j=1}^{H}c_p\mathbb{1}_{a\mathbf{n}+j\equiv pb[ap]}g_1(a\mathbf{n}+j)g_2(a\mathbf{n}+j+ph)\mathbb{1}_{n\equiv 0[a]}\right]\gg\frac{\varepsilon H}{\log(H)}.$$

Si $j+ph\notin[1,H]$ alors $j\in[1,|h|\,\varepsilon^2 H]$ ou $j\in[(1-\varepsilon^2|h|)H,H]$ et dans ce cas on a :

$$\mathbf{E}\left[\sum_{p\in\mathcal{P}_H}\sum_{\substack{j\in[1,H]\\ j+ph\notin[1,H]}}c_p\mathbb{1}_{a\mathbf{n}+j\equiv pb[ap]}g_1(a\mathbf{n}+j)g_2(a\mathbf{n}+j+ph)\mathbb{1}_{n\equiv 0[a]}\right]\ll\frac{\varepsilon^2 H}{\log(H)}.$$

On en déduit le résultat annoncé par inégalité triangulaire. □

On va maintenant effectuer une discrétisation du problème. On définit $g_{i,\varepsilon^2}(n)$ pour $i=1,2$ l'élément de $\varepsilon^2\mathbb{Z}[i]$ le plus proche de $g_i(n)$ (s'il y a égalité, on en prend un arbitrairement). Cette fonction n'est pus multiplicative mais prend au plus $\mathcal{O}_\varepsilon(1)$ valeurs, est toujours par 1 et vérifie : $g_{i,\varepsilon^2}=g_i+\mathcal{O}\left(\varepsilon^2\right)$ pour $i=1,2$. D'après la proposition III.13 et l'inégalité triangulaire :

$$\mathbf{E}\left[\sum_{p\in\mathcal{P}_H}\sum_{\substack{j:\\ j\in[1,H]\\ j+pH\in[1,H]}}c_p\mathbb{1}_{a\mathbf{n}+j\equiv pb[ap]}g_{1,\varepsilon^2}(a\mathbf{n}+j)g_{2,\varepsilon^2}(a\mathbf{n}+j+ph)\right]\gg\frac{\varepsilon H}{\log(H)}.$$

On réécrit cette inégalité sous la forme : $\mathbf{E}\left[F(\mathbf{X}_H,\mathbf{Y}_H)\right]\gg\varepsilon\dfrac{H}{\log(H)}$, où :

$$\mathbf{X}_H:=(g_{i,\varepsilon^2}(a\mathbf{n}+j))_{i=1,2;\ j=1,\dots,H}\quad\text{et}\quad\mathbf{Y}_H=\mathbf{n}\bmod P_H,$$

et où F désigne la fonction suivante :

$$F:$$

3. On a aussi utilisé l'équivalent suivant : $\log\log(x)-\log\log\left(\frac{x}{2}\right)\underset{x\to+\infty}{\sim}\sqrt{2}\frac{1}{\log(x)}$.

Notons bien ici que la fonction est bien définie car si on prend deux représentants $\mathbf{Y}_1$ et $\mathbf{Y}_2$ de la même classe dans $\mathbb{Z}/P_H\mathbb{Z}$, alors $a\mathbf{Y}_1 \equiv a\mathbf{Y}_2[ap]$ pour tout $p \in \mathcal{P}_H$ car $(P_H, a) = 1$ [4]

III.6 Point de vue entropique et conclusion

Dans cette partie on va utiliser la notion d'entropie de Shannon vu en I.1.4 pour en arriver à une absurdité et donc en déduire le théorème III.10 et donc la conjecture d'Elliott en moyenne logarithmique (III.9).

III.6.1 Argument de décrément d'entropie

Avant d'appliquer l'argument, on va d'abord voir des estimations élémentaires de l'entropie de Shannon des variables aléatoires $\mathbf{X}_H$ et $\mathbf{Y}_H$ introduites précédemment :

Proposition III.14

Soit $H \in [H_-, H_+]$. On a les propriétés suivantes :

$$0 \leqslant H(\mathbf{X}_H) \ll_\varepsilon H \quad \text{et} \quad H(\mathbf{Y}_H) \ll H.$$

Démonstration : Chaque composante de $\mathbf{X}_H$ prend $\mathcal{O}_\varepsilon(1)$ valeurs donc on a $H(\mathbf{X}_H) \ll_\varepsilon H$ d'après le lemme I.13. Pour la deuxième inégalité, on a par un développement limité :

$$\forall \ \frac{X}{\omega} \leqslant k \leqslant x : \quad \mathbf{P}(\mathbf{Y}_H = k) = \frac{1}{P_H} + o_{H\to+\infty}(1).$$

Ainsi $H(\mathbf{Y}_H) = \log(P_H) - o_{A\to+\infty}(1)$. On obtient le résultat par le théorème des nombres premiers. □

L'idée principale de l'argument de décrément d'entropie se base sur la proposition suivante :

Proposition III.15

Soient $H_- \leqslant H \leqslant kH \leqslant H_+$. On a l'estimation :

$$\frac{H(\mathbf{X}_{kH})}{kH} \leqslant \frac{H(\mathbf{X}_H)}{H} - \frac{I(\mathbf{X}_H, \mathbf{Y}_H)}{H} + \mathcal{O}\left(\frac{1}{k}\right).$$

Démonstration : On définit temporairement la variante de $\mathbf{X}_H$ suivante :

$$\forall H_- \leqslant H_1, H_2 \leqslant H_+ : \quad \mathbf{X}_{H_1,H_1+H_2} := X\left(g_{i,\varepsilon^2}(\mathbf{n}+j)\right)_{i=1,2;j=H_1+1,\dots,H_1+H_2}.$$

Soient $H_- \leqslant H_1, H_2 \leqslant H_+$. D'après le lemme III.12 :

$$H\left(\mathbf{X}_{H_1,H_1+H_2}|\overline{\mathbf{n}}^{P_H}\right) = H\left(\mathbf{X}_{H_1,H_1+H_2}|\overline{\mathbf{n}+H_1}^{P_H}\right) = H\left(\mathbf{X}_{H_2}|\overline{\mathbf{n}}^{P_H}\right) + o_{A\to+\infty}(1).$$

On en déduit par sous-additivité relative de l'entropie (I.11) :

$$H\left(\mathbf{X}_{H_1+H_2}|\mathbf{Y}_H\right) \leqslant H(\mathbf{X}_{H_1}|\mathbf{Y}_{H_1}) + H(\mathbf{X}_{H_2}|\mathbf{Y}_{H_2}) + o_{A\to+\infty}(1).$$

4. On rappelle ici qu'on a H suffisamment grand devant $\frac{1}{\varepsilon^2}$ et donc que pour tout $p \in \mathcal{P}_H$ est premier avec a.

En itérant ce procédé, on conclut que :

$$\forall k, H \text{ tels que } H_- \leqslant H \leqslant kH \leqslant H_+ : \quad H(\mathbf{X}_{kH}|\mathbf{Y}_H) \leqslant kH(\mathbf{X}_H|\mathbf{Y}_H) + o_{A\to+\infty}(1).$$

On en déduit facilement que pour de tels k, H que

$$H(\mathbf{X}_{kH}) \leqslant kH(\mathbf{X}_H) - kI(\mathbf{X}_H, \mathbf{Y}_H) + H(\mathbf{Y}_H) + o_{A\to+\infty}(1).$$

On obtient le fait annoncé par la proposition précédente. □

Cette inégalité nous dit que la présence d'information mutuelle entre $\mathbf{X}_H$ et $\mathbf{Y}_H$ amène à une diminution du taux d'entropie de $\mathbf{X}_H$, ainsi si cette information mutuelle est taux grande alors elle amènerait à une trop grande diminution du taux d'entropie et qu'elle devienne ainsi négative. Cette idée est confirmée par le théorème suivant :

Théorème III.16: (Décrément d'entropie)

Il existe $H \in [H_-, H_+]$, multiple de a tel que :

$$I(\mathbf{X}_H, \mathbf{Y}_H) \leqslant \frac{H}{\log(H)\log\log\log(H)}.$$

Démonstration : Supposons par l'absurde que pour tout $H \in [H_-, H_+]$ multiple de a :

$$I(\mathbf{X}_H, \mathbf{Y}_H) > \frac{H}{\log(H)\log\log\log(H)}.$$

Soient C_0 un nombre suffisamment grand devant H_-, J suffisamment grand devant C_0, H_-, ε. On peut supposer que H_+ est suffisamment grand en fonction de H_-, C_0, J. On définit par récurrence une famille $(H_j)_{j\in\{1,\dots,J\}}$ par $H_1 = aH_-$ et pour tout $j \geqslant 1$: $H_{j+1} = H_j \lfloor C_0 \log(H_j)\log\log\log(H_j)\rfloor$ et avec J l'entier maximal tel que $H_J \in [H_-, H_+]$. D'après la proposition précédente appliquée à $(H, k) = (H_j, \lfloor C_0 \log(H_j)\log\log\log(H_j)\rfloor)$ pour tout $j \in \{1, \dots, J\}$:

$$\frac{H(\mathbf{X}_{H_{j+1}})}{H_{j+1}} \leqslant \frac{H(\mathbf{X}_{H_j})}{H_j} - \frac{I(\mathbf{X}_{H_j}, \mathbf{Y}_{H_j})}{H_j} + \mathcal{O}\left(\frac{1}{\lfloor C_0 \log(H_j)\log\log\log(H_j)\rfloor}\right).$$

Donc par la proposition I.13 :

$$\frac{H(\mathbf{X}_{H_{j+1}})}{H_{j+1}} \leqslant \frac{H(\mathbf{X}_{H_j})}{H_j} - \frac{1}{2\log(H_j)\log\log\log(H_j)}.$$

Montrons que :

$$\text{Il existe } B \text{ (dépendant de } C_0, H_-) \text{ tel que pour tout } j \geqslant 2 : H_j \leqslant e^{Bj\log(j)}.$$

On pose pour tout j : $M_j = Bj\log(j)$ et $R_j = \log(H_j)$. On a :

$$R_{j+1} \leqslant R_j + \log(C_0) + \log(R_j) + \log\log\log(R_j).$$

Ainsi pour $j \geqslant j_0$ assez grand devant C_0, H_- : $R_{j+1} \leqslant R_j + 2\log(R_j)$. Soit B tel que pour tout $j \leqslant j_0$: $H_j \leqslant e^{Bj\log(j)}$. Procédons maintenant par récurrence, par le choix de B le fait que l'on veut montrer est

vrai pour $j \leqslant j_0$. Supposons celui-ci vrai pour $j \in \mathbb{N}$ et montrons le fait pour $j+1$. Par choix de j_0 et par hypothèse de récurrence :

$$R_{j+1} \leqslant M_j + 2\log(M_j) \leqslant M_j + \frac{M_j}{j},$$

où la deuxième inégalité est obtenue par l'inégalité vraie pour tout $x, y \geqslant 4 : 4x + 2y \leqslant e^y x$. On obtient ainsi $R_{j+1} \leqslant M_j + \frac{M_j}{j} \leqslant M_{j+1}$. D'où l'hérédité. On obtient donc l'inégalité par récurrence. Ainsi :

$$\frac{H(\mathbf{X}_{H_{j+1}})}{H_{j+1}} \leqslant \frac{H(\mathbf{X}_{H_j})}{H_j} - \frac{1}{2Bj\log(j)\log\log(Bj\log(j))}.$$

Donc par la proposition III.15 et par télescopage on obtient :

$$\sum_{j=2}^{J-1} \frac{1}{2Bj\log(j)\log\log(Bj\log(j))} \ll_\varepsilon 1.$$

Or la somme de gauche diverge pour $J \to +\infty$, on obtient donc une contradiction lorsque J est suffisamment grand. □

Ce résultat utilise ce qu'on appelle un argument de décrément d'entropie, en fait l'idée est de montrer que le rapport d'entropie $\frac{H(\mathbf{X}_H)}{H}$ décroît quand $H \to +\infty$. Le fait de supposer que l'information $I(\mathbf{X}_H, \mathbf{Y}_H)$ est suffisamment grande pour un H amène au fait que le rapport $\frac{H(\mathbf{X}_H)}{H}$ va décroître trop vite et ainsi devenir négatif pour un certain H. Désormais, on considéra le H amené par ce théorème.

III.6.2 Quelques conséquences

Définition III.17

Dans cette section et uniquement dans cette section, on dira que x est une ***bonne*** *valeur si :*

$$H(\mathbf{Y}_H) - H(\mathbf{Y}_H|\mathbf{X}_H = x) = o_{H_-\to+\infty}\left(\frac{H}{\log(H)}\right).$$

Le premier intérêt à considérer les bonnes valeurs est qu'elles représentes presque toutes les valeurs prises par $\mathbf{X}_H$:

Lemme III.18

La probabilité que $\mathbf{X}_H$ prenne une bonne valeur est $1 - o_{H_-\to+\infty}(1)$.

Démonstration : On a par définition d'une bonne valeur on a pour tout $\varepsilon > 0$ fixé :

$$\sum_x \mathbf{P}(\mathbf{X}_H = x)\left(H(\mathbf{Y}_H) - H(\mathbf{Y}_H \mid \mathbf{X}_H = x)\right)$$
$$\geqslant \sum_{x \text{bonne valeur}} \mathbf{P}(\mathbf{X}_H = x)\left(H(\mathbf{Y}_H) - H(\mathbf{Y}_H \mid \mathbf{X}_H = x)\right) + \varepsilon\frac{H}{\log(H)}.$$

Or par le théorème III.16 : $I(\mathbf{X}_H, \mathbf{Y}_H) = o_{H_- \to +\infty}\left(\frac{H}{\log(H)}\right)$. Ainsi :

$$\sum_x \mathbf{P}(\mathbf{X}_H = x)\left(H(\mathbf{Y}_H) - H(\mathbf{Y}_H \mid \mathbf{X}_H = x)\right) = o_{H_- \to +\infty}\left(\frac{H}{\log(H)}\right).$$

On en déduit le résultat annoncé □

De manière intuitive, si x est bon alors $\mathbf{Y}_H$ reste presque uniformément distribuée sur $\mathbb{Z}/P_H\mathbb{Z}$ (comme on l'a vu dans la preuve de la proposition III.15) même après conditionnement par rapport à l'évènement $\mathbf{X}_H = x$. Plus précisément on a :

Lemme III.19

Soit x une bonne valeur. Soit E_x une partie de $\mathbb{Z}/P_H\mathbb{Z}$ de cardinal au plus $\exp\left(-\varepsilon^7 \frac{H}{\log(H)}\right) P_H$, alors on a :

$$\mathbf{P}(\mathbf{Y}_H \in E_x \mid \mathbf{X}_H = x) = o_{H_- \to +\infty}(1).$$

Démonstration : Par définition de l'entropie conditionnelle, on a :

$$H(\mathbf{Y}_H|(\mathbf{X}_H = x, \mathbb{1}_{E_x}(\mathbf{Y}_H))) \geqslant H(\mathbf{Y}_H|\mathbf{X}_H = x) - H(\mathbb{1}_{E_x}(\mathbf{Y}_H)|\mathbf{X}_H = x).$$

Ainsi comme x est une bonne valeur on en déduit que :

$$\begin{aligned}\mathbf{P}(\mathbf{Y}_H \in E_x|\mathbf{X}_H = x)H(\mathbf{Y}_H|(\mathbf{X}_H = x, \mathbf{Y}_H \in E_x)) + \mathbf{P}(\mathbf{Y}_H \notin E_x|\mathbf{X}_H = x)H(\mathbf{Y}_H|(\mathbf{X}_H = x, \mathbf{Y}_H \notin E_x))\\ \geqslant H(\mathbf{Y}_H) - H(\mathbb{1}_{E_x}(\mathbf{Y}_H)|\mathbf{X}_H = x) - o_{H_- \to +\infty}\left(\frac{H}{\log(H)}\right).\end{aligned}$$

Or $\mathbb{1}_{E_x}(\mathbf{Y}_H)|\mathbf{X}_H = x$ ne prend que deux valeurs possibles, ainsi par la proposition I.13 :

$$H(\mathbb{1}_{E_x}(\mathbf{Y}_H)|\mathbf{X}_H = x) = o_{H_- \to +\infty}\left(\frac{H}{\log(H)}\right).$$

Et d'après la proposition I.11 :

$$H(\mathbf{Y}_H|(\mathbf{X}_H = x, \mathbf{Y}_H \notin E_x)) \leqslant H(\mathbf{Y}_H).$$

On en déduit que :

$$\mathbf{P}(\mathbf{Y}_H \in E_x|\mathbf{X}_H = x)\left(H(\mathbf{Y}_H) - H(\mathbf{Y}_H|(\mathbf{X}_H = x, \mathbf{Y}_H \in E_x))\right) \leqslant o_{H_- \to +\infty}\left(\frac{H}{\log(H)}\right).$$

Or par hypothèse $\mathbf{Y}_H|\mathbf{X}_H = x, \mathbf{Y}_H \in E_x$ prend au maximum $|E_x|$ valeurs donc par la proposition I.13 :

$$H(\mathbf{Y}_H|\mathbf{X}_H = x, \mathbf{Y}_H \in E_x) \leqslant \log(|E_x|) \leqslant \log(P_H) - \varepsilon^7 \frac{H}{\log(H)}.$$

Or d'après la proposition

Proposition III.20

Soient a, H comme au dessus. On note pour $\alpha \in \mathbb{R}/\mathbb{Z}$:

$$S_H(\alpha) = \sum_{p \in \mathcal{P}_H} \frac{c_p(\alpha p)}{p},$$

et notons :

$$\Xi_H = \left\{ \xi \in \mathbb{Z}/H\mathbb{Z} \text{ tel qu'il existe } \eta \in \mathbb{Z}/a\mathbb{Z} \text{ tel que : } \left| S_H\left(-\frac{(b+h)\eta}{a} - \frac{h \cdot \xi}{H} \right) \right| \geqslant \frac{\varepsilon^2}{\log(H)} \right\}.$$

On introduit $(\mathbf{X}_{i,j})_{i=1,2,j=1,\dots,H}$ une famille d'éléments dans $\mathbf{S}^1$. On a :

$$\sum_{p \in \mathcal{P}_H} \frac{c_p}{p} \sum_{j : j, j+ph \in [1,H]} \mathbb{1}_{j \equiv pb[a]} \mathbf{X}_{1,j} \mathbf{X}_{2,j+ph} \ll_{x,h} \frac{H}{\log(H)} \left(\varepsilon^2 + \sum_{\xi \in \Xi_H} \frac{1}{H} \left| \sum_{j=1}^{H} \mathbf{X}_{1,j} e\left(-\frac{j\varepsilon}{H} \right) \right| \right).$$

Démonstration : On étend $\mathbf{X}_{1,j}$ et $\mathbf{X}_{2,j}$ de manière périodique (*i.e* $\mathbf{X}_{1,H+j} := \mathbf{X}_{1,j}$ pour tout j). Si on retire de la somme de gauche la contraire $j + ph \in [1, H]$ cela occasionnera une erreur de :

$$\mathcal{O}\left(\sum_{p \in \mathcal{P}} \frac{1}{p} |ph| \right) = \mathcal{O}_h\left(\varepsilon^2 \frac{H}{\log(H)} \right),$$

ce qui est acceptable par rapport à ce que l'on veut montrer. Maintenant en re-faisant la confusion d'un élément de $[1, H]$ par sa projection dans $\mathbb{Z}/H\mathbb{Z}$, on peut dire que la quantité que l'on cherche à estimer est :

$$\sum_{p \in \mathcal{P}_H} \frac{c_p}{p} \sum_{j \mathbb{Z}/H\mathbb{Z}} \mathbb{1}_{j \equiv pb[a]} \mathbf{X}_{1,j} \mathbf{X}_{2,j+ph}.$$

Par inversion de Fourier, on écrit pour $i = 1, 2; j \in \mathbb{Z}/H\mathbb{Z}$:

$$\mathbf{X}_{i,j} = \sum_{\xi \mathbb{Z}/H\mathbb{Z}} G_i(\xi) e\left(\frac{j\xi}{H} \right) \quad \text{où pour tout } \xi \in \mathbb{Z}/H\mathbb{Z} : \quad G_i(\xi) = \frac{1}{H} \sum_{j \in \mathbb{Z}/H\mathbb{Z}} \mathbf{X}_{i,j} e\left(-\frac{j\xi}{H} \right).$$

Des manipulations élémentaires permettent d'écrire :

$$\sum_{p \in \mathcal{P}_H} \frac{c_p}{p} \sum_{j \mathbb{Z}/H\mathbb{Z}} \mathbb{1}_{j \equiv pb[a]} \mathbf{X}_{1,j} \mathbf{X}_{2,j+ph} = \frac{H}{a} \sum_{\eta \in \mathbb{Z}/a\mathbb{Z}} \sum_{\xi \in \mathbb{Z}/H\mathbb{Z}} G_1(\xi) G_2\left(-\xi - \frac{H}{a}\eta \right) S_H\left(-\frac{(b+h)\eta}{a} - \frac{h \cdot \xi}{H} \right).$$

Or par l'inégalité de Cauchy-Scwharz et par l'identité de Plancherel on a :

$$\sum_{\xi \in \mathbb{Z}/H\mathbb{Z}} \left| G_1(\xi) G_2\left(-\xi - \frac{H}{a}\eta \right) \right| \ll 1.$$

De plus, on a par le deuxième théorème de Mertens (I.15) : $S_H\left(-\frac{(b+h)\eta}{a} - \frac{h \cdot \xi}{H} \right) \ll \frac{1}{\log(H)}$. On obtient

donc par définition de Ξ_H et en majorant trivialement pour $\xi \in \Xi_H$, $\left|G_2\left(-\xi - \frac{H}{a}\eta\right)\right| \ll 1$:

$$\sum_{p\in\mathcal{P}_H} \frac{c_p}{p} \sum_{j\mathbb{Z}/H\mathbb{Z}} \mathbb{1}_{j\equiv pb[a]} \mathbf{X}_{1,j}\mathbf{X}_{2,j+ph} \ll \frac{H}{a} \sum_{\eta\in\mathbb{Z}/a\mathbb{Z}} \sum_{\xi\in\Xi_H} \frac{|G_1(\xi)|}{\log(H)} + \frac{H}{a} \sum_{\eta\in\mathbb{Z}/a\mathbb{Z}} \frac{\varepsilon^2}{\log(H)}$$
$$\ll \frac{H}{\log(H)} \left(\varepsilon^2 + \sum_{\xi\in\Xi_H} |G_1(\xi)|\right).$$

On en déduit le résultat annoncé par définition de G_1. □

Lemme III.21

Pour les mêmes quantités introduites que précédemment, on a : $|\Xi_H| \ll_{a,h,\varepsilon} 1$.

Démonstration : Des manipulations élémentaires nous permettent d'obtenir :

$$\sum_{k\in\mathbb{Z}/a\mathbb{Z}} \left|S_H\left(\frac{k}{aH}\right)\right|^4 = aH \sum_{\substack{p_1,p_2,p_3,p_4\in\mathcal{P}_H \\ p_1+p_2-p_3-p_4=0}} \frac{c_{p_1}c_{p_2}\overline{c_{p_3}c_{p_4}}}{p_1p_2p_3p_4} \ll_a \frac{1}{H^3} \sum_{\substack{p_1,p_2,p_3,p_4\in\mathcal{P}_H \\ p_1+p_2-p_3-p_4=0}} 1.$$

Et l'estimation suivante provenant d'un crible de Selberg nous permet de conclure

$$\sum_{k\in\mathbb{Z}/a\mathbb{Z}} \left|S_H\left(\frac{k}{aH}\right)\right|^4 \ll_{a,\varepsilon} \frac{1}{\log(H)^4} \ll_{a,\varepsilon,h} 1.$$

□

Conclusion : Essayons maintenant de conclure à une absurdité. Grâce aux deux lemmes précédents, on a

$$\sum_{\xi\in\Xi_H} \mathbf{E}\left[\frac{1}{H}\left|\sum_{j=1}^{H} g_1(a\mathbf{n}+j)e\left(-\frac{j\xi}{H}\right)\right|\right] \gg_{a,h} \varepsilon.$$

On en déduit d'après III.2 que $\varepsilon \ll_{a,h} o_{H_-\to+\infty}(|\Xi_H|)$. Avec le lemme précédent, on conclut à une absurdité.

IV Problème de la discrépance d'Erdös

IV.1 Présentation du problème

Commençons par nous mettre en situation. Imaginons que nous devons guider par message quelqu'un qui se trouve sur la voie du milieu d'un pont à trois voies. Notre objectif est d'avancer sur ce pont sans jamais tomber, tout en respectant une règle essentielle : il est interdit d'avancer en ligne droite. Pour mieux visualiser la situation, nous pouvons subdiviser le pont comme suit :

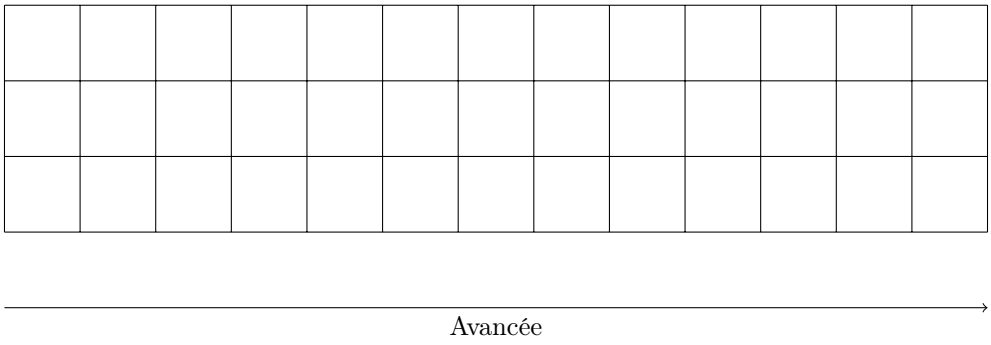


FIGURE 2 – Schématisation du pont

Une solution très simple pour progresser sur le pont sans tomber consiste à avancer en zigzag, comme illustré dans le schéma ci-dessous :

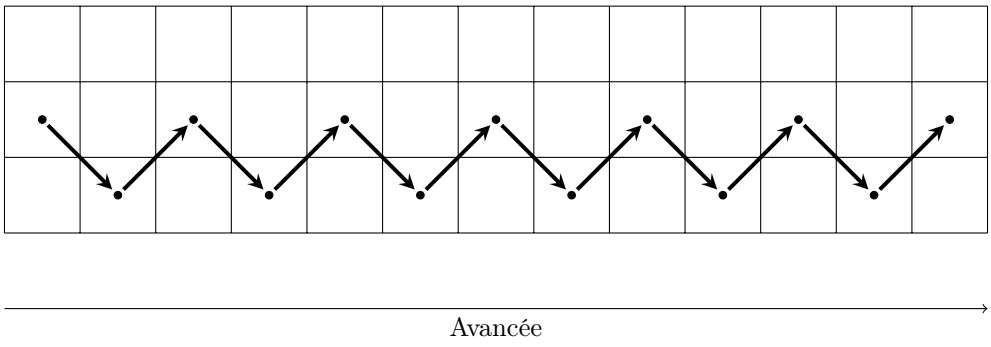


FIGURE 3 – Schématisation du déplacement en zigzag.

Nous représentons ce déplacement par la séquence suivante : $DGDGDGDG\ldots$, que l'on lit de gauche à droite. La lettre $'D'$ indique un déplacement en diagonale vers la droite, tandis que la lettre $'G'$ indique un déplacement en diagonale vers la gauche. C'est ainsi que nous communiquons ces instructions à la personne se trouvant sur le pont.
Maintenant, supposons qu'il soit difficile de communiquer directement avec cette personne, ce qui fait qu'elle pourrait ne comprendre qu'une lettre sur deux du message. Il est donc essentiel de s'assurer que si elle ne reçoit qu'une lettre sur deux, le message lui permettra tout de même de rester sur le pont sans jamais tomber.

Prenons l'exemple du message précédent et considérons uniquement une lettre sur deux. Nous obtenons alors le message suivant : $GGGGGG\ldots$. Sur le schéma ci-contre cela revient à dire qu'en suivant uniquement les flèches rouges, la personne tombera du pont.

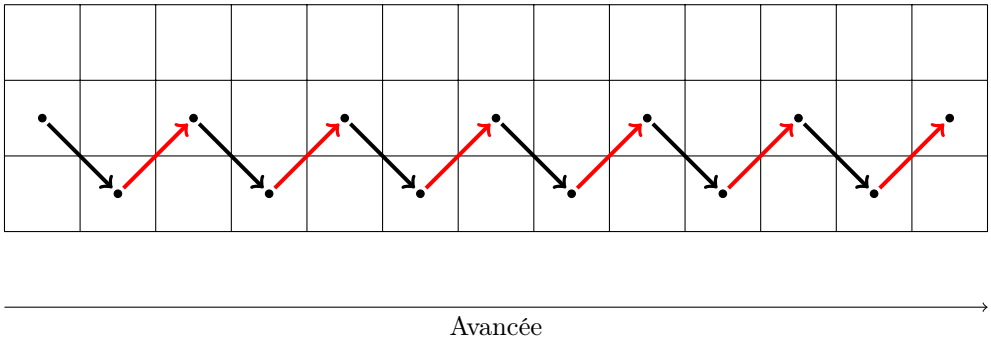


FIGURE 4 – Schéma illustrant que le déplacement "trivial" sur le pont n'est pas adapté si on ne considère qu'un déplacement sur deux.

Par conséquent, il est nécessaire de réfléchir plus en profondeur. Nous allons donc procéder étape par étape :

— Le premier déplacement est sans importance, nous choisissons donc D.

— Pour rester sur le pont, nous sommes obligés de choisir G pour le deuxième déplacement.

— Si nous choisissons D pour le troisième déplacement, en prenant un déplacement sur deux, le chemin devient DD, ce qui fera tomber la personne du pont. Nous sommes donc contraints de choisir G pour le troisième déplacement.

— Pour rester sur le pont, nous devons choisir D pour le quatrième déplacement.

— Nous continuons à prolonger ce chemin de manière périodique.

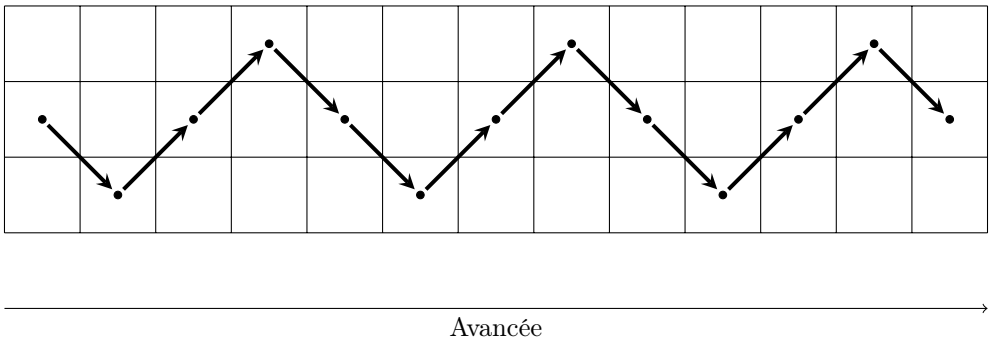


FIGURE 5 – Schématisation d'un chemin qui permet à la personne de rester sur le pont même un déplacement sur deux.

Ainsi, nous obtenons le chemin $DGGDDGGDDGGDDGG\ldots$, qui permettra à la personne de rester sur le pont s'il est transmis intact. Si le chemin est transmis en prenant une lettre sur deux, il devient $DGDGDGDG\ldots$, qui est évidemment un chemin permettant de rester sur le pont (voir figure 6).

Nous allons compliquer davantage le processus en supposant que la personne sur le pont ne peut comprendre aussi qu'un déplacement sur trois du message. Il est donc crucial de s'assurer que si nous considérons un déplacement sur trois, le chemin obtenu permettra à la personne de rester sur le pont. En examinant le message précédemment obtenu, le chemin résultant en prenant un déplacement sur trois commence par GG, ce qui fera tomber la personne du pont.

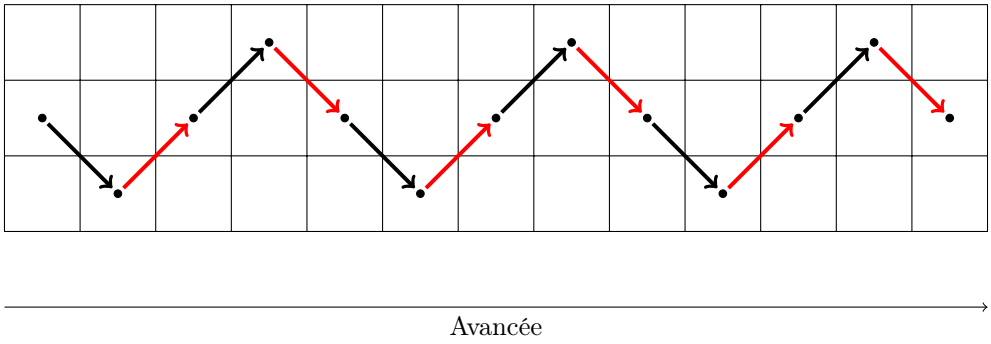


FIGURE 6 – Vérification qu'en considérant un déplacement sur deux, la personne sur deux restera sur le pont.

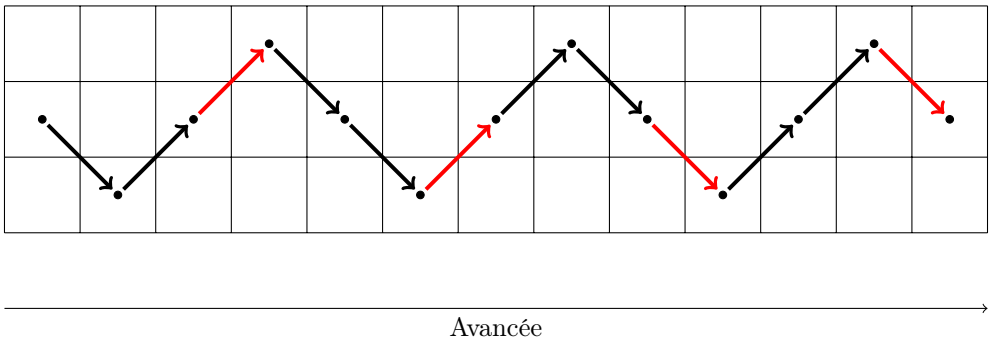


FIGURE 7 – En prenant le même chemin et en ne considérant qu'un déplacement sur trois alors la personne tombera du pont.

Nous devons donc trouver un nouveau chemin en poursuivant le même procédé.

— Nous sommes contraints de reprendre les quatre premiers déplacements identiques.

— Si nous choisissons D pour le cinquième déplacement, nous serons contraints de choisir G pour le sixième, ce qui entraînera le problème décrit précédemment. Nous devons donc choisir G pour le cinquième déplacement et D pour le sixième.

— Le chemin doit donc commencer par $DGGDGD$. Ensuite, si nous choisissons G pour le septième déplacement, le huitième sera D et le chemin deviendra $DGGDGDGD$, ce qui ne respecte pas les conditions si l'on considère un déplacement sur deux. Nous devons donc prendre D comme septième déplacement et G comme huitième déplacement.

— Le chemin doit donc commencer par $DGGDGDDG$. Si nous choisissons D comme neuvième déplacement, le dixième sera forcément G, et comme indiqué précédemment, le onzième sera G et le douzième sera D (sinon le chemin ne fonctionnerait pas si nous ne considérons qu'un déplacement sur deux). Nous obtenons donc le chemin qui commence par $DGGDGDDGDGGD$. En considérant un déplacement sur trois, nous obtenons $GDDD$, ce qui fera tomber la personne du pont dans ce cas.

— Le chemin doit donc commencer par $DGGDGDDGGD$. Cependant, si nous continuons avec un déplacement sur trois en choisissant D comme neuvième déplacement, le dixième sera inévitablement G, et comme mentionné précédemment, le onzième sera également G et le douzième sera D (sinon le chemin ne fonctionnerait pas si nous ne considérons qu'un déplacement sur deux). Cela nous donnerait un chemin commençant par $DGGDGDDGDGGD$. En prenant un déplacement sur trois, nous obtenons $GDGG$, ce qui fera également tomber la personne du pont dans ce cas. Ainsi,

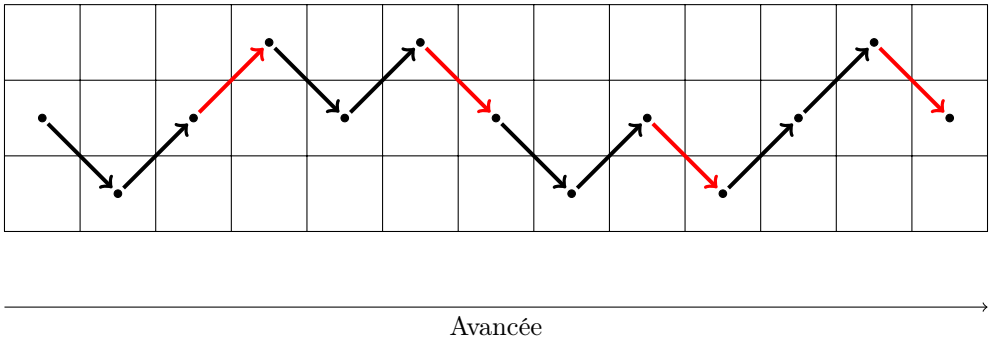


FIGURE 8 – Premier cas non concluant

nous pouvons conclure qu'il n'existe aucun chemin satisfaisant à nos conditions avec une longueur minimale de 12 déplacements. En réalité, la situation est symétrique lorsque nous sommes à l'étape 8, et il y a autant de déplacements à droite qu'à gauche, que nous prenions tous les déplacements, un déplacement sur deux ou un déplacement sur trois. Par conséquent, nous aurions pu éviter de considérer ce cas. Ce qui pose problème ici, c'est que nous sommes contraints de raisonner par groupes de quatre afin d'éviter d'avoir deux fois le même déplacement en prenant un déplacement sur deux. Cependant, dans le quadruplet d'éléments supérieurs à huit, il y a deux multiples de trois qui se retrouveront nécessairement dans la même direction si nous raisonnons ainsi. Cela conduit à l'impossibilité de former un chemin d'une longueur supérieure ou égale à douze qui respecte les conditions énoncées.

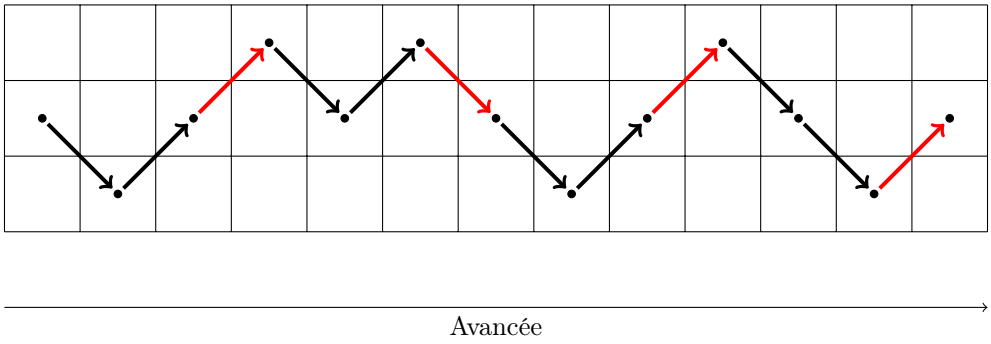


FIGURE 9 – Deuxième cas aussi non concluant : il est impossible de construire un chemin de longueur au moins 12 vérfiant les conditions demandées

—- On pourrait se demander comment résoudre le problème lorsque le pont est de taille plus grande et en considérant également toutes les progressions arithmétiques homogènes [5]. Cette question a été soulevée par Erdos au début des années 1930, et il a émis l'hypothèse que la réponse serait également négative dans le cas général, c'est-à-dire qu'il n'existerait pas de tels chemins. Pour formaliser davantage la situation, nous pouvons considérer le n-ème déplacement comme un élément $f(n) \in -1, +1$, qui représente les directions droite et gauche. Ainsi, la position de la personne sur le pont après le n-ème déplacement peut être exprimée par $\sum_{k=1}^{n} f(n)$. On peut dire que la personne tombera du pont si la valeur absolue de cette quantité devient trop grande le long de toute progression arithmétique homogène, on appelle cette quantité la discrépance de la $(f(n))$ [6] :

5. C'est-à-dire qu'il faudrait également tenir compte des suites de déplacements sur quatre, cinq, etc.
6. Le terme "discrépance" est peu fréquent en français ; en fait, il n'est même pas répertorié dans les dictionnaires

Définition IV.1

Soit $(f(n))$ une suite de réels, on appelle ***discrépance*** *de la suite $(f(n))$ la quantité* $\sup_{d,n\in\mathbb{N}^*} \sum_{j=1}^{n} f(jd)$.

La conjecture d'Erdos peut finalement être formulée de la manière suivante :

Conjecture IV.2: (Conjecture d'Erdos)

Pour toute suite d'éléments $(f(n))_{n\in\mathbb{N}^*}$ et pour tout $C \geqslant 0$, on a :

$$\sup_{n,d\in\mathbb{N}^*} \left|\sum_{k=1}^{n} f(dj)\right| \geqslant C.$$

On a montré précédemment que cette conjecture était vrai si $C = 2$. On peut même énoncer cette conjecture dans un cadre encore plus général.

Conjecture IV.3: (Conjecture d'Erdos générale)

Soit $(H, \|\cdot\|_H)$ un espace vectoriel normé. Pour toute suite d'éléments $(f(n))_{n\in\mathbb{N}^*}$ de norme 1 et pour tout $C \geqslant 0$, on a :

$$\sup_{n,d\in\mathbb{N}^*} \left\|\sum_{k=1}^{n} f(dj)\right\|_H \geqslant C.$$

C'est cette version que l'on va démontrer dans la suite de cette partie.

IV.2 Quelques remarques sur le problème

Avant de passer à des éléments de démonstration, il est intéressant de s'intéresser un peu à ce qu'il y a autour. Cette démonstration a été établie par Terence Tao en 2015 [**taoErdos**] alors que le problème a été posé plus de 80 ans auparavant. Roth a montré en 1964 dans [6] que le résultat est vrai si l'on considère la borne supérieure sur toutes les progressions arithmétiques (et non seulement les progressions arithmétiques homogènes comme dans la conjecture d'Erdos). En 2010, le projet PolyMath5 avait pour projet d'arriver à bout de cette conjecture, mais sans réel succès (malgré l'efficacité des différents projets PolyMath) même si on retrouve quelques éléments mis au point pendant ce projet dans la démonstration de Tao. Les seuls réels succès avant la preuve de Tao en toute généralité étaient des preuves par ordinateurs dans les cas où $C = 3$ et $C = 4$ [7] de 2014 seulement !

Essayons de se rendre compte de la difficulté du problème en cherchant des presque contre-exemples lorsque l'on modifie très peu les hypothèses.

Soit $\chi : \mathbb{N} \to \mathbb{C}$ un caractère de Dirichlet modulo q non principal. Par définition, χ est complètement multiplicatif et a pour moyenne 0 sur tout intervalle de longueur q, ainsi on a :

$$\forall j, n \in \mathbb{N}^* : \quad \left|\sum_{j=1}^{n} \chi(jd)\right| = |\chi(d)| \left|\sum_{j=1}^{n} \chi(j)\right| \leqslant q.$$

Ceci ne constitue évidemment pas un contre-exemple à la discrépance d'Erdos puisque que $\chi(n) = 0$ quand n et q ne sont pas premiers entre eux.

traditionnels. Il dérive du latin "discrepantia" et peut être synonyme de "dissonance". Au sens figuré, il peut également signifier "divergence, dissemblance, discordance".

Maintenant ce que l'on peut faire c'est d'adapter un peu le contre-exemple précédent pour qu'il rentre dans les hypothèses du problème de la discrépance d'Erdos. On prend χ_3 le caractère de Dirichlet vérifiant $\chi_3(n) = +1$ ni $n \equiv 1[3]$, 0 si $n \equiv 0[3]$ et -1 si $n \equiv -1[3]$.

Soit H un espace de Hilbert de base orthonormale $e_0, e_1, \ldots$ et soit $f : \mathbb{N} \to H$ définie par $f(3^a m) = \chi_3(m) e_a$ quand $a = 0, 1, 2, \ldots$ et $(m, 3) = 1$, f prend des valeurs dans $S_H(0, 1)$. Prenons $n = 1+3+\ldots+3^k$ pour k assez grand. On remarque que :

$$\sum_{j=1}^{n} f(j) = e_0 + \ldots + e_k.$$

Donc :

$$\left\| \sum_{j=1}^{n} f(j) \right\|_H = \sqrt{k+1} \gg \sqrt{\log(n)}.$$

De plus, on remarque par le théorème de Pythagore :

$$\left\| \sum_{j=1}^{n} f(jd) \right\|_H \ll \sqrt{\log(n)}.$$

Ainsi en modifiant un peu le contre-exemple précédant on obtient une discrépance d'ordre $\sqrt{\log(N)}$.

On n'a pas trouvé de fonctions arithmétiques qui admettent des discrépances beaucoup plus petites. D'ailleurs, cela peut nous donner une estimation (pas très fondée certes) du nombre de la taille d'une séquence $(f(n))$ de discrépance fixée. En effet les preuves par ordinateurs ont permis de montrer qu'il n'existe pas de chemin de longueur strictement plus grande que 1160 de discrépance d'au moins 3 et qu'il n'existe pas de chemin de longueur strictement plus grande que 127645 (!) de discrépance d'au moins 4. Et on remarque que $\sqrt{\log(1160)} \simeq 2.66 \in [2, 3]$ et $\sqrt{\log(127645)} \simeq 3.43 \in [3, 4]$. Ainsi on peut imaginer que la plus longue suite qui admettrait une discrépance C aurait environ $\exp\left(C^2\right)$ termes. Ainsi si on veut construire une très grande suite de discrépance 14, il faudrait sûrement s'aventurer à construire une suite possédant $\exp\left(14^2\right) \simeq 1.10^{85}$ ce qui environ autant de particules dans l'univers observable. Bref cela montre à quel point les chances pour obtenir une preuve exhaustive de la conjecture d'Erdos par ordinateur est vouée à l'échec avec les techniques les plus sophistiquées trouvées en 2014. C'est ici que la théorie analytique et probabiliste des nombres vient à la rescousse pour démontrer la conjecture d'Erdos, on utilisera dans la démonstration de cette conjecture la conjecture d'Elliott en moyenne logarithmique qui est une conséquence des travaux de Matomäki et Radziwiłł, les résultats qu'ils ont démontrés forment le point de départ de la démonstration.

IV.3 Forme équivalente du problème

Le but de cette partie est de montrer que les théorèmes suivants sont des formes équivalentes de la conjecture d'Erdos (IV.2).

Théorème IV.4: (Forme probabiliste de la conjecture d'Erdos)

Soit $\mathbf{g} : \mathbb{N} \to \mathbf{S}^1$ une fonction stochastique complètement multiplicative. Alors :

$$\sup_{n\in\mathbb{N}} \mathbf{E}\left[\left(\sum_{j=1}^{n} \mathbf{g}(j)\right)^2\right] = +\infty.$$

Théorème IV.5: (Formulation en théorie de la mesure de la conjecture d'Erdos)

Soient (Ω, μ) un espace probabilisé, $g : \Omega \to \left(\mathbf{S}^1\right)^{\mathbb{N}}$ une fonction mesurable telle que pour μ-presque tout $\omega \in \Omega$, $g(\omega)$ est complètement multiplicative, alors :

$$\sup_{n\in\mathbb{N}} \int_\Omega \left|\sum_{j=1}^{n} g(\omega)(j)\right|^2 \mathrm{d}\mu(\omega) = +\infty.$$

Ces deux formulations sont très clairement équivalentes.

Proposition IV.6

Le théorème IV.3 implique le théorème IV.5

Démonstration : Soient (Ω, μ) et g comme dans le théorème IV.5, on note H l'espace de Hilbert sont $L^2(\Omega, \mu)$. Pour tout $n \in \mathbb{N}$, on pose $f(n) \in H$ l'application $\omega \mapsto g(\omega)(n)$. Comme $g(\omega)(n) \in \mathbf{S}^1$ pour μ-presque tout $\omega \in \Omega$ et que μ est une mesure de probabilité on a pour tout $n \in \mathbb{N}$: $\|f\|_H = 1$. De plus, comme g est μ-presque partout complètement multiplicative, pour toute progression arithmétique homogène de raison d on a :

$$\left\|\sum_{j=1}^{n} f(jd)\right\|_H^2 = \int_\Omega \left|\sum_{j=1}^{n} g(\omega)(jd)\right|^2 \mathrm{d}\mu(\omega) = \int_\Omega \left|\sum_{j=1}^{n} g(\omega)(j)\right|^2 \mathrm{d}\mu(\omega).$$

Ainsi si on suppose le théorème IV.3 vrai et en prenant la borne supérieure sur n dans la relation précédente, on obtient :

$$\sup_{n\in\mathbb{N}} \int_\Omega \left|\sum_{j=1}^{n} g(\omega)(j)\right|^2 \mathrm{d}\mu(\omega) = \sup_{n\in\mathbb{N}} \left\|\sum_{j=1}^{n} f(jd)\right\|_H^2 = +\infty.$$

On obtient alors dans ce cas que le théorème IV.5 est vrai. □

L'implication réciproque est bien plus difficile, on aura besoin d'un résultat d'analyse fonctionnelle qui nécessite la définition suivante :

Définition IV.7

Soit X un espace métrique et soient $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesures de probabilités sur X. On dit que la suite $(\mu_n)_{n\in\mathbb{N}}$ ***converge étroitement*** *vers μ lorsque pour tout fonction f continue bornée sur X, on a :*

$$\int_X f(x)\,\mathrm{d}\mu_n(x) \underset{n\to+\infty}{\to} \int_X f(x)\,\mathrm{d}\mu(x).$$

Lemme IV.8: (Théorème de Prokhorov - Cas compact)

Soit X un espace métrique compact et μ_n une suite de mesures de probas sur X. Alors il existe une sous-suite de μ_n qui converge étroitement vers un autre mesure de proba μ.

Démonstration : On sait que si X est un espace métrique compact alors $\mathcal{C}^0(X,\mathbb{C})$ est séparable. Donc par le théorème de Banach-Alaoglu la sphère unité de $\mathcal{M}(X)$, le dual de $\mathcal{C}^0(X,\mathbb{C})$ est séquentiellement compact pour la topologie faible-$*$.

Or pour tout $n\in\mathbb{N}$, on peut poser l'application continue suivante $T_{\mu_n} f \in \mathcal{C}^0(X,\mathbb{C}) \mapsto \int_X f\,\mathrm{d}\mu_n$.

Cette application est clairement continue de norme d'application égale à 1 (car les μ_n sont des mesures de probabilités sur X). Ainsi il existe une extractrice φ telle que $T_{\mu_{\varphi(n)}}$ converge faible-$*$ vers $T\in\mathcal{M}(X)$ de norme égale à 1. Or par le théorème de représentation de Riesz-Radon, il existe une mesure de probabilité μ telle que $T = T_\mu$. Cela revient à dire que $(\mu_{\varphi(n)})$ converge étroitement vers μ. □

Voici un second lemme qui consiste en l'application du théorème de Prokhorov.

Lemme IV.9

Supposons que pour tout $X \geqslant 1$, il existe une fonction stochastique complètement multiplicative $\mathbf{g}_X$ [a] qui soit à valeurs dans $\mathbf{S}^1$ et telle que :

$$\forall n \leqslant X : \mathbf{E}\left[\left|\sum_{j=1}^{n} \mathbf{g}_X(j)\right|\right]^2 \ll_C 1.$$

Alors il existe une fonction $\mathbf{g}$ stochastique complètement multiplicative à valeurs dans $\mathbf{S}^1$ et telle que :

$$\forall n \in \mathbb{N} : \mathbf{E}\left[\left|\sum_{j=1}^{n} \mathbf{g}(j)\right|^2\right] \ll_C 1.$$

a. L'univers sur lequel est définie $\mathbf{g}_X$ peut dépendre de X

Démonstration : Dans cette preuve on notera $\mathcal{P}$ l'ensemble des nombres premiers. Soit $\mathcal{M}$ l'ensemble des fonctions complètement multiplicatives à valeurs dans $\mathbf{S}^1$. Comme les éléments de $\mathcal{M}$ sont déterminées uniquement par leurs valeurs sur les nombres premiers, on voit facilement que $\mathcal{M}$ est isomorphe à $\left(\mathbf{S}^1\right)^{\mathcal{P}}$. Comme produit dénombrable de compacts, $\mathcal{M}$ est compact pour la topologie produit. Or comme énoncé précédemment, on peut voir chaque $\mathbf{g}_X$ comme une application mesurable $f_X : \Omega_X \to \mathcal{M}$ telle que :

$$\forall n \leqslant X : \quad \int_{\Omega_X} \left|\sum_{j=1}^{n} f_X(\omega)(j)\right|^2 \mathrm{d}\mu_X(\omega) \ll_C 1.$$

On note ν_X la mesure image de μ_X par f_X, c'est une mesure sur $\mathcal{M}$ qui vérifie :

$$\forall F \in \mathcal{C}^0(\mathcal{M}, \mathbb{C}) : \int_{\mathcal{M}} F(g)\, \mathrm{d}\nu_X(g) = \mathbf{E}\left[F(g_X)\right] = \int_{\Omega_X} F(f_x(\omega))\, \mathrm{d}\mu_X(\omega).$$

On remarque que les fonctions $g \mapsto \left(\sum_{j=1}^{n} g(j)\right)^2$ sont continues sur $\mathcal{M}$, en effet on peut décomposer ces applications en polynômes en les $g \mapsto g(p)$ où p est premier, ces applications étant continues par définition de la topologie produit. Ainsi en appliquant la dernière estimation à ces fonctions, on obtient que :

$$\forall n \leqslant X : \int_{\mathcal{M}} \left|\sum_{j=1}^{n} g(j)\right|^2 \mathrm{d}\nu_X \ll_C 1.$$

Par le lemme de Prokhorov (IV.8), on peut trouver une sous-suite ν_{X_j} de ν_X telle que $(\nu_{X_j})_{j\in\mathbb{N}}$ converge étroitement vers une mesure ν de probabilité sur $\mathcal{M}$. Donc on conclut que :

$$\forall n \in \mathbb{N} : \int_{\mathcal{M}} \left|\sum_{j=1}^{n} g(j)\right|^2 \mathrm{d}\nu(g) \ll_C 1.$$

On définit ensuite une fonction stochastique complètement multiplicative $\mathbf{g} : \mathbb{N} \to \mathbf{S}^1$ en choisissant $(\mathcal{M}, \nu)$ comme espace de probabilité ambiant et l'identité $g \mapsto g$ comme application mesurable. On obtient alors dans cet espace probabilisé :

$$\forall n \in \mathbb{N} : \mathbf{E}\left[\left|\sum_{j=1}^{n} \mathbf{g}(j)\right|^2\right] \ll_C 1.$$

□

On peut enfin démontrer la proposition suivante :

Proposition IV.10

Le théorème IV.4 implique le théorème *IV.3*

Démonstration : Procédons par contraposée, on suppose qu'il existe $f : \mathbb{N} \to H$ à valeurs dans $S_H(0,1)$ et $C \geqslant 0$ tels que

$$\forall d \in \mathbb{N}^* : \quad \left\|\sum_{j=1}^{n} f(jd)\right\|_H \leqslant C. \tag{IV.1}$$

Pour conclure notre preuve, il suffit de construire une fonction stochastique complètement multiplicative $\mathbf{g}$ à valeurs dans $\mathbf{S}^1$ telle que :

$$\forall n \in \mathbb{N} : \mathbf{E}\left[\left|\sum_{j=1}^{n} \mathbf{g}(j)\right|^2\right] \ll_C 1.$$

D'après le lemme IV.9, il suffit de construire pour tout $X \geqslant 1$, une fonction stochastique complètement

multiplicative $\mathbf{g}_X$ qui soit à valeurs dans $\mathbf{S}^1$ et telle que :

$$\forall n \leqslant X : \mathbf{E}\left[\left|\sum_{j=1}^{n} \mathbf{g}_X(j)\right|\right]^2 \ll_C 1.$$

Soit $X \geqslant 1$, notons $p_1, \dots, p_r$ les nombres premiers inférieurs ou égaux à X rangés par ordre croissant. Soit $M \geqslant X$ un entier naturel que l'on suppose suffisamment grand devant C, X. On définit les fonctions suivantes : [7]

$$F : \begin{cases} (\mathbb{Z}/M\mathbb{Z})^r & \to H \\ (a_1[M], \dots, a_r[M]) & \mapsto f\left(p_1^{a_1} \dots p_r^{a_r}\right) \end{cases} \quad \text{et} \quad \pi : \begin{cases} [1, X] & \to (\mathbb{Z}/M\mathbb{Z})^r \\ p_1^{a_1} \dots p_r^{a_r} & \mapsto (a_1[M], \dots, a_r[M]) \end{cases}.$$

En appliquant l'équation IV.1 pour tout $n \leqslant X$ et pour tout $d \in \mathbb{N}$ de la forme $d = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ avec pour tout $i \in \{1, \dots, r\} : 1 \leqslant \alpha_i \leqslant M - X$:

$$\left\|\sum_{j=1}^{n} f(jp_1^{\alpha_1} \dots p_r^{\alpha_r})\right\|_H \ll_C 1,$$

en notant $j = p_1^{\beta_{j_1}} \dots p_r^{\beta_{j_r}}$. Or pour tout $x = (x_1, \dots, x_r) \in (\mathbb{Z}/M\mathbb{Z})^r$:

$$\left\|\sum_{j=1}^{n} F(x + \pi(j)\right\|_H = \left\|\sum_{j=1}^{n} f(p_1^{x_1+\beta_{j_1}} \dots p_r^{x_r+\beta_{j_r}})\right\|_H.$$

Il y a donc $(M-X)^{r-1}$ choix de x (correspondants aux $(M-X)^{r-1}$ choix de d possibles) qui permettent d'avoir :

$$\left\|\sum_{j=1}^{n} F(x + \pi(j))\right\|_H \ll_C 1.$$

On obtient alors :

$$\frac{1}{M^r} \sum_{x \in (\mathbb{Z}/M\mathbb{Z})^r} \left\|\sum_{j=1}^{n} F(x + \pi(j))\right\|_H^2 \ll_C 1.$$

Or, on peux écrire pour tout $x \in (\mathbb{Z}/M\mathbb{Z})^r$:

$$F(x) = \sum_{\xi \in (\mathbb{Z}/M\mathbb{Z})^r} \hat{F}(\xi) e\left(\frac{x \cdot \xi}{M}\right) \quad \text{où} \quad \hat{F}(\xi) = \frac{1}{M^r} \sum_{\omega \in (\mathbb{Z}/M\mathbb{Z})^r} F(\omega) e\left(-\frac{\omega.\xi}{M}\right).$$

Ainsi par le théorème de Plancherel :

$$\frac{1}{M^r} \sum_{x \in (\mathbb{Z}/M\mathbb{Z})^r} \left\|\sum_{j=1}^{n} F(x + \pi(j))\right\|_H^2 = \sum_{\xi \in (\mathbb{Z}/M\mathbb{Z})^r} \left\|\hat{F}(\xi)\right\|_H^2 \left|\sum_{j=1}^{n} e\left(\frac{\pi(j).\xi}{M}\right)\right|^2.$$

7. Pour s'assurer que F est bien définie on considère que les a_i sont des éléments de $\{0, \dots, M-1\}$. Notons aussi que π est bien définie pour $M \geqslant X$

De plus, aussi d'après le théorème de Plancherel :

$$\sum_{\xi\in(\mathbb{Z}/M\mathbb{Z})^r} \left\|\hat{F}(\xi)\right\|_H^2 = 1.$$

Donc on peut interpréter $\left\|\hat{F}(\xi)\right\|_H^2$ comme une densité de probabilité d'une fréquence $\xi = (\xi_1, \dots, xi_r) \in (\mathbb{Z}/M\mathbb{Z})^r$. Ainsi avec cette interprétation :

$$\forall n \leqslant X : \mathbf{E}\left[\left|\sum_{j=1}^{n} e\left(\frac{\pi(j)\cdot\xi}{M}\right)\right|\right] \ll_C 1.$$

Si on définit la fonction stochastique complètement multiplicative $\mathbf{g}_X$ définie par : $\mathbf{g}_X(p_j) = e\left(\frac{\xi_j}{M}\right)$ pour $j \in \{1, \dots, r\}$ et $g_X(p)$ pour les autres nombres premiers (il n'interviennent pas ensuite), on obtient :

$$\forall n \leqslant X : \mathbf{E}\left[\left|\sum_{j=1}^{n} \mathbf{g}_X(j)\right|^2\right] \ll_C 1.$$

D'où le résultat. □

Remarque : Un point intéressant à remarquer est de voir notre preuve ne marche plus si on considère un caractère de Dirichlet χ. Il est bien de module souvent à valeurs dans le cercle unité mais pour $q \leqslant p_r$, la fonction qui à $(a_1, \dots, a_r)$ associe $\chi\left(p_1^{a_1} \dots p_r^{a_r}\right)$ est presque toujours nulle (comme l'argument est très souvent un multiple de q). Ainsi :

$$\sum_{\xi\in(\mathbb{Z}/M\mathbb{Z})^r} \left\|\hat{F}(\xi)\right\| \ll 1.$$

On ne peut donc pas définir une fonction $\mathbf{g}_X$ telle que dans le preuve ci-dessus.

IV.4 Application de la conjecture d'Elliott logarithmique

Proposition IV.11

Supposons que $\mathbf{g} : \mathbb{N} \to \mathbf{S}^1$ soit une fonction stochastique complètement multiplicative telle qu'il existe $C > 0$:

$$\forall n \in \mathbb{N} : \mathbf{E}\left[\left|\sum_{j=1}^{n} \mathbf{g}(j)\right|^2\right] \leqslant C^2.$$

Soit $\varepsilon > 0$ et supposons que X est suffisamment grand devant ε, C. Alors avec probabilité $1 - \mathcal{O}(\varepsilon)$, il existe un caractère de Dirichlet (stochastique) de période $\mathbf{q} = \mathcal{O}_{C,\varepsilon}(1)$ et un nombre réel (stochastique) $\mathbf{t} = \mathcal{O}_{C,\varepsilon}(X)$ tels que :

$$\sum_{p\leqslant X} \frac{1 - \mathcal{R}e\left(\mathbf{g}(p)\overline{\chi(p)}p^{-it}\right)}{p} \ll_{C,\varepsilon} 1.$$

Démonstration : Soient $\mathbf{g}, C, \varepsilon$ comme introduits ci-dessus. Soit $H \geqslant 1$ un nombre modérément large devant ε à choisir plus tard. On suppose que X est suffisamment grand devant X, ε. Par hypothèse et

par inégalité triangulaire, on a pour X suffisamment grand :

$$\mathbf{E}\left[\sum_{\sqrt{X}\leqslant n\leqslant X}\frac{1}{n}\left|\sum_{j=n+1}^{n+H}\mathbf{g}(j)\right|^2\right]\ll_C \log(X).$$

Ainsi par l'inégalité de Markov, avec probabilité $1-\mathcal{O}(\varepsilon)$:

$$\sum_{h_1,h_2\in[1,H]}\sum_{\sqrt{X}\leqslant n\leqslant X}\frac{\mathbf{g}(n+h_1)\mathbf{g}(n+h_2)}{n}=\sum_{\sqrt{X}\leqslant n\leqslant X}\frac{1}{n}\left|\sum_{h=1}^{H}\mathbf{g}(n+h)\right|^2\ll_{C,\varepsilon}\log(X).$$

Or dans la somme sur h_1,h_2 il est faicle de voir que le terme diagonal contribue pour $\gg H\log(X)$. Ainsi pour H suffisamment grand devant C,ε et d'après le principe des tiroirs, on peut trouver deux entiers (stochastiques) et distincts $h_1,h_2\in[1,X]$ tels que :

$$\left|\sum_{\sqrt{X}\leqslant n\leqslant X}\frac{\mathbf{g}(n+h_1)\mathbf{g}(n+h_2)}{n}\right|\gg_{C,\varepsilon,H}\log(X).$$

Donc par réciproque de la conjecture d'Elliott en moyenne logarithmique III.9, il existe un caractère de Dirichlet stochastique de période $\mathbf{q}=\mathcal{O}_{C,\varepsilon}(1)$ et un nombre réel stochastique $\mathbf{t}=\mathcal{O}_{C,\varepsilon}(X)$ tels que :

$$\sum_{p\leqslant X}\frac{1-\mathcal{R}e\left(\mathbf{g}(p)\overline{\chi(p)}p^{-it}\right)}{p}\ll_{C,\varepsilon}1.$$

Détaillons juste comment on peut s'arranger pour que $(\chi,\mathbf{t})$ soient pris mesurables (et ainsi stochastiques). Par le théorème III.9, on sait qu'il existe plusieurs couples (χ,t) tels que :

$$\sum_{p\leqslant x}\frac{1-g(p)\overline{\chi(p)}p^{-it}}{p}<a.$$

Et s'il existe $t_0\in\mathbb{R}$ tel que l'inégalité précédente est lieu pour $t=t_0$ alors par densité de $\mathbb{Q}$ dans $\mathbb{R}$, on peut trouver $t_0^{'}\in\mathbb{Q}$ tel que cette inégalité est aussi lieu pour $t=t_0^{'}$. Ainsi on peut s'arranger pour considérer un ensemble dénombrable de (χ,t) qui vérifie l'inégalité. Comme il est dénombrable, on peut lui construire un bon ordre sur cet ensemble. Or l'application $g\mapsto\left(\sum_{p\leqslant x}\frac{1-g(p)\overline{\chi}(p)p^{-it}}{p}\right)_{(\chi,t)}$ est continue donc mesurable et l'application qui à une suite (x_n) de réels associe l'unique $n\in\mathbb{N}$ tel que $x_n\leqslant a$ et pour tout $m\leqslant n: x_n>a$ est aussi mesurable[8]. Ainsi on peut trouver $(\chi,\mathbf{t})$ qui vérifie :

$$\sum_{p\leqslant X}\frac{1-\mathcal{R}e\left(\mathbf{g}(p)\overline{\chi(p)}p^{-it}\right)}{p}\ll_{C,\varepsilon}1.$$

□

8. On rappelle ici qu'on considère la tribu $\mathcal{P}(\mathbb{N})$ sur $\mathbb{N}$ et la tribu produit dans les espaces produits.

IV.5 Preuve de la conjecture d'Erdos

Dans cette partie, on va prouver le théorème IV.4 ce qui démontrera la conjecture d'Erdos. Pour cela, on va supposer par l'absurde qu'il est faux. Il existe donc une constante $C > 0$ et une fonction $\mathbf{g}$ stochastique complètement multiplicative à valeurs dans $\mathbf{S}^1$ telle que :

$$\forall n \in \mathbb{N} : \mathbf{E}\left[\left|\sum_{j=1}^{n} \mathbf{g}(j)\right|^2\right] \leqslant C^2.$$

On va maintenant considérer que toutes les constantes peuvent dépendre de C, ainsi on pourra écrire :

$$\forall n \in \mathbb{N} : \mathbf{E}\left[\left|\sum_{j=1}^{n} \mathbf{g}(j)\right|^2\right] \ll 1.$$

On introduit désormais les grandeurs suivantes $\varepsilon, H, \delta, k, X$ vérifiant :

$$C \ll \frac{1}{\varepsilon} \ll H \ll \frac{1}{\delta}, k \ll X.$$

D'après la proposition IV.11, avec probabilité $1 - \mathcal{O}(\varepsilon)$, il existe un caractère de Dirichlet (stochastique) de période $\mathbf{q} = \mathcal{O}_{C,\varepsilon}(1)$ et un nombre réel (stochastique) $\mathbf{t} = \mathcal{O}_{C,\varepsilon}(X)$ tels que :

$$\sum_{p \leqslant X} \frac{1 - \mathcal{R}e\left(\mathbf{g}(p)\overline{\chi(p)}p^{-it}\right)}{p} \ll_{C,\varepsilon} 1.$$

Quitte à réduire χ, on peut supposer qu'il est primitif (car si χ' est un caractère primitif qui induit χ alors$\chi(p)$ et $\chi'(p)$ ne diffèrent qu'en les premiers p qui divisent le module de χ, dont il y en a un nombre fini).

Lemme IV.12

Avec probabilité $1 - \mathcal{O}(\varepsilon)$, $\mathbf{t} = \mathcal{O}_{\varepsilon}\left(X^{\delta}\right)$.

Démonstration : Preuve admise, elle se base sur des résultats d'analyse complexe qu'on a pas eu le temps d'aborder pendant le stage. On pourra en trouver une démonstration dans (Lemme 4.1, [5]). □

On se conditionne maintenant sur ces deux évènements de probabilité $1 - \mathcal{O}(\varepsilon)$.

On écrit pour tout $n \in \mathbb{N}$: $\mathbf{g}(n) = \tilde{\chi(n)}n^{it}\mathbf{h}(n)$ où on définit :

— pour tout p premier ne divisant pas $\mathbf{q}$: $\tilde{\chi}(p) = \chi(p)$ et $\mathbf{h}(p) = \mathbf{g}(p)\overline{\chi(p)}p^{it}$,

— pour tout p premier divisant $\mathbf{q}$: $\tilde{\chi}(p) = \mathbf{g}(p)p^{-it}$ et $\mathbf{h}(p) = 1$.

Avec ces notations on peut écrire :

$$\left|\sum_{p \leqslant X} \frac{1 - \mathcal{R}e\left(\mathbf{h}(p)\right)}{p}\right| \ll_{\varepsilon} 1. \tag{IV.2}$$

Lemme IV.13

Conditionné à l'évènement de probabilité $1-\mathcal{O}(\varepsilon)$, on a :

$$\frac{1}{H}\sum_{H<H'\leqslant 2H}\sum_{n\in\mathbb{N}}\frac{1}{n^{1+\frac{1}{\log(X)}}}\left|\sum_{m=1}^{H'}\tilde{\chi}(n+m)\mathbf{h}(n+m)\right|^2 \ll_\varepsilon \log(X).$$

Démonstration : Par hypothèse sur $\mathbf{g}$ et par inégalité triangulaire :

$$\forall n\in\mathbb{N}:\mathbf{E}\left[\frac{1}{H}\sum_{H<H'\leqslant 2H}\left|\sum_{m=1}^{H'}\mathbf{g}(n+m)\right|^2\right]\ll 1.$$

Soit $n\geqslant X^{2\delta}$, d'après le lemme IV.12 on sait que $\mathbf{t}=\mathcal{O}_\varepsilon\left(X^\delta\right)$ donc par un simple développement limité :

$$\forall m\leqslant H^{'}:(n+m)^{i\mathbf{t}}=n^{i\mathbf{t}}+\mathcal{O}_{\varepsilon,H,\delta}\left(X^{-\delta}\right).$$

Ainsi on a :

$$\forall n\geqslant X^{2\delta}:\mathbf{E}\left[\frac{1}{H}\sum_{H<H'\leqslant 2H}\left|\sum_{m=1}^{H'}\tilde{\chi}(n+m)n^{i\mathbf{t}}\mathbf{h}(n+m)\right|^2\right]\ll 1.$$

Si $n<X^{2\delta}$ alors on majore cette même quantité par H^2 trivialement. Ainsi pour δ suffisamment petit on obtient en moyennisant par la densité $\frac{1}{n^{1+\frac{1}{\log(X)}}}$:

$$\mathbf{E}\left[\frac{1}{H}\sum_{H<H'\leqslant 2H}\sum_{n\in\mathbb{N}}\frac{1}{n^{1+\frac{1}{\log(X)}}}\left|\sum_{m=1}^{H'}\tilde{\chi}(n+m)\mathbf{h}(n+m)\right|^2\right]\ll\log(X).$$

On obtient le résultat demandé par l'inégalité de Markov. □

Définition IV.14

Dans cette partie et uniquement dans cette partie, on dira qu'une classe a modulo $\mathbf{q}^k$ est ***mauvaise*** *si pour tout $m\in\{1,\dots,2H\}$, il existe un facteur premier p de $\mathbf{q}$ tel que $a+m$ soit divisible par p^k. Sinon, on dira que cette classe est* ***bonne****.*

Lemme IV.15

Avec les définitions précédentes, on a :

$$\frac{1}{\mathbf{q}^k}\sum_{a\text{ bonne}}\frac{1}{H}\sum_{H<H'\leqslant 2H}\left|\sum_{m=1}^{H'}\tilde{\chi}(a+m)\right|^2\ll_\varepsilon 1.$$

Démonstration : On reprend le résultat du lemme précédent en se restreignant à l'évènement de

probabilité $1-\mathcal{O}(\varepsilon)$ et en ne sommant que sur les bonnes classes :

$$\frac{1}{H}\sum_{H<H'\leqslant 2H}\sum_{\substack{a \text{ bonne}\\ n\equiv a[\mathbf{q}^k]}}\frac{1}{n^{1+\frac{1}{\log(X)}}}\left|\sum_{m=1}^{H'}\tilde{\chi}(n+m)\mathbf{h}(n+m)\right|^2 \ll_\varepsilon \log(X).$$

Donc d'après l'inégalité de Cauchy-Scwharz :

$$\frac{1}{H}\sum_{H<H'\leqslant 2H}\sum_{a \text{ bonne}}\left|\sum_{n\equiv a[\mathbf{q}^k]}\frac{1}{n^{1+\frac{1}{\log(X)}}}\sum_{m=1}^{H'}\tilde{\chi}(n+m)\mathbf{h}(n+m)\right|^2 \ll_\varepsilon \frac{\log(X)^2}{\mathbf{q}^k}.$$

Or on remarque que si n appartient à une bonne classe a alors $\tilde{\chi}(n+m)=\tilde{\chi}(a+m)$ (c'est en fait pour cela qu'on les a introduites). Ainsi on obtient facilement :

$$\frac{1}{H}\sum_{H<H'\leqslant 2H}\sum_{a \text{ bonne}}\left|\sum_{m=1}^{H'}\tilde{\chi}(a+m)\sum_{n\equiv a+m[\mathbf{q}^k]}\frac{1}{n^{1+\frac{1}{\log(X)}}}\mathbf{h}(n)\right|^2 \ll_\varepsilon \frac{\log(X)^2}{\mathbf{q}^k}.$$

On va maintenant estimer $\displaystyle\sum_{n\equiv a+m[\mathbf{q}^k]}\frac{1}{n^{1+\frac{1}{\log(X)}}}\mathbf{h}(n)$. On remarque tout d'abord par le second théorème de Mertens I.15 et par hypothèse sur $\mathbf{h}$ IV.2 que :

$$\log(X)\ll_\varepsilon \mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)\ll_\varepsilon \log(X). \tag{IV.3}$$

Maintenant on va regarder cette quantité quand χ est un caractère de Dirichlet : $\displaystyle\sum_{n\in\mathbb{N}}\frac{\chi(n)\mathbf{h}(n)}{n^{1+\frac{1}{\log(X)}}}$.

— Supposons que χ_1 est un caractère de Dirichlet non-principal de période divisant $\mathbf{q}^k$ alors par analycité en 1 de la série de Dirichlet associée à χ_1 :

$$\mathcal{D}\chi_1\left(1+\frac{1}{\log(X)}\right)\ll_{\mathbf{q},k} 1.$$

De plus en développant en produit eulérien on a l'égalité suivante :

$$\sum_{n\in\mathbb{N}}\frac{\chi_1(n)\mathbf{h}(n)}{n^{1+\frac{1}{\log(X)}}}=\mathcal{D}\chi_1\left(1+\frac{1}{\log(X)}\right)\prod_p\left(\left(1-\frac{\mathbf{h}(\mathbf{p})\chi_{\mathbf{1}}(\mathbf{p})}{p^{1+\frac{1}{\log(X)}}}\right)^{-1}\left(1-\frac{\chi_1(p)}{p^{1+\frac{1}{\log(X)}}}\right)\right).$$

On en déduit l'estimation suivante (en ne gardant que les termes dans le produit sur p en $\dfrac{1}{p^{1+\frac{1}{\log(X)}}}$) :

$$\sum_{n\in\mathbb{N}}\frac{\chi_1(n)\mathbf{h}(n)}{n^{1+\frac{1}{\log(X)}}}\ll_{\mathbf{q},k}\exp\left(\sum_p\frac{|1-\mathbf{h}(p)|}{p^{1+\frac{1}{\log(X)}}}\right).$$

Ainsi comme la série de terme général $\frac{|1-\mathbf{h}(p)|}{p^{1+\frac{1}{\log(X)}}}$ est convergente, on a :

$$\exp\left(\sum_{p}\frac{|1-\mathbf{h}(p)|}{p^{1+\frac{1}{\log(X)}}}\right)=\exp\left(\sum_{p\leqslant X}\frac{|1-\mathbf{h}(p)|}{p^{1+\frac{1}{\log(X)}}}\right)+\mathcal{O}(1)\ll\exp\left(\sum_{p\leqslant X}\frac{|1-\mathbf{h}(p)|}{p}\right).$$

On en déduit par l'inégalité de Cauchy-Scwharz puis le deuxième théorème de Mertens (I.15) et enfin l'hypothèse IV.2 sur $\mathbf{h}$:

$$\begin{aligned}\sum_{n\in\mathbb{N}}\frac{\chi_1(n)\mathbf{h}(n)}{n^{1+\frac{1}{\log(X)}}}&\ll_{\mathbf{q},k}\exp\left(\sum_{p\leqslant X}\frac{\mathcal{O}\left(1-\mathcal{R}e\left(\mathbf{h}(p)\right)\right)^{\frac{1}{2}}}{p}\right),\\&\ll_{\mathbf{q},k}\exp\left(\mathcal{O}\left(\log\log(X)\sum_{p\leqslant X}\frac{(1-\mathcal{R}e(\mathbf{h}(p))}{p}\right)^{\frac{1}{2}}\right),\\&\ll_{\mathbf{q},k}\exp\left(\mathcal{O}_{\varepsilon}\left(\sqrt{\log\log(X)}\right)\right).\end{aligned}$$

— Maintenant supposons que χ_0 soit un caractère principal de période $r|\mathbf{q}^k$. On peut alors écrire en développant en produit eulérien et en se rappelant que $\mathbf{h}(p)=1$ pour tout p divisant r (donc $\mathbf{q}^k$) :

$$\begin{aligned}\sum_{n\in\mathbb{N}}\frac{\chi_0(n)\mathbf{h}(n)}{n^{1+\frac{1}{\log(X)}}}&=\mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)\prod_{p|r}\left(1-\frac{1}{p^{1+\frac{1}{\log(X)}}}\right)\\&=\mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)\left(1+\mathcal{O}_{\varepsilon}\left(\frac{1}{\log(X)}\right)\right)\prod_{p|r}\left(1-\frac{1}{p}\right)\\&=\frac{\varphi(r)}{r}\mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)+\mathcal{O}_{\varepsilon}\left(\frac{\varphi(r)}{r}\mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)\frac{1}{\log(X)}\right)\\&=\frac{\varphi(r)}{r}\mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)+\mathcal{O}_{\varepsilon}(1)\end{aligned}$$

où la dernière estimation à été obtenue grâce à l'estimation IV.3.

De ces deux points par orthogonalité des caractères, on en déduit pour toute classe de résidus primitives $b[r]$ ($i.e$ $(b,r)=1$), on a :

$$\sum_{n\equiv b[r]}\frac{\mathbf{h}(n)}{n^{1+\frac{1}{\log(X)}}}=\frac{1}{r}\mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)+\mathcal{O}_{\mathbf{q},k}\left(\exp\left(\mathcal{O}_{\varepsilon}(\sqrt{\log\log(X)})\right)\right).$$

Si $b[r]$ est une classe de résidus non-primitive alors en notant $r'=\frac{r}{(b,r)}$ et $b'=\frac{b}{(b,r)}$:

$$\sum_{n\equiv b'[r']}\frac{\mathbf{h}(n)}{n^{1+\frac{1}{\log(X)}}}=\frac{1}{r}\mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)+\mathcal{O}_{\mathbf{q},k}\left(\exp\left(\mathcal{O}_{\varepsilon}(\sqrt{\log\log(X)})\right)\right).$$

Donc en utilisant l'estimation IV.3, on obtient quelque soit la classe de résidus $b[r]$:

$$\sum_{n\equiv b[r]}\frac{\mathbf{h}(n)}{n^{1+\frac{1}{\log(X)}}}=\frac{1}{r}\mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)+\mathcal{O}_{\mathbf{q},k}\left(\exp\left(\mathcal{O}_{\varepsilon}(\sqrt{\log\log(X)})\right)\right).$$

Ainsi en réinsérant cette expression dans celle de départ :

$$\frac{1}{H}\sum_{H<H'\leqslant 2H}\sum_{a\text{ bonne}}\left|\sum_{m=1}^{H'}\frac{\tilde{\chi}(a+m)}{\mathbf{q}^k}\mathcal{D}\mathbf{h}\left(1+\frac{1}{\log(X)}\right)+\mathcal{O}_{\mathbf{q},k}\left(\exp\left(\mathcal{O}_\varepsilon(\sqrt{\log\log(X)})\right)\right)\right|^2\ll_\varepsilon\frac{\log(X)^2}{\mathbf{q}^k}.$$

Or la contribution de $\mathcal{O}_{\mathbf{q},k}\left(\exp\left(\mathcal{O}_\varepsilon(\sqrt{\log\log(X)})\right)\right)$ est $\ll_\varepsilon \frac{\log(X)^2}{\mathbf{q}^k}$ pour X suffisamment grand. On en déduit le fait annoncé en utilisant IV.3. □

Lemme IV.16

Pour tous $d_1\neq d_2$ diviseurs de $\mathbf{q}^{k-1}$ et $m_1,m_2\in\{1,\dots,H'\}$, on a :

$$\sum_{\substack{a=1\\ d_1\mid a+m_1\\ d_2\mid a+m_2}}^{\mathbf{q}^k}\chi\left(\frac{a+m_1}{d_1}\right)\overline{\chi\left(\frac{a+m_2}{d_2}\right)}=0.$$

Démonstration : Nous n'allons pas faire en détails cette preuve. D'après les résultats du paragraphe I.1.3, on peut dire que le terme que l'on cherche à rendre nul est une combinaison linéaire de termes de la forme : $e\left(\frac{\xi a}{d_1\mathbf{q}}\right)$ pour $(\xi,d_1\mathbf{q})=1$ et $e\left(\frac{\xi a}{d_2\mathbf{q}}\right)$ pour $(\xi,d_2\mathbf{q})=1$. Comme $d_1\neq d_2$ toutes les fréquences qui vont apparaître seront non nulles et ainsi leur somme sera elle nulle. □

Fin de la démonstration du théorème IV.4 On va maintenant utiliser les deux lemmes précédents pour démontrer la démonstration du théorème IV.4. On se rappelle que l'on procède par l'absurde.

Si on est dans l'évènement $\mathbf{q}=1$ alors χ est constant égal à 1 et donc d'après le lemme IV.15 on a le fait suivant :

$$H\ll\frac{1}{H}\sum_{H<H'\leqslant 2H}\left(H'\right)^2\ll_\varepsilon 1,$$

ce qui est absurde pour H suffisamment grand devant ε. On se restreint maintenant à l'évènement $\mathbf{q}>1$, ainsi χ sera non-principal. Encore d'après le lemme IV.15 :

$$\frac{1}{H}\sum_{H<H'\leqslant 2H}\sum_{m_1,m_2=1,\dots,H'}\sum_{a\text{ bonne}}\tilde{\chi}(a+m_1)\overline{\tilde{\chi}(a+m_2)}\ll_\varepsilon\mathbf{q}^k.$$

On écrit $d_1=(a+m_1,\mathbf{q}^k)$ et $d_2=(a+m_2,\mathbf{q}^k)$ ainsi d_1,d_2 divisent $\mathbf{q}^{k-1}$. Ainsi :

$$\sum_{d_1,d_2\mid\mathbf{q}^{k-1}}\sum_{H<H'\leqslant 2H}\frac{\tilde{\chi}(d_1)\overline{\tilde{\chi}(d_2)}}{H}\sum_{m_1,m_2=1,\dots,H'}\sum_{\substack{a\text{ bonne}\\ d_1=(a+m_1,\mathbf{q}^k)\\ d_2=(a+m_2,\mathbf{q}^k)}}\chi\left(\frac{a+m_1}{d_1}\right)\overline{\chi\left(\frac{a+m_2}{d_2}\right)}\ll_\varepsilon\mathbf{q}^k.$$

Or on remarque que le nombre de mauvaises classes a est au plus $H\sum p|\mathbf{q}\left(\frac{\mathbf{q}}{p}\right)^k\ll H2^{-k}\mathbf{q}^k$. Ainsi on peut écrire :

$$\sum_{d_1,d_2\mid\mathbf{q}^{k-1}}\sum_{H<H'\leqslant 2H}\frac{\tilde{\chi}(d_1)\overline{\tilde{\chi}(d_2)}}{H}\sum_{m_1,m_2=1,\dots,H'}\sum_{\substack{a\\ d_1\mid a+m_1\\ d_2\mid a+m_2}}\chi\left(\frac{a+m_1}{d_1}\right)\overline{\chi\left(\frac{a+m_2}{d_2}\right)}\ll_\varepsilon\mathbf{q}^k.$$

D'après le lemme IV.16 on obtient :

$$\frac{1}{H}\sum_{d|\mathbf{q}^{k-1}}\sum_{H<H'\leqslant 2H}\sum_{m_1,m_2=1,\dots,H'}\sum_{\substack{a\\ d|a+m_1\\ d|a+m_2}}\chi\left(\frac{a+m_1}{d}\right)\overline{\chi\left(\frac{a+m_2}{d}\right)}\ll_\varepsilon \mathbf{q}^k.$$

On peut réécrire ceci sous cette forme :

$$\frac{1}{H}\sum_{d|\mathbf{q}^{k-1}}\sum_{H<H'\leqslant 2H}\sum_{a}\left|\sum_{\substack{m\in[1,H']\\ d|a+m}}\chi\left(\frac{a+m}{d}\right)\right|^2\ll_\varepsilon \mathbf{q}^k.$$

On voit ici que tous les termes de la somme sur d sont positifs ainsi on peut se restreindre à ceux qui s'écrivent $\mathbf{q}^i$ avec $\mathbf{q}^i<\sqrt{H}$:

$$\frac{1}{H}\sum_{i:\mathbf{q}^i<\sqrt{H}}\sum_{H<H'\leqslant 2H}\sum_{a}\left|\sum_{\substack{m\in[1,H']\\ \mathbf{q}^i|a+m}}\chi\left(\frac{a+m}{\mathbf{q}^i}\right)\right|^2\ll_\varepsilon \mathbf{q}^k.$$

Observons que par inégalité triangulaire on a :

$$\sum_{i:\mathbf{q}^i<\sqrt{H}}\frac{1}{H}\sum_{H'\in\left[H,\frac{3}{2}H\right]}\sum_{a}\left|\sum_{\substack{H'<m\leqslant H'+\mathbf{q}^i\\ \mathbf{q}^i|a+m}}\chi\left(\frac{a+m}{\mathbf{q}^i}\right)\right|^2\ll_\varepsilon \mathbf{q}^k.$$

Donc il existe $H'\in\left[H,\dfrac{3H}{2}\right]$ tel que :

$$\sum_{i:\mathbf{q}^i<\sqrt{H}}\sum_{a}\left|\sum_{\substack{H'<m\leqslant H'+\mathbf{q}^i\\ \mathbf{q}^i|a+m}}\chi\left(\frac{a+m}{\mathbf{q}^i}\right)\right|^2\ll_\varepsilon \mathbf{q}^k.$$

Et on peut le réécrire sous la forme :

$$\sum_{i:\mathbf{q}^i<\sqrt{H}}\sum_{a\leqslant\frac{\mathbf{q}^k}{2}}\left|\sum_{\substack{H'<m\leqslant H'+\mathbf{q}^i\\ \mathbf{q}^i|a+m}}\chi\left(\frac{a+m}{\mathbf{q}^i}\right)\right|^2\ll_\varepsilon \mathbf{q}^k.$$

Or on observe que la condition $0<m\leqslant\mathbf{q}^i$
$\mathbf{q}^i|a+m$ n'est vérifiée que par un seul entier m et il vérifie : $\dfrac{a+m}{\mathbf{q}^i}=\left\lfloor\dfrac{a}{\mathbf{q}^i}\right\rfloor+1$. Donc par changement de variable (et en retirant quelques termes parasites positifs) :

$$\sum_{i:\mathbf{q}^i<\sqrt{H}}\sum_{b\in[1,\mathbf{q}^{\frac{k-i}{4}}]}|\chi(b)|^2\ll_\varepsilon \mathbf{q}^k.$$

Or $b \mapsto |\chi(b)|^2 \gg_\varepsilon \mathbf{q}^{k-i}$ est périodique de période $\mathbf{q}$ et de moyenne $\gg_\varepsilon 1$. On en déduit :

$$\sum_{i:\mathbf{q}^i<\sqrt{H}} \sum_{b\in[1,\mathbf{q}^{\frac{k-i}{4}}]} |\chi(b)|^2 \ll_\varepsilon \sum_{i:\mathbf{q}^i<\sqrt{H}} 1.$$

Des deux dernières estimations on en déduit que :

$$\sum_{i:\mathbf{q}^i<\sqrt{H}} 1 \ll_\varepsilon 1.$$

Et ceci est absurde quand H est suffisamment grand. On en déduit le théorème IV.3 et donc la conjecture d'Erdos est démontrée.

Références

[1] G. Peyré. *L'algèbre discrète de la transformée de Fourier : niveau M1*. Mathématiques à l'université : cours et exercices corrigés. Ellipses, 2004. ISBN : 9782729818678. URL : https://books.google.fr/books?id=WFp8AAAACAAJ.

[2] Gérald Tenenbaum. *Introduction à la théorie analytique et probabiliste des nombres : Cours et exercices*. Dunod, 2022.

[3] Kannan Soundararajan. "The Liouville function in short intervals [after Matomaki and Radziwill]". In : *arXiv preprint arXiv :1606.08021* (2016).

[4] S. Chowla. *The Riemann Hypothesis and Hilbert's Tenth Problem*. Mathematics and its applications. Gordon et Breach, 1966. ISBN : 9780677001401.

[5] Terence Tao. "The Logarithmically averaged Chowla and Elliott conjectures for two-point correlations". In : *Forum of Mathematics, Pi* 4 (2016), e8. DOI : 10.1017/fmp.2016.6.

[6] K. Roth. "Remark concerning integer sequences". eng. In : *Acta Arithmetica* 9.3 (1964), p. 257-260. URL : http://eudml.org/doc/207480.

[7] Boris Konev et Alexei Lisitsa. "Computer-aided proof of Erd**H**os discrepancy properties". In : *Artificial Intelligence* 224 (2015), p. 103-118.

[8] Terence Tao. *Elementary multiplicative number theory*. https://terrytao.wordpress.com/2014/11/23/254a-notes-1-elementary-multiplicative-number-theory/. 2022.

[9] Terence Tao. *Mean values of nonpretentious multiplicative functions*. https://terrytao.wordpress.com/2019/12/17/254a-notes-10-mean-values-of-nonpretentious-multiplicative-functions/. 2019.

[10] Terence Tao. *Second moment and entropy methods*. https://terrytao.wordpress.com/2019/11/12/254a-notes-9-second-moment-and-entropy-methods/. 2019.

[11] Terence Tao. *A cheap version of Halasz's inequality*. https://terrytao.wordpress.com/2015/11/23/a-cheap-version-of-halaszs-inequality/. 2015.

[12] Olivier Bordellès. *Arithmetic tales*. Springer, 2012.

[13] Jérémy Bettinger. *Le théorème des nombres premiers*.

[14] Kaisa Matomäki et Maksym Radziwiłł. "Multiplicative functions in short intervals". In : *Annals of Mathematics* (2016), p. 1015-1056.

[15] Adrian Dudek. "An Elementary Proof of an Asymptotic Formula of Ramanujan". In : *arXiv preprint arXiv :1401.1514* (2014).

[16] Kaisa Matomäki, Maksym Radziwiłł, Terence Tao et al. "An averaged form of Chowla's conjecture". In : *Algebra Number Theory* 9.9 (2015), p. 2167-2196.

[17] Hugh L Montgomery et Robert C Vaughan. *Multiplicative number theory I : Classical theory*. 97. Cambridge university press, 2007.

[18] Henryk Iwaniec et Emmanuel Kowalski. *Analytic number theory*. T. 53. American Mathematical Soc., 2004.

[19] Terence Tao. *The Erdos discrepancy problem*. 2017. arXiv : 1509.05363 [math.CO].