THE DISTRIBUTION OF PRIME NUMBERS

Andrew Granville and K. Soundararajan

0. Preface: Seduced by zeros

0.1. The prime number theorem, a brief history. A *prime* number is a positive integer which has no positive integer factors other than 1 and itself. It is difficult to determine directly from this definition whether there are many primes, indeed whether there are infinitely many.

Euclid described in his *Elements*, an ancient Greek proof that there are infinitely many primes, a proof by contradiction, that today highlights for us the depth of abstract thinking in antiquity. So the next question is to quantify how many primes there are up to a given point.

By studying tables of primes, Gauss understood, as a boy of 15 or 16 (in 1792 or 1793), that the primes occur with density $\frac{1}{\log x}$ at around x. In other words

$$\pi(x) := \#\{\text{primes} \le x\} \text{ is approximately } \sum_{n\le x} \frac{1}{\log n} \approx \text{Li}(x) \text{ where } \text{Li}(x) := \int_2^x \frac{dt}{\log t}.$$

This leads to the conjecture that $\pi(x)/\text{Li}(x) \to 1$ as $x \to \infty$, which is hard to interpret since $\text{Li}(x)$ is not such a natural function. One can show that $\text{Li}(x)/\frac{x}{\log x} \to 1$ as $x \to \infty$, so we can rephrase our conjecture as $\pi(x)/\frac{x}{\log x} \to 1$ as $x \to \infty$, or in less cumbersome notation that

$$\pi(x) \sim \frac{x}{\log x}. \tag{0.1.1}$$

It is not easy to find a strategy to prove such a result.

Since primes are those integers with no prime factors less than or equal to their square-root, one obvious approach to counting the number of primes in $(\sqrt{x}, x]$ is to try to estimate the number of integers up to x, with no prime factors $\le \sqrt{x}$. One might proceed by removing the integers divisible by 2 from the integers $\le x$, then those divisible by 3, etc, and keeping track of how many integers are left at each stage. No one has ever succeeded in getting a sharp estimate for $\pi(x)$ with such a *sieving* strategy, though it is a good way to get upper bounds (see §3).

In 1859, Riemann wrote his only article in number theory, a nine page memoir containing an extraordinary plan to estimate $\pi(x)$. Using ideas seemingly far afield from the elementary question of counting prime numbers, Riemann brought in deep ideas from the

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

theory of complex functions to formulate a "program" to prove (0.1.1) that took others forty years to bring to fruition. Riemann's approach begins with the *Riemann zeta-function*,

$$\zeta(s) := \sum_{n\geq 1} \frac{1}{n^s},$$

which is well-defined when the sum is absolutely convergent, that is when $\mathrm{Re}(s) > 1$. Note also that by the Fundamental Theorem of Arithmetic one can factor each n in a unique way, and so

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1}, \tag{0.1.2}$$

when $\mathrm{Re}(s) > 1$.

The Riemann zeta-function can be extended, in a unique way, to a function that is analytic in the whole complex plane (except at $s = 1$ where it has a pole of order 1).[1] As we describe in more detail in §0.8, Riemann gave the following remarkable identity for a weighted sum over the prime powers $\leq x$:

$$\sum_{\substack{p \text{ prime} \\ p^m \leq x \\ m\geq 1}} \log p = x - \sum_{\rho:\ \zeta(\rho)=0} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)}, \tag{0.1.3}$$

counting a zero with multiplicity m_ρ exactly m_ρ times in this sum. By gaining a good understanding of the sum over the zeros ρ on the right side of (0.1.3) one can deduce that

$$\sum_{\substack{p \text{ prime} \\ p^m \leq x \\ m\geq 1}} \log p \sim x, \tag{0.1.4}$$

which is equivalent to (0.1.1). To be more precise, the *Riemann Hypothesis*, that all such ρ have real part $\leq \frac{1}{2}$, implies that each $\left|\frac{x^\rho}{\rho}\right| \leq \frac{x^{1/2}}{|\rho|}$, and one can then deduce from (0.1.3) that

$$\left|\sum_{p^m\leq x} \log p - x\right| \leq 2\sqrt{x}\log^2 x,$$

for all $x \geq 100$; or, equivalently,

$$|\pi(x) - \mathrm{Li}(x)| \leq 3\sqrt{x}\log x.$$

These estimates, and hence the Riemann Hypothesis, are far from proved. However we do not need such a strong bound on the real part of zeros of the Riemann zeta-function to

[1] In other words, there is a unique Taylor series for $(s-1)\zeta(s)$ around every point in the complex plane.

deduce (0.1.1). Indeed we shall see in §0.8 how one can deduce the *prime number theorem*, that is (0.1.1), from (0.1.3) simply by knowing that there are no zeros very close to the 1-line,[2] more precisely that there are no zeros $\rho = \beta + it$ with $\beta > 1 - 1/|t|^{1/3}$. Note that there are no zeros ρ with $\mathrm{Re}(\rho) > 1$, by (0.1.2).

Clever people near the end of the nineteenth century were able to show that the prime number theorem would follow if one could show that $\zeta(1+it) \neq 0$ for all $t \in \mathbb{R}$; that is there are no zeros of the Riemann zeta-function actually **on** the 1-line. This was proved by de la Vallée Poussin and Hadamard in 1896.

Exercises. 0.1) Show that $\mathrm{Li}(x) = x/\log x + O(x/(\log x)^2)$ and then give an asymptotic series expansion for $\mathrm{Li}(x)$. (Hint: Integrate by parts, and be careful about convergence issues).

0.2) Let $p_1 = 2 < p_2 = 3 < \dots$ be the sequence of primes. Show that the prime number theorem, (0.1.1), is equivalent to the assertion that $p_n \sim n \log n$ as $n \to \infty$. Give a much more accurate estimate for p_n assuming that the Riemann Hypothesis holds.

0.3) Show that the prime number theorem, (0.1.1), is equivalent to the assertion

$$\theta(x) := \sum_{p \le x} \log p \sim x,$$

where we weight each prime by $\log p$. (Hint: Restrict attention to the primes $> x/\log x$.)

0.4) Show that the prime number theorem, (0.1.1), is equivalent to the assertion

$$\psi(x) := \sum_{\substack{p \text{ prime} \\ p^m \le x \\ m \ge 1}} \log p \sim x. \tag{0.1.4}$$

0.2. Seduced by zeros. The birth and life of analytic number theory. The formula (0.1.3) allows one to show that the accuracy of Gauss's guesstimate for $\pi(x)$ depends on what zero-free regions for $\zeta(s)$ have been established; and vice-versa. For instance, if $\frac{1}{2} \le \alpha < 1$ then

$$\zeta(\beta + it) \neq 0 \text{ for } \beta \ge \alpha \quad \text{if and only if } |\pi(x) - \mathrm{Li}(x)| \le x^{\kappa} \text{ for } x \ge x_{\kappa},$$

for any fixed $\kappa > \alpha$ where x_κ is some sufficiently large constant. More pertinent to what is known unconditionally is that

$$\zeta(\beta + it) \neq 0 \text{ for } \beta \ge 1 - \frac{1}{(\log x)^{\alpha}} \quad \text{if and only if } |\pi(x) - \mathrm{Li}(x)| \le \frac{x}{\exp(c_3 (\log x)^{\kappa})},$$

where $\kappa = 1/(1+\alpha)$. The best result known unconditionally is that one can take any $\alpha > \frac{2}{3}$ and hence any $\kappa < \frac{3}{5}$. This result is over fifty years old – the subject is desperately in need of new ideas.

These equivalences can be viewed as expressing a tautology, between questions about the distribution of prime numbers, and questions about the distribution of zeros of the Riemann zeta-function, and whereas we have few tools to approach the former, the theory

[2] The "β-line" is defined to be those complex numbers s with $\mathrm{Re}(s) = \beta$.

of complex functions allows any number of attacks and insights into the Riemann zeta-function. For more than 150 years we have seen many beautiful observations about $\zeta(s)$ emerge, indeed it is at the center of a web of conjectures that cover all of number theory, and many questions throughout mathematics. If one believes that the charm of mathematics lies in finding surprising, profound conjectures between hitherto completely different areas, then Riemann's is the ultimate such result.

If one asks about the distribution of primes in arithmetic progressions then there are analogous zeta-functions to work with, and indeed an analogous *Generalized Riemann Hypothesis*.

0.3. Can there be analytic number theory without zeros? Is it really *necessary* to go to the theory of complex functions to count primes? And to work there with the zeros of an analytic continuation of a function, not even the function itself? This was something that was initially hard to swallow in the 19th century but gradually people came to believe it, seeing in (0.1.3) an equivalence, more-or-less, between questions about the distribution of primes and questions about the distribution of zeros of $\zeta(s)$. This is discussed in the introduction to Ingham's book [I1]: "*Every known proof of the prime number theorem is based on a certain property of the complex zeros of $\zeta(s)$, and this conversely is a simple consequence of the prime number theorem itself. It seems therefore clear that this property must be used (explicitly or implicitly) in any proof based on $\zeta(s)$, and it is not easy to see how this is to be done if we take account only of real values of s.* For these reasons, it was long believed that it was impossible to give an elementary proof of the prime number theorem.[3]

In 1948 Selberg found an elementary proof of a formula that counts not a weighted sum of primes up to x, but a weighted sum of those integers that are either prime or the product of two primes, namely:[4]

$$\sum_{\substack{p \text{ prime} \\ p\leq x}} \log^2 p + \sum_{\substack{p,q \text{ prime} \\ pq\leq x}} \log p \ \log q = 2x\log x + O(x). \tag{0.3.1}$$

We will give Selberg's proof of (0.3.1) in section 4.2. Such a formula is so close to the prime number theorem that it would seem to be impossible to prove without zeros of $\zeta(s)$. But what Selberg had done was to construct a formula in which the influence of any zeros close to the 1-line is muted, and hence can be proved in an elementary way.[5] From the

---

[3]An even better quote is due to Hardy: "*No elementary proof of the prime number theorem is known and one may ask whether it is reasonable to expect one. Now we know that the theorem is roughly equivalent to a theorem about an analytic function... A proof of such a theorem, not fundamentally dependent on the theory of functions, seems to me extraordinarily unlikely. It is rash to assert that a mathematical theorem cannot be proved in a particular way; but one thing seems clear. We have certain views about the logic of the theory; we think that some theorems, as we say, "lie deep" and others nearer to the surface. If anyone produces an elementary proof of the prime number theorem, he will show that these views are wrong, that the subject does not hang together in the way we have supposed, and that it is time for the books to be cast aside and for the theory to be rewritten.*"

[4]Here we introduce the "Big Oh" notation. That $f(x) = O(g(x))$ if there exists a constant $c > 0$ such that $|f(x)| \leq cg(x)$ for all $x \geq 1$. We also can write $f(x) \ll g(x)$.

[5]See Ingham's insightful Math Review of Selberg's article for a detailed discussion.

brilliance of this formula, Erdős quickly deduced the prime number theorem, followed by a proof of Selberg shortly thereafter.

Other elementary proofs have appeared, most using some formula like (0.3.1). There is another approach using a non-obvious re-formulation of (0.1.1):

0.4. The Möbius function. The Möbius function, $\mu(n)$, is given by the coefficients of the inverse of the Riemann zeta-function; that is by

$$\frac{1}{\zeta(s)} = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right) = \sum_{n \geq 1} \frac{\mu(n)}{n^s}, \tag{0.4.1}$$

where $\mathrm{Re}(s) > 1$. More explicitly $\mu(n) = 0$ if n is divisible by the square of a prime, and $\mu(n) = (-1)^k$ if n is the product of k distinct primes. Note that $\mu(n)$ is an example of a *multiplicative function*, that is a function $f : \mathbb{N} \to \mathbb{C}$ for which $f(mn) = f(m)f(n)$ whenever $(m, n) = 1$; in particular $f(p_1^{e_1} p_2^{e_2} \dots p_k^{e_k}) = f(p_1^{e_1}) f(p_2^{e_2}) \dots f(p_k^{e_k})$ if $p_1, \dots, p_k$ are distinct primes.

One can easily predict the number of $\mu(n)$ that are 0, but it seems far less obvious how many equal $+1$, and how many -1. Since multiplying n by one more (new) prime causes $\mu(n)$ to change sign, one might guess that there are roughly equal numbers of $+1$ and -1 amongst the $\mu(n)$. This guess can be formulated as

$$\lim_{x \to \infty} \frac{1}{x} \sum_{n \leq x} \mu(n) \text{ exists and equals } 0,$$

which can be re-written more simply as

$$\sum_{n \leq x} \mu(n) = o(x). \tag{0.4.2}$$

What is surprising is that this statement is easily shown to be equivalent to (0.1.1),[6] and several elementary proofs of the prime number theorem focus on proving this statement. Moreover the formulation (0.4.2) is key to the theory developed in this book.

0.5. Halász's Theorem. Let $f(n)$ be any multiplicative function for which $|f(n)| \leq 1$ for all integers $n \geq 1$. We are interested in determining when the mean value of $f(n)$ up to x is "large", that is $> \delta$ in absolute value for some fixed $\delta > 0$. There are some obvious examples: If $f(n) = 1$ then the mean value is 1. The correct generalization of this is to take $f(n) = n^{it}$, for some fixed real number t. Then

$$\frac{1}{x} \sum_{n \leq x} n^{it} \approx \frac{1}{x} \int_{u=0}^{x} u^{it} du = \frac{1}{x} \frac{x^{1+it}}{1+it} = \frac{x^{it}}{1+it}.$$

Now this has absolute value $1/\sqrt{1+t^2}$, so we must have $|t| = O(1/\delta)$ (which we will also write as $|t| \ll 1/\delta$) to obtain a mean value of size $> \delta$. Hence the absolute value of the

[6] Which will be proved in section 1.5.

mean-value tends to a limit as $x \to \infty$. Surprisingly this does not imply that the mean-value tends to a limit as $x \to \infty$: Indeed in our last example the mean value rotates around the origin slowly, since x^{it} does as x increases, whereas the size of the mean value tends to a limit.

What other examples have large mean value? An obvious class of examples come from minor alterations to the ones we have above. For instance if $f(p) = 1$ for all primes $p \neq 3$, and $f(3) = -1$, then f has mean value $\to \frac{1}{2}$ as $x \to \infty$. In general, if $f(n)$ is close to n^{it}, the mean value can be large. In this case we say that $f(n)$ is "*n^{it}-pretentious*", in that it is pretending to be that simple function. We will need to give a formal definition of this a little later, but that is complicated so for now we remain deliberately vague.

Halász's great theorem states that if the mean-value of f is large in absolute value, then f must be n^{it}-pretentious for some real number t for which $|t| \ll 1/\delta$.

0.6. Sketch of a proof of the prime number theorem. Now let us apply this to the mean value of the Möbius function: If (0.4.2) is false then there exists a fixed $\delta > 0$ such that there are arbitrarily large x for which $|\sum_{n\leq x} \mu(n)| \geq \delta x$. By Halász's theorem we deduce that there exists a real number t, with $|t| \ll 1/\delta$, for which $\mu(n)$ is n^{it}-pretentious. We will now give a heuristic that we will develop into a proof in the main part of this book: $\mu(n)/n^{it}$ is 1-pretentious, if and only if

$$\sum_{n\geq 1} \frac{\mu(n)}{n^{1+it}} \text{ behaves much like } \sum_{n\geq 1} \frac{1}{n}, \text{which diverges to } \infty.$$

Now $1/\zeta(1+it)$ should be much like $\sum_{n\geq 1} \mu(n)/n^{1+it}$ according to (0.4.1).[7] This suggests that $\mu(n)/n^{it}$ is 1-pretentious if and only if $\zeta(1+it) = 0$. Hence we need to show that $\zeta(1+it) \neq 0$, as did Hadamard and de la Vallée Poussin did in 1896 (see section 0.1 above).

In the proofs of Hadamard and de la Vallée Poussin one shows that if $\zeta(1+it) = 0$ then $\zeta(1+2it) = \infty$, which is impossible as $\zeta(s)$ is analytic at all $s \neq 1$. Most textbooks give an easy proof of the first deduction via an inequality of Mertens: If $\sigma > 1$ then

$$\left|\zeta(\sigma)^3\zeta(\sigma+it)^4\zeta(\sigma+2it)\right| \geq 1. \tag{0.5.1}$$

We know that $(\sigma-1)\zeta(\sigma) \to c_1$ as $\sigma \to 1$, since $\zeta(s)$ has a pole of order 1, at $s = 1$, for some constant $c_1 \neq 0$. As $\zeta(s)$ is analytic at $s = 1+it$, and as $\zeta(1+it) = 0$, therefore $\zeta(\sigma+it)/(\sigma-1) \to c_2$ as $\sigma \to 1$. Inserting these into (0.6.1) we deduce that $|(\sigma-1)\zeta(\sigma+2it)| \gg 1$ as $\sigma \to 1$, and hence $\zeta(s)$ has a pole at $s = 1+2it$.

Our proof of this deduction is less magical than the use of Mertens' inequality, but arguably more straightforward: If $\mu(n)$ is n^{it}-pretentious then $\mu(n)^2$ is n^{2it}-pretentious. This can happen if and only if

$$\sum_{n\geq 1} \frac{\mu(n)^2}{n^{1+2it}} \text{ behaves much like } \sum_{n\geq 1} \frac{\mu(n)^2}{n}, \text{which diverges to } \infty.$$

[7]But notice that (0.4.1) is only a valid identity when $\text{Re}(s) > 1$ and hence not truly valid in our current situation.

Now $\sum_{n\geq 1} \mu(n)^2/n^{1+2it}$ should be much like $\zeta(1+2it)/\zeta(2+4it)$, and $\zeta(2+4it)$ converges to a non-zero constant. Therefore $\mu(n)^2$ is n^{2it}-pretentious if and only if $\zeta(1+2it)$ diverges.

There are various ways in which one can show that $\mu(n)^2$ cannot be n^{2it}-pretentious. Our proof will use upper bounds on the number of primes in short intervals, proved using sieve methods, to establish that p^{2it} rotates around as p varies, and so cannot almost always be pointing more-or-less in the positive real direction.

Exercise. 0.5) Use (0.1.2) to show that $\zeta(s)$ has no zeros with $\mathrm{Re}(s) > 1$. (Hint: Consider the Euler product for $\zeta(s)$ in this range.)

0.7. Multiplicative Number Theory is the title of Davenport's classic book on the distribution of prime numbers, though the contents of that book mostly stem out of Riemann's seminal idea.[8] In this text we rework the basic results on the distribution of primes to be a consequence of results on the distribution of mean values of multiplicative functions, stemming mostly from the fundamental idea of Halász. As in Davenport's book we will prove theorems on $\pi(x)$ and $\pi(x; q, a)$, the number of primes up to x that are $\equiv a \pmod q$, focussing on uniformity in x, including the Bombieri-Vinogradov theorem, and a new simpler proof of Linnik's theorem as well as Vinogradov's three primes theorem. We will prove an improved Polya-Vinogradov Theorem, as well as Burgess's Theorem. Qualitatively we get all the same results, often with substantially easier proofs, quantitatively we often get poorer error terms. The biggest advantage of our approach is that our results are applicable to all multiplicative functions with values inside the unit disk; the biggest disadvantage is that we have not yet proved the Siegel-Walfisz Theorem, and this lack of uniformity is a substantial impediment to several applications of this theory.

We make no claims about giving an elementary proof of the prime number theorem. Our proof of Halász's theorem does use complex analysis: Fourier transforms and Plancherel's formula. This last may be regarded as the simplest non-trivial result in the area, and in our applications could easily be proved without complex analysis, though it seems artificial to try to do so. Since our results are so general, the proofs, as one might expect, use less tools designed for a particular problem (like zeta functions that satisfy a "functional equation").

We mostly choose to work with multiplicative functions with values inside the unit disk. Many of the technical results about multiplicative functions can be extended to wider classes, though not all, and not without some significant complications. Books by Elliott [El] and Tenenbaum [Te] are probably the best sources for advanced material in this area.

0.8. More details / more sketch for the proof of the prime number theorem.

The existing data lends support to Gauss's belief that $\pi(x)$ is well-approximated by $\mathrm{Li}(x)$.

---

[8] In his preface Davenport calls his book "*... a connected account of analytic number theory in so far as it relates to problems of a multiplicative character...*"

x	$\pi(x) = \#\{\text{primes} \le x\}$	Overcount: $[\mathrm{Li}(x) - \pi(x)]$
10^8	5761455	753
10^9	50847534	1700
10^{10}	455052511	3103
10^{11}	4118054813	11587
10^{12}	37607912018	38262
10^{13}	346065536839	108970
10^{14}	3204941750802	314889
10^{15}	29844570422669	1052618
10^{16}	279238341033925	3214631
10^{17}	2623557157654233	7956588
10^{18}	24739954287740860	21949554
10^{19}	234057667276344607	99877774
10^{20}	2220819602560918840	222744643
10^{21}	21127269486018731928	597394253
10^{22}	201467286689315906290	1932355207
10^{23}	1925320391606803968923	7250186214

Table 1. The number of primes up to various x.

One may make more precise guesses from the data in Table 1. For example one can see that the entries in the final column are always positive and are always about half the width of the entries in the middle column. So perhaps Gauss's guess is always an overcount by about $\sqrt{x}$? We have seen that if the Riemann Hypothesis is true then the difference between $\mathrm{Li}(x)$ and $\pi(x)$ is never much bigger than $\sqrt{x}$; however Gauss's guess is not always an overcount. In 1914 Littlewood showed that the difference changes sign infinitely often, it probably first goes negative at around 10^{316} (which is far beyond where we can explicitly compute all primes in the foreseeable future). Littlewood's proof involves zeros far away from the 1-line and we are currently unable to propose a proof using our methods.[9]

The *Prime Number Theorem*, was proved in 1896, by Hadamard and de la Vallée Poussin, following a program of study laid out almost forty years earlier by Riemann. Riemann's idea was to use a formula of Perron to extend the sum in (0.1.3) to be over all prime powers p^m, while picking out only those that are $\le x$. The special case of Perron's formula that we need here is

$$\frac{1}{2i\pi}\int_{s:\ \mathrm{Re}(s)=2} \frac{t^s}{s}\, ds = \begin{cases} 0 & \text{if } t < 1, \\ 1 & \text{if } t > 1, \end{cases}$$

for positive real t. We apply this with $t = x/p^m$, when x is not itself a prime power, which

[9]Our methods work best when the classical proof proceeds by showing that the zeta function zeros that are close to the 1-line are sparse.

gives us a characteristic function for integers $p^m < x$. Hence

$$\sum_{\substack{p^m \le x \\ p \text{ prime} \\ m \ge 1}} \log p = \sum_{\substack{p \text{ prime} \\ m \ge 1}} \log p \cdot \frac{1}{2i\pi} \int_{s:\ \mathrm{Re}(s)=2} \frac{(x/p^m)^s}{s}\, ds$$

$$= \frac{1}{2i\pi} \int_{s:\ \mathrm{Re}(s)=2} \sum_{p \text{ prime}} \frac{\log p}{p^{ms}} \frac{x^s}{s}\, ds.$$

Here we were able to safely swap the infinite sum and the infinite integral since the terms are sufficiently convergent as $\mathrm{Re}(s) = 2$. Recognizing that

$$\sum_{p \text{ prime}} \sum_{m \ge 1} \frac{\log p}{p^{ms}} = -\frac{\zeta'(s)}{\zeta(s)},$$

at least for $\mathrm{Re}(s) > 1$, we obtain the closed formula

$$\psi(x) = \sum_{\substack{p \text{ prime} \\ p^m \le x \\ m \ge 1}} \log p = -\frac{1}{2i\pi} \int_{s:\ \mathrm{Re}(s)=2} \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, ds. \tag{0.8.1}$$

To evaluate (0.8.1), Riemann proposed moving the contour from the line $\mathrm{Re}(s) = 2$, far to the left, and using the theory of residues to evaluate the integral. What a beautiful idea! However before one can possibly succeed with that plan one needs to know many things, for instance whether $\zeta(s)$ makes sense to the left, that is one needs an *analytic continuation* of $\zeta(s)$. Riemann was able to do this based on an extraordinary identity of Jacobi. Next, to use the residue theorem, one needs to be able to identify the poles of $\zeta'(s)/\zeta(s)$, that is the zeros and poles of $\zeta(s)$. The poles are not so hard, there is just the one, a simple pole at $s = 1$ with residue 1, so the contribution of that pole to the above formula is

$$-\lim_{s \to 1} (s-1) \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} = -\lim_{s \to 1} (s-1) \left(\frac{-1}{(s-1)} \right) \frac{x^1}{1} = x,$$

the expected main term. The locations of the zeros of $\zeta(s)$ are much more mysterious. Moreover, even if we do have some idea of where they are, in order to complete Riemann's plan, one needs to be able to bound the contribution from the discarded contour when one moves the main line of integration to the left, and hence one needs bounds on $|\zeta(s)|$ throughout the plane. We do this in part by having a pretty good idea of how many zeros there are of $\zeta(s)$ up to a certain height, and there are many other details besides. These all had to be worked out (see, eg [Da] or [Ti], for further details), after Riemann's initial plan – this is what took forty years! At the end, if all goes well, then one has the *exact formula*

$$\psi(x) = \sum_{\substack{p \text{ prime} \\ p^m \le x \\ m \ge 1}} \log p = x - \sum_{\rho:\ \zeta(\rho)=0} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)}. \tag{0.1.3}$$

Amazing! A precise formula for the weighted sum of prime powers, in terms of the zeros of an analytic continuation of a function. What an unexpected and delightful identity. This clearly runs deep and is so profound that it must lead to all sorts of insights. Indeed this has been the basis of much of analytic number theory for the last 150 years.

Using the right-side of (0.1.3) is, in practice, easier said than done. For one thing, there are infinitely many zeros of $\zeta(s)$ that effect the sum – it seems odd to deal with an infinite sum to understand a finite problem, that is the number of primes up to x. We can address this problem by truncating the sum over zeros at a given height T, and to have real part ≥ 0, that is consider only those ρ with[10] $0 \leq \text{Re}(\rho) \leq 1$ and $|\text{Im}(\rho)| \leq T$. One then has the approximation,

$$\psi(x) = x - \sum_{\substack{\rho:\ \zeta(\rho)=0 \\ 0\leq \text{Re}(\rho)\leq 1 \\ |\text{Im}(\rho)|\leq T}} \frac{x^\rho}{\rho} + O\left(\frac{x\log^2(xT)}{T}\right), \tag{0.8.2}$$

in the range $1 \leq T \leq x$, and the sum is known to be over only finitely many zeros.

As we discussed above, we bound the sum on the right side of (0.8.2) simply by taking the absolute value of each term, so we miss out on any potential cancelation (and one might guess that there will be quite a bit). Hence if $\beta_T \geq \text{Re}(\rho)$ for all ρ for which $\zeta(\rho) = 0$ and $|\text{Im}(\rho)| \leq T$, then our sum is $\leq x^{\beta_T} \sum_\rho \frac{1}{|\rho|}$ and it can be shown that this sum over the zeros in this box is $\ll \log^2 T$.

Exercise. Assume that if $\zeta(\beta + it) = 0$ then $1 - \beta \geq 1/|t|^{1/3}$. Deduce the prime number theorem (using the above discussion).

Selecting $T = x^{1-\beta}\log^2 x$ we deduce that

$$|\theta(x) - x| \ll x^\beta((1-\beta)\log x + \log\log x)^2. \tag{0.8.3}$$

Let $\pi(x; q, a)$ denote the number of primes $\leq x$ that are $\equiv a \pmod q$. A proof analogous to that proposed by Riemann, reveals that if $(a, q) = 1$ then

$$\pi(x; q, a) \sim \frac{\pi(x)}{\varphi(q)}\ , \tag{0.8.4}$$

once x is sufficiently large. However in many application one wants to know just how large x needs to be for the primes to be equi-distributed in arithmetic progressions mod q. Calculations reveal that the primes up to x are equi-distributed amongst the arithmetic progressions mod q, once x is just a tiny bit larger than q, say $x \geq q^{1+\delta}$ for any fixed $\delta > 0$ (once q is sufficiently large). However the best proven results have x bigger than the exponential of a power of q, far larger than what we expect. If we are prepared to assume the unproven *Generalized Riemann Hypothesis* we do much better, being able to prove that the primes up to $q^{2+\delta}$ are equally distributed amongst the arithmetic progressions mod q, for q sufficiently large, though notice that this is still somewhat larger than what we expect to be true.

[10]We already saw that $\zeta(s)$ has no zeros ρ for which $\text{Re}(\rho) > 1$. Moreover the zeros with $\text{Re}(\rho) < 0$ are easily found: These "trivial" zeros lie at $\rho = -2, -4, -6, \ldots$ and have little effect on the formulas above.

1. Introduction

1.1. The prime number theorem. As a boy Gauss determined 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

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

Less intuitive, but simpler, is to weight each prime with $\log p$; and, as we have seen, it is natural to throw the prime powers into this sum, which has little impact on the size, so that, defining

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

we conjecture that

$$\psi(x) := \sum_{n\le x} \Lambda(n) \sim x. \tag{1.1.2}$$

The equivalent estimates (1.1.1) and (1.1.2), known as the *prime number theorem*, are difficult to prove. Our primary goal at the beginning of this book, is to give a new proof of the prime number theorem that highlights the techniques that we develop herein.

Short of (1.1.1), there are several ways to obtain good bounds on the number of primes up to x. Perhaps the easiest is to note that all of the primes in $(N, 2N]$ divide the numerator of the binomial coefficient $\binom{2N}{N}$, and so

$$\prod_{N<p\le 2N} p \le \binom{2N}{N} \le 4^N;$$

from which it is not hard to deduce that

$$\theta(N) = \sum_{p\le N} \log p \le (\log 4)\ N, \tag{1.1.3}$$

and that

$$\pi(N) \le (\log 4 + o(1))\frac{N}{\log N}. \tag{1.1.4}$$

Lower bounds for $\pi(N)$ of the right order can be obtained by a modification of this method:

Exercise. (i) Show that there are $[N/q]$ integers $\le N$ that are divisible by q, and hence the difference in the number of integers in the numerator and denominator of $\binom{2N}{N}$ that are divisible by q is $[2N/q] - 2[N/q]$, which equals either 0 or 1.
(ii) Deduce that if p^k divides $\binom{2N}{N}$ then $p^k \le 2N$. Moreover show that p divides $\binom{2N}{N}$ if $N < p \le 2N$, but does not if $2N/3 < p \le N$.

(iii) Prove that the largest $\binom{2N}{k}$ occurs when $k=N$, so that $\binom{2N}{N} \geq \frac{4^N}{2N+1}$.
(iv) Deduce that $N\log 4 + O(\log N) \leq \theta(2N)-\theta(N)+\theta(2N/3)+(\log N)\pi(\sqrt{2N})$.
(v) Use (1.1.3) and (1.1.4) to deduce that $\theta(2N)-\theta(N) \geq \frac{N}{3}\log 4 + O(\sqrt{N})$, and hence that

$$\pi(N) \geq \left(\frac{\log 4}{3}+o(1)\right)\frac{N}{\log N}.$$

Since the Riemann-zeta function is absolutely convergent for $\mathrm{Re}(s)>1$, we can manipulate this series, more-or-less at will in this range. Various functions of $\zeta(s)$ will be of importance to us, in particular

$$\begin{aligned} -\zeta'(s) &= \sum_{n\geq 1}\frac{\log n}{n^s}, \\ -\frac{\zeta'(s)}{\zeta(s)} &= \sum_{n\geq 1}\frac{\Lambda(n)}{n^s}, \\ \frac{1}{\zeta(s)} &= \sum_{n\geq 1}\frac{\mu(n)}{n^s}, \end{aligned}$$

where we (again) define the Möbius function,

$$\mu(n) := \begin{cases} 0 & \text{if } p^2 \text{ divides } n, \text{ for some prime } p \\ (-1)^k & \text{if } n=p_1p_2\dots p_k, \text{ where } p_1,p_2,\dots,p_k \text{ are distinct primes.} \end{cases}$$

The Möbius function is an example of a multiplicative function: That is a function $f(.)$ with the property that

$$f(mn)=f(m)f(n) \text{ whenever } (m,n)=1.$$

One can show that if $n=p_1^{a_1}p_2^{a_2}\dots p_k^{a_k}$, where $p_1,p_2,\dots,p_k$ are distinct primes, then

$$f(n)=f(p_1^{a_1})f(p_2^{a_2})\dots f(p_k^{a_k}).$$

Our main goal in this section will be to prove that the prime number theorem (1.1.1) is equivalent to the conjecture that

$$M(x)=\sum_{n\leq x}\mu(n)=o(x). \tag{1.1.5}$$

That is that the mean value of a certain multiplicative function, that lives inside the unit disc, tends to 0. The main study in this book are the mean values of such multiplicative functions in some generality.

1.2. Integrals of monotone functions.

Exercises. 1.2.1) Define $s_N := \sum_{n=1}^N \frac{1}{n} - \log N$. Since $1/t$ is a decreasing function one sees that

$$\int_{n-1}^{n} \frac{dt}{t} > \frac{1}{n} > \int_{n}^{n+1} \frac{dt}{t}.$$

Use this to show that $s_N > 0$ for all $N \geq 1$, and that if $N > M > 1$ then $0 < s_M - s_N < \frac{1}{M}$. Deduce that

$$\sum_{n=1}^{N} \frac{1}{n} = \log N + \gamma + O\left(\frac{1}{N}\right), \tag{1.2.1}$$

where we define the Euler-Mascheroni constant

$$\gamma := \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n} - \log N.$$

1.2.2) Use the fact that $\log t$ is an increasing function to deduce, by analogous arguments, that

$$\log N! = N \log N - N + O(\log N). \tag{1.2.2}$$

Improve this to $\log N! = N\log N - N + \frac{1}{2}\log N + c + O(1/N)$, for some constant c, by showing that $\log N = \int_{N-1/2}^{N+1/2} \log t \; dt + O(1/N^2)$. (Establishing that $c = \frac{1}{2}\log 2\pi$ yields Stirling's formula.)

Another approach to this formula is to use the identity

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

to deduce that

$$\log N! = \sum_{n\leq N} \sum_{d|n} \Lambda(d) = \sum_{d\leq N} \Lambda(d) \left[\frac{N}{d}\right] = N \sum_{d\leq N} \frac{\Lambda(d)}{d} + O\left(\sum_{d\leq N} \Lambda(d)\right).$$

By (1.1.3) and (1.2.2) we deduce that

$$\sum_{p\leq N} \frac{\log p}{p} = \log N + O(1). \tag{1.2.3}$$

Exercises. 1.2.3) Use (1.2.3) to show that if there exists a constant $c > 0$ such that $\psi(x) \sim cx$ then $c = 1$.
1.2.4) Use partial summation on (1.2.3) to show that

$$\sum_{y<p\leq x} \frac{1}{p} = \log\left(\frac{\log x}{\log y}\right) + O\left(\frac{1}{\log y}\right);$$

and then use this to show that there exists a constant c such that

$$\sum_{p\leq x} \frac{1}{p} = \log\log x + c + O\left(\frac{1}{\log x}\right). \tag{1.2.4}$$

From this deduce that there exists a constant γ such that

$$\prod_{p\leq x} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log x}. \tag{1.2.5}$$

(In fact γ *is* the Euler-Mascheroni constant. There does not seem to be a straightforward, intuitive proof known that it is indeed *this* constant.)

If $f(n)$ is any function with $|f(n)| \leq 1$ for all n then, by (1.2.2),

$$\left|\sum_{n\leq N} f(n)\log(N/n)\right| \leq \sum_{n\leq N} \log(N/n) = N\log N - \log N! \leq N + O(\log N). \tag{1.2.6}$$

1.3. Dirichlet series, convergence and Möbius inversion..

Exercises. Begin by establishing that

$$\sum_{ab=n} \mu(a) = \begin{cases} 1 & \text{if } n=1 \\ 0 & \text{otherwise.} \end{cases} \tag{1.3.1}$$

Remind yourself what the Möbius inversion formula is and prove it using (1.3.1).

Since $\tau(n) = \sum_{d|n} 1$, use Möbius inversion to obtain that

$$1 = \sum_{ab=n} \mu(a)\tau(b) \tag{1.3.2}$$

Similarly starting with

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

use Möbius inversion to obtain von Mangoldt's formula

$$\Lambda(n) = \sum_{ab=n} \mu(a) \log b. \tag{1.3.3}$$

Now $\sum_{a|n} \mu(a) \log n = 0$ by (1.3.1), so writing $b = n/a$ in (1.3.3) we have $\Lambda(n) = \sum_{a|n} \mu(a) \log 1/a$. Therefore, by Möbius inversion,

$$\mu(n) \log 1/n = \sum_{ab=n} \mu(a)\Lambda(b). \tag{1.3.4}$$

Exercises. Suppose that, for every $\epsilon > 0$ we have $|f_n|, |g_n| \ll n^\epsilon$. Prove that $\sum_{n\geq 1} f_n/n^s$ is absolutely convergent for $\text{Re}(s) > 1$. Deduce that if

$$\sum_{n\geq 1} \frac{h_n}{n^s} = \sum_{a\geq 1} \frac{f_a}{a^s} \cdot \sum_{b\geq 1} \frac{g_b}{b^s} \text{ then } h_n = \sum_{\substack{ab=n \\ a,b\geq 1}} f_a g_b \text{ for all } n \geq 1.$$

Use this to establish the four identities in displayed equations in this subsection.

We call

$$\sum_{n\geq 1} \frac{f(n)}{n^s}$$

a *Dirichlet series*. The product of two Dirichlet series, whose coefficients are given by the above formula

$$h(n) = \sum_{\substack{ab=n \\ a,b\geq 1}} f(a)g(b) = \sum_{d|n} f(d)g(n/d),$$

is known as a (Dirichlet) *convolution*, and is often denoted by $h = f * g$.

Exercise. Prove that if f is multiplicative and $|f(n)| \ll_\epsilon n^\epsilon$ then

$$\sum_{n\geq 1} \frac{f(n)}{n^s} = \prod_{p \text{ prime}} \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \dots\right)$$

for any s for which $\text{Re}(s) > 1$.

1.4. Dirichlet's hyperbola trick. Divisors of an integer come in pairs. That is $\{a, b\}$ where $ab = n$. Evidently the smaller of the two is $\leq \sqrt{n}$ and therefore

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

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

$$\sum_{n\leq x} \tau(n) = 2 \sum_{n\leq x} \sum_{\substack{d|n \\ d<\sqrt{n}}} 1 + \sum_{\substack{n\leq x \\ n=d^2}} 1 = \sum_{d<\sqrt{x}} \left(1 + 2 \sum_{\substack{d^2<n\leq x \\ d|n}} 1 \right) = \sum_{d<\sqrt{x}} \left(2[x/d] - 2d + 1\right),$$

and so

$$\sum_{n\leq x} \tau(n) = 2x \sum_{d<\sqrt{x}} \frac{1}{d} - x + O(\sqrt{x}) = x \log x - x + 2\gamma x + O(\sqrt{x}),$$

by (1.2.1). We deduce, using (1.2.2), that

$$\sum_{n\leq x} (\log n + 2\gamma - \tau(n)) = O(\sqrt{x}). \tag{1.4.1}$$

1.5. The prime number theorem and multiplicative functions.

If we sum up the identity (1.3.1) over all $n \leq x$ we obtain

$$1 = \sum_{ab\leq x} \mu(a) = \sum_{a\leq x} \mu(a) \left[\frac{x}{a}\right].$$

Now $0 \leq x/a - [x/a] < 1$ for each a and so

$$\left| x \sum_{a\leq x} \frac{\mu(a)}{a} \right| \leq 1 + \sum_{a\leq x} |\mu(a)| \leq x,$$

for $x \geq 4$. Therefore, verifying the cases $x = 1, 2, 3$ by hand, we have

$$\left| \sum_{a\leq x} \frac{\mu(a)}{a} \right| \leq 1 \tag{1.5.1}$$

for all $x \geq 1$. With this tool in hand we shall prove:

Theorem 1.1. *The estimates (1.1.1) and (1.1.5) are equivalent. That is $\psi(x) - x = o(x)$ if and only if $M(x) = o(x)$.*

Summing the identities in (1.3.3), (1.3.2) and (1.3.1) over all $n \le x$, yields

$$\psi(x) - x = \sum_{n \le x} (\Lambda(n) - 1) = \sum_{ab \le x} \mu(a)(\log b - \tau(b) + 2\gamma) - 2\gamma.$$

We separate this sum into two parts. Those $b \le B$, and the rest. By (1.4.1), the rest yields

$$\ll \sum_{a \le x/B} |\mu(a)| \sqrt{x/a} \ll x/\sqrt{B},$$

and so

$$\psi(x) - x = \sum_{b \le B} (\log b - \tau(b) + 2\gamma) M(x/b) + O\left(\frac{x}{\sqrt{B}} \right). \tag{1.5.2}$$

We deduce that if $M(x) = o(x)$ then for any fixed B we have $\psi(x) - x \ll x/\sqrt{B}$, and so $\psi(x) - x = o(x)$ letting $B \to \infty$ slowly enough with x.

Summing (1.3.4) over all $n \le x$ yields

$$\sum_{n \le x} \mu(n) \log 1/n = \sum_{ab \le x} \mu(a) \Lambda(b).$$

By (1.2.6) and (1.5.1), we deduce that

$$M(x) \log x = - \sum_{a \le x} \mu(a) \left(\psi\left(\frac{x}{a}\right) - \frac{x}{a} \right) + O(x). \tag{1.5.3}$$

Now, by (1.1.3),

$$\sum_{A < a \le x} \mu(a) \left(\psi\left(\frac{x}{a}\right) - \frac{x}{a} \right) \ll \sum_{A < a \le x} \frac{x}{a} \ll x \log(x/A).$$

We deduce that if $\psi(x) - x = o(x)$ then by splitting the sum in (1.5.3) into those $a \le A := x/\log x$, and those $a > A$ (as in the line above), we have $M(x) = o(x)$.

Remark. It is worth noting that the identity at the beginning of this proof can be seen to be the sum of the coefficients of the Dirichlet series

$$\sum_{n \ge 1} \frac{\Lambda(n) - 1}{n^s} = \frac{\zeta'(s)}{\zeta(s)} - \zeta(s) = \frac{1}{\zeta(s)}(-\zeta'(s) - \zeta(s)^2) = \sum_{a \ge 1} \frac{\mu(a)}{a^s} \sum_{b \ge 1} \frac{\log b - \tau(b)}{b^s}.$$

Exercises. Modify the above proof to show that

i) If $M(x) \ll \frac{x}{(\log x)^A}$ then $\psi(x) - x \ll \frac{x}{(\log x)^A} (\log\log x)^2$;

ii) If $\psi(x) - x \ll \frac{x}{(\log x)^A}$ then $M(x) \ll \frac{x}{(\log x)^{\min\{1,A\}}}$.

1.6. Other equivalent forms of the prime number theorem. We have seen that the estimate $\psi(x) \sim x$ is equivalent to $M(x) = o(x)$. In [TM, Theorem 5] two more equivalent forms are given:[11]

$$\sum_{n \geq 1} \frac{\mu(n)}{n} = 0$$

which is stronger than (1.5.1). Also, that

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log x - \gamma + o(1),$$

a strong form of (1.2.3). Note that Merten's Theorem, (1.2.5) yields

$$\sum_{n \leq x} \frac{\Lambda(n)}{n \log n} = \log \log x + \gamma + o(1). \tag{1.6.1}$$

[11]With elegant elementary proofs of their equivalence.

2. Sieving

2.1. Integers coprime to m. One first encounters a multiplicative function when counting the integers in an interval that are coprime with a given integer m, for this is the sum of $f(n)$ where $f(p) = 1$ unless $p|m$, in which case $f(p) = 0$.

The classic way to work on this is via the inclusion-exclusion formula:

$$\sum_{\substack{n\le x\\(n,m)=1}} 1 = \sum_{n\le x}\sum_{d|(m,n)} \mu(d) = \sum_{d|m}\mu(d)\left[\frac{x}{d}\right];$$

and then approximating $[x/d] = x/d + O(1)$ we obtain

$$\prod_{p|m}\left(1-\frac{1}{p}\right)x + O(\tau(m)).$$

The problem is that if m has $\gg \log x$ prime factors then the error term here is much larger than the main term, so we need to be more subtle about how we apply inclusion-exclusion. The trick is to use the inequalities:

$$\sum_{\substack{d|m\\ \omega(d)\le 2k}} \mu(d) \ge \sum_{d|m}\mu(d) \ge \sum_{\substack{d|m\\ \omega(d)\le 2k+1}} \mu(d),$$

which are valid for all $k \ge 0$ (Exercise). Hence we have the approximation

$$\sum_{j=0}^{J-1}\sum_{\substack{d|m\\ \omega(d)=j}} \mu(d)\left[\frac{x}{d}\right] = \sum_{j=0}^{J-1}\sum_{\substack{d|m\\ \omega(d)=j}} \mu(d)\frac{x}{d} + O\left(\sum_{\substack{d|m\\ \omega(d)\le J-1}} 1\right)$$

for the number of integers $\le x$ that are coprime with m, with error

$$\le \sum_{\substack{d|m\\ \omega(d)=J}} \left[\frac{x}{d}\right] \le \sum_{\substack{d|m\\ \omega(d)=J}} \frac{x}{d} \le \frac{x}{J!}\left(\sum_{p|m}\frac{1}{p}\right)^J. \tag{2.1.1}$$

If we repeat this argument with $x = m$ then we find that

$$\prod_{p|m}\left(1-\frac{1}{p}\right) = \sum_{j=0}^{J-1}\sum_{\substack{d|m\\ \omega(d)=j}} \frac{\mu(d)}{d} + O\left(\frac{1}{J!}\left(\sum_{p|m}\frac{1}{p}\right)^J\right).$$

Combining these estimates we obtain

$$\sum_{\substack{n\le x\\(n,m)=1}} 1 - x\prod_{p|m}\left(1-\frac{1}{p}\right) \ll \sum_{i=1}^{J-1}\frac{\omega(m)^i}{i!} + \frac{x}{J!}\left(\sum_{p|m}\frac{1}{p}\right)^J. \tag{2.1.2}$$

where $\omega(m)$ denotes the number of distinct prime factors of m. So if all prime factors of m are $\leq y = x^{1/u}$ then, selecting $J = [u]$ we obtain, by (1.2.4) and then (1.2.2),

$$\sum_{\substack{n\leq x\\(n,m)=1}} 1 = x\prod_{p|m}\left(1-\frac{1}{p}\right) + O\left(x\left(\frac{e\log\log y}{u}\right)^{u-1}\right). \tag{2.1.3}$$

This generalizes to give the *Fundamental Lemma of the Sieve* (see any book on sieves)...

2.2. Mean values of multiplicative functions: Heuristic. Given a multiplicative function f with $|f(n)| \leq 1$ for all n, we are interested in the mean-value of f up to x, that is $\frac{1}{x}\sum_{n\leq x} f(n)$. A simple heuristic suggests that

$$\frac{1}{x}\sum_{n\leq x} f(n) \to \text{Prod}(f,\infty) \quad \text{as } x\to\infty. \tag{2.2.1}$$

where

$$\text{Prod}(f,x) := \prod_{p\leq x}\Big(1+\frac{f(p)}{p}+\frac{f(p^2)}{p^2}+\dots\Big)\Big(1-\frac{1}{p}\Big).$$

Erdős and Wintner conjectured that is true when f is real-valued, which was proved by Wintner in 1944 when $\text{Prod}(f,\infty)\neq 0$, and Wirsing in 1967 when $\text{Prod}(f,\infty) = 0$).

The heuristic goes as follows: Define g so that $f(n) = \sum_{d|n} g(d)$, and therefore $g(n) = \sum_{d|n}\mu(n/d)f(d)$. Then

$$\sum_{n\leq N} f(n) = \sum_{n\leq N}\sum_{d|n} g(d) = \sum_{d\leq N} g(d)\sum_{\substack{n\leq N\\ d|n}} 1 = \sum_{d\leq N} g(d)\left[\frac{N}{d}\right].$$

Now each $[N/d] = N/d + O(1)$ so this becomes

$$N\sum_{d\leq N}\frac{g(d)}{d} + O\left(\sum_{d\leq N}|g(d)|\right).$$

One can easily invent restrictive hypotheses to ensure that this tends to a limit; and note that

$$\sum_{\substack{d\geq 1\\ p|d \implies p\leq N}} \frac{g(d)}{d} = \text{Prod}(f,N),$$

so we can hope to obtain (2.2.1). It turns out that such a heuristic is easiest to turn into a justified argument when f is real-valued (or is "close" to a real-valued function); and the example of section 0.5 (the mean-value of n^{it}) shows that (2.2.1) is certainly not always true.

Proposition 2.2.1. (i) *If* $\sum_p \frac{|1-f(p)|}{p}$ *converges then (2.2.1) holds. In particular*

(ii) *If* f *is real-valued and Prod*$(f;x)$ *converges then (2.2.1) holds.*

(iii) *If* $0 \le f(n) \le 1$ *for all* n *then (2.2.1) holds in all cases.*

(iv) *If* f *is real-valued and* $1 \le f(p) \le f(p^2) \le \dots$ *for all primes* $p \le x$*, then* $\sum_{n\le x} f(n) \le x\mathrm{Prod}(f;x)$.

Erdős and Wintner conjectured (2.2.1) when f is real-valued, which was proved by Wintner in 1944 when $\mathrm{Prod}(f,\infty) \ne 0$ (that is, Proposition 2.2.1(ii)), and by Wirsing in 1967 when $\mathrm{Prod}(f,\infty) = 0$ (which is a little more than Proposition 2.2.1(iii)).

Proof. Fix $\epsilon > 0$ and select y so that $\sum_{p>y} \frac{|1-f(p)|}{p} < \epsilon$. Let $g(p^a) = f(p^a)$ if $p \le y$, and $g(p^a) = 1$ if $p > y$. We observe that

$$\mathrm{Prod}(g,x) = \mathrm{Prod}(g,y) = \mathrm{Prod}(f,y) = \mathrm{Prod}(f,x)e^{O(\epsilon)} = \mathrm{Prod}(f,x) + O(\epsilon).$$

Let $f = g * h$ so that

$$\sum_{n\le x}(f(n)-g(n)) = \sum_{dm\le x} h(d)g(m) - \sum_{m\le x} g(m) = \sum_{1<d\le x} h(d) \sum_{m\le x/d} g(m).$$

Taking absolute values this is

$$\begin{aligned} &\le \sum_{1<d\le x} |h(d)| \frac{x}{d} \le x \prod_{p\le x}\left(1 + \frac{|h(p)|}{p} + \frac{|h(p^2)|}{p^2} + \dots\right) - x \\ &\ll x\left(\exp\left(\sum_{y<p\le x} \frac{|1-f(p)|}{p} + O\left(\frac{1}{p^2}\right)\right) - 1\right) \ll \epsilon x. \end{aligned}$$

Hence $\sum_{n\le x} f(n) = \sum_{n\le x} g(n) + O(\epsilon x)$.

Now let $g = 1 * \ell$ so that

$$\begin{aligned} \sum_{n\le x} g(n) &= \sum_{n\le x}\sum_{d|n} \ell(d) = \sum_{d\le x} \ell(d)\left[\frac{x}{d}\right] = x\sum_{d\le x}\frac{\ell(d)}{d} + O\left(\sum_{d\le x}|\ell(d)|\right) \\ &= x\sum_{d\ge 1}\frac{\ell(d)}{d} + O\left(\sum_{d\le x}|\ell(d)| + x\sum_{d>x}\frac{|\ell(d)|}{d}\right) \end{aligned}$$

since $\ell(p^a) = g(p^a) - g(p^{a-1})$. The main term here is $x\mathrm{Prod}(g,x) = x\mathrm{Prod}(f,x) + O(\epsilon x)$. Now notice that if $\ell(d) \ne 0$ then d is y-smooth; that is all prime factors of d are $\le y$. Now if $0 < \alpha < 1$ then $(x/d)^\alpha \ge 1$ if $d \le x$, and $(d/x)^{1-\alpha} \ge 1$ if $d > x$ so that

$$\begin{aligned} \sum_{d\le x}|\ell(d)| + x\sum_{d>x}\frac{|\ell(d)|}{d} &\le \sum_{d\le x}|\ell(d)|\left(\frac{x}{d}\right)^\alpha + x\sum_{d>x}\frac{|\ell(d)|}{d}\left(\frac{d}{x}\right)^{1-\alpha} \\ &= x^\alpha \prod_{p\le y}\left(1 + \frac{|\ell(p)|}{p^\alpha} + \frac{|\ell(p^2)|}{p^{2\alpha}} + \dots\right) \end{aligned}$$

Let $x = y^u$ and select $\alpha = 1 - \frac{A}{\log y}$ for some $A > 1$ with $\alpha > 2/3$ so that the above is $\ll (x/e^{Au}) \exp\left(2\sum_{p\le y} \frac{1}{p^\alpha}\right)$, and the sum over prime numbers is, by (1.2.4) and (1.1.4),

$$\le \sum_{p\le y^{1/A}} \frac{1}{p} + \sum_{y^{1/A}<p\le y} \frac{1}{p^\alpha} \le \log\left(\frac{\log y}{A}\right) + O(1) + (\log 4 + o(1))\frac{e^A}{A}.$$

We select $A = \log(u \log u)$ and therefore our error term is

$$\ll \left(\frac{4 + o(1)}{u \log u}\right)^u x \log y.$$

We pick x big enough so that this is $\le \epsilon x$. Collecting the above estimates we deduce that $\sum_{n\le x} f(n) = x\text{Prod}(f, y) + O(\epsilon x) = x\text{Prod}(f, x)e^{O(\epsilon)} + O(\epsilon x) = x\text{Prod}(f, x) + O(\epsilon x)$, which yields the first two parts of the result.

Suppose that $0 \le f(n) \le 1$ for all n, and that $\text{Prod}(f; \infty) = 0$. Select y so that $\text{Prod}(f, y) < \epsilon$, and then select $g(n)$ as above. As $0 \le f(n) \le g(n)$ for all n we have $0 \le \sum_{n\le x} f(n) \le \sum_{n\le x} g(n) \le x\text{Prod}(f, y) + O(\epsilon x) \ll \epsilon x$, and thus the third part of the result is proved.

Let $g = f * \mu$, so that $f = 1 * g$. The hypothesis of iv) implies that $g(n) \ge 0$ for all n. Therefore

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

$$\le x\prod_{p\le x}\left(1 + \frac{g(p)}{p} + \frac{g(p^2)}{p^2} + \dots\right) = x\text{Prod}(f; x). \tag{2.2.2}$$

2.3. The Brun-Titchmarsh Theorem.

The Brun-Titchmarsh Theorem. *If $(a, q) = 1$ then we have, uniformly,*

$$\sum_{\substack{x<p\le x+yq \\ p\equiv a \pmod q}} 1 \le \{2 + o(1)\}\frac{q}{\varphi(q)}\,\frac{y}{\log y}.$$

Let m be a given integer. Let A denote the integers $n \equiv a \pmod q$ for which $x < n \le x + yq$. Selberg observed that if $\lambda_d, d \ge 1$ are real and $\lambda_1 = 1$ then

$$\left(\sum_{d|(m,n)} \lambda_d\right)^2 \ge \begin{cases} 1 & \text{if } (n, m) = 1 \\ 0 & \text{otherwise,} \end{cases}$$

no matter what the choice of the λ_d's (where we suppose that m is the product of some primes $\le z$, with $(m,q)=1$). Hence

$$\begin{aligned} N := \sum_{\substack{n\in A\\ (n,m)=1}} 1 &\le \sum_{n\in A}\Bigg(\sum_{d|(m,n)} \lambda_d\Bigg)^2 \\ &= \sum_{d_1,d_2|m} \lambda_{d_1}\lambda_{d_2} \sum_{\substack{n\in A\\ [d_1,d_2]|n}} 1 \\ &\le \sum_{d_1,d_2|m} \lambda_{d_1}\lambda_{d_2} \frac{y}{[d_1,d_2]} + \sum_{d_1,d_2|m} |\lambda_{d_1}|\,|\lambda_{d_2}|. \end{aligned} \tag{2.3.1}$$

Now $\frac{1}{[d_1,d_2]} = \frac{(d_1,d_2)}{d_1d_2} = \frac{1}{d_1d_2}\sum_{r|(d_1,d_2)}\varphi(r)$ so the first term is

$$y\sum_{d_1,d_2|m} \frac{\lambda_{d_1}\lambda_{d_2}}{d_1d_2} \sum_{r|(d_1,d_2)} \varphi(r) = y\sum_{r|m}\varphi(r)\Bigg(\sum_{\substack{d|m\\ r|d}} \frac{\lambda_d}{d}\Bigg)^2$$

We now define $\lambda_d = 0$ if $d > z$ and $d \nmid m$; with

$$G_k(x) := \sum_{\substack{n\le x\\ n|m\\ (n,k)=1}} \frac{\mu^2(n)}{\varphi(n)}, \text{ and } \lambda_d = \frac{\mu(d)d}{\varphi(d)}\frac{G_d(z/d)}{G(z)}$$

if $d\le z$ and $d|m$, and $G(x) = G_1(x)$. If $d|m$ then

$$\frac{d}{\varphi(d)}G_d(z/d) = \sum_{g|d}\frac{\mu^2(g)}{\varphi(g)} \sum_{\substack{n\le z/d\\ n|m\\ (n,d)=1}} \frac{\mu^2(n)}{\varphi(n)} \le \sum_{\substack{r\le z\\ r|m}} \frac{\mu^2(r)}{\varphi(r)} = G(z)$$

writing $r = gn$, and noting that there is at most one such factorization of r. Hence $|\lambda_d|\le 1$, and the second term in (2.3.1) is $\le z^2$.

The choice of the λ_d gives that if $r|m$ and $r\le z$ then

$$\begin{aligned} \sum_{\substack{d|m\\ r|d}} \frac{\lambda_d}{d} &= \frac{1}{G(z)}\sum_{\substack{d|m\\ r|d}}\frac{\mu(d)}{\varphi(d)} \sum_{\substack{n\le z/d\\ n|m\\ (n,d)=1}} \frac{\mu^2(n)}{\varphi(n)} \\ &= \frac{1}{G(z)}\sum_{\substack{\ell\le z\\ \ell|m}}\frac{\mu^2(\ell)}{\varphi(\ell)} \sum_{d:\ r|d|\ell} \mu(d) = \frac{1}{G(z)}\frac{\mu(r)}{\varphi(r)}, \end{aligned}$$

writing $\ell = dn$. Hence the first term in (2.3.1) is

$$\frac{y}{G(z)^2} \sum_{\substack{r \le z \\ r|m}} \frac{\mu^2(r)}{\varphi(r)} = \frac{y}{G(z)}.$$

Let Q be the product of the primes $\le z$ that do not divide m. Now $\frac{n}{\varphi(n)} = \sum_\ell \frac{1}{\ell}$, where the sum is over all integers ℓ whose prime factors all divide n, and so

$$\frac{Q}{\varphi(Q)}\ G(z) = \sum_r \frac{1}{r}\ \sum_\ell \frac{1}{\ell} \ge \sum_{n \le z} \frac{1}{n} \ge \log z$$

by exercise 1.2.1. where our sums over all integers r whose prime factors all divide Q, and all integers ℓ whose largest squarefree divisor divides m and is $\le z$. We deduce that

$$N \le \frac{Q}{\varphi(Q)}\ \frac{y}{\log z} + z^2.$$

The Brun-Titchmarsh Theorem follows by taking $Q = q$ so that m is the product of all of the primes $\le z$ that do not divide q and $z = y^{1/2}/\log y$. More generally, by (1.2.5) we deduce that if $(m, q) = 1$ then

$$\sum_{\substack{x < n \le x+yq \\ n \equiv a \pmod q \\ (n,m)=1}} 1 \le \{e^\gamma + o(1)\} y \prod_{\substack{p \le \sqrt{y} \\ p|m}} \left(1 - \frac{1}{p}\right). \tag{2.3.2}$$

Exercise: For any given η, $\frac{1}{\log y} \ll \eta < 1$, show that

$$\sum_{p \le y} \frac{1}{p^{1-\eta}} \le \log(1/\eta) + O\left(\frac{y^\eta}{\log(y^\eta)}\right). \tag{2.3.3}$$

(Hint: Compare the sum for the primes with $p^\eta \ll 1$ to the sum of $1/p$ in the same range. Use upper bounds on $\pi(x)$ for those primes for which $p^\eta \gg 1$.)

3. Multiplicative functions

3.1. Mean values of multiplicative functions. In general one has

$$S(x) \log x = \sum_{p \le x} f(p) \log p\ S\left(\frac{x}{p}\right) + O(x). \tag{3.1.1}$$

where

$$S(x) := \sum_{n \le x} f(n),$$

for any multiplicative f for which $|f(n)| \le 1$. To prove this we use (1.2.6) to show that

$$S(x)\log x + O(x) = \sum_{n\le x} f(n)\log n = \sum_{n\le x} f(n)\sum_{d|n}\Lambda(d) = \sum_{d\le x}\Lambda(d)\sum_{\substack{n\le x\\ d|n}} f(n). \tag{3.1.2}$$

This last sum has size $\le x/d$. We discard the terms with $d = p^b, b \ge 2$; and replace the terms $f(n)$ by $f(pn/p^k)$ when $d = p$ is prime and $p^k \| n$ where $k \ge 2$. These operations yield an error of $\ll x/p^2$, and hence (3.1.1) follows.

Now, for $z = y + y/\log^2 y$, using the Brun-Titchmarsh theorem,

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

and

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

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

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

which implies, by (3.1.1), that

$$\begin{aligned}|S(x)| &\ll \frac{1}{\log x}\int_1^x \left|S\left(\frac{x}{t}\right)\right| dt + \frac{x}{\log x} \\ &= \frac{x}{\log x}\int_1^x |S(t)|\frac{dt}{t^2} + \frac{x}{\log x}.\end{aligned} \tag{3.1.3}$$

Note that if $f(.) \ge 0$ then

$$\int_1^x |S(t)|\frac{dt}{t^2} = \int_1^x \sum_{n\le t} f(n)\frac{dt}{t^2} = \sum_{n\le x} f(n)\int_n^x \frac{dt}{t^2} = \sum_{n\le x}\frac{f(n)}{n} - \frac{1}{x}\sum_{n\le x} f(n),$$

which can be inserted into (3.1.3), though we will do better in the next subsection.

3.2. Upper bounds by averaging further. Suppose that $0 \le h(p^a) \ll C^a$ for all prime powers p^a, where $C < 2$.

Exercise: Use this hypothesis to show that $\sum_{p^a\le x} h(p^a)\log p^a \ll x$. Give an example to show that this fails for $C = 2$.

Therefore

$$\sum_{n\le x} h(n)\log n = \sum_{n\le x} h(n)\sum_{p^a\|n}\log p^a = \sum_{m\le x} h(m)\sum_{\substack{p^a\le x/m\\ p\nmid m}} h(p^a)\log p^a \ll x\sum_{m\le x}\frac{h(m)}{m},$$

by the Brun-Titchmarsh theorem. Moreover, since $\log(x/n)\le x/n$ whenever $n\le x$, hence

$$\sum_{n\le x} h(n)\log(x/n)\le x\sum_{m\le x}\frac{h(m)}{m}$$

and adding these together gives

$$\sum_{n\le x} h(n)\ll \frac{x}{\log x}\sum_{m\le x}\frac{h(m)}{m}. \tag{3.2.1}$$

Using partial summation we deduce from (3.2.1) that for $1\le y\le x^{1/2}$,

$$\sum_{x/y<n\le x}\frac{h(n)}{n}\ll\left\{\frac{1}{\log x}-\log\left(1-\frac{\log y}{\log x}\right)\right\}\sum_{n\le x}\frac{h(n)}{n}. \tag{3.2.2}$$

3.3. Smooth numbers, I. In section $1\frac{1}{2}$.1 we discussed the most basic multiplicative functions, the integers coprime to given integer m. Here we introduce a function that occurs all over analytic number theory, which counts the integers that only have "small" prime factors; that is

$$\Psi(x,y) := \sum_{\substack{n\le x\\ p|n \implies p\le y}} 1.$$

We call an integer *y-smooth* if all of its prime factors are $\le y$. We can recover the above as a question about multiplicative functions by taking $f(p)=1$ if $p\le y$, and $f(p)=0$ otherwise. The key result is that $\Psi(x,x^{1/u})/x$ tends to a (non-zero) limit as $x\to\infty$, for any fixed u: Define the *Dickman-de Bruijn function*, $\rho(u)$, as follows: $\rho(u)=1$ if $0\le u\le 1$, and $\rho(u)=1-\log u$ if $1\le u\le 2$. In fact, for any $u>1$, $\rho(u)$ can be defined as an average of its recent history:

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

Theorem 3.3.1. *We have*

$$\Psi(x,y)=x\rho(u)+O\left(\frac{x}{\log x}\right)$$

where $x=y^u$, and $\rho(.)$ is the Dickman-de Bruijn function.

Proof. If $0 \le u \le 1$ then $\Psi(x,y) = \Psi(x,x) = [x] = x + O(1)$ and the result follows. Any integer counted by $[x] - \Psi(x,y)$ has a prime factor $> y$, and so if $1 \le u \le 2$ then each such integer can be written uniquely as mp where p is a prime $> y$. Therefore

$$\Psi(x,y) = [x] - \sum_{\substack{y<p\le x \\ p \text{ prime}}} \sum_{m \le x/p} 1 = [x] - \sum_{y<p\le x} [x/p] = x\left(1 - \sum_{y<p\le x} \frac{1}{p}\right) + O\left(\pi(x)\right).$$

By exercise 1.2.4, and the bound (1.1.4) for the error term we deduce that

$$\Psi(x,y) = x\left(1 - \log u\right) + O\left(\frac{x}{\log y}\right),$$

as desired, since $\rho(u) = 1 - \log u$ in this range.

Now if f is the characteristic function for the y-smooth integers, then (3.1.1) yields

$$\Psi(x,y)\log x = \sum_{p\le y} \log p \ \Psi\left(\frac{x}{p}, y\right) + O(x) \tag{3.3.1}$$

By partial summation we have, letting $E(t) := \sum_{p\le t} \frac{\log p}{p} - \log t$,

$$\sum_{y^\alpha < p \le y^\beta} \log p \ \frac{x}{p} \ \rho\left(\frac{\log(x/p)}{\log y}\right) = x\int_{y^\alpha}^{y^\beta} \rho\left(u - \frac{\log t}{\log y}\right) d(\log t + E(t)).$$

The first part of the integral equals, writing $t = y^\nu$,

$$x\log y \int_\alpha^\beta \rho\left(u - \nu\right) d\nu.$$

Now $E(t) \ll 1$ by (1.2.3) and so the second part is, after integrating by parts,

$$x\left[\rho\left(u - \frac{\log t}{\log y}\right) E(t)\right]_{y^\alpha}^{y^\beta} - x\int_{y^\alpha}^{y^\beta} \frac{d}{dt}\rho\left(u - \frac{\log t}{\log y}\right) E(t)dt \ll x\rho(u-\alpha)$$

since $\rho(.)$ is monotonically decreasing. Now if we write $\Psi(x,y) = x\rho(u)(1+\epsilon(u))$ then substituting the above into (3.3.1) yields

$$x\rho(u)\epsilon(u)\log x = \sum_{p\le y} (x/p)\log p \ \rho(u_p)\epsilon(u_p) + O(x) \tag{3.3.2}$$

where $x/p = y^{u_p}$, using the integral equation defining $\rho(u)$.

We shall prove our theorem by induction on $n \ge 1$, for $n/2 < u \le (n+1)/2$. The result is already proved for $n \le 3$, so now assume that the result is proved for $n-1$. Let

$\kappa(n) := \max_{n/2\le v\le (n+1)/2} |\Psi(y^v,y)/y^v\rho(v)-1|$, and select u where this maximum occurs. Hence for $x=y^u$, we can deduce from (3.3.2) that

$$\kappa(n)x\rho(u)\log x \le \sum_{p<x/y^{n/2}} (x/p)\log p\ \rho(u_p)\epsilon(u_p) + \sum_{x/y^{n/2}\le p\le y} (x/p)\log p\ \rho(u_p)\epsilon(u_p) + O(x)$$
$$\le \kappa(n)x\log y\int_{n/2}^{u}\rho(\nu)\,d\nu + \kappa(n-1)x\log y\int_{u-1}^{n/2}\rho(\nu)\,d\nu + O(x).$$

Moving the first term from the right side to the left side, using the functional equation for $\rho(u)$, and noting that $\int_{u-1}^{n/2}\rho(\nu)\,d\nu \ge \int_{u-1}^{u-1/2}\rho(\nu)\,d\nu \ge u\rho(u)/2$, we now have

$$\kappa(n)\le\kappa(n-1)+O(1/u\rho(u)\log y),$$

and the result follows, by induction.

There is a very nice technique, due to Rankin, to find an upper bound on $\Psi(x,y)$ as follows: Select any $\sigma>0$, so that

$$\Psi(x,y)=\sum_{\substack{n\le x\\ P(n)\le y}} 1 \le \sum_{\substack{n\le x\\ P(n)\le y}} \left(\frac{x}{n}\right)^{\sigma} = x^{\sigma}\prod_{p\le y}\left(1-\frac{1}{p^{\sigma}}\right)^{-1}.$$

This can be extended to consider, for $0<\sigma<1$

$$x\int_x^{\infty}\frac{\Psi(t,y)}{t^2}dt = \sum_{\substack{n\le x\\ P(n)\le y}} 1 + \sum_{\substack{n>x\\ P(n)\le y}} \frac{x}{n}$$
$$\le \sum_{\substack{n\le x\\ P(n)\le y}} \left(\frac{x}{n}\right)^{\sigma} + \sum_{\substack{n>x\\ P(n)\le y}} \frac{x}{n}\cdot\left(\frac{n}{x}\right)^{1-\sigma} = x^{\sigma}\prod_{p\le y}\left(1-\frac{1}{p^{\sigma}}\right)^{-1}.$$

Let $\sigma = 1-\frac{\log(u\log u)}{\log y}$, and use (2.3), to deduce that

$$\int_x^{\infty}\frac{\Psi(t,y)}{t^2}dt \le \left(\frac{C}{u\log u}\right)^{u}\log y \tag{3.3.3}$$

for a certain constant $C>0$.

For small u we get a better bound for $\Psi(x,y)$, by using (3.2.1): Let $X\le x$ and use the upper bound $1\le (n/X)^{\eta}$ for $n\ge X$, with $\eta = c/\log y$ and $c<\log 2$, so that $p^{\eta}\le y^{\eta}=e^c<2$. Hence, by (3.2.1) we have

$$\Psi(x,y)\le X+\sum_{\substack{n\le x\\ P(n)\le y}} (n/X)^{\eta} \ll X + \frac{x}{X^{\eta}\log x}\sum_{\substack{m\le x\\ P(m)\le y}} \frac{1}{m^{1-\eta}}$$

Since η is a far smaller than $1/2$, we have

$$\log\left(\frac{1}{\log y}\sum_{\substack{m\le x\\ P(m)\le y}}\frac{1}{m^{1-\eta}}\right)\le O(1)+\sum_{p\le y}\left(\frac{1}{p^{1-\eta}}-\frac{1}{p}\right)\le \eta\sum_{p\le y}\frac{\log p}{p}+O(1)\ll 1.$$

Therefore choosing $X=x^{1-\eta}$, we have the upper bound $\Psi(x,y)\ll x^{1-\eta+\eta^2}\ll x/e^{2u/3}$, and also

$$\int_x^\infty \frac{\Psi(t,y)}{t^2}dt \ll \int_x^\infty \frac{dt}{t^{1+\eta-\eta^2}}\ll \frac{1}{\eta x^{\eta-\eta^2}}\ll e^{-2u/3}$$

3.4. Bounding the tail of a sum.

Lemma 3.4.1. *If f and g are totally multiplicative, with $0\le f(p)\le g(p)\le p$ for all primes p, then*

$$\prod_{p\le y}\left(1-\frac{f(p)}{p}\right)\sum_{\substack{n\le x\\ P(n)\le y}}\frac{f(n)}{n}\le \prod_{p\le y}\left(1-\frac{g(p)}{p}\right)\sum_{\substack{n\le x\\ P(n)\le y}}\frac{g(n)}{n}$$

Proof. We prove this in the case that $f(q)<g(q)$ and $g(p)=f(p)$ otherwise, since then the result follows by induction. Define h so that $g=f*h$, so that $h(q^{b+1})=(g(q)-f(q))g(q^b)$ for all $b\ge 0$, and $h(p^a)=0$ otherwise. The left hand side above equals $\prod_{p\le y}\left(1-\frac{g(p)}{p}\right)$ times

$$\sum_{m\ge 1}\frac{h(m)}{m}\sum_{\substack{n\le x\\ P(n)\le y}}\frac{f(n)}{n}\ge \sum_{\substack{N\le x\\ P(N)\le y}}\sum_{mn=N}\frac{h(m)}{m}\cdot\frac{f(n)}{n}=\sum_{\substack{n\le x\\ P(n)\le y}}\frac{g(n)}{n},$$

as desired.

Corollary 3.4.2. *Suppose that f is a totally multiplicative function, with $0\le f(p)\le 1$ for all primes p. Then*

$$\prod_{p\le y}\left(1-\frac{f(p)}{p}\right)\sum_{\substack{n>x\\ p|n\implies p\le y}}\frac{f(n)}{n}\ll\left(\frac{C}{u\log u}\right)^u,$$

where $x=y^u$.

Proof. If take $x=\infty$, both sides equal 1 in the Lemma. Hence if we subtract both sides from 1, and let $g=1$, we obtain

$$\prod_{p\le y}\left(1-\frac{f(p)}{p}\right)\sum_{\substack{n>x\\ p|n\implies p\le y}}\frac{f(n)}{n}\le \prod_{p\le y}\left(1-\frac{1}{p}\right)\sum_{\substack{n>x\\ p|n\implies p\le y}}\frac{1}{n}.$$

By Mertens' theorem and this is

$$\lesssim \frac{e^{-\gamma}}{\log y}\int_x^\infty \frac{d\Psi(t,y)}{t}\le \frac{e^{-\gamma}}{\log y}\int_x^\infty \frac{\Psi(t,y)}{t^2}dt,$$

and the result follows from (3.3.3).

3.5. Elementary proofs of the prime number theorem.

Selberg's formula (as discussed in (0.3.1)) can be written as

$$(\psi(x)-x)\log x = -\sum_{p\le x}\log p\ \left(\psi\left(\frac{x}{p}\right)-\frac{x}{p}\right)+O(x). \tag{3.5.1}$$

There is an analogous formula for $\mu(n)$, derived from (3.1.1):

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

Exercise: Show that if $F(x)$ is any function for which

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

holds for all x, then

$$\liminf_{x\to\infty}\frac{F(x)}{x}+\limsup_{x\to\infty}\frac{F(x)}{x}=0,$$

and so if $\lim_{x\to\infty}F(x)/x$ exists then it equals 0. Note that these deductions apply to $M(x)$ and $\psi(x)-x$, given the formulas in the last two displayed equations.

Proof of Selberg's formula. Let $\Lambda_2(n):=\Lambda(n)\log n+\sum_{\ell m=n}\Lambda(\ell)\Lambda(m)$, so that

$$\begin{aligned}\sum_{d|n}\Lambda_2(n) &= \sum_{\ell|n}\Lambda(\ell)\log\ell+\sum_{\ell m|n}\Lambda(\ell)\Lambda(m)\\ &= \sum_{\ell|n}\Lambda(\ell)\left(\log\ell+\sum_{m|(n/\ell)}\Lambda(m)\right)=(\log n)^2,\end{aligned}$$

and therefore

$$\Lambda_2(n)=\sum_{d|n}\mu(d)(\log n/d)^2$$

by Möbius inversion.

Now

$$\begin{aligned}\sum_{m\le x}\ \sum_{n\le x/m}\Lambda_2(n) &= \sum_{mn\le x}\ \sum_{dr=n}\mu(d)(\log r)^2=\sum_{dmr\le x}\mu(d)(\log r)^2\\ &= \sum_{Nr\le x}(\log r)^2\sum_{d|N}\mu(d)=\sum_{r\le x}(\log r)^2\\ &= x(\log^2 x-2\log x+2)+O(\log^2 x).\end{aligned}$$

Moreover by (1.2.1) we obtain

$$2\sum_{m\le x}\frac{x}{m}\log\frac{x}{m}=2x\int_1^x\sum_{m\le t}\frac{1}{m}\frac{dt}{t}=x(\log^2 x+2\gamma\log x+c_0)+O(1),$$

for some constant c_0, and

$$\sum_{m\leq x}\left(c_1\frac{x}{m}+c_2\right)=x(c_1\log x+2-c_0)+O(1),$$

with $c_1=-2(1+\gamma)$ and $c_2=2-c_0-c_1\gamma$.

Therefore if $A(x):=\sum_{n\leq x}\Lambda_2(n)-2x\log x-c_1x-c_2$, and then

$$B(x):=\sum_{m\leq x}A(x/m)\ll\log^2 x,$$

so that

$$A(x)=\sum_{n\leq x}\mu(n)B(x/n)\ll\sum_{n\leq x}\log^2(x/n)\ll x.$$

Therefore we have proved

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

The result, (0.3.1) or (3.5.1), follows from (1.2.3) and the bounds

$$\sum_{n\leq x}\Lambda(n)\log x/n\ll x\quad\text{and}\quad\sum_{\substack{\ell\leq x\\ \ell=p^b,\ b\geq 2}}\Lambda(\ell)\psi\left(\frac{x}{\ell}\right)\ll x,$$

which follow from (1.1.4).

4. Distances

4.1. Distance functions. Throughout we define

$$\mathbb{U} := \{|z| \leq 1\} \text{ and } \mathbb{U}^{\mathbb{N}} = \{\mathbf{z} = (z_1, z_2, \dots) : \text{Each } z_i \in \mathbb{U}\}.$$

Lemma 4.1.1. *The function* $\eta : \mathbb{U} \to \mathbb{R}_{\geq 0}$ *given by* $\eta(z)^2 = 1 - \text{Re}(z)$ *satisfies*

$$\eta(yz) \leq \eta(y) + \eta(z) \text{ for all } y, z \in \mathbb{U}.$$

Proof. Let $y = e^{2i\pi\varphi}$ and $z = e^{2i\pi\theta}$. Since $1 - \text{Re}(e^{2i\pi\alpha}) = 2\sin^2(\pi\alpha)$, for any real α we deduce that

$$\begin{aligned}\eta(yz)/\sqrt{2} &= |\sin(\pi(\theta+\varphi))| \leq |\sin(\pi\theta)\cos(\pi\varphi)| + |\sin(\pi\varphi)\cos(\pi\theta)| \\ &\leq |\sin(\pi\theta)| + |\sin(\pi\varphi)| = (\eta(y) + \eta(z))/\sqrt{2}.\end{aligned}$$

This settles the case where $|z| = |w| = 1$, and (Exercise) one can extend this to all pairs $z, w \in \mathbb{U}$.

We can define a product in $\mathbb{U}^{\mathbb{N}}$ by multiplying componentwise: that is,

$$\mathbf{y} \times \mathbf{z} = (y_1 z_1, y_2 z_2, \dots).$$

Lemma 4.1.2. *Suppose we have a sequence of functions*

$$\eta_j : \mathbb{U} \to \mathbb{R}_{\geq 0} \text{ for which } \eta_j(yz) \leq \eta_j(y) + \eta_j(z) \text{ for any } y, z \in \mathbb{U}.$$

Then we may define a 'norm' on $\mathbb{U}^{\mathbb{N}}$ *by setting*

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

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

$$\|\mathbf{y} \times \mathbf{z}\| \leq \|\mathbf{y}\| + \|\mathbf{z}\|. \tag{4.1.1}$$

Proof. Indeed we have

$$\begin{aligned}\|\mathbf{y} \times \mathbf{z}\|^2 &= \sum_{j=1}^{\infty} \eta_j(y_j z_j)^2 \leq \sum_{j=1}^{\infty} (\eta_j(y_j)^2 + \eta_j(z_j)^2 + 2\eta_j(y_j)\eta_j(z_j)) \\ &\leq \|\mathbf{y}\|^2 + \|\mathbf{z}\|^2 + 2\Big(\sum_{j=1}^{\infty} \eta_j(y_j)^2\Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{\infty} \eta_j(z_j)^2\Big)^{\frac{1}{2}} = (\|\mathbf{y}\| + \|\mathbf{z}\|)^2,\end{aligned}$$

using the Cauchy-Schwarz inequality, which implies (4.1.1).

A nice class of examples is provided by taking $\eta_j(z_j)^2 = a_j(1-\mathrm{Re}(z_j))$ (as in Lemma 4.1.1) where the a_j are non-negative constants with $\sum_{j=1}^{\infty} a_j < \infty$. This last condition ensures the convergence of the sum in the definition of the norm. A specific case of this, for a multiplicative function $f(n)$, is to define the coefficient $a_j = 1/p$ and let $z_j = f(p)$ for each prime $p \le x$. We therefore have the norm

$$\|f_x\|^2 := \sum_{p\le x} \frac{1-\mathrm{Re}(f(p))}{p},$$

where f_x corresponds to f truncated at x. We can extend this to define the distance (up to x) between the multiplicative functions f and g as

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

By Lemmas 4.1.1 and 4.1.2 this satisfies the triangle inequality

$$\mathbb{D}(f_1,g_1;x)+\mathbb{D}(f_2,g_2;x) \ge \mathbb{D}(f_1g_1,f_2g_2;x). \tag{4.1.2}$$

(We might alternately use $\mathbb{D}^+(f,g;x)^2 = \sum_{p^k\le x} \frac{1-\mathrm{Re}\ f(p^k)\overline{g(p^k)}}{p^k}$, though this can lead to complications.)

For $x\ge 3$ and $T \ge 1$ define $t(x,T) = t_f(x,T)$ to be a value of t with $|t|\le T$ for which $\mathbb{D}(f(n),n^{it};x)^2$ is minimized; and then define

$$M(x,T) = M_f(x,T) := \min_{|t|\le T} \mathbb{D}(f(n),n^{it};x)^2 = \mathbb{D}(f(n),n^{it(x,T)};x)^2$$

Now

$$M(x,T) = \sum_{p\le x} \frac{1-\mathrm{Re}\ f(p)p^{-it(x,T)}}{p} \le \sum_{p\le x} \frac{1-\mathrm{Re}\ f(p)p^{-it}}{p}$$

for all t with $|t|\le T$. Therefore if $x>y$ then

$$\sum_{y<p\le x} \frac{1-\mathrm{Re}\ f(p)p^{-it(x,T)}}{p} \le M(x,T)-M(y,T) \le \sum_{y<p\le x} \frac{1-\mathrm{Re}\ f(p)p^{-it(y,T)}}{p},$$

and so, by (1.2.4),

$$|M(x,T)-M(y,T)| \le 2\log\left(\frac{\log x}{\log y}\right) + O(1). \tag{4.1.3}$$

We will need some further definitions for a given multiplicative function f: For any complex number s with $\mathrm{Re}(s)>0$, let

$$F(s) = F(s;x) := \prod_{p\le x}\left(1+\frac{f(p)}{p^s}+\frac{f(p^2)}{p^{2s}}+\dots\right).$$

Exercise: Show that $|F(1,x)| \asymp (\log x)e^{-\|f_x\|^2}$; and that $|F(1+it,x)| \asymp (\log x)e^{-\mathbb{D}(f(n),n^{it};x)^2}$.

Now define

$$L = L(x,T) := \frac{1}{\log x}\Big(\max_{|t|\le T} |F(1+it)|\Big).$$

Exercise: Show that $M(x,T) = \log(1/L(x,T)) + O(1)$.

4.2. A lower bound on a key distance.

Lemma 4.2.1. *If* $|t| \leq x^{o(1)}$ *then*

$$\mathbb{D}^2(\mu(n), n^{it}; x) \geq \left\{1 - \frac{2}{\pi} + o(1)\right\} \log\left(\frac{\log x}{\log(2+|t|)}\right).$$

Proof. Fix $\alpha \in [0,1)$ and $\epsilon > 0$. Let P be the set of primes for which there exists an integer n such that $p \in I_n := [e^{2\pi(n+\alpha)/|t|}, e^{2\pi(n+\alpha+\epsilon)/|t|})$, so that $\mathrm{Re}(p^{it})$ lies between $\cos(2\pi\alpha)$ and $\cos(2\pi(\alpha+\epsilon))$. We partition the intervals I_n into subintervals of the form $[y, y+z]$, where $z = o(y)$ and $\log z \sim \log y$, which is possible provided $|t| = o(n/\log n)$ (Exercise). The Brun-Titchmarsh Theorem implies that the number of primes in each such interval is $\leq \{2+o(1)\}z/\log y$, and so $\sum_{p\in I_n} 1/p \leq \{2+o(1)\}\log(1+\frac{\epsilon}{n+\alpha})$, from which we deduce

$$\sum_{\substack{x_0<p\leq x\\ p\in I_n \text{ for some } n}} \frac{1}{p} \leq \{2\epsilon + o(1)\}\log\left(\frac{\log x}{\log x_0}\right) + O(\epsilon),$$

where $x_0 := (2+|t|)^{\log u} e^{2\pi/|t|}$ and $2+|t| = x^{1/u}$, as $u \to \infty$. Combining this with (1.2.4), we deduce (exercise) that

$$\begin{aligned}\sum_{x_0<p\leq x} \frac{1+\cos(t\log p)}{p} &\geq \{2+o(1)\}\log\left(\frac{\log x}{\log x_0}\right)\int_{1/4}^{3/4}(1+\cos(2\pi\alpha))d\alpha + O(1)\\ &\geq \left\{1-\frac{2}{\pi}+o(1)\right\}\log\left(\frac{\log x}{\log x_0}\right) + O(1).\end{aligned}$$

The result follows if $|t| \geq 1$. If $|t| < 1$ then $\log\left(\frac{\log x}{\log x_0}\right) \sim \log(|t|\log x)$. However, we also have

$$\sum_{p\leq e^{2\pi/3|t|}} \frac{1+\cos(t\log p)}{p} \geq (1+\cos(2\pi/3)) \sum_{p\leq e^{2\pi/3|t|}} \frac{1}{p} \geq \frac{1}{2}\log\frac{1}{|t|} + O(1),$$

by (1.2.4), and then adding these lower bounds gives the result.

5. Zeta Functions and Dirichlet series: A minimalist discussion

5.1. Dirichlet characters and Dirichlet L-functions.

`This section maybe may be mostly discarded, though we may have to wait to see how other things pan out.`

Define the Dirichlet characters, especially the role of primitive characters, and L-functions to the right of the 1-line. Note that

(5.1.1) $$\frac{1}{\varphi(q)}\sum_{\chi \pmod q}\chi(m) = \begin{cases} 1 & \text{if } m=1,\\ 0 & \text{otherwise,}\end{cases}$$

and

$$\frac{1}{\varphi(q)} \sum_{b \pmod q} \chi(b) = \begin{cases} 1 & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise,} \end{cases} \tag{5.1.2}$$

where χ_0 is the principal character mod q.

We will need to add a proof of Dirichlet's class number formula, perhaps a uniform version? (Since this can be used to establish the connection between small class number and small numbers of primes in arithmetic progressions). We also need to discuss the theory of binary quadratic forms, at least enough for the class number formula and to understand prime values of such forms.

Lemma 5.1.1. *For any non-principal Dirichlet character* $\chi \pmod q$ *and any complex number* s *with real part* > 0, *we can define*

$$L(s,\chi) = \lim_{N\to\infty} \sum_{n=1}^{N} \frac{\chi(n)}{n^s},$$

since this limit exists.

The content of this result is that the right-side of the equation converges. One usually uses the idea of analytic continuation to state that this equals the left-side.

Proof sketch. We will prove this by suitably bounding

$$\sum_{n=N+1}^{\infty} \frac{\chi(n)}{n^s},$$

for $N \geq (q(1+|s|))^{2/\sigma}$, where $s = \sigma + it$. If $n = N+j$ we replace the n in the denominator by N, incurring an error of

$$\left| \frac{1}{(N+j)^s} - \frac{1}{N^s} \right| \ll \frac{1}{N^\sigma} \frac{|s|j}{N} \ll \frac{|s|q}{N^{1+\sigma}},$$

for $1 \leq j \leq q$. Summing this over all n in the interval $(N, N+q]$, gives $N^{-s} \sum_n \chi(n) + O(|s|q^2/N^{1+\sigma}) \ll |s|q^2/N^{1+\sigma}$. Summing now over $N, N+q, N+2q, \ldots$, we obtain a total error of $\ll |s|q/\sigma N^\sigma$, which implies the result.

Exercise. Actually this proof is not complete. Find the error and correct it.

Lemma 5.1.2. *For any complex number* s *with real part* > 0 *we can define*

$$\zeta(s) = \frac{1}{s-1} - s \int_1^\infty \frac{\{t\}}{t^{s+1}} dt,$$

where $\{t\}$ *is the fractional part of* t, *since this integral is absolutely convergent.*

Proof. We simply use partial summation so that for $\mathrm{Re}(s) > 1$ we have

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = \int_1^\infty \frac{d[t]}{t^s} = \int_1^\infty \frac{d(t - \{t\})}{t^s} = \frac{1}{s-1} - \int_1^\infty \frac{d\{t\}}{t^s}.$$

Integrating by parts

$$\int_1^\infty \frac{d\{t\}}{t^s} = \left[\frac{\{t\}}{t^s} \right]_1^\infty + s \int_1^\infty \frac{\{t\}}{t^{s+1}} dt = s \int_1^\infty \frac{\{t\}}{t^{s+1}} dt,$$

and the result follows.

5.2. Dirichlet series just to the right of the 1-line.

Lemma 5.2.1. *For $x \geq 2$ we have*

$$\exp\left(\sum_{p\leq x}\frac{f(p)}{p^{1+it}}\right) \asymp \sum_{n\geq 1}\frac{f(n)}{n^{1+\frac{1}{\log x}+it}}$$

Proof. Let $s = 1 + \frac{1}{\log x} + it$. By (1.2.3) we deduce that

$$\left|\sum_{p>x}\frac{f(p)}{p^s}\right| \leq \sum_{p>x}\frac{1}{p^{1+1/\log x}} \ll 1,$$

and

$$\left|\sum_{p\leq x}\left(\frac{f(p)}{p^s} - \frac{f(p)}{p^{1+it}}\right)\right| \leq \sum_{p\leq x}\frac{1}{p}(1-p^{-1/\log x}) \ll \frac{1}{\log x}\sum_{p\leq x}\frac{\log p}{p} \ll 1,$$

and so

$$\sum_{p\leq x}\frac{f(p)}{p^{1+it}} = \sum_{p\ \text{prime}}\frac{f(p)}{p^s} + O(1).$$

The result follows by, adding the relevant terms for $p^k, k \geq 2$ to the right side, which converge, and then taking the exponential of each side.

Corollary 5.2.2. *Suppose that there exists an integer $k \geq 1$ such that $f(p)^k = 1$ for all primes p. Then $\mathbb{D}(f(n), n^{it}; \infty) = \infty$ for every non-zero real t.*

Examples of this include $f = \mu$ the Möbius function, χ a Dirichlet character, and even $\mu\chi$.

Proof. Suppose that there exists a real number $t \neq 0$ such that $\mathbb{D}(f(n), n^{it}; \infty) < \infty$. Then $\mathbb{D}(1, n^{ikt}; \infty) \leq k\mathbb{D}(f(n), n^{it}; \infty) < \infty$ by the triangle inequality (4.1.2). Let $s = 1 + \frac{1}{\log x} + ikt$. By Lemma 5.2.1, we have

$$\log\zeta(s) = \sum_{p\leq x}\frac{1}{p^{1+ikt}} + O(1),$$

and so

$$\begin{aligned}\log|\zeta(s)| = \mathrm{Re}(\log\zeta\,(s)) &= \sum_{p\leq x}\frac{\mathrm{Re}(p^{ikt})}{p} + O(1)\\ &= \sum_{p\leq x}\frac{1}{p} - \mathbb{D}(1, n^{ikt}; x) + O(1) = \log\log x + O_t(1),\end{aligned}$$

and therefore $|\zeta(s)| \gg \log x$. However Lemma 5.1.2 yields that

$$\zeta(s) = \frac{1}{s-1} + O(1+|t|) = \frac{1}{it} + O\left(1 + |t| + \frac{1}{|t|^2\log x}\right),$$

a contradiction.

Lemma 5.2.3. *Let a_n be a sequence of complex numbers such that $\sum_{n=1}^{\infty}\frac{|a_n|}{n}<\infty$, so that $A(s)=\sum_{n=1}^{\infty}a_n n^{-s}$ is absolutely convergent in $Re(s)\geq 1$. For all real numbers $T\geq 1$, and all $0\leq\alpha\leq 1$ we have*

$$\max_{|t|\leq T}|A(1+\alpha+it)|\leq\max_{|u|\leq 2T}|A(1+iu)|+O\Big(\frac{\alpha}{T}\sum_{n=1}^{\infty}\frac{|a_n|}{n}\Big). \tag{5.2.1}$$

Exercise. Prove that, for any integer $n\geq 1$, we have

$$n^{-\alpha}=\frac{1}{\pi}\int_{-T}^{T}\frac{\alpha}{\alpha^2+\xi^2}n^{-i\xi}d\xi+O\Big(\frac{\alpha}{T}\Big).$$

(Hint: Show that $\frac{2\alpha}{\alpha^2+\xi^2}$ is the Fourier transform of $e^{-\alpha|z|}$.)

Proof. Multiplying the result in this exercise through by a_n/n^{1+it}, and summing over all n, we obtain

$$A(1+\alpha+it)=\frac{1}{\pi}\int_{-T}^{T}\frac{\alpha}{\alpha^2+\xi^2}A(1+it+i\xi)d\xi+O\Big(\frac{\alpha}{T}\sum_{n=1}^{\infty}\frac{|a_n|}{n}\Big)$$

which yields the result when $|t|\leq T$, since then $|u|\leq|t|+|\xi|\leq 2T$ for $u=t+\xi$, and as $\frac{1}{\pi}\int_{-T}^{T}\frac{\alpha}{\alpha^2+\xi^2}d\xi\leq\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\alpha}{\alpha^2+\xi^2}d\xi=1$ by the exercise with $n=1$.

Lemma 5.2.4. *If χ is a character mod q, with $x\geq y\geq q$ and $|t|\leq y^{O(1)}$, then*

$$\left|L\left(1+\frac{1}{\log x}+it,\chi\right)\right|\ll\left|L\left(1+\frac{1}{\log y}+it,\chi\right)\right|$$

`It would be good to have a proof of this that stays to the right of the 1-line, and does not use the analytic continuation. Here are two proofs with not too much in them.`

Proof # 1. Note that $1<1+\frac{1}{\log x}\leq 1+\frac{1}{\log y}\leq 1+\frac{1}{\log q}$. By (2) and the next two displayed equations of Chapter 14 of Davenport, we know that for $s=\sigma+it$ where $\sigma>1$, we have[12]

$$-\text{Re }\frac{L'(s,\chi)}{L(s,\chi)}\ll\log(q(2+|t|)). \tag{5.2.2}$$

Therefore

$$\log\left|\frac{L\left(1+\frac{1}{\log x}+it,\chi\right)}{L\left(1+\frac{1}{\log y}+it,\chi\right)}\right|=\text{Re}\left(\log\left(\frac{L\left(1+\frac{1}{\log x}+it,\chi\right)}{L\left(1+\frac{1}{\log y}+it,\chi\right)}\right)\right)$$
$$=-\int_{1/\log x}^{1/\log y}\text{Re }\frac{L'(1+v+it,\chi)}{L(1+v+it,\chi)}dv\ll\frac{\log(q(2+|t|))}{\log y}\ll 1.$$

[12]This is the proof from Lemma 1 [Elliott6]. The key is (5.2.2) – can we prove it without zeros? In Lemma 14 he gives a proof with limited analytic continuation.

Proof # 2. It is well-known that the completed Dirichlet L-function has a Hadamard factorization; that is if $\delta = (1-\chi(-1))/2$ then

$$\Lambda(s,\chi) := \left(\frac{\pi}{q}\right)^{-\frac{s+\delta}{2}} \Gamma\left(\frac{s+\delta}{2}\right) L(s,\chi) = e^{A+Bs} \prod_{\rho:\ \Lambda(\rho,\chi)=0} \left(1-\frac{s}{\rho}\right) e^{s/\rho}$$

where $\mathrm{Re}(B + \sum_\rho 1/\rho) = 0$ (as in Chapter 12 of Davenport). Since all zeros ρ of $L(s,\chi)$ have $\mathrm{Re}(\rho) \leq 1$ we see that $\left|1+\frac{1}{\log y}+it-\rho\right| \geq \left|1+\frac{1}{\log x}+it-\rho\right|$, by applying the triangle inequality. Hence $\left|\Lambda(1+\frac{1}{\log y}+it,\chi)\right| \geq \left|\Lambda(1+\frac{1}{\log x}+it,\chi)\right|$ by multiplying over all zeros. Inserting this into the definition of $\Lambda(s,\chi)$, we deduce the result from the fact that $\Gamma'(s)/\Gamma(s) = \log s + O(1/|s|)$ (as in (6) of Chapter 10 of Davenport), which implies that the ratio of the Gamma factors is $\ll \log|t|/\log y \ll 1$.

6. Halász's Theorem

6.1. Halász's Theorem in context. In §$1\frac{1}{2}$.2 we gave a simple heuristic that for any real-valued multiplicative function f with $|f(n)| \leq 1$ for all n, we have

$$\frac{1}{x}\sum_{n\leq x} f(n) \to \mathrm{Prod}(f,\infty) \quad \text{as } x\to\infty. \tag{2.2.1}$$

However, not all complex valued multiplicative functions have a mean value tending to a limit. For example, the function $f(n) = n^{it}$, with $t \in \mathbb{R}\setminus\{0\}$, since

$$\frac{1}{x}\sum_{n\leq x} n^{it} = \frac{1}{x}\int_{u=1}^{x} u^{it}du + O(1) \sim \frac{x^{it}}{1+it}.$$

Notice that for large x this mean value rotates around a path getting closer and closer to the circle of radius $1/(1+t^2)^{1/2}$. And it is not only the functions n^{it} whose mean value does not tend to a limit – indeed you might expect a similar phenomena if you make some minor alterations to n^{it}. For example, if we take $f(2) = -2^{it}$ and $f(p) = p^{it}$ for all odd primes p, then the mean-value up to x is $\sim x^{it}/3(1+it)$. So we find that any f that is "close" to n^{it} also has this property. We measure "close" by using the distance function, that is $\mathbb{D}(f(n), n^{it}; \infty)$ is bounded. In this case we say that $f(n)$ is n^{it}*-pretentious*. Are there any non-pretentious multiplicative functions whose mean values do not exist?

In the early seventies, Gábor Halász [8,9] brilliantly proved that the answer is "no":

Halász's Theorem. *Suppose that $f(.)$ is a multiplicative function with $|f(n)| \leq 1$ for all n. If $\lim_{x\to\infty}\frac{1}{x}\sum_{n\leq x} f(n)$ does not exist then $f(n)$ is n^{it}-pretentious for some real number t. That is,*

$$\mathbb{D}(f(n), n^{it}; \infty) \quad \textit{is bounded.}$$

Halász's theorem gives more, both qualitatively and quantitatively:

$$\lim_{x\to\infty}\frac{1}{x}\Big|\sum_{n\leq x} f(n)\Big| \ \textit{exists.}$$

If $f(n)$ is not n^{it}-pretentious then $\lim_{x\to\infty}\frac{1}{x}\sum_{n\leq x} f(n) = 0$.

Exercise: Prove that there can be at most one value of t for which $f(n)$ is n^{it}-pretentious.

In fact, if $f(n)$ is n^{it}-pretentious then

$$\frac{1}{x}\sum_{n\leq x} f(n) \sim \frac{x^{it}}{1+it}\,\mathrm{Prod}(f_t, x)$$

where $f_t(n) := f(n)/n^{it}$. This converges in absolute value, and the angle varies slowly.[13]

Although Halász's result is a little technical, it does indicate how rapidly mean values converge: We have

$$\frac{1}{x}\Big|\sum_{n\leq x} f(n)\Big| \ll (1+M_f(x,T))e^{-M_f(x,T)} + \frac{1}{T} + \frac{\log\log x}{\log x} \tag{6.1.2}$$

[13] That is the argument for the mean values at x and at x^c, differ by $< \epsilon$, once x is sufficiently large.

This can be formulated a little differently. At the end of section 4.1 we saw that $M_f(x,T) = \log(1/L(x,T)) + O(1)$, so that (6.1.2) is equivalent to

$$\frac{1}{x}\Big|\sum_{n\le x} f(n)\Big| \ll L\log\frac{2}{L} + \frac{1}{T} + \frac{\log\log x}{\log x}, \tag{6.1.3}$$

where $L = L(x,T)$.

Exercise: Taking $f(n) = \chi(n)$, a Dirichlet character, deduce the following: For $0 < \epsilon < \log\log x/\log x$ we have

$$\Big|\sum_{n\le x}\chi(n)\Big| \le \epsilon x$$

except perhaps if $|L(1+it,\chi)| \gg \epsilon' \log x$ for some $t, |t| \ll 1/\epsilon$ where $\epsilon' = \epsilon/\log(1/\epsilon)$.

6.2. Inverse and Hybrid results, etc.

Theorem 6.2. *There exists a constant $c > 0$ such that there exists y in the range $x^{\eta/|\log\eta|} \le y \le x$ for which*

$$\left|\sum_{n\le y} f(n)\right| > \eta y, \text{ where } \eta = \eta(x,T) := \frac{cL(x,T)}{1+|t|}.$$

and $t = t(x,T)$.

Proof. Let $\varphi = \log(1/\eta)$, $\tau = \eta/\varphi$ and $\delta = \varphi/\log x$. Now

$$\begin{aligned}\left|\sum_{n\ge 1}\frac{f(n)}{n^{1+\delta+it}}\right| &= \left|\prod_{p\le x}\left(1-\frac{f(p)}{p^{1+\delta+it}}\right)^{-1}\right| \\ &\asymp \prod_{p\le x}\left(1-\frac{1}{p^{1+\delta}}\right)^{-1}\exp\left(-\sum_{p\le x}\frac{1-\mathrm{Re}(f(p)/p^{it})}{p^{1+\delta}}\right) \\ &\asymp \frac{\log x}{\varphi}\, e^{-M(x,T)} \asymp \frac{\log x}{\varphi} L(x,T),\end{aligned}$$

by the prime number theorem. On the other hand

$$\sum_{n\ge 1}\frac{f(n)}{n^{1+\delta+it}} = (1+\delta+it)\int_1^\infty \frac{1}{y^{2+\delta+it}}\sum_{n\le y} f(n)dy.$$

Assuming that $|\sum_{n\le y} f(n)| \le \eta y$ for all $x^{\eta/\varphi} \le y \le x$, and using the the trivial bound $|\sum_{n\le y} f(n)| \le y$ otherwise, we find that the integral here is

$$\begin{aligned}&\le \int_1^{x^{\eta/\varphi}} \frac{dy}{y^{1+\delta}} + \eta\int_{x^{\eta/\varphi}}^x \frac{dy}{y^{1+\delta}} + \int_x^\infty \frac{dy}{y^{1+\delta}} \\ &= \frac{\log x}{\varphi}\left((1-e^{-\eta}) + \eta(e^{-\eta}-e^{-\varphi}) + e^{-\varphi}\right) \le 3\eta\,\frac{\log x}{\varphi},\end{aligned}$$

which yields a contradiction if c is chosen sufficiently small.

There is a version of Halász's theorem that takes into account the point $1 + it$:

Theorem 6.3. *Let $t = t(x, \log x)$ and let $L = L(x, \log x)$. Then*

$$\frac{1}{x}\Big|\sum_{n\le x} f(n)\Big| \ll \frac{L}{1+|t|}\log\frac{2}{L} + \frac{\log\log x}{(\log x)^{2-\sqrt{3}}}.$$

We require the following lemma, which relates the mean value of $f(n)$ to the mean-value of $f(n)n^{it}$.

Lemma 6.4. *Suppose $f(n)$ is a multiplicative function with $|f(n)| \le 1$ for all n. Then for any real number t we have*

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

Corollary 6.5. *Suppose $f(n)$ is a multiplicative function with $|f(n)| \le 1$ for all n. If $t = t_f(x, \log x)$ then*

$$\sum_{n\le x} f(n) = \frac{x^{it}}{1+it}\sum_{n\le x}\frac{f(n)}{n^{it}} + O\bigg(\frac{x\log\log x}{(\log x)^{2-\sqrt{3}}}\bigg).$$

Proof of Lemma 6.4. Let g and h denote the multiplicative functions defined by $g(n) = f(n)/n^{it}$, and $h(p^k) = g(p^k) - g(p^{k-1})$, so that $g(n) = \sum_{d|n} h(d)$. Then

$$\sum_{n\le x} f(n) = \sum_{n\le x} g(n)n^{it} = \sum_{n\le x} n^{it}\sum_{d|n} h(d) = \sum_{d\le x} h(d)d^{it}\sum_{m\le x/d} m^{it}. \tag{6.2.1}$$

By partial summation it is easy to see that

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

We use the first estimate above in (6.2.1) when $d \le x/(1+t^2)$, and the second estimate when $x/(1+t^2) \le d \le x$. This gives

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

Applying (2.4.5) and (2.4.6) we deduce that

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

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.

Then, by the Cauchy-Schwarz inequality,

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

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

Proof of Corollary 6.5 and Theorem 6.3. We may assume that $M := M_f(x,\log x) > (2-\sqrt{3})\log\log x$ else Corollary 6.5 follows immediately from Lemma 6.4. Moreover, in this case $\sum_{n\le x} f(n) \ll x\log\log x/(\log x)^{2-\sqrt{3}}$ by (6.1.2). Now let $g(n)=f(n)/n^{it}$. If $|t| > \frac{1}{2}\log x$ then $|(x^{it}/(1+it))\sum_{n\le x} g(n)| \le x/(1+|t|) \ll x/\log x$ and Corollary 6.5 follows. But if $|t| > \frac{1}{2}\log x$ then $t_g(x,\frac{1}{2}\log x)=0$, so that $M_g(x,\frac{1}{2}\log x)=M$, and Corollary 6.5 follows from (6.1.2) applied to g.

Finally Theorem 6.3 follows from Corollary 6.5 by the definition of L.

6.4. Halász's key Proposition. The main result is the following:

Proposition 1. *Let f be a multiplicative function with $|f(n)|\le 1$ for all n. Let $x\ge 3$ and $T\ge 1$ be real numbers. F be as in Theorem 1. Then*

$$\frac{1}{x}\Big|\sum_{n\le x} f(n)\Big| \ll \frac{1}{\log x}\int_{1/\log x}^{1}\max_{|t|\le T}|F(1+\alpha+it)|\ \frac{d\alpha}{\alpha}+\frac{1}{T}+\frac{\log\log x}{\log x}\ .$$

To evaluate the integral we use Lemma 2.7.

Proof of (6.1.3). We will bound the terms in the integral in Proposition 1 using Lemma 2.7. Let a_n be the multiplicative function with $a_{p^k}=f(p^k)$ if $p\le x$ and $a_{p^k}=0$ so that $\sum_n |a_n|/n \le \prod_{p\le x}(1-1/p)^{-1} \ll \log x$

Let $M=\max_{|t|\le 2T}|F(1+it)|$ so that $M=L\log x$. If $M<1$ then (6.1.3) follows immediately by Proposition 1 and Lemma 2.7. If $M>1$ then, by Lemma 2.7, for $1/\log x \le \alpha \le 1/M$, we have

$$\max_{|t|\le T}|F(1+\alpha+it)| \le M+O\Big(\frac{\alpha\log x}{T}\Big).$$

Moreover, for any real t and $1/M<\alpha\le 1$, we have

$$|F(1+\alpha+it)| \le \zeta(1+\alpha)=\frac{1}{\alpha}+O(1)\ll\frac{1}{\alpha},$$

by taking the absolute value of each summand. The result follows from Proposition 1.

6.5. The proof of Proposition 1.

Proof of Proposition 1. By (1.2.6) we know that

$$S(N) := \sum_{n\le N} f(n) = \frac{1}{\log N}\sum_{n\le N} f(n)\log n + O\left(\frac{N}{\log N}\right)$$
$$\ll \int_{1/\log x}^{1} \left|\sum_{n\le N} f(n)\log n\right| \frac{d\alpha}{N^{2\alpha}} + \frac{N}{\log N}$$

whenever $x \ge \max\{N, 4\}$. Therefore

$$\int_2^x \frac{|S(y)|}{y^2}dy \ll \int_{1/\log x}^{1}\left(\int_2^x \Big|\sum_{n\le y} f(n)\log n\Big|\frac{dy}{y^{2+2\alpha}}\right)d\alpha + \log\log x.$$

Applying Cauchy's inequality twice we obtain, for $\alpha \ge 1/\log x$,

$$\left(\int_2^x \Big|\sum_{n\le y} f(n)\log n\Big|\frac{dy}{y^{2+2\alpha}}\right)^2 \le \Big(\int_1^x \frac{dy}{y^{1+2\alpha}}\Big)\Big(\int_2^x \Big|\sum_{n\le y} f(n)\log n\Big|^2\frac{dy}{y^{3+2\alpha}}\Big)$$
$$\ll \frac{1}{\alpha}\int_1^\infty \Big|\sum_{n\le y} f(n)\log n\Big|^2\frac{dy}{y^{3+2\alpha}}$$
$$= \frac{1}{\alpha}\int_0^\infty \Big|\sum_{n\le e^t} f(n)\log n\Big|^2 e^{-2(1+\alpha)t}dt$$
$$= \frac{1}{2\pi\alpha}\int_{-\infty}^\infty \Big|\frac{F'(1+\alpha+iy)}{1+\alpha+iy}\Big|^2 dy,$$

by Plancherel's formula.

The integral in the region $|y| \le T$ is evidently

$$\le \max_{|y|\le T}|F(1+\alpha+iy)|^2\int_{-T}^{T}\Big|\frac{(F'/F)(1+\alpha+iy)}{1+\alpha+iy}\Big|^2 dy.$$

This integral is, by Plancherel,

$$\le \int_{-\infty}^{\infty}\Big|\frac{(F'/F)(1+\alpha+iy)}{1+\alpha+iy}\Big|^2 dy = \int_1^\infty \Big|\sum_{n\le y} f(n)\Lambda(n)\Big|^2\frac{dy}{y^{3+2\alpha}} \ll \int_1^\infty \frac{dy}{y^{1+2\alpha}} \ll \frac{1}{\alpha}.$$

For $|y| > T$, we expand the integral to obtain

$$\sum_{m,n\ge 1}\frac{f(m)\overline{f(n)}\log m\log n}{(mn)^{1+\alpha}}\int_{|y|>T}\frac{1}{|1+\alpha+iy|^2(m/n)^{iy}}dy. \tag{6.5.1}$$

If $m = n$ the integral is $\ll 1/T$. Otherwise, partitioning the range into intervals of length $2\pi/|\log(m/n)|$ we deduce that the integral is $\ll 1/T^2|\log(m/n)|$. Hence the above is

$$\ll \frac{1}{T}\sum_{n\geq 1}\frac{(\log n)^2}{n^{2+2\alpha}} + \frac{1}{T^2}\sum_{m>n\geq 1}\frac{\log m \log n}{(mn)^{1+\alpha}\log(m/n)}.$$

The first sum is bounded. For the second sum we consider the sum over m for n fixed, breaking the sum into those with $m = n + j, 1 \leq j \leq n$, then $m = in + \ell, 1 \leq \ell \leq n$, and finally $m \geq n^2$:

$$\begin{aligned}\sum_{m>n\geq 1}\frac{\log m}{m^{1+\alpha}\log(m/n)} &\ll \sum_{j=1}^{n}\frac{\log n}{jn^\alpha} + n\sum_{i=1}^{n}\frac{\log n}{(in)^{1+\alpha}\log 2i} + \sum_{m>n^2}\frac{1}{m^{1+\alpha}} \\ &\ll \frac{(\log n)^2}{n^\alpha} + \frac{1}{\alpha n^{2\alpha}}.\end{aligned}$$

Hence in total we have

$$\ll \sum_{n\geq 1}\frac{(\log n)^3}{n^{1+2\alpha}} + \frac{\log n}{\alpha n^{1+3\alpha}} \ll \frac{1}{\alpha^4}, \tag{6.5.2}$$

yielding a bound of $\ll \frac{1}{T} + \frac{1}{\alpha^4 T^2}$ for (6.5.1).

Substituting this all in above yields

$$\int_2^x \frac{|S(y)|}{y^2}dy \ll \int_{1/\log x}^{1}\max_{|t|\leq T}|F(1+\alpha+it)|\ \frac{d\alpha}{\alpha} + \frac{(\log x)^{3/2}}{T} + \log\log x,$$

which implies Proposition 1, with $\frac{(\log x)^{1/2}}{T}$ in place of $\frac{1}{T}$, using (2.4.4).

To improve the error term in Proposition 1 from $\frac{(\log x)^{1/2}}{T}$ to $\frac{1}{T}$, we now improve (6.5.2) to $\ll 1/\alpha^3$

Strong Hilbert's Inequality. *If $a_1, a_2, \ldots$ is a sequence of complex numbers, and $x_1, x_2, \ldots$ are distinct real numbers then*

$$\left|\sum_{r\neq s}\frac{a_r\overline{a_s}}{x_r - x_s}\right| \ll \sum_r \frac{|a_r|^2}{\min_{s\neq r}|x_s - x_r|},$$

provided the right side converges.

Corollary. *If* $a_1, a_2, \dots$ *is a sequence of complex numbers then*

$$\int_0^T \left| \sum_{n\geq 1} \frac{a_n}{n^{it}} \right|^2 dt \ll \sum_{n\geq 1} (T+n)|a_n|^2,$$

provided the right side converges.

Proof. If we expand the left side we get

$$T\sum_n |a_n|^2 + \sum_{r\neq s} a_r \overline{a_s} \int_0^T (s/r)^{it} dt.$$

The second term equals

$$\sum_{r\neq s} \frac{a_r\overline{a_s}}{i\log(s/r)}((s/r)^{iT}-1) = i\sum_{r\neq s} \frac{a_r\overline{a_s}}{\log s - \log r} - i\sum_{r\neq s} \frac{a_r r^{-iT}\overline{a_s s^{-iT}}}{\log s - \log r}.$$

Applying the Strong Hilbert's Inequality to each sum yields our result.

Reworking the last part of the proof of Proposition 1. By the Corollary above we have

$$\int_{kT\leq |y|\leq (k+1)T} \Big| \frac{F'(1+\alpha+iy)}{1+\alpha+iy} \Big|^2 dy \ll \frac{1}{1+k^2T^2} \sum_{n=1}^{\infty} \frac{|f(n)|^2 \log^2 n}{n^{2+2\alpha}} (T+n).$$

Summing over all $k\geq 1$ gives, in total,

$$\ll \frac{1}{T}\sum_{n=1}^{\infty} \frac{\log^2 n}{n^{2+2\alpha}} + \frac{1}{T^2} \sum_{\substack{n=1 \\ p|n \implies p\leq x}}^{\infty} \frac{\log^2 n}{n^{1+2\alpha}} \ll \frac{1}{T} + \frac{1}{\alpha^3 T^2}.$$

Proposition 1 now follows by substituting this into the argument above.

7. The prime number theorem

7.1. Pretentious proofs of the prime number theorem.

Proof of the Prime Number Theorem. Take $f = \mu$ and $T = \log x$ in (6.1.2). By Lemma 2.2 we deduce that

$$\Big|\sum_{n\le x} \mu(n)\Big| \ll \frac{x}{(\log x)^{1-\frac{2}{\pi}+o(1)}};$$

and then

$$\psi(x) = x + O\left(\frac{x}{(\log x)^{1-\frac{2}{\pi}+o(1)}}\right) \tag{7.1}$$

by exercise (i) of §1.5.

The classical proof of the Prime Number Theorem yields a much better error term than in (7.1); indeed something like

$$\psi(x) = x + O\left(x\exp\left(-(\log x)^{3/5+o(1)}\right)\right).$$

There are elementary proofs of the prime number theorem that yield an error term of $O\left(x\exp\left(-(\log x)^{1/2+o(1)}\right)\right)$. We believe that some of the ideas that come up below indicate that we will not able to improve the exponent $1-\frac{2}{\pi}$ in (7.1) even to 1. That is our methods are very far, quantitatively, from what can be obtained by several other methods. Hence to get good error terms with our methods one will need to incorporate unpretentious ideas.

7.2. Lower bounds on distances, II.

Lemma 7.2.1. *For any f, and any real numbers t_1, t_2 with $|t_1 - t_2| \le \log x$ we have*

$$\max_{j=1,2} \mathbb{D}(f, n^{it_j}; x)^2 \ge \left(1-\frac{2}{\pi}\right)\left(\log\log x - \log\left(\frac{1}{|t_1-t_2|}+1\right)\right) + O(1).$$

Proof. We may assume $\beta := |t_1 - t_2|/2 > 1/\log x$, else the result is trivial. Then

$$\max_{j=1,2} \mathbb{D}(f, n^{it_j}; x)^2 \ge \frac{1}{2}\sum_{j=1}^{2} \mathbb{D}(f, n^{it_j}; x)^2 = \sum_{p\le x} \frac{1 - \text{Re } f(p)(p^{-it_1}+p^{-it_2})/2}{p}$$
$$\ge \log\log x - \sum_{p\le x}\frac{|\cos(\beta\log p)|}{p} + O(1).$$

By partial summation this equals

$$\log\log x - \int_{\beta}^{\beta\log x} \frac{|\cos u|}{u}du + O\left(1 + \int_2^x \frac{|\psi(v)-v|}{v^2\log v}dv\right), \tag{7.2.1}$$

and the error term here is $O(1)$, by the prime number theorem (7.1).

Exercise: Show that $\int_t^{t+2\pi} \frac{|\cos u|}{u} du = \frac{2}{\pi} \int_t^{t+2\pi} \frac{du}{u} + O\left(\frac{1}{t^2}\right)$ for any $t \geq 1$.
Show that $\int_\beta^1 \frac{|\cos u|}{u} du = \log(1/\beta) + O(1)$. (Hint: Compare the left side to $\int_\beta^1 \frac{1}{u} du$.)

Hence, if $\beta \geq 1$ then $\int_\beta^{\beta \log x} \frac{|\cos u|}{u} du = \frac{2}{\pi} \int_\beta^{\beta \log x} \frac{du}{u} + O(1) = \frac{2}{\pi} \log\log x + O(1)$, and the result follows.

If $\beta < 1$ then $\int_\beta^{\beta \log x} \frac{|\cos u|}{u} du = \int_1^{\beta \log x} \frac{|\cos u|}{u} du + \int_\beta^1 \frac{|\cos u|}{u} du = \frac{2}{\pi} \log(\beta \log x) + \log(1/\beta) + O(1)$, so that (7.2.1) becomes $\left(1 - \frac{2}{\pi}\right) \log(\beta \log x) + O(1)$ and the result follows.

7.3. Ellenberg's problem. Suppose that f is a multiplicative function, with $|f(n)| = 1$ for all $n \geq 1$. Define

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

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

Distribution Theorem. *Let f be a completely multiplicative function such that each $f(p)$ is on the unit circle. Either the $f(n)$ are uniformly distributed on the unit circle, or there exists a positive integer k for which $(1/N) \sum_{n \leq N} f(n)^k \not\to 0$. If k is the smallest such integer then*

$$R(N, \alpha + \tfrac{1}{k}, \beta + \tfrac{1}{k}) - R(N, \alpha, \beta) \to 0 \text{ for all } 0 \leq \alpha < \beta < 1.$$

Moreover $R_f(N, \alpha, \beta) - \frac{1}{k} R_{f^k}(N, k\alpha, k\beta) \to 0$ for $0 \leq \alpha < \beta < 1$

The last parts of the result tell us that if f is not uniformly distributed on the unit circle, then its distribution function is k copies of the distribution function for f^k, a multiplicative function whose mean value does not $\to 0$. It is easy to construct examples of such functions $f^k = g$ whose distribution function is not uniform: Let $g(p) = 1$ for all odd primes p and $g(2) = e(\sqrt{2})$, where g is completely multiplicative.

To prove the 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 $(1/N) \sum_{n \leq N} \xi_n^m$ exists and equals 0, for each non-zero integer m.*

We warm up for the proof of the distribution theorem by proving the following result:

Corollary. *Let f be a completely multiplicative function such that each $f(p)$ is on the unit circle. The following statements are equivalent:*

(i) The $f(n)$ are uniformly distributed on the unit circle.
(ii) Fix any $t \in \mathbb{R}$. The $f(n)n^{it}$ are uniformly distributed on the unit circle.
(iii) For each fixed non-zero integer k, we have $\sum_{n \leq N} f(n)^k = o(N)$.

Proof. That (i) is equivalent to (iii) is given by Weyl's equidistribution theorem. By (3.1.2) we find that (iii) does not hold for some given $k \neq 0$ if and only if $f(n)^k$ is n^{iu}-pretentious for some fixed u. But this holds if and only if $(f(n)n^{it})^k$ is $n^{i(u+kt)}$-pretentious for some fixed u. But then, by Theorem 6.2, we see that (iii) does not hold with $f(n)$ replaced by $f(n)n^{it}$, and hence the $f(n)n^{it}$ are not uniformly distributed on the unit circle.

Proof of the distribution theorem. The first part of the result follows from the above Corollary. If k is the smallest positive integer for which $\sum_{n\le N} f(n)^k \gg N$ then, by Halasz's Theorem we know that there exists $u_k \ll 1$ such that $\mathbb{D}(f(n)^k, n^{iku_k}, \infty) < \infty$, and that $\mathbb{D}(f^j, n^{iu}, \infty) = \infty$ for $1 \le j \le k-1$, whenever $|u| \ll 1$.[14] Write $f(p) = r(p)p^{iu_k}g(p)$, where $r(p)$ is chosen to be the nearest kth root of unity to $f(p)p^{-iu_k}$, so that $|\arg(g(p))| \le \pi/k$, and hence $1 - \mathrm{Re}(g(p)) \le 1 - \mathrm{Re}(g(p)^k)$. Therefore $\mathbb{D}(1, g, \infty) \le \mathbb{D}(g^k, 1, \infty) = \mathbb{D}(f(n)^k, n^{iku_k}, \infty) < \infty$.

By further use of the triangle inequality, $\mathbb{D}(f^{mk}, n^{ikmu_k}, \infty) \le m\mathbb{D}(f^k, n^{iku_k}, \infty) < \infty$, and $\mathbb{D}(f^{mk+j}, n^{iu}, \infty) \ge D(f^j, n^{iv}\infty) - \mathbb{D}(f^{mk}, n^{ikmu_k}, \infty) = \infty$, where $v = u - kmu_k$ for $1 \le j \le k-1$ and any $|u| \ll 1$, and so $\sum_{n\le N} f(n)^\ell = o_\ell(N)$ if $k \nmid \ell$.

The characteristic function of the interval (α, β) is

$$\sum_{m\in\mathbb{Z}} \frac{e(m\alpha) - e(m\beta)}{2i\pi m}\, e(mt).$$

We can take this sum in the range $1 \le |m| \le M$ with an error $\le \epsilon$. Hence

$$\begin{aligned} R(N, \alpha, \beta) &= \sum_{1\le|m|\le M} \frac{e(m\alpha) - e(m\beta)}{2i\pi m} \frac{1}{N} \sum_{n\le N} f(n)^m + O(\epsilon) \\ &= \sum_{1\le|r|\le R} \frac{e(kr\alpha) - e(kr\beta)}{2i\pi kr}\, \frac{1}{N} \sum_{n\le N} f(n)^{kr} + O(\epsilon) \end{aligned}$$

writing $m = kr$ (since the other mean values are 0) and $R = [M/k]$. This formula does not change value when we change $\{\alpha, \beta\}$ to $\{\alpha + \frac{1}{k}, \beta + \frac{1}{k}\}$, nor when we change $\{f, \alpha, \beta\}$ to $\frac{1}{k}$ times the formula for $\{f^k, k\alpha, k\beta\}$ and hence the results.

[14] Note that the sum for $-k$ is the complex conjugate of the sum for k, so we can restrict attention to positive integers k.

8. The large sieve

Let $a_1, a_2, \dots$ be a sequence of complex numbers. We are interested in how they are distributed in arithmetic progressions. By (2.2.2), when $(b,q)=1$, we have

$$\sum_{n\equiv b \pmod q} a_n = \frac{1}{\varphi(q)} \sum_{\chi \pmod q} \overline{\chi}(b) \sum_n a_n\chi(n),$$

Therefore, by using (2.2.3), we deduce that

$$\sum_{(b,q)=1} \Bigg| \sum_{n\equiv b \pmod q} a_n \Bigg|^2 = \frac{1}{\varphi(q)} \sum_{\chi \pmod q} \Bigg| \sum_n a_n\chi(n) \Bigg|^2. \tag{8.1}$$

Now

$$\begin{aligned} \sum_{(b,q)=1} \Bigg| \sum_{\substack{n\le N \\ n\equiv b \pmod q}} a_n \Bigg|^2 &\le \sum_{(b,q)=1} \left(\frac{N}{q}+1\right) \sum_{\substack{n\le N \\ n\equiv b \pmod q}} |a_n|^2 \\ &= \left(\frac{N}{q}+1\right) \sum_n |a_n|^2, \end{aligned}$$

so by (8.1) we deduce that

$$\frac{q}{\varphi(q)} \sum_{\chi \pmod q} \Bigg| \sum_{n\le N} a_n\chi(n) \Bigg|^2 \le (q+N) \sum_{n\le N} |a_n|^2. \tag{8.3}$$

Note that if $a_n = \overline{\chi}(n)$ for all n, then the term on the left-side of (8.3) corresponding to the character χ has size $\frac{\varphi(q)}{q}N^2$, whereas the right-side of (8.3) is about $(q+N)\frac{\varphi(q)}{q}N$. Hence if $q = o(N)$ and then (8.3) is best possible and any of the terms on the left-side could be as large as the right side. It thus makes sense to remove the largest term on the right side (or largest few terms) to determine whether we can get a significantly better upper bound for the remaining terms.

The same argument used to prove (8.1) yields, for any choice of $\chi_1, \dots, \chi_k$,

(8.2)

$$\sum_{(b,q)=1} \Bigg| \sum_{n\equiv b \pmod q} a_n - \frac{1}{\varphi(q)} \sum_{i=1}^{k} \overline{\chi}(b) \sum_n a_n\chi_i(n) \Bigg|^2 = \frac{1}{\varphi(q)} \sum_{\chi\ne\chi_1,\dots,\chi_k} \Bigg| \sum_n a_n\chi(n) \Bigg|^2.$$

Summing the left-side of (8.3) over $q \le Q$ is important in applications, which yields a right-side with coefficient $Q^2/2 + QN$. Using some simple linear algebra we will improve this to

$$\sum_{q\le Q} \frac{q}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \Bigg| \sum_{n=M+1}^{M+N} a_n\chi(n) \Bigg|^2 \le (N + 3Q^2\log Q) \sum_{n=M+1}^{M+N} |a_n|^2. \tag{8.4}$$

This is also known to be true with $3Q^2 \log Q$ replaced by $Q^2 - 1$, and we will (?) use Hilbert's inequality to do this.

The Duality Principle. *Let $x_{m,n} \in \mathbb{C}$ for $1 \le m \le M,\ 1 \le n \le N$. For any constant c we have*

$$\sum_n \left| \sum_m a_m x_{m,n} \right|^2 \le c \sum_n |a_n|^2$$

for all $a_m \in \mathbb{C},\ 1 \le m \le M$ if and only if

$$\sum_m \left| \sum_n b_n x_{m,n} \right|^2 \le c \sum_m |b_m|^2$$

for all $b_n \in \mathbb{C},\ 1 \le n \le N$.

Proof. Assume that the first inequality is true. Given $b_n \in \mathbb{C},\ 1 \le n \le N$ define $a_m = \sum_n b_n x_{m,n}$, so that

$$\sum_m \left| \sum_n b_n x_{m,n} \right|^2 = \sum_m \overline{a}_m \sum_n b_n x_{m,n} = \sum_n b_n \sum_m \overline{a}_m x_{m,n},$$

so by the Cauchy-Schwarz inequality, the above squared is

$$\left(\sum_n |a_n|^2 \right)^2 \le \sum_n |b_n|^2 \cdot \sum_n \left| \sum_m \overline{a}_m x_{m,n} \right|^2 \le \sum_m |b_m|^2 \cdot c \sum_n |a_n|^2,$$

using the hypothesis, and the result follows. The reverse implication is completely analogous.

Proposition 8.1. *Let $a_n, M+1 \le n \le M+N$ be a set of complex numbers, and $x_r, 1 \le r \le R$ be a set of real numbers. Let $\delta := \min_{r \ne s} \|x_r - x_s\| \in [0, 1/2]$, where $\|t\|$ denotes the distance from t to the nearest integer. Then*

$$\sum_r \left| \sum_{n=M+1}^{M+N} a_n e(nx_r) \right|^2 \le \left(N + \frac{\log(e/\delta)}{\delta} \right) \sum_{n=M+1}^{M+N} |a_n|^2$$

where $e(t) = e^{2i\pi t}$.

Proof. For any $b_r \in \mathbb{C},\ 1 \le r \le R$, we have

$$\sum_n \left| \sum_r b_r e(nx_r) \right|^2 = \sum_{r,s} b_r \overline{b}_s \sum_{n=M+1}^{M+N} e(n(x_r - x_s)) = N\|b\|^2 + E,$$

since the inner sum is N if $r = s$, where, for $L := M + \frac{1}{2}(N+1)$,

$$E \le \sum_{r \ne s} b_r \overline{b}_s e(L(x_r - x_s)) \frac{\sin(\pi N(x_r - x_s))}{\sin(\pi(x_r - x_s))}.$$

Taking absolute values we obtain

$$|E| \le \sum_{r\neq s} \frac{|b_r\overline{b}_s|}{|\sin(\pi(x_r-x_s))|} \le \sum_{r\neq s} \frac{|b_r\overline{b}_s|}{2\|x_r-x_s\|} \le \sum_r |b_r|^2 \sum_{s\neq r} \frac{1}{2\|x_r-x_s\|}$$

since $2|b_r\overline{b}_s| \le |b_r|^2+|b_s|^2$. Now, for each x_r the nearest two x_s are at distance at least δ away, the next two at distance at least 2δ away, etc, and so

$$|E| \le \sum_r |b_r|^2 \sum_{j=1}^{[1/\delta]} \frac{2}{2j\delta} \le \frac{\log(e/\delta)}{\delta} \sum_m |b_m|^2,$$

so that

$$\sum_n \left|\sum_r b_r e(nx_r)\right|^2 \le \left(N+\frac{\log(e/\delta)}{\delta}\right)\sum_m |b_m|^2.$$

The result follows by the duality principle.

`We should improve the result, getting a constant` $N+O(1/\delta)$ `by using the strong Hilbert inequality in the proof above.`

For a character $\chi \pmod q$, let the *Gauss sum* $g(\chi)$ be defined as

$$g(\chi) := \sum_{a \pmod q} \chi(a) e\left(\frac{a}{q}\right).$$

For fixed m with $(m,q)=1$ we change the variable a to mb, as b varies through the residues mod q, coprime to q, so that

$$\overline{\chi}(m)g(\chi) = g(\chi,m), \text{ where } g(\chi,m) := \sum_{a \pmod q} \chi(a) e\left(\frac{am}{q}\right). \tag{8.5}$$

It is known that if χ is primitive then $|g(\chi)|=\sqrt{q}$, Using this we can establish (8.4):

Proof. By (8.5) we have

$$\sum_{n=M+1}^{M+N} a_n\chi(n) = \frac{1}{g(\overline{\chi})} \sum_{b \pmod q} \overline{\chi}(b) \sum_{n=M+1}^{M+N} a_n e\left(\frac{bn}{q}\right).$$

Therefore, using (8.1)

$$\sum_{\substack{\chi \pmod q\\ \chi \text{ primitive}}} \left|\sum_{n=M+1}^{M+N} a_n\chi(n)\right|^2 \le \frac{1}{q} \sum_{\substack{\chi \pmod q\\ \chi \text{ primitive}}} \left|\sum_{b \pmod q} \overline{\chi}(b) \sum_{n=M+1}^{M+N} a_n e\left(\frac{bn}{q}\right)\right|^2$$

$$\le \frac{\varphi(q)}{q} \sum_{\substack{b \pmod q\\ (b,q)=1}} \left|\sum_{n=M+1}^{M+N} a_n e\left(\frac{bn}{q}\right)\right|^2.$$

We deduce that the left side of (8.4) is

$$\sum_{q \le Q} \sum_{\substack{b \pmod q \\ (b,q)=1}} \left| \sum_{n=M+1}^{M+N} a_n e\left(\frac{bn}{q}\right) \right|^2$$

We now apply Proposition 8.1 with $\{x_r\} = \{b/q : \ (b,q)=1, \ q \le Q\}$, so that

$$\delta \ge \min_{q,q' \le Q} \min_{\substack{b,b' \\ b/q \ne b'/q'}} \left| \frac{b}{q} - \frac{b'}{q'} \right| \ge \min_{q \ne q' \le Q} \frac{1}{qq'} \ge \frac{1}{Q(Q-1)},$$

and (8.4) follows.

8.2. Other forms of the large sieve. `Needs verifying`

Hildebrand used the large sieve in the form (?):

$$\sum_{\substack{p \le Q \\ p \text{ prime}}} \left| \sum_{\substack{n \le N \\ n \equiv 0 \pmod p}} a_n - \frac{1}{p} \sum_{n \le N} a_n \right| \le ?? \sum_{n \le N} |a_n|^2$$

Elliott [MR962733] proved that for $Q < x^{1/2-\epsilon}$, and f multiplicative with $|f(n)| \le 1$,

$$\sum_{p \le Q}{}' (p-1) \max_{y \le x} \max_{(a,p)=1} \left| \sum_{\substack{n \le y \\ n \equiv a \pmod p}} f(n) - \frac{1}{p-1} \sum_{\substack{n \le y \\ (n,p)=1}} f(n) \right|^2 \ll \frac{x}{\log^A x},$$

where the sum is over all p except one where there might be an exceptional character.

8.3. Consequences of the large sieve.

1/ Barban-Dav-Halb
2/ Bombieri-Vinogradov
3/ Least quadratic non-residue.

9. The Small Sieve

9.1. Shiu's Theorem and the proof of Lemma 9.1.

We wish to prove that if $0 \le f(n) \le 1$ and $(a,q)=1$ then

$$\sum_{\substack{x<n\le x+qy \\ n\equiv a \pmod q}} f(n) \ll y \exp\left(-\sum_{\substack{p\le y\\ p\nmid q}} \frac{1-f(p)}{p} \right)$$

For now assume that f is totally multiplicative. Write $n = p_1p_2\dots$ with $p_1 \le p_2 \le \dots$, and let $d = p_1p_2\dots p_k$ where $d \le y^{1/2} < dp_{k+1}$. Therefore $n = dm$ where the smallest prime factor of m is at least $\max\{P(d), y^{1/2}/d\}$, where $P(d)$ is the largest prime factor of d, and so equals p_k. For any such d we have m in an interval $(x/d, x/d + qy/d]$ of an arithmetic progression $a/d \pmod q$ containing $y/d + O(1)$ terms. By the small sieve the number of such m is $\ll \frac{qy/d}{\varphi(q)\log(P(d)+y^{1/2}/d)}$. Since $f(n) \le f(d)$ we deduce that

$$\sum_{\substack{x<n\le x+qy \\ n\equiv a \pmod q}} f(n) \le \frac{qy}{\varphi(q)} \sum_{\substack{d\le y^{1/2}\\ (d,q)=1}} \frac{f(d)}{d\log(P(d)+y^{1/2}/d)}.$$

If we consider just those terms with $d \le y^{1/2-\epsilon}$ or $P(d) > y^{\epsilon}$, so that $\log(P(d)+y^{1/2}/d) \ge \epsilon \log y$, then we get

$$\ll \frac{qy}{\varphi(q)} \prod_{p\le y}\left(1-\frac{1}{p}\right) \sum_{\substack{d\le y^{1/2}\\ (d,q)=1}} \frac{f(d)}{d} \ll y \prod_{\substack{p\le y\\ p\nmid q}} \left(1-\frac{1}{p}\right)\left(1-\frac{f(p)}{p}\right)^{-1}, \tag{9.1}$$

the upper bound claimed above. We are left with the $d > y^{1/2-\epsilon}$ with $P(d) \asymp 2^r$ for some $1 \le r \le k = [\epsilon \log y]$. Hence we obtain an upper bound:

$$\frac{qy}{\varphi(q)} \sum_{r=1}^{k} \frac{1}{r} \sum_{\substack{d>y^{1/2-\epsilon}\\ (d,q)=1\\ P(d)\asymp 2^r}} \frac{f(d)}{d} \ll \frac{qy}{\varphi(q)} \left(\frac{1}{k} \sum_{\substack{d>y^{1/2-\epsilon}\\ (d,q)=1\\ P(d)\le 2^k}} \frac{f(d)}{d} + \sum_{r=1}^{k} \frac{1}{r^2} \sum_{\substack{d>y^{1/2-\epsilon}\\ (d,q)=1\\ P(d)\le 2^r}} \frac{f(d)}{d} \right).$$

For the first term we proceed as above. For the remaining terms we use Corollary 3.4.2 to obtain

$$\ll \frac{qy}{\varphi(q)} \sum_{r=1}^{k} \frac{1}{r^2} \prod_{\substack{p\le 2^r\\ p\nmid q}} \left(1-\frac{f(p)}{p}\right)^{-1} \frac{1}{u_r^{u_r+1}},$$

where $u_r := (1/2-\epsilon)\log y/(r\log 2)$. Now this is (9.1) times

$$\ll \prod_{p\le y}\left(1-\frac{1}{p}\right)^{-1}\sum_{r=1}^{k}\frac{1}{r^2u_r^{u_r+1}}\prod_{\substack{2^r<p\le y\\ p\nmid q}}\left(1-\frac{f(p)}{p}\right)\ll\sum_{r=1}^{k}\frac{1}{ru_r^{u_r}}\ll \epsilon^{O(1/\epsilon)}.$$

To prove this last step, consider those r in an interval $R<r\le 2R$ and write $u=(1/2-\epsilon)\log y/(R\log 2)$ so that u is in a dyadic interval also. Hence this sum is about $1/u^u$ and we sum over $u=u_0,2u_0,4u_0,\dots$ with $u_0=1/\epsilon$.

9.2. Small sieve type results. Define

$$\rho_q(f):=\prod_{\substack{p\le q\\ p\nmid q}}\left(1-\frac{1}{p}\right)\left(1+\frac{|f(p)|}{p}\right)\text{ and }\rho_q'(f)=\frac{\varphi(q)}{q}\rho_q(f).$$

We also define

$$\log_S(n):=\sum_{\substack{d\in S\\ d|n}}\Lambda(d),$$

where S might be an interval $[a,b]$, and we might write "$\le Q$" in place of "'$[2,Q]$", or "$\ge R$" in place of "'$[R,\infty)$". Note that $\log n=\log_{[2,n]}n$.

Lemma 9.1. *Suppose that $x\ge Q^{2+\epsilon}$ and $Q\ge q$. Then, for any character $\chi \pmod q$,*

$$\left|\sum_{\substack{n\in\mathcal{N}\\ n\equiv a \pmod q}}f(n)\overline{\chi}(n)\mathcal{L}(n)\right|\ll\rho_q(f)\frac{x}{q}=\rho_q'(f)\frac{x}{\varphi(q)},$$

where $\mathcal{L}(n)=1,\log(x/n)$, $\frac{\log_{\le Q}n}{\log Q}$ or $\frac{\log_{\ge x/Q}n}{\log Q}$, and $\mathcal{N}=\{n:Y<n\le Y+x\}$ for $Y=0$ in the second and fourth cases, and for any Y in the other two cases.

Proof. For the first estimate, the small sieve (`reference?`) yields that if $x\ge q^{1+\epsilon}$ then

$$\left|\sum_{\substack{n\in\mathcal{N}\\ n\equiv a \pmod q}}f(n)\overline{\chi}(n)\right|\le\sum_{\substack{n\in\mathcal{N}\\ n\equiv a \pmod q}}|f(n)|\ll\rho_q(f)\frac{x}{q}.$$

The second estimate follows similarly. If d is a power of the prime p then let $f_d(n)$ denote $f(n/p^a)$ where $p^a\|n$, so that if $n=dm$ then $|f(n)|\le|f_d(m)|$. Therefore if $x>Qq^{1+\epsilon}$ then, for the third estimate, times $\log Q$, we have

$$\le\sum_{\substack{Y<md\le Y+x\\ md\equiv a \pmod q\\ d\le Q}}|f(md)|\Lambda(d)\le\sum_{\substack{d\le Q\\ (d,q)=1}}\Lambda(d)\sum_{\substack{Y/d<m\le(Y+x)/d\\ m\equiv a/d \pmod q}}|f_d(m)|$$

$$\ll\sum_{\substack{d\le Q\\ (d,q)=1}}\frac{\Lambda(d)}{d}\,\rho_q(f_d)\frac{x}{q}\ll\rho_q(f)\frac{x}{q}\log Q.$$

In the final case, if $x > Q(q^{1+\epsilon} + \log Q) + q^2$ then writing $n = mp$ where p is a prime $> x/Q$, we have

$$\begin{aligned}
&\le \sum_{\substack{m\le Q\\ (m,q)=1}} |f(m)| \sum_{\substack{x/Q<p\le x/m\\ p\equiv a/m \pmod q}} \log p + \sum_{x/Q<p\le\sqrt{x}} \sum_{\substack{n\le x\\ n\equiv a \pmod q\\ p^2|n}} \log n\\
&\ll \sum_{\substack{m\le Q\\ (m,q)=1}} |f(m)| \frac{x/m}{\varphi(q)} + \log x \sum_{x/Q<p\le\sqrt{x}} \left(\frac{x}{qp^2}+1\right)\\
&\ll \frac{x}{q}\rho_q(f)\log Q + \frac{Q}{q}\log x + \sqrt{x} \ll \rho_q(f)\frac{x}{q}\log Q.
\end{aligned}$$

by the Brun-Titchmarsh theorem.

By (8.1) we immediately deduce

Corollary 9.2. *Suppose that $x \ge Q^{2+\epsilon}$ and $Q \ge q$. Then*

$$\sum_{\chi \pmod q} \left| \sum_{n\le x} f(n)\overline{\chi}(n)\mathcal{L}(n)\right|^2 \ll (\rho_q'(f)x)^2,$$

where $\mathcal{L}(n) = 1,\ \log(x/n),\ \frac{\log_{\le Q} n}{\log Q}$ or $\frac{\log_{\ge x/Q} n}{\log Q}$, and $\mathcal{N} = \{n : Y < n \le Y + x\}$ for $Y = 0$ in the second and fourth cases, and for any Y in the other two cases.

Lemma 9.3. *If $\Delta > q^{1+\epsilon}$ then for any $D \ge 0$ we have*

$$\sum_{\chi \pmod q} \left| \sum_{D\le d\le D+\Delta} f(d)\overline{\chi}(d)\Lambda(d)\right|^2 \ll \Delta^2.$$

Proof. By (8.1), when we expand the sum on the left hand side we obtain $\varphi(q)$ times

$$\sum_{(b,q)=1} \left| \sum_{\substack{d\equiv b \pmod q\\ D\le d\le D+\Delta}} f(d)\Lambda(d)\right|^2 \le \sum_{(b,q)=1} \left| \sum_{\substack{d\equiv b \pmod q\\ D\le d\le D+\Delta}} \Lambda(d)\right|^2 \ll \varphi(q)\left(\frac{\Delta}{\varphi(q)}\right)^2$$

by the Brun-Titchmarsh theorem, and the result follows.

10. The Pretentious Large Sieve

10.1. Mean values of multiplicative functions, on average. Our goal is to produce an averaged version of (3.1.3) for f twisted by all the characters $\chi \pmod q$, but with a better error term. Define

$$S_\chi(x) := \sum_{n\le x} f(n)\overline{\chi}(n).$$

Throughout we let $\mathcal{C}_q$ be any subset of the set of characters $\pmod q$, and define

$$L = L(\mathcal{C}_q) := \frac{1}{\log x} \max_{\chi\in\mathcal{C}_q} \max_{|t|\le \log^2 x} |F_\chi(1+it)|,$$

where

$$F_\chi(s) := \prod_{p\le x}\left(1+\frac{f(p)\overline{\chi}(p)}{p^s}+\frac{f(p^2)\overline{\chi}(p^2)}{p^{2s}}+\dots\right).$$

We will let $L_k(x;q)$ be the minimum of $L(\mathcal{C}_q)$ as we vary over all sets $\mathcal{C}_q$ of k distinct characters modulo q. Our main result is the following:

Theorem 10.1. *Suppose that $x\ge Q^{2+\epsilon}$ and $Q\ge q^{2+\epsilon}\log x$. Then*

$$\sum_{\chi\in\mathcal{C}_q}\left|\frac{1}{x}S_\chi(x)\right|^2 \ll \left(\left(L(\mathcal{C}_q)+\rho_q'(f)\frac{\log Q}{\log x}\right)\log\left(\frac{\log x}{\log Q}\right)\right)^2.$$

Lemma 10.2. *Suppose that $x\ge Q^{2+\epsilon}$ and $Q\ge q^{2+\epsilon}\log x$. Then*

$$\log^2 x \sum_{\chi\in\mathcal{C}_q}\left|\frac{1}{x}S_\chi(x)\right|^2 \ll \sum_{\chi\in\mathcal{C}_q}\left(\int_Q^x\left|\frac{1}{t}S_\chi(t)\right|\frac{dt}{t}\right)^2+\rho_q'(f)^2\left(\log Q+\log\left(\frac{\log x}{\log Q}\right)\right)^2.$$

Proof. We set $z=\frac{x}{Q}, h = x\ \frac{\log Q}{\log x}$ and $T = x\left(\frac{\log Q}{\log x}\right)^2$. By Corollary 9.2 we have

$$\sum_{\chi \pmod q}\left|\sum_{n\le x} f(n)\overline{\chi}(n)\log(x/n)\right|^2 \ll (\rho_q'(f)x)^2,$$

and

$$\sum_{\chi}\left|\sum_{n\le x} f(n)\overline{\chi}(n)\log_{>z} n\right|^2 \ll (\rho_q'(f)x\log Q)^2.$$

For $g=f\overline{\chi}$ we have

$$\left|\sum_{n\le x} g(n)\log_{\le z} n\right| \le \frac{1}{h}\int_x^{x+h}\left|\sum_{n\le t} g(n)\log_{\le z} n\right| dt+\frac{1}{h}\int_x^{x+h}\left|\sum_{x<n\le t} g(n)\log_{\le z} n\right| dt, \tag{10.1}$$

by the triangle inequality. When we square and sum over characters χ, with $g = f\overline{\chi}$, the sum of the second terms on the right side of (10.1) is

$$\leq 2\sum_{\chi}\left|\frac{1}{h}\int_x^{x+h}\left|\sum_{x<n\leq t} f(n)\overline{\chi}(n)\log_{\leq z} n\right| dt\right|^2$$
$$\leq \frac{2}{h}\int_x^{x+h}\sum_{\chi}\left|\sum_{x<n\leq t} f(n)\overline{\chi}(n)\log_{\leq z} n\right|^2 dt \ll (\rho_q'(f)h\log z)^2 = (\rho_q'(f)x\log Q)^2,$$

by Corollary 9.2. For the first term on the right side of (10.1) we have

$$\sum_{n\leq x} g(n)\log_{\leq z} n = \sum_{\substack{md\leq x\\ d\leq z}} g(m)g(d)\Lambda(d) \leq \sum_{d\leq z}\Lambda(d)\left|\sum_{m\leq x/d} g(m)\right|,$$

and so

$$\frac{1}{h}\int_x^{x+h}\left|\sum_{n\leq t} g(n)\log_{\leq z} n\right| dt \leq \frac{1}{h}\int_x^{x+h}\sum_{d\leq z}\Lambda(d)\left|\sum_{m\leq t/d} g(m)\right| dt$$
$$\leq \sum_{d\leq z} d\Lambda(d)\frac{1}{h}\int_{x/d}^{(x+h)/d}\left|\sum_{m\leq t} g(m)\right| dt.$$

Note that $t \geq x/d \geq x/z = Q$. We split the integral into two parts. The first part of the integral is where $Q \leq t \leq T$, and we get

$$\frac{1}{h}\int_Q^T\left|\sum_{m\leq t} g(m)\right|\sum_{\substack{x/t\leq d\leq (x+h)/t\\ d\leq z}} d\Lambda(d)dt \ll x\int_Q^T\left|\sum_{m\leq t} g(m)\right|\frac{\log(x/t)}{\log(h/t)}\frac{dt}{t^2}$$
$$\ll x\int_Q^T\left|\sum_{m\leq t} g(m)\right|\frac{dt}{t^2}$$

by the Brun-Titchmarsh theorem, as $h/T = (x/T)^{1/2}$. To obtain the result we will extend the upper limit of the integral from T to x which is okay since the integrand is non-negative. For the second part we have

$$\sum_{d\leq (x+h)/T} d\Lambda(d)\frac{1}{h}\int_{\max\{T,x/d\}}^{(x+h)/d}\left|\sum_{m\leq t} g(m)\right| dt;$$

and when we square and sum over characters χ, with $g = f\overline{\chi}$, the sum of the second parts is, Cauchying twice,

$$\ll \sum_{d\le(x+h)/T} \frac{\Lambda(d)}{d} \sum_{d\le(x+h)/T} d^3\Lambda(d) \sum_{\chi} \left(\frac{1}{h} \int_{x/d}^{(x+h)/d} \left| \sum_{m\le t} f(m)\overline{\chi}(m) \right| dt \right)^2$$
$$\ll \log(x/T) \sum_{d\le(x+h)/T} d^2\Lambda(d) \frac{1}{h} \int_{x/d}^{(x+h)/d} \sum_{\chi} \left| \sum_{m\le t} f(m)\overline{\chi}(m) \right|^2 dt$$
$$\ll \log(x/T) \sum_{d\le(x+h)/T} d^2\Lambda(d) \frac{1}{h} \int_{x/d}^{(x+h)/d} (\rho_q'(f)t)^2 dt$$
$$\ll \rho_q'(f)^2 \log(x/T) \sum_{d\le(x+h)/T} d\Lambda(d)(x/d)^2 \ll (\rho_q'(f)x\log(x/T))^2 \ll \left(\rho_q'(f)x\log\left(\frac{\log x}{\log Q}\right)\right)^2,$$

by Corollary 9.2. The result follows from collecting up the estimates given, using that $\log x = \log(x/n) + \log_{\le z} n + \log_{>z} n$.

Lemma 10.3. *Suppose that $x \ge Q^{2+\epsilon}$ and $Q \ge q^{2+\epsilon}\log x$. Then*

$$\sum_{\chi\in\mathcal{C}_q} \left(\int_Q^x \left| \frac{1}{t} S_\chi(t) \right| \frac{dt}{t} \right)^2 \ll \sum_{\chi\in\mathcal{C}_q} \left(\int_Q^x \left| \sum_{n\le t} f(n)\overline{\chi}(n) \log_{>Q/q} n \right| \frac{dt}{t^2\log t} \right)^2 + \left(\rho_q'(f)\log Q \cdot \log\left(\frac{\log x}{\log Q}\right)\right)^2.$$

Proof. We expand using the fact that $\log t = \log(t/n) + \log_{\le Q/q} n + \log_{>Q/q} n$; and the Cauchy-Schwarz inequality so that, for any function $c_\chi(t)$,

$$\sum_{\chi} \left(\int_Q^x c_\chi(t) \frac{dt}{t^2\log t} \right)^2 \le \int_Q^x \frac{dt}{t\log t} \cdot \int_Q^x \sum_{\chi} c_\chi(t)^2 \frac{dt}{t^3\log t}$$

By Corollary 9.2 we then have

$$\int_Q^x \sum_{\chi} \left| \sum_{m\le t} f(m)\overline{\chi}(m)\log(t/m) \right|^2 \frac{dt}{t^3\log t} \ll \rho_q'(f)^2 \int_Q^x \frac{dt}{t\log t} \ll \rho_q'(f)^2 \log\left(\frac{\log x}{\log Q}\right)$$

and

$$\int_Q^x \sum_{\chi} \left| \sum_{m\le t} f(m)\overline{\chi}(m)\log_{\le Q/q} m \right|^2 \frac{dt}{t^3\log t} \ll \int_Q^x \left(\rho_q'(f)t\log(Q/q)\right)^2 \frac{dt}{t^3\log t},$$

and the result follows.

Proposition 10.4. *If $x > Q^{1+\epsilon}$ and $Q \geq q^{1+\epsilon}$ then*

$$\sum_{\chi \in \mathcal{C}_q} \left(\int_{Qq}^{x} \left| \sum_{Q \leq n \leq t} f(n)\chi(n) \log_{>Q} n \right| \frac{dt}{t^2 \log t} \right)^2$$

$$\ll \log\left(\frac{\log x}{\log Q}\right) \left(M^2 \log\left(\frac{\log x}{\log Q}\right) + \frac{\varphi(q)}{T}\frac{\log Q}{Q} + \frac{\log^3 x}{T^2} \right)$$

where $M := \max_{\chi \in \mathcal{C}_q} \ \max_{|u| \leq 2T} |F_\chi(1+iu)|$.

Proof, by revisiting the proof of Proposition 1. For a given $g = f\chi$ and Q we define

$$h(n) = \sum_{\substack{md=n \\ d>Q}} g(m)g(d)\Lambda(d),$$

so that $G(s)(G'_{>Q}(s)/G_{>Q}(s)) = -\sum_{n \geq 1} h(n)/n^s$ for $\mathrm{Re}(s) > 1$. Now

$$\left| \sum_{n \leq t} g(n) \log_{>Q} n - \sum_{n \leq t} h(n) \right| \leq 2 \sum_{p^b > Q} \log p \sum_{\substack{n \leq t \\ p^{b+1} | n}} 1 \leq 2t \sum_{b \geq 1} \sum_{\substack{p^b > Q \\ p^{b+1} \leq t}} \frac{\log p}{p^{b+1}} \ll \frac{t \log t}{Q},$$

by the prime number theorem. This substitution leads to a total error, in our estimate, of

$$\ll |\mathcal{C}_q| \left(\int_{Qq}^{x} \frac{t \log t}{Q} \frac{dt}{t^2 \log t} \right)^2 \ll \frac{q}{Q^2} \log^2\left(\frac{\log x}{\log Q}\right) \ll \frac{1}{q} \log^2\left(\frac{\log x}{\log Q}\right),$$

which is smaller than the first term in the given upper bound, since $M \gg 1/\log q$.

Now we use the fact that

$$\frac{1}{\log t} \ll \int_{1/\log x}^{1/\log Q} \frac{d\alpha}{t^{2\alpha}}$$

whenever $x \geq t \geq Q$, as $x > Q^{1+\epsilon}$, so that

$$\int_2^x \left| \sum_{n \leq t} h(n) \right| \frac{dt}{t^2 \log t} \ll \int_{1/\log x}^{1/\log Q} \left(\int_2^x \left| \sum_{n \leq t} h(n) \right| \frac{dt}{t^{2+2\alpha}} \right) d\alpha.$$

Now, Cauchying, but otherwise proceeding as in the proof of Proposition 1 (with $h(n)$ here replaced by $f(n) \log n$ there), the square of the left side is

$$\ll \int_{1/\log x}^{1/\log Q} \frac{d\alpha}{\alpha} \cdot \int_{1/\log x}^{1/\log Q} \alpha \cdot \frac{1}{2\pi\alpha} \int_{-\infty}^{\infty} \left| \frac{G(G'_{>Q}/G_{>Q})(1+\alpha+it)}{1+\alpha+it} \right|^2 dt d\alpha.$$

The integral in the region with $|t| \le T$ is now

$$\le \max_{|t|\le T} |G(1+\alpha+it)|^2 \int_1^\infty \Big| \sum_{Q<n\le t} g(n)\Lambda(n)\Big|^2 \frac{dt}{t^{3+2\alpha}}.$$

If we take $g = f\chi$ and sum this over all characters $\chi \in \mathcal{C}_q$ then we obtain an error

$$\le \max_{\substack{|t|\le T\\ \chi\in\mathcal{C}_q}} |F_\chi(1+\alpha+it)|^2 \int_Q^\infty \sum_{\chi \pmod q} \Big| \sum_{Q<n\le t} f(n)\chi(n)\Lambda(n)\Big|^2 \frac{dt}{t^{3+2\alpha}}$$
$$\ll \max_{\substack{|t|\le T\\ \chi\in\mathcal{C}_q}} |F_\chi(1+\alpha+it)|^2 \int_Q^\infty \frac{dt}{t^{1+2\alpha}} \ll \frac{1}{\alpha} \max_{\substack{|t|\le T\\ \chi\in\mathcal{C}_q}} |F_\chi(1+\alpha+it)|^2,$$

by Lemma 9.3 as $t \ge Q \ge q^{1+\epsilon}$.

For that part of the integral with $|t| > T$, summed over all twists of f by characters $\chi \pmod q$, we now proceed as in the proof of Proposition 1. We obtain $\varphi(q)$ times (3.5.1), with $f(\ell)\log\ell$ replaced by $h(\ell)$ for $\ell = m$ and n, but now with the sum over $m \equiv n \pmod q$ with $m, n \ge Q$. Observing that $|h(\ell)| \le \log\ell$, we proceed analogously to obtain, in total

$$\ll \frac{\varphi(q)}{T}\frac{(\log Q)^2}{Q} + \frac{\varphi(q)}{q}\cdot\frac{1}{\alpha^4T^2}.$$

The result follows by collecting the above, and by Lemma 2.7.

Proof of Theorem 10.1. By combining the last three results, one immediately deduces that

$$\log^2 x \sum_{\chi\in\mathcal{C}_q} \Big|\frac{1}{x}S_\chi(x)\Big|^2 \ll \log\left(\frac{\log x}{\log Q}\right)\left(M^2\log\left(\frac{\log x}{\log Q}\right) + \frac{1}{T} + \frac{\log^3 x}{T^2}\right)$$
$$+ \quad \rho_q'(f)^2\left(\log Q\log\left(\frac{\log x}{\log Q}\right)\right)^2$$

where $M := \max_{\substack{|u|\le 2T\\ \chi\in\mathcal{C}_q}} |F_\chi(1+iu)|$. Now letting $T = \frac{1}{2}\log^2 x$ gives the result as $\rho_q'(f) \gg 1/\log q$.

10.2. Lower bounds on distances, II.

Lemma 10.5. *Order the characters* $\psi_j \pmod q$ *so that the* $M_{f\overline{\psi}_j}(x,T)$ *are organized in ascending order. If* $Q \geq q$ *then*

$$\exp\left(-M_{f\overline{\psi}_j}(x,T)\right) \ll e^{O(\sqrt{k})}\rho_q'(f)\left(\frac{\log Q}{\log x}\right)^{1-\frac{1}{\sqrt{k}}},$$

where the implicit constants are independent of f. *If* f *is real and* ψ_1 *is not then we can extend this to* $k=1$ *with exponent* $1-\frac{1}{\sqrt{2}}$ *on the right side of the equation.*

We shall order the characters mod q in two different ways:

Order the characters $\chi_j \pmod q$ so that the $S_{\chi_j}(x)$ are organized in descending order. Define $C_{\chi,k} = \{\chi_j : \ j \geq k\}$.

Order the characters $\psi_j \pmod q$ so that the $\max_{|t|\leq \log^2 x} |F_{\chi_j}(1+it)|$ are in descending order. Define $C_{\psi,k} = \{\psi_j : \ j \geq k\}$.

Combining this with Theorem 10.1 we obtain the following:

Corollary 10.6. *Suppose that* $x \geq Q^{2+\epsilon}$ *and* $Q \geq q^{2+\epsilon}\log x$. *For any fixed* $k \geq 1$, *we have*

$$\sum_{\substack{\chi \pmod q\\ \chi\neq\chi_1,\chi_2,\cdots,\chi_{k-1}}} \left|\frac{1}{x}S_\chi(x)\right|^2 \ll \left(e^{O(\sqrt{k})}\rho_q'(f)\left(\frac{\log Q}{\log x}\right)^{1-\frac{1}{\sqrt{k}}}\log\left(\frac{\log x}{\log Q}\right)\right)^2,$$

where the implicit constants are independent of f. *If* f *is real and* ψ_1 *is not then in the case* $k=1$ *we can replace the exponent* 0 *with* $1-\frac{1}{\sqrt{2}}$ *on the right side of the equation.*

Proof. Since $\sum_{\chi\in C_{\chi,k}} |S_\chi(x)|^2 \leq \sum_{\chi\in C_{\psi,k}} |S_\chi(x)|^2$ we apply Theorem 10.1, and then Lemma 10.5 (since $M = \log(1/L) + O(1)$ – see section 3.1).

Proof of Lemma 10.5. We begin by noting that if $Q = q\log x$ then

$$\exp\left(-\sum_{p<Q}\frac{1-\mathrm{Re}(f(p)\overline{\psi}(p)/p^{it})}{p}\right) \leq \exp\left(-\sum_{p<q}\frac{1-|f(p)|}{p}\right) \asymp \rho_q'(f).$$

Let $t_j = t_{f\overline{\psi}_j}(x,T)$ for each j. Moreover

$$\begin{aligned}\left|\sum_{j=1}^k\sum_{Q<p\leq x}\frac{f(p)\overline{\psi}_j(p)}{p^{1+it_j}}\right|^2 &\leq \left(\sum_{Q<p\leq x}\frac{1}{p}\left|\sum_{j=1}^k\frac{\psi_j(p)}{p^{it_j}}\right|\right)^2\\ &\leq \sum_{Q<p\leq x}\frac{1}{p}\cdot\left(k\sum_{Q<p\leq x}\frac{1}{p}+\sum_{1\leq i\neq j\leq k}\sum_{Q<p\leq x}\mathrm{Re}\left(\frac{\psi_i(p)\overline{\psi}_j(p)}{p^{1+i(t_i-t_j)}}\right)\right),\end{aligned}$$

by Cauchying. Now by Lemma 2.4,

$$\sum_{Q<p\le x} \frac{\psi(p)}{p^{1+i\tau}} = \log\left(\frac{L\left(1+\frac{1}{\log x}+i\tau,\psi\right)}{L\left(1+\frac{1}{\log Q}+i\tau,\psi\right)}\right) + O(1) \ll 1$$

by Lemma 10.7, taking $Q=(q\log x)^{O(1)}$. Hence the above is

$$\le \left(\log\left(\frac{\log x}{\log Q}\right)+o(1)\right)\left(k\log\left(\frac{\log x}{\log Q}\right)+O\left(k^2\right)\right),$$

Taking the square root we deduce that

$$\sum_{Q<p\le x} \operatorname{Re}\left(\frac{f(p)\overline{\psi}_k(p)}{p^{1+it_k}}\right) \le \frac{1}{\sqrt{k}}\log\left(\frac{\log x}{\log Q}\right) + O\left(\sqrt{k}\right),$$

and hence, exponentiating, gives the first result. If f is real-valued and ψ_1 is not, then $\operatorname{Re}(f(p)\overline{\psi}_1(p)/p^{it_1}) = \operatorname{Re}(f(p)\psi_1(p)/p^{-it_1})$ and $\overline{\psi}_1=\psi_j$ for some $j\ge 2$, and the final part of the result follows.

10.3. Linear averaging. By (2.1.3) we can generalize Corollary 10.6 to

Corollary 10.8. *Order the characters $\chi_j \pmod q$ so that the $S_{\chi_j}(x)$ are organized in descending order. Fix $\epsilon>0$. There exists an integer $k\ll 1/\epsilon^2$ such that if $x\ge q^{4+5\epsilon}$ then*

$$\sum_{\substack{\chi \pmod q\\ \chi\ne\chi_1,\chi_2,\cdots,\chi_k}} \left|\frac{1}{y}S_{\chi_j}(y)\right|^2 \ll e^{O(1/\epsilon)}\left(\rho_q'(f)\left(\frac{\log q+\log\log y}{\log y}\right)^{1-\epsilon}\right)^2, \tag{10.3.1}$$

for any y in the range

$$\log x \ge \log y \ge \log x \Big/ \left(\frac{\log x}{\log q+\log\log x}\right)^{\epsilon/2},$$

where the implicit constants are independent of f.

Proof. Select k to be the smallest integer for which $1/\sqrt{k}<3\epsilon$. Let $\mathcal{C}_q$ be the set of all characters mod q except $\chi_1,\chi_2,\ldots,\chi_k$. Let $Q=q^{2+\epsilon}\log x$ and write $x=Q^B$, so that $y=Q^C$, where $B\ge C\ge B^{1-\epsilon/2}$. We apply Theorem 10.1 with $x=y$. Noting that

$$L_y \ll L_x\left(\frac{\log x}{\log y}\right)^2 \ll V^{O(1/\epsilon)}\rho_q'(f)\frac{1}{B^{1-3\epsilon}}B^{\epsilon} \ll e^{O(1/\epsilon)}\rho_q'(f)\frac{1}{C^{1-2\epsilon}}$$

by (2.1.3) and Lemma 10.5, and the result follows. Note that we bound L_y by a function of L_x so that we can have the same exceptional characters $\chi_1,\chi_2,\ldots,\chi_k$ for each y in our range.

We use this to deduce the following technical tool.

Proposition 10.9. *Fix $\varepsilon > 0$. For given $x = q^A$ there exists $K \ll \epsilon^{-3} \log\log A$ such that if $x \geq X \geq x^{1/2}$ and $X/q^{4+5\epsilon} \geq U > L \geq q^{2+\epsilon} \log x$ then*

$$\frac{1}{\log x} \sum_{\substack{\chi \pmod q \\ \chi \neq \chi_j,\ j=1,\dots,K}} \left| \frac{1}{X} \sum_{n \leq X} f(n)\overline{\chi}(n) \log_{[L,U]} n \right| \ll e^{O(1/\epsilon)} \rho_q'(f) \left(\frac{\log q + \log\log x}{\log x} \right)^{1-\epsilon}.$$

Remark. Note that the ordering of the χ_j is defined by the values of the sums $S_{\chi_j}(x)$, and in this result we are using the same order for each X in our range.

Proof. We may assume that $\left(\frac{\log x}{\log q + \log\log x}\right)^{\epsilon/2} \geq 4$ else the result is trivial. We apply Corollary 10.8 with $x = x_i$, where $\log x_i = 2^{(1+\epsilon/3)^i + 1} \log q$ for $0 \leq i \leq I \ll (1/\epsilon) \log\log A$, with I chosen to be the smallest integer for which $y_I > x/L$. The characters excluded from the sum will be, say, $\chi_{j,i}$, $1 \leq j \leq k$ for $1 \leq i \leq I$: Let $\chi_1, \chi_2, \dots, \chi_K$ be the union of these sets of characters, so that $K \leq k(I+1) \ll \epsilon^{-3} \log\log A$. Hence (10.3.1) holds for all $y \in [x/U, x/L]$.

We can rewrite the sum in the Proposition as

$$\sum_{\substack{\chi \pmod q \\ \chi \neq \chi_j,\ j=1,\dots,K}} \left| \sum_{\substack{dm \leq X \\ L \leq d \leq U}} f(m)\overline{\chi}(m) f(d)\overline{\chi}(d)\Lambda(d) \right|.$$

We split this into subsums, depending on the size of d. For a given D in the range $L \leq D \leq U$ let $\Delta = \frac{D \log(q \log(X/D))}{q \log(X/D)}$. Then we get an upper bound of a sum of sums of the form

$$\sum_{\substack{\chi \pmod q \\ \chi \neq \chi_j,\ j=1,\dots,K}} \left| \sum_{D \leq d \leq D+\Delta} f(d)\overline{\chi}(d)\Lambda(d) \sum_{m \leq X/d} f(m)\overline{\chi}(m) \right|. \tag{10.3.2}$$

If we approximate the final sum by the sum with $m \leq X/D$ then the new version of (10.3.2) is the square root of

$$\left(\sum_{\substack{\chi \pmod q \\ \chi \neq \chi_j,\ j=1,\dots,K}} \left| \sum_{D \leq d \leq D+\Delta} f(d)\overline{\chi}(d)\Lambda(d) \sum_{m \leq X/D} f(m)\overline{\chi}(m) \right| \right)^2$$

$$\leq \sum_{\chi \pmod q} \left| \sum_{D \leq d \leq D+\Delta} f(d)\overline{\chi}(d)\Lambda(d) \right|^2 \sum_{\substack{\chi \pmod q \\ \chi \neq \chi_j,\ j=1,\dots,k-1}} \left| \sum_{m \leq X/D} f(m)\overline{\chi}(m) \right|^2$$

(10.3.4)

$$\ll e^{O(1/\epsilon)} \left(\Delta \cdot \rho_q'(f) \frac{X}{D} \left(\frac{\log(q \log X)}{\log(X/D)} \right)^{1-\epsilon} \right)^2,$$

by Lemma 9.3 and Corollary 10.8. The error in the approximation of the dth term is

$$\left|\sum_{X/d<m\le X/D} f(m)\overline{\chi}(m)\right| \ll \rho_q'(f)X\ \frac{\Delta}{D^2}$$

by Lemma 9.1, yielding a total error of the square root of

$$\sum_{\chi \pmod q}\left|\sum_{D\le d\le D+\Delta} f(d)\overline{\chi}(d)\Lambda(d)\right|^2 \cdot \varphi(q)\left(\rho_q'(f)\frac{X}{D}\ \frac{\Delta}{D}\right)^2 \ll \varphi(q)\left(\Delta\cdot\rho_q'(f)\frac{X}{D}\ \frac{\Delta}{D}\right)^2,$$

which is negligible compared to the bound in (10.3.4). Hence (10.3.2) is

$$\ll e^{O(1/\epsilon)}\Delta\rho_q'(f)\frac{X}{D}\left(\frac{\log(q\log x)}{\log(X/D)}\right)^{1-\epsilon}.$$

Summing over $\asymp D/\Delta$ intervals $[D, D+\Delta)$ that give a partition of $[D, 2D]$ yields

$$\ll e^{O(1/\epsilon)}\rho_q'(f)X\left(\frac{\log(q\log x)}{\log(X/D)}\right)^{1-\epsilon}.$$

Finally, summing over D of the form $U/2^j$ with $L\le D\le U$ yields

$$\ll e^{O(1/\epsilon)}\rho_q'(f)X\log x\left(\frac{\log(q\log x)}{\log x}\right)^{1-\epsilon},$$

and the result follows.

10.4. Multiplicative functions in arithmetic progressions.

Define

$$E_f^{(k)}(y;q,a) := \sum_{\substack{n\le y\\ n\equiv a \pmod q}} f(n) - \frac{1}{\varphi(q)}\sum_{j=1}^{k}\chi_j(a)\sum_{n\le y} f(n)\overline{\chi}_j(n).$$

By bounding each term in these sums, we have $|E_f^{(k)}(y;q,a)| \ll (k+1)\rho_q'(f)y/\varphi(q)$ which we now improve.

Theorem 10.10. *For any given $k\ge 2$ and sufficiently large x, if $x\ge X\ge \max\{x^{1/2}, q^{6+7\epsilon}\}$ then*

$$|E_f^{(k-1)}(X;q,a)| \ll e^{C\sqrt{k}}\ \frac{\rho_q'(f)X}{\varphi(q)}\ \left(\frac{\log(q\log x)}{\log x}\right)^{1-\frac{1}{\sqrt{k}}}\log\left(\frac{\log x}{\log(q\log x)}\right),$$

where the implicit constants are independent of f and k. If f is real and χ_1 is not then we can extend this to $k=1$ with exponent $1-\frac{1}{\sqrt{2}}$.

Remark. Note that the ordering of the χ_j is defined by the values of the sums $S_{\chi_j}(x)$, and in this result we are using the same order for each X in our range.

Proof. Fix $\epsilon > 0$ sufficiently small with $1/\sqrt{k} > \epsilon$. Let $L = q^{2+\epsilon}\log x$ and $U = x/q^{4+5\epsilon}$. By applying Lemma 9.1, with $\chi = \chi_0$ we have

$$\log x \sum_{\substack{n\le x\\ n\equiv a \pmod q}} f(n) = \sum_{\substack{n\le x\\ n\equiv a \pmod q}} f(n)\log_{[L,U]} n + O\left(\rho_q'(f)\frac{x}{\varphi(q)}\log(q\log x)\right).$$

Then by (2.2.2), and Lemma 9.1 summed over all $a \pmod q$, we obtain

$$\begin{aligned} E_f^{(K)}(x;q,a) &= \frac{1}{\varphi(q)}\sum_{j=K+1}^{\varphi(q)}\chi_j(a)\sum_{n\le x} f(n)\overline{\chi}_j(n)\frac{\log_{[L,U]} n}{\log x} + O\left(K\rho_q'(f)\frac{x}{\varphi(q)}\frac{\log(q\log x)}{\log x}\right)\\ &\ll e^{O(1/\epsilon)}\frac{\rho_q'(f)x}{\varphi(q)}\left(\frac{\log(q\log x)}{\log x}\right)^{1-\epsilon}, \end{aligned}$$

by Proposition 10.9, where $K \ll \epsilon^{-3}\log\log A$. By Cauchying and then Corollary 10.6 with $Q = L$, we obtain

$$\begin{aligned} |E_f^{(k)}(x;q,a) - E_f^{(K)}(x;q,a)| &\le \frac{1}{\varphi(q)}\sum_{j=k+1}^{K}\left|\sum_{n\le x} f(n)\overline{\chi}_j(n)\right|\\ &\le \frac{1}{\varphi(q)}\left((K-k)\sum_{j=k+1}^{K}\left|\sum_{n\le x} f(n)\overline{\chi}_j(n)\right|^2\right)^{1/2}\\ &\ll e^{O(\sqrt{k})}\rho_q'(f)\frac{x}{\varphi(q)}\left(\frac{\log(q\log x)}{\log x}\right)^{1-\frac{1}{\sqrt{k}}}, \end{aligned}$$

since $K \ll \log\log A$, and $\frac{1}{\sqrt{k+1}} < \frac{1}{\sqrt{k}}$. Applying the same argument again, we also obtain

$$|E_f^{(k-1)}(x;q,a) - E_f^{(k)}(x;q,a)| \ll e^{C\sqrt{k}}\,\frac{\rho_q'(f)x}{\varphi(q)}\,\left(\frac{\log(q\log x)}{\log x}\right)^{1-\frac{1}{\sqrt{k}}}\log\left(\frac{\log x}{\log(q\log x)}\right).$$

The result follows from using the triangle inequality and adding the last three inequalities.

11. Primes in arithmetic progression

Theorem 11.1. *For any $k \geq 2$ and $x \geq q^2$ there exists an ordering $\chi_1, \dots$ of the characters $\chi \pmod q$ such that*

$$\sum_{\substack{n \leq y \\ n \equiv a \pmod q}} \Lambda(n) - \frac{1}{\varphi(q)} \sum_{n \leq y} \Lambda(n) - \frac{1}{\varphi(q)} \sum_{j=1}^{k-1} \chi_j(a) \sum_{n \leq y} \Lambda(n) \overline{\chi}_j(n)$$

$$\ll e^{C\sqrt{k}} \frac{x}{\varphi(q)} \left(\frac{\log(q \log x)}{\log x} \right)^{1 - \frac{1}{\sqrt{k}}} \log^3 \left(\frac{\log x}{\log(q \log x)} \right).$$

Remark. Note that we can take $\chi_{\varphi(q)}$ to be χ_0 the principal character since the distance function

Corollary 11.2. *There exists a character $\chi \pmod q$ such that if $x \geq q^2$ then*

$$\sum_{\substack{n \leq x \\ n \equiv a \pmod q}} \Lambda(n) - \frac{1}{\varphi(q)} \sum_{n \leq x} \Lambda(n) - \frac{\chi(a)}{\varphi(q)} \sum_{n \leq x} \Lambda(n) \overline{\chi}(n) \ll \frac{x}{\varphi(q)} \left(\frac{\log(q \log x)}{\log x} \right)^{1 - \frac{1}{\sqrt{2}} - \varepsilon}.$$

We may remove the χ term unless χ is a real-valued character.

Remark. It would be nice to obtain the error in term of $|L(1+it, \chi)^{-1}| / \log x$, which is probably possible. In the case that χ is real one can then probably deduce that $t = 0$.

Proof of Corollary 11.2. We let $k = 2$ in Theorem 11.1 to deduce the first part. If χ is not real valued, then we know that

$$\left| \sum_{n \leq x} \Lambda(n) \overline{\chi}(n) \right| = \left| \sum_{n \leq x} \Lambda(n) \chi(n) \right| \leq |E_\Lambda^{(3)}(x; q, a) - E_\Lambda^{(2)}(x; q, a)|$$

and the result follows from Theorem 11.1.

Proof of Theorem 11.1. We may assume that $x \geq q^B$ for B sufficiently large, else the result follows from the Brun-Titchmarsh Theorem.

Let χ_0 be the principal character $\pmod q$. Let $Q = q \log x$ and $\nu = \log \left(\frac{\log x}{\log Q} \right)$. Let f be the totally multiplicative function for which $f(p) = 0$ for $p \leq Q$, $f(p) = -1$ for $p > Q$ and $f(p^b) = 0$ for all $b \geq 2$, and let $g(.)$ be the totally multiplicative function for which $g(p) = 0$ for $p \leq Q$ and $g(p) = 1$ for $p > Q$. We use the following variant of von Mangoldt's formula (1.3.3),

$$\Lambda_Q(n) := \sum_{dm = n} f(d) g(m) \log m = \begin{cases} \Lambda(n) & \text{if } p | n \implies p > Q, \\ 0 & \text{otherwise.} \end{cases}$$

Now

$$\sum_{\substack{n\le x\\ n\equiv b \pmod q}} (\Lambda(n)-\Lambda_Q(n)) \le \sum_{\substack{n\le x\\ p|n \implies p\le Q}} \Lambda(d) \ll \sum_{p\le Q}\log x \ll Q\frac{\log x}{\log Q}.$$

by the Brun-Titchmarsh theorem. Denote the left side of the equation in the Theorem as $E^{(k-1)}_{\Lambda,+}(x;q,a)$, and note that all of these sums can be expressed as mean-values of $\sum_{\substack{n\le x\\ n\equiv b \pmod q}} \Lambda(n)$, as b varies. Hence

$$E^{(k-1)}_{\Lambda,+}(x;q,a) - E^{(k-1)}_{\Lambda_Q,+}(x;q,a) \ll Q\frac{\log x}{\log Q}.$$

Now

$$\sum_{\substack{n\le x\\ n\equiv a \pmod q}} \Lambda_Q(n) = \sum_{d\le x} f(d) \sum_{\substack{m\le x/d\\ m\equiv a/d \pmod q}} g(m)\log m. \tag{11.1.1}$$

By the fundamental lemma of the sieve (see §$1\frac{1}{2}$) the sum

$$\sum_{\substack{m\le M\\ m\equiv b \pmod q}} g(m)\log m,$$

equals a main term that is independent of b plus an error term that is

$$\ll \frac{M\log M}{\varphi(q)\log Q}\cdot\frac{1}{u^{u+2}} + \frac{M}{q}e^{-c\sqrt{\log(M/q)}}, \tag{11.1.2}$$

where $M/q = Q^u$. Decomposing the other terms of $E^{(k-1)}_{\Lambda_Q}$ similarly, leads to inner sums

$$\frac{1}{\varphi(q)}\sum_{m\le x/d} g(m)\log m = \frac{1}{\varphi(q)}\sum_{(b,q)=1}\sum_{\substack{m\le x/d\\ m\equiv b \pmod q}} g(m)\log m,$$

$$\text{and } \frac{1}{\varphi(q)}\sum_{m\le x/d} g(m)\chi_j(m)\log m = \frac{1}{\varphi(q)}\sum_{(b,q)=1}\chi_j(m)\sum_{\substack{m\le x/d\\ m\equiv b \pmod q}} g(m)\log m,$$

and so we may apply the same sieve estimate to all of these inner sums, and thus the main terms cancel and we end up with an error term of k times (11.1.2) for $M = x/d$. We sum this up over all d in a range $x/Q^{2u} < d \le x/Q^u$ with $f(d)\ne 0$, to get

$$\ll \frac{x}{\varphi(q)}\cdot\frac{1}{u^u} + \frac{x}{q}\cdot e^{-c\sqrt{v\log Q}}.$$

When we sum this up to get a bound over all $d \le x/Q^\nu$ with $f(d) \ne 0$, we obtain

$$\ll \frac{x}{q}\left(\frac{\log Q}{\log x}\right)^2. \tag{11.1.3}$$

We are left with those $d > x/Q^\nu$ so that $m < Q^\nu$. The remaining sum in (11.1.1) is

$$\sum_{\substack{m<Q^\nu\\(m,q)=1}} g(m)\log m \sum_{\substack{x/Q^\nu<d\le x/m\\ d\equiv a/m \pmod q}} f(d).$$

There are analogous sums for the remaining terms in $E^{(k-1)}_{\Lambda_Q,+}(x;q,a)$ and so we need to bound

$$\sum_{m<Q^\nu} g(m)\log m\ |E^{(k-1)}_{f,+}(x/m;q,a/m) - E^{(k-1)}_{f,+}(x/Q^\nu;q,a/m)|,$$

where

$$E^{(k-1)}_{f,+}(y;q,a) := \sum_{\substack{n\le y\\ n\equiv a \pmod q}} f(n) - \frac{1}{\varphi(q)}\sum_{n\le y} f(n) - \frac{1}{\varphi(q)}\sum_{j=1}^{k-1}\chi_j(a)\sum_{n\le y} f(n)\overline{\chi}_j(n).$$

This is, in absolute value, $\le |E^{(k-1)}_f(y;q,a)|$ plus $\frac{1}{\varphi(q)}\left|\sum_{n\le y} f(n)\right|$. Now $\frac{1}{\varphi(q)}\left|\sum_{n\le y} f(n)\right|$ is appropriately small by Halász's theorem. For the remaining terms we use Theorem 10.10 to obtain the bound

$$\sum_{m<Q^\nu} g(m)\log m\ |E^{(k-1)}_{f}(x/m;q,a/m) - E^{(k-1)}_{f}(x/Q^\nu;q,a/m)|$$

$$\ll e^{C\sqrt{k}}\ \frac{\rho_q'(f)x}{\varphi(q)}\ \left(\frac{\log Q}{\log x}\right)^{1-\frac{1}{\sqrt{k}}} \log\left(\frac{\log x}{\log Q}\right)\sum_{m<Q^\nu} g(m)\frac{\log m}{m}$$

$$\ll e^{C\sqrt{k}}\ \frac{\nu^2 x}{\varphi(q)}\ \left(\frac{\log Q}{\log x}\right)^{1-\frac{1}{\sqrt{k}}} \log\left(\frac{\log x}{\log Q}\right).$$

since $\rho_q'(f) \ll 1/\log Q$. The result follows.

11.2. Linnik's Theorem.

Proposition 11.3. *Suppose that $x \geq q^A$ where A is chosen sufficiently large. If*

$$\left| \sum_{\substack{n\leq x \\ n\equiv a \pmod q}} \Lambda(n) - \frac{x}{\varphi(q)} \right| \gg \frac{x}{\varphi(q)} \Big/ \log^B\left(\frac{\log x}{\log Q}\right)$$

for some $B \geq 0$ where $Q = q\log x$, then there exists a real character $\chi \pmod q$ such that

$$\sum_{Q<p\leq x} \frac{1+\chi(p)}{p} \ll \log\log\left(\frac{\log x}{\log Q}\right).$$

Corollary 11.4. *If there are no primes $\leq x$ that are $\equiv a \pmod q$, where $x \geq q^A$ and A is chosen sufficiently large, then there exists a real character $\chi \pmod q$ such that $\chi(a) = -1$, and*

$$\sum_{\substack{q<p\leq x \\ \chi(p)=1}} \frac{1}{p} \ll 1. \tag{11.1.4}$$

Proof of Proposition 11.3. By Corollary 11.2 we immediately deduce from the hypothesis that

$$\left|\sum_{n\leq x} \Lambda(n)\overline{\chi}(n)\right| \gg \frac{x}{\varphi(q)} \Big/ \log^B\left(\frac{\log x}{\log Q}\right). \tag{11.1.5}$$

If we trace through the proof of Theorem 11.1, then we see that this implies that there exists y in the range $x^{1/2} < y \leq x$ for which

$$\left|\sum_{n\leq y} f(n)\overline{\chi}(n)\right| \gg \frac{y}{\log Q}\left(\frac{\log Q}{\log x}\right)^\kappa \Big/ \log^{B+2}\left(\frac{\log x}{\log Q}\right),$$

for some y in the range $x^{1/2} < y \leq x$. We observe that there are $\asymp \frac{y}{\log Q}$ non-zero terms in the sum on the right, so we see that $f(n)\chi(n) = 1$ for significantly more than half the values of $n \leq y$ for which $f(n)\chi(n) \neq 0$.

We do not get a useful bound on the sum from Halász's Theorem, but we can improve (3.1.2) by taking into account the fact that $f(p) = 0$ if prime $p \leq Q$. First note that $S(N) = 1$ for all $N \leq Q$, so we can reduce the range in the integral for α, throughout the proof of Proposition 1, to $\frac{1}{\log x} \leq \alpha \leq \frac{1}{\log Q}$. Moreover in the first displayed equation we can change the error term from $\ll \frac{N}{\log N}$ to $\ll \frac{1}{\log Q}\frac{N}{\log N}$ for $N \geq Q$. This allows us to replace the error term in the second displayed equation from $\ll \log\log x$ to $\ll 1+\frac{1}{\log Q}\log\left(\frac{\log x}{\log Q}\right)$. Hence we can restate Proposition 1 with the range for α and the $\log\log x$ changed in this way. Now we use the bound $|F(1+\alpha+it)| \le |F(1+iu)| + O\left(\frac{\alpha}{T}\frac{\log x}{\log Q}\right)$ throughout this range, as in Lemma 2.7, to obtain

$$\Big|\sum_{n\le y} f(n)\overline{\chi}(n)\Big| \ll \frac{y}{\log Q}\log\left(\frac{\log x}{\log Q}\right)\exp\left(-\min_{|u|\le \log^2 x}\sum_{Q<p\le x}\frac{1-\text{Re }(f(p)\overline{\chi}(p)p^{-iu})}{p}\right).$$

Combining the last two displayed equations yields that

$$\sum_{Q<p\le x}\frac{1+\text{Re }(\chi(p)p^{iu})}{p} \ll \log\log\left(\frac{\log x}{\log Q}\right).$$

Note that if $|u| \gg 1/\log x$ then this sum is $\gg \log\left(\frac{\log x}{\log Q}\right)$, and hence we may assume that $u=0$. The result follows.

Proof of Corollary 11.4. By Corollary 11.2 we know that for all y in the range $Q := q^A \le y \le x$ we have

$$\sum_{n\le y}\Lambda(n)\chi(n) = -\chi(a)y + O\left(y\,\left(\frac{\log(q\log y)}{\log y}\right)^{\kappa}\right)$$

where $0<\kappa<1-\frac{1}{\sqrt{2}}$. By partial summation, we deduce that

$$\sum_{Q<p\le x}\frac{\chi(a)+\chi(p)}{p} \ll 1.$$

Comparing this to the conclusion of Proposition 11.3, we deduce the result.

Proposition 11.5. *If (11.1.4) holds for $x=q^A$ where A is sufficiently large, and if $\chi(a)=1$ then there are primes $\le x$ that are $\equiv a \pmod q$.*

12. Exponential Sums

Proposition 12.1.

$$\sum_{\substack{n\le x\\ (n,q)=1}} f(n)e\left(\frac{an}{q}\right) - \frac{1}{\varphi(q)}\sum_{i=1}^{k-1}\overline{\chi}_i(a)g(\chi_i)\sum_{n\le x} f(n)\overline{\chi}_i(n)$$
$$\ll \frac{\sqrt{q}}{\varphi(q)}e^{O(\sqrt{k})}\rho_q'(f)x\left(\frac{\log Q}{\log x}\right)^{1-\frac{1}{\sqrt{k}}} + \rho_q'(f)x\frac{\log Q}{\log x}.$$

Note that the second error term dominates unless $q \ll (\log x)^{2+o(1)}$, a fact that is well-known from the paper of Montgomery and Vaughan. We will therefore assume that

Corollary 12.2. *If $q \le (\log x)^{2+o(1)}$ then there exists a primitive character $\psi \pmod r$ for some $r|q$ such that*

$$\sum_{n\le x} f(n)e\left(\frac{an}{q}\right) = \frac{\overline{\psi}(a)g(\psi)}{\varphi(r)}\ \psi(q_r)\ \sum_{n\le x} f(n)\overline{\psi}(n_r)h(n)$$
$$+O\left(\frac{1}{\sqrt{q}}\prod_{p|q}\left(1+\frac{3}{\sqrt{p}}\right)\rho_q'(f)x\left(\frac{\log\log x}{\log x}\right)^{1-\frac{1}{\sqrt{2}}} + \left(\frac{q}{\varphi(q)}\right)^2\rho_q'(f)x\frac{\log\log x}{\log x}\right).$$

where n_r is the largest divisor of n that is coprime to r, and h is the multiplicative function with $h(p^b)=1$ if $p\nmid q$, and

$$h(p^b) = \begin{cases} 1 & \text{if } r_p \ge 1 \text{ and } b = q_p - r_p;\\ 1 & \text{if } r_p = 0 \text{ and } b \ge q_p;\\ -1/(p-1) & \text{if } r_p = 0 \text{ and } b = q_p - 1;\\ 0 & \text{otherwise}\end{cases}$$

where $p^{q_p}\|q$ and $p^{r_p}\|r$. If f is real then we may assume that ψ is real.

Proof of Proposition 12.1. The same proof as in Lemma 9.1 yields that if $x \ge Q^{2+\epsilon}$ and $Q \ge q$ then

$$\left|\sum_{\substack{n\le x\\ n\equiv a \pmod q}} f(n)\overline{\chi}(n)e(\alpha n)\mathcal{L}(n)\right| \ll \rho_q'(f)\frac{x}{\varphi(q)},$$

where $\mathcal{L}(n) = 1, \log(x/n)$, $\frac{\log_{\le Q} n}{\log Q}$ or $\frac{\log_{\ge x/Q} n}{\log Q}$, and then summing over all a that

$$\left|\sum_{n\le x} f(n)\overline{\chi}(n)e(\alpha n)\mathcal{L}(n)\right| \ll \rho_q'(f)x.$$

Let $U = X/q^{4+5\epsilon}$, $L = q^{2+\epsilon}\log x$ and $Q = q\log x$. Therefore

$$\log x \sum_{n\le x} f(n)\overline{\chi}(n)e(\alpha n) = \sum_{n\le x} f(n)\overline{\chi}(n)e(\alpha n)\log_{[L,U]} n + O\left(\rho_q'(f)x\log Q\right). \tag{12.1}$$

Taking $\alpha = a/q$ and $\chi = \chi_0$ in (12.1), we obtain.

$$\log x \sum_{\substack{n\le x\\(n,q)=1}} f(n)e\left(\frac{an}{q}\right) = \sum_{\substack{n\le x\\(n,q)=1}} f(n)e\left(\frac{an}{q}\right)\log_{[L,U]} n + O\left(\rho_q'(f)x\log Q\right).$$

Using the expansion

$$e\left(\frac{b}{q}\right) = \frac{1}{\varphi(q)}\sum_{\chi \bmod q}\overline{\chi}(b)g(\chi)$$

when $(b,q)=1$, we deduce that

$$\sum_{\substack{n\le x\\(n,q)=1}} f(n)e\left(\frac{an}{q}\right)\log_{[L,U]} n = \frac{1}{\varphi(q)}\sum_{\chi \bmod q}\overline{\chi}(a)g(\chi)\sum_{n\le x} f(n)\overline{\chi}(n)\log_{[L,U]} n.$$

Taking $\alpha = 0$ in (12.1) yields

$$\log x\sum_{n\le x} f(n)\overline{\chi}(n) = \sum_{n\le x} f(n)\overline{\chi}(n)\log_{[L,U]} n + O\left(\rho_q'(f)x\log Q\right).$$

Combining the last three displayed equations yields that

$$\begin{aligned}&\sum_{\substack{n\le x\\(n,q)=1}} f(n)e\left(\frac{an}{q}\right) - \frac{1}{\varphi(q)}\sum_{i=1}^{K}\overline{\chi}_i(a)g(\chi_i)\sum_{n\le x} f(n)\overline{\chi}_i(n)\\ &\le \frac{\sqrt{q}}{\varphi(q)}\frac{1}{\log x}\sum_{\substack{\chi \bmod q\\ \chi\ne\chi_1,\cdots,\chi_K}}\left|\sum_{n\le x} f(n)\overline{\chi}(n)\log_{[L,U]} n\right| + O\left(\rho_q'(f)x\frac{\log Q}{\log x}\right),\end{aligned} \tag{12.2}$$

as each $|g(\chi)|\le\sqrt{q}$, assuming $K\le q^{1/3}$.

We now proceed much as in the proof of Theorem 10.10. First apply Proposition 10.9. Let $x = q^A$ and $K \asymp \epsilon^{-3}\log\log A$, so that the first part of (12.2) is

$$\ll \frac{\sqrt{q}}{\varphi(q)}e^{O(1/\epsilon)}\rho_q'(f)x\left(\frac{\log Q}{\log x}\right)^{1-\epsilon}.$$

For the terms with $k+1\le i\le K$ we use Cauchy-Schwarz and then Corollary 10.6 to get

$$\ll \frac{\sqrt{q}}{\varphi(q)}e^{O(\sqrt{k})}\sqrt{K}\rho_q'(f)x\left(\frac{\log Q}{\log x}\right)^{1-\frac{1}{\sqrt{k+1}}}.$$

If $K > q^{1/3}$ we use this same method to bound all the terms with $k+1 \leq i \leq \varphi(q)$, so we would replace $K^{1/2}$ in this upper bound by $K^{3/2}$. We bound the kth directly using Halász's theorem. This then implies the result.

Proof of Corollary 12.2. We take $k = 2$ in Proposition 12.1, with $\chi = \chi_1$, to obtain

$$\sum_{\substack{n \leq x \\ (n,q)=1}} f(n) e\left(\frac{an}{q}\right) - \frac{\overline{\chi}(a) g(\chi)}{\varphi(q)} \sum_{n \leq x} f(n)\overline{\chi}(n)$$

$$\ll \frac{\sqrt{q}}{\varphi(q)} e^{O(1)} \rho_q'(f) x \left(\frac{\log\log x}{\log x}\right)^{1-\frac{1}{\sqrt{2}}} + \rho_q'(f) x \frac{\log\log x}{\log x}.$$

Now if $d|q$ then, writing $n = mq/d$ when $(q/d)|n$,

$$\sum_{\substack{n \leq x \\ (n,q)=q/d}} f(n) e\left(\frac{an}{q}\right) = \sum_{\substack{m \leq x/(q/d) \\ (m,d)=1}} f(mq/d) e\left(\frac{am}{d}\right).$$

We apply (12.3) to this and sum over all $d|q$. We claim that each χ can be assumed to be the same, for if we have $\psi \pmod d$, and the character induced by $\psi \pmod q$ is not χ, then we can use the lower bound on $M_{f\overline{\psi}}(x,T)$ implicit from the proof, to get a good lower bound on $M_{f\overline{\psi}}(x/d,T)$ (note that we can ignore the prime factors of q in this calculation since they are $\leq Q$). To calculate the main term, note that if $\chi \pmod d$ is induced from a character $\psi \pmod r$ where $r|d|q$, then $g(\chi) = \mu(d/r)\psi(d/r)g(\psi)$. The first part of the result follows.

Finally if f is real then $\left|\sum_{n\leq x} f(n)\overline{\chi}(n)\right| = \left|\sum_{n\leq x} f(n)\chi(n)\right|$, so this is only bigger than the error term if χ is real.

Evaluation. By Theorem 4 of [Decay] and Corollary 3.5, we deduce that

$$\frac{\psi(q_r)}{\varphi(r)} \sum_{n \leq x} f(n)\overline{\psi}(n_r) h(n) = \frac{\kappa}{\varphi(q)} \frac{x^{it}}{1+it} \sum_{\substack{n \leq x \\ (n,q)=1}} \frac{f(n)\overline{\psi}(n)}{n^{it}} + O\left(\frac{2^{\omega(q)-\omega(r)}}{\varphi(q)} \frac{x\log\log x}{(\log x)^{2-\sqrt{3}}}\right)$$

where

$$\kappa := \frac{1}{(q/r)^{it}} \prod_{\substack{p^{q_p} \| q/r \\ p|r}} f(p^{q_p}) \prod_{\substack{p^{q_p} \| q/r \\ p\nmid r}} \left(\sum_{j \geq 0} \frac{F_p(p^j) - F_p(p^{j-1})}{p^j}\right)$$

and $F_p(p^j) = f(p^{q_p+j})\overline{\psi}(p^j)/p^{ijt}$. Now, since $2 - \sqrt{3} < 1 - \frac{1}{\sqrt{2}}$ we deduce from this and Corollary 12.2 that if $q \leq (\log x)^{2+o(1)}$ then

$$\sum_{n \leq x} f(n) e\left(\frac{an}{q}\right) = \kappa \frac{\overline{\psi}(a) g(\psi)}{\varphi(q)} \frac{x^{it}}{1+it} \sum_{\substack{n \leq x \\ (n,q)=1}} \frac{f(n)\overline{\psi}(n)}{n^{it}}$$

$$+ O\left(\frac{3^{\omega(q)}}{\sqrt{q}} \frac{x\log\log x}{(\log x)^{2-\sqrt{3}}} + \left(\frac{q}{\varphi(q)}\right)^2 \rho_q'(f) x \frac{\log\log x}{\log x}\right).$$

If f is real then we may assume that ψ is real.

13. Primes in progressions, on average

Suppose that the character $\chi \pmod q$ is induced from the primitive character $\psi \pmod r$. Then we write $\text{cond}\,\chi = q$ and $\text{cond}^*\chi = r$.

13.1. The Barban-Davenport-Halberstam-Montgomery-Hooley Theorem. We begin with a technical lemma; most of the proof is left as an exercise.

Lemma 13.1. *Let* $c := \prod_p \left(1 + \frac{1}{p(p-1)}\right)$ *and* $\gamma, := \gamma - \sum_p \frac{\log p}{p^2-p+1}$. *Then*

$$\sum_{r\le R} \frac{1}{\varphi(r)} = c\log R + c\gamma' + O\left(\frac{\log R}{R}\right),$$
$$\sum_{r\le R} \frac{r}{\varphi(r)} = cR + O(\log R),$$
$$\sum_{r\le R} \frac{r^2}{\varphi(r)} = \frac{c}{2}R^2 + O(R\log R).$$

Also

$$\sum_{\substack{r\le R\\ m|r}} \frac{1}{\varphi(r)} = \frac{1}{\varphi(m)} \prod_{p\nmid m}\left(1+\frac{1}{p(p-1)}\right)\left(\log\frac{R}{m} + \gamma - \sum_{p\nmid m}\frac{\log p}{p^2-p+1}\right) + O\left(\frac{\log R/m}{R}\right).$$

Proof. We can write $r/\varphi(r) = \sum_{d|r} \mu^2(d)/\varphi(d)$ to obtain in the first case

$$\sum_{r\le R}\frac{1}{\varphi(r)} = \sum_{r\le R}\frac{1}{r}\sum_{d|r}\frac{\mu^2(d)}{\varphi(d)} = \sum_{d\le R}\frac{\mu^2(d)}{\varphi(d)}\sum_{\substack{r\le R\\ d|r}}\frac{1}{r} = \sum_{d\le R}\frac{\mu^2(d)}{d\varphi(d)}(\log\frac{R}{d} + \gamma + O(\frac{d}{R}))$$
$$= c(\log R + \gamma') + O\left(\frac{\log R}{R}\right),$$

by (1.2.1). The next two estimates follow analogously but more easily. The last estimate is an easy generalization of the first.

Proposition 13.2.

$$\sum_{q\le Q}\frac{1}{\varphi(q)}\sum_{\substack{\chi \pmod q\\ \text{cond}^*\chi\ge R}}\left|\sum_{n=X+1}^{X+N} a_n\chi(n)\right|^2 \ll \left(\frac{N}{R}\log Q + Q\right)\log\log Q\sum_{n=X+1}^{X+N}|a_n|^2.$$

Proof. Suppose that the character $\chi \pmod q$ is induced from the primitive character $\psi \pmod r$. Let m be the product of the the primes that divide q but not r and write $q = rm\ell$ so that $(r,m) = 1$, and $p|\ell \implies p|rm$. Hence $\varphi(q) = \varphi(r)\varphi(m)\ell$ and

$$\sum_n a_n\chi(n) = \sum_{n:\ (n,m)=1} a_n\psi(n);$$

and therefore the left side of the above equation equals

$$\sum_{m\le Q}\frac{\mu^2(m)}{\varphi(m)}\sum_{\substack{R\le r\le Q/m\\ (r,m)=1}}\frac{1}{\varphi(r)}\sideset{}{^*}\sum_{\psi \pmod r}\Bigg|\sum_{\substack{X<n\le X+N\\ (n,m)=1}}a_n\psi(n)\Bigg|^2\sum_{\substack{\ell\le Q/rm\\ p|\ell\implies p|rm}}\frac{1}{\ell}.$$

The last sum is $\le \frac{r}{\varphi(r)}\cdot\frac{m}{\varphi(m)}$. We partition the sum over r into dyadic intervals $y<r\le 2y$; in such an interval we have $\frac{r}{\varphi(r)^2}\ll\frac{r}{\varphi(r)}\cdot\frac{\log\log y}{y}$, and so by (8.4) the above becomes

$$\ll \log\log Q\sum_{m\le Q}\frac{\mu^2(m)m}{\varphi(m)^2}\sum_{\substack{y=2^iR,\ i=0,\dots I\\ 2^IR:=Q/m}}\frac{1}{y}(N+y^2)\sum_{n=X+1}^{X+N}|a_n|^2$$
$$\ll \log\log Q\sum_{m\le Q}\frac{\mu^2(m)m}{\varphi(m)^2}\left(\frac{N}{R}+\frac{Q}{m}\right)\sum_{n=X+1}^{X+N}|a_n|^2,$$

which implies the result.

Let
$$\psi^{(R)}(x;q,a)=\psi(x;q,a)-\frac{1}{\varphi(q)}\sum_{\substack{r\le R\\ r|q}}\sum_{\substack{\chi \pmod q\\ \text{cond}^*\chi=r}}\overline{\chi}(a)\sum_{n\le x}\chi(n)\Lambda(n),$$

so that $\psi^{(1)}(N;q,a)=\psi(N;q,a)-\frac{\psi(N)}{\varphi(q)}$.

Corollary 13.3. *For* $\log N\le R\le\sqrt{Q}$ *with* $Q\le N$ *we have*

$$\sum_{q\le Q}\sum_{(a,q)=1}\left|\psi(N;q,a)-\frac{\psi(N)}{\varphi(q)}\right|^2\ll\frac{\log Q}{\log R}\sum_{q\le R^2}\sum_{(a,q)=1}\left|\psi(N;q,a)-\frac{\psi(N)}{\varphi(q)}\right|^2$$
$$+O\left(\frac{N^2\log^3 N}{R}+QN\log N\log\log N\right).$$

Proof. By (8.2), and taking $a_n=\Lambda(n), X=0$ in Proposition 13.2, we deduce that

$$\sum_{q\le Q}\sum_{(a,q)=1}\left|\psi^{(R)}(N;q,a)\right|^2\ll\left(\frac{N}{R}\log N+Q\right)N\log N\log\log N.$$

by using the prime number theorem. Now, if $\chi \pmod q$ is induced from $\psi \pmod r$ then

$$\sum_{n\le N}\chi(n)\Lambda(n)=\sum_{n\le N}\psi(n)\Lambda(n)-\sum_{\substack{p^a\le N\\ p|q,\ p\nmid r}}\psi(p^a)\log p,$$

hence the error term in replacing χ by ψ is $\ll (\omega(q)-\omega(r))N\log N$, and so in total is

$$\ll \sum_{r\le R}\sum_{\substack{q\le Q\\ r|q}} \frac{\omega(q)-\omega(r)}{\varphi(q)} N\log N \ll N\log N\log Q\log R\log\log Q \ll N(\log N)^{2+\epsilon},$$

which is smaller than the above.

What remains is, by (8.2),

$$\sum_{q\le Q}\frac{1}{\varphi(q)}\sum_{\substack{r\le R\\ r|q}}\ \sideset{}{^*}\sum_{\psi \pmod r}\left|\sum_{n\le N}\psi(n)\Lambda(n)\right|^2 = \sum_{r\le R}\frac{1}{\varphi(r)}\ \sideset{}{^*}\sum_{\psi \pmod r}\left|\sum_{n\le N}\psi(n)\Lambda(n)\right|^2\sum_{\substack{q\le Q\\ r|q}}\frac{\varphi(r)}{\varphi(q)}.$$

By Lemma 13.1 this last sum is $\asymp \log 2Q/r$. Replacing Q here by R^2, we see that this quantity is

$$\asymp \frac{\log Q}{\log R}\sum_{q\le R^2}\frac{1}{\varphi(q)}\sum_{\substack{r\le R\\ r|q}}\ \sideset{}{^*}\sum_{\psi \pmod r}\left|\sum_{n\le N}\psi(n)\Lambda(n)\right|^2$$

and our result follows, by replacing Q with R^2 in the arguments above.

The Siegel-Walfisz Theorem states that for any fixed $A,B>0$ one has

$$\psi(N;q,a)-\frac{\psi(N)}{\varphi(q)} \ll \frac{N}{\varphi(q)\log^B N},$$

uniformly for $q\ll \log^A x$ and $(a,q)=1$.

Proposition 13.4. *Assume the Siegel-Walfisz Theorem. If $N/(\log N)^C\le Q\le N$ then*

$$\sum_{q\le Q}\sum_{(a,q)=1}\left|\psi(N;q,a)-\frac{\psi(N)}{\varphi(q)}\right|^2 = NQ\log N + O\left(NQ\log(N/Q)\right).$$

Proof. Let $Q'=Q/\log^2 N$ and $R=(N\log^3 N)/Q$, and assume the Siegel-Walfisz Theorem with $A=2C+6$ and $B=C+1$ so that Corollary 13.3 yields

$$\sum_{q\le Q'}\sum_{(a,q)=1}\left|\psi(N;q,a)-\frac{\psi(N)}{\varphi(q)}\right|^2 \ll QN.$$

We are left with the sum for $Q'<q\le Q$, which we will treat as the sum for $Q'<q\le N$, minus the sum for $Q<q\le N$. We describe only how we manipulate the second sum, as the first is entirely analogous.

Now the qth term in our sum equals

$$\sum_{p\le N}\log^2 p+2\sum_{\substack{p_1<p_2\le N\\ p_2\equiv p_1 \pmod r}}\log p_1\log p_2-\frac{\psi(N)^2}{\varphi(q)}, \tag{13.5}$$

plus a small, irrelevant error term made up of contributions from prime powers that divide q. We will sum the middle sum over all q in the range $Q<q\le N$ and then subtract this from the sum for $Q'<q\le N$. We write $p_2=p_1+qr$ then $r\le N/q<N/Q$, so that $p_2\equiv p_1 \pmod r$ with $N\ge p_2\ge p_1+Qr$, and therefore the sum equals

$$\begin{aligned}\sum_{r\le N/Q}\sum_{p\le N-Qr}&\{\psi(N;r,p)-\psi(p+Qr;r,p)\}\ \log p\\ &=\sum_{r\le N/Q}\frac{1}{\varphi(r)}\sum_{p\le N-Qr}(N-p-Qr)\log p+O\left(\sum_{r\le N/Q}\frac{N^2}{\varphi(r)\log^B N}\right)\\ &=\frac12\sum_{r\le N/Q}\frac{(N-Qr)^2}{\varphi(r)}+O\left(NQ\right),\end{aligned}$$

by the Siegel-Walfisz theorem. We now subtract this from the same expression with Q replaced by Q', add the sum, for $Q'<q\le Q$, of the outer terms in (13.5), which are both evaluated using the prime number theorem. By estimating the various sums that arise using Lemma 13.1, we obtain our result, after some remarkable cancelations.

Remark. We really could do with Siegel's theorem and then the Siegel Walfisz theorem. These probably can all be expressed as theorems about $L(1,\chi)$ so it should be doable.

We can surely incorporate the discussion on binary quadratic forms in my Italy paper.

We can get a good prime number theorem from the elementary proofs since they helpfully avoid any zeros! We could just quote this or we could find our own version.

13.2. The Bombieri-Vinogradov Theorem.

Proposition 13.5.

$$\begin{aligned}\sum_{q\le Q}\frac{1}{\varphi(q)}&\sum_{\substack{\chi \pmod q\\ \mathrm{cond}^*\chi\ge R}}\left|\sum_{m=X+1}^{X+M}a_m\chi(m)\right|\left|\sum_{n=Y+1}^{Y+N}b_n\chi(n)\right|\\ &\ll\left(\frac{\sqrt{MN}}{R}\log Q+Q+(\sqrt{M}+\sqrt{N})\log^2 Q\right)\log\log Q\sqrt{\sum_{m=X+1}^{X+M}|a_m|^2\cdot\sum_{n=Y+1}^{Y+N}|b_n|^2}.\end{aligned}$$

Proof. We proceed analogously to the proof of Proposition 13.2 to obtain

$$\ll\log\log Q\sum_{\ell\le Q}\frac{\mu^2(\ell)\ell}{\varphi(\ell)^2}\sum_{\substack{y=2^iR,\ i=0,\dots I\\ 2^IR:=Q/\ell}}\frac1y\sum_{y<r\le 2y}\frac{r}{\varphi(r)}\sideset{}{^*}\sum_{\psi \pmod r}\left|\sum_{\substack{m=X+1\\(m,\ell)=1}}^{X+M}a_m\chi(m)\right|\left|\sum_{\substack{n=Y+1\\(n,\ell)=1}}^{Y+N}b_n\chi(n)\right|.$$

The square of the sum over r may be bounded, after Cauchying, by

$$\sum_{y<r\le 2y} \frac{r}{\varphi(r)} \sum_{\psi \pmod r}^{*} \left| \sum_{\substack{m=X+1\\(m,\ell)=1}}^{X+M} a_m \chi(m) \right|^2 \cdot \sum_{y<r\le 2y} \frac{r}{\varphi(r)} \sum_{\psi \pmod r}^{*} \left| \sum_{\substack{n=Y+1\\(n,\ell)=1}}^{Y+N} b_n \chi(n) \right|^2$$
$$\ll (M+y^2) \sum_{m=X+1}^{X+M} |a_m|^2 \cdot (N+y^2) \sum_{n=Y+1}^{Y+N} |b_n|^2$$

by applying (8.4). Hence the above is $\sqrt{\sum_{m=X+1}^{X+M} |a_m|^2 \cdot \sum_{n=Y+1}^{Y+N} |b_n|^2}$ times

$$\ll \log\log Q \sum_{\ell \le Q} \frac{\mu^2(\ell)\ell}{\varphi(\ell)^2} \sum_{\substack{y=2^i R,\ i=0,\dots I\\ 2^I R:=Q/\ell}} \left(\frac{\sqrt{MN}}{y} + \sqrt{M} + \sqrt{N} + y \right),$$

and the result follows.

Proposition 13.6. *Suppose that a_n, b_n are given sequences with $a_n, b_n = 0$ for $n \le R^2$, and $|a_n| \le a_0,\ |b_n| \le b_0$ for all $n \le x$. If $c_N := \sum_{mn=N} a_m b_n$ then*

$$\sum_{q\le Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod q\\ \text{cond}^*\chi \ge R}} \left| \sum_{N\le x} c_N \chi(N) \right| \ll a_0 b_0 \left(\frac{x}{R} \log^2 x + Q\sqrt{x} \log x \log R \right) \log\log x.$$

Proof. We begin by noting that

$$\sum_{N\le x} c_N \chi(N) = \sum_{mn\le x} a_m \chi(m) \cdot b_n \chi(n).$$

We will partition the pairs m, n with $mn \le x$ in order to apply Proposition 13.5 to many different sums. For the intervals $X < m \le X+M,\ Y < n \le Y+N$, Proposition 13.5 yields the upper bound

$$a_0 b_0 \sqrt{MN} \left(\frac{\sqrt{MN}}{R} \log Q + Q + (\sqrt{M} + \sqrt{N}) \log^2 Q \right) \log\log Q$$

We now describe the partition for m in the range $X < m \le 2X$. Let $Y = x/X$. We begin with all $X < m \le 2X,\ n \le Y/2$. Then in step k, with $k = 1, 2, \dots K$, we take

$$\left(1 + \frac{2j}{2^k}\right) X < m \le \left(1 + \frac{2j+1}{2^k}\right) X, \quad Y\Big/\left(1 + \frac{2j+2}{2^k}\right) < n \le Y\Big/\left(1 + \frac{2j+1}{2^k}\right),$$

for $0 \le j \le 2^{k-1}-1$. The total upper bound from all these is $a_0 b_0 \log\log Q$ times

$$\sqrt{XY}\left(\frac{\sqrt{XY}}{R}\log Q + Q + (\sqrt{X}+\sqrt{Y})\log^2 Q\right) + \sum_{k=1}^{K} 2^{k-1}\cdot\frac{\sqrt{XY}}{2^k}\left(\frac{\sqrt{XY}}{2^k R}\log Q + Q + \frac{(\sqrt{X}+\sqrt{Y}}{2^{k/2}}\right.$$

$$\ll \sqrt{XY}\left(\frac{\sqrt{XY}}{R}\log Q + KQ + (\sqrt{X}+\sqrt{Y})\log^2 Q\right).$$

Now, for each such m the number of n not yet accounted for is $\ll Y/2^K$, thus contributing a total of $\ll a_0 b_0 XY/2^K$. This is negligible compared to the main term, taking $2^K \approx R$. We now sum up the upper bound over $X = 2^j R^2$ for $j = 0, 1, 2, \ldots, J$ where $2^J = x/R^2$ (since if $m < R^2$ then $b_m = 0$, and if $m > x/R^2$ then $n < R^2$ and so $c_n = 0$), to obtain the claimed upper bound.

Corollary 13.7a. *If $R \le x^{1/4}$ and $Q \le x^{1/2}$ then*

$$\sum_{q \le Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ \text{cond}^*\chi \ge R}} |\psi(x,\chi) - G(x,\chi)| \ll \left(\frac{x}{R}\log x + Q\sqrt{x}\log R\right)\log^2 x \log\log x,$$

where $G(x,\chi) := \sum'_{n \le x} \chi(n)\log n$, the sum over integers whose prime factors are all $> R^2$.

Proof. Let $a_1 = 0$ with $a_n = f(n)$ for $n > 1$ and $b_m = g(m)\log m$ in Proposition 13.6, where f and g are as in the proof of Theorem 11.1, for the decomposition of Λ_{R^2}. The missing contribution of the powers of the primes $\le R^2$ is $\le Q\sum_{p \le R^2} \log x \ll QR^2 \log x/\log R$.

Remark. We should be able to get rid of the sum of the $|G(x,\chi)|$, surely! Each $G(x,\chi)$ with χ non-principal is a sum over $\varphi(q)$ arithmetic progressions, where the main term cancels. Hence by (11.1.2) we have

$$|G(x,\chi)| \ll \frac{x\log x}{\log R}\cdot\frac{1}{u^{u+2}} + xe^{-8c\sqrt{\log(x/q)}}$$

which is not small enough. Another idea is to write $G(x,\chi)$ as a convolution: $G(x,\chi) = \sum_{bd \le x} \mu(d)\chi(bd)\log bd$, where the d are restricted to integers that are R^2-smooth. We can again apply Proposition 13.6, so we need only worry about the $b, d \le R^2$. If $d \le R^2$ we are taking sums like $\sum_{b \le x/d} \chi(b)\log bd$ which is $\ll \sqrt{q}\log q \log x$ by the Polya-Vinogradov Theorem (and partial summation). We sum this up to get $\ll Q^{3/2}R^2\log^2 x$ which is $\ll x^{3/4+\epsilon}$ so easily acceptable. It is when $b \le R^2$ that we have problems, that is with sums $\sum_{d \le x/b} \mu(d)\chi(d)\log bd$; you would think it would be easy! We again get the sort of terms we have in the previous display.

We now prove a version of the Bombieri-Vinogradov Theorem:

Corollary 13.7b. *If $R \le e^{c\sqrt{\log x}}$ and $Q \le x^{1/2}$ then*

$$\sum_{q \le Q} \max_{(a,q)=1} |\psi^{(R)}(x;q,a)| \ll \left(\frac{x}{R}\log x + Q\sqrt{x}\log R\right)\log^2 x \log\log x.$$

Proof. The estimate in (11.1.2) yields

$$\sum_{q\le Q}\max_{(a,q)=1}|G^{(R)}(x;q,a)| \ll \sum_{q\le Q}\left(1+\sum_{\substack{r\le R\\ r|q}}\varphi^*(r)\right)\left(\frac{x\log x}{\varphi(q)\log R}\cdot\frac{1}{u^{u+2}}+\frac{x}{q}e^{-8c\sqrt{\log(x/q)}}\right)$$
$$\ll R\log Q\left(\frac{x\log x}{\log R}\cdot\frac{1}{u^{u+2}}+xe^{-4c\sqrt{\log x}}\right)\ll\frac{x}{R},$$

where $x^{1/2}=R^{2u}$, since $x/q\ge x/Q\ge x^{1/2}$, and $\varphi^*(r)$ is the number of primitive characters of conductor r (here $\varphi^*(p)=p-2$ and $\varphi^*(p^b)=p^{b-2}(p-1)^2$). Combining this with Corollary 13.7a gives the result.

Remark. We should try to improve the small sieve argument for the contribution of g. If we can then we can get some amazing consequences for moduli q without small prime factors (e.g. prime q), because if q has no prime factor $\le R$ then $\psi^{(R)}(x;q,a)=\psi(x;q,a)-\psi(x)/\varphi(q)$. At the moment we can claim that for primes q in the range $e^{c\sqrt{\log x}}<q<\sqrt{x}/e^{c\sqrt{\log x}}$ we win in the error term by a factor of $e^{c\sqrt{\log x}}$.

Corollary 13.8. *Assume the Siegel-Walfisz Theorem. Fix $A>0$. If $x^{1/2}/\log^A x<Q\le x^{1/2}$ then*

$$\sum_{q\le Q}\max_{(a,q)=1}\left|\psi(x;q,a)-\frac{\psi(x)}{\varphi(q)}\right|\ll Q\sqrt{x}(\log x\log\log x)^2$$

Proof. Let $R=\log^{A+1}x$ in Corollary 13.7b.

Corollary 13.9. *Assume the Siegel-Walfisz Theorem. If $x^{1/2}/e^{(c/2)\sqrt{\log x}}<Q\le x^{1/2}$ then*

$$\sum_{q\le Q}\max_{(a,q)=1}\left|\psi(x;q,a)-\frac{\psi(x)}{\varphi(q)}-\overline{\chi}_1(a)\frac{\psi(x,\chi_1)}{\varphi(q)}\right|\ll Q\sqrt{x}\log^3 x\log\log x,$$

which χ_1 (mod m) is the most pretentious character for all conductors $\le Q$ (and the term is only included if $m|q$.

Proof. [Da] §20, (8) states that there exists a constant $c>0$ such that for all χ (mod q) with $q\le e^{c\sqrt{\log x}}$, except when $\chi=\chi_0$ and perhaps $\chi=\chi_1$ for some exceptional χ_1, we have $\psi(x,\chi)\ll x/e^{c\sqrt{\log x}}$. This implies, dealing with the prime powers dividing q as we did in the proof of Corollary 13.7a,

$$\sum_{q\le Q}\frac{1}{\varphi(q)}\sum_{\substack{\chi \pmod q\\ 1<\mathrm{cond}^*\chi\le R\\ \chi\ne\chi_1}}|\psi(x,\chi)|\ll\sum_{1<r\le R}\frac{1}{\varphi(r)}\sum_{\substack{\mathrm{cond}\chi=r\\ \chi\ne\chi_1}}|\psi(x,\chi)|\log Q+R\log x\log\log Q$$
$$\ll\frac{Rx\log Q}{e^{c\sqrt{\log x}}}.$$

Letting $R\le\sqrt{x}/Q$, and combining this with Corollary 13.7b, yields the result.

Using [Da], §19, (13) with $T = R^2$ we have that $\psi(x,\chi) \ll \frac{x}{R^2}\ \log^2 x$ provided $\chi \notin \Xi_R$, where Ξ_R is the set of characters χ with conductors $\le R$ for which $L(s,\chi)$ has no zeros ρ with $|\mathrm{Im}(\rho)| \le R^2$ and $\mathrm{Re}(\rho) > 1 - \frac{2\log R}{\log x}$. Now the remark to Theorem 14 of [Bo] implies that $|\Xi_R| \ll \exp\left(23\ \frac{\log^2 R}{\log x}\right)$. Hence by Corollary 13.7a we have that if $R \le x^{1/4}$ then

$$\sum_{q\le x^{1/2}/R} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ \chi\notin\Xi_R}} |\psi(x,\chi)| \ll \frac{x}{R}\log^3 x \log\log x,$$

provided that we can obtain an appropriate estimate for

$$\sum_{q\le Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ \mathrm{cond}^*\chi\ge R}} |G(x,\chi)|\,.$$

14. The Polya-Vinogradov Theorem

15. Burgess's Theorem

Proof of B's Theorem via additive combinatorics
Strange results using Halasz.
The number of chi for which the chaa=racter sum is large.

16. Subconvexity

17. Explicit sieve constructions

17.1. Long gaps between primes (Erdös-Rankin).

17.2. Short gaps between primes (Erdös-Rankin).

17.3. Daniel Shiu's Theorem.

18. Mean-values of multiplicative functions

This should include the decomposition theorem into small and large primes.
A study of differential delay equations, especially Buchstab's function.
A discussion of Spectra

19. Limitations to Equi-distribution

19.1. Maier's Theorem.

19.2. G-S generalization.

Ideas?

Compare $\zeta(1+it)^{-1}$ for $|t| \le T = x^c$ — can we replace RH by a hypothesis on this? ie to get an error term of $O(\sqrt{x})$.

References

Books

[Bo] E. Bombieri, *Le grand crible dans la théorie analytique des nombres*, Astérisque **18** (1987/1974), 103 pp.

[CP] R. Crandall and C. Pomerance, *Prime numbers: A computational perspective*, Springer Verlag, New York, 2001.

[Da] H. Davenport, *Multiplicative number theory*, Springer Verlag, New York, 1980.

[El] P.D.T.A. Elliott, *Duality in analytic number theory*, Cambridge, 1997.

[HR] H. Halberstam and H.-E. Richert,, *Sieve Methods*, Academic Press, London, New York, San Francisco, 1974.

[HW] G.H. Hardy and E.M. Wright, *Introduction to the theory of numbers*, Oxford, 1932.

[I1] A.E. Ingham, *The distribution of prime numbers*, Cambridge Math Library, Cambridge, 1932.

[IK] H. Iwaniec and E. Kowalski, Analytic Number Theory, AMS Colloquium Publications, vol. 53, American Mathematical Society, Providence, Rhode Island, 2004.

[M1] H.L. Montgomery, *Topics in multiplicative Number Theory*, Lecture notes in Math, Springer, 197?.

[M2] H.L. Montgomery, *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis*, C.B.M.S. Regional Conference Ser. Math, vol. 84, Amer. Math. Soc, 1994.

[MV] H. Montgomery and R.C. Vaughan, *Multiplicative number theory I: Classical theory*, Cambridge University Press, 2006.

[Nar] W. Narkiewicz, *The development of prime number theory*, Springer monographs in mathematics, Springer-Verlag, Berlin, 2000.

[Te] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge University Press, 1995.

[TM] G. Tenenbaum and M. Mendès France, *The prime numbers and their distribution*, American Mathematical Society, 2000.

[Ti] E.C. Titchmarsh, *The Theory of the Riemann Zeta-Function (2nd ed.)*, Oxford University Press, 1986.

Key Articles

[4] P.D.T.A. Elliott, *Multiplicative functions on arithmetic progressions*, Mathematika **34** (1987), 199-206.

[5] P.D.T.A. Elliott, *Multiplicative functions on arithmetic progressions II*, Mathematika **35** (1988), 38-50.

3. P.D.T.A. Elliott, *Extrapolating the mean-values of multiplicative functions*, Indag. Math **51** (1989), 409-420.

4. P.D.T.A. Elliott, *Some remarks about multiplicative functions of modulus* ≤ 1, Analytic number theory (Allerton Park, IL, 1989), Progr. Math, vol. 85, Birkhäuser Boston, Boston, MA, 1990, pp. 159-164.

E1. P.D.T.A. Elliott, *Multiplicative functions on arithmetic progressions. VII. Large moduli*, J. London Math. Soc. **66** (2002), 14–28.

E2. P.D.T.A. Elliott, *The least prime primitive root and Linnik's theorem*, Number theory for the millennium, I (Urbana, IL, 2000), A K Peters, Natick, MA, 2002, pp. 393–418.

G5. D. A. Goldston, J. Pintz, C. Y. Yildirim, *Primes in Tuples I*, Ann. of Math. **170** (2009), 819–862.

6. A. Granville and K. Soundararajan, *The spectrum of multiplicative functions*, Annals of Math **153** (2001), 407-470.

7. A. Granville and K. Soundararajan, *An upper bound for the unsieved integers up to* x (to appear).

16. G.H. Hardy and J.E. Littlewood, *Some problems of Parititio Numerorum (III): On the expression of a number as a sum of primes*, Acta Math. **44** (1922), 1–70.

8. G. Halász, *On the distribution of additive and mean-values of multiplicative functions*, Stud. Sci. Math. Hungar **6** (1971), 211-233.

9. G. Halász, *On the distribution of additive arithmetic functions*, Acta Arith. **27** (1975), 143-152.

I2. A.E. Ingham, *MR0029410/29411*, Mathematical Reviews **10** (1949), 595–596.

L1. U.V. Linnik, *On the least prime in an arithmetic progression. II. The Deuring-Heilbronn phenomenon*, Rec. Math. [Mat. Sb.] N.S. **15** (1944), 347-368.

M1. H. Maier, *Primes in short intervals*, Michigan Math. J. **32** (1985), 221–225.

17. H.L. Montgomery and R.C. Vaughan, *Exponential sums with multiplicative coefficients*, Invent. Math **43** (1977), 69-82.

18. A. Selberg, *An elementary proof of the prime number theorem for arithmetic progressions*, Can. J. Math **2** (1950), 66-78.