

On the Sequence of Prime Numbers

COSMIN DAVIDESCU

064140

Mathematicians have tried in vain to this day to discover some order in the sequence of prime numbers, and we have reason to believe that it is a mystery into which the mind will never penetrate.

– LEONHARD EULER

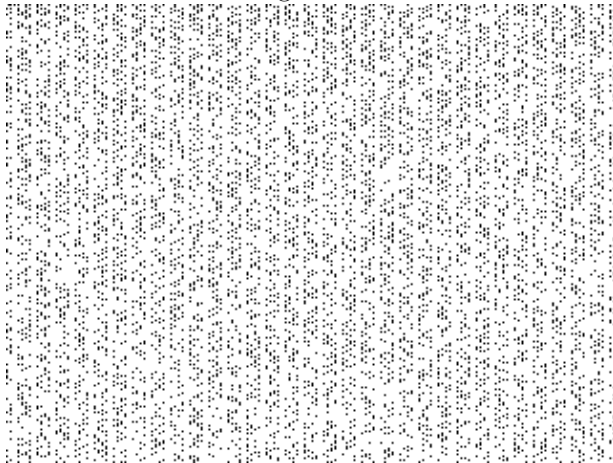
The prime numbers constitute at the same time one of the most basic and one of the most bewildering concepts in the whole of mathematics. They are, according to the fundamental theorem of arithmetic, the “atoms” or “building blocks” of the integers and have been studied ever since mankind started studying mathematics itself. Indeed, even today, you would be hard pressed to find an elementary number theory book or course that does not mention them at the beginning; Euclid’s proof of the infinity of primes is among the very first proofs any young mathematician encounters. At the same time, the primes are at the heart of many of modern mathematics’ hardest and most puzzling problems, from the famous Goldbach and Twin Prime conjectures to the Riemann Hypothesis, widely considered one of the most important unsolved problems of the whole of mathematics. Although they are among the “purest” of mathematical objects, their theory has applications in areas of cryptography and computer science. We think to know them well yet they constantly surprise us with their properties and behaviour. The whole theory of prime numbers is filled with such conflicting truths and indeed our very attitudes towards them have ranged from adoration (as was the case of the Pythagoreans) to confusion (as is echoed by Euler’s quote, even though few in the history of mathematics have been as acquainted with them as he was).

It is a similarly ambivalent statement that I wish to convey in the following lines: despite the apparent randomness and irregularity with which the primes are distributed among the integers, when one looks closer, they exhibit stunning order and regularity.

1 A First Look

Let us look at the first few primes in the sequence, 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, . . . ; we have included below (Figure 1) a representation of the numbers from 1 to 76800, each represented by a “pixel”: white if the number is composite and black if it is prime. Looking at the sequence, the primes seem to appear quite sporadically and without any discernable pattern

Figure 1:



other than the fact that they seem to thin out as we go farther in the integers. Even so, as Euclid has shown us long ago, they never disappear completely:

Theorem. *There are infinitely many primes.*

Proof. We shall prove this with Euclid's famous argument (which appears in Book IX of the *Elements*): assume there is only a finite number of primes p_1, p_2, \dots, p_n . The number $p_1 p_2 \cdots p_n + 1$ is not one of the p_i and clearly not divisible by any of them even though every number should either be a prime or be divisible by one: a contradiction. \square

The fact that the primes are infinite in number, however, does not tell the whole story. With a little extra work, we can obtain the following (relatively surprising) result, initially proved by Euler (however, our proof is an elegant argument due to Paul Erdős in [6][9]):

Theorem. *The series $\sum_{p \in \mathbb{P}} \frac{1}{p}$ diverges.*

Proof. Suppose that the series converges. Then, for a certain $k \in \mathbb{N}$, we have:

$$\sum_{i > k} \frac{1}{p_i} < \frac{1}{2}.$$

We define $\lambda_k(n)$ as the number of integers less than n divisible *only* by the first k primes. Taking an arbitrary such number m , we can express it as $m = p_1^{\omega_1} \cdots p_k^{\omega_k} \cdot s^2$, $\omega_i \in \{0, 1\}$ (i.e. a square multiplied by a square-free number). There are 2^k ways to choose the square-free part and clearly $s \leq \sqrt{n}$, so $\lambda_k(n) \leq 2^k \sqrt{n}$.

Now, the number of integers divisible by p_i less than n is $\lfloor \frac{n}{p_i} \rfloor$, so the number of integers less than n divisible by primes bigger than p_k (which we shall write $\Lambda_k(n)$) is bounded above as follows:

$$\Lambda_k(n) \leq \sum_{i > k} \left\lfloor \frac{n}{p_i} \right\rfloor \leq \sum_{i > k} \frac{n}{p_i} < \frac{n}{2},$$

the last inequality due to our initial remark.

However, by their definitions, $\lambda_k(n) + \Lambda_k(n) = n$ for all $n \in \mathbb{N}$ and so it is sufficient to find an n such that $\lambda_k(n) \leq \frac{n}{2}$ for a contradiction and, using the previous bound for $\lambda_k(n)$ (that is, $\lambda_k(n) \leq 2^k \sqrt{n}$), we see that $n = 2^{2k+2}$ works. \square

This is an interesting result as it shows us the primes are not as rare as we might think. Indeed, even though the sum diverges very slowly¹, it shows that the primes are more “dense” than the squares, for instance, since $\sum_n \frac{1}{n^2}$ converges to $\frac{\pi^2}{6}$.

Another natural question that we can try to answer in order to find some sort of structure in the primes is how they behave in terms of integer progressions, especially arithmetic progressions. One obvious result is that all the primes, except 2, are found in the arithmetic progression $(2k + 1)_{k \geq 1}$; moreover, these primes can only be of the form $4m + 1$ or $4m - 1$: can we say, as before, that all but a finite number of primes lie in one of the two categories? The answer is no:

Theorem. *There are an infinite number of primes in both arithmetic progressions $(4k - 1)_{k \geq 1}$ and $(4k + 1)_{k \geq 1}$.*

We shall prove these statements separately. The first is rather easy and requires a relatively small adjustment of Euclid’s argument:

Proof. $(4m - 1)$ Suppose the primes of the form $4m - 1$ are in finite number, namely q_1, q_2, \dots, q_n . Then, the number $Q := q_1 q_2 \cdots q_n - 1$ is not divisible by any of the q_i ; however, since $Q \equiv -1 \pmod{4}$ and any product of primes of the form $4m + 1$ is congruent to 1 mod 4, Q has to be divisible by a prime of the form $4m - 1$, a contradiction. \square

The second is slightly harder but can also be turned into a similar argument by using the well known fact that for $p \in \mathbb{P}$ there exists $x \in \mathbb{Z}_p$ such that $x^2 \equiv -1 \pmod{p}$ if and only if $p = 4m + 1$ for some m (-1 is a quadratic residue mod p). With this, we can prove the second part of the theorem:

Proof. $(4m + 1)$ Suppose, analogously, the primes of the form $4m + 1$ form a finite set and name them q_1, q_2, \dots, q_n . Now, the number $Q := (q_1 q_2 \cdots q_n - 1)^2 + 1$ isn’t divisible by any of the q_i . Let $p | Q$ be a prime, we thus have $Q \equiv 0 \pmod{p}$ so we have $(q_1 q_2 \cdots q_n)^2 \equiv -1 \pmod{p}$, which by the fact stated earlier implies $p = 4m + 1$ for some m , and we have our contradiction. \square

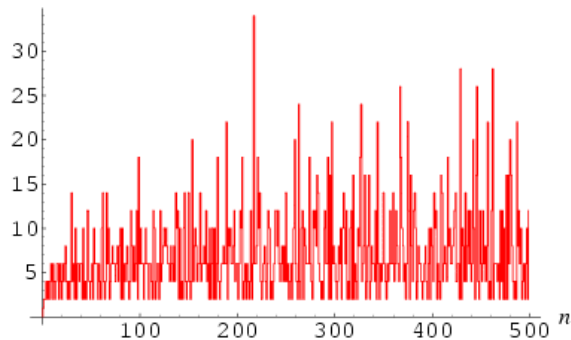
These are in fact special cases of a much more general (and notoriously difficult) theorem by Dirichlet which states that one can find an infinity of primes in any arithmetic progression of the form $(ak + b)_{k \geq 1}$ with a coprime to b . Little is known, however, for any other type of progression or similar sequence (even polynomial). Also noteworthy is a recent result of Green and Tao which proves the existence of (finite) arithmetic progressions of length k of primes for all k .

¹Summing over $p \in \mathbb{P}$, $\sum_{p < x} \frac{1}{p} = \log \log x + C + o(1)$ for a small constant C . Summing up to the largest currently known prime, $2^{32582657} - 1$ (a 9808358 digit long number), gives us an estimate of $\log(\log(2^{32582657} - 1)) + C \approx 17.1943 \dots$ which is very small considering how many terms we have summed.

2 Gaps in the Sequence of Primes

One interesting aspect to which we shall devote more attention is the size of the gaps between consecutive primes, $p_{n+1} - p_n$, as this is likely to be relevant to finding any sort of regularity within the sequence of primes. Looking at the

Figure 2: $p_{n+1} - p_n$ as a function of n



numbers for small n , the gaps seem distinctly irregular, although the average gap tends to get larger as we progress. It is indeed easy to see that the size of the gaps is unbounded: for any $n \in \mathbb{N}$, the $n - 1$ consecutive numbers $n! + 2, n! + 3, \dots, n! + n$ are all composite. There are, however, also upper bounds on this gap and we shall consider the most famous of these here: Bertrand's Postulate.

Theorem. *There is always a prime between n and $2n$.*

This of course implies that $p_{n+1} - p_n \leq p_n$. It was proven by Chebyshev (whom we shall soon meet again) in 1850, but our proof is, again, by the brilliant Paul Erdős [6], as exposed in [2].

Proof. The proof mainly relies on getting a clever estimate for the binomial coefficient $\binom{2n}{n} = \frac{(2n)!}{(n!)^2}$. Our first step is checking that Bertrand's Postulate holds for $n \leq 4000$: this is easily done by checking that the numbers 2, 3, 5, 7, 13, 24, 43, 83, 163, 317, 631, 1259, 2503, 4001 are all prime – each number in the list is less than two times the previous number (so that the intervals p to $2p$ overlap), which means that for any $n \leq 4000$, one of these will be between n and $2n$.

Next, we will need the inequality

$$\prod_{p \leq x} p \leq 4^{x-1} \quad \text{for all } x \geq 2,$$

with the product being over primes p . Note that if q is the largest prime smaller than x , it suffices to show that the inequality holds for $x = q := 2m + 1$ an odd prime (the inequality holds trivially for $x = 2$). We shall prove this by induction: having already verified the initial case $x = 2$, we may assume the inequality to hold for all integers up to $2m$ and prove that it holds for $2m + 1$:

$$\prod_{p \leq 2m+1} p = \left(\prod_{p \leq m+1} p \right) \cdot \left(\prod_{m+1 < p \leq 2m+1} p \right) = 4^m \cdot \binom{2m+1}{m} \leq 4^m \cdot 2^{2m} = 4^{2m}.$$

This is true since:

- $\prod_{p \leq m+1} p = 4^m$ holds by induction.
- $\prod_{m+1 < p \leq 2m+1} p = \binom{2m+1}{m}$ is true because $\binom{2m+1}{m} = \frac{(2m+1)!}{m!(m+1)!}$ and each of the primes in the product divides the numerator but not the denominator.
- $\binom{2m+1}{m} \leq 2^{2m}$ works as both $\binom{2m+1}{m} = \binom{2m+1}{m+1}$ appear in (and are thus smaller than) the sum $\sum_{k=0}^{2m+1} \binom{2m+1}{k} = 2^{2m+1}$.

Our next step requires an easy but useful result known as Legendre's theorem, which states that if the largest prime power of a prime p dividing $n!$ is p^α , then $\alpha = \sum_{k \geq 1} \lfloor \frac{n}{p^k} \rfloor$ (the proof follows easily from the fact that $\lfloor \frac{n}{p^k} \rfloor$ is the number of multiples of p^k less than n : we are counting every multiple of p once, every multiple of p^2 twice, etc.). Now, this result makes it clear that $\binom{2n}{n}$ contains the prime factor p exactly

$$\sum_{k \geq 1} \left(\left\lfloor \frac{2n}{p^k} \right\rfloor - 2 \left\lfloor \frac{n}{p^k} \right\rfloor \right)$$

times. Using the properties of the floor function, we have that

$$\left\lfloor \frac{2n}{p^k} \right\rfloor - 2 \left\lfloor \frac{n}{p^k} \right\rfloor < \frac{2n}{p^k} - 2 \left(\frac{n}{p^k} - 1 \right) = 2$$

and hence each term of the sum is either 1 or 0. Moreover, since each term with $p^k > 2n$ is 0, we get that $\binom{2n}{n}$ contains p at most r times if p^r is the largest prime power smaller than $2n$. As such, the largest prime power dividing $\binom{2n}{n}$ is not larger than $2n$, primes $p > \sqrt{n}$ appear at most once and, finally, primes p such that $\frac{2}{3}n < p \leq n$ do not divide $\binom{2n}{n}$ at all! In effect, $3p > 2n$ implies (for $n \geq 3$ and thus $p \geq 3$) that p and $2p$ are the only factors appearing in the numerator of $\binom{2n}{n} = \frac{(2n)!}{(n!)^2}$ while we have two appearances of p in the denominator.

We are almost done. Let us estimate $\binom{2n}{n}$ using what we have learned:

$$\frac{4^n}{2n} \leq \binom{2n}{n} \leq \left(\prod_{p \leq \sqrt{2n}} 2n \right) \cdot \left(\prod_{\sqrt{2n} \leq p \leq \frac{2}{3}n} p \right) \cdot \left(\prod_{n < p \leq 2n} p \right).$$

The first inequality comes from the fact that $\sum_{k \geq 1} \binom{2n}{k} = 2^{2n} = 4^n$ and the sum has n terms which are all bounded above by $\binom{2n}{n}$ – the middle binomial coefficient in the sequence. Using the fact that there are no more than $\sqrt{2n}$ primes smaller than $\sqrt{2n}$, our inequality becomes

$$4^n \leq (2n)^{\sqrt{2n}+1} \cdot \left(\prod_{\sqrt{2n} \leq p \leq \frac{2}{3}n} p \right) \cdot \left(\prod_{n < p \leq 2n} p \right).$$

To end the proof, assume that there is no prime $n < p \leq 2n$, rendering the last product in the previous inequality 1. Using the inequality we had established at the beginning of the proof for the remaining product, we now have:

$$4^n \leq (2n)^{\sqrt{2n+1}} \cdot 4^{\frac{2}{3}n},$$

i.e.

$$4^{\frac{n}{3}} \leq (2n)^{\sqrt{2n+1}},$$

which will be false for n large enough. Using the fact that $x + 1 < 2^x$ for $x \geq 2$, we have

$$2n = (\sqrt[6]{2n})^6 < (\lfloor \sqrt[6]{2n} \rfloor + 1)^6 < 2^{6\lfloor \sqrt[6]{2n} \rfloor} < 2^{6\sqrt[6]{2n}}$$

and, for $n \geq 50$ (and thus $2\sqrt{2n} > 18$), combining the last two inequalities gives us:

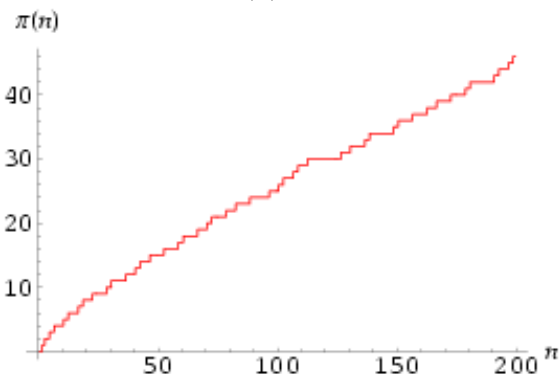
$$2^{2n} \leq (2n)^{3(\sqrt{2n}+1)} < 2^{\sqrt{2n}(18+18\sqrt{2n})} < 2^{20\sqrt[6]{2n}\sqrt{2n}} = 2^{20(2n)^{2/3}},$$

which implies $(2n)^{1/3} < 20$ and thus (surprise!) $n < 4000$; we are done. \square

3 The Prime Number Theorem

We come here to the crux of the problem, the single most important theorem regarding the distribution of prime numbers. We need to new function, $\pi(x)$, the so-called prime counting function, which denotes the number of primes under a given real number x (it is obviously a step function with an increment of 1 at every prime). Formally speaking, we can write it as $\pi(x) = \sum_{p \leq x} 1$ where the sum is indexed by primes p . This function quantifies the proportion of primes

Figure 3: $\pi(n)$ up to $n = 200$



up to a certain number and it's behaviour and rate of growth shows us roughly how the primes are distributed among the integers. Its study would thus provide a vast array of information about the primes and, justifiably, mathematicians during the 19th century began to try to find simpler estimates for the function. Two stood out in this search: one was Legendre, who conjectured in his *Theorie des Nombres* that $\pi(x) \sim \frac{x}{\log x - 1.08\dots}$ (where the constant seems to have been chosen by trial and error). The second, more importantly, was Gauss, whom, upon receiving tables of prime numbers when he was only around 15 years old,

had set out to find an estimate for $\pi(x)$ himself and conjectured (though, as so often with Gauss, did not publish) that $\pi(x) \sim \frac{x}{\log x}$ and also that $\pi(x) \sim \text{Li}(x)$, where

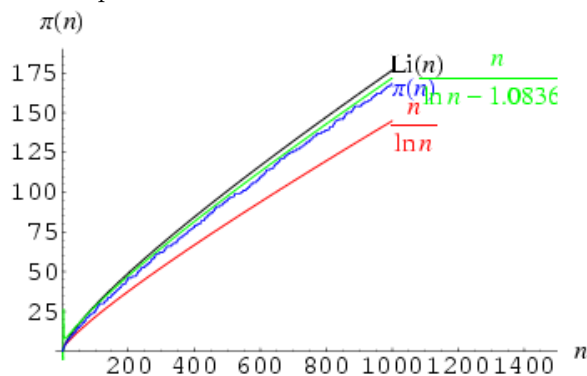
$$\text{Li}(x) := \int_2^x \frac{1}{\log t} dt.$$

These three estimates are in fact equivalent, although Gauss' $\text{Li}(x)$ is more accurate than the two others (and it turns out Legendre's constant is of no use). Indeed, integrating $\text{Li}(x)$ by parts gives us:

$$\begin{aligned} \text{Li}(x) &= \int_2^x \frac{1}{\log t} dt = \left[\frac{t}{\log t} \right]_2^x + \int_2^x \frac{1}{(\log t)^2} dt \\ &= \frac{x}{\log x} + \frac{x}{(\log x)^2} + \int_2^x \frac{1}{(\log t)^3} dt + O(1), \end{aligned}$$

which gives us $x/\log x$ as the first term of a series converging to $\text{Li}(x)$ that we can calculate as far as we like.

Figure 4: A comparison of the three estimates for small values of n .



These estimates later became known as the *Prime Number Theorem* (abbreviated *PNT*), which we shall restate here under its' $\text{Li}(x)$ form (this being the most interesting one):

Theorem. $\pi(x) \sim \text{Li}(x)$.

This beautiful theorem tells us that despite their apparent local randomness, the primes are in fact quite regularly distributed among the integers when one sees the “big picture”, as their growth is approximated by a very natural and predictable function.

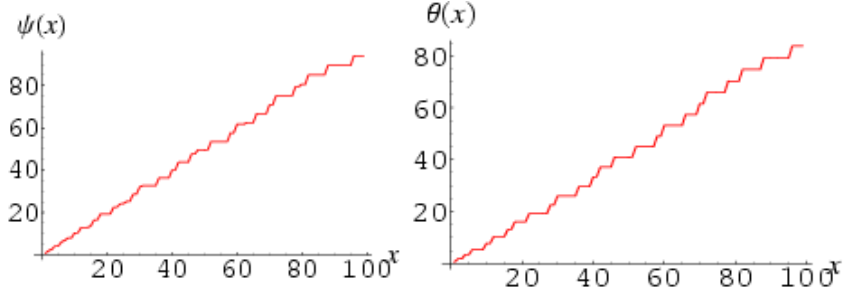
Let us pause for a moment and examine how Gauss came to conjecture this estimate and by the same occasion provide an interesting interpretation of it. In effect, what Gauss was doing was counting the proportion of primes in intervals of 1000, i.e. the “probability” of a number in one of those intervals of being prime. What he found was that the probability that a number around x is prime is roughly $1/\log x$ and summed all the intervals up by integrating (thus effectively using $1/\log x$ as a probability distribution for the primes with $\text{Li}(x)$ being its' probability density function) to find an estimate for the number of primes under a given real number.

This conjecture soon became a holy grail of sorts for mathematicians. The first to make significant progress towards a proof of it was the aforementioned Pafnuty Chebyshev in 1954. What Chebyshev proved was that if $\lim_{x \rightarrow \infty} \frac{\pi(x)}{x/\log x}$ exists, then it is equal to 1 but came short of proving the limit existed. He did, however, find good enough bounds $C_- < \lim_{x \rightarrow \infty} \pi(x)/\log x < C_+$ (he had $C_- \approx 0.922$ and $C_+ \approx 1.105$) which he used to prove Bertrand's Postulate and even proved that the relative error in approximating $\pi(x)$ by $\text{Li}(x)$ was smaller than 0.11 but couldn't prove that it tends to 0. He also introduced two new functions which would turn out to be very useful:

$$\psi(x) := \sum_{p^k \leq x} \log p \quad \text{and} \quad \theta(x) := \sum_{p \leq x} \log p,$$

where the first sum is taken over all prime powers $p^k \leq x$ and the second over all primes $p \leq x$. Note that we can write $\psi(x) = \text{lcm}\{1, 2, 3, 4, \dots, [x]\}$ as well. These are, in fact, weighed equivalents of the prime counting function,

Figure 5: The Chebyshev Functions



with estimates of their own: $\psi(x) \sim x$ and $\theta(x) \sim x$, as one could conjecture by looking at the graphs. The usefulness of these functions stems from the following theorem:

Theorem.
$$\lim_{x \rightarrow \infty} \frac{\theta(x)}{x} = \lim_{x \rightarrow \infty} \frac{\psi(x)}{x} = \lim_{x \rightarrow \infty} \frac{\pi(x)}{x/\log x}.$$

Proof. To begin with, note that if $p^k < x$ then k is the largest integer such that $k < \frac{\log x}{\log p}$, i.e. $k = \left\lfloor \frac{\log x}{\log p} \right\rfloor$. This implies the fact that

$$\psi(x) = \sum_{p \leq x} \left\lfloor \frac{\log x}{\log p} \right\rfloor \log p.$$

We thus have the inequality:

$$\theta(x) \leq \psi(x) = \sum_{p \leq x} \left\lfloor \frac{\log x}{\log p} \right\rfloor \log p \leq \sum_{p \leq x} \frac{\log x}{\log p} \log p = \log x \sum_{p \leq x} 1 = \log x \cdot \pi(x)$$

which, dividing by x , becomes:

$$\frac{\theta(x)}{x} \leq \frac{\psi(x)}{x} \leq \frac{\pi(x)}{x/\log x}.$$

Taking $x \rightarrow \infty$, we see that the three (possibly infinite) limits, which we shall name $\ell_\theta, \ell_\psi, \ell_\pi$, obey the same inequality.

Let $0 < \alpha < 1$ and $x > 1$, we then have

$$\theta(x) \geq \sum_{x^\alpha < p \leq x} \log p \geq \log(x^\alpha) \sum_{x^\alpha < p \leq x} 1 = \log(x^\alpha)(\pi(x) - \pi(x^\alpha)).$$

Since $\pi(x^\alpha) < x^\alpha$, we have $\theta(x) \geq \log(x^\alpha)(\pi(x) - x^\alpha)$ and thus

$$\frac{\theta(x)}{x} \geq \alpha \frac{\pi(x) \log x - x^\alpha \log x}{x} = \alpha \left(\frac{\pi(x)}{x/\log x} - \frac{\log x}{x^{1-\alpha}} \right).$$

As $x \rightarrow \infty$, $\frac{\log x}{x^{1-\alpha}} \rightarrow 0$, which shows that $\ell_\theta \geq \alpha \ell_\pi$ and since we can make α as close to 1 as we please, $\ell_\theta \geq \ell_\pi$. Combining this with the earlier $\ell_\theta \leq \ell_\psi \leq \ell_\pi$, we get $\ell_\theta = \ell_\psi = \ell_\pi$, which was our desired result. \square

In short, these functions provide other perspectives for solving the PNT since proving one limit amounts to proving them all. This is significant since, as we shall soon see, $\psi(x)$ is somewhat easier to handle than $\pi(x)$ and it is thus easier to prove the prime number theorem under its $\psi(x) \sim x$ form.

4 Riemann's Zeta Function

As it turns out, Chebyshev's real analytical methods were not enough to find a proof of the PNT. Jacques Hadamard, one of the two mathematicians who did later prove the PNT, is quoted as having said "*The shortest path between two truths in the real domain sometimes passes through the complex domain.*" This turns out to be very true in the case of the PNT², but it took the genius of Bernhard Riemann to make the necessary connection. In an epoch-making memoir that Riemann presented to the Berlin Academy of Sciences in 1859 (his *only* paper on number theory), Riemann introduced the ideas that would spark the whole field of analytic number theory. His crucial idea was to take a real function known as $\zeta(s)$ (which is closely related to the prime numbers) extend it to the whole complex plane (minus one point) and unleash the full power of complex analysis on the study of primes. The Riemann zeta function, which was in fact known to Euler (who proved a lot of its' fundamental properties) is one of the most beautiful and mindbendingly difficult objects in all of mathematics. Its definition, however, is rather simple:

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

The series is defined and converges (by the integral test) for all real $s > 1$; in fact, letting $s = \sigma + it$ be a complex number, we have $|n^{-s}| = |e^{-(\sigma+it)\log n}| = |e^{-\sigma \log n}| = n^{-\sigma}$, which means the series converges absolutely for all complex numbers with $\Re[s] > 1$. Using the Weierstrass M-test, we can see that it actually converges uniformly in the half-plane $\Re[s] > 1$ and so defines an analytic function there.

²Proofs that use only real analysis have been found – the first one by Erdős and Selberg in 1949. These so-called *elementary* proofs are ironically much harder than the ones using complex analysis.

The fundamental property linking the zeta function to the primes was found by Euler and is therefore known as the Euler product:

Theorem.
$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \in \mathbb{P}} \left(1 - \frac{1}{p^s}\right)^{-1}$$

Proof. Consider the following equality

$$(1 - 2^{-s})\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} - \sum_{n=1}^{\infty} \frac{1}{(2n)^s} = \sum_{n \text{ odd}} \frac{1}{n^s} = 1 + \sum_{p|n \Rightarrow p > 2} \frac{1}{n^s},$$

where the last sum is over all n with all prime factors greater than 2. Repeating this argument inductively for the first k primes gives us

$$\prod_{i=1}^k (1 - p_i^{-s}) \zeta(s) = 1 + \sum_{p|n \Rightarrow p > p_k} \frac{1}{n^s},$$

where the last sum is again over all n with all prime factors greater than p_k : this constitutes a subsum of the tail of the series defining $\zeta(s)$ and must hence tend to 0 if we make $k \rightarrow \infty$, leaving us with:

$$\prod_{p \in \mathbb{P}} (1 - p^{-s}) \zeta(s) = 1;$$

all that is left to do is divide by the product on both sides and we are done. \square

Riemann's idea was to extend the zeta function to a holomorphic function on $\mathbb{C} \setminus \{1\}$ by analytic continuation. In order to do this, he started with the following identity (which was also first proven by Euler):

Lemma.
$$\zeta(s)\Gamma(s) = \int_0^{\infty} \frac{u^{s-1}}{e^u - 1} du, \text{ where } \Gamma(s) \text{ is the well known special function defined by } \Gamma(s) := \int_0^{\infty} t^{s-1} e^{-t} dt.$$

Proof. Start with the integral definition of gamma and make the substitution $t = nu$ to get:

$$\Gamma(s) = \int_0^{\infty} (nu)^{s-1} e^{-nu} n du = n^s \int_0^{\infty} u^{s-1} e^{-nu} du.$$

Hence, we have

$$\frac{1}{n^s} = \frac{1}{\Gamma(s)} \int_0^{\infty} u^{s-1} e^{-nu} du.$$

and, summing over all n ,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{\Gamma(s)} \int_0^{\infty} u^{s-1} e^{-nu} du = \frac{1}{\Gamma(s)} \int_0^{\infty} u^{s-1} \sum_{n=1}^{\infty} e^{-nu} du.$$

Summing the geometric series, we are left with:

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^{\infty} u^{s-1} \frac{e^{-u}}{1 - e^{-u}} du = \frac{1}{\Gamma(s)} \int_0^{\infty} \frac{u^{s-1}}{e^u - 1} du.$$

\square

Riemann next considers the contour integral

$$\mathcal{J}(s) := \frac{1}{2\pi i} \int_C \frac{(-u)^{s-1}}{e^u - 1} du$$

on a contour which goes from $R + i\infty$ just above and parallel the real axis, circles the origin on a semi-circle of radius R and goes back to $-R + i\infty$ just below and parallel to the real axis. Computing $\mathcal{J}(s)$ and letting $R \rightarrow 0$, he obtains the aforementioned identity in the form

$$\pi \mathcal{J}(s) = \sin(\pi s) \int_0^\infty \frac{u^{s-1}}{e^u - 1} du = \sin(\pi s) \zeta(s) \Gamma(s).$$

Using $\Gamma(s)$'s complement formula $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$ in the last formula, he finally obtains:

$$\zeta(s) = \Gamma(1-s) \mathcal{J}(s) = \frac{\Gamma(1-s)}{2\pi i} \int_C \frac{(-u)^{s-1}}{e^u - 1} du,$$

a formula for $\zeta(s)$ which is valid for all $s \in \mathbb{C} \setminus \{1\}$. Using a similar technique (the same integral on a different contour), he also gets the following useful functional equation for $\zeta(s)$:

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

which can also be written in the somewhat more elegant form $F(s) = F(1-s)$ for $F(s) := \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s)$.

How is all this related to the PNT, you may ask? Riemann goes on to define, in the next part of his paper, the function³ $J(x) := \sum_{n \geq 1} \frac{1}{n} \pi\left(x^{\frac{1}{n}}\right)$, another weighed prime counting function. This function is related to the original prime counting function $\pi(x)$ by $\pi(x) = \sum_{n \geq 1} \frac{\mu(n)}{n} J\left(x^{\frac{1}{n}}\right)$, by using a common number theoretic technique known as Möbius inversion ($\mu(n)$ is the Möbius function, which is 0 if n is divisible by a square and $(-1)^k$ if n is a product of k distinct prime factors). By taking the logarithm of the Euler Product and applying a technique known as *Fourier Inversion*, Riemann obtains (without rigorous proof) an exact formula for $J(x)$ (and thus $\pi(x)$):

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

where the second sum is over the *non-trivial zeros* of the zeta function⁴. This formula can itself be used to solve the Prime Number Theorem, which turns out to be equivalent to the statement that the non-trivial zeroes of the zeta function all have real part less than 1, but we shall instead use an equivalent formula for Chebyshev's ψ function, proved by Von Mangoldt some 40 years after Riemann's paper.

³It is easy to check that all this function does is add $1/n$ for each n^{th} power of a prime it passes; moreover it is a finite sum as the terms are all 0 from a certain point onwards as $\pi(x) = 0$ for $x < 2$.

⁴The zeta function has two types of roots: the trivial zeros occurring at each even negative integer – they appear because of the cosine term in the functional equation – and the non-trivial zeros which are all located in the *critical strip* $0 \leq \Re[s] \leq 1$.

Von Mangoldt starts similarly by taking the logarithm of the Euler product:

$$\log \zeta(s) = \log \prod_{p \in \mathbb{P}} \left(1 - \frac{1}{p^s}\right)^{-1} = - \sum_{p \in \mathbb{P}} \log(1 - p^{-s}),$$

which he then differentiates, obtaining

$$\frac{\zeta'(s)}{\zeta(s)} = - \sum_{p \in \mathbb{P}} \frac{p^{-s} \log p}{1 - p^{-s}} = - \sum_{p \in \mathbb{P}} \sum_{k=1}^{\infty} \frac{\log p}{p^{ks}}.$$

Using Fourier inversion, he isolates the logarithmic part of the sum and obtains

$$\psi(x) = \sum_{p^k \leq x} \log p = - \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} ds$$

for a suitable closed contour \mathcal{C} containing all the poles of the integrand. Evaluating the contour integral by residues finally yields four types of residues:

- The residue at 0 (pole of the x^s/s term) is $\frac{\zeta'(0)}{\zeta(0)} = \log(2\pi)$.
- The residue at 1 (as 1 is a pole of $\zeta'(s)$) is $-x$.
- The residues at the trivial zeroes of zeta (poles of $(\zeta(s))^{-1}$) are $-(2k)^{-1}x^{-2k}$ for integer $k \geq 1$. Summing them all up gives the Taylor series for $\frac{1}{2} \log(1 - x^2)$.
- The residues at the non-trivial zeroes ρ of zeta (poles of $(\zeta(s))^{-1}$) give residues $\frac{x^\rho}{\rho}$.

Summing them all up and multiplying by $2\pi i$, we finally get the explicit formula Von Mangoldt proved:

$$\psi(x) = x - \sum_{\rho} \frac{x^\rho}{\rho} - \log(2\pi) - \frac{1}{2} \log(1 - x^2).$$

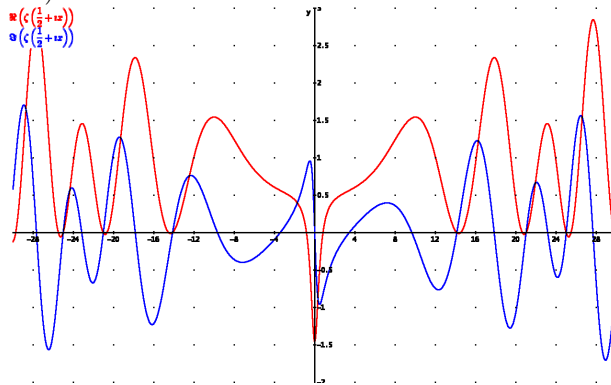
Recall that we need $\psi(x) \sim x$ for the PNT to be true and indeed, since the last two terms are negligible compared to the first two, if $0 < \Re[\rho] < 1$ then the x term will be the dominating one and the PNT will be true (we have $|x^\rho| = x^{\Re[\rho]}$). A short time after Von Mangoldt provided rigorous proofs of the results which Riemann didn't prove in his paper (including the explicit formula), Jacques Hadamard and Charles de la Vallée Poussin independantly proved this result, settling the Prime Number Theorem at last.

5 The Riemann Hypothesis

The story is nevertheless not over yet, as a more precise determination of the location of the zeta function's zeros would provide a better error term for the PNT. The celebrated Riemann Hypothesis (conjectured by Riemann in his 1859 paper), possibly the greatest unsolved problem in mathematics today, states that all the zeros of the Riemann zeta function lie on the line $\Re[s] = \frac{1}{2}$, right in the middle of the critical strip. This effectively reduces the real part of the

zeros as much as is allowed (which in turn reduces the order of the ρ sums in the explicit formulas for $\psi(x)$ and $\pi(x)$, bringing them closer to their asymptotic equivalents x and $\text{Li}(x)$), since the non-trivial roots obey two symmetries in the critical strip: if ρ is a non-trivial zero, then so will $\bar{\rho}$, $1 - \rho$ (which is the point symmetrical to ρ by the critical line $\Re[s] = \frac{1}{2}$) and $\overline{1 - \rho}$; as such, the existence of a root with real part less than $\frac{1}{2}$ guarantees the existence of one with real part more than $\frac{1}{2}$.

Figure 6: Plots of $\Re[\zeta(\frac{1}{2} + it)]$ and $\Im[\zeta(\frac{1}{2} + it)]$ as functions of t (note the non-trivial zeros).



Accordingly, Helge von Koch proved in 1901 that if the Riemann Hypothesis holds, then we have $|\text{Li}(x) - \pi(x)| = O(\sqrt{x} \log x)$, which is known to be more or less the best possible error term. This would thus imply amazing regularity from the seemingly random sequence of primes and has justifiably attracted many mathematicians despite its immense difficulty. Just as astonishing is the number of consequences that the hypothesis has all over mathematics or the number of equivalent statements which seem perfectly elementary, ranging from the Farey Sequence to the number of maximal order elements in the symmetric group. It has for example recently been shown that the Riemann Hypothesis is equivalent to $\sigma(n) \leq H_n + \log(H_n)e^{H_n}$ for all n , where H_n is the n^{th} term of the harmonic series and $\sigma(n)$ is the sum of the divisors of n . There are in fact many statements about the growth of certain arithmetic functions equivalent to or stronger than the Riemann Hypothesis. A famous one (which the Dutch mathematician Stieltjes had erroneously thought he had solved) concerns the Möbius $\mu(n)$ function which we have encountered earlier:

Theorem. Let the Mertens function $M(x)$ be defined by $\sum_{n=1}^{\lfloor x \rfloor} \mu(n)$. The fact that $M(x) = O(\sqrt{x})$ implies the Riemann Hypothesis.

Proof. We start with the identity

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} \quad \text{for } \Re[s] > 1.$$

This is a well-known expression deriving from Möbius inversion as well (it is quite easy to prove using the theory of Dirichlet convolutions of arithmetic

functions but we shall simply assume it here). Setting $M(0) := 0$, we have

$$\begin{aligned}
\frac{1}{\zeta(s)} &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \sum_{n=1}^{\infty} \frac{M(n) - M(n-1)}{n^s} \\
&= \sum_{n=1}^{\infty} \frac{M(n)}{n^s} - \sum_{n=1}^{\infty} \frac{M(n-1)}{n^s} = \sum_{n=1}^{\infty} \frac{M(n)}{n^s} - \sum_{n=1}^{\infty} \frac{M(n)}{(n+1)^s} \\
&= \sum_{n=1}^{\infty} M(n) \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right) = \sum_{n=1}^{\infty} M(n) \int_r^{r+1} \frac{s}{x^{s+1}} dx \\
&= s \sum_{n=1}^{\infty} \int_r^{r+1} \frac{M(x)}{x^{s+1}} dx = s \int_1^{\infty} \frac{M(x)}{x^{s+1}} dx,
\end{aligned}$$

since $M(x) = M(n)$ on each interval $[n, n+1)$. Suppose now that $M(x) = O(\sqrt{x})$, i.e. $|M(x)| < C\sqrt{x}$ for some positive C . If so, we have

$$\left| \frac{M(x)}{x^{s+1}} \right| < \left| \frac{C\sqrt{x}}{x^{s+1}} \right| = \frac{C}{\sqrt{x}} \left| \frac{1}{x^s} \right| = \frac{C}{\sqrt{x}} \frac{1}{x^{\Re[s]}} = \frac{C}{x^{\Re[s] + \frac{1}{2}}}$$

As such, that last integral will converge if $\Re[s] + \frac{1}{2} > 1$ meaning that $\Re[s] > \frac{1}{2}$ which would give an analytic continuation of $\frac{1}{\zeta(s)}$ in the half plane $\Re[s] > \frac{1}{2}$ which implies the fact that $\frac{1}{\zeta(s)}$ has no poles for $\Re[s] > \frac{1}{2}$ and thus $\zeta(s)$ has no zeros in the region. By the symmetry of zeta's roots, there would be no zeros with $\Re[s] < \frac{1}{2}$ either: they must all be on the critical line! \square

So far, however, we are not even close to a solution of the hypothesis, although some progress has been made. For example, Hardy has proved (not without some difficulty) that an infinite number of the non-trivial zeros do lie on the critical line and Selberg later strengthened this to the fact that the zeros on the critical line have positive density. Despite the sheer amount of calculated evidence in favor of it or the number of results proven assuming its truth, no one is certain that the hypothesis should be true and some of the main actors in its history, including Selberg and Littlewood, have believed it to be false. Nevertheless, most mathematicians remain optimistic that the Riemann hypothesis is true and shall be proven. As much as has already been said on the distribution of the primes, there is still much more to be said. We would indeed prefer to believe David Hilbert rather than the great Euler on the issue: *Wir müssen wissen. Wir werden wissen.*

References

- [1] T. APOSTOL: *An Introduction to Analytic Number Theory*, Springer-Verlag (1976).
- [2] M. AIGNER & G. M. ZIEGLER: *Proofs from THE BOOK*, Third edition, Springer-Verlag (2004).
- [3] H. COHEN: *Number Theory - Volume II: Analytic and Modern Tools*, Springer-Verlag (2007).

- [4] H. DAVENPORT: *Multiplicative Number Theory*, Third edition, Springer-Verlag (2000).
- [5] H. M. EDWARDS: *Riemann's Zeta Function*, Academic Press (1974).
- [6] P. ERDŐS: *Über die Reihe $\sum \frac{1}{p}$* , *Mathematica* (1938).
- [7] P. ERDŐS: *Beweis eines Satzes von Tschebyschef*, *Acta Sci. Math.* (1932).
- [8] G. EVERETT & T. WARD: *An Introduction to Number Theory*, Springer-Verlag (2005).
- [9] G. H. HARDY & E. M. WRIGHT: *An Introduction to the Theory of Numbers*, Fifth edition, Oxford University Press (1979).
- [10] J. HAVIL: *Gamma: Exploring Euler's Constant*, Princeton University Press (2003).
- [11] G. A. JONES & J. M. JONES: *Elementary Number Theory*, Springer-Verlag (1998).
- [12] J. STOPPLE: *A Primer of Analytic Number Theory*, Cambridge University Press (2003).

Notation

- \mathbb{P} — the set of all primes.
- p_n — the n^{th} prime (p without an index shall also always denote a prime).
- $\lfloor x \rfloor$ — the floor function, i.e. greatest integer smaller than x .
- $\binom{n}{k} := \frac{n!}{k!(n-k)!}$ — a binomial coefficient.
- \mathbb{Z}_p — the finite field with p elements.
- $f(x) = O(g(x))$ — “big-oh” notation: there exists $C > 0$ such that $|f(x)| < C|g(x)|$ for large enough x .
- $f(x) = o(g(x))$ — “small-oh” notation: $\lim_{x \rightarrow \infty} f(x)/g(x) = 0$.
- $f(x) \sim g(x)$ — asymptotic equivalence: $\lim_{x \rightarrow \infty} f(x)/g(x) = 1$.
- $\Re[s]$ — the real part of $s \in \mathbb{C}$.
- $\Im[s]$ — the imaginary part of $s \in \mathbb{C}$.
- ρ — a non-trivial zero of the Riemann zeta function.
- $\pi(x) := \sum_{p \leq x} 1$ — the prime counting function.
- $\theta(x) := \sum_{p \leq x} \log p$ — Chebyshev's θ function.
- $\psi(x) := \sum_{p^k \leq x} \log p$ — Chebyshev's ψ function.
- $\zeta(s) := \sum_{n \geq 1} n^{-s}$ — Riemann's zeta function.

- $\text{Li}(x) := \int_2^x \frac{dt}{\log t}$ — the Logarithmic Integral function.
- $\mu(n) := \begin{cases} 0 & \text{if } p^2 | n \text{ for some } p \\ (-1)^k & \text{if } n \text{ is a product of } k \text{ distinct primes} \end{cases}$ — Möbius' function.
- $M(x) := \sum_{n=1}^{\lfloor x \rfloor} \mu(n)$ — Mertens' function.
- $H_n := \sum_{k=1}^n \frac{1}{k}$ — the n^{th} harmonic number.
- $\sigma(n) := \sum_{d|n} d$ — the sum of all divisors of n .