1
Introduction

1.1 The sieve problem

Let $\mathcal{P}$ be a finite set of primes $\{p\}$ (the symbol p denotes a prime throughout Part I of this book) and let

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

(Later, starting in Chapter 3, we shall let $\mathcal{P}$ denote an infinite set of primes and use $\mathcal{P}_z$ to denote the finite set $\mathcal{P}\cap[2,\, z)$, i.e., $\mathcal{P}$ truncated at z.) The indicator function of the set of all integers n coprime with P, that is, having no divisors in $\mathcal{P}$, is expressed in terms of the Moebius μ function by

$$\sum_{d|(n,P)} \mu(d) = \begin{cases} 1, & (n,P)=1, \\ 0, & \text{otherwise.} \end{cases} \tag{1.1}$$

We call $\mathcal{P}$ a *sieve* and say that $\mathcal{P}$ *sifts out* an integer n if $(n,P)>1$.

Let $\mathcal{A}$ be a finite integer sequence, taking account of possible repetitions. An example of such a sequence is

$$\mathcal{A} = \{n^2+1 : -9 \le n \le 11\}.$$

When we apply the sieve $\mathcal{P}$ to $\mathcal{A}$—we might say alternatively, when we put, or filter, $\mathcal{A}$ through $\mathcal{P}$, or *sift* $\mathcal{A}$ by $\mathcal{P}$—the elements of $\mathcal{A}$ that remain *unsifted* are those that are coprime with P, and their number

$S(\mathcal{A},\mathcal{P})$ is given by

$$S(\mathcal{A},\mathcal{P}) := |\{a\in\mathcal{A}\colon (a,P)=1\}| = \sum_{a\in\mathcal{A}}\sum_{d\mid(a,P)}\mu(d)$$
$$= \sum_{a\in\mathcal{A}}\sum_{\substack{d\mid a\\ d\mid P}}\mu(d) = \sum_{d\mid P}\mu(d)\sum_{\substack{a\in\mathcal{A}\\ d\mid a}}1\,.$$

Writing $\mathcal{A}_d := \{a\in\mathcal{A}\colon d\mid a\}$, we have the *Eratosthenes–Legendre formula*

$$S(\mathcal{A},\mathcal{P}) = \sum_{d\mid P}\mu(d)|\mathcal{A}_d|\,. \tag{1.2}$$

Example 1.1. Take $\mathcal{A}=\{n\in\mathbb{N}\colon n\le x\}$ and take $\mathcal{P}$ to be the set of all primes p not exceeding $x^{1/2}$. Then $|\mathcal{A}_d|=[x/d]$ and, by the famous observation of Eratosthenes, the identity for $S(\mathcal{A},\mathcal{P})$ yields the prime counting formula

$$\pi(x)-\pi(x^{1/2})+1 = \sum_{d\mid P}\mu(d)\Big[\frac{x}{d}\Big], \qquad P=\prod_{p\le x^{1/2}}p.$$

We can think of two natural ways to write the sum: either as

$$x\prod_{p\le x^{1/2}}\Big(1-\frac{1}{p}\Big)+\sum_{d\mid P}\mu(d)\Big(\Big[\frac{x}{d}\Big]-\frac{x}{d}\Big),$$

or as

$$x\sum_{\substack{d\mid P\\ d\le x}}\frac{\mu(d)}{d}+\sum_{\substack{d\mid P\\ d\le x}}\mu(d)\Big(\Big[\frac{x}{d}\Big]-\frac{x}{d}\Big).$$

In the first way, the leading term does suggest the correct order of magnitude of $\pi(x)$, but it turns out that the sum of the "remainders" has the same order of magnitude. The second way appears to be more promising, but it turns out that here we do not know how to handle either sum!

1.2 Some basic hypotheses

In the above example we know, of course, how the sequence $\mathcal{A}$ is distributed in the residue classes $0 \mod d$, $d\mid P$; in fact, the corresponding information is available for many integer sequences $\mathcal{A}$ occurring in the

literature and takes the form, which henceforward we assume, that *there exists a convenient approximation X to $|\mathcal{A}|$ and a non-negative multiplicative arithmetic function $\omega(\cdot)$ such that*

$$0 \le \omega(p) < p \quad (p \in \mathcal{P}), \qquad \omega(p) = 0 \quad (p \notin \mathcal{P}), \tag{1.3}$$

and such that the remainder terms

$$r_{\mathcal{A}}(d) := |\mathcal{A}_d| - \frac{\omega(d)}{d} X \qquad (d \mid P) \tag{1.4}$$

are suitably small, at least on average, over some restricted range of values of d. (In a naive sense, the number $\omega(p)/p$ is the probability that the prime p of $\mathcal{P}$ is a divisor of elements in $\mathcal{A}$.) With this assumption, we obtain

$$\begin{aligned} S(\mathcal{A},\mathcal{P}) &= X \sum_{d \mid P} \mu(d) \frac{\omega(p)}{p} + \sum_{d \mid P} \mu(d) r_{\mathcal{A}}(d) \\ &= X \prod_{p \in \mathcal{P}} \Big(1 - \frac{\omega(p)}{p}\Big) + \sum_{d \mid P} \mu(d) r_{\mathcal{A}}(d). \end{aligned}$$

Again, unless $\mathcal{P}$ is *very* sparse, we expect the remainder sum to be too large to derive asymptotics for $S(\mathcal{A},\mathcal{P})$, but we have the impression nevertheless that $S(\mathcal{A},\mathcal{P})$ should be measured in terms of the magnitude of the "leading" term

$$X \prod_{p \in \mathcal{P}} \Big(1 - \frac{\omega(p)}{p}\Big) =: XV(\mathcal{P}),$$

say. *The aim of a sieve method is to modify the Moebius function in the indicator function* (1.1) *in a way that allows us to approximate $S(\mathcal{A},\mathcal{P})$ from above, and sometimes from below, with some accuracy, and to obtain asymptotics for $S(\mathcal{A},\mathcal{P})$ when $\mathcal{P}$ is sparse.*

It is instructive to see why we assume that $\omega(p) < p$ holds for all $p \in \mathcal{P}$. Otherwise—that is, if there existed a prime $p^* \in \mathcal{P}$ for which $\omega(p^*)/p^*$ equals (or is very near to) 1—we would have

$$|\mathcal{A}_{p^*}| - X \approx |\mathcal{A}_{p^*}| - X\omega(p^*)/p^*,$$

and the last quantity is small by hypothesis, as is $X - |\mathcal{A}|$ as well. It follows that $|\mathcal{A}| - |\mathcal{A}_{p^*}|$ is small, i.e., most members of $\mathcal{A}$ are multiples of p^*. After these elements are sifted out, little would be left in $\mathcal{A}$—or for us to say.

Appeal to probabilistic thinking is often helpful in arithmetic investigations but tends to fall short when it comes to supplying proofs. The usual reason is that such thinking is based upon a probabilistic model involving a sequence of *independent* events, whereas the actual arithmetical "events" being modeled—in our case, "divisibility of elements of $\mathcal{A}$ by primes p from $\mathcal{P}$"—have a poor independence relation for sets of primes whose products have a size comparable to X. If these events were independent, then indeed we should expect $XV(\mathcal{P})$ to be a true measure of $S(\mathcal{A},\mathcal{P})$; instead, we have seen in the classical case of sifting the interval $[1,x]$ by the primes not exceeding $x^{1/2}$, that there

$$XV(\mathcal{P}) = x \prod_{p\le x^{1/2}} \Big(1-\frac{1}{p}\Big) \sim x\frac{e^{-\gamma}}{\log x^{1/2}} = 2e^{-\gamma}\frac{x}{\log x} \quad \text{as} \quad x\to\infty,$$

by the well-known Mertens' product formula ([HW79], Theorem 429) and $2e^{-\gamma} = 1.122918\ldots$, whereas by the Prime Number Theorem

$$\pi(x) \sim \frac{x}{\log x} \quad \text{as} \quad x\to\infty.$$

In contrast, suppose we sift $[1,x]$ by a "thin" infinite sequence of primes $\mathcal{P}\colon p_1 < p_2 < \ldots$ such that

$$\sum_{j=1}^{\infty}\frac{1}{p_j} < \infty.$$

In this case the density of integers divisible by none of the primes of $\mathcal{P}$ is indeed

$$\prod_{j=1}^{\infty}\Big(1-\frac{1}{p_j}\Big).$$

1.3 Prime g-tuples

Before we begin our account in earnest, we consider another example more relevant to our main objective. The inspiration for this example is the famous *twin prime conjecture*, which asserts that there are infinitely many pairs of positive integers $(n, n+2)$, which are both prime numbers. The sieve method of Brun broke new ground by producing an upper bound for the number of pairs of twin primes in any interval $[1, x]$, but the original conjecture remains unproved.

Except for the example (2, 3), there is no other pair of primes of the form $(n, n+a)$ for a an odd number, since one member of the pair is then even. Similar reasoning shows that (3, 5, 7) is the only triple of primes of the form $(n, n+2, n+4)$. There are analogues of the twin prime conjecture for pairs or triples of primes that are not ruled out by congruential reasoning, such as $(n, n+4)$ or $(n, n+2, n+6)$. More generally, the *prime g-tuples conjecture* asserts that, absent any congruential obstruction, there exist infinitely many prime g-tuples $(n, n+a_1, \dots, n+a_{g-1})$ (with fixed integers $a_1, \dots, a_{g-1}$).

As a first attempt at detecting twin primes, take

$$\mathcal{A} = \{n(n+2) : 1 \le n \le X\}$$

and $\mathcal{P}$ as the set of all primes. The number of twin primes $(p, p+2)$ with $\sqrt{X+2} < p \le X$ is provided by

$$S(\mathcal{A}, \mathcal{P}, \sqrt{X+2}),$$

where $S(\mathcal{A}, \mathcal{P}, z)$ denotes the number of elements in $\mathcal{A}$ coprime with the primes of $\mathcal{P}$ that are less than z. Here, as in Example 1.1, we are not able to approximate the S expression effectively. However, it provides a framework for our investigations.

Example 1.2. Let

$$L(n) := \prod_{i=1}^{g} (a_i n + b_i),$$

where the coefficients are integers satisfying $(a_i, b_i) = 1$ $(i = 1, \dots, g)$ and the discriminant

$$\Delta = \prod_{i=1}^{g} a_i \prod_{1 \le r < s \le g} (a_r b_s - a_s b_r)$$

is non-zero. The non-vanishing of Δ ensures that the linear factors of L are not constant and that none is a linear multiple of another. Now let $\mathcal{P}$ be the set of all primes less than z and

$$\mathcal{A} = \{L(n) : x - y < n \le x\}, \qquad 1 < y \le x.$$

Here $X = y$, $\omega(d)$ is the number of incongruent solutions modulo d of the congruence $L(n) \equiv 0 \mod d$, and $|r_{\mathcal{A}}(d)| \le \omega(d)$. From elementary number theory, $\omega(p) \le g$ for all primes p, with equality when $p \nmid \Delta$.

When $p \mid \Delta$, $\omega(p)$ may take on any integer value in $[0, g)$. Let $\nu(d)$ denote the number of distinct prime divisors of d. Then, for squarefree d,

$$\omega(d) \le g^{\nu(d)}$$

with equality when $(d, \Delta) = 1$. We shall come back to this example, basic to the "prime g-tuples" conjecture, and estimate $S(\mathcal{A}, \mathcal{P})$ in several applications later. It would be a great triumph for sieve theory to show that $L(n) = P_{g+\ell}$ infinitely often for some positive integer $\ell < g$; for that would imply that one of the factors $a_i n + b_i$ is a prime!

1.4 The $\Omega(\kappa)$ condition

We introduce at this point a weak average condition on $\omega(\cdot)$ that is to hold throughout.

Definition 1.3. We say that a sieve problem satisfies *the* $\boldsymbol{\Omega(\kappa)}$ *condition* provided there exist constants $\kappa \ge 1$, $A > 1$ such that

$$\prod_{w_1 \le p < w} \Big(1 - \frac{\omega(p)}{p}\Big)^{-1} \le \Big(\frac{\log w}{\log w_1}\Big)^{\kappa} \Big(1 + \frac{A}{\log w_1}\Big), \quad 2 \le w_1 < w. \tag{1.5}$$

The parameter κ is clearly not unique—if $\boldsymbol{\Omega(\kappa)}$ holds for some number κ, then it holds for any $\kappa' > \kappa$. Nevertheless, in most sieve problems the minimal κ is known and we refer to it as the *dimension*, or *sifting density*, of the problem. Problems of dimension 1 are especially important and we refer to them as *linear*. Note that $\boldsymbol{\Omega(\kappa)}$ implies that

$$\prod_{w_1 \le p < w} \Big(1 - \frac{\omega(p)}{p}\Big)^{-1} \le \prod_{2 \le p < w} \Big(1 - \frac{\omega(p)}{p}\Big)^{-1} \ll (\log w)^{\kappa}.$$

We pause here to check that $\boldsymbol{\Omega(\kappa)}$ holds in Example 1.2 with $\kappa = g$. By adjusting the bound A if necessary, we may assume that $w_1 \ge g + 1$. Then, since $\omega(p) \le g$, we have

$$\begin{aligned} \prod_{w_1 \le p < w} \Big(1 - \frac{\omega(p)}{p}\Big)^{-1} &\le \exp\Big\{ \sum_{w_1 \le p < w} -\log\Big(1 - \frac{g}{p}\Big) \Big\} \\ &= \exp\Big\{ g \sum_{w_1 \le p < w} \frac{1}{p} + \sum_{w_1 \le p < w} \sum_{r=2}^{\infty} \frac{1}{r} \Big(\frac{g}{p}\Big)^r \Big\}. \end{aligned}$$

Thus

$$\begin{aligned}\prod_{w_1\le p<w}\Big(1-\frac{\omega(p)}{p}\Big)^{-1} &\le \exp\Big\{g\sum_{w_1\le p<w}\frac{1}{p}+\frac{1}{2}g^2\sum_{p\ge w_1}\frac{1}{p(p-g)}\Big\}\\ &\le \exp\Big\{g\sum_{w_1\le p<w}\frac{1}{p}+\frac{1}{2}g^2(g+1)\sum_{p\ge w_1}\frac{1}{p^2}\Big\}\\ &= \exp\Big\{g\log\frac{\log w}{\log w_1}+O\Big(\frac{1}{\log w_1}\Big)+O\Big(\frac{1}{w_1}\Big)\Big\}\\ &= \Big(\frac{\log w}{\log w_1}\Big)^g\exp\Big\{O\Big(\frac{1}{\log w_1}\Big)\Big\},\end{aligned}$$

which implies $\Omega(g)$; at the next to last stage we used Mertens' sum formula ([HW79], Theorems 427, 428) that

$$\sum_{w_1\le p<w}\frac{1}{p}=\log\frac{\log w}{\log w_1}+O\Big(\frac{1}{\log w_1}\Big),\quad 2\le w_1<w. \tag{1.6}$$

Note that the preceding argument shows incidentally that $\Omega(\kappa)$ holds with $\kappa=A_0$ whenever $\omega(p)\le A_0$ holds for all primes $p\in\mathcal{P}$.

As an immediate consequence of $\Omega(\kappa)$, on taking logarithms, we have

$$\sum_{w_1\le p<w}\frac{\omega(p)}{p}\le\kappa\log\Big(\frac{\log w}{\log w_1}\Big)+\frac{A}{\log w_1},\quad 2\le w_1<w. \tag{1.7}$$

Several useful variants of the last inequality follow by partial summation, and we note them here for later use.

Lemma 1.4. *Assume $\Omega(\kappa)$, and let f be a continuous nonnegative monotone function on an interval $[w_1,w]$, $w_1\ge 2$. If f is increasing on $[w_1, w]$, then*

$$\sum_{w_1\le p<w}\frac{\omega(p)}{p}f(p)\le\frac{Af(w)}{\log w}+\int_{w_1}^{w}f(t)\Big(\frac{\kappa}{\log t}+\frac{A}{\log^2 t}\Big)\frac{dt}{t}. \tag{1.8}$$

If f is decreasing on $[w_1, w]$, then

$$\sum_{w_1\le p<w}\frac{\omega(p)}{p}f(p)\le\frac{Af(w_1)}{\log w_1}+\kappa\int_{w_1}^{w}\frac{f(t)\,dt}{t\log t}. \tag{1.9}$$

Proof. We have

$$L(s,t):=\sum_{s\le p<t}\frac{\omega(p)}{p}\le\kappa\log\Big(\frac{\log t}{\log s}\Big)+\frac{A}{\log s},\quad 2\le s\le t.$$

For f increasing,

$$\begin{aligned}\sum_{w_1\le p<w}\frac{\omega(p)}{p}f(p)-\int_{w_1}^{w}f(t)\Big(\frac{\kappa\,dt}{t\log t}+\frac{A\,dt}{t\log^2 t}\Big)&\\ =-\int_{w_1}^{w}f(t)\,d\Big\{L(t,w)-\kappa\log\Big(\frac{\log w}{\log t}\Big)-\frac{A}{\log t}\Big\}&\\ =-f(t)\Big\{L(t,w)-\kappa\log\Big(\frac{\log w}{\log t}\Big)-\frac{A}{\log t}\Big\}\Big|_{w_1}^{w}&\\ +\int_{w_1}^{w}\Big\{L(t,w)-\kappa\log\Big(\frac{\log w}{\log t}\Big)-\frac{A}{\log t}\Big\}df(t)&\\ \le A\,f(w)/\log w.&\end{aligned}$$

For f decreasing,

$$\begin{aligned}\sum_{w_1\le p<w}\frac{\omega(p)}{p}f(p)-\int_{w_1}^{w}\frac{f(t)\,dt}{t\log t}&\\ =\int_{w_1}^{w}f(t)\,d\Big\{L(w_1,t)-\kappa\log\Big(\frac{\log t}{\log w_1}\Big)-\frac{A}{\log w_1}\Big\}&\\ =f(t)\Big\{L(w_1,t)-\kappa\log\Big(\frac{\log t}{\log w_1}\Big)-\frac{A}{\log w_1}\Big\}\Big|_{w_1}^{w}&\\ -\int_{w_1}^{w}\Big\{L(w_1,t)-\kappa\log\Big(\frac{\log t}{\log w_1}\Big)-\frac{A}{\log w_1}\Big\}df(t)&\\ \le Af(w_1)/\log w_1.&\end{aligned}$$

□

Corollary 1.5. *Assume* $\boldsymbol{\Omega(\kappa)}$ *and* $2\le w_1<w$. *Then*

$$\sum_{w_1\le p<w}\frac{\omega(p)}{p}\log p\le\kappa\log\frac{w}{w_1}+A\Big(1+\log\frac{\log w}{\log w_1}\Big), \tag{1.10}$$

$$\sum_{p<w}\frac{\omega(p)}{p}(p^{\epsilon}-1)\le\frac{\kappa(w^{\epsilon}-1)}{\epsilon\log w}+\frac{Aw^{\epsilon}}{\log w}+O\Big(\frac{w^{\epsilon}}{1+\epsilon^2\log^2 w}\Big),\quad 0<\epsilon\le 1, \tag{1.11}$$

$$\sum_{w_1\le p<w}\frac{\omega(p)}{p\log p}\le\frac{A}{\log^2 w_1}+\frac{\kappa}{\log w_1}-\frac{\kappa}{\log w}. \tag{1.12}$$

Proof. The first and third inequalities follow at once from the lemma. We show that the second inequality holds uniformly for $\epsilon>0$. The first term on the right side of (1.8) is bounded above by $Aw^{\epsilon}/\log w$; it

remains to estimate the integral, which is in this case

$$\mathcal{I} := \int_2^w (t^\epsilon - 1)\Big\{\frac{\kappa}{\log t} + \frac{A}{\log^2 t}\Big\}\frac{dt}{t} = \int_{\epsilon\log 2}^{\epsilon\log w} (e^v - 1)\Big(\frac{\kappa}{v} + \frac{A\epsilon}{v^2}\Big)dv$$
$$\le \kappa\int_0^{\epsilon\log w} \frac{e^v - 1}{v}\,dv + A\epsilon\int_{\epsilon\log 2}^{\epsilon\log w} \frac{e^v - 1}{v^2}dv =: \mathcal{I}_1 + \mathcal{I}_2\,, \quad \text{say}.$$

We estimate the integrals explicitly, for possible numerical applications.

$$\mathcal{I}_1 = \kappa\frac{(e^v - v - 1)}{v}\Big|_0^{\epsilon\log w} + \kappa\int_0^{\epsilon\log w} \frac{e^v - v - 1}{v^2}\,dv$$
$$= \frac{\kappa(w^\epsilon - 1 - \epsilon\log w)}{\epsilon\log w} + \int_0^{\epsilon\log w} \kappa\sum_{r=1}^{\infty}\frac{v^{r-1}}{(r+1)!}\,dv.$$

The last integral equals

$$\kappa\sum_{r=1}^{\infty}\frac{(\epsilon\log w)^r}{(r+1)!\,r} = \frac{\kappa}{(\epsilon\log w)^2}\sum_{r=1}^{\infty}\frac{(\epsilon\log w)^{r+2}}{(r+2)!}\,\frac{r+2}{r}$$
$$\le \frac{3\kappa(w^\epsilon - 1 - \epsilon\log w)}{(\epsilon\log w)^2}.$$

Thus

$$\mathcal{I}_1 \le \frac{\kappa(w^\epsilon - 1)}{\epsilon\log w} + \frac{3\kappa(w^\epsilon - 1 - \epsilon\log w)}{(\epsilon\log w)^2}.$$

In the same manner,

$$\mathcal{I}_2 \le A\epsilon\frac{(e^v - 1 - v)}{v^2}\Big|_0^{\epsilon\log w} + 2A\epsilon\int_{\epsilon\log 2}^{\epsilon\log w} \frac{e^v - 1 - v}{v^3}\,dv$$
$$\le A\epsilon\frac{(w^\epsilon - 1 - \epsilon\log w)}{(\epsilon\log w)^2} + \frac{2A}{\log 2}\int_0^{\epsilon\log w} \frac{e^v - 1 - v}{v^2}\,dv$$
$$\le \frac{A(w^\epsilon - 1 - \epsilon\log w)}{\epsilon\log^2 w} + \frac{6A(w^\epsilon - 1 - \epsilon\log w)}{(\log 2)\,\epsilon^2\log^2 w},$$

(the last by using the integral estimate from $\mathcal{I}_1$). The error term of (1.11) covers the cases of both small and large values of $\epsilon\log w$. □

1.5 Notes on Chapter 1

With minor exceptions, we use the notation introduced in [HR74].

Overviews of sieve methods, useful examples, and many problems are given in the books [HR74], [BaD04], and [MV06].

We shall treat the case $\kappa = 1$ in Chapter 7. However, our main thrust is to deal with integer or half integer dimensions that exceed 1, and we analyze that case in Chapter 9.

Bateman and Horn [BH62] conjectured that

$$|\{n \le x : \Omega(L(n)) = g\}| \sim Cx(\log x)^{-g}, \quad x \to \infty,$$

with an explicit constant C depending on the coefficients a_i and b_j. This conjecture has not been confirmed for any $g \ge 2$. Approximations take the form

$$|\{n \le x : \Omega(L(n)) \le r_g\}| \gg x(\log x)^{-g},$$

where $r_g \sim g \log g$ ([HR74], Theorem 10.5). Better values for r_g for small g are given in Table 11.1 below.

In connection with the remarks following Example 1.2 on prime g-tuples, there are the recent spectacular results of Goldston *et al.* ([GPY, GPY06, GGPY]) about gaps between primes and many related results, some conditional. These results will be the subject of a forthcoming book by those authors.

The condition $\boldsymbol{\Omega(\kappa)}$ could be weakened slightly by replacing the factor $1 + A/\log w_1$ with $\exp(A/\log w_1)$, as some authors have done. We retain the original formulation of Iwaniec.