*Proc. Steklov Inst. Math.*
**314** (2021), 264–278.

# Effective Erdős-Wintner theorems

Gérald Tenenbaum & Johann Verwee

*In memory of Ivan Matveievitch Vinogradov,*
*who bequeathed to us*
*elegance in theorems and vision in methods.*

**Abstract.** The classical theorem of Erdős & Wintner furnishes a criterion for the existence of a limiting distribution for a real, additive arithmetical function. This work is devoted to providing an effective estimate for the remainder term under the assumption that the conditions in the criterion are fulfilled. We also investigate the case of a conditional distribution.

**Keywords:** distribution of real additive functions, mean values of complex multiplicative function, Erdős-Wintner theorem, effective averages, number of prime factors.

**2010 Mathematics Subject Classification:** 11N25, 11N37, 11N60.

## 1. Introduction and statement of results

The classical theorem of Erdős & Wintner [3], [5], is the analogue in probabilistic number theory of Kolmogorov's three series theorem in probability theory. It asserts that a real, additive arithmetical function $f$ possesses a limiting distribution if, and only if,

$$\sum_{p\in\mathcal{P}} \frac{\min\big(1, f(p)^2\big)}{p} < \infty, \quad \text{and} \quad \sum_{\substack{p\in\mathcal{P}\\ |f(p)|\leqslant 1}} \frac{f(p)}{p} \text{ converges;} \tag{1·1}$$

here and in the sequel, $\mathcal{P}$ denotes the set of primes. Moreover it follows from a theorem of Lévy [8] that the limit law is continuous if, and only if,

$$\sum_{f(p)\neq 0} \frac{1}{p} = \infty, \tag{1·2}$$

while a well-known theorem of Jessen and Wintner [6] tells us that this limit law is necessarily pure. See, e.g., [10; ch. III.4] for proofs and historical comments.

In this work, our first aim is to exploit a recent result of the first author [11] on mean values of complex multiplicative functions in order to provide an effective version of the Erdős–Wintner theorem, or, in other words, to furnish an effective estimate for the supremum norm

$$\|F_x - F\|_\infty := \sup_{y\in\mathbb{R}} |F_x(y) - F(y)| \qquad (x \geqslant 1)$$

where, for each $x \geqslant 1$,

$$F_x(y) := \frac{1}{x} \sum_{\substack{n\leqslant x\\ f(n)\leqslant y}} 1 \qquad (y \in \mathbb{R})$$

is the empirical distribution function and $F$ is the limiting distribution. It is well known that $F$ has characteristic function

$$\varphi_F(\tau) := \int_{\mathbb{R}} \mathrm{e}^{i\tau y}\,\mathrm{d}F(y) = \prod_p \Big(1 - \frac{1}{p}\Big) \sum_{\nu\geqslant 0} \frac{\mathrm{e}^{i\tau f(p^\nu)}}{p^\nu} \qquad (\tau \in \mathbb{R}). \tag{1·3}$$

We state our results in this direction as two separate theorems, corresponding respectively to the discrete and the continuous case.

Let us first consider the situation when (1·1) is realised but (1·2) is not. We then define a multiplicative function $u_f$ by its values on prime powers

$$u_f(p^\nu) := \begin{cases} 1 & \text{if } f(p^\nu) \neq 0, \\ 0 & \text{if } f(p^\nu) = 0, \end{cases} \tag{1·4}$$

and, given a prime $p \in \mathcal{P}$, write

$$S_p = S_p(f) := \sum_{\nu \geqslant 1} \frac{u_f(p^\nu)}{p^\nu}, \quad w_p = w_p(f) := \Big(1 - \frac{1}{p}\Big) S_p(f), \tag{1·5}$$

so that the convergence of the series on the left-hand side of (1·2) implies the absolute convergence of $\sum_p w_p$. We also plainly have

$$\alpha_f(y) := \sum_{p > y} \frac{u_f(p)}{p} \to 0, \quad \beta_f(y) := \frac{1}{\log y} \int_1^y \frac{\alpha_f(t)}{t} \,\mathrm{d}t \to 0 \quad (y \to \infty). \tag{1·6}$$

Writing

$$h_f(m) := u_f(m) \prod_{p \mid m} \frac{1 - 1/p}{1 - w_p} \quad (m \geqslant 1),$$

we easily check that

$$F(y) := \prod_p (1 - w_p) \sum_{f(m) \leqslant y} \frac{h_f(m)}{m} \qquad (y \in \mathbb{R})$$

is a distribution function, indeed

$$\sum_{m \geqslant 1} \frac{h_f(m)}{m} = \prod_p \Big(1 + \frac{1 - 1/p}{1 - w_p} S_p\Big) = \prod_p \Big(1 + \frac{w_p}{1 - w_p}\Big) = \prod_p \frac{1}{1 - w_p}.$$

With these notations, we can state our first result. Here and in the sequel, we let $\log_k$ denote the $k$-fold iterated logarithm.

**Theorem 1.1.** *Let $f$ be a real additive function satisfying* (1·1) *but not* (1·2). *Then, uniformly for $x \geqslant 2$, we have*

$$\|F_x - F\|_\infty \ll R_x := \alpha_f\big(x^{1/\log_2 x}\big) + \beta_f\big(\sqrt{x}\big)^{1/4} + \frac{1}{(\log x)^{1/6}}.$$

*Examples.* (i) Let $\kappa > 0$ be a parameter and consider an additive function $f$ such that $f(p) = 1$ if $2^n < p \leqslant 2^n(1 + 1/(\log n)^\kappa)$ for some $n \geqslant 3$ and $f(p) = 0$ otherwise. In this setting, the limit law is atomic—i.e. (1·2) fails—if, and only if, $\kappa > 1$. We then have

$$\alpha_f(y) \asymp \beta_f(y) \asymp \frac{1}{(\log_2 y)^{\kappa - 1}}, \quad R_x \asymp \frac{1}{(\log_2 x)^{(\kappa - 1)/4}}.$$

(ii) Assume now that $f(p) = 1$ if $2^n < p \leqslant 2^n(1 + 1/n^\kappa)$ for some $n \geqslant 1$, while $f(p) = 0$ otherwise. Then the series (1·2) converges for all $\kappa > 0$ and we have

$$\alpha_f(y) \asymp \frac{1}{(\log y)^\kappa}, \quad \beta_f(y) \asymp \frac{(\log_2 y)^{\delta_{1\kappa}}}{(\log y)^{\min(1,\kappa)}}, \quad R_x \asymp \frac{1}{(\log x)^{\min(2/3,\kappa)/4}},$$

with Kronecker's notation $\delta_{1\kappa}$.

(iii) More generally, when the non-zero values of $f(p)$ are distributed with sufficient regularity, a simple criterion for the continuity of the limit law may be stated and subsequent estimates for $R_x$ may then be easily computed. Indeed, writing

$$\{p \in \mathcal{P} : f(p) \neq 0\} = \mathcal{P} \cap \big(\cup_{k \geqslant 1} ]a_k, b_k]\big)$$

where the $a_k$, $b_k$ are integers, $2 \leqslant a_k < b_k$, we first observe that this set is certainly infinite provided

$$b_k > a_k + a_k^{1-c} \qquad (k \geqslant 1) \tag{1·7}$$

for sufficiently small, positive $c$: it follows from [1] that, with $c = 0.475$, we have $\pi(x+y) - \pi(x) \asymp y/\log x$ for $x^{1-c} \leqslant y \leqslant x$—the sharpest estimate of Hoheisel type to date. Appealing to this result and to the prime number theorem in the form

$$\sum_{a<p\leqslant b} \frac{1}{p} = \log\Big(\frac{\log b}{\log a}\Big) + O\Big(\mathrm{e}^{-\sqrt{\log a}}\Big) \qquad (b \geqslant a \geqslant 3),$$

it is a simple matter to deduce that, assuming (1·7), condition (1·2) holds if, and only if,

$$\sum_{k\geqslant 1} \log\Big(\frac{\log b_k}{\log a_k}\Big) = \infty.$$

We next turn our attention to the case when (1·1) and (1·2) are both satisfied, which implies that the limiting distribution $F$ is continuous. We then let $\eta_f(y)$ denote any continuous, non-increasing function tending to 0 at infinity and such that

$$\Bigg|\sum_{\substack{p>y\\ |f(p)|\leqslant 1}} \frac{f(p)}{p}\Bigg| \leqslant \eta_f(y), \quad \sum_{p^\nu > y} \frac{\min(1, f(p^\nu)^2)}{p^\nu} \leqslant \eta_f(y) \qquad (y \geqslant 1). \tag{1·8}$$

For $x \geqslant 2$, we consider a quantity $\varepsilon_x$ such that $1/\sqrt{\log x} < \varepsilon_x = o(1)$, and assume henceforth that $\varepsilon_x$ approaches 0 so slowly that

$$\eta_f\big(x^{\varepsilon_x}\big) = o\Big(\varepsilon_x^{1/3}\Big) \qquad (x \to \infty). \tag{1·9}$$

We write furthermore

$$B_f(v)^2 := 2 + \sum_{p^\nu \leqslant v} \frac{f(p^\nu)^2}{p^\nu} \qquad (v \geqslant 1),$$

and let $\ell \mapsto Q_F(\ell) := \sup_{y\in\mathbb{R}} \big\{F(y+\ell) - F(y)\big\}$ denote the concentration function associated to $F$. Since $F$ is continuous, we know that $Q_F(\ell) \to 0$ as $\ell \to 0$. Effective upper bounds, depending explicitly on the sequence $\{f(p)\}_{p\in\mathcal{P}}$ or on $\varphi_F$ are available in the literature: see, e.g., [2], [4], [7], [9], and [10; ch. III.2]. For instance, the Kolmogorov–Rogozin inequality implies

$$Q_F(\ell) \ll \frac{1}{\sqrt{1 + \sum_{|f(p)|>\ell} 1/p}}, \tag{1·10}$$

while a simple computation (see, e.g., [10; lemma III.2.9]) provides

$$Q_F(\ell) \ll \ell \int_{-1/\ell}^{1/\ell} |\varphi_F(\tau)| \,\mathrm{d}\tau. \tag{1·11}$$

**Theorem 1.2.** *Uniformly for all real additive functions $f$ satisfying* (1·1) *and* (1·2), *and all $R \geqslant 3$, $T \geqslant 3$, $x \geqslant 3$, such that*

$$2\log_2 R + \tfrac{1}{2}T^2\eta_f(R) + 7 \leqslant \tfrac{1}{4}\log(1/\varepsilon_x), \quad T^2\eta_f(x^{\varepsilon_x}) \ll \varepsilon_x^{1/3}, \tag{1·12}$$

*we have*

$$\|F_x - F\|_\infty \ll Q_F\Big(\frac{1}{T}\Big) + \varepsilon_x^{1/6} \log\Big(\frac{TB_f(R)}{\varepsilon_x}\Big) + \eta_f(R). \tag{1·13}$$

*Remarks.* (i) When, for instance, $\xi > 1$ and $f$ is the strongly additive function defined by $f(p) = 1/(\log p)^\xi$, an estimate of Koukoulopoulos [7] sharpening a result of La Bretèche & Tenenbaum [2] yields $Q_F(\ell) \asymp \ell^{1/\xi}$ $(0 < \ell \leqslant 1/3)$. Then, $B_f(R) \asymp 1$, $\eta_f(y) \asymp 1/(\log y)^\xi$, the choice

$$\varepsilon_x = 2/\sqrt{\log x}, \quad R = \mathrm{e}^{c(\log x)^{1/16}}, \quad T = (\log x)^{\xi/32}$$

is admissible for suitably small $c > 0$, and we get, ignoring some negative powers of $\log_2 x$,

$$\|F_x - F\|_\infty \ll \frac{1}{(\log x)^{1/32}}.$$

(ii) The general estimate (1·13) may be superseded by specific known results when $f(p)$ shows rapid and smooth decrease. For instance, if $f(p) = 1/p^\xi$ with $\xi > 0$, $f(p^\nu) = 0$ $(\nu \geqslant 2)$, we have $Q(\ell) \asymp 1/|\log \ell|$ $(0 < \ell \leqslant 1/3)$ by [2; cor. 1.3]. The optimal choice is then

$$\varepsilon_x \asymp 1/\sqrt{\log x}, \quad R = \mathrm{e}^{c(\log x)^{1/16}}, \quad \log T \asymp (\log x)^{1/24},$$

and we only get

$$\|F_x - F\|_\infty \ll \frac{1}{(\log x)^{1/24}},$$

while the left-hand side is actually $\ll (\log_2 x)/\{(\log x)\log_3 x\}$, in view of [2; Cor. 1.5], which includes the cases of the additive functions $\log\{n/\varphi(n)\}$ and $\log\{\sigma(n)/n\}$ where $\varphi$ is Euler's totient and $\sigma(n) := \sum_{d|n} d$. This lack of precision may be traced back to the use of the general upper bounds (4·5) and (4·6) *infra*, which only integrate partial information on the distribution of the $f(p)$: when $f(p)$ is quickly decreasing, a direct bound for the difference of the characteristic functions furnishes the stated sharpening.

The technique involved in the proofs of the above results is actually fairly flexible. As an illustration, we present a further effective theorem, describing how the distribution of an additive function fluctuates when restricting the support to integers with a fixed number of prime factors. To avoid technicalities we focus on the case of a strongly additive function with continuous distribution, but a completely general statement could be achieved by the same method.

Let $\omega(n)$ denote the number of distinct prime factors of an integer $n$ and, for $x \geqslant 1$, let $\pi_k(x)$ represent the cardinality of the level set $\mathcal{E}(x;k) := \{n \leqslant x : \omega(n) = k\}$. We have, classicaly (see e.g. [10; ch. II.6])

$$\pi_k(x) \asymp \frac{x(\log_2 x)^{k-1}}{(k-1)!\log x} \qquad (1 \leqslant k \ll \log_2 x),$$

and we may replace $k - 1$ by $k$ when $k \asymp \log_2 x$.

Given a strongly additive function $f$ satisfying (1·1), we consider for each $r > 0$ the characteristic function

$$\varphi(\tau; r) := \prod_p \Big(1 + \frac{r\mathrm{e}^{i\tau f(p)}}{p-1}\Big)\Big(1 + \frac{r}{p-1}\Big)^{-1} \qquad (\tau \in \mathbb{R}), \tag{1·14}$$

and denote as $\mathcal{F}_r$ the corresponding distribution function.

Our estimate depends on the function $\eta_f$ defined in (1·8). We furthermore introduce parameters $v$, $T$ and $R$ such that

$$\begin{gathered}\frac{1}{\log_2 x} \leqslant v \leqslant c_0, \quad 3 \leqslant R \leqslant \mathrm{e}^{1/v}, \quad T \geqslant 1,\\ T^2\eta_f(R) \leqslant \log(1/v), \quad T^2\eta_f(x^w) \ll w \quad (w := v^{c_1}),\end{gathered} \tag{1·15}$$

where $c_0$ and $c_1$ denote strictly positive constants, depending at most on $\kappa$, $c_0$ being sufficiently small and $c_1$ sufficiently large.

**Theorem 1.3.** *Let $\kappa \in ]0,1[$ and let $f$ be a real, strongly additive function. Assume* (1·1) *and* (1·2) *hold. Then, uniformly for $\kappa \leqslant r := k/\log_2 x \leqslant 1/\kappa$, $y \in \mathbb{R}$, and $v$, $T$, $R$ satisfying* (1·15), *we have*

$$\frac{1}{\pi_k(x)} \sum_{\substack{n\in\mathcal{E}(x;k)\\ f(n)\leqslant y}} 1 = \mathcal{F}_r(y) + O\big(\mathfrak{R}\big) \tag{1·16}$$

*with*

$$\mathfrak{R} := Q_{\mathcal{F}_r}\Big(\frac{1}{T}\Big) + \Big(v + \frac{\log(1/v)}{\sqrt{k}}\Big)\log\Big(\frac{TB_f(R)}{v}\Big) + \eta_f(R)^{r/(r+1)}.$$

Due to the generality of the hypotheses, this statement turns out as rather technical. Indeed an optimal choice of the parameters heavily depends on the sequence $\{f(p)\}_{p\in\mathcal{P}}$. However, an explicit estimate easily follows in non-pathological situations. As an example, consider the case when $f(p) := 1/(\log p)^\xi$ with $0 < \xi < r$. It is then easy to show (see, e.g., [12; Exercise 259]) that $|\varphi(\tau;r)| = |\tau|^{-r/\xi}(\log|\tau|)^{O(1)}$ as $|\tau| \to \infty$ and hence, by (1·11), that $Q_{\mathcal{F}_r}(\ell) \ll \ell$ as $\ell \to 0$. We may therefore select

$$v := 1/\log_2 x, \quad R := \log x, \quad T := (\log_2 x)^{\xi/2},$$

and infer that $\mathfrak{R} \ll (\log_3 x)^2/\sqrt{\log_2 x} + 1/(\log_2 x)^{\xi\min\{1/2,r/(r+1)\}}$.

## 2. The key argument

Our approach rests on the following recent result of the first author [11; th. 1.2], for the statement of which we introduce further notation. We let $\mathcal{M}(A,B)$ designate the class of those complex-valued multiplicative functions $g$ such that

$$\max_p |g(p)| \leqslant A, \qquad \sum_{p,\,\nu\geqslant 2} \frac{|g(p^\nu)|\log p^\nu}{p^\nu} \leqslant B, \tag{2·1}$$

and, for $\mathfrak{b} \in \mathbb{R}$, we write

$$\beta = \beta(\mathfrak{b},A) := 1 - \frac{\sin(2\pi\mathfrak{b}/A)}{2\pi\mathfrak{b}/A}\cdot \tag{2·2}$$

Moreover, given any complex-valued function $g$, we put $\mathfrak{c}_g := 1$ if $g$ is real, $\mathfrak{c}_g := 2$ otherwise, and consider

$$M(x;g) := \sum_{n\leqslant x} g(n), \qquad Z(x;g) := \sum_{p\leqslant x}\frac{g(p)}{p}\cdot$$

**Theorem 2.1 ([11]).** *Let*

$$\mathfrak{a} \in ]0,\tfrac14], \quad \mathfrak{b} \in [\mathfrak{a},\tfrac12[, \quad \mathfrak{h} := (1-\mathfrak{b})/\mathfrak{b}, \quad A \geqslant 2\mathfrak{b}, \quad B > 0, \quad \beta := \beta(\mathfrak{b},A),$$
$$2\mathfrak{b} \leqslant \varrho \leqslant A, \quad x \geqslant 2, \quad 1/\sqrt{\log x} < \varepsilon \leqslant \tfrac12,$$

*and let the multiplicative functions $g$, $r$, such that $r \in \mathcal{M}(2A,B)$, $|g| \leqslant r$, satisfy the conditions*

$$\sum_{p\leqslant x}\frac{r(p) - \Re e\, g(p)}{p} \leqslant \tfrac12\beta\mathfrak{b}\log(1/\varepsilon), \tag{2·3}$$

$$\sum_{x^\varepsilon<p\leqslant y}\frac{\{r(p)-\Re e\, g(p)\}^{\mathfrak{h}}\log p}{p} \ll \varepsilon^{\mathfrak{c}_g\delta\mathfrak{h}}\log y \qquad (x^\varepsilon < y \leqslant x), \tag{2·4}$$

$$\sum_{p\leqslant y}\frac{(r(p)-\varrho)\log p}{p} \ll \varepsilon\log y \qquad (x^\varepsilon < y \leqslant x) \tag{2·5}$$

*with $\delta \in ]0, 2\beta\mathfrak{b}/(3\mathfrak{c}_g)\big]$.*

*We then have*

$$M(x;g)=\frac{\mathrm{e}^{-\gamma\varrho}x}{\Gamma(\varrho)\log x}\Big\{\prod_p\sum_{p^\nu\leqslant x}\frac{g(p^\nu)}{p^\nu}+O\Big(\varepsilon^\delta\mathrm{e}^{Z(x;g)}\Big)\Big\},\tag{2·6}$$

*where* $\gamma$ *denotes Euler's constant. The implicit constant in* (2·6) *depends at most upon* $A$*,* $B$*,* $\mathfrak{a}$*, and* $\mathfrak{b}$*.*

## 3. Proof of Theorem 1.1

Let $u_f$ be defined by (1·4) and let $v_f$ be the multiplicative function defined of prime powers by $v_f(p^\nu):=1-u_f(p^\nu)$. Then any integer $n\geqslant 1$ may be uniquely represented as a product $n=md$ with $u_f(m)=v_f(d)=1$, $(m,d)=1$ and $f(n)=f(m)$. Therefore

$$F_x(y)=\frac{1}{x}\sum_{\substack{m\leqslant x\\ f(m)\leqslant y}}u_f(m)V_m\Big(\frac{x}{m}\Big)\tag{3·1}$$

with

$$V_m(t):=\sum_{\substack{d\leqslant t\\ (d,m)=1}}v_f(d)=\sum_{d\leqslant t}v_f(d;m)\qquad(t\geqslant 1),$$

say, where, for each $m$, $v_f(d;m)$ is the multiplicative function of $d$ defined on prime powers by $v_f(p^\nu;m)=v_f(p^\nu)$ if $p\nmid m$ and $=0$ otherwise. For $t\geqslant 3$, $m\leqslant t$, we have, by the convergence of (1·2),

$$\sum_{p\leqslant t}\frac{1-v_f(p;m)}{p}\leqslant\sum_{p|m}\frac{1}{p}+O(1)\leqslant\log_3 t+O(1)$$

and, similarly,

$$\begin{aligned}\sum_{p\leqslant t}\frac{\{1-v_f(p;m)\}\log p}{p}&\leqslant\sum_{p\leqslant t}\frac{\{1-v_f(p)\}\log p}{p}+\sum_{p|m}\frac{\log p}{p}\\&=\int_1^t\sum_{u<p\leqslant t}\frac{u_f(p)}{p}\frac{\mathrm{d}u}{u}+O(\log_2 t)\leqslant\beta_f(t)\log t+O(\log_2 t),\end{aligned}$$

with notation (1·6).

Hence, for $\varepsilon:=\beta_f(t)^{3/4}+1/\sqrt{\log t}$, we have, for $t^\varepsilon<s\leqslant t$,

$$\sum_{t^\varepsilon<p\leqslant s}\frac{\{1-v_f(p;m)\}\log p}{p}\ll\beta_f(t)^{1/4}\log s+O(\log_2 s)\ll\varepsilon^{1/3}\log t.$$

We may therefore estimate $V_m(t)$, uniformly in $m\leqslant t$, by applying Theorem 2.1 to $g:=v_f(\cdot;m)$ with

$$\mathfrak{b}=\tfrac12,\quad \varrho=A=\mathfrak{c}_g=\beta=\mathfrak{h}=1,\quad \varepsilon:=\beta_f(t)^{3/4}+1/\sqrt{\log t},\quad \delta=\tfrac13.$$

We get, for $1\leqslant m\leqslant t$,

$$V_m(t)=\Big\{1+O\Big(\frac{1}{\log 2t}\Big)\Big\}t\psi_m(t)+O\Big(t\beta_f(t)^{1/4}+\frac{t}{(\log 2t)^{1/6}}\Big),$$

with

$$\psi_m(t) := \prod_{p\leqslant t}\Big(1-\frac{1}{p}\Big)\prod_{\substack{p\leqslant t\\ p\nmid m}}\Big(\frac{1}{1-1/p}-S_p+O\Big(\frac{1}{t}\Big)\Big)$$
$$=\Big\{1+O\Big(\frac{1}{\log 2t}\Big)\Big\}\prod_{p\leqslant t}(1-w_p)\prod_{p\mid m}\Big(\frac{1-1/p}{1-w_p}\Big),$$

where $S_p$, $w_p$, are defined in (1·5) and we have taken into account that $1/(1-1/p)-S_p\geqslant 1$. Since $w_p\leqslant 1/p$, we have $\log(1-w_p)\geqslant -2w_p$, whence

$$\prod_{p>t}(1-w_p)\geqslant \exp\Big\{-2\sum_{p>t}w_p\Big\}\geqslant 1-O\big(\alpha_f(t)\big),$$

where $\alpha_f$ is defined in (1·6). This yields

$$u_f(m)V_m(t)=\prod_p(1-w_p)th_f(m)+O\big(tu_f(m)R_0(t)\big)\qquad (t\geqslant m\geqslant 1)$$

with $R_0(t):=\alpha_f(t)+\beta_f(t)^{1/4}+1/(\log 2t)^{1/6}$. Splitting the sum in (3·1) at $m=\lfloor\sqrt{x}\rfloor$ and considering that $V_m(t)\leqslant t$, we readily obtain, uniformly for $y\in\mathbb{R}$,

$$F_x(y)=F(y)+O\big(E_1+E_2\big)$$

with

$$E_1:=\sum_{m>\sqrt{x}}\frac{u_f(m)}{m},\quad E_2:=R_0\big(\sqrt{x}\big)\sum_{m\leqslant\sqrt{x}}\frac{u_f(m)}{m}\ll R_0\big(\sqrt{x}\big),$$

where the last bound follows from the convergence of the series in (1·2).

In order to bound $E_1$, we introduce a parameter $T\geqslant 2$ and split the summation according to whether the largest prime factor of $m$, say $P^+(m)$, exceeds $T$ or not. We obtain, for any $\sigma\in]0,\frac{1}{3}[$

$$E_1\leqslant \sum_{\substack{m>\sqrt{x}\\ P^+(m)\leqslant T}}\frac{1}{m}+\sum_{p>T}\sum_{\nu\geqslant 1}\frac{u(p^\nu)}{p^\nu}\sum_{m\geqslant 1}\frac{u_f(m)}{m}$$
$$\ll\frac{1}{x^{\sigma/2}}\prod_{p\leqslant T}\Big(1+\frac{1}{p^{1-\sigma}}\Big)+\alpha_f(T)+\frac{1}{T}$$

For large $T$, we select $\sigma:=4/\log T$. The last $p$-product is then $\ll\log T$, and so

$$E_1\ll x^{-2/\log T}\log T+\alpha_f(T)+\frac{1}{T}.$$

The required estimate follows by selecting $T:=x^{1/\log_2 x}$.

## 4. Proof of Theorem 1.2

Given $R\geqslant 3$, we define the additive function $f_R$ by

$$f_R(p^\nu):=\begin{cases} f(p^\nu) & \text{if } p^\nu\leqslant R \text{ or } |f(p^\nu)|\leqslant 1,\\ 0 & \text{in all other cases.}\end{cases}\tag{4·1}$$

Denote by $F_x(y;R)$ the distribution function of $f_R$ on the set of integers not exceeding $x$ and by $F(y;R)$ that of the limit law. We first observe that, when $x\in\mathbb{N}^*$,

$$|F_x(y;R)-F_x(y)|\leqslant\sum_{\substack{p^\nu>R\\ |f(p^\nu)|>1}}\frac{1}{p^\nu}\leqslant\eta_f(R)\qquad(y\in\mathbb{R}),\tag{4·2}$$

the same bound being valid for $|F(y;R)-F(y)|$ since the inequality is independent of $x$. We may hence restrict to evaluating $F_x(y;R)-F(y;R)$ with the perspective of ultimately optimising the parameter $R$.

Note that, for $3 \leqslant R \leqslant x$,

$$\sum_{p^\nu \leqslant x} \frac{f_R(p^\nu)}{p^\nu} = \sum_{p^\nu \leqslant R} \frac{f(p^\nu)}{p^\nu} + \sum_{\substack{R<p^\nu \leqslant x \\ |f(p^\nu)|\leqslant 1}} \frac{f(p^\nu)}{p^\nu} \ll B_f(R)\sqrt{\log_2 R},$$

$$\sum_{p^\nu \leqslant x} \frac{f_R(p^\nu)^2}{p^\nu} = \sum_{p^\nu \leqslant R} \frac{f(p^\nu)^2}{p^\nu} + \sum_{\substack{R<p^\nu \leqslant x \\ |f(p^\nu)|\leqslant 1}} \frac{f(p^\nu)^2}{p^\nu} \ll B_f(R)^2,$$

where we used (1·1) to bound the last sum. By the Turán-Kubilius inequality, it follows, still for $3 \leqslant R \leqslant x$, that

$$\begin{aligned}\frac{1}{x}\sum_{n\leqslant x}\big(\mathrm{e}^{i\tau f_R(n)} - 1\big) &= \frac{i\tau}{x}\sum_{n\leqslant x} f_R(n) + O\Big(\frac{\tau^2}{x}\sum_{n\leqslant x} f_R(n)^2\Big) \\ &\ll |\tau| B_f(R)\sqrt{\log_2 R} + \tau^2 B_f(R)^2 \log_2 R.\end{aligned} \tag{4·3}$$

Writing

$$\varphi_x(\tau;R) := \frac{1}{x}\sum_{n\leqslant x} \mathrm{e}^{i\tau f_R(n)}, \qquad \varphi(\tau;R) := \int_{\mathbb{R}} \mathrm{e}^{i\tau y}\,\mathrm{d}F(y;R) = \prod_p \Big(1-\frac{1}{p}\Big)\sum_{\nu\geqslant 0}\frac{\mathrm{e}^{i\tau f_R(p^\nu)}}{p^\nu},$$

and considering that the upper bound in (4·3) does not depend on $x$, we hence see that

$$\frac{\varphi_x(\tau;R) - \varphi(\tau;R)}{\tau} \ll B_f(R)\sqrt{\log_2 R} + |\tau| B_f(R)^2 \log_2 R \qquad (\tau\in\mathbb{R}). \tag{4·4}$$

This estimate will be used for dealing with small values of $|\tau|$.

Next we evaluate $\varphi_x(\tau;R)$ when $|\tau|$ is not too close to 0, $|\tau| \leqslant T$, and assuming (1·12). We have, for large $x$,

$$\begin{aligned}\sum_{p\leqslant x}\frac{1-\cos(\tau f_R(p))}{p} &\leqslant \sum_{p\leqslant R}\frac{2}{p} + \sum_{\substack{R<p\leqslant x\\ |f(p)|\leqslant 1}} \frac{\tau^2 f(p)^2}{2p} \qquad (\tau\in\mathbb{R}),\\ &\leqslant 2\log_2 R + 7 + \tfrac12 T^2 \eta_f(R) \leqslant \tfrac14 \log(1/\varepsilon_x),\end{aligned} \tag{4·5}$$

(where we used the estimate $\sum_{p\leqslant y} 1/p \leqslant \log_2 y + 7/2$ $(y\geqslant 2)$ which follows by partial summation from Mertens' first theorem in the form given for instance in [10; th. I.1.8]) and similarly, for $|\tau|\leqslant T$, since $z_x := x^{\varepsilon_x} \geqslant R$ by (1·12),

$$\begin{aligned}\sum_{z_x<p\leqslant y} \frac{\{1-\cos(\tau f_R(p))\}\log p}{p} &\leqslant \tfrac12\tau^2 \sum_{\substack{z_x<p\leqslant y\\ |f(p)|\leqslant 1}} \frac{f(p)^2}{p}\log p \qquad (x^{\varepsilon_x} < y \leqslant x).\\ &\ll T^2\eta_f(z_x)\log y \ll \varepsilon_x^{1/3}\log y\end{aligned} \tag{4·6}$$

We may hence apply Theorem 2.1 to $g := \mathrm{e}^{i\tau f_R}$, with $A = \varrho = 1$, $\mathfrak{b} = \frac12$, $\mathfrak{h} = 1$, $\beta = 1$, $r = \mathbf{1}$, and $\varepsilon = \varepsilon_x$. This yields

$$\begin{aligned}\varphi_x(\tau;R) &= \prod_{p\leqslant x}\Big(1-\frac1p\Big)\sum_{p^\nu\leqslant x}\frac{\mathrm{e}^{i\tau f_R(p^\nu)}}{p^\nu} + O\big(\varepsilon_x^{1/6}\big)\\ &= \prod_{p\leqslant x}\Big(1-\frac1p\Big)\sum_{\nu\geqslant 0}\frac{\mathrm{e}^{i\tau f_R(p^\nu)}}{p^\nu} + O\big(\varepsilon_x^{1/6}\big),\end{aligned}$$

where we used the inequality $|\prod_p(u_p+v_p) - \prod_p u_p| \leqslant \sum_p |v_p|$, valid for all $u_p$, $v_p$ such that $|u_p|\leqslant 1$, $|u_p+v_p|\leqslant 1$.

Since

$$\begin{aligned}\prod_{p>x}\Big(1-\frac{1}{p}\Big)\sum_{\nu\geqslant 0}\frac{\mathrm{e}^{i\tau f_R(p^\nu)}}{p^\nu}&=\prod_{p>x}\Big(1+\frac{\mathrm{e}^{i\tau f_R(p)}-1}{p}+O\Big(\frac{1}{p^2}\Big)\Big)\\&=\exp\Bigg\{\sum_{\substack{p>x\\|f(p)|\leqslant 1}}\frac{i\tau f(p)}{p}+O\Big(\frac{\tau^2 f(p)^2}{p}\Big)+O\Big(\frac{1}{x\log x}\Big)\Bigg\}\\&=\exp\Big\{O\Big(\eta_f(x)(1+\tau^2)+\frac{1}{x\log x}\Big)\Big\},\end{aligned}$$

(using $|\tau|\leqslant 1+\tau^2$) we eventually obtain, for $|\tau|\leqslant T$,

$$\varphi_x(\tau;R)=\varphi(\tau;R)\Big\{1+O\Big(\varepsilon_x^{1/6}\Big)\Big\}+O\big(\varepsilon_x^{1/6}\big),$$

and so

$$\varphi_x(\tau;R)=\varphi(\tau;R)+O\big(\varepsilon_x^{1/6}\big). \tag{4·7}$$

This enables an appeal to the Berry-Esseen inequality

$$\|F_x(\cdot;R)-F(\cdot;R)\|_\infty\ll Q\Big(\frac{1}{T};R\Big)+\int_{-T}^{T}\left|\frac{\varphi_x(\tau;R)-\varphi(\tau;R)}{\tau}\right|\mathrm{d}\tau,$$

where $Q(\cdot;R)$ is the concentration function associated to $F(\cdot;R)$. Taking (4·2) into account, we get

$$\|F_x-F\|_\infty\ll Q_F\Big(\frac{1}{T}\Big)+\int_{-T}^{T}\left|\frac{\varphi_x(\tau;R)-\varphi(\tau;R)}{\tau}\right|\mathrm{d}\tau+\eta_f(R).$$

To bound the last integral, say $I$, from above, we introduce a parameter $u\in]0,1[$ and apply (4·4) for $|\tau|\leqslant u$, then (4·7) for $u<|\tau|\leqslant T$. This yields

$$\begin{aligned}I&\ll uB_f(R)\sqrt{\log_2 R}+u^2B_f(R)^2\log_2 R+\varepsilon_x^{1/6}\log(T/u)\\&\ll\varepsilon_x^{1/6}\log\Big(\frac{TB_f(R)}{\varepsilon_x}\Big)\end{aligned}$$

for the quasi-optimal choice $u:=\varepsilon_x^{1/6}/\{B_f(R)\sqrt{\log_2 R}\}$.

## 5. Proof of Theorem 1.3

Let $f$ be strongly additive, satisfying (1·1) and (1·2), and for $R\geqslant 3$ let $f_R$ be defined by (4·1). We start with a lemma showing that, for large $R$, we have $f_R(n)=f(n)$ for most integers $n\in\mathcal{E}(x;k)$. We recall the notation $r:=k/\log_2 x$ and put

$$\sigma_f(R):=\eta_f(R)^{r/(r+1)}+1/(\log x)^{r/(r+1)}.$$

**Lemma 5.1.** *Let $\kappa\in]0,1[$. Uniformly for $\kappa\leqslant r:=k/\log_2 x\leqslant 1/\kappa$, $3\leqslant R\leqslant x$, we have*

$$\sum_{\substack{n\in\mathcal{E}(x;k)\\f_R(n)\neq f(n)}}1\ll\sigma_f(R)\pi_k(x). \tag{5·1}$$

*Proof.* We may plainly assume $x$ to be large and hence that $k \geqslant 2$. Put

$$\mathcal{P}_R := \{p \in \mathcal{P} : p > R,\ |f(p)| > 1\}, \quad E_R(x) := \sum_{\substack{p \leqslant x \\ p \in \mathcal{P}_R}} \frac{1}{p} \leqslant \eta_f(R).$$

The quantity to be bounded does not exceed the number of those integers $n \in \mathcal{E}(x;k)$ having at least one prime divisor in $\mathcal{P}_R$.

From the classical Hardy-Ramanujan estimate for $\pi_k(y)$ (see e.g. [10; Ex. 264]) the left-hand side of (5·1) is, for an absolute constant $a$,

$$\leqslant \sum_{n \in \mathcal{E}(x;k)} \sum_{\substack{p^\nu \| n \\ p \in \mathcal{P}_R}} 1 \ll \sum_{\substack{p^\nu \leqslant x \\ p \in \mathcal{P}_R}} \pi_{k-1}\Big(\frac{x}{p^\nu}\Big) \ll \sum_{\substack{p^\nu \leqslant x \\ p \in \mathcal{P}_R}} \frac{x\{\log_2(3x/p^\nu) + a\}^{k-2}}{p^\nu (k-2)! \log(2x/p^\nu)}.$$

Put $v := \sigma_f(R)^{1/r}$. The subsum corresponding to $p^\nu \leqslant x^{1-v}$ is plainly

$$\ll \frac{x(\log_2 x)^{k-2} E_R(x)}{v(k-2)! \log x} \ll \frac{\sigma_f(R) x (\log_2 x)^{k-2}}{(k-2)! \log x} \ll \sigma_f(R)\pi_k(x).$$

The complementary subsum may be dealt with by partial summation. By the prime number theorem, it is

$$\begin{aligned} &\ll \frac{1}{(k-2)!} \int_{x^{1-v}}^{x} \frac{x\{\log_2(3x/t) + a\}^{k-2}}{t \log(2x/t) \log t}\,\mathrm{d}t \asymp \frac{x}{(k-2)! \log x} \int_1^{x^v} \frac{(\log_2 3u + a)^{k-2}}{u \log 2u}\,\mathrm{d}u \\ &\ll \frac{x\{\log_2 x - \log(1/v) + a\}^{k-1}}{(k-1)! \log x} \ll \pi_k(x)\Big\{1 - \frac{\log(1/v)}{\log_2 x}\Big\}^k \ll \pi_k(x) v^r = \sigma_f(R)\pi_k(x). \end{aligned}$$

□

Our next lemma consists in obtaining a uniform upper bound for

$$S_R(x;\tau,z) := \sum_{n \leqslant x} z^{\omega(n)} \mathrm{e}^{i\tau f_R(n)} \qquad (x \geqslant 1,\ |z| = r).$$

**Lemma 5.2.** *Let* $\kappa \in ]0,1[$. *Uniformly for* $3 \leqslant R \leqslant \log x$, $\kappa \leqslant r \leqslant 1/\kappa$, $z = r\mathrm{e}^{i\vartheta}$, $|\vartheta| \leqslant \pi$, $|\tau| \leqslant T$, *we have*

$$S_R(x;\tau,z) \ll x(\log x)^{r-1}\Big\{\frac{\mathrm{e}^{9rT^2\eta_f(R)}}{(\log x)^{r\vartheta^2/60}} + \frac{1}{\sqrt{\log x}}\Big\}. \tag{5·2}$$

*Proof.* By [11; cor. 2.1], the left-hand side of (5·2) is

$$\ll x(\log x)^{r-1}\Big\{\frac{1 + m_f(x;\tau)}{\mathrm{e}^{m_f(x;\tau)}} + \frac{1}{\sqrt{\log x}}\Big\} \tag{5·3}$$

with

$$m_f(x;\tau) = r \min_{|t| \leqslant \log x} \sum_{p \leqslant x} \frac{1 - \cos\{\vartheta + t\log p + \tau f_R(p)\}}{p}.$$

Let $\|a\|$ denote the distance of the real number $a$ to the set of integers. The elementary inequality $\|a+b\|^2 \geqslant \frac{1}{2}\|a\|^2 - b^2$ and the standard lower bound $1 - \cos a \geqslant 8\|a/2\pi\|^2$ yield

$$m_f(x;\tau) \geqslant 4r \min_{|t| \leqslant \log x} \lambda_f(x;t) - 8r\tau^2\eta_f(R) \tag{5·4}$$

with

$$\lambda_f(x,t) := \sum_{R < p \leqslant x} \frac{1}{p} \Big\|\frac{\vartheta + t\log p}{2\pi}\Big\|^2.$$

Now by [10; lemma III.4.13], we have, restricting the $p$-sum to $y<p\leqslant x$ with $y>R$,

$$\lambda_f(x;t)\geqslant \tfrac{1}{12}\log\Big(\frac{\log x}{\log y}\Big)+O\Big(\frac{1}{|t|\log y}+\frac{1+|t|}{\mathrm{e}^{\sqrt{\log y}}}\Big)\qquad(2\leqslant y\leqslant x).\tag{5·5}$$

If $1\leqslant |t|\leqslant \log x$, we select $y:=\exp\big\{(\log_2 x)^2\big\}$ to get

$$\lambda_f(x;t)\geqslant \tfrac{1}{12}\log_2 x+O(\log_3 x).$$

Let us then define $\nu:=(\log x)^{(\vartheta^2/2\pi^2)-1}$. If $\nu\leqslant |t|\leqslant 1$, we select $y=\mathrm{e}^{1/\nu}$ in (5·5) and obtain

$$\lambda_f(x;t)\geqslant \frac{\vartheta^2}{24\pi^2}\log_2 x+O(1).$$

Finally, if $|t|\leqslant\nu$, we have

$$\begin{aligned}\lambda_f(x;t)&\geqslant\sum_{\log x<p\leqslant \mathrm{e}^{1/\nu}}\frac{\vartheta^2/4\pi^2+O(\nu\log p)}{p}\geqslant\frac{\vartheta^2(1-\vartheta^2/2\pi^2)}{4\pi^2}\log\Big(\frac{\log x}{\log_2 x}\Big)+O(1)\\&\geqslant\frac{\vartheta^2}{8\pi^2}\log\Big(\frac{\log x}{\log_2 x}\Big)+O(1).\end{aligned}\tag{5·6}$$

Carrying back into (5·4) and (5·3) yields the stated estimate since $1/6\pi^2>1/60$. □

We now deduce from Theorem 2.1 an asymptotic formula with remainder for $S_R(x;\tau,z)$ when $z$ belongs to a neighbourhood of the real point $r$ on the circle $|z|=r$.

**Lemma 5.3.** *Let $\kappa\in]0,1[$. For a suitable constant $c_0\in]0,1]$, arbitrary large $c_1>0$, and uniformly under the assumptions $\kappa\leqslant r\leqslant 1/\kappa$ and*

$$\begin{aligned}&z=r\mathrm{e}^{i\vartheta},\ \frac{1}{\log_2 x}\leqslant v\leqslant c_0,\ |\vartheta|\leqslant\vartheta_x:=\sqrt{\frac{c_1\log(1/v)}{\log_2 x}},\ 1+|\tau|\leqslant T,\\&3\leqslant R\leqslant \mathrm{e}^{1/v},\quad T^2\eta_f(R)\leqslant\log(1/v),\quad T^2\eta_f(x^w)\ll w\quad(w:=v^{c_1}),\end{aligned}\tag{5·7}$$

*we have*

$$S_R(x;\tau,z)=\frac{x\mathrm{e}^{-\gamma r}}{\log x}\Big\{\prod_{p\leqslant x}\Big(1+\frac{z\mathrm{e}^{i\tau f_R(p)}}{p-1}\Big)+O\Big((|\vartheta|+v^2)(\log x)^r\Big)\Big\},\tag{5·8}$$

*Proof.* We apply Theorem 2.1 with $r(n):=r^{\omega(n)}$, $g(n):=z^{\omega(n)}\mathrm{e}^{i\tau f_R(n)}$, $\mathfrak{b}:=\frac12\min(1,r)$, $A:=\max(1,r)$, $\varrho:=r$, $\delta:=c_2\beta\mathfrak{b}$, and $\varepsilon:=(|\vartheta|+v^2)^{1/\delta}$. We select $c_2$ so small to ensure that $2\delta\mathfrak{h}\leqslant 1$, where $\mathfrak{h}=(1-\mathfrak{b})/\mathfrak{b}$.

Since

$$\begin{aligned}\sum_{p\leqslant x}\frac{r(p)-\Re e\, g(p)}{p}&=\sum_{p\leqslant R}\frac{r-\Re e\, g(p)}{p}+r\sum_{\substack{R<p\leqslant x\\|f(p)|\leqslant 1}}\frac{1-\cos\big(\vartheta+\tau f(p)\big)}{p}+O(1)\\&\leqslant 2r\log_2 R+r\vartheta^2\log_2 x+rT^2\eta_f(R)+O(1)\\&\leqslant (c_1+3)r\log(1/v)+O(1),\end{aligned}$$

we see that condition (2·3) is satisfied for an appropriate choice of $c_0$ and $c_2$: indeed, this is clear if $v^2>|\vartheta|$ for then $1/v<\sqrt{2}/\varepsilon^{\delta/2}$, and, if $v^2\leqslant|\vartheta|$, we have $\varepsilon^\delta\leqslant 2\vartheta_x$ whence $\log(1/v)\leqslant\log_3 x\ll\log 1/\vartheta_x\ll\delta\log(1/\varepsilon)$.

Next, since $v^{c_1} \leqslant \varepsilon$ provided $c_1 \geqslant 2/\delta$, we have, for $x^\varepsilon < y \leqslant x$,

$$\sum_{\substack{x^\varepsilon < p \leqslant y\\ |f(p)|\leqslant 1}} \frac{r(1-\cos\{\vartheta+\tau f(p)\})^{\mathfrak{h}} \log p}{p} \ll \big(|\vartheta|^{2\mathfrak{h}} + T^2\eta_f(x^\varepsilon)\big)\log y \ll \varepsilon^{2\delta\mathfrak{h}}\log y,$$

and so condition (2·4) is also satisfied. Considering the fact that (2·5) holds trivially, we obtain

$$S_R(x;\tau,z) = \frac{x\mathrm{e}^{-\gamma r}}{\log x}\bigg\{\prod_{p\leqslant x}\Big(1+\frac{z\mathrm{e}^{i\tau f_R(p)}}{p-1}\Big) + O\Big(\varepsilon^\delta \mathrm{e}^{z\sum_{p\leqslant x}\mathrm{e}^{i\tau f_R(p)}/p}\Big)\bigg\}.$$

The required estimate hence follows from a trivial estimate for the last sum over $p$. □

We are now in a position to embark on the final part of the proof.
Define

$$L(\tau;x) := \sum_{p\leqslant x}\frac{\mathrm{e}^{i\tau f_R(p)}}{p-1},\qquad G_\tau(z;x) := \mathrm{e}^{-zL(\tau;x)}\prod_{p\leqslant x}\Big(1+\frac{z\mathrm{e}^{i\tau f_R(p)}}{p-1}\Big).$$

Under conditions (5·7) for $\tau$, we have

$$\begin{aligned}|L(0;x)-L(\tau;x)| &\leqslant 2\log_2 R + T\eta_f(R) + \tfrac12 T^2\eta_f(R) + O(1)\\ &\leqslant 4\log(1/v) + O(1) \ll \log k,\end{aligned}\tag{5·9}$$

in particular $L(\tau;x) = \log_2 x + O(\log k)$, while of course $L(0;x) = \log_2 x + O(1)$. Moreover, $G_\tau(z;x)$ is an entire function of $z$ which is uniformly bounded on any compact subset with respect to $\tau$ and $x$, so we have for instance

$$G_\tau^{(j)}(0;x)/j! \ll 1/(1+r)^j \quad (j\geqslant 0).\tag{5·10}$$

We now apply Cauchy's integral formula to $S_R(x;\tau,z)$ for the circle $|z| = r = k/\log_2 x$, under hypotheses (1·15). This yields

$$\sum_{\substack{n\leqslant x\\ \omega(n)=k}} \mathrm{e}^{i\tau f_R(n)} = \frac{x\mathrm{e}^{-\gamma r}(I_1 - I_2 + I_3)}{\log x} + I_4$$

with

$$I_1 := \frac{1}{2\pi i}\oint_{|z|=r}\mathrm{e}^{zL(\tau;x)}G_\tau(z;x)\frac{\mathrm{d}z}{z^{k+1}},\quad I_2 := \frac{1}{2\pi i}\int_{\substack{|z|=r\\ \vartheta_x<|\vartheta|\leqslant\pi}}\mathrm{e}^{zL(\tau;x)}G_\tau(z;x)\frac{\mathrm{d}z}{z^{k+1}},$$

$$I_3 := \frac{\mathrm{e}^k}{2\pi i}\int_{\substack{|z|=r\\ |\vartheta|\leqslant\vartheta_x}} O(|\vartheta|+v^2)\frac{\mathrm{d}z}{z^{k+1}},\quad I_4 := \frac{1}{2\pi i}\int_{\substack{|z|=r\\ \vartheta_x<|\vartheta|\leqslant\pi}} S_R(x;\tau,z)\frac{\mathrm{d}z}{z^{k+1}}.$$

The main term is provided by $I_1$, equal to the coefficient of $z^k$ in $\mathrm{e}^{zL(\tau;x)}G_\tau(z;x)$. We thus have

$$I_1 = \sum_{0\leqslant j\leqslant k}\frac{L(\tau;x)^{k-j}G_\tau^{(j)}(0;x)}{(k-j)!j!} = \frac{L(\tau;x)^k}{k!}\Big\{G_\tau(r;x) + O\Big(\frac{\log k}{k}\Big)\Big\},\tag{5·11}$$

by (5·9) and (5·10), after a short computation involving truncating the sum at $\lfloor\sqrt{k}\rfloor$, for instance, and noting that

$$\frac{1}{(k-j)!L(\tau;x)^j} = \frac{r^j}{k!}\Big\{1+O\Big(\frac{j^2+j\log k}{k}\Big)\Big\}\qquad (0\leqslant j\leqslant\sqrt{k}).$$

The integrals $I_j$ $(2 \leqslant j \leqslant 4)$ are treated as error terms. We first have, by (5·9),

$$I_2 \ll \int_{\vartheta_x}^{\pi} \mathrm{e}^{-rL(0;x)\cos\vartheta}\,\mathrm{d}\vartheta \ll \frac{\sqrt{k}(\log_2 x)^k}{v^{4r}k!}\int_{\vartheta_x}^{\pi} \mathrm{e}^{-2k\vartheta^2/\pi^2}\,\mathrm{d}\vartheta$$
$$\ll \frac{\mathrm{e}^{-k\vartheta_x^2/\pi^2}(\log_2 x)^k}{v^{4r}k!} \ll \frac{v^{(c_1/\pi^2-4)r}(\log_2 x)^k}{k!} \ll \frac{v(\log_2 x)^k}{k!},$$

since $c_1$ may be taken arbitrarily large in terms of $\kappa$. Next

$$I_3 \ll \frac{\mathrm{e}^k}{r^k}\{\vartheta_x^2 + v^2\vartheta_x\} \ll \frac{x(\log_2 x)^k}{k!\log x}\Big\{v + \frac{\log 1/v}{\sqrt{k}}\Big\} \asymp \pi_k(x)\Big\{v + \frac{\log 1/v}{\sqrt{k}}\Big\},$$

and finally, by (5·2), still with a suitable choice of $c_1$,

$$I_4 \ll \frac{x\mathrm{e}^k}{r^k\log x}\Big\{\frac{v^{c_1 r/60}}{v^{9r}\sqrt{k}} + \frac{1}{\sqrt{\log x}}\Big\} \ll v\pi_k(x).$$

Gathering our estimates, we arrive at

$$\sum_{\substack{n\leqslant x\\ \omega(n)=k}} \mathrm{e}^{i\tau f_R(n)} = \frac{x\mathrm{e}^{-\gamma r}L(\tau;x)^k G_\tau(r;x)}{k!\log x} + O\Big(\pi_k(x)\Big(v + \frac{\log(1/v)}{\sqrt{k}}\Big)\Big),$$

since the error term of (5·11) may be absorbed by the other remainders. Applying this with $\tau = 0$, we get

$$\pi_k(x) = \frac{x\mathrm{e}^{-\gamma r}L(0;x)^k G_0(r;x)}{k!\log x}\Big\{1 + O\Big(v + \frac{\log(1/v)}{\sqrt{k}}\Big)\Big\},$$

and so, recalling definition (1·14),

$$\varphi_x(\tau;k) := \frac{1}{\pi_k(x)}\sum_{\substack{n\leqslant x\\ \omega(n)=k}} \mathrm{e}^{i\tau f_R(n)}$$
$$= \frac{L(\tau;x)^k G_\tau(r;x)}{L(0;x)^k G_0(r;x)}\Big\{1 + O\Big(v + \frac{\log(1/v)}{\sqrt{k}}\Big)\Big\} + O\Big(v + \frac{\log(1/v)}{\sqrt{k}}\Big)$$
$$= \frac{L(\tau;x)^k G_\tau(r;x)}{L(0;x)^k G_0(r;x)} + O\Big(v + \frac{\log(1/v)}{\sqrt{k}}\Big) = \varphi(\tau;r) + O\Big(v + \frac{\log(1/v)}{\sqrt{k}}\Big),$$

in view of (5·9) and since

$$\Big|\sum_{p>x}\frac{\mathrm{e}^{i\tau f_R(p)}-1}{p-1}\Big| \leqslant T\eta_f(x) + \tfrac{1}{2}T^2\eta_f(x) + O\Big(\frac{1}{x}\Big) \ll v.$$

It remains to apply the Berry-Esseen inequality, taking into account a variant of (4·3) conditioned to $\omega(n) = k$ in order to handle the contribution of small $\tau$. Assuming (1·15), we get, for $0 < u \leqslant T$,

$$\sup_{y\in\mathbb{R}}\Big|\frac{1}{\pi_k(x)}\sum_{\substack{n\in\mathcal{E}(x;k)\\ f_R(n)\leqslant y}} 1 - \mathcal{F}_r(y)\Big| \ll Q_{\mathcal{F}_r}\Big(\frac{1}{T}\Big) + \int_{-T}^{T}\Big|\frac{\varphi_x(\tau;k)-\varphi(\tau;r)}{\tau}\Big|\,\mathrm{d}\tau$$
$$\ll Q_{\mathcal{F}_r}\Big(\frac{1}{T}\Big) + uB_f(R)\sqrt{\log_2 R} + u^2B_f(R)^2 + \Big(v + \frac{\log(1/v)}{\sqrt{k}}\Big)\log\frac{T}{u}$$
$$\ll Q_{\mathcal{F}_r}\Big(\frac{1}{T}\Big) + \Big(v + \frac{\log(1/v)}{\sqrt{k}}\Big)\log\Big(\frac{TB_f(R)}{v}\Big),$$

with the quasi-optimal choice $u = v/\{B_f(R)\sqrt{\log_2 R}\}$. The required estimate now follows from (5·1).

*Acknowledgements.* The authors wish to thank the referee for a thorough checking of this work and useful suggestions. The first author takes pleasure in expressing warm thanks to Régis de la Bretèche for useful remarks on a preliminary draft.

## References


[1] R.C. Baker, G. Harman & J. Pintz, The difference between consecutive primes, II, *Proc. London Math. Soc.* (3) **83** n° 3 (2001), 532–562.
[2] R. de la Bretèche & G. Tenenbaum, Sur la concentration de certaines fonctions additives, *Math. Proc. Camb. Phil. Soc.* **152**, n° 1 (2012), 179–189; erratum *ibid.*, 191.
[3] P. Erdős, On the density of some sequences of numbers I, *J. London Math. Soc.* **10** (1935), 120–125; II, *ibid.* **12** (1937), 7–11 ; III, *ibid.* **13** (1938), 119–127.
[4] P. Erdős & I. Kátai, On the concentration of distribution of additive functions, *Acta Sci. Math. (Szeged)* **41** n$^{\text{os}}$ 3-4 (1979), 295–305.
[5] P. Erdős & A. Wintner, Additive arithmetical functions and statistical independence, *Amer. J. Math.* **61** (1939), 713–721.
[6] B. Jessen & A. Wintner, Distribution functions and the Riemann Zeta function, *Trans. Amer. Math. Soc.* **38** (1935), 48–88.
[7] D. Koukoulopoulos, On the concentration of certain additive functions, *Acta Arith.* **162**, no. 3 (2014), 223–241.
[8] P. Lévy, Sur les séries dont les termes sont des variables éventuelles indépendantes, *Studia Math.* **3** (1931), 119–155.
[9] I.Z. Ruzsa, On the concentration of additive functions, *Acta Math. Acad. Sci. Hungar.* **36** (1980), no. 3-4, 215–232.
[10] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, 3rd ed., Graduate Studies in Mathematics 163, Amer. Math. Soc. 2015.
[11] G. Tenenbaum, Moyennes effectives de fonctions multiplicatives complexes, *Ramanujan J.* **44**, no. 3 (2017), 641–701; corrigendum, *ibid.* **52** no. 4 (2020).
[12] G. Tenenbaum & J. Wu (coll.), *Théorie analytique et probabiliste des nombres, 307 exercices corrigés*, coll. Échelles, Belin, 2014, 347 pp.



Gérald Tenenbaum
Institut Élie Cartan
Université de Lorraine
BP 70239
54506 Vandœuvre-lès-Nancy Cedex
France
gerald.tenenbaum@univ-lorraine.fr

Johann Verwee
Institut Élie Cartan
Université de Lorraine
BP 70239
54506 Vandœuvre-lès-Nancy Cedex
France
johann.verwee@univ-lorraine.fr