Multiplicative number theory: The pretentious approach

Andrew Granville

K. Soundararajan

To Marci and Waheeda

Preface

Riemann's seminal 1860 memoir showed how questions on the distribution of prime numbers are more-or-less equivalent to questions on the distribution of zeros of the Riemann zeta function. This was the starting point for the beautiful theory which is at the heart of analytic number theory. Until now there has been no other coherent approach that was capable of addressing all of the central issues of analytic number theory.

In this book we present the pretentious view of analytic number theory; allowing us to recover the basic results of prime number theory *without* use of zeros of the Riemann zeta-function and related L-functions, and to improve various results in the literature. This approach is certainly more flexible than the classical approach since it allows one to work on many questions for which L-function methods are not suited. However there is no beautiful explicit formula that promises to obtain the strongest believable results (which is the sort of thing one obtains from the Riemann zeta-function). So why *pretentious*?

• It is an intellectual challenge to see how much of the classical theory one can reprove without recourse to the more subtle L-function methodology (For a long time, top experts had believed that it is impossible is prove the prime number theorem without an analysis of zeros of analytic continuations. Selberg and Erdős refuted this prejudice but until now, such methods had seemed *ad hoc*, rather than part of a coherent theory).

• Selberg showed how sieve bounds can be obtained by optimizing values over a wide class of combinatorial objects, making them a very flexible tool. Pretentious methods allow us to introduce analogous flexibility into many problems where the issue is not the properties of a very specific function, but rather of a broad class of functions.

• This flexibility allows us to go further in many problems than classical methods alone, as we shall see in the latter chapters of this book.

The Riemann zeta-function $\zeta(s)$ is defined when $\mathrm{Re}(s) > 1$; and then it is given a value for each $s \in \mathbb{C}$ by the theory of analytic continuation. Riemann pointed to the study of the zeros of $\zeta(s)$ on the line where $\mathrm{Re}(s) = 1/2$. However we have few methods that truly allow us to say much so far away from the original domain of definition. Indeed almost all of the unconditional results in the literature are about understanding zeros with $\mathrm{Re}(s)$ very close to 1. Usually the methods used to do so, can be viewed as an extrapolation of our strong understanding of $\zeta(s)$ when $\mathrm{Re}(s) > 1$. This suggests that, in proving these results, one can perhaps dispense with an analysis of the values of $\zeta(s)$ with $\mathrm{Re}(s) \leq 1$, which is, in effect, what we do.

Our original goal in the first part of this book was to recover all the main results of Davenport"s *Multiplicative Number Theory* MR1790423 [?] by pretentious methods, and then to prove as much as possible of the result of classical literature, such as the results in MR891718 [?]. It turns out that pretentious methods yield a much easier proof of Linnik's Theorem, and quantitatively yield much the same quality of results throughout the subject.

However Siegel's Theorem, giving a lower bound on $|L(1,\chi)|$, is one result that we have little hope of addressing without considering zeros of L-functions. The difficulty is that all proofs of his lower bound run as follows: Either the Generalized

Riemann Hypothesis (GRH) is true, in which case we have a good lower bound, or the GRH is false, in which case we have a lower bound in terms of the first counterexample to GRH. Classically this explains the inexplicit constants in analytic number theory (evidently Siegel's lower bound cannot be made explicit unless another proof is found, or GRH is resolved) and, without a fundamentally different proof, we have little hope of avoiding zeros. Instead we give a proof, due to Pintz, that is formulated in terms of multiplicative functions and a putative zero.

Although this is the first coherent account of this theory, our work rests on ideas that have been around for some time, and the contributions of many authors. The central role in our development belongs to Halász's Theorem. Much is based on the results and perspectives of Paul Erdős and Atle Selberg. Other early authors include Wirsing, Halász, Daboussi and Delange. More recent influential authors include Elliott, Hall, Hildebrand, Iwaniec, Montgomery and Vaughan, Pintz, and Tenenbaum. In addition, Tenenbaum's book MR1366197 [?] gives beautiful insight into multiplicative functions, often from a classical perspective.

Our own thinking has developed in part thanks to conversations with our collaborators John Friedlander, Régis de la Bréteche, and Antal Balog. We are particularly grateful to Dimitris Koukoulopoulos and Adam Harper who have been working with us while we have worked on this book, and proved several results that we needed, when we needed them! Various people have contributed to our development of this book by asking the right questions or making useful mathematical remarks – in this vein we would like to thank Jordan Ellenberg, Hugh Montgomery.

The exercises: In order to really learn the subject the keen student should try to fully answer the exercises. We have marked several with † if they are difficult, and occasionally †† if extremely difficult. The † questions are probably too difficult except for well-prepared students. Some exercises are embedded in the text and need to be completed to fully understand the text; there are many other exercises at the end of each chapter. At a minimum the reader might attempt the exercises embedded in the text as well as those at the end of each chapter with are marked with *.

Contents

Part 1

Introductory results

In the the first four chapters we introduce well-known results of analytic number theory, from a perspective that will be useful in the remainder of the book.

CHAPTER 1.1

The prime number theorem

As a boy Gauss determined, from studying the primes up to three million, that the density of primes around x is $1/\log x$, leading him to conjecture that the number of primes up to x is well-approximated by the estimate

PNT
$$\pi(x) := \sum_{p \le x} 1 \sim \frac{x}{\log x}. \tag{1.1.1}$$

It is less intuitive, but simpler, to weight each prime with $\log p$; and to include the prime powers in the sum (which has little impact on the size). Thus we define the *von Mangoldt function*

vM
$$\Lambda(n) := \begin{cases} \log p & \text{if } n = p^m, \text{ where } p \text{ is prime, and } m \ge 1 \\ 0 & \text{otherwise,} \end{cases} \tag{1.1.2}$$

and then, in place of (1.1.1), we conjecture that

PNT2
$$\psi(x) := \sum_{n \le x} \Lambda(n) \sim x. \tag{1.1.3}$$

The equivalent estimates (1.1.1) and (1.1.3), known as the *prime number theorem*, are difficult to prove. In this chapter we show how the prime number theorem is equivalent to understanding the mean value of the Möbius function. This will motivate our study of multiplicative functions in general, and provide new ways of looking at many of the classical questions in analytic number theory.

1.1.1. Partial Summation

Given a sequence of complex numbers a_n, and some function $f : \mathbb{R} \to \mathbb{C}$, we wish to determine the value of

$$\sum_{n=A+1}^{B} a_n f(n)$$

from estimates for the partial sums $S(t) := \sum_{k \le t} a_k$. Usually f is continuously differentiable on $[A, B]$, so we can replace our sum by the appropriate Riemann-Stieltjes integral, and then integrate by parts as follows:[1]

PS2
$$\begin{aligned} \sum_{A < n \le B} a_n f(n) &= \int_{A^+}^{B^+} f(t)d(S(t)) = [S(t)f(t)]_A^B - \int_A^B S(t)f'(t)dt \\ &= S(B)f(B) - S(A)f(A) - \int_A^B S(t)f'(t)dt. \end{aligned} \tag{1.1.4}$$

(Note that (1.1.4) continues to hold for all non-negative real numbers $A < B$).

[1]The notation "t^+" denotes a real number "marginally" larger than t.

In Abel's approach one does not need to make any assumption about f: Simply write $a_n = S(n) - S(n-1)$, so that

$$\sum_{n=A+1}^{B} a_n f(n) = \sum_{n=A+1}^{B} f(n)(S(n) - S(n-1)),$$

and with a little rearranging we obtain

PS1
$$\sum_{n=A+1}^{B} a_n f(n) = S(B)f(B) - S(A)f(A) - \sum_{n=A}^{B-1} S(n)(f(n+1) - f(n)). \tag{1.1.5}$$

If we now suppose that f is continuously differentiable on $[A, B]$ (as above) then we can rewrite (1.1.5) as (1.1.4).

Exercise 1.1.1. Use partial summation to show that (1.1.1) is equivalent to

PNT3
$$\theta(x) = \sum_{p \le x} \log p = x + o(x); \tag{1.1.6}$$

and then show that both are equivalent to (1.1.3).

The Riemann zeta function is given by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p} \left(1 - \frac{1}{p^s}\right)^{-1} \quad \text{for } \mathrm{Re}(s) > 1.$$

This definition is restricted to the region $\mathrm{Re}(s) > 1$, since it is only there that this Dirichlet series and this Euler product both converge absolutely (see the next subsection for definitions).

zeta

Exercise 1.1.2. (i) Prove that for $\mathrm{Re}(s) > 1$

$$\zeta(s) = s \int_1^{\infty} \frac{[y]}{y^{s+1}} dy = \frac{s}{s-1} - s \int_1^{\infty} \frac{\{y\}}{y^{s+1}} dy.$$

where throughout we write $[t]$ for the integer part of t, and $\{t\}$ for its fractional part (so that $t = [t] + \{t\}$).
The right hand side is an analytic function of s in the region $\mathrm{Re}(s) > 0$ except for a simple pole at $s = 1$ with residue 1. Thus we have an analytic continuation of $\zeta(s)$ to this larger region, and near $s = 1$ we have the Laurent expansion

$$\zeta(s) = \frac{1}{s-1} + \gamma + c_1(s-1) + \dots.$$

(The value of the constant γ is given in exercise 1.1.4.)

(ii) Deduce that $\zeta(1 + \frac{1}{\log x}) = \log x + \gamma + O_{x\to\infty}(\frac{1}{\log x})$.
(iii) † Adapt the argument in Exercise 1.1.5 to obtain an analytic continuation of $\zeta(s)$ to the region $\mathrm{Re}(s) > -1$.
(iv) † Generalize.

1.1.2. Chebyshev's elementary estimates

Chebyshev made significant progress on the distribution of primes by showing that there are constants $0 < c < 1 < C$ with

$$(c + o(1))\frac{x}{\log x} \le \pi(x) \le (C + o(1))\frac{x}{\log x}. \tag{1.1.7}$$

Moreover he showed that if

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x}$$

exists, then it must equal 1.

The key to obtaining such information is to write the prime factorization of n in the form

$$\log n = \sum_{d|n} \Lambda(d).$$

Summing both sides over n (and re-writing "$d|n$" as "$n = dk$"), we obtain that

$$\sum_{n\le x} \log n = \sum_{n\le x}\sum_{n=dk} \Lambda(d) = \sum_{k=1}^{\infty} \psi(x/k). \tag{1.1.8}$$

Using Stirling's formula, Exercise 1.1.5, we deduce that

$$\sum_{k=1}^{\infty} \psi(x/k) = x\log x - x + O(\log x). \tag{1.1.9}$$

Exercise 1.1.3. Use (1.1.9) to prove that

$$\limsup_{x\to\infty} \frac{\psi(x)}{x} \ge 1 \ge \liminf_{x\to\infty} \frac{\psi(x)}{x},$$

so that if $\lim_{x\to\infty} \psi(x)/x$ exists it must be 1.

To obtain Chebyshev's estimates (1.1.7), take (1.1.8) at $2x$ and subtract twice that relation taken at x. This yields

$$x\log 4 + O(\log x) = \psi(2x) - \psi(2x/2) + \psi(2x/3) - \psi(2x/4) + \ldots,$$

and upper and lower estimates for the right hand side above follow upon truncating the series after an odd or even number of steps. In particular we obtain that

$$\psi(2x) \ge x\log 4 + O(\log x),$$

which gives the lower bound of (1.1.7) with $c = \log 2$ a permissible value. And we also obtain that

$$\psi(2x) - \psi(x) \le x\log 4 + O(\log x),$$

which, when used at $x/2, x/4, \ldots$ and summed, leads to $\psi(x) \le x\log 4 + O((\log x)^2)$. Thus we obtain the upper bound in (1.1.7) with $C = \log 4$ a permissible value.

Returning to (1.1.8), we may recast it as

$$\sum_{n\le x} \log n = \sum_{d\le x} \Lambda(d) \sum_{k\le x/d} 1 = \sum_{d\le x} \Lambda(d)\Big(\frac{x}{d} + O(1)\Big).$$

Using Stirling's formula, and the recently established $\psi(x) = O(x)$, we conclude that

$$x\log x + O(x) = x\sum_{d\le x} \frac{\Lambda(d)}{d},$$

or in other words

Pavg

$$\sum_{p\le x}\frac{\log p}{p}=\sum_{n\le x}\frac{\Lambda(n)}{n}+O(1)=\log x+O(1). \tag{1.1.10}$$

1.1.3. Multiplicative functions and Dirichlet series

The main objects of study in this book are *multiplicative functions*. These are functions $f:\mathbb{N}\to\mathbb{C}$ satisfying $f(mn)=f(m)f(n)$ for all coprime integers m and n. If the relation $f(mn)=f(m)f(n)$ holds for all integers m and n we say that f is *completely multiplicative*. If $n=\prod_j p_j^{\alpha_j}$ is the prime factorization of n, where the primes p_j are distinct, then $f(n)=\prod_j f(p_j^{\alpha_j})$ for multiplicative functions f. Thus a multiplicative function is specified by its values at prime powers and a completely multiplicative function is specified by its values at primes.

One can study the multiplicative function $f(n)$ using the *Dirichlet series*,

$$F(s)=\sum_{n=1}^{\infty}\frac{f(n)}{n^s}=\prod_p\Big(1+\frac{f(p)}{p^s}+\frac{f(p^2)}{p^{2s}}+\dots\Big).$$

The product over primes above is called an *Euler product*, and viewed formally the equality of the Dirichlet series and the Euler product above is a restatement of the unique factorization of integers into primes. If we suppose that the multiplicative function f does not grow rapidy – for example, that $|f(n)|\ll n^A$ for some constant A – then the Dirichlet series and Euler product will converge absolutely in some half-plane with $\mathrm{Re}(s)$ suitably large.

Given any two functions f and g from $\mathbb{N}\to\mathbb{C}$ (not necessarily multiplicative), their *Dirichlet convolution* $f*g$ is defined by

$$(f*g)(n)=\sum_{ab=n}f(a)g(b).$$

If $F(s)=\sum_{n=1}^{\infty}f(n)n^{-s}$ and $G(s)=\sum_{n=1}^{\infty}g(n)n^{-s}$ are the associated Dirichlet series, then the convolution $f*g$ corresponds to their product:

$$F(s)G(s)=\sum_{n=1}^{\infty}\frac{(f*g)(n)}{n^s}.$$

The basic multiplicative functions and their associated Dirichlet series are:

• The function $\delta(1)=1$ and $\delta(n)=0$ for all $n\ge 2$ has the associated Dirichlet series 1.

• The function $1(n)=1$ for all $n\in\mathbb{N}$ has the associated Dirichlet series $\zeta(s)$ which converges absolutely when $\mathrm{Re}(s)>1$, and whose analytic continuation we discussed in Exercise 1.1.2.

• For a natural number k, the k-divisor function $d_k(n)$ counts the number of ways of writing n as $a_1\cdots a_k$. That is, d_k is the k-fold convolution of the function $1(n)$, and its associated Dirichlet series is $\zeta(s)^k$. The function $d_2(n)$ is called the *divisor function* and denoted simply by $d(n)$. More generally, for any complex number z, the z-th divisor function $d_z(n)$ is defined as the coefficient of $1/n^s$ in the Dirichlet series, $\zeta(s)^z$.[2]

[2] To explicitly determine $\zeta(s)^z$ it is easiest to expand each factor in the Euler product using the generalized binomial theorem, so that $\zeta(s)^z=\prod_p\Big(1+\sum_{k\ge1}\binom{-z}{k}(-p^{-s})^k\Big)$.

• The *Möbius function* $\mu(n)$ is defined to be 0 if n is divisible by the square of some prime and, if n is square-free, $\mu(n)$ is 1 or -1 depending on whether n has an even or odd number of prime factors. The associated Dirichlet series $\sum_{n=1}^{\infty} \mu(n) n^{-s} = \zeta(s)^{-1}$ so that μ is the same as d_{-1}. We deduce that $\mu * 1 = \delta$.

• The von Mangoldt function $\Lambda(n)$ is not multiplicative, but is of great interest to us. We write its associated Dirichlet series as $L(s)$. Since

$$\log n = \sum_{d|n} \Lambda(d) = (1 * \Lambda)(n)$$

hence $-\zeta'(s) = L(s)\zeta(s)$, that is $L(s) = (-\zeta'/\zeta)(s)$. Writing this as

$$\frac{1}{\zeta(s)} \cdot (-\zeta'(s))$$

we deduce that

Lammu

$$\Lambda(n) = (\mu * \log)(n) = \sum_{ab=n} \mu(a) \log b. \tag{1.1.11}$$

As mentioned earlier, our goal in this chapter is to show that the prime number theorem is equivalent to a statement about the mean value of the multiplicative function μ. We now formulate this equivalence precisely.

PNTM

[PNT and the mean of the Mőbius function] The prime number theorem, namely $\psi(x) = x + o(x)$, is equivalent to

Mx

$$M(x) = \sum_{n \le x} \mu(n) = o(x). \tag{1.1.12}$$

In other words, half the non-zero values of $\mu(n)$ equal 1, the other half -1.

Before we can prove this, we need one more ingredient: namely, we need to understand the average value of the divisor function.

PrS4

1.1.4. The average value of the divisor function and Dirichlet's hyperbola method

We wish to evaluate asymptotically $\sum_{n \le x} d(n)$. An immediate idea gives

$$\begin{aligned}\sum_{n \le x} d(n) &= \sum_{n \le x} \sum_{d|n} 1 = \sum_{d \le x} \sum_{\substack{n \le x \\ d|n}} 1 \\ &= \sum_{d \le x} \left[\frac{x}{d}\right] = \sum_{d \le x} \left(\frac{x}{d} + O(1)\right) \\ &= x \log x + O(x).\end{aligned}$$

Dirichlet realized that one can substantially improve the error term above by pairing each divisor a of an integer n with its complementary divisor $b = n/a$; one minor exception is when $n = m^2$ and the divisor m cannot be so paired. Since a or n/a must be $\le \sqrt{n}$ we have

$$d(n) = \sum_{d|n} 1 = 2 \sum_{\substack{d|n \\ d<\sqrt{n}}} 1 + \delta_n,$$

where $\delta_n = 1$ if n is a square, and 0 otherwise. Therefore

$$\begin{aligned}\sum_{n\le x} d(n) &= 2\sum_{n\le x}\sum_{\substack{d|n\\ d<\sqrt{n}}} 1 + \sum_{\substack{n\le x\\ n=d^2}} 1\\ &= \sum_{d\le\sqrt{x}}\Big(1+2\sum_{\substack{d^2<n\le x\\ d|n}} 1\Big)\\ &= \sum_{d\le\sqrt{x}}(2[x/d]-2d+1),\end{aligned}$$

and so

DD

$$\sum_{n\le x} d(n) = 2x\sum_{d\le\sqrt{x}}\frac{1}{d} - x + O(\sqrt{x}) = x\log x - x + 2\gamma x + O(\sqrt{x}), \tag{1.1.13}$$

by Exercise 1.1.4.

The method described above is called the *hyperbola method* because we are trying to count the number of lattice points (a,b) with a and b non-negative and lying below the hyperbola $ab = x$. Dirichlet's idea may be thought of as choosing parameters A, B with $AB = x$, and dividing the points under the hyperbola according to whether $a \le A$ or $b \le B$ or both. We remark that an outstanding open problem, known as the Dirichlet divisor problem, is to show that the error term in (1.1.13) may be improved to $O(x^{\frac14+\epsilon})$ (for any fixed $\epsilon > 0$).

For our subsequent work, we use Exercise 1.1.5 to recast (1.1.13) as

divest

$$\sum_{n\le x}(\log n - d(n) + 2\gamma) = O(\sqrt{x}). \tag{1.1.14}$$

Primes5

1.1.5. The prime number theorem and the Möbius function: proof of Theorem 1.1.3

First we show that the estimate $M(x) = \sum_{n\le x}\mu(n) = o(x)$ implies the prime number theorem $\psi(x) = x + o(x)$.

Define the arithmetic function $a(n) = \log n - d(n) + 2\gamma$, so that

$$a(n) = (1 * (\Lambda - 1))(n) + 2\gamma 1(n).$$

When we form the Dirichlet convolution of a with the Möbius function we therefore obtain

$$(\mu * a)(n) = (\mu * 1 * (\Lambda - 1))(n) + 2\gamma(\mu * 1)(n) = (\Lambda - 1)(n) + 2\gamma\delta(n),$$

where $\delta(1) = 1$, and $\delta(n) = 0$ for $n > 1$. Hence, when we sum $(\mu * a)(n)$ over all $n \le x$, we obtain

$$\sum_{n\le x}(\mu * a)(n) = \sum_{n\le x}(\Lambda(n) - 1) + 2\gamma = \psi(x) - x + O(1).$$

On the other hand, we may write the left hand side above as

$$\sum_{dk\le x}\mu(d)a(k),$$

and, as in the hyperbola method, split this into terms where $k \le K$ or $k > K$ (in which case $d \le x/K$). Thus we find that

$$\sum_{dk\le x} \mu(d)a(k) = \sum_{k\le K} a(k)M(x/k) + \sum_{d\le x/K} \mu(d) \sum_{K<k\le x/d} a(k).$$

Using (1.1.14) we see that the second term above is

$$= O\Big(\sum_{d\le x/K} \sqrt{x/d} \Big) = O(x/\sqrt{K}).$$

Putting everything together, we deduce that

$$\psi(x) - x = \sum_{k\le K} a(k)M(x/k) + O(x/\sqrt{K}).$$

Now suppose that $M(x) = o(x)$. Fix $\epsilon > 0$ and select K to be the smallest integer $> 1/\epsilon^2$, and then let $\alpha_K := \sum_{k\le K} |a(k)|/k$. Finally choose y_ϵ so that $|M(y)| \le (\epsilon/\alpha_k)y$ whenever $y \ge y_\epsilon$. Inserting all this into the last line for $x \ge Ky_\epsilon$ yields $\psi(x) - x \ll (\epsilon/\alpha_k)x \sum_{k\le K} |a(k)|/k + \epsilon x \ll \epsilon x$. We may conclude that $\psi(x) - x = o(x)$, the prime number theorem.

Now we turn to the converse. Consider the arithmetic function $-\mu(n)\log n$ which is the coefficient of $1/n^s$ in the Dirichlet series $(1/\zeta(s))'$. Since

$$\Big(\frac{1}{\zeta(s)}\Big)' = -\frac{\zeta'(s)}{\zeta(s)^2} = -\frac{\zeta'}{\zeta}(s) \cdot \frac{1}{\zeta(s)},$$

we obtain the identity $-\mu(n)\log n = (\mu * \Lambda)(n)$. As $\mu * 1 = \delta$, we find that

$$\sum_{n\le x} (\mu * (\Lambda - 1))(n) = -\sum_{n\le x} \mu(n)\log n - 1. \tag{1.1.15}$$

The right hand side of (1.1.15) is

$$\begin{aligned} -\log x \sum_{n\le x} \mu(n) + \sum_{n\le x} \mu(n)\log(x/n) - 1 &= -(\log x)M(x) + O\Big(\sum_{n\le x} \log(x/n)\Big) \\ &= -(\log x)M(x) + O(x), \end{aligned}$$

upon using Exercise 1.1.5. The left hand side of (1.1.15) is

$$\sum_{ab\le x} \mu(a)(\Lambda(b) - 1) = \sum_{a\le x} \mu(a)\Big(\psi(x/a) - x/a\Big).$$

Now suppose that $\psi(x) - x = o(x)$, the prime number theorem, so that, for given $\epsilon > 0$ we have $|\psi(t) - t| \le \epsilon t$ if $t \ge T_\epsilon$. Suppose that $T \ge T_\epsilon$ and $x > T^{1/\epsilon}$. Using this $|\psi(x/a) - x/a| \le \epsilon x/a$ for $a \le x/T$ (so that $x/a > T$), and the Chebyshev estimate $|\psi(x/a) - x/a| \ll x/a$ for $x/T \le a \le x$, we find that the left hand side of (1.1.15) is

$$\ll \sum_{a\le x/T} \epsilon x/a + \sum_{x/T\le a\le x} x/a \ll \epsilon x \log x + x \log T.$$

Combining these observations, we find that

$$|M(x)| \ll \epsilon x + x\frac{\log T}{\log x} \ll \epsilon x,$$

if x is sufficiently large. Since ϵ was arbitrary, we have demonstrated that $M(x) = o(x)$.

1.1.6. Selberg's formula

The elementary techniques discussed above were brilliantly used by Selberg to get an asymptotic formula for a suitably weighted sum of primes and products of two primes. *Selberg's formula* then led Erdős and Selberg to discover elementary proofs of the prime number theorem. We will not discuss these elementary proofs of the prime number theorem here, but let us see how Selberg's formula follows from the ideas developed so far.

Selberg [Selberg's formula] We have

$$\sum_{p\le x}(\log p)^2+\sum_{pq\le x}(\log p)(\log q)=2x\log x+O(x).$$

Proof. We define $\Lambda_2(n):=\Lambda(n)\log n+\sum_{\ell m=n}\Lambda(\ell)\Lambda(m)$. Thus $\Lambda_2(n)$ is the coefficient of $1/n^s$ in the Dirichlet series

$$\Big(\frac{\zeta'}{\zeta}(s)\Big)'+\Big(\frac{\zeta'}{\zeta}(s)\Big)^2=\frac{\zeta''(s)}{\zeta(s)},$$

so that $\Lambda_2=(\mu*(\log)^2)$.

In the previous section we exploited the fact that $\Lambda=(\mu*\log)$ and that the function $d(n)-2\gamma$ has the same average value as $\log n$. Now we search for a divisor type function which has the same average as $(\log n)^2$.

By partial summation we find that

$$\sum_{n\le x}(\log n)^2=x(\log x)^2-2x\log x+2x+O((\log x)^2).$$

Using Exercise k-div 1.1.14 we may find constants c_2 and c_1 such that

$$\sum_{n\le x}(2d_3(n)+c_2d(n)+c_1)=x(\log x)^2-2x\log x+2x+O(x^{2/3+\epsilon}).$$

Set $b(n)=(\log n)^2-2d_3(n)-c_2d(n)-c_1$ so that the last two displayed equations give

Pr61 (1.1.16)
$$\sum_{n\le x}b(n)=O(x^{2/3+\epsilon}).$$

Now consider $(\mu*b)(n)=\Lambda_2(n)-2d(n)-c_2-c_1\delta(n)$, and summing this over all $n\le x$ we get that

$$\sum_{n\le x}(\mu*b)(n)=\sum_{n\le x}\Lambda_2(n)-2x\log x+O(x).$$

The left hand side is

$$\sum_{k\le x}\mu(k)\sum_{l\le x/k}b(l)\ll\sum_{k\le x}(x/k)^{2/3+\epsilon}\ll x$$

by (Pr61 1.1.16), and we conclude that

$$\sum_{n\le x}\Lambda_2(n)=2x\log x+O(x).$$

The difference between the left hand side above and the left hand side of our desired formula is the contribution of the prime powers, which is easily shown to be $\ll\sqrt{x}\log x$, and so our Theorem follows. $\square$

1.1.7. Exercises

ex:harmonic

EXERCISE 1.1.4. * (i) Using partial summation, prove that for any $x \geq 1$

$$\sum_{1\leq n\leq x} \frac{1}{n} = \log x + \frac{[x]}{x} - \int_1^x \frac{\{t\}}{t^2} dt.$$

(ii) Deduce that for any $x \geq 1$ we have the approximation

$$\left| \sum_{n\leq x} \frac{1}{n} - (\log x + \gamma) \right| \leq \frac{1}{x},$$

where γ is the *Euler-Mascheroni constant*,

$$\gamma := \lim_{N\to\infty} \Big(\sum_{n=1}^{N} \frac{1}{n} - \log N \Big) = 1 - \int_1^\infty \frac{\{t\}}{t^2} dt.$$

ex:stirling

EXERCISE 1.1.5. (i) For an integer $N \geq 1$ show that

$$\log N! = N \log N - N + 1 + \int_1^N \frac{\{t\}}{t} dt.$$

(ii) Deduce that $x - 1 \geq \sum_{n\leq x} \log(x/n) \geq x - 2 - \log x$ for all $x \geq 1$.

(iii) Using that $\int_1^x (\{t\} - 1/2)dt = (\{x\}^2 - \{x\})/2$ and integrating by parts, show that

$$\int_1^N \frac{\{t\}}{t} dt = \frac{1}{2} \log N - \frac{1}{2} \int_1^N \frac{\{t\} - \{t\}^2}{t^2} dt.$$

(iv) Conclude that $N! = C\sqrt{N}(N/e)^N \{1 + O(1/N)\}$, where

$$C = \exp\Big(1 - \frac{1}{2} \int_1^\infty \frac{\{t\} - \{t\}^2}{t^2} dt\Big). \text{ In fact } C = \sqrt{2\pi},$$

and the resulting asymptotic for $N!$, namely $N! \sim \sqrt{2\pi N}(N/e)^N$, is known as *Stirling's formula.*

zeta2

EXERCISE 1.1.6. *

(i) Prove that for $\mathrm{Re}(s) > 0$ we have

$$\sum_{n=1}^{N} \frac{1}{n^s} - \int_1^N \frac{dt}{t^s} = \zeta(s) - \frac{1}{s-1} + s \int_N^\infty \frac{\{y\}}{y^{s+1}} dy.$$

(ii) Deduce that, in this same range but with $s \neq 1$, we can define

$$\zeta(s) = \lim_{N\to\infty} \left\{ \sum_{n=1}^{N} \frac{1}{n^s} - \frac{N^{1-s}}{1-s} \right\}.$$

ex:Bertrand

EXERCISE 1.1.7. * Using that $\psi(2x) - \psi(x) + \psi(2x/3) \geq x \log 4 + O(\log x)$, prove Bertrand's postulate that there is a prime between N and $2N$, for N sufficiently large.

ex:Cheb

EXERCISE 1.1.8. (i) Using (Cheb2 1.1.8), prove that if $L(x) := \sum_{n\leq x} \log n$ then

$$\psi(x) - \psi(x/6) \leq L(x) - L(x/2) - L(x/3) - L(x/5) + L(x/30) \leq \psi(x).$$

(ii) Deduce, using (1.1.9), that with

$$\kappa = \frac{\log 2}{2} + \frac{\log 3}{3} + \frac{\log 5}{5} - \frac{\log 30}{30} = 0.9212920229\ldots,$$

we have $\kappa x + O(\log x) \le \psi(x) \le \frac{6}{5}\kappa x + O(\log^2 x)$.

(iii) † Improve on these bounds by similar methods.

Pavg+

EXERCISE 1.1.9. (i) Use partial summation to prove that if

$$\lim_{N\to\infty} \sum_{n\le N} \frac{\Lambda(n) - 1}{n} \quad \text{exists,}$$

then the prime number theorem, in the form $\psi(x) = x + o(x)$, follows.

(ii) † Prove that the prime number theorem implies that this limit holds.

(iii) Using exercise 1.1.2, prove that $-(\zeta'/\zeta)(s) - \zeta(s)$ has a Taylor expansion $-2\gamma + c_1'(s-1) + \ldots$ around $s = 1$.

(iv) Explain why we cannot then deduce that

$$\lim_{N\to\infty} \sum_{n\le N} \frac{\Lambda(n) - 1}{n} = \lim_{s\to 1^+} \sum_{n\ge 1} \frac{\Lambda(n) - 1}{n^s}, \text{ which exists and equals } -2\gamma.$$

exmertens

EXERCISE 1.1.10. *

(i) Use (1.1.10) and partial summation show that there is a constant c such that

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

(ii) Deduce *Mertens' Theorem*, that there exists a constant γ such that

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

In the two preceding exercises the constant γ is in fact the Euler-Mascheroni constant, but this is not so straightforward to establish. The next exercise gives one way of obtaining information about the constant in Exercise 1.1.10.

exmertens2

EXERCISE 1.1.11. † In this exercise, put $\sigma = 1 + 1/\log x$.

(i) Show that

$$\sum_{p>x} \log\left(1 - \frac{1}{p^\sigma}\right)^{-1} = \sum_{p>x} \frac{1}{p^\sigma} + O\Big(\frac{1}{x}\Big) = \int_1^\infty \frac{e^{-t}}{t} dt + O\Big(\frac{1}{\log x}\Big).$$

(ii) Show that

$$\sum_{p\le x} \left(\log\left(1 - \frac{1}{p^\sigma}\right)^{-1} - \log\left(1 - \frac{1}{p}\right)^{-1}\right) = -\int_0^1 \frac{1 - e^{-t}}{t} dt + O\Big(\frac{1}{\log x}\Big).$$

(iii) Conclude, using exercise 1.1.2, that the constant γ in exercise 1.1.10(ii) equals

$$\int_0^1 \frac{1 - e^{-t}}{t} dt - \int_1^\infty \frac{e^{-t}}{t} dt.$$

That this equals the Euler-Mascheroni constant is established in [?].

ex2.7

Exercise 1.1.12. * Uniformly for η in the range $\frac{1}{\log y} \ll \eta < 1$, show that

$$\sum_{p\le y} \frac{\log p}{p^{1-\eta}} \ll \frac{y^\eta}{\eta} \ ;$$

and

$$\sum_{p\le y} \frac{1}{p^{1-\eta}} \le \log(1/\eta) + O\Big(\frac{y^\eta}{\log(y^\eta)}\Big).$$

Hint: Split the sum into those primes with $p^\eta \ll 1$, and those with $p^\eta \gg 1$.

Exercise 1.1.13. * If f and g are functions from $\mathbb{N}$ to $\mathbb{C}$, show that the relation $f = 1 * g$ is equivalent to the relation $g = \mu * f$. (Given two proofs.) This is known as *Möbius inversion*.

k-div

Exercise 1.1.14. (i) Given a natural number k, use the hyperbola method together with induction and partial summation to show that

$$\sum_{n\le x} d_k(n) = xP_k(\log x) + O(x^{1-1/k+\epsilon})$$

where $P_k(t)$ denotes a polynomial of degree $k-1$ with leading term $t^{k-1}/(k-1)!$.
(ii) Deduce, using partial summation, that if $R_k(t) + R_k'(t) = P_k(t)$ then

$$\sum_{n\le x} d_k(n)\log(x/n) = xR_k(\log x) + O(x^{1-1/k+\epsilon}).$$

(iii) Deduce, using partial summation, that if $Q_k(u) = P_k(u) + \int_{t=0}^{u} P_k(t)dt$ then

$$\sum_{n\le x} \frac{d_k(n)}{n} = Q_k(\log x) + O(1).$$

Analogies of these estimates hold for any real $k > 0$, in which case $(k-1)!$ is replaced by $\Gamma(k)$.

ex:MobiusEquiv

Exercise 1.1.15. Modify the above proof to show that
(i) If $M(x) \ll x/(\log x)^A$ then $\psi(x) - x \ll x(\log\log x)^2/(\log x)^A$.
(ii) Conversely, if $\psi(x) - x \ll x/(\log x)^A$ then $M(x) \ll x/(\log x)^{\min(1,A)}$.

MobPNT

Exercise 1.1.16. (i) * Show that

$$M(x)\log x = -\sum_{p\le x} \log p \ M(x/p) + O(x).$$

(ii) Deduce that

$$\liminf_{x\to\infty} \frac{M(x)}{x} + \limsup_{x\to\infty} \frac{M(x)}{x} = 0.$$

(iii) Use Selberg's formula to prove that

$$(\psi(x) - x)\log x = -\sum_{p\le x} \log p \ \left(\psi\left(\frac{x}{p}\right) - \frac{x}{p}\right) + O(x).$$

(iv) Deduce that

$$\liminf_{x\to\infty} \frac{\psi(x)-x}{x} \quad + \quad \limsup_{x\to\infty} \frac{\psi(x)-x}{x} \ = \ 0.$$

Compare!

CHAPTER 1.2

First results on multiplicative functions

C2

We have just seen that understanding the mean value of the Möbius function leads to the prime number theorem. Motivated by this, we now begin a more general study of mean values of multiplicative functions.

S2.1

1.2.1. A heuristic

In Section 1.1.4 we saw that one can estimate the mean value of the k-divisor function by writing d_k as the convolution $1 * d_{k-1}$. Given a multiplicative function f, let us write f as $1 * g$ so that g is also multiplicative. Then

$$\sum_{n \le x} f(n) = \sum_{n \le x} \sum_{d|n} g(d) = \sum_{d \le x} g(d) \Big[\frac{x}{d} \Big].$$

Since $[z] = z + O(1)$ we have

E2.1

$$\sum_{n \le x} f(n) = x \sum_{d \le x} \frac{g(d)}{d} + O\Big(\sum_{d \le x} |g(d)| \Big). \tag{1.2.1}$$

In several situations, for example in the case of the k-divisor function treated earlier, the remainder term in (1.2.1) may be shown to be small. Omitting this term, and approximating $\sum_{d \le x} g(d)/d$ by $\prod_{p \le x}(1 + g(p)/p + g(p^2)/p^2 + \ldots)$ we arrive at the following heuristic:

E2.2

$$\sum_{n \le x} f(n) \approx x \; \mathcal{P}(f; x) \tag{1.2.2}$$

where "≈" is interpreted as "is roughly equal to", and

E2.3

(1.2.3)

$$\mathcal{P}(f;x) = \prod_{p \le x} \Big(1 + \frac{g(p)}{p} + \frac{g(p^2)}{p^2} + \ldots \Big) = \prod_{p \le x} \Big(1 - \frac{1}{p} \Big) \Big(1 + \frac{f(p)}{p} + \frac{f(p^2)}{p^2} + \ldots \Big).$$

In the special case that $0 \le f(p) \le f(p^2) \le \ldots$ for all primes p (so that $g(d) \ge 0$ for all d), one easily gets an upper bound of the correct order of magnitude: If $f = 1 * g$ then $g(d) \ge 0$ for all $d \ge 1$ by assumption, and so

$$\sum_{n \le x} f(n) = \sum_{d \le x} g(d) \Big[\frac{x}{d} \Big] \le \sum_{d \le x} g(d) \frac{x}{d} \le x \; \mathcal{P}(f; x)$$

(as in (1.2.3)).

In the case of the k-divisor function, the heuristic (1.2.2) predicts that

$$\sum_{n \le x} d_k(n) \approx x \prod_{p \le x} \Big(1 - \frac{1}{p} \Big)^{-(k-1)} \sim x (e^\gamma \log x)^{k-1},$$

which is off from the correct asymptotic formula, $\sim x(\log x)^{k-1}/(k-1)!$, by only a constant factor (see exercise 1.1.14(i)). Moreover $d_k(p^j) \geq d_k(p^{j-1})$ for all p^j so this yields an (unconditional) upper bound.

One of our aims will be to obtain results that are uniform over the class of all mutiplicative functions. Thus for example we could consider x to be large and consider the multiplicative function f with $f(p^k) = 0$ for $p \leq \sqrt{x}$ and $f(p^k) = 1$ for $p > \sqrt{x}$. In this case, we have $f(n) = 1$ if n is a prime between $\sqrt{x}$ and x and $f(n) = 0$ for other $n \leq x$. Thus, the heuristic suggests that

$$\pi(x) - \pi(\sqrt{x}) + 1 = \sum_{n\leq x} f(n) \approx x \prod_{p\leq\sqrt{x}} \left(1 - \frac{1}{p}\right) \sim x\frac{e^{-\gamma}}{\log\sqrt{x}} \sim \frac{2e^{-\gamma}x}{\log x}.$$

Comparing this to the prime number theorem, the heuristic is off by a constant factor again, this time $2e^{-\gamma} \approx 1.1...$.

This heuristic suggests that the sum of the Möbius function,

$$M(x) = \sum_{n\leq x} \mu(n) \qquad \text{is comparable with} \qquad x\prod_{p\leq x}\left(1-\frac{1}{p}\right)^2 \sim \frac{xe^{-2\gamma}}{(\log x)^2}.$$

However $M(x)$ is known to be much smaller. The best bound that we know unconditionally is that $M(x) \ll x\exp(-c(\log x)^{\frac{3}{5}-\epsilon})$ (see chapter ??), and we expect $M(x)$ to be as small as $x^{\frac{1}{2}+\epsilon}$ (as this is equivalent to the unproved *Riemann Hypothesis*). In any event, the heuristic certainly suggests that $M(x) = o(x)$, which is equivalent to the prime number theorem, as we saw in Theorem 1.1.3.

S2.2

1.2.2. Multiplicative functions and Dirichlet series

Given a multiplicative function $f(n)$ we define $F(s) := \sum_{n\geq 1} \frac{f(n)}{n^s}$ as usual, and now define the coefficients $\Lambda_f(n)$ by

$$-\frac{F'(s)}{F(s)} = \sum_{n\geq 1} \frac{\Lambda_f(n)}{n^s}.$$

Comparing the coefficient of $1/n^s$ in $-F'(s) = F(s)\cdot(-F'(s)/F(s))$ we have

ConvolEqNew (1.2.4)
$$f(n)\log n = \sum_{d|n} \Lambda_f(d) f(n/d).$$

LambdaF EXERCISE 1.2.1. Let f be a multiplicative function. and fix $\kappa > 0$

(i) Show that $\Lambda_f(n) = 0$ unless n is a prime power.
(ii) Show that if f is totally multiplicative then $\Lambda_f(n) = f(n)\Lambda(n)$.
(iii) Show that $\Lambda_f(p) = f(p)\log p$, $\Lambda_f(p^2) = (2f(p^2) - f(p)^2)\log p$, and that every $\Lambda_f(p^k)$ equals $\log p$ times some polynomial in $f(p), f(p^2), \ldots, f(p^k)$.
(iv) Show that if $|\Lambda_f(n)| \leq \kappa\Lambda(n)$ for all n, then $|f(n)| \leq d_\kappa(n)$.

ex2.1 EXERCISE 1.2.2. Suppose that f is a non-negative arithmetic function, and that $F(\sigma) = \sum_{n=1}^{\infty} f(n)n^{-\sigma}$ is convergent for some $\sigma > 0$.

(i) Prove that $\sum_{n\leq x} f(n) \leq x^\sigma F(\sigma)$.
(ii) Moreover show that if $0 < \sigma < 1$ then

$$\sum_{n\leq x} f(n) + x\sum_{n>x}\frac{f(n)}{n} \leq x^\sigma F(\sigma).$$

This technique is known as *Rankin's trick*, and is surprisingly effective. The values $f(p^k)$ for $p^k > x$ appear in the Euler product for $F(\sigma)$ and yet are irrelevant to the mean value of $f(n)$ for n up to x. However, for a given x, we can take $f(p^k) = 0$ for every $p^k > x$, to minimize the value of $F(\sigma)$ above.

S2.2

1.2.3. Multiplicative functions close to 1

The heuristic (1.2.2) is accurate and easy to justify when the function g is small in size, or in other words, when f is close to 1. We give a sample such result which will lead to several applications.

pr2.1 Proposition 1.2.1. *Let $f = 1 * g$ be a multiplicative function. If*

$$\sum_{d=1}^{\infty} \frac{|g(d)|}{d^\sigma} = \widetilde{G}(\sigma)$$

is convergent for some σ, $0 \le \sigma \le 1$, then

$$\Big| \sum_{n\le x} f(n) - x\mathcal{P}(f) \Big| \le x^\sigma \widetilde{G}(\sigma),$$

where $\mathcal{P}(f) := \mathcal{P}(f;\infty)$, and

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\le x} f(n) = \mathcal{P}(f).$$

If $\widetilde{G}(\sigma)$ converges then $\widetilde{G}(1)$ does. If each $|f(n)| \le 1$ then $\widetilde{G}(1)$ converges if and only if $\sum_p \frac{|1-f(p)|}{p} < \infty$.

Proof. The argument giving (1.2.1) yields that

$$\Big| \sum_{n\le x} f(n) - x \sum_{d\le x} \frac{g(d)}{d} \Big| \le \sum_{d\le x} |g(d)|.$$

Since $\mathcal{P}(f) = \sum_{d\ge 1} g(d)/d$ we have that

$$\Big| \sum_{d\le x} \frac{g(d)}{d} - \mathcal{P}(f) \Big| \le \sum_{d>x} \frac{|g(d)|}{d}.$$

Combining these two inequalities yields

SweetBound (1.2.5)
$$\Big| \sum_{n\le x} f(n) - x\mathcal{P}(f) \Big| \le \sum_{d\le x} |g(d)| + x \sum_{d>x} \frac{|g(d)|}{d}. \tag{1.2.5}$$

We now use Rankin's trick: we multiply the terms in the first sum by $(x/d)^\sigma \ge 1$, and in the second sum by $(d/x)^{1-\sigma} > 1$, so that the right hand side of (1.2.5) is

$$\le \sum_{d\le x} |g(d)| \left(\frac{x}{d}\right)^\sigma + x \sum_{d>x} \frac{|g(d)|}{d} \left(\frac{d}{x}\right)^{1-\sigma} = x^\sigma \widetilde{G}(\sigma),$$

the first result in the lemma. This immediately implies the second result for $0 \le \sigma < 1$.

One can rewrite the right hand side of (1.2.5) as

$$\int_0^x \sum_{n>t} \frac{|g(n)|}{n} dt = o_{x\to\infty}(x),$$

because $\sum_{n>t}|g(n)|/n$ is bounded, and tends to zero as $t\to\infty$. This implies the second result for $\sigma=1$. □

sec:Non-neg

1.2.4. Non-negative multiplicative functions

Let us now consider our heuristic for the special case of non-negative multiplicative functions with suitable growth conditions. Here we shall see that right side of our heuristic (1.2.2) is at least a good upper bound for $\sum_{n\le x} f(n)$.

Prop2.1 PROPOSITION 1.2.2. *Let f be a non-negative multiplicative function, and suppose there are constants A and B for which*

eq:sumps
$$\sum_{m\le z}\Lambda_f(m)\le Az+B, \tag{1.2.6}$$

for all $z\ge 1$. Then for $x\ge e^{2B}$ we have

$$\sum_{n\le x} f(n)\le \frac{(A+1)x}{\log x+1-B}\sum_{n\le x}\frac{f(n)}{n}$$

PROOF. We begin with the decomposition

$$\begin{aligned}\sum_{n\le x} f(n)\log x &= \sum_{n\le x} f(n)\log n+\sum_{n\le x} f(n)\log(x/n)\\ &\le \sum_{n\le x} f(n)\log n+\sum_{n\le x} f(n)\Big(\frac{x}{n}-1\Big),\end{aligned}$$

which holds since $0\le \log t\le t-1$ for all $t\ge 1$. For the first term we have

$$\begin{aligned}\sum_{n\le x} f(n)\log n = \sum_{n\le x}\sum_{n=mr} f(r)\Lambda_f(m) &\le \sum_{r\le x} f(r)\sum_{m\le x/r}\Lambda_f(m)\\ &\le \sum_{r\le x} f(r)\Big(\frac{Ax}{r}+B\Big).\end{aligned}$$

The result follows by combining these two inequalities. □

Proposition 1.2.2 establishes the heuristic (1.2.3) for many common multiplicative functions:

cor2.3 COROLLARY 1.2.3. *Let f be a non-negative multiplicative function for which either $0\le f(n)\le 1$ for all n, or $|\Lambda_f(n)|\le\kappa\Lambda(n)$ for all n, for some given constant $\kappa>1$. Then*

E2.5
$$\frac{1}{x}\sum_{n\le x} f(n)\ll_{A,B}\mathcal{P}(f;x)\ll \exp\Big(-\sum_{p\le x}\frac{1-f(p)}{p}\Big). \tag{1.2.7}$$

Moreover if $0\le f(n)\le 1$ for all n then

$$\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x} f(n)=\mathcal{P}(f).$$

PROOF. The hypothesis implies that (1.2.6) holds: If $|f(n)|\le 1$ then this follows by exercise 1.2.5(iii). If each $|\Lambda_f(n)|\le\kappa\Lambda(n)$ then the Chebyshev estimates give that

$$\sum_{n\le z}|\Lambda_f(n)|\le\kappa\sum_{n\le z}\Lambda(n)\le Az+B,$$

any constant $A > \kappa \log 4$ being permissible.

So we apply Proposition 1.2.2, and bound the right-hand side using Mertens' Theorem, and

$$\sum_{n\le x} \frac{f(n)}{n} \le \prod_{p\le x} \Big(1 + \frac{f(p)}{p} + \frac{f(p^2)}{p^2} + \dots\Big),$$

to obtain the first inequality. The second inequality then follows from exercise 1.2.6 with $\epsilon = \frac{1}{2}$.

If $\sum_p (1-f(p))/p$ diverges, then (1.2.7) shows that

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\le x} f(n) = 0 = \mathcal{P}(f).$$

Suppose now that $\sum_p (1-f(p))/p$ converges. If we write $f = 1 * g$ then this condition assures us that $\sum_{p^k} |g(p^k)|/p^k$ converges, which in turn is equivalent to the convergence of $\sum_n |g(n)|/n$ by exercise 1.2.7. The second statement in Proposition 1.2.1 now finishes our proof. $\square$

In the coming chapters we will establish appropriate generalizations of Corollary 1.2.3. For example, for real-valued multiplicative functions with $-1 \le f(n) \le 1$, Wirsing proved that $\sum_{n\le x} f(n) \sim \mathcal{P}(f)x$. This implies that $\sum_{n\le x} \mu(n) = o(x)$ and hence the prime number theorem, by Theorem 1.1.3. We will go on to study Halász's seminal result on the mean values of complex-valued multiplicative functions which take values in the unit disc.

Proposition 1.2.2 also enables us to prove a preliminary result indicating that mean values of multiplicative functions vary slowly. The result given here is only useful when f is "close" to 1, but we shall see a more general such result in Chapter ??.

FirstLip

PROPOSITION 1.2.4. *Let f be a multiplicative function with $|f(n)| \le 1$ for all n. Then for all $1 \le y \le \sqrt{x}$ we have*

$$\Big| \frac{1}{x} \sum_{n\le x} f(n) - \frac{y}{x} \sum_{n\le x/y} f(n) \Big| \ll \frac{\log(ey)}{\log x} \exp\Big(\sum_{p\le x} \frac{|1-f(p)|}{p} \Big).$$

PROOF. Write $f = 1 * g$, so that g is a multiplicative function with each $g(p) = f(p) - 1$, and each $\Lambda_g(p) = \Lambda_f(p) - \Lambda(p)$ (so that (1.2.6) holds by exercise 1.2.5(iii)). Recall that

$$\Big| \frac{1}{x} \sum_{n\le x} f(n) - \sum_{d\le x} \frac{g(d)}{d} \Big| \le \frac{1}{x} \sum_{d\le x} |g(d)|,$$

so that

FirstLip1

$$\Big| \frac{1}{x} \sum_{n\le x} f(n) - \frac{y}{x} \sum_{n\le x/y} f(n) \Big| \ll \frac{1}{x} \sum_{d\le x} |g(d)| + \frac{y}{x} \sum_{d\le x/y} |g(d)| + \sum_{x/y<d\le x} \frac{|g(d)|}{d}. \tag{1.2.8}$$

Appealing to Proposition 1.2.2 we find that for any $z \ge 3$

$$\sum_{n\le z} |g(n)| \ll \frac{z}{\log z} \sum_{n\le z} \frac{|g(n)|}{n} \ll \frac{z}{\log z} \exp\Big(\sum_{p\le z} \frac{|1-f(p)|}{p} \Big).$$

From this estimate and partial summation we find that the right hand side of (1.2.8) is

$$\ll \frac{\log(ey)}{\log x} \exp\Big(\sum_{p\le x} \frac{|1-f(p)|}{p} \Big),$$

proving our Proposition. $\square$

1.2.5. Logarithmic means

In addition to the natural mean values $\frac{1}{x}\sum_{n\le x} f(n)$, we have already encountered logarithmic means $\frac{1}{\log x}\sum_{n\le x} f(n)/n$ several times in our work above. We now prove the analogy to Proposition 1.2.1 for logarithmic means:

NasProp

Proposition 1.2.5 (Naslund). *Let $f = 1 * g$ be a multiplicative function and $\sum_d |g(d)|d^{-\sigma} = \widetilde{G}(\sigma) < \infty$ for some $\sigma \in [0,1)$. Then*

$$\left| \sum_{n\le x} \frac{f(n)}{n} - \mathcal{P}(f)\left(\log x + \gamma - \sum_{n\ge 1} \frac{\Lambda_f(n)-\Lambda(n)}{n} \right) \right| \le \frac{x^{\sigma-1}}{1-\sigma}\ \widetilde{G}(\sigma).$$

Proof. We start with

$$\sum_{n\le x} \frac{f(n)}{n} = \sum_{n\le x} \frac{1}{n}\sum_{d|n} g(d) = \sum_{d\le x}\frac{g(d)}{d} \sum_{m\le x/d} \frac{1}{m}$$

and then, using exercise 1.1.4, we deduce that

$$\left| \sum_{n\le x} \frac{f(n)}{n} - \sum_{d\le x}\frac{g(d)}{d}\Big(\log\frac{x}{d} + \gamma\Big) \right| \le \sum_{d\le x}\frac{|g(d)|}{d}\cdot\frac{d}{x} = \frac{1}{x}\sum_{d\le x}|g(d)|.$$

Since $g(n)\log n$ is the coefficient of $1/n^s$ in $-G'(s) = G(s)(-G'/G)(s)$, thus $g(n)\log n = (g*\Lambda_g)(n)$, and we note that $\Lambda_f = \Lambda + \Lambda_g$. Hence

$$\sum_{n\ge 1}\frac{g(n)\log n}{n} = \sum_{a,b\ge 1}\frac{g(a)\Lambda_g(b)}{ab} = \mathcal{P}(f)\sum_{m\ge 1}\frac{\Lambda_f(m)-\Lambda(m)}{m}.$$

Therefore $\sum_{d\ge 1}\frac{g(d)}{d}\Big(\log\frac{x}{d}+\gamma\Big) = \mathcal{P}(f)\Big(\log x + \gamma - \sum_{n\ge 1}\frac{\Lambda_f(n)-\Lambda(n)}{n}\Big)$, and so the error term in our main result is

$$\le \frac{1}{x}\sum_{d\le x}|g(d)| + \sum_{d>x}\frac{|g(d)|}{d}\Big|\log\frac{x}{d}+\gamma\Big|.$$

Since $1/(1-\sigma)\ge 1$ we can use the inequalities $1 \le (x/d)^\sigma \le (x/d)^\sigma/(1-\sigma)$ for $d\le x$, and

$$|\log(x/d)+\gamma| \le 1+\log(d/x) \le \frac{(d/x)^{1-\sigma}}{1-\sigma}$$

for $d>x$, to get a bound on the error term of $\frac{x^{\sigma-1}}{1-\sigma}\ \widetilde{G}(\sigma)$ as claimed. $\square$

GOUB

Proposition 1.2.6. *If f is a multiplicative function with $|f(n)|\le 1$ for all n, then*

$$\frac{1}{\log x}\Big|\sum_{n\le x}\frac{f(n)}{n}\Big| \ll\ \exp\Big(-\frac{1}{2}\sum_{p\le x}\frac{1-\mathrm{Re}(f(p))}{p}\Big).$$

Proof. Let $h = 1 * f$, so that

$$\sum_{n\le x} h(n) = \sum_{n\le x}\sum_{d|n} f(d) = \sum_{d\le x} f(d)\Big(\frac{x}{d} + O(1)\Big) = x\sum_{d\le x}\frac{f(d)}{d} + O(x).$$

We deduce, applying Proposition 1.2.2 (since (1.2.6) is satisfied as $\Lambda_h = \Lambda + \Lambda_f$, and then by exercise 1.2.5(iii)), that

$$\begin{aligned}\frac{1}{\log x}\Big|\sum_{n\le x}\frac{f(n)}{n}\Big| &\le \frac{1}{x\log x}\sum_{n\le x}|h(n)| + O\Big(\frac{1}{\log x}\Big)\\ &\ll \frac{1}{\log^2 x}\sum_{n\le x}\frac{|h(n)|}{n} + \frac{1}{\log x}\\ &\ll \exp\left(\sum_{p\le x}\frac{|1+f(p)|-2}{p}\right) + \frac{1}{\log x}\end{aligned}$$

using Mertens' theorem. Now $\frac12(1-\mathrm{Re}(z)) \le 2 - |1+z| \le 1 - \mathrm{Re}(z)$ whenever $|z|\le 1$, and so the result follows. □

We expect that, for non-negative real multiplicative functions f, the quantity

$$\mathcal{R}(f;x) := \sum_{n\le x}\frac{f(n)}{n}\Big/\prod_{p\le x}\Big(1 + \frac{f(p)}{p} + \frac{f(p^2)}{p^2} + \dots\Big),$$

should typically be bounded, based on the heuristic discussion above. For example $\mathcal{R}(d_\kappa;x) \sim (e^{-\gamma})^\kappa/|\Gamma(\kappa+1)|$ by exercise 1.1.14(iii) and Mertens' Theorem.

SumCompare

Exercise 1.2.3. Suppose that f and g are real multiplicative functions with $f(n), g(n) \ge 0$ for all $n \ge 1$.

(i) Prove that $0 \le \mathcal{R}(f;x) \le 1$.
(ii) Prove that $\mathcal{R}(f;x) \ge \mathcal{R}(f;x)\cdot\mathcal{R}(g;x) \ge \mathcal{R}(f*g;x)$.
(iii) Deduce that if f is totally multiplicative and $0 \le f(n) \le 1$ for all $n \ge 1$ then $1 \ge \mathcal{R}(f;x) \ge \mathcal{R}(1;x) \sim e^{-\gamma}$.
(iv) Suppose that f is *supported* only on squarefree integers (that is, $f(n) = 0$ if $p^2|n$ for some prime p). Let g be the totally multiplicative function with $g(p) = f(p)$ for each prime p. Prove that $\mathcal{R}(f;x) \ge \mathcal{R}(g;x)$.

1.2.6. Exercises

Exercise 1.2.4. * Prove that if $f(.)$ is multiplicative with $-1 \le f(p^k) \le 1$ for each prime power p^k then $\lim_{x\to\infty}\mathcal{P}(f;x)$ exists and equals $\mathcal{P}(f)$

ex:WhichHypo

Exercise 1.2.5. (i) Prove that if $|f(p^k)| \le B^k$ for all prime powers p^k then $|\Lambda_f(p^k)| \le (2^k-1)B^k\log p$ for all prime powers p^k.
(ii) Show this is best possible (Hint: Try $f(p^k) = -(-B)^k$).
(iii) Show that $|f(n)| \le 1$ for all n then there exist constant A, C such that $\sum_{m\le z}|\Lambda_f(m)| \le Az + C$, for all $z \ge 1$.
(iv) Give an example of an f where $B > 1$, for which $\sum_{n\le x}|\Lambda_f(n)| \gg x^{1+\delta_B}$.

This explains why, when we consider f with values outside the unit circle, we prefer working with the hypothesis $|\Lambda_f(n)| \le \kappa\Lambda(n)$ rather than $|f(p^k)| \le B$.

SizeP(f,x)

Exercise 1.2.6. (i) Let f be a real-valued multiplicative function for which there exist constants $\kappa \geq 1$ and $\epsilon > 0$, such that $|f(p^k)| \leq d_\kappa(p^k)(p^k)^{\frac{1}{2}-\epsilon}$ for every prime power p^k. Prove that

$$\mathcal{P}(f;x) \ll_{\kappa,\epsilon} \exp\Big(-\sum_{p\leq x}\frac{1-f(p)}{p}\Big).$$

This should be interpreted as telling us that, in the situations which we are interested in, the values of $f(p^k)$ with $k>1$ have little effect on the value of $\mathcal{P}(f;x)$.
(ii) Show that if, in addition, there exists a constant $\delta > 0$ for which

$$\Big|1+\frac{f(p)}{p}+\frac{f(p^2)}{p^2}+\dots\Big| \geq \delta$$

for every prime p then

$$\mathcal{P}(f;x) \asymp_{\kappa,\delta,\epsilon} \exp\Big(-\sum_{p\leq x}\frac{1-f(p)}{p}\Big).$$

(iii)* Prove that if $|\Lambda_f(n)| \leq \Lambda(n)$ for all n then the above hypotheses hold with $\kappa = 1,\ \epsilon = \frac{1}{2}$ and $\delta = \frac{1}{4}$.

ex2.0

Exercise 1.2.7. * Show that if $g(.)$ is multiplicative then $\sum_{n\geq 1}|g(n)|/n^\sigma < \infty$ if and only if $\sum_{p^k}|g(p^k)|/p^{k\sigma} < \infty$.

ex:Prop2.1

Exercise 1.2.8. * Deduce, from Proposition 1.2.1 and the previous exercise, that if $\sum_{p^k}|f(p^k)-f(p^{k-1})|/p^k < \infty$ then $\sum_{n\leq x} f(n) \sim x\mathcal{P}(f)$ as $x\to\infty$.

ex2.2

Exercise 1.2.9. * For any natural number q, prove that for any $\sigma \geq 0$ we have

$$\Big|\sum_{\substack{n\leq x\\(n,q)=1}} 1 - \frac{\phi(q)}{q}x\Big| \leq x^\sigma \prod_{p|q}\Big(1+\frac{1}{p^\sigma}\Big).$$

Taking $\sigma = 0$, we obtain the sieve of Eratosthenes bound of $2^{\omega(q)}$.[1] Prove that the bound is optimized by the solution to $\sum_{p|q}(\log p)/(p^\sigma+1) = \log x$, if that solution is ≥ 0. Explain why the bound is of interest only if $0 \leq \sigma < 1$.

ex:near1

Exercise 1.2.10. Suppose that f is a multiplicative function "close to 1", that is $|f(p^k)-f(p^{k-1})| \leq \frac{1}{2p^k}\binom{k+r}{r}$ for all prime powers p^k, for some integer $r \geq 0$. Prove that

$$\sum_{n\leq x} f(n) = x\mathcal{P}(f) + O((\log x)^{r+1}).$$

(Hint: Use Proposition 1.2.1 with $\sigma = 0$, the Taylor expansion for $(1-t)^{-r-1}$ and Mertens' Theorem.)

ex2.3

Exercise 1.2.11. * Let $\sigma(n) = \sum_{d|n} d$. Prove that

$$\sum_{n\leq x}\frac{\mu(n)^2\sigma(n)}{\phi(n)} = \frac{15}{\pi^2}x + O(\sqrt{x}\log x).$$

ex2.5

Exercise 1.2.12. † Let f be multiplicative and write $f = d_k * g$ where $k \in \mathbb{N}$ and d_k deontes the k-divisor function. Assuming that $|g|$ is small, as in Proposition 1.2.1, develop an asymptotic formula for $\sum_{n\leq x} f(n)$.

[1] Where $\omega(q)$ denotes the number of distinct primes dividing q.

ex:Pr2.1RightConstant

EXERCISE 1.2.13. Fix $\kappa > 0$. Assume that f is a non-negative multiplicative function and that each $|\Lambda_f(n)| \leq \kappa\Lambda(n)$.

(i) In the proof of Proposition 1.2.2, modify the bound on $f(n)\log(x/n)$ using exercise 1.1.14 , to deduce that for any $A > \kappa$,

$$\sum_{n\leq x} f(n) \leq \frac{x}{\log x + O(1)}\Big(A\sum_{n\leq x}\frac{f(n)}{n} + O\Big(\frac{(\log x)^{\kappa-1}}{\Gamma(\kappa)}\Big)\Big)$$

(ii) Deduce that $\frac{1}{x}\sum_{n\leq x} f(n) \leq \kappa(e^\gamma + o(1))\ \mathcal{P}(f;x) + O((\log x)^{\kappa-2})$.

The bound in (i) is essentially "best possible" since exercise 1.1.14 implies that

$$\sum_{n\leq x} d_\kappa(n) \sim \kappa\frac{x}{\log x}\sum_{n\leq x}\frac{d_\kappa(n)}{n}.$$

ex:WeightL(x/n)

EXERCISE 1.2.14. Let f be a multiplicative function with each $|f(n)| \leq 1$.

(i) Show that $\sum_{n\leq x} f(n)\log\frac{x}{n} = \int_1^x \frac{1}{t}\sum_{n\leq t} f(n)\ dt$.

(ii) Deduce, using Proposition 1.2.4, that

$$\sum_{n\leq x} f(n)\log\frac{x}{n} - \sum_{n\leq x} f(n) \ll \frac{x}{\log x}\exp\Big(\sum_{p\leq x}\frac{|1-f(p)|}{p}\Big).$$

ex:f(n)f(n+1)

EXERCISE 1.2.15. Suppose that f and g are multiplicative functions with each $|f(n)|, |g(n)| \leq 1$. Define $\mathcal{P}_p(f) := \Big(1-\frac{1}{p}\Big)\Big(1+\frac{f(p)}{p}+\frac{f(p^2)}{p^2}+\dots\Big)$, and then $\mathcal{P}_p(f,g) = \mathcal{P}_p(f)+\mathcal{P}_p(g)-1$. Finally let $\mathcal{P}(f,g) = \prod_p \mathcal{P}_p(f,g)$. Prove that if $\sum_p \frac{|1-f(p)|}{p},\ \sum_p\frac{|1-g(p)|}{p} < \infty$ then

$$\lim_{x\to\infty}\frac{1}{x}\sum_{n\leq x} f(n)g(n+1) = \mathcal{P}(f,g).$$

CHAPTER 1.3

Integers without large prime factors

ch:smooths

1.3.1. "Smooth" or "friable" numbers

Let $p(n)$ and $P(n)$ be the smallest and largest prime factors of n, respectively. Given a real number $y \geq 2$, the integers, n, all of whose prime factors are at most y (that is, for which $P(n) \leq y$) are called "y-smooth" or "y-friable".[1] Smooth numbers appear all over analytic number theory. For example most factoring algorithms search for smooth numbers (in an intermediate step) which appear in a certain way, since they are relatively easy to factor. Moreover all smooth numbers n may be factored as ab, where $a \in (A/y, A]$ for any given A, $1 \leq A \leq n$. This "well-factorability" is useful in attacking Waring's problem and in finding small gaps between consecutive primes (see chapter ch:MaynardTao ??). However, counting the y-smooth numbers up to x can be surprisingly tricky. Define

$$\Psi(x,y) := \sum_{\substack{n\leq x\\ P(n)\leq y}} 1.$$

We can formulate this as a question about multiplicative functions by considering the multiplicative function given by $f(p^k) = 1$ if $p \leq y$, and $f(p^k) = 0$ otherwise.

If $x \leq y$ then clearly $\Psi(x,y) = [x] = x + O(1)$. Next suppose that $y \leq x \leq y^2$. If $n \leq x$ is not y-smooth then it must be divisible by a unique prime $p \in (y, x]$. Thus, by exercise exmertens 1.1.10(i),

$$\begin{aligned}\Psi(x,y) &= [x] - \sum_{y<p\leq x}\sum_{\substack{n\leq x\\ p|n}} 1 = x + O(1) - \sum_{y<p\leq x}\Big(\frac{x}{p} + O(1)\Big)\\ &= x\Big(1 - \log\frac{\log x}{\log y}\Big) + O\Big(\frac{x}{\log y}\Big).\end{aligned}$$

This formula tempts one to write $x = y^u$, and then, for $1 \leq u \leq 2$, we obtain

$$\Psi(y^u, y) = y^u(1 - \log u) + O\Big(\frac{y^u}{\log y}\Big).$$

We can continue the process begun above, using the principle of inclusion and exclusion to evaluate $\Psi(y^u, y)$ by subtracting from $[y^u]$ the number of integers which are divisible by a prime larger than y, adding back the contribution from integers divisible by two primes larger than y, and so on.[2] The estimate for $\Psi(y^u, y)$ involves the Dickman-de Bruijn function $\rho(u)$ defined as follows:

---

[1] "Friable" is French (and also appears in the O.E.D.) for "crumbly". Its usage, in this context, is spreading, because the word "smooth" is already overused in mathematics.

[2] A result of this type for small values of u may be found in Ramanujan's unpublished manuscripts (collected in *The last notebook*), but the first published uniform results on this problem are due to Dickman and de Bruijn.

For $0 \le u \le 1$ let $\rho(u) = 1$, and let $\rho(u) = 1 - \log u$ for $1 \le u \le 2$. For $u > 1$ we define ρ by means of the differential-difference equation

$$u\rho'(u) = -\rho(u-1);$$

indeed there is a unique continuous solution given by the (equivalent) integral (delay) equation

$$u\rho(u) = \int_{u-1}^{u} \rho(t)dt.$$

The integral equation implies (by induction) that $\rho(u) > 0$ for all $u \ge 0$, and then the differential equation implies that $\rho'(u) < 0$ for all $u \ge 1$, so that $\rho(u)$ is decreasing in this range. The integral equation implies that $u\rho(u) \le \rho(u-1)$, and iterating this we find that $\rho(u) \le 1/[u]!$.

smooth THEOREM 1.3.1. *Uniformly for all $u \ge 1$ we have*

$$\Psi(y^u, y) = \rho(u)y^u + O\Big(\frac{y^u}{\log y} + 1\Big).$$

In other words, if we fix $u > 1$ then the proportion of the integers $\le x$ that have all of their prime factors $\le x^{1/u}$, tends to $\rho(u)$, as $x \to \infty$.

PROOF. Let $x = y^u$, and we start with

$$\Psi(x,y)\log x = \sum_{\substack{n \le x \\ P(n) \le y}} \log n + O\Big(\sum_{n \le x} \log(x/n)\Big) = \sum_{\substack{n \le x \\ P(n) \le y}} \log n + O(x).$$

Using $\log n = \sum_{d|n} \Lambda(d)$ we have

$$\sum_{\substack{n \le x \\ P(n) \le y}} \log n = \sum_{\substack{d \le x \\ P(d) \le y}} \Lambda(d)\Psi(x/d, y) = \sum_{p \le y} (\log p)\Psi(x/p, y) + O(x),$$

since the contribution of prime powers p^k (with $k \ge 2$) is easily seen to be $O(x)$. Thus

E2.10 (1.3.1)
$$\Psi(x,y)\log x = \sum_{p \le y} \log p \ \Psi\Big(\frac{x}{p}, y\Big) + O(x).$$

(Compare this to the formulae in Exercise MobPNT 1.1.16.)

Now we show that a similar equation is satisfied by what we think approximates $\Psi(x,y)$, namely $x\rho(u)$. Put $E(t) = \sum_{p \le t} \frac{\log p}{p} - \log t$ so that $E(t) = O(1)$ by (Pavg 1.1.10). Now

$$\sum_{p \le y} \frac{\log p}{p} \rho\Big(\frac{\log(x/p)}{\log y}\Big) = \int_1^y \rho\Big(u - \frac{\log t}{\log y}\Big) d(\log t + E(t)),$$

and making a change of variables $t = y^\nu$ we find that

$$\int_1^y \rho\Big(u - \frac{\log t}{\log y}\Big) d(\log t) = (\log y)\int_0^1 \rho(u-\nu)d\nu = (\log x)\rho(u).$$

Moreover, since $E(t) \ll 1$ and ρ is monotone decreasing, integration by parts gives

$$\int_1^y \rho\Big(u - \frac{\log t}{\log y}\Big) d(E(t)) \ll \rho(u-1) + \int_1^y \Big|\frac{d}{dt}\rho\Big(u - \frac{\log t}{\log y}\Big)\Big| dt \ll \rho(u-1).$$

Thus we find that

E2.11
$$(x\rho(u))\log x=\sum_{p\le y}\log p\Big(\frac{x}{p}\rho\Big(\frac{\log(x/p)}{\log y}\Big)\Big)+O(\rho(u-1)x). \tag{1.3.2}$$

Subtracting (1.3.2) from (1.3.1) we arrive at

E2.12
$$|\Psi(x,y)-x\rho(u)|\log x\le\sum_{p\le y}\log p\Big|\Psi\Big(\frac{x}{p},y\Big)-\frac{x}{p}\rho\Big(\frac{\log x/p}{\log y}\Big)\Big|+Cx, \tag{1.3.3}$$

for a suitable constant C.

Suppose that the Theorem has been proved for $\Psi(z,y)$ for all $z\le x/2$, and we now wish to establish it for x. We may suppose that $x\ge y^2$, and our induction hypothesis is that for all $t\le x/2$ we have

$$\Big|\Psi(t,y)-t\rho\Big(\frac{\log t}{\log y}\Big)\Big|\le C_1\Big(\frac{t}{\log y}+1\Big),$$

for a suitable constant C_1. From (1.3.3) we obtain that

$$|\Psi(x,y)-x\rho(u)|\log x\le C_1\sum_{p\le y}\log p\Big(\frac{x}{p\log y}+1\Big)+Cx\le C_1x+O\Big(\frac{x}{\log y}+y\Big)+Cx.$$

Assuming, as we may, that $C_1\ge 2C$ and that y is sufficiently large, the right hand side above is $\le 2C_1x$, and we conclude that $|\Psi(x,y)-x\rho(u)|\le C_1x/\log y$ as $u\ge 2$. This completes our proof. $\Box$

Now $\rho(u)\le 1/[u]!$ decreases very rapidly. Therefore the main term in Theorem 1.3.1 dominates the remainder term only in the narrow range when $u^u\ll\log y$. However the asymptotic $\Psi(y^u,y)\sim\rho(u)y^u$ has been established in a much wider range than in Theorem 1.3.1 by Hildebrand [?],[3] who showed that

HildPsi
$$\Psi(y^u,y)=\rho(u)y^u\left\{1+O\left(\frac{\log(u+1)}{\log y}\right)\right\} \tag{1.3.4}$$

for $y\ge\exp((\log\log x)^2)$ where $x=y^u$. This is an extraordinarily wide range, given that Hildebrand also showed that this asymptotic holds in the only slightly larger range $y\ge(\log x)^{2+o(1)}$ if and only if the Riemann Hypothesis is true.

One can prove Theorem 1.3.1 in a number of ways. The key to the proof that we gave is the identity (1.3.1), but there are other identities that one can use. Indeed few are more elegant than de Bruijn's identity:

EDeB
$$\Psi(x,y)=[x]-\sum_{y<p\le x}\Psi\Big(\frac{x}{p},p\Big)+O(x). \tag{1.3.5}$$

However, this works out less successfully than (1.3.1), perhaps because only the X-variable in $\Psi(X,Y)$ varies in (1.3.1), whereas both variables vary in (1.3.5).

How does the result in Theorem 1.3.1 compare to the heuristic of chapter 1.2? If $f(p^k)=1$ if prime $p\le y$ and $f(p^k)=0$ otherwise then $\Psi(x,y)=\sum_{n\le x}f(n)$. The heuristic of chapter 1.2 then proposes the asymptotic $x\prod_{y<p\le x}(1-\frac{1}{p})\sim x/u$ by Mertens' Theorem. This is far larger than the actual asymptotic $\sim x\rho(u)$ of Theorem 1.3.1, since $\rho(u)\le 1/[u]!$ (and a more precise estimate is given in exercise

[3]Hildebrand's proof uses a strong form of the prime number theorem, which we wish to avoid, since one of our goals is provide a different, independent proof of a strong prime number theorem.

1.3.6). Hence, removing the multiples of the small primes leaves far fewer integers than the heuristic suggests.

1.3.2. Rankin's trick and beyond, with applications

Good upper bounds may be obtained for $\Psi(x,y)$, uniformly in a wide range, by a simple application of Rankin's trick (recall Exercise 1.2.2). Below we shall write

$$\zeta(s,y)=\prod_{p\le y}\left(1-\frac{1}{p^s}\right)^{-1}=\sum_{\substack{n\ge 1\\ P(n)\le y}} n^{-s},$$

where the product and the series are both absolutely convergent in the half-plane $\mathrm{Re}(s)>0$.

ex2.6 EXERCISE 1.3.1. * (i) Show that, for any real numbers $x\ge 1$ and $y\ge 2$, the function $x^\sigma\zeta(\sigma,y)$ for $\sigma\in(0,\infty)$ attains its minimum when $\sigma=\alpha=\alpha(x,y)$ satisfying

$$\log x=\sum_{p\le y}\frac{\log p}{p^\alpha-1}.$$

(ii) Use Rankin's trick (see Exercise 1.2.2) to show that

$$\Psi(x,y)\le\sum_{\substack{n\ge 1\\ P(n)\le y}}\min\left\{1,\frac{x}{n}\right\}\le x^\alpha\zeta(\alpha,y)=\min_{\sigma>0}x^\sigma\zeta(\sigma,y).$$

(iii) Establish a wide range in which

$$\sum_{\substack{n\ge 1\\ P(n)\le y}}\min\left\{1,\frac{x}{n}\right\}\sim x\log y\cdot\int_u^\infty\rho(t)dt.$$

By a more sophisticated argument, using the saddle point method, Hildebrand and Tenenbaum [?] established an asymptotic formula for $\Psi(x,y)$ uniformly in $x\ge y\ge 2$:

HilTen (1.3.6)
$$\Psi(x,y)=\frac{x^\alpha\zeta(\alpha,y)}{\alpha\sqrt{2\pi\phi_2(\alpha,y)}}\Big(1+O\Big(\frac{1}{u}\Big)+O\Big(\frac{\log y}{y}\Big)\Big),$$

with α as in Exercise 1.3.1(i), $\phi(s,y)=\log\zeta(s,y)$ and $\phi_2(s,y)=\frac{d^2}{ds^2}\phi(s,y)$. This work implies that if $y\ge(\log x)^{1+\delta}$ then the (easy) upper bound obtained in Exercise 1.3.1(ii) is larger than $\Psi(x,y)$ by a factor of about $\sqrt{u}\log y$, that is $\Psi(x,y)\asymp x^\alpha\zeta(\alpha,y)/(\sqrt{u}\log y)$. However, in Exercise 1.3.1(ii), we saw that Rankin's method really gives an upper bound on $\min\{1,\frac{x}{n}\}$, summed over *all* y-smooth n. The result of Exercise 1.3.1(ii) then implies that the upper bound is too large by a factor of only $\asymp\sqrt{u}\log u$.

We now improve Rankin's upper bound, yielding an upper bound for $\Psi(x,y)$ which is also too large by a factor of only $\asymp\sqrt{u}\log u$.

Rankin2 PROPOSITION 1.3.2. *Let $x\ge y\ge 3$ be real numbers. There is an absolute constant C such that for any $0<\sigma\le 1$ we have*

$$\Psi(x,y)\le C\frac{y^{1-\sigma}}{\sigma\log x}x^\sigma\zeta(\sigma,y).$$

PROOF. We consider

$$\sum_{\substack{n\le x\\ P(n)\le y}} \log n = \sum_{\substack{n\le x\\ P(n)\le y}} \sum_{n=dm} \Lambda(d) = \sum_{\substack{m\le x\\ P(m)\le y}} \sum_{\substack{d\le x/m\\ P(d)\le y}} \Lambda(d).$$

The inner sum over d is

$$\sum_{p\le \min(y,x/m)} \log p\Big[\frac{\log(x/m)}{\log p}\Big] \le \sum_{p\le \min(y,x/m)} \log(x/m) = \min(\pi(y),\pi(x/m))\log(x/m),$$

and so we find that

$$\begin{aligned}\Psi(x,y)\log x &= \sum_{\substack{n\le x\\ P(n)\le y}} \Big(\log n + \log(x/n)\Big)\\ &\le \sum_{\substack{n\le x\\ P(n)\le y}} \Big(\min(\pi(y),\pi(x/n))+1\Big)\log(x/n).\end{aligned}$$

We now use the Chebyshev bound $\pi(x) \ll x/\log x$ (see (Cheb1 1.1.7)), together with the observation that for any $0<\sigma\le 1$ and $n\le x$ we have

$$\frac{y^{1-\sigma}(x/n)^\sigma}{\sigma} \ge \begin{cases} x/n & \text{if } x/y\le n\le x\\ y\log(x/n)/\log y & \text{if } n\le x/y.\end{cases}$$

Thus we obtain that

$$\Psi(x,y)\log x \ll \sum_{\substack{n\le x\\ P(n)\le y}} \frac{y^{1-\sigma}}{\sigma}\Big(\frac{x}{n}\Big)^\sigma \le \frac{y^{1-\sigma}}{\sigma} x^\sigma \zeta(\sigma,y),$$

as desired. □

1.3.3. Large gaps between primes

We now apply our estimates for smooth numbers to construct large gaps between primes. The gaps between primes get arbitrarily large since each of $m!+2, m!+3, \dots, m!+m$ are composite, so if p is the largest prime $\le m!+1$, and q the next prime, then $q-p\ge m$. Note that $m \sim \log p/(\log\log p)$ by Stirling's formula (Exercise ex:stirling 1.1.5), whereas we expect, from (PNT 1.1.1), gaps as large as $\log p$. Can such techniques establish that there are gaps between primes that are substantially larger than $\log p$ (and substantially smaller)? That is, if $p_1=2<p_2=3<\dots$ is the sequence of prime numbers then

LargePrimeGaps

$$\limsup_{n\to\infty} \frac{p_{n+1}-p_n}{\log p_n} = \infty. \tag{1.3.7}$$

In section ch:MaynardTao ?? we will return to such questions and prove that

SmallPrimeGaps

$$\liminf_{n\to\infty} \frac{p_{n+1}-p_n}{\log p_n} = 0. \tag{1.3.8}$$

long gaps

THEOREM 1.3.3. *There are arbitrarily large p_n for which*

$$p_{n+1}-p_n \ge \frac{1}{2}\log p_n \frac{(\log\log p_n)\log\log\log\log p_n}{(\log\log\log p_n)^2}.$$

In particular (LargePrimeGaps 1.3.7) holds.

PROOF. The idea is to construct a long sequence of integers, each of which is known to be composite since it divisible by a small prime. Let $m = \prod_{p \le z} p$. Our goal is to show that there exists an interval $(T, T+x]$ for which $(T+j, m) > 1$ for each $j, 1 \le j \le x$, with $T > z$ (so that every element of the interval is composite). Erdős formulated an easy way to think about this problem

The Erdős shift: There exists an integer T for which $(T+j, m) > 1$ for each $j, 1 \le j \le x$ if and only if for every prime $p|m$ there exists a residue class $a_p \pmod p$ such that for each $j, 1 \le j \le x$ there exists a prime $p|m$ for which $j \equiv a_p \pmod p$.

PROOF OF THE ERDŐS SHIFT. Given T, let each $a_p = -T$, since if $(T+j, m) > 1$ then there exists a prime $p|m$ with $p|T+j$ and so $j \equiv -T \equiv a_p \pmod p$. In the other direction select $T \equiv -a_p \pmod p$ for each $p|m$, using the Chinese Remainder Theorem, and so if $j \equiv a_p \pmod p$ then $T + j \equiv (-a_p) + a_p \equiv 0 \pmod p$ and so $p|(T+j, m)$. □

The y-smooth integers up to x, can be viewed as the set of integers up to x, with the integers in the residue classes $0 \pmod p$ *sieved out*, by each prime p in the range $y < p \le x$. The proportion of the integers that are left unsieved is $\rho(u)$ (as we proved above), which is roughly $1/u^u$. This is far smaller than the proportion suggested by the usual heuristic:[4]

$$\prod_{y<p\le x}\left(1 - \frac{1}{p}\right) \sim \frac{\log y}{\log x} = \frac{1}{u},$$

by Mertens' Theorem.

To construct as long an interval as possible in which every integer has a small prime factor, we need to sieve as efficiently as possible, and so we adapt the smooth numbers to our purpose. This is the key to the Erdős-Rankin construction (and indeed, it is for this purpose, that Rankin introduced his moment method). We will partition the primes up to x into three parts, those $\le y$, those in $(y, \epsilon z]$, and those in $(\epsilon z, z]$ where ϵ is a very small constant. We select y and z to be optimal in the proof below; good choices turn out to be

$$x = y^u \text{ with } u = (1+\epsilon)\frac{\log\log x}{\log\log\log x}; \text{ and } z = \frac{x}{\log x} \cdot \frac{(\log\log x)^2}{\log\log\log x}.$$

Notice that $y \cdot \epsilon z \ge x$, and that $\Psi(x, y) = o(x/\log x)$ by Exercise ex2.8 1.3.5.

(I) We select the congruence classes $a_p = 0 \pmod p$ for each prime $p \in (y, \epsilon z]$. Let

$$\mathcal{N}_0 := \{n \le x : \; n \not\equiv 0 \pmod p \text{ for all } p \in (y, \epsilon z]\}.$$

The integers n counted in $\mathcal{N}_0$ either have a prime factor $p > \epsilon z$ or not. If they do then we can write $n = mp$ so that $m = n/p \le x/\epsilon z \le y$ and therefore m is composed only of prime factors $\le y$. Otherwise if n does not have a large prime factor then all of its prime factors are $\le y$. By this decomposition, (Cheb1 1.1.7) and then

[4]For a randomly chosen interval, the proportion of integers removed when we sieve by the prime p is $\frac{1}{p}$; and the different primes act "independently".

Exercise exmertens 1.1.10, we have

$$\#\mathcal{N}_0 = \sum_{\epsilon z<p<x} [x/p] + \Psi(x,y) = \sum_{\epsilon z<p<x} \frac{x}{p} + O\left(\frac{x}{\log x}\right)$$
$$= x\log\left(\frac{\log x}{\log \epsilon z}\right) + O\left(\frac{x}{\log x}\right) \sim x\frac{\log\log x}{\log x}.$$

(II) Now for each consecutive prime $p_j \le y$ let

$$\mathcal{N}_j := \{n \in \mathcal{N}_0 :\ n \notin a_p \pmod p \text{ for all } p = p_1, \dots, p_j\}$$
$$= \{n \in \mathcal{N}_{j-1} :\ n \notin a_p \pmod p \text{ for } p = p_j\}.$$

We select a_p for $p = p_j$ so as to maximize $\#\{n \in \mathcal{N}_{j-1} :\ n \equiv a_p \pmod p\}$, which must be at least the average $\frac{1}{p}\ \#\mathcal{N}_{j-1}$. Hence $\#\mathcal{N}_j \le (1-\frac{1}{p_j})\#\mathcal{N}_{j-1}$, and so if p_k is the largest prime $\le y$ then, by induction, we obtain that

$$r := \#\mathcal{N}_k \le \prod_{p\le y}\left(1-\frac{1}{p}\right)\#\mathcal{N}_0 \sim \frac{e^{-\gamma}}{\log y}x\frac{\log\log x}{\log x} \sim e^{-\gamma}(1+\epsilon)\frac{z}{\log z}$$

using Mertens' Theorem. This implies that $r < \#\{p \in (\epsilon z, z]\}$ using (Cheb1 1.1.7) (which we proved there with constant $c = \log 2$), since $e^{-\gamma} < \log 2$.

(III) Let $\mathcal{N}_k = \{b_1, \dots, b_r\}$, and let $p_{\ell+1} < p_{\ell+2} < \dots < p_{\ell+r}$ be the r smallest primes in $(\epsilon z, z]$. Now let $a_p = b_j$ for $p = p_{\ell+j}$ for $j =, 2, \dots, r$. Hence every integer $n \le x$ belongs to an arithmetic progression $a_p \pmod p$ for some $p \le z$.

We have now shown how to choose $a_p \pmod p$ for each $p \le z$ so that every $n \le x$ belongs to at least one of these arithmetic progressions. By the Erdős shift we know that there exists $T \pmod m$, where $m = \prod_{p\le z} p$ for which $(T+j, m) > 1$ for $1 \le j \le x$. We select $T \in (m, 2m]$ to guarantee that every element of the interval $(T, T+x]$ is greater than any of the prime factors of m. Hence if p_n is the largest prime $\le T$, then $p_{n+1} - p_n > x$.

We need to determine how big this gap is compared to the size of the primes involved. Now $p_n \le 2m$ and $\log m \le \psi(z) \le z\log 4 + O(\log z)$ by (Cheb1 1.1.7), so that $z \ge \frac{2}{3}\log p_n$. This implies the theorem. $\square$

EXERCISE 1.3.2. * Assuming the prime number theorem, improve the constant $\frac{1}{2}$ in this lower bound to $e^{\gamma} + o(1)$. [5]

The Erdős shift for arithmetic progressions: It is not difficult to modify the above argument to obtain large gaps between primes in any given arithmetic progression. However there is a direct connection between strings of consecutive composite numbers, and strings of consecutive composite numbers in an arithmetic progression: Let m be the product of a finite set of primes that do not divide q. Select integer r for which $qr \equiv 1 \pmod m$. Hence

$$(a+jq, m) = (ar + jqr, m) = (ar + j, m),$$

and so, for $T = ar$,

ErdosShift (1.3.9) $$\#\{1 \le j \le N :\ (a+jq, m) = 1\} = \#\{1 \le j \le N :\ (T+j, m) = 1\}.$$

[5] Using additional ideas, this has recently FGKT, Mayn2 [?, ?] been improved to allow any constant $c > 0$ in place of $\frac{1}{2}$, resolving the great Paul Erdős's favourite challenge problem. We shall return to this in chapter SelbergWtsAgain ??.

In other words, the sieving problem in an arithmetic progression is equivalent to sieving an interval.

1.3.4. Additional exercises

ex:3.4

EXERCISE 1.3.3. Suppose that f is a non-negative multiplicative function, for which $f(p^k)=0$ if $p>y$, and $\sum_{d\le D}\Lambda_f(d)\ll \min\{D, y^{\frac{\log D}{\log y}}\}$ for all $D\geq 1$. Prove that

$$\sum_{n\le x} f(n)\ll \frac{y^{1-\sigma}}{\sigma\log x}x^\sigma F(\sigma)$$

for any $0<\sigma\le 1$. When is this an improvement on the bound in Exercise 1.2.2?

EXERCISE 1.3.4. Prove that if f is a non-negative arithmetic function, and $F(\sigma)$ is convergent for some $0<\sigma<1$ then

$$\sum_{n\le x} f(n)+x\sum_{n>x}\frac{f(n)}{n}\le \frac{x^\sigma}{\sigma\log x}(F(\sigma)-\sigma F'(\sigma)).$$

(Hints: Either study the coefficient of each $f(n)$; or bound $\sum_{n\le x} f(n)\log(x/n)$ by integrating by parts, using the first part of Exercise 1.2.2, and then apply the second part of Exercise 1.2.2 for $(-F')$.)

ex2.8

EXERCISE 1.3.5. † For $x=y^u$ with $y>(\log x)^2$, let $\sigma=1-\frac{\log(u\log u)}{\log y}$. If $y\le(\log x)^2$, let $\sigma=\frac12+\epsilon$.
(i) Deduce from Proposition 1.3.2 and exercise 1.2.2(ii), together with exercise 1.1.12, that there exists a constant $C>0$ such that

$$\Psi(x,y)+\prod_{p\le y}\left(1-\frac1p\right)\sum_{\substack{n>x\\ P(n)\le y}}\frac{x}{n}\ll x\left(\frac{C}{u\log u}\right)^u+x^{1/2+o(1)}.$$

(Hint: For small y, show that $\zeta(\sigma,y)\ll x^{o(1)}$.)
(ii) Suppose that f is a multiplicative function with $0\le f(n)\le 1$ for all integers n, supported only on the y-smooth integers. Prove that

$$\sum_{\substack{n>x\\ P(n)\le y}}\frac{f(n)}{n}\ll\left(\left(\frac{C}{u\log u}\right)^u+\frac{1}{x^{1/2+o(1)}}\right)\prod_{p\le y}\left(1+\frac{f(p)}{p}+\frac{f(p^2)}{p^2}+\dots\right),$$

where $x=y^u$ with $u\geq 1$. (Hint: Prove the result for totally multiplicative f, by using exercise 1.2.3 to bound $\mathcal{R}(f;\infty)-\mathcal{R}(f;x)$ in terms of the analogous sum for the characteristic function for the y-smooth integers. Then extend this result to all such f.)
(iii) Suppose now that f is a multiplicative function with $0\le f(n)\le d_\kappa(n)$ for all integers n, supported only on the y-smooth integers. State and prove a result analogous to (ii). (Hint: One replaces C by C^κ. One should treat the primes $p\le 2\kappa$ by a separate argument)

ex2.9

EXERCISE 1.3.6. Prove that

$$\rho(u)=\left(\frac{e+o(1)}{u\log u}\right)^u.$$

(Hint: Select c maximal such that $\rho(u) \gg (c/u \log u)^u$. By using the functional equation for ρ deduce that $c \geq e$. Take a similar approach for the implicit upper bound.)

ex2.10

EXERCISE 1.3.7. * A permutation $\pi \in S_n$ is *m-smooth* if its cycle decomposition contains only cycles with length at most m. Let $N(n,m)$ denote the number of m-smooth permutations in S_n. (i) Prove that

$$n\frac{N(n,m)}{n!} = \sum_{j=1}^{m} \frac{N(n-j,m)}{(n-j)!}.$$

(ii) Deduce that $N(n,m) \geq \rho(n/m)n!$ holds for all $m,\ n \geq 1$.
(iii)† Prove that there is a constant C such that for all $m,\ n \geq 1$, we have

$$\frac{N(n,m)}{n!} \leq \rho\Big(\frac{n}{m}\Big) + \frac{C}{m}.$$

(One can take $C = 1$ in this result.)
Therefore, a random permutation in S_n is n/u-smooth with probability $\to \rho(u)$ as $n \to \infty$.

CHAPTER 1.4

Selberg's sieve applied to an arithmetic progression

In order to develop the theory of mean-values of multiplicative functions, we shall need an estimate for the number of primes in short intervals. We need only an upper estimate for the number of such primes, and this can be achieved by a simple sieve method, and does not need results of the strength of the prime number theorem. We describe a beautiful method of Selberg which works well in this and many other applications. In fact, several different sieve techniques would also work; see, e.g., Friedlander and Iwaniec's *Opera de Cribro* for a thorough treatment of sieves and their many applications.

1.4.1. Selberg's sieve

C71 Let $\mathcal{I}$ be the set of integers in the interval $(x, x+y]$, that are $\equiv a \pmod q$. For a given integer P which is coprime to q, we wish to estimate the number of integers in $\mathcal{I}$ that are coprime to P; that is, the integers that remain when $\mathcal{I}$ is *sieved* (or *sifted*) by the primes dividing P. Selberg's sieve yields a good upper bound for this quantity. Note that this quantity, plus the number of primes dividing P, is an upper bound for the number of primes in $\mathcal{I}$; selecting P well will give us the Brun-Titchmarsh theorem. When P is the product of the primes $\le x^{1/u}$, other than those that divide q, we will obtain (for suitably large u) strong upper and C73 lower bounds for the size of the sifted set; this result, which we develop in Section 1.4.2, is a simplified version of the *fundamental lemma of sieve theory*.

Let $\lambda_1 = 1$ and let λ_d be any sequence of real numbers for which $\lambda_d \neq 0$ only when $d \in S(R, P)$, which is the set of integers $d \le R$ such that d is composed entirely of primes dividing P (where R is a parameter to be chosen later). We say that λ is *supported* on $S(R, P)$. Selberg's sieve is based on the simple idea that squares of real numbers are ≥ 0, and so

$$\Big(\sum_{d|n} \lambda_d\Big)^2 \text{ is } \begin{cases} = 1 & \text{if } (n, P) = 1 \\ \ge 0 & \text{always.} \end{cases}$$

Therefore we obtain that

$$\sum_{\substack{n \in \mathcal{I} \\ (n,P)=1}} 1 \le \sum_{n \in \mathcal{I}} \Big(\sum_{d|n} \lambda_d\Big)^2.$$

Expanding out the inner sum over d, the first term on the right hand side above is

$$\sum_{d_1, d_2} \lambda_{d_1} \lambda_{d_2} \sum_{\substack{x < n \le x+y \\ n \equiv a \pmod q \\ [d_1, d_2] | n}} 1,$$

where $[d_1, d_2]$ denotes the l.c.m. of d_1 and d_2. Since P is coprime to q, we have $\lambda_d = 0$ whenever $(d, q) \neq 1$. Therefore the inner sum over n above is over one congruence class $\pmod{q[d_1, d_2]}$, and so within 1 of $y/(q[d_1, d_2])$. We conclude that

E3.1

$$\begin{aligned}\sum_{\substack{n\in\mathcal{I}\\(n,P)=1}} 1 &\leq \frac{y}{q}\sum_{d_1,d_2}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} + \sum_{d_1,d_2}|\lambda_{d_1}\lambda_{d_2}| \\ &= \frac{y}{q}\sum_{d_1,d_2}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} + \Big(\sum_d |\lambda_d|\Big)^2 \end{aligned} \tag{1.4.1}$$

The second term here is obtained from the accumulated errors obtained when we estimated the number of elements of $\mathcal{I}$ in given congruence classes. In order that each error is small compared to the main term, we need that 1 is small compared to $y/(q[d_1, d_2])$, that is $[d_1, d_2]$ should be small compared to y/q. Now if d_1, d_2 are coprime and close to R then this forces the restriction that $R \ll \sqrt{y/q}$.

The first term in (1.4.1) is a quadratic form in the variables λ_d, which we wish to minimize subject to the linear constraint $\lambda_1 = 1$. Selberg made the remarkable observation that this quadratic form can be elegantly diagonalized, which allowed him to determine the optimal choices for the λ_d: Since $[d_1, d_2] = d_1 d_2/(d_1, d_2)$, and $(d_1, d_2) = \sum_{\ell|(d_1,d_2)} \phi(\ell)$ we have

E3.2

$$\sum_{d_1,d_2}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} = \sum_\ell \phi(\ell)\sum_{\substack{\ell|d_1\\ \ell|d_2}}\frac{\lambda_{d_1}}{d_1}\frac{\lambda_{d_2}}{d_2} = \sum_\ell\frac{\phi(\ell)}{\ell^2}\Big(\sum_d\frac{\lambda_{d\ell}}{d}\Big)^2 = \sum_\ell\frac{\phi(\ell)}{\ell^2}\xi_\ell^2, \tag{1.4.2}$$

where each

$$\xi_\ell = \sum_d \frac{\lambda_{d\ell}}{d}.$$

So we have diagonalized the quadratic form. Note that if $\xi_\ell \neq 0$ then $\ell \in S(R, P)$, just like the λ_d's.

We claim that (1.4.2) provides the desired diagonalization of the quadratic form. To prove this, we must show that this change of variables is invertible, which is not difficult using the fact that $\mu * 1 = \delta$. Thus

$$\lambda_d = \sum_\ell \frac{\lambda_{d\ell}}{\ell}\sum_{r|\ell}\mu(r) = \sum_r \mu(r)\sum_{r|\ell}\frac{\lambda_{d\ell}}{\ell} = \sum_r\frac{\mu(r)}{r}\xi_{dr}.$$

In particular, the constraint $\lambda_1 = 1$ becomes

E3.3

$$1 = \sum_r \frac{\mu(r)}{r}\xi_r. \tag{1.4.3}$$

We have transformed our problem to minimizing the diagonal quadratic form in (1.4.2) subject to the constraint in (1.4.3). Calculus reveals that the optimal choice is when ξ_r is proportional to $\mu(r)r/\phi(r)$ for each $r \in S(R, P)$ (and 0 otherwise). The constant of proportionality can be determined from (1.4.3) and we conclude that the optimal choice is to take (for $r \in S(R, P)$)

E3.4

$$\xi_r = \frac{1}{L(R;P)}\frac{r\mu(r)}{\phi(r)} \qquad \text{where} \qquad L(R;P) := \sum_{\substack{r\leq R\\ p|r \implies p|P}}\frac{\mu(r)^2}{\phi(r)}. \tag{1.4.4}$$

For this choice, the quadratic form in (1.4.2) attains its minimum value, which is $1/L(R;P)$. Note also that for this choice of ξ, we have (for $d \in S(R,P)$)

$$\lambda_d = \frac{1}{L(R;P)} \sum_{\substack{r \le R/d \\ p|r \implies p|P}} \frac{d\mu(r)\mu(dr)}{\phi(dr)},$$

and so

E3.5

$$\sum_{d\le R} |\lambda_d| \le \frac{1}{L(R;P)} \sum_{\substack{d,r \\ dr \le R \\ p|dr \implies p|P}} \frac{\mu(dr)^2 d}{\phi(dr)} = \frac{1}{L(R;P)} \sum_{\substack{n \le R \\ p|n \implies p|P}} \frac{\mu(n)^2 \sigma(n)}{\phi(n)}, \tag{1.4.5}$$

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

Putting these observations together, we arrive at the following Theorem.

thm7.1

Theorem 1.4.1. *Suppose that $(P,q) = 1$. The number of integers from the interval $[x, x+y]$ that are in the arithmetic progression $a \pmod q$, and are coprime to P, is bounded above by*

$$\frac{y}{qL(R;P)} + \frac{1}{L(R;P)^2} \Big(\sum_{\substack{n \le R \\ p|n \implies p|P}} \frac{\mu(n)^2 \sigma(n)}{\phi(n)} \Big)^2$$

for any given $R \ge 1$, where $L(R;P)$ is as in (1.4.4).

C73

1.4.2. The Fundamental Lemma of Sieve Theory

We will need estimates for the number of integers in an interval of an arithmetic progression that are left unsieved by a subset of the primes up to some bound. Sieve theory provides a strong estimate for this quantity, and indeed the *fundamental Lemma of sieve theory* provides an extraordinarily precise answer for a big generalization of this question. Given our limited needs we will provide a self-contained proof, though note that it is somewhat weaker than what follows from the strongest known versions of the fundamental lemma.

FLS

Theorem 1.4.2 (The Fundamental Lemma of Sieve Theory). *Let P be an integer with $(P,q) = 1$, such that every prime factor of P is $\le (y/q)^{1/u}$ for some given $u \ge 1$. Then, uniformly, we have*

$$\sum_{\substack{x < n \le x+y \\ (n,P)=1 \\ n \equiv a \pmod q}} 1 = \frac{y}{q}\frac{\phi(P)}{P}\Big(1 + O(u^{-u/2})\Big) + O\Big(\Big(\frac{y}{q}\Big)^{3/4+o(1)}\Big).$$

As mentioned already, one can obtain stronger results by other methods. In particular, the error terms above may be improved to $O(u^{-u})$ in place of $O(u^{-u/2})$, and $O((y/q)^{1/2+o(1)})$ in place of $O((y/q)^{3/4+o(1)})$.

We will obtain the upper bound of the Fundamental Lemma by directly applying Theorem 1.4.1 and using our understanding of multiplicative functions to evaluate the various terms there.

We will deduce the lower bound from the upper bound, via a *sieve identity*, which is a technique that often works in sieve theory. We have already seen sieve identities in the previous chapter (e.g. (1.3.1) and (1.3.5)), and they are often used to turn upper bounds into lower bounds. In this case we wish to count the number of integers in a given set $\mathcal{I}$ that are coprime to a given integer P. We begin by writing $P = p_1 \cdots p_k$ with $p_1 < p_2 < \ldots < p_k$, and $P_j = \prod_{i=1}^{j-1} p_i$ for each $j > 1$, with P_1 being interpreted as 1. Since every element in I is either coprime to P, or its common factor with P has a smallest prime factor p_j for some j, we have

SieveId1
$$\#\{n \in \mathcal{I} : (n, P) = 1\} = \#\mathcal{I} - \sum_{j=1}^{k} \#\{n \in \mathcal{I} : p_j | n \text{ and } (n, P_j) = 1\}. \tag{1.4.6}$$

Good upper bounds on each $\#\{n \in \mathcal{I} : p_j|n \text{ and } (n, P_j) = 1\}$ will therefore yield a good lower bound on $\#\{n \in \mathcal{I} : (n, P) = 1\}$.

Proof. We again let $\mathcal{I} := \{n \in (x, x + y] : \ n \equiv a \pmod q\}$. We prove the upper bound using Theorem 1.4.1 with $R = \sqrt{y/q}$. Therefore if $p|P$ then $p \leq y := R^{2/u}$, and so

$$\begin{aligned} L(R; P) &= \sum_{\substack{r \geq 1 \\ p|r \implies p|P}} \frac{\mu(r)^2}{\phi(r)} + O\Big(\sum_{\substack{r > R \\ p|r \implies p|P}} \frac{\mu(r)^2}{\phi(r)} \Big) \\ &= \frac{P}{\phi(P)} \Big\{ 1 + O\Big(\Big(\frac{C}{u \log u} \Big)^{u/2} + \frac{1}{R^{1/2+o(1)}} \Big) \Big\} \end{aligned}$$

by exercise 1.3.5(iii) with $\kappa = 2$ for the error term. Moreover, by the Cauchy-Schwarz inequality, and then exercises 1.2.10 and 1.3.5(i), we have

$$\Big(\sum_{\substack{n \leq R \\ p|n \implies p|P}} \frac{\mu(n)^2 \sigma(n)}{\phi(n)} \Big)^2 \leq \Big(\sum_{n \leq R} \frac{\sigma(n)^2}{\phi(n)^2} \Big) \Psi(R, R^{2/u}) \ll R^2 \Big(\frac{C}{u \log u} \Big)^{u/2}.$$

Inserting these estimates into the bound of Theorem 1.4.1, yields the upper bound

$$\sum_{\substack{n \in \mathcal{I} \\ (n,P)=1}} 1 \leq \frac{y}{q} \frac{\phi(P)}{P} \Big(1 + O\Big(\Big(\frac{C}{u \log u} \Big)^{u/2} \Big) \Big) + O\Big(\Big(\frac{y}{q} \Big)^{3/4+o(1)} \Big),$$

which implies the upper bound claimed, with improved error terms.

We now prove the lower bound using (1.4.6), and that $\#\mathcal{I} = y/q + O(1)$. The upper bound that we just proved implies that

$$\begin{aligned} \sum_{\substack{n \in \mathcal{I} \\ p_j | n \\ (n, P_j) = 1}} 1 &= \sum_{\substack{x/p_j < n \leq (x+y)/p_j \\ (n, P_j) = 1 \\ n \equiv a p_j^{-1} \pmod q}} 1 \\ &\leq \frac{y}{q p_j} \frac{\phi(P_j)}{P_j} \Big(1 + O\Big(\Big(\frac{C}{u_j \log u_j} \Big)^{u_j/2} \Big) \Big) + O\Big(\Big(\frac{y}{q p_j} \Big)^{3/4+o(1)} \Big), \end{aligned}$$

where $u_j := \log(y/qp_j)/\log p_j$. Inserting this into (1.4.6), for the main term we have

$$1 - \sum_{j=1}^{k} \frac{1}{p_j} \frac{\phi(P_j)}{P_j} = \frac{\phi(P)}{P}.$$

Since the second error term is larger than the first only when $u \to \infty$, hence when we sum over all p_j, the second error term remains $\ll (y/q)^{3/4+o(1)}$. For the first error term we begin by noting that $u_j = \log(y/q)/\log p_j - 1 \geq u - 1$ and so

$(u_j/u)^2(C/u_j \log u_j)^{u_j/2} \ll (C'/u \log u)^{u/2}$ for some constant $C' > 0$. We deal with the sum over j by then noting that $\phi(P_j)/P_j \ll (u_j/u)\phi(P)/P$ and so

$$\sum_{j=1}^{k} \frac{1}{p_j} \frac{\phi(P_j)}{P_j} \frac{u^2}{u_j^2} \ll \frac{\phi(P)}{P} \frac{u}{\log(y/q)} \sum_{p \le (y/q)^{1/u}} \frac{\log p}{p} \ll \frac{\phi(P)}{P},$$

by (1.1.10). This completes our proof. $\square$

FLS3

Exercise 1.4.1. * Suppose that y, z, q are integers for which $\log q \le z \le y/q$, and let $m = \prod_{p \le z} p$. Use the Fundamental Lemma of Sieve Theory to prove that if $(a, q) = 1$ then

$$\sum_{\substack{x < n \le x+y \\ n \equiv a \pmod q \\ (n,m)=1}} 1 \ll \frac{y}{\phi(q) \log z}.$$

Taking the special case here with $z = (y/q)^{1/2}$, and trivially bounding the number of primes $\le z$ that are $\equiv a \pmod q$, we deduce the most interesting corollary to Theorem 1.4.2:

BTfromFLS

Corollary 1.4.3 (The Brun-Titchmarsh Theorem). *Let $\pi(x; q, a)$ denote the number of primes $p \le x$ with $p \equiv a \pmod q$. There exists a constant $\kappa > 0$ such that*

$$\pi(x + y; q, a) - \pi(x; q, a) \le \frac{\kappa y}{\phi(q) \log(y/q)}.$$

1.4.3. A stronger Brun-Titchmarsh Theorem

We have just seen that sieve methods can give an upper bound for the number of primes in an interval $(x, x + y]$ that belong to the arithmetic progression $a \pmod q$. The smallest explicit constant κ known for Corollary 1.4.3 is $\kappa = 2$, due to Montgomery and Vaughan, which we prove in this section using the Selberg sieve:

BTstrong

Theorem 1.4.4. *There is a constant $C > 1$ such that if $y/q \ge C$ then*

OurBT

$$\pi(x + y; q, a) - \pi(x; q, a) \le \frac{2y}{\phi(q) \log(y/q)}, \tag{1.4.7}$$

for any arithmetic progression $a \pmod q$ with $(a, q) = 1$.

Since $\pi(x + y; q, a) - \pi(x; q, a) \le y/q + 1$, we deduce (1.4.7) for $q \le y \le q \exp(q/\phi(q))$.

One can considerably simplify proofs in this area using *Selberg's monotonicity principle*: For given integers $\omega(p) < p$, for each prime p, and any integer N, define

$$\mathcal{S}^+(N, \{\omega(p)\}_p) := \max_{\substack{\mathcal{I} \text{ an interval} \\ \#(\mathcal{I} \cap \mathbb{Z}) = N}} \max_{\substack{\Omega(p) \subset \mathbb{Z}/p\mathbb{Z} \ \forall p \\ \#\Omega(p) = \omega(p)}} \frac{\#\{n \in \mathcal{I} : \ n \notin \Omega(p) \text{ for all primes } p\}}{\#\mathcal{I} \ \prod_p \left(1 - \frac{\omega(p)}{p}\right)}$$

where the first "max" is over all intervals containing exactly N integers, and the second "max" is over all sets $\Omega(p)$ of $\omega(p)$ residue classes mod p, for each prime p. We can analogously define $\mathcal{S}^-(N, \{\omega(p)\}_p)$ as the minimum.

Selbergmonotone

Lemma 1.4.5 (Selberg's monotonicity principle). *If $\omega_1(p) \le \omega_2(p)$ for all primes p then, for all integers $N \ge 1$,*

$$\mathcal{S}^+(N, \{\omega_2(p)\}_p) \ge \mathcal{S}^+(N, \{\omega_1(p)\}_p) \ge \mathcal{S}^-(N, \{\omega_1(p)\}_p) \ge \mathcal{S}^-(N, \{\omega_2(p)\}_p).$$

Proof. We shall establish the result when $\omega'(p)=\omega(p)$ for all primes $p\neq q$, and $\omega'(q)=\omega(q)+1$, and then the full result follows by induction. So given the sets $\{\Omega(p)\}_p$ and an interval $\mathcal{I}$, let $\mathcal{N}:=\{n\in\mathcal{I}:\ n\notin\Omega(p)$ for all primes $p\}$. Let m be the product of all primes $p\neq q$ with $\omega(p)\neq 0$, and then define $\mathcal{I}_j:=\mathcal{I}+jm$ for $j=0,1,\dots,q-1$. Define $J:=\{j\in[0,q-1]:\ -jm\notin\Omega(q)\}$ so that $\#J=q-\omega(q)$. Let $\Omega_j(p)=\Omega(p)$ for all $p\neq q$ and $\Omega_j(q)=(\Omega(q)+jm)\cup\{0\}$; notice that $\#\Omega_j(q)=\#\Omega(q)+1$ whenever $j\in J$. Moreover, letting $\mathcal{N}_j:=\{n+jm\in\mathcal{I}_j:\ n+jm\notin\Omega_j(p)$ for all primes $p\}$ we have

$$\#\mathcal{N}_j=\#\mathcal{N}\setminus\#\{n\in\mathcal{N}:\ n\equiv -jm \pmod q\}.$$

We sum this equality over every $j\in J$. Notice that each $n\in\mathcal{N}$ satisfies $n\equiv -jm$ (mod q) for a unique $j\in J$, and hence $\sum_{j\in J}\#\mathcal{N}_j=(\#J-1)\#\mathcal{N}$, which implies that

$$\#\mathcal{N}\le\frac{(1-\omega(q)/q)}{(1-\omega'(q)/q)}\ \max_{j\in J}\#\mathcal{N}_j;$$

and therefore $\mathcal{S}^+(N,\{\omega(p)\}_p)\le\mathcal{S}^+(N,\{\omega'(p)\}_p)$. The last step can be reworked, analogously, to also yield $\mathcal{S}^-(N,\{\omega(p)\}_p)\ge\mathcal{S}^-(N,\{\omega'(p)\}_p)$. $\square$

Proof of Theorem 1.4.4. Let P be the set of primes $\le R$ so that Proposition 1.2.5 (with $\sigma=\frac34$ say) yields

$$L(R;P)\ge\log R+\gamma'+o(1)$$

where $\gamma':=\gamma+\sum_p\frac{\log p}{p(p-1)}$; and Exercise 1.2.11 gives that

$$\sum_{n\le R}\frac{\mu(n)^2\sigma(n)}{\phi(n)}=\frac{15}{\pi^2}R+o(R).$$

Inserting these estimates into Theorem 1.4.1 with $R:=\frac{\pi^2}{15}\sqrt{\frac{y}{2}}$ we deduce that

FirstSieve
$$\#\{n\in[x,x+y):\ (n,P)=1\}\le\frac{2y}{\log y+c+o(1)}\tag{1.4.8}$$

where $c:=2\gamma'-1-\log 2+2\log(\pi^2/15)=0.1346\dots$ This implies (1.4.7) for $q=1$ when $y\ge C$, for some constant C (given by when $c+o(1)>0$).

Given y and q, let $Y=y/q$ and let m be the product of the primes $\le R$ that do not divide q. Suppose that $Y\ge C$.

Let $\{a+jq:\ 1\le j\le N\}$ be the integers $\in(x,x+y]$ in the arithmetic progression $a \pmod q$ (so that $N=Y+O(1)$). By (1.3.9) we know that the number of these integers that are coprime to m, equals exactly the number of integers in some interval of length N that are coprime to m, and this is $\le\mathcal{S}^+(N,\{\omega_1(p)\}_p)$, by definition, where $\Omega_1(p)=\{0\}$ for each $p|m$ and $\Omega_1(p)=\emptyset$ otherwise. Now suppose that $\Omega_2(p)=\{0\}$ for each $p|P$ and $\Omega_2(p)=\emptyset$ otherwise, so that Selberg's monotonicity principle implies that $\mathcal{S}^+(N,\{\omega_1(p)\}_p)\le\mathcal{S}^+(N,\{\omega_2(p)\}_p)$. In other words

$$\max_x\#\{n\in(x,x+N]:\ (n,m)=1\}\le\frac{P/m}{\phi(P/m)}\cdot\max_T\#\{n\in(T,T+N]:\ (n,P)=1\},$$

and the result follows from (1.4.8) since P/m divides q.

$\square$

1.4.4. Sieving complex-valued functions

In our subsequent work we shall need estimates for

$$\sum_{\substack{n\le x\\(n,P)=1}} n^{it},$$

where t is some real number, and P is composed of primes smaller than some parameter y. It is perhaps unusual to sieve the values of a complex valued function (since the core of every sieve methods involves sharp inequalities). In this section we show that the estimates developed so far allow such a variant of the fundamental lemma.

FLit PROPOSITION 1.4.6. *Let t and y be real numbers with $y \ge 1+|t|$ and let $x = y^u$ with $u \ge 1$. Let P be an integer composed of primes smaller than y. Then*

$$\sum_{\substack{n\le x\\(n,P)=1}} n^{it} = \frac{x^{1+it}}{1+it}\frac{\phi(P)}{P} + O\Big(x\frac{\phi(P)}{P}u^{-u/2} + x^{\frac{3}{4}+\epsilon}\Big).$$

PROOF. Let λ_d be weights as in Selberg's sieve, supported on the set $S(R,P)$. Since $(\sum_{d|n}\lambda_d)^2$ is at least 1 if $(n,P)=1$ and non-negative otherwise, it follows that

FLit1
$$\sum_{\substack{n\le x\\(n,P)=1}} n^{it} = \sum_{n\le x} n^{it}\Big(\sum_{d|n}\lambda_d\Big)^2 + O\Big(\sum_{n\le x}\Big(\sum_{d|n}\lambda_d\Big)^2 - \sum_{\substack{n\le x\\(n,P)=1}} 1\Big). \tag{1.4.9}$$

The error term here is precisely that considered in the proof of Theorem 1.4.2 and so we can use the bound from there.

A straightforward argument using partial summation shows that

$$\sum_{n\le N} n^{it} = \frac{N^{1+it}}{1+it} + O((1+|t|)\log N),$$

and therefore for any d

$$\sum_{\substack{n\le N\\d|n}} n^{it} = d^{it}\sum_{m\le N/d} m^{it} = \frac{1}{d}\cdot\frac{N^{1+it}}{1+it} + O((1+|t|)\log N).$$

Therefore the main term in (1.4.9) equals

$$\sum_{d_1,d_2}\lambda_{d_1}\lambda_{d_2}\sum_{\substack{n\le x\\ [d_1,d_2]|n}} n^{it} = \frac{x^{1+it}}{1+it}\sum_{d_1,d_2}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} + O\Big((1+|t|)\log x\Big(\sum_d|\lambda_d|\Big)^2\Big).$$

We have seen the sum in the main term in (1.4.2), and that it equals $1/L(R;P)$. The error term is bounded by using (1.4.5). These can both be evaluated using the estimates proved (for this purpose) in the proof of Theorem 1.4.2. □

1.4.5. Multiplicative functions that only vary at small prime factors

C2a

The characteristic function of the integers that are coprime to P, is given by the totally multiplicative function f with $f(p)=0$ when $p|P$, and $f(p)=1$ otherwise. Hence Theorem 1.4.2 (with $x=a=0,\ q=1$) can be viewed as a mean value theorem for a certain class of multiplicative functions (those which only take values

0 and 1, and equal 1 on all primes $p > y$). We now deduce a result of this type for a wider class of multiplicative functions:

GenFundLem

PROPOSITION 1.4.7. *Suppose that* $|f(n)| \leq 1$ *for all* n, *and* $f(p^k) = 1$ *for all* $p > y$. *If* $x = y^u$ *then*

$$\frac{1}{x} \sum_{n \leq x} f(n) = \mathcal{P}(f; x) + O\Big(u^{-u/3+o(u)} + x^{-1/6+o(1)}\Big).$$

This result is weaker than desirable since if u is bounded then the first error term is bigger than the main term unless $\sum_{p \leq x} \frac{1-f(p)}{p}$ is very small. We would prefer an estimate like $\mathcal{P}(f;x)\{1 + O(u^{-c_1 u})\} + O(x^{-c_2})$ for some $c_1, c_2 > 0$. When each $f(p) = 0$ or 1 this is essentially the Fundamental lemma of the sieve (Theorem 1.4.2). However it is false, in general, as one may see in Proposition ?? and even for real-valued f, as may be seen by taking $f(p) = -1$ for all $p \leq y$ (though we only prove this later in chapter ??). We guess that one does have an estimate $\mathcal{P}(f;x)\{1 + O(u^{-cu})\} + O(\frac{1}{\log x})$, for real f with each $f(p) \in [-1, 1]$, a challenging open problem.

PROOF OF PROPOSITION 1.4.7. We may write each integer n as ab where $P(a) \leq y$, and $p|b \implies p > y$, so that $f(n) = f(a)f(b) = f(a)$, and thus

$$\sum_{n \leq x} f(n) = \sum_{\substack{a \leq x \\ P(a) \leq y}} f(a) \sum_{\substack{b \leq x/a \\ p|b \implies p > y}} 1.$$

If $a \geq x/y$ then the inner sum equals 1, as it only counts the integer 1. Otherwise we apply Theorem 1.4.2 with $P = \prod_{p \leq y} p$ (and taking there x, y, a, q as $0, x, 0, 1$, respectively). If $A = x^{1/3} < a < x/y$ then we deduce the crude upper bound $\ll x/(a \log y)$ for the inner sum, by Merten's Theorem. Finally if $a \leq x^{1/3}$ then $\log(x/a)/\log y \geq 2u/3$, giving $\frac{\phi(P)}{P}\frac{x}{a}(1 + O(u^{-u/3+o(1)})) + O((\frac{x}{a})^{3/4+o(1)})$ for the inner sum. Combining these estimates, we now have a main term of

$$\frac{\phi(P)}{P}\, x \sum_{\substack{a \geq 1 \\ P(a) \leq y}} \frac{f(a)}{a} = \mathcal{P}(f;x)\ x,$$

and an error term which is

$$\ll u^{-u/3+o(1)} x \frac{\phi(P)}{P} \sum_{\substack{a \geq 1 \\ P(a) \leq y}} \frac{1}{a} + \sum_{a \leq x^{1/3}} \left(\frac{x}{a}\right)^{3/4+o(1)} + \frac{x}{\log y} \sum_{\substack{a > x^{1/3} \\ P(a) \leq y}} \frac{1}{a} + \sum_{\substack{x/y \leq a \leq x \\ P(a) \leq y}} 1$$

$$\ll u^{-u/3+o(1)} x + x^{5/6+o(1)}$$

as desired, using exercise 1.3.5(i) to bound the last two sums. □

1.4.6. Additional exercises

EXERCISE 1.4.2. * Prove that our choice of λ_d (as in section 1.4.1) is only supported on squarefree integers d and that $0 \leq \mu(d)\lambda_d \leq 1$.

EXERCISE 1.4.3. * (i) Prove the following *reciprocity law*: If $L(d)$ and $Y(r)$ are supported only on the squarefree integers then

$$Y(r) := \mu(r) \sum_{m:\ r|m} L(m) \text{ for all } r \geq 1 \text{ if and only if } L(d) = \mu(d) \sum_{n:\ d|n} Y(n) \text{ for all } d \geq 1.$$

(ii) Deduce the relationship, given in Selberg's sieve, between the sequences λ_d/d and $\mu(r)\xi_r/r$.
(iii) Suppose that g is a multiplicative function and $f = 1 * g$. Prove that

$$\sum_{d_1,d_2\geq 1} L(d_1)L(d_2)f((d_1,d_2)) = \sum_{n\geq 1} g(n)Y(n)^2.$$

(iv) Suppose that L is supported only on squarefree integers in $S(R,P)$. Show that to maximize the expression in (iii), where each $f(p) > 1$, subject to the constraint $L(1) = 1$, we have that Y is supported only on $S(R,P)$, and then $Y(n) = c/g(n)$ where $c = \sum_n 1/g(n)$. Use this to determine the value of each $L(m)$ in terms of g.
(v) Prove that $0 \leq f(m)\mu(m)L(m) \leq 1$ for all m; and if $R = \infty$ then $L(m) = \mu(m)/f(m)$ for all $m \in S(P)$.

FLS1 Exercise 1.4.4. * Show that if $(am,q) = 1$ and all of the prime factors of m are $\leq (x/q)^{1/u}$ then

$$\sum_{\substack{n\leq x\\ (n,m)=1\\ n\equiv a \pmod q}} \log n = \frac{\phi(m)}{m}\frac{x}{q}(\log x - 1)\Big\{1 + O(u^{-u/2})\Big\} + O\Big(\Big(\frac{x}{q}\Big)^{3/4+o(1)}\log x\Big).$$

Exercise 1.4.5. † Fill in the final computational details of the proof of Theorem 1.4.4 to determine a value for C.

Exercise 1.4.6. Use Selberg's monotonicity principle, and exercise 1.2.9 with $q = \prod_{p\leq z} p$ where $z = (y/q)^{1/u}$ (and exercise 1.1.12) to prove the Fundamental Lemma of Sieve Theory in the form

$$\sum_{\substack{x<n\leq x+y\\ (n,P)=1\\ n\equiv a \pmod q}} 1 = \frac{y}{q}\frac{\phi(P)}{P} + O\Big(\frac{y}{q}\,\Big(\frac{e+o(1)}{u\log u}\Big)^u \cdot \log y\Big).$$

Exercise 1.4.7. Prove that if P is the set of all primes $\leq y$, and $0 < |t| \leq y$ then for any x we have

$$\sum_{\substack{n\leq x\\ (n,P)=1}} \frac{1}{n^{1+it}} \ll 1 + \frac{1}{|t|\log y}.$$

Exercise 1.4.8. Suppose that $f(n)$ is a multiplicative function with each $|f(n)| \leq 1$. Prove that

$$\sum_{\substack{n\leq x\\ (n,P)=1}} f(n) - \frac{\phi(P)}{P}\sum_{n\leq x} f(n)$$

$$\ll x\frac{\phi(P)}{P}u^{-u/2} + x^{\frac{3}{4}+\epsilon} + \sum_{d\leq R^2}\mu^2(d)3^{\omega(d)}\left|\sum_{\substack{n\leq x\\ d|n}} f(n) - \frac{1}{d}\sum_{n\leq x} f(n)\right|,$$

where $\omega(d)$ denotes the number of prime factors of d. (Hint: Modify the technique of Proposition 1.4.6.)

CHAPTER 1.5

The structure of mean values

MeanF(n)

We have encountered two basic types of mean values of multiplicative functions:

• In Chapter 1.2 we gave a heuristic which suggested that the mean value of f up to x, should be $\sim \mathcal{P}(f;x)$. We were able to show this when $\sum_{p\leq x}|1-f(p)|/p$ is small, and in particular in the case that $f(p)=1$ for all "large" primes, that is. for the primes $p>y$ (Proposition 1.4.7).

• In Chapter 1.3 we considered an example in which the mean value is far smaller than the heuristic, in this case $f(p)=1$ for all "large" primes, that is. for the primes $p\leq y$.

These behaviours are very different, though arise from quite different types of multiplicative functions (the first varies from 1 on the "small primes", the second on the "large primes"). In the next two sections we study the latter case in more generality, and then consider multiplicative functions which vary on both the small and large primes. The error terms in most of the results proved in this chapter will be improved later once we have established some fundamental estimates of the subject.

1.5.1. Some familiar Averages

Let f be a multiplicative function with each $|f(n)|\leq 1$, and then let

$$S(x)=\sum_{n\leq x} f(n) \quad \text{and} \quad -\frac{F'(s)}{F(s)}=\sum_{n\geq 1}\Lambda_f(n)/n^s.$$

Looking at the coefficients of $-F'(s)=F(s)\cdot(-\frac{F'(s)}{F(s)})$ we obtain that

$$f(n)\log n=\sum_{ab=n} f(a)\Lambda_f(b).$$

Summing this over all $n\leq x$, and using exercise 1.2.14(i), we deduce that

$$S(x)\log x=\sum_{n\leq x}\Lambda_f(n)S(x/n)+\int_1^x \frac{S(t)}{t}dt.$$

Now, as $|S(t)|\leq t$ the last term is $O(x)$. The terms in the sum for which n is a prime power also contribute $O(x)$, and hence

HildIdentity

$$S(x)\log x=\sum_{p\leq x} f(p)\log p\ S(x/p)+O(x). \tag{1.5.1}$$

This is a generalization of the identities in exercise 1.1.16 (i, iii), and (1.3.1).

1.5.2. Multiplicative functions that vary only the large prime factors

Our goal is to use the identity in (1.5.1) to gain an understanding of $S(x)$ in the spirit of chapter 1.3. To proceed we define functions

$$s(u) := y^{-u} S(y^u) = \frac{1}{y^u} \sum_{n \le y^u} f(n) \quad \text{and} \quad \chi(u) := \frac{1}{y^u} \sum_{m \le y^u} \Lambda_f(m).$$

Using the definitions, we now evaluate, for $x = y^u$, the integral

$$\begin{aligned} \frac{1}{u} \int_0^u s(u-t)\chi(t)dt &= \frac{1}{u} \int_0^u \frac{1}{y^{u-t}} \sum_{a \le y^{u-t}} f(a) \cdot \frac{1}{y^t} \sum_{b \le y^t} \Lambda_f(b) dt \\ &= \frac{1}{x} \sum_{ab \le x} f(a)\Lambda_f(b) \frac{1}{u} \int_{\frac{\log b}{\log y}}^{u - \frac{\log a}{\log y}} 1 dt \\ &= \frac{1}{x} \sum_{n \le x} f(n) \log n \left(1 - \frac{\log n}{\log x}\right). \end{aligned}$$

The difference between this and $\frac{1}{x}\sum_{n \le x} f(n) \log \frac{x}{n}$ is

$$\le \frac{\log x}{x} \sum_{n \le x} |f(n)| \left(1 - \frac{\log n}{\log x}\right)^2 \le \frac{\log x}{x} \sum_{n \le x} \left(1 - \frac{\log n}{\log x}\right)^2 \ll \frac{1}{\log x}.$$

Combining this with exercise 1.2.14(ii) we deduce that

s-Identity
$$s(u) = \frac{1}{u} \int_0^u s(u-t)\chi(t)dt + O\left(\frac{1}{\log x} \exp\left(\sum_{p \le x} \frac{|1 - f(p)|}{p}\right)\right). \tag{1.5.2}$$

The integral $\int_0^u g(u-t)h(t)dt$ is known as the *(integral) convolution* of g and h, and is denoted by $(g * h)(u)$.

In the particular case that $f(p^k) = 1$ for all $p \le y$ we have $S(x) = [x]$ for $x \le y$, and so $s(t) = 1 + o(y^{-t})$ for $0 \le t \le 1$. Moreover (1.5.2) becomes

s-Id2
$$s(u) = \frac{1}{u} \int_0^u s(u-t)\chi(t)dt + O\left(\frac{u}{\log y}\right). \tag{1.5.3}$$

This suggests that if we define a continuous function σ with $\sigma(t) = 1$ for $0 \le t \le 1$ and then

IntDelEqn
$$\sigma(u) = \frac{1}{u} \int_0^u \sigma(u-t)\chi(t)dt \quad \text{for all } u \ge 1, \tag{1.5.4}$$

then we must have, for $x = y^u$

IntDel1
$$\frac{1}{x} \sum_{m \le x} f(m) = \sigma(u) + O\left(\frac{\log u}{\log y}\right). \tag{1.5.5}$$

We will deduce this, later, once we have proved the prime number theorem (which is relevant since it implies that $\chi(t) = 1 + o(1)$ for $0 < t \le 1$, and $|\chi(t)| \le 1 + o(1)$ for all $t > 0$) but, for now, we observe that a result like (1.5.5) shows that the mean value of every multiplicative function which only varies on the large primes, can be determined in terms of an integral delay equation like (1.5.4). This is quite different from the mean value of multiplicative functions that only vary on the small primes, which can be determined by the Euler product $\mathcal{P}(f; x)$.

1.5.3. A first Structure Theorem

We have seen that the mean value of a multiplicative function which only varies on its small primes is determined by an Euler product, whereas the the mean value of a multiplicative function which only varies on its large primes is determined by an integral delay equation. What about multiplicative functions which vary on both? In the next result we show how the mean value of a multiplicative function can be determined as the product of the mean values of the multiplicative functions given by its value on the small primes, and by its value on the large primes.

Structure1 THEOREM 1.5.1. *Let f be a multiplicative function with $|f(n)| \le 1$ for all n. For any given y, we can write $1 * f = g * h$ where g only varies (from 1) on the primes $> y$, and h only varies on the primes $\le y$:*

$$g(p^k) = \begin{cases} 1 & \text{if } p \le y \\ f(p^k) & \text{if } p > y \end{cases} \quad \text{and} \quad h(p^k) = \begin{cases} f(p^k) & \text{if } p \le y \\ 1 & \text{if } p > y. \end{cases}$$

Then, for $x = y^u$ we have

1stStructure
$$\frac{1}{x}\sum_{n\le x} f(n) = \frac{1}{x}\sum_{n\le x} g(n)\cdot \frac{1}{x}\sum_{n\le x} h(n) + O\Big(\frac{1}{u}\ \exp\Big(\sum_{p\le x}\frac{|1-f(p)|}{p}\Big)\Big). \tag{1.5.6}$$

If u is sufficiently large (as determined by the size of $\sum_{p\le x}\frac{|1-f(p)|}{p}$) then the error term here is $o(1)$, and hence

StructAsymp
$$\frac{1}{x}\sum_{n\le x} f(n) = \frac{1}{x}\sum_{n\le x} g(n)\cdot \frac{1}{x}\sum_{n\le x} h(n) + o(1). \tag{1.5.7}$$

In Theorem StructThm ?? we will prove that (StructAsymp 1.5.7) holds whenever $u \to \infty$. This is "best possible" as will be discussed in Chapter C20 ??.

PROOF. Let $H = \mu * h$ so that $h = 1 * H$ and $f = g * H$. Therefore

$$\frac{1}{x}\sum_{n\le x} f(n) = \frac{1}{x}\sum_{ab\le x} H(a)g(b) = \sum_{a\le x}\frac{H(a)}{a}\frac{1}{x/a}\sum_{b\le x/a} g(b).$$

By Proposition FirstLip 1.2.4 this is

$$\sum_{a\le x}\frac{H(a)}{a}\cdot\frac{1}{x}\sum_{b\le x} g(b) + O\Big(\sum_{a\le x}\frac{|H(a)|}{a}\frac{\log(2a)}{\log x}\ \exp\Big(\sum_{p\le x}\frac{|1-g(p)|}{p}\Big)\Big).$$

We may extend both sums over a, to be over *all* integers $a \ge 1$ since the error term is trivially bigger than the main term when $a > x$. Now

$$\begin{aligned}\sum_{a\ge 1}\frac{|H(a)|}{a}\log a &= \sum_{a\ge 1}\frac{|H(a)|}{a}\sum_{p^k\|a} k\log p \\ &\le 2\sum_{\substack{p\le y\\ k\ge 1}}\frac{k\log p}{p^k}\sum_{A\ge 1}\frac{|H(A)|}{A} \ll \log y\cdot\exp\Big(\sum_{p\le}\frac{|H(p)|}{p}\Big),\end{aligned}$$

writing $a = p^k A$ with $(A,p) = 1$ and then extending the sum to all A, since $|H(p^k)| \le 2$. Now

$$\sum_{p\le x}\frac{|1-g(p)|+|H(p)|}{p} = \sum_{p\le x}\frac{|1-f(p)|}{p},$$

and so the error term above is acceptable. Finally we note that

$$\sum_{a\le x}\frac{H(a)}{a}=\frac{1}{x}\sum_{n\le x}h(n)+O\Big(u^{-u/3+o(u)}+x^{-1/6+o(1)}\Big)$$

by applying Proposition 1.4.7, and the result follows. $\square$

1.5.4. An upper bound on averages

For any multiplicative function f with $|f(n)|\le 1$ for all n we have $|\chi(t)|\ll 1$ for all $t>0$. We can then take absolute values in (1.5.2) to obtain the upper bound

$$|s(u)|\ll\frac{1}{u}\int_0^u|s(t)|dt+\frac{1}{\log x}\exp\Big(\sum_{p\le x}\frac{|1-f(p)|}{p}\Big).$$

In this section we will improve this upper bound using the Brun-Titchmarsh Theorem to

propHal1 (1.5.8)
$$|s(u)|\ll\frac{1}{u}\int_0^u|s(t)|dt+\frac{1}{\log x}.$$

If we could assume the prime number theorem then we could obtain this result with "$\ll$" replaced by "$\le$".

Proof of (1.5.8). Now, for $z=y+y^{1/2}+y^2/x$, using the Brun-Titchmarsh theorem,

$$\begin{aligned}\sum_{y<p\le z}\log p\,\Big|S\Big(\frac{x}{p}\Big)\Big|&\le\sum_{y<p\le z}\log p\max_{y\le u\le z}\Big|S\Big(\frac{x}{u}\Big)\Big|\ll(z-y)\max_{y\le u\le z}\Big|S\Big(\frac{x}{u}\Big)\Big|\\&\le\int_y^z\Big|S\Big(\frac{x}{t}\Big)\Big|dt+(z-y)\max_{y\le t,u\le z}\Big|S\Big(\frac{x}{t}\Big)-S\Big(\frac{x}{u}\Big)\Big|,\end{aligned}$$

and if $y\le t,u\le z$ then

$$\Big|S\Big(\frac{x}{t}\Big)-S\Big(\frac{x}{u}\Big)\Big|\le\Big|\frac{x}{y}-\frac{x}{z}\Big|=x\cdot\frac{z-y}{y^2}.$$

Summing over such intervals between y and $2y$ we obtain

$$\sum_{y<p\le 2y}\log p\,\Big|S\Big(\frac{x}{p}\Big)\Big|\ll\int_y^{2y}\Big|S\Big(\frac{x}{t}\Big)\Big|dt+\frac{x}{y^{1/2}}+y.$$

We sum this over each dyadic interval between 1 and x. By (1.5.1) this implies that

$$\begin{aligned}|S(x)|\log x&\le\sum_{p\le x}\log p\,\Big|S\Big(\frac{x}{p}\Big)\Big|+O(x)\\&\ll\int_1^x\Big|S\Big(\frac{x}{t}\Big)\Big|dt+x=x\int_1^x\frac{|S(w)|}{w^2}dw+x.\end{aligned}$$

Taking $w=x^t$ and dividing through by $x\log x$, yields (1.5.8). $\square$

By partial summation, we have

$$\begin{aligned}\sum_{n\le x}\frac{f(n)}{n}&=\frac{1}{x}\sum_{n\le x}f(n)+\int_1^x\sum_{n\le w}f(n)\frac{dw}{w^2}=\frac{S(x)}{x}+\int_1^x\frac{S(w)}{w^2}dw\\&=s(u)+\log y\int_0^u s(t)dt.\end{aligned}$$

Using (1.5.8), and that $s(t) \geq 1/2$ for $0 \leq t \leq 1/2 \log y$, we deduce the same upper bound for the logarithmic mean of f that we had for the mean of f (in (1.5.8)):

`propHal1log`

$$\frac{1}{\log x}\Big|\sum_{n\leq x}\frac{f(n)}{n}\Big| \leq \frac{1}{u}\int_0^u |s(t)|dt\left(1+O\left(\frac{1}{\log x}\right)\right). \tag{1.5.9}$$

`IterateAverages`

1.5.5. Iterating identities

In this section we develop further identities, involving multi-convolutions of multiplicative functions, which turn out to be useful. We have already seen that $f(n)\log n = \sum_{ar=n}\Lambda_f(a)f(r)$, so iterating this twice yields

$$f(n)\log n - \Lambda_f(n) = \sum_{\substack{ar=n\\ r>1}}\frac{\Lambda_f(a)}{\log r}f(r)\log r = \sum_{\substack{ar=n\\ r>1}}\frac{\Lambda_f(a)}{\log r}\sum_{bm=r}\Lambda_f(b)f(m).$$

The $\log r$ in the denominator is difficult to deal with but can be replaced using the identity $\frac{1}{\log r} = \int_0^\infty r^{-\alpha}d\alpha$, and so

$$f(n)\log n - \Lambda_f(n) = \int_0^\infty \sum_{abm=n}\Lambda_f(a)\frac{\Lambda_f(b)}{b^\alpha}\frac{f(m)}{m^\alpha}d\alpha$$

(the condition $r>1$ disappears because $\Lambda_f(1)=0$). If we now sum the left hand side over all $n\leq x$ then we change the condition on the sum on the right-hand side to $abm \leq x$.

There are several variations possible on this basic identity. If we iterate (1.5.1) then we have $\log(x/p)$ in the denominator. We remove this, as above, to obtain

$$S(x)\log x = \int_0^\infty \sum_{pq\leq x}(f(p)p^\alpha\log p)(f(q)\log q)x^{-\alpha}S\left(\frac{x}{pq}\right)d\alpha + O(x\log\log x),$$

though some effort is needed to deal with the error terms. One useful variant is to restrict the primes p and q to the ranges $Q\leq p\leq x/Q,\ q>Q$ at the cost an extra $O(x\log Q)$ in the error term.

1.5.6. Exercises

`ex:ConvolutionId`

Exercise 1.5.1. Prove that

$$\frac{1}{u}\int_0^u s(u-t)\chi(t)dt = \frac{\log y}{u}\int_0^u s(t)(2t-u)y^t dt$$

Exercise 1.5.2. Define $\chi^*(u) := \frac{1}{\psi(y^u)}\sum_{m\leq y^u}\Lambda_f(m)$, so that if $|\Lambda_f(m)| \leq \kappa\Lambda(m)$ for all m then $|\chi^*(u)|\leq\kappa$. Prove that if $\kappa=1$ and $\psi(x)=x+O(x/(\log x)^2)$ then $\int_0^u s(u-t)\chi^*(t)dt = \int_0^u s(u-t)\chi(t)dt + O(1/\log y)$.

Exercise 1.5.3. Convince yourself that the functional equation for estimating smooth numbers, that we gave earlier, is a special case of (1.5.2).

Exercise 1.5.4. Improve (1.5.8) to $|s(u)| \leq \frac{1}{u}\int_0^u |s(t)|dt + o(1)$ assuming the prime number theorem. Moreover improve the error term to $O(\frac{\log\log x}{\log x})$ assuming that $\theta(x) = x + O(\frac{x}{(\log x)^2})$.

Part 2

Mean values of multiplicative functions

We introduce the main results in the theory of mean values of multiplicative functions. We begin with results as we look at the mean up to x, as $x \to \infty$. Then we introduce and prove Halász's Theorem, which allows us to obtain results that are uniform in x. The subtle proof of Halász's Theorem requires a chapter of its own.

CHAPTER 2.1

Distances. The Theorems of Delange, Wirsing and Halász

In Chapter 1.2 we considered the heuristic that the mean value of a multiplicative function f might be approximated by the Euler product $\mathcal{P}(f;x)$ (see (1.2.2) and (1.2.3)). We proved some elementary results towards this heuristic and were most successful when f was "close to 1" (see §1.2.3) or when f was non-negative (see §1.2.4). Even for nice non-negative functions the heuristic is not entirely accurate, as revealed by the example of smooth numbers discussed in Chapter ??. We now continue our study of this heuristic, and focus on whether the mean value can be bounded above by something like $|\mathcal{P}(f;x)|$. We begin by making precise the geometric language, already employed in §1.2.3, of one multiplicative function being "close" to another.

2.1.1. The distance between two multiplicative functions

The notion of a distance between multiplicative functions makes most sense in the context of functions whose values are restricted to the unit disc $\mathbb{U} = \{|z| \le 1\}$. In thinking of the distance between two such multiplicative functions f and g, naturally we may focus on the difference between $f(p^k)$ and $g(p^k)$ on prime powers. An obvious candidate for quantifying this distance is

$$\sum_{p^k \le x} \frac{|f(p^k) - g(p^k)|}{p^k},$$

as it is used in Propositions 1.2.1, 1.2.4, 1.2.5 and 1.2.6. However, it turns out that a better notion of distance involves $1 - \mathrm{Re}(f(p^k)\overline{g(p^k)})$ in place of $|f(p^k) - g(p^k)|$.

lem4.1 Lemma 2.1.1. *Suppose we have a sequence of functions $\eta_j : U \to \mathbb{R}_{\ge 0}$ satisfying the triangle inequality*

$$\eta_j(z_1, z_3) \le \eta_j(z_1, z_2) + \eta_j(z_2, z_3),$$

for all z_1, z_2, $z_3 \in U$. Then we may define a metric $U^{\mathbb{N}} = \{\mathbf{z} = (z_1, z_2, \ldots) :$ each $z_j \in U\}$ by setting

$$d(\mathbf{z}, \mathbf{w}) = \Big(\sum_{j=1}^{\infty} \eta_j(z_j, w_j)^2 \Big)^{\frac{1}{2}},$$

assuming that the sum converges. This metric satisfies the triangle inequality

$$d(\mathbf{z}, \mathbf{w}) \le d(\mathbf{z}, \mathbf{y}) + d(\mathbf{y}, \mathbf{w}).$$

Proof. Expanding out we have

$$d(\mathbf{z},\mathbf{w})^2 = \sum_{j=1}^{\infty} \eta_j(z_j,w_j)^2 \le \sum_{j=1}^{\infty} (\eta_j(z_j,y_j)+\eta_j(y_j,w_j))^2$$

by the assumed triangle inequality for η_j. Now, using Cauchy-Schwarz, we have

$$\begin{aligned}\sum_{j=1}^{\infty}(\eta_j(z_j,y_j)+\eta_j(y_j,w_j))^2 &= d(\mathbf{z},\mathbf{y})^2 + d(\mathbf{y},\mathbf{w})^2 + 2\sum_{j=1}^{\infty}\eta_j(z_j,y_j)\eta_j(y_j,w_j)\\ &\le d(\mathbf{z},\mathbf{y})^2 + d(\mathbf{y},\mathbf{w})^2 + 2\Big(\sum_{j=1}^{\infty}\eta_j(z_j,y_j,)^2\Big)^{\frac12}\Big(\sum_{j=1}^{\infty}\eta_j(y_j,w_j)^2\Big)^{\frac12}\\ &= (d(\mathbf{z},\mathbf{y})+d(\mathbf{y},\mathbf{w}))^2,\end{aligned}$$

which proves the triangle inequality. □

A nice class of examples is provided by taking $\eta_j(z) = a_j(1-\mathrm{Re}\ (z_j))$ for non-negative a_j, with $U=\mathbb{U}$. We now check that this satisfies the hypothesis of Lemma lem4.1 2.1.1.

lem4.0 Lemma 2.1.2. *Define $\eta : \mathbb{U}\times\mathbb{U}\to\mathbb{R}_{\ge 0}$ by $\eta(z,w)^2 = 1-\mathrm{Re}(z\overline{w})$. Then for any w, y, z in $\mathbb{U}$ we have*

$$\eta(w,y)\le \eta(w,z)+\eta(z,y).$$

Proof. (Terry Tao) Any point u on the unit disk is the midpoint of the line between two points u_1,u_2 on the unit *circle*, and thus their average (that is $u=\frac12(u_1+u_2)$).[1] Therefore

$$\begin{aligned}\frac18\sum_{i,j=1}^{2}|t_i-u_j|^2 &= \frac14\sum_{i,j=1}^{2}(1-\mathrm{Re}(t_i\overline{u_j}))\\ &= 1-\mathrm{Re}\left(\frac12\sum_{i=1}^{2}t_i\cdot\frac12\sum_{j=1}^{2}\overline{u_j}\right) = \mathrm{Re}(1-t\overline{u}) = \eta(t,u)^2.\end{aligned}$$

Define the four dimensional vectors $v(w,z) := (w_1-z_1, w_1-z_2, w_2-z_2, w_2-z_1)$ and $v(z,y) := (z_1-y_1, z_2-y_2, z_2-y_1, z_1-y_2)$, with $v(w,y) := v(w,z)+v(z,y)$, so that $\eta(t,u) = \frac{1}{\sqrt8}|v(t,u)|$ where t,u is any pair from w,y,z. Using the usual triangle inequality, we deduce that

$$\eta(w,y) = \frac{1}{\sqrt8}|v(w,y)| \le \frac{1}{\sqrt8}(|v(w,z)|+|v(z,y)|) = \eta(w,z)+\eta(z,y).$$

□

We can use the above remarks to define distances between multiplicative functions taking values in the unit disc. If we let $a_j = 1/p$ for each prime $p\le x$ then we may define the distance (up to x) between the multiplicative functions f and g by

$$\mathbb{D}(f,g;x)^2 = \sum_{p\le x}\frac{1-\mathrm{Re}\ f(p)\overline{g(p)}}{p}.$$

[1] To see this, draw the line L from the origin to u and then the line perpendicular to L, going through u. This meets the unit circle at u_1 and u_2. If u was on the unit circle to begin with then $u_1 = u_2 = u$.

By Lemma 2.1.1 this satisfies the triangle inequality

triangle1 (2.1.1)
$$\mathbb{D}(f,g;x)+\mathbb{D}(g,h;x)\geq \mathbb{D}(f,h;x).$$

EXERCISE 2.1.1. (i) Determine when $\mathbb{D}(f,g;x)=0$.
(ii) Determine when $\mathbb{D}(f,g;x)+\mathbb{D}(g,h;x)=\mathbb{D}(f,h;x)$.

EXERCISE 2.1.2. It is natural to multiply multiplicative functions together, and to ask if f_1 and g_1 are close to each other, and f_2 and g_2 are close to each other, is f_1f_2 is close to g_1g_2? Indeed prove this variant of the triangle inequality:

triangle2 (2.1.2)
$$\mathbb{D}(f_1,g_1;x)+\mathbb{D}(f_2,g_2;x)\geq \mathbb{D}(f_1f_2,g_1g_2;x).$$

There are several different distances that one may take. There are advantages and disadvantages to including the prime powers in the definition of $\mathbb{D}$ (see, e.g exercise ??),

$$\mathbb{D}^*(f,g;x)^2=\sum_{p^k\leq x}\frac{1-\text{Re }f(p^k)\overline{g(p^k)}}{p^k};$$

but either way the difference between two such notions of distance is bounded by a constant. Another alternative is to define a distance $\mathbb{D}_\alpha$, defined by taking the coefficients $a_j=1/p^\alpha$ and $z_j=f(p)$, as p runs over all primes for any fixed $\alpha>1$, which satisfies the analogies to (2.1.1) and (2.1.2).

ex:MertensIneq EXERCISE 2.1.3. Combine the last two variants of distance to form $\mathbb{D}^*_\alpha$. Use the triangle inequality (and exponentiate) to deduce *Mertens inequality*: For all $\sigma>1$ and all $t\in\mathbb{R}$,

$$\zeta(\sigma)^3|\zeta(\sigma+it)|^4|\zeta(\sigma+2it)|\geq 1;$$

as well as $\zeta(\sigma)^3|\zeta(\sigma+2it)|\geq|\zeta(\sigma+it)|^4$.

lem4.3 EXERCISE 2.1.4. Prove that if each $|a_p|\leq 2$ and $\alpha=1+1/\log x$ then

$$\sum_{p\leq x}\frac{a_p}{p}=\sum_{p\ \text{prime}}\frac{a_p}{p^\alpha}+O(1)\ .$$

(Hint: Consider the primes $p\leq x$, and those $>x$, separately.) Deduce that for any multiplicative functions f and g taking values in the unit disc we have

$$\mathbb{D}(f,g;x)^2=\sum_{p\ \text{prime}}\frac{1-\text{Re }f(p)\overline{g(p)}}{p^\alpha}+O(1)$$

EXERCISE 2.1.5. Suppose that f is a multiplicative function taking values in the unit disc and $\text{Re}(s)>1$. Recall that $F(s):=\sum_{n\geq1}f(n)/n^s$. Prove that

$$\log F(s)=\sum_{p\ \text{prime}}\frac{\Lambda_f(n)/\log n}{n^s}=\sum_{p\ \text{prime}}\frac{f(p)}{p^s}+O(1).$$

Deduce from this and the previous exercise that

TruncRight (2.1.3)
$$\left|F\left(1+\frac{1}{\log x}+it\right)\right|\asymp\log x\ \exp\Big(-\mathbb{D}(f(n),n^{it};x)^2\Big).$$

2.1.2. Delange's Theorem

We are interested in when the mean value of f up to x is close to its "expected" value of $\mathcal{P}(f;x)$, or even $\mathcal{P}(f)$. Proposition pr2.1 1.2.1 implies (as in exercise ex:Prop2.1 1.2.8) that if f is a multiplicative function taking values in the unit disc $\mathbb{U}$ and $\sum_p |1-f(p)|/p < \infty$ then $\sum_{n\le x} f(n) \sim x\mathcal{P}(f)$ as $x\to\infty$. Delange's theorem, which follows, is therefore a refinement of Proposition pr2.1 1.2.1.

Delange

THEOREM 2.1.3. **(Delange's theorem)** *Let f be a multiplicative function taking values in the unit disc $\mathbb{U}$. Suppose that*

$$\mathbb{D}(1,f;\infty)^2 = \sum_p \frac{1-\text{Re } f(p)}{p} < \infty.$$

Then

$$\sum_{n\le x} f(n) \sim x\mathcal{P}(f;x) \quad \text{as } x\to\infty.$$

We shall prove Delange's Theorem in the next chapter. Delange's Theorem is not exactly what we asked for in the discussion above, so the question now is whether $\lim_{x\to\infty}\mathcal{P}(f;x)$ exists and equals $\mathcal{P}(f)$. It is straightforward to deduce the following:

DelangeCor

COROLLARY 2.1.4. *Let f be a multiplicative function taking values in the unit disc $\mathbb{U}$. Suppose that*

$$\lim_{x\to\infty}\sum_{p\le x}\frac{1-f(p)}{p} \quad \text{converges (to a finite value).}$$

Then

$$\sum_{n\le x} f(n) \sim x\mathcal{P}(f) \quad \text{as } x\to\infty.$$

We postpone the proof of Delange's theorem to the next chapter.

2.1.3. A key example: the multiplicative function $f(n) = n^{i\alpha}$

Delange's theorem gives a satisfactory answer in the case of multiplicative functions at a bounded distance from 1, and we are left to ponder what happens when $\mathbb{D}(1,f;x)\to\infty$ as $x\to\infty$. One would be tempted to think that in this case $\frac{1}{x}\sum_{n\le x} f(n)\to 0$ as $x\to\infty$ were it not for the following important counter example. Let $\alpha\ne 0$ be a fixed real number and consider the completely multiplicative function $f(n)=n^{i\alpha}$. By partial summation we find that

eq:4.1

$$\sum_{n\le x} n^{i\alpha} = \int_{0^+}^{x^+} y^{i\alpha}d[y] \sim \frac{x^{1+i\alpha}}{1+i\alpha}. \tag{2.1.4}$$

The mean-value at x then is $\sim x^{i\alpha}/(1+i\alpha)$ which has magnitude $1/|1+i\alpha|$ but whose argument varies with x. In this example it seems plausible enough that $\mathbb{D}(1,p^{i\alpha};x)\to\infty$ as $x\to\infty$ and we now supply a proof of this important fact. We begin with a useful Lemma on the Riemann zeta function.

lem4.3.0

LEMMA 2.1.5. *If $s=\sigma+it$ with $\sigma>1$ then*

$$\left|\zeta(s)-\frac{s}{s-1}\right| \le \frac{|s|}{\sigma}\ .$$

If in addition we have $|s-1| \gg 1$ then

$$|\zeta(s)| \ll \log(2+|s|).$$

PROOF. The first assertion follows easily from Exercise 1.1.2. To prove the second assertion, we deduce from Exercise 1.1.6 that, for any integer $N \geq 1$, we have

$$\zeta(s) = \sum_{n=1}^{N} \frac{1}{n^s} + \frac{N^{1-s}}{s-1} - s\int_N^\infty \frac{\{y\}}{y^{s+1}} dy.$$

Choose $N = [|s|] + 1$, and bound the sum over n trivially to deduce the stated bound for $|\zeta(s)|$. ☐

Zetaprimebound

EXERCISE 2.1.6. Use similar ideas to prove that if $s = \sigma + it$ with $\sigma > 1$ and $|s-1| \gg 1$ then $|\zeta'(s)| \ll \log^2(2+|s|)$.

lem4.3.1

LEMMA 2.1.6. *Let α be any real number. Then for all $x \geq 3$ we have*

$$\mathbb{D}(1, p^{i\alpha}; x)^2 = \log(1 + |\alpha| \log x) + O(1),$$

in the case $|\alpha| \leq 100$. When $|\alpha| \geq 1/100$ we have

DLowerTBig

$$\mathbb{D}(1, p^{i\alpha}; x)^2 \geq \log\log x - \log\log(2+|\alpha|) + O(1), \tag{2.1.5}$$
$$\textit{and}\ \ \mathbb{D}(1, p^{i\alpha}; x)^2 \leq \log\log x + 8\log\log(2+|\alpha|) + O(1)$$

PROOF. We take $f(n) = 1$ in (2.1.3). The first two estimates follow directly from the bounds of Lemma 2.1.5, and are equivalent to

EquivUpperBounds

$$\sum_{p\leq x} \frac{\mathrm{Re}(p^{i\alpha})}{p} \begin{cases} = \log(1/|\alpha|) + O(1), & \text{if } 1/\log x \leq |\alpha| \leq 100; \\ \leq \log\log(2+|\alpha|) + O(1), & \text{if } |\alpha| \geq 1/100. \end{cases} \tag{2.1.6}$$

The first estimate yields the third estimate for $1/100 \leq |\alpha| \leq 100$ so henceforth we assume $|\alpha| > 100$. The second estimate of (2.1.6) implies that for $u \geq y := \exp((\log|\alpha|)^8)$ we have

$$\sum_{p\leq u} \frac{\cos^2(\alpha\log p)}{p} = \frac{1}{2}\sum_{p\leq u}\frac{1+\cos(2\alpha\log p)}{p} \leq \frac{9}{16}\log\log u + O(1).$$

This implies that

$$\sum_{p\leq u} \frac{-\cos(\alpha\log p)}{p} \leq \Big(\sum_{p\leq u}\frac{1}{p}\Big)^{1/2}\Big(\sum_{p\leq u}\frac{\cos^2(\alpha\log p)}{p}\Big)^{1/2} \leq \frac{3}{4}\log\log u + O(1),$$

and therefore, by (2.1.3),

$$\Big|1/\zeta\Big(1+\frac{1}{\log u}+i\alpha\Big)\Big| \asymp \exp\Big\{-\sum_{p\leq u}\frac{\cos(\alpha\log p)}{p}\Big\} \ll (\log u)^{3/4}.$$

Combining this with the bound $|\zeta'(1+\frac{1}{\log u}+i\alpha)| \ll \log^2|\alpha|$ obtained from exercise 2.1.6, we deduce that for $x > y$,

$$\Big|\sum_{y<p\le x}\frac{1}{p^{1+i\alpha}}\Big| = \Big|\log\Big\{\zeta\Big(1+\frac{1}{\log x}+i\alpha\Big)\Big\} - \log\Big\{\zeta\Big(1+\frac{1}{\log y}+i\alpha\Big)\Big\}\Big| + O(1)$$
$$= \Big|\int_y^x \frac{-\zeta'}{\zeta}\Big(1+\frac{1}{\log u}+i\alpha\Big)\frac{du}{u\log^2 u}\Big| + O(1)$$
$$\ll 1 + \int_y^x \frac{\log^2|\alpha|(\log u)^{3/4}}{u\log^2 u}du \ll 1.$$

The result then follows, since

$$-\sum_{p\le x}\frac{\mathrm{Re}(p^{i\alpha})}{p} + O(1) \le \sum_{p\le y}\frac{1}{p} + O(1) \le 8\log\log|\alpha| + O(1).$$

$\square$

One important consequence of Lemma 2.1.6 and the triangle inequality is that a multiplicative function cannot pretend to be like two different problem examples, $n^{i\alpha}$ and $n^{i\beta}$.

cor:repulsive

Corollary 2.1.7. *Let α and β be two real numbers and let f be a multiplicative function taking values in the unit disc. If $\delta = |\alpha-\beta|$ then*

$$\Big(\mathbb{D}(f,p^{i\alpha};x)+\mathbb{D}(f,p^{i\beta};x)\Big)^2 \ge \begin{cases}\log(1+\delta\log x)+O(1), & \text{if } \delta\le 1/10;\\ \log\log x - \log\log(2+\delta)+O(1), & \text{if } \delta \ge 1/10.\end{cases}$$

Proof. Indeed the triangle inequality gives that $\mathbb{D}(f,p^{i\alpha};x)+\mathbb{D}(f,p^{i\beta};x) \ge \mathbb{D}(p^{i\alpha},p^{i\beta};x) = \mathbb{D}(1,p^{i(\alpha-\beta)};x)$ and we may now invoke Lemma 2.1.6. $\square$

An useful consequence of Lemma 2.1.6 when working with Dirichlet characters (see Chapter ?? for the definition) is the following:

cor:ftothekbound

Corollary 2.1.8. *Suppose that there exists an integer $k\ge 1$ such that $f(p)^k = 1$ for all primes p. For any fixed non-zero real α we have*

$$\mathbb{D}(f(p),p^{i\alpha};x)^2 \ge \frac{1}{k^2}\log\log x + O_{k,\alpha}(1).$$

Examples of this include $f=\mu$ the Möbius function, or indeed any $f(n)$ which only takes values -1 and 1, as well as $f=\chi$ a Dirichlet character (though one needs to modify the result to deal with the finitely many primes p for which $\chi(p)=0$), and even $f=\mu\chi$.

Proof of Corollary 2.1.8. By the triangle inequality, we have $k\mathbb{D}(f(p),p^{i\alpha};x) \ge \mathbb{D}(1,p^{ik\alpha};x)$ and the result then follows immediately from Lemma 2.1.6. $\square$

The problem example $n^{i\alpha}$ discussed above takes on complex values, and one might wonder if there is a real valued multiplicative function f taking values in $[-1,1]$ for which $\mathbb{D}(1,f;x)\to\infty$ as $x\to\infty$ but for which the mean value does not tend to zero. A lovely theorem of Wirsing shows that this does not happen.

Wirsing

THEOREM 2.1.9 (Wirsing's Theorem). *Let f be a real valued multiplicative function with $|f(n)| \leq 1$ and $\mathbb{D}(1, f; x) \to \infty$ as $x \to \infty$. Then as $x \to \infty$*

$$\frac{1}{x}\sum_{n\leq x} f(n) \to 0.$$

Wirsing's theorem applied to $\mu(n)$ immediately yields the prime number theorem (using Theorem 1.1.3). We shall not directly prove Wirsing's theorem, but instead deduce it as a consequence of the important theorem of Halász, which we discuss in the next section (see Corollary ?? for a quantitative version of Theorem 2.1.9).

2.1.4. Halász's theorem; the qualitative version

We saw in the previous section that the function $f(n) = n^{i\alpha}$ has a large mean value even though $\mathbb{D}(1, f; x) \to \infty$ as $x \to \infty$. We may tweak such a function at a small number of primes and expect a similar result to hold. More precisely, one can ask if an analogy to Delange's result holds: that is if f is multiplicative with $\mathbb{D}(f(p), p^{i\alpha}; \infty) < \infty$ for some α, can we understand the behavior of $\sum_{n\leq x} f(n)$? This is the content of the qualitative version of Halász's theorem.

Hal1

THEOREM 2.1.10. **(Qualitative Halász theorem)** *Let f be a multiplicative function with $|f(n)| \leq 1$ for all integers n.*
(i) *Suppose that there exists $\alpha \in \mathbb{R}$ for which $\mathbb{D}(f, p^{i\alpha}; \infty) < \infty$. Write $f(n) = g(n)n^{i\alpha}$. Then, as $x \to \infty$,*

$$\sum_{n\leq x} f(n) = \frac{x^{1+i\alpha}}{1+i\alpha}\ \mathcal{P}(g; x) + o(x).$$

(ii) *Suppose that $\mathbb{D}(f, p^{i\alpha}; \infty) = \infty$ for all $\alpha \in \mathbb{R}$. Then, as $x \to \infty$,*

$$\frac{1}{x}\sum_{n\leq x} f(n) \to 0.$$

ex:fLimit

EXERCISE 2.1.7. Deduce that if f is a multiplicative function with $|f(n)| \leq 1$ for all integers n then $\frac{1}{x}\sum_{n\leq x} f(n) \to 0$ if and only if
Either (i) $\mathbb{D}(f, p^{i\alpha}; \infty) = \infty$ for all $\alpha \in \mathbb{R}$,
Or (ii) $\mathbb{D}(f, p^{i\alpha}; \infty) < \infty$ for some $\alpha \in \mathbb{R}$ and $f(2^k) = -(2^k)^{i\alpha}$ for all $k \geq 1$.
Establish that (ii) cannot happen if f is completely multiplicative.

ex:SlowVary

EXERCISE 2.1.8. If f is a multiplicative function with $|f(n)| \leq 1$ show that $\mathcal{P}(f; y)$ is slowly varying, that is $\mathcal{P}(f; y) = \mathcal{P}(f; x) + O(\log(ex/y)/\log x)$ if $y \leq x$.

PROOF OF THEOREM 2.1.10(I). We will deduce (i) from Delange's Theorem 2.1.3 and exercise 2.1.8. By partial summation we have

$$\sum_{n\leq x} f(n) = \int_1^x t^{i\alpha} d\Big(\sum_{n\leq t} g(n)\Big) = x^{i\alpha}\sum_{n\leq x} g(n) - i\alpha\int_1^x t^{i\alpha-1}\sum_{n\leq t} g(n)dt.$$

Now $\mathbb{D}(1, g; \infty) = \mathbb{D}(f, p^{i\alpha}; \infty) < \infty$ and so by Delange's theorem, if t is sufficiently large then

$$\sum_{n\leq t} g(n) = t\mathcal{P}(g; t) + o(t).$$

Substituting this into the equation above, and then applying exercise 2.1.8, we obtain

$$\sum_{n\le x} f(n) = x^{1+i\alpha}\mathcal{P}(g;x) - i\alpha\int_1^x t^{i\alpha}\mathcal{P}(g;x)dt + o(x) = \frac{x^{1+i\alpha}}{1+i\alpha}\ \mathcal{P}(g;x) + o(x).$$

□

We will deduce Part (ii) of Theorem 2.1.10 from the quantitative version of Halasz's Theorem, which we will state only in section 7.

Applying Theorem 2.1.10(i) with f replaced by $f(n)/n^{i\alpha}$ we obtain the following:

cor:Compare

Corollary 2.1.11. *Let f be multiplicative function with $|f(n)| \le 1$ and suppose there exists $\alpha \in \mathbb{R}$ such that $\mathbb{D}(f, p^{i\alpha}; \infty) < \infty$. Then as $x \to \infty$*

$$\frac{1}{x}\sum_{n\le x} f(n) = \frac{x^{i\alpha}}{1+i\alpha}\cdot\frac{1}{x}\sum_{n\le x}\ \frac{f(n)}{n^{i\alpha}} + o(1).$$

This will be improved considerably in Theorem ??. Taking absolute values in both parts of Theorem 2.1.10 we deduce:

LimAbsVal

Corollary 2.1.12. *If f is multiplicative with $|f(n)| \le 1$ then*

$$\lim_{x\to\infty}\frac{1}{x}\left|\sum_{n\le x} f(n)\right| \quad \text{exists.}$$

2.1.5. A better comparison theorem

The following quantitative result, relating the mean value of $f(n)$ to the mean-value of $f(n)n^{it}$ for any t, improves the error term in Corollary 2.1.11 to (better than) $O(x/(\log x)^{1+o(1)})$, and provides an alternative proof of Theorem 2.1.10, assuming Delange's Theorem.

AsympT1

Lemma 2.1.13. *Suppose $f(n)$ is a multiplicative function with $|f(n)| \le 1$ for all n. Then for any real number t with $|t| \le x^{1/3}$ we have*

$$\sum_{n\le x} f(n) = \frac{x^{it}}{1+it}\sum_{n\le x}\frac{f(n)}{n^{it}} + O\left(\frac{x}{\log x}\log(2+|t|)\exp\Big(\mathbb{D}(f(n), n^{it}; x)\sqrt{2\log\log x}\Big)\right).$$

ex:mtotheit

Exercise 2.1.1. Prove that if $|t| \ll m$ and $|\delta| \le 1/2$ then $2m^{it} = (m-\delta)^{it} + (m+\delta)^{it} + O(|t|/m^2)$. Deduce that

$$\sum_{m\le z} m^{it} = \begin{cases} \frac{z^{1+it}}{1+it} + O(1+t^2) \\ O(z). \end{cases}$$

Proof of Lemma 2.1.13. Let g and h denote the multiplicative functions defined by $g(n) = f(n)/n^{it}$, and $g = 1 * h$, so that $h = \mu * g$. Then

$$\sum_{n\le x} f(n) = \sum_{n\le x} g(n)n^{it} = \sum_{n\le x} n^{it}\sum_{d|n} h(d) = \sum_{d\le x} h(d)d^{it}\sum_{m\le x/d} m^{it}.$$

We use the first estimate in exercise ex:mtotheit 2.1.1 when $d \leq x/(1+t^2)$, and the second estimate when $x/(1+t^2) \leq d \leq x$. This gives

$$\sum_{n\leq x} f(n) = \frac{x^{1+it}}{1+it}\sum_{d\leq x}\frac{h(d)}{d} + O\bigg((1+t^2)\sum_{d\leq x/(1+t^2)}|h(d)| + x\sum_{x/(1+t^2)\leq d\leq x}\frac{|h(d)|}{d}\bigg).$$

Applying Proposition Prop2.1 1.2.2 and partial summation, we deduce that

$$\begin{aligned}\sum_{n\leq x} f(n) &= \frac{x^{1+it}}{1+it}\sum_{d\leq x}\frac{h(d)}{d} + O\bigg(\frac{x}{\log x}\log(2+|t|)\sum_{d\leq x}\frac{|h(d)|}{d}\bigg)\\ &= \frac{x^{1+it}}{1+it}\sum_{d\leq x}\frac{h(d)}{d} + O\bigg(\frac{x}{\log x}\log(2+|t|)\exp\bigg(\sum_{p\leq x}\frac{|1-g(p)|}{p}\bigg)\bigg).\end{aligned}$$

We use this estimate twice, once as it is, and then with $f(n)$ replaced by $f(n)/n^{it}$, and t replaced by 0, so that g and h are the same in both cases.

By the Cauchy-Schwarz inequality,

$$\bigg(\sum_{p\leq x}\frac{|1-g(p)|}{p}\bigg)^2 \leq 2\sum_{p\leq x}\frac{1}{p}\cdot\sum_{p\leq x}\frac{1-\mathrm{Re}(g(p))}{p} \leq 2\mathbb{D}(g(n),1;x)^2(\log\log x + O(1)),$$

and the result follows, since $\mathbb{D}(f(n),n^{it};x)^2 = \mathbb{D}(g(n),1;x)^2 \ll \log\log x$. □

DistI

2.1.6. Distribution of values of a multiplicative function, I

Given a multiplicative function f with $|f(n)| \leq 1$ for all n, one can ask how the values $f(n)$ are distributed in the unit disc. For example, classical work of Erdős determined the distribution of the values of $\phi(n)/n$ in the unit interval $[0,1]$. In this section we will look at the distribution of angles – these do not change when we replace each $f(n)$ by $f(n)/|f(n)|$ so we may assume that each $f(n)$ lies *on* the unit circle, i.e. $|f(n)| = 1$. To this end, let

$$R_f(N,\alpha,\beta) := \frac{1}{N}\ \#\left\{n \leq N:\ \frac{1}{2\pi}\arg(f(n)) \in (\alpha,\beta]\right\} - (\beta-\alpha).$$

We say that the $f(n)$ are *uniformly distributed* on the unit circle if $R_f(N,\alpha,\beta)\to 0$ for all $0 \leq \alpha < \beta < 1$. Jordan Ellenberg asked us whether the values $f(n)$ are necessarily equi-distributed on the unit circle according to some measure, and if not whether their distribution is entirely predictable. We prove the following (though see Proposition **??** for more on this).

equidist

Theorem 2.1.14. *Let f be a completely multiplicative function such that each $f(p)$ is on the unit circle. Either*
(i) *The $f(n)$ are uniformly distributed on the unit circle; or*
(ii) *There exists a positive integer m, and $\alpha\in\mathbb{R}$ for which $\mathbb{D}(f(p)^m, p^{im\alpha};\infty) < \infty$.*

This leads to the rather surprising (immediate) consequence:

equidistit

Corollary 2.1.15. *Let f be a completely multiplicative function such that each $f(p)$ is on the unit circle. Then the $f(n)$ are uniformly distributed on the unit circle if and only if the $f(n)/n^{it}$ are uniformly distributed on the unit circle for every $t\in\mathbb{R}$.*

To prove our distribution theorem we use

Weyl's equidistribution theorem *Let $\{\xi_n : n \geq 1\}$ be any sequence of points on the unit circle. The set $\{\xi_n : n \geq 1\}$ is uniformly distributed on the unit circle if and only if $\lim_{N\to\infty} \frac{1}{N}\sum_{n\leq N} \xi_n^m$ exists and equals 0, for each non-zero integer m.*

Proof **of Theorem 2.1.14** By Weyl's equidistribution theorem the $f(n)$ are uniformly distributed on the unit circle if and only if $\lim_{N\to\infty} \frac{1}{N}\sum_{n\leq N} f(n)^m$ exists and equals 0, for each non-zero integer m. Now each $f^m(n)$ is a completely multiplicative function such that each $f^m(p)$ is on the unit circle, so by exercise 2.1.7, this holds if and only if $\mathbb{D}(f^m(p), p^{imt}; \infty) = \infty$ for all $t \in \mathbb{R}$. The result follows. □

2.1.7. Additional exercises

Exercise 2.1.9. Prove that $\eta(z,w) := |1 - z\overline{w}|$ also satisfies the triangle inequality inside $\mathbb{U}$; i.e. $|1-z\overline{w}| \leq |1-z\overline{y}| + |1-y\overline{w}|$ for $w, y, z \in \mathbb{U}$. Prove that we get equality if and only if $z = y$, or $w = y$, or $|w| = |z| = 1$ and y is on the line segment connecting z and w. (Hint: $|1 - z\overline{w}| \leq |1 - z\overline{y}| + |z\overline{y} - z\overline{w}| \leq |1 - z\overline{y}| + |y - w| \leq |1 - z\overline{y}| + |1 - y\overline{w}|$.)

This last notion comes up in many arguments and so it is useful to compare the two quantities:

Exercise 2.1.10. By showing that $\frac{1}{2}|1-z|^2 \leq 1 - \mathrm{Re}(z) \leq |1-z|$ whenever $|z| \leq 1$, deduce that

$$\frac{1}{2}\sum_{p\leq x} \frac{|1 - f(p)\overline{g(p)}|^2}{p} \leq \mathbb{D}(f,g;x)^2 \leq \sum_{p\leq x} \frac{|1 - f(p)\overline{g(p)}|}{p} .$$

We define $\mathbb{D}(f,g;\infty) := \lim_{x\to\infty} \mathbb{D}(f,g;x)$. In the next exercise we relate distance to the product $\mathcal{P}(f;x)$, which is the heuristic mean value of f up to x:

ex:PasD Exercise 2.1.11. Suppose that f is a multiplicative function for which $|\Lambda_f(n)| \leq \Lambda(n)$ for all n. Prove that $\lim_{x\to\infty} \mathbb{D}(f,g;x)$ exists. Show that $\log|\mathcal{P}(f;x)| = -\mathbb{D}(1,f;x)^2 + O(1)$; and then deduce that $\lim_{x\to\infty} |\mathcal{P}(f;x)|$ exists if and only if $\mathbb{D}(1,f;\infty) < \infty$. Show that $|\mathcal{P}(f;x)| = 1 + O(\mathbb{D}^*(1,f;x)^2)$.

Exercise 2.1.12. Come up with an example of f, with $|f(n)| \leq 1$ for all n, for which $\mathbb{D}(1,f;\infty)$ converges but $\sum_p (1 - f(p))/p$ diverges.

ex:4.4.1 Exercise 2.1.13. If f is a multiplicative function with $|f(n)| \leq 1$ show that there is at most one real number α with $\mathbb{D}(f, p^{i\alpha}; \infty) < \infty$.

ex:Wirsing Exercise 2.1.14. Deduce Wirsing's Theorem (Theorem 2.1.9) from Theorem 2.1.10(ii). (Hint: You might use the Brun-Titchmarsh Theorem.)

ex:f(n)f(n+1)2 Exercise 2.1.15. Suppose that f is a multiplicative function with $-1 \leq f(n) \leq 1$ for each integer n.

(i) Prove that if $\frac{1}{x}\sum_{n\leq x} f(n) \not\to 0$ then

$$\lim_{x\to\infty} \frac{1}{x}\sum_{n\leq x} f(n)f(n+1) = \mathcal{P}(f,f).$$

(Hint: Use exercise 1.2.15 and Wirsing's Theorem.)

(ii) Prove that this is non-zero unless $\mathcal{P}_p(f) = \frac{1}{2}$ for some prime p.

(iii) Prove that if $f(n)$ only takes on values 1 and -1, and $\frac{1}{x}\sum_{n\le x} f(n) \not\to 0$, then $\lim_{x\to\infty} \frac{1}{x}\sum_{n\le x} f(n) > 0$, and $\lim_{x\to\infty} \frac{1}{x}\sum_{n\le x} f(n)f(n+1) > 0$ unless $\mathcal{P}_2(f,f) \le 0$ or $\mathcal{P}_3(f,f) \le 0$.