term be paired with its "twin" $\rho \leftrightarrow 1 - \rho$,

$$\sum_{\operatorname{Im}\rho>0}\left[\log\left(1-\frac{s}{\rho}\right)+\log\left(1-\frac{s}{1-\rho}\right)\right], \tag{3}$$

because this sum converges absolutely. The proof of the absolute convergence of (3) is roughly as follows.

To prove the absolute convergence of

$$\sum_{\operatorname{Im}\rho>0}\log\left[\left(1-\frac{s}{\rho}\right)\left(1-\frac{s}{1-\rho}\right)\right]=\sum_{\operatorname{Im}\rho>0}\log\left[1-\frac{s(1-s)}{\rho(1-\rho)}\right],$$

it suffices to prove the absolute convergence of

$$\sum_{\operatorname{Im}\rho>0}\frac{1}{\rho(1-\rho)}. \tag{4}$$

(In other words, to prove the absolute convergence of a product $\prod(1 + a_i)$, it suffices to prove the absolute convergence of the sum $\sum a_i$.) But the estimate of the distribution of the roots ρ given in the preceding section indicates that their density is roughly

$$d\left(\frac{T}{2\pi}\log\frac{T}{2\pi}-\frac{T}{2\pi}\right)=\frac{1}{2\pi}\log\frac{T}{2\pi}\,dT.$$

Hence

$$\sum_{\operatorname{Im}\rho>0}\frac{1}{\rho(1-\rho)}\sim\int^{\infty}\frac{1}{T^2}\frac{1}{2\pi}\log\frac{T}{2\pi}\,dT<\infty,$$

or, in short, the terms are like T^{-2} and their density is like $\log T$ so their sum converges. As will be seen in Chapter 2, the only serious difficulty in making this into a rigorous proof of the absolute convergence of (3) is the proof that the vertical density of the roots ρ is in some sense a constant times $\log(T/2\pi)$. Riemann merely states this fact without proof.

Riemann then goes on to say that the function defined by (3) grows only as fast as $s \log s$ for large s; hence, because it differs from $\log \xi(s)$ by an even function of $s - \frac{1}{2}$ [and because $\log \xi(s)$ also grows like $s \log s$ for large s], this difference must be constant because it can contain no terms in $(s - \frac{1}{2})^2$, $(s - \frac{1}{2})^4, \ldots$. It will be shown in Chapter 2 that the steps in this argument can all be filled in more or less as Riemann indicates, but it must be admitted that Riemann's sketch is so abbreviated as to make it virtually useless in constructing a proof of (2).

The first proof of the product representation (2) of $\xi(s)$ was published by Hadamard [H1] in 1893.

1.11 THE CONNECTION BETWEEN $\zeta(s)$ AND PRIMES

The essence of the relationship between $\zeta(s)$ and prime numbers is the Euler product formula

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}} \qquad (\operatorname{Re} s > 1) \tag{1}$$

in which the product on the right is over all prime numbers p. Taking the log of both sides and using the series $\log(1 - x) = -x - \frac{1}{2}x^2 - \frac{1}{3}x^3 - \cdots$ puts this in the form

$$\log \zeta(s) = \sum_p \left[\sum_n (1/n)p^{-ns} \right] \qquad (\operatorname{Re} s > 1).$$

Since the double series on the right is absolutely convergent for $\operatorname{Re} s > 1$, the order of summation is unimportant and the sum can be written simply

$$\log \zeta(s) = \sum_p \sum_n (1/n)p^{-ns} \qquad (\operatorname{Re} s > 1). \tag{2}$$

It will be convenient in what follows to write this sum as a Stieltjes integral

$$\log \zeta(s) = \int_0^\infty x^{-s}\, dJ(x) \qquad (\operatorname{Re} s > 1) \tag{3}$$

where $J(x)$ is† the function which begins at 0 for $x = 0$ and increases by a jump of 1 at primes p, by a jump of $\frac{1}{2}$ at prime squares p^2, by a jump of $\frac{1}{3}$ at prime cubes, etc. As is usual in the theory of Stieltjes integrals, the value of $J(x)$ at each jump is defined to be halfway between its new value and its old value. Thus $J(x)$ is zero for $0 \le x < 2$, is $\frac{1}{2}$ for $x = 2$, is 1 for $2 < x < 3$, is $1\frac{1}{2}$ for $x = 3$, is 2 for $3 < x < 4$, is $2\frac{1}{4}$ for $x = 4$, is $2\frac{1}{2}$ for $4 < x < 5$, is 3 for $x = 5$, is $3\frac{1}{2}$ for $5 < x < 7$, etc. A formula for $J(x)$ is

$$J(x) = \frac{1}{2}\left[\sum_{p^n < x} \frac{1}{n} + \sum_{p^n \le x} \frac{1}{n} \right].$$

Riemann did not, of course, have the vocabulary of Stieltjes integration available to him, and he stated (3) in the slightly different form

$$\log \zeta(s) = s \int_0^\infty J(x) x^{-s-1}\, dx \qquad (\operatorname{Re} s > 1) \tag{4}$$

which can be obtained from (3) by integration by parts. [As $x \downarrow 0$, clearly $x^{-s}J(x) = 0$ because $J(x) \equiv 0$ for $x < 2$. On the other hand, $J(x) < x$ for all x, so $x^{-s}J(x) \to 0$ as $x \to \infty$ for $\operatorname{Re} s > 1$.] The integral in (4) can be con-

†Riemann denotes this function $f(x)$, and most other writers denote it $\Pi(x)$. Since $f(x)$ now is commonly used to denote a generic function and since $\Pi(x)$ in this book denotes the factorial function, I have taken the liberty of introducing a new notation $J(x)$ for this function.

sidered to be an ordinary Riemann integral and the formula itself can be derived without using Stieltjes integration by setting

$$p^{-ns} = s\int_{p^n}^{\infty} x^{-s-1}\,dx \qquad (\operatorname{Re} s > 1)$$

in (2), which is Riemann's derivation of (4).

Formulas (2)–(4) should all be thought of as minor variations of the Euler product formula (1) which is the basic idea connecting $\zeta(s)$ and primes.

1.12 FOURIER INVERSION

Riemann was a master of Fourier analysis and his work in developing this theory must certainly be counted among his greatest contributions to mathematics. It is not surprising, therefore, that he immediately applies Fourier inversion to the formula

$$\text{(1)} \qquad \frac{\log \zeta(s)}{s} = \int_0^{\infty} J(x)x^{-s-1}\,dx \qquad (\operatorname{Re} s > 1)$$

to conclude

$$\text{(2)} \qquad J(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \log \zeta(s) x^s \frac{ds}{s} \qquad (a > 1).$$

Then using an alternative formula for $\log \zeta(s)$, he obtains an alternative formula for $J(x)$ which is the main result of the paper.

[The improper integral in (2) is only conditionally convergent and an "order of summation" must be specified. Here it is understood that the integral in (2) means the limit as $T \to \infty$ of the integral over the vertical line segment from $a - iT$ to $a + iT$. More generally, conditionally convergent integrals and series are very common in Fourier analysis, and it is always understood that such integrals and series are summed in their "natural order"; for example,

$$\sum_{n=-\infty}^{\infty} c_n e^{inx} \qquad \text{means} \qquad \lim_{N\to\infty} \sum_{n=-N}^{N} c_n e^{inx},$$

$$\int_{-\infty}^{\infty} f(y)e^{iyx}\,dy \qquad \text{means} \qquad \lim_{T\to\infty} \int_{-T}^{T} f(y)e^{iyx}\,dy,$$

etc. This is analogous to the convention that discontinuous functions such as $J(x)$ assume the middle value $J(x) = \frac{1}{2}[J(x-\epsilon) + J(x+\epsilon)]$ at any jump x, that divergent integrals such as $\operatorname{Li}(x)$ (see Section 1.14 below) are taken to mean the Cauchy principal value, and that the product $\prod [1 - (s/\rho)]$ is ordered in such a way as to pair ρ with $1 - \rho$, or, later on, ordered by $|\operatorname{Im} \rho|$.]

In deriving (2) from (1) Riemann makes use of "Fourier's theorem," by which he means† the Fourier inversion formula

$$\phi(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\phi(\lambda)e^{i(x-\lambda)\mu}\,d\lambda\right]d\mu. \tag{3}$$

Otherwise stated, "Fourier's theorem" is the statement that in order to write a given function $\phi(x)$ as a superposition of exponentials

$$\phi(x) = \int_{-\infty}^{\infty}\Phi(\mu)e^{i\mu x}\,d\mu,$$

it is necessary and sufficient (under suitable conditions) that the "coefficients" $\Phi(\mu)$ of the expansion be defined by

$$\Phi(\mu) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(\lambda)e^{-i\lambda\mu}\,d\lambda.$$

This statement of Fourier's theorem brings out the analogy with Fourier series

$$f(x) = \sum_{-\infty}^{\infty}a_n e^{inx} \Longleftrightarrow a_n = \frac{1}{2\pi}\int_0^{2\pi}f(\lambda)e^{-in\lambda}\,d\lambda,$$

and in fact theorem (3) for Fourier integrals follows formally from a passage to the limit in the theorem for Fourier series.

To derive (2) from (1), let $s = a + i\mu$, where a is a constant $a > 1$ and μ is a real variable, let $\lambda = \log x$, and let $\phi(x) = 2\pi J(e^x)e^{-ax}$. Then (1) becomes

$$\begin{aligned}\frac{\log\zeta(a+i\mu)}{a+i\mu} &= \int_{-\infty}^{\infty}J(e^\lambda)e^{-(a+i\mu)\lambda}\,d\lambda \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(\lambda)e^{-i\mu\lambda}\,d\lambda \qquad (a > 1),\end{aligned}$$

and when this function is taken to be $\Phi(\mu)$, Fourier's theorem gives

$$\begin{aligned}2\pi J(e^x)e^{-ax} &= \int_{-\infty}^{\infty}\frac{\log\zeta(a+i\mu)}{a+i\mu}e^{i\mu x}\,d\mu, \\ J(y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\log\zeta(a+i\mu)}{a+i\mu}y^{a+i\mu}\,d\mu,\end{aligned}$$

from which (2) follows immediately.

Riemann completely ignores the question of the applicability of Fourier's theorem to the function $J(e^x)e^{-ax}$ and states simply that (2) holds in "complete generality." However, $J(e^x)e^{-ax}$ is a very well-behaved function—it has simple well-behaved jumps, it is identically zero for $x < 0$, and it goes to zero

†See Riemann [R2, p. 86].

faster than $e^{-(a-1)x}$ as $x \longrightarrow \infty$—and the very simplest theorems† on Fourier integrals suffice to prove rigorously Riemann's statement that (2) holds in complete generality.

1.13 METHOD FOR DERIVING THE FORMULA FOR $J(x)$

The two formulas for $\xi(s)$, namely,

$$\xi(s) = \Pi\left(\frac{s}{2}\right)\pi^{-s/2}(s-1)\zeta(s) \qquad \text{and} \qquad \xi(s) = \xi(0)\prod_{\rho}\left(1 - \frac{s}{\rho}\right),$$

combine to give

$$\begin{aligned}\log \zeta(s) &= \log \xi(s) - \log \Pi\left(\frac{s}{2}\right) + \frac{s}{2}\log \pi - \log(s-1)\\ &= \log \xi(0) + \sum_{\rho} \log\left(1 - \frac{s}{\rho}\right) - \log \Pi\left(\frac{s}{2}\right)\\ &\quad + \frac{s}{2}\log \pi - \log(s-1).\end{aligned}$$

Riemann's formula for $J(x)$, which is the main result of his paper, is obtained essentially by substituting this formula for $\log \zeta(s)$ in the formula

$$J(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \log \zeta(s) x^s \frac{ds}{s} \qquad (a > 1)$$

of the preceding section and integrating termwise. However, because a direct substitution leads to divergent integrals [the term $(s/2)\log \pi$, for example, leads to an integral which is a constant times $(i)^{-1}\int x^s\, ds = e^a \int e^{iu \log x}\, du$ which oscillates and does not converge even conditionally], Riemann first integrates by parts to obtain

$$(1) \qquad J(x) = -\frac{1}{2\pi i}\cdot\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty} \frac{d}{ds}\left[\frac{\log \zeta(s)}{s}\right] x^s\, ds \qquad (a > 1)$$

before substituting the above expression for $\log \zeta(s)$. The validity of the integration by parts by which (1) is obtained depends merely on showing that

$$(2) \qquad \lim_{T\to\infty} \frac{\log \zeta(a \pm iT)}{a \pm iT} x^{a \pm iT} = 0,$$

†See, for example, Taylor [T2]. The particular form of Fourier inversion that Riemann uses here—which is essentially Fourier analysis on the multiplicative group of positive reals rather than on the additive group of all real numbers—is often called Mellin inversion. Riemann's work precedes that of Mellin by 40 years.

which follows easily from the inequality

$$\begin{aligned} |\log \zeta(a \pm iT)| &= \left|\sum_n \sum_p (1/n) p^{-n(a \pm iT)}\right| \\ &\le \sum_n \sum_p (1/n) p^{-na} = \log \zeta(a) = \text{const} \end{aligned} \tag{3}$$

because this shows that the numerator in (2) is bounded while the denominator goes to infinity.

The substitution of

$$\log \zeta(s) = \log \xi(0) + \sum_\rho \log\left(1 - \frac{s}{\rho}\right) - \log \Pi\left(\frac{s}{2}\right) + \frac{s}{2} \log \pi - \log(s-1)$$

into (1) expresses $J(x)$ as a sum of five terms (the integral of a finite sum is always the sum of the integrals provided the latter converge) and the derivation of Riemann's formula for $J(x)$ depends now on the evaluation of these five definite integrals.

It should be noted that for any fixed s there is some ambiguity in the definition of $\log[1-(s/\rho)]$ for those roots ρ which are not large relative to s. In order to remove this ambiguity in $\operatorname{Re} s > 1$ let $\log[1-(s/\rho)]$ be defined to be $\log(s-\rho) - \log(-\rho)$; this is meaningful because none† of the ρ's are real and greater than or equal to 0. In this way $\log[1-(s/\rho)]$ is unambiguously defined throughout $\operatorname{Re} s > 1$ and, in particular, on the path of integration $\operatorname{Re} s = a > 1$.

1.14 THE PRINCIPAL TERM OF $J(x)$

It will be seen below that the principal term in the formula for $J(x)$ is the term corresponding to the term $-\log(s-1)$ of the expansion of $\log \zeta(s)$. This term is

$$\frac{1}{2\pi i} \frac{1}{\log x} \int_{a-i\infty}^{a+i\infty} \frac{d}{ds}\left[\frac{\log(s-1)}{s}\right] x^s \, ds \qquad (a > 1).$$

Riemann shows that for $x > 1$ the value of this definite integral is the logarithmic integral

$$\operatorname{Li}(x) = \lim_{\epsilon \downarrow 0} \left[\int_0^{1-\epsilon} \frac{dt}{\log t} + \int_{1+\epsilon}^x \frac{dt}{\log t}\right];$$

†See Section 2.3, or observe that the series $1 - 2^{-s} + 3^{-s} - 4^{-s} + 5^{-s} - \cdots$ converges to a positive number for $s > 0$ and that this number is

$$\zeta(s) - 2 \cdot 2^{-s}(1 + 2^{-s} + 3^{-s} + \cdots) = (1 - 2^{1-s})\zeta(s).$$

that is, it is the Cauchy principal value of the divergent integral $\int_0^x (dt/\log t)$. His argument is as follows:

Fix $x > 1$ and consider the function of β defined by

$$F(\beta) = \frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty} \frac{d}{ds}\left\{\frac{\log[(s/\beta) - 1]}{s}\right\} x^s\, ds$$

so that the desired number is $F(1)$. The definition of $F(\beta)$ can be extended to all real or complex numbers β other than real numbers $\beta \leq 0$ by taking $a >$ Re β and defining $\log[(s/\beta) - 1]$ to be $\log(s - \beta) - \log \beta$, where, as usual, $\log z$ is defined for all z other than real $z \leq 0$ by the condition that it be real for real $z > 0$. The integral $F(\beta)$ converges absolutely because

$$\left|\frac{d}{ds}\left\{\frac{\log[(s/\beta) - 1]}{s}\right\}\right| \leq \frac{|\log[(s/\beta) - 1]|}{|s|^2} + \frac{1}{|s(s - \beta)|}$$

is integrable while x^s oscillates on the line of integration. Because

$$\frac{d}{d\beta}\left\{\frac{\log[(s/\beta) - 1]}{s}\right\} = \frac{1}{(\beta - s)\beta},$$

differentiation under the integral sign and integration by parts give

$$\begin{aligned} F'(\beta) &= \frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty} \frac{d}{ds}\left[\frac{1}{(\beta - s)\beta}\right] x^s\, ds \\ &= -\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \frac{x^s}{(\beta - s)\beta}\, ds. \end{aligned}$$

This last integral can be evaluated by applying Fourier inversion to the formula

$$\frac{1}{s - \beta} = \int_1^\infty x^{-s}x^{\beta-1}\, dx \qquad [\mathrm{Re}(s - \beta) > 0],$$

$$\frac{1}{a + i\mu - \beta} = \int_0^\infty e^{-i\lambda\mu}e^{\lambda(\beta - a)}\, d\lambda \qquad [a > \mathrm{Re}\,\beta],$$

to obtain

$$\int_{-\infty}^{\infty} \frac{1}{a + i\mu - \beta} e^{i\mu x}\, d\mu = \begin{cases} 2\pi e^{x(\beta - a)} & \text{if } x > 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{1}$$

from which it follows that

$$\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \frac{1}{s - \beta} y^s\, ds = \begin{cases} y^\beta & \text{if } y > 1, \\ 0 & \text{if } y < 1, \end{cases} \tag{2}$$

provided $a >$ Re β. Since $x > 1$ by assumption, this gives $F'(\beta) = x^\beta/\beta$.

Now let C^+ denote the contour in the complex t-plane which consists of the line segment from 0 to $1 - \epsilon$ (where ϵ is a small positive number), followed by the semicircle in the upper halfplane Im $t \geq 0$ from $1 - \epsilon$ to $1 + \epsilon$, followed by the line segment from $1+\epsilon$ to x, and let

$$G(\beta)=\int_{C+}\frac{t^{\beta-1}}{\log t}\,dt.$$

Then

$$G'(\beta)=\int_{C+}t^{\beta-1}\,dt=\frac{t^{\beta}}{\beta}\bigg|_0^x=F'(\beta).$$

Now $G(\beta)$ is defined and analytic for $\operatorname{Re}\beta>0$ (if $\operatorname{Re}\beta<0$, then the integral which defines G diverges at $t=0$) as is $F(\beta)$; hence they differ by a constant (which might depend on x) throughout $\operatorname{Re}\beta>0$. Riemann states that this constant can be evaluated by holding $\operatorname{Re}\beta$ fixed and letting $\operatorname{Im}\beta\longrightarrow+\infty$ in both $F(\beta)$ and $G(\beta)$, but he does not carry out this evaluation.

To evaluate the limit of $G(\beta)$, set $\beta=\sigma+i\tau$, where σ is fixed and $\tau\longrightarrow\infty$. The change of variable $t=e^u$, $u=\log t$ puts $G(\beta)$ in the form

$$\int_{i\delta-\infty}^{i\delta+\log x}\frac{e^{\beta u}}{u}\,du+\int_{i\delta+\log x}^{\log x}\frac{e^{\beta u}}{u}\,du,$$

where the path of integration has been altered slightly using Cauchy's theorem. The changes of variable $u=i\delta+v$ in the first integral and $u=\log x+iw$ in the second put this in the form

$$G(\beta)=e^{i\delta\sigma}e^{-\delta\tau}\int_{-\infty}^{\log x}\frac{e^{\sigma v}}{i\delta+v}e^{i\tau v}\,dv-ix^{\beta}\int_0^{\delta}\frac{e^{-\tau w}e^{\sigma i w}}{\log x+iw}\,dw.$$

Both integrals in this expression approach zero as $\tau\longrightarrow\infty$, the first because $e^{-\delta\tau}\longrightarrow0$ and the second because $e^{-\tau w}\longrightarrow0$ except at $w=0$. Thus *the limit of $G(\beta)$ as $\tau\longrightarrow\infty$ is zero.* (Note, however, that this argument would not be valid if C^+ were changed to follow the lower semicircle because then $e^{-\delta\tau}$ would be replaced by $e^{\delta\tau}$ and $e^{-\tau w}$ would be replaced by $e^{\tau w}$.)

To evaluate the limit of $F(\beta)$ let

$$H(\beta)=\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left\{\frac{\log[1-(s/\beta)]}{s}\right\}x^s\,ds$$

where $a>\operatorname{Re}\beta$ and where $\log[1-(s/\beta)]$ is defined for all complex numbers β other than real numbers $\beta\geq0$ to be $\log(s-\beta)-\log(-\beta)$. The difference $H(\beta)-F(\beta)$ is defined for all complex numbers β other than the real axis, and in the upper halfplane $\operatorname{Im}\beta>0$ it is equal to

$$\begin{aligned}H(\beta)-F(\beta)&=\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left[\frac{\log\beta-\log(-\beta)}{s}\right]x^s\,ds,\\&=\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left[\frac{i\pi}{s}\right]x^s\,ds\\&=-\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\frac{i\pi}{s}x^s\,ds=-i\pi\end{aligned}$$

by the case $\beta = 0$ of (2). Thus $F(\beta) = H(\beta) + i\pi$ throughout the upper halfplane, and it will suffice to evaluate the limit of $H(\beta)$ as $\tau \to \infty$ ($\beta = \sigma + i\tau$). Now $1 - (s/\beta) \to 1$; hence its log goes to zero and it appears plausible therefore that $H(\beta)$ also goes to zero. This can be proved by carrying out the differentiation

$$\frac{d}{ds}\left\{\frac{\log[1-(s/\beta)]}{s}\right\} = -\frac{\log[1-(s/\beta)]}{s^2} + \frac{1}{s(s-\beta)}$$
$$= -\frac{\log[1-(s/\beta)]}{s^2} + \frac{1}{\beta(s-\beta)} - \frac{1}{\beta s},$$

multiplying by $x^s\,ds/2\pi i$, and integrating from $a - i\infty$ to $a + i\infty$ (in the usual sense, namely, the limit as $T \to \infty$ of the integral from $a - iT$ to $a + iT$). Because of the s^2 in the denominator of the first integral, it is not difficult to show, using the Lebesgue bounded convergence theorem (see Edwards [E1]), that the limit of this integral as $\tau \to \infty$ is the integral of the limit, namely, zero. The remaining two integrals can be evaluated using (2) to find they are $x^\beta/\beta - x^0/\beta = (x^\beta - 1)/\beta$. Since the numerator is bounded and $|\beta| \to \infty$, this approaches zero; hence $H(\beta)$ approaches zero and $F(\beta)$ therefore approaches $i\pi$. Hence $F(\beta) = G(\beta) + i\pi$ in the halfplane $\operatorname{Re}\beta > 0$. Thus the desired number $F(1)$ is

$$F(1) = \int_0^{1-\epsilon} \frac{dt}{\log t} + \int_{1-\epsilon}^{1+\epsilon} \frac{(t-1)}{\log t}\cdot\frac{dt}{t-1} + \int_{1+\epsilon}^{x} \frac{dt}{\log t} + i\pi,$$

where the second integral is over the semicircle in the upper halfplane; as $\epsilon \downarrow 0$, the quotient $(t-1)/\log t$ approaches 1 along this semicircle, and hence the integral approaches $\int_{1-\epsilon}^{1+\epsilon} dt/(t-1) = -i\pi$. Thus the limit as $\epsilon \downarrow 0$ of the above formula is

$$F(1) = \operatorname{Li}(x)$$

as was to be shown.

1.15 THE TERM INVOLVING THE ROOTS ρ

Consider next the term in the formula for $J(x)$ arising from the term $\sum \log[1 - (s/\rho)]$ in the formula for $\log \zeta(s)$, namely,

$$(1)\qquad -\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty} \frac{d}{ds}\left\{\frac{\sum \log[1-(s/\rho)]}{s}\right\} x^s\,ds.$$

If the operation of summation over ρ can be interchanged with the differentiation and the integration, *then* this is equal to $-\Sigma H(\rho)$, where $H(\rho)$ is defined as in the preceding section. Now it was shown that $H(\rho) \equiv G(\rho)$ for ρ in the first quadrant ($\operatorname{Re}\rho > 0$, $\operatorname{Im}\rho > 0$) and in exactly the same way it can be shown that for ρ in the fourth quadrant ($\operatorname{Re} \rho > 0$, $\operatorname{Im} \rho < 0$) the value of $H(\rho)$ is equal to the integral $G(\rho)$ except that the integral must be over the contour C^- which goes over the lower semicircle from $1 - \epsilon$ to $1 + \epsilon$ rather than over the upper semicircle as C^+ did. Thus, pairing terms of the sum over ρ in the usual way, the integral (1) would be

$$(2) \qquad -\sum_{\operatorname{Im}\rho>0} \left(\int_{C^+} \frac{t^{\rho-1}}{\log t}\, dt + \int_{C^-} \frac{t^{-\rho}}{\log t}\, dt \right)$$

if it could be evaluated termwise. Now if β is real and positive, then the change of variable $u = t^\beta$, which implies $\log t = \log u/\beta$, $dt/t = du/u\beta$, gives

$$\int_{C^+} \frac{t^{\beta-1}}{\log t}\, dt = \int_0^{x^\beta} \frac{du}{\log u} = \operatorname{Li}(x^\beta) - i\pi,$$

where the second integral is over a path which passes above the singularity at $u = 1$. Since the integral on the left converges throughout the halfplane $\operatorname{Re} \beta > 0$, this formula gives the analytic continuation of $\operatorname{Li}(x^\beta)$ to this half-plane (when x is, as always, a fixed number $x > 1$). In the same way

$$\int_{C^-} \frac{t^{\beta-1}}{\log t}\, dt = \operatorname{Li}(x^\beta) + i\pi,$$

and (2) becomes

$$(3) \qquad -\sum_{\operatorname{Im}\rho>0} [\operatorname{Li}(x^\rho) + \operatorname{Li}(x^{1-\rho})].$$

Thus, if termwise evaluation is valid, the desired integral (1) is equal to (3).

Riemann states that termwise evaluation is valid and that (3) is indeed the desired value (1) but that the series (3) is only conditionally convergent—even though the terms ρ, $1 - \rho$ are paired—and that *it must be summed in the order of increasing*† $\operatorname{Im} \rho$. He concedes that the validity of this termwise evaluation of (1) requires "a more exact discussion of the function ξ," but says that this is "easy" and passes on to the next point.

One other small remark about the sum (3) is necessary. The computations above assume $\operatorname{Re} \rho > 0$, but it has not been shown that this is true for all roots ρ. Although Hadamard later proved that there are no roots ρ on the line $\operatorname{Re} \rho = 0$ (see Section 4.2), Riemann has not excluded this possibility and he is therefore not justified in ignoring the point as he does.

†It is interesting to note that Riemann writes $\rho = \frac{1}{2} + i\alpha$ and says first that the sum (3) is over all *positive* values of α in order of size before then adding parenthetically that it is over all α's with $\operatorname{Re}(\alpha) > 0$ in order of size. Thus he admits, albeit parenthetically, the possibility that the Riemann hypothesis is false.

1.16 THE REMAINING TERMS

One of the three remaining terms in the formula for $J(x)$, namely, the term arising from $(s/2)\log\pi$, drops out when it is divided by s and differentiated with respect to s. The term arising from the constant $\log\xi(0)$ is

$$-\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left(\frac{\log\xi(0)}{s}\right)x^s\,ds$$
$$=\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\frac{\log\xi(0)}{s}x^s\,ds=\log\xi(0)$$

using (2) of Section 1.14 in the case $\beta=0$. Now $\xi(0)=\Pi(0)\pi^{-0}(0-1)\zeta(0)$ $=-\zeta(0)=\frac{1}{2}$ so $\log\xi(0)=-\log 2$ is the numerical value of this term.

Riemann writes $\log\xi(0)$ instead of $-\log 2$, but since he uses ξ to denote a different function—namely, the function $\xi(\frac{1}{2}+it)$ of t—his $\xi(0)$ denotes $\xi(\frac{1}{2})\neq\frac{1}{2}$ and thus his formula is in error. It is hard to guess what the source of this trivial error might be, other than to say that it arises from some confusion between the product formula

$$\xi(s)=\xi(0)\prod_\rho[1-(s/\rho)]$$

in the form it is given above and the product formula

$$\begin{aligned}\xi\left(\frac{1}{2}+it\right)&=\xi(0)\prod\left(1-\frac{\frac{1}{2}+it}{\frac{1}{2}+i\alpha}\right)\\&=\xi(0)\prod\left(\frac{i\alpha-it}{\frac{1}{2}+i\alpha}\right)\\&=\xi(0)\prod\left(\frac{i\alpha}{\frac{1}{2}+i\alpha}\right)\prod\left(1-\frac{it}{i\alpha}\right)\\&=\xi(0)\prod\left(1-\frac{\frac{1}{2}}{\frac{1}{2}+i\alpha}\right)\prod\left(1-\frac{t}{\alpha}\right)\\&=\xi\left(\frac{1}{2}\right)\prod_{\operatorname{Re}\alpha>0}\left(1-\frac{t^2}{\alpha^2}\right)\end{aligned}$$

in the form given by Riemann, and a concomitant confusion of the integral

$$\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left\{\frac{\log[1-(s/\rho)]}{s}\right\}x^s\,ds$$

which he evaluates, with the integral

$$\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left\{\frac{\log[1-(s-\frac{1}{2})/i\alpha]}{s}\right\}x^s\,ds$$

which differs from it by a constant. Whatever the source of the error, Riemann makes the same error in the letter quoted by the editors in the notes which follow the paper in the collected works, and his unpublished papers [R1a] include a computation of $\log\xi(\frac{1}{2})$ to several decimal places, so it was definitely not a typographical error as the editors of the collected works suppose. The error was noticed by Genocchi [G4] during Riemann's lifetime.

This leaves only one term

$$\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left[\frac{\log \Pi(s/2)}{s}\right]x^s\,ds \tag{1}$$

to be evaluated. Now by formula (4) of Section 1.3

$$\log \Pi\left(\frac{s}{2}\right)=\sum_{n=1}^{\infty}\left[-\log\left(1+\frac{s}{2n}\right)+\frac{s}{2}\log\left(1+\frac{1}{n}\right)\right].$$

Using this formula in (1) and assuming that termwise integration is valid puts (1) in the form

$$-\sum_{n=1}^{\infty}\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-i\infty}^{a+i\infty}\frac{d}{ds}\left\{\frac{\log[1+(s/2n)]}{s}\right\}x^s\,ds=-\sum_{n=1}^{\infty}H(-2n),$$

where H is as in Section 1.14. The previous formulas for $H(\beta)$ apply only in the halfplane $\operatorname{Re}\beta>0$. To obtain a formula for H in $\operatorname{Re}\beta<0$ set

$$E(\beta)=-\int_x^{\infty}\frac{t^{\beta-1}}{\log t}\,dt.$$

Then $E(\beta)$ converges for $\operatorname{Re}\beta<0$ and satisfies

$$E'(\beta)=-\int_x^{\infty}t^{\beta-1}\,dt=\frac{x^\beta}{\beta}=F'(\beta)=H'(\beta)$$

so $E(\beta)$ differs from $H(\beta)$ by a constant throughout $\operatorname{Re}\beta<0$. Since both E and H approach zero as $\beta\longrightarrow-\infty$, the constant is zero and $E\equiv H$. Thus (1) becomes

$$\sum_{n=1}^{\infty}\int_x^{\infty}\frac{t^{-2n-1}}{\log t}\,dt=\int_x^{\infty}\frac{1}{t\log t}\left(\sum t^{-2n}\right)dt=\int_x^{\infty}\frac{dt}{t(t^2-1)\log t}$$

provided termwise integration is valid. The proof that termwise integration is valid, which Riemann (tacitly) leaves to the reader, can be given as follows.

Note first that the series

$$\frac{d}{ds}\left[\frac{\log \Pi(s/2)}{s}\right]=-\sum_{n=1}^{\infty}\frac{d}{ds}\left\{\frac{\log[1+(s/2n)]}{s}\right\}$$

converges uniformly in any disk $|s|\le K$. [For large n the series expansion $\log(1+x)=x-\frac{1}{2}x^2+\frac{1}{3}x^3-\cdots$ can be used, and the summand on the right contains only terms in $n^{-2}, n^{-3}, \ldots$.] This justifies the termwise differentiation and also justifies termwise integration over any finite interval

$$\begin{aligned}&\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-iT}^{a+iT}\frac{d}{ds}\left[\frac{\log \Pi(s/2)}{s}\right]x^s\,ds\\&\quad=-\sum_{n=1}^{\infty}\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-iT}^{a+iT}\frac{d}{ds}\left\{\frac{\log[1+(s/2n)]}{s}\right\}x^s\,ds.\end{aligned}\tag{2}$$

To estimate the nth term of the sum on the right set $v=(s-a)/2n$, $b=a/2n$,

$c = T/2n$, so $s = 2n(v + b)$ and the nth term is minus

$$\frac{1}{2\pi i}\frac{1}{\log x}\int_{-ic}^{ic}\frac{d}{2n\,dv}\left[\frac{\log(1+v+b)}{2n(v+b)}\right]x^{2nv+a}2n\,dv$$
$$= \frac{1}{2\pi i}\frac{x^a}{2n\log x}\int_{-ic}^{ic}\frac{d}{dv}\left[\frac{\log(1+v+b)}{v+b}\right]x^{2nv}\,dv.$$

Integration by parts puts this in the form

$$= \frac{1}{2\pi i}\frac{x^a}{2n\log x}\cdot\frac{1}{2n\log x}\left(\left\{\frac{d}{dv}\left[\frac{\log(1+v+b)}{v+b}\right]x^{2nv}\right|_{v=-ic}^{v=ic}\right.$$
$$\left.-\int_{-ic}^{ic}\frac{d^2}{dv^2}\left[\frac{\log(1+v+b)}{v+b}\right]x^{2nv}\,dv\right).$$

Now b is a real number $0 \leq b \leq a$, the function

$$\frac{d}{dv}\left[\frac{\log(1+v+b)}{v+b}\right] = \frac{1}{(v+b)(v+b+1)} - \frac{\log(1+v+b)}{(v+b)^2}$$

is bounded on the imaginary axis, and its derivative is absolutely integrable over $(-i\infty, i\infty)$, from which it follows that the modulus of the nth term of the series on the right side of (2) is at most a constant times n^{-2} *for all* T. Thus the series converges uniformly in T and one can pass to the limit $T \to \infty$ termwise, as was to be shown.

This completes the evaluation of the terms in the formula for $J(x)$. Combining them gives the final result

$$(3)\qquad J(x) = \mathrm{Li}(x) - \sum_{\operatorname{Im}\rho>0}[\mathrm{Li}(x^\rho) + \mathrm{Li}(x^{1-\rho})]$$
$$+ \int_x^\infty \frac{dt}{t(t^2-1)\log t} + \log\xi(0) \qquad (x > 1)$$

which is Riemann's formula [except that, as noted above, $\log\xi(0)$ equals $\log(\frac{1}{2})$ and not $\log\xi(\frac{1}{2})$ as in Riemann's notation it should]. This analytic formula for $J(x)$ is the principal result of the paper.

1.17 THE FORMULA FOR $\pi(x)$

Of course Riemann's goal was to obtain a formula not for $J(x)$ but for the function $\pi(x)$, that is, for the number of primes less than any given magnitude x. Since the number of prime squares less than x is obviously equal to the number of primes less than $x^{1/2}$, that is, equal to $\pi(x^{1/2})$, and since in the same way the number of prime nth powers p^n less than x is $\pi(x^{1/n})$, it follows that J and π are related by the formula

$$(1)\qquad J(x) = \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\pi(x^{1/3}) + \cdots + \frac{1}{n}\pi(x^{1/n}) + \cdots.$$

The series in this formula is finite for any given x because $x^{1/n} < 2$ for n sufficiently large, which implies $\pi(x^{1/n}) = 0$. Riemann inverts this relationship by means of the Möbius inversion formula† (see Section 10.9) to obtain

$$\pi(x) = J(x) - \frac{1}{2}J(x^{1/2}) - \frac{1}{3}J(x^{1/3}) - \frac{1}{5}J(x^{1/5}) + \frac{1}{6}J(x^{1/6}) + \cdots + \frac{\mu(n)}{n}J(x^{1/n}) + \cdots, \tag{2}$$

where $\mu(n)$ is 0 if n is divisible by a prime square, 1 if n is a product of an even number of distinct primes, and -1 if n is a product of an odd number of distinct primes. The series (2) is a finite series for any fixed x and when combined with the analytical formula for $J(x)$

$$J(x) = \operatorname{Li}(x) - \sum_{\rho} \operatorname{Li}(x^{\rho}) - \log 2 + \int_x^{\infty} \frac{dt}{t(t^2-1)\log t} \qquad (x > 1), \tag{3}$$

it gives an analytical formula for $\pi(x)$ as desired.

The formula for $\pi(x)$ which results from substituting (3) in the (finite) series (2) consists of three kinds of terms, namely, those which do not grow as x grows [arising from the last two terms of (3)], those which grow as x grows but which oscillate in sign [the terms arising from $\operatorname{Li}(x^{\rho})$ which Riemann calls "periodic"], and those which grow steadily as x grows [the terms arising from $\operatorname{Li}(x)$]. If all but the last type are ignored, the terms in the formula for $\pi(x)$ are just

$$\operatorname{Li}(x) - \tfrac{1}{2}\operatorname{Li}(x^{1/2}) - \tfrac{1}{3}\operatorname{Li}(x^{1/3}) - \tfrac{1}{5}\operatorname{Li}(x^{1/5}) + \tfrac{1}{6}\operatorname{Li}(x^{1/6}) - \tfrac{1}{7}\operatorname{Li}(x^{1/7}) + \cdots.$$

Now *empirically* this is found to be a good approximation to $\pi(x)$. In fact, the first term alone is essentially Gauss's approximation

$$\pi(x) \sim \int_2^x \frac{dt}{\log t} = \operatorname{Li}(x) - \operatorname{Li}(2)$$

[$\operatorname{Li}(2) = 1.04\ldots$] and the first two terms indicate that

$$\pi(x) \sim \operatorname{Li}(x) - \tfrac{1}{2}\operatorname{Li}(x^{1/2})$$

†Very simply this inversion is effected by performing successively for each prime $p = 2, 3, 5, 7, 11, \ldots$ the operation of replacing the functions $f(x)$ on each side of the equation with the functions $f(x) - (1/p)f(x^{1/p})$. This gives successively

$$J(x) - \tfrac{1}{2}J(x^{1/2}) = \pi(x) + \tfrac{1}{3}\pi(x^{1/3}) + \tfrac{1}{5}\pi(x^{1/5}) + \cdots,$$

$$J(x) - \tfrac{1}{2}J(x^{1/2}) - \tfrac{1}{3}J(x^{1/3}) + \tfrac{1}{6}J(x^{1/6}) = \pi(x) + \tfrac{1}{5}\pi(x^{1/5}) + \tfrac{1}{7}\pi(x^{1/7}) + \cdots,$$

etc., where at each step the sum on the left consists of those terms of the right side of (2) for which the factors of n contain *only* the primes already covered and the sum on the right consists of those terms of the right side of (1) for which the factors of n contain *none* of the primes already covered. Once p is sufficiently large, the latter are all zero except for $\pi(x)$.

which gives, for example,

$$\pi(10^6) \sim 78{,}628 - \tfrac{1}{2}\cdot 178 = 78{,}539$$

which is better than Gauss's approximation and which becomes still better if the third term is used. The extent to which Riemann's suggested approximation

$$\pi(x) \sim \mathrm{Li}(x) + \sum_{n=2}^{\infty} \frac{\mu(n)}{n} \mathrm{Li}(x^{1/n}) \tag{4}$$

is better than $\pi(x) \sim \mathrm{Li}(x)$ is stunningly illustrated by one of Lehmer's tables [L9], an extract of which is given in Table III.

TABLE III[a]

x	Riemann's error	Gauss's error
1,000,000	30	130
2,000,000	−9	122
3,000,000	0	155
4,000,000	33	206
5,000,000	−64	125
6,000,000	24	228
7,000,000	−38	179
8,000,000	−6	223
9,000,000	−53	187
10,000,000	88	339

[a]From Lehmer [L9].

Of course Riemann did not have such extensive empirical data at his disposal, but he seems well aware of the fact that (4) is a better approximation, as well as a more natural approximation, to $\pi(x)$.

Riemann was also well aware, however, of the defects of the approximation (4) and of his analysis of it. Although he has succeeded in giving an exact analytical formula for the error

$$\pi(x) - \sum_{n=1}^{N} \frac{\mu(n)}{n} \mathrm{Li}(x^{1/n}) = \sum_{n=1}^{N} \sum_{\rho} \mathrm{Li}(x^{\rho/n}) + \text{lesser terms}$$

(where N is large enough that $x^{1/(N+1)} < 2$) he has no estimate at all of the size of these "periodic" terms $\sum \sum \mathrm{Li}(x^{\rho/n})$. Actually, the empirical fact that they are as small as Lehmer found them to be is somewhat surprising in view of the fact that the series $\sum [\mathrm{Li}(x^{\rho}) + \mathrm{Li}(x^{1-\rho})]$ is only conditionally convergent—hence the smallness of its sum for any x depends on wholesale cancellation of signs among the terms—and in view of the fact that the individual terms $\mathrm{Li}(x^\rho)$ grow in magnitude like $|x^\rho/\log x^\rho| = x^{\mathrm{Re}\,\rho}/|\rho|\log x$ (see Section 5.5) so that many of them grow at least as fast as $x^{1/2}/\log x \sim 2\,\mathrm{Li}(x^{1/2}) > \mathrm{Li}(x^{1/3})$ and would therefore be expected to be as significant for large x as the term $-\frac{1}{2}\,\mathrm{Li}(x^{1/2})$ and more significant than any of the following terms of (4). On these subjects Riemann restricts himself to the statement that it would be interesting in later counts of primes to study the effect of the particular "periodic" terms on their distribution.

In short, although formulas (2) and (3) combine to give an analytical formula for $\pi(x)$, the validity of the new approximation (4) to $\pi(x)$ to which it leads is based, like that of the old approximation $\pi(x) \sim \mathrm{Li}(x)$, solely on empirical evidence.

1.18 THE DENSITY *dJ*

A simple formulation of the main result

$$J(x) = \mathrm{Li}(x) - \sum_\rho \mathrm{Li}(x^\rho) - \log 2 + \int_x^\infty \frac{dt}{t(t^2-1)\log t} \tag{1}$$

can be obtained by differentiating to find

(2)
$$dJ = \left[\frac{1}{\log x} - \sum_{\mathrm{Re}\,\alpha>0} \frac{2\cos(\alpha\log x)}{x^{1/2}\log x} - \frac{1}{x(x^2-1)\log x}\right] dx \qquad (x > 1),$$

where α ranges over all values such that $\rho = \frac{1}{2} + i\alpha$—in other words $\alpha = -i(\rho - \frac{1}{2})$, where ρ ranges over the roots—so that

$$x^{\rho-1} + x^{-\rho} = x^{-1/2}[x^{i\alpha} + x^{-i\alpha}] = 2x^{-1/2}\cos(\alpha\log x).$$

[The Riemann hypothesis is that the α's are all real. In writing formula (2) in this form Riemann is clearly thinking of the α's as being real since otherwise the natural form would be $x^{\rho-1} + x^{\bar\rho-1} = 2x^{\beta-1}\cos(\gamma\log x)$, where $\rho = \beta + i\gamma$.]

By the definition of J, the measure dJ is dx times the density of primes plus $\frac{1}{2}$ the density of prime squares, plus $\frac{1}{3}$ the density of prime cubes plus, etc. Thus $1/\log x$ should be considered to be an approximation not to the density of primes as Gauss suggested but rather to dJ, that is, to the density of primes *plus* $\frac{1}{2}$ the density of prime squares, plus, etc.

Given two large numbers $a < b$ the approximation obtained by taking a finite number of the α's

$$J(b) - J(a) \sim \int_a^b \frac{dt}{\log t} - 2\sum \int_a^b \frac{\cos(\alpha\log t)\,dt}{t^{1/2}\log t} \tag{3}$$

should be a fairly good approximation because the omitted term $\int dx/x(x^2-1)$

log x is entirely negligible and because the integrals involving the large α's oscillate very rapidly for large x and therefore should make very small contributions. In fact, the basic formula (1) implies immediately that the error in (3) approaches the negligible omitted term as more and more of the α's are included in the sum.

It is in the sense of investigating the number of α's which are significant in (3) that Riemann meant to investigate empirically the influence of the "periodic" terms on the distribution of primes. So far as I know, no such investigation has ever been carried out.

1.19 QUESTIONS UNRESOLVED BY RIEMANN

Riemann himself, in a letter quoted in the notes which follow this paper in his collected works, singles out two statements of the paper as not having been fully proved as yet, namely, the statement that the equation $\xi(\frac{1}{2} + i\alpha) = 0$ has approximately $(T/2\pi)\log(T/2\pi)$ *real* roots α in the range $0 < \alpha < T$ and the statement that the integral of Section 1.15 can be evaluated termwise. He expresses no doubt about the truth of these statements, however, and says that they follow from a new representation of the function ξ which he has not yet simplified sufficiently to publish. Nonetheless, as was stated in Section 1.9, the first of these two statements—at least if it is understood to mean that the relative error in the approximation approaches zero as $T \to \infty$—has never been proved. The second was proved by von Mangoldt in 1895, but by a method completely different from that suggested by Riemann, namely by proving first that Riemann's formula for $J(x)$ is valid and by concluding from this that the termwise value of the integral in Section 1.15 must be correct.

Riemann evidently believed that he had given a proof of the product formula for $\xi(s)$, but, at least from the reading of the paper given above, one cannot consider his proof to be complete, and, in particular, one must question Riemann's estimate of the number of roots ρ in the range $\{0 \leq \operatorname{Im} \rho \leq T\}$ on which this proof is based. It was not until 1893 that Hadamard proved the product formula, and not until 1905 that von Mangoldt proved the estimate of the number of roots in $\{0 \leq \operatorname{Im} \rho \leq T\}$.

Next, the original question of the validity of the approximation $\pi(x) \sim \int_2^x (dt/\log t)$ remained entirely unresolved by Riemann's paper. It can be shown that the relative error of this approximation approaches zero as $x \to \infty$ if and only if the same is true of the relative error in Riemann's approximation $J(x) \sim \operatorname{Li}(x)$, so the original question is equivalent to the question of whether $\sum \operatorname{Li}(x^\rho)/\operatorname{Li}(x) \to 0$, but this unfortunately does not bring the problem any nearer to a solution. It was not until 1896 that Hadamard and, independently, de la Vallée Poussin proved the prime number theorem to the effect that the relative error in $\pi(x) \sim \int_2^x (dt/\log t)$ does approach zero as $x \to \infty$.

Finally, the paper raised a question much greater than any question it answered, the question of the truth or falsity of the Riemann hypothesis.

The remainder of this book is devoted to the subsequent history of these six questions. In summary, they are as follows:

(a) Is Riemann's estimate of the number of roots ρ on the line segment from $\frac{1}{2}$ to $\frac{1}{2} + iT$ correct as $T \to \infty$? (Unknown.)

(b) Is termwise evaluation of the integral of Section 1.15 valid? (Yes, von Mangoldt, 1895.)

(c) Is the product formula for $\xi(s)$ valid? (Yes, Hadamard, 1893.)

(d) Is Riemann's estimate of the number of roots ρ in the strip $\{0 \leq \operatorname{Im} \rho \leq T\}$ correct? (Yes, von Mangoldt, 1905.)

(e) Is the prime number theorem true? [Yes, Hadamard and de la Vallée Poussin (independently), 1896.]

(f) Is the Riemann hypothesis true? (Unknown.)

Chapter 2

The Product Formula for ξ

2.1 INTRODUCTION

In 1893 Hadamard published a paper [H1] in which he studied entire functions (functions of a complex variable which are defined and analytic at all points of the complex plane) and their representations as infinite products. One consequence of the general theory which he developed in this paper is the fact that the product formula

$$\xi(s) = \xi(0) \prod_{\rho} \left(1 - \frac{s}{\rho}\right) \tag{1}$$

is valid; here ξ is the entire function defined in Section 1.8, ρ ranges over all roots ρ of $\xi(\rho) = 0$, and the infinite product is understood to be taken in an order which pairs each root ρ with the corresponding root $1 - \rho$. Hadamard's proof of the product formula for ξ was called by von Mangoldt [M1] "the first real progress in the field in 34 years," that is, the first since Riemann's paper.

This chapter is devoted to the proof of the product formula (1). Since only the specific function ξ is of interest here, Hadamard's methods for general entire functions can, of course, be considerably specialized and simplified† for this case, and in the end the proof which results is closer to the one outlined by Riemann than to Hadamard's proof. The first step of the proof is to make an estimate of the distribution of the roots ρ. This estimate, which is that the number of roots ρ in the disk $|\rho - \frac{1}{2}| < R$ is less than a constant times $R \log R$ as $R \longrightarrow \infty$, is based on Jensen's theorem and is much less exact than Riemann's estimate that the number of roots in the strip $\{0 < \mathrm{Im}$

†A major simplification is the use of Jensen's theorem, which was not known at the time Hadamard was writing.

$\rho < T\}$ is $(T/2\pi)\log(T/2\pi) - (T/2\pi)$ with a relative error which is of the order of magnitude of T^{-1}. It is exact enough, however, to prove the convergence of the product (1). Once it has been shown that this product converges, the rest of the proof can be carried out more or less as Riemann suggests.

2.2 JENSEN'S THEOREM

Theorem Let $f(z)$ be a function which is defined and analytic throughout a disk $\{|z| \le R\}$. Suppose that $f(z)$ has no zeros on the bounding circle $|z| = R$ and that inside the disk it has the zeros $z_1, z_2, \ldots, z_n$ (where a zero of order k is included k times in the list). Suppose, finally, that $f(0) \neq 0$. Then

$$\log\left| f(0)\cdot\frac{R}{z_1}\cdot\frac{R}{z_2}\cdot\cdots\cdot\frac{R}{z_n}\right| = \frac{1}{2\pi}\int_0^{2\pi}\log|f(Re^{i\theta})|\,d\theta. \tag{1}$$

***Proof*†** If $f(z)$ has no zeros inside the disk, then the equation is merely

$$\log|f(0)| = \frac{1}{2\pi}\int_0^{2\pi}\log|f(Re^{i\theta})|\,d\theta; \tag{2}$$

that is, the equation is the statement that the value of $\log|f(z)|$ at the center of the disk is equal to its average value on the bounding circle. This can be proved either by observing that $\log|f(z)|$ is the real part of the analytic function $\log f(z)$ and is therefore a harmonic function, or by taking the real part of the Cauchy integral formula

$$\log f(0) = \frac{1}{2\pi i}\int_{|z|=R}\frac{\log f(z)}{z}\,dz = \frac{1}{2\pi}\int_0^{2\pi}\log f(Re^{i\theta})\,d\theta,$$

where $\log f(z)$ is defined in the disk to be

$$\log|f(0)| + \int_0^z \frac{f'(t)}{f(t)}\,dt.$$

Applying this formula (2) to the function

$$F(z) = f(z)\frac{R^2 - \bar{z}_1 z}{R(z - z_1)}\cdot\frac{R^2 - \bar{z}_2 z}{R(z - z_2)}\cdots\frac{R^2 - \bar{z}_n z}{R(z - z_n)}$$

gives

$$\log|F(0)| = \frac{1}{2\pi}\int_0^{2\pi}\log|F(Re^{i\theta})|\,d\theta$$

because $F(z)$ is analytic and has no zeros in the disk. But this is the formula

†See Ahlfors [A3]. This method of proof of Jensen's threorem is to be found in Backlund's 1918 paper on the Lindelöf hypothesis [B3] (see Section 9.4).

of Jensen's theorem (1) because

$$\left|\frac{R^2 - \bar{z}_j \cdot 0}{R(0 - z_j)}\right| = \left|\frac{R}{z_j}\right|$$

and because by a basic formula in the theory of conformal mapping

$$\left|\frac{R^2 - \bar{z}_j z}{R(z - z_j)}\right| = 1 \qquad \text{when} \quad |z| = R.$$

(To prove this formula multiply the numerator by $\bar{z}/R$. This does not change the modulus if $|z| = R$ and it makes the numerator into the complex conjugate of the denominator.) This completes the proof of Jensen's theorem (1).

2.3 A SIMPLE ESTIMATE OF $|\xi(s)|$

Theorem For all sufficiently large values of R the estimate $|\xi(s)| \le R^R$ holds throughout the disk $|s - \frac{1}{2}| \le R$.

Proof It was shown in Section 1.8 that $\xi(s)$ can be expanded as a power series in $(s - \frac{1}{2})$:

$$\xi(s) = a_0 + a_2(s - \tfrac{1}{2})^2 + \cdots + a_{2n}(s - \tfrac{1}{2})^{2n} + \cdots,$$

where

$$a_{2n} = 4\int_1^\infty \frac{d}{dx}[x^{3/2}\psi'(x)]\, x^{-1/4}\frac{(\frac{1}{2}\log x)^{2n}}{(2n)!}\, dx.$$

The fact that the coefficients a_n are *positive* follows immediately from

$$\begin{aligned}\frac{d}{dx}[x^{3/2}\psi'(x)] &= \frac{d}{dx}\left(-\sum_{n=1}^\infty x^{3/2}n^2\pi e^{-n^2\pi x}\right)\\ &= \sum_{n=1}^\infty \left(n^4\pi^2 x - \frac{3}{2}n^2\pi\right)x^{1/2}e^{-n^2\pi x}\end{aligned}$$

because this shows that the integrand in the integral for a_{2n} is positive for $x \ge 1$. Thus the largest value of $\xi(s)$ on the disk $|s - \frac{1}{2}| \le R$ occurs at the point $s = \frac{1}{2} + R$, and to prove the theorem it suffices to show that $\xi(\frac{1}{2} + R) \le R^R$ for all sufficiently large R. Now

$$\xi(s) = \Pi(s/2)\pi^{-s/2}(s-1)\zeta(s)$$

and $\zeta(s)$ decreases to 1 as $s \to +\infty$, so if R is given and if N is chosen so that $\frac{1}{2} + R \le 2N < \frac{1}{2} + R + 2$, it follows that

$$\begin{aligned}\xi(\tfrac{1}{2} + R) \le \xi(2N) &= (N!)\pi^{-N}(2N-1)\zeta(2N)\\ &\le N^N\pi^{-0}(2N)\zeta(2)\\ &= \text{const } N^{N+1}\\ &\le \text{const } (\tfrac{1}{2}R + 2)^{(R/2)+3} < R^R\end{aligned}$$

for all sufficiently large R, which completes the proof of the theorem.

2.4 THE RESULTING ESTIMATE OF THE ROOTS ρ

Theorem Let $n(R)$ denote the number of roots ρ of $\xi(\rho) = 0$ which lie inside or on the circle $|s - \frac{1}{2}| = R$ (counted with multiplicities). Then $n(R) \leq 2R \log R$ for all sufficiently large R.

Proof Jensen's theorem applied to $\xi(s)$ on the disk $|s - \frac{1}{2}| \leq 2R$ gives

$$\log \xi\left(\frac{1}{2}\right) + \sum_{|\rho-1/2|<2R} \log \frac{2R}{|\rho - \frac{1}{2}|} \leq \log[(2R)^{2R}].$$

The terms of the sum over ρ are all positive and the terms corresponding to roots ρ inside the circle $|\rho - \frac{1}{2}| \leq R$ are all at least $\log 2$; hence,

$$\begin{aligned} n(R) \log 2 &\leq 2R \log 2R - \log \xi(\tfrac{1}{2}) \\ n(R) &\leq \frac{2}{\log 2} R \log R + 2\mathrm{R} - \frac{\log \xi(\frac{1}{2})}{\log 2} \\ &\leq 2R \log R \end{aligned}$$

for all sufficiently large R, as was to be shown. If there are roots ρ on the circle $|s - \frac{1}{2}| = 2R$, so that Jensen's theorem is not applicable, one can apply the above to the circle with radius $R + \epsilon$ and let $\epsilon \to 0$.

2.5 CONVERGENCE OF THE PRODUCT

As was noted in Section 1.10, in order to prove the convergence of the product

$$\prod \left(1 - \frac{s}{\rho}\right) = \prod_{\operatorname{Im} \rho > 0} \left[1 - \frac{s(1 - s)}{\rho(1 - \rho)}\right] \tag{1}$$

for all s, it suffices to prove the convergence of the sum $\sum |\rho(1 - \rho)|^{-1}$. Since all but a finite number of roots ρ satisfy the inequality

$$\frac{1}{|\rho(1 - \rho)|} = \frac{1}{|(\rho - \frac{1}{2})^2 - \frac{1}{4}|} < \frac{1}{|\rho - \frac{1}{2}|^2},$$

it suffices therefore to prove the convergence of the sum $\sum |\rho - \frac{1}{2}|^{-2}$; here the sum can be considered either as a sum over roots ρ in the upper halfplane $\operatorname{Im} \rho > 0$ or as a sum over all roots since the first of these is merely twice the second. The convergence of the product (1) is therefore a consequence of the case $\epsilon = 1$ of the following theorem.

Theorem For any given $\epsilon > 0$ the series

$$\sum \frac{1}{|\rho - \frac{1}{2}|^{1+\epsilon}}$$

converges, where ρ ranges over all roots ρ of $\xi(\rho) = 0$.

[Note that this theorem would follow immediately from Riemann's observation that the vertical density of the roots ρ is a constant times $(\log T)\,dT$ and from the fact that $\int^{\infty} T^{-1-\epsilon}(\log T)\,dT$ converges. This is Riemann's first step in his "proof" of the product formula for ξ.]

Proof Let the roots ρ be numbered $\rho_1, \rho_2, \rho_3, \ldots$ in order of increasing $|\rho - \frac{1}{2}|$. Furthermore, let $R_1, R_2, R_3, \ldots$ be the sequence of positive real numbers defined implicitly by the equation $3R_n \log R_n = n$. Then by the theorem of the preceding section there are at most $2n/3$ roots ρ inside the circle $|s - \frac{1}{2}| = R_n$; hence the nth root is not in this circle, that is, $|\rho_n - \frac{1}{2}| > R_n$. Thus

$$\sum \frac{1}{|\rho_n - \frac{1}{2}|^{1+\epsilon}} \le \sum \frac{1}{R_n^{1+\epsilon}} = \sum \frac{(3 \log R_n)^{1+\epsilon}}{n^{1+\epsilon}}$$
$$= \sum \frac{1}{n^{1+(\epsilon/2)}} \cdot \frac{(3 \log R_n)^{1+\epsilon}}{n^{\epsilon/2}}.$$

Now $\log n = \log R_n + \log 3 + \log \log R_n > \log R_n$. Hence $(3 \log R_n)^{1+\epsilon} < 9 (\log n)^2 < n^{\epsilon/2}$ for all sufficiently large n and

$$\sum \frac{1}{|\rho - \frac{1}{2}|^{1+\epsilon}} < \text{const} + \sum \frac{1}{n^{1+(\epsilon/2)}} < \infty$$

as was to be shown.

2.6 RATE OF GROWTH OF THE QUOTIENT

Riemann states that $\log \xi(s) - \sum \log [1 - (s/\rho)]$ grows no faster than $s \log s$, from which he concludes, since it is an even function, that it must be a constant. In this section the weaker result that the growth of its real part is no faster than $|s|^{1+\epsilon}$ will be proved. This still permits one to conclude, as will be shown in the next section, that it is constant.

Theorem Let $\epsilon > 0$ be given. Then

$$\operatorname{Re} \log \frac{\xi(s)}{\prod_\rho \left(1 - \frac{s - \frac{1}{2}}{\rho - \frac{1}{2}}\right)} \le \left| s - \frac{1}{2} \right|^{1+\epsilon}$$

for all sufficiently large $|s - \frac{1}{2}|$.

Proof Let R be given and let the function being estimated be written as a sum of two functions

$$\operatorname{Re} \log \frac{\xi(s)}{\prod \left(1 - \frac{s - \frac{1}{2}}{\rho - \frac{1}{2}}\right)} = u_R(s) + v_R(s),$$

where

$$u_R(s) = \text{Re}\log \frac{\xi(s)}{\prod_{|\rho-1/2|\le 2R}\left(1 - \frac{s-\frac{1}{2}}{\rho-\frac{1}{2}}\right)}$$

$$v_R(s) = \text{Re}\log \frac{1}{\prod_{|\rho-1/2|>2R}\left(1 - \frac{s-\frac{1}{2}}{\rho-\frac{1}{2}}\right)}.$$

These logarithms are defined only up to multiples of $2\pi i$, but their real parts are well defined except at the points $s = \rho$ for $|\rho - \frac{1}{2}| > 2R$ (at which points u_R is $-\infty$ and v_R is $+\infty$). It will suffice to show that for large R both $u_R(s)$ and $v_R(s)$ are at most $R^{1+\epsilon}$ on $|s - \frac{1}{2}| = R$ since then, when ϵ is decreased slightly to ϵ', it follows that $u_R(s) + v_R(s) \le 2R^{1+\epsilon'} \le |s - \frac{1}{2}|^{1+\epsilon}$ on $|s - \frac{1}{2}| = R$ for all R large enough that $u_R \le R^{1+\epsilon'}$, $v_R \le R^{1+\epsilon'}$, and $2 \le R^{\epsilon-\epsilon'}$.

First consider $u_R(s)$. On the circle $|s - \frac{1}{2}| = 4R$ the factors in the denominator are all at least 1; therefore

$$\begin{aligned} u_R(s) &\le \text{Re}\log\xi(s) = \log|\xi(s)| \\ &\le \log[(4R)^{4R}] = 4R\log 4R \le R^{1+\epsilon} \end{aligned}$$

on the circle $|s - \frac{1}{2}| = 4R$, for large R (large enough that $4\log 4R < R^{\epsilon}$). Now u_R is a harmonic function on the disk $|s - \frac{1}{2}| \le 4R$ except at the points $s = \rho$ in the range $2R < |s - \frac{1}{2}| \le 4R$. But near these singular points $s = \rho$ the value of u_R is near $-\infty$, so the maximum value of the harmonic function u_R on the disk $|s - \frac{1}{2}| \le 4R$ must occur on the outer boundary $|s - \frac{1}{2}| = 4R$. Thus the maximum of u_R on the disk, and in particular on the circle $|s - \frac{1}{2}| = R$, is at most $R^{1+\epsilon}$ as was to be shown.

Now consider $v_R(s)$. For complex x in the disk $|x| \le \frac{1}{2}$ the inequality

$$\begin{aligned} \text{Re}\log\frac{1}{1-x} &= -\text{Re}\log(1-x) = \text{Re}\int_0^x \frac{dt}{1-t} \\ &\le \left|\int_0^x \frac{dt}{1-t}\right| \le |x|\max\frac{1}{|1-t|} = 2|x| \end{aligned}$$

holds. Thus for $|s - \frac{1}{2}| = R$ the inequality

$$\begin{aligned} v_R(s) &= \text{Re}\log\frac{1}{\prod_{|\rho-1/2|>2R}\left(1 - \frac{(s-\frac{1}{2})^2}{(\rho-\frac{1}{2})^2}\right)} \\ &\le 2\sum_{|\rho-1/2|>2R} \frac{R^2}{|\rho-\frac{1}{2}|^2} \\ &= 2\sum\left(\frac{R}{|\rho-\frac{1}{2}|}\right)^{1-\epsilon}\left(\frac{R}{|\rho-\frac{1}{2}|}\right)^{1+\epsilon} \\ &\le 2\sum\left(\frac{1}{2}\right)^{1-\epsilon}\frac{R^{1+\epsilon}}{|\rho-\frac{1}{2}|^{1+\epsilon}} \\ &= 2^{\epsilon}R^{1+\epsilon}\sum_{|\rho-1/2|>2R}\frac{1}{|\rho-\frac{1}{2}|^{1+\epsilon}} \end{aligned}$$

holds. Now the sum in this expression converges by the theorem of Section 2.5, and it decreases to zero as R increases. Thus $v_R(s) \leq R^{1+\epsilon}$ on $|s - \frac{1}{2}| = R$ for all sufficiently large R as was to be shown. This completes the proof.

2.7 RATE OF GROWTH OF EVEN ENTIRE FUNCTIONS

Theorem Let $f(s)$ be an analytic function, defined in the entire s-plane, which is even in the sense that $f(-s) \equiv f(s)$ and which grows more slowly than $|s|^2$ in the sense that for every $\epsilon > 0$ there is an R such that $\operatorname{Re} f(s) < \epsilon |s|^2$ at all points s satisfying $|s| \geq R$. Then f must be constant.

Proof The subtle point of the theorem is that only the *upward* growth of the *real part* of f is limited. The main step in the proof is the following lemma, which shows that this implies that the growth of the *modulus* of f is also limited.

Lemma Let $f(s)$ be an analytic function on the disk $\{|s| \leq r\}$, let $f(0) = 0$, and let M be the maximum value of $\operatorname{Re} f(s)$ on the bounding circle $|s| = r$ (and hence on the entire disk). Then for $r_1 < r$ the modulus of f on the smaller disk $\{|s| \leq r_1\}$ is bounded by

$$|f(s)| \leq 2r_1 M/(r - r_1) \qquad (|s| \leq r_1).$$

Proof of the Lemma Consider the function

$$\phi(s) = f(s)/s[2M - f(s)].$$

If $u(s)$ and $v(s)$ denote the real and imaginary parts, respectively, of f, then $|2M - u(s)| \geq M \geq u(s)$ on the circle $|s| = r$; so the modulus of ϕ on this circle is at most

$$|\phi(s)| = \frac{(u^2 + v^2)^{1/2}}{r[(2M - u)^2 + v^2]^{1/2}} \leq \frac{(u^2 + v^2)^{1/2}}{r(u^2 + v^2)^{1/2}} = \frac{1}{r}$$

which implies that $|\phi(s)| \leq r^{-1}$ throughout the disk $\{|s| \leq r\}$. But $f(s)$ can be expressed in terms of $\phi(s)$ as

$$\phi(s)s[2M - f(s)] = f(s), \qquad f(s) = \frac{2Ms\phi(s)}{1 + s\phi(s)}$$

which shows that for $|s| = r_1$ the modulus of $f(s)$ is at most

$$|f(s)| \leq 2Mr_1 r^{-1}/(1 - r_1 r^{-1}) = 2Mr_1/(r - r_1).$$

Hence the same inequality holds throughout the disk $\{|s| \leq r_1\}$ as was to be shown.

Now to prove the theorem let $f(s) = \sum_{n=0}^{\infty} a_n s^n$ be the power series expansion of a function $f(s)$ satisfying the conditions of the theorem. Note first that it can be assumed without loss of generality that $a_0 = 0$ because $f(s)$ satisfies the growth condition of the theorem if and only if $f(s) - f(0)$ does and because $f(s)$ is constant if and only if $f(s) - f(0)$ is. Now Cauchy's integral formula for the coefficients is

$$a_n = \frac{1}{2\pi i}\int_{\partial D}\frac{f(s)\,ds}{s^{n+1}},$$

where D is any domain containing the origin. Let ϵ, R be as in the statement of the theorem and let D be the disk $\{|s| \le \frac{1}{2}R\}$. Then the above formula gives

$$\begin{aligned}|a_n| &= \left|\frac{1}{2\pi i}\int_0^{2\pi}\frac{f(\frac{1}{2}\mathrm{Re}^{i\theta})}{(\frac{1}{2}\mathrm{Re}^{i\theta})^n}i\,d\theta\right| \\ &\le \frac{1}{2\pi}\int_0^{2\pi}\frac{2^n\,|f(\frac{1}{2}\mathrm{Re}^{i\theta})|}{R^n}\,d\theta.\end{aligned}$$

The right side is the average value of a function whose value is by the lemma at most

$$\frac{2^n}{R^n}\frac{2(\epsilon R^2)(\frac{1}{2}R)}{R - (\frac{1}{2}R)} = \frac{2^{n+1}\epsilon}{R^{n-2}}.$$

If $n \ge 2$, this is at most $2^{n+1}\epsilon$, and since ϵ is arbitrary, a_n must be zero for $n \ge 2$. Thus $f(s) = a_1 s$. However a_1 must be zero by the evenness condition $f(s) \equiv f(-s)$. Therefore $f(s) \equiv 0$ which is constant, as was to be shown.

2.8 THE PRODUCT FORMULA FOR ξ

The function $F(s) = \xi(s)/\prod_\rho [1 - (s - \frac{1}{2})/(\rho - \frac{1}{2})]$ is analytic in the entire s-plane and is an even function of $s - \frac{1}{2}$. Moreover, it has no zeros, so its logarithm is well defined up to an additive constant $2\pi ni$ (n an integer) by the formula $\log F(s) = \int_0^s F'(z)\,dz/F(z) + \log F(0)$, where $\log F(0)$ is determined to within an additive constant $2\pi ni$. The results of the preceding two sections then combine to give $\log F(s) =$ const, and therefore upon exponentiation

$$\xi(s) = c\prod\left(1 - \frac{s - \frac{1}{2}}{\rho - \frac{1}{2}}\right),$$

where c is a constant. Dividing this by the particular value

$$\xi(0) = c\prod\left(1 - \frac{-\frac{1}{2}}{\rho - \frac{1}{2}}\right)$$

gives

$$\frac{\xi(s)}{\xi(0)} = \prod\left(1 - \frac{s - \frac{1}{2}}{\rho - \frac{1}{2}}\right)\left(1 - \frac{-\frac{1}{2}}{\rho - \frac{1}{2}}\right)^{-1}.$$

The factors on the right are linear functions of s which are 0 when $s = \rho$ and 1 when $s = 0$; hence they are $1 - (s/\rho)$ and the formula is the desired formula

$$\xi(s) = \xi(0) \prod_{\rho} \left(1 - \frac{s}{\rho}\right),$$

where, as always, it is understood that the factors ρ and $1 - \rho$ are paired.†

†The same argument proves the validity of the product formula for the sine

$$\sin \pi s = \pi s \prod_{n=1}^{\infty} \left(1 - \frac{s^2}{n^2}\right)$$

mentioned in Section 1.3. The only other unproved statement in Section 1.3 which is not elementary is the equivalence of the two definitions (2) and (3) of $\Pi(s)$.

Chapter 3

Riemann's Main Formula

3.1 INTRODUCTION

Soon after Hadamard proved the product formula for $\xi(s)$, von Mangoldt [M1] proved Riemann's main formula

$$J(x) = \mathrm{Li}(x) - \sum_\rho \mathrm{Li}(x^\rho) - \log 2 + \int_x^\infty \frac{dt}{t(t^2-1)\log t} \qquad (x > 1). \tag{1}$$

Von Mangoldt also recast this formula in a simpler form which has virtually replaced Riemann's original statement (1) in the subsequent development of the theory. This simpler form of (1) can be derived as follows.

The essence of Riemann's derivation of (1) is the inversion of the relationship

$$\log \zeta(s) = \int_0^\infty x^{-s}\, dJ(x) \tag{2}$$

for $J(x)$ in terms of $\log \zeta(s)$ and then the use of

$$\Pi\left(\frac{s}{2}\right)\pi^{-s/2}(s-1)\zeta(s) = \frac{1}{2}\prod_\rho\left(1 - \frac{s}{\rho}\right) \tag{3}$$

to express $\log \zeta(s)$ in terms of elementary functions and in terms of the roots ρ. Now the function $\log \zeta(s)$ has logarithmic singularities at all the roots ρ, and as a function of a complex variable it is very awkward outside the half-plane $\mathrm{Re}\, s > 1$. On the other hand, its derivative $\zeta'(s)/\zeta(s)$ is analytic in the entire plane except for poles at the roots ρ, the pole 1, and the zeros $-2n$. This might well lead one to begin not with formula (2) but with its derivative

$$\frac{\zeta'(s)}{\zeta(s)} = -\int_0^\infty x^{-s}(\log x)\, dJ(x). \tag{4}$$

The measure $(\log x)\, dJ(x)$ is a point measure which assigns the weight $\log(p^n)\cdot(1/n)$ to prime powers p^n and the weight 0 to all other points. Thus it can be written as a Stieltjes measure $d\psi(x)$, where $\psi(x)$ is the step function which begins at 0 and has a jump of $\log(p^n)\cdot(1/n) = \log p$ at each prime power p^n. In other words

$$\psi(x) = \sum_{p^n < x} \log p$$

except when x is a prime power; at the jumps $x = p^n$ the value of ψ is defined, as usual, to be halfway between the new and old values $\psi(x) = \frac{1}{2}[\psi(x - \epsilon) + \psi(x + \epsilon)]$. This function $\psi(x)$ had already been considered by Chebyshev,† who named it $\psi(x)$ and who proved‡ among other things that the prime number theorem is essentially the same as the statement that $\psi(x) \sim x$ with a relative error which approaches zero as $x \to \infty$. In terms of $\psi(x)$ formula (4) becomes

$$-\frac{\zeta'(s)}{\zeta(s)} = \int_0^\infty x^{-s}\, d\psi(x). \tag{5}$$

In other words, if J is replaced by ψ in the original formula (2), then the awkward function $\log \zeta(s)$ is replaced by the more tractable function $-\zeta'(s)/\zeta(s)$.

Now the argument by which Riemann went from formula (2) to the formula (1) for $J(x)$ can be applied equally well to go from formula (5) to a new formula for $\psi(x)$. The explicit computations of this argument are given in the next section. However, even without the explicit computations, one can guess the formula for $\psi(x)$ as follows. The simplest formulation of Riemann's result is his formula (see Section 1.18)

$$dJ = \left(\frac{1}{\log x} - \sum_\rho \frac{x^{\rho-1}}{\log x} - \frac{1}{x(x^2-1)\log x}\right) dx \qquad (x > 1)$$

which gives

$$\begin{aligned} d\psi &= (\log x)\, dJ \\ &= \left(1 - \sum_\rho x^{\rho-1} - \sum_n x^{-2n-1}\right) dx \qquad (x > 1) \end{aligned}$$

and leads to the guess

$$\psi(x) = x - \sum_\rho \frac{x^\rho}{\rho} + \sum_n \frac{x^{-2n}}{2n} + \text{const} \qquad (x > 1). \tag{6}$$

This is von Mangoldt's reformulation of (1) referred to in the first paragraph. [The value of the constant is given in Section 3.2. It is assumed in von Mangoldt's formula (6), as it is in Riemann's formula (1), that the terms of the sum over ρ are taken in the order of increasing $|\text{Im}\, \rho|$; these sums converge

†This work of Chebyshev [C3] in 1850 preceded Riemann's paper.

‡Actually Chebyshev does not state this result explicitly, but it follows trivially from the techniques he introduces for deducing estimates of π from estimates of ψ.

only conditionally—even when the terms ρ, $1-\rho$ are paired—so their order is essential.]

The main part of this chapter is devoted to von Mangoldt's proof of the formula (6) for $\psi(x)$. In Section 3.2 the formula is derived from the termwise evaluation of certain definite integrals, and in the following three sections the validity of these termwise evaluations is rigorously proved. Von Mangoldt's proof of Riemann's original formula (1) is outlined in the next two sections, and the last section deals with the numerical evaluation of the constant $\zeta'(0)/\zeta(0)$.

3.2 DERIVATION OF VON MANGOLDT'S FORMULA FOR $\psi(x)$

The technique of Section 1.12 applied to the formula $-\zeta'(s)/\zeta(s) = s\int_0^\infty \psi(x)x^{-s-1}\,dx$ instead of to the formula $\log\zeta(s) = s\int_0^\infty J(x)x^{-s-1}\,dx$ puts $\psi(x)$ in the form of a definite integral

$$\text{(1)} \qquad \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\left[-\frac{\zeta'(s)}{\zeta(s)}\right]x^s\frac{ds}{s} \qquad (a>1).$$

Von Mangoldt's method of proving the formula for $\psi(x)$ is to evaluate this definite integral in two different ways, one of which gives the value $\psi(x)$ and the other $x - \sum(x^\rho/\rho) + \sum(x^{-2n}/2n) + \text{const.}$ (Neither of these evaluations uses Fourier's theorem, so the use of Fourier's theorem in Section 1.12 can be regarded as purely heuristic.)

The first method of evaluating the definite integral (1) is as follows. Beginning with the formula†

$$\text{(2)} \qquad -\frac{\zeta'(s)}{\zeta(s)} = \int_0^\infty x^{-s}\,d\psi(x),$$

let $\Lambda(n)$ denote the weight assigned to the integer n by the measure $d\psi$—that is, $\Lambda(n)$ is zero unless n is a prime power, in which case $\Lambda(n)$ is the log of the prime of which n is a power—so that the integral (2) can be written as a sum

$$\text{(3)} \qquad -\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=2}^\infty \Lambda(n)n^{-s} \qquad (\operatorname{Re} s > 1).$$

Substitute this formula in (1) and assume termwise integration is valid. This gives the value

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

†This formula is essentially the Euler product formula (see the concluding remarks of Section 1.11). In fact, logarithmic differentiation of $\zeta(s) = \prod(1-p^{-s})^{-1}$ gives (3) immediately.

for (1). Now formula (2) of Section 1.14 (with $\beta = 0$) shows that the integral corresponding to n in this sum is 1 if $x/n > 1$ and 0 if $x/n < 1$; hence the sum is just

$$\sum_{n<x} \Lambda(n) = \psi(x)$$

as was to be shown.

To justify this sequence of steps leading to the value $\psi(x)$ for the definite integral (1), note first that the series (3) converges uniformly in any halfplane $\operatorname{Re} s \geq K > 1$ [by comparison with the convergent series $\sum (\log n)n^{-K}$]. This proves both that the termwise differentiation of $-\log \zeta(s) = \sum \log (1 - p^{-s})$ is valid (a series can be differentiated termwise if the result is uniformly convergent) and that the integral over any *finite* path can be computed termwise:

$$\text{(4)} \qquad \frac{1}{2\pi i}\int_{a-ih}^{a+ih}\left[-\frac{\zeta'(s)}{\zeta(s)}\right]x^s\frac{ds}{s} = \sum_{n=1}^{\infty} \Lambda(n)\cdot\frac{1}{2\pi i}\int_{a-ih}^{a+ih}\left(\frac{x}{n}\right)^s\frac{ds}{s}$$

when $a > 1$. Now *if* it can be shown that *the limit as $h \to \infty$ of the sum on the right is equal to the sum of the limits*, then the fact that the integral (1) has the value $\psi(x)$ will follow immediately from the formula

$$\text{(5)} \qquad \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} y^s\frac{ds}{s} = \begin{cases} 0 & \text{if } 0 < y < 1 \\ \frac{1}{2} & \text{if } y = 1 \\ 1 & \text{if } y > 1 \end{cases} \qquad (a > 0).$$

This formula, which was deduced from Fourier's theorem in Section 1.14, will be proved directly in the next section. To summarize, then, the proof that the definite integral (1) is equal to $\psi(x)$ depends on the proof that the limit as $h \to \infty$ of the sum in (4) is the sum of the limits, and on the proof of the integral formula (5).

The second method of evaluating the integral (1) is as follows. Differentiate logarithmically the formula

$$\Pi\left(\frac{s}{2}\right)\pi^{-s/2}(s-1)\zeta(s) = \xi(0)\prod_{\rho}\left(1-\frac{s}{\rho}\right)$$

to find

$$\frac{d}{ds}\log\Pi\left(\frac{s}{2}\right) - \frac{1}{2}\log\pi + \frac{1}{s-1} + \frac{\zeta'(s)}{\zeta(s)} = \sum_{\rho}\frac{1}{1-(s/\rho)}\cdot\left(-\frac{1}{\rho}\right).$$

Using the expression of $\Pi(x)$ as an infinite product [(4) of Section 1.3], and differentiating termwise then gives

$$(6)\qquad -\frac{\zeta'(s)}{\zeta(s)} = \frac{1}{s-1} - \sum_\rho \frac{1}{s-\rho} + \sum_{n=1}^{\infty}\left[-\frac{1}{s+2n} + \frac{1}{2}\log\left(1+\frac{1}{n}\right)\right] - \frac{1}{2}\log\pi.$$

With $s = 0$ this gives

$$-\frac{\zeta'(0)}{\zeta(0)} = -1 - \sum_\rho\left(-\frac{1}{\rho}\right) + \sum_{n=1}^{\infty}\left[-\frac{1}{2n} + \frac{1}{2}\log\left(1+\frac{1}{n}\right)\right] - \frac{1}{2}\log\pi,$$

so subtraction gives

$$-\frac{\zeta'(s)}{\zeta(s)} + \frac{\zeta'(0)}{\zeta(0)} = \left[\frac{1}{s-1} + 1\right] - \sum_\rho\left[\frac{1}{s-\rho} + \frac{1}{\rho}\right] - \sum_{n=1}^{\infty}\left[\frac{1}{s+2n} - \frac{1}{2n}\right]$$

and finally

$$(7)\qquad -\frac{\zeta'(s)}{\zeta(s)} = \frac{s}{s-1} - \sum_\rho \frac{s}{\rho(s-\rho)} + \sum_{n=1}^{\infty}\frac{s}{2n(s+2n)} - \frac{\zeta'(0)}{\zeta(0)}.$$

Substitute this formula in the definite integral (1) and assume that termwise integration is valid. This gives the value

$$\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} x^s\frac{ds}{s-1} - \sum_\rho \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} x^s\frac{ds}{\rho(s-\rho)} + \sum_n \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} x^s\frac{ds}{2n(s+2n)} + \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} x^s\left[-\frac{\zeta'(0)}{\zeta(0)}\right]\frac{ds}{s}$$

for (1). Now the change of variable $t = s - \beta$ in the previous integral formula (5) gives

$$(8)\qquad \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} x^s\frac{ds}{s-\beta} = \frac{1}{2\pi i}\int_{a-\beta-i\infty}^{a-\beta+i\infty} x^\beta x^t\frac{dt}{t} = x^\beta\frac{1}{2\pi i}\int_{\mathrm{Re}(a-\beta)-i\infty}^{\mathrm{Re}(a-\beta)+i\infty} x^t\frac{dt}{t} = x^\beta$$

provided $x > 1$ and $a > \mathrm{Re}\,\beta$. (The middle equation here, in which the limits of integration are switched from $a - \beta \pm i\infty$ to $\mathrm{Re}(a-\beta) \pm i\infty$, is not trivial because these two integrals when written as limits as $h \to \infty$ are different. However, the difference between them is two integrals of $x^t\,dt/t$ over intervals of the form $[a \pm ib, a \pm i(b+c)]$, where c is fixed and $b \to \infty$; thus the difference is less than a constant times $c/b \to 0$.) Thus the value of

(1) reduces, when $x > 1$, to

$$x - \sum_\rho \frac{x^\rho}{\rho} + \sum_n \frac{x^{-2n}}{2n} - \frac{\zeta'(0)}{\zeta(0)} \qquad (x > 1)$$

as desired.

To justify this sequence of steps leading to the value $x - \Sigma(x^\rho/\rho) + \Sigma(x^{-2n}/2n) - \zeta'(0)/\zeta(0)$ of the definite integral (1), note first that both of the infinite series in (6) converge uniformly in any disk $|s| \le K$. (The series in n converges uniformly because

$$\begin{aligned}
&|(s+2n)^{-1} - \tfrac{1}{2}\log(1+n^{-1})| \\
&\quad = |(s+2n)^{-1} - (2n)^{-1} + (2n)^{-1} \\
&\qquad - \tfrac{1}{2}(n^{-1} - \tfrac{1}{2}n^{-2} + \tfrac{1}{3}n^{-3} - \cdots)| \\
&\quad \le |s(s+2n)^{-1}(2n)^{-1}| + |\tfrac{1}{4}n^{-2} - \tfrac{1}{6}n^{-3} + \cdots| \\
&\quad \le K(2n)^{-2} + n^{-2} \le \text{const}/n^2
\end{aligned}$$

for all sufficiently large n, and the series in ρ converges uniformly because when the terms ρ and $1-\rho$ are paired

$$\begin{aligned}
&|(s-\rho)^{-1} + [s-(1-\rho)]^{-1}| \\
&\quad = \left|\left[\left(s-\frac{1}{2}\right) - \left(\rho - \frac{1}{2}\right)\right]^{-1} + \left[\left(s-\frac{1}{2}\right) + \left(\rho-\frac{1}{2}\right)\right]^{-1}\right| \\
&\quad = \left|\frac{2(s-\frac{1}{2})}{(s-\frac{1}{2})^2 - (\rho-\frac{1}{2})^2}\right| \le \text{const}\left|\rho - \frac{1}{2}\right|^{-2}
\end{aligned}$$

for all sufficiently large ρ once K is fixed and because $\sum|\rho - \frac{1}{2}|^{-2}$ converges by the theorem of Section 2.5.) This proves that the termwise differentiation by which (6) was obtained is valid. Then it follows by an elementary theorem $[\sum (a_n + b_n) = \sum a_n + \sum b_n$ when $\sum a_n, \sum b_n$ both converge$]$ that (7) is valid—except at the zeros and poles $1, \rho, -2n$ of ζ—and that the series it contains both converge uniformly in $|s| \le K$. Thus it can be integrated termwise over *finite* intervals and the integral (1) is therefore equal to

$$\begin{aligned}
(9) \quad &\lim_{h\to\infty} \frac{1}{2\pi i}\int_{a-ih}^{a+ih}\left[\frac{s}{s-1} - \sum_\rho \frac{s}{\rho(s-\rho)} + \sum_n \frac{s}{2n(s+2n)} - \frac{\zeta'(0)}{\zeta(0)}\right]x^s\frac{ds}{s} \\
&= x - \lim_{h\to\infty}\sum_\rho \frac{1}{2\pi i}\int_{a-ih}^{a+ih}\frac{x^s\,ds}{\rho(s-\rho)} \\
&\quad + \lim_{h\to\infty}\sum_n \frac{1}{2\pi i}\int_{a-ih}^{a+ih}\frac{x^s\,ds}{2n(s+2n)} - \frac{\zeta'(0)}{\zeta(0)}
\end{aligned}$$

[where use is made of the fact that the limit of a finite sum is the sum of the limits and the first and last terms are evaluated using (8) and (5)]. Now *if* the limits of these two sums are equal to the sums of the limits, then the rest of the argument is elementary and the value $x - \sum(x^\rho/\rho) + \sum(x^{-2n}/2n) - \zeta'(0)/\zeta(0)$ for (1) will be proved. Thus, in addition to the basic formula (5), the rigorous proof that the definite integral (1) has this value depends on the validity of the termwise evaluation of the two limits in (9).

Thus von Mangoldt's formula

$$\psi(x) = x - \sum_\rho \frac{x^\rho}{\rho} + \sum_n \frac{x^{-2n}}{2n} - \frac{\zeta'(0)}{\zeta(0)} \qquad (x > 1)$$

depends on the validity of the integral formula (5) and of the three interchanges of $\lim_{h\to\infty}$ with infinite sums in (4) and (9). The following three sections are devoted to proving that all are indeed valid. The numerical value of the constant $\zeta'(0)/\zeta(0) = \log 2\pi$ is found in Section 3.8.

3.3 THE BASIC INTEGRAL FORMULA

This section is devoted to the evaluation of

$$\text{(1)} \qquad \lim_{h\to\infty} \frac{1}{2\pi i} \int_{a-ih}^{a+ih} \frac{x^s\,ds}{s} \qquad (x > 0, \quad a > 0).$$

Since the arguments of Section 3.2 deal with infinite sums of such limits, it will be necessary to find, in addition to the limit (1), the *rate* at which this limit is approached. For the case $0 < x < 1$ this is accomplished by the estimate

$$\text{(2)} \qquad \left| \frac{1}{2\pi i} \int_{a-ih}^{a+ih} \frac{x^s\,ds}{s} \right| \le \frac{x^a}{\pi h\,|\log x|} \qquad (a > 0, \quad 0 < x < 1),$$

which can be proved as follows. Because $a > 0$, the function x^s/s has no singularity in the rectangle $\{a \le \operatorname{Re} s \le K, -h \le \operatorname{Im} s \le h\}$, where K is a large constant. Hence by Cauchy's theorem the integral of $x^s\,ds/s$ around the boundary of this rectangle is zero, which gives

$$\frac{1}{2\pi i} \int_{a-ih}^{a+ih} \frac{x^s\,ds}{s} = -\frac{1}{2\pi i} \int_{a+ih}^{K+ih} \frac{x^s\,ds}{s} + \frac{1}{2\pi i} \int_{a-ih}^{K-ih} \frac{x^s\,ds}{s} + \frac{1}{2\pi i} \int_{K-ih}^{K+ih} \frac{x^s\,ds}{s}.$$

The last integral has modulus at most $(2\pi)^{-1}(x^K/K)(2h)$. Each of the other two integrals on the right has modulus at most

$$\frac{1}{2\pi} \int_a^K \frac{x^\sigma\,d\sigma}{h} = \frac{1}{2\pi} \left[\frac{x^\sigma}{h \log x} \right]_{\sigma=a}^{\sigma=K}$$

which then gives

$$\left| \frac{1}{2\pi i} \int_{a-ih}^{a+ih} \frac{x^s\,ds}{s} \right| \le \frac{1}{\pi} \frac{|x^K - x^a|}{h\,|\log x|} + \frac{1}{2\pi} \frac{x^K}{K} 2h.$$

When $0 < x < 1$, the limit of x^K as $K \to \infty$ is zero and (2) follows. Thus, in particular, the limit (1) is zero when $0 < x < 1$.

In the case $x = 1$ no estimate of the rate of approach to the limit will be needed and it suffices to note that (1) is

$$\lim_{h\to\infty} \frac{1}{2\pi} \int_{-h}^{h} \frac{dt}{a+it} = \lim_{h\to\infty} \left(\frac{1}{2\pi} \int_{-h}^{h} \frac{a\,dt}{a^2+t^2} - i\frac{1}{2\pi} \int_{-h}^{h} \frac{t\,dt}{a^2+t^2}\right)$$
$$= \lim_{h\to\infty} \frac{1}{2\pi} \int_{-h/a}^{h/a} \frac{du}{1+u^2} = \frac{1}{2}$$

using the well-known† formula $\int_{-\infty}^{\infty} (1+u^2)^{-1}\,du = \pi$. Thus for $x = 1$ the limit (1) is $\frac{1}{2}$.

For $x > 1$ the estimate analogous to (2) is obtained by considering the integral of $x^s\,ds/(2\pi is)$ around the boundary of a rectangle of the form $\{-K \le \operatorname{Re} s \le a,\ -h \le \operatorname{Im} s \le h\}$. Since x^s is analytic in this rectangle and $s = 0$ lies inside the rectangle, the Cauchy integral formula states that this integral is $x^0 = 1$, hence

$$\frac{1}{2\pi i} \int_{a-ih}^{a+ih} \frac{x^s\,ds}{s} + \frac{1}{2\pi i} \int_{a+ih}^{-K+ih} \frac{x^s\,ds}{s}$$
$$+ \frac{1}{2\pi i} \int_{-K+ih}^{-K-ih} \frac{x^s\,ds}{s} + \frac{1}{2\pi i} \int_{-K-ih}^{a-ih} \frac{x^s\,ds}{s} = 1,$$

$$\left|\frac{1}{2\pi i} \int_{a-ih}^{a+ih} \frac{x^s\,ds}{s} - 1\right| \le \frac{1}{2\pi} \int_{-K}^{a} \frac{x^\sigma\,d\sigma}{h} + \frac{1}{2\pi} \frac{x^{-K}}{K} \cdot 2h + \frac{1}{2\pi} \int_{-K}^{a} \frac{x^\sigma\,d\sigma}{h}$$
$$= \frac{1}{\pi} \frac{x^a - x^{-K}}{h \log x} + \frac{1}{\pi} \frac{x^{-K}h}{K}.$$

Letting $K \to \infty$ then gives

$$\left|\frac{1}{2\pi i} \int_{a-ih}^{a+ih} \frac{x^s\,ds}{s} - 1\right| \le \frac{x^a}{\pi h \log x} \qquad (a > 0, \quad x > 1) \tag{3}$$

which is the desired estimate. In particular, the limit (1) is 1 when $x > 1$.

One other estimate of integrals is used in the proof of von Mangoldt's formula, namely, von Mangoldt's estimate

$$\left|\frac{1}{2\pi i} \int_{a+ic}^{a+id} \frac{x^s\,ds}{s}\right| \le K \frac{x^a}{(a+c)\log x} \qquad (x > 1, \quad a > 0, \quad d > c \ge 0), \tag{4}$$

where K is a constant which may be taken to be $(4+\pi)/(2\pi\sqrt{2})$. To prove this formula, note that integration by parts

$$\int \frac{x^s\,ds}{s} = \frac{x^s}{s \log x} + \int \frac{x^s\,ds}{s^2 \log x}$$

gives

$$\left|\int_{a+ic}^{a+id} \frac{x^s\,ds}{s}\right| \le \left|\frac{x^{a+id}}{(a+id)\log x}\right| + \left|\frac{x^{a+ic}}{(a+ic)\log x}\right|$$
$$+ \frac{x^a}{\log x} \left|\int_c^d \frac{x^{it}\,dt}{(a+it)^2}\right|.$$

†See Edwards [E1, p. 65].