Foreword

> The gold in 'them there hills' is not always buried deep. Much of it is within easy reach. Some of it is right on the surface to be picked up by any searcher with a keen eye for detail and an eagerness to explore. As in any treasure hunt, the involvement grows as the hunt proceeds and each success whether small or great adds the fuel of excitement to the exploration. – A. E. Ross

Number theory is one of the few areas of mathematics where problems of substantial interest can be described to someone possessing scant mathematical background. It sometimes proves to be the case that a problem which is simple to state requires for its resolution considerable mathematical preparation; e.g., this appears to be the case for Fermat's conjecture regarding integer solutions to the equation $x^n + y^n = z^n$. But this is by no means a universal phenomenon; many engaging problems can be successfully attacked with little more than one's "mathematical bare hands". In this case one says that the problem can be solved in an *elementary* way (even though the elementary solution may be far from simple). Such elementary methods and the problems to which they apply are the subject of this book.

Because of the nature of the material, very little is required in terms of prerequisites: The reader is expected to have prior familiarity with number theory at the level of an undergraduate course. The nececssary background can be gleaned from any number of excellent texts, such as Sierpiński's charmingly discursive *Elementary Theory of Numbers* or LeVeque's lucid and methodical *Fundamentals of Number Theory*. Apart from this, a rigorous course in calculus, some facility with manipulation of estimates (in

particular, big-Oh and little-oh notation), and a first course in modern algebra (covering groups, rings, and fields) should suffice for the majority of the text. A course in complex variables is *not* required, provided that the reader is willing to overlook some motivational remarks made in Chapter 7.

Rather than attempt a comprehensive account of elementary methods in number theory, I have focused on topics that I find particularly attractive and accessible:

- Chapters 1, 3, 4, and 7 collectively provide an overview of prime number theory, starting from the infinitude of the primes, moving through the elementary estimates of Chebyshev and Mertens, then the theorem of Dirichlet on primes in prescribed arithmetic progressions, and culminating in an elementary proof of the prime number theorem.
- Chapter 2 contains a discussion of Gauss's arithmetic theory of the roots of unity (*cyclotomy*), which was first presented in the final section of his *Disquisitiones Arithmeticae*. After developing this theory to the extent required to prove Gauss's characterization of constructible regular polygons, we give a cyclotomic proof of the quadratic reciprocity law, followed by a detailed account of a little-known cubic reciprocity law due to Jacobi.
- Chapter 5 is a 12-page interlude containing Dress's proof of the following result conjectured by Waring in 1770 and established by Hilbert in 1909: For each fixed integer $k \geq 2$, every natural number can be expressed as the sum of a bounded number of nonnegative kth powers, where the bound depends only on k.
- Chapter 6 is an introduction to combinatorial sieve methods, which were introduced by Brun in the early twentieth century. The best-known consequence of Brun's method is that if one sums the reciprocals of each prime appearing in a twin prime pair $p, p+2$, then the answer is finite. Our treatment of sieve methods is robust enough to establish not only this and other comparable 'upper bound' results, but also Brun's deeper "lower bound" results. For example, we prove that there are infinitely many n for which both n and $n+2$ have at most 7 prime factors, counted with multiplicity.
- Chapter 8 summarizes what is known at present about *perfect numbers*, numbers which are the sum of their proper divisors.

At the end of each chapter (excepting the interlude) I have included several nonroutine exercises. Many are based on articles from the mathematical literature, including both research journals and expository publications like the *American Mathematical Monthly*. Here, as throughout the text, I have

made a conscious effort to document original sources and thus encourage conformance to Abel's advice to "read the masters".

While the study of elementary methods in number theory is one of the most accessible branches of mathematics, the lack of suitable textbooks has been a repellent to potential students. It is hoped that this modest contribution will help to reverse this injustice.

Paul Pollack

Notation

While most of our notation is standard and should be familiar from an introductory course in number theory, a few of our conventions deserve explicit mention: The set $\mathbf{N}$ of natural numbers is the set $\{1, 2, 3, 4, \dots\}$. Thus 0 is *not* considered a natural number. Also, if $n \in \mathbf{N}$, we write "$\tau(n)$" (instead of "$d(n)$") for the number of divisors of n. This is simply to avoid awkward expressions like "$d(d)$" for the number of divisors of the natural number d. Throughout the book, we reserve the letter p for a prime variable.

We remind the reader that "$A = O(B)$" indicates that $|A| \le c|B|$ for some constant $c > 0$ (called the *implied constant*); an equivalent notation is "$A \ll B$". The notation "$A \gg B$" means $B \ll A$, and we write "$A \asymp B$" if both $A \ll B$ and $A \gg B$. If A and B are functions of a single real variable x, we often speak of an estimate of this kind holding as "$x \to a$" (where a belongs to the two-point compactification $\mathbf{R} \cup \{\pm\infty\}$ of $\mathbf{R}$) to mean that the estimate is valid on some deleted neighborhood of a. Subscripts on any of these symbols indicate parameters on which the implied constants (and, if applicable, the deleted neighborhoods) may depend. The notation "$A \sim B$" means $A/B \to 1$ while "$A = o(B)$" means $A/B \to 0$; here subscripts indicate parameters on which the rate of convergence may depend.

If S is a subset of the natural numbers $\mathbf{N}$, the (*asymptotic*, or *natural*) *density* of S is defined as the limit

$$\lim_{x \to \infty} \frac{1}{x} \#\{n \in S : n \le x\},$$

provided that this limit exists. The *lower density* and *upper density* of S are defined similarly, with lim inf and lim sup replacing lim (respectively). We say that a statement holds for *almost all natural numbers* n if it holds on a subset of $\mathbf{N}$ of density 1.

If f and G are defined on a closed interval $[a, b] \subset \mathbf{R}$, with f' piecewise continuous there, we define

$$\int_a^b f(t)\, dG(t) := G(b)f(b) - G(a)f(a) - \int_a^b f'(t)G(t)\, dt, \tag{0.1}$$

provided that the right-hand integral exists. (Experts will recognize the right-hand side as the formula for integration by parts for the Riemann–Stieltjes integral, but defining the left-hand side in this manner allows us to avoid assuming any knowledge of Riemann–Stieltjes integration.) We will often apply partial summation in the following form, which is straightforward to verify directly: *Suppose that a and b are real numbers with $a \leq b$ and that we are given complex numbers a_n for all natural numbers n with $a < n \leq b$. Put $S(t) := \sum_{a<n\leq t} a_n$. If f' is piecewise continuous on $[a, b]$, then*

$$\sum_{a<n\leq b} a_n f(n) = \int_a^b f(t)\, dS(t).$$

In order to paint an accurate portrait of the mathematical landscape without straying off point, it has been necessary on occasion to state certain theorems without proof; such results are marked with a star (★). For some of these results, proofs are sketched in the corresponding chapter exercises.

Acknowledgements

There are many people without whom this book could not have been written and many others without whom this book would not be worth reading.

Key members of the first group include my middle and high-school teachers Daniel Phelon, Sharon Bellak, and Jeff Miller. It is thanks to their tireless efforts that I was prepared to attend the Ross Summer Mathematics Program at Ohio State University in 1998. There Arnold Ross, assisted by my counselor Noah Snyder and my seminar instructor Daniel Shapiro, impressed upon me the importance of grappling with mathematical ideas for oneself. I regard this as the most important lesson I have learned so far on my mathematical journey. As an undergraduate, I was the fortunate recipient of generous mentoring from Andrew Granville and Matt Baker, and I had the privilege of attending A. J. Hildebrand's 2002 REU in number theory. My subsequent graduate experience at Dartmouth College ranks as one of the happiest times of my life, due in large measure to the wise guidance of my advisor, Carl Pomerance.

My family — my father Lawrence, my mother Lolita, and my brother Michael — has done so much for me over the years that it would be impossible (and inappropriate!) for me to express the extent of my appreciation in this brief space. Another friend for whom I am grateful beyond words is Susan Roth, who for the last decade has accompanied me on many of my (mis)adventures in genre television.

Mits Kobayashi cheerfully donated his time to prepare many of the figures included in the text. Both he and Enrique Treviño pointed out several

typographical errors and inaccuracies in earlier versions of the manuscript. I am grateful for both their help and their friendship.

This text served as the basis for a graduate topics course taught by the author during the Spring 2009 semester at the University of Illinois at Urbana-Champaign. I am grateful to the U of I for allowing me this opportunity. Almost concurrently, Carl Pomerance used a preliminary version of these notes to teach a quarter-long course at Dartmouth College. This manuscript is better for his numerous insightful suggestions.

Finally, I would like to thank the American Mathematical Society, especially Ed Dunne, Cristin Zanella, and Luann Cole, for their encouragement of this project at every stage.

Chapter 1

Elementary Prime Number Theory, I

> Prime numbers are more than any assigned multitude of prime numbers. – Euclid
>
> No prime minister is a prime number – A. Plantinga

1. Introduction

Recall that a natural number larger than 1 is called *prime* if its only positive divisors are 1 and itself, and *composite* otherwise. The sequence of primes begins

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, \ldots$$

Few topics in number theory attract more attention, popular or professional, than the theory of prime numbers. It is not hard to see why. The study of the distribution of the primes possesses in abundance the very features that draw so many of us to mathematics in the first place: intrinsic beauty, accessible points of entry, and a lingering sense of mystery embodied in numerous unpretentious but infuriatingly obstinate open problems.

Put

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

Prime number theory begins with the following famous theorem from antiquity:

Theorem 1.1. *There are infinitely many primes, i.e.,* $\pi(x) \to \infty$ *as* $x \to \infty$.

The first half of this chapter is a survey of the many proofs that have been given for Theorem 1.1. The second half of this chapter is devoted to the theme of prime-producing formulas and the occurrence of primes in various natural sequences.

2. Euclid and his imitators

We begin with the classic proof from Euclid's *Elements* (circa 300 BC):

Proof. Suppose that $p_1, p_2, \ldots, p_k$ is any finite list of primes. Let P denote the product of the p_i and consider the integer $P+1$. Since $P+1 \equiv 1 \pmod{p_i}$ for each $1 \le i \le k$, none of the p_i divide $P+1$. But since $P+1 > 1$, it must have some prime divisor p. It follows that there is always a prime missing from any finite list, or, as Euclid put it, "prime numbers are more than any assigned multitude of primes." □

There are many trivial variants; for instance, we can easily show that for every integer m there is a prime $p > m$ by taking p to be any prime divisor of $m! + 1$.

In this section we collect several Euclidean proofs for Theorem 1.1. All of these start with a finite list of primes and then produce an integer > 1 that is coprime to every prime on the list. Stieltjes's proof is typical:

Stieltjes's proof, 1890. Suppose that $p_1, \ldots, p_k$ is a finite list of distinct primes with product P and let $P = AB$ be any decomposition of P into two positive factors. Suppose that p is one of the p_i. Then $p \mid AB$, so that either $p \mid A$ or $p \mid B$. If p divides both A and B, then p^2 divides P, which is false. Consequently, p divides exactly one of A and B. It follows that $p \nmid A + B$. So $A + B$ is divisible by none of the p_i; but as $A + B \ge 2$, it has some prime divisor. So again we have discovered a prime not on our original list. □

Euler's second proof (published posthumously). This proof is based on the multiplicativity of the Euler totient function: Let $p_1, \ldots, p_k$ be a list of distinct primes with product P. By said multiplicativity,

$$\varphi(P) = \prod_{i=1}^{k} (p_i - 1) \ge 2^{k-1} \ge 2,$$

provided that our list contains at least two primes (as we may assume). It follows that there is an integer in the interval $[2, P]$ that is coprime to P; but such an integer has a prime factor distinct from all of the p_i. □

Proof of Braun (1897), Métrod (1917). Let $p_1, \dots, p_k$ be a list of $k \geq 2$ distinct primes and let $P = p_1 p_2 \cdots p_k$. Consider the integer

$$N := P/p_1 + P/p_2 + \cdots + P/p_k.$$

For each $1 \leq i \leq k$, we have

$$N \equiv P/p_i = \prod\nolimits_{j \neq i} p_j \not\equiv 0 \pmod{p_i},$$

so that N is divisible by none of the p_i. But $N \geq 2$, and so it must possess a prime factor not on our list. □

3. Coprime integer sequences

Suppose we know an infinite sequence of pairwise relatively prime positive integers

$$2 \leq n_1 < n_2 < \cdots .$$

Then we may define a sequence of primes p_i by selecting arbitrarily a prime divisor of the corresponding n_i; the terms of this sequence are pairwise distinct because the n_i are pairwise coprime.

If we can exhibit such a sequence of n_i without invoking the infinitude of the primes, then we have a further proof of Theorem 1.1. An argument of this nature was given by Goldbach:

Proof (Goldbach). Let $n_1 = 3$, and for $i > 1$ inductively define

$$n_i = 2 + \prod_{1 \leq j < i} n_j.$$

The following assertions are all easily verified in succession:

(i) Each n_i is odd.

(ii) When $j > i$, we have $n_j \equiv 2 \mod n_i$.

(iii) We have $\gcd(n_i, n_j) = 1$ for $i \neq j$.

Theorem 1.1 now follows from the above remarks. □

A straightforward induction shows that

$$n_i = 2^{2^{i-1}} + 1, \tag{1.1}$$

and this is how Goldbach presented the proof.

Before proceeding, we pause to note that the above proof implies more than simply the infinitude of the primes. First, it gives us an upper bound for the nth prime, $2^{2^{n-1}} + 1$; this translates into a lower bound of the shape

$$\pi(x) \gg \log\log x \quad (x \to \infty).$$

Second, it may be used to prove that certain arithmetic progressions contain infinitely many primes. To see this, suppose that $p \mid n_i$ and note that by (1.1), we have

$$2^{2^{i-1}} \equiv -1 \pmod{p}, \quad \text{so that} \quad 2^{2^i} \equiv (2^{2^{i-1}})^2 \equiv 1 \pmod{p}.$$

Hence the order of 2 modulo p is precisely 2^i. Thus $2^i \mid (\mathbf{Z}/p\mathbf{Z})^\times = p-1$, so that $p \equiv 1 \pmod{2^i}$. As a consequence, for any fixed k, there are infinitely many primes $p \equiv 1 \pmod{2^k}$: choose a prime p_i dividing n_i for each $i \geq k$. In §9.1 we will prove the more general result that for each $m \geq 1$, there are infinitely many primes $p \equiv 1 \pmod{m}$.

A related method of proving the infinitude of the primes is as follows: Let $a_1 < a_2 < a_3 < \cdots$ be a sequence of positive integers with the property that

$$\gcd(i,j) = 1 \quad \Longrightarrow \quad \gcd(a_i, a_j) = 1.$$

Moreover, suppose that for some prime p, the integer a_p has at least two distinct prime divisors. Then if $p_1, \ldots, p_k$ were a list of all the primes, the integer

$$a_{p_1} a_{p_2} \cdots a_{p_k}$$

would possess at least $k+1$ prime factors: indeed, each factor exceeds 1, the factors are pairwise relatively prime, and one of the factors is divisible by two distinct primes. So there are $k+1 > k$ primes, a contradiction.

It remains to construct such a sequence. We leave to the reader the easy exercise of showing that $a_n = 2^n - 1$ has the desired properties (note that $a_{11} = 23 \cdot 89$). The original version of this argument, where a_n is instead chosen as the nth Fibonacci number, is due to Wunderlich [**Wun65**]. The generalization presented here is that of Hemminger [**Hem66**].

Saidak [**Sai06**] has recently given a very simple argument making use of coprimality. Start with a natural number $n > 1$. Because n and $n+1$ are coprime, the number $N_2 := n(n+1)$ must have at least two distinct prime factors. By the same reasoning,

$$N_3 := N_2(N_2+1) = n(n+1)(n(n+1)+1)$$

must have at least three distinct prime factors. In general, having constructed N_j with at least j different prime factors, the number $N_{j+1} := N_j(N_j+1)$ must have at least $j+1$.

4. The Euler-Riemann zeta function

For complex numbers s with real part greater than 1, define the zeta function by putting

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

(The condition that $\Re(s) > 1$ guarantees convergence of the series.) In the analytic approach to prime number theory, this function occupies a central position. Because of this text's emphasis on elementary methods, the zeta function will not play a large role for us, but it should be stressed that in many of the deeper investigations into the distribution of primes, the zeta function is an indispensable tool.

Riemann introduced the study of $\zeta(s)$ as a function of a complex variable in an 1859 memoir on the distribution of primes [**Rie59**]. But the connection between the zeta function and prime number theory goes back earlier. Over a hundred years prior to Riemann's study, Euler had looked at the same series for real s and had shown that [**Eul37**, Theorema 8]

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1 - \frac{1}{p^s}} \quad (s > 1). \tag{1.2}$$

This is often called an analytic statement of unique factorization. To see why, notice that formally (i.e., disregarding matters of convergence)

$$\prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots\right) = \sum_{n=1}^{\infty} \frac{a_n}{n^s},$$

where a_n counts the number of factorizations of n into prime powers. Thus unique factorization, the statement that $a_n = 1$ for all n, is equivalent to the statement that (1.2) holds as a formal product of Dirichlet series.[1] This, in turn, is equivalent to the validity of (1.2) for all real $s > 1$ (or even a sequence of s tending to ∞) by a standard result in the theory of Dirichlet series (see, e.g., [**Apo76**, Theorem 11.3]).

Euler's product expansion of the zeta function is the first example of what is now called an *Euler factorization*. We now prove (following [**Hua82**]) a theorem giving general conditions for the validity of such factorizations.

Theorem 1.2 (Euler factorizations)**.** *Let f be a multiplicative function. Then*

$$\sum_{n=1}^{\infty} f(n) = \prod_p \left(1 + f(p) + f(p^2) + \cdots\right) \tag{1.3}$$

if either of the following two conditions holds:

(i) $\sum_{n=1}^{\infty} |f(n)|$ *converges.*
(ii) $\prod_p \left(1 + |f(p)| + |f(p^2)| + \cdots\right)$ *converges.*

[1] Here a *Dirichlet series* is a series of the form $F(s) = \sum_{n=1}^{\infty} c_n/n^s$, where each c_n is a complex number.

Remark. Without imposing a condition such as (i) or (ii), it is possible for either the series or the product in (1.3) to converge while the other diverges, or for both to converge without being equal. See [**Win43**, §15] for explicit examples.

If f is not merely multiplicative but completely multiplicative, then the factors in (1.3) form a geometric series whose convergence is implied by either of the above conditions. Thus we have the following consequence:

Corollary 1.3. *Let f be a completely multiplicative function. Then*

$$\sum_{n=1}^{\infty} f(n) = \prod_{p} \frac{1}{1 - f(p)}$$

subject to either of the two convergence criteria of Theorem 1.2.

The factorization (1.2) of the zeta function is immediate from this corollary: One takes $f(n) = 1/n^s$ and observes that for $s > 1$, condition (i) holds (for example) by the integral test.

Proof of Theorem 1.2. Suppose that condition (i) holds and set $S_0 := \sum_{n=1}^{\infty} |f(n)|$. For each prime p, the series $\sum_{k=0}^{\infty} f(p^k)$ converges absolutely, since $\sum_{k=0}^{\infty} |f(p^k)| \le S_0$. Therefore

$$P(x) = \prod_{p \le x} \big(1 + f(p) + f(p^2) + \cdots\big)$$

is a finite product of absolutely convergent series. It follows that

$$P(x) = \sum_{n : p|n \Rightarrow p \le x} f(n).$$

If we now set $S = \sum_{n=1}^{\infty} f(n)$ (which converges absolutely), we have

$$S - P(x) = \sum_{n : p|n \text{ for some } p > x} f(n),$$

which shows

$$|S - P(x)| \le \sum_{n > x} |f(n)| \to 0$$

as $x \to \infty$. Thus $P(x) \to S$ as $x \to \infty$, which is the assertion of (1.3).

Now suppose that (ii) holds. We shall show that (i) holds as well, so that the theorem follows from what we have just done. To see this, let

$$P_0 = \prod_{p} \big(1 + |f(p)| + |f(p^2)| + \cdots\big),$$

and let

$$P_0(x) := \prod_{p\le x} \left(1 + |f(p)| + |f(p^2)| + \cdots\right)$$
$$= \sum_{n:p|n\Rightarrow p\le x} |f(n)| \ge \sum_{n\le x} |f(n)|.$$

Since $P_0(x) \le P_0$ for all x, the partial sums $\sum_{n\le x} |f(n)|$ form a bounded increasing sequence. Thus $\sum |f(n)|$ converges, proving (i). □

We can now present Euler's first proof of the infinitude of the primes.

Euler's first proof of Theorem 1.1. Let f be defined by $f(n) = 1/n$ for every n. Assuming that there are only finitely many primes, condition (ii) of Theorem 1.3 is trivially satisfied, as the product in question only has finitely many terms. It follows that

$$\sum_{n=1}^{\infty} \frac{1}{n} = \prod_{p} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right) < \infty,$$

in contradiction with the well-known divergence of the harmonic series. □

As pointed out by Euler, this proof gives a much stronger result than that asserted in Theorem 1.1.

Theorem 1.4. *The series $\sum \frac{1}{p}$ diverges, where the sum extends over all primes p.*

Proof. Suppose not and let $C = \sum 1/p$. As in the last proof, we take $f(n) = 1/n$ and apply Theorem 1.2. Let us check that condition (ii) of that theorem holds here. First, notice that

$$\prod_{p\le x} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right) = \prod_{p\le x} \frac{1}{1 - \frac{1}{p}} = \prod_{p\le x} \left(1 + \frac{1}{p-1}\right) \le \prod_{p\le x} \left(1 + \frac{2}{p}\right).$$

Now recall that $e^t \ge 1 + t$ for every nonnegative t; this is clear from truncating the Taylor expansion $e^t = 1 + t + t^2/2! + \ldots$. It follows that

$$\prod_{p\le x} \left(1 + \frac{2}{p}\right) \le \prod_{p\le x} e^{2/p} = \exp\left(\sum_{p\le x} 2/p\right) \le \exp(2C).$$

Consequently, the partial products

$$\prod_{p\le x} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right)$$

form a bounded, increasing sequence, which shows that we have condition (ii). We conclude that

$$\sum_{n=1}^{\infty} \frac{1}{n} = \prod_p \frac{1}{1-\frac{1}{p}} \leq \exp(2C),$$

a contradiction. □

Tweaking this argument, it is possible to derive an explicit lower bound on the partial sums $\sum_{p\leq x} 1/p$: Note that for $x \geq 2$,

$$\prod_{p\leq x} \frac{1}{1-\frac{1}{p}} = \sum_{n:p|n\Rightarrow p\leq x} \frac{1}{n} \geq \sum_{n\leq x} \frac{1}{n} \geq \log x. \tag{1.4}$$

From the upper bound $(1-1/p)^{-1} = (1+1/(p-1)) \leq \exp((p-1)^{-1})$, we deduce (taking the logarithm of (1.4)) that $\sum_{p\leq x} (p-1)^{-1} \geq \log\log x$. To derive a lower bound for $\sum_{p\leq x} 1/p$ from this, note that

$$\begin{aligned}\sum_{p\leq x} \frac{1}{p} &= \sum_{p\leq x} \frac{1}{p-1} - \sum_{p\leq x} \left(\frac{1}{p-1} - \frac{1}{p}\right) \\ &\geq \sum_{p\leq x} \frac{1}{p-1} - \sum_{n\geq 2} \left(\frac{1}{n-1} - \frac{1}{n}\right) = \left(\sum_{p\leq x} \frac{1}{p-1}\right) - 1 \geq \log\log x - 1.\end{aligned} \tag{1.5}$$

The next two proofs also make use of the zeta function and its Euler factorization, but in a decidedly different manner.

Proof of J. Hacks. We need the well-known result, also due to Euler, that $\zeta(2) = \pi^2/6$; a proof is sketched in Exercise 5 (for alternative arguments see [**AZ04**, Chapter 7], [**Cha02**]). Plugging $s = 2$ into the Euler factorization (1.2) we obtain

$$\frac{\pi^2}{6} = \zeta(2) = \prod_p \frac{1}{1-\frac{1}{p^2}}.$$

If there are only finitely many primes, then the product appearing here is a finite product of rational numbers, so that $\pi^2/6$ must also be a rational number. But this is impossible, since π is well known to be a *transcendental number*, i.e., not the root of any nonzero polynomial with rational coefficients. A weaker result, which suffices for the current argument, is the subject of Exercise 6 (cf. [**AZ04**, Chapter 6, Theorem 2]). □

One can give a similar argument avoiding irrationality considerations:

Proof. We use not only that $\zeta(2) = \pi^2/6$ but also that $\zeta(4) = \pi^4/90$. (Again see Exercise 5.) Thus $\zeta(2)^2/\zeta(4) = 5/2$. The Euler factorization (1.2) implies that

$$\frac{5}{2} = \frac{\zeta(2)^2}{\zeta(4)} = \prod_p (1-p^{-4})(1-p^{-2})^{-2} = \prod_p \frac{p^4-1}{p^4}\frac{p^4}{(p^2-1)^2} = \prod_p \frac{p^2+1}{p^2-1},$$

so that

$$\frac{5}{2} = \frac{5}{3}\cdot\frac{10}{8}\cdot\frac{26}{24}\cdots.$$

If there are only finitely many primes, then the product on the right-hand side is a finite one and can be written as M/N, where $M = 5\cdot 10\cdot 26\cdots$ and $N = 3\cdot 8\cdot 24\cdots$. Then $M/N = 5/2$, so $2M = 5N$. Since $3 \mid N$, it must be that $3 \mid M$. But this cannot be: M is a product of numbers of the form k^2+1, and no such number is a multiple of 3. □

Wagstaff has asked whether one can give a more elementary proof that $5/2 = \prod_p \frac{p^2+1}{p^2-1}$. The discussion of this (open) question in [**Guy04**, B48] was the motivation for the preceding proof of Theorem 1.1.

5. Squarefree and smooth numbers

Recall that a natural number n is said to be *squarefree* if it is not divisible by the square of any integer larger than 1. The fundamental theorem of arithmetic shows that there is a bijection

$$\{\text{finite subsets of the primes}\} \longleftrightarrow \{\text{squarefree positive integers}\},$$

given by sending

$$S \longmapsto \prod_{p\in S} p.$$

So to prove the infinitude of the primes, it suffices to prove that there are infinitely many positive squarefree integers.

J. Perott's proof, 1881. We sieve out the non-squarefree integers from $1,\dots,N$ by removing those divisible by 2^2, then those divisible by 3^2, etc. The number of removed integers is bounded above by

$$\sum_{k=2}^{\infty} \lfloor N/k^2 \rfloor \le N\sum_{k=2}^{\infty} k^{-2} = N(\zeta(2)-1),$$

so that the number of squarefree integers up to N, say $A(N)$, satisfies

$$A(N) \ge N - N(\zeta(2)-1) = N(2-\zeta(2)). \tag{1.6}$$

At this point Perott uses the evaluation $\zeta(2) = \pi^2/6$. However, it is simpler to proceed as follows: Since t^{-2} is a decreasing function of t on the positive real axis,

$$\zeta(2) = 1 + \sum_{n=2}^{\infty} \frac{1}{n^2} < 1 + \sum_{n=1}^{\infty} \int_n^{n+1} \frac{dt}{t^2} = 1 + \int_1^{\infty} \frac{dt}{t^2} = 2.$$

Referring back to (1.6), we see that $A(N)/N$ is bounded below by a positive constant. In particular, it must be that $A(N) \to \infty$ as $N \to \infty$. □

Remark. As observed by Dressler [**Dre75**], Perott's argument also yields a lower bound on $\pi(N)$. Note that since every squarefree number $\le N$ is a product of some subset of the $\pi(N)$ primes up to N, we have $2^{\pi(N)} \ge A(N)$. The argument above establishes that $A(N) \ge cN$ for $c = 2 - \zeta(2) > 0$, and so $\pi(N) \ge \log N/\log 2 + O(1)$.

For the next proof we need the following simple lemma:

Lemma 1.5. *Every natural number n can be written in the form rs^2, where r and s are natural numbers and r is squarefree.*

Proof. Choose the positive integer s so that s^2 is the largest perfect square dividing n, and put $r = n/s^2$. We claim that r is squarefree. Otherwise $p^2 \mid r$ for some prime p. But then $(ps)^2 \mid n$, contrary to the choice of s. □

Erdős's proof of Theorem 1.1. Let N be a positive integer. There are at most $\sqrt{N}$ squares not exceeding N and at most $2^{\pi(N)}$ squarefree integers below this bound. So Lemma 1.5 implies that

$$2^{\pi(N)}\sqrt{N} \ge N.$$

Dividing by $\sqrt{N}$ and taking logarithms yields the lower bound $\pi(N) \ge \log N/\log 4$. □

A modification of this argument leads to another proof that $\sum \frac{1}{p}$ diverges:

Erdős's proof of Theorem 1.4. Suppose that $\sum 1/p$ converges. Then we can choose an M for which

$$\sum_{p>M} \frac{1}{p} < \frac{1}{2}. \tag{1.7}$$

Keep this M fixed.

Let N be an arbitrary natural number. The estimate (1.7) implies that most integers up to N factor completely over the primes not exceeding M.

Indeed, the number of integers not exceeding N that have a prime factor $p > M$ is bounded above by

$$\sum_{M<p\le N} \left\lfloor \frac{N}{p} \right\rfloor \le N \sum_{p>M} \frac{1}{p} < N/2,$$

so that more than $N/2$ of the natural numbers not exceeding N are divisible only by primes $p \le M$.

We now show that there are too few integers divisible only by primes $p \le M$ for this to be possible. There are at most $\sqrt{N}$ squares not exceeding N and at most $C := 2^{\pi(M)}$ squarefree numbers composed only of primes not exceeding M. Thus there are at most $C\sqrt{N}$ natural numbers $\le N$ having all their prime factors $\le M$. But $C\sqrt{N} < N/2$ once $N > 4C^2$. □

In the last argument we needed an estimate for the number of integers up to a given point with only small prime factors. This motivates the following definition: Call a natural number *y-smooth* if all of its prime factors are bounded by y. We let $\Psi(x, y)$ denote the number of y-smooth numbers not exceeding x; i.e.,

$$\Psi(x, y) := \#\{n \le x : p \mid n \Rightarrow p \le y\}. \tag{1.8}$$

Smooth numbers are important auxiliary tools in many number-theoretic investigations, and so there has been quite a bit of work on estimating the size of $\Psi(x, y)$ in various ranges of x and y. (For a survey of both the applications and the estimates, see [**Gra08b**].) A trivial estimate yields an easy proof of Theorem 1.1.

Lemma 1.6. *For $x \ge 1$ and $y \ge 2$, we have*

$$\Psi(x, y) \le \left(1 + \frac{\log x}{\log 2}\right)^{\pi(y)}.$$

Proof. Let $k = \pi(y)$. By the fundamental theorem of arithmetic, $\Psi(x, y)$ is the number of k-tuples of nonnegative integers $e_1, \dots, e_k$ with

$$p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k} \le x.$$

This inequality requires $p_i^{e_i} \le x$, so that

$$e_i \le \log x / \log p_i \le \log x / \log 2,$$

so that there are at most $1 + \lfloor \log x / \log 2 \rfloor$ possibilities for each e_i. □

Since every positive integer not exceeding N is a (possibly empty) product of primes not exceeding N,

$$N = \Psi(N, N) \le (1 + \log N / \log 2)^{\pi(N)}.$$

It follows that

$$\pi(N) \geq \frac{\log N}{\log(1 + \log N/\log 2)}.$$

Taking some care to estimate the denominator, we obtain the lower bound

$$\pi(N) \geq (1 + o(1))\frac{\log N}{\log\log N},$$

which tends to infinity. Similar proofs of Theorem 1.1 have been given by Thue (1897), Auric (1915), Schnirelmann [**Sch40**, pp. 44–45], Chernoff [**Che65**], and Rubinstein [**Rub93**]. See also Exercise 17.

6. Sledgehammers!

In the spirit of the saying, "nothing is too simple to be made complicated," we finish off the first half of this chapter with two proofs of Theorem 1.1 that dip into the tool chest of higher mathematics.

The following "topological proof" is due to Furstenberg ([**Fur55**]):

Proof. We put a topology on $\mathbf{Z}$ by taking as a basis for the open sets all arithmetic progressions, infinite in both directions. (This is permissible since the intersection of two such progressions is either empty or is itself an arithmetic progression.) Then each arithmetic progression is both open and closed: it is open by choice of the basis, and it is closed since its complement is the union of the other arithmetic progressions with the same common difference. For each prime p, let $A_p = p\mathbf{Z}$, and define $A := \bigcup_p A_p$. The set $\{-1, 1\} = \mathbf{Z} \setminus A$ is not open. (Indeed, each open set is either empty or contains an arithmetic progression, so must be infinite.) It follows that A is not closed. On the other hand, if there are only finitely many primes, then A is a finite union of closed sets, and so it *is* closed. □

Our next proof, due to L. Washington (and taken from [**Rib96**]) uses the machinery of commutative algebra. Recall that a *Dedekind domain* is an integral domain R with the following three properties:

(i) R is *Noetherian*: if $I_1 \subset I_2 \subset I_3 \subset \cdots$ is an ascending chain of ideals of R, then there is an n for which

$$I_n = I_{n+1} = I_{n+2} = \cdots.$$

(ii) R is *integrally closed*: if K denotes the fraction field of R and $\alpha \in K$ is the root of a monic polynomial with coefficients in R, then in fact $\alpha \in R$.

(iii) Every nonzero prime ideal of R is a maximal ideal.

Proof. We use the theorem that a Dedekind domain with finitely many nonzero prime ideals is a principal ideal domain (see, e.g., [**Lor96**, Proposition III.2.12]) and thus also a unique factorization domain. The ring of integers $\mathfrak{O}_K$ of a number field K is always a Dedekind domain; consequently, if K does not possess unique factorization, then $\mathfrak{O}_K$ has infinitely many nonzero prime ideals. Each such prime ideal lies above a rational prime p, and for each prime p there are at most $[K : \mathbf{Q}]$ prime ideals lying above it. It follows that there are infinitely many primes p, provided that there is a single number field K for which $\mathfrak{O}_K$ does not possess unique factorization. And there is: If $K = \mathbf{Q}(\sqrt{-5})$, then

$$6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$$

is a well-known instance of the failure of unique factorization in $\mathfrak{O}_K = \mathbf{Z}[\sqrt{-5}]$. □

7. Prime-producing formulas

> A mathematician is a conjurer who gives away his secrets. – J. H. Conway

Now that we know there are infinitely many primes, the next question is: Where are they hiding? Or, to ask a question that has ensnared many who have flirted with number theory: Is there a formula for producing primes? This line of inquiry, as natural as it seems, has not been very productive.

The following 1952 result of Sierpiński [**Sie52**] is representative of many in this subject. Let p_n denote the nth prime number. Define a real number ξ by putting

$$\xi := \sum_{n=1}^{\infty} p_n 10^{-2^n} = 0.0203000500000007000000000000011\ldots.$$

★ **Theorem 1.7.** *We have*

$$p_n = \left\lfloor 10^{2^n} \xi \right\rfloor - 10^{2^{n-1}} \left\lfloor 10^{2^{n-1}} \xi \right\rfloor.$$

This is, in the literal sense, a formula for primes. But while it may have some aesthetic merit, it must be considered a complete failure from the standpoint of utility; determining the number ξ seems to require us to already know the sequence of primes. A similar criticism can be leveled against a result of Mills [**Mil47**], which asserts the existence of a real number $A > 1$ with the property that $\lfloor A^{3^n} \rfloor$ is prime for each natural number n.

A more surprising way of generating primes was proposed by J. H. Conway [**Con87**]. Consider the following list of 14 fractions:

A	B	C	D	E	F	G	H	I	J	K	L	M	N
$\frac{17}{91}$	$\frac{78}{85}$	$\frac{19}{51}$	$\frac{23}{38}$	$\frac{29}{33}$	$\frac{77}{29}$	$\frac{95}{23}$	$\frac{77}{19}$	$\frac{1}{17}$	$\frac{11}{13}$	$\frac{13}{11}$	$\frac{15}{2}$	$\frac{1}{7}$	$\frac{55}{1}$

Now run the following algorithm: Beginning with the number 2, look for the first (leftmost) fraction which can be multiplied by the current number to give an integer. Perform the multiplication and repeat. Whenever you reach a power of 2, output the exponent. The first several (19) steps of the algorithm are

$$2 \mapsto 15 \mapsto 825 \mapsto 725 \mapsto 1925 \mapsto 2275 \mapsto 425 \mapsto 390 \mapsto 330 \mapsto 290 \mapsto 770$$
$$\mapsto 910 \mapsto 170 \mapsto 156 \mapsto 132 \mapsto 116 \mapsto 308 \mapsto 364 \mapsto 68 \mapsto 4 = 2^2,$$

and so the first output is 2. Fifty more steps yield

$$2^2 \mapsto 30 \mapsto 225 \mapsto 12375 \mapsto \cdots \mapsto 232 \mapsto 616 \mapsto 728 \mapsto 136 \mapsto 8 = 2^3,$$

and so the second output is 3. After another 212 steps, we arrive at $32 = 2^5$, and so our third output is 5.

★ **Theorem 1.8** (Conway)**.** *The sequence of outputs is exactly the sequence of primes in increasing order.*

This is rather striking; the sequence of primes, which seems random in so many ways, is the output of a deterministic algorithm involving 14 fractions. But perhaps this should not come as such a shock. Most anyone who has experimented with programming knows that the primes are the output of a deterministic algorithm: Test the numbers $2, 3, 4, \ldots$ successively for primality, using (say) trial division for the individual tests. And actually, underneath the surface, this is exactly what is being done in Conway's algorithm. This sequence of 14 fractions encodes a simple computer program: The number n is tested for divisibility first by $d = n - 1$, then $d = n - 2$, etc; as soon as a divisor is found, n is incremented by 1 and the process is repeated. The game is rigged so that a power of 2 arises only when d reaches 1, i.e., when n is prime. Moreover, there is nothing special in Theorem 1.8 about the sequence of primes; an analogue of Theorem 1.8 can be proved for any recursive set. (Here a set of natural numbers S is called *recursive* if there is an algorithm for determining whether a natural number belongs to S.) We conclude that while Conway's result *is* genuinely surprising, the surprise is that one can simulate computer programs with lists of fractions, and is in no way specific to the prime numbers.

8. Euler's prime-producing polynomial

The prime-producing functions we have been considering up to now have all been rather complicated. In some sense this is necessary; one can show that

any function which produces only primes cannot have too simple a form. We give only one early example of a result in this direction. (See [**War30**], [**Rei43**] for more theorems of this flavor.)

Theorem 1.9 (Goldbach)**.** *If $F(T) \in \mathbf{Z}[T]$ is a nonconstant polynomial with positive leading coefficient, then $F(n)$ is composite for infinitely many natural numbers n.*

Proof. Suppose F is nonconstant but that $F(n)$ is prime for all $n \geq N_0$, where N_0 is a natural number. Let $p = F(N_0)$; then p divides $F(N_0 + kp)$ for every positive integer k. But since F has a positive leading coefficient, $F(N_0 + kp) > p$ for every sufficiently large integer k, and so $F(N_0 + kp)$ is composite, contrary to the choice of N_0. □

Theorem 1.9 does not forbid the existence of polynomials F which assume prime values over impressively long stretches. And indeed these do exist; a famous example is due to Euler, who observed that if $f(T) = T^2 + T + 41$, then $f(n)$ is prime for all integers $0 \leq n < 40$.

It turns out that Euler's observation, rather than being an isolated curiosity, is intimately connected with the theory of imaginary quadratic fields. We will prove the following theorem:

Theorem 1.10. *Let $A \geq 2$, and set $D := 1 - 4A$. Then the following are equivalent:*

(i) *$n^2 + n + A$ is prime for all $0 \leq n < A - 1$,*

(ii) *$n^2 + n + A$ is prime for all $0 \leq n \leq \frac{1}{2}\sqrt{\frac{|D|}{3}} - \frac{1}{2}$,*

(iii) *the ring $\mathbf{Z}[(-1 + \sqrt{D})/2]$ is a unique factorization domain.*

The equivalence (i) ⇔ (iii) is proved by Rabinowitsch in [**Rab13**], and is usually referred to as *Rabinowitsch's theorem.*

Remark. Since $n^2 + n + A = (n + 1/2)^2 + (4A - 1)/4$, (ii) can be rephrased as asserting that $(n + 1/2)^2 + |D|/4$ is prime for every integer n for which $|n + 1/2| \leq \frac{1}{2}\sqrt{\frac{|D|}{3}}$. We will use this observation in the proof of Theorem 1.10.

Cognoscenti will recognize that $\mathbf{Z}[(-1 + \sqrt{D})/2]$ is an order in the quadratic field $\mathbf{Q}(\sqrt{D})$. However, the proof of Theorem 1.10 presented here, due to Gyarmati (née Lanczi) [**Lán65**], [**Gya83**] and Zaupper [**Zau83**], requires neither the vocabulary of algebraic number theory nor the theory of ideals.

We begin the proof of Theorem 1.10 by observing that the bound on n in (ii) is always at least as strict as the bound on n in (i), which makes clear that (i) implies (ii). So it is enough to show that (ii) implies (iii)

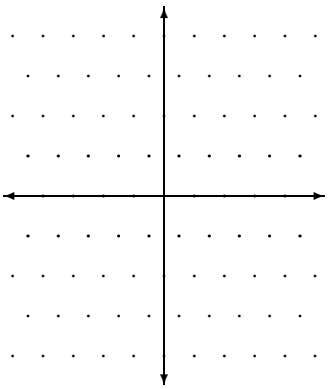

Figure 1. The lattice $\mathbf{Z} + \mathbf{Z}\eta$ sitting inside $\mathbf{C}$. Here $A = 2$ so that $D = -7$.

and that (iii) implies (i). To continue we need some preliminary results on the arithmetic of the rings $\mathbf{Z}[(-1 + \sqrt{D})/2]$. These will be familiar to students of algebraic number theory, but we include full proofs for the sake of completeness.

Let $A \geq 2$ be an integer, and fix a complex root η of $x^2 + x + A$, so that (for an appropriate choice of the square root) $\eta = (-1 + \sqrt{D})/2$. Since $\eta^2 = -\eta - A$, it follows that

$$\mathbf{Z}[\eta] = \mathbf{Z} + \mathbf{Z}\eta = \{x + y\eta : x, y \in \mathbf{Z}\}.$$

For $\alpha \in \mathbf{Z}[\eta]$, we denote its complex conjugate by $\overline{\alpha}$. Observe that $\overline{\eta} = -1 - \eta$; consequently, $\mathbf{Z}[\eta]$ is closed under complex-conjugation. We define the *norm* of the element $\alpha = x + y\eta \in \mathbf{Z}[\eta]$ by

$$\begin{aligned} \mathcal{N}(\alpha) :&= |\alpha|^2 \\ &= \alpha\overline{\alpha} = x^2 - xy + Ay^2. \end{aligned}$$

Notice that the norm of $\alpha \in \mathbf{Z}[\eta]$ is always an integer and is positive whenever $\alpha \neq 0$. Moreover, since the complex absolute value is multiplicative, it is immediate that

$$\mathcal{N}(\alpha\beta) = \mathcal{N}(\alpha) \cdot \mathcal{N}(\beta) \quad \text{for all} \quad \alpha, \beta \in \mathbf{Z}[\eta].$$

We now recall the requisite definitions from ring theory: If $\alpha, \beta \in \mathbf{Z}[\eta]$, we say that α *divides* β if $\beta = \alpha\gamma$ for some $\gamma \in \mathbf{Z}[\eta]$. A nonzero element $\alpha \in \mathbf{Z}[\eta]$ is called a *unit* if α divides 1. A nonunit element $\alpha \in \mathbf{Z}[\eta]$ is *irreducible* if whenever $\alpha = \beta\gamma$ with $\beta, \gamma \in \mathbf{Z}[\eta]$, then either β is a unit or γ is a unit. Finally, $\pi \in \mathbf{Z}[\eta]$ is called *prime* if whenever π divides $\beta\gamma$ for $\beta, \gamma \in \mathbf{Z}[\eta]$, then either π divides β or π divides γ.

Lemma 1.11. *An element $\alpha \in \mathbf{Z}[\eta]$ is a unit precisely when $\mathcal{N}(\alpha) = 1$. The only units in $\mathbf{Z}[\eta]$ are ± 1.*

Proof. If α is a unit, then $\mathcal{N}(\alpha) \cdot \mathcal{N}(\alpha^{-1}) = 1$. Moreover, both $\mathcal{N}(\alpha)$ and $\mathcal{N}(\alpha^{-1})$ are positive integers, so that $\mathcal{N}(\alpha) = \mathcal{N}(\alpha^{-1}) = 1$. Conversely, if $\mathcal{N}(\alpha) = 1$, then $\alpha\overline{\alpha} = 1$, and so α is a unit. Finally, notice that if $y \neq 0$, then

$$\mathcal{N}(x + y\eta) = x^2 - xy + Ay^2 = (x - y/2)^2 + \frac{1}{4}(4A - 1)y^2 \geq \frac{4A - 1}{4} > \frac{7}{4} > 1.$$

So $x + y\eta$ can be a unit only when $y = 0$. In this case we must have $\mathcal{N}(x) = x^2 = 1$, and this occurs exactly when $x = \pm 1$. □

Lemma 1.12. *If α is a nonzero, nonunit element of $\mathbf{Z}[\eta]$, then α can be written as a product of irreducible elements of $\mathbf{Z}[\eta]$.*

Proof. If the claim fails, there is a nonzero, nonunit α of smallest norm for which it fails. Clearly α is not irreducible, and so we can write $\alpha = \beta\gamma$, where β and γ are nonzero nonunits. Hence $\mathcal{N}(\alpha) = \mathcal{N}(\beta)\mathcal{N}(\gamma)$. Since $\mathcal{N}(\beta)$ and $\mathcal{N}(\gamma)$ are each larger than 1, both $\mathcal{N}(\beta)$ and $\mathcal{N}(\gamma)$ must be smaller than $\mathcal{N}(\alpha)$. So by the choice of α, both β and γ factor as products of irreducibles, and thus α does as well. This contradicts the choice of α. □

We can now prove one of the two outstanding implications:

Proof that (iii) ⇒ (i). Let $\eta = (-1 + \sqrt{D})/2$. Suppose $0 \leq n < A - 1$. We have

$$n^2 + n + A = (n - \eta)(n - \bar{\eta}) = (n - \eta)(n + 1 + \eta). \tag{1.9}$$

Let p be a prime dividing $n^2 + n + A$. We claim that p is not irreducible in $\mathbf{Z}[\eta]$. Indeed, since $\mathbf{Z}[\eta]$ is a unique factorization domain by hypothesis, if p were irreducible, then p would be prime. So from (1.9), we would have that p divides $n - \eta$ or $n + 1 + \eta$. But this is impossible, since neither $n/p - \eta/p$ nor $(n + 1)/p + \eta/p$ belongs to $\mathbf{Z}[\eta] = \mathbf{Z} + \mathbf{Z}\eta$.

Hence we can write $p = \alpha\beta$, where $\alpha, \beta \in \mathbf{Z}[\eta]$ and neither α nor β is a unit. Taking norms, we deduce that $p^2 = \mathcal{N}(p) = \mathcal{N}(\alpha)\mathcal{N}(\beta)$. Since α and β are not units, we must have $\mathcal{N}(\alpha) = \mathcal{N}(\beta) = p$.

Write $\alpha = x + y\eta$ for integers x, y. Then $y \neq 0$ (since p is a rational prime), and so

$$p = N(\alpha) = x^2 - xy + Ay^2 = (x - y/2)^2 + (A - 1/4)y^2 \geq A - 1/4.$$

Thus (since p is an integer) $p \geq A$. Moreover, since $0 \leq n < A - 1$,

$$n^2 + n + A < (A - 1)^2 + (A - 1) + A = (A - 1)A + A = A^2.$$

This shows that every prime divisor of $n^2 + n + A$ exceeds its square root, so that $n^2 + n + A$ is prime. □

The proof of the remaining implication requires one more preliminary result:

Lemma 1.13. *If π is an element of $\mathbf{Z}[\eta]$ whose norm is a rational prime p, then π is prime in $\mathbf{Z}[\eta]$.*

Proof. We claim that $\mathbf{Z}[\eta]/(\pi)$ is isomorphic to $\mathbf{Z}/p\mathbf{Z}$. Since $\mathbf{Z}/p\mathbf{Z}$ is a field, this implies that π generates a prime ideal of $\mathbf{Z}[\eta]$, which in turn implies that π is prime. Let $\psi\colon \mathbf{Z} \to \mathbf{Z}[\eta]/(\pi)$ be the ring homomorphism defined by mapping n to $n \bmod \pi$. Since $p = \pi\overline{\pi} \equiv 0 \pmod{\pi}$, the kernel of ψ contains the ideal $p\mathbf{Z}$. Since $p\mathbf{Z}$ is a maximal ideal, either ψ is identically zero or the kernel of ψ is precisely $p\mathbf{Z}$. Since π is not a unit in $\mathbf{Z}[\eta]$, $\psi(1)$ is nonzero, and so the kernel of ψ is precisely $p\mathbf{Z}$. Hence $\mathbf{Z}/p\mathbf{Z}$ is isomorphic to the image of ψ. So the proof will be complete if we show that ψ is surjective.

Write $\pi = r + s\eta$ for integers r and s, and let $x + y\eta$ be an arbitrary element of $\mathbf{Z}[\eta]$. We can choose integers a and b for which

$$m := x + y\eta - \pi(a + b\eta) \in \mathbf{Z}.$$

Indeed, a short computation shows that this containment holds precisely when

$$b(r - s) + as = y,$$

which is a solvable linear Diophantine equation in a and b since $\gcd(r-s, s) = \gcd(r, s) = 1$. Then $m \equiv x + y\eta \pmod{\pi}$, and so $\psi(m) = x + y\eta \bmod \pi$. Since $x + y\eta$ was arbitrary, ψ is surjective as claimed. □

Proof that (ii) ⇒ (iii). Suppose that $n^2 + n + A$ is prime for all

$$0 \le n \le \frac{1}{2}\sqrt{\frac{|D|}{3}} - \frac{1}{2}.$$

We are to prove that $\mathbf{Z}[\eta]$ possesses unique factorization. Suppose otherwise, and let α be a nonzero, nonunit of minimal norm with two distinct factorizations into irreducibles, say

$$\alpha = \pi_1 \cdots \pi_k = \rho_1 \cdots \rho_j.$$

(Here *distinct* means that either $k \ne j$, or that $k = j$, but there is no way to reorder the π_i so that each π_i is a unit multiple of ρ_i.) By the minimality of $\mathcal{N}(\alpha)$, it is easy to see that none of the irreducibles in the first factorization can be a unit multiple of an irreducible in the second factorization. Consequently, none of the irreducibles appearing in either factorization can be prime in $\mathbf{Z}[\eta]$.

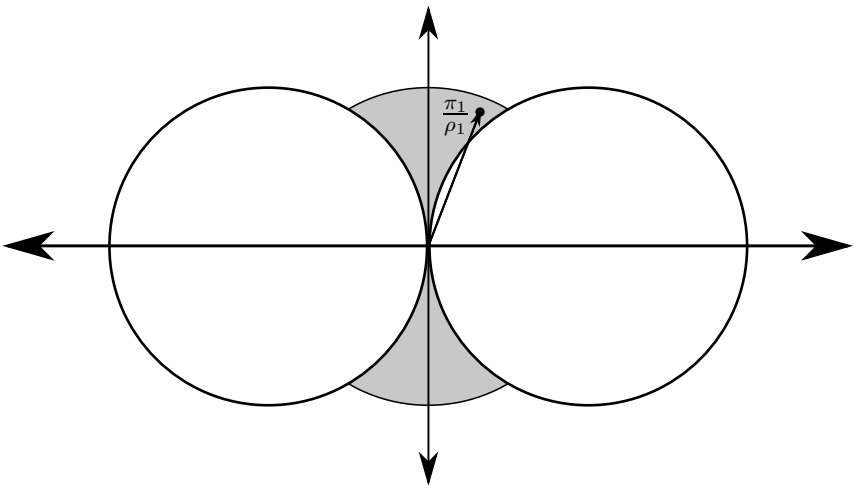

Figure 2. (Based on [**Zau83**].)

We can assume that $\mathcal{N}(\pi_1) \leq \mathcal{N}(\rho_1)$. (If this does not hold initially, interchange the two factorizations.) For $\xi, \gamma \in \mathbf{Z}[\eta]$ still to be chosen, define

$$\alpha' := (\rho_1\xi - \pi_1\gamma)\rho_2 \cdots \rho_j. \tag{1.10}$$

Then

$$\begin{aligned}\alpha' &= \alpha\xi - \pi_1\frac{\alpha}{\rho_1}\gamma \\ &= \pi_1(\pi_2 \cdots \pi_k\xi - \rho_2 \cdots \rho_j\gamma).\end{aligned}$$

Factoring the parenthetical expression, we deduce that α' has a factorization into irreducibles where one of the irreducibles is π_1. We will choose ξ and γ so that $\pi_1 \nmid \rho_1\xi$. Then $\pi_1 \nmid \rho_1\xi - \pi_1\gamma$, and so we may deduce from (1.10) that α' has a factorization into irreducibles, none of which is a unit multiple of π_1. So α' possesses two distinct factorizations into irreducibles. If further, γ and ξ satisfy

$$\mathcal{N}(\rho_1\xi - \pi_1\gamma) < \mathcal{N}(\rho_1),$$

then $\mathcal{N}(\alpha')$ is smaller than $\mathcal{N}(\alpha)$, and so we have a contradiction to our choice of α.

So it remains to show that it is possible to choose $\xi, \gamma \in \mathbf{Z}[\eta]$ with the following two properties:

(P1) $\pi_1 \nmid \rho_1\xi$,

(P2) $\mathcal{N}(\rho_1\xi - \pi_1\gamma) < \mathcal{N}(\rho_1)$, or equivalently, $\left|\xi - \frac{\pi_1}{\rho_1}\gamma\right| < 1$.

Since $\mathcal{N}(\pi_1) \leq \mathcal{N}(\rho_1)$, the complex number π_1/ρ_1 lies on or inside the unit circle. Suppose first that π_1/ρ_1 lies outside the shaded region indicated in

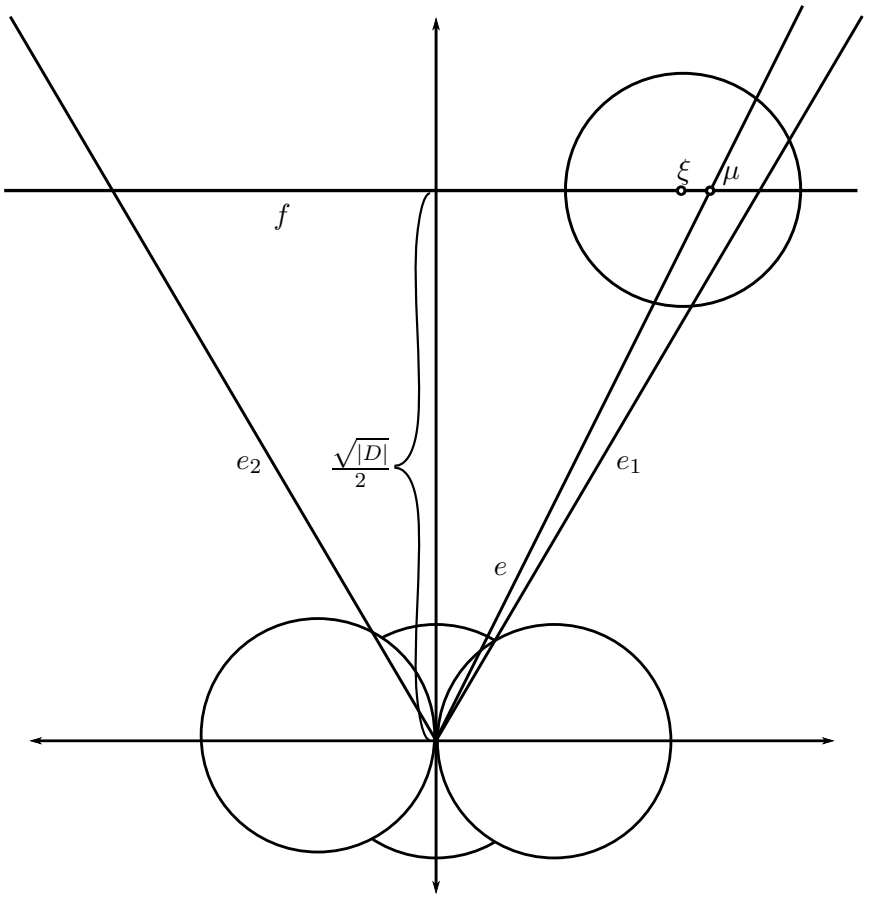

Figure 3. (Based on [**Zau83**].)

Figure 2. Then for either $\xi = 1$ or $\xi = -1$, we have

$$|\xi - \pi_1/\rho_1| < 1.$$

Then (P1) and (P2) hold if we choose this value of ξ and take $\gamma = 1$. Note that $\pi_1 \nmid \pm\rho_1$, since otherwise π_1 and ρ_1 would be unit multiples of each other, which we have already argued is not the case.

So we may assume that π_1/ρ_1 lies within the shaded region. Let e_1 be the ray from the origin making an angle of 60° with the x-axis, and let e_2 be the ray from the origin making an angle of 120° with that axis. Then the ray e (say) from the origin through π_1/ρ_1 is contained within the 60° angle determined by e_1 and e_2.[2] Let f be the horizontal line consisting of those complex numbers with imaginary part $\sqrt{|D|}/2$; thus f is the first horizontal line above the x-axis containing points of the lattice $\mathbf{Z} + \mathbf{Z}\eta$. Let μ be the

[2]Here the *angle determined by* e_1 *and* e_2 means the closed set of points between e_1 and e_2.

complex number corresponding to the intersection of e and f. The angle determined by e_1 and e_2 cuts f into a segment of length $\sqrt{|D|/3} > 1$, and so there is a point of $\mathbf{Z} + \mathbf{Z}\eta$ on f within this angle. We choose such a point ξ for which the distance from ξ to μ is as small as possible. See Figure 3.

We claim that the distance from ξ to e is strictly smaller than $\sqrt{3}/2$. This is clear if both $\xi + 1$ and $\xi - 1$ fall within the angle determined by e_1 and e_2, since in that case, the distance from ξ to μ must be at most $1/2$. So suppose that $\xi + 1$ falls outside this angle; the case when $\xi - 1$ falls outside is analogous. Then $\xi - 1$ must lie within the given angle. Now, if ξ is to the right of μ, then in order that ξ be at least as close to μ as $\xi - 1$, it must be that the distance from ξ to μ is at most $1/2$. So we can assume that ξ falls to the left of μ. This is the scenario depicted in Figure 3. In this case we we use the following argument: Let ν represent the intersection of e_1 and f; then the distance between ξ and ν is smaller than 1. Since e_1 makes an angle of 60° with f, elementary trigonometry shows that the distance from ξ to e_1 is strictly smaller than $\sqrt{3}/2$. But the perpendicular line segment from ξ to e_1 meets e. So the distance from ξ to e is also strictly smaller than $\sqrt{3}/2$.

It follows that the unit disc centered at ξ intersects e in a segment of total length > 1. (Indeed, let τ be the point on e for which the line from ξ to τ is perpendicular to e, so that the distance from ξ to τ is strictly smaller than $\sqrt{3}/2$. Then by the Pythagorean theorem, τ divides the segment in question into two parts, each of length $> 1/2$.) Since $|\pi_1/\rho_1| \leq 1$, it follows that we can choose a rational integer γ so that $\gamma\pi_1/\rho_1$ lies within the open unit disc centered at ξ.

We claim that with the above choices of ξ and γ, both (P1) and (P2) hold. Condition (P2) is guaranteed by the choice of γ, so it remains only to verify (P1). For this it is enough to prove that ξ is prime. Indeed, suppose that ξ is prime but (P1) fails. Then

$$\rho_1 \xi = \pi_1 \kappa$$

for some κ. Since ξ is prime, it must divide either π_1 or κ. But ξ cannot divide π_1; if it did, then since π_1 is irreducible, we would have that π_1 is a unit multiple of ξ. But then π_1 would be prime since ξ is prime. This contradicts the observation made above that none of the π_i are prime. So ξ must divide κ; but then dividing through by ξ we find that π_1 divides ρ_1. That implies that π_1 and ρ_1 are unit multiples of each other, which again contradicts our initial observations.

Why should ξ be prime? Since ξ is a point of the lattice $\mathbf{Z} + \mathbf{Z}\eta$ lying on f, we have $\xi = n + \eta$ for some integer n. Moreover, since ξ belongs to

the 60° angle determined by e_1 and e_2, we find that

$$|(n-1)+1/2| = |n-1/2| \le \frac{1}{2}\sqrt{|D|/3}.$$

But now (ii) of Theorem 1.10 implies that

$$\begin{aligned}\mathcal{N}(\xi) &= n^2 - n + A \\ &= (n-1)^2 + (n-1) + A\end{aligned}$$

is prime, so that ξ is a prime element of $\mathbf{Z}[\eta]$ by Lemma 1.13. □

A small amount of computation shows that condition (ii) of Theorem 1.10 holds for the values $A = 2, 3, 5, 11, 17$, and 41. This yields the following corollary:

Corollary 1.14. $\mathbf{Z}[(-1+\sqrt{D})/2]$ *is a unique factorization domain for* $D = -7, -11, -19, -43, -67, -163$.

Checking larger values of A does not appear to yield any more examples satisfying the conditions of Theorem 1.10. Whether or not the list in Corollary 1.14 is complete is known as the *class number* 1 *problem*; an equivalent question appears in Gauss's *Disquisitiones* (see [**Gau86**, Art. 303]). In 1933, Lehmer showed [**Leh33**] that any missing value of A is necessarily large, in that $|D| > 5 \cdot 10^9$. In 1934, Heilbronn & Linfoot [**HL34**] showed that there is at most one missing value of A. Finally, in 1952, Heegner settled the problem, using new techniques from the theory of modular functions:

Theorem 1.15 (Heegner)**.** *If* $A > 41$*, then* $\mathbf{Z}[\eta]$ *does not have unique factorization. Hence if* $A \ge 2$ *is an integer for which* $n^2 + n + A$ *is prime for all* $0 \le n < A - 1$*, then* $A \le 41$.

For a modern account of Heegner's proof, see [**Cox89**, §12].

9. Primes represented by general polynomials

The result of the previous section leaves a very natural question unresolved: Does Euler's polynomial $T^2 + T + 41$, which does such a marvelous job of producing primes at the first several natural numbers n, represent infinitely many primes as n ranges over the set of all positive integers? More generally, what can one say about the set of prime values assumed by a polynomial $F(T) \in \mathbf{Z}[T]$? In this section we survey the known results in this direction.

9.1. The linear case. Suppose first that $F(T)$ is linear, say $F(T) = a + mT$, where $m > 0$. Asking whether $F(n)$ is prime for infinitely many natural numbers n amounts to asking whether the infinite arithmetic progression

$$a+m, \quad a+2m, \quad a+3m, \quad a+4m, \quad \ldots$$

contains infinitely many primes — or, phrased in terms of congruences, whether or not there are infinitely many primes $p \equiv a \pmod{m}$.

This question is sometimes easy to answer. Let $d = \gcd(a, m)$. If $d > 1$, then there are at most finitely many primes in the above progression, since every term is divisible by d, and so we have a negative answer to our query. So let us suppose that $d = 1$. Then certain special cases can easily be settled in the affirmative. For example, if $a = -1$ and $m = 4$, then we are asking for infinitely many primes $p \equiv -1 \pmod{4}$, and now we can mimic Euclid: If there are only finitely many such primes, say $p_1, \ldots, p_k$, form the number $N := 4p_1 \cdots p_k - 1$. Since $N \equiv -1 \pmod{4}$, it must have at least one prime divisor $p \equiv -1 \pmod{4}$. But p cannot be any of $p_1, \ldots, p_k$, and we have a contradiction. A similar argument works when $a = -1$ and $m = 3$.

The general case of our problem is much more difficult. It turns out that whenever $\gcd(a, m) = 1$, there *are* infinitely many primes $p \equiv a \pmod{m}$. This was proved by Dirichlet in 1837, by analytic methods. (One can view his argument as a far-reaching generalization of Euler's proof that the sum of the reciprocals of the primes diverges.) We will give a proof of Dirichlet's theorem in Chapter 4.

For now we content ourselves with some special cases of Dirichlet's theorem that follow from algebraic arguments. We noted above that an easy variant of Euclid's proof shows that there are infinitely many primes p for which the residue class of p avoids the trivial subgroup of the unit group $(\mathbf{Z}/4\mathbf{Z})^\times$, and similarly for $(\mathbf{Z}/3\mathbf{Z})^\times$. As observed by A. Granville (unpublished), we have the following general result:

Theorem 1.16. *If H is a proper subgroup of $(\mathbf{Z}/m\mathbf{Z})^\times$, then there are infinitely many primes p for which $p \bmod m \notin H$.*

Proof. Let $\mathcal{P}$ be the set of primes p for which $p \bmod m \notin H$, and let $\mathcal{P}'$ be the set of such primes not dividing m. Assuming $\mathcal{P}$ is finite, let P be the product of the elements of $\mathcal{P}'$. Fix an integer a coprime to m with $a \bmod m \notin H$ (which is possible since H is a *proper* subgroup), and then choose a positive integer n satisfying the congruences $n \equiv 1 \pmod{P}$ and $n \equiv a \pmod{m}$. (Such a choice of n is possible by the Chinese remainder theorem.) Since n is coprime to mP, none of its prime divisors can come from $\mathcal{P}$, so that every prime p dividing n must be such that $p \bmod m \in H$. But since H is closed under multiplication, this implies that $n \bmod m \in H$. This contradicts the choice of a. □

If $F(T)$ is a nonzero polynomial with integer coefficients, we say that the prime p is a *prime divisor* of F if p divides $F(n)$ for some integer n. The following useful lemma is due to Schur [**Sch12**]:

Lemma 1.17. *Let $F(T)$ be a nonconstant polynomial with integer coefficients. Then F has infinitely many prime divisors.*

Proof. If $F(0) = 0$, then every prime is a prime divisor of F. So we can assume that the constant term c_0 (say) of $F(T)$ is nonzero. Then $F(c_0T) = c_0G(T)$ for some nonconstant polynomial $G(T)$ with constant term 1. It is enough to show that G has infinitely many prime divisors. Suppose that $p_1, \dots, p_k$ is a list of prime divisors of G. For m sufficiently large, we have $|G(mp_1 \cdots p_k)| > 1$, so that there must be some prime p dividing $G(mp_1 \cdots p_k)$. Then p is a prime divisor of G and p is not equal to any of the p_i, since $G(mp_1 \cdots p_k) \equiv 1 \pmod{p_i}$ for each $1 \le i \le k$. So no finite list of prime divisors of G can be complete. □

For example, let $F(T) = T^2 + 1$. If p divides $n^2 + 1$, then $n^2 \equiv -1 \pmod{p}$, and so either $p = 2$ or $p \equiv 1 \pmod{4}$. So Lemma 1.17 implies that there are infinitely many primes $p \equiv 1 \pmod{4}$. Similarly, if $F(T) = T^2+T+1$, then any prime divisor p of F is such that $p \equiv 1 \pmod{3}$, and so there are infinitely many primes $p \equiv 1 \pmod{3}$. Combining this with our earlier results, we have proved Dirichlet's theorem for all progressions modulo 3 and modulo 4.

These examples are special cases of the following construction: Recall that the *mth cyclotomic polynomial* is defined by

$$\Phi_m(T) = \prod_{\substack{1 \le k \le m \\ \gcd(k,m)=1}} \left(T - e^{2\pi i k/m}\right),$$

i.e., $\Phi_m(T)$ is the monic polynomial in $\mathbf{C}[T]$ whose roots are precisely the primitive mth roots of unity, each occurring with multiplicity 1. For example, $\Phi_4(T) = T^2 + 1$ and $\Phi_3(T) = T^2 + T + 1$.

We will apply Lemma 1.17 to Φ_m to deduce that there are infinitely many primes $p \equiv 1 \pmod{m}$. To apply Lemma 1.17, we need that the coefficients of $\Phi_m(T)$ are not merely complex numbers, but in fact integers.

Lemma 1.18. *For each positive integer m, the polynomial $\Phi_m(T)$ has integer coefficients.*

Proof. For each m we have the factorization

$$T^m - 1 = \prod_{d \mid m} \Phi_d(T). \tag{1.11}$$

To see this, note that $T^m - 1 = \prod_{\zeta^m=1}(T - \zeta)$. Since the set of mth roots of unity is the disjoint union of the primitive dth roots of unity, taken over

those d dividing m, we have (1.11). Applying Möbius inversion to (1.11) yields

$$\Phi_m(T) = \prod_{d|m} \Big(T^d - 1\Big)^{\mu(m/d)} = \frac{\prod_{d|m,\,\mu(m/d)=1} \big(T^d - 1\big)}{\prod_{d|m,\,\mu(m/d)=-1} \big(T^d - 1\big)} = \frac{F}{G},$$

say. Now F and G are *monic* polynomials in $\mathbf{Z}[T]$ with $G \neq 0$, and so we can write

$$F = GQ + R, \tag{1.12}$$

where $Q, R \in \mathbf{Z}[T]$ and $\deg R < \deg Q$. Of course (1.12) remains valid over $\mathbf{C}[T]$ and expresses in that ring one result of division by G. But we know that over $\mathbf{C}[T]$, we have $F = G\Phi_m$, so that G goes into F with no remainder. By the uniqueness of quotient and remainder in the division algorithm for polynomials, we must have $R = 0$ above. Consequently, $\Phi_m = F/G = Q \in \mathbf{Z}[T]$. □

Lemma 1.19. *If p is a prime divisor of Φ_m, then either $p \mid m$ or $p \equiv 1 \pmod{m}$.*

Proof. If p is a prime divisor of Φ_m, then p divides $\Phi_m(n)$ for some integer n. Since the cyclotomic polynomials have integer coefficients, it follows from (1.11) that $p \mid \prod_{d|m} \Phi_d(n) = n^m - 1$, so that the order of n modulo p is a divisor of m.

Suppose now that p does not divide m. We claim that in this case, m is the precise order of n modulo p. Thus m divides $p - 1$, whence $p \equiv 1 \pmod{m}$. To prove the claim, suppose for the sake of contradiction that $f < m$ is the exact order of n mod p. Then f is a proper divisor of m. Moreover, p divides $n^f - 1 = \prod_{e|f} \Phi_e(n)$, so that p divides $\Phi_e(n)$ for some $e \mid f$. Hence the residue class n mod p is a zero of both $\Phi_e(T)$ and $\Phi_m(T)$. The polynomials Φ_e and Φ_m both appear in the factorization (1.11) of $T^m - 1$, so that $T^m - 1$ has a zero of order ≥ 2 over $\mathbf{Z}/p\mathbf{Z}$. But $T^m - 1$ has no multiple roots over $\mathbf{Z}/p\mathbf{Z}$, since $T^m - 1$ has no roots in common with its derivative mT^{m-1}. □

Since only finitely many primes divide m, Lemmas 1.17 and 1.19 have the following corollary:

Corollary 1.20. *For each natural number m, there are infinitely many primes $p \equiv 1 \pmod{m}$.*

This proof of Corollary 1.20 is essentially due to Wendt [**Wen95**].

How far can one take this algebraic approach? The following result is due to Schur (op. cit.).

★ **Theorem 1.21.** *Let m be a positive integer and let H be a subgroup of $(\mathbf{Z}/m\mathbf{Z})^\times$. There is a nonconstant polynomial $F(T) \in \mathbf{Z}[T]$ with the following property: Every prime divisor p of F, with finitely many exceptions, satisfies $p \bmod m \in H$. Consequently, there are infinitely many primes p for which $p \bmod m \in H$.*

When H is the trivial subgroup we have just seen that $F := \Phi_m$ satisfies the conclusion of Theorem 1.21.

Schur gave an elementary proof of Theorem 1.21 requiring only familiarity with the theory of finite fields. A less elementary proof is outlined in Exercise 20. When m is a prime number, Theorem 1.21 is contained in the results of Chapter 2 (see, in particular, Theorem 2.15).

Suppose that a and m satisfy $a^2 \equiv 1 \pmod{m}$, where $a \not\equiv 1 \pmod{m}$. Applying Theorem 1.21 to the 2-element subgroup of $(\mathbf{Z}/m\mathbf{Z})^\times$ generated by $a \bmod m$, we obtain a polynomial $F(T)$ all of whose prime divisors (with finitely many exceptions) satisfy either $p \equiv 1 \pmod{m}$ or $p \equiv a \pmod{m}$. Schur showed (op. cit.) that if there is a single, suitably large prime $p \equiv a \pmod{m}$, then the polynomial F he constructs cannot have all (or even all but finitely many) of its prime divisors from the progression $1 \bmod m$. (See the first example below for an illustration of how this works.) So F must have infinitely many prime divisors $p \equiv a \pmod{m}$.

Since Dirichlet's theorem is true, there is always a suitably large prime $p \equiv a \pmod{m}$ to be used in Schur's argument, and so in principle, it is possible to give a purely algebraic proof of Dirichlet's theorem for any progression $a \bmod m$ satisfying $a^2 \equiv 1 \pmod{m}$. Moreover, this is best possible in the following sense:

★ **Theorem 1.22** (Murty [**Mur88, MT06**])**.** *Suppose m is a positive integer. If F is a nonconstant polynomial with the property that every prime divisor p of F, with finitely many exceptions, satisfies $p \equiv 1 \pmod{m}$ or $p \equiv a \pmod{m}$, then $a^2 \equiv 1 \pmod{m}$.*

The proof of Theorem 1.22 rests on rather deep results in algebro-analytic number theory. The principal tool required is the *Chebotarev density theorem*, which is a far-reaching generalization of Dirichlet's theorem. See [**SL96**] for a down-to-earth discussion of Chebotarev's result.

Example. As an easy example of Schur's method, consider the problem of showing that there are infinitely many primes $p \equiv 3 \pmod{8}$. We start by taking $F(T) := T^2 + 2$. From the elementary theory of quadratic residues we have that each odd prime divisor of $F(T)$ satisfies $p \equiv 1$ or $3 \pmod{8}$. Now we observe that there is at least one prime in the residue class $3 \pmod{8}$,

namely 11. We replace T by $4T + 3$ and so obtain from F the polynomial

$$G(T) = F(4T + 3) = 16T^2 + 24T + 11 = 8(2T^2 + 3T) + 11.$$

Then every prime divisor of G belongs to either the residue class 1 mod 8 or 3 mod 8. Moreover, for each positive integer n, there is at least one prime $p \equiv 3 \pmod 8$ for which $p \mid G(n)$, since $G(n) \equiv 3 \pmod 8$. We will show that G (and hence also F) must have infinitely many prime divisors from the residue class 3 mod 8. Suppose otherwise, and let $p_1, p_2, \ldots, p_k$ be a complete list of the prime divisors $p \equiv 3 \pmod 8$ of G. For each p_i, choose an integer n_i for which $G(n_i) \not\equiv 0 \pmod{p_i}$. (This is possible since G has at most two roots modulo p_i.) If n is a positive integer chosen by the Chinese remainder theorem to satisfy $n \equiv n_i \pmod{p_i}$ for all $1 \le i \le k$, then $G(n)$ cannot be divisible by any of $p_1, \ldots, p_k$. So $G(n)$ must have a prime divisor from the residue class 3 mod 8 other than $p_1, \ldots, p_k$, a contradiction.

Example. Since every integer a coprime to 24 satisfies $a^2 \equiv 1 \pmod{24}$, it is in principle possible to give an algebraic proof of Dirichlet's theorem for progressions with common difference 24. The details in this case have been completely worked out by Bateman & Low [**BL65**]. We leave to the reader the task of showing that 24 is the largest modulus m with the property that $a^2 \equiv 1 \pmod m$ for each a coprime to m.

9.2. Hypothesis H.

> I do not mean to deny that there are mathematical truths, morally certain, which defy and will probably to the end of time continue to defy proof, as, *e.g.*, that every indecomposable polynomial function must represent an infinitude of primes. – J. J. Sylvester [**Syl88**]

There are two natural directions we might head in if we hope to generalize Dirichlet's result: First, we might inquire about simultaneous prime values of several linear polynomials. One has to be careful here, of course. For example, we cannot hope that there are infinitely many n for which both n and $n + 1$ are prime, because one of these two numbers is always even! However, if instead of n and $n + 1$ we consider n and $n + 2$, then this obstruction disappears, and we arrive at the following famous conjecture:

Conjecture 1.23 (Twin prime conjecture)**.** *There are infinitely many natural numbers n for which both n and $n + 2$ are prime.*

Alternatively, we might accept the restriction of working with a single polynomial, but hope to treat polynomials of higher degree. The following conjecture of Euler, which appears in correspondence with Goldbach, fits nicely into this framework:

Conjecture 1.24 (Euler)**.** *There are infinitely many natural numbers n for which n^2+1 is prime.*

Similarly, it seems reasonable to conjecture that our old friend, T^2+T+41, represents infinitely many primes. Once again, formulating conjectures of this type requires some care; if n^2+1 or n^2+n+41 is replaced by n^2+n+2, then the statement corresponding to Euler's conjecture is false, since n^2+n+2 is always even.

Suppose more generally that $F_1(T),\dots,F_r(T)\in\mathbf{Z}[T]$ are nonconstant polynomials, each with positive leading coefficient. We can ask when it is the case that $F_1(n),\dots,F_r(n)$ are simultaneously prime for infinitely many natural numbers n. Evidently if this is to be the case, then we must suppose that each F_i is irreducible over $\mathbf{Z}$. The example of $r=2$ and $F_1(T)=T$, $F_2(T)=T+1$ shows that this is not sufficient, as does the example of $r=1$ and $F_1(T)=T^2+T+2$. What goes wrong in these examples is that there is a *local obstruction*: If we put $G(T):=\prod_{i=1}^r F_i(T)$, then $G(n)$ is always even. In 1958, Schinzel conjectured (see [**SS58**]) that these are the only remaining obstructions to be accounted for:

Conjecture 1.25 (Schinzel's "Hypothesis H")**.** *Suppose $F_1(T),\dots,F_r(T)\in\mathbf{Z}[T]$ are nonconstant and irreducible and that each F_i has a positive leading coefficient. Put $G(T):=\prod_{i=1}^r F_i(T)$, and suppose that there is no prime p which divides $G(n)$ for every integer n. Then $F_1(n)$, $F_2(n)$, …, $F_r(n)$ are simultaneously prime for infinitely many natural numbers n.*

The hypothesis on G is necessary: Suppose that p is a (fixed) prime which divides $G(n)$ for each n. Then p divides some $F_i(n)$ for each n. But for large n, each $F_i(n)>p$, and so for large n, some $F_i(n)$ is composite.

The twin prime conjecture corresponds to choosing $r=2$, $F_1(T)=T$, and $F_2(T)=T+2$ in Hypothesis H. Taking instead $r=1$ and $F_1(T)=T^2+1$, we recover Euler's Conjecture 1.24. Despite substantial attention, both the twin prime conjecture and Conjecture 1.24 remain open. Even more depressing, no case of Hypothesis H has ever been shown to hold except when $r=1$ and $F_1(T)$ is linear, when Hypothesis H reduces to Dirichlet's theorem!

Sieve methods, which we introduce in Chapter 6, can be used to obtain certain approximations to Hypothesis H. We give two examples: A theorem of Chen [**Che73**] asserts that there are infinitely many primes p for which $p+2$ is either prime or the product of two primes. And Iwaniec [**Iwa78**] has shown that there are infinitely many n for which n^2+1 is either prime or the product of two primes. (This latter result applies also to n^2+n+41, and in fact to any quadratic obeying the conditions of Hypothesis H.)

10. Primes and composites in other sequences

We conclude by discussing the occurrence of primes in other sequences of interest. Results in this area are rather thin on the ground, and so we content ourselves with a smattering of problems and results meant to showcase our collective ignorance.

One sequence that has received much attention is that of the *Mersenne numbers* $2^n - 1$. The occurrence of primes in this sequence has long been of interest in view of Euclid's result that if $2^n - 1$ is prime, then $2^{n-1}(2^n - 1)$ is a perfect number. (Here a number is called *perfect* if it is the sum of its proper divisors.) Since $2^d - 1$ divides $2^n - 1$ whenever d divides n, for $2^n - 1$ to be prime it is necessary that n be prime. At first glance it appears that $2^p - 1$ is often prime; 7 of the first 10 primes p have this property. However, the tide quickly turns: Of the 78498 primes p up to 10^6, only 31 yield primes. As of February 2009, there are 46 known primes of the form $2^p - 1$, the largest corresponding to $p = 43112609$. It is not clear from this data whether or not we should expect infinitely many primes of this form, but probabilistic considerations to be discussed in Chapter 3 suggest that we should:

Conjecture 1.26. *For infinitely many primes p, the number $2^p - 1$ is prime.*

Unfortunately, this conjecture seems far beyond reach. In fact, we know disturbingly little about the numbers $2^p - 1$; perhaps the most striking illustration of this is that even the following modest conjecture remains unproved:

Conjecture 1.27. *For infinitely many primes p, the number $2^p - 1$ is composite.*

We may also change the "$-$" sign to a "$+$" and consider primes of the form $2^n + 1$. Since $2^d + 1$ divides $2^n + 1$ when n/d is odd, we see that $2^n + 1$ can be prime only if n is a power of 2. This leads us to consider the *Fermat numbers* $F_m = 2^{2^m} + 1$. The attentive reader will recall that these numbers appeared already in Goldbach's proof of Theorem 1.1. For $m = 0, 1, 2, 3$, and 4, the numbers F_m are prime:

$$2^{2^0} + 1 = 3, \quad 2^{2^1} + 1 = 5, \quad 2^{2^2} + 1 = 17, \quad 2^{2^3} + 1 = 257, \quad 2^{2^4} + 1 = 65537.$$

Fermat was intuitively certain that F_m is prime for all $m \geq 0$, and expressed this belief in letters to his contemporaries; but in 1732 Euler discovered the factorization

$$2^{2^5} + 1 = 641 \cdot 6700417.$$

It is now known that F_m is composite for $5 \le m \le 32$, and (for the same probabilistic reasons alluded to above) it is widely believed that F_m is composite for every $m \ge 5$. So much for intuition! Despite this widespread belief, the following conjecture appears intractable:

Conjecture 1.28. *The Fermat number F_m is composite for infinitely many natural numbers m.*

Similarly, for each even natural number a, one can look for primes in the sequence $a^{2^m}+1$. Again we believe that there should be at most finitely many, but again the analogue of Conjecture 1.28 seems impossibly difficult! Indeed, there is no specific even number a for which we can prove that $a^{2^m}+1$ is composite infinitely often. This is a somewhat odd state of affairs in view of the following amusing theorem of Schinzel [**Sch63**]:

Theorem 1.29. *Suppose that infinitely many of the Fermat numbers F_j are prime. If $a > 1$ is an integer not of the form 2^{2^r} (where $r \ge 0$), then $a^{2^m}+1$ is composite for infinitely many natural numbers m.*

Proof. Fix an integer $a > 1$ not of the form 2^{2^r}. Let M_0 be an arbitrary positive integer. We will show that $a^{2^m}+1$ is composite for some $m \ge M_0$.

Let F_j be a prime Fermat number not dividing $a(a^{2^{M_0}}-1)$. Since a is coprime to F_j, Fermat's little theorem implies that

$$a^{F_j-1} = a^{2^{2^j}} \equiv 1 \pmod{F_j}.$$

Since $F_j \nmid a^{2^{M_0}}-1$, we must have $M_0 < 2^j$. So we can write

$$\begin{aligned} a^{F_j-1}-1 &= a^{2^{2^j}}-1 \\ &= (a^{2^{M_0}}-1)(a^{2^{M_0}}+1)(a^{2^{M_0+1}}+1)(a^{2^{M_0+2}}+1)\cdots(a^{2^{2^j-1}}+1). \end{aligned}$$

Since F_j divides $a^{F_j-1}-1$ but not $a^{2^{M_0}}-1$, it must be that F_j divides $a^{2^m}+1$ for some $M_0 \le m < 2^j$. We cannot have $a^{2^m}+1 = F_j$, since a is not of the form 2^{2^r}, and so $a^{2^m}+1$ is composite. □

In connection with Fermat-type numbers the following result of Shapiro & Sparer [**SS72**] merits attention (cf. [**Sha83**, Theorem 5.1.5]). It shows (in particular) that the doubly exponential sequences $a^{2^m}+1$ are unusually difficult to handle among sequences of the same general shape:

★ **Theorem 1.30.** *Suppose a, b, and c are integers, and that $a, b > 1$. If c is odd, then*

$$a^{b^m}+c$$

is composite for infinitely many $m \in \mathbf{N}$, except possibly in the case when a is even, $c = 1$, and $b = 2^k$ for some $k \ge 1$. If c is even, there are infinitely many such m except possibly when a is odd and $c = 2$.

The reader should note that the Shapiro–Sparer paper contains several other attractive results on composite numbers in various sequences.

We close this section by considering the sequence of shifted factorials $n! + 1$. Here we can easily obtain infinitely many composite terms, since Wilson's theorem implies that $(p-1)! + 1$ is composite for each $p > 3$. The following pretty theorem of Schinzel [**Sch62b**] generalizes this result:

Theorem 1.31. *Let α be a positive rational number. Then there are infinitely many n for which $\alpha \cdot n! + 1$ is composite.*

Lemma 1.32. *Let p be a prime and let r and s be positive integers. Then for $0 \le i \le p-1$, we have*

$$p \mid si! + (-1)^{i+1} r \iff p \mid r(p-1-i)! + s.$$

Proof. By Wilson's theorem,

$$\begin{aligned} -1 \equiv (p-1)! &= (p-1)(p-2)\cdots(p-i)(p-i-1)! \\ &\equiv (-1)^i i!(p-1-i)! \pmod{p}, \end{aligned}$$

so that $(p-1-i)!i! \equiv (-1)^{i+1} \pmod{p}$. Since p and $(p-1-i)!$ are coprime,

$$\begin{aligned} p \mid si! + (-1)^{i+1} r &\iff p \mid s(p-1-i)!i! + (-1)^{i+1} r(p-1-i)! \\ &\iff p \mid (-1)^{i+1} s + (-1)^{i+1} r(p-1-i)! \\ &\iff p \mid s + r(p-1-i)!. \end{aligned}$$

□

Proof of Theorem 1.31. Write $\alpha = r/s$, where r and s are relatively prime positive integers. Assume $l \in \mathbf{N}$ and $l \ge r/2$. Then $(4l)!\alpha^{-1}$ is an integer divisible by both 4 and r. Since $4 \mid (4l)!\alpha^{-1}$, we can choose a prime $p_l \equiv -1 \pmod{4}$ with

$$p_l \mid (4l)!\alpha^{-1} - 1.$$

Because $r \mid (4l)!\alpha^{-1}$, necessarily $p_l \nmid r$. Since

$$p_l \mid r\left((4l)!\alpha^{-1} - 1\right) = s(4l)! - r, \tag{1.13}$$

we must have $p_l > 4l$. From Lemma 1.32 (with $i = 4l$) and (1.13), we find that

$$p_l \mid r(p_l - 4l - 1)! + s. \tag{1.14}$$

Since $p_l \nmid r$, (1.14) implies that $p_l \nmid s$, and so

$$p_l \mid N_l := \alpha(p_l - 4l - 1)! + 1$$

whenever N_l is an integer. This happens for all large l: Indeed, from (1.14) we have $N_l \ge p_l/s \ge 4l/s$, so that $N_l \to \infty$ with l, which is only possible if $p_l - 4l - 1 \to \infty$ with l. But N_l is an integer whenever $p_l - 4l - 1 \ge s$.

Finally, notice that for large l, we cannot have $p_l = N_l$, since $p_l \equiv -1 \pmod{4}$ while $N_l \equiv 1 \pmod{4}$. Thus N_l is a composite integer of the

form $\alpha \cdot n! + 1$. Letting $l \to \infty$, we obtain infinitely many composite numbers of this form. □

Notes

Most of the proofs discussed for the infinitude of the primes may be found in [**Dic66**, Chapter XVIII] or [**Nar00**, §1.1]. For other compilations, see [**Rib96**, Chapter 1], [**FR07**, Chapter 3], and [**Moh79**]. An amusing version of Euclid's proof, couched in the language of nonstandard analysis, is presented in [**Gol98**, pp. 57–58]. Additional elementary proofs of the stronger result that $\sum 1/p$ diverges may be found in [**Bel43**], [**Mos58**], and the survey [**VE80**].

The following result of Matijasevich and Putnam provides an interesting contrast to Goldbach's theorem (Theorem 1.9): *There is a polynomial with integral coefficients such that the set of primes coincides with the set of positive values assumed by this polynomial, as the variables range over the nonnegative integers.* (An explicit example of such a polynomial, in 26 variables, was produced by Jones et al. [**JSWW76**].) Yet upon inspection we realize we are once again looking at a result that properly belongs not to number theory but to computability theory (or logic); an analogous statement is true if we replace the set of primes with any *listable set*. Here a set of positive integers S is called *listable* if there is a computer program which, when left running forever, outputs precisely the elements of S. A very approachable introduction to this circle of ideas is Matijasevich's article [**Mat99**]; for complete details see [**Mat93**].

In connection with the results of §8, we cannot resist pointing out the remarkable identity

$$e^{\pi\sqrt{163}} = 262537412640768743.99999999999925\ldots,$$

which shows that $e^{\pi\sqrt{163}}$ is very nearly an integer. We sketch the explanation, which comes from the theory of modular functions; for details one may consult [**Cox89**, §11]. Every lattice $L \subset \mathbf{C}$ has a so-called j-invariant $j(L)$, and $j(L_1) = j(L_2)$ precisely when L_1 and L_2 are homothetic, i.e., when one can be obtained from the other by rotation and scaling. We view j as a function on the upper half-plane $\{z \in \mathbf{C} : \Im(z) > 0\}$ by defining $j(\tau)$ as $j(L)$, where L is the lattice spanned by 1 and τ. It turns out that j is then holomorphic on the upper half-plane. Moreover, since 1 and τ determine the same lattice as 1 and $\tau + 1$, we have $j(\tau) = j(\tau + 1)$. This shows that $j(\tau)$ is holomorphic as a function of $q = e^{2\pi i\tau}$ in the punctured disc $0 < |q| < 1$, and so j has a Laurent expansion. It turns out that this expansion starts

$$j(\tau) = \frac{1}{q} + 744 + 196884q + \cdots,$$

so that $j(\tau) \approx 1/q + 744$ for small q. Now for the coup de grâce: One can show that if K is an imaginary quadratic field with integral basis $1, \tau$, then $j(\tau)$ is an algebraic integer of degree exactly $h(K)$, the class number of K. In particular, if K has class number 1, then $j(\tau)$ is a rational integer. The main theorem of §8 implies that $K = \mathbf{Q}(\sqrt{-163})$ has class number 1, and so $j(\tau) \in \mathbf{Z}$ for $\tau = \frac{1+i\sqrt{163}}{2}$. This value of τ corresponds to $q = -1/\exp(\pi\sqrt{163})$, so that

$$e^{\pi\sqrt{163}} \approx j(\tau) - 744 \in \mathbf{Z}.$$

We remark that $e^{\pi\sqrt{163}}$ is actually transcendental, as may be deduced from the following theorem of Gelfond and Schneider (noting that $e^{\pi\sqrt{163}} = (-1)^{i\sqrt{163}}$): *If α and β are algebraic numbers, where $\alpha \neq 0$ and β is irrational, then α^β is transcendental.* Here "α^β" stands for $\exp(\beta \log \alpha)$, and any nonzero value of $\log \alpha$ is permissible. For a proof of the Gelfond–Schneider result, see, e.g., [**Hua82**, §17.9].

There are many sequences not discussed in §10 where it would be of interest to decide if they contain infinitely many primes, or composites. For example, fix a nonintegral rational number $\alpha > 1$, and consider the sequence of numbers $\lfloor \alpha^n \rfloor$. Whiteman has conjectured that this sequence always contains infinitely many primes. If we drop the rationality condition, then from a very general theorem of Harman [**Har97**] we have that each sequence $\lfloor \alpha^n \rfloor$ contains infinitely many primes as long as $\alpha > 1$ avoids a set of measure zero. (Of course since the rational numbers have measure zero, this has no direct consequence for Whiteman's conjecture.) Very little is known about the sequences considered by Whiteman. For the particular numbers $\alpha = 3/2$ and $\alpha = 4/3$, Forman & Shapiro [**FS67**] present ingenious elementary arguments showing that the sequence $\lfloor \alpha^n \rfloor$ contains infinitely many composite numbers. Some extensions of their results have been obtained by Dubickas & Novikas [**DN05**]; e.g., these authors prove that if $\xi > 0$ and $\alpha \in \{2, 3, 4, 6, 3/2, 4/3, 5/4\}$, then the sequence $\lfloor \xi\alpha^n \rfloor$ contains infinitely many composites.

Exercises

1. (Harris [**Har56**]) Let b_0, b_1, b_2 be positive integers with b_0 coprime to b_2. Define A_k for $k = 0, 1$ and 2 as the numerator when the finite continued fraction

$$b_0 + \cfrac{1}{b_1 + \cfrac{1}{\ddots + \cfrac{1}{b_k}}}$$

is put in lowest terms. For $k = 3, 4, \ldots$, inductively define b_k and A_k by

$$b_k = A_0 A_1 \cdots A_{k-3}$$

and A_k by the rule given above. Prove that the A_i form an increasing sequence of pairwise coprime positive integers.

2. (Aldaz & Bravo [**AB03**]) Let p_i denote the ith prime. Euclid's argument shows that for each r, there is a prime in the interval $(p_r, \prod_1^r p_i + 1]$. Prove that the number of primes in the (smaller) interval $(p_r, \prod_2^r p_i + 1]$ tends to infinity with r. *Suggestion:* With $P = \prod_2^r p_i$, show that $P - 2, P - 2^2, \ldots, P - 2^k$ are > 1 and pairwise coprime for fixed k and large r; then choose a prime factor of each.

3. (Chowdhury [**Cho89**]) It is trivial that for $n \geq 1$, the number $n! + 1$ has a prime divisor exceeding n. Show that for $n \geq 6$, the same holds for each of the numbers $n! + k$, where $2 \leq k \leq n$.

4. (Hegyvári [**Heg93**]) Suppose $a_1 < a_2 < a_3 < \ldots$ is an increasing sequence of natural numbers for which $\sum 1/a_i$ diverges. Show that the real number $\alpha := 0.a_1 a_2 a_3 \ldots$ formed by concatenating the decimal expansions of the a_i is irrational. In particular, $0.235711131719\ldots$ is irrational. *Hint:* First show that every finite sequence of decimal digits appears in the expansion of α.

Remark. Suppose that in place of our divergence hypotheses, we assume that for each fixed $\theta < 1$, the number of $a_i \leq x$ exceeds x^θ for all sufficiently large x. Then Copeland & Erdős [**CE46**] have proved that the number α constructed above is *normal* (in base 10); in other words, not only does every finite digit string appear in the expansion of α, but each string of length k appears with the expected frequency 10^{-k}.

5. (Euler) In courses in complex analysis, it is often proved that $\sin x$ possesses the following Weierstrass factorization (valid for all $x \in \mathbf{C}$):

$$\sin x = x \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2\pi^2}\right); \tag{1.15}$$

see, e.g., [**Pri01**] for a short, direct proof of this identity. A proof using only real-variable methods appears in [**Kob84**, Chapter II].

(a) Starting from (1.15), show that

$$x \cot x = 1 - 2\sum_{m=1}^{\infty} \zeta(2m) \frac{x^{2m}}{\pi^{2m}},$$

where ζ denotes the Euler-Riemann zeta function. *Hint:* Take the logarithmic derivative of both sides.

(b) Computing by hand the first few coefficients in the Taylor series for $x \cot x$ about $x = 0$, check that $\zeta(2) = \pi^2/6$ and $\zeta(4) = \pi^4/90$.

6. (J. D. Dixon) We outline Dixon's proof [**Dix62**] that π is not the root of a polynomial over $\mathbf{Z}$ of degree ≤ 2. The method is that employed by Niven to show π is irrational (see [**Niv47**]). Suppose for the sake of contradiction that π is a root of $P(T) = aT^2 + bT + c$, where a, b and c are integers, not all vanishing.

Given a polynomial $f(T) \in \mathbf{R}[T]$, define

$$F(T) := f(T) - f^{(2)}(T) + f^{(4)}(T) - f^{(6)}(T) + \cdots. \tag{1.16}$$

Then $F(T) \in \mathbf{R}[T]$. View F as a function of a real variable x.

(a) Check that

$$\frac{d}{dx}\left(F'(x)\sin x - F(x)\cos x\right) = f(x)\sin(x),$$

and conclude that

$$\int_0^{\pi} f(x)\sin x\, dx = F(\pi) + F(0). \tag{1.17}$$

(b) With n a positive integer to be chosen shortly, let f be the polynomial

$$f(T) := \frac{1}{n!}P(T)^{2n}(P(T) - P(0))^{2n}.$$

Show that the left-hand side of (1.17) is strictly between 0 and 1 if n is sufficiently large.

We now fix such an n and derive a contradiction by showing that the right-hand side of (1.17) is an integer.

(c) Show that $f^{(r)}(0) = f^{(r)}(\pi) = 0$ for all $0 \leq r < 2n$.

(d) If e and r are nonnegative integers and r is even, show that there is an expansion of the form

$$\frac{d^r}{dx^r}\left(P(x)^e\right) = \sum_{j=r/2}^{r} c_j j! \binom{e}{j} P(x)^{e-j}$$

for certain integers c_j.

(e) Use the result of part (d) to show that if e is a nonnegative integer and $r \geq 2n$ is even, then $\frac{1}{n!}\frac{d^r}{dx^r}(P(x)^e)$ is a polynomial in $P(x)$ with integer coefficients. Conclude that $f^{(r)}(0)$ and $f^{(r)}(\pi)$ are integers.

(f) Referring back to definition (1.16), deduce that $F(\pi) + F(0) \in \mathbf{Z}$.

7. In this exercise we present a proof similar to that of J. Hacks (on p. 8) but relying on the irrationality of π in place of π^2. Let

$$\chi(n) = \begin{cases} (-1)^{(n-1)/2} & \text{if } 2 \nmid n, \\ 0 & \text{otherwise.} \end{cases}$$

a) Show that $\chi(n)$ is a completely multiplicative function, i.e.,

$$\chi(ab) = \chi(a)\chi(b)$$

for every pair of positive integers a, b.

b) Assume that there are only finitely many primes. Show that for every $s > 0$,

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

c) Take $s = 1$ and obtain a contradiction to the irrationality of π. You may assume that $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$.

8. Say that a natural number n is *squarefull* if $p^2 \mid n$ whenever $p \mid n$, i.e., if every prime showing up in the factorization of n occurs with multiplicity > 1. Every perfect power is squarefull, but there are many other examples, such as $864 = 2^5 \cdot 3^3$. Using Theorem 1.2, show that $\sum' n^{-1}$ converges to $\frac{\zeta(2)\zeta(3)}{\zeta(6)}$, where the $'$ indicates that the sum is restricted to squarefull n. Determine the set of real α for which $\sum' n^{-\alpha}$ converges.

9. (Continuation) Show that every squarefull number has a unique representation in the form u^2v^3, where u and v are positive integers with v squarefree. Deduce that for $x \geq 1$,

$$\sum_{\substack{n \leq x \\ n \text{ squarefull}}} 1 = \frac{\zeta(3/2)}{\zeta(3)} x^{1/2} + O(x^{1/3}).$$

10. (Ramanujan) Assuming $\zeta(2) = \pi^2/6$ and $\zeta(4) = \pi^4/90$, show that

$$\sum{}' \frac{1}{n^2} = \frac{9}{2\pi^2},$$

where the $'$ indicates that the sum ranges over positive squarefree integers n with an odd number of prime divisors.

11. (Cf. Porubský [**Por01**]) If R is a commutative ring, its *Jacobson radical* $J(R)$ is the intersection of all of its maximal ideals. Show that

$$J(R) = \{x \in R : 1 - xy \text{ is invertible for all } y \in R\}.$$

Deduce that if R is an integral domain with finitely many units, then $J(R) = \{0\}$. Use this to prove that if R is a principal ideal domain with finitely many units, then either R is a field or R contains an infinite set of pairwise nonassociated primes.

12. By carefully examining the proof of Theorem 1.10, show that the theorem remains correct when $A = 1$, provided that in condition (ii) we replace "prime" with "prime or equal to 1".

13. Suppose that $a_1 < a_2 < a_3 < \cdots$ is an increasing sequence of natural numbers, and put $A(x) := \sum_{a_i \le x} 1$. Prove that if $(\log x)^{-k} A(x) \to \infty$ for each fixed k, then infinitely many primes p divide some a_i. Use this to give another proof of Lemma 1.17.

14. Prove the following theorem of Bauer [**Bau06**]:

Theorem. *If $F(T) \in \mathbf{Z}[T]$ is a nonconstant polynomial with at least one real root, then for every $m \ge 3$, there exist infinitely many prime divisors p of F with $p \not\equiv 1 \pmod{m}$.*

Proceed by showing that each of the following conditions on F is sufficient for the conclusion of the theorem to hold:

(a) F has a positive leading coefficient and constant term -1.
(b) F has a positive leading coefficient and negative constant term.
(c) F has a positive leading coefficient and $F(a) < 0$ for some $a \in \mathbf{Z}$.
(d) F has a positive leading coefficient and $F(a) < 0$ for some $a \in \mathbf{Q}$.
(e) F has a positive leading coefficient and $F(a) < 0$ for some $a \in \mathbf{R}$.
(f) F has a positive leading coefficient and $F(a) = 0$ for some $a \in \mathbf{R}$.
Hint for (f): Reduce to the case when F has no multiple roots.

15. Let F be a field of characteristic not dividing m. By carefully examining the proof of Lemma 1.19, show that the roots of $\Phi_m(T)$ in the algebraic closure of F are precisely the primitive mth roots of unity there, and that all these roots are simple.

16. (Continuation; Kronecker [**Kro88**], Dirichlet, Bauer [**Bau06**]) Define $\Phi_m(X,Y)$ as the homogenization of $\Phi_m(T)$, so that

$$\Phi_m(X,Y) = \prod_{\substack{\zeta^m=1 \\ \zeta^j\neq 1 \text{ if } 1\le j<m}} (X-\zeta Y).$$

(a) Suppose $m>2$. Show that $\Phi_m(X+Y,X-Y) = G_m(X,Y^2)$ for some polynomial G_m (say) with integer coefficients. Show also that $\prod_{d|m} d^{\mu(m/d)}$ is the coefficient of $X^{\varphi(m)}$ in $\Phi_m(X+Y,X-Y)$.

(b) Let F be a field of characteristic not dividing m. Suppose s is a nonsquare integer, and let $\sqrt{s}$ denote a fixed square root of s from the algebraic closure of F. Show that the roots of $G_m(T,s)\in\mathbf{Z}[T]$ in the algebraic closure of F are precisely the elements

$$\sqrt{s}\frac{\zeta+1}{\zeta-1},$$

where ζ runs through the primitive mth roots of unity.

(c) Suppose s is as in (b), and let p be a prime for which $p\nmid 2ms$. Show that p is a prime divisor of $G_m(T,s)$ if and only if $p\equiv\left(\frac{s}{p}\right) \pmod m$.

(d) Show that if $p\equiv -1 \pmod 4$ is a prime divisor of $G_m(T,-1)$ which does not divide m, then $p\equiv -1 \pmod m$. Use Exercise 14 to show that $G_m(T,-1)$ has infinitely many such prime divisors, and deduce that there are infinitely many primes $p\equiv -1 \pmod m$.

17. (M. Hirschhorn [**Hir02**]) Let $p_1<p_2<p_3<\cdots$ denote the sequence of *odd* primes.

(a) Let $N\in\mathbf{N}$. Prove that the number of odd positive integers $\le N$ which can be written in the form $p_1^{e_1}\cdots p_k^{e_k}$ does not exceed

$$\prod_{i=1}^{k}\left(\frac{\log N}{\log p_i}+1\right) < (\log(p_k N))^k < \sqrt{2k!}\sqrt{p_k N}.$$

Hint: Show that $(\log u)^k u^{-1/2}\le (2k/e)^k$ whenever $u\ge 1$. Now invoke the inequality $m!\ge (m/e)^m$, valid for every integer $m\ge 0$.

(b) Supposing that $p_1,\ldots,p_k$ exist (i.e., that there are at least k odd primes), prove that p_{k+1} exists and satisfies $p_{k+1}\le 4(2k!)p_k+1$.

18. Suppose that A is a commutative monoid (written multiplicatively) and that P is a system of generators for A, so that each element of A can be written in the form $\prod_{p\in P} p^{e_p}$, where each $e_p\ge 0$ and only finitely many of the e_p are nonzero. (We do *not* require that this representation be unique.) Suppose also that there is a function $\|\cdot\|: A\to\mathbf{N}$ with the following two properties:

(a) $\|\cdot\|$ respects multiplication, i.e., $\|ab\| = \|a\|\|b\|$ for all $a,b\in A$.

(b) For some real number x_0 and constants $c_1, c_2 > 0$, we have

$$c_1 x \le \#\{a \in A : \|a\| \le x\} \le c_2 x \quad \text{for all} \quad x > x_0. \tag{1.18}$$

Prove that P is infinite, and that in fact $\sum_{p \in P} \frac{1}{\|p\|}$ diverges.

19. (Continuation)
    (a) For each nonzero Gaussian integer α put $\|\alpha\| = |\alpha|^2$. Show that $\sum_\pi \|\pi\|^{-1}$ diverges, where the sum is over all Gaussian primes π. Deduce that $\sum_{p \equiv 1 \pmod 4} p^{-1}$ diverges, where the sum is over rational primes $p \equiv 1 \bmod 4$.
    (b) For each nonzero polynomial $F(T) \in \mathbf{F}_q[T]$, put $\|F\| := q^{\deg F}$. Show that $\sum \|P\|^{-1}$ diverges, where P ranges over the irreducible elements of $\mathbf{F}_q[T]$.

20. This exercise outlines a proof of Theorem 1.21 via algebraic number theory. Let m be a positive integer, and let ζ be a primitive mth root of unity. Put $K = \mathbf{Q}(\zeta_m)$, and identify $\mathrm{Gal}(K/\mathbf{Q})$ with $(\mathbf{Z}/m\mathbf{Z})^\times$. Let H be a subgroup of $(\mathbf{Z}/m\mathbf{Z})^\times$, and let $L \subset K$ be the fixed field of H.
    (a) Say that two sets of rational primes $\mathcal{P}_1$ and $\mathcal{P}_2$ eventually coincide if their symmetric difference is finite; in this case we write $\mathcal{P}_1 \doteq \mathcal{P}_2$. Prove that $\mathcal{P}_1 \doteq \mathcal{P}_2$, where $\mathcal{P}_1$ is the set of primes for which $p \bmod m \in H$ and $\mathcal{P}_2$ is the set of primes which split completely in L. *Hint:* If p is a prime not dividing m, analyze how the Frobenius element of p in $\mathrm{Gal}(K/\mathbf{Q})$ behaves upon restriction to L.
    (b) Let θ be an algebraic integer for which $L = \mathbf{Q}(\theta)$. Let F be the minimal polynomial of θ. Prove that $\mathcal{P}_2$, and hence also $\mathcal{P}_1$, eventually coincides with the set of prime divisors of F. *Hint:* $L/\mathbf{Q}$ is Galois, so an unramified rational prime splits completely in L exactly when it has a degree 1 prime factor; now apply the Kummer-Dedekind theorem.

21. (Pólya [**Pól21**]; see also [**MS00**]) Suppose that a and b are nonzero integers and $a \neq \pm 1$. Let $\mathcal{P}$ be the set of primes for which the exponential congruence $a^k \equiv b \pmod p$ has a positive integer solution k. In other words, $\mathcal{P}$ is the set of primes which divide some term of the sequence

$$a - b, \quad a^2 - b, \quad a^3 - b, \quad a^4 - b, \ldots.$$

This exercise outlines a proof that $\mathcal{P}$ is always an infinite set.

We may suppose that b is not a power of a, as otherwise $\mathcal{P}$ contains every prime. We assume for the sake of contradiction that $\mathcal{P}$ is finite.

(a) For each $p \in \mathcal{P}$ and each $k \ge 1$, define integers $v_{p,k} \ge 0$ by writing

$$a^k - b = \pm \prod_{p \in P} p^{v_{p,k}}.$$

For each $p \in \mathcal{P}$, set $v_p := \sup_{k\geq 1} v_{p,k}$. We let $\mathcal{P}_1 := \{p \in \mathcal{P} : v_p < \infty\}$ and we put $\mathcal{P}_2 := \mathcal{P} \setminus \mathcal{P}_1$. Show that if $p \in \mathcal{P}_2$, then $p \nmid a$.

(b) Suppose $p \in \mathcal{P}_2$, and let l_p be the order of a modulo p. (This exists by part (a).) Define e_p so that $p^{e_p} \parallel a^{l_p} - 1$. Show that if k is a positive integer for which $p^{e_p+1} \mid a^k - b$, then k belongs to a fixed residue class modulo p.

(c) Show that there is an infinite arithmetic progression of integers k which avoid all the residue classes mod p ($p \in \mathcal{P}_2$) determined in (b). Prove that $a^k - b$ is uniformly bounded for such k, contradicting that $|a^k - b| \to \infty$ as $k \to \infty$.

Remark. In the opposite direction, one can ask when the set $\mathcal{P}$ defined above omits infinitely many primes. Using the Chebotarev density theorem, Schinzel [**Sch60**] has shown that this holds unless $b = a^k$ for some nonnegative integer k. See also [**MS00**].

22. (Křížek et al. [**KLS02**]) Let $F_n = 2^{2^n} + 1$ be the nth Fermat number. Suppose $N \in \mathbf{N}$.

(a) Show that there are fewer than 2^N distinct prime divisors of the product $F_0 \cdots F_{N-1}$.

(b) Show that for each $x > 0$, the number of primes $p \leq x$ which divide F_n for some $n \geq N$ is at most $x/2^{N+1}$.

(c) Making an appropriate choice of N, deduce from (a) and (b) that there are $\ll \sqrt{x}$ primes $p \leq x$ which divide a term of the sequence $F_0, F_1, F_2, \ldots$.

(d) Deduce that if $\lambda > 1/2$, then $\sum' p^{-\lambda} < \infty$, where the $'$ indicates that the sum is restricted to primes dividing at least one Fermat number. When $\lambda = 1$, this confirms a conjecture of Golomb [**Gol55**].

23. (Erdős & Turán [**ET34**]) For $n > 1$, write $P(n)$ for the largest prime factor of n. In this exercise we show that if S is an infinite set of natural numbers, then

$$\{P(a+b) : a, b \in S\} \quad \text{is unbounded.} \tag{1.19}$$

For each prime p, let v_p be the *p-adic valuation*, defined so that $p^{v_p(n)} \parallel n$ for every natural number n.

(a) Let S be an arbitrary infinite set of natural numbers. Show that for each odd prime p, we can determine an infinite subset $S' \subset S$ with the property that whenever $a, b \in S'$,

$$v_p(a+b) = \min\{v_p(a), v_p(b)\}. \tag{1.20}$$

Hint: First treat the case when no element of S is divisible by p.

(b) Suppose, for the sake of contradiction, that S is infinite but (1.19) fails. Using part (a), argue that we may assume (1.20) holds for

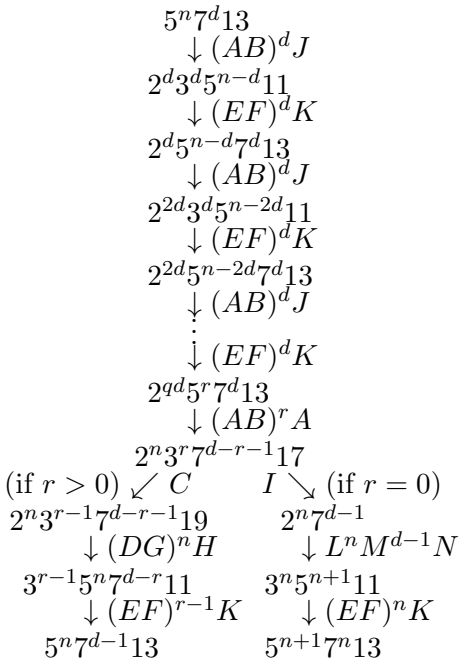

Figure 4. The action of Conway's prime-producing machine when started with 5^n7^d13, where $0 < d < n$. The variables q and d are defined by the division algorithm: $n = dq + r$ where $0 \le r < d$.

every pair $a, b \in S$ and every odd prime p. We make this assumption from now on.

(c) Now argue that $v_2(a) = v_2(b)$ for every pair of elements $a, b \in S$. Thus, dividing through by a suitable power of 2, we may (and do) assume that all the elements of S are odd.

(d) Finally, show that for each pair of elements $a, b \in S$, we have

$$a + b = 2^{v_2(a+b)} \prod_{p>2} p^{\min\{v_p(a),v_p(b)\}}.$$

Show that this equation leads to a contradiction if a and b are chosen to be congruent modulo 4.

24. Figure 4, based on Conway's article [**Con87**], describes the action of Conway's prime-producing machine. Decipher this figure and explain how it proves Theorem 1.8. For a more detailed explanation of the workings of Conway's prime-producing machine, see Guy's expository article [**Guy83**].

25. (Schinzel [**Sch62a**]) In 1857, Bunyakovsky conjectured [**Bun57**] that if $F(T) \in \mathbf{Z}[T]$ is an irreducible polynomial with positive leading coefficient and D is the largest positive integer dividing $F(n)$ for each $n \in \mathbf{Z}$, then $F(n)/D$ is prime for infinitely many natural numbers n. Show that this would follow from Hypothesis H.

26. (Granville; see, e.g., [**Mol97**, Theorem 2.1]) Assume Hypothesis H. Show that for every natural number N_0, one can find a positive integer A with the property that $n^2 + n + A$ assumes prime values for all $0 \le n \le N_0$. *Hint:* Apply Hypothesis H to the N_0 linear polynomials $T, T + (1^2 + 1), T + (2^2 + 2), \dots, T + (N_0^2 + N_0)$.

27. (Schinzel & Sierpiński [**SS58**]) Assume Hypothesis H. Show that if $n > 1$ and r is a positive integer divisible by all primes $p \le n$, then there are infinitely many arithmetic progressions of length n and common difference r consisting of consecutive primes.

    Remark. The weaker claim that there are arbitrarily long arithmetic progressions of primes was recently proved in a technical tour de force by Green & Tao [**GT08**], using ideas borrowed from ergodic theory (and several other fields). For some striking elementary consequences of the Green–Tao result, see [**Gra08a**].

28. (Cf. Chang & Lih [**CL77**]) Show that for every $N \in \mathbf{N}$, there is a polynomial $F(T) \in \mathbf{Z}[T]$ for which $\{F(k)\}_{k=0}^N$ is a sequence of $N + 1$ distinct primes. *Hint:* For $0 \le k \le N$, put $c_k(T) = \prod_{0 \le i \le N, i \ne k} (T - i)$. Using Corollary 1.20, choose integers $r_0, r_1, \dots, r_N$ for which $\{1 + r_k c_k(k)\}_{k=0}^N$ is a sequence of $N + 1$ distinct primes. Put $F(T) := 1 + \sum_{i=0}^N r_i c_i(T)$.

29. (Clement [**Cle49**], Cucurezeanu [**Cuc68**]) Let k and n be integers with $n > k \ge 2$. Suppose that n has no prime divisors $< k$. Show that n and $n + k$ are simultaneously prime if and only if

$$k \cdot k!((n-1)! + 1) + (k! - (-1)^k)n \equiv 0 \pmod{n(n+k)}.$$

30. (Shanks [**Sha64**]) Let $F(z) = \sum_{n=0}^\infty z^{n(n+1)/2}$ and define

$$G(z) := (F(z) - 1)^2 - (F(z) - 1).$$

    Prove that there are infinitely many primes of the form $\frac{n^2+1}{2}$ (with $n \in \mathbf{N}$) if and only if the power series expansion of G has infinitely many negative coefficients.

31. Suppose $p \equiv 3 \pmod 4$ is prime. Prove that if $2p + 1$ is also prime, then $2p + 1 \mid 2^p - 1$. Deduce that Hypothesis H implies Conjecture 1.27.

32. (Selfridge; cf. [**Erd50b**]) Let $n \in \mathbf{N}$. Show that $78557 \cdot 2^n + 1$ is divisible by some prime number from the set $\{3, 5, 7, 13, 19, 37, 73\}$. In particular, $78557 \cdot 2^n + 1$ is always composite.

Table 1. Mann-Shanks criterion: Columns containing only bold entries are indexed by prime numbers.

	0	1	**2**	**3**	4	**5**	6	**7**	8	9	10	**11**	12	**13**
0	1													
1			**1**	**1**										
2					1	**2**	1							
3							1	**3**	**3**	1				
4									1	**4**	6	**4**	1	
5											1	**5**	**10**	**10**
6													1	**6**

33. (Louisiana State University Problem Solving Group [**PSG02**]) Prove that $5^{4n} + 5^{3n} + 5^{2n} + 5^n + 1$ is composite for every natural number n.

    If you know some algebraic number theory, establish the following generalization: If $q > 1$ is a squarefree natural number with $q \equiv 1 \pmod 4$, then $\Phi_q(q^n)$ is composite for every natural number n.

    Hint (due to J. A. Rouse): $q^n - \zeta$ is a difference of squares in $\mathbf{Z}[\zeta]$, where ζ denotes a primitive qth root of unity.

34. Table 1 illustrates a primality criterion discovered by Mann & Shanks [**MS72**]: Place the rows of Pascal's triangle in an infinite table, where the zeroth row (consisting of the single element 1) is placed in column 0. Each successive row is shifted two units right. An element of the nth row is written in boldface when it is divisible by n. Then the column number is prime exactly when all entries in its column are written in boldface. Prove this!

35. (Hayes [**Hay65**]) Suppose that R is a principal ideal domain with infinitely many prime ideals. Show that every nonconstant polynomial A over R can be written as the sum of two irreducible polynomials of the same degree as A. *Hint:* Arrange for both summands to satisfy the Eisenstein criterion with respect to the same prime.

Chapter 6

Sieve Methods

> Brun's [sieve] method ... is perhaps our most powerful elementary tool in number theory. – P. Erdős [**Erd65**]

1. Introduction

1.1. The sieve of Eratosthenes. Granville has pointed out [**Gra95**] that ancient Greek mathematics produced two results in prime number theory that have proved of first importance in subsequent thought. The first is Euclid's proof of the infinitude of the primes, which was discussed in Chapter 1. The second is the sieve of Eratosthenes.

Eratosthenes' method allows one to determine the primes not exceeding x assuming only knowledge of the primes not exceeding $\sqrt{x}$. In this procedure one begins with a list of all positive integers in the interval $[2, x]$. For each prime $p \leq \sqrt{x}$, we cross out all the multiples of p on the list; the numbers remaining are exactly the primes in the interval $(\sqrt{x}, x]$. We illustrate this with $x = 30$, sieving by the primes $2, 3$, and 5:

	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~~

This procedure is remarkable not only insofar as it gives a fast algorithm for listing primes, but also in that it suggests the useful viewpoint of the primes as the integers surviving a "sieving process".

1.2. Legendre's formula. Let us attempt to count how many integers remain after Eratosthenes' sieving procedure is carried out. More generally, let us count the number of positive integers up to x remaining after the

deletion (or "sifting out") of the multiples of all primes not exceeding z, where z is a parameter at our disposal (in Eratosthenes' sieve, $z = \sqrt{x}$). We use $\pi(x, z)$ to denote this quantity, i.e.,

$$\pi(x, z) := \#\{n \le x : p \mid n \Rightarrow p > z\}.$$

Then for every $z > 0$,

$$\pi(x) \le z + \pi(x, z),$$

and

$$\pi(x, x^{1/2}) = \pi(x) - \pi(\sqrt{x}) + 1.$$

Our estimate of $\pi(x, z)$ proceeds in several stages. We begin with the total number $\lfloor x \rfloor$ of positive integers not exceeding x, and then for each prime $p_1 \le z$ we subtract the number of multiples of p_1:

$$\lfloor x \rfloor - \sum_{p_1 \le z} \left\lfloor \frac{x}{p_1} \right\rfloor.$$

This counts correctly those n with at most one prime divisor $p \le z$, but those n with two or more prime factors $p \le z$ have been subtracted off twice. Hence, we add these back in, to obtain our next approximation,

$$\lfloor x \rfloor - \sum_{p_1 \le z} \left\lfloor \frac{x}{p_1} \right\rfloor + \sum_{p_1 < p_2 \le z} \left\lfloor \frac{x}{p_1 p_2} \right\rfloor.$$

But now those integers divisible by three primes $p \le z$ have been added back in too many times; for instance, if n has exactly three prime divisors not exceeding z, it is counted with weight $1 - 3 + 3 > 0$. Thus we should subtract a term corresponding to the integers divisible by three primes $p \le z$; we would then find ourselves needing to add a term corresponding to integers divisible by four such p, etc. Continuing in this manner, we are led to the formula

$$\pi(x, z) = \lfloor x \rfloor - \sum_{p_1 \le z} \left\lfloor \frac{x}{p_1} \right\rfloor + \cdots + (-1)^r \sum_{p_1 < \cdots < p_r \le z} \left\lfloor \frac{x}{p_1 \cdots p_r} \right\rfloor, \tag{6.1}$$

where $r = \pi(z)$. If we set

$$P := \prod_{p \le z} p,$$

we can put (6.1) in the alternative form

$$\pi(x, z) = \sum_{d \mid P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor. \tag{6.2}$$

This reasoning, due to Legendre, can be tightened into a proof of (6.1). For the time being we assume (6.1), postponing a rigorous justification to §3, where we will establish a more general result.

1.3. Consequences. We now have an exact formula for $\pi(x,z)$. Unfortunately this exact formula is a bit unsatisfying because it leaves the most natural question unanswered: How large is $\pi(x,z)$?

What does our formula (6.2) have to say about this question? Sums involving the greatest-integer function are generally difficult to work with, so we drop the greatest integer signs in (6.2) and transfer the incurred error to a separate sum:

$$\pi(x,z) = x\sum_{d|P}\frac{\mu(d)}{d} + \sum_{d|P}\mu(d)\left(\left\lfloor\frac{x}{d}\right\rfloor - \frac{x}{d}\right).$$

The first sum can be written as the product $\prod_{p\le z}(1-1/p)$. The second sum (which we view as the error term) is bounded in absolute value by $2^{\pi(z)}$, since there are $2^{\pi(z)}$ divisors d of P, and for each of these the corresponding summand has absolute value at most 1. Thus

$$\pi(x,z) = x\prod_{p\le z}\left(1-\frac{1}{p}\right) + O\left(2^{\pi(z)}\right). \tag{6.3}$$

How useful is this estimate? Suppose first that z is fixed while x is tending to infinity; then the error term in (6.3) is $O_z(1)$ and we obtain the asymptotic formula $\pi(x,z) \sim x\prod_{p\le z}(1-1/p)$. The same asymptotic estimate holds if z is not fixed, but instead tends to infinity with x sufficiently slowly.

Whenever $z = z(x) \to \infty$, Mertens' theorem implies that

$$x\prod_{p\le z}\left(1-\frac{1}{p}\right) \sim e^{-\gamma}\frac{x}{\log z} \quad (x\to\infty). \tag{6.4}$$

If $z = z(x)$ also satisfies $z \le \log x$ once x is sufficiently large, then the O-term in (6.3) is $\ll 2^z \le x^{\log 2}$, which is of smaller order than $x/\log z$. Consequently, $\pi(x,z) \sim e^{-\gamma}x/\log z$. Taking $z = \log x$, we obtain that

$$\pi(x) \le \pi(x,\log x) + \log x \le (e^{-\gamma}+o(1))\frac{x}{\log\log x}, \tag{6.5}$$

which provides another proof that the set of primes has density zero.

We have yet to treat the case corresponding to Eratosthenes' sieve, that of $z = \sqrt{x}$. In this case the "main term" in (6.3) is

$$x\prod_{p\le x^{1/2}}\left(1-\frac{1}{p}\right) \sim 2e^{-\gamma}\frac{x}{\log x} = (1.229\ldots)\frac{x}{\log x}. \tag{6.6}$$

Unfortunately our bound of $2^{\pi(\sqrt{x})}$ for the "error term" dwarfs the value of this main term. (For example, by Chebyshev's results, $2^{\pi(\sqrt{x})} > 2^{\sqrt[3]{x}}$ for large x, and $2^{\sqrt[3]{x}}$ grows faster than any fixed power of x.) So (6.3) does

not give us the asymptotic formula $\pi(x, x^{1/2}) \sim x \prod_{p \le x^{1/2}} (1 - 1/p)$. And in fact, by the prime number theorem,

$$\pi(x, x^{1/2}) = \pi(x) - \pi(\sqrt{x}) + 1 \sim x/\log x, \tag{6.7}$$

so that *it is not even true* that $\pi(x, x^{1/2}) \sim x \prod_{p \le x^{1/2}} (1 - 1/p)$. This points to a limitation of our method for approximating $\pi(x, z)$; in §1.5 we will discuss to what extent difficulties of this sort can be overcome.

1.4. General sieving situations. The problem treated in the last section is of the following form: Given a finite sequence[1] of integers $\mathcal{A}$ and a finite set of primes $\mathcal{P}$, estimate the number $S(\mathcal{A}, \mathcal{P})$ of terms of $\mathcal{A}$ divisible by no prime $p \in \mathcal{P}$. For example, if

$$\mathcal{A} := \{n \le x\} \quad \text{and} \quad \mathcal{P} := \{p \le z\}, \tag{6.8}$$

then $S(\mathcal{A}, \mathcal{P})$ is what we have been calling $\pi(x, z)$.

Many problems in number theory fit into this framework. For example, suppose $x, z > 0$. Set

$$\mathcal{A} := \{n(n+2) : n \le x\}, \quad \mathcal{P} := \{p \le z\}. \tag{6.9}$$

If both n and $n + 2$ are prime, then either $n \le z$ or both n and $n + 2$ have only prime factors exceeding z. Consequently,

$$\pi_2(x) \le S(\mathcal{A}, \mathcal{P}) + z.$$

Moreover, n and $n + 2$ are both prime if all of their prime factors exceed $\sqrt{x+2}$. So if we take $z = \sqrt{x+2}$, then

$$0 \le \pi_2(x) - S(\mathcal{A}, \mathcal{P}) \le z. \tag{6.10}$$

Estimates for $S(\mathcal{A}, \mathcal{P})$ are thus intimately connected with the quantitative version of the twin prime conjecture introduced in Chapter 3, §5.

In order to prove any general theorems on the size of $S(\mathcal{A}, \mathcal{P})$, it is necessary to make some further assumptions. We will assume that $\mathcal{A}$ has "approximate" length X and that divisibility by distinct primes $p \in \mathcal{P}$ constitute "approximately" independent events, each occurring with "approximate" probability $\alpha(p)$. (All of this will be made precise in §2.) In this case, it is natural to expect that

$$S(\mathcal{A}, \mathcal{P}) \approx X \prod_{p \in \mathcal{P}} (1 - \alpha(p)). \tag{6.11}$$

From our perspective in this chapter, the goal of sieve theory is to quantify and then to justify such approximations, in as wide a range of circumstances as possible.

[1]The sole rationale for insisting that $\mathcal{A}$ be a sequence instead of a set is to ensure that duplicate elements are counted with multiplicity. Notationally, we will treat $\mathcal{A}$ below as if it were a set, but the reader should understand that $\mathcal{A}$ is actually a multiset.

In the classical situation described by (6.8), it is reasonable to approximate the count of natural numbers $n \le x$ by x and the probability that such an integer is divisible by p by $1/p$. Then (6.11) is the guess that $\pi(x,z) \approx x\prod_{p\le z}(1-1/p)$. We have seen that when z is constant or slow-growing, this approximation holds as an asymptotic formula, but that for $z=\sqrt{x}$ (the case originally of interest), the approximation is off by a constant factor. Nevertheless, (6.11) is still correct if read as the assertion that both sides have the same order of magnitude.

For another example, consider the situation described by (6.9). Again the length of $\mathcal{A}$ is approximately x. The probability that a term of $\mathcal{A}$ is divisible by the prime p is approximately $\nu(p)/p$, where $\nu(2)=1$ and $\nu(p)=2$ for $p>2$. (So that $\nu(p)$ is the number of solutions to $n(n+2)\equiv 0 \pmod p$.) The prediction (6.11) is that

$$S(\mathcal{A},\mathcal{P}) \approx x\prod_{p\le z}\left(1-\frac{\nu(p)}{p}\right). \tag{6.12}$$

If $z=z(x)\to\infty$ as $x\to\infty$, it is an easy deduction from Mertens' theorem (given in §3.2 below) that

$$x\prod_{p\le z}\left(1-\frac{\nu(p)}{p}\right) \sim 2C_2e^{-2\gamma}\frac{x}{(\log z)^2}, \tag{6.13}$$

where $C_2=\prod_{p>2}\left(1-(p-1)^{-2}\right)$ is the twin prime constant. Arguing as in §1.3, one can show that $S(\mathcal{A},\mathcal{P})$ is asymptotic to the right-hand side of (6.13) when z is quite small (say $z\le \frac12\log x$) and x tends to infinity. Probably no method can establish the same if $z=\sqrt{x+2}$; indeed, referring back to (6.10), we see that this would contradict the quantitative form of the twin prime conjecture (Conjecture 3.18). Note that even if $z=\sqrt{x+2}$, we still expect that the right-hand side of (6.13) has the same order of magnitude as $\pi_2(x)$; it is only off from what is conjecturally correct by a factor of $(2e^{-\gamma})^2$; cf. Exercise 28.

1.5. Legendre, Brun, and Hooley; oh my! We have already stated that the goal of sieve theory, for us, is to quantify and to justify estimates of the form

$$S(\mathcal{A},\mathcal{P}) \approx X\prod_{p\in\mathcal{P}}(1-\alpha(p)).$$

We can get a feel for the respective power of the three sieve methods considered in this chapter if we reflect on what they say about the particular estimate $\pi(x,z)\approx x\prod_{p\le z}(1-1/p)$ corresponding to our initial problem. As noted above, Legendre's method of successive approximation can be developed to show that this approximation is asymptotically correct when $z=\log x$. The first improvement on Legendre's methods, known as Brun's

pure sieve, shows that this remains true in a wider range: We need only assume that $z = z(x) \to \infty$ subject to the inequality $z(x) \le x^{1/(10 \log\log x)}$ (for large x). In particular, choosing z as large as possible and referring to (6.4), we find that

$$\pi(x) \le \pi(x,z) + z \ll \frac{x}{\log x} \log\log x, \tag{6.14}$$

which is considerably sharper than (6.5).

The final method to be developed in this chapter, known as the Brun–Hooley sieve, allows one to obtain upper and lower bounds for $\pi(x,z)$ when z is as large as a (small) fixed power of x. From its upper bound we recover Chebyshev's estimate $\pi(x) \ll x/\log x$. (But one should take this with a grain of salt — in the derivation, we require the results of Mertens, which in turn rest on those of Chebyshev.) The lower bound aspect is also interesting, and allows one to deduce bounds of the shape $\pi(x, x^{1/1000}) \gg x/\log x$. Such a lower bound does *not* translate into a lower bound on $\pi(x)$; but because an integer up to x all of whose prime factors exceed $x^{1/1000}$ can have at most 1000 prime factors, it *does* give us a lower bound on the number of 1000-almost primes up to x. Here an *r-almost prime* is an integer with no more than r prime divisors, counting multiplicity.

All of this might seem a bit silly because we have known the correct order of magnitude of $\pi(x)$ since Chapter 3. But the general sieve framework is rather flexible, and therein lies the potential of this approach. We have already seen that sieve methods can be adapted to yield information about the twin prime conjecture. Developing these ideas, Brun used his pure sieve to prove (in analogy with (6.14)) that

$$\pi_2(x) \ll \frac{x}{(\log x)^2} (\log\log x)^2. \tag{6.15}$$

This is off by a factor of $(\log\log x)^2$ from the conjectured order of magnitude, but it still has profound implications. One consequence is that $\sum_p 1/p$, restricted to primes p which belong to a twin prime pair, is either a finite sum or a convergent infinite series.

Brun succeeded in removing the unwanted factor $(\log\log x)^2$ from (6.15) but required a rather complicated combinatorial apparatus to do so. We will reach the same goal by making use of simple ideas of Hooley. The same method will allow us to prove the following two deep theorems of Brun ([**Bru20**]; see [**Wan84**] for an English translation), approximations to the twin prime and Goldbach conjectures respectively:

- There are infinitely many pairs of 9-almost primes $n, n+2$.
- Every large even integer N is a sum of two 9-almost primes.

In the next section we formally introduce some notions and notation arising in the general sieving situation. We then discuss the first sieve method, that of Eratosthenes–Legendre. This is just a general version of Legendre's method of successive approximation, seen above. After giving a few elementary applications, we turn to a discussion of Brun's pure sieve. This method gets its name from its origin in the purely combinatorial observation that the approximations in Legendre's method are alternately over and underestimates. Brun's pure sieve is much more powerful than Legendre's method, which we illustrate by proving the aforementioned theorem of Brun on the sum of the reciprocals of the twin primes. We then describe Hooley's elegant and surprisingly powerful "almost-pure" sieve, basing our treatment on Hooley's original article [**Hoo94**] and the exposition of Ford & Halberstam [**FH00**]. We conclude the chapter with a striking application of sieve methods to the Goldbach problem, found by Schnirelmann.

2. The general sieve problem: Notation and preliminaries

> Probability is not a notion of pure mathematics, but of philosophy or physics. – G. H. Hardy & J. E. Littlewood [**HL23**]

The general sieve problem takes the following form: Given a finite sequence of integers $\mathcal{A} = \{a_i\}$ and a finite set of primes $\mathcal{P}$, estimate the quantity

$$S(\mathcal{A}, \mathcal{P}) := \#\{a \in \mathcal{A} : \gcd(a, P) = 1\},$$

where $P := \prod_{p \in \mathcal{P}} p$.

In many situations, the sieving set $\mathcal{P}$ is obtained by truncating an infinite set of primes at a point z. Consequently, it is expedient to allow the set $\mathcal{P}$ to be infinite and to introduce special notation indicating that we sieve only by those primes $p \in \mathcal{P}$ with $p \le z$. We therefore define

$$S(\mathcal{A}, \mathcal{P}, z) := \#\{a \in \mathcal{A} : \gcd(a, P(z)) = 1\},$$

where

$$P(z) := \prod_{\substack{p \in \mathcal{P} \\ p \le z}} p.$$

Hence $S(\mathcal{A}, \mathcal{P}, z) = S(\mathcal{A}, \mathcal{P} \cap [2, z])$.

We use the notation A_d to denote the number of terms of $\mathcal{A}$ divisible by d, i.e.,

$$A_d := \#\{a \in \mathcal{A} : d \mid a\}.$$

The letter X denotes an approximation to the size of $\mathcal{A}$. We assume the existence of a multiplicative function α taking values in $[0, 1]$ for which

$$A_d = X\alpha(d) + r(d) \tag{6.16}$$

for each $d \mid P$ (or each $d \mid P(z)$, as the case may be). In practice, we *choose* X and α, and we *define* $r(d)$, for $d \mid P$, so that (6.16) holds.

3. The sieve of Eratosthenes–Legendre and its applications

3.1. The principle of inclusion-exclusion. Any rigorous study of sieve methods begins with the following fundamental result from enumerative combinatorics:

Theorem 6.1 (Principle of inclusion-exclusion)**.** *Let X be a nonempty, finite set of N objects, and let $P_1, \dots, P_r$ be properties that elements of X may have. For each subset $I \subset \{1, 2, \dots, r\}$, let $N(I)$ denote the number of elements of X that have each of the properties indexed by the elements of I. Then with N_0 denoting the number of elements of X with none of these properties, we have*

$$\begin{aligned} N_0 &= \sum_{k=0}^{r} (-1)^k \sum_{\substack{I \subset \{1,2,\dots,r\} \\ |I|=k}} N(I) \\ &= \sum_{I \subset \{1,2,\dots,r\}} (-1)^{|I|} N(I). \end{aligned} \tag{6.17}$$

Proof. Suppose $x \in X$ has exactly l of the properties $P_1, \dots, P_r$. If $l = 0$, then x is counted only once in (6.17), in the term $N(\emptyset)$. On the other hand, if $1 \le l \le r$, then the number of k-element sets $I \subset \{1, 2, 3, \dots, r\}$ for which x is counted in $N(I)$ is exactly $\binom{l}{k}$, and the total weight with which x is counted is

$$\sum_{k=0}^{l} (-1)^k \binom{l}{k} = (1-1)^l = 0,$$

by the binomial theorem. □

3.2. A first sieve result. The principle of inclusion-exclusion can be applied immediately to the situation of §2:

Theorem 6.2 (Sieve of Eratosthenes–Legendre)**.**

$$S(\mathcal{A}, \mathcal{P}) = X \prod_{p \in \mathcal{P}} (1 - \alpha(p)) + \sum_{d \mid P} \mu(d) r(d).$$

Proof. Let $p_1, \dots, p_r$ be a list of the primes in $\mathcal{P}$, and for each $1 \le i \le r$, let P_i be the property of being divisible by p_i. For every $d \mid P$, there are

$X\alpha(d) + r(d)$ terms $a \in \mathcal{A}$ divisible by d. So by the principle of inclusion-exclusion, the number of $a \in \mathcal{A}$ divisible by none of the primes of $\mathcal{P}$ is

$$\sum_{k=0}^{r}(-1)^k \sum_{\substack{I\subset\{1,2,\dots,r\}\\ |I|=k}} N(I) = \sum_{k=0}^{r}(-1)^k \sum_{\substack{d|P\\ \omega(d)=k}} A_d$$

$$= \sum_{k=0}^{r} \sum_{\substack{d|P\\ \omega(d)=k}} \mu(d)\left(X\alpha(d) + r(d)\right) = X\sum_{d|P}\mu(d)\alpha(d) + \sum_{d|P}\mu(d)r(d)$$

$$= X\prod_{p\in\mathcal{P}}(1-\alpha(p)) + \sum_{d|P}\mu(d)r(d). \quad \square$$

Example. Let $\mathcal{A} = \{n \le x\}$ and let $\mathcal{P} = \{p \le z\}$. Then $S(\mathcal{A},\mathcal{P})$ is what we referred to in the introduction as $\pi(x,z)$. For each d, we have $A_d = \lfloor x/d \rfloor$. So if we set $X = x$ and $\alpha(d) = 1/d$, and define $r(d)$ by (6.16), then $r(d) = -\{x/d\}$. In particular, $|r(d)| \le 1$ for each d. So applying Theorem 6.2 with this choice of X and α, we recover the estimate (6.3), which was derived in a nonrigorous fashion in the introduction.

Example. Let $\mathcal{A} = \{n(n+2) : n \le x\}$ and let $\mathcal{P} = \{p \le z\}$. As pointed out in (6.10), for this choice of A and $\mathcal{P}$, $S(\mathcal{A},\mathcal{P})$ is related to the twin-prime counting function $\pi_2(x)$. In order to decide on a reasonable choice of X and α in this situation, let us attempt to get a feel for the numbers A_d. The condition that d divides $n(n+2)$ is a condition on n modulo d, so we set

$$\nu(d) := \#\{n \pmod d : n(n+2) \equiv 0 \pmod d\}. \tag{6.18}$$

Then each block of d consecutive integers contains precisely $\nu(d)$ solutions of the congruence $n(n+2) \equiv 0 \pmod d$. Hence $A_d \approx (x/d)\nu(d)$, which suggests that we choose $X = x$ and $\alpha(d) = \nu(d)/d$. (Note that ν, and hence α, is multiplicative by the Chinese remainder theorem.) In fact, since the interval $[1,x]$ contains the first $\lfloor x/d \rfloor$ blocks of d consecutive natural numbers, and is contained in the first $\lceil x/d \rceil$ such blocks, with this choice of X and α we have

$$\lfloor x/d \rfloor \nu(d) \le A_d \le \lceil x/d \rceil \nu(d), \quad \text{so that} \quad |r(d)| = |A_d - x\nu(d)/d| \le \nu(d).$$

We now apply Theorem 6.2, with z a function of x tending slowly to infinity. The coefficient of $X = x$ in the main term of Theorem 6.2 is

$$\prod_{p\le z}(1-\alpha(p)) = \frac{1}{2}\prod_{2<p\le z}\left(1-\frac{2}{p}\right) = \left(2\prod_{2<p\le z}\frac{1-\frac{2}{p}}{\left(1-\frac{1}{p}\right)^2}\right)\prod_{p\le z}\left(1-\frac{1}{p}\right)^2.$$

Estimating the last product here by Mertens' theorem, we find that the main term is asymptotic to

$$2C_2e^{-2\gamma}x/(\log z)^2.$$

The error term is bounded by

$$\sum_{d|P}\nu(d)=\prod_{p\le z}(1+\nu(p))\le 3^{\pi(z)}\le 3^z,$$

which is negligible in comparison with the main term if (e.g.) $z=\frac{1}{2}\log x$. We will return to this example in §4.4.

Example. Here is an example different from those alluded to in the introduction, due to Nagell [**Nag22**]. Let $\pi_{T^2+1}(x)$ denote the number of $n\le x$ for which n^2+1 is prime. Let $\mathcal{A}=\{n^2+1:n\le x\}$ and let $\mathcal{P}$ be the set of all primes. Then for any choice of positive numbers x and z, we have

$$\pi_{T^2+1}(x)\le S(\mathcal{A},\mathcal{P},z)+z^{1/2}. \tag{6.19}$$

The congruence $n^2+1\equiv 0 \pmod d$ is satisfied precisely when n falls into one of $\nu(d)$ (say) residue classes modulo d. As in the preceding example, this suggests we take $X=x$ and $\alpha(d)=\nu(d)/d$; with this choice of X and α, the numbers $r(d)$ defined by (6.16) satisfy $|r(d)|\le\nu(d)$.

Now $\nu(2)=1$, while if p is an odd prime, $\nu(p)=0$ or 2, depending on whether $p\equiv 3 \pmod 4$ or $1 \pmod 4$, respectively. So by (6.19) and Theorem 6.2, if $x>0$ and $z\ge 2$, then

$$\begin{aligned}\pi_{T^2+1}(x)&\le S(\mathcal{A},\mathcal{P},z)+z^{1/2}\\&\le\frac{1}{2}x\prod_{\substack{p\le z\\ p\equiv 1 \pmod 4}}\left(1-\frac{2}{p}\right)+O\left(\sum_{d|P_z}\nu(d)\right)+z^{1/2}.\end{aligned}$$

To understand the main term, note that

$$\prod_{\substack{p\le z\\ p\equiv 1 \pmod 4}}\left(1-\frac{2}{p}\right)\le\exp\left(-2\sum_{\substack{p\le z\\ p\equiv 1 \pmod 4}}\frac{1}{p}\right)\ll\frac{1}{\log z}.$$

Here we have used that $\sum_{p\le z,p\equiv 1 \pmod 4}\frac{1}{p}=\frac{1}{2}\log\log z+O(1)$, which follows by partial summation from the results of Chapter 4. For the O-term, we have $\sum_{d|P_z}\nu(d)=\prod_{p\le z}(1+\nu(p))<3^z$. Inserting these estimates above and choosing $z=\frac{1}{2}\log x$, we find that $\pi_{T^2+1}(x)\ll x/\log\log x$. In particular, the set of numbers n for which n^2+1 is prime has density zero.

The following simple consequence of Theorem 6.2 is often useful:

Corollary 6.3. *Let $\mathcal{P}$ be a set of prime numbers, and let $M(\mathcal{P})$ denote the set of $n \in \mathbf{N}$ divisible by some prime $p \in \mathcal{P}$. Then $M(\mathcal{P})$ has asymptotic density $1 - \prod_{p\in\mathcal{P}}(1-1/p)$. In particular, $M(\mathcal{P})$ has density 1 precisely when $\sum_{p\in\mathcal{P}} p^{-1}$ diverges.*

Proof. Let $M' = \mathbf{N} \setminus M(\mathcal{P})$ be the set of natural numbers n divisible by none of the elements of $\mathcal{P}$, and write $M'(x)$ for the associated counting function. Put $\mathcal{A} := \{n \le x\}$. Then for any choice of z, we have

$$M'(x) \le S(\mathcal{A}, \mathcal{P}, z). \tag{6.20}$$

The right-hand side will be estimated with the aid of Theorem 6.2. We take $X = x$ and let α be the multiplicative function with $\alpha(n) := 1/n$ for every $n \in \mathbf{N}$. With this choice of X and α, we have $|r(d)| \le 1$ for every $d \mid P$. Now put $z = \log x$. By Theorem 6.2,

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, z) &= x \prod_{\substack{p\in\mathcal{P}\\ p\le\log x}} (1-1/p) + O(2^{\log z}) \\ &= (C + o(1))x, \quad \text{where} \quad C := \prod_{p\in\mathcal{P}} (1-1/p). \end{aligned} \tag{6.21}$$

If $C = 0$, then we obtain from (6.20) and (6.21) that M' has density zero, so that $M(\mathcal{P})$ has density 1, which is the assertion of the corollary in this case. If $C \ne 0$, then $\sum_{p\in\mathcal{P}} p^{-1}$ converges, and so

$$\begin{aligned} M'(x) &\ge S(\mathcal{A}, \mathcal{P}, z) - \sum_{\substack{p\in\mathcal{P}\\ p>z}} \frac{x}{p} \\ &= (C + o(1))x + o(x) = (C + o(1))x. \end{aligned}$$

With (6.21), this shows that M' has asymptotic density C, so that $M(\mathcal{P})$ has density $1 - C$, as desired. □

Suppose that in Corollary 6.3 we take $\mathcal{P}$ to be the entire set of prime numbers. Then $M(\mathcal{P})$ consists of every natural number $n > 1$ and so has density 1. Thus $\prod_{p\in\mathcal{P}}(1-1/p) = 0$. This gives another proof of Euler's result from Chapter 1 that $\sum_p \frac{1}{p}$ diverges (cf. [**Pin09**]).

3.3. Three applications. We pause to give three further applications of Corollary 6.3. None of the results we prove are the best of their kind, but the proofs are simple and the statements fairly striking.

Theorem 6.4. *Each of the following sets has density zero:*

(i) *the set of integers $n > 1$ for which the equation*

$$4/n = 1/a + 1/b + 1/c \tag{6.22}$$

has no solution in positive integers a, b, c,

(ii) *the set of natural numbers expressible as a sum of two squares,*

(iii) *the set of odd perfect numbers.*

Remark. The set in (iii) is famously conjectured to be empty; we discuss this conjecture at length in Chapter 8. Erdős & Straus (see [**Erd50a**]) believe that the same holds for the set in (i), that is, that $4/n$ can always be written as a sum of three unit fractions (for $n > 1$). For example,

$$\frac{4}{301} = \frac{1}{76} + \frac{1}{7626} + \frac{1}{87226188}.$$

Of course, the analogous conjecture is trivial if "three" is replaced by "four". It has been verified by computer that the set in (i) contains no $n \le 10^{14}$.

As regards (ii), Landau [**Lan08**] has proved that the number of $n \le x$ expressible as a sum of two squares is

$$\sim \frac{1}{\sqrt{2}} \left(\prod_{p \equiv 3 \pmod 4} \left(1 - \frac{1}{p^2}\right) \right)^{-1/2} \frac{x}{\sqrt{\log x}}.$$

The simplest proof of Landau's result seems to be that of Selberg [**Sel91**, pp. 183–185].

Lemma 6.5. *The set of positive integers divisible by no prime* $p \equiv 3 \pmod 4$ *has density* 0.

Proof. From Chapter 4, we have that $\sum'_{p \le x} p^{-1} \log p = \frac{1}{2} \log x + O(1)$, where the $'$ indicates that the sum is restricted to primes $p \equiv 3 \pmod 4$. So by partial summation, $\sum'_{p \le x} p^{-1} \sim \frac{1}{2} \log\log x$. In particular, $\sum'_p p^{-1}$ diverges. So the result follows from Corollary 6.3. □

Proof of Theorem 6.4(i). It suffices to show that (6.22) is solvable if n possesses a prime divisor $p = 4k - 1 \equiv 3 \pmod 4$. In this case write $n = (4k-1)q$. Then

$$\frac{4}{n} = \frac{4}{q(4k-1)} = \frac{1}{2qk} + \frac{1}{2qk} + \frac{1}{q(4k^2-k)}.$$

This argument also shows that $4/n$ can almost always be written as a sum of *two* unit fractions, since $1/(2qk) + 1/(2qk) = 1/(qk)$. □

Proof of Theorem 6.4(ii). Let $R(x)$ be the number of $n \le x$ which can be written as a sum of two squares, and let $A(x)$ be the number of $n \le x$ which have a *primitive* representation of this form, i.e., a representation as a sum of two coprime squares. As shown by Euler, the n counted by $A(x)$

are precisely those divisible by neither 4 nor any prime $p \equiv 3 \pmod 4$. Moreover,

$$R(x) \le A\left(\frac{x}{1^2}\right) + A\left(\frac{x}{2^2}\right) + A\left(\frac{x}{3^2}\right) + \cdots . \tag{6.23}$$

By Lemma 6.5, we have $A(x) = o(x)$. Now given $\epsilon > 0$, choose an $N \in \mathbf{N}$ for which $A(x) < \epsilon x/4$ whenever $x > N$. Thinking of x as large, we split the sum in (6.23) into two parts according to whether $x/k^2 > N$ or $x/k^2 \le N$. The first of the two resulting sums is bounded by

$$\frac{1}{4}\epsilon \sum_{k\ge 1} \frac{x}{k^2} = \epsilon \frac{\zeta(2)}{4} x < \frac{\epsilon}{2} x.$$

Every term in the second sum is bounded by $A(N)$, and there are no more than $\sqrt{x}$ nonzero terms. Thus,

$$R(x) \le \epsilon x/2 + A(N)\sqrt{x} < \epsilon x$$

for large x. As $\epsilon > 0$ was arbitrary, it follows that $R(x) = o(x)$. □

Proof of Theorem 6.4(iii). It has been known since Euler that every odd perfect number n can be written in the form pa^2, where $p \equiv 1 \pmod 4$ is prime. (We will prove a stronger version of this result in Chapter 8; see Theorem 8.2 there.) Since such integers are sums of two squares, the result follows from that of part (ii). □

4. Brun's pure sieve

In the derivation of Legendre's formula for $\pi(x, z)$ given in §1.2 above, we begin with the total number of positive integers not exceeding x. For each prime $p \le z$, we take away the number of multiples of p. Then, for each pair of primes $p < q \le z$, we add back the number n divisible by both p and q. Continuing we eventually converge on the exact value of $\pi(x, z)$. It is intuitively clear (and we will prove it below) that after each even (addition) step what we have is an overestimate for $\pi(x, z)$, and after each odd (subtraction) step we have an underestimate. A suitable generalization of this fact forms the heart of Brun's pure sieve.

4.1. Preparation. To prove the appropriate generalization, it is convenient to first establish a technical lemma on alternating sums of symmetric functions.

If $a_1, \ldots, a_n$ is a (possibly empty) sequence of $n \ge 0$ elements belonging to a commutative ring, we define (for $k \ge 0$) the kth elementary symmetric function $\sigma_k(a_1, \ldots, a_n)$ as the sum of all possible $\binom{n}{k}$ products of the a_i taken k at a time. We adopt the usual conventions about empty sums and

products, so that for $n = 0$, we have $\sigma_0 = 1$ and $\sigma_k = 0$ for $k > 0$. To take a less pathological example, when $n = 2$, one has

$$\sigma_0(a_1, a_2) = 1, \quad \sigma_1(a_1, a_2) = a_1 + a_2, \quad \sigma_2(a_1, a_2) = a_1 a_2,$$

and $\sigma_k(a_1, a_2) = 0$ for $k > 2$. The following lemma can be found, e.g., in [**Hoo94**]:

Lemma 6.6. *Suppose $0 \le a_1, \dots, a_n \le 1$, where n is nonnegative. Then*

$$\sum_{k=0}^{m} (-1)^k \sigma_k(a_1, \dots, a_n) - \prod_{j=1}^{n} (1 - a_j) \tag{6.24}$$

is nonnegative or nonpositive according to whether m is even or odd, respectively.

Remark. Note that (6.24) vanishes when $m \ge n$.

Proof. We induct on the length n of the sequence. When $n = 0$, the product $P := \prod_{i=1}^{n}(1 - a_i)$ appearing in (6.24) is empty, so equal to 1, while

$$\sum_{k=0}^{m} (-1)^k \sigma_k = 1 - 0 + 0 - \cdots \pm 0 = 1.$$

Hence (6.24) vanishes for every m, confirming the result in this case. Now assume that the result holds for each sequence of n real numbers in $[0, 1]$ and each m, and consider an arbitrary sequence $0 \le a_1, \dots, a_{n+1} \le 1$ of length $n + 1$. By the induction hypothesis, it suffices to prove that

$$\begin{aligned} &\left(\sum_{k=0}^{m} (-1)^k \sigma_k(a_1, \dots, a_{n+1}) - \prod_{i=1}^{n+1} (1 - a_i) \right) \\ &\qquad - \left(\sum_{k=0}^{m} (-1)^k \sigma_k(a_1, \dots, a_n) - \prod_{i=1}^{n} (1 - a_i) \right) \end{aligned} \tag{6.25}$$

is nonnegative or nonpositive according to whether m is even or odd respectively. This is easily seen to hold for $m = 0$, since then (6.25) simplifies to Pa_{n+1}, which is nonnegative. When $m > 0$, we can rewrite (6.25) as

$$\begin{aligned} &\sum_{k=1}^{m} (-1)^k \left(\sigma_k(a_1, \dots, a_{n+1}) - \sigma_k(a_1, \dots, a_n) \right) + Pa_{n+1} \\ &\qquad = \sum_{k=1}^{m} (-1)^k a_{n+1} \sigma_{k-1}(a_1, \dots, a_n) + Pa_{n+1} \\ &\qquad = a_{n+1} \left(P - \sum_{k=0}^{m-1} (-1)^k \sigma_k(a_1, \dots, a_n) \right). \end{aligned}$$

The claim in this case now follows from the induction hypothesis. □

An important special case occurs when $n \in \mathbf{N}$ and $a_1 = a_2 = \cdots = a_n = 1$. Then $\prod_{i=1}^{n}(1 - a_i) = (1-1)^n = 0$, while $\sigma_k(1, \ldots, 1) = \binom{n}{k}$. So from Lemma 6.6 we obtain the following:

Lemma 6.7. *Let n be a positive integer. Then the alternating sum*

$$\sum_{k=0}^{m} (-1)^k \binom{n}{k}$$

is nonnegative or nonpositive according to whether m is even or odd.

Remark. While Lemma 6.6 will be important in our treatment of the Brun–Hooley sieve, for Brun's pure sieve we only need Lemma 6.7. Thus it is of interest that Lemma 6.7 admits a simple proof independent of Lemma 6.6: Indeed, by induction on m, one easily finds that

$$\sum_{k=0}^{m} (-1)^k \binom{n}{k} = (-1)^m \binom{n-1}{m}, \tag{6.26}$$

which makes Lemma 6.7 obvious. Alternatively, (6.26) follows by comparing the coefficient of x^m in both sides of the power series identity $(1-x)^{n-1} = (1-x)^{-1}(1-x)^n$.

Lemma 6.7 implies the following variant of Theorem 6.1:

Theorem 6.8 (Bonferroni inequalities)**.** *Let X be a nonempty, finite set of N objects, and let $P_1, \ldots, P_r$ be properties that elements of X may have. For each subset $I \subset \{1, 2, \ldots, r\}$, let $N(I)$ denote the number of elements of X that have each of the properties indexed by the elements of I. Let N_0 denote the number of elements of X with none of these properties. Then if m is a nonnegative even integer,*

$$N_0 \le \sum_{k=0}^{m} (-1)^k \sum_{\substack{I \subset \{1,2,\ldots,r\} \\ |I|=k}} N(I), \tag{6.27}$$

while if m is a nonnegative odd integer,

$$N_0 \ge \sum_{k=0}^{m} (-1)^k \sum_{\substack{I \subset \{1,2,\ldots,r\} \\ |I|=k}} N(I). \tag{6.28}$$

Proof. Suppose that $x \in X$ has exactly l of the properties $P_1, \ldots, P_r$. If $l = 0$, then x is counted once by both N_0 and the common right-hand side of (6.27) and (6.28) (corresponding to $I = \emptyset$). If $l \ge 1$, then x is not counted at all by N_0, and is counted by this right-hand sum with weight

$$\sum_{k=0}^{m} (-1)^k \binom{l}{k} \begin{cases} \ge 0 & \text{if } m \text{ is even,} \\ \le 0 & \text{otherwise.} \end{cases}$$

Summing over $x \in X$ gives the theorem. □

4.2. A working version.

Corollary 6.9 (Brun's pure sieve, general form)**.** *With the notation of §2, we have for every nonnegative even integer m,*

$$\sum_{d|P,\omega(d)\le m-1} \mu(d)A_d \le S(\mathcal{A},\mathcal{P}) \le \sum_{d|P,\omega(d)\le m} \mu(d)A_d.$$

Proof. As in the proof of Theorem 6.2, let $p_1,\dots,p_r$ be a list of the primes $p \in \mathcal{P}$, and let P_i be the property of being divisible by p_i. We aim to estimate the number $S(\mathcal{A},\mathcal{P})$ of elements of $\mathcal{A}$ possessing none of the P_i. The upper bound for $S(\mathcal{A},\mathcal{P})$ in the corollary is just (6.27). If $m = 0$, then the lower bound is trivial, while if $m > 0$, then $m-1$ is a nonnegative odd integer, and the lower bound follows from (6.28). □

To obtain a result suitable for applications, we substitute $A_d = X\alpha(d) + r(d)$. With a bit of manipulation, we arrive at the following theorem:

Theorem 6.10 (Brun's pure sieve)**.** *For every even integer $m \ge 0$,*

$$S(\mathcal{A},\mathcal{P}) = X\prod_{p\in\mathcal{P}}(1-\alpha(p)) + O\Big(\sum_{d|P,\omega(d)\le m}|r(d)|\Big) + O\Big(X\sum_{d|P,\omega(d)\ge m}\alpha(d)\Big).$$

Here the implied constants are absolute.

Proof. From Corollary 6.9,

$$\begin{aligned} S(\mathcal{A},\mathcal{P}) &= \sum_{\substack{d|P\\ \omega(d)\le m}} \mu(d)A_d + O\Big(\sum_{\substack{d|P\\ \omega(d)=m}} A_d\Big) \\ &= \sum_{\substack{d|P\\ \omega(d)\le m}} \mu(d)(X\alpha(d)+r(d)) + O\Big(\sum_{\substack{d|P\\ \omega(d)=m}} A_d\Big) \\ &= X\sum_{\substack{d|P\\ \omega(d)\le m}} \mu(d)\alpha(d) + O\Big(\sum_{\substack{d|P\\ \omega(d)\le m}}|r(d)|\Big) + O\Big(\sum_{\substack{d|P\\ \omega(d)=m}} A_d\Big). \end{aligned}$$

Writing $A_d = X\alpha(d) + r(d)$, we see that the last of these error terms is

$$\ll X\sum_{d|P,\omega(d)=m}\alpha(d) + \sum_{d|P,\omega(d)=m}|r(d)|;$$

hence,

$$S(\mathcal{A},\mathcal{P}) = X \sum_{\substack{d\mid P\\ \omega(d)\le m}} \mu(d)\alpha(d) + O\Big(\sum_{\substack{d\mid P\\ \omega(d)\le m}} |r(d)|\Big) + O\Big(X \sum_{\substack{d\mid P\\ \omega(d)=m}} \alpha(d)\Big). \tag{6.29}$$

In order to factor the sum appearing in the main term, we extend the sum to all $d \mid P$; the main term can then be expressed as $X\prod_{p\in\mathcal{P}}(1-\alpha(p))$, but we have introduced a new error of

$$\ll X \sum_{d\mid P,\omega(d)>m} \alpha(d).$$

If this is combined with the last error term of (6.29), we find that

$$S(\mathcal{A},\mathcal{P}) = X\prod_{p\in\mathcal{P}}(1-\alpha(p)) + O\Big(\sum_{d\mid P,\omega(d)\le m} |r(d)|\Big) + O\Big(X \sum_{d\mid P,\omega(d)\ge m} \alpha(d)\Big),$$

exactly as the theorem asserts. □

4.3. Application to the twin prime problem. The most famous application of Brun's pure sieve is Brun's own 1919 contribution [**Bru19a**] to the twin prime problem:

Theorem 6.11. *As $x\to\infty$,*

$$\pi_2(x) \ll \frac{x}{(\log x)^2}(\log\log x)^2.$$

The upper bound differs from what is expected by a factor of $(\log\log x)^2$. We shall later remedy this defect. Nevertheless, it is worth noting that the estimate of Theorem 6.11 is already sharp enough to imply the following striking result:

Corollary 6.12. *If there are infinitely many primes p such that $p+2$ is also prime, then the sum*

$$\sum_p \frac{1}{p},$$

taken over all such primes, converges.

Proof. By Theorem 6.13, $\pi_2(x) \ll x/(\log x)^{3/2}$ as $x\to\infty$. It follows that the same estimate holds, with perhaps a different implied constant, in the range $x \ge 3$. Letting p_n denote the nth prime p for which $p+2$ is also prime, we see that for $n \ge 1$,

$$n = \pi_2(p_n) \ll p_n/(\log p_n)^{3/2},$$

so that

$$p_n \gg n(\log p_n)^{3/2} \geq \frac{1}{2}(n+1)(\log(n+1))^{3/2}.$$

The comparison and integral tests together now imply that $\sum_{n=1}^{\infty} p_n^{-1}$ converges, which is the assertion of the corollary. □

Remark. For historical reasons, in place of the series appearing in Corollary 6.12 one usually sees the slight variant

$$\left(\frac{1}{3}+\frac{1}{5}\right)+\left(\frac{1}{5}+\frac{1}{7}\right)+\left(\frac{1}{11}+\frac{1}{13}\right)+\cdots.$$

Of course this series converges (by comparison with that of the corollary), and its value B is known as *Brun's constant*. Computing the value of B to any precision seems to be difficult; while constants like π and e are known to billions of decimal digits, the sharpest known bounds on B are (roughly)

$$1.830 < B < 2.347.$$

Thus we do not know B to even one significant digit! The lower bound here is due to Sebah [**SG**], who computed all the twin prime pairs up to 10^{16} and summed their reciprocals. The upper bound is due to Crandall & Pomerance ([**CP05**, pp. 16-17], see also [**Kly07**, Chapter 3]), who bound the sum of the twin prime pairs past 10^{16} using an explicit upper estimate of Riesel and Vaughan [**RV83**] for the number of twin prime pairs. Much sharper estimates for Brun's constant are available if one assumes a suitable quantitative version of the twin prime conjecture; e.g., it is plausible that

$$B = 1.902160583121 \pm 4.08 \times 10^{-8}.$$

This last estimate is taken from the Ph.D. thesis of Klyve [**Kly07**], which the reader should consult for references to earlier work.

With $\mathcal{A} := \{n(n+2) : n \leq x\}$ and $\mathcal{P} := \{p \leq z\}$, put $\pi_2(x,z) := S(\mathcal{A},\mathcal{P})$. Theorem 6.11 is an easy consequence of the following estimate:

Theorem 6.13. *Suppose $z = z(x) \to \infty$ as $x \to \infty$ and that $z(x) \leq x^{1/(20\log\log x)}$ for all large x. Then $\pi_2(x,z) \sim 2C_2 e^{-2\gamma} x/(\log z)^2$ as $x \to \infty$, where C_2 is the twin prime constant.*

Proof of Theorem 6.11 assuming Theorem 6.13. Relation (6.10) tells us that $\pi_2(x) \leq z + \pi_2(x,z)$. Take $z = x^{1/(20\log\log x)}$. Theorem 6.13 implies that as $x \to \infty$,

$$\pi_2(x) \ll x^{1/(20\log\log x)} + \frac{x}{(\log x)^2}(\log\log x)^2 \ll \frac{x}{(\log x)^2}(\log\log x)^2. \quad \square$$

4.4. Proof of Theorem 6.13. Estimates for $\pi_2(x,z)$ were discussed in the second example of §3.2; the difference here is that we now have Brun's pure sieve at our disposal. As in that example, we take $X = x$ and $\alpha(d) = \nu(d)/d$, where ν is defined by (6.18). Then $|r(d)| \le \nu(d) \le 2^{\omega(d)}$ for all d. So by Theorem 6.10,

$$\pi_2(x,z) = x\prod_{p\le z}(1-\alpha(p)) + O\Big(\sum_{d|P,\omega(d)\le m} 2^{\omega(d)}\Big) + O\Big(x\sum_{d|P,\omega(d)\ge m}\alpha(d)\Big), \tag{6.30}$$

for each even number $m \ge 0$. We take

$$m := 10\,\lfloor \log\log z \rfloor .$$

Note that as x goes to infinity, so does z and hence also m. In §3.2, we calculated that the main term of (6.30) is asymptotic to

$$2C_2 e^{-2\gamma} x/(\log z)^2$$

as $x \to \infty$. So to prove Theorem 6.13, it is enough to establish the following two estimates:

(i) With $E_1 := \sum_{d|P,\omega(d)\le m} 2^{\omega(d)}$, we have $E_1 = o(x/(\log z)^2)$.

(ii) With $E_2 := x\sum_{d|P,\omega(d)\ge m}\alpha(d)$, we have $E_2 = o(x/(\log z)^2)$.

Proof of (i). For large x,

$$\begin{aligned} E_1 = \sum_{d|P,\omega(d)\le m} 2^{\omega(d)} &= \sum_{k=0}^{m} 2^k \binom{\pi(z)}{k} \le \sum_{k=0}^{m} (2\pi(z))^k \\ &\le \sum_{k=-\infty}^{m} (2\pi(z))^k = (2\pi(z))^m \frac{1}{1-\frac{1}{2\pi(z)}} \\ &\le 2(2\pi(z))^m \le 2z^m, \end{aligned}$$

since $\pi(z) \le z/2$ for large x. Hence

$$E_1 \le 2z^{10\log\log z} \le 2z^{10\log\log x} \le 2x^{1/2}.$$

This upper bound is certainly $o(x/(\log z)^2)$, since as $x \to \infty$,

$$\frac{x^{1/2}}{x/(\log z)^2} \le \frac{x^{1/2}}{x/(\log x)^2} = \frac{(\log x)^2}{x^{1/2}} \to 0. \qquad \square$$

Proof of (ii). We can write $E_2 = x\sum_{k\ge m}\sum_{d|P,\omega(d)=k}\alpha(d)$. For the inner sum we have

$$\sum_{\substack{d|P\\ \omega(d)=k}} \alpha(d) = \sum_{p_1<p_2<\cdots<p_k\le z} \alpha(p_1)\alpha(p_2)\cdots\alpha(p_k) \le \frac{1}{k!}\Big(\sum_{p\le z}\alpha(p)\Big)^k.$$

Here the upper bound comes from the multinomial theorem: In the expansion of $(\sum_{p\le z}\alpha(p))^k$, every term $\alpha(p_1)\cdots\alpha(p_k)$ appears with coefficient $k!$. From Mertens' first theorem, we have $\sum_{p\le z}p^{-1}\le \log\log z + c$ for $z\ge 3$, where c is an absolute constant. Since $\alpha(p)\le 2/p$ for every prime p,

$$\sum_{k\ge m}\frac{1}{k!}\Big(\sum_{p\le z}\alpha(p)\Big)^k \le \sum_{k\ge m}\frac{1}{k!}(2\log\log z + 2c)^k. \tag{6.31}$$

The ratio of the $(k+1)$th term in the right-hand series to the kth is given by

$$\frac{2\log\log z + 2c}{k+1} \le \frac{2\log\log z + 2c}{10\,\lfloor\log\log z\rfloor + 1} \le 1/2,$$

for large enough z, and hence also for large enough x. So, for such x the right-hand sum in (6.31) is bounded by twice its first term. Because

$$e^m = 1 + m + m^2/2! + m^3/3! + \cdots \ge m^m/m!,$$

we have $m! \ge (m/e)^m$, so that

$$\sum_{k\ge m}\frac{1}{k!}(2\log\log z + 2c)^k \le 2\left(\frac{2e\log\log z + 2ce}{m}\right)^m.$$

Since $m = 10\,\lfloor\log\log z\rfloor$, the parenthetical expression on the right is eventually smaller than any constant exceeding $2e/10$; in particular, it is eventually smaller than $3/5$. It follows that for large x,

$$\begin{aligned} E_2 \le 2x(3/5)^m &= 2x(3/5)^{10\lfloor\log\log z\rfloor} \\ &\ll x(3/5)^{10\log\log z} \ll x/(\log z)^5, \end{aligned}$$

since $10\log\frac{3}{5} < -5$. So $E_2 = o(x/(\log z)^2)$. □

5. The Brun–Hooley sieve

5.1. The sifting function perspective. Before we discuss the Brun–Hooley method, it is worthwhile for us to revisit some of the earlier results of this chapter from a a slightly different perspective. Keeping the notation of §2, we introduce the *sifting function*

$$s(n) := \begin{cases} 1 & \text{if } \gcd(n,P) = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6.32}$$

Then

$$S(\mathcal{A},\mathcal{P}) = \sum_{a\in A} s(a). \tag{6.33}$$

Since $\sum_{d|m}\mu(d)$ vanishes for each natural number $m>1$, the sifting function $s(n)$ has the following important representation:

$$s(n)=\sum_{d|n,d|P}\mu(d). \tag{6.34}$$

Substituting this into (6.33) and interchanging the order of summation, we easily arrive at Theorem 6.2 (the sieve of Eratosthenes–Legendre). In the same way, Brun's pure sieve is a consequence of the following lemma:

Lemma 6.14. *Let n be a natural number. The expression*

$$\sum_{\substack{d|n,d|P\\ \omega(d)\le m}}\mu(d)-\sum_{d|n,d|P}\mu(d) \tag{6.35}$$

is nonnegative or nonpositive according to whether the integer $m\ge 0$ is even or odd.

The proof of Lemma 6.14 is essentially the one already given for the Bonferroni inequalities. Namely, if we suppose that n is divisible by exactly l primes $p\in\mathcal{P}$, then by Lemma 6.6,

$$\sum_{\substack{d|n,d|P\\ \omega(d)\le m}}\mu(d)=\sum_{k=0}^{m}(-1)^k\binom{l}{k}\begin{cases}=1 & \text{if } l=0 \text{ (i.e., if } \gcd(n,P)=1),\\ \ge 0 & \text{if } l\ge 1, m \text{ even},\\ \le 0 & \text{if } l\ge 1, m \text{ odd}.\end{cases}$$

For later use we note the following consequence of Lemma 6.14:

Lemma 6.15. *If n is a natural number and $m\ge 0$ is even, then*

$$0\le\sum_{\substack{d|n,d|P\\ \omega(d)\le m}}\mu(d)-\sum_{d|n,d|P}\mu(d)\le\sum_{\substack{d|n,d|P\\ \omega(d)=m+1}}1.$$

5.2. The upper bound. The Brun–Hooley method takes two forms, depending on whether we are after upper or lower bounds. Here we describe the simpler upper bound method. We suppose the sifting set $\mathcal{P}$ to be partitioned into r disjoint sets, say $\mathcal{P}=\dot{\bigcup}_{j=1}^{r}\mathcal{P}_j$. Then n is divisible by no prime $p\in\mathcal{P}$ precisely when n is divisible by no prime $p\in\mathcal{P}_j$ for every $1\le j\le r$. Consequently, setting $P_j:=\prod_{p\in\mathcal{P}_j}p$, and invoking Lemma 6.14 (with $\mathcal{P}_j, P_j$ in place of $\mathcal{P}, P$) we see that

$$\begin{aligned}s(n)=\sum_{d|n,d|P}\mu(d)&=\prod_{j=1}^{r}\sum_{d_j|n,d_j|P_j}\mu(d_j)\\ &\le\prod_{j=1}^{r}\sum_{\substack{d_j|n,d_j|P_j\\ \omega(d_j)\le m_j}}\mu(d_j),\end{aligned}$$

for any choice of nonnegative even integers $m_1, \dots, m_r$. Referring to (6.33), we obtain the upper bound

$$\begin{aligned} S(\mathcal{A},\mathcal{P}) &\le \sum_{\substack{d_1,\dots,d_r\\ d_j\mid P_j,\omega(d_j)\le m_j}} \mu(d_1)\cdots\mu(d_r)A_{d_1\cdots d_r} \\ &= X \sum_{\substack{d_1,\dots,d_r\\ d_j\mid P_j,\omega(d_j)\le m_j}} \mu(d_1)\cdots\mu(d_r)\alpha(d_1)\cdots\alpha(d_r) \\ &\qquad + \sum_{\substack{d_1,\dots,d_r\\ d_j\mid P_j,\omega(d_j)\le m_j}} \mu(d_1)\cdots\mu(d_r)r(d_1\cdots d_r). \end{aligned} \tag{6.36}$$

Hence $S(\mathcal{A},\mathcal{P})$ is bounded above by

$$X\prod_{j=1}^{r}\sum_{\substack{d_j\mid P_j\\ \omega(d_j)\le m_j}}\mu(d_j)\alpha(d_j) + \sum_{\substack{d_1,\dots,d_r\\ d_j\mid P_j,\omega(d_j)\le m_j}}\mu(d_1)\cdots\mu(d_r)r(d_1\cdots d_r). \tag{6.37}$$

This is the upper bound of the Brun–Hooley method. To facilitate applications, we replace the first term of (6.37), which we think of as the main term, with something more easily compared with $X\prod_{p\in\mathcal{P}}(1-\alpha(p))$. This can be accomplished by replacing the jth term of the product in (6.37) with something more easily compared with $\prod_{p\in\mathcal{P}_j}(1-\alpha(p))$. For this, we utilize Lemma 6.6, which implies that for each $1\le j\le r$,

$$0 \le \sum_{\substack{d_j\mid P_j\\ \omega(d_j)\le m_j}}\mu(d_j)\alpha(d_j) - \prod_{p\in\mathcal{P}_j}(1-\alpha(p)) \le \sum_{\substack{d_j\mid P_j\\ \omega(d_j)=m_j+1}}\alpha(d_j).$$

Thus, if we set

$$\prod\nolimits^{(j)} := \prod_{p\in\mathcal{P}_j}(1-\alpha(p)), \quad \sum\nolimits^{(j)} := \sum_{\substack{d_j\mid P_j\\ \omega(d_j)=m_j+1}}\alpha(d_j), \tag{6.38}$$

then

$$\begin{aligned} X\prod_{j=1}^{r}\sum_{\substack{d_j\mid P_j\\ \omega(d_j)\le m_j}}\mu(d_j)\alpha(d_j) &\le X\prod_{j=1}^{r}\Big(\prod\nolimits^{(j)} + \sum\nolimits^{(j)}\Big) \\ &= X\prod_{p\in\mathcal{P}}(1-\alpha(p))\prod_{j=1}^{r}\Big(1+\sum\nolimits^{(j)}\Big/\prod\nolimits^{(j)}\Big), \end{aligned}$$

provided the division makes sense, i.e., provided $\alpha(p)<1$ for each $p\in\mathcal{P}$. Henceforth, we assume (as will be the case in all our applications) this condition on α.

Recalling that $1 + t \le \exp(t)$, after estimating the remainder term of (6.37) trivially, we arrive at the following theorem:

Theorem 6.16 (Brun–Hooley sieve, upper bound)**.** *Let $\mathcal{P} = \dot{\bigcup}_{j=1}^{r} \mathcal{P}_j$ be a partition of $\mathcal{P}$. Suppose that $\alpha(p) < 1$ for each $p \in \mathcal{P}$. For any choice of nonnegative even integers $m_1, \dots, m_r$, we have*

$$S(\mathcal{A}, \mathcal{P}) \le X \prod_{p \in \mathcal{P}} (1 - \alpha(p)) \exp\left(\sum_{j=1}^{r} \Big(\sum\nolimits^{(j)} \Big/ \prod\nolimits^{(j)} \Big) \right) + O\Big(\sum_{\substack{d_1, \dots, d_r \\ d_j \mid P_j, \omega(d_j) \le m_j}} |r(d_1 \cdots d_r)| \Big), \tag{6.39}$$

where $\prod^{(j)}$ and $\sum^{(j)}$ are defined, for $1 \le j \le r$, by (6.38)*, and the implied constant is absolute.*

5.3. Applications of the upper bound. Define $R(N)$ as the number of (ordered) representations of N as a sum of two primes, or equivalently, as the number of ordered prime pairs $(p, N-p)$. In Chapter 3, we conjectured that as $N \to \infty$ through *even* integers,

$$R(N) \sim 2C_2 \frac{N}{(\log N)^2} \prod_{p \mid N, p > 2} \frac{p-1}{p-2}.$$

We now use the Brun–Hooley sieve to establish an upper bound for $R(N)$ of the conjecturally correct order of magnitude:

Theorem 6.17. *For every even natural number N,*

$$R(N) \ll \frac{N}{(\log N)^2} \prod_{p \mid N} \left(1 + \frac{1}{p}\right).$$

Let N be an even natural number and define $\mathcal{A} := \{n(N-n) : 1 \le n \le N\}$. Letting $\mathcal{P}$ be the set of all primes, we have for each choice of $z > 0$,

$$R(N) \le 2z + S(\mathcal{A}, \mathcal{P}, z).$$

Indeed, if $N = n + (N-n)$ is a representation of N as a sum of two primes, then either at least one of n or $N-n$ lies in $[2, z]$ or both n and $N-n$ have no prime factors $\le z$. The former case occurs for no more than $2z$ values of n, and the n for which the latter holds (which necessarily satisfy $2 \le n \le N-2$) are counted by $S(\mathcal{A}, \mathcal{P}, z)$.

We now choose our sifting parameters: Let $X = N$, and let $\alpha(d) = \nu(d)/d$, where

$$\nu(d) := \#\{n \bmod d : n(N-n) \equiv 0 \pmod d\};$$

then

$$\alpha(p) = \begin{cases} 1/p & \text{if } p \mid N, \\ 2/p & \text{if } p \nmid N. \end{cases} \tag{6.40}$$

Because N is even, $\alpha(p) < 1$ for every prime p. Moreover,

$$A_d = X\alpha(d) + r(d) \quad \text{where} \quad |r(d)| \le \nu(d) \quad \text{for all } d \mid P(z). \tag{6.41}$$

We think of $X = N$ as heading off towards infinity while $u > 1$ is fixed. Our immediate goal is to show that if u is fixed large enough, then

$$S(\mathcal{A}, \mathcal{P}, z) \ll X \prod_{p \le z} (1 - \alpha(p)) \quad (X \to \infty), \quad \text{where } z := X^{1/u}.$$

To apply the Brun–Hooley sieve to this situation we need a partition of $\mathcal{P} \cap [2, z]$. We introduce the notation

$$\eta = \log\log X$$

and the choice of parameters

$$K := 1.57, \quad K_1 := 1.571. \tag{6.42}$$

For the present discussion it is only important that $1 < K < K_1$, but this choice will be particularly effective for the lower bound applications of §5.5.

For large X, we have $\eta < z = X^{1/u}$, so that if we define R as the minimal integer with

$$z^{1/K^R} < \eta,$$

then $R \ge 1$. (Indeed, $R \to \infty$ with X.) For such X, we define

$$z_j = \begin{cases} z^{1/K^j} & \text{for } 0 \le j \le R-1, \\ \eta & \text{for } j = R, \\ 1 & \text{for } j = R+1. \end{cases}$$

We partition $\mathcal{P} \cap [2, z]$ into the $r := R + 1$ sets

$$\mathcal{P}_j := \{p \in \mathcal{P} : z_j < p \le z_{j-1}\} \quad (1 \le j \le R+1),$$

and we define the corresponding nonnegative even integers $m_1, \dots, m_{R+1}$ by putting

$$m_j = 2j \quad (j = 1, \dots, R) \quad \text{and} \qquad m_{R+1} = \infty;$$

here "∞" indicates that m_{R+1} is chosen at least as large as the cardinality of $\mathcal{P}_{R+1}$. For definiteness, we take m_{R+1} as the smallest even integer with this property. With this choice of m_{R+1}, the condition on a divisor d of P_{R+1} that it has no more than m_{R+1} prime divisors becomes vacuous.

We are finally in a position to apply the upper bound (6.39) to our problem. By our choice of m_{R+1},

$$\sum^{(R+1)} = \sum_{\substack{d_{R+1}|P_{R+1} \\ \omega(d_{R+1})=m_{R+1}+1}} \alpha(d_{R+1}) = 0. \tag{6.43}$$

Hence $\sum^{(j)} / \prod^{(j)}$ vanishes at $j = R+1$, and to estimate the main term of (6.39) it suffices to estimate the ratio $\sum^{(j)} / \prod^{(j)}$ for $j = 1, \dots, R$. The denominator is handled by the following lemma:

Lemma 6.18. *As $x \to \infty$, we have*

$$\prod_{x<p\le y} \left(1 - \frac{2}{p}\right) = \frac{(\log x)^2}{(\log y)^2} \left(1 + O\left(\frac{1}{\log x}\right)\right)$$

uniformly for $y \ge x$.

Proof. Suppose $x \ge 4$; then $2/p \le 1/2$ for each $p \ge x$, so that $\log(1-2/p) = -2/p + O((-2/p)^2)$ with an absolute implied constant, and

$$\begin{aligned}\sum_{x<p\le y} \log\left(1 - \frac{2}{p}\right) &= -2 \sum_{x<p\le y} \frac{1}{p} + O\left(\sum_{x<p\le y} \frac{1}{p^2}\right) \\ &= -2\left(\log \frac{\log y}{\log x} + O\left(\frac{1}{\log x}\right)\right) + O\left(\frac{1}{x}\right) \\ &= \log \frac{(\log x)^2}{(\log y)^2} + O\left(\frac{1}{\log x}\right).\end{aligned}$$

Exponentiating gives the result. □

As $X \to \infty$, so do each of $z_1, \dots, z_R$ (since each is at least η). Consequently, Lemma 6.18 implies that for large X (and each $j = 1, 2, \dots, R$),

$$\begin{aligned}\prod^{(j)} &= \prod_{z_j<p\le z_{j-1}} (1 - \alpha(p)) \ge \prod_{z_j<p\le z_{j-1}} \left(1 - \frac{2}{p}\right) \\ &= \frac{(\log z_j)^2}{(\log z_{j-1})^2} \left(1 + O\left(\frac{1}{\log z_j}\right)\right) \ge \frac{1}{K^2} \left(1 + O\left(\frac{1}{\log \eta}\right)\right) \ge \frac{1}{K_1^2}.\end{aligned} \tag{6.44}$$

Moreover, for $1 \le j \le R$, we have

$$\sum\nolimits^{(j)} = \sum_{\substack{d_j \mid P_j \\ \omega(d_j)=m_j+1}} \alpha(d_j) \le \frac{1}{(m_j+1)!}\left(\sum_{p\in\mathcal{P}_j}\alpha(p)\right)^{m_j+1}$$

$$\le \frac{1}{(m_j+1)!}\left(\sum_{p\in\mathcal{P}_j}\frac{2}{p}\right)^{m_j+1} \le \frac{(2\log K_1)^{m_j+1}}{(m_j+1)!} \tag{6.45}$$

provided X is large enough, since in that case

$$\sum_{z_j<p\le z_{j-1}}\frac{2}{p} = 2\log\frac{\log z_{j-1}}{\log z_j} + O\left(\frac{1}{\log z_j}\right) \le 2\log K + O\left(\frac{1}{\log\eta}\right) \le 2\log K_1.$$

Putting (6.44) and (6.45) together and recalling (6.43), we find that for large X,

$$\sum_{j=1}^{R+1}\left(\sum\nolimits^{(j)} \Big/ \prod\nolimits^{(j)}\right) \le K_1^2 \sum_{j=1}^{R}\frac{(2\log K_1)^{2j+1}}{(2j+1)!} \le K_1^2\exp(2\log K_1).$$

This shows that the main term of (6.39) is bounded above by a constant multiple of $X\prod_{p\le z}(1-\alpha(p))$. For any fixed $u>1$,

$$X\prod_{p\le X^{1/u}}(1-\alpha(p)) \ge \frac{1}{2}X\prod_{2<p\le X^{1/u}}(1-2/p) \asymp X/(\log X)^2 \quad (X\to\infty), \tag{6.46}$$

so that to obtain the estimate $S(\mathcal{A},\mathcal{P},z) \ll X\prod_{p\le z}(1-\alpha(p))$ we need only ensure that the sum appearing in the expression for the remainder term,

$$\sum_{\substack{d_1,\dots,d_{R+1} \\ d_j\mid P_j, \omega(d_j)\le m_j}} |r(d_1\cdots d_{R+1})|, \tag{6.47}$$

is of smaller order than $X/(\log X)^2$. We will show that for an appropriate choice of u, this sum is $\ll X^\delta$ for a constant $\delta<1$.

Observe that any product $d_1\cdots d_{R+1}$ appearing as an argument of $r(\cdot)$ in the sum (6.47) satisfies

$$\begin{aligned} d_1\cdots d_{R+1} &\le \left(\prod_{j=1}^{R} z_{j-1}^{m_j}\right)\eta^\eta \\ &= X^{\frac{1}{u}\left(\sum_{j=1}^{R} m_j/K^{j-1}\right)} X^{\log\log X\log\log\log X/\log X}. \end{aligned}$$

Also,

$$\sum_{j=1}^{R}\frac{m_j}{K^{j-1}} \le \sum_{j=1}^{\infty}\frac{2j}{K^{j-1}} = \frac{2K^2}{(K-1)^2} = 15.173\ldots.$$

We fix a choice of u exceeding $15.173\ldots$, say $u = 16$ for definiteness. Then for large enough X, we have $d_1 \cdots d_{R+1} \le X^{15.2/16}$ for every such product $d_1 \cdots d_{R+1}$. For every d dividing $P(z)$,

$$|r(d)| \le \nu(d) = \prod_{p|d} \nu(p) \le 2^{\omega(d)} \le \tau(d).$$

Since each integer admits at most one representation in the form $d_1 \cdots d_{R+1}$ (since the d_i are supported on disjoint sets of primes), the sum (6.47) above is bounded by

$$\sum_{n \le X^{15.2/16}} \tau(n) = \sum_{n \le X^{15.2/16}} \sum_{e|n} 1 \le X^{15.2/16} \sum_{e \le X^{15.2/16}} \frac{1}{e} \ll X^{15.2/16} \log X.$$

It follows that for all large X,

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, X^{\frac{1}{16}}) &\ll X \prod_{p \le X^{\frac{1}{16}}} (1 - \alpha(p)) \\ &= X \prod_{\substack{p \le X^{\frac{1}{16}} \\ p \nmid N}} \left(1 - \frac{2}{p}\right) \prod_{\substack{p \le X^{\frac{1}{16}} \\ p | N}} \left(1 - \frac{1}{p}\right). \end{aligned}$$

Since $(1 - 2/p) \le (1 - 1/p)^2$, we find that

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, X^{\frac{1}{16}}) &\le X \prod_{\substack{p \le X^{\frac{1}{16}} \\ p \nmid N}} \left(1 - \frac{1}{p}\right)^2 \prod_{\substack{p \le X^{\frac{1}{16}} \\ p | N}} \left(1 - \frac{1}{p}\right) \\ &= X \prod_{p \le X^{\frac{1}{16}}} \left(1 - \frac{1}{p}\right)^2 \prod_{\substack{p \le X^{\frac{1}{16}} \\ p | N}} \left(1 - \frac{1}{p}\right)^{-1} \\ &\ll \frac{X}{(\log X)^2} \prod_{p|N} \left(1 - \frac{1}{p}\right)^{-1}. \end{aligned}$$

Noting that

$$\prod_{p|N} \left(1 - \frac{1}{p}\right)^{-1} \Big/ \prod_{p|N} \left(1 + \frac{1}{p}\right) = \prod_{p|N} \left(1 - \frac{1}{p^2}\right)^{-1} \le \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$

we conclude that for large X,

$$S(\mathcal{A}, \mathcal{P}, X^{1/16}) \ll \frac{X}{(\log X)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right). \tag{6.48}$$

Consequently, for all large positive even numbers N,

$$\begin{aligned} R(N) &\le S(\mathcal{A}, \mathcal{P}, X^{1/16}) + 2X^{1/16} \\ &\ll \frac{X}{(\log X)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right) = \frac{N}{(\log N)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right). \end{aligned}$$

This gives the assertion of Theorem 6.17 for sufficiently large N, but for bounded N the theorem is trivial.

The proof we have given applies mutatis mutandis to the generalized prime twin problem, i.e., the problem of estimating

$$\pi_N(x) := \#\{p \le x : p, p + N \text{ are both prime}\}.$$

Indeed, let N be a positive even integer, and define the sequence

$$\mathcal{A} := \{n(n + N) : 1 \le n \le x\}.$$

Then

$$\pi_N(x) \le z + S(\mathcal{A}, \mathcal{P}, z)$$

for any choice of positive z. To estimate $S(\mathcal{A}, \mathcal{P}, z)$, we take $X = x$ and choose $\alpha(d) = \nu(d)/d$, where $\nu(d)$ is the number of solutions to the congruence $n(N + n) \equiv 0 \pmod{d}$. Then $\alpha(d)$ is again given by (6.40). If we now choose z, the z_j, the partition $\mathcal{P}_j$, and the m_j exactly as before, the same proof as above shows that (6.48) holds for all sufficiently large X, say $X \ge x_0$. Moreover, both x_0 and the implied constant in (6.48) are independent of N. So, for $x \ge x_0$,

$$\begin{aligned} \pi_N(x) &\ll X^{1/16} + S(\mathcal{A}, \mathcal{P}, X^{1/16}) \\ &\ll x^{1/16} + \frac{x}{(\log x)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right) \ll \frac{x}{(\log x)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right), \end{aligned}$$

uniformly in N. Since $\pi_N(x)$ is trivially bounded by x_0 for $2 \le x \le x_0$, the same upper estimate for $\pi_N(x)$ remains valid for all $x \ge 2$ and all even natural numbers N (with perhaps a different implied constant). So we have proved:

Theorem 6.19. *Let N be a positive even integer. Then for $x \ge 2$,*

$$\pi_N(x) \ll \frac{x}{(\log x)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right),$$

where the implied constant is absolute.

5.4. The lower bound. We turn now to the problem of bounding $S(\mathcal{A},\mathcal{P})$ from below. A natural temptation here is to simply parallel what we did in the upper bound case: If we suppose $m_1,\dots,m_r$ to be r odd natural numbers, then for each j,

$$\sum_{\substack{d_j\mid n,d_j\mid P_j\\ \omega(d_j)\le m_j}}\mu(d_j)\le\sum_{d_j\mid n,d_j\mid P_j}\mu(d_j).$$

But since it is (generally) not the case that for every $1\le j\le r$, both sides of this inequality are nonnegative, we cannot simply take the product of both sides over j and expect the inequality to be preserved.

So we require a different approach. By Lemma 6.15 (with $\mathcal{P},P$ replaced by $\mathcal{P}_j,P_j$), for any choice of nonnegative even integers $m_1,\dots,m_r$, we have

$$0\le\sum_{\substack{d_j\mid n,d_j\mid P_j\\ \omega(d_j)\le m_j}}\mu(d_j)-\sum_{d_j\mid n,d_j\mid P}\mu(d_j)\le\sum_{\substack{d_j\mid n,d_j\mid P\\ \omega(d_j)=m_j+1}}1\quad(1\le j\le r).\tag{6.49}$$

These bounds allow us to a coax a lower bound for the sifting function

$$s(n)=\prod_{j=1}^{r}\sum_{d_j\mid n,d_j\mid P_j}\mu(d_j)\tag{6.50}$$

out of the following general inequality:

Lemma 6.20 ([**FH00**, Lemma 1])**.** *Suppose that* $0\le x_j\le y_j$ *for* $1\le j\le r$. *Then*

$$x_1\cdots x_r\ge y_1\cdots y_r-\sum_{l=1}^{r}(y_l-x_l)\prod_{\substack{j=1\\ j\ne l}}^{r}y_j.$$

Proof. The result holds with equality when $r=1$. If the lemma holds for $r-1$ for a certain $r\ge 2$, then

$$\begin{aligned}y_1\cdots y_r-x_1\cdots x_r&=(y_1\cdots y_{r-1}-x_1\cdots x_{r-1})y_r+(x_1\cdots x_{r-1})(y_r-x_r)\\&\le(y_1\cdots y_{r-1}-x_1\cdots x_{r-1})y_r+(y_1\cdots y_{r-1})(y_r-x_r)\\&\le\sum_{l=1}^{r-1}(y_l-x_l)\prod_{\substack{j=1\\ j\ne l}}^{r}y_j+(y_r-x_r)\prod_{\substack{j=1\\ j\ne r}}^{r}y_j,\end{aligned}$$

which is just $\sum_{l=1}^{r}(y_l-x_l)\prod_{\substack{j=1\\ j\ne l,}}^{r}y_j$. So the result follows by induction. □

Assuming $m_1, \dots, m_r$ are nonnegative even integers, we apply Lemma 6.20 with

$$x_j := \sum_{d_j \mid n, d_j \mid P_j} \mu(d_j), \quad y_j := \sum_{\substack{d_j \mid n, d_j \mid P_j \\ \omega(d_j) \le m_j}} \mu(d_j).$$

Equation (6.49) implies that the hypotheses of Lemma 6.20 are satisfied and gives us an upper bound on the terms $y_l - x_l$. Using this bound in Lemma 6.20 and recalling (6.50), we obtain

$$s(n) \ge \prod_{j=1}^{r} \sum_{\substack{d_j \mid n, d_j \mid P_j \\ \omega(d_j) \le m_j}} \mu(d_j) - \sum_{l=1}^{r} \Bigg(\sum_{\substack{d_l \mid n, d_l \mid P_l \\ \omega(d_l) = m_l + 1}} 1 \Bigg) \prod_{\substack{j=1 \\ j \ne l}}^{r} \Bigg(\sum_{\substack{d_j \mid n, d_j \mid P_j \\ \omega(d_j) \le m_j}} \mu(d_j) \Bigg).$$

Summing over $n \in \mathcal{A}$ shows that

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}) \ge & \sum_{\substack{d_1, \dots, d_r \\ d_j \mid P_j, \omega(d_j) \le m_j}} \mu(d_1) \cdots \mu(d_r) A_{d_1 \cdots d_r} \\ & - \sum_{l=1}^{r} \sum_{\substack{d_1, \dots, d_r \\ d_j \mid P_j, \omega(d_j) \le m_j (j \ne l) \\ d_l \mid P_l, \omega(d_l) = m_l + 1}} \frac{\mu(d_1) \cdots \mu(d_r)}{\mu(d_l)} A_{d_1 \cdots d_r}. \end{aligned} \tag{6.51}$$

Writing $A_d = X\alpha(d) + r(d)$, the right-hand side of (6.51) becomes

$$X \prod_{j=1}^{r} \sum_{\substack{d_j \mid P_j \\ \omega(d_j) \le m_j}} \mu(d_j)\alpha(d_j) - X \sum_{l=1}^{r} \sum_{\substack{d_l \mid P_l \\ \omega(d_l) = m_l + 1}} \alpha(d_l) \prod_{j \ne l} \sum_{\substack{d_j \mid P_j \\ \omega(d_j) \le m_j}} \mu(d_j)\alpha(d_j), \tag{6.52}$$

up to an error term that is (with an absolute implied constant)

$$\ll \sum_{\substack{d_j \mid P_j \ (1 \le j \le r) \\ \theta_{d_1, \dots, d_r}}} |r(d_1 \cdots d_r)|.$$

Here $\theta_{d_1, \dots, d_r}$ denotes the condition that there exist $r-1$ indices j, $1 \le j \le r$, for which $\omega(d_j) \le m_j$, while the remaining index satisfies $\omega(d_j) \le m_j + 1$.

Assume, as we did for the upper bound, that $\alpha(p) < 1$ for each $p \in \mathcal{P}$. Lemma 6.6 implies that for each $1 \le j \le r$,

$$\sum_{\substack{d_j \mid P_j \\ \omega(d_j) \le m_j}} \mu(d_j)\alpha(d_j) \ge \prod_{p \in \mathcal{P}_j} (1 - \alpha(p)) > 0,$$

so that the main term in (6.52) is

$$X\bigg(1-\sum_{1\le l\le r}\frac{\sum_{d_l\mid P_l,\omega(d_l)=m_l+1}\alpha(d_l)}{\sum_{d_l\mid P_l,\omega(d_l)\le m_l}\mu(d_l)\alpha(d_l)}\bigg)\prod_{j=1}^{r}\sum_{\substack{d_j\mid P_j\\ \omega(d_j)\le m_j}}\mu(d_j)\alpha(d_j)$$

$$\ge X\prod_{p\in\mathcal{P}}(1-\alpha(p))\bigg(1-\sum_{1\le l\le r}\bigg(\sum_{\substack{d_l\mid P_l\\ \omega(d_l)=m_l+1}}\alpha(d_l)\Big/\prod_{p\in\mathcal{P}_l}(1-\alpha(p))\bigg)\bigg).$$

Summarizing, we have proved the following theorem:

Theorem 6.21 (Brun–Hooley sieve, lower bound)**.** *Let $\mathcal{P}=\dot{\bigcup}_{j=1}^{r}\mathcal{P}_j$ be a partition of $\mathcal{P}$. Suppose that $\alpha(p)<1$ for each $p\in\mathcal{P}$. For any choice of nonnegative even integers $m_1,\dots,m_r$, we have*

$$S(\mathcal{A},\mathcal{P})\ge X\prod_{p\in\mathcal{P}}(1-\alpha(p))\bigg(1-\sum_{j=1}^{r}\Big(\sum\nolimits^{(j)}\Big/\prod\nolimits^{(j)}\Big)\bigg)$$
$$+O\bigg(\sum_{\substack{d_j\mid P_j(1\le j\le r)\\ \theta_{d_1,\dots,d_r}}}|r(d_1\cdots d_r)|\bigg),$$

where $\prod^{(j)}$ and $\sum^{(j)}$ are defined, for $1\le j\le r$, by (6.38)*, and the implied constant is absolute.*

5.5. Applications of the lower bound. We now prove the two remarkable theorems of Brun mentioned in the introduction: Every large even integer is a sum of two 9-almost primes, and there exist infinitely many pairs of 9-almost primes differing by 2.

Our setup for attacking these problems is the same as that used in attacking the analogous upper bound problems considered in §5.3. For the first of these, we assume N is an even natural number, and we take $\mathcal{A}:=\{n(N-n):1\le n\le N\}$. As before, we let $\mathcal{P}$ be the set of all primes.

Suppose that we have a positive even integer N and a $u>1$ for which

$$S(\mathcal{A},\mathcal{P},N^{1/u})>0. \tag{6.53}$$

Then there exists an n, $1\le n\le N$, such that both n and $N-n$ have all their prime divisors exceeding $N^{1/u}$; since both n and $N-n$ are bounded by N, each must have at most u prime divisors. We will show that if we choose u large enough, (6.53) holds for all sufficiently large N (depending on u). Brun's results then follow from a quantitative determination of which u are "large enough".

For the most part, we may choose our sieving parameters as in §5.3, so that $X = N$ and α is given by (6.40). With u a parameter to be chosen later, we define the partition of $\mathcal{P} \cap [2, z]$ into sets $\mathcal{P}_j$ as in §5.3. However, the choice of the corresponding m_j requires more care.

To describe this choice, suppose for the moment that we have constructed a sequence $\{n_i\}_{i=1}^{\infty}$ of nonnegative even integers satisfying the two inequalities

$$\sum_{j=1}^{\infty} \frac{(2\log K_1)^{n_j+1}}{(n_j+1)!} < \frac{1}{K_1^2}, \tag{6.54}$$

$$\Gamma := 1 + \sum_{j=1}^{\infty} \frac{n_j}{K^{j-1}} < \infty, \tag{6.55}$$

where K and K_1 are given by (6.42). We fix $u > \Gamma$ and define (with same meaning of "∞" as in §5.3)

$$m_j = n_j \quad (1 \le j \le R), \qquad m_{R+1} = \infty.$$

Then for all large X, we have (recalling (6.43), (6.44), (6.45))

$$\sum_{j=1}^{R+1} \Big(\sum\nolimits^{(j)} \Big/ \prod\nolimits^{(j)} \Big) = \sum_{j=1}^{R} \Big(\sum\nolimits^{(j)} \Big/ \prod\nolimits^{(j)} \Big)$$
$$\le K_1^2 \sum_{j=1}^{R} \sum\nolimits^{(j)} \le K_1^2 \sum_{j=1}^{R} \frac{(2\log K_1)^{m_j+1}}{(m_j+1)!} \le 1 - \epsilon$$

for a positive constant ϵ, by (6.54). This implies that the main term in the lower bound

$$S(\mathcal{A}, \mathcal{P}) \ge X \prod_{p \in \mathcal{P}} (1 - \alpha(p)) \left(1 - \sum_{1 \le j \le R+1} \Big(\sum\nolimits^{(j)} \Big/ \prod\nolimits^{(j)} \Big) \right) + O\Bigg(\sum_{\substack{d_j \mid P_j (1 \le j \le R+1) \\ \theta_{d_1,\dots,d_{R+1}}}} |r(d_1 \cdots d_{R+1})| \Bigg), \tag{6.56}$$

is (cf. (6.46))

$$\gg X \prod_{p \le X^{1/u}} (1 - \alpha(p)) \gg X/(\log X)^2 \quad (X \to \infty).$$

The O-term can be treated much as in §5.3: The largest value of $d_1 \cdots d_{R+1}$ appearing as an argument of $r(\cdot)$ is bounded above by

$$X^{\frac{1}{u}\left(1 + \sum_{j=1}^{R} m_j/K^{j-1}\right)} X^{\log\log X \log\log\log X/\log X} \le X^{\Gamma/u + o(1)} \le X^{\delta}$$

for all large X, where $\delta := \frac{1}{2}(1+\Gamma/u)$. Notice that $\delta < 1$. The argument of §5.3 shows that the O-term in (6.56) is $\ll X^\delta \log X$, which is $o(X/(\log X)^2)$. So with this choice of parameters, we obtain (6.53) in the stronger form

$$S(\mathcal{A}, \mathcal{P}, X^{1/u}) \gg X/(\log X)^2 \quad (X \to \infty).$$

It remains to construct a suitable sequence $\{n_i\}$. It is not hard to see that (6.54) and (6.55) will be satisfied with the simple choice $n_i = b+2(i-1)$ $(i \geq 1)$, if we pick b to be a suitably large even natural number. However, this construction leads to an unnecessarily bloated value of Γ, so that while we still obtain a statement of the form "every large even N is a sum of two numbers with $O(1)$ prime factors", the $O(1)$ term dictating the number of summands is larger than we might like. We do better if we use the greedy algorithm to pick the first several n_i (which play the largest role in determining the size of Γ): Choose as many of the initial n_i to be 2 as (6.54) allows, then as many of the subsequent n_i to be 4 as allowed, etc.

Using a calculator or computer, we find that the sequence obtained in this way begins

$$n_1 = n_2 = n_3 = 2, \quad n_4 = \cdots = n_{10} = 4, \quad n_{11} = \cdots = n_{24} = 6.$$

Instead of continuing in this manner, we make the simple choice

$$n_{25} = 8 + 2(j-25) \quad (j \geq 25).$$

Then, setting $L := 2\log K_1$,

$$\begin{aligned}
&\frac{1}{K_1^2} - \sum_{j=1}^{\infty} \frac{(2\log K_1)^{n_j+1}}{(n_j+1)!} \\
&\quad \geq \frac{1}{K_1^2} - \sum_{j=1}^{3} \frac{L^3}{3!} - \sum_{j=4}^{10} \frac{L^5}{5!} - \sum_{j=11}^{24} \frac{L^7}{7!} - \sum_{j=25}^{\infty} \frac{L^{9+2(j-25)}}{(9+2(j-25))!} \\
&\quad \geq \frac{1}{K_1^2} - 3\frac{L^3}{3!} - 7\frac{L^5}{5!} - 14\frac{L^7}{7!} - \frac{L^9/9!}{1 - L^2/(11\cdot 10)} = 0.00003\ldots > 0,
\end{aligned}$$

so that (6.54) holds in this case. Also,

$$\begin{aligned}
\Gamma &= 1 + \sum_{j=1}^{3} \frac{2}{K^{j-1}} + \sum_{j=4}^{10} \frac{4}{K^{j-1}} + \sum_{j=11}^{24} \frac{6}{K^{j-1}} + \sum_{j=25}^{\infty} \frac{8+2(j-25)}{K^{j-1}} \\
&= 1 + \sum_{j=1}^{3} \frac{2}{K^{j-1}} + \sum_{j=4}^{10} \frac{4}{K^{j-1}} + \sum_{j=11}^{24} \frac{6}{K^{j-1}} + \frac{2(4K-3)}{K^{23}(K-1)^2} = 7.993\ldots.
\end{aligned}$$

Thus (6.55) holds. Moreover, we can take $u = 7.995$, say. Doing so, we obtain an even stronger theorem than that stated in the introduction: Every large enough even N may be represented as a sum of two natural numbers

each of which has no more than 7 prime divisors, and the number of such representations is $\gg X/(\log X)^2 = N/(\log N)^2$ as $N \to \infty$.

In like manner, one can show that there are $\gg x/(\log x)^2$ positive integers $n \le x$ for which both n and $n+N$ have no prime divisor $\le x^{1/7.995}$, uniformly in the choice of the even natural number N. Suppose now that N is fixed; then for large enough x, we have

$$n \le n + N \le x + N < (x^{1/7.995})^8;$$

it follows that there are

$$\gg_N x/(\log x)^2 \quad (x \to \infty)$$

integers $n \le x$ for which both n and $n+N$ have no more than 7 prime divisors. When $N = 2$ we obtain Brun's statement (with 9 replaced by the superior constant 7).

Note that K and K_1 in (6.42) were chosen to minimize the quantity Γ, which is the limiting factor in how small we are allowed to select u. Their numerical values were found by computer (cf. [**FH00**, pp. 347-348]).

6. An application to the Goldbach problem

While sieve methods are now part of the standard tool chest of analytic number theory, this was not always the case. In the monograph of Halberstam & Richert [**HR74**, p. 6], the story is told of how Landau left Brun's manuscript untouched in a drawer for six years until hearing of a striking application made by the Russian mathematician Schnirelmann [**Sch33**]:

Theorem 6.22. *There is an absolute constant S with the following property: Every integer $n > 1$ can be written as a sum of at most S prime numbers.*

Our objective in this section is to prove Theorem 6.22.

6.1. Schnirelmann density. Write $\mathbf{N}_0$ for the set of nonnegative integers. In what follows we use script letters to denote subsets of $\mathbf{N}_0$ and use the corresponding Roman letters for their counting functions. Even though such sets may contain zero, it is convenient to define our counting functions so that only positive elements are tallied; thus, e.g.,

$$A(n) = \#\{a \in \mathcal{A} : 1 \le a \le n\}.$$

If $\mathcal{A}, \mathcal{B} \subset \mathbf{N}$, we define the *sumset* $\mathcal{A} \oplus \mathcal{B}$ by

$$\mathcal{A} \oplus \mathcal{B} := \{a + b : a \in \mathcal{A}, b \in \mathcal{B}\}.$$

For $h \in \mathbf{N}$, we put

$$h\mathcal{A} := \overbrace{\mathcal{A} \oplus \cdots \oplus \mathcal{A}}^{h \text{ summands}}.$$

We say that $\mathcal{A}$ is a *basis of finite order* if $h\mathcal{A} = \mathbf{N}_0$ for some $h \in \mathbf{N}$. In this case the smallest such h is called the *order* of the basis. For example, if $\mathcal{A} = \{n^2 : n \in \mathbf{Z}\}$, then $\mathcal{A}$ is a basis of order 4. In fact, if k is any integer with $k \geq 2$, then $\{n^k : n \in \mathbf{N}_0\}$ is a basis of finite order by the Hilbert–Waring Theorem considered in Chapter 5.

For each subset $A \subset \mathbf{N}_0$, we define the *Schnirelmann density* $\delta(\mathcal{A})$ of $\mathcal{A}$ by

$$\delta(\mathcal{A}) := \inf_{n=1,2,3,\ldots} \frac{A(n)}{n}.$$

This definition is a bit odd; unlike (e.g.) the notion of asymptotic density, the presence (or absence) of small numbers in $\mathcal{A}$ has a disproportionate impact. The most extreme instance of this is that $\mathcal{A}$ automatically has Schnirelmann density zero whenever $1 \notin \mathcal{A}$. Moreover, the only way that a set $\mathcal{A}$ can have Schnirelmann density 1 is if $\mathcal{A}$ contains every natural number. Despite these peculiarities, the Schnirelmann density is a very convenient measure of size for questions in additive number theory. Indeed, Schnirelmann succeeded in proving the following very useful criterion for a set to be a basis of finite order:

Theorem 6.23 (Schnirelmann's basis theorem)**.** *Let $\mathcal{A}$ be a subset of $\mathbf{N}_0$ with $0 \in \mathcal{A}$ and $\delta(\mathcal{A}) > 0$. Then $\mathcal{A}$ is a basis of finite order.*

The proof requires two simple lemmas.

Lemma 6.24. *If $\mathcal{A}$ and $\mathcal{B}$ are sets of nonnegative integers, each containing 0, and $\delta(\mathcal{A}) + \delta(\mathcal{B}) \geq 1$, then $\mathcal{A} \oplus \mathcal{B} = \mathbf{N}_0$. In particular, if $0 \in \mathcal{A}$ and $\delta(\mathcal{A}) \geq 1/2$, then $2\mathcal{A} = \mathbf{N}_0$.*

Proof. We will show that each $n \in \mathbf{N}_0$ belongs to the sumset $\mathcal{A} \oplus \mathcal{B}$. Suppose that $a_0 = 0 < a_1 < a_2 < \cdots$ is an enumeration of $\mathcal{A}$ and that $0 = b_0 < b_1 < b_2 < \cdots$ is an enumeration of $\mathcal{B}$. Let $n \in \mathbf{N}_0$, and consider the following list of nonnegative integers from $[0, n]$:

$$0 = a_0, a_1, \ldots, a_{A(n)}, n = n - b_0, n - b_1, \ldots, n - b_{B(n)}.$$

This list has length

$$(A(n) + 1) + (B(n) + 1) \geq \delta(\mathcal{A})n + \delta(\mathcal{B})n + 2 \geq n + 2 > n + 1.$$

Since there are only $n + 1$ integers in the interval $[0, n]$, it must be that for some pair of i and j with $0 \leq i \leq A(n)$ and $0 \leq j \leq B(n)$, we have $a_i = n - b_j$. But then $n = a_i + b_j \in \mathcal{A} \oplus \mathcal{B}$. □

Lemma 6.25. *If $\mathcal{A}$ and $\mathcal{B}$ are sets of nonnegative integers, each containing 0, then $\delta(\mathcal{A} \oplus \mathcal{B}) \geq \delta(\mathcal{A}) + \delta(\mathcal{B}) - \delta(\mathcal{A})\delta(\mathcal{B})$.*

Proof. Let $n \in \mathbf{N}$, and let $0 < a_1 < a_2 < \cdots < a_{A(n)} \le n$ be a list of the elements of $\mathcal{A} \cap [1, n]$. Define intervals I_j for $0 \le j \le A(n)$ by putting $I_0 = (0, a_1)$, $I_1 = (a_1, a_2)$, $I_2 = (a_2, a_3)$, $\ldots$, $I_{A(n)-1} = (a_{A(n)-1}, a_{A(n)})$, and $I_{A(n)} = (a_{A(n)}, n]$. We now estimate $\#(\mathcal{A} \oplus \mathcal{B}) \cap I_j$ for each j.

For $j = 0$, we have $\#(\mathcal{A} \oplus \mathcal{B}) \cap I_0 \ge B(a_1 - 1)$, since if $b \in \mathcal{B} \cap [1, a_1 - 1]$, then $0 + b \in (\mathcal{A} \oplus \mathcal{B}) \cap I_0$. Similarly, for $1 \le j < A(n)$, we have $\#(\mathcal{A} \oplus \mathcal{B}) \cap I_j \ge B(a_{j+1} - a_j - 1)$, since if $b \in \mathcal{B} \cap [1, a_{j+1} - a_j - 1]$, then $a_j + b \in (\mathcal{A} \oplus \mathcal{B}) \cap I_j$. Finally, $\#(\mathcal{A} \oplus \mathcal{B}) \cap I_{A(n)} \ge B(n - a_{A(n)})$, since if $b \in \mathcal{B} \cap [1, n - a_{A(n)}]$, then $a_{A(n)} + b \in (\mathcal{A} \oplus \mathcal{B}) \cap I_{A(n)}$. Moreover, since $0 \in \mathcal{B}$, we know also that $\mathcal{A} \oplus \mathcal{B} \supset \mathcal{A}$. Hence,

$$\begin{aligned}(A \oplus B)(n) &\ge A(n) + \sum\nolimits_{i=0}^{A(n)} \#(\mathcal{A} \oplus \mathcal{B}) \cap I_j \\ &\ge A(n) + B(a_1 - 1) + \sum_{i=1}^{A(n)-1} B(a_{i+1} - a_i - 1) + B(n - a_n).\end{aligned}$$

Since $B(m) \ge \delta(\mathcal{B})m$ for each $m \in \mathbf{N}_0$, this is at least

$$\begin{aligned}A(n) + \delta(\mathcal{B}) &\left((a_1 - 1) + \sum_{i=1}^{A(n)-1} (a_{i+1} - a_i - 1) + n - a_{A(n)} \right) \\ &= A(n) + \delta(\mathcal{B})(n - A(n)) = A(n)(1 - \delta(\mathcal{B})) + \delta(\mathcal{B}).\end{aligned}$$

But $A(n) \ge \delta(\mathcal{A})n$, so that

$$\begin{aligned}(A \oplus B)(n) &\ge \delta(\mathcal{A})n(1 - \delta(\mathcal{B})) + \delta(\mathcal{B})n \\ &= n(\delta(\mathcal{A}) + \delta(\mathcal{B}) - \delta(\mathcal{A})\delta(\mathcal{B})).\end{aligned}$$

Since n was arbitrary, the assertion of the lemma follows from the definition of Schnirelmann density. □

Proof of Theorem 6.23. Taking $\mathcal{A} = \mathcal{B}$ in Lemma 6.25, we find $\delta(2\mathcal{A}) \ge 2\delta(\mathcal{A}) - \delta(\mathcal{A})^2$. Said differently, $1 - \delta(2\mathcal{A}) \le (1 - \delta(\mathcal{A}))^2$. Starting from this inequality, an easy induction shows that for every $k \ge 1$,

$$1 - \delta(2^k\mathcal{A}) \le (1 - \delta(\mathcal{A}))^{2^k}.$$

Since $\delta(\mathcal{A}) > 0$, we can choose a natural number k for which the right-hand side of this inequality is at most $1/2$. Then $\delta(2^k\mathcal{A}) \ge 1/2$, and so $2^{k+1}\mathcal{A} = \mathbf{N}_0$ by Lemma 6.24. So $\mathcal{A}$ is a basis of order at most 2^{k+1}. □

Remark. A theorem of Mann [**Man42**], strengthening Lemma 6.25, asserts that if $\mathcal{A}$ and $\mathcal{B}$ are subsets of $\mathbf{N}_0$ with $0 \in \mathcal{A} \cap \mathcal{B}$, then $\delta(\mathcal{A} \oplus \mathcal{B}) \ge \min\{1, \delta(\mathcal{A}) + \delta(\mathcal{B})\}$. This had been conjectured by Landau & Schnirelmann. An immediate consequence of Mann's theorem is that under the hypotheses of Theorem 6.23, $\mathcal{A}$ is a basis of order at most $\lceil 1/\delta(\mathcal{A}) \rceil$. For a discussion of

Mann's theorem and subsequent related developments (including the important work of Kneser), see the volumes of Ostmann mentioned in the notes at the end of this chapter. There is also some discussion of these results in the appealing survey [**PS95**].

6.2. Proof of Theorem 6.22. Observe that if $\mathcal{A} \subset \mathbf{N}_0$ has positive lower density, in the sense that

$$\liminf_{x\to\infty} \frac{A(x)}{x} > 0, \tag{6.57}$$

then $\mathcal{B} := \{0,1\} \cup \mathcal{A}$ has positive Schnirelmann density. Indeed, (6.57) implies that for some $\delta_0 > 0$ and $N_0 \in \mathbf{N}$, we have $A(N) \geq \delta_0 N$ for all $N \geq N_0$. But then $\delta(\mathcal{B}) \geq \min\{\delta_0, 1/N_0\} > 0$. Since also $0 \in \mathcal{B}$, we may apply Theorem 6.23 to deduce that $\mathcal{B}$ is a basis of finite order. We will shortly make use of these observations for an appropriately chosen set $\mathcal{A}$.

Recall that for a natural number N, the number of ordered representations of N as a sum of two primes is denoted by $R(N)$. For each $N \geq 2$, we have

$$R(N) \ll \frac{N}{(\log N)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right). \tag{6.58}$$

(This was proved in §5.3 when N is even. If N is odd, then $R(N) \leq 2$ and so (6.58) is trivial.) We now let

$$\mathcal{A} := \{N \in \mathbf{N} : R(N) > 0\}.$$

We will prove the following:

Theorem 6.26. *The set $\mathcal{A}$ has positive lower density.*

Once this is proved, Theorem 6.23 follows easily. Indeed, let $\mathcal{B} = \mathcal{A} \cup \{0,1\}$, so that from the above discussion $\mathcal{B}$ is a basis of finite order $h \geq 1$, say. Then for every integer $n \geq 2$, we can write

$$n - 2 = p_1 + p_2 + \cdots + p_{2k} + \overbrace{1 + 1 + \cdots + 1}^{l \text{ summands}},$$

say, where the p_i are primes, k and l are nonnegative integers, and $k + l \leq h$. Then

$$n = p_1 + \cdots + p_{2k} + (l + 2).$$

Since $l + 2 \geq 2$, it can be written as a sum of 2s and 3s, where the number of summands is at most $(l + 2)/2 \leq h/2 + 1$. This means that n has a representation as a sum of at most $2k + h/2 + 1 \leq 5h/2 + 1$ primes. Theorem 6.23 follows with $S = 5h/2 + 1$.

The main tool needed in the proof of Theorem 6.26 is the upper bound (6.58). It is initially surprising that an upper bound for $R(N)$ would be of

use in establishing a lower density result. But this seeming paradox is easily explained: As we will see shortly, it is a simple matter to obtain a lower bound for $\sum_{N\le x} R(N)$. If, as (6.58) asserts, $R(N)$ is never too big, then the only way to account for the size of this lower bound is for there to be many terms for which $R(N)$ is nonzero. In other words, $\mathcal{A}$ must be fairly dense. We now make this precise.

Lemma 6.27. *As $x\to\infty$, we have $\sum_{N\le x} R(N) \gg x^2/(\log x)^2$.*

Proof. By Chebyshev's results from Chapter 3, we have $\pi(x/2) \gg x/\log x$ as $x\to\infty$. Thus

$$\sum_{N\le x} R(N) = \sum_{N\le x}\sum_{p+q=N} 1 = \sum_{p+q\le x} 1 \ge \left(\sum_{p\le x/2} 1\right)^2 \gg \frac{x^2}{(\log x)^2}. \qquad \square$$

Lemma 6.28. *As $x\to\infty$, we have $\sum_{N\le x} R(N)^2 \ll x^3/(\log x)^4$.*

Proof. From (6.58),

$$\begin{aligned}\sum_{N\le x} R(N)^2 &\ll \sum_{2\le N\le x}\left(\frac{N}{(\log N)^2}\prod_{p\mid N}\left(1+\frac{1}{p}\right)\right)^2\\ &\ll \frac{x^2}{(\log x)^4}\sum_{2\le N\le x}\left(\prod_{p\mid N}\left(1+\frac{1}{p}\right)\right)^2\\ &\ll \frac{x^2}{(\log x)^4}\sum_{2\le N\le x}\left(\sum_{d\mid N}\frac{1}{d}\right)^2.\end{aligned}$$

It remains to show that the outer sum is $O(x)$. For this, observe that for any natural numbers d_1 and d_2,

$$[d_1,d_2] \ge \max\{d_1,d_2\} \ge (d_1d_2)^{1/2},$$

so that

$$\begin{aligned}\sum_{N\le x}\left(\sum_{d\mid N}\frac{1}{d}\right)^2 &= \sum_{N\le x}\sum_{d_1\mid N}\sum_{d_2\mid N}\frac{1}{d_1d_2} = \sum_{d_1,d_2\le x}\frac{1}{d_1d_2}\sum_{\substack{N\le x\\ d_1\mid N, d_2\mid N}} 1\\ &\le \sum_{d_1,d_2\le x}\frac{1}{d_1d_2}\frac{x}{[d_1,d_2]} \le x\sum_{d_1,d_2\le x}\frac{1}{(d_1d_2)^{\frac{3}{2}}} \le x\left(\sum_{d=1}^{\infty} d^{-\frac{3}{2}}\right)^2 \ll x. \quad \square\end{aligned}$$

Proof of Theorem 6.26. Writing $R(N) = R(N) \cdot 1$, the Schwarz inequality and Lemmas 6.27 and 6.28 yield that

$$\frac{x^4}{(\log x)^4} \ll \left(\sum_{N \le x} R(N) \right)^2 = \left(\sum_{\substack{N \le x \\ R(N)>0}} R(N) \cdot 1 \right)^2$$
$$\le \sum_{\substack{N \le x \\ R(N)>0}} R(N)^2 \sum_{\substack{N \le x \\ R(N)>0}} 1 \ll \frac{x^3}{(\log x)^4} A(x),$$

so that $A(x) \gg x$ as $x \to \infty$. In other words, $\mathcal{A}$ has positive lower density. □

Notes

The results of this chapter barely begin to scratch the surface of modern sieve theory. Encyclopedic accounts of this subject include the monographs of Halberstam & Richert [**HR74**] and Greaves [**Gre01**]. The introductory texts of Schwarz [**Sch74**] and Cojocaru & Murty [**CM06**] take a more discursive approach. Another treatment of the Brun–Hooley sieve can be found in the the introduction to analytic number theory written by Bateman & Diamond [**BD04**].

Excellent references for additive number theory include Ostmann's two-volume work [**Ost56**] and Nathanson's book [**Nat96**]. Nathanson's text includes a proof of the following theorem of Vinogradov which should be compared with Theorem 6.22:

★ **Theorem 6.29** (Three primes theorem)**.** *Let $R_3(N)$ denote the number of ways of writing N as an ordered sum of three primes. As $N \to \infty$ through odd integers, we have*

$$R_3(N) \sim \prod_p \left(1 + \frac{1}{(p-1)^3}\right) \prod_{p \mid N} \left(1 - \frac{1}{p^2 - 3p + 3}\right) \frac{N^2}{2(\log N)^3}.$$

In particular, every sufficiently large odd integer is a sum of three primes.

It follows from Vinogradov's result that every large enough natural number is the sum of at most 4 primes. While Vinogradov's theorem has a similar flavor to Theorem 6.22, the proof, which depends on the circle method, requires substantially deeper input from prime number theory.

See [**KT05**] for a thorough survey of additive prime number theory.

Exercises

1. (Gandhi [**Gan71**], Golomb [**Gol74**]) For each set of natural numbers S, put $w(S) := \sum_{n\in S} 2^{-n}$. For each natural number k, let p_k denote the kth prime.
   (a) If S is the set of natural numbers coprime to $p_1 \cdots p_k$, show that $w(S) = \frac{1}{2} + \frac{1}{2^{p_{k+1}}} + E$ where $0 < E < \frac{1}{2^{p_{k+1}}}$.
   (b) Show that for the set S in (a), we have $w(S) = \sum_{d|p_1\cdots p_k} \frac{\mu(d)}{2^d-1}$.
   (c) Deduce that p_{k+1} is the unique integer for which
   $$1 < 2^{p_{k+1}} \left(\sum_{d|p_1\cdots p_k} \frac{\mu(d)}{2^d-1} - \frac{1}{2} \right) < 2.$$
2. (Cf. Nagell [**Nag22**, §3])
   (a) Let D be an integer that is not a square. Using the law of quadratic reciprocity, prove that there is a collection S (say) of $\frac{1}{2}\phi(4|D|)$ residue classes modulo $4|D|$ with the property that for each prime $p \nmid 4D$, $\left(\frac{D}{p}\right) = 1 \iff p \bmod 4|D| \in S$.
   (b) Deduce from (a) and the results of Chapter 4 that
   $$\sum_{p\le x,\ \left(\frac{D}{p}\right)=1} \frac{\log p}{p} = \frac{1}{2}\log x + O(1),$$
   where the implied constant may depend on D. (Thus, in a certain average sense, D is a square modulo precisely $\frac{1}{2}$ of all primes.)
   (c) Let $F(T)$ be a quadratic polynomial with integer coefficients. Using the sieve of Eratosthenes–Legendre, show that as $x \to \infty$, the number of $n \le x$ with $|F(n)|$ prime is $\ll_F x/\log\log x$. (The case when $F(T) = T^2 + 1$ is the third example of §3.2; cf. Exercise 22.)
3. Use the inclusion-exclusion principle to establish each of the following assertions about squarefree numbers:
   (a) The number of squarefree $n \le x$ is asymptotic to $\frac{1}{\zeta(2)}x = \frac{6}{\pi^2}x$ as $x \to \infty$.
   (b) The number of pairs of squarefree integers $n, n+2$ with $1 \le n \le x$ is asymptotic to $x\prod_p(1 - 2/p^2)$ as $x \to \infty$.
   (c) The number of ordered representations of a natural number N as a sum of two positive squarefree integers is asymptotic to
   $$N \prod_p \left(1 - \frac{2}{p^2}\right) \prod_{p^2|N} \frac{p^2-1}{p^2-2} \qquad (N \to \infty).$$

Hint: For each of (a)–(c), first sieve out the multiples of p^2 for $p \le z$, where $z = z(x) \to \infty$ slowly enough to keep the error term in check. To conclude, observe that almost no n are divisible by p^2 for some prime $p > z$, since $\sum_{p>z} \frac{1}{p^2}$ is $o(1)$.

4. (Rényi [**Rén55**])
   (a) Show that for each fixed integer $j \ge 0$, the set of natural numbers n with $\Omega(n) - \omega(n) = j$ possesses an asymptotic density d_j (say). Check that $\sum_{j=0}^{\infty} d_j = 1$.
   (b) Show that for all complex numbers z with $|z| < 2$, we have
   $$\sum_{j=0}^{\infty} d_j z^j = \frac{1}{\zeta(2)} \prod_p \left(1 - \frac{z}{p+1}\right)\left(1 - \frac{z}{p}\right)^{-1}.$$

5. (Hooley [**Hoo76**], Rieger [**Rie77**]) If m is an odd natural number, write $l(m)$ for the order of 2 modulo m.
   (a) Suppose $m \in \mathbf{N}$ is odd and squarefree and put $M := \operatorname{lcm}[m, l(m)]$. Show that $n \cdot 2^n$ runs through every residue class modulo m exactly M/m times as n runs over the integers $1, 2, 3, \ldots, M$.
   (b) Using the result of (a) and the sieve of Eratosthenes–Legendre, show that the set of $n \in \mathbf{N}$ for which $n \cdot 2^n + 1$ is prime has density zero. (Primes of the form $n \cdot 2^n + 1$ are called *Cullen primes*; the first several examples correspond to $n = 1$, 141, 4713, 5795, 6611, 18496, 32292.)

6. Let A and B be subsets of the natural numbers defined by
   $$A = \{n : n \mid 2^k - 1 \text{ for some positive integer } k\},$$
   $$B = \{n : n \mid 2^k + 1 \text{ for some positive integer } k\}.$$
   Prove that A has asymptotic density $\frac{1}{2}$ and B has asymptotic density 0.

7. (Cf. Luca [**Luc06**, Problem 190]) Let F_n denote the nth Fibonacci number, so that $F_0 = 0$, $F_1 = 1$, and for $n > 1$, $F_n = F_{n-1} + F_{n-2}$. Show that the set of n for which F_n can be written as a sum of two coprime squares has asymptotic density 1/2.

8. Show that for each $d \in \mathbf{N}$, the set of natural numbers n for which $d \mid \varphi(n)$ has asymptotic density 1. Deduce that the set of n for which $\gcd(n, \varphi(n)) = 1$ has density zero.

9. (Continuation; cf. Pillai [**Pil29**]) Let $\mathcal{V} := \{\varphi(m) : m \in \mathbf{N}\}$ be the image of the Euler φ-function, and let $V(x)$ be the number of $n \le x$ belonging to $\mathcal{V}$. Show that $V(x) = o(x)$. *Hint:* Divide the elements n of $\mathcal{V}$ into two classes, depending on whether or not n has a preimage m with only a "small" number of distinct odd prime divisors.

Remark. Maier & Pomerance [**MP88**] showed in 1988 that

$$V(x) = \frac{x}{\log x} \exp((C + o(1))(\log\log\log x)^2)$$

for a constant $C = 0.81781464640\ldots$. This improved upon earlier results of Erdős, Hall, and Pomerance. The (somewhat complicated) exact order of magnitude of $V(x)$ was subsequently determined by Ford [**For98a, For98b**].

10. (Blecksmith, Erdős & Selfridge [**BES99**]) Say that a prime p is *a cluster prime* if every even natural number $n < p-2$ can be written in the form $q - q'$, where q and q' are primes $\leq p$.
    (a) Check (perhaps with the aid of a computer) that every prime $p < 97$ is a cluster prime, but that $p = 97$ is not.
    (b) Show that if p is a cluster prime, then for every integer $3 \leq t \leq p-3$, the number of primes in the closed interval $[p-t, p]$ is $\gg \log t$, where the implied constant is absolute. In other words, the primes to the left of p have to "cluster" around p.
    (c) Show that contrary to what one might expect from (a), the cluster primes are comparatively rare: For every k, the number of cluster primes up to x is $O_k(x/(\log x)^k)$ as $x \to \infty$.
11. (Cf. Erdős [**Erd36**]) For each $r \in \mathbf{N}$, define a function $p_r\colon \mathbf{N} \to \{\text{primes}\} \cup \{\infty\}$ by setting $p_r(n)$ equal to the rth smallest prime factor of n if n has at least r distinct prime factors and putting $p_r(n) = \infty$ otherwise. Observe that $p_1(n) < p_1(n+1)$ precisely when n is even. In particular, $p_1(n) < p_1(n+1)$ on a set of asymptotic density $1/2$. Show that for each fixed r, we have $p_r(n) < p_r(n+1)$ on a set of asymptotic density $1/2$.

    Remark. For each $n > 1$, put $P(n)$ equal to the largest prime factor of n, and put $P(1) = 0$. In the 1930s, Erdős conjectured that $P(n) < P(n+1)$ on a set of asymptotic density $1/2$. This remains open. Erdős & Pomerance have shown that each of the inequalities $P(n) > P(n+1)$ and $P(n) < P(n+1)$ holds for a positive proportion of the natural numbers [**EP78**].

12. For each prime p, let p' be the prime immediately following p. Show that for each $\epsilon > 0$, there is a $K > 0$ for which the following holds: For large x, all but at most $\epsilon x/\log x$ primes $p \leq x$ satisfy

$$\frac{1}{K}\log x \leq p' - p \leq K \log x.$$

    Remark. It is conjectured (see, e.g., [**Sou07**, Conjecture 1]) that for each fixed $K > 0$, the number of $p \leq x$ with $p' - p \leq K \log x$ is asymptotically $(1 - e^{-K})x/\log x$ as $x \to \infty$.

13. Call a prime p *M-reclusive* if $|q-p| > M$ for every prime $q \neq p$. Show that for every $M > 0$ and every $k \in \mathbf{N}$, there are infinitely many k-tuples of consecutive primes all of which are M-reclusive. (This strengthens the result of Exercise 4.12.)

14. (Erdős & Nathanson [**EN96**]) Let p_n be the nth prime number (in the usual, increasing order). Use Theorem 6.19 to show that for each $\lambda > 2$, the series
$$\sum_{n=1}^{\infty} \frac{1}{n(\log\log 3n)^{\lambda}(p_{n+1}-p_n)}$$
converges. It is conjectured that this result is the best possible, in the sense that the series diverges when $\lambda = 2$.

15. For each even natural number N, let $R^*(N)$ be the number of *unordered* representations of N as a sum of two primes. Then
$$R^*(N) \leq \pi(N-2) - \pi((N-1)/2),$$
with equality holding exactly when $N-p$ is prime for each prime p with $N/2 \leq p \leq N-2$. Use the estimate (3.21) in conjunction with Theorem 6.17 to prove that this upper bound is attained for only finitely many N.

    Remark. It has been shown by Deshouillers et al. [**DGNP93**] that $N = 210$ is the largest value for which the upper bound is achieved.

16. By modifying the argument of §5.5, show that the number of representations of an even natural number N as a sum of two 7-almost primes is $\gg \frac{N}{(\log N)^2} \prod_{p|N, p>2} \frac{p-1}{p-2}$, as $N \to \infty$.

17. (Brun) Prove the following theorems of Brun, announced in [**Bru19b**]:
    (a) Every infinite arithmetic progression $a \bmod m$ with $\gcd(a,m) = 1$ contains infinitely many 5-almost primes. (Naturally, Dirichlet's theorem is off-limits here.)
    (b) If x is sufficiently large, there is always an 11-almost prime in the interval $(x, x+\sqrt{x}]$.

    Suggestion: Imitate the lower bound applications of the text, including the selection of the first several m_j by the greedy algorithm, but begin instead with the values $K = 2.49, K_1 = 2.50$.

18. (A general version of Brun's method) Fix a natural number k.
    (a) Let $A > 0$. Suppose that to each prime $p \leq x^A$, we associate $k_p \leq k$ residue classes modulo p. Show that the number of natural numbers $n \leq x$ avoiding all of these residue classes is
$$\ll_{k,A} x \prod_{p \leq x^A} \left(1 - \frac{k_p}{p}\right) \quad (\text{for } x > 0),$$

where the implied constant is independent of the particular choice of residue classes.

(b) Show that there is a constant $B > 0$, depending only on k, with the following property: If we choose $k_p \le k$ residue classes modulo p for each prime $p \le x^B$, then the number of natural numbers $n \le x$ avoiding all these classes is

$$\gg_k x \prod_{p \le x^B} \left(1 - \frac{k_p}{p}\right) \quad (\text{for } x \to \infty),$$

again uniformly in the particular choice of residue classes.

Hint: Use the Chinese remainder theorem to construct a polynomial F for which $p \mid F(n)$ precisely when n falls into one of the k_p chosen residue classes mod p.

Remark. From (a) and (b) we may rederive the results given in the text regarding the twin prime and Goldbach problems, with a slight loss of precision (in that in our lower bound applications, we obtain r-almost primes with an unspecified constant r in place of $r = 7$). For the twin prime problem, the forbidden residue classes are 0 and $-2 \bmod p$. For the Goldbach problem, the forbidden classes are 0 and $N \bmod p$.

When one sees references to "Brun's method" in the literature, often the author has the results of (a) and (b) in mind.

N. B. The results of Problem 18 suffice to handle all the sieving situations that arise in the remaining exercises in this chapter.

19. Suppose that $y = y(x)$ is a positive-valued function of x for which $\frac{\log y}{\log x} \to 0$ as $x \to \infty$. Show that as $x \to \infty$, all but $o(x)$ of the natural numbers $n \le x$ have a prime factor $> y$. In other words, $\Psi(x, y) = o(x)$.

20. (Hardy & Littlewood [**HL23**]) Show that $\pi(y + x) - \pi(y) \ll \frac{x}{\log x}$ for $y \ge 0$ and $x \ge 2$, where the implied constant is absolute.

21. ("Brun–Titchmarsh inequality" [**Tit30**]) Let $x \ge 2$. Suppose that a and m are coprime integers with $1 \le m < x$. Prove that

$$\pi(x; m, a) \ll \frac{x}{\varphi(m) \log \frac{x}{m}},$$

where the implied constant is absolute. (Recall that $\pi(x; m, a)$ denotes the number of primes $p \le x$ with $p \equiv a \pmod m$.) Is this still true without the assumption that a and m are relatively prime?

22. Suppose $F(T) \in \mathbf{Z}[T]$ is irreducible over $\mathbf{Q}$ and that the leading coefficient of $F(T)$ is positive. For each natural number d, let $\nu(d)$ denote the number of roots of F modulo d.

(a) A theorem of Landau (cf. [**Lan02**, eq. (67)]) asserts that for $x \geq 3$,

$$\sum_{p \leq x} \frac{\nu(p)}{p} = \log\log x + C_F + O_F\left(\frac{1}{\log x}\right),$$

where C_F is a constant depending on F. Deduce from this result and the Brun–Hooley sieve that the number of $n \leq x$ for which $F(n)$ is prime is $\ll_F x/\log x$, again for $x \geq 3$.

(b) Now impose the additional hypothesis that there is no prime p that divides $F(n)$ for every $n \in \mathbf{Z}$. Show that there is an $r \in \mathbf{N}$, depending only on the degree g of F, with the property that $F(n)$ is an r-almost prime for infinitely many natural numbers n.

Remark. Richert [**Ric69**] has shown that one can take $r = g + 1$.

23. (Yang [**Yan82**]; see also Webb [**Web70**]) Using the identities

$$\frac{4}{n} = \begin{cases} \frac{1}{n(k+1)k} + \frac{1}{n(k+1)} + \frac{1}{qk} & \text{if } n = (4k-1)q, \\ \frac{1}{nk} + \frac{1}{nqk} + \frac{1}{qk} & \text{if } n+1 = (4k-1)q, \\ \frac{1}{nk} + \frac{1}{nk(qk-1)} + \frac{1}{qk-1} & \text{if } n+4 = (4k-1)q, \\ \frac{1}{nk} + \frac{1}{k(qk-n)} + \frac{1}{n(qk-n)} & \text{if } 4n+1 = (4k-1)q, \end{cases}$$

show that the number of $n \leq x$ for which (6.22) is unsolvable is $\ll x/(\log x)^2$ as $x \to \infty$. Deduce that the sum of the reciprocals of all n of this kind converges.

Remark. Vaughan [**Vau70**] has shown that the number of $n \leq x$ for which (6.22) is unsolvable is $\ll x \exp(-c(\log x)^{2/3})$ for a positive constant c.

24. (Erdős [**Erd35c**]) In Exercise 3.23, we proved that a typical natural number $n \leq x$ has about $\log\log x$ prime factors. One may wonder whether such a result continues to hold if one restricts n to certain special classes of numbers. Here we treat numbers of the form $p - 1$, where p is prime. (Such numbers are important, for example, in the study of the Euler φ-function.) We show that we do indeed have such a result, and that in fact for each $\epsilon > 0$,

$$\#\{p \leq x : |\omega(p-1) - \log\log x| > \epsilon \log\log x\} \ll_\epsilon x/(\log x)^{1+\delta},$$

where $\delta > 0$ depends on ϵ.

(a) Assume $x \geq 3$. Show that all but $O(x/(\log x)^2)$ natural numbers $n \leq x$ possess both of the following properties:

(i) the largest prime factor $P(n)$ (say) of n satisfies $P(n) > x^{1/(6\log\log x)}$,

(ii) n is not divisible by $P(n)^2$,

Hint: Use the result of Exercise 3.32 to handle condition (i).

(b) For each nonnegative integer k, let N_k be the number of primes $p \le x$ for which $p-1$ has both properties (i) and (ii) and satisfies $\omega(p-1) = k$. Show that

$$N_k \le \sum_{\substack{a \le x^{1-1/(6\log\log x)} \\ \omega(a)=k-1}} \sum_{\substack{p \le x \\ a \mid p-1 \text{ and } \frac{p-1}{a} \text{ is prime}}} 1.$$

(c) Show that for each natural number $a < x$,

$$\sum_{\substack{p \le x \\ a \mid p-1 \text{ and } \frac{p-1}{a} \text{ is prime}}} 1 \ll \frac{x}{\varphi(a)(\log \frac{x}{a})^2},$$

with an absolute implied constant.

(d) Convince yourself that

$$\sum_{\substack{a \le x \\ \omega(a)=k-1}} \frac{1}{\varphi(a)} \le \frac{1}{(k-1)!} \left(\sum_{p^l \le x} \frac{1}{\varphi(p^l)} \right)^{k-1},$$

where the right-hand sum is over primes and prime powers $p^l \le x$.

(e) Show that for a certain absolute constant C,

$$N_k \ll \frac{x(\log\log x)^2}{(\log x)^2} \frac{(\log\log x + C)^{k-1}}{(k-1)!},$$

uniformly in k. Complete the proof by summing this estimate over $k < (1-\epsilon)\log\log x$ and $k > (1+\epsilon)\log\log x$.

25. (Erdős, ibid.) Prove that if $\epsilon > 0$ is sufficiently small, then the following holds: As $x \to \infty$, there are $\gg_\epsilon x/\log x$ primes $p \le x$ for which the largest prime divisor of $p+1$ is bounded by $x^{1-\epsilon}$. *Hint:* If $p+1 \le x$ and $p+1$ has a prime divisor at least $x^{1-\epsilon}$, then $(p+1)/a$ is prime for some natural number $a \le x^\epsilon$.

26. (Luca [**Luc07**]) Show that the number of natural numbers not exceeding x which can be written in the form $p^2 - q^2$, where p and q are primes, is $\ll x/\log x$.

27. Call the natural number n *twinnish* if $d + n/d + 1$ is prime for every d dividing n. If p is the smaller member of a twin prime pair, then p is twinnish, but there are many other such n, for example $n = 21$ and (less obviously) $n = 190757 = 7^2 \cdot 17 \cdot 229$. Prove or disprove: $\sum \frac{1}{n} < \infty$, where the sum is extended over all twinnish numbers n.

28. (Hardy & Littlewood [**HL23**], cf. Landau [**Lan00**]) Let $R(N)$ be the number of ordered representations of N as a sum of two primes. Conjecture 3.19 asserts that as $N \to \infty$ through even numbers,

$$R(N) = (A + o(1)) \left(\prod_{p \mid N} \frac{p-1}{p-2} \right) \frac{N}{(\log N)^2}, \tag{6.59}$$

where

$$A = 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2} \right). \tag{6.60}$$

This differs from what a naive sieve argument would suggest, namely that (6.59) holds with

$$A = 8 \exp(-2\gamma) \prod_{p>2} \left(1 - \frac{1}{(p-1)^2} \right). \tag{6.61}$$

In this exercise we outline a proof that (6.61) cannot be correct. In fact, we show that if an asymptotic relation of the form (6.59) holds, then A must be given by (6.60).

(a) Use the prime number theorem to show that $\sum_{N \le x} R(N) \sim \frac{1}{2} \frac{x^2}{(\log x)^2}$ as $x \to \infty$.

(b) Deduce from (a) that as $x \to \infty$,

$$\sum_{2 \le N \le x} \frac{R(N)}{N/(\log N)^2} \sim x.$$

(c) Put $g(N) := \prod_{p \mid N, p>2} \frac{p-1}{p-2}$ for each N, and define an arithmetic function h by the relation $g(N) = \sum_{d \mid N} h(d)$. Show that h is supported on odd, squarefree positive integers, and that as $x \to \infty$,

$$\frac{1}{x} \sum_{\substack{N \le x \\ N \text{ even}}} g(N) \to \frac{1}{2} \sum_{d \text{ odd}} \frac{h(d)}{d} = \frac{1}{2} \prod_{p>2} \frac{(p-1)^2}{p(p-2)}.$$

(d) Use the result of (c) and the purported relation (6.59) to derive another asymptotic formula for $\sum_{2 \le N \le x} \frac{R(N)}{N/(\log N)^2}$ which, when compared with that of (b), proves (6.60).

Remark. The methods used to prove Vinogradov's three primes theorem can be employed to show that in fact the relation (6.59) with A given by (6.60) holds for almost all even natural numbers N (see, e.g, [**Vau97**, §3.2]). More precisely, (6.59) holds (with this A) as $N \to \infty$ through even numbers, provided we exclude a particular set of even numbers N of asymptotic density zero.

29. (Landau [**Lan30**]) Show that under the hypotheses of Theorem 6.23, the set $\mathcal{A}$ is a basis of order at most $2\lfloor 1/\delta(\mathcal{A})\rfloor$.

30. Say that a set $\mathcal{A} \subset \mathbf{N}_0$ is an *asymptotic basis of finite order* if $\mathbf{N} \setminus h\mathcal{A}$ is finite for some $h \in \mathbf{N}$.
    (a) Show that if $a_1, \dots, a_k \in \mathbf{N}$ and $\gcd(a_1, \dots, a_k) = 1$, then every sufficiently large natural number can be written in the form $\sum_{i=1}^{k} a_i x_i$, where each $x_i \in \mathbf{N}_0$.
    (b) Let $\mathcal{A}$ be a subset of $\mathbf{N}_0$. Suppose that $0 \in \mathcal{A}$, that $\mathcal{A}$ has positive lower density (i.e., (6.57) holds), and that there is no integer $d > 1$ dividing each $a \in \mathcal{A}$. Show that $\mathcal{A}$ is an asymptotic basis of finite order.

31. (Landau, ibid.; see also Nathanson [**Nat87a**]) Suppose $\mathcal{P}$ is a set of primes with the property that
$$\liminf_{x\to\infty} \frac{\#\{p \in \mathcal{P} : p \le x\}}{x/\log x} > 0.$$
Show that there is a constant $S_{\mathcal{P}}$ with the property that every sufficiently large natural number is the sum of at most $S_{\mathcal{P}}$ primes all of which belong to $\mathcal{P}$.

32. (Prachar [**Pra52**]) Show that for large x, there are $\gg x$ natural numbers $n \le x$ that can be written in the form $q - p$, where $p, q \le x$ are primes. *Hint:* Adapt the second-moment method appearing in the proof of Schnirelmann's theorem.

33. (Continuation) For each prime p, write p' for the prime immediately following p. Show that for some constant $K > 0$, the following holds: For all large x, there are $\gg \log x$ natural numbers $n \le K \log x$ which can be written in the form $p' - p$ for some prime $p \le x$. *Hint:* Use Exercise 12.

34. (Romanov [**Rom34**]) Let $r(n)$ be the number of representations of n in the form $2^k + p$, where p is prime and $k \ge 1$. In this exercise and the next, we sketch a proof that $r(n) > 0$ on a set of positive lower density. In Exercise 36, we prove the complementary result that $r(n) = 0$ on a set of odd numbers of positive density.
    (a) Show that for all natural numbers n, we have $\sum_{d|n} \frac{1}{d} \ll \log\log 3n$.
    (b) For each odd integer d, let $l(d)$ denote the order of 2 modulo d. Show that if $l(d) \le x$, then d divides $D := \prod_{1\le k\le x}(2^k - 1)$. Deduce from (a) that $\sum_{l(d)\le x} d^{-1} \ll \log(2x)$ for $x \ge 1$.
    (c) Using partial summation, prove that $\sum_{\substack{d \text{ odd}\\ d\ge 1}} \frac{1}{d\cdot l(d)} < \infty$.

35. (Continuation)
    (a) Show that $\sum_{n\le x} r(n) \gg x$ as $x \to \infty$.

(b) Show that $\sum_{n\le x} r(n)^2$ does not exceed the number of solutions (p_1, p_2, k_1, k_2) to

$$p_2 - p_1 = 2^{k_1} - 2^{k_2},$$

where p_1, p_2 are primes $\le x$ and $1 \le k_1, k_2 \le \log x/\log 2$.

(c) Show that the number of solutions as in (b) is $\ll x$. *Hint:* To estimate the number of solutions with $k_1 \ne k_2$, use Theorem 6.19 and the result of Exercise 34(c).

(d) Deduce from (a)–(c), and the Cauchy–Schwarz inequality that there are $\gg x$ natural numbers $n \le x$ for which $r(n) > 0$.

36. (Continuation; Erdős [**Erd50b**], following [**Sie88**, Chapter XII])

(a) Check that every integer k belongs to at least one of the congruence classes 0 mod 2, 0 mod 3, 1 mod 4, 3 mod 8, 7 mod 12, 23 mod 24.

(b) Suppose $n \equiv 1 \pmod 3$, $n \equiv 1 \pmod 7$, $n \equiv 2 \pmod 5$, $n \equiv 2^3 \pmod{17}$, $n \equiv 2^7 \pmod{13}$, and $n \equiv 2^{23} \pmod{241}$. Show that for every integer $k \ge 0$, the number $n - 2^k$ is divisible by some prime from the set $\{3, 5, 7, 13, 17, 241\}$.

(c) Suppose that in addition to the congruences in (b), we require also that $n \equiv 1 \pmod 2$ and $n \equiv 3 \pmod{31}$. Show that the positive n satisfying all of these congruences comprise an infinite arithmetic progression of odd integers n with $r(n) = 0$.