L'Enseignement Mathématique (2) 61 (2015), 45–69 DOI 10.4171/LEM/61-1/2-3

On Karamata's proof of the Landau–Ingham Tauberian theorem

Michael Müger

Abstract. This is a self-contained exposition of (a generalization of) Karamata's little known elementary proof of the Landau–Ingham Tauberian theorem, a result in real analysis from which the Prime Number Theorem follows in a few lines.

Mathematics Subject Classification (2010). Primary: 11N05; Secondary: 40E05.

Keywords. Prime number theorem, elementary proof, Tauberian theorem.

1. Introduction

The aim of this paper is to give a self-contained, accessible and 'elementary' proof of of the following theorem, which we call the Landau-Ingham Tauberian theorem:

Theorem 1.1. *Let* $f : [1,\infty) \to \mathbb{R}$ *be non-negative and non-decreasing and assume that*

(1.1)

$$F(x) := \sum_{n\le x} f\left(\frac{x}{n}\right) \quad \text{satisfies} \quad F(x) = Ax\log x + Bx + C\frac{x}{\log x} + o\left(\frac{x}{\log x}\right).$$

Then $f(x) = Ax + o(x)$, *equivalently* $f(x) \sim Ax$.

The interest of this theorem derives from the fact that, while ostensibly it is a result firmly located in classical real analysis, the prime number theorem (PNT) $\pi(x) \sim \frac{x}{\log x}$ can be deduced from it by a few lines of Chebychev-style reasoning (cf. the Appendix).

Versions of Theorem 1.1 were proven by Landau [Lan, §160] as early as 1909, Ingham [Ing, Theorem 1], Gordon [Gor] and Ellison [Ell, Theorem 3.1], but none of these proofs was from scratch. Landau used as input the identity

$\sum_n \frac{\mu(n)\log n}{n} = -1$. But the latter easily implies $M(x) = \sum_{n\leq x} \mu(n) = o(x)$ which (as also shown by Landau) is equivalent to the PNT. Actually, $\sum_n \frac{\mu(n)\log n}{n} = -1$ is 'stronger' than the PNT in the sense that it cannot be deduced from the latter (other than by elementarily reproving the PNT with a sufficiently strong remainder estimate). In this sense, Gordon's version of Theorem 1.1 is an improvement, in that he uses as input exactly the PNT (in the form $\psi(x) \sim x$) and thereby shows that Theorem 1.1 is not 'stronger' than the PNT. Ellison's version assumes $M(x) = o(x)$ (and an $O(x^\beta)$ remainder with $\beta < 1$ in (1.1)). It is thus clear that none of these approaches provides a proof of the PNT. Ingham's proof, on the other hand, starts from the information that $\zeta(1+it) \neq 0$ (which can be deduced from the PNT, but also be proven *ab initio*). Thus his proof is not 'elementary', but arguably it is one of the nicer and more conceptual deductions of the PNT from $\zeta(1+it) \neq 0$ – though certainly not the simplest (which is [Zag]) given that the proof requires Wiener's L^1-Tauberian theorem.

Our proof of Theorem 1.1 will essentially follow the elementary Selberg-style proof given by Karamata [1] [Karl] under the assumption that f is the summatory function of an arithmetic function, i.e. constant between successive integers. We will remove this assumption. For the proof of the PNT, this generality is not needed, but from an analysis perspective it seems desirable, and it brings us fairly close to Ingham's version of the theorem, which differed only in having $o(x)$ instead of $C\frac{x}{\log x} + o(\frac{x}{\log x})$ in the hypothesis.

Unfortunately, Karamata's paper [Karl] seems to be essentially forgotten: There are so few references to it that we can discuss them all. It is mentioned in [EI] by Erdős and Ingham and in the book [Ell] of Ellison and Mendès-France. (Considering that the latter authors know Karamata's work, one may find it surprising that for their elementary proof of the PNT they chose the somewhat roundabout route of giving a Selberg-style proof of $M(x) = o(x)$, using this to prove a weak version of Theorem 1.1, from which then $\psi(x) \sim x$ is deduced.) Even the two books [BGT, Kor] on Tauberian theory only briefly mention Karamata's [Karl] (or just the survey paper [Kar2]) but then discuss in detail only Ingham's proof. Finally, [Karl, Kar2] are cited in the recent historical article [Nik], but its emphasis is on other matters. We close by noting that Karamata is not even mentioned in the only other paper pursuing an elementary proof of a Landau-Ingham theorem, namely Balog's [Ba], where a version of Theorem 1.1 with a (fairly weak) error term in the conclusion is proven.

Our reason for advertising Karamata's approach is that, in our view, it is the conceptually cleanest and simplest of the Selberg–Erdős style proofs of the PNT,

[1] Note des éditeurs : Jovan Karamata, né près de Belgrade en 1902 et mort à Genève en 1967, fut professeur à Genève dès 1951 et directeur de L'Enseignement Mathématique de 1954 à 1967. Voir M. Tomić, *Jovan Karamata (1902-1967)*, Enseignement Math. (2) **15**, 1–20 (1969).

cf. [Sel, Erd1] and followers, e.g. [PR, Nev, Kal, Lev, Sch, Pol]. For $f = \psi$ and $f(x) = M(x) + \lfloor x \rfloor$, Theorem 1.1 readily implies $\psi(x) = x + o(x)$ and $M(x) = o(x)$, respectively. Making these substitutions in advance, the proof simplifies only marginally, but it becomes less transparent (in particular for $f = \psi$) due to an abundance of non-linear expressions. By contrast, Theorem 1.1 is linear w.r.t. f and F. To be sure, also the proof given below has a non-linear core, cf. (3.2) and Proposition 3.14, but by putting the latter into evidence, the logic of the proof becomes clearer. One is actually led to believe that the non-linear component of the proof is inevitable, as is also suggested by Theorem 2 in Erdős' [Erd2], to wit

$$a_k \geq 0 \ \forall k \geq 1 \ \wedge \ \sum_{k=1}^{N} k a_k + \sum_{k+l \leq N} a_k a_l = N^2 + O(1) \Rightarrow \sum_{k=1}^{N} a_k = N + O(1),$$

from which the PNT can be deduced with little effort. (Cf. [HT] for more in this direction.)

Another respect in which [Karl] is superior to most of the later papers, including V. Nevanlinna's [Nev] (whose approach is adopted by several books [Sch, Pol]), concerns the Tauberian deduction of the final result from a Selberg-style integral inequality. In [Karl], this is achieved by a theorem (Theorem 2.4 below, attributed to Erdős) with clearly identified, obviously minimal hypotheses and an elegant proof. This advantage over other approaches like [Nev], which tend to use further information about the discontinuities of the function under consideration, is essential for our generalization to arbitrary non-decreasing functions. However, we will have to adapt the proof (not least in order to work around an obscure issue).

In our exposition we make a point of avoiding the explicit summations over (pairs of) primes littering many elementary proofs, almost obtaining a proof of the PNT free of primes! This is achieved by defining the Möbius and von Mangoldt functions μ and Λ in terms of the functional identities they satisfy and using their explicit computation only to show that they are bounded and non-negative, respectively. Some of the proofs are formulated in terms of parametric Stieltjes integrals, typically of the form $\int f(x/t)dg(t)$ and integration by parts. We also do this in situations where f and g may both be discontinuous. Since our functions will always have bounded variation, thus at most countably many discontinuities, this can be justified by observing that the resulting identities hold for all x outside a countable set. Alternatively, we can replace $f(x)$ at every point of discontinuity by $(f(x+0)+f(x-0))/2$ without changing the asymptotics. For such functions, integration by parts always holds in the theory of Lebesgue-Stieltjes integration, cf. [Hew, HS].

The author hopes that the proof of Theorem 1.1 given below will help dispelling the prejudice that the elementary proofs of the PNT are (conceptionally and/or technically) difficult. Indeed he thinks that this is the most satisfactory of the elementary (and in fact of all) proofs of the PNT in that, besides not invoking complex analysis or Riemann's ζ-function, it minimizes number theoretic reasoning to a very well circumscribed minimum. One may certainly dispute that this is desirable, but we will argue elsewhere that it is.

The author is of course aware of the fact that the more direct elementary proofs of the PNT give better control of the remainder term. (Cf. the review [Dia] and the very recent paper [Kou], which provides a "a new and largely elementary proof of the best result known on the counting function of primes in arithmetic progressions".) It is not clear whether this is necessarily so.

Acknowledgments. The author would like to thank the referees for constructive comments that led to several improvements, in particular a better proof of Corollary 4.3.

2. First steps and strategy

Proposition 2.1. *Let* $f : [1, \infty) \to \mathbb{R}$ *be non-negative and non-decreasing and assume that* $F(x) = \sum_{n \le x} f(x/n)$ *satisfies* $F(x) = Ax \log x + Bx + o(x)$. *Then*

(i) $f(x) = O(x)$.

(ii) $\displaystyle\int_{1-0}^{x} \frac{df(t)}{t} = A \log x + O(1)$.

(iii) $\displaystyle\int_{1}^{x} \frac{f(t) - At}{t^2} dt = O(1)$.

Proof. (i) Following Ingham [Ing], we define f to be 0 on $[0, 1)$ and compute

$$\begin{aligned} f(x) - f\left(\frac{x}{2}\right) + f\left(\frac{x}{3}\right) - \cdots &= F(x) - 2F\left(\frac{x}{2}\right) \\ &= Ax \log x + Bx - 2\left(A\frac{x}{2} \log \frac{x}{2} + B\frac{x}{2}\right) + o(x) \\ &= Ax \log 2 + o(x). \end{aligned}$$

With positivity and monotonicity of f, this gives $f(x) - f(x/2) \le Kx$ for some $K > 0$. Adding these inequalities for $x, \frac{x}{2}, \frac{x}{4}, \ldots$, we find $f(x) \le 2Kx$. Together with $f \ge 0$, this gives (i).

(ii) We compute

$$F(x) = \sum_{n \le x} f\left(\frac{x}{n}\right) = \int_{1-0}^{x} f\left(\frac{x}{t}\right) d\lfloor t \rfloor$$

$$
\begin{aligned}
&= \left[\lfloor t \rfloor f\left(\frac{x}{t}\right) \right]_{t=1}^{t=x} - \int_{1-0}^{x} \lfloor t \rfloor \, df\left(\frac{x}{t}\right) \\
&= (\lfloor x \rfloor f(1) - f(x)) - \int_{1-0}^{x} t \, df\left(\frac{x}{t}\right) + \int_{1-0}^{x} (t - \lfloor t \rfloor) \, df\left(\frac{x}{t}\right) \\
&= O(x) + \int_{1-0}^{x} \frac{x}{u} df(u) + \int_{1-0}^{x} (t - \lfloor t \rfloor) \, df\left(\frac{x}{t}\right).
\end{aligned}
$$

In view of $0 \le t - \lfloor t \rfloor < 1$ and the weak monotonicity of f, the last integral is bounded by $|\int_1^x df(x/t)| = f(x) - f(1)$, which is $O(x)$ by (i). Using the hypothesis about F, we have

$$
Ax \log x + Bx + o(x) = O(x) + x \int_{1-0}^{x} \frac{df(t)}{t} + O(x),
$$

and division by x proves the claim.

(iii) Integrating by parts, we have

$$
\begin{aligned}
\int_1^x \frac{f(t) - At}{t^2} dt &= -\frac{f(x)}{x} + \int_{1-0}^{x} \frac{df(t)}{t} - \int_1^x \frac{A}{t} dt \\
&= O(1) + (A \log x + O(1)) - A \log x = O(1),
\end{aligned}
$$

where we used (i) and (ii). □

Remark 2.2. 1. The proposition can be proven under the weaker assumption $F(x) = Ax \log x + O(x)$, but we don't bother since later we will need the stronger hypothesis anyway.

2. Theorem 1.1, which we ultimately want to prove, implies a strong form of (iii): $\int_1^\infty \frac{f(t)-At}{t^2} dt = B - \gamma A$, cf. [Ing]. Conversely, existence of the improper integral already implies $f(x) \sim Ax$, cf. [Zag].
3. Putting $f = \psi$ and using (A.1), the above proofs of (i) and (ii) reduce to those of Chebychev and Mertens, respectively. □

The following two theorems will be proven in Sections 3 and 4, respectively:

Theorem 2.3. *Let f, F be as in Theorem 1.1. Then $g(x) = f(x) - Ax$ satisfies*

$$
\frac{|g(x)|}{x} \le \frac{1}{\log x} \int_1^x \frac{|g(t)|}{t^2} dt + o(1) \quad \text{as} \quad x \to \infty. \tag{2.1}
$$

Theorem 2.4. *For $g : [1, \infty) \to \mathbb{R}$, assume that there are $M, M' \ge 0$ such that*

$$
x \mapsto g(x) + Mx \quad \textit{is non-decreasing}, \tag{2.2}
$$

$$\left| \int_1^x \frac{g(t)}{t^2} dt \right| \leq M' \quad \forall x \geq 1. \tag{2.3}$$

Then

$$S := \limsup_{x\to\infty} \frac{|g(x)|}{x} < \infty, \tag{2.4}$$

and when $S > 0$ *we have*

$$\limsup_{x\to\infty} \frac{1}{\log x} \int_1^x \frac{|g(t)|}{t^2} dt < S. \tag{2.5}$$

Remark 2.5. 1. Note that (2.2) implies that g is Riemann integrable over finite intervals.

2. In our application, (2.4) already follows from Proposition 2.1 so that we do not need the corresponding part of the proof of Theorem 2.4. It will be proven nevertheless in order to give Theorem 2.4 an independent existence. □

Proof of Theorem 1.1 assuming Theorems 2.3 and 2.4. Since f is nondecreasing, it is clear that $g(x) = f(x) - Ax$ satisfies (2.2) with $M = A$, and (2.3) is implied by Proposition 2.1(iii). Now $S = \limsup |g(x)|/x$ is finite, by either Proposition 2.1(i) or the first conclusion of Theorem 2.4. Furthermore, $S > 0$ would imply (2.5). But combining this with the result (2.1) of Theorem 2.3, we would have the absurdity

$$S = \limsup_{x\to\infty} \frac{|g(x)|}{x} \leq \limsup_{x\to\infty} \frac{1}{\log x} \int_1^x \frac{|g(t)|}{t^2} dt < S.$$

Thus $S = 0$ holds, which is equivalent to $\frac{g(x)}{x} = \frac{f(x)-Ax}{x} \to 0$, as was to be proven. □

The next two sections are dedicated to the proofs of Theorems 2.3 and 2.4. The statements of both results are free of number theory, and this is also the case for the proof of the second. The proof of Theorem 2.3, however, uses a very modest amount of number theory, but nothing beyond Möbius inversion and the divisibility theory of $\mathbb{N}$ up to the fundamental theorem of arithmetic.

3. Proof of Theorem 2.3

3.1. Arithmetic. The aim of this subsection is to collect the basic arithmetic results that will be needed. We note that this is very little.

We begin by noting that $(\mathbb{N}, \cdot, 1)$ is an abelian monoid. Given $n, m \in \mathbb{N}$, we call m a divisor of n if there is an $r \in \mathbb{N}$ such that $mr = n$, in which case we write $m|n$. In view of the additive structure of the semiring $\mathbb{N}$, it is clear that the monoid $\mathbb{N}$ has cancellation ($ab = ac \Rightarrow b = c$), so the quotient r above is unique, and that the set of divisors of any n is finite.

Calling a function $f : \mathbb{N} \to \mathbb{R}$ an arithmetic function, the facts just stated allow us to define:

Definition 3.1. *If $f, g : \mathbb{N} \to \mathbb{R}$ are arithmetic functions, their Dirichlet convolution $f \star g$ denotes the function*

$$(f \star g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right) = \sum_{\substack{a,b \\ ab=n}} f(a) g(b).$$

It is easy to see that Dirichlet convolution is commutative and associative. It has a unit given by the function δ defined by $\delta(1) = 1$ and $\delta(n) = 0$ if $n \neq 1$.

By $\mathbf{1}$ we denote the constant function $\mathbf{1}(n) = 1$. Clearly, $(f \star \mathbf{1})(n) = \sum_{d|n} f(d)$.

Lemma 3.2. *There is a unique arithmetic function μ, called the Möbius function, such that $\mu \star \mathbf{1} = \delta$.*

Proof. μ must satisfy $\sum_{d|n} \mu(d) = \delta(n)$. Taking $n = 1$ we see that $\mu(1) = 1$. For $n > 1$ we have $\sum_{d|n} \mu(d) = 0$, which is equivalent to

$$\mu(n) = -\sum_{\substack{d|n \\ d<n}} \mu(d).$$

This uniquely determines $\mu(n) \in \mathbb{Z}$ inductively in terms of $\mu(m)$ with $m < n$. □

Proposition 3.3. (i) *μ is multiplicative, i.e. $\mu(nm) = \mu(n)\mu(m)$ whenever $(n, m) = 1$.*

(ii) *If p is a prime then $\mu(p) = -1$, and $\mu(p^k) = 0$ if $k \geq 2$.*

(iii) *$\mu(n) = O(1)$, i.e. μ is bounded.*

Proof. (i) Since $\mu(1) = 1$, $\mu(nm) = \mu(n)\mu(m)$ clearly holds if $n = 1$ or $m = 1$. Assume, by way of induction, that $\mu(uv) = \mu(u)\mu(v)$ holds whenever $(u, v) = 1$ and $uv < nm$, and let $n \neq 1 \neq m$ be relatively prime. Since every divisor of nm is of the form st with $s|n, t|m$, we have

$$\begin{aligned} 0 = \sum_{d|nm} \mu(d) &= \mu(nm) + \sum_{\substack{s|n, t|m \\ st<nm}} \mu(st) = \mu(nm) + \sum_{\substack{s|n, t|m \\ st<nm}} \mu(s)\mu(t) \\ &= \mu(nm) + \sum_{s|n} \mu(s) \sum_{t|m} \mu(t) - \mu(n)\mu(m) = \mu(nm) - \mu(n)\mu(m), \end{aligned}$$

which is the inductive step. (ii) For $k \geq 1$, we have $\mu(p^k) = -\sum_{i=0}^{k-1} \mu(p^i)$, inductively implying $\mu(p) = -1$ and $\mu(p^k) = 0$ if $k \geq 2$. Thus $\mu(p^k) \in \{0, -1\}$, which together with multiplicativity (i) gives $\mu(n) \in \{-1, 0, 1\}$ for all n, thus (iii). □

Proposition 3.4. (i) *The arithmetic function* $\Lambda := \log \star \mu$ *is the unique solution of* $\Lambda \star \mathbf{1} = \log$.

(ii) $\Lambda(n) = -\sum_{d|n} \mu(d) \log d$. *In particular,* $\Lambda(1) = 0$.

(iii) $\Lambda(n) = \log p$ *if* $n = p^k$ *where* p *is prime and* $k \geq 1$, *and* $\Lambda(n) = 0$ *otherwise.*

(iv) $\Lambda(n) \geq 0$.

Proof. (i) Existence: $\log \star \mu \star \mathbf{1} = \log \star \delta = \log$. Uniqueness: If $\Lambda_1 \star \mathbf{1} = \log = \Lambda_2 \star \mathbf{1}$ then $\Lambda_1 = \Lambda_1 \star \delta = \Lambda_1 \star \mathbf{1} \star \mu = \Lambda_2 \star \mathbf{1} \star \mu = \Lambda_2 \star \delta = \Lambda_2$.

(ii) $\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d} = \sum_{d|n} \mu(d)(\log n - \log d) = \log n \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \log d$. Now use $\sum_{d|n} \mu(d) = \delta(n)$. $\Lambda(1) = 0$ is obvious.

(iii) Using (ii), we have $\Lambda(p^k) = -\sum_{l=0}^{k} \mu(p^l) l \log p$, which together with Proposition 3.3(ii) implies $\Lambda(p^k) = \log p \ \forall k \geq 1$. If $n, m > 1$ and $(n, m) = 1$ then by the multiplicativity of μ,

$$\begin{aligned}\Lambda(nm) &= -\sum_{s|n} \sum_{t|m} \mu(st) \log(st) = -\sum_{s|n} \sum_{t|m} \mu(s)\mu(t)(\log s + \log t) \\ &= \sum_{s|n} \mu(s) \log s \sum_{t|m} \mu(t) + \sum_{t|m} \mu(t) \log t \sum_{s|n} \mu(s) = 0.\end{aligned}$$

(iv) Obvious consequence of (iii). □

Remark 3.5. The only properties of μ and Λ that will be used in the proof of Theorem 1.1 are the defining ones ($\mu \star \mathbf{1} = \delta$, $\Lambda \star \mathbf{1} = \log$), the trivial consequence (ii) in the above proposition, and the boundedness of μ and the non-negativity of Λ.

In particular, the explicit computations of $\mu(n)$ and $\Lambda(n)$ in terms of the prime factorization of n were only needed to prove the latter two properties. (Of course, these properties of μ and Λ would be obvious if one defined them by the explicit formulae proven above, but this would be ad hoc and ugly, and one would still need to use the fundamental theorem of arithmetic for proving that $\mu \star \mathbf{1} = \delta$ and $\Lambda \star \mathbf{1} = \log$.)

Note that prime numbers will play no rôle whatsoever before we turn to the actual proof of the prime number theorem in the Appendix, where the computation of $\Lambda(n)$ will be used again. □

3.2. The (weighted) Möbius transform.

Definition 3.6. *Given a function* $f : [1, \infty) \to \mathbb{R}$*, its 'Möbius transform' is defined by*

$$F(x) = \sum_{n \le x} f\left(\frac{x}{n}\right).$$

Lemma 3.7. *The Möbius transform* $f \mapsto F$ *is invertible, the inverse Möbius transform being given by*

$$f(x) = \sum_{n \le x} \mu(n) F\left(\frac{x}{n}\right).$$

Proof. We compute

$$\begin{aligned}\sum_{n \le x} \mu(n) F\left(\frac{x}{n}\right) &= \sum_{n \le x} \mu(n) \sum_{m \le x/n} f\left(\frac{x}{nm}\right) = \sum_{nm \le x} \mu(n) f\left(\frac{x}{nm}\right) \\ &= \sum_{r \le x} f\left(\frac{x}{r}\right) \sum_{s|r} \mu(s) = \sum_{r \le x} f\left(\frac{x}{r}\right) \delta(r) = f(x),\end{aligned}$$

where we used the defining property $\sum_{d|n} \mu(d) = \delta(n)$ of μ. □

Remark 3.8. Since the point of Theorem 1.1 is to deduce information about f from information concerning its Möbius transform F, it is tempting to appeal to Lemma 3.7 directly. However, in order for this to succeed, we would need control over $M(x) = \sum_{n \le x} \mu(n)$, at least as good as $M(x) = o(x)$. But then one is back in Ellison's approach mentioned in the introduction. The essential idea of the Selberg–Erdős approach to the PNT, not entirely transparent in the early papers but clarified soon after [TI], is to consider weighted Möbius inversion formulae as follows. □

Lemma 3.9. *Let* $f : [1, \infty) \to \mathbb{R}$ *be arbitrary and* $F(x) = \sum_{n \le x} f(x/n)$*. Then*

$$f(x) \log x + \sum_{n \le x} \Lambda(n) f\left(\frac{x}{n}\right) = \sum_{n \le x} \mu(n) \log \frac{x}{n} F\left(\frac{x}{n}\right). \tag{3.1}$$

Proof. We compute

$$\sum_{n \le x} \mu(n) \log \frac{x}{n} F\left(\frac{x}{n}\right) = \log x \sum_{n \le x} \mu(n) F\left(\frac{x}{n}\right) - \sum_{n \le x} \mu(n) \log n \, F\left(\frac{x}{n}\right).$$

By Lemma 3.7, the first term equals $f(x) \log x$, whereas for the second we have

$$\sum_{n\le x}\mu(n)\log n\, F\left(\frac{x}{n}\right)=\sum_{n\le x}\mu(n)\log n\sum_{m\le x/n} f\left(\frac{x}{nm}\right)=\sum_{nm\le x}\mu(n)\log n\, f\left(\frac{x}{nm}\right)$$
$$=\sum_{s\le x}\left(\sum_{n|s}\mu(n)\log n\right) f\left(\frac{x}{s}\right)=-\sum_{s\le x}\Lambda(s)\, f\left(\frac{x}{s}\right),$$

the last equality being Proposition 3.4(ii). Putting everything together, we obtain (3.1). □

Remark 3.10. 1. Eq. (3.1) is known as the 'Tatuzawa-Iseki formula', cf. [TI, (8)] (and [Kar1, p. 24]).

2. Without the factor $\log(x/n)$ on the right hand side, (3.1) reduces to Möbius inversion. Thus (3.1) is a sort of weighted Möbius inversion formula. The presence of the sum involving $f(x/n)$ is very much wanted, since it will allow us to obtain the integral inequality (2.1) involving all $f(t), t\in[1,x]$. In order to do so, we must get rid of the explicit appearance of the function $\Lambda(n)$, which is very irregular and about which we know little. This requires some preparation. □

Lemma 3.11. *For any arithmetic function $f:\mathbb{N}\to\mathbb{R}$ we have*

$$f(n)\log n+\sum_{d|n}\Lambda(d)f\left(\frac{n}{d}\right)=\sum_{d|n}\mu(d)\,\log\frac{n}{d}\sum_{m|(n/d)} f(m).$$

In particular, we have Selberg's identity:

$$\Lambda(n)\log n+\sum_{d|n}\Lambda(d)\Lambda\left(\frac{n}{d}\right)=\sum_{d|n}\mu(d)\log^2\frac{n}{d}. \tag{3.2}$$

Proof. If f is an arithmetic function, i.e. defined only on $\mathbb{N}$, we extend it to $\mathbb{R}$ as being 0 on $\mathbb{R}\backslash\mathbb{N}$. With this extension,

$$F(n)=\sum_{m\le n} f\left(\frac{n}{m}\right)=\sum_{m|n} f\left(\frac{n}{m}\right)=\sum_{m|n} f(m),$$

so that (3.1) becomes the claimed identity. Taking $f(n)=\Lambda(n)$ and using $\sum_{d|n}\Lambda(d)=\log n$, Selberg's formula follows. □

3.3. Preliminary estimates.

Lemma 3.12. *The following elementary estimates hold as* $x \to \infty$*:*

$$\sum_{n\le x} \frac{1}{n} = \log x + \gamma + O\left(\frac{1}{x}\right), \tag{3.3}$$

$$\sum_{n\le x} \frac{\log n}{n} = \frac{\log^2 x}{2} + c + O\left(\frac{1+\log x}{x}\right), \tag{3.4}$$

$$\sum_{n\le x} \log n = x\log x - x + O(\log x), \tag{3.5}$$

$$\sum_{n\le x} \log \frac{x}{n} = x + O(\log x), \tag{3.6}$$

$$\sum_{n\le x} \log^2 n = x(\log^2 x - 2\log x + 1) + O(\log^2 x), \tag{3.7}$$

$$\sum_{n\le x} \log^2 \frac{x}{n} = x + O(\log^2 x). \tag{3.8}$$

Here, γ *is Euler's constant and* $c > 0$.

Proof. (3.3): We have

$$\sum_{n=1}^{N} \frac{1}{n} - \int_1^N \frac{dt}{t} = \int_{1-0}^{N} \frac{d(\lfloor t\rfloor - t)}{t} = \left[\frac{\lfloor t\rfloor - t}{t}\right]_1^N + \int_1^N \frac{t-\lfloor t\rfloor}{t^2}dt.$$

Since $0 \le t - \lfloor t\rfloor < 1$, the integral on the r.h.s. converges as $N \to \infty$ to some number γ (Euler's constant) strictly between 0 and $1 = \int_1^\infty dt/t^2$. Thus

$$\sum_{n=1}^{N} \frac{1}{n} = \int_1^N \frac{dt}{t} + \gamma - \int_N^\infty \frac{t-\lfloor t\rfloor}{t^2}dt = \log N + \gamma + O\left(\frac{1}{N}\right).$$

(3.4): Similarly to the proof of (3.3), we have

$$\begin{aligned}\sum_{n=1}^{N} \frac{\log n}{n} - \int_1^N \frac{\log t}{t}dt &= \int_{1-0}^{N} \frac{\log t}{t}\, d(\lfloor t\rfloor - t)\\ &= \left[\frac{(\lfloor t\rfloor - t)\log t}{t}\right]_1^N + \int_1^N \frac{(t-\lfloor t\rfloor)\log t}{t^2}dt.\end{aligned}$$

The final integral converges to some $c > 0$ as $N \to \infty$ since $(\log t)/t^2 = O(t^{-2+\varepsilon})$. Using

$$\int_1^x \frac{\log t}{t}dt = \frac{\log^2 x}{2}, \qquad \int_N^\infty \frac{\log t}{t^2}dt = -\left[\frac{\log t}{t}\right]_N^\infty + \int_N^\infty \frac{dt}{t^2} = \frac{1+\log N}{N}$$

we have

$$\sum_{n=1}^{N}\frac{\log n}{n}=\int_1^x\frac{\log t}{t}dt+c-\int_N^\infty\frac{(t-\lfloor t\rfloor)\log t}{t^2}dt=\frac{\log^2 x}{2}+c+O\left(\frac{1+\log x}{x}\right).$$

(3.5): By monotonicity, we have

$$\int_1^x\log t\,dt\le\sum_{n\le x}\log n\le\int_1^{x+1}\log t\,dt.$$

Combining this with $\int_1^x\log t\,dt=x\log x-x+1$, (3.5) follows.

(3.6): Using (3.5), we have

$$\begin{aligned}\sum_{n\le x}\log\frac{x}{n}&=\lfloor x\rfloor\log x-\sum_{n\le x}\log n=\big(x+O(1)\big)\log x-\big(x\log x-x+O(\log x)\big)\\&=x+O(\log x).\end{aligned}$$

(3.7): By monotonicity,

$$\int_1^x\log^2 t\,dt\le\sum_{n\le x}\log^2 n\le\int_1^{x+1}\log^2 t\,dt.$$

Now,

$$\int_1^x\log^2 t\,dt=\int_0^{\log x}e^u u^2du=[e^u(u^2-2u+1)]_0^{\log x}=x(\log^2 x-2\log x+1)-1.$$

Combining these two facts, (3.7) follows.

(3.8): Using (3.5) and (3.7), we compute

$$\begin{aligned}\sum_{n\le x}\log^2\frac{x}{n}&=\sum_{n\le x}(\log x-\log n)^2\\&=\lfloor x\rfloor\log^2 x-2\log x\big(x\log x-x+O(\log x)\big)\\&\qquad+x(\log^2 x-2\log x+1)+O(\log^2 x)\\&=x+O(\log^2 x)\,.\end{aligned}$$

■

Proposition 3.13. *The following estimates involving the Möbius function hold:*

$$\sum_{n\le x}\frac{\mu(n)}{n}=O(1),\tag{3.9}$$

$$\sum_{n\le x}\frac{\mu(n)}{n}\log\frac{x}{n}=O(1),\tag{3.10}$$

$$\sum_{n\le x}\frac{\mu(n)}{n}\log^2\frac{x}{n}=2\log x+O(1).\tag{3.11}$$

Proof. (3.9): If $f(x) = 1$ then $F(x) = \lfloor x \rfloor$. Möbius inversion (Lemma 3.7) gives

$$1 = \sum_{n\le x} \mu(n) \left\lfloor \frac{x}{n} \right\rfloor = \sum_{n\le x} \mu(n) \left(\frac{x}{n} + O(1) \right) = x \sum_{n\le x} \frac{\mu(n)}{n} + \sum_{n\le x} O(1), \tag{3.12}$$

where we used $\mu(n) = O(1)$ (Proposition 3.3(iii)). In view of $\sum_{n\le x} O(1) = O(x)$, we have $\sum_{n\le x} \mu(n)/n = O(x)/x = O(1)$.

(3.10): If $f(x) = x$ then $F(x) = \sum_{n\le x} x/n = x \log x + \gamma x + O(1)$ by (3.3). By Möbius inversion,

$$x = \sum_{n\le x} \mu(n) \left(\frac{x}{n} \log \frac{x}{n} + \gamma \frac{x}{n} + O(1) \right) = x \sum_{n\le x} \frac{\mu(n)}{n} \log \frac{x}{n} + xO(1) + O(x),$$

where we used (3.9) and Proposition 3.3(iii). From this we easily read off (3.10).

(3.11): If $f(x) = x \log x$ then

$$\begin{aligned} F(x) &= \sum_{n\le x} \frac{x}{n} \log \frac{x}{n} = \sum_{n\le x} \frac{x}{n} (\log x - \log n) = x \log x \sum_{n\le x} \frac{1}{n} - x \sum_{n\le x} \frac{\log n}{n} \\ &= x \log x \left(\log x + \gamma + O\left(\frac{1}{x} \right) \right) - x \left(\frac{\log^2 x}{2} + c + O\left(\frac{1 + \log x}{x} \right) \right) \\ &= \frac{1}{2} x \log^2 x + \gamma x \log x - cx + O(1 + \log x), \end{aligned}$$

by (3.3) and (3.4). Now Möbius inversion gives

$$\begin{aligned} x \log x &= \sum_{n\le x} \mu(n) \left(\frac{x}{2n} \log^2 \frac{x}{n} + \gamma \frac{x}{n} \log \frac{x}{n} - c \frac{x}{n} + O(1 + \log \frac{x}{n}) \right) \\ &= \frac{x}{2} \sum_{n\le x} \frac{\mu(n)}{n} \log^2 \frac{x}{n} + xO(1) + xO(1) + O(x), \end{aligned}$$

where we used (3.9), (3.10) and (3.6), and division by $x/2$ gives (3.11). □

Proposition 3.14 (Selberg, Erdős–Karamata [EK]). *Defining*

$$K(1) = 0, \qquad K(n) = \frac{1}{\log n} \sum_{d\,|\,n} \Lambda(d) \Lambda\left(\frac{n}{d} \right) \quad \text{if} \quad n \ge 2,$$

we have $K(n) \ge 0$ *and*

$$\sum_{n\le x} \big(\Lambda(n) + K(n) \big) = 2x + O\left(\frac{x}{\log x} \right). \tag{3.13}$$

Proof. The first claim is obvious in view of Proposition 3.4(iv). We estimate

$$\begin{aligned} U(x) &:= \sum_{n\le x}\sum_{d\mid n}\mu(d)\log^2\frac{n}{d} = \sum_{n\le x}\mu(n)\sum_{m\le x/n}\log^2 m \\ &= \sum_{n\le x}\mu(n)\left(\frac{x}{n}\Big(\log^2\frac{x}{n} - 2\log\frac{x}{n} + 1\Big) + O\Big(\log^2\frac{x}{n}\Big)\right) \\ &= x\big(2\log x + O(1)\big) - 2xO(1) + xO(1) + O(x) = 2x\log x + O(x). \end{aligned}$$

Here we used (3.7), (3.11), (3.10), (3.9), the fact $\mu(n) = O(1)$, and (3.8). Comparing (3.13) and (3.2), we have

$$\begin{aligned} \sum_{n\le x}\Lambda(n) + K(n) &= \sum_{2\le n\le x}\frac{1}{\log n}\sum_{d\mid n}\mu(d)\log^2\frac{n}{d} \\ &= \int_{2-0}^{x}\frac{dU(t)}{\log t} = \left[\frac{U(t)}{\log t}\right]_2^x + \int_2^x\frac{U(t)}{t\log^2 t}dt \\ &= 2x + O\left(\frac{x}{\log x}\right) + \int_2^x\frac{dt}{\log t} + O\left(\int_2^x\frac{dt}{\log^2 t}\right). \end{aligned}$$

In view of the estimate

$$\int_2^x\frac{dt}{\log t} = \int_2^{\sqrt{x}}\frac{dt}{\log t} + \int_{\sqrt{x}}^x\frac{dt}{\log t} \le \frac{\sqrt{x}}{\log 2} + \frac{x}{\log\sqrt{x}} = O\left(\frac{x}{\log x}\right),$$

we are done. □

Remark 3.15. In view of (3.2), the above estimate $U(x) = 2x\log x + O(x)$ is equivalent to

$$\sum_{n\le x}\Lambda(n)\log n + \sum_{ab\le x}\Lambda(a)\Lambda(b) = 2x\log x + O(x),$$

which is used in most Selberg-style proofs. (It would lead to (3.20) with $k = 2$.) □

3.4. Conclusion.

Proposition 3.16. *If $g : [1,\infty) \to \mathbb{R}$ is such that*

$$G(x) = \sum_{n\le x} g\left(\frac{x}{n}\right) = Bx + C\frac{x}{\log x} + o\left(\frac{x}{\log x}\right) \tag{3.14}$$

then

$$g(x)\log x + \sum_{n\le x}\Lambda(n)\,g\left(\frac{x}{n}\right) = o(x\log x). \tag{3.15}$$

Proof. In view of Lemma 3.9, all we have to do is estimate

$$\sum_{n\le x}\mu(n)\log\frac{x}{n}\left(B\frac{x}{n}+C\frac{x}{n\log\frac{x}{n}}+o\left(\frac{x}{n\log\frac{x}{n}}\right)\right)=S_1+S_2+S_3.$$

The three terms are

$$S_1=Bx\sum_{n\le x}\frac{\mu(n)}{n}\log\frac{x}{n}=xO(1)=O(x),$$

$$S_2=Cx\sum_{n\le x}\frac{\mu(n)}{n}=xO(1)=O(x),$$

$$S_3=\sum_{n\le x}\mu(n)o\left(\frac{x}{n}\right)=\sum_{n\le x}o\left(\frac{x}{n}\right)=o\Big(x\sum_{n\le x}\frac{1}{n}\Big)=o(x\log x),$$

where we used (3.10), (3.9), and $\mu(n)=O(1)$, respectively. □

Proof of Theorem 2.3. In view of $g(x)=f(x)-Ax$ and Proposition 2.1 (i), (ii), we immediately have

$$g(x)=O(x),\qquad \int_1^x\frac{dg(u)}{u}=O(1). \tag{3.16}$$

Furthermore, since f satisfies (1.1), and (3.3) gives $\sum_{n\le x}Ax/n=Ax\log x+A\gamma x+O(1)$, the Möbius transform G of $g(x)=f(x)-Ax$ satisfies (3.14) (with a different B), so that Proposition 3.16 applies and (3.15) holds.

Writing $N(x)=\sum_{n\le x}\Lambda(n)+K(n)$, by Proposition 3.14 we have $N(x)=2x+\omega(x)$ with $\omega(x)=o(x)$. Now,

$$\begin{aligned}\sum_{n\le x}\big(\Lambda(n)+K(n)\big)g\left(\frac{x}{n}\right)&=\int_{1-0}^{x}g\left(\frac{x}{t}\right)dN(t)\\&=\Big[N(t)g\left(\frac{x}{t}\right)\Big]_1^x-\int_1^x N(t)dg\left(\frac{x}{t}\right)\\&=\big(N(x)g(1)-N(1)g(x)\big)+\int_1^x N\left(\frac{x}{u}\right)dg(u)\\&=O(x)+\int_1^x\left(\frac{2x}{u}+o\left(\frac{2x}{u}\right)\right)dg(u)\\&=O(x)+2x\int_1^x\frac{dg(u)}{u}+o\left(x\int_1^x\frac{dg(u)}{u}\right)\\&=O(x)+O(x)+o(x)=O(x),\end{aligned} \tag{3.17}$$

where we used (3.16). On the other hand,

$$\begin{aligned}\sum_{n\le x}(\Lambda(n)+K(n))\left|g\left(\frac{x}{n}\right)\right| &= \int_{1-0}^{x}\left|g\left(\frac{x}{t}\right)\right| dN(t)\\ &= 2\int_1^x\left|g\left(\frac{x}{t}\right)\right| dt + \int_{1-0}^{x}\left|g\left(\frac{x}{t}\right)\right| d\omega(t)\\ &= 2x\int_1^x\frac{|g(t)|}{t^2}dt - \int_{1-0}^{x}\omega(t)\,d\left|g\left(\frac{x}{t}\right)\right|\\ &\qquad + \left[g\left(\frac{x}{t}\right)\omega(t)\right]_{t=1-0}^{t=x}\\ &= 2x\int_1^x\frac{|g(t)|}{t^2}dt + \int_1^{x+0}\omega\left(\frac{x}{t}\right)d|g(t)|\\ &\qquad + g(1)\omega(x) - g(x+0)\omega(1-0).\end{aligned}\tag{3.18}$$

In view of $g(x)=O(x)$ and $\omega(x)=o(x)$, the sum of the last two terms is $O(x)$. Furthermore,

$$\begin{aligned}\int_1^{x+0}\omega\left(\frac{x}{t}\right)d|g(t)| &= o\left(x\int_1^{x+0}\frac{|d|g(t)||}{t}\right)\le o\left(x\int_1^{x+0}\frac{|dg(t)|}{t}\right)\\ &\le o\left(x\int_1^{x+0}\frac{df+Adt}{t}\right) = o(x\log x),\end{aligned}$$

where we used $g(x)=f(x)-Ax$ and $df=|df|$ (since f is non-decreasing) to obtain $|dg|=|df-Adt|\le|df|+Adt=df+Adt$ and Proposition 2.1(ii). Introducing this into (3.18), we have

$$\sum_{n\le x}(\Lambda(n)+K(n))\left|g\left(\frac{x}{n}\right)\right| = 2x\int_1^x\frac{|g(t)|}{t^2}dt + o(x\log x).\tag{3.19}$$

After these preparations, we can conclude quickly: Subtracting (3.17) from (3.15) we obtain

$$g(x)\log x = \sum_{n\le x}K(n)g\left(\frac{x}{n}\right) + o(x\log x).$$

Taking absolute values of this and of (3.15) while observing that Λ and K are non-negative, we have the inequalities

$$\begin{aligned}|g(x)|\log x &\le \sum_{n\le x}\Lambda(n)\left|g\left(\frac{x}{n}\right)\right| + o(x\log x),\\ |g(x)|\log x &\le \sum_{n\le x}K(n)\left|g\left(\frac{x}{n}\right)\right| + o(x\log x).\end{aligned}$$

Adding these inequalities and comparing with (3.19) we have

$$\begin{aligned}2|g(x)|\log x &\le \sum_{n\le x}(\Lambda(n)+K(n))\left|g\left(\frac{x}{n}\right)\right| + o(x\log x)\\ &= 2x\int_1^x\frac{|g(t)|}{t^2}dt + o(x\log x),\end{aligned}$$

so that (2.1), and with it Theorem 2.3, is obtained on dividing by $2x\log x$. □

Remark 3.17. 1. We did not use the full strength of Proposition 3.14, but only an $o(x)$ remainder.

2. Inequality (2.1) is the special case $k = 1$ of the more general integral inequality

$$\frac{|g(x)|}{x}\log^k x \le k\int_1^x \frac{|g(t)|\log^{k-1} t}{t^2}dt + O(\log^{k-c} x) \quad \forall k \in \mathbb{N} \tag{3.20}$$

proven in [Ba], assuming a $O\left(\frac{x}{\log^2 x}\right)$ in (1.1) instead of $C\frac{x}{\log x} + o\left(\frac{x}{\log x}\right)$. □

4. Proof of Theorem 2.4

The proof will be based on the following proposition, to be proven later:

Proposition 4.1. *If $s : [0,\infty) \to \mathbb{R}$ satisfies*

$$e^{t'}s(t') - e^t s(t) \ge -M(e^{t'} - e^t) \quad \forall t' \ge t \ge 0, \tag{4.1}$$

$$\left|\int_0^x s(t)dt\right| \le M' \quad \forall x \ge 0, \tag{4.2}$$

and $S = \limsup|s(x)| > 0$ then there exist numbers $0 < S_1 < S$ and $e, h > 0$ such that

$$\mu(E_{x,h,S_1}) \ge e \quad \forall x \ge 0, \quad \text{where} \quad E_{x,h,S_1} = \{t \in [x, x+h] \mid |s(t)| \le S_1\}, \tag{4.3}$$

and μ denotes the Lebesgue measure.

Proof of Theorem 2.4 assuming Proposition 4.1. It is convenient to replace $g : [1,\infty) \to \mathbb{R}$ by $s : [0,\infty) \to \mathbb{R}$, $s(t) = e^{-t}g(e^t)$. Now s is locally integrable, and the assumptions (2.2) and (2.3) become (4.1) and (4.2), respectively, whereas the conclusions (2.4) and (2.5) assume the form

$$S = \limsup_{t\to\infty}|s(t)| < \infty, \tag{4.4}$$

$$S > 0 \quad \Rightarrow \quad \limsup_{x\to\infty}\frac{1}{x}\int_0^x |s(t)|dt < S. \tag{4.5}$$

The proof of (4.4) is easy: Dividing (4.1) by $e^{t'}$ and integrating over $t' \in [t, t+h]$, where $h > 0$, one obtains

$$\int_t^{t+h} s(t')dt' - s(t)(1-e^{-h}) \geq -Mh + M(1-e^{-h}),$$

and using $|\int_a^b s(t)dt| \leq |\int_0^a s(t)dt| + |\int_0^b s(t)dt| \leq 2M'$ by (4.2), we have the upper bound

$$s(t) \leq \frac{2M' + M(e^{-h} - 1 + h)}{1 - e^{-h}}.$$

Similarly, dividing (4.1) by e^t and integrating over $t \in [t'-h, t']$, one obtains the lower bound

$$-\frac{2M' + M(e^h - 1 - h)}{e^h - 1} \leq s(t),$$

thus (4.4) holds.

Assuming $S > 0$, let S_1, h, e be as required by Proposition 4.1. For each $\widehat{S} > S$ there is x_0 such that $x \geq x_0 \Rightarrow |s(x)| \leq \widehat{S}$. Given $x \geq x_0$ and putting $N = \lfloor \frac{x-x_0}{h} \rfloor$, we have

$$\begin{aligned}
\int_0^x |s(t)|dt &= \int_0^{x_0} |s(t)|dt + \sum_{n=1}^{N} \int_{x_0+(n-1)h}^{x_0+nh} |s(t)|dt + \int_{x_0+Nh}^{x} |s(t)|dt \\
&\leq 2M' + N[\widehat{S}(h-e) + S_1 e] + 2M' \\
&= \left(\frac{x-x_0}{h} + O(1)\right) h \left[\left(1 - \frac{e}{h}\right)\widehat{S} + \frac{e}{h} S_1\right] + 4M' \\
&= x\left[\left(1 - \frac{e}{h}\right)\widehat{S} + \frac{e}{h} S_1\right] + O(1).
\end{aligned}$$

Thus

$$\limsup_{x\to\infty} \frac{1}{x} \int_0^x |s(t)|dt \leq \left(1 - \frac{e}{h}\right)\widehat{S} + \frac{e}{h} S_1.$$

Since $S_1 < S$ and since $\widehat{S} > S$ can be chosen arbitrarily close to S, (4.5) holds and thus Theorem 2.4. □

In order to make plain how the assumptions (4.1) and (4.2) are used to prove Proposition 4.1, we prove two intermediate results that each use only one of the assumptions. For the first we need a "geometrically obvious" lemma of isoperimetric character:

Lemma 4.2. *Let $t_1 < t_2$, $C_1 > C_2 > 0$ and $k : [t_1, t_2] \to \mathbb{R}$ be non-decreasing with $k(t_1) \geq C_1 e^{t_1}$ and $k(t_2) \leq C_2 e^{t_2}$. Then*

$$\mu\left(\{t \in [t_1, t_2] \mid C_2 e^t \leq k(t) \leq C_1 e^t\}\right) \geq \log \frac{C_1}{C_2}.$$

Proof. As a non-decreasing function, k has left and right limits $k(t \pm 0)$ everywhere and $k(t-0) \le k(t) \le k(t+0)$. The assumptions imply $t_1 \in A := \{t \in [t_1, t_2] \mid k(t) \ge C_1 e^t\}$, thus we can define $T_1 = \sup(A)$. Quite obviously we have $t > T_1 \Rightarrow k(t) < C_1 e^t$, which together with the non-decreasing property of k and the continuity of the exponential function implies $k(T_1 + 0) \le C_1 e^{T_1}$ (provided $T_1 < t_2$). We have $T_1 \in A$ if and only if $k(T_1) \ge C_1 e^{T_1}$. If $T_1 \notin A$ then $T_1 > t_1$, and every interval $(T_1 - \varepsilon, T_1)$ (with $0 < \varepsilon < T_1 - t_1$) contains points t such that $k(t) \ge C_1 e^t$. This implies $k(T_1 - 0) \ge C_1 e^{T_1}$. Now assume $T_1 = t_2$. If $T_1 \in A$ then $C_1 e^{T_1} \le k(T_1) \le C_2 e^{T_1}$. If $T_1 \notin A$ then $C_1 e^{T_1} \le k(T_1 - 0) \le k(T_1) \le C_2 e^{t_2}$. In both cases we arrive at a contradiction since $C_2 < C_1$. Thus $T_1 < t_2$. If $T_1 \in A$ (in particular if $T_1 = t_1$) then $C_1 e^{T_1} \le k(T_1) \le k(T_1 + 0) \le C_1 e^{T_1}$. Thus k is continuous from the right at T_1 and $k(T_1) = C_1 e^{T_1}$. If $T_1 \notin A$ then $T_1 > t_1$ and $C_1 e^{T_1} \le k(T_1 - 0) \le k(T_1 + 0) \le C_1 e^{T_1}$. This implies $k(T_1) = C_1 e^{T_1}$, thus the contradiction $T_1 \in A$. Thus we always have $T_1 \in A$, thus $k(T_1) = C_1 e^{T_1}$.

Now let $B = \{t \in [T_1, t_2] \mid k(t) \le C_2 e^t\}$. We have $t_2 \in B$, thus $T_2 = \inf(B)$ is defined and $T_2 \ge T_1$. Arguing similarly as before we have $t < T_2 \Rightarrow k(t) > C_2 e^t$, implying $k(T_2 - 0) \ge C_2 e^{T_2}$. And if $T_2 < t_2$ and $T_2 \notin B$ then $k(T_2 + 0) \le C_2 e^{T_2}$. If $T_2 \in B$ (in particular if $T_2 = t_2$) then $C_2 e^{T_2} \le k(T_2 - 0) \le k(T_2) \le C_2 e^{T_2}$, implying $k(T_2 - 0) = k(T_2) = C_2 e^{T_2}$ so that k is continuous from the left at T_2. If $T_2 \notin B$ then $T_2 < t_2$ and $C_2 e^{T_2} \le k(T_2 - 0) \le k(T_2 + 0) \le C_2 e^{T_2}$, implying $k(T_2) = C_2 e^{T_2}$ and thus a contradiction. Thus we always have $T_2 \in B$, thus $k(T_2) = C_2 e^{T_2}$.

By the above results, we have $C_2 e^t \le k(t) \le C_1 e^t \ \forall t \in [T_1, T_2]$ and thus

$$\mu\left(\{t \in [t_1, t_2] \mid C_2 e^t \le k(t) \le C_1 e^t\}\right) \ge T_2 - T_1. \tag{4.6}$$

Using once more that k is non-decreasing, we have

$$C_1 e^{T_1} = k(T_1) \le k(T_2) = C_2 e^{T_2},$$

implying $T_2 - T_1 \ge \log \frac{C_1}{C_2}$, and combining this with (4.6) proves the claim. □

Corollary 4.3. *Assume that* $s : [0, \infty) \to \mathbb{R}$ *satisfies* (4.1) *and* $s(t_1) \ge S_1 \ge S_2 \ge s(t_2)$, *where* $S_2 + M > 0$. *Then*

$$\mu(\{t \in [t_1, t_2] \mid s(t) \in [S_2, S_1]\}) \ge \log \frac{S_1 + M}{S_2 + M}.$$

Proof. We note that (4.1) is equivalent to the statement that the function $k : t \mapsto e^t(s(t) + M)$ is non-decreasing. The assumption $s(t_1) \ge S_1 \ge S_2 \ge s(t_2)$ implies $k(t_1) \ge (S_1 + M)e^{t_1}$ and $k(t_2) \le (S_2 + M)e^{t_2}$. Now the claim follows directly by an application of the preceding lemma. □

Lemma 4.4. *Let $s : [0,\infty) \to \mathbb{R}$ be integrable over bounded intervals, satisfying (4.2). Let $e > 0$ and $0 < S_2 < S_1$ be arbitrary, and assume*

$$h \geq 2\left(e + \frac{M'}{S_1} + \frac{M'}{S_2}\right). \tag{4.7}$$

Then every interval $[x, x+h]$ satisfies at least one of the following conditions:

(i) *$\mu(E_{x,h,S_1}) \geq e$, where E_{x,h,S_1} is as in (4.3),*

(ii) *there exist t_1, t_2 such that $x \leq t_1 < t_2 \leq x + h$ and $s(t_1) \geq S_1$ and $s(t_2) \leq S_2$.*

Proof. It is enough to show that falsity of (i) implies (ii). Define

$$T = \sup\{t \in [x, x+h] \mid s(t) \leq S_2\},$$

with the understanding that $T = x$ if $s(t) > S_2$ for all $t \in [x, x+h]$. Then $s(t) > S_2$ $\forall t \in (T, x+h]$, which implies

$$(x + h - T)S_2 \leq \int_T^{x+h} s(t)dt \leq 2M'$$

and therefore

$$x + h - T \leq \frac{2M'}{S_2}. \tag{4.8}$$

We observe that (4.8) with $T = x$ would contradict (4.7). Thus $x < T \leq x + h$, so we can indeed find a $t_2 \in [x, x+h]$ with $s(t_2) \leq S_2$. Since we do not assume continuity of s, we cannot claim that we may take $t_2 = T$, but by definition a t_2 can be found in $(T - \varepsilon, T]$ for every $\varepsilon > 0$.

Now we claim that there is a point $t_1 \in [x, t_2]$ such that $s(t_1) \geq S_1$. Otherwise, we would have $s(t) < S_1$ for all $t \in [x, t_2]$. By definition, $|s(t)| \leq S_1$ for $t \in E_{x,h,S_1}$, thus $|s| > S_1$ on the complement of E_{x,h,S_1}. Combined with $s(t) < S_1$ for $t \in [x, t_2]$, this means $s(t) < -S_1$ whenever $t \in [x, t_2] \setminus E_{x,h,S_1}$. Thus

$$\begin{aligned}\int_x^{t_2} s(t)dt &\leq S_1\mu([x, t_2] \cap E_{x,h,S_1}) - S_1\mu([x, t_2] \setminus E_{x,h,S_1}) \\ &= -S_1(t_2 - x) + 2S_1\mu([x, t_2] \cap E_{x,h,S_1}) \\ &= S_1(x - t_2 + 2\mu([x, t_2] \cap E_{x,h,S_1}))\,.\end{aligned}$$

In view of (4.8) and $t_2 > T - \varepsilon$ (with $\varepsilon > 0$ arbitrary), we have $x - t_2 < x - T + \varepsilon \leq 2M'/S_2 - h + \varepsilon$, thus we continue the preceding inequality as

$$\cdots < S_1\left(\frac{2M'}{S_2} - h + \varepsilon + 2\mu([x, t_2] \cap E_{x,h,S_1})\right).$$

By our assumption that (i) is false, we have $\mu([x,t_2]\cap E_{x,h,S_1}) \le \mu(E_{x,h,S_1}) < e$. Thus choosing ε such that $0<\varepsilon<2(e-\mu([x,t_2]\cap E_{x,h,S_1}))$, we have

$$\cdots < S_1\left(\frac{2M'}{S_2} - h + 2e\right).$$

Combining this with (4.7), we finally obtain $\int_x^{t_2} s(t)dt < -2M'$, which contradicts the assumption (4.2). Thus there is a point $t_1 \in [x,t_2]$ such that $s(t_1) \ge S_1$. In view of $s(t_1) \ge S_1 > S_2 \ge s(t_2)$, we have $t_1 \ne t_2$, thus $t_1 < t_2$. □

Proof of Proposition 4.1. Assuming that $S = \limsup |s(x)| > 0$, choose S_1, S_2 such that $0 < S_2 < S_1 < S$. Then $e := \log \frac{S_1+M}{S_2+M} > 0$. Let h satisfy (4.7). Assume that there is an $x \ge 0$ such that $\mu(E_{x,h,S_1}) < e$. Then Lemma 4.4 implies the existence of t_1, t_2 such that $x \le t_1 < t_2 \le x+h$ and $s(t_1) \ge S_1$, $s(t_2) \le S_2$. But then Corollary 4.3 gives $\mu([t_1,t_2]\cap s^{-1}([S_2,S_1])) \ge \log \frac{S_1+M}{S_2+M}$. Since $[t_1,t_2]\cap s^{-1}([S_2,S_1]) \subset E_{x,h,S_1}$, we have $\mu(E_{x,h,S_1}) \ge \log \frac{S_1+M}{S_2+M} = e$, which is a contradiction. □

Remark 4.5. The author did not succeed in making full sense of the proof in [Kar1] corresponding to that of Corollary 4.3. It seems that there is a logical mistake in the reasoning, which is why we resorted to the above more topological approach. □

A. The Prime Number Theorem

Proposition A.1. *Defining* $\psi(x) := \sum_{n\le x} \Lambda(n)$, *we have* $\psi(x) \sim x$.

Proof. Since $\Lambda(x) \ge 0$, we have that ψ is non-negative and non-decreasing. Furthermore,

$$\begin{aligned}\sum_{n\le x} \psi\left(\frac{x}{n}\right) &= \sum_{n\le x}\sum_{m\le x/n} \Lambda(m) = \sum_{r\le x}\sum_{s|r}\Lambda(s)\\ &= \sum_{r\le x} \log r = x\log x - x + O(\log x)\end{aligned} \tag{A.1}$$

by Proposition 3.4(i) and (3.5). Now Theorem 1.1 implies $\psi(x) = x + o(x)$, or $\psi(x) \sim x$. □

Note that we still used only (i) of Proposition 3.4, but we will need now (iii):

Theorem A.2. *Let $\pi(x)$ be the number of primes $\leq x$ and p_n the n-th prime. Then*

$$\pi(x) \sim \frac{x}{\log x},$$
$$p_n \sim n \log n.$$

Proof. Using Proposition 3.4(iii), we compute

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = \sum_{p^k \leq x} \log p = \sum_{p \leq x} \log p \left\lfloor \frac{\log x}{\log p} \right\rfloor \leq \pi(x) \log x.$$

If $1 < y < x$ then

$$\pi(x) - \pi(y) = \sum_{y < p \leq x} 1 \leq \sum_{y < p \leq x} \frac{\log p}{\log y} \leq \frac{\psi(x)}{\log y}.$$

Thus $\pi(x) \leq y + \psi(x)/\log y$. Taking $y = x/\log^2 x$ this gives

$$\frac{\psi(x)}{x} \leq \frac{\pi(x) \log x}{x} \leq \frac{\psi(x)}{x} \frac{\log x}{\log(x/\log^2 x)} + \frac{1}{\log x},$$

thus $\psi(x) \sim \pi(x) \log x$. Together with Proposition A.1, this gives $\pi(x) \sim x/\log x$.

Taking logarithms of $\pi(x) \sim x/\log x$, we have $\log \pi(x) \sim \log x - \log\log x \sim \log x$ and thus $\pi(x) \log \pi(x) \sim x$. Taking $x = p_n$ and using $\pi(p_n) = n$ gives $n \log n \sim p_n$. □

Remark A.3. Karamata's proof of the Landau-Ingham theorem is obviously modeled on Selberg's original elementary proof [Sel] of the prime number theorem. However, Selberg worked with $f = \psi$ from the beginning. Most later proofs follow Selberg's approach, but there are some that work with M instead of ψ. Cf. the papers [PR, Kal] and the textbooks [GL, Ell]. As mentioned in the introduction, the result for M also follows easily from Theorem 1.1: □

Proposition A.4. *Defining $M(x) = \sum_{n \leq x} \mu(n)$, we have $M(x) = o(x)$.*

Proof. We define $f(x) = M(x) + \lfloor x \rfloor$, which is non-negative and non-decreasing. Now

$$F(x) = \sum_{n \leq x} M\left(\frac{x}{n}\right) + \left\lfloor \frac{x}{n} \right\rfloor = \sum_{n \leq x} \sum_{m \leq x/n} (\mu(m) + 1)$$
$$= \sum_{m \leq x} (\mu(m) + 1) \left\lfloor \frac{x}{m} \right\rfloor = 1 + \sum_{m \leq x} \left\lfloor \frac{x}{m} \right\rfloor,$$

where the last identity is just the first in (3.12). The remaining sum is known from Dirichlet's divisor problem and can be computed in elementary fashion,

$$\sum_{m\le x}\left\lfloor\frac{x}{m}\right\rfloor = x\log x + (2\gamma-1)x + O(\sqrt{x}), \tag{A.2}$$

cf. e.g. [TFM]. Thus $F(x) = x\log x + (2\gamma-1)x + O(\sqrt{x})$, and Theorem 1.1 implies $f(x) = x + o(x)$, thus $M(x) = o(x)$. □

Remark A.5. Note that we had to define $f(x) = M(x) + \lfloor x\rfloor$ and use (A.2) since $f(x) = M(x) + x$ is non-negative, but not non-decreasing. One can generalize Theorem 1.1 somewhat so that it applies to functions like $f(x) = M(x) + x$ weakly violating monotonicity. But the additional effort would exceed that for the easy proof of (A.2). □

References

[Ba] A. Balog, An elementary Tauberian theorem and the prime number theorem. *Acta Math. Acad. Sci. Hung.* **37** (1981), 285–299. Zbl 0418.10042 MR 0616898

[BGT] N. H. Bingham, C. M. Goldie and J. L. Teugels, *Regular variation*. Cambridge University Press, 1987. Zbl 0617.26001 MR 0898871

[Dia] H. G. Diamond, Elementary methods in the study of the distribution of prime numbers. *Bull. Amer. Math. Soc.* **7**, 553–589 (1982). Zbl 0505.10021 MR 0670132

[Ell] W. & F. Ellison, *Prime numbers*. Wiley, Hermann, 1985. (Original French version: W. Ellison, M. Mendès France: *Les nombres premiers*. Hermann, 1975.) Zbl 0624.10001 MR 0814687

[Erd1] P. Erdős, On a new method in elementary number theory which leads to an elementary proof of the prime-number theorem. *P.N.A.S.* **35** (1949), 374–384. Zbl 0034.31403 MR 0029411

[Erd2] P. Erdős, On a Tauberian theorem connected with the new proof of the prime number theorem. *J. Indian Math. Soc.* (N.S.) **13** (1949), 131–144. Zbl 0034.31501 MR 0033309

[EI] P. Erdős and A. E. Ingham, Arithmetical Tauberian theorems. *Acta Arith.* **9** (1964), 341–356. Zbl 0127.27101 MR 0176960

[EK] P. Erdős and J. Karamata, Sur la majorabilité C des suites des nombres réels. *Publ. Inst. Math. Acad. Serbe* **10** (1956), 37–52. Zbl 0075.04701 MR 0081981

[GL] A. O. Gelfond and Yu. V. Linnik, *Elementary methods in the analytic theory of numbers*. Pergamon Press, 1966. Zbl 0142.01403 MR 0201368

[Gor] B. Gordon, On a Tauberian theorem of Landau. *Proc. Amer. Math. Soc.* **9** (1958), 693–696. Zbl 0092.05002 MR 0104084

[Hew] E. HEWITT, Integration by parts for Stieltjes integrals. *Amer. Math. Monthly* **67** (1960), 419–423. Zbl 0095.26702 MR 0112937

[HS] E. HEWITT and K. STROMBERG, *Real and abstract analysis*. Springer, 1965. Zbl 0137.03202 MR 0188387

[HT] A. HILDEBRAND and G. TENENBAUM, On some Tauberian theorems related to the prime number theorem. *Compos. Math.* **90** (1994), 315–349. Zbl 0804.11052 MR 1272104

[Ing] A. E. INGHAM, Some Tauberian theorems connected with the prime number theorem. *J. London Math. Soc.* **20** (1945), 171–180. Zbl 0061.12802 MR 0017392

[Kal] M. KALECKI, A simple elementary proof of $M(x) = \sum_{n \le x} \mu(n) = o(x)$. *Acta Arithm.* **13** (1967), 1–7. Zbl 0159.06104 MR 0219493

[Kar1] J. KARAMATA, Sur les inversions asymptotiques de certains produits de convolution. *Bull. Acad. Serbe Sci.* (*N.S.*) *Cl. Sci. Math.-Nat. Sci. Math.* **3** (1957), 11–32. Zbl 0144.28302 MR 0103178

[Kar2] J. KARAMATA, Sur les procédés de sommation intervenant dans la théorie des nombres. pp. 12–31. *Colloque sur la théorie des suites tenu à Bruxelles du 18 au 20 décembre 1957.* Gauthier-Villars, 1958. Zbl 0102.03402 MR 0100745

[Kor] J. KOREVAAR, *Tauberian theory. A century of developments*. Springer, 2004. Zbl 1056.40002 MR 2073637

[Kou] D. KOUKOULOPOULOS, Pretentious multiplicative functions and the prime number theorem for arithmetic progressions. *Compos. Math.* **149** (2013), 1129–1149. Zbl 06192538 MR 3078641

[Lan] E. LANDAU, *Handbuch der Lehre von der Verteilung der Primzahlen*. Teubner, 1909. JFM 40.0232.08 MR 0068565

[Lev] N. LEVINSON, A motivated account of an elementary proof of the prime-number theorem. *Amer. Math. Monthly* **76** (1969), 224–245. Zbl 0172.06001 MR 0241372

[Nev] V. NEVANLINNA, *Über die elementaren Beweise der Primzahlsätze und deren äquivalente Fassungen*. Univ. of Helsinki Thesis, 1964. Zbl 0124.27703 MR 0168539

[Nik] A. NIKOLIĆ, The story of majorizability as Karamata's condition of convergence for Abel summable series. *Hist. Math.* **36** (2009), 405–419. Zbl 1187.01023 MR 2569877

[Pol] P. POLLACK, *Not always buried deep. A Second Course in Elementary Number Theory*. Amer. Math. Soc., 2009. Zbl 1187.11001 MR 2555430

[PR] A. G. POSTNIKOV and N. P. ROMANOV, A simplification of Selberg's elementary proof of the asymptotic law of distribution of primes. (In Russian.) *Uspek. Mat. Nauk*. 10:4 **66** (1955), 75–87. Zbl 0065.27701 MR 0074450

[Sch] W. SCHWARZ, *Einführung in Methoden und Ergebnisse der Primzahltheorie*. Bibliographisches Institut, 1969. Zbl 0217.31601 MR 0263750

[Sel] A. Selberg, An elementary proof of the prime-number theorem. *Ann. Math.* **50** (1949), 305–313. Zbl 0036.30604 MR 0029410

[TI] T. Tatuzawa and K. Iseki, On Selberg's elementary proof of the prime number theorem. *Proc. Japan Acad.* **27** (1951), 340–342. Zbl 0043.27401 MR 0046382

[TFM] G. Tenenbaum and M. Mendès France, *The prime numbers and their distribution*. Amer. Math. Soc., 2000. (Translation of *Les nombres premiers*, Presses Universitaires de France, 1997.) Zbl 0867.11002 MR 1756233

[Zag] D. Zagier, Newman's short proof of the prime number theorem. *Amer. Math. Monthly* **104** (1997), 705–708. Zbl 0887.11039 MR 1476753

(*Reçu le 20 novembre 2013*)

Michael Müger, Institute for Mathematics, Astrophysics and Particle Physics, Radboud University, Postbus 9010, 6500 GL Nijmegen, The Netherlands

e-mail: mueger@math.ru.nl