

## Chapter 3

### Principles and First Examples of Sieve Methods

#### 1. Initiation

The aim of sieve theory is to construct estimates for the number of integers remaining in a set after members of certain arithmetic progressions have been discarded. If  $k$  is given, then the asymptotic density of the set of integers relatively prime to  $k$  is  $\varphi(k)/k$ ; with the aid of sieves we can estimate how quickly this asymptotic behaviour is approached. Throughout this chapter we let  $S(x, y; k)$  denote the numbers of integers  $n$  in the interval  $x < n \leq x + y$  for which  $(n, k) = 1$ . A first (weak) result is provided by

**Theorem 1.** (Eratosthenes–Legendre) *For any real  $x$ , and any  $y \geq 0$ ,*

$$S(x, y; k) = \frac{\varphi(k)}{k}y + O(2^{\omega(k)}).$$

Of course if  $y$  is an integral multiple of  $k$  then the above holds with no error term. Since  $2^{\omega(k)} \leq d(k) \ll k^\varepsilon$ , the main term above is larger than the error term if  $y \geq k^\varepsilon$ ; thus the reduced residues are roughly uniformly distributed in the interval  $(0, k]$ .

**Proof.** From the characteristic property (1.20) of the Möbius  $\mu$ -function, and the fact that  $d|(n, k)$  if and only if  $d|n$  and  $d|k$ , we see that

$$\begin{aligned} S(x, y; k) &= \sum_{x < n \leq x+y} \sum_{\substack{d|n \\ d|k}} \mu(d) \\ &= \sum_{d|k} \mu(d) \sum_{\substack{x < n \leq x+y \\ d|n}} 1 \\ (1) \quad &= \sum_{d|k} \mu(d) \left( \left[ \frac{x+y}{d} \right] - \left[ \frac{x}{d} \right] \right). \end{aligned}$$

Removing the square brackets, we see that this is

$$= y \sum_{d|k} \frac{\mu(d)}{d} + O\left( \sum_{d|k} |\mu(d)| \right),$$

which is the desired result.

The identity (1) can be considered to be an instance of Sylvester's principle of inclusion-exclusion, which in general asserts that if  $\mathcal{S}$  is a finite set and  $\mathcal{S}_1, \dots, \mathcal{S}_R$  are subsets of  $\mathcal{S}$ , then

$$(2) \quad \text{card} \left( \mathcal{S} \setminus \bigcup_{r=1}^R \mathcal{S}_r \right) = \text{card}(\mathcal{S}) - \Sigma_1 + \Sigma_2 - \dots + (-1)^R \Sigma_R$$

where

$$\Sigma_s = \sum_{1 \leq r_1 < \dots < r_s \leq R} \text{card} \left( \bigcap_{j=1}^s \mathcal{S}_{r_j} \right).$$

To obtain (1) we take  $\mathcal{S} = \{n \in \mathbb{Z} : x < n \leq x + y\}$ ,  $R = \omega(k)$ , we let  $p_1, \dots, p_R$  be the distinct primes dividing  $k$ , and we put  $\mathcal{S}_r = \{n : x < n \leq x + y, p_r | n\}$ . Here we see that the Möbius  $\mu$ -function has an important combinatorial significance, namely that it enables us to present the inclusion-exclusion identity in a compact manner, in arithmetic situations such as (1) above.

To prove (2) it suffices to note that if an element of  $\mathcal{S}$  is not in any of the  $\mathcal{S}_r$  then it is counted once on the right hand side, while if it is in precisely  $t > 0$  of the sets  $\mathcal{S}_r$  then it is counted  $\binom{t}{s}$  times in  $\Sigma_s$ , and hence it contributes altogether

$$\sum_{s=0}^k (-1)^s \binom{t}{s} = \sum_{s=0}^t (-1)^s \binom{t}{s} = (1 - 1)^t = 0.$$

If  $p$  is a prime then either  $p|k$  or  $(p, k) = 1$ . Hence

$$(3) \quad \pi(x + y) - \pi(x) \leq \omega(k) + S(x, y; k),$$

so that a bound for  $S(x, y; k)$  can be used to bound the number of prime numbers in an interval. In view of the main term in Theorem 1, it is reasonable to expect that it will be best to take  $k$  of the form

$$(4) \quad k = \prod_{p \leq z} p.$$

On taking  $z = \log y$ , we see immediately that

$$\pi(x + y) - \pi(x) \leq \left( e^{-C_0} + \varepsilon(y) \right) \frac{y}{\log \log y}$$

where  $\varepsilon(y) \rightarrow 0$  as  $y \rightarrow \infty$ . This bound is very weak, but has the interesting property of being uniform in  $x$ . Since the bound for the error term in Theorem 1 is very crude, we might expect that more is true, so that perhaps

$$S(x, y; k) \sim \frac{\varphi(k)}{k} y$$

even when  $z$  is fairly large. However, as we have already noted in our remarks following Theorem 2.11, this asymptotic formula fails when  $z = y^{1/2}$ .

In order to derive a sharper estimate for  $S(x, y; k)$ , we replace  $\mu(d)$  by a more general arithmetic function  $\lambda_d$  that in some sense is a truncated approximation to  $\mu(d)$ . This is reminiscent of our derivation of the Chebyshev bounds, but in fact the specific properties required of the  $\lambda_d$  are now rather different. Suppose that we seek an upper bound for  $S(x, y; k)$ . Let  $\lambda_n^+$  be a function such that

$$(5) \quad \sum_{d|n} \lambda_d^+ \geq \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Such a  $\lambda_d^+$  we call an ‘upper bound sifting function’, and by arguing as in the proof of Theorem 1 we see that

$$(6) \quad S(x, y; k) \leq \sum_{x < n \leq x+y} \sum_{\substack{d|n \\ d|k}} \lambda_d^+ = y \sum_{d|k} \lambda_d^+ / d + O\left(\sum_{d|k} |\lambda_d^+|\right).$$

This will be useful if  $\sum_{d|k} \lambda_d^+ / d$  is not much larger than  $\varphi(k)/k$ , and if  $\sum_{d|k} |\lambda_d^+|$  is much smaller than  $2^{\omega(k)}$ . Brun (1915) was the first to succeed with an argument of this kind. He took his  $\lambda_n^+$  to be of the form

$$\lambda_n^+ = \begin{cases} \mu(n) & \text{if } n \in \mathcal{D}^+, \\ 0 & \text{otherwise,} \end{cases}$$

where  $\mathcal{D}^+$  is a judiciously chosen set of integers. A sieve of this kind is called ‘combinatorial’. With Brun’s choice of  $\mathcal{D}^+$  it is easy to verify (5), and it is not hard to bound  $\sum_{d|k} |\lambda_d^+|$ , but the determination of the asymptotic size of the main term  $\sum_{d|k} \lambda_d^+ / d$  presents some technical difficulties. We do not develop a detailed account of Brun’s method, but the spirit of the approach can be appreciated by considering the following simple choice of  $\mathcal{D}^+$ : Let  $r$  be an integer at our disposal, and put

$$\mathcal{D}^+ = \{n : \omega(n) \leq 2r\}.$$

We observe that

$$\sum_{d|k} \lambda_d^+ = \sum_{j=0}^{2r} \sum_{\substack{d|k \\ \omega(d)=j}} \mu(d) = \sum_{j=0}^{2r} (-1)^j \binom{\omega(k)}{j}.$$

Then (5) follows on taking  $J = 2r$ ,  $h = \omega(k)$  in the binomial coefficient identity

$$\sum_{j=0}^J (-1)^j \binom{h}{j} = (-1)^J \binom{h-1}{J}.$$

This identity can in turn be proved by induction, or by equating coefficients in the power series identity

$$\left(\sum_{i=0}^{\infty} x^i\right) \left(\sum_{j=0}^h (-1)^j \binom{h}{j} x^j\right) = (1-x)^{h-1} = \sum_{J=0}^{h-1} (-1)^J \binom{h-1}{J} x^J.$$

Lower bounds for  $S(x, y; k)$  can be derived in a parallel manner, by introducing a lower bound sifting function  $\lambda_n^-$ . That is,  $\lambda_n^-$  is an arithmetic function such that

$$(7) \quad \sum_{d|n} \lambda_d^- \leq \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Corresponding to the upper bound (6) we have

$$(8) \quad S(x, y; k) \geq y \sum_{d|k} \lambda_d^- / d - O\left(\sum_{d|k} |\lambda_d^-|\right).$$

Unfortunately, this lower bound may be negative, in which case it is useless, since trivially  $S(x, y; k) \geq 0$ . Brun determined  $\lambda_d^-$  combinatorially by constructing a set  $\mathcal{D}^-$  similar to his  $\mathcal{D}^+$ . Indeed, an admissible set can be obtained by taking

$$\mathcal{D}^- = \{n : \omega(n) \leq 2r - 1\}.$$

By Brun's method it can be shown that

$$(9) \quad \pi(x+y) - \pi(x) \ll \frac{y}{\log y}.$$

When  $x = 0$  this is merely a weak form of the Chebyshev upper bound. The main utility of the above is that it holds uniformly in  $x$ . We shall establish a refined form of (9) in the next section (cf Corollary 4).

### 3.1. Exercises

1. (Charles Dodgson) In a very hotly fought battle, at least 70% of the combatants lost an eye, at least 75% an ear, at least 80% an arm, and at least 85% a leg. What can you say about the percentage that lost all four members?
2. (P. T. Bateman) Would you believe a market investigator who reports that of 1000 people, 816 like candy, 723 like ice cream, 645 like cake, while 562 like both candy and ice cream, 463 like both candy and cake, 470 like both ice cream and cake, while 310 like all three?
3. (Erdős (1946)) For  $x > 0$  write

$$\sum_{\substack{1 \leq n \leq x \\ (n, k) = 1}} 1 = \frac{\varphi(k)}{k} x + E_k(x).$$

(a) Show that if  $k > 1$  then

$$E_k(x) = - \sum_{d|k} \mu(d) B_1(\{x/d\})$$

where  $B_1(z) = z - 1/2$  is the first Bernoulli polynomial. Let  $E_k(x)$  be defined by this formula when  $x < 0$ .

(b) Show that if  $k > 1$  then  $E_k(x)$  is periodic with period  $k$ , that  $E_k(x)$  is an odd function (apart from values at discontinuities), and that

$$\int_0^k E_k(x) dx = 0.$$

(c) By using the result of Exercise B.10, or otherwise, show that if  $d|k$  and  $e|k$  then

$$\int_0^k B_1(\{x/d\}) B_1(\{x/e\}) dx = \frac{(d, e)^2}{12de} k.$$

(d) Show that if  $k > 1$  then

$$\int_0^k E_k(x)^2 dx = \frac{1}{12} 2^{\omega(k)} \varphi(k).$$

(e) Deduce that if  $k > 1$  then

$$\max_x |E_k(x)| \gg 2^{\omega(k)/2} \left( \frac{\varphi(k)}{k} \right)^{1/2}.$$

**4.** (Lehmer (1955); cf Vijayaraghavan (1951)) Let  $E_k(x)$  be defined as above.

(a) Show that  $|E_k(x)| \leq 2^{\omega(k)-1}$  for all  $k > 1$ .

(b) Suppose that  $k$  is composed of distinct primes  $p \equiv 3 \pmod{4}$ , and that  $\omega(k)$  is even. Show that if  $d|k$  then  $\mu(d) B_1(\{k/(4d)\}) = -1/4$ .

(c) Show that there exist infinitely many numbers  $k$  for which

$$\max_x |E_k(x)| \geq 2^{\omega(k)-2}.$$

**5.** (Behrend (1948); cf Heilbronn (1937), Rohrbach (1937), Chung (1941), van der Corput (1958)) Let  $a_1, \dots, a_J$  be positive integers, and let  $T(a_1, \dots, a_J)$  denote the asymptotic density of the set of those positive integers that are not divisible by any of the  $a_i$ .

(a) Show that  $T(a_1, \dots, a_J) = \sum_{j=0}^J (-1)^j \Sigma_j$  where

$$\Sigma_j = \sum_{1 \leq i_1 < \dots < i_j \leq J} \frac{1}{[a_{i_1}, \dots, a_{i_j}]}.$$

(b) Show that if  $a_1, \dots, a_J$  are pairwise relatively prime, then

$$T(a_1, \dots, a_J) = \prod_{j=1}^J \left(1 - \frac{1}{a_j}\right).$$

(c) Show if  $(d, v_s) = 1$  for  $1 \leq s \leq S$  then

$$\begin{aligned} T(du_1, \dots, du_R, v_1, \dots, v_S) &= \frac{1}{d} T(u_1, \dots, u_R, v_1, \dots, v_S) \\ &\quad + \left(1 - \frac{1}{d}\right) T(v_1, \dots, v_S). \end{aligned}$$

(d) Suppose that  $d|a_j$  for  $1 \leq j \leq j_0$ , that  $(d, a_j) = 1$  for  $j > j_0$ , that  $d|b_k$  for  $1 \leq k \leq k_0$ , and that  $(d, b_k) = 1$  for  $k_0 < k \leq K$ . Put  $a'_j = a_j/d$  for  $1 \leq j \leq j_0$ , and  $b'_k = b_k/d$  for  $1 \leq k \leq k_0$ . Explain why

$$\begin{aligned} &T(a_1, \dots, a_J) T(b_1, \dots, b_K) \\ &= \frac{1}{d} T(a'_1, \dots, a'_{j_0}, a_{j_0+1}, \dots, a_J) T(b'_1, \dots, b'_{k_0}, b_{k_0+1}, \dots, b_K) \\ &\quad + \left(1 - \frac{1}{d}\right) T(a_{j_0+1}, \dots, a_J) T(b_{k_0+1}, \dots, b_K) \\ &\quad - \frac{1}{d} \left(1 - \frac{1}{d}\right) \left( T(a_{j_0+1}, \dots, a_J) - T(a'_1, \dots, a'_{j_0}, a_{j_0+1}, \dots, a_J) \right) \\ &\quad \cdot \left( T(b_{k_0+1}, \dots, b_K) - T(b'_1, \dots, b'_{k_0}, b_{k_0+1}, \dots, b_K) \right) \end{aligned}$$

(e) Explain why the factors that constitute the last term above are all non-negative.

(f) Show that

$$T(a_1, \dots, a_J, b_1, \dots, b_K) \geq T(a_1, \dots, a_J) T(b_1, \dots, b_K).$$

(g) Show that

$$T(a_1, \dots, a_J) \geq \prod_{j=1}^J \left(1 - \frac{1}{a_j}\right).$$

## 2. The Selberg Lambda-square method

Let  $\Lambda_n$  be a real-valued arithmetic function such that  $\Lambda_1 = 1$ . Then

$$\left( \sum_{d|n} \Lambda_d \right)^2 \geq \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

This simple observation can be used to obtain an upper bound for  $S(x, y; k)$ ; namely

$$S(x, y; k) \leq \sum_{x < n \leq x+y} \left( \sum_{\substack{d|n \\ d|k}} \Lambda_d \right)^2$$

$$\begin{aligned}
&= \sum_{\substack{d|k \\ e|k}} \Lambda_d \Lambda_e \sum_{\substack{x < n \leq x+y \\ d|n, e|n}} 1 \\
&= \sum_{\substack{d|k \\ e|k}} \Lambda_d \Lambda_e \left( \left[ \frac{x+y}{[d, e]} \right] - \left[ \frac{x}{[d, e]} \right] \right) \\
(10) \quad &= y \sum_{\substack{d|k \\ e|k}} \frac{\Lambda_d \Lambda_e}{[d, e]} + O\left( \left( \sum_{d|k} |\Lambda_d| \right)^2 \right).
\end{aligned}$$

In the general framework of the preceding section this amounts to taking

$$\lambda_n^+ = \sum_{\substack{d, e \\ [d, e] = n}} \Lambda_d \Lambda_e,$$

since it then follows that

$$\sum_{d|n} \lambda_d^+ = \left( \sum_{d|n} \Lambda_d \right)^2.$$

We now suppose that  $\Lambda_n = 0$  for  $n > z$  where  $z$  is a parameter at our disposal, in the hope that this will restrict the size the error term. As for the main term, we see that we wish to minimize a quadratic form subject to the constraint  $\Lambda_1 = 1$ . In fact we can diagonalize this quadratic form and determine the optimal  $\Lambda_n$  exactly; this permits us to prove

**Theorem 2.** *Let  $x, y$ , and  $z$  be real numbers such that  $y > 0$  and  $z \geq 1$ . For any positive integer  $k$  we have*

$$S(x, y; k) \leq \frac{y}{L_k(z)} + O(z^2 L_k(z)^{-2})$$

where

$$L_k(z) = \sum_{\substack{n \leq z \\ n|k}} \frac{\mu(n)^2}{\varphi(n)}.$$

**Proof.** Clearly we may assume that  $k$  is squarefree. Since  $[d, e](d, e) = de$  and  $\sum_{d|n} \varphi(d) = n$ , we see that

$$\frac{1}{[d, e]} = \frac{(d, e)}{de} = \frac{1}{de} \sum_{f|d, f|e} \varphi(f).$$

Hence

$$\begin{aligned}
\sum_{d|k, e|k} \frac{\Lambda_d \Lambda_e}{[d, e]} &= \sum_{f|k} \varphi(f) \sum_{\substack{d \\ f|d|k}} \frac{\Lambda_d}{d} \sum_{\substack{e \\ f|e|k}} \frac{\Lambda_e}{e} \\
&= \sum_{f|k} \varphi(f) y_f^2
\end{aligned}$$

where

$$(11) \quad y_f = \sum_{f|d|k} \frac{\Lambda_d}{d}.$$

This linear change of variables, from  $\Lambda_d$  to  $y_f$ , is non-singular. That is, if the  $y_f$  are given then there exist unique  $\Lambda_d$  such that the above holds. Indeed, by a form of the Möbius inversion formula (cf Exercise 2.1.6) the above is equivalent to the relation

$$(12) \quad \Lambda_d = d \sum_{\substack{f \\ d|f|k}} y_f \mu(f/d).$$

Moreover, from these formulæ we see that  $\Lambda_d = 0$  for all  $d > z$  if and only if  $y_f = 0$  for all  $f > z$ . Thus we have diagonalized the quadratic form in (10), and by (12) we see that the constraint  $\Lambda_1 = 1$  is equivalent to the linear condition

$$(13) \quad \sum_{f|k} y_f \mu(f) = 1.$$

We determine the value of the constrained minimum by completing squares. If the  $y_f$  satisfy (13) then

$$(14) \quad \sum_{\substack{f|k \\ f \leq z}} \varphi(f) y_f^2 = \sum_{\substack{f|k \\ f \leq z}} \varphi(f) \left( y_f - \frac{\mu(f)}{\varphi(f) L_k(z)} \right)^2 + \frac{1}{L_k(z)}.$$

Here the right hand side is minimized by taking

$$(15) \quad y_f = \frac{\mu(f)}{\varphi(f) L_k(z)}$$

for  $f \leq z$ , and we note that these  $y_f$  satisfy (13). Hence the minimum of the quadratic form in (10), subject to  $\Lambda_1 = 1$ , is precisely  $1/L_k(z)$ ; this gives the main term.

We now treat the error term. Since  $k$  is squarefree, from (12) and (15) we see that

$$(16) \quad \Lambda_d = \frac{d}{L_k(z)} \sum_{\substack{f \\ d|f|k \\ f \leq z}} \frac{\mu(f) \mu(f/d)}{\varphi(f)} = \frac{d \mu(d)}{L_k(z) \varphi(d)} \sum_{\substack{m|k \\ (m,d)=1 \\ m \leq z/d}} \frac{\mu(m)^2}{\varphi(m)};$$

here we have put  $m = f/d$ . Thus

$$\sum_{d \leq z} |\Lambda_d| \leq \frac{1}{L_k(z)} \sum_{d \leq z} \frac{d}{\varphi(d)} \sum_{m \leq z/d} \frac{1}{\varphi(m)} = \frac{1}{L_k(z)} \sum_{m \leq z} \frac{1}{\varphi(m)} \sum_{d \leq z/m} \frac{d}{\varphi(d)}.$$

Since  $d/\varphi(d) = \sum_{r|d} \mu^2(r)/\varphi(r)$ , it follows by the method of §2.1 that

$$\sum_{d \leq y} \frac{d}{\varphi(d)} = \sum_{r \leq y} \frac{\mu^2(r)}{\varphi(r)} [y/r] \leq y \sum_r \frac{\mu^2(r)}{r\varphi(r)} \ll y.$$

On inserting this in our former estimate, we find that

$$(17) \quad \sum_{d \leq z} |\Lambda_d| \ll \frac{z}{L_k(z)} \sum_{m \leq z} \frac{1}{m\varphi(m)} \ll \frac{z}{L_k(z)}.$$

This gives the stated error term, so the proof is complete.

In order to apply Theorem 2, we require a lower bound for the sum  $L_k(z)$ . To this end we show that

$$(18) \quad \sum_{n \leq z} \frac{\mu(n)^2}{\varphi(n)} > \log z$$

for all  $z \geq 1$ . Let  $s(n)$  denote the largest squarefree number dividing  $n$  (sometimes called the ‘squarefree part of  $n$ ’). Then for squarefree  $n$ ,

$$\frac{1}{\varphi(n)} = \frac{1}{n} \prod_{p|n} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right) = \sum_{\substack{m \\ s(m)=n}} \frac{1}{m},$$

so that the sum in (18) is

$$\sum_{\substack{m \\ s(m) \leq z}} \frac{1}{m}.$$

Since  $s(m) \leq m$ , this latter sum is

$$\geq \sum_{m \leq z} \frac{1}{m} > \log z.$$

Here the last inequality is obtained by the integral test. With more work one can derive an asymptotic formula for the the sum in (18) (recall Exercise 2.1.17).

By taking  $z = y^{1/2}$  in Theorem 2, and appealing to (18), we obtain

**Theorem 3.** *Let  $k = \prod_{p \leq \sqrt{y}} p$ . Then for any  $x$  and any  $y \geq 2$ ,*

$$S(x, y; k) \leq \frac{2y}{\log y} \left(1 + O\left(\frac{1}{\log y}\right)\right).$$

By combining the above with (3) we obtain an immediate application to the distribution of prime numbers.

**Corollary 4.** For any  $x \geq 0$  and any  $y \geq 2$ ,

$$\pi(x+y) - \pi(x) \leq \frac{2y}{\log y} \left(1 + O\left(\frac{1}{\log y}\right)\right).$$

In Theorem 3 we consider only a very special sort of  $k$ , but the following lemma enables us to obtain corresponding results for more general  $k$ .

**Lemma 5.** Put  $M(y; k) = \max_x S(x, y; k)$ . If  $(k, q) = 1$  then

$$M(y; k) \leq \frac{q}{\varphi(q)} M(y; qk).$$

**Proof.** It suffices to show that

$$(19) \quad \varphi(q)S(x, y; k) = \sum_{m=1}^q S(x + km, y; qk),$$

since the right hand side is bounded above by  $qM(y; qk)$ . Suppose that  $x + km < n \leq x + km + y$  and that  $(n, qk) = 1$ . Put  $r = n - km$ . Then  $x < r \leq x + y$ ,  $(r, k) = 1$ , and  $(r + km, q) = 1$ . Thus the right hand side above is

$$\sum_m \sum_r 1 = \sum_{\substack{x < r \leq x+y \\ (r, k)=1}} \sum_{\substack{1 \leq m \leq q \\ (r+km, q)=1}} 1.$$

Since  $(k, q) = 1$ , the map  $m \mapsto r + km$  permutes the residue classes (mod  $q$ ). Hence the inner sum above is  $\varphi(q)$ , and we have (19).

**Theorem 6.** For any real  $x$  and any  $y \geq 2$ ,

$$S(x, y; k) \leq e^{C_0} y \left( \prod_{\substack{p|k \\ p \leq \sqrt{y}}} \left(1 - \frac{1}{p}\right) \right) \left(1 + O\left(\frac{1}{\log y}\right)\right).$$

**Proof.** Let

$$k_1 = \prod_{\substack{p|k \\ p \leq \sqrt{y}}} p, \quad q_1 = \prod_{\substack{p \nmid k \\ p \leq \sqrt{y}}} p.$$

Theorem 3 provides an upper bound for  $M(y; q_1 k_1)$ , and hence by Lemma 5 we have an upper bound for  $M(y; k_1)$ . To complete the argument it suffices to note that  $S(x, y; k) \leq S(x, y; k_1) \leq M(y; k_1)$ , and to appeal to Mertens' formula (Theorem 2.7(e)).

We note that Theorem 3 is a special case of Theorem 6. Although we have taken great care to derive uniform estimates, for many purposes it is enough to know that

$$(20) \quad S(x, y; k) \ll y \prod_{\substack{p|k \\ p \leq y}} \left(1 - \frac{1}{p}\right).$$

This follows from Theorem 6 since  $\prod_{\sqrt{y} < p \leq y} (1 - 1/p)^{-1} \ll 1$  by Mertens' formula. To obtain an estimate in the opposite direction, write  $k = k_1 q_1$  where  $k_1$  is composed entirely of primes  $> y$ , and  $q_1$  is composed entirely of primes  $\leq y$ . Since the integers in the interval  $(0, y]$  have no prime factor  $> y$ , we see that  $M(y; k_1) \geq [y]$ . Hence by Lemma 5,

$$(21) \quad M(y; k) \geq [y] \prod_{\substack{p|k \\ p \leq y}} \left(1 - \frac{1}{p}\right).$$

Thus the bound (20) is of the correct order of magnitude.

The advantage of Theorem 6 lies in its uniformity. On the other hand, the use of Lemma 5 is wasteful if the  $k$  in Theorem 6 is much smaller than in Theorem 3. For example, if  $k = \prod_{p \leq y^{1/4}} p$  then by Theorem 6 we find that

$$S(x, y; k) \leq \frac{cy}{\log y} \left(1 + O\left(\frac{1}{\log y}\right)\right)$$

with  $c = 4$ , whereas by Theorem 2 with  $z = y^{1/2}$  we obtain the above with the better constant  $c = \frac{4}{3-2\log 2} = 2.4787668\dots$ . To see this, we note that

$$(22) \quad L_k(z) = \sum_{n \leq z} \frac{\mu(n)^2}{\varphi(n)} - \sum_{z^{1/2} < p \leq z} \frac{1}{p-1} \sum_{n \leq z/p} \frac{\mu(n)^2}{\varphi(n)}.$$

Then by Exercise 2.1.17 and Mertens' estimates (Theorem 2.7) it follows that this is  $\frac{1}{4}(3 - 2\log 2)\log y + O(1)$ .

### 3.2. Exercises

1. Let  $\Lambda_d$  be defined as in the proof of Theorem 2.

(a) Show that

$$\Lambda_d \ll \frac{d}{L_k(z)\varphi(d)} \log \frac{2z}{d}$$

for  $d \leq z$ .

(b) Use the above to give a second proof of (17).

2. Show that for  $y \geq 2$  the number of prime powers  $p^k$  in the interval  $(x, x + y]$  is

$$\leq \frac{2y}{\log y} \left(1 + O\left(\frac{1}{\log y}\right)\right).$$

3. (Chowla (1932)) Let  $f(n)$  be an arithmetic function, put

$$g(n) = \sum_{[d,e]=n} f(d)\overline{f(e)},$$

and let  $\sigma_c$  denote the abscissa of convergence of the Dirichlet series  $\sum g(n)n^{-s}$ .

(a) Show that if  $\sigma > \max(1, \sigma_c)$  then

$$\zeta(s) \sum_{d, e} \frac{f(d)\overline{f(e)}}{[d, e]^s} = \sum_{n=1}^{\infty} \left| \sum_{d|n} f(d) \right|^2 n^{-s}.$$

(b) Show that

$$\sum_{d, e} \frac{\mu(d)\mu(e)}{[d, e]^2} = \frac{6}{\pi^2}.$$

(c) Show that

$$\sum_{\substack{d, e \\ [d, e]=n}} \mu(d)\mu(e) = \mu(n)$$

for all positive integers  $n$ .

**4.** Let  $f(n)$  be an arithmetic function such that  $f(1) = 1$ . Show that  $f$  is multiplicative if and only if  $f(m)f(n) = f((m, n))f([m, n])$  for all pairs of positive integers  $m, n$ .

**5.** (Hensley (1978)) (a) Let  $k = \prod_{p \leq \sqrt{y}} p$ . Show that the number of  $n$ ,  $x < n \leq x + y$ , such that  $\Omega(n) = 2$ , is

$$\leq S(x, y; k) + \sum_{p \leq \sqrt{y}} \left( \pi\left(\frac{x+y}{p}\right) - \pi\left(\frac{x}{p}\right) \right).$$

(b) By using Theorem 3 and Corollary 4, show that for  $y \geq 2$ ,

$$\sum_{\substack{x < n \leq x+y \\ \Omega(n)=2}} 1 \leq \frac{2y \log \log y}{\log y} \left( 1 + O\left(\frac{1}{\log \log y}\right) \right).$$

**6.** (Brun (1919); cf Riesel & Vaughan (1983)) (a) Let  $g(n)$  be the multiplicative function such that  $g(2) = 1$ ,  $g(4) = 2$ ,  $g(2^a) = 4$  for  $a \geq 3$ , and  $g(p^a) = 2$  when  $p > 2$ . Show that

$$\sum_{\substack{x < n \leq x+y \\ d|n(n+2)}} 1 = g(d) \left( \frac{y}{d} + O(1) \right).$$

(b) Suppose that  $\Lambda_1 = 1$ , that  $\Lambda_n \in \mathbb{R}$ , that  $\Lambda_n = 0$  for  $n > z$ , and that  $\Lambda_n = 0$  if  $n \nmid k$ , where  $k$  is some given integer. Show that

$$\sum_{\substack{x < n \leq x+y \\ (n(n+2), k)=1}} 1 \leq y \sum_{d, e} \frac{\Lambda_d \Lambda_e}{[d, e]} g([d, e]) + O\left( \left( \sum_d |\Lambda_d| g(d) \right)^2 \right).$$

(c) Show that  $m/g(m) = \sum_{f|m} h(f)$  where

$$h(f) = \prod_{p^t || f} \left( \frac{p^t}{g(p^t)} - \frac{p^{t-1}}{g(p^{t-1})} \right).$$

Deduce that the sum in the main term in (b) is

$$\sum_{f|k} h(f) \left( \sum_{\substack{d \\ f|d}} \frac{g(d)\Lambda_d}{d} \right)^2.$$

(d) Show that the minimum of the above sum, subject to  $\Lambda_1 = 1$ , is  $1/L$  where  $L = \sum_{f|k, f \leq z} \mu(f)^2/h(f)$ , and that this is attained when

$$\Lambda_d = \frac{\mu(d)}{L} \left( \prod_{p|d} \left( \frac{p}{p-g(p)} \right) \right) \sum_{\substack{m|k \\ m \leq z/d \\ (m,d)=1}} \frac{\mu(m)^2}{h(m)}$$

for  $d|k$ ,  $d \leq z$ .

(e) Let  $j(n)$  be the unique multiplicative function such that

$$\frac{\mu(n)^2}{h(n)} = \frac{1}{n} \sum_{d|n} d(d)j(n/d).$$

Show that  $\sum_n |j(n)|/n^\sigma < \infty$  for  $\sigma > 1/2$ , that  $\sum_{n>u} |j(n)|/n \ll u^{-1/2+\varepsilon}$ , and that

$$\sum_{n=1}^{\infty} \frac{j(n)}{n} = \frac{1}{2} \prod_{p>2} \left( 1 + \frac{1}{p(p-2)} \right) = \frac{1}{c},$$

say.

(f) Show that  $\sum_{n \leq r} d(n)/n = \frac{1}{2}(\log r)^2 + O(\log r)$  for  $r \geq 2$ .

(g) Show that

$$\sum_{n \leq z} \frac{\mu(n)^2}{h(n)} = \frac{1}{2c}(\log z)^2 + O(\log z)$$

for  $z \geq 2$ . (For a sharper estimate of this sum, see Riesel & Vaughan (1983).)

(h) Show that the extremal  $\Lambda_d$  satisfy

$$\Lambda_d \ll \left( \frac{d}{\varphi(d)} \right)^2 \left( \frac{\log 2z/d}{\log 2z} \right)^2.$$

Deduce that  $\sum_{d \leq z} |\Lambda_d|g(d) \ll z/\log 2z$  for  $z \geq 1$ .

(i) Suppose that  $k = \prod_{p \leq \sqrt{y}} p$ . By taking  $z = y^{1/2}/\log y$ , show that

$$\sum_{\substack{x < n \leq x+y \\ (n(n-2), k)=1}} 1 \leq \frac{8cy}{(\log y)^2} \left( 1 + O\left( \frac{\log \log y}{\log y} \right) \right)$$

for  $y \geq 2$ .

(j) Conclude that the number of primes  $p$ ,  $x < p \leq x + y$ , such that  $p + 2$  is prime, satisfies the above bound. Note that  $c$  can be written  $c = 2 \prod_{p>2} (1 - (p-1)^{-2})$ . Hardy and Littlewood (1922) conjectured that the number of primes  $p \leq x$  such that  $p + 2$  is prime is  $\sim cx(\log x)^{-2}$ .

(k) Show that the sum of the reciprocals of the twin primes converges.

7. (H.-E. Richert, unpublished) (a) Show that

$$\sum_{x < n \leq x+y} \left( \sum_{d^2 | n} \Lambda_d \right)^2 = y \sum_{d, e} \frac{\Lambda_d \Lambda_e}{[d, e]^2} + O\left( \left( \sum_d |\Lambda_d| \right)^2 \right).$$

(b) Let  $f(n) = n^2 \prod_{p|n} (1 - p^{-2})$ . Show that  $\sum_{d|n} f(d) = n^2$ .

(c) For  $1 \leq d \leq z$  let  $\Lambda_d$  be real numbers such that  $\Lambda_1 = 1$ . Show that the minimum of  $\sum_{d, e} \Lambda_d \Lambda_e / [d, e]^2$  is  $1/L$  where  $L = \sum_{n \leq z} \mu(n)^2 / f(n)$ . Show also that  $\Lambda_d \ll 1$  for the extremal  $\Lambda_d$ .

(d) Show that  $\zeta(2) - 1/z \leq L \leq \zeta(2)$ .

(e) Let  $Q(x)$  denote the number of squarefree numbers not exceeding  $x$ . Show that for  $x \geq 0$ ,  $y \geq 1$ ,

$$Q(x+y) - Q(x) \leq \frac{y}{\zeta(2)} + O(y^{2/3}).$$

8. Let  $m(y; k) = \min_x S(x, y; k)$ . Show that if  $(q, k) = 1$  then

$$m(y; k) \geq \frac{q}{\varphi(q)} m(y; qk).$$

9. (N. G. de Bruijn, unpublished; cf van Lint & Richert (1964)) Let  $\mathcal{M}$  be an arbitrary set of natural numbers, and let  $s(n)$  denote the largest squarefree divisor of  $n$ . Show that

$$0 \leq \sum_{\substack{n \leq x \\ n \in \mathcal{M}}} \frac{\mu(n)^2}{\varphi(n)} - \sum_{\substack{n \leq x \\ s(n) \in \mathcal{M}}} \frac{1}{n} \leq \sum_{n \leq x} \frac{\mu(n)^2}{\varphi(n)} - \sum_{n \leq x} \frac{1}{n} \ll 1.$$

10. (van Lint & Richert (1965)) (a) Show that

$$\sum_{n \leq z} \frac{\mu(n)^2}{\varphi(n)} \leq \left( \sum_{d|q} \frac{\mu(d)^2}{\varphi(d)} \right) \left( \sum_{\substack{m \leq z \\ (m, q)=1}} \frac{\mu(m)^2}{\varphi(m)} \right).$$

(b) Deduce that

$$\sum_{\substack{n \leq z \\ (n, q)=1}} \frac{\mu(n)^2}{\varphi(n)} \geq \frac{\varphi(q)}{q} \sum_{n \leq z} \frac{\mu(n)^2}{\varphi(n)}.$$

**11.** (Hooley (1972), Montgomery & Vaughan (1979)) (a) Let  $\lambda_d^+$  be an upper bound sifting function such that  $\lambda_d^+ = 0$  for all  $d > z$ . Show that for any  $q$ ,

$$0 \leq \frac{\varphi(q)}{q} \sum_{\substack{d \\ (d,q)=1}} \frac{\lambda_d^+}{d} \leq \sum_d \frac{\lambda_d^+}{d}.$$

(Hint: Multiply both sides by  $k/\varphi(k) = \sum 1/m$  where  $m$  runs over all integers composed of the primes dividing  $k$ , and  $k = \prod_{p \leq z} p$ .)

(b) Let  $\Lambda_d$  be real with  $\Lambda_d = 0$  for  $d > z$ . Show that for any  $q$ ,

$$0 \leq \frac{\varphi(q)}{q} \sum_{\substack{d,e \\ (de,q)=1}} \frac{\Lambda_d \Lambda_e}{[d,e]} \leq \sum_{d,e} \frac{\Lambda_d \Lambda_e}{[d,e]}.$$

(c) Let  $\lambda_d^-$  be a lower bound sifting function such that  $\lambda_d^- = 0$  for  $d > z$ . Show that for any  $q$ ,

$$\frac{\varphi(q)}{q} \sum_{\substack{d \\ (d,q)=1}} \frac{\lambda_d^-}{d} \geq \sum_d \frac{\lambda_d^-}{d}.$$

**12.** For each prime  $p$  let  $\mathcal{B}(p)$  be the union of  $b(p)$  ‘bad’ arithmetic progressions with common difference  $p$ , and suppose that  $b(p) = 0$  for all sufficiently large  $p$ . Put  $\mathcal{B} = \bigcup_p \mathcal{B}(p)$ , let

$$S(N; \mathcal{B}) = \text{card}\{n : 1 \leq n \leq x, \quad n \notin \mathcal{B}\},$$

and set

$$m(N; \mathbf{b}) = \min_{\mathcal{B}} S(N; \mathcal{B}),$$

$$M(N; \mathbf{b}) = \max_{\mathcal{B}} S(N; \mathcal{B})$$

where the extrema are over all choices of the  $\mathcal{B}(p)$  with  $b(p)$  fixed. Show that if  $\mathbf{b}_1(p) \leq \mathbf{b}_2(p)$  for all  $p$  then

$$M(N; \mathbf{b}_1) \prod_p \left(1 - \frac{b_1(p)}{p}\right)^{-1} \leq M(N; \mathbf{b}_2) \prod_p \left(1 - \frac{b_2(p)}{p}\right)^{-1},$$

and similarly for  $m(N; \mathbf{b})$  with the inequality reversed.

**13.** In the above notation, suppose that  $b(p) \leq 2$  for all  $p$ . Deduce from the above and Exercise 6(i) that

$$S(N; \mathcal{B}) \leq \left( \prod_{p \leq \sqrt{N}} \left(1 - \frac{b(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-2} \right) \frac{8N}{(\log N)^2} \left(1 + O\left(\frac{\log \log N}{\log N}\right)\right).$$

14. Show that the number of primes  $p$ ,  $x < p \leq x + y$ , such that  $p + 2k$  is also prime, is

$$\leq \left( \prod_{\substack{p|k \\ p>2}} \frac{p-1}{p-2} \right) \frac{8cy}{(\log y)^2} \left( 1 + O\left( \frac{\log \log y}{\log y} \right) \right)$$

uniformly for  $k > 0$ ,  $x \geq x$ ,  $y \geq 4$ . Here  $c$  is defined as in Exercise 6.

15. Show that the number of primes  $p \leq 2n$  such that  $2n - p$  is prime is

$$\leq 8c \left( \prod_{\substack{p|n \\ p>2}} \frac{p-1}{p-2} \right) \frac{2n}{(\log 2n)^2} \left( 1 + O\left( \frac{\log \log 4n}{\log 2n} \right) \right).$$

16. Show that

$$\sum_{n \leq x} \prod_{\substack{p|n \\ p>2}} \left( \frac{p-1}{p-2} \right) = \frac{2x}{c} + O(\log x)$$

where  $c$  is defined as in Exercise 6.

### 3. Sifting an arithmetic progression

Thus far we have sifted only the zero residue class from a set of consecutive integers. We now widen the situation slightly.

**Lemma 7.** *Let  $k$  be a positive integer, and for each prime  $p$  dividing  $k$  suppose that one particular residue class  $a_p$  has been chosen. Let  $S'(x, y; k)$  denote the number of integers  $m$ ,  $x < m \leq x + y$ , such that for each  $p|k$ ,  $m \not\equiv a_p \pmod{p}$ . Then*

$$\max_x S'(x, y; k) = \max_x S(x, y; k).$$

Since  $S'(x, y; k)$  reduces to  $S(x, y; k)$  when we take  $a_p = 0$  for all  $p|k$ , we see that there is no loss of generality in sifting only the zero residue class, when the initial set of numbers consists of consecutive integers. Also, we note that the value of the maximum taken above is independent of the choice of the  $a_p$ .

**Proof.** By the Chinese Remainder Theorem there is a number  $c$  such that  $c \equiv a_p \pmod{p}$  for every  $p|k$ . Put  $n = m - c$ . Thus the inequality  $x < m \leq x + y$  is equivalent to  $x - c < n \leq x - c + y$ , and the condition that  $p|k$  implies  $m \not\equiv a_p \pmod{p}$  is equivalent to  $(n, k) = 1$ . Hence  $S'(x, y; k) = S(x - c, y; k)$ , so that

$$\max_x S'(x, y; k) = \max_x S(x - c, y; k) = \max_x S(x, y; k),$$

and the proof is complete.

**Theorem 8.** *Suppose that  $(a, q) = 1$ , that  $(k, q) = 1$ , and that  $x$  and  $y$  are real numbers with  $y \geq 2q$ . The number of  $n$ ,  $x < n \leq x + y$ , such that  $n \equiv a \pmod{q}$  and  $(n, k) = 1$  is*

$$\leq e^{C_0} \frac{y}{q} \left( \prod_{\substack{p|k \\ p \leq \sqrt{y/q}}} \left(1 - \frac{1}{p}\right) \right) \left(1 + O\left(\frac{1}{\log y/q}\right)\right).$$

**Proof.** Write  $n = mq + a$ , so that  $x' < m \leq x' + y'$  where  $x' = (x - a)/q$  and  $y' = y/q$ . For each  $p|k$  let  $a_p$  be the unique residue class  $\pmod{p}$  such that  $a_p q + a \equiv 0 \pmod{p}$ . Thus  $p|n$  if and only if  $m \equiv a_p \pmod{p}$ . Hence the number of  $n$  in question is  $S'(x', y'; k)$ , in the language of Lemma 7. The stated bound now follows from this lemma and Theorem 6.

Using the estimate above, we generalize Corollary 4 to arithmetic progressions.

**Theorem 9.** (Brun–Titchmarsh) *Let  $a$  and  $q$  be integers with  $(a, q) = 1$ , and let  $x$  and  $y$  be real numbers with  $x \geq 0$  and  $y \geq 2q$ . Then*

$$(23) \quad \pi(x + y; q, a) - \pi(x; q, a) \leq \frac{2y}{\varphi(q) \log y/q} \left(1 + O\left(\frac{1}{\log y/q}\right)\right).$$

**Proof.** Take  $k$  to be the product of those primes  $p \leq \sqrt{y/q}$  such that  $p \nmid q$ . Then

$$\begin{aligned} \prod_{p|k} \left(1 - \frac{1}{p}\right) &= \prod_{\substack{p|q \\ p \leq \sqrt{y/q}}} \left(1 - \frac{1}{p}\right)^{-1} \prod_{p \leq \sqrt{y/q}} \left(1 - \frac{1}{p}\right) \\ &\leq \prod_{p|q} \left(1 - \frac{1}{p}\right)^{-1} \prod_{p \leq \sqrt{y/q}} \left(1 - \frac{1}{p}\right). \end{aligned}$$

By Mertens' estimate this is

$$= \frac{q}{\varphi(q)} \cdot \frac{2e^{-C_0}}{\log y/q} \left(1 + O\left(\frac{1}{\log y/q}\right)\right).$$

Thus by Theorem 8, the number of primes  $p$ ,  $x < p \leq x + y$ , such that  $p \equiv a \pmod{q}$  and  $(p, k) = 1$  satisfies the bound (23). To complete the proof it remains to note that the number of primes  $p$ ,  $x < p \leq x + y$ , such that  $p \equiv a \pmod{q}$  and  $p|k$  is at most  $\omega(k) \leq \sqrt{y/q}$ , which can be absorbed in the error term in (23).

### 3. Notes

The modern era of sieve methods began with the work of Brun (1915, 1919). The  $\Lambda^2$  method of Selberg (1947) provides only upper bounds, but lower bounds can also be derived from it by using ideas of Buchstab (1938).

In contrast to the elegance of the Selberg  $\Lambda^2$  method, the further study of sieves leads us to construct asymptotic estimates for complicated sums over integers whose prime factors are distributed in certain ways. In this connection, the argument (22) is a simple foretaste of more complicated things to come. Hence further discussion of sieves is possible only after the appropriate technical tools are in place.

In this chapter we have applied the sieve only to arithmetic progressions, but it can be shown that the sieve is applicable to much more general sets. This makes sieves very versatile, but it also means that they are subject to certain unfortunate limitations. In order to estimate the number of elements of a set  $\mathcal{S}$  that remain after sifting, it suffices to have a reasonably precise estimate of the number  $X_d$  of multiples of  $d$  in the set, say of the form  $X_d = f(d)X/d + O(R_d)$  where  $X$  is an estimate for the cardinality of  $\mathcal{S}$ , and  $f$  is a multiplicative function. Thus Theorem 3 can be generalized to much more general sets, and in that more general setting it is known that the constant 2 is best-possible. It may be true that the constant 2 can be improved in the special case that one is sieving an interval, but this has not been achieved thus far.

When sifting an interval, the error terms can be avoided by using Fourier analysis as in Selberg (1991, §§19–22), or by using the large sieve as in Montgomery & Vaughan (1973). In particular, the number of integers in  $[M + 1, M + N]$  remaining after sifting is at most  $N/L$  where

$$(24) \quad L = \sum_{q \leq Q} \frac{\mu(q)^2}{1 + \frac{3}{2}qQ/N} \prod_{p|q} \frac{h(p)}{p - h(p)}.$$

Here  $h(p)$  is the number of residue classes modulo  $p$  that are deleted. This is both a generalization and a sharpening of Theorem 2.

Suppose that  $d_1, \dots, d_k$  are distinct integers, and let  $h(p)$  denote the number of distinct residue classes modulo  $p$  found among the  $d_i$ . The prime  $k$ -tuple conjecture asserts that if  $h(p) < p$  for every prime number  $p$ , then there exist infinitely many positive integers  $n$  such that the  $k$  numbers  $n + d_i$  are all prime. Hardy & Littlewood (1922) put this in a quantitative form. They conjectured that if  $h(p) < p$  for all  $p$ , then the number of  $n \leq N$  for which the  $k$  numbers  $n + d_i$  are all prime is

$$(25) \quad = (\mathfrak{S}(\mathbf{d}) + o(1)) \frac{N}{(\log N)^k}$$

as  $N \rightarrow \infty$  where

$$(26) \quad \mathfrak{S}(\mathbf{d}) = \prod_p \left(1 - \frac{h(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-k}.$$

This product is absolutely convergent, since  $h(p) = k$  for all sufficiently large primes  $p$ . Although this remains unproved, by sieving we can obtain an upper bound of the expected

order of magnitude. In particular, from (24) it can be shown that the number of  $n$ ,  $M + 1 \leq n \leq M + N$ , for which the numbers  $n + d_i$  are all prime is

$$(27) \qquad \leq (2^k k! \mathfrak{S}(\mathbf{d}) + o(1)) \frac{N}{(\log N)^k}.$$

Corollary 3 and Exercise 3.2.6(j) are special cases of this.

Titchmarsh (1930) used Brun's method to obtain Theorem 9, but with a larger constant instead of 2. Montgomery & Vaughan (1973) showed that Corollary 4 and Theorem 9 are still valid when the error terms are omitted. See also Selberg (1991, §22). The first significant improvement of Theorem 9 was obtained by Motohashi (1973). Other improvements of various kinds have been derived by Motohashi (1974), Hooley (1972, 1975), Goldfeld (1975), Iwaniec (1982), and Friedlander & Iwaniec (1997).

In Lemma 5 and in Exercises 3.2.8, 3.2.10, 3.2.11, 3.2.12 we see evidence of a monotonicity principle that permeates sieve theory; cf Selberg (1991, pp. 72–73).

Hooley (1996) has shown that quite sharp sieve bounds can be derived using the interrupted inclusion-exclusion idea that Brun started with. This approach has been developed further by Ford & Halberstam (2000). An exposition of sieves based on these ideas is given by Diamond & Halberstam (2004, Chapters 12, 13). Still more extensive accounts of sieve methods have been given by Greaves (2001), Halberstam & Richert (1974), Iwaniec & Kowalski (2004, Chapter 6), Motohashi (1983), and Selberg (1971, 1991). In addition, a collection of applications of sieves to arithmetic problems has been given by Hooley (1976), and additional sieve ideas are found in Bombieri (1977), Bombieri, Friedlander & Iwaniec (1986, 1987, 1989), Fouvry & Iwaniec (1997), Friedlander & Iwaniec (1998a, 1998b), and Iwaniec (1978, 1980a, 1980b, 1981).

### 3. Literature

- N. C. Ankeny & H. Onishi (1964/1965). *The general sieve*, Acta Arith. **10**, 31–62.
- P. T. Bateman & H. Diamond (2004). *Analytic number theory*, WorldScientific
- F. A. Behrend (1948). *Generalization of an inequality of Heilbronn and Rohrbach*, Bull. Amer. Math. Soc. **54**, 681–684.
- E. Bombieri (1977). *The asymptotic sieve*, Rend. Accad. Naz. XL (5) **1/2** (1975/76), 243–269.
- E. Bombieri, J. B. Friedlander & H. Iwaniec (1986). *Primes in arithmetic progressions to large moduli*, Acta Math. **156**, 203–251.
- (1987). *Primes in arithmetic progressions to large moduli, II*, Math. Ann. **277**, 361–393.
- (1989). *Primes in arithmetic progressions to large moduli, III*, J. Amer. Math. Soc. **2**, 215–224.
- V. Brun (1915). *Über das Goldbachsche Gesetz und die Anzahl der Primzahlpaare*, Archiv for Math. og Naturvid. B **34**, no. 8, 19 pp.

- (1919). *La série  $1/5 + 1/7 + 1/11 + 1/13 + 1/17 + 1/19 + 1/29 + 1/31 + 1/41 + 1/43 + 1/59 + 1/61 + \dots$  où les dénominateurs sont “nombres premiers jumeaux” est convergente ou finie*, Bull. Sci. Math. (2) **43**, 100–104; 124–128.
- (1967). *Reflections on the sieve of Eratosthenes*, Norske Vid. Selsk. Skr. Trondheim, no. 1, 9 pp.
- A. A. Buchstab (1938). *New improvements in the method of the sieve of Eratosthenes*, Math. Sbornik (N. S.) **4** (46), 375–387.
- S. Chowla (1932). *Contributions to the analytic theory of numbers*, Math. Z. **35**, 279–299.
- K.-L. Chung (1941). *A generalization of an inequality in the elementary theory of numbers*, J. Reine Angew. Math. **183**, 193–196.
- J. G. van der Corput (1958). *Inequalities involving least common multiple and other arithmetical functions*, Nederl. Akad. Wetensch. Proc. Ser. A **61** (= Indag. Math. **20**), 5–15.
- P. Erdős (1946). *On the coefficients of the cyclotomic polynomial*, Bull. Amer. Math. Soc. **52**, 179–184.
- K. Ford & H. Halberstam (2000). *The Brun–Hooley sieve*, J. Number Theory **81**, 335–350.
- E. Fouvry & H. Iwaniec (1997). *Gaussian primes*, Acta Arith. **79** (1997), 249–287.
- J. B. Friedlander & H. Iwaniec (1997). *The Brun–Titchmarsh theorem*, Analytic number theory (Kyoto, 1996), London Math. Soc. Lecture Note Ser. **247**, Cambridge Univ. Press (Cambridge), pp. 85–93.
- (1998a). *The polynomial  $X^2 + Y^4$  captures its primes*, Ann. of Math. (2) **148**, 945–1040.
- (1998b). *Asymptotic sieve for primes*, Ann. of Math. (2) **148**, 1041–1065.
- D. M. Goldfeld (1975). *A further improvement of the Brun–Titchmarsh theorem*, J. London Math. Soc. (2) **11**, 434–444.
- G. Greaves (2001). *Sieves in number theory*, Springer (Berlin), xii+304 pp.
- H. Halberstam (1985). *Lectures on the linear sieve*, Topics in analytic number theory (Austin, 1982), Univ. Texas Press (Austin), pp. 165–220.
- H. Halberstam & H.-E. Richert (1973). *Brun’s method and the fundamental lemma*, Acta Arith. **24**, 113–133.
- (1974). *Sieve Methods*, Academic Press (London), xiv+364.
- (1975). *Brun’s method and the fundamental lemma. II*, Acta Arith. **27**, 51–59.
- G. H. Hardy & J. E. Littlewood (1922). *Some problems of ‘Partitio Numerorum’: III. On the expression of a number as a sum of primes*, Acta Math. **44**, 1–70; *Collected Papers*, Vol. I, Oxford Univ. Press (London), 1966, pp. 561–630.
- H. Heilbronn (1937). *On an inequality in the elementary theory of numbers*, Proc. Cambridge Philos. Soc. **33**, 207–209.
- D. Hensley (1978). *An almost-prime sieve*, J. Number Theory **10**, 250–262; *Corrigendum*, **12**, (1980), 437.
- C. Hooley (1972). *On the Brun–Titchmarsh theorem*, J. Reine Angew. Math. **255**, 60–79.

- (1975). *On the Brun–Titchmarsh theorem, II*, Proc. London Math. Soc. (3) **30**, 114–128.
- (1976). *Applications of sieve methods to the theory of prime numbers*, Cambridge Tract **70**, Cambridge University Press (Cambridge), xiv+122 pp.
- (1994). *An almost pure sieve*, Acta Arith. **66**, 359–368.
- H. Iwaniec (1978). *Almost-primes represented by quadratic polynomials*, Invent. Math. **47**, 171–188.
- (1980a). *Rosser’s sieve*, Acta arith. **36**, 171–202.
- (1980b). *A new form of the error term in the linear sieve*, Acta Arith. **37**, 307–320.
- (1981). *Recent progress in analytic number theory*, Vol. 1, Academic Press (), pp. 203–230.
- (1982). *On the Brun–Titchmarsh theorem*, J. Math. Soc. Japan **34**, 95–123.
- H. Iwaniec & E. Kowalski (2004). *Analytic number theory*, Colloquium Publications 53, Amer. Math. Soc. (Providence), xii+605 pp.
- W. B. Jurkat & H.-E. Richert (1965). *An improvement in Selberg’s sieve method, I*, Acta Arith. **11**, 217–240.
- D. H. Lehmer (1955). *The distribution of totatives*, Canad. J. Math. **7**, 347–357.
- J. H. van Lint & H.-E. Richert (1964). *Über die Summe  $\sum_{\substack{n \leq x \\ p(n) < y}} \frac{\mu^2(n)}{\varphi(n)}$*  Nederl. Akad. Wetensch. Proc. Ser. A **67** (= Indag. Math. **26**), 582–587.
- (1965). *On primes in arithmetic progressions*, Acta Arith. **11**, 209–216.
- H. L. Montgomery (1968). *A note on the large sieve*, J. London Math. Soc. **43**, 93–98.
- H. L. Montgomery & R. C. Vaughan (1973). *The large sieve*, Mathematika **20**, 119–134.
- (1979). *Mean values of character sums*, Canad. J. Math. **31**, 476–487.
- Y. Motohashi (1973). *On some improvements of the Brun–Titchmarsh theorem, II*, Research of analytic number theory (Proc. Sympos., Res. Inst. Math. Sci., Kyoto, 1973), Sōurikaiseikikenkyōusho Kōkyōuroku, No. 193, 97–109.
- (1974). *On some improvements of the Brun–Titchmarsh theorem*, J. Math. Soc. Japan **26**, 306–323.
- (1975). *On some improvements of the Brun–Titchmarsh theorem, III*, J. Math. Soc. Japan **27**, 444–453.
- (1983). *Lectures on sieve methods and prime number theory*, Tata Institute of Fundamental Research (Bombay), Springer–Verlag (Berlin), xiii+205.
- H. Riesel & R. C. Vaughan (1983). *On sums of primes*, Ark. Mat. **21**, 46–74.
- H. Rohrbach (1937). *Beweis einer zahlentheoretischen Ungleichung*, J. Reine Angew. Math. **177**, 193–196.
- A. Selberg (1947). *On an elementary method in the theory of primes*, Norske Vid. Selsk. Forh. Trondhjem **19**, no. 18, 64–67.

——— (1952). *On elementary methods in primenumber-theory and their limitations*, Den 11te Skandinaviske Matematikerkongress (Trondheim, 1949), Johan Grundt Tanums Forlag (Oslo), pp. 13–22.

——— (1952). *The general sieve-method and its place in prime-number theory*, Proceedings of the International Congress of Mathematicians (Cambridge MA, 1950), Vol. 1, Amer. Math. Soc. (Providence), pp. 286–292.; *Selected Papers*, Vol. I, Springer–Verlag (Berlin), 1989, pp. 411–417.

——— (1971). *Sieve methods*, Proc. Sympos. Pure Math (SUNY Stony Brook, 1969), Vol. XX, Amer. Math. Soc. (Providence), 311–351.; *Selected Papers*, Vol. I, Springer–Verlag (Berlin), 1989, pp. 568–608.

——— (1972). *Remarks on sieves*, Proceedings of the Number Theory Conference (Boulder CO Aug. 14–18), pp. 205–216.; *Selected Papers*, Vol. I, Springer–Verlag (Berlin), 1989, pp. 609–615.

——— (1989). *Sifting problems, sifting density and sieves*, Number theory, trace formulas, and discrete groups, (Oslo, 1987), K. E. Aubert, E. Bombieri, D. Goldfeld, eds., Academic Press (Boston), pp. 467–484; *Selected Papers*, Vol. I, Springer–Verlag (Berlin), 1989, pp. 675–694.

——— (1991). *Lectures on sieves*, in *Selected Papers*, Vol. II, Springer–Verlag (Berlin), pp. 65–247.

E. C. Titchmarsh (1930). *A divisor problem*, Rend. Circ. Math. Palermo **54**, 414–429.

K. M. Tsang (1989). *Remarks on the sieving limit of the Buchstab–Rosser sieve*, Number theory, trace formulas and discrete groups (Oslo, 1987), Academic Press (Boston), pp. 485–502.

T. Vijayaraghavan (1951). *On a problem in elementary number theory*, J. Indian Math. Soc. (N.S.) **15**, 51–56.