1

A Brief History of Prime

Most of the results in this book grow out of one theorem that has probably been known in some form since antiquity.

Theorem 1.1. [FUNDAMENTAL THEOREM OF ARITHMETIC] *Every integer greater than* 1 *can be expressed as a product of prime numbers in a way that is unique up to order.*

For the moment, we are using the term *prime* in its most primitive form – to mean an irreducible integer greater than one. Thus a positive integer p is prime if $p > 1$ and the factorization $p = ab$ into positive integers implies that either $a = 1$ or $b = 1$. The expression "up to order" means simply that we regard, for example, the two factorizations $6 = 2 \cdot 3 = 3 \cdot 2$ as the same.

Theorem 1.1, the Fundamental Theorem of Arithmetic, will reverberate throughout the text. The fact that the primes are the building blocks for all integers already suggests they are worth particular study, rather in the way that scientists study matter at an atomic level. In this case, we need a way of looking for primes and methods to construct them, identify them, and even quantify their appearance if possible. Some of these quests took thousands of years to fulfill, and some are still works in progress. At the end of this chapter, we will give a proof of Theorem 1.1, but for now we want to get on with our main theme.

1.1 Euclid and Primes

The first consequence of the Fundamental Theorem of Arithmetic for the primes is that there must be infinitely many of them.

Theorem 1.2. [EUCLID] *There are infinitely many primes.*

To emphasize the diversity of approaches to number theory, we will give several proofs of this famous result.

Euclid's Proof in Modern Form. If there are only finitely many primes, we can list them as $p_1, \ldots, p_r$. Let

$$N = p_1 \cdots p_r + 1 > 1.$$

By the Fundamental Theorem of Arithmetic, N can be factorized, so it must be divisible by some prime p_k of our list. Since p_k also divides $p_1 \cdots p_r$, it must divide the difference

$$N - p_1 \cdots p_r = 1,$$

which is impossible, as $p_k > 1$. □

Euler's Analytic Proof. Assume that there are only finitely many primes, so they may be listed as $p_1, \ldots, p_r$. Consider the product

$$X = \prod_{k=1}^{r} \left(1 - \frac{1}{p_k}\right)^{-1}.$$

The product is finite since 1 is not a prime and by hypothesis there are only finitely many primes. Now expand each factor into a convergent geometric series,

$$\frac{1}{1-\frac{1}{p}} = 1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \cdots .$$

For any fixed K, we deduce that

$$\frac{1}{1-\frac{1}{p}} \geqslant 1 + \frac{1}{p} + \frac{1}{p^2} + \cdots + \frac{1}{p^K}.$$

Putting this into the equation for X gives

$$\begin{aligned} X &\geqslant \left(1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^K}\right) \cdot \left(1 + \frac{1}{3} + \frac{1}{3^2} + \cdots + \frac{1}{3^K}\right) \\ &\quad \cdot \left(1 + \frac{1}{5} + \frac{1}{5^2} + \cdots + \frac{1}{5^K}\right) \cdots \left(1 + \frac{1}{p_r} + \frac{1}{p_r^2} + \cdots + \frac{1}{p_r^K}\right) \\ &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots \\ &= \sum_{n \in \mathcal{N}(K)} \frac{1}{n}, \end{aligned} \tag{1.1}$$

where

$$\mathcal{N}(K) = \{n \in \mathbb{N} \mid n = p_1^{e_1} \cdots p_r^{e_r}, e_i \leqslant K \text{ for all } i\}$$

denotes the set of all natural numbers with the property that each prime factor appears no more than K times. Notice that the identity (1.1) requires the Fundamental Theorem of Arithmetic. Given any number $n \in \mathbb{N}$, if K is large enough, then $n \in \mathcal{N}(K)$, so we deduce that

$$X \geqslant \sum_{n=1}^{\infty} \frac{1}{n}.$$

The series on the right-hand side (known as the *harmonic series*) diverges to infinity, but X is finite. Again we have reached a contradiction from the assumption that there are finitely many primes. □

Let us recall why the harmonic series diverges to infinity. As with Theorem 1.2, there are many ways to prove this; the first is elementary, while the second compares the series with an integral.

Elementary Proof. Notice that

$$\begin{aligned} 1 + \frac{1}{2} &\geqslant \frac{1}{2}, \\ \frac{1}{3} + \frac{1}{4} &\geqslant \frac{1}{2}, \\ \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} &\geqslant \frac{1}{2}, \end{aligned}$$

and so on. For any $k \geqslant 1$,

$$\frac{1}{2^k+1} + \frac{1}{2^k+2} + \cdots + \frac{1}{2^{k+1}} \geqslant 2^k \cdot \frac{1}{2^{k+1}} = \frac{1}{2}.$$

This means that

$$\sum_{n=1}^{2^{k+1}} \frac{1}{n} \geqslant \frac{k}{2} \text{ for all } k \geqslant 1,$$

and it follows that $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. □

Hidden in the last argument is some indication of the *rate* at which the harmonic series diverges. Since the sum of the first 2^{k+1} terms exceeds $k/2$, the sum of the first N terms must be approximately $C \log N$ for some positive constant C. The second proof improves on this: Equation (1.2) gives a sharper lower bound as well as an upper bound.

Exercise 1.1. Try to prove that $\sum_{n=1}^{\infty} \frac{1}{n^2}$ diverges using the same technique of grouping terms together. Of course, this will not work since this series converges, but you will see something mildly interesting. In particular, can you use this to estimate the sum?

Using the Integral Test. Compare $\sum_{n=1}^{N} \frac{1}{n}$ with the integral

$$\int_1^N \frac{1}{x}\, dx = \log N.$$

Figure 1.1 shows $\sum_{n=1}^{6} \frac{1}{n}$ trapped between $\int_0^6 \frac{1}{x+1}\, dx$ and $1 + \int_1^6 \frac{1}{x}\, dx$; in general, it follows that

$$\log(N+1) \leqslant \sum_{n=1}^{N} \frac{1}{n} \leqslant 1 + \log N. \tag{1.2}$$

This shows again that the harmonic series diverges and that the partial sum of the first N terms is approximately $\log N$.

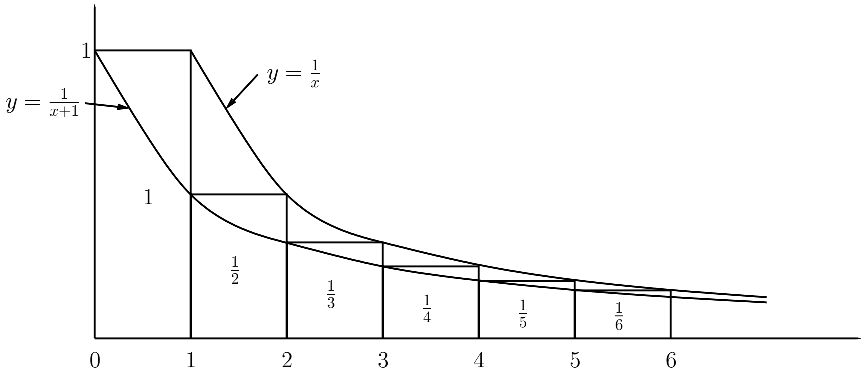

Figure 1.1. Graphs of $y = \frac{1}{x}$ and $y = \frac{1}{x+1}$ trapping the harmonic series.

□

This proof is a harbinger of more subtle results. Comparing series with integrals is a powerful technique; more generally, using *analytic* techniques to study properties of numbers has been one of the most important ideas in number theory.

Exercise 1.2. Extend the method illustrated in Figure 1.1 to show that the sequence (a_n) defined by

$$a_n = \sum_{m=1}^{n} \frac{1}{m} - \log n$$

is decreasing (that is, $a_{n+1} \leqslant a_n$ for all n) and nonnegative. Deduce that it converges to some number γ, and estimate γ to three digits. This number is known as the Euler–Mascheroni constant. It is not known if γ is rational, although it is expected not to be.

1.2 Summing Over the Primes

We begin this section with yet another proof that there are infinitely many primes. Recall that $\mathbb{P}$ denotes the set of prime numbers.

Theorem 1.3. *The series* $\displaystyle\sum_{p\in\mathbb{P}}\frac{1}{p}$ *diverges.*

Several proofs are offered; each one provides different insights. We adopt the convention that p always denotes a prime so, for example, $\displaystyle\sum_{p>N} a_p$ denotes $\displaystyle\sum_{p\in\mathbb{P},p>N} a_p$.

Notice that Theorem 1.3 tells us something about the sequence (p_n) of primes that begins $p_1 = 2$, $p_2 = 3$, $p_3 = 5, \dots$. For example, the sequence $\big(n^{1+\varepsilon}/p_n\big)$ cannot be bounded for any $\varepsilon > 0$.

First Proof of Theorem 1.3. We argue by contradiction: Assume that the series converges. Then there is some N such that

$$\sum_{p>N}\frac{1}{p} < \frac{1}{2}.$$

Let

$$Q = \prod_{p\leqslant N} p$$

be the product of all the primes less than or equal to N. The numbers

$$1 + nQ, \quad n \in \mathbb{N},$$

are never divisible by primes less than N because such primes do divide Q. Now consider

$$P = \sum_{t=1}^{\infty}\left(\sum_{p>N}\frac{1}{p}\right)^t < \sum_{t=1}^{\infty}\frac{1}{2^t} = 1.$$

We claim that

$$\sum_{n=1}^{\infty}\frac{1}{1+nQ} \leqslant \sum_{t=1}^{\infty}\left(\sum_{p>N}\frac{1}{p}\right)^t$$

because every term on the left-hand side appears on the right-hand side at least once. (Convince yourself of this claim by taking $N = 11$ and finding some terms on the right-hand side.) It follows that

$$\sum_{n=1}^{\infty}\frac{1}{1+nQ} \leqslant 1. \tag{1.3}$$

However, the series in Equation (1.3) diverges since

$$\sum_{n=1}^{K} \frac{1}{1+nQ} \geqslant \frac{1}{2Q} \sum_{n=1}^{K} \frac{1}{n}$$

for any K, and the right-hand side diverges as $K \to \infty$. This contradiction proves the theorem. □

Second Proof of Theorem 1.3. We will prove a stronger result, namely

$$\sum_{p \leqslant N} \frac{1}{p} > \log\log N - 2. \tag{1.4}$$

Fix N and let

$$\mathfrak{N}(N) = \{n \in \mathbb{N} : \text{ all prime factors of } n \text{ are less than or equal to } N\}.$$

Then (just as in Euler's analytic proof of Theorem 1.2 on p. 8)

$$\begin{aligned}\sum_{n \in \mathfrak{N}(N)} \frac{1}{n} &= \prod_{p \leqslant N} \left(1 + p^{-1} + p^{-2} + p^{-3} + \cdots\right) \\ &= \prod_{p \leqslant N} \left(1 - p^{-1}\right)^{-1}.\end{aligned}$$

If $n \leqslant N$, then certainly $n \in \mathfrak{N}(N)$, so

$$\sum_{n \leqslant N} \frac{1}{n} \leqslant \sum_{n \in \mathfrak{N}(N)} \frac{1}{n}.$$

It follows by Equation (1.2) that

$$\log N \leqslant \sum_{n \in \mathfrak{N}(N)} \frac{1}{n} = \prod_{p \leqslant N} \left(1 - p^{-1}\right)^{-1}. \tag{1.5}$$

In order to estimate the right-hand side of Equation (1.5), we need the following bound. For any $v \in [0, 1/2]$,

$$\frac{1}{1-v} \leqslant e^{v+v^2}. \tag{1.6}$$

To see why the bound (1.6) holds, let $\mathsf{f}(v) = (1-v)\exp(v+v^2)$. Then

$$\mathsf{f}'(v) = v(1-2v)\exp(v+v^2) \geqslant 0 \text{ for } v \in [0, \tfrac{1}{2}],$$

so the fact that $\mathsf{f}(0) = 1$ implies that $\mathsf{f}(v) \geqslant 1$ for all $v \in [0, 1/2]$.

For any prime p, $v = \frac{1}{p} \leqslant \frac{1}{2}$, so by the bound (1.6)

$$\prod_{p\leqslant N}\left(1-p^{-1}\right)^{-1}\leqslant\prod_{p\leqslant N}\exp\left(p^{-1}+p^{-2}\right).$$

Combining this with Equation (1.5) and taking logarithms gives

$$\log\log N\leqslant\sum_{p\leqslant N}\left(p^{-1}+p^{-2}\right). \tag{1.7}$$

Finally, we observe that

$$\sum_{p}\frac{1}{p^2}<\sum_{n=2}^{\infty}\frac{1}{n^2}<1, \tag{1.8}$$

so the contribution to the right-hand side of Equation (1.7) from $\sum_{p\leqslant N}p^{-2}$ is bounded independently of N. This completes the second proof of Theorem 1.3. □

Exercise 1.3. Prove the second inequality in Equation (1.8) using the integral test: Show that

$$\sum_{n=2}^{N}\frac{1}{n^2}<\int_2^N\frac{1}{(x-1)^2}\,dx\leqslant 1\quad\text{for all}\quad N\geqslant 2.$$

In fact, an estimate stronger than Equation (1.4) holds. Mertens showed that there is a constant A (approximately 0.261) such that

$$\sum_{p\leqslant N}\frac{1}{p}=\log\log N+A+\mathrm{O}\left(\frac{1}{\log N}\right). \tag{1.9}$$

Exercise 1.4. Is it possible to prove Equation (1.9) with $\mathrm{O}(1)$ in place of

$$A+\mathrm{O}(\frac{1}{\log N})$$

using only the methods of the second proof of Theorem 1.3?

The third proof of Theorem 1.3 extends the relationship between products such as $\prod_{p\in\mathbb{P}}\left(1-p^{-1}\right)^{-1}$ and the harmonic series to a factorization of a *function* that will later turn out to have a starring role.

Definition 1.4. *The* Riemann zeta function *is defined by*

$$\zeta(\sigma)=\sum_{n=1}^{\infty}\frac{1}{n^\sigma}$$

wherever this makes sense.

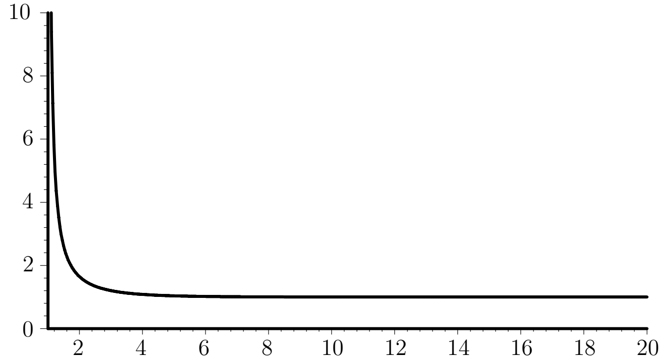

Figure 1.2. The graph of $\zeta(\sigma)$ for $1 < \sigma \leqslant 20$.

Understanding the properties of this function turns out to be the key to many deeper properties of the prime numbers. For now, we simply think of σ as being a real number and note that the series defining $\zeta(\sigma)$ converges by the integral test for $\sigma > 1$ to a positive sum and diverges at $\sigma = 1$. For $\sigma > 1$, $\zeta(\sigma)$ is a decreasing function of σ.

Viewed as a real function of a real variable, the zeta function does not look particularly subtle or useful. Figure 1.2 shows the graph of $\zeta(\sigma)$ for $1 < \sigma \leqslant 20$. Some indication of just how complicated this function really is appears when it is viewed as a complex-valued function of a complex variable. It is clear that the series defining the zeta function converges for $s = \sigma + \mathrm{i}t$ when $\sigma > 1$ (see p. 166 for more on this). Figure 1.3 shows the function $\Re(\zeta(\frac{3}{2} + \mathrm{i}t))$ for $0 \leqslant t \leqslant 60$, giving the first insight into the complex properties of the zeta function.

In Chapter 8, the Riemann zeta function is extended to a complex analytic function defined on the whole complex plane with the exception of a single pole, and this opens up the most mysterious aspect of the zeta function – its behavior along the line $\Re(s) = \frac{1}{2}$. Figure 9.1 on p. 186 gives some idea of how complicated this is.

Recall that p will be used to denote a prime number, so a product over the variable p means a product over $p \in \mathbb{P}$.

The first step in understanding the zeta function is the *Euler product representation*, which is a factorization of the zeta function into terms corresponding to primes. The idea of factorizing a function will be discussed again at the start of Chapter 9.

Theorem 1.5. [Euler Product Representation] *For any* $\sigma > 1$*,*

$$\zeta(\sigma) = \prod_{p} \left(1 - p^{-\sigma}\right)^{-1}.$$

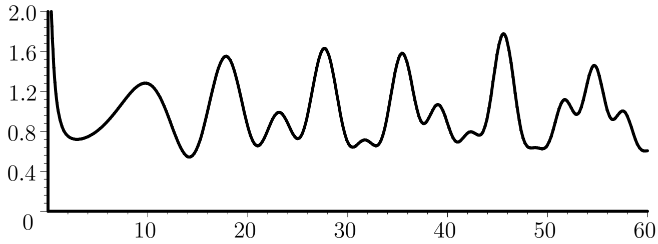

Figure 1.3. The graph of $\Re(\zeta(\frac{3}{2}+\mathrm{i}t))$ for $0 \leqslant t \leqslant 60$.

Proof. For any $\sigma > 1$,

$$\begin{aligned}\left(1-2^{-\sigma}\right)\zeta(\sigma) &= \sum_{n=1}^{\infty}\frac{1}{n^{\sigma}} - \sum_{n=1}^{\infty}\frac{1}{(2n)^{\sigma}}\\ &= \sum_{n \text{ odd}}\frac{1}{n^{\sigma}}\\ &= 1 + \sum_{p|n \Rightarrow p>2}\frac{1}{n^{\sigma}},\end{aligned}$$

where the last sum is taken over those n with all prime factors greater than 2 (that is, the odd numbers greater than 2).

Now let P be a large prime and repeat the same argument with each of the primes $3, 5, \dots, P$ in turn. This gives

$$\left(1-2^{-\sigma}\right)\left(1-3^{-\sigma}\right)\left(1-5^{-\sigma}\right)\cdots\left(1-P^{-\sigma}\right)\zeta(\sigma) = 1 + \sum_{p|n \Rightarrow p>P}\frac{1}{n^{\sigma}}.$$

The last sum ranges over those n with the property that all the prime factors of n are greater than P. Thus the last sum is a subsum of the tail of the convergent series defining $\zeta(\sigma)$, and in particular it must tend to zero as P goes to infinity. It follows that

$$\lim_{P\to\infty}\left(1-2^{-\sigma}\right)\left(1-3^{-\sigma}\right)\left(1-5^{-\sigma}\right)\cdots\left(1-P^{-\sigma}\right)\zeta(\sigma) = 1,$$

so

$$\zeta(\sigma) = \prod_{p}\left(1-p^{-\sigma}\right)^{-1}.$$

□

Remark 1.6. An infinite product is defined to be convergent if the corresponding partial products form a convergent sequence, *that does not converge to zero.* The nonzero condition is imposed to allow us to take logarithms of infinite products, thereby connecting infinite products and infinite sums in a meaningful way.

Third Proof of Theorem 1.3. Taking logarithms of the Euler product representation shows that, for any $\sigma > 1$,

$$\begin{aligned}\log \zeta(\sigma) &= -\sum_p \log\left(1 - p^{-\sigma}\right)\\ &= -\sum_p \sum_{m=1}^{\infty} \frac{-1}{mp^{m\sigma}} = \sum_p \frac{1}{p^\sigma} + \sum_p \sum_{m=2}^{\infty} \frac{1}{mp^{m\sigma}}. \qquad (1.10)\end{aligned}$$

Notice that the series involved converge absolutely, so rearrangement is permissible. For any prime p,

$$1 - \frac{1}{p^\sigma} \geqslant \frac{1}{2},$$

so

$$\begin{aligned}\sum_p \sum_{m=2}^{\infty} \frac{1}{mp^{m\sigma}} &< \sum_p \sum_{m=2}^{\infty} \frac{1}{p^{m\sigma}}\\ &= \sum_p \frac{1}{p^{2\sigma}} \frac{1}{1 - p^{-\sigma}}\\ &\leqslant 2\sum_p \frac{1}{p^{2\sigma}} \leqslant 2\zeta(2\sigma) < 2\zeta(2),\end{aligned}$$

which shows that the last double sum in Equation (1.10) is bounded. The bound $2\zeta(2)$ holds for any $\sigma \geqslant 1$, and the double sum converges for $\sigma > \frac{1}{2}$.

Thus

$$\log \zeta(\sigma) = \sum_p \frac{1}{p^\sigma} + \mathrm{O}(1).$$

The left-hand side goes to infinity as σ tends to 1 from above, so the sum on the right-hand side must do the same. □

1.3 Listing the Primes

Early in the history of the subject, Eratosthenes[1] devised a kind of sieve for listing the primes. To illustrate his method – the *sieve of Eratosthenes* – we consider the problem of finding all the primes up to 50. First arrange all the integers between 1 and 50 in a grid.

[1] Eratosthenes of Cyrene (276 B.C.–194 B.C.) was born in what is now Libya. He made major contributions to many subjects, including finding surprisingly accurate estimates for the circumference of the Earth and the distances from the Earth to the Sun and the Moon.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50

Now do the sieving: Eliminate 1, then start with 2 and cross out all numbers greater than 2 and divisible by 2. Then take the next surviving number 3 and cross out all the multiples of 3 that are greater than 3. Repeat with the next surviving number and continue until the numbers divisible by 7 are crossed out.

Exercise 1.5. Why can you stop sieving once you get to 7?

The remaining numbers are the prime numbers below 50, as shown below.

□	2	3	□	5	□	7	□	□	□
11	□	13	□	□	□	17	□	19	□
□	□	23	□	□	□	□	□	29	□
31	□	□	□	□	□	37	□	□	□
41	□	43	□	□	□	47	□	□	□

Understanding the patterns of the surviving numbers remains one of the great challenges facing mathematics two thousand years after Eratosthenes.

This method has great value, allowing people throughout history to rapidly create lists of primes. It fails to meet our longer-term objectives however. It elegantly and efficiently produces lists of primes without having to do trial divisions but does not help to decide if a given large number (with hundreds of digits, for example) is prime.

Table 1.1. Early prime hunters.

Name	Date	Bound
Pietro Cataldi	1588	750
T. Brancker	1688	100000
Felkel Kulik	1876	100330200
Derrick Henry Lehmer	1909	10006721

Table 1.1 is a short list of some of the calculations of prime tables in recent history; in each case all the primes up to the bound were listed. A rather different problem is to find exactly how many primes there are below a certain bound (without finding them all). Kulik listed the smallest factors of all the integers up to his bound and in particular found all the primes up to his bound. Lehmer's table was widely distributed and as a result was very influential (despite being shorter than Kulik's table).

1.3.1 Functions that Generate Primes.

In the seventeenth century attention turned to finding formulas that would generate the primes. Euler pointed out the following polynomial example.

Example 1.7. The polynomial $x^2 + x + 41$ yields prime values for $0 \leqslant x \leqslant 39$, but $x = 40, 41$ do not yield primes.

What is striking about this example is that it is prime for many values in succession relative to the size of the coefficients and the degree.

Exercise 1.6. (a) [Goldbach 1752] Prove that if $f \in \mathbb{Z}[x]$ has the property that $f(n)$ is prime for all $n \geqslant 1$, then f must be a constant.
(b) Extend your argument to show that if $f \in \mathbb{Z}[x]$ has the property that $f(n)$ is prime for all $n \geqslant N$ for some N, then f must be a constant.
(c) Let $P \in \mathbb{Z}[x_1, \dots, x_k]$ be a polynomial in $k \geqslant 2$ variables with integer coefficients. Define a function f by $f(n) = P(n, 2^n, 3^n, \dots, (k-1)^n)$, and assume that $f(n) \to \infty$ as $n \to \infty$. Show that $f(n)$ is composite for infinitely many values of n.

Remarkably, there is an explicit integral polynomial in several variables whose set of *positive* values as the variables run through the nonnegative integers coincides with the primes. This polynomial was discovered as a by-product of research into Hilbert's 10th Problem, which asked if there could be an algorithm to determine if a polynomial Diophantine[2] problem has a solution. However, once again, this is useless with regard to the aim of finding ways to generate primes efficiently.

There are ingenious "formulas" for the primes. Many of these require knowledge of the first $(n-1)$ primes to produce the nth prime, and none of them seem to be computationally useful. We will prove one striking result of this kind here, and two further results in Exercise 1.24 on p. 33 and in Exercise 8.9 on p. 163. The result proved here rests on Bertrand's Postulate, which is the first of many results that say something about how the prime numbers appear and how the next prime compares in size with the previous prime. The arguments below are intricate but elementary, and the basic contradiction arrived at in the proof of Theorem 1.9 is similar to one that will be used to prove Zsigmondy's Theorem (Theorem 1.15) in Section 8.3.1.

We need a lemma that says something about the growth in the product of all the primes up to n. As usual p will be used to denote a prime.

Lemma 1.8. *For any $n \geqslant 1$,*

$$\sum_{p \leqslant n} \log p < 2n \log 2. \tag{1.11}$$

[2] Diophantine problems are discussed in Chapter 2. The term is used to denote problems involving equations in which only integer solutions are sought.

Proof. Let

$$M = \binom{2m+1}{m} = \frac{(2m+1)(2m)\cdots(m+2)}{m!}.$$

This is a binomial coefficient, so it is an integer (see Exercise 1.10 for a stronger form of this). The coefficient M appears twice in the binomial expansion of $2^{2m+1} = (1+1)^{2m+1}$, so $M < 2^{2m}$. If $m+1 < p \leqslant 2m+1$ for some prime p, then p divides the numerator of M but does not divide the denominator, so

$$\prod_{p \in A(m)} p \quad \text{divides} \quad M,$$

where $A(m)$ denotes the set of primes p with $m+1 < p \leqslant 2m+1$. It follows that

$$\sum_{p \leqslant 2m+1} \log p - \sum_{p \leqslant m+1} \log p = \sum_{p \in A(m)} \log p \leqslant \log M < 2m \log 2. \tag{1.12}$$

We now prove Equation (1.11) by induction. It holds for $n \leqslant 2$, so suppose it holds for all $n \leqslant k-1$. If k is even, then

$$\sum_{p \leqslant k} \log p = \sum_{p \leqslant k-1} \log p < 2(k-1)\log 2 < 2k \log 2$$

by the inductive hypothesis. If k is odd, write $k = 2m+1$ and then

$$\begin{aligned} \sum_{p \leqslant 2m+1} \log p &= \sum_{p \leqslant 2m+1} \log p - \sum_{p \leqslant m+1} \log p + \sum_{p \leqslant m+1} \log p \\ &< 2m \log 2 + 2(m+1)\log 2 \\ &= 2(2m+1)\log 2 = 2k \log 2, \end{aligned}$$

since $m+1 < k$. Thus the inequality (1.11) holds for all n by induction. □

Theorem 1.9. [Bertrand's Postulate] *If $n \geqslant 1$, then there is at least one prime p with the property that*

$$n < p \leqslant 2n. \tag{1.13}$$

Proof. For any real number x, let $\lfloor x \rfloor$ denote the integer part of x. Thus $\lfloor x \rfloor$ is the greatest integer less than or equal to x. Let p be any prime. Then

$$\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots$$

is the largest power of p dividing $n!$ (see Exercise 8.7(a) on p. 162). Fix $n \geqslant 1$ and let

$$N = \prod_{p \leqslant 2n} p^{k(p)}$$

be the prime decomposition of $N = (2n)!/(n!)^2$. The number of times that a given prime p divides N is the difference between the number of times it divides $(2n)!$ and $(n!)^2$, so

$$k(p) = \sum_{m=1}^{\infty} \left(\left\lfloor \frac{2n}{p^m} \right\rfloor - 2 \left\lfloor \frac{n}{p^m} \right\rfloor \right), \tag{1.14}$$

and each of the terms in the sum is either 0 or 1, depending on whether $\lfloor \frac{2n}{p^m} \rfloor$ is odd or even. If $p^m > 2n$ the term is certainly 0, so

$$k(p) \leqslant \left\lfloor \frac{\log 2n}{\log p} \right\rfloor. \tag{1.15}$$

Now the proof of the theorem proceeds by a contradiction argument. Assume there is some $n \geqslant 1$ for which there is no prime satisfying the inequality (1.13), and let p be a prime factor of $N = (2n)!/(n!)^2$. Thus $p < n$ by our assumption, and $k(p) \geqslant 1$. If

$$\frac{2}{3}n < p \leqslant n$$

then

$$2p \leqslant 2n < 3p \text{ and } p^2 > \frac{4}{9}n^2 > 2n,$$

so Equation (1.14) becomes

$$k(p) = \left\lfloor \frac{2n}{p} \right\rfloor - 2 \left\lfloor \frac{n}{p} \right\rfloor = 2 - 2 = 0.$$

We deduce that $p \leqslant \frac{2}{3}n$ for every prime factor p of N. It follows that

$$\sum_{p|N} \log p \leqslant \sum_{p \leqslant 2n/3} \log p \leqslant \frac{4}{3} n \log 2 \tag{1.16}$$

by Lemma 1.8. Now if $k(p) \geqslant 2$ then by the bound (1.15),

$$2 \log p \leqslant k(p) \log p \leqslant \log 2n,$$

so $p \leqslant \sqrt{2n}$ and thus there are at most $\sqrt{2n}$ possible values of p. Hence

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

Together with the inequality (1.16), this shows that

$$\begin{aligned}\log N &\leqslant \sum_{k(p)=1} \log p \ + \sum_{k(p)\geqslant 2} k(p)\log p \\ &\leqslant \sum_{p\mid N} \log p \ + \sqrt{2n}\log 2n \\ &\leqslant \frac{4}{3}\log 2 \ + \sqrt{2n}\log 2n. \end{aligned} \tag{1.17}$$

Now N is the largest coefficient (namely the middle one) in the binomial expansion of

$$2^{2n} = (1+1)^{2n},$$

so

$$2^{2n} = 2 + \binom{2n}{1} + \binom{2n}{2} + \cdots + \binom{2n}{2n-1} \leqslant 2nN.$$

Substituting this estimate into the inequality (1.17) gives

$$2n\log 2 \leqslant \frac{4}{3}n\log 2 + \log 2n + \sqrt{2n}\log 2n. \tag{1.18}$$

It is clear that the inequality (1.18) cannot hold for large values of n; a simple calculation shows that (1.18) implies that n does not exceed 500.

It follows that if $n > 500$, then there is a prime satisfying the inequality (1.13). A calculation confirms that (1.13) also holds for all $n \leqslant 500$, completing the proof of the theorem. □

Notice that a consequence of Equation (1.13) is that if the primes are listed in order as $p_1, p_2, \dots$, then

$$p_{n+1} < 2p_n \ \text{ for all } \ n \geqslant 1. \tag{1.19}$$

It is clear that Theorem 1.9 gives another proof that there must be infinitely many primes. In each interval of the form $(n, 2n]$ there is at least one. This gives us a bound for the prime counting function

$$\pi(X) = |\{p \leqslant X \mid p \in \mathbb{P}\}.$$

The proof of Euclid's Theorem 1.2 already says a little more than the purely qualitative statement that $\pi(X) \to \infty$ as $X \to \infty$: from the proof of Theorem 1.2 we see that

$$p_{n+1} \leqslant p_1 p_2 \cdots p_n + 1.$$

This tells us something about $\pi(X)$. Define a sequence (u_n) by setting $u_1 = 2$ and $u_{n+1} = u_1 \cdots u_n + 1$ for $n \geqslant 1$. Then

$$\pi(X) \geqslant \min\{n \mid u_n \geqslant X\}.$$

This is an extremely slowly growing sequence, and the bound obtained for $\pi(X)$ is very far from the truth.

Theorem 1.9 says more: there are at least N primes in the interval

$$(1, 2^N] = (1, 2] \cup (2, 4] \cup (4, 8] \cup \cdots \cup (2^{N-1}, 2^N],$$

so $\pi(2^N) > N$. It follows that $\pi(X)$ is larger than $C \log(X)$ for some positive constant C, infinitely often. Something closer to the truth about the asymptotic behavior of $\pi(X)$ is the Prime Number Theorem (Theorem 8.1). Finding more refined estimates for $\pi(X)$ generally involves deep problems in analytic number theory. An exception is the result of Tchebychef, described in Exercise 8.7 on p. 162, which uses elementary methods to give better bounds for $\pi(X)$.

Bertrand's Postulate is enough to exhibit a striking but impractical formula for the primes. More importantly, the bound (1.13) immediately motivates the question of whether the upper estimate $2n$ could be reduced, perhaps for all large n only, and this is the subject of ongoing research.

Corollary 1.10. *There exists a real number θ with the property that*

$$\left\lfloor 2^{2^{2^{\cdot^{\cdot^{\cdot^{\theta}}}}}} \right\rfloor$$

is a prime number for any number of iterations of the exponential.

Proof. Let q_1 be any prime, and choose a sequence of primes (q_n) with the property that

$$2^{q_n} < q_{n+1} < 2^{q_n+1}. \tag{1.20}$$

This is possible by Bertrand's Postulate. Now define functions $f^{(1)}, f^{(2)}, \dots$ by $f^{(1)}(x) = \log_2(x)$ and $f^{(n+1)}(x) = \log_2(f^{(n)}(x))$ for $n \geqslant 1$. Define sequences (u_n) and (v_n) by

$$u_n = f^{(n)}(q_n) \text{ and } v_n = f^{(n)}(q_n + 1).$$

By the inequality (1.20),

$$q_n < f^{(1)}(q_{n+1}) < f^{(1)}(q_{n+1} + 1) < q_n + 1,$$

so by applying the increasing function $f^{(n)}$ we have

$$u_n < u_{n+1} < v_{n+1} < v_n.$$

It follows that the sequence (u_n) is increasing and bounded above, so it converges. Let

$$\theta = \lim_{n\to\infty} u_n.$$

Define functions $g^{(n)}$ by $g^{(1)}(x) = 2^x$ and $g^{(n+1)}(x) = 2^{g^{(n)}(x)}$ for all $n \geqslant 1$. Then

$$g^{(n)}(u_n) < g^{(n)}(\theta) < g^{(n)}(v_n),$$

so

$$q_n < g^{(n)}(\theta) < q_n + 1 \text{ for all } n \geqslant 1$$

as required. □

Exercise 1.7. [Mills] A deep result of Ingham improves Equation (1.13) to say that there is a constant C such that

$$p_{n+1} - p_n < Cp_n^{5/8}.$$

Assuming this result, modify the proof of Corollary 1.10 to show that there is a real number θ with the property that $\lfloor \theta^{3^n} \rfloor$ is a prime for all $n \geqslant 1$.

Exercise 1.8. [Richert] Use Theorem 1.9 to show that every integer greater than 6 is a sum of distinct primes. (Hint: Show this is true for the numbers 7 to 19, then use Theorem 1.9 to see that we can keep adding new primes to the set of sums obtained without missing out any integers).

Exercise 1.9. [Dressler] (a) Modify the proof of Theorem 1.9 to show that

$$p_{n+1} < 2p_n - 10 \text{ for all } n > 6.$$

(Hint: Assume there is an integer $n \geqslant 1000$ for which no prime p has the property $n < p < 2n - 10$, and consider the primes dividing $N = \binom{2n-10}{n-10}$.)
(b)*Use your result to prove that every positive integer apart from 1, 2, 4, 6 and 9 can be written as a sum of distinct odd primes.

1.3.2 Mersenne Primes

Mersenne[3] noticed that $2^2 - 1 = 3$, $2^3 - 1 = 7$, $2^5 - 1 = 31$, and $2^7 - 1 = 127$ are all primes. He suggested on the basis of experiments that $2^p - 1$ would be a prime whenever p is a prime that exceeds by 3 or less an even power of 2.

Lemma 1.11. *If $2^n - 1$ is prime, then n is prime.*

Proof. We prove the contrapositive statement that n being composite forces $2^n - 1$ to be composite. If $n = ab$ with $a, b > 1$, then

$$2^n - 1 = (2^a - 1)(2^{n-a} + 2^{n-2a} + \cdots + 2^a + 1),$$

so $2^n - 1$ is composite. □

The list of primes noticed by Mersenne does not continue uninterrupted because $2^{11}-1$ is composite. A prime of the form 2^p-1 is known as a *Mersenne*

[3] Marin Mersenne (1588–1648) was a French friar in the religious order of the Minims. He defended Descartes and Galileo against their theological critics and worked to undermine alchemy and astrology. He wrote on music as part of his studies in physics and mathematics.

prime. The next few Mersenne primes are $2^{13}-1$, $2^{17}-1$ and $2^{19}-1$. It is not known if there are infinitely many Mersenne primes. That $2^{19}-1$ is prime was known to Cataldi in 1588, and this was the largest known prime for 150 years. Fermat discovered that $2^{23}-1$ is not prime in 1640; in 1732 Euler knew that $2^{29}-1$ is not prime but that $2^{31}-1$ is prime.

It is worth pausing to say something about how this knowledge, which potentially requires the factorization of ten-digit numbers, accrued. Generally this involved a mixture of improving technique with congruences, some guile, and some heroic calculations. The first of several theoretical advances was discovered by Fermat and is now known as Fermat's Little Theorem.

Theorem 1.12. [Fermat's Little Theorem] *For any prime p and any integer a,*

$$a^p \equiv a \pmod{p}.$$

In keeping with our philosophy about differing approaches, we present two proofs of Fermat's Little Theorem.

Combinatorial Proof. It is enough to prove the statement when a is a positive integer, so we use induction. The result is true for $a=1$ because both sides are 1. Assume it is true for $a=b$. Now

$$(b+1)^p = b^p + pb^{p-1} + \cdots + pb + 1 = \sum_{j=0}^{p} \binom{p}{j} b^j$$

by the Binomial Theorem. For $0<j<p$, $\binom{p}{j} = \frac{p!}{j!(p-j)!}$ has a numerator divisible by p and denominator not divisible by p; the Fundamental Theorem of Arithmetic then shows that $\binom{p}{j}$ is divisible by p for $j=1,\dots,p-1$. So

$$(b+1)^p \equiv b^p + 1 \equiv b+1 \pmod{p}$$

by the inductive hypothesis. Thus Fermat's Little Theorem is proved. □

Exercise 1.10. Prove that the product of any n successive integers is divisible by $n!$.

A second, and often more useful, version of Fermat's Little Theorem can be written as follows. Integers a and b are said to be *coprime* if $\gcd(a,b)=1$. For all $a \in \mathbb{Z}$ that are coprime to p,

$$a^{p-1} \equiv 1 \pmod{p}. \tag{1.21}$$

This form is easily seen to be equivalent to Theorem 1.12 as follows:

$$a^p - a = a(a^{p-1}-1),$$

so when p does not divide a the Fundamental Theorem of Arithmetic shows that $p\big|(a^{p-1}-1)$ if and only if $p\big|(a^p-a)$.

The second proof of Fermat's Little Theorem proves the congruence (1.21) and uses slightly more sophisticated ideas from group theory. The virtue of this second proof is that it is quicker and (as we shall see) is better suited to generalization. It does require some properties of modular arithmetic (see Exercise 1.28 on p. 38).

Proof Using Group Theory. Work in the group $G = (\mathbb{Z}/p\mathbb{Z})^*$ of nonzero residues modulo p under multiplication. The residue of a generates a cyclic subgroup of G whose order must divide that of G by Lagrange's Theorem. Since the order of G is $(p-1)$, we deduce Equation (1.21). □

This proof is something of an anachronism: Lagrange's Theorem generalized Fermat's Little Theorem. However, thinking of residues using group theory is a powerful tool and gives rise to many more results, so it is useful to begin thinking in those terms now. Exercise 3.6 on p. 62 gives a good example where a proof using group theory can be favourably compared with a proof that only uses congruences.

Exercise 1.11. Fermat's Little Theorem says that, for any prime p, $2^{p-1}-1$ is divisible by p. It sometimes happens that $2^{p-1}-1$ is divisible by p^2. Find all the primes p with this property for $p < 10^6$. Such primes are called *Wieferich primes*, and it is not known if there are infinitely many of them.

Exercise 1.12. *A pair of congruences that arises in the Catalan problem (see p. 57) for odd primes p, q is

$$p^{q-1} \equiv 1 \pmod{q^2} \text{ and } q^{p-1} \equiv 1 \pmod{p^2}. \tag{1.22}$$

A pair of odd primes satisfying Equation (1.22) is called a *Wieferich pair.* Find all the Wieferich pairs with $p, q < 10^4$.

Exercise 1.13. An integer n is called a *perfect number* if it is equal to the sum of its proper divisors. Thus $6 = 1 + 2 + 3$ is a perfect number.
(a) If $q = 2^p - 1$ is a Mersenne prime, prove that $2^{p-1}q$ is a perfect number.
(b) Prove that if n is an *even* perfect number, then n has the form $2^{p-1}(2^p-1)$ for some prime of the form $2^p - 1$.

It is not known if there are any odd perfect numbers, but there are certainly no odd perfect numbers smaller than 10^{400}.

Write $M_n = 2^n - 1$ for the nth Mersenne number. The Mersenne numbers have special properties that make them particularly suitable for primality testing. The next result is the first of a series of results showing that divisors of M_n are quite prescribed when n is prime.

Lemma 1.13. *Suppose p is a prime and q is a nontrivial prime divisor of M_p. Then $q \equiv 1$ modulo p.*

Again, we give two proofs.

Proof Using the Euclidean Algorithm. The condition that q divides M_p amounts to

$$2^p \equiv 1 \pmod{q}.$$

By Fermat's Little Theorem, $2^{q-1} \equiv 1$ modulo q. Let $d = \gcd(p, q-1)$. If $d = p$, then $p\big|(q-1)$ as required. The only other possibility is $d = 1$ since p is prime. By Theorem 1.23 (see p. 35), in this case there are integers a and b with $1 = pa + (q-1)b$. Notice that one of a and b must be negative. Now

$$2 \equiv 2^1 \equiv 2^{pa+(q-1)b} \equiv (2^p)^a (2^{(q-1)})^b \equiv 1^a 1^b \equiv 1 \pmod{q}, \tag{1.23}$$

which is impossible as $q > 1$, so the result is proved. □

In the preceding argument, we have made use of negative exponents of expressions modulo q, but only in the form

$$1^{-a} \equiv 1 \pmod{q} \text{ for } a > 0. \tag{1.24}$$

Proof Using Group Theory. Work in the group G of nonzero residues modulo q. In this group 2 generates a cyclic subgroup whose order divides p since $2^p - 1 \equiv 0$ modulo q. Since 2 is not the identity and p is prime, the order of 2 must be p. Again, by Lagrange's Theorem, this order must divide the order of the group G, which is $(q-1)$. □

Example 1.14. Lemma 1.13 is a significant help in factorizing M_n. To see how this works, we present Fermat's proof from 1640 that $2^{23} - 1$ is not prime. If q is a prime dividing $2^{23} - 1$, then $q \equiv 1$ modulo 23. Now $23n + 1$ is a prime smaller than $\sqrt{2^{23} - 1}$ only for

$$n = 2, 12, 20, 26, 30, 36, 42, 44, 50, 56, 60, 62, 72, 84, 86, 102, 104, 110.$$

Trial division shows that M_{23} is divisible by the first of the resulting numbers, 47. In general, there is no reason to expect the smallest possible candidate to be a divisor, but even if the largest were the first such divisor, only 18 trial divisions are involved.

In 1876, Lucas discovered a test for proving the primality of Mersenne numbers. Using this test, he proved that

$$2^{127} - 1 = 170141183460469231731687303715884105727$$

is prime, but $2^{67} - 1$ is not. This disproved the suggestion of Mersenne.

The latter number occupies a special place in the history (and folklore) of mathematics. First, Lucas showed it is not prime but was not able to exhibit a nontrivial factor, which might seem a remarkable idea. In fact, it is something we will encounter again in the computational number theory sections. Second, this number was the subject of a famous talk given by Prof. F. N. Cole to the American Mathematical Society in 1903 entitled "On the Factorization of Large Numbers." On one blackboard, he wrote out the decimal expansion of $2^{67} - 1$ and on another he proceeded to compute the product of 193707721 and 761838257287, thereby showing them to be equal. The legend goes that after this silent lecture he sat down to "prolonged applause."

The specific arithmetic properties of Mersenne numbers mean that results on the primality of later terms in the sequence sometimes predated results on earlier terms. For example, $2^{127}-1$ was shown to be prime in 1876 while $2^{89}-1$ and $2^{107} - 1$ were shown to be prime in 1914.

Exercise 1.14. *[Lucas–Lehmer Test] Define an integer sequence by

$$S_1 = 4 \quad \text{and} \quad S_{n+1} = S_n^2 - 2 \quad \text{for} \quad n \geqslant 2.$$

Let p be an odd prime. Prove that $M_p = 2^p - 1$ is a prime if and only if $S_{p-1} \equiv 0$ modulo M_p.

1.3.3 Zsigmondy's Theorem

Although the proof of the conjecture that there are infinitely many Mersenne primes seems a long way off, it is known that the sequence starts to produce *new prime factors* very quickly. A prime p is a *primitive divisor* of M_n if p divides M_n but does not divide M_m for any $m < n$. Table 1.2 shows the prime factorization of M_n for $2 \leqslant n \leqslant 24$, with primitive divisors shown in bold.

The pattern that seems to emerge from Table 1.2 turns out to reflect something genuine. Sequences such as the Mersenne sequence, after a few initial terms, always have primitive divisors.

Theorem 1.15. [Zsigmondy] *Let $M_n = 2^n-1$. Then for every $n \neq 6$, $n > 1$, the term M_n has a primitive divisor.*

As seen in Table 1.2, M_6 does not have a primitive divisor, so this result is optimal. The proof of Theorem 1.15 is presented in Section 8.3.1 on p. 167, after we have proved the Möbius inversion formula (Theorem 8.15). A basic result that will be needed for the proof can be proved now, using the Binomial Theorem. Notice that this result, proved as the next exercise, already shows that the divisors of the sequence (M_n) have a special structure.

Exercise 1.15. Let p denote a prime, and for any integer N, define $\text{ord}_p(N)$ to be the exact power of p that divides N. Thus $\text{ord}_p(N) = a$ means $p^a \big| N$ but $p^{a+1} \not\big| N$.
(a) Prove that ord_p behaves like a logarithm in the sense that

$$\text{ord}_p(xy) = \text{ord}_p(x) + \text{ord}_p(y)$$

for all integers x, y.
(b) Prove that if $p \big| M_n$ then $\text{ord}_p(M_{kn}) = \text{ord}_p(M_n) + \text{ord}_p(k)$.
(c) Deduce that $\gcd(M_n, M_m) = M_{\gcd(n,m)}$ for all m, n.

Table 1.2. Primitive divisors of (M_n).

n	M_n	Factorization
2	3	$\mathbf{3}$
3	7	$\mathbf{7}$
4	15	$3 \cdot \mathbf{5}$
5	31	$\mathbf{31}$
6	63	$3^2 \cdot 7$
7	127	$\mathbf{127}$
8	255	$3 \cdot 5 \cdot \mathbf{17}$
9	511	$7 \cdot \mathbf{73}$
10	1023	$3 \cdot \mathbf{11} \cdot 31$
11	2047	$\mathbf{23} \cdot \mathbf{89}$
12	4095	$3 \cdot 5 \cdot 7 \cdot \mathbf{13}$
13	8191	$\mathbf{8191}$
14	16383	$3 \cdot \mathbf{43} \cdot 127$
15	32767	$7 \cdot 31 \cdot \mathbf{151}$
16	65535	$3 \cdot 5 \cdot 17 \cdot \mathbf{257}$
17	131071	$\mathbf{131071}$
18	262143	$3^3 \cdot 7 \cdot \mathbf{19} \cdot 73$
19	524287	$\mathbf{524287}$
20	1048575	$3 \cdot 5^2 \cdot 11 \cdot 31 \cdot \mathbf{41}$
21	2097151	$7 \cdot 127 \cdot \mathbf{337}$
22	4194303	$3 \cdot 23 \cdot 89 \cdot \mathbf{683}$
23	8388607	$\mathbf{47} \cdot \mathbf{178481}$
24	16777215	$3 \cdot 5 \cdot 7 \cdot 13 \cdot 17 \cdot \mathbf{241}$

Exercise 1.16. (a) Show that if q is a prime then every prime divisor of M_q is a primitive divisor.
(b) If M_n does not have a primitive divisor show that M_n divides the quantity

$$n \prod_{\substack{p|n,\\ p<n}} M_{n/p}.$$

(c) Deduce that for $n > 6$, every term M_n has a primitive divisor if n has only two distinct prime divisors. (Hint: take logarithms of the quantities in (b) and compare the growth rates of both sides.)
(d) What can you deduce if n has three distinct prime divisors?

Zsigmondy's Theorem holds in greater generality, though we will not prove the following result here.

Theorem 1.16. [Zsigmondy] *Let $a_n = c^n - d^n$, where $c > d$ are positive coprime integers. Then a_n always has a primitive divisor unless*

(1) $c = 2, d = 1$ *and* $n = 6$*; or*
(2) $c + d = 2^k$ *and* $n = 2$.

Exercise 1.17. Find some nontrivial examples of case (2) of the theorem.

A more general result is considered in Exercise 8.19 on p. 169.

Exercise 1.18. Prove that the sequence (u_n) does not satisfy a Zsigmondy Theorem in each of the following cases. This means that for every N there is a term u_n, $n > N$, which does not have a primitive divisor.
(a) $u_n = an + b$ for integers a and b;
(b) $u_n = n^2 + an + b$ for integers a and b with the property that the zeros of $x^2 + ax + b$ are integers;
(c)*$u_n = n^2 + an + b$ for integers a and b.

Exercise 1.19. *Can any polynomial $u_n = n^d + a_{d-1}n^{d-1} + \cdots + a_0$ for integers $a_0, \ldots, a_{d-1}$ have the property that the sequence (u_n) satisfies a Zsigmondy Theorem?

1.3.4 Mersenne Primes in the Computer Age

The arrival of electronic computers extended the limits of large Mersenne prime-hunting dramatically.

Table 1.3 is a short list showing how the size of the largest known Mersenne prime has grown over recent years; $\#M_p$ denotes the number of decimal digits in M_p. In 1978, Nickol and Noll were 18-year-old students. We do not distinguish here between a Mersenne prime that is the largest known at the time from a Mersenne prime for which all smaller Mersenne primes are known; see the references for a more detailed discussion. In Table 1.3, (G) denotes GIMPS and (P) denotes PrimeNet; these are distributed computer searches using idle time on many thousands of computers all over the world. Because of the special properties of Mersenne numbers (and related numbers of special shape), it has usually been the case that the largest explicitly known prime number is a Mersenne prime.

1.4 Fermat Numbers

Fermat noticed that the expression $F_n = 2^{2^n} + 1$ takes prime values for the first few values of n:

$$F_0 = 3, \quad F_1 = 5, \quad F_2 = 17, \quad F_3 = 257, \quad \text{and} \quad F_4 = 65537.$$

He believed the sequence might always take prime values. Euler in 1732 gave the first counterexample, when he showed that $641 \big| F_5$.

Euler, in common with Fermat and many others, was able to perform these impressive calculations through a good use of technique to minimize the amount of calculation required. Since Euler's time, many other Fermat numbers have been investigated and shown to be composite. No prime values

Table 1.3. Largest known prime values of M_p (from Caldwell's Prime Pages [25]).

p	$\#M_p$	Date	Discoverer
17	6	1588	Cataldi
19	6	1588	Cataldi
31	10	1772	Euler
61	19	1883	Pervushin
89	27	1911	Powers
107	33	1914	Powers
127	39	1876	Lucas
521	157	1952	Robinson
607	183	1952	Robinson
1279	386	1952	Robinson
2203	664	1952	Robinson
2281	687	1952	Robinson
3217	969	1957	Riesel
4253	1281	1961	Hurwitz
4423	1332	1961	Hurwitz
9689	2917	1963	Gillies
9941	2993	1963	Gillies
11213	3376	1963	Gillies
19937	6002	1971	Tuckerman
21701	6533	1978	Nickol and Noll
23209	6987	1979	Noll
44497	13395	1979	Nelson and Slowinski
86243	25962	1982	Slowinski
110503	33265	1988	Colquitt and Welsh
132049	39751	1983	Slowinski
216091	65050	1985	Slowinski
756839	227832	1992	Slowinski and Gage
859433	258716	1994	Slowinski and Gage
1257787	378632	1996	Slowinski and Gage
1398269	420921	1996	Armengaud, Woltman et al. (G)
2976221	895932	1997	Spence, Woltman et al. (G)
3021377	909526	1998	Clarkson, Woltman, Kurowski et al. (G, P)
6972593	2098960	1999	Hajratwala, Woltman, Kurowski et al. (G, P)
13466917	4053946	2001	Cameron, Woltman, Kurowski et al. (G, P)
20996011	6320430	2003	Shafer, Woltman, Kurowski et al. (G, P)
24036583	7235733	2004	Findley, Woltman, Kurowski et al. (G)

of F_n with $n > 4$ have been discovered, and it is generally expected that only finitely many terms of the sequence (F_n) are prime.

To begin, we return to Euler's result that 641 divides F_5. First, notice that $640 = 5 \cdot 2^7 \equiv -1$ modulo 641 so working modulo 641,

$$1 = (-1)^4 \equiv (5 \cdot 2^7)^4 = 5^4 \cdot 2^{28}.$$

Now $5^4 = 625 \equiv -16$ modulo 641 and $16 = 2^4$. Hence

$$1 \equiv -2^{32} \equiv -2^{2^5} \pmod{641}.$$

Of course, this elegant argument is useful only once we suspect that 641 is a factor of F_5. Euler also used some cunning to reach that point.

Lemma 1.17. *Suppose p is a prime with $p \big| F_n$. Then $p = 2^{n+1}k + 1$ for some $k \in \mathbb{N}$.*

Example 1.18. When $n = 5$, Lemma 1.17 shows that if p is a prime dividing F_5, then $p = 2^6 k + 1 = 64k + 1$ for some k. Thus the list of possible divisors is greatly reduced. We only have to test F_5 for divisibility by

$$65, 129, 193, 257, 321, 385, 449, 513, 577, 641, \dots,$$

of which $65, 129, 321, 385, 513, \dots$ are not primes. Therefore we only have to test $193, 257, 449, 577, 641, \dots$ and so on. At the fifth attempt, we find that $641 \big| F_5$.

Proof of Lemma 1.17. Suppose p is a prime with $p \big| F_n$, so $2^{2^n} \equiv -1$ modulo p and p is odd. Hence

$$2^{2^{n+1}} = (2^{2^n})^2 \equiv (-1)^2 \equiv 1 \pmod{p}.$$

Let $d = \gcd(2^{n+1}, p-1)$, and write $d = 2^{n+1}a + (p-1)b$ for integers a and b using Theorem 1.23. Just as in Equation (1.23) one of a and b will be negative, so we again use Equation (1.24) to argue that

$$2^d = 2^{2^{n+1}a + (p-1)b} \equiv (2^{2^{n+1}})^a (2^{p-1})^b \equiv 1 \pmod{p}.$$

Since $d \big| 2^{n+1}$, $d = 2^c$ for some $0 \leqslant c \leqslant n+1$ so

$$2^{2^c} = 2^d \equiv 1 \pmod{p}.$$

However, $2^{2^n} \equiv -1$ modulo p and $-1 \not\equiv 1$ modulo p, so the smallest possibility for c is $(n+1)$. Hence $d = 2^{n+1}$. On the other hand, $d \big| (p-1)$ so $p - 1 = k2^{n+1}$ as claimed. □

Exercise 1.20. Strengthen Lemma 1.17 by showing that any prime p dividing F_n must have the form $2^{n+2}k + 1$ for some $k \in \mathbb{N}$.

1.5 Primality Testing

We have covered enough ground to take a first look at the challenges thrown up by primality testing. Given a small integer, one can determine if it is prime by testing for divisibility by known small primes. This method becomes totally unfeasible very quickly. We are really trying to factorize. The ability to rapidly factorize large integers remains the Holy Grail of computational number theory. Later we will look at some more sophisticated techniques and estimate the range of integers for which they are applicable.

For now, we concentrate on properties of primes that can be used to help determine primality. Fermat's Little Theorem is an example, although it does not give a necessary and sufficient condition for primality, just a necessary one. The next result does give a necessary and sufficient condition; it is known as Wilson's Theorem because of a remark to this effect allegedly made by John Wilson in 1770 to the mathematician Edward Waring. An early proof was published by Lagrange in 1772. The theorem first seems to have been noted by al-Haytham[4] some 750 years before Wilson.

Theorem 1.19. *An integer $n > 1$ is prime if and only if*

$$(n-1)! \equiv -1 \pmod{n}.$$

PROOF OF 'ONLY IF' DIRECTION. We prove that the congruence is satisfied when n is prime and leave the converse as an exercise. Assume that $n = p$ is an odd prime. (The congruence is clear for $n = 2$.)

Each of the integers $1 < a < p-1$ has a unique multiplicative inverse distinct from a modulo p (see Corollary 1.25). Uniqueness is obvious; for distinctness, note that $a^2 \equiv 1$ modulo p implies $p \big| (a+1)(a-1)$, forcing $a \equiv \pm 1$ modulo p by primality. Thus in the product

$$(p-1)! = (p-1)(p-2)\cdots 3 \cdot 2 \cdot 1,$$

all the terms cancel out modulo p except the first and the last. Their product is clearly -1 modulo p. □

Exercise 1.21. Prove the converse: If $n > 1$ and $(n-1)! \equiv -1$ modulo n, then n is prime.

Exercise 1.22. [GAUSS] Prove the following generalization of Theorem 1.19. Let

$$P_n = \prod_{\substack{m<n,\\ \gcd(m,n)=1}} m$$

be the product of all positive integers less than n and coprime to n. Then P_n+1 is divisible by n if n is equal to 4, p^k, or $2p^k$ for some odd prime p, and $P_n - 1$ is divisible by n if n is not of that form.

[4] Abu Ali al-Hasan ibn al-Haytham (964–1040) lived in Persia and Egypt. He is most famous for *Alhazen's Problem*: Find the point on a spherical mirror where a light will be reflected to an observer. In number theory, in addition to proving what we often call Wilson's Theorem, al-Haytham worked on perfect numbers (see Exercise 1.13).

Exercise 1.23. [Clement] (a) Use al-Haytham's Theorem (Theorem 1.19) to prove that, for $n > 1$, n and $n+2$ are both prime if and only if

$$4\big((n-1)!+1\big)+n \equiv 0 \pmod{n(n+2)}.$$

(b) Prove that, for $n > 13$, the triple n, $n+2$, and $n+6$ are all prime if and only if

$$4320\Big(4\big((n-1)!+1\big)+n\Big)+361n(n+2) \equiv 0 \bmod \big(n(n+2)(n+6)\big).$$

(c) Find a similar characterization of prime triples of the form n, $n+4$, and $n+6$.

Primes p for which $p+2$ is also a prime are called *twin primes*, and it is a long-standing conjecture that there are infinitely many twin primes. A remarkable result of Brun from 1919 is that the reciprocals of the twin primes (whether there are infinitely many or not) are summable:

$$\sum_{p,p+2\in\mathbb{P}} \frac{1}{p} = B < \infty. \tag{1.25}$$

Numerical estimation of *Brun's constant* B is very difficult.

Exercise 1.24. Theorem 1.19 gives another 'formula' for the primes. Show that $(n-2)!$ is congruent to 1 or 0 modulo n depending on whether n is prime or not, for $n \geqslant 3$.
(a) Deduce that the prime counting function $\pi(X) = |\{p \in \mathbb{P} \mid p \leqslant X\}|$ may be written

$$\pi(X) = 1 + \sum_{j=3}^{X}\left((j-2)! - j\left\lfloor\frac{(j-2)!}{j}\right\rfloor\right), \; X \geqslant 3,$$

with $\pi(1) = 0$, $\pi(2) = 1$.
(b) Define a function f by $f(x,x) = 0$ and

$$f(x,y) = \frac{1}{2}\left(1 + \frac{x-y}{|x-y|}\right) \quad \text{for} \quad x \neq y.$$

Use Theorem 1.9 to prove that

$$p_n = 1 + \sum_{j=1}^{2^n} f(n, \pi(j)).$$

In principle, Theorem 1.19 seems to offer a general primality test because the condition is necessary and sufficient. The problem is that in practice it is impossible to compute $(n-1)!$ modulo n in a reasonable amount of time for any integer that is not quite small. In Chapter 12 we will seek to give a better understanding of what counts as “small” or “large” in terms of modern computing.

Fermat’s Little Theorem offers another hope. Taking $a = 2$, Fermat’s Little Theorem implies that

$$2^{p-1} \equiv 1 \pmod{p} \text{ whenever } p \text{ is prime.} \tag{1.26}$$

At various times in history, it has been thought that a kind of converse might be true: If n is odd and $2^{n-1} \equiv 1$ modulo n, might it follow that n is prime? Calculations tend to support this, and for $n < 341$ this does indeed successfully detect primality.

Example 1.20. Testing the congruence $2^{n-1} \equiv 1$ modulo n fails to detect the fact that $n = 341 = 11 \cdot 31$ is composite. By Fermat’s Little Theorem, $2^{10} \equiv 1$ modulo 11 so $2^{340} \equiv 1^{34} \equiv 1$ modulo 11. Also $2^5 = 32 \equiv 1$ modulo 31, so

$$2^{340} = (2^5)^{68} \equiv 1^{68} = 1 \pmod{31}.$$

Thus $2^{340} - 1$ is divisible by the coprime numbers 11 and 31, and hence by their product 341, so $2^{340} \equiv 1$ modulo 341.

However, Fermat’s Little Theorem says more than Equation (1.26): It gives the congruence

$$a^{p-1} \equiv 1 \pmod{p}$$

for *any base* a, not just $a = 2$. Taking $a = 3$ in Example 1.20, we quickly find

$$3^{340} \equiv 56 \pmod{341},$$

which contradicts Fermat’s Little Theorem with $a = 3$, showing that 341 cannot be prime. Notice the recurrence of a phenomenon encountered before: Using $a = 3$, we have shown that a number is not prime without exhibiting a nontrivial factor.

This method suggests the following as a primality test. Given an integer n, choose numbers a at random with $1 < a < n$ and test to see if $a^{n-1} \equiv 1$ modulo n. If not, then n is definitely composite. If the congruence is satisfied for several such a, we might view this as compelling evidence that n must be prime. Unfortunately, this also fails as a primality test.

Exercise 1.25. Prove that $n = 561$ is a composite number that satisfies Fermat’s Little Theorem for every possible base by showing that $a^{560} \equiv 1$ modulo 561 for every a, $1 < a < n$ with $\gcd(a, 561) = 1$. (Hint: Use Fermat’s Little Theorem on each of the factors 3, 11, and 17 of 561.)

A composite integer that satisfies the congruence of Fermat’s Little Theorem for all bases coprime to itself is known as a *Carmichael number*; these will be discussed in more detail in Section 12.5. It was not known whether there are infinitely many Carmichael numbers until 1994, when Alford, Granville, and Pomerance not only proved that there are infinitely many but gave some measure of how many there are asymptotically. The existence of infinitely many Carmichael numbers renders the test based on Fermat's Little Theorem test too unreliable. Later, we will see however that a more sophisticated version is salvageable as a primality test.

1.6 Proving the Fundamental Theorem of Arithmetic

We uncover Euclid's real genius once we try to prove the Fundamental Theorem of Arithmetic. There are two parts to it: existence and uniqueness. The existence part is not difficult. Let $n > 1$ be an integer, and choose r with $2^r > n$. If n itself is not divisible by any a with $1 < a < n$, then nothing else needs to be said. Otherwise, we can write $n = ab$ with $1 < a, b < n$. Again, if a and b cannot be factorized, further then we are done. If this is not the case then at least one of them can be factorized. Once we have done this r times, we have $n = a_1 \cdots a_r$ with each $1 < a_i < n$. This implies $n \geqslant 2^r$, giving a contradiction. Thus n must be a product of no more than r prime factors.

It is when we come to the uniqueness part of the proof that we uncover a subtlety – namely, that the definition of prime as an irreducible element is not really adequate to prove the Fundamental Theorem of Arithmetic. Suppose we try to argue as follows: Consider two factorizations for n into primes, say

$$p_1 \cdots p_r = n = q_1 \cdots q_s.$$

We would like to say that because p_1 divides the right-hand side, it must divide one of the q_i. However, if we are working with the definition of *prime* as *irreducible*, then we need a result that tells us that being irreducible forces this divisibility property. Such a result may be found using the Euclidean Algorithm.

Later, we will see examples in rings that are closely related to $\mathbb{Z}$ whose elements have genuinely different factorizations into irreducibles.

Exercise 1.26. Let

$$A = \{n \in \mathbb{N} \mid n \equiv 1 \pmod 4\},$$

and call $n \neq 1$ an *A-prime* if the only divisors of n in A are 1 and n.
(a) Show that every element of A except 1 factorizes as a finite product of A-primes.
(b) Show that this factorization into A-primes is not unique.

1.6.1 The Euclidean Algorithm

Given $a, b > 0$ in $\mathbb{Z}$, we can always find q and r with $a = bq + r$ and $0 \leqslant r < b$. Indeed, for q we can simply take the integer part $\lfloor a/b \rfloor$ of a/b and then show that by defining $r = a - bq$ we must have $0 \leqslant r < b$.

Something very interesting happens when we iterate this process. It will help to define $q = q_1$ and $r = r_1$ and continue to find quotients and remainders as follows:

$$\begin{aligned} a &= bq_1 + r_1, & 0 &\leqslant r_1 < b\\ b &= r_1 q_2 + r_2, & 0 &\leqslant r_2 < r_1\\ &\vdots & &\vdots\\ r_{n-3} &= r_{n-2}q_{n-1} + r_{n-1}, & 0 &\leqslant r_{n-1} < r_{n-2}\\ r_{n-2} &= r_{n-1}q_n + r_n, & 0 &\leqslant r_n < r_{n-1}\\ r_{n-1} &= r_n q_{n+1} + 0. \end{aligned}$$

The sequence of remainders is decreasing and each term is nonnegative, so the sequence must terminate. We have written r_n for the last nonzero remainder, so $r_n \big| r_{n-1}$. We claim that r_n is the greatest common divisor of a and b.

Example 1.21. Let $a = 17$ and $b = 11$. Then the Euclidean Algorithm gives the equations

$$\begin{aligned} 17 &= 11 \cdot 1 + 6,\\ 11 &= 6 \cdot 1 + 5,\\ 6 &= 5 \cdot 1 + 1,\\ 5 &= 1 \cdot 5 + 0. \end{aligned}$$

The last nonzero remainder is the greatest common divisor of 17 and 11, which is clearly 1.

To prove that $r_n = \gcd(a, b)$, we need a better notion of greatest common divisor than the intuitive one.

Definition 1.22. *If a and b in $\mathbb{Z}$ are not both zero, d is said to be a* greatest common divisor *of a and b if*

(1) *$d \big| a$ and $d \big| b$; and*
(2) *if d' is any number with $d' \big| a$ and $d' \big| b$, then $d' \big| d$.*

The first condition says d is a common divisor of a and b, while the second says it is the greatest such divisor.

Note that we say "a" greatest common divisor rather than "the" greatest common divisor because if d satisfies this condition then $-d$ will also satisfy the definition. If we work in $\mathbb{N}$, then the greatest common divisor will be unique. The notation $\gcd(a, b)$ denotes the unique nonnegative greatest common divisor of a and b. If $\gcd(a, b) = 1$, then we will call a and b *coprime.*

Exercise 1.27. Using Definition 1.22, show that $r_n = \gcd(a, b)$. (Hint: Work your way up and then down the chain of equations to verify the two properties.)

The next result is fundamental to the structure of the integers; it is an easy consequence of the Euclidean Algorithm and is sometimes referred to as Bezout's Lemma.

Theorem 1.23. *If $d = \gcd(a, b)$ with $a, b \in \mathbb{Z}$ not both zero, then there are numbers $x, y \in \mathbb{Z}$ with*

$$d = ax + by. \tag{1.27}$$

Proof. The idea is to work your way up the chain of equations in the Euclidean Algorithm, always expressing the remainder in terms of the previous two remainders. Writing $*$ for an integer, we get

$$\begin{aligned}
\gcd(a, b) = r_n &= r_{n-2} - r_{n-1}q_n \\
&= r_{n-2}(1 + q_n q_{n-1}) - r_{n-3}q_n \\
&= r_{n-3} \cdot * + r_{n-4} \cdot * \\
&\vdots \\
&= b \cdot * + r_1 \cdot * \\
&= a \cdot * + b \cdot *.
\end{aligned}$$

□

Example 1.24. Using the equations from Example 1.21 we find that

$$\begin{aligned}
1 &= 6 - 5 \\
&= 6 - (11 - 6) \\
&= 2 \cdot 6 - 11 \\
&= 2(17 - 11) - 11 \\
&= 2 \cdot 17 - 3 \cdot 11.
\end{aligned}$$

Corollary 1.25. *Let $n > 1$ and a denote elements of $\mathbb{Z}$. Then a and n are coprime if and only if there exists x with*

$$ax \equiv 1 \pmod{n}.$$

That is, $\gcd(a, n) = 1$ if and only if a is invertible modulo n.

Proof. The congruence is equivalent to the existence of an integer y with

$$ax + ny = 1.$$

If a and n have a factor in common then that factor will also divide 1, so the congruence implies a and n are coprime. Conversely, if a and n are coprime then 1 is a greatest common divisor of a and n so we can use Theorem 1.23 to see that there are integers x and y with $ax + ny = 1$, which translates into the congruence. □

Exercise 1.28. Let p be a prime. Prove that the set $(\mathbb{Z}/p\mathbb{Z})^*$ of nonzero elements in $\mathbb{Z}/p\mathbb{Z}$ forms a group under multiplication modulo p.

One of the remarkable things about the Euclidean Algorithm is that it finds the greatest common divisor of two integers without factorizing either of them. We will see later how this has been exploited in powerful ways by computational number theory in recent years.

Exercise 1.29. Prove the Fundamental Theorem of Arithmetic using Theorem 1.23. (Hint: This is done in greater generality on p. 47.)

1.6.2 An Inductive Proof of Theorem 1.1

We wish to prove that any natural number n has a decomposition $n = p_1 \cdots p_r$ into primes uniquely up to rearrangement of the prime factors.

For $n = 2$, the theorem is clearly true. We proceed by induction. Suppose that the Fundamental Theorem of Arithmetic holds for all natural numbers strictly less than some $a > 1$. We want to deduce the Fundamental Theorem of Arithmetic for a. Let

$$D = \{d \mid d > 1, d\big|a\}$$

denote the set of non-identity divisors of a. The set D is nonempty since it contains a, so it has a smallest element, which we denote p. This smallest element must be a prime because if it had a nontrivial divisor that would be a smaller element of D. Thus we have a decomposition

$$a = pb, p \text{ prime}, b < a.$$

Since $b < a$, by the inductive hypothesis, the Fundamental Theorem of Arithmetic holds for b, so there is a prime decomposition

$$b = p_1 \cdots p_s$$

into primes uniquely up to rearrangement. It follows that

$$a = p \cdot p_1 \cdots p_s$$

is a prime decomposition of a, and a has no other prime decomposition *involving the prime p.*

Suppose that a has another prime decomposition,

$$a = q_1 \cdots q_r,$$

in which the prime p does not appear. In particular, $q_1 \neq p$. Moreover, by the definition of p, $q_1 > p$ since $q_1 \in D$, $1 \leqslant q_1 - p < q_1$. Let $c = q_2 \cdots q_r$, and define

$$a_0 = a - pc = p(b - c) = (q_1 - p)c. \tag{1.28}$$

Now $1 \leqslant a_0 < a$ and the divisors $(b-c), (q_1 - p)$, and c are all less than a. By the inductive hypothesis, the numbers a_0, $(b-c)$, $(q_1 - p)$, and c all have unique prime decompositions. By Equation (1.28), the prime p must appear in any prime decomposition of a_0 and therefore (by uniqueness) must also appear in the decomposition of $(q_1 - p)$ or that of c.

Now p cannot appear in a prime decomposition of $(q_1 - p)$ because that would require $p \big| q_1$, which is impossible, as p and q_1 are distinct primes. Nor can p appear in a prime decomposition of $c = q_2 \cdots q_r$ by assumption. Thus the assumption of a second prime decomposition for a leads to a contradiction, completing the proof of the Fundamental Theorem of Arithmetic.

1.7 Euclid's Theorem Revisited

In this section, three further proofs of Theorem 1.2 are given, each interesting and suggestive in its own right.

1.7.1 What Did Euclid Really Prove?

First, we return to the master's proof. The following is a translation of Euclid's proof taken from Joyce's Web translation of Euclid's *Elements*. In Euclid's time, numbers were thought of as relatively concrete lengths of line segments. Thus, for example, a number A *measures* a number B if a stick of length A could be used to fit into a stick of length B a whole number of times. In modern terminology, A divides B. We start with Euclid's Theorem in (an approximation of) Euclid's language:

Οἱ πρῶτοι ἀριθμοὶ πλείους εἰσὶ παντὸς τοῦ
προτεθέντος πλήθους πρώτων ἀριθμῶν.

A translation of this is the following theorem, which is Proposition 20 of Book IX in Euclid's *Elements*.

Theorem 1.26. *The prime numbers are more than any assigned multitude of prime numbers.*

Proof. Let A, B, and C be the assigned prime numbers. I say that there are more prime numbers than A, B, and C. Take the least number DE measured by A, B, and C. Add the unit DF to DE.

Then EF is either prime or not.

First, let it be prime. Then the prime numbers A, B, C, and EF have been found, which are more than A, B, and C.

Next, let EF not be prime. Therefore, it is measured by some prime number. Let it be measured by the prime number G. I say that G is not the same as any of the numbers A, B, and C.

If possible, let it be so.

Now A, B, and C measure DE, and therefore G also measures DE. But it also measures EF. Therefore G, being a number, measures the remainder, the unit DF, which is absurd.

Therefore G is not the same as any one of the numbers A, B, and C, and by hypothesis it is prime. Therefore, the prime numbers A, B, C, and G have been found, which are more than the assigned multitude of A, B, and C. Therefore, prime numbers are more than any assigned multitude of prime numbers. □

There is little between this argument and Euclid's proof in modern form on p. 8. Euclid did not have our modern notion of infinity, so he proved that there are more primes than any prescribed number. He also often stated proofs using examples (in this case, what he really proves is that there are more than three primes), but it is clear he understood the general case. It is possible that part of the reason for this is the notational difficulties involved in dealing with arbitrarily large finite lists of objects.

1.7.2 A Topological Proof of Theorem 1.2

In 1955, Furstenberg gave a completely different type of proof of the infinitude of the primes using ideas from topology.

Furstenberg's Topological Proof of Theorem 1.2. Define a topology on the integers $\mathbb{Z}$ by taking as a basis the arithmetic progressions. For each prime p, let S_p denote the arithmetic progression $p\mathbb{Z}$. Since

$$S_p = \mathbb{Z}\backslash\big((p\mathbb{Z}+1) \cup \cdots \cup (p\mathbb{Z}+(p-1))\big),$$

the set S_p is the complement of an open set, and thus is closed. Let $S = \bigcup_p S_p$ be the union of all the sets S_p as p varies over the primes. If there are only finitely many primes, then S is a finite union of closed sets, and thus is closed. However, every integer except ±1 is in some S_p, so the complement of S is $\{1, -1\}$, which is clearly not open. It follows that S cannot be closed and therefore cannot be a finite union, so there must be infinitely many primes. □

In contrast with the other proofs of Theorem 1.2, this is qualitative – all it tells us about the prime counting function is that $\pi(X) \to \infty$ as $X \to \infty$.

1.7.3 Goldbach's Proof

Goldbach showed how one may use a sequence of integers with the property that an infinite subsequence are pairwise coprime to give a different proof.

Goldbach's Proof of Theorem 1.2. We claim that the Fermat numbers $F_n = 2^{2^n}+1$ are pairwise coprime:.

$$m \neq n \implies \gcd(F_m, F_n) = 1. \tag{1.29}$$

The first step is to show by induction that

$$F_m - 2 = F_0 F_1 \cdots F_{m-1} \text{ for all } m \geqslant 1. \tag{1.30}$$

To see why this is true, first note that $F_1 - 2 = F_0$ and assume that Equation (1.30) holds for $m \leqslant k$. Then

$$\begin{aligned} F_0 F_1 \cdots F_{k-1} F_k &= (F_k - 2) F_k \\ &= \left(2^{2^k} - 1\right)\left(2^{2^k} + 1\right) \\ &= 2^{2^{k+1}} - 1 = F_{k+1} - 2, \end{aligned}$$

showing Equation (1.30) by induction. Thus for $m > n$,

$$d \big| F_m, d \big| F_n \implies d \big| F_m - 2 \implies d \big| 2,$$

which forces d to be 1 since all the F_n are odd numbers. This proves Equation (1.29).

This in turn means there must be infinitely many primes. By Theorem 1.1, each F_n has a prime factor p_n, say, and by Equation (1.29) these are all distinct. □

The proof using Fermat numbers actually does a little more than prove there are infinitely many primes. It also gives some insight into how many primes there are that are smaller than a given number. By the time we reach the number F_n, we must have seen at least n different primes, so

$$\pi(X) \geqslant \frac{1}{\log 2} \log\left(\frac{\log(X-1)}{\log 2}\right),$$

which is approximately proportional to $\log\log X$. This is far weaker than the remark on p. 21.

Notes to Chapter 1: The exact history of Theorem 1.1 is not clear, and it is likely that it was known and used long before it was explicitly stated. The earliest precise formulation and proof seems to be due to Gauss [67], but it could be argued that Euclid certainly knew that if a prime p divides a product ab, then p must divide a or b, and that his geometrical formalism and approach to exposition did not require him to consider products of more than three terms (see Section 1.7.1). Many of the proofs of Euclid's Theorem are featured in the Prime Pages Web site [25]; Ribenboim's book [125] describes no fewer than 11 proofs. Example 1.7 is related to subtle problems in algebraic number theory; see Ribenboim's book [125] for a discussion and detailed references. That the positive values of a polynomial in several variables could coincide with the primes is essentially a by-product of Matijasevič's solution to one of Hilbert's famous problems. Some of the history and references and two explicit polynomials are given in accessible form in the paper [85] of Jones, Sato, Wada and Wiens. The proofs of Lemma 1.8 and Theorem 1.9 are those of

Erdös [51] and Kalmar, and may be found in Hardy and Wright [75]; that of Corollary 1.10 follows a survey paper of Dudley [46]. Bertrand's Postulate (Theorem 1.9) was first proved by Tchebychef [151, Tome I, pp. 49–70, 63]. He also proved that for any $e > \frac{1}{5}$, there is a prime between x and $(1+e)x$ for x sufficiently large. The deep result of Ingham [80] has been improved a great deal — for example, Baker, Harman and Pintz [8] have shown that there is a prime in the interval $[x - x^{0.525}, x]$ for x sufficiently large. Exercise 1.7 is due to Mills [107]. Exercise 1.8 comes from a paper of Richert [127]; Exercise 1.9 from a paper of Dressler [45]. Further material on Mersenne primes – and on large primes in general – may be found on Caldwell's Prime Pages Web site [25]; Table 1.3 is taken from his Web site. A recent account of the GIMPS record-breaking prime is in Ziegler's short article [167]. Zsigmondy's Theorems 1.15 and 1.16 appeared first in his paper [168]; a more accessible proof may be found in a short paper by Roitman [132]. Deep recent work has extended this to a larger class of sequences: Bilu, Hanrot and Voutier have shown that for $n > 30$ the nth term of any Lucas or Lehmer sequence has a primitive divisor in their paper [15]. The current status of Fermat numbers and their factorization may be found on Keller's Web site [88]. Parts of the intricate connection between group theory and the origins of modern number theory, and in particular a discussion of how Gauss used group-theoretic concepts long before they were formalized, are in a paper of Wußing [164]. For more on the very special numbers found in Exercise 1.11 see Ribenboim's popular article [123]. The inductive proof of Theorem 1.1 in Section 1.6.2 is taken from Hasse's classic text [76] and is attributed there to Zermelo. Hasse's text is also the source of the statement of Euclid's Theorem in Greek in Section 1.7.1. We thank David Joyce for permission to use the translation in Section 1.7 from his Web site [86]; this Web site is based on several translations of Euclid's work, but the primary and most accessible source remains the translation by Heath [53]. Exercise 1.24 is taken from Hardy and Wright [75]. Furstenberg's proof of Euclid's Theorem appeared in [63]. Exercise 1.23 is taken from Clement's paper [31]. Brun's result in Equation (1.25) appeared originally in his paper [24]; a modern proof may be found in the book of LeVeque [100]. Finally, we make some remarks concerning Section 1.7.2. Using topology in this setting might seem odd, but perhaps Euler's proof using the harmonic series seemed odd when it first appeared. We don't wish to stretch the point, but it could just be that Furstenburg's proof points forward to new ways of looking at arithmetic in just the same way as Euler's did. Profound structures in the integers have certainly been uncovered using methods from ergodic theory, combinatorics, functional analysis, and Fourier analysis; see a survey paper of Bergelson [11], the book by Furstenberg [64], and a new approach in a paper of Gowers [72] for some of these startling results. In a similar vein, Green and Tao [73] have recently proved the deep result that the primes contain arbitrarily long arithmetic progressions.