Chapter 1

Arithmetic Functions

1.1. The method of Chebyshev

Around 1792 J. K. F. Gauss counted primes in successive blocks of a thousand integers, and noticed that the sequence of primes seems to thin out according to a definite law. One formulation of his law is that the intervals $(x-h, x]$ contain about $h/\log(x)$ primes when x is large and h is not too small. Another way to express the same empirical observation is that the chance of a randomly chosen large integer n being prime is approximately $1/\log(n)$. Gauss went on to integrate this density to obtain the approximation

$$\pi(x) \stackrel{\text{def}}{=} \sum_{p \le x} 1 \approx \mathrm{li}(x) \stackrel{\text{def}}{=} \int_2^x \frac{du}{\log(u)}$$

for the counting function $\pi(x)$ of the primes. Here $\mathrm{li}(x)$ is the *integral logarithm.* By his counts of primes, Gauss guessed that $\lim_{x\to+\infty} \pi(x)/\mathrm{li}(x) = 1$. This is simply the statement that the relative error $|\pi(x) - \mathrm{li}(x)|/\pi(x)$ in the approximation goes to zero as $x \to +\infty$. Using the notation

$$f(x) \sim g(x) \quad \stackrel{\text{def}}{\iff} \quad \lim_{x\to+\infty} \frac{f(x)}{g(x)} = 1$$

of asymptotic equality familiar from analysis, the conjecture can be expressed as $\pi(x) \sim \mathrm{li}(x)$. This is the Prime Number Theorem (abbreviated PNT) first proved by J. S. Hadamard and C. G. J. N. de la Vallée Poussin in 1896. Since $\mathrm{li}(x) \sim x/\log(x)$ by l'Hôpital's rule, the PNT also has the formulation $\pi(x) \sim x/\log(x)$, showing that about $1/\log(x)$ of the positive integers up to x are prime.

Because the density of the primes near n is approximately $1/\log(n)$, it may be more natural to count each prime p with weight $\log(p)$ rather than

with weight 1. This gives the weighted counting function

$$\vartheta(x) \stackrel{\text{def}}{=} \sum_{p \le x} \log(p)$$

introduced by P. L. Chebyshev. To obtain nice formulas that are easier to analyze, it is advantageous also to count prime powers p^k with weight $\log(p)$. The *von Mangoldt function* Λ given by $\Lambda(p^k) = \log(p)$ when the argument is a prime power, and zero otherwise, serves this purpose. The weighted counting function

$$\psi(x) \stackrel{\text{def}}{=} \sum_{p^k \le x} \log(p) = \sum_{n \le x} \Lambda(n)$$

was also introduced by Chebyshev. Since prime powers with exponent higher than one are quite sparse, $\psi(x)$ mainly counts primes with weight $\log(p)$. The functions $\psi(x)$ and $\vartheta(x)$ are thus approximately equal, and both are closely related to the counting function $\pi(x)$ of the primes. In particular it will turn out that the Prime Number Theorem may equally well be expressed as one of the asymptotic relations $\psi(x) \sim x$ or $\vartheta(x) \sim x$. These formulations are often more convenient.

The integers $n = pm$ in the interval $0 < n \le N$ that are divisible by a prescribed prime p are given by the integer solutions m of the inequality $0 < m \le N/p$. The largest integer less than or equal to a real number x is denoted by $[x]$. It is called the *integer part* of x or the *Gauss bracket*. Clearly $[N/p]$ is the number of integers n as above. The same reasoning shows that, of these integers, exactly $[N/p^k]$ are divisible by p^k. This observation allows us to write down the prime factorization

$$N! = \prod_{p^k} p^{[N/p^k]}$$

of the factorial, due to A.-M. Legendre. The product is taken over all prime powers, but has only finitely many factors different from 1 because $[N/p^k] = 0$ when $p^k > N$. The importance of the identity lies in the fact that the left-hand side does not contain the primes explicitly, and is susceptible of being estimated analytically. Taking the logarithm on both sides of the Legendre identity yields

$$\sum_{n \le N} \Lambda(n) \left[\frac{N}{n}\right] = \sum_{p^k} \log(p) \left[\frac{N}{p^k}\right] = \log(N!).$$

Now

$$\sum_{p^k} \log(p) \left[\frac{N}{p^k}\right] = \sum_{p^k} \log(p) \sum_{mp^k \le N} 1 = \sum_{m \le N} \sum_{p^k \le N/m} \log(p)$$

by interchanging the order of summation in the double sum. Then

$$\sum_{m\le N}\sum_{n\le N/m}\Lambda(n)=\log(N!)=\sum_{n\le N}\log(n).$$

Any mapping $f:\mathbb{N}\to\mathbb{C}$ from the positive integers into the complex numbers is called an *arithmetic function.* The functions Λ and log are examples of arithmetic functions. Every arithmetic function f has a *summatory function*

$$F(x)=\sum_{n\le x}f(n).$$

Thus ψ is the summatory function of Λ. The summatory function

$$T(x)\stackrel{\text{def}}{=}\sum_{n\le x}\log(n)$$

of log is also important. The identity

$$\sum_{m\le x}\psi\Big(\frac{x}{m}\Big)=\sum_{n\le x}\Lambda(n)\Big[\frac{x}{n}\Big]=T(x)$$

holds for nonnegative x since $\log(N!)=T(x)$ where $N=[x]$. This identity is the starting point for the method of Chebyshev.

Proposition 1.1. *The inequalities*

$$\log(2)x-\log(4x)\le\psi(x)\le 2\log(2)x+\frac{\log^2(x)}{\log(2)}$$

hold for $x\ge 1$.

Proof. The terms in the last sum in the computation

$$\begin{aligned}T(x)-2\,T\Big(\frac{x}{2}\Big)&=\sum_{n\le x}\log(n)-2\sum_{m\le x/2}\log(m)\\&=\sum_{n\le x}\log(n)-2\sum_{2m\le x}\log(2m)+2\sum_{2m\le x}\log(2)\\&=\sum_{n\le x}(-1)^{n-1}\log(n)+2\Big[\frac{x}{2}\Big]\log(2)\end{aligned}$$

alternate in sign and increase in magnitude. So

$$\Big|T(x)-2\,T\Big(\frac{x}{2}\Big)-2\Big[\frac{x}{2}\Big]\log(2)\Big|\le\log([x])$$

for $x\ge 1$. Thus

$$\log(2)x-\log(4x)\le T(x)-2\,T\Big(\frac{x}{2}\Big)\le\log(2)x+\log(x).$$

Substituting the expression

$$T(x)=\sum_{n\le x}\psi\Big(\frac{x}{n}\Big)$$

into $T(x) - 2\,T(x/2)$ yields

$$\psi(x) - \psi\Big(\frac{x}{2}\Big) + \psi\Big(\frac{x}{3}\Big) - \cdots = T(x) - 2\,T\Big(\frac{x}{2}\Big).$$

Then

$$\psi(x) \geq \log(2)x - \log(4x)$$

since ψ is an increasing and nonnegative function, and

$$\psi(x) - \psi\Big(\frac{x}{2}\Big) \leq \log(2)x + \log(x)$$

for the same reason. Adding up the inequalities

$$\psi\Big(\frac{x}{2^j}\Big) - \psi\Big(\frac{x}{2^{j+1}}\Big) \leq \log(2)2^{-j}x + \log(x)$$

for $j = 0, 1, 2, \ldots, [\log(x)/\log(2)] - 1$ yields

$$\psi(x) \leq 2\log(2)x + \left[\frac{\log(x)}{\log(2)}\right]\log(x) \leq 2\log(2)x + \frac{\log^2(x)}{\log(2)}$$

since $\psi(x/2^{j+1}) = 0$ when $x/2^{j+1} < 2$. □

The inequalities in Proposition 1.1 and the limits $\log^2(x)/(x\log(2)) \to 0$ and $\log(4x)/x \to 0$ as $x \to +\infty$ imply that for every $\varepsilon > 0$ there exists some $x_0(\varepsilon)$ so that $\log(2) - \varepsilon < \psi(x)/x < 2\log(2) + \varepsilon$ for $x \geq x_0(\varepsilon)$. Such ε-$x_0(\varepsilon)$ inequalities are often expressed in a somewhat different but equivalent way, using concepts from analysis. The *limit superior* $\limsup_{x\to+\infty} f(x)$ of a bounded real function $f(x)$ on an interval $[a, \infty)$ is the unique real number σ such that $f(x) < \sigma+\varepsilon$ holds for all x sufficiently large, while $f(x) < \sigma-\varepsilon$ fails for some x arbitrarily large, no matter how small $\varepsilon > 0$ is taken. Similarly the *limit inferior* $\liminf_{x\to+\infty} f(x)$ is the unique real number ι such that $f(x) > \iota - \varepsilon$ holds for all x sufficiently large, while $f(x) > \iota + \varepsilon$ fails for some x arbitrarily large, no matter how small $\varepsilon > 0$ is taken. That σ and ι must necessarily exist is a consequence of the completeness property of the real number system. Define

$$\mathbf{a} \stackrel{\text{def}}{=} \liminf_{x\to+\infty} \frac{\psi(x)}{x} \quad \text{and} \quad \mathbf{A} \stackrel{\text{def}}{=} \limsup_{x\to+\infty} \frac{\psi(x)}{x}.$$

Then the ε-$x_0(\varepsilon)$ bounds for $\psi(x)/x$ can be reformulated as the statement that $\log(2) \leq \mathbf{a} \leq \mathbf{A} \leq 2\log(2)$.

The definitions of ψ and ϑ yield

$$\psi(x) = \sum_{p^k \leq x} \log(p) = \sum_{k=1}^{\infty} \sum_{p \leq x^{1/k}} \log(p) = \vartheta(x) + \vartheta(x^{1/2}) + \vartheta(x^{1/3}) + \cdots,$$

and so

$$\psi(x) - 2\psi(x^{1/2}) = \vartheta(x) - \vartheta(x^{1/2}) + \vartheta(x^{1/3}) - \cdots.$$

The inequality $\psi(x)-2\psi(x^{1/2}) \le \vartheta(x) \le \psi(x)$ follows since ϑ is an increasing and nonnegative function, and then

$$\log(2) \le \liminf_{x\to+\infty} \frac{\vartheta(x)}{x} \le \limsup_{x\to+\infty} \frac{\vartheta(x)}{x} \le 2\log(2)$$

by our estimates on $\psi(x)$.

Clearly $\vartheta(x) \le \pi(x)\log(x)$, and so $\pi(x)/(x/\log(x)) \ge \vartheta(x)/x$. Finding an upper bound for $\pi(x)$ is only slightly more challenging. The inequality

$$\vartheta(x) \ge \sum_{y<p\le x} \log(p) \ge (\pi(x)-\pi(y))\log(y)$$

yields

$$\frac{\pi(x)}{x/\log(y)} \le \frac{\vartheta(x)}{x} + \frac{\pi(y)}{x/\log(y)} \le \frac{\vartheta(x)}{x} + \frac{y}{x/\log(y)}.$$

To obtain the desired upper bound, we must let y increase fast enough with x so that the left-hand side of the inequality is close to $\pi(x)/(x/\log(x))$ for large x, while the second term on the right-hand side should become negligible in comparison. The choice $y = x/\log^2(x)$ works well, giving

$$\frac{\pi(x)}{x/(\log(x)-2\log\log(x))} \le \frac{\vartheta(x)}{x} + \frac{x/\log^2(x)}{x/(\log(x)-2\log\log(x))}.$$

The second term on the right-hand side tends to zero, and

$$\frac{\frac{\pi(x)}{x/(\log(x)-2\log\log(x))}}{\frac{\pi(x)}{x/\log(x)}} = \frac{\log(x)-2\log\log(x)}{\log(x)} \to 1$$

as $x \to +\infty$. Thus

$$\liminf_{x\to+\infty} \frac{\vartheta(x)}{x} \le \liminf_{x\to+\infty} \frac{\pi(x)}{x/\log(x)} \le \limsup_{x\to+\infty} \frac{\pi(x)}{x/\log(x)} \le \limsup_{x\to+\infty} \frac{\vartheta(x)}{x}$$

and so

$$\log(2) \le \liminf_{x\to+\infty} \frac{\pi(x)}{x/\log(x)} \le \limsup_{x\to+\infty} \frac{\pi(x)}{x/\log(x)} \le 2\log(2).$$

Since $\lim_{x\to+\infty} \mathrm{li}(x)/(x\log(x)) = 1$, the last inequality shows that $\pi(x)$ and $\mathrm{li}(x)$ have the same order of growth.

Choosing $y = x^a$ with $0 < a < 1$ and a otherwise, arbitrarily yields

$$\frac{\vartheta(x)}{x} \le \frac{\pi(x)}{x/\log(x)} \le \frac{1}{a}\frac{\vartheta(x)}{x} + \frac{x^a}{x/\log(x)};$$

thus we see that $\pi(x) \sim x/\log(x)$, and $\pi(x) \sim \mathrm{li}(x)$, and $\vartheta(x) \sim x$, and $\psi(x) \sim x$, and $\mathbf{a} = \mathbf{A} = 1$, are equivalent formulations of the Prime Number Theorem.

1.2. Bertrand's Postulate

The method of Chebyshev does not prove the Prime Number Theorem, but it does show that $\psi(x)$ has the expected order of growth. By a more elaborate version of his method Chebyshev obtained bounds such as $.921{\cdot}x \le \psi(x) \le 1.106{\cdot}x$ for all x sufficiently large. He used these bounds to give the first proof of Bertrand's Postulate. Today Bertrand's Postulate is taken as the statement that for all $x \ge 2$ there is at least one prime in the interval $(x/2, x]$. The weaker bounds in Proposition 1.1 do not suffice to prove this with Chebyshev's approach. But we shall obtain the result anyway using an idea of S. A. Ramanujan.

Proposition 1.2 (Bertrand's Postulate)**.** *For every $x \ge 2$ there exists at least one prime p with $x/2 < p \le x$.*

Proof. For $2 \le x \le 797$ the interval $(x/2, x]$ contains a prime by a trick of E. G. H. Landau: The chain $2, 3, 5, 7, 11, 17, 31, 59, 107, 211, 401, 797$ consists of primes and each is smaller than twice its predecessor. To detect primes in the intervals $(x/2, x]$ for $x \ge 797$ we show that the difference $\vartheta(x) - \vartheta(x/2)$ is positive.

Fetch the inequality

$$\psi(x) - \psi\Big(\frac{x}{2}\Big) + \psi\Big(\frac{x}{3}\Big) \ge T(x) - 2\,T\Big(\frac{x}{2}\Big)$$

from the proof of Proposition 1.1. Retention of the term $\psi(x/3)$ is an idea due to Ramanujan. Now

$$\begin{aligned}
\vartheta(x) - \vartheta\Big(\frac{x}{2}\Big) &\ge \psi(x) - 2\psi(x^{1/2}) - \psi\Big(\frac{x}{2}\Big) \\
&\ge T(x) - 2\,T\Big(\frac{x}{2}\Big) - \psi\Big(\frac{x}{3}\Big) - 2\psi(x^{1/2}) \\
&\ge \log(2)x - \log(4x) - 2\log(2)\frac{x}{3} \\
&\quad - \frac{\log^2(x/3)}{\log(2)} - 4\log(2)x^{1/2} - 2\frac{\log^2(x^{1/2})}{\log(2)} = f(x)
\end{aligned}$$

with

$$f(x) = \frac{\log(2)}{3}x - \log(4x) - \frac{3\log^2(x)}{2\log(2)} - 4\log(2)x^{1/2},$$

by Proposition 1.1 and its proof. The derivative

$$f'(x) = \frac{\log(2)}{3} - \frac{1}{x} - \frac{3\log(x)}{x\log(2)} - \frac{2\log(2)}{x^{1/2}}$$

is an increasing function on $x \ge 797$ since $\log(x)/x$ is decreasing on $x \ge e$. Then $f'(x) > 0$ on $x \ge 797$ because $f'(797) = 0.14$. Thus $f(x)$ is increasing on this interval and hence positive there because $f(797) = 1.2$. □

1.3. Simple estimation techniques

In analytic number theory arithmetical questions are characteristically answered by finding estimates that imply the desired conclusions. The proof of Bertrand's Postulate is typical; we wanted to know that every interval $(x/2, x]$ for $x \geq 2$ contains at least one prime, but to obtain this result we detoured through a lower bound for $\vartheta(x) - \vartheta(x/2)$. Carrying out an estimate often requires long calculations with inequalities, for which the usual notation proves cumbersome. The big-O notation of P. G. H. Bachmann is well adapted for efficient calculations with long chains of inequalities. The statement $f(x) = O(g(x))$ means that there exist some unspecified constants $C > 0$ and x_0 such that $|f(x)| \leq Cg(x)$ on the interval $x \geq x_0$. The bound $\psi(x) \leq 2\log(2)x + \log^2(x)/\log(2)$ for $x \geq 1$ would be expressed as $\psi(x) = O(x)$ in the big-O notation. The latter statement is less precise, but the kind of information suppressed in the big-O notation is often unimportant anyway. The following result is the key to labor-saving calculations using the big-O notation.

Proposition 1.3. *If $f_1(x) = O(g(x))$ and $f_2(x) = O(g(x))$ then $f_1(x) + f_2(x) = O(g(x))$. If also $f_3(x) = O(h(x))$ then $f_1(x)f_3(x) = O(g(x)h(x))$.*

Proof. There are $C_1, C_2, C_3 > 0$ and x_1, x_2, x_3 such that $|f_1(x)| \leq C_1 g(x)$ for $x \geq x_1$, $|f_2(x)| \leq C_2 g(x)$ for $x \geq x_2$ and $|f_3(x)| \leq C_3 h(x)$ for $x \geq x_3$. Then

$$|f_1(x) + f_2(x)| \leq |f_1(x)| + |f_2(x)| \leq C_1 g(x) + C_2 g(x) = (C_1 + C_2)g(x)$$

for $x \geq \max(x_1, x_2)$ and

$$|f_1(x)f_3(x)| = |f_1(x)||f_2(x)| \leq C_1 g(x) C_3 h(x) = (C_1 C_3) g(x) h(x)$$

for $x \geq \max(x_1, x_3)$. □

Note in particular that the largest of the O-terms in a sum absorbs everything smaller. The notation $f(x) \ll g(x)$ due to I. M. Vinogradov is synonymous with $f(x) = O(g(x))$. The notation $f(x) \asymp g(x)$ of G. H. Hardy means that $f(x) \ll g(x)$ and $g(x) \ll f(x)$. There is also a small-o notation due to Landau. The statement $f(x) = o(g(x))$ means that $\lim_{x\to+\infty} f(x)/g(x) = 0$, or equivalently, for any $\varepsilon > 0$ there exists some $x_0(\varepsilon)$ so that $|f(x)| \leq \varepsilon g(x)$ for $x \geq x_0(\varepsilon)$.

If the function f in an estimate $f = O(g)$ depends on one or more parameters, say $f_a(x) = O_a(g(x))$, the question of uniformity of the estimate is often of considerable importance. If it is possible to choose both C and x_0 in the statement that

$$|f_a(x)| \leq Cg(x) \quad \text{for} \quad x \geq x_0$$

independently of the parameter a, then the estimate is said to be *uniform* in a. Neglect of uniformity is a recognized source of error in number theoretical arguments. The case where C is independent of a while x_0 depends on a in some (unobvious) way is especially insidious; the expected inequality may hold on *some* such interval for *each* value of a, but there may be *no* such interval on which the inequality holds for *all* values of a.

From the next section onward, and all through the book, we frequently encounter integrals that must be estimated. That is to say, be shown to be small in absolute value, or at least not too big. So some remarks on this topic are in order at this point.

Let $f : I \to \mathbb{C}$ be a continuous function on some interval $I \subseteq \mathbb{R}$. Consider the problem of bounding

$$\left| \int_I f(x)\, dx \right|$$

from above. The inequality

$$\left| \int_I f(x)\, dx \right| \leq \int_I |f(x)|\, dx$$

is basic here, but we will briefly describe some techniques that go a little further. Note that

$$\left| \int_I f(x)\, dx \right| \leq \int_I |f(x)|\, dx \leq \int_I M\, dx = M\ell(I)$$

if $|f(x)| \leq M$ for $x \in I$. Here $\ell(I)$ denotes the length of I. An extension of this argument yields

$$\left| \int_I f(x)g(x)\, dx \right| \leq M \int_I |g(x)|\, dx$$

if $|f(x)| \leq M$ for $x \in I$. For integrals where the integrand is the product of two functions, one of which is oscillatory, integration by parts is often useful. We have

$$\int_a^b f(x)g(x)\, dx = f(b)G(b) - \int_a^b f'(x)G(x)\, dx$$

assuming f continuously differentiable on $[a, b]$ and putting

$$G(x) = \int_a^x g(u)\, du.$$

If $g(x)$ oscillates on $[a, b]$, there is a good possibility that $G(x)$ will grow slowly on the interval. If in addition $f(x)$ changes fairly slowly on $[a, b]$, its derivative $f'(x)$ will be small in magnitude, and integration by parts may yield a very favorable estimate of the integral of $f(x)g(x)$ over $[a, b]$. This is a standard technique for estimating Fourier transforms.

Proposition 1.4 (Hölder inequality)**.** *If I is an interval and f and g are continuous functions on I then*

$$\int_I |f(x)g(x)|\,dx \le \left(\int_I |f(x)|^p\,dx\right)^{1/p} \left(\int_I |g(x)|^q\,dx\right)^{1/q},$$

where $1 < p, q < \infty$ with $1/p + 1/q = 1$.

Proof. The case where

$$\int_I |f(x)|^p\,dx = 0 \quad \text{or} \quad \int_I |g(x)|^q\,dx = 0$$

is trivial. Multiplying f and g by suitable positive real numbers, we may therefore assume that

$$\int_I |f(x)|^p\,dx = \int_I |g(x)|^q\,dx = 1$$

by homogeneity. Now

$$|f(x)g(x)| = (|f(x)|^p)^{1/p}\,(|g(x)|^q)^{1/q} = (|f(x)|^p)^{1/p}\,(|g(x)|^q)^{1-1/p},$$

so

$$\log(|f(x)g(x)|) = \frac{1}{p}\log\left(|f(x)|^p\right) + \left(1 - \frac{1}{p}\right)\log\left(|g(x)|^q\right).$$

But the logarithm function is strictly concave, that is to say, all the chords lie strictly below the graph except for their endpoints. Hence

$$\begin{aligned}\frac{1}{p}\log\left(|f(x)|^p\right) + \left(1 - \frac{1}{p}\right)\log\left(|g(x)|^q\right)& \\ \le \log\left(\frac{1}{p}\log\left(|f(x)|^p\right) + \left(1 - \frac{1}{p}\right)\log\left(|g(x)|^q\right)\right),&\end{aligned}$$

and so

$$|f(x)g(x)| \le \frac{1}{p}\log\left(|f(x)|^p\right) + \left(1 - \frac{1}{p}\right)\log\left(|g(x)|^q\right).$$

Integrating over I yields

$$\int_I |f(x)g(x)|\,dx \le \frac{1}{p}\cdot 1 + \left(1 - \frac{1}{p}\right)\cdot 1 = 1,$$

and this proves the inequality. □

The case $p = 2$ is especially important; this is the Cauchy-Schwarz inequality. The Hölder inequality extends by induction to estimate integrals of products of more than two functions. We have

$$\int_I |f_1(x)\cdots f_n(x)|\,dx \le \left(\int_I |f_1(x)|^{p_1}\,dx\right)^{1/p_1} \cdots \left(\int_I |f_n(x)|^{p_n}\,dx\right)^{1/p_n}$$

if $1 < p_1, \ldots, p_n < \infty$ with $1/p_1 + \cdots + 1/p_n = 1$.

1.4. The Mertens estimates

Sums

$$\sum_{p\le x} f(p)$$

over primes occur frequently in analytic number theory. In the simpler cases, f is a positive, continuous, monotone function of a real variable that does not change rapidly. If the sum diverges as $x \to +\infty$, its asymptotic behavior may be guessed from the heuristic

$$\sum_{p\le x} f(p) \sim \sum_{2\le n\le x} \frac{f(n)}{\log(n)} \sim \int_2^x \frac{f(u)\,du}{\log(u)},$$

which is inspired by the observation that the density of the primes near n is close to $1/\log(n)$. The latter statement is a formulation of the Prime Number Theorem, and indeed the PNT with a good estimate for the error term is a natural tool with which to estimate such sums. As an example of the heuristic in action, consider the sum of $\log(p)/p$ over primes $p \le x$. The guess for the asymptotic behavior is

$$\sum_{p\le x} \frac{\log(p)}{p} \sim \sum_{2\le n\le x} \frac{\log(n)/n}{\log(n)} \sim \int_2^x \frac{du}{u} \sim \log(x),$$

and this is actually correct. Indeed, F. C. J. Mertens proved in 1874 that the absolute error in the asymptotic approximation is bounded. This may be expressed as

$$\sum_{p\le x} \frac{\log(p)}{p} = \log(x) + O(1)$$

by means of the big-O notation for the error term.

The heuristic for guessing the asymptotic behavior of sums over primes will not perform satisfactorily if f does not have the nice properties assumed. If, for example, f changes sign, there will be cancellation in the sum, and the underlying rationale for the heuristic does not take account of this.

Partial summation is perhaps the tool most frequently applied in analytic number theory. The basic version is the identity

$$\sum_{m=1}^{n} a_m b_m = b_n \sum_{m=1}^{n} a_m - \sum_{m=1}^{n-1} (b_{m+1} - b_m) \sum_{k=1}^{m} a_k.$$

This is an analogue, for sums, of integration by parts. The partial summation identity is easily proved by observing that $b_j(a_1 + \cdots + a_j) = b_j(a_1 + \cdots + a_{j-1}) + a_j b_j$ and applying mathematical induction. The partial summation identity yields a formula that is very convenient for estimating weighted sums of arithmetic functions.

Proposition 1.5 (Partial summation)**.** *Let f be an arithmetic function and g a continuous function with piecewise continuous derivative on $[1,\infty)$. Then*

$$\sum_{n\le x} f(n)g(n) = F(x)g(x) - \int_1^x F(u)g'(u)\,du,$$

where F is the summatory function of f.

Proof. Calculate

$$\begin{aligned}\sum_{n\le x} f(n)g(n) &= F(x)g([x]) - \sum_{n=1}^{[x]-1}(g(n+1)-g(n))F(n)\\ &= F(x)g([x]) - \sum_{n=1}^{[x]-1} F(n)\int_n^{n+1} g'(u)\,du\\ &= F(x)g([x]) - \sum_{n=1}^{[x]-1}\int_n^{n+1} F(u)g'(u)\,du\\ &= F(x)g([x]) - \int_1^{[x]} F(u)g'(u)\,du\end{aligned}$$

by the partial summation identity and the Fundamental Theorem of Calculus. Then replace $[x]$ by x in the last step, for the resulting changes cancel. □

The partial summation formula is best understood in terms of the Stieltjes integral, but we eschew this refinement. The following bound for integrals involving the sawtooth function $S(x) = x - [x] - 1/2$ is sometimes useful.

Proposition 1.6. *If $1 \le a \le b$ the estimate*

$$\left|\int_a^b \frac{S(x)}{x^s}\,dx\right| \le \left(\frac{1}{8} + \frac{|s|}{16\sigma}\right) a^{-\sigma}$$

holds for any complex number $s = \sigma + it$ with positive real part σ.

Proof. If

$$\beta(x) = \frac{1}{16} + \int_0^x S(u)\,du$$

then $|\beta(x)| \le 1/16$. Integration by parts gives

$$\begin{aligned}\left|\int_a^b \frac{S(x)}{x^s}\,dx\right| &= \left|\frac{\beta(x)}{x^s}\Big|_a^b - \int_a^b (-s)\frac{\beta(x)}{x^{s+1}}\,dx\right| \le 2\cdot\frac{1/16}{a^\sigma} + \int_a^b |s|\frac{1/16}{x^{\sigma+1}}\,dx\\ &= \frac{1}{8a^\sigma} - \frac{|s|}{16\sigma}\,\frac{1}{x^\sigma}\Big|_a^b \le \frac{1}{8a^\sigma} + \frac{|s|}{16\sigma}\frac{1}{a^\sigma},\end{aligned}$$

since $S(x)$ is piecewise continuous. □

The last inequality yields asymptotic estimates for the summatory functions of $\log(m)$ and $1/m$. The first of these is a weak version of Stirling's formula.

Proposition 1.7. *The estimates*

$$\sum_{m=1}^{n} \log(m) = n\log(n) - n + \frac{\log(n)}{2} + 1 + R_n$$

and

$$\sum_{m=1}^{n} \frac{1}{m} = \log(n) + \gamma + \frac{1}{2n} + S_n$$

hold for all positive integers n with $|R_n| \le 3/16$ and $|S_n| \le 3/(16n^2)$.

Proof. The partial summation formula of Proposition 1.5 gives

$$\begin{aligned}\sum_{m=1}^{n} \log(m) &= n\log(n) - \int_1^n [u]\frac{du}{u} \\ &= n\log(n) - n + 1 + \frac{1}{2}\log(n) + \int_1^n S(u)\frac{du}{u}\end{aligned}$$

when $f(n) \equiv 1$ and $g(x) = \log(x)$. The last integral is bounded by $3/16$ in absolute value by Proposition 1.6. Using the partial summation formula again yields

$$\begin{aligned}\sum_{m=1}^{n} \frac{1}{m} &= n\cdot\frac{1}{n} - \int_1^n [u]\left(-\frac{du}{u^2}\right) \\ &= 1 + \log(n) + \frac{1}{2n} - \frac{1}{2} - \int_1^\infty S(u)\frac{du}{u^2} + \int_n^\infty S(u)\frac{du}{u^2}.\end{aligned}$$

The last term is bounded by $3/(16n^2)$ by Proposition 1.6. □

The real number $\gamma = 0.5772\ldots$ is known as the Euler-Mascheroni constant. It is unknown whether this is irrational. Note that Proposition 1.7 yields the version $T(x) = x\log(x) - x + O(\log(x))$ of Stirling's formula that is most commonly applied in analytic number theory.

Proposition 1.8 (Euler-Maclaurin summation formula)**.** *If $A < B$ are integers and f a continuous function on the interval $[A, B]$ with f' piecewise continuous there, then*

$$\sum_{n=A}^{B} f(n) = \int_A^B f(u)\,du + \frac{f(A) + f(B)}{2} + \int_A^B S(u)f'(u)\,du$$

with $S(u) = u - [u] - 1/2$ the sawtooth function.

Proof. Partial summation yields

$$\sum_{n=1}^{B} f(n) = Bf(B) - \int_1^B [u]f'(u)\,du$$

and

$$\sum_{n=1}^{A} f(n) = Af(A) - \int_1^A [u]f'(u)\,du.$$

Then

$$\int_A^B [u]f'(u)\,du = uf(u)\Big|_A^B - \sum_{n=A+1}^{B} f(n)$$

after taking the difference. Now

$$\int_A^B \left(u - \frac{1}{2}\right) f'(u)\,dt = \left(u - \frac{1}{2}\right) f(u)\Big|_A^B - \int_A^B f(u)\,du$$

by integration by parts. Subtracting the next to last formula from the last formula yields the Euler-Maclaurin summation formula. □

Despite the fact that anything obtainable from the Euler-Maclaurin summation formula may also be obtained by partial summation, resort to the former is sometimes more convenient. Moreover, repeated integration by parts in the Euler-Maclaurin summation formula yields a technique for obtaining precise approximations to sums. We make no use of this, so it is not covered here.

The next two results are due to Mertens. These depend on the Chebyshev bound $\psi(x) = O(x)$ in an essential way.

Proposition 1.9. *The estimates*

$$\sum_{p\le x} \frac{\log(p)}{p} = \sum_{n\le x} \frac{\Lambda(n)}{n} + O(1) = \log(x) + O(1)$$

hold.

Proof. First

$$x \sum_{n\le x} \frac{\Lambda(n)}{n} = \sum_{n\le x} \Lambda(n)\left[\frac{x}{n}\right] + O\left(\sum_{n\le x} \Lambda(n)\right)$$

because $0 \le x - [x] < 1$. Now

$$T(x) = \sum_{m\le x} \Lambda(m)\left[\frac{x}{m}\right]$$

yields

$$\sum_{n\le x}\frac{\Lambda(n)}{n}=\frac{T(x)}{x}+O\biggl(\frac{\psi(x)}{x}\biggr)=\log(x)+O(1)$$

by Stirling's formula and Proposition 1.1. The series

$$\sum_{p}\sum_{k=2}^{\infty}\frac{\log(p)}{p^k}$$

converges, and $\Lambda(n)$ is zero off the prime powers. □

Sometimes it is necessary to remove a factor from the terms of a sum. This is an important application of the partial summation formula, and is illustrated in the proof of the next result.

Proposition 1.10. *The estimate*

$$\sum_{p\le x}\frac{1}{p}=\log\log(x)+a+O\biggl(\frac{1}{\log(x)}\biggr)$$

holds with some constant a.

Proof. First

$$\begin{aligned}\sum_{p\le x}\frac{1}{p}&=\sum_{p\le x}\frac{\log(p)}{p}\frac{1}{\log(p)}\\&=\biggl(\sum_{p\le x}\frac{\log(p)}{p}\biggr)\frac{1}{\log(x)}-\int_2^x\biggl(\sum_{p\le u}\frac{\log(p)}{p}\biggr)\frac{(-1)\,du}{u\log^2(u)}\end{aligned}$$

by partial summation. Then

$$\begin{aligned}\sum_{p\le x}\frac{1}{p}&=1+O\biggl(\frac{1}{\log(x)}\biggr)+\int_2^x\frac{du}{u\log(u)}\\&\qquad+\int_2^x\biggl(\sum_{p\le u}\frac{\log(p)}{p}-\log(u)\biggr)\frac{du}{u\log^2(u)}\\&=1+O\biggl(\frac{1}{\log(x)}\biggr)+\log\log(x)-\log\log(2)\\&\qquad+\int_2^\infty\biggl(\sum_{p\le u}\frac{\log(p)}{p}-\log(u)\biggr)\frac{du}{u\log^2(u)}+\int_x^\infty\frac{O(1)}{u\log^2(u)}\,du\\&=\log\log(x)+a+O\biggl(\frac{1}{\log(x)}\biggr)\end{aligned}$$

by Proposition 1.9 and integration by parts. □

The next result is of considerable significance in prime number theory for various considerations of a probabilistic nature. The fact that the constant b in the formula is positive is important in such contexts. Actually b equals the Euler-Mascheroni constant γ, though we won't prove this.

Proposition 1.11 (Mertens' formula)**.** *The estimate*

$$\prod_{p\le x}\left(1-\frac{1}{p}\right)\sim\frac{e^{-b}}{\log(x)}$$

holds with some constant b.

Proof. First

$$\log\left(\prod_{p\le x}\left(1-\frac{1}{p}\right)\right)=\sum_{p\le x}\log\left(1-\frac{1}{p}\right)=-\sum_{p\le x}\sum_{k=1}^{\infty}\frac{1}{kp^k}$$

where

$$-\sum_{p\le x}\sum_{k=1}^{\infty}\frac{1}{kp^k}=-\sum_{p\le x}\frac{1}{p}-\sum_{p}\sum_{k=2}^{\infty}\frac{1}{kp^k}+\sum_{p>x}\sum_{k=2}^{\infty}\frac{1}{kp^k}.$$

The first term on the right-hand side may be estimated by means of Proposition 1.10, the second term is a convergent infinite series, and the third term tends to zero as $x\to+\infty$. Thus

$$\prod_{p\le x}\left(1-\frac{1}{p}\right)\sim\exp(-\log\log(x)-b)=\frac{e^{-b}}{\log(x)}$$

by exponentiating. □

About half of all the integers n with $y<n\le x$ for x and $x-y$ large are even, one third are divisible by three, and so forth. A suggestive way of phrasing this observation is to say that the chance of a randomly chosen large integer n being divisible by a prime p is $1/p$. An integer $n\ge 2$ that is not divisible by any prime $p\le\sqrt{n}$ is itself prime. So if for $\sqrt{x}\le y<n\le x$ the events $p|n$ and $q|n$ for distinct primes $p,q\le\sqrt{x}$ are independent in the sense of probability theory, the chance of n being prime should be

$$\prod_{p\le\sqrt{x}}\left(1-\frac{1}{p}\right)\sim\frac{e^{-\gamma}}{\log(\sqrt{x})}=\frac{2e^{-\gamma}}{\log(x)}>\frac{1.12}{\log(x)}.$$

But the density of the primes near x is close to $1/\log(x)$ by the Prime Number Theorem. We conclude that the events $p|n$ and $q|n$ are not independent. It is easy to persuade oneself that independence must hold for pairs of distinct primes that are very small compared with n. Thus Mertens' formula reveals an aspect of divisibility of integers by comparatively large primes.

1.5. Sums over divisors

We remind the reader that an arithmetic function f is a mapping $f : \mathbb{N} \to \mathbb{C}$. Some arithmetic functions such as log arise by restricting functions of a real variable to the positive integers. But in most cases of interest $f(n)$ is determined by arithmetical information about the integer n. The *divisor function* $d(n)$ given as the number of positive divisors of the positive integer n is an example. Each positive integer n has a unique factorization

$$n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$$

into primes and the divisors d of n are the integers of the form

$$d = p_1^{\beta_1} \cdots p_r^{\beta_r}$$

where the β_j are integers satisfying $0 \leq \beta_j \leq \alpha_j$ for $j = 1, 2, \ldots, r$. Hence $d(n) = (\alpha_1 + 1)(\alpha_2 + 1) \cdots (\alpha_r + 1)$ is a formula for the divisor function $d(n)$ given in terms of the prime factorization of n.

An arithmetic function f is *additive* if $f(mn) = f(m) + f(n)$ whenever $\gcd(m, n) = 1$. It is *multiplicative* if $f \not\equiv 0$ and $f(mn) = f(m)f(n)$ whenever $\gcd(m, n) = 1$. It is *totally additive* or *totally multiplicative* if the corresponding property holds without requiring the condition $\gcd(m, n) = 1$. A multiplicative or additive function can be unambiguously prescribed by giving its values on the prime powers, and a totally multiplicative or totally additive function by giving its values on the primes. Note also that $f(1) = 1$ if f is multiplicative.

The function $\log(n)$ and the function $\Omega(n)$ that counts the prime divisors of n with multiplicity are totally additive. The function $\omega(n)$ that counts the distinct prime divisors of n is additive, but not totally additive. The identity function id given by $n \mapsto n$ is totally multiplicative. So is the *Liouville function*

$$\lambda(n) = (-1)^{\Omega(n)}.$$

The divisor function $d(n)$ is multiplicative, but not totally multiplicative. The Euler phi-function $\phi(n)$ is also multiplicative. Another multiplicative arithmetic function is the *radical*

$$\operatorname{rad}(n) = \prod_{p|n} p.$$

It is also called the *squarefree kernel.*

The von Mangoldt function $\Lambda(n)$ is an important arithmetic function that is neither additive nor multiplicative.

The product of two multiplicative functions is multiplicative and the sum of two additive functions is additive. But a more important algebraic

operation on arithmetic functions is the *Dirichlet convolution.* If f and g are arithmetic functions then

$$(f * g)(n) = \sum_{d|n} f(d)g\Big(\frac{n}{d}\Big) = \sum_{km=n} f(k)g(m)$$

is their Dirichlet convolution. It is a straightforward exercise to see that under addition and Dirichlet convolution, the arithmetic functions form a commutative ring with multiplicative neutral element e where $\mathrm{e}(1) = 1$ and $\mathrm{e}(n) = 0$ for $n \geq 2$. This is called the *Dirichlet ring.* Denoting the constant function equal to 1 by 1 we note $d = 1 * 1$ as an example of Dirichlet convolution. Another convolution identity is $1 * \Lambda = \log$. This is easily proved by observing that

$$(1 * \Lambda)(n) = \sum_{d|n} \Lambda(d) = \sum_{p^k|n} \log(p) = \log(n),$$

since Λ is zero off the prime powers. This can replace the Legendre identity as the point of entry for the method of Chebyshev.

Proposition 1.12. *If f and g are multiplicative, so is $f * g$.*

Proof. If $\gcd(m, n) = 1$, then the divisors $d|mn$ are precisely those positive integers of the form $d = bc$ where $b|m$ and $c|n$. Hence

$$\begin{aligned}(f * g)(mn) &= \sum_{d|mn} f(d)g\Big(\frac{mn}{d}\Big) = \sum_{b|m,c|n} f(bc)g\Big(\frac{mn}{bc}\Big)\\ &= \sum_{b|m,c|n} f(b)f(c)g\Big(\frac{m}{b}\Big)g\Big(\frac{n}{c}\Big)\\ &= \sum_{b|m} f(b)g\Big(\frac{m}{b}\Big)\sum_{c|n} f(c)g\Big(\frac{n}{c}\Big)\\ &= (f * g)(m)(f * g)(n)\end{aligned}$$

by the multiplicativity of f and g. □

Since 1 is multiplicative, Proposition 1.12 shows that d is also multiplicative. The *sum-of-divisors function*

$$\sigma(n) = \sum_{d|n} d$$

is given by the Dirichlet convolution $\sigma = 1 * \mathrm{id}$, so σ is multiplicative, because id is.

Part of the significance of Proposition 1.12 is that for Dirichlet convolutions of multiplicative functions it affords a straightforward means of calculation; it is enough to calculate their values on prime powers. Let

us, for example, calculate the Dirichlet convolution $1 * \phi$. Both factors are multiplicative, so the calculation

$$(1 * \phi)(p^\alpha) = \sum_{p^\beta | p^\alpha} 1 \cdot \phi(p^\beta) = 1 + \sum_{\beta=1}^{\alpha} (p-1)p^{\beta-1} = p^\alpha$$

yields the convolution identity $1 * \phi = \mathrm{id}$ of Gauss.

The *Möbius mu-function* μ is the unique multiplicative arithmetic function with values $\mu(p) = -1$ on the primes p, and values $\mu(p^k) = 0$ on the prime powers p^k with $k \geq 2$. The Möbius function has a strong combinatorial flavor. It is closely connected to the principle of inclusion and exclusion, and to the fact that the integers form a partially ordered set under the relation of divisibility. The importance of the Möbius function is due to the convolution identity

$$\sum_{d|n} \mu(d) = \mathrm{e}(n).$$

There are many ways to establish that $1 * \mu = \mathrm{e}$, but the quickest is to recall that μ is multiplicative. Then so is $1 * \mu$, and thus $(1 * \mu)(p^\alpha) = 1 + (-1) + 0 + 0 + \cdots = 0$ yields $(1 * \mu)(n) = 0$ for all $n \geq 2$.

Proposition 1.13 (First Möbius inversion formula)**.** *If $g = 1 * f$ then $f = \mu * g$ and conversely.*

Proof. If $g = 1 * f$, then $\mu * g = \mu * (1 * f) = (\mu * 1) * f = \mathrm{e} * f = f$, and if $f = \mu * g$ then $1 * f = 1 * (\mu * g) = (1 * \mu) * g = \mathrm{e} * g = g$. □

This shows that 1 is a unit in the Dirichlet ring, and μ is its multiplicative inverse. An arithmetic function f is a unit if and only if $f(1) \neq 0$. Under this condition a Dirichlet inverse g for f may be constructed incrementally from

$$g(n) = \mathrm{e}(n) - \sum_{n \neq d|n} f\Big(\frac{n}{d}\Big) g(d).$$

The relation $f(1)g(1) = \mathrm{e}(1) = 1$ shows that the condition $f(1) \neq 0$ is necessary. Constructing an explicit Dirichlet inverse is usually infeasible, except in the very important case when f is multiplicative. The first Möbius inversion formula is often formulated as the statement

$$\text{“}f(n) = \sum_{d|n} \mu(d) g\Big(\frac{n}{d}\Big) \quad \text{if and only if} \quad g(n) = \sum_{d|n} f(d)\text{”}$$

about divisor sums.

Proposition 1.14 (Second Möbius inversion formula)**.** *Suppose that F is a function on the interval $[1,\infty)$. If*

$$G(x)=\sum_{n\le x}F\Big(\frac{x}{n}\Big)$$

then

$$F(x)=\sum_{n\le x}\mu(n)G\Big(\frac{x}{n}\Big)$$

and conversely on this interval.

Proof. First

$$\begin{aligned}\sum_{n\le x}\mu(n)G\Big(\frac{x}{n}\Big)&=\sum_{n\le x}\mu(n)\sum_{m\le x/n}F\bigg(\frac{x/n}{m}\bigg)=\sum_{mn\le x}\mu(n)F\Big(\frac{x}{mn}\Big)\\&=\sum_{N\le x}F\Big(\frac{x}{N}\Big)\sum_{n|N}\mu(n)=F(x)\end{aligned}$$

and then

$$\begin{aligned}\sum_{n\le x}F\Big(\frac{x}{n}\Big)&=\sum_{n\le x}\sum_{m\le x/n}\mu(m)G\bigg(\frac{x/n}{m}\bigg)=\sum_{mn\le x}\mu(m)G\Big(\frac{x}{mn}\Big)\\&=\sum_{N\le x}G\Big(\frac{x}{N}\Big)\sum_{m|N}\mu(m)=G(x)\end{aligned}$$

since $1*\mu=e$. □

The second Möbius inversion formula throws light on the method of Chebyshev. The relation

$$\psi(x)=\sum_{n\le x}\mu(n)T\Big(\frac{x}{n}\Big)$$

holds by Möbius inversion. Since $T(x)$ is quite precisely known, it might seem possible to estimate $\psi(x)$ fairly accurately by means of this formula. The problem here is that there is a great deal of cancellation in the sum, due to the oscillation of sign of $\mu(n)$. Too little is known about the behavior of $\mu(n)$ for this approach to promise much success. But the estimates of Chebyshev may be obtained by replacing $\mu(n)$ by an approximation of a particular kind. The approximation associated with the proof of Proposition 1.1 is $\mu(n)\approx \mathrm{e}_1(n)-2\mathrm{e}_2(n)$ where $\mathrm{e}_k(n)$ is the arithmetic function that equals

1 for $n = k$ and is zero otherwise. The calculation

$$\begin{aligned}\sum_{n\le x}(\mathrm{e}_1(n)-2\mathrm{e}_2(n))T\Big(\frac{x}{n}\Big) &= \sum_{n\le x}(\mathrm{e}_1(n)-2\mathrm{e}_2(n))\sum_{m\le x/n}\psi\Big(\frac{x/n}{m}\Big)\\ &= \sum_{k\le x}\psi\Big(\frac{x}{k}\Big)\sum_{d|k}(\mathrm{e}_1(d)-2\mathrm{e}_2(d))\\ &= \sum_{k\le x}(-1)^{k-1}\psi\Big(\frac{x}{k}\Big)\end{aligned}$$

may be taken as the framework of the proof of Proposition 1.1. It can be generalized by replacing $\mathrm{e}_1-2\mathrm{e}_2$ with a more complicated linear combination f of $\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_N$. Then $1 * f$ is required to take only the values $0, \pm 1$, and the nonzero values of this function should start with $(1*f)(1) = 1$ and alternate. Chebyshev chose the linear combination $f = \mathrm{e}_1 - \mathrm{e}_2 - \mathrm{e}_3 - \mathrm{e}_5 + \mathrm{e}_{30}$ to obtain his better estimates.

We exhibit an example of the use of Möbius inversion to establish an arithmetically significant estimate. We find a quite precise bound for the error term in an asymptotic estimate for the summatory function

$$\Phi(x) \stackrel{\text{def}}{=} \sum_{n\le x}\phi(n)$$

of the Euler totient. The convolution identity of Gauss yields

$$\begin{aligned}\sum_{n\le x}\Phi\Big(\frac{x}{n}\Big) &= \sum_{n\le x}\sum_{d\le x/n}\phi(d) = \sum_{nd\le x}\phi(d) = \sum_{m\le x}\sum_{d|m}\phi(d)\\ &= \sum_{m\le x}(1*\phi)(m) = \sum_{m\le x} m = \frac{1}{2}[x]([x]+1),\end{aligned}$$

and then

$$\begin{aligned}\Phi(x) &= \sum_{n\le x}\mu(n)\frac{1}{2}\Big[\frac{x}{n}\Big]\Big(\Big[\frac{x}{n}\Big]+1\Big) = \frac{1}{2}\sum_{n\le x}\mu(n)\Big(\frac{x}{n}+O(1)\Big)\Big(\frac{x}{n}+O(1)\Big)\\ &= \frac{x^2}{2}\sum_{n\le x}\frac{\mu(n)}{n^2} + O\bigg(x\sum_{n\le x}\frac{1}{n}\bigg) = \frac{x^2}{2}\sum_{n=1}^{\infty}\frac{\mu(n)}{n^2} + O\bigg(x^2\sum_{n>x}\frac{1}{n^2}\bigg)\\ &\quad + O(x\log(x)) = \frac{x^2}{2}\sum_{n=1}^{\infty}\frac{\mu(n)}{n^2} + O(x\log(x))\end{aligned}$$

by the second Möbius inversion formula. Absolutely convergent series may be multiplied together to yield absolutely convergent series, and so

$$\bigg(\sum_{m=1}^{\infty}\frac{1}{m^2}\bigg)\bigg(\sum_{n=1}^{\infty}\frac{\mu(n)}{n^2}\bigg) = \sum_{N=1}^{\infty}\frac{1}{N^2}\sum_{mn=N}1\cdot\mu(n) = \sum_{N=1}^{\infty}\frac{\mathrm{e}(N)}{N^2} = 1.$$

Then

$$\Phi(x) = \frac{3}{\pi^2}x^2 + O(x\log(x))$$

by the famous formula

$$\sum_{m=1}^{\infty} \frac{1}{m^2} = \frac{\pi^2}{6}$$

of Euler.

An Introduction to the Theory of Numbers by G. H. Hardy and E. M. Wright contains interesting material on arithmetic functions and applications of elementary techniques in number theory. Another good source for such material is *Introduction to the Theory of Numbers* by H. N. Shapiro.

1.6. The hyperbola method

Many arithmetic functions fluctuate rapidly and substantially, but we may still want precise information about their growth. There are various ways to approach such questions, differing not just in the methods used, but more fundamentally in the kind of statement at which one aims. A positive function g is a *maximal order* for an arithmetic function f if for any $\varepsilon > 0$ the inequality $|f(n)| \leq (1+\varepsilon)g(n)$ holds for all n sufficiently large, while the inequality $|f(n)| \leq (1-\varepsilon)g(n)$ fails for infinitely many n no matter the choice of $\varepsilon > 0$. For the concept to be useful, the maximal order should be some simple function that grows reasonably evenly. Otherwise we could just choose $g \equiv f$ and be done. Naturally there is also an analogous concept of minimal orders for arithmetic functions, though for functions that fluctuate greatly, minimal orders are often of little interest. An easy example showing the strengths and weaknesses of this approach is the function

$$\Omega(n) = \sum_{p^\alpha | n} 1$$

that counts the prime divisors of n with multiplicity. To see how large $\Omega(n)$ could be for given n, it is natural to look at integers that have many prime factors for their size. Because repeated prime factors are counted, this leads us to the powers of 2. Indeed the inequality

$$n = p_1^{\alpha_1} \cdots p_r^{\alpha_r} \geq 2^{\alpha_1 + \cdots + \alpha_r} = 2^{\Omega(n)}$$

is strict unless n is a power of 2, and it shows that $\Omega(n) \leq k$ for $2^k \leq n < 2^{k+1}$. So $\log(n)/\log(2)$ is a maximal order for $\Omega(n)$. The advantage here is that the statement is valid for all n individually; the disadvantage is that $\Omega(n)$ is actually very much smaller than the maximal order $\log(n)/\log(2)$ for most n; see Proposition 2.7. The analogous question for the divisor function $d(n)$ lies a little deeper. This function does not itself have a tractable maximal order, but its logarithm does.

Proposition 1.15. $\log(d(n))$ *has maximal order* $\log(2)\log(n)/\log\log(n)$.

Proof. Write the prime factorization of an arbitrary positive integer n in logarithmic form

$$\log(n) = \alpha_1 \log(p_1) + \cdots + \alpha_r \log(p_r).$$

The inequality $\alpha_k \log(2) \le \alpha_k \log(p_k) \le \log(n)$ is an immediate consequence, so $\alpha_k \le \log(n)/\log(2)$. Furthermore

$$\log(d(n)) = \log(\alpha_1 + 1) + \cdots + \log(\alpha_r + 1),$$

and the inequality $\log(\alpha_k + 1) \le \alpha_k \log(2)$ also holds. Apply the first inequality when $\log(p_k)$ is comparatively small, and the second inequality otherwise. Suppose $\log(p_k) \le c$ precisely when $1 \le k \le m$, where c is a parameter. Then

$$\begin{aligned}
\log(d(n)) &= \sum_{k=1}^{m} \log(\alpha_k + 1) + \sum_{k=m+1}^{r} \log(\alpha_k + 1) \\
&\le e^c \log\left(\frac{\log(n)}{\log(2)} + 1\right) + \sum_{k=m+1}^{r} \alpha_k \log(2) \\
&\le e^c \log\left(\frac{\log(n)}{\log(2)} + 1\right) + \frac{\log(2)}{c} \sum_{k=m+1}^{r} \alpha_k \log(p_k) \\
&\le e^c \log\left(\frac{\log(n)}{\log(2)} + 1\right) + \frac{\log(2)\log(n)}{c}
\end{aligned}$$

since $m \le \exp(c)$. Now choose $c = (1-\delta)\log\log(n)$ with $0 < \delta < 1$. Now

$$\log(d(n)) \le \log^{1-\delta}(n) \log\left(\frac{\log(n)}{\log(2)} + 1\right) + (1-\delta)^{-1} \frac{\log(2)\log(n)}{\log\log(n)}.$$

Since δ may be chosen arbitrarily close to 0, for every $\varepsilon > 0$ there is some $n(\varepsilon)$ so that $\log(d(n)) < (1+\varepsilon)\log(2)\log(n)/\log\log(n)$ for $n \ge n(\varepsilon)$.

Let p be any prime so large that $\vartheta(p) \ge p/e$ and let $n = 2{\cdot}3\cdots p$ be the product of all the primes up to and including p. Then

$$\begin{aligned}
\frac{\log(d(n))}{\frac{\log(2)\log(n)}{\log\log(n)}} &= \frac{\pi(p)\log(2)}{\frac{\log(2)\vartheta(p)}{\log(\vartheta(p))}} \ge \frac{\pi(p)\log(2)}{\frac{\log(2)\pi(p)\log(p)}{\log(\vartheta(p))}} \\
&= \frac{\log(\vartheta(p))}{\log(p)} \ge \frac{\log(p/e)}{\log(p)} = 1 - \frac{1}{\log(p)}.
\end{aligned}$$

But p can be taken arbitrarily large. □

The above result immediately yields the weaker bound $d(n) \ll_\varepsilon n^\varepsilon$, valid for any $\varepsilon > 0$. This bound is usually more convenient in applications.

Another approach to study the growth of the rapidly fluctuating arithmetic function $d(n)$ is to consider a *local average* such as

$$\frac{1}{h}\sum_{x-h<n\le x} d(n).$$

The fluctuations of $d(n)$ are smoothed out by the process of averaging. Then

$$\frac{1}{h}\sum_{x-h<n\le x} d(n) = \frac{D(x)-D(x-h)}{h}$$

where $D(x)$ is the summatory function of the divisor function. The use of local averaging to study the growth of arithmetic functions can be traced back to article 301 in the *Disquisitiones Arithmeticae.* Gauss was interested in the growth of the class number and the number of genera for binary quadratic forms, as (irregularly fluctuating) arithmetic functions of the discriminant. He quoted results on the rate of growth of their local averages, but judged the proofs to be too difficult to include in the *Disquisitiones.*

To calculate a local average by differencing an estimate for the associated summatory function is not always efficient. When h is small, one is apt to run into the same kind of problem as one does in numerics when subtracting floating-point numbers that are nearly equal.

The bijection $d \mapsto n/d$ on the set of divisors d of an integer n is called the *Dirichlet interchange.* Since $d < \sqrt{n}$ is equivalent to $n/d > \sqrt{n}$ it is clear that

$$d(n) = 2\sum_{\sqrt{n}>d|n} 1$$

unless n is a square. In the latter case the divisor $\sqrt{n}$ is missing and it is necessary to add 1 on the right-hand side. Applying this formula to the definition of $D(x)$ gives

$$\begin{aligned} D(x) &= [\sqrt{x}] + \sum_{n\le x} 2\sum_{\sqrt{n}>d|n} 1 = [\sqrt{x}] + 2\sum_{d\le\sqrt{x}}\sum_{d^2<kd\le x} 1 \\ &= [\sqrt{x}] + 2\sum_{d\le\sqrt{x}} \left(\left[\frac{x}{d}\right] - d\right) = 2\sum_{m\le\sqrt{x}} \left[\frac{x}{m}\right] - [\sqrt{x}]^2. \end{aligned}$$

The latter formula is due to D. F. E. Meissel.

Proposition 1.16. *The estimate*

$$D(x) = x\log(x) + (2\gamma-1)x + O(x^{1/2})$$

holds.

Proof. We have

$$\begin{aligned} D(x) &= 2\sum_{n\le\sqrt{x}}\left[\frac{x}{n}\right] - [\sqrt{x}]^2 \\ &= 2\sum_{n\le\sqrt{x}}\left(\frac{x}{n}+O(1)\right) - (\sqrt{x}+O(1))^2 \\ &= 2x(\log(\sqrt{x})+\gamma+O(1/\sqrt{x})) - x + O(x^{1/2}) \\ &= x\log(x) + (2\gamma-1)x + O(x^{1/2}) \end{aligned}$$

by Proposition 1.7. □

This estimate is due to J. P. G. Lejeune Dirichlet. It implies that the arithmetic average of $d(n)$ over the range $1 \le n \le x$ is asymptotic to $\log(x)$ as $x \to +\infty$. One says that $d(n)$ has *average order* $\log(x)$. From Proposition 1.15 it is easy to see that $d(n)$ is sometimes larger than any fixed power of $\log(n)$. But Proposition 1.16 implies that $d(n)$ is only rarely so large. The notation $\Delta(x) = D(x) - x\log(x) - (2\gamma-1)x$ is traditional for the error term in the estimate in Proposition 1.16. The problem of bounding $\Delta(x)$ is known as the *Dirichlet divisor problem.* More precisely, the divisor problem is to find the least ϑ for which an estimate $\Delta(x) = O(x^{\vartheta+\varepsilon})$ holds for all $\varepsilon > 0$. The result just proved shows that $\vartheta \le 1/2$. For x large and h rather smaller than x, say $h < x/2$, one obtains

$$\begin{aligned} \frac{1}{h}\sum_{x-h<n\le x} d(n) &= \frac{D(x)-D(x-h)}{h} \\ &= \frac{x\log(x)+(2\gamma-1)x+\Delta(x)}{h} \\ &\quad - \frac{(x-h)\log(x-h)+(2\gamma-1)(x-h)+\Delta(x-h)}{h} \\ &= \log(x) + 2\gamma + O\left(\frac{h}{x}\right) + O\left(\frac{x^{1/2}}{h}\right) \end{aligned}$$

by the estimate $\Delta(x) = O(x^{1/2})$. The error is a sum of two terms, one of which dominates when h is large and the other when h is small. In such situations one would usually try to choose the parameter optimally to obtain a small error term overall. Minimizing $h/x + x^{1/2}/h$ over h for x fixed, one sees that $h = x^{3/4}$ is an optimal choice. Thus

$$\frac{1}{h}\sum_{x-h<n\le x} d(n) = \log(x) + 2\gamma + O(x^{-1/4}), \qquad h = x^{3/4}.$$

The asymptotic law of growth $\log(x) + 2\gamma$ for the local average of $d(n)$ was Dirichlet's main application in his 1849 paper on the divisor problem. The

asymptotic estimate for the local average may be improved in two different ways: By shortening the interval or reducing the error term. Modifying the estimate as it stands by shortening the interval as much as possible, we lose the error term and obtain the asymptotic estimate

$$\frac{1}{h}\sum_{x-h<n\leq x} d(n) \sim \log(x) + 2\gamma, \qquad h = o(x^{1/2}/\log(x)).$$

Using a better estimate in the divisor problem, the estimate for the local average may be improved both by reducing the error term and shortening the interval.

One may also prove Proposition 1.16 from Dirichlet's formula

$$D(x) = \sum_{dk\leq x} 1 = \sum_{n\leq x} \left[\frac{x}{n}\right].$$

That $D(x) = x\log(x) + O(x)$ is immediate from this formula. The better estimate for the error term may be obtained by observing that the sum equals the number of integer lattice points in the region of the uv-plane given by the inequalities $u \geq 1$, $v \geq 1$ and $uv \leq x$. One can then recover the formula of Meissel by observing that the union of the two subregions obtained by imposing the inequalities $u \leq \sqrt{x}$ and $v \leq \sqrt{x}$ equals the original region, while their intersection equals the square given by $1 \leq u \leq \sqrt{x}$ and $1 \leq v \leq \sqrt{x}$. The interpretation of $D(x)$ in terms of the number of lattice points under a hyperbola is very important for more advanced work on the divisor problem. See Figure 2 on page 26 for an illustration.

The Dirichlet interchange and the approach to the Meissel formula based on counting lattice points under a hyperbola are closely connected. The technique is usually called the *Dirichlet hyperbola method.* It has other applications and so we exhibit a more general formulation due to H. G. Diamond.

Proposition 1.17 (Dirichlet hyperbola method)**.** *If f is an arithmetic function with summatory function F and g an arithmetic function with summatory function G then*

$$\sum_{n\leq x}(f*g)(n) = \sum_{k\leq y} f(k)G\Big(\frac{x}{k}\Big) + \sum_{m\leq x/y} g(m)F\Big(\frac{x}{m}\Big) - F(y)G\Big(\frac{x}{y}\Big)$$

for $1 \leq y \leq x$.

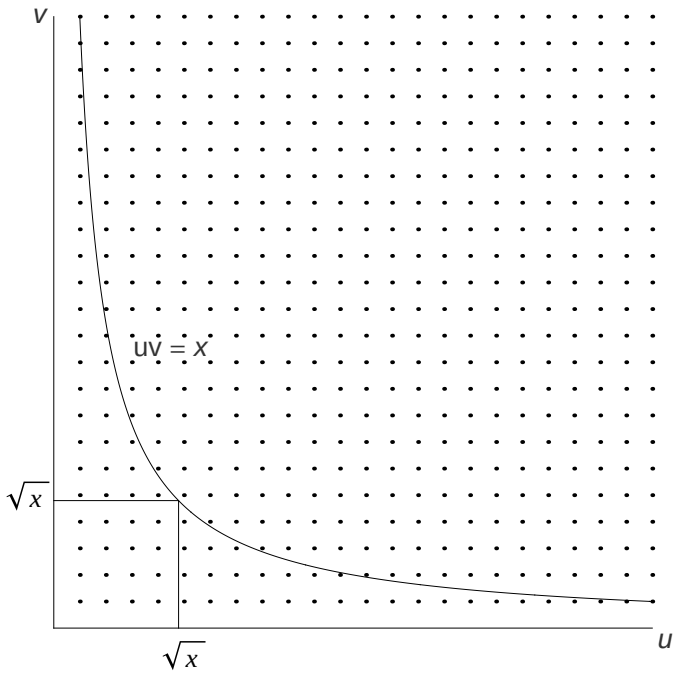

Figure 2. Lattice points in the divisor problem

Proof. We have

$$\begin{aligned}
\sum_{n\le x}(f*g)(n) &= \sum_{n\le x}\sum_{km=n} f(k)g(m)\\
&= \sum_{k\le y} f(k) \sum_{km\le x} g(m) + \sum_{k>y}\sum_{km\le x} f(k)g(m)\\
&= \sum_{k\le y} f(k)G\Big(\frac{x}{k}\Big) + \sum_{m\le x/y} g(m) \sum_{y<k\le x/m} f(k)\\
&= \sum_{k\le y} f(k)G\Big(\frac{x}{k}\Big) + \sum_{m\le x/y} g(m)\left(\sum_{k\le x/m} f(k) - \sum_{k\le y} f(k)\right)\\
&= \sum_{k\le y} f(k)G\Big(\frac{x}{k}\Big) + \sum_{m\le x/y} g(m)F\Big(\frac{x}{m}\Big) - F(y)G\Big(\frac{x}{y}\Big)
\end{aligned}$$

for any y with $1 \le y \le x$. □

Questions about divisors more delicate than their gross count have also been studied. The monograph *Divisors* by R. R. Hall and G. Tenenbaum is a good source for information of this kind.

Elementary Methods in the Analytic Theory of Numbers by A. O. Gelfond and Y. V. Linnik is a classic that covers some of the material of this chapter in greater depth, and other topics as well. *Elementary methods in number theory* by M. B. Nathanson and *Not Always Buried Deep* by P. Pollack also treat elementary techniques in analytic number theory. Material of this kind may also be found in *Introduction to the Theory of Numbers* by G. H. Hardy and E. M. Wright, *Lectures on Elementary Number Theory* by Hans Rademacher, and *Introduction to the Theory of Numbers* by H. N. Shapiro.

1.7. Notes

Gauss never published his empirical investigations on the distribution of the primes, but these are known from a letter that he wrote on Christmas Eve of 1849 to a former student of his, the astronomer J. F. F. Encke, and also from cryptic jottings in his research diary and on a flyleaf of a logarithm table. See pages 444–447 of volume II and pages 11–18 of volume X of his collected works [**Gau33**].

Legendre published the prime factorization of the factorial in the 1808 edition of his treatise *Essai sur la Théorie des Nombres* [**Leg08**]. There he also proposed

$$\pi(x) \approx \frac{x}{\log(x) - 1.08366}$$

as an excellent approximation. Today it is known that $\mathrm{li}(x)$ is a much better approximation for very large x, and that the asymptotically best approximation to $\pi(x)$ of the form $x/(\log(x)-A)$ is obtained for $A = 1$. But Legendre's approximation is better than Gauss' approximation in the interval between $x = 10^2$ and $x = 4{\cdot}10^6$, which stretches beyond the range of the tables of primes available in the early nineteenth century. Gauss makes a comment in his letter to Encke on the approximation of Legendre, to the effect that he does not care to commit himself as to what limit $A(x)$ in

$$\pi(x) = \frac{x}{\log(x) - A(x)},$$

may tend to as $x \to +\infty$.

Legendre made yet a third discovery of great importance to the development of prime number theory. Since antiquity an algorithm had been known for efficiently constructing tables of primes. The algorithm is called the Sieve of Eratosthenes, after the Hellenistic scholar Eratosthenes of Cyrene. We will explain how his algorithm may be used to construct a table of primes up to 30. Start with a list of the integers n with $2 \le n \le 30$. Keep the integer 2 but strike out all its proper multiples. Then keep the next integer 3, but strike out all its proper multiples. Next keep 5, but strike out all its proper multiples. The integers left in the list are the primes $2 \le p \le 30$. For every composite integer $n \le 30$ has some prime divisor $p \le \sqrt{30} < 5.5$. Note that we obtain the primes in the interval $[6, 30]$ by removing from the set of integers in that interval those that lie in the three arithmetic progressions $2\mathbb{Z}$, $3\mathbb{Z}$ and $5\mathbb{Z}$. Legendre reformulated the Sieve of Eratosthenes in terms of the principle of exclusion and inclusion from combinatorics, potentially

making it available to count primes analytically. A modern version of Legendre's sieve formula is

$$\pi(x) = \pi(\sqrt{x}) - 1 + \sum_{d|P} \mu(d) \left[\frac{x}{d}\right], \qquad P = \prod_{p \le \sqrt{x}} p.$$

Though the Legendre sieve formula gives $\pi(x)$ exactly, it is difficult to extract strong information about the distribution of the primes from it and many years passed before the idea of sieving had any impact on number theory beyond the preparation of tables and other computational work. The main objective of modern sieve theory is to find upper and lower bounds on the number of elements of a finite set of integers remaining after those elements that lie in some prescribed arithmetic progressions have been removed. This is such a flexible framework that a wide variety of arithmetical problems are amenable to sieve methods to some extent.

The young French mathematician J. Merlin began making progress on sieve theory around 1911, but he fell in the Great War. From 1915 V. Brun obtained results on the primes by means of sieving, that were not accessible by any other method. Since then, a vast amount of work has been done to improve sieve methods and develop new and more powerful ones. They have had a very broad impact on analytic number theory, leading to proofs of many important results, often in combination with ideas from outside sieve theory. Some of the most striking results obtained by means of sieves are:

- The series of reciprocals of primes p for which $p + 2$ is prime, is finite or convergent (Brun [**Bru15**] in 1915.)
- There are infinitely many primes p for which $p+2$ has at most two prime factors (J. R. Chen [**Che73**] in 1966.)
- For every sufficiently large even integer n there is some prime p and some integer m with at most two prime factors so that $n = p + m$ (Chen [**Che73**] in 1966.)
- There are infinitely many integers m for which $m^2 + 1$ has at most two prime factors (H. Iwaniec [**Iwa78**] in 1978.)
- The polynomial $m^2 + n^4$ takes infinitely many prime values (J. B. Friedlander and H. Iwaniec [**FI98**] in 1998.)
- The polynomial m^3+2n^3 takes infinitely many prime values (D. R. Heath-Brown [**HB01**] in 2001.)

The first two results are related to the twin prime problem, to prove that there are infinitely many pairs of primes that differ by 2. The third is related to the binary Goldbach problem, asserting that every even integer $n \ge 4$ is the sum of two primes, and the next two are related to the conjecture that the polynomial $m^2 + 1$ takes infinitely many prime values. All three problems are old, and all are unsolved. The first of the six results above is nowadays not very difficult to prove. Indeed M. Ram Murty and N. Saradha [**MS87**] discovered that even the sieve of Eratosthenes suffices, when combined with an elementary device due to R. A. Rankin. There is an exposition of their proof in *An Introduction to Sieve Methods and their Applications* by A. C. Cojocaru and M. Ram Murty [**CM06**]. The proofs of the other five results are much more difficult.

Chebyshev was the first mathematician to actually prove any results on the distribution of the primes of the kind envisioned by Gauss and Legendre. He published two papers on this topic, in 1848 and in 1850. In the first paper [**Che48**] Chebyshev shows that the limit

$$\lim_{s\to 1^+} \sum_{n=2}^{\infty} \left(\pi(n+1) - \pi(n) - \frac{1}{\log(n)}\right) \frac{\log^m(n)}{n^s}$$

exists and is finite for any choice of the integer m. This is a weak formulation of the statement that the density of the primes is $1/\log(x)$. To prove this result, he uses the Euler product formula that we shall consider in Chapter 3. He then deduces that for any $\varepsilon > 0$ and any integer m each of the inequalities $\pi(x) > \mathrm{li}(x) - \varepsilon x/\log^m(x)$ and $\pi(x) < \mathrm{li}(x) + \varepsilon x/\log^m(x)$ has arbitrarily large solutions x. Now it is immediately clear that *if* the ratio $\pi(x)/\mathrm{li}(x)$ tends to a limit, that limit must be 1, so that then the relative error in the approximation of Gauss tends to zero as $x \to \infty$. Chebyshev goes on to conclude that if we put $\pi(x) = x/(\log(x) - A(x))$ and $A(x)$ *has* a limit as $x \to +\infty$, then the limit must be 1. So if there is an asymptotically best approximation of the form $\pi(x) \approx x/(\log(x) - A)$ at all, that must be the approximation with $A = 1$ rather than with $A = 1.08366$ as proposed by Legendre. In his 1850 paper [**Che50**] Chebyshev proved an unconditional estimate that is equivalent to $0.921 \cdot \mathrm{li}(x) \le \pi(x) \le 1.106 \cdot \mathrm{li}(x)$ for all x sufficiently large, by a rather more elaborate version of the method that we used to prove Proposition 1.16. Even better upper and lower bounds were obtained by J. J. Sylvester [**Syl81**] using Chebyshev's method. Long after this work Diamond and Erdös [**DE80**] showed by means of the Prime Number Theorem that with sufficient calculation the method will yield estimates $c_1 \cdot \mathrm{li}(x) \le \pi(x) \le c_2 \cdot \mathrm{li}(x)$ with c_1 and c_2 arbitrarily close to 1.

The formula relating ψ and T was discovered independently of Chebyshev by A. de Polignac [**dP51**].

Chebyshev's proof of Bertrand's postulate in his 1850 paper [**Che50**] inaugurated the study of the local distribution of primes. Any nontrivial bound for the error term in the Prime Number Theorem implies existence of primes in short intervals. How short one can take the intervals by this approach depends on the quality of the bound on the error term in the PNT. However, in 1930 G. Hoheisel [**Hoh30**] by means of a different, analytic kind of argument succeeded in proving that the interval $(x - x^\theta, x]$ contains a prime for all x sufficiently large, with the exponent $\theta = 1 - 1/33000$. Even today, a bound for the error term in the Prime Number Theorem strong enough to allow this conclusion is not known. The exponent θ was gradually reduced over the years, by H. A. Heilbronn [**Hei33**] ($\theta = 0.996$ in 1933), N. G. Chudakov [**Chu36**] ($\theta = 3/4 + \varepsilon$ in 1936), A. E. Ingham [**Ing37**] ($\theta = 5/8 + \varepsilon$ in 1937), H. L. Montgomery [**Mon71**] ($\theta = 3/5 + \varepsilon$ in 1971), M. N. Huxley [**Hux72**] ($\theta = 7/12 + \varepsilon$ in 1972), H. Iwaniec and M. Jutila [**IJ79**] ($\theta = 13/23$ in 1979), Heath-Brown and Iwaniec [**HBI79**] ($\theta = 11/20$ in 1979), J. Pintz [**Pin84**] ($\theta = 17/31 - c$ for some small computable $c > 0$ in 1984), Iwaniec and Pintz [**IP84**] ($\theta = 23/42$ in 1984), C. J. Mozzochi [**Moz86**] ($\theta = 11/20 - 1/384$ in 1986), S. T. Lou and Q. Yao [**LY92, LY93**] ($\theta = 6/11$ in 1992 and $\theta = 6/11$ in 1993), R. C. Baker and G. Harman [**BH96**] ($\theta = .535$ in 1996), and Baker, Harman and Pintz [**BHP01**] ($\theta = 0.525$ in 2001.) There are heuristic arguments in favor of stronger conclusions. See page 422 of *Multiplicative Number Theory I. Classical Theory* by

H L. Montgomery and R. C. Vaughan [**MV07**], where the possibility that $\theta = \varepsilon$ with $\varepsilon > 0$ arbitrary is discussed. Also see the paper [**Gon93**] by S. M. Gonek. An even stronger conclusion follows if one accepts a probabilistic model of the distribution of the primes originated by H. Cramér [**Cra36**], and modified by A. Granville [**Gra95**] after work of H. Maier [**Mai85**]. This indicates that there should be some constant $C > 2e^{-\gamma}$ such that the interval $(x - C\log^2(x), x]$ contains a prime for all x sufficiently large, and that possibly any $C > 2e^{-\gamma}$ will do.

Defining $d_k = p_{k+1} - p_k$ one may, in view of the above considerations, ask how large d_k can become, say in terms of p_k. It is an immediate consequence of the Prime Number Theorem that $\limsup_{k\to+\infty} d_k/\log(p_k) \geq c$ with $c = 1$. R. J. Backlund [**Bac29**] showed in 1929 that one can take $c = 2$, and A. T. Brauer and H. Zeitz [**BZ30**] achieved $c = 4$ the year after. E. Westzynthius [**Wes31**] proved in 1933 that

$$\limsup_{k\to+\infty} \frac{d_k}{\frac{\log(p_k)\log_3(p_k)}{\log_4(p_k)}} \geq 2e^{\gamma},$$

where $\log_m(x)$ denotes the m times iterated logarithm. This was improved to

$$\limsup_{k\to+\infty} \frac{d_k}{\log(p_k)\log_3(p_k)} > 0$$

by G. Ricci [**Ric34**]. Further progress was made by Erdös [**Erd35**], who showed that

$$\limsup_{k\to+\infty} \frac{d_k}{\frac{\log(p_k)\log_2(p_k)}{(\log_3(p_k))^2}} > 0,$$

and by Rankin [**Ran38**], who showed in 1938 that

$$\limsup_{k\to+\infty} \frac{d_k}{\frac{\log(p_k)\log_2(p_k)\log_4(p_k)}{(\log_3(p_k))^2}} \geq c$$

with $c = 1/3$. Since then, only improvements in the constant c have been obtained, by A. Schönhage [**Sch63**], ($c = e^{\gamma_0/2}$ in 1963), Rankin [**Ran63**], ($c = e^{\gamma_0}$ in 1963), H. Maier and C. B. Pomerance [**MP90**], ($c = 1.312\ldots e^{\gamma_0}$ in 1990), and Pintz [**Pin97**], ($c = 2e^{\gamma_0}$ in 1997.)

The question of how small d_k can be in the long run is also of interest. If there are infinitely many twin primes, then $d_k = 2$ infinitely often. Defining $E = \liminf_{n\to\infty} d_k/\log(p_k)$, it is again an immediate consequence of the Prime Number Theorem that $E \leq 1$. Upper bounds for E were obtained by Erdös [**Erd40**] ($E < 1 - c$ for some small computable $c > 0$ in 1940), Rankin [**Ran50**] ($E \leq 42/43$ in 1950), Ricci [**Ric54a**] ($E \leq 15/16$ in 1954), E. Bombieri and H. Davenport [**BD66**] ($E \leq (2+\sqrt{3})/8$ in 1966), G. Z. Pil'tjaĭ [**Pil72**] ($E \leq (2\sqrt{2}-1)/4$ in 1972), Huxley [**Hux73, Hux77, Hux84**] ($E \leq 1/4 + \pi/16$ in 1973, $E \leq 0.4425\ldots$ in 1977 and $E \leq 0.4393\ldots$ in 1984), É. Fouvry and F. Grupp [**FG86**] ($E \leq 0.4342\ldots$ in 1986), H. Maier [**Mai88**] ($E \leq 0.2484\ldots$ in 1988) and finally D. A. Goldston, J. Pintz and C. Y. Yıldırım [**GPY09**] ($E = 0$ in 2005.) Recently Y. Zhang [**Zha14**] proved that $d_k \leq 70{\cdot}10^6$ infinitely often, which was a great advance. In work to appear in Annals of Mathematics, J. Maynard has shown that $p_{k+m} - p_k \leq c_m$ infinitely often for each positive integer m, with $c_1 = 600$ admissible.

The proof of Proposition 1.2 is modeled on the one given by Ramanujan [**Ram19**].

Landau gave sufficient conditions for the heuristic mentioned at the beginning of Section 1.4 to hold true; see page 201 of his treatise [**Lan74**] on prime number theory, or the paper [**Lan00**]. Partial summation goes back to a paper of N. H. Abel [**Abe26**] on power series. The Euler-Maclaurin summation formula [**Eul38, Mac42**] dates to the first half of the eighteenth century. The work of Mertens is in [**Mer74a, Mer74b**].

The formula for $d(n)$ dates to 1673 and is due to J. Kersey [**Ker73**]. Obscure today, to his contemporaries he was known as an author of well-regarded textbooks.

It seems that R. Descartes [**Des79**] was the first to explicitly note an arithmetic function multiplicative; in a posthumous manuscript he stated that the sum-of-divisors function $\sigma(n)$ has this property. From a letter [**Des98**] to M. Mersenne it seems likely that he knew this by 1638.

The von Mangoldt function, though not the notation for it used today, and the convolution identity $1 * \Lambda = \log$, are due to N. W. Bugaev [**Bug73**] and E. Césaro [**Cés88**].

The convolution identity $1 * \phi = \mathrm{id}$ was proved by Gauss in article 39 of the *Disquisitiones,* and this may well be the first instance of a divisor sum of an arithmetic function.

E. T. Bell [**Bel15**] and M. Cipolla [**Cip15**] independently in 1915 considered the set of arithmetic functions as an algebraic structure and gave Proposition 1.12. But a good many particular convolutions of multiplicative functions were known in the nineteenth century, so this result may well have been appreciated earlier.

A. F. Möbius [**Mőb31**] introduced the function named after him in 1831. But already Euler had considered infinite series whose terms involved values of the Möbius function.

Proposition 1.13 is due to J. W. R. Dedekind [**Ded57**] and J. Liouville [**Lio57**] independently in 1857, and Proposition 1.14 to Möbius [**Mőb31**].

The asymptotic estimate for $\Phi(x)$ is due to Mertens [**Mer74b**]. Dirichlet had obtained a similar estimate with x^{ε} in place of $\log(x)$ in [**Dir49**].

The maximal order of $\log(d(n))$ was found by S. Wigert [**Wig07**] around 1906 using the Prime Number Theorem. The dependence on the PNT was removed some years later by Ramanujan [**Ram15**]. The proof of Proposition 1.15 is a modified version of the one given by Wigert.

In two papers [**Dir38a, Dir38b**] of 1838 Dirichlet considers the question of how one can study arithmetic functions that fluctuate irregularly. It is not unreasonable to consider these papers as founding the theory of arithmetic functions, but Dirichlet himself refers to earlier work, and in particular to remarks of Gauss in article 301 of the *Disquisitiones.* What Dirichlet set out to do in the first of these papers was to find 'das asymptotische Gesetz' in the sense of Gauss, of the divisor function $d(n)$. He indicated an argument, based on the Lambert series expansion

$$\sum_{n=1}^{\infty} d(n)x^n = \sum_{m=1}^{\infty} \frac{x^m}{1-x^m},$$

that $\log(x) + 2\gamma$ is the asymptotic law of growth of $d(n)$ in the sense that the local average approaches $\log(x)+2\gamma$ as $x \to +\infty$ and $y \to +\infty$ in a suitable way. Dirichlet [**Dir49**] returned to the divisor problem in 1849 and gave his classical bound $O(\sqrt{x})$ for the error term. Using Dirichlet's result L. Kronecker in his lectures [**Kro01**] optimized the length of the interval to minimize the error bound for the local average of $d(n)$.

The Dirichlet divisor problem has a rich history, with a couple of obscure turns in the early stages. The first of these is a letter from Dirichlet to Kronecker dated July 23, 1858. Dirichlet had visited Kronecker for a few days in Ilsenburg, a resort by the Harz mountains, where the latter was spending his summer vacation. From the letter it is clear that they had discussed the divisor problem, and Dirichlet writes that he has now managed to substantially improve on his result of 1849 (*'.... die Summe ... bedeutend in die Enge zu treiben.'*) Dirichlet died on May 5, 1859 and his improvement never appeared. Apparently his Nachlass did not contain any material bearing on the problem, and nothing is known of the nature of the new method of which Dirichlet hints in the letter to Kronecker, nor of the extent of the improvement.

In 1903 G. F. Voronoi [**Vor03**] showed that $\vartheta \leq 1/3$ in the Dirichlet divisor problem. The proof is nearly forty pages long, and is based on the interpretation of $D(x)$ in terms of lattice points under a hyperbola. By means of Farey fractions Voronoi closely approximated the hyperbola by a polygon, and then he used the Euler summation formula to obtain the estimate $\Delta(x) = O(x^{1/3}\log(x))$. The following year he gave a different and much more analytic proof [**Vor04**].

As late as 1917 Voronoi's result was judged *'one of the deepest in the analytic theory of numbers'* by Hardy and Ramanujan [**HR17a**] in their paper on the normal number of prime factors of an integer. But in the same year I. M. Vinogradov [**Vin18a**] found a much easier proof. In 1922 J. G. van der Corput [**Cor22**] proved $\vartheta \leq 33/100$. His proof required estimates for exponential sums, and since that time further progress has depended on better estimates for such sums and on related techniques of counting lattice points.

The exponent in the Dirichlet divisor problem was slowly reduced over a long period, by van der Corput ($\vartheta = 27/82$ in 1928), T.-T. Chih [**Chi50**] ($\vartheta = 15/46$ in 1950), G. A. Kolesnik [**Kol69, Kol74, Kol82, Kol85**] ($\vartheta = 12/37$ in 1969, $\vartheta = 346/1067$ in 1973, $\vartheta = 35/108$ in 1982 and $\vartheta = 139/429$ in 1985), Iwaniec and Mozzochi [**IM88**] ($\vartheta = 7/22$ in 1988), and Huxley [**Hux93, Hux02, Hux03**] ($\vartheta = 23/73$ in 1993 and $\vartheta = 131/416$ in 2000.) In the other direction, Hardy [**Har15b, Har15a**] proved in 1914 that $\vartheta \geq 1/4$, and it is generally believed that $\vartheta = 1/4$ holds.

The formula of Meissel is in [**Mei54**]. Diamond's version of the hyperbola method is in his survey paper [**Dia82**] on elementary methods in prime number theory. The special case $y = x^{1/2}$ was noted by J. Franel [**Fra99**] in 1899.

Exercises

(1) a) Let $C(M,N)$ denote the binomial coefficient M choose N. Find the prime factorization of $C(2N,N)$.
b) Find a good upper bound for $C(2N,N)$.
c) Use a) and b) to show that $\vartheta(x) = O(x)$.

(2) Calculate $\operatorname{lcm}[1,2,\ldots,N]$ in terms of $\psi(N)$.

(3) You are offered to make bets in favor of integers being squares. Integers n are drawn at random from the interval $x - x^{3/4} < n \le x$ for some fixed very large x. Each time n is a square, you win $1.5\,x^{1/2}$ dollars. Each time it is not, you lose one dollar. Should you accept these bets?

(4) Prove the infinitude of primes by first proving the infinitude of squarefree numbers (J. Perrott.)

(5) Use the formula

$$\sum_{n\le x} \psi\Big(\frac{x}{n}\Big) = T(x)$$

to prove that $\mathbf{a} \le 1 \le \mathbf{A}$ and thus show that if $\psi(x)/x$ has a limit as $x \to \infty$, that limit must equal 1 (Chebyshev.)

(6) Let $A \subseteq \mathbb{Z}$ and denote by $N_A(x)$ the number of elements $a \in A$ with $|a| \le x$. Show that if $N_A(x)$ grows faster than any power of $\log(x)$, the total set of prime divisors of the elements of A is infinite. Show that if c is any positive constant, the total set of prime divisors of the sequence $[\exp(\log^c(n))]$ for $n = 1,2,3,\ldots$ is infinite. (Hint: The Fundamental Theorem of Arithmetic in logarithmic form.)

(7) † Let $N(a,b,c)$ denote the number of solutions of the inequality $ak + bm \le c$ in nonnegative integers k and m. Express $N(a,b,c)$ in terms of sums involving the integer part function and establish the estimate

$$\left| N(a,b,c) - \frac{1}{2}\Big(\frac{c}{a}+1\Big)\Big(\frac{c}{b}+1\Big)\right| \le \frac{1}{2} + \frac{3c}{4a} + \frac{3c}{4b}$$

when a, b and c are positive real numbers. About how many positive integers less than or equal to x are there of the form $2^k\cdot 3^m$ when x is large?

(8) Show that $n!$ is never a square for $n \ge 2$.

(9) Let n be an arbitrary positive integer. Show that counted with multiplicity every prime occurs at least as often in total as a factor of the integers in the interval $[n+1, 2n]$ as it does of the integers in the interval $[1,n]$.

(10) Counting prime factors without multiplicity, show that for every integer $n \geq 2$ there is some integer n' with $n < n' < 2n$ so that n and n' have the same number of prime factors.

(11) Use divisibility properties of the factorial to show that the sequence of primes has unbounded gaps. Then use a Chebyshev-type estimate to find a constant $c > 0$ so that there exist arbitrarily large pairs p_k, p_{k+1} of consecutive primes with $p_{k+1} - p_k \geq c \log(p_k)$.

(12) † a) Compute the normalized differences $(p_{k+1} - p_k)/\log(p_k)$ of pairs p_k, p_{k+1} of successive primes in intervals

$$(1000, 2000],\ (10000, 11000],\ (100000, 101000],\ (1000000, 1001000].$$

Sort and plot the normalized differences on each interval. Comment on: (i) the family resemblance of the various plots disregarding scale, (ii) the shape of the plots with special attention to the prevalence of small and large differences, and (iii) lack of "smoothness" of the plots.
b) Compute ratios $(p_{k+1} - p_k)/\log^2(p_k)$ to investigate an old conjecture of C. H. Cramér to the effect that this ratio is bounded. Since extreme values of the difference $p_{k+1} - p_k$ are sought, the computational strategy should be different from part a). Determine approximately the largest prime for which the time involved in computing the ratio is 1/10 of a second, and calculate a thousand ratios for randomly chosen and well scattered primes of about the same order of magnitude. Sort and plot the ratios as in a). (Here a computer algebra system is needed, or else some programming. In the latter case, *Prime Numbers. A computational perspective* by Richard Crandall and Carl Pomerance, or *Prime numbers and computer methods for factorization* by Hans Riesel, may prove helpful.)

(13) Show that

$$\prod_{p \leq n} p \leq 5^n$$

for all $n \geq 1$. According to P. Erdös and L. Kalmár one can take 4 in the inequality, and according to D. Hanson one can take 3.

(14) a) Let $A_x = (x/2, x] \cup (x/4, x/3] \cup (x/6, x/5] \cup \cdots$. Show that

$$\sum_{n \in A_x} \Lambda(n) = \log(2)x + O(\log(x)),$$

and that the error term cannot be improved.
b) Deduce that the interval $(x/2, x]$ contains at least $x/(5\log(x))$ primes for all x sufficiently large.

(15) Show that

$$\int_{x_0}^{x} f(u)\,du = O\left(\int_{x_0}^{x} g(u)\,du\right)$$

if $f(x) = O(g(x))$ and f and g are integrable on bounded intervals.

(16) Prove that if f and g are real-valued continuous functions on a closed and bounded interval $[a, b]$ and g is positive there, then

$$\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx$$

for some $c \in (a, b)$. This is called the *first mean value theorem for integrals.*

(17) a) Prove that if f and g are real-valued continuous functions on a closed and bounded interval $[a, b]$ and f is monotone there, then

$$\int_a^b f(x)g(x)\,dx = f(a)\int_a^c g(x)\,dx + f(b)\int_c^b g(x)\,dx$$

for some $c \in (a, b)$. This is called the *second mean value theorem for integrals.* The second mean value theorem is useful for the estimation of integrals with an oscillatory factor in the integrand.

b) Show that

$$\left|\int_0^{2\pi} f(x)\cos(x)\,dx\right| \leq |f(0) - f(2\pi)|$$

if f is a continuous monotone function.

(18) Show that

$$\lim_{h\to 0}\int_a^b \frac{f(x+h) - f(x)}{h}\,dx = f(b) - f(a)$$

if f is a continuous function on an open interval containing the closed interval $[a, b]$. Establish an integration by parts formula for two continuous functions, one of which is monotone.

(19) Make a guess for the asymptotic behavior of the sum

$$\sum_{p\leq x}\frac{1}{\sqrt{p}}$$

and use partial summation and estimates of Chebyshev to show that your guess has the right order of growth.

(20) Show that

$$\sum_{pq\leq x}\log(p)\log(q) \asymp x\log(x)$$

where p and q denote primes.

(21) a) Use integration by parts to show that

$$\int_n^{n+1} f(u)\,du = \frac{f(n)}{2} + \frac{f(n+1)}{2} - \int_n^{n+1} \left(u - n - \frac{1}{2}\right) f'(u)\,du.$$

Here n is an integer and f a continuous function on the interval $[n, n+1]$ with f' piecewise continuous there.
b) Use part a) to establish the Euler-Maclaurin summation formula.

(22) Prove the estimate

$$\sum_{n=-\infty}^{\infty} e^{-\pi n^2 u} = \frac{1}{\sqrt{u}} + O(1)$$

as $u \to 0^+$. This series comes from the theory of elliptic theta functions, and satisfies a functional equation that yields a far more precise estimate.

(23) † Establish the estimate

$$\sum_{n \le x} \log^m \left(\frac{x}{n}\right) = m!x + O(m! \log^m(x))$$

uniformly in positive integers m.

(24) Show without integration that $T(n) = n\log(n) + O(n)$ by subdividing the interval $[1, n]$ between successive powers of 2.

(25) Show that

$$\prod_{p \le x} \left(1 + \frac{1}{p}\right) \sim c\log(x)$$

as $x \to +\infty$, for some positive constant c.

(26) Show that

$$\sum_{p \le x} \frac{\log^2(p)}{p} = \frac{1}{2}\log^2(x) + O(\log(x))$$

as $x \to +\infty$.

(27) Show that the real number 1 is a point of accumulation of the sequence of ratios $(p_{k+1}/p_k)_{k=1}^{\infty}$ of successive primes.

(28) Show that the series

$$\sum_p \frac{1}{p(\log\log(p))^2}$$

converges.

(29) A divisor d of an integer n is called a *block divisor* if it is coprime to its complementary divisor n/d. In group theory such divisors are called *Hall divisors.* Count the number of block divisors of n.

(30) Show that the multiplicative arithmetic functions under Dirichlet convolution constitute a subgroup of the group of all arithmetic functions that have a Dirichlet inverse.

(31) Calculate $1 * \lambda$, $1 * \varrho$ and $d * \lambda$ where ϱ is the indicator function of the squares.

(32) Calculate $\mathrm{id} * \mathrm{id}$, $\phi * \sigma$ and $d * \phi$.

(33) Calculate $f * f$ for f totally multiplicative.

(34) Use the Binomial Theorem to show that $1 * \mu = \mathrm{e}$.

(35) Show that the sum of the primitive n-th roots of unity equals $\mu(n)$.

(36) Deduce the formula for T in terms of ψ by means of $1 * \Lambda = \log$.

(37) Show that $1 * (\mu \cdot \log) = -\Lambda$.

(38) Show that

$$\sum_{n \le x} \log(\mathrm{rad}(n)) = x \log(x) + O(x)$$

where rad is the radical.

(39) a) Show that the number of rationals in the interval $[0, 1]$ written in lowest terms and with denominator $\le x$ is asymptotic to $3x^2/\pi^2$.
b) Show that the chance that two randomly chosen large integers be coprime is $6/\pi^2$.
c) A lattice point (j, k) *occults* the lattice point (m, n) if both lie on the same ray from the origin and (j, k) lies closer to the origin. Show that the proportion of lattice points in $\mathbb{Z} \times \mathbb{Z}$ visible from the origin is $6/\pi^2$.

(40) Show that

$$\sum_{d \mid n} 2^{\omega(d)} = d(n^2),$$

and generalize (Dirichlet).

(41) Establish the estimate

$$\left| \sum_{n \le x} d(\gcd(m, n)) - \frac{\sigma(m)}{m} x \right| \le d(m)$$

for any fixed positive integer m.

(42) Show that

$$\sum_{n \le x} \frac{\sigma(n)}{n} = \frac{\pi^2}{6} x + O(\log(x))$$

as $x \to +\infty$.

(43) Let $P^+(n)$ denote the largest prime factor of n. Show that the infinite series

$$\sum_{P^+(n) \le x} \frac{\mu(n)}{n}$$

converges absolutely for every $x \ge 1$ and that the sum is always positive. Find the limit of the sum as $x \to +\infty$.

(44) An arithmetic function h is given, which is never zero. It is known that $h = fg$ where f is multiplicative and g is additive. Determine f and g from h.

(45) a) The sieve of Eratosthenes-Legendre: Show that

$$\sum_{d|P} \mu(d)\Big[\frac{x}{d}\Big] = \pi(x) - \pi(\sqrt{x}) + 1$$

if $P_{\sqrt{x}}$ denotes the product of the primes $p \le \sqrt{x}$.
b) Show that

$$\pi(x) \le \pi(z) - 1 + \sum_{d|P_z} \mu(d)\Big[\frac{x}{d}\Big]$$

if P_z denotes the product of the primes $p \le z$.
c) Show that $\pi(x) \ll x/\log\log(x)$ by a choice of z in terms of x in part b). This is weaker than what we already have from Section 1.1, but see the next part.
d) In the sieve of Eratosthenes-Legendre, we calculated the number of primes in the interval $\sqrt{x} < n \le x$ by removing those integers that lie in arithmetic progressions $p\mathbb{Z}$ with $p \le \sqrt{x}$. Apply the same idea on the interval $x - h < n \le x$ to show that a bound $\pi(x) - \pi(x-h) \ll h/\log\log(h)$ holds uniformly in x. Does this follow from the Chebyshev theory of Section 1.1?

(46) Find a maximal order and a minimal order of $\log(\phi(n))$.

(47) Use the Dirichlet interchange to show that $1 * \lambda = \varrho$, where ϱ is the characteristic function of the squares.

(48) Use the Dirichlet interchange to calculate $1 * \log$.

(49) A natural number n is called *perfect* if $\sigma(n) = 2n$. Use the Dirichlet interchange and congruences modulo 3 to show that $n \not\equiv 2 \pmod 3$ if n is perfect (J. Touchard).

(50) Find the arithmetic average of the fractional part $\{x/n\} = x/n - [x/n]$ of x/n over the interval $1 \le n \le x$ as $x \to +\infty$ (Dirichlet).

(51) Show that for every $\varepsilon > 0$ the equation $xu - yv = 1$ has $O_\varepsilon(R^{2+\varepsilon})$ solutions in integers in the ball $x^2 + y^2 + u^2 + v^2 \le R^2$.

(52) a) Show that

$$\sum_{n\le x} f(n)G\Big(\frac{x}{n}\Big) = \sum_{m\le x} (G(m) - G(m-1))F\Big(\frac{x}{m}\Big)$$

if $f(n)$ is an arithmetic function with summatory function $F(x)$ and $G(x)$ is the summatory function of some arithmetic function (Dirichlet).

b) Find a constant $c(\theta)$ such that

$$\sum_{n\le x}[x/n]^{-\theta} = c(\theta)x + O(1)$$

holds uniformly for $\theta \ge 0$. Uniformity means that the constants C and x_0 that are implicit in the O-term do not depend on the parameter θ.

(53) Show that

$$\sum_{n\le x}\log(\phi(n)) = x\log(x) + (c-1)x + O(\log(x))$$

where

$$c = \sum_p \frac{1}{p}\log\Big(1-\frac{1}{p}\Big).$$

Find the average order of $\log(\phi(n))$. Calculate the local average and choose the length h of the interval in terms of x so as to minimize the error term.

(54) a) Use the identity

$$\sum_{n\le x}\Big(\Big[\frac{x}{n}\Big] - \psi\Big(\frac{x}{n}\Big) - 2\gamma\Big) = D(x) - T(x) - 2\gamma[x]$$

and the Dirichlet bound in the divisor problem to show that $\psi(x) = O(x)$ independently of the Chebyshev method. This idea is due to Nina Spears. Her approach may be refined to yield close upper and lower numerical bounds for $\psi(x)/x$. Though unfortunately only at the cost of extensive computation.

Note that this establishes the bound $\vartheta(x) = O(x)$ and Proposition 1.9 without reliance on the method of Chebyshev.

b) Show that

$$\sum_{x/K<p\le x}\frac{\log(p)}{p} = \log(K) + O(1)$$

where $K \ge 1$ is an arbitrary constant, and the error is uniform in K.

c) Apply b) to show that $\vartheta(x) \asymp x$ without reliance on the method of Chebyshev (H. N. Shapiro).

(55) † Find an asymptotic estimate for the sum

$$\sum_{n\le x}\frac{d(n)}{\log(n)+2\gamma},$$

with a bound for the error term.

(56) † Let $c \geq 1$ be a fixed real number, and let $f_c(n)$ be the number of divisors d of n satisfying the inequality $1/c \leq n/d^2 \leq c$. Show that

$$\sum_{n \leq x} f_c(n) = x \log(c) + O(\sqrt{cx})$$

where the constants implied in the O-term do not depend on c. (The estimate is *uniform* in c.) Then conclude that 50% of the divisors d of the integers n in the interval $1 \leq n \leq x$ satisfy the inequality

$$\frac{1}{\sqrt{x}} \leq \frac{n}{d^2} \leq \sqrt{x}$$

as $x \to +\infty$. An estimate that contains a parameter frequently becomes much more useful if we can establish that the estimate is uniform in the parameter.