Prime Number Theory and the Riemann Zeta-Function

D.R. Heath-Brown

1 Primes

An integer $p \in \mathbb{N}$ is said to be "prime" if $p \neq 1$ and there is no integer n dividing p with $1 < n < p$. (This is not the algebraist's definition, but in our situation the two definitions are equivalent.)

The primes are multiplicative building blocks for $\mathbb{N}$, as the following crucial result describes.

Theorem 1 (The Fundamental Theorem of Arithmetic.) *Every $n \in \mathbb{N}$ can be written in exactly one way in the form*

$$n = p_1^{e_1} p_2^{e_2} \dots p_k^{e_k},$$

with $k \geq 0$, $e_1, \dots, e_k \geq 1$ and primes $p_1 < p_2 < \dots < p_k$.

For a proof, see Hardy and Wright [5, Theorem 2], for example. The situation for $\mathbb{N}$ contrasts with that for arithmetic in the set

$$\{m + n\sqrt{-5} : m, n \in \mathbb{Z}\},$$

where one has, for example,

$$6 = 2 \times 3 = (1 + \sqrt{-5}) \times (1 - \sqrt{-5}),$$

with $2, 3, 1 + \sqrt{-5}$ and $1 - \sqrt{-5}$ all being "primes".

A second fundamental result appears in the works of Euclid.

Theorem 2 *There are infinitely many primes.*

This is proved by contradiction. Assume there are only finitely many primes, $p_1, p_2, \dots, p_n$, say. Consider the integer $N = 1 + p_1 p_2 \dots p_n$. Then $N \geq 2$, so that N must have at least one prime factor p, say. But our list of primes was supposedly complete, so that p must be one of the primes p_i,

say. Then p_i divides $N-1$, by construction, while $p = p_i$ divides N by assumption. It follows that p divides $N-(N-1)=1$, which is impossible. This contradiction shows that there can be no finite list containing all the primes.

There have been many tables of primes produced over the years. They show that the detailed distribution is quite erratic, but if we define

$$\pi(x) = \#\{p \le x : p \text{ prime}\},$$

then we find that $\pi(x)$ grows fairly steadily. Gauss conjectured that

$$\pi(x) \sim \mathrm{Li}(x),$$

where

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t},$$

that is to say that

$$\lim_{x\to\infty} \frac{\pi(x)}{\mathrm{Li}(x)} = 1.$$

The following figures bear this out.

$$\begin{array}{rcrrcl} \pi(10^8) & = & 5{,}776{,}455 & \frac{\pi(x)}{\mathrm{Li}(x)} & = & 0.999869147\ldots, \\ \pi(10^{12}) & = & 37{,}607{,}912{,}018 & \frac{\pi(x)}{\mathrm{Li}(x)} & = & 0.999989825\ldots, \\ \pi(10^{16}) & = & 279{,}238{,}341{,}033{,}925 & \frac{\pi(x)}{\mathrm{Li}(x)} & = & 0.999999989\ldots. \end{array}$$

It is not hard to show that in fact

$$\mathrm{Li}(x) \sim \frac{x}{\log x},$$

but it turns out that $\mathrm{Li}(x)$ gives a better approximation to $\pi(x)$ than $x/\log x$ does. Gauss' conjecture was finally proved in 1896, by Hadamard and de la Vallée Poussin, working independently.

Theorem 3 (The Prime Number Theorem.) *We have*

$$\pi(x) \sim \frac{x}{\log x}$$

as $x \to \infty$.

One interesting interpretation of the Prime Number Theorem is that for a number n in the vicinity of x the "probability" that n is prime is asymptotically $1/\log x$, or equivalently, that the "probability" that n is prime is asymptotically $1/\log n$. Of course the event "n is prime" is deterministic — that is to say, the probability is 1 if n is prime, and 0 otherwise. None the

less the probabilistic interpretation leads to a number of plausible heuristic arguments. As an example of this, consider, for a given large integer n, the probability that $n+1, n+2, \ldots, n+k$ are all composite. If k is at most n, say, then the probability that any one of these is composite is about $1 - 1/\log n$. Thus if the events were all independent, which they are not, the overall probability would be about

$$\left(1 - \frac{1}{\log n}\right)^k.$$

Taking $k = \mu(\log n)^2$ and approximating

$$\left(1 - \frac{1}{\log n}\right)^{\log n}$$

by e^{-1}, we would have that the probability that $n+1, n+2, \ldots, n+k$ are all composite, is around $n^{-\mu}$.

If E_n is the event that $n+1, n+2, \ldots, n+k$ are all composite, then the events E_n and E_{n+1} are clearly not independent. However we may hope that E_n and E_{n+k} are independent. If the events E_n were genuinely independent for different values of n then an application of the Borel-Cantelli lemma would tell us that E_n should happen infinitely often when $\mu < 1$, and finitely often for $\mu \geq 1$. With more care one can make this plausible even though E_n and $E_{n'}$ are correlated for nearby values n and n'. We are thus led to the following conjecture.

Conjecture 1 *If p' denotes the next prime after p then*

$$\limsup_{p \to \infty} \frac{p' - p}{(\log p)^2} = 1.$$

Numerical evidence for this is hard to produce, but what there is seems to be consistent with the conjecture.

In the reverse direction, our simple probabilistic interpretation of the Prime Number Theorem might suggest that the probability of having both n and $n+1$ prime should be around $(\log n)^{-2}$. This is clearly wrong, since one of n and $n+1$ is always even. However, a due allowance for such arithmetic effects leads one to the following.

Conjecture 2 *If*

$$c = 2 \prod_{p > 2} \left(1 - \frac{1}{(p-1)^2}\right) = 1.3202\ldots,$$

the product being over primes, then

$$\#\{n \leq x : n, n+2 \text{ both prime}\} \doteqdot c \int_2^x \frac{dt}{(\log t)^2}. \tag{1}$$

The numerical evidence for this is extremely convincing.

Thus the straightforward probabilistic interpretation of the Prime Number Theorem leads to a number of conjectures, which fit very well with the available numerical evidence. This probabilistic model is known as "Cramér's Model" and has been widely used for predicting the behaviour of primes.

One further example of this line of reasoning shows us however that the primes are more subtle than one might think. Consider the size of

$$\pi(N+H) - \pi(N) = \#\{p : N < p \leq N+H\},$$

when H is small compared with N. The Prime Number Theorem leads one to expect that

$$\pi(N+H) - \pi(N) \doteqdot \int_N^{N+H} \frac{dt}{\log t} \sim \frac{H}{\log N}.$$

However the Prime Number Theorem only says that

$$\pi(x) = \int_2^x \frac{dt}{\log t} + o(\frac{x}{\log x}),$$

or equivalently that

$$\pi(x) = \int_2^x \frac{dt}{\log t} + f(x),$$

where

$$\frac{f(x)}{x/\log x} \to 0$$

as $x \to \infty$. Hence

$$\pi(N+H) - \pi(N) = \int_N^{N+H} \frac{dt}{\log t} + f(N+H) - f(N).$$

In order to assert that

$$\frac{f(N+H) - f(N)}{H/\log N} \to 0$$

as $N \to \infty$ we need $cN \leq H \leq N$ for some constant $c > 0$. None the less, considerably more subtle arguments show that

$$\pi(N+H) - \pi(N) \sim \frac{H}{\log N}$$

even when H is distinctly smaller than N.

A careful application of the Cramér Model suggests the following conjecture. "

Conjecture 3 *" Let $\kappa > 2$ be any constant. Then if $H = (\log N)^\kappa$ we should have*

$$\pi(N+H) - \pi(N) \sim \frac{H}{\log N}$$

as $N \to \infty$.

This is supported by the following result due to Selberg in 1943 [15].

Theorem 4 *Let $f(N)$ be any increasing function for which $f(N) \to \infty$ as $N \to \infty$. Assume the Riemann Hypothesis. Then there is a subset $\mathcal{E}$ of the integers $\mathbb{N}$, with*

$$\#\{n \in \mathcal{E} : n \leq N\} = o(N)$$

as $N \to \infty$, such that

$$\pi(n + f(n)\log^2 n) - \pi(n) \sim f(n)\log n$$

for all $n \notin \mathcal{E}$.

Conjecture 3 would say that one can take $\mathcal{E} = \emptyset$ if $f(N)$ is a positive power of $\log N$.

Since Cramér's Model leads inexorably to Conjecture 3, it came as quite a shock to prime number theorists when the conjecture was disproved by Maier [9] in 1985. Maier established the following result.

Theorem 5 *For any $\kappa > 1$ there is a constant $\delta_\kappa > 0$ such that*

$$\limsup_{N\to\infty} \frac{\pi(N + (\log N)^\kappa) - \pi(N)}{(\log N)^{\kappa-1}} \geq 1 + \delta_\kappa$$

and

$$\liminf_{N\to\infty} \frac{\pi(N + (\log N)^\kappa) - \pi(N)}{(\log N)^{\kappa-1}} \leq 1 - \delta_\kappa.$$

The values of N produced by Maier, where $\pi(N + (\log N)^\kappa) - \pi(N)$ is abnormally large, (or abnormally small), are very rare. None the less their existence shows that the Cramér Model breaks down. Broadly speaking one could summarize the reason for this failure by saying that arithmetic effects play a bigger rôle than previously supposed. As yet we have no good alternative to the Cramér model.

2 Open Questions About Primes, and Important Results

Here are a few of the well-known unsolved problems about the primes.

(1) Are there infinitely many "prime twins" $n, n+2$ both of which are prime? (Conjecture 2 gives a prediction for the rate at which the number of such pairs grows.)

(2) Is every even integer $n \geq 4$ the sum of two primes? (Goldbach's Conjecture.)

(3) Are there infinitely many primes of the form $p = n^2 + 1$?

(4) Are there infinitely many "Mersenne primes" of the form $p = 2^n - 1$?

(5) Are there arbitrarily long arithmetic progressions, all of whose terms are prime?

(6) Is there always a prime between any two successive squares?

However there have been some significant results proved too. Here are a selection.

(1) There are infinitely many primes of the form $a^2 + b^4$. (Friedlander and Iwaniec [4], 1998.)

(2) There are infinitely many primes p for which $p + 2$ is either prime or a product of two primes. (Chen [2], 1966.)

(3) There is a number n_0 such that any even number $n \geq n_0$ can be written as $n = p + p'$ with p prime and p' either prime or a product of two primes. (Chen [2], 1966.)

(4) There are infinitely many integers n such that $n^2 + 1$ is either prime or a product of two primes. (Iwaniec [8], 1978.)

(5) For any constant $c < \frac{243}{205} = 1.185\ldots$, there are infinitely many integers n such that $[n^c]$ is prime. Here $[x]$ denotes the integral part of x, that is to say the largest integer N satisfying $N \leq x$. (Rivat and Wu [14], 2001, after Piatetski-Shapiro, [11], 1953.)

(6) Apart from a finite number of exceptions, there is always a prime between any two consecutive cubes. (Ingham [6], 1937.)

(7) There is a number n_0 such that for every $n \geq n_0$ there is at least one prime in the interval $[n\,,\, n + n^{0.525}]$. (Baker, Harman and Pintz, [1], 2001.)

(8) There are infinitely many consecutive primes $p,'p$ such that $p' - p \leq (\log p)/4$. (Maier [10], 1988.)

(9) There is a positive constant c such that there are infinitely many consecutive primes $p,'p$ such that

$$p' - p \geq c \log p \frac{(\log\log p)(\log\log\log\log p)}{(\log\log\log p)^2}.$$

(Rankin [13], 1938.)

(10) For any positive integer q and any integer a in the range $0 \leq a < q$, which is coprime to q, there are arbitrarily long strings of consecutive primes, all of which leave remainder a on division by q. (Shiu [16], 2000.)

By way of explanation we should say the following. The result (1) demonstrates that even though we cannot yet handle primes of the form $n^2 + 1$, we can say something about the relatively sparse polynomial sequence $a^2 + b^4$. The result in (5) can be viewed in the same context. One can think of $[n^c]$ as being a "polynomial of degree c" with $c > 1$. Numbers (2), (3) and (4) are approximations to, respectively, the prime twins problem, Goldbach's problem, and the problem of primes of the shape $n^2 + 1$. The theorems in (6) and (7) are approximations to the conjecture that there should be a prime between consecutive squares. Of these (7) is stronger, if less elegant. Maier's result (8) shows that the difference between consecutive primes is sometimes smaller than average by a factor $1/4$, the average spacing being $\log p$ by the Prime Number Theorem. (Of course the twin prime conjecture would be a much stronger result, with differences between consecutive primes sometimes being as small as 2.) Similarly, Rankin's result (9) demonstrates that the gaps between consecutive primes can sometimes be larger than average, by a factor which is almost $\log\log p$. Again this is some way from what we expect, since Conjecture 1 predict gaps as large as $(\log p)^2$. Finally, Shiu's result (10) is best understood by taking $q = 10^7$ and $a = 7,777,777$, say. Thus a prime leaves remainder a when divided by q, precisely when its decimal expansion ends in 7 consecutive 7's. Then (10) tells us that a table of primes will somewhere contain a million consecutive entries, each of which ends in the digits 7,777,777.

3 The Riemann Zeta-Function

In the theory of the zeta-function it is customary to use the variable $s = \sigma + it \in \mathbb{C}$. One then defines the complex exponential

$$n^{-s} := \exp(-s \log n), \quad \text{with} \quad \log n \in \mathbb{R}.$$

The Riemann Zeta-function is then

$$\zeta(s) := \sum_{n=1}^{\infty} n^{-s}, \quad \sigma > 1. \tag{2}$$

The sum is absolutely convergent for $\sigma > 1$, and for fixed $\delta > 0$ it is uniformly convergent for $\sigma \geq 1 + \delta$. It follows that $\zeta(s)$ is holomorphic for $\sigma > 1$. The function is connected to the primes as follows.

Theorem 6 (The Euler Product.) *If $\sigma > 1$ then we have*

$$\zeta(s) = \prod_{p} (1 - p^{-s})^{-1},$$

where p runs over all primes, and the product is absolutely convergent.

This result is, philosophically, at the heart of the theory. It relates a sum over all positive integers to a product over primes. Thus it relates the additive structure, in which successive positive integers are generated by adding 1, to the multiplicative structure. Moreover we shall see in the proof that the fact that the sum and the product are equal is exactly an expression of the Fundamental Theorem of Arithmetic.

To prove the result consider the finite product

$$\prod_{p \leq X} (1 - p^{-s})^{-1}.$$

Since $\sigma > 1$ we have $|p^{-s}| < p^{-1} < 1$, whence we can expand $(1 - p^{-s})^{-1}$ as an absolutely convergent series $1 + p^{-s} + p^{-2s} + p^{-3s} + \ldots$. We may multiply together a finite number of such series, and rearrange them, since we have absolute convergence. This yields

$$\prod_{p \leq X} (1 - p^{-s})^{-1} = \sum_{n=1}^{\infty} \frac{a_X(n)}{n^s},$$

where the coefficient $a_X(n)$ is the number of ways of writing n in the form

$$n = p_1^{e_1} p_2^{e_2} \ldots p_r^{e_r} \quad \text{with} \quad p_1 < p_2 < \ldots < p_r \leq X.$$

By the Fundamental Theorem of Arithmetic we have $a_X(n) = 0$ or 1, and if $n \leq X$ we will have $a_X(n) = 1$. It follows that

$$|\sum_{n=1}^{\infty} n^{-s} - \sum_{n=1}^{\infty} \frac{a_X(n)}{n^s}| \leq \sum_{n>X} |\frac{1}{n^s}| = \sum_{n>X} \frac{1}{n^\sigma}.$$

As $X \to \infty$ this final sum must tend to zero, since the infinite sum $\sum_{n=1}^{\infty} n^{-\sigma}$ converges. We therefore deduce that if $\sigma > 1$, then

$$\lim_{X\to\infty} \prod_{p\leq X} (1 - p^{-s})^{-1} = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

as required. Of course the product is absolutely convergent, as one may see by taking $s = \sigma$.

One important deduction from the Euler product identity comes from taking logarithms and differentiating termwise. This can be justified by the local uniform convergence of the resulting series.

Corollary 1 *We have*

$$-\frac{\zeta'}{\zeta}(s) = \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n^s}, \quad (\sigma > 1), \tag{3}$$

where

$$\Lambda(n) = \begin{cases} \log p, & n = p^e, \\ 0, & \textit{otherwise.} \end{cases}$$

The function $\Lambda(n)$ is known as the von Mangoldt function.

4 The Analytic Continuation and Functional Equation of $\zeta(s)$

Our definition only gives a meaning to $\zeta(s)$ when $\sigma > 1$. We now seek to extend the definition to all $s \in \mathbb{C}$. The key tool is the Poisson Summation Formula.

Theorem 7 (The Poisson Summation Formula.) *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is twice differentiable and that f, f' and f'' are all integrable over $\mathbb{R}$. Define the Fourier transform by*

$$\hat{f}(t) := \int_{-\infty}^{\infty} f(x)e^{-2\pi itx}dx.$$

Then

$$\sum_{-\infty}^{\infty} f(n) = \sum_{-\infty}^{\infty} \hat{f}(n),$$

both sides converging absolutely.

There are weaker conditions under which this holds, but the above more than suffices for our application. The reader should note that there are a number of conventions in use for defining the Fourier transform, but the one used here is the most appropriate for number theoretic purposes.

The proof (see Rademacher [12, page 71], for example) uses harmonic analysis on $\mathbb{R}^+$. Thus it depends only on the additive structure and not on the multiplicative structure.

If we apply the theorem to $f(x) = \exp\{-x^2\pi v\}$, which certainly fulfils the conditions, we have

$$\begin{aligned}
\hat{f}(n) &= \int_{-\infty}^{\infty} e^{-x^2\pi v} e^{-2\pi i n x} dx \\
&= \int_{-\infty}^{\infty} e^{-\pi v(x+in/v)^2} e^{-\pi n^2/v} dx \\
&= e^{-\pi n^2/v} \int_{-\infty}^{\infty} e^{-\pi v y^2} dy \\
&= \frac{1}{\sqrt{v}}\, e^{-\pi n^2/v},
\end{aligned}$$

providing that v is real and positive. Thus if we define

$$\theta(v) := \sum_{-\infty}^{\infty} \exp(-\pi n^2 v),$$

then the Poisson Summation Formula leads to the transformation formula

$$\theta(v) = \frac{1}{\sqrt{v}}\, \theta(1/v).$$

The function $\theta(v)$ is a *theta-function*, and is an example of a *modular form*. It is the fact that $\theta(v)$ not only satisfies the above transformation formula when v goes to $1/v$ but is also periodic, that makes $\theta(v)$ a modular form.

The "Langlands Philosophy" says that all reasonable generalizations of the Riemann Zeta-function are related to modular forms, in a suitably generalized sense.

We are now ready to consider $\zeta(s)$, but first we introduce the function

$$\psi(v) = \sum_{n=1}^{\infty} e^{-n^2\pi v}, \tag{4}$$

so that $\psi(v) = (\theta(v) - 1)/2$ and

$$2\psi(v) + 1 = \frac{1}{\sqrt{v}}\{2\psi(\frac{1}{v}) + 1\}. \tag{5}$$

We proceed to compute that, if $\sigma > 1$, then

$$\begin{aligned}
\int_0^{\infty} x^{s/2-1}\psi(x)dx &= \sum_{n=1}^{\infty}\int_0^{\infty} x^{s/2-1}e^{-n^2\pi x}dx \\
&= \sum_{n=1}^{\infty}\frac{1}{(n^2\pi)^{s/2}}\int_0^{\infty} y^{s/2-1}e^{-y}dy \\
&= \sum_{n=1}^{\infty}\frac{1}{(n^2\pi)^{s/2}}\Gamma(\frac{s}{2}) \\
&= \zeta(s)\pi^{-s/2}\Gamma(\frac{s}{2}),
\end{aligned}$$

on substituting $y = n^2\pi x$. The interchange of summation and integration is justified by the absolute convergence of the resulting sum.

We now split the range of integration in the original integral, and apply the transformation formula (5). For $\sigma > 1$ we obtain the expression

$$\begin{aligned}
\zeta(s)\pi^{-s/2}\Gamma(\frac{s}{2}) &= \int_1^{\infty} x^{s/2-1}\psi(x)dx + \int_0^1 x^{s/2-1}\psi(x)dx \\
&= \int_1^{\infty} x^{s/2-1}\psi(x)dx + \int_0^1 x^{s/2-1}\{\frac{1}{\sqrt{x}}\psi(\frac{1}{x}) + \frac{1}{2\sqrt{x}} - \frac{1}{2}\}dx \\
&= \int_1^{\infty} x^{s/2-1}\psi(x)dx + \int_0^1 x^{s/2-3/2}\psi(\frac{1}{x})dx + \frac{1}{s-1} - \frac{1}{s} \\
&= \int_1^{\infty} x^{s/2-1}\psi(x)dx + \int_1^{\infty} y^{(1-s)/2-1}\psi(y)dy - \frac{1}{s(1-s)},
\end{aligned}$$

where we have substituted y for $1/x$ in the final integral.

We therefore conclude that

$$\zeta(s)\pi^{-s/2}\Gamma(\frac{s}{2}) = \int_1^{\infty}\{x^{s/2-1} + x^{(1-s)/2-1}\}\psi(x)dx - \frac{1}{s(1-s)}, \tag{6}$$

whenever $\sigma > 1$. However the right-hand side is meaningful for all values $s \in \mathbb{C} - \{0, 1\}$, since the integral converges by virtue of the exponential decay of $\psi(x)$. We may therefore use the above expression to define $\zeta(s)$ for all $s \in \mathbb{C} - \{0, 1\}$, on noting that the factor $\pi^{-s/2}\Gamma(s/2)$ never vanishes. Indeed, since $\Gamma(s/2)^{-1}$ has a zero at $s = 0$ we see that the resulting expression for $\zeta(s)$ is regular at $s = 0$. Finally we observe that the right-hand side of (6) is invariant on substituting s for $1 - s$. We are therefore led to the the following conclusion.

Theorem 8 (Analytic Continuation and Functional Equation.) *The function $\zeta(s)$ has an analytic continuation to $\mathbb{C}$, and is regular apart from a simple pole at $s = 1$, with residue 1. Moreover*

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

Furthermore, if $a \leq \sigma \leq b$ and $|t| \geq 1$, then $\pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s)$ is bounded in terms of a and b.

To prove the last statement in the theorem we merely observe that

$$|\pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s)| \leq 1 + \int_1^\infty (x^{b/2-1} + x^{(1-a)/2-1})\psi(x)dx.$$

5 Zeros of $\zeta(s)$

It is convenient to define

$$\xi(s) = \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s) = (s-1)\pi^{-s/2}\Gamma(1+\frac{s}{2})\zeta(s), \qquad (7)$$

so that $\xi(s)$ is entire. The functional equation then takes the form $\xi(s) = \xi(1-s)$. It is clear from (3) that $\zeta(s)$ can have no zeros for $\sigma > 1$, since the series converges. Since $1/\Gamma(z)$ is entire, the function $\Gamma(s/2)$ is non-vanishing, so that $\xi(s)$ also has no zeros in $\sigma > 1$. Thus, by the functional equation, the zeros of $\xi(s)$ are confined to the "critical strip" $0 \leq \sigma \leq 1$. Moreover any zero of $\zeta(s)$ must either be a zero of $\xi(s)$, or a pole of $\Gamma(s/2)$. We then see that the zeros of $\zeta(s)$ lie in the critical strip, with the exception of the "trivial zeros" at $s = -2, -4, -6, \ldots$ corresponding to poles of $\Gamma(s/2)$.

We may also observe that if ρ is a zero of $\xi(s)$ then, by the functional equation, so is $1-\rho$. Moreover, since $\overline{\xi(s)} = \xi(\overline{s})$, we deduce that $\overline{\rho}$ and $1-\overline{\rho}$ are also zeros. Thus the zeros are symmetrically arranged about the real axis, and also about the "critical line" given by $\sigma = 1/2$. With this picture in mind we mention the following important conjectures.

Conjecture 4 (The Riemann Hypothesis.) *We have $\sigma = 1/2$ for all zeros of $\xi(s)$.*

Conjecture 5 *All zeros of $\xi(s)$ are simple.*

In the absence of a proof of Conjecture 5 we adopt the convention that in any sum or product over zeros, we shall count them according to multiplicity.

6 The Product Formula

There is a useful product formula for $\xi(s)$, due to Hadamard. In general we have the following result, for which see Davenport [3, Chapter 11] for example.

Theorem 9 *Let $f(z)$ be an entire function with $f(0) \neq 0$, and suppose that there are constants $A > 0$ and $\theta < 2$ such that $f(z) = O(\exp(A|z|^\theta))$ for all complex z. Then there are constants α and β such that*

$$f(z) = e^{\alpha+\beta z} \prod_{n=1}^{\infty} \{(1 - \frac{z}{z_n})e^{z/z_n}\},$$

where z_n runs over the zeros of $f(z)$ counted with multiplicity. The infinite sum $\sum_{n=1}^{\infty} |z_n|^{-2}$ converges, so that the product above is absolutely and uniformly convergent in any compact set which includes none of the zeros.

We can apply this to $\xi(s)$, since it is apparent from Theorem 8, together with the definition (7) that

$$\xi(0) = \xi(1) = \frac{1}{2}\pi^{-1/2}\Gamma(\frac{1}{2})\mathrm{Res}\{\zeta(s); s = 1\} = \frac{1}{2}.$$

For $\sigma \geq 2$ one has $\zeta(s) = O(1)$ directly from the series (2), while Stirling's approximation yields $\Gamma(s/2) = O(\exp(|s| \log |s|))$. It follows that $\xi(s) = O(\exp(|s| \log |s|))$ whenever $\sigma \geq 2$. Moreover, when $\frac{1}{2} \leq \sigma \leq 2$ one sees from Theorem 8 that $\xi(s)$ is bounded. Thus, using the functional equation, we can deduce that $\xi(s) = O(\exp(|s| \log |s|))$ for all s with $|s| \geq 2$.

We may therefore deduce from Theorem 9 that

$$\xi(s) = e^{\alpha+\beta s} \prod_{\rho} \{(1 - \frac{s}{\rho})e^{s/\rho}\},$$

where ρ runs over the zeros of $\xi(s)$. Thus, with appropriate branches of the logarithms, we have

$$\log \xi(s) = \alpha + \beta s + \sum_{\rho} \{\log(1 - \frac{s}{\rho}) + \frac{s}{\rho}\}.$$

We can then differentiate termwise to deduce that

$$\frac{\xi'}{\xi}(s) = \beta + \sum_{\rho}\{\frac{1}{s-\rho} + \frac{1}{\rho}\},$$

the termwise differentiation being justified by the local uniform convergence of the resulting sum. We therefore deduce that

$$\frac{\zeta'}{\zeta}(s) = \beta - \frac{1}{s-1} + \frac{1}{2}\log\pi - \frac{1}{2}\frac{\Gamma'}{\Gamma}(\frac{s}{2}+1) + \sum_{\rho}\{\frac{1}{s-\rho} + \frac{1}{\rho}\}, \qquad (8)$$

where, as ever, ρ runs over the zeros of ξ counted according to multiplicity. In fact, on taking $s \to 1$, one can show that

$$\beta = -\frac{1}{2}\gamma - 1 - \frac{1}{2}\log 4\pi,$$

where γ is Euler's constant. However we shall make no use of this fact.

7 The Functions $N(T)$ and $S(T)$

We shall now investigate the frequency of the zeros ρ. We define

$$N(T) = \#\{\rho = \beta + i\gamma : 0 \le \beta \le 1, 0 \le \gamma \le T\}.$$

The notation $\beta = \Re(\rho)$, $\gamma = \Im(\rho)$ is standard. In fact one can easily show that $\psi(x) < (2\sqrt{x})^{-1}$, whence (6) suffices to prove that $\zeta(s) < 0$ for real $s \in (0,1)$. Thus we have $\gamma > 0$ for any zero counted by $N(T)$.

The first result we shall prove is the following.

Theorem 10 *If T is not the ordinate of a zero, then*

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} + S(T) + O(1/T),$$

where

$$S(T) = \frac{1}{\pi}\arg\zeta(\frac{1}{2} + iT),$$

is defined by continuous variation along the line segments from 2 to $2+iT$ to $\frac{1}{2}+iT$.

We shall evaluate $N(T)$ using the Principle of the Argument, which shows that

$$N(T) = \frac{1}{2\pi}\Delta_R \arg\xi(s),$$

providing that T is not the ordinate of any zero. Here R is the rectangular path joining 2, $2+iT$, $-1+iT$, and -1. To calculate $\Delta_R \arg \xi(s)$ one starts with any branch of $\arg \xi(s)$ and allows it to vary continuously around the path. Then $\Delta_R \arg \xi(s)$ is the increase in $\arg \xi(s)$ along the path. Our assumption about T ensures that $\xi(s)$ does not vanish on R.

Now $\xi(s) = \xi(1-s)$ and $\xi(1-s) = \overline{\xi(1-\overline{s})}$, whence $\xi(\frac{1}{2}+a+ib)$ is conjugate to $\xi(\frac{1}{2}-a+ib)$. (In particular this shows that $\xi(\frac{1}{2}+it)$ is always real.) It follows that

$$\Delta_R \xi(s) = 2\Delta_P \xi(s),$$

where P is the path $\frac{1}{2} \to 2 \to 2+iT \to \frac{1}{2}+iT$. On the first line segment $\xi(s)$ is real and strictly positive, so that the contribution to $\Delta_P \xi(s)$ is zero. Let L be the remaining path $2 \to 2+iT \to \frac{1}{2}+iT$. Then

$$\Delta_L \xi(s) = \Delta_L \{\arg(s-1)\pi^{-s/2}\Gamma(\frac{s}{2}+1)\} + \Delta_L \arg \zeta(s).$$

Now on L the function $s-1$ goes from 1 to $-\frac{1}{2}+iT$, whence

$$\Delta_L \arg(s-1) = \arg(-\frac{1}{2}+iT) = \frac{\pi}{2} + O(T^{-1}).$$

We also have

$$\arg \pi^{-s/2} = \Im \log \pi^{-s/2} = \Im(-\frac{s}{2}\log \pi),$$

so that $\arg \pi^{-s/2}$ goes from 0 to $-(T \log \pi)/2$ and

$$\Delta_L \arg \pi^{-s/2} = -\frac{T}{2}\log \pi.$$

Finally, Stirling's formula yields

$$\log \Gamma(z) = (z-\frac{1}{2})\log z - z + \frac{1}{2}\log(2\pi) + O(|z|^{-1}), \quad (|\arg(z)| \le \pi - \delta), \tag{9}$$

whence

$$\begin{aligned} \Delta_L \arg \Gamma(\frac{s}{2}+1) &= \Im \log \Gamma(\frac{\frac{1}{2}+iT}{2}+1) \\ &= \Im\{(\frac{3}{4}+i\frac{T}{2})\log(\frac{5}{4}+i\frac{T}{2}) - (\frac{5}{4}+i\frac{T}{2}) + \frac{1}{2}\log(2\pi)\} \\ &\qquad +O(1/T) \\ &= \frac{T}{2}\log\frac{T}{2} - \frac{T}{2} + \frac{3\pi}{8} + O(1/T), \end{aligned}$$

since

$$\log(\frac{5}{4}+i\frac{T}{2}) = \log\frac{T}{2} + i\frac{\pi}{2} + O(1/T).$$

These results suffice for Theorem 10

We now need to know about $S(T)$. Here we show the following.

Theorem 11 *We have* $S(T) = O(\log T)$.

Corollary 2 (The Riemann – von Mangoldt Formula). *We have*

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T).$$

We start the proof by taking $s = 2 + iT$ in (3) and noting that

$$|\frac{\zeta'}{\zeta}(s)| \leq \sum_{n=2}^{\infty}\frac{\Lambda(n)}{n^2} = O(1).$$

Thus the partial fraction decomposition (8) yields

$$\sum_{\rho}\{\frac{1}{2+iT-\rho} + \frac{1}{\rho}\} = \frac{1}{2}\frac{\Gamma'}{\Gamma}(2+\frac{iT}{2}) + O(1).$$

We may differentiate (9), using Cauchy's formula for the first derivative, to produce

$$\frac{\Gamma'}{\Gamma}(z) = \log z + O(1), \quad (|\arg(z)| \leq \pi - \delta), \tag{10}$$

and then deduce that

$$\sum_{\rho}\{\frac{1}{2+iT-\rho} + \frac{1}{\rho}\} = O(\log(2+T)). \tag{11}$$

We have only assumed here that $T \geq 0$, not that $T \geq 2$. In order to get the correct order estimate when $0 \leq T \leq 2$ we have therefore written $O(\log(2+T))$, which is $O(1)$ for $0 \leq T \leq 2$.

Setting $\rho = \beta + i\gamma$ we now have

$$\Re\frac{1}{2+iT-\rho} = \frac{2-\beta}{(2-\beta)^2+(T-\gamma)^2} \geq \frac{1}{4+(T-\gamma)^2}$$

and

$$\Re\frac{1}{\rho} = \frac{\beta}{\beta^2+\gamma^2} \geq 0,$$

since $0 \leq \beta \leq 1$. We therefore produce the useful estimate

$$\sum_{\rho}\frac{1}{4+(T-\gamma)^2} = O(\log(2+T)), \tag{12}$$

which implies in particular that

$$\#\{\rho : T-1 \leq \gamma \leq T+1\} = O(\log(2+T)). \tag{13}$$

We now apply (8) with $s = \sigma + iT$ and $0 \leq \sigma \leq 2$, and subtract (11) from it to produce

$$\frac{\zeta'}{\zeta}(\sigma + iT) = -\frac{1}{\sigma + iT - 1} + \sum_{\rho} \{\frac{1}{\sigma + iT - \rho} - \frac{1}{2 + iT - \rho}\} + O(\log(2 + T)).$$

Terms with $|\gamma - T| > 1$ have

$$\begin{aligned} |\frac{1}{\sigma + iT - \rho} - \frac{1}{2 + iT - \rho}| &= |\frac{2 - \sigma}{(\sigma + iT - \rho)(2 + iT - \rho)}| \\ &\leq \frac{2}{|\gamma - T|.|\gamma - T|} \\ &\leq \frac{2}{\frac{1}{5}\{4 + (T - \gamma)^2\}}. \end{aligned}$$

Thus (12) implies that

$$\sum_{\rho:\ |\gamma - T| > 1} \{\frac{1}{\sigma + iT - \rho} - \frac{1}{2 + iT - \rho}\} = O(\log(2 + T)),$$

and hence that

$$\begin{aligned} \frac{\zeta'}{\zeta}(\sigma + iT) &= -\frac{1}{\sigma + iT - 1} + \sum_{\rho:\ |\gamma - T| \leq 1} \{\frac{1}{\sigma + iT - \rho} - \frac{1}{2 + iT - \rho}\} \\ &\quad + O(\log(2 + T)). \end{aligned}$$

However we also have

$$|\frac{1}{2 + iT - \rho}| \leq \frac{1}{2 - \beta} \leq 1,$$

whence (13) produces

$$\sum_{\rho:\ |\gamma - T| \leq 1} \frac{1}{2 + iT - \rho} = O(\log(2 + T)).$$

We therefore deduce the following estimate.

Lemma 1 *For $0 \leq \sigma \leq 2$ and $T \geq 0$ we have*

$$\frac{\zeta'}{\zeta}(\sigma + iT) = -\frac{1}{\sigma + iT - 1} + \sum_{\rho:\ |\gamma - T| \leq 1} \frac{1}{\sigma + iT - \rho} + O(\log(2 + T)).$$

We are now ready to complete our estimation of $S(T)$. Taking $T \geq 2$, we have

$$\arg \zeta(\frac{1}{2} + iT) = \Im \log \zeta(\frac{1}{2} + iT) = \Im \int_2^{1/2+iT} \frac{\zeta'}{\zeta}(s)ds,$$

the path of integration consisting of the line segments from 2 to $2 + iT$ and from $2 + iT$ to $1/2 + iT$. Along the first of these we use the formula (3), which yields

$$\int_2^{2+iT} \frac{\zeta'}{\zeta}(s)ds = \left[\sum_{n=2}^{\infty} \frac{\Lambda(n)}{n^s \log n}\right]_2^{2+iT} = O(1).$$

For the remaining range we use Lemma 1, which produces

$$\begin{aligned}
\Im \int_{2+iT}^{1/2+iT} \frac{\zeta'}{\zeta}(s)ds &= \sum_{\rho:\, |\gamma-T|\leq 1} \Im \int_{2+iT}^{1/2+iT} \frac{ds}{s-\rho} + O(\log T) \\
&= \sum_{\rho:\, |\gamma-T|\leq 1} \Im\{\log(\frac{1}{2} + iT - \rho) - \log(2 + iT - \rho)\} \\
&\qquad\qquad +O(\log T) \\
&= \sum_{\rho:\, |\gamma-T|\leq 1} \{\arg(\frac{1}{2} + iT - \rho) - \arg(2 + iT - \rho)\} \\
&\qquad\qquad +O(\log T) \\
&= \sum_{\rho:\, |\gamma-T|\leq 1} O(1) + O(\log T) \\
&= O(\log T),
\end{aligned}$$

by (13). This suffices for the proof of Theorem 11.

8 The Non-Vanishing of $\zeta(s)$ on $\sigma = 1$

So for we know only that the non-trivial zeros of $\zeta(s)$ lie in the critical strip $0 \leq \sigma \leq 1$. Qualitatively the only further information we have is that there are no zeros on the boundary of this strip.

Theorem 12 *(Hadamard and de la Vallée Poussin, independently, 1896.) We have $\zeta(1 + it) \neq 0$, for all real t.*

This result was the key to the proof of the Prime Number Theorem. Quantitatively one can say a little more.

Theorem 13 *(De la Vallée Poussin.) There is a positive absolute constant c such that for any $T \geq 2$ there are no zeros of $\zeta(s)$ in the region*

$$\sigma \geq 1 - \frac{c}{\log T}, \quad |t| \leq T.$$

In fact, with much more work, one can replace the function $c/\log T$ by one that tends to zero slightly more slowly, but that will not concern us here. The proof of Theorem 13 uses the following simple fact.

Lemma 2 *For any real θ we have*

$$3 + 4\cos\theta + \cos 2\theta \geq 0.$$

This is obvious, since

$$3 + 4\cos\theta + \cos 2\theta = 2\{1 + \cos\theta\}^2.$$

We now use the identity (3) to show that

$$\begin{aligned} -3\frac{\zeta'}{\zeta}(\sigma) - 4\Re\frac{\zeta'}{\zeta}(\sigma + it) - \Re\frac{\zeta'}{\zeta}(\sigma + 2it) & \\ = \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n^\sigma}\{3 + 4\cos(t\log n) + \cos(2t\log n)\} & \\ \geq 0, & \end{aligned}$$

for $\sigma > 1$. When $1 < \sigma \leq 2$ we have

$$-\frac{\zeta'}{\zeta}(\sigma) = \frac{1}{\sigma - 1} + O(1),$$

from the Laurent expansion around the pole at $s = 1$. For the remaining two terms we use Lemma 1, to deduce that

$$\frac{3}{\sigma - 1} + O(1) - 4\Re \sum_{\rho:\ |\gamma - t| \leq 1} \frac{1}{\sigma + it - \rho} - \Re \sum_{\rho:\ |\gamma - 2t| \leq 1} \frac{1}{\sigma + 2it - \rho} + O(\log T)$$

$$\geq 0$$

for $1 < \sigma \leq 2$, $T \geq 2$, and $|t| \leq T$. Suppose we have a zero $\rho_0 = \beta_0 + i\gamma_0$, say, with $0 \leq \gamma_0 \leq T$. Set $t = \gamma_0$. We then observe that for any zero we have

$$\Re\frac{1}{\sigma + it - \rho} = \frac{\sigma - \beta}{(\sigma - \beta)^2 + (t - \gamma)^2} \geq 0,$$

since $\sigma > 1 \geq \beta$, and similarly

$$\Re \frac{1}{\sigma + 2it - \rho} \geq 0.$$

We can therefore drop all terms from the two sums above, with the exception of the term corresponding to $\rho = \rho_0$, to deduce that

$$\frac{4}{\sigma - \beta_0} \leq \frac{3}{\sigma - 1} + O(\log T).$$

Suppose that the constant implied by the $O(\ldots)$ notation is c_0. This is just a numerical value that one could calculate with a little effort. Then

$$\frac{4}{\sigma - \beta_0} \leq \frac{3}{\sigma - 1} + c_0 \log T$$

whenever $1 < \sigma \leq 2$. If $\beta_0 = 1$ we get an immediate contradiction by choosing $\sigma = 1 + (2c_0 \log T)^{-1}$. If $\beta_0 < 3/4$ the result of Theorem 13 is immediate. For the remaining range of β_0 we choose $\sigma = 1 + 4(1 - \beta_0)$, which will show that

$$\frac{4}{5(1 - \beta_0)} \leq \frac{3}{4(1 - \beta_0)} + c_0 \log T.$$

Thus

$$\frac{1}{20(1 - \beta_0)} \leq c_0 \log T,$$

and hence

$$1 - \beta_0 \geq \frac{1}{20 c_0 \log T}.$$

This completes the proof of Theorem 13.

The reader should observe that the key feature of the inequality given in Lemma 2 is that the coefficients are non-negative, and that the coefficient of $\cos\theta$ is strictly greater than the constant term. In particular, the inequality

$$1 + \cos\theta \geq 0$$

just fails to work.

Theorem 13 has a useful corollary.

Corollary 3 *Let c be as in Theorem 13, and let $T \geq 2$. Then if*

$$1 - \frac{c}{2 \log T} \leq \sigma \leq 2$$

and $|t| \leq T$, we have

$$\frac{\zeta'}{\zeta}(\sigma + it) = -\frac{1}{\sigma + it - 1} + O(\log^2 T).$$

For the proof we use Lemma 1. The sum over zeros has $O(\log T)$ terms, by (13), and each term is $O(\log T)$, since

$$\sigma - \beta \geq \frac{c}{2\log T},$$

by Theorem 13.

9 Proof of the Prime Number Theorem

Since our argument is based on the formula (3), it is natural to work with $\Lambda(n)$. We define

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = \sum_{p^k \leq x} \log p. \tag{14}$$

This is not the same function as that defined in (4)! Our sum $\psi(x)$ is related to $\pi(x)$ in the following lemma.

Lemma 3 *For $x \geq 2$ we have*

$$\pi(x) = \frac{\psi(x)}{\log x} + \int_2^x \frac{\psi(t)}{t \log^2 t} dt + O(x^{1/2}).$$

For the proof we begin by setting

$$\theta(x) = \sum_{p \leq x} \log p.$$

Then

$$\begin{aligned} \int_2^x \frac{\theta(t)}{t \log^2 t} dt &= \int_2^x \sum_{p \leq t} \frac{\log p}{t \log^2 t} dt \\ &= \sum_{p \leq x} \int_p^x \frac{\log p}{t \log^2 t} dt \\ &= \sum_{p \leq x} \left[-\frac{\log p}{\log t} \right]_p^x \\ &= \pi(x) - \frac{\theta(x)}{\log x}, \end{aligned}$$

so that

$$\pi(x) = \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t \log^2 t} dt. \tag{15}$$

However it is clear that terms in (14) with $k \geq 2$ have $p \leq x^{1/2}$, and there are at most $x^{1/2}$ such p. Moreover $k \leq \log x / \log p$, whence the total contribution from terms with $k \geq 2$ is $O(x^{1/2} \log x)$. Thus

$$\psi(x) = \theta(x) + O(x^{1/2} \log x).$$

If we substitute this into (15) the required result follows.

We will use contour integration to relate $\psi(x)$ to $\zeta'(s)/\zeta(s)$. This will be done via the following result.

Lemma 4 *Let $y > 0$, $c > 1$ and $T \geq 1$. Define*

$$I(y,T) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{y^s}{s} ds.$$

Then

$$I(y,T) = \left\{ \begin{array}{ll} 0, & 0 < y < 1 \\ 1, & y > 1 \end{array} \right\} + O(\frac{y^c}{T|\log y|}).$$

When $0 < y < 1$ we replace the path of integration by the line segments $c - iT \to N - iT \to N + iT \to c + iT$, and let $N \to \infty$. Then

$$\int_{N-iT}^{N+iT} \frac{y^s}{s} ds \to 0,$$

while

$$\int_{c-iT}^{N-iT} \frac{y^s}{s} ds = O(\int_c^N \frac{y^\sigma}{T}\, d\sigma) = O(\frac{y^c}{T|\log y|}),$$

and similarly for the integral from $N + iT$ to $c + iT$. It follows that

$$I(y,T) = O(\frac{y^c}{T|\log y|})$$

for $0 < y < 1$. The case $y > 1$ can be treated analogously, using the path $c - iT \to -N - iT \to -N + iT \to c + iT$. However in this case we pass a pole at $s = 0$, with residue 1, and this produces the corresponding main term for $I(y,T)$.

We can now give our formula for $\psi(x)$.

Theorem 14 *For $x - \frac{1}{2} \in \mathbb{N}$, $\alpha = 1 + 1/\log x$ and $T \geq 1$ we have*

$$\psi(x) = \frac{1}{2\pi i} \int_{\alpha-iT}^{\alpha+iT} \{-\frac{\zeta'}{\zeta}(s)\} \frac{x^s}{s} ds + O(\frac{x \log^2 x}{T}).$$

For the proof we integrate termwise to get

$$\begin{aligned}\frac{1}{2\pi i}\int_{\alpha-iT}^{\alpha+iT}\{-\frac{\zeta'}{\zeta}(s)\}\frac{x^s}{s}ds &= \sum_{n=2}^{\infty}\Lambda(n)I(\frac{x}{n},T)\\ &= \sum_{n\le x}\Lambda(n)+O(\sum_{n=2}^{\infty}\Lambda(n)(\frac{x}{n})^{\alpha}\frac{1}{T|\log x/n|}).\end{aligned}$$

Since we are taking $x-\frac{1}{2}\in\mathbb{N}$ the case $x/n=1$ does not occur. In the error sum those terms with $n\le x/2$ or $n\ge 2x$ have $|\log x/n|\ge\log 2$. Such terms therefore contribute

$$\begin{aligned}O(\sum_{n=2}^{\infty}\Lambda(n)\frac{x^{\alpha}}{Tn^{\alpha}}) &= O(\frac{x^{\alpha}}{T}|\frac{\zeta'}{\zeta}(\alpha)|)\\ &= O(\frac{x^{1+1/\log x}}{T}\frac{1}{\alpha-1})\\ &= O(\frac{x\log x}{T}).\end{aligned}$$

When $x/2<n<2x$ we have

$$|\log x/n|\ge\frac{1}{2}\frac{|x-n|}{x}$$

and

$$\Lambda(n)(\frac{x}{n})^{\alpha}=O(\log x).$$

These terms therefore contribute

$$\sum_{x/2<n<2x}O(\frac{x\log x}{T|x-n|})=O(\frac{x\log^2 x}{T})$$

on bearing in mind that $x-\frac{1}{2}\in\mathbb{N}$. The theorem now follows.

We are now ready to prove the following major result.

Theorem 15 *There is a positive constant c_0 such that*

$$\psi(x)=x+O(x\exp\{-c_0\sqrt{\log x}\}) \tag{16}$$

for all $x\ge 2$. Moreover we have

$$\pi(x)=\mathrm{Li}(x)+O(x\exp\{-c_0\sqrt{\log x}\})$$

for all $x\ge 2$.

The error terms here can be improved slightly, but with considerably more work.

It clearly suffices to consider the case in which $x - \frac{1}{2} \in \mathbb{N}$. To prove the result we set

$$\mu = 1 - \frac{c}{2\log T}, \quad T \geq 2,$$

as in Lemma 3, and replace the line of integration in Theorem 14 by the path $\alpha - iT \to \mu - iT \to \mu + iT \to \alpha + iT$. The integrand has a pole at $s = 1$ with residue x, arising from the pole of $\zeta(s)$, but no other singularities, by virtue of Theorem 13. On the new path of integration Lemma 3 shows that

$$\frac{\zeta'}{\zeta}(s) = O(\log^2 T).$$

We therefore deduce that

$$\psi(x) = x + O(\frac{x\log^2 x}{T}) + O(\int_\mu^\alpha \frac{\log^2 T}{T} x^\sigma d\sigma) + O(\int_{-T}^T \frac{\log^2 T}{|\mu + it|} x^\mu dt),$$

where the first error integral corresponds to the line segments $\alpha - iT \to \mu - iT$ and $\mu + iT \to \alpha + iT$, and the second error integral to the segment $\mu - iT \to \mu + iT$. These integrals are readily estimated to yield

$$\psi(x) = x + O(\frac{x\log^2 x}{T}) + O(\frac{\log^2 T}{T} x^\alpha) + O(x^\mu \log^3 T).$$

Of course $x^\alpha = O(x)$ here. Thus if $T \leq x$ we merely get

$$\psi(x) = x + O(x\log^3 x\{\frac{1}{T} + x^{\mu-1}\}).$$

We now choose

$$T = \exp\{\sqrt{\log x}\},$$

whence

$$\psi(x) = x + O(x(\log x)^3 \exp\{-\min(1, \frac{c}{2})\sqrt{\log x}\}).$$

We may therefore choose any positive constant $c_0 < \min(1, \frac{c}{2})$ in Theorem 15. This establishes (16). To prove the remaining assertion, it suffices to insert (16) into Lemma 3.

Finally we should stress that the success of this argument depends on being able to take $\mu < 1$, since there is an error term which is essentially of order x^μ. Thus it is crucial that we should at least know that $\zeta(1 + it) \neq 0$.

If we assume the Riemann Hypothesis, then we may take any $\mu > \frac{1}{2}$ in the above analysis. This leads to the following estimates.

Theorem 16 *On the Riemann Hypothesis we have*

$$\psi(x) = x + O(x^\theta)$$

and

$$\pi(x) = \mathrm{Li}(x) + O(x^\theta)$$

for any $\theta > \frac{1}{2}$ *and all* $x \geq 2$.

One cannot reduce the exponent below 1/2, since there is a genuine contribution to $\pi(x)$ arising from the zeros of $\zeta(s)$.

10 Explicit Formulae

In this section we shall argue somewhat informally, and present results without proof.

If $f : (0, \infty) \to \mathbb{C}$ we define the Mellin transform of f to be the function

$$F(s) := \int_0^\infty f(x) x^{s-1} dx.$$

By a suitable change of variables one sees that this is essentially a form of Fourier transform. Indeed all the properties of Mellin transforms can readily be translated from standard results about Fourier transforms. In particular, under suitable conditions one has an inversion formula

$$f(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} F(s) x^{-s} ds.$$

Arguing purely formally one then has

$$\begin{aligned} \sum_{n=2}^\infty \Lambda(n) f(n) &= \sum_{n=2}^\infty \Lambda(n) \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} F(s) n^{-s} ds \\ &= \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \{-\frac{\zeta'}{\zeta}(s)\} F(s) ds. \end{aligned}$$

If one now moves the line of integration to $\Re(s) = -N$ one passes poles at $s = 1$ and at $s = \rho$ for every non-trivial zero ρ, as well as at the trivial zeros $-2n$. Under suitable conditions the integral along $\Re(s) = -N$ will tend to 0 as $N \to \infty$. This argument leads to the following result.

Theorem 17 *Suppose that* $f \in C^2(0, \infty)$ *and that* $\mathrm{supp}(f) \subseteq [1, X]$ *for some* X. *Then*

$$\sum_{n=2}^\infty \Lambda(n) f(n) = F(1) - \sum_\rho F(\rho) - \sum_{n=1}^\infty F(-2n).$$

One can prove such results subject to weaker conditions on f. If x is given, and

$$f(t) = \begin{cases} 1, & t \leq x \\ 0, & t > x, \end{cases}$$

then the conditions above are certainly not satisfied, but we have the following related result.

Theorem 18 (The Explicit Formula.) *Let $x \geq T \geq 2$. Then*

$$\psi(x) = x - \sum_{\rho:\, |\gamma| \leq T} \frac{x^\rho}{\rho} + O(\frac{x \log^2 x}{T}).$$

For a proof of this see Davenport [3, Chapter 17], for example. There are variants of this result containing a sum over all zeros, and with no error term, but the above is usually more useful.

The explicit formula shows exactly how the zeros influence the behaviour of $\psi(x)$, and hence of $\pi(x)$. The connection between zeros and primes is even more clearly shown by the following result of Landau.

Theorem 19 *For fixed positive real x define $\Lambda(x) = 0$ if $x \notin \mathbb{N}$ and $\Lambda(x) = \Lambda(n)$ if $x = n \in \mathbb{N}$. Then*

$$\Lambda(x) = -\frac{2\pi}{T} \sum_{\rho:\, 0 < \gamma \leq T} x^\rho + O_x(\frac{\log T}{T}),$$

where $O_x(\ldots)$ indicates that the implied constant may depend on x.

This result shows that the zeros precisely determine the primes. Thus, for example, one can reformulate the conjecture (1) as a statement about the zeros of the zeta-function. All the unevenness of the primes, for example the behaviour described by Theorem 5, is encoded in the zeros of the zeta-function. It therefore seems reasonable to expect that the zeros themselves should have corresponding unevenness.

11 Dirichlet Characters

We now turn to the simplest type of generalization of the Riemann Zeta-function, namely the Dirichlet L-functions. In the remainder of these notes we shall omit most of the proofs, being content merely to describe what can be proved.

A straightforward example of a Dirichlet L-function is provided by the infinite series

$$1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \frac{1}{9^s} - \frac{1}{11^s} + \dots . \tag{17}$$

We first need to describe the coefficients which arise.

Definition *Let $q \in \mathbb{N}$. A "(Dirichlet) character χ to modulus q" is a function $\chi : \mathbb{Z} \to \mathbb{C}$ such that*

(i) $\chi(mn) = \chi(m)\chi(n)$ *for all* $m, n \in \mathbb{Z}$;

(ii) $\chi(n)$ *has period* q;

(iii) $\chi(n) = 0$ *whenever* $(n, q) \neq 1$; *and*

(iv) $\chi(1) = 1$.

Part (iv) of the definition is necessary merely to rule out the possibility that χ is identically zero.

As an example we can take the function

$$\chi(n) = \begin{cases} 1, & n \equiv 1 \pmod 4, \\ -1, & n \equiv 3 \pmod 4, \\ 0, & n \equiv 0 \pmod 2. \end{cases} \tag{18}$$

This is a character modulo 4, and is the one generating the series (17). A second example is the function

$$\chi_0(n) := \begin{cases} 1, & (n, q) = 1, \\ 0, & (n, q) \neq 1. \end{cases}$$

This produces a character for every modulus q, known as the *principal character* modulo q.

A number of key facts are gathered together in the following theorem.

Theorem 20 *(i) We have $|\chi(n)| = 1$ for every n coprime to q.*

(ii) If χ_1 and χ_2 are two characters to modulus q, then so is $\chi_1\chi_2$, where we define $\chi_1\chi_2(n) = \chi_1(n)\chi_2(n)$.

(iii) There are exactly $\phi(q)$ different characters to modulus q.

(iv) If $n \not\equiv 1 \pmod q$ then there is at least one character χ modulo q for which $\chi(n) \neq 1$.

In part (iii) the function $\phi(q)$ is the number of positive integers $n \leq q$ for which n and q are coprime.

To prove part (i) we note that the sequence $n^k \pmod q$ must eventually repeat when k runs through $\mathbb{N}$. Thus there exist $k < j$ with $\chi(n^k) = \chi(n^j)$, and hence $\chi(n)^k = \chi(n)^j$. Since n is coprime to q we have $\chi(n) \neq 0$, so that $\chi(n)^{j-k} = 1$.

Part (ii) of the theorem is obvious, but parts (iii) and (iv) are harder, and we refer the reader to Davenport [3, Chapter 4] for the details. As an example of (iii) we note that $\phi(4) = 2$, and we have already found two characters modulo 4. There are no others.

One further fact may elucidate the situation. Consider a general finite abelian group G — In our case we will have $G = (\mathbb{Z}/q\mathbb{Z})^\times$. Define $\widehat{G}$ to be the group of homomorphisms $\theta : G \to \mathbb{C}^\times$, where the group action is given by $(\theta_1\theta_2)(g) := \theta_1(g)\theta_2(g)$. In our case these homomorphisms are, in effect, the characters. Then the groups G and $\widehat{G}$ are isomorphic, and part (iii) above is an immediate consequence. The details can be found in Ireland and Rosen [7, pages 253 and 254], for example. Indeed there is a duality between G and $\widehat{G}$. The isomorphism between them is not "natural", but there is a natural isomorphism

$$G \simeq \widehat{\widehat{G}}.$$

There are two orthogonality relations satisfied by the characters to a given modulus q. The first of these is the following.

Theorem 21 *If a and q are coprime then*

$$S := \sum_{\chi \pmod q} \chi(n)\overline{\chi(a)} = \begin{cases} \phi(q), & n \equiv a \pmod q, \\ 0, & n \not\equiv a \pmod q. \end{cases}$$

When $n \equiv a \pmod q$ this is immediate since then $\chi(n)\overline{\chi(a)} = 1$ for all χ. In the remaining case, choose an element b with $ab \equiv 1 \pmod q$. By (iv) of Theorem 20 there exists a character χ_1 such that $\chi_1(nb) \neq 1$. Then

$$\chi_1(nb)S = \sum_{\chi \pmod q} \chi_1(n)\chi(n)\chi_1(b)\overline{\chi(a)}.$$

However

$$\chi_1(b)\chi_1(a) = \chi_1(ab) = \chi_1(1) = 1,$$

whence $\chi_1(b) = \overline{\chi_1(a)}$. We therefore deduce that

$$\begin{aligned} \chi_1(nb)S &= \sum_{\chi \pmod q} \chi_1(n)\chi(n)\overline{\chi_1(a)\chi(a)} \\ &= \sum_{\chi \pmod q} \chi_1\chi(n)\overline{\chi_1\chi(a)}. \end{aligned}$$

As χ runs over the complete set of characters to modulus q the product $\chi_1\chi$ does as well, since $\chi_1\chi = \chi_1\chi'$ implies $\chi = \chi'$. Thus

$$\sum_{\chi \pmod q} \chi_1\chi(n)\overline{\chi_1\chi(a)} = S$$

and hence $\chi_1(nb)S = S$. Since $\chi_1(nb) \neq 1$ we deduce that $S = 0$, as required.

The second orthogonality relation is the following.

Theorem 22 *If $\chi \neq \chi_0$ then $\sum_{n=1}^{q} \chi(n) = 0$.*

The proof is analogous to the previous result, and is based on the obvious fact that if $\chi \neq \chi_0$ then there is some integer n coprime to q such that $\chi(n) \neq 1$. The details are left as an exercise for the reader.

If q has a factor r and χ is a character modulo r we can define the character ψ modulo q which is "induced by" χ. This is done by setting

$$\psi(n) = \begin{cases} \chi(n), & (n,q) = 1, \\ 0, & (n,q) \neq 1. \end{cases}$$

For example, we may take χ to be the character modulo 4 given by (18). Then if $q = 12$ we induce a character ψ modulo 12, as in the following table.

	1	2	3	4	5	6	7	8	9	10	11	12
χ	1	0	-1	0	1	0	-1	0	1	0	-1	0
ψ	1	0	0	0	1	0	-1	0	0	0	-1	0

A character $\chi \pmod q$ which *cannot* be produced this way from some divisor $r < q$ is said to be "primitive". The principal character is induced by the character $\chi_0 \pmod 1$, that is to say by the character which is identically 1. If q is prime, then all the characters except for the principal character are primitive. In general there will be both primitive and imprimitive characters to each modulus. Imprimitive characters are a real nuisance!!

12 Dirichlet L-functions

For any character χ to modulus q we will define the corresponding Dirichlet L-function by setting

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}, \quad (\sigma > 1).$$

We content ourselves here with describing the key features of these functions, and refer the reader to Davenport [3], for example, for details.

The sum is absolutely convergent for $\sigma > 1$ and is locally uniformly convergent, so that $L(s,\chi)$ is holomorphic in this region. If χ is the principal character modulo q then the series fails to converge when $\sigma \leq 1$. However for non-principal χ the series is conditionally convergent when $\sigma > 0$, and the series defines a holomorphic function in this larger region.

There is an Euler product identity

$$L(s,\chi) = \prod_p (1 - \chi(p)p^{-s})^{-1}, \quad (\sigma > 1).$$

This can be proved in the same way as for $\zeta(s)$ using the multiplicativity of the function χ.

Suppose that χ is primitive, and that $\chi(-1) = 1$. If we apply the Poisson summation formula to

$$f(x) = e^{-(a+qx)^2\pi v/q},$$

multiply the result by $\chi(a)$, and sum for $1 \leq a \leq q$, we find that

$$\theta(v,\chi) = \frac{\tau(\chi)}{\sqrt{q}} \frac{1}{\sqrt{v}} \theta(\frac{1}{v}, \overline{\chi}),$$

where

$$\theta(v,\chi) := \sum_{n=-\infty}^{\infty} \chi(n) e^{-n^2\pi v/q}$$

is a generalisation of the theta-function, and

$$\tau(\chi) := \sum_{a=1}^{q} \chi(a) e^{2\pi i a/q}$$

is the *Gauss sum*.

When χ is primitive and $\chi(-1) = -1$ the function $\theta(v,\chi)$ vanishes identically. Instead we use

$$\theta_1(v,\chi) := \sum_{n=-\infty}^{\infty} n\chi(n) e^{-n^2\pi v/q},$$

for which one finds the analogous transformation formula

$$\theta_1(v,\chi) = \frac{i\tau(\chi)}{\sqrt{q}} \frac{1}{v^{3/2}} \theta_1(\frac{1}{v}, \overline{\chi}).$$

These two transformation formulae then lead to the analytic continuation and functional equation for $L(s,\chi)$. The conclusion is that, if χ is primitive then $L(s,\chi)$ has an analytic continuation to the whole complex plane, and is regular everywhere, except when χ is identically 1, (in which case $L(s,\chi)$ is just the Riemann Zeta-function $\zeta(s)$). Moreover, still assuming that χ is primitive, with modulus q, we set

$$\xi(s,\chi) = (\frac{q}{\pi})^{(s+a)/2}\Gamma(\frac{s+a}{2})L(s,\chi),$$

where

$$a = a(\chi) := \begin{cases} 0, & \chi(-1) = 1, \\ 1, & \chi(-1) = -1. \end{cases}$$

Then

$$\xi(1-s,\overline{\chi}) = \frac{i^a q^{1/2}}{\tau(\chi)}\xi(s,\chi).$$

Notice in particular that, unless the values taken by χ are all real, this functional equation relates $L(s,\chi)$ not to the same function at $1-s$ but to the conjugate L-function, with character $\overline{\chi}$.

It follows from the Euler product and the functional equation that there are no zeros of $\xi(s,\chi)$ outside the critical strip. The zeros will be symmetrically distributed about the critical line $\sigma = 1/2$, but unless χ is real they will not necessarily be symmetric about the real line. Hence in general it is appropriate to define

$$N(T,\chi) := \#\{\rho : \xi(\rho,\chi) = 0 \; |\gamma| \le T\},$$

counting zeros both above and below the real axis. We then have

$$\frac{1}{2}N(T,\chi) = \frac{T}{2\pi}\log\frac{qT}{2\pi} - \frac{T}{2\pi} + O(\log qT)$$

for $T \ge 2$, which can be seen as an analogue of the Riemann – von Mangoldt formula. This shows in particular that the interval $[T, T+1]$ contains around

$$\frac{1}{2\pi}\log\frac{qT}{2\pi}$$

zeros, on average.

The work on regions without zeros can be generalized, but there are serious problems with possible zeros on the real axis. Thus one can show that there is a constant $c > 0$, which is independent of q, such that if $T \ge 2$ then $L(s,\chi)$ has no zeros in the region

$$\sigma \ge 1 - \frac{c}{\log qT}, \quad 0 < |t| \le T.$$

If χ is not a real-valued character then we can extend this result to the case $t = 0$, but is a significant open problem to deal with the case in which χ is real. However in many other important aspects techniques used for the Riemann Zeta-function can be successfully generalized to handle Dirichlet L-functions.

References

[1] R.C. Baker, G. Harman, and J. Pintz, The difference between consecutive primes. II., *Proc. London Math. Soc. (3)*, 83 (2001), 532-562.

[2] J.-R. Chen, on the representation of a large even integer as a sum of a prime and a product of at most two primes, *Kexue Tongbao*, 17 (1966), 385-386.

[3] H. Davenport, *Multiplicative number theory,* Graduate Texts in Mathematics, 74. (Springer-Verlag, New York-Berlin, 1980).

[4] J.B. Friedlander and H. Iwaniec, The polynomial $X^2 + Y^4$ captures its primes, *Annals of Math. (2)*, 148 (1998), 945-1040.

[5] G.H. Hardy and E.M. Wright, *An introduction to the theory of numbers*, (Oxford University Press, New York, 1979).

[6] A.E. Ingham, On the difference between consecutive primes, *Quart. J. Math. Oxford*, 8 (1937), 255-266.

[7] K. Ireland and M. Rosen, *A classical introduction to modern number theory*, Graduate Texts in Math., 84, (Springer, Heidelberg-New York, 1990).

[8] H. Iwaniec, Almost-primes represented by quadratic polynomials, *Invent. Math.*, 47 (1978), 171-188.

[9] H. Maier, Primes in short intervals, *Michigan Math. J.*, 32 (1985), 221-225.

[10] H. Maier, Small differences between prime numbers, *Michigan Math. J.*, 35 (1988), 323-344.

[11] I.I. Piatetski-Shapiro, On the distribution of prime numbers in sequences of the form $[f(n)]$, *Mat. Sbornik N.S.*, 33(75) (1953), 559-566.

[12] H. Rademacher, *Topics in analytic number theory*, Grundlehren math. Wiss., 169, (Springer, New York-Heidelberg, 1973).

[13] R.A. Rankin, The difference between consecutive prime numbers, *J. London Math. Soc.*, 13 (1938), 242-247.

[14] J. Rivat and J. Wu, Prime numbers of the form $[n^c]$, *Glasg. Math. J.*, 43 (2001), 237-254.

[15] A. Selberg, On the normal density of primes in small intervals, and the difference between consecutive primes, *Arch. Math. Naturvid.*, 47, (1943), 87-105.

[16] D.K.L. Shiu, Strings of congruent primes, *J. London Math. Soc. (2)*, 61 (2000), 359-373.

Mathematical Institute,
24-29, St. Giles',
Oxford OX1 3LB

rhb@maths.ox.ac.uk