N° d'ordre : 3277

THÈSE

présentée à

L'UNIVERSITÉ BORDEAUX I

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET INFORMATIQUE

par **Arnaud CHADOZEAU**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : **Mathématiques pures**

SUR LA RÉPARTITION DES ENTIERS PREMIERS À UN ENTIER DONNÉ

Soutenue le 7 décembre 2006 à l'Institut de Mathématiques de Bordeaux

Après avis de :

R. de la BRETÈCHE, Professeur	Université Paris VII	**Rapporteur**
H. L. MONTGOMERY, Professeur	Université du Michigan	**Rapporteur**

Devant la commission d'examen composée de :

M. BALAZARD, Chargé de Recherches	Université Bordeaux I	**Directeur**
K. BELABAS, Professeur	Université Bordeaux I	
R. de la BRETÈCHE, Professeur	Université Paris VII	**Rapporteur**
E. KOWALSKI, Professeur	Université Bordeaux I	
H. L. MONTGOMERY, Professeur	Université du Michigan	**Rapporteur**
E. ROYER, Professeur	Université de Clermont-Ferrand II	**Président**
É. SAIAS, Maître de Conférences	Université Paris VI	

– 2006 –

Arnaud Chadozeau

SUR LA RÉPARTITION DES ENTIERS PREMIERS À UN ENTIER DONNÉ

Arnaud Chadozeau
Université Bordeaux I, A2X, 351 cours de la Libération, 33 405 Talence cedex.
E-mail : Arnaud.Chadozeau@math.u-bordeaux1.fr

En premier lieu, j'aimerais adresser mes profonds remerciements à mon directeur de thèse, Michel Balazard, pour son regard expert, son écoute attentive, ses conseils avisés et pour son amitié.
Je remercie également les rapporteurs, Régis de la Bretèche et Hugh L. Montgomery, de l'attention soutenue qu'ils ont portée à mon travail, ainsi qu'aux autres membres du jury, Karim Belabas, Emmanuel Kowalski, Emmanuel Royer et Éric Saias, d'avoir accepté de participer à ma soutenance.
Je suis particulièrement reconnaissant à Régis de la Bretèche et à Éric Saias, qui ont orienté mes choix mathématiques initiaux et qui ont influencé le cours de cette thèse. D'autres personnes n'ont pas non plus compté leur temps pour discuter de ce travail et ont contribué à sa forme actuelle, j'aimerais leur signifier ma gratitude : Laurent Habsieger, Kevin Ford, Pierre Mazet, Ramachandran Balasubramanian, Gérald Tenenbaum.

Je tiens également à saluer tous les membres et le personnel de l'Institut de Mathématiques de Bordeaux, qui font chaque jour de cet endroit un cadre agréable de travail et un lieu d'échanges fertiles.

J'aimerais enfin adresser ces derniers remerciements à mes proches pour leur soutien et leur compréhension, et tout spécialement à Isabel pour sa foi indéfectible.

7 décembre 2006

TABLE DES MATIÈRES

Notations et avertissements

Dans ce travail, par commodité on utilisera la notation classique de Vinogradov $\ll$ (resp. $\ll_{\varepsilon,\delta}$ si la constante implicite dépend des variables ε et δ) étendue de la façon suivante $\ll^k$ (resp. $\ll^k_{\varepsilon,\delta}$) pour indiquer que la dépendance en k de la constante implicite peut être choisie sous la forme $A\cdot C^k$, où la constante C sera appelée la magnitude de l'estimation. Par exemple, on a $\binom{k}{m} \ll^k 1$, où l'on peut choisir 2 comme magnitude, et $k \ll^k 1$ où la constante de magnitude peut être choisie aussi proche de 1 que souhaitée. Il est à noter que dans $\ll^k_\varepsilon$, les rôles de ε et de k ne sont pas symétriques. En effet, le symbole $\ll_\varepsilon$ traduit un énoncé du genre « $\forall\varepsilon, \exists C, \ldots \leqslant C\cdot\ldots$ », alors que $\ll^k$ sous-entend plutôt « $\exists C, \forall k, \ldots \leqslant C^k\cdot\ldots$ ». Il faut donc voir la relation $\ll^k_\varepsilon$ comme $\big(\ll_\varepsilon\big)^k$, soit un énoncé du genre « $\forall\varepsilon, \exists C, \forall k, \ldots \leqslant C^k\cdot\ldots$ ». Cette notation peu standard est contestable, elle ne permet d'économiser qu'une seule lettre ; aussi nous efforcerons-nous de ne pas l'utiliser dans les énoncés de nos principaux résultats. Cependant, à l'instar de la notation de Vinogradov, elle permet tout de même d'économiser une lettre, dont la signification peut varier au cours du calcul : cette abréviation se montre particulièrement pratique pour une suite de majorations (puisque la relation $\ll^k$ est transitive) et nous l'emploierons sans retenue au cours de nos démonstrations.

On utilisera également la notation plus classique de la double factorielle, définie sur les entiers naturels positifs par la relation de récurrence $(2n+2)!! = (2n+1)\cdot(2n)!!$ et par les valeurs initiales $0!! = 1$ et $1!! = 0$. La double factorielle possède une interprétation combinatoire : la quantité $n!!$ compte le nombre d'appariements parmi n objets, c'est-à-dire le nombre d'involutions sans point fixe de $\mathfrak{S}_n$. Cette fonction possède également une interprétation probabiliste puisqu'il s'agit des moments de la loi normale centrée réduite

$$n!! \coloneqq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t^n e^{-t^2/2}\,\mathrm{d}t = \begin{cases} \frac{n!}{2^{n/2}(n/2)!} & \text{si } n \text{ est pair ;}\\ 0 & \text{si } n \text{ est impair.}\end{cases}$$

Il sera donc naturel de retrouver ce facteur dans l'estimation de moments centrés. Cependant, il est utile de considérer également les moments absolus de la loi normale qui valent

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} |t|^n e^{-t^2/2}\,\mathrm{d}t = \Gamma\Big(\frac{n+1}{2}\Big)\frac{2^{n/2}}{\sqrt{\pi}} = \frac{\Gamma(n+1)}{2^{n/2}\Gamma(n/2+1)}.$$

On définit donc la fonction $\mathbb{\Gamma}(s) \coloneqq \Gamma\big(\frac{s}{2}\big)\frac{2^{s/2}}{\sqrt{2\pi}}$ méromorphe sur $\mathbb{C}$ avec un pôle simple en chaque entier pair négatif et vérifiant $\mathbb{\Gamma}(n+1) = n!!$ pour tout entier pair positif. Par la formule de Stirling, on a $\mathbb{\Gamma}(t+1) \sim \sqrt{2}\,(t/e)^{t/2}$ pour t réel tendant vers $+\infty$.

On utilise également la notation $[\![a,b]\!]$ pour $\mathbb{Z}\cap[a,b]$ et on parlera d'intervalle entier de longueur $b-a+1$. Les notations plus classiques $P^-(q)$ et $P^+(q)$ pour le plus petit et le plus grand diviseur premier d'un entier $q>1$ seront aussi employées, avec la convention $P^-(1)=\infty$ et $P^+(1)=1$.

Enfin, on utilise la notation d'Iverson $[\mathcal{P}]$ qui à une propriété $\mathcal{P}$ associe la valeur 1 si elle est vérifiée, et 0 sinon.

INTRODUCTION

1. La fonction de Jacobsthal

En 1961 et 1962, le mathématicien berlinois Ernst Jacobsthal, alors âgé de 80 ans, publie une série de cinq articles [**23**] sur une fonction qui porte aujourd'hui son nom.

Soit $q > 1$ un entier. Jacobsthal s'est intéressé aux entiers g qui sont tels que tout intervalle entier $[\![a+1, a+g]\!]$ de longueur g (pour $a \in \mathbb{Z}$) contienne au moins un entier premier à q. Il définit alors la fonction $g(q)$, aujourd'hui connue comme la « fonction de Jacobsthal », comme le plus petit de ces entiers g, à savoir

$$g(q) \coloneqq \min\Big\{g; \forall a \in \mathbb{Z}, \min_{1 \leqslant i \leqslant g} (a+i, q) = 1\Big\}.$$

On peut remarquer que $g(q) = g(q')$ si les entiers q et q' ont les mêmes facteurs premiers : l'étude peut donc se restreindre aux entiers sans facteur carré. Plus exactement, Jacobsthal a surtout étudié la fonction

$$C(r) \coloneqq \max_{\omega(q)=r} g(q) - 1,$$

où le maximum porte sur tous les entiers q possédant r facteurs premiers distincts. Il en calcula les premières valeurs ($C(1) = 1$, $C(2) = 3$, $C(3) = 5$, $C(4) = 9$, $C(5) = 13$, $C(6) = 21$, $C(7) = 25$, ...) et les premiers encadrements ($r \ll C(r) \ll r2^r$). Les premières valeurs semblent montrer que $C(r) = 2p_{r-1} - 1$, où p_r désigne le r-ième nombre premier (avec $p_0 = 1$). L'étude que fait Jacobsthal montre également que pour ces petites valeurs de r, le maximum définissant $C(r)$ est atteint pour $q = p_1 p_2 \cdots p_r$, le produit des r plus petits nombres premiers. Ainsi, en notant

$$C_0(r) = g(p_1 \cdots p_r) - 1,$$

il obtint $C_0(r) = C(r)$ pour $r \leqslant 7$. La postérité du travail de Jacobsthal tient surtout aux deux conjectures qu'il a formulées :

Conjecture 1 : On a $C(r) = C_0(r)$ pour tout $r \geqslant 1$.

Conjecture 2 : On a $C(r) \ll r^2$ pour $r \geqslant 1$.

C'est cette deuxième conjecture, énoncée dans une lettre adressée à Erdős et reprise dans [**10**], qui connut un réel développement.

2. Historique

Dans la première édition de l'*Essai sur la théorie des nombres*, en 1798*, Legendre démontre qu'il existe une infinité de nombres premiers dans toute progression arithmétique dont la raison est première à l'un des termes. Cette démonstration se présente en trois étapes, que nous pouvons réécrire ainsi :

a). On a $C_0(r) \leqslant 2p_{r-1} - 1$ pour tout $r \geqslant 1$;
b). On a $C(r) = C_0(r)$ pour tout $r \geqslant 1$;
c). La majoration $g(q) \leqslant 2p_{\omega(q)-1} - 1$ implique la propriété annoncée.

Il est amusant de constater que Legendre avait conscience de la profondeur de cet énoncé et s'étonnait naïvement de la simplicité de sa preuve.

> « Toute progression arithmétique dont le premier terme et la raison sont « premiers entre eux, contient une infinité de nombres premiers. »
>
> Cette proposition, qui est très-utile dans la théorie des nombres, avait été indiquée dans les Mémoires de l'Académie des Sciences, an. 1785 ; mais jusqu'à présent sa démonstration n'était point encore connue et paraissait offrir de grandes difficultés.

Sa preuve était si simple que l'Académie des Sciences vint à en douter et mit au concours sa vérification ou son infirmation. Dupré† parvint à donner une série de contre-exemples au point a) pour $15 \leqslant r \leqslant 31$: il toucha la récompense de 3 000 Francs. La véracité de l'énoncé ne fut par contre jamais mise en cause. Celui-ci fut d'ailleurs établi rigoureusement et quantitativement en 1839 par Dirichlet, qui reconnut avoir essayé de formaliser l'argument de Legendre avant de l'abandonner pour une méthode inspirée de la preuve d'Euler de l'infinitude des entiers premiers. L'assertion b), qui correspond à la conjecture 1, a simplement été admise par Legendre. Nous reparlerons du point c) plus loin.

Après de nombreuses infirmations quantitatives de la majoration de Legendre, Rankin [**33**, th. II] dans son travail sur les grandes différences entre nombres premiers consécutifs, énonce une minoration de $C_0(r)$ dont on tire immédiatement, comme le remarque Erdős [**10**], la minoration

$$C(r) \geqslant C_0(r) \gg r(\log r)^2 \frac{\log_3 r}{(\log_2 r)^2} \gg_\varepsilon p_r(\log r)^{1-\varepsilon}. \tag{0.1}$$

Cette estimation, à la valeur de la constante implicite près, n'a pas été améliorée depuis 1938.‡

S'il n'a pas réussi à fournir de bornes précises pour $C(r)$, le travail de Jacobsthal a permis de remettre l'ouvrage sur le métier. La poursuite de cette étude et sa popularisation fut alors effectuée par Erdős : dans [**10**], il remarque que l'on peut obtenir

*Voir également le théorème 405 de *Essai sur la théorie des nombres, 2^e éd.,* p. 404, Paris, Courcier, 1808 ; ou le théorème 410 de *Théorie des nombres, 3^e éd.,* vol. II, p. 76, Paris, Firmin Didot, 1830.
†A. Dupré, *Examen d'une proposition de Legendre relative à la théorie des nombres,* suivi d'un *Mémoire sur la résolution des équations numériques,* Paris, Mallet-Bachelier, 1859.
‡La meilleure constante implicite connue est $2e^\gamma + o(1)$. [J. Pintz, Very large gaps between consecutive primes, *J. Num. Theory,* **63** (1997), 286–301]

par une application directe du crible de Brun la majoration

$$C(r) \ll r^C, \tag{0.2}$$

pour une certaine constante C. Erdős y donne surtout la minoration

$$g(q) \geqslant \big(1+o(1)\big)\omega(q)\frac{q}{\varphi(q)} \tag{0.3}$$

et l'ordre normal

$$g(q) = \big(1+o(1)\big)\omega(q)\frac{q}{\varphi(q)} \qquad \text{pour presque tout } q. \tag{0.4}$$

La majoration $g(q) \ll \omega(q)\frac{q}{\varphi(q)}$ n'est cependant pas valable pour $q \geqslant 1$, puisque celle-ci induirait $C(r) \ll r\log r$ et contredirait la minoration de l'encadrement (0.1). On peut voir ici un argument en faveur de la conjecture 1 de Jacobsthal : pour les entiers normaux, on a $g(q) \ll \omega(q)\log\omega(q)$ alors que cette majoration est fausse pour les produits de petits nombres premiers. On voit également que la majoration $g(q) < 2p_{\omega(q)-1}$ « démontrée » par Legendre n'est pas totalement infondée.

Le problème de la majoration de $g(q)$ et de $C(r)$ se traduit en termes de minorations par des méthodes de crible ; c'est la philosophie de la majoration (0.2). Dans un célèbre article [**21**], Iwaniec formule les inégalités optimales du crible linéaire et déduit ainsi la majoration

$$C_0(r) \ll r^2(\log r)^2, \tag{0.5}$$

et remarque aussi que l'on peut choisir $C = 2+\varepsilon$ dans (0.2). Il affinera ce résultat dans [**22**] en montrant

$$g(q) \ll \frac{q}{\varphi(q)}\omega(q)^2\log 2\omega(q) \tag{0.6}$$

et donc fournit une conclusion quasi-complète à la seconde conjecture de Jacobsthal, puisque la majoration (0.6) induit le résultat suivant

$$C(r) \ll r^2(\log r)^2. \tag{0.7}$$

Vaughan, qui avait établi peu de temps avant Iwaniec dans [**38**] une version moins précise de (0.6), à savoir $g(q) \ll \omega(q)^2\big(\log 2\omega(q)\big)^4$, analyse de façon très pertinente les raisons de la présence de la puissance carrée de $\omega(q)$:

> Any further substantial progress towards [$g(q) \ll \omega(q)^{1+\varepsilon}$] probably requires a fundamental new idea. That 2 is the best that can be done by the method below is due to the usual inability of the one dimensional sieve to handle moduli larger than about the square root of the length of the interval under examination. Moreover, that any further improvement must lie very deep is indicated by the work of Kanold [...] which shows that Linnik's celebrated theorem on the least prime in arithmetic progression follows easily from [$g(q) \ll \omega(q)^C$] with $C < 2$.

En effet, dans un de ses articles consacrés à la fonction de Jacobsthal, Kanold [**24**] montre que la majoration $g(q) \ll \omega(q)^C$ avec $C < 2$ induirait l'existence d'une infinité de nombres premiers dans une progression arithmétique admissible et le théorème de Linnik avec une constante de Linnik $L = 2/(2-C)$. Il s'agit en fait

d'une redécouverte et d'une généralisation de l'argument de Legendre : l'idée principale est que pour tout triplet (a, d, m) avec $(a, d) = (d, m) = 1$, la progression arithmétique $\{a + d, a + 2d, \dots, a + g(m)d\}$ contient au moins un nombre premier à dm. On rappelle que Linnik [**26**] a montré qu'il existait une constante absolue L telle qu'on ait uniformément pour toute paire d'entiers $1 \leqslant a \leqslant d$ avec $(a, d) = 1$ la majoration

$$P(d, a) \ll_\varepsilon d^{L+\varepsilon},$$

où $P(d, a)$ est le plus petit entier premier congru à a modulo d. Chowla [**4**] a montré que sous l'hypothèse de Riemann généralisée, on pouvait choisir $L = 2$. On précise que la valeur conjecturale de la constante de Linnik communément admise est $L = 1$ et que la meilleure majoration connue $L = 5,5$ est due à Heath-Brown [**18**].

Une seconde conséquence importante de l'inégalité $g(q) \ll \omega(q)^C$ avec $C < 2$, est la majoration des différences entre nombres premiers consécutifs $p_{n+1} - p_n \ll n^{C/2}$. Cramér [**6, 7**] a démontré $p_{n+1} - p_n \ll {p_n}^{1/2} \log p_n$ en supposant l'hypothèse de Riemann, avant de proposer $p_{n+1} - p_n \ll (\log p_n)^2$ dans [**7**]. Le meilleur résultat dans cette direction est aujourd'hui dû à Baker, Harman et Pintz [**1**] qui ont prouvé $p_{n+1} - p_n \ll {p_n}^{21/40+\varepsilon}$.

Il est courant de conjecturer $g(q) \ll \omega(q)^{1+\varepsilon}$, et même sous une forme optimiste

$$g(q) \ll \frac{q}{\varphi(q)} \log q. \tag{0.8}$$

En plus de l'infinitude des nombres premiers dans les suites arithmétiques, cette conjecture induirait également le théorème de Linnik avec $L = 2$ ainsi que la majoration $p_{n+1} - p_n \ll {p_n}^{1/2} \log p_n$. Il n'est pas raisonnable de conjecturer mieux, puisque la minoration de Rankin (0.1) indique que pour une infinité d'entiers, on a $g(q) \gg \big(\frac{q}{\varphi(q)}\big)^{1-o(1)} \log q$.

3. Un autre point de vue

L'intérêt d'Erdős pour la fonction de Jacobsthal vient probablement du fait qu'elle est directement reliée à l'une de ses vieilles conjectures préférées[§] formulée en 1940 [**9**] : si l'on pose $1 = a_1 < \ldots < a_{\varphi(n)} = q - 1$ la suite exhaustive des entiers premiers à q et inférieurs à q, on a

$$\sum_{i=1}^{\varphi(q)-1} (a_{i+1} - a_i)^\gamma \ll_\gamma \varphi(q) \Big(\frac{q}{\varphi(q)}\Big)^\gamma, \tag{0.9}$$

pour $\gamma = 2$. Le lien avec la fonction de Jacobsthal vient de l'identité $g(q) = \max_i(a_{i+1} - a_i)$, identité simple mais qui permet d'envisager différemment le problème de la majoration de $g(q)$. Hooley [**19**] apporta du crédit à cette conjecture en prouvant la relation (0.9) pour tout réel $0 \leqslant \gamma < 2$. Erdős [**11**] proposa alors une série de conjectures menant implicitement à (0.8) : d'abord, la relation (0.9), en plus d'être

[§] « one of my favorite old conjectures », [**11**, p. 3].

vérifiée pour $\gamma = 2$, est vraie pour tout réel $\gamma \geqslant 0$; ensuite, il existe un réel $x > 0$ tel que

$$\sum_{i=1}^{\varphi(q)-1} e^{x(a_{i+1}-a_i)\frac{\varphi(q)}{q}} \ll \varphi(q). \tag{0.10}$$

Il est évident que la majoration (0.10) mène à (0.8), puisque $g(q) = \max_i(a_{i+1} - a_i)$.

La réalisation de la première partie de ce programme fut accomplie par Montgomery et Vaughan [**30**]. Comme chez Erdős [**10**] et Hooley [**19**], l'étude passe par la majoration des moments centrés

$$M_k(h;q) \coloneqq \sum_{n=1}^{q}\Big(\sum_{i=1}^{h}\big[(n+i,q)=1\big] - h\frac{\varphi(q)}{q}\Big)^k,$$

où $\big[(n,q)=1\big]$ vaut 1 si n et q sont premiers entre eux et 0 sinon. Le résultat principal menant à (0.9) pour tout γ s'énonce alors ainsi

$$M_k(h;q) \ll_k q\Big(h\frac{\varphi(q)}{q}\Big)^{k/2}, \tag{0.11}$$

pour tout $k \in \mathbb{N}$ et pour $h\frac{\varphi(q)}{q} \geqslant 1$. Pour obtenir un tel résultat, Montgomery et Vaughan montrent premièrement que l'étude peut se réduire de façon indépendante au cas où q est sans grand facteur premier et au cas où q est sans petit facteur premier. Si q ne possède que des grands facteurs premiers (grands par rapport à une certaine fonction de h et de k), la suite des a_i est statistiquement proche d'une suite prise au hasard moyenne, dont le moment est facilement calculable. Si q ne possède que des petits facteurs premiers, l'estimation souhaitée provient d'une analyse harmonique et d'un lemme fondamental miraculeux, explicitement décrit dans [**31**].

Nous nous proposons de continuer — sans parvenir à l'achever — le programme d'Erdős en vue d'obtenir (0.10), et donc (0.8). Pour cela nous suivons les pas de Montgomery et de Vaughan, en faisant constamment attention à cette dépendance en k, améliorant ainsi nombre de leurs résultats.

Dans un premier temps, nous réalisons une étude détaillée des liens entre des énoncés du type (0.8) ou (0.10) et des versions uniformes de relations du type (0.9) ou (0.11), et ce pour toute suite croissante d'entiers $(a_i)_{i\in\mathbb{Z}}$ généralisant la suite des entiers premiers à un entier donné. En particulier, on montre qu'un énoncé comme celui de la conjecture que nous formulons permet de prouver les relations conjecturales (0.10) et (0.8).

Conjecture I. — *On a la majoration*

$$M_k(h;q) \ll q(ck)^{k/2}\Big(k + h\frac{\varphi(q)}{q}\Big)^{k/2}$$

uniformément en les entiers naturels k, h et q et pour une constante absolue $c > 0$.

Dans ce premier chapitre, nous présentons également un outil général pour décomposer sous certaines conditions un moment associé à une suite complexe en une somme de moments de suites plus simples.

L'analogue probabiliste du moment $\frac{1}{q}M_k(h;q)$, qui permet d'étudier d'étudier la répartition d'une suite générale $(a_i)_{i\in\mathbb{Z}}$ en moyenne, est le moment centré $\mu_k(h,P)$ d'ordre k d'une loi binomiale de paramètres (h,P), où P est la densité des valeurs de la suite $(a_i)_{i\in\mathbb{Z}}$. Nous établissons l'équivalent probabiliste de la conjecture I.

Théorème II. — *On a la majoration*

$$\mu_k(h,P) \ll (ck)^{k/2}\big(k+hP(1-P)\big)^{k/2}$$

uniformément en les entiers naturels k et h et en la probabilité $P\in[0,1]$ pour une constante absolue $c>0$.

Nous discutons également le caractère optimal de la majoration du théorème II ; en particulier, le changement de comportement asymptotique traduit par le terme $k+hP(1-P)$ est une réalité.

Proposition III. — *On a uniformément en l'entier $k\geqslant 2$ et en le réel $\bar{\mu}_2$*

$$\lim_{\substack{(h,P)\to(+\infty,0)\\ hP(1-P)=\bar{\mu}_2}} \mu_k(h,P) \gg \begin{cases} \big(\frac{1}{\log(k/\bar{\mu}_2)}\big)^{2k}(ck)^k & \text{si } \bar{\mu}_2\leqslant k/2,\\ (ck)^{k/2}\bar{\mu}_2^{\lfloor k/2\rfloor} & \text{sinon,}\end{cases}$$

pour une constante absolue $c>0$.

Sous les conditions $hP(1-P)\gg k^3$ et k pair, nous obtenons même un équivalent uniforme en les trois variables qui n'était auparavant établi qu'à k fixé.

Proposition IV. — *Lorsque $hP(1-P)/k^3$ tend vers l'infini, on a*

$$\mu_k(h,P)\sim \frac{k!}{2^{k/2}(k/2)!}\big(hP(1-P)\big)^{k/2},$$

uniformément en k pair, en $h\geqslant 1$ et en $P\in[0,1]$.

L'argument principal est que la quantité à étudier $\mu_k(h,P)$ est un polynôme de degré $\lfloor k/2\rfloor$ en la variable $hP(1-P)$ et dont les coefficients sont des éléments de $\mathbb{Z}[P]$. En étudiant précisément ces coefficients (que nous appelons polynômes de Romanovsky), nous parvenons à démontrer ces estimations sur $\mu_k(h,P)$.

Nous utilisons ensuite ces résultats pour montrer que la relation de la conjecture I est vérifiée sous la condition supplémentaire que tous les diviseurs premiers de q sont supérieurs à h.

Théorème V. — *Pour une constante absolue $c>0$, on a la majoration*

$$M_k(h;q)\ll q(ck)^{k/2}\Big(k+h\frac{\varphi(q)}{q}\Big)^{k/2}$$

uniformément en les entiers naturels k, h et q pour peu que tous les diviseurs premiers de q soient supérieurs à h.

Cette condition est la limite naturelle de cette méthode probabiliste : en effet, les évènements « n est premier à q » et « $n+p$ est premier à q » ne sont pas physiquement indépendants si p est un diviseur premier de q ; si ce facteur premier p est strictement inférieur à h, le nombre d'entiers premiers à q ne peut être modélisé par une variable aléatoire suivant une loi binomiale. Cette comparaison entre $M_k(h;q)$ et son analogue probabiliste a été déjà utilisée par Montgomery et Vaughan qui ont montré à k fixé

$$M_k(h;q) \sim q\mu_k\big(h, \varphi(q)/q\big)$$

sous la condition que les diviseurs de q soient tous supérieurs à $h^{k/2}$. Nous avons pu améliorer ce résultat en utilisant un polynôme généralisant le polynôme $\mu_k(h,P)$ sous forme de convolution, et en étudiant les annulations entre les termes $\prod_{p|q}(1-\frac{1}{p})^{-t}(1-\frac{t}{p})$. Ce dernier point a été traité de façon très générale.

Théorème VI. — *Soient $\varepsilon \in]0,1/2[$ et $K>0$ deux réels. Pour tout vecteur complexe $\mathbf{x}=(x_1,\dots,x_n)\in\mathbb{C}^n$, on pose*

$$N(\mathbf{x}) \coloneqq \max\Big(|x_1|,\dots,|x_n|,\sum_{i=1}^{n}|x_i|^2\Big).$$

Uniformément pour tout vecteur complexe $\mathbf{x}=(x_1,\dots,x_n)\in\mathbb{C}^n$ vérifiant $|1-x_i|\geqslant\varepsilon$ pour tout i et pour tout entier naturel t vérifiant $tN(\mathbf{x})\leqslant K$ on a

$$\sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}\prod_{i=1}^{n}\frac{1-sx_i}{(1-x_i)^s} \ll (c_{K,\varepsilon})^t\big(tN(\mathbf{x})\big)^{\lceil t/2\rceil},$$

où $c_{K,\varepsilon}$ est une constante positive dépendant seulement de ε et de K.

La preuve de ce théorème a nécessité la résolution d'un problème vraisemblablement nouveau d'énumération de partitions, que nous ne détaillerons pas dans cette introduction.

Dans le travail de Montgomery et Vaughan, le second outil essentiel est une inégalité, généralisation multidimensionnelle de l'inégalité de Cauchy-Schwarz, dont deux démonstrations différentes ont été présentées. Nous réinterprétons ces démonstrations avec l'aide de la théorie des graphes, et proposons une nouvelle preuve de cette relation, dont les anciennes sont des cas très spécifiques. Cette nouvelle construction nous permet de penser que dans l'inégalité

$$\Bigg|\sum_{\substack{a_i\in[\![1,d_i]\!]\\(1\leqslant i\leqslant k)\\ \sum_i a_i/d_i\in\mathbb{Z}}}\prod_{i=1}^{k}G_i(a_i/d_i)\Bigg| \leqslant \frac{\sqrt{\prod_{i=1}^{k}d_i}}{d}\prod_{i=1}^{k}\Big(\sum_{a_i=1}^{d_i}\big|G_i(a_i/d_i)\big|^2\Big)^{1/2},$$

où les d_i sont k entiers naturels de p.p.c.m. d, les G_i des fonctions complexes 1-périodiques, le terme $(\prod_{i=1}^{k}d_i)^{1/2}/d$ n'est pas essentiel. Pour créditer cette hypothèse, nous montrons que dans le cas où les fonctions G_i sont définis par $G_i(\rho)=\frac{\sin\pi h\rho}{\sin\pi\rho}$ si ρ est une fraction de la forme a_i/d_i avec $(a_i,d_i)=1$ et par $G_i(\rho)=0$ sinon — cas qui est celui qui apparaît dans l'étude de $M_k(h;q)$ — ce facteur $(\prod_{i=1}^{k}d_i)^{1/2}/d$ n'est pas toujours présent.

Proposition VII. — *On a*

$$\left| \sum_{\substack{a_i\in[\![1,d_i]\!]\\(a_i,d_i)=1\\(1\leqslant i\leqslant k)\\ \sum_i a_i/d_i\in\mathbb{Z}}} \prod_{i=1}^{k} \frac{\sin \pi h a_i/d_i}{\sin \pi a_i/d_i} \right| \leqslant \prod_{i=1}^{k} \sigma(d_i).$$

Cette nouvelle majoration permet d'obtenir de meilleurs résultats pour les « petits » d_i, et nous n'avons pas réussi à trouver une relation équivalente pour les « grands » d_i. De toute façon, même en faisant disparaître le facteur $(\prod_{i=1}^{k} d_i)^{1/2}/d$, cette majoration ne permet pas d'éviter le phénomène d'explosion combinatoire dû au grand nombre de k-uplets de d_i considéré, et permettrait tout au plus de l'atténuer.

Le travail effectué ne permet donc pas d'établir la conjecture I, mais permet de restreindre les cas pour lesquels celle-ci reste à vérifier. Notamment, l'étude peut à présent se limiter aux entiers q sans facteur premier supérieur à h (dans ce cas $q/\varphi(q) \ll \log h$) et aux entiers $k \leqslant h\varphi(q)/q$. Grâce à la proposition VII, et en notant $S_{\mathbf{d}}(h)$ la somme majorée en valeur absolue dans cette proposition VII, on peut se restreindre à établir

$$\sum_{\substack{\sqrt{kh}(\varphi(q)/q)^{3/2}<d_i|q\\(1\leqslant i\leqslant k)}} \prod_{i=1}^{k} \frac{\mu(d_i)}{\varphi(d_i)} S_{\mathbf{d}}(h) \ll \Big(c\, kh \frac{q}{\varphi(q)}\Big)^{k/2}$$

uniformément en k, h et q sous les conditions citées et pour $c > 0$ une constante absolue. Cependant, une idée nouvelle reste à trouver pour gérer les annulations parmi la multitude des termes de la somme.

Indépendamment des conjectures de Jacobsthal, nous avons également étudié un problème lié au graphe divisoriel. On note $\mathcal{D}(x)$ le graphe simple dont les sommets sont les entiers positifs inférieurs à x et dont les arêtes sont les couples d'entiers distincts dont l'un est multiple de l'autre. Si la plus longue chaîne de $\mathcal{D}(x)$ a fait l'objet de nombreuses publications, le problème de partitionner en chaînes le graphe $\mathcal{D}(x)$ est moins populaire. Initialement posée par Erdős en 1993 et développée par Erdős et Saias, cette étude a connu un développement récent grâce à Mazet, qui après avoir reformulé le problème dans un cadre naturel plus large (une sorte de généralisation aux entiers de Beurling) a su montrer que le nombre minimal de chaînes nécessaires à une partition de $\mathcal{D}(x)$ possède un comportement asymptotique en Cx, où C est une constante non calculée. De façon spécifique, sa preuve est basée sur le fait que pour tout réel y, le graphe $\mathcal{D}(x, y)$, graphe divisoriel restreint au entiers inférieurs à x dont les facteurs premiers sont inférieurs à y, est recouvert par une unique chaîne pour peu que x soit assez grand. Nous rendons cet énoncé explicite, c'est-à-dire que nous explicitons une fonction $X(y)$ telle que le graphe $\mathcal{D}(x, y)$ soit hamiltonien dès que $x \geqslant X(y)$, et améliorons la construction de Mazet pour obtenir une meilleure fonction $X(y)$. Cela permet notamment d'établir le résultat suivant.

Proposition VIII. — *Il existe une constante $C \in [0,1]$ telle que le nombre minimal de chaînes nécessaires à une partition de $\mathcal{D}(x)$ vaut*

$$Cx\Big(1+O\Big(\frac{1}{\log_2 x \log_3 x}\Big)\Big).$$

L'ouvrage est conclu par deux appendices, le premier contenant un mémorandum sur les nombres de Stirling, le second un résultat sur les sommes de Ramanujan $c_d(n)$ en moyenne, qui vient compléter une question soulevée au cours du chapitre 4. Nous y montrons notamment l'encadrement suivant en discutant le caractère optimal de la majoration.

Théorème IX. — *Pour d sans facteur carré, on a*

$$d \ll \sup_{x<y}\Big|\sum_{x<n\leqslant y} c_d(n)\Big| \ll d\min\Big(\prod_{\substack{p\mid d\\ p\leqslant \log d}}\Big(1-\frac{1}{p}\Big)^{-1},\ \prod_{\substack{p\nmid d\\ p\leqslant \log d}}\Big(1-\frac{1}{p}\Big)^{-1}\Big).$$

CHAPITRE 1

ÉTUDE GÉNÉRALE DE LA RÉPARTITION D'UN SOUS-ENSEMBLE DE $\mathbb{Z}/q\mathbb{Z}$

Dans ce chapitre, nous essayons de relier de façon généralisée les problèmes de type Jacobsthal — à savoir estimer les différences maximales entre valeurs consécutives d'une suite — et les problèmes de type Erdős — c'est-à-dire estimer l'irrégularité de la répartition d'une suite à l'aide d'outils statistiques, comme les énoncés (0.9) et (0.11). En cela, nous sommes très proches de la philosophie de Hooley dans [**20**]. L'objectif est de comprendre quelle forme uniforme en k de la majoration (0.11) permet d'obtenir des majorations de la fonction de Jacobsthal du type (0.8).

1.1. Cadre général du problème

Soient q un entier naturel et $\mathbf{a}$ une partie de $\mathbb{Z}/q\mathbb{Z}$ de cardinal $r \geqslant 1$. On trie les relèvements dans $\mathbb{Z}$ des éléments de $\mathbf{a}$ de la façon suivante : $0 < a_1 < a_2 < \ldots < a_r \leqslant q$. On étend le domaine de définition de la suite des a_i par la relation $a_{rm+i} = qm + a_i$ pour tout $1 \leqslant i \leqslant r$ et tout $m \in \mathbb{Z}$. Par commodité, on identifiera le sous-ensemble $\mathbf{a}$ et la suite des a_i qui est à « croissance périodique »*. On définit la **densité** P de $\mathbf{a}$ par $P \coloneqq r/q$ et la suite $(e_n)_{n\in\mathbb{Z}}$ détectant la présence d'une valeur de $\mathbf{a}$ par $e_n \coloneqq [\exists i,\, n = a_i]$, qui vaut 1 si n est une valeur prise par la suite $\mathbf{a}$ et 0 sinon.

Pour évaluer la répartition locale de ces entiers, on se donne quatre fonctions de comptage élémentaires :

- $E_n(l) \coloneqq \mathrm{card}\{a_i \in [\![n, n+l-1]\!]\} = \sum_{i=0}^{l-1} e_{n+i}$, avec $n \in \mathbb{Z}$ et $l \geqslant 1$,
- $F_j(l) \coloneqq \mathrm{card}\{n \in [\![1, q]\!]\,;\, E_n(l) = j\}$, avec $l \geqslant 1$ et $0 \leqslant j \leqslant l$,
- $n(l) \coloneqq \mathrm{card}\{i \in [\![1, r]\!]\,;\, a_{i+1} - a_i = l\}$, avec $l \geqslant 1$,
- $N(l) \coloneqq \mathrm{card}\{i \in [\![1, r]\!]\,;\, a_{i+1} - a_i \geqslant l\} = \sum_{k\geqslant l} n(k)$, avec $l \geqslant 1$.

Les deux premières sont liées au nombre de valeurs de $\mathbf{a}$ dans un intervalle de longueur l, les deux dernières s'intéressent quant à elles aux écarts entre valeurs consécutives de la suite $\mathbf{a}$. On peut cependant établir des liens entre ces deux groupes de fonctions.

*On dira d'une suite de $\mathbb{Z}$ dans $\mathbb{Z}$ qu'elle est à croissance périodique si la suite de terme $a_{i+1} - a_i$ est strictement positive et périodique.

Lemme 1.1.1. — *Soit $l \geqslant 1$ un entier. Les quatre conditions suivantes sont équivalentes :*

a). *Pour tout $n \in \mathbb{Z}$, $E_n(l) \geqslant 1$;*
b). $F_0(l) = 0$ *;*
c). *Pour tout $j > l$, $n(j) = 0$;*
d). $N(l+1) = 0$.

Démonstration. — En effet,

$$\begin{aligned}\exists n \in \mathbb{Z},\ E_n(l) = 0 &\Leftrightarrow \exists n \in \mathbb{Z}, \exists i \in \mathbb{Z},\ a_{i+1} \geqslant n + l \geqslant n - 1 \geqslant a_i \\ &\Leftrightarrow \exists i \in \mathbb{Z},\ a_{i+1} - a_i \geqslant l + 1 \\ &\Leftrightarrow N(l+1) \geqslant 1.\end{aligned}$$

□

Il existe donc un seul entier l_0 tel que les $q+2$ entiers $E_1(l_0), \ldots, E_n(l_0), n(l_0), N(l_0)$ soient tous non nuls. Cet entier représente l'écart maximal entre deux valeurs de a_i consécutives. On définit alors la généralisation naturelle de la fonction de Jacobsthal

$$g(\mathbf{a}) \coloneqq \max_i\ a_{i+1} - a_i.$$

En outre, on peut donner une forme quantitative du lemme précédent.

Proposition 1.1.2. — *Soit $l \geqslant 1$ un entier. On a*

$$F_0(l) = \sum_{j>l} (j-l)\, n(j) = \sum_{j>l} N(j).$$

Lemme 1.1.3. — *Soient $l \geqslant 1$, n et n' trois entiers vérifiant $E_n(l) = 0$, $E_{n'}(l) = 0$ et $n \leqslant n' \leqslant n + l$. On a $E_m(l) = 0$ pour tout entier m vérifiant $n \leqslant m \leqslant n'$.*

Démonstration. — Comme $E_n(l) = 0$ et $E_{n'}(l) = 0$, aucun terme a_i n'est dans l'ensemble $[\![n, n+l-1]\!] \cup [\![n', n'+l-1]\!]$. Comme $n \leqslant n' \leqslant n+l$, ce dernier ensemble est $[\![n, n'+l-1]\!]$ et contient donc $[\![m, m+l-1]\!]$ pour tout $m \in [\![n, n']\!]$. □

Démonstration de la proposition 1.1.2. — Il vient du lemme 1.1.2 que, vu comme sous-ensemble de $\mathbb{Z}/q\mathbb{Z}$, l'ensemble $\{n \in [\![1, q]\!]; E_n(l) = 0\}$ est une réunion d'intervalles disjoints $[\![m+1, m+j]\!]$ (avec $j \geqslant 1$) et séparés l'un de l'autre par au moins l éléments. À chacun de ces intervalles $[\![m+1, m+j]\!]$ correspond un couple de valeurs consécutives de a_i — à savoir $a_i = m$ et $a_{i+1} = m + j + l$ — dont la différence vaut $j + l$. Il est clair que cette correspondance est biunivoque. On obtient donc

$$\begin{aligned}\operatorname{card}\{n \in [\![1, q]\!]\ ;\ E_n(l) = 0\} &= \operatorname{card} \bigcup_{m,j} [\![m+1, m+j]\!] \\ &= \sum_j \sum_{\substack{m = a_i \\ m+l+j = a_{i+1}}} j \\ &= \sum_j j\, n(l+j),\end{aligned}$$

ce qui donne le résultat en translatant la variable de sommation. □

Dans le but d'obtenir des informations sur la fonction g, on définit deux outils de mesure de répartition locale des valeurs de la suite $\mathbf{a}$, qui ont un comportement analytique plus commode et qui sont les généralisations des fonctions rencontrées dans les énoncés (0.9) et (0.11) :

– le moment centré d'ordre k :

$$M_k(l) \coloneqq \sum_{n=1}^{q} \big(E_n(l) - lP\big)^k = \sum_{j=0}^{l} (j - lP)^k F_j(l)\,,$$

ainsi que le moment absolu d'ordre κ, avec $\kappa \in \mathbb{R}_+$:

$$M_\kappa^+(l) \coloneqq \sum_{n=1}^{q} \big|E_n(l) - lP\big|^\kappa = \sum_{j=0}^{l} |j - lP|^\kappa F_j(l)\,;$$

– la variance de répartition d'ordre γ :

$$V_\gamma \coloneqq \sum_{i=1}^{r} (a_{i+1} - a_i)^\gamma = \sum_{l} l^\gamma\, n(l).$$

On a clairement $M_k(l) \leqslant M_k^+(l)$ pour tout entier naturel k, avec une égalité si k est pair. Tout ces fonctions dépendent de la suite $\mathbf{a}$. Dans cette partie, cette dépendance sera implicite ; mais on se réserve le droit de la rendre explicite par la suite.

Lemme 1.1.4. — *Soient β et β' deux réels positifs vérifiant $\beta + \beta' = 1$. Alors pour tous les réels positifs κ et κ', on a*

$$M^+_{\beta\kappa+\beta'\kappa'}(l) \leqslant \big(M_\kappa^+(l)\big)^\beta \big(M_{\kappa'}^+(l)\big)^{\beta'}.$$

De même, pour tous les réels γ et γ', on a

$$V_{\beta\gamma+\beta'\gamma'} \leqslant \big(V_\gamma\big)^\beta \big(V_{\gamma'}\big)^{\beta'}.$$

Il s'agit d'une version *ad hoc* de l'inégalité de Hölder. On va relier ces deux grandeurs à la fonction $F_0(l) = \sum_{j>l} N(j)$ intervenant dans la proposition 1.1.2.

Proposition 1.1.5. — *Pour tout réel $\kappa \geqslant 0$ et tout entier $l \geqslant 1$, on a*

$$F_0(l) \leqslant (lP)^{-\kappa}\, M_\kappa^+(l).$$

Pour tout réel $\gamma \geqslant 2$ et tout entier $L \geqslant 1$, on a

$$V_\gamma \leqslant rL^\gamma + \gamma(L+1)^{\gamma-1} F_0(L) + \gamma(\gamma-1) \sum_{l>L} (l+1)^{\gamma-2} F_0(l).$$

On rappelle que $V_0 = r$ et $V_1 = q$.

Démonstration. — La première inégalité ne pose pas de difficulté :

$$M_\kappa^+(l) = \sum_{j=0}^{l} |j - lP|^\kappa F_j(l) \geqslant (lP)^\kappa F_0(l).$$

Pour la seconde, on utilise deux transformations d'Abel consécutives

$$\begin{aligned} V_\gamma &= \sum_l l^\gamma\, n(l) \\ &= \sum_{l\leqslant L} l^\gamma\, n(l) + N(L+1)\, L^\gamma + \sum_{l>L} N(l)\, \big(l^\gamma - (l-1)^\gamma\big) \\ &= \sum_{l\leqslant L} l^\gamma\, n(l) + N(L+1)\, L^\gamma + F_0(L)\, \big((L+1)^\gamma - L^\gamma\big) \\ &\qquad\qquad + \sum_{l>L} F_0(l)\, \big((l+1)^\gamma + (l-1)^\gamma - 2l^\gamma\big). \end{aligned}$$

Le théorème des accroissements finis permet d'établir $\gamma l^{\gamma-1} \leqslant (l+1)^\gamma - l^\gamma \leqslant \gamma(l+1)^{\gamma-1}$ et $\gamma(\gamma-1)\,(l-1)^{\gamma-2} \leqslant (l+1)^\gamma + (l-1)^\gamma - 2l^\gamma \leqslant \gamma(\gamma-1)\,(l+1)^{\gamma-2}$. Enfin, on a

$$\sum_{l\leqslant L} n(l) l^\gamma + N(L+1)L^\gamma \leqslant \Big(\sum_{l\leqslant L} n(l) + N(L+1)\Big) L^\gamma = rL^\gamma,$$

ce qui permet d'obtenir la majoration proposée. □

Remarque. — Les définitions des fonctions $E_n(l)$, $F_j(l)$ et $M_k(l)$ peuvent être étendues aux ensembles d'entiers I par

- $E_n(I) \coloneqq \operatorname{card}\{i \in \mathbb{Z}\,;\, \exists j \in I,\, a_i = n + j\} = \sum_{i\in I} e_{n+i}$,
- $F_j(I) \coloneqq \operatorname{card}\{n \in [\![1, q]\!]\,;\, E_n(I) = j\}$, pour $j \leqslant \operatorname{card} I$,
- $M_k(I) \coloneqq \sum_{n=1}^{q} \big(E_n(I) - lP\big)^k = \sum_{j=0}^{l} (j - lP)^k F_j(I)$, où $l = \operatorname{card} I$.

Ces définitions recouvrent les définitions originelles dans le sens où l'on doit voir l'entier l comme représentant l'ensemble $I = [\![0, l-1]\!]$ de cardinal l. Cependant, cette spécificité des ensembles permet de définir une variante de $F_j(I)$: pour tout $J \subset I$, on pose

$$F_J^*(I) \coloneqq \operatorname{card}\{n \in [\![1, q]\!]\,;\, E_n(I) = \operatorname{card} J = E_n(J)\}, \tag{1.1}$$

si bien que pour $j \leqslant \operatorname{card} I$

$$F_j(I) = \sum_{\substack{J\subset I \\ \operatorname{card} J = j}} F_J^*(I). \tag{1.2}$$

On a également

$$F_J^*(I) = \sum_{n=1}^{q} \prod_{i\in J} e_{n+i} \prod_{i\in I\smallsetminus J} (1 - e_{n+i}) = \sum_{J'\subset I\smallsetminus J} (-1)^{\operatorname{card} J'} F_{J\cup J'}^*(J \cup J'). \tag{1.3}$$

1.2. Familles équiréparties

À partir de cet instant, lorsque l'on considère une suite à croissance périodique $\mathbf{a}$, on sous-entend également la donnée de q et de r (qui, sinon, seraient définis à un multiple commun près) ainsi que la donnée de P, de la suite e_n, etc.

On dit d'une famille $\mathcal{F}$ de suites à croissance périodique $\mathbf{a}$ qu'elle est $(\alpha_1, \alpha_2, \alpha_3)$-**équirépartie** si pour une constante absolue $c > 0$, l'estimation suivante

$$M_k(l) \ll c^k q k^{\alpha_1 k} (lP)^{\alpha_2 k} \big(\max(k, lP)\big)^{\alpha_3 k},$$

est uniformément vérifiée pour $\mathbf{a} \in \mathcal{F}$ et pour l et k deux entiers naturels[†], où α_1, α_2 et α_3 sont trois réels ($M_k(l)$, q et P dépendent implicitement de la suite $\mathbf{a}$). On ne considérera que le cas où les α_i sont positifs. À la valeur de la constante c près, il est équivalent de demander

$$M_k(l) \ll c^k q k^{\alpha_1 k} (lP)^{\alpha_2 k} \big(k^{\alpha_3 k} + (lP)^{\alpha_3 k}\big)$$

ou encore

$$M_k(l) \ll c^k q k^{\alpha_1 k} (lP)^{\alpha_2 k} (k + lP)^{\alpha_3 k},$$

et nous emploierons indistinctement ces différentes définitions.

Cette définition peut sembler artificielle ; elle l'est en partie. Cependant, sa forme et son utilisation sont assez souples pour supporter des modifications. L'essence de cette définition réside dans le changement de comportement asymptotique observé autour de l'égalité $k = lP$ (le terme à la puissance α_3). On verra que ce phénomène arrive tout naturellement : par exemple, le cas moyen s'assimile à une $(1/2, 0, 1/2)$-équirépartition (proposition 2.2.5 *infra*), et l'on ne peut faire mieux ni pour $k \ll lP$, ni pour $k \gg lP$.

Lemme 1.2.1. — *Une famille $\mathcal{F}$ de suites $\mathbf{a}$ est $(\alpha_1, \alpha_2, \alpha_3)$-équirépartie si et seulement si pour une constante absolue $c > 0$ on a uniformément*

$$M_\kappa^+(l) \ll c^\kappa q \kappa^{\alpha_1 \kappa} (lP)^{\alpha_2 \kappa} \big(\max(\kappa, lP)\big)^{\alpha_3 \kappa},$$

pour tout $\kappa \geqslant \log 2$ réel.

Démonstration. — On suppose qu'il existe une constante $C > 0$ telle qu'uniformément

$$M_k(l) \ll C^k q k^{\alpha_1 k} (lP)^{\alpha_2 k} \big(\max(k, lP)\big)^{\alpha_3 k}.$$

On choisit $k = 2\lceil \kappa/2 \rceil \geqslant \kappa$ et $\beta = \kappa/k \geqslant (\log 2)/2$, deux valeurs dépendant de κ. On a par le lemme 1.1.4

$$M_\kappa^+(l) \ll q^{1-\beta} \big(M_k^+(l)\big)^\beta$$

comme k est pair $M_k^+(l)$ vaut $M_k(l)$ et l'on utilise la majoration précédente

$$\ll C^{\beta k} q k^{\alpha_1 \beta k} (lP)^{\alpha_2 \beta k} \big(\max(k, lP)\big)^{\alpha_3 \beta k}$$

or $k = \kappa/\beta$, donc on peut majorer $\max(k, lP)$ par $\beta^{-1} \max(\kappa, lP)$

$$\begin{aligned} &\ll C^\kappa \beta^{-(\alpha_1+\alpha_3)\kappa} q \kappa^{\alpha_1 \kappa} (lP)^{\alpha_2 \kappa} \big(\max(\kappa, lP)\big)^{\alpha_3 \kappa} \\ &\ll \big(C(2/\log 2)^{\alpha_1+\alpha_3}\big)^\kappa q \kappa^{\alpha_1 \kappa} (lP)^{\alpha_2 \kappa} \big(\max(\kappa, lP)\big)^{\alpha_3 \kappa}, \end{aligned}$$

puisque $\beta^{-1} \leqslant 2/\log 2$. On obtient donc la condition nécessaire. La condition suffisante est évidente puisqu'on a $|M_k(l)| \leqslant M_k^+(l)$. □

[†]On pourra ajouter des conditions sur ces deux paramètres, celles-ci seront alors précisées.

Lemme 1.2.2. — *Soit $\mathcal{F}$ une famille $(\alpha_1, \alpha_2, \alpha_3)$-équirépartie. On a pour tout $\kappa \geqslant \log 2$ réel*

$$F_0(l) \ll c^\kappa q \kappa^{\alpha_1 \kappa} (lP)^{(\alpha_2 - 1)\kappa} \big(\max(\kappa, lP) \big)^{\alpha_3 \kappa}$$

uniformément pour $l \geqslant 1$ entier et pour $\mathbf{a} \in \mathcal{F}$, où la constante c est absolue.

Démonstration. — La proposition 1.1.5 et le lemme 1.2.1 fournissent le résultat. □

Proposition 1.2.3. — *Soit $\mathcal{F}$ une famille $(\alpha_1, \alpha_2, \alpha_3)$-équirépartie, avec $\alpha_2 + \alpha_3 < 1$. Posons*

$$\beta := \begin{cases} \frac{\alpha_1}{1-\alpha_2-\alpha_3} & \text{si } \alpha_1 + \alpha_2 + \alpha_3 \geqslant 1, \\ \frac{\alpha_1+\alpha_3}{1-\alpha_2} & \text{sinon.} \end{cases}$$

Alors, on a

$$g(\mathbf{a}) \ll P^{-1} (\log q)^\beta,$$

uniformément pour $\mathbf{a} \in \mathcal{F}$.

Dans cette proposition, on exclut l'étude des suites $\mathbf{a}$ avec $q = 1$ (il n'y en a qu'une !), qui est triviale et inintéressante.

Démonstration. — Par le lemme 1.2.2, on sait qu'il existe une constante absolue C telle que pour tout $\mathbf{a} \in \mathcal{F}$ et tout $\kappa \geqslant \log 2$ on a

$$F_0(l) \leqslant \Big(e^{\log q / \kappa} C \kappa^{\alpha_1} (lP)^{\alpha_2 - 1} \big(\max(\kappa, lP) \big)^{\alpha_3} \Big)^\kappa.$$

On choisit $\kappa = \log q \geqslant \log 2$, et on se pose la question de savoir à partir de quelle valeur de l le terme entre parenthèses devient strictement inférieur à 1. Supposons d'abord que $\alpha_1 + \alpha_2 + \alpha_3 \geqslant 1$; on a en particulier $\beta \geqslant 1$. Si $lP \geqslant \kappa$, il nous faut résoudre

$$(lP)^{1-\alpha_2-\alpha_3} \gg (\log q)^{\alpha_1},$$

ce qui est vérifié dès que $l \gg P^{-1} (\log q)^\beta$, condition qui n'est pas en contradiction avec l'hypothèse $lP \geqslant \kappa$. Ainsi on a bien

$$g(\mathbf{a}) \ll P^{-1} (\log q)^{\frac{\alpha_1}{1-\alpha_2-\alpha_3}}$$

si $\alpha_1 + \alpha_2 + \alpha_3 \geqslant 1$. On se place à présent dans le cas inverse ; en particulier, on a $\beta < 1$. Si $lP \leqslant \kappa$, il nous faut résoudre

$$(lP)^{1-\alpha_2} \gg (\log q)^{\alpha_1+\alpha_3},$$

qui est vérifié dès que $lP \gg (\log q)^{(\alpha_1+\alpha_3)/(1-\alpha_2)}$, ce qui est possible comme on a supposé $lP \leqslant \log q$. Dans ce cas, on a

$$g(\mathbf{a}) \ll P^{-1} (\log q)^{\frac{\alpha_1+\alpha_3}{1-\alpha_2}}.$$ □

Proposition 1.2.4. — *Soit $\mathcal{F}$ une famille $(\alpha_1, \alpha_2, \alpha_3)$-équirépartie, avec $\alpha_2 + \alpha_3 < 1$. Alors, on a*

$$V_\gamma \ll c^\gamma \gamma^{\gamma \max\left(1, \frac{\alpha_1}{1-\alpha_2-\alpha_3}\right)} q P^{-\gamma+1},$$

uniformément pour $\mathbf{a} \in \mathcal{F}$ et pour $\gamma \geqslant 0$, où la constante c est absolue.

En utilisant un simple argument de convexité (lemme 1.1.4) identique à celui du lemme 1.2.1, on peut montrer que ce résultat est également vrai pour $\gamma \geqslant 1$ réel.

Démonstration. — Les cas $\gamma = 0$ et $\gamma = 1$ sont triviaux, on peut supposer $\gamma \geqslant 2$. Des trois termes intervenant dans la majoration de V_γ apparaissant dans la proposition 1.1.5, le troisième est le plus délicat, car il faut que la somme en l converge. Quand l devient très grand (notamment devant κ/P), on a par le lemme 1.2.2

$$F_0(l) \ll^\kappa q\kappa^{\alpha_1\kappa}(lP)^{(\alpha_2+\alpha_3-1)\kappa};$$

pour que la somme en l soit convergente, il faut que $\gamma - 2 + (\alpha_2 + \alpha_3 - 1)\kappa < -1$, soit $\kappa > (\gamma-1)/(1-\alpha_2-\alpha_3) \geqslant 1$. On a donc

$$\begin{aligned}\sum_{l\geqslant\kappa/P} F_0(l)\,(l+1)^{\gamma-2} &\ll^\gamma \sum_{l\geqslant\kappa/P} F_0(l)\,l^{\gamma-2}\\ &\ll^{\kappa,\gamma} q\kappa^{\alpha_1\kappa}P^{(\alpha_2+\alpha_3-1)\kappa}\int_{\kappa/P}^{+\infty}\frac{x^{\gamma-2}}{x^{(1-\alpha_2-\alpha_3)\kappa}}\,\mathrm{d}x\\ &\ll^{\kappa,\gamma} q\kappa^{\alpha_1\kappa}P^{(\alpha_2+\alpha_3-1)\kappa}\frac{(\kappa/P)^{(\alpha_2+\alpha_3-1)\kappa+\gamma-1}}{(1-\alpha_2-\alpha_3)\kappa-\gamma+1}\\ &\ll^{\kappa,\gamma} qP^{-\gamma+1}\frac{\kappa^{(\alpha_1+\alpha_2+\alpha_3-1)\kappa+\gamma-1}}{(1-\alpha_2-\alpha_3)\kappa-\gamma+1}.\end{aligned}$$

Ici on fixe $\kappa = \gamma/(1-\alpha_2-\alpha_3)$. Ainsi, on a en posant $c_1 = 1/(1-\alpha_2-\alpha_3)$

$$\gamma(\gamma-1)\sum_{l\geqslant c_1\gamma/P} F_0(l)\,(l+1)^{\gamma-2} \ll^\gamma \gamma^{\alpha_1 c_1\gamma} qP^{-\gamma+1}.$$

Ceci constitue le noyau dur de l'estimation. Pour le reste, on peut simplement choisir $L = 1$ et utiliser la majoration $F_0(l) \leqslant q$. On a alors

$$\begin{aligned}rL^\gamma + \gamma(L+1)^{\gamma-1}F_0(L) + \gamma(\gamma-1)\sum_{L<l\leqslant c_1\gamma/P}(l+1)^{\gamma-2}F_0(l)&\\ \ll q\gamma(\gamma-1)\sum_{l\leqslant c_1\gamma/P}(l+1)^{\gamma-2} \quad &\ll^\gamma qP^{-\gamma+1}\gamma^\gamma.\end{aligned}$$

Ce terme n'est prépondérant que si $\alpha_1+\alpha_2+\alpha_3 < 1$. □

Il est certainement possible d'améliorer la majoration de la proposition 1.2.4 pour $\alpha_1+\alpha_2+\alpha_3 < 1$, et trouver un $\gamma^{\alpha_1 c_1\gamma}$ à la place du γ^γ ; mais la version proposée est tout de même assez précise pour établir le résultat suivant.

Proposition 1.2.5. — *Soit $\mathcal{F}$ une famille $(\alpha_1,\alpha_2,\alpha_3)$-équirépartie, avec $\alpha_2+\alpha_3 < 1$ et $\alpha_1+\alpha_2+\alpha_3 \leqslant 1$. Alors il existe un réel $x > 0$ tel que*

$$\sum_{i=1}^{r} e^{(a_{i+1}-a_i)Px} \ll r, \qquad \textit{uniformément pour } \mathbf{a}\in\mathcal{F}.$$

En particulier, on a

$$g(\mathbf{a}) \ll P^{-1}\log r, \qquad \textit{uniformément pour } \mathbf{a}\in\mathcal{F}.$$

Démonstration. — Il s'agit donc de montrer que la série entière

$$\frac{1}{r}\sum_{\gamma\geqslant 0}\frac{V_\gamma P^\gamma}{\gamma!}x^\gamma$$

possède un rayon de convergence non nul et ne dépendant pas de la suite $\mathbf{a}$, et que l'on peut en borner les valeurs indépendamment de la suite $\mathbf{a}$. On sait par la proposition 1.2.4 qu'il existe une constante absolue C telle que soit vérifée uniformément la majoration

$$V_\gamma \leqslant C^\gamma \gamma^{\gamma \max\left(1,\frac{\alpha_1}{1-\alpha_2-\alpha_3}\right)}\, qP^{-\gamma+1} \leqslant C^\gamma\gamma^\gamma rP^{-\gamma}.$$

En effet, comme $\alpha_1+\alpha_2+\alpha_3\leqslant 1$, on a $\alpha_1/(1-\alpha_2-\alpha_3)\leqslant 1$. Combinée au fait que pour tout $\gamma\geqslant 0$ on a $\gamma!\geqslant \big(\gamma/e\big)^\gamma$, cette estimation donne

$$\frac{1}{r}\sum_{\gamma\geqslant 0}\frac{V_\gamma P^\gamma}{\gamma!}x^\gamma \leqslant \sum_{\gamma\geqslant 0}(Ce)^\gamma x^\gamma.$$

En prenant par exemple $x=(2Ce)^{-1}$, on a uniformément pour $\mathbf{a}\in\mathcal{F}$

$$\sum_{i=1}^{r} e^{(a_{i+1}-a_i)Px} \ll r.$$

Un des r termes de la somme vaut $e^{g(\mathbf{a})Px}$, d'où la seconde estimation. □

1.3. Suites décomposables

À l'instar des caractères de Dirichlet, il est possible de construire une « suite produit » à partir de deux suites $\mathbf{a}^{(1)}$ et $\mathbf{a}^{(2)}$.

Lemme 1.3.1. — *Soit κ un réel positif. Pour tous les x et y réels positifs, on a*

$$(x+y)^\kappa \leqslant 2^\kappa(x^\kappa+y^\kappa).$$

Démonstration. — En effet, on a

$$(x+y)^\kappa \leqslant \big(2\max(x,y)\big)^\kappa = 2^\kappa \max(x^\kappa,y^\kappa)\leqslant 2^\kappa(x^\kappa+y^\kappa).$$ □

Évidemment, par convexité, on peut gagner un facteur 2 si $\kappa\geqslant 1$. Mais cela n'est pas nécessaire pour nos applications.

Proposition 1.3.2. — *Soit κ un réel positif. Pour tous les l, l' entiers naturels et toutes les suites $\mathbf{a}$, on a*

$$M_\kappa^+(l+l')\leqslant 2^\kappa\Big(M_\kappa^+(l)+M_\kappa^+(l')\Big),$$

et si $l'\leqslant l$,

$$M_\kappa^+(l-l')\leqslant 2^\kappa\Big(M_\kappa^+(l)+M_\kappa^+(l')\Big).$$

On dit d'une suite $\mathbf{a}$ (pour laquelle sont définies les entiers r et q, ainsi que la densité P et la fonction de détection $(e_n)_{n\in\mathbb{Z}}$) qu'elle est **décomposable** en deux suites $\mathbf{a}^{(1)}$ (associée à r_1, q_1, P_1 et $(e_n^{(1)})_{n\in\mathbb{Z}}$) et $\mathbf{a}^{(2)}$ (associée à r_2, q_2, P_2 et $(e_n^{(2)})_{n\in\mathbb{Z}}$) si sont verifiées les conditions suivantes :

a). On a pour tout entier $n \in \mathbb{Z}$ l'équivalence

$$\exists i,\ n = a_i \Leftrightarrow \exists i_1, i_2,\ n = a_{i_1}^{(1)} = a_{i_2}^{(2)}.$$

b). On a $q = q_1q_2$, $q_1 > 1$, $q_2 > 1$ et $(q_1, q_2) = 1$.

La condition a) est équivalente à l'égalité $e_n = e_n^{(1)}e_n^{(2)}$, pour tout $n \in \mathbb{Z}$.
La condition b) sert essentiellement à se prévenir de décompositions triviales. Mais il est à remarquer qu'elle implique les décompositions $r = r_1r_2$ et $P = P_1P_2$.

Proposition 1.3.3. — *Soient l un entier naturel, κ un réel positif et $\mathbf{a}$ une suite décomposable en deux suites suites $\mathbf{a}^{(1)}$ et $\mathbf{a}^{(2)}$. On conserve les notations précédemment introduites. Pour m entier on pose $I_m = \{i \in [\![0, l-1]\!]\,;\, e_{m+i}^{(1)} = 1\}$. Alors on a*

$$M_\kappa^+(l;\mathbf{a}) \leqslant 2^\kappa\Big(q_2{P_2}^\kappa M_\kappa^+(l;\mathbf{a}^{(1)}) + \sum_{m=1}^{q_1} M_\kappa^+(I_m;\mathbf{a}^{(2)})\Big).$$

Démonstration. — Puisque q_1 et q_2 sont premiers entre eux, on sait par le théorème chinois des restes que la relation de congruence $n \equiv n_1c_2q_2 + n_2c_1q_1\,[q]$, où c_1 (resp. c_2) est l'inverse de q_1 (resp. q_2) modulo q_2 (resp. q_1), met en correspondance biunivoque les entiers n de $[\![1, q]\!]$ et les couples d'entiers (n_1, n_2) de $[\![1, q_1]\!] \times [\![1, q_2]\!]$. En particulier, elle met en correspondance l'entier $n = 1$ avec le couple $(n_1, n_2) = (1, 1)$. Ainsi, pour tout entier i, on a

$$n + i \equiv (n_1 + i)c_2q_2 + (n_2 + i)c_1q_1\ [q],$$

ce qui donne $e_{n+i} = e_{n_1+i}^{(1)}e_{n_2+i}^{(2)}$. On peut donc écrire

$$\begin{aligned}\sum_{i=0}^{l-1} e_{n+i}\ - lP &= \sum_{i=0}^{l-1} e_{n_1+i}^{(1)}e_{n_2+i}^{(2)}\ - lP \\ &= \Big(P_2\sum_{i=0}^{l-1} e_{n_1+i}^{(1)}\ - lP\Big) + \Big(\sum_{i=0}^{l-1} e_{n_1+i}^{(1)}e_{n_2+i}^{(2)}\ - P_2\sum_{i=0}^{l-1} e_{n_1+i}^{(1)}\Big).\end{aligned}$$

On a $P = P_1P_2$. On note l_{n_1} le cardinal de I_{n_1}. On obtient alors

$$\sum_{i=0}^{l-1} e_{n_1+i}^{(1)}\ - lP = P_2\Big(\sum_{i=0}^{l-1} e_{n_1+i}^{(1)}\ - lP_1\Big) + \Big(\sum_{i\in I_{n_1}} e_{n_2+i}^{(2)}\ - l_{n_1}P_2\Big).$$

Par le lemme 1.3.1, on a en passant à la valeur absolue puis à la puissance κ, et en sommant sur n

$$M_\kappa^+(l;\mathbf{a}) \leqslant 2^\kappa\Big(q_2{P_2}^k M_\kappa^+(l;\mathbf{a}^{(1)}) + \sum_{n_1=1}^{q_1} M_\kappa^+(I_{n_1};\mathbf{a}^{(2)})\Big). \qquad \square$$

Pour traiter le second terme, il est nécessaire d'en savoir un peu plus sur la suite $\mathbf{a}^{(2)}$.

Proposition 1.3.4. — *Soient $y > 1$ un réel, k un entier naturel pair, $\mathbf{a}$ une suite décomposable en deux suites $\mathbf{a}^{(1)}$ et $\mathbf{a}^{(2)}$. On suppose que*

- *on a $M_k(I;\mathbf{a}^{(2)}) = M_k(\operatorname{card} I;\mathbf{a}^{(2)})$ pour tout $I \subset [\![0, y-1]\!]$;*
- *il existe $(\alpha_1,\alpha_2,\alpha_3) \in [0,1]^3$ et $C > 0$ un réel tels que $\alpha_2+\alpha_3 \leqslant 1$ et*

$$M_k(l;\mathbf{a}^{(2)}) \leqslant Cq_2k^{\alpha_1 k}(lP_2)^{\alpha_2 k}\big(k^{\alpha_3 k} + (lP_2)^{\alpha_3 k}\big),$$

uniformément pour tout entier naturel $l \leqslant y$.

Pour tout $l \leqslant y$, on a, en notant $c_k \coloneqq c_k(l,\mathbf{a}^{(1)}) = \big({q_1}^{-1}(lP_1)^{-k}M_k(l;\mathbf{a}^{(1)})\big)^{1/k}$,

$$M_k(l;\mathbf{a}) \leqslant 2^k\,{q_2}{P_2}^k M_k(l;\mathbf{a}^{(1)}) + 4^k C\,qk^{\alpha_1 k}(lP)^{\alpha_2 k}\big(k^{\alpha_3 k} + (lP)^{\alpha_3 k}\big) + 4^k C\,qk^{\alpha_1 k}(c_k\,lP)^{\alpha_2 k}\big(k^{\alpha_3 k} + (c_k\,lP)^{\alpha_3 k}\big).$$

Remarque. — Les deux hypothèses sont que, sous la condition « l assez petit », les moments centrés ne dépendent pas de la « géométrie » de la fenêtre I — c'est-à-dire de la façon dont sont répartis les $|I|$ éléments de I dans $[\![0, y-1]\!]$ — et l'on a une famille $(\alpha_1,\alpha_2,\alpha_3)$-équirépartie, avec $\alpha_2+\alpha_3 \leqslant 1$. C'est exactement ce que notre raisonnement heuristique va mettre en évidence au chapitre suivant.
On remarque également que le dernier des trois termes est absorbé par le deuxième (à la valeur de la constante C près) si $M_k(l;\mathbf{a}^{(1)}) \leqslant C' q_1 (lP_1)^k$: on a alors $c_k \leqslant C'^{1/k}$.

Démonstration. — On utilise les notations et les résultats de la proposition 1.3.3 : il s'agit de majorer la somme

$$\sum_{n_1=1}^{q_1} M_k(I_{n_1};\mathbf{a}^{(2)}).$$

On suppose $l \leqslant y$. On a $I_{n_1} \subset [\![0, l-1]\!]$ pour tout n_1, on peut donc appliquer la première des deux hypothèses et obtenir

$$\sum_{n_1=1}^{q_1} M_k(I_{n_1};\mathbf{a}^{(2)}) = \sum_{n_1=1}^{q_1} M_k(l_{n_1};\mathbf{a}^{(2)}),$$

où $l_{n_1} = \operatorname{card} I_{n_1}$ et donc $l_{n_1} \leqslant l \leqslant y$. On peut à présent appliquer la seconde hypothèse, à savoir qu'il existe une constante absolue C telle qu'uniformément pour $l \leqslant y$, on a

$$M_k(l_{n_1};\mathbf{a}^{(2)}) \leqslant Cq_2k^{\alpha_1 k}(l_{n_1}P_2)^{\alpha_2 k}\big(k^{\alpha_3 k} + (l_{n_1}P_2)^{\alpha_3 k}\big).$$

Or pour tout $\alpha \in [0,1]$, on a par le lemme 1.3.1 puis par inégalité de Hölder (k est pair)

$$\begin{aligned}\sum_{n_1=1}^{q_1} {l_{n_1}}^{\alpha k} &\leqslant 2^{\alpha k}\Big(\sum_{n_1=1}^{q_1} |l_{n_1} - lP_1|^{\alpha k} + q_1(lP_1)^{\alpha k}\Big)\\ &\leqslant 2^{\alpha k}\Big({q_1}^{1-\alpha}\big(M_k(l;\mathbf{a}^{(1)})\big)^{\alpha} + q_1(lP_1)^{\alpha k}\Big).\end{aligned}$$

On obtient en utilisant cette majoration avec $\alpha = \alpha_2$ et avec $\alpha = \alpha_2 + \alpha_3 \leqslant 1$

$$\sum_{n_1=1}^{q_1} M_k(I_{n_1}; \mathbf{a}^{(2)}) \leqslant 2^k C q k^{\alpha_1 k} (lP)^{\alpha_2 k} \big(k^{\alpha_3 k} + (lP)^{\alpha_3 k}\big) + 2^k C q k^{\alpha_1 k} \Big(\frac{1}{q_1} M_k(l; \mathbf{a}^{(1)}) {P_2}^k\Big)^{\alpha_2} \Big(k^{\alpha_3 k} + \big(\tfrac{1}{q_1} M_k(l; \mathbf{a}^{(1)}) {P_2}^k\big)^{\alpha_3}\Big),$$

qui fournit l'inégalité proposée en la replaçant dans le résultat de la proposition 1.3.3. □

Remarque. — La première hypothèse est une sorte de condition d'indépendance combinatoire locale pour la suite $\mathbf{a}^{(2)}$: on demande en effet que, en moyenne, le nombre de $m \in [\![1, n_2]\!]$ vérifiant $e^{(2)}_{m+i_1} = e^{(2)}_{m+i_2} = \cdots = e^{(2)}_{m+i_l} = 1$ pour $0 < i_1 < i_2 < \cdots < i_l \leqslant y$ ne dépende pas du choix des i_j mais seulement de leur nombre l. Il est possible d'établir une version de la proposition 1.3.4 assouplissant cette condition mais faisant apparaître en contrepartie un quatrième terme d'erreur. Nous ne l'avons pas fait car c'est sous cette version-ci que nous utiliserons cette proposition.

La seconde condition imposée à la suite $\mathbf{a}^{(2)}$ est bien plus classique : il s'agit d'un version affaiblie et précisée à la fois de la $(\alpha_1, \alpha_2, \alpha_3)$-équirépartition, affaiblie car elle n'est nécessaire que pour de petits l ($l \leqslant y$), précisée car on impose $\alpha_2 + \alpha_3 \leqslant 1$[‡]. Par exemple, si l'on demande aux suites $\mathbf{a}^{(1)}$ et $\mathbf{a}^{(2)}$ d'être $(1/2, 0, 1/2)$-équiréparties, alors la suite $\mathbf{a}$ le sera également ; cet argument sera exploité plus loin pour obtenir la proposition 5.1.1 dans le cas de la répartition des entiers premiers à un entier fixé.

[‡]Cette condition est de toute façon nécessaire pour appliquer les propositions 1.2.3, 1.2.4 et 1.2.5.

CHAPITRE 2

MODÈLE PROBABILISTE

Nous continuons notre étude générale du problème, mais cette fois sous l'angle probabiliste. Le but est de préciser notre intuition à propos du moment M_k. Nous allons étudier le comportement moyen de ce moment, en espérant que nos suites particulières vérifient les mêmes propriétés. Pour ce, nous allons faire l'étude d'une suite de polynômes particuliers, que nous retrouvons également au chapitre suivant, avant de calculer le comportement asymptotique des moments centrés d'une loi binomiale.

2.1. Polynômes de Romanovsky

Nous consacrons cette section à l'étude d'une suite de polynômes, que nous baptisons « polynômes de Romanovsky »*, en occultant entièrement leur signification en termes de probabilités. Cette interprétation probabiliste sera accomplie en temps voulu, justifiant et motivant ainsi leur étude.

On définit les polynômes $R_{k,j} \in \mathbb{Z}[X]$ par les valeurs initiales

$$R_{0,j} = [j = 0],\ R_{1,j} = 0 \text{ et } R_{k,0} = [k = 0], \text{ pour tous les entiers naturels } j \text{ et } k$$

et par la formule de récurrence valable pour $j \geqslant 1$ et $k \geqslant 1$

$$R_{k+1,j}(X) = kR_{k-1,j-1}(X) + j(1-2X)R_{k,j}(X) + X(1-X)R'_{k,j}(X).$$

Un simple raisonnement par récurrence grâce à la formule définissant les $R_{k,j}$ permet de prouver les propriétés simples qui apparaissent sur les premières valeurs, que nous réunissons dans une table pour le confort du lecteur.

Lemme 2.1.1. — *Soient j et k deux entiers naturels.*

a). $R_{k,j}(1-X) = (-1)^k R_{k,j}(X)$.

b). *Si k est pair, on a $R_{k,j} \in \mathbb{Z}\big[X(1-X)\big]$, et si k est impair, on a $R_{k,j} \in (1-2X)\cdot\mathbb{Z}\big[X(1-X)\big]$.*

c). *Si $1 \leqslant j \leqslant k/2$, on a $\deg R_{k,j} = k - 2j$.*

*Cette dénomination est contestable ; elle doit être interprétée plus comme un hommage que comme issue d'une réalité historique.

d). *Si $j > k/2$, on a $R_{k,j}(X) = 0$.*

TABLE 1. Valeurs des premiers polynômes de Romanovsky $R_{k,j}$

$k \diagdown j$	0	1	2	3
0	1	0	0	0
1	0	0	0	0
2	0	1	0	0
3	0	$1-2X$	0	0
4	0	$1-6X(1-X)$	3	0
5	0	$(1-2X)\big(1-12X(1-X)\big)$	$10(1-2X)$	0
6	0	$1-30X(1-X)+120X^2(1-X)^2$	$25-130X(1-X)$	15

2.1.1. Majoration en norme. — Nous allons estimer la norme $\|\cdot\|_1$ de ces polynômes. On rappelle que cette norme est définie par $\|\sum_i c_i X^i\|_1 = \sum_i |c_i|$, que c'est une norme d'algèbre de $\mathbb{R}[X]$ (*i.e.* on a $\|P\cdot Q\|_1 \leqslant \|P\|_1 \cdot \|Q\|_1$) et qu'elle vérifie l'inégalité $\|P'\|_1 \leqslant \deg P\, \|P\|_1$.

Lemme 2.1.2. — *Pour tous les entiers $j \geqslant 1$ et $k \geqslant 1$, on a*

$$\big\|R_{k+1,j}\big\|_1 \leqslant k\big\|R_{k-1,j-1}\big\|_1 + (2k-j)\big\|R_{k,j}\big\|_1.$$

Démonstration. — On utilise la relation de récurrence définissant les polynômes de Romanovsky pour majorer $\big\|R_{k,j}\big\|_1$; on obtient donc

$$\begin{aligned}\big\|R_{k+1,j}(X)\big\|_1 &= \big\|kR_{k-1,j-1}(X) + j(1-2X)R_{k,j}(X) + X(1-X)R'_{k,j}(X)\big\|_1\\ &\leqslant k\big\|R_{k-1,j-1}\big\|_1 + j\|1-2X\|_1\big\|R_{k,j}\big\|_1 + \|X(1-X)\|_1\big\|R'_{k,j}\big\|_1\\ &\leqslant k\big\|R_{k-1,j-1}\big\|_1 + \big(3j+2(k-2j)\big)\big\|R_{k,j}\big\|_1\\ &\leqslant k\big\|R_{k-1,j-1}\big\|_1 + (2k-j)\big\|R_{k,j}\big\|_1.\end{aligned}$$

□

Pour tous les entiers naturels j et k, on pose

$$T_{k,j} \coloneqq \mathbb{\Gamma}(2k-2j+1)\binom{k-j-1}{j-1}. \tag{2.1}$$

En particulier si $k \geqslant j$, on a $T_{k,j} = (2k-2j)!!\binom{k-j-1}{j-1}$ et $T_{k,j} = 0$ sinon. On rappelle que les coefficients binomiaux $\binom{n}{k}$ valent $\frac{n!}{k!(n-k)!}$ si $0 \leqslant k \leqslant n$, $(-1)^{n-k}\frac{(-k-1)!}{(-n-1)!(n-k)!}$ si $k \leqslant n < 0$ et valent 0 sinon, si bien que les relations $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$ et $\binom{n}{k} = (-1)^{n-k}\binom{-k-1}{-n-1}$ sont vérifiées pour tous les n et k dans $\mathbb{Z}$. On sait que $\binom{n}{k} \leqslant 2^n$ si $n \in \mathbb{N}$ et, en utilisant les approximations de Stirling, que $(2n)!! \leqslant \sqrt{2}(2n/e)^n$ pour tout $n \in \mathbb{N}$; on a donc

$$T_{k,j} \leqslant \frac{1}{\sqrt{2}}\Big(\frac{4(k-j)}{e}\Big)^{k-j}, \qquad \text{pour } k \geqslant j \geqslant 1. \tag{2.2}$$

TABLE 2. Premières valeurs de $\|R_{k,j}\|_1$ et de $T_{k,j}$

$k \diagdown j$	0	1	2	3
0	1	0	0	0
1	0	0	0	0
2	0	1	0	0
3	0	3	0	0
4	0	13	3	0
5	0	75	30	0
6	0	541	285	15

$k \diagdown j$	0	1	2	3
0	1	0	0	0
1	0	0	0	0
2	0	1	0	0
3	0	3	0	0
4	0	15	3	0
5	0	105	30	0
6	0	945	315	15

Lemme 2.1.3. — *On a $T_{k,0} = [k = 0]$ pour tout k entier naturel. On a $T_{k,j} = 0$ si $j > k/2$. Pour $k \geqslant 1$ et $j \geqslant 1$, on a*

$$T_{k+1,j} = k\, T_{k-1,j-1} + (2k - j)\, T_{k,j}.$$

Démonstration. — Pour $j = 0$, on a bien $T_{k,0} = (2k)!!\binom{k-1}{-1} = (2k)!![k = 0] = [k = 0]$. Pour $j > k/2$, on a $k - j - 1 < j - 1$, donc $\binom{k-j-1}{j-1}$ est nul, et $T_{k,j}$ aussi.
Soient $k \geqslant 1$ et $j \geqslant 1$. On pose $m = k - j$. On a donc

$$\binom{m-1}{j-1} = \frac{1}{\mathbb{\Gamma}(2m+1)} T_{j+m,j}.$$

La relation d'absorption $(j-1)\binom{m-1}{j-1} = (m-j+1)\binom{m-1}{j-2}$ se traduit par $(j-1)\frac{1}{\mathbb{\Gamma}(2m+1)} T_{j+m,j} = (m-j+1)\frac{1}{\mathbb{\Gamma}(2m+1)} T_{j+m-1,j-1}$, soit encore

$$(j-1)T_{k,j} = (k-2j+1)T_{k-1,j-1}.$$

On effectue le même raisonnement pour la relation de récurrence $\binom{m}{j-1} = \binom{m-1}{j-2} + \binom{m-1}{j-1}$, qui se traduit alors par

$$\frac{1}{\mathbb{\Gamma}(2m+3)} T_{j+m+1,j} = \frac{1}{\mathbb{\Gamma}(2m+1)} T_{j+m-1,j-1} + \frac{1}{\mathbb{\Gamma}(2m+1)} T_{j+m,j},$$

soit en utilisant la relation $\mathbb{\Gamma}(s+2) = s\mathbb{\Gamma}(s)$, et en revenant sur le changement de variable

$$\begin{aligned} T_{k+1,j} &= (2k - 2j + 1)\left(T_{k-1,j-1} + T_{k,j}\right) \\ &= k\, T_{k-1,j-1} + (2k - j)\, T_{k,j} + \Big((k-2j+1)T_{k-1,j-1} - (j-1)T_{k,j}\Big), \end{aligned}$$

où le terme entre parenthèses est effectivement nul. □

Proposition 2.1.4. — *Pour tous les entiers naturels j et k, on a*

$$\|R_{k,j}\|_1 \leqslant T_{k,j},$$

où $T_{k,j}$ est la valeur définie en (2.1).

Démonstration. — Par récurrence, on déduit des lemmes 2.1.2 et 2.1.3 que $\|R_{k,j}\|_1 \leqslant T_{k,j}$ pour tous les entiers naturels j et k. □

2.1.2. Minoration des termes constants. — La relation de récurrence des polynômes de Romanovsky se simplifie pour les termes constants en

$$R_{k+1,j}(0) = kR_{k-1,j-1}(0) + jR_{k,j}(0).$$

On rappelle qu'on connaît les premières valeurs : $R_{0,j}(0) = [j = 0]$, $R_{1,j}(0) = 0$ et $R_{k,0}(0) = [k = 0]$ pour tous les entiers naturels j et k. On peut en tirer les premières conséquences.

Lemme 2.1.5. — *Soient j et k deux entiers naturels.*

a). *On a $R_{k,j}(0) \geqslant 0$.*
b). *Si $j > k/2$, on a $R_{k,j}(0) = 0$.*
c). *On a $R_{2k,k} = (2k)!!$.*
d). *Si $k \geqslant 2$, on a $R_{k,1}(0) = 1$.*

On pose pour l et j entiers naturels

$$r_{l,j} := \frac{1}{\mathbb{\Gamma}(2j+l+1)} R_{2j+l,j}(0). \tag{2.3}$$

TABLE 3. Premières valeurs de $R_{k,j}(0)$ et de leur normalisation $r_{l,j}$

$k \diagdown j$	0	1	2	3
0	1	0	0	0
1	0	0	0	0
2	0	1	0	0
3	0	1	0	0
4	0	1	3	0
5	0	1	10	0
6	0	1	25	15

$l \diagdown j$	0	1	2	3
0	1	1	1	1
1	0	$1/\sqrt{2\pi}$	$5/2\sqrt{2\pi}$	...
2	0	$1/3$	$5/3$	...
3	0	$1/4\sqrt{2\pi}$	...	...
4	0	$1/15$	...	...
5	0	...	...	...
6	0	...	...	...

Le facteur $1/\sqrt{2\pi}$ provient du fait que $\mathbb{\Gamma}(2) = \sqrt{2/\pi}$. Cette normalisation nécessite de modifier la relation de récurrence.

Lemme 2.1.6. — *Soient l et j deux entiers naturels. On a $r_{0,j} = 1$ et $r_{l,0} = [l = 0]$. Si $l \geqslant 1$ et $j \geqslant 1$, on a*

$$r_{l,j} = r_{l,j-1} + j\frac{\mathbb{\Gamma}(2j+l)}{\mathbb{\Gamma}(2j+l+1)} r_{l-1,j}.$$

Démonstration. — Par le lemme 2.1.5, on a $r_{0,j} = \frac{1}{(2j)!!} R_{2j,j}(0) = 1$ et par définition des valeurs initiales des polynômes de Romanovsky, on a $r_{l,0} = \frac{1}{\mathbb{\Gamma}(l+1)} R_{l,0}(0) =$

$\frac{1}{\mathbb{\Gamma}(l+1)}[l=0]=[l=0]$. En outre, pour $l,j\geqslant 1$ on a

$$\begin{aligned}
r_{l,j}&=\frac{1}{\mathbb{\Gamma}(2j+l+1)}R_{2j+l,j}(0)\\
&=\frac{1}{\mathbb{\Gamma}(2j+l+1)}\Big((2j+l-1)R_{2j+l-2,j-1}(0)+jR_{2j+l-1,j}(0)\Big)\\
&=\frac{1}{\mathbb{\Gamma}(2j+l-1)}R_{2j+l-2,j}(0)+j\frac{\mathbb{\Gamma}(2j+l)}{\mathbb{\Gamma}(2j+l+1)}\frac{1}{\mathbb{\Gamma}(2j+l)}R_{2j+l-1,j}(0)\\
&=r_{l,j-1}+j\frac{\mathbb{\Gamma}(2j+l)}{\mathbb{\Gamma}(2j+l+1)}r_{l-1,j}.
\end{aligned}$$

□

Proposition 2.1.7. — *Pour une constante absolue $c>0$, on a pour tout $\alpha\in]0,1]$ et pour $j\geqslant\alpha l\geqslant 0$*

$$r_{l,j}\geqslant c^l\alpha^{l/2}\frac{j^{3l/2}}{l!}.$$

Démonstration. — Par itération de la relation du lemme 2.1.6, on a pour $l\geqslant 1$

$$r_{l,j}=\sum_{i\leqslant j}i\frac{\mathbb{\Gamma}(2i+l)}{\mathbb{\Gamma}(2i+l+1)}\,r_{l-1,i}.$$

Nous avons par la formule de Stirling que pour $t\to+\infty$, on a $\mathbb{\Gamma}(t+1)/\mathbb{\Gamma}(t)\sim\sqrt{t/e}$. Posons K le maximum de $\sqrt{t/3}\,\mathbb{\Gamma}(t)/\mathbb{\Gamma}(t+1)$ pour $t\geqslant 0$. On obtient donc

$$r_{l,j}\geqslant\sum_{i\leqslant j}i\frac{K\sqrt{3}}{\sqrt{2i+l}}\,r_{l-1,i}.$$

En oubliant les premiers termes de la somme, on a pour tout $\alpha\in]0,1]$

$$r_{l,j}\geqslant\sum_{\alpha l\leqslant i\leqslant j}i\frac{K\sqrt{3}}{\sqrt{2i+l}}\,r_{l-1,i}\geqslant\sum_{\alpha l\leqslant i\leqslant j}K\sqrt{\alpha i}\sqrt{\frac{3}{2\alpha+1}}\,r_{l-1,i}\geqslant K\sqrt{\alpha}\sum_{\alpha l\leqslant i\leqslant j}\sqrt{i}\,r_{l-1,i}.$$

Nous allons prouver par récurrence sur l que pour tout réel $\alpha\in]0,1]$ et pour tout $j\geqslant\alpha l$ on a

$$r_{l,j}\geqslant\left(K\frac{1-e^{-3/2}}{3/2}\right)^l\alpha^{l/2}\frac{j^{3l/2}}{l!}.$$

Pour $l=0$, on a par le lemme 2.1.6 l'inégalité (et même l'égalité!) pour tout $j\geqslant 0$, puisqu'alors $r_{0,j}=1$.
Soit à présent $l\geqslant 1$. Soit $\alpha\in]0,1]$ un réel. Si pour $i\geqslant\alpha(l-1)$ il est vrai que

$$r_{l-1,i}\geqslant\left(K\frac{1-e^{-3/2}}{3/2}\right)^{l-1}\alpha^{(l-1)/2}\frac{i^{3(l-1)/2}}{(l-1)!},$$

on a pour $j\geqslant\alpha l$

$$r_{l,j}\geqslant K\left(K\frac{1-e^{-3/2}}{3/2}\right)^{l-1}\alpha^{l/2}\sum_{\alpha l\leqslant i\leqslant j}\frac{i^{3l/2-1}}{(l-1)!}.$$

On a $\lceil \alpha l \rceil \geqslant 1$, donc

$$\begin{aligned} r_{l,j} &\geqslant K\left(K\frac{1-e^{-3/2}}{3/2}\right)^{l-1}\alpha^{l/2}\int_{\lceil \alpha l\rceil -1}^{j}\frac{x^{3l/2-1}}{(l-1)!}\,\mathrm{d}x \\ &= K\left(K\frac{1-e^{-3/2}}{3/2}\right)^{l-1}\alpha^{l/2}\frac{j^{3l/2}}{\frac{3}{2}l!}\left(1-\left(\frac{\lceil \alpha l\rceil -1}{j}\right)^{3l/2}\right) \end{aligned}$$

or $j \geqslant \lceil \alpha l \rceil$ et $l \geqslant \lceil \alpha l \rceil$, ce qui donne $\left((\lceil \alpha l\rceil -1)/j\right)^l \leqslant \left(1-1/\lceil \alpha l\rceil\right)^l \leqslant e^{-l/\lceil \alpha l\rceil} \leqslant e^{-1}$, donc en replaçant cela dans notre expression

$$\geqslant \left(K\frac{1-e^{-3/2}}{3/2}\right)^{l}\alpha^{l/2}\frac{j^{3l/2}}{l!}.$$

Ainsi on a bien uniformément pour tout réel $\alpha \in]0,1]$ et pour tout $j \geqslant \alpha l \geqslant 0$

$$r_{l,j} \geqslant c^l \alpha^{l/2}\frac{j^{3l/2}}{l!},$$

où l'on peut choisir $c = K\frac{1-e^{-3/2}}{3/2}$ comme constante. □

Puisque l'estimation de la proposition 2.1.7,il peut sembler plus pratique de disposer d'une estimation sans faire apparaître de paramètre α. Ainsi en choisissant $\alpha = \min(1/2, j/l) \geqslant j/(2j+l)$, on obtient le résultat suivant.

Corollaire 2.1.8. — *Pour une constante absolue $c > 0$, on a pour tout $j \geqslant 1$ et pour tout $l \geqslant 0$*

$$r_{l,j} \geqslant c^l \frac{j^{2l}}{(l+2j)^{l/2}l!}.$$

Cependant, ce résultat n'est intéressant que si j est assez grand devant l. Par exemple, si l'on choisit $j = 1$, le corollaire 2.1.8 nous donne $r_{l,1} \gg c^l l^{-3l/2}$ alors qu'il est clair que $r_{l,1} = \frac{1}{\Gamma(l+3)} \gg c^l l^{-l/2}$. Ce corollaire est plus esthétique que la proposition 2.1.7, mais c'est purement cosmétique, puisqu'il ne fournit aucun bon comportement asymptotique pour l grand devant j.

L'estimation du corollaire 2.1.8 peut se traduire directement en terme de polynômes de Romanovsky, à partir de la relation (2.3).

Corollaire 2.1.9. — *Pour une constante absolue $c > 0$, on a pour tout $j \geqslant 1$ et pour tout $k \geqslant 2j$*

$$R_{k,j}(0) \geqslant c^k k^j \frac{j^{2(k-2j)}}{(k-2j)!}.$$

Là encore, le résultat n'est intéressant que lorsque k et j sont du même ordre de grandeur.

2.2. Moments centrés d'une loi binomiale

On cherche à modéliser de façon probabiliste une suite à croissance périodique **a** de densité P. Il est naturel de considérer q variables aléatoires indépendantes X_1, …, X_q de même loi de Bernoulli de paramètre P, représentant les valeurs de e_1, …, e_q ; l'évènement $X_n = 1$ doit donc être interprété comme la réalisation de la condition suivante : il existe un indice i tel que $a_i = n$. Les indices seront considérés modulo q, *i.e.* $X_{n+q} = X_n$ pour tout $n \in \mathbb{Z}$. Dans ce cadre, le moment centré d'ordre k d'une suite **a** « moyenne » correspond à l'espérance

$$\mathbb{E}\Big(\sum_{n=1}^{q}\big(\sum_{i=1}^{h} X_{n+i} - hP\big)^k\Big) = \sum_{n=1}^{q}\mathbb{E}\Big(\big(\sum_{i=1}^{h} X_{n+i} - hP\big)^k\Big) = q\mu_k(h,P), \tag{2.4}$$

où $\mu_k(h,P)$ est le moment centré d'ordre k de la loi binomiale de paramètre (h,P). La loi binomiale de paramètre (h,P) est obtenue en sommant h variables indépendantes de même loi de Bernoulli de paramètre P.

Il faut noter ici que les hypothèses faites sur les X_n sont bien plus fortes que ce qui est nécessaire pour établir l'identité (2.4). En effet, il n'est pas nécessaire que toutes les variables X_n soient indépendantes deux à deux : l'indépendance est seulement utilisée pour obtenir une loi binomiale. Cela signifie que d'un modèle d'indépendance « globale » des q variables X_i, on peut se restreindre à un modèle d'indépendance « locale » : celle des h variables X_{n+1}, …, X_{n+h}, et ce pour tout n.

Remarque. — Cette indépendance « locale » est bien pratique d'un point de vue statistique. En effet, une fois donnée une suite **a**, difficile de dire à quel point cette suite est la réalisation de q variables indépendantes ; le fait de se restreindre à l'indépendance de h variables permet une étude statistique sur la population de taille q des données $(e_{n+1}, \dots, e_{n+h})$, pour n allant de 1 à q. On peut donc calculer les coefficients de corrélation pour ces h variables, et prédire si le moment $M_k(h;\mathbf{a})$ est susceptible de se comporter ou non comme $q\mu_k(h,P)$.

On rappelle qu'une variable aléatoire X suit la loi binomiale de paramètre (h,P) si l'on a pour tout entier j la probabilité

$$\mathbb{P}(X=j) = \binom{h}{j} P^j (1-P)^{h-j}.$$

Le moment centré d'ordre k d'une loi binomiale est un sujet d'étude courant. On a les expressions suivantes

$$\begin{aligned}
\mu_k(h,P) &= \mathbb{E}\big(X - \mathbb{E}(X)\big)^k \\
&= \sum_{j=0}^{h} (j-hP)^k \binom{h}{j} P^j (1-P)^{h-j} \\
&= \sum_{s} \binom{k}{s} (-hP)^{k-s} \sum_{t} t! \binom{h}{t} \left\{ {s \atop t} \right\} P^t. \qquad (2.5)
\end{aligned}$$

La dernière identité provient du lemme suivant.

Lemme 2.2.1. — *Pour tous les entiers naturels h et k, on a l'identité polynomiale suivante*

$$\sum_{j=0}^{h}(j-hX)^k\binom{h}{j}Y^j(1-Y)^{h-j}=\sum_s\binom{k}{s}(-hX)^{k-s}\sum_t t!\binom{h}{t}\left\{{s\atop t}\right\}Y^t.$$

Démonstration. — On transforme notre expression en utilisant la formule de Newton

$$\sum_{j=0}^{h}(j-hX)^k\binom{h}{j}Y^j(1-Y)^{h-j}$$
$$=\sum_{j=0}^{h}(j-hX)^k\binom{h}{j}\sum_{t=0}^{h-j}(-1)^t\binom{h-j}{t}Y^{t+j}$$

puis en translatant de j la variable t

$$=\sum_{j=0}^{h}\sum_{t=j}^{h}(j-hX)^k(-1)^t(-1)^j\binom{h}{j}\binom{h-j}{t-j}y^t$$

enfin en interpolant les binomiales $\binom{h}{t}\binom{t}{j}=\binom{h}{j}\binom{h-j}{t-j}$ et en permutant les sommes

$$=\sum_{t=0}^{h}(-1)^t\binom{h}{t}\left(\sum_{j=0}^{t}(-1)^j\binom{t}{j}(j-hX)^k\right)Y^t.$$

On examine la somme intérieure en utilisant de nouveau la formule de Newton

$$\sum_{j=0}^{t}(-1)^j\binom{t}{j}(j-hX)^k=\sum_{j=0}^{t}\sum_{s=0}^{k}(-1)^j\binom{t}{j}\binom{k}{s}j^s(-hX)^{k-s}$$

puis on traite la somme sur j grâce à l'identité (A.9)

$$=\sum_{s=0}^{k}\binom{k}{s}(-hX)^{k-s}(-1)^t t!\left\{{s\atop t}\right\}.$$

En replaçant ce terme dans l'expression précédente, on obtient bien l'identité attendue. □

La dernière des trois formes de $\mu_k(h,P)$ est particulièrement importante : elle permet de considérer ce moment comme un polynôme en les variables h et P. Elle permet également de démontrer les deux propositions suivantes.

Proposition 2.2.2 **(Romanovsky,[34])**. — *Pour tout $k\geqslant 1$, on a*

$$\mu_{k+1}(h,P)=khP(1-P)\mu_{k-1}(h,P)+P(1-P)\frac{\partial\mu_k}{\partial P}(h,P).$$

Proposition 2.2.3. — *Pour tout $k \geqslant 0$, on a*

$$\mu_{2k}(h,P) = [k=0] + \sum_{i=1}^{k} \big(hP(1-P)\big)^i Q_i^{2k}\big(P(1-P)\big),$$
$$\mu_{2k+1}(h,P) = (1-2P)\sum_{i=1}^{k} \big(hP(1-P)\big)^i Q_i^{2k+1}\big(P(1-P)\big),$$

où les Q_i^k sont des polynômes de degré $\lfloor k/2 \rfloor - i$.

On préférera à ces deux propositions le corollaire suivant, qui justifie l'introduction des polynômes de Romanovsky, qui sont d'ailleurs plus simples à manipuler que les polynômes Q_i^k.

Corollaire 2.2.4. — *Pour tout $k \geqslant 0$, on a*

$$\mu_k(h,P) = \sum_{j\geqslant 0} \big(hP(1-P)\big)^j R_{k,j}(P),$$

où les polynômes $R_{k,j}$ sont les polynômes de Romanovsky.

Démonstration. — Grâce à la proposition 2.2.3, on sait qu'il existe des polynômes $\tilde{R}_{k,j}$ de $\mathbb{Z}[P]$ tels que la formule annoncée soit vraie. Il reste à montrer que ces polynômes sont bien les polynômes de Romanovsky. Puisque $\mu_0 = 1$, on constate que les polynômes $\tilde{R}_{0,j}$ sont bien égaux aux polynômes de Romanovsky $R_{0,j}$. La formule de Romanovsky de la proposition 2.2.2 permet de montrer par récurrence que pour $k \geqslant 1$, le polynôme $\mu_k(h,P)$ est un multiple de $hP(1-P)$, donc le polynôme $\tilde{R}_{k,0}$ est nul, tout comme l'est le polynôme de Romanovsky $R_{k,0}$. On a donc $\tilde{R}_{k,j} = R_{k,j}$ si $j = 0$ ou si $k = 0$. De plus, la relation de la proposition 2.2.2 traduite en terme des polynômes $\tilde{R}_{k,j}$ montre que les polynômes $\tilde{R}_{k,j}$ vérifie la même formule de récurrence que les polynômes de Romanovsky. On a donc bien $\tilde{R}_{k,j} = R_{k,j}$ pour tous les entiers j et k. □

On peut à présent utiliser nos connaissances sur les polynômes de Romanovsky pour estimer $\mu_k(h,P)$.

Proposition 2.2.5. — *Pour tout $k \geqslant 0$, on a*

$$\mu_k(h,P) \ll c^k k^{k/2}\, hP(1-P)\Big(\max\big(k, hP(1-P)\big)\Big)^{\lfloor k/2\rfloor - 1},$$

où c est une constante absolue qui peut être choisie strictement supérieure à $\frac{4}{e}$.

Démonstration. — Puisque P est une probabilité, on a $P \in [0,1]$ et donc $\big|R_{k,j}(P)\big| \leqslant \big\|R_{k,j}\big\|_1$. Grâce aux propositions 2.1.4 et 2.2.3 et à la majoration (2.2), on peut écrire

$$\begin{aligned}
\mu_k(h,P) &\leqslant \sum_{j=1}^{\lfloor k/2\rfloor} \big(hP(1-P)\big)^j T_j^k \\
&\leqslant \sum_{j=1}^{\lfloor k/2\rfloor} \big(hP(1-P)\big)^j \big(\tfrac{4}{e}(k-j)\big)^{k-j} \\
&\leqslant \big(\tfrac{4}{e}\big)^k k^k \sum_{j=1}^{\lfloor k/2\rfloor} \big(hP(1-P)/k\big)^j \\
&\leqslant k^2 \big(\tfrac{4}{e}\big)^k k^{k/2}\, hP(1-P) \Big(\max\big(k, hP(1-P)\big)\Big)^{\lfloor k/2\rfloor - 1}.
\end{aligned}$$

□

On peut être étonné de ce changement de comportement asymptotique qui a lieu pour $hP(1-P) \approx k$. On peut cependant obtenir des minorations qui font état de cette déviation.

Proposition 2.2.6. — *On a uniformément pour $hP(1-P)/k^3 \to \infty$*

$$\mu_{2k}(h,P) \sim (2k)!!\big(hP(1-P)\big)^k.$$

Démonstration. — On a égalité si $k = 0$, on suppose donc $k \geqslant 1$. En utilisant la propriété c) du lemme 2.1.5 et l'identité du corollaire 2.2.4, on a

$$\frac{\mu_{2k}(h,P)}{(2k)!!\big(hP(1-P)\big)^k} - 1 = \frac{1}{(2k)!!} \sum_{j=1}^{k-1} \frac{R_{2k,j}(P)}{\big(hP(1-P)\big)^{k-j}}.$$

Or la formule de Stirling permet d'établir $(2k)!! = \frac{(2k)!}{2^k k!} \sim \sqrt{2}\frac{(2k)^k}{e^k}$. Ainsi, on obtient grâce à cette remarque et à la proposition 2.1.4

$$\begin{aligned}
\frac{\mu_{2k}(h,P)}{(2k)!!\big(hP(1-P)\big)^k} - 1 &\ll \frac{e^k}{(2k)^k} \sum_{j=1}^{k-1} \frac{(4k-2j)^{2k-j}}{e^{2k-j}\big(hP(1-P)\big)^{k-j}} \binom{2k-j-1}{j-1} \\
&= \sum_{j=1}^{k-1} \Big(1 + \frac{k-j}{k}\Big)^k \Big(\frac{4k-2j}{e\,hP(1-P)}\Big)^{k-j} \binom{2k-j-1}{j-1}
\end{aligned}$$

or $1 + \frac{k-j}{k} \leqslant e^{(k-j)/k}$, donc

$$\leqslant \sum_{j=1}^{k-1} \Big(\frac{4k-2j}{hP(1-P)}\Big)^{k-j} \binom{2k-j-1}{j-1}$$

en changeant de variable

$$= \sum_{j=1}^{k-1} \Big(\frac{2k+2j}{hP(1-P)}\Big)^{j} \binom{k+j-1}{k-j-1}$$

or $\binom{k+j-1}{k-j-1} = \binom{k+j-1}{2j} \leqslant \frac{(2k)^{2j}}{(2j)!}$

$$\leqslant \sum_{j=1}^{k-1} \frac{1}{(2j)!} \Big(\frac{16k^3}{hP(1-P)} \Big)^j,$$

ce qui est bien $= o(1)$ si $k^{-3} hP(1-P) \to +\infty$. □

Ainsi la majoration de la proposition 2.2.5 ne peut pas être améliorée (à la valeur de la constante de magnitude près) pour $hP(1-P) \gg k^3$, puisque cela contredirait le résultat de la proposition 2.2.6 précédente. Il serait intéressant de pouvoir affaiblir cette condition jusqu'à $hP(1-P) \gg k$, qui serait la limite naturelle ; la proposition 2.2.7 suivante permet cependant d'établir un résultat partiel dans ce sens. De plus, nous verrons que la majoration de la proposition 2.2.5 est aussi essentiellement optimale pour $hP(1-P) \ll k$, sous peine de contredire la proposition 2.2.8 *infra.*

Plus expressément, on peut montrer que de toute majoration uniforme

$$(\star) \qquad \mu_k(h,P) \ll c^k k^{\alpha_1 k} \big(hP(1-P) \big)^{\alpha_2 k} \big(k + hP(1-P) \big)^{\alpha_3 k},$$

le meilleur choix du triplet $(\alpha_1, \alpha_2, \alpha_3)$ est $(1/2, 0, 1/2)$. Pour réaliser cela, nous allons raisonner à variance constante. On fixe donc un réel $\bar{\mu}_2 \geqslant 0$, et on considère l'ensemble des couples $(h,P) \in \mathbb{R}_+ \times [0,1]$ qui vérifie $hP(1-P) = \bar{\mu}_2$. On note alors $\bar{\mu}_k$ la limite de $\mu_k(h,P)$ pour P tendant vers 0 et pour (h,P) dans l'ensemble sus-cité. Les valeurs de $\bar{\mu}_k$ dépendent donc implicitement de $\bar{\mu}_2$, et les deux définitions de $\bar{\mu}_2$ sont cohérentes entre elles. On a en passant à la limite dans l'expression du corollaire 2.2.4 la formule suivante

$$\bar{\mu}_k = \sum_j R_{k,j}(0) \bar{\mu}_2^j = \mathbb{\Gamma}(k+1) \sum_j \bar{\mu}_2^j \, r_{k-2j,j}, \tag{2.6}$$

puisque les termes $r_{l,j}$ sont définies par (2.3).

Proposition 2.2.7. — *On a uniformément pour $k \geqslant 2$*

$$\bar{\mu}_k \gg c^k \mathbb{\Gamma}(k+1) \, \bar{\mu}_2^{\lfloor k/2 \rfloor},$$

où la constante c est absolue.

Démonstration. — Par le corollaire 2.1.8, on peut écrire pour une constante absolue $c > 0$

$$\bar{\mu}_k \geqslant c^k \mathbb{\Gamma}(k+1) \sum_{2j+l=k} \bar{\mu}_2^j \frac{j^{2l}}{k^{l/2} l!}.$$

On minore la somme par un seul de ses termes que l'on choisit par $j_0 = \lfloor k/2 \rfloor \geqslant 1$ et $l_0 = 2\{k/2\} \leqslant 1$. Ainsi on a $l_0! = 1$ et ${j_0}^2/\sqrt{k} \gg k^{3/2}$, ainsi on a bien

$$\bar{\mu}_k \gg k^{3/2} c^k \mathbb{\Gamma}(k+1) \, \bar{\mu}_2^{\lfloor k/2 \rfloor}.$$ □

Remarque. — La proposition précédente prouve que, dans le cas où $2 \leqslant k \ll hP(1-P)$, il est impossible de trouver une meilleure majoration en termes de k et $hP(1-P)$ que celle fournie par la proposition 2.2.5, à savoir

$$\mu_k(h,P) \ll c^k \mathbb{\Gamma}(k+1) \big(hP(1-P) \big)^{\lfloor k/2 \rfloor},$$

pour une constante absolue $c > 0$. Malheureusement, nous n'avons pas prouvé que

$$\mu_k(h,P) \gg c^k \mathbb{\Gamma}(k+1)\big(hP(1-P)\big)^{\lfloor k/2 \rfloor},$$

pour $k \ll hP(1-P)$, et où $c > 0$ est une constante absolue éventuellement différente de la constante précédente. Cependant, cette proposition montre que les triplets admissibles dans l'équation $(\star)$ doivent vérifier les inégalités $\alpha_1+\alpha_2+\alpha_3 \geqslant 1$, $\alpha_1+\alpha_3 \geqslant 1/2$ et $\alpha_2+\alpha_3 \geqslant 1/2$.

La proposition suivante permet de prouver que l'on doit avoir $\alpha_1+\alpha_3 \geqslant 1-\varepsilon$, pour tout $\varepsilon > 0$ assez petit. Ainsi le meilleur choix possible est bien $(1/2, 0, 1/2)$, qui est celui que fournit la proposition 2.2.5. On a bien un changement de comportement asymptotique, notamment lorsque P est petit, et ce changement se fait autour de l'égalité $k = hP(1-P)$. Nous n'avons malheureusement pas d'explications heuristiques de ce phénomène.

Proposition 2.2.8. — *On a uniformément pour $k \geqslant 0$, pour $\bar{\mu}_2 > 0$ et pour $\varepsilon \in]0, 1/3]$*

$$\bar{\mu}_k \gg c^k \varepsilon^{2k} k^k (\bar{\mu}_2/k)^{\varepsilon k},$$

où la constante c est absolue.

Démonstration. — On pose $\alpha = \varepsilon/(1-2\varepsilon) \in]0,1]$. On rappelle que l'on note $r_{l,j} = R_{2j+l,j}(0)/\mathbb{\Gamma}(2j+l+1)$, si bien que

$$\bar{\mu}_k = \mathbb{\Gamma}(k+1) \sum_j \bar{\mu}_2^j \, r_{k-2j,j}.$$

On a par la proposition 2.1.7 la minoration uniforme pour $j \geqslant \alpha l \geqslant 0$

$$r_{l,j} \geqslant C^l \alpha^{l/2} \frac{j^{3l/2}}{l!},$$

où $C > 0$ est une constante absolue. Ainsi, on a

$$\bar{\mu}_k \geqslant \mathbb{\Gamma}(k+1) \sum_{\substack{2j+l=k \\ j \geqslant \alpha l}} C^l \alpha^{l/2} \bar{\mu}_2^j \frac{j^{3l/2}}{l!}.$$

Les conditions de sommation imposent $j \geqslant \frac{\alpha}{1+2\alpha} k = \varepsilon k$. On minore la somme par un seul de ses termes que l'on choisit par $j_0 = \lceil \varepsilon k \rceil \geqslant \varepsilon k$ et $l_0 = k - 2j_0$. On a l'encadrement

$$\frac{\varepsilon}{\alpha} k - 2 = k - 2(\varepsilon k + 1) \leqslant l_0 \leqslant k - 2(\varepsilon k) = \frac{\varepsilon}{\alpha} k.$$

On en déduit $l_0! \leqslant \big(\frac{\varepsilon}{\alpha} k\big)^{k\varepsilon/\alpha}$, ${j_0}^{3l_0/2} \geqslant k^{-3}\big((\varepsilon k)^{k\varepsilon/\alpha}\big)^{3/2}$, $\bar{\mu}_2^{j_0} \geqslant \bar{\mu}_2^{\varepsilon k}$ et $\alpha^{l_0/2} \geqslant (\alpha^{\varepsilon/\alpha})^{k/2}$. On obtient

$$\bar{\mu}_k \geqslant \mathbb{\Gamma}(k+1) C^{k\varepsilon/\alpha} k^{-3} (\varepsilon^{1/2}\alpha^{3/2})^{k\varepsilon/\alpha} \bar{\mu}_2^{\varepsilon k} \big(k^{k\varepsilon/\alpha}\big)^{1/2} \gg c^k (\varepsilon^{1/2}\alpha^{3/2})^{k\varepsilon/\alpha} k^k (\bar{\mu}_2/k)^{\varepsilon k},$$

puisque $\varepsilon/\alpha = 1-2\varepsilon$ et $\mathbb{\Gamma}(k+1) \gg (k/e)^{k/2}$, et pour une constante absolue $c < C/\sqrt{e}$. Enfin, en remarquant les minorations suivantes $\varepsilon^{1-2\varepsilon} \geqslant \varepsilon$ et $\alpha^{\varepsilon/\alpha} \geqslant \varepsilon$, on obtient la relation souhaitée. □

En choisissant $\varepsilon = \frac{\log 2}{3\log(k/\bar{\mu}_2)}$ dans la proposition 2.2.8, on obtient le corollaire suivant.

Corollaire 2.2.9. — *On a uniformément pour* $0 < \bar{\mu}_2 \leqslant k/2$

$$\bar{\mu}_k \gg \Big(\frac{c}{\log(k/\bar{\mu}_2)}\Big)^{2k} k^k,$$

où la constante c est absolue.

On voit bien que la majoration obtenue dans ce cas par la proposition 2.2.5, à savoir $\bar{\mu}_k \ll c^k k^k$, est presque optimale. Il faut cependant remarquer que lorsque $\bar{\mu}_2$ est vraiment très petit, la minoration obtenue par le corollaire 2.2.9 est moins bonne que la minoration triviale $\bar{\mu}_k \geqslant \bar{\mu}_2$ obtenue en ne considérant que le premier terme de la somme de l'identité (2.6), c'est-à-dire $R_{k,1}(0)\bar{\mu}_2$ et en remarquant que $R_{k,1}(0) = 1$ (*cf.* lemme 2.1.5d).

CHAPITRE 3

LE CAS DES ENTIERS SANS PETITS FACTEURS PREMIERS

Ce chapitre est entièrement consacré à la preuve de la relation

$$M_k(h;q) \ll c^k q\, k^{k/2}\, (k+hP)^{k/2},$$

uniformément pour k, h et q sous la condition que tous les diviseurs premiers de q soient plus grands que h, et où la constante c est absolue. Ce résultat est à comparer avec les résultats du chapitre précédent pour notre modèle probabiliste.

Pour parvenir à ce résultat, nous remarquons dans la proposition 3.1.2 que sous la condition $P^-(q) \geqslant h$, le moment $\frac{1}{q}M_k(h;q)$ possède une expression comparable à celle (2.5) de $\mu_k(h,P)$, mais où certaines puissances $P^j = \prod_{p|q}(1-1/p)^j$ ont été remplacées par leurs analogues « linéarisés » $\prod_{p|q}(1-j/p)$, que nous noterons $\mathbb{E}_{\mathrm{stat}}(Y^h)$, *cf.* la définition (3.1). Nous introduisons à cet effet par la formule (3.2) un nouveau polynôme $m_k(h,X,Y)$, variante de l'expression polynomiale (2.5) de $\mu_k(h,P)$ où l'on a distingué les puissances à « linéariser » des autres. Nous montrons dans la proposition 3.1.3 que ce polynôme $m_k(h,X,Y)$ s'écrit comme une convolution des polynômes $\mu_k(h,Y)$ et $(X-Y)^k$; on obtient le résultat souhaité en estimant après « linéarisation » (concrétisée par l'opérateur $\mathbb{E}_{\mathrm{stat}}$) les quantités $\mu_k(h,Y)$ et $(X-Y)^k$. Si l'estimation de la première est rendue facile par l'étude des polynômes de Romanovsky réalisée au chapitre précédent, l'estimation de la seconde énoncée dans le corollaire 3.1.7 nécessite une analyse détaillée qui fera l'objet de la deuxième section de ce chapitre.

Au cours de cette deuxième section, nous étudions en réalité le problème plus général d'évaluer des sommes du type

$$(*) \qquad \sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}\prod_{i=1}^{n}\frac{1-sx_i}{(1-x_i)^s}$$

avec des complexes x_i assez petits en module devant $1/t$. Dans les applications, ces x_i seront les $1/p$, pour p facteur premier de q. On sait que chacun des produits présents dans la somme est proche de 1, avec un terme d'erreur parfaitement connu. La somme des termes principaux est donc nulle, mais en première approximation, la somme des termes d'erreur est assez grande, alors que nous prouvons que les annulations sont

nombreuses parmi ces termes d'erreur. Pour en arriver à cette conclusion, nous transformons dans la proposition 3.2.5 la somme (*) en polynôme en les variables $x_i/(1-x_i)$ et faisons apparaître parmi les coefficents de ce polynôme une quantité à l'aspect compliqué introduite par la définition (3.5)

$$\omega(\mathbf{s}, m) := \frac{\prod_{i=1}^{n} s_i!}{m!} \sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} \prod_{i=1}^{n} \binom{k}{s_i},$$

mais dont la majoration va conduire en deux temps, au travers des lemmes 3.2.10 et 3.2.11, à notre estimation principale de la somme (*) par $c^t(t\|\mathbf{x}\|_\infty + t\|\mathbf{x}\|_2^2)^{t/2}$ pour une constante absolue $c > 0$, comme énoncé dans le théorème 3.1.4.

La majoration de la quantité $\omega(\mathbf{s}, m)$ est l'objet de la troisième et dernière section de ce chapitre. Celle-ci est obtenue grâce à une interprétation énumérative de $\omega(\mathbf{s}, m)$ en termes des partitions d'ensemble.

Ce chapitre est ainsi découpée en trois sections gigognes, la deuxième section détaillant une boîte noire utilisée au cours de la première, et la troisième pareillement vis-à-vis de la deuxième. La lecture de chacune de ces parties peut cependant se faire indépendamment ou dans un ordre différent de celui que nous proposons.

3.1. Répartition des entiers premiers relativement à un entier sans petit facteur premier

On concentre à présent notre intérêt sur une famille de suites particulières, celle des suites exhaustives des entiers premiers relativement à q, pour q variant parmi les entiers sans facteur carré. Pour chaque entier q, nous définissons donc la suite $1 = a_1 < a_2 < \ldots, a_{\varphi(q)} \leqslant q$, où chaque a_i vérifie $(a_i, q) = 1$, et on étend le domaine de définition de cette suite par la relation $a_{\varphi(q)+i} = q + a_i$. Cette suite ne dépendant que de l'ensemble des facteurs premiers de q, l'entier q sera supposé sans facteur carré. Comme chaque suite ainsi créée est caractérisée par l'entier q dont elle dépend, on notera la dépendance d'une grandeur à la suite par une dépendance en q : par exemple, nous utiliserons les notations $M_k(h; q)$ et $V_\gamma(q)$. Il est à remarquer que, avec la formalisme du chapitre 1, l'on a $r = \varphi(q)$ et $P = \prod_{p|q}(1 - 1/p)$.

Nous avons déjà remarqué qu'il existe un critère heuristique pour se persuader qu'une suite risque d'adopter un comportement moyen, au moins dans l'évaluation de son moment centré d'ordre k. Il s'agit en effet de vérifier que, statistiquement, la suite **a** possède la propriété d'indépendance « locale ». La question est de savoir si l'on peut considérer la population de taille q des données $(e_{n+1}, \ldots, e_{n+h})$ comme des réalisations de la variable aléatoire $(Y_1, \ldots, Y_h)$ où les Y_i sont des variables aléatoires indépendantes suivant la même loi de Bernoulli de paramètre P ; cela revient à se demander si les covariances statistiques sont faibles.

$$\begin{aligned} \mathrm{Cov}_{\mathrm{stat}}(Y_i, Y_j) &= \frac{\mathrm{Var}_{\mathrm{stat}}(Y_i Y_j)}{\sqrt{\mathrm{Var}_{\mathrm{stat}}(Y_i)\mathrm{Var}_{\mathrm{stat}}(Y_j)}} \\ &= \frac{1}{P(1-P)} \left(\frac{1}{q} \sum_{n=1}^{q} e_{n+i} e_{n+j} - P^2 \right) \end{aligned}$$

Évaluer cette somme est un simple exercice d'arithmétique :

$$\begin{aligned}\mathrm{Cov}_{\mathrm{stat}}(Y_i,Y_j) &= \frac{1}{P(1-P)}\Big(\prod_{\substack{p|q\\ p|i-j}}\Big(1-\frac{1}{p}\Big)\prod_{\substack{p|q\\ p\nmid i-j}}\Big(1-\frac{2}{p}\Big)-\prod_{p|q}\Big(1-\frac{1}{p}\Big)^2\Big)\\ &= \frac{P}{1-P}\Big(\prod_{\substack{p|q\\ p|i-j}}\Big(1+\frac{1}{p-1}\Big)\prod_{\substack{p|q\\ p\nmid i-j}}\Big(1-\frac{1}{(p-1)^2}\Big)-1\Big)\end{aligned}$$

Le premier des deux produits grandit relativement vite avec $(q, i-j)$, et les petits facteurs de q qui ne diviseraient pas $i-j$ empêcheraient cependant le second produit de s'approcher de 1. Le cas idéal est donc celui d'un q sans facteur premier inférieur à h pour éviter tout phénomène de résonance arithmétique. Dans ce cas, on a pour tout $1 \leqslant i < j \leqslant h \leqslant P^-(q)$

$$\mathrm{Cov}_{\mathrm{stat}}(Y_i,Y_j) \ll \tfrac{P}{1-P}\big(P^-(q)\log P^-(q)\big)^{-1}.^*$$

Le modèle heuristique nous permet d'espérer obtenir $M_k(h,q) \approx q\mu_k(h,P)$ pour des q sans facteur premier inférieur à h.

Cette indépendance statistique est d'ailleurs tout aussi bien vérifiée si l'on généralise au cas d'un ensemble d'entiers I de cardinal h. Pour peu qu'on ait $I \subset [\![1,l]\!]$ avec l assez petit devant $P^-(q)$, la condition d'indépendance « locale » est statistiquement vérifiée : on peut donc espérer

$$M_k(I;q) = \sum_{n=1}^{q}\big(e_n(I)-hP\big)^k \approx q\mu_k(h,P),$$

sous les conditions indiquées sur I et q.

3.1.1. Géométrie de la fenêtre. — On commence par montrer que tant que $I \subset [\![1,P^-(q)]\!]$, la valeur du moment $M_k(I;q)$ ne dépend que du cardinal h de I, et non de la répartition des éléments de l'ensemble I. Si cette propriété n'aide pas à établir le comportement asymptotique de $M_k(h;q)$, elle est nécessaire pour utiliser la proposition 1.3.4 au moment voulu, et sa démonstration permet d'introduire des outils qui nous serviront par la suite : la fonction d'espérance statistique $\mathbb{E}_{\mathrm{stat}}$ et le polynôme $m_k(h,X,Y)$.

Lemme 3.1.1. — *Soient q un entier naturel et $I \subset [\![1,P^-(q)]\!]$. On a*

$$\frac{1}{q}F_I^*(I) = \prod_{p|q}\Big(1-\frac{\operatorname{card} I}{p}\Big).$$

*On peut évidemment obtenir une borne plus précise mais cette formule est essentiellement la meilleure en termes de $P^-(q)$ uniquement.

Démonstration. — La quantité $F_I^*(I)$, définie par (1.1), compte le nombre de translations $n+I$ de l'ensemble I, qui soient telles que chaque entier $n+i$ de $n+I$ soit premier à q, autrement dit

$$F_I^*(I) = \sum_{n=1}^{q} \prod_{i\in I} [(n+i,q)=1].$$

Or la fonction qui à l'entier q associe le produit $\prod_{i\in I}[(n+i,q)=1]$ est clairement une fonction multiplicative ; on a donc

$$F_I^*(I) = \prod_{p|q} \Big(\sum_{n=1}^{p} \prod_{i\in I} [(n+i,p)=1] \Big).$$

Puisque la longueur de I est inférieur à p, et ce pour chaque p divisant q, les éléments i de I sont tous distincts modulo p ; le terme du produit correspondant à p vaut donc bien $p - \operatorname{card} I$. □

En quelque sorte, cela établit que l'espérance statistique de $\prod_{i\in I} Y_i$ ne dépend que du cardinal de l'ensemble I, pourvu que la longueur de celui-ci ne soit plus grande qu'aucun diviseur premier de q. Ainsi on définit une forme linéaire (dépendant de q) sur l'espace des polynômes de degré au plus $P^-(q)$ par

$$\mathbb{E}_{\mathrm{stat}}(Y^h) \coloneqq \prod_{p|q} \Big(1 - \frac{h}{p}\Big), \tag{3.1}$$

pour tout $h \leqslant P^-(q)$ et l'on a, quel que soit l'ensemble I vérifiant $\operatorname{card} I = h$ et $I \subset [\![1, P^-(q)]\!]$

$$\mathbb{E}_{\mathrm{stat}}(Y^h) = \mathbb{E}_{\mathrm{stat}}\Big(\prod_{i\in I} Y_i\Big).$$

Proposition 3.1.2. — *Soient q un entier naturel et $I \subset [\![1, P^-(q)]\!]$. On pose $h = \operatorname{card} I$. On a*

$$M_k(I;q) = q\, \mathbb{E}_{\mathrm{stat}}\Big(\sum_j (j-hP)^k \binom{h}{j} Y^j (1-Y)^{h-j} \Big).$$

En particulier, on a $M_k(I;q) = M_k(h;q)$.

En d'autres termes cette proposition indique que la valeur du moment centré ne dépend pas de la géométrie de la fenêtre. La fonction intervenant dans cette proposition nous sera d'une grande utilité par la suite ; nous définissons donc ce polynôme en X et en Y pour $h, k \in \mathbb{N}$ par

$$m_k(h,X,Y) \coloneqq \sum_j (j-hX)^k \binom{h}{j} Y^j (1-Y)^{h-j}. \tag{3.2}$$

Démonstration. — Soit $J \subset I$ de cardinal $j \leqslant h$. On a par l'équation (1.3)

$$F_J^*(I) = \sum_{J' \subset I \smallsetminus J} (-1)^{\operatorname{card} J'} F_{J\cup J'}^*(J \cup J')$$

et par le lemme 3.1.1 qui calcule les $F_I^*(I)$,

$$\begin{aligned}
&= \sum_{J' \subset I \smallsetminus J} (-1)^{\operatorname{card} J'} q\, \mathbb{E}_{\mathrm{stat}}\big(Y^{\operatorname{card} J \cup J'}\big) \\
&= q \sum_{j'} \binom{h-j}{j'} (-1)^{j'} \mathbb{E}_{\mathrm{stat}}\big(Y^{j+j'}\big) \\
&= q\, \mathbb{E}_{\mathrm{stat}}\big(Y^j (1-Y)^{h-j}\big).
\end{aligned}$$

Or on a par l'équation (1.2)

$$\begin{aligned}
F_j(I) &= \sum_{\substack{J \subset I \\ \operatorname{card} J = j}} F_J^*(I) \\
&= q\, \mathbb{E}_{\mathrm{stat}}\big(Y^j (1-Y)^{h-j}\big) \sum_{\substack{J \subset I \\ \operatorname{card} J = j}} 1 \\
&= q \binom{h}{j} \mathbb{E}_{\mathrm{stat}}\big(Y^j (1-Y)^{h-j}\big),
\end{aligned}$$

d'où

$$\begin{aligned}
M_k(I;q) &= \sum_{n=1}^{q} \big(E_n(I) - hP\big)^k \\
&= \sum_{j=0}^{h} (j - hP)^k F_j(I) \\
&= q \sum_{j=0}^{h} (j - hP)^k \binom{h}{j} \mathbb{E}_{\mathrm{stat}}\big(Y^j (1-Y)^{h-j}\big).
\end{aligned}$$

□

3.1.2. Équirépartition. — Le polynôme $m_k(h, X, Y)$ possède beaucoup de points communs avec le moment centré de la loi binomiale : en effet, il est immédiat que

$$m_k(h, P, P) = \mu_k(h, P). \tag{3.3}$$

D'ailleurs, lui aussi peut s'écrire sous forme d'un polynôme de $\mathbb{Z}[h, X, Y]$ grâce au lemme 2.2.1

$$m_k(h, X, Y) = \sum_s \binom{k}{s} (-hX)^{k-s} \sum_t t! \binom{h}{t} \left\{ {s \atop t} \right\} Y^t.$$

Ainsi, la symétrie naturelle $m_k(h, 1-X, 1-Y) = (-1)^k m_k(h, X, Y)$ est une identité polynomiale. On peut également établir une version du résultat de Romanovsky [**34**], mais nous préférons établir une formule de convolution qui lie les polynômes m_k aux polynômes μ_k.

Proposition 3.1.3. — *Soit k un entier naturel. On a l'identité polynomiale suivante*

$$m_k(h, X, Y) = \sum_m \binom{k}{m} \big(h(Y - X)\big)^m \mu_{k-m}(h, Y).$$

Démonstration. — On a pour $h \in \mathbb{N}$

$$\begin{aligned} m_k(h,X,Y) &= \sum_j (j-hX)^k \binom{h}{j} Y^j (1-Y)^{h-j} \\ &= \sum_j \sum_m \binom{k}{m} (j-hY)^{k-m} (hY-hX)^m \binom{h}{j} Y^j (1-Y)^{h-j} \\ &= \sum_m \binom{k}{m} \big(h(Y-X)\big)^m m_{k-m}(h,Y,Y), \end{aligned}$$

et $m_{k-m}(h,Y,Y) = \mu_{k-m}(h,Y)$ par (3.3). Comme les deux termes de l'identité sont des éléments de $\mathbb{Z}[h,X,Y]$, l'identité a un sens en tant qu'identité polynomiale. □

On souhaite exploiter cette identité en profitant de la bonne connaissance des termes $\mu_k(h,P)$ acquise au chapitre 2, notamment au travers des polynômes de Romanovsky. Pour se faire, il nous faut également étudier les termes correspondant aux facteurs $\big(h(Y-X)\big)^m$ de l'identité de la proposition 3.1.3. Le cœur de notre raisonnement est l'évaluation du moment statistique centré $\mathbb{E}_{\mathrm{stat}}\big((Y-P)^m\big)$, du moins lorsque m est petit par rapport à $P^-(q)$. Pour établir un tel résultat, nous allons utiliser le théorème suivant, qui s'énonce dans un cadre plus général et dont la démonstration est renvoyée à la section suivante.

Théorème 3.1.4. — *Soient $\varepsilon \in]0,1/2[$ et $K>0$ deux réels. Pour tout vecteur complexe $\mathbf{x} = (x_1,\dots,x_n) \in \mathbb{C}^n$, on pose*

$$N(\mathbf{x}) \coloneqq \max\Big(|x_1|,\dots,|x_n|, \sum_{i=1}^n |x_i|^2\Big).$$

Uniformément pour tout vecteur complexe $\mathbf{x} = (x_1,\dots,x_n) \in \mathbb{C}^n$ vérifiant $|1-x_i| \geqslant \varepsilon$ pour tout i et pour tout entier naturel t vérifiant $tN(\mathbf{x}) \leqslant K$ on a

$$\sum_{s=0}^t (-1)^{t-s} \binom{t}{s} \prod_{i=1}^n \frac{1-sx_i}{(1-x_i)^s} \ll (c_{K,\varepsilon})^t \big(tN(\mathbf{x})\big)^{\lceil t/2 \rceil},$$

où $c_{K,\varepsilon}$ est une constante dépendant uniquement de K et de ε.

Nous n'utilisons de ce résultat que la version *ad hoc* suivante.

Corollaire 3.1.5. — *Uniformément pour tout entier naturel m et pour tout entier q sans facteur carré vérifiant $m \leqslant P^-(q)$, on a*

$$\mathbb{E}_{\mathrm{stat}}\big((Y/P-1)^m\big) \ll c^m \frac{m^{m/2}}{\big(P^-(q)\big)^{m/2}},$$

où la constante c est absolue.

Démonstration. — Par la formule du binôme et par la définition (3.1), il est clair que l'expression à majorer est

$$\mathbb{E}_{\mathrm{stat}}\big((Y/P-1)^m\big)=\sum_{s=0}^{m}(-1)^{m-s}\binom{m}{s}P^{-s}\mathbb{E}_{\mathrm{stat}}(Y^s)$$
$$=\sum_{s=0}^{m}(-1)^{m-s}\binom{m}{s}\prod_{p\mid q}\frac{1-s/p}{(1-1/p)^s}.$$

Cette expression est une spécialisation au vecteur $\mathbf{x}=(1/p_1,\dots,1/p_n)$ du moment considéré par le théorème 3.1.4, où p_1, p_2, ..., p_n sont les facteurs premiers de q ordonnés de façon croissante, *i.e.* $q=p_1p_2\cdots p_n$ et $p_1<p_2<\ldots<p_n$. On a clairement $\|\mathbf{x}\|^2=\sum_i 1/{p_i}^2\ll 1/{p_1}^{\dagger}$ et donc $N(\mathbf{X})\asymp 1/p_1$. Par le théorème 3.1.4, on a sous la condition $m\leqslant P^-(q)$ et $\mu(q)\neq 0$

$$\sum_{s=0}^{m}(-1)^{m-s}\binom{m}{s}\prod_{p\mid q}\frac{1-s/p}{(1-1/p)^s}\ll c^m\left(\frac{m}{P^-(q)}\right)^{\lceil m/2\rceil},$$

où c est une constante absolue, ce qui fournit bien la majoration désirée. □

Corollaire 3.1.6. — *Uniformément pour tout couple d'entiers naturels (m,n) et pour tout entier q sans facteur carré vérifiant $m+n\leqslant P^-(q)$ on a*

$$\mathbb{E}_{\mathrm{stat}}\big((Y-P)^mY^n\big)\ll c^{m+n}P^{m+n}\frac{m^{m/2}}{\big(P^-(q)\big)^{m/2}},$$

où la constante c est absolue.

Démonstration. — On a

$$\mathbb{E}_{\mathrm{stat}}\big((Y-P)^mY^n\big)=P^{m+n}\sum_{k=0}^{n}\binom{n}{k}\mathbb{E}_{\mathrm{stat}}\big((Y/P-1)^{m+k}\big)$$

et grâce au corollaire 3.1.5, on sait qu'il existe une constante absolue C telle qu'on a

$$\ll P^{m+n}\sum_{k=0}^{n}\binom{n}{k}C^{m+k}\frac{(m+k)^{(m+k)/2}}{\big(P^-(q)\big)^{(m+k)/2}}$$
$$\ll P^{m+n}C^m\frac{m^{m/2}}{\big(P^-(q)\big)^{m/2}}\sum_{k=0}^{n}\binom{n}{k}C^k\Big(1+\frac{k}{m}\Big)^{m/2}\Big(\frac{m+k}{P^-(q)}\Big)^{k/2}$$

[†]On peut bien évident faire mieux (gagner un facteur $1/\log p_1$) mais dans la définition de $N(\mathbf{x})$ le terme en $1/p_1$ est de toute façon limitant. Cette amélioration peut avoir une utilité si l'on essaye de calculer la constante implicite, ce qui n'est pas notre intérêt principal.

or on a $m+k \leqslant m+n \leqslant P^-(q)$ et $(1+k/m)^m \leqslant e^k$, donc

$$\begin{aligned}&\ll P^{m+n}C^m(Ce^{1/2}+1)^n\frac{m^{m/2}}{\big(P^-(q)\big)^{m/2}}\\&\ll c^{m+n}P^{m+n}\frac{m^{m/2}}{\big(P^-(q)\big)^{m/2}},\end{aligned}$$

avec $c = Ce^{1/2}+1$ comme choix de constante. □

Corollaire 3.1.7. — *Uniformément pour tout couple d'entiers naturels (m,n), tout polynôme R de $\mathbb{R}[X]$ de degré d et pour tout entier q sans facteur carré vérifiant $m+n+d \leqslant P^-(q)$ on a*

$$\mathbb{E}_{\mathrm{stat}}\big((Y-P)^mY^nR(Y)\big) \ll c^{m+n+d}P^{m+n}\|R\|_1\frac{m^{m/2}}{\big(P^-(q)\big)^{m/2}},$$

où la constante c est absolue.

Démonstration. — En posant $R = \sum_{i=0}^d r_iX^i$, on a

$$\mathbb{E}_{\mathrm{stat}}\big((Y-P)^mY^nR(Y)\big) = \sum_{i=0}^d r_i\mathbb{E}_{\mathrm{stat}}\big((Y-P)^mY^{n+i}\big)$$

puisque $m+n+i \leqslant m+n+d \leqslant P^-(q)$, on sait par le corollaire 3.1.6 qu'il existe une constante absolue $C \geqslant 1$ telle qu'on a

$$\begin{aligned}&\ll \sum_{i=0}^d |r_i|C^{m+n+i}P^{m+n+i}\frac{m^{m/2}}{\big(P^-(q)\big)^{m/2}}\\&\ll C^{m+n+d}P^{m+n}\Big(\sum\nolimits_{i=0}^d |r_i|\Big)\frac{m^{m/2}}{\big(P^-(q)\big)^{m/2}}.\end{aligned}$$ □

Nous pouvons à présent établir le résultat annoncé au début de ce chapitre.

Théorème 3.1.8. — *Uniformément pour tout entier h et pour tout entier q sans facteur carré vérifiant $h \leqslant P^-(q)$ on a*

$$M_k(h;q) \ll^k qk^{k/2}\max(k,hP)^{k/2}.$$

Démonstration. — Si $h \leqslant k$, on a *a fortiori* $hP \leqslant k$ et donc

$$M_k(h;q) = \sum_{n=1}^q (e_n(h)-hP)^k \leqslant qh^k \leqslant qk^k = qk^{k/2}\max(k,hP)^{k/2}.$$

On peut donc se placer dans le cas où $k \leqslant h \leqslant P^-(q)$. On replace l'expression du corollaire 2.2.4 dans le résultat de la proposition 3.1.3 et donc

$$m_k(h,X,Y) = \sum_s \binom{k}{s}\big(h(Y-X)\big)^s\sum_j \big(h\,Y(1-Y)\big)^jR_{k-s,j}(Y),$$

où le polynôme $R_{k-s,j}$ est de degré inférieur à $k-s-2j$ (proposition 2.2.3) et vérifie $\|R_{k-s,j}\|_1 \leqslant (2k-2s-2j)!!\binom{k-s-j-1}{j-1}$ (proposition 2.1.4). Puisque $h \leqslant P^-(q)$, on a par la proposition 3.1.2

$$\begin{aligned} M_k(h;q) &= q\,\mathbb{E}_{\mathrm{stat}}\big(m_k(h,P,Y)\big) \\ &= q\sum_s\sum_j \binom{k}{s} h^{j+s}\mathbb{E}_{\mathrm{stat}}\big((Y-P)^s Y^j R_{k-s,j}(Y)\,(1-Y)^j\big). \end{aligned}$$

Puisque le degré en Y est inférieur à $s+j+(k-s-j)=k\leqslant P^-(q)$, on peut appliquer le corollaire 3.1.7 à notre situation. On note que

$$\|R_{k-s,j}(Y)\,(1-Y)^j\|_1 \leqslant 2^j(2k-2s-2j)!!\binom{k-s-j-1}{j-1} \ll^k (k-s-j)^{k-s-j},$$

ainsi on a

$$\mathbb{E}_{\mathrm{stat}}\big((Y-P)^s Y^j R_{k-s,j}(Y)\,(1-Y)^j\big) \ll^k P^{s+j}\frac{s^{s/2}}{\big(P^-(q)\big)^{s/2}}(k-s-j)^{k-s-j}.$$

On rappelle que les variables de sommation j et s vérifient $s+2j\leqslant k$. à présent, le moment peut être majoré

$$M_k(h;q) \ll^k q\sum_{s+2j\leqslant k}\frac{(hP)^{s+j}}{\big(P^-(q)\big)^{s/2}}s^{s/2}(k-s-j)^{k-s-j}$$

or $P^-(q)\geqslant h\geqslant hP$ et $s\leqslant k$, donc

$$\ll^k q\sum_{s+2j\leqslant k}(hP)^{s/2+j}k^{k-s/2-j}$$

et à n fixé, il y a au plus $k\ll^k 1$ couples (j,s) qui vérifient $s+2j=n$

$$\ll^k qk^k\sum_{n\leqslant k}\Big(\frac{hP}{k}\Big)^{n/2} \ll^k qk^k\,\max\Big(1,\frac{hP}{k}\Big)^{k/2},$$

ce qu'il fallait démontrer. □

3.2. Défaut de potentialité

Cette section est indépendante de la précédente, et est entièrement dédiée à la démonstration du théorème 3.1.4. En plus du soin apporté à la constante en k de l'estimation du théorème 3.1.8, ce résultat représente notre principal apport à la preuve de Montgomery et Vaughan. Avant d'en présenter la preuve, nous souhaitons également en souligner l'enjeu.

Lorsque le réel x est très proche de 0, on sait que $1-tx$ se comporte comme $(1-x)^t$, pour t entier positif ; mais il y a différentes façons de quantifier ce phénomène.

Lemme 3.2.1. — *Soient $t\in\mathbb{N}$ un entier et x un réel. On a les propriétés suivantes :*

a). *Si $x\leqslant 1$, on a $(1-x)^t\geqslant 1-tx$;*

b). *Pour $|x|\leqslant 1/t$, on a uniformément $(1-x)^t = 1-tx+O(t^2x^2)$;*

c). *On a*

$$\sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}(1-x)^{t-s}(1-sx) = (1-t)x^t.$$

Démonstration. — Pour la propriété *a),* un simple argument de convexité suffit. Pour la *b),* on peut développer

$$\big|(1-x)^t-(1-tx)\big| \leqslant \sum_{s=2}^{t}\binom{t}{s}|x|^s = t(t-1)x^2\sum_{s=0}^{t-2}\binom{t-2}{s}\frac{|x|^s}{(s+1)(s+2)},$$

ce qui est bien $\leqslant \frac{1}{2}t^2x^2(1+|x|)^{t-2} \leqslant \frac{e}{2}t^2x^2 \ll t^2x^2$.
Enfin, on a

$$\begin{aligned}\sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}(1-x)^{t-s}(1-sx) &= \sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}(1-x)^{t-s}\\ &\quad -\sum_{s=1}^{t}(-1)^{t-s}s\binom{t}{s}(1-x)^{t-s}x\\ &= x^t - tx\sum_{s=0}^{t-1}(-1)^{t-1-s}\binom{t-1}{s}(1-x)^{t-1-s}\\ &= (1-t)x^t,\end{aligned}$$

ce qui établit la propriété *c).* □

La propriété *b)* mesure le fait que les quantités $1-tx$ et $(1-x)^t$ sont proches lorsque x est assez petit, et le mesure additivement, en estimant leur différence. Il existe toutefois d'autres façons de mesurer la proximité de deux valeurs, ou plus exactement dans ce cas, deux ensembles de valeurs. La propriété *c)* en est un exemple : elle exploite le fait que la suite des $1-tx$ se comporte, pour des valeurs du paramètre t assez petites, comme la suite géométrique des $(1-x)^t$. Cette estimation est de nature plus multiplicative, et en devient plus efficace : si la propriété *b)* s'apparente à l'estimation du polynôme X^t-Y^t, avec X et Y très proche, la propriété *c)* est à relier au polynôme $(X-Y)^t$, qui s'annule avec une multiplicité supérieure en $X=Y$.

Notre objectif est de donner une version multidimensionnelle du lemme 3.2.1, c'est-à-dire savoir estimer

$$\prod_{i=1}^{n}(1-tx_i) - \Big(\prod_{i=1}^{n}(1-x_i)\Big)^t,$$

et *in fine* les quantités

$$\sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}\Big(\prod_{i=1}^{n}(1-x_i)\Big)^{t-s}\prod_{i=1}^{n}(1-sx_i). \tag{3.4}$$

On sait déjà que si $|x_i| \leqslant 1/t$, par le lemme 3.2.1, qu'il existe une constante absolue C telle que

$$1 \geqslant \frac{1-tx_i}{(1-x_i)^t} \geqslant 1-Ct^2{x_i}^2,$$

et donc si $\max_i |x_i| \leqslant 1/t$, on a

$$1 \geqslant \prod_{i=1}^{n} \frac{1-tx_i}{(1-x_i)^t} \geqslant \prod_{i=1}^{n} \big(1 - Ct^2 {x_i}^2\big) \geqslant 1 - Ct^2 \sum_{i=1}^{n} {x_i}^2.$$

Lemme 3.2.2. — *Soient $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ un vecteur et $t \geqslant 1$ un entier. Si $\|\mathbf{x}\|_\infty \leqslant 1/t$, on a pour une constante absolue C*

$$\prod_{i=1}^{n}(1-tx_i) - \Big(\prod_{i=1}^{n}(1-x_i)\Big)^t \ll C^n t^2 \|\mathbf{x}\|_2{}^2.$$

C'est ce type de majoration qu'utilise Montgomery et Vaughan pour estimer le moment $M_k(h;q)$ à partir des moments centrés de loi binomiale $\mu_k(h,P)$. La conséquence est que le domaine de validité de leur estimation asymptotique de $M_k(h;q)$ est sensiblement moins bon que le notre, qui est moralement optimal pour des arguments probabilistes.

Le gain est obtenu par l'estimation des quantités (3.4) plutôt que des quantités du lemme 3.2.2, estimation qui peut être intégrée dans nos calculs grâce à la formule de convolution de la proposition 3.1.3.

On rappelle que dans l'énoncé du résultat motivant cette section, le théorème 3.1.4, on doit introduire une distance particulière sur l'espace $\mathbb{C}^n$. On définit la *norme mixte* d'un vecteur $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{C}^n$ par

$$N(\mathbf{x}) = \max\Big(|x_1|, \dots, |x_n|, \sum_{i=1}^{n} |x_i|^2\Big).$$

La fonction N n'est pas à proprement parler une norme, mais c'est un compromis entre la norme absolue et la norme euclidienne. Par souci de concision, nous noterons à présent x_∞ la norme absolue du vecteur $\mathbf{x}$ et $\|\mathbf{x}\|$ sa norme euclidienne.

3.2.1. Égalités polynomiales. — Nous transformons la quantité (3.4) afin d'obtenir une expression polynomiale plus simple à évaluer.

Lemme 3.2.3. — *Soit $x \in \mathbb{C} \smallsetminus \{1\}$. En posant $X = x/(1-x)$, on a pour tout entier $t \in \mathbb{N}$*

$$\frac{1-tx}{(1-x)^t} = 1 - X^2 \sum_{s=0}^{t-2} (s+1) \binom{t}{s+2} X^s.$$

Démonstration. — On a

$$\frac{1-tx}{(1-x)^t} = \frac{(1-x+x)^{t-1}\big(1-x-(t-1)x\big)}{(1-x)^t} = (1+X)^{t-1}\big(1-(t-1)X\big)$$

ce qui vaut en développant

$$= \sum_s \binom{t-1}{s} X^s \big(1-(t-1)X\big) = \sum_s \Big(\binom{t-1}{s} - (t-1)\binom{t-1}{s-1}\Big) X^s.$$

En retranchant à l'identité d'absorption $s\binom{t}{s} = t\binom{t-1}{s}$, la formule d'addition $\binom{t}{s} = \binom{t-1}{s} + \binom{t-1}{s-1}$, on obtient $(s-1)\binom{t}{s} = (t-1)\binom{t-1}{s-1} - \binom{t-1}{s}$, ce qui permet enfin d'écrire

$$= -\sum_s (s-1)\binom{t}{s}X^s = 1 - X^2 \sum_{s\geqslant 0}(s+1)\binom{t}{s+2}X^s. \qquad \square$$

Lemme 3.2.4. — *Soit* $\mathbf{x} = (x_1, \dots, x_n) \in \big(\mathbb{C} \smallsetminus \{1\}\big)^n$. *En posant* $X_i = x_i/(1-x_i)$, *on a pour tout entier* $t \in \mathbb{N}$

$$\prod_{i=1}^n \frac{1-tx_i}{(1-x_i)^t} = \sum_{J\subset[\![1,n]\!]} (-1)^{|J|}\Big(\prod_{j\in J} X_j\Big)^2 \sum_{\substack{s_j\geqslant 0\\(j\in J)}} \prod_{j\in J}(s_j+1)\binom{s}{s_j+2}{X_j}^{s_j}.$$

Démonstration. — Il s'agit de développer le produit obtenu en utilisant le lemme 3.2.3 pour chacune des n variables x_i. $\square$

On introduit, de façon purement fortuite ici mais justifiée par la suite, le coefficient

$$\omega(\mathbf{s}, m) \coloneqq \frac{\prod_i s_i!}{m!}\sum_{k=0}^m (-1)^{m-k}\binom{m}{k}\prod_{i\geqslant 1}\binom{k}{s_i}. \tag{3.5}$$

Proposition 3.2.5. — *Soit* $\mathbf{x} = (x_1, \dots, x_n) \in \big(\mathbb{C} \smallsetminus \{1\}\big)^n$. *En posant* $X_i = x_i/(1-x_i)$, *on a pour tout entier* $t \in \mathbb{N}$

$$\sum_{s=0}^t (-1)^{t-s}\binom{t}{s}\prod_{i=1}^n \frac{1-sx_i}{(1-x_i)^s} = \sum_{J\subset[\![1,n]\!]} (-1)^{|J|}\Big(\prod_{j\in J} X_j\Big)^2 t! \sum_{\substack{s_j\geqslant 0\\(j\in J)}} \omega(\mathbf{s}+\mathbf{2}, t)\prod_{j\in J}\frac{s_j+1}{(s_j+2)!}{X_j}^{s_j}.$$

Démonstration. — On a par le lemme 3.2.4

$$\sum_{s=0}^t (-1)^{t-s}\binom{t}{s}\prod_{i=1}^n \frac{1-sx_i}{(1-x_i)^s} = \sum_{J\subset[\![1,n]\!]} (-1)^{|J|}\Big(\prod_{j\in J} X_j\Big)^2 \sum_{\substack{s_j\geqslant 0\\(j\in J)}} \prod_{j\in J}(s_j+1){X_j}^{s_j} \sum_{s=0}^t (-1)^{t-s}\binom{t}{s}\prod_{j\in J}\binom{s}{s_j+2},$$

ce qui fournit le résultat souhaité. $\square$

3.2.2. Quelques majorations combinatoires. — La proposition 3.2.5 nous pousse donc à majorer le terme

$$\sum_{\substack{s_j\geqslant 0\\(j\in J)}} \omega(\mathbf{s}+\mathbf{2}, t)\prod_{j\in J}\frac{s_j+1}{(s_j+2)!}{X_j}^{s_j}.$$

Ces majorations se basent sur une estimation de la quantité $\omega(\mathbf{s}, m)$ à l'aide des nombres de Stirling de seconde espèce qui sera présentée de façon indépendante dans la section suivante. Pour l'instant, nous admettons que l'encadrement suivant est vérifié pour tout $\mathbf{s} \in \mathbb{N}^J$ et tout $m \geqslant 0$:

$$0 \leqslant \omega(\mathbf{s}, m) \leqslant \left\{ \begin{matrix} \sum_j s_j \\ m \end{matrix} \right\}. \tag{3.6}$$

Lemme 3.2.6. — *En posant $X_\infty = \max_{j\in J} |X_j|$, on a*

$$\left| \sum_{\substack{s_j \geqslant 0 \\ (j\in J)}} \omega(\mathbf{s}+\mathbf{2}, t) \prod_{j\in J} \frac{s_j+1}{(s_j+2)!} {X_j}^{s_j} \right| \leqslant \sum_{s\geqslant 0} \frac{|J|!}{(s+|J|)!} \left\{ \begin{matrix} s+2|J| \\ t \end{matrix} \right\} \left\{ \begin{matrix} s+|J| \\ |J| \end{matrix} \right\} {X_\infty}^s.$$

Démonstration. — En utilisant l'encadrement (3.6), en remarquant que $\frac{s_j+1}{(s_j+2)!} \leqslant \frac{1}{(s_j+1)!}$ et en majorant les $|X_j|$ par X_∞, on a

$$\sum_{\substack{s_j \geqslant 0 \\ (j\in J)}} \omega(\mathbf{s}+\mathbf{2}, t) \prod_{j\in J} \frac{s_j+1}{(s_j+2)!} {X_j}^{s_j} \leqslant \sum_{\substack{s_j \geqslant 0 \\ (j\in J)}} \left\{ \begin{matrix} \sum_J s_j + 2|J| \\ t \end{matrix} \right\} \frac{{X_\infty}^{\sum_J s_j}}{\prod_{j\in J}(s_j+1)!}.$$

En posant $s = \sum_J s_j$ et en faisant apparaître un coefficient multinomial, on obtient une nouvelle expression du majorant en

$$\sum_{s\geqslant 0} \left\{ \begin{matrix} s+2|J| \\ t \end{matrix} \right\} \frac{{X_\infty}^s}{(s+|J|)!} \sum_{\substack{s_j \geqslant 0 \\ (j\in J) \\ \sum_j s_j = s}} \binom{s+|J|}{\mathbf{s}+\mathbf{1}},$$

et la formule (A.13) permet de transformer celle-ci sous la forme attendue. □

Similairement au cas à une unique variable, nous allons supposer $X_\infty \ll 1/t$, ce qui équivaut moralement à supposer $\max |x_i| \ll 1/t$. Pour expliquer par la suite ce « moralement » de façon rigoureuse, nous allons fixer un réel $K > 0$ et nous allons supposer que $tX_\infty \leqslant K$.

Lemme 3.2.7. — *Sous la condition $tX_\infty \leqslant K$, on a*

$$\sum_{s\geqslant 0} \frac{|J|!}{(s+|J|)!} \left\{ \begin{matrix} s+2|J| \\ t \end{matrix} \right\} \left\{ \begin{matrix} s+|J| \\ |J| \end{matrix} \right\} {X_\infty}^s \leqslant \frac{t^{2|J|}}{t!} \Big(\frac{e^K - 1}{K} \Big)^{|J|}.$$

Démonstration. — On utilise la majoration (A.8) pour faire disparaître le premier nombre de Stirling :

$$\sum_{s\geqslant 0}\frac{|J|!}{(s+|J|)!}\left\{{s+2|J|\atop t}\right\}\left\{{s+|J|\atop |J|}\right\}{X_\infty}^s \leqslant \frac{t^{2|J|}}{t!}\sum_{s\geqslant 0}\frac{|J|!}{(s+|J|)!}\left\{{s+|J|\atop |J|}\right\}(tX_\infty)^s = \frac{t^{2|J|}}{t!}\sum_{s\geqslant 0}\frac{|J|!}{s!}\left\{{s\atop |J|}\right\}(tX_\infty)^{s-|J|}$$

où, malgré l'ajout de $|J|$ termes à la somme consécutivement au changement de variable, l'égalité est justifiée par la nullité des nombres de Stirling $\left\{{s\atop |J|}\right\}$ lorsque $s<|J|$. On utilise l'identité (A.5) pour transformer le majorant

$$\frac{t^{2|J|}}{t!}\sum_{s\geqslant 0}\frac{|J|!}{s!}\left\{{s\atop |J|}\right\}(tX_\infty)^{s-|J|} = \frac{t^{2|J|}}{t!}\Big(\frac{e^{tX_\infty}-1}{tX_\infty}\Big)^{|J|}.$$

Par convexité de la fonction exponentielle, la condition $tX_\infty\leqslant K$ fournit la majoration désirée. □

Cette estimation n'exploite pas entièrement la forme particulière du majorant dans l'encadrement (3.6) ; nous verrons que son utilisation n'est pas suffisante pour obtenir de bonnes majorations du moment (3.4). En effet, si $|J|$ est petit devant t, le nombre de Stirling $\left\{{s+2|J|\atop t}\right\}$ est nul pour les premières valeurs de s. Ce fait n'est pas pris en compte dans la démonstration du lemme 3.2.7. Le lemme suivant est énoncé à ce dessein.

Lemme 3.2.8. — *Sous les conditions $tX_\infty\leqslant K$ et $t\geqslant 2|J|$, on a*

$$\sum_{s\geqslant 0}\frac{|J|!}{(s+|J|)!}\left\{{s+2|J|\atop t}\right\}\left\{{s+|J|\atop |J|}\right\}{X_\infty}^s \leqslant \frac{|J|!\,|J|^t}{|J|^{|J|}\,t!}\Big(2\frac{e^K-1}{K}\Big)^t {X_\infty}^{t-2|J|}.$$

Démonstration. — On utilise la majoration (A.8) pour faire disparaître le second nombre de Stirling :

$$\begin{aligned}\sum_{s\geqslant 0}\frac{|J|!}{(s+|J|)!}\left\{{s+2|J|\atop t}\right\}\left\{{s+|J|\atop |J|}\right\}{X_\infty}^s &\leqslant \frac{|J|^{|J|}}{|J|!}\sum_{s\geqslant 0}\frac{|J|!}{(s+|J|)!}\left\{{s+2|J|\atop t}\right\}(|J|X_\infty)^s\\ &= |J|^{|J|}\sum_{s\geqslant 0}\frac{|J|!}{(s+2|J|)!}\binom{s+2|J|}{|J|}\left\{{s+2|J|\atop t}\right\}(|J|X_\infty)^s\\ &\leqslant |J|^{|J|}\sum_{s\geqslant 0}\frac{|J|!}{s!}\binom{s}{|J|}\left\{{s\atop t}\right\}(|J|X_\infty)^{s-2|J|},\end{aligned}$$

où la dernière majoration est une égalité lorsque $t \geqslant 2|J|$, malgré l'ajout de $2|J|$ termes, grâce à la nullité des nombres de Stirling $\left\{ {s \atop t} \right\}$. On utilise la majoration triviale $\binom{s}{|J|} \leqslant 2^s$ et l'identité (A.5) pour obtenir le majorant

$$|J|^{|J|} \sum_{s \geqslant 0} \frac{|J|!}{s!} 2^s \left\{ {s \atop t} \right\} (|J| X_\infty)^{s-2|J|} = \frac{|J|!}{|J|^{|J|} t!} \frac{(e^{2|J|X_\infty} - 1)^t}{{X_\infty}^{2|J|}}.$$

Par convexité de la fonction exponentielle, la condition $2|J|X_\infty \leqslant tX_\infty \leqslant K$ fournit la majoration désirée. □

Ce lemme 3.2.8 améliore le lemme 3.2.7 puisque l'on a :

$$\begin{aligned}\left[\frac{|J|!\,|J|^t}{|J|^{|J|}\,t!}\Big(2\frac{e^K-1}{K}\Big)^t {X_\infty}^{t-2|J|}\right] \Big/ \left[\frac{t^{2|J|}}{t!}\Big(\frac{e^K-1}{K}\Big)^{|J|}\right] &= \frac{|J|!}{|J|^{|J|}}\Big(\frac{K}{e^K-1}\Big)^{|J|}\left(\frac{e^K-1}{K}\right)^t \left(\frac{2|J|}{t}\right)^t (tX_\infty)^{t-2|J|} \\ &\leqslant {c_1}^{|J|} {c_2}^t \left(\frac{2|J|}{t}\right)^t (tX_\infty)^{t-2|J|}.\end{aligned}$$

En négligeant l'influence des deux constantes c_1 et c_2 (dépendentes de K), on voit que le rapport des deux estimations est très petit si $|J|$ est assez petit devant t ou si tX_∞ est assez petit devant K.

3.2.3. Estimations finales. — On rappelle que la norme notée $\|\cdot\|$ est la norme euclidienne et qu'on utilise également la norme mixte N, dont la définition précède l'énoncé du théorème 3.1.4. Nous débutons par trois petits lemmes très généraux, dont le premier est généralement attribué à Erdős.

Lemme 3.2.9. — *Soit k un entier naturel. On a*

$$\sum_{\substack{J \subset [\![1,n]\!] \\ |J|=k}} \Big| \prod_{j \in J} X_j \Big|^2 \leqslant \frac{\|\mathbf{X}\|^{2k}}{k!}.$$

Démonstration. — On a

$$\Big(\sum_{i=1}^n |X_i|^2\Big)^k = \sum_{\substack{i_j \in [\![1,n]\!] \\ (j \in [\![1,k]\!])}} \Big| \prod_{j=1}^k X_{i_j} \Big|^2 \geqslant \sum_{\substack{i_j \in [\![1,n]\!] \\ (j \in [\![1,k]\!]) \\ \forall j \neq j',\, i_j \neq i_{j'}}} \Big| \prod_{j=1}^k X_{i_j} \Big|^2 = k! \sum_{\substack{i_j \in [\![1,n]\!] \\ (j \in [\![1,k]\!]) \\ i_1 < \ldots < i_k}} \Big| \prod_{j=1}^k X_{i_j} \Big|^2.$$

En posant $J = \big\{ i_j ;\, j \in [\![1,k]\!] \big\}$, on retrouve bien l'inégalité annoncée. □

Lemme 3.2.10. — *Soit $\kappa > 0$ un réel. On a*

$$\sum_{J \subset [\![1,n]\!]} \kappa^{|J|} \Big| \prod_{j \in J} X_j \Big|^2 \leqslant \exp\big(\kappa \|\mathbf{X}\|^2\big).$$

Démonstration. — C'est un corollaire du lemme 3.2.9 précédent. On peut aussi remarquer que la somme à majorer est le produit

$$\prod_{i=1}^{n}\bigl(1+\kappa|X_i|^2\bigr)\leqslant\exp\Bigl(\kappa\sum_{i=1}^{n}|X_i|^2\Bigr).$$ □

Lemme 3.2.11. — *Soient $\kappa>0$ un réel et k un entier naturel. On a*

$$\sum_{\substack{J\subset[\![1,n]\!]\\|J|\geqslant k}}\kappa^{|J|}\Bigl|\prod_{j\in J}X_j\Bigr|^2\leqslant\kappa^k\exp\bigl(\kappa\|\mathbf{X}\|^2\bigr)\frac{\|\mathbf{X}\|^{2k}}{k!}.$$

Démonstration. — Il s'agit d'un corollaire du lemme 3.2.9 précédent si l'on remarque que $\sum_{n\geqslant k}\frac{x^n}{n!}\leqslant\frac{x^k}{k!}e^x$. On peut aussi remarquer qu'un ensemble J à plus de k éléments peut s'écrire plus d'une fois comme réunion disjointe d'un ensemble J_1 à k élément et d'un autre ensemble J_2, donc

$$\sum_{\substack{J\subset[\![1,n]\!]\\|J|\geqslant k}}\kappa^{|J|}\Bigl|\prod_{j\in J}X_j\Bigr|^2\leqslant\Biggl[\kappa^k\sum_{\substack{J_1\subset[\![1,n]\!]\\|J_1|=k}}\Bigl|\prod_{j\in J_1}X_j\Bigr|^2\Biggr]\times\Biggl[\sum_{J_2\subset[\![1,n]\!]}\kappa^{|J_2|}\Bigl|\prod_{j\in J_2}X_j\Bigr|^2\Biggr].$$

Les lemmes 3.2.9 et 3.2.10 permettent de conclure. □

On peut à présent donner une majoration précise de la quantité (3.4), de laquelle se déduit facilement le théorème 3.1.4.

Proposition 3.2.12. — *Soit $K>0$ un réel, t un entier naturel et $\mathbf{x}\in\bigl(\mathbb{C}\smallsetminus\{1\}\bigr)^n$ un vecteur complexe. On définit le vecteur complexe associé $\mathbf{X}=\bigl(\frac{x_1}{1-x_1},\dots,\frac{x_n}{1-x_n}\bigr)$. Sous la condition $tN(\mathbf{X})\leqslant K$, on a*

$$\sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}\prod_{i=1}^{n}\frac{1-sx_i}{(1-x_i)^s}\leqslant{C_1}^t\bigl(tN(\mathbf{X})\bigr)^{\lceil t/2\rceil}+C_2{C_3}^t\bigl(t\|\mathbf{X}\|^2\bigr)^{\lfloor t/2\rfloor+1},$$

avec

$$C_1=\sqrt{2+K}\frac{e^K-1}{K},\qquad C_2=2\frac{e^K-1}{K}\quad\text{et}\quad C_3=e^{e^K}\sqrt{\frac{2}{e}\frac{e^K-1}{K}}.$$

Démonstration. — On sépare la somme de la proposition 3.2.5 en deux, selon la valeur de $|J|$ par rapport à $t/2$, et l'on applique respectivement à l'une et l'autre des

sous-sommes obtenues, les majorations des lemmes 3.2.8 et 3.2.7 :

$$\begin{aligned}
&\sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}\prod_{i=1}^{n}\frac{1-sx_i}{(1-x_i)^s}\\
&\qquad=\sum_{\substack{J\subset[\![1,n]\!]\\|J|\leqslant t/2}}(-1)^{|J|}\Big(\prod_{j\in J}X_j\Big)^2 t!\sum_{\substack{s_j\geqslant 0\\(j\in J)}}\omega(\mathbf{s}+\mathbf{2},t)\prod_{j\in J}\frac{s_j+1}{(s_j+2)!}{X_j}^{s_j}\\
&\qquad\quad+\sum_{\substack{J\subset[\![1,n]\!]\\|J|>t/2}}(-1)^{|J|}\Big(\prod_{j\in J}X_j\Big)^2 t!\sum_{\substack{s_j\geqslant 0\\(j\in J)}}\omega(\mathbf{s}+\mathbf{2},t)\prod_{j\in J}\frac{s_j+1}{(s_j+2)!}{X_j}^{s_j}\\
&\qquad\leqslant\sum_{\substack{J\subset[\![1,n]\!]\\|J|\leqslant t/2}}\Big|\prod_{j\in J}X_j\Big|^2\frac{|J|!}{|J|^{|J|}}|J|^t\Big(2\frac{e^K-1}{K}\Big)^t{X_\infty}^{t-2|J|}\\
&\qquad\qquad\qquad+\sum_{\substack{J\subset[\![1,n]\!]\\|J|>t/2}}\Big|\prod_{j\in J}X_j\Big|^2 t^{2|J|}\Big(\frac{e^K-1}{K}\Big)^{|J|}.
\end{aligned}$$

On pose par commodité $k=|J|$ et $c_1=(e^K-1)/K$. On traite d'abord la première somme Σ_1 grâce au lemme 3.2.9

$$\begin{aligned}
\Sigma_1&\leqslant 2^t{c_1}^t\sum_{k=0}^{\lfloor t/2\rfloor}k^{t-k}\|\mathbf{X}\|^{2k}{X_\infty}^{t-2k}\\
&=2^t{c_1}^t\sum_{k=0}^{\lfloor t/2\rfloor}k^{\lceil t/2\rceil}(kX_\infty)^{\lfloor t/2\rfloor-k}\|\mathbf{X}\|^{2k}{X_\infty}^{\lceil t/2\rceil-k}
\end{aligned}$$

or on a $\|\mathbf{X}\|^{2k}{X_\infty}^{\lceil t/2\rceil-k}\leqslant\max\big(\|\mathbf{X}\|^2,X_\infty\big)^{\lceil t/2\rceil}=N(\mathbf{X})^{\lceil t/2\rceil}$, $k\leqslant t/2$ et $kX_\infty\leqslant K/2$, donc

$$\begin{aligned}
&\leqslant 2^{\lfloor t/2\rfloor}{c_1}^t\big(tN(\mathbf{X})\big)^{\lceil t/2\rceil}\sum_{k=0}^{\lfloor t/2\rfloor}\Big(\frac{K}{2}\Big)^k\\
&\leqslant(2+K)^{t/2}{c_1}^t\big(tN(\mathbf{X})\big)^{\lceil t/2\rceil}\\
&\leqslant{C_1}^t\big(tN(\mathbf{X})\big)^{\lceil t/2\rceil}.
\end{aligned}$$

Pour la seconde somme, on utilise le lemme 3.2.11

$$\Sigma_2\leqslant(c_1t^2)^{\lfloor t/2\rfloor+1}\exp\big(c_1t^2\|\mathbf{X}\|^2\big)\frac{\|\mathbf{X}\|^{2\lfloor t/2\rfloor+2}}{\big(\lfloor t/2\rfloor+1\big)!}$$

et grâce aux inégalités $n! \geqslant (n/e)^n e$ et $t\|\mathbf{X}\|^2 \leqslant K$

$$\begin{aligned}
&\leqslant e^{-1}\big(c_1 e\tfrac{t^2}{\lfloor t/2\rfloor+1}\|\mathbf{X}\|^2\big)^{\lfloor t/2\rfloor+1} e^{c_1 K t}\\
&\leqslant 2c_1\big(2c_1 e^{2c_1K+1}\big)^{t/2}\big(t\|\mathbf{X}\|^2\big)^{\lfloor t/2\rfloor+1}\\
&= C_2{C_3}^t\big(t\|\mathbf{X}\|^2\big)^{\lfloor t/2\rfloor+1}.
\end{aligned}$$

□

En remarquant que $C_1 \leqslant C_3$ et que

$$\big(t\|\mathbf{X}\|^2\big)^{\lfloor t/2\rfloor+1} \leqslant \big(tN(\mathbf{X})\big)^{\lceil t/2\rceil}\big(tN(\mathbf{X})\big)^{[2|t]} \leqslant (1+K)\big(tN(\mathbf{X})\big)^{\lceil t/2\rceil},$$

on peut donner une version simplifiée du majoration de la proposition sous la forme

$$\sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}\prod_{i=1}^{n}\frac{1-sx_i}{(1-x_i)^s} \leqslant \big((1+K)C_2+1\big)\,{C_3}^t\,\big(tN(\mathbf{X})\big)^{\lceil t/2\rceil},$$

ce que l'on note de façon compacte (mais moins précise)

$$\sum_{s=0}^{t}(-1)^{t-s}\binom{t}{s}\prod_{i=1}^{n}\frac{1-sx_i}{(1-x_i)^s} \ll (c_K)^t\big(tN(\mathbf{X})\big)^{\lceil t/2\rceil},$$

où la constante c_K ne dépend que de K.

Cette proposition 3.2.12 précise le théorème 3.1.4, qui en découle immédiatement. En effet, si l'on souhaite se passer de la norme $N(\mathbf{X})$ pour majorer le moment (3.4), il est nécessaire d'avoir $N(\mathbf{X}) \ll_K N(\mathbf{x})$. Pour cela, il faut s'assurer que les composantes x_i sont relativement éloignés de 1. Cela se déduit de la condition $tN(\mathbf{x}) \leqslant K$ dès que $t > K$. Dans le cas inverse, on est contraint d'introduire la condition explicite $|1-x_i| \geqslant \varepsilon$ du théorème 3.1.4.

3.3. Un lemme combinatoire

Cette section est consacrée à l'étude et à la majoration de la quantité $\omega(\mathbf{s}, m)$ dont nous rappelons ici la définition (3.5) :

$$\omega(\mathbf{s}, m) = \frac{\prod_i s_i!}{m!}\sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}\prod_{i\geqslant 1}\binom{k}{s_i},$$

qui apparaît au cours de la section précédente, et dont la majoration (3.6) permet d'obtenir les théorèmes 3.1.4 puis 3.1.8.

On commence par réécrire cette valeur en termes de nombres de Stirling. Nous avons regroupés dans l'appendice un ensemble de définitions et de formules, qui sont classiques en combinatoire mais qui peuvent être utile à un lecteur néophyte.

Lemme 3.3.1. — *Soit $\mathbf{s} \in \mathbb{N}^J$ et m un entier naturel. On a*

$$\omega(\mathbf{s}, m) = \sum_{\substack{t_j \leqslant s_j\\ (j\in J)}}\left\{\begin{matrix}\sum_J t_j\\ m\end{matrix}\right\}(-1)^{\sum_J(s_j-t_j)}\prod_{j\in J}\begin{bmatrix}s_j\\ t_j\end{bmatrix}.$$

Démonstration. — Par l'identité $\binom{k}{s_j} = \frac{1}{s_j!}\sum_{t_j}(-1)^{s_j-t_j}\left[\begin{smallmatrix}s_j\\t_j\end{smallmatrix}\right]k^{t_j}$ issue de la formule (A.10), on peut écrire

$$\sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}\prod_{j\in J}\binom{k}{s_j}$$
$$= \frac{1}{\prod_j s_j!}\sum_{\substack{0\leqslant t_j\leqslant s_j\\(j\in J)}}\prod_{i\geqslant 1}(-1)^{s_j-t_j}\begin{bmatrix}s_j\\t_j\end{bmatrix}\sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}k^{\sum_J t_j}$$

puis par l'identité $\sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}k^{\sum_J t_j} = m!\left\{\begin{smallmatrix}\sum_J t_j\\m\end{smallmatrix}\right\}$ issue de la formule (A.9)

$$= \frac{m!}{\prod_J s_j!}\sum_{\substack{t_j\leqslant s_j\\(j\in J)}}\left\{\begin{matrix}\sum_J t_j\\m\end{matrix}\right\}(-1)^{\sum_J(s_j-t_j)}\prod_{j\in J}\begin{bmatrix}s_j\\t_j\end{bmatrix},$$

ce qui donne bien la relation proposée. □

La quantité étudiée dans ce lemme — à savoir $\sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}\prod_J\binom{k}{s_j}$ — possède une interprétation combinatoire : elle compte le nombre de suites $(E_j)_{j\in J}$ de sous-ensembles de $[\![1,m]\!]$ vérifiant les conditions $\operatorname{card}E_j = s_j$ pour tout $j\in J$ et $\bigcup_{j\in J}E_j = [\![1,m]\!]$. Cette remarque permet déjà de fournir un majorant pratique de $\omega(\mathbf{s},m)$. Mais nous allons utiliser un argument énumératif un peu plus complexe pour obtenir la borne annoncée.

3.3.1. Un peu de vocabulaire sur les partitions et les partages. — Pour éviter toute confusion entre les deux vocables, nous donnons une définition précise de chacun d'entre eux.

On appelle **partition** d'un ensemble fini E tout ensemble de sous-ensembles non-vides de E qui vérifie les conditions suivantes :
- la réunion de ces sous-ensembles vaut E ;
- l'intersection de toute paire de sous-ensembles est vide.

Les sous-ensembles sont appelés les parties ou **composantes** de la partition.

On appelle **partage** d'un entier n toute collection[‡] finie d'entiers strictement positifs dont la somme vaut n. Ces entiers sont appelés les **parts** du partage. Par commodité on simulera l'absence d'ordre dans la collection des entiers en considérant des suites décroissantes d'entiers positifs dont la somme vaut n. Ainsi on représentera souvent le partage $\boldsymbol{\lambda}$ sous la forme $\lambda_1\geqslant\lambda_2\geqslant\ldots\geqslant\lambda_k>0$.

Pour une partition π de l'ensemble E de cardinal n, on appelle **partage cardinal** de π le partage de n composé des cardinaux de chaque partie de π.

Nous choisissons $[\![1,n]\!]$ comme représentant d'un ensemble de cardinal n générique. La plupart des constructions considérées peuvent s'étendre à tout ensemble fini, mais notre propos n'est pas là. On construit deux types d'objets sur $[\![1,n]\!]$:

[‡]avec répétitions, sans ordre.

– l'ensemble des permutations des éléments de $[\![1, n]\!]$, que l'on note $\mathfrak{S}_n$; on peut le voir comme l'ensemble des bijections de $[\![1, n]\!]$ dans lui-même, et la composition lui confère une structure de groupe : on l'appelle *le groupe symétrique* ;
– l'ensemble des partitions de $[\![1, n]\!]$, que l'on note Π_n ; on peut le voir comme l'ensemble des relations d'équivalence sur $[\![1, n]\!]$, et la conjonction et la disjonction des relations lui confèrent une structure de treillis : on l'appelle classiquement *le treillis des partitions.*

On considère l'action de $\mathfrak{S}_n$ sur $[\![1, n]\!]$; celle-ci induit une action sur l'ensemble des parties de $[\![1, n]\!]$, dont les orbites sont caractérisées par le cardinal des parties, à savoir que deux parties de $[\![1, n]\!]$ sont dans la même orbite sous $\mathfrak{S}_n$ si et seulement si elles ont même cardinal. Cette action sur les parties induit une action sur les partitions de $[\![1, n]\!]$; les orbites sous cette action sont caractérisées par les cardinaux des parties de la partition, c'est-à-dire par le partage cardinal des partitions, qui est un partage de n. Ainsi deux partitions de $[\![1, n]\!]$ sont équivalentes sous l'action de $\mathfrak{S}_n$ si et seulement si elles ont même partage cardinal.

On a vu que Π_n est un treillis si on le munit de la conjonction et de la disjonction des relations d'équivalence. Plus concrètement, pour deux partitions π_1 et π_2, leur conjonction $\pi_1 \wedge \pi_2$ est la partition dont les parties sont les intersections non vides des parties de π_1 avec les parties de π_2, leur disjonction $\pi_1 \vee \pi_2$ est obtenue en réunissant toutes les parties non disjointes de π_1 et π_2. On observe que l'action de $\mathfrak{S}_n$ conserve ces notions, c'est-à-dire que pour $\sigma \in \mathfrak{S}_n$ on a $(\sigma \cdot \pi_1) \wedge (\sigma \cdot \pi_2) = \sigma \cdot (\pi_1 \wedge \pi_2)$ et pareillement pour la disjonction.

Un treillis peut également se définir par une relation d'ordre partiel $\leqslant$, définie par $\pi_1 \leqslant \pi_2 \Leftrightarrow \pi_1 \wedge \pi_2 = \pi_1 \Leftrightarrow \pi_1 \vee \pi_2 = \pi_2$, si bien que la conjonction de deux partitions en est la borne inférieure et la disjonction, la borne supérieure. Cet ordre recouvre la notion intuitive de sous-partition. Le treillis des partitions admet une partition minimale $\pi_{\min}$, la partition en singletons, élément absorbant de la loi $\wedge$ et élément neutre de la loi $\vee$, et une partition maximale $\pi_{\max}$, la partition en une seule partie, élément neutre de la loi $\wedge$ et élément absorbant de la loi $\vee$. Le partage cardinal de $\pi_{\max}$ est le partage en une part de valeur n, le partage cardinal de $\pi_{\min}$ est le partage en n parts valant 1 chacune.

On peut alors se poser différentes questions : pour une partition π_1 donnée, combien y a-t-il de partitions π_2 vérifiant $\pi_1 \wedge \pi_2 = \pi_{\min}$? Combien vérifiant $\pi_1 \vee \pi_2 = \pi_{\max}$? Grieser [**15**] s'est intéressé au nombre de partitions π_2 vérifiant les deux conditions à la fois. J'ignore si on sait plus de choses sur l'une de ces trois questions. Pour notre étude, nous n'allons traiter que la première, et ce d'une façon utile à notre propos. Nous dirons donc que deux partitions π_1 et $\pi_2 \in \Pi_n$ sont *perpendiculaires* si elles vérifient $\pi_1 \wedge \pi_2 = \pi_{\min}$.

Pour se faire, on considère pour une partition $\pi \in \Pi_n$ fixée et pour un entier naturel m, le nombre de partitions $\pi' \in \Pi_n$ en m parties vérifiant $\pi \wedge \pi' = \pi_{\min}$. Comme toutes ces notions sont stables sous l'action de $\mathfrak{S}_n$, on peut définir ce nombre non sur les partitions de Π_n mais sur les orbites de Π_n sous l'action de $\mathfrak{S}_n$, c'est-à-dire sur les partages $\boldsymbol{\lambda}$ de l'entier n. Donc pour $\boldsymbol{\lambda}$ partage de l'entier n et m entier naturel,

on pose $\tilde{\omega}(\boldsymbol{\lambda}, m)$ le nombre de partitions $\pi' \in \Pi_n$ en m parties, vérifiant $\pi \wedge \pi' = \pi_{\min}$, où $\pi \in \Pi_n$ est une partition fixée quelconque de partage cardinal $\boldsymbol{\lambda}$.

TABLE 1. Valeurs de $\tilde{\omega}(\boldsymbol{\lambda}, m)$ pour $\boldsymbol{\lambda}$ partages des entiers 2, 3, 4 et 5

$\boldsymbol{\lambda} \diagdown m$	1	2
2		1
1, 1	1	1

$\boldsymbol{\lambda} \diagdown m$	1	2	3
3			1
2, 1		2	1
1, 1, 1	1	3	1

$\boldsymbol{\lambda} \diagdown m$	1	2	3	4
4				1
3, 1			3	1
2, 2		2	4	1
2, 1, 1		4	5	1
1, 1, 1, 1	1	7	6	1

$\boldsymbol{\lambda} \diagdown m$	1	2	3	4	5
5					1
4, 1				4	1
3, 2			6	6	1
3, 1, 1			9	7	1
2, 2, 1		4	14	8	1
2, 1, 1, 1		8	19	9	1
1, 1, 1, 1, 1	1	15	25	10	1

Lemme 3.3.2. — *Soient m et n deux entiers naturels et $\boldsymbol{\lambda}$ un partage de l'entier n. On a les propriétés suivantes :*

a). *On a* $0 \leqslant \tilde{\omega}(\boldsymbol{\lambda}, m) \leqslant \left\{ {n \atop m} \right\}$;
b). *Si* $m > n$, *on a* $\tilde{\omega}(\boldsymbol{\lambda}, m) = 0$;
c). *Si* $m = n$, *on a* $\tilde{\omega}(\boldsymbol{\lambda}, m) = 1$;
d). *Si* $m < \lambda_1$, *on a* $\tilde{\omega}(\boldsymbol{\lambda}, m) = 0$;
e). *Si* $\boldsymbol{\lambda}$ *est le partage en* 1 *part, on a* $\tilde{\omega}(\boldsymbol{\lambda}, m) = [m = n]$;
f). *Si* $\boldsymbol{\lambda}$ *est le partage en* n *parts, on a* $\tilde{\omega}(\boldsymbol{\lambda}, m) = \left\{ {n \atop m} \right\}$;

Démonstration. — Comme la fonction $\tilde{\omega}(\boldsymbol{\lambda}, m)$ compte des partitions en m parties de l'ensemble $[\![1, n]\!]$, on obtient bien l'encadrement du *a)*. Le *b)* s'en déduit.
Il n'y a qu'une seule partition dans Π_n en n parties, c'est la partition minimale $\pi_{\min}$, et comme elle est l'élément absorbant de la loi $\wedge$, on obtient bien le *c)*.
Soit une partition $\pi \in \Pi_n$ de partage cardinal $\boldsymbol{\lambda}$: l'une de ses parties est de cardinal λ_1, on la note I. Par le principe des tiroirs, pour toute partition $\pi' \in \Pi_n$ en m parties, avec $m < \lambda_1$, il existe deux éléments de la partie I présents dans une même partie de la partition $\pi \wedge \pi'$: on a donc $\pi \wedge \pi' \neq \pi_{\min}$ et le *d)* s'en déduit.
Si $\boldsymbol{\lambda}$ est le partage en une part, on a $\lambda_1 = n$, et les points *b), c)* et *d)* établissent *e)*.
La seule partition ayant pour partage cardinal le partage en n parts est la partition minimale $\pi_{\min}$, élément absorbant de la loi $\wedge$. On en déduit *f)*. □

Par commodité, on étend le domaine de définition de $\tilde{\omega}(\,\cdot\,, m)$ à l'ensemble $\mathbb{N}^{(\mathbb{N}^*)}$ des suites d'entiers naturels tendant vers 0. C'est un monoïde de base canonique $(\delta_j)_{j \in \mathbb{N}^*}$ définie par $\delta_j(n) = [j = n]$. On dispose sur cet ensemble $\mathbb{N}^{(\mathbb{N}^*)}$

– d'un ordre partiel $\leqslant$: défini par $\mathbf{s} \leqslant \mathbf{t} \Leftrightarrow \forall i \geqslant 1,\ s_i \leqslant t_i$;
– d'une fonction *longueur* ℓ : définie par $\ell\mathbf{s} = \mathrm{card}\{i \in \mathbb{N}^*;\ s_i \geqslant 1\}$, le nombre de termes non nuls de la suite ;
– d'une fonction *somme* Σ : définie par $\Sigma\mathbf{s} = \sum_{i\in\mathbb{N}^*} s_i$, la somme des termes de la suite.

Quitte à changer l'ordre des termes et à retirer les termes nuls, on peut associer à une suite $\mathbf{s}$ un unique partage $\boldsymbol{\sigma}$ de $\Sigma\mathbf{s}$ en $\ell\mathbf{s}$ parts ; on pose $\tilde{\omega}(\mathbf{s}, m) \coloneqq \tilde{\omega}(\boldsymbol{\sigma}, m)$.

Lemme 3.3.3. — *Soient $\mathbf{s} \in \mathbb{N}^{(\mathbb{N}^*)}$ une suite et $j,\ m \geqslant 1$ deux entiers. On a*

$$\tilde{\omega}(\mathbf{s} + \delta_j, m) = (m - s_j)\,\tilde{\omega}(\mathbf{s}, m) + \tilde{\omega}(\mathbf{s}, m-1).$$

Démonstration. — On pose $\boldsymbol{\sigma}$ comme le partage associé à $\mathbf{s}$, et $\boldsymbol{\sigma'}$ comme le partage associé à $\mathbf{s}+\delta_j$. On fixe une partition $\pi_1 \in \Pi_{\Sigma\mathbf{s}}$ de partage cardinal $\boldsymbol{\sigma}$, et on construit la partition $\pi_1' \in \Pi_{\Sigma\mathbf{s}+1}$ en ajoutant l'élément $\Sigma\mathbf{s}+1$ à une partie de π_1 de cardinal s_j. On note I cette partie de π_1. Le partage cardinal de π_1' est donc $\boldsymbol{\sigma'}$. Pour toute partition $\pi_2' \in \Pi_{\Sigma\mathbf{s}+1}$ en m parties vérifiant $\pi_1' \wedge \pi_2' = \pi_{\min}$, on construit la partition $\pi_2 \in \Pi_{\Sigma\mathbf{s}}$ en retirant l'élément $\Sigma\mathbf{s}+1$. On a alors $\pi_1 \wedge \pi_2 = \pi_{\min}$. Le nombre de parties de la partition π_2 est m ou $m-1$.

Le nombre de parties de π_2 est $m-1$ si et seulement si le singleton $\{\Sigma\mathbf{s}+1\}$ est une partie de la partition π_2'. Il y a donc une correspondance biunivoque entre les partitions en m parties perpendiculaires à π_1' dont une partie est le singleton $\{\Sigma\mathbf{s}+1\}$, et les partitions en $m-1$ parties perpendiculaires à π_1. Le cardinal commun de ces ensembles est $\tilde{\omega}(\mathbf{s}, m-1)$.

Dans le cas où le singleton $\{\Sigma\mathbf{s}+1\}$ n'est pas une partie π_2', la partition π_2 est en m parties. Pour construire à partir d'une des $\tilde{\omega}(\mathbf{s}, m)$ partitions π_2 en m parties perpendiculaires à π_1 une partition p_2' en m parties, il faut et suffit d'ajouter l'élément $\Sigma\mathbf{s}+1$ à l'une des m parties de π_2. Cela fait donc m possibilités. Cependant, une conditon nécessaire et suffisante pour la perpendicularité entre π_1' et π_2', est que la partie à laquelle on rajoute $\Sigma\mathbf{s}+1$ ne contienne pas déjà un élément i de I, car dans le cas contraire, une partie de la partition $\pi_1' \wedge \pi_2'$ contiendrait simultanément l'élément i et l'élément $\Sigma\mathbf{s}+1$. Or la condition $\pi_1 \wedge \pi_2 = \pi_{\min}$ indique que les s_j éléments de I sont placés dans s_j parties distinctes des m parties de π_2. Il y a donc exactement s_j parties à éviter pour placer l'élément $\Sigma\mathbf{s}+1$: cela laisse donc $m-s_j$ choix. Ces $m-s_j$ constructions de π_2' à partir d'une des $\tilde{\omega}(\mathbf{s}, m)$ partitions π_2 sont toutes distinctes et permettent de construire toutes les partitions p_2' en m parties perpendiculaires à π_1' pour lesquelles le singleton $\{\Sigma\mathbf{s}+1\}$ n'est pas une partie : elles sont donc au nombre de $(m-s_j)\,\tilde{\omega}(\mathbf{s}, m)$. □

Cette relation de récurrence fait penser à un mélange des deux relations de récurrence (A.3) sur les nombres de Stirling.
On rappelle que pour une suite $\mathbf{s} \in \mathbb{N}^{(\mathbb{N}^*)}$ et un entier $m \in \mathbb{N}$, on a grâce au lemme 3.3.1

$$\omega(\mathbf{s}, m) = \sum_{\mathbf{t}\leqslant\mathbf{s}} \left\{ {\Sigma\mathbf{t} \atop m} \right\} (-1)^{\Sigma\mathbf{s}-\Sigma\mathbf{t}} \prod_{i\geqslant 1} \begin{bmatrix} s_i \\ t_i \end{bmatrix}.$$

Le produit infini n'est pas gênant car pour i assez grand on a $s_i = t_i = 0$ et donc $\begin{bmatrix} s_i \\ t_i \end{bmatrix} = 1$.

Lemme 3.3.4. — *Soient* $\mathbf{s} \in \mathbb{N}^{(\mathbb{N}^*)}$ *une suite et* $m \geqslant 1$ *un entier. On a les propriétés suivantes :*

a). *Si* $m > \Sigma\mathbf{s}$, *on a* $\omega(\mathbf{s}, m) = 0$;
b). *Si* $\ell\mathbf{s} = 1$, *on a* $\omega(\mathbf{s}, m) = [m = \Sigma\mathbf{s}]$;
c). *Si* $\ell\mathbf{s} = \Sigma\mathbf{s}$, *on a* $\omega(\mathbf{s}, m) = \left\{ {\Sigma\mathbf{s} \atop m} \right\}$.

Il faut rapprocher ces trois propriétés respectivement des propriétés *b), e)* et *f)* du lemme 3.3.2.

Démonstration. — Pour tout $\mathbf{t}$ intervenant dans la somme, on a $\Sigma\mathbf{t} \leqslant \Sigma\mathbf{s}$. Ainsi lorsque $m > \Sigma\mathbf{s}$, on a $\left\{ {\Sigma\mathbf{t} \atop m} \right\} = 0$ pour toute suite $\Sigma\mathbf{t} \leqslant \Sigma\mathbf{s}$. Par conséquent, on a $\omega(\mathbf{s}, m) = 0$.
Pour tout $\mathbf{t}$ intervenant dans la somme, on a $\ell\mathbf{t} \leqslant \ell\mathbf{s}$. Ainsi lorsque $\ell\mathbf{s} = 1$, on a $\ell\mathbf{t} \leqslant 1$ pour toute suite $\Sigma\mathbf{t} \leqslant \Sigma\mathbf{s}$. On note s l'unique valeur non vide de $\mathbf{s}$. On a alors $s = \Sigma\mathbf{s}$. Ainsi on peut écrire

$$\omega(\mathbf{s}, m) = \sum_{t=0}^{s} \left\{ {t \atop m} \right\} (-1)^{s-t} \begin{bmatrix} s \\ t \end{bmatrix},$$

qui vaut $[m = s]$ selon la formule d'inversion (A.11).
Si $\ell\mathbf{s} = \Sigma\mathbf{s}$, la suite $\mathbf{s}$ prend ses valeurs dans l'ensemble $\{0, 1\}$, et pareillement pour toute suite $\mathbf{t}$ vérifiant $\Sigma\mathbf{t} \leqslant \Sigma\mathbf{s}$. Or on a $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1$ et $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0$, si bien que $\prod_{i \geqslant 1} \begin{bmatrix} s_i \\ t_i \end{bmatrix} = [\mathbf{s} = \mathbf{t}]$. On a donc dans ce cas $\omega(\mathbf{s}, m) = \left\{ {\Sigma\mathbf{s} \atop m} \right\}$. □

Lemme 3.3.5. — *Soit* $\mathbf{s} \in \mathbb{N}^{(\mathbb{N}^*)}$ *une suite et* j, $m \geqslant 1$ *deux entiers. On a*

$$\omega(\mathbf{s} + \delta_j, m) = (m - s_j)\, \omega(\mathbf{s}, m) + \omega(\mathbf{s}, m - 1).$$

Démonstration. — En sommant d'abord sur $t = \Sigma\mathbf{t}$, on peut écrire

$$\omega(\mathbf{s}, m) = \sum_t (-1)^t \left\{ {t \atop m} \right\} \sum_{\sum_i t_i = t} (-1)^{\Sigma\mathbf{s}} \prod_{i \geqslant 1} \begin{bmatrix} s_i \\ t_i \end{bmatrix},$$

où il n'est pas nécessaire de préciser les bornes de sommation puisqu'elles sont imposées par les nombres de Stirling. On va transformer le terme

$$\omega(\mathbf{s} + \delta_j, m) = \sum_t (-1)^t \left\{ {t \atop m} \right\} \sum_{\sum_i t_i = t} (-1)^{\Sigma\mathbf{s}+1} \begin{bmatrix} s_j + 1 \\ t_j \end{bmatrix} \prod_{i \neq j} \begin{bmatrix} s_i \\ t_i \end{bmatrix}$$

en utilisant la relation de récurrence (A.3) $\begin{bmatrix} s_j + 1 \\ t_j \end{bmatrix} = s_j \begin{bmatrix} s_j \\ t_j \end{bmatrix} + \begin{bmatrix} s_j \\ t_j - 1 \end{bmatrix}$ pour obtenir

$$s_j \sum_t (-1)^t \left\{ {t \atop m} \right\} \sum_{\sum_i t_i = t} (-1)^{\Sigma\mathbf{s}+1} \prod_{i \geqslant 1} \begin{bmatrix} s_i \\ t_i \end{bmatrix} + \sum_t (-1)^t \left\{ {t \atop m} \right\} \sum_{\sum_i t_i = t} (-1)^{\Sigma\mathbf{s}+1} \begin{bmatrix} s_j \\ t_j - 1 \end{bmatrix} \prod_{i \neq j} \begin{bmatrix} s_i \\ t_i \end{bmatrix}.$$

Le premier terme vaut $-s_j\,\omega(\mathbf{s},m)$. On réécrit le second terme en effectuant une translation d'une unité sur les variables t et t_j pour obtenir

$$\omega(\mathbf{s}+\delta_j,m)+s_j\,\omega(\mathbf{s},m)=\sum_t(-1)^{t+1}\left\{{t+1\atop m}\right\}\sum_{\sum_i t_i=t}(-1)^{\Sigma\mathbf{s}+1}\prod_{i\geqslant 1}\begin{bmatrix}s_i\\t_i\end{bmatrix}.$$

On utilise à présent la relation de récurrence (A.3) $\left\{{t+1\atop m}\right\}=m\left[{t\atop m}\right]+\left[{t\atop m-1}\right]$ pour obtenir

$$m\sum_t(-1)^t\left\{{t\atop m}\right\}\sum_{\sum_i t_i=t}(-1)^{\Sigma\mathbf{s}}\prod_{i\geqslant 1}\begin{bmatrix}s_i\\t_i\end{bmatrix}$$
$$+\sum_t(-1)^t\left\{{t\atop m-1}\right\}\sum_{\sum_i t_i=t}(-1)^{\Sigma\mathbf{s}}\begin{bmatrix}s_j\\t_j-1\end{bmatrix}\prod_{i\neq j}\begin{bmatrix}s_i\\t_i\end{bmatrix},$$

ce qui vaut $m\,\omega(\mathbf{s},m)+\omega(\mathbf{s},m-1)$. En regroupant les différents termes, on obtient bien la formule souhaitée. □

On obtient donc grâce aux lemmes 3.3.2, 3.3.3, 3.3.4 et 3.3.5 l'égalité pour toute suite $\mathbf{s}\in\mathbb{N}^{(\mathbb{N}^*)}$ et tout entier $m\in\mathbb{N}$

$$\omega(\mathbf{s},m)=\tilde{\omega}(\mathbf{s},m).$$

On en déduit grâce aux propriétés *a)*, *c)* et *d)* du lemme 3.3.2 l'encadrement suivant, qui lui-même implique l'encadrement (3.6) utilisé au cours de la preuve du théorème 3.1.4.

Proposition 3.3.6. — *Soient $n\geqslant 1$ un entier et $(s_1,\dots,s_n)\in\mathbb{N}^n$ un vecteur d'entiers. On pose $s=\max_i s_i$ et $S=\sum_i s_i$. On a pour tout entier $m\in\mathbb{N}$ on a*

$$[m=S]\leqslant\sum_{\substack{0\leqslant t_i\leqslant s_i\\(1\leqslant i\leqslant n)}}\left\{{\sum_i t_i\atop m}\right\}(-1)^{S-\sum_i t_i}\prod_{i=1}^n\begin{bmatrix}s_i\\t_i\end{bmatrix}\leqslant\left\{{S\atop m}\right\}[s\leqslant m\leqslant S].$$

CHAPITRE 4

DU LEMME FONDAMENTAL

Dans ce chapitre, nous changeons de point de vue sur la question de la répartition des entiers premiers relativement à un entier q donné. Plutôt que le point de vue probabiliste dont nous avons la sensation d'avoir tiré aux chapitres précédents des résultats satisfaisants, nous considérons le point vue harmonique, développé également dans l'article de Montgomery et Vaughan [**30**]. Nous donnons d'abord une preuve nouvelle de leur *lemme fondamental*, unifiant les précédentes démonstrations. Dans un cas spécifique, nous montrons que cette majoration peut-être précisée.

4.1. Analyse harmonique

L'une des grandes nouveautés du travail de Montgomery et Vaughan est d'avoir su tirer profit d'une identité obtenue de façon élémentaire grâce à l'analyse harmonique discrète. Ce calcul peut en fait se généraliser facilement à l'étude de la répartition des valeurs d'une fonction paire modulo q.

Soient donc q un entier naturel positif et f une fonction paire modulo q, c'est-à-dire telle que pour tout entier n, on a $f(n) = f\big((n,q)\big)$. On note $P \coloneqq \frac{1}{q}\sum_{n=1}^{q} f(n)$ sa moyenne et on s'intéresse au k-ième moment centré

$$M_k(h;q) = \sum_{n\in\mathbb{Z}/q\mathbb{Z}} \Big(\sum_{i=1}^{h} f(n+i) - hP\Big)^k.$$

On sait* que l'espace vectoriel des fonctions paires modulo q admet pour base les fonctions de Ramanujan c_d, avec d un diviseur de q. On a plus précisément

$$f(n) = \sum_{d|q} \frac{a_d}{\varphi(d)} c_d(n),$$

où les coefficients sont

$$a_d = \frac{1}{q}\sum_{n=1}^{q} f(n)c_d(n).$$

*voir le chapitre 2 de l'ouvrage de McCarthy.

Dans le cas qui nous intéresse, on a $a_d = P\mu(d)$. On rappelle que les sommes de Ramanujan sont définies de la façon suivante

$$c_d(n) \coloneqq \sum_{\rho \in \mathcal{R}(d)} \mathrm{e}(n\rho), \tag{4.1}$$

où $\mathcal{R}(d)$ désigne l'ensemble des fractions du tore $\mathbb{Q}/\mathbb{Z}$ (c'est-à-dire avec identification modulo 1) dont le dénominateur est exactement d. Un ensemble de représentants naturel pour $\mathcal{R}(d)$ est $\{a/d\,; 1 \leqslant a \leqslant d, (a,d) = 1\}$. On rappelle également que ces sommes vérifient un nombre important d'identités qui nous seront utiles :

$$c_d(n) = \sum_{t \mid d} \mu(d/t) t [t \mid n] \tag{4.2}$$

$$= \mu\big(\tfrac{d}{(d,n)}\big) \frac{\varphi(d)}{\varphi\big(\frac{d}{(d,n)}\big)}. \tag{4.3}$$

On remarque que dans le cas général, on a $a_1 = P$. On peut alors écrire le moment sous la forme

$$M_k(h;q) = q \sum_{\substack{1 \leqslant m_i \leqslant h \\ (1 \leqslant i \leqslant k)}} \sum_{\substack{1 < d_i \mid q \\ (1 \leqslant i \leqslant k)}} \prod_{i=1}^{k} \frac{a_{d_i}}{\varphi(d_i)} C_{\mathbf{d}}(\mathbf{m}),$$

où

$$C_{\mathbf{d}}(\mathbf{m}) \coloneqq \frac{1}{q} \sum_{n=1}^{q} \prod_{i=1}^{k} c_{d_i}(n + m_i) = \sum_{\substack{\rho_i \in \mathcal{R}(d_i) \\ (1 \leqslant i \leqslant k) \\ \sum \rho_i \equiv 0}} \mathrm{e}(\textstyle\sum m_i \rho_i).^{\dagger} \tag{4.4}$$

La condition de congruence $\sum \rho_i \equiv 0$ se fait modulo $\mathbb{Z}$. Nous omettrons généralement cette précision lorsqu'il n'y a pas de risque de confusion.

Se pose alors naturellement la question d'estimer la somme

$$S_{\mathbf{d}}(h) = \sum_{\substack{1 \leqslant m_i \leqslant h \\ (1 \leqslant i \leqslant k)}} C_{\mathbf{d}}(\mathbf{m}).$$

Comme la fonction $C_{\mathbf{d}}(\mathbf{m})$ peut s'écrire sous la forme de somme d'exponentielles — *cf.* la formule (4.4) —, on peut écrire

$$S_{\mathbf{d}}(h) = \sum_{\substack{\rho_i \in \mathcal{R}(d_i) \\ (1 \leqslant i \leqslant h) \\ \sum \rho_i \equiv 0}} \prod_{i=1}^{k} \frac{\sin \pi h \rho_i}{\sin \pi \rho_i}.$$

[†]Les bornes de sommation entre parenthèses ($1 \leqslant i \leqslant k$) signifie qu'il s'agit d'une somme multiple sur k variables, indexées pour leur définition par un indice i, qui est par conséquent une variable muette hors de la définition de l'ensemble de sommation. Il n'y a donc pas de téléscopage avec d'autres emplois de cet indice i dans une même formule.

On voit ici que le nombre de termes est majoré par $\frac{\prod_{i=1}^k d_i}{d}$ et que chaque terme est majoré par h^k. On obtient ainsi une estimation triviale

$$|S_{\mathbf{d}}(h)| \leqslant \frac{\prod_{i=1}^k d_i}{d} h^k, \tag{4.5}$$

où d est le p.p.c.m. des d_i. Mais cette estimation est bien trop grossière pour parvenir à une estimation non triviale du moment M_k.

4.2. Le lemme fondamental

Pour aboutir à une estimation plus efficace que la majoration (4.5), Montgomery et Vaughan considèrent le problème plus général d'évaluer la somme

$$S_{\mathbf{d}}^{(\rho)} = \sum_{\substack{\rho_i \in \mathcal{C}(d_i) \\ (1 \leqslant i \leqslant h) \\ \sum \rho_i \equiv \rho}} \prod_{i=1}^k G_i(\rho_i),$$

où $\mathcal{C}(d_i)$ est l'ensemble des fractions ρ_i modulo 1 (c'est-à-dire un sous-ensemble de $\mathbb{Q}/\mathbb{Z}$) telles que $d_i\rho_i \in \mathbb{Z}$ — en particulier $\mathcal{R}(d_i) \subset \mathcal{C}(d_i)$ —, où la fonction G_i est une fonction de $\mathcal{C}(d_i)$ à valeurs complexes et où ρ est un élément de $\mathcal{C}(d)$, où d est le p.p.c.m. des d_i. On remarque que l'on obtient la somme $S_{\mathbf{d}}(h)$ en choisissant $\rho \equiv 0$ et $G_i(\rho_i) = \frac{\sin \pi h \rho_i}{\sin \pi \rho_i}[\rho_i \in \mathcal{R}(d_i)]$.

Ici les entiers d_i sont supposés sans facteur carré. Le théorème chinois des restes montre que si s et t sont premiers entre eux, l'application qui au couple $(\sigma, \tau) \in \mathcal{C}(s) \times \mathcal{C}(t)$ associe la fraction $\sigma + \tau \in \mathcal{C}(st)$ définit une bijection de $\mathcal{C}(s) \times \mathcal{C}(t)$ vers $\mathcal{C}(st)$. La restriction de cette application à l'ensemble $\mathcal{R}(s) \times \mathcal{R}(t)$ définit une bijection vers $\mathcal{R}(st)$. Ainsi, pour tout diviseur s de d, on parlera de σ comme de la **composante** dans $\mathcal{C}(s)$ (resp. dans $\mathcal{R}(s)$) de l'élément ρ de $\mathcal{C}(d)$ (resp. de $\mathcal{R}(d)$).
On définit la norme euclidienne $\|\cdot\|_2$ par $\|G_i\|_2 = \big(\sum_{\rho_i \in \mathcal{C}(d_i)} |G_i(\rho_i)|^2\big)^{1/2}$.

Si un entier premier p ne divise qu'un seul des d_i, la condition $\sum \rho_i \equiv \rho$ n'est vérifiée que si les composantes de ρ_i et de ρ dans $\mathcal{C}(p)$ sont égales. Ainsi en translatant la fonction G_i, on peut remplacer les entiers d_i et d respectivement par d_i/p et d/p. On se réduira donc au cas des k-uplets $\mathbf{d}$ qui sont tels que chaque diviseur premier du p.p.c.m. d divise au moins deux des d_i. Il est équivalent d'imposer la condition $d^2 \mid d_1 \cdots d_k$. On parlera dans ce cas de **configurations cohérentes**. Supposer cette condition évite d'effectuer des majorations trop grossières. En outre, lorsque les fonctions G_i sont supportées par les ensembles $\mathcal{R}(d_i)$ et que $\rho \equiv 0$, cette condition de cohérence devient cruciale, puisque la composante en $\mathcal{R}(p)$ de $\rho_i \in \mathcal{R}(d_i)$ ne peut être égale à la composante en $\mathcal{C}(p)$ de ρ, qui est nulle. On obtient donc le lemme suivant.

Lemme 4.2.1. — *Soit un k-uplet $\mathbf{d}$ d'entiers sans facteur carré qui n'est pas en configuration cohérente. Les quantités $C_{\mathbf{d}}(\mathbf{n})$ et $S_{\mathbf{d}}(h)$ sont nulles.*

Le résultat de Montgomery et Vaughan s'énonce ainsi.

Lemme 4.2.2. — *Soit un k-uplet $\mathbf{d}$ d'entiers sans facteur carré en configuration cohérente. On a pour tout $\rho \in \mathcal{C}(d)$*

$$\Big| \sum_{\substack{\rho_i \in \mathcal{C}(d_i) \\ (1 \leqslant i \leqslant k) \\ \sum \rho_i \equiv \rho}} \prod_{i=1}^{k} G_i(\rho_i) \Big| \leqslant \frac{\sqrt{\prod_{i=1}^{k} d_i}}{d} \prod_{i=1}^{k} \|G_i\|_2.$$

On remarque que ce lemme est optimal dans le sens où il peut y avoir égalité. En effet, il y a égalité si k est pair, si $\rho \equiv 0$ et s'il existe un appariement[‡] σ tel que pour tout i, on a $d_i = d_{\sigma(i)}$ et $G_i = \overline{G}_{\sigma(i)}$ et tel que pour toute paire $i \neq j$ vérifiant $\sigma(i) \neq j$ on a $(d_i, d_j) = 1$.

Ces conditions peuvent paraître restrictives, mais dans nos applications les conditions sur ρ et sur les fonctions G_i seront toujours vérifiées ; ne restent alors que les conditions sur le k-uplet d'entiers $\mathbf{d}$. On dira que le k-uplet $\mathbf{d}$ est en **configuration diagonale** si k est pair et s'il existe un appariement σ tel que pour tout i on a $d_i = d_{\sigma(i)}$ et tel que pour toute paire $i \neq j$ vérifiant $\sigma(i) \neq j$ on a $(d_i, d_j) = 1$. En particulier, on a $d^2 = d_1 \cdots d_k$; les configurations diagonales sont donc cohérentes.

Ce lemme fondamental 4.2.2 est obtenu en appliquant par récurrence l'inégalité de Cauchy-Schwarz. Cependant la construction qui permet cette récurrence est subtile. Montgomery et Vaughan en donne deux versions bien différentes[§], exploitant de chacune des propriétés diverses. Nous présentons ici une construction qui a l'avantage de généraliser ces deux preuves, précisant ainsi la compréhension de ces deux constructions, mais qui n'a cependant pas permis de tirer de nouvelles informations sur ce type de somme.

Pour simplifier les notations de cette construction compliquée, nous utiliserons comme ensemble d'indices un ensemble fini I générique plutôt que l'ensemble habituel $\{1, 2, \ldots, k\}$. Le k-uplet $\mathbf{d}$ sera donc une collection d'entiers naturels non nuls sans facteur carré indexée par I. Pour un sous-ensemble J de I, on notera $[\mathbf{d}]_J$ le p.p.c.m. des d_j pour j parcourant J. Et l'on souhaite estimer la somme

$$\sum_{\substack{\rho_i \in \mathcal{C}(d_i) \\ (i \in I) \\ \sum_I \rho_i \equiv \rho}} \prod_{i \in I} G_i(\rho_i).$$

Nous nous placerons dans le cas la collection d'entiers $\mathbf{d}$ est en configuration cohérente. Un bon point de vue pour considérer cette construction est celui des graphes. On associe à une telle collection d'entiers $\mathbf{d}$ un graphe simple $G_{\mathbf{d}}$ dont I est l'ensemble des sommets, et tel que deux sommets sont liés si et seulement les deux entiers d'indices correspondants ne sont pas premiers entre eux. On donne alors un poids à chacune des ces arêtes : l'arête joignant i à j aura pour poids $(d_i, d_j) > 1$.[¶]

[‡] C'est-à-dire une involution sans point fixe de $\mathfrak{S}_k$.

[§] Pour ces deux versions, voir dans [**30**] et dans [**31**].

[¶] On peut également considérer ce même graphe complété par des arêtes de poids 1.

Puisque $\mathbf{d}$ est cohérente, l'entier d_i peut se calculer à partir de ce graphe à poids : l'entier d_i est le p.p.c.m. des poids des arêtes joignant le sommet i.

Les deux constructions à partir de ces graphes qu'il faut garder en tête sont les suivantes :

- Pour un sous-ensemble de sommets J, on peut considérer le sous-graphe de $G_{\mathbf{d}}$ où l'on a conservé les sommets de J et les arêtes joignant deux sommets de J ; cette construction correspond à la notion classique de **sous-graphe induit** par un sous-ensemble de sommets.
- Pour un sous-ensemble de sommets J, on peut également considérer le sous-graphe de $G_{\mathbf{d}}$ où l'on a supprimé les arêtes joignant deux sommets de J et où l'on a contracté les sommets de J en un seul et même sommet, que l'on appellera J. Cependant, cette notion classique de contraction d'un sous-ensemble de sommets n'est bien définie que dans le cas de multigraphe (où deux sommets peuvent être reliés par plusieurs arêtes), alors que nous souhaitons obtenir un graphe simple. Ainsi, nous remplacerons les éventuelles arêtes multiples par des arêtes simples, dont chacune est affectée du poids obtenu en prenant le p.p.c.m. des poids des arêtes qu'elle remplace. On parlera, malgré l'abus de langage, du graphe obtenu par **contraction** d'un sous-ensemble de sommets.

Ainsi à partir d'un sous-ensemble de sommets J, on construit deux sous-graphes de $G_{\mathbf{d}}$; et de la même façon qu'on peut retrouver les valeurs des entiers d_i à partir des poids du graphe $G_{\mathbf{d}}$, on peut associer à chacun de ces deux sous-graphes une collection d'entiers. C'est l'idée derrière la définition des deux « sous-collections » $\mathbf{d}'$ et $\mathbf{d}^*$:

- Pour tout $j \in J$, on pose d'_j comme le produit des diviseurs premiers p de d_j tels qu'il existe un indice j' de J différent de j vérifiant $p \mid d_{j'}$, si bien que $d'_j \coloneqq (d_j, [\mathbf{d}]_{J \smallsetminus \{j\}})$. On pose alors $d' \coloneqq [\mathbf{d}']_J$ et pour chaque $j \in J$ on pose $t_j \coloneqq d_j/d'_j$.
- On définit l'ensemble d'indices $I^* \coloneqq (I \smallsetminus J) \cup \{J\}$. Pour tout $i \in I \smallsetminus J$ on pose $d^*_i \coloneqq d_i$, et on définit d^*_J comme le produit des entiers premiers p tels qu'il existe une paire d'indices $(i, j) \in (I \smallsetminus J) \times J$ vérifiant $p \mid (d_i, d_j)$, si bien que $d^*_J \coloneqq ([\mathbf{d}]_J, [\mathbf{d}]_{I \smallsetminus J})$. On pose alors $d^* \coloneqq [\mathbf{d}^*]_{I^*}$ et $s \coloneqq (d', d^*)$.

Nous allons d'abord réécrire ces définitions d'une autre façon :
Pour chaque diviseur premier p de d, on note I_p l'ensemble des indices i de I tels que $p \mid d_i$. La condition de cohérence de $\mathbf{d}$ se traduit par le fait que chacun de ces ensembles contiennent au moins deux éléments. On peut alors faire les descriptions suivantes :

- Les diviseurs premiers de d^* sont exactement les p tels que $|I_p \smallsetminus J| \geqslant 1$: en effet, il faut et suffit qu'il existe un $i \in I \smallsetminus J$ tel que $p \mid d_i = d^*_i$ pour que $p \mid d^*$.
- Les diviseurs premiers de d^*_J sont exactement les p tels que $|I_p \smallsetminus J| \geqslant 1$ et $|I_p \cap J| \geqslant 1$: c'est la définition même de d^*_J.
- Les diviseurs premiers de d' sont exactement les p tels que $|I_p \cap J| \geqslant 2$: cela provient de la définition des d'_j pour $j \in J$.
- Les diviseurs premiers de s sont exactement les p tels que $|I_p \smallsetminus J| \geqslant 1$ et $|I_p \cap J| \geqslant 2$: comme $s = (d^*, d')$, il faut et suffit que soient vérifiées simultanément les

conditions caractérisant les diviseurs premiers de d' et celles caractérisant les diviseurs premiers de d^*.

Lemme 4.2.3. — *Soient* $\mathbf{d}$ *une collection d'entiers naturels sans facteur carré indexée par l'ensemble* I *en configuration cohérente, et* J *un sous-ensemble de* I.

a). *Les collections d'entiers* $\mathbf{d}'$ *et* $\mathbf{d}^*$ *sont en configuration cohérente.*

b). *Les entiers* t_j*, pour* $j \in J$*, sont premiers entre eux deux à deux ; on note* t *leur produit. On a* $d_J^* = ts$ *et* $d = [d', d^*]$*. En particulier, pour tout* $\rho \in \mathcal{C}(d)$*, il existe* $\rho' \in \mathcal{C}(d')$ *et* $\rho^* \in \mathcal{C}(d^*)$*, définis à* $\mathcal{C}(s)$ *près, tels que* $\rho \equiv \rho' + \rho^*$.

c). *Le graphe* $G_{\mathbf{d}'}$ *est le sous-graphe de* $G_{\mathbf{d}}$ *induit par* J *et le graphe* $G_{\mathbf{d}^*}$ *est le graphe obtenu de* $G_{\mathbf{d}}$ *par contraction de* J.

d). *Soit* $\rho \in \mathcal{C}(d)$*. Soient* $\rho' \in \mathcal{C}(d')$ *et* $\rho^* \in \mathcal{C}(d^*)$ *comme définis dans b). On a*

$$\sum_{\substack{\rho_i \in \mathcal{C}(d_i) \\ (i \in I) \\ \sum_I \rho_i \equiv \rho}} \prod_I G_i(\rho_i) = \sum_{\substack{\rho_i^* \in \mathcal{C}(d_i^*) \\ (i \in I^*) \\ \sum_{I^*} \rho_i^* \equiv \rho^*}} \prod_{i \in I^*} G_i(\rho_i^*),$$

où la fonction G_J *est définie sur* $\mathcal{C}(d_J^*)$ *par*

$$G_J(\rho_J^*) = \sum_{\substack{\rho_j' \in \mathcal{C}(d_j') \\ (j \in J) \\ \sum_J \rho_J' \equiv \sigma + \rho'}} \prod_{j \in J} G_j(\rho_j' + \tau_j),$$

où σ *et* τ_j *sont les composantes de* $\rho_J^* \in \mathcal{C}(d_J^*)$ *respectivement dans* $\mathcal{C}(s)$ *et dans* $\mathcal{C}(t_j)$*, pour tout* $j \in J$.

Les démonstrations de Montgomery et Vaughan correspondent au cas où J est de cardinal 2 (dans [**30**]) et au cas où J est le complémentaire dans I d'un singleton (dans [**31**]).

Démonstration. — On prouve la proposition point par point.

a). De la définition même de la collection $\mathbf{d}'$, on sait que pour tout entier premier p divisant d_j' avec $j \in J$, il existe un indice $j' \in J$, différent de j, tel que $p \mid d_{j'}$. Mais il existe un indice $j'' \in J$, différent de j', tel que $p \mid d_{j''}$, par exemple $j'' = j$, donc p divise $d_{j'}'$. La collection $\mathbf{d}'$ est donc cohérente.

Soient p un entier premier et i un élément de $I \smallsetminus J$ tels que $p \mid d_i^* = d_i$. Puisque $\mathbf{d}$ est cohérente, il existe un $i' \in I$ différent de i, tel que p divise $d_{i'}$: si $i' \in J$, alors p divise d_J^*, sinon, $i' \in I \smallsetminus J$ et $p \mid d_{i'} = d_{i'}^*$. Maintenant, soit p un diviseur premier de d_J^*, alors, par définition de d_J^*, $p \mid [\mathbf{d}]_{I \smallsetminus J}$ donc il existe un $i \in I \smallsetminus J$ tel que $p \mid d_i = d_i^*$. La collection $\mathbf{d}^*$ est donc cohérente.

b). Soient $j \in J$ et p un diviseur premier de t_j. Celui-ci n'est donc pas un diviseur de d_j', il n'existe donc pas de $j' \in J$ différent de j tel que $p \mid d_{j'}$, et donc $p \nmid t_{j'}$. Les t_j, pour $j \in J$ sont donc bien premiers entre eux deux à deux. On pose $t \coloneqq \prod_J t_j = [\mathbf{t}]_J$.
En utilisant les ensembles I_p, on voit que les diviseurs premiers de t sont exactement les p tels que $|I_p \cap J| = 1$: les t_j, qui divisent les d_j correspondants ($|I_p \cap J| \neq 0$), sont premiers entre eux deux à deux ($|I_p \cap J| < 2$). Ils ne sont

donc pas des diviseurs de s (pour lesquels on a $|I_p \cap J| \geqslant 2$). Les entiers s et t sont donc premiers entre eux et les diviseurs premiers de st sont exactement les p tels que $|I_p \cap J| = 1$ et donc $|I_p \smallsetminus J| \geqslant 1$, ou bien tels que $|I_p \cap J| \geqslant 2$ et $|I_p \smallsetminus J| \geqslant 1$, c'est-à-dire ceux tels que $|I_p \cap J| \geqslant 1$ et $|I_p \smallsetminus J| \geqslant 1$, qui sont exactement les diviseurs premiers de d^*_J. Donc $d^*_J = st$. On voit également que pour tous les diviseurs premiers p de d, si I_p n'a pas d'éléments dans $I \smallsetminus J$ (si $p \nmid d^*$), alors I_p a tous ses éléments (c'est-à-dire au moins deux) dans J (donc $p \mid d'$). Cette remarque triviale traduit la relation $d = [d', d^*]$.

c). On va montrer que pour toute paire $i \neq i'$ d'éléments de $I \smallsetminus J$, on a $(d^*_i, d^*_{i'}) = (d_i, d_{i'})$ et que pour tout $i \in I \smallsetminus J$, on a $(d^*_i, d^*_J) = [(d_i, d_j)]_{j\in J}$. Cela prouve que le graphe $G_{\mathbf{d}^*}$ est bien obtenu de $G_{\mathbf{d}}$ par contraction de J, notamment que le sommet i est relié au sommet J dans $G_{\mathbf{d}^*}$ si et seulement si le sommet i est relié dans $G_{\mathbf{d}}$ à un sommet de J. Le poids de cette arête est bien le p.p.c.m. des poids des arêtes de $G_{\mathbf{d}}$ entre i et les sommets de J.
Soient $i \neq i'$ deux éléments de $I \smallsetminus J$, on a $d_i = d^*_i$ et $d_{i'} = d^*_{i'}$, d'où $(d^*_i, d^*_{i'}) = (d_i, d_{i'})$. Soit à présent i un élément de $I \smallsetminus J$, on a $[(d_i, d_j)]_{j\in J} = (d^*_i, [\mathbf{d}]_J)$, or il est clair que les diviseurs premiers p de $[\mathbf{d}]_J$ qui vérifient $I_p \subset J$ n'interviennent pas, donc n'interviennent les diviseurs premiers p de $[d_j]_{j\in J}$ (c'est-à-dire tels que $|I_p \cap J| \geqslant 1$) qui vérifient $|I_p \smallsetminus J| \geqslant 1$, c'est-à-dire les diviseurs premiers de d^*_J, donc $[(d_i, d_j)]_{j\in J} = (d^*_i, d^*_J)$.

On va montrer que pour toute paire $j \neq j'$ d'éléments de J, on a $(d'_j, d'_{j'}) = (d_j, d_{j'})$. Cela prouve que le graphe $G_{\mathbf{d}'}$ est bien le sous-graphe de $G_{\mathbf{d}}$ induit par J. Soient $j \neq j'$ deux éléments de J, on a $d'_j \mid d_j$ et $d'_{j'} \mid d_{j'}$, d'où $(d'_j, d'_{j'}) \mid (d_j, d_{j'})$. Soit en outre un diviseur premier p de $(d_j, d_{j'})$, celui-ci divise alors d'_j et $d'_{j'}$ (*cf.* la définition de la collection $\mathbf{d}'$), et donc également $(d'_j, d'_{j'})$.

d). Soit $\rho \in \mathcal{C}(d)$. Le fait que $d = [d', d^*]$ (voir *b)*) permet de fixer $\rho' \in \mathcal{C}(d')$ et $\rho^* \in \mathcal{C}(d^*)$ tel que $\rho' + \rho^* \equiv \rho$. Soient $\rho_i \in \mathcal{C}(d_i)$ avec $i \in I$ tels que $\sum_I \rho_i \equiv \rho$. Pour $j \in J$, on pose τ_j et ρ'_j les composantes de $\rho_i \in \mathcal{C}(d_i)$ respectivement dans $\mathcal{C}(t_j)$ et dans $\mathcal{C}(d'_j)$, si bien que $\rho_i \equiv \rho'_i + \tau_i$. On a $\rho^* - \sum_{I\smallsetminus J} \rho_i \in \mathcal{C}(d^*)$ et en notant σ sa composante dans $\mathcal{C}(s)$, on a $\rho^* - \sum_{I\smallsetminus J} \rho_i - \sigma \in \mathcal{C}(d^*/s) = \mathcal{C}(d/d')$. En considérant les composantes dans $\mathcal{C}(d/d')$ des termes de l'égalité $\sum_I \rho_i \equiv \rho$, on obtient l'égalité $\sum_{I\smallsetminus J} \rho_i + \sigma + \sum_J \tau_j \equiv \rho^*$ et par soustraction à la première égalité, $\sum_J \rho'_j \equiv \sigma + \rho'$. On peut donc décomposer la somme que l'on souhaite estimer

$$\sum_{\substack{\rho_i \in \mathcal{C}(d_i)\\ (i\in I)\\ \sum_I \rho_i \equiv \rho}} \prod_I G_i(\rho_i)$$

en une succession de sommes sur les différentes composantes σ, τ_j et ρ'_j que l'on a considérées

$$\sum_{\substack{\sigma\in\mathcal{C}(s)\\ \rho_i\in\mathcal{C}(d_i)\\ (i\in I\smallsetminus J)}} \sum_{\substack{\tau_j\in\mathcal{C}(t_j)\\ (j\in J)\\ \sum_{I\smallsetminus J}\rho_i+\sigma+\sum_J\tau_j\equiv\rho^*}} \sum_{\substack{\rho'_j\in\mathcal{C}(d'_j)\\ (j\in J)\\ \sum_J\rho'_j\equiv\sigma+\rho'}} \prod_{i\in I\smallsetminus J} G_i(\rho_i) \prod_{j\in J} G_j(\rho'_j + \tau_j).$$

Le *b)* permet de remplacer la somme sur $\sigma \in \mathcal{C}(s)$ et $\tau_j \in \mathcal{C}(t_j)$ pour $j \in J$ par une somme sur $\rho_J^* \in \mathcal{C}(d_J^*)$. En isolant le terme dépendant de ρ_J^*, on obtient bien la forme voulue. □

Il ne reste plus qu'à exploiter cette construction pour utiliser l'inégalité de Cauchy-Schwarz par récurrence.

Démonstration du lemme 4.2.2. — Pour $k = 1$, la condition de cohérence impose les égalités $d_1 = d = 1$, et la majoration est triviale.
Pour $k = 2$, la condition de cohérence impose $d_1 = d_2 = d$, et pour tout $\rho \in \mathcal{C}(d)$, l'inégalité de Cauchy-Schwarz fournit

$$\Big| \sum_{\rho_1 \in \mathcal{C}(d)} G_1(\rho_1) G_2(\rho - \rho_1) \Big| \leqslant \Big(\sum_{\rho_1 \in \mathcal{C}(d_1)} \big|G_1(\rho_1)\big|^2 \Big)^{1/2} \Big(\sum_{\rho_2 \in \mathcal{C}(d_2)} \big|G_2(\rho_2)\big|^2 \Big)^{1/2}$$

qui est bien égal à $\frac{\sqrt{d_1 d_2}}{d} \|G_1\|_2 \|G_2\|_2$.
Pour $k \geqslant 3$, on suppose que le lemme est vérifié pour toute collection cohérente $\mathbf{d}$ comprenant strictement moins de k entiers. On choisit de façon arbitraire un sous-ensemble J de l'ensemble $I = \{1, 2, \ldots, k\}$ dont le cardinal est compris entre 2 et $k-1$ (on remarque que $k \geqslant 3$). On construit à partir de cet ensemble les deux collections d'entiers $\mathbf{d}'$ et $\mathbf{d}^*$ respectivement indexés par les ensembles J et I^*, dont les cardinaux sont compris entre 2 et $k-1$. On peut donc appliquer l'hypothèse de récurrence à la collection $\mathbf{d}^*$, qui est cohérente par le *a)* du lemme 4.2.3, et on obtient grâce au point *d)*

$$\Big| \sum_{\substack{\rho_i \in \mathcal{C}(d_i) \\ (i \in I) \\ \sum_I \rho_i \equiv \rho}} \prod_I G_i(\rho_i) \Big| = \Big| \sum_{\substack{\rho_i^* \in \mathcal{C}(d_i^*) \\ (i \in I^*) \\ \sum_{I^*} \equiv \rho^*}} \prod_{i \in I^*} G_i(\rho_i^*) \Big| \leqslant \frac{\sqrt{d_J^* \prod_{i \in I \smallsetminus J} d_i}}{d^*} \|G_J\|_2 \prod_{i \in I \smallsetminus J} \|G_i\|_2.$$

On souhaite estimer $\|G_J\|_2$. Comme la collection $\mathbf{d}'$ est cohérente par le *a)* du lemme 4.2.3, on utilise l'hypothèse de récurrence pour majorer $|G_J^*(\rho_J^*)|$ pour chaque $\rho_J^* \in \mathcal{C}(d_J^*)$

$$\Big| \sum_{\substack{\rho'_j \in \mathcal{C}(d'_j) \\ (j \in J) \\ \sum_J \rho'_J \equiv \sigma + \rho'}} \prod_{j \in J} G_j(\rho'_j + \tau_j) \Big| \leqslant \frac{\sqrt{\prod_{j \in J} d'_j}}{d'} \prod_{j \in J} \big\| G_j(\,\cdot + \tau_j) \big\|_2,$$

où les fonctions $G_j(\,\cdot + \tau_j)$ sont définies sur $\mathcal{C}(d'_j)$ et où, à l'aide du point *c)* du lemme 4.2.3, les fractions σ et τ_j sont les composantes de ρ_J^* respectivement dans $\mathcal{C}(s)$ et dans $\mathcal{C}(t_j)$, pour chaque $j \in J$, et donc où l'on a $\rho_J^* \equiv \sigma + \sum_J \tau_j$. On a donc une

majoration de $\|G_J\|_2{}^2$

$$\sum_{\rho_J^*\in\mathcal{C}(d_J^*)}\big|G_J^*(\rho_J^*)\big|^2 \leqslant \sum_{\sigma\in\mathcal{C}(s)}\sum_{\substack{\tau_j\in\mathcal{C}(t_j)\\(j\in J)}}\Big(\frac{\prod_{j\in J}d'_j}{(d')^2}\prod_{j\in J}\Big(\sum_{\rho'_j\in\mathcal{C}(d'_j)}\big|G_j(\rho'_j+\tau_j)\big|^2\Big)\Big)$$
$$= s\frac{\prod_{j\in J}d'_j}{(d')^2}\prod_{j\in J}\Big(\sum_{\tau_j\in\mathcal{C}(t_j)}\sum_{\rho'_j\in\mathcal{C}(d'_j)}\big|G_j(\rho'_j+\tau_j)\big|^2\Big).$$

Comme $d_j = d'_j t_j$, on obtient $\|G_J\|_2 \leqslant \frac{\sqrt{s\prod_J d'_j}}{d'}\prod_J\|G_j\|_2$. En regroupant ces informations, on trouve

$$\Big|\sum_{\substack{\rho_i\in\mathcal{C}(d_i)\\(i\in I)\\\sum_I\rho_i\equiv\rho}}\prod_I G_i(\rho_i)\Big| \leqslant \frac{\sqrt{sd_J^*\prod_{i\in I\smallsetminus J}d_i\prod_{j\in J}d'_j}}{d^*d'}\prod_{i\in I}\|G_i\|_2.$$

Par le point *b)* du lemme 4.2.3, on a $d_J^* = st = s\prod_J t_j$; et comme $d_j = d'_j t_j$, on a $sd_J^*\prod_{I\smallsetminus J}d_i\prod_J d'_j = s^2\prod_I d_i$. En outre, on a $d'd^* = (d',d^*)[d',d^*] = sd$, toujours à l'aide du point *b)* du lemme 4.2.3. On obtient la majoration annoncée. □

Cette analyse permet d'observer que le facteur inattendu $\sqrt{\prod d_i}/d$ du lemme 4.2.2 est la racine carrée du produit des différents facteurs s qui apparaissent dans chaque opération d'induction-contraction, et ce quel que soit le choix des sous-ensembles J successifs.

Lors d'une opération d'induction-contraction, le facteur s est le produit des entiers premiers qui se trouvent simultanément dans les deux sous-graphes construits. Chacun de ces facteurs premiers possède une multiplicité parmi les deux collections $\mathbf{d}'$ et $\mathbf{d}^*$ strictement supérieure à celle observée dans la collection $\mathbf{d}$. En effet, on a

$$s = \frac{\prod_{j\in J}d'_j\cdot\prod_{i\in I^*}d_i^*}{\prod_{i\in I}d_i}.$$

Ce facteur s est donc le coût de cette opération d'induction-contraction, l'artifice qui permet de séparer l'estimation d'une grosse somme en deux sommes plus petites. On peut donc considérer que la majoration conjecturale

$$\Big|\sum_{\substack{\rho_i\in\mathcal{C}(d_i)\\(1\leqslant i\leqslant h)\\\sum\rho_i\equiv\rho}}\prod_{i=1}^k G_i(\rho_i)\Big| \leqslant \prod_{i=1}^k\|G_i\|_2 \tag{4.6}$$

n'est pas complètement infondée. On remarque que, pareillement à l'inégalité du lemme 4.2.2, cette majoration est une égalité dans le cas de configuration diagonale.

4.3. Estimation de $S_{\mathbf{d}}(h)$

Une application directe du lemme fondamental 4.2.2 fournit

$$|S_{\mathbf{d}}(h)| \leqslant \frac{\sqrt{\prod_{i=1}^k d_i}}{d} \prod_{i=1}^k \Big(\sum_{\rho_i \in \mathcal{R}(d_i)} \Big(\frac{\sin \pi h \rho_i}{\sin \pi \rho_i} \Big)^2 \Big)^{1/2}. \tag{4.7}$$

On sait préciser un peu le terme de ce produit.

Lemme 4.3.1. — *Soient d et h deux entiers naturels.*

a). *On a*

$$\sum_{\rho \in \mathcal{R}(d)} \Big(\frac{\sin \pi h \rho}{\sin \pi \rho} \Big)^2 \leqslant 2dh$$

b). *On a*

$$\sum_{\rho \in \mathcal{R}(d)} \Big(\frac{\sin \pi h \rho}{\sin \pi \rho} \Big)^2 = \sum_{1 \leqslant m,m' \leqslant h} c_d(m-m') = \varphi(d)h + O(h^3).$$

c). *On a pour $d > 1$*

$$\sum_{\rho \in \mathcal{R}(d)} \Big(\frac{\sin \pi h \rho}{\sin \pi \rho} \Big)^2 = -\sum_{t \mid d} \mu(d/t) t^2 \left\{ \frac{h}{t} \right\} \Big(1 - \left\{ \frac{h}{t} \right\} \Big) \leqslant \frac{15}{4\pi^2} d^2.$$

Le point *a)* donne la majoration que nous allons principalement utiliser, le point *b)* montre que celle-ci n'est pas trop loin de la vérité lorsque $d \gg h^{2+\varepsilon}$ et le point *c)* permet d'améliorer cette majoration pour d petit par rapport à h.

Démonstration. — On remarque que $|\sin \pi h\rho / \sin \pi \rho| \leqslant h$ et que $2\|\rho\| \leqslant |\sin \pi \rho| \leqslant 1$, où $\| \cdot \|$ désigne la distance à l'entier le plus proche. Ainsi en définissant la fonction paire et 1-périodique $E_h(\rho) \coloneqq \min\big(h, 1/2\|\rho\|\big)$, on a

$$\sum_{\rho \in \mathcal{R}(d)} \Big(\frac{\sin \pi h \rho}{\sin \pi \rho} \Big)^2 \leqslant \sum_{\rho \in \mathcal{R}(d)} E_h(\rho)^2 \leqslant 2 \sum_{n=1}^{\lfloor d/2 \rfloor} E_h(n/d)^2.$$

Comme la fonction E_h est décroissante sur $[0, 1/2]$, on a

$$\frac{1}{d} \sum_{n=1}^{\lfloor d/2 \rfloor} E_h(n/d)^2 \leqslant \int_0^{1/2} E_h(\rho)^2 \, \mathrm{d}\rho \leqslant \int_0^{1/2h} h^2 \, \mathrm{d}\rho + \int_{1/2h}^{1/2} \frac{\mathrm{d}\rho}{4\rho^2} \leqslant h.$$

On utilise le fait que $|\sum_{m=1}^h \mathrm{e}(m\rho)| = |\sin \pi h\rho \,/\, \sin \pi \rho|$ et l'identité (4.1) pour déduire

$$\sum_{\rho \in \mathcal{R}(d)} \Big(\frac{\sin \pi h \rho}{\sin \pi \rho} \Big)^2 = \sum_{\rho \in \mathcal{R}(d)} \sum_{1 \leqslant m,m' \leqslant h} \mathrm{e}\big((m-m')\rho\big) = \sum_{1 \leqslant m,m' \leqslant h} c_d(m-m').$$

La contribution des termes avec $m = m'$ est exactement $h\varphi(d)$. Pour les autres termes (au plus h^2), on utilise l'identité (4.3) pour déduire que $|c_d(n)| \leqslant |n|$ pour $n \neq 0$. Ainsi le terme d'erreur est bien $O(h^3)$.

Enfin, en utilisant l'équation (4.2) on obtient

$$\sum_{\rho\in\mathcal{R}(d)}\Big(\frac{\sin\pi h\rho}{\sin\pi\rho}\Big)^2=\sum_{1\leqslant m,m'\leqslant h}c_d(m-m')=\sum_{t|d}\mu(d/t)t\sum_{1\leqslant m,m'\leqslant h}[t\mid m-m'].$$

La dernière somme est facile à calculer et vaut $h^2/t-t\{h/t\}\big(1-\{h/t\}\big)$. Le premier terme fournit dans la somme un terme $h^2\delta(d)$, nul lorsque $d>1$, et on obtient l'identité annoncée. On remarque que l'on a $0\leqslant\{x\}\big(1-\{x\}\big)\leqslant 1/4$, donc

$$\sum_{\rho\in\mathcal{R}(d)}\Big(\frac{\sin\pi h\rho}{\sin\pi\rho}\Big)^2=-\sum_{t|d}\mu(d/t)t^2\left\{\frac{h}{t}\right\}\left(1-\left\{\frac{h}{t}\right\}\right)\leqslant\frac{1}{4}\sum_{t|d}t^2\leqslant\frac{1}{4}d^2\frac{\zeta(2)}{\zeta(4)}.\qquad\square$$

On utilise à présent des estimations en moyennes de sommes de Ramanujan pour améliorer la majoration (4.7) et donner du crédit à la forme conjecturale (4.6) du lemme fondamental. Nous avons réalisé une étude plus fine des moyennes de sommes de Ramanujan, mais ces estimations ne conduisant pas à des résultats sensiblement meilleurs, nous avons regroupés ces travaux dans l'appendice B.

Lemme 4.3.2. — *Soit h un entier et $\mathbf{d}$ un k-uplet cohérent d'entiers. On a*

$$\big|S_{\mathbf{d}}(h)\big|\leqslant\prod_{i=1}^{k}\sigma(d_i).$$

Dans certains cas, ce lemme est essentiellement meilleur que le lemme fondamental 4.2.2 associé aux estimations du lemme 4.3.1 qui fournissent le résultat suivant.

Lemme 4.3.3. — *Soit h un entier et $\mathbf{d}$ un k-uplet cohérent d'entiers. On a*

$$\big|S_{\mathbf{d}}(h)\big|\leqslant c^k\frac{\prod_{i=1}^{k}d_i}{d}\sqrt{\prod_{i=1}^{k}\min(d_i,h)},$$

où la constante c est absolue.

En effet, si h est supérieur aux d_i, ou bien si $d\leqslant h^{k/2}$, le majorant est supérieur à $\prod_i d_i$, qui est essentiellement le majorant obtenu dans ce lemme. Ce résultat représente aussi un élément de preuve en faveur de l'estimation (4.6). Dans cet optique, il serait intéressant d'obtenir également un majorant de $S_{\mathbf{d}}(h)$ de la forme $\big(\prod_i d_i\big)^{1/2}h^{k/2}$.

Démonstration du lemme 4.3.2. — Par l'identité (4.4), on a

$$S_{\mathbf{d}}(h)=\frac{1}{q}\sum_{n=1}^{q}\prod_{i=1}^{k}\left(\sum_{m_i=1}^{h}c_{d_i}(n+m_i)\right)\leqslant\prod_{i=1}^{k}A(d_i).$$

On utilise la proposition B.1.1.*a)* pour obtenir la majoration annoncée. $\square$

En utilisant le théorème B.3.3, si l'on a $P^+(d_i)\leqslant h$ pour tout i, on obtient l'estimation qui peut parfois se révéler légèrement plus précise

$$\big|S_{\mathbf{d}}(h)\big|\leqslant c^k\prod_{i=1}^{k}d_i(\log h)^{k/2},$$

où la constante c est absolue.

CHAPITRE 5

ÉTAT DES LIEUX

Au cours des chapitres précédents, on a remarqué que l'heuristique basée sur le fait que les événements « n est premier à q » et « n' est premier à q » sont des évenements presque indépendants lorsque n et n' sont des entiers assez proches fournit de bonnes estimations uniformes du moment $M_k(h;q)$ pour des entiers q sans petit facteur premier, mais n'est pas vérifiée si q possède de petits facteurs premiers. Les progrès de la méthode harmonique sont trop minces pour espérer obtenir de bonnes estimations uniformes du moment $M_k(h;q)$ lorsque q ne possède que des petits facteurs premiers.

Dans ce paragraphe, nous listons les avancées obtenues en direction de la conjecture 5.1.2 *infra*, pour mieux cerner les cas qu'il reste à traiter. Nous analysons les limites de la méthode employée mais également les raisons de croire en notre majoration conjecturale.

5.1. Résumé des épisodes précédents

Le dessein de cette étude est de démontrer à terme la conjecture suivante.

Conjecture 5.1.1. — Pour tout entier naturel n, on a

$$g(n) \ll \frac{\varphi(n)}{n} \log n.$$

Nous avons montré de façon détaillée dans la proposition 1.2.3 que cette conjecture pouvait se déduire de cette seconde conjecture.

Conjecture 5.1.2. — Il existe une constante absolue $C > 0$ telle qu'on a

$$M_k(h;q) \ll C^k q k^{k/2} (k + hP)^{k/2},$$

uniformément pour $h \geqslant 1$, $q \geqslant 1$ et $k \geqslant 1$.

On peut noter qu'une spécialisation de $h = g(q) - 1$ et $k = \frac{1}{eC^2} hP$ fournit bien immédiatement $g(q) \leqslant 2eC^2 \frac{\varphi(q)}{q} \big(\log q + O(1)\big)$.

Le travail fait au cours des précédents chapitres permet de se restreindre à vérifier la conjecture sous des conditions plus agréables.

Proposition 5.1.1. — *On suppose qu'il existe une constante absolue $C > 0$ telle qu'on ait*

$$M_k(h;q) \ll C^k q(k\,hP)^{k/2},$$

uniformément pour $h \geqslant 1$, q vérifiant $P^+(q) \leqslant h$ et $k \leqslant hP$. Les conjectures 5.1.1 et 5.1.2 sont alors vraies.

Démonstration. — Soient h, q et k trois entiers naturels. Il est clair que $M_k(h;q) = q/\bar{q}\; M_k(h;\bar{q})$, où l'on a noté $\bar{q}$ le noyau sans facteur carré de q. On suppose donc q sans facteur carré. On pose $q_\sharp$ le produit des facteurs premiers de q qui sont strictement inférieurs à h, et $q_\flat$ le produit de ceux qui sont supérieurs à h. Il est clair que $q = q_\sharp q_\flat$ et que $(q_\sharp, q_\flat) = 1$. La suite des entiers premiers à q est donc décomposable en la suite des entiers premiers à $q_\sharp$ et de celle des entiers premiers à $q_\flat$. Puisque l'on a $P^+(q_\sharp) < h$, on sait qu'il existe une constante $c_\sharp \geqslant 1$ telle que tout intervalle de longueur h contient moins de $c_\sharp h\varphi(q_\sharp)/q_\sharp$ entiers premiers à $q_\sharp$. Il s'agit ici d'une application immédiate du crible combinatoire*. En posant $P_\sharp = \varphi(q_\sharp)/q_\sharp$, on a

$$M_k(h;q_\sharp) \leqslant {c_\sharp}^k q_\sharp (hP_\sharp)^k. \tag{5.1}$$

En adjoignant ce résultat à la supposition faite dans l'énoncé, on prouve l'existence d'une constante $C_\sharp$ telle que

$$M_k(h;q_\sharp) \leqslant {C_\sharp}^k q_\sharp \min(k, hP_\sharp)^{k/2}(hP_\sharp)^{k/2}.$$

D'autre part, les conditions du théorème 3.1.8 sont vérifiées par $q_\flat$ puisque $P^-(q_\flat) \geqslant h$. Il existe donc une constante $C_\flat$ telle que

$$M_k(h;q_\flat) \leqslant {C_\flat}^k q_\flat k^{k/2}(k + hP_\flat)^{k/2}.$$

De plus, la proposition 3.1.2 s'applique. La suite des entiers premiers à $q_\flat$ remplit les conditions de la proposition 1.3.4. On remarque que l'inégalité (5.1) permet de remplacer avantageusement les coefficient « $c_k(l, \mathbf{a}^{(1)})$ » par $C_\sharp$ dans la majoration de la proposition 1.3.4. On a donc

$$\begin{aligned} M_k(h,q) \leqslant & (2C_\sharp)^k q {P_\flat}^{k/2} \min(k, hP_\sharp)^{k/2}(hP)^{k/2} + (4C_\flat)^k q k^{k/2}(k+hP)^{k/2} \\ & (4C_\flat C_\sharp)^k q k^{k/2}(k+hP)^{k/2} \\ \leqslant & C^k q k^{k/2}(k+hP)^{k/2}, \end{aligned}$$

où $C = 2C_\sharp + 4C_\flat + 4C_\flat C_\sharp$, ce qui établit le théorème. □

Remarque. — On peut instruire le procès de l'hypothèse formulée dans la proposition 5.1.1. La charge la plus sérieuse est l'œuvre de Granville et Soundararajan [**14**], qui ont montré qu'on ne peut pas disposer d'une bonne majoration de $M_k(h;q)$ de la forme $C^k q(k\,hP)^{k/2}$ si k est trop grand. En effet, on peut perturber légèrement

*Comme indiqué dans l'ouvrage d'Halberstam et Richert, cette phrase sous-entend l'utilisation du théorème 2.2 de H. Halberstam & H.-E. Richert, *Sieve Methods,* Academic Press, Londres, 1974

la suite **a** des entiers premiers à q pour la rendre équirépartie sur les petits intervalles, et quantifier cette déviation à l'aide de $M_k(h;q)$. Si ce moment n'est pas trop grand, la nouvelle suite préserve les propriétés de **a** d'équirépartition dans les progressions arithmétiques. Le principe d'incertitude de Granville et Soundararajan, basée sur la méthode de la matrice de Maier et sur les oscillations en moyenne de fonctions multiplicatives, implique qu'une suite donnée ne peut pas être bien équirépartie simultanément dans les petits intervalles et dans les progressions arithmétiques. Il serait intéressant de voir à quel point ces arguments permettraient de fournir une minoration uniforme de $M_k(h;q)$.

Il reste donc à établir la majoration uniforme dans les cas où les conditions suivantes sont vérifiées :
- l'entier q est sans facteur carré et tous ses diviseurs premiers sont strictement inférieurs à h ;
- l'entier k est inférieur à hP.

On se place à présent sous ces conditions. On peut utiliser le résultat du lemme 4.3.2 pour s'affranchir des d_i trop petits. Malheureusement, les améliorations obtenues sont maigres. De plus, l'on verra que le point faible de cette méthode réside dans le fait que les estimations pour les « grands » d_i ne sont pas assez précises pour gérer leur grand nombre.

Proposition 5.1.2. — *On a uniformément pour $h \geqslant 1$, q sans facteur carré et k pair*

$$M_k(h;q) \leqslant 2^k q k^{k/2} (hP)^{k/2} + 2^k q P^k \sum_{\substack{\sqrt{khP^3} < d_i \mid q \\ (1 \leqslant i \leqslant k)}} \prod_{i=1}^{k} \frac{\mu(d_i)}{\varphi(d_i)} S_{\mathbf{d}}(h)$$

Démonstration. — On a en utilisant la décomposition en sommes de Ramanujan de la fonction paire modulo q, caractéristique de l'ensemble des entiers premiers à q,

$$M_k(h;q) = P^k \sum_{n=1}^{q} \Big(\sum_{\substack{1 < d \mid q \\ d \leqslant \sqrt{khP^3}}} \sum_{m=1}^{q} \frac{\mu(d)}{\varphi(d)} c_d(n+m) + \sum_{\sqrt{khP^3} < d \mid q} \sum_{m=1}^{q} \frac{\mu(d)}{\varphi(d)} c_d(n+m) \Big)^k,$$

et l'on applique le lemme trivial 1.3.1 pour obtenir

$$M_k(h;q) \leqslant 2^k P^k \sum_{n=1}^{q} \Big(\sum_{\substack{1 < d \mid q \\ d \leqslant \sqrt{khP^3}}} \sum_{m=1}^{q} \frac{\mu(d)}{\varphi(d)} c_d(n+m) \Big)^k + 2^k P^k \sum_{n=1}^{q} \Big(\sum_{\sqrt{khP^3} < d \mid q} \sum_{m=1}^{q} \frac{\mu(d)}{\varphi(d)} c_d(n+m) \Big)^k.$$

En menant un calcul similaire à celui établissant l'identité (5.3), on trouve que le majorant vaut exactement

$$2^k qP^k \sum_{\substack{1<d_i|q\\ d_i\leqslant\sqrt{khP^3}\\ (1\leqslant i\leqslant k)}} \prod_{i=1}^{k} \frac{\mu(d_i)}{\varphi(d_i)} S_{\mathbf{d}}(h) + 2^k qP^k \sum_{\substack{\sqrt{khP^3}<d_i|q\\ (1\leqslant i\leqslant k)}} \prod_{i=1}^{k} \frac{\mu(d_i)}{\varphi(d_i)} S_{\mathbf{d}}(h).$$

On utilise alors le lemme 4.3.2 dans le terme de gauche, qui est alors majoré par

$$2^k qP^k \sum_{\substack{1<d_i|q\\ d_i\leqslant\sqrt{khP^3}\\ (1\leqslant i\leqslant k)}} \prod_{i=1}^{k} \frac{\sigma(d_i)}{\varphi(d_i)} \leqslant 2^k qP^{-k}(khP^3)^{k/2}.$$

On a utilisé ici le fait que $\sigma(d)/\varphi(d) \leqslant P^{-2}$ et l'on a pas tenu compte des conditions de cohérence, qui comme on l'a vu, sont généralement peu restrictives. □

5.2. Explosion combinatoire

On tire du lemme 4.3.1 l'estimation suivante

$$\big|S_{\mathbf{d}}(h)\big| \leqslant 2^k \frac{\prod_{i=1}^{k} d_i}{d} h^{k/2}, \tag{5.2}$$

que l'on peut comparer à la majoration (4.5) et qui une fois replacée dans l'expression

$$M_k(h;q) = qP^k \sum_{\substack{1<d_i|q\\ (1\leqslant i\leqslant k)}} \prod_{i=1}^{k} \frac{\mu(d_i)}{\varphi(d_i)} S_{\mathbf{d}}(h), \tag{5.3}$$

fournit l'estimation

$$M_k(h;q) \leqslant 2^k q(hP)^{k/2} P^{k/2-2^k}.$$

Cette estimation est très utile à k fixé, lorsque l'on ne s'intéresse pas à une quelconque uniformité en k. En effet, puisque les facteurs premiers de q ont été supposés inférieurs à h, on a $P^{-1} \ll \log h$. Ainsi pour tout $\varepsilon > 0$, on a

$$M_k(h;q) \ll_{k,\varepsilon} q(hP)^{k/2+\varepsilon}.$$

Le tour de force de Montgomery et Vaughan [**30**] et dans un deuxième temps de Montgomery et Soundararajan [**29**] est d'avoir précisé cette majoration en retirant le ε dans un premier temps, puis en identifiant la constante dépendant de k.

Théorème 5.2.1 **(Montgomery & Soundararajan)**. — *Soit k un entier fixé. On a pour $hP \to \infty$*

$$M_k(h;q) = \big(k!! + o(1)\big)\, q\big(\tfrac{1}{q} M_2(h;q)\big)^{k/2}.$$

On note que $M_2(h;q) \leqslant qhP$.[†]

Ce résultat est évidemment à rapprocher de sa version heuristique, la proposition 2.2.6. Cependant, si ce théorème va dans le sens de la conjecture 5.1.2, il n'aide en rien à sa résolution. En effet, tapi dans le terme d'erreur, le terme P^{-2^k} annihile nos espoirs d'obtenir l'uniformité souhaitée en la variable k.

Pour obtenir ce résultat, Montgomery et Vaughan montrent qu'il est possible d'améliorer très légèrement la majoration (5.2) lorsque le k-uplet $\mathbf{d}$ est « loin » d'être en configuration diagonale. Leur résultat est trop complexe et trop peu explicite pour être précisé ici. Il faut retenir de leur résultat la philosophie générale : les k-uplets en configuration diagonale contribuent essentiellement au terme principal dans l'évaluation de $M_k(h;q)$.

Lemme 5.2.2. — *Soit k un entier pair. La contribution des k-uplets en configuration diagonale dans la somme* (5.3) *est inférieure à*

$$k!!\, q\big(\tfrac{1}{q}M_2(h;q)\big)^{k/2}.$$

Démonstration. — Pour un k-uplet $\mathbf{d}$ en configuration diagonale, on a

$$S_{\mathbf{d}}(h) = \frac{\sqrt{\prod_{i=1}^k d_i}}{d} \prod_{i=1}^k \Big(\sum_{\rho_i \in \mathcal{R}(d_i)} \Big(\frac{\sin \pi h\rho_i}{\sin \pi \rho_i}\Big)^2 \Big)^{1/2}.$$

Il est équivalent de se donner un k-uplet $\mathbf{d}$ en configuration diagonale avec $1 < d_i \mid q$ ou de se donner un appariement σ et $k/2$ entier $t_{\{i,\sigma(i)\}} > 1$, où $\{i,\sigma(i)\}$ parcourt l'ensemble $\mathcal{O}_\sigma$ des orbites de σ, avec $\prod_{I\in\mathcal{O}_\sigma} t_I \mid q$. On obtient donc

$$S_{\mathbf{d}}(h) = \prod_{I\in\mathcal{O}_\sigma} \Big(\sum_{\rho_I \in \mathcal{R}(t_I)} \Big(\frac{\sin \pi h\rho_I}{\sin \pi \rho_I}\Big)^2 \Big).$$

Ainsi, la contribution des k-uplets en configuration diagonale dans la somme (5.3) vaut exactement

$$qP^k \sum_{\sigma \text{appariement}} \sum_{\substack{1<t_I\mid q \\ (I\in\mathcal{O}_\sigma) \\ \prod_I t_I \mid q}} \prod_{I\in\mathcal{O}_\sigma} \Big(\frac{1}{\varphi(t_I)^2} \sum_{\rho_I\in\mathcal{R}(t_I)} \Big(\frac{\sin \pi h\rho_I}{\sin \pi \rho_I}\Big)^2 \Big)$$

$$\leqslant k!!\, q \Big(P^2 \sum_{1<t\mid q} \frac{1}{\varphi(t)^2} \sum_{\rho\in\mathcal{R}(t)} \Big(\frac{\sin \pi h\rho}{\sin \pi \rho}\Big)^2 \Big)^{k/2},$$

ce qui conclut la preuve grâce à l'identité (5.3). □

On s'attend en réalité à ce que ces termes diagonaux soient les principaux responsables de ce terme principal ; les autres configurations seraient essentiellement responsables du terme de reste. C'est la philosophie sous-jacente aux résultats de Montgomery et Vaughan et de Montgomery et Soundararajan.

[†]Voir l'article de Hausman et Shapiro [**17**].

Reste malgré tout dans leur termes de reste ce P^{-2^k}, qui certes peut être compensé par une petite puissance de h quand $h \gg_k 1$ (car $P^{-1} \ll \log h$ dans notre cas), mais qui empêche d'obtenir l'uniformité tant espérée. À défaut d'éliminer ce terme nous allons essayer de comprendre le pourquoi de sa présence. En utilisant l'estimation (5.2) dans l'expression (5.3) on obtient un majorant

$$\big|M_k(h;q)\big| \leqslant qP^k \sum_{\substack{1<d_i|q\\(1\leqslant i\leqslant k)\\ \mathbf{d}\text{ cohérent}}} \frac{1}{d} h^{k/2} \prod_{i=1}^{k} \frac{d_i}{\varphi(d_i)}.$$

On peut majorer $d_i/\varphi(d_i)$ par P^{-1} sans faire trop de sacrifices. De façon générale, nous n'essaierons pas d'épargner des puissances de P^{-1} ou de $\log h$, tant que l'exposant est linéaire en k. On obtient ainsi un majorant de nature combinatoire

$$\big|M_k(h;q)\big| \leqslant qh^{k/2} \sum_{1<d|q} \frac{1}{d} \operatorname{card}\big\{\mathbf{d} \in \mathbb{N}^k\,;\, \forall i, 1<d_i \mid d\,;\, d^2 \mid \textstyle\prod d_i\,;\, [d_1,\dots,d_k]=d\big\}.$$

Par le principe d'inclusion-exclusion on peut se libérer de la condition $d_i > 1$ pour obtenir

$$qh^{k/2} \sum_{s=0}^{k} \binom{k}{s} (-1)^{k-s} \sum_{d|q} \frac{1}{d} \operatorname{card}\big\{\mathbf{d} \in \mathbb{N}^s\,;\, \forall i, d_i \mid d\,;\, d^2 \mid \textstyle\prod d_i\,;\, [d_1,\dots,d_s]=d\big\}.$$

Mais à présent, chaque élément de $\big\{\mathbf{d} \in \mathbb{N}^s\,;\, \forall i, d_i \mid d\,;\, d^2 \mid \prod d_i\,;\, [d_1,\dots,d_s]=d\big\}$ est exactement la donnée pour chaque diviseur premier p de d d'un sous-ensemble de $\{1,\dots,s\}$ possédant au moins deux éléments. Ainsi, on a

$$\operatorname{card}\big\{\mathbf{d} \in \mathbb{N}^s\,;\, \forall i, d_i \mid d\,;\, d^2 \mid \textstyle\prod d_i\,;\, [d_1,\dots,d_s]=d\big\} = (2^s - s - 1)^{\omega(d)}.$$

Le majorant ainsi obtenu devient

$$\big|M_k(h;q)\big| \leqslant qh^{k/2} \sum_{s=0}^{k} \binom{k}{s} (-1)^{k-s} \prod_{p|q} \Big(1 + \frac{2^s - s - 1}{p}\Big).$$

Il est maintenant évident que le terme dominant dans cette somme est celui pour lequel $s = k$. On obtient donc bien

$$\big|M_k(h;q)\big| \leqslant qh^{k/2} P^k P^{-2^k}.$$

On aurait obtenu également ce P^{-2^k} si nous avions directement majorer le cardinal de $\big\{\mathbf{d} \in \mathbb{N}^k\,;\, \forall i, 1 < d_i \mid d\,;\, d^2 \mid \prod d_i\,;\, [d_1,\dots,d_k] = d\big\}$ par le cardinal de l'ensemble $\big\{\mathbf{d} \in \mathbb{N}^k\,;\, \forall i, d_i \mid d\big\}$. C'est dans ce sens que l'on prétend qu'il y a explosion combinatoire : le nombre d'objets considérés croît tout simplement trop vite.

Cependant, les raisons de croire à cette conjecture 5.1.2 sont nombreuses. En plus des arguments apportés, on peut mettre cette conjecture dans un contexte plus général. Les travaux de Montgomery et Soundararajan [**29**][‡] relient les comportements

[‡]Nous tirons de cet article le contenu de ce paragraphe.

asymptotiques de $M_k(h;q)$ et des moments $\sum_{n=0}^{N}\big(\psi(n+h)-\psi(n)-h\big)^k$ liés à la répartition des entiers premiers dans un intervalle de longueur h à l'aide d'une version uniforme et optimiste de la conjecture des k-uplets de Hardy et Littlewood. Sous l'hypothèse de Riemann, il est connu[§] qu'une bonne estimation du moment d'ordre 2 est équivalent à la conjecture de corrélation des paires énoncée par Montgomery[¶]. Cet argument a été généralisé par Chan[‖] aux ordres supérieurs, montrant ainsi une relation avec l'estimation asymptotique des moments $\int_1^X\big(\sum_{\zeta(1/2+i\gamma)=0}^{0<\gamma\leqslant T}\cos(\gamma\log x)\big)^k\,\mathrm{d}x$ liés à la répartition locale des zéros non triviaux de la fonction ζ de Riemann sur la droite critique. Grâce aux conjectures issues de l'étude des polynômes caractéristiques d'une matrice unitaire aléatoire, on a une bonne idée de ce comportement asymptotique. Évidemment, il y a ici beaucoup de conjectures, et aucun énoncé, même conjectural, ne fait mention d'une quelconque uniformité en l'ordre k. Cependant, toutes font apparaître un coefficient $k!!+o(1)$, ce qui traduit le fait qu'on attend des comportements proches d'une loi normale dans chacun des cas. Cela peut accréditer le fait que dans le cas de $M_k(h;q)$, cette constante soit également correcte.

De toute façon, face au problème de l'uniformité en k et aux importantes conséquences qui en découleraient, on peut supposer qu'il manque encore une idée fondamentalement nouvelle pour résoudre la conjecture 5.1.2.

[§] *cf.* D. A. Goldston & H. L. Montgomery, On pair correlations and primes in short intervals, Analytic Number Theory and Diophantine Problems (Stillwater, OK, July 1984) (A. C. Adolphson, J. B. Conrey, A. Ghosh, R. I. Yager, éds), *Prog. Math.* **70**, Birkäuser, Boston, 1987, p. 183–203.

[¶] *cf.* H. L. Montgomery, The pair correlation of zeros of the zeta function, Analytic Number Theory (St Louis Univ., 1972), *Proc. Sympos. Pure Math.* **24**, Amer. Math. Soc., Providence, RI, 1973, p. 181–193.

[‖] *cf.* T. H. Chan, *Pair correlation and distribution of prime numbers,* Thèse de doctorat, Ann Arbor, MI, 2002.

CHAPITRE 6

AUTOUR DU GRAPHE DIVISORIEL

Ce chapitre est totalement indépendant des précédents. Il regroupe les travaux effectués sur le graphe divisoriel à la suite d'un exposé initiatique d'Éric Saias. L'ensemble des résultats ont été obtenus après de fructueuses discussions avec Éric Saias et Pierre Mazet au Laboratoire de Probabilités et Modèles Aléatoires de Paris VII.

6.1. Motivations

6.1.1. Le graphe divisoriel et ses recouvrements hamiltoniens. — Le graphe divisoriel $\mathcal{D}$ est le graphe simple (non orienté, sans boucle ni arête double) dont les sommets sont les éléments de $\mathbb{N}^*$ et dont les arêtes sont les couples d'entiers distincts dont l'un est multiple de l'autre. Il est courant de considérer des restrictions suivantes de $\mathcal{D}$:

- le sous-graphe $\mathcal{D}(x)$ restreint aux entiers inférieurs à x, qui possède $\lfloor x \rfloor$ sommets ;
- le sous-graphe $\mathcal{D}(x;y)$ restreint aux entiers inférieurs à x dont les facteurs premiers sont inférieurs à y, qui possède $\Psi(x,y)$ sommets*.

Différentes études ont été développées dans ce cadre ; la plus notable est celle de la plus longue chaîne de $\mathcal{D}(x)$, qui a principalement été menée par Erdős, Pomerance [**32**], Tenenbaum [**37**] ou encore Saias [**35**], qui a montré que la longueur de la plus longue chaîne de $\mathcal{D}(x)$ est du même ordre de magnitude que $\asymp x/\log x$.

On rappelle qu'une **chaîne** d'un graphe est une suite de sommets tous distincts dont chaque paire de sommets consécutifs sont reliés par une arête ; le **support** d'une chaîne est l'ensemble des sommets qui font partie de cette suite et la **longueur** d'une chaîne est le cardinal de son support. Si le support d'une chaîne coïncide avec l'ensemble des sommets du graphe, cette chaîne est dite **hamiltonienne**. Si un graphe admet un chaîne hamiltonienne dont les deux extrémités sont liées par une arêtes (il s'agit alors d'un **cycle**), on dit que ce graphe est **hamiltonien**.

*La fonction $\Psi(x,y)$ est commune en théorie des nombres et désigne le nombre d'entiers inférieurs à x dont les facteurs premiers sont inférieurs à y.

La question qui nous motive, initialement posée par Erdős en 1993 et développée par Erdős et Saias dans [**12**], est celle des recouvrements hamiltoniens de $\mathcal{D}(x)$ en un minimum de chaînes. On appelle **recouvrement hamiltonien** d'un graphe toute famille de chaînes dont les supports forment une partition de l'ensemble des sommets du graphe. On appelle **indice de Hamilton** $H(G)$ d'un graphe G le nombre minimal de chaînes présentes dans un recouvrement hamiltonien de G. Même si, par son énoncé, ce problème semble proche de celui de la plus longue chaîne de $\mathcal{D}(x)$, il en est assez indépendant, tant au niveau des méthodes qu'au niveau des résultats. En particulier, on ne doit pas s'attendre à obtenir au cours de ce chapitre de nouveaux résultats concernant la longueur de la plus longue chaîne de $\mathcal{D}(x)$.

Un raisonnement simple montre que $\lfloor x/6 \rfloor \leqslant H\big(\mathcal{D}(x)\big) \leqslant x/3$ pour $x \geqslant 2$. En utilisant la majoration

$$H\big(\mathcal{D}(x)\big) \leqslant \sum_{\substack{m \geqslant 1 \\ P^+(m) > y}} H\big(\mathcal{D}(x/m; y)\big)$$

valable pour tout réel y, Saias [**36**] réussit à améliorer la majoration de la façon suivante :

$$H\big(\mathcal{D}(x)\big) < x/4,$$

pour x assez grand. Récemment, Mazet [**27**] a montré qu'il existe une constante C telle que

$$H\big(\mathcal{D}(x)\big) \sim Cx, \tag{6.1}$$

pour $x \to +\infty$, et que cette constante vérifie $0,1706 \leqslant C \leqslant 0,2289$. L'argument principal est que pour tout réel y, il existe un réel $X_0 := X_0(y)$ tel que si $x \geqslant X_0$, le graphe $\mathcal{D}(x; y)$ est hamiltonien.

Nous allons voir comment d'une version effective de cet argument, comme le sont les propositions 6.2.1 ou 6.3.1, on peut déduire une version effective de la relation (6.1).

6.1.2. Effectivité des résultats. — Pour $n \in \mathbb{N}$, on définit x_n une suite croissante de réels strictement positifs vérifiant la condition suivante :

Pour tout $x \geqslant x_n$, on a $H\big(\mathcal{D}(x; p_n)\big) = 1$.

On suppose que cette suite est log-convexe, c'est-à-dire que pout tous les entiers m et n et tout réel $\alpha \in [0; 1]$ vérifiant $\alpha m + (1-\alpha)n \in \mathbb{N}$, on a

$$\log x_{\alpha m + (1-\alpha)n} \leqslant \alpha \log x_m + (1-\alpha) \log x_n.$$

Il est possible d'interpoler la suite x_n sur $\mathbb{R}_+$ de telle façon que la fonction $t \mapsto \log x_t$ soit convexe.

On suppose que cette suite de terme x_n n'est pas constante, donc il existe un réel $\kappa > 1$ tel que $x_n \gg \kappa^n$. Pour chaque $x \geqslant 1$, on pose n_x le plus petit $m \geqslant 0$ qui vérifie la condition suivante :

Pour tout $n > m$, on a $x_n > x$.

Il s'agit d'une fonction croissante de x. De plus, si $x \geqslant x_0$ on a $x \geqslant x_{n_x}$. Pour tout $\delta \in [0; 1/2]$ on a l'encadrement valable pour $x \geqslant {x_0}^2$

$$n_x \geqslant n_{x^{1-\delta}} \geqslant \big\lfloor (1-2\delta) n_x \big\rfloor. \tag{6.2}$$

En effet, la première inégalité provient de la croissance de la fonction $x \mapsto n_x$ et par log-convexité on a

$$\log x_{(1-2\delta)n} \leqslant (1-2\delta)\log x_n + 2\delta \log x_0 \leqslant (1-\delta)\log x_n.$$

On en déduit que $x_{(1-2\delta)n_x} \leqslant {x_{n_x}}^{1-\delta} \leqslant x^{1-\delta}$, ce qui fournit la seconde inégalité.

Nous allons démontrer le résultat la proposition suivante

Proposition 6.1.1. — *Soit x_n une suite croissante log-convexe de réels vérifiant la condition :*

Pour tout $x \geqslant x_n$, on a $H\big(\mathcal{D}(x;p_n)\big) = 1$.

Il existe une constante C telle que

$$\big|H\big(\mathcal{D}(x)\big) - Cx\big| \leqslant \frac{2+o(1)}{n_x \log n_x} x,$$

où pour x assez grand, on a $n_x = \max\{m; x_m \leqslant x\}$.

6.1.2.1. Un lemme de théorie des nombres. — Pour x et y deux réels positifs, on pose

$$S(x;y) = \{n \leqslant x\,;\, P^+(n) \leqslant y\} \qquad \text{et} \qquad T(x;y) = \{n \leqslant x\,;\, P^-(n) > y\},$$

et on pose $\Psi(x,y)$ et $\Phi(x,y)$ les cardinaux respectifs de ces deux ensembles. On rappelle deux estimations classiques de ces valeurs, uniformes pour $x \geqslant y \geqslant 2$, loin d'être optimales mais suffisantes pour nos besoins :

$$\Phi(x,y) = x \prod_{p \leqslant y} \Big(1 - \frac{1}{p}\Big) + O\big(\Psi(x,y)\big) \qquad \text{et} \qquad \Psi(x,y) \ll x e^{-\frac{\log x}{2\log y}}.$$

Lemme 6.1.2. — *Soit $\varepsilon > 0$ un réel. On a uniformément pour les réels $x \geqslant y \geqslant 2$ et pour l'entier $N \geqslant 1$ vérifiant $\log(x/N) \geqslant (2+\varepsilon)(\log y)^2$*

$$\frac{1}{x} \sum_{n \leqslant N} \Psi(x/n, y) = o_\varepsilon\Big(\frac{1}{y}\Big).$$

Démonstration. — Pour tout n, on a $x/n \geqslant 2$. On a donc

$$\sum_{n \leqslant N} \Psi(x/n, y) \ll \sum_{n \leqslant N} \Big(\frac{x}{n}\Big)^{1-\frac{1}{2\log y}}.$$

Or pour $\delta > 0$, on a $\sum_{n \leqslant N} n^{\delta-1} \leqslant 1 + \int_1^N t^{\delta-1}\,\mathrm{d}t = 1 - 1/\delta + N^\delta/\delta$. Ainsi, on a

$$\sum_{n \leqslant N} \Psi(x/n, y) \ll x \log y\, (x/N)^{-\frac{1}{2\log y}}.$$

Par hypothèse, on a $-\log(x/N)\,/\,2\log y \leqslant -(1+\varepsilon/2)\log y$, ce qui une fois replacé dans la majoration précédente fournit le résultat. □

6.1.2.2. Un résultat général de Mazet. — Nous retraduisons en terme de graphe divisoriel un lemme général contenu dans l'article de Mazet, qui est l'argument essentiel de la preuve de la proposition 6.1.1.

Pour alléger les notations, on pose $F(x) \coloneqq H\big(\mathcal{D}(x)\big)$, $F_n(x) \coloneqq H\big(\mathcal{D}(x;p_n)\big)$ et $G_n(x) \coloneqq \sum_{m\in T(x;p_n)} F_n(x/m)$.

Lemme 6.1.3. — *Pour tous les entiers $x \geqslant 1$ et $n \geqslant 1$, on a*

$$0 \leqslant \frac{G_n(x)}{x} - \frac{F(x)}{x} \leqslant \frac{2}{p_{n+1}}.$$

Démonstration. — On remarque que $G_n(x)$ est l'indice de Hamilton du graphe

$$\mathcal{G}_n(x) \coloneqq \coprod_{m\in T(x;p_n)} m\,\mathcal{D}(x/m;p_n).$$

Ce graphe possède les mêmes sommets que $\mathcal{D}(x)$, et toutes ces arêtes sont des arêtes de $\mathcal{D}(x)$: tout recouvrement hamiltonien de ce graphe est un recouvrement hamiltonien de $\mathcal{D}(x)$. Ainsi, on a bien $G_n(x) \geqslant F(x)$. Inversement, de tout recouvrement hamiltonien de $\mathcal{D}(x)$ on peut construire un recouvrement hamiltonien du graphe $\mathcal{G}_n(x)$ en y retirant toutes les arêtes absentes du graphe $\mathcal{G}_n(x)$. De telles arêtes rejoignent deux entiers dont le rapport est dans $T(x;p_n) \smallsetminus \{1\}$; l'un de ces entiers est donc inférieur à x/p_{n+1}. Dans un recouvrement hamiltonien, chaque sommet est l'extrémité d'au plus deux arêtes. Ainsi, le nombre d'arêtes retirées à un recouvrement hamiltonien de $\mathcal{D}(x)$ pour le rendre recouvrement hamiltonien de $\mathcal{G}_n(x)$ est inférieur à $2x/p_{n+1}$, ce qui signifie que $G_n(x) \leqslant F(x) + 2x/p_{n+1}$. □

6.1.2.3. Un lemme spécifique au graphe divisoriel. —

Lemme 6.1.4. — *Soit n un entier naturel. Il existe pour $1 \leqslant k \leqslant x_n$ des éléments λ_k de $\{-1,0,1\}$ tel que pour tout réel x on a*

$$F_n(x) = \sum_{1\leqslant k\leqslant x_n} \lambda_k [x \geqslant k],$$

où $x \mapsto [x \geqslant k]$ est la fonction caractéristique de $[k,+\infty[$.

Démonstration. — La fonction F_n est une fonction en escalier, dont les discontinuités ont lieu en des valeurs entières de la variable. De plus $F_n(x)$ est nulle si $x < 1$ et vaut 1 si $x \geqslant x_n$. Par sommation d'Abel, on a bien

$$F_n(x) = \sum_{1\leqslant k\leqslant x_n} \big(F_n(k) - F_n(k-1)\big)[x \geqslant k].$$

En particulier on a $\sum_k \lambda_k = 1$. Il ne reste qu'à montrer que pour tout k on a

$$-1 \leqslant F_n(k) - F_n(k-1) \leqslant 1.$$

Les graphes $\mathcal{D}(k;p_n)$ et $\mathcal{D}(k-1;p_n)$ différent éventuellement d'un sommet et des arêtes liées à ce sommet. Plaçons-nous dans le cas où ces graphes ne sont pas identiques. En retirant le sommet k à un recouvrement hamiltonien $\mathcal{D}(k;p_n)$, on obtient un recouvrement hamiltonien de $\mathcal{D}(k-1;p_n)$ avec éventuellement une composante en plus : ainsi, on a $F_n(k-1) \leqslant F_n(k) + 1$. Inversement, en rajoutant le sommet k à

un recouvrement hamiltonien de $\mathcal{D}(k-1;p_n)$, on obtient un recouvrement hamiltonien $\mathcal{D}(k;p_n)$ qui possède une composante de plus : ainsi, on a $F_n(k) \leqslant F_n(k-1)+1$. On obtient bien l'encadrement annoncé. □

6.1.2.4. Estimation principale. — Nous allons à présent fournir la preuve de la proposition 6.1.1 en établissant

$$\Big|\frac{F(x)}{x} - C\Big| \leqslant \frac{2+o(1)}{n_x \log n_x}.$$

Démonstration de la proposition 6.1.1. — En utilisant le lemme 6.1.4 dans la définition de G_n, on obtient

$$G_n(x) = \sum_{m\in T(x;p_n)} F_n(x/m) = \sum_{1\leqslant k\leqslant x_n} \lambda_k \sum_{m\in T(x;p_n)} [x \geqslant km] = \sum_{1\leqslant k\leqslant x_n} \lambda_k \Phi(x/k, p_n).$$

On définit pour chaque entier n la constante C_n par

$$C_n \coloneqq \prod_{i=1}^{n}\Big(1-\frac{1}{p_i}\Big) \sum_{1\leqslant k\leqslant x_n} \frac{\lambda_k}{k}.$$

On a uniformément pour les entiers x et n vérifiant $\log(x/x_n) \geqslant 3(\log p_n)^2$

$$\Big|\frac{G_n(x)}{x} - C_n\Big| = \frac{1}{x}\Big|\sum_{1\leqslant k\leqslant x_n} \lambda_k \Big(\Phi(x/k,p_n) - \frac{x}{k}\prod_{i=1}^{n}\Big(1-\frac{1}{p_i}\Big)\Big)\Big|$$

or par le lemme 6.1.4 on a $|\lambda_k| \geqslant 1$ et on rappelle la majoration uniforme $\Phi(x,y) - x\prod_{p\leqslant y}(1-1/p) \ll \Psi(x,y)$

$$\ll \frac{1}{x}\sum_{1\leqslant k\leqslant x_n} \Psi(x/k, p_n)$$

et comme les conditions du lemme 6.1.2 sont satisfaites avec $\varepsilon = 1$,

$$= o\Big(\frac{1}{p_{n+1}}\Big).$$

Ainsi, pour tout entier n, le rapport $G_n(x)/x$ converge vers un réel C_n lorsque x tend vers ∞. Le lemme 6.1.3 entraîne que pour toute paire d'entiers $n < m$ on a $0 \leqslant C_n - C_m \leqslant 2/p_{n+1}$. La suite des C_n converge donc vers un réel C qui vérifie $0 \leqslant C_n - C \leqslant 2/p_{n+1}$. Pour tous les entiers n et x, on a

$$\frac{F(x)}{x} - C = \Big(\frac{F(x)}{x} - \frac{G_n(x)}{x}\Big) + \Big(\frac{G_n(x)}{x} - C_n\Big) + (C_n - C).$$

Par une nouvelle utilisation du lemme 6.1.3, on obtient

$$\Big|\frac{F(x)}{x} - C\Big| \leqslant \Big|\frac{G_n(x)}{x} - C_n\Big| + \frac{2}{p_{n+1}} = \frac{2+o(1)}{p_{n+1}}. \tag{6.3}$$

Cette majoration est valable uniformément en n tant que $\log(x/x_n) \geqslant 3(\log p_n)^2$. On souhaite donc choisir cet entier n de la meilleure façon possible. Puisque l'on doit avoir $x_n \leqslant xe^{-3(\log p_n)^2} \leqslant x$, on doit choisir $n \leqslant n_x$. Ainsi, il suffit que soit

vérifiée $x_n \leqslant xe^{-3(\log p_{n_x})^2}$. Or, comme $x_n \gg \kappa^n$, on a $n_x \ll \log x$ et donc pour x assez grand on a $3(\log p_{n_x})^2 \leqslant \frac{1}{2}\log x$ et on peut utiliser l'encadrement (6.2) pour voir que l'on peut choisir n vérifiant simultanément $\log(x/x_n) \geqslant 3(\log p_n)^2$ et

$$\Big(1 - 6\frac{(\log p_{n_x})^2}{\log x}\Big)n_x - 1 \leqslant n.$$

Avec un tel choix de n et à l'aide du théorème des nombres premiers, on a

$$p_{n+1} \geqslant \Big(1 + O\big((\log p_{n_x})^2/\log x\big)\Big)n_x \log n_x = \big(1 + o(1)\big)n_x \log n_x.$$

Et puisque $\log(x/x_n) \geqslant 3(\log p_n)^2$ est toujours vérifiée, on peut replacer cette estimation dans la majoration (6.3) pour en déduire le résultat annoncé. □

6.2. Le théorème de Mazet

Un des intérêts du travail de Mazet est que les résultats qu'il obtient dépassent largement le cadre du graphe divisoriel mais s'appliquent à tout monoïde (ou semi-groupe) libre commutatif « plongé » dans $\mathbb{R}_+$ de façon raisonnable. Cela revient à considérer le problème du graphe divisoriel pour les entiers généralisés de Beurling. Par souci de clarté et de cohérence avec le travail de Mazet, nous utiliserons pour nos monoïdes le formalisme additif. Cependant, on traduira régulièrement nos résultats dans le formalisme multiplicatif, et dans le cas particulier du graphe divisoriel.

Soit $(M, +)$ un monoïde libre commutatif de base $(e_x)_{x\in X}$ et ν un morphisme de monoïdes de $(M, +)$ dans $(\mathbb{R}_+, \cdot)$. Pour tout élément $m \in M$ et tout $x \in X$, on définit par $\langle m|e_x\rangle \in \mathbb{N}$ la coordonnée de m en e_x, si bien que $m = \sum_X \langle m|e_x\rangle e_x$. On pose également $\nu_x := \nu(e_x) \in \mathbb{R}_+$ pour $x \in X$. Ainsi, le morphisme ν est entièrement déterminé par la suite des ν_x puisque l'on a

$$\nu(m) = \sum_{x\in X} \langle m|e_x\rangle \nu_x$$

pour tout $m \in M$.

Notre résultat s'énonce pour les monoïdes libres de type fini, *i.e.* tel que $\operatorname{card} X$ soit un entier n. Dans ce cas, on effectue un changement d'indices afin que $X = \{1, 2, \ldots, n\}$ et que la suite $\nu_1, \nu_2, \ldots, \nu_n$ soit croissante. Mais les conséquences que permettent d'en tirer le travail de Mazet s'énoncent pour tout monoïde libre M envoyé par ν dans $\mathbb{R}_+$ tel que l'ensemble des $\{\nu_x\}_{x\in X}$ ne possède pas de point d'accumulation dans $\mathbb{R}_+$ (ce qui est vérifié si X est fini) ; on dira alors que ν est un morphisme **régulier**. Ainsi, si X n'est pas fini, il est nécessairement dénombrable : on peut alors effectuer un changement d'indices afin que $X = \{1, 2, 3, \ldots\}$ et que la suite $\nu_1, \nu_2, \nu_3, \ldots$ soit croissante. Par la suite, toute donnée d'un monoïde libre induira la donnée d'un morphisme régulier ν et d'une indexation de sa base par des entiers naturels.

En outre, puisque l'on utilisera le raisonnement par récurrence, il sera commode de considérer tout monoïde de type fini comme le sous-monoïde d'un monoïde de type non-fini, ou du moins d'un monoïde de dimension supérieure. Ainsi pour tout monoïde libre M muni d'un morphisme régulier ν et pour tout entier n strictement inférieur

à la dimension de M, on définit M_n comme sous-monoïde de M engendré par e_1, e_2, … , e_n et on le munit du morphisme ν restreint à M_n.

Pour tout monoïde libre M et pour tout réel t, on définit le monoïde tronqué

$$M(t) \coloneqq \{m \in M ; \nu(m) \leqslant t\}.$$

L'hypothèse de régularité du morphisme ν est équivalente au fait que pour tout réel t, le cardinal de $M(t)$ est fini. On note $|M(t)|$ ce cardinal.

Enfin, indépendamment du morphisme ν, on considère la relation d'ordre partielle sur les monoïdes : pour m et m' deux éléments de M, on dit que m est inférieure à m' si $m' - m$ est un élément du monoïde. Dans le cas d'un monoïde libre, cela est équivalent au fait que, par rapport à chaque vecteur de la base, la coordonnée de m soit inférieure à celle de m'. On définit alors $\mathcal{M}$ le graphe simple de sommets M et dont les arêtes relient les paires de sommets distincts m et m' qui sont ordonnées. On note $\mathcal{M}(t)$ pour $t \in \mathbb{R}$ le sous-graphe de $\mathcal{M}$ restreint aux éléments de $M(t)$.

On utilisera des versions étoilées de ces objets pour signifier que l'on a retiré l'origine.

En posant $\nu_n = \log p_n$, le graphe $\mathcal{M}(t)$ est clairement isomorphe au graphe $\mathcal{D}(e^t)$ et que le graphe $\mathcal{M}_n(t)$ est isomorphe au graphe $\mathcal{D}(e^t ; p_n)$. En adaptant le choix de ν_n, on obtient le graphe divisoriel d'entiers généralisés de Beurling.

Nous donnons ici, pour comparer, une version effective adaptée directement du théorème de Mazet.

Pour $n \geqslant 2$ un entier et t_* un réel, on dit qu'un monoïde libre M de dimension supérieure à n vérifie la condition $\mathfrak{H}_n(t_*)$ si pour tout $t \geqslant t_*$, le graphe $\mathcal{M}_n^*(t)$ possède une chaîne hamiltonienne d'extrémités e_1 et e_n. On dit que M vérifie $\mathfrak{H}_1(t_*)$ si pour tout $t \geqslant t_*$, le graphe $\mathcal{M}_1^*(t)$ possède une chaîne hamiltonienne d'extrémités e_1 et $2e_1$.

Proposition 6.2.1. — *Soit $n \geqslant 2$ un entier et soit M un monoïde libre de dimension supérieure à n. S'il existe un réel t_{n-1} tel que M vérifie $\mathfrak{H}_{n-1}(t_{n-1})$, alors M vérifie $\mathfrak{H}_n(t_n)$ pour $t_n = t_{n-1} + \nu_n\big(1 + |M_n(t_{n-1})|\big)$.*

Démonstration. — La démonstration est faite pour $n \geqslant 3$, et évite la spécificité de la condition $\mathfrak{H}_1$. Pour $n = 2$, il suffit de remplacer formellement tout occurrence de « e_{n-1} » par le terme « $2e_1$ ».

Soit $t \geqslant t_n$. On pose $\kappa \coloneqq \lfloor (t - t_{n-1})/\nu_n \rfloor$, si bien que sont vérifiées les inégalités

$$\kappa \geqslant |M_n(t_{n-1})| + 1, \qquad t - (\kappa+1)\nu_n < t_{n-1} \qquad \text{et} \qquad t - \kappa\nu_n \geqslant t_{n-1}.$$

On considère la partition suivante de $M_n^*(t)$ en trois parties.

- La partie $\mathcal{R}$ ($\mathcal{R}$ comme « réserve ») est l'ensemble des ke_n pour $1 \leqslant k \leqslant \kappa$.
- La partie $\mathcal{S}$ ($\mathcal{S}$ comme « strates ») est l'ensemble des m de $M_n^*(t)$ tel que $\langle m | e_n \rangle \leqslant \kappa$ et tel que leur projection sur $M_{n-1}(t)$ soit non nul, *i.e.* $m - \langle m | e_n \rangle e_n \neq 0$. On peut stratifier cet ensemble selon la valeur k de $\langle m | e_n \rangle$, la strate de hauteur k étant isomorphe à $M_{n-1}^*(t - \nu_n k)$. On écrira cet ensemble sous la forme

$$\mathcal{S} = \coprod_{k=0}^{\kappa} \big(M_{n-1}^*(t - \nu_n k) + ke_n\big).$$

– La partie $\mathcal{T}$ ($\mathcal{T}$ comme « toit ») est l'ensemble des m de $M_n^*(t)$ tel que $\langle m|e_n\rangle > \kappa$. Cet ensemble est isomorphe à $M_n\big(t-\nu_n(\kappa+1)\big)$. Plus précisément, on a

$$\mathcal{T} = M_n\big(t-\nu_n(\kappa+1)\big) + (\kappa+1)e_n,$$

donc $|\mathcal{T}| = \big|M_n\big(t-\nu_n(\kappa+1)\big)\big| \leqslant |M_n(t_{n-1})|$ par notre choix de κ.

L'idée est de recouvrir chaque strate de $\mathcal{S}$ par une chaîne hamiltonienne. Cela est possible en appliquant l'hypothèse de récurrence à chaque strate ; on a effet choisit κ tel que $t-\kappa\nu_n \geqslant t_{n-1}$. La strate de hauteur k est alors recouverte par une chaîne d'extrémités e_1+ke_n et $e_{n-1}+ke_n$. On relie ces chaînes tête-bêche en utilisant les arêtes $\{e_{n-1}+ke_n, e_{n-1}+(k+1)e_n\}$ pour les k pairs et $\{e_1+ke_n, e_1+(k+1)e_n\}$ pour les k impairs. On a donc recouvert $\mathcal{S}$ par une chaîne hamiltonienne ayant e_1 pour extrémité ; l'autre extrémité est soit $e_1+\kappa e_n$ soit $e_{n-1}+\kappa e_n$, selon que κ est impair ou pair. Pour simplifier, on note $e_{\mathcal{S}}$ cette dernière extrémité.

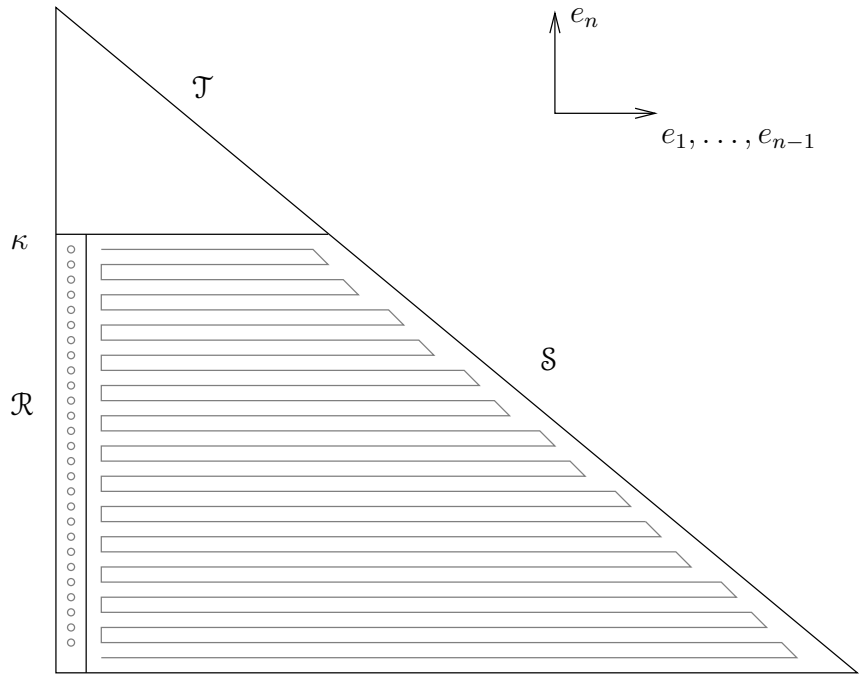

FIGURE 1. La construction de Mazet

L'intérêt de la réserve $\mathcal{R}$ est que chacun de ces points est en relation avec $e_{\mathcal{S}}$ et avec l'ensemble des points de $\mathcal{T}$ et de $\mathcal{R}$. Ainsi, puisque $|\mathcal{R}| = \kappa \geqslant |\mathcal{T}|+1$, de toute énumération $t_1, t_2, \ldots, t_{|\mathcal{T}|}$ des points de $\mathcal{T}$ on construit la chaîne en alternant points

de $\mathcal{R}$ et point de $\mathcal{T}$

$$\kappa e_n \,-\, t_1 \,-\, (\kappa-1)e_n \,-\, t_2 \,-\, (\kappa-2)e_n \,-\, \ldots \,-\, (\kappa-|\mathcal{T}|+1)e_n \,-\, t_{|\mathcal{T}|} \,- \\ -\, (\kappa-|\mathcal{T}|)e_n \,-\, (\kappa-|\mathcal{T}|-1)e_n \,-\, \ldots \,-\, 2e_n \,-\, e_n.$$

Cette chaîne couvre entièrement $\mathcal{R}$ et $\mathcal{T}$, et a pour extrémités e_n et κe_n. En utilisant l'arête $\{e_{\mathcal{S}}, \kappa e_n\}$, on peut relier la chaîne couvrant $\mathcal{S}$ à la chaîne couvrant $\mathcal{R}$ et $\mathcal{T}$, et ainsi former une chaîne hamiltonienne de $\mathcal{M}_n^*(t)$ ayant pour extrémités e_1 et e_n. □

Ce résultat n'est pas très efficace pour majorer t_n le plus petit t tel que M vérifie $\mathfrak{H}_n(t)$. En effet, il ne permet pas de fournir une majoration meilleure que $t_n \ll (\sum_1^n \nu_i)e^{n!}$ dans les cas où la suite ν_i ne croît pas trop vite. Par exemple, on se place dans le cas où il existe un réel $C > 0$ tel que pour tout $n > 1$ on ait $\nu_n \leqslant C\sum_{i=1}^{n-1}\nu_i$, si bien que $\sum_{i=1}^n \nu_i \leqslant (1+C)\sum_{i=1}^{n-1}\nu_i$. Il est clair que le nombre de points de $M_n(t)$ est supérieur au volume du simplexe de $\mathbb{R}^n$ d'extrémités $(0,0,\ldots,0)$, $(t/\nu_1,0,\ldots,0)$, $(0,t/\nu_2,\ldots,0)$, $\ldots$, $(0,0,\ldots,t/\nu_n)$, ainsi $|M_n(t)| \geqslant \frac{1}{n!}t^n/\prod_{i=1}^n \nu_i$. De plus, pour tout i, on a $(n-i)\nu_i \leqslant \sum_1^{n-1}\nu_i$, ainsi on peut écrire

$$t_n/(\textstyle\sum_1^n \nu_i) \geqslant \dfrac{\sum_1^{n-1}\nu_i}{\sum_1^n \nu_i}\dfrac{1}{n}\big(t_{n-1}/(\textstyle\sum_1^{n-1}\nu_i)\big)^n \geqslant \dfrac{1}{n(1+C)}\big(t_{n-1}/(\textstyle\sum_1^{n-1}\nu_i)\big)^n.$$

Donc si on a $t_{n-1} \geqslant 2(1+C)^{1/(n-1)}\big(\sum_1^{n-1}\nu_i\big)e^{(n-1)!}$, on a $t_n \geqslant 2(1+C)^{1/n}\big(\sum_1^n \nu_i\big)e^{n!}$. Dans l'autre sens, il est possible de montrer que sous des conditions analogues, on a

$$t_n \ll \nu_1 2^{(n^n)}.$$

Application au graphe divisoriel. — On note $x_n := e^{t_n}$ le majorant fourni par le lemme de Mazet pour le plus petit x qui vérifie que pour tout $m \geqslant x$ le graphe divisoriel $\mathcal{D}(m; p_n)$ est hamiltonien. On a donc par la proposition 6.2.1

$$x_n = x_{n-1}{p_n}^{1+\Psi(x_{n-1},p_n)} \quad \text{et} \quad x_n \gg e^{(e^{n!})}\prod_{p\leqslant p_n} p.$$

Pour simplifier et sans faire de concession, on peut choisir $x_n = e^{(e^{(n^n)})}$, ce qui est bien log-convexe et on obtient

$$n_x \sim \log_3 x/\log_4 x.$$

On peut donc prouver grâce à la proposition 6.1.1 la relation

$$H\big(\mathcal{D}(x)\big) = \big(C + O(1/\log_3 x)\big)x.$$

6.3. Amélioration de la construction de Mazet

La construction de Mazet ne cherchait pas à être effective, il est clair que l'on doit pouvoir faire beaucoup mieux. Nous proposons la conjecture suivante.

Conjecture 6.3.1. — Soit M un monoïde libre de dimension finie de base B. Le graphe $\mathcal{M}(t)$ est hamiltonien si le vecteur $\sum_{e\in B} e$ appartient à $M(t)$.

D'une façon équivalente, cela signifie que $t_n \leqslant \sum_{i=1}^{n} \nu_i$. Cette borne est optimale dans le cas où l'on choisit $\nu_1 = \nu_2 = \ldots = \nu_n = 1$. En effet, dans ce cas, l'ensemble des sommets extrémaux m de $\mathcal{M}_n(t)$, *i.e.* tels que pour tout vecteur m' de M_n^* on a $m+m' \notin M_n(t)$, est exactement l'ensemble $M_n(t) \smallsetminus M_n(t-1)$, de cardinal $\binom{\lfloor t \rfloor + n}{n} - \binom{\lfloor t \rfloor + n - 1}{n} = \binom{\lfloor t \rfloor + n - 1}{n-1}$, et cet ensemble est complètement déconnecté dans $\mathcal{M}_n(t)$, qui lui possède $\binom{\lfloor t \rfloor + n}{n}$ sommets. Une condition nécessaire au fait que $\mathcal{M}_n(t)$ soit hamiltonien est donc que $\binom{\lfloor t \rfloor + n}{n} \geqslant 2\binom{\lfloor t \rfloor + n - 1}{n-1}$, soit $t \geqslant n$.

Nous allons présenter une construction qui permet de montrer que le graphe $\mathcal{M}(t)$ est hamiltonien pour peu que $t_n \gg e^n$. Ce résultat représente donc un « gain d'une exponentielle » par rapport à la preuve originale de Mazet une fois rendue effective (on a la condition $t_n \gg e^{(e^{(n^{1+\varepsilon})})}$ à partir de la proposition 6.2.1, et la conjecture 6.3.1 correspond à la condition $t_n \geqslant n$. On se trouve donc en quelque sorte « au milieu du chemin ».)

On définit la propriété $\mathfrak{P}_n(t_n)$ de la façon suivante : pour tout $t \geqslant t_n$ et pour tout $\mathcal{I} \subset M_n(t - t_n)$, il existe une chaîne de $M_n(t)$ ayant pour extrémités $\lfloor t/\nu_1 \rfloor\, e_1$ et $\lfloor t/\nu_n \rfloor\, e_n$ et pour sommets l'ensemble $M_n(t) \smallsetminus \mathcal{I}$.

Proposition 6.3.1. — *Soit $n \geqslant 3$ un entier et M un monoïde libre de dimension supérieure à n. S'il existe un réel t_{n-1} tel que M possède la propriété $\mathfrak{P}_{n-1}(t_{n-1})$, alors M vérifie $\mathfrak{P}_n(t_n)$ pour tout $t_n = 2\max(t_{n-1}, \nu_n) + 2\nu_n + \nu_{n-1}$.*

Démonstration. — On pose $t' \coloneqq \max(t_{n-1}, \nu_n)$ et $t_n \coloneqq 2t' + 2\nu_n + \nu_{n-1}$. Soit $t \geqslant t_n$ un réel et soit $\mathcal{I} \subset M_n(t - t_n)$. On pose $\kappa \coloneqq \lfloor (t - t')/\nu_n \rfloor$ et $\kappa' \coloneqq \lfloor (t - t_n)/\nu_n \rfloor + 1$. On remarque comme $t_n \geqslant t' + 3\nu_n$, on a $\kappa' + 1 < \kappa$. On a également $t - \kappa\nu_n \geqslant t'$ et $\kappa'\nu_n > t - t_n$.

Pour $0 \leqslant j < \kappa$ on définit l'élément p_j qui servira de point de colle par

$$p_j \coloneqq \begin{cases} \lfloor (t - t' - j\nu_n)/\nu_{n-1} \rfloor\, e_{n-1} + j e_n & \text{si } j \text{ est pair}\,; \\ \lfloor (t - t' - j\nu_n)/\nu_1 \rfloor\, e_1 + j e_n & \text{si } j \text{ est impair.} \end{cases}$$

Ces points sont tous différents (par leur composante en e_n) et les valeurs de $\nu(p_j)$ vérifient les encadrements suivants :

$$\begin{aligned} t - t' - \nu_{n-1} < \nu_{n-1} \lfloor (t - t' - j\nu_n)/\nu_{n-1} \rfloor + \nu_n j \leqslant t - t' \qquad & \text{si } j \text{ est pair}\,; \\ t - t' - \nu_1 < \nu_1 \lfloor (t - t' - j\nu_n)/\nu_1 \rfloor + \nu_n j \leqslant t - t' \qquad & \text{si } j \text{ est impair.} \end{aligned}$$

Ainsi les points de colle p_j vérifient tous $t - t' - \nu_{n-1} < \nu(p_j) \leqslant t - t'$, donc pour tout j, $p_j + e_n$ vérifie $\nu(p_j + e_n) \leqslant t - t' + \nu_n \leqslant t$, et donc $p_j + e_n \in M_n(t)$. On note $\mathcal{P}$ l'ensemble de ces points de colle. Comme $t - t' - \nu_{n-1} \geqslant t - t_n$, les ensembles $\mathcal{P}$ et $\mathcal{I}$ sont disjoints

On définit la réserve $\mathcal{R}$ comme l'ensemble

$$\mathcal{R} \coloneqq \big(M_{n-1}^*(t') + \kappa' e_n\big) \amalg \big(M_{n-1}(t') + (\kappa' + 1)e_n\big).$$

Les points $r \in \mathcal{R}$ vérifient

$$\nu(r) \leqslant t' + (\kappa' + 1)\nu_n \leqslant t' + t - t_n + 2\nu_n \leqslant t - t' - \nu_{n-1}$$

donc les ensembles $\mathcal{P}$ et $\mathcal{R}$ sont disjoints et

$$\nu(r) > \kappa'\nu_n > t - t_n,$$

donc les ensembles $\mathcal{R}$ et $\mathcal{I}$ sont disjoints.

On pose $\mathcal{Q} := \mathcal{I} \amalg \mathcal{P} \amalg \mathcal{R}$. Pour $0 \leqslant j \leqslant \kappa$, on définit $\mathcal{Q}_j$ l'ensemble des éléments q de $\mathcal{Q}$ tels que $\langle q|e_n\rangle = j$. Il est clair que ces différents ensembles sont deux à deux disjoints. De plus, on a remarqué que tout élément q de $\mathcal{Q}$ (qu'il soit dans $\mathcal{P}$, dans $\mathcal{R}$ ou dans $\mathcal{I}$) vérifie $\nu(q) \leqslant t - t'$, et donc $\langle q|e_n\rangle \leqslant \lfloor (t-t')/\nu_n \rfloor = \kappa$. On a donc

$$\mathcal{Q} = \coprod_{j=0}^{\kappa} \mathcal{Q}_j.$$

Enfin, on définit $\mathcal{T}$ comme l'ensemble des points m de $M_n(t) \smallsetminus M_n(t-t')$ vérifiant $\langle m|e_n\rangle > \kappa$ (donc $\mathcal{T}$ est disjoint de $\mathcal{Q}$) et on pose $\mathcal{S} := M_n(t) \smallsetminus (\mathcal{T} \amalg \mathcal{Q})$.

On définit pour $0 \leqslant j \leqslant \kappa$ l'ensemble $\mathcal{S}_j$ des éléments s de $\mathcal{S}$ vérifiant $\langle s|e_n\rangle = j$, si bien que

$$\mathcal{S} = \coprod_{j=0}^{\kappa} \mathcal{S}_j.$$

Pour chaque j entre 0 et κ, on a $\mathcal{S}_j \amalg \mathcal{Q}_j = je_n + M_{n-1}(t - j\nu_n)$ et $\mathcal{Q}_j \subset M_{n-1}(t - t' - j\nu_n)$. Puisque $t - j\nu_n \geqslant t - \kappa\nu_n \geqslant t' \geqslant t_{n-1}$, on peut utiliser la propriété $\mathfrak{P}_{n-1}(t_{n-1})$ pour déduire qu'il existe une chaîne hamiltonienne C_j de $\mathcal{S}_j$ d'extrémités $je_n + \lfloor (t - j\nu_n)/\nu_{n-1} \rfloor e_{n-1}$ et $je_n + \lfloor (t - j\nu_n)/\nu_1 \rfloor e_1$. Pour $0 \leqslant j \leqslant \kappa$, l'élément $je_n + \lfloor (t - j\nu_n)/\nu_1 \rfloor e_1$ est désigné par a_j si j est pair et par b_j si j est impair, et l'élément $je_n + \lfloor (t - j\nu_n)/\nu_{n-1} \rfloor e_{n-1}$ est désigné par b_j si j est pair et par a_j si j est impair ; les deux extrémités de la chaîne C_j sont a_j et b_j. De plus, pour tout $0 \leqslant j < \kappa$, le point de colle p_j est inférieur à b_j (pour la relation d'ordre de M_n) et à $p_j + e_n$, qui est lui-même inférieur ou égal à a_{j+1}. On peut ainsi lier les chaînes C_j tête-bêche, de la façon suivante

$$\underbrace{a_0 - \ldots - b_0}_{C_0} - p_0 - \underbrace{a_1 - \ldots - b_1}_{C_1} - p_1 - \ldots - p_{\kappa-1} - \underbrace{a_\kappa - \ldots - b_\kappa}_{C_\kappa}.$$

On a donc construit une chaîne hamiltonienne de $\mathcal{S} \amalg \mathcal{P}$ ayant a_0 et b_κ pour extrémités.

On utilise à présent la propiété $\mathfrak{P}_{n-1}(t_{n-1})$ pour recouvrir $M^*_{n-1}(t')$. On note γ le cardinal de cet ensemble. On sait qu'il existe une chaîne hamiltonienne de $M^*_{n-1}(t')$, qu'on note $v_1 - v_2 - \ldots - v_\gamma$ où l'on a imposé $v_1 = \lfloor t'/\nu_{n-1} \rfloor e_{n-1}$ et $v_\gamma = \lfloor t'/\nu_1 \rfloor e_1$ si κ est pair, et $v_1 = \lfloor t'/\nu_1 \rfloor e_1$ et $v_\gamma = \lfloor t'/\nu_{n-1} \rfloor e_{n-1}$ si κ est impair. On pose $v_0 = 0$. On construit à partir de cette chaîne une hamiltonienne de $\mathcal{R}$, comme suit :

$$\begin{array}{ll} (\kappa'+1)e_n + \ldots & \quad v_0 - v_1 \quad v_2 - v_3 \quad v_4 - \ldots \\ & \qquad\; | \quad\;\; | \quad\;\; | \quad\;\; | \\ \kappa' e_n + \ldots & \qquad v_1 - v_2 \quad v_3 - v_4 \end{array}$$

$$\text{et ce jusqu'à} \quad \begin{array}{c} v_\gamma \\ | \\ \ldots - v_\gamma \end{array} \quad \text{ou} \quad \begin{array}{c} \ldots - v_\gamma \\ | \\ v_\gamma \end{array} \quad \text{selon la parité de } \gamma.$$

Nous donnons quelques précisions sur ce schéma.
Les sommets de la première ligne sont compris comme étant la somme de $(\kappa'+1)e_n$ et d'un élément de $M_{n-1}(t')$; les arêtes horizontales au sein de la première ligne sont justifiées car elles proviennent d'arêtes provenant de $M_{n-1}(t')$. Ces sommets décrivent exactement l'ensemble $(\kappa'+1)e_n + M_{n-1}(t')$. Les sommets de la seconde ligne sont eux compris comme étant la somme de $\kappa' e_n$ et d'un élément de $M^*_{n-1}(t')$; les arêtes horizontales sont justifiées pour une raison similaire. Ces sommets décrivent l'ensemble $\kappa' e_n + M^*_{n-1}(t')$. Les arêtes verticales correspondent à l'ajout (ou la soustraction) de e_n à un vecteur, elles sont donc justifiées. On a bien ainsi une chaîne hamiltonienne de $\mathcal{R}$.

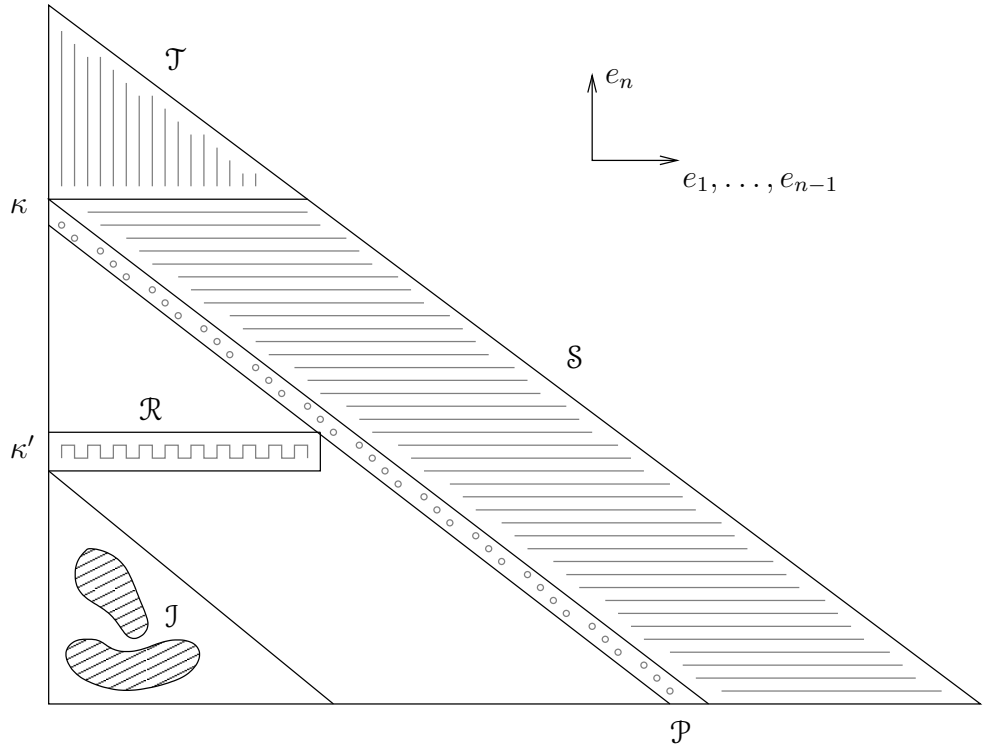

FIGURE 2. Une amélioration de la construction de Mazet

On considère pour tout point m de $\mathcal{T}$ le projection $\pi(m)$ sur M_{n-1}. Plus précisément, on a $\pi(m) = m - \langle m|e_n\rangle e_n$. On a

$$\nu\big(\pi(m)\big) = \nu(m) - \langle m|e_n\rangle\nu_n \leqslant t - (\kappa+1)\nu_n < t'.$$

On rappelle que les éléments v_i décrivent $M_{n-1}(t')$. Pour tout $0 \leqslant i \leqslant \gamma$, on définit l'ensemble $\mathcal{T}_i$ des éléments m de $\mathcal{T}$ vérifiant $\pi(m) = v_i$: on obtient ainsi une partition de $\mathcal{T}$. Entre deux éléments d'une même composante $\mathcal{T}_i$, seule la coordonnée en e_n

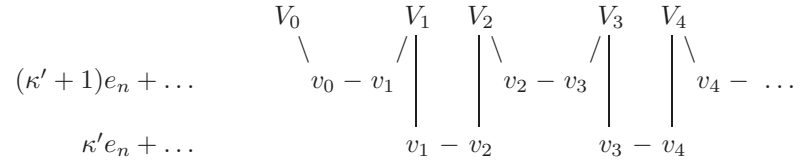

FIGURE 3. La technique du serpent

change : les ensembles $\mathcal{T}_i$ sont donc chacun complètement ordonnés. Pour tout $0 \leqslant i \leqslant \gamma$, on classe les éléments de $\mathcal{T}_i$ du plus petit c_i au plus grand d_i, on obtient ainsi une chaîne hamiltonienne V_i de $\mathcal{T}_i$ ayant pour extrémités c_i et d_i. Ces chaînes peuvent être réduites à un point ou même être vides. On remarque que $d_0 = \lfloor t/\nu_n \rfloor\, e_n$.

Puisque $\kappa > \kappa'$, on peut replacer les chaînes V_i au milieu de la chaîne recouvrant $\mathcal{R}$ de la façon indiquée par la figure 3. Le sens dans lequel sont collées les chaînes V_i n'importe pas, la seule chose qui est demandée est que ce soit c_0 qui soit relié à v_0. La chaîne ainsi créée est une chaîne hamiltonienne de $\mathcal{R} \amalg \mathcal{T}$ et possède pour extrémités d'un côté l'élément $d_0 = \lfloor t/\nu_n \rfloor\, e_n$, et de l'autre, l'élément $v_\gamma + (\kappa'+1)e_n$ ou $v_\gamma + \kappa' e_n$ selon la parité de γ.

Nous disposons de deux chaînes, l'une couvrant $\mathcal{S} \amalg \mathcal{P}$, dont les extrémités sont $a_0 = \lfloor t/\nu_1 \rfloor\, e_1$ et b_κ, l'autre couvrant $\mathcal{R} \amalg \mathcal{T}$, dont les extrémités sont $d_0 = \lfloor t/\nu_n \rfloor\, e_n$ et $v_\gamma + (\kappa'+1)e_n$ ou $v_\gamma + \kappa' e_n$ selon la parité de γ. On a $t' \leqslant t - \kappa\nu_n$, donc quelque soit la parité de κ, l'élément $v_\gamma + \kappa e_n$ est inférieur à $b\kappa$. On a également $\kappa' < \kappa$, donc quelque soit la parité de γ, l'élément $v_\gamma + \kappa e_n$ est supérieur à l'extrémité $v_\gamma + (\kappa'+1)e_n$ ou $v_\gamma + \kappa' e_n$. Ainsi ces deux chaînes peuvent être reliée par une arête et former une chaîne dont l'ensemble des sommets est $\mathcal{T} \amalg \mathcal{S} \amalg \mathcal{P} \amalg \mathcal{R} = M_n(t) \smallsetminus \mathcal{I}$ et dont les extrémités sont $\lfloor t/\nu_1 \rfloor\, e_1$ et $\lfloor t/\nu_n \rfloor\, e_n$. Une telle construction est possible pour tout $t \geqslant t_n$, et pour tout $\mathcal{I} \subset M_n(t - t_n)$. Le monoïde libre M possède donc la propriété $\mathfrak{P}_n(t_n)$. □

La preuve présentée ici est structurellement similiaire à celle de la proposition 6.2.1. Techniquement, le gain réalisé par cette nouvelle preuve provient d'un changement de forme de la réserve. Alors que la réserve était de dimension 1 dans la preuve de Mazet, celle-ci est (dans le cadre où on veut montrer que le $\mathcal{M}_n(t)$ est hamiltonien) essentiellement de dimension $n-1$. On ajoute également à la « technique des points de colles » de Mazet une nouvelle technique dite « du serpent ». On pourrait parvenir à une amélioration s'il était possible de placer la réserve herpétomorphe au-dessus des strates. Cela pose cependant un problème technique nouveau, qu'il faut encore résoudre.

Application au graphe divisoriel. — Cette proposition 6.3.1 permet précisément de montrer que le graphe $\mathcal{M}_n(t)$ est hamiltonien pour tout $t \geqslant 5\big(\sum_{i=1}^{n} 2^{n-i}\nu_i\big)$ et l'on ne peut trouver mieux (à la constante multiplicative près) à partir de cette proposition.

Posons $c \coloneqq \sum_n 2^{-n} \log(p_n) \approx 1,\, 105\,568\,246$. Dans le cadre du graphe divisoriel, la proposition 6.3.1 permet d'établir le caractère hamiltonien de $\mathcal{D}(x; p_n)$ pour $x \geqslant \exp(5c2^n)$, et ainsi on peut prouver la proposition suivante à l'aide de la proposition 6.1.1.

Proposition 6.3.2. — *Il existe un réel $C > 0$ tel que*

$$H\big(\mathcal{D}(x)\big) = \big(C + O(1/\log_2 x \log_3 x)\big)x.$$

La conjecture 6.3.1 permettrait d'obtenir le terme d'erreur suivant

$$H\big(\mathcal{D}(x)\big) = \big(C + O(1/\log x \log_2 x)\big)x.$$

APPENDICE A

AUTOUR DES NOMBRES DE STIRLING

Dans cette section nous allons introduire quelques notions classiques de combinatoire, notamment les nombres de Stirling. Nous en faisons à dessein une présentation très énumérative : le lecteur intéressé pourra se référer au chapitre correspondant de [**5**], ou encore à l'ouvrage [**13**], auquel nous empruntons également les notations*.

A.1. Permutations et partitions

Soit E un ensemble de cardinal n. On construit deux types d'objets sur E :

- l'ensemble des permutations des éléments de E, que l'on note $\mathfrak{S}(E)$, où $\mathfrak{S}_n$ si $E = [\![1, n]\!]$; on peut le voir comme l'ensemble des bijections de E dans lui-même, et la composition lui confère une structure de groupe : on l'appelle *le groupe symétrique* ;
- l'ensemble des partitions des E, que l'on note $\Pi(E)$, où Π_n si $E = [\![1, n]\!]$; on peut le voir comme l'ensemble des relations d'équivalence sur E, et la conjonction et la disjonction des relations lui confèrent une structure de treillis : on l'appelle classiquement *le treillis des partitions.*

On peut décomposer chaque permutation en un produit de cycles à supports disjoints, de façon unique à l'ordre des termes près. Il revient au même de décomposer E en orbites et de considérer les cycles induits par la permutation sur chacune de ces orbites. Ainsi, on dispose d'une flèche

$$\mathfrak{S}(E) \to \Pi(E)$$

qui à une permutation associe la partition formée par ses orbites. Cette flèche est surjective car pour une partition de E on peut définir sur chacune de ses parties un cycle qui définira une permutation antécédente par cette flèche.

On note $\left[\begin{smallmatrix} n \\ k \end{smallmatrix}\right]$ le nombre de permutations de $\mathfrak{S}_n$ possédant k orbites et $\left\{\begin{smallmatrix} n \\ k \end{smallmatrix}\right\}$ le nombres de partitions de $[\![1, n]\!]$ en k parties. Il s'agit ici des nombres de Stirling de première et

*Les auteurs de cet ouvrage attribuent à Jovan Karamata la paternité de cette notation.

seconde espèces. La flèche surjective précédemment définie envoyant une permutation à k orbites sur une partition en k parties, on a la relation pour tout $n \geqslant 0$ et tout $k \geqslant 0$

$$\begin{bmatrix} n \\ k \end{bmatrix} \geqslant \begin{Bmatrix} n \\ k \end{Bmatrix}. \tag{A.1}$$

Cette relation est vérifiée pour $n = 0$ car on a dans ce cas $\begin{bmatrix} 0 \\ k \end{bmatrix} = \begin{Bmatrix} 0 \\ k \end{Bmatrix} = [k = 0]$. On a également

$$\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} = \operatorname{card} \mathfrak{S}_n = n!. \tag{A.2}$$

Le cardinal de Π_n, qui vaut lui $\sum_{k=0}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix}$, est le n^{e} nombre de Bell.

TABLE 1. Premières valeurs de $\begin{bmatrix} n \\ k \end{bmatrix}$

$k \diagdown n$	0	1	2	3	4	5	6	7	8	9
0	1									
1		1								
2		1	1							
3		2	3	1						
4		6	11	6	1					
5		24	50	35	10	1				
6		120	274	225	85	15	1			
7		720	1764	1624	735	175	21	1		
8		5040	13068	13132	6769	1960	322	28	1	
9		40320	109584	118124	67284	22449	4536	546	36	1

TABLE 2. Premières valeurs de $\begin{Bmatrix} n \\ k \end{Bmatrix}$

$k \diagdown n$	0	1	2	3	4	5	6	7	8	9
0	1									
1		1								
2		1	1							
3		1	3	1						
4		1	7	6	1					
5		1	15	25	10	1				
6		1	31	90	65	15	1			
7		1	63	301	350	140	21	1		
8		1	127	966	1701	1050	266	28	1	
9		1	255	3025	7770	6951	2646	462	36	1

A.2. Propriétés analytiques

On a les relations de récurrence suivantes pour $n \geqslant 0$ et $k \geqslant 1$

$$\text{(A.3)} \qquad \begin{bmatrix} n+1 \\ k \end{bmatrix} = n\begin{bmatrix} n \\ k \end{bmatrix} + \begin{bmatrix} n \\ k-1 \end{bmatrix} \quad \text{et} \quad \begin{Bmatrix} n+1 \\ k \end{Bmatrix} = k\begin{Bmatrix} n \\ k \end{Bmatrix} + \begin{Bmatrix} n \\ k-1 \end{Bmatrix}.$$

La relation pour les nombres de Stirling de seconde espèce reflète le fait que l'on peut construire une partition en k parties de $[\![1, n+1]\!]$ de deux façons :

- soit en ajoutant à une partition en $k-1$ parties de $[\![1, n]\!]$ le singleton $\{n+1\}$ comme k^{e} partie, ce qui représente $\left\{ {n \atop k-1} \right\}$ partitions ;
- soit en partant d'une partition en k parties de $[\![1, n]\!]$ et en ajoutant $n+1$ à l'une des k parties de cette partition, cela fait k choix pour chacune des $\left\{ {n \atop k} \right\}$ partitions possibles.

De même, on peut expliquer la relation sur les nombres de Stirling de première espèce en construisant une permutation à k cycles $\sigma' \in \mathfrak{S}_{n+1}$ à partir d'une permutation $\sigma \in \mathfrak{S}_n$:

- soit en choisissant σ à $k-1$ cycles et en rajoutant le cycle trivial $[n+1]$, ce qui donne
$$\sigma'(i) = \begin{cases} \sigma(i) & \text{si } i \in [\![1, n]\!], \\ n+1 & \text{si } i = n+1, \end{cases}$$
et cela représente $\left[{n \atop k-1} \right]$ premutations ;
- soit en choisissant σ à k cycles et en intercalant $n+1$ dans l'un de ces cycles, ce qui revient à choisir $j \in [\![1, n]\!]$ et à définir σ' par
$$\sigma'(i) = \begin{cases} \sigma(i) & \text{si } i \in [\![1, n]\!] \text{ et } i \neq j, \\ n+1 & \text{si } i = j, \\ \sigma(j) & \text{si } i = n+1, \end{cases}$$
et cela fait n choix pour chacune des $\left[{n \atop k} \right]$ permutations σ possibles.

Ces relations (A.3) induisent des équations différentielles sur les fonctions génératrices exponentielles (f.g.e.) des nombres de Stirling $s_k(x) = \sum_n \left[{n \atop k} \right] \frac{x^n}{n!}$ et $S_k(x) = \sum_n \left\{ {n \atop k} \right\} \frac{x^n}{n!}$. On rappelle des propriétés basiques des f.g.e. : si on pose $A(x) = \sum_n a_n \frac{x^n}{n!}$, on a $A'(x) = \sum_n a_{n+1} \frac{x^n}{n!}$ et $xA(x) = \sum_n n a_{n-1} \frac{x^n}{n!}$. Ainsi, on a pour $k \geqslant 1$

$$\text{(A.4)} \qquad s_k'(x) = x s_k'(x) + s_{k-1}(x) \quad \text{et} \quad S_k'(x) = k S_k(x) + S_{k-1}(x)$$

avec $s_k(0) = S_k(0) = 0$, mais aussi avec $s_0(x) = S_0(x) = 1$. Il alors facile de vérifier que l'on a

$$\text{(A.5)} \qquad \sum_{n \geqslant 0} \begin{bmatrix} n \\ k \end{bmatrix} \frac{x^n}{n!} = \frac{1}{k!}\big(-\log(1-x)\big)^k \quad \text{et} \quad \sum_{n \geqslant 0} \begin{Bmatrix} n \\ k \end{Bmatrix} \frac{x^n}{n!} = \frac{1}{k!}\big(e^x - 1\big)^k.$$

A.3. Inversions binomiale et de Stirling

Soit X un ensemble de cardinal t. On considère l'ensemble produit X^n de cardinal t^n, et l'ensemble des arrangements $X^{\underline{n}} = \{(x_1, \ldots, x_n);\ \forall i \neq j, x_i \neq x_j\}$ de

cardinal $t^{\underline{n}} = \prod_{i=0}^{n-1}(t-i) = n!\binom{t}{n}$. À un vecteur $\mathbf{x} = (x_1, \ldots, x_n)$ de X^n, on peut associer une partition de l'ensemble $[\![1, n]\!]$ des indices où i et j sont dans la même partie si et seulement si $x_i = x_j$. Cela fournit une application de X^n dans Π_n. Le nombre d'antécédents d'une partition $\pi \in \Pi_n$ en k parties vaut alors $t^{\underline{k}}$. On a donc la relation

$$t^n = \sum_{k=0}^{n} \left\{ n \atop k \right\} t^{\underline{k}} = \sum_{k=0}^{n} \left\{ n \atop k \right\} k! \binom{t}{k}. \tag{A.6}$$

On peut interpréter ce résultat en terme d'applications d'un ensemble E de cardinal n dans l'ensemble X de deux façons différentes. D'abord en suivant le même schéma de raisonnement que précédemment, mais en remarquant que X^n est isomorphe à l'ensemble des applications de E dans X et que $X^{\underline{n}}$ est isomorphe à l'ensemble des injections de E dans X. Le raisonnement précédent peut alors se condenser en l'identité suivante

$$F(E, X) = \coprod_{\pi \in \Pi(E)} I(\pi, X),$$

où $F(E, X)$ (respectivement $I(E, X)$) est l'ensemble des applications (respectivement des injections) de E dans X.
On peut effectuer un comptage équivalent en utilisant des surjections. En effet, on a

$$F(E, X) = \coprod_{Y \subset X} S(E, Y),$$

où $S(E, X)$ est l'ensemble des surjections de E dans Y. On note k le cardinal de Y. À une surjection $s \in S(E, Y)$, on peut associer la partition $\pi_s = \{s^{-1}(y); y \in Y\} \in \Pi(E)$ en k parties. Une surjection de E dans Y correspond à une partition de E dont les k parties sont indexées par les éléments de Y : on a donc

$$\operatorname{card} S(E, Y) = k! \left\{ n \atop k \right\}. \tag{A.7}$$

Cette relation permet de retrouver la formule (A.6), mais permet également de donner une première estimation des nombres de Stirling puisque

$$\operatorname{card} S(E, Y) = k! \left\{ n \atop k \right\} \leqslant k^n = \operatorname{card} F(E, Y). \tag{A.8}$$

On peut améliorer cette estimation en une identité, au moyen de l'inversion binomiale. Cette inversion, avatar du principe d'inclusion-exclusion, affirme que

$$\Big(\forall t \in \mathbb{N},\ f_t = \sum_k \binom{t}{k} g_k\Big) \Longleftrightarrow \Big(\forall k \in \mathbb{N},\ g_k = \sum_t (-1)^{k-t} \binom{k}{t} f_t\Big).$$

Cette identité est basée sur l'égalité $\sum_k (-1)^{n-k} \binom{n}{k} \binom{k}{t} = [n = t]$ et permet d'inverser la formule (A.6) en

$$k! \left\{ n \atop k \right\} = \sum_{t=0}^{k} (-1)^{k-t} \binom{k}{t} t^n. \tag{A.9}$$

Cette expression a donc été obtenue en « inversant » les coefficients binomiaux. Il est en fait possible de réaliser une inversion mais cette fois sur les nombres de Stirling de seconde espèce.

On considère à nouveau les ensembles X^n et $X^{\underline{n}}$. On fait agir $\mathfrak{S}_n$ sur X^n par permutations des coordonnées. Pour un élément $\mathbf{x} \in X^n$, on considère son stabilisateur $H_{\mathbf{x}} = \{\sigma \in \mathfrak{S}_n;\ \sigma \cdot \mathbf{x} = \mathbf{x}\}$, sous-groupe de $\mathfrak{S}_n$. On note $\varepsilon(\sigma)$ la signature de la permutation $\sigma \in \mathfrak{S}_n$. On a les propriétés suivantes :

– si $\mathbf{x} \in X^{\underline{n}}$, on a $H_{\mathbf{x}} = \{1\}$, et donc

$$\sum_{\sigma \in H_{\mathbf{x}}} \varepsilon(\sigma) = 1;$$

– si $\mathbf{x} \notin X^{\underline{n}}$, il y au moins une transposition τ dans $H_{\mathbf{x}}$, et on a donc

$$\sum_{\sigma \in H_{\mathbf{x}}} \varepsilon(\sigma) = \sum_{\sigma \in H_{\mathbf{x}}} \varepsilon(\tau\sigma) = \varepsilon(\tau) \sum_{\sigma \in H_{\mathbf{x}}} \varepsilon(\sigma) = -\sum_{\sigma \in H_{\mathbf{x}}} \varepsilon(\sigma);$$

ainsi, on a

$$\sum_{\sigma \in H_{\mathbf{x}}} \varepsilon(\sigma) = 0.$$

Ces deux cas peuvent se résumer en une unique relation

$$\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)[\sigma \cdot \mathbf{x} = \mathbf{x}] = [\mathbf{x} \in X^{\underline{n}}].$$

Les éléments de X^n fixés par une permutation σ ont les valeurs de leur coordonnées égales pour des indices appartenant à la même orbite de σ. Ainsi, si σ possède k orbites (ce qui représente $\begin{bmatrix} n \\ k \end{bmatrix}$ cas), il existe exactement t^k éléments de X^n fixés par σ, et on a $\varepsilon(\sigma) = (-1)^{n-k}$. Ainsi on a

(A.10)

$$n!\binom{t}{n} = t^{\underline{n}} = \sum_{\mathbf{x} \in X^n} [\mathbf{x} \in X^{\underline{n}}] = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \sum_{\mathbf{x} \in X^n} [\sigma \cdot \mathbf{x} = \mathbf{x}] = \sum_{k=0}^{n} (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} t^k.$$

On obtient bien ici une inversion de la formule (A.6), mais cette fois-ci par inversion du nombre de Stirling de seconde espèce en nombre de Stirling de première espèce. Les relations (A.6) et (A.10) pouvant être vues comme des identités polynomiales sur $\mathbb{R}[t]$, on obtient rapidement les formules d'inversion

$$\text{(A.11)} \qquad \sum_k (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} \begin{Bmatrix} k \\ m \end{Bmatrix} = [m = n] \quad \text{et} \quad \sum_k (-1)^{k-m} \begin{bmatrix} k \\ m \end{bmatrix} \begin{Bmatrix} n \\ k \end{Bmatrix} = [m = n],$$

qui permettent de formaliser l'inversion de Stirling

$$\text{(A.12)} \qquad \Big(\forall n \in \mathbb{N},\ f_n = \sum_k \begin{Bmatrix} n \\ k \end{Bmatrix} g_k\Big) \iff \Big(\forall k \in \mathbb{N},\ g_k = \sum_n (-1)^{k-n} \begin{bmatrix} k \\ n \end{bmatrix} f_n\Big).$$

A.4. Partages et partitions

Un partage de l'entier n en k parts est la donnée d'entiers strictement positifs $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_k$ vérifiant $\sum_{i=1}^{k} \lambda_i = n$.

On considère l'action de $\mathfrak{S}_n$ sur $[\![1, n]\!]$; celle-ci induit une action sur l'ensemble des parties de $[\![1, n]\!]$, dont les orbites sont caractérisées par le cardinal des parties, à savoir que deux parties de $[\![1, n]\!]$ sont dans la même orbite sous $\mathfrak{S}_n$ si et seulement si elles ont même cardinal. Cette action sur les parties induit une action sur les partitions de $[\![1, n]\!]$; les orbites sous cette action sont caractérisées par les cardinaux des parties de la partition. Une partition est de partage cardinal $\boldsymbol{\lambda}$, où $\boldsymbol{\lambda}$ est un partage de n, si les cardinaux des parties de la partition, ordonnés par ordre décroissant, sont les parts λ_i. Deux partitions de Π_n sont dans la même orbite sous l'action de $\mathfrak{S}_n$ si et seulement si elles ont même partage cardinal[†].

La décroissance des parts d'un partage de n est un artifice pratique pour concrétiser le fait que les parties d'une partition ne sont pas étiquetées[‡]. Pour les partitions étiquetées, l'action de $\mathfrak{S}_n$ est la même, mais les orbites sont caractérisées par les partages cardinaux *étiquetés* des partitions : deux partitions étiquetées $(E_1, \ldots, E_k)$ et $(E'_1, \ldots, E'_{k'})$ sont dans la même orbite si et seulement si $k = k'$ et pour tout $0 \leqslant i \leqslant k$, on a $\operatorname{card} E_i = \operatorname{card} E'_i$; leur partage cardinal commun est le partage de n en parts étiquetées $(\operatorname{card} E_1, \ldots, \operatorname{card} E_n)$.

Un partage de l'entier n en k parts étiquetées est la donnée du vecteur d'entiers strictement positifs $(l_1, \ldots, l_k)$ vérifiant $\sum_{i=1}^{k} l_i = n$. Les partages étiquetés sont donc bien plus facile à énumérer que ceux qui ne le sont pas : par exemple, le nombre de partitions étiquetées de partage cardinal $\boldsymbol{l}$ est le coefficient multinomial $\frac{n!}{l_1! \cdots l_k!} = \binom{n}{\boldsymbol{l}}$. Ainsi, en sommant sur les partages étiquetés de n en k parts, on obtient le nombre de partitions étiquetées en k parties, soit

$$k! \left\{ {n \atop k} \right\} = \sum_{\substack{l_i \geqslant 1 \\ (1 \leqslant i \leqslant k) \\ \sum_i l_i = n}} \binom{n}{l_1, \ldots, l_k}. \tag{A.13}$$

Il est alors facile à partir de cette formule de retrouver l'estimation (A.8).

[†] Il est à remarquer que le partage cardinal d'une partition n'est pas le nombre de parties de cette partition.

[‡] On emploie également « colorées », ou même « ordonnées ».

APPENDICE B

MOYENNES DE SOMMES DE RAMANUJAN

Le contenu de cet appendice à fait l'objet d'une note [**3**] aux Comptes-Rendus de l'Académie des Sciences.

On s'intéresse ici aux moyennes des sommes de Ramanujan* sur les intervalles. Le problème n'étant pas pertinent pour les sommes $c_q(n)$ avec $q = 1$, on supposera $q > 1$. On définit la fonction de sommation partielle pour $x \in \mathbb{R}_+$

$$s_q(x) \coloneqq \sum_{0<m\leqslant x} c_q(m).$$

Cette fonction est q-périodique, et se définit donc pour tout $x \in \mathbb{R}$. On s'intéresse donc à l'amplitude de ses variations

$$A(q) \coloneqq \sup_{x<y} \Big| \sum_{x<m\leqslant y} c_q(m) \Big| = \sup_{x<y} \big| s_q(y) - s_q(x) \big|.$$

B.1. Propriétés élémentaires

La fonction s_q a hérité des sommes de Ramanujan une série de propriétés que nous listons sans démonstration.

Propriété 1. — *Pour $x \in \mathbb{R} \smallsetminus \mathbb{Z}$, on a $s_q(x) + s_q(-x) + \varphi(q) = 0$.*

Propriété 2. — *Pour $\alpha \in \mathbb{R}$, on a $s_q(q\alpha)/q = s_{\bar{q}}(\bar{q}\alpha)/\bar{q}$, où $\bar{q}$ est le noyau sans facteur carré de q.*

Propriété 3. — *Pour $x \in \mathbb{R}$, on a*

$$s_q(x) = \sum_{d|q} \mu\Big(\frac{q}{d}\Big) d \, \Big\lfloor \frac{x}{d} \Big\rfloor = -\sum_{d|q} \mu\Big(\frac{q}{d}\Big) d \, \Big\{ \frac{x}{d} \Big\},$$

*La définition et quelques propriétés de ces sommes ont déjà été énoncées au chapitre 4, équation (4.1) *sqq*.

et pour tout q' entier premier à q, on a

$$s_{q'q}(x) = \sum_{d'|q'} \mu\Big(\frac{q'}{d'}\Big) d' s_q\Big(\frac{x}{d'}\Big).$$

De ces propriétés nous déduisons immédiatement une série de conséquences sur la fonction A.

Proposition B.1.1. — *Soit $q > 1$ un entier.*

a). *On a l'encadrement $\varphi(q) \leqslant A(q) \leqslant \sigma(q)$.*
b). *On a la relation $A(q)/q = A(\bar{q})/\bar{q}$.*
c). *Pour tout q' entier premier à q, on a $\varphi(q')A(q) \leqslant A(q'q) \leqslant \sigma(q')A(q)$.*

L'encadrement trivial du *a)* aurait été suffisant pour les applications futures ; l'on va cependant préciser cet encadrement.

Démonstration. — Il est clair que $A(q) \geqslant c_q(0) = \varphi(q)$ et par la propriété 3, on a

$$|s_q(y) - s_q(x)| = \Big|\sum_{d|q} \mu\Big(\frac{q}{d}\Big) d\Big(\Big\{\frac{y}{d}\Big\} - \Big\{\frac{x}{d}\Big\}\Big)\Big| \leqslant \sum_{d|q} d = \sigma(d).$$

Le point *b)* se déduit immédiatement de la propriété 2.
De la propriété 3, on tire immédiatement

$$|s_{q'q}(y) - s_{q'q}(x)| = \Big|\sum_{d'|q'} \mu\Big(\frac{q'}{d'}\Big) d'\big(s_q(x/d') - s_q(y/d')\big)\Big| \leqslant \sum_{d'|q'} d' A(q) = \sigma(q')A(q).$$

Mais on a aussi en choisissant q' comme un entier premier p ne divisant pas q

$$\Big|p\big(s_q(x/p) - s_q(y/p)\big)\Big| \leqslant \big|s_{pq}(x) - s_{pq}(y)\big| + \big|s_q(x) - s_q(y)\big| \leqslant A(pq) + A(q).$$

Ainsi, on a bien $A(pq) \geqslant (p-1)A(q)$. En itérant ce raisonnement sur les facteurs premiers de q', et en utilisant le point *b)*, on obtient la conclusion attendue. □

Malgré sa simplicité, la proposition B.1.1.*c)* est particulièrement importante. Elle est, avec la relation $A\big(\prod_{p\leqslant y} p\big) \ll \prod_{p\leqslant y} p$[†], l'ingrédient principal du théorème B.3.3. Cette dernière relation s'établit grâce à un puissant théorème de P-convergence dû à la Bretèche et Tenenbaum [**2**, théorème 2.1].

B.2. Séries de Fourier et minoration

La propriété 1 énonçant la symétrie de la fonction s_q nous conduit à définir la fonction $\tilde{s}_q$ par $\tilde{s}_q(x) = \frac{1}{2}\big(s_q(x) - s_q(-x)\big)$, qui est une translation de $s_q(x)$ sur sa moyenne $-\frac{1}{2}\varphi(q)$ pour $x \notin \mathbb{Z}$. Pour $x \in \mathbb{Z}$, donc aux points éventuels de discontinuité, on a $2\tilde{s}_q(x) = \tilde{s}_q(x^+) + \tilde{s}_q(x^-)$. On a donc $A(q) = 2\sup|\tilde{s}_q|$ et que $\tilde{s}_q$ est somme de sa série de Fourier.

[†] *cf.* proposition B.3.2 *infra.*

Lemme B.2.1. — *Soient $q > 1$ un entier et α un réel. On a*

$$\frac{\tilde{s}_q(q\alpha)}{q} = -\sum_{d|q} \frac{\mu(d)}{d} \mathrm{B}_1(d\alpha) = \sum_{\substack{n\geqslant 1\\(n,q)=1}} \frac{\sin 2\pi n\alpha}{\pi n}.$$

La seconde somme n'est pas absolument convergente ; sauf mention contraire, l'ordre de sommation dans ce genre de somme est l'ordre croissant.

Démonstration. — La première identité s'obtient immédiatement de la propriété 3. Le fait de connaître la série de Fourier de la première fonction de Bernoulli

$$\mathrm{B}_1(x) = \sum_{n\geqslant 1} \frac{\sin 2\pi nx}{\pi n}$$

permet de calculer par linéarité les termes de la série de Fourier de $\tilde{s}_q$. Comme cette dernière est convenablement normalisée, elle est somme de sa série de Fourier. □

Davenport propose dans [**8**] de déterminer l'éventuel sens analytique de la relation formelle

$$\sum_{n=1}^{\infty} \mathrm{B}_1(n\alpha)\, \frac{f(n)}{n} = -\sum_{m=1}^{\infty} g(m) \frac{\sin 2\pi m\alpha}{\pi m},$$

où les fonctions f et g sont liées par $g(n) = \sum_{d|n} f(d)$. Le lemme B.2.1 n'énonce que la convergence ponctuelle d'un cas très simple d'identité de Davenport, associée au couple de fonctions (δ_q, μ_q), où $\delta_q(n) = \delta\big((n,q)\big)$ et où $\mu_q = \delta_q * \mu$, c'est-à-dire $\mu_q(n) = \mu(n)$ si $n \mid q$ et $\mu_q(n) = 0$ sinon. Ceci permet tout de même d'obtenir une minoration.

Proposition B.2.2. — *Pour tout entier $q > 1$ on a*

$$A(q) \geqslant \sqrt{\frac{\sigma(q)\varphi(q)}{3}}.$$

En particulier, on a $A(q) \gg q$.

Démonstration. — On a par la formule de Parseval et le lemme B.2.1

$$\frac{A(q)^2}{4q^2} \geqslant \Big(\frac{\sup|\tilde{s}_q|}{q}\Big)^2 \geqslant \int_0^1 \Big|\frac{\tilde{s}_q(q\alpha)}{q}\Big|^2 \,\mathrm{d}\alpha = \sum_{\substack{n\geqslant 1\\(n,q)=1}} \frac{1}{2\pi^2 n^2} = \frac{\sigma(2)}{2\pi^2} \prod_{p|q} \Big(1 - \frac{1}{p^2}\Big). \quad \square$$

B.3. P-convergence et majoration

Pour obtenir de meilleures informations sur $\tilde{s}_q$, nous allons prouver la P-convergence uniforme de sa série de Fourier.

Le théorème de P-convergence de [**2**, théorème 2.1], qui établit uniformément en $y \geq 2$ et en $\alpha \in \mathbb{R}$

$$\sum_{P^+(n)\leq y} \frac{\mu(n)}{n} \mathrm{B}_1(n\alpha) + \frac{\sin(2\pi\alpha)}{\pi} \ll \frac{1}{\log y}, \tag{B.1}$$

se traduit directement, grâce au lemme B.2.1, en termes de la fonction $\tilde{s}_q$.

Proposition B.3.1. — *On a uniformément pour $q > 1$, $\alpha \in \mathbb{R}$ et $y \geqslant 2$*

$$\frac{\tilde{s}_q(q\alpha)}{q} = \sum_{\substack{(n,q)=1\\ P^+(n)\leqslant y}} \frac{\sin(2\pi n\alpha)}{\pi n} + O\Big(\prod_{\substack{p\nmid q\\ p\leqslant y}} \Big(1-\frac{1}{p}\Big)^{-1} \frac{1}{\log y}\Big).$$

Démonstration. — On remarque avec la Bretèche et Tenenbaum [**2**, (1.22)] que pour toute fonction multiplicative f, on a sous réserve de convergence

$$\sum_{P^+(m)\leqslant y} \frac{f(m)}{\pi m}\sin(2\pi m\alpha) + \sum_{P^+(n)\leqslant y} \frac{(f*\mu)(n)}{n}\mathrm{B}_1(n\alpha)$$
$$= \sum_{P^+(m)\leqslant y} \frac{f(m)}{m}\left(\frac{\sin(2\pi m\alpha)}{\pi} + \sum_{P^+(n)\leqslant y} \frac{\mu(n)}{n}\mathrm{B}_1(nm\alpha)\right),$$

et en utilisant l'équation (B.1) pour traiter le terme entre parenthèses,

$$\ll \sum_{P^+(m)\leqslant y} \frac{|f(m)|}{m}\frac{1}{\log y}.$$

Il suffit à présent de spécialiser $f = \delta_q$ pour obtenir la relation annoncée. □

Corollaire B.3.2. — *Posons $q_y = \prod_{p\leqslant y} p$. On a $A(q_y) \sim \frac{2}{\pi} q_y$ pour $y \to +\infty$.*

Il s'agit d'une conséquence directe de l'équation (B.1) et du lemme B.2.1.

Théorème B.3.3. — *On a uniformément pour $q > 1$*

$$A(q) \leqslant \big(1+o(1)\big) q \min\Big(\prod_{\substack{p\nmid q\\ p\leqslant \log q}} \Big(1-\frac{1}{p}\Big)^{-1}, \frac{2}{\pi}\prod_{\substack{p\mid q\\ p\leqslant \log q}} \Big(1-\frac{1}{p}\Big)^{-1}\Big).$$

En particulier, on a $A(q) \leqslant \big(\sqrt{2e^\gamma/\pi} + o(1)\big) q\sqrt{\log_2 q}$.

On a $\sqrt{2e^\gamma/\pi} \approx 1,064\,831\,403$.

Lemme B.3.4. — *On a uniformément pour $q \geqslant 3$ entier*

$$\prod_{\substack{p\mid q\\ \log q<p}} \Big(1-\frac{1}{p}\Big)^{-1} = 1 + O\Big(\frac{1}{\log_2 q}\Big).$$

Nous n'avons pas cherché ici le meilleur terme d'erreur possible. Il est probable qu'avec un peu plus de travail, on puisse l'améliorer de façon significative.

Démonstration. — Pour tout x vérifiant $\pi(x) - \pi(\log q) \geqslant \omega(q)$, on a la majoration

$$\prod_{\substack{p\mid q\\ \log q<p}} \Big(1-\frac{1}{p}\Big)^{-1} \leqslant \prod_{\log q<p\leqslant x} \Big(1-\frac{1}{p}\Big)^{-1}.$$

On sait grâce aux encadrements de Tchébychev qu'il existe deux constantes $0 < c_1 < 1 < c_2$ telles que $c_1 y/\log y \leqslant \pi(y) \leqslant c_2 y/\log y$ pour $y \geqslant 2$. Il existe également une constante $c_3 > 1$ telle que $\omega(q) \leqslant c_3 \log q/\log_2 q$. En prenant $x = \frac{c_2+c_3}{c_1}\log q$, on a

$$\pi(x) \geqslant c_1 \frac{x}{\log x} \geqslant (c_2+c_3)\frac{\log q}{\log_2 q} \geqslant \pi(\log q) + \omega(q).$$

Ainsi, posant $C = \frac{c_2+c_3}{c_1}$, on a par la formule de Mertens

$$\prod_{\log q < p \leqslant C\log q} \Big(1-\frac{1}{p}\Big)^{-1} = \frac{\log\big(C\log q\big)}{\log_2(q)}\Big(1+O\Big(\frac{1}{\log_2 q}\Big)\Big) = 1+O\Big(\frac{1}{\log_2 q}\Big). \qquad \square$$

Démonstration du théorème B.3.3. — On peut, par la proposition B.1.1.*b)*, supposer q sans facteur carré. En utilisant la proposition B.1.1.*a)* puis le lemme B.3.4, on a

$$A(q) \leqslant \sigma(q) \leqslant q\prod_{p\mid q}\Big(1-\frac{1}{p}\Big)^{-1} \sim q \prod_{\substack{p\mid q\\ p\leqslant \log q}}\Big(1-\frac{1}{p}\Big)^{-1}.$$

D'un autre côté, en posant $q^\sharp = \prod_{p>\log q}^{p\mid q} p$, $q^\flat = \prod_{p\leqslant \log q}^{p\mid q} p$ et $Q = \prod_{p\leqslant \log q} p$, on a grâce à la proposition B.1.1.*c)*, à la proposition B.3.1 et au lemme B.3.4

$$A(q) \leqslant \sigma(q^\sharp)A(q^\flat) \leqslant \sigma(q^\sharp)\frac{A(Q)}{\varphi(Q/q^\flat)} \leqslant \big(\tfrac{2}{\pi}+o(1)\big)q \prod_{\substack{p\nmid q\\ p\leqslant \log q}}\Big(1-\frac{1}{p}\Big)^{-1}.$$

On en déduit notamment que

$$A(q) \leqslant \big(1+o(1)\big)q\Big(\tfrac{2}{\pi}\prod_{p\leqslant \log q}\big(1-\tfrac{1}{p}\big)^{-1}\Big)^{1/2} \sim \sqrt{\tfrac{2e^\gamma}{\pi}}\, q\sqrt{\log_2 q}$$

par la formule de Mertens. $\square$

B.4. Quelques exemples

Nous commençons par donner une version réduite du théorème, mais qui nous paraît plus « visuelle ». On dit qu'un ensemble d'entiers premiers $\mathcal{E}$ possède une densité logarithmique forte $\kappa \in [0,1]$ (ou plus sommairement, que $\mathcal{E}$ est de densité κ) si l'on a $\sum_{p\leq y}^{p\in\mathcal{E}} p^{-1} = \kappa \log_2 y + O(1)$ pour $y \to +\infty$.

Proposition B.4.1. — *Soit $\mathcal{E} \subset \mathcal{P}$ un ensemble d'entiers premiers de densité $\kappa \in [0,1]$. Posons $q_y = \prod_{p\leq y}^{p\in\mathcal{E}} p$. On a $A(q_y) \ll q_y\big(\log_2 q_y\big)^{\min(\kappa,1-\kappa)}$.*

Nous montrons que pour certains ensembles d'entiers premiers bien choisis, la proposition B.4.1 est essentiellement optimale.

Proposition B.4.2. — *Soit $n \geqslant 3$ un entier.*

a). *Posons $q_y = \prod_{p\equiv -1\,[n]}^{p\leqslant y} p$. On a $A(q_y) \asymp q_y\big(\log_2 q_y\big)^{1/\varphi(n)}$ pour $y \geqslant 3$.*

b). *Posons $q_y = \prod_{p\not\equiv 1\,[n]}^{p\leqslant y} p$. On a $A(q_y) \asymp q_y\big(\log_2 q_y\big)^{1/\varphi(n)}$ pour $y \geqslant 3$.*

Lemme B.4.3. — *Soit $n \geqslant 3$ un entier.*

a). *Soit $p_1 < p_2 < \ldots < p_k$ des entiers premiers congrus à -1 modulo n. Posons $q = p_1 p_2 \cdots p_k$. On a pour tout $a \in \mathbb{Z}$*

$$\tilde{s}_q\Big(q\frac{a}{n}\Big) = -\sigma(q)\mathrm{B}_1\Big(\frac{a}{n}\Big).$$

b). *Soit $p_1 < p_2 < \ldots < p_k$ des entiers premiers congrus à 1 modulo n. Posons $q = \big(\prod_{p\leqslant p_k} p\big)/(p_1 p_2 \cdots p_k)$. On a pour tout $a \in \mathbb{Z}$*

$$\tilde{s}_q\big(q\frac{a}{n}\big) = \frac{1}{\pi}\, q \prod_{i=1}^{k}\Big(1-\frac{1}{p_i}\Big)^{-1}\bigg(\sin\Big(2\pi\frac{a}{n}\Big) + O\Big(\frac{1}{\log p_k}\Big)\bigg),$$

où la constante implicite est absolue pour n fixé.

Démonstration. — Les deux parties se démontrent de façon similaire et symétrique.

a). Du lemme B.2.1, on a pour tout $a \in \mathbb{Z}$

$$\tilde{s}_q\Big(q\frac{a}{n}\Big) = -\sum_{d|q}\mu(d)\frac{q}{d}\mathrm{B}_1\Big(d\frac{a}{n}\Big).$$

Puisque $p_i \equiv -1\ [n]$, on a $d \equiv \mu(d)\ [n]$ pour tout diviseur d de q, et donc $d\frac{a}{n} \equiv \mu(d)\frac{a}{n}\ [1]$. Par imparité et 1-périodicité de la première fonction de Bernoulli B_1, on obtient

$$\tilde{s}_q\Big(q\frac{a}{n}\Big) = -\sum_{d|q}\frac{q}{d}\mathrm{B}_1\Big(\frac{a}{n}\Big) = -\sigma(q)\mathrm{B}_1\Big(\frac{a}{n}\Big).$$

b). De la proposition B.3.1, on a pour tout $a \in \mathbb{Z}$

$$\tilde{s}_q\Big(q\frac{a}{n}\Big) = q\sum_{\substack{(m,q)=1\\ P^+(m)\leqslant p_k}}\frac{\sin(2\pi m\frac{a}{n})}{\pi m} + O\bigg(q\prod_{\substack{p\nmid q\\ p\leqslant p_k}}\Big(1-\frac{1}{p}\Big)^{-1}\frac{1}{\log p_k}\bigg).$$

Les p intervenant dans le produit sont exactement les p_i ; de même, chaque m intervenant dans la somme est composé uniquement de facteurs premiers parmi ces p_i et vérifie donc $m \equiv 1\ [n]$ et donc $m\frac{a}{n} \equiv \frac{a}{n}\ [1]$. On conclut grâce à la 1-périodicité de la fonction $x \mapsto \sin(2\pi x)$. □

Démonstration de la proposition B.4.2. — Si $q_y = \prod_{p\equiv -1\,[n]}^{p\leqslant y} p$, le lemme B.4.3.*a)* spécialisé en $a = 1$, fournit la minoration

$$A(q_y) \gg q_y \prod_{\substack{p\equiv -1\,[n]\\ p\leqslant y}}\Big(1+\frac{1}{p}\Big) \asymp q_y \exp\bigg(\sum_{\substack{p\equiv -1\,[n]\\ p\leqslant \log q_y}}\frac{1}{p}\bigg) \asymp q_y\big(\log_2 q_y\big)^{1/\varphi(n)},$$

où l'on a utilisé le lemme B.3.4 pour la seconde estimation et où la dernière estimation est due au résultat classique[‡]

$$\sum_{\substack{p\leqslant y\\ p\equiv -1\,[n]}}\frac{1}{p} = \frac{1}{\varphi(n)}\log_2 y + O(1).$$

[‡]Voir par exemple Landau [**25**, § 111].

La majoration est due au théorème B.3.3. Le cas $q_y = \prod_{p \not\equiv 1\,[n]}^{p \leqslant y} p$ se traite de façon similaire. □

Ces exemples fournissent donc bien l'optimalité de l'exposant $\min(\kappa, 1-\kappa)$ dans la proposition B.4.1 pour $\kappa = 1/\varphi(n)$ et $\kappa = 1 - 1/\varphi(n)$ pour $n \geqslant 3$, par exemple $\kappa = \frac{1}{6}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, \frac{5}{6}, \ldots$ Mais le lemme B.4.3 permet de choisir plus généralement des sous-ensembles de $\{p,\, p \equiv -1\ [n]\}$ ou des sur-ensembles de $\{p,\, p \not\equiv 1\ [n]\}$. Ainsi, grâce au lemme suivant, on peut construire pour tout $\kappa \in [0,1]$ des ensembles d'entiers premiers de densité κ tels que $A(q_y) \asymp q_y \big(\log_2 q_y\big)^{\min(\kappa, 1-\kappa)}$.

Lemme B.4.4. — *Soit $\mathcal{E} \subset \mathcal{P}$ un ensemble d'entiers premiers de densité logarithmique forte $\kappa \in [0,1]$. Pour tout $0 < \kappa' < \kappa$, il existe $\mathcal{E}' \subset \mathcal{E}$ un ensemble d'entiers premiers de densité logarithmique forte κ'.*

Ce lemme se prouve en définissant récursivement

$$e'_0 \coloneqq \min \mathcal{E}, \quad e'_{n+1} \coloneqq \min \big\{ e \in \mathcal{E},\, e > e'_n,\, \kappa' \log_2 e \geq \textstyle\sum_{i \leq n} 1/e'_i \big\},$$

et en montrant que l'ensemble $\mathcal{E}' \coloneqq \{e'_n;\ n \in \mathbb{N}\}$ possède effectivement la densité κ' voulue.

À l'inverse, il n'est pas clair que pour tout $\kappa \in [0,1]$ on puisse trouver un ensemble de nombres premiers $\mathcal{E}$ vérifiant $\sum_{p \leqslant y}^{p \in \mathcal{E}} p^{-1} = \kappa \log_2 y + O(1)$ avec $A(q_y) \asymp q_y$.

BIBLIOGRAPHIE

[1] R. C. Baker, G. Harman & J. Pintz, The difference between consecutive primes, II, *Proc. London Math. Soc. (3)* **83** (2001), no. 3, 532–562.

[2] R. de la Bretèche & G. Tenenbaum, Séries trigonométriques à coefficients arithmétiques, *J. Anal. Math.* **92** (2004), 1–79.

[3] A. Chadozeau, Une remarque sur les sommes de Ramanujan, *C. R. Math. Acad. Sci. Paris* **341** (2005), no. 7, 399–404.

[4] S. Chowla, On the least prime in an arithmetical progression, *J. Indian Math. Soc. (2)* **1** (1934), 1-3.

[5] L. Comtet, *Analyse combinatoire,* 2 tomes, PUF, Paris, 1970.

[6] H. Cramér, Some theorems concerning prime numbers, *Ark. Mat. Astronom. Fys.* **15** (1920), no. 5, 1–32.

[7] H. Cramér, On the order of magnitude of the difference between consecutive prime numbers, *Acta Arith.* **2** (1936), 23–46.

[8] H. Davenport, On some infinite series involving arithmetical functions, *Quart. J. Math. Oxford,* **8** (1937), 8–13.

[9] P. Erdős, The difference of consecutive primes, *Duke Math.* **6** (1940), 438–441.

[10] P. Erdős, On the integers relatively prime to n and on a number-theoretic function considered by Jacobsthal, *Math. Scand.* **10** (1962), 163–170.

[11] P. Erdős, Problems and results in number theory, *in* : *Recent Progress in Analytic Theory,* vol. 1, éd. : H. Halberstam & C. Hooley, 1–13, Academic Press, Londres, 1981.

[12] P. Erdős & É. Saias, Sur le graphe divisoriel, *Acta Arith.* **73** (1995), 189–198.

[13] R. L. Graham, D. E. Knuth & O. Patashnik, *Concrete Mathematics,* (2e éd.), Addison-Wesley, Reading MA, 1994 (traduction française : A. Denise, Mathématiques concrètes, International Thomson Publishing, Paris, 1998).

[14] A. Granville & K. Soundararajan, An uncertainty principle for arithmetic sequences, *Ann. of Math.*, 39 p., à paraître. Prépublication disponible : `http://arxiv.org/abs/math.NT/0406018`

[15] D. Grieser, Counting complements in the partition lattice, and hypertrees, *J. Combin. Theory Ser. A* **57** (1991), no. 1, 144–150.

[16] H. Halberstam & H.-E. Richert, *Sieve Methods,* Academic Press, London, 1974.

[17] M. Hausman & H. N. Shapiro, On the mean square distribution of primitive roots of unity, *Comm. Pure Appl. Math.* **29** (1976), no. 3, 323–341.

[18] D. R. Heath-Brown, Zero-free regions for Dirichlet L-functions, and the least prime in an arithmetic progression, *Proc. London Math. Soc. (3)* **64** (1992), no. 2, 265–338.

[19] C. Hooley, On the difference of consecutive numbers prime to n, *Acta Arith.* **8** (1963), 343–347 ; II, *Publ. Math. Debrecen* **12** (1965), 39–49 ; III, *Math. Z.* **90** (1965), 355–364.

[20] C. Hooley, On the intervals between consecutive terms of sequences, *Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972),* 129–140, Amer. Math. Soc., Providence, RI, 1973.

[21] H. Iwaniec, On the error term in the linear sieve, *Acta Arith.* **19** (1971), 1–30.

[22] H. Iwaniec, On the problem of Jacobsthal, *Demostratio Math.* **11** (1978), 225–231.

[23] E. Jacobsthal, Über Sequenzen ganzer Zahlen, von denen keine zu n teilerfremd ist, I–III, *Norske Vid. Selsk. Forhdl. (Trondheim)* **33** (1961), 117–124, 125–131, 132–139 ; IV–V, *ibid.* **34** (1962), 1–7, 110–115.

[24] H.-J. Kanold, Über Primzahlen in arithmetischen Folgen, *Math. Ann.* **156** (1964), 393–395 ; II, *ibid.* **157** (1965), 358–362.

[25] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen,* 3eéd., Chelsea, New York (1974).

[26] U. V. Linnik, On the least prime in an arithmetic progression, I–II, *Rec. Math. [Mat. Sbornik] N.S.* **15(57)** (1944), 139–178, 347–368.

[27] P. Mazet, Recouvrements hamiltoniens de certains graphes. *European J. Combin.* **27** (2006), no. 5, 739–749.

[28] P. J. McCarthy, *Introduction to arithmetic functions,* Springer Verlag, New York, 1986.

[29] H. L. Montgomery & K. Soundararajan, Primes in short intervals, *Comm. Math. Phys.* **252** (2004), no. 1–3, 589–617.

[30] H. L. Montgomery & R. C. Vaughan, On the distribution of reduced residues, *Ann. Math.* **123** (1986), 311–333.

[31] H. L. Montgomery & R. C. Vaughan, A basic inequality, *Congress in Number Theory (Zarautz, 1984),* 163–175, Bilbao, Universidad del País Vasco, 1989.

[32] C. Pomerance, On the longest simple path in the divisor graph, *in* : Proceedings of the fourteenth Southeastern conference on combinatorics, graph theory and computing (Boca Raton, Fla., 1983), *Congr. Numer.* **40** (1983), 291–304.

[33] R. A. Rankin, The difference between consecutive prime numbers, *J. London Math. Soc.* **13** (1938), 242–247; V, *Proc. Edin. Math. Soc.* **13** (1962–63), 331–332.

[34] V. Romanovsky, *Note on the moments of a binomial* $(p+q)^n$ *about its mean,* Biometrika **15** (1923), 410.

[35] É. Saias, Applications des entiers à diviseurs denses, *Acta Arith.* **83** (1998), no. 3, 225–240.

[36] É. Saias, Étude du graphe divisoriel, III, *Rend. Circ. Mat. Palermo (2)* **52** (2003), no. 3, 481–488.

[37] G. Tenenbaum, Sur un problème de crible et ses applications, II, Corrigendum et étude du graphe divisoriel, *Ann. Sci. École Norm. Sup. (4)* **28** (1995), no. 2, 115–127.

[38] R. C. Vaughan, On the order of magnitude of Jacobsthal's function, *Proc. Edin. Math. Soc.* **20** (1976/77), 329–331.

Résumé

Suivant une idée de Erdős, nous montrons qu'une estimation uniforme en k, h et q de $M_k(h;q)$, moment centré d'ordre k de la répartition dans un intervalle glissant de longueur h des entiers premiers à l'entier q, permet de majorer la fonction de Jacobsthal, qui mesure l'écart maximal entre entiers consécutifs premiers à q. Montgomery et Vaughan ont étudié ce moment à k fixé : nous suivons leur argumentation. En établissant le comportement asymptotique du moment centré d'ordre k de la loi binomiale de paramètres (h, P) uniformément en ces trois variables, nous déduisons que l'estimation conjecturée est valable dès que les facteurs premiers de q sont supérieurs à h. Pour les cas restants (q sans grand facteur premier et k « petit »), nous analysons le lemme fondamental de Montgomery et Vaughan, et l'améliorons dans certains cas grâce à l'étude des sommes de Ramanujan.
Indépendamment, d'autres problèmes sont traités, concernant notamment l'étude du graphe divisoriel et de ses recouvrements en chaînes disjointes.

Mots-clefs : Répartition de suite d'entiers, méthode indirecte, loi binomiale, sommes de Ramanujan ; graphe divisoriel.

Abstract

By following the lead of Erdős, we prove that an estimate uniform in h, k and q of $M_k(h;q)$, centered moment of order k of the distribution of integers coprime to q in a sliding interval of length h, leads to an upper bound of the Jacobsthal function, which is the maximal gap between integers coprime to q. Montgomery and Vaughan studied this moment with k fixed: we follow their arguments. By stating the asymptotic behaviour of the k^{th} centered moment of a binomial distribution with h trials and success probability P uniformly in these three parameters, we deduce that the conjectured bound for $M_k(h;q)$ is valid unless q has a prime factor less than h. In the other cases (if q is free of large prime factor and k is small enough), we examine the fundamental lemma due to Montgomery and Vaughan and we improve it in some cases by studying Ramanujan sums.
Separately, we investigate some other problems, *e.g.* about the divisorial graph and its coverings with disconnected chains.

Keywords: Distribution of integer sequences, indirect method, binomial distribution, Ramanujan sums ; divisorial graph.