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

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

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.

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

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 $n+1, n+2, \dots, n+k$ 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 $n+1, n+2, \dots, n+k$ are all composite, is around $n^{-\mu}$.

If E_n is the event that $n+1, n+2, \dots, n+k$ 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*

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

the product being over primes, then

$$\#\{n \le x : n, n+2 \text{ both prime}\} \div c \int_2^x \frac{dt}{(\log t)^2}. \tag{1.1}$$

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

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

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

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

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 $cN \le H \le N$ 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

$$\zeta(s) := \sum_{n=1}^{\infty} n^{-s}, \quad \sigma > 1. \tag{3.1}$$

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

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

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

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

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

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

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

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

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

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

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*

$$-\frac{\zeta'}{\zeta}(s) = \sum_{n=2}^{\infty}\frac{\Lambda(n)}{n^s}, \quad (\sigma > 1), \tag{3.2}$$

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*

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

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

$$\psi(v) = \sum_{n=1}^{\infty} e^{-n^2\pi v}, \tag{4.1}$$

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

$$2\psi(v) + 1 = \frac{1}{\sqrt{v}}\{2\psi(\frac{1}{v}) + 1\}. \tag{4.2}$$

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 (4.2). 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

$$\zeta(s)\pi^{-s/2}\Gamma(\frac{s}{2}) = \int_1^{\infty} \{x^{s/2-1} + x^{(1-s)/2-1}\}\psi(x)dx - \frac{1}{s(1-s)}, \tag{4.3}$$