Prime Numbers

A Nesin Mathematics Village course

David Pierce

August 20–26, 2018
Last edited August 29, 2018

Matematik Bölümü
Mimar Sinan Güzel Sanatlar Üniversitesi
dpierce@msgsu.edu.tr
mat.msgsu.edu.tr/~dpierce/
polytropy.com

Preface

In six days on the Prime Number Theorem, here is what I tried to cover. Students knew the needed real analysis, but not the complex analysis. I tried to give this. The course might then be considered as a practical introduction to that subject.

Monday: Chapter 1 and its Theorems 1 and 2, that $\pi(x)$ grows beyond all bounds and, more precisely, exceeds $\log\log x$. Begin Chapter 2 with, hurriedly, Theorem 3, giving the important bound,

$$\vartheta(x) \leqslant 2x \log x.$$

Tuesday: Review the proof of Theorem 3. Prove Theorem 4, giving the prime factorization of n-factorial. Start proving Theorem 5, Bertrand's Postulate,

$$\pi(2n) - \pi(n) > 0$$

for all n in $\mathbb{N}$; but take a break at (2.19). Finish after the break.

Wednesday: In Chapter 3, prove Theorem 6, Chebyshev's Theorem,

$$\pi(x) \log x \asymp x.$$

First reach (3.5), $\pi(x) \log x = \mathrm{O}(x)$, before introducing ψ. Take a break after reaching (3.10), without having given (3.7) or (3.9) or discussed strategy. After the break, a student suggests the line of argument in (3.8). Finish Chapter 3 with time to spare. Skip to Chapter 5 by defining the Riemann zeta function. One student (of six now) knows $\zeta(s)$ is defined when the real part σ of s is greater than 1. State Theorem 18, the product formula for ζ.

Friday: Prove Theorem 18 and then Theorem 24, that

$$\pi(x) \log x \sim \vartheta(x).$$

From Chapter 4, define holomorphic and analytic functions, prove Theorem 7 (the Cauchy–Riemann Equations), and state Theorem 16 (holomorphic functions are analytic). Prove Theorem 19, that $\zeta(s) - (s-1)^{-1}$ extends holomorphically to $\sigma > 0$. Summarize the steps leading to Theorem 25, the Prime Number Theorem.

Saturday: Prove Theorem 23, that $f(x) \sim x$ when f is increasing and $\int_1^\infty \bigl(f(t) - t\bigr)\,\mathrm{d}\,t/t^2$ converges. Derive Theorem 22, that $\int_1^\infty (\vartheta(x) - x)\,\mathrm{d}\,x/x^2$ converges, from Theorem 21, Newman's "analytic theorem." With some handwaving about residues (Theorem 17), and without proving $\zeta(s) \neq 0$ when $\sigma = 1$, prove Theorem 20, that $\Phi(s) - (s-1)^{-1}$ extends holomorphically to $\sigma \geqslant 1$. End with the simplest case of Theorem 14, that $\int_\gamma \mathrm{d}\,z/(z-a) = 2\mathrm{i}\,\pi$.

Sunday: State what remains: the assertion about zeros of $\zeta(s)$, and Theorem 21. A crash course in complex analysis (Chapter 4), not stating Theorems 8 and 9, but spelling out the rest, up to Theorems 15 (Cauchy's Integral Formula) and 16. I use mostly English today, since the three students present at the beginning (more come later) are comfortable with it. One comments at the end how quickly we have gone.

Contents

List of Figures

1 Introduction

These notes concern the **counting numbers,** or *positive* integers; the set of them is $\mathbb{N}$. The **natural numbers** are the *non-negative* integers; the set of them is ω. We shall have reason to consider also real and complex numbers, composing the sets $\mathbb{R}$ and $\mathbb{C}$ respectively. Some real analysis is discussed later in this Introduction. In Chapter 4 is the complex analysis needed for our ultimate goal.

That goal is to prove the Prime Number Theorem (Theorem 25, page 56). We shall fill in the details of the proof in Don Zagier's four-page article, "Newman's Short Proof of the Prime Number Theorem" [13]. Newman's own treatment is in his own four-page article, "Simple analytic proof of the prime number theorem" [10].

Before the PNT, we shall prove Bertrand's Postulate and Chebyshev's Theorem, all discussed in this Introduction. A general reference is Hardy and Wright, *Introduction to the Theory of Numbers* [6, p. 12].

In the present document, the real numbers e and π, and the complex number i, related by the equation

$$\mathrm{e}^{\mathrm{i}\,\pi} + 1 = 0,$$

will be thus written, in upright, roman shape. This in principle means that the slanted, italic letters e, i, and π are available to serve as variables. We shall give π also another meaning as a function, which will be the object of our study.

Used mathematically, *the italic letter p will always denote a prime number.* This is a prime number in the original sense of Euclid, thus an element of $\mathbb{N}$. We are going to use the roman or upright letter π not only for the ratio of the circumference of a circle to its diameter, but also for the function of a real argument giving the number of primes that are no greater than the argument. Thus

$$\pi(x) = \sum_{p \leqslant x} 1 = |\{p \colon p \leqslant x\}| . \tag{1.1}$$

Theorem 1 (Euclid [5, IX.20]). $\lim_{x \to \infty} \pi(x) = \infty$.

Proof. The prime numbers form a list $(p_0, p_1, p_2, \dots)$, where

$$p_0 < p_1 < p_2 < \cdots$$

We show that the list never terminates; that is, p_n exists for each n in ω. We use strong induction. Suppose p_k exists when $k < n$. Then

$$\prod_{k<n} p_k + 1 \geqslant 2.$$

(Note that this is true, even if $n = 0$; the empty product has the value 1.) If $j < n$, then

$$\prod_{k<n} p_k + 1 \equiv 1 \pmod{p_j},$$

so p_j is not a factor of $\prod_{k<n} p_k + 1$. However, thus sum must *have* a prime factor. The least such is p_n. Therefore p_n does exist for all natural numbers n. □

One can say a little more, as Hardy and Wright do [6, p. 12]. In the following, as always, the base of logarithms is e, unless otherwise specified.

Theorem 2. *For sufficiently large real numbers x,*

$$\pi(x) \geqslant \log\log x. \tag{1.2}$$

Proof. As in the proof of Theorem 1, and *by* the theorem, the primes form the increasing sequence $(p_n : n \in \omega)$. The proof also shows

$$\bigwedge_{j<n} p_j \leqslant 2^{2^j} \implies p_n \leqslant \prod_{j<n} p_j + 1 \leqslant \prod_{j<n} 2^{2^j} + 1$$
$$= 2^{\sum_{j<n} 2^j} + 1 = 2^{2^n - 1} + 1 \leqslant 2^{2^n},$$

and so, by strong induction, for all n in ω,

$$p_n \leqslant 2^{2^n}, \qquad\qquad n \leqslant \pi(2^{2^n}).$$

If $x \geqslant 2$, then for some n in ω,

$$2^{2^{n+1}} > x \geqslant 2^{2^n},$$

and therefore

$$\begin{aligned}\pi(x) \geqslant \pi(2^{2^n}) \geqslant n &= \log_2\log_2(2^{2^{n+1}}) - 1 \\ &> \log_2\log_2 x - 1 = \frac{\log\log x - \log\log 2 - \log 2}{\log 2} \\ &= \log\log x + \frac{(1 - \log 2)\log\log x - \log\log 2 - \log 2}{\log 2}.\end{aligned}$$

Since $1 > \log 2$, we have the desired result. □

Specifically, (1.2) holds if

$$\log\log x \geqslant \frac{\log\log 2 + \log 2}{1 - \log 2}.$$

To handle lesser x by hand, we may note

$$\mathrm{e}^3 > \left(\frac{27}{10}\right)^3 = \frac{19683}{1000} > 16 = 2^4,$$

so that

$$n \geqslant 4 \implies \mathrm{e}^{\mathrm{e}^{n-1}} > \mathrm{e}^{2^n} > 2^{2^n},$$

and therefore

$$\mathrm{e}^{\mathrm{e}^n} \geqslant x > \mathrm{e}^{\mathrm{e}^{n-1}} \;\&\; n \geqslant 4$$
$$\implies \pi(x) \geqslant \pi(\mathrm{e}^{\mathrm{e}^{n-1}}) \geqslant \pi(2^{2^n}) \geqslant n \geqslant \log\log x.$$

This proves (1.2) when $x > \mathrm{e}^{\mathrm{e}^3}$. Going further, we have

$$\mathrm{e}^{\mathrm{e}^3} \geqslant x \geqslant 5 \implies \pi(x) \geqslant 3 \geqslant \log\log x.$$

Moreover,

$$\mathrm{e}^{\mathrm{e}^1} > \left(\frac{27}{10}\right)^2 = \frac{729}{100} > 5,$$

so

$$5 > x \geqslant 2 \implies \pi(x) \geqslant 1 \geqslant \log\log x.$$

Finally, $\mathrm{e}^{\mathrm{e}^0} > 2$, so $2 > x \geqslant 1 \implies \pi(x) = 0 \geqslant \log\log x$.

Our first big result will be **Bertrand's Postulate** (Theorem 5, page 17), that for every counting number n,

$$\pi(2n) - \pi(n) > 0.$$

I do not know why the result is called a postulate, but apparently Bertrand verified it in 1845 when n was less than three million. My source here is Dickson's *History of the Theory of Numbers* [3, p. 435]. This was first published in 1919, when some results that we shall consider were fresh or not yet known. For details, Dickson defers to Landau's *Handbuch [der Lehre von der] Verteilung der Primzahlen, I,* of 1919 (which I haven't got).

Given two positive-valued functions f and g, defined on either $\mathbb{N}$ or an interval (a, ∞), if the quotient $f(x)/g(x)$ is bounded above, we shall use the standard notation (due to Landau, I believe)

$$f = \mathrm{O}(g).$$

One must recognize that $\mathrm{O}(g)$ means nothing in isolation. In case both $f = \mathrm{O}(g)$ and $g = \mathrm{O}(f)$, we shall write

$$f \asymp g, \tag{1.3}$$

using the notation of Hardy and Wright [6, p. 7]. In this case,

- f **is of the same order of magnitude as** g, in the terminology of Hardy and Wright;
- f and g **go to infinity at similar rates,** for Mazur and Stein in *Prime Numbers and the Riemann Hypothesis* [9, ch. 17, p. 56].

The latter terminology is potentially misleading, since a constant function does not "go to infinity," although, as we shall presently confirm,

$$s > 1 \implies \int_1^x \frac{\mathrm{d}\,t}{t^s} \asymp 1.$$

Here $\int_1^x \mathrm{d}\,t/t^s$ denotes the function that takes this value at x. For arbitrary positive s different from 1,

$$\int_1^x \frac{\mathrm{d}\,t}{t^s} = \frac{1}{(1-s)t^{s-1}}\bigg|_1^x = \frac{1}{1-s}\left(\frac{1}{x^{s-1}} - 1\right).$$

Thus, as we shall want to know for Theorem 19 on page 46,

$$s > 1 \implies \lim_{x\to\infty} \int_1^x \frac{\mathrm{d}\,t}{t^s} = \frac{1}{s-1}. \tag{1.4}$$

Similarly,

$$0 < s < 1 \implies \int_1^x \frac{\mathrm{d}\,t}{t^s} \asymp x^{1-s},$$

because

$$0 < s < 1 \implies \lim_{x\to\infty} \frac{\int_1^x \mathrm{d}\,t/t^s}{x^{1-s}} = \frac{1}{1-s}. \qquad (1.5)$$

More difficult than Bertrand's Postulate will be **Chebyshev's Theorem** (Theorem 6, page 23),

$$\pi(x) \asymp \frac{x}{\log x},$$

established around 1850. In such an expression, again, we understand the symbolism $f(x)$ as another name for f. In Chebyshev's Theorem, the base of logarithms is irrelevant. The statement of the Theorem does not imply Bertrand's Postulate, but the proofs will use lemmas in common.

In addition to (1.3), we define the expression

$$f \prec g,$$

which means $f(x)/g(x)$ tends to 0 as x grows large. By an easy application of L'Hôpital's Rule,

$$s > 0 \implies \log x \prec x^s. \qquad (1.6)$$

We shall use this in proving Chebyshev's Theorem.

Under the condition not only that $f(x)/g(x)$ stays within positive bounds, but that it has unity as a limit, or equivalently

$$\lim_{x\to\infty} \frac{f(x)-g(x)}{g(x)} = 0,$$

we shall write

$$f \sim g;$$

this means

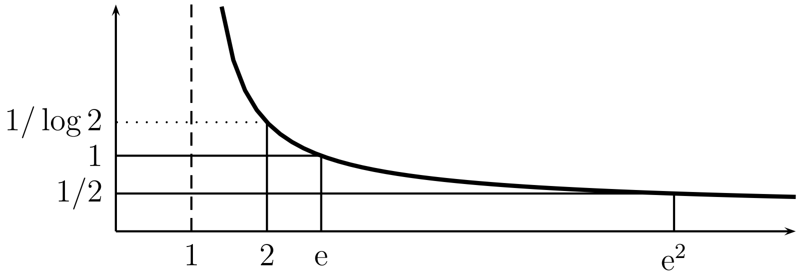

Figure 1.1: $y \log x = 1$

- f **is asymptotic to** g [6, p. 8];
- f and g **go to infinity at the same rate** [9, ch. 17, p. 56].

For example, by (1.4) and (1.5),

$$s > 1 \implies \int_1^x \frac{\mathrm{d}\,t}{t^s} \sim \frac{1}{s-1},$$
$$0 < s < 1 \implies \int_1^x \frac{\mathrm{d}\,t}{t^s} \sim \frac{x^{1-s}}{1-s}.$$

These are easy results, because we can calculate the integrals. With more work, when we define

$$\mathrm{Li}(x) = \int_2^x \frac{\mathrm{d}\,t}{\log t}, \tag{1.7}$$

we obtain

$$\mathrm{Li}(x) \sim \frac{x}{\log x} \tag{1.8}$$

by L'Hôpital's Rule. See Figure 1.1. First we compute

$$\frac{\mathrm{Li}(x) - x/\log x}{x/\log x} = \frac{\log x \cdot \mathrm{Li}(x) - x}{x}.$$

If the latter is f/g, then, since g grows without bound, we compute

$$\frac{f'}{g'} = \frac{1}{x}\,\mathrm{Li}(x) = \frac{\mathrm{Li}(x)}{x},$$

and if *this* is f/g, then f'/g' is $1/\log x$, which tends to 0 as x grows without bound.

The **Prime Number Theorem** (Theorem 25, page 56) is

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

established in 1896 by Hadamard and de la Vallée-Poussin independently. The Theorem implies Bertrand's Conjecture.

The *Riemann Hypothesis* is a refinement of the Prime Number Theorem.

2 Bertrand's Postulate

Our proof of Bertrand's Postulate, Theorem 5 below, will be based on that of Hardy and Wright [6, §22.3], who attribute it to Paul Erdős [4]. Erdős's paper appeared in 1932, and Erdős was born in 1913. An earlier proof, from 1919, is due to Srinivasa Ramanujan [11]; this proof is very short (2 pages), but makes use of the Gamma function and Stirling's approximation to it. Hardy and Wright attribute the earliest proof of Bertrand's Postulate to Tchebyshef (Chebyshev) in 1850.

We shall make use of the auxiliary function ϑ given by

$$\vartheta(x) = \sum_{p \leqslant x} \log p. \tag{2.1}$$

Then $\vartheta(x) = \log \prod_{p \leqslant x} p$, so that, for example,

$$2 < 3 < 5 \;\; \& \;\; \vartheta(5) = \log(2 \cdot 3 \cdot 5).$$

We establish the following bound [6, Thm 441, p. 341].

Theorem 3. *For all positive real numbers x,*

$$\vartheta(x) < 2x \log 2. \tag{2.2}$$

In particular, $\vartheta(x) = \mathrm{O}(x)$.

Proof. Since

$$\vartheta(x) = \vartheta\bigl([x]\bigr),$$

it is enough to prove the claim when x is a positive integer. Primes will enter the argument as follows. If $0 \leqslant k \leqslant n$, by definition

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!}. \tag{2.3}$$

One way to prove this an integer is to use the Binomial Theorem,

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k; \tag{2.4}$$

here the coefficients in the expansion of $(x+y)^n$ must be integers. The proof of (2.4) will use the rules

$$\binom{n}{0} = 1, \qquad \binom{n}{n} = 1,$$

and

$$0 \leqslant k < n \implies \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1};$$

from these, one can prove directly (by induction) that each $\binom{n}{k}$ is an integer.

It is a standard exercise to show

$$0 < k < p \implies p \,\Bigg|\, \binom{p}{k} \tag{2.5}$$

using Euclid's Lemma [5, Proposition VII.30], that a prime measuring a product measures one of the factors; symbolically,

$$p \mid ab \ \ \& \ \ p \nmid a \implies p \mid b.$$

Thus, in particular, since

$$p \,\Bigg|\, \binom{p}{k} \cdot k! \cdot (p-k)!,$$

but

$$0 < k < p \implies p \nmid k! \cdot (p-k)!,$$

we can conclude (2.5). In the same way,

$$n - k \leqslant k \implies \prod_{k<p\leqslant n} p \;\Big|\; \binom{n}{k}.$$

Since

$$\log \prod_{k<p\leqslant n} p = \vartheta(n) - \vartheta(k),$$

we conclude

$$n - k \leqslant k \implies \vartheta(n) - \vartheta(k) \leqslant \log \binom{n}{k}. \tag{2.6}$$

We can now proceed by strong induction. We assume (2.2) holds when x is an integer less than n. The case where $n = 1$ is trivial. In case $n = 2$, we compute immediately

$$\theta(n) = \theta(2) = \log 2 < 4 \log 2 = 2n \log 2.$$

In case $n = 2m$, where $m > 1$, then, $2m$ being composite,

$$\vartheta(n) = \vartheta(2m) = \vartheta(2m-1) < 2(2m-1)\log 2 < 2n \log 2.$$

We suppose finally $n = 2m + 1$, where again $m > 1$. As a special case of (2.6),

$$\vartheta(n) = \vartheta(2m+1) \leqslant \log \binom{2m+1}{m+1} + \vartheta(m+1). \tag{2.7}$$

We have also

$$\binom{2m+1}{m+1} = \binom{2m+1}{m},$$

and these are distinct terms in the expansion of $(1+1)^{2m+1}$, by (2.4); so

$$2\binom{2m+1}{m+1} \leqslant 2^{2m+1}, \qquad \binom{2m+1}{m+1} \leqslant 2^{2m}.$$

Plugging this, and the strong inductive hypothesis

$$\vartheta(m+1) < 2(m+1)\log 2,$$

into (2.7), we obtain

$$\vartheta(n) \leqslant 2(2m+1)\log 2 = \log 2n \log 2. \qquad \square$$

Erdős attributes the following to Legendre.

Theorem 4. *For all positive integers n,*

$$\log(n!) = \sum_{p \leqslant n} \log p \cdot \sum_{j \in \mathbb{N}} \left[\frac{n}{p^j}\right]. \tag{2.8}$$

Proof. We are counting the number of times that each p is a factor of $n!$. Such p must be no greater than n. To the number of times that some fixed p is a factor of $n!$,

- each multiple ℓp of p, where $\ell p \leqslant n$, contributes one;
- each multiple ℓp^2 of p, where $\ell p^2 \leqslant n$, contributes an additional one, besides the factor contributed because $(\ell p)p \leqslant n$;
- each multiple ℓp^3 of p, where $\ell p^3 \leqslant n$, contributes an additional one, besides the factors contributed because $(ellp)p \leqslant n$ and $(\ell p^2)p \leqslant n$;

and so on. The result is the sum over all j in $\mathbb{N}$ of those ℓp^j such that $\ell p^j \leqslant n$. The number of such multiples ℓp^j is $[n/p^j]$. $\square$

The sum in (2.8) has a term $\log p \cdot [n/p^j]$ precisely for those prime-powers (that is, powers of primes) p^j that are no greater than n. Thus we might write

$$\log(n!) = \sum_{p^j \leqslant n} \log p \cdot \left[\frac{n}{p^j}\right].$$

Alternatively,

$$\log(n!) = \sum_{k \leqslant n} \Lambda(k) \cdot \left[\frac{n}{k}\right],$$

where Λ is the **von Mangoldt function,** given by

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^m \text{ for some positive } m; \\ 0, & \text{otherwise.} \end{cases}$$

We shall use this function in (3.1) on page 23. Meanwhile, we may observe

$$\sum_{m \mid n} \Lambda(m) = \log n.$$

I would call this the Fundamental Theorem of Arithmetic; however, in *Multiplicative Number Theory,* Davenport remarks [2, p. 55],

> Although this can be proved directly, the simplest way of deriving this and similar identities is by comparing coefficients in two Dirichlet series which have the same sum.

We just pass on to the following.

Theorem 5 (Bertrand's Postulate)**.** *For every positive integer n there is a prime p such that*

$$n < p \leqslant 2n. \tag{2.9}$$

Proof. No prime factor of $\binom{2n}{n}$ exceeds $2n$. Thus it is enough to show that $\binom{2n}{n}$ has a prime factor exceeding n. There are exponents $n(p)$ such that

$$\binom{2n}{n} = \prod_{p \leqslant 2n} p^{n(p)}.$$

Thus

$$n(p) \geqslant 1 \iff p \,\Big|\, \binom{2n}{n}, \tag{2.10}$$

$$\log \binom{2n}{n} = \sum_{p \leqslant 2n} n(p) \log p. \tag{2.11}$$

Since also, by (2.3),

$$\log \binom{2n}{n} = \log\big((2n)!\big) - 2\log(n!),$$

we have, by Theorem 4,

$$n(p) = \sum_{j=1}^{\infty} \left(\left[\frac{2n}{p^j}\right] - 2\left[\frac{n}{p^j}\right] \right). \tag{2.12}$$

Here, in each case,

$$0 \leqslant \left(\left[\frac{2n}{p^j}\right] - 2\left[\frac{n}{p^j}\right] \right) \leqslant 1, \tag{2.13}$$

for,

$$2m \leqslant \frac{2n}{k} < 2m+2 \implies \left[\frac{n}{k}\right] = m,$$

$$2m \leqslant \frac{2n}{k} < 2m+1 \implies \left[\frac{2n}{k}\right] = 2m,$$

$$2m+1 \leqslant \frac{2n}{k} < 2m+2 \implies \left[\frac{2n}{k}\right] = 2m+1.$$

Moreover,

$$k > 2n \implies \left[\frac{2n}{k}\right] - 2\left[\frac{n}{k}\right] = 0, \tag{2.14}$$

while

$$p^j > 2n \iff j \log p > \log(2n). \tag{2.15}$$

Plugging (2.13), (2.14), and (2.15) into (2.12) gives

$$n(p) \leqslant \sum_{j \leqslant \log(2n)/\log p} 1 \leqslant \frac{\log(2n)}{\log p}. \tag{2.16}$$

Therefore

$$2 \leqslant n(p) \implies 2\log p \leqslant \log(2n) \implies p \leqslant \sqrt{2n}. \tag{2.17}$$

From (2.11) and (2.10),

$$\log\binom{2n}{n} = \sum_{n(p)\geqslant 1} n(p)\log p = \sum_{n(p)=1} \log p + \sum_{n(p)\geqslant 2} n(p)\log p. \tag{2.18}$$

Here, again by (2.10),

$$\sum_{n(p)=1} \log p \leqslant \sum_{p \mid \binom{2n}{n}} \log p,$$

while by (2.16) and (2.17),

$$\sum_{n(p)\geqslant 2} n(p)\log p \leqslant \sum_{n(p)\geqslant 2} \log(2n) \leqslant \sqrt{2n}\log(2n).$$

Thus (2.18) becomes

$$\log\binom{2n}{n} \leqslant \sum_{p \mid \binom{2n}{n}} \log p + \sqrt{2n}\log(2n). \tag{2.19}$$

Thus we have an upper bound on $\log\binom{2n}{n}$. We introduce a lower bound by noting that, since

$$2^{2n} = \sum_{j=0}^{2n} \binom{2n}{j} = 2 + \sum_{j=1}^{2n-1} \binom{2n}{j},$$

where there are $2n$ terms, the greatest being $\binom{2n}{n}$,

$$2^{2n} \leqslant 2n\binom{2n}{n},$$

$$2n \log 2 \leqslant \log(2n) + \log\binom{2n}{n}. \tag{2.20}$$

Combining (2.19) and (2.20) yields

$$2n \log 2 \leqslant \sum_{p \mid \binom{2n}{n}} \log p + (1 + \sqrt{2n}) \log(2n). \tag{2.21}$$

The left-hand side dominates the second term on the right, by (1.6). We shall show then that $\binom{2n}{n}$ must have prime factors exceeding n, to compensate. By Theorem 3,

$$\sum_{p \left| \binom{2n}{n} \right. \,\&\, p \leqslant n} \log p \leqslant \sum_{p \leqslant n} \log p = \vartheta(n) \leqslant 2n \log 2. \tag{2.22}$$

Since this is just the left-hand side of (2.21), there is no problem so far. However,

$$\frac{2n}{3} < p \leqslant n \implies 2p \leqslant 2n < 3p$$
$$\implies \left[\frac{2n}{p}\right] - 2\left[\frac{n}{p}\right] = 2 - 2 \cdot 1 = 0,$$

and also

$$\frac{2n}{3} < p \ \ \& \ \ n \geqslant 5 \implies p^2 > \frac{4n^2}{9} = \frac{2n}{9} \cdot 2n > 2n$$
$$\implies \left[\frac{2n}{p^2}\right] = 0,$$

and therefore

$$\frac{2n}{3} < p \leqslant n \ \ \& \ \ n \geqslant 5 \implies n(p) = 0.$$

Thus, assuming $n \geqslant 5$, in the manner of (2.22) we have

$$\sum_{p \left| \binom{2n}{n} \right. \ \& \ p \leqslant n} \log p \leqslant \sum_{p \leqslant 2n/3} \log p = \vartheta\left(\frac{2n}{3}\right) \leqslant \frac{4n}{3} \log 2.$$

Combining with (2.21) gives

$$\sum_{p \left| \binom{2n}{n} \right. \ \& \ n<p} \log p \geqslant \frac{2n}{3} \log 2 - \left(1 + \sqrt{2n}\right) \log(2n). \qquad (2.23)$$

We already know that the right-hand side is positive when n is large enough. It is enough to show 631 is large enough, since there is a sequence

$$2,\ 3,\ 5,\ 7,\ 13,\ 23,\ 43,\ 83,\ 163,\ 317,\ 631$$

of primes, where each successive term is the greatest prime that is less than twice the previous term. Since $631 > 512 = 2^9$, let us assume now

$$n > 2^9, \qquad 2n > 2^{10} = 1024, \qquad \sqrt{2n} > 2^5 = 32.$$

Multiplying the right-hand side of (2.23) by what will turn out to be a convenient factor, we compute

$$\begin{aligned}
\frac{3}{\sqrt{2n}\log\sqrt{2n}\log 2}&\left(\frac{2n}{3}\log 2-\bigl(1+\sqrt{2n}\bigr)\log(2n)\right)\\
&=\frac{\sqrt{2n}}{\log\sqrt{2n}}-\frac{2\cdot 3}{\log 2}\left(\frac{1}{\sqrt{2n}}+1\right)\\
&\geqslant\frac{\sqrt{2n}}{\log\sqrt{2n}}-\frac{2\cdot 3}{\log 2}\left(\frac{1}{32}+1\right)\\
&=\frac{\sqrt{2n}}{\log\sqrt{2n}}-\frac{2\cdot 3\cdot 33}{2^5\log 2}\\
&\geqslant\frac{\sqrt{2n}}{\log\sqrt{2n}}-\frac{2\cdot 100\cdot 2^5}{2^{10}\log 2}\\
&\geqslant\frac{\sqrt{2n}}{\log\sqrt{2n}}-\frac{2^5}{5\log 2}>0,
\end{aligned}$$

since $x/\log x$ is an increasing function of x on $[\mathrm{e},\infty)$, since its derivative is $(\log x-1)/(\log x)^2$. □

3 Chebyshev's Theorem

Chebyshev's Theorem is stated [6, Thm 7, p. 9] and proved [6, §§22.1–4, p. 340–6] by Hardy and Wright; we shall follow their proof, which uses the auxiliary function ϑ, as given by (2.1), and also ψ, given by

$$\psi(x) = \sum_{p^m \leqslant x} \log p = \sum_{n \leqslant x} \Lambda(n). \tag{3.1}$$

Hence $\psi(x) = \log \prod_{p^m \leqslant x} p$, so that, for example,

$$2 < 3 < 2^2 < 5 \ \ \& \ \ \psi(x) = \log(2 \cdot 3 \cdot 2 \cdot 5).$$

Also

$$\psi(x) = \log \operatorname{lcm}([x]!).$$

Landau also proves the theorem [8, Thm 112, pp. 88–91].

Theorem 6 (Chebyshev)**.**

$$\pi(x) \asymp \frac{x}{\log x}.$$

Proof. We have to bound $\pi(x)$ from both sides. From Theorem 3, we have an upper bound $2x \log 2$ on $\vartheta(x)$. We shall relate this to $\pi(x)$. We observe

$$\vartheta(x) \geqslant \sum_{x \geqslant p > \sqrt{x}} \log \sqrt{x} = \bigl(\pi(x) - \pi(\sqrt{x})\bigr) \log \sqrt{x}. \tag{3.2}$$

Using here $\log\sqrt{x} = \frac{1}{2}\log x$ and

$$\pi(\sqrt{x}) \leqslant \sqrt{x} = \frac{x}{\sqrt{x}} \leqslant \frac{x}{\log x},$$

we obtain

$$\vartheta(x) \geqslant \left(\pi(x) - \frac{x}{\log x}\right)\frac{\log x}{2}, \tag{3.3}$$

$$\frac{2\vartheta(x)+x}{\log x} \geqslant \pi(x), \tag{3.4}$$

$$(4\log 2 + 1)\frac{x}{\log x} \geqslant \pi(x) \tag{3.5}$$

by Theorem 3. This is half of Chebyshev's Theorem. For the other half, we first work out a complement of (3.2), namely

$$\vartheta(x) \leqslant \sum_{p\leqslant x}\log x = \pi(x)\log x, \tag{3.6}$$

so that we have, complementing (3.4),

$$\frac{\vartheta(x)}{\log x} \leqslant \pi(x). \tag{3.7}$$

We want to bound $\vartheta(x)$ below by a multiple of x. We shall do this by

(i) bounding $\psi(x)$ in the same way, and

(ii) showing $\psi(x) - \vartheta(x) \prec x$.

If we have (ii), and (i) in the form $\psi(x) \geqslant Ax$, then for x large enough,

$$\frac{A}{2} \geqslant \frac{\psi(x)-\vartheta(x)}{x}, \qquad \vartheta(x) \geqslant \frac{A}{2}x. \tag{3.8}$$

To achieve (ii), we shall use an upper bound on $\vartheta(x)$:

$$\vartheta(x) \leqslant x \log x. \tag{3.9}$$

This will be useful, because

$$\psi(x) = \sum_m \vartheta(x^{1/m}),$$

$$2 \leqslant x^{1/m} \iff m \leqslant \frac{\log x}{\log 2},$$

so that

$$\psi(x) - \vartheta(x) = \sum_{m=2}^{[\log x/\log 2]} \vartheta(x^{1/m}). \tag{3.10}$$

From (3.9),

$$2 \geqslant m \implies \vartheta(x^{1/m}) \leqslant \vartheta(\sqrt{x}) \leqslant \sqrt{x} \log \sqrt{x} \leqslant \sqrt{x} \log x,$$

so that

$$\sum_{m=2}^{[\log x/\log 2]} \vartheta(x^{1/m}) \leqslant \frac{\sqrt{x}(\log x)^2}{\log 2} \prec x.$$

It remains to bound $\psi(x)$ below by a multiple of x. To do this, we let

$$\left[\frac{x}{2}\right] = n.$$

If $x \geqslant 2$, we have

$$\frac{x}{3} \leqslant n.$$

In the notation of, and by, the proof of Theorem 5, we have

$$\log \binom{2n}{n} = \sum_{p \leqslant 2n} n(p) \log p \leqslant \sum_{p \leqslant 2n} \left[\frac{\log(2n)}{\log p}\right] \log p$$
$$= \psi(2n) \leqslant \psi(x).$$

Moreover,

$$2^n \leqslant \prod_{1\leqslant j\leqslant n} \frac{n+j}{j} = \binom{2n}{n},$$

which means now

$$\frac{\log 2}{2} x \leqslant n \log 2 \leqslant \log \binom{2n}{n}.$$

Thus we have the desired bound. □

We can now pass immediately to Theorem 24. Meanwhile, in the process of proving Chebyshev's Theorem, we have established also

$$\vartheta(x) \asymp x \tag{3.11}$$

as well as

$$\psi(x) \asymp x.$$

4 Complex Analysis

A standard text for complex analysis is Ahlfors's *Complex Analysis* [1]. But there is a lot that we shall not need. A terse source is "A first course in complex variables," Problem **4-33** of Spivak's *Calculus on Manifolds* [12, pp. 105–6].

4.1 Holomorphic Functions

We can define limits and derivatives of complex-valued functions as we do for real-valued functions. To do so, we replace finite open intervals with open disks. Each of these is an open *ball* of the appropriate dimension.

If $a \in \mathbb{C}$, and ε is real and positive, the **open disk** having center a and radius ε is the subset

$$\{z \in \mathbb{C}\colon\ |z - a| < \varepsilon\}$$

of $\mathbb{C}$; we may denote this disk by

$$\mathrm{B}(a; \varepsilon).$$

A **neighborhood** of a is a subset of $\mathbb{C}$ that includes some disk $\mathrm{B}(a; \varepsilon)$. A subset of $\mathbb{C}$ is called **open** if it includes a neighborhood of its every point. We shall generally let Ω be an open subset of $\mathbb{C}$, and $f\colon \Omega \to \mathbb{C}$.

An element L of $\mathbb{C}$ is the **limit** of f at a point a in Ω, and we write

$$L = \lim_{z \to a} f(z),$$

provided that, for all positive ε in $\mathbb{R}$, for some positive δ in $\mathbb{R}$, for all z in Ω,

$$0 < |z - a| < \delta \implies |f(z) - L| < \varepsilon.$$

Here the value of $f(a)$ is irrelevant (and we may allow it to be undefined). If however

$$\lim_{z \to a} f(z) = f(a),$$

then f is **continuous** at a.

An element D of $\mathbb{C}$ is the **derivative** of f at a, and we write

$$D = f'(a),$$

if

$$\lim_{z \to a} \frac{f(z) - f(a)}{z - a} = D.$$

If f has a derivative at every point of Ω, then f is called **holomorphic.** In place of $f'(z)$, we may write

$$\mathrm{D}_z\, f(z),$$

especially if we have not got a letter like f in isolation, but only $f(z)$, as in (4.12) below.*

The rules of differentiation familiar from calculus—the constant, multiplication, division, and power rules—still apply to complex differentiation. Moreover, the identity $z \mapsto z$ is holomorphic, with derivative 1. Thus all rational functions are holomorphic on their domains.

---

*According to Kline in *Mathematical Thought from Ancient to Modern Times* [7, p. 642], the French form of the term "holomorphic" was introduced by Briot (1817–82) and Bouquet (1819–85) to replace the term *synectique,* which had been introduced by Cauchy in 1851.

The fundamental example of a function that is *not* holomorphic is **complex conjugation,** $z \mapsto \overline{z}$, where

$$\overline{x + \mathrm{i}\, y} = x - \mathrm{i}\, y.$$

For, the function $z \mapsto (\overline{z} - \overline{a})/(z - a)$ has no limit at a, since

$$\frac{\overline{z} - \overline{a}}{z - a} = \frac{\overline{z - a}}{z - a} = \frac{\overline{z - a}^2}{|z - a|^2},$$

and this is on the unit circle and can be anywhere on that circle.

Ahlfors uses the term *analytic* for the functions that we are calling holomorphic [1, p. 24], For Spivak, an analytic function is a function that is holomorphic in our sense *and* has continuous derivative. Ahlfors remarks [1, p. 101],

> it is only quite recently that it became possible to prove, without resorting to complex integrals or similar tools, that the derivative of an analytic function is continuous, or that the higher derivatives exist. At present the integration-free proofs are, to say the least, much more difficult than the classical proofs.

The same might be said for analysis-free proofs of the Prime Number Theorem.

In another approach to taking derivatives, we may consider $\mathbb{C}$ merely as the two-dimensional real vector-space $\mathbb{R}^2$, where

$$1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathrm{i} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

If there is a linear transformation T of $\mathbb{R}^2$ such that

$$\lim_{z \to a} \frac{|f(z) - f(a) - T(z - a)|}{|z - a|} = 0,$$

then let us write

$$T = \mathrm{D}\, f(a).$$

If $\mathrm{D}\, f(a)$ exists for all a in Ω, we may say that f is **differentiable.**

We can express $\mathrm{D}\, f$ in terms of partial derivatives. If we define

$$f = g + \mathrm{i}\, h,$$

where g and h are real-valued functions, then

$$Df(a)(x + \mathrm{i}\, y) = \begin{pmatrix} \partial_x g(a) & \partial_y g(a) \\ \partial_x h(a) & \partial_y h(a) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{4.1}$$

where for example

$$\partial_y g(a) = \lim_{h \to 0} \frac{g(a + \mathrm{i}\, h) - g(a)}{h},$$

the h here being real. Complex conjugation is differentiable, its derivative being given by the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In general, we may write

$$\partial_x g(a) + \mathrm{i}\, \partial_x h(a) = \partial_x f(a), \quad \partial_y g(a) + \mathrm{i}\, \partial_y h(a) = \partial_y f(a).$$

For f to be differentiable, it is necessary for $\partial_x f$ and $\partial_y f$ to exist, and sufficient that they be continuous [12, p. 31].

In the following, (4.4) are the **Cauchy–Riemann Equations.**

Theorem 7. *For the function f on Ω to be holomorphic, it is necessary and sufficient that f be differentiable and*

$$\mathrm{i}\, \partial_x f = \partial_y f. \tag{4.2}$$

Proof. If $f'(a)$ exists, this means

$$\lim_{z\to a}\frac{f(z)-f(a)-f'(a)(z-a)}{z-a}=0. \tag{4·3}$$

Here multiplication by $f'(a)$ is a *complex*-linear transformation of $\mathbb{C}$, and therefore a real-linear transformation of $\mathbb{R}^2$. In particular,

$$(b+\mathrm{i}\,c)(x+\mathrm{i}\,y)=bx-cy+\mathrm{i}\,(cx+by)=\begin{pmatrix}b&-c\\c&b\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$

Comparison with (4.1) shows that $Df(a)$ is complex-linear at each a in Ω if and only if, on Ω,

$$\partial_x g=\partial_y h, \qquad \partial_x h=-\partial_y g. \tag{4·4}$$

These equations are summarized as (4.2). □

The theorem follows also from the observation that, if $f'(a)$ exists, then

$$f'(a)=\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}=\partial_x f(a),$$

where h is real, but also

$$\mathrm{i}\,f'(a)=\lim_{h\to 0}\frac{f(a+\mathrm{i}\,h)-f(a)}{h}=\partial_y f(a).$$

4.2 Analytic Functions

If, for every a in Ω, there is an infinite sequence $(c_n\colon n\in\omega)$, and there is some positive ε, such that

$$|z-a|<\varepsilon \implies f(z)=\sum_{n\in\omega}c_n(z-a)^n,$$

then f is called **analytic.** In this case, f has all higher derivatives $f^{(n)}(a)$ at every point a of Ω, and moreover

$$\mathrm{B}(a;\varepsilon) \subseteq \Omega \ \ \& \ \ z \in \mathrm{B}(a;\varepsilon) \implies f(z) = \sum_{n\in\omega} \frac{f^{(n)}(a)}{n!}(z-a)^n.$$

In particular, analytic functions are holomorphic. We shall prove the converse as Theorem 16.

Suppose now $\gamma\colon [a,b] \to \Omega$. We may make the analysis

$$\gamma = \gamma_0 + \mathrm{i}\,\gamma_1, \tag{4.5}$$

where $\gamma_e\colon [a,b] \to \mathbb{R}$. Now we can define

$$\int_a^b \gamma = \int_a^b \gamma_0 + \mathrm{i} \int_a^b \gamma_1,$$

provided the integrals on the right exist.

Suppose further that γ is continuously differentiable (which means that each γ_e is continuously differentiable) and one-to-one. Then γ can be said to have the **initial point** $\gamma(a)$ and the **terminal point** $\gamma(b)$, and γ itself can be called an **arc** or **path** or **line.** If the range of γ is included in the domain Ω of the holomorphic function f, then by definition,

$$\int_\gamma f = \int_a^b (f \circ \gamma)\gamma'. \tag{4.6}$$

We may also write this as

$$\int_\gamma f(z)\,\mathrm{d}\,z = \int_a^b f(\gamma(t))\gamma'(t)\,\mathrm{d}\,t.$$

This is a **line integral,** and it is computed by the formal substitution

$$z = \gamma(t), \qquad \mathrm{d}\,z = \gamma'(t)\,\mathrm{d}\,t.$$

We may confuse a path like γ with its range and its initial and terminal points, by the following.

Theorem 8. *If δ is an arc with the same range, initial point, and terminal point as γ, then*

$$\int_\gamma f = \int_\delta f,$$

Proof. Suppose the domain of δ is $[c, d]$. By the substitution

$$\gamma(t) = \delta(u), \qquad \gamma'(t)\,\mathrm{d}\,t = \delta'(u)\,\mathrm{d}\,u,$$

we obtain

$$\int_a^b f(\gamma(t))\gamma'(t)\,\mathrm{d}\,t = \int_a^b f(\delta(u))\delta'(u)\,\mathrm{d}\,u. \qquad \square$$

We can allow arcs to be merely "piecewise" continuously differentiable, that is, continuously differentiable at all but finitely many points of their domains.

Understanding z as $x + \mathrm{i}\,y$, we may write

$$\int_\gamma f = \int_\gamma f(z)\,\mathrm{d}\,z = \int_\gamma \bigl(f(z)\,\mathrm{d}\,x + \mathrm{i}\,f(z)\,\mathrm{d}\,y\bigr). \tag{4.7}$$

Writing f as $f_0 + \mathrm{i}\,f_1$, we obtain

$$\int_\gamma f = \int_\gamma \bigl(f_0(z)\,\mathrm{d}\,x - f_1(z)\,\mathrm{d}\,y\bigr) + \mathrm{i}\int_\gamma \bigl(f_1(z)\,\mathrm{d}\,x + f_0(z)\,\mathrm{d}\,y\bigr). \tag{4.8}$$

This is just notation for now, but it suggests an alternative approach to defining the complex integral. That is, given two functions g and h from Ω to $\mathbb{R}$, we can define

$$\int_\gamma \big(g(z)\,\mathrm{d}\,x + h(z)\,\mathrm{d}\,y\big) = \int_a^b \big((g\circ\gamma){\gamma_0}' + (h\circ\gamma){\gamma_1}'\big),$$

when the latter exists. Using this twice in (4.8), we recover (4.6).

We also define an integral with respect to arc length:

$$\int_\gamma f(z)\,|\mathrm{d}\,z| = \int_a^b (f\circ\gamma)\,|\gamma'|\,.$$

Theorem 9.

$$\left|\int_\gamma f\right| \leqslant \int_\gamma |f(z)|\,|\mathrm{d}\,z|\,.$$

Proof. For any real θ,

$$\mathrm{e}^{\mathrm{i}\,\theta}\int_\gamma f = \int_\gamma \mathrm{e}^{\mathrm{i}\,\theta} f = \int_a^b \mathrm{e}^{\mathrm{i}\,\theta}(f\circ\gamma)\gamma',$$

and so

$$\Re\left(\mathrm{e}^{\mathrm{i}\,\theta}\int_\gamma f\right) = \int_a^b \Re\left(\mathrm{e}^{\mathrm{i}\,\theta}(f\circ\gamma)\gamma'\right)$$
$$\leqslant \int_a^b |f\circ\gamma|\,|\gamma'| = \int_\gamma |f(z)|\,|\mathrm{d}\,z|\,.$$

Finally, for appropriate θ,

$$\left|\int_\gamma f\right| = \Re\left(\mathrm{e}^{\mathrm{i}\,\theta}\int_\gamma f\right).$$

□

If $\int_\gamma f$ depends only on the initial and terminal points of γ, we may say that the integral is **path independent.**

Theorem 10. *The line integral of a continuous derivative of a holomorphic function is path independent.*

Proof. If $f = F'$, then by the Chain Rule,

$$(F \circ \gamma)' = (F' \circ \gamma)\gamma' = (f \circ \gamma)\gamma',$$

and so by (4.6) and the Fundamental Theorem of Calculus,

$$\int_\gamma f = \int_a^b (F \circ \gamma)' = F(\gamma(b)) - F(\gamma(a)). \qquad \square$$

Theorem 11. *If the line integrals of a continuous function are path independent, then the function is the derivative of a holomorphic function.*

Proof. Suppose line integrals of a continuous function f from Ω to $\mathbb{C}$ are path independent. Fix a point c of Ω. If, for a point d of Ω, there are arcs in Γ with initial point c and terminal point d, let γ be one of them. We can unambiguously define

$$F(d) = \int_\gamma f. \tag{4.9}$$

This gives us a function F defined on the set of points to which there is a path from c. Supposing γ has domain $[a, b]$, we can write (4.9) as

$$(F \circ \gamma)(b) = \int_a^b (f \circ \gamma)\gamma'.$$

Since Ω is open, it includes a neighborhood of b, and we can therefore extend γ beyond d. In particular, for some positive

δ, on $[a, b+\delta)$ we have

$$(F \circ \gamma)(x) = \int_a^x (f \circ \gamma)\gamma'.$$

By the Fundamental Theorem, $F \circ \gamma$ is differentiable at b, and

$$(F \circ \gamma)'(b) = f(\gamma(b))\gamma'(b) = f(d)\gamma'(b). \tag{4.10}$$

Near b, we can require γ to be either of

$$t \mapsto d + t - b, \qquad t \mapsto d + \mathrm{i}\,(t - b),$$

yielding respectively

$$\begin{aligned} \gamma'(b) &= 1, & \gamma'(b) &= \mathrm{i}\,, \\ (F \circ \gamma)'(b) &= \partial_x F(d), & (F \circ \gamma)'(b) &= \partial_y F(d). \end{aligned}$$

Then from (4.10) we obtain

$$\partial_x F(d) = f(d), \qquad \partial_y F(d) = \mathrm{i}\, f(d). \tag{4.11}$$

Thus F satisfies the Cauchy–Riemann Equations, in the combined form (4.2). In particular, if F is differentiable, then it is holomorphic, and $F' = f$. Again, a sufficient condition for the differentiability of F (given the existence of the partial derivatives) is continuity of f. □

If we suppose only that the real part of line integrals of f are path independent, this means by (4.8) that the integrals

$$\int_\gamma \bigl(f_0(z)\,\mathrm{d}\,x - f_1(z)\,\mathrm{d}\,y\bigr)$$

are path independent. We obtain F as before, but now real-valued; and in place of (4.11),

$$\partial_x F(d) = f_0(d), \qquad \partial_y F(d) = -f_1(d).$$

One then refers to the integrand $f_0(z)\,\mathrm{d}\,x - f_1(z)\,\mathrm{d}\,y$ as an **exact differential.** One may refer to the integrand $f(z)\,\mathrm{d}\,x + \mathrm{i}\,f(z)\,\mathrm{d}\,y$ in (4.7) in the same way.

The last two theorems are summarized by Ahlfors [1, p. 107]:

> *The integral $\int_\gamma f\,\mathrm{d}\,z$, with continuous f, depends only on the endpoints of γ if and only if f is the derivative of an analytic function on Ω.*
>
> Under these circumstance we shall prove later that $f(z)$ is itself analytic.

To this end, we prove the following.

Theorem 12. *The integral of a holomorphic function along the boundary of a rectangle (whose sides are parallel to the coordinate axes) is zero.*

Proof. Say the function is f and the rectangle is traced by γ_0. By bisecting the rectangle vertically and horizontally, we obtain four sub-rectangles. For one of them, if γ_1 traces its boundary, then

$$\left|\int_{\gamma_1} f\right| \geqslant \frac{1}{4}\left|\int_{\gamma_0} f\right|.$$

Continuing in this way, we get a sequence of γ_k, where

$$\left|\int_{\gamma_k} f\right| \geqslant \frac{1}{4^k}\left|\int_{\gamma_0} f\right|,$$

and all of the γ_k surround a unique point a. If n is large enough, then inside γ_n,

$$\left|\frac{f(z)-f(a)}{z-a}-f'(a)\right|\leqslant\varepsilon,$$
$$|f(z)-f(a)-(z-a)f'(a)|\leqslant\varepsilon\,|z-a|\,.$$

Since, by Theorem 10,

$$\int_{\gamma_n} f(a)=0, \qquad \int_{\gamma_n}(z-a)f'(a)\,\mathrm{d}\,z=0,$$

we have now, by Theorem 9,

$$\begin{aligned}\left|\int_{\gamma_n} f\right|&=\left|\int_{\gamma_n}\big(f(z)-f(a)-(z-a)f'(a)\big)\,\mathrm{d}\,z\right|\\&\leqslant\int_{\gamma_n}|f(z)-f(a)-(z-a)f'(a)|\,|\mathrm{d}\,z|\\&\leqslant\varepsilon\int_{\gamma_n}|z-a|\,|\mathrm{d}\,z|\leqslant\frac{\varepsilon dL}{4^n},\end{aligned}$$

where d is the diagonal of the rectangle surrounded by γ_0, and L is the perimeter of the rectangle. Thus

$$\left|\int_{\gamma_0} f\right|\leqslant\varepsilon dL.$$

This being true for all positive ε, $\int_{\gamma_0} f=0$. □

Theorem 13. *Holomorphic functions on convex regions are derivatives of holomorphic functions.*

Proof. Let Ω be convex, f holomorphic on Ω, and $a \in \Omega$. We define F on Ω so that, if $b \in \Omega$, then

$$F(b) = \int_\gamma f,$$

where γ is the path from a to b that is first horizontal, then vertical. By Theorem 12, γ may also be first vertical, then horizontal. This shows, as in the proof of Theorem 11, that F is holomorphic with derivative f. $\square$

An arc whose initial and terminal points are the same is a **closed curve.** The line integrals of a function are path independent if and only if the integrals around closed curves are path independent. The latter case means those integrals are zero.

For example, $(z-a)^n$ is holomorphic on its domain, which is $\mathbb{C}$ if $n \geqslant 0$, and $\{z \in \mathbb{C} \colon z \neq a\}$, if $n < 0$. Also

$$\mathrm{D}_z(z-a)^n = n(z-a)^{n-1}. \tag{4.12}$$

Therefore, letting $\oint$ denote an integral along a counterclockwise closed curve in the domain of $(z-a)^n$, we have

$$n \neq 0 \implies \oint (z-a)^{n-1}\,\mathrm{d}\,z = 0.$$

Theorem 14. *If γ describes a counterclockwise loop around a, then*

$$\int_\gamma \frac{\mathrm{d}\,z}{z-a} = 2\mathrm{i}\,\pi.$$

Proof. We may assume $a = 0$. If δ is $t \mapsto \mathrm{e}^{\mathrm{i}\,t}$ on $[0, 2\pi]$, we compute

$$\int_\delta \frac{\mathrm{d}\,z}{z} = \int_0^{2\pi} \frac{\mathrm{i}\,\mathrm{e}^{\mathrm{i}\,t}\,\mathrm{d}\,t}{e^{\mathrm{i}\,t}} = \mathrm{i}\int_0^{2\pi} \mathrm{d}\,t = 2\mathrm{i}\,\pi.$$

The general case follows from Theorem 10, since we can analyze $\delta - \gamma$ as a sum of closed curves, each surrounding a region where $1/z$ is the derivative of a holomorphic function (which we ambiguously call $\log z$). □

In the following, (4.13) is **Cauchy's Integral Formula.**

Theorem 15. *If f is holomorphic on an open neighborhood of a, and γ describes a counterclockwise loop around a within that neighborhood, then*

$$f(a) = \frac{1}{2\mathrm{i}\,\pi} \int_\gamma \frac{f(z)}{z-a} \,\mathrm{d}\, z. \tag{4.13}$$

Proof. Again we may assume $a = 0$. By Theorem 14,

$$f(0) = \frac{1}{2\mathrm{i}\,\pi} \int_\gamma \frac{f(0)}{z} \,\mathrm{d}\, z. \tag{4.14}$$

As in the proof of Theorem 14, we may adjust γ, now shrinking it to a circle of radius δ around 0. we may let δ be small enough that

$$|f(z) - f(0)| < \varepsilon$$

on γ. Using (4.14), we compute

$$\begin{aligned}\left|\frac{1}{2\mathrm{i}\,\pi}\int_\gamma \frac{f(z)}{z}\,\mathrm{d}\, z - f(0)\right| &= \left|\frac{1}{2\mathrm{i}\,\pi}\int_\gamma \frac{f(z)-f(0)}{z}\,\mathrm{d}\, z\right| \\ &= \frac{1}{2\pi}\left|\int_0^1 \frac{f(\delta \mathrm{e}^{2\mathrm{i}\,\pi t}) - f(0)}{\delta \mathrm{e}^{2\mathrm{i}\,\pi t}} 2\mathrm{i}\,\pi\delta \mathrm{e}^{2\mathrm{i}\,\pi t}\,\mathrm{d}\, t\right| \\ &\leqslant \int_0^1 \left|f(\delta \mathrm{e}^{2\mathrm{i}\,\pi t}) - f(0)\right|\mathrm{d}\, t < \varepsilon. \qquad \square\end{aligned}$$

Theorem 16. *Holomorphic functions are analytic.*

Proof. Let γ describe a circle in the domain of a holomorphic function f. We may assume the center of the circle is 0. Let w be a point inside the circle. By Cauchy's Integral Formula (4.13), and then the rule (5.3) for geometric series,

$$\begin{aligned} 2\mathrm{i}\,\pi f(w) &= \int_\gamma \frac{f(z)}{z-w}\,\mathrm{d}\,z = \int_\gamma \frac{f(z)}{z(1-w/z)}\,\mathrm{d}\,z \\ &= \int_\gamma \frac{f(z)}{z} \sum_{n\in\omega} \left(\frac{w}{z}\right)^n \mathrm{d}\,z = \int_\gamma \sum_{n\in\omega} \frac{f(z)w^n}{z^{n+1}}\,\mathrm{d}\,z. \end{aligned}$$

Since $f(z)$ is bounded on γ (this being compact), the convergence of the series is absolute, so we can interchange the integration and summation:

$$f(w) = \sum_{n\in\omega} \frac{1}{2\mathrm{i}\,\pi} \left(\int_\gamma \frac{f(z)\,\mathrm{d}\,z}{z^{n+1}} \right) w^n. \qquad \square$$

To say that f is **analytic at** a means f is analytic on some neighborhood of a. If $n \in \mathbb{N}$, and $(z-a)^n f(z)$ is analytic at a, but $(z-a)^{n-1} f(z)$ is not, then f has a **pole of order** n at a. In this case, the **residue** of f at a is the coefficient of $(z-a)^{-1}$ in the power-series expansion of f; thus the residue is

$$\frac{1}{2\mathrm{i}\,\pi} \oint f,$$

the integral being taken around a along a curve not encompassing any other poles. If $f'(a)$ exists for all a in Ω, except for poles, then f is **meromorphic.**

If $f(z)/(z-a)^n$ is analytic at a, but $f(z)/(z-a)^{n-1}$ is not, then f has a **zero of order** n at a. Thus a zero of order n of f at a is a pole of order n of $1/f$ at a.

Theorem 17. *The order of a zero, or the negative of the order of a pole, of f at a is the residue of f'/f at a; and here, f'/f has a pole of order* 1.

Proof. From

$$f(z) = c_n(x-a)^n + c_{n+1}(x-a)^{n+1} + \cdots,$$
$$f'(z) = nc_n(x-a)^{n-1} + (n+1)c_{n+1}(x-a)^n + \cdots,$$

we obtain

$$\frac{f'(z)}{f(z)} = n(x-a)^{-1} + \cdots. \qquad \square$$

Note that, in general,

$$\frac{f'(z)}{f(z)} = \mathrm{D}_z \log(f(z));$$

we may refer to this as the **logarithmic derivative** of f.

5 The Prime Number Theorem

We shall prove the Prime Number Theorem as Theorem 25 below. There are other formulations. With the definition of (1.7), Mazur and Stein [9, chh. 13 & 17, pp. 49 & 58] give first

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

This is equivalent, by (1.8), to (5.13), which is the formulation of Hardy and Wright [6, Thm 6, p. 9], who prove it over the course of 27 pages [6, §§22.1–17, pp. 340–67].

We shall follow and expand Don Zagier's short article "Newman's Short Proof of the Prime Number Theorem" [13]. The argument there is highly concentrated and proceeds in six numbered steps, our Theorems 18, 19, 3, 20, 22, and 23 respectively. The ultimate conclusion, Theorem 25, needs one more proof, namely that of our Theorem 24. Theorem 22 uses an "analytic theorem," our Theorem 21; this is Newman's contribution, reformulated by Zagier, who says he will both

1) reproduce "an ingenious short proof" that the Riemann zeta function has no zeros with real part 1 (this is part of Theorem 20), and
2) "describe" the analytic proof of the Prime Number Theorem from this, using "a very simple version [found by D. J. Newman] of the Tauberian argument needed."

As they are formulated here, we can prove Theorems 18, 19, 21, 23, and 24 in any order. Theorem 20 needs Theorems 18 and 19; Theorem 22 needs Theorems 20 and 21; Theorem 25 needs Theorems 22, 23, and 24. See Figure 5.1. We now

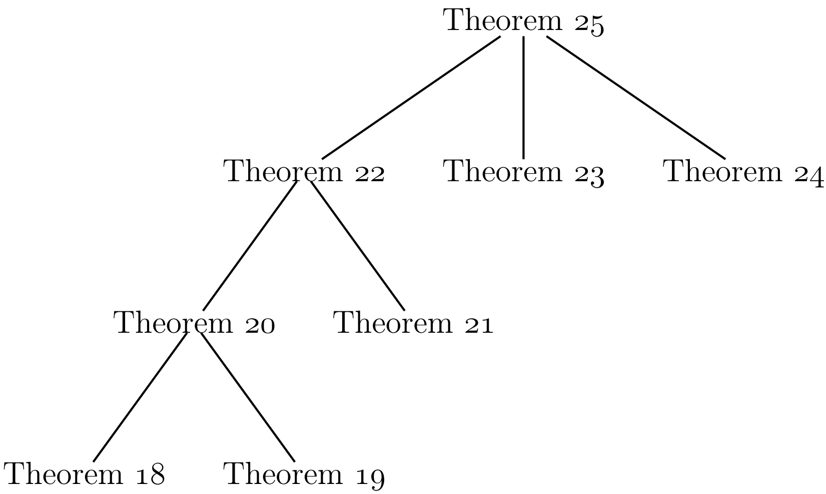

Figure 5.1: Dependence of theorems

proceed.

The **Riemann zeta function** is the function ζ given on the half-plane $\{s \in \mathbb{C} \colon \Re(s) > 1\}$ by

$$\zeta(s) = \sum_{n \in \mathbb{N}} \frac{1}{n^s}. \tag{5.1}$$

To simplify such expressions, we may use Riemann's notation,

$$s = \sigma + \mathrm{i}\, t,$$

so that, in particular,

$$\Re(s) = \sigma.$$

Speaking elliptically, we say that ζ is the function $\sum_{n\in\mathbb{N}} 1/n^s$, where $\sigma > 1$. The convergence of the sum is absolute, by the Integral Test, since

$$n^s = \mathrm{e}^{s \log n} = \mathrm{e}^{\mathrm{i}\, t \log n} \mathrm{e}^{\sigma \log n} = \mathrm{e}^{\mathrm{i}\, t \log n} n^{\sigma},$$

and so

$$\left|\frac{1}{n^s}\right| = \frac{1}{n^\sigma},$$

while $\int_1^\infty \mathrm{d}\,t/t^\sigma$ converges when $\sigma > 1$, by (1.4).

Theorem 18. *When $\sigma > 1$, then*

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}}. \tag{5.2}$$

Proof. The factors are sums of geometric series:

$$\frac{1}{1 - z^{-1}} = \sum_{k\in\omega} \frac{1}{z^k}. \tag{5.3}$$

Consequently, if P is a finite set of primes, then

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

where ω^P is the set of functions from P to ω. In the sum, n^{-s} appears, provided n is the product of primes coming only from P. Let N_P be the set of such n. By the Fundamental Theorem of Arithmetic, a term n^{-s} appears only once in the sum above. Thus

$$\prod_{p\in P} \frac{1}{1 - p^{-s}} = \sum_{n\in N_P} \frac{1}{n^{-s}}. \tag{5.4}$$

Since the sum in (5.1) converges absolutely, no matter how P grows to encompass all primes, the product in (5.4) approaches the sum in (5.1) as a limit. □

Theorem 19. *The function ζ has a pole of order 1 at 1, but the function*

$$\zeta(s) - \frac{1}{s-1}$$

extends holomorphically to the half-plane $\sigma > 0$ (and thus the residue of ζ at 1 is 1).

Proof. When $\sigma > 1$, we have, by (1.4) on page 9,

$$\zeta(s) - \frac{1}{s-1} = \sum_{n \in \mathbb{N}} \frac{1}{n^s} - \int_1^\infty \frac{\mathrm{d}\,x}{x^s} = \sum_{n \in \mathbb{N}} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) \mathrm{d}\,x,$$

and the last series converges absolutely when $\sigma > 0$ (and thus can be used to define $\zeta(s) - (s-1)^{-1}$ then), since

$$\frac{1}{n^s} - \frac{1}{x^s} = s \int_n^x \frac{\mathrm{d}\,u}{u^{s+1}},$$

and this, on $[n, n+1]$, is bounded absolutely by $|s| / n^{\sigma+1}$. □

We now introduce the function Φ, given on the half-plane $\sigma > 1$ by

$$\Phi(s) = \sum_p \frac{\log p}{p^s}.$$

Theorem 20. *The function*

$$\Phi(s) - \frac{1}{s-1}$$

extends holomorphically to $\sigma \geqslant 1$, and ζ is non-zero here.

Proof. We already know from (5.2) that ζ is nonzero on $\sigma > 1$. Therefore, on this domain, the logarithmic derivative ζ'/ζ of ζ is holomorphic. From

$$\log \zeta(s) = \sum_p \log\bigl((1 - p^{-s})^{-1}\bigr),$$

using $p^{-s} = \mathrm{e}^{-s \log p}$, we compute

$$\begin{aligned} \frac{\zeta'(s)}{\zeta(s)} &= \sum_p (1 - p^{-s})(-1)(1 - p^{-s})^{-2} p^{-s} \log p \\ &= -\sum_p \frac{\log p}{p^s(1 - p^{-s})} = -\sum_p \frac{\log p}{p^s - 1}. \end{aligned}$$

Since

$$\frac{1}{x-1} = \frac{1}{x} + \frac{1}{x(x-1)},$$

we compute now

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \left(\frac{\log p}{p^s} + \frac{\log p}{p^s(p^s - 1)} \right) = \Phi(s) + \sum_p \frac{\log p}{p^s(p^s - 1)}.$$

The last series converges absolutely when $\sigma > 1/2$, since for large p,

$$\frac{1}{p^{2s} - p^s} < \frac{2}{p^{2s}}.$$

Thus Φ extends *meromorphically* to $\sigma > 1/2$, as ζ'/ζ does. Since ζ has a pole of order 1 at 1 by Theorem 19, so does $-\zeta'/\zeta$, and its residue is 1, by Theorem 17. Then also Φ has at 1 a pole of order 1 and the residue 1. Thus $\Phi(s) - (s-1)^{-1}$ is analytic at 1 as well as meromorphic on $\sigma \geqslant 1$, and it will have poles only at zeroes of ζ.

By the product formula (5.2), ζ has no zero when $\sigma > 1$. Suppose if possible ζ has a zero of order μ at $1 + \mathrm{i}\, a$. Then

$$-\mu = -\lim_{z \to 1+\mathrm{i}\, a} \frac{(z - 1 - \mathrm{i}\, a)\zeta'(z)}{\zeta(z)} \\ = -\lim_{z \to 0} \frac{z\zeta'(1 + \mathrm{i}\, a + z)}{\zeta(1 + \mathrm{i}\, a + z)} = \lim_{z \to 0} z\Phi(1 + \mathrm{i}\, a + z).$$

Also, because of the pole and the residue at 1,

$$1 = \lim_{z \to 0} z\Phi(z).$$

Now, ζ will have a zero of *some* order ν, possibly 0, at $1 + 2\mathrm{i}\, a$. Since

$$\overline{\Phi(s)} = \Phi(\overline{s}),$$

so that

$$2\Re(\Phi(s)) = \Phi(s) + \Phi(\overline{s}),$$

we have, when ε is a small positive real number,

$$\sum_{j=-2}^{2} \binom{4}{2+j} \Phi(1 + j\mathrm{i}\, a + \varepsilon) \\ = 2\Re(\Phi(1 + 2\mathrm{i}\, a + \varepsilon)) + 8\Re(\Phi(1 + \mathrm{i}\, a + \varepsilon)) \\ + 6\Re(\Phi(1 + \varepsilon)). \quad (5.5)$$

Also,

$$\sum_{j=-2}^{2}\binom{4}{2+j}\Phi(1+j\,\mathrm{i}\,a+\varepsilon)$$
$$=\sum_p \frac{\log p}{p^{1+\varepsilon}}\sum_{j=-2}^{2}\binom{4}{2+j}\frac{1}{p^{j\,\mathrm{i}\,a}}$$
$$=\sum_p \frac{\log p}{p^{1+\varepsilon}}\left(\frac{1}{p^{\mathrm{i}\,a/2}}+\frac{1}{p^{-\mathrm{i}\,a/2}}\right)^4$$
$$=\sum_p \frac{\log p}{p^{1+\varepsilon}}\left(2\Re\Bigl(\frac{1}{p^{\mathrm{i}\,a/2}}\Bigr)\right)^4 \geqslant 0.$$

In the limit at ε, the product with ε of the sum in (5.5) is $-2\nu-8\mu+6$. Since this is not negative, μ must be 0. □

The "analytic theorem" is the following; we use it solely to derive Theorem 22.

Theorem 21. *Any function f that is bounded and locally integrable on $[0,\infty)$ is globally integrable, provided the function g given on $\sigma>0$ by*

$$g(s)=\int_0^\infty f(t)\mathrm{e}^{-st}\,\mathrm{d}\,t$$

extends holomorphically to $\sigma \geqslant 0$. In this case, moreover,

$$g(0)=\int_0^\infty f(t)\,\mathrm{d}\,t.$$

Proof. Defining

$$g_x(z)=\int_0^x f(t)\mathrm{e}^{-zt}\,\mathrm{d}\,t,$$

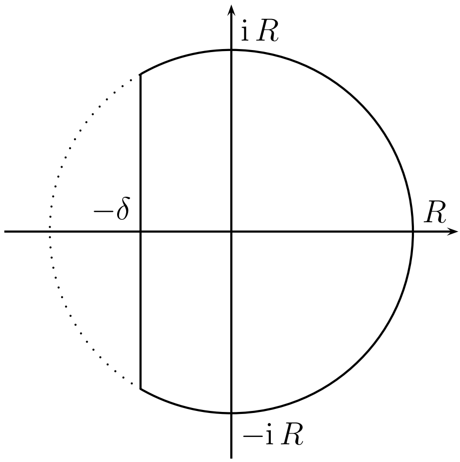

Figure 5.2: A contour for integration

we want to prove

$$g(0) = \lim_{x \to \infty} g_x(0). \tag{5.6}$$

Given large positive R, we can find positive δ so that g is holomorphic on the region $-\delta \leqslant \sigma \;\; \& \;\; |s| \leqslant R$ shown in Figure 5.2. Let γ be a counterclockwise path around this region. By Cauchy's Integral Formula,

$$g(0) - g_x(0) = \frac{1}{2\mathrm{i}\,\pi} \int_\gamma \bigl(g(z) - g_x(z)\bigr) \frac{\mathrm{d}\,z}{z}. \tag{5.7}$$

Now let

$$h_x(s) = \mathrm{e}^{sx}\left(1 + \frac{s^2}{R^2}\right),$$

so h_x is holomorphic on $\mathbb{C}$. The innovation of Newman [10] is to multiply $g(z) - g_x(z)$ by $h_x(z)$, so that, since $h_x(0) = 1$, (5.7) becomes

$$g(0) - g_x(0) = \frac{1}{2\mathrm{i}\,\pi} \int_\gamma \big(g(z) - g_x(z)\big) \frac{h_x(z)}{z} \,\mathrm{d}\, z.$$

We shall bound the integral. We have

$$\left|\frac{h_x(s)}{s}\right| = \frac{\mathrm{e}^{\sigma x}}{R} \left|\frac{R}{s} + \frac{s}{R}\right|,$$

and so

$$|s| = R \implies \left|\frac{h_x(s)}{s}\right| \leqslant \frac{2\mathrm{e}^{\sigma x}\,|\sigma|}{R^2}. \tag{5.8}$$

Since f is assumed to be bounded, we may let

$$B = \sup_{0 \leqslant t} |f(t)|\,,$$

so that

$$\sigma > 0 \implies |g(s) - g_x(s)| = \left|\int_x^\infty f(t)\mathrm{e}^{-st}\,\mathrm{d}\,t\right| \leqslant B \int_x^\infty \mathrm{e}^{-\sigma t}\,\mathrm{d}\,t = \frac{B}{\mathrm{e}^{\sigma x}\sigma}. \tag{5.9}$$

Combining the two estimates (5.8) and (5.9), letting γ_+ be the restriction of γ so that the range is in $\sigma > 0$, and thus the length of γ_+ is πR, we have

$$\left|\int_{\gamma_+} \big(g(z) - g_x(z)\big) \frac{h_x(z)}{z}\,\mathrm{d}\,z\right| \leqslant \pi R \cdot \frac{B}{\mathrm{e}^{\sigma x}\sigma} \cdot \frac{2\mathrm{e}^{\sigma x}\sigma}{R^2} = \frac{2\pi B}{R}. \tag{5.10}$$

Letting γ_- be the other part of γ, we have

$$\left|\int_{\gamma_-}\bigl(g(z)-g_x(z)\bigr)\frac{h_x(z)}{z}\,\mathrm{d}\,z\right| \leqslant \left|\int_{\gamma_-} g(z)\frac{h_x(z)}{z}\,\mathrm{d}\,z\right| + \left|\int_{\gamma_-} g_x(z)\frac{h_x(z)}{z}\,\mathrm{d}\,z\right|. \quad (5.11)$$

For the last integral, since g_x is entire, we can replace γ_- with $\gamma_-{}'$ having the same endpoints, other points having negative real part and absolute value R. Since, as in (5.9),

$$\sigma < 0 \implies |g_x(s)| = \left|\int_0^x f(t)\mathrm{e}^{-st}\,\mathrm{d}\,t\right| \leqslant B\int_{-\infty}^x \mathrm{e}^{-\sigma t}\,\mathrm{d}\,t = \frac{B}{\mathrm{e}^{\sigma x}\,|\sigma|},$$

combining with (5.8) gives, as in (5.10),

$$\left|\int_{\gamma_-} g_x(z)\frac{h_x(z)}{z}\,\mathrm{d}\,z\right| = \left|\int_{\gamma_-{}'} g_x(z)\frac{h_x(z)}{z}\,\mathrm{d}\,z\right| \leqslant \pi R\cdot\frac{B}{\mathrm{e}^{\sigma x}\,|\sigma|}\cdot\frac{2\mathrm{e}^{\sigma x}\,|\sigma|}{R^2} = \frac{2\pi B}{R}.$$

Thus we have a bound of $4\pi B/R$ on everything so far, and we can make this bound as small as we like, by letting R grow large. One integral remains to consider from (5.11). We have

$$\int_{\gamma_-} g(z)\frac{h_x(z)}{z}\,\mathrm{d}\,z = \int_{\gamma_-}\mathrm{e}^{zx}g(z)\left(1+\frac{z^2}{R^2}\right)\frac{\mathrm{d}\,z}{z}.$$

Here x occurs only in the factor e^{zx}. For some large N, we analyze γ_- into components γ_N and γ_N', according to whether the real part of a point is greater than (that is, to the right of) $-\delta/N$ or not. First,

$$\left|\int_{\gamma_N} \mathrm{e}^{zx} g(z)\left(1+\frac{z^2}{R^2}\right)\frac{\mathrm{d}\,z}{z}\right| \leqslant \left|\int_{\gamma_N} g(z)\left(1+\frac{z^2}{R^2}\right)\frac{\mathrm{d}\,z}{z}\right|.$$

We can make this bound, which is independent of x, as small as we like, by making N large enough. Moreover,

$$\left|\int_{\gamma'_N} \mathrm{e}^{zx} g(z)\left(1+\frac{z^2}{R^2}\right)\frac{\mathrm{d}\,z}{z}\right| \leqslant \mathrm{e}^{-\delta x/N}\left|\int_{\gamma'_N} g(z)\left(1+\frac{z^2}{R^2}\right)\frac{\mathrm{d}\,z}{z}\right|,$$

and we can make *this* as small as we like, given N, by making x large enough. Thus for all large R, for all positive ε, for sufficiently large x,

$$|g(0) - g_x(0)| \leqslant \frac{4\pi B}{R} + \varepsilon.$$

This implies (5.6). □

We immediately apply Theorem 21.

Theorem 22. *The integral*

$$\int_1^\infty \frac{\vartheta(x) - x}{x^2}\,\mathrm{d}\,x$$

converges.

Proof. As in the proof of Theorem 2, we consider the primes as forming the increasing sequence $(p_n\colon n \in \omega)$. Now

$$\log p_0 = \vartheta(p_0), \qquad \log p_{n+1} = \vartheta(p_{n+1}) - \vartheta(p_n),$$

so that, when $\sigma > 1$,

$$\Phi(s) = \sum_{n\in\omega} \frac{\log p_n}{{p_n}^s} = \frac{\vartheta(p_0)}{{p_0}^s} + \sum_{n\in\omega} \frac{-\vartheta(p_n) + \vartheta(p_{n+1})}{{p_{n+1}}^s}$$
$$= \sum_{n\in\omega} \vartheta(p_n) \left(\frac{1}{{p_n}^s} - \frac{1}{{p_{n+1}}^s}\right) = \sum_{n\in\omega} \vartheta(p_n) \int_{p_n}^{p_{n+1}} \frac{s\,\mathrm{d}\,x}{x^{s+1}}$$
$$= s \sum_{n\in\omega} \int_{p_n}^{p_{n+1}} \frac{\vartheta(x)}{x^{s+1}}\,\mathrm{d}\,x = s \int_1^\infty \frac{\vartheta(x)}{x^{s+1}}\,\mathrm{d}\,x,$$

and therefore

$$\frac{\Phi(s+1)}{s+1} - \frac{1}{s} = \int_1^\infty \frac{\vartheta(x)}{x^{s+2}}\,\mathrm{d}\,x - \int_1^\infty \frac{\mathrm{d}\,x}{x^{s+1}} = \int_1^\infty \frac{\vartheta(x) - x}{x^{s+2}}\,\mathrm{d}\,x.$$

By Theorem 20, the left-hand side extends holomorphically to $\sigma \geqslant 0$. We want to show that the equation still holds when $s = 0$. To apply Theorem 21, we use the substitution

$$x = \mathrm{e}^t, \qquad \mathrm{d}\,x = \mathrm{e}^t\,\mathrm{d}\,t,$$

obtaining

$$\int_1^\infty \frac{\vartheta(x) - x}{x^{s+2}}\,\mathrm{d}\,x = \int_0^\infty \mathrm{e}^{-st} \left(\frac{\vartheta(\mathrm{e}^t)}{\mathrm{e}^t} - 1\right) \mathrm{d}\,t.$$

Since $\vartheta(\mathrm{e}^t)/\mathrm{e}^t$ is bounded by Theorem 3, we are done. □

The following is a straightforward theorem of calculus, needing nothing special that we have done. However, with Theorem 22, it will allow us to refine Theorem (3.11), $\vartheta(x) \asymp x$.

Theorem 23. *If f is an increasing function such that the integral*

$$\int_1^\infty \frac{f(t) - t}{t^2}\,\mathrm{d}\,t$$

converges, then

$$f(x) \sim x.$$

Proof. Let f be an increasing function such that $f(x) \nsim x$. There are two ways this can happen.

1. Suppose first for some λ in $(1, \infty)$, for arbitrarily large x,

$$f(x) \geqslant \lambda x.$$

For such x, since f is increasing,

$$\int_x^{\lambda x} \frac{f(t) - t}{t^2} \, \mathrm{d}\, t \geqslant \int_x^{\lambda x} \frac{\lambda x - t}{t^2} \, \mathrm{d}\, t = I$$

for some I. Letting $t = xu$, so that $\mathrm{d}\, t = x \, \mathrm{d}\, u$, we have

$$I = \int_1^{\lambda} \frac{\lambda - u}{u^2} \, \mathrm{d}\, u > 0,$$

so $\int_1^x (f(t) - t) \, \mathrm{d}\, t / t^2$ has no limit as x goes to ∞.

2. Similarly, if for λ in $(0, 1)$, for arbitrarily large x,

$$f(x) \leqslant \lambda x,$$

then for such x,

$$\int_{\lambda x}^{x} \frac{f(t) - t}{t^2} \, \mathrm{d}\, t \leqslant \int_{\lambda x}^{x} \frac{\lambda x - t}{t^2} \, \mathrm{d}\, t = \int_{\lambda}^{1} \frac{\lambda - u}{u^2} \, \mathrm{d}\, u < 0,$$

so again $\int_1^x (f(t) - t) \, \mathrm{d}\, t / t^2$ has no limit as x goes to ∞. □

The final ingredient is the following, which we obtain by refining the proof of Chebyshev's Theorem, Theorem 6 (page 23).

Theorem 24.

$$\vartheta(x) \sim \pi(x) \log x. \tag{5.12}$$

Proof. As a variant of (3.2) and (3.3), if $0 < \varepsilon < 1/2$, we have

$$\vartheta(x) \geqslant \sum_{x^{1-\varepsilon} < p \leqslant x} \log(x^{1-\varepsilon}) \geqslant (1-\varepsilon)(\pi(x) - x^{1-\varepsilon}) \log x,$$

$$\vartheta(x) + (1-\varepsilon) x \frac{\log x}{x^{\varepsilon}} \geqslant (1-\varepsilon)\pi(x) \log x,$$

$$\frac{\vartheta(x)}{\pi(x)\log x} + (1-\varepsilon)\frac{x}{\pi(x)\log x} \cdot \frac{\log x}{x^{\varepsilon}} \geqslant 1 - \varepsilon.$$

While $x/\pi(x)\log x$ is bounded above by Theorem 6, while $\log x \prec x^{\varepsilon}$ by (1.6), when x is large enough we have

$$\frac{\vartheta(x)}{\pi(x)\log x} \geqslant 1 - 2\varepsilon.$$

Thus

$$\liminf_{x\to\infty} \frac{\vartheta(x)}{\pi(x)\log x} \geqslant 1.$$

With (3.6), namely $\vartheta(x) \leqslant \pi(x)\log x$, we are done. □

Theorem 25 (The Prime Number Theorem)**.**

$$\pi(x) \sim \frac{x}{\log x}. \tag{5.13}$$

Proof. By Theorems 22, 23, and 24,

$$x \sim \vartheta(x) \sim \pi(x) \log x.$$ □

Bibliography

[1] Lars Ahlfors. *Complex Analysis.* McGraw Hill, New York, second edition, 1966.

[2] Harold Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics.* Springer-Verlag, New York-Berlin, second edition, 1980. Revised by Hugh L. Montgomery.

[3] Leonard Eugene Dickson. *History of the Theory of Numbers. Volume I: Divisibility and Primality.* Dover, Mineola, New York, 2005. Unabridged republication of the work originally published as Publication Number 256 in Washington D.C. by the Carnegie Institute of Washington in 1919.

[4] P. Erdős. Beweis eines Satzes von Tschebyschef (in German). *Acta Litt. Sci. Szeged*, 5:194–198, 1932. Available at `www.renyi.hu/~p_erdos/1932-01.pdf` (as of December 3, 2010).

[5] Euclid. *Euclid's Elements.* Green Lion Press, Santa Fe, NM, 2002. All thirteen books complete in one volume. The Thomas L. Heath translation, edited by Dana Densmore.

[6] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers.* Clarendon Press, Oxford, fifth edition, 1979. First edition 1938. Reprinted 1990.

[7] Morris Kline. *Mathematical Thought from Ancient to Modern Times*. Oxford University Press, New York, 1972.

[8] Edmund Landau. *Elementary Number Theory*. Chelsea Publishing, New York, 1958. Originally part of *Vorlesungen über Zahlentheorie* (Leipzig, 1927). Translated by J. E. Goodman.

[9] Barry Mazur and William Stein. *Prime Numbers and the Riemann Hypothesis*. Cambridge University Press (to appear), September 2015. Draft at `wstein.org/rh/` as of December 29, 2015.

[10] D. J. Newman. Simple analytic proof of the prime number theorem. *Amer. Math. Monthly*, 87(9):693–696, 1980.

[11] Srinivasa Ramanujan. A proof of Bertrand's postulate. *Journal of the Indian Mathematical Society*, XI:181–2, 1919. Available at `www.imsc.res.in/~rao/ramanujan/CamUnivCpapers/Cpaper24/page1.htm` (as of December 3, 2010).

[12] Michael Spivak. *Calculus on manifolds. A modern approach to classical theorems of advanced calculus*. W. A. Benjamin, New York–Amsterdam, 1965.

[13] D. Zagier. Newman's short proof of the prime number theorem. *Amer. Math. Monthly*, 104(8):705–708, 1997.