

## Chapter 10

# The Prime Number Theorem

In the wake of Euclid's proof of the infinitude of the primes, the question of how the primes were distributed among the integers became central — a question that has intrigued and challenged mathematicians ever since. The sieve of Eratosthenes provided a simple but very inefficient means of identifying which integers were prime, but attempts to find explicit, closed formulas for the  $n$ th prime, or for the number  $\pi(x)$  of primes less than or equal to a given number  $x$ , proved fruitless. Eventually extensive tables of integers and their least factors were compiled, detailed examination of which suggested that the apparently unpredictable occurrence of primes in the sequence of integers nonetheless exhibited some statistical regularity. In particular, in 1792 Euler asserted that for large values of  $x$ ,  $\pi(x)$  was approximately given by  $\frac{x}{\ln x}$ ; six years later, Legendre suggested  $\frac{x}{\ln x - 1}$  and (wrongly)  $\frac{x}{\ln x - 1.0836}$  as better approximations; and in 1849, in a letter to his student Encke (translated in the appendix to Goldstein 1973), Gauss mentioned his apparently long-held belief that the logarithmic integral

$$\text{li}(x) = \int_2^x \frac{1}{\ln t} dt$$

gave a still better approximation.<sup>1</sup> Using the notation  $f(x) \sim g(x)$  to denote the equivalence relation defined by  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1$ , those conjectures may be expressed in asymptotic form by the statements

$$\text{(PNT)} \quad \pi(x) \sim \frac{x}{\ln x}, \quad \pi(x) \sim \frac{x}{\ln x - 1}, \quad \text{and} \quad \pi(x) \sim \text{li}(x).$$

<sup>1</sup>Some texts instead define  $\text{li}(x)$  as  $\lim_{\epsilon \rightarrow 0} (\int_0^{1-\epsilon} 1/\ln t dt + \int_{1+\epsilon}^x 1/\ln t dt)$ , which adds a constant (approximately 1.04) to  $\text{li}(x)$  as defined above, but does not affect asymptotic arguments.

Clearly  $\frac{x}{\ln x} \sim \frac{x}{\ln x - 1}$ , and integration by parts can be used to show that  $\frac{x}{\ln x} \sim \text{li}(x)$  as well. For

$$\begin{aligned} \text{li}(x) &= \int_2^x \frac{1}{\ln t} dt = K + \int_e^x \frac{1}{\ln t} dt = K + \frac{x}{\ln x} - e + \int_e^x \frac{1}{(\ln t)^2} dt \\ &= (K - e) + \frac{x}{\ln x} + \left[ \int_e^{\sqrt{x}} \frac{1}{(\ln t)^2} dt + \int_{\sqrt{x}}^x \frac{1}{(\ln t)^2} dt \right], \end{aligned}$$

where  $K$  is a constant. Then since  $1/(\ln t)^2$  is less than or equal to 1 on the interval  $[e, \sqrt{x}]$  and less than or equal to  $1/(\ln \sqrt{x})^2 = 4/(\ln x)^2$  on the interval  $[\sqrt{x}, x]$ ,

$$\text{li}(x) \leq (K - e) + \frac{x}{\ln x} + \sqrt{x} + \frac{4x}{(\ln x)^2}.$$

But also,  $\text{li}(x) \geq (x - 2)/\ln x$ , since  $1/\ln t$  is decreasing on  $[2, x]$ . Hence

$$\frac{x - 2}{x} \leq \frac{\ln x}{x} \text{li}(x) \leq \frac{(K - e) \ln x}{x} + 1 + \frac{\ln x}{\sqrt{x}} + \frac{4}{\ln x}.$$

It follows immediately by L'Hopital's rule that  $\lim_{x \rightarrow \infty} \frac{\ln x}{x} \text{li}(x) = 1$ ; that is,  $\text{li}(x) \sim \frac{x}{\ln x}$ .

The three statements (PNT) are thus logically equivalent. *Ipso facto*, however, they do not capture the full strength of the conjectures made by Legendre and Gauss, since they do not indicate how accurately each of the three formulas approximates  $\pi(x)$ . For that, estimates of the absolute or relative errors involved in those approximations are also needed.

The truth of those conjectures, together with error estimates, was finally established in 1896, independently and nearly simultaneously, by Jacques Hadamard and Charles de la Vallée Poussin, using techniques of complex contour integration discussed further below. Their long and complicated proofs of what has ever since been called the Prime Number Theorem were distinct, and have been analyzed in detail in Narkiewicz (2000). Later, proofs were devised that avoided some of the delicate issues involved in contour integration over infinite paths by invoking Norbert Wiener's Tauberian theory for Fourier integrals; but the arguments remained complex. Eventually, more than 80 years after the original proofs, Donald Newman found a short proof involving integration only over finite contours (Newman 1980); and in the meantime, against the expectation of most of the mathematical community, Atle Selberg and Paul Erdős (partly independently and again nearly simultaneously) published so-called 'elementary' proofs, involving no recourse to complex-analytic methods (Erdős 1949; Selberg 1949).<sup>2</sup>

---

<sup>2</sup>In part for those proofs, Selberg was awarded a Fields Medal in 1950 and Erdős the 1951 Cole Prize in Number Theory.

Because of the length and complexity of all but Newman's proof, the discussion that follows, unlike that in the preceding chapters, does not give full details of the various proofs, but rather focuses on their essential ideas and the differences among them. Readers seeking further details may consult the recent monograph (Jameson 2003), which provides a self-contained presentation of a descendant of the original proofs, as well as Newman's proof and an elementary proof based on that given in Levinson (1969), all presented at a level accessible to students who have had only basic courses in real and complex analysis. An overview of work on the Prime Number Theorem in the century since its first proofs is given in Bateman and Diamond (1996), while Diamond (1982) and Goldfeld (2004) discuss the history of elementary proofs of the theorem. Narkiewicz (2000) provides a comprehensive history of the overall development of prime number theory up to the time of Hardy and Littlewood, including alternative proofs of many results.

## 10.1 Steps toward the goal: Prior results of Dirichlet, Chebyshev, and Riemann

As noted in Narkiewicz (2000), p. 49, "the first use of analytic methods in number theory was made by P.G. Dirichlet" during the years 1837–39. In particular, Dirichlet proved that if  $k$  and  $l$  are relatively prime integers with  $k < l$ , then the arithmetic progression  $\{nk + l\}$ , where  $n$  ranges over the positive integers, must contain infinitely many primes. In doing so Dirichlet considered the series now named after him (those of the form  $\sum_{n=1}^{\infty} \frac{f(n)}{n^s}$ ) and gave an argument that could be adapted to show that if  $f$  is a completely multiplicative complex-valued function (that is,  $f$  is not identically zero and  $f(mn) = f(m)f(n)$  for all integers  $m$  and  $n$ ) and  $\sum_{n=1}^{\infty} \frac{f(n)}{n^{s_0}}$  converges absolutely for some real value  $s_0$ , then it converges absolutely for all complex  $s$  with  $\operatorname{Re} s \geq s_0$  and

$$(28) \quad \sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_{p \text{ prime}} \sum_{j=0}^{\infty} \frac{f(p^j)}{p^{js}} = \prod_{p \text{ prime}} \frac{1}{1 - f(p)p^{-s}}.^3$$

With  $s$  restricted to real values, the special case  $f(n) = 1$  is Euler's product formula (cf. Chapter 7, footnote 6):

$$(29) \quad \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}$$

(from which it follows that  $\zeta(s) \neq 0$  for  $\operatorname{Re} s > 1$ ).

<sup>3</sup>In fact, for any absolutely convergent series  $\sum_{n=1}^{\infty} f(n)$  of non-zero terms, if  $f(n)$  is completely multiplicative and no  $f(n) = -1$ , then  $\sum_{n=1}^{\infty} f(n) = \prod_{p \text{ prime}} (1 - f(p))^{-1}$ .

Nine years later, in the first of two papers on  $\pi(x)$  (Chebyshev 1848), Pafnuty Chebyshev proved that for all integers  $k > 1$  and any constant  $C > 0$ , there are infinitely many integers  $m$  and  $n$  for which

$$(30) \quad \pi(m) < \text{li}(m) + C \frac{m}{\ln^k m} \quad \text{and} \quad \pi(n) > \text{li}(n) - C \frac{n}{\ln^k n}.$$

(For further details, see Narkiewicz (2000), pp. 98–102.) Consequently,

$$\pi(m) \frac{\ln m}{m} < \text{li}(m) \frac{\ln m}{m} + \frac{C}{\ln^{k-1} m} \quad \text{and} \quad \pi(n) \frac{\ln n}{n} > \text{li}(n) \frac{\ln n}{n} - \frac{C}{\ln^{k-1} n}$$

for arbitrarily large integers  $m$  and  $n$ . Therefore, since  $\text{li}(x) \sim \frac{x}{\ln x}$ , if the  $\lim_{x \rightarrow \infty} \frac{\pi(x) \ln(x)}{x}$  exists, then it must equal 1 (that is,  $\pi(x) \sim \frac{x}{\ln x}$ ).

In proving (30), Chebyshev made use of the Gamma function, defined for  $\text{Re } s > 0$  by  $\Gamma(s) = \int_0^\infty e^{-t} t^{s-1} dt$ . In his second paper (Chebyshev 1850), however, he introduced two number-theoretic functions related to  $\pi(x)$ , namely

$$\theta(x) = \sum_{\substack{p \text{ prime} \\ p \leq x}} \ln p \quad \text{and} \quad \psi(x) = \sum_{\substack{p \text{ prime} \\ p^n \leq x}} \ln p = \sum_{n=1}^m \theta(x^{1/n}),$$

where  $m$  is the largest integer for which  $2^m \leq x$ , and obtained important bounds on their values using only elementary means. In particular, since  $\ln x$  is an increasing function,  $\theta(x) \leq \pi(x) \ln x \leq x \ln x$  and  $\psi(x) \leq \pi(x) \ln x$ ; and since  $m \leq \ln x / \ln 2$ ,

$$\begin{aligned} \psi(x) - \theta(x) &= \sum_{n=2}^m \theta(x^{1/n}) \leq \theta(x^{1/2}) + m\theta(x^{1/3}) \\ &\leq \frac{1}{2} x^{1/2} \ln x + \frac{\ln x}{\ln 2} \frac{1}{3} x^{1/3} \ln x \\ &\leq \frac{1}{2} x^{1/2} \ln x + \frac{\ln x}{\ln 2} \frac{x^{1/2} \ln x}{x^{1/6}} \\ &\leq 4x^{1/2} \ln x, \end{aligned}$$

because the maximum value of  $\frac{\ln x}{\ln 2} \frac{1}{x^{1/6}}$  occurs for  $x = e^6$  and is less than 3.2. Thus  $\psi(x) - 4\sqrt{x} \ln x \leq \theta(x) \leq \psi(x)$ .

Hence if  $\lim_{x \rightarrow \infty} \frac{\psi(x)}{x}$  exists, then so does  $\lim_{x \rightarrow \infty} \frac{\theta(x)}{x}$ , and they are equal. Furthermore, if  $\lim_{x \rightarrow \infty} \frac{\theta(x)}{x}$  exists, then so must  $\lim_{x \rightarrow \infty} \frac{\pi(x) \ln(x)}{x}$ , because for  $0 < \epsilon < 1$ ,

$$\theta(x) \geq \sum_{x^{1-\epsilon} \leq p \leq x} \ln p \geq (\pi(x) - \pi(x^{1-\epsilon}))(1-\epsilon) \ln x \geq (\pi(x) - x^{1-\epsilon})(1-\epsilon) \ln x;$$

so

$$\frac{\theta(x)}{x} \leq \frac{\pi(x) \ln x}{x} \leq \frac{1}{1-\epsilon} \frac{\theta(x)}{x} + \frac{\ln x}{x^\epsilon}.$$

As already noted,  $\lim_{x \rightarrow \infty} \frac{\pi(x) \ln(x)}{x}$  must then equal 1. So to prove the Prime Number Theorem it suffices to prove that either  $\lim_{x \rightarrow \infty} \frac{\psi(x)}{x}$  or  $\lim_{x \rightarrow \infty} \frac{\theta(x)}{x}$  exists.

The principal theorem of Chebyshev (1850) was that there exist constants  $C_1$ ,  $C_2$ , and  $C_3$  such that for all  $x \geq 2$ ,

$$(31) \quad (i) \quad C_1 x \leq \psi(x) \leq C_2 x \quad \text{and} \quad (ii) \quad C_3 x \leq \theta(x) \leq C_2 x.$$

Indeed, Chebyshev's methods (involving Stirling's formula) showed that  $C_2$  can be taken to be 1.1224 and, if  $x \geq 37$ , that  $C_3$  can be taken to be .73 (see Narkiewicz 2000, p. 111); and from the latter estimate Chebyshev deduced Bertrand's postulate (that for every  $x > 1$  the interval  $(x, 2x]$  contains some prime) as an immediate corollary. For direct calculation shows that Bertrand's postulate holds for  $1 \leq x < 37$ , and if it failed for some  $x \geq 37$ , the estimate would imply the contradiction that  $1.46x \leq \theta(2x) = \theta(x) \leq 1.13x$ .

Without using Stirling's formula, considerations involving binomial coefficients yield the weaker inequality  $\theta(x) \leq x \ln 4 \approx 1.3863x$ . The argument, given in Jameson (2003), p. 36, uses no number-theoretic facts beyond Euclid's lemma and its corollary that if distinct primes divide  $n$ , so does their product. Namely, for any given  $n$  the expansion of  $(1+1)^{2n+1}$  includes the two equal binomial coefficients  $\binom{2n+1}{n}$  and  $\binom{2n+1}{n+1}$ , so  $\binom{2n+1}{n} < 2^{2n} = 4^n$ . If  $p_i, \dots, p_{i+j}$  are all the primes between  $n+2$  and  $2n+1$ , inclusive, then none of those primes divides the denominator  $n!$  of  $\binom{2n+1}{n}$ , but each of them, and hence the product of all of them, divides its numerator,  $n! \binom{2n+1}{n}$ . Therefore that product must divide  $\binom{2n+1}{n}$  itself and so must be less than  $\binom{2n+1}{n}$ . Consequently,

$$\theta(2n+1) - \theta(n+1) = \sum_{k=0}^j \ln p_{i+k} = \ln(p_i \cdots p_{i+j}) \leq \ln \binom{2n+1}{n} \leq n \ln 4.$$

Then since  $\theta(n) \leq n \ln 4$  for  $n = 2, 3, 4$ , we may assume by induction on  $m$  that  $\theta(k) \leq k \ln 4$  for  $k \leq 2m$ . The assumption holds for  $m = 2$ , and  $m+1 < 2m$  for  $m \geq 2$ . So  $\theta(m+1) \leq (m+1) \ln 4$ , whence by the last displayed inequality,  $\theta(2m+1) \leq (2m+1) \ln 4$ ; and since  $2m+2$  is not prime,  $\theta(2m+2) = \theta(2m+1)$ . Thus  $\theta(2m+2) \leq (2m+1) \ln 4 < (2m+2) \ln 4$ , finishing the induction.

The inequality (31)(ii) also entails corresponding bounds on  $\pi(x)$ . For  $\pi(x)$  can be expressed as the sum  $\sum_{2 \leq n \leq x} (f(n)/\ln n)$ , where  $f(n) = \ln n$  if  $n$  is prime and 0 otherwise, and Abel summation<sup>4</sup> then gives

$$\pi(x) = \frac{\theta(x)}{\ln x} + \int_2^x \frac{\theta(t)}{t(\ln t)^2},$$

so (31)(ii) yields

$$\frac{C_3 x}{\ln x} + C_3 \int_2^x \frac{1}{(\ln t)^2} dt \leq \pi(x) \leq \frac{C_2 x}{\ln x} + C_2 \int_2^x \frac{1}{(\ln t)^2} dt.$$

On the other hand, integration by parts shows that

$$\text{li}(x) + \frac{2}{\ln 2} = \frac{x}{\ln x} + \int_2^x \frac{1}{(\ln t)^2} dt.$$

Hence

$$C_3 \left( \text{li}(x) + \frac{2}{\ln 2} \right) \leq \pi(x) \leq C_2 \left( \text{li}(x) + \frac{2}{\ln 2} \right).$$

Similar arguments (see Jameson (2003), p. 35) show that for any  $\epsilon > 0$ , an  $x_0$  must exist such that  $(C_3 - \epsilon)\text{li}(x) \leq \pi(x) \leq (C_2 + \epsilon)\text{li}(x)$  for all  $x > x_0$ . The Prime Number Theorem would follow if it could be shown that the bounds in Chebyshev's estimate can be taken to be  $C_2 = C_3 = 1$ . But Chebyshev's methods were inadequate for that task.

A key breakthrough came a decade later in Bernhard Riemann's memoir "Ueber die Anzahl der Primzahlen unter einer gegebener Größe" (Riemann 1860). Riemann began by recalling Euler's product formula (29), but took  $s$  therein to be a *complex* number and initially defined  $\zeta(s)$  to be the function of  $s$  given by the expressions on each side of Euler's formula whenever both expressions converged — that is, when  $\text{Re } s > 1$ . So defined,  $\zeta(s)$  is analytic on the half plane  $\text{Re } s > 1$ , but Riemann went on immediately to extend it to a function analytic on all of  $\mathbb{C}$  except  $s = 1$ . To do so he employed the Gamma function (actually, the function  $\Pi(s) = \Gamma(s + 1)$ ) together with the substitution  $t = nx$  to obtain

$$\frac{\Gamma(s)}{n^s} = \int_0^\infty e^{-nx} x^{s-1} dx$$

---

<sup>4</sup>The result of applying the formula

$$\sum_{2 \leq n \leq x} a(n)g(n) = \left[ \sum_{n \leq x} a(n) \right] g(x) - \int_2^x \left[ \sum_{n \leq t} a(n) \right] g'(t) dt,$$

valid whenever  $a(1) = 0$  and  $g(x)$  has a continuous derivative on  $[2, x]$ .

as the  $n$ th term of the geometric series for  $\Gamma(s)\zeta(s)$ . Then

$$\Gamma(s)\zeta(s) = \sum_{n=1}^{\infty} \int_0^{\infty} e^{-nx} x^{s-1} dx = \int_0^{\infty} x^{s-1} \sum_{n=1}^{\infty} e^{-nx} dx = \int_0^{\infty} \frac{x^{s-1}}{e^x - 1} dx.$$

For  $s \neq 1$  he equated  $e^{-\pi si} - e^{\pi si}$  times the latter integral to the contour integral

$$\int_C \frac{(-x)^{s-1}}{e^x - 1} dx,$$

“taken from  $+\infty$  to  $+\infty$  in a positive sense” around a region containing “in its interior the point 0 but no other point of discontinuity of the integrand.” He concluded that  $\zeta(s)$  could then be defined for all  $s \neq 1$  by the equation

$$(32) \quad 2 \sin(\pi s) \Gamma(s) \zeta(s) = i \int_C \frac{(-x)^{s-1}}{e^x - 1} dx,$$

and that so defined  $\zeta(s)$  would be analytic except for a simple pole at  $s = 1$ ; that is,  $\zeta(s) - 1/(s - 1)$  would be an entire function.

Riemann noted that for  $\operatorname{Re} s < 0$  the same integral could be taken in the reverse direction, surrounding the region exterior to the curve  $C$ . It would then be “infinitely small for all  $s$  of infinitely large modulus,” and the integrand would be discontinuous only at the points  $2n\pi i$ . Its value would thus be  $\sum_{n=1}^{\infty} (-n2\pi i)^{s-1} (-2\pi i)$ , whence

$$2 \sin(\pi s) \Gamma(s) \zeta(s) = (2\pi)^s \sum_{n=1}^{\infty} n^{s-1} ((-i)^{s-1} + i^{s-1}) = (2\pi)^s \zeta(1-s) 2 \sin \frac{\pi s}{2},$$

which reduces to the functional equation

$$\zeta(1-s) = \frac{2}{(2\pi)^s} \cos\left(\frac{\pi s}{2}\right) \Gamma(s) \zeta(s).$$

In the remainder of his memoir, Riemann made some assertions concerning  $\pi(x)$  and the non-real zeros of the zeta function — assertions equivalent to the following statements:

- (i) The number  $N(T)$  of non-real roots  $\rho$  of  $\zeta(s)$  for which  $|\operatorname{Im} \rho| \leq T$  (with multiple roots counted according to their multiplicity) is asymptotically equal to

$$(33) \quad \frac{T}{2\pi} \ln \frac{T}{2\pi} - \frac{T}{2\pi}.$$

- (ii) It is “very likely” that all non-real zeros of  $\zeta(s)$  satisfy  $\operatorname{Re} s = 1/2$ .

(iii) If  $F(x) = 1/2k + \sum_{n=1}^{\infty} \frac{\pi(x^{1/n})}{n}$  whenever  $x = p^k$  for some prime  $p$  and  $F(x) = \sum_{n=1}^{\infty} \frac{\pi(x^{1/n})}{n}$  otherwise, then for  $x > 1$ ,

$$F(x) = \text{li}(x) - \sum_{\rho} \text{li}(x^{\rho}) + \int_x^{\infty} \frac{dx}{(x^2 - 1)x \ln x} - \ln 2,$$

where  $\rho$  ranges over all non-real roots of  $\zeta(s)$ .

Assertions (i) and (iii) were later proved rigorously in von Mangoldt (1895), while the conjecture that all non-real zeros of  $\zeta(s)$  satisfy  $\text{Re } s = 1/2$  is the still unproven Riemann Hypothesis.

Beyond the specific results it contained, three aspects of Riemann's memoir were seminal for the subsequent proofs of the Prime Number Theorem: the idea of regarding the series  $\sum_{n=1}^{\infty} n^{-s}$  as a function of a *complex* argument; the extension of that function to a function analytic in the region  $s \neq 1$ , to which Cauchy's theory of contour integration could be applied; and the revelation that the distribution of prime numbers was intimately related to the location of the non-real zeros of that extended zeta function.<sup>5</sup>

## 10.2 Hadamard's proof

Just a year after receiving his degree of Docteur ès Sciences, Jacques Hadamard published the first of a series of papers concerning properties of the Riemann zeta function. The series culminated in his paper Hadamard (1896b), in which he proved the Prime Number Theorem in the form  $\lim_{x \rightarrow \infty} \frac{\theta(x)}{x} = 1$ .

Hadamard began that paper by noting that  $\zeta(s)$  was analytic except for a simple pole at  $s = 1$ , and though it had infinitely many zeros with real part between 0 and 1 (a consequence of results in his earlier paper Hadamard 1893), it was non-zero for all  $s$  with  $\text{Re } s > 1$ . His first goal was then to show that  $\zeta(s)$  was also non-zero for all  $s$  with  $\text{Re } s = 1$ .

To establish that he considered  $\ln(\zeta(s))$ , which by Euler's product theorem and the Maclaurin series for  $\ln(1 - x)$  is equal to  $\sum_{k=1}^{\infty} \sum_{p \text{ prime}} \frac{1}{kp^{ks}}$ . He split that sum into two parts, so that

$$(34) \quad \ln(\zeta(s)) = \sum_{p \text{ prime}} \frac{1}{p^s} + \sum_{k=2}^{\infty} \sum_{p \text{ prime}} \frac{1}{kp^{ks}} = S(s) + \sum_{k=2}^{\infty} \sum_{p \text{ prime}} \frac{1}{kp^{ks}},$$

<sup>5</sup>A detailed commentary on developments stemming from Riemann's classic paper is given in Edwards (1974).



and noted that since  $(s - 1)\zeta(s)$  is an entire function, it is bounded in any closed neighborhood of  $s = 1$ . Thus in any such neighborhood the difference  $\ln(\zeta(s)) - \ln \frac{1}{s-1} = \ln((s-1)\zeta(s))$  is a bounded function, so if  $s$  approaches 1 from the right along the real axis,  $\ln(\zeta(s))$  must approach  $+\infty$ . But by (34), in any sufficiently small neighborhood of  $s = 1$ ,  $\ln(\zeta(s))$  also differs from  $S(s)$  by a bounded expression, since the second term on the right of (34) is analytic for  $\text{Re } s > 1/2$ . So  $S(s)$  also approaches  $+\infty$  as  $s$  approaches 1 from the right along the real axis.

On the other hand, consider  $\ln(|\zeta(s)|)$ . By (34), for  $s = \sigma + it$ ,  $\ln(|\zeta(s)|) = \text{Re } \ln(\zeta(s))$  differs from

$$P(s) = \sum_{p \text{ prime}} \text{Re } p^{-s} = \sum_{p \text{ prime}} \frac{\cos(t \ln p)}{p^\sigma}$$

by a bounded function; and if  $\zeta(1 + it_0) = 0$  for some  $t_0 \neq 0$ , then  $1 + it_0$  must be a simple root of  $\zeta(s)$ , whence  $\ln(|\zeta(\sigma + it)|)$  must differ from  $\ln(\sigma - 1)$  by a bounded function. So  $\ln(|\zeta(\sigma + it_0)|)$  and  $P(\sigma + it_0)$  must both approach  $-\infty$  as  $\sigma$  approaches 1 from the right, and  $P(\sigma + it_0)$  must differ from  $-S(\sigma + it_0)$  by a bounded function.

Now suppose  $0 < \rho = 1 - \epsilon < 1$ ,  $0 < \alpha < 1$ , and let  $P_n$  and  $S_n$  denote the partial sums of  $P$  and  $S$  for  $p \leq n$ . Hadamard divided the set of primes  $p \leq n$  into two subsets, according to whether or not, for some integer  $k$ , the prime  $p$  satisfied the inequality  $|t_0 \ln p - (2k + 1)\pi| < \alpha$ . Writing  $S_n = S'_n + S''_n$  and  $P_n = P'_n + P''_n$  to correspond to that division, he used elementary inequalities to conclude that there must exist an integer  $N_\rho$  such that for all  $n \geq N_\rho$ ,  $\rho_n(\sigma) = S'_n(\sigma + it_0)/S_n(\sigma + it_0) > \rho$ , since otherwise  $P(\sigma + it_0) \geq -\theta S(\sigma)$ , where  $\theta = 1 - \epsilon + \epsilon \cos \alpha < 1$ . Then, by the result of the previous paragraph, for some function  $F$  bounded throughout the half-plane  $\text{Re } s > 1/2$ ,  $-S(\sigma + it_0) + F(\sigma + it_0) \geq -\theta S(\sigma)$ . That is,  $F(\sigma + it_0) \geq (1 - \theta)S(\sigma)$ , with  $1 - \theta > 0$ . But as noted above,  $\lim_{\sigma \rightarrow 1^+} S(\sigma) = +\infty$ .

The argument just given, which rests on the assumption that  $\zeta(1 + it_0) = 0$ , applies to any  $\rho$  satisfying  $0 < \rho < 1$ . If  $\rho$  is further required to satisfy  $\frac{1}{1 + \cos(2\alpha)} < \rho < 1$ , then similar manipulations of inequalities show that for  $n \geq N_\rho$  and  $s = \sigma + i(2t_0)$ ,

$$P_n(s) \geq \rho_n S_n(\sigma) \cos(2\alpha) + (\rho_n - 1)S_n(\sigma) \\ > \rho S_n(\sigma) \cos(2\alpha) + (\rho - 1)S_n(\sigma) = \Theta S_n(\sigma),$$

with  $\Theta = \rho[(1 + \cos(2\alpha)) - 1] > 0$ . Hence  $P(s) = \lim_{n \rightarrow +\infty} P_n(s) \geq \Theta S(\sigma)$ . Once again,  $\lim_{\sigma \rightarrow 1} S(\sigma) = +\infty$ , so  $P(\sigma + i(2t_0))$ , and thus also  $\ln|\zeta(\sigma + i(2t_0))|$ , must approach  $+\infty$  as  $\sigma$  approaches 1 from the right. But that cannot be, since  $\zeta(s)$  is analytic at the point  $1 + i(2t_0)$ . The assumption that  $\zeta(1 + it_0) = 0$  is thereby refuted.

For the rest of his proof Hadamard drew upon ideas of E. Cahen, who in his doctoral dissertation at the École Normale Supérieure had unsuccessfully attempted to prove the Prime Number Theorem.

Given real numbers  $a$  and  $x$ , with  $0 < x \neq 1$ , Cahen had considered the contour integrals

$$\frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{x^z}{z} dz, \quad \text{for } a > 0, \quad \text{and} \quad -\frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{x^z \zeta'(z)}{z \zeta(z)} dz, \quad \text{for } a > 1.$$

Hadamard considered instead the integrals

$$\frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{x^z}{z^\mu} dz, \quad \text{for } a > 0, \quad \text{and} \quad -\frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{x^z \zeta'(z)}{z^\mu \zeta(z)} dz, \quad \text{for } a > 1,$$

where  $\mu > 0$ , which he denoted by  $J_\mu$  and  $\psi_\mu$ , respectively.

In order to evaluate the integrals  $J_\mu$ , Hadamard distinguished three cases:  $\mu$  is an integer,  $n$ ;  $\mu$  is non-integral and  $x < 1$ ; or  $\mu$  is non-integral and  $x > 1$ . For his proof of the Prime Number Theorem, however, only the first case was needed (indeed, just the case  $n = 2$ ), which he established by integrating by parts  $n - 1$  times, using the identity  $1/z^n = \frac{(-1)^{n-1}}{(n-1)!} \frac{d^{n-1}}{dz^{n-1}}(1/z)$ .<sup>6</sup> That gave

$$J_n = \frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{x^z}{z^n} dz = \frac{1}{2\pi i} \frac{(\ln x)^{n-1}}{(n-1)!} \int_{a-\infty i}^{a+\infty i} \frac{x^z}{z} dz,$$

whence

$$(35) \quad J_n = \begin{cases} 0 & \text{if } x < 1, \\ \frac{(\ln x)^{n-1}}{(n-1)!} & \text{if } x > 1, \end{cases}$$

since von Mangoldt had shown the year before that

$$(36) \quad \frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{x^z}{z} dz = \begin{cases} 0 & \text{if } x < 1, \\ 1/2 & \text{if } x = 1, \\ 1 & \text{if } x > 1, \end{cases} \quad (\text{von Mangoldt 1895}).$$

---

<sup>6</sup>In the other cases Hadamard used the identity  $1/z^\mu = \frac{(-1)^{\mu-1}}{\Gamma(\mu)} \frac{d^{\mu-1}}{dz^{\mu-1}}(1/z)$ , together with Cauchy's integral theorem, to obtain the general formula

$$J_\mu = \frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{x^z}{z^\mu} dz = \begin{cases} 0 & \text{if } x < 1, \\ \frac{(\ln x)^{\mu-1}}{\Gamma(\mu)} & \text{if } x > 1. \end{cases}$$

To evaluate the integrals  $\psi_\mu$ , Hadamard noted that by (29),  $\ln(\zeta(s)) = -\sum_p \ln(1 - p^{-s})$ , so logarithmic differentiation gives

$$(37) \quad \frac{\zeta'(s)}{\zeta(s)} = -\sum_p \frac{(\ln p)p^{-s}}{1 - p^{-s}} = -\sum_p \ln p \sum_{k=1}^{\infty} \frac{1}{p^{ks}} = -\sum_{k=1}^{\infty} \frac{\Lambda(k)}{k^s},$$

where  $\Lambda$  denotes the von Mangoldt function, defined by  $\Lambda(k) = \ln p$  if  $k = p^m$  for some prime  $p$  and integer  $m$  and  $\Lambda(k) = 0$  otherwise.

Then for  $\mu = 2$ ,

$$\begin{aligned} \psi_2(x) &= -\frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s^2} ds \\ &= \sum_p \ln p \frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \sum_{k=1}^{\infty} \frac{x^s}{p^{ks}} \frac{1}{s^2} ds \\ &= \sum_p \ln p \sum_{k=1}^{\infty} \frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{(x/p^k)^s}{s^2} ds, \end{aligned}$$

where by (35) the integrals in the last member are equal to 0 if  $x/p^k < 1$  (that is, if  $x^{1/k} < p$ ) and equal to  $\ln\left(\frac{x}{p^k}\right)$  otherwise. Hence

$$\psi_2(x) = \sum_{k=1}^{\infty} \sum_{p \leq x^{1/k}} \ln p \ln\left(\frac{x}{p^k}\right) = \sum_{p \leq x} \ln p \ln\left(\frac{x}{p}\right) + \sum_{k=2}^{\infty} \sum_{p \leq x^{1/k}} \ln p \ln\left(\frac{x}{p^k}\right).$$

The double sum in the last term of the equation above is only apparently infinite, since the inner sum is vacuous for  $k > \ln x / \ln 2$ . Thus finally

$$(38) \quad \psi_2(x) = \sum_{p \leq x} \ln p \ln\left(\frac{x}{p}\right) + \sum_{k=2}^{\lfloor \frac{\ln x}{\ln 2} \rfloor} \sum_{p \leq x^{1/k}} \ln p \ln\left(\frac{x}{p^k}\right),$$

where the brackets above the penultimate summation symbol denote the greatest integer function.

In the second term of (38) the first summation involves no more than  $\ln x / \ln 2$  summands and the second summation no more than  $\sqrt{x}$ , the largest of which is that for  $k = 2$ . Consequently,

$$\psi_2(x) \leq \sum_{p \leq x} \ln p \ln\left(\frac{x}{p}\right) + \sqrt{x} \frac{\ln x}{\ln 2} \ln(\sqrt{x}) \ln x = \sum_{p \leq x} \ln p \ln\left(\frac{x}{p}\right) + \sqrt{x} \frac{\ln^3 x}{2 \ln 2}.$$

When divided by  $x$ , the last term above approaches 0 as  $x$  approaches  $\infty$ . Hadamard's final goals were then to show

(i) that  $\lim_{x \rightarrow \infty} \frac{1}{x} \sum_{p \leq x} \ln p \ln \left( \frac{x}{p} \right) = 1$

and

(ii) that (i) implies that  $\lim_{x \rightarrow \infty} \frac{\theta(x)}{x} = 1$ .

Hadamard established (ii) via elementary but rather involved  $\epsilon - \delta$  computations, showing that for every  $\epsilon > 0$ ,  $x - 2\epsilon x \leq \theta(x) \leq x + \frac{5\epsilon}{2}x$ . (For details, see Narkiewicz 2000, pp. 202–204.) To prove (i) he showed that  $\lim_{x \rightarrow \infty} \frac{\psi_2(x)}{x} = 1$ , using Cauchy's integral theorem, von Mangoldt's theorem justifying Riemann's asymptotic estimate (33) of the number  $N(T)$  of non-real roots  $\rho$  of  $\zeta(s)$  for which  $|\operatorname{Im} \rho| \leq T$ , and two results from his earlier paper Hadamard (1896a)—in particular,

(A) If  $\rho_1, \rho_2, \dots$  are the non-real roots of  $\zeta(s)$  ordered according to increasing absolute value, then  $\sum_n 1/|\rho_n|^2$  converges.

and

(B) If  $s \neq 1$  is not a root of  $\zeta(s)$ , then for some constant  $K$

$$(39) \quad \frac{\zeta'(s)}{\zeta(s)} = \frac{1}{1-s} + \sum_{\rho} \left( \frac{1}{s-\rho} + \frac{1}{\rho} \right) + K,$$

where  $\rho$  ranges over all roots of  $\zeta(s)$ , ordered according to increasing absolute value.

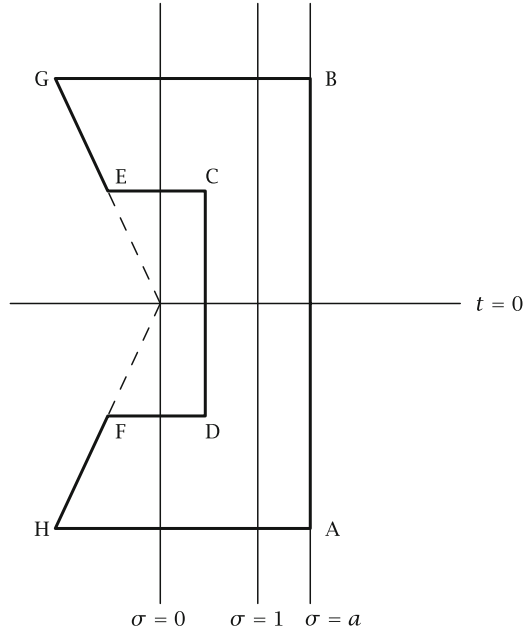
It follows from (A) that for any  $\epsilon > 0$  there is an integer  $M$  such that  $\sum_{n > M} 1/|\rho_n|^2 < \epsilon$ . Let  $I$  be the maximum of  $|\operatorname{Im} \rho_n|$  for  $n \leq M$  and  $R$  the maximum of  $|\operatorname{Re} \rho_n|$  for  $n \leq M$ . Since  $\zeta(s)$  has no roots  $\rho$  with  $\operatorname{Re} \rho \geq 1$ ,  $R < 1$ . Then, given  $a > 1$ , to compute

$$\psi_2(x) = -\frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s^2} ds$$

Hadamard considered the infinite family of polygons  $\Gamma = \mathbf{ABGECDFHA}$  defined in terms of a real parameter  $y$  as follows (see Figure 10.1): Choose  $d > I$  such that no root of  $\zeta(s)$  lies on the line  $\operatorname{Im}(s) = d$ , take  $c$  to be a real number satisfying  $R < c < 1$ , and let  $\mathbf{C}$  be the point  $c + di$  and  $\mathbf{D}$  the point  $c - di$ . For  $y > d$ , let  $\mathbf{A}$  be the point  $a - yi$  and  $\mathbf{B}$  the point  $a + yi$ . Finally, fix  $e < 0$ , let  $\mathbf{E}$  be the point  $e + di$  and  $\mathbf{F}$  the point  $e - di$ , and denote by  $\mathbf{G}$  and  $\mathbf{H}$  the points where the lines  $\operatorname{Im}(s) = \pm y$  intersect the lines from the origin that pass through  $\mathbf{E}$  and  $\mathbf{F}$ .

**Fig. 10.1** Contour used by Hadamard to evaluate the integral

$$\frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{\zeta'(z)}{\zeta(z)} \frac{x^z}{z^2} dz$$



Hadamard used (33) and (39) to show that the integral of  $\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s^2}$  along each of the segments **BG** and **AH** approaches 0 as  $y$  approaches  $\infty$ .

In particular, he concluded from (33) that for any  $A > 1$  and any positive integer  $\lambda$  the number of roots  $\rho$  of  $\zeta(s)$  for which  $|\text{Im } \rho|$  lies between  $A^{3\lambda+3}$  and  $A^{3\lambda}$  does not exceed  $K\lambda A^{3\lambda}$ , where  $K$  is a constant. The same bound applies *a fortiori* to the number of roots  $N_r$  for which  $|\text{Im } \rho|$  lies between  $3\lambda + 1$  and  $3\lambda + 2$ . On the other hand, the asymptotic expression for  $N(T)$  given in (33) approaches  $\infty$  as  $\lambda$  does, so for sufficiently large  $\lambda$  there must be two consecutive roots  $\rho_j$  and  $\rho_{j+1}$  whose imaginary parts  $\beta_j$  and  $\beta_{j+1}$  differ by at least  $\frac{A^{3\lambda+2} - A^{3\lambda+1}}{K\lambda A^{3\lambda}} = \frac{A(A-1)}{K\lambda}$ . Consequently, for  $y_\lambda = (\beta_j + \beta_{j+1})/2$ , any root of  $\zeta(s)$  must lie above or below the line  $\text{Im}(s) = y_\lambda$  by at least  $\frac{A(A-1)}{2K\lambda}$ .

That being so, if **BG** lies along  $\text{Im}(s) = y_\lambda$ , the summation in (39) may be estimated by splitting it into two parts, the first sum ranging over all roots  $\rho$  satisfying  $A^{3\lambda} \leq \text{Im } \rho \leq A^{3\lambda+3}$  and the second over all other roots. The definition of  $y_\lambda$  entails that the first sum is bounded by  $C_1 y_\lambda \ln^2 y_\lambda$  and the second by  $C_2 y_\lambda \ln y_\lambda$ , for some constants  $C_1$  and  $C_2$ , so  $|\zeta'(s)/\zeta(s)|$  is itself bounded by a constant multiple of  $y_\lambda \ln^2 y_\lambda$ . Hence

$$\left| \int_{\text{BG}} \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s^2} ds \right| \leq C_3 \frac{y_\lambda \ln^2 y_\lambda}{y_\lambda^2} \left| \int_{\text{BG}} x^s ds \right| = C_4 \frac{\ln^2 y_\lambda}{y_\lambda},$$

which approaches 0 as  $y_\lambda$  approaches  $\infty$ . Similar considerations show that the integral along  $\mathbf{AH}$  also approaches 0 as  $y_\lambda$  approaches  $\infty$ .

Cauchy's integral theorem yields that

$$I_\Gamma = \frac{1}{2\pi i} \int_\Gamma \frac{\zeta'(s) x^s}{\zeta(s) s^2} ds$$

is equal to the sum of the residues at the poles of  $\zeta'(s)$  and the roots of  $\zeta(s)$  that lie within  $\Gamma$ . The only such pole occurs at  $s = 1$ , where the residue is  $-x$ , and by construction, the sum of the residues of the roots inside  $\Gamma$  cannot exceed  $\epsilon x$ . Thus, writing  $I_\Gamma = I_{\mathbf{AB}} + I_{\mathbf{BG}} + I_{\mathbf{GE}} + I_{\mathbf{ECDF}} + I_{\mathbf{FH}} + I_{\mathbf{HA}}$  and taking the limit as  $\lambda$  approaches  $\infty$ ,

$$-x \leq -\psi_2(x) + 0 + \lim_{\lambda \rightarrow \infty} I_{\mathbf{GE}} + I_{\mathbf{ECDF}} + \lim_{\lambda \rightarrow \infty} I_{\mathbf{FH}} + 0 \leq -x + \epsilon x.$$

The quantity  $I_{\mathbf{ECDF}}$  is a constant, since the boundary segment  $\mathbf{ECDF}$  is fixed regardless of the value of  $\lambda$ . So  $\lim_{x \rightarrow \infty} (I_{\mathbf{ECDF}}/x) = 0$ , and to show that  $\lim_{x \rightarrow \infty} (\psi_2(x)/x) = 1$ , it remains only to show that

$$\lim_{x \rightarrow \infty} \frac{\lim_{\lambda \rightarrow \infty} I_{\mathbf{GE}}}{x} = \lim_{x \rightarrow \infty} \frac{\lim_{\lambda \rightarrow \infty} I_{\mathbf{FH}}}{x} = 0.$$

For that, Hadamard noted that for  $s$  on  $\mathbf{GE}$  and  $\mathbf{FH}$  (regardless of the value of  $\lambda$ , which determines the position of  $\mathbf{G}$  and  $\mathbf{H}$ ) and any root  $\rho$  of  $\zeta(s)$ , the ratio  $\left| \frac{s - \rho}{\rho} \right|$  must be greater than a fixed constant, so by (39),  $\left| \frac{1}{s} \frac{\zeta'(s)}{\zeta(s)} \right|$  is finite. Therefore for some constant  $K$ ,  $I_{\mathbf{GE}}$  is less than  $K \int_{\mathbf{GE}} |x^s|/|s| ds$  — a finite quantity — and likewise for  $I_{\mathbf{FH}}$ . So  $\frac{I_{\mathbf{GE}}}{x}$  and  $\frac{I_{\mathbf{FH}}}{x}$  both approach 0 as  $x$  approaches  $\infty$ .

### 10.3 The proof of de la Vallée Poussin

In a note appended to the end of Hadamard (1896b), Hadamard remarked that while correcting the proofs for that paper he had been notified of de la Vallée Poussin's paper on the same topic (de la Vallée Poussin 1896), and he acknowledged that their proofs, found independently, had some points in common. In particular, both involved showing that  $\zeta(s)$  had no roots on the line  $\operatorname{Re} s = 1$ . But their methods for doing so were entirely different, and Hadamard judged that his own method was simpler.<sup>7</sup>

<sup>7</sup>In a statement quoted on page 198 of Narkiewicz (2000), de la Vallée Poussin agreed, but nevertheless claimed priority for the proof of that result.

Indeed, of the twenty-one pages in Hadamard (1896b), only two and a half are devoted to his proof that  $\zeta(s)$  has no root of the form  $1 + \beta i$ ; and as Hadamard noted, that proof rested on only two simple properties of  $\zeta(s)$ : that its logarithm could be expressed as a series of the form  $\sum a_n e^{-\lambda_n s}$ , for some positive constants  $a_n$  and  $\lambda_n$ , and that  $\zeta$  itself was analytic on the limiting line of convergence of that series, with the exception of a single simple pole.

In contrast, de la Vallée Poussin's proof that  $\zeta(s)$  had no roots on the line  $\text{Re } s = 1$  was the subject of the third chapter of his paper de la Vallée Poussin (1896), which took up 18 of its 74 pages. Like Hadamard's proof, de la Vallée Poussin's was by contradiction and used the fact that any root  $s = 1 + \beta i$  of  $\zeta(s)$  would have to be a simple root. But unlike Hadamard's argument, which rested on boundedness considerations, de la Vallée Poussin's was based on the uniqueness of certain Fourier series expansions. Specifically, he considered complex-valued functions  $f(y)$  of a real variable  $y$  that for  $y > 1$  have the form  $L(y) + P(y)$ , where  $L(y)$  denotes a function that approaches a finite limit  $A$  as  $y$  approaches  $\infty$  and  $P(y)$  denotes an infinite series of the form

$$\sum_{n=1}^{\infty} [c_n \cos(\alpha_n \ln y) + d_n \sin(\alpha_n \ln y)],$$

in which the coefficients  $\alpha_n$  do not approach zero as  $n$  approaches  $\infty$  and the series  $\sum c_n$  and  $\sum d_n$  are absolutely convergent. He proved that in any such representation the values  $A$ ,  $c_n$ ,  $d_n$  and  $\alpha_n$  are all uniquely determined by  $f(y)$ . On the other hand, if  $1 + \beta i$  were a root of  $\zeta(s)$ , he showed by a long and delicate argument that the function  $\frac{1 + \cos(\beta \ln y)}{y} \sum_{p < y} \ln p$ , where  $p$  ranges over primes, would have two distinct representations of the form  $L(y) + P(y)$ . Salient details of the computations involved are given on pp. 208–214 of Narkiewicz (2000).

Two other differences between the tools used by Hadamard and those employed by de la Vallée Poussin are worth noting.

First, although de la Vallée Poussin referred to Riemann's functional equation for the  $\zeta$ -function and observed that it could be used to define  $\zeta(s)$  throughout  $\mathbb{C} - \{1\}$ , he did not make use of that extension in his proof of the Prime Number Theorem. Instead, he noted that integration by parts yields

$$\int_0^1 \frac{dx}{(n+x)^s} = \frac{1}{(n+1)^s} + s \int_0^1 \frac{xdx}{(n+x)^{s+1}},$$

from which it follows that for  $\text{Re } s > 1$ ,

$$\zeta(s) = 1 + \sum_{n=1}^{\infty} \left( \int_0^1 \frac{dx}{(n+x)^s} - s \int_0^1 \frac{xdx}{(n+x)^{s+1}} \right).$$

But

$$\sum_{n=1}^{\infty} \int_0^1 \frac{dx}{(n+x)^s} = \int_1^{\infty} \frac{dx}{x^s} = \frac{1}{s-1},$$

so

$$(40) \quad \zeta(s) = 1 + \frac{1}{s-1} - s \sum_1^\infty \int_0^1 \frac{x dx}{(n+x)^{s+1}};$$

and since the sum on the right side of (40) converges absolutely for  $\text{Re } s > 0$ , that equation may be used to define  $\zeta(s)$  in that half-plane, except for a simple pole at  $s = 1$  with residue 1.

Second, in place of the integral  $\psi_2(x)$  employed by Hadamard, both in his proof that no  $s = 1 + \beta i$  is a root of  $\zeta(s)$  and in his proof that  $\pi(x) \sim x / \ln x$  de la Vallée Poussin made use of the integral

$$I_{u,v}(y) = \frac{1}{2\pi i} \int_{a-\infty i}^{a+\infty i} \frac{\zeta'(s)}{\zeta(s)} \frac{y^s ds}{(s-u)(s-v)},$$

in which neither  $u$  nor  $v$  are poles or zeros of  $\zeta(s)$ ,  $y$  is a real number greater than 1, and  $a$  is a real number greater than any of 1,  $\text{Re } u$  and  $\text{Re } v$ .

Replacing the fraction  $\zeta'(s)/\zeta(s)$  in the integrand of  $I_{u,v}$  by the expression on the right side of (39) and integrating the result, de la Vallée Poussin obtained the identity

$$(41) \quad I_{u,v}(y) = \frac{1}{u-v} \left( y^u \frac{\zeta'(u)}{\zeta(u)} - y^v \frac{\zeta'(v)}{\zeta(v)} \right) - \frac{y}{(u-1)(v-1)} + \sum_\rho \frac{y^\rho}{(u-\rho)(v-\rho)} + \sum_{m=1}^\infty \frac{y^{-2m}}{(2m+u)(2m+v)},$$

where  $\rho$  ranges over the roots of  $\zeta(s)$ , ordered according to increasing absolute value.

On the other hand, using (37) and (36), de la Vallée Poussin found that

$$(42) \quad I_{u,v}(y) = -\frac{1}{u-v} \left( y^u \sum_{n < y} \frac{\Lambda(n)}{n^u} - y^v \sum_{n < y} \frac{\Lambda(n)}{n^v} \right),$$

valid for all  $u, v$  whenever  $y > 1$ .

Equating those two expressions for  $I_{u,v}(y)$ , solving for the quantity in parentheses in (42), setting  $v = 0$ , dividing by  $y^u$  and letting  $u$  approach 1 yields

$$(43) \quad \sum_{n < y} \frac{\Lambda(n)}{n} - \frac{1}{y} \sum_{n < y} \Lambda(n) = \ln y - \lim_{u \rightarrow 1} \left( \frac{\zeta'(u)}{\zeta(u)} + \frac{u}{u-1} \right) + \frac{1}{y} \frac{\zeta'(0)}{\zeta(0)} - \sum_\rho \frac{y^{\rho-1}}{\rho(\rho-1)} - \sum_{m=1}^\infty \frac{y^{-2m-1}}{2m(2m+1)},$$

where the term  $\ln y$  is obtained by writing  $\frac{u}{1-u} y^{1-u}$  as  $\frac{u}{1-u} [1 + (y^{1-u} - 1)]$  and applying l'Hopital's Rule to  $\lim_{u \rightarrow 1} \frac{y^{1-u} - 1}{1-u}$ .



Recalling the definition of  $\Lambda(n)$ , the left member of (43) can be written as  $\sum_{p^m < y} \frac{\ln p}{p^m} - \frac{1}{y} \sum_{p^m < y} \ln p$ . In the right member the fraction  $\frac{u}{u-1}$  can be split up as  $\frac{1}{u-1} + 1$ , and both summations there approach zero as  $y$  approaches  $\infty$ . (In particular, since the real part of any root  $\rho$  of  $\zeta(s)$  has been shown to be less than one, the real part of  $\rho - 1$  must be negative. The terms of the first summation are thus dominated by those of the absolutely convergent series  $\sum_{\rho} \frac{1}{\rho(\rho-1)}$ , so that sum must converge *uniformly* to zero as  $y$  increases without bound.) Consequently, setting  $u = s$ , (43) takes the form

$$(44) \quad \sum_{p^m < y} \frac{\ln p}{p^m} - \frac{1}{y} \sum_{p^m < y} \ln p = \ln y - 1 - \lim_{s \rightarrow 1} \left[ \frac{\zeta'(s)}{\zeta(s)} + \frac{1}{s-1} \right] + f(y),$$

where  $f(y)$  approaches 0 as  $y$  approaches  $\infty$ .

De la Vallée Poussin next noted that the difference  $\sum_{p < y} \frac{\ln p}{p-1} - \sum_{p^m < y} \frac{\ln p}{p^m}$  approaches 0 as  $y$  approaches  $\infty$ , and proved that  $\frac{1}{y} \sum_{p^m < y} \ln p - \frac{1}{y} \sum_{p < y} \ln p$  does so as well. He also proved that

$$\lim_{s \rightarrow 1} \left[ \frac{\zeta'(s)}{\zeta(s)} + \frac{1}{s-1} \right] = \text{Euler's constant } \gamma,$$

defined as  $\gamma = \lim_{n \rightarrow \infty} (\sum_{k=1}^n 1/k - \ln n)$ .

Equation (44) can therefore be replaced by

$$(45) \quad \sum_{p < y} \frac{\ln p}{p-1} - \frac{1}{y} \sum_{p < y} \ln p = \ln y - 1 - \gamma + g(y),$$

where  $g(y)$  approaches 0 as  $y$  approaches  $\infty$ .

Equation (45) is the key to the final two steps in de la Vallée Poussin's derivation of the prime number theorem, namely

(i) showing that

$$(46) \quad \int_1^x \frac{1}{y} \sum_{p < y} \ln p \, dy = x(1 + h(x)), \text{ where } \lim_{x \rightarrow \infty} h(x) = 0,$$

and

(ii) showing that (46) implies that  $\lim_{x \rightarrow \infty} \frac{\theta(x)}{x} = 1$ .

To establish (46) de la Vallée Poussin integrated (45) from 1 to  $x$  and then multiplied each term by  $1/x$ , thus obtaining

$$(47) \quad \frac{1}{x} \int_1^x \sum_{p < y} \frac{\ln p}{p-1} \, dy - \frac{1}{x} \int_1^x \frac{1}{y} \sum_{p < y} \ln p \, dy = \frac{1}{x} \int_1^x \ln y \, dy - \gamma - 1 + j(x),$$

where  $j(x)$  approaches 0 as  $x$  approaches  $\infty$ . He then noted that the first term in the left member of (47) can be rewritten as

$$\begin{aligned}
 (48) \quad \frac{1}{x} \int_1^x \sum_{p < y} \frac{\ln p}{p-1} dy &= \frac{1}{x} \sum_{p < x} \frac{\ln p}{p-1} \int_p^x dy \\
 &= \sum_{p < x} \frac{\ln p}{p-1} - \frac{1}{x} \sum_{p < x} \frac{p}{p-1} \ln p \\
 &= \sum_{p < x} \frac{\ln p}{p-1} - \frac{1}{x} \sum_{p < x} \ln p + \frac{1}{x} \sum_{p < x} \frac{\ln p}{p-1}.
 \end{aligned}$$

The first two terms in the last member of (48) make up the left member of (45), with  $x$  in place of  $y$ . Moreover, as Franz Mertens had shown in 1874, it follows from Chebyshev's upper bound (31) for  $\psi(x)$  that  $\sum_{p < y} \frac{\ln p}{p-1}$  is less than a constant multiple of  $\ln y$ . Consequently,

$$\frac{1}{x} \int_1^x \sum_{p < y} \frac{\ln p}{p-1} dy = \ln x - \gamma - 1 + k(x),$$

where  $k(x)$  approaches 0 as  $x$  approaches  $\infty$ . Substituting the expression on the right of this equation for the first integral in (47) and carrying out the integration in the right member of (47) then yields

$$\ln x - \gamma - 1 + k(x) - \frac{1}{x} \int_1^x \frac{1}{y} \sum_{p < y} \ln p dy = \frac{1}{x} [x \ln x - x] - \gamma - 1 + j(x),$$

which after cancellation of terms common to both members is (i), with  $h(x) = k(x) - j(x)$ .

To show that  $\lim_{x \rightarrow \infty} \frac{\theta(x)}{x} = 1$ , take  $\epsilon > 0$ , and use (46) to evaluate

$$\begin{aligned}
 (49) \quad \int_x^{(1+\epsilon)x} \frac{1}{y} \sum_{p < y} \ln p dy &= \int_1^{(1+\epsilon)x} \frac{1}{y} \sum_{p < y} \ln p dy - \int_1^x \frac{1}{y} \sum_{p < y} \ln p dy \\
 &= (1+\epsilon)x [1 + h((1+\epsilon)x)] - x [1 + h(x)] \\
 &= \epsilon x - h(x)x + (1+\epsilon)xh((1+\epsilon)x).
 \end{aligned}$$

Dividing (49) by  $\epsilon x$  gives

$$\frac{1}{\epsilon x} \int_x^{(1+\epsilon)x} \frac{1}{y} \sum_{p < y} \ln p dy = 1 + \frac{(1+\epsilon)h((1+\epsilon)x) - h(x)}{\epsilon},$$

and since  $\theta(y) = \sum_{p \leq y} \ln p$ , upper and lower bounds on the integrand give

$$\frac{\theta(x)}{x} \frac{\ln(1 + \epsilon)}{\epsilon} \leq 1 + \frac{(1 + \epsilon)h((1 + \epsilon)x) - h(x)}{\epsilon} \leq \frac{\theta((1 + \epsilon)x)}{x} \frac{\ln(1 + \epsilon)}{\epsilon},$$

that is,

$$(50) \quad \frac{\theta(x)}{x} \leq \frac{\epsilon}{\ln(1 + \epsilon)} + \frac{(1 + \epsilon)h((1 + \epsilon)x) - h(x)}{\ln(1 + \epsilon)} \leq \frac{\theta((1 + \epsilon)x)}{x}.$$

Since both  $h(x)$  and  $h((1 + \epsilon)x)$  approach zero as  $x$  approaches  $\infty$ , (50) shows that

$$\limsup_{x \rightarrow \infty} \frac{\theta(x)}{x} \leq \frac{\epsilon}{\ln(1 + \epsilon)} \leq \liminf_{x \rightarrow \infty} \frac{\theta((1 + \epsilon)x)}{x};$$

and by replacing  $x$  in the latter inequality by  $\frac{x}{1 + \epsilon}$ , it follows that

$$\frac{\epsilon}{\ln(1 + \epsilon)} \leq \liminf_{x \rightarrow \infty} \frac{\theta(x)}{x} (1 + \epsilon).$$

By L'Hopital's rule,  $\lim_{\epsilon \rightarrow 0} \frac{\epsilon}{\ln(1 + \epsilon)} = 1$ , so finally,  $\lim_{x \rightarrow \infty} \frac{\theta(x)}{x} = 1$ .

## 10.4 Later refinements

In the wake of Hadamard's and de la Vallée Poussin's proofs, various simplifications, generalizations, and improvements of their arguments were developed, by Edmund Landau, Franz Mertens, de la Vallée Poussin himself (who in 1899 obtained the error bound  $\pi(x) - \text{li}(x) \leq Kxe^{-c\sqrt{\ln x}}$  for some positive constants  $K$  and  $c$  — a result not bettered for a quarter of a century thereafter<sup>8</sup>), and others. Modern proofs of the Prime Number Theorem that are descendants of the classical ones incorporate many of those refinements and also make use of other tools such as the Riemann-Lebesgue lemma, integral transforms, and Tauberian theorems (discussed further below).

The proof of the Prime Number Theorem given in Jameson (2003) may be taken as an exemplar of such proofs. Here, in outline, is its structure:

1. The comparison test is used to prove Dirichlet's result that if

$$\sum_{n=1}^{\infty} |a_n/n^\alpha|$$

<sup>8</sup>Cf. the discussion in Bateman and Diamond (1996), p. 736.

converges for some real  $\alpha$ , then the corresponding Dirichlet series  $\sum_{n=1}^{\infty} a_n/n^s$  converges absolutely for all  $s = \sigma + it$  with  $\sigma \geq \alpha$ . In particular, if  $|a_n| \leq 1$  for all  $n$ , then  $\sum_{n=1}^{\infty} a_n/n^s$  converges absolutely when  $\operatorname{Re} s > 1$ , so  $\sum_{n=1}^{\infty} 1/n^s$  can be used to define  $\zeta(s)$  in that region.

2. It is shown that if  $F(s) = \sum_{n=1}^{\infty} a_n/n^s$  converges whenever  $\operatorname{Re} s > c$ , then  $F(s)$  is analytic at all points  $s$  with  $\operatorname{Re} s > c$  and  $F'(s) = -\sum_{n=1}^{\infty} a_n \ln n/n^s$ . In particular,  $\zeta'(s) = -\sum_{n=1}^{\infty} \ln n/n^s$  for  $\operatorname{Re} s > 1$ .
3. After defining the Möbius function  $\mu(n)$  by  $\mu(1) = 1$ ;  $\mu(n) = 0$  if  $p^2$  divides  $n$  for some prime  $p$ ; and  $\mu(n) = (-1)^k$  if  $n = p_1 p_2 \dots p_k$  for distinct primes  $p_1, p_2, \dots, p_k$ , the generalized Euler product identity (stated in footnote 3 above) is invoked and inverted to obtain

$$\frac{1}{\sum_{n=1}^{\infty} f(n)} = \prod_{p \text{ prime}} (1 - f(p)) = \sum_{n=1}^{\infty} \mu(n) f(n).$$

Consequently  $\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$  for  $\operatorname{Re} s > 1$ .

4. It is verified that for  $|s| < 1$  the series  $\sum_{n=1}^{\infty} s^n/n$  defines an analytic function  $h(s)$  that is a logarithm of  $\frac{1}{1-s}$ . The Euler product identity then implies that the logarithm of  $\zeta(s)$  should be given by  $H(s) = \sum_{p \text{ prime}} h(p^{-s}) = \sum_{p \text{ prime}} \sum_{n=1}^{\infty} \frac{1}{np^{ns}}$ . It is proved that that double series converges when  $\operatorname{Re} s >$

1 and that its sum is equal to that of the series  $\sum_{n=1}^{\infty} \frac{c(n)}{n^s}$ , where  $c(n) = 1/m$  if  $n = p^m$  for some prime  $p$ , and  $c(n) = 0$  otherwise.

5. Consequently,  $H'(s) = \frac{\zeta'(s)}{\zeta(s)} = -\sum_{n=1}^{\infty} \frac{c(n) \ln n}{n^s} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$ , where  $\Lambda(n)$  is the von Mangoldt function.

6. Abel summation is used to show that for any sequence  $a(n)$  and corresponding summation function  $A(x) = \sum_{n \leq x} a(n)$ , if  $X > 1$  then

$$\sum_{n \leq X} \frac{a(n)}{n^s} = \frac{A(x)}{X^s} + s \int_1^X \frac{A(x)}{x^{s+1}} dx.$$

Furthermore, if  $s \neq 0$ ,  $A(x)/x^s$  approaches 0 as  $x$  approaches  $\infty$  and the Dirichlet series  $\sum_{n=1}^{\infty} \frac{a(n)}{n^s}$  converges, then the Dirichlet integral

$$s \int_1^{\infty} \frac{A(x)}{x^{s+1}} dx$$

converges to the same value. Since  $\psi(x)$  is the summation function for  $\Lambda(n)$  and  $\psi(x)/x^s$  approaches 0 as  $x$  approaches  $\infty$  if  $\operatorname{Re} s > 1$ , it follows from 5 that

$$(51) \quad \frac{\zeta'(s)}{\zeta(s)} = -s \int_1^\infty \frac{\psi(x)}{x^{s+1}} dx.$$

7. A simplification due to Edmund Landau is used to extend the domain of definition of  $\zeta(s)$  without using the functional equation for  $\zeta(s)$ .<sup>9</sup> Specifically, for  $\text{Re } s > 0$ ,  $\zeta(s)$  may be defined by

$$(52) \quad \zeta(s) = \frac{1}{s-1} + 1 - s \int_1^\infty \frac{x - [x]}{x^{s+1}} dx,$$

where  $[x]$  denotes the greatest integer not exceeding  $x$ , and differentiation under the integral sign shows that  $\zeta(s) - \frac{1}{s-1}$  is analytic at  $s = 1$ , so  $\zeta(s)$  has a simple pole there. Furthermore,

$$\lim_{s \rightarrow 1} \left( \zeta(s) - \frac{1}{s-1} \right) = 1 - \int_1^\infty \frac{x - [x]}{x^2} dx = \text{Euler's constant } \gamma.$$

8. Hence  $\zeta(s)$ ,  $\frac{1}{\zeta(s)}$  and  $\frac{\zeta'(s)}{\zeta(s)}$  are represented by Laurent series of the forms

$$\begin{aligned} \zeta(s) &= \frac{1}{s-1} + \gamma + \sum_{n=1}^\infty c_n (s-1)^n \\ \frac{1}{\zeta(s)} &= (s-1) - \gamma(s-1)^2 + \dots \quad \text{and} \\ \frac{\zeta'(s)}{\zeta(s)} &= -\frac{1}{s-1} + a_0 + a_1(s-1) + \dots, \end{aligned}$$

all converging in some disk with center  $s = 1$ .

9. Since

$$\left| \int_N^\infty \frac{x - [x]}{x^{s+1}} dx \right| \leq \int_N^\infty \frac{1}{x^{\sigma+1}} dx = \frac{1}{\sigma N^\sigma},$$

Euler's summation formula for finite sums<sup>10</sup> yields

$$\zeta(s) = \sum_{n=1}^N \frac{1}{n^s} + \frac{N^{1-s}}{s-1} - s \int_N^\infty \frac{x - [x]}{x^{s+1}} dx \leq \sum_{n=1}^N \frac{1}{n^s} + \frac{N^{1-s}}{s-1} + \frac{|s|}{\sigma N^\sigma}.$$

Straightforward calculations using the last inequality then show that when  $\sigma \geq 1$  and  $t \geq 2$ ,  $|\zeta(\sigma + it)| \leq \ln t + 4$  and  $|\zeta'(\sigma + it)| \leq \frac{1}{2}(\ln t + 3)^2$ .

<sup>9</sup>According to Bateman and Diamond (1996), p. 737, Landau was the first to prove the PNT without recourse to that functional equation.

<sup>10</sup> $\sum_{n=2}^N f(n) = \int_1^N f(x) dx + \int_1^N (x - [x])f'(x) dx.$

Those inequalities can in turn be used to show that  $1/|\zeta(\sigma + it)| \leq 4(\ln t + 5)^7$  for  $\sigma > 1$  and  $t \geq 2$ . (For details see Jameson 2003, pp. 108–109.)

10. The proof that  $\zeta(s) \neq 0$  when  $\operatorname{Re} s = 1$  is carried out by a simplified argument based on Hadamard's approach and later ideas of de la Vallée Poussin and Franz Mertens. It rests on the trigonometric identity  $0 \leq 2(1 + \cos \theta)^2 = 3 + 4 \cos \theta + \cos(2\theta)$ . For if a Dirichlet series  $\sum_{n=1}^{\infty} \frac{a(n)}{n^s}$  with positive real coefficients  $a(n)$  converges for  $\sigma > \sigma_0$  to a function  $f(s)$ , then for such  $\sigma$

$$3f(\sigma) + 4f(\sigma + it) + f(\sigma + 2it) = \sum_{n=1}^{\infty} \frac{a(n)}{n^{\sigma}} (3 + 4n^{-it} + n^{-2it})$$

has real part

$$\sum_{n=1}^{\infty} \frac{a(n)}{n^{\sigma}} \operatorname{Re}(3 + 4n^{-it} + n^{-2it}) = \sum_{n=1}^{\infty} \frac{a(n)}{n^{\sigma}} (3 + 4 \cos(t \ln n) + \cos(2t \ln n)) \geq 0.$$

In particular, by 4. the Dirichlet series for  $\ln(\zeta(s))$  has positive coefficients and converges when  $\sigma > 1$ , so for such  $\sigma$  and all  $t$ ,

$$\begin{aligned} 3 \ln(\zeta(\sigma)) + 4 \operatorname{Re} \zeta(\sigma + it) + \operatorname{Re} \zeta(\sigma + 2it) \\ = \ln [\zeta(\sigma)^3 |\zeta(\sigma + it)|^4 |\zeta(\sigma + 2it)|] \geq 0 \end{aligned}$$

(because  $\operatorname{Re} \ln z = \ln |z|$ ), that is

$$(53) \quad \zeta(\sigma)^3 |\zeta(\sigma + it)|^4 |\zeta(\sigma + 2it)| \geq 1.$$

Suppose then that  $\zeta(1 + it_0) = 0$  for some  $t_0 \neq 0$ . Then

$$\begin{aligned} (54) \quad \zeta(\sigma)^3 |\zeta(\sigma + it_0)|^4 |\zeta(\sigma + 2it_0)| \\ = [(\sigma - 1)\zeta(\sigma)]^3 \left( \frac{|\zeta(\sigma + it_0)|}{\sigma - 1} \right)^4 (\sigma - 1) |\zeta(\sigma + 2it_0)|. \end{aligned}$$

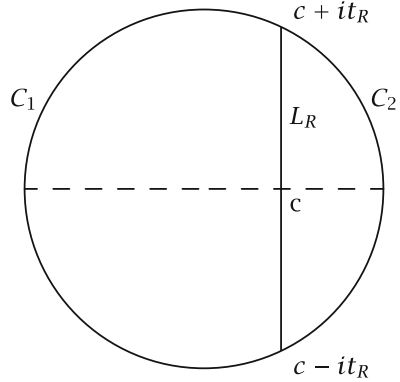
But as  $\sigma$  approaches  $1^+$ ,  $(\sigma - 1)\zeta(\sigma)$  approaches 1, while  $\frac{\zeta(\sigma + it_0)}{\sigma - 1}$  approaches  $\zeta'(1 + it_0)$  and  $\zeta(\sigma + 2it_0)$  approaches  $\zeta(1 + 2it_0)$ . So by (54), the product  $\zeta(\sigma)^3 |\zeta(\sigma + it_0)|^4 |\zeta(\sigma + 2it_0)|$  approaches 0, contrary to (53).

11. Cauchy's integral theorem is used to evaluate three infinite contour integrals, namely

$$(55) \quad \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^s}{s^2} ds = S(x) \ln x \quad \text{for } x > 0 \text{ and } c > 0$$

$$(56) \quad \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^s}{s(s-1)} ds = (x-1)S(x) \quad \text{for } x > 0 \text{ and } c > 1$$

**Fig. 10.2** Contour used to evaluate the integrals (53) through (55). Adapted from *The Prime Number Theorem*, p. 116, by G.J.O. Jameson. Reprinted with the permission of Cambridge University Press



$$(57) \quad \frac{1}{2\pi i} \int_{c-it_R}^{c+it_R} \frac{x^{s-1}}{s(s-1)} f(s) ds = \int_1^x \frac{A(y)}{y^2} dy \quad \text{for } x > 1 \text{ and } c > 1,$$

where  $S(x)$  denotes the step-function equal to 0 for  $x < 1$  and 1 for  $x \geq 1$  and  $A(x) = \sum_{n \leq x} a(n)$  is the summation function corresponding to a Dirichlet series  $\sum_{n=1}^{\infty} a(n)/n^s$  that converges absolutely to  $f(s)$  whenever  $\text{Re } s > 1$ .

In all three cases the infinite contour integrals are obtained as the limit, as  $R$  approaches  $\infty$ , of integrals taken along the paths  $\Gamma_1 = C_1 \cup L_R$  or  $\Gamma_2 = C_2 \cup L_R$ , where  $C_1$  and  $C_2$  are the arcs of the circle of radius  $R$  centered at  $s = 0$  that lie, respectively, to the left and right of the vertical line segment  $L_R$  with endpoints  $c - t_R$  and  $c + t_R$  on the circle. (See Figure 10.2.)

The integrand of (55) may be rewritten as

$$\frac{x^s}{s^2} = \frac{e^{ls}}{s^2} = \frac{1}{s^2} \sum_{n=0}^{\infty} \frac{l^n s^n}{n!},$$

where  $l = \ln x$ . Since the exponential series converges uniformly on any closed interval, the integration may then be carried out term by term. For  $x \geq 1$  the path  $\Gamma_1$  is taken, enclosing the pole at  $s = 0$  of  $x^s/s^2$ , where the residue is  $l$ . For  $0 < x < 1$  the path  $\Gamma_2$  is taken instead, which encloses no poles of  $x^s/s^2$ ; so that integral is 0. In the first case,  $|s| = R$  and  $|x^s| = x^\sigma \leq x^c$  on  $C_1$ , so the integral along  $C_1$  has absolute value less than or equal to  $\frac{1}{2\pi} \frac{x^c}{R^2} 2\pi R = \frac{x^c}{R}$ , which approaches 0 as  $R$  approaches  $\infty$ . In the second case,  $x^\sigma \leq x^c$  on  $C_2$  as well, since for  $0 < x < 1$ ,  $x^\sigma$  decreases as  $\sigma$  increases. The integral along  $C_2$  thus also approaches 0 as  $R$  approaches  $\infty$ . The proof of (56) is similar, using  $|s(s-1)| \geq R(R-1)$  and writing  $\frac{x^s}{s(s-1)}$  as  $\frac{x^s}{s-1} - \frac{x^s}{s}$ .

To obtain (57), first note that  $x^s f(s) = \sum_{n \leq x} a(n) \left(\frac{x}{n}\right)^s + \sum_{n > x} a(n) \left(\frac{x}{n}\right)^s$ . The first sum is finite, and so can be integrated term by term. By (56), the result is  $\sum_{n \leq x} S\left(\frac{x}{n}\right) \left(\frac{x}{n} - 1\right) = \sum_{n \leq x} a(n) \left(\frac{x}{n} - 1\right)$ . The second term,

on the other hand, is an analytic function of  $s$  that for  $\operatorname{Re} s \geq c$  is bounded by  $\sum_{n>x} |a(n)|(x/n)^c$ . By Cauchy's theorem, its integral over  $\Gamma_2$  is zero, and by the same reasoning as before, its integral over  $C_2$  tends to zero as  $R$  approaches  $\infty$ . Consequently, after dividing by  $x$ ,

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s-1}}{s(s-1)} f(s) ds = \sum_{n \leq x} a(n) \left( \frac{1}{n} - \frac{1}{x} \right).$$

By Abel summation applied to the function  $1/y$ , the last sum is equal to  $\int_1^x A(y)/y^2 dy$ .

12. Finally, the Prime Number Theorem in the form  $\lim_{x \rightarrow \infty} \frac{\psi(x)}{x} = 1$  is obtained as a special case of the following much more general result.

**Theorem:** *Suppose the function  $f(s)$  is analytic throughout the region  $\operatorname{Re} s \geq 1$ , except perhaps at  $s = 1$ , and satisfies the following conditions:*

(C1)  $f(s) = \sum_{n=1}^{\infty} \frac{a(n)}{n^s}$  converges absolutely when  $\operatorname{Re} s > 1$ .

(C2)  $f(s) = \frac{\alpha}{s-1} + \alpha_0 + (s-1)h(s)$ , where  $h$  is differentiable at  $s = 1$ .

(C3) *There is a function  $P(t)$  such that  $|f(\sigma+it)| \leq P(t)$  for  $\sigma \geq 1$  and  $t \geq t_0 \geq 1$ , and  $\int_1^{\infty} P(t)/t^2 dt$  is convergent.*

*Then  $\int_1^{\infty} \frac{A(x) - \alpha x}{x^2} dx$  converges to  $\alpha_0 - \alpha$  where  $A(x) = \sum_{n \leq x} a(n)$ .*

*If, furthermore,  $A(x)$  is increasing and non-negative, then  $\lim_{x \rightarrow \infty} \frac{A(x)}{x} = \alpha$ .*

To prove the first claim, note that  $\frac{1}{s-1} = \frac{s}{s-1} - 1$ , so  $(s-1)h(s) = f(s) - \alpha \frac{s}{s-1} - (\alpha_0 - \alpha)$ , and  $|(s-1)h(s)| \leq P(t) + |\alpha| + |\alpha_0|$  when  $|\sigma| \geq 1$  and  $|t| \geq t_0$ . Then for  $x > 1$  and  $c > 1$ , (55), (56), and (57) give

$$\begin{aligned} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s-1}h(s)}{s} ds &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s-1}}{s(s-1)} f(s) ds \\ &\quad - \frac{\alpha}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s-1}}{(s-1)^2} ds - \frac{\alpha_0 - \alpha}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s-1}}{s(s-1)} ds \\ &= \int_1^x \frac{A(y)}{y^2} dy - \alpha \ln x - (\alpha_0 - \alpha) \left( 1 - \frac{1}{x} \right) \\ &= \int_1^x \frac{A(y) - \alpha y}{y^2} dy - (\alpha_0 - \alpha) \left( 1 - \frac{1}{x} \right). \end{aligned}$$

Careful examination (see Jameson 2003, p. 121) shows that the result just obtained also holds for  $c = 1$ . That is,

$$\frac{1}{2\pi i} \int_{1-i\infty}^{1+i\infty} \frac{x^{s-1}h(s)}{s} ds = \int_1^x \frac{A(y) - \alpha y}{y^2} dy - (\alpha_0 - \alpha) \left( 1 - \frac{1}{x} \right).$$



The path of integration for the contour integral on the left is the vertical line  $\operatorname{Re} s = 1$ , where  $s = 1 + it$ , so that integral may be rewritten as

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{x^{it} h(1 + it)}{1 + it} dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{it \ln x} h(1 + it)}{1 + it} dt$$

(which converges absolutely since  $\left| \frac{h(s)}{s} \right| \leq \frac{P(t) + |\alpha| + |\alpha_0|}{t^2}$ ). That the latter integral approaches 0 as  $x$  approaches  $\infty$  then follows from the Riemann-Lebesgue lemma, which states that

If  $\phi(t)$  is a continuously differentiable function from  $\mathbb{R}$  to  $\mathbb{C}$  and  $\int_{-\infty}^{+\infty} |\phi(t)| dt$  converges, then the integral

$$\int_{-\infty}^{+\infty} e^{i\lambda t} \phi(t) dt$$

approaches 0 as  $\lambda$  approaches  $\infty$ .

The convergence of  $\int_1^\infty (A(x) - \alpha x)/x^2 dx$  implies the remaining claim (that if  $A(x)$  is increasing and non-negative, then  $\lim_{x \rightarrow \infty} \frac{A(x)}{x} = \alpha$ ). For such convergence means that for any  $\delta > 0$ , there is an  $R$  such that

$$\left| \int_{x_0}^{x_1} \frac{A(x) - x}{x^2} dx \right| < \delta \quad \text{whenever } x_1 > x_0 > R.$$

Then for  $0 < \delta < 1/2$ , the assumption either that  $A(x_0) > (1 + \delta)x_0$  for some  $x_0 > R$  (and hence for all  $x \geq x_0$ , since  $A(x)$  is increasing), or that  $A(x_0) < (1 - \delta)x_0$  for some  $x_0 \geq 2R$ , leads to a contradiction. (See Jameson 2003, p. 131.)

The Prime Number Theorem follows from the Theorem by taking  $f(s) = \frac{\zeta'(s)}{\zeta(s)}$ : For the result of step 5. shows that  $f(s) = \sum_{n=1}^\infty \frac{\Lambda(n)}{n^s}$  (whence (C1) is satisfied);  $\psi(x) = \sum_{n \leq x} \Lambda(n)$  is increasing; (C2) holds by the third equation in step 8., with  $\alpha = 1$ ; and the inequalities found in step 9. show that  $|f(s)| = |f(\sigma + it)| < P(t) = 2(\ln t + 5)^9$ . Since  $\int_1^\infty P(t)/t^2 dt$  converges, (C3) is thus also satisfied.

## 10.5 Tauberian theorems and Newman's proof

The Theorem stated in the previous section is an example of a so-called Tauberian theorem, broadly defined (as in Edwards 1974, p. 279) as a theorem that “permits a conclusion about one kind of average [in this case,  $A(x)/x$ ] given information about another kind of average [here, the integral from 1 to  $\infty$  of  $(A(x) - \alpha x)/x^2$ ].”

The Prime Number Theorem was first deduced from a Tauberian theorem by Edmund Landau in his paper Landau (1908).<sup>11</sup> Seven years later Hardy and Littlewood gave another such proof (Hardy and Littlewood 1915), and in the 1930s Tauberian theorems based on Norbert Wiener's methods in the theory of Fourier transforms were employed to deduce the Prime Number Theorem from the non-vanishing of  $\zeta(s)$  on the line  $\operatorname{Re} s = 1$ . In particular, the Prime Number Theorem was so deduced using

**The Wiener-Ikehara Theorem:** Suppose that  $f$  is a non-decreasing real-valued function on  $[1, \infty)$  for which  $\int_1^\infty |f(u)|u^{-\sigma-1} du$  converges for all real  $\sigma > 1$ . If in addition  $\int_1^\infty f(u)u^{-s-1} du = \alpha/(s-1) + g(s)$  for some real  $\alpha$ , where  $g(s)$  is the restriction to  $\operatorname{Re} s > 1$  of a function that is continuous on  $\operatorname{Re} s \geq 1$ , then  $\lim_{u \rightarrow \infty} f(u)/u = \alpha$ .

But proofs of the Wiener-Ikehara Theorem are themselves difficult.

In 1935 A.E. Ingham proved another Tauberian theorem that related the Laplace transform  $\int_0^\infty f(t)e^{-zt} dt$  of a function  $f(t)$  defined on  $[0, \infty)$  to the integral of  $f(t)$  itself over that interval.<sup>12</sup> But that proof, too, was complicated, and was also based on results from Fourier analysis (Ingham 1935). In 1980, however, D.J. Newman found a way to prove a variant of Ingham's theorem and to derive the Prime Number Theorem from it without resort either to Fourier techniques or to contour integrals over infinite paths.

Newman's original proof (Newman 1980) was couched in terms of Dirichlet series: he proved that if  $\sum_{n=1}^\infty a_n n^{-s}$  converges to an analytic function  $f(s)$  for all  $s$  with  $\operatorname{Re} s > 1$ , if  $|a_n| \leq 1$  for every  $n$ , and if  $f(s)$  is also analytic when  $\operatorname{Re} s = 1$ , then  $\sum_{n=1}^\infty a_n n^{-s}$  converges for all  $s$  with  $\operatorname{Re} s = 1$ . Subsequent refinements of that proof, as given in Korevaar (1982), Zagier (1997), chapter 7 of Lax and Zalzman (2012), and Jameson (2003), recast it as an alternative proof of Ingham's Tauberian theorem. The formulation of that result given in Lax and Zalzman (2012) reads:

*Let  $f$  be a bounded measurable function on  $[0, \infty)$ . Suppose that the Laplace transform*

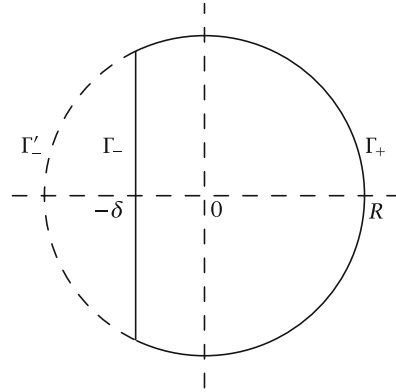
$$g(z) = \int_0^\infty f(t)e^{-zt} dt,$$

*which is defined and analytic on the open half plane  $\{z : \operatorname{Re} z > 0\}$ , extends analytically to an open set containing  $\{z : \operatorname{Re} z \geq 0\}$ . Then the improper integral  $\int_0^\infty f(t) dt = \lim_{T \rightarrow \infty} \int_0^T f(t) dt$  converges and coincides with  $g(0)$ , the value of the analytic extension of  $g$  at  $z = 0$ .*

<sup>11</sup>Discussed in detail in Narkiewicz (2000), pp. 298–302.

<sup>12</sup>Ingham's theorem may alternatively be stated in terms of the Mellin transform  $\int_1^\infty f(t)t^{-s} dt$ . See, e.g., Korevaar (1982) or Jameson (2003), pp. 124–129.

**Fig. 10.3** Contours used in Newman's proof



The proof there proceeds as follows:

Say  $|f(t)| \leq M$  for all  $t \geq 0$ . For  $T > 0$  let  $g_T(z)$  be the entire function defined by  $\int_0^T f(t)e^{-zt} dt$ . The theorem then asserts that  $\lim_{T \rightarrow \infty} |g(0) - g_T(0)| = 0$ .

To establish that, choose  $R > 0$  and  $\delta(R)$  sufficiently small that  $g$  is analytic throughout the region  $D = \{z : |z| \leq R \text{ and } \operatorname{Re} z \geq -\delta(R)\}$ . Let  $\Gamma$  be the boundary of  $D$  (shown as the solid curve in Figure 10.3), traversed counterclockwise, and consider

$$(58) \quad \frac{1}{2\pi i} \int_{\Gamma} [g(z) - g_T(z)] e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{1}{z} dz.$$

By Cauchy's theorem, the value of (58) is  $g(0) - g_T(0)$ . Writing  $\sigma = \operatorname{Re} z$ , if  $\sigma > 0$  then

$$(59) \quad |g(z) - g_T(z)| = \left| \int_T^{\infty} f(t)e^{-zt} dt \right| \leq M \int_T^{\infty} e^{-\sigma t} dt = \frac{Me^{-\sigma T}}{\sigma}.$$

Also, when  $|z| = R$ ,

$$(60) \quad \left| e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{1}{z} \right| = e^{\sigma T} \left| \left(\frac{1}{z} + \frac{z}{|z|^2}\right) \right| = e^{\sigma T} \left| \frac{1}{z} + \frac{1}{\bar{z}} \right| = e^{\sigma T} \frac{2|\sigma|}{R^2}.$$

Now let  $\Gamma_+$  be the semicircle  $\Gamma \cap \{\operatorname{Re} z > 0\}$ , let  $\Gamma_-$  denote  $\Gamma \cap \{\operatorname{Re} z < 0\}$ , and let  $\Gamma'_-$  be the semicircle  $\{z : |z| = R \text{ and } \operatorname{Re} z < 0\}$ . By (59) and (60), for  $z \in \Gamma_+$ , the absolute value of the integrand in (58) is bounded by  $2M/R^2$ , so

$$(61) \quad \left| \frac{1}{2\pi i} \int_{\Gamma_+} [g(z) - g_T(z)] e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{1}{z} dz \right| \leq \frac{M}{R}.$$

For  $z \in \Gamma_-$ , first consider  $g_T$ . As an entire function, its integral over  $\Gamma_-$  is the same as its integral over  $\Gamma'_-$ . Since for  $\sigma < 0$ ,

$$(62) \quad |g_T(z)| = \left| \int_0^T f(t)e^{-zt} dt \right| \leq M \int_{-\infty}^T e^{-\sigma t} dt = \frac{Me^{-\sigma T}}{|\sigma|},$$

it follows from (60) that

$$(63) \quad \left| \frac{1}{2\pi i} \int_{\Gamma'_-} g_T(z)e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{1}{z} dz \right| \leq \frac{M}{R}.$$

Next consider  $g(z)$ . It is analytic on  $\Gamma_-$ , so the quantity  $\left| g(z) \left(1 + \frac{z^2}{R^2}\right) \frac{1}{z} \right|$  is bounded on  $\Gamma_-$  by some constant  $K$  (whose value depends on  $\delta$  and  $R$ ). Likewise,  $e^{zT}$  is bounded on  $\Gamma_-$ , and converges uniformly to 0 on compact subsets of  $\{\operatorname{Re} z < 0\}$  as  $T$  approaches  $\infty$ . Consequently,

$$(64) \quad \lim_{T \rightarrow \infty} \left| \frac{1}{2\pi i} \int_{\Gamma_-} g(z)e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{1}{z} dz \right| = 0.$$

Recalling that  $g(0) - g_T(0)$  is given by the integral in (58), it follows from (61), (63), and (64) that

$$\limsup_{T \rightarrow \infty} |g(0) - g_T(0)| \leq \frac{2M}{R}.$$

Since that holds for arbitrarily large values of  $R$ ,  $\lim_{T \rightarrow \infty} |g(0) - g_T(0)| = 0$ .

The Prime Number Theorem in the form  $\lim_{x \rightarrow \infty} \psi(x)/x = 1$  is then deduced, as in the previous section, from the convergence of the improper integral  $\int_1^\infty [\psi(x) - x]/x^2 dx$ . That, in turn, follows from Ingham's Tauberian theorem by setting  $x = e^t$  and taking  $f(t) = \psi(e^t)e^{-t} - 1$ . For, by equation (51) above,

$$-\frac{\zeta'(s)}{\zeta(s)} = s \int_1^\infty \frac{\psi(x)}{x^{s+1}} dx = s \int_0^\infty \psi(e^t)e^{-st} dt,$$

so

$$\begin{aligned} g(s) &= \int_0^\infty f(t)e^{-st} dt = \int_0^\infty [\psi(e^t)e^{-t} - 1]e^{-st} dt \\ &= \int_0^\infty \psi(e^t)e^{-(s+1)t} dt - \int_0^\infty e^{-st} dt \\ &= \frac{1}{s+1} \left[ -\frac{\zeta'(s+1)}{\zeta(s+1)} \right] - \frac{1}{s} = \frac{1}{s+1} \left[ -\frac{\zeta'(s+1)}{\zeta(s+1)} - \frac{1}{s} - 1 \right]. \end{aligned}$$

Chebyshev's upper bound for  $\psi(x)$  shows that  $f(t)$  is bounded, and the Laurent series for  $\zeta'(s)/\zeta(s)$  given in step 8 of the previous section shows that the expression in the last member of the equation above can be extended to an analytic function on the half plane  $\operatorname{Re} s \geq 0$ . Hence the hypotheses of Ingham's theorem are satisfied.

## 10.6 Elementary proofs

In 1909 Edmund Landau published an influential handbook (Landau 1909) that "presented in accessible form nearly everything that was then known about the distribution of prime numbers" (Bateman and Diamond 1996, p. 737). It popularized use of the  $O$ -notation<sup>13</sup> in statements concerning growth rates of functions, and drew attention to the power of complex-analytic methods in number theory.

In particular, in contrast to the elementary methods of Chebyshev, those of complex analysis had yielded the Prime Number Theorem. The question thus arose: Were such methods *essential* to the proof of that theorem?

Many leading number theorists came to believe that they were. G.H. Hardy, for example, in an address to the Mathematical Society of Copenhagen in 1921, declared

No elementary proof of the prime number theorem is known, and one may ask whether it is reasonable to expect one. [For] ... we know that ... theorem is roughly equivalent to ... the theorem that Riemann's zeta function has no roots on a certain line.<sup>14</sup> A proof of such a theorem, not fundamentally dependent on the theory of functions, [thus] seems to me extraordinarily unlikely." (Quoted from Goldfeld 2004.)

In 1948, however, Atle Selberg and Paul Erdős, independently but each using results of the other, found ways to prove the Prime Number Theorem without reference to the  $\zeta$ -function or complex variables and without resort to methods of Fourier analysis. Their proofs, however, were 'elementary' only in that technical sense. Indeed, in Edwards (1974) Harold Edwards expresses the widely shared opinion that "Since 1949 many variations, extensions and refinements of [Selberg's and Erdős's] elementary proof[s] have been given, but none of them seems very straightforward or natural, nor does any of them give much insight into the theorem." Furthermore, as noted in Jameson (2003), p. 207, no elementary proof so far devised has given error estimates for the approximation of  $\pi(x)$  by  $\operatorname{li}(x)$  that are "as strong

<sup>13</sup>Whereby  $f(x) = O(g(x))$  for  $x > x_1 \geq x_0$  means that  $f$  is eventually dominated by  $g$ , that is, that  $f$  and  $g$  are both defined for  $x > x_0$ ,  $g(x) > 0$  for  $x > x_0$ , and there is a constant  $K$  such that  $|f(x)| \leq Kg(x)$  for all  $x > x_1$ .

<sup>14</sup>For as noted in the preceding section, the Wiener-Ikehara Theorem implies that the Prime Number Theorem follows from the absence of zeroes of the  $\zeta$ -function on the line  $\operatorname{Re} s = 1$ , a fact that is also implied by the Prime Number Theorem. (See, for example, Diamond 1982, pp. 572–573.)

as that of de la Vallée Poussin.” The interest in such proofs would thus seem to stem primarily from concern for purity of method.

The basis of most elementary proofs of the Prime Number Theorem is a growth estimate for a logarithmic summation found by Selberg. In one form it states that for  $x > 1$

$$(65) \quad \sum_{p \leq x} (\ln p)^2 + \sum_{pq \leq x} \ln p \ln q = 2x \ln x + O(x),$$

where  $p$  and  $q$  denote prime numbers. Other statements equivalent to (65), shown to be so in Diamond (1982), p. 566 and Jameson (2003), pp. 214–215, are

$$(66) \quad \theta(x) \ln x + \sum_{p \leq x} \theta(x/p) \ln p = 2x \ln x + O(x)$$

$$(67) \quad \sum_{n \leq x} [\Lambda(n) \ln n + (\Lambda * \Lambda)(n)] = 2x \ln x + O(x)$$

$$(68) \quad \psi(x) \ln x + \sum_{n \leq x} \Lambda(n) \psi(x/n) = 2x \ln x + O(x) \quad \text{and}$$

$$(69) \quad R(x) \ln x + \sum_{n \leq x} \Lambda(n) R(x/n) = O(x) \quad (\text{where } R(x) = \psi(x) - x),$$

in which the symbol  $*$  denotes the Dirichlet convolution operation on arithmetic sequences, defined by  $a * b = \sum_{jk=n} a(j)b(k) = \sum_{j|n} a(j)b(n/j)$ .

Selberg’s original proof derived the equation  $\lim_{x \rightarrow \infty} \theta(x)/x = 1$  from (66) using a consequence of (66) discovered by Erdős. In particular, denoting  $\liminf \theta(x)/x$  by  $a$  and  $\limsup \theta(x)/x$  by  $A$ , Selberg deduced from (66) that  $a + A = 2$ . Meanwhile, unaware of that fact, Erdős used (66) to show that for any  $\delta > 0$  there is a constant  $K(\delta)$  such that for sufficiently large values of  $x$  there are more than  $K(\delta)x/\ln x$  primes in the interval  $(x, x + K(\delta)x)$ . Erdős communicated his proof of that fact to Selberg, who then, via an intricate argument, used Erdős’s result to prove that  $A \leq a$ . Consequently,  $A = a = 1$ . (See Erdős 1949 for details of all those proofs. The proof given in Selberg 1949 is a later, more direct one that does not use Erdős’s result.<sup>15</sup>)

The most accessible elementary proof of the Prime Number Theorem is probably that in Levinson (1969), whose very title (“A motivated elementary proof of the Prime Number Theorem”) suggests that the strategies underlying elementary proofs

---

<sup>15</sup>Regrettably, the interaction between Erdős and Selberg in this matter was a source of lasting bitterness between them. Goldfeld (2004) provides a balanced account of the dispute, based on primary sources. As noted there, the issue was not one of priority of discovery, but “arose over the question of whether a joint paper (on the entire proof) or separate papers (on each individual contribution) should appear”.

of that theorem are not perspicuous.<sup>16</sup> Variants of Levinson's proof are also given in Edwards (1974) and Jameson (2003). Here, in outline, is the structure of the latter version:

1. The goal is to show that  $\lim_{x \rightarrow \infty} \psi(x)/x = 1$ . Toward that end, three related functions whose behavior is easier to study are defined, namely

$$R(x) = \begin{cases} 0 & \text{if } x < 1 \\ \psi(x) - x & \text{if } x \geq 1 \end{cases}, \quad S(x) = \int_0^x \frac{R(t)}{t} dt \quad \text{and} \quad W(x) = \frac{S(e^x)}{e^x}.$$

Then  $\lim_{x \rightarrow \infty} \psi(x)/x = 1$  if and only if  $\lim_{x \rightarrow \infty} R(x)/x = 0$ .

2. It follows from Chebyshev's result  $\psi(x) \leq 2x$  that  $|R(x)| \leq x$  for  $x > 0$ . Furthermore,  $\int_1^x R(t)/t^2 dt = \int_1^x \psi(t)/t^2 dt - \ln x$ , and Abel summation yields

$$\int_1^x \frac{\psi(t)}{t^2} dt = \sum_{n \leq x} \frac{\Lambda(n)}{n} - \frac{\psi(x)}{x}.$$

On the other hand, Mertens in 1874 applied Chebyshev's bound to obtain that  $\sum_{n \leq x} \Lambda(n)/n = \ln x + O(1)$ . (See Jameson 2003, p. 90 for details.) Consequently, applying Chebyshev's bound once more,  $\int_1^x R(t)/t^2 dt$  is bounded for  $x > 1$ .

3. Since, by 2., the absolute value of the integrand in the definition of  $S$  is bounded by 1,  $S$  satisfies the Lipschitz condition  $|S(x_2) - S(x_1)| \leq x_2 - x_1$  for  $x_2 > x_1 > 0$ . That, in turn, together with the inequality  $e^{-x} \geq 1 - x$ , shows that  $W$  likewise satisfies the Lipschitz condition  $|W(x_2) - W(x_1)| \leq 2(x_2 - x_1)$  for  $x_2 > x_1 > 0$ .
4. The Lipschitz condition on  $S$  gives  $|S(x)| \leq x$  for  $x > 0$ , that is,  $|S(x)/x| \leq 1$  for  $x > 0$ . Then  $|W(x)| \leq 1$  and

$$\begin{aligned} \int_1^x \frac{S(t)}{t^2} dt &= \int_1^x \frac{1}{t^2} \int_1^t \frac{R(u)}{u} du dt \\ &= \int_1^x \frac{R(u)}{u} \int_u^x \frac{1}{t^2} dt du \\ &= \int_1^x \frac{R(u)}{u} \left( \frac{1}{u} - \frac{1}{x} \right) du \\ &= \int_1^x \frac{R(u)}{u^2} du - \frac{S(x)}{x}, \end{aligned}$$

---

<sup>16</sup>Levinson's paper won the Mathematical Association of America's Chauvenet Prize for exposition in 1971. Nevertheless, after reading it, the number theorist Harold Stark commented "Well, Norman tried, but the thing is as mysterious as ever." (Quoted in Segal 2009, p. 99.)

which is bounded for all  $x > 1$  by the result of step 2. above. Consequently,  $\int_0^x W(t) dt$  is bounded for all  $x > 0$ .

5. Straightforward arguments with inequalities yield the following Tauberian theorem:

If  $A(x) \geq 0$ ,  $A(x)$  is increasing for  $x > 1$ , and  $\frac{1}{x} \int_1^x \frac{A(t)}{t} dt \rightarrow 1$  as  $x \rightarrow \infty$ ,  
then  $\frac{A(x)}{x} \rightarrow 1$  as  $x \rightarrow \infty$ .

Then since  $S(x)/x = (1/x) \int_1^x \psi(t)/t dt - 1 - 1/x$ , taking  $A(x) = \psi(x)$  in the Tauberian theorem shows that to prove that  $\psi(x)/x \rightarrow 1$  as  $x \rightarrow \infty$  it suffices to prove that  $S(x)/x \rightarrow 0$  as  $x \rightarrow \infty$ .

6. Equivalently, it suffices to show that  $W(x) \rightarrow 0$  as  $x \rightarrow \infty$ . For that purpose, let

$$\alpha = \limsup_{x \rightarrow \infty} |W(x)| \leq 1 \quad \text{and} \quad \beta = \limsup_{x \rightarrow \infty} \frac{1}{x} \int_0^x |W(t)| dt.$$

Then  $\beta \leq \alpha$ . Crucially, however,

$$(70) \quad \beta = \alpha \quad \text{only if} \quad \alpha = 0;$$

so to prove the Prime Number Theorem it suffices to show that  $\alpha \leq \beta$ .

To prove (70), assume  $\alpha > 0$ . Since  $\int_0^x W(t) dt$  is bounded for all  $x > 0$ , there is a constant  $B$  such that  $|\int_{x_1}^{x_2} W(t) dt| \leq B$  for all  $x_2 > x_1 > 0$ . Also, by the definition of  $\alpha$ , for any  $a > \alpha$  there is some  $x_a$  such that  $|W(x)| \leq a$  for  $x > x_a$ . So suppose  $\alpha < a \leq 2\alpha$  and consider  $\int_{x_1}^{x_1+h} |W(t)| dt$ , where  $x_1 \geq x_a$  and  $h \geq 2\alpha$  is to be determined. If  $W(x)$  changes sign within the interval  $[x_1, x_1 + h]$ , the intermediate-value theorem yields the existence of a point  $z$  in that interval where  $W(z) = 0$ . The Lipschitz condition on  $W$  then gives  $|W(x)| \leq 2|x - z|$ , and since  $h \geq a$ , at least one of the points  $z \pm \frac{a}{2}$  lies between  $x_1$  and  $x_1 + h$ . So either the interval  $[z, z + a/2]$  lies within  $[x_1, x_1 + h]$  or  $[z - a/2, z]$  does. Whichever does, call it  $I$ . (If both do, pick one.) Then  $\int_I |W(x)| dx \leq \int_I 2|x - z| dx = a^2/4$ . The part of  $[x_1, x_1 + h]$  lying outside  $I$  has length  $h - a/2$ , and there  $|W(x)| \leq a$ . So, finally,

$$(71) \quad \int_{x_1}^{x_1+h} |W(x)| dx \leq \frac{a^2}{4} + a \left( h - \frac{a}{2} \right) = a \left( h - \frac{a}{4} \right) < a \left( h - \frac{\alpha}{4} \right).$$

By choosing  $h$  to be the greater of  $\alpha$  and  $B/\alpha + \alpha/4$ , (71) can be ensured to hold as well if  $W(x)$  does not change sign within  $[x_1, x_1 + h]$ .

To complete the proof, note that for any  $x \geq x_1 + h$  there is an integer  $n$  such that  $x_a + nh \leq x < x_a + (n + 1)h$ . Then



$$\begin{aligned} \int_0^x |W(x)| dx &\leq \int_0^{x_a} |W(x)| dx + (n + 1) \left( h - \frac{\alpha}{4} \right) ah \\ &= C + (n + 1) \left( h - \frac{\alpha}{4} \right) ah, \end{aligned}$$

where  $C$  is constant and  $n \rightarrow \infty$  as  $x$  does. Hence, since  $x > nh$ ,

$$\frac{1}{x} \int_0^x |W(x)| dx \leq \frac{C}{x} + \left( 1 + \frac{1}{n} \right) \left( 1 - \frac{\alpha}{4h} \right) a.$$

As  $x \rightarrow \infty$ , the right member of that inequality approaches  $(1 - \alpha/4h)a$ . Consequently  $\beta \leq (1 - \alpha/4h)a$  for any  $a > \alpha$ . Therefore  $\beta \leq (1 - \alpha/4h)\alpha < \alpha$ .

7. It is at this point that Selberg’s inequality enters in. In Levinson (1969) and Jameson (2003) that inequality, in the forms (68) and (69), is derived as a corollary to the *Tatuzawa-Iseki identity*, which states that if  $F$  is a function defined on the interval  $[1, \infty)$  and  $G(x) = \sum_{n \leq x} F(x/n)$ , then for  $x \geq 1$ :

$$(72) \quad \sum_{k \leq x} \mu(k) \ln \frac{x}{k} G\left(\frac{x}{k}\right) = F(x) \ln x + \sum_{n \leq x} \Lambda(n) F\left(\frac{x}{n}\right),$$

where  $\mu$  denotes the Möbius function and  $\Lambda$  the von Mangoldt function.

The *Tatuzawa-Iseki identity* is a variant of the *Möbius inversion formula*, which states that under the same hypotheses,  $F(x) = \sum_{n \leq x} \mu(n)G(x/n)$ . It is obtained by multiplying the Möbius formula by  $\ln x$  to get the first term in the right member of (72), rewriting the factor  $\ln(x/k)$  in the left member as  $\ln x - \ln k$ , replacing  $G\left(\frac{x}{k}\right)$  in  $\sum_{k \leq x} \mu(k) \ln k G\left(\frac{x}{k}\right)$  by  $\sum_{j \leq x/k} F\left(\frac{x}{jk}\right)$ , interchanging the order of summation in the double sum, and expressing the result in terms of  $\Lambda(n)$ .<sup>17</sup>

To obtain inequality (68), the *Tatuzawa-Iseki identity* is applied with  $F(x) = R(x) + \gamma + 1$ , where  $\gamma$  is Euler’s constant. One proves that then  $|G(x)| \leq \ln x + 2$  for  $x \geq 1$ . The derivation is completed by invoking the integral test for series together with Chebyshev’s upperbound for  $\psi(x)$ .

8. Inequality (67) is deduced as a corollary to (68), using Abel summation and Chebyshev’s result that  $\psi(x) \leq 2x$ . Since the integral test implies that  $\sum_{n \leq x} \ln n = x \ln x + O(x)$ , a further equivalent of Selberg’s formula is  $\sum_{n \leq x} [\Lambda(n) \ln n + (\Lambda * \Lambda)(n)] - 2 \ln n = O(x)$ .
9. Dividing (69) by  $t$ , where  $1 \leq t \leq x$ , gives

$$R(t) \frac{\ln t}{t} + \frac{1}{t} \sum_{n \leq x} \Lambda(n) R\left(\frac{t}{n}\right) = O(1).$$

---

<sup>17</sup>The Möbius inversion formula by itself does not suffice to give the desired bound on  $R(x)$ , and that, in Levinson’s opinion, accounts for “the long delay in the discovery of an elementary proof” of the Prime Number Theorem. (Levinson 1969, p. 235)

Integrating and using  $|S(t)/t| \leq 1$  from step 4, then shows that for  $x \geq 1$ ,

$$(73) \quad S(x) \ln x + \sum_{n \leq x} \Lambda(n) S(x/n) = O(x),$$

which is (69) with  $R$  replaced by  $S$ .

10. Three technical lemmas are now proved, using (68), (69) and the Lipschitz condition on  $S$ , respectively. First, for all  $x \geq 1$  there is a constant  $K_1$  such that

$$(74) \quad |S(x)|(\ln x)^2 \leq \sum_{n \leq x} [\Lambda(n) \ln n + (\Lambda * \Lambda)(n)] \left| S\left(\frac{x}{n}\right) \right| + K_1 x \ln x.$$

Second,

$$(75) \quad \sum_{n \leq x} [\Lambda(n) \ln n + (\Lambda * \Lambda)(n)] \left| S\left(\frac{x}{n}\right) \right| = 2 \sum_{n \leq x} \ln n \left| S\left(\frac{x}{n}\right) \right| + O(x \ln x);$$

and third, for all  $x \geq 1$  there is a constant  $K_2$  such that

$$(76) \quad \sum_{n \leq x} \ln n \left| S\left(\frac{x}{n}\right) \right| \leq \int_1^x \ln t \left| S\left(\frac{x}{t}\right) \right| dt + K_2 x.$$

11. Together, (74), (75), and (76) yield that for all  $x \geq 1$  there is a constant  $K_3$  such that

$$|S(x)|(\ln x)^2 \leq 2 \int_1^x \ln t \left| S\left(\frac{x}{t}\right) \right| dt + K_3 x \ln x, \quad \text{so that}$$

$$(77) \quad |W(x)| \leq \frac{2}{x^2} \int_0^x (x-u) |W(u)| du + \frac{K_3}{x}.$$

12. Finally, for  $\alpha$  and  $\beta$  as defined in step 6., (77) implies that  $\alpha \leq \beta$ . For, by the definition of  $\beta$ , for every  $\epsilon > 0$  there exists an  $x_1$  such that for every  $x \geq x_1$ ,  $\int_0^x |W(t)| dt \leq (\beta + \epsilon)x$ . In order to apply (77), consider  $\int_0^x (x-u) |W(u)| du$ , which may be rewritten as

$$\int_0^x |W(u)| \int_u^x dt du = \int_0^x \int_0^t |W(u)| du dt.$$

Then for  $x \geq x_1$ ,

$$\int_{x_1}^x \int_0^t |W(u)| du dt \leq \int_{x_1}^x (\beta + \epsilon)t dt = \frac{1}{2}(\beta + \epsilon)(x^2 - x_1^2).$$

On the other hand,  $|W(u)| \leq 1$  for all  $u$  by step 4. So

$$\int_0^{x_1} \int_0^t |W(u)| \, du \, dt \leq \int_0^{x_1} t \, dt = \frac{1}{2}x_1^2.$$

Therefore

$$\begin{aligned} \int_0^x (x-u)|W(u)| \, du &= \int_0^{x_1} (x-u)|W(u)| \, du + \int_{x_1}^x (x-u)|W(u)| \, du \\ &\leq \frac{1}{2}(\beta + \epsilon)x^2 + \frac{1}{2}x_1^2. \end{aligned}$$

Then by (77),  $|W(x)| \leq \beta + \epsilon + \frac{x_1^2}{x^2} + \frac{K_3}{x}$ . By definition of  $\alpha$ , that means  $\alpha \leq \beta + \epsilon$ ; and since  $\epsilon$  can be chosen to be arbitrarily small,  $\alpha \leq \beta$ .    q.e.d.

## 10.7 Overview

The five proofs of the Prime Number Theorem considered here employ a wide range of methodologies, including analytic continuation, Abel summation, Dirichlet convolution, contour integration, Fourier analysis, Laplace transforms, and Tauberian theorems. They differ from one another in many respects, including the ways in which the domain of definition of  $\zeta(s)$  is extended and whether or not recourse is made to the functional equation for the  $\zeta$ -function.

The proofs also exemplify several of the different motivations discussed in Chapter 2. For example, Riemann's program for proving the Prime Number Theorem by examining the behavior of the complex  $\zeta$ -function proposed bringing the methods of complex analysis to bear on the seemingly remote field of number theory; so the successful carrying out of that program by Hadamard and de la Vallée Poussin may be deemed instances of benchmarking. And the marked differences in the independent and nearly simultaneous proofs of the theorem that Hadamard and de la Vallée Poussin gave — each of which was based upon the same corpus of earlier work (especially Chebyshev's results) and followed the same basic steps in Riemann's program (proving that  $\zeta(s)$  has no roots of the form  $1 + it$  and then applying complex contour integration to expressions involving  $\zeta'/\zeta$ ) — are surely attributable to differences in their individual patterns of thought.

In the case of Hadamard, a further motive was that of correcting deficiencies in Cahen's earlier proof attempt. Indeed, at the beginning of section 12 of his memoir, Hadamard noted explicitly that Cahen had claimed to have proved that  $\theta(x) \sim x$ , but that his demonstration was based on an unsubstantiated claim by Stieltjes (who thought he had proved the Riemann Hypothesis). Nonetheless, Hadamard declared,

“we will show that the same result can be obtained in a completely rigorous way” via “an easy modification” of Cahen’s analysis.<sup>18</sup>

The modern descendants of those classical proofs illustrate how simplifications, generalizations, and refinements of mathematical arguments gradually evolve and are incorporated into later proofs. Examples include the proof that  $\zeta(s)$  has no roots with real part equal to one, which was significantly shortened and simplified through the use of a trigonometric identity; the deduction of the Prime Number Theorem from a general Tauberian theorem, which showed it to be a particular instance of a family of such theorems concerning Dirichlet series;<sup>19</sup> and the use in Newman’s proof of a contour integral that is much more easily evaluated than the classical ones. Newman’s proof also exhibits economy of means with regard to conceptual prerequisites, making it comprehensible to those having only a rudimentary knowledge of complex analysis.

The ‘elementary’ proofs, on the other hand, are no simpler than the classical analytic ones, are generally regarded as *less* perspicuous (giving little or no insight into *why* the Prime Number Theorem is true), and do not yield as sharp error bounds as those obtained by analytic means. The esteem nonetheless accorded them by the mathematical community, as reflected in the prizes awarded to their discoverers, may thus be taken as a quintessential manifestation of the high regard mathematicians have for purity of method in and of itself.

## References

- Bateman, P., Diamond, H.: A hundred years of prime numbers. *Amer. Math. Monthly* **103**, 729–741 (1996)
- Chebyshev, P.: Sur la fonction qui détermine la totalité des nombres premiers inférieurs à une limite donnée. *Mémoires des savants étrangers de l’Acad. Sci. St.Pétersbourg* **6**, 1–19 (1848)
- Chebyshev, P.: Mémoire sur nombres premiers. *Mémoires des savants étrangers de l’Acad. Sci. St.Pétersbourg* **7**, 17–33 (1850)
- de la Vallée Poussin, C.: Recherches analytiques sur la théorie des nombres premiers. *Ann. Soc. Sci. Bruxelles* **20**, 183–256 (1896)
- Diamond, H.: Elementary methods in the study of the distribution of prime numbers. *Bull. Amer. Math. Soc.* **7**, 553–589 (1982)
- Edwards, H.M.: *Riemann’s Zeta Function*. Academic Press, New York and London (1974)
- Erdős, P.: On a new method in elementary number theory which leads to an elementary proof of the prime number theorem. *Proc. Nat. Acad. Sci. U.S.A.* **35**, 374–384 (1949)

<sup>18</sup>“Nous allons voir qu’en modifiant légèrement l’analyse de l’auteur on peut établir le même résultat en toute rigueur.”

<sup>19</sup>Three other instances given in Jameson (2003) are  $\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0$ , where  $\mu$  denotes the Möbius function,  $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^{1+it}} = \frac{1}{\zeta(1+it)}$ , and  $\sum_{n=1}^{\infty} \frac{(-1)^{\Omega(n)}}{n} = 0$ , where  $\Omega(n)$  denotes the number of prime factors of  $n$ , each counted according to its multiplicity.

- Goldfeld, D.: The elementary proof of the prime number theorem: an historical perspective. *Number Theory* (New York Seminar, 2003). Springer, New York, 179–192 (2004)
- Goldstein, L.J.: A history of the prime number theorem. *Amer. Math. Monthly* **80**, 599–615 (1973)
- Hadamard, J.: Étude sur les propriétés des fonctions entières et en particulier d'une fonction considérée par Riemann. *J. math. pures appl. (4)* **9**, 171–215 (1893)
- Hadamard, J.: Sur les zéros de la fonction  $\zeta(s)$  de Riemann. *Comptes Rendus Acad. Sci. Paris* **122**, 1470–1473 (1896a)
- Hadamard, J.: Sur la distribution des zéros de la fonction  $\zeta(s)$  et ses conséquences arithmétiques. *Bull. Soc. Math. France* **24**, 199–220 (1896b)
- Hardy, G.H., Littlewood, J.: New proofs of the prime number theorem and similar theorems. *Quart. J. Math., Oxford ser.* **46**, 215–219 (1915)
- Ingham, A.E.: On Wiener's method in Tauberian theorems. *Proc. London Math. Soc. (2)* **38**, 458–480 (1935)
- Jameson, G.J.O.: *The Prime Number Theorem*. London Mathematical Society Student Texts **53**. Cambridge U.P., Cambridge (2003)
- Korevaar, J.: On Newman's quick way to the Prime Number Theorem. *Math. Intelligencer* **4:3**, 108–115.
- Landau, E.: Zwei neue Herleitungen für die asymptotische Anzahl der Primzahlen unter einer gegebenen Grenze. *SBer. Kgl. Preuß. Akad. Wiss. Berlin*, 746–764 (1908)
- Landau, E.: *Handbuch der Lehre von der Verteilung der Primzahlen*. Teubner, Leipzig (1909)
- Lax, P.D., Zalcman, L.: *Complex Proofs of Real Theorems*. University Lecture Series 58. Amer. Math. Soc., Providence (2012)
- Levinson, N.: A motivated account of an elementary proof of the prime number theorem. *Amer. Math. Monthly* **76**, 225–245 (1969)
- Narkiewicz, W.: *The Development of Prime Number Theory*. Springer, Berlin (2000)
- Newman, D.J.: Simple analytic proof of the prime number theorem. *Amer. Math. Monthly* **87**, 693–696 (1980)
- Riemann, B.: Ueber die Anzahl der Primzahlen unter einer gegebenen Größe. *Monatsberichte der Königlich Preußische Akademie der Wissenschaften zu Berlin*, 671–680 (1860)
- Segal, J.: *Recountings: Conversations with MIT Mathematicians*. A K Peters, Natick, Mass. (2009)
- Selberg, A.: An elementary proof of the prime number theorem. *Ann. Math. (2)* **50**, 305–313 (1949)
- von Mangoldt, H.: Zu Riemanns Abhandlung "Ueber die Anzahl der Primzahlen unter einer gegebenen Größe". *J. reine u. angewandte Math.* **114**, 255–305 (1895)
- Zagier, D.: Newman's short proof of the Prime Number Theorem. *Amer. Math. Monthly* **104**, 705–708 (1997)