Séminaire BOURBAKI Mars 2006
58ème année, 2005-2006, n^{o} 959

ÉCARTS ENTRE NOMBRES PREMIERS SUCCESSIFS [d'après Goldston, Pintz, Yıldırım, ...]

par **Emmanuel KOWALSKI**

1. INTRODUCTION

Rien n'est plus commun que de remarquer l'importance de la suite infinie

$$2 = p_1 < 3 = p_2 < \cdots < p_n < p_{n+1} < \cdots$$

des nombres premiers dans l'imaginaire des mathématiciens et des arithméticiens tout particulièrement. Parmi les belles questions qui restent complètement ouvertes concernant les nombres premiers, celles qui ont trait aux *écarts* entre nombres premiers consécutifs

$$\gamma(n) = p_{n+1} - p_n \text{ pour } n \geqslant 1$$

ont exercé depuis longtemps une fascination particulière.

Le point de départ d'une analyse rigoureuse de la répartition des valeurs de $\gamma(n)$ est forcément le théorème des nombres premiers. Notons $\pi(X)$, comme d'habitude, le nombre de nombres premiers $p \leqslant X$; on a alors

$$\pi(X) \sim \frac{X}{\ln X}, \text{ quand } X \to +\infty.$$

De manière équivalente, le n-ème nombre premier p_n vérifie

$$p_n \sim n \ln n, \text{ quand } n \to +\infty,$$

et donc, « en moyenne », la distance $\gamma(n)$ entre nombres premiers consécutifs est de l'ordre de $\ln n$: on a

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{2\leqslant n\leqslant N} \frac{\gamma(n)}{\ln n} = 1,$$

d'où en particulier :

$$\liminf_{n\to+\infty} \frac{\gamma(n)}{\ln n} \leqslant 1.$$

L'autre source de conjectures – qui n'a pas attendu la preuve du théorème des nombres premiers – consiste à « regarder » ce qui se passe dans des tables de nombres premiers. Il avait été observé(1) qu'un nombre relativement grand de paires (p_n, p_{n+1}) vérifient $p_{n+1} - p_n = 2$.

(1) Dickson attribue cela à Polignac.

En conséquence il a été conjecturé qu'une infinité de telles paires existent. C'est la conjecture des nombres premiers jumeaux : avec la notation ci-dessus, elle affirme que

$$\liminf_{n\to+\infty} \gamma(n) = 2. \tag{1}$$

Le thème particulier de ce rapport est celui de l'existence de petits écarts entre nombres premiers (par opposition, par exemple, avec la question de borner $\gamma(n)$ pour tout n). Après des progrès sporadiques, dus à Erdös, Rankin, Bombieri-Davenport [2], Huxley, Maier, en particulier, l'étude des petits écarts entre nombres premiers a été bouleversée durant l'année 2005. Goldston, Pintz et Yıldırım [11], [13], [14] ont en effet démontré :

Théorème 1.1 (Goldston, Pintz, Yıldırım). — *On a*

$$\liminf_{n\to+\infty} \frac{\gamma(n)}{\ln n} = 0.$$

Le meilleur résultat précédemment connu était dû à Maier :

$$\liminf_{n\to+\infty} \frac{\gamma(n)}{\ln n} \leqslant 0,2484\ldots$$

De plus, en admettant certaines hypothèses usuelles concernant la répartition des nombres premiers dans les progressions arithmétiques, on a des résultats vraiment spectaculaires :

Théorème 1.2 (Goldston, Pintz, Yıldırım). — (1) *Si l'exposant de répartition θ* au sens fort *de la suite des nombres premiers vérifie $\theta > 1/2$, on a*

$$\liminf_{n\to+\infty} \gamma(n) < +\infty.$$

(2) *Si $\theta = 1$, alors*

$$\liminf_{n\to+\infty} \gamma(n) \leqslant 16.$$

L'exposant de répartition θ est défini ci-dessous (Définition 2.2). On sait que $\theta \geqslant 1/2$: c'est le célèbre théorème de Bombieri-Vinogradov.

Dans la Section 7, nous présenterons des énoncés plus forts et d'autres résultats annoncés, concernant par exemple les écarts entre nombres de la forme p_1p_2 avec $p_1 \neq p_2$ (c'est à dire, ayant exactement deux facteurs premiers).

Mais, pour commencer, nous allons expliquer la démonstration des Théorèmes 1.1 et 1.2, après avoir présenté la base de la méthode telle qu'elle apparait à l'heure actuelle.

Il faut signaler que le sujet est en pleine ébullition : tous les résultats sont encore sous forme de prépublications, et différentes variantes de la méthode apparaissent dans chacun d'entre eux. Malgré la malencontreuse aventure de 2003, aucun doute ne subsiste cependant quand à la validité des deux énoncés ci-dessus.

Notations. Rappelons quelques notations de théorie analytique des nombres.
– p désignera toujours un nombre premier et n, m des entiers $\geqslant 1$.
– $\Lambda(n)$ désigne la fonction de von Mangoldt ; $\Lambda(n) = 0$ si n n'est pas une puissance d'un nombre premier p et $\Lambda(p^k) = \ln p$. De plus,

$$\psi(x; q, a) = \sum_{\substack{n\leqslant x \\ n\equiv a \pmod q}} \Lambda(n).$$

– $\mu(n)$ désigne la fonction de Möbius, $\tau(n)$ la fonction « nombre de diviseurs » et $\varphi(n)$ la fonction d'Euler. Pour $n = p^k$, $k \geqslant 1$, on a $\mu(p^k) = -1$ si $k = 1$, 0 sinon, $\tau(p^k) = k + 1$ et $\varphi(p^k) = p^k - p^{k-1}$.
– $f \star g$ désigne la convolution arithmétique de f et g :

$$f \star g(n) = \sum_{d \mid n} f(d)g(n/d) = \sum_{ab=n} f(a)g(b).$$

– $\sum^{\flat}$ désigne une somme restreinte aux entiers n sans facteurs carrés.
– Les notations de Landau $f = O(g)$ et de Vinogradov $f \ll g$ sont considérées comme synonymes : $f(x) = O(g(x))$ pour tout $x \in D$ signifie qu'il existe une constante « implicite » $C \geqslant 0$ (le plus souvent, fonction d'autres paramètres explicitement mentionnés) telle que $|f(x)| \leqslant Cg(x)$ pour tout $x \in D$. Cette définition *diffère* de celle, topologique, de Bourbaki [4, Chap. V]. Par contre, nous utilisons les notations $f(x) \sim g(x)$ et $f = o(g)$ dans le sens asymptotique de loc. cit.

Remerciements. Je remercie D. Goldston, J. Pintz, C. Yıldırım, J. Sivak, R. de la Bretèche et É. Fouvry pour leurs remarques et corrections concernant les différents brouillons de ce texte.

2. RAPPELS DE QUELQUES NOTIONS DE CRIBLE

Cette section rappelle quelques définitions relatives aux méthodes de crible (voir par exemple [15] pour un traitement classique très complet, [18, Ch. 6], [19, Ch. 4] pour des exposés assez courts comportant les preuves des assertions de base, et [20], [5] pour deux survols dans ce séminaire).

Soit $\mathcal{A} = (a_n)$, $n \geqslant 1$, une suite de nombres réels positifs. On note

$$\mathcal{A}(x) = \sum_{n \leqslant x} a_n$$

sa fonction sommatoire.

Soit P un entier, qui sera le plus souvent le produit de certains nombres premiers bien choisis, par exemple le produit des nombres premiers $p < z$. Les *sommes criblées* correspondantes sont

$$\mathcal{A}(x, P) = \sum_{\substack{n \leqslant x \\ (n,P)=1}} a_n.$$

L'objectif général du crible est d'estimer $\mathcal{A}(x, P)$ avec la plus grande généralité possible.

Pour aborder l'étude des sommes criblées $\mathcal{A}(x, P)$, les méthodes de « petit » crible sont inspirées par la formule de Legendre(2) : on a

$$\mathcal{A}(x, P) = \sum_{d \mid P} \mu(d)\mathcal{A}_d(x), \qquad \text{avec} \qquad \mathcal{A}_d(x) = \sum_{\substack{n \leqslant x \\ n \equiv 0 \pmod{d}}} a_n.$$

(2)Que l'on peut voir, au choix, comme l'expression du principe d'inclusion–exclusion, ou de la formule $\sum_{d \mid n} \mu(d) = 0$ si $n \neq 1$.

Bien qu'elle ne soit pas exploitable directement en général car le nombre de termes (i.e., le nombre de diviseurs de P) est trop grand, cette formule suggère d'étudier les suites $\mathcal{A}_d = (a_{nd})$. Dans beaucoup de cas intéressants du point de vue de la théorie multiplicative des nombres, il se dégage naturellement une approximation de $\mathcal{A}_d(x)$ de la forme

$$\mathcal{A}_d(x) = g(d)X + r_d(\mathcal{A}; x) \tag{2}$$

où $X \geqslant 0$ est une fonction de x et $d \mapsto g(d)$ est une fonction arithmétique que l'on suppose multiplicative[(3)] et $r_d(\mathcal{A}; x)$ est un « reste ».

En pratique, cette approximation est assez précise pour des valeurs de d assez grandes, disons $d < D$, c'est-à-dire que le reste total

$$\sum_{d<D} |r_d(\mathcal{A}; x)|$$

peut être estimé « suffisamment bien ». Plus précisément, la notion *d'exposant de répartition* est classiquement définie :

Définition 2.1. — *La suite $\mathcal{A}$ a un exposant de répartition* au sens faible $\geqslant \theta$ *si, pour tout $A > 0$ et tout $\varepsilon > 0$, on a*

$$\sum_{d<D} |r_d(\mathcal{A}; x)| \ll \frac{X}{(\ln X)^A} \tag{3}$$

pour $x \geqslant 2$ et $D \leqslant x^{\theta-\varepsilon}$, la constante implicite dépendant au plus de A, ε et $\mathcal{A}$.

La méthode de Goldston, Pintz et Yıldırım requiert crucialement des informations, non seulement pour une suite $\mathcal{A} = (a_n)$, mais pour les suites « décalées » (a_{n+t}), où $t \geqslant 0$ est un entier quelconque. Cela mène à un renforcement naturel de la notion d'exposant de répartition.

Pour la définir, on généralise donc (2) en écrivant

$$\sum_{\substack{n\leqslant x\\ n\equiv t\ (d)}} a_n = g_t(d)X + r_d(\mathcal{A}; x, t) \tag{4}$$

où X est encore une fonction de x et $d \mapsto g_t(d)$ est pour tout t une fonction multiplicative dont la valeur ne dépend que de t modulo d.

Définition 2.2. — *Une suite $\mathcal{A} = (a_n)$ de réels positifs a un exposant de répartition* au sens fort $\geqslant \theta$ *si pour tout $A > 0$ et tout $\varepsilon > 0$, on a*

$$\sum_{d\leqslant D} \max_{t \,(\mathrm{mod}\ d)} \max_{y\leqslant x} |r_d(\mathcal{A}; y, t)| \ll \frac{X}{(\ln X)^A}, \tag{5}$$

pour $x \geqslant 2$ et $D \leqslant x^{\theta-\varepsilon}$, la constante implicite pouvant dépendre de A, ε et $\mathcal{A}$.

Pour le Théorème 1.2, la suite concernée est la suite $(\Lambda(n))$, ou bien la fonction caractéristique des nombres premiers.

On va voir dans la preuve du Théorème 1.2 qu'une notion plus faible se dégage. Comme cela pourrait fournir une voie plus accessible à la preuve inconditionnelle de ce résultat, il semble utile de l'isoler.

[(3)] Au sens où $g(nm) = g(n)g(m)$ si $(n, m) = 1$.

Définition 2.3. — *Une suite $\mathcal{A}=(a_n)$ de réels positifs a un P-*exposant de répartition $\geqslant \theta$ *si pour tout $A>0$, tout $\varepsilon>0$ et tout polynôme primitif $F\in\mathbf{Z}[X]$ de degré $\geqslant 1$ on a*

$$\sum_{d\leqslant D}^{\flat}\sum_{\substack{\nu\,(\mathrm{mod}\ d)\\ F(\nu)=0}}\max_{y\leqslant x}|r_d(\mathcal{A};y,\nu)|\ll\frac{X}{(\ln X)^A},$$

pour $x\geqslant 2$ et $D\leqslant x^{\theta-\varepsilon}$, la constante implicite pouvant dépendre de A, ε, $\mathcal{A}$ et F.

Remarque 2.4. — Une prépublication de Pintz et Motohashi [21], parue après la première version de ce survol, introduit explicitement une variante de cette dernière notion dans le cas de la suite des nombres premiers, avec une plus grande flexibilité potentiellement intéressante.

Le cas le plus important pour notre propos est celui de la suite $a_n=\Lambda(n)$. Dans ce cas on a pour $t\geqslant 0$

$$\sum_{\substack{n\leqslant x\\ n\equiv t\ (d)}}a_n=\sum_{\substack{n\leqslant x\\ n\equiv t\ (d)}}\Lambda(n)=\psi(x;d,t),$$

et si $(d,t)=1$ on a l'approximation

$$\sum_{\substack{n\leqslant x\\ n\equiv t\ (d)}}a_n=\frac{x}{\varphi(d)}+r_d(\mathcal{A};x,t)\tag{6}$$

avec la borne

$$r_d(\mathcal{A};x,t)\ll x(\ln x)^{-A}$$

valide pour tout $x\geqslant 2$, $d\geqslant 1$, t modulo d (tel que $(d,t)=1$) et tout $A>0$, la constante implicite (ineffective) dépendant seulement de A ; c'est le théorème de la progression arithmétique sous la forme donnée par Siegel-Walfisz ([18, Cor. 5.29] par exemple).

Tenant compte des cas $(d,t)\neq 1$, on peut définir

$$X=x,\quad g_t(d)=\begin{cases}0 & \text{si } (d,t)\neq 1,\\ \frac{1}{\varphi(d)} & \text{si } (d,t)=1,\end{cases}$$

pour avoir (4) avec

$$r_d(\mathcal{A};x,t)\ll x(\ln x)^{-A}$$

dans tous les cas.

Tel quel, cela ne donne qu'un exposant de répartition nul. L'ingrédient crucial pour cribler efficacement une telle suite est le théorème de Bombieri-Vinogradov (voir par exemple [18, Ch. 17]) :

Théorème 2.5. — *Pour tout $A>0$, il existe $B>0$ tel que*

$$\sum_{d\leqslant D}\max_{\substack{(a,d)=1\\ a\,(\mathrm{mod}\ d)}}\max_{y\leqslant x}\Big|\psi(y;d,a)-\frac{y}{\varphi(d)}\Big|\ll\frac{x}{(\ln x)^A}$$

pour $x\geqslant 2$ et $D\leqslant x^{1/2}(\ln x)^{-B}$, la constante implicite dépendant uniquement de A.

La restriction à $(a, d) = 1$ étant inoffensive, cela implique que la suite $(\Lambda(n))$ a un exposant de répartition $\geqslant 1/2$. La limite à $1/2$ est bien comprise : c'est ce que donnerait trivialement l'Hypothèse de Riemann Généralisée pour les fonctions L de Dirichlet. Aller au delà revient donc, d'une manière ou d'une autre, à obtenir des informations sur la répartition des zéros de $L(\chi, s)$ par rapport à χ variable (voir la Section 8). Dit d'une autre manière, analytiquement, il faut aller au delà de l'inégalité générale du Grand Crible qui est l'un des ingrédients cruciaux de la preuve (un autre étant (6)).

Remarque 2.6. — (1) On appellera, suivant l'usage, *Conjecture d'Elliott-Halberstam* la conjecture qui énonce[(4)] que $(\Lambda(n))$ a un exposant de répartition (au sens fort[(5)]) égal à 1.

(2) Les généralisations du théorème de Bombieir- Voir la dernière section de ce rapport pour quelques remarques concernant des suites pour lesquelles on connaît un exposant de répartition $> 1/2$ (faible ou fort).

(3) Le P-exposant de répartition semble être une notion nouvelle. Bien entendu, si une suite a un exposant de répartition $\geqslant \theta$ au sens fort, il en est de même pour le P-exposant. Trouver des suites non triviales pour lesquelles le P-exposant de répartition dépasse la limite (universelle) du grand crible semble un problème intéressant.

L'exemple de crible qui est sous-jacent à la méthode de Goldston, Pintz et Yıldırım est lié à la « conjecture des k-uplets » de Hardy et Littlewood. Soit $k \geqslant 2$ et $\boldsymbol{h} = (h_1, \ldots, h_k)$ un k-uplet d'entiers $\geqslant 0$ distincts. Le problème est de trouver les entiers n tels que $n + h_1$, ..., $n + h_k$ soient tous premiers, autrement dit de trouver les entiers $n \geqslant 1$ tels que

$$F_{\boldsymbol{h}}(n) = (n + h_1) \cdots (n + h_k)$$

possède k facteurs premiers exactement. On notera $\pi(X, \boldsymbol{h})$ le nombre des $n \leqslant X$ vérifiant cette condition.

Puisque tout entier de ce type assez grand (en fonction de $\varepsilon > 0$) a tous ses facteurs premiers $> n > F_{\boldsymbol{h}}(n)^{1/k-\varepsilon}$, les méthodes de crible fournissent aisément une borne supérieure (voir par exemple [18, Th. 6.7], [15, Th. 5.7]) :

Théorème 2.7. — *On a*

$$\pi(X, \boldsymbol{h}) \leqslant 2^k k! \mathfrak{S}(\boldsymbol{h}) X (\ln X)^{-k} \Big(1 + O\Big(\frac{\ln \ln X}{\ln X}\Big)\Big)$$

pour $X \geqslant 3$*, où*

$$\mathfrak{S}(\boldsymbol{h}) = \prod_p \Big(1 - \frac{\nu_{\boldsymbol{h}}(p)}{p}\Big)\Big(1 - \frac{1}{p}\Big)^{-k}, \tag{7}$$

avec $\nu_{\boldsymbol{h}}(p)$ *la cardinalité de l'ensemble* $\{h_1 \pmod p, \ldots, h_k \pmod p\}$ *des réductions modulo* p *des composantes de* $\boldsymbol{h}$*, ce produit étant absolument convergent.*

Comme dans [11], [12], [13], on dira que $\boldsymbol{h}$ est *admissible* si les composantes de $\boldsymbol{h}$ sont distinctes et si $\mathfrak{S}(\boldsymbol{h}) \neq 0$. Cela signifie qu'il n'existe pas de nombre premier p avec $\nu_{\boldsymbol{h}}(p) = p$, c'est-à-dire qu'il existe pas de nombre premier p qui divise toujours l'un au moins des entiers $n + h_1$, ..., $n + h_k$. Des raisonnements heuristiques (effectués, dans ce cas, par Hardy et Littlewood, voir [18, 13.1] ou [20, 2], par exemple, pour une présentation moderne) suggèrent

(4) Ce n'est pas la version originale de Elliott et Halberstam, qui est fausse, comme l'ont démontré Friedlander et Granville.

(5) Dans la suite, on omettra parfois de préciser « au sens fort », qui sera sous-entendu.

que la borne supérieure obtenue est du bon ordre de grandeur[6] si $\boldsymbol{h}$ est admissible, et plus précisément :

Conjecture 2.8 (Hardy–Littlewood). — *Si $\boldsymbol{h}$ est admissible, on a*

$$\pi(X,\boldsymbol{h}) \sim \mathfrak{S}(\boldsymbol{h})X(\ln X)^{-k} \text{ quand } X \to +\infty.$$

Bien entendu, le cas $\boldsymbol{h}=(0,2)$ correspond à la conjecture des nombres premiers jumeaux.

Le facteur[7] $\mathfrak{S}(\boldsymbol{h})$ joue un rôle important dans la partie inconditionnelle des travaux de Goldston, Pintz et Yıldırım (le Théorème 1.1). Plus précisément, il est nécessaire de connaître sa valeur moyenne lorsque $\boldsymbol{h}$ parcourt les k-uplets admissibles $\boldsymbol{h}=(h_i)$ avec $1 \leqslant h_i \leqslant h$ pour k fixé. Celle-ci a été calculée par Gallagher [9].

Proposition 2.9. — *Pour $h \geqslant 1$ et $k \geqslant 1$ des entiers, soit*

$$\mathsf{S}(h,k) = \frac{1}{h^k} \sideset{}{^*}\sum_{\substack{\boldsymbol{h}=(h_1,\ldots,h_k)\\ 1\leqslant h_i\leqslant h}} \mathfrak{S}(\boldsymbol{h}) \tag{8}$$

où $\sum^*$ signifie que la somme est restreinte aux k-uplets admissibles[8].

Pour tout k fixé, on a

$$\lim_{h\to+\infty} \mathsf{S}(h,k) = 1.$$

Cela veut dire que la valeur moyenne de $\mathfrak{S}(\boldsymbol{h})$ est 1, ce qui est cohérent avec ce qu'on attend d'un produit eulérien. L'estimation uniforme suivante est également utile : il existe une constante $b_k \geqslant 0$ telle que

$$\mathfrak{S}(\boldsymbol{h}) \ll (\ln\ln 10h)^{b_k}, \tag{9}$$

pour tout k-uplet $\boldsymbol{h}$ tel que $1 \leqslant h_i \leqslant h$ pour tout i, la constante implicite ne dépendant que de k (voir par exemple [13, Lemma 6]).

Remarque 2.10. — Ce résultat de Gallagher est lié à une conjecture concernant la répartition de $\gamma(n)$, à savoir que la suite $(\gamma(n)/\ln n)$ est, pour $n \to +\infty$, équirépartie pour la mesure $e^{-t}dt$ sur $[0,+\infty[$ (« distribution de Poisson », correspondant à des nombres purement aléatoires et sans corrélations). On s'attend donc à avoir

$$\lim_{N\to+\infty} \frac{1}{N}|\{n \leqslant N \mid \alpha \leqslant \frac{\gamma(n)}{\ln n} \leqslant \beta\}| = \int_\alpha^\beta e^{-t}dt \tag{10}$$

pour tout $0 \leqslant \alpha < \beta$. Le lien avec la Proposition 2.9 est que celle-ci permet de démontrer qu'une forme forte et uniforme de la conjecture des k-uplets implique (10), en calculant les moments nécessaires.

(6) Cela peut se « deviner » en disant que chaque $n+h_i$ a une « probabilité » $1/(\ln n)$ d'être premier et en supposant les différentes conditions indépendantes.

(7) Appelé parfois la « série singulière » par référence à la terminologie employée dans la méthode du cercle de Hardy et Littlewood.

(8) En particulier, dont les composantes h_i sont distinctes ; par contre, les composantes ne sont pas ordonnées.

La loi de Poisson est aussi la répartition que l'on attend pour la suite des écarts normalisés entre valeurs propres du laplacien sur la courbe modulaire $SL(2, \mathbf{Z})\backslash\mathbf{H}$; rappellons [20] que la répartition des écarts normalisés des zéros de $\zeta(s)$ est (conjecturalement) très différente.(9)

3. PRINCIPE DE LA MÉTHODE

Dans cette section, nous allons présenter – comme l'auteur de ce rapport est parvenu à le comprendre – le principe de base de la démonstration des Théorèmes 1.1 et 1.2. La description tente de présenter le raisonnement le plus naturellement du monde, mais n'en respecte pas nécessairement la chronologie exacte. Le préprint [12] fournit plus de détails à ce sujet.

La première chose à faire est de considérer que la conjecture des k-uplets est « plus fondamentale » que l'énoncé du Théorème 1.1 par exemple. On va donc chercher à aborder des formes plus faibles de cette conjecture. Il est classique d'essayer de le faire en cherchant des k-uplets « presque premiers ». La première idée nouvelle est *d'affaiblir différemment la conjecture.* Soit donc $\boldsymbol{h}$ un k-uplet admissible. On cherchera des n tels que le k-uplet $(n + h_1, \ldots, n + h_k)$ contienne *au moins* deux nombres premiers. Lorsque c'est le cas, on a évidemment « trouvé » deux nombres premiers tels que $p < p'$ et $p' - p \leqslant \max(h_i - h_j)$. Le Théorème 1.2 en sera une conséquence ; quand au Théorème 1.1, un seul k-uplet ne suffisant pas, on combinera à l'aide de la Proposition 2.9 l'effet de tous ceux tels que $h_i \leqslant h$, avec $h = \delta \ln n$...

Un intérêt de cet affaiblissement (qui n'en est un que si $k \geqslant 3$, d'ailleurs) est qu'un « détecteur » manipulable de la condition voulue peut se décrire assez facilement. Notons $n \lhd \boldsymbol{h}$ la relation

« Il y a au moins deux (puissances de) nombres premiers parmi $n + h_1, \ldots, n + h_k$ »,

et posons

$$\vartheta_{\boldsymbol{h}}(n) = \sum_{1 \leqslant i \leqslant k} \Lambda(n + h_i).$$

On a alors, si $h_i \leqslant n$:

$$\vartheta_{\boldsymbol{h}}(n) > \ln 2n \text{ implique } n \lhd \boldsymbol{h}. \tag{11}$$

On va tenter de comparer(10) deux sommes, à savoir

$$Q_g = \sum_{N < n \leqslant 2N} \vartheta_{\boldsymbol{h}}(n) E(n) \quad \text{ et } \quad Q_p = (\ln 3N) \sum_{N < n \leqslant 2N} E(n)$$

où $N \geqslant 1$ et les poids $E(n) \geqslant 0$ sont à notre disposition. Le détecteur ci-dessus fournit la minoration :

$$(k-1)(\ln 3N) \sum_{\substack{N < n \leqslant 2N \\ n \lhd \boldsymbol{h}}} E(n) \geqslant (Q_g - Q_p),$$

(9) L'analogue du Théorème 1.1 pour l'espacement moyen des parties imaginaires des zéros de $\zeta(s)$ n'est pas connu ; d'après un résultat – en fait beaucoup plus fort – de Conrey et Iwaniec, cela impliquerait qu'il n'existe pas de zéros de Laudau-Siegel pour les fonctions L de Dirichlet.

(10) Sous la forme plus précise qui sera abordée ci-dessous, Goldston, Pintz et Yıldırım attribuent cette formulation de leur méthode à Granville et Soundararajan.

en particulier, si $Q_g > Q_p$, il existe au moins une valeur de n qui convient dans l'intervalle $N < n \leqslant 2N$.

C'est dans le choix des coefficients $E(n)$ que se trouve la clé du succès des travaux de Goldston, Pintz et Yıldırım. Il n'y a rien d'évident à ce qu'il existe un choix impliquant un résultat aussi fort que le Théorème 1.2, et cette méthode de détection, suivant le problème, est plus ou moins efficace.(11) *A posteriori*, l'approche peut sembler « évidente » (si le rédacteur de ce rapport a bien fait son travail...), mais comme souvent en théorie analytique des nombres, il faut combiner deux ou trois idées significatives pour obtenir un nouveau résultat vraiment convaincant.

On cherche donc alors à minorer Q_g et majorer Q_p. On ne peut réellement espérer qu'une chose : que Q_g et Q_p soit du même ordre grandeur, mais avec

$$\rho = \limsup_{N\to+\infty} \frac{Q_g}{Q_p} > 1.$$

Autrement dit : *la réussite dépendra d'une inégalité numérique*(12). Même si, avec le type de coefficients $E(n)$ discutés ci-dessous, il est assez aisé de deviner la valeur de ρ (sous la Conjecture 2.8, et éventuellement d'autres) et donc de justifier de manière plausible le choix effectué, la difficulté est de parvenir à démontrer rigoureusement que $\rho > 1$.

Remarque 3.1. — La méthode de Bombieri et Davenport [2] (réinterprétée par Goldston [10]) peut essentiellement se ramener à ce cadre, en changeant le détecteur et avec $E(n) = \Lambda(n)$: il s'agit grosso modo de démontrer que

$$Q_1 = \sum_{N<n\leqslant 2N} \Bigl(\sum_{k=1}^{h} \alpha(k)\Lambda(n+2k)\Bigr)\Lambda(n),$$

pour certains coefficients $\alpha(k)$ bien choisis, est suffisamment proche de sa valeur attendue

$$Q_2 = \sum_{N<n\leqslant 2n} \Bigl(\sum_{k=1}^{h} \alpha(k)\mathfrak{S}(0,2k)\Bigr)\Lambda(n) \sim \Bigl(\sum_{k=1}^{h} \alpha(k)\mathfrak{S}(0,2k)\Bigr)N$$

(donnée par la conjecture de Hardy-Littlewood) pour impliquer l'existence de nombres premiers à distance $\leqslant 2h$, lorsque h est de l'ordre de $\frac{1}{4}\log N$. Il faut noter qu'ici les deux sommes Q_1 et Q_2 sont asymptotiquement égales (conjecturalement), c'est-à-dire que $\rho = 1$.

Le coefficient 1/4 est lié à l'exposant de répartition (au sens fort) des suites de nombres premiers décalés, et est donné par le Théorème de Bombieri-Vinogradov. En particulier, comme l'observe Goldston [10], cette méthode démontrait déjà qu'une certaine forme de la Conjecture d'Elliott-Halberstam implique le Théorème 1.1 (*pas* le Théorème 1.2).

La somme Q_g porte sur un support limité aux entiers n tels que $\{n+h_1,\dots,n+h_k\}$ contienne un nombre premier (ou une puissance...) au moins, alors que Q_p n'a pas cette

(11) Par exemple, l'analogue du Théorème 1.1 pour les écarts $p_{n+2} - p_n$, que l'on est tenté d'aborder en considérant $\vartheta_h(n) - 2\ln 3N$, n'est pas connu inconditionnellement ; il l'est cependant sous la Conjecture d'Elliott-Halberstam, voir Théorème 7.3.

(12) Cela peut être contrasté avec la méthode de « mollification » utilisée, par exemple, par Selberg pour minorer le nombre de zéros critiques de $\zeta(s)$. Le succés ne dépend alors que d'avoir le bon ordre de grandeur du second moment mollifié.

restriction a priori. Il est donc naturel de demander à $E(n)$ d'imposer une restriction similaire, afin d'ajuster la « densité » relative espérée pour Q_p à celle de Q_g.

Bien entendu, prendre

$$E'(n) = \Lambda(n+h_1)\cdots\Lambda(n+h_k)$$

est le meilleur choix en principe : conjecturalement, la somme Q_g construite avec $E'(n)$ compte exactement (avec multiplicité k) les k-uplets de nombres premiers. Mais – c'est toute la question que l'on cherche à résoudre – on ne saurait pas démontrer rigoureusement les estimations nécessaires. On va donc introduire des approximations de $E'(n)$ qui seront plus manipulables.(13)

La recherche de la bonne approximation a procédé par tâtonnements. Il est classique d'approcher la fonction Λ en écrivant

$$\Lambda(n) = \sum_{d\mid n} \mu(d)\Big(\ln \frac{n}{d}\Big)$$

et en utilisant une convolution tronquée

$$\Lambda(n,x) = \sum_{\substack{d\mid n\\ d\leqslant x}} \mu(d)\Big(\ln \frac{x}{d}\Big),$$

où $x \geqslant 1$ est un paramètre, que l'on souhaite pouvoir prendre le plus grand possible.

La tentation est grande de choisir comme coefficients

$$E_0(n) = \Lambda(n+h_1,x)\cdots\Lambda(n+h_k,x).$$

Un premier obstacle apparaît : il n'est plus vrai que $E_0(n) \geqslant 0$; cela rend impossible de déduire d'une inégalité $Q_g > Q_p$ qu'il existe n tel que $n \lhd \boldsymbol{h}$. On peut vouloir contourner le problème en prenant $E_0(n)^2$; c'est en effet la précédente approche de Goldston et Yıldırım. Elle s'avère insuffisante pour démontrer le Théorème 1.1.(14)

Il s'avère que trouver une bonne approximation nécessite de considérer le problème plus globalement en rapport avec l'objectif recherché par la méthode considérée : puisque on se contente (maintenant) d'avoir au moins deux nombres premiers dans les k-uplets considérés, il peut être (et il est) préférable de faire appel à un poids qui tienne compte de ce relâchement. Plutôt que $E'(n)$, qui « compte » des k-uplets d'entiers tous premiers, on utilise donc comme référence un poids qui sélectionne les k-uplets $(n+h_1,\ldots,n+h_k)$ d'entiers *dont le produit n'a que peu de facteurs premiers, mais éventuellement plus que* k.

Précisément, l'innovation critique de Goldston, Pintz et Yıldırım est donc d'utiliser un poids (de type « crible ») qui sélectionne essentiellement les entiers n tels que $F_{\boldsymbol{h}}(n)$ ait au plus $k+\ell$ facteurs premiers, où

$$F_{\boldsymbol{h}}(n) = (n+h_1)\cdots(n+h_k)$$

et où le paramètre $\ell \geqslant 1$ est également disponible pour être optimisé. Il est clair que si ℓ reste suffisamment petit par rapport à k, une grande proportion des k-uplets ainsi sélectionnés vérifieront la relation $n \lhd \boldsymbol{h}$. (Voir la remarque ci-dessous pour une illustration numérique, qui vaut ce qu'elle vaut...)

(13)Là aussi, on peut penser à la méthode de mollification pour l'étude des zéros de fonctions L.

(14)Mais c'est un des outils arithmétiques de la preuve du théorème de Green-Tao [17].

Pour concrétiser cela, [11] utilise le poids

$$E(n) = \Lambda_{k+\ell}(F_{\boldsymbol{h}}(n), x)$$

où $\Lambda_m(n,x)$, pour tout entier $m \geqslant 1$, est une troncation de la m-ème fonction de von Mangoldt Λ_m, par analogie avec le cas $\Lambda = \Lambda_1$:

$$\Lambda_m(n) = (\mu \star \ln^m)(n) = \sum_{d \mid n} \mu(d)\Bigl(\ln \frac{n}{d}\Bigr)^m \quad \text{et} \quad \Lambda_m(n,x) = \sum_{\substack{d \mid n \\ d \leqslant x}} \mu(d)\Bigl(\ln \frac{x}{d}\Bigr)^m.$$

Le choix de $\Lambda_m(n)$ dans l'étude de la répartition des nombres « presque premiers » dans une suite est largement justifié par la théorie du crible, en particulier par le « crible asymptotique » de Bombieri où ces fonctions ont fait leur apparition. Il est facile de voir que Λ_m est supportée sur les entiers ayant au plus m facteurs premiers (sans multiplicité)(15), et Bombieri est parvenu à démontrer (entre autres) que la Conjecture d'Elliott-Halberstam implique la formule asymptotique

$$\sum_{n \leqslant X} \Lambda(n)\Lambda_m(n+2) \sim m\mathfrak{S}(0,2)X(\log X)^{m-1} \qquad \text{quand } X \to +\infty,$$

pour tout $m \geqslant 2$.(16)

Remarque 3.2. — La remarque qui suit est assez gratuite, mais puisque l'arithmétique est une science expérimentale, il serait dommage de s'en passer. Considérons le 6-uplet admissible $\boldsymbol{h} = (7, 11, 13, 17, 19, 23)$. Pour $X = 10,000,000$, et $1 \leqslant k \leqslant 6$, voici le nombre N_k d'entiers $n \leqslant X$ tels que $\{n+7, \ldots, n+23\}$ contienne au moins k nombres premiers, et pour $6 \leqslant \omega \leqslant 11$, le nombre M_ω d'entiers $n \leqslant X$ tels que $F_{\boldsymbol{h}}(n)$ ait ω facteurs premiers (*sans* multiplicité) :

k	N_k	k	N_k	ω	M_ω	ω	M_ω
1	2,917,127	2	890,130	6	19	7	381
3	160,592	4	18,479	8	3281	9	16955
5	1,092	6	17	10	58947	11	150090

Remarque 3.3. — Concernant l'introduction cruciale du paramètre ℓ, un travail de Heath-Brown [16], lui-même inspiré par des idées de Selberg [22, §18, §23], est cité comme influence dans [11], [12].

Peut-être une raison de l'efficacité de ce paramètre, même petit – sous la conjecture d'Elliott-Halberstam, on prendra $\ell = 1$ –, est-elle à chercher dans le fait que demander que $F_{\boldsymbol{h}}(n)$ ait $k+\ell$ facteurs premiers avec $\ell \neq 0$ n'impose plus, en particulier, que la parité du nombre de facteurs premiers soit constante.

Dans [11], Goldston, Pintz et Yıldırım évaluent donc

$$Q_g = \sum_{N < n \leqslant 2N} \vartheta_{\boldsymbol{h}}(n)\Lambda_{k+\ell}(F_{\boldsymbol{h}}(n), x)^2$$

et

$$Q_p = (\ln 3N) \sum_{N < n \leqslant 2N} \Lambda_{k+\ell}(F_{\boldsymbol{h}}(n), x)^2.$$

(15) Cela découle par récurrence de $\Lambda_1 = \Lambda$ et $\Lambda_{m+1} = (\ln)\Lambda_m + \Lambda \star \Lambda_m$.

(16) Friedlander et Iwaniec ont considérablement développé ce crible asymptotique, en fournissant des variantes qui peuvent traiter le cas $m = 1$ en ajoutant des hypothèses bilinéaires aux conditions classiques, voir [20].

Plutôt que de présenter ces calculs[17], nous allons employer un poids différent qu'ils emploient avec Graham dans [13]. Son avantage est de permettre une évaluation *élémentaire* de Q_g et Q_p, alors que dans [11], il est nécessaire de faire appel à des intégrales multiples de fractions rationnelles en la fonction zéta de Riemann et conséquemment à la zone sans zéro de celle-ci, ce qui est relativement délicat.

L'idée dans [13] est de voir le problème de maximiser Q_g/Q_p comme proche du type de problème d'optimisation caractéristique du crible de Selberg, et c'est d'ailleurs – pour un problème relativement similaire – ce qu'avait fait Selberg lui-même. On cherche alors

$$E(n) = \lambda(F_{\boldsymbol{h}}(n))^2 \quad \text{où} \quad \lambda(n) = \sum_{d|n} \lambda_d,$$

les coefficients λ_d étant à notre disposition, avec la seule contrainte que $\lambda_d = 0$ pour $d > x$, où x (le « niveau », dans le langage du crible) est un paramètre jouant évidemment le même rôle que dans $\Lambda_m(n, x)$ – qui est de la forme postulée ici.

On voit que Q_g et Q_p deviennent des formes quadratiques :

$$Q_g = \sum_{N<n\leqslant 2N} \vartheta_{\boldsymbol{h}}(n) \Big(\sum_{d|F_{\boldsymbol{h}}(n)} \lambda_d\Big)^2, \tag{12}$$

$$Q_p = (\ln 3N) \sum_{N<n\leqslant 2N} \Big(\sum_{d|F_{\boldsymbol{h}}(n)} \lambda_d\Big)^2. \tag{13}$$

La situation typique du crible de Selberg est de minimiser une forme quadratique avec une contrainte linéaire. Ici, le problème de maximiser une forme quadratique par rapport à une autre est différent et n'admet pas de solution « formelle » aussi générale.

Remarque 3.4. — Il est amusant de noter que cette méthode de comparaison de formes quadratiques se trouve dans les travaux de Selberg et Heath-Brown déjà mentionnés, et qu'une variante a aussi été introduite récemment par Soundararajan pour le problème de trouver des grandes valeurs de $|\zeta(\frac{1}{2} + it)|$ ou de $|L(f, \frac{1}{2})|$ lorsque f parcourt une famille de fonctions L. Il y a aussi un goût de la méthode d'amplification de Duke, Friedlander et Iwaniec. Puisqu'il est important de donner un nom aux choses, on peut suggérer « méthode de résonance »[18], ou – plus poétiquement – « méthode d'écrémage »[19] pour ce procédé de détection de valeurs extrêmes.

Dans son cas, Selberg avait procédé en minimisant la « petite » forme quadratique sous la contrainte $\lambda_1 = 1$, puis avait évalué la « grande » forme quadratique pour ce choix de valeurs. La stratégie dans [13] est légèrement différente ; précisément, on obtient un $\lambda_{d,0}$ de cette manière, mais ensuite il est modifié « brutalement » pour incorporer la souplesse du paramètre ℓ. (Le choix initial correspondant en effet à $\ell = 0$). Il est utile de préciser que ce paramètre est spécifique à la situation des écarts entre nombres premiers, et n'a pas d'analogue dans les autres cas mentionnés.

[17] Les résultats sont asymptotiquement identiques à ceux qui sont obtenus ci-dessous.
[18] Où les coefficients formeraient un résonateur ; Soundararajan utilise ce nom.
[19] Où les coefficients formeraient un écrémoir.

4. CHOIX ET ESTIMATION DES DEUX FORMES QUADRATIQUES

Considérons d'abord de manière assez générale une suite $\mathcal{A} = (a_n)$ de réels positifs tels que la série $\sum a_n$ converge, et un k-uplet admissible $\boldsymbol{h}$. On va étudier l'expression

$$Q = \sum_n a_n \Big(\sum_{d \mid F_{\boldsymbol{h}}(n)} \lambda_d \Big)^2, \tag{14}$$

vue comme forme quadratique en les variables λ_d, $d < x$.

On fait l'hypothèse que λ_d soit supportée sur les entiers $d < x$ *sans facteurs carrés.* Cette restriction est justifiée par l'expérience du crible de Selberg, et simplifie techniquement les manipulations qui vont suivre.[20]

Dans cette section, on suppose qu'un analogue de (15) est valide, sous la forme suivante :

$$\sum_{n \equiv t \ (d)} a_n = g_t(d) X + r_d(\mathcal{A}, t), \tag{15}$$

pour $d \geqslant 1$ et t modulo d, où les fonctions arithmétiques $d \mapsto g_t(d)$ sont multiplicatives et de plus $g_t(d)$ ne dépend que de t modulo d. Pour commencer, cette décomposition est formelle, mais on verra plus bas quelles hypothèses sont faites sur les restes $r_d(\mathcal{A}, t)$ pour les contrôler.

Exemple 4.1. — Pour les applications en vue, on considèrera essentiellement deux cas de suites (a_n), supportées sur $N < n \leqslant 2N$ où $N \geqslant 1$ est un entier :

$$a_n = 1, \quad X = N, \quad g_t(d) = \frac{1}{d},$$

$$a_n = \Lambda(n+j), \quad X = N, \quad g_t(d) = \begin{cases} \dfrac{1}{\varphi(d)} & \text{si } (d, t+j) = 1, \\ 0 & \text{sinon} \end{cases}$$

(avec j fixé).

Remarque 4.2. — Dans ce qui suit, on peut considérer le k-uplet $\boldsymbol{h}$ *fixé* pour le besoin du Théorème 1.2 ; autrement dit, les « constantes implicites » peuvent alors dépendre de $\boldsymbol{h}$ (et le Théorème A.1 de l'Appendice peut remplacer le Théorème A.2). Pour le Théorème 1.1, inconditionnel, les termes de reste doivent être *uniformes* par rapport à $\boldsymbol{h}$, pour k et ℓ fixés ; on supposera que $\boldsymbol{h}$ vérifie

$$0 \leqslant h_i \leqslant h \tag{16}$$

et ces estimations uniformes seront exprimées en fonction du paramètre h, que l'on supposera être $\geqslant 10$ (pour que $\log\log h > 0$).

On commence par une transformation formelle de la forme quadratique.

[20] En particulier, une fonction arithmétique f multiplicative supportée sur les entiers sans facteurs carrés est caractérisée par $f(p)$ pour p premier, avec

$$f(d) = \mu^2(d) \prod_{p \mid d} f(p)$$

et l'inverse de f pour la convolution arithmétique vérifie $f^{\star(-1)}(d) = \mu(d) f(d)$ pour d sans facteur carré.

Lemme 4.3. — *Avec les notations et hypothèses ci-dessus, on a*

$$Q = XH + R$$

où

$$H = \sum_{d,e<x} g_{\boldsymbol{h}}([d,e])\lambda_d\lambda_e, \qquad R = \sum_{d,e<x} \lambda_d\lambda_e R_{[d,e]}(\mathcal{A},\boldsymbol{h}),$$

avec

$$g_{\boldsymbol{h}}(d) = \sum_{\substack{\nu \,(\mathrm{mod}\ d) \\ F_{\boldsymbol{h}}(\nu)=0}} g_\nu(d), \qquad R_d(\mathcal{A},\boldsymbol{h}) = \sum_{\substack{\nu \,(\mathrm{mod}\ d) \\ F_{\boldsymbol{h}}(\nu)=0}} r_d(\mathcal{A},\nu)$$

pour tout d sans facteurs carrés.

Démonstration. — Il suffit de développer le carré et d'inverser l'ordre de sommation, ce qui fait apparaître la somme intérieure

$$\sum_{\substack{n \\ F_{\boldsymbol{h}}(n)\equiv 0 \ ([d,e])}} a_n = \sum_{\substack{\nu \,(\mathrm{mod}\ d) \\ F_{\boldsymbol{h}}(\nu)=0}} \ \sum_{n\equiv \nu \ ([d,e])} a_n,$$

et il ne reste plus qu'à appliquer (15) à la somme sur n. □

Puisque $g_\nu(d)$ ne dépend que de ν modulo d, le théorème chinois implique que $d \mapsto g_{\boldsymbol{h}}(d)$ est encore multiplicative. La forme quadratique H est typique dans le crible de Selberg. Le reste R, pour sa part, fait intervenir les quantités $r_d(\mathcal{A},\nu)$ et pas seulement $r_d(\mathcal{A})$; pour le contrôler il est donc nécessaire de comprendre la répartition de $\mathcal{A}$ dans des progressions arithmétiques de module $[d,e] < x^2$ pour des classes de congruence non fixées, et c'est de là que vient l'insuffisance de l'exposant de répartition au sens faible.

Quoi qu'il en soit, suivant le synopsis du crible de Selberg, on procède à la diagonalisation de H.

Lemme 4.4. — *Pour toute fonction multiplicative $d \mapsto g(d)$, on a*

$$\sum_{d,e<x} g([d,e])\lambda_d\lambda_e = \sum_r{}^{\flat}\, t(r) y_r^2$$

où on a, pour tout r sans facteurs carrés,

$$t(r) = \sum_{ab=r} g(a)\mu(b)g(b)^2 = g(r)\prod_{p\mid r}(1-g(p)), \qquad et \qquad y_r = \sum_d g(d)\lambda_{rd}.$$

De plus, on a la formule d'inversion

$$\lambda_d = \sum_r \mu(r)g(r)y_{rd}. \tag{17}$$

Démonstration. — On pose $a = (d,e)$, donc $[d,e] = a\frac{d}{a}\frac{e}{a}$. Pour d, e sans facteurs carrés on a $(a, d/a) = (a, e/a) = 1$. On somme sur a d'abord, puis sur d et e divisibles par a et premiers entre eux. Cette condition $(d,e) = 1$ est détectée à l'aide de la fonction de Möbius

$$\sum_{b\mid (d,e)} \mu(b) = \begin{cases} 1 & \text{si } (d,e) = 1, \\ 0 & \text{sinon,} \end{cases}$$

et il vient

$$\begin{aligned}\sum_{d,e<x} g([d,e])\lambda_d\lambda_e &= \sum_a g(a) \sum_{\substack{d,e<x/a\\(d,e)=1}} g(d)g(e)\lambda_{ad}\lambda_{ae}\\ &= \sum_a g(a)\sum_b \mu(b)g(b)^2 \sum_{d,e<x/ab} g(d)g(e)\lambda_{abd}\lambda_{abe}\\ &= \sum_r{}^{\flat}\, t(r)y_r^2.\end{aligned}$$

Par inversion de Möbius on retrouve λ_d en fonction des variables y_r à l'aide de l'inverse de convolution j de g, à savoir

$$\lambda_d = \sum_r{}^{\flat}\, j(r)y_{rd} = \sum_r \mu(r)g(r)y_{rd}$$

puisque $j(r) = \mu(r)g(r)$ pour r sans facteurs carrés. □

Il est facile de minimiser H sous une contrainte linéaire. Ici, nous le faisons pour le cas $a_n = 1$ pour $N < n \leqslant 2N$. Dans ce cas, $X = N$, $g(d) = 1/d$ et $g_{\boldsymbol{h}}(d) = \nu_{\boldsymbol{h}}(d)/d$ où $\nu_{\boldsymbol{h}}(d)$ est le nombre de racines de $F_{\boldsymbol{h}}$ modulo d, ce qui coïncide avec la description de $\nu_{\boldsymbol{h}}(p)$ dans le Théorème 2.7. On note $t_{\boldsymbol{h}}(r)$ la fonction apparaissant dans le changement de variable $(\lambda_d) \mapsto (y_r)$:

$$t_{\boldsymbol{h}}(r) = g_{\boldsymbol{h}}(r)\prod_{p|r}(1-g_{\boldsymbol{h}}(p)) = \frac{\nu_{\boldsymbol{h}}(r)}{r}\prod_{p|r}\Big(1-\frac{\nu_{\boldsymbol{h}}(p)}{p}\Big). \tag{18}$$

La contrainte linéaire la plus simple est $\lambda_1 = 1$. On déduit alors de l'inégalité de Cauchy-Schwarz que si $\sum \mu(r)g_{\boldsymbol{h}}(r)y_r = \lambda_1 = 1$, on a

$$\begin{aligned}1 = \Big(\sum_r \mu(r)g_{\boldsymbol{h}}(r)y_r\Big)^2 &\leqslant \Big(\sum_r{}^{\flat}\frac{g_{\boldsymbol{h}}(r)^2}{t_{\boldsymbol{h}}(r)}\Big)\Big(\sum_r{}^{\flat}\, t_{\boldsymbol{h}}(r)y_r^2\Big)\\ &= \tilde{H}H(\lambda), \qquad \text{avec } \tilde{H} = \sum_r{}^{\flat}\frac{g_{\boldsymbol{h}}(r)^2}{t_{\boldsymbol{h}}(r)},\end{aligned}$$

l'égalité étant obtenue, et donc H minimimale, lorsque

$$y_r = \frac{1}{\tilde{H}}\frac{\mu(r)g_{\boldsymbol{h}}(r)}{t_{\boldsymbol{h}}(r)}, \text{ pour } 1 \leqslant r < x.$$

Comme expliqué dans la précédente section, on modifie ce choix en insérant un facteur logarithme visant à « mimer » la $(k+\ell)$-ème fonction de von Mangoldt. On pose

$$y_{r,\ell} = \frac{\mu(r)g_{\boldsymbol{h}}(r)}{t_{\boldsymbol{h}}(r)}\Big(\ln\frac{x}{r}\Big)^{\ell} = \mu(r)\prod_{p|r}\Big(1-\frac{\nu_{\boldsymbol{h}}(p)}{p}\Big)^{-1}\Big(\ln\frac{x}{r}\Big)^{\ell} \tag{19}$$

pour $r < x$, $\ell \geqslant 1$; on a pu omettre le facteur $1/\tilde{H}$ pour simplifier, en raison de l'homogénéité des formules considérées. Les deux formes quadratiques Q_g et Q_p vont maintenant être évaluées asymptotiquement pour ce choix des variables y_r (et le choix correspondant de λ_d).

Commençons par le terme « principal » H_p correspondant à Q_p.

Théorème 4.5. — *Soit $h \geqslant 3$ un entier, $\boldsymbol{h}$ un k-uplet admissible vérifiant* (16), *$\ell \geqslant 0$ un entier. Soit*

$$H_p = \sum_{d,e<x} g_{\boldsymbol{h}}([d,e])\lambda_d\lambda_e = \sum_{d,e<x} \frac{\nu_{\boldsymbol{h}}([d,e])}{[d,e]}\lambda_d\lambda_e = \sum_r{}^{\flat}\, t_{\boldsymbol{h}}(r)y_r^2$$

la forme quadratique ci-dessus. Lorsque (λ_d) est donné par (19), *on a*

$$H_p(y_{r,\ell}) = \frac{(2\ell)!}{(k+2\ell)!}\frac{1}{\mathfrak{S}(\boldsymbol{h})}(\log x)^{k+2\ell}\Big(1 + O\Big(\frac{\mathfrak{S}(\boldsymbol{h})(\ln\ln h)}{\ln x}\Big)\Big),$$

pour $x \geqslant 2$, la constante implicite ne dépendant que de k et ℓ.

Démonstration. — Il s'agit d'évaluer

$$H_p(y_{r,\ell}) = \sum_{r<x}{}^{\flat}\, t_{\boldsymbol{h}}(r)y_{r,\ell}^2 = \sum_{r<x}{}^{\flat}\, \frac{g_{\boldsymbol{h}}(r)^2}{t_{\boldsymbol{h}}(r)}\Big(\ln\frac{x}{r}\Big)^{2\ell}. \tag{20}$$

Pour $\ell = 0$, il s'agit de la moyenne d'une fonction multiplicative sur $r < x$, qui est un sujet bien balisé, et le facteur lisse $(\ln x/r)^{2\ell}$ peut être traité par sommation par parties.

Plus précisément, on a pour r sans facteurs carrés

$$\frac{g_{\boldsymbol{h}}(r)^2}{t_{\boldsymbol{h}}(r)} = \frac{\nu_{\boldsymbol{h}}(r)}{r}\prod_{p\mid r}\Big(1 - \frac{\nu_{\boldsymbol{h}}(p)}{p}\Big)^{-1} = \prod_{p\mid r}\frac{\nu_{\boldsymbol{h}}(p)}{p-\nu_{\boldsymbol{h}}(p)}. \tag{21}$$

On a

$$\sum_{p<x}\frac{\nu_{\boldsymbol{h}}(p)}{p}\ln p = k\ln x + O(1)$$

et comme $\nu_{\boldsymbol{h}}(p) < p$, on en déduit

$$\sum_{p<x}\frac{\nu_{\boldsymbol{h}}(p)}{p-\nu_{\boldsymbol{h}}(p)}\ln p = k\ln x + O(1),$$

pour $x \geqslant 2$, la constante implicite dépendant de k. Cela permet d'appliquer le Théorème A.2 de l'Appendice ; après avoir vérifié que l'on peut prendre $L \ll \ln\ln h$ dans (31), on termine la preuve en vérifiant que la constante c de loc. cit. est bien égale à $1/\mathfrak{S}(\boldsymbol{h})$. □

Passons à l'évaluation du terme principal de Q_g. Il s'agit donc de calculer

$$\sum_{i=1}^{k}\sum_{N<n\leqslant 2N}\Lambda(n+h_i)\Big(\sum_{d\mid F_{\boldsymbol{h}}(n)}\lambda_d\Big)^2,$$

et on le fait pour chaque i séparément. Pour la suite $a_n = \Lambda(n+h_i)$, i fixé, $N < n \leqslant 2N$, on a (15) avec $X = N$; dans le Lemme 4.3, on a

$$\sum_{F_{\boldsymbol{h}}(\nu)\equiv 0\ (p)} g_\nu(p) = \frac{\nu_{\boldsymbol{h}}(p)-1}{p-1},$$

pour p premier (voir Exemple 4.1), et donc le rôle de $g_{\boldsymbol{h}}(d)$ est joué ici par

$$g^{\sharp}_{\boldsymbol{h}}(d) = \frac{1}{\varphi(d)}\prod_{p\mid d}(\nu_{\boldsymbol{h}}(p)-1), \tag{22}$$

pour d sans facteurs carrés.

Théorème 4.6. — *Avec les notations et hypothèses ci-dessus, et avec H_g la forme quadratique du Lemme 4.3 pour la suite $a_n = \Lambda(n+h_i)$ pour $N < n \leqslant 2N$, on a*

$$H_g(y_{r,\ell}) = \frac{1}{(\ell+1)^2}\frac{(2\ell+2)!}{(k+2\ell+1)!}\frac{1}{\mathfrak{S}(\boldsymbol{h})}(\ln x)^{k+2\ell+1}\Big(1+O\Big(\frac{\mathfrak{S}(\boldsymbol{h})(\ln\ln h)}{\ln x}\Big)\Big),$$

pour $x \geqslant 2$, où la constante implicite ne dépend que de k et ℓ.

Démonstration. — L'idée est d'appliquer le Lemme 4.4 à H_g, et d'exprimer les nouvelles variables, disons z_r, qui y apparaissent, en fonction des variables y_r qui interviennent dans la diagonalisation de H_p. Cela permet d'abord d'évaluer z_r, puis ensuite H_g. Certaines simplifications viennent faciliter le travail, mais c'est essentiellement un exercice (tant que $\boldsymbol{h}$ est fixé).

Tout d'abord, on a la diagonalisation

$$H_g = \sum_r{}^{\flat}\, t^\sharp_{\boldsymbol{h}}(r) z_r^2$$

avec

$$z_r = \sum_d g^\sharp_{\boldsymbol{h}}(d)\lambda_{rd} \quad \text{et} \quad t^\sharp_{\boldsymbol{h}}(r) = g^\sharp_{\boldsymbol{h}}(r)\prod_{p\mid r}(1-g^\sharp_{\boldsymbol{h}}(p)).$$

D'après l'expression (17) qui fournit λ_d en fonction des y_r, on a donc

$$z_r = \sum_b\Big(\sum_{de=b} g^\sharp_{\boldsymbol{h}}(d)\mu(e)g_{\boldsymbol{h}}(e)\Big)y_{br}.$$

Pour $b=p$, la fonction multiplicative qui apparaît est égale à

$$g^\sharp_{\boldsymbol{h}}(p) - g_{\boldsymbol{h}}(p) = \frac{\nu_{\boldsymbol{h}}(p)-1}{p-1} - \frac{\nu_{\boldsymbol{h}}(p)}{p} = -\frac{p-\nu_{\boldsymbol{h}}(p)}{p(p-1)}.$$

D'après (19) on trouve pour $(y_r) = (y_{r,\ell})$:

$$z_r = \sum_{b<x/r}\frac{\mu(b)}{\varphi(b)}\prod_{p\mid b}\Big(1-\frac{\nu_{\boldsymbol{h}}(p)}{p}\Big)\frac{\mu(br)g_{\boldsymbol{h}}(br)}{t_{\boldsymbol{h}}(br)}\Big(\ln\frac{x}{br}\Big)^{\ell}.$$

Les termes qui contribuent vérifient $(b,r)=1$, ce qui permet de séparer le produit br. En utilisant (18) on trouve

$$z_r = \frac{\mu(r)g_{\boldsymbol{h}}(r)}{t_{\boldsymbol{h}}(r)}\sum_{\substack{b<x/r\\(r,b)=1}}{}^{\flat}\,\frac{1}{\varphi(b)}\Big(\ln\frac{x}{br}\Big)^{\ell}.$$

Une application très simple du Théorème A.2 (avec $\kappa=1$) donne

$$z_r = \frac{1}{\ell+1}\frac{\mu(r)g_{\boldsymbol{h}}(r)}{t_{\boldsymbol{h}}(r)}\frac{\varphi(r)}{r}\Big(\ln\frac{x}{r}\Big)^{\ell+1} + O\Big((\ln\ln r)\Big(\ln\frac{x}{r}\Big)^{\ell}\Big).$$

Le calcul de $H_g(y_{r,\ell})$ est de nouveau une application de ce même résultat. On a

$$t^\sharp_{\boldsymbol{h}}(r)z_r^2 = \frac{1}{(\ell+1)^2}t^\sharp_{\boldsymbol{h}}(r)\frac{g_{\boldsymbol{h}}(r)^2}{t_{\boldsymbol{h}}(r)^2}\Big(\frac{\varphi(r)}{r}\Big)^2\Big(\ln\frac{x}{r}\Big)^{2\ell+2} + (\text{reste}).$$

Pour $r=p$ premier, il vient

$$\begin{aligned} t^\sharp_{\boldsymbol{h}}(r)\frac{g_{\boldsymbol{h}}(r)^2}{t_{\boldsymbol{h}}(r)^2}\Big(\frac{\varphi(r)}{r}\Big)^2 &= \frac{\nu_{\boldsymbol{h}}(p)-1}{p-1}\Big(1-\frac{\nu_{\boldsymbol{h}}(p)-1}{p-1}\Big)\Big(1-\frac{\nu_{\boldsymbol{h}}(p)}{p}\Big)^{-2}\Big(\frac{p-1}{p}\Big)^2 \\ &= \frac{\nu_{\boldsymbol{h}}(p)-1}{p-\nu_{\boldsymbol{h}}(p)}. \end{aligned}$$

Le Théorème A.1 (avec $\kappa = k-1$ maintenant) et de nouvelles estimations, longues mais élémentaires, de la contribution du reste donnent

$$H_g(y_{r,\ell}) = \frac{1}{(\ell+1)^2}\frac{(2\ell+2)!}{(k+2\ell+1)!}\frac{1}{\mathfrak{S}(\boldsymbol{h})}(\ln x)^{k+2\ell+1}\Big(1+O\Big(\frac{\mathfrak{S}(\boldsymbol{h})(\ln\ln h)}{\ln x}\Big)\Big),$$

car la constante c de loc. cit. est (encore !)

(23) $$c = \prod_p \Big(1-\frac{1}{p}\Big)^{k-1}\Big(1+\frac{\nu_{\boldsymbol{h}}(p)-1}{p-\nu_{\boldsymbol{h}}(p)}\Big) = \prod_p \Big(1-\frac{1}{p}\Big)^{k-1}\frac{p-1}{p-\nu_{\boldsymbol{h}}(p)} = \frac{1}{\mathfrak{S}(\boldsymbol{h})}.$$

□

Pour être prêt à démontrer le Théorème 1.2, il ne reste qu'à contrôler les termes de reste provenant du Lemme 4.3. On suppose pour cela que $\mathcal{A}$ s'obtient par « localisation dyadique » d'une suite infinie $\mathbf{A}$, c'est à dire que pour une suite $\mathbf{A} = (a_n)$ vérifiant (4), on pose $\mathcal{A} = (a_n)_{N<n\leqslant 2N}$ pour un certain $N \geqslant 1$. On écrit alors (15) avec les choix « évidents » des paramètres, provenant de (4), en particulier :

(24) $$r_d(\mathcal{A},t) = r_d(\mathcal{A};2N,t) - r_d(\mathcal{A};N,t).$$

Lemme 4.7. — *Soit $\boldsymbol{h}$ un k-uplet. Soit $\mathbf{A} = (a_n)$ une suite de réels positifs telle que $\mathbf{A}$ ait un P-exposant de répartition $\geqslant \theta$, par exemple telle que $\mathbf{A}$ ait un exposant de répartition au sens fort $\geqslant \theta$. Supposons de plus que (λ_d) vérifie*

(25) $$|\lambda_d| \leqslant \tau(d)^{C_1}(\ln 2x)^{C_2}$$

pour tout $d < x$, et que

(26) $$|r_d(\mathbf{A};N,\nu)| \leqslant d^{-1}\tau(d)^{C_3}N(\ln 2N)^{C_4}$$

pour $N \geqslant 1$, d sans facteurs carrés et ν modulo d, $C_1,\ldots,C_4$ étant des constantes $\geqslant 0$.

Alors pour $\mathcal{A} = (a_n)_{N<n\leqslant 2N}$, et avec $r_d(\mathcal{A},t)$ défini comme ci-dessus, on a

$$R = \sum_{d,e<x} \lambda_d\lambda_e R_{[d,e]}(\mathcal{A},\boldsymbol{h}) \ll N(\ln N)^{-A}$$

pour tout $x \geqslant 2$ tel que $x^2 < N^{\theta-\varepsilon}$ avec $\varepsilon > 0$, et tout $A \geqslant 1$. La constante implicite dépend de $(\mathbf{A},\varepsilon,A,C_1,C_2,C_3,C_4)$.

Démonstration. — On a

$$|R| \leqslant \sideset{}{^\flat}\sum_{d<x^2} \tau_x(d) \sum_{\substack{\nu \pmod d\\ F_{\boldsymbol{h}}(\nu)=0}} |r_d(\mathcal{A},\nu)|$$

où

(27) $$\tau_x(d) = \sum_{[a,b]=d} |\lambda_a\lambda_b| \ll \tau(d)^{2+2C_1}(\ln 2x)^{2C_2}.$$

On applique alors l'inégalité de Cauchy-Schwarz pour obtenir

$$|R|^2 \leqslant \Big(\sideset{}{^\flat}\sum_{d<x^2} \tau_x(d)^2 \sum_{\substack{\nu \pmod d\\ F_{\boldsymbol{h}}(\nu)=0}} |r_d(\mathcal{A},\nu)|\Big)\Big(\sideset{}{^\flat}\sum_{d<x^2} \sum_{\substack{\nu \pmod d\\ F_{\boldsymbol{h}}(\nu)=0}} |r_d(\mathcal{A},\nu)|\Big).$$

On borne $|r_d(\mathcal{A},\nu)|$ à l'aide de (24) ; pour le premier facteur, on applique (26) et (27) ; la somme avec des fonctions diviseurs et des logarithmes fait « perdre » un facteur de type $(\ln N)^{C_5}$ avec C_5 une constante assez grande. L'hypothèse $x^2 < N^{\theta-\varepsilon}$ permet d'appliquer au second facteur la borne (5) définissant le P-exposant de répartition, et cette puissance de logarithme égarée est « récupérée » avec toute la marge nécessaire. $\square$

Noter que (26) n'est pas une condition onéreuse ; par exemple si $g_\nu(p) \leqslant kp^{-1}$ et a_n est bornée par une puissance de logarithme, elle sera aussitôt vérifiée. De même pour l'estimation individuelle (25). Pour les λ_d discutés précédemment, on a

Lemme 4.8. — *Soit* $\boldsymbol{h}$ *un* k*-uplet admissible,* $\ell \geqslant 0$ *et* (λ_d) *donné par* (19) *et* (17) *pour* $d < x$. *On a alors*

$$\lambda_d \ll \frac{1}{\mathfrak{S}(\boldsymbol{h})}(\ln x)^{k+\ell}$$

pour tout $x \geqslant 2$ *et* $d \geqslant 2$ *sans facteurs carrés, la constante implicite ne dépendant que de* k *et* ℓ.

Voir [13, Lemma 9] ; pour $\boldsymbol{h}$ fixé, c'est à peu près évident.

Les lecteurs sont encouragés à passer à la section suivante où les résultats ci-dessus sont appliqués à la preuve du Théorème 1.2. Mais pour celle (inconditionnelle) du Théorème 1.1, un ingrédient supplémentaire similaire à ce qui précède est requis ; il s'agit de considérer la forme quadratique (14) avec $a_n = \Lambda(n+j)$ pour $N < n \leqslant 2N$ et $j \geqslant 0$, mais cette fois sous l'hypothèse que j *n'est pas* une composante de $\boldsymbol{h}$. On note Q_i cette somme. Comme il s'agit de la même suite a_n que pour Q_g on a la même forme de (15) que dans ce cas. On écrit donc

$$Q_i = NH_i + R_i$$

selon le Lemme 4.3. Si l'on note $\boldsymbol{j}$ le $(k+1)$-uplet $(\boldsymbol{h}, j)$, on s'aperçoit alors que H_i est donnée par

$$H_i = \sum_{d,e<x} g^\sharp_{\boldsymbol{j}}([d,e])\lambda_d\lambda_e,$$

où la fonction $d \mapsto g^\sharp_{\boldsymbol{j}}(d)$ est définie par (22), pour $\boldsymbol{j}$ au lieu de $\boldsymbol{h}$.

Théorème 4.9. — *Soit* $h \geqslant 3$, $\boldsymbol{h}$ *un* k*-uplet admissible,* $j \geqslant 0$ *un entier qui n'est pas une composante de* h *tel que* $\boldsymbol{j} = (\boldsymbol{h}, j)$ *soit admissible et vérifie* (16). *Soit* $\ell \geqslant 1$ *un entier et* (λ_d) *donné par* (19), *appliqué à* $\boldsymbol{h}$ *toujours. On a alors*

$$H_i(y_{r,\ell}) = \frac{(2\ell)!}{(k+2\ell)!}\frac{\mathfrak{S}(\boldsymbol{j})}{\mathfrak{S}(\boldsymbol{h})^2}(\ln x)^{k+2\ell}\Big(1 + O\Big(\frac{\mathfrak{S}(\boldsymbol{h})(\ln\ln h)}{\ln x}\Big)\Big),$$

pour $x \geqslant 2$, *la constante implicite ne dépendant que de* k *et* ℓ.

Démonstration. — Formellement, cela ressemble beaucoup à ce qui précède. Le Lemme 4.4 permet de diagonaliser

$$H_i = \sum_r{}^\flat\, t^\sharp_{\boldsymbol{j}}(r) w_r^2$$

avec

$$w_r = \sum_d g^\sharp_{\boldsymbol{j}}(d)\lambda_{rd} = \sum_b \Big(\sum_{de=b} g^\sharp_{\boldsymbol{j}}(d)\mu(e)g_{\boldsymbol{h}}(d)\Big)y_{br}.$$

En insérant la valeur de $y_{r,\ell}$, on trouve

$$w_r = \frac{\mu(r)g_{\boldsymbol{h}}(r)}{t_{\boldsymbol{h}}(r)} \sum_{\substack{b<x/r\\(r,b)=1}} \frac{\mu(b)g_{\boldsymbol{h}}(b)}{t_{\boldsymbol{h}}(b)}\Big(\sum_{de=b} g^\sharp_{\boldsymbol{j}}(d)\mu(e)g_{\boldsymbol{h}}(d)\Big)\Big(\ln\frac{x}{br}\Big)^\ell.$$

Le point où le calcul diffère du précédent est que la série

$$\sum_{\substack{b\geqslant 1\\(r,b)=1}}\frac{\mu(b)g_{\boldsymbol{h}}(b)}{t_{\boldsymbol{h}}(b)}\Big(\sum_{de=b}g^{\sharp}_{\boldsymbol{j}}(d)\mu(e)g_{\boldsymbol{h}}(d)\Big)$$

converge ; en effet, en calculant formellement (puis en justifiant sans peine), elle est égale à

$$\prod_{p\nmid r}\Big(1-\frac{1}{1-g_{\boldsymbol{h}}(p)}(g^{\sharp}_{\boldsymbol{j}}(p)-g_{\boldsymbol{h}}(p))\Big)=\prod_{p\nmid r}\frac{1-g^{\sharp}_{\boldsymbol{j}}(p)}{1-g_{\boldsymbol{h}}(p)},$$

dont la convergence provient du fait que $g^{\sharp}_{\boldsymbol{j}}(p)$ et $g_{\boldsymbol{h}}(p)$ sont toute deux de l'ordre de k/p en moyenne.

Il n'est pas très difficile de déduire de cela que

$$\begin{aligned}
w_r&=\frac{\mu(r)g_{\boldsymbol{h}}(r)}{t_{\boldsymbol{h}}(r)}\prod_{p\nmid r}\frac{1-g^{\sharp}_{\boldsymbol{j}}(p)}{1-g_{\boldsymbol{h}}(p)}\Big(\ln\frac{x}{r}\Big)^{\ell}+(\text{reste}),\\
&=\mu(r)\prod_{p\mid r}\frac{1}{1-g_{\boldsymbol{h}}(p)}\prod_{p\nmid r}\frac{1-g^{\sharp}_{\boldsymbol{j}}(p)}{1-g_{\boldsymbol{h}}(p)}\Big(\ln\frac{x}{r}\Big)^{\ell}+(\text{reste}),\\
&=\mu(r)\prod_{p\mid r}(1-g^{\sharp}_{\boldsymbol{j}}(p))^{-1}\frac{\mathfrak{S}(\boldsymbol{j})}{\mathfrak{S}(\boldsymbol{h})}\Big(\ln\frac{x}{r}\Big)^{\ell}+(\text{reste})\\
&=\frac{\mu(r)g^{\sharp}_{\boldsymbol{j}}(r)}{t^{\sharp}_{\boldsymbol{j}}(r)}\frac{\mathfrak{S}(\boldsymbol{j})}{\mathfrak{S}(\boldsymbol{h})}\Big(\ln\frac{x}{r}\Big)^{\ell}+(\text{reste}).
\end{aligned}$$

Finalement, on a

$$\sum_{r}{}^{\flat}\, t^{\sharp}_{\boldsymbol{j}}(r)w_r^2=\frac{\mathfrak{S}(\boldsymbol{j})^2}{\mathfrak{S}(\boldsymbol{h})^2}\sum_{r<x}{}^{\flat}\,\frac{g^{\sharp}_{\boldsymbol{j}}(r)^2}{t^{\sharp}_{\boldsymbol{j}}(r)}\Big(\ln\frac{x}{r}\Big)^{2\ell}+(\text{reste}).$$

Comme dans le Théorème 4.5 (voir (20) et (21)), on a d'après le Théorème A.2

$$\sum_{r}{}^{\flat}\, t^{\sharp}_{\boldsymbol{j}}(r)w_r^2=\frac{\mathfrak{S}(\boldsymbol{j})^2}{\mathfrak{S}(\boldsymbol{h})^2}\frac{1}{\mathfrak{S}(\boldsymbol{j})}\frac{(2\ell)!}{(k+2\ell)!}\Big(\ln\frac{x}{r}\Big)^{2\ell+1}+(\text{reste})$$

(pour le calcul de la constante c, voir (23) appliqué à $\boldsymbol{j}$), ce qui donne le résultat, au terme de reste près dont il faut s'assurer qu'il est tel que décrit (c'est relativement long mais pas du tout difficile). □

5. PREUVE DU THÉORÈME 1.2

Pour démontrer le Théorème 1.2, il ne reste qu'à combiner les informations obtenues dans la section précédente. Fixons un k-uplet admissible $\boldsymbol{h}$, avec $k\geqslant 2$, et supposons que le P-exposant de répartition de la suite des nombres premiers (donc de ses translatées) soit $\geqslant\theta$. Prenons de plus $\ell\geqslant 1$.

Pour $x = N^\beta$ avec $2\beta < \theta - \varepsilon$, la combinaison des Théorèmes 4.5 et 4.6 (appliqué pour chaque $\Lambda(n+h_i)$, $1 \leqslant i \leqslant k$) et du Lemme 4.7 permet d'affirmer qu'il existe (λ_d) tels que les formes quadratiques Q_g et Q_p définies par (12) et (13) vérifient

$$Q_p = \frac{(2\ell)!}{(k+2\ell)!}\frac{\beta^{k+2\ell}}{\mathfrak{S}(\boldsymbol{h})}N(\ln N)^{k+2\ell+1}\Big(1+O\Big(\frac{1}{\ln N}\Big)\Big),$$
$$Q_g = \frac{k}{(\ell+1)^2}\frac{(2\ell+2)!}{(k+2\ell+1)!}\frac{\beta^{k+2\ell+1}}{\mathfrak{S}(\boldsymbol{h})}N(\ln N)^{k+2\ell+1}\Big(1+O\Big(\frac{1}{\ln N}\Big)\Big)$$

pour tout $N \geqslant 2$, les constantes implicites dépendant ici de k et $\boldsymbol{h}$. Donc

$$(28) \qquad Q_g - Q_p \geqslant \frac{1}{2}\Big\{\frac{2k\beta(2\ell+1)}{(\ell+1)(k+2\ell+1)} - 1\Big\}\frac{\beta^{k+2\ell}}{\mathfrak{S}(\boldsymbol{h})}\frac{(2\ell)!}{(k+2\ell)!}N(\ln N)^{k+2\ell+1},$$

pour N assez grand.

Déjà lorsque $k = 7$, $\ell = 1$, le terme de droite est > 0 si $2\beta > 20/21$, ce qui démontre que la conjecture d'Elliott-Halberstam implique que pour tout 7-uplet admissible $\boldsymbol{h}$, il existe une infinité de n tels que $n \lhd \boldsymbol{h}$. Puisque $\boldsymbol{h} = (11, 13, 17, 19, 23, 29, 31)$ est admissible, on a(21) $\liminf \gamma(n) \leqslant 20$. Pour obtenir la première partie du Théorème 1.2 (où 16 remplace 20), une astuce supplémentaire est requise (on crée une combinaison des cas $\ell = 0$ et $\ell = 1$) ; voir [13, p. 44]. Noter aussi que $\ell = 0$ est toujours insuffisant.

De plus la fraction rationnelle entre accolades dans (28) tend vers $2\theta - 1$ si $k, \ell \to +\infty$ avec $\ell = o(k)$ (par exemple, $\ell = [\sqrt{k}]$). Si $\theta > 1/2$, on peut choisir k et ℓ tels que $Q_g > Q_p$ pour tout N assez grand, ce qui démontre la première partie du Théorème 1.2.(22)

6. PREUVE DU THÉORÈME 1.1

Inconditionnellement, on peut appliquer (28) pour tout $\beta < 1/4$, mais on échoue alors tout juste. Il faut une autre idée, pas du tout évidente non plus, pour passer la difficulté. Il s'agit de donner un petit coup de pouce à Q_g en rajoutant les sommes du type

$$Q_i = \sum_{N<n\leqslant 2N}\Lambda(n+j)\Big(\sum_{d\mid F_{\boldsymbol{h}}(n)}\lambda_d\Big)^2$$

considérées à la fin de la Section 4, où cette fois j n'apparaît pas dans $\boldsymbol{h}$. Il n'y a aucune corrélation alors à espérer entre le fait que $n + j$ soit premier et que $F_{\boldsymbol{h}}(n)$ soit presque premier ; cela signifie, comme on l'a constaté, que le terme principal dans Q_i est plus petit par un facteur logarithmique de celui dans Q_g. Mais toute somme de type Q_i apporte une contribution « infinitésimale » telle que la combinaison d'un grand nombre d'entre elles finit par faire déborder le vase...

Voici le fonctionnement précis de cette idée. On considère $h \geqslant 1$ et on pose

$$\vartheta'(n,h) = \sum_{1\leqslant j\leqslant h}\Lambda(n+j),$$

(21) On doit aussi remarquer que la contribution à $Q_g - Q_p$ des puissances de nombres premiers également détectées par $\vartheta_{\boldsymbol{h}}(n)$ est $\ll N^{1/2+\varepsilon}$ pour tout $\varepsilon > 0$.

(22) Même remarque que ci-dessus.

puis on définit les quantités

$$Q'_g = \frac{1}{h^k} \sideset{}{^*}\sum_{\substack{\boldsymbol{h}=(h_1,\dots,h_k)\\ h_i\leqslant h}} \sum_{N<n\leqslant 2n} \vartheta'(n,h)\Big(\sum_{d\mid F_{\boldsymbol{h}}(n)} \lambda_{d,\boldsymbol{h}}\Big)^2,$$

$$Q'_p = \frac{(\ln 3N)}{h^k} \sideset{}{^*}\sum_{\substack{\boldsymbol{h}=(h_1,\dots,h_k)\\ h_i\leqslant h}} \sum_{N<n\leqslant 2N} \Big(\sum_{d\mid F_{\boldsymbol{h}}(n)} \lambda_{d,\boldsymbol{h}}\Big)^2,$$

le symbole $\sum^*$ signifiant que la somme est restreinte aux k-uplets admissibles. Noter que l'on peut s'autoriser à choisir des coefficients dépendant de $\boldsymbol{h}$, ce qui est important puisque ceux définis dans la section précédente en dépendent effectivement. Précisément, pour tout $\boldsymbol{h}$ on va prendre

$$\lambda_{d,\boldsymbol{h}} = \mathfrak{S}(\boldsymbol{h})\lambda_d \tag{29}$$

où λ_d est donné par (17) et (19) ; l'insertion du facteur $\mathfrak{S}(\boldsymbol{h})$ permet par homogénéité de simplifier la moyenne sur $\boldsymbol{h}$ en la ramenant à celle considérée par Gallagher (Proposition 2.9).

On note encore que si $Q'_g > Q'_p$, il existe n tel que $\vartheta'(n,h) > \ln 2N$, et alors (si $h \leqslant N$) il existe deux (puissances de) nombres premiers $p < p'$, $N < p \leqslant 2N$, tels que $p' - p \leqslant h$.

Théorème 6.1. — *Soit $h \geqslant 1$, $k \geqslant 1$ et $\ell \geqslant 1$ des entiers. Pour chaque $\boldsymbol{h}$, soit $\lambda_{d,\boldsymbol{h}}$ défini par* (29). *On a alors pour tout $\varepsilon > 0$*

$$Q'_p = \frac{(2\ell)!}{(k+2\ell)!} N(\ln x)^{k+2\ell}(\ln 3N)\mathsf{S}(h,k) + O(N(\ln x)^{k+2\ell-1+\varepsilon})$$

et

$$Q'_g = \frac{(2\ell)!}{(k+2\ell)!} N(\ln x)^{k+2\ell+1}\Big\{\frac{2k\mathsf{S}(h,k)}{(\ell+1)(k+2\ell+1)} + \frac{h\mathsf{S}(h,k+1)}{\ln x}\Big\} + O(N(\ln x)^{k+2\ell+\varepsilon})$$

où $\mathsf{S}(h,k)$ est défini par (8)*, pour tout $N \geqslant 2$, $x \geqslant 2$ tel que $h \leqslant \ln x \leqslant \ln N$, la constante implicite dépendant de k, ℓ et ε seulement.*

Démonstration. — Le calcul de chaque terme dans Q'_p provient directement du Théorème 4.5 ; la somme sur $\boldsymbol{h}$ fait apparaître exactement le terme $\mathsf{S}(h,k)$ d'après sa définition.

Si l'on développe $\vartheta'(n,h)$ dans Q'_g, la contribution de chacun des k-uplets $\boldsymbol{h}$ dans lesquels j apparaît est donnée par le Théorème 4.6, et il apparaît le coefficient

$$\sum_{j=1}^{k} \sideset{}{^*}\sum_{\substack{\boldsymbol{h}=(h_1,\dots,h_k)\\ 1\leqslant h_i\leqslant h\\ h_i\neq j}} \mathfrak{S}(\boldsymbol{h}) = k\mathsf{S}(h,k)$$

en inversant l'ordre de sommation.

La dernière contribution est celle des k-uplets $\boldsymbol{h}$ et des j tels que j n'apparaît pas dans $\boldsymbol{h}$. Chaque terme relève du Théorème 4.9, et il apparaît le facteur

$$\sum_{1\leqslant j\leqslant k} \sideset{}{^*}\sum_{\substack{\boldsymbol{h}=(h_1,\dots,h_k)\\ j\neq h_i\leqslant h}} \mathfrak{S}(\boldsymbol{h},j) = \sideset{}{^*}\sum_{\substack{\boldsymbol{j}=(h_1,\dots,h_k,h_{k+1})\\ h_i\leqslant h}} \mathfrak{S}(\boldsymbol{j}) = \mathsf{S}(h,k+1).$$

□

On conclut alors de la manière suivante : en faisant la différence de $Q'_g - Q'_p$ avec $x = N^\beta$ le terme principal est maintenant

$$\Xi \frac{(2\ell)!}{(k+2\ell)!} \beta^{k+2\ell} N (\ln N)^{k+2\ell+1},$$

où

$$\Xi = \frac{2k\beta(2\ell+1)\mathsf{S}(h,k)}{(\ell+1)(k+2\ell+1)} + \frac{h\mathsf{S}(h,k+1)}{\ln N} - 1.$$

Puisque pour k fixé on a $\mathsf{S}(h,k) \to 1$ et $\mathsf{S}(h,k+1) \to 1$, d'après le résultat de Gallagher (Proposition 2.9), on a approximativement[23]

$$\Xi \approx \Xi' = \frac{2k\beta(2\ell+1)}{(\ell+1)(k+2\ell+1)} + \frac{h}{\ln N} - 1,$$

et lorsque h est grand, cela montre qu'on a gagné un facteur $h/\ln N$ par rapport au cas précédent (voir (28)). Puisque ce dernier argument échouait tout juste, il suffit que $h/\ln N$ soit strictement positif pour tout N assez grand (mais éventuellement arbitrairement petit) pour pouvoir conclure.

Précisément, fixons $\delta > 0$ arbitrairement petit et posons $h = \delta \ln N$, pour $N \geqslant \exp(\delta^{-1})$. La limite de Ξ' quand $k, \ell \to +\infty$ avec $\ell^2 \leqslant k$ est alors égale à $4\beta - 1 + \delta$. On peut trouver $\beta < 1/4$ tel que $4\beta + \delta > 1$, et cela fait il existe k et ℓ (dépendant seulement de δ et β) tels que $\Xi' > 0$ pour tout N suffisamment grand (N est maintenant la seule variable). On en déduit par la Proposition 2.9 que $\Xi > 0$ également pour tout N assez grand puisque $h \to +\infty$ lorsque $N \to +\infty$.

Faisant finalement appel au Théorème de Bombieri-Vinogradov pour estimer comme précédemment les termes de reste, on peut donc conclure que pour tout N assez grand, on a

$$Q'_g - Q'_p \geqslant \frac{1}{2} \Xi \frac{(2\ell)!}{(k+2\ell)!} \beta^{k+2\ell} N (\ln N)^{k+2\ell+1} > 0.$$

Par conséquent, il existe une infinité d'entiers tels que $\gamma(n) \leqslant h \leqslant \delta \ln N$, ce qui signifie que

$$\liminf_{n \to +\infty} \frac{\gamma(n)}{\ln n} \leqslant \delta.$$

Finalement, le Théorème 1.1 est obtenu puisque δ est arbitrairement petit.

Remarque 6.2. — Il est à remarquer que dans cette preuve *l'uniformité* des estimations de Q'_p et Q'_g par rapport au k-uplet $\boldsymbol{h}$ est absolument essentielle.

7. AUTRES RÉSULTATS

La méthode utilisée par Goldston, Pintz et Yıldırım est robuste et a un fort potentiel d'applications au delà de la preuve des Théorèmes 1.1 et 1.2. Nous allons ici simplement mentionner quelques-uns des résultats annoncés, référant aux prépublications disponibles et à venir pour plus de détails.

[23] Le symbole $\approx$ est heuristique seulement.

En premier lieu, dans [13], Goldston, Graham, Pintz et Yıldırım démontrent que l'exposant de répartition $1/2$ n'est pas nécessairement une barrière intrinsèque : concernant les écarts entre entiers $n = p_1 p_2$ avec p_i des nombres premier distincts,[24] ils démontrent :

Théorème 7.1. — *Soient*

$$e_1 = 6 < e_2 = 10 < e_3 = 15 < \cdots < e_{278} = 7 \cdot 137 < \cdots < e_n < \cdots$$

la suite des entiers qui sont produit de deux facteurs premiers distincts. Alors on a

$$\liminf_{n \to +\infty} (e_{n+1} - e_n) \leqslant 26 < +\infty.$$

Pour mettre ce résultat en contexte, on peut remarquer que la suite (e_n) est plus dense que celle des nombres premiers : on a

$$e_n \sim \frac{n \ln n}{\ln \ln n} \text{ quand } n \to +\infty,$$

ce qui suggère que l'espacement entre les e_n est généralement moindre qu'entre nombres premiers.[25] Pour ce théorème, l'analogue du théorème de Bombieri-Vinogradov pour la répartition des e_n dans les progressions arithmétiques est suffisant.

Le résultat suivant est annoncé dans [13] et [11] :

Théorème 7.2. — *On a*

$$\liminf_{n \to +\infty} \frac{\gamma(n)}{\sqrt{\ln n}(\ln \ln n)^2} < +\infty.$$

Au vu de la preuve du Théorème 1.1, il est naturel d'essayer d'obtenir un tel raffinement, car tous les paramètres disponibles n'ont pas été employées optimalement ; autrement dit, il s'agit de préciser toutes les estimations pour les rendre *uniformes par rapport à k*. Il s'agit d'une tâche extrêmement délicate, et nous n'en dirons pas plus.

Finalement, il est naturel de s'enquérir des espacements $\gamma_r(n) = p_{n+r} - p_n$ pour tout $r \geqslant 1$. La méthode, avec des adaptations, fournit des résultats, mais pas l'analogue du Théorème 1.1. Goldston, Pintz et Yıldırım [11] démontrent :

Théorème 7.3. — *Soit $\theta \geqslant 0$ tel que la suite $(\Lambda(n))$ ait un exposant de répartition $\geqslant \theta$; alors on a*

$$\liminf_{n \to +\infty} \frac{\gamma_r(n)}{\ln n} \leqslant (\sqrt{r} - \sqrt{2\theta})^2,$$

pour $r \geqslant 1$, et en particulier

$$\liminf_{n \to +\infty} \frac{\gamma_r(n)}{\ln n} \leqslant (\sqrt{r} - 1)^2.$$

[24] À ne pas confondre avec les nombres « P_2 » du crible classique, qui *incluent* les nombres premiers et leurs carrés.

[25] Toujours afin de préciser le contexte, rappelons que la plupart des nombres P_2 obtenus par les méthodes de crible vérifient une propriété additionnelle plus forte : ils n'ont « pas de petits facteurs premiers », ce qui ramène leur densité moyenne à celle des nombres premiers.

Pour $r \geqslant 11$, ce n'est cependant pas le meilleur résultat connu, qui est

$$\liminf_{n\to+\infty} \frac{\gamma_r(n)}{\ln n} \leqslant e^{-\gamma}\Big(r - \frac{\sqrt{r}}{2}\Big)$$

où γ est la constante d'Euler. Ce résultat de J. Sivak [23] est basé sur la combinaison de la méthode de Goldston-Yıldırım et de celle de Maier qui fournit des intervalles (clairsemés) d'entiers contenant plus de nombres premiers qu'attendu. Goldston, Pintz et Yıldırım [11] ont annoncé pouvoir incorporer les idées de Maier à leurs arguments, ce qui devrait mener à

$$\liminf_{n\to+\infty} \frac{\gamma_r(n)}{\ln n} \leqslant e^{-\gamma}(\sqrt{r}-1)^2.$$

8. QUEL EST L'EXPOSANT DE RÉPARTITION DES NOMBRES PREMIERS ?

Dès la preuve du théorème de Bombieri-Vinogradov et de ses généralisations à d'autres suites, l'importance potentielle d'une amélioration de l'exposant de répartition d'une suite au-delà de la limite du grand crible avait été reconnue. Depuis environ 1980 (le premier exemple est l'article [8] de Fouvry et Iwaniec, puis la thèse d'É. Fouvry est une étape importante), la recherche de telles améliorations pour diverses suites naturelles a fait l'objet d'un effort concerté de grande ampleur de la part essentiellement de Fouvry, Iwaniec, Friedlander et Bombieri.

Au vu du Théorème 1.2, la question se pose de nouveau de manière particulièrement brûlante, et nous allons faire ici rapidement un survol de ce qui est connu concernant ce problème.

La première remarque essentielle est de bien distinguer entre exposant répartition *au sens faible* et *au sens fort* (ce n'est pas toujours le cas dans la littérature). En effet, rappelons que pour un problème de crible classique, c'est l'exposant de répartition au sens faible qui intervient naturellement, alors que pour la méthode de Goldston, Pintz et Yıldırım, il ne peut suffir.

Dans le cas important des nombres premiers, bien que l'exposant au sens fort soit celui le plus discuté dans la littérature, la notion de P-exposant de répartition mérite peut-être l'attention car, en limitant le problème d'uniformité à un ensemble de classes modulo q assez spécial – et, en particulier, de taille raisonnable puisque de l'ordre de $(\ln q)^{k-1}$ en moyenne, – il ne parait pas impossible qu'il soit plus accessible.

Voici déjà quelques bonnes raisons de croire que l'exposant de répartition des nombres premiers puisse être $> 1/2$:

– D'une part, l'exposant 1/2 découle, comme on l'a dit, de l'Hypothèse de Riemann Généralisée pour les fonctions L de Dirichlet *de manière triviale* : on part d'une « formule explicite » (voir par exemple [18, 5.66]) telle que

$$\psi(x;q,a) = \frac{x}{\varphi(q)} + \frac{\sqrt{x}}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\ q)} \bar{\chi}(a) \sum_{|\gamma_\chi|\leqslant T} \frac{x^{i\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} + O\Big(\frac{x(\ln qx)^2}{T}\Big),$$

pour $1 \leqslant T \leqslant x \leqslant X$, $q \geqslant 1$, la constante implicite étant absolue ($\frac{1}{2}+i\gamma_\chi$ parcourt les zéros non triviaux de $L(\chi,s)$), et on estime « trivialement » la somme sur les zéros avec $T = x$. Il ne semble pas déraisonnable de penser que la somme sur χ donne lieu à des compensations significatives permettant d'obtenir un exposant de répartition des nombres premiers $> 1/2$.[26]

– À l'heure actuelle, le premier point relève du « wishful thinking ». Il semble beaucoup plus convaincant à l'auteur qu'il soit *connu* (essentiellement) que l'exposant de répartition *au sens faible* de la suite des nombres premiers est $> 1/2$. Précisément, on a le théorème suivant :

THÉORÈME 8.1. — *Il existe $\theta_0 > 1/2$ tel que pour tout $X \geqslant 2$, tout $x \leqslant X$, tout $A > 0$ et tout $B > 0$ on ait*

$$\sum_{\substack{d \leqslant D \\ (a,d)=1}} \gamma_d \Big\{ \psi(x; d, a) - \frac{x}{\varphi(d)} \Big\} \ll \frac{X}{(\ln X)^A}$$

uniformément pour tout $D \leqslant X^{\theta_0 - \varepsilon}$ avec $\varepsilon > 0$, tout a tel que $1 \leqslant |a| \leqslant (\ln X)^B$, et toute fonction arithmétique bien factorisable *γ_d de niveau D et d'ordre fixé. La constante implicite dépend au plus de A, B et ε.*

Avec $\theta_0 = 9/17$, cela est dû à Fouvry et Iwaniec, et le meilleur résultat connu, $\theta_0 = 4/7$, est dû à Bombieri, Friedlander et Iwaniec [3]. Dans les deux cas, un ingrédient crucial est la théorie spectrale des formes automorphes et ses applications à l'estimation des sommes de sommes de Kloosterman, développées par Deshouillers et Iwaniec.[27]

La notion de fonction bien factorisable[28] est due à Iwaniec ; bien que l'énoncé ne permette pas *a priori* de prendre pour γ_d le signe de $\psi(x; d, a) - \frac{x}{\varphi(d)}$, Iwaniec a démontré que ces fonctions permettent d'exprimer le terme d'erreur du crible « linéaire ». Il en résulte que cet énoncé est – du point de vue des applications usuelles – équivalent à dire que l'exposant de répartition au sens faible (pour a fixé) est $\geqslant \theta_0 > 1/2$.

Voici une liste (sans doute incomplète) de fonctions dont l'exposant de répartition θ au sens faible est $> 1/2$. Un survol du crible qui met en valeur cet aspect et ses applications a été écrit par Fouvry [7], auquel nous renvoyons pour les références.

– Pour tout $\varepsilon > 0$ fixé, la fonction caractéristique des $n \geqslant 1$ tels que $p \mid n$ implique $p > n^{1/6-\varepsilon}$ (Fouvry).

– La multiplicité de représentation de n sous la forme $n = a^2 + b^2$, a appartenant à n'importe quelle suite d'entiers « assez dense » (Fouvry et Iwaniec) ou $n = a^2 + b^4$ (Friedlander et Iwaniec) ; ici l'exposant de répartition au sens faible est optimal.

– La multiplicité de représentation de n sous la forme $n = a^3 + 2b^3$ (Heath-Brown).

Concernant l'exposant de répartition au sens fort, on a :

– La fonction diviseur $\tau(n)$, avec $\theta \geqslant 2/3$ (Selberg, Linnik, Hooley).[29]

– La fonction $\tau_3(n)$, nombre d'écritures $n = abc$ avec a, b, $c \geqslant 1$ (Friedlander–Iwaniec).[30]

---

[26] Dans le monde des fonctions L sur les corps finis, des résultats (relativement) similaires sont connus.

[27] L'uniformité pour $|a| \leqslant (\ln X)^B$ n'est pas énoncée explicitement. Il est aussi possible de remplacer la classe fixée a par la classe racine d'une équation linéaire donnée $ma + n = 0 \pmod d$, où $m \geqslant 1$ et n est un entier tel que $(m, n) = 1$; cela peut se voir comme un premier pas vers un P-exposant de répartition. Je dois ces remarques à É. Fouvry.

[28] Une fonction arithmétique f est bien factorisable de niveau D entier et d'ordre k si $f(d) = 0$ pour $d > D$, $0 \leqslant f(d) \leqslant \tau(d)^k$, et si pour toute factorisation $D = D_1 D_2$, il existe f_1, f_2 d'ordre k et de niveau D_1 et D_2 respectivement telles que $f = f_1 \star f_2$.

[29] Les termes de reste sont même majorés individuellement ici.

[30] Même remarque.

– Finalement, le seul (?) exemple connu de « fonction caractéristique » dont l'exposant de répartition *au sens fort* soit $> 1/2$ est la suite (a_n), fonction caractéristique des entiers dont la somme des chiffres en base 2 est pair, et ses variantes évidentes en d'autres bases (Fouvry–Mauduit).

9. QUESTIONS

Comme tout progrés significatif, les idées de Goldston, Pintz et Yıldırım soulèvent un certain nombre de nouvelles questions ; nous terminons en indiquant les plus évidentes.

– Peut-on prouver le Théorème 1.2 inconditionnellement ? Ce serait un progrès spectaculaire, et cela semble moins inimaginable qu'il y a quelques mois. Une raison d'espérer est que l'essentiel de la méthode décrite jusqu'à présent se situe au niveau des « termes principaux », et ne procède à aucun travail nouveau au niveau des termes de reste dans (4). Cela semble laisser un grand potentiel.

– Conditionnellement, quelle est la limite de la méthode ? Peut-on démontrer la conjecture des nombres premiers jumeaux (sous la forme (1)) en supposant valide la conjecture d'Elliott-Halberstam ?

– La méthode peut-elle s'étendre aux écarts entre nombres premiers vérifiant des propriétés supplémentaires ? Comme cela n'a de sens que si l'on sait qu'il existe une infinité de tels nombres premiers, le cas le plus naturel est celui des p vérifiant $p \equiv a \pmod q$ (avec q et a fixés) ; cela peut se traiter (c'est annoncé dans [12]) en considérant $G_{\boldsymbol{h}}(n) = (qn + a + h_1) \cdots (qn + a + h_k)$ au lieu de $F_{\boldsymbol{h}}(n)$. En particulier, il ne semble pas évident *a priori* de considérer les écarts entre nombres premiers vérifiant une condition galoisienne (par exemple, p totalement scindé dans une extension galoisienne $K/\mathbf{Q}$ fixée), ce qui serait pourtant tout à fait intéressant.

APPENDICE A : SOMMES DE FONCTIONS MULTIPLICATIVES

Voici un des théorèmes standard permettant d'évaluer asymptotiquement une somme de fonction multiplicative.

Théorème A.1. — *Soit g une fonction multiplicative positive, $\kappa \geqslant 0$ un entier tel que*

$$\sum_{p<x} g(p) \ln p = \kappa \ln x + O(1).$$

Alors pour tout entier $\ell \geqslant 0$ et tout $x \geqslant 2$ on a

$$\sum_{d<x}^{\flat} g(d) \Big(\ln \frac{x}{d} \Big)^{\ell} = c \frac{\ell!}{(\ell + \kappa)!} (\ln x)^{\kappa+\ell} + O((\ln x)^{\ell+\kappa-1})$$

avec

$$c = \prod_p \Big(1 - \frac{1}{p} \Big)^{\kappa} (1 + g(p)),$$

la constante implicite dépendant seulement de g et ℓ.

Expliquons en quelques mots d'où vient la constante c, et pourquoi elle est naturelle. On peut essayer de démontrer ce résultat en procédant par intégration complexe, en écrivant

$$\sum_{d<x}^{\flat} g(d)\Big(\ln\frac{x}{d}\Big)^{\ell} = \frac{\ell!}{2i\pi}\int_{(1)} D_g^{\flat}(s)x^s\frac{ds}{s^{\ell+1}}$$

où l'intégrale complexe est prise sur la droite verticale $\mathrm{Re}(s)=1$ et

$$D_g^{\flat}(s) = \sum_{n\geqslant 1}^{\flat} g(n)n^{-s} = \prod_p (1+g(p)p^{-s}).$$

L'hypothèse peut s'interpréter en disant que $g(p)$ est égal à κ/p « en moyenne », ce qui suggère de comparer $D_g^{\flat}(s)$ à $\zeta(s+1)^{\kappa}$, et de définir une fonction $E_g(s)$ en posant

$$D_g^{\flat}(s) = \zeta(s+1)^{\kappa}E_g(s).$$

C'est effectivement possible pour $\mathrm{Re}(s)>0$; si $E_g(s)$ se prolonge analytiquement légèrement à gauche de $\mathrm{Re}(s)=0$ avec une croissance modérée, on peut déplacer le contour d'intégration vers la droite verticale $\mathrm{Re}(s)=\delta<0$, faisant apparaître un pôle d'ordre $\ell+\kappa+1$ en $s=0$. Le résidu se calcule alors facilement :

$$\begin{aligned}\operatorname*{res}_{s=0}\Big(\zeta(s+1)^{\kappa}E_g(s)\frac{x^s}{s^{\ell+1}}\Big) &= \frac{1}{(\kappa+\ell)!}E_g(0)(\ln x)^{\kappa+\ell} + O((\ln x)^{\kappa+\ell-1})\\ &= \frac{c}{(\kappa+\ell)!}(\ln x)^{\kappa+\ell} + O((\ln x)^{\kappa+\ell-1}),\end{aligned}$$

ce qui explique la forme du terme principal ci-dessus.

En pratique, le prolongement de $E_g(s)$ n'est pas si évident, et il est en fait possible de démontrer l'estimation annoncée élémentairement, suivant des idées de Wirsing. Voir [18, Th. 1.1] ou [15, Lemma 5.4] pour les détails dans le cas $\ell=0$ et [13, p. 11–12]) pour le cas général.

Terminons enfin avec une version de ce théorème uniforme par rapport à g ; elle est nécessaire pour l'uniformité des estimations par rapport à $\boldsymbol{h}$ dans la Section 4.

Théorème A.2. — *Soit g une fonction multiplicative, $\kappa\geqslant 0$ un entier. On note*

$$\rho(p) = \frac{g(p)}{1+g(p)}$$

et on suppose qu'il existe $A_1>0$, $A_2\geqslant 1$, $L\geqslant 1$ tels que

$$0\leqslant g(p)\leqslant A_1 \tag{30}$$

$$-L\leqslant \sum_{w<p\leqslant z}\rho(p)\ln p - \kappa\ln\frac{z}{w}\leqslant A_2. \tag{31}$$

Alors on a

$$\sum_{d<x}^{\flat} g(d)\Big(\ln\frac{x}{d}\Big)^{\ell} = c\frac{\ell!}{(\ell+\kappa)!}(\ln x)^{\kappa+\ell}\Big\{1+O\Big(\frac{L}{\ln x}\Big)\Big\}$$

pour tout entier $\ell\geqslant 0$ et tout $x\geqslant 2$, avec

$$c = \prod_p\Big(1-\frac{1}{p}\Big)^{\kappa}(1+g(p)),$$

la constante implicite dépendant seulement de κ, ℓ, A_1, A_2 et L.

C'est en fait ce qui est démontré (à des changements de notation près) dans [15, Lemma 5.4] et [13, Lemma 5].

RÉFÉRENCES

[1] E. BOMBIERI – *On twin almost primes*, Acta. Arith. 28 (1975–76), 177–193 ; Correction, id., 457–461.

[2] E. BOMBIERI et H. DAVENPORT – *Small differences between prime numbers*, Proc. Roy. Soc. A 293 (1966), 1–18 ; aussi dans The Collected Works of H. Davenport, vol. IV, 1639–1656.

[3] E. BOMBIERI, J. FRIEDLANDER et H. IWANIEC – *Primes in arithmetic progressions to large moduli*, Acta Math. 156 (1986), 203–251.

[4] N. BOURBAKI – *Fonctions d'une variable réelle*, Paris, Hermann, 1976.

[5] J.M. DESHOUILLERS – *Progrès récents des petits cribles* (*d'après Chen, Iwaniec*), Séminaire Bourbaki, Exp. 520 (1980).

[6] É. FOUVRY – *Autour du théorème de Bombieri-Vinogradov*, Acta Math. 152 (1984), 219–244.

[7] É. FOUVRY – *Cinquante ans de théorie analytique des nombres*, dans Development of mathematics 1950–2000, 485–514, Birkhäuser, Basel, 2000.

[8] É. FOUVRY et H. IWANIEC – *On a theorem of Bombieri-Vinogradov type*, Mathematika 27 (1980), 135–152.

[9] P.X. GALLAGHER – *On the distribution of primes in short intervals*, Mathematika 23 (1976), 4–9.

[10] D.A. GOLDSTON – *On Bombieri and Davenport's theorem conerning small gaps between primes*, Mathematika 39 (1992), 10–17.

[11] D.A. GOLDSTON, J. PINTZ et C.Y. YILDIRIM – *Primes in tuples, I.* Prépublication `arXiv:math.NT/0508185`

[12] D.A. GOLDSTON, J. PINTZ et C.Y. YILDIRIM – *The Path to Recent Progress on Small Gaps Between Primes.* Prépublication `arXiv:math.NT/0512436`

[13] D.A. GOLDSTON, S.W. GRAHAM, J. PINTZ et C.Y. YILDIRIM – *Small gaps between primes or almost primes.* Prépublication `arXiv:math.NT/0506067`

[14] D.A. GOLDSTON, Y. MOTOHASHI, J. PINTZ et C.Y. YILDIRIM – *Small Gaps between Primes Exist.* Prépublication `arXiv:math.NT/0505300`

[15] H. HALBERSTAM et H. RICHERT – *Sieve methods*, Academic Press (1974).

[16] D. R. HEATH-BROWN – *Almost prime k-tuples*, Mathematika 44 (1997), 245–266.

[17] B. HOST – *Progressions arithmétiques dans les nombres premiers*, Séminaire Bourbaki, Exp. 944 (2005).

[18] H. IWANIEC et E. KOWALSKI – *Analytic Number Theory*, A.M.S Colloquium Publications vol. 53 (2004).

[19] E. KOWALSKI – *Un cours de théorie analytique des nombres*, S.M.F Cours Spécialisé 13, (2004).

[20] P. MICHEL – *Progrés recents du crible et applications* (*d'après Duke, Fouvry, Friedlander, Iwaniec*), Séminaire Bourbaki, Exp. 842 (1998).

[21] J. PINTZ et Y. MOTOHASHI – *A smoothed GPY sieve.* Prépublication `arXiv:math.NT/0602599`

[22] A. SELBERG – *Lectures on sieves*, dans Collected Papers, vol. II, Springer Verlag 1992 (pp. 65–247).

[23] J. SIVAK – *Méthodes de crible appliquées aux sommes de Kloosterman et aux petits écarts entre nombres premiers*, Thèse de doctorat (Université Paris Sud, décembre 2005).

Emmanuel KOWALSKI

Université de Bordeaux I
Laboratoire d'algorithmique arithmétique
et expérimentale (A2X)
UMR 5465 du CNRS
351, Cours de la Libération
F–33405 TALENCE
E-mail : `kowalski@math.u-bordeaux1.fr`