Small Sieves, Part II: The Selberg Sieve

1. Introduction. In this chapter we introduce the Selberg Upper Bound Sieve. We proceed by showing how one could hope to be led to Selberg's method by attempting to place the ideas of Brun (which we dicussed in the previous chapter) in a more general context. To keep the exposition as simple as possible we again treat just one example, leaving the general case until later. The example will be the so-called Brun-Titchmarsh problem of estimating the number of primes in the interval $A = (M, M+N]$.

2. The Brun-Titchmarsh Problem Let $\mathcal{P}$ be the set of all primes, and write $S(A, \mathcal{P}, z)$ for the number of elements of A which are not divisible by any $p \in \mathcal{P}$ with $p < z$. In our discussion of the Brun Sieve we observed that

$$S(A, \mathcal{P}, z) \; = \; \sum_{d|P_z} \mu(d) A_d \tag{1}$$

where $P_z = \prod_{p\in\mathcal{P}, p<z} p$ and A_d is the number of elements of A which are divisible by d. The main idea of Brun was to replace $\mu(d)$ by a function $\mu^+(d)$ for which it was the case that

$$S(A, \mathcal{P}, z) \; \leq \; \sum_{d|P_z} \mu^+(d) A_d,$$

but which allowed us to handle the error terms (which resulted from our poor knowledge of A_d for large d) more effectively. Such a function was found by the rather special combinatorial device of the Bonferroni inequalities, or inclusion-exclusion principle. Selberg starts by observing that one can, instead of (1), write

$$S(A, \mathcal{P}, z) \; = \; \sum_{a\in A} \sum_{d|(a, P_z)} \mu(d). \tag{2}$$

This is an easy consequence of the fact that $\sum_{d|n} \mu(d) = 0$ when $n \geq 2$. He now observes that if $\lambda : \mathbb{N} \to \mathbb{R}$ is any function with $\lambda(1) = 1$ then one has, for all k,

$$\sum_{d|k} \mu(d) \; \leq \; \left(\sum_{d|k} \lambda(d) \right)^2.$$

This rather vacuous-looking statement is the conceptial heart of Selberg's method. Substituting into (2) gives

$$\begin{aligned} S(A, \mathcal{P}, z) \; &\leq \; \sum_{a\in A} \left(\sum_{d|(a,P_z)} \lambda(d) \right)^2 \\ &= \; \sum_{a\in A} \sum_{d|(a,P_z)} \rho(d) \\ &= \; \sum_{d|P_z} \rho(d) A_d, \end{aligned} \tag{3}$$

where

$$\rho(d) \;=\; \sum_{\mathrm{lcm}(d_1,d_2)=d} \lambda(d_1)\lambda(d_2). \tag{4}$$

This gives us a whole family of functions ρ which perform the rôle taken by μ^+ in the Brun method. We will have the freedom to choose any λ we like later on (subject only to $\lambda(1) = 1$), but before commenting on this we specialise (3) to the Brun-Titchmarsh situation. In this case (where $A = [M+1, M+N))$ one has $A_d = \frac{N}{d} + R_d$, where $|R_d| \leq 1$. Thus (3) becomes

$$\begin{aligned} S(A, \mathcal{P}, z) \;&\leq\; N\sum_{d|P_z} \frac{\rho(d)}{d} \;+\; \sum_{d|P_z} \rho(d) R_d \\ &=\; NQ + E. \end{aligned} \tag{5}$$

The term NQ will be thought of as the main term, whilst E has the status of an error term. We shall choose λ so as to minimise Q, whilst keeping an eye on E by insisting that $\lambda(d) = 0$ for $d > w$, where w will be chosen later. This problem turns out to be equivalent to a simple optimisation problem involving a quadratic form, although as we are about to see this is not quite obvious. Let us begin by writing Q in the form

$$Q \;=\; \sum_{d_1,d_2|P_z} \frac{\lambda(d_1)\lambda(d_2)}{d_1 d_2}(d_1, d_2). \tag{6}$$

The key to rewriting this as a sort of quadratic form lies in the well-known fact that

$$k \;=\; \sum_{d|k} \phi(d) \tag{7}$$

for all k. Substituting this into (6) gives

$$\begin{aligned} Q \;&=\; \sum_{d_1,d_2|P_z} \frac{\lambda(d_1)\lambda(d_2)}{d_1 d_2} \sum_{\delta|(d_1,d_2)} \phi(\delta) \\ &=\; \sum_{k|P_z} \phi(k) \left(\sum_{k|d, d|P_z} \frac{\lambda(d)}{d} \right)^2 \\ &=\; \sum_{k|P_z} \phi(k) y_k^2, \end{aligned} \tag{8}$$

where

$$y_k \;=\; \sum_{k|d, d|P_z} \frac{\lambda(d)}{d}. \tag{9}$$

We have, in (8), expressed Q as a diagonal quadratic form. In order to minimise it we need to translate the condition $\lambda(1) = 1$ into a relation between the y_k. This can be done with the aid of the following slightly fancy version of the Möbius inversion formula, whose proof is left as an exercise to the reader.

Lemma 1 *Suppose that $\mathcal{D} \subseteq \mathbb{N}$ is divisor closed (by which we mean that $d \in \mathcal{D}$ and $d'|d$ imply $d' \in \mathcal{D}$). Suppose that functions $f, g : \mathbb{N} \to \mathbb{C}$ are related by*

$$f(n) \;=\; \sum_{n|d,\, d\in\mathcal{D}} g(d).$$

Then one has

$$g(n) \;=\; \sum_{n|d,\, d\in\mathcal{D}} \mu\left(\frac{n}{d}\right) f(d).$$

It follows immediately from the lemma and the definition of y_k that

$$\frac{\lambda(k)}{k} \;=\; \sum_{k|d,\, d|P_z} \mu\left(\frac{d}{k}\right) y_d. \tag{10}$$

In particular the condition $\lambda(1) = 1$ becomes simply

$$\sum_{d|P_z} \mu(d) y_d \;=\; 1.$$

It is easy to see from (9) and (10) that $\lambda(d) = 0$ for $d > w$ if and only if $y_d = 0$ for $d > w$. Hence we set about minimising

$$Q \;=\; \sum_{d|P_z,\, d\le w} \phi(d) y_d^2$$

subject to

$$\sum_{d|P_z,\, d\le w} \mu(d) y_d \;=\; 1.$$

This is rather easy, and we find that the minimum value of Q is $1/D$, where

$$D \;=\; D(z,w) \;=\; \sum_{d|P_z,\, d\le w} \frac{\mu^2(d)}{\phi(d)} \;=\; \sum_{d|P_z,\, d\le w} \frac{1}{\phi(d)}, \tag{11}$$

and that this occurs when $y_d = \mu(d)/D\phi(d)$ for all d such that $d|P_z$ and $d \le w$. Using (10), we see that this corresponds to choosing

$$\lambda(k) \;=\; \frac{k}{D} \sum_{k|d,\, d|P_z,\, d\le w} \frac{\mu\left(\frac{d}{k}\right)\mu(d)}{\phi(d)}. \tag{12}$$

Logically we could have simply declared this choice of the $\lambda(k)$ at a much earlier stage, since the purpose of much of the above analysis has simply been to motivate our choice. Assuming that the reader has been sufficiently motivated, we suppose from now on that λ is given by (12).

Lemma 2 $E \le (wD)^2$.

Proof Recall that we have $E = \left|\sum_{d|P_z} \rho(d) R_d\right|$, where $|R_d| \le 1$ and $\rho(d) = \sum_{\mathrm{lcm}(d_1,d_2)=d} \lambda(d_1)\lambda(d_2)$. It follows that

$$E \;\le\; \sum_{d|P_z} |\rho(d)| \;\le\; \left(\sum_{d|P_z} |\lambda(d)|\right)^2 \;=\; \left(\sum_{d|P_z, d\le w} |\lambda(d)|\right)^2. \tag{13}$$

However

$$\begin{aligned} |\lambda(k)| \;&=\; \left|\frac{k}{D}\sum_{k|d, d|P_z, d\le w} \frac{\mu\left(\frac{d}{k}\right)\mu(d)}{\phi(d)}\right| \\ &\le\; \frac{k}{D}\sum_{k|d, d|P_z, d\le w} \frac{1}{\phi(d)} \\ &\le\; \frac{k}{D\phi(k)}\sum_{d|P_z, d\le w} \frac{1}{\phi(d)} \\ &=\; \frac{k}{\phi(k)}, \end{aligned}$$

the last step being an application of (11). The lemma follows immediately from (13) and a further use of (11). $\square$

Now recall (5), which said that $S(A, \mathcal{P}, z) \le NQ + E$. The above lemma, together with the fact that $Q = 1/D$, implies that

$$S(A, \mathcal{P}, z) \;\le\; N/D + (wD)^2. \tag{14}$$

This reduces the problem of bounding $S(A, \mathcal{P}, z)$ from above to the rather simpler-looking problem of estimating $D = D(z, w)$ above and below.

Lemma 3 *Let $D(z, w) = \sum_{d|P_z, d\le w} 1/\phi(d)$. Then for $w \ge z \ge 2$ we have*

$$\log z \;\le\; D(z, w) \;\le\; C \log z,$$

where C is an absolute constant.

Proof For the upper bound we recall the result of Mertens to the effect that

$$\prod_{p\le z} \left(1 - p^{-1}\right)^{-1} \;\ll\; \log z. \tag{15}$$

To use this, note that

$$D(z, w) \;\le\; \sum_{d|P_z} \frac{1}{\phi(d)} \;\le\; \prod_{p\le z}\left(1 + \frac{1}{\phi(p)}\right),$$

an expression which is equal to the object studied by Mertens. For the lower bound, observe that one can write

$$\frac{1}{\phi(d)} = \prod_{p|d}\left(\frac{1}{p}+\frac{1}{p^2}+\dots\right)$$

when d is squarefree. Hence

$$\begin{aligned} D(z,z) &= \sum_{d|P_z, d\le z}\frac{1}{\phi(d)} \\ &= 1+\sum_{2\le d\le z, \mu(d)\neq 0}\prod_{p|d}\left(\frac{1}{p}+\frac{1}{p^2}+\dots\right) \\ &\ge \sum_{m\le z}\frac{1}{m} \\ &\ge \log z. \end{aligned}$$

Noting that $w \ge z$ implies that $D(z,w) \ge D(z,z)$, the proof of the lemma is complete. $\square$

It may seem as if this is only part of the story, since one might well wish to take $w < z$. However in this case one has $D(z,w) = D(w,w)$, and so D can in fact be bounded using the lemma (though this means that taking $w < z$ would not be very sensible). Choosing $z = w = N^{1/2-\epsilon}$ and applying the lemma and (14), we get the following result.

Theorem 4 (Brun-Titchmarsh) *Let $\epsilon > 0$. Then for all $N > N(\epsilon)$ and all M one has*

$$\pi(M+N) - \pi(M) \le \frac{(2+\epsilon)N}{\log N}.$$

This follows, of course, because any prime in the set A is certainly counted by $S(A,\mathcal{P},z)$. The reader may care to convince his or herself that there is no choice of z, w which significantly improves on this theorem.

3. The General Sieve Although in the last section we restricted ourselves to a special case, the ideas and even most of the technical details developed above are valid in a much more general setting. This we now describe.

Take a polynomial h with integer coefficients and consider the set $A = \{h(M), h(M+1), \dots, h(M+N-1)\}$ of size N. We wish to find an upper bound for the number of primes in A, which we do by estimating $S(A,\mathcal{P},z)$, the number of elements of A not divisible by any $p \le z$. The Brun-Titchmarsh situation is thus the simplest possible instance of this problem, in which $h(x) = x$. We succeeded there because we had reasonably strong information about A_d, the number of elements of A which are divisible by d. Something similar is true here. Let d be squarefree (the case of interest) and suppose that the number of elements of $\{h(1), \dots, h(d)\}$ which are divisible by d is $\omega(d)$. Then is easy to see that

$$A_d = |A|\frac{\omega(d)}{d} + R_d,$$

where $|R_d| \le \omega(d)$, and that ω may be extended to a completely multiplicative function on $\mathbb{N}$. These two properties alone allow us to carry through the entire analysis above to get the following analogue of (14).

Theorem 5 (The Selberg Sieve) *In the situation just described we have*

$$S(A, \mathcal{P}, z) \;\le\; \frac{N}{D} + (wD)^2,$$

where

$$D \;=\; D(z, w) \;=\; \sum_{d | P_z, d \le w} \frac{1}{f(d)}$$

and

$$f(k) \;=\; \sum_{d|k} \frac{d\mu\left(\frac{k}{d}\right)}{\omega(d)}.$$

The only slightly novel feature here is the appearance of the function f in place of ϕ from the Brun-Titchmarsh problem. This arises because it is no longer possible to just "notice" a function which plays the rôle taken by ϕ in (7), so one must instead construct f. This is quite easy using Möbius inversion.

For this result to be at all useful one needs to be able to estimate $D(z, w)$. One way of doing this is to follow through the proof of Lemma 3, replacing ϕ with f wherever it occurs. Working through the details and putting $z = w$ gives the following result.

Theorem 6 *We have the inequality*

$$S(A, \mathcal{P}, z) \;\le\; N \left(\sum_{m \le z} \frac{\omega(m)}{m} \right)^{-1} + z^2 \prod_{p \le z} \left(1 - \frac{\omega(p)}{p} \right)^{-2}.$$

It is possible, with a small amount of effort, to apply this result as it stands to a wide range of problems. So that the reader may get a feel for how the Selberg Sieve is applied as quickly as possible, we do this in the next section. After that we give a slightly technical derivation of a general result which is weaker but far more easily applicable than Theorem 6.

4. The Classical Problems In this section we show how Theorem 6 sheds light on two classical problems – the Goldbach and Twin Prime conjectures. For the twin prime problem we consider $h(n) = n(n + 2)$. If $p \le N$ is a twin prime then either $p \le N^{1/3}$ or else $h(p)$ has no prime factors $q \le N^{1/3}$. It follows immediately that $\pi_2(N)$, the number of twin primes less than N, is bounded above by

$$\pi_2(N) \;\le\; S(A, \mathcal{P}, N^{1/3}) + N^{1/3}. \tag{16}$$

We estimate $S(A,\mathcal{P},N^{1/3})$ using the Selberg Sieve as formulated in Theorem 6. In the case at hand we have $\omega(2)=1$ and $\omega(p)=2$ for $p\geq 2$. It is not hard to see that we have

$$\prod_{p\leq N^{1/3}}\left(1-\frac{\omega(p)}{p}\right)^{-1} \ll (\log N)^2.$$

Indeed if $p>5$ then one has

$$\left(1-\frac{2}{p}\right)^{-1} \leq \left(1-\frac{1}{p}\right)^{-2}\left(1-\frac{2}{p^2}\right)^{-1},$$

and our inequality follows from Merten's theorem (15) and the observation that $\prod_p\left(1-2p^{-2}\right)$ converges. To bound the term $\sum_{m\leq N^{1/3}}\frac{\omega(m)}{m}$ below one can observe that $\omega(n)\geq d(n)$, the number of divisors of n, for all odd n. This follows by writing $n=p_1^{\alpha_1}\dots p_k^{\alpha_k}$, so that $\omega(n)=2^{\alpha_1}\dots 2^{\alpha_k}$ and $d(n)=(\alpha_1+1)\dots(\alpha_k+1)$. From this observation one can deduce that

$$\begin{aligned}\sum_{\substack{m\leq N^{1/3}\\(2,m)=1}}\frac{\omega(m)}{m} &\geq \sum_{\substack{m\leq N^{1/3}\\(2,m)=1}}\frac{d(m)}{m}\\ &\geq \left(\sum_{\substack{m\leq N^{1/6}\\(2,m)=1}}\frac{1}{m}\right)^2\\ &\gg (\log N)^2.\end{aligned}$$

This implies, using (16), the following distinct improvement of the result we obtained using Brun's Sieve.

Proposition 7 *The number of twin primes less than N, $\pi_2(N)$, satisfies*

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

To deal with the Goldbach Problem we consider, for some even integer N, the polynomial $h(n)=n(N-n)$. If p is such that $N-p$ is also prime then either p lies in $(0,N^{1/3})\cup(N-N^{1/3},N)$ or else $h(p)$ has no prime factor $q\leq N^{1/3}$. It follows that $r(N)$, the number of representations of N as a sum of two primes, is bounded by

$$r(N) \leq S(A,\mathcal{P},N^{1/3})+2N^{1/3}. \tag{17}$$

In this case we have $\omega(p)=2$ unless $p|N$, in which case $\omega(p)=1$. Once again we apply Theorem 6. It is easy to see, exactly as before, that

$$\prod_{p\leq N^{1/3}}\left(1-\frac{\omega(p)}{p}\right)^{-1} \ll (\log N)^2. \tag{18}$$

To bound the term $\sum_{m \leq N^{1/3}} \frac{\omega(m)}{m}$ below suppose that $m = p_1^{\alpha_1} \dots p_s^{\alpha_s} q_1^{\beta_1} \dots q_t^{\beta_t}$, where $p_1, \dots, p_s$ are the primes dividing N. Then $\omega(m)$ is equal to $2^{\beta_1} \dots 2^{\beta_t}$, which is at least $(\beta_1 + 1) \dots (\beta_t + 1)$. This, however, is simply the number of divisors of m which are coprime with N. It follows that

$$\prod_{m \leq N^{1/3}} \frac{\omega(m)}{m} \geq \left(\sum_{l \leq N^{1/6}} \frac{1}{l} \right) \left(\sum_{\substack{k \leq N^{1/6} \\ (k,N)=1}} \frac{1}{k} \right).$$

However it is not hard to see that

$$\left(\sum_{\substack{k \leq N^{1/6} \\ (k,N)=1}} \frac{1}{k} \right) \prod_{p|N} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \dots \right) \gg \log N,$$

and so

$$\sum_{m \leq N^{1/3}} \frac{\omega(m)}{m} \gg (\log N)^2 \prod_{p|N} \left(1 - \frac{1}{p} \right).$$

Using the fact that $\prod_p (1 - p^{-2})$ converges, it follows from this, (17) and (18) that

Proposition 8

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

In exactly the same way we can generalise Proposition 7 to

Proposition 9 *The number of primes $p \leq N$ for which $p + M$ is also prime is at most*

$$\frac{N}{(\log N)^2} \prod_{p|M} \left(1 + \frac{1}{p} \right).$$

Proposition 9 has a somewhat surprising application – we can use it to show that the gap between consecutive primes is smaller than expected infinitely often.

Proposition 10 *There is an absolute constant $A < 1$ such that $d_i = p_{i+1} - p_i$, the gap between the ith and $(i+1)$st primes, is at most $A \log i$ infinitely often.*

Proof Consider the set $\{p_i | i \in I\}$ of all primes $p \in (N/2, N]$, together with the differences d_i. We shall suppose for a contradiction that $d_i \geq A \log N$ for all $i \in I$, for some $A < 1$ to be chosen later. The contradiction will follow by choosing an appropriate $B > 1$ and considering the equation

$$\sum_{i \in I} d_i = \sum_{\substack{i \in I \\ A \log N \leq d_i \leq B \log N}} d_i + \sum_{\substack{i \in I \\ d_i > B \log N}} d_i. \tag{19}$$

The left hand sum is evidently $\frac{N}{2}(1+o(1))$. Denoting by $Q = Q(N)$ the number of terms in the first sum on the RHS, we can bound the RHS below by

$$AQ \log N + B \log N \left(\pi(N) - \pi(N/2) - Q\right) \;=\; \frac{BN}{2}(1+o(1)) - (B-A)Q\log N. \tag{20}$$

The Selberg Sieve enters the picture via Proposition 9, which may be used to bound Q below. Indeed if $d_i = m$ then both p_i and $p_i + m$ are prime, and so Q is bounded above by a constant multiple of

$$\begin{aligned} & \frac{N}{(\log N)^2} \sum_{m=A\log N}^{B\log N} \prod_{p|m}\left(1+\frac{1}{p}\right) \\ \leq\; & \frac{N}{(\log N)^2} \sum_{m=A\log N}^{B\log N} \sum_{d|m} \frac{1}{d} \\ \leq\; & \frac{N}{(\log N)^2} \sum_{d\leq B\log N} \frac{1}{d}\left\lceil \frac{(B-A)\log N}{d} \right\rceil \\ \ll\; & \frac{(B-A)N}{\log N}. \end{aligned}$$

Substituting into (20) it follows that the RHS of (19) is at most

$$\left(\tfrac{1}{2} + o(1)\right)\left(B - C(B-A)^2\right) N$$

for some absolute constant C. Dividing (19) through by N and letting $N \geq N(\epsilon)$ be sufficiently large gives

$$1 + \epsilon \;\geq\; B - C(B-A)^2.$$

Taking $\epsilon = 1/20C$, $A = 1 - \frac{1}{10C}$ and $B = 1 + \frac{1}{10C}$ gives a contradiction, as can easily be checked. □

5. A Useful Upper Bound Our aim in this section is to prove a general result which allows one to obtain the results of §4 without the need for the *ad hoc* estimates we were forced to use there. As motivation let us suppose that we have a polynomial h with integer coefficients, and as usual let $\omega(p)/p$ denote the proportion of integers n for which $h(n)$ is divisible by p. An upper bound for the number of primes in the set $\{h(1), \dots, h(N)\}$ gives information about a number of classical problems as we showed in the previous section. Let us write $S(h, \mathcal{P}, N)$ for this quantity. Now let X be a uniform random variable on $\{1, \dots, N\}$ and write E_p for the event that $p|h(X)$. Let us just suppose that the events E_p, $p \leq N^{1/2}$, were independent. The only way that $h(X)$ can be prime is if either $h(X) \leq N^{1/2}$ or if $X \in \bigwedge_p \overline{E}_p$, and this happens with probability at most

$$O\left(N^{-1/2}\right) + \prod_{p\leq N^{1/2}} \left(1 - \frac{\omega(p)}{p}\right).$$

Thus it is not unreasonable to expect that

$$S(h, \mathcal{P}, N) \ \ll \ N \prod_{p \leq N^{1/2}} \left(1 - \frac{\omega(p)}{p}\right).$$

It turns out that the heuristic we used in deriving this result is completely inaccurate. The reader may care to verify that the estimate such an argument gives for $\pi(x)$ is out by a constant factor (she may need Merten's result $\prod_{p \leq x} \left(1 - p^{-1}\right) \sim e^{\gamma} \log x$). However the result itself is broadly correct, and provides a useful guide to what sieve theory is capable of.

In what follows we shall assume that $\omega(p) \leq C$ for some $C \geq 1$, a rather simple condition which is satisfied in all the applications we have discussed. This may be a good time to mention κ, the dimension of a sieve. Roughly speaking κ is the average value of $\omega(p)$ (there are various ways of making this precise), and it is one of the most important quantities associated with a sieve. When $\kappa = 1$ we have what is called a *linear* sieve. Of the sieves we have encountered so far only the one used in the Brun-Titchmarsh problem was linear.

Our aim now is to find a lower bound for $D(z, w)$. The reader may care to refresh his memory of this quantity and the other notation used in §3. Observe that if p is prime then

$$f(p) \ = \ \frac{p}{\omega(p)} - 1$$

and so

$$\sum_{d | P_z} \frac{1}{f(d)} \ = \ \prod_{p \leq z} \left(1 - \frac{\omega(p)}{p}\right)^{-1}. \tag{21}$$

To bound $D(z, w)$ below, then, it is sufficient to bound the quantity

$$E(z, w) \ = \ \sum_{d | P_z, d \geq w} \frac{1}{f(d)}$$

above. The trick to doing this is to observe that for any $0 < s < 1$ we have

$$\begin{aligned} E(z, w) \ &\leq \ w^{s-1} \sum_{d | P_z} \frac{d^{1-s}}{f(d)} \\ &= \ w^{s-1} \prod_{p \leq z} \left(1 + \frac{\omega(p)}{p^s \left(1 - \omega(p)/p\right)}\right) \end{aligned}$$

and so

$$\begin{aligned} E(z, w) \cdot \prod_{p \leq z} \left(1 - \frac{\omega(p)}{p}\right) \ &\leq \ w^{s-1} \prod_{p \leq z} \left(1 - \frac{\omega(p)}{p} + \frac{\omega(p)}{p^s}\right). \\ &\leq \ \exp\left(-(1-s) \log w + \sum_{p \leq z} \omega(p) \left(\frac{1}{p^s} - \frac{1}{p}\right)\right). \end{aligned} \tag{22}$$

Now by the mean value theorem we have

$$\frac{1}{p^s} - \frac{1}{p} \;\leq\; \frac{(1-s)\log p}{p^s}$$

and so

$$\sum_{p \leq z} \left(\frac{1}{p^s} - \frac{1}{p} \right) \;\leq\; 2z^{1-s}$$

for z sufficiently large. Substituting into (22) and making the eminently sensible choice of s given by $1 - s = (\log z)^{-1}$ yields

$$E(z,w) \cdot \prod_{p \leq z} \left(1 - \frac{\omega(p)}{p} \right) \;\leq\; \exp\left(-\frac{\log w}{\log z} + 2eC \right).$$

In particular we have

$$E(z, z^{7C}) \;\leq\; \tfrac{1}{2} \prod_{p \leq z} \left(1 - \frac{\omega(p)}{p} \right)^{-1}$$

which immediately implies, from (21), that

$$D(z, z^{7C}) \;\geq\; \tfrac{1}{2} \prod_{p \leq z} \left(1 - \frac{\omega(p)}{p} \right)^{-1}.$$

Let us combine our findings with Theorem 5 in the form of a theorem.

Theorem 11 *Suppose, in our earlier notation, that $\omega(p) \leq C$ for all primes p. Then*

$$S(A, \mathcal{P}, N^{1/16C}) \;\ll\; N \prod_{p \leq N^{1/16C}} \left(1 - \frac{\omega(p)}{p} \right),$$

where the implied constant depends only on C.

Proof Put $w = z^{7C}$ in Theorem 5. We must deal with the error term $(wD)^2$ of that theorem. We have

$$D(z,w)^2 \;\leq\; \prod_{p \leq z} \left(1 - \frac{\omega(p)}{p} \right)^{-2} \;\ll\; (\log z)^{2C}$$

by Mertens' result, and this is dominated by z for large z. Thus the error term is at most z^{15C}, and when $z = N^{1/16C}$ this is dominated by the main term (which is of order at least $N(\log N)^{-C}$ by another application of Mertens' result). $\square$

References All of this material is standard, but it is not easy to find a really good reference. In preparing the above I used Nathanson's *Additive Number Theory: The Classical Bases*, Halberstam and Roth's *Sequences*, Halberstam and Richert's *Sieve Methods* and unpublished lecture notes of Alan Baker. Our proof of Proposition 10 is taken from an exercise in Parent's *Exercices de théorie des nombres*.