Chapter 3

The Prime Number Theorem

This chapter gives without proof the two basic results of analytic number theory.

3.1 The Theorem

Recall that if $f(x), g(x)$ are two real-valued functions, we write

$$f(x) \sim g(x)$$

to mean

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

Recall too that $\pi(x)$ denotes the number of primes $\leq x$.

Theorem 3.1. *As $x \to \infty$,*

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

This tells us that if we are looking at numbers around 10^6 then we may expect roughly 1 number in

$$\ln 10^6 = 6 \ln 10 \approx 14$$

to be prime. (Recall that $\ln x$ denotes the logarithm to base e.)

Note that this includes even numbers, which of course we know are not prime. So approximately 1 odd number in 7 around 1 million is prime.

When we get to 10^{12} the density is halved; about 1 number in 28 is prime.

3.2 A little history

The history of the prime number theorem is very interesting.

Although the result had been conjectured earlier by Lagrange and others, Gauss seems to have been the first to make a serious statistical study, starting when he was a teenager, and eventually examining a million primes or more.

He actually expressed the conjecture slightly differently, as

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

where $\mathrm{Li}(x)$ is the 'logarithmic integral'

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\ln t}.$$

It is reasonably easy to see, on integrating by parts, that

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

so this is equivalent to the form we gave. In fact $\mathrm{Li}(x)$ gives a slightly better estimate.

But the great break-through came in 1859 — so this is the 150th anniversary — when an 8-page paper by Riemann, "On the number of primes less than a given number", was published. [1]

Riemann's technique must have seemed almost like magic at the time. It was based on Euler's Product Formula,

$$1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots = \prod \left(1 - \frac{1}{p^s}\right)^{-1}$$

which we considered earlier, but Riemann now considered s as a *complex* variable. This was the start of analytic number theory — the use of complex functions in number theory.

The function

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots$$

on the left is now known as the Riemann zeta function, or just the zeta function if there is no ambiguity. (There are many different zeta functions in use today.) By convention, the letter s is used rather than z in analytic number theory, because that is what Riemann used.

Note that there is no difficulty in defining the terms in the series:

$$\frac{1}{n^s} = e^{-(\ln n)s}.$$

[1] Thanks to David Wilkins, you can find this paper in both German and English at `http://www.maths.tcd.ie/pub/HistMath/People/Riemann/Zeta/`

The series for $\zeta(s)$ only converges for $\Re(s) > 1$. But Riemann showed that $\zeta(s)$ can be extended to a meromorphic function (a function with only isolated poles) in the whole complex plane.

The function can be extended easily to the half-plane $\Re(s) > 0$ by noting that

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

Also it is not difficult to see that $\zeta(s)$ has a pole with residue 1 at $s = 1$, and that this is the only pole in the half-plane $\Re(s) > 0$:

$$\zeta(s) = \frac{1}{s-1} + f(s),$$

where $f(s)$ is homomorphic in the half-plane.

Riemann established — this is much more difficult — a 'functional equation' relating $\zeta(s)$ and $\zeta(1-s)$ by applying a variant of the Fourier transform. Note that the map $s \mapsto 1 - s$ is reflection in the point $1/2$, and sends the line $\Re(s) = 1/2$ into itself.

Riemann was led by his study of this relation to the conclusion that the zeros of $\zeta(s)$ in the 'critical strip' $0 \leq \Re(s) \leq 1$ all lie on the 'critical line' $\Re(s) = 1/2$. This is the *Riemann Hypothesis*, perhaps the most famous unsolved problem in mathematics.

Riemann showed that the Prime Number Theorem would follow from his Conjecture.

Riemann didn't prove the Theorem, but it was proved 40 years later using Riemann's method by two French mathematicians, Hadamard and de la Vallée Poussin. They showed that it was in fact sufficient to establish that $\zeta(s)$ had no zeros on the line $\Re(s) = 1$, which is relatively easy.

The reason why the *zeros* of $\zeta(s)$ are so important in that in studying the distribution of the primes we take the logarithmic derivative $\zeta'(s)/\zeta(s)$ (as in the proof that $\sum 1/p$ diverges in the last Chapter), and the zeros of $\zeta(s)$ give poles of this function. In general, the behaviour of a complex function is largely determined by its poles.

It has been shown that the first 10^{13} zeros all lie on the critical line. The technique is relatively simple. If $s = \sigma + it$ is a zero then so is its complex conjugate $\bar{s} = \sigma - it$, and it follows from the functional relation that $1 - \bar{s} = (1 - \sigma) + it$ is also a zero. Thus if s is not on the critical line then there is a pair of zeros with the same imaginary part t. It is (relatively) easy to compute the number of zeros with $0 \leq t \leq T$, say, and show that there is no sudden increase by 2.

3.3 Alternative version

Recall that p_n denotes the nth prime.

Proposition 3.1. *As $n \to \infty$,*

$$p_n \sim n \ln n.$$

Thus the millionth prime is around

$$1000000 \cdot 6 \cdot ln10 \approx 14000000.$$

We will show that this result is equivalent to the Prime Number Theorem, ie either result can be derived from the other.

Proof. This is a little exercise in the use of the asymptotic operator $\sim$.

Let

$$f(x) = \frac{x}{\ln x}, \quad g(x) = x \ln x.$$

These functions are "asymptotically inverse" in the sense that

$$f(g(x)) \sim x, \quad g(f(x)) \sim x.$$

For

$$\begin{aligned} f(g(x)) &= \frac{x \ln x}{\ln x + \ln \ln x} \\ &= x \frac{\ln x}{\ln x + \ln \ln x} \\ &\sim x, \end{aligned}$$

since

$$\frac{\ln x + \ln \ln x}{\ln x} = 1 + \frac{\ln x}{\ln \ln x} \sim 1.$$

Similarly

$$\begin{aligned} g(f(x)) &= \frac{x(\ln x - \ln \ln x)}{\ln x} \\ &= x \left(1 - \frac{\ln x}{\ln \ln x}\right) \\ &\sim x. \end{aligned}$$

Now suppose the Prime Number Theorem holds: By definition,

$$\pi(p_n) = n.$$

It follows that

$$g(\pi(p_m)) = g(n).$$

By hypothesis,

$$\pi(p_n) \sim f(p_n).$$

Hence

$$g(\pi(p_m)) \sim g(f(p_n)) \sim p_n,$$

from above. Thus

$$p_n \sim g(n) = n \ln n.$$

Now suppose

$$p_n \sim g(n).$$

Given x, let

$$p_n \leq x < p_{n+1}.$$

By definition,

$$\pi(x) = n.$$

Hence

$$f(p_n) \leq f(x) < f(p_{n+1}).$$

By hypothesis,

$$p_n \sim g(n).$$

Hence

$$f(p_n) \sim f(g(n)) \sim n,$$

and similarly

$$f(p_{n+1}) \sim n + 1 \sim n.$$

It follows that

$$\pi(x) = n \sim f(x) = \frac{x}{\ln x},$$

as required.

□

3.4 Dirichlet's Theorem

Theorem 3.2. *There are an infinity of primes in each arithmetic sequence*

$$an + b \qquad (\gcd(a, b) > 1).$$

(If a, b are *not* coprime, then there is at most one prime in the sequence, namely $\gcd(a, b)$.)

We are not going to give a proof of this result, but will give some indication of Dirichlet's argument.

His proof followed Riemann's method, with the addition of a technique for singling out primes in an arithmetic sequence, using *Dirichlet characters*.

If $m \in \mathbb{N}$, a Dirichlet character $\chi \mod m$ is a map

$$\chi : \mathbb{Z} \to \mathbb{C}$$

such that

1. $s \equiv t \bmod m \implies \chi(s) = \chi(t)$;
2. $\gcd(s, m) > 1 \implies \chi(s) = 0$;
3. $\chi(st) = \chi(s)\chi(t)$;
4. $\chi(1) = 1$.

(From (3),

$$\chi(1)^2 = \chi(1) \implies \chi(1) = 0 \text{ or } 1.$$

If $\chi(1) = 0$ then $\chi(r) = 0$ for all r. So condition (4) is added to exclude this trivial case.)

For those familiar with group representations, a Dirichlet character $\chi \bmod m$ can be identified with a character of the finite abelian group $(\mathbb{Z}/m)^*$ formed by the residue classes mod m coprime to m.

The *principal character* mod m is that given by

$$\chi(r) = \begin{cases} 1 \text{ if } \gcd(r, m) = 1, \\ 0 \text{ if } \gcd(r, m) > 1. \end{cases}$$

To the Dirichlet character χ we associate the *L-series*

$$L_\chi(s) = \sum_{n \geq 1} \frac{\chi(n)}{n^s}.$$

Because of the multiplicative property of the Dirichlet character, the Euler Product Formula extends to the L-series:

$$L_\chi(s) = 1 + \frac{\chi(2)}{2^s} + \frac{\chi(3)}{3^s} + \cdots = \prod \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

This allows Riemann's technique to be applied.

As with $\zeta(s)$, it is (relatively) easy to show that the principal L-function has a pole at $s = 1$, and no other pole in $\Re(s) > 0$, while the other L-functions have no poles at all in this region.

It is also relatively easy to show that if $\gcd(r, m) = 1$ then we can find a linear combination of Dirichlet characters mod m taking value 1 at r and 0 elsewhere:

$$c_1\chi_1(t) + c_2\chi_2(t) + \cdots c_d\chi_d(t) = \begin{cases} 1 \text{ if } t \equiv r \bmod m, \\ 0 \text{ otherwise.} \end{cases}$$

It follows that the corresponding linear combination

$$c_1L_{\chi_1}(s) + c_2L_{\chi_2}(s) + \cdots c_dL_{\chi_d}(s) = \sum_{n \equiv r \bmod m} \frac{1}{n^s}.$$

This only gives a taste of the proof of Dirichlet's Theorem. If you would like to learn more there is a highly-regarded (and not too difficult) book on the subject of zeta- and L-functions: "Riemann's Zeta Function" by Harold M. Edwards (available for under €10 from Amazon).

Like $\zeta(s)$, each L-function satsifies a 'functional equation' relating $L(s)$ and $L(1 - s)$. This allows $L_\chi(s)$ to be extended to the whole complex plane.

The *Generalised Riemann Hypothesis* (often abbreviated to GRH) says that the zeros of each L-functions in the critical strip $0 < \Re(s) < 1$ lie on the critical line $\Re(s) = 1/2$.

The argument used to prove the Prime Number Theorem allows Dirichlet's Theorem to be greatly strengthened: the primes are distributed evenly among the arithmetic sequences modulo m. More precisely, let $\pi(m, r, x)$ denote the number of primes

$$p \equiv r \bmod m,\ p \le x.$$

Then if $\gcd(r, m) = 1$,

$$\pi(m, r, x) \sim \frac{1}{\phi(m)} \frac{x}{\ln x},$$

where $\phi(m)$ is the number of remainders $r \in (0, m)$ coprime to m. (This is Euler's totient function, which we shall meet later.)

As an example, it follows that

$$\frac{\text{Number of primes } p \le x,\ p \equiv 1 \bmod 4}{\text{Number of primes } p \le x,\ p \equiv -1 \bmod 4} \to 1 \text{ as } x \to \infty.$$

Or as we can say crudely, primes $\equiv 1 \bmod 4$ appear roughly as often as primes $\equiv -1 \bmod 4$.