

Sur l'inégalité de Turán–Kubilius friable

R. de la Bretèche & G. Tenenbaum

Abstract. We obtain a new form, uniform with respect to all parameters, of the friable (i.e. relevant to integers free of large prime factors) Turán–Kubilius inequality, comparing the empirical variance of an additive arithmetical function with friable support to that of its probabilistic model. Several applications are developed, significantly improving on previously known results for small values of the friability parameter.

Keywords: Friable integers, additive functions, Kubilius model, Turán–Kubilius inequality.

Résumé. Nous établissons une version uniforme, relativement à tous les paramètres, de l'inégalité de Turán–Kubilius friable (i.e. relative aux entiers sans grand facteur premier), comparant la variance empirique d'une fonction arithmétique additive à support friable à celle de son modèle probabiliste. Nous développons plusieurs applications améliorant significativement des résultats antérieurement connus pour les petites valeurs du paramètre de friabilité.

Mots clefs : Entiers friables, fonctions additives, modèle de Kubilius, inégalité de Turán–Kubilius.

1. Introduction et énoncé

La théorie des entiers friables, ou sans grand facteur premier, occupe une place grandissante au sein de théorie analytique et probabiliste des nombres. L'un des outils les plus productifs en termes d'applications est l'extension à ce cadre de l'inégalité de Turán–Kubilius. Abordée dans [1], puis [10], [11], cette étude a été poursuivie et uniformisée dans [3], [8], [6], [4]. Nous nous proposons ici d'apporter une réponse à l'une des questions fondamentales restant en suspens et de développer quelques conséquences.

Soit \mathbb{A} la classe des fonctions arithmétiques additives complexes. Pour $x \geq y \geq 2$, nous désignons par $S(x, y)$ l'ensemble des entiers y -friables n'excédant pas x ,⁽¹⁾ et par $\alpha = \alpha(x, y)$ le point-selle associé à l'intégrale de Perron pour $\Psi(x, y) := |S(x, y)|$, défini par l'équation transcendante

$$(1.1) \quad \sum_{p \leq y} \frac{\log p}{p^\alpha - 1} = \log x.$$

Dans tout ce travail, nous réservons la lettre p pour désigner un nombre premier.

Dans [3], nous avons proposé, pour chaque fonction f de \mathbb{A} , un modèle probabiliste $Z_f = Z_{f, x, y}$ de la restriction de f à $S(x, y)$, défini par la formule

$$Z_f := \sum_{p \leq y} \xi_p$$

où $\xi_p = \xi_p(f)$ est une variable aléatoire géométrique telle que

$$(1.2) \quad \mathbb{P}(\xi_p = f(p^\nu)) = g_p(\alpha) / p^{\nu\alpha} \quad (\nu \in \mathbb{N}),$$

1. Autrement dit l'ensemble de tous les entiers naturels $n \leq x$ dont le plus grand facteur premier est $\leq y$.

les ξ_p étant supposées indépendantes, avec la convention que $f(p^\nu) = 0$ si $p^\nu > x$. Ici et dans la suite, nous utilisons la notation

$$g_m(s) := \prod_{p|m} (1 - 1/p^s) \quad (m \in \mathbb{N}^*, s \in \mathbb{C}).$$

Par convention, si plusieurs valeurs (éventuellement en nombre infini) $f(p^\nu)$ sont égales, la probabilité correspondante figurant au membre de gauche de (1.2) est définie comme la somme des quantités apparaissant au membre de droite.

Comme dans [3], nous définissons la variance semi-empirique V_f de f sur $S(x, y)$ par la formule

$$(1.3) \quad V_f = V_f(x, y) := \frac{1}{\Psi(x, y)} \sum_{n \in S(x, y)} |f(n) - \mathbb{E}(Z_f)|^2,$$

avec donc

$$(1.4) \quad \mathbb{E}(Z_f) = \sum_{p^\nu \in S(x, y)} \frac{g_p(\alpha) f(p^\nu)}{p^{\nu\alpha}}.$$

Dans son acception la plus forte, l'inégalité de Turán–Kubilius friable consiste en une majoration de $V_f(x, y)$ par un multiple constant de $\mathbb{V}(Z_f)$. Nous avons

$$(1.5) \quad \mathbb{V}(Z_f) := \sum_{p^\nu \in S(x, y)} \frac{g_p(\alpha)}{p^{\nu\alpha}} |f(p^\nu)|^2 - \sum_{p \leq y} \left| \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha) f(p^\nu)}{p^{\nu\alpha}} \right|^2,$$

où nous avons posé $\nu_p = \nu_p(x) := \lfloor (\log x) / \log p \rfloor$ ($p \leq y$).

Les résultats de [3] et [4], fournissent la majoration souhaitée

$$(1.6) \quad V_f \ll \mathbb{V}(Z_f)$$

dans le domaine défini par la condition $\sqrt{\log x} \log_2 x \leq y \leq x$. De plus, le théorème 1.1 de [3] implique (1.6) lorsque les variables aléatoires ξ_p sont centrées, autrement dit lorsque

$$(1.7) \quad \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha) f(p^\nu)}{p^{\nu\alpha}} = 0 \quad (p \in \mathcal{P}).$$

Ici et dans la suite, \mathcal{P} désigne l'ensemble des nombres premiers.

Le présent travail est consacré à l'extension du domaine de validité de (1.6). Nous obtenons le résultat suivant.

Théorème 1.1. *La majoration (1.6) est valable uniformément pour $f \in \mathbb{A}$ et $x \geq y \geq 2$.*

Posons

$$(1.8) \quad \begin{aligned} B_f^2 &= B_f(x, y)^2 := \sum_{p^\nu \in S(x, y)} \frac{g_p(\alpha)}{p^{\nu\alpha}} |f(p^\nu)|^2, \\ (B_f^-)^2 &= B_f^-(x, y)^2 := \sum_{p^\nu \in S(x, y)} \frac{g_p(\alpha)^2}{p^{\nu\alpha}} |f(p^\nu)|^2. \end{aligned}$$

L'inégalité de Cauchy-Schwarz permet d'écrire

$$(1.9) \quad (1 - 2^{-\alpha}) B_f^2 \leq (B_f^-)^2 \leq \mathbb{V}(Z_f) \leq B_f^2.$$

Notre approche est fondée sur le fait que l'inégalité

$$(1.10) \quad V_f(x, y) \ll B_f(x, y)^2 \quad (f \in \mathbb{A}, x \geq y \geq 2),$$

établie dans [3] en toute généralité, permet de ramener la preuve de (1.6) au cas d'une fonction fortement additive. Sous l'hypothèse $y \ll \log x$, une estimation générale, prouvée dans [2], du comportement local de $\Psi(x, y)$ fournit alors rapidement l'estimation souhaitée.

Il est à noter que le modèle (1.2) comporte un biais systématique, dû au fait que les probabilités ainsi définies ne sont pas nulles lorsque $p^\nu > x$. Ce phénomène est particulièrement évident lorsque f est fortement additive : notant

$$(1.11) \quad w_p = w_p(x, y) := p^{-\alpha \nu_p},$$

nous avons alors

$$\mathbb{P}(\xi_p = f(p)) = \frac{1 - w_p}{p^\alpha}$$

alors que (voir en particulier [3] ou (3.4) *infra*) l'approximation au premier ordre de $\Psi(x/p, y)/\Psi(x, y)$ est $1/p^\alpha$. Nous avons — cf., par exemple, [2], formule (3.4) —

$$(1.12) \quad x^{-\alpha} \leq w_p \leq \min\{(y/x)^\alpha, x^{-\alpha/2}\}, \quad (xy)^{-\alpha} \leq w_p/p^\alpha \leq x^{-\alpha},$$

de sorte que le biais n'a qu'une influence marginale dès que $y \rightarrow \infty$. En effet, notant classiquement

$$\pi(y) := \sum_{p \leq y} 1, \quad \vartheta(y) := \sum_{p \leq y} \log p \asymp y \quad (y \geq 2),$$

et posant traditionnellement $u := (\log x)/\log y$, nous observons, à fins de référence ultérieure, que la définition (1.1) implique immédiatement

$$(1.13) \quad \left(1 + \frac{\vartheta(y)}{\log x}\right)^u \leq x^\alpha \leq e^{\pi(y)} \quad (x \geq y \geq 2).$$

Pour améliorer la précision des résultats, nous introduisons les variables aléatoires géométriques indépendantes non biaisées $\xi_p^* = \xi_p^*(f)$ vérifiant

$$(1.14) \quad \mathbb{P}(\xi_p^* = 0) = g_p(\alpha), \quad \mathbb{P}(\xi_p^* = f(p^\nu)) = \frac{g_p(\alpha)}{(1 - w_p)p^{\nu\alpha}} \quad (1 \leq \nu \leq \nu_p),$$

avec la même convention que précédemment pour les cas d'égalité des valeurs prises, et nous posons

$$(1.15) \quad Z_f^* := \sum_{p \leq y} \xi_p^*.$$

Notons que, lorsque f est fortement additive, les ξ_p^* sont des variables de Bernoulli définies par

$$(1.16) \quad \mathbb{P}(\xi_p^* = 0) = g_p(\alpha), \quad \mathbb{P}(\xi_p^* = f(p)) = \frac{1}{p^\alpha}.$$

Dans le cas général, nous avons

$$(1.17) \quad \begin{aligned} \mathbb{E}(Z_f^*) &= \sum_{p^\nu \in S(x, y)} \frac{g_p(\alpha) f(p^\nu)}{(1 - w_p) p^{\nu\alpha}}, \\ \mathbb{V}(Z_f^*) &:= \sum_{p^\nu \in S(x, y)} \frac{g_p(\alpha) |f(p^\nu)|^2}{(1 - w_p) p^{\nu\alpha}} - \sum_{p \leq y} \left| \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha) f(p^\nu)}{(1 - w_p) p^{\nu\alpha}} \right|^2, \end{aligned}$$

la variance semi-empirique associée à ce modèle étant définie par

$$(1.18) \quad V_f^* = V_f^*(x, y) := \frac{1}{\Psi(x, y)} \sum_{n \in S(x, y)} |f(n) - \mathbb{E}(Z_f^*)|^2.$$

Notre approche fonctionne sans changement dans ce cadre : parallèlement au Théorème 1.1, nous obtenons le résultat suivant.

Théorème 1.2. *Nous avons*

$$(1.19) \quad V_f^*(x, y) \ll \mathbb{V}(Z_f^*)$$

uniformément pour toute fonction additive complexe f et tous x, y tels que $x \geq y \geq 2$.

Il est à noter que, lorsque f est fortement additive, la variance de Z_f^* prend la forme particulièrement agréable

$$(1.20) \quad \mathbb{V}(Z_f^*) = \sum_{p \leq y} \frac{g_p(\alpha) |f(p)|^2}{p^\alpha}.$$

2. Applications

Nous présentons ici quelques exemples d'applications ouvertes par les Théorèmes 1.1 et 1.2.

On sait classiquement que la forme duale de l'inégalité de Turán–Kubilius s'apparente à une inégalité de grand crible. Nous avons énoncé un tel résultat dans [3], comme conséquence de la majoration universelle (1.10). Comme le principe de dualité nécessite, au moins dans l'approche standard, une majoration par une forme quadratique diagonale, le recours à (1.6) ou (1.19) au lieu de (1.10) n'apporte pas d'amélioration immédiate. Cependant, dans le cas des fonctions fortement additives, la variance du modèle se présente sous forme diagonale. Nous obtenons le résultat suivant.

Théorème 2.1. *Il existe une constante absolue $C > 0$ telle que, pour toute suite $\{a_n\}_{n=1}^\infty$ de nombres complexes, et uniformément pour $x \geq y \geq 2$, on ait*

$$(2.1) \quad \sum_{p \leq y} \frac{p^\alpha}{g_p(\alpha)} \left| \sum_{\substack{n \in S(x, y) \\ p|n}} a_n - \frac{1}{p^\alpha} \sum_{n \in S(x, y)} a_n \right|^2 \leq C \Psi(x, y) \sum_{n \in S(x, y)} |a_n|^2.$$

Cela représente une amélioration significative par rapport au résultat analogue issu de (1.10), dans lequel les dénominateurs $g_p(\alpha)$ sont absents : lorsque $y \leq \log x$, on a $g_p(\alpha) \asymp \pi(y)(\log p)/\log x \ll y/\log x$ et le gain est patent.

Il est également à noter que l'utilisation du modèle non biaisé Z_f^* apporte aussi une précision supplémentaire : avec le modèle biaisé Z_f , nous n'aurions obtenu qu'une inégalité de type (2.1) où les dénominateurs $g_p(\alpha)$ sont remplacés par $g_p(\alpha) + w_p$.

Notre seconde application concerne la structure multiplicative d'un entier friable. Désignons par $\{p_j(n)\}_{j=1}^{\omega(n)}$ la suite croissante des facteurs premiers distincts d'un entier générique n , qui coïncide donc avec l'ensemble des sauts de la fonction

$$t \mapsto \omega_t(n) := \sum_{\substack{p|n \\ p \leq t}} 1 \quad (2 \leq t \leq y).$$

Plus précisément, nous avons $\omega_t(n) = j$ si, et seulement si, $p_j(n) \leq t < p_{j+1}(n)$, avec la convention que $p_{j+1}(n) = \infty$ lorsque $j = \omega(n)$. Comme $\omega_t(n)$ dépend additivement de n pour chaque valeur de t fixée, il est naturel d'attendre qu'une majoration de variance fournisse, via un argument de type Bienaymé–Tchébychev, une évaluation de $p_j(n)$ pour une sous-suite dense de $S(x, y)$. Nous avons développé une telle application dans [3] comme conséquence de (1.10). L'inégalité (1.19) permet de préciser le théorème 1.6 de [3].

Posons, pour $\alpha = \alpha(x, y)$,

$$v_t := \frac{t^{1-\alpha} - 1}{(1-\alpha)\log t}, \quad M_t = M_t(x, y) := \sum_{p \leq t} \frac{1}{p^\alpha} \quad (2 \leq t \leq y \leq x).$$

Nous avons établi dans [3] que

$$(2.2) \quad M_t = \log_2 2t + v_t + O\left(\frac{v_t}{\log 2v_t}\right) \quad (2 \leq t \leq y \leq x).$$

Nous obtenons le résultat suivant.

Théorème 2.2. *Pour toute constante $b > 1$ et toute fonction positive ϑ_x tendant vers l'infini avec x , il existe une fonction ε_x telle que $\lim_{x \rightarrow \infty} \varepsilon_x = 0$ ayant la propriété suivante : pour tous les entiers n de $S(x, y)$ sauf au plus $\varepsilon_x \Psi(x, y)$ d'entre eux et uniformément pour $\vartheta_x \leq t \leq y \leq x$, nous avons*

$$(2.3) \quad |\omega_t(n) - M_t| \leq \begin{cases} \alpha^{1/3} M_t^{2/3} \{\log(3M_t/\alpha)\}^{2b/3} + 1 & (2 \leq t \leq z), \\ M_t^{2/3} (\log M_t)^{b/3} & (z < t \leq y), \end{cases}$$

où l'on a posé $z := \min\{y, \exp(1/\alpha)\}$.

Si, de plus, $\alpha = o(1)$, on peut choisir ε_x de sorte que $\omega_t(n) = \pi(t)$ dès que $t = o(1/\alpha)$.

Remarque. La formule (2.3) est essentiellement équivalente à

$$(2.4) \quad |\omega_t(n) - M_t| \leq \frac{2\alpha^{1/3} M_t^{2/3} \{\log(3M_t/\alpha)\}^{2b/3}}{1 + \{\alpha \log(3M_t/\alpha)\}^{1/3}} + 1 \quad (2 \leq t \leq y).$$

Le Théorème 2.2 permet une estimation de la répartition normale de la suite $\{p_j(n)\}_{j=1}^{\omega(n)}$. Qualitativement, lorsque, par exemple, $y \leq (\log x)^k$, où $k > 1$ est une constante arbitraire fixée, nous mettons en évidence un remarquable effet de tassement, au point que $p_j(n)$ devient graduellement comparable, lorsque y décroît, au j -ème nombre premier, noté p_j . Pour les plus grandes valeurs de y , nous avons $\alpha \gg 1$, et seule la seconde éventualité de (2.3) se présente. La conclusion du corollaire 1.7 de [3] fournit alors une description satisfaisante de la situation.

Nous posons $z_1 := \min(y, e^{1/2\alpha})$ et notons que $\alpha \leq 1 - 1/k + o(1)$ pour $y \leq (\log x)^k$.

Corollaire 2.3. *Soit $k > 1$. Pour toute constante $b > 1$ et toute fonction positive J_x tendant vers l'infini avec x , il existe une fonction ε_x telle que $\lim_{x \rightarrow \infty} \varepsilon_x = 0$ et vérifiant la propriété suivante : pour tous les entiers n de $S(x, y)$ sauf au plus $\varepsilon_x \Psi(x, y)$ d'entre eux et pour $J_x \leq j \leq \omega(n)$, $2 \leq y \leq (\log x)^k$, nous avons*

$$(2.5) \quad \begin{aligned} 0 \leq p_j(n) - p_j &\ll \alpha^{1/3} p_j^{2/3} \{\log(3p_j/\alpha)\}^{2b/3} + \alpha p_j + 1 \quad (p_j \leq z_1) \\ p_j(n) &= p_j^{1/(1-\alpha)} e^{O(\alpha)} \quad (z_1 < p_j \leq y). \end{aligned}$$

De plus, lorsque $\alpha = o(1)$, on a $p_j(n) = p_j$ dès que $p_j = o(1/\alpha)$.

Il est à noter que l'énoncé précédent implique en particulier que presque tous les entiers de $S(x, y)$ vérifient

$$(2.6) \quad p_j(n) \sim p_j \quad (1 \leq j \leq \omega(n))$$

dès que $y = o(\log x)$.

3. Démonstration du Théorème 1.1

Soit $f \in \mathbb{A}$. Étant donné x, y tels que $x \geq y \geq 2$, nous définissons une fonction fortement additive $h = h_{x,y}$ par

$$(3.1) \quad \frac{h(p)}{p^\alpha} (1 - w_p) = \sum_{1 \leq \nu \leq \nu_p} g_p(\alpha) \frac{f(p^\nu)}{p^{\nu\alpha}} \quad (p \in \mathcal{P}),$$

où nous avons utilisé les notations introduites en (1.11).

Nous avons donc $f = g + h$, où g vérifie (1.7). Pour un couplage convenable, nous avons également $Z_f = Z_g + Z_h$: il suffit en effet de choisir, avec la convention indiquée dans l'introduction,

$$\left\{ \begin{array}{l} \mathbb{P}(\xi_p(h) = 0, \xi_p(g) = 0) := g_p(\alpha) + \frac{w_p}{p^\alpha}, \\ \mathbb{P}(\xi_p(h) = h(p), \xi_p(g) = g(p^\nu)) := \frac{g_p(\alpha)}{p^{\nu\alpha}} \quad (1 \leq \nu \leq \nu_p). \end{array} \right.$$

Ainsi qu'il a été rappelé plus haut, il résulte du théorème 1.1 de [3] que

$$V_g = V_g(x, y) \ll \mathbb{V}(Z_g).$$

De plus,

$$V_f \leq 2V_g + 2V_h, \\ \mathbb{V}(Z_h) = \sum_{p \leq y} (1 - w_p) \frac{|h(p)|^2}{p^\alpha} \left(1 - \frac{1 - w_p}{p^\alpha}\right) = \sum_{p \leq y} (1 - w_p) \frac{|h(p)|^2}{p^\alpha} \left(g_p(\alpha) + \frac{w_p}{p^\alpha}\right).$$

Observons alors, d'une part, que, pour tout nombre premier $p \leq y$,

$$\begin{aligned} (1 - w_p) \frac{|h(p)|^2}{p^\alpha} &\leq \frac{p^\alpha}{1 - w_p} \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha)}{p^{\nu\alpha}} \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha)}{p^{\nu\alpha}} |f(p^\nu)|^2 \\ &= \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha)}{p^{\nu\alpha}} |f(p^\nu)|^2. \end{aligned}$$

alors que, d'autre part, par un emploi standard de l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} \mathbb{V}(Z_f) &= \sum_{p^\nu \in S(x,y)} \frac{g_p(\alpha)}{p^{\nu\alpha}} |f(p^\nu)|^2 - \sum_{p \leq y} \left| \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha)}{p^{\nu\alpha}} f(p^\nu) \right|^2 \\ &\geq \sum_{p^\nu \in S(x,y)} \frac{g_p(\alpha)}{p^{\nu\alpha}} |f(p^\nu)|^2 \left(g_p(\alpha) + \frac{w_p}{p^\alpha}\right). \end{aligned}$$

Il s'ensuit que donc

$$\mathbb{V}(Z_h) \leq \mathbb{V}(Z_f).$$

Sous l'hypothèse $V_h \ll \mathbb{V}(Z_h)$, nous déduisons donc de ce qui précède que nous avons

$$V_f \leq 2V_g + 2V_h \ll \mathbb{V}(Z_f - Z_h) + \mathbb{V}(Z_h) \ll \mathbb{V}(Z_f) + \mathbb{V}(Z_h) \ll \mathbb{V}(Z_f),$$

c'est-à-dire (1.6).

Nous pouvons ainsi, sans restreindre la généralité de l'argument, supposer dans la suite que f est une fonction fortement additive.

Nous pouvons également supposer f réelle puisque

$$V_f = V_{\Re f} + V_{\Im f}, \quad \mathbb{V}(Z_f) = \mathbb{V}(Z_{\Re f}) + \mathbb{V}(Z_{\Im f})$$

dans le cas général.

À fins de référence ultérieure, nous notons immédiatement que

$$(3.2) \quad \mathbb{V}(Z_f) = \sum_{p \leq y} (1 - w_p) \frac{f(p)^2}{p^\alpha} \left(g_p(\alpha) + \frac{w_p}{p^\alpha} \right),$$

et rappelons la formule (1.20) pour $\mathbb{V}(Z_f^*)$.

Sous les hypothèses supplémentaires mentionnées plus haut, nous pouvons écrire

$$(3.3) \quad V_f(x, y) = \frac{1}{\Psi(x, y)} \sum_{n \in S(x, y)} \left(\sum_{p|n} f(p) - \mathbb{E}(Z_f) \right)^2 = M_2 - 2M_1 \mathbb{E}(Z_f) + \mathbb{E}(Z_f)^2,$$

avec

$$\mathbb{E}(Z_f) = \sum_{p \leq y} \frac{f(p)(1 - w_p)}{p^\alpha}, \quad M_j := \frac{1}{\Psi(x, y)} \sum_{n \in S(x, y)} f(n)^j \quad (j = 1, 2),$$

de sorte que

$$M_1 = \sum_{p \leq y} f(p) \frac{\Psi(x/p, y)}{\Psi(x, y)}, \quad M_2 = \sum_{p \leq y} f(p)^2 \frac{\Psi(x/p, y)}{\Psi(x, y)} + \sum_{\substack{p, q \leq y \\ p \neq q}} f(p)f(q) \frac{\Psi(x/pq, y)}{\Psi(x, y)}.$$

Il est manifeste que des estimations relatives au comportement local de $\Psi(x, y)$ sont nécessaires pour évaluer M_1 et M_2 . Nous ferons appel à la formule générale suivante, établie au théorème 2.4 de [2] et valable uniformément sous les conditions $x \geq y \geq 2$, $1 \leq d \leq x$, $\log d \ll \log y$, $\omega(m) \ll 1$:

$$(3.4) \quad \Psi_m\left(\frac{x}{d}, y\right) := \sum_{\substack{n \in S(x/d, y) \\ (n, m) = 1}} 1 = \frac{g_m(\alpha)}{d^\alpha} \Psi(x, y) \left\{ 1 + O\left(\frac{1}{\bar{u}}\right) \right\},$$

où l'on a posé

$$(3.5) \quad \bar{u} := \min(u, y/\log y).$$

Introduisant

$$r_d = r_d(x, y) := \frac{\Psi(x/d, y)}{\Psi(x, y)} - \frac{1}{d^\alpha} \quad (x \geq y \geq 2, d \geq 1),$$

nous obtenons en appliquant (3.4) avec $m = 1$, $d = p$, puis $m = p$, $d = 1$,

$$(3.6) \quad r_p = \frac{\Psi(x/p, y)}{\Psi(x, y)} - \frac{1}{p^\alpha} = g_p(\alpha) - \frac{\Psi_p(x, y)}{\Psi(x, y)} \ll \frac{g_p(\alpha)}{p^\alpha \bar{u}},$$

uniformément pour $x \geq y \geq 2$, $p \leq y$.

Compte tenu des résultats rappelés dans l'introduction, nous pouvons supposer

$$(3.7) \quad 2 \leq y \leq \log x,$$

de sorte que $y^\alpha \asymp 1$ — cf., par exemple, [7] ou (5.15) *infra*. Observons d'emblée qu'à toute fin utile nous pouvons remplacer, dans le membre de droite de (3.3), l'espérance du modèle Z_f par celle du modèle non biaisé Z_f^* , défini en (1.15). En effet, puisque

$$(3.8) \quad \mathbb{V}(Z_f) = \sum_{p \leq y} (1 - w_p) \frac{f(p)^2}{p^\alpha} \left(g_p(\alpha) + \frac{w_p}{p^\alpha} \right) \geq \{1 - (y/x)^\alpha\} \sum_{p \leq y} \frac{f(p)^2 w_p}{p^\alpha},$$

nous avons, d'une part,

$$(3.9) \quad \{\mathbb{E}(Z_f) - \mathbb{E}(Z_f^*)\}^2 \leq \sum_{p \leq y} \frac{f(p)^2 w_p}{p^\alpha} \sum_{p \leq y} \frac{w_p}{p^\alpha} \ll \frac{\pi(y) x^{-\alpha}}{1 - (y/x)^\alpha} \mathbb{V}(Z_f) \ll \frac{y \mathbb{V}(Z_f)}{x^\alpha \log y}$$

et, d'autre part, au vu de (3.6), puisque $\alpha \leq \pi(y)/\log x$ d'après (1.13) et donc $p^\alpha \ll 1$ pour $p \leq y$ dans le domaine (3.7),

$$(3.10) \quad M_1 - \mathbb{E}(Z_f^*) = \sum_{p \leq y} f(p) r_p \ll \frac{1}{\pi(y)} \sqrt{\sum_{p \leq y} \frac{f(p)^2 g_p(\alpha)}{p^\alpha} \sum_{p \leq y} \frac{g_p(\alpha)}{p^\alpha}} \ll \sqrt{\frac{\mathbb{V}(Z_f^*)}{u}},$$

$$(3.11) \quad M_1 - \mathbb{E}(Z_f) = \sum_{p \leq y} f(p) \left\{ r_p + \frac{w_p}{p^\alpha} \right\} \ll \sqrt{\frac{\mathbb{V}(Z_f^*)}{u}} + \sqrt{\frac{\pi(y) \mathbb{V}(Z_f)}{x^\alpha}}.$$

Puisque $V_f^* = V_f^*(x, y) = M_2 - 2M_1 \mathbb{E}(Z_f^*) + \mathbb{E}(Z_f^*)^2$, il s'ensuit que

$$(3.12) \quad \begin{aligned} V_f - V_f^* &= \{\mathbb{E}(Z_f^*) - \mathbb{E}(Z_f)\} \{2M_1 - \mathbb{E}(Z_f) - \mathbb{E}(Z_f^*)\} \\ &\ll \left\{ \sqrt{\frac{y}{x^\alpha \log x}} + \frac{y}{x^\alpha \log y} \right\} \mathbb{V}(Z_f) \ll \mathbb{V}(Z_f). \end{aligned}$$

Estimons donc V_f^* . Nous avons

$$\begin{aligned} M_2 &= \sum_{p \leq y} \frac{f(p)^2}{p^\alpha} + \sum_{p \leq y} f(p)^2 r_p + \sum_{\substack{p, q \leq y \\ p \neq q}} \frac{f(p) f(q)}{p^\alpha q^\alpha} + \sum_{\substack{p, q \leq y \\ p \neq q}} f(p) f(q) r_{pq} \\ &= \mathbb{V}(Z_f^*) + \mathbb{E}(Z_f^*)^2 + \sum_{p \leq y} f(p)^2 r_p + \sum_{\substack{p, q \leq y \\ p \neq q}} f(p) f(q) r_{pq}, \end{aligned}$$

d'où, compte tenu de l'égalité de (3.10),

$$(3.13) \quad V_f^* = \mathbb{V}(Z_f^*) - G_f + D_f$$

avec

$$(3.14) \quad \mathbb{V}(Z_f^*) = \sum_{p \leq y} \frac{g_p(\alpha) f(p)^2}{p^\alpha} \leq \frac{\mathbb{V}(Z_f)}{1 - (y/x)^\alpha} = \left\{ 1 + O\left(\frac{1}{x^\alpha}\right) \right\} \mathbb{V}(Z_f),$$

$$(3.15) \quad G_f := \sum_{p \leq y} f(p)^2 r_p \left(\frac{2}{p^\alpha} - 1 \right), \quad D_f := \sum_{\substack{p, q \leq y \\ p \neq q}} f(p) f(q) \left\{ r_{pq} - \frac{r_p}{q^\alpha} - \frac{r_q}{p^\alpha} \right\},$$

où (3.14) découle de (3.8).

Il résulte immédiatement de (3.6) que

$$(3.16) \quad G_f \ll \frac{\mathbb{V}(Z_f^*)}{\pi(y)}.$$

Nous avons également

$$(3.17) \quad r_{pq} - \frac{r_p}{q^\alpha} - \frac{r_q}{p^\alpha} = \frac{1}{\Psi(x, y)} \left\{ \Psi\left(\frac{x}{pq}, y\right) - \frac{\Psi(x/p, y)}{q^\alpha} - \frac{\Psi(x/q, y)}{p^\alpha} + \frac{\Psi(x, y)}{p^\alpha q^\alpha} \right\}.$$

La quantité entre accolades vaut

$$(3.18) \quad \begin{aligned} \sum_{t|pq} \mu(t) \left(\frac{t}{pq}\right)^\alpha \Psi\left(\frac{x}{t}, y\right) &= \sum_{t|pq} \mu(t) \sum_{d|pq/t} g_d(\alpha) \mu(d) \sum_{\substack{n \in S(x, y) \\ t|n}} 1 \\ &= \sum_{d|pq} \mu(d) g_d(\alpha) \sum_{n \in S(x, y)} \sum_{t|(n, pq/d)} \mu(t) \\ &= \sum_{d|pq} \mu(d) g_d(\alpha) \Psi_{pq/d}(x, y). \end{aligned}$$

Compte tenu de (3.4), le membre de droite de (3.18) peut être évalué par

$$\sum_{d|pq} \frac{\mu(d) d^\alpha g_d(\alpha) g_{pq/d}(\alpha)}{d^\alpha p^\alpha q^\alpha} \Psi(x, y) + O\left(\frac{g_{pq}(\alpha)}{\pi(y) p^\alpha q^\alpha} \Psi(x, y)\right) \ll \frac{g_{pq}(\alpha)}{\pi(y) p^\alpha q^\alpha} \Psi(x, y).$$

Nous obtenons donc

$$D_f \ll \frac{1}{\pi(y)} \left(\sum_{p \leq y} g_p(\alpha) \frac{|f(p)|}{p^\alpha} \right)^2 \leq \frac{\mathbb{V}(Z_f^*)}{\pi(y)} \sum_{p \leq y} \frac{g_p(\alpha)}{p^\alpha}.$$

La dernière somme est $\ll \alpha y \ll y\pi(y)/\log x$. Il suit

$$(3.19) \quad D_f \ll \frac{y}{\log x} \mathbb{V}(Z_f^*).$$

La relation (3.13) et les estimations (3.12), (3.14), (3.16) et (3.19) fournissent alors

$$(3.20) \quad V_f(x, y) = \mathbb{V}(Z_f) \left\{ 1 + O\left(\frac{\log y}{y} + \frac{y}{\log x}\right) \right\} \quad (2 \leq y \leq \log x),$$

ce qui complète la preuve du Théorème 1.1.

À fins de référence ultérieure, nous observons que la démonstration précédente fournit également

$$(3.21) \quad V_f^*(x, y) = \mathbb{V}(Z_f^*) \left\{ 1 + O\left(\frac{\log y}{y} + \frac{y}{\log x}\right) \right\} \quad (2 \leq y \leq \log x).$$

4. Démonstration du Théorème 1.2

Comme précédemment, il suffit de considérer le cas des fonctions fortement additives. En effet, étant donnés une fonction réelle $f \in \mathbb{A}$ et x, y tels que $x \geq y \geq 2$, la fonction $g := f - h$, où $h = h_{x,y}$ est définie par (3.1), vérifie $\mathbb{E}(\xi_p(g)) = \mathbb{E}(\xi_p^*(g)) = 0$ pour tout $p \leq y$. Nous pouvons donc appliquer le théorème 1.1 de [3], qui fournit

$$V_g(x, y) \ll \mathbb{V}(Z_g) \asymp \mathbb{V}(Z_g^*)$$

puisque $1 - w_p \asymp 1$ pour $p \leq y$. De plus, pour un couplage convenable, nous avons la relation entre variables aléatoires $Z_f^* = Z_g^* + Z_h^*$, de sorte que $\mathbb{V}(Z_g^*) \leq 2\mathbb{V}(Z_h^*) + 2\mathbb{V}(Z_f^*)$.

Maintenant

$$\begin{aligned} V_g^* - 2V_g &\leq 2 |\mathbb{E}(Z_g) - \mathbb{E}(Z_g^*)|^2 \ll \sum_{p \leq y} \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha) |g(p^\nu)|^2}{p^{\nu\alpha}} \sum_{p \leq y} w_p^2 \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha)}{p^{\nu\alpha}} \\ &\ll \mathbb{V}(Z_g^*) \frac{\pi(y)}{x^{2\alpha}} \ll \mathbb{V}(Z_g^*), \end{aligned}$$

d'où

$$V_g^* \ll \mathbb{V}(Z_g^*) \ll \mathbb{V}(Z_h^*) + \mathbb{V}(Z_f^*).$$

De plus,

$$\begin{aligned} \mathbb{V}(Z_h^*) &= \sum_{p \leq y} \frac{g_p(\alpha) |h(p)|^2}{p^\alpha} \leq \sum_{p \leq y} \frac{g_p(\alpha)}{(1 - w_p)^2} \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha) |f(p^\nu)|^2}{p^{\nu\alpha}} \sum_{1 \leq \nu \leq \nu_p} \frac{g_p(\alpha)}{p^{(\nu-1)\alpha}} \\ &\leq \sum_{p^\nu \in S(x, y)} \frac{g_p(\alpha)^2 |f(p^\nu)|^2}{(1 - w_p) p^\alpha} \ll \mathbb{V}(Z_f^*). \end{aligned}$$

Sous l'hypothèse $V_h^* \ll V(Z_h^*)$, il suit donc

$$V_f^* \leq 2V_g^* + 2V_h^* \ll \mathbb{V}(Z_g^*) + \mathbb{V}(Z_h^*) \ll \mathbb{V}(Z_f^*) + \mathbb{V}(Z_h^*) \ll \mathbb{V}(Z_f^*).$$

Il reste à établir que $V_h^* \ll V(Z_h^*)$. Cela résulte de (3.21) lorsque $2 \leq y \leq \log x$. Lorsque $\log x < y \leq x$, nous avons $w_p/p^\alpha \ll g_p(\alpha)$ pour tout $p \leq y$, donc $\mathbb{V}(Z_h) \asymp \mathbb{V}(Z_h^*)$, et, grâce au théorème 1.1 de [3]

$$\begin{aligned} V_h^* &\leq 2V_h + 2\{\mathbb{E}(Z_h) - \mathbb{E}(Z_h^*)\}^2 \ll \mathbb{V}(Z_h) + \sum_{p \leq y} \frac{w_p^2}{p^\alpha} \sum_{p \leq y} \frac{|f(p)|^2}{p^\alpha} \\ &\ll \mathbb{V}(Z_h) + \frac{\pi(y)}{x^{3\alpha/2}} B_h(x, y)^2 \ll \mathbb{V}(Z_h) \ll \mathbb{V}(Z_h^*), \end{aligned}$$

où l'avant-dernière estimation résulte de (1.9) et (1.13).

5. Le cas des fonctions fortement additives

5.1. Énoncé du résultat principal et remarques

La démonstration du Théorème 1.1 fait apparaître le rôle essentiel joué par la classe des fonctions fortement additives dans l'étude du cas général. Nous rassemblons nos résultats relatifs à cette hypothèse dans l'énoncé suivant.

Nous posons

$$(5.1) \quad \begin{aligned} \eta &= \eta(x, y) := \frac{1}{u} + \frac{y^2 \pi(y)}{(\log x)^2} + \frac{y^3 \pi(y)^2}{(\log x)^3}, \\ \mathcal{R} &= \mathcal{R}(x, y) := \min \left\{ \frac{\log y}{y} + \frac{y}{\log x}, \eta(x, y) \right\}, \end{aligned}$$

et notons que $\eta = o(1)$ si, et seulement si, $y = o((\log x)^{3/5}(\log_2 x)^{2/5})$. Nous rappelons également la notation de Tchébychev

$$\vartheta(y) := \sum_{p \leq y} \log p.$$

Par ailleurs, nous désignons par \mathbb{A}_0 la classe des fonctions arithmétiques réelles fortement additives.

Théorème 5.1. *Soit $f \in \mathbb{A}_0$.*

(i) *Nous avons*

$$(5.2) \quad V_f^*(x, y) \leq \{1 + o(1)\} \mathbb{V}(Z_f^*) \quad (u \rightarrow \infty).$$

Si, en outre, $y = o(\log x)$, alors

$$(5.3) \quad V_f^*(x, y) = \{1 + o(1)\} \mathbb{V}(Z_f^*).$$

(ii) *Lorsque y et u tendent vers l'infini, nous avons également*

$$(5.4) \quad V_f(x, y) \leq \{1 + o(1)\} \mathbb{V}(Z_f).$$

(iii) *Si y et $(\log x)/y$ tendent vers l'infini, alors*

$$(5.5) \quad \frac{1}{1 + (\log x)/e^{\pi(y)}\pi(y) \log 2} + o(1) \leq \frac{V_f(x, y)}{\mathbb{V}(Z_f)} \leq \frac{1}{1 + (\log x)/ye^{\pi(y)}} + o(1).$$

(iv) *Sous l'hypothèse $2 \leq y \leq \log x$, nous avons plus précisément*

$$(5.6) \quad 1 - \frac{\vartheta(y)}{\log x} + O(\mathcal{R}) \leq \frac{V_f^*(x, y)}{\mathbb{V}(Z_f^*)} \leq 1 + O(\mathcal{R}).$$

Remarques. (i) La formule (5.3) met en évidence l'indépendance asymptotique, dans le domaine $y = o(\log x)$, des antécédents arithmétiques des variables ξ_p^* associées aux fonctions fortement additives selon le modèle non biaisé.

(ii) Martin et Tenenbaum ont montré — [8], formule (13.6) — que, pour tout $u \geq 1$ fixé, nous avons

$$(5.7) \quad \limsup_{x \rightarrow \infty} \sup_{f \in \mathbb{A}_0} \frac{V_f(x, x^{1/u})}{\mathbb{V}(Z_f)} = \lambda(u)$$

pour une application convenable $\lambda : [1, \infty[\rightarrow \mathbb{R}^+$, précisément définie dans le même travail et vérifiant $\lambda(u) - 1 \sim 1/8u$ ($u \rightarrow \infty$). Les valeurs de $\lambda(u)$ ont été tabulées dans [6], et vérifient $\lambda(u) > 1$ pour tout $u > 1$. Cela implique en particulier l'optimalité du domaine de validité de (5.2).

(iii) La formule (5.4) est une conséquence immédiate de (5.2), (3.12) et (3.14). Dans le domaine $y \leq \log x$, on obtient en fait une estimation effective en ajoutant $\pi(y)e^{-\pi(y)}$ aux termes d'erreur de (5.6).

(iv) Lorsque $y = 2$, il est aisé d'estimer $V_f(x, y)/\mathbb{V}(Z_f)$. Pour des raisons d'homogénéité, nous pouvons supposer $f(2) = 1$. Posant $N := \Psi(x, 2) = 1 + \nu_2(x)$, nous avons $\alpha = (1 + O(1/N))/N \log 2$ et

$$\begin{aligned}\mathbb{E}(Z_f) &= 2^{-\alpha}(1 - 2^{-N\alpha}) = 1 - e^{-1} + O(1/N), \\ V_f &= e^{-2} + O(1/N), \\ \mathbb{V}(Z_f) &= e^{-1} - e^{-2} + O(1/N),\end{aligned}$$

de sorte que

$$\frac{V_f(x, y)}{\mathbb{V}(Z_f)} = \frac{1}{e-1} + O\left(\frac{1}{N}\right) < 1.$$

(v) Désignons, conformément à l'usage, par $\omega(n)$ le nombre des facteurs premiers distincts d'un entier naturel n . Comme nous l'établirons au paragraphe 5.4, nous avons lorsque $(\log x)/y$ et y tendent vers l'infini,

$$(5.8) \quad V_\omega = \left(\frac{1}{1 + (\log x)/ye^{\pi(y)}} + o(1) \right) \mathbb{V}(Z_\omega).$$

Cela implique l'optimalité de la majoration de (5.5) dans la généralité énoncée.

(vi) Le choix de la fonction indicatrice des nombres pairs permet de montrer que la minoration de (5.5) est également optimale dans la généralité énoncée : il suffit de faire appel à (3.2), (1.20) et (5.15), (5.27) *infra*. Nous omettons les détails de la vérification.

5.2. Lemmes

Commençons par affiner les estimations (3.16) et (3.19) dans le domaine (3.7). Rappelons la notation (5.1).

Lemme 5.2. *Dans le domaine (3.7), uniformément pour $f \in \mathbb{A}_0$, nous avons*

$$(5.9) \quad G_f \ll \eta \mathbb{V}(Z_f^*)$$

et, pour une constante absolue convenable K ,

$$(5.10) \quad -\mathbb{V}(Z_f^*) \left\{ \frac{\vartheta(y)}{\log x} + K\eta \right\} \leq D_f \leq K\eta \mathbb{V}(Z_f^*).$$

Remarque. L'estimation (5.9) fournit une amélioration de (3.16) dès que

$$y \ll \sqrt{\log x \log_2 x}.$$

Démonstration. Considérons d'abord (5.9). Posons

$$\delta = \delta(x, y) := \frac{1}{u} + \frac{y^2 \pi(y)}{(\log x)^2}.$$

Dans le domaine (3.7) et pour $2 \leq p \leq y$, le corollaire 1.8 de [5] implique

$$(5.11) \quad \Psi_p(x, y) = g_p(\alpha) \Psi(x, y) \{1 + O(\delta)\},$$

d'où

$$(5.12) \quad r_p \ll g_p(\alpha) \delta,$$

et finalement (5.9).

Nous procédons similairement pour établir la majoration de (5.10). Une forme affaiblie du corollaire 1.8 de [5] permet d'écrire

$$(5.13) \quad \Psi_{pq}(x, y) = g_{pq}(\alpha) \Psi(x, y) \left\{ 1 - \frac{1}{\pi(y)} + O(\delta) \right\},$$

d'où, compte tenu de (3.17) et (3.18),

$$\begin{aligned} r_{pq} - \frac{r_p}{q^\alpha} - \frac{r_q}{p^\alpha} &= \sum_{d|pq} \mu(d) g_d(\alpha) \frac{\Psi_{pq/d}(x, y)}{\Psi(x, y)} \\ &= g_{pq}(\alpha) \left\{ -\frac{1}{\pi(y)} + O(\delta) \right\} = \frac{-g_{pq}(\alpha)}{p^\alpha q^\alpha \pi(y)} + O\left(\frac{g_{pq}(\alpha)\delta}{p^\alpha q^\alpha}\right), \end{aligned}$$

où nous avons utilisé l'estimation $p^\alpha q^\alpha = 1 + O(y/\log x)$, qui résulte de (1.13).

Reportons cette estimation dans l'expression de D_f donnée en (3.15). Le terme d'erreur fournit une contribution $\ll \eta \mathbb{V}(Z_f^*)$. Il s'ensuit que, dans le domaine (3.7), nous avons

$$(5.14) \quad \begin{aligned} D_f &= -\frac{1}{\pi(y)} \left(\sum_{p \leq y} \frac{f(p) g_p(\alpha)}{p^\alpha} \right)^2 + \frac{1}{\pi(y)} \sum_{p \leq y} \frac{f(p)^2 g_p(\alpha)^2}{p^{2\alpha}} + O\left(\alpha y \delta \mathbb{V}(Z_f^*)\right) \\ &\leq K \eta \mathbb{V}(Z_f^*), \end{aligned}$$

pour une constante absolue convenable K , puisque $g_p(\alpha)/\pi(y) \ll 1/u \ll \eta$ dans le domaine considéré.

Pour établir la minoration de (5.10), nous appliquons l'inégalité de Cauchy-Schwarz sous la forme

$$\begin{aligned} \left(\sum_{p \leq y} \frac{f(p) g_p(\alpha)}{p^\alpha} \right)^2 &\leq \mathbb{V}(Z_f^*) \sum_{p \leq y} \frac{g_p(\alpha)}{p^\alpha} \leq \mathbb{V}(Z_f^*) \alpha \sum_{p \leq y} \log p \\ &\leq \mathbb{V}(Z_f^*) \frac{\pi(y) \vartheta(y)}{\log x} \left\{ 1 + O\left(\frac{y}{\log x}\right) \right\}, \end{aligned}$$

où nous avons utilisé l'estimation (voir la formule (2.14) de [5])

$$(5.15) \quad \alpha = \frac{\pi(y)}{\log x} \left\{ 1 + O\left(\frac{y}{\log x}\right) \right\} \quad (2 \leq y \leq \log x).$$

Nous en déduisons l'inégalité souhaitée en reportant dans l'égalité de (5.14). \square

Lemme 5.3. Soit $h : [1, \infty[\rightarrow \mathbb{R}^+$ une fonction telle que $\lim_{x \rightarrow \infty} h(x) = \infty$. Sous la condition

$$(5.16) \quad h(x) \sqrt{\log x} \log_2 x \leq y \leq x^{1/h(x)},$$

et uniformément pour $f \in \mathbb{A}_0$, nous avons

$$(5.17) \quad V_f(x, y) \leq \{1 + o(1)\} \mathbb{V}(Z_f) \quad (x \rightarrow \infty).$$

Démonstration. Lorsque (5.16) est réalisée et, sous la condition supplémentaire $u > \sqrt{\log y}$, la preuve du théorème principal de [4] fournit en fait l'existence d'une constante $K > 0$

telle que pour toute fonction f additive réelle

$$V_f(x, y) - \mathbb{V}(Z_f) \leq K \frac{B^2(x, y)}{\bar{u}}. \quad (2)$$

Lorsque f est fortement additive, l'inégalité $\frac{1}{2}\alpha \log 2 \leq g_p(\alpha)$ permet d'en déduire que

$$V_f(x, y) - \mathbb{V}(Z_f) \leq \frac{K}{\bar{u}} \sum_{p \leq y} \frac{f(p)^2}{p^\alpha} \leq \frac{2K}{\bar{u}\alpha \log 2} \mathbb{V}(Z_f^*).$$

La minoration classique $\alpha \log y \gg \log(1 + y/\log x)$ implique alors, dans les hypothèses effectuées,

$$\alpha \bar{u} \gg \log \left(1 + \frac{y}{\log x} \right) \frac{\min(y, \log x)}{(\log y)^2} \gg \min \left(h(x), \frac{\log x}{(\log_2 x)^2} \right),$$

d'où (5.17).

Lorsque $u \leq \sqrt{\log y}$, et donc $e^{(\log x)^{2/3}} \leq y \leq x^{1/h(x)}$, nous pouvons faire appel au corollaire 5.2 de [3] qui fournit directement (5.17). \square

Dans l'énoncé suivant, nous posons

$$E := \frac{(y + \log x) \log y}{x^\alpha y} \ll \frac{\log x}{\pi(y)} \left(1 + \frac{\vartheta(y)}{\log x} \right)^{1-u},$$

$$F := \frac{(y + \log x) \sqrt{\bar{u} + \log_2 y}}{x^\alpha \bar{u} \sqrt{\log x}} \ll \frac{\sqrt{\log x} \sqrt{\bar{u} + \log_2 y}}{\bar{u} y} \left(1 + \frac{\vartheta(y)}{\log x} \right)^{1-u}.$$

où nous avons fait appel à (1.13).

Lemme 5.4. Uniformément pour $x \geq y \geq 2$ et $f \in \mathbb{A}_0$, nous avons

$$(5.18) \quad \frac{V_f(x, y)}{\mathbb{V}(Z_f)} = \frac{V_f^*(x, y)}{\mathbb{V}(Z_f^*)} \{1 + O(E)\} + O(F).$$

En particulier, pour tout $\varepsilon > 0$, nous avons, lorsque $x \rightarrow \infty$ et $y \geq (1 + \varepsilon) \log_2 x \log_3 x$,

$$(5.19) \quad \frac{V_f(x, y)}{\mathbb{V}(Z_f)} = \frac{V_f^*(x, y)}{\mathbb{V}(Z_f^*)} \{1 + o(1)\} + o(1).$$

De plus, dans le domaine $2 \leq y \leq \log x$, nous avons

$$(5.20) \quad V_f(x, y) - V_f^*(x, y) \ll \mathbb{V}(Z_f) \left(\frac{y^{9/2}}{x^{\alpha/2} (\log x)^{3/2}} + \frac{y}{x^\alpha \log y} \right).$$

Remarques. (i) Compte tenu de (5.7), nous pouvons déduire de (5.19) que, pour tout $u \geq 1$ fixé,

$$(5.21) \quad \limsup_{x \rightarrow \infty} \sup_{f \in \mathbb{A}_0} \frac{V_f^*(x, x^{1/u})}{\mathbb{V}(Z_f^*)} = \lambda(u).$$

(ii) La minoration de x^α indiquée en (1.13), implique l'existence d'une constante positive convenable c telle que l'on ait $x^\alpha \gg e^{cy/\log y}$ dans le domaine $2 \leq y \leq \log x$. Cela permet de remplacer le membre de droite de (5.20) par $\mathbb{V}(Z_f) e^{-c_1 y/\log y}$ avec $c_1 > 0$.

2. En effet, la formule (2.1) de [4] s'écrit, avec les notations de [4],

$$V_f(x, y) = \mathbb{V}(Z_f) + T_f(x, y) + V_f^*(x, y) - U_f(x, y),$$

où, d'après la dernière formule centrée de l'article, $U_f(x, y) \ll B^2(x, y)/\bar{u}$.

L'inégalité $T_f(x, y) \leq C_1 B^2(x, y)/\bar{u}$, où C_1 est une constante absolue, est établie au milieu de la page 263.

Enfin, la formule (2.4) de [4] implique $V_f^*(x, y) = S^* + O(B^2(x, y)/\bar{u})$ où $S^* \leq C_2 B^2(x, y)/\bar{u}$ et C_2 désigne une constante absolue.

Démonstration. Nous avons vu en (3.12) que

$$(5.22) \quad V_f(x, y) - V_f^*(x, y) = \{\mathbb{E}(Z_f^*) - \mathbb{E}(Z_f)\} \{2(M_1 - \mathbb{E}(Z_f^*)) + \mathbb{E}(Z_f) - \mathbb{E}(Z_f^*)\}.$$

Compte tenu de la majoration (3.6), nous pouvons écrire

$$\{M_1 - \mathbb{E}(Z_f^*)\}^2 = \left\{ \sum_{p \leq y} f(p)r_p \right\}^2 \ll \frac{1}{\bar{u}^2} \sum_{p \leq y} \frac{f(p)^2 g_p(\alpha)}{p^\alpha} \sum_{p \leq y} \frac{g_p(\alpha)}{p^\alpha} \ll \mathbb{V}(Z_f^*) \frac{\bar{u} + \log_2 y}{\bar{u}^2},$$

où la dernière somme en p a été évaluée comme indiqué au lemme 3.2 de [3]. La majoration $w_p \ll (p/x)^\alpha$ mentionnée en (1.12) fournit de plus

$$\begin{aligned} \{\mathbb{E}(Z_f) - \mathbb{E}(Z_f^*)\}^2 &= \left\{ \sum_{p \leq y} f(p) \frac{w_p}{p^\alpha} \right\}^2 \leq \frac{1}{x^{2\alpha}} \sum_{p \leq y} \frac{f(p)^2 g_p(\alpha)}{p^\alpha} \sum_{p \leq y} \frac{p^\alpha}{g_p(\alpha)} \\ &\ll \frac{\pi(y)y^\alpha}{x^{2\alpha}} \left(\frac{1}{\alpha \log y} + 1 \right) \mathbb{V}(Z_f^*) \ll \frac{y^\alpha \{u + \pi(y)\}}{x^{2\alpha}} \mathbb{V}(Z_f^*). \end{aligned}$$

Or, d'après les formules (7.6) et (7.8) de [7], nous avons

$$y^\alpha \ll \left(1 + \frac{y}{\log x} \right) \log y \ll \frac{y + \log x}{u},$$

donc

$$\{\mathbb{E}(Z_f) - \mathbb{E}(Z_f^*)\}^2 \ll \mathbb{V}(Z_f^*) \frac{(y + \log x)^2}{x^{2\alpha} \log x}.$$

Il s'ensuit que

$$(5.23) \quad V_f(x, y) - V_f^*(x, y) \ll \mathbb{V}(Z_f^*) \frac{(y + \log x) \sqrt{(\bar{u} + \log_2 y)}}{x^\alpha \bar{u} \sqrt{\log x}} = \mathbb{V}(Z_f^*) F.$$

Par ailleurs, la majoration

$$\frac{w_p}{p^\alpha g_p(\alpha)} \ll \frac{1}{x^\alpha \alpha} \ll \frac{(\log x + y) \log y}{x^\alpha y},$$

implique

$$(5.24) \quad \mathbb{V}(Z_f^*) - \mathbb{V}(Z_f) = \sum_{p \leq y} \frac{f(p)^2 w_p}{p^{2\alpha}} \ll \mathbb{V}(Z_f^*) \frac{(y + \log x) \log y}{x^\alpha y} = \mathbb{V}(Z_f^*) E.$$

Compte tenu de l'identité

$$\frac{V_f(x, y)}{\mathbb{V}(Z_f)} \left(1 + \frac{\mathbb{V}(Z_f) - \mathbb{V}(Z_f^*)}{\mathbb{V}(Z_f^*)} \right) - \frac{V_f^*(x, y)}{\mathbb{V}(Z_f^*)} = \frac{V_f(x, y) - V_f^*(x, y)}{\mathbb{V}(Z_f^*)}.$$

les majorations (5.23) et (5.24) impliquent bien l'estimation (5.18) annoncée.

Pour établir (5.19), nous observons que, sous les hypothèses effectuées, nous avons certainement

$$\begin{aligned} E &\ll (\log x) e^{-(1-\varepsilon/3)y/\log y} \ll (\log x)^{-\varepsilon(1-\varepsilon)/2}, \\ F &\ll \sqrt{\log x} \left(1 + \frac{\vartheta(y)}{\log x} \right)^{-u} \ll (\log x)^{-(1-\varepsilon)/2}. \end{aligned}$$

Il reste à établir (5.20) dans le domaine (3.7). D'après (3.10) et (5.12), nous avons sous cette hypothèse

$$M_1 - \mathbb{E}(Z_f^*) = \sum_{p \leq y} f(p)r_p \ll \sum_{p \leq y} g_p(\alpha) \frac{|f(p)|}{p^\alpha} \left(\frac{1}{u} + \frac{y^2 \pi(y)}{(\log x)^2} \right),$$

d'où

$$(5.25) \quad \begin{aligned} \{M_1 - \mathbb{E}(Z_f^*)\}^2 &\ll \mathbb{V}(Z_f^*) \sum_{p \leq y} \frac{g_p(\alpha)}{p^\alpha} \left(\frac{1}{u} + \frac{y^2 \pi(y)}{(\log x)^2} \right)^2 \\ &\ll \mathbb{V}(Z_f^*) \left\{ \frac{y^2 \log y}{(\log x)^3} + \frac{y^5 \pi(y)^3}{(\log x)^5} \right\}. \end{aligned}$$

En reportant dans (5.22) les majorations (3.9) et (5.25), nous obtenons bien (5.20) sous les hypothèses effectuées. \square

5.3. Démonstration du Théorème 5.1.

Commençons par établir l'assertion (iv). En reportant (5.9) et (5.10) dans (3.13), nous obtenons bien (5.6) dans le domaine (3.7). Nous observons incidemment que (5.2) en découle dans le sous-domaine $y = o(\log x)$.

Considérons ensuite le point (iii). Selon l'hypothèse effectuée, il existe une fonction $h(x) \rightarrow \infty$ telle que

$$(5.26) \quad h(x) \leq y \leq (\log x)/h(x).$$

Nous avons alors $V_f^*(x, y) = \mathbb{V}(Z_f^*)\{1 + o(1)\}$ d'après (5.6). En reportant cette inégalité dans (5.20), nous obtenons

$$(5.27) \quad V_f(x, y) = \mathbb{V}(Z_f^*) + o(\mathbb{V}(Z_f)).$$

Or, en comparant les identités données en (3.8) pour $\mathbb{V}(Z_f)$ et en (3.14) pour $\mathbb{V}(Z_f^*)$, nous obtenons, lorsque y tend vers l'infini,

$$\frac{\mathbb{V}(Z_f)}{\mathbb{V}(Z_f^*)} \leq 1 + \max_{p \leq y} \left(\frac{p^\alpha}{x^\alpha g_p(\alpha)} \right) = 1 + \frac{1 + o(1)}{x^\alpha \alpha \log 2},$$

et

$$\frac{\mathbb{V}(Z_f)}{\mathbb{V}(Z_f^*)} \geq \{1 + o(1)\} \min_{p \leq y} \left(1 + \frac{p^\alpha}{x^\alpha g_p(\alpha)} \right) = \{1 + o(1)\} \frac{1 + x^\alpha \alpha \log y}{x^\alpha \alpha \log y}.$$

L'estimation (5.15) de α fournit bien (5.5) dans le domaine (5.26).

Pour établir l'assertion (ii), i.e. (5.4) dans le domaine

$$(5.28) \quad h(x) \leq y \leq x^{1/h(x)}$$

dès que $h(x) \rightarrow \infty$, nous observons d'abord que la conclusion souhaitée découle de (5.5) sous la condition (5.26). Comme le Lemme 5.3 fournit (5.4) sous la condition (5.16), nous obtenons bien (5.4) dans le domaine (5.28) quitte à choisir $h(x) \leq \log_2 x$ dans (5.26) et (5.16).

Il reste à prouver l'assertion (i). En vertu de (5.6), l'inégalité requise (5.2) est en fait une égalité sous la condition $y = o(\log x)$. Nous pouvons donc supposer

$$(5.29) \quad (\log x)/h(x) \leq y \leq x^{1/h(x)},$$

où $h(x)$ est une fonction qui tend vers $+\infty$. Or, dans ce domaine, la relation (5.19) implique que, si le rapport $V_f(x, y)/\mathbb{V}(Z_f)$ est $\leq 1 + o(1)$, alors il en va de même pour $V_f^*(x, y)/\mathbb{V}(Z_f^*)$. L'assertion (ii) fournit donc le résultat requis.

5.4. Le cas $f = \omega$

Établissons ici (5.8) sous l'hypothèse que y et $(\log x)/y$ tendent vers l'infini. D'après (3.2) et (1.20), nous avons

$$\mathbb{V}(Z_\omega) = \sum_{p \leq y} \frac{1 - w_p}{p^\alpha} \left(g_p(\alpha) + \frac{w_p}{p^\alpha} \right), \quad \mathbb{V}(Z_\omega^*) = \sum_{p \leq y} \frac{g_p(\alpha)}{p^\alpha}.$$

Comme (5.15) implique, dans la circonstance considérée, que l'on a $w_p = o(1)$ et $p^\alpha = 1 + o(1)$ uniformément pour $p \leq y$, il s'ensuit que

$$(5.30) \quad \mathbb{V}(Z_\omega) - \mathbb{V}(Z_\omega^*) = \{1 + o(1)\} \sum_{p \leq y} \frac{w_p}{p^\alpha} + o(\mathbb{V}(Z_\omega^*)).$$

Si, par exemple, $y > (\log_2 x)^2$, alors $\sup_{p \leq y} w_p/g_p(\alpha) = o(1)$. Nous obtenons donc $V_\omega \sim \mathbb{V}(Z_\omega^*)$, et (5.8) résulte de (5.27).

Dans le cas contraire, nous avons $\alpha \log y = o(1)$ et $\alpha \log x = \pi(y) + o(1)$ en vertu de (5.15), donc $w_p/p^\alpha \sim e^{-\pi(y)}$ uniformément pour $p \leq y$. Il vient donc

$$\sum_{p \leq y} \frac{w_p}{p^\alpha} \sim \pi(y) e^{-\pi(y)}$$

alors que

$$\mathbb{V}(Z_\omega^*) \sim \alpha \vartheta(y) \sim \frac{y\pi(y)}{\log x}.$$

Par (5.30), nous en déduisons que

$$\mathbb{V}(Z_\omega) \sim \left\{ 1 + \frac{\log x}{ye^{\pi(y)}} \right\} \mathbb{V}(Z_\omega^*).$$

En reportant dans (5.27), nous obtenons à nouveau (5.8).

5.5. Démonstration du Théorème 2.1

Soit f une fonction arithmétique additive. On peut écrire, lorsque $n \in S(x, y)$,

$$f(n) - \mathbb{E}(Z_f^*) = \sum_{p \leq y} c_{n,p} y_p$$

avec $y_p := f(p) \sqrt{g_p(\alpha)/p^\alpha}$ et

$$\begin{cases} c_{n,p} := \sqrt{p^\alpha/g_p(\alpha)} - \sqrt{g_p(\alpha)/p^\alpha} & \text{si } p \mid n, \\ c_{n,p} := -\sqrt{g_p(\alpha)/p^\alpha} & \text{si } p \nmid n. \end{cases}$$

La majoration (1.19) équivaut donc à la validité de

$$\sum_{n \in S(x,y)} \left| \sum_{p \leq y} c_{n,p} y_p \right|^2 \leq C \Psi(x, y) \sum_{p \leq y} |y_p|^2$$

pour tout choix des nombres complexes y_p . Le principe de dualité exprimant l'égalité des normes d'un opérateur d'espaces de Hilbert et de son adjoint (cf., par exemple, le lemme I.4.5.1 de [9]) montre alors que cette majoration est équivalente à la validité de l'inégalité

$$\sum_{p \leq y} \left| \sum_{n \in S(x,y)} c_{n,p} a_n \right|^2 \leq C \Psi(x, y) \sum_{n \in S(x,y)} |a_n|^2$$

pour toute suite $\{a_n\}_{n=1}^\infty$ de nombres complexes.

6. La répartition des facteurs premiers d'un entier friable

6.1. Preuve du Théorème 2.2

Posons

$$W_t := \mathbb{V}(Z_{\omega_t}^*) = \sum_{p \leq t} \frac{g_p(\alpha)}{p^\alpha} \ll \min(1, \alpha \log t) M_t \quad (2 \leq t \leq y \leq x).$$

Il résulte du Théorème 1.2 que, pour chaque t fixé dans $[2, y]$, $y \leq \log x$,

$$(6.1) \quad \sum_{\substack{n \in S(x, y) \\ |\omega_t(n) - M_t| > h\sqrt{W_t}}} 1 \ll \frac{\Psi(x, y)}{h^2} \quad (x \geq y \geq 2, h > 0).$$

Observons d'emblée que, si $\alpha = o(1)$, alors, pour tout entier $T = T(x, y)$ tel que $\alpha T = o(1)$, nous pouvons choisir, par exemple, $h := (\alpha T)^{-1/4}$ pour obtenir $\omega_T(n) = M_T + o(1)$ pour tous les entiers n de $S(x, y)$ sauf au plus $o(\Psi(x, y))$ d'entre eux. Comme $M_T = \pi(T) + O(\alpha T)$, cela implique en fait $\omega_T(n) = \pi(T)$, avec le même nombre d'exceptions.

Pour traiter le cas général, nous introduisons une constante positive c assez petite, rappelons la notation $z := \min\{y, \exp(1/\alpha)\}$ et, dans un premier temps, introduisons l'ensemble $\mathcal{K}(x, y)$ des entiers naturels k tels que

$$M_2 \leq c\alpha k^3 (\log k)^{1+b} \leq M_z.$$

Pour tout k de $\mathcal{K}(x, y)$, il existe alors un $t_k \in [2, z]$ tel que

$$M_{t_k} \leq c\alpha k^3 (\log k)^{1+b} \leq M_{t_k} + 1.$$

Cela résulte du fait que la fonction $t \mapsto M_t$ croît par sauts n'excédant pas l'unité.

Pour tous les entiers k de $\mathcal{K}(x, y)$, nous avons $W_{t_k} \ll \alpha M_{t_k} \log t_k$ et, en vertu de (2.2), $M_{t_k} \asymp t_k / \log t_k$, donc $t_k \ll \alpha k^3 (\log k)^{b+2}$. En spécialisant $t = t_k$ et $h \asymp k^{1/2} (\log k)^{b/2}$ dans (6.1), nous obtenons que l'estimation

$$(6.2) \quad |\omega_{t_k}(n) - M_{t_k}| \leq c^2 \alpha k^2 (\log k)^{1+b} \leq c\alpha^{1/3} M_{t_k}^{2/3} \{\log(3M_{t_k}/\alpha)\}^{2b/3}$$

a lieu, avec un choix convenable de c , pour tous les entiers n de $S(x, y)$ sauf peut-être pour au plus $\ll \Psi(x, y) / \{k(\log k)^b\}$ d'entre eux. Comme

$$M_{t_{k+1}} - M_{t_k} \leq c^{3/4} \alpha k^2 (\log k)^{1+b} + 1 \leq c^{1/12} \alpha^{1/3} M_{t_k}^{2/3} \{\log(3M_{t_k}/\alpha)\}^{2b/3} + 1,$$

et compte tenu la croissance en t de $\omega_t(n)$, il s'ensuit, par sommation sur k , que nous obtenons bien la validité de la première estimation (2.3) dans les conditions de l'énoncé.

Dans le cas $z < t \leq y$, et donc $\alpha > 1/\log y$, l'inégalité (2.3) a été établie au théorème 1.6 de [3]. Pour la commodité du lecteur, rappelons rapidement les détails. Nous introduisons des points-tests $s_k \in [z, y]$ tels que $M_{s_k} \leq ck^3 (\log k)^b < M_{s_k} + 1$. Pour le même choix de h que précédemment dans (6.1), nous obtenons alors que

$$\omega_{s_k}(n) - M_{s_k} \ll c^{1/3} M_{s_k}^{2/3} (\log M_{s_k})^{b/3}$$

pour toutes les valeurs admissibles de k . Cela implique bien le résultat annoncé.

6.2. Preuve du Corollaire 2.3

Appliquons (2.3) avec $t = p_j(n)$ et donc $\omega_t(n) = j$. Compte tenu de l'approximation $M_t - \pi(t) \ll \alpha t \ll \pi(t)$, la première majoration fournit $p_j(n) \leq e^{1/\alpha}$ sous l'hypothèse $p_j \leq e^{1/2\alpha}$ dès que x , et donc J_x , est assez grand. Il suit

$$\pi(p_j(n)) - j = \pi(p_j(n)) - \pi(p_j) \ll \alpha^{1/3} j^{2/3} \{\log(3j/\alpha)\}^{2b/3} + 1.$$

Cela implique $\log p_j(n) \asymp \log p_j$ et donc

$$p_j(n) - p_j \ll \{\pi(p_j(n)) - \pi(p_j)\} \log p_j.$$

On en déduit immédiatement la première estimation de (2.5).

Observons par ailleurs que (2.3) implique

$$\omega_t(n) - M_t \ll M_t^{2/3} (\log M_t)^{2b/3} \quad (2 \leq t \leq y).$$

En spécialisant $t = p_j(n)$ et en tenant compte de (2.2), nous obtenons, pour j assez grand,

$$(6.3) \quad \log w_j + \frac{e^{w_j} - 1}{w_j} + O\left(\frac{e^{w_j} - 1}{w_j^2}\right) = j$$

où l'on a posé $w_j := (1 - \alpha) \log p_j(n)$. Sous l'hypothèse $p_j > e^{1/2\alpha}$, et donc $\alpha \log j \gg 1$, le dernier terme d'erreur est $\ll \alpha j$. En inversant le terme principal du membre de gauche de (6.3) comme indiqué dans [3], nous obtenons,

$$w_j = \log(j \log j) + \frac{\log_2 j}{\log j} + O(\alpha) = \log p_j + O(\alpha),$$

où nous avons fait appel à l'estimation classique $p_j = j \log(j \log j) + O(j)$. Cela implique la seconde estimation de (2.5).

La précision finale de l'énoncé découle immédiatement de la dernière assertion du Théorème 2.2.

Bibliographie

- [1] K. Alladi, The Turán–Kubilius inequality for integers without large prime factors, *J. reine angew. Math.* **335** (1982), 180–196.
- [2] R. de la Bretèche & G. Tenenbaum, Propriétés statistiques des entiers friables, *Ramanujan J.* **9** (2005), 139–202.
- [3] R. de la Bretèche & G. Tenenbaum, Entiers friables : inégalité de Turán–Kubilius et applications, *Invent. Math.* **159** (2005), 531–588.
- [4] R. de la Bretèche & G. Tenenbaum, On the friable Turán–Kubilius inequality, Proceedings of the Fifth International Conference *Analytic and Probabilistic Methods in Number Theory* held on the occasion of Professor Jonas Kubilius' 90th birthday (2012), E. Manstavičius et al. (eds), TEV Vilnius 2012, 259–265.
- [5] R. de la Bretèche & G. Tenenbaum, Une nouvelle approche dans la théorie des entiers friables, *Composition Math.* **153** (2017), 453–473.
- [6] G. Hanrot, B. Martin & G. Tenenbaum, Constantes de Turán–Kubilius friables : étude numérique, *Exp. Math. Experiment. Math.* **19**, n° 3 (2010), 345–361.
- [7] A. Hildebrand & G. Tenenbaum, On integers free of large prime factors, *Trans. Amer. Math. Soc.* **296** (1986), n° 1, 265–290.

- [8] B. Martin & G. Tenenbaum, Sur l'inégalité de Turán-Kubilius friable, *J. reine angew. Math.* **647** (2010), 175–234.
- [9] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, troisième édition, coll. Échelles, Belin, 2008, 592 pp.
- [10] T.Z. Xuan, The Turán-Kubilius inequality for integers free of large prime factors, *J. Number Theory* **43** (1993), 82–87.
- [11] T.Z. Xuan, The Turán-Kubilius inequality for integers free of large prime factors (II), *Acta Arith.* **65** (1993), 329–352.

Régis de la Bretèche
Université Paris Diderot-Paris 7
Sorbonne Paris Cité, UMR 7586
Institut de Mathématiques de Jussieu-PRG
Case 7012, F-75013 Paris
France

`regis.delabreteche@imj-prg.fr`

Gérald Tenenbaum
Institut Élie Cartan
Université de Lorraine
BP 70239
F-54506 Vandœuvre-lès-Nancy Cedex
France

`gerald.tenenbaum@univ-lorraine.fr`