Part II

Sieves large and small

Chapter Six

Sums used in sieve theory

The purpose of this section is to estimate some sums that we will need in our study of sieves. The reader should feel free to skip ahead and refer back to this chapter as needed.

The sums $L(x)$, $L_q(x)$ defined in (6.1) will be familiar to those who have worked through or applied Selberg's sieve. They also tend to appear in applications of the large sieve, as will actually be the case for us. In §6.1, we go over some simple bounds, and also some asymptotically optimal bounds due to Ramaré. We shall later see how to use the two kinds of bounds together.

The remaining sections are on sums necessary for our treatment of a quadratic sieve in Ch. 7. First, in §6.2, we estimate the sum $\tilde{m}_d$, related to the sum $\check{m}_d$ from §5.3. Indeed, our estimates on $\tilde{m}_d$ will be based in part on our bounds on $\check{m}_d$ and in part on computations.

Our work in Ch. 7 will rely crucially on estimating of $g_v(y_1, y_2)$, defined in (6.39) in terms of $\tilde{m}_d$. We will discuss various possible approaches in §6.3. Contour integration gives us the right idea but seems impractical as an explicit method. Instead, we will proceed, again, by combining computations and mainly real-analytic methods, together with our bounds on $\check{m}_d$.

We begin in §6.4 by seeing how g_v relates to the sum h_v, which we obtain by taking out the constant term from the sums $\tilde{m}_{dv}$ in the definition of g_v. The sum $h_v(y_1, y_2)$ can be bounded trivially by a sum $\mathbf{h}_v(y_1, y_2)$. In §6.5, we will see how to bound $\mathbf{h}_v(y_1, y_2)$ using our estimates on $\tilde{m}(t)$, as well as some computation intertwined with some analysis.

In §6.6, we will compute $h_v(y_1, y_2)$ for y_1, y_2 bounded by a fairly large constant Y, and $h_v(y, y)$ for y bounded by a larger constant Y. Some of the algorithms here are not completely trivial.

We finish by §6.7, which is an optional section – not used elsewhere – that is devoted to proving an estimate (Prop. 6.30) on the unsmoothed sum considered in [DIT83]. We can easily fill this gap in the literature using our arguments. Thus we see that some of the work we did towards estimating g_v can be used elsewhere in number theory.

6.1 SUMS OF $\mu^2(n)/\phi(n)$

The sums

$$L(x) = \sum_{n\leq x} \frac{\mu^2(n)}{\phi(n)}, \qquad L_q(x) = \sum_{\substack{n\leq x\\(n,q)=1}} \frac{\mu^2(n)}{\phi(n)} \tag{6.1}$$

appear time and again in analytic number theory – for instance, in the denominator of the main estimate in Selberg's sieve for primes (as in [IK04, §6.6]), in the large sieve (as in [MV73], [Ram09] or of course this chapter), and also in previous work on representing integers as sums of primes [RV83], [Ram95].

We can figure out the asymptotics of $L(x)$ by examining the behavior of the corresponding series $\sum_n \frac{\mu^2(n)}{\phi(n)} n^{-s}$ as $s \to 0^+$:

$$\begin{aligned}\frac{\sum_n \frac{\mu^2(n)}{\phi(n)} n^{-s}}{\zeta(s+1)} &= \prod_p \left(1 + \frac{p^{-s}}{p-1}\right)(1-p^{-s-1}) = \prod_p \left(1 + \frac{p^{-s}-p^{-2s}}{p(p-1)}\right)\\ &= 1 + \left(\sum_p \frac{\log p}{p(p-1)}\right) s + \dots,\end{aligned} \tag{6.2}$$

because $p^{-s} - p^{-2s} = e^{-s\log p} - e^{-2s\log p} = (1 - s\log p + \dots) - (1 - 2s\log p + \dots)$. Since $\zeta(s+1) = 1/s + \gamma + (\dots)s$, it follows that $\sum_n \frac{\mu^2(n)}{\phi(n)} n^{-s}$ has the expansion $1/s + c_0 + (\dots)s$ for s around 0, where

$$c_0 = \gamma + \sum_p \frac{\log p}{p(p-1)} = 1.33258227\dots, \tag{6.3}$$

and so $L(x)$ asymptotes to $\log x + c_0$ as $x \to \infty$. (The numerical value of c_0 is as in [RS62, (2.11)].) In the same way, we can show that, for any fixed q, $L_q(x)$ asymptotes to $(\phi(q)/q)(\log x + c_0 + \sum_{p|q} (\log p)/p)$.

Giving good explicit bounds is, as usual, a different and more complicated matter, especially when q is unbounded; at the same time, the matter is not really difficult – thanks ultimately to the fact that the infinite products in (6.2) converge for $\Re s > 1/2$. We will go over several kinds of bounds. Simple bounds (§6.1.1) will prove useful for many purposes; indeed, the classical bounds up to (6.6) will be enough for Chapter 8. We will also prove further elementary bounds (§6.1.1, after (6.6)). We will later combine them with rather sharp estimates due to Ramaré (§6.1.2) to estimate ratios of the form $L_q(R/q)/L_q(R_0/q)$ in Chapter 9.

6.1.1 Elementary bounds

We may start with the following bound, simple and well-known: for $x > 0$,

$$\begin{aligned} L(x) &= \sum_{n\le x} \frac{\mu^2(n)}{\phi(n)} = \sum_{n\le x} \frac{\mu^2(n)}{n} \prod_{p|n} \left(1 - \frac{1}{p}\right)^{-1} \\ &= \sum_{n\le x} \frac{\mu^2(n)}{n} \prod_{p|n} \sum_{j\ge 1} \frac{1}{p^j} \ge \sum_{n\le x} \frac{1}{n} > \log x. \end{aligned} \tag{6.4}$$

While the bound (6.4) is tight for $x \to 1^{-1}$, it is clear that one can improve it under the assumption of lower bounds on x. For one thing, we could use (3.25) in the last step, and so obtain $L(x) \ge \log x + \gamma - 1/x$. The elementary procedure in [MV73, Lemmas 5–7] can be used to go further, permitting an explicit estimate of the form $L(x) > \log x + \gamma + \sum_{p\le C} (\log p)/p(p-1)$ for any C and all x greater than a constant depending only on C. In particular, [MV73, Lemmas 7] combines such an explicit estimate with a small computation to show that $L(x) \ge \log x + 1.07$ for all $x \ge 6$; the same method (applied with $C = 5$) can be used to yield $L(x) \ge \log x + 1.114$. We will see more precise bounds (due to Ramaré) in §6.1.2.

Let us now discuss the relation between $L(x)$ and $L_q(x)$. Since, for q, r coprime,

$$\frac{q}{\phi(q)} L_{qr}(x) = \sum_{d|q} \frac{\mu^2(d)}{\phi(d)} \sum_{\substack{n\le x \\ (n,qr)=1}} \frac{\mu^2(n)}{\phi(n)} = \sum_{d|q} \sum_{\substack{n\le x \\ (n,qr)=1}} \frac{\mu^2(dn)}{\phi(dn)},$$

it is clear that

$$L_r(x) \le \frac{q}{\phi(q)} L_{qr}(x) \le L_r(qx). \tag{6.5}$$

In particular,

$$L_q(x) > \frac{\phi(q)}{q} \log x \tag{6.6}$$

for all $x > 0$. These inequalities can already be found in [vLR65, §1].

We can continue the same train of thought as follows:

$$\begin{aligned} \frac{q}{\phi(q)} L_q(x) - L(x) &= \sum_{d|q} \frac{\mu^2(d)}{\phi(d)} \sum_{\substack{x/d<n\le x \\ (n,q)=1}} \frac{\mu^2(n)}{\phi(n)} \\ &= \sum_{d|q} \frac{\mu^2(d)}{\phi(d)} \left(L_q(x) - L_q(x/d)\right) \le \sum_{d|q} \frac{\mu^2(d)}{\phi(d)} \left(L(x) - L(x/d)\right). \end{aligned}$$

Thus, if we assume that $L(x) \le \log x + C_1$, we obtain

$$L_q(x) \le \frac{\phi(q)}{q} \cdot (\log x + C_{1,q}), \tag{6.7}$$

where

$$\begin{aligned} C_{1,q} &= C_1 + \sum_{\substack{d|q\\ d>1}} \frac{\mu^2(d)}{\phi(d)}(\log d + C_1) \\ &= \frac{q}{\phi(q)}C_1 + \sum_{p|q}(\log p)\sum_{\substack{d|q\\ p|d}} \frac{\mu^2(d)}{\phi(d)} = \frac{q}{\phi(q)}\left(C_1 + \sum_{p|q}\frac{\log p}{p}\right). \end{aligned} \tag{6.8}$$

In exactly the same way, we can show that, for any $r|q$,

$$\frac{q}{\phi(q)}L_q(x) - \frac{r}{\phi(r)}L_r(x) \le \sum_{\substack{d|q\\ (d,r)=1}} \frac{\mu^2(d)}{\phi(d)}\frac{r}{\phi(r)}\left(L_r(x) - L_r(x/d)\right),$$

and so, if we assume $L_r(x) \le (\phi(r)/r)\cdot(\log x + C_r)$, we obtain, using (6.6), that

$$L_q(x) \le \frac{\phi(q)}{q}\cdot(\log x + C_{r,q}), \tag{6.9}$$

where

$$C_{r,q} = \frac{q/r}{\phi(q)/\phi(r)}\left(C_r + \sum_{\substack{p|q\\ p\nmid r}}\frac{\log p}{p}\right). \tag{6.10}$$

6.1.2 Further elementary estimates: the convolution method

The following is a sharper version of the estimate [Ram95, Lem. 3.4] used in the first version of the present chapter. (That older estimate would be enough for our purposes.)

Proposition 6.1 ([RA17, Thm. 1.1])**.** *Let $q \ge 1$, $x > 0$. Let L_q be as in (9.5). Then*

$$L_q(x) = \frac{\phi(q)}{q}\left(\log x + c_0 + \sum_{p|q}\frac{\log p}{p}\right) + O^*\left(\frac{5.9j(q)}{\sqrt{x}}\right), \tag{6.11}$$

where c_0 is as in (6.3) and

$$j(q) = \prod_{\substack{p|q\\ p\neq 2}} \frac{p^{3/2} + p - \sqrt{p} - 1}{p^{3/2} - \sqrt{p} + 1} \prod_{2|q} \frac{21}{25}.$$

Ramaré calls the basic method of proof the *convolution method*. Since (as might be expected) the proof of [RA17, Thm. 1.1] involves several refinements, let us expound instead the proof of the older result [Ram95, Lem. 3.4].

Lemma 6.2. *Let $q \geq 1$, $x > 0$. Let L_q be as in (9.5). Then*

$$L_q(x) = \frac{\phi(q)}{q}\left(\log x + c_0 + \sum_{p|q}\frac{\log p}{p}\right) + O^*\left(\frac{7.35985 j_{1/3}(q)}{x^{1/3}}\right), \tag{6.12}$$

where c_0 is as in (6.3) and

$$j_{1/3}(q) = \prod_{p|q}(1+p^{-2/3})\left(1+\frac{p^{1/3}+p^{2/3}}{p(p-1)}\right)^{-1}.$$

This bound is exactly [Ram95, Lem. 3.4], except that [Ram95] gives the constant 7.284 instead of 7.35985, due to a numerical error.

Proof. We recall that $L_q(x) = \sum_{n\leq x} f(n)$, where $f(n) = \mu^2(n)/\phi(n)$ if $(n,q) = 1$ and $f(n) = 0$ otherwise. We may write f as a Dirichlet convolution:

$$f(n) = h(n) * \frac{1}{n},$$

where $h(n)$ is the multiplicative function defined by

$$h(p) = \frac{1}{p(p-1)}, \quad h(p^2) = -\frac{1}{p(p-1)}, \quad h(p^k) = 0 \quad \text{if } k \geq 3$$

for every prime $p \nmid q$, and

$$h(p) = -\frac{1}{p}, \quad h(p^k) = 0 \quad \text{if } k \geq 2$$

for $p|q$. Then

$$L_q(x) = \sum_m h(m) \sum_{n \leq x/m} \frac{1}{n}.$$

Note that no condition of the form $m \leq x$ is needed; we are taking a sum over all positive integers m.

We are now free to use any estimate of the form

$$\sum_{n\leq y}\frac{1}{n} = \log y + \gamma + O^*\left(\frac{c_\alpha}{y^\alpha}\right), \qquad \alpha > 0, \tag{6.13}$$

provided that it is valid for all $y > 0$ (and not just $y \geq 1$). It follows easily from (3.32) that, for $0 < \alpha \leq 1$ and all $0 < y \leq 1$, (6.13) holds with $c_\alpha = \max(\gamma, \alpha^{-1}e^{-\alpha\gamma-1})$. (For such c_α, the bound (3.32) is stronger than (6.13) for all $y \geq 1$, and we also have $\log y + \gamma \leq c_\alpha/y$ for $0 < y \leq 1$; thus, it remains only to check that $\log y + \gamma \geq -c_\alpha/y$, as we can easily do by taking derivatives.)

We obtain

$$L_q(x) = \sum_m h(m) \log \frac{x}{m} + \gamma \sum_m h(m) + O^*\left(\frac{c_\alpha}{x^\alpha} \sum_m |h(m)| m^\alpha\right).$$

It is easy to see that $\sum_m h(m) = H(0)$ and $-\sum_m h(m) \log m = H'(0)$, where $H(s) = \sum_m h(m) m^{-s}$. Expanding $H(s)$ as an Euler product $H(s) = \prod_p (1 + h(p)p^{-s} + h(p^2)p^{-2s} + \dots)$, we see that

$$H(0) = \prod_{p|q} \left(1 - \frac{1}{p}\right) = \frac{\phi(q)}{q},$$

$$\frac{H'(0)}{H(0)} = (\log H(s))'|_{s=0} = \sum_{p\nmid q} \frac{\log p}{p(p-1)} + \sum_{p|q} \frac{(\log p)/p}{1 - 1/p} = \sum_p \frac{\log p}{p(p-1)} + \sum_{p|q} \frac{\log p}{p}.$$

Hence

$$\begin{aligned} \sum_m h(m) \log \frac{x}{m} + \gamma \sum_m h(m) &= \frac{\phi(q)}{q} \left(\log x + \gamma + \frac{H'(0)}{H(0)}\right) \\ &= \frac{\phi(q)}{q} \left(\log x + c_0 + \sum_{p|q} \frac{\log p}{p}\right). \end{aligned}$$

Let $\overline{H}(s) = \sum |h(m)| m^{-s}$. Evidently, $\overline{H}(-\alpha) = \sum_m |h(m)| m^\alpha$, assuming convergence. This series does converge provided that $\alpha < 1/2$. Its value is

$$\begin{aligned} &\prod_{p\nmid q} \left(1 + \frac{p^\alpha + p^{2\alpha}}{p(p-1)}\right) \cdot \prod_{p|q} \left(1 + \frac{p^\alpha}{p}\right) \\ =&C_\alpha \cdot \prod_{p|q} (1 + p^{\alpha-1}) \left(1 + \frac{p^\alpha + p^{2\alpha}}{p(p-1)}\right)^{-1}, \end{aligned}$$

where $C_\alpha = \prod_p (1 + (p^\alpha + p^{2\alpha})/p(p-1))$. Thus

$$L_q(x) = \frac{\phi(q)}{q} \left(\log x + c_0 + \sum_{p|q} \frac{\log p}{p}\right) + O^*\left(\frac{c_\alpha C_\alpha j_\alpha(q)}{x^\alpha}\right)$$

for any $0 < \alpha < 1/2$, where $j_\alpha(q) = \prod_{p|q} (1 + p^{\alpha-1})(1 + (p^\alpha + p^{2\alpha})/p(p-1))$. It is easy to check that, for instance, for $\alpha = 1/3$,

$$c_\alpha = 0.91047\dots, \quad C_\alpha = 8.08355\dots, \quad c_\alpha C_\alpha = 7.35984\dots.$$

Here, of course, we compute C_α as in §4.4, noting that

$$C_\alpha = \zeta(2 - 2\alpha)\zeta(2 - \alpha) \prod_p \left(1 + \frac{P_2(p^{-1}, p^\alpha)}{1 - p^{-1}}\right),$$

where $P_2(x, y) = x^6(y^5 + y^4) - x^5 y^3 - x^4(y^4 + y^3 + y^2) + x^3(y^2 + y)$. □

We will now use Proposition 6.1 to prove upper and lower bounds on $L_q(R)$ for given q: we simply use Prop. 6.1 for R larger than a constant, and check all smaller R computationally. We would obtain exactly the same bounds from Lemma 6.2, as the bounds are tight, being reached by small values of R; we would simply need more computation, though still not much.

Define $\mathrm{err}_{q,R}$ so that

$$L_q(R) = \frac{\phi(q)}{q}\left(\log R + c_0 + \sum_{p|q} \frac{\log p}{p}\right) + \mathrm{err}_{q,R}, \tag{6.14}$$

where c_0 as in (6.3). In other words, $\mathrm{err}_{q,R}$ equals $L_q(R)$ minus what we know to be the asymptotic main term of $L_q(R)$.

Then we obtain, by Prop. 6.1 and a little computation, that, for $R \geq 1$,

$$\begin{aligned} &-1.33259 \leq \mathrm{err}_{1,R} \leq 0.13818, \quad -0.83958 \leq \mathrm{err}_{2,R} \leq 0.16043,\\ &-1.13253 \leq \mathrm{err}_{3,R} \leq 0.40538, \quad -1.32358 \leq \mathrm{err}_{5,R} \leq 0.29754,\\ &-0.68179 \leq \mathrm{err}_{6,R} \leq 0.31822, \quad -1.38049 \leq \mathrm{err}_{7,R} \leq 0.33372, \end{aligned} \tag{6.15}$$

and so forth. We also have the lower bound $\mathrm{err}_{q,R} > -c_0 - \sum_{p|q} (\log p)/p$ (valid for $R > 0$, and tight for $R \to 1^+$), by (6.6).

For $R \geq 200$,

$$\begin{aligned} &-0.02003 \leq \mathrm{err}_{1,R} \leq 0.02123, \quad -0.00906 \leq \mathrm{err}_{2,R} \leq 0.00925,\\ &-0.01502 \leq \mathrm{err}_{3,R} \leq 0.02157, \quad -0.01924 \leq \mathrm{err}_{5,R} \leq 0.01652,\\ &-0.00738 \leq \mathrm{err}_{6,R} \leq 0.00707, \quad -0.01483 \leq \mathrm{err}_{7,R} \leq 0.02133, \end{aligned} \tag{6.16}$$

and so forth. In the end, it will turn out that we only need to compute lower bounds on $\mathrm{err}_{r,R}$ for $r|2310$ and $R \geq 200$, and upper bounds on $\mathrm{err}_{r,R}$ for $r|2310$ and R satisfying some small lower bounds ($R \geq 1$, $R \geq 2$ and so forth).

6.2 A SMOOTHED SUM OF $\mu(N)/\sigma(N)$

Let us estimate the sum $\tilde{m}_d(x)$ defined by

$$\tilde{m}_d(x) = \sum_{\substack{n \leq x\\ (n,d)=1}} \frac{\mu(n)}{\sigma(n)} \log \frac{x}{n}, \tag{6.17}$$

where $\sigma(n)$ denotes the sum of divisors of n, as usual. The sum $\tilde{m}_d(x)$ plays an important role in a quadratic sieve (Chapter 7).

Our procedure for taking the coprimality condition $(n, d) = 1$ into account will be much as in §5.3.4. Our task is somewhat more involved this time, since we will again be relying on (5.49), which is an estimate on $\check{m}(t)$, not $\tilde{m}(t)$, for large t.

6.2.1 The case of x bounded

For x and d bounded, we can compute $\tilde{m}_d(x)$ directly.

Lemma 6.3. *Let $0 < x \leq 10^{12}$. Let $\tilde{m}_d$ be as in (6.17). Then, for $v = 1, 2$,*

$$\left|\tilde{m}_v(x) - \frac{\sigma(v)}{v}\zeta(2)\right| \leq \frac{c_{v,+}}{\sqrt{x}}, \tag{6.18}$$

where

$$c_{1,+} = \frac{\pi^2}{6}, \qquad c_{2,+} = 2e^{\pi^2/8-1}, \tag{6.19}$$

and, moreover, for $v = 1, 2$,

$$\left|\tilde{m}_v(x) - \frac{\sigma(v)}{v}\zeta(2)\right| \leq \frac{c_{v,0}}{\sqrt{x}} + \frac{c_{v,1}}{x^{3/4}}, \tag{6.20}$$

where

$$\begin{aligned} c_{1,0} &= 0.014429, & c_{1,1} &= 1.8, \\ c_{2,0} &= 0.006534, & c_{2,1} &= 3.9877. \end{aligned} \tag{6.21}$$

Proof. By a computation, as in the proof of Lemmas 5.8 or 5.9. A segmented sieve of Eratosthenes can be easily adapted to compute $\sigma(n)$; the procedure is similar to that for $\mu(n)$.

The low ranges have to be tested separately. For $v = 1$ and $1 \leq x \leq 2$, we verify (6.18) using the fact that the derivative of $\pi^2/6\sqrt{x} - (\pi^2/6 - \log x)$ is positive on $[1, 2]$. For $v = 2$, $1 \leq x \leq e$, note that the derivative of $f(x) = (\pi^2/4 - \log x)\sqrt{x}$ vanishes at $x_0 = \exp(\pi^2/4 - 2)$, and that the value of $f(x)$ at x_0 is $2\exp(\pi^2/8 - 1)$. We verify (6.18) for $v = 2$, $e \leq x \leq 3$ and (6.20) for $v = 1, 2$, $1 \leq x \leq v + 1$ by trivial bisection.

□

For x bounded and general d, we opt to derive bounds for $\tilde{m}_d(x)$ from our bounds on $\tilde{m}_1(x)$ and $\tilde{m}_2(x)$.

Lemma 6.4. *Let $\tilde{m}_d$ be as in (6.17). Let $d, v \in \mathbb{Z}^+$ be square-free and coprime. Assume that, for all $0 < t \leq x$,*

$$\left|\tilde{m}_v(t) - \frac{\sigma(v)}{v}\zeta(2)\right| \leq \sum_{i=1}^{I} c_i t^{-\alpha_i},$$

where $I \geq 1$, $c_i \geq 0$ and $0 \leq \alpha_i < 1$. Then

$$\left|\tilde{m}_{dv}(x) - \frac{\sigma(dv)}{dv}\zeta(2)\right| \leq \sum_{i=1}^{I} \frac{c_i}{x^{\alpha_i}} \prod_{p|d} \frac{p+1}{p+1-p^{\alpha_i}}.$$

Proof. By (5.73) with $h(n) = (\log^+ x/n)/\prod_{p|n}(1+1/p)^{v_p(n)}$ for $(n,v) = 1$ and $h(n) = 0$ for $(n,v) \neq 1$, and q and d exchanged,

$$\begin{aligned}\check{m}_{dv}(x) &= \sum_{q|d^\infty} \frac{1}{q} \sum_{\substack{n\\(n,v)=1}} \frac{\mu(n)}{n} \frac{1}{\prod_{p|qn}(1+1/p)^{v_p(qn)}} \log^+ \frac{x/q}{n} \\ &= \sum_{q|d^\infty} \frac{1}{\prod_{p|q}(p+1)^{v_p(q)}} \check{m}_v(x/q),\end{aligned} \tag{6.22}$$

and so

$$\begin{aligned}\left|\check{m}_{dv}(x) - \frac{\sigma(dv)}{dv}\zeta(2)\right| &= \left|\sum_{q|d^\infty} \frac{1}{\prod_{p|q}(p+1)^{v_p(q)}} \check{m}_v(x/q) - \sum_{q|d^\infty} \frac{\sigma(v)\zeta(2)/v}{\prod_{p|q}(p+1)^{v_p(q)}}\right| \\ &\leq \sum_{q|d^\infty} \frac{1}{\prod_{p|q}(p+1)^{v_p(q)}} \sum_{i=1}^{I} c_i \left(\frac{q}{x}\right)^{\alpha_i} \\ &\leq \sum_{i=1}^{I} \frac{c_i}{x^{\alpha_i}} \prod_{p|d} \sum_{k=0}^{\infty} \left(\frac{p^{\alpha_i}}{p+1}\right)^k \leq \sum_{i=1}^{I} \frac{c_i}{x^{\alpha_i}} \prod_{p|d} \frac{p+1}{p+1-p^{\alpha_i}}.\end{aligned} \tag{6.23}$$

□

Corollary 6.5. *Let $\check{m}_d$ be as in (6.17). Let $d, v \in \mathbb{Z}^+$ be square-free and coprime. Then, for $x \leq 10^{12}$,*

$$\left|\check{m}_{dv}(x) - \frac{\sigma(dv)}{dv}\zeta(2)\right| \leq \prod_{p|dv} \frac{p+1}{p+1-p^\alpha} \cdot \frac{c_{v,+}}{x^\alpha} \tag{6.24}$$

for $\alpha \in [0, 1/2]$ arbitrary, where $c_{v,+}$ is as in (6.19). Moreover,

$$\left|\check{m}_{dv}(x) - \frac{\sigma(dv)}{dv}\zeta(2)\right| \leq \frac{0.014429}{\sqrt{x}} \prod_{p|d} \frac{p+1}{p+1-\sqrt{p}} + \frac{1.8}{x^{3/4}} \prod_{p|d} \frac{p+1}{p+1-p^{3/4}}, \tag{6.25}$$

where $c_{v,i}$, $i = 0, 1$, are as in (6.21).

Proof. Immediate from Lemmas 6.3 and 6.4. □

6.2.2 Bounds for x arbitrary

We also need bounds on $\check{m}_d(x)$ for x arbitrary. We will derive them from our bounds on $\check{m}(x) = \check{m}_1(x)$. Let us then begin by establishing the relation between the two sums.

We will first need to prove a simple lemma. The proof and the statement are of a well-known kind, in that they involve expressing a function as a Dirichlet convolution of two other functions. (It is the same idea as in, say, the beginning of the proof of Lemma 6.2; we can say that what follows is another instance of the *convolution method*.)

Lemma 6.6. *For any function $g : \mathbb{Z}^+ \to \mathbb{C}$ of compact support and any $w \in \mathbb{Z}^+$,*

$$\sum_{\substack{m \\ (m,w)=1}} \frac{\mu(m)}{\sigma(m)} g(m) = \sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',w)=1}} \frac{1}{d'\sigma(d')} \sum_{\substack{d'' \\ d''|(d'w)^\infty}} \frac{1}{d''} \sum_r \frac{\mu(r)}{r} g(d'd''r).$$

It will be obvious from the proof that the condition of compact support can be replaced by a condition of fast decay, if needed.

Proof. Since $m/\sigma(m) = \prod_{p|m}(p/(p+1)) = \sum_{d'|m} \prod_{p|d'}(p/(p+1) - 1)$,

$$\begin{aligned}
\sum_{\substack{m \\ (m,w)=1}} \frac{\mu(m)}{\sigma(m)} g(m) &= \sum_{\substack{m \\ (m,w)=1}} \frac{\mu(m)g(m)}{m} \sum_{d'|m} \prod_{p|d'} \frac{-1}{p+1} \\
&= \sum_{\substack{d' \\ \mu(d')^2=1}} \left(\prod_{p|d'} \frac{-1}{p+1} \right) \sum_{\substack{m \\ (m,w)=1 \\ d'|m}} \frac{\mu(m)g(m)}{m} \\
&= \sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',w)=1}} \frac{1}{d'\sigma(d')} \sum_{\substack{r \\ (r,d'w)=1}} \frac{\mu(r)}{r} g(d'r).
\end{aligned}$$

We then use (5.73) on the inner sum, with $h(n) = g(d'n)$. □

By Lemma 6.6 with $w = d$ and

$$g(r) = \begin{cases} \log^+ \frac{x}{r} & \text{if } (r,v) = 1 \\ 0 & \text{if } (r,v) \neq 1, \end{cases}$$

we see that

$$\tilde{m}_{dv}(x) = \sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',dv)=1}} \frac{1}{d'\sigma(d')} \sum_{\substack{d''|(dd')^\infty \\ (dd'',v)=1}} \frac{1}{d''} \check{m}_v\left(\frac{x}{d'd''}\right) \tag{6.26}$$

for $x > 0$, $v > 0$ arbitrary.

This equality enables us to make clear what we were saying about taking out main terms. We can see from (5.49) and (5.76) that we should regard $\check{m}_v(y)$ as having $v/\phi(v)$ as its leading term, in the sense that $\check{m}_v(y) - v/\phi(v) \to 0$ as $y \to \infty$. We will determine the leading term of $\tilde{m}_{dv}(x)$ by the following lemma.

Lemma 6.7. *For $d, v \in \mathbb{Z}$ coprime*

$$\sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',dv)=1}} \frac{1}{d'\sigma(d')} \sum_{\substack{d''|(dd')^\infty \\ (d'',v)=1}} \frac{1}{d''} = \zeta(2) \prod_{p|d} \left(1 + \frac{1}{p}\right) \prod_{p|v} \left(1 - \frac{1}{p^2}\right). \tag{6.27}$$

Proof. We express infinite sums as infinite products:

$$\begin{aligned} \sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',dv)=1}} \frac{1}{d'\sigma(d')} \sum_{\substack{d''|(dd')^\infty \\ (d'',v)=1}} \frac{1}{d''} &= \sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',dv)=1}} \frac{1}{d'\sigma(d')} \prod_{\substack{p|dd' \\ p\nmid v}} \left(1 - \frac{1}{p}\right)^{-1} \\ &= \prod_{\substack{p|d \\ p\nmid v}} \left(1 - \frac{1}{p}\right)^{-1} \sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',dv)=1}} \frac{1}{\prod_{p|d'}(p+1)(p-1)} \\ &= \prod_{p|d} \frac{p}{p-1} \prod_{p\nmid dv} \frac{p^2}{(p+1)(p-1)} = \prod_{p} \frac{p^2}{p^2-1} \prod_{p|d} \frac{p+1}{p} \prod_{p|v} \frac{p^2-1}{p^2}. \end{aligned} \tag{6.28}$$

□

By (6.26) and (6.27), for d and v coprime,

$$\check{m}_{dv}(x) - \frac{\sigma(dv)}{dv}\zeta(2) = \sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',dv)=1}} \frac{1}{d'\sigma(d')} \sum_{\substack{d''|(dd')^\infty \\ (d'',v)=1}} \frac{1}{d''} \left(\check{m}_v\left(\frac{x}{d'd''}\right) - \frac{v}{\phi(v)}\right). \tag{6.29}$$

Since the leading term of $\check{m}_v$ is $v/\phi(v)$, we now see that $(\sigma(dv)/dv)\zeta(2)$ is the leading term of $\check{m}_{dv}(x)$.

We can now prove our main estimate.

Proposition 6.8. *Let $\check{m}_d$ be as in (6.17). Let $d \in \mathbb{Z}^+$ be square-free. Then, for any $x > 0$,*

$$\begin{aligned} \left|\check{m}_d(x) - \frac{\sigma(d)}{d}\zeta(2)\right| &\le \prod_{p|d} \frac{p+1}{p+1-p^{\alpha_1}} \cdot \frac{k(1-\alpha_1)}{x^{\alpha_1}} \\ &+ \prod_{p|d} \frac{p+1}{p+1-p^{\alpha_2}} \cdot \frac{k(1-\alpha_2)}{389\log x}, \end{aligned}$$

for $\alpha_1 \in [0, 1/2]$ arbitrary, $\alpha_2 = 1/\log 10^{12}$ and

$$\begin{aligned} \kappa_p(\sigma) &= 1 + \frac{p^{-(\sigma+1)}}{(1+p^{-1})(1-p^{-\sigma})} \\ \kappa(\sigma) &= \prod_p \kappa_p(\sigma). \end{aligned} \tag{6.30}$$

Proof. For $x \leq 10^{12}$, the statement follows from (6.24) by $\kappa(1-\alpha_1) \geq \kappa(1) = \zeta(2)$. (It is clear that $\sigma \mapsto \kappa(\sigma)$ is a decreasing function of σ, simply because, for every prime p, the function $\sigma \mapsto \kappa_p(\sigma) = 1 + (1/(1-p^{-\sigma}) - 1)/(p+1)$ is decreasing.) Assume $x \geq 10^{12}$.

By Lemma 5.13, for any $\alpha_1 \in [0,1]$ and $\alpha_2 = 1/\log 10^{12}$,

$$|\check{m}(t) - 1| \leq \frac{1}{t^{\alpha_1}} + \frac{1}{389} \frac{x^{\alpha_2}}{\log x} \frac{1}{t^{\alpha_2}} \tag{6.31}$$

for all $0 < t \leq x$. We can set ourselves, in general, the task of bounding

$$S_{f,d}(x) = \sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',d)=1}} \frac{1}{d'\sigma(d')} \sum_{d''|(dd')^\infty} \frac{1}{d''} f\left(\frac{x}{d'd''}\right), \tag{6.32}$$

where $f : \mathbb{R}^+ \to \mathbb{C}$ is such that $|f(t)|$ is bounded by $\sum_{i=1}^{I} c_i t^{-\alpha_i}$ for $0 < t \leq x$, where $\alpha_i \in [0,1]$. We know that

$$\check{m}_d(x) - \frac{\sigma(d)}{d}\zeta(2) = S_{\check{m}-1,d}(x)$$

by (6.29).

By definition,

$$\begin{aligned}
|S_{f,d}(x)| &\leq \sum_{\substack{d' \\ \mu(d')^2=1 \\ (d',d)=1}} \frac{1}{d'\sigma(d')} \sum_{d''|(dd')^\infty} \frac{1}{d''} \sum_{i=1}^{I} c_i \left(\frac{d'd''}{x}\right)^{\alpha_i} \\
&= \sum_{i=1}^{I} \frac{c_i}{x^{\alpha_i}} \sum_{\substack{d' \\ (d',d)=1}} \frac{\mu^2(d')}{\sigma(d')} \sum_{d''|(dd')^\infty} (d'd'')^{\alpha_i - 1} \\
&= \sum_{i=1}^{I} \frac{c_i/x^{\alpha_i}}{\prod_{p|d}(1-p^{\alpha_i-1})} \sum_{\substack{d' \\ (d',d)=1}} \frac{\mu^2(d')}{(d')^{2-\alpha_i}} \prod_{p|d'} \frac{1}{(1+p^{-1})(1-p^{\alpha_i-1})}.
\end{aligned}$$

We can express the inner sum in terms of a product:

$$\sum_{\substack{d' \\ (d',d)=1}} \frac{\mu^2(d')}{(d')^{2-\alpha_i}} \prod_{p|d'} \frac{1}{(1+p^{-1})(1-p^{\alpha_i-1})} = \prod_{p\nmid d} \left(1 + \frac{p^{-(2-\alpha_i)}}{(1+p^{-1})(1-p^{\alpha_i-1})}\right). \tag{6.33}$$

We conclude that

$$|S_{f,d}(x)| \leq \sum_{i=1}^{I} \frac{c_i}{x^{\alpha_i}} \prod_{p|d} \frac{1+p^{-1}}{1+p^{-1}-p^{\alpha_i-1}} \prod_{p} \kappa_p(1-\alpha_i),$$

where κ_p is as in (6.30).

In our case, by (6.31), we can set $I = 2$, $c_1 = 1$, $c_2 = x^{\alpha_2}/(389 \log x)$. We obtain

$$\left|\tilde{m}_d(x) - \frac{\sigma(d)}{d}\zeta(2)\right| \leq \prod_{p|d} \frac{1+p^{-1}}{1+p^{-1}-p^{\alpha_1 - 1}} \cdot \frac{\kappa(1-\alpha_1)}{x^{\alpha_1}} + \prod_{p|d} \frac{1+p^{-1}}{1+p^{-1}-p^{\alpha_2 - 1}} \cdot \frac{\kappa(1-\alpha_2)}{389 \log x}.$$

□

We will need to estimate the infinite product $\kappa(\sigma)$ for several values of σ.

Lemma 6.9. *Let $\kappa(\sigma)$ be as in (6.30). Then*

$$\kappa(1) = \zeta(2) = 1.644934\ldots, \tag{6.34}$$

$$\kappa\left(1 - \frac{1}{\log 10^{12}}\right) \leq 1.692392, \tag{6.35}$$

$$\kappa(4/7) \leq 2.942346, \tag{6.36}$$

$$\kappa(8/15) \leq 3.243633, \tag{6.37}$$

$$\kappa(1/2) \leq 3.574861. \tag{6.38}$$

Proof. This is exactly the example carried out in §4.4. Set, e.g., $k = 2$, $N = 30000$. □

Incidentally, in [Helb], infinite products like $\kappa(\sigma)$ (and the more complicated infinite products $\kappa(\sigma_1, \sigma_2)$, to be seen in §6.5) were estimated differently. The bounds were suboptimal, and also relied on inequalities that, while elementary, were tricky to prove; as we remarked in §4.7, an automated theorem-prover (QEPCAD) was used to prove one of them. It turns out to be much better to follow the procedure in §4.4.

6.3 DEFINITION OF $g_v(y_1, y_2)$. POSSIBLE APPROACHES.

Let us define

$$g_v(y_1, y_2) = \sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{\sigma(d)^2} \tilde{m}_{dv}(y_1/d) \tilde{m}_{dv}(y_2/d), \tag{6.39}$$

where $\tilde{m}_d(x)$ is as in (6.17). We will need explicit estimates on $g_v(y_0, y_1)$ for our work on what we shall call the natural quadratic sieve (Ch. 7). There are several ways in which one might consider proceeding.

a) **Contour integration.** Define

$$G_v(s_1,s_2) = \sum_{\substack{d\\(d,v)=1}} \frac{\mu(d)}{\sigma(d)^2} \sum_{\substack{d_1'\\(d_1',dv)=1}} \frac{\mu(d_1')}{\sigma(d_1')} \frac{1}{(dd_1')^{s_1}} \sum_{\substack{d_2'\\(d_2',dv)=1}} \frac{\mu(d_2')}{\sigma(d_2')} \frac{1}{(dd_2')^{s_2}}.$$

The following identity is much like Perron's formula, only it is nicer, as a consequence of smoothing:

$$\frac{1}{2\pi i}\int_{\Re s=\sigma} y^s s^{-2} ds = \log^+ y \tag{6.40}$$

for $\sigma > 0$. It can be easily checked by shifting the contour all the way to the left, picking up the contribution of the residue at $s = 0$. (To put it otherwise: as we said in §2.5.1, for $f(x) = \log^+ u/x$, the Mellin transform Mf equals u^s/s^2, and that is equivalent to (6.40).)
Hence

$$g_v(y_1,y_2) = \frac{1}{(2\pi i)^2}\int_{\Re(s_1)=\sigma}\int_{\Re(s_2)=\sigma} y_1^{s_1} y_2^{s_2} G_v(s_1,s_2) s_1^{-2} s_2^{-2} ds_2 ds_1 \tag{6.41}$$

for $\sigma > 0$. What we could do now is shift the lines of integration slightly to the left of $\sigma = 0$. Clearly

$$G_v(s_1,s_2) = \sum_{\substack{d\\(d,v)=1}} \frac{\mu(d)}{\sigma(d)^2 d^{s_1+s_2}} \prod_{p\nmid dv}\left(1-\frac{1}{p^{s_1}(p+1)}\right)\left(1-\frac{1}{p^{s_2}(p+1)}\right)$$

$$= \prod_{p\nmid v}\left(1-\frac{1}{p^{s_1}(p+1)}\right)\prod_{p\nmid v}\left(1-\frac{1}{p^{s_2}(p+1)}\right)$$

$$\cdot \prod_{p\nmid v}\left(1-\frac{1}{(p+1)^2 p^{s_1+s_2}}\left(1-\frac{1}{p^{s_1}(p+1)}\right)^{-1}\left(1-\frac{1}{p^{s_2}(p+1)}\right)^{-1}\right).$$

It is helpful to examine the ratio of $G_v(s_1,s_2)$ to

$$\prod_{p\nmid v}\left(1-\frac{1}{p^{s_1+1}}\right)\left(1-\frac{1}{p^{s_2+1}}\right)$$

$$= \zeta(s_1+1)^{-1}\zeta(s_2+1)^{-1}\prod_{p|v}\left(1-\frac{1}{p^{s_1+1}}\right)^{-1}\left(1-\frac{1}{p^{s_2+1}}\right)^{-1}.$$

As $s_1 \to 0$, $s_2 \to 0$, this ratio tends to

$$\begin{aligned}&\prod_{p\nmid v}\left(\frac{1-1/(p+1)}{1-1/p}\right)^2\prod_{p\nmid v}\left(1-\frac{1}{(p+1)^2}\left(1-\frac{1}{p+1}\right)^{-2}\right)\\&=\prod_{p\nmid v}\left(\frac{1}{1-1/p^2}\right)^2\prod_{p\nmid v}\left(1-\frac{1}{p^2}\right)\\&=\prod_{p\nmid v}\left(1-\frac{1}{p^2}\right)^{-1}=\zeta(2)\prod_{p|v}\left(1-\frac{1}{p^2}\right).\end{aligned}$$

Since $\zeta(s)$ has a simple pole at $s = 1$ with residue 1, the Taylor expansion of $\zeta(s+1)^{-1}$ around $s = 0$ is of the form $s + (\dots)s^2$. Thus, the contribution to (6.41) from the poles at $s_1 = 0$, $s_2 = 0$ is thus

$$\zeta(2)\prod_{p|v}\left(1-\frac{1}{p}\right)^{-2}\left(1-\frac{1}{p^2}\right)=\zeta(2)\prod_{p|v}\frac{p+1}{p-1}=\begin{cases}\zeta(2) & \text{if } v=1,\\ 3\zeta(2) & \text{if } v=2\end{cases} \tag{6.42}$$

times the contribution of the poles at $s_1 = 0$, $s_2 = 0$ to

$$\frac{1}{(2\pi i)^2}\int_{\Re(s_1)=\sigma}\int_{\Re(s_2)=\sigma} y_1^{s_1}y_2^{s_2}s_1^{-1}s_2^{-1}ds_2ds_1.$$

After some work, the main term of that contribution turns out to be 1. Hence, the main term of $g_v(y_1, y_2)$ as $y_1, y_2 \to \infty$ equals

$$\zeta(2) \quad \text{if } v = 1, \qquad 3\zeta(2) \quad \text{if } v = 2.$$

It is obviously useful to know what the main term should be. At the same time, we will not actually be able to follow this approach, for the reason already explained in §5.3: we would have to estimate integrals involving $1/\zeta(s)$ for $\Re s < 1$, and this turns out to be rather hard to do explicitly and well – considerably harder than for integrals involving $\zeta'(s)/\zeta(s)$, say. (Compare the bounds on $1/\zeta(s)$ and $\zeta'(s)/\zeta(s)$ in (3.65).)

If it were not for the fact that we need explicit results, the approach by contour integration would be feasible. In fact, it seems to have been the approach followed in [Mot74] (apud [Jut79b]), and also forms the basis for [GY02].

b) We could use (5.74) to bound the inner sums in (6.39). Such a procedure would lead to a constant larger than the true one in front of the main term, since we would be losing the cancellation due to $\mu(d)$ in (6.39). We would also be losing cancellation in (7.30), and that would lead to a loss of two factors of log.

c) We can carry out an analysis based on what we did in §5.3. Many of the estimates there are based on Ramaré's bound (5.49), combined with numerical bounds for small x. This will be our chosen path. It will involve a mixture of analytic and computational work.

6.4 THE FINITE SUM $g_v(y_1, y_2)$ AND THE INFINITE SUM $h_v(y_1, y_2)$

Our task is to bound the sum $g_v(y_1, y_2)$. The fact that this sum is finite can be convenient for computations, but not necessarily for asymptotics. We will define infinite sums $h_v(y_1, y_2)$ that are more convenient for asymptotic analysis; they are essentially $g_v(y_1, y_2)$ with the main term taken out. We will then see how to go back and forth between $g_v(y_1, y_2)$ and $h_v(y_1, y_2)$.

We define

$$h_v(y_1, y_2) = \sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{\sigma(d)^2} \left(\tilde{m}_{dv}(y_1/d) - \frac{\sigma(dv)}{dv}\zeta(2)\right) \left(\tilde{m}_{dv}(y_2/d) - \frac{\sigma(dv)}{dv}\zeta(2)\right). \tag{6.43}$$

We will write $h_v(y)$ for $h_v(y, y)$.

We can express $g_v(y)$ in terms of $h_v(y_1, y_2)$: we see from (7.30), or directly from (7.28), that

$$\begin{aligned} g_v(y) &= \frac{1}{\left(\log \frac{U_1}{U_0}\right)^2} \sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{\sigma(d)^2} \left(\tilde{m}_{dv}\left(\frac{y_0}{d}\right) - \tilde{m}_{dv}\left(\frac{y_1}{d}\right)\right)^2 \\ &= \frac{h_v(y/U_0, y/U_0) - 2h_v(y/U_0, y/U_1) + h_v(y/U_1, y/U_1)}{(\log U_1/U_0)^2}. \end{aligned} \tag{6.44}$$

Let us now look at how we can express $g_v(y_1, y_2)$ in terms of $h_v(y_1, y_2)$ and viceversa.

Lemma 6.10. *Let v be square-free. Let g_v be as in (6.39); let h_v be as in (6.43). Then, for $y_1, y_2 > 0$ arbitrary,*

$$\begin{aligned} g_v(y_1, y_2) &= h_v(y_1, y_2) + \frac{\sigma(v)}{\phi(v)}\zeta(2) \\ &\quad + \frac{\sigma(v)}{v}\zeta(2)\left(\left(\check{m}_v(y_1) - \frac{v}{\phi(v)}\right) + \left(\check{m}_v(y_2) - \frac{v}{\phi(v)}\right)\right). \end{aligned} \tag{6.45}$$

Proof. Clearly,

$$\begin{aligned} g_v(y_1,y_2) = &\sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{\sigma(d)^2}\left(\tilde{m}_{dv}\left(\frac{y_1}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right)\left(\tilde{m}_{dv}\left(\frac{y_2}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right) \\ &+ \sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{\sigma(d)^2}\tilde{m}_{dv}\left(\frac{y_1}{d}\right)\frac{\sigma(dv)}{dv}\zeta(2) + \sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{\sigma(d)^2}\tilde{m}_{dv}\left(\frac{y_2}{d}\right)\frac{\sigma(dv)}{dv}\zeta(2) \\ &- \left(\frac{\sigma(v)}{v}\zeta(2)\right)^2 \sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{d^2}. \end{aligned}$$

It is also easy to see that

$$\frac{\sigma(v)}{v}\zeta(2)\sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{d^2} = \frac{\sigma(v)}{v}\zeta(2)\prod_{p\nmid v}\left(1-\frac{1}{p^2}\right) = \frac{v}{\phi(v)}.$$

It remains to show that

$$\sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{\sigma(d)^2}\tilde{m}_{dv}\left(\frac{y}{d}\right)\frac{\sigma(d)}{d} = \sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{\sigma(d)d}\left(\sum_{\substack{d' \\ (d',dv)=1}} \frac{\mu(d')}{\sigma(d')}\log^+\frac{y}{dd'}\right) \tag{6.46}$$

equals $\check{m}_v(y)$. This is so because the contribution to the right side of (6.46) of all terms with $dd' = n$ for given $n \le y$ is

$$\sum_{\substack{d|n \\ (d,n/d)=1}} \frac{\mu(d)}{\sigma(d)d}\frac{\mu(n/d)}{\sigma(n/d)}\log\frac{y}{n} = \frac{\mu(n)}{\sigma(n)}\log\frac{y}{n}\sum_{d|n}\frac{1}{d} = \frac{\mu(n)}{n}\log\frac{y}{n}$$

if $(n,v) = 1$ and 0 otherwise. □

6.5 BOUNDS ON $\mathbf{h}_v$

Let us define

$$\mathbf{h}_v(y_1,y_2) = \sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2}\left|\tilde{m}_{dv}\left(\frac{y_1}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right|\left|\tilde{m}_{dv}\left(\frac{y_2}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right|. \tag{6.47}$$

It is obvious from definition (6.43) that

$$|h_v(y_1,y_2)| \le \mathbf{h}_v(y_1,y_2).$$

It is also clear that bounding $|h_v(y_1, y_2)|$ by $\mathbf{h}_v(y_1, y_2)$ amounts to foregoing cancellation on the d variable.

We will now bound $\mathbf{h}_v(y_1, y_2)$. We will use the material in this section when a bound on $\mathbf{h}_v(y_1, y_2)$ suffices, and a more precise estimate is not feasible, either because the variables y_1, y_2 are large or because too many different values of y_1, y_2 would need to be checked individually.

Our starting point will be the following identity. As can be guessed from its form, we will later use it in combination with Corollary 6.5 and Proposition 6.8.

Lemma 6.11. *Let $\alpha_1, \alpha_2 \in [0, 1)$ with $\alpha_1 + \alpha_2 < 1$. Then, for any $y_1, y_2 > 0$,*

$$\sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} \prod_{j=1}^{2} d^{\alpha_j} \prod_{p|d} \frac{p+1}{p+1-p^{\alpha_j}} = \kappa_{\backslash v}(1-\alpha_1, 1-\alpha_2),$$

where

$$\begin{aligned} \kappa_{\backslash v}(\beta_1, \beta_2) &= \prod_{p \nmid v} \kappa_p(\beta_1, \beta_2), \\ \kappa_p(\beta_1, \beta_2) &= 1 + \frac{p^{-(\beta_1+\beta_2)}}{(1-p^{-\beta_1}+p^{-1})(1-p^{-\beta_2}+p^{-1})}. \end{aligned} \tag{6.48}$$

The sum and the product both converge because $\alpha_1 + \alpha_2 < 1$.

Proof. This is just an exercise:

$$\begin{aligned} \sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} \prod_{j=1}^{2} d^{\alpha_j} \prod_{p|d} \frac{p+1}{p+1-p_j^{\alpha}} &= \sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{d^{2-(\alpha_1+\alpha_2)}} \prod_{j=1}^{2} \prod_{p|d} \frac{1}{1+p^{-1}-p^{\alpha_j-1}} \\ &= \prod_{p \nmid v} \left(1 + \frac{p^{\alpha_1+\alpha_2-2}}{\prod_{j=1}^{2}(1+p^{-1}-p^{\alpha_j-1})}\right). \end{aligned}$$

□

Before we apply Lemma 6.11, we will see what a direct computational approach gives us for y_1, y_2 small (§6.5.1). We will then set out an analytic method with computational inputs (§6.5.2); we shall use that method for y_1, y_2 up to a moderate range. Lastly, we will use Lemma 6.11 together with analytic estimates for y_1, y_2 large (§6.5.3).

6.5.1 Bounds on $\mathbf{h}_v(y_1, y_2)$ for y_1, y_2 small

A simple computation yields the following result.

Lemma 6.12. *Let $\mathbf{h}_v(y_1, y_2)$ be as in (6.47). Then, for $y \le 10^6$,*

$$\mathbf{h}_v(y, y) \le \frac{1}{y} \cdot \begin{cases} \min(4.11234 + 0.858 \log y, 9.39557) & \text{if } v = 1, \\ \min(7.71695 + 1.75 \log y, 18.3362) & \text{if } v = 2. \end{cases} \tag{6.49}$$

A straightforward algorithm takes roughly quadratic time on Y to compute $\mathbf{h}_v(y_1, y_2)$ for all $y_1, y_2 \leq Y$. We will use such a simple algorithm, as quadratic time is acceptable for $Y = 10^6$.

Proof. To carry out the computation, it is enough to note that

$$\mathbf{h}_v(y, y) = \mathbf{g}_v(y, y) + \prod_{p|v} \frac{(p+1)^2}{p^2+1} \cdot \frac{\zeta^3(2)}{\zeta(4)} - 2\frac{\sigma(v)}{v}\zeta(2) \sum_{\substack{n \\ (n,v)=1}} \frac{\mu(n)\phi(n)}{n\sigma(n)} \log^+ \frac{y}{n},$$

where $\mathbf{g}_v(y_1, y_2)$ is the finite sum

$$\mathbf{g}_v(y_1, y_2) = \sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} \tilde{m}_{dv}(y_1/d)\tilde{m}_{dv}(y_2/d).$$

We compute $\tilde{m}_{dv}(n)$ and $\dot{m}_{dv}(n)$ for all $d \leq 10^6$ and all $n \leq 10^6/d$ with $(n, v) = 1$ at the very beginning. (As always, $\tilde{m}_{dv}(n)$ and $\dot{m}_{dv}(n)$ give us an expression for $\tilde{m}_{dv}(y)$ for $n \leq y \leq n + v$.) □

In the next subsection, we shall see how to obtain estimates on a much broader range by means of a mixed analytic-computational approach. The bounds will not be optimal, but they will not be much weaker than Lemma 6.12, particularly in the critical range of y small.

We will later see how one can compute $h_v(y, y)$ in time roughly linear on y by means of a less simple-minded algorithm (§6.6.2). It is presumably also possible to give a similar algorithm for computing $\mathbf{h}_v(y, y)$ for y moderate. However, the bounds in §6.5.2, together with those we have just seen, will be more than enough.

6.5.2 Bounds on $\mathbf{h}_v(y_1, y_2)$ for y_1, y_2 small or moderate

By "small or moderate", we will mean "$\leq 10^{12}$". We first derive a simple bound, though we shall actually use the one we will prove right thereafter.

Lemma 6.13. *Let $\mathbf{h}_v$ be as in (6.47), with $v = 1$ or $v = 2$. Then, for $0 < y_1, y_2 \leq 10^{12}$ and any $\epsilon \in (0, 1)$,*

$$\mathbf{h}_v(y_1, y_2) \leq \kappa_{\setminus v}(\beta, \beta) \cdot \frac{c_{v,+}^2}{(y_1 y_2)^{\frac{1-\epsilon}{2}}},$$

where $\beta = (1 + \epsilon)/2$, $c_{v,+}$ is as in (6.19) and $\kappa_{\setminus v}(\beta_1, \beta_2)$ is as in (6.48).

We can compute $\kappa_{\setminus v}(\beta, \beta)$ for $\beta = (1 + \epsilon)/2$ by factoring out $\zeta(1 + \epsilon)$, in the style of the proof of Lemma 6.9. It is easy to see that $\zeta(1 + \epsilon)$ asymptotes to $1/\epsilon$ as $\epsilon \to 0^+$.

Proof. Let $\alpha = (1-\epsilon)/2 < 1/2$. By Corollary 6.5,

$$\mathbf{h}_v(y_1, y_2) \leq \left(\frac{\sigma(v)}{v}\right)^2 \sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} \frac{\zeta(2)^2}{(y_1/d)^\alpha (y_2/d)^\alpha} \left(\prod_{p|d} \frac{p+1}{p+1-p^\alpha}\right)^2.$$

Lemma 6.11 gives us that

$$\sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} d^{2\alpha} \left(\prod_{p|dv} \frac{p+1}{p+1-p^\alpha}\right)^2 = \kappa_{\setminus v}(\beta,\beta) \prod_{p|v} \left(\frac{p+1}{p+1-p^\alpha}\right)^2$$

for $\beta = 1-\alpha$. □

We could have reduced the task to the case $y_1 = y_2$ by Cauchy-Schwarz:

$$\mathbf{h}_v(y_1, y_2) \leq \sqrt{\mathbf{h}_v(y_1, y_1)\mathbf{h}_v(y_2, y_2)}. \tag{6.50}$$

Clearly, using this inequality does not worsen this kind of bound, as it is symmetric in y_1 and y_2. Indeed, we will use (6.50) in the future.

Lemma 6.13 amounts to an instance of Rankin's trick: we use $\epsilon > 0$ in order to bound a sum by an infinite sum that can be expressed as an infinite product. Let us make this matter as clear as possible. Since $\tilde{m}_v(t) = 0$ for $0 < t < 1$,

$$\begin{aligned} \mathbf{h}_v(y,y) = &\sum_{\substack{d \leq y \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} \left| \tilde{m}_{dv}\left(\frac{y}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right|^2 \\ &+ \left(\frac{\sigma(v)}{v}\right)^2 \zeta(2)^2 \sum_{\substack{d > y \\ (d,v)=1}} \frac{\mu^2(d)}{d^2}. \end{aligned} \tag{6.51}$$

To get a clear notion of what we did in the proof of Lemma 6.13, and to see how far Lemma 6.13 is from an optimal bound, consider the simpler sum

$$\frac{1}{y}\sum_{d \leq y} \frac{1}{d} + \sum_{d > y} \frac{1}{d^2}.$$

This sum asymptotes to $(\log y + \gamma)/y + 1/y$. Rankin's trick, in this case, consists in bounding this sum from above by

$$\frac{1}{y}\sum_{d \leq y} \left(\frac{y}{d}\right)^\epsilon \frac{1}{d} + \sum_{d > y} \frac{1}{d^2} \leq \frac{\zeta(1+\epsilon)}{y^{1-\epsilon}} \sim \frac{1}{\epsilon y^{1-\epsilon}}.$$

The expression on the right is minimal for $\epsilon = 1/\log y$. We then have $1/\epsilon y^{1-\epsilon} = e(\log y)/y$. In other words, using Rankin's trick here, we lose a factor of about e.

There are several alternatives. For one thing, Rankin's trick is really a poor man's Tauberian theorem. Thus, one option is to try to use an actual Tauberian theorem, that is, a result that estimates the growth of a sum $\sum_{n\leq x} a_n$ in terms of the properties of $\sum_n a_n n^{-s}$ as a function of s, under some conditions. In the end, this strategy seems to lead us back to option a) ("Contour integration") in §7.3. If $F(s) = \sum_n a_n n^{-s}$ can be continued analytically, as is the case here, then contour integration provides a straightforward way to prove a Tauberian theorem. Conversely, if, as we already saw it to be the case here, contour integration involves significant technical difficulties in the estimation of error terms, analogous difficulties may appear in a Tauberian treatment.

We will take another route. We will bound the expression within absolute values in (6.51) using our bounds in Cor. 6.5. Then we will compute the resulting expression numerically. We can do this computation for all $y \leq Y$ in time essentially linear on Y. Since the sum in (6.51) is finite, we can apply (6.24) with $\alpha = 1/2$; in the proof of Lemma 6.13, we had to use an exponent $\alpha < 1/2$ so as to make our sum – which was then infinite – converge. In fact, we can use the sharper bound (6.25) in Cor. 6.5 instead of (6.24), if we so wish.

As we shall see at the end of §6.6.2, (a) a purely computational approach is in fact feasible for Y small or moderate, (b) the precise estimates we would get then are better than the bounds we are about to obtain, but not by all that much. For most purposes, Prop. 6.14 will do nicely.

Proposition 6.14. *Let $\mathbf{h}_v$ be as in (6.47), with $v = 1$ or $v = 2$. Then, for $0 < y \leq 10^{12}$,*

$$\mathbf{h}_v(y, y) \leq \frac{1}{y} \cdot \begin{cases} 4.89606 + 3.83717 \log y & \text{if } v = 1, \\ 9.57182 + 4.99703 \log y & \text{if } v = 2, \end{cases} \tag{6.52}$$

and

$$\mathbf{h}_v(y, y) \leq \frac{1}{y} \cdot \begin{cases} 74.554 & \text{if } v = 1, \\ 147.6449 & \text{if } v = 2, \end{cases} \tag{6.53}$$

Proof. By Cor. 6.5 and (6.51),

$$\begin{aligned} \mathbf{h}_v(y, y) \leq \frac{c_{v,+}^2}{y} \sum_{\substack{d\leq y \\ (d,v)=1}} \frac{\mu^2(d)}{d} \prod_{p|dv} \left(\frac{p}{p - \sqrt{p} + 1}\right)^2 \\ + \left(\frac{\sigma(v)}{v}\zeta(2)\right)^2 \sum_{\substack{d>y \\ (d,v)=1}} \frac{\mu^2(d)}{d^2} \end{aligned} \tag{6.54}$$

and also

$$\mathbf{h}_v(y,y) \le \frac{c_{v,1}^2}{y} \sum_{\substack{d\le y\\(d,v)=1}} \frac{\mu^2(d)}{d} \prod_{p|d} \left(\frac{p}{p-\sqrt{p}+1}\right)^2$$
$$+ \frac{2c_{v,1}c_{v,2}}{y^{5/4}} \sum_{\substack{d\le y\\(d,v)=1}} \frac{\mu^2(d)}{d^{3/4}} \prod_{p|d} \left(\frac{p}{p-\sqrt{p}+1}\right)\left(\frac{p}{p-p^{3/4}+1}\right)$$
$$+ \frac{c_{v,2}^2}{y^{3/2}} \sum_{\substack{d\le y\\(d,v)=1}} \frac{\mu^2(d)}{\sqrt{d}} \prod_{p|d} \left(\frac{p}{p-p^{3/4}+1}\right)^2 + \left(\frac{\sigma(v)}{v}\right)^2 \zeta(2)^2 \sum_{\substack{d>y\\(d,v)=1}} \frac{\mu^2(d)}{d^2}, \tag{6.55}$$

where $c_{1,1} = 0.014429$, $c_{1,2} = 1.8$, $c_{2,1} = 0.006534$, $c_{2,2} = 3.9877$.

We compute that, for $2 \le y \le 10^{12}$,

$$\sum_{\substack{d\le y\\(d,v)=1}} \frac{\mu^2(d)}{d} \prod_{p|d} \left(\frac{p}{p-\sqrt{p}+1}\right)^2 \le \begin{cases} 1+1.41812\log y & \text{if } v=1,\\ 1+0.78282\log y & \text{if } v=2,\end{cases} \tag{6.56}$$

$$\sum_{\substack{d\le y\\(d,v)=1}} \frac{\mu^2(d)}{d^{3/4}} \prod_{p|d} \left(\frac{p}{p-\sqrt{p}+1}\right)\left(\frac{p}{p-p^{3/4}+1}\right) \le \begin{cases} 14.1958y^{1/4} & \text{if } v=1,\\ 7.2578y^{1/4} & \text{if } v=2,\end{cases} \tag{6.57}$$

$$\sum_{\substack{d\le y\\(d,v)=1}} \frac{\mu^2(d)}{d^{1/2}} \prod_{p|d} \left(\frac{p}{p-p^{3/4}+1}\right)^2 \le \begin{cases} 22.1043\sqrt{y} & \text{if } v=1,\\ 10.2836\sqrt{y} & \text{if } v=2.\end{cases} \tag{6.58}$$

By Lemmas 5.1 and 5.2,

$$\left(\frac{\sigma(v)}{v}\zeta(2)\right)^2 \sum_{\substack{d>y\\(d,v)=1}} \frac{\mu^2(d)}{d^2} \le \begin{cases} \frac{2.19025}{y} & \text{if } v=1 \text{ and } y\ge 2,\\ \frac{3.18845}{y} & \text{if } v=2 \text{ and } y\ge 3.\end{cases} \tag{6.59}$$

We obtain, using (6.54), that, in total,

$$y\mathbf{h}_v(y,y) \le \begin{cases} 4.89606+3.83717\log y & \text{if } v=1 \text{ and } 2\le y\le 10^{12},\\ 9.57182+4.99703\log y & \text{if } v=2 \text{ and } 3\le y\le 10^{12}.\end{cases} \tag{6.60}$$

Using (6.55) instead, we obtain

$$y\mathbf{h}_v(y,y) \le \begin{cases} 74.54579+0.000296\log y & \text{if } v=1 \text{ and } 2\le y\le 10^{12},\\ 167.09396+0.000034\log y & \text{if } v=2 \text{ and } 3\le y\le 10^{12}.\end{cases} \tag{6.61}$$

Note that (6.60) is always better than (6.61) for $v = 2$ and $y \leq 10^{12}$: it implies that $\mathbf{h}_2(y, y) \leq 147.6449$. (In contrast, for $y = 10^{12}$ and $v = 1$, (6.60) gives us $\mathbf{h}_1(y, y) \leq 110.921$, whereas (6.61) gives us that $\mathbf{h}_1(y, y) \leq 74.554$ for all $2 \leq y \leq 10^{12}$.) We have just obtained (6.52) and (6.53) for $y \geq 2$ and $v = 1$, and for $y \geq 3$ and $v = 2$ as well.

For $1 \leq y < 2$ and $v = 1$, or $1 \leq y < 3$ and $v = 2$,

$$\mathbf{h}_v(y, y) = \left|\log y - (\sigma(v)/v)\zeta(2)\right|^2. \tag{6.62}$$

A trivial calculation shows that (6.62) implies the bounds in (6.52) and (6.53) in those ranges. For $0 < y < 1$, of course, $\mathbf{h}_v(y, y) = 0$. □

6.5.3 Bounds on $\mathbf{h}_v(y_1, y_2)$ for y_1, y_2 large

Bounding $\mathbf{h}_v(y_1, y_2)$ for y_1, y_2 both large is not hard given what we know.

Lemma 6.15. *Let $\mathbf{h}_v$ be as in (6.47), with $v = 1$ or $v = 2$. Then, for $y_1, y_2 \geq 10^{12}$ and any $0 < \epsilon \leq 1/2$,*

$$\begin{aligned}\mathbf{h}_v(y_1, y_2) \leq\ & \frac{\kappa_{\cdot v}(1/2)\kappa_{\cdot v}(1/2+\epsilon)}{y_1^{1/2} y_2^{1/2-\epsilon}} \kappa_{\backslash v}(1/2, 1/2+\epsilon) \\ & + \frac{\kappa_{\cdot v}(1/2)\kappa_{\cdot v}(1-\alpha)}{389}\left(\frac{1}{\sqrt{y_1}\log y_2} + \frac{1}{\sqrt{y_2}\log y_1}\right)\kappa_{\backslash v}\left(\frac{1}{2}, 1-\alpha\right) \\ & + \frac{\kappa_{\cdot v}(1-\alpha)^2}{389^2 \log y_1 \log y_2}\kappa_{\backslash v}(1-\alpha, 1-\alpha),\end{aligned} \tag{6.63}$$

where $\alpha = 1/\log 10^{12}$, $\kappa(\beta)$, $\kappa_{\backslash v}(\beta_1, \beta_2)$ are as in (6.30) and (6.48), and

$$\kappa_{\cdot v}(\beta) = \kappa(\beta) \cdot \prod_{p|v} \frac{p+1}{p+1-p^{1-\beta}}. \tag{6.64}$$

Proof. By Prop. 6.8 and Lemma 5.11,

$$\begin{aligned}\left|\tilde{m}_{dv}\left(\frac{y}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right| \leq & \prod_{p|dv} \frac{p+1}{p+1-p^{\alpha'}} \cdot \frac{\kappa(1-\alpha')}{y^{\alpha'}} d^{\alpha'} \\ & + \prod_{p|dv} \frac{p+1}{p+1-p^{\alpha}} \cdot \frac{\kappa(1-\alpha) d^{\alpha}}{389 \log y}\end{aligned} \tag{6.65}$$

for $y/d \geq 10^{12}$, where $\alpha = 1/\log 10^{12}$ and $\alpha' \in [0, 1/2]$ is arbitrary. Recall that, as we noted in the proof of Prop. 6.8, $\kappa(1-\alpha) \geq \kappa(1) = \zeta(2)$. Thus, applying (6.24), we obtain that (6.65) also holds for $y/d < 10^{12}$.

Hence, setting $\alpha' = 1/2$, we obtain that $\mathbf{h}_v(y_1, y_2)$ is bounded by

$$\sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} \prod_{p|dv} \frac{p+1}{p+1-\sqrt{p}} \cdot \frac{\kappa(1/2)}{\sqrt{y_1}} \sqrt{d} \left|\tilde{m}_{dv}\left(\frac{y_2}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right| \tag{6.66}$$

plus

$$\sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} \prod_{p|dv} \frac{p+1}{p+1-p^{\alpha}} \cdot \frac{\kappa(1-\alpha)d^{\alpha}}{389\log y_1} \left| \tilde{m}_{dv}\left(\frac{y_2}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2) \right| . \tag{6.67}$$

Applying (6.65) again, once to (6.66), with $\alpha' = 1/2 - \epsilon$, and once to (6.67), with $\alpha' = 1/2$, we obtain that $\mathbf{h}_v(y_1, y_2)$ is at most

$$\sum_{i=1}^{4} c_i \sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} \prod_{j=1}^{2} d^{\alpha_{i,j}} \prod_{p|dv} \frac{p+1}{p+1-p^{\alpha_{i,j}}} \tag{6.68}$$

for

$$\begin{aligned}
c_1 &= \frac{\kappa(1/2)}{y_1^{1/2}} \frac{\kappa(1/2+\epsilon)}{y_2^{1/2-\epsilon}}, \qquad \alpha_{1,1} = \frac{1}{2}, \quad \alpha_{1,2} = \frac{1}{2} - \epsilon, \\
c_2 &= \frac{\kappa(1/2)}{y_1^{1/2}} \frac{\kappa(1-\alpha)}{389\log y_2}, \qquad \alpha_{2,1} = \frac{1}{2}, \quad \alpha_{2,2} = \alpha, \\
c_3 &= \frac{\kappa(1-\alpha)}{389\log y_1} \frac{\kappa(1/2)}{y_2^{1/2}}, \qquad \alpha_{3,1} = \alpha, \quad \alpha_{3,2} = \frac{1}{2}, \\
c_4 &= \frac{\kappa(1-\alpha)}{389\log y_1} \cdot \frac{\kappa(1-\alpha)}{389\log y_2}, \qquad \alpha_{4,1} = \alpha_{4,2} = \alpha.
\end{aligned}$$

By Lemma 6.11 and (6.64), we conclude that (6.63) holds. The factors $(p+1)/(p+1-p^{1-\beta})$, $p|v$, in (6.63) come from the contribution of primes $p|v$ to the innermost product in (6.68). □

Let us extract explicit estimates from Lemma 6.15. We first need to estimate the infinite product $\kappa(\beta_1, \beta_2)$.

Lemma 6.16. *Let $\kappa(\beta_1, \beta_2) = \prod_p \kappa_p(\beta_1, \beta_2)$, where $\kappa_p(\beta_1, \beta_2)$ is as in (6.48). Then*

$$\kappa\left(1, 1 - \frac{1}{\log 10^{12}}\right) \leq 1.552738, \tag{6.69}$$

$$\kappa\left(1 - \frac{1}{\log 10^{12}}, 1 - \frac{1}{\log 10^{12}}\right) \leq 1.588493, \tag{6.70}$$

$$\kappa\left(1, \frac{1}{2}\right) \leq 2.620185, \tag{6.71}$$

$$\kappa\left(1 - \frac{1}{\log 10^{12}}, \frac{1}{2}\right) \leq 2.791671, \tag{6.72}$$

$$\kappa\left(1, \frac{1}{4}\right) \leq 6.761469, \tag{6.73}$$

$$\kappa\left(\frac{8}{15},\frac{8}{15}\right) \le 19.82414, \tag{6.74}$$

$$\kappa\left(\frac{4}{7},\frac{1}{2}\right) \le 18.50841. \tag{6.75}$$

Proof. The issue to address is that the Euler product defining $\kappa(\beta_1, \beta_2)$ can converge rather slowly. We proceed as in §4.4. Analogously to (4.21),

$$\left(1 + \frac{y_1 y_2}{(1-y_1+x)(1-y_2+x)}\right)\prod_{j=1}^{k}(1-y_1y_2^j) = 1 + \frac{R_k(x,y_1,y_2)}{(1-y_1+x)(1-y_2+x)},$$

where the polynomial R_k is congruent to $y_1 y_2^{k+1}$ modulo the ideal generated by xy_1y_2 and $y_1^2 y_2$, as can be shown by induction on k. Thus, $\kappa(\beta_1, \beta_2)$ equals

$$\prod_{j=1}^{k}\zeta(\beta_1 + j\beta_2)\cdot\prod_p\left(1 + \frac{R_k\left(p^{-1},p^{-\beta_1},p^{-\beta_2}\right)}{(1-p^{-\beta_1}+p^{-1})(1-p^{-\beta_2}+p^{-1})}\right). \tag{6.76}$$

Assuming $0 < \beta_1, \beta_2 \le 1$, we let $\sigma = \min(2\beta_1+\beta_2, \beta_1+(k+1)\beta_2)$, and bound

$$\frac{\left|R_k\left(t,t^{\beta_1},t^{\beta_2}\right)/t^{\sigma}\right|}{(1-t^{\beta_1}+t)(1-t^{\beta_2}+t)} \le \frac{C_k}{(1-t^{\beta_1}+t)(1-t^{\beta_2}+t)}, \tag{6.77}$$

where C_k is the sum of all positive coefficients of monomials in R_k. We can then bound the tail, that is, the product of the terms in (6.76) with $p > N$: use (6.77) together with the fact that $x \mapsto 1 - x^{-\beta} + x^{-1}$ is increasing on x for $x \ge 4$, $\beta \ge 1/2$.

The above procedure is enough to prove (6.69)–(6.73). To prove the other inequalities on $\kappa(\beta_1, \beta_2)$ in our statement, we need to be able to cope with the effect of β_1, β_2 being both significantly smaller than 1. We divide $\kappa(\beta_1, \beta_2)$ by

$$\zeta(\beta_1+\beta_2)\zeta(2\beta_1+\beta_2)\zeta(3\beta_1+\beta_2)\zeta(\beta_1+2\beta_2)\zeta(\beta_1+3\beta_2),$$

and obtain

$$\prod_p\left(1 + \frac{R(p^{-1},p^{-\beta_1},p^{-\beta_2})}{(1-p^{-\beta_1}+p^{-1})\,(1-p^{-\beta_2}+p^{-1})}\right),$$

where $R(x, y_1, y_2)$ equals

$$\begin{aligned}&(1-y_1y_2)\prod_{j=2}^{3}(1-y_1^jy_2)(1-y_1y_2^j)\cdot((1-y_1+x)(1-y_2+x)+y_1y_2)\\ &-(1-y_1+x)(1-y_2+x).\end{aligned}$$

Note that all monomials in $R(x, y_1, y_2)$ are multiples of $x^2y_1y_2$, $xy_1^2y_2$, $xy_1y_2^2$, $y_1y_2^4$ or $y_1^4y_2$, except for two negative terms (multiples of xy_1y_2 and $y_1^3y_2^3$). Thus, for $\beta_1, \beta_2 \ge$

$1/2$ and $0 < t < 1$, we can bound $R(t, t^{\beta_1}, t^{\beta_2})/t^{\sigma}$ from above by the sum C of all positive coefficients of monomials in R, where

$$\sigma = \max(2 + \beta_1 + \beta_2, 1 + 2\beta_1 + \beta_2, 1 + \beta_1 + 2\beta_2, 4\beta_1 + \beta_2, \beta_1 + 4\beta_2).$$

We can then bound the tail as before. □

Now we can use Lemma 6.15 to obtain explicit bounds on $\mathbf{h}_v(y_1, y_2)$ for y_1, y_2 large.

Proposition 6.17. *Let* $\mathbf{h}_v(y_1, y_2)$ *be as in (6.47), where* $v = 1$ *or* $v = 2$*. If* $y_1, y_2 \geq 10^{12}$*, then*

$$\mathbf{h}_v(y_1, y_2) \leq \begin{cases} \frac{0.000033536}{\log y_1 \log y_2} & \text{if } v = 1, \\ \frac{0.0000615022}{\log y_1 \log y_2} & \text{if } v = 2. \end{cases} \tag{6.78}$$

Proof. We simply have to apply Lemma 6.15 together with the estimates on values of k and κ in Lemmas 6.9 and 6.16.

We choose to work with $\epsilon = 1/14$. Then,

$$\kappa(1/2)\kappa(1/2 + \epsilon)\kappa(1/2, 1/2 + \epsilon) \leq 194.680303, \tag{6.79}$$

$$\frac{\kappa(1/2)\kappa(1-\alpha)}{389}\kappa(1/2, 1-\alpha) \leq \frac{16.889795}{389} \leq 0.0434185, \tag{6.80}$$

$$\frac{\kappa(1-\alpha)^2}{389^2}\kappa(1-\alpha, 1-\alpha) \leq \frac{4.549747}{389^2} \leq 0.00003006686 \tag{6.81}$$

for $\alpha = 1/\log 10^{12}$. For $y \geq 10^{12}$, $(\log y)/\sqrt{y} \leq 0.00002763103$ and $(\log y)/y^{1/2-\epsilon} \leq 0.00019885651$. We conclude, by Lemma 6.15, that

$$\begin{aligned} \mathbf{h}_1(y_1, y_2) &\leq \frac{0.00003006686}{\log y_1 \log y_2} + 2 \cdot \frac{0.0434185}{\log y_1 \log y_2} \cdot 0.00002763103 \\ &+ \frac{194.680303}{\log y_1 \log y_2} \cdot 0.00002763103 \cdot 0.00019885651 \\ &\leq \frac{0.000033536}{\log y_1 \log y_2}. \end{aligned} \tag{6.82}$$

To estimate $\mathbf{h}_2(y_1, y_2)$, notice that

$$\kappa_{\cdot 2}(1/2) \leq 1.891806\kappa(1/2), \qquad \kappa_{\cdot 2}(1/2 + \epsilon) \leq 1.813676\kappa(1/2 + \epsilon),$$
$$\kappa_{\cdot 2}(1-\alpha) \leq 1.519298\kappa(1-\alpha),$$

$$\kappa_{\backslash 2}(1/2, 1/2+\epsilon) \leq \frac{\kappa(1/2, 1/2+\epsilon)}{1.72564}, \qquad \kappa_{\backslash 2}(1/2, 1-\alpha) \leq \frac{\kappa(1/2, 1-\alpha)}{1.463112},$$
$$\kappa_{\backslash 2}(1-\alpha, 1-\alpha) \leq \frac{\kappa(1-\alpha, 1-\alpha)}{1.269669}.$$

Hence

$$\kappa_{\cdot 2}(1/2)\kappa_{\cdot 2}(1/2+\epsilon)\kappa_{\backslash 2}(1/2,1/2+\epsilon) \leq \frac{1.891806 \cdot 1.813676}{1.72564} \cdot 194.680303 \leq 387.08659, \tag{6.83}$$

$$\frac{\kappa_{\cdot 2}(1/2)\kappa_{\cdot 2}(1-\alpha)}{389}\kappa_{\backslash 2}(1/2,1-\alpha) \leq \frac{1.891806 \cdot 1.519298}{1.463112} \cdot 0.0434185 \leq 0.0852937, \tag{6.84}$$

$$\frac{\kappa_{\cdot 2}(1-\alpha)^2}{389^2}\kappa_{\backslash 2}(1-\alpha, 1-\alpha) \leq \frac{1.519298^2}{1.269669} \cdot 0.00003006686 \leq 0.00005466175 \tag{6.85}$$

and so

$$\begin{aligned} \mathbf{h}_2(y_1,y_2) &\leq \frac{0.00005466175}{\log y_1 \log y_2} + 2 \cdot \frac{0.0852937}{\log y_1 \log y_2} \cdot 0.00002763103 \\ &+ \frac{387.08659}{\log y_1 \log y_2} \cdot 0.00002763103 \cdot 0.00019885651 \\ &\leq \frac{0.0000615022}{\log y_1 \log y_2}. \end{aligned} \tag{6.86}$$

□

6.5.4 Bounds on $\mathbf{h}_v(y_1, y_2)$ for y_1 large or moderate and y_2 small

We also need to bound $\mathbf{h}_v(y_1, y_2)$ when y_1, y_2 are distinct, one of them is small, and the other one may be large or moderate.

We can obtain several useful bounds simply by using bounds for $\mathbf{h}_v(y, y)$ together with Cauchy-Schwarz.

Corollary 6.18. *Let $\mathbf{h}_v(y_1, y_2)$ be as in (6.47), where $v = 1$ or $v = 2$. If $0 < y_1 \leq 10^{12} \leq y_2$, then*

$$\mathbf{h}_v(y_1,y_2) \leq \begin{cases} \frac{0.05001}{\sqrt{y_1}\log y_2} & \text{if } v=1, \\ \frac{0.0953}{\sqrt{y_1}\log y_2} & \text{if } v=2 \end{cases} \tag{6.87}$$

and

$$\mathbf{h}_v(y_1,y_2) \leq \begin{cases} \frac{\sqrt{0.0001642+0.0001287\log y_1}}{\sqrt{y_1}\log y_2} & \text{if } v=1, \\ \frac{\sqrt{0.0005887+0.0003074\log y_1}}{\sqrt{y_1}\log y_2} & \text{if } v=2. \end{cases} \tag{6.88}$$

Proof. By Prop. 6.14, Prop. 6.17 and (6.50) (Cauchy-Schwarz). □

Lemma 6.19. *Let $\mathbf{h}_v(y_1, y_2)$ be as in (6.47), where $v = 1$ or $v = 2$. If $0 < y_1 \leq 10^6$, $y_2 \geq 10^{12}$, then*

$$\mathbf{h}_v(y_1,y_2) \leq \begin{cases} \frac{\min\left(\sqrt{0.000138+0.00002878\log y_1},0.017751\right)}{\sqrt{y_1}\log y_2} & \text{if } v=1, \\ \frac{\min\left(\sqrt{0.0004747+0.00010763\log y_1},0.033582\right)}{\sqrt{y_1}\log y_2} & \text{if } v=2. \end{cases} \tag{6.89}$$

Proof. By Prop. 6.17, (6.49) and Cauchy-Schwarz. □

The following bound can be stronger than what we would obtain from Cauchy-Schwarz, at least in the critical range of y_2 rather small.

Lemma 6.20. *Let $\mathbf{h}_v(y_1, y_2)$ be as in (6.47), where $v = 1$ or $v = 2$. Suppose $y_1 \leq 10^{12}$, $y_2 \leq 10^6$. Then $\mathbf{h}_v(y_1, y_2)$ is at most*

$$\begin{cases} \dfrac{0.06219}{y_1^{1/2}} + \dfrac{20.01991}{y_1^{3/4}} & \text{if } v = 1, \\ \dfrac{0.029216}{y_1^{1/2}} + \dfrac{40.61753}{y_1^{3/4}} & \text{if } v = 2. \end{cases} \tag{6.90}$$

Proof. By Cor. 6.5,

$$\left| \tilde{m}_{dv}\left(\frac{y_1}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2) \right| \leq \sum_{j=0}^{1} c_{v,j} \prod_{p|d} \frac{p+1}{p+1-p^{1-\beta_j}} \left(\frac{d}{y_1}\right)^{1-\beta_j},$$

where $\beta_0 = 1/2$, $\beta_1 = 1/4$, $c_{1,0} = 0.014429$, $c_{1,1} = 1.8$, $c_{2,0} = 0.006534$ and $c_{2,1} = 3.9877$. Also by Cor. 6.5, this time with $\alpha = 0$,

$$\left| \tilde{m}_{dv}\left(\frac{y_2}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2) \right| \leq \frac{\sigma(dv)}{dv}\zeta(2).$$

Hence,

$$\mathbf{h}_v(y_1, y_2) \leq c_{v,0}\frac{K_v(1/2)}{y_1^{1/2}} + c_{v,1}\frac{K_v(1/4)}{y_1^{3/4}},$$

where

$$K_v(\beta) = \zeta(2)\frac{\sigma(v)}{v} \sum_{\substack{d \\ (d,v)=1}} \frac{\mu^2(d)}{\sigma(d)^2} d^{1-\beta} \left(\prod_{p|d} \frac{p+1}{p+1-p^{1-\beta}} \right) \frac{\sigma(d)}{d} \tag{6.91}$$

for $\beta = 1/2$ and $\beta = 1/4$. By Lemma 6.11,

$$K_v(\beta) \leq \zeta(2)\frac{\sigma(v)}{v}\kappa_{\setminus v}(\beta, 1),$$

where $\kappa_{\setminus v}(\beta, 1) = \kappa_{\setminus v}(1, \beta)$ is as in (6.48). Thus, by Lemma 6.16,

$$\begin{aligned} c_{1,0}K_1(1/2) &\leq 0.014429 \cdot \zeta(2) \cdot 2.620185 \leq 0.06219 \\ c_{1,1}K_1(1/4) &\leq 1.8 \cdot \zeta(2) \cdot 6.761469 \leq 20.01991, \\ c_{2,0}K_2(1/2) &\leq 0.006534 \cdot \frac{3}{2}\zeta(2) \cdot \frac{2.620185}{1 + \frac{2^{-3/2}}{1-2^{-1/2}+1/2}} \leq 0.029216 \\ c_{2,1}K_2(1/4) &\leq 3.9877 \cdot \frac{3}{2}\zeta(2) \cdot \frac{6.761469}{1 + \frac{2^{-5/4}}{1-2^{-1/4}+1/2}} \leq 40.61753. \end{aligned}$$

□

6.6 COMPUTING h_v FOR SMALL ARGUMENTS

In the previous section, we proved bounds on $h_v(y_1, y_2)$. For y_1, y_2 large, our bounds were quite satisfactory, and, at any rate, we do not know how to do better. We should now prove sharper bounds on $h_v(y_1, y_2)$ for y_1, y_2 small ($\leq 10^6$) and on $h_v(y, y)$ for y moderate ($\leq 10^{10}$).

Before we start, we should explain why we will be working with h_v directly, and not with g_v, which is a finite sum, and thus would seem simpler to compute. The reason is the following common computational issue.

If two quantities a, b differ by a small amount $\delta = b - a \sim 2^{-d}$, and we have determined a and b with c bits of precision, then we obtain $\delta = b - a$ with only $c - d$ bits of precision. It is thus a good idea to start by taking out main terms from our expressions. Now, h_v is precisely g_v with main terms taken out. We took the main terms out to make our asymptotic analysis easier, but it so happens that we also prepared the ground for greater precision in our computations thereby.

The same computational issue implies, for instance, that it is wise to apply an implementation of $\log(1 + x)$ as a function of x small, rather than use $\log x$ for $x \sim 1$. Actually, the same principle applies to our work in §6.2. As we already saw in Lemma 6.4, it makes sense to see $(\sigma(dv)/dv)\zeta(2)$ as the main term of $\tilde{m}_{dv}(y)$ for dv square-free. Thus, when we compute $\tilde{m}_{dv}(n)$ for $n = 1, 2, 3, \ldots$ – and we are going to carry out additional computations of that kind in this section – we should actually compute and store $\tilde{m}_{dv}(n) - (\sigma(dv)/dv)\zeta(2)$, so as to obtain more precise results.

6.6.1 Computing $h_v(y_1, y_2)$ for y_1, y_2 small

Our first task is to compute $h_v(y_1, y_2)$ for all $y_1, y_2 \leq Y$ (say) in time $O\left(Y^2\right)$. First of all, we should make it clear that the task can be accomplished in finite time. While $h_v(y_1, y_2)$ does not depend on $\lfloor y_1 \rfloor$, $\lfloor y_2 \rfloor$ alone, it equals an expression of the form

$$\sum_{j_1=0}^{1} \sum_{j_2=0}^{1} \kappa_{j_1,j_2,v}(\lfloor y_1 \rfloor, \lfloor y_2 \rfloor) \left(\log \frac{y_1}{\lfloor y_1 \rfloor}\right)^{j_1} \left(\log \frac{y_2}{\lfloor y_2 \rfloor}\right)^{j_2}, \tag{6.92}$$

for $y_1, y_2 \geq 1$, as can be seen from its definition (6.43). Indeed, it must equal

$$\sum_{j_1=0}^{1} \sum_{j_2=0}^{1} \kappa_{j_1,j_2,v}(n_1, n_2) \left(\log \frac{y_1}{n_1}\right)^{j_1} \left(\log \frac{y_2}{n_2}\right)^{j_2}, \tag{6.93}$$

where $n_1 = \lfloor y_1 \rfloor$, $n_2 = \lfloor y_2 \rfloor$ if $v = 1$, and, if $v = 2$, n_i is the largest odd number $\leq y_i$, for $i = 1, 2$. (The meaning of $\kappa_{j_1,j_2,v}$ is the same in (6.92) and (6.93).) Our task is to compute $\kappa_{j_1,j_2,v}(n_1, n_2)$ for $n_1, n_2 \leq Y$ and $j_1, j_2 = 0, 1$.

Since $\log^+$ is a continuous function, the dependence of $h_v(y_1, y_2)$ on y_1 is continuous. Hence, for any n_1, n_2, $\lim_{y_1 \to n_1^-} h_v(y_1, n_2) = h_v(n_1, n_2)$. For n_1, n_2 coprime to v, using (6.93), we conclude that

$$\kappa_{0,0,v}(n_1 + v, n_2) = \kappa_{0,0,v}(n_1, n_2) + \kappa_{1,0,v}(n_1, n_2) \log \frac{n_1 + v}{n_1}.$$

In fact, $\lim_{y_1\to n_1^-} h_v(y_1,y_2) = h_v(n_1,y_2)$ for every $n_2 \leq y_2 < n_2+v$, and so what we have is a linear function on $\log y_2$ converging to another linear function on $\log y_2$ (namely, $h_v(n_1,y_2)$). The coefficients of the first linear function must then converge to the coefficients of the second linear function; in other words,

$$\kappa_{0,j_2,v}(n_1+v,n_2) = \kappa_{0,j_2,v}(n_1,n_2) + \kappa_{1,j_2,v}(n_1,n_2)\log\frac{n_1+v}{n_1}, \tag{6.94}$$

for $j_2 = 0,1$.

Proceeding in exactly the same way with the roles of y_1 and y_2 switched, we obtain that

$$\kappa_{j_1,0,v}(n_1,n_2+v) = \kappa_{j_1,0,v}(n_1,n_2) + \kappa_{j_1,1,v}(n_1,n_2)\log\frac{n_2+v}{n_2} \tag{6.95}$$

for $j_2 = 1$, where, just as in (6.94), $v\in\{1,2\}$ and n_1 and n_2 coprime to v.

By (6.43),

$$\begin{aligned}\kappa_{0,0,v}(1,1) &= \sum_{\substack{d\\(d,v)=1}} \frac{\mu(d)}{\sigma(d)^2}\left(\frac{\sigma(dv)}{dv}\zeta(2)\right)^2\\ &= \left(\frac{\sigma(v)}{v}\zeta(2)\right)^2 \sum_{\substack{d\\(d,v)=1}} \frac{\mu(d)}{d^2} = \frac{\sigma(v)}{\phi(v)}\zeta(2).\end{aligned} \tag{6.96}$$

Applying (7.28) backwards with $\varrho = \log^+$, we see that

$$g_v(y_1,y_2) = \sum_{\substack{r_1,r_2\\(r_1,r_2)=1\\(r_1r_2,v)=1}} \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)}\left(\log^+\frac{y_1}{r_1}\right)\left(\log^+\frac{y_2}{r_2}\right) \tag{6.97}$$

and so, by (6.43) or Lemma 6.10,

$$\kappa_{1,1,v}(n_1,n_2) = \sum_{\substack{r_1\leq n_1\\(r_1,r_2)=1\\(r_1r_2,v)=1}}\sum_{r_2\leq n_2} \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)} \tag{6.98}$$

for $n_1,n_2\geq 1$. Again by Lemma 6.10, we see that, for $n\geq 1$,

$$\kappa_{1,0,v}(n,1) = \kappa_{0,1,v}(1,n) = -\frac{\sigma(v)}{v}\zeta(2)\sum_{\substack{d\leq n\\(d,v)=1}}\frac{\mu(d)}{d}. \tag{6.99}$$

(It is clear from (6.97) that, for $1\leq y_2 < 1+v$, all terms are proportional to $\log y_2$ or $(\log y_1)(\log y_2)$, and thus the term $\kappa_{1,0,v}(n,1)\cdot\log y_1$ in $h_v(y_1,y_2)$ and the term proportional to $\log y_1$ in the last line of (6.45) must cancel out. Exactly the same happens with $\kappa_{0,1,v}(1,n)$, with y_1 and y_2 switched.)

Recurrence relations (6.94) and (6.95), together with the initial values (so to speak) that we determined in (6.96) and (6.99), reduce the problem of determining the coefficients $\kappa_{j_1,j_2,v}(n_1,n_2)$, $j_1, j_2 = 0, 1$, for all $n_1, n_2 \leq Y$, to the problem of determining $\kappa_{1,1,v}(n_1,n_2)$ for all $n_1, n_2 \leq Y$. It remains to see how to compute $\kappa_{1,1,v}(n_1,n_2)$ quickly.

A naïve approach would consist in computing the sum (6.98) for each pair (n_1,n_2), thus taking time at least $O(n_1 n_2)$ for each such pair. It is easy to do much better. We shall use the following procedure with $a_{r_1,r_2} = \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)}$ for r_1, r_2 with $(r_1,r_2) = 1$, $(r_1 r_2, v) = 1$ and $a_{r_1,r_2} = 0$ otherwise.

Lemma 6.21. *Let $a_{r_1,r_2} \in \mathbb{C}$ for $1 \leq r_1 \leq n_1$, $1 \leq r_2 \leq n_2$. Then*

$$s(n_1,n_2) = \sum_{r_1 \leq n_1} \sum_{r_2 \leq n_2} a_{r_1,r_2}$$

can be computed for all $n_1 \leq N_1$, $n_2 \leq N_2$, in time $O(N_1 N_2)$ and space $O(N_1)$.

The values of $s(n_1,n_2)$ are output in the following order:

$$s(1,1), s(2,1), \ldots, s(N_1,1), s(1,2), \ldots, s(N_1,2), \ldots, s(N_1,N_2).$$

It is understood that a_{r_1,r_2} is not necessarily given by an array, but may be given by a procedure call. Then the time $O(N_1 N_2)$ is multiplied by the time the procedure takes. In our case, we can compute $\mu(r)$, $\sigma(r)$ at the beginning; the time then taken to compute a_{r_1,r_2} is essentially constant.

Proof. We compute $s(n_1,1)$ for $1 \leq n_1 \leq N_1$ by $s(n_1,1) = s(n_1-1,1) + a_{n_1,1}$. We store the result ($b_{n_1} \leftarrow s(n_1,1)$) and output it.

Now, for $n \geq 2$,

$$s(n_1,n) = s(n_1,n-1) + \sum_{r_1 \leq n_1} a_{r_1,n_2}.$$

We can of course compute $d_{n_1,n_2} = \sum_{r_1 \leq n_1} a_{r_1,n_2}$ for n_2 fixed and successive n_1 by $d_{n_1,n_2} = d_{n_1-1,n_2} + a_{n_1,n_2}$. So, we go over all $1 \leq n_1 \leq N_1$, $1 \leq n_2 \leq N_2$, with n_2 in the outer loop; that is to say, we go over the pairs in the order

$$(1,1), (1,2), \ldots, (1,n_2), (2,1), (2,2), \ldots, (3,1), (3,2), \ldots, (n_1,n_2).$$

We compute (and store and output) $s(n_1,n_2)$ for $1 \leq n_1 \leq N_1$ and n_2 fixed by $b_{n_1} \leftarrow b_{n_1} + d_{n_1,n_2}$. □

Corollary 6.22. *Let $h_v(y_1,y_2)$ be as in (6.43). Let $v = 1$ or $v = 2$. We can compute $h_v(y_1,y_2)$ for all $y_1, y_2 \leq Y$ in time $O(Y^2)$ and space $O(Y)$.*

Proof. Compute $\mu(r)$, $\sigma(r)$ by means of a sieve of Eratosthenes (§5.3.2), and apply Lemma 6.21. □

We can thus easily verify the following statement. It will later become clear why it is useful to put our bounds on $h_v(x,y)$ in such a way.

Corollary 6.23. *Let $h_v(y_1,y_2)$ be as in (6.43). Let $v=1$ or $v=2$. For $r\geq 1$, $1\leq x\leq 10^6/r$, let*

$$f_{r,v}(x)=\inf_{rx\leq y\leq 10^6}\sqrt{xy}\cdot h_v(x,y).$$

Then

$$f_{r,v}(x)\geq -F_{r,v}(x),$$

where $F_{r,v}:[0,\infty)\to\mathbb{R}$ is a continuous, non-decreasing, non-negative function of compact support, satisfying

$$\int_0^\infty \frac{F_{r,v}(x)}{x^2}dx\leq\begin{cases}0.86894 & \text{if } v=1,\\ 1.03489 & \text{if } v=2,\end{cases}\tag{6.100}$$

and, for $v=2$,

$$\max_{x\geq 1}\frac{F_{r,v}(x)}{x}\leq 0.43102.\tag{6.101}$$

Moreover, when $v=1$ and $r\geq 5$, 0.86894 can be replaced by 0.74957.

The function $f_{r,v}$ is far from being non-increasing. We will simply find a minorant such as $-F_{r,v}$ useful for computing an error term later. The loss incurred in using such a majorant is around 10 percent for $v=1$, and less than 3 percent for $v=2$.

Proof. By a computation as in Corollary 6.22, with $Y=10^6$. We simply let

$$F_{r,v}(x)=\max_{t\leq x}\max\left(-f_{r,v}(t),0\right).$$

The basic procedure to follow is clear: for given x, we bound $f_{r,v}(x,y)$ by computing $\sqrt{xy}\cdot h_v(x,y)$ over many points $y\in[rx,10^6]$; then, to approximate $F_{r,v}(x)$ for all $x\in[1,10^6/r]$, we compute $f_{r,v}(x,y)$ at many points $x\in[1,10^6/r]$. The question now is how to choose the values of x and y at which to carry out such computations, and what to do so that taking samples in this fashion results in actual bounds, rather than mere empirical estimates.

Interval arithmetic supplies an easy answer: if we evaluate our expression (6.93) by inputing into it *intervals* y_1, y_2 contained in the intervals $[n_1,n_1+v]$ and $[n_2,n_2+v]$ (say), then what interval arithmetic provides is a lower bound (and an upper bound) on $h_v(x,y)$ for all $x\in y_1$, $y\in y_2$.

For $n\leq 5000$, we subdivide each segment of the form $[n,n+v]$, $(n,v)=1$, into $\sim 10000/n$ intervals; for $n>5000$, we simply take $[n,n+v]$, $(n,v)=1$ as our intervals, since (as it turns out) we need less precision for n large. We use the same subdivision into intervals for the variable x and the variable y.

Proceeding in this way, we obtain lower bounds on $f_{r,v}(x)$ and, thus, upper bounds on $F_{r,v}(x)$. The bounds on $F_{r,v}(x)$ satisfy (6.100) and (6.101). ☐

6.6.2 Computing $h_v(y)$ for y small or moderate

We would like to have more precise estimates on $h_v(y) = h_v(y,y)$ for y below a certain bound.

Our task is to compute $h_v(y)$ for all $y \leq Y$ (say) in time roughly linear on Y, that is, in particular, without having to compute $h_v(y_1, y_2)$ for all $y_1, y_2 \leq Y$. We would also like to compute $h_v'(y)$ for all non-integer $y \leq Y$ in time roughly linear on Y. (If y is an integer, h_v may not be differentiable at y, as $\log^+$ is not differentiable at 1.)

By (6.92), $h_v(y)$ equals an expression of the form

$$h_v(y) = \sum_{j=0}^{2} \kappa_{j,v}(n) \left(\log \frac{y}{n}\right)^j, \tag{6.102}$$

where n is the largest integer $\leq y$ coprime to v. Once we determine $\kappa_{j,v}(n)$, $j = 0, 1, 2$, we will have determined not just $h_v(y)$, but also $h_v'(y)$, since

$$h_v'(y) = \frac{\kappa_{1,v}(\lfloor y \rfloor)}{y} + \frac{2\kappa_{2,v}(\lfloor y \rfloor)}{y} \log \frac{y}{\lfloor y \rfloor} \tag{6.103}$$

for y not an integer. It is easy to tell that

$$\kappa_{0,v}(n) = \kappa_{0,0,v}(n,n), \qquad \kappa_{1,v}(n) = \kappa_{0,1,v}(n,n) + \kappa_{1,0,v}(n,n),$$
$$\kappa_{2,v}(n) = \kappa_{1,1,v}(n,n).$$

Much as in (6.94)–(6.95), since $\log^+$ is continuous, the dependence of $h_v(y)$ on y is continuous, and so, for $n \geq 1$ coprime to v,

$$\kappa_{0,v}(n+v) = \kappa_{0,v}(n) + \kappa_{1,v}(n) \log \frac{n+v}{n} + \kappa_{2,v}(n) \left(\log \frac{n+v}{n}\right)^2. \tag{6.104}$$

Comparing (6.102) with (6.92), we see that

$$\kappa_{0,v}(n) = \kappa_{0,0,v}(n,n), \qquad \kappa_{1,v}(n) = \kappa_{1,0,v}(n,n) + \kappa_{0,1,v}(n,n),$$
$$\kappa_{2,v}(n) = \kappa_{1,1,v}(n,n).$$

Thus, in particular, by (6.96),

$$\kappa_{0,v}(1) = \frac{\sigma(v)}{\phi(v)} \zeta(2). \tag{6.105}$$

It remains to determine $\kappa_{1,v}(n)$ and $\kappa_{2,v}(n)$ for $n \leq Y$. By (6.98),

$$\kappa_{2,v}(n) = \sum_{\substack{r_1 \leq n \\ (r_1,r_2)=1 \\ (r_1 r_2, v)=1}} \sum_{r_2 \leq n} \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)} \tag{6.106}$$

for $n \geq 1$. By (6.97) and Lemma 6.10,

$$\kappa_{1,v}(n) = 2 \sum_{\substack{r_1 \leq n \\ (r_1,r_2)=1 \\ (r_1 r_2, v)=1}} \sum_{r_2 \leq n} \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)} \log \frac{n}{r_1} - 2\frac{\sigma(v)}{v}\zeta(2) m_v(n). \tag{6.107}$$

It is plausible to say that the procedure we will follow to compute $\kappa_{1,v}(n)$ and $\kappa_{2,v}(n)$ is a simple case of what is called "dynamic programming" (an often-used and not particularly well-defined term in computer science). What this means for is essentially that we will be keeping and continually updating the results of intermediate computations. This involves significant memory usage; we may say that we trade space for time.

In our case, we can keep memory usage very moderate by some number-theoretical tricks. Let us first see how well can we do if we do not use any such tricks.

Lemma 6.24. *Let $h_v(y) = h_v(y, y)$ be as in (6.43). Let $v = 1$ or $v = 2$. We can compute $h_v(y, y)$ for all $y \leq Y$ in time $O(Y \log Y)$ and space $O(Y)$.*

Proof. Our task is to compute $\kappa_{j,v}(n)$, $j = 0, 1, 2$, for $n \leq Y$. Starting from (6.106), and proceeding as in (7.28), we see that

$$\kappa_{2,v}(n) = \sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{(\sigma(d))^2} \check{m}_{dv} \left(\frac{n}{d}\right)^2,$$

where

$$\check{m}_w(y) = \sum_{\substack{r \leq y \\ (r,w)=1}} \frac{\mu(r)}{\sigma(r)}. \tag{6.108}$$

(In fact, this is literally an application of (7.28), with ϱ set equal to the brutal truncation $\varrho(x) = 1$ for $x \geq 1$, $\varrho(x) = 0$ for $x < 1$.) We can see that $\check{m}_w(y)$ depends only on w and $\lfloor y \rfloor$. Hence

$$\kappa_{2,v}(n) - \kappa_{2,v}(n-1) = \sum_{\substack{d|n \\ (d,v)=1}} \frac{\mu(d)}{(\sigma(d))^2} \left(\check{m}_{dv} \left(\frac{n}{d}\right)^2 - \check{m}_{dv} \left(\frac{n}{d} - 1\right)^2 \right). \tag{6.109}$$

Both sides vanish unless n is square-free and coprime to v. (If n is square-free, then $\check{m}_{dv}(n/d) = \check{m}_{dv}(n/d - 1)$ whether n/d is square-free or not, for a different reason in each case.)

Let n be square-free and coprime to v; let d be a divisor of n. Then, obviously, $\check{m}_{dv}(n/d) = \check{m}_{dv}(n/d - 1)\mu(n/d)/\sigma(n/d)$, and so

$$\check{m}_{dv} \left(\frac{n}{d}\right)^2 - \check{m}_{dv} \left(\frac{n}{d} - 1\right)^2 = \frac{1}{\sigma(n/d)^2} + 2 \cdot \frac{\mu(n/d)}{\sigma(n/d)} \check{m}_{dv} \left(\frac{n}{d} - 1\right). \tag{6.110}$$

Hence, by (6.109), for n square-free and coprime to v,

$$\begin{aligned}\kappa_{2,v}(n) - \kappa_{2,v}(n-1) &= \sum_{d|n} \frac{\mu(d)}{\sigma(n)^2} + \sum_{d|n} \frac{2\mu(d)\mu(n/d)}{\sigma(d)^2\sigma(n/d)} \dot{m}_{dv}\left(\frac{n}{d}-1\right) \\ &= \delta_{n,1} + \frac{2\mu(n)}{\sigma(n)} \sum_{d|n} \frac{1}{\sigma(d)} \dot{m}_{dv}\left(\frac{n}{d}-1\right),\end{aligned} \tag{6.111}$$

where, we recall, $\delta_{x,y}$ denotes the Kronecker delta: here, $\delta_{n,1} = 1$ if $n = 1$ and $\delta_{n,1} = 0$ otherwise.

Similarly,

$$\kappa_{1,v}(n) = 2 \sum_{\substack{d \\ (d,v)=1}} \frac{\mu(d)}{(\sigma(d))^2} \dot{m}_{dv}\left(\frac{n}{d}\right)\left(\tilde{m}_{dv}\left(\frac{n}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right),$$

and, since $\tilde{m}_{dv}(n/d) = \tilde{m}_{dv}((n-1)/d) + \dot{m}_{dv}((n-1)/d)\log(n/(n-1))$, we see that

$$\kappa_{1,v}(n) - \kappa_{1,v}(n-1) = 2\kappa_{2,v}(n-1)\log\frac{n}{n-1} \tag{6.112}$$

if n is not coprime to v, whereas, if n is coprime to v, $\kappa_{1,v}(n) - \kappa_{1,v}(n-1)$ equals $2\kappa_{2,v}(n-1)\log\frac{n}{n-1}$ plus

$$\begin{aligned}&2\sum_{d|n} \frac{\mu(d)}{(\sigma(d))^2}\frac{\mu(n/d)}{\sigma(n/d)}\tilde{m}_{dv}\left(\frac{n}{d}\right) - 2\sum_{d|n}\frac{\mu(d)}{(\sigma(d))^2}\frac{\mu(n/d)}{\sigma(n/d)}\frac{\sigma(dv)}{dv}\zeta(2) \\ =&2\frac{\mu(n)}{\sigma(n)}\sum_{d|n}\frac{1}{\sigma(d)}\left(\tilde{m}_{dv}\left(\frac{n}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right).\end{aligned} \tag{6.113}$$

Clearly,

$$\tilde{m}_w(t) = \tilde{m}_w(r) + \dot{m}_w(r)\log\frac{t}{r} \tag{6.114}$$

for all $t \in [r, r+1]$, r an integer; thus we can (a) determine $\tilde{m}_w(t)$ for t not an integer and (b) determine $\tilde{m}_w(r+1)$, given $\tilde{m}_w(r)$.

We compute $\kappa_{j,v}(n)$, $j = 1, 2$, by using (6.111) and (6.113) repeatedly. (To compute $\kappa_{0,v}(n)$, it suffices to use (6.104) and (6.105).) What we will do, then, is keep $\dot{m}_{dv}(n/d) = \dot{m}_{dv}(\lfloor n/d \rfloor)$ and $\tilde{m}_{dv}(\lfloor n/d \rfloor)$ in storage for all $d \leq n$. Every time we increase n by 1, we update $\dot{m}_{dv}(\lfloor n/d \rfloor)$ and $\tilde{m}_{dv}(\lfloor n/d \rfloor)$ for all $d|n$. (Their values for $d \nmid n$ do not change.) Since the average number of divisors d of an integer n is $\sim \log n$, the times it takes to compute $\kappa_{j,v}(n) - \kappa_{j,v}(n-1)$ in this way is, on the average, proportional to $(\log n)$, assuming that $\mu(n)$, $\sigma(n)$, $\sigma(n/d)$, and the list of divisors of n have been all computed in one go or can be computed quickly.

We can compute and store $\mu(m)$, $\sigma(m)$ and all prime divisors of m for $m \leq Y$ in time $O(Y \log Y)$ and space $O(Y \log\log Y)$. (The average number of prime divisors of a number $m \leq Y$ is $\sim \sum_{p \leq Y} 1/p \sim \log\log Y$.) A moment's thought shows that we need only know the factorization of the integer n we are working with at any given moment, in that the expressions (6.111), (6.113) involve $\sigma(d)$ and $\mu(d)$ only for divisors d of n, and such quantities can be computed quickly given the factorization of n. We can use a segmented sieve (§4.5) to factor our integers n, and use space $O(\sqrt{Y}\log\log Y)$ to store factorizations of integers n in intervals of length $\sqrt{Y}$. Thus, the lion's share of space is taken by our keeping $\check{m}_{dv}(n/d)$ and $\tilde{m}_{dv}(\lfloor n/d \rfloor)$ in storage for all $d \leq n$. Since n goes up to Y, total space consumption is $O(Y)$. □

What we meant by "dynamic programming" should now be clear: we are constantly updating stored information.

It is in our interest to reduce space further, particularly because space usage in the proof we have just seen is not sequential. In practice, a program that uses a large amount of memory non-sequentially – that is, not going through storage in some sort of simple order – will be slowed down considerably. Let us now think like number theorists so as to save on space.

Corollary 6.25. *Let $h_v(y) = h_v(y, y)$ be as in (6.43). Let $v = 1$ or $v = 2$. We can compute $h_v(y)$ for all $y \leq Y$ in time $O(Y \log Y)$ and space $O(\sqrt{Y}\log\log Y)$.*

Proof. The issue is how to decrease the space consumption of the procedure in the proof of Lemma 6.24. As we already saw, using a segmented sieve of Eratosthenes takes care of only part of the issue: yes, we can factorize our increasing variable n in this fashion, but we have also been keeping information associated to each $d \leq n$. Let us manipulate our expressions so that we will need to keep information for far fewer d.

Assume $1 < n \leq Y$. Then, by (6.111), when n is square-free and coprime to v, the difference $\kappa_{2,v}(n) - \kappa_{2,v}(n-1)$ equals

$$\frac{2\mu(n)}{\sigma(n)} \left(\sum_{\substack{d|n \\ d \leq \sqrt{Y}}} \frac{1}{\sigma(d)} \check{m}_{dv}\left(\frac{n}{d} - 1\right) + \sum_{\substack{d|n \\ d < n/\sqrt{Y}}} \frac{\sigma(d)}{\sigma(n)} \check{m}_{nv/d}(d-1) \right),$$

where we are changing variables by $d \mapsto n/d$ in the second sum. By (6.108),

$$\begin{aligned} \check{m}_{nv/d}(d-1) &= \sum_{\substack{r \leq d-1 \\ (r,v)=1}} \frac{\mu(r)}{\sigma(r)} \sum_{\ell | \left(r, \frac{n}{d}\right)} \mu(\ell) \\ &= \sum_{\ell | \frac{n}{d}} \frac{\mu(\ell)^2}{\sigma(\ell)} \sum_{\substack{r \leq \frac{d-1}{\ell} \\ (r,\ell v)=1}} \frac{\mu(r)}{\sigma(r)} = \sum_{\ell | \frac{n}{d}} \frac{\mu(\ell)^2}{\sigma(\ell)} \check{m}_{\ell v}\left(\frac{d-1}{\ell}\right). \end{aligned} \tag{6.115}$$

Clearly $\check{m}_{\ell v}((d-1)/\ell) = 0$ for $\ell \geq \sqrt{Y}$, since then $\ell > d - 1$.

Similarly, by (6.113), for $1 < n \leq Y$ square-free and coprime to v, the difference

$\kappa_{1,v}(n) - \kappa_{1,v}(n-1)$ equals $2\kappa_{2,v}(n-1)\log(n/(n-1))$ plus

$$\begin{aligned} &2\frac{\mu(n)}{\sigma(n)} \sum_{\substack{d|n \\ d\leq\sqrt{Y}}} \frac{1}{\sigma(d)} \left(\tilde{m}_{dv}\left(\frac{n}{d}\right) - \frac{\sigma(dv)}{dv}\zeta(2)\right) \\ &+2\frac{\mu(n)}{\sigma(n)} \sum_{\substack{d|n \\ d<n/\sqrt{Y}}} \frac{\sigma(d)}{\sigma(n)} \left(\tilde{m}_{nv/d}(d) - \frac{\sigma(nv/d)}{nv/d}\zeta(2)\right). \end{aligned} \tag{6.116}$$

Just as in (6.115), for n coprime to v,

$$\begin{aligned} \tilde{m}_{nv/d}(d) &= \sum_{\substack{r\leq d \\ (r,v)=1}} \frac{\mu(r)}{\sigma(r)} \log\frac{d}{r} \sum_{\ell|(r,\frac{n}{d})} \mu(\ell) \\ &= \sum_{\ell|\frac{n}{d}} \frac{\mu(\ell)^2}{\sigma(\ell)} \sum_{\substack{r\leq d/\ell \\ (r,\ell v)=1}} \frac{\mu(r)}{\sigma(r)} \log\frac{d}{\ell r} = \sum_{\ell|\frac{n}{d}} \frac{\mu(\ell)^2}{\sigma(\ell)} \tilde{m}_{\ell v}(d/\ell), \end{aligned} \tag{6.117}$$

and so

$$\tilde{m}_{nv/d}(d) - \frac{\sigma(nv/d)}{nv/d}\zeta(2) = \sum_{\ell|\frac{n}{d}} \frac{\mu(\ell)^2}{\sigma(\ell)} \left(\tilde{m}_{\ell v}(d/\ell) - \frac{\sigma(\ell v)}{\ell v}\zeta(2)\right).$$

Again, it is clear that $\tilde{m}_{\ell v}(d/\ell) = 0$ for $\ell \geq \sqrt{Y}$, since then $\ell > d$.

Thus, in order to verify the proposition for all $x \leq Y$, it is enough to do the following:

1. Start by computing $\mu(r)$, $\sigma(r)$ for all $r \leq \sqrt{Y}$ by a sieve of Eratosthenes.
2. Still at the beginning of the procedure, compute and store $\check{m}_{\ell v}(m)$ and $\tilde{m}_{\ell v}(m) - (\sigma(\ell v)/\ell v)\zeta(2)$ for all $\ell < \sqrt{Y}$, $m < \sqrt{Y}/\ell$. This step takes space and time $O(\sqrt{Y}\log Y)$.
3. Now, as n goes from 1 to Y, factor n. This step takes time $O(Y\log\log Y)$ and space $O(\sqrt{Y}\log\log Y)$, in that we split $[1, Y]$ into intervals of length about $\sqrt{Y}$ and apply the sieve of Eratosthenes to each of them.
4. Keep track of $\check{m}_{dv}(n/d)$ and $\tilde{m}_{dv}(n/d) - (\sigma(dv)/dv)\zeta(2)$, $d \leq \sqrt{Y}$, as n increases. This takes takes space $O(\sqrt{Y})$ and time $O(Y\log Y)$, since, as we already saw, we need to update $h_{dv}(n/d)$ only for d dividing n and n square-free. We can compute $\mu(n/d)$ and $\sigma(n/d)$ immediately from $\mu(n)$, $\sigma(n)$, $\mu(d)$ and $\sigma(d)$, and compute $\mu(n)$, $\sigma(n)$ quickly using the factorization of n.
5. As n increases, keep computing $\kappa_{j,v}(n) - \kappa_{j,v}(n-1)$ for $j = 0, 1, 2$.

Steps (3) and (4) are to be carried out together rather than consecutively. In total, the procedure takes time $O(Y\log Y)$ and space $O(\sqrt{Y}\log\log Y)$. □

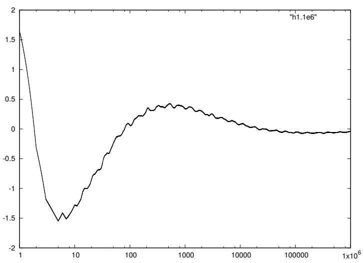

Figure 6.1: $th_1(t)$ on $[1, 10^6]$

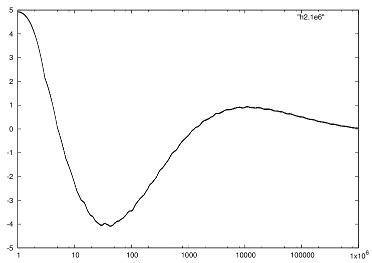

Figure 6.2: $th_2(t)$ on $[1, 10^6]$

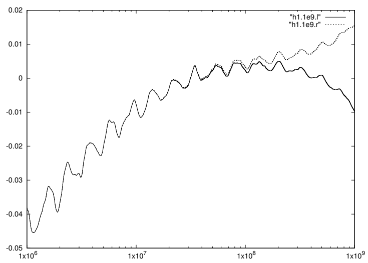

Figure 6.3: $th_1(t)$ on $[10^6, 10^9]$

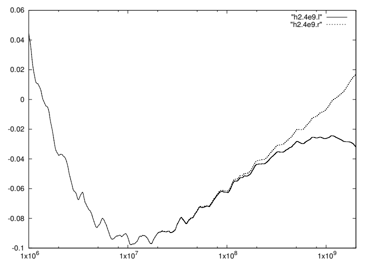

Figure 6.4: $th_2(t)$ on $[10^6, 2 \cdot 10^9]$

Proposition 6.26. *Let $h_v(y) = h_v(y, y)$, where $h_v(y, y)$ is as in (6.43) and $v = 1$ or $v = 2$. Then*

$$\int_1^{10^6} h_v(t)\frac{dt}{t} \leq \begin{cases} -0.049510004 & \text{if } v = 1 \\ 2.634812714 & \text{if } v = 2. \end{cases} \tag{6.118}$$

and, for $10^6 \leq t \leq 10^9$,

$$|h_v(t)| \leq \begin{cases} \frac{0.0455}{t} & \text{if } v = 1, \\ \frac{0.0978}{t} & \text{if } v = 2. \end{cases} \tag{6.119}$$

The algorithm is good enough that time is not the main issue, in comparison with precision loss; going up to $T \sim 10^{10}$ takes a day or two on a single processor core. As can be seen in Figures 6.3 and 6.4, precision degenerates visibly starting at about $t \sim 10^8$. The higher and lower curve in each graph both depict $h_v(t)$: they represent the upper and lower bounds on $h_v(t)$ given by interval arithmetic based on IEEE double precision (64 bits).

Proof. We apply the algorithm in the proof of Cor. 6.25. Note that $\int_1^T h_v(t)dt/t$ equals

$$\sum_{n=1}^{n_0-1} \kappa_{0,v}(n) \log \frac{n+v}{n} + \frac{\kappa_{1,v}(n)}{2} \left(\log \frac{n+v}{n}\right)^2 + \frac{\kappa_{2,v}(n)}{3} \left(\log \frac{n+v}{n}\right)^3$$
$$+ \kappa_{0,v}(n_0) \log \frac{T}{n_0} + \frac{\kappa_{1,v}(n_0)}{2} \left(\log \frac{T}{n_0}\right)^2 + \frac{\kappa_{2,v}(n_0)}{3} \left(\log \frac{T}{n_0}\right)^3, \tag{6.120}$$

where n_0 is the largest integer $\le T$ coprime to v. Again, it is important to use an implementation of $\log(1+x)$, rather than of $\log x$.

We obtain, in fact, that

$$\int_1^{10^6} h_v(t) \frac{dt}{t} \in \begin{cases} [-0.04951001463, -0.04951000438] & \text{if } v = 1 \\ [2.63481249177, 2.63481271383] & \text{if } v = 2. \end{cases} \tag{6.121}$$

□

We will use the following estimates once later, in order to treat a small error term. Thus, somewhat rough bounds are enough.

Corollary 6.27. *Let $h_v(y) = h_v(y, y)$, where $h_v(y, y)$ is as in (6.43) and $v = 1$ or $v = 2$. Then, for all $1 \le T \le 10^9$,*

$$\int_1^T \left| \frac{d}{dt} t h_v(t) \right| dt \le \begin{cases} 7.035 & \text{if } v = 1, \\ 17.61 & \text{if } v = 2, \end{cases} \tag{6.122}$$

and

$$\int_1^T \left| \frac{d}{dt} t h_v(t) \right| dt + |T h_v(T)| \le \begin{cases} 7.05 & \text{if } v = 1, \\ 17.64 & \text{if } v = 2. \end{cases} \tag{6.123}$$

Proof. The function $t \mapsto t h_v(t)$ is continuous everywhere and differentiable outside the integers. For $n < t < n + v$, where $(n, v) = 1$,

$$\frac{d}{dt} t h_v(t) = h_v(t) + t h_v'(t)$$
$$= (\kappa_{0,v} + \kappa_{1,v})(\lfloor t \rfloor) + (\kappa_{1,v} + 2\kappa_{2,v})(\lfloor t \rfloor) \log \frac{t}{\lfloor t \rfloor} + \kappa_{2,v}(\lfloor t \rfloor) \left(\log \frac{t}{\lfloor t \rfloor}\right)^2, \tag{6.124}$$

We evaluate this expression with integer arithmetic, replacing $t/\lfloor t \rfloor$ by $1 + v/n$. Call the resulting interval I.

If I does not contain 0, then $(t h_v(t))'$ does not change sign in $(n, n+v)$. Hence,

$$\int_n^{n+v} \left| \frac{d}{dt} t h_v(t) \right| dt = \left| \int_n^{n+v} \left(\frac{d}{dt} t h_v(t) \right) dt \right| = (n+v) h_v(n+v) - n h_v(n)$$
$$= v h_v(n) + (n+v)(h_v(n+v) - h_v(n))$$
$$= v \kappa_{0,v}(n) + (n+v) \left(\kappa_{1,v}(n) \log \frac{n+v}{n} + \kappa_{2,v}(n) \log^2 \frac{n+v}{n} \right).$$

If I contains 0, we take absolute values in (6.123):

$$\left|\frac{d}{dt}th_v(t)\right| = |(\kappa_{0,v}+\kappa_{1,v})(n)| + |(\kappa_{1,v}+2\kappa_{2,v})(n)|\log\frac{t}{n} + |\kappa_{2,v}(n)|\left(\log\frac{t}{n}\right)^2,$$

where $n < t < n+v$. Hence

$$\int_n^{n+v}\left|\frac{d}{dt}th_v(t)\right|dt = |(\kappa_{0,v}+\kappa_{1,v})(n)|\cdot v + |(\kappa_{1,v}+2\kappa_{2,v})(n)|\int_n^{n+v}\log\frac{t}{n}dt + |\kappa_{2,v}(n)|\int_n^{n+v}\left(\log\frac{t}{n}\right)^2 dt.$$

Clearly

$$\int_n^{n+v}\log\frac{t}{n}dt = n\int_0^{v/n}\log(1+\delta)d\delta = n\left(\left(1+\frac{v}{n}\right)\log\left(1+\frac{v}{n}\right)-\frac{v}{n}\right), \quad (6.125)$$

$$\int_n^{n+v}\left(\log\frac{t}{n}\right)^2 dt = n\int_0^{v/n}(\log(1+\delta))^2 \le n\left(\left(1+\frac{v}{n}\right)\log^2\left(1+\frac{v}{n}\right) - 2\left(1+\frac{v}{n}\right)\log\left(1+\frac{v}{n}\right) + 2\frac{v}{n}\right). \quad (6.126)$$

If n is large (say, $n > 10^6$), we use the bounds

$$\int_0^{v/n}\log(1+\delta)d\delta \le \int_0^{v/n}\delta d\delta = \frac{v^2}{2n^2},$$
$$\int_0^{v/n}(\log(1+\delta))^2 d\delta \le \int_0^{v/n}\delta^2 d\delta = \frac{v^3}{3n^3}.$$

In this way, summing over $n = 1, 1+v, 1+2v, \ldots$, we obtain that (6.122) and (6.123) hold for $T = 10^9$.

Lastly: the integral in (6.122) is obviously a non-increasing function of T. The same holds for the left side of (6.123): it increases when $(th_v)'(t)$ is non-zero and not of sign opposite to that of $th_v(t)$, and remains constant otherwise. Hence it is enough to evaluate the left sides of (6.122) and (6.123) for $T = 10^9$.

□

If we wanted more precise estimates in (6.122) and (6.123), we could use the fact that the integral of $|(th_v(t))'|$ from 1 to T equals

$$|t_1h_v(t_1) - h_v(1)| + |t_2h_v(t_2) - t_1h_v(t_1)| + \ldots + |t_kh_v(t_k) - t_{k-1}h_v(t_k-1)| + |Th_v(T) - t_kh_v(t_k)|, \quad (6.127)$$

where $t_1, t_2, \ldots, t_k$ are the local maxima and minima of $th_v(t)$ in $(1, T)$, or any superset thereof. The expression on the second line of (6.124) is quadratic on $\log(t/|t|)$, so we can find maxima and minima simply by solving a quadratic equation on $\log(t/n)$ for every n, or for each n for which interval arithmetic indicates (as above) that $(th_v(t))'$ might vanish inside $(n, n+v)$.

6.7 EXCURSUS ON THE DRESS-IWANIEC-TENENBAUM CONSTANT

Dress, Iwaniec and Tenenbaum [DIT83] considered the sum

$$\sum_{m\leq X}\left(\sum_{\substack{d\leq U\\ d|m}}\mu(d)\right)^2, \tag{6.128}$$

which can arise naturally when Vaughan's identity is applied without smoothing. Part of their motivation seems to have been to improve on the power of log in the Bombieri-Vinogradov theorem. Then U can be chosen significantly smaller than $\sqrt{X}$. (In the ternary Goldbach problem, on the other hand, one cannot easily make such an assumption, though one can assume that U is significantly *larger* than X.) It is then useful to write

$$\sum_{m\leq X}\left(\sum_{\substack{d\leq U\\ d|m}}\mu(d)\right)^2 = XS(U)+O(U^2),$$

where

$$S(U)=\sum_{m,n\leq U}\frac{\mu(m)\mu(n)}{[m,n]}.$$

The main theorem in [DIT83] states that $S(U)$ tends to a limit L as $U\to\infty$, and that, moreover,

$$L=\frac{6}{\pi^2}\sum_{j=1}^{\infty}\log\frac{j+1}{j}\sum_{m,n\leq j}\frac{\mu(mn)}{\sigma(m)\sigma(n)}.$$

Let

$$\dot{h}(y)=\sum_{d\leq y}\frac{\mu(d)}{\sigma(d)^2}\left(\dot{m}_d\left(\frac{y}{d}\right)\right)^2, \tag{6.129}$$

where $\dot{m}_d$ is as in (6.108). Then

$$L=\frac{6}{\pi^2}\int_1^{\infty}\dot{h}(y)\frac{dy}{y},$$

since, as we can easily show by applying (7.28) with $v=1$ and $\varrho=1_{[0,1]}$,

$$\dot{h}(y)=\sum_{\substack{m,n\leq y\\ (m,n)=1}}\frac{\mu(m)\mu(n)}{\sigma(m)\sigma(n)}.$$

As was stated in [DIT83, §3], computations suggested that $L\sim 0.440729$, but the state of knowledge at the time on $m(x)$ (defined in §5.3) left no hope of actually proving this, or anything close to this. (It is also reported in [DIT83] that R. R. Hall had shown that $L\leq 0.947$ by a different method, in unpublished work.)

Thanks in part to our approach and in part to the development of estimates on $m(x)$ since then, we can actually prove a good estimate on L, confirming what Dress, Iwaniec and Tenenbaum conjectured.

We first need some bounds on $\check{m}_v(x)$.

Lemma 6.28. *Let $\check{m}_d(x)$ be as in (6.108). Then, for $x \leq 10^{12}$,*

$$|\check{m}_d(x)| \leq \sqrt{\frac{2}{x}} \prod_{p|d} \frac{p+1}{p+1-\sqrt{p}} \tag{6.130}$$

for d arbitrary.

Moreover, for all $x > 0$,

$$\begin{aligned} |\check{m}_d(x)| \leq \prod_{p|d} \frac{p+1}{p+1-p^{\alpha_1}} \cdot \frac{2^{\alpha_1}\kappa(1-\alpha_1)}{x^{\alpha_1}} \\ + 0.0144 \prod_{p|d} \frac{p+1}{p+1-p^{\alpha_2}} \cdot \frac{\kappa(1-\alpha_2)}{\log x}, \end{aligned} \tag{6.131}$$

for $\alpha_1 \in [0, 1/2]$ arbitrary, $\alpha_2 = 1/\log 10^{12}$ and $\kappa(\beta)$ as in (6.30).

Proof. Let $\check{m}(x) = \check{m}_1(x)$. We establish

$$|\check{m}(x)| \leq \sqrt{\frac{2}{x}}, \tag{6.132}$$

for $x \leq 10^{12}$ by direct computation. Proceeding just as in the proof of Lemma 6.4, we obtain that, for d arbitrary,

$$\begin{aligned} |\check{m}_d(x)| &= \left| \sum_{q|d^\infty} \frac{1}{\prod_{p|q}(p+1)^{v_p(q)}} \check{m}(x/q) \right| \\ &\leq \sum_{q|d^\infty} \frac{\sqrt{2}}{\prod_{p|q}(p+1)^{v_p(q)}} \sqrt{\frac{q}{x}} \\ &= \sqrt{\frac{2}{x}} \prod_{p|d} \sum_{k=0}^{\infty} \left(\frac{\sqrt{p}}{p+1}\right)^k = \sqrt{\frac{2}{x}} \prod_{p|d} \frac{p+1}{p+1-\sqrt{p}}. \end{aligned}$$

For $x > 10^{12}$, we proceed just as in the proof of Prop. 6.8, starting from Lemma 5.12 instead of Lemma 5.13. We obtain (6.131) for $x > 10^{12}$. For $x \leq 10^{12}$, (6.131) follows from (6.130). □

Let us now prove a bound on $\dot{h}(y)$ analogous to the bounds on $h_v(y, y)$ we derived in §6.5.

Proposition 6.29. *Let $\dot{h}(y)$ be as in (6.129). If $y \geq 10^{12}$, then*

$$\dot{h}(y) \leq \frac{0.00096448}{(\log y)^2}. \tag{6.133}$$

If $y \leq 10^{12}$, then

$$|\dot{h}(y)| \leq \frac{2 + 2.83624 \log y}{y}. \tag{6.134}$$

We could derive a bound much sharper than (6.133) (but less sharp than (6.134)) for $10^{12} < y \leq 10^{14}$ if we used Lemma 5.10.

Proof. We will bound $\dot{h}(y)$ by $\dot{\mathbf{h}}(y)$, where

$$\dot{\mathbf{h}}(y) = \sum_{d \leq y} \frac{\mu^2(d)}{\sigma(d)^2} \left(\dot{m}_d \left(\frac{y}{d} \right) \right)^2.$$

Hence, by Lemma 6.28,

$$\dot{\mathbf{h}}(y) \leq \frac{2}{y} \sum_{d \leq y} \frac{\mu^2(d)}{d} \prod_{p|d} \left(\frac{p}{p - \sqrt{p} + 1} \right)^2 \tag{6.135}$$

for $y \leq 10^{12}$. By (6.56), we conclude that

$$\dot{\mathbf{h}}(y) \leq \frac{2 + 2.83624 \log y}{y}$$

for $y \leq 10^{12}$.

Assume now that $y \geq 10^{12}$. We proceed just as in the proof of Lemma 6.15, and obtain that $\dot{\mathbf{h}}(y)$ is at most

$$\sum_{i=1}^{3} c_i \sum_d \frac{\mu^2(d)}{\sigma(d)^2} \prod_{j=1}^{2} d^{\alpha_{i,j}} \prod_{p|d} \frac{p+1}{p+1-p^{\alpha_{i,j}}}$$

for $\alpha = 1/\log 10^{12}$, where

$$c_1 = 2^{1-\epsilon} \frac{\kappa(1/2)}{\sqrt{y_1}} \frac{\kappa(1/2+\epsilon)}{y_2^{1/2-\epsilon}}, \qquad \alpha_{1,1} = \frac{1}{2}, \quad \alpha_{1,2} = \frac{1}{2} - \epsilon,$$

$$c_2 = 2 \cdot 2^{1/2} \cdot 0.0144 \frac{\kappa(1/2)}{\sqrt{y}} \frac{\kappa(1-\alpha)}{\log y}, \qquad \alpha_{2,1} = \frac{1}{2}, \quad \alpha_{2,2} = \alpha,$$

$$c_3 = 0.0144^2 \frac{\kappa(1-\alpha)^2}{(\log y)^2}, \qquad \alpha_{3,1} = \alpha_{3,2} = \alpha$$

and $\epsilon \in (0, 1/2)$. We apply Lemma 6.11, and conclude that

$$\begin{aligned} \dot{\mathbf{h}}(y) &\leq 2^{1-\epsilon} \frac{\kappa(1/2)\kappa(1/2+\epsilon)}{y^{1-\epsilon}} \kappa(1/2, 1/2+\epsilon) \\ &+ 0.0288\sqrt{2}\frac{\kappa(1/2)\kappa(1-\alpha)}{\sqrt{y}\log y}\kappa\left(\frac{1}{2}, 1-\alpha\right) + 0.0144^2 \frac{\kappa(1-\alpha)^2}{(\log y)^2}\kappa(1-\alpha, 1-\alpha). \end{aligned} \tag{6.136}$$

We apply (6.79)–(6.81), and obtain

$$\begin{aligned} \dot{\mathbf{h}}(y) &\leq \frac{370.552602}{y^{1-1/14}} + \frac{0.0288 \cdot 23.885778}{\sqrt{y}\log y} + \frac{0.0144^2 \cdot 4.549747}{(\log y)^2} \\ &\leq \frac{0.00096448}{(\log y)^2}. \end{aligned} \tag{6.137}$$

□

Proposition 6.30. *Let*

$$S(U) = \sum_{d_1, d_2 \leq U} \frac{\mu(d_1)\mu(d_2)}{[d_1, d_2]}.$$

Then $\lim_{U\to\infty} S(U) = L$, *where*

$$L = 0.440729 + O^*(0.0000213).$$

The six digits in the main term here are likely to be correct; most of the error term here comes from our (suboptimal) bound on an integral from 10^{14} to ∞.

Proof. As we said, the main theorem in [DIT83] gives us that

$$L = \frac{6}{\pi^2}\int_1^\infty \dot{h}(y)\frac{dy}{y},$$

where $\dot{h}(y)$ is as in (6.108).

By Prop. 6.29,

$$\begin{aligned} \int_{10^{12}}^\infty \dot{h}(y)\frac{dy}{y} &\leq \int_{10^{12}}^\infty \frac{0.00096448}{(\log y)^2}\frac{dy}{y} = \frac{0.00096448}{\log 10^{12}} \\ &\leq 0.00003491, \end{aligned} \tag{6.138}$$

$$\begin{aligned} &\int_{10^9}^{10^{12}} \dot{h}(y)\frac{dy}{y} \leq \int_{10^9}^{10^{12}} \frac{2 + 2.83624\log y}{y}\frac{dy}{y} \\ &= 2\cdot\left(\frac{1}{10^9} - \frac{1}{10^{12}}\right) + 2.83624\cdot\left(\frac{1+\log 10^9}{10^9} - \frac{1+\log 10^{12}}{10^{12}}\right) \\ &\leq 2.019\cdot 10^{-9}. \end{aligned} \tag{6.139}$$

The same computation as in Prop. 6.26 (or, rather, part of the computation) gives us that

$$\int_1^{10^9} \dot{h}(y)\frac{dy}{y} \leq 0.7249702 + O^*(6 \cdot 10^{-8}). \tag{6.140}$$

Hence

$$\begin{aligned} L &= \frac{6}{\pi^2}(0.7249702 + O^*(0.0000349721)) \\ &= 0.440729 + O^*(0.0000213). \end{aligned} \tag{6.141}$$

□

Chapter Seven

A natural upper-bound sieve

7.1 INTRODUCTION

Our aim in this chapter is to find cancellation in sums of the form

$$\sum_{d|m} \mu(d)\varrho(d), \tag{7.1}$$

where ϱ is a smoothing function, equal to a constant for m larger than some parameter. What we want is an upper bound on the ℓ^2 norm of (7.1); that is, we want to bound

$$\sum_{m\in I\cap\mathbb{Z}} \left(\sum_{d|m} \mu(d)\varrho(d)\right)^2,$$

where I is an interval.

Experts will recognize this as a situation of small sieve type. This may come as a surprise, since the expression we have to bound may seem to have arisen almost out of its own volition.

Some words on small sieves are in order here. It is of course clear that, if ϱ were the constant function $\varrho(d) = 1$, then (7.1) would be 1 for $m = 1$, and 0 otherwise. If, instead, $\varrho(d) = \log d$, then (7.1) equals $-\Lambda(m)$, and so is non-zero only at the primes and their powers. A *small sieve* is, in essence, a clever choice of function ϱ, so that, while ϱ obeys some relatively strong constraints that $d \mapsto \log d$ and the constant function do not, (7.1) still captures the primes approximately, in the sense of, say, being supported on a relatively small superset of the primes, or some other related sense.

If, to be more precise, we aim at bounding (7.1) in ℓ^2 norm from above, we say we have a quadratic sieve. The best we can usually hope for is an upper bound that is no larger than a constant times the number of primes in I. As is well-known, such a bound is achieved by *Selberg's sieve*, where ϱ is defined as a function from $\mathbb{Z}^+$ to $\mathbb{R}$ given as the solution to an optimization problem. We need to work with ϱ defined on $\mathbb{R}$ – and preferably monotonic, unlike the choice of ϱ in Selberg's sieve.

The first version of the present proof of ternary Goldbach simply used a brutal truncation ϱ, with $\varrho(d) = 0$ for $d \leq U$, $\varrho(d) = 1$ for $d > U$. The ℓ^2 norm is not then bounded by a constant times the number of primes in I; rather, it asymptotes to a constant times the number of integers in I. A numerical value for the constant had already been conjectured by [DIT83]. What the relevant part of [Helb] contained was a proven estimate, with a small, explicit error term.

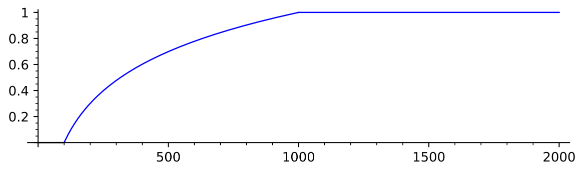

Figure 7.1: The smoothing function $\varrho(t)$ with $U_0 = 100$, $U_1 = 1000$

Here, we will work with a continuous ϱ, and show that it is essentially as good as Selberg's sieve. This is within an existing line of work that may reasonably be said to be related to work done by Selberg *before* his sieve, when he mollified the Riemann zeta function [Sel42]; see the discussion in [FI10, §7.2]. The particular smoothing function ϱ we will use was already studied in this context by Barban-Vehov [BV68], Motohashi [Mot74] (see also [Mot76]) and Graham [Gra78].

Theorem 7.1. *Let $U_1 > U_0 > 0$, $v \in \{1, 2\}$, $\beta \in \{0, 1/2\}$. For $t > 0$, let*

$$\varrho(t) = \frac{\log^+(t/U_0) - \log^+(t/U_1)}{\log U_1/U_0} = \begin{cases} 0 & \text{if } t < U_0, \\ \frac{\log t/U_0}{\log U_1/U_0} & \text{if } U_0 \le t \le U_1, \\ 1 & \text{if } t > U_1, \end{cases} \tag{7.2}$$

where $\log^+ x = \max(\log x, 0)$. Let

$$S_{v,\beta}(X) = \sum_{\substack{\beta X < m \le X \\ (m,v)=1}} \left(\sum_{d|m} \mu(d)\varrho(d) \right)^2 .$$

(a) For $U_0 \le X \le U_1$,

$$S_{v,\beta}(X) \le \left((1-\beta) \log \frac{X}{U_0} + R_{v,\beta}\left(\frac{U_0}{\sqrt{X}} \right) \right) \cdot \frac{X}{\log^2 \frac{U_1}{U_0}}, \tag{7.3}$$

where

$$R_{v,\beta}(t) = \frac{\kappa_{v,1}}{t} + \frac{\kappa_{v,2}}{t^{3/2}}, \tag{7.4}$$

$$\kappa_{1,1} = 34.39 - 13.75\beta, \quad \kappa_{1,2} = 13.99 - 6.16\beta,$$
$$\kappa_{2,1} = 4.93 - 1.64\beta, \quad \kappa_{2,2} = 2.54 - 0.84\beta.$$

(b) For $X \ge U_1$,

$$S_{v,\beta}(X) \le \frac{(1-\beta)X}{\log \frac{U_1}{U_0}} + R_{v,\beta}\left(\frac{U_0}{\sqrt{X}} \right) \cdot \frac{X}{\log^2 \frac{U_1}{U_0}}, \tag{7.5}$$

where $R_{v,\beta}(t)$ is as in (7.4), provided that $U_1/U_0 \ge c_{v,\beta}$, where

$$c_{1,1/2} = 9, \quad c_{2,1/2} = 9, \quad c_{1,0} = 8, \quad c_{2,0} = 6.$$

If $X < U_0$, then $S_{v,\beta}$ clearly equals 0.

A few remarks are in order. Here, we give a remainder term $R_{\beta,v}$ that is small for U_0 greater than $\sqrt{X}$; as we shall discuss in §7.2, it is most likely also possible to give a remainder term that is small for U_1 smaller than $\sqrt{X}$. In fact, it is a priori more noteworthy that – as has been known since [BV68] and [Gra78] – one can give useful bounds when U_0 greater than $\sqrt{X}$, since such is not the case for most sieves. (Simply changing variables $d \mapsto m/d$ would result in a mess.)

We are allowing a parameter β for the sake of generality. A reader wishing to apply Theorem 7.1 to sums over $m \le X$ would set $\beta = 0$. We will later apply Theorem 7.1 with $\beta = 1/2$, as we are interested in a sum over a dyadic interval, i.e., a sum over $X/2 < m \le X$. In this chapter, we carry out all our work for a general $\beta \in [0,1]$ until the very end, when we focus on $\beta = 0$ and $\beta = 1/2$ to optimize our estimates.

It is unsurprising that, if we let $v = 2$ rather than $v = 1$, the main term does not change, yet the remainder-term bound decreases sharply. Setting $v = 2$ amounts to sieving out the even numbers in a very simple way before we apply our sieve. It was already known (since the work of Selberg himself) that it can be useful to combine Selberg's sieve with a relatively simple sieve for small prime numbers; see [FI10, §7.6–7.7].

The assumption $U_1/U_0 \ge c_{v,\beta}$ in Thm. 7.1((b)) is not unduly restrictive, in that using the smoothing ϱ in (7.2) with $U_1/U_0 < c_{v,\beta}$ would make little sense: the denominator $\log^2(U_1/U_0)$ in (7.3) and (7.5) would be small. Moreover, precisely because $\log^2(U_1/U_0)$ is small, it is likely possible to prove Thm. 7.1((b)) for $1 \le U_1/U_0 \le c_{v,\beta}$ starting from main-term estimates for the "poor man's sieve", that is, an analogue of Thm. 7.1 for the brutal truncation $\varrho = 1_{[1,\infty)}$. We do not study the poor man's sieve here for reasons of space. See, however, [Helb, §4.1].

Refinements. We actually give a bound on $S_{v,\beta}$ a little sharper than that in (7.3) and (7.5); see Theorem 7.23, a technically stronger version of Theorem 7.1. On a relatively broad range, the improvement in L is enough to swallow the remainder term $R_{\beta,v}$. Thus we will have, for instance,

$$\sum_{\substack{X/2<m\le X\\(m,v)=1}} \left(\sum_{d|m} \mu(d)\varrho(d)\right)^2 \le \frac{\log\frac{X}{U_0}}{\log^2\frac{U_1}{U_0}} \cdot \frac{X}{2} \tag{7.6}$$

for $\beta = 1/2$, $v = 1$ and (say) $100\sqrt{X} \le U_0 \le X/3$, $U_1 \ge X$, or for $\beta = 1/2$, $v = 2$ and (say) $14\sqrt{X} \le U_0 \le X/3$, $U_1 \ge X$.

Comparison to previous versions. Prior work was non-explicit, except for the recent preprint [Bet19], which proves a bound corresponding roughly to

$$S_{1,0} \le \frac{166X}{\log(U_1/U_0)}$$

in the special case $U_1 = U_0^2$. The leading-order terms in (7.3) (for $U_0 \le X \le U_1$) and (7.5) (for $X > U_1$) had been known since Graham [Gra78].

While previous work towards Thm. 7.1 was in principle effective, it was based on estimates of the form $\sum_{m\le M} \mu(m) \le C_A M/(\log M)^A$. As soon as A is about, say, 2 or larger, such estimates are known only with C_A much too large to be practical. We will base our work instead on the results in [Ram13b] (see §5.3), which rely on a bound for $A = 1$. In other words, we face the challenge of working using tools based on an input that is better than trivial only by a factor of log times a constant.

Comparison to Selberg's sieve. Selberg's sieve, with $\rho : \mathbb{Z}^+ \to \mathbb{R}$ supported on integers $\leq U$, gives us a bound of the form

$$\sum_{m\leq x} \left(\sum_{d|m} \mu(d)\rho(d) \right)^2 \leq \frac{x}{h(U)} + \text{remainder term}, \tag{7.7}$$

where $h(U) = \sum_{d\leq U} \mu^2(d)/\phi(d)$ (see, say, [IK04, Thm. 6.4]), and the remainder term is $O(U^2/\log^2 U)$ (as in [IK04, §6.7]). Here $h(U) > \log U$ for all U and $h(U) = \log U + 1.33258\ldots + o(1)$ ([Ram95, Lem. 3.4]).

The estimate to be compared to Selberg's sieve would be the variant of Thm. 7.1 to which we referred in the above – namely, a variant where the remainder term is small for U_1 smaller than $\sqrt{X}$; as we will see in §7.2, the remainder term would be $\lesssim U_1^2/(\zeta(2) \log U_1/U_0)^2 = (36/\pi^4)U^2/\log^2 U$ for $U_1 = U$, $U_0 = 1$. As we shall see, there is still work remaining to be done on the main term of such a possible variant. While that main term cannot be quite as low as the main term $x/h(U)$ in (7.7) (since Selberg's sieve is optimal), it seems reasonable, in the light of Theorem 7.1((a)), to hope to bound it by $x/\log U$, or even by $x/\log U - (c + o(1))/(\log U)^2$, where c is some constant close to $1.33258\ldots$; see Theorem 7.23 and the brief discussion therafter (§7.8.1).

The main reason to be interested in a sieve like that in Theorem 7.1 for $U_1 \ll \sqrt{X}$ (call such a sieve "low-range and continuous") is essentially the same as our reason for proving Theorem 7.1 ("high-range and continuous"): there are contexts in which a continuous, monotonic, bounded ϱ is helpful or necessary. (The optimal sequence $\rho(d)$ in Selberg's quadratic sieve is neither continuous nor bounded, nor the restriction of a sensible continuous function to the integers.) In our context, the function ϱ arises as a smoothing applied to Vaughan's identity. Of course, if ϱ is continuous, monotonic and bounded, so is $1 - \varrho$, and having a continuous, monotonic and bounded weight $1 - \rho$ is very useful for the estimation of other terms of Vaughan's identity.

Further perspectives. The quadratic sieve considered here has been generalized (starting with [Mot78], [Mot83] and then [Jut79b], [Jut79a], [Mot04]) and applied in [GPY09], [Zha14], [May15] and related work. It may be worthwhile to attempt to adapt our explicit work to such more general situations. What is being sieved for in such cases is not primes, but products of few primes, and the choice of function ϱ would then reflect this fact. We use a function based on log; to sieve for products of few primes, what is used is functions based on powers of log.

See also the variant in [Ram12], which is in some sense intermediate between Selberg's sieve and the quadratic sieve considered here.

Can one hope to give a good error term when either U_0 or U_1 is close to $\sqrt{X}$? The recent result [dlBDT] is encouraging.

On a different note: is the function ϱ in Theorem 7.1 optimal within its class? That is to say: is L (in (7.3) or (7.151)) minimal, when we let ϱ range among monotonic, continuous functions such that $\varrho(t) = 0$ for $t < U_0$ and $\varrho(t) = 1$ for $t > U_1$? (The bound on L to be considered here is not that in (7.3), but the best possible one, as discussed at the end of §7.8.1.) The answer is unknown as of the time of writing. (It has been clear since [BV68] that the choice of ϱ used here gives a result of the optimal order of magnitude. See also [GKM16].)

7.2 SEPARATING THE MAIN TERM AND THE REMAINDER

Let us first discuss how to separate the remainder term in (7.3) from the main term.

The fact that the remainder term in (7.3) is essentially proportional to $(\sqrt{X}/U_0)\cdot X$ may seem odd. This is a deliberate feature: in our main application, we will be able to ensure that U_0 is considerably larger than $\sqrt{X}$, and thus $(\sqrt{X}/U_0)\cdot X$ is considerably smaller than X.

It would be possible to give instead a remainder term essentially proportional to U_1^2. (We see that the range $U_0 \lesssim \sqrt{X}$, $U_1 \gtrsim \sqrt{X}$ would still be problematic.) Such a remainder term would be what is usual in sieve theory. Let us see how a treatment in such a direction would begin, and where would it face non-trivial difficulties similar to the ones we will overcome.

By (2.1), $\sum_{d|m}\mu(d)\varrho(d) = -\sum_{d|m}\mu(d)(1-\varrho(d))$ for $m > 1$. Write λ_t for $1-\varrho(t)$. Then

$$\begin{aligned}\sum_{\substack{X/2<m\leq X\\(m,v)=1}}\left(\sum_{d|m}\mu(d)\varrho(d)\right)^2 &= \sum_{d_1}\sum_{d_2}\sum_{\substack{X/2<m\leq X\\(m,v)=1\\d_1|m,d_2|m}}\mu(d_1)\mu(d_2)\lambda_{d_1}\lambda_{d_2}\\ &= \sum_{\substack{d_1,d_2\\(d_1d_2,v)=1}}\mu(d_1)\mu(d_2)\lambda_{d_1}\lambda_{d_2}\sum_{\substack{X/2<m\leq X\\(m,v)=1\\ [d_1,d_2]|m}}1\\ &= \frac{X}{2v}\sum_{\substack{d_1,d_2\\(d_1d_2,v)=1}}\frac{\mu(d_1)\mu(d_2)}{[d_1,d_2]}\lambda_{d_1}\lambda_{d_2} + O^*\left(\sum_{\substack{d_1,d_2\\(d_1d_2,v)=1}}\mu^2(d_1)\mu^2(d_2)\lambda_{d_1}\lambda_{d_2}\right).\end{aligned} \tag{7.8}$$

The last term here equals $(\sum_{d:(d,v)=1}\mu^2(d)\lambda_d)^2$. For our choice of ϱ, it is easy to see that

$$\sum_{\substack{d\\(d,v)=1}}\mu^2(d)\lambda_d \sim \frac{6v}{\pi^2\sigma(v)}\int_0^{U_1}\lambda_t dt = \frac{6v}{\pi^2\sigma(v)}\frac{U_1-U_0}{\log U_1/U_0},$$

and thus the remainder is $\sim (U_1-U_0)^2/(\zeta(2)\log U_1/U_0)^2$. (The case $U_0 = 1$ then gives us $\sim U_1^2/(\zeta(2)\log U_1)^2$, much as in Selberg's sieve.)

That was very easy. What is more complicated is to estimate the main term. It is tempting to write, as in [Gra78, §3],

$$\begin{aligned}\sum_{\substack{d_1,d_2\\(d_1d_2,v)=1}} \frac{\mu(d_1)\mu(d_2)}{[d_1,d_2]}\lambda_{d_1}\lambda_{d_2} &= \sum_{\substack{d_1,d_2\\(d_1d_2,v)=1}} \frac{\mu(d_1)\mu(d_2)}{d_1d_2}\lambda_{d_1}\lambda_{d_2}\sum_{r|(d_1,d_2)}\phi(r)\\ &= \sum_{\substack{r\leq U_1\\(r,v)=1}} \frac{\mu^2(r)\phi(r)}{r^2}\left(\sum_{\substack{d\\(d,rv)=1}}\frac{\mu(d)}{d}\lambda_{rd}\right)^2.\end{aligned} \tag{7.9}$$

If r is considerably smaller than U_1 yet larger than U_0, or considerably smaller than U_0, we can find a fair deal of cancellation in the sum inside the square, proceeding as in §5.3. The problem is precisely what to do when r is slightly smaller than either U_0 or U_1. If r is slightly smaller than U_1, the sum inside the square in (7.9) is very short, and we are thus unlikely to be able to find much cancellation in it. In part for this reason, the error term given in [Gra78, §3] is of size $O(1/\log(U_1/U_0))$ times the main term. We need to do better.

* * *

As we already said, we will really be working with U_0 larger than $\sqrt{X}$, and so should aim at a remainder term proportional to $(X/U_0)^2$ (or smaller), not to U_1^2 or $U_1^2/(\log U_1)^2$. This is actually a less intuitive case at first, since the variables d_1, d_2 are large rather than small; we cannot extract the remainder term as easily as in (7.8).

Instead of starting by switching from ϱ to $1-\varrho$, we apply what is essentially the same procedure as in [Gra78, p. 91].

Lemma 7.2. *Let $X > X_0 \geq 0$ and $v \in \mathbb{Z}^+$. Let $\varrho : \mathbb{R}^+ \to \mathbb{C}$. Then*

$$\sum_{\substack{X_0<m\leq X\\(m,v)=1}}\left(\sum_{d|m}\mu(d)\varrho(d)\right)^2$$

equals

$$\sum_{\substack{s\\(s,v)=1}}\sum_{\substack{r_1\\(r_1,r_2)=1\\(r_1r_2,v)=1}}\sum_{r_2}\mu(r_1)\mu(r_2)\sum_{\substack{\frac{X_0}{r_1r_2s}<l\leq\frac{X}{r_1r_2s}\\(l,r_1r_2v)=1,\mu(l)^2=1}}\varrho(r_1l)\varrho(r_2l). \tag{7.10}$$

Proof. We expand and change variables:

$$\sum_{\substack{X_0<m\leq X\\(m,v)=1}}\left(\sum_{d|m}\mu(d)\varrho(d)\right)^2 = \sum_{\substack{X_0<m\leq X\\(m,v)=1}}\sum_{d_1,d_2|m}\mu(d_1)\mu(d_2)\varrho(d_1)\varrho(d_2)$$
$$= \sum_{\substack{r_1\\(r_1,r_2)=1\\(r_1r_2,v)=1}}\sum_{r_2}\sum_{\substack{l\\(l,r_1r_2v)=1\\\mu^2(l)=1}}\mu(r_1)\mu(r_2)\varrho(r_1l)\varrho(r_2l)\sum_{\substack{X_0<m\leq X\\r_1r_2l|m\\(m,v)=1}}1, \tag{7.11}$$

where $d_1 = r_1l$, $d_2 = r_2l$, $l = (d_1, d_2)$. Let $s = m/(r_1r_2l)$. Changing the order of summation, we obtain (7.10). □

We now diverge from [Gra78]. We will study the inner triple sum in (7.10), detaching a remainder term from it. We proceed by standard techniques: we reformulate our expression so as to free a variable, over which we then sum, thus eliminating the variable. We then obtain our main term by completing a sum, that is, making it into an infinite sum, which we are then able to simplify.

Proposition 7.3. *Let $z, u > 0$, $0 \leq \beta \leq 1$, $v \in \{1, 2\}$. Let $\varrho : \mathbb{R} \to \mathbb{C}$ be a non-decreasing function such that ϱ, ϱ' are in L^1 and $\varrho(t) = 0$ for $t < u$. Then*

$$\sum_{\substack{r_1\\(r_1,r_2)=1\\(r_1r_2,v)=1}}\sum_{r_2}\mu(r_1)\mu(r_2)\sum_{\substack{\frac{\beta z}{r_1r_2}<l\leq\frac{z}{r_1r_2}\\(l,r_1r_2v)=1,\ \mu(l)^2=1}}\varrho(r_1l)\varrho(r_2l), \tag{7.12}$$

equals

$$\frac{6z}{\pi^2}\cdot\frac{v}{\sigma(v)}\int_\beta^1 g_v(tz)dt$$
$$+ O^*\left(c_{v,1}\sqrt{z}\cdot R_{1,1/2,u,v,\varrho}(z/u) + c_{v,2}z^{1/4}\cdot R_{1/2,1/4,u,v,\varrho}(z/u)\right), \tag{7.13}$$

where

$$g_v(y) = \sum_{\substack{r_1\\(r_1,r_2)=1\\(r_1r_2,v)=1}}\sum_{r_2}\frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)}\varrho\left(\frac{y}{r_1}\right)\varrho\left(\frac{y}{r_2}\right), \tag{7.14}$$

$$R_{\alpha_1,\alpha_2,u,v,\varrho}(y) = \sum_{\substack{d_1,d_2\leq y\\\mu^2(d_1l_1d_2l_2v)=1}}\sum_{l_i\leq y/d_i}\frac{\varrho\left(\frac{yu}{d_1l_1}\right)\varrho\left(\frac{yu}{d_2l_2}\right)}{(d_1d_2)^{\alpha_1}(l_1l_2)^{\alpha_2}}, \tag{7.15}$$

$$c_{v,1} = \begin{cases}\left(\frac{5}{3}-\frac{2\beta}{3}\right)\frac{6}{\pi^2} & \text{if } v=1,\\ \left(\frac{3}{2}-\frac{\beta}{2}\right)\frac{4}{\pi^2} & \text{if } v=2,\end{cases}\qquad c_{v,2} = \begin{cases}\sqrt{3}\left(1-\frac{6}{\pi^2}\right)+(1-\beta)\left(\sqrt{8}-\frac{8^{3/2}}{\pi^2}\right) & \text{if } v=1,\\ 1-\frac{4}{\pi^2}+\frac{1-\beta}{2}\left(1-\frac{4}{\pi^2}\right) & \text{if } v=2.\end{cases}$$

Proof. By Möbius inversion in the sense of (2.1), the expression in (7.12) equals

$$\sum_{\substack{r_1\\(r_1,r_2)=1\\(r_1r_2,v)=1}}\sum_{r_2}\mu(r_1)\mu(r_2)\sum_{\substack{\frac{\beta z}{r_1r_2}<l\le\frac{z}{r_1r_2}\\(l,v)=1}}\varrho(r_1l)\varrho(r_2l)\sum_{\substack{d_1|r_1,d_2|r_2\\d_1d_2|l}}\mu(d_1)\mu(d_2)\sum_{\substack{m^2|l\\(m,r_1r_2)=1}}\mu(m). \tag{7.16}$$

We now change the order of summation, putting the sum over l at the very end, and introducing the variable $n = l/d_1d_2m^2$. The inner triple sum in (7.16) then becomes

$$\sum_{d_1|r_1,d_2|r_2}\mu(d_1)\mu(d_2)\sum_{\substack{m\\(m,r_1r_2v)=1}}\mu(m)\sum_{\substack{\frac{\beta z}{r_1r_2d_1d_2m^2}<n\le\frac{z}{r_1r_2d_1d_2m^2}\\(n,v)=1}}\varrho(r_1d_1d_2m^2n)\varrho(r_2d_1d_2m^2n). \tag{7.17}$$

We estimate the innermost sum by (3.4) and (3.10); we obtain that it is

$$\frac{z}{vr_1r_2d_1d_2m^2}\int\Upsilon_{z/r_1,z/r_2}(t)dt + O^*\left(\frac{1}{2}\left|\Upsilon'_{z/r_1,z/r_2}\right|_1\right), \tag{7.18}$$

where

$$\Upsilon_{y_1,y_2}(t) = \varrho(y_1t)\varrho(y_2t)\cdot 1_{(\beta,1]}.$$

Of course, the innermost sum is actually 0 if $r_1r_2d_1d_2m^2 > z$ or either r_1 or r_2 is $> z/u$. We will set the conditions $r_1, r_2 \le z/u$ throughout. Our error terms will come from the error term in (7.18), when $m \le \sqrt{z/r_1r_2d_1d_2}$, and the main term in (7.18), when $m > \sqrt{z/r_1r_2d_1d_2}$; that is, we complete the sum on m in the main term.

Our main term is thus

$$z\int\sum_{\substack{r_1,r_2\le z/u\\(r_1,r_2)=1\\(r_1r_2,v)=1}}\frac{\mu(r_1)\mu(r_2)}{r_1r_2}\Upsilon_{z/r_1,z/r_2}(t)\sum_{d_1|r_1,d_2|r_2}\sum_{\substack{m\\(m,r_1r_2v)=1}}\frac{\mu(d_1)\mu(d_2)\mu(m)}{vd_1d_2m^2}dt.$$

For r_1, r_2 satisfying $(r_1, r_2) = 1$, $(r_1r_2, v) = 1$,

$$\begin{aligned}\sum_{d_1|r_1,d_2|r_2}\sum_{\substack{m\\(m,r_1r_2v)=1}}\frac{\mu(d_1)\mu(d_2)\mu(m)}{vd_1d_2m^2} &= \frac{1}{v}\prod_{p|r_1r_2}\left(1-\frac{1}{p}\right)\prod_{\substack{p\nmid r_1r_2\\p\nmid v}}\left(1-\frac{1}{p^2}\right)\\ &= \frac{v}{\sigma(v)}\cdot\frac{6}{\pi^2}\frac{1}{\prod_{p|r_1r_2}(1+1/p)}.\end{aligned}$$

(Here we are using the fact that $v = 1$ or $v = 2$, but the factor $v/\sigma(v)$ is correct in general: the dependence on v in (7.18) would look different for $v > 2$.) Hence, our main term simplifies to

$$\frac{6z}{\pi^2} \cdot \frac{v}{\sigma(v)} \int \sum_{\substack{r_1,r_2 \le z/u \\ (r_1,r_2)=1 \\ (r_1r_2,v)=1}} \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)} \Upsilon_{z/r_1,z/r_2}(t)\, dt.$$

It remains to consider the error terms. The terms coming from completing the sum add up to

$$\sum_{\substack{r_1,r_2 \le z/u \\ (r_1,r_2)=1 \\ (r_1r_2,v)=1}} \frac{z\mu(r_1r_2)}{vr_1r_2} \int \Upsilon_{z/r_1,z/r_2}(t)dt \sum_{d_1|r_1, d_2|r_2} \frac{\mu(d_1)\mu(d_2)}{d_1d_2} \sum_{\substack{m > \sqrt{\frac{z}{r_1r_2d_1d_2}} \\ (m,r_1,r_2v)=1}} \frac{\mu(m)}{m^2}, \tag{7.19}$$

whereas the terms coming from the error term in (7.18) contribute at most

$$\sum_{\substack{r_1,r_2 \le z/u \\ (r_1,r_2)=1 \\ (r_1r_2,v)=1}} \mu^2(r_1r_2) \sum_{d_1|r_1, d_2|r_2} \frac{1}{2} \left|\Upsilon'_{z/r_1,z/r_2}\right|_1 \sum_{\substack{m \le \sqrt{\frac{z}{r_1r_2d_1d_2}} \\ (m,r_1,r_2v)=1}} \mu^2(m) \tag{7.20}$$

in absolute value.

Since ϱ is non-decreasing,

$$\int \Upsilon_{z/r_1,z/r_2}(t)dt = \int_\beta^1 \varrho\left(\frac{zt}{r_1}\right) \varrho\left(\frac{zt}{r_2}\right) dt \le (1-\beta)\varrho\left(\frac{z}{r_1}\right) \varrho\left(\frac{z}{r_2}\right)$$

and

$$\frac{1}{2}\left|\Upsilon'_{z/r_1,z/r_2}\right|_1 \le \varrho\left(\frac{z}{r_1}\right) \varrho\left(\frac{z}{r_2}\right).$$

We now apply Lemmas 5.1–5.3, and obtain that the absolute value of the contribution of (7.19) plus that of (7.20) is at most

$$\begin{aligned} &c_{v,1} \sum_{\substack{r_1,r_2 \le z/u \\ (r_1,r_2)=1 \\ (r_1r_2,v)=1}} \mu^2(r_1r_2) \varrho\left(\frac{z}{r_1}\right) \varrho\left(\frac{z}{r_2}\right) \sum_{d_1|r_1, d_2|r_2} \sqrt{\frac{z}{r_1r_2d_1d_2}} \\ &+ c_{v,2} \sum_{\substack{r_1,r_2 \le z/u \\ (r_1,r_2)=1 \\ (r_1r_2,v)=1}} \mu^2(r_1r_2) \varrho\left(\frac{z}{r_1}\right) \varrho\left(\frac{z}{r_2}\right) \sum_{d_1|r_1, d_2|r_2} \left(\frac{z}{r_1r_2d_1d_2}\right)^{1/4}, \end{aligned} \tag{7.21}$$

where

$$c_{v,1} = \begin{cases} \frac{6}{\pi^2} + (1-\beta)\frac{4}{\pi^2} & \text{if } v = 1, \\ \frac{4}{\pi^2} + \frac{1-\beta}{2}\frac{4}{\pi^2} & \text{if } v = 2, \end{cases}$$

$$c_{v,2} = \begin{cases} \sqrt{3}\left(1-\frac{6}{\pi^2}\right) + (1-\beta)\left(\sqrt{8}-\frac{8^{3/2}}{\pi^2}\right) & \text{if } v=1, \\ 1-\frac{4}{\pi^2}+\frac{1-\beta}{2}\left(1-\frac{4}{\pi^2}\right) & \text{if } v=2. \end{cases}$$

□

Corollary 7.4. *Let $X, u > 0$, $0 \le \beta \le 1$, $v \in \{1,2\}$. Let $\varrho : \mathbb{R} \to \mathbb{C}$ be a non-decreasing function such that ϱ, ϱ' are in L^1 and $\varrho(t) = 0$ for $t < u$. Then*

$$\sum_{\substack{\beta X < m \le X \\ (m,v)=1}} \left(\sum_{d|m} \mu(d)\varrho(d)\right)^2 = \frac{6X}{\pi^2}\frac{v}{\sigma(v)} \int_\beta^1 \sum_{\substack{s \le tX/u \\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{tX}{s}\right) dt$$
$$+ O^*\left(\sum_{j=1}^2 c_{v,j} \sum_{\substack{s \le X/u \\ (s,v)=1}} R_{1/j,1/2j,u,v,\varrho}(X/su) \cdot \left(\frac{X}{s}\right)^{\frac{1}{2j}}\right), \tag{7.22}$$

where g_v, $R_{\alpha_1,\alpha_2,u,v,\varrho}$, $c_{v,1}$ and $c_{v,2}$ are as in the statement of Prop. 7.3.

Proof. By Lemma 7.2 and Prop. 7.3 with $X_0 = \beta X$ and $z = X/s$. Since $g_v(x) = 0$ for $x < u$,

$$\sum_{\substack{s \le X/u \\ (s,v)=1}} \frac{1}{s} \int_\beta^1 g_v\left(\frac{tX}{s}\right) dt = \int_\beta^1 \sum_{\substack{s \le tX/u \\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{tX}{s}\right) dt.$$

□

We have here an expression whose remainder terms will be small when u is larger than $\sqrt{X}$ by at least a somewhat large constant factor. When u is rather close to $\sqrt{X}$, it may be better to modify the proof of Prop. 7.3, estimating (7.19) and (7.20) more coarsely; the remainder term would then involve only $R_{1,1/2,u,v,\varrho}(X/su)\sqrt{X/s}$, times a constant larger than $c_{v,1}$. We will not bother, since, in our application, u will in fact be a fair bit larger than $\sqrt{X}$.

Remark. Part (b) of the main theorem in [DIT83] states that, as $U \to \infty$,

$$\sum_{d_1,d_2 \le U} \frac{\mu(d_1)\mu(d_2)}{[d_1,d_2]} \tag{7.23}$$

tends to

$$\frac{6}{\pi^2} \int_1^\infty \left(\sum_{\substack{r_1,r_2 \le t \\ (r_1,r_2)=1}} \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)}\right) \frac{dt}{t}. \tag{7.24}$$

Notice the similarity, on one side, between (7.24) and the leading term in (7.13), and, on the other, between (7.23) and the expression on the left of (7.9), viz.,

$$\sum_{\substack{d_1,d_2 \\ (d_1 d_2, v)=1}} \frac{\mu(d_1)\mu(d_2)}{[d_1, d_2]} \lambda_{d_1} \lambda_{d_2}, \tag{7.25}$$

where $\lambda_x = 1 - \varrho(x)$. This similarity suggests that the argument sketched in [DIT83, p. 56] can be adapted to give an explicit estimate for (7.25) involving the same quantity $g_v(x)$ as in Prop. 7.3 and Cor. 7.4. The final result would be analogous to Theorem 7.1, with a remainder term proportional to $(U_1 - U_0)^2/(\log U_1/U_0)^2$.

Indeed, we may write (7.25) in the form

$$\sum_{\substack{l \\ (l,v)=1}} \frac{\mu^2(l)}{l} \sum_{\substack{r_1,r_2 \\ (r_1 r_2, lv)=1}} \frac{\mu(r_1 r_2)}{r_1 r_2} \lambda_{l r_1} \lambda_{l r_2}. \tag{7.26}$$

Changing the order of summation, we see that (7.26) equals

$$\sum_{\substack{r_1,r_2 \\ (r_1 r_2, v)=1}} \frac{\mu(r_1 r_2)}{r_1 r_2} \sum_{\substack{l \\ (l, r_1 r_2 v)=1, \mu^2(l)=1}} \frac{\lambda_{l r_1} \lambda_{l r_2}}{l}.$$

We see that this is a sum of very similar type to (7.12). It can be estimated similarly, and yields

$$\frac{6}{\pi^2} \frac{v}{\sigma(v)} \int_1^\infty \sum_{\substack{r_1,r_2 \\ (r_1 r_2, v)=1}} \frac{\mu(r_1 r_2)}{\sigma(r_1)\sigma(r_2)} \lambda_{r_1 t} \lambda_{r_2 t} \frac{dt}{t} \tag{7.27}$$

as the main term. The lower-order terms seem harder to estimate than in Prop. 7.3.

We will not follow on this matter here, save to remark that, for our choice of ϱ, we have $\varrho(x) = \varrho(U_0 U_1/x)$, and so

$$\sum_{\substack{r_1,r_2 \\ (r_1 r_2, v)=1}} \frac{\mu(r_1 r_2)}{\sigma(r_1)\sigma(r_2)} \lambda_{r_1 t} \lambda_{r_2 t} = g_v\left(\frac{U_0 U_1}{t}\right).$$

Setting $s = U_0 U_1/t$, we see that (7.27) equals

$$\frac{6}{\pi^2} \frac{v}{\sigma(v)} \int_1^\infty \sum_{\substack{r_1,r_2 \\ (r_1 r_2, v)=1}} \frac{\mu(r_1 r_2)}{\sigma(r_1)\sigma(r_2)} \lambda_{r_1 t} \lambda_{r_2 t} \frac{dt}{t} = \frac{6}{\pi^2} \frac{v}{\sigma(v)} \int_{U_0}^{U_0 U_1} g_v(s) \frac{ds}{s}.$$

Thus, one should be able to proving an analogue of Theorem 7.1 by means of a detailed study of $g_v(s)$. Our proof of Theorem 7.1 will be based on precisely such a study.

7.3 EXPLICIT BOUNDS ON A SUM INVOLVING μ: GENERAL PLAN

Our task is now to estimate

$$g_v(y) = \sum_{\substack{r_1,r_2\\(r_1,r_2)=1\\(r_1r_2,v)=1}} \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)} \varrho\left(\frac{y}{r_1}\right)\varrho\left(\frac{y}{r_2}\right).$$

It is not too hard to show that it tends to 0. Obtaining good bounds is a more delicate matter. For our purposes, we will need the expression to converge to 0 at least as fast as $1/\log^2$, with a good constant – preferably optimal – in front.

We begin as one might expect, using the Möbius function to remove the link between the variables r_1, r_2:

$$\begin{aligned} g_v(y) &= \sum_{\substack{r_1,r_2\\(r_1r_2,v)=1}} \left(\sum_{d|(r_1,r_2)} \mu(d)\right) \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)} \varrho\left(\frac{y}{r_1}\right)\varrho\left(\frac{y}{r_2}\right) \\ &= \sum_{\substack{d\\(d,v)=1}} \mu(d) \sum_{\substack{r_1,r_2\\d|(r_1,r_2)\\(r_1r_2,v)=1}} \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)} \varrho\left(\frac{y}{r_1}\right)\varrho\left(\frac{y}{r_2}\right) \\ &= \sum_{\substack{d\\(d,v)=1}} \frac{\mu(d)}{(\sigma(d))^2} \left(\sum_{\substack{d'\\(d',dv)=1}} \frac{\mu(d')}{\sigma(d')} \varrho\left(\frac{y}{dd'}\right)\right)^2. \end{aligned} \tag{7.28}$$

It is now time to specify our smoothing function ϱ. We set ϱ as in (7.2):

$$\varrho(t) = \frac{\log^+(t/U_0) - \log^+(t/U_1)}{\log U_1/U_0}. \tag{7.29}$$

This choice is standard, and as far as the leading term will be concerned, optimal. Higher powers of log (as in [Jut79b], [Jut79a]) are better when the sieving problem is of higher "dimension", in the sense of, say, [FI10, §5.5].

Clearly, then, by (7.28),

$$g_v(y) = \frac{g_v(y/U_0, y/U_0) - 2g_v(y/U_0, y/U_1) + g_v(y/U_1, y/U_1)}{(\log U_1/U_0)^2}, \tag{7.30}$$

where $g_v(y_0, y_1)$ is as in (6.39). We will thus be able to use our bounds from §6.4–6.6.1.

Remark. One might ask: why not proceed as in [Gra78]? The procedure there assumes that we have bounds such as $\sum_{n\leq N} \mu(n)/n \ll_A (\log N)^{-A}$ for relatively high powers A. There are simply no good explicit bounds of that quality currently available. If we could bound $1/\zeta(s)$ well for $\Re s < 1$, we would have such bounds, but then proceeding by contour integration as we have just explained would presumably be more efficient.

It does seem possible to derive explicit bounds of the form $\sum_{m\leq N}\mu(n)/n \ll_A (\log N)^{-A}$ by examining the behavior of $\zeta(s)$ for $s \to 1^+$, as in [Iwa14, Ch. IV]. However, the constants seem likely to be too large for such bounds to be practical for $A > 3$, and [Gra78] does require $A > 3$.

There are also bounds of the form $\sum_{m\leq N}\mu(n) \leq c_A N(\log N)^{-A}$, A arbitrary, derived from estimates for $\sum_{m\leq N}\Lambda(n)$ [EM95, Thms. 3–4]. Here c_A is very large, even for A fairly small; for $A = 2$, $c_A = 362.7$. These bounds can be used to prove bounds of the form $\sum_{m\leq N}\mu(n)/n \ll_A (\log N)^{-A}$, but the implied constants are of course very large as well. (The bounds from [Ram13b] and [Ram15] we have been using are very careful, highly optimized estimates of this same kind, with $A = 1$.)

Let us return to what we will actually do. We will bound some tail terms and cross-terms in §7.4. Then, in §7.5, we will be able to estimate our main term

$$\frac{6X}{\pi^2}\frac{v}{\sigma(v)}\sum_{\substack{s\leq X/u\\(s,v)=1}}\frac{1}{s}\int_\beta^1 g_v\left(\frac{tX}{s}\right)dt,$$

from (7.22). Finally, in §7.7, we will bound our remainder terms $R_{\alpha_1,\alpha_2,u,v,\varrho}$. For this last task, a complex-analytic approach does prove feasible.

Before we proceed, let us now look at how $g_v(y)$ behaves in different ranges. If $y \leq U_0$, then it is clear from (7.30) that $g_v(y) = 0$. (Notice that $g_v(y_1, y_2) = 0$ if $y_1 \leq 1$ or $y_2 \leq 1$.) If $U_0 < y \leq U_1$, then it follows from (7.30) that

$$g_v(y) = \frac{g_v(y/U_0, y/U_0)}{(\log U_1/U_0)^2}.$$

In this case, we apply Lemma 6.10 to express $g_v(y/U_0, y/U_0)$ as the sum of $h_v(y/U_0) = h_v(y/U_0, y/U_0)$ and some other terms. It is the last term, $(\sigma(v)/v)\zeta(2)v/\phi(v)$, that will give us the main term overall: it is a constant, and, since $g_v(y)$ will be inside an integral of the form $\int g_v(y)dy/y$, the contribution of this constant will be of size proportional to $\int_{U_0}^{U_1} dy/y = \log U_1/U_0$.

Finally, if $y > U_1$, we will use (6.44) to express $g_v(y)$ in terms of $h_v(y/U_0, y/U_0)$, $h_v(y/U_1, y/U_1)$ and $h_v(y/U_0, y/U_1) = h_v(y/U_1, y/U_0)$. The first and last quantities will be small, whereas $h_v(y/U_1, y/U_1)$ will make a contribution similar to the one that $h_v(y/U_0, y/U_0)$ made before.

7.4 TAIL TERMS AND CROSS-TERMS

Let us see how to bound some lower-order terms that will arise in our estimation of the main quantity we wish to estimate, that is, the sum in the right side of (7.22).

Propositions 6.14 and 6.17 are enough to let us estimate the tail of the sum we mean to study.

Proposition 7.5. *Let $h_v(y) = h_v(y, y)$, where $h_v(y_1, y_2)$ is as in (6.43) and $v = 1$ or $v = 2$. Let $Y \geq 10^9$. Then*

$$\sum_{\substack{1 \leq s < \frac{Y}{10^9} \\ (s,v)=1}} \frac{1}{s} \left| h_v\left(\frac{Y}{s}\right) \right| \leq \begin{cases} 1.289 \cdot 10^{-6} & \text{if } v = 1, \\ 1.261 \cdot 10^{-6} & \text{if } v = 2. \end{cases} \tag{7.31}$$

Proof. By Prop. 6.14,

$$\sum_{\substack{\max(1,Y/10^{12}) \leq s < \frac{Y}{10^9} \\ (s,v)=1}} \frac{1}{s} \left| h_v\left(\frac{Y}{s}\right) \right| \leq \sum_{\substack{\max(1,Y/10^{12}) \leq s < \frac{Y}{10^9} \\ (s,v)=1}} \frac{1}{s} \frac{c_{0,v}}{(Y/s)} \leq \frac{c_{0,v}}{10^9}, \tag{7.32}$$

where $c_{0,1} = 74.554$ and $c_{0,2} = 147.6449$. By Prop. 6.17, if $Y > 10^{12}$, then

$$\sum_{\substack{1 \leq s < \frac{Y}{10^{12}} \\ (s,v)=1}} \frac{1}{s} \left| h_v\left(\frac{Y}{s}\right) \right| \leq \sum_{\substack{1 \leq s < \frac{Y}{10^{12}} \\ (s,v)=1}} \frac{1}{s} \frac{c_{1,v}}{\left(\log \frac{Y}{s}\right)^2},$$

where $c_{1,1} = 0.000033536$, $c_{1,2} = 0.0000615022$. The function $s \mapsto 1/s \log(Y/s)^2$ is decreasing for $s \leq Y/e^2$. Hence, by Euler-Maclaurin ((3.4), (3.10)),

$$\begin{aligned}
\sum_{\substack{1 \leq s < \frac{Y}{10^{12}} \\ (s,v)=1}} \frac{1}{s\left(\log \frac{Y}{s}\right)^2} &= \frac{1}{v} \int_1^{Y/10^{12}} \frac{ds}{s\left(\log \frac{Y}{s}\right)^2} + O^*\left(\frac{1}{\log^2 Y}\right) \\
&= \frac{1}{v} \frac{1}{\log Y/s} \Big|_1^{Y/10^{12}} + O^*\left(\frac{1}{\log^2 Y}\right) \\
&= \frac{1}{v}\left(\frac{1}{\log 10^{12}} - \frac{1}{\log Y}\right) + O^*\left(\frac{1}{\log^2 Y}\right) \leq \frac{1}{v} \frac{1}{\log 10^{12}},
\end{aligned}$$

since $Y > 10^{12} > e^2$. Therefore,

$$\sum_{\substack{1 \leq s < \frac{Y}{10^{12}} \\ (s,v)=1}} \frac{1}{s} \left| h_v\left(\frac{Y}{s}\right) \right| \leq \frac{c_{1,v}}{v} \frac{1}{\log 10^{12}} \leq \begin{cases} 1.21371 \cdot 10^{-6} & \text{if } v = 1, \\ 1.11292 \cdot 10^{-6} & \text{if } v = 2. \end{cases} \tag{7.33}$$

We conclude that, for $Y > 10^9$,

$$\sum_{\substack{1 \leq s < \frac{Y}{10^9} \\ (s,v)=1}} \frac{1}{s} \left| h\left(\frac{Y}{s}\right) \right| \leq \begin{cases} 1.289 \cdot 10^{-6} & \text{if } v = 1, \\ 1.261 \cdot 10^{-6} & \text{if } v = 2. \end{cases} \tag{7.34}$$

□

We should now study cross-terms, namely, terms of the form $h_v(y_1, y_2)$, where $y_1 \neq y_2$, one of y_1, y_2 is small, and the other one may be large or moderate. We may assume without loss of generality that $y_1 > y_2$.

These terms arise in such a way that $y_1 = ry_2$ and $y_2 \geq 1$, where $r = U_1/U_0$. We aim at bounds that are good when U_1/U_0 is larger than a constant; otherwise there would be little point in using the quadratic sieve as opposed to a simpler sum on μ such as the one appearing in [DIT83] and [Helb]. Our bounds on $h_v(y_1, y_2)$ will later go into a sum of the form

$$\sum_{1\leq s\leq S} \frac{1}{s} h_v\left(r\frac{S}{s}, \frac{S}{s}\right).$$

The most crucial range is that of $y_2 = S/s$ very small, between 1 and a constant c, as it is that range where $h_v(ry_2, y_2) = h_v(rS/s, S/s)$ will be largest.

We first note a coarse auxiliary inequality we will use twice. For $0 \leq \sigma < 1$ and $x > 0$,

$$\sum_{\substack{n\leq x\\ n \text{ odd}}} \frac{1}{n^\sigma} \leq x^{1-\sigma} \cdot \begin{cases} \frac{1/2}{1-\sigma} & \text{if } \sigma \geq 1/2, \\ 1 & \text{if } \sigma < 1/2. \end{cases} \tag{7.35}$$

The proof is easy: the inequality is certainly true for $x = 1$; for $x \geq 3$ odd, $1/x^\sigma < \int_{x-2}^{x} dt/t^\sigma = (1/2)x^{1-\sigma}/(1-\sigma)$, and so the inequality holds for x arbitrary.

Lemma 7.6. *Let $h_v(y_1, y_2)$ be as in (6.43) and $v = 1$ or $v = 2$. Let $r, S \geq 1$. Then*

$$\sum_{\substack{s\leq \min\left(S, \frac{rS}{10^6}\right)\\ (s,v)=1}} \frac{1}{s}\left|h_v\left(\frac{rS}{s}, \frac{S}{s}\right)\right| \leq \begin{cases} 0.001977 & \text{if } v = 1, \\ 0.002391 & \text{if } v = 2. \end{cases} \tag{7.36}$$

Proof. By Prop. 6.17, if $S/s > 10^{12}$,

$$\left|h_v\left(\frac{rS}{s}, \frac{S}{s}\right)\right| \leq \frac{c_{1,v}}{\log\frac{rS}{s}\log\frac{S}{s}} \leq \frac{c_{1,v}}{\left(\log\frac{S}{s}\right)^2},$$

where $c_{1,1} = 0.000033536$, $c_{1,2} = 0.0000615022$. We proceed as in the second half of the proof of Prop. 7.5, and obtain

$$\sum_{\substack{1\leq s\leq \frac{S}{10^{12}}\\ (s,v)=1}} \frac{1}{s}\left|h_v\left(\frac{rS}{s}, \frac{S}{s}\right)\right| \leq \begin{cases} 1.21371 \cdot 10^{-6} & \text{if } v = 1, \\ 1.11292 \cdot 10^{-6} & \text{if } v = 2. \end{cases}$$

We can thus focus on the contribution of $s \geq S/10^{12}$. We will have to consider different ranges separately.

Let us first consider the range

$$\min(S/10^6, rS/10^{12}) < s \leq \min(\max(S/10^6, rS/10^{12}), S).$$

It could be that $S/10^6$ is smaller than $rS/10^{12}$. By Lemma 6.19,

$$\sum_{\substack{\frac{S}{10^6}<s\leq \min\left(\frac{rS}{10^{12}},S\right)\\ (s,v)=1}} \frac{1}{s}\left|h_v\left(\frac{rS}{s},\frac{S}{s}\right)\right| \leq \sum_{\substack{\frac{S}{10^{12}}<s\leq \min\left(\frac{rS}{10^{12}},S\right)\\ (s,v)=1}} \frac{1}{s}\frac{\sqrt{c_{2,v}+c_{3,v}\log\frac{S}{s}}}{\log\frac{rS}{s}\sqrt{\frac{S}{s}}}$$

$$\leq \frac{1}{\log 10^{12}}\sum_{\substack{1\leq s\leq S\\ (s,v)=1}} \frac{\sqrt{c_{2,v}+c_{3,v}\log\frac{S}{s}}}{\sqrt{Ss}},$$

where $c_{2,1}=0.000138$, $c_{3,1}=0.00002878$, $c_{2,2}=0.0004747$, $c_{3,2}=0.00010763$. Now, for $a,b>0$,

$$\sum_{1\leq s\leq S} \frac{\sqrt{a+b\log\frac{S}{s}}}{\sqrt{Ss}} \leq \int_0^S \frac{\sqrt{a+b\log\frac{S}{s}}}{\sqrt{Ss}}ds \tag{7.37}$$

because the integrand is non-increasing, and

$$\sum_{\substack{1\leq s\leq S\\ s \text{ odd}}} \frac{\sqrt{a+b\log\frac{S}{s}}}{\sqrt{Ss}} \leq \frac{1}{2}\int_0^S \frac{\sqrt{a+b\log\frac{S}{s}}}{\sqrt{Ss}}ds + \frac{\sqrt{a}}{2S} \tag{7.38}$$

if $(3/4)a^2+ab-b^2/4\geq 0$, since the integrand is then not just non-increasing but convex for $0\leq t\leq S$:

$$\frac{d^2}{ds^2}\frac{\sqrt{a+b\log\frac{S}{s}}}{\sqrt{s}} = \frac{\frac{3}{4}b^2\log^2\frac{S}{s}+\left(\frac{3}{2}ab+b^2\right)\log\frac{S}{s}+\frac{3}{4}a^2+ab-\frac{b^2}{4}}{s^{\frac{5}{2}}\left(a+b\log\frac{S}{s}\right)^{\frac{3}{2}}} > 0.$$

(We are using (3.2) to establish (7.38).) We verify that $(3/4)c_{2,v}^2+c_{2,v}c_{3,v}-c_{3,v}^2\geq 0$ for $v=1,2$.

Moreover, if f is non-increasing and convex on an interval containing S, then $s\mapsto \max(f(x),f(S))$ is also convex on that interval.

By a change of variables $t=\sqrt{a+b\log(S/s)}$,

$$\int_0^S \frac{\sqrt{a+b\log\frac{S}{s}}}{\sqrt{Ss}}ds = \frac{2}{b}e^{\frac{a}{2b}}\int_{\sqrt{a}}^{\infty} t^2 e^{\frac{-t^2}{2b}}dt$$

$$= 2\sqrt{a}+\sqrt{2\pi b}\,e^{\frac{a}{2b}}\left(1-\operatorname{erf}\left(\sqrt{\frac{a}{2b}}\right)\right).$$

Incidentally, it is easy to see that, if $a\geq b$, then, for $S<3$, $(1/2)\cdot 2\sqrt{a}=\sqrt{a}$ is an upper bound for the left side of (7.38). For $S\geq 3$, the term $\sqrt{a}/2S$ on the right side of (7.38) is at most $\sqrt{a}/6$.

Hence, since $c_{2,2} > c_{3,2}$,

$$\frac{1}{\log 10^{12}} \sum_{\substack{1\leq s\leq S\\(s,v)=1}} \frac{\sqrt{c_{2,v}+c_{3,v}\log\frac{S}{s}}}{\sqrt{Ss}} \leq \frac{1}{\log 10^{12}} \cdot \begin{cases} 0.0277151 & \text{if } v=1,\\ 0.0260012 + \frac{0.0217877}{6} & \text{if } v=2.\end{cases}$$
$$\leq \begin{cases} 0.0010031 & \text{if } v=1,\\ 0.0010725 & \text{if } v=2.\end{cases} \tag{7.39}$$

It could also happen that $S/10^6 \geq rS/10^{12}$. By (6.52), Cauchy-Schwarz and the fact that $y \mapsto (\log y)/y$ is decreasing for $y > e$,

$$\sum_{\substack{\frac{rS}{10^{12}}<s\leq\frac{S}{10^6}\\(s,v)=1}} \frac{1}{s}\left|h_v\left(\frac{rS}{s},\frac{S}{s}\right)\right| \leq \sum_{\substack{s\leq\frac{S}{10^6}\\(s,v)=1}} \sqrt{\frac{c_{4,v}+c_{5,v}\log\frac{rS}{s}}{rS}}\sqrt{\frac{c_{4,v}+c_{5,v}\log\frac{S}{s}}{S}}$$
$$\leq \sum_{\substack{s\leq\frac{S}{10^6}\\(s,v)=1}} \frac{c_{4,v}+c_{5,v}\log\frac{S}{s}}{S},$$

where $c_{4,1} = 4.89606$, $c_{5,1} = 3.83717$, $c_{4,2} = 9.57182$, $c_{5,2} = 4.99703$.

Since the numerator here is a decreasing function of s,

$$\sum_{s\leq\frac{S}{10^6}} \frac{c_{4,1}+c_{5,1}\log\frac{S}{s}}{S} \leq \frac{1}{S}\int_0^{S/10^6}\left(c_{4,1}+c_{5,1}\log\frac{S}{s}\right)ds$$
$$= \frac{c_{4,1}+c_{5,1}(1+\log 10^6)}{10^6} \leq 6.17457\cdot 10^{-5},$$

whereas, since $s \mapsto \log S/s$ is decreasing and convex,

$$\sum_{\substack{s\leq\frac{S}{10^6}\\ s \text{ odd}}} \frac{c_{4,2}+c_{5,2}\log\frac{S}{s}}{S}$$
$$\leq \frac{1}{2S}\left(\int_0^{S/10^6}\left(c_{4,2}+c_{5,2}\log\frac{S}{s}\right)ds + c_{4,2}+c_{5,2}\log 10^6\right)$$
$$\leq \frac{c_{4,2}+c_{5,2}(1+\log 10^6)}{2\cdot 10^6} + \frac{c_{4,2}+c_{5,2}\log 10^6}{2\cdot S}$$
$$\leq 4.18027\cdot 10^{-5} + \frac{39.30418}{S} \leq 8.11069\cdot 10^{-5},$$

where we assume $S \geq 10^6$, as the sum we are considering is otherwise empty. We see that our bound on the contribution of this range is much smaller than that in (7.39).

We are done considering the range

$$\min(S/10^6, rS/10^{12}) < s \le \min(\max(S/10^6, rS/10^{12}), S);$$

its contribution is at most

$$\begin{array}{ll} 0.0010031 & \text{if } v = 1, \\ 0.0010725 & \text{if } v = 2, \end{array}$$

as in (7.39). We treated the range $1 \le s \le S/10^{12}$ at the very beginning. Let us now look at the remaining ranges.

Whether $S/10^6$ is larger or smaller than $rS/10^{12}$,

$$\begin{aligned} \sum_{\substack{\frac{S}{10^{12}} \le s \le \min\left(\frac{S}{10^6}, \frac{rS}{10^{12}}\right) \\ (s,v)=1}} \frac{1}{s} \left| h_v\left(\frac{rS}{s}, \frac{S}{s}\right)\right| &\le \sum_{\substack{\frac{S}{10^{12}} \le s \le \min\left(\frac{S}{10^6}, \frac{rS}{10^{12}}\right) \\ (s,v)=1}} \frac{1}{s} \frac{c_{6,v}}{\log \frac{rS}{s} \sqrt{\frac{S}{s}}} \\ &\le \frac{1}{\log 10^{12}} \sum_{\substack{1 \le s \le \frac{S}{10^6} \\ (s,v)=1}} \frac{c_{6,v}}{\sqrt{Ss}} \end{aligned}$$

by Cor. 6.18, where $c_{6,1} = 0.05001$ and $c_{6,2} = 0.0953$. By (7.35), if $v = 2$, or by the fact that $s \to 1/\sqrt{s}$ is decreasing, if $v = 1$,

$$\sum_{\substack{1 \le s \le T \\ (s,v)=1}} \frac{1}{\sqrt{s}} \le \frac{2}{v}\sqrt{T}$$

for any $T > 0$. Hence

$$\frac{1}{\log 10^{12}} \sum_{\substack{1 \le s \le \frac{S}{10^6} \\ (s,v)=1}} \frac{c_{6,v}}{\sqrt{Ss}} \le \frac{c_{6,v}}{\log 10^{12}} \cdot \frac{2}{v} \frac{1}{10^3} \le \begin{cases} 3.7 \cdot 10^{-6} & \text{if } v = 1, \\ 3.5 \cdot 10^{-6} & \text{if } v = 2. \end{cases}$$

Lastly, by Lemma 6.20,

$$\sum_{\substack{\max\left(\frac{S}{10^6}, \frac{rS}{10^{12}}\right) < s \le \min\left(S, \frac{rS}{10^6}\right) \\ (s,v)=1}} \frac{1}{s} \left| h_v\left(\frac{rS}{s}, \frac{S}{s}\right)\right|$$

is at most

$$\sum_{\substack{s \le \frac{rS}{10^6} \\ (s,v)=1}} \frac{1}{s} \left(\frac{c_{7,v}}{(rS/s)^{1/2}} + \frac{c_{8,v}}{(rS/s)^{3/4}}\right),$$

where $c_{7,1} = 0.06219$, $c_{8,1} = 20.01991$, $c_{7,2} = 0.029216$ and $c_{8,2} = 40.61753$. For any $0 < \alpha < 1$,

$$\sum_{s \le \frac{rS}{10^6}} \frac{1}{s} \frac{1}{(rS/s)^\alpha} = \frac{1}{(rS)^\alpha} \sum_{s \le \frac{rS}{10^6}} \frac{1}{s^{1-\alpha}} \le \frac{1/\alpha}{(rS)^\alpha} \left(\frac{rS}{10^6}\right)^\alpha = \frac{1/\alpha}{10^{6\alpha}},$$

and, by (7.35), for $1/2 \le \alpha \le 1$,

$$\sum_{\substack{s \le \frac{rS}{10^6} \\ s \text{ odd}}} \frac{1}{s} \frac{1}{(rS/s)^\alpha} = \frac{1}{(rS)^\alpha} \sum_{\substack{s \le \frac{rS}{10^6} \\ s \text{ odd}}} \frac{1}{s^{1-\alpha}} \le \frac{1}{(rS)^\alpha} \cdot \left(\frac{rS}{10^6}\right)^\alpha = \frac{1}{10^{6\alpha}}.$$

Hence

$$\sum_{\substack{s \le \frac{rS}{10^6} \\ (s,v)=1}} \frac{1}{s} \left(\frac{c_{7,v}}{(rS/s)^{1/2}} + \frac{c_{8,v}}{(rS/s)^{3/4}} \right) \le \begin{cases} 0.0009685 & \text{if } v = 1, \\ 0.0013137 & \text{if } v = 2. \end{cases}$$

Taking totals, we conclude that the sum on the left side of (7.36) is at most

$$\begin{aligned} 1.21371 \cdot 10^{-6} + 0.0010031 + 3.7 \cdot 10^{-6} + 0.0009685 &\le 0.001977 \quad \text{if } v = 1, \\ 1.11292 \cdot 10^{-6} + 0.0010725 + 3.5 \cdot 10^{-6} + 0.0013137 &\le 0.002391 \quad \text{if } v = 2. \end{aligned}$$

□

Lemma 7.7. *Let $h_v(y_1, y_2)$ be as in (6.43) and $v = 1$ or $v = 2$. Let $r, S \ge 1$. Then*

$$\sum_{\substack{\frac{rS}{10^6} < s \le S \\ (s,v)=1}} \frac{1}{s} h_v \left(\frac{rS}{s}, \frac{S}{s} \right) \ge \frac{1}{\sqrt{r}} \cdot \begin{cases} -0.86894 & \text{if } v = 1, \\ -0.73296 & \text{if } v = 2. \end{cases} \tag{7.40}$$

For $v = 1$, if we assume $r \ge 5$, then

$$\sum_{\substack{\frac{rS}{10^6} < s \le S \\ (s,v)=1}} \frac{1}{s} h_v \left(\frac{rS}{s}, \frac{S}{s} \right) \ge \frac{0.74957}{\sqrt{r}}. \tag{7.41}$$

Proof. Let $f_{r,v}$ and $F_{r,v}$ be as in Corollary 6.23. Then, for $s \in [rS/10^6, S]$,

$$\begin{aligned} \frac{1}{s} h_v \left(\frac{rS}{s}, \frac{S}{s} \right) &= \frac{1}{s} h_v \left(\frac{S}{s}, \frac{rS}{s} \right) \\ &\ge \frac{1}{s} \frac{f_{r,v}(S/s)}{\sqrt{rS/s} \cdot \sqrt{S/s}} = \frac{f_{r,v}(S/s)}{\sqrt{r} \cdot S} \ge \frac{-F_{r,v}(S/s)}{\sqrt{r} \cdot S}, \end{aligned}$$

since $f_{r,v}(x) \ge -F_{r,v}(x)$. As $F_{r,v}(x)$ is non-decreasing and non-negative, $s \mapsto -F_{r,v}(S/s)$ is non-decreasing and non-positive. Thus,

$$\begin{aligned} \sum_{\frac{rS}{10^6} < s \le S} \frac{1}{s} h_1 \left(\frac{rS}{s}, \frac{S}{s} \right) &\ge -\frac{1}{\sqrt{r} \cdot S} \sum_{1 \le s \le S} F_1(S/s) \ge -\frac{1}{\sqrt{r} \cdot S} \int_0^S F_1(S/s) ds \\ &= -\frac{1}{\sqrt{r}} \int_1^\infty \frac{F_1(u)}{u^2} du \ge -\frac{0.86894}{\sqrt{r}}, \end{aligned}$$

where we apply Corollary 6.23. Moreover, again by Corollary 6.23 when $r \geq 5$, the constant 0.86894 can be replaced by 0.74957.

Again because $s \mapsto -F_{r,v}(S/s)$ is non-decreasing and non-positive,

$$\begin{aligned}\sum_{\substack{\frac{rS}{10^6}<s\leq S\\ s \text{ odd}}} \frac{1}{s} h_2\left(\frac{rS}{s}, \frac{S}{s}\right) &\geq -\frac{1}{\sqrt{r}\cdot S} \sum_{\substack{1\leq s\leq S\\ s \text{ odd}}} F_2(S/s)\\ &\geq -\frac{1}{2\sqrt{r}\cdot S}\int_0^S F_2(S/s)ds - \frac{1}{2\sqrt{r}\cdot S} F_2(S)\\ &\geq -\frac{1}{2\sqrt{r}}\int_1^\infty \frac{F_2(u)}{u^2}du - \frac{1}{2\sqrt{r}}\max_{S\geq 1}\frac{F_2(S)}{S}\\ &\geq -\frac{1}{2}\cdot\frac{1.03489+0.43102}{\sqrt{r}} = -\frac{0.73296}{\sqrt{r}}.\end{aligned}$$

□

7.5 ESTIMATING THE MAIN TERM

We will now be able to bound the sum in the main term of (7.22), viz.,

$$\sum_{\substack{s\leq tX/U_0\\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{tX}{s}\right) \tag{7.42}$$

for t varying in an interval $[\beta, 1]$.

Let us first split the sum into contributions of different kinds from different ranges. In what follows, we will use repeatedly the following identity, an easy application of Möbius inversion: for $Y \geq 1$, $v \in \mathbb{Z}^+$ arbitrary,

$$\sum_{\substack{s\leq Y\\ (s,v)=1}} \frac{1}{s}\check{m}_v\left(\frac{Y}{s}\right) = \sum_{\substack{s\\ (s,v)=1}} \frac{1}{s}\sum_{\substack{d\\ (d,v)=1}} \frac{\mu(d)}{d}\log^+\frac{Y/s}{d} = \log Y. \tag{7.43}$$

Proposition 7.8. *Let*

$$g_v(y) = \sum_{\substack{r_1\\ (r_1,r_2)=1\\ (r_1r_2,v)=1}}\sum_{r_2} \frac{\mu(r_1)\mu(r_2)}{\sigma(r_1)\sigma(r_2)}\varrho\left(\frac{y}{r_1}\right)\varrho\left(\frac{y}{r_2}\right), \tag{7.44}$$

where ϱ is as in (7.29) for some $U_1 > U_0 > 0$. Let

$$G_v(X, U_0, U_1) = \sum_{\substack{s\leq X/U_0\\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{X}{s}\right). \tag{7.45}$$

Then, for any $X \geq U_1$, $G_v(X, U_0, U_1)$ equals

$$-\frac{\sigma(v)}{\phi(v)}\zeta(2) \sum_{\substack{X/U_1 < s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s} + 2\frac{\sigma(v)}{v}\zeta(2) \sum_{\substack{X/U_1 < s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s}\check{m}_v\left(\frac{X}{sU_0}\right)$$
$$+ \sum_{\substack{s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{X}{sU_0}\right) + \sum_{\substack{s \leq X/U_1 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{X}{sU_1}\right) - 2\sum_{\substack{s \leq X/U_1 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{X}{sU_0}, \frac{X}{sU_1}\right) \tag{7.46}$$

divided by $(\log U_1/U_0)^2$. *For $U_0 \leq X \leq U_1$, $G_v(X, U_0, U_1)$ equals*

$$\frac{\sigma(v)}{v}\zeta(2)\left(2\log\frac{X}{U_0} - \frac{v}{\phi(v)}\sum_{\substack{s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s}\right) + \sum_{\substack{s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{X}{sU_0}\right) \tag{7.47}$$

divided by $(\log U_1/U_0)^2$. *Finally, for $X < U_0$, $G_v(X, U_0, U_1)$ equals* 0.

Proof. It is clear that (7.45) equals 0 for $X < U_0$: the sum is then empty. It is also clear that $g_v(y) = 0$ for $y \leq U_0$.

By (7.30), for $U_0 \leq y \leq U_1$, $g_v(y)$ equals

$$g_v\left(\frac{y}{U_0}, \frac{y}{U_0}\right) / (\log U_1/U_0)^2,$$

which, by Lemma 6.10, equals

$$\frac{1}{(\log U_1/U_0)^2}\left(h_v\left(\frac{y}{U_0}\right) + \frac{\sigma(v)}{v}\zeta(2)\left(2 \cdot \check{m}_v\left(\frac{y}{U_0}\right) - \frac{v}{\phi(v)}\right)\right). \tag{7.48}$$

We apply (7.43) with $Y = X/U_0$, and obtain that (7.47) holds for $U_0 \leq X \leq U_1$.

For $y \geq U_1$, we can use the expression (6.44) for $g_v(y)$:

$$g_v(y) = \frac{h_v(y/U_0, y/U_0) - 2h_v(y/U_0, y/U_1) + h_v(y/U_1, y/U_1)}{(\log U_1/U_0)^2}.$$

We obtain that $G_v(X, U_0, U_1)(\log U_1/U_0)^2$ equals

$$\sum_{\substack{s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{X}{sU_0}\right) - 2\sum_{\substack{s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{X}{sU_0}, \frac{X}{sU_1}\right) + \sum_{\substack{s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{X}{sU_1}\right). \tag{7.49}$$

We realize that

$$\sum_{\substack{s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{X}{sU_1}\right) = \sum_{\substack{s \leq X/U_1 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{X}{sU_1}\right) + \sum_{\substack{X/U_1 < s \leq X/U_0 \\ (s,v)=1}} \frac{1}{s}\frac{\sigma(v)}{\phi(v)}\zeta(2),$$

since $h_v(y, y) = h_v(y) = (\sigma(v)/\phi(v))\zeta(2)$ for $y \leq 1$, by (6.45). Also by (6.45), for $X/U_1 < s \leq X/U_0$,

$$h_v\left(\frac{X}{sU_0}, \frac{X}{sU_1}\right) = -\frac{\sigma(v)}{v}\zeta(2) \cdot \left(\check{m}_v\left(\frac{X}{sU_0}\right) - \frac{v}{\phi(v)}\right).$$

□

The leading term of either (7.46) or (7.47) is the first one. In order to estimate the other sums in the statement of Prop. 7.8, we will use Euler-Maclaurin, in the following form.

Lemma 7.9. *Let $Y > 0$, $T_1 \geq T_0 > 0$, $v \in \{1, 2\}$. Let $f : [T_0, T_1] \to \mathbb{C}$ be of bounded variation. Then*

$$\sum_{\substack{\frac{Y}{T_1} < s \leq \frac{Y}{T_0} \\ (s,v)=1}} \frac{1}{s} f\left(\frac{Y}{s}\right) = \frac{\phi(v)}{v} \int_{T_0}^{T_1} f(t)\frac{dt}{t}$$
$$- B_1\left(\left\{\frac{Y}{vT_0} - \frac{1}{v}\right\}\right)\frac{T_0}{Y} f(T_0) + B_1\left(\left\{\frac{Y}{vT_1} - \frac{1}{v}\right\}\right)\frac{T_1}{Y} f(T_1)$$
$$+ \frac{1}{2Y} O^*\left(\int_{T_0}^{T_1} \left|\frac{d}{dt} t f(t)\right| dt\right). \tag{7.50}$$

If f' is also of bounded variation and well-defined at T_0 and T_1,

$$\sum_{\substack{\frac{Y}{T_1} < s \leq \frac{Y}{T_0} \\ (s,v)=1}} \frac{1}{s} f\left(\frac{Y}{s}\right) = \frac{\phi(v)}{v} \int_{T_0}^{T_1} f(t)\frac{dt}{t}$$
$$+ B_1\left(\left\{\frac{Y}{T_1}\right\}\right)\frac{T_1}{Y} f(T_1) - B_1\left(\left\{\frac{Y}{T_0}\right\}\right)\frac{T_0}{Y} f(T_0)$$
$$+ \frac{v}{16} \cdot O^*\left(\frac{|(tf)'(T_1)|}{(Y/T_1)^2} + \frac{|(tf)'(T_0)|}{(Y/T_0)^2} + \frac{1}{Y^2}\int_{T_0}^{T_1} \left|t(t^2 f)''(t)\right| dt\right). \tag{7.51}$$

Here, as usual, if f has any discontinuities, f' is to be understood in the sense of distributions.

Proof. By Euler-Maclaurin to first order (Lemma 3.2) applied to the function $G(x) =$

$f(Y/x)/x$ (or, if $v=2$, to $x \mapsto G(2x+1)$ instead),

$$\sum_{\substack{\frac{Y}{T_1} < s \leq \frac{Y}{T_0} \\ (s,v)=1}} \frac{1}{s} f\left(\frac{Y}{s}\right) = \frac{\phi(v)}{v} \int_{Y/T_1}^{Y/T_0} G(x)dx$$
$$+ B_1\left(\left\{\frac{Y}{vT_1} - \frac{1}{v}\right\}\right) G\left(\frac{Y}{T_1}\right) - B_1\left(\left\{\frac{Y}{vT_0} - \frac{1}{v}\right\}\right) G\left(\frac{Y}{T_0}\right)$$
$$+ \frac{1}{2} O^*\left(\int_{Y/T_1}^{Y/T_0} |G'(x)|dx\right),$$

and the same is true if the inequality $s > Y/T_1$ is replaced by $s \geq Y/T_1$.

Now,

$$\int_{Y/T_1}^{Y/T_0} G(x)dx = \int_{Y/T_1}^{Y/T_0} f\left(\frac{Y}{x}\right)\frac{dx}{x} = \int_{T_0}^{T_1} f(t)\frac{dt}{t}$$

with the change of variables $t = Y/x$. Clearly

$$G'(x)dx = dG = d\left(\frac{tf(t)}{Y}\right) = \frac{1}{Y}(tf)'(t)dt,$$

and so

$$\int_{\frac{Y}{T_1}}^{\frac{Y}{T_0}} |G'(x)|dx = \frac{1}{Y}\int_{T_0}^{T_1} |(tf)'(t)|dt.$$

Suppose now that f is continuous on $[T_0, T_1]$ and continuously differentiable outside a discrete set of points. We apply second-order Euler-Maclaurin (Lemma 3.2) to G (or to $x \mapsto G(2x+1)$), using the function $F(x) = x^2 - x + 1/8$ instead of $B_2(x)$ (see the comment after the statement of Lemma 3.2). We obtain

$$\sum_{\substack{\frac{Y}{T_1} < s \leq \frac{Y}{T_0} \\ (s,v)=1}} \frac{1}{s} f\left(\frac{Y}{s}\right) = \frac{\phi(v)}{v} \int_{Y/T_1}^{Y/T_0} G(x)dx$$
$$+ B_1\left(\left\{\frac{Y}{vT_1} - \frac{1}{v}\right\}\right) G\left(\frac{Y}{T_1}\right) - B_1\left(\left\{\frac{Y}{vT_0} - \frac{1}{v}\right\}\right) G\left(\frac{Y}{T_0}\right)$$
$$+ \frac{v}{16} O^*\left(\left|G'\left(\frac{Y}{T_1}\right)\right| + \left|G'\left(\frac{Y}{T_0}\right)\right| + \frac{1}{Y^2}\int_{T_0}^{T_1} \left|t(t^2 f)''(t)\right| dt\right).$$

We use the facts that

$$G'(x) = -\frac{f(Y/x)}{x^2} - \frac{Y}{x}\frac{f'(Y/x)}{x^2}$$

and

$$G''(x) = 2\frac{f(Y/x)}{x^3} + 4\frac{Y}{x}\frac{f'(Y/x)}{x^3} + \frac{Y^2}{x^2}\frac{f''(Y/x)}{x^3}.$$

□

Lemma 7.10. *Let $h_v(y) = h_v(y,y)$, where $h_v(y,y)$ is as in (6.43) and $v = 1$ or $v = 2$. Let $Y \geq 1$. Then*

$$\sum_{\substack{1\leq s\leq Y\\(s,v)=1}} \frac{1}{s} h_v\left(\frac{Y}{s}\right) \leq \begin{cases} -0.0495 + 1.6945/Y & \text{if } v = 1,\\ 1.31742 + 3.6174/Y & \text{if } v = 2.\end{cases} \tag{7.52}$$

Moreover,

$$\begin{aligned}&\sum_{\substack{1\leq s\leq Y\\(s,v)=1}} \frac{1}{s}\left(h_v\left(\frac{Y}{s}\right) - \frac{\sigma(v)}{\phi(v)}\zeta(2)\right)\\ &\quad\leq -\frac{\sigma(v)}{v}\frac{\pi^2}{6}\log Y - \begin{cases} 0.998982\left(1 - \frac{1}{Y^{\frac{3}{2}}}\right) & \text{if } v = 1,\\ 1.817075\left(1 - \frac{1}{Y^{\frac{4}{3}}}\right) & \text{if } v = 2.\end{cases}\end{aligned} \tag{7.53}$$

There is nothing particularly meaningful about the exponents $\frac{3}{2}$ and $\frac{4}{3}$ in (7.53). Indeed, the computations below suffice to establish a slightly stronger version of (7.53), with 1.53 and 1.3578 instead of $\frac{3}{2}$ and $\frac{4}{3}$. We choose to work with $\frac{3}{2}$ and $\frac{4}{3}$ for the sake of simplicity.

Proof. By Prop. 7.5, if $Y \geq 10^9$,

$$\sum_{\substack{1\leq s<\frac{Y}{10^9}\\(s,v)=1}} \frac{1}{s}\left|h_v\left(\frac{Y}{s}\right)\right| \leq \begin{cases} 1.289\cdot 10^{-6} & \text{if } v = 1,\\ 1.261\cdot 10^{-6} & \text{if } v = 2.\end{cases}$$

We can thus let $T_1 = \min(10^9, Y)$, and focus on the terms with $Y/T_1 \leq s \leq Y$.

By Lemma 7.9,

$$\begin{aligned}\sum_{\substack{\frac{Y}{T_1}\leq s\leq Y\\(s,v)=1}} \frac{1}{s} h_v\left(\frac{Y}{s}\right) &= \frac{\phi(v)}{v}\int_1^{T_1} h_v(t)\frac{dt}{t} + \frac{T_1}{2Y}O^*(h_v(T_1)) + \frac{1}{2Y}O^*(h_v(1))\\ &\quad+ \frac{1}{2Y}O^*\left(\int_1^{T_1}\left|\frac{d}{dt}th_v(t)\right|dt\right).\end{aligned}$$

By Lemma 6.10, for $t \geq 1$,

$$h_v(t) = g_v(t,t) - 2\frac{\sigma(v)}{v}\zeta(2)\check{m}_v(t) + \frac{\sigma(v)}{\phi(v)}\zeta(2),$$

and so $h_v(1) = (\sigma(v)/\phi(v))\zeta(2)$. By Cor. 6.27,

$$T_1|h_v(T_1)| + \int_1^{T_1}\left|\frac{d}{dt}th_v(t)\right|dt \leq \begin{cases} 7.05 & \text{if } v = 1,\\ 17.64 & \text{if } v = 2.\end{cases} \tag{7.54}$$

By Prop. 6.26, for $10^6 \le T \le 10^9$,

$$\frac{\phi(v)}{v}\int_1^T h_v(t)\frac{dt}{t} \le \begin{cases} -0.0495099 & \text{if } v=1,\\ 1.3174065 & \text{if } v=2.\end{cases} \tag{7.55}$$

Hence

$$\sum_{\substack{1\le s\le Y\\(s,v)=1}} \frac{1}{s} h_v\left(\frac{Y}{s}\right) \le \begin{cases} -0.049508 + 4.348/Y & \text{if } v=1 \text{ and } Y\ge 10^6,\\ 1.317408 + 11.288/Y & \text{if } v=2 \text{ and } Y\ge 10^6.\end{cases} \tag{7.56}$$

By (3.25) and (3.27), we conclude that, for $Y \ge 10^6$,

$$\begin{aligned}&\sum_{\substack{1\le s\le Y\\(s,v)=1}} \frac{1}{s}\left(h_v\left(\frac{Y}{s}\right) - \frac{\sigma(v)}{\phi(v)}\zeta(2)\right)\\ &\qquad\le -\frac{\sigma(v)}{v}\frac{\pi^2}{6}\log Y + \begin{cases} -0.998989 + 5.993/Y & \text{if } v=1,\\ -1.8170866 + 16.223/Y & \text{if } v=2.\end{cases}\end{aligned} \tag{7.57}$$

It is clear that, for $Y \ge 10^6$, the bounds in (7.56) are stronger than those in (7.52), and the first bound in (7.57) is stronger than that the first bound in (7.53); for $Y \ge 1.5\cdot 10^6$, the second bound in (7.57) is stronger than that in (7.53). Now it just remains to treat the case of $Y < 10^6$ (or $Y < 1.5\cdot 10^6$). Let us see how to compute the sum on the left of (7.56) – call it $H_v(Y)$ - for Y bounded. Computing $\sum_{1\le s\le Y:\,(s,v)=1} 1/s$ for Y bounded is of course trivial.

The first thing to notice here is that, since $h_v(y)$ can be written in the form (6.102), and since, for s a given positive integer, $\lfloor Y/s\rfloor$ equals $\lfloor \lfloor Y\rfloor/s\rfloor$, which depends only on $\lfloor Y\rfloor$ and s,

$$\begin{aligned} H_v(Y) &= \sum_{\substack{1\le s\le \lfloor Y\rfloor\\(s,v)=1}} \frac{1}{s}\sum_{j=0}^{2} \kappa_{j,v}\left(\left\lfloor\frac{Y}{s}\right\rfloor\right)\left(\log\frac{Y/s}{\lfloor\lfloor Y\rfloor/s\rfloor}\right)^j\\ &= \sum_{\substack{1\le s\le \lfloor Y\rfloor\\(s,v)=1}} \frac{1}{s}\sum_{j=0}^{2} \kappa_{j,v}\left(\left\lfloor\frac{\lfloor Y\rfloor}{s}\right\rfloor\right)\left(\log\frac{\lfloor Y\rfloor/s}{\lfloor\lfloor Y\rfloor/s\rfloor} + \log\frac{Y}{\lfloor Y\rfloor}\right)^j,\end{aligned} \tag{7.58}$$

where $\kappa_{v,j}$ are as in (6.102). In particular, we can express $H_v(Y)$ as a linear combination of 1, $\log(Y/\lfloor Y\rfloor)$ and $\log^2(Y/\lfloor Y\rfloor)$. We can compute $\kappa_{v,j}$ as we explained in the proof of Prop. 6.26. In this way we obtain computationally that, for $2 \le Y \le 1.5\cdot 10^6$, (7.52) and (7.53) hold. The lower-order terms in (7.52) (namely, $1.6945/Y$ and $3.6174/Y$) are not tight in that range; they will be close to tight when $Y \to 1^+$.

It remains to check (7.52) and (7.53) for $1 \le Y \le 2$. The matter is routine. We use the fact that $g_v(y) = (\log y)^2$ for $1 \le y \le 2$ and so, by Lemma 6.10,

$$h_v(y) = (\log y)^2 - 2\frac{\sigma(v)}{v}\zeta(2)\log y + \frac{\sigma(v)}{\phi(v)}\zeta(2). \tag{7.59}$$

Of course, since $1 \leq Y \leq 2$,

$$\sum_{\substack{1\leq s\leq Y\\(s,v)=1}} \frac{1}{s} h_v\left(\frac{Y}{s}\right) = h_v(y).$$

We show that the right side of (7.52) is greater than the right side of (7.59) throughout the interval $[1,2]$ using a trivial version of the bisection method (§4.1.1). Since the comparison is rather tight at $y=0$, we could instead compare their values at $y=0$ and then use the same version of the bisection method to compare their derivatives throughout $[0,1]$. To show that the right side of (7.53) (minus $(\sigma(v)/\phi(v))\zeta(2)$) is greater than or equal than the right side of (7.59) on $[1,2]$, we combine the two approaches: since the two functions being compared equal each other at 1, we apply the bisection method twice, once to compare their derivatives on, say, $[1, 9/8]$, and once to compare the two functions on the complement, namely, $[9/8, 2]$. □

The following estimation will be a little delicate – in some sense, more delicate than that of sums involving h_v: we are generally relying on the estimate (5.49) on $\check{m}_v(t) - v/\phi(v)$ for large values of t, and that bound is of the form $c/\log t$. We will have to bound

$$\int_1^T \left(\check{m}_v(t) - \frac{v}{\phi(v)}\right)\frac{dt}{t}, \tag{7.60}$$

and we do not want to bound it by c times

$$\int_1^T \frac{1}{\log t}\frac{dt}{t}$$

because that integral diverges as T goes to infinity. We will deal with this situation by proceeding differently for large and small values of T; for large T, we will use an identity to rephrase the integral in (7.60). Let us do the case of large T first.

We have made frequent use of the function $\check{m}_q$, defined in (5.45). Now we will also need m_q and $\check{\check{m}}_q$, the other functions defined in (5.45):

$$m_q(x) = \sum_{\substack{n\leq x\\(n,q)=1}} \frac{\mu(n)}{n}, \qquad \check{\check{m}}_q(x) = \sum_{\substack{n\leq x\\(n,q)=1}} \frac{\mu(n)}{n}\left(\log\frac{x}{n}\right)^2. \tag{7.61}$$

Proposition 7.11. *Let $\check{m}_v(y)$ be as in (5.45), where $v=1$ or $v=2$. Let $10^{12} \leq T \leq Y$. Then*

$$\sum_{\substack{\frac{Y}{T}<s\leq Y\\(s,v)=1}} \frac{1}{s}\check{m}_v\left(\frac{Y}{s}\right) + \frac{v}{\phi(v)} \sum_{\substack{s\leq Y/T\\(s,v)=1}} \frac{1}{s} = \log Y + \mathrm{err}_v(T,Y), \tag{7.62}$$

where

$$|\mathrm{err}_v(T,Y)| \leq \begin{cases} \frac{0.006299}{\log T} \leq 0.00022797 & \text{if } v=1\\ \frac{0.008092}{\log T} \leq 0.00029286 & \text{if } v=2. \end{cases} \tag{7.63}$$

It is tempting to start by applying (7.43), and obtain

$$\sum_{\substack{\frac{Y}{T}<s\leq Y\\(s,v)=1}} \frac{1}{s}\check{m}_v\left(\frac{Y}{s}\right) + \frac{v}{\phi(v)}\sum_{\substack{s\leq Y/T\\(s,v)=1}}\frac{1}{s} = \log Y - \sum_{\substack{s\leq Y/T\\(s,v)=1}}\frac{1}{s}\left(\check{m}_v\left(\frac{Y}{s}\right)-\frac{v}{\phi(v)}\right). \tag{7.64}$$

As we were saying, following this route would lead us to a divergent sum when Y is very large, and so we must find another way. However, we can already see here that the main term will be $\log Y$. We will use (7.64) later, for small T – and, in an auxiliary role, even in the proof below, for T very close to Y.

Proof. By the first inequality in Lemma 7.9,

$$\sum_{\substack{\frac{Y}{T}<s\leq Y\\(s,v)=1}} \frac{1}{s}\left(\check{m}_v\left(\frac{Y}{s}\right)-\frac{v}{\phi(v)}\right) = \frac{\phi(v)}{v}\int_1^T f_v(t)\frac{dt}{t} + \frac{1}{2Y}O^*\left(\int_1^T\left|\frac{d}{dt}tf_v(t)\right|dt\right) + \frac{T}{Y}B_1\left(\left\{\frac{Y}{vT}-\frac{1}{v}\right\}\right)f_v(T) + \frac{1}{Y}B_1\left(\left\{\frac{Y}{v}-\frac{1}{v}\right\}\right) \tag{7.65}$$

where $f_v(t) = \check{m}_v(t) - v/\phi(v)$. We know that, for $m \leq y \leq m+1$,

$$\check{m}_v(y) = \check{m}_v(m) + m_v(m)\log\frac{y}{m}.$$

Hence

$$\frac{d}{dt}tf_v(t) = f_v(t) + tf_v'(t) = f_v(t) + m_v(\lfloor t\rfloor),$$

and so

$$\int_1^T\left|\frac{d}{dt}tf_v(t)\right|dt = \int_1^T |f_v(t)+m_v(\lfloor t\rfloor)|\,dt = \int_1^T |\check{m}_v(t)-1+m_v(t)|\,dt.$$

Now

$$\int_1^T \check{m}_v(t)\frac{dt}{t} = \left(\sum_{\substack{n\leq T\\(n,v)=1}}\frac{\mu(n)}{n}\int_n^T\frac{1}{t}\log\frac{t}{n}dt\right) = \frac{1}{2}\sum_{\substack{n\leq T\\(n,v)=1}}\frac{\mu(n)}{n}\left(\log\frac{T}{n}\right)^2 = \frac{1}{2}\check{\check{m}}_v(T),$$

and so

$$\int_1^T f_v(t)\frac{dt}{t} = \int_1^T\left(\check{m}_v(t)-\frac{v}{\phi(v)}\right)\frac{dt}{t} = \frac{1}{2}\check{\check{m}}_v(T) - \frac{v}{\phi(v)}\log T.$$

By (3.30) and (3.33),

$$\frac{v}{\phi(v)} \sum_{\substack{s\leq Y\\(s,v)=1}} \frac{1}{s} = \log Y + \gamma + \log v - \frac{B_1(\{Y/v - 1/v\})}{Y} + O^*\left(\frac{v}{Y^2}\right). \tag{7.66}$$

The terms $(1/Y)B_1(\{Y/v - 1/v\})$ in (7.65) and (7.66) cancel each other. We conclude that

$$\begin{aligned}&\sum_{\substack{\frac{Y}{T}<s\leq Y\\(s,v)=1}} \frac{1}{s}\check{m}_v\left(\frac{Y}{s}\right) + \sum_{\substack{s\leq \frac{Y}{T}\\(s,v)=1}} \frac{1}{s}\frac{v}{\phi(v)}\\ &= \sum_{\substack{\frac{Y}{T}<s\leq Y\\(s,v)=1}} \frac{1}{s}\left(\check{m}_v\left(\frac{Y}{s}\right) - \frac{v}{\phi(v)}\right) + \frac{v}{\phi(v)} \sum_{\substack{s\leq Y\\(s,v)=1}} \frac{1}{s}\\ &= \frac{\phi(v)}{2v}\check{\check{m}}_v(T) + \log\frac{Y}{T} + \gamma + \log v\\ &+ \frac{1}{2Y}O^*\left(\int_1^T |\check{m}_v(t) - 1 + m_v(t)|\,dt\right) + O^*\left(\frac{1}{2Y} + \frac{v}{Y^2}\right).\end{aligned}$$

We can write

$$\mathrm{err}_{1,v}(T) = \frac{\phi(v)}{2v}\check{\check{m}}_v(T) - \log\frac{T}{v} + \gamma \tag{7.67}$$

and

$$\mathrm{err}_{2,v}(T) = \frac{1}{2T}\int_1^T |\check{m}_v(t) - 1 + m_v(t)|\,dt. \tag{7.68}$$

We have just shown that

$$\begin{aligned}\sum_{\substack{\frac{Y}{T}<s\leq Y\\(s,v)=1}} \frac{1}{s}\check{m}_v\left(\frac{Y}{s}\right) + \frac{v}{\phi(v)} \sum_{\substack{s\leq \frac{Y}{T}\\(s,v)=1}} \frac{1}{s}&\\ = \log Y + \mathrm{err}_{1,v}(T)&\\ + O^*\left(\frac{\mathrm{err}_{2,v}(T)}{Y/T}\right) + O^*\left(\frac{v}{Y^2}\right),&\end{aligned} \tag{7.69}$$

so it just remains to estimate $\mathrm{err}_{1,v}(T)$ and $\mathrm{err}_{2,v}(T)$.

By (5.50) and (5.91),

$$|\mathrm{err}_{1,v}(T)| \leq \frac{1}{\log T}\cdot\begin{cases}\frac{1}{206} & \text{if } v=1 \text{ and } T\geq 9,\\ \frac{1}{200} & \text{if } v=2 \text{ and } T\geq 2209.\end{cases} \tag{7.70}$$

It remains to bound $\mathrm{err}_{2,v}(T)$. Let $C \geq 96955$. By (5.48), (5.49), (5.89) and (5.90), for $t \geq C$,

$$|f_v(t)| \leq \frac{k_{1,v}}{\log t}, \qquad |m_v(t)| \leq \frac{k_{2,v}}{\log t}, \tag{7.71}$$

where

$$k_{1,v} = \begin{cases} \frac{1}{389} & \text{if } v = 1, \\ \frac{2}{379} & \text{if } v = 2, \end{cases} \qquad k_{2,v} = \begin{cases} 0.0144 & \text{if } v = 1, \\ 0.0296 & \text{if } v = 2. \end{cases}$$

Hence, for $T \geq C$,

$$\int_C^T |f_v(t)| dt \leq k_{1,v} \int_C^T \frac{dt}{\log t}, \qquad \int_C^T |m_v(t)| dt \leq k_{2,v} \int_C^T \frac{dt}{\log t}.$$

We will bound the logarithmic integrals here as follows: for $t \geq C$ and $c \geq 1$,

$$\left(\frac{t}{\log t} + \frac{ct}{(\log t)^2} \right)' = \frac{1}{\log t} - \frac{1}{(\log t)^2} + \frac{c}{(\log t)^2} - \frac{2c}{(\log t)^3} \geq \frac{1}{\log t}$$

provided that $c - 1 \geq 2c/\log C$, and so, for $T \geq C$,

$$\int_C^T \frac{dt}{\log t} \leq \frac{t}{\log t} \left(1 + \frac{c}{\log t} \right) \Big|_C^T. \tag{7.72}$$

We let $c = 1/(1 - 2/\log C)$, $C = 10^{12}$.

For $11 \leq R \leq 10^{12}$, by (5.54) and (5.62),

$$\begin{aligned} \int_{11}^R |\check{m}_1(t) - 1 + m_1(t)| \, dt &\leq \int_{11}^R \left(\frac{0.0234188}{\sqrt{x}} + \frac{0.569449}{\sqrt{x}} \right) dx \\ &\leq 1.18574 \left(\sqrt{R} - \sqrt{11} \right). \end{aligned} \tag{7.73}$$

By a simple calculation,

$$\begin{aligned} \int_1^{11} |\check{m}_1(t) - 1 + m_1(t)| \, dt &\leq \sum_{n=1}^{10} \int_n^{n+1} \left| \check{m}_1(n) - 1 + m_1(n) \left(1 + \log \frac{t}{n} \right) \right| dt \\ &\leq \sum_{n=1}^{10} \left(|\check{m}_1(n) - 1 + m_1(n)| + |m_1(n)| \int_n^{n+1} \log \frac{t}{n} dt \right) \leq 1.21379, \end{aligned} \tag{7.74}$$

where we express the integral as in (6.125).

Hence, for $T \geq C = 10^{12}$, $|\operatorname{err}_{2,1}(T)|$ is at most

$$\frac{1.21379 + 1.18574 \left(\sqrt{C} - \sqrt{11} \right)}{2T} + \frac{k_{1,1} + k_{2,1}}{2 \log t} \left(1 + \frac{c}{\log t} \right) \Big|_C^T + \frac{k_{1,1}}{2 \log T}. \tag{7.75}$$

Since, for $T \geq C = 10^{12}$, as a quick calculation shows,

$$\frac{1.21379 + 1.18574 \left(\sqrt{C} - \sqrt{11} \right)}{2T} < \frac{k_{1,1} + k_{2,1}}{2 \log C} \left(1 + \frac{c}{\log C} \right), \tag{7.76}$$

the expression in (7.75) is at most

$$\frac{1}{2\log T}\left(2k_{1,1}+k_{2,1}+(k_{1,1}+k_{2,1})\frac{c}{\log T}\right)\leq\frac{c_1}{\log T}, \tag{7.77}$$

where $c_1 = (2k_{1,1} + k_{2,1} + (k_{1,1} + k_{2,1}) \cdot c/\log C)/2 \leq 0.01011$.

Likewise, by (5.79) and (5.80), for $R \geq 2001$,

$$\begin{aligned}\int_1^R |\check{m}_2(t)-2+m_2(t)|\,dt &\leq \int_1^{2001} |\check{m}_2(t)-2+m_2(t)|\,dt\\ &\quad+\int_{2001}^R\left(\frac{0.068199}{\sqrt{x}}+\frac{0.390056}{\sqrt{x}}\right)dx\\ &\leq 10.68573+0.91651(\sqrt{R}-\sqrt{2001}),\end{aligned} \tag{7.78}$$

where we bound the integral from 1 to 2001 much as in (7.74). Just as in (7.76), a quick calculation shows that, for $T \geq C = 10^{12}$,

$$\frac{10.68573+0.91651\left(\sqrt{C}-\sqrt{2001}\right)}{2T}<\frac{k_{1,2}+k_{2,2}}{2\log C}\left(1+\frac{c}{\log C}\right), \tag{7.79}$$

We conclude, just as before, that, for $T \geq 10^{12}$,

$$|\operatorname{err}_{2,2}(T)|\leq\frac{c_2}{\log T},$$

where $c_2 = (2k_{1,2} + k_{2,2} + (k_{1,2} + k_{2,2})c/\log C)/2 \leq 0.02076$.

We must now set out to estimate our total error term

$$\operatorname{err}_v(T,Y)=\sum_{\substack{\frac{Y}{T}<s\leq Y\\(s,v)=1}}\frac{1}{s}\check{m}_v\left(\frac{Y}{s}\right)+\frac{v}{\phi(v)}\sum_{\substack{s\leq Y/T\\(s,v)=1}}\frac{1}{s}-\log Y. \tag{7.80}$$

If $Y/r < T \leq Y/(r-v)$, where $(r,v) = 1$, then, by (7.64),

$$\begin{aligned}\sum_{\substack{\frac{Y}{T}<s\leq Y\\(s,v)=1}}\frac{1}{s}\check{m}_v\left(\frac{Y}{s}\right)+\frac{v}{\phi(v)}\sum_{\substack{s\leq Y/T\\(s,v)=1}}\frac{1}{s}&=\log Y-\sum_{\substack{s<r\\(s,v)=1}}\frac{1}{s}\left(\check{m}_v\left(\frac{Y}{s}\right)-\frac{v}{\phi(v)}\right)\\ &=\log Y+O^*\left(\frac{k_{1,v}}{\log T}\cdot\sum_{\substack{s<r\\(s,v)=1}}\frac{1}{s}\right),\end{aligned}$$

where we use the first bound in (7.71) (which quotes (5.49) and (5.90)). It is easy to see that, for $r \leq 7$, the error term here is

$$\begin{aligned}&0.006299/\log T \quad \text{if } v=1,\\ &0.008092/\log T \quad \text{if } v=2.\end{aligned} \tag{7.81}$$

If, on the other hand, $T \le Y/7$, then, we use our bounds on $\mathrm{err}_{1,v}(T)$ and $\mathrm{err}_{2,v}(T)$: by (7.69), (7.70), and $|\,\mathrm{err}_{2,v}(T)| \le c_v/\log T$,

$$\begin{aligned}|\mathrm{err}_v(T,Y)| &\le |\mathrm{err}_{1,v}(T)| + \left|\frac{1}{7}\,\mathrm{err}_{2,v}(T)\right| + \left|\frac{v}{(7T)^2}\right| \\ &\le \begin{cases} \frac{1}{\log T}\left(\frac{1}{206} + \frac{0.01011}{7}\right) + \frac{1}{49T^2} & \text{if } v = 1, \\ \frac{1}{\log T}\left(\frac{1}{200} + \frac{0.02076}{7}\right) + \frac{2}{49T^2} & \text{if } v = 2, \end{cases}\end{aligned}$$

which, given that $T \ge 10^{12}$, is a stronger bound than (7.81) for either $v = 1$ or $v = 2$. □

Now we should prove a result complementary to that of Prop. 7.11 – that is, we must estimate the same quantity as in Prop. 7.11, but for T small. The following simple auxiliary result will be helpful.

Lemma 7.12. *Let $X \ge 1$, $\alpha \in (0,1)$. Then*

$$\int_\alpha^1 \frac{1}{u} \sum_{\substack{n \le uX \\ n \text{ odd}}} 1 \cdot du = \frac{1-\alpha}{2} X + O^*\left(\frac{1+1/\alpha}{8X}\right).$$

The idea is that, instead of estimating a sum first and then integrating the result, we can use the integral as a form of smoothing, and then estimate a smoothed sum.

Proof. Clearly

$$\begin{aligned}\int_\alpha^1 \frac{1}{u} \sum_{\substack{n \le uX \\ n \text{ odd}}} 1 \cdot du &= \sum_{n \text{ odd}} \int_{\max(\alpha, n/X)}^1 \frac{du}{u} \\ &= \sum_{n \text{ odd}} f(n) = \frac{1}{2} \sum_{\substack{n \in \mathbb{Z} \\ n \text{ odd}}} f(n),\end{aligned}$$

where

$$f(t) = \begin{cases} 0 & \text{if } |t| \ge X, \\ \log(X/|t|) & \text{if } \alpha X \le |t| \le X, \\ \log(1/\alpha) & \text{if } |t| \le \alpha X. \end{cases}$$

(As usual, in a sum over "n odd", n ranges over odd positive numbers; we write $n \in \mathbb{Z}$ when we want to specify that n takes negative values as well.) The function f is continuous and piecewise differentiable. Its derivative $f'(t)$ equals $-1/t$ for $\alpha X \le |t| \le X$ and 0 otherwise. Hence, $|f''|_1 = 4/\alpha X$.

Applying Euler-Maclaurin to second order (in the form (3.11)), we obtain

$$\begin{aligned}\frac{1}{2} \sum_{\substack{n \in \mathbb{Z} \\ n \text{ odd}}} f(n) &= \frac{1}{4} \int_{-\infty}^{\infty} f(x)dx + O^*\left(\frac{1}{16}|f''|_1\right) \\ &= \frac{1-\alpha}{2} X + O^*\left(\frac{1}{4\alpha X}\right).\end{aligned}$$

□

Now we prove a result complementing Prop. 7.11.

Proposition 7.13. *Let $\check{m}_v(y)$ be as in (5.45), where $v = 1$ or $v = 2$. Let $1 \leq T \leq \min\left(Y, 10^{12}\right)$. Then*

$$\sum_{\substack{\frac{Y}{T} < s \leq Y \\ (s,v)=1}} \frac{1}{s} \check{m}_v\left(\frac{Y}{s}\right) + \frac{v}{\phi(v)} \sum_{\substack{s \leq Y/T \\ (s,v)=1}} \frac{1}{s} = \log Y + \mathrm{err}_v(T, Y), \tag{7.82}$$

where

$$|\mathrm{err}_v(T,Y)| \leq \begin{cases} \frac{1}{T} + \frac{0.0399648}{\sqrt{T}} + 0.00022797 & \text{if } v = 1, \\ \frac{2.9}{T} + \frac{0.017351}{\sqrt{T}} + 0.00029286 & \text{if } v = 2. \end{cases} \tag{7.83}$$

Moreover, for $v = 2$ and $\beta \in \{0, 1/2\}$ and $1 \leq T \leq \min(X, 10^{12})$,

$$\left| \int_{\max(\beta, T/X)}^{1} \mathrm{err}_v(T, tX) dt \right| \tag{7.84}$$

is at most

$$(1-\beta) \cdot \left(\frac{0.017351}{\sqrt{T}} + 0.00029286 + \frac{2.9}{T} \cdot \begin{cases} \log 2 & \text{if } \beta = 1/2 \\ 1/2 & \text{if } \beta = 0 \end{cases} \right). \tag{7.85}$$

As we can see, one of the error terms is improved by a factor of $\log 2$ (if $\beta = 1/2$) or $1/2$ (if $\beta = 1$) when we integrate, relative to what one would expect from (7.83). The reason is that, as we were saying before, integration naturally introduces a form of smoothing, and thus helps with error terms.

Proof. If $Y < 10^{12}$, we use (7.64). If $Y \geq 10^{12}$, we remark first that

$$\begin{aligned} \sum_{\substack{\frac{Y}{T} < s \leq Y \\ (s,v)=1}} \frac{1}{s} \check{m}_v\left(\frac{Y}{s}\right) + \frac{v}{\phi(v)} \sum_{\substack{s \leq Y/T \\ (s,v)=1}} \frac{1}{s} &= \sum_{\substack{\frac{Y}{10^{12}} < s \leq Y \\ (s,v)=1}} \frac{1}{s} \check{m}_v\left(\frac{Y}{s}\right) + \frac{v}{\phi(v)} \sum_{\substack{s \leq \frac{Y}{10^{12}} \\ (s,v)=1}} \frac{1}{s} \\ &\quad - \sum_{\substack{\frac{Y}{10^{12}} < s \leq \frac{Y}{T} \\ (s,v)=1}} \frac{1}{s} \left(\check{m}_v\left(\frac{Y}{s}\right) - \frac{v}{\phi(v)} \right), \end{aligned}$$

and apply Proposition 7.11 with 10^{12} instead of T. In either case, we must bound

$$\sum_{\substack{\frac{Y}{10^{12}} < s \leq \frac{Y}{T} \\ (s,v)=1}} \frac{1}{s} \left(\check{m}_v\left(\frac{Y}{s}\right) - \frac{v}{\phi(v)} \right). \tag{7.86}$$

By (5.55) and (5.81), for $x \leq 10^{12}$,

$$\left|\check{m}_v(x) - \frac{v}{\phi(v)}\right| \leq \frac{k_{3,v}}{x} + \frac{k_{4,v}}{\sqrt{x}},$$

where $k_{3,1} = 1$, $k_{4,1} = 0.023418$, $k_{3,2} = 2.9$, $k_{4,2} = 0.014519$. Hence the sum in (7.86) is at most

$$\sum_{\substack{s \leq \frac{Y}{T} \\ (s,v)=1}} \left(\frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Ys}}\right). \tag{7.87}$$

For $v = 1$, we simply observe that, since $s \mapsto 1/\sqrt{s}$ is decreasing,

$$\sum_{s \leq \frac{Y}{T}} \left(\frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Ys}}\right) \leq \int_0^{Y/T} \left(\frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Ys}}\right) ds = \frac{k_{3,v}}{T} + \frac{2k_{4,v}}{\sqrt{T}}.$$

For $v = 2$ and $T \leq Y/3$,

$$\sum_{\substack{s \leq \frac{Y}{T} \\ (s,v)=1}} \left(\frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Ys}}\right) \leq \frac{1}{2}\int_0^{Y/T} \left(\frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Ys}}\right) ds + \frac{1}{2}\left(\frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Y}}\right)$$

$$= \frac{k_{3,v}}{2T} + \frac{k_{4,v}}{\sqrt{T}} + \frac{1}{2}\left(\frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Y}}\right) \leq \frac{1+1/3}{2} \cdot \frac{k_{3,v}}{T} + \frac{2 + \frac{1}{\sqrt{3}}}{2} \cdot \frac{k_{4,v}}{\sqrt{T}},$$

where we use the convexity of $s \mapsto 1/\sqrt{s}$. Finally, if $v = 2$ and $Y/3 < T \leq Y$,

$$\sum_{\substack{s \leq \frac{Y}{T} \\ (s,v)=1}} \left(\frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Ys}}\right) = \frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Y}} \leq \frac{k_{3,v}}{T} + \frac{k_{4,v}}{\sqrt{T}}.$$

Hence, for $v = 2$ and any $1 \leq T \leq \min(Y, 10^{12})$,

$$\sum_{\substack{s \leq \frac{Y}{T} \\ (s,v)=1}} \left(\frac{k_{3,v}}{Y} + \frac{k_{4,v}}{\sqrt{Ys}}\right) \leq \frac{k_{3,v}}{T} + \left(1 + \frac{1/2}{\sqrt{3}}\right)\frac{k_{4,v}}{\sqrt{T}}.$$

We have obtained (7.82)–(7.83). Let us prove the bound (7.85) on (7.84). Given our work above, it is enough to give a bound on

$$\int_\alpha^1 \sum_{\substack{s \leq \frac{tX}{T} \\ (s,v)=1}} \frac{k_{3,v}}{tX} dt \tag{7.88}$$

for $v = 2$, where $\alpha = \max(\beta, T/X)$. Applying Lemma 7.12, we see that (7.88) is at most

$$\frac{k_{3,v}}{X}\left(\frac{1-\alpha}{2}\frac{X}{T} + O^*\left(\frac{1}{4\alpha X/T}\right)\right) \leq \left(\frac{1-\alpha}{2} + \frac{1}{4\alpha r^2}\right)\frac{k_{3,v}}{T}, \tag{7.89}$$

where $r = X/T$.

Let us consider first the case $r \geq 3$. If $\beta \geq T/X$, then $\alpha = \beta$, and so we get that the expression in (7.88) is at most

$$\left(\frac{1-\beta}{2} + \frac{1}{4\beta \cdot 3^2}\right)\frac{k_{3,v}}{T}.$$

If $\beta < T/X$, then $\alpha = T/X = 1/r \leq 1/3$, and so we obtain that (7.88) is at most

$$\left(\frac{1-\alpha}{2} + \frac{\alpha}{4}\right)\frac{k_{3,v}}{T} < \frac{k_{3,v}}{2T}. \tag{7.90}$$

Assume now that $1 \leq r < 3$. Then there is only one odd integer $s \leq r = X/T$, namely, 1. Hence, (7.88) equals

$$\int_\alpha^1 \frac{k_{3,v}}{tX}dt = \frac{k_{3,v}}{X}|\log \alpha|.$$

If $\beta \geq T/X$, then $|\log \alpha| = |\log \beta|$, and so

$$\frac{k_{3,v}}{X}|\log \alpha| = \frac{T}{X}\frac{k_{3,v}}{T}|\log \beta| \leq \frac{k_{3,v}}{T}\beta|\log \beta|.$$

For $\beta = 1/2$, this bound equals $((\log 2)/2)k_{3,v}/T$, which is actually greater than the bound in (7.89) for $\beta = 1/2$.

If $\beta < T/X$, then

$$\int_\alpha^1 \frac{k_{3,v}}{tX}dt = \int_{T/X}^1 \frac{k_{3,v}}{tX}dt = \frac{k_{3,v}}{X}\log r = \frac{k_{3,v}}{T}\frac{\log r}{r}.$$

The maximum of $(\log r)/r$ on $[1,3]$ is $1/e < 1/2$; thus, the upper bound in (7.90) is still valid. For $\beta = 1/2$, since $r < 1/\beta$ and $r \mapsto (\log r)/r$ is increasing on $[1,2]$, the bound $((\log 2)/2)k_{3,v}/T$ is also valid.

We conclude that (7.88) is always at most $((\log 2)/2)k_{3,v}/T$, if $\beta = 1/2$, and at most $k_{3,v}/2T$, if $\beta = 0$. □

7.6 MAIN-TERM TOTALS

We can now finish the task we started in Proposition 7.8. Let us first bound our expression in the intermediate range, namely, $U_0 \leq X \leq U_1$. For this purpose, we do not actually need the very last estimates we proved (Prop. 7.11–7.13).

Proposition 7.14. *Let $g_v(y)$ be as in (7.14), where $v \in \{1,2\}$ and ϱ is as in (7.29) for some $U_1 > U_0 > 0$. Then, for $U_0 \leq Y \leq U_1$,*

$$\left(\log \frac{U_1}{U_0}\right)^2 \sum_{\substack{s \leq Y/U_0 \\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{Y}{s}\right) \tag{7.91}$$

is at most

$$\frac{\pi^2}{6}\frac{\sigma(v)}{v} \log \frac{Y}{U_0} - \begin{cases} 0.998982\left(1 - \frac{1}{(Y/U_0)^{3/2}}\right) & \text{if } v = 1, \\ 1.817075\left(1 - \frac{1}{(Y/U_0)^{4/3}}\right) & \text{if } v = 2. \end{cases} \tag{7.92}$$

Proof. We just apply (7.47). We use (7.53) to estimate

$$\sum_{\substack{1 \leq s \leq Y/U_0 \\ (s,v)=1}} \frac{1}{s} h_v\left(\frac{Y/U_0}{s}\right) - \frac{\sigma(v)}{\phi(v)}\zeta(2) \sum_{\substack{1 \leq s \leq Y/U_0 \\ (s,v)=1}} \frac{1}{s}.$$

□

Corollary 7.15. *Let $g_v(y)$ be as in (7.14), where $v \in \{1,2\}$ and ϱ is as in (7.29) for some $U_1 > U_0 > 0$.*

Then, for $U_0 \leq X \leq U_1$ and $\beta = 1/2$,

$$\frac{6X}{\pi^2}\frac{v}{\sigma(v)} \int_\beta^1 \sum_{\substack{s \leq tX/U_0 \\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{tX}{s}\right) dt \tag{7.93}$$

is at most $(X/2)/(\log U_1/U_0)^2$ times

$$\log \frac{X}{U_0} - \left(1 - \frac{U_0}{X}\right) \cdot \begin{cases} 0.91415 & \text{if } v = 1, \\ 1 & \text{if } v = 2. \end{cases} \tag{7.94}$$

For $U_0 \leq X \leq U_1$ and $\beta = 0$, the expression in (7.93) is at most $X/(\log U_1/U_0)^2$ times

$$\log \frac{X}{U_0} - \left(1 - \frac{U_0}{X}\right) - c_v \delta_v(X/U_0), \tag{7.95}$$

where

$$\delta_v(t) = 1 - \frac{1}{t} + \frac{1}{\alpha_v - 1}\left(\frac{1}{t} - \frac{1}{t^{\alpha_v}}\right) \tag{7.96}$$

and $c_1 = 0.607308$, $c_2 = 0.736432$, $\alpha_1 = \frac{3}{2}$, $\alpha_2 = \frac{4}{3}$.

Note that $\delta_v(t) \geq 0$ for all $t \geq 1$, as will be clear from the proof.

Proof. Assume first that $\beta = 1/2$ and $X/U_0 \geq 2$. Then, by Prop. 7.14,

$$\left(\log \frac{U_1}{U_0}\right)^2 \frac{6}{\pi^2} \frac{v}{\sigma(v)} \int_{1/2}^{1} \sum_{\substack{s \leq tX/U_0 \\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{tX}{s}\right) dt \tag{7.97}$$
$$\leq \int_{1/2}^{1} \log \frac{tX}{U_0} dt - c_v \int_{1/2}^{1} \left(1 - \frac{1}{(tX/U_0)^{\alpha_v}}\right) dt.$$

Clearly

$$\int_{1/2}^{1} \log \frac{tX}{U_0} dt = \frac{1}{2} \log \frac{X}{U_0} + \frac{\log 2 - 1}{2}$$

and

$$\int_{1/2}^{1} \left(1 - \frac{1}{(tX/U_0)^{\alpha_v}}\right) dt = \frac{1}{2} - \frac{1}{(X/U_0)^{\alpha_v}} \frac{2^{\alpha_v - 1} - 1}{\alpha_v - 1}. \tag{7.98}$$

Hence, (7.97) is at most $1/2$ times

$$\log \frac{X}{U_0} - \begin{cases} 0.91416 - \frac{1.00623}{(X/U_0)^{\frac{3}{2}}} & \text{if } v = 1, \\ 1.04328 - \frac{1.14849}{(X/U_0)^{\frac{4}{3}}} & \text{if } v = 2. \end{cases} \tag{7.99}$$

Using $X/U_0 \geq 2$, we immediately derive the cleaner-looking bound

$$\frac{1}{2} \log \frac{X}{U_0} - \frac{1}{2}\left(1 - \frac{U_0}{X}\right) \cdot \begin{cases} 0.91416 & \text{if } v = 1, \\ 1.04328 & \text{if } v = 2, \end{cases} \tag{7.100}$$

since $1.00623/2^{\frac{3}{2}} < 0.91416$ and $1.14849/2^{\frac{4}{3}} < 1.04328$.

Consider now the case $\beta = 1/2$, $1 \leq X/U_0 < 2$. For $1 \leq y \leq v$, we obtain immediately from the definition (6.39) of $g_v(y_0, y_1)$ that

$$g_v(y, y) = (\log y)^2.$$

Hence, by the definition (7.30) of $g_v(y)$,

$$g_v(y) = \frac{(\log y)^2}{(\log U_1/U_0)^2}$$

for $U_0 \leq y \leq \min(U_1, vU_0)$, whereas $g_v(y) = 0$ for $y < U_0$. Therefore,

$$\begin{aligned}
&\left(\log \frac{U_1}{U_0}\right)^2 \int_{1/2}^{1} \sum_{\substack{s \leq tX/U_0 \\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{tX}{s}\right) dt \\
&\quad = \left(\log \frac{U_1}{U_0}\right)^2 \int_{U_0/X}^{1} \sum_{\substack{s \leq tX/U_0 \\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{tX}{s}\right) dt \\
&\quad = \int_{U_0/X}^{1} \left(\log \frac{tX}{U_0}\right)^2 dt = \frac{1}{X/U_0} \int_1^{X/U_0} (\log u)^2 du \\
&\quad = \left(\log \frac{X}{U_0}\right)^2 - 2\log \frac{X}{U_0} + 2 - \frac{2}{X/U_0}.
\end{aligned}$$

It is easy to check that

$$\frac{6}{\pi^2} (\log t)^2 \leq \left(2 \cdot \frac{6}{\pi^2} + \frac{1}{2}\right) \left(\log t - 1 + \frac{1}{t}\right)$$

for all $1 \leq t \leq 2$. (The inequality holds (non-strictly) at $t = 1$, the derivative of the left side at $t = 2$ is less than the derivative of the right side at $t = 2$, and the left side is convex on $[1, 2]$, whereas the right side is concave.) Hence,

$$\left(\log \frac{U_1}{U_0}\right)^2 \frac{6}{\pi^2} \frac{v}{\sigma(v)} \int_{1/2}^{1} \sum_{\substack{s \leq tX/U_0 \\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{tX}{s}\right) dt$$

is at most

$$\frac{1}{2} \log \frac{X}{U_0} - \frac{1}{2}\left(1 - \frac{U_0}{X}\right)$$

for $1 \leq X/U_0 \leq 2$. We conclude that (7.100) holds whether $X/U_0 \geq 1$ or $1 \leq X/U_0 \leq 2$.

Lastly, let $\beta = 0$, $X/U_0 \geq 1$ arbitrary. Then, by Prop. 7.14,

$$\begin{aligned}
&\left(\log \frac{U_1}{U_0}\right)^2 \frac{6}{\pi^2} \frac{v}{\sigma(v)} \int_{\beta}^{1} \sum_{\substack{s \leq tX/U_0 \\ (s,v)=1}} \frac{1}{s} g_v\left(\frac{tX}{s}\right) dt \\
&\qquad \leq \int_{U_0/X}^{1} \log \frac{tX}{U_0} dt - c_v \int_{U_0/X}^{1} \left(1 - \frac{1}{(tX/U_0)^{\alpha_v}}\right) dt.
\end{aligned} \tag{7.101}$$

Now

$$\int_{U_0/X}^{1} \log \frac{tX}{U_0} dt = \log \frac{X}{U_0} - \left(1 - \frac{U_0}{X}\right).$$

To finish, note that

$$\int_{U_0/X}^{1}\left(1-\frac{1}{(tX/U_0)^{\alpha_v}}\right)dt = 1-\frac{U_0}{X}-\frac{1}{\alpha_v-1}\left(\frac{U_0}{X}-\left(\frac{U_0}{X}\right)^{\alpha_v}\right), \tag{7.102}$$

and so the first line of (7.101) equals

$$\log\frac{X}{U_0}-(1+c_v)\left(1-\frac{U_0}{X}\right)+\frac{c_v}{\alpha_v-1}\left(\frac{U_0}{X}-\left(\frac{U_0}{X}\right)^{\alpha_v}\right).$$

We have obtained (7.95). The left side of (7.102) is non-negative because the integrand is non-negative. Hence $\delta_v(t)$, defined in (7.96), is non-negative for $t \geq 1$. □

The following bound is really made for $Y \geq U_1$, but we will soon find the fact that it is valid for all $Y > 0$ to be convenient.

Proposition 7.16. *Let $g_v(y)$ be as in (7.14), where $v \in \{1,2\}$ and ϱ is as in (7.29) for some $U_1 > U_0 > 0$. Then, for $Y \geq U_1$,*

$$\left(\log\frac{U_1}{U_0}\right)^2 \sum_{\substack{s\leq Y/U_0\\(s,v)=1}} \frac{1}{s} g_v\left(\frac{Y}{s}\right) \tag{7.103}$$

is at most

$$\begin{aligned}&\frac{\pi^2}{6}\frac{\sigma(v)}{v}\log\frac{U_1}{U_0}+2\frac{\pi^2}{6}\frac{\sigma(v)}{v}\left|\mathrm{err}_v\left(\frac{U_1}{U_0},\frac{Y}{U_0}\right)\right|\\&+\begin{cases}-0.998982\left(2-\left(\frac{U_0}{Y}\right)^{3/2}-\left(\frac{U_1}{Y}\right)^{3/2}\right)+0.003954+\frac{1.73788}{\sqrt{U_1/U_0}} & \text{if } v=1,\\ -1.817075\left(2-\left(\frac{U_0}{Y}\right)^{4/3}-\left(\frac{U_0}{Y}\right)^{4/3}\right)+0.004782+\frac{1.46592}{\sqrt{U_1/U_0}} & \text{if } v=2,\end{cases}\end{aligned} \tag{7.104}$$

where

$$\mathrm{err}_v(T,Z) = -\log Z + \sum_{\substack{\frac{Z}{T}<s\leq Z\\(s,v)=1}} \frac{1}{s}\check{m}_v\left(\frac{Z}{s}\right) + \frac{v}{\phi(v)}\sum_{\substack{s\leq Z/T\\(s,v)=1}}\frac{1}{s}.$$

The terms in (7.104) proportional to $1/\sqrt{U_1/U_0}$ can be omitted if $U_1/U_0 \geq 10^6$, and 1.73788 can be replaced by 1.49914 if $v = 1$ and $U_1/U_0 \geq 5$.

Proof. We start from (7.46). Lemma 7.10 gives us that

$$-\frac{\sigma(v)}{\phi(v)}\zeta(2)\sum_{\substack{Y/U_1<s\leq Y/U_0\\(s,v)=1}}\frac{1}{s}+\sum_{\substack{s\leq Y/U_0\\(s,v)=1}}\frac{1}{s}h_v\left(\frac{Y}{sU_0}\right)+\sum_{\substack{s\leq Y/U_1\\(s,v)=1}}\frac{1}{s}h_v\left(\frac{Y}{sU_1}\right) \tag{7.105}$$

is at most

$$2\frac{\sigma(v)}{\phi(v)}\zeta(2)\sum_{\substack{s\le Y/U_1\\(s,v)=1}}\frac{1}{s}-\frac{\sigma(v)}{v}\frac{\pi^2}{6}\left(\log\frac{Y}{U_0}+\log\frac{Y}{U_1}\right)$$
$$-\begin{cases}0.998982\left(2-\left(\frac{U_0}{Y}\right)^{3/2}-\left(\frac{U_1}{Y}\right)^{3/2}\right) & \text{if } v=1,\\ 1.817075\left(2-\left(\frac{U_0}{Y}\right)^{4/3}-\left(\frac{U_1}{Y}\right)^{4/3}\right) & \text{if } v=2.\end{cases} \tag{7.106}$$

By the definition of err_v,

$$2\frac{\sigma(v)}{\phi(v)}\zeta(2)\sum_{\substack{s\le Y/U_1\\(s,v)=1}}\frac{1}{s}+2\frac{\sigma(v)}{v}\zeta(2)\sum_{\substack{Y/U_1<s\le Y/U_0\\(s,v)=1}}\frac{1}{s}\check{m}_v\left(\frac{Y}{sU_0}\right)$$

equals

$$2\frac{\sigma(v)}{v}\zeta(2)\left(\log\frac{Y}{U_0}+\mathrm{err}_v\left(\frac{U_1}{U_0}\right)\right).$$

Obviously,

$$-\frac{\sigma(v)}{v}\frac{\pi^2}{6}\left(\log\frac{Y}{U_0}+\log\frac{Y}{U_1}\right)+2\frac{\sigma(v)}{v}\zeta(2)\log\frac{Y}{U_0}=\frac{\pi^2}{6}\frac{\sigma(v)}{v}\log\frac{U_1}{U_0}.$$

Lastly, by Lemma 7.6 and Lemma 7.7,

$$-2\sum_{\substack{s\le Y/U_1\\(s,v)=1}}\frac{1}{s}h_v\left(\frac{Y}{sU_0},\frac{Y}{sU_1}\right)\le\begin{cases}0.003954+\frac{1.73788}{\sqrt{U_1/U_0}} & \text{if } v=1,\\ 0.004782+\frac{1.46592}{\sqrt{U_1/U_0}} & \text{if } v=2,\end{cases}$$

where the second terms on the right can be omitted if $U_1/U_0\ge 10^6$, and 1.73788 can be replaced by 1.49914 if $v=1$ and $U_1/U_0\ge 5$.

□

Corollary 7.17. *Let $g_v(y)$ be as in (7.14), where $v\in\{1,2\}$ and ϱ is as in (7.29) for some $U_1>U_0>0$. Let $\beta=0$ or $\beta=1/2$. Then, for $X\ge U_1$,*

$$\frac{6X}{\pi^2}\frac{v}{\sigma(v)}\int_\beta^1\sum_{\substack{s\le tX/U_0\\(s,v)=1}}g_v\left(\frac{tX}{s}\right)dt \tag{7.107}$$

is at most

$$\frac{(1-\beta)X}{\log\frac{U_1}{U_0}}\left(1-\frac{1}{\log\frac{U_1}{U_0}}d_{v,\beta}\left(\frac{U_1}{U_0},\frac{X}{U_1}\right)\right), \tag{7.108}$$

where $d_{v,\beta}(t_1,t_2)$ is at least

$$\begin{cases}0.607309\cdot(\delta_{1,\beta}(t_2)+\delta_{1,\beta}(t_1t_2))-0.00286-\frac{1.13644}{\sqrt{t_1}}-\frac{2}{t_1} & \textit{if } v=1,\\ 0.736433\cdot(\delta_{2,\beta}(t_2)+\delta_{2,\beta}(t_1t_2))-0.00253-\frac{0.62882}{\sqrt{t_1}}-\frac{k_\beta}{t_1} & \textit{if } v=2,\end{cases} \tag{7.109}$$

and where $\delta_{v,0}(t) = \delta_v(t)$ *(defined as in (7.96)),*

$$\delta_{v,1/2}(t) = \max\left(1 - \frac{2^{\alpha_v} - 2}{\alpha_v - 1} t^{-\alpha_v}, 0\right),$$

$\alpha_v = (v+2)/(v+1)$, $k_{1/2} = 4.02026$ *and* $k_0 = 2.9$. *If* $U_1/U_0 \geq 5$, *then the constant* 1.13644 *may be replaced by* 0.9913.

It follows easily from the lower bound (7.109) that $d_v(t_1, t_2)$ is non-negative when $t_1 \geq b_{v,\beta}$, where

$$b_{1,1/2} = 8.72, \qquad b_{2,1/2} = 8.78, \qquad b_{1,1} = 7.54, \qquad b_{2,1} = 5.68. \tag{7.110}$$

Proof. Apply Propositions 7.14 and 7.16. The main term from either proposition contributes the main term

$$\frac{(1-\beta)X}{\log \frac{U_1}{U_0}}$$

in (7.108).

Then there are the following terms for $i = 1, 2$:

$$-\kappa_v \int_{\max(\beta, U_i/X)}^{1} \left(1 - \left(\frac{U_i}{X}\right)^{\alpha_v}\right) dt, \tag{7.111}$$

where $\kappa_1 = 0.998982$ and $\kappa_2 = 1.817075$. It is clear that the integral in (7.111) is always non-negative.

Much as in (7.98),

$$\int_{1/2}^{1} \frac{dt}{(tX/U)^{\alpha_v}} = \frac{1}{(X/U)^{\alpha_v}} \frac{2^{\alpha_v - 1} - 1}{\alpha_v - 1}$$

and so

$$\sum_{i=0}^{1} \int_{\max(1/2, X/U_i)}^{1} \left(1 - \left(\frac{U_i}{tX}\right)^{\alpha_v}\right) \geq \sum_{i=0}^{1} \max\left(\frac{1}{2} - \frac{2^{\alpha_v - 1} - 1}{\alpha_v - 1} \left(\frac{U_i}{X}\right)^{\alpha_v}, 0\right).$$

On the other hand,

$$\int_{U/X}^{1} \frac{dt}{(tX/U)^{\alpha_v}} = \frac{1}{(X/U_0)^{\alpha_v}} \frac{(X/U)^{\alpha_v - 1} - 1}{\alpha_v - 1} = \frac{1}{\alpha_v - 1} \left(\frac{U}{X} - \left(\frac{U}{X}\right)^{\alpha_v}\right)$$

and so

$$\sum_{i=0}^{1} \int_{U_i/X}^{1} \left(1 - \left(\frac{U_i}{tX}\right)^{\alpha_v}\right) dt = \delta_v\left(\frac{X}{U_0}\right) + \delta_v\left(\frac{X}{U_1}\right),$$

where $\delta_v(t)$ is as in (7.96). We do not forget to multiply by $(6/\pi^2)(v/\sigma(v))\kappa_v/(1-\beta)$, where the factor of $(1-\beta)$ is there to compensate to eliminate $(1-\beta)$ in (7.108).

The last terms in (7.104) are independent of t, and thus must simply be multiplied by $(6/\pi^2)v/\sigma(v)$. They give us

$$\begin{aligned} &0.002404 + \frac{1.056505}{\sqrt{U_1/U_0}} && \text{if } v = 1 \\ &0.001939 + \frac{0.594115}{\sqrt{U_1/U_0}} && \text{if } v = 2, \end{aligned}$$

where 1.056505 can be replaced by 0.911368 if $U_1/U_0 \geq 5$. (Of course we could also take advantage of the fact that these terms are not there when $t < U_1/X$, but we will not bother.)

It remains to estimate

$$\frac{2}{1-\beta}\left|\int_{\alpha}^{1} \mathrm{err}_v\left(\frac{U_1}{U_0}, \frac{tX}{U_0}\right) dt\right|, \tag{7.112}$$

where $\alpha = \max(\beta, U_1/X)$. If $v = 1$, then, by (7.83), the expression in (7.112) is at most

$$\frac{2}{U_1/U_0} + \frac{0.0799296}{\sqrt{U_1/U_0}} + 0.00045594.$$

If $v = 2$, we apply the bound (7.85) instead, and obtain that (7.112) is at most

$$\frac{0.034702}{\sqrt{U_1/U_0}} + 0.00058572 + \frac{2.9}{U_1/U_0} \cdot \begin{cases} 2\log 2 & \text{if } \beta = 1/2, \\ 1 & \text{if } \beta = 0. \end{cases}$$

□

7.7 REMAINDER TERMS

We should now bound the contribution of the remainder terms

$$R_{\alpha_1,\alpha_2,u,v,\varrho}(y) = \sum_{\substack{d_1,d_2 \leq y \\ \mu^2(d_1 l_1 d_2 l_2 v)=1}} \sum_{\substack{l_1 \leq y/d_1 \\ l_2 \leq y/d_2}} \frac{\varrho\left(\frac{yu}{d_1 l_1}\right)\varrho\left(\frac{yu}{d_2 l_2}\right)}{(d_1 d_2)^{\alpha_1}(l_1 l_2)^{\alpha_2}} \tag{7.113}$$

in Proposition 7.3 and Corollary 7.4.

Let ϱ be as in the statement of Theorem 7.1. Let $u = U_0$. Then

$$\sum_{d \leq y} \sum_{l \leq y/d} \frac{\varrho\left(\frac{yu}{dl}\right)}{d^{\alpha_1} l^{\alpha_2}} = \frac{D_{\alpha_1,\alpha_2}(y) - D_{\alpha_1,\alpha_2}\left(\frac{y}{U_1/U_0}\right)}{\log U_1/U_0},$$

where

$$D_{\alpha_1,\alpha_2}(y) = \sum_{d \leq y} \sum_{l \leq y/d} \frac{\log \frac{y}{dl}}{d^{\alpha_1} l^{\alpha_2}}. \tag{7.114}$$

Let us, then, estimate $D_{\alpha_1,\alpha_2}(y)$.

Proposition 7.18. *Let* $0 < \alpha_1, \alpha_2 \leq 1$, $\alpha_1 \neq \alpha_2$. *Let* $D_{\alpha_1,\alpha_2}(y)$ *be as in (7.114). Then, for all* $y > 0$ *and every* $0 \leq \beta < \min(\alpha_1, \alpha_2)$,

$$D_{\alpha_1,\alpha_2}(y) = \frac{y^{1-\alpha_1}}{(1-\alpha_1)^2}\zeta\left(1-\alpha_1+\alpha_2\right) + \frac{y^{1-\alpha_2}}{(1-\alpha_2)^2}\zeta\left(1-\alpha_2+\alpha_1\right) \\ +\zeta\left(\alpha_1\right)\zeta\left(\alpha_2\right)\log y + \zeta'\left(\alpha_1\right)\zeta\left(\alpha_2\right) + \zeta\left(\alpha_1\right)\zeta'\left(\alpha_2\right) + O^*\left(c_{\beta,\alpha_1,\alpha_2}y^{-\beta}\right) \tag{7.115}$$

if $\alpha_1, \alpha_2 \neq 1$, *and, if* $\alpha_1 = 1$, $\alpha_2 = \alpha < 1$, *or* $\alpha_2 = 1$, $\alpha_1 = \alpha < 1$,

$$\begin{aligned} D_{\alpha_1,\alpha_2}(y) = &\frac{y^{1-\alpha}}{(1-\alpha)^2}\zeta\left(2-\alpha\right) \\ &+ \frac{\zeta(\alpha)}{2}(\log y)^2 + \left(\zeta'(\alpha)+\gamma\zeta(\alpha)\right)\log y \\ &+ \frac{\zeta''(\alpha)}{2} + \gamma\zeta'(\alpha) - \gamma_1\zeta(\alpha) + O^*\left(c_{\beta,\alpha_1,\alpha_2}y^{-\beta}\right), \end{aligned} \tag{7.116}$$

where $c_{\beta,\alpha_1,\alpha_2}$ *is an explicitly computable constant depending only on* β, α_1 *and* α_2.

Here γ (Euler-Mascheroni constant) and γ_1 (first Stieltjes constant) are the first coefficients in the expansion $\zeta(1+s) = 1/s + \gamma - \gamma_1 s + \ldots$ of ζ around 1. We know that $\gamma = 0.57721\ldots > 0$; it will be useful to know as well that $\gamma_1 = -0.072815\ldots < 0$.

We will give a complex-analytic proof and sketch a real-analytic proof. The first proof follows a suggestion of Mathoverflow contributor Lucia.

Proof. By (2.34), for any $\sigma > 1$,

$$D_{\alpha_1,\alpha_2}(y) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty}\frac{Z(s)}{s^2}y^s ds,$$

where

$$\begin{aligned} Z(s) &= \sum_n \left(\sum_{d|n}\frac{1}{d^{\alpha_1}(n/d)^{\alpha_2}}\right)n^{-s} \\ &= \sum_d d^{-(s+\alpha_1)}\cdot\sum_l l^{-(s+\alpha_2)} = \zeta(s+\alpha_1)\zeta(s+\alpha_2). \end{aligned}$$

Shifting the line of integration to the left, we obtain, for $\alpha_1, \alpha_2 \neq 1$,

$$\begin{aligned} D_{\alpha_1,\alpha_2}(y) = &\frac{\zeta(1-\alpha_1+\alpha_2)}{(1-\alpha_1)^2}y^{1-\alpha_1}\cdot\mathrm{Res}_{s=1-\alpha_1}\zeta(s+\alpha_1) \\ &+ \frac{\zeta(1-\alpha_2+\alpha_1)}{(1-\alpha_2)^2}y^{1-\alpha_2}\cdot\mathrm{Res}_{s=1-\alpha_2}\zeta\left(s+\alpha_2\right) \\ &+ \mathrm{Res}_{s=0}\frac{\zeta(s+\alpha_1)\zeta(s+\alpha_2)}{s^2}y^s + \frac{1}{2\pi i}\int_{R_\beta}\frac{\zeta(s+\alpha_1)\zeta(s+\alpha_2)}{s^2}y^{-s}ds, \end{aligned}$$

where R_β is a path of our choice going from $-\beta - i\infty$ to $-\beta + i\infty$ such that $\Re s \leq -\beta$ for every point s on R_β and $-\beta < \min(1-\alpha_1, 1-\alpha_2, 1)$. The last condition on β is needed so that the path is to the left of all singularities.

If $\alpha_1 = 1$, $\alpha_2 = \alpha < 1$, or vice versa,

$$D_{\alpha_1,\alpha_2}(y) = \frac{\zeta(2-\alpha)}{(1-\alpha)^2} y^{1-\alpha} \cdot \mathrm{Res}_{s=1-\alpha}\, \zeta(s+\alpha) + \mathrm{Res}_{s=0} \frac{\zeta(s+1)\zeta(s+\alpha)}{s^2} y^s + \frac{1}{2\pi i}\int_{R_\beta} \frac{\zeta(s+\alpha_1)\zeta(s+\alpha_2)}{s^2} y^{-s} ds.$$

The residues here give us the main terms in (7.115) and (7.116). (Of course, $\mathrm{Res}_{s=1-\alpha}\, \zeta(s+\alpha) = \mathrm{Res}_{s=1}\, \zeta(s) = 1$.) We let $c_{\beta,\alpha_1,\alpha_2}$ be equal to

$$\frac{1}{2\pi}\int_{R_\beta} \frac{|\zeta(s+\alpha_1)||\zeta(s+\alpha_2)|}{|s|^2}|ds| \tag{7.117}$$

and obtain (7.115) and (7.116). It remains to see that this integral is finite and can be computed. For either of these two purposes, it is enough to show how to bound its tails.

For any $T > 0$, by Cauchy-Schwarz

$$\frac{1}{2\pi}\int_{-\beta-i\infty}^{-\beta-iT} + \int_{-\beta+iT}^{-\beta+i\infty} \frac{|\zeta(s+\alpha_1)||\zeta(s+\alpha_2)|}{|s|^2}|ds| \leq \sqrt{I(T,\alpha_1,\beta)I(T,\alpha_2,\beta)}, \tag{7.118}$$

where

$$\begin{aligned} I(T,\alpha,\beta) &= \frac{1}{2\pi}\left(\int_{-\beta-i\infty}^{-\beta-iT} + \int_{-\beta+iT}^{-\beta+i\infty} \frac{|\zeta(s+\alpha)|^2}{|s|^2}|ds|\right) \\ &\leq \kappa_{\alpha,\beta}\cdot\frac{1}{2\pi}\left(\int_{\alpha-\beta-i\infty}^{\alpha-\beta-iT} + \int_{\alpha-\beta+iT}^{\alpha-\beta+i\infty} \frac{|\zeta(s)|^2}{|s|^2}|ds|\right), \end{aligned} \tag{7.119}$$

where $\kappa_{\alpha,\beta} = 1$ if $|\beta| \geq |\alpha-\beta|$, and $\kappa_{\alpha,\beta} = (T^2 + (\alpha-\beta)^2)/(T^2+\beta^2)$ if $|\beta| < |\alpha-\beta|$. By the assumptions $\alpha_j \leq 1$ and $0 \leq \beta < \alpha_j$, we know that $0 < \alpha_j - \beta \leq 1$. Hence, Prop. 3.14 tells us that $I(T,\alpha_j,\beta)$ is finite for $j = 1, 2$ and shows us how to bound it explicitly.

□

Sketch of real-analytic proof. We can estimate sums of the form

$$\sum_{n\leq x} \frac{\left(\log\frac{x}{n}\right)^k}{n^\alpha}$$

by Euler-Maclaurin. We cannot simply apply the estimates we obtain to the inner sum in (7.114), however, since they are not very precise when $x = y/d$ is very small.

The obvious solution is to split the range and change the order of summation:

$$D_{\alpha_1,\alpha_2}(y) = \sum_{d\leq w}\sum_{l\leq\frac{y}{d}} \frac{\log\frac{y}{dl}}{d^{\alpha_1}l^{\alpha_2}} + \sum_{l\leq\frac{y}{w}}\sum_{w<d\leq\frac{y}{l}} \frac{\log\frac{y}{ld}}{l^{\alpha_2}d^{\alpha_1}}$$

for w arbitrary. Further analysis shows that this is suboptimal, and that what we really need is a split with smoothing:

$$D_{\alpha_1,\alpha_2}(y) = \sum_{d\le c_2 w}\sum_{l\le \frac{y}{d}} \frac{\log \frac{y}{dl}}{d^{\alpha_1} l^{\alpha_2}}\eta\left(\frac{d}{w}\right) + \sum_{l\le \frac{y}{c_1 w}}\sum_{d\le \frac{y}{l}} \frac{\log \frac{y}{ld}}{l^{\alpha_2} d^{\alpha_1}}\left(1-\eta\left(\frac{d}{w}\right)\right),$$

where $\eta : [0,\infty) \to \mathbb{R}$ is continuous and vanishes outside $[c_1, c_2\infty)$, say. The choice $\eta(t) = 1$ for $t \le 1$, $\eta(t) = 0$ for $t \ge 2$, $\eta(t) = \log(2/t)/\log 2$ for $1 \le t \le 2$ is good and convenient. (This function is of course of the same kind as ϱ; notice, however, that $\eta(t)$ goes from 1 to 0 much more quickly – within the bounded interval $[1, 2]$.) We set w optimally and, after much cancellation, obtain (7.115) for some $\beta > 0$ (though, with this choice of η, not for all $0 < \beta < \min(\alpha_1, \alpha_2)$). □

We still have to see how to take advantage of the coprimality conditions in the sum defining $R_{\alpha_1,\alpha_2,u,v,\varrho}$ in (7.113). Here are some possible approaches.

a) Derive an estimate for $R_{\alpha_1,\alpha_2,u,v,\varrho}$ from our estimate for $(D_{\alpha_1,\alpha_2}(y))^2$ by sieving for common factors of d_1, d_2, l_1 and l_2. Such an approach is in principle possible – it amounts to a more complicated version of the naïve square-free sieve at the beginning of §5.1 – but it is rather unwieldy.
b) Estimate

$$D^*_{\alpha_1,\alpha_2}(y) = \sum_{\substack{d\le y\\ \mu^2(dl)=1}}\sum_{l\le y/d} \frac{\log \frac{y}{dl}}{d^{\alpha_1} l^{\alpha_2}},$$

directly, much as we estimated $D_{\alpha_1,\alpha_2}(y)$, by means of the integral

$$\frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \frac{\zeta(s+\alpha_1)\zeta(s+\alpha_2)}{\zeta(2s+2\alpha_1)\zeta(2s+2\alpha_2)} L(s)\frac{y^s}{s^2}ds,$$

where the factor

$$L(s) = \prod_p \frac{(1-p^{-s-\alpha_1})(1-p^{-s-\alpha_2})}{1-p^{-s-\alpha_1}-p^{-s-\alpha_2}}$$

is there because of the coprimality condition $(d, l) = 1$ implied by $\mu^2(dl) = 1$.
The main difficulty here is that, when we shift the path of integration to the left, we cannot go beyond the vertical line $\Re s = 1/2 - \min(\alpha_1, \alpha_2)$, and the result is a an error bound that is not as good as it could be.
Obtaining a tight estimate for $R_{\alpha_1,\alpha_2,u,v,\varrho}$ from an estimate for $\left(D^*_{\alpha_1,\alpha_2}(y)\right)^2$ does not seem much simpler than the sieving procedure in a).
c) We can proceed as we are about to do, taking advantage of the coprimality conditions only for small primes p. The constant in front of the main term will not be optimal, but, given that what we are estimating is a remainder term, we can accept this situation.

Proposition 7.19. *For $y > 0$, $z \geq 1$ and $0 < \alpha_1, \alpha_2 \leq 1$, let*

$$D_{\alpha_1,\alpha_2,z}(y) = \sum_{\substack{d\leq y \\ p\leq z \Rightarrow p\nmid dl}} \sum_{l\leq y/d} \frac{\log \frac{y}{dl}}{d^{\alpha_1} l^{\alpha_2}}. \tag{7.120}$$

Then, for all $y > 0$,

$$\begin{aligned} D_{1,1/2,19}(y) &\leq 4\zeta(3/2) \prod_{p\leq 19} \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{p^{3/2}}\right) \cdot y^{1/2} \\ &= 0.744908\ldots \cdot y^{1/2}. \end{aligned} \tag{7.121}$$

The inequality would not be true for all y if 19 were replaced by a larger prime.

Proof. We proceed as in the proof of Prop. 7.18, using

$$Z_z(s) = \zeta(s+\alpha_1)\zeta(s+\alpha_2) \cdot \mathbf{P}_z(s+\alpha_1)\mathbf{P}_z(s+\alpha_2) \tag{7.122}$$

instead of $Z(s)$, where $\mathbf{P}_z(s) = \prod_{p\leq z}(1-p^{-s})$. We obtain

$$\begin{aligned} D_{1,\alpha,z}(y) = &\frac{y^{1-\alpha}}{(1-\alpha)^2}\zeta(2-\alpha)\mathbf{P}_z(1)\mathbf{P}_z(2-\alpha) \\ &+ \frac{\zeta(\alpha)}{2}\mathbf{P}_z(\alpha)\mathbf{P}_z(1)(\log y)^2 \\ &+ \left((\zeta\cdot\mathbf{P}_z)'(\alpha)\mathbf{P}_z(1) + \zeta(\alpha)\mathbf{P}_z(\alpha)(\gamma\mathbf{P}_z(1) + \mathbf{P}_z'(1))\right)\log y \\ &+ \left(\frac{(\zeta\cdot\mathbf{P}_z)''(\alpha)}{2} + \gamma\cdot(\zeta\cdot\mathbf{P}_z)'(\alpha) - \gamma_1\cdot(\zeta\cdot\mathbf{P}_z)(\alpha)\right)\mathbf{P}_z(1) \\ &+ \left((\zeta\cdot\mathbf{P}_z)'(\alpha) + \gamma\zeta(\alpha)\mathbf{P}_z(\alpha)\right)\mathbf{P}_z'(1) \\ &+ \zeta(\alpha)\mathbf{P}_z(\alpha)\frac{\mathbf{P}_z''(1)}{2} + O^*(c_{\beta,1,\alpha,z}y^{-\beta}), \end{aligned} \tag{7.123}$$

where

$$c_{\beta,\alpha_1,\alpha_2,z} = \frac{1}{2\pi}\int_{R_\beta} |\mathbf{P}_z(s+\alpha_1)\mathbf{P}_z(s+\alpha_2)| \frac{|\zeta(s+\alpha_1)||\zeta(s+\alpha_2)|}{|s|^2}|ds| \tag{7.124}$$

and R_β is any path of our choice going from $-\beta - i\infty$ to $-\beta + i\infty$ and staying on the half-plane $\Re s \leq -\beta$. An easy interval-arithmetic computation (via ARB) gives us that, for $\alpha = 1/2$ and $z = 19$, the expression on the right of (7.123) is at most

$$\begin{aligned} &4\zeta(3/2)\mathbf{P}_{19}(1)\mathbf{P}_{19}(3/2)\sqrt{y} \\ &- 0.001566(\log y)^2 - 0.048512(\log y) - 0.326737 + O^*(c_{\beta,1,\alpha,z}y^{-\beta}). \end{aligned} \tag{7.125}$$

Here

$$4\zeta(3/2)\mathbf{P}_{19}(1)\mathbf{P}_{19}(3/2) = 0.744908\ldots.$$

For $\beta = 1/4$ and any $T > 0$,

$$\frac{1}{2\pi}\int_{-\beta-i\infty}^{-\beta-iT} + \int_{-\beta+iT}^{-\beta+i\infty} |\mathbf{P}_z(s+\alpha_1)\mathbf{P}_z(s+\alpha_2)| \frac{|\zeta(s+\alpha_1)||\zeta(s+\alpha_2)|}{|s|^2}|ds| \leq B_z(-\beta+\alpha_1)B_z(-\beta+\alpha_2)\sqrt{I(T,\alpha_1,\beta)I(T,\alpha_2,\beta)}, \tag{7.126}$$

where $I(T,\alpha,\beta)$ is as in (7.119) and $B_z(\alpha) = \prod_{p\leq z}(1+p^{-\alpha})$. Now,

$$B_{19}(3/4)B_{19}(1/4) \leq 277.7841,$$

and, for $T = 40000$, by Prop. 3.14,

$$I(T,1/2,1/4) \leq 0.0052897, \qquad I(T,1,1/4) \leq 0.0000285109. \tag{7.127}$$

Hence, for $T = 40000$,

$$B_{19}(3/4)B_{19}(1/4)\sqrt{I(T,1/2,1/4)I(T,1,1/4)} \leq 0.10788. \tag{7.128}$$

Let R' consist of a straight segment from $-1/4 - iT$ to $-1/4 + iT$. ARB[1] shows rigorously that

$$\frac{1}{2\pi}\int_{R'} |\mathbf{P}_{19}(s+1/2)\mathbf{P}_{19}(s+1)| \frac{|\zeta(s+1/2)||\zeta(s+1)|}{|s|^2}|ds| = 1.136393 + O^*(5\cdot 10^{-7}). \tag{7.129}$$

We conclude that

$$c_{1/4,1,1/2,19} \leq 1.136394 + 0.10788 = 1.244274. \tag{7.130}$$

Since

$$0.001566(\log 60)^2 + 0.048512\log 60 + 0.326737 = 0.55161\ldots \tag{7.131}$$

and $1.244274 \cdot 60^{-1/4} = 0.44708\ldots < 0.55161$, it follows from (7.125) that (7.121) holds for $y \geq 60$. For $1 \leq y \leq 23$, $D_{1,1/2,19}(y) = \log y$, and (7.121) follows from $\max_{y\geq 1}(\log y)/y^{1/2} = 2/e$ and $2/e < 0.744908$. For $23 \leq y \leq 60$ (or indeed for $23 \leq y < 19^2$), we note that

$$D_{1,1/2,19}(y) = \log y + \sum_{23\leq p\leq y}\left(\frac{1}{\sqrt{p}} + \frac{1}{p}\right)\log\frac{y}{p},$$

and so, for y not a prime,

$$D'_{1,1/2,19}(y) = \frac{1}{y}\cdot\left(1 + \sum_{23\leq p\leq y}\left(\frac{1}{\sqrt{p}} + \frac{1}{p}\right)\right). \tag{7.132}$$

[1] When this section was written (late 2017 – early 2018), computing path integrals quickly in ARB required some special coding, particularly for $\Im s$ large. F. Johansson – the developer of the interval-arithmetic package ARB – kindly did most of the coding at the author's request. ARB has seen further development in that direction [Joh18].

It is easy to check that $1+\sum_{23\le p\le y}(1/\sqrt{p}+1/p)$ is less than $0.744908\sqrt{y}/2$ for every prime y in the interval $[23,60]$, and hence for every real number in the interval $[23,60]$. Hence $D'_{1,1/2,19}(y) \le 0.744908/2\sqrt{y}$ for every $23 \le y \le 60$ that is not a prime, and so $D_{1,1/2,19}(y) \le 0.744908\sqrt{y}$ for every $23 \le y < 60$, and so we are done. □

Lemma 7.20. *Let $D_{\alpha_1,\alpha_2,z}(y)$ be as in (7.120). Then, for all $y > 0$,*

$$\begin{aligned} D_{1/2,1/4,11}(y) &\le \frac{16}{9}\zeta(5/4)\prod_{p\le 11}\left(1-\frac{1}{p}\right)\left(1-p^{-5/4}\right)\cdot y^{3/4} \\ &= 0.551488\cdot y^{3/4}. \end{aligned} \tag{7.133}$$

Proof. The proof is essentially the same as that of Prop. 7.19. We start as in the proof of Prop. 7.18, with $Z_z(s)$ (defined in (7.122)) taking the place of $Z(s)$. We obtain

$$\begin{aligned} D_{1/2,1/4,11}(y) = &\frac{16}{9}\zeta(5/4)\mathbf{P}_{11}(1)\mathbf{P}_{11}(5/4)y^{3/4} + 4\zeta(3/4)\mathbf{P}_{11}(3/4)\mathbf{P}_{11}(1)y^{1/2} \\ &+\left(\frac{\zeta'(1/4)}{\zeta(1/4)}+\frac{\zeta'(1/2)}{\zeta(1/2)}+\frac{P'_{11}(1/4)}{P_{11}(1/4)}+\frac{P'_{11}(1/2)}{P_{11}(1/2)}+\log y\right) \\ &\cdot\zeta(1/4)\zeta(1/2)\mathbf{P}_{11}(1/4)\mathbf{P}_{11}(1/2) + O^*(c_{\beta,1/4,1/2,11}y^{-\beta}), \end{aligned} \tag{7.134}$$

where $c_{\beta,\alpha_1,\alpha_2,z}$ is as in (7.124). It is simple to compute that the right side of (7.134) is at most

$$\begin{aligned} &\frac{16}{9}\zeta(5/4)\mathbf{P}_{11}(1)\mathbf{P}_{11}(5/4)y^{3/4} - 0.29222\sqrt{y} \\ &+ 0.00007764\log y + 0.0021647 + O^*(c_{\beta,1/4,1/2,11}y^{-\beta}). \end{aligned} \tag{7.135}$$

By (7.126), for $\beta = 0$, the contribution to $c_{\beta,1/4,1/2,11}$ of values of s with $|\Im s| \ge T$ is at most

$$B_{11}(1/2)B_{11}(1/4)\sqrt{I(T,1/2,0)I(T,1/4,0)}, \tag{7.136}$$

where $B_z(\alpha) = \prod_{p\le z}(1+p^{-\alpha})$ and $I(T,\alpha,\beta)$ is as in (7.119). By Prop. 3.14,

$$I(T,1/2,0) \le 0.0002203834$$

for $T = 40000$. By the definition (7.119) of $I(T,\alpha,\beta)$,

$$I(T,1/4,0) \le \frac{T^2+(1/4)^2}{T^2} I(T,1/2,1/4).$$

Hence, for $T = 40000$, $I(T,1/4,0) \le 0.00529$ by (7.127), and so the expression in (7.136) is at most

$$6.98888\cdot 13.52339\cdot\sqrt{0.0002203834\cdot 0.00529} \le 0.10205.$$

Let R'_1 consist of straight segments from $-40000i$ to $-200i$ and from $200i$ to $40000i$; let R'_2 consist of straight segments from $-200i$ to -0.005 and from -0.005 to

$200i$. Then, by ARB,

$$\frac{1}{2\pi i}\int_{R_1'} |\mathbf{P}_{11}(s+1/2)\mathbf{P}_{11}(s+1/4)| \frac{|\zeta(s+1/2)||\zeta(s+1/4)|}{|s|^2}|ds| = 0.009269 + O^*(3\cdot 10^{-6}), \tag{7.137}$$

$$\frac{1}{2\pi i}\int_{R_2'} |\mathbf{P}_{11}(s+1/2)\mathbf{P}_{11}(s+1/4)| \frac{|\zeta(s+1/2)||\zeta(s+1/4)|}{|s|^2}|ds| = 0.685046\ldots. \tag{7.138}$$

Hence,

$$c_{0,1/2,1/4,11} \leq 0.009269 + 3\cdot 10^{-6} + 0.685047 + 0.10205 \leq 0.79637. \tag{7.139}$$

Since

$$0.29222\sqrt{13} - 0.00007764\log 13 - 0.0021647 = 1.05126\ldots > 0.79637, \tag{7.140}$$

it follows that (7.133) holds for $y \geq 13$. For $1 \leq y < 13$, $D_{1/2,1/4,11} = \log y$, and we know that $\log y \leq 0.551488 y^{3/4}$ because $(\log y)/y^{3/4} \leq (4/3)/e < 0.551488$. □

Corollary 7.21. *For any $y > 0$, $v \geq 1$, $0 < \alpha_1, \alpha_2 \leq 1$, let*

$$R^*_{\alpha_1,\alpha_2,v}(y) = \sum_{d_1\leq y}\sum_{l_1\leq y/d_1}\sum_{d_2\leq y}\sum_{\substack{l_2\leq y/d_2\\ \mu^2(d_1l_1d_2l_2v)=1}} \frac{\log\frac{y}{d_1l_1}}{d_1^{\alpha_1}l_1^{\alpha_2}}\frac{\log\frac{y}{d_2l_2}}{d_2^{\alpha_1}l_2^{\alpha_2}}. \tag{7.141}$$

Then, for all $y > 0$,

$$R^*_{1,1/2,v}(y) \leq \begin{cases} 12.9893\cdot y & \text{if } v = 1,\\ 4.7983\cdot y & \text{if } v = 2,\end{cases} \tag{7.142}$$

$$R^*_{1/2,1/4,v}(y) \leq \begin{cases} 5.8645\cdot y^{3/2} & \text{if } v = 1,\\ 2.0644\cdot y^{3/2} & \text{if } v = 2,\end{cases} \tag{7.143}$$

Proof. Let $D_{\alpha_1,\alpha_2,z}$ be as in (7.120). Let $P(z) = \prod_{p\leq z} p$. For any $v \in \{1,2\}$ and $z \geq v$, we can write

$$\sum_{d_1\leq y}\sum_{l_1\leq y/d_1}\sum_{d_2\leq y}\sum_{\substack{l_2\leq y/d_2\\ \mu^2((d_1l_1d_2l_2v,P(z)^2))=1}} \frac{\log\frac{y}{d_1l_1}}{d_1^{\alpha_1}l_1^{\alpha_2}}\frac{\log\frac{y}{d_2l_2}}{d_2^{\alpha_1}l_2^{\alpha_2}} \tag{7.144}$$

in the form

$$\sum_{\substack{d_1',d_2',l_1',l_2'\\ d_1'd_2'l_1'l_2'|P(z)\\ v\nmid d_1'd_2'l_1'l_2'}} \frac{D_{\alpha_1,\alpha_2,z}\left(\frac{y}{d_1'l_1'}\right) D_{\alpha_1,\alpha_2,z}\left(\frac{y}{d_2'l_2'}\right)}{(d_1'd_2')^{\alpha_1}(l_1'l_2')^{\alpha_2}}. \tag{7.145}$$

Since $R^*_{\alpha_1,\alpha_2,v}(y)$ is clearly bounded from above by the expression in (7.144), it must be bounded from above by the expression in (7.145).

Thus, if $D_{\alpha_1,\alpha_2,z}(y) \leq cy^\eta$ for all $y > 0$,

$$\begin{aligned} R^*_{\alpha_1,\alpha_2,v}(y) &\leq \sum_{\substack{d_1',d_2',l_1',l_2' \\ d_1'd_2'l_1'l_2'|P(z) \\ v\nmid d_1'd_2'l_1'l_2'}} \frac{c^2 y^{2\eta}}{(d_1'd_2')^{\alpha_1+\eta}(l_1'l_2')^{\alpha_2+\eta}} \\ &\leq c^2 y^{2\eta} \cdot \prod_{v<p\leq z} \left(1 + \frac{2}{p^{\alpha_1+\eta}} + \frac{2}{p^{\alpha_2+\eta}}\right). \end{aligned}$$

By Prop. 7.19, we conclude that

$$\begin{aligned} R^*_{1,1/2,v}(y) \leq y \cdot &\left(4\zeta(3/2) \prod_{p\leq 19}\left(1-\frac{1}{p}\right)\left(1-\frac{1}{p^{3/2}}\right)\right)^2 \\ &\cdot \prod_{v<p\leq 19}\left(1+\frac{2}{p}+\frac{2}{p^{3/2}}\right), \end{aligned}$$

and, by Lemma 7.20,

$$\begin{aligned} R^*_{1/2,1/4,v}(y) \leq y^{3/2} \cdot &\left(\frac{16}{9}\zeta(5/4) \prod_{p\leq 11}\left(1-\frac{1}{p}\right)\left(1-\frac{1}{p^{5/4}}\right)\right)^2 \\ &\cdot \prod_{v<p\leq 11}\left(1+\frac{2}{p}+\frac{2}{p^{5/4}}\right). \end{aligned}$$

□

Corollary 7.22. *Let $X > 0$, $v \in \{1,2\}$. Let ϱ be as in the statement of Theorem 7.1 for some $U_1 > U_0 > 0$. Let $R_{\alpha_1,\alpha_2,u,v,\varrho}$ be as in (7.113).*

Then

$$\sum_{\substack{s\leq X/U_0 \\ (s,v)=1}} R_{1,1/2,U_0,v,\varrho}\left(\frac{X}{sU_0}\right)\sqrt{\frac{X}{s}} = \frac{X/(U_0/\sqrt{X})}{\left(\log\frac{U_1}{U_0}\right)^2} \cdot \begin{cases} 33.933 & \text{if } v=1, \\ 8.1032 & \text{if } v=2. \end{cases} \tag{7.146}$$

$$\sum_{\substack{s\leq X/U_0 \\ (s,v)=1}} R_{1/2,1/4,U_0,v,\varrho}\left(\frac{X}{sU_0}\right)\cdot\left(\frac{X}{s}\right)^{1/4} = \frac{X/(U_0/\sqrt{X})^{3/2}}{\left(\log\frac{U_1}{U_0}\right)^2} \cdot \begin{cases} 11.5081 & \text{if } v=1, \\ 2.8467 & \text{if } v=2. \end{cases} \tag{7.147}$$

Proof. We bound $R_{\alpha_1,\alpha_2,U_0,v,\varrho}(y)$ brutally by $R^*_{\alpha_1,\alpha_2,v}(y)/(\log U_1/U_0)^2$, defined as in (7.141). Then, by Corollary 7.21,

$$\sum_{\substack{s\leq y\\(s,v)=1}} R_{1,1/2,U_0,v,\varrho}\left(\frac{y}{s}\right)\sqrt{\frac{y}{s}} \leq c_{1,v} \sum_{\substack{s\leq y\\(s,v)=1}} \left(\frac{y}{s}\right)^{3/2}$$
$$\leq y^{3/2}\cdot \begin{cases} c_{1,1}\zeta(3/2) & \text{if } v=1,\\ c_{1,2}\zeta(3/2)\left(1-2^{-3/2}\right) & \text{if } v=2,\end{cases} \tag{7.148}$$

where $c_{1,1} = 12.9893$ and $c_{1,2} = 4.7983$. Again by Corollary 7.21,

$$\sum_{\substack{s\leq y\\(s,v)=1}} R_{1/2,1/4,U_0,v,\varrho}\left(\frac{y}{s}\right)\left(\frac{y}{s}\right)^{1/4} \leq c_{2,v} \sum_{\substack{s\leq y\\(s,v)=1}} \left(\frac{y}{s}\right)^{7/4}$$
$$\leq y^{7/4}\cdot \begin{cases} c_{2,1}\zeta(7/4) & \text{if } v=1,\\ c_{2,2}\zeta(7/4)\left(1-2^{-7/4}\right) & \text{if } v=2,\end{cases} \tag{7.149}$$

where $c_{2,1} = 5.8645$ and $c_{2,2} = 2.0644$.

We let $y = X/U_0$. We need to multiply $\sqrt{y/s}$ (on the left side of (7.148)) by $\sqrt{U_0}$ to obtain $\sqrt{X/s}$ (on the left side of (7.146)); we also need to multiply $(y/s)^{1/4}$ (on the left side of (7.149)) by $U_0^{1/4}$ to obtain $(X/s)^{1/4}$ (on the left side of (7.147)). Obviously, $(X/U_0)^{3/2}\sqrt{U_0} = X/(U_0/\sqrt{X})$ and $(X/U_0)^{7/4}U_0^{1/4} = X/(U_0/\sqrt{X})^{3/2}$, and so we obtain (7.146) and (7.147). □

7.8 CONCLUSION

7.8.1 Proof of main result

Proving Theorem 7.1 is now just a matter of collecting statements. We will, in fact, prove the following stronger version.

Theorem 7.23. *Let $U_1 > U_0 > 0$. For $t > 0$, let*

$$\varrho(t) = \frac{\log^+(t/U_0) - \log^+(t/U_1)}{\log U_1/U_0} = \begin{cases} 0 & \text{if } t < U_0,\\ \frac{\log t/U_0}{\log U_1/U_0} & \text{if } U_0 \leq t \leq U_1,\\ 1 & \text{if } t > U_1.\end{cases} \tag{7.150}$$

a) *For* $U_0 \leq X \leq U_1$, $v \in \{1,2\}$ *and* $\beta = 1/2$ *or* $\beta = 1$,

$$\sum_{\substack{\beta X < m \leq X \\ (m,v)=1}} \left(\sum_{d|m} \mu(d)\varrho(d) \right)^2 \leq \left((1-\beta) L_{-,v,\beta}\left(\frac{X}{U_0}\right) + R_{v,\beta}\left(\frac{U_0}{\sqrt{X}}\right) \right) \frac{X}{\log^2 \frac{U_1}{U_0}}, \tag{7.151}$$

where

$$R_{v,\beta}(t) = \frac{\kappa_{v,1}}{t} + \frac{\kappa_{v,2}}{t^{3/2}}, \tag{7.152}$$

$$\kappa_{1,1} = 34.39 - 13.75\beta, \quad \kappa_{1,2} = 13.99 - 6.16\beta,$$
$$\kappa_{2,1} = 4.93 - 1.64\beta, \quad \kappa_{2,2} = 2.54 - 0.84\beta.$$

$$\begin{aligned} L_{-,v,1/2}(t) &\leq \log t - \left(1 - \frac{1}{t}\right) \cdot \begin{cases} 0.91415 & \text{if } v = 1, \\ 1 & \text{if } v = 2, \end{cases} \\ L_{-,v,0}(t) &\leq \log t - \left(1 - \frac{1}{t}\right) - c_v \delta_{v,0}(t), \end{aligned} \tag{7.153}$$

$$\delta_v(t) = 1 - \frac{1}{t} + \frac{1}{\alpha_v - 1}\left(\frac{1}{t} - \frac{1}{t^{\alpha_v}}\right) \tag{7.154}$$

and $c_1 = 0.607308$, $c_2 = 0.736432$, $\alpha_1 = \frac{3}{2}$, $\alpha_2 = \frac{4}{3}$.

b) *For* $X \geq U_1$ *and* $v \in \{1,2\}$,

$$\sum_{\substack{\beta X < m \leq X \\ (m,v)=1}} \left(\sum_{d|m} \mu(d)\varrho(d) \right)^2 \leq \left((1-\beta) L_{+,v,\beta} + R_{v,\beta}\left(\frac{U_0}{\sqrt{X}}\right) \right) \frac{X}{\log^2 \frac{U_1}{U_0}}, \tag{7.155}$$

where

$$L_{+,v,\beta} = \log \frac{U_1}{U_0} - d_v\left(\frac{U_1}{U_0}, \frac{X}{U_1}\right),$$

the remainder term $R_{v,\beta}$ *is as in* (7.152) *and* $d_v(t_1,t_2)$ *is at least*

$$\begin{aligned} &0.607309 \cdot (\delta_{1,\beta}(t_1) + \delta_{1,\beta}(t_1 t_2)) - 0.00286 - \frac{1.13644}{\sqrt{t_1}} - \frac{2}{t_1} && \text{if } v = 1, \\ &0.736433 \cdot (\delta_{2,\beta}(t_1) + \delta_{2,\beta}(t_1 t_2)) - 0.00253 - \frac{0.62882}{\sqrt{t_1}} - \frac{k_\beta}{t_1} && \text{if } v = 2, \end{aligned} \tag{7.156}$$

where $\delta_{v,0}(t)$ *is as above,* $k_{1/2} = 4.02026$, $k_0 = 2.9$ *and*

$$\delta_{v,1/2}(t) = \max\left(1 - \frac{2^{\alpha_v} - 2}{\alpha_v - 1} t^{-\alpha_v}, 0\right).$$

If $U_1/U_0 \geq 5$, *then the constant* 1.13644 *may be replaced by* 0.9913.

Note that $\delta_{v,0}(t)$ is non-negative for all $t \geq 1$, and $d_v(t_1, t_2)$ is non-negative when $t_1 \geq b_{v,\beta}$, where

$$b_{1,1/2} = 8.72, \qquad b_{2,1/2} = 8.78, \qquad b_{1,1} = 7.54, \qquad b_{2,1} = 5.68. \tag{7.157}$$

Proof. Immediate from Corollaries 7.4, 7.15, 7.17 and 7.22.

We compute our remainder-term bounds multiplying the constants defined in Prop. 7.3 by those defined in Cor. 7.22:

$$33.933 \cdot \left(\frac{5}{3} - \frac{2\beta}{3}\right)\frac{6}{\pi^2} \leq 34.39 - 13.75\beta,$$

$$8.1032 \cdot \left(\frac{3}{2} - \frac{\beta}{2}\right)\frac{4}{\pi^2} \leq 4.93 - 1.64\beta,$$

$$11.5081 \cdot \left(\sqrt{3}\left(1 - \frac{6}{\pi^2}\right) + (1-\beta)\left(\sqrt{8} - \frac{8^{3/2}}{\pi^2}\right)\right) \leq 13.99 - 6.16\beta,$$

$$2.8467 \cdot \left(1 - \frac{4}{\pi^2} + \frac{1-\beta}{2}\left(1 - \frac{4}{\pi^2}\right)\right) \leq 2.54 - 0.84\beta.$$

□

Theorem 7.1 follows immediately from Theorem 7.23, thanks to the remarks above on the non-negativity of $\delta_{v,0}(t)$ and $d_v(t_0, t_1)$.

The asymptotics of our bounds are clear:

1. as $t \to \infty$, $L_{-,v,1/2}(t)$ asymptotes to $\log t - 0.91415$ if $v = 1$ and to $\log t - 1$ if $v = 2$; the proof of Corollary 7.15 makes it clear how to obtain a bound with an asymptotic of $\log t - 1.04328$ if $v = 2$;
2. $L_{-,v,1/2}(t)$ asymptotes to $\log t - 1 - c_v$,
3. the bound on $L_{+,v,1/2}(t)$ asymptotes to $\log t - 2 \cdot 0.607309 + 0.00286 = \log t - 1.211758$ for $v = 1$, and to $\log t - 2 \cdot 0.736433 + 0.00253 = \log t - 1.470336$ for $v = 2$.

The proofs in §7.5 suffice to show that $\log t - 0.91415$, $\log t - 1.04328$ and $\log t - 1 - c_v$ are the true asymptotics for the optimal bounds on $L_{-,v,\beta}(t)$ for $U_0 \leq X \leq U_1$, up to an error of the order of 10^{-5}. For $X \geq U_1$, the terms 0.00286, 0.00253 are not there in reality; that is, the asymptotics should be $\log t - 1.21461\ldots$ for $v = 1$ and $\log t - 1.47286\ldots$ for $v = 2$, up to an error of the order of 10^{-5}. (In both cases, this small error comes from Lemma 7.10.)

Are these asymptotics optimal? That is, are there other smoothing functions ϱ satisfying $\varrho(t) = 0$ for $t < U_0$, $0 \leq \varrho(t) \leq 1$ for $U_0 \leq t \leq U_1$ and $\varrho(t) = 1$ for $t > U_1$ such that the asymptotics of the sum

$$\sum_{\substack{\beta X < m \leq X \\ (m,v)=1}} \left(\sum_{d|m} \mu(d)\varrho(d)\right)^2$$

are better than those given by the asymptotics above? Which smoothing functions would give optimal asymptotics? These questions remain open as of the time of writing.

7.8.2 Final remarks

There remain several matters to be looked into. First of all, there is what we already discussed in §7.2: the task of obtaining a similar bound when ϱ is such that it is supported on integers $\ll \sqrt{X}$, rather than on integers $\gg \sqrt{X}$; or, what would amount to the same, obtaining a bound when ϱ is as we defined it, but $U_1 \ll \sqrt{X}$ (whereas our current bound is good when $U_0 \gg \sqrt{X}$). It would also be interesting to obtain a bound valid for U_0, U_1 arbitrary, as in [Gra78], but with explicit constants and optimized error terms. The worst case is probably that where any of U_0^2, U_1^2 or U_0U_1 is close to X.

It goes almost without saying that the case $U_1 \ll \sqrt{X}$ would give a bound that could be stated in a traditional sieve framework. Instead of sieving all the integers in the interval $X/2 < n \leq X$, one could sieve any sequence that is well-distributed in arithmetic progressions of modulus up to U_1^2. It would admittedly be absurd to assume as much for $U_1 \gg \sqrt{X}$. What one could do with the results in the range we have studied is state them for an arbitrary smoothing. Of course, one can simply derive such a result from what we have – we did almost all of our work on an arbitrary interval $(\beta X, X]$, and chose to focus on $\beta = 1/2$ and $\beta = 0$ at the very end – but that might not be the most efficient option.

Finally, a remark on our tools. Because of our need for strong explicit estimates, we relied on results on sums of $\mu(n)$ based mainly on estimates of the form $\psi(x) = (1 + O^*(\epsilon))x$, with $\epsilon > 0$ fixed. Now, the estimates we used are based on finite verifications of the Riemann Hypothesis and on zero-free regions. However, estimates of the same kind, though with much larger values of ϵ, can be obtained through elementary means. (The oldest such bound is due to Chebyshev; see, e.g., [IK04, §2.2].) As a consequence, it is likely that qualitatively good (though numerically inferior) bounds can be obtained by purely elementary means, following the general procedure in this chapter.

Chapter Eight

The large sieve: smoothing and scattering

In a key part of our work (minor arcs, type II sums), we will need to bound a sum of the form

$$\sum_{M_0<m\leq M_1} \left| \sum_{p>V} (\log p) e(\alpha m p) \eta(p/W) \right|^2, \tag{8.1}$$

where $\eta : \mathbb{R} \to \mathbb{R}$ is a smoothing function.

From a modern perspective, the task of estimating a sum such as (8.1) is clearly a case for a large sieve. It is also clear that we ought to try to apply a large sieve for sequences of prime support.

For a large sieve to work, the frequencies that it is given – here: αm, for m in an interval – must be distinct, and not too close together. We may start by taking a Diophantine approximation $\alpha = a/q + \delta/x$.

If δ is zero and q is not too large, it is clear that all goes well: for m in a subinterval of $(M_0, M_1]$ of length q, the frequencies $\alpha m = am/q$ are distinct fractions a'/q, and thus any two of them are separated by at least $1/q$. In general, as we shall see in §8.2, we will manage to keep our angles αm apart from each other by using not just a/q, but δ/x. Our procedure will be compatible with taking advantage of prime support.

We will begin by reviewing what is meant by a large sieve and how to take a smoothing function such as η into account. We will then (§8.2) explain how to use the error term δ/x, besides reviewing Montgomery's method for taking advantage of prime support.

* * *

On saving a logarithmic factor. Methods for taking advantage of prime support in the large sieve to gain a factor of log are not something new; they are well-known to the specialists. Strangely enough, they seem to be rare in the literature on Goldbach's problem. Perhaps this oversight is due to the fact that proofs of Vinogradov's result given in textbooks often follow Linnik's dispersion method, rather than the large sieve. However, it should be emphasized that Ramaré did in fact use a large sieve for primes in [Ram95, §7].

The expression that Ramaré had to bound by a large sieve arose in a somewhat different way from the one here. Moreover, his treatment of the large sieve for primes is in the spirit of [BD69]. Our treatment of the large sieve will follow the basic lines set by Montgomery and Montgomery-Vaughan [MV73, (1.6)]. What we have to do, besides allowing for smoothing, is to show how that way to win a factor of log is compatible with the way we will exploit the error term δ/x.

8.1 THE LARGE SIEVE AND SMOOTHING

We recall that Plancherel's theorem states that the Fourier transform is an isometry: it preserves the L^2 norm. In particular, for $f : \mathbb{Z} \to \mathbb{C}$ in ℓ^1,

$$\int_{\mathbb{R}/\mathbb{Z}} |\widehat{f}(\alpha)|^2 d\alpha = \sum_{n\in\mathbb{Z}} |f(n)|^2 \qquad \text{(Parseval's theorem).}$$

Now, what if we wish to estimate, not the integral of $|\widehat{f}(\alpha)|^2$, but the sum of $|\widehat{f}(\alpha_i)|^2$ over a finite collection of points $\alpha_i \in \mathbb{R}/\mathbb{Z}$?

The remarkable thing is that it is possible to give a good bound for such a sum under rather loose conditions. Assume that f has support on $I \cap \mathbb{Z}$ for some interval $I \subset \mathbb{R}$ of length x and that the points $\{\alpha_i\}$ are separated by at least $\gamma > 0$, i.e., $\alpha_i - \alpha_j \notin (-\gamma, \gamma) \bmod \mathbb{Z}$ for all i, j with $i \neq j$. A *large sieve inequality* states that

$$\sum_i |\widehat{f}(\alpha_i)|^2 \leq c(x, \gamma) \sum_i |a_i|^2. \tag{8.2}$$

It is known that this is true with $c(x, \gamma) = x + 1/\gamma$ ([MV74], [Sel91, §20]), and that this coefficient $c(x, \gamma)$ is in general optimal.

The name "large sieve" has historical reasons – in its very first version, due to Linnik [Lin41], the large sieve served to estimate the size of a set out of which a large number of congruence classes had been excluded ("sifted out"). This is not the use we shall give to it. However, it is true and unsurprising that, thanks to our use of the large sieve, we will not need to use small sieves to win a factor of log in our estimates; indeed, we do not use small sieves anywhere at all.

There are many kinds of generalizations of the large sieve; see [IK04, Ch. VII]. Many consider transforms taken, not against additive characters $e(\alpha n)$, as in the Fourier transform $\widehat{f}(\alpha) = \sum_n f(n) e(\alpha n)$, but rather against multiplicative characters, or against much more general coefficients coming from automorphic forms. We will briefly touch upon a large sieve for multiplicative characters in §9.2. Now we have to discuss a different kind of generalization. We will still work with additive characters $e(\alpha n)$, but we also use a weight, or smoothing function, η: we want to bound

$$\sum_i \left| \sum_n a_n e(\alpha_i n) \eta(n/x) \right|^2,$$

for any $a_n \in \mathbb{C}$ and a given function $\eta : \mathbb{R} \to \mathbb{C}$ that is, if not smooth, at least continuous or piecewise continuous.

One might ask why one cannot simply set $f(n) = a_n \eta(n/x)$, apply (8.2) and be done with it. Of course one can do that, but that is often suboptimal, especially when the range within which a_n varies stays more or less constant as n grows. (This is typically the case in applications: a_n can be something like the characteristic function of a set of arithmetical significance, such as the primes.)

Two possible approaches suggest themselves. One is to follow the basic approach of Selberg [Sel91, §20]. This is what is done in [GV81] for the smoothing function

$$\eta(x) = \begin{cases} e^{-\lambda x} & \text{for } x \geq 0, \\ 0 & \text{for } x < 0, \end{cases}$$

and some variants thereof. The virtue of this approach is that, if conducted very carefully, it can give an optimal bound. The disadvantage is that it requires a great deal of ad-hoc work and ingenuity to optimize matters – and that for one specific smoothing function η. While [CV10a], [CV10b] do solve the optimization problem for a more general class of functions, it is a class that does not include the functions we will be interested in. There is a completely general result in [Vaa85, §4], but it is not optimal.

The alternative approach is what we do here: we take an intermediate result towards (8.2) (from the approach in [MV74], not that in [Sel91]) and then we derive from it a result for general η. The result is better than what one would get by naïvely setting $f(n) = a_n \eta(n/x)$, and indeed what is usually the main term seems to be optimal. (Other terms are not optimal; the total bound seems to match that in [Vaa85, §4].)

Using a result on sums without smoothing to derive a result on sums with smoothing may seem to defeat the purpose of smoothing. At the same time – what we have in the literature is carefully optimized bounds for the large sieve without smoothing. One advantage that results without smoothing do have is that they tend to imply results with smoothing, whereas the inverse is not in general the case.

What truly decides the issue is that we will have to work with versions of the large sieve that take advantage of situations where some points α_i are more isolated than others. Then there are delicate estimates in the literature for the case without smoothing, and nothing at all, apparently, for the case with smoothing. Hence, we now commit to deriving our results with smoothing from the best results known without smoothing.

* * *

Most proofs of large-sieve inequalities go through the duality principle. As we saw in §2.3.2, this principle asserts that the operator norm of a linear operator equals the operator norm of its dual (Lemma 2.3). This gives us the following.

Lemma 8.1. *Let $\{\alpha_i\}_{i\in S}$, $\{c_i\}_{i\in S}$ be finite collections of elements $\alpha_i \in \mathbb{R}/\mathbb{Z}$, $c_i \in [0,\infty)$. Let $\{\varphi_n\}_{n\in\mathbb{Z}}$, $\varphi_n \geq 0$, also be given. Assume that, for all $\{b_i\}_{i\in S}$, $b_i \in \mathbb{C}$, the inequality*

$$\sum_{n\in\mathbb{Z}} \left| \sum_{i\in S} b_i e(\alpha_i n) \right|^2 \varphi_n \leq \sum_{i\in S} c_i |b_i|^2 \tag{8.3}$$

holds. Then, for all $\{a_n\}_{n\in\mathbb{Z}}$, $a_n \in \mathbb{C}$, lying in ℓ^1 with respect to the measure φ_n,

$$\sum_{i\in S} c_i^{-1} \left|\sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\varphi_n\right|^2 \le \sum_{n\in\mathbb{Z}} |a_n|^2 \varphi_n. \tag{8.4}$$

Proof. Let $A : \mathbb{C}^S \to \mathbb{C}^{\mathbb{Z}}$ be the linear operator

$$\{b_i\}_{i\in S} \mapsto \left\{\sum_{i\in S} b_i e(-\alpha_i n)\right\}_{n\in\mathbb{Z}}.$$

Endow $\mathbb{C}^S$ with the weighted ℓ^2-norm $|\{b_i\}| = \sqrt{\sum_i |b_i|^2 c_i}$ and $\mathbb{C}^{\mathbb{Z}}$ with the weighted ℓ^2-norm $|\{a_n\}| = \sqrt{\sum_n |a_n|^2 \varphi_n}$. We write, as is usual, $\ell^2(\mathbb{Z})$ for the subspace of $\mathbb{C}^{\mathbb{Z}}$ consisting of elements of finite ℓ^2 norm; since S is finite, every element of $\mathbb{C}^S$ has finite ℓ^2 norm, and so $\ell^2(S) = \mathbb{C}^S$. By (8.3) applied to $\overline{b_i}$ instead of b_i,

$$\sum_{n\in\mathbb{Z}} \left|\sum_{i\in S} b_i e(-\alpha_i n)\right|^2 \varphi_n \le \sum_{i\in S} c_i |b_i|^2,$$

and so the operator A is bounded; in other words, A is actually an operator from $V = \ell^2(S)$ to $W = \ell^2(\mathbb{Z})$, where the ℓ^2 norms are weighted as above.

Define the inner product $\langle a_n, a_n'\rangle = \sum_n \overline{a_n} a_n' \varphi_n$ on $W \times W$ and the inner product $\langle b_i, b_i'\rangle = \sum_i c_i \overline{b_i} b_i'$ on $V \times V$. They induce the ℓ^2-norms we have just described. The dual $A^* : W \to V$ of A with respect to these inner products is the conjugate transpose

$$\{a_n\}_{n\in\mathbb{Z}} \mapsto \left\{c_i^{-1} \sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\phi_n\right\}_{i\in S}.$$

By (8.3), the operator A is not only bounded, but of norm $|A| \le 1$. By the duality principle (Lemma 2.3), $|A^*| = |A|$. Hence $|A^*| \le 1$. That means precisely that (8.4) holds. (The assumption $\{a_n\} \in \ell^1$ implies that A is well-defined at $\{a_n\}$.) □

The following estimates are at the core of the Montgomery-Vaughan approach.

Lemma 8.2. *Let $\gamma > 0$. Let $\{\alpha_i\}_{i\in S}$ be a finite collection of elements $\alpha_i \in \mathbb{R}/\mathbb{Z}$ such that $\alpha_i - \alpha_j \notin (-\gamma, \gamma) \bmod \mathbb{Z}$ for all $i, j \in S$ with $i \ne j$. Then, for all $\{b_i\}_{i\in S}$, $b_i \in \mathbb{C}$,*

$$\left|\sum_{i\in S}\sum_{\substack{j\in S\\ i\ne j}} \frac{b_i \overline{b_j}}{\sin \pi(\alpha_i - \alpha_j)}\right| \le \gamma^{-1} \sum_{i\in S} |b_i|^2. \tag{8.5}$$

Moreover, if $\{\gamma_i\}_{i\in S}$, $\gamma_i \in \mathbb{R}^+$, are such that $\alpha_i - \alpha_j \notin (-\gamma_i, \gamma_i)$ for all $i, j \in S$, $i \ne j$,

$$\left|\sum_{i\in S}\sum_{\substack{j\in S\\ i\ne j}} \frac{b_i \overline{b_j}}{\sin \pi(\alpha_i - \alpha_j)}\right| \le \frac{4}{3} \sum_{i\in S} \gamma_i^{-1} |b_i|^2. \tag{8.6}$$

Proof. Inequality (8.5) is part of [MV74, Thm. 1]. Inequality (8.6) is [Pre84, Thm]; it is a refinement of the other part of [MV74, Thm. 1].

The main tool behind (8.5) is Montgomery and Vaughan's following generalization of Hilbert's inequality:

$$\left| \sum_{i\in S} \sum_{\substack{j\in S \\ i\neq j}} \frac{b_i \overline{b_j}}{\lambda_i - \lambda_j} \right| \leq \frac{\pi}{\gamma} \sum_{i\in S} |b_i|^2$$

for any finite or infinite collection of elements $\lambda_i \in \mathbb{R}$ such that $|\lambda_i - \lambda_j| \geq \gamma$ for all i, j with $i \neq j$. (The classical form of Hilbert's inequality is the special case $\lambda_i = i$.) See either [MV74] or the exposition in [IK04, §7.4]. □

We will need a simple lemma expressing a function as a linear combination of characteristic functions of intervals; it is a variant of what is known as a "layer-cake decomposition", though the usual layer cake – consisting of "layers" $\{x \in \mathbb{R} : f(x) \geq t\}$ – would not do, as its "layers" are not in general intervals. Here, as usual, $1_{(a,b]}$ is the characteristic function of the interval $(a, b]$, i.e., $1_{(a,b]}(x)$ is 1 for $x \in (a, b]$ and 0 otherwise.

Lemma 8.3. *Let $f : \mathbb{R} \to [0, \infty)$ be integrable and of bounded variation. Then there are non-decreasing functions $a_0, a_1 : (0, t_0) \to \mathbb{R}$, where $t_0 = |f'|_1/2$ and $a_0(t) \leq a_1(t)$ for all $t \in (0, t_0)$, such that*

$$f(x) = \int_0^{t_0} 1_{(a_0(t),a_1(t)]}(x)dt$$

for all x at which f is left-continuous.

Here, as usual (see §2.3.3), if f is not in C^1, f' is understood in the sense of measures or distributions; in other words, $|f'|_1$ always stands for the total variation $\|df\|$.

If we were using intervals $[a_0(t), a_1(t))$, we would require f to be right-continuous at x instead. Needless to say, a function of bounded variation is both left- and right-continuous outside a countable set.

Proof. Since f is integrable and of bounded variation,

$$\lim_{t\to\infty} f(t) = \lim_{t\to-\infty} f(t) = 0.$$

Let

$$f_+(x) = \int_{-\infty}^{x} \max(f'(y), 0)dy, \qquad f_-(x) = \int_{-\infty}^{x} \max(-f'(y), 0)dy.$$

Then, by the fundamental theorem of calculus,

$$f(x) = \int_{-\infty}^{x} f'(y)dy = f_+(x) - f_-(x).$$

(Here, again, we are using $f'(y)dy$ as shorthand for df; this is just the integral of the Jordan decomposition $df = df^+ - df^-$, as in, e.g., [Rud74, §6.6].)

It is clear that $f_+(x)$ and $f_-(x)$ are both non-decreasing, and that

$$\lim_{x\to-\infty} f_+(x) = \lim_{x\to-\infty} f_-(x) = 0,$$
$$\lim_{x\to\infty} f_+(x) = \lim_{x\to\infty} f_-(x) = t_0.$$

For $t \in (0, t_0)$, let $a_0(t) = \inf\{x : f_+(x) > t\}$ and $a_1(t) = \inf\{x : f_-(x) > t\}$. It is clear that $a_0(t)$, $a_1(t)$ are well defined and $> -\infty$. Since $f(x) \geq 0$ for all x, we have $f_+(x) \geq f_-(x)$; hence, $a_0(t) \leq a_1(t)$ for all $t \in (0, t_0)$.

For any x,

$$\begin{aligned}\int_0^{t_0} 1_{(a_0(t),a_1(t)]}(x)dt &= \sup\{t : a_0(t) < x\} - \inf\{t : a_1(t) < x\} \\ &= \sup\{t : f_+(x) > t \wedge \exists x' < x \text{ s.t. } f_+(x') > t\} \\ &\quad - \inf\{t : f_-(x) > t \wedge \exists x' < x \text{ s.t. } f_-(x') > t\}.\end{aligned}$$

Now, if f is left-continuous at x, so are f_+ and f_-. Hence, $f_+(x) > t$ implies that $\exists x' < x$ such that $f_+(x') > t$, and dthe same is true for f_-. Hence, when f is left-continuous at x,

$$\begin{aligned}\int_0^{t_0} 1_{(a_0(t),a_1(t)]}(x)dt &= \sup\{t : f_+(x) > t\} - \inf\{t : f_-(x) > t\} \\ &= f_+(x) - f_-(x) = f(x).\end{aligned}$$

□

We can finally prove a large sieve with smoothing. As we already discussed, it is not that we actually take advantage of smoothing, as that we manage not to lose anything by it, and that we allow for a fully general smoothing function η.

Proposition 8.4 (A large sieve with smoothing)**.** *Let $\gamma > 0$. Let $\{\alpha_i\}_{i\in S}$ be a finite collection of elements $\alpha_i \in \mathbb{R}/\mathbb{Z}$ such that $\alpha_i - \alpha_j \notin (-\gamma, \gamma) \bmod \mathbb{Z}$ for all $i, j \in S$ with $i \neq j$. Let $\eta : \mathbb{R} \to [0, \infty)$ be integrable and of bounded variation. Let $x > 0$. Then, for all $\{a_n\}_{n\in\mathbb{Z}}$ such that $\sum_n |a_n|\eta(n/x) < \infty$,*

$$\sum_{i\in S} \left|\sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\eta(n/x)\right|^2 \leq \left(|\eta|_1 x + \frac{|\eta'|_1}{2\gamma}\right) \sum_{n\in\mathbb{Z}} |a_n|^2 \eta(n/x). \qquad (8.7)$$

Moreover, if $\{\gamma_i\}_{i\in S}$, $\gamma_i \in \mathbb{R}^+$, are such that $\alpha_i - \alpha_j \notin (-\gamma_i, \gamma_i)$ for all $i, j \in S$,

$$\sum_{i\in S} \left(|\eta|_1 x + \frac{2|\eta'|_1}{3\gamma_i}\right)^{-1} \left|\sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\eta(n/x)\right|^2 \leq \sum_{n\in\mathbb{Z}} |a_n|^2 \eta(n/x). \qquad (8.8)$$

Strictly speaking, we should also assume that $\sum_n |a_n|^2\eta(n/x) < \infty$, but, if that condition does not hold, then the conclusions (8.7) and (8.8) are at any rate empty as written; so, it does not matter.

Proof. Let us first prove (8.7) with $|\eta|_1 x+|\eta'|_1(1/2+1/2\gamma)$ instead of $|\eta|_1 x+|\eta'|_1/2\gamma$. By Lemma 8.1 with $\varphi_n = \eta(n/x)$, it is enough to prove that

$$\sum_{n\in\mathbb{Z}}\left|\sum_{i\in S} b_i e(\alpha_i n)\right|^2 \eta(n/x) \leq \left(|\eta|_1 x + \frac{|\eta'|_1}{2} + \frac{|\eta'|_1}{2\gamma}\right)\sum_{i\in S}|b_i|^2 \tag{8.9}$$

for all $\{b_i\}_{i\in S}$, $b_i \in \mathbb{C}$. (Notice that, since η is of bounded variation, it is also bounded.)

This is a known inequality for η the characteristic function of an interval (see the proof of [MV74, Thm. 1], or [IK04, §7.4]), so we could just use Lemma 8.3 (in the way we are about to display) to deduce it for general η. Let us actually derive the result we want from Lemma 8.2, since the process is instructive.

Expanding the square and reversing the order of summation, we see that the left side of (8.9) is at most

$$\sum_{i\in S}|b_i|^2\sum_n \eta(n/x) + \sum_{i\in S}\sum_{\substack{j\in S\\ i\neq j}} b_i\overline{b_j}\sum_n e((\alpha_i-\alpha_j)n)\eta(n/x). \tag{8.10}$$

By first-order Euler-Maclaurin (3.4),

$$\sum_n \eta(n/x) = x\int_{\mathbb{R}}\eta(x)dx + O^*(|\eta'|_1/2) \leq x|\eta|_1 + |\eta'|_1/2. \tag{8.11}$$

We can assume without loss of generality that η is continuous at every point n/x, $n\in\mathbb{Z}$, since we can dilate each such point to an interval of length ϵ and contracting all intervals $(n/x,(n+1)/x)$ slightly. (This does not change $|\eta'|_1$, and changes $|\eta|_1$ by at most $\epsilon|\eta'|_1$.) Hence, by Lemma 8.3 with $f(t)=\eta(t/x)$,

$$\sum_n e((\alpha_i-\alpha_j)n)\eta(n/x) = \sum_n e((\alpha_i-\alpha_j)n)\int_0^{t_0} 1_{(a_0(t),a_1(t)]}(n)\,dt,$$

where $t_0 = |f'|_1/2 = |\eta'|_1/2$. Since $|\eta|_1 < \infty$, we can restrict the sum to $-N \leq n \leq N$, and later let $N\to\infty$. Thus, and using as well the fact that a_0 and a_1 are non-decreasing, we justify the exchange of variables (i.e., Fubini's theorem):

$$\begin{aligned}\sum_n e((\alpha_i-\alpha_j)n)\int_0^{t_0} 1_{(a_0(t),a_1(t)]}(n)\,dt &= \int_0^{t_0}\sum_n e((\alpha_i-\alpha_j)n)1_{(a_0(t),a_1(t)]}(n)\,dt\\ &= \int_0^{t_0}\sum_{a_0(t)<n\leq a_1(t)} e((\alpha_i-\alpha_j)n)\,dt.\end{aligned}$$

Hence

$$\sum_{i\in S}\sum_{\substack{j\in S\\ i\neq j}} b_i\overline{b_j}\sum_n e((\alpha_i-\alpha_j)n)\eta(n/x)$$
$$= \int_0^{|\eta'|_1/2} \sum_{i\in S}\sum_{\substack{j\in S\\ i\neq j}} b_i\overline{b_j} \sum_{\lfloor a_0(t)\rfloor < n \leq \lfloor a_1(t)\rfloor} e((\alpha_i-\alpha_j)n)\; dt. \tag{8.12}$$

Now, for any integers m_0, m_1 and any $\alpha \in \mathbb{R}/\mathbb{Z}$, the sum $\sum_{m_0<n\leq m_1} e(\alpha n)$ equals

$$\frac{e((m_1+1/2)\alpha)-e((m_0+1/2)\alpha)}{e(\alpha/2)-e(-\alpha/2)} = \frac{\frac{i}{2}\cdot(e((m_0+1/2)\alpha)-e((m_1+1/2)\alpha))}{\sin\pi\alpha}.$$

Therefore, (8.12) equals

$$\sum_{k=0,1}\frac{(-1)^k i}{2}\int_0^{|\eta'|_1/2}\sum_{i\in S}\sum_{\substack{j\in S\\ i\neq j}}\frac{b_{i,k}(t)\overline{b_{j,k}(t)}}{\sin\pi(\alpha_i-\alpha_j)}dt, \tag{8.13}$$

where $b_{i,0}(t) = b_i e((\lfloor a_0(t)\rfloor + 1/2)\alpha_i)$ and $b_{i,1}(t) = b_i e((\lfloor a_1(t)\rfloor + 1/2)\alpha_i)$. We apply (8.5) in Lemma 8.2 for each k and t, and obtain that (8.13) is at most

$$\sum_{k=0,1}\frac{1}{2}\int_0^{|\eta'|_1/2}\gamma^{-1}\sum_{i\in S}|b_i|^2 dt = \frac{|\eta'|_1}{2\gamma}\sum_{i\in S}|b_i|^2 dt.$$

Putting this together with (8.10) and (8.11), we conclude that (8.9) holds. Therefore,

$$\sum_{i\in S}\left|\sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\eta(n/x)\right|^2 \leq \left(|\eta|_1 x + \frac{|\eta'|_1}{2} + \frac{|\eta'|_1}{2\gamma}\right)\sum_{n\in\mathbb{Z}}|a_n|^2\eta(n/x). \tag{8.14}$$

In order to remove the term $|\eta'|_1/2$, use the following trick, credited in [IK04, §7.4] to Paul Cohen. Let $R \geq 1$. Apply (8.14) to the sequence $a'_n = a_{n/R}$ if $R|n$, $a'_n = 0$ otherwise, with γ/R instead of γ, Rx instead of x, and the collection of points $\{(\alpha_i + r)/R\}_{i\in S, 0\leq r<R}$ instead of $\{\alpha_i\}_{i\in S}$. We obtain

$$\begin{aligned} R\sum_{i\in S}\left|\sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\eta(n/x)\right|^2 &= \sum_{r=0}^{R-1}\sum_{i\in S}\left|\sum_{n\in\mathbb{Z}} a'_n e\left(\frac{\alpha_i + r}{R}\cdot n\right)\eta(n/Rx)\right|^2 \\ &\leq \left(|\eta|_1\cdot Rx + \frac{|\eta'|_1}{2} + \frac{|\eta'|_1}{2(\gamma/R)}\right)\sum_{n\in\mathbb{Z}}|a'_n|^2\eta(n/Rx) \\ &= R\left(|\eta|_1 x + \frac{|\eta'|_1}{2R} + \frac{|\eta'|_1}{2\gamma}\right)\sum_{n\in\mathbb{Z}}|a_n|^2\eta(n/x). \end{aligned}$$

Dividing by R and letting $R \to \infty$, we obtain (8.7).

In order to prove (8.8), we proceed analogously. First, using (8.6) instead of (8.5), we follow the same steps as above, and obtain

$$\sum_{n\in\mathbb{Z}} \left|\sum_{i\in S} b_i e(\alpha_i n)\right|^2 \eta(n/x) \leq \sum_{i\in S} \left(|\eta|_1 x + \frac{|\eta'|_1}{2} + \frac{4}{3}\frac{|\eta'|_1}{2\gamma_i}\right) |b_i|^2 \tag{8.15}$$

for all $\{b_i\}_{i\in S}$, $b_i \in \mathbb{C}$. By Lemma 8.1 with $\varphi_n = \eta(n/x)$, this implies that

$$\sum_{i\in S} \left(|\eta|_1 x + \frac{|\eta'|_1}{2} + \frac{2|\eta'|_1}{3\gamma_i}\right)^{-1} \left|\sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\eta(n/x)\right|^2 \leq \sum_{n\in\mathbb{Z}} |a_n|^2 \eta(n/x).$$

We remove the term $|\eta'|_1/2$ by the same trick as before. □

As we said before, it has been known since [Vaa85] that the approach pioneered by Selberg can be used to prove a general result of the same shape as (8.7). It is possible to do qualitatively better when – as is the case here – η has no discontinuities. To be precise, via Selberg's majorant method, [Lit06] should imply that, for $k \geq 0$, if $\eta^{(k)}$ has bounded variation, then $|\eta'|_1/2\gamma$ in (8.7) can be replaced by $|\eta^{(k+1)}|/\gamma^{k+1}$ times a constant. (Thanks are due to E. Carneiro for this reference.) A similar result should be obtainable in the more traditional way followed here; in fact, the inequality that is then needed instead of Hilbert's inequality (in Lemma 8.2) is rather easy, provided that one does not care to obtain optimal constants. (All that we really need is the fact that $\sum_n 1/n^2$ converges.) Of course, we do care about good constants here, and Lemma 8.2 will do nicely.

8.2 SCATTERING: TAILS AND PRIMES

We will be bounding sums of the form

$$\sum_{M_0<m\leq M_1} \left|\sum_{n\in\mathbb{Z}} a_n e(\alpha m n)\eta(n/x)\right|^2, \tag{8.16}$$

of which the sum in (8.1) is a special case.

If α is exactly a/q, or very close to a/q, then it is clear how to proceed: split the sum over m into segments of length q; in any such segment $M < m \leq M + q$, the angles $\alpha_m = \alpha m$ will be separated by $1/q$ (or almost $1/q$); hence, we can apply the large sieve (Proposition 8.4).

The one issue here is that the bound that the large sieve gives must be multiplied by the number of segments to which the large sieve is applied, namely, $\lceil (M_1 - M_0)/q\rceil$, and that number may be large. There is no way out of this: if we took more than q consecutive values of m, the angles $\alpha m = am/q \bmod \mathbb{Z}$ would start repeating.

$m\alpha$, $m = 0, 1, \ldots, q-1$ Going beyond $m = q-1$

Figure 8.1: Separation between the angles $m\alpha$, where $\alpha = 1/q + \delta/x$

Suppose now that $\alpha = a/q + \delta/x$, where δ is small, but not too small. Then, while everything is a little less tidy, the angles do not repeat so readily: $\alpha(m+q)$ is not very close to αm – to be precise, $\alpha(m+q) - \alpha m = \delta q/x$. We can hence work with much longer segments without incurring in repetitions. The separation in the input to the large sieve is $\delta q/x$ rather than $1/q$; since, in our applications, $\delta q/x \leq 1/q$, the resulting bound from the large sieve will not be quite as good as before, but it will be multiplied by a much smaller number of segments. Thus, we will be able to give a much better total bound.

It can be helpful to look first at the case $\alpha = 1/q + \delta/x$ to apprehend quickly what is going on here. See Figure 8.1.

There is also the fact that we will be working with sequences a_n such that $a_n = 0$ when n is not a prime. There is a well-known technique for saving a factor of $\log x$ in this context. Lemmas of this kind were first given in [BD69] and [Mon68]; we will follow the latter approach.

Lemma 8.5 (Montgomery)**.** *Let $\{a_n\}_{n\in\mathbb{Z}}$ be in ℓ^1. Write $S(\alpha) = \sum_n e(\alpha n) a_n$. Let q be square-free. Assume that $a_n = 0$ for every n such that n and q are not coprime. Then*

$$|S(0)|^2 \leq \phi(q) \sum_{a \in (\mathbb{Z}/q\mathbb{Z})^*} \left| S\left(\frac{a}{q}\right) \right|^2. \tag{8.17}$$

The statement in the literature is more general than this: it is enough for there to be a forbidden congruence class r_p for each $p|q$, i.e., an r_p such that $a_n = 0$ if $n \equiv r_p \bmod p$; if there are several such congruence classes, the bound in (8.17) is improved accordingly.

Proof. Montgomery's original proof [Mon68] consists of the observation that, ignoring all terms but that corresponding the trivial character χ_T in (3.71), we obtain

$$\sum_{a \in (\mathbb{Z}/q\mathbb{Z})^*} \left| S\left(\frac{a}{q}\right) \right|^2 \geq \frac{1}{\phi(q)} \left| \sum_n a_n \chi_T(n) \right|^2 = \frac{|S(0)|^2}{\phi(q)}.$$

Alternatively, we may proceed as Huxley did ([Hux68], or [IK04, Lemma 7.15]): prove statement (8.17) for q prime using Plancherel and Cauchy-Schwarz; derive the statement for general q from this, by induction on the number of prime factors of q. □

We can use Lemma 8.5 in the following way, which is essentially standard.[1] Say that a_n in (8.16) is 0 not only whenever n is not prime, but also whenever $n \leq D$ for for a certain D. This implies that $a_n = 0$ for any n having at least one prime divisor $p \leq D$. Assume first, for simplicity, that $\alpha = a/q$. For every $q' \leq D$ and every $r \in \mathbb{Z}/q\mathbb{Z}$, we apply (8.17) with q' instead of q and $a_n e(rn/q)\eta(n/x)$ instead of a_n. We obtain

$$\frac{1}{\phi(q')}\left|\sum_{n\in\mathbb{Z}} a_n e(rn/q)\eta\left(\frac{n}{x}\right)\right|^2 \leq \sum_{a'\in(\mathbb{Z}/q'\mathbb{Z})^*}\left|\sum_{n\in\mathbb{Z}} a_n e\left(\frac{rn}{q}+\frac{a'n}{q'}\right)\eta\left(\frac{n}{x}\right)\right|^2$$

for every square-free $q' \leq D$. Summing over r and over q' square-free and coprime to q, we obtain

$$\begin{aligned}&\left(\sum_{\substack{q'\leq D\\ \mu^2(q')=1,(q',q)=1}} \frac{1}{\phi(q')}\right)\cdot\sum_{r=0}^{q-1}\left|\sum_{n\in\mathbb{Z}} a_n e(rn/q)\eta\left(\frac{n}{x}\right)\right|^2\\ &\leq \sum_{\substack{q'\leq D\\ \mu^2(q')=1,(q',q)=1}}\sum_{r=0}^{q-1}\sum_{a'\in(\mathbb{Z}/q'\mathbb{Z})^*}\left|\sum_{n\in\mathbb{Z}} a_n e\left(\left(\frac{r}{q}+\frac{a'}{q'}\right)n\right)\eta\left(\frac{n}{x}\right)\right|^2.\end{aligned} \tag{8.18}$$

Now, the angles $r/q + a'/q'$ (q fixed and q' varying over all $q' \leq D$ coprime to q) are separated by at least $1/qD^2$ from each other. Hence, Proposition 8.4 gives us that the right side of (8.18) is at most

$$\left(|\eta|_1 x + \frac{|\eta'|_1}{2/qD^2}\right)\sum_{n\in\mathbb{Z}}|a_n|^2\eta(n/x).$$

We thus conclude that

$$\sum_{r=0}^{q-1}\left|\sum_{n\in\mathbb{Z}} a_n e(rn/q)\eta\left(\frac{n}{x}\right)\right|^2 \leq \frac{|\eta|_1 x + |\eta'|_1 qD^2/2}{L_q(D)}\sum_{n\in\mathbb{Z}}|a_n|^2\eta(n/x), \tag{8.19}$$

where $L_q(D) = \sum_{q'\leq D:(q',q)=1}\mu^2(q')/\phi(q')$. As we saw in (6.5)–(6.6), it is easy to show that $L_q(D) > (\phi(q)/q)\log D$.

Choosing D so that $|\eta'|_1 qD^2/2$ is substantially smaller than $|\eta|_1 x$, we obtain an improvement by a factor of almost $L_q(D)$ over what we would have obtained from a direct application of Proposition 8.4.

[1]To be precise: what follows is usually done for $q = 1$, $a = 0$, as a way to estimate $\sum_n a_n$. This gives a result akin to a small sieve, as in [IK04, Thm. 7.14]. The author got the idea of using Lemma 8.5 to estimate sums of squares of $f(\alpha) = \sum_n a_n e(\alpha n)$ (as in (8.19)) from Heath-Brown, who showed him how to use Lemma 8.5 to bound an integral of $|f(\alpha)|^2$ over major arcs.

8.3 USING DIFFERENT KINDS OF SCATTERING

Let us start with two remarks. First – as is well-known, Montgomery's lemma (Lemma 8.5) can be applied very efficiently in combination with (8.8), rather than (8.7). (Inequality (8.8) can be superior to (8.7) when some angles α_i are separated from all others by more than the average.) This idea goes back to [MV73], where it was used to improve on the Brun-Titchmarsh theorem. We will soon see how to apply it to obtain an even larger improvement on a direct application of the large sieve than the one we already obtained in (8.19). What will happen is that, in effect, the term $|\eta'|_1 qD^2/2$ on the right side of (8.19) will become smaller, and thus we will be able to take a larger D, close to $\sqrt{x/q}$.

Second, it is possible, and in fact very straightforward, to combine Montgomery's lemma with the idea of using δ to let angles scatter, which we saw illustrated in Figure 8.1. The idea is that these are two kinds of scattering that can be allowed to happen at the same time: the scattering by δ works on a very small scale, whereas Montgomery's lemma – also a form of scattering – leaves large separations between the resulting angles. In other words, we will have clusters of many points (with points in each cluster separated by a small angle ν) around well-separated angles. Let us see exactly how this will work.

Lemma 8.6. *Let $q \geq 1$, $\nu, \upsilon > 0$, with $\nu + \upsilon \leq 1/q$. Let $\{\alpha_i\}_{i\in S}$ be a finite collection of elements $\alpha_i \in \mathbb{R}/\mathbb{Z}$ of the form $\alpha_i = a_i/q + \upsilon_i$, $0 \leq a_i < q$, where*

- *the elements $\upsilon_i \in \mathbb{R}$ all lie in an interval of length υ, and*
- *for any $i, j \in S$ with $i \neq j$, if $a_i = a_j$, then $|\upsilon_i - \upsilon_j| > \nu > 0$.*

Let $\eta : \mathbb{R} \to [0, \infty)$ be integrable and of bounded variation. Let $x > 0$. Then, for all $\{a_n\}_{n\in\mathbb{Z}}$ such that $\sum_n |a_n|\eta(n/x) < \infty$ and such that $a_n = 0$ whenever n has a prime factor $\leq 1/\sqrt{q(\nu + \upsilon)}$,

$$\sum_{i\in S}\left|\sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\eta(n/x)\right|^2 \leq \min\left(1, \frac{2q/\phi(q)}{\log\left((q(\nu+\upsilon))^{-1}\right)}\right) \cdot \left(|\eta|_1 x + \frac{|\eta'|_1}{2\nu}\right)\sum_{n\in\mathbb{Z}}|a_n|^2\eta(n/x). \tag{8.20}$$

Proof. For any distinct i, j, the angles α_i, α_j are separated by at least ν (if $a_i = a_j$) or at least $1/q - |\upsilon_i - \upsilon_j| \geq 1/q - \upsilon \geq \nu$ (if $a_i \neq a_j$). Hence we can apply Proposition 8.4 (the large sieve with smoothing) and obtain that the left side of (8.20) is at most

$$\left(|\eta|_1 x + \frac{|\eta'|_1}{2\nu}\right)\sum_{n\in\mathbb{Z}}|a_n|^2\eta(n/x).$$

We can also apply Lemma 8.5 with q' instead of q and $a_n e(\alpha_i n)\eta(n/x)$. Summing over all $i \in S$ and over all $q' \leq D$ square-free and coprime to q, much as in the exposition in §8.2, we get that, if $a_n = 0$ for all n with prime divisors $\leq D$,

$$\begin{aligned}&\left(\sum_{\substack{q'\leq D\\ \mu^2(q')=1,(q',q)=1}} \frac{1}{\phi(q')}\right)\cdot\sum_{i\in S}\left|\sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\eta\left(\frac{n}{x}\right)\right|^2\\ &\leq \sum_{\substack{q'\leq D\\ \mu^2(q')=1,(q',q)=1}}\ \sum_{a'\in(\mathbb{Z}/q'\mathbb{Z})^*}\sum_{i\in S}\left|\sum_{n\in\mathbb{Z}} a_n e\left(\left(\alpha_i+\frac{a'}{q'}\right)n\right)\eta\left(\frac{n}{x}\right)\right|^2.\end{aligned} \tag{8.21}$$

When we add all possible fractions of the form a'/q', $q'\leq D$, $(q',q)=1$, to the fractions a_i/q, we obtain fractions that are separated by at least $1/qD^2$. If $\nu+\upsilon\leq 1/qD^2$, then the resulting angles $\alpha_i + a'/r$ are still separated by at least ν. Set, then, $D = 1/\sqrt{(\nu+\upsilon)q}$, and apply Proposition 8.4 to bound the right side of (8.21).

We obtain

$$\sum_{i\in S}\left|\sum_{n\in\mathbb{Z}} a_n e(\alpha_i n)\eta\left(\frac{n}{x}\right)\right|^2 \leq \frac{|\eta|_1 x + \frac{|\eta'|_1}{2\nu}}{L_q(D)}\sum_{n\in\mathbb{Z}}|a_n|^2\eta(n/x),$$

where $L_q(D) = \sum_{q'\leq D:(q',q)=1}\mu^2(q')/\phi(q')$. By (6.6), $L_q(D) > (\phi(q)/q)\log D = (\phi(q)/2q)\log(((\nu+\upsilon)q)^{-1})$. □

Let us now see how to use (8.8). What we are about to do is what we will apply when $\alpha = a/q + \delta/x$, δ small, whereas what we just saw (Lemma 8.6) will be applied when δ is not so small. Here, just as in [MV73], we will exploit the fact that the Farey fractions (rationals with bounded denominator) are not equidistributed.

Lemma 8.7. *Let $\alpha = a/q + O^*(1/qQ)$, $q\leq Q$. Let $\eta:\mathbb{R}\to[0,\infty)$ be integrable and of bounded variation. Let $m_0\in\mathbb{Z}$. Let $0<x\leq Q/\kappa$, $\kappa>0$. Then, for all $\{a_n\}_{n\in\mathbb{Z}}$ such that (a) $\sum_n |a_n|\eta(n/x) < \infty$ and (b) $a_n = 0$ whenever n has a prime factor $< \sqrt{\kappa x/q}$,*

$$\sum_{m_0<m\leq m_0+q}\left|\sum_{n\in\mathbb{Z}} a_n e(\alpha m n)\eta(n/x)\right|^2 \leq \frac{2q}{\phi(q)}\frac{|\eta|_1 x}{\log\frac{x}{c\tau_{\eta,\kappa}q}}\sum_{n\in\mathbb{Z}}|a_n|^2\eta(n/x), \tag{8.22}$$

where $c = e^{-0.50136}$ and $\tau_{\eta,\kappa} = 1/\kappa + 2|\eta'|_1/3|\eta|_1$, provided that $x\geq 4\tau_{\eta,\kappa}q$.

Proof. We begin by using Montgomery's lemma (Lemma 8.5), much as before, with q' instead of q and $a_n e(\alpha m n)\eta(n/x)$ instead of a_n. Let $D < \sqrt{\kappa x/q}$. Then, summing over all $m_0 < m \leq m_0 + q$ all $q'\leq D$ square-free and coprime to q, we see that, for any choice of weights $w_{q'} \geq 0$,

$$\left(\sum_{\substack{q' \leq D \\ \mu^2(q')=1,(q',q)=1}} \frac{w_{q'}}{\phi(q')}\right) \cdot \sum_{m_0 < m \leq m_0+q} \left|\sum_{n \in \mathbb{Z}} a_n e(\alpha m n) \eta\left(\frac{n}{x}\right)\right|^2$$
$$\leq \sum_{\substack{q' \leq D \\ \mu^2(q')=1 \\ (q',q)=1}} \sum_{a' \in (\mathbb{Z}/q'\mathbb{Z})^*} \sum_{m_0 < m \leq m_0+q} w_{q'} \left|\sum_{n \in \mathbb{Z}} a_n e\left(\left(m\alpha + \frac{a'}{q'}\right) n\right) \eta\left(\frac{n}{x}\right)\right|^2. \tag{8.23}$$

What we will use now is the fact that an angle $a_1/q + a_1'/q_1'$ is separated from other angles $a_2/q + a_2'/q_2'$ ($q_1', q_2' \leq D$) by at least $1/qq_1'D$ rather than just by $1/qD^2$. This will give us a substantial gain when q_1' is considerably less than D. Of course, α is not exactly a/q; rather, $\alpha = a/q + \delta/x$, where $|\delta/x| \leq 1/qQ$. Still, for any two angles $m_1\alpha + a_1'/q_1'$, $m_2\alpha + a_2'/q_2'$ with $(q_i', q) = 1$, $a_i \in (\mathbb{Z}/q\mathbb{Z})^*$ and (m_1, a_1, q_1), (m_2, a_2, q_2) distinct,

$$\left|\left(m_1\alpha + \frac{a_1'}{q_1'}\right) - \left(m_2\alpha + \frac{a_2'}{q_2'}\right)\right| = \left|(m - m')\left(\frac{a}{q} + \frac{\delta}{x}\right) + \frac{a_1'}{q_1'} - \frac{a_2'}{q_2'}\right|$$
$$\geq \left|(m - m')\frac{a}{q} + \frac{a_1'}{q_1'} - \frac{a_2'}{q_2'}\right| - q \cdot \frac{\delta}{x} \geq \frac{1}{qq_1'q_2'} - \frac{1}{Q} \geq \frac{1}{qq_1'D} - \frac{1}{Q}.$$

Since $qD^2 \leq \kappa x < Q$, we know that $1/qq_1'D - 1/Q$ is always positive. Hence, we can apply Proposition 8.4, and obtain, from conclusion (8.8), that

$$\sum_{\substack{q' \leq D \\ \mu^2(q')=1,(q',q)=1}} \sum_{a' \in (\mathbb{Z}/q'\mathbb{Z})^*} \sum_{m_0 < m \leq m_0+q} w_{q'} \left|\sum_{n \in \mathbb{Z}} a_n e\left(\left(m\alpha + \frac{a'}{q'}\right) n\right) \eta\left(\frac{n}{x}\right)\right|^2$$
$$\leq \sum_{n \in \mathbb{Z}} |a_n|^2 \eta(n/x)$$

for

$$w_{q'} = \left(|\eta|_1 x + \frac{2|\eta'|_1/3}{\frac{1}{qq_1'D} - \frac{1}{Q}}\right)^{-1}.$$

Hence

$$\sum_{m_0 < m \leq m_0+q} \left|\sum_{n \in \mathbb{Z}} a_n e(\alpha m n) \eta\left(\frac{n}{x}\right)\right|^2 \leq \frac{1}{S} \sum_{n \in \mathbb{Z}} |a_n|^2 \eta(n/x)$$

where

$$S = \sum_{\substack{q' \leq D \\ \mu^2(q')=1,(q',q)=1}} \frac{w_{q'}}{\phi(q')} = \sum_{\substack{q' \leq D \\ (q',q)=1}} \left(|\eta|_1 x + \frac{2|\eta'|_1/3}{\frac{1}{qq'D} - \frac{1}{Q}}\right)^{-1} \frac{\mu^2(q')}{\phi(q')}.$$

What remains to do, then, is to bound S from below.

We will take $D = \sqrt{x/\tau q}$, where $\tau > 1/\kappa$ will be set later. Since $x \leq Q/\kappa$,

$$\frac{1}{Q} \leq \frac{1}{\kappa x} = \frac{1}{\kappa\tau}\frac{\tau}{x} = \frac{1}{\kappa\tau}\frac{1}{qD^2} \leq \frac{1}{\kappa\tau}\frac{1}{qq'D}.$$

Hence,

$$\begin{aligned}|\eta|_1 x + \frac{2|\eta'|_1/3}{\frac{1}{qq'D} - \frac{1}{Q}} &= |\eta|_1 x + \frac{2|\eta'|_1}{3}\frac{qq'D}{1-\frac{1}{\kappa\tau}} \\ &= |\eta|_1 x + \frac{2|\eta'|_1 q'}{3\left(1-\frac{1}{\kappa\tau}\right)D}\frac{x}{\tau} \\ &= \left(1 + \frac{2|\eta'|_1/3|\eta|_1}{\tau - 1/\kappa}\frac{q'}{D}\right)\cdot |\eta|_1 x,\end{aligned}$$

and so

$$\begin{aligned}S &\geq \frac{1}{|\eta|_1 x}\sum_{\substack{q' \leq D\\ (q',q)=1}} \left(1 + \frac{2|\eta'|_1/3|\eta|_1}{\tau - 1/\kappa}\frac{q'}{D}\right)^{-1}\frac{\mu^2(q')}{\phi(q')} \\ &= \frac{1}{|\eta|_1 x}\sum_{\substack{q' \leq D\\ (q',q)=1}} (1+q'D^{-1})^{-1}\frac{\mu^2(q')}{\phi(q')},\end{aligned}$$

where we have let $\tau = 1/\kappa + 2|\eta'|_1/3|\eta|_1$. Much as at the end of the proof of Lemma 8.6,

$$\sum_{\substack{q' \leq D\\ (q',q)=1}} \frac{1}{1+q'D^{-1}}\frac{\mu^2(q')}{\phi(q')} = \frac{\phi(q)}{q}\sum_{\substack{q' \leq D\\ (q',q)=1}} \frac{1}{1+q'D^{-1}}\frac{\mu^2(q')}{\phi(q')}\prod_{p|q}\left(1+\frac{1}{p-1}\right)$$

$$= \frac{\phi(q)}{q}\sum_{\substack{d\\ \frac{d}{(d,q)} \leq D}} \left(1 + \frac{d}{(d,q)}D^{-1}\right)^{-1}\frac{\mu^2(d)}{\phi(d)} \geq \frac{\phi(q)}{q}\sum_{d \leq D}\left(1+dD^{-1}\right)^{-1}\frac{\mu^2(d)}{\phi(d)}.$$

Now, for $D \geq 2$,

$$\sum_{d \leq D}(1+dD^{-1})^{-1}\frac{\mu^2(d)}{\phi(d)} > \log D + 0.25068. \tag{8.24}$$

This inequality is true for $D \geq 100$ by [MV73, Lemma 8]; it is easily verifiable numerically for $2 \leq D < 100$. (It suffices to verify the inequality for D integer with $d < D$ instead of $d \leq D$, as that is the worst case.) Our assumption $x \geq 4\tau_{\eta,k} q = 4\tau q$ ensures that $D \geq 2$.

Therefore,

$$S \geq \frac{1}{|\eta|_1 x}\frac{\phi(q)}{q}(\log D + 0.25068) = \frac{1}{|\eta|_1 x}\frac{\phi(q)}{q}\cdot\frac{1}{2}\log\frac{x}{c\tau q},$$

where $c = e^{-2\cdot 0.25068}$, $\tau = 1/\kappa + 2|\eta'|_1/3|\eta|_1$, and so we are done. □

Chapter Nine

The L^2 norm and the large sieve for primes

9.1 INTRODUCTION

Let $\{a_n\}_{n\geq 1}$, $a_n \in \mathbb{C}$, be given. We would like to bound

$$\int_{\mathfrak{M}} |S(\alpha)|^2 d\alpha, \tag{9.1}$$

where $S(\alpha)$ is the exponential sum $S(\alpha) = \sum_n a_n e(\alpha n)$ and $\mathfrak{M} = \mathfrak{M}(Q_0) \subset \mathbb{R}/\mathbb{Z}$ is a union of arcs around rationals a/q with denominator $q \leq Q_0$. In general, the best we can do is to use the trivial bound coming from Plancherel's identity:

$$\int_{\mathfrak{M}} |S(\alpha)|^2 d\alpha \leq \int_{\mathbb{R}/\mathbb{Z}} |S(\alpha)|^2 d\alpha = \sum_n |a_n|^2.$$

Assume from now on that a_n has prime support, that is, $a_n = 0$ for all non-prime n. Then we can hope to do better.

There are at least three possible approaches.

1. The integral (9.1) is similar to sums of the form

$$\sum_{q\leq Q_0} \sum_{\substack{a \\ (a,q)=1}} |S(a/q)|^2,$$

which can, of course, be bounded by a large sieve. As is well-known – and as we saw in §8.2 – large-sieve estimates can be improved by a factor when a_n has prime support. For a_n with support on the primes in $[1, N]$, the upper bound given by the large sieve is divided by

$$\sum_{\substack{q\leq Q/Q_0 \\ p|q \Rightarrow p>Q_0}} \frac{\mu^2(q)}{\phi(q)} \approx \sum_{\substack{q\leq Q/Q_0 \\ p|q \Rightarrow p>Q_0}} \frac{1}{q}, \tag{9.2}$$

where Q is somewhat smaller than $\sqrt{N}$. For $Q \to \infty$ and Q_0 such that $\log Q_0 = o(\log Q)$, it can be shown by exclusion-inclusion that the second sum in (9.2) asymptotes to

$$\prod_{p\leq Q_0} \left(1 - \frac{1}{p}\right) \cdot \sum_{q\leq Q/Q_0} \frac{1}{q} \sim e^{-\gamma} \frac{\log Q/Q_0}{\log Q_0} \approx \frac{\log N/Q_0^2}{2e^\gamma \log Q_0}, \tag{9.3}$$

where γ is Euler's constant. (Proving this asymptotic is a technical exercise. One natural approach is by the *fundamental lemma of sieve theory* (see, e.g., [IK04, Lemma 6.3]), which states, in essence, that small sieves give the right asymptotic when $\log Q_0 = o(\log Q)$, where Q is the length of an interval being sieved by primes $p \leq Q_0$.)

Can one improve on the trivial bound on (9.1) by the same factor, using analogous methods? Yes – this is an observation of Heath-Brown's, communicated to the author many years ago. That technique was then used by Tao in his work on sums of five primes (see [Tao14, Lem. 4.6] and adjoining comments), and also by the author, in the first version of [Helb].

The downside is that the factor in (9.2) is not actually optimal. In fact, as a comparison of (9.3) with a later asymptotic will make clear, having expression (9.3) in the denominator leads to bounds that are suboptimal by a factor of at least e^{γ}.

2. The second approach is like the first one, in that it also takes the proof of an existing result on the large sieve for a_n of prime support, and adapts it to bound the integral (9.1). This time, instead of using the method in §8.2, we follow the proof of a result of Ramaré's on the large sieve ([Ram09, Thm. 2.1]; see also [Ram09, Thm. 5.2]).

   The main difficulty lies not in adapting the method, but in completing it; in [Ram09], Ramaré carries out his analysis in full for Q_0 bounded, whereas we will need to work with Q_0 as large as a fractional power of N. Of course, we also need to give completely explicit results.

3. The work of Harper and Soundararajan [HS17] also gives estimates on integrals of the same type as (9.1). Their goal is to give non-trivial lower bounds on the integral of $|S(\alpha)|$ over the *complement* of $\mathfrak{M}$. They attain this aim by means of a multiplier $\sum_n \tilde{a}_n e(\alpha n)$, where $\tilde{a}_n$ is defined in terms of the natural quadratic sieve studied in §7.

   While giving lower bounds on the integral over the complement of $\mathfrak{M}$ amounts to the same as giving upper bounds on the integral over $\mathfrak{M}$, the task in [HS17] and our task are somewhat different. We want our estimate to be as good as possible, in fact asymptotically optimal, for Q_0 up to a certain level, but, on the other hand, we do not need bounds so fine as to give non-trivial lower bounds on the complement for very large Q_0.

   There are also purely practical difficulties: it seems hard work to carry out the approach in [HS17] explicitly.

We will take the second approach. It remains the case that an explicit version of the third approach would be interesting, and could supplement the results we will give.

9.2 THE LARGE SIEVE FOR PRIMES AND ITS ANALOGUE FOR AN INTEGRAL

We begin by proving Ramaré's inequality ([Ram09, Thm. 2.1]).

Proposition 9.1 (Ramaré). *Let $\{a_n\}_{n=1}^{\infty}$, $a_n \in \mathbb{C}$, be in ℓ^1. Let $1 \leq Q_0 \leq Q$. Assume that $a_n = 0$ whenever n has a prime factor $\leq Q$.*

Let $S(\alpha) = \sum_n a_n e(\alpha n)$ for $\alpha \in \mathbb{R}/\mathbb{Z}$. Then

$$\sum_{q \leq Q_0} \sum_{\substack{a \bmod q \\ (a,q)=1}} |S(a/q)|^2 \leq \left(\max_{q \leq Q_0} \frac{L_q(Q_0/q)}{L_q(Q/q)} \right) \sum_{q \leq Q} \sum_{\substack{a \bmod q \\ (a,q)=1}} |S(a/q)|^2, \tag{9.4}$$

where

$$L_q(x) = \sum_{\substack{n \leq x \\ (n,q)=1}} \frac{\mu^2(n)}{\phi(n)}. \tag{9.5}$$

Proof. Let us begin by expressing our sums in terms of primitive multiplicative characters. Write

$$s_\chi(q^*) = \sum_{\chi \bmod q^*}^{*} \left| \sum_n a_n \chi(n) \right|^2. \tag{9.6}$$

By Lemma 3.16, for any $1 \leq R \leq Q$,

$$\begin{aligned} \sum_{q \leq R} \sum_{\substack{a \bmod q \\ (a,q)=1}} |S(a/q)|^2 &= \sum_{q \leq R} \sum_{\substack{q^* | q \\ (q^*, q/q^*)=1 \\ \mu^2(q/q^*)=1}} \frac{q^*}{\phi(q)} \cdot s_\chi(q^*) \\ &= \sum_{q^* \leq R} L_{q^*}\left(\frac{R}{q^*}\right) \cdot \frac{q^*}{\phi(q^*)} s_\chi(q^*), \end{aligned} \tag{9.7}$$

since, for any q^*,

$$\sum_{\substack{q \leq R: q^* | q \\ (q^*, q/q^*)=1 \\ \mu^2(q/q^*)=1}} \frac{q^*}{\phi(q)} = \frac{q^*}{\phi(q^*)} \sum_{\substack{n \leq R/q^* \\ (n,q^*)=1}} \frac{\mu^2(n)}{\phi(n)} = \frac{q^*}{\phi(q^*)} L_{q^*}\left(\frac{R}{q^*}\right).$$

We now compare this last expression for $R = Q_0$ and for $R = Q$. Clearly

$$\begin{aligned} &\sum_{q^* \leq Q_0} L_{q^*}\left(\frac{Q_0}{q^*}\right) \cdot \frac{q^*}{\phi(q^*)} s_\chi(q^*) \\ &\qquad \leq \left(\max_{q \leq Q_0} \frac{L_q(Q_0/q)}{L_q(Q/q)} \right) \cdot \sum_{q^* \leq Q_0} L_{q^*}\left(\frac{Q}{q^*}\right) \frac{q^*}{\phi(q^*)} s_\chi(q^*) \\ &\qquad \leq \left(\max_{q \leq Q_0} \frac{L_q(Q_0/q)}{L_q(Q/q)} \right) \cdot \sum_{q^* \leq Q} L_{q^*}\left(\frac{Q}{q^*}\right) \frac{q^*}{\phi(q^*)} s_\chi(q^*). \end{aligned} \tag{9.8}$$

Applying (9.7) with $R = Q_0$ and $R = Q$ together with (9.8), we obtain (9.4). □

For a_n with support on an interval of length $\leq x$, we can, of course, bound the double sum on the right side of (9.4) by

$$(x+Q^2)\sum_n |a_n|^2 \tag{9.9}$$

by the large sieve, in the optimal form due to Montgomery-Vaughan and Selberg. (Take inequality (8.7), with η the characteristic function of an interval of unit length.)

We will not actually use (9.9), as we have other goals in mind. We would like to bound

$$\int_M |S(\alpha)|^2 d\alpha$$

for M a union of arcs around the fractions a/q in $\mathbb{R}/\mathbb{Z}$, where $q \leq Q_0$ and $(a,q)=1$. Consider first the case where $M = U_{Q_0,\beta}$, where $U_{R,\beta}$ is a union of arcs $(a/q - \beta, a/q+\beta)$ of equal length $\beta > 0$ around all a/q, $q \leq R$, $(a,q)=1$. Then, integrating inequality (9.4) (applied with $a_n e(\delta n)$, $\delta \in (-\beta,\beta)$, instead of a_n), we obtain

$$\int_M |S(\alpha)|^2 d\alpha \leq \left(\max_{q\leq Q_0} \frac{L_q(Q_0/q)}{L_q(Q/q)}\right)\int_{U_{Q,\beta}} |S(\alpha)|^2 d\alpha,$$

provided that the arcs in $U_{Q,\beta}$ do not overlap. Then we can simply bound

$$\int_{U_{Q,\beta}} |S(\alpha)|^2 d\alpha \leq \int_{\mathbb{R}/\mathbb{Z}} |S(\alpha)|^2 d\alpha = \sum_n |a_n|^2.$$

The arcs in $U_{Q,\beta}$ will not overlap provided that $\beta \leq 1/2Q^2$.

As it turns out, we can do better under the same assumptions: we can let M be a union of arcs broader than those in $U_{R,\beta}$ around rationals with small denominator.

Proposition 9.2. *Let $\{a_n\}_{n=1}^{\infty}$, $a_n \in \mathbb{C}$, be in ℓ^1. Let $1 \leq Q_0 \leq Q$. Assume that $a_n = 0$ whenever n has a prime factor $\leq Q$. Let $S(\alpha) = \sum_n a_n e(\alpha n)$ for $\alpha \in \mathbb{R}/\mathbb{Z}$.*

For $R \geq 1$, $\beta > 0$, let

$$M_{R,\beta} = \bigcup_{q\leq R} \bigcup_{\substack{a \bmod q\\(a,q)=1}} \left(\frac{a}{q} - \frac{R}{q}\beta, \frac{a}{q} + \frac{R}{q}\beta\right). \tag{9.10}$$

Then, for any $Q_0 \leq Q$ and $\beta \leq 1/2Q^2$,

$$\int_{M_{Q_0,\beta}} |S(\alpha)|^2\, d\alpha \leq \left(\max_{\substack{q\in\mathbb{Z}^+\\q\leq Q_0}} \max_{\substack{s\in\mathbb{R}\\1\leq s\leq \frac{Q_0}{q}}} \frac{L_q(Q_0/sq)}{L_q(Q/sq)}\right)\sum_n |a_n|^2,$$

where $L_q(x)$ is as in (9.5).

We are now stating that $q \in \mathbb{Z}^+$ for clarity, as s is a real variable. Before, q was also understood to be a positive integer variable, as is usual for us.

Proof. For R such that $\beta \leq 1/2R^2$, the arcs in $M_{R,\beta}$ do not overlap, and so

$$\begin{aligned}\int_{M_{R,\beta}} |S(\alpha)|^2 \, d\alpha &= \sum_{q\leq R} \int_{-\frac{R}{q}\beta}^{\frac{R}{q}\beta} \sum_{\substack{a \bmod q \\ (a,q)=1}} \left| S\left(\frac{a}{q}+\alpha\right)\right|^2 d\alpha \\ &= \int_{-\beta R}^{\beta R} \sum_{q\leq \min\left(R, \frac{\beta R}{|\alpha|}\right)} \sum_{\substack{a \bmod q \\ (a,q)=1}} \left| S\left(\frac{a}{q}+\alpha\right)\right|^2 d\alpha.\end{aligned} \tag{9.11}$$

Thus, by (9.7) applied with $a_n e(\alpha n)$ instead of a_n,

$$\int_{M_{R,\beta}} |S(\alpha)|^2 \, d\alpha = \int_{-\beta R}^{\beta R} \sum_{q\leq \min\left(R, \frac{\beta R}{|\alpha|}\right)} L_q \left(\frac{\min\left(R, \frac{\beta R}{|\alpha|}\right)}{q}\right) \frac{q}{\phi(q)} s_\chi(q,\alpha) d\alpha, \tag{9.12}$$

where

$$s_\chi(q,\alpha) = \sum_{\chi \bmod q}^{*} \left| \sum_n a_n e(\alpha n) \chi(n) \right|^2. \tag{9.13}$$

Now we compare the expression on the right of (9.12) for $R = Q_0$ and $R = Q$:

$$\begin{aligned}&\int_{-\beta Q_0}^{\beta Q_0} \sum_{q\leq \min\left(Q_0, \frac{\beta Q_0}{|\alpha|}\right)} L_q \left(\frac{\min\left(Q_0, \frac{\beta Q_0}{|\alpha|}\right)}{q}\right) \frac{q}{\phi(q)} s_\chi(q,\alpha) d\alpha \\ \leq K \cdot &\int_{-\beta Q_0}^{\beta Q_0} \sum_{q\leq \min\left(Q_0, \frac{\beta Q_0}{|\alpha|}\right)} L_q \left(\frac{\min\left(Q, \frac{\beta Q}{|\alpha|}\right)}{q}\right) \frac{q}{\phi(q)} s_\chi(q,\alpha) d\alpha \\ \leq K \cdot &\int_{-\beta Q}^{\beta Q} \sum_{q\leq \min\left(Q, \frac{\beta Q}{|\alpha|}\right)} L_q \left(\frac{\min\left(Q, \frac{\beta Q}{|\alpha|}\right)}{q}\right) \frac{q}{\phi(q)} s_\chi(q,\alpha) d\alpha,\end{aligned}$$

where

$$K = \max_{q\leq Q_0} \max_{\substack{s\in\mathbb{R} \\ 1\leq s\leq \frac{Q_0}{q}}} \frac{L_q(Q_0/sq)}{L_q(Q/sq)}.$$

Hence

$$\int_{M_{Q_0,\beta}} |S(\alpha)|^2 \, d\alpha \leq K \int_{M_{Q,\beta}} |S(\alpha)|^2 \, d\alpha \leq K \int_{\mathbb{R}/\mathbb{Z}} |S(\alpha)|^2 \, d\alpha = K \sum_n |a_n|^2.$$

□

9.3 BOUNDING THE QUOTIENT

As we saw in §9.2, following the same approach as in Ramaré's work on the large sieve leads us to estimates (Props. 9.1 and 9.2) involving the quotient

$$\max_{1\leq q\leq Q_0}\max_{1\leq s\leq Q_0/q}\frac{L_q(Q_0/sq)}{L_q(Q/sq)}, \tag{9.14}$$

where L_q is the quantity we defined in (9.5) and studied in §6.1.

We will see how to bound such a quotient in a way that is essentially optimal, not just asymptotically, but also in the ranges that are most relevant to us. Of course, besides being necessary for our work, our bounds will be immediately applicable to Ramaré's result on the large sieve itself, and thus ought to be useful elsewhere.

We start with a very easy algebraic manipulation.

Lemma 9.3. *Let L_q be as in (9.5). Let $Q, Q_0 \geq 1$. Let $1 \leq q \leq Q_0$, $1 \leq s \leq Q_0/q$, with $q \in \mathbb{Z}$ and $s \in \mathbb{R}$. Let c_0 be as in (6.3) and $\mathrm{err}_{q,R}$ as in (6.14).*

Then

$$\frac{L_q(Q_0/sq)}{L_q(Q/sq)} \leq \frac{\log Q_0 + c_0}{\log Q + c_0} \tag{9.15}$$

if and only if

$$\log sq \geq \sum_{p|q}\frac{\log p}{p} + \frac{q}{\phi(q)}\left((1+\eta)\,\mathrm{err}_{q,\frac{Q_0}{sq}} - \eta\,\mathrm{err}_{q,\frac{Q}{sq}}\right), \tag{9.16}$$

where $\eta = (\log Q_0 + c_0)/(\log Q/Q_0)$.

Proof. By the definition of $\mathrm{err}_{q,R}$, (9.15) holds if and only if

$$\begin{aligned}&\left(\log Q_0 - \log sq + c_0 + \sum_{p|q}\frac{\log p}{p} + \frac{q}{\phi(q)}\mathrm{err}_{q,\frac{Q_0}{sq}}\right)(\log Q + c_0)\\ &\leq \left(\log Q - \log sq + c_0 + \sum_{p|q}\frac{\log p}{p} + \frac{q}{\phi(q)}\mathrm{err}_{q,\frac{Q}{sq}}\right)(\log Q_0 + c_0),\end{aligned} \tag{9.17}$$

that is, if and only if

$$\left(\log sq - \sum_{p|q}\frac{\log p}{p}\right)\log\frac{Q}{Q_0} \geq \frac{q}{\phi(q)}\left((\log Q + c_0)\,\mathrm{err}_{q,\frac{Q_0}{sq}} - (\log Q_0 + c_0)\,\mathrm{err}_{q,\frac{Q}{sq}}\right),$$

which is equivalent to (9.16). □

The following would seem to be an inviting strategy. By the simple lower bound (6.6),

$$\mathrm{err}_{q,\frac{Q}{sq}} \geq -\frac{\phi(q)}{q}\left(c_0 + \sum_{p|q}\frac{\log p}{p}\right),$$

whereas, by the upper bound (6.7) with $C_{1,q}$ as in (6.8) and $C_1 = 1.47077$ (from (6.3) and (6.15)),

$$\mathrm{err}_{q,\frac{Q_0}{sq}} \leq C_1 + \sum_{p|q} \frac{\log p}{p} - \frac{\phi(q)}{q}\left(c_0 + \sum_{p|q}\frac{\log p}{p}\right).$$

It is easy to show that $\sum_{p|q}(\log p)/p \ll \log\log q$. It follows that $(q/\phi(q))\,\mathrm{err}_{q,Q/sq} \geq -c\log\log q$ and $(q/\phi(q))\,\mathrm{err}_{q,Q_0/sq} \leq c(\log\log q)^2$ for some constant c. We conclude that, for η bounded, (9.16) holds for all $q > c_\eta$, where $c_\eta > 0$ depends only on η. We could then hope to deal with $q \leq c_\eta$ by brute force.

The main issue with this approach is that, even for small values of η – say, η around 2 – the value of c_η is very large. We must thus refine this approach, applying (6.5) (together with the bounds in §6.1.2) instead of (6.8) to establish a better upper bound on $\mathrm{err}_{q,Q_0/sq}$, and also using (6.5) (together with a condition of the form $Q/Q_0 \geq C$ and the bounds in §6.1.2) to give a better lower bound on $\mathrm{err}_{q,Q/sq}$.

Lemma 9.4. *Let $q \in \mathbb{Z}^+$, q square-free. Let $R_0, R \geq 1$. Let c_0 be as in (6.3) and $\mathrm{err}_{q,R}$ as in (6.14).*

Then, for any $r|q$, and any $0 \leq \eta \leq \eta_0$,

$$\sum_{p|q}\frac{\log p}{p} + \frac{q}{\phi(q)}\left((1+\eta)\,\mathrm{err}_{q,R_0} - \eta\,\mathrm{err}_{q,R}\right) \tag{9.18}$$

is at most

$$(1+\eta_0)\frac{q}{\phi(q)}\left(\frac{\phi(r)}{r}\sum_{p|q}\frac{\log p}{p} + \kappa_{1,r}\right) - c_0 - \eta_0\kappa_{2,r},$$

where

$$\kappa_{1,r} = \frac{\phi(r)}{r}c_0 + \max(\mathrm{err}_{r,R_0}, 0), \qquad \kappa_{2,r} = c_0 + \sum_{p|r}\frac{\log p}{p} + \frac{r}{\phi(r)}\min\left(\mathrm{err}_{r,R}, 0\right).$$

Proof. By (6.14), for any R,

$$\mathrm{err}_{q,R} = L_q(R) - \frac{\phi(q)}{q}\left(\log R + c_0 + \sum_{p|q}\frac{\log p}{p}\right). \tag{9.19}$$

Write $q = rq'$. Now, by (6.5), (6.16) and (9.19),

$$\begin{aligned}\mathrm{err}_{q,R} &\geq \frac{\phi(q')}{q'}L_r(R) - \frac{\phi(q)}{q}\left(\log R + c_0 + \sum_{p|q}\frac{\log p}{p}\right)\\ &= \frac{\phi(q)}{q}\sum_{p|r}\frac{\log p}{p} + \frac{\phi(q')}{q'}\,\mathrm{err}_{r,R} - \frac{\phi(q)}{q}\sum_{p|q}\frac{\log p}{p}\\ &= \frac{\phi(q')}{q'}\,\mathrm{err}_{r,R} - \frac{\phi(q)}{q}\sum_{p|q'}\frac{\log p}{p}.\end{aligned}$$

On the other hand, by (6.9), (6.10) and (9.19),

$$\mathrm{err}_{q,R_0} \le \frac{\phi(r)}{r}\left(C_r(R_0) + \sum_{p|q'} \frac{\log p}{p}\right) - \frac{\phi(q)}{q}\left(c_0 + \sum_{p|q}\frac{\log p}{p}\right),$$

where

$$C_r(R_0) = \frac{r}{\phi(r)} L_r(R_0) - \log R_0 = c_0 + \sum_{p|r}\frac{\log p}{p} + \frac{r}{\phi(r)}\mathrm{err}_{r,R_0}.$$

Hence

$$\mathrm{err}_{q,R_0} \le \left(\frac{\phi(r)}{r} - \frac{\phi(q)}{q}\right)\left(c_0 + \sum_{p|q}\frac{\log p}{p}\right) + \mathrm{err}_{r,R_0}.$$

We thus see that, for $\eta \ge 0$,

$$\begin{aligned}(1+\eta)\,\mathrm{err}_{q,R_0} - \eta\,\mathrm{err}_{q,R} &= (1+\eta)\frac{\phi(r)}{r}\left(c_0 + \sum_{p|q}\frac{\log p}{p}\right) - \frac{\phi(q)}{q}\sum_{p|q'}\frac{\log p}{p}\\ &- (1+\eta)\frac{\phi(q)}{q}\left(c_0 + \sum_{p|r}\frac{\log p}{p}\right) + (1+\eta)\,\mathrm{err}_{r,R_0} - \frac{\phi(q')}{q'}\eta\,\mathrm{err}_{r,R}.\end{aligned}$$

Hence, the expression in (9.18) is at most

$$(1+\eta)\frac{q}{\phi(q)}\left(\frac{\phi(r)}{r}\sum_{p|q}\frac{\log p}{p} + \kappa_{1,r}\right) - c_0 - \eta\kappa_{2,r}$$

where

$$c_{1,r} = \frac{\phi(r)}{r}c_0 + \mathrm{err}_{r,R_0}, \quad c_{2,r} = c_0 + \sum_{p|r}\frac{\log p}{p} + \frac{r}{\phi(r)}\mathrm{err}_{r,R}.$$

The bound obviously still holds if we replace err_{r,R_0} by $\max(\mathrm{err}_{r,R_0}, 0)$ and $\mathrm{err}_{r,R}$ by $\min(\mathrm{err}_{r,R}, 0)$. After that, a quick examination shows that the coefficient of η must be positive, and thus, if we replace η by η_0, the bound is still valid. □

Our aim is to prove inequality (9.15). Before doing the general case, let us prove it in the special case of $q > 1$ not divisible by any prime $p \le 7$, so as to make the basic method clear.

Lemma 9.5. *Let L_q be as in (9.5). Let c_0 be as in (6.3). Let $Q_0 \ge 1$, $Q \ge 200Q_0$, with $\log Q/Q_0 \ge (3/10)(\log Q_0 + c_0)$. Let $1 < q \le Q_0$, $1 \le s \le Q_0/q$, with $q \in \mathbb{Z}$ and $s \in \mathbb{R}$. Assume that q is not divisible by any prime $p \le 7$.*

Then

$$\frac{L_q(Q_0/sq)}{L_q(Q/sq)} \le \frac{\log Q_0 + c_0}{\log Q + c_0}.$$

Proof. We can assume without loss of generality that q is square-free: if $p^k|q$, we can redefine q as q/p^{k-1} and s as sp^{k-1}.

By assumption, $\eta = (\log Q_0 + c_0)/(\log Q/Q_0) \geq 10/3$. Then, by Lemmas 9.3 and 9.4 with $r = 1$, inequality (9.15) will be true provided that

$$\log q \geq \frac{13}{3}\frac{q}{\phi(q)}\left(\sum_{p|q}\frac{\log p}{p} + c_{1,1}\right) - c_0 - \frac{10}{3}\kappa_{2,1}, \tag{9.20}$$

where

$$\kappa_{1,1} = c_0 + \max(\mathrm{err}_{1,Q_0/qs}, 0), \qquad \kappa_{2,1} = c_0 + \min\left(\mathrm{err}_{1,Q/qs}, 0\right).$$

By (6.15)–(6.16),

$$\mathrm{err}_{1,Q_0/qs} \leq 0.13818, \qquad \mathrm{err}_{1,Q/qs} \geq -0.02003.$$

Thus

$$\kappa_{1,1} \leq 1.47077, \qquad \kappa_{2,1} \geq 1.31255.$$

We must now show that (9.20) holds for $q > 1$ not divisible by any prime $p \leq 7$. For q having r prime factors, the right side of (9.20) is maximized when q is the product of the first r primes < 7, whereas the left side is then maximized. Thus, it is enough to consider q of the form $q = P(z)/P(7)$, $z \geq 11$, where $P(z) = \prod_{p\leq z} p$. Then $\log q = \sum_{p\leq z}\log p - \log 210$, $\sum_{p|q}(\log p)/p = \sum_{7<p\leq z}(\log p)/p$ and $q/\phi(q) = (\phi(210)/210)\prod_{p\leq z} p/(p-1)$.

For $7 < z \leq 41$ (say), we verify (9.20) directly. For larger z, we may use the bounds in [RS62, (3.24), (3.30)]:

$$\sum_{p\leq z}(\log p)/p < \log z, \qquad \prod_{p\leq z}\frac{p}{p-1} < e^{\gamma}\left(\log z + \frac{1}{\log z}\right), \tag{9.21}$$

valid for $z > 1$, and the bound, valid for $z \geq 17$,

$$\vartheta(z) = \sum_{p\leq z}\log p > 0.662865z, \tag{9.22}$$

which is true by (5.17) together with a simple check for $17 \leq z \leq 10^5$.

It is now clearly enough to verify that

$$0.662865z - \log 210 \tag{9.23}$$

is at least

$$\frac{13}{3}\frac{\phi(210)}{210}e^{\gamma}\left(\log z + \frac{1}{\log z}\right)\left(\log z - \sum_{p\leq 7}\frac{\log p}{p} + 1.47077\right) - c_0 - \frac{10}{3}\cdot 1.31255. \tag{9.24}$$

This inequality is easily seen to hold for $z \geq 41$: it holds for $z = 41$, and the derivative of the expression in (9.23) is greater than the derivative of the expression in (9.24) for $z = 41$, and the expression in (9.24) is a concave function (for $z \geq 41$, indeed for $z > e$), while the second derivative of the expression in (9.23) is of course 0. We conclude that inequality (9.15) holds. □

We will find it convenient to consider separately the case of Q_0/sq very small. In particular, if Q_0/sq is ≥ 1 but smaller than the smallest prime not dividing q, we know that $L_q(Q_0/sq) = 1$, and so we can use the following bound.

Lemma 9.6. *Let L_q be as in (9.5); let c_0 be as in (6.3). Let $Q_0 \geq 100$, $Q \geq 200Q_0$, with $\log Q/Q_0 \geq (\log Q_0 + c_0)/4$. Let $1 < q \leq Q_0$, $1 \leq s \leq Q_0$, with $q \in \mathbb{Z}$ and $s \in \mathbb{R}$.*

Then

$$\frac{1}{L_q(Q/sq)} \leq \frac{\log Q_0 + c_0}{\log Q + c_0}. \tag{9.25}$$

Much as in Lemma 9.5, we have not bothered to make the constant $1/4$ in the condition $\log Q/Q_0 \geq (\log Q_0 + c_0)/4$ quite as small as possible. (We will later use a stricter inequality in other cases.) It is clear that some condition of the form $\log Q/Q_0 \gg (\log Q_0 + c_0)$ is needed, as otherwise q could be so large that no integer $1 < m \leq Q/sq$ is coprime to q; in that case, we would have $L_q(Q/sq) = 1$, and (9.25) would be false.

Proof. The worst-case scenario here is clearly $Q_0 = \max(qs, 100)$; we will assume as much from now on. Write $R = Q/qs$. For q and R fixed, the left side of (9.25) is fixed. If $q \leq 100$, the right side is minimized when $s = 100/q$. Thus, it is then necessary and sufficient to show that

$$L_q(R) \geq \frac{\log 100 + \log R + c_0}{\log 100 + c_0}.$$

For $q \geq 100$, the right side of (9.25) is minimized for $s = 1$, and it is thus necessary and sufficient to show that

$$L_q(R) \geq \frac{\log q + \log R + c_0}{\log q + c_0}.$$

Hence, in either case, it is enough to show that

$$L_q(R) \geq 1 + \frac{\log R}{\log q' + c_0} \tag{9.26}$$

for $R \geq \max(200, (e^{c_0} q)^{1/4})$, where $q' = \min(q, 100)$.

Assume first that q is odd. It will be enough to use the lower bound

$$L_q(R) \geq \frac{\phi(q)}{q}(\log R + c_1)$$

with $c_1 = c_0 + \max_{R\geq 200} \mathrm{err}_{1,R} \geq 1.31255$, valid by (6.5) and (6.16). Thanks to this bound, (9.26) holds when

$$\left(\frac{\phi(q)}{q} - \frac{1}{\log q' + c_0}\right) \log R \geq 1 - \frac{\phi(q)}{q} c_1.$$

It is easy to check that this inequality holds for $q \leq 1000$ odd and $R = 200$ (and thus for any higher R). We must also check it for $q > 1000$ and $R = \max\left(200, (e^{c_0} q)^{1/4}\right)$. It is clearly enough if

$$\left(\frac{\phi(q)}{q} - \frac{1}{\log q + c_0}\right) \cdot \frac{1}{4}(\log q + c_0) \geq 1 - \frac{\phi(q)}{q} c_1,$$

i.e.,

$$\frac{\phi(q)}{q}\left(\frac{1}{4}(\log q + c_0) + c_1\right) \geq 1 + \frac{1}{4}.$$

By [RS62, Thm. 15], $2q/\phi(2q) > \eth(2q)$, where $\eth(r) = e^{\gamma} \log\log r + 2.50637/\log\log r$. Of course, since q is odd, $q/\phi(q) = (2q/\phi(q))/2$. Hence, it suffices to check that

$$\frac{z + c_0}{4} + c_1 \geq \frac{5}{8}\left(e^{\gamma} + \frac{2.50637}{(\log\log 2000)^2}\right) \log(z + \log 2)$$

for $z \geq \log 1000$. We do so by comparing the values and first derivatives of both sides for $z = \log 1000$, and noting that the right side is concave.

Let us now consider q even. In general, for r fixed and q square-free and divisible by r, we deduce from (6.5) and (6.14) that

$$L_q(R) \geq \frac{\phi(q/r)}{q/r} L_r(r) \geq \frac{\phi(q)}{q}(\log R + c_r)$$

with $c_r = c_0 + \sum_{p|r} (\log p)/p + (r/\phi(r)) \max_{R\geq 200} \mathrm{err}_{r,R}$. Just as before, we see it is enough to show that

$$\left(\frac{\phi(q)}{q} - \frac{1}{\log q' + c_0}\right) \log R \geq 1 - \frac{\phi(q)}{q} c_r.$$

For each even $r|2310$, we verify this bound for all $q \leq 30000r$ such that $\gcd(q, 2310) = r$. We also verify that, for every even $r|2310$ and $r' = 2310/r$,

$$\frac{z + c_0}{4} + c_r \geq \frac{5}{4}\frac{\phi(r')}{r'}\left(e^{\gamma} + \frac{2.50637}{(\log\log 30000rr')^2}\right) \log\left(z + \log r'\right)$$

for $z \geq \log 30000r$, in that we compare values and derivatives at $z = \log 30000r$. □

We can now consider the general case.

Proposition 9.7. *Let L_q be as in (9.5). Let c_0 be as in (6.3). Let $Q_0 \geq 360$, $Q \geq 200Q_0$, with $\log Q/Q_0 \geq (2/5)(\log Q_0 + c_0)$. Let $1 < q \leq Q_0$, $1 \leq s \leq Q_0/q$, with $q \in \mathbb{Z}$ and $s \in \mathbb{R}$.*

Then

$$\frac{L_q(Q_0/sq)}{L_q(Q/sq)} \leq \frac{\log Q_0 + c_0}{\log Q + c_0}. \tag{9.27}$$

As we will see, the conditions $Q_0 \geq 360$ and $Q \geq 200Q_0$ are essentially tight, though one can be loosened if the other one is strengthened. We could also loosen the condition $\log Q/Q_0 \geq (2/5)(\log Q_0 + c_0)$ somewhat, at the cost of more ad hoc work.

Proof. As always, we can assume that q is square-free. Let $r|2310$. By Lemmas 9.3 and 9.4, inequality (9.15) will be true for q with $r|\gcd(q, 2310)$ provided that-

$$\log q \geq \frac{7}{2} \frac{q}{\phi(q)} \left(\frac{\phi(r)}{r} \sum_{p|q} \frac{\log p}{p} + \kappa_{1,r} \right) - c_0 - \frac{5}{2}\kappa_{2,r}, \tag{9.28}$$

where

$$\kappa_{1,r} = \frac{\phi(r)}{r} c_0 + \max(\mathrm{err}_{r,Q_0/sq}, 0).$$

$$\kappa_{2,r} = c_0 + \sum_{p|r} \frac{\log p}{p} + \frac{r}{\phi(r)} \min(\mathrm{err}_{r,Q/qs}, 0).$$

We can assume $Q_0/qs \geq p_r$, where p_r is the least prime not dividing r, as otherwise the result follows from Lemma 9.6. Just as in (6.15)–(6.16), we establish an upper bound on $\max_{R_0 \geq p_r} \mathrm{err}_{r,R_0}$ and a lower bound on $\max_{R \geq 200} \mathrm{err}_{r,R}$ by Proposition 6.1 and some computation. We thus obtain an upper bound $c_{1,r}$ on $\kappa_{1,r}$ and a lower bound $c_{2,r}$ on $\kappa_{2,r}$ for each $r|2310$. Here $c_{2,r}$ is the same as c_r in the proof of Lemma 9.6.

We must now prove that (9.28) holds for all q with $\gcd(q, 2310) = r$. For the same reason as in the proof of Lemma 9.5, it is enough to consider q of the form $q = r \cdot \prod_{11 < p \leq z} p$, $q > 1$. Again, we can use the bounds (9.21)–(9.22); thanks to them, for $z \geq 17$, it is enough to show that

$$0.662865z - \log \frac{2310}{r}$$

is at least

$$\frac{7}{2} \frac{\phi(2310/r)}{2310/r} e^{\gamma} \left(\log z + \frac{1}{\log z} \right) \left(\frac{\phi(r)}{r} \left(\log z - \sum_{p \leq 11} \frac{\log p}{p} \right) + c_{1,r} \right) - c_0 - \frac{5}{2} c_{2,r}.$$

For each $r|2310$, a simple computation shows that this inequality holds for $z = 30$, and that the derivative of the first expression is also greater than the derivative of the second one at $z = 30$; since the second expression is concave, it follows that the inequality holds for all $z \geq 30$.

We check (9.28) for all $q = r \cdot \prod_{11 < p \leq z} p$, $z < 30$, directly. The only failures are for

$$q \in \{1, 2, 3, 5, 10, 15, 21, 39, 70, 105, 165, 195, 1365\}.$$

What is more, (9.28) holds also for all such q other than $q = 2, 5$ if we set $c_{1,r} = c_0 + \max_{R_0 \geq p'_r} \mathrm{err}_{r,R}$ instead of $c_{1,r} = c_0 + \max_{R_0 \geq p_r} \mathrm{err}_{r,R}$ as an upper bound on $\kappa_{1,r}$ for $r = \gcd(q, 2310)$, where p'_r is the least prime that does not divide r and is larger than 3. We then have to consider the case $p_r \leq R_0 < p_r'$ separately, but we can easily do so: we just have to check, much as in the proof of Lemma 9.6, that

$$L_q(R) \geq \left(1 + \frac{\log R}{\log 360 + c_0}\right) \cdot \begin{cases} 2 & \text{if } q \in \{3, 15, 21, 39, 105, 165, 195, 1365\}, \\ 3/2 & \text{if } q \in \{10, 70\}. \end{cases}$$

For all cases except 105, 165 and 1365, the inequality follows immediately from

$$L_q(R) \geq \frac{\phi(q)}{q}(\log R + c_{1,r}),$$

where $r = \gcd(q, 2310)$, and $R \geq 200$. For $q \in \{105, 165, 1365\}$, the inequality follows in the same way for $R \geq 275$, and can be checked case-by-case for $200 \leq R \leq 275$.

It remains to consider $q = 2$ and $q = 5$. Inequality (9.16) is certainly satisfied for $\eta = 5/2$, $q = 2$ and $sq \geq 4$, given the bounds $\mathrm{err}_{2,Q/sq} \geq \max_{R\geq 200} \mathrm{err}_{2,R} \geq -0.00906$ and $\mathrm{err}_{2,Q_0/sq} \leq \max_{R_0 \geq 3} \mathrm{err}_{2,R_0} \leq 0.11112$. If $sq < 4$, then $Q_0/sq > 360/4 = 90$, and so we can use the better bound $\mathrm{err}_{2,Q_0/sq} \leq \max_{R_0\geq 90} \mathrm{err}_{2,R_0} \leq 0.01947$, which also gives us that (9.16) is satisfied for $\eta = 5/2$. Hence, in either case, Lemma 9.3 gives us the conclusion we want.

The case $q = 5$, $\eta = 5/2$ is similar: for $sq \geq 6$, the bounds $\mathrm{err}_{2,Q/sq} \geq \max_{R\geq 200} \mathrm{err}_{2,R} \geq -0.01924$ and $\mathrm{err}_{2,Q_0/sq} \leq \max_{R_0\geq 2} \mathrm{err}_{2,R_0} \leq 0.29754$ give us inequality (9.16); for $sq < 6$, we have $Q_0/sq > 360/6 = 60$, and so we use the better bound $\mathrm{err}_{5,Q_0/sq} \leq \max_{R_0 \geq 60} \mathrm{err}_{5,R_0} \leq 0.03324$ to establish (9.16). □

We now pass to the main result of the section. The constant $c_0 = 1.33258\ldots$ in the numerator of the right side of inequality (9.27) will now be replaced by $c_+ = 1.35$. The reason is that, while (9.27) still holds asymptotically for $q = 1$, errors terms exist.

Proposition 9.8. *Let L_q be as in (9.5). Let c_0 be as in (6.3). Let $Q_0 \geq 360$, $Q \geq 200Q_0$ with $Q_0 \leq Q^{2/3}$. Let $1 \leq q \leq Q_0$, $1 \leq s \leq Q_0/q$, with $q \in \mathbb{Z}$ and $s \in \mathbb{R}$. Then*

$$\frac{L_q(Q_0/sq)}{L_q(Q/sq)} \leq \frac{\log Q_0 + c_+}{\log Q + c_0}, \tag{9.29}$$

where $c_+ = 1.35$.

As should be clear by now, none of the constants in the conditions $Q_0 \geq 360$, $Q \geq 200Q_0$, $Q_0 \leq Q^{2/3}$ is in itself meaningful: some can be relaxed if others are strengthened. The value of c_+ could also be lowered a little.

Proof. Since $Q_0 \geq 360$ and $Q_0 \leq Q^{2/3}$, we know that $\log Q_0 > 4c_0$, and so $\log Q/Q_0 \geq (\log Q_0)/2 \geq (2/5)(\log Q_0 + c_0)$. We can thus apply Proposition 9.7, and obtain that (9.29) holds for $q > 1$.

Assume now that $q = 1$. The case $s \geq 2$ is easy: by (6.15)–(6.16),

$$\log 2 > \frac{7}{2} \max_{R\geq 1} \mathrm{err}_{1,R} - \frac{5}{2} \max_{R\geq 200} \mathrm{err}_{1,R},$$

and so Lemma 9.3 tells us we are done. The case $3/2 \leq s < 2$ is the same, with the difference that we use that

$$\log \frac{3}{2} > \frac{7}{2} \max_{R \geq 180} \mathrm{err}_{1,R} - \frac{5}{2} \max_{R \geq 200} \mathrm{err}_{1,R}.$$

Consider, then, $1 \leq s \leq 3/2$. By definition of $\mathrm{err}_{q,R}$, for $s \leq 3/2$,

$$\frac{L_1(Q_0/s)}{L_1(Q/s)} \leq \frac{\log Q_0/s + c_0 + \mathrm{err}_{1,Q_0/s}}{\log Q/s + c_0 + \mathrm{err}_{1,Q/s}} \leq \frac{\log Q_0 + c_0 + \delta_1}{\log Q + c_0 - \delta_2},$$

where

$$\delta_1 = \max_{R \geq 240} \mathrm{err}_{1,R} \leq 0.01379,$$

$$\delta_2 = - \max_{R \geq 48000} \mathrm{err}_{1,R} \leq 0.001.$$

(As a larger computation establishes, the true value of δ_2 is lower, but the bound here is enough.) Then, to establish the upper bound $(\log Q_0 + c_+)/(\log Q + c_0)$ on $L_1(Q_0/s)/L_1(Q/s)$, it suffices to show that, for $\delta = c_+ - c_0$,

$$(\log Q_0 + c_0 + \delta)(\log Q + c_0 - \delta_2) \geq (\log Q_0 + c_0 + \delta_1)(\log Q + c_0).$$

Since $(\log Q_0 + c_0) \leq (5/7)(\log Q + c_0)$, this inequality follows from

$$\left(\delta - \delta_1 - \frac{5}{7}\delta_2\right)(\log Q + c_0) \geq \delta\delta_2 \tag{9.30}$$

Since $\delta - \delta_1 - (5/7)\delta_2 \geq 0.00291$, $\log Q + c_0 \geq \log 48000 + c_0 \geq 12.11153$ and $\delta\delta_2 \leq 0.00002$, inequality (9.30) holds, and so we are done. □

9.4 CONCLUSIONS

We first state our final result on the large sieve, as it is applicable elsewhere. It combines Ramaré's result (Prop. 9.1) with the work in §6.1–9.3 (including bounds using [RA17, Thm. 1.1]) and the large sieve.

Theorem 9.9. *Let $\{a_n\}_{n=1}^{\infty}$, $a_n \in \mathbb{C}$, have support on an interval $[x_0, x_0 + x]$. Let $Q_0 \geq 360$, $Q \geq 200Q_0$ with $Q_0 \leq Q^{2/3}$. Assume that $a_n = 0$ whenever n has a prime factor $\leq Q$.*

Let $S(\alpha) = \sum_n a_n e(\alpha n)$ for $\alpha \in \mathbb{R}/\mathbb{Z}$. Then

$$\sum_{q \leq Q_0} \sum_{\substack{a \bmod q \\ (a,q)=1}} |S(a/q)|^2 \leq \frac{\log Q_0 + c_+}{\log Q + c_0} \cdot (x + Q^2) \sum_n |a_n|^2, \tag{9.31}$$

where $c_+ = 1.35$ and $c_0 = \gamma + \sum_p \frac{\log p}{p(p-1)} = 1.33258227\ldots$.

Proof. By Propositions 9.1 and 9.8,

$$\sum_{q\leq Q_0} \sum_{\substack{a \bmod q \\ (a,q)=1}} |S(a/q)|^2 \leq \frac{\log Q_0 + c_+}{\log Q + c_0} \sum_{q\leq Q} \sum_{\substack{a \bmod q \\ (a,q)=1}} |S(a/q)|^2 .$$

By the large sieve (as in [MV74] or [Sel91, §20]; see the exposition in [IK04, §7.4], or apply Prop. 8.4 with $\eta = 1_{[x_0/x, 1+x_0/x]}$),

$$\sum_{q\leq Q} \sum_{\substack{a \bmod q \\ (a,q)=1}} |S(a/q)|^2 \leq (x+Q^2) \sum_n |a_n|^2.$$

□

Let us finally prove the result we will actually use.

Theorem 9.10. *Let* $\{a_n\}_{n=1}^{\infty}$, $a_n \in \mathbb{C}$, *be in* ℓ^1. *Let* $Q_0 \geq 360$, $Q \geq 200Q_0$ *with* $Q_0 \leq Q^{2/3}$. *Assume that* $a_n = 0$ *whenever* n *has a prime factor* $\leq Q$.

Let $S(\alpha) = \sum_n a_n e(\alpha n)$ *for* $\alpha \in \mathbb{R}/\mathbb{Z}$. *For* $R \geq 1$, $\beta > 0$, *let*

$$M_{R,\beta} = \bigcup_{q\leq R} \bigcup_{\substack{a \bmod q \\ (a,q)=1}} \left(\frac{a}{q} - \frac{R}{q}\beta, \frac{a}{q} + \frac{R}{q}\beta\right). \tag{9.32}$$

Then, for any $\beta \leq 1/2Q^2$,

$$\int_{M_{Q_0,\beta}} |S(\alpha)|^2 \, d\alpha \leq \frac{\log Q_0 + c_+}{\log Q + c_0} \sum_n |a_n|^2, \tag{9.33}$$

where $c_+ = 1.35$ *and* $c_0 = \gamma + \sum_p \frac{\log p}{p(p-1)} = 1.33258227\ldots$.

Proof. By Propositions 9.2 and 9.8. □

Final remarks. (a) It would be interesting to see what one could do with a variant of Prop. 9.1 that reduces matters to (8.6) instead of (8.5). Presumably the final bound would improve for the habitual reason why (8.6) is sometimes used instead of (8.5): (8.6) lets one exploit the fact that the rationals a/q, $q \leq Q$ are unevenly spaced.

(b) As inequality (9.27) is particularly clean, it would be tempting to hope for a clean, short proof. It is in fact the case that conditions have been relaxed and the need for computation has been lessened in comparison to the first version of this chapter in [Helc, §5] – thanks in part to the use of improved bounds from [RA17], but also because the proof has been completely redone.

However, we probably should not hope for a short, analysis-free proof of the fact that the minimum of $L_q(Q_0/sq)/L_q(Q/sq)$ is reached when $s = q = 1$. For one thing, as we already made clear in the comment after the statement of Prop. 9.7, our conditions on Q_0 and Q are fairly tight. Here is another, simple example showing that some conditions on Q_0 and Q are needed: by (6.11) or (6.12), for $Q_0 = 2$ and Q large,

$$\frac{L_2\left(\frac{Q_0}{2}\right)}{L_2\left(\frac{Q}{2}\right)} \sim \frac{1}{\frac{1}{2}\left(\log\frac{Q}{2} + c_0 + \frac{\log 2}{2}\right)} = \frac{2}{\log Q + c_0 - \frac{\log 2}{2}} > \frac{2}{\log Q + c_0} \sim \frac{L(Q_0)}{L(Q)}.$$

It may thus be best to accept that (9.27) is an analytic inequality, and not hope for an algebraic proof. The same then goes for Theorems 9.9 and 9.10.