Math 539: Analytic Number Theory
Lecture Notes

Lior Silberman

ABSTRACT. These are rough notes for the Spring 2014 course. Problem sets and solutions were posted on an internal website.

Contents

Introduction (Lecture 1, 6/1/14)

Lior Silberman, `lior@Math.UBC.CA`, `http://www.math.ubc.ca/~lior`
Office: Math Building 229B
Phone: 604-827-3031

0.1. Administrivia

- Problem sets will be posted on the course website.
  - To the extent I have time, solutions may be posted on Connect.
  - I will do my best to mark regularly.
- Textbooks
  - Davenport [**3**]
  - Montgomery–Vaughn [**7**]
  - Iwaniec–Kowalski [**6**]

0.2. Course plan (subject to revision)

- Elementary counting ("change the order of summation")
- Exponential sums
- Counting primes, primes in arithmetic progressions
- Other topics if time permits.

0.3. Introduction

DEFINITION 1 (Caricature). Number Theory tries to find integer solutions to polynomial equations.

- Algebraic Number Theory: study individual solutions.
  - Solve $x^2+y^2=p$, and $x^2+y^2=n$ using prime factorization in the Gaussian integers.
  - Solve $x^3+y^3=z^3$ using prime factorization in the Eisenstein integers.
  - Solve $a^p+b^p=c^p$ using the Frey curve $y^2=x(x-a^p)(x-b^p)$.
- Analytic Number Theory: count the solutions.
  - (Gauss circle) What is the average number of ways to represent an integer at most x as a sum of two squares?
  - (Roth) Let A be a dense subset of $[n]$. Then A must have many solutions to $x+z=2y$.
  - Primes
    * (Mertens) $\sum_{p\le x}\frac{1}{p}=\log\log x+C+O(\frac{1}{\log x})$.
    * (Gauss; Riemann+dvP/Hadamard) $\sum_{p\le x}\log p=x+O\left(x\exp\left\{-\sqrt{\log x}\right\}\right)$, hence $\sum_{p\le x}1\sim\frac{x}{\log x}$.
    * (Twin primes conj) $\sum_{p\le x,p+2\text{ prime}}1\sim 2C_2\frac{x}{\log^2 x}$

– (Vinogradov) Let n be large enough and odd. Then the equation $p_1+p_2+p_3=n$ has about $\frac{n^2}{\log^3 n}$ solutions.
– (Green) Let A be a dense subset of the primes. Then A must have many solutions to $x+z=2y$.

THEOREM 2 (Helfgott 2013). *For all odd $N>10^{28}$ there is x for which*

$$\sum_{n_1+n_2+n_3=N}\prod_{i=1}^{3}\Lambda(n_i)\eta_i\left(\frac{n_i}{x}\right)>0,$$

where η_i are appropriate (positive) smooth functions.

COROLLARY 3. *(Adding numerics of Helfgott–Platt) Every odd integer $N>5$ is the sum of three primes.*

THEOREM 4 (Zhang 2013). *There is a weight function $\nu(n)>0$ and a finite set $\mathcal{H}$ of positive integers such that for all large enough x*

$$\sum_{\substack{x\le n\le 2x\\ n\equiv b\,(W(x))}}\left(\sum_{h\in\mathcal{H}}\theta\,(n+h)-\log 3x\right)\nu(n)>0,$$

where $W(x)$ is some slowly growing function of x and b is chosen appropriately.

COROLLARY 5. *For every x large enough there is $x\le n\le 2x$ and distinct $h_1,h_2\in\mathcal{H}$ such that $n+h_1,n+h_2$ are prime. In particular, there are arbitrarily large pairs of prime numbers whose difference is at most* $\max\mathcal{H}-\min\mathcal{H}$.

REMARK 6. Zhang obtained the bound $7\cdot 10^7$ for the gap $\max\mathcal{H}-\min\mathcal{H}$. Further work by Polymath8, Motohashi–Pintz and Maynard has reduced the gap to 272.

Math 539: Problem Set 0 (due 15/1/2013)

Real analysis

1. Some asymptotics
   (a) Let f,g be functions such that $f(x),g(x)>2$ for x large enough. Show that $f \ll g$ implies $\log f \ll \log g$. Give a counterexample under the weaker hypothesis $f(x),g(x)>1$.
   (b) For all $A>0$, $0<b<1$ and $\varepsilon>0$ show that for $x \geq 1$,
$$\log^A x \ll \exp\left(\log^b x\right) \ll x^{\varepsilon}.$$

2. Set $\log_1 x = \log x$ and for x large enough, $\log_{k+1} x = \log\left(\log_k x\right)$. Fix $\varepsilon>0$.
   (PRAC) Find the interval of definition of $\log_k x$. For the rest of the problem we suppose that $\log_k x$ is defined at N.
   (a) Show that $\sum_{n=N}^{\infty} \frac{1}{n\log n \log_2 n \cdots \log_{k-1} n (\log_k n)^{1+\varepsilon}}$ converges.
   (b) Show that $\sum_{n=N}^{\infty} \frac{1}{n\log n \log_2 n \cdots \log_{k-1} n (\log_k n)^{1-\varepsilon}}$ diverges.

3. (Stirling's formula)
   (a) Show that $\int_{k-1/2}^{k+1/2} \log t \, \mathrm{d}t - \log k = O(\frac{1}{k^2})$.
   (b) Show that there is a constant C such that
$$\log(n!) = \sum_{k=1}^{n} \log k = \left(n+\frac{1}{2}\right)\log n - n + C + O\left(\frac{1}{n}\right).$$
   RMK $C = \frac{1}{2}\log(2\pi)$, but this is largely irrelevant.

4. Let $\{a_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty} \subset \mathbb{C}$ be sequences with partial sums $A_n = \sum_{k=1}^{n} a_k$, $B_n = \sum_{k=1}^{n} b_k$.
   (a) (Abel summation formula) $\sum_{n=1}^{N} a_n b_n = A_N b_N - \sum_{n=1}^{N-1} A_n (b_{n+1} - b_n)$
   – (Summation by parts formula) Show that $\sum_{n=1}^{N} a_n B_n = A_N B_N - \sum_{n=1}^{N-1} A_n b_{n+1}$.
   (b) (Dirichlet's criterion) Suppose that $\{A_n\}_{n=1}^{\infty}$ are uniformly bounded and that $b_n \in \mathbb{R}_{>0}$ decrease monotonically to zero. Show that $\sum_{n=1}^{\infty} a_n b_n$ converges.

Supplementary problem: Review of Arithmetic functions

A. (a) The set of arithmetic functions with pointwise addition and Dirichlet convolution forms a commutative ring. The identity element is the function $\delta(n) = \begin{cases} 1 & n=1 \\ 0 & n>1 \end{cases}$.
   (b) f is invertible in this ring iff $f(1)$ is invertible in $\mathbb{C}$.
   (c) If f,g are multiplicative so is $f*g$.
   DEF $I(n)=1$, $N(n)=n$, $\varphi(n) = \left|(\mathbb{Z}/n\mathbb{Z})^{\times}\right|$, $\mu(n) = (-1)^r$ if n is a product of $r \geq 0$ distinct primes, $\mu(n)=0$ otherwise (i.e. if n is divisible by some p^2).
   (d) Show that $I*\mu = \delta$ by explicitly evaluating the convolution at $n=p^m$ and using (c).
   (e) Show that $\varphi * I = N$: (i) by explcitly evaluating the convolution at $n=p^m$ and using (c); (ii) by a combinatorial argument.
   (f) Show that pointwise multiplication by an arithmetic function $L(n)$ is a derivation in the ring iff $L(n)$ is completely additive: $L(de) = L(d) + (e)$ for all $d,e \geq 1$.

Supplementary problems: the Mellin transform and the Gamma function

For a function ϕ on $(0,\infty)$ its *Mellin transform* is given by $\mathcal{M}\phi(s) = \int_0^\infty \phi(x)x^s \frac{\mathrm{d}x}{x}$ whenver the integral converges absolutely.

B. Let ϕ be a bounded measurable function on $(0,\infty)$.

(a) Suppose that for some $\alpha > 0$ we have $\phi(x) = O(x^{-\alpha})$ as $x \to \infty$. Show that the $\mathcal{M}\phi$ defines a holomorphic function in the strip $0 < \Re(s) < \alpha$.

For the rest of the problem assume that $\phi(x) = O(x^{-\alpha})$ holds for all $\alpha > 0$.

(b) Suppose that ϕ is smooth in some interval $[0,b]$ (that is, there $b > 0$ and is a function $\psi \in C^\infty([0,b])$ such that $\psi(x) = \phi(x)$ with $0 < x \leq b$). Show that $\tilde{\phi}(s)$ extends to a meromorphic function in $\Re(s) < \alpha$, with at most simple poles at $-m$, $m \in \mathbb{Z}_{\geq 0}$ where the residues are $\frac{\phi^{(m)}(0)}{m!}$ (in particular, if this derivative vanishes there is no pole).

(c) Extend the result of (b) to ϕ such that $\phi(x) - \sum_{i=1}^r \frac{a_i}{x^i}$ is smooth in an interval $[0,b]$.

(d) Let $\Gamma(s) = \int_0^\infty e^{-t}t^s \frac{\mathrm{d}t}{t}$. Show that $\Gamma(s)$ extends to a meromorphic function in $\mathbb{C}$ with simple poles at $\mathbb{Z}_{\leq 0}$ where the residue at $-m$ is $\frac{(-1)^m}{m!}$.

C. (The Gamma function) Let $\Gamma(s) = \int_0^\infty e^{-t}t^s \frac{\mathrm{d}t}{t}$, defined initially for $\Re(s) > 0$.

FACT A standard integration by parts shows that $s\Gamma(s) = \Gamma(s+1)$ and hence $\Gamma(n) = (n-1)!$ for $n \in \mathbb{Z}_{\geq 1}$.

(a) Let $Q_N(s) = \int_0^N \left(1 - \frac{x}{N}\right)^N x^s \frac{\mathrm{d}x}{x}$. Show that $Q_N(s) = \frac{N!}{s(s+1)\cdots(s+N)}N^s$. Show that $0 \leq \left(1 - \frac{x}{N}\right)^N \leq e^{-x}$ holds for $0 \leq x \leq N$, and conclude that $\lim_{N\to\infty} \frac{N!}{s(s+1)\cdots(s+N)}N^s = \Gamma(s)$ for on $\Re s > 0$ (for a quantitative argument show instead $0 \leq e^{-x} - \left(1 - \frac{x}{N}\right)^N \leq \frac{x^2}{N}e^{-x}$)

(b) Define $f(s) = se^{\gamma s}\prod_{n=1}^\infty \left(1 + \frac{s}{n}\right)e^{-s/n}$ where $\gamma = \lim_{n\to\infty}\left(\sum_{i=1}^n \frac{1}{i} - \log n\right)$ is Euler's constant. Show that the product converges locally uniformly absolutely and hence defines an entire function in the complex plane, with zeros at $\mathbb{Z}_{\leq 0}$. Show that $f(s+1) = \frac{1}{s}f(s)$.

(c) Let $P_N(s) = se^{\gamma s}\prod_{n=1}^N \left(1 + \frac{s}{n}\right)e^{-s/n}$. Show that for $\alpha \in (0,\infty)$,$\lim_{N\to\infty} Q_N(\alpha)P_N(\alpha) = 1$ and conclude (without using problem B) that $\Gamma(s)$ extends to a meromorphic function in $\mathbb{C}$ with simple poles at $\mathbb{Z}_{\leq 0}$, that $\Gamma(s) \neq 0$ for all $s \in \mathbb{C} \setminus \mathbb{Z}_{\leq 0}$ and that the Weierstraß product representation

$$\Gamma(s) = \frac{e^{-\gamma s}}{s}\prod_{n=1}^\infty \left(1 + \frac{s}{n}\right)^{-1} e^{s/n}$$

holds.

(d) Let $F(s) = \frac{\Gamma'(s)}{\Gamma(s)}$ be the Digamma function. Using the Euler–Maclaurin summation formula $\sum_{n=0}^{n=N} f(n) = \int_0^N f(x)\,\mathrm{d}x + \frac{1}{2}(f(0) + f(N)) + \frac{1}{12}(f'(0) - f'(N)) + R$, with $|R| \leq \frac{1}{12}\int_0^N |f''(x)|\,\mathrm{d}x$, show that if $-\pi + \delta \leq \arg(s) \leq \pi + \delta$ and s is non-zero then

$$F(s) = \log s - \frac{1}{2s} + O_\delta\left(|s|^{-2}\right).$$

Integrating on an appropriate contour, obtain *Stirling's Approximation*: there is a constant c such that if $\arg(s)$ is as above then

$$\log\Gamma(s) = \left(s - \frac{1}{2}\right)\log s - s + c + O_\delta\left(\frac{1}{|s|}\right).$$

(e) Show *Euler's reflection formula*

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}\,.$$

Conclude that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ and hence that $\int_{-\infty}^{+\infty} e^{-\alpha x^2}\,\mathrm{d}x = \sqrt{\frac{\pi}{\alpha}}$.

(f) Setting $s = \frac{1}{2} + it$ in the reflection formula and letting $t \to \infty$, show that $c = \frac{1}{2}\log(2\pi)$ in Stirling's Approximation.

(g) Show *Legendre's duplication formula*

$$\Gamma(\frac{s}{2})\Gamma(\frac{s+1}{2}) = \sqrt{\pi}2^{1-s}\Gamma(s)\,.$$

CHAPTER 1

Elementary counting

1.1. Basic tools

1.1.1. Stirling's formula (PS0).

1.1.2. Abel Summation (PS0).

1.1.3. Arithmetic functions (PS0) (Lecture 2; 8/6/2014).

DEFINITION 7. An *arithmetic function* is a function $f\colon \mathbb{Z}_{>0} \to \mathbb{C}$.

EXAMPLE 8. $\delta(n) = \begin{cases} 1 & n=1 \\ 0 & n>1 \end{cases}$, $I(n) = 1$; $N(n) = n$. The *divisor function* $\tau(n) = \sum_{d|n} 1$ and sum-of-divisors function $\sigma(n) = \sum_{d|n} d$. The *Euler totient* $\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^{\times}$. For $n = \prod_{i=1}^{r} p_i^{e_i}$ set $\omega(n) = r$, $\Omega(n) = \sum_{i=1}^{r} e_i$ (so ω is additive, Ω completely additive), *Möbius function* $\mu(n) = \begin{cases} (-1)^{\omega(n)} & n \text{ squarefree} \\ 0 & n \text{ squarefull} \end{cases}$, *Liouville* function $\lambda(n) = (-1)^{\Omega(n)}$.

DEFINITION 9. The *Dirichlet convolution* (or *multiplicative convolution*) of f,g is the function

$$(f*g)(n) = \sum_{de=n} f(d)g(e)\,.$$

EXAMPLE 10. $\tau = I*I$, $\sigma = I*N$, $I*\mu = \delta$.

COROLLARY 11 (Möbius inversion formula). *If $F = G*I$ then $G = F*\mu$.*

LEMMA 12. *The set of arithmetic functions with pointwise addition and Dirichlet convolution forms a commutative ring with identity δ. f is invertible iff $f(1)$ is invertible in $\mathbb{C}$ (note $f \mapsto f(1)$ is ring hom to C).*

The Chinese Remainder Theorem says: if $(m,n) = 1$ then $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}) \simeq (\mathbb{Z}/nm\mathbb{Z})$ as rings. This forces some relations. For example, $\varphi(nm) = \varphi(n)\varphi(m)$, $\tau(nm) = \tau(n)\tau(m)$, $\sigma(nm) = \sigma(n)\sigma(m)$.

DEFINITION 13. Call f *multiplicative* if $f(nm) = f(n)f(m)$ if $(n,m) = 1$, *completely multiplicative* if $f(nm) = f(n)f(m)$ for all n,m.

LEMMA 14. *If f,g are multiplicative so is $f*g$. If $f(1) \neq 0$ then f is multiplicative iff f^{-1} is.*

EXAMPLE 15. I,N hence τ,σ,μ,λ .

Multiplicative f are determined by values at prime powers.

- To an arithmetic function associate the (formal) Dirichlet series $D_f(s) = \sum_{n\geq 1} f(n)n^{-s}$.
- Multiplication given by Dirichlet convolution – gives isomorphic link.

EXAMPLE 16. $\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_p (1-p^{-s})^{-1}$. Then $\zeta(s)^{-1} = \prod_p (1-p^{-s}) = \sum_n \mu(n) n^{-s}$ where $\mu(n) = \begin{cases} (-1)^{\omega(n)} & n \text{ squarefree} \\ 0 & \text{otherwise} \end{cases}$. In other words (Möbius inversion)

$$I * \mu = \mu * I = \delta .$$

EXERCISE 17. $\varphi * I = N$.

EXAMPLE 18. Formal differentiation gives $-\zeta'(s) = \sum_{n \geq 1} L(n) n^{-s}$ with $L(n) = \log n$. Multiplication by L (or any additive function) is a derivation in the ring. Formally differentaiting the Euler product also gives

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n \geq 1} \Lambda(n) n^{-s}$$

where

$$\Lambda(n) = \begin{cases} \log p & n = p^k \\ 0 & \text{otherwise} \end{cases}$$

is the *von Mangoldt function*. Note the identity above: $\zeta(s) \sum_{n \geq 1} \Lambda(n) n^{-s} = -\zeta'(s)$, that is

$$I * \Lambda = L .$$

1.2. Averages of arithmetic functions (Lecture 3, 10/1/2014)

- Goal: how big $f(n)$ is "on average".

1.2.1. Idea: convolutions are smoothing. Suppose $f = g * h$. Then

$$\begin{aligned} \sum_{n \leq x} f(n) &= \sum_{n \leq x} \sum_{d | n} g(d) h(\frac{n}{d}) \\ &= \sum_{d \leq x} g(d) \sum_{m \leq \frac{x}{d}} h(m) . \end{aligned}$$

Now if h is "smooth" then $\sum_{m \leq \frac{x}{d}} h(m)$ may be nice enough to evaluate.

EXAMPLE 19 (Elementary calculations). (1) The divisor function

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

Thus

$$\frac{1}{x} \sum_{n \leq x} \tau(n) = \log x + O(1) .$$

(2) The totient function.

$$\begin{aligned}
\sum_{n\le x}\varphi(n) &= \sum_{n\le x}\sum_{d|n}\mu(d)\frac{n}{d}=\sum_{d\le x}\mu(d)\sum_{d|n\le x}\frac{n}{d}\\
&= \sum_{d\le x}\mu(d)\sum_{m\le \frac{x}{d}} m=\sum_{d\le x}\mu(d)\left(\frac{x^2}{2d^2}+O(\frac{x}{d})\right)\\
&= x^2\sum_{d\le x}\frac{\mu(d)}{d^2}+O\left(x\sum_{d\le x}\frac{1}{d}\right)\\
&= x^2\left(\zeta^{-1}(2)-O(\frac{1}{x})\right)+O(x\log x)\\
&= \frac{x^2}{\zeta(2)}+O(x\log x)\,.
\end{aligned}$$

Thus

$$\frac{1}{x}\sum_{n\le x}\varphi(n)=\frac{x}{\zeta(2)}+O(\log x)\,.$$

(3) The normalized totient function

$$\begin{aligned}
\sum_{n\le x}\frac{\varphi(n)}{n} &= \sum_{d\le x}\frac{\mu(d)}{d}\left[\frac{x}{d}\right]\\
&= x\sum_{d\le x}\frac{\mu(d)}{d^2}+O(\log x)\\
&= \frac{x}{\zeta(2)}+O(\log x)\,.
\end{aligned}$$

Thus

$$\frac{1}{x}\sum_{n\le x}\frac{\varphi(n)}{n}=\frac{1}{\zeta(2)}+O\left(\frac{\log x}{x}\right)\,.$$

1.2.2. The Gauss Circle Problem.

DEFINITION 20. Let $r_k(n)=\#\left\{\underline{a}\in\mathbb{Z}^k\mid\sum_{i=1}^k a_i^2=n\right\}$ be the numebr of representations of n as a sum of k squares.

Then $\sum_{n\le x} r_k(n)=\#\left(\mathbb{Z}^k\cap B_{\mathbb{R}^k}(\sqrt{x})\right)$. Now tile the plane with units cubes centered at the lattice points of $\mathbb{Z}^k$ and let d be the diameter of the unit cube. Then

$$B_{\mathbb{R}^k}\left(\sqrt{x}-d\right)\subset\bigcup_{\underline{a}\in\mathbb{Z}^k\cap B_{\mathbb{R}^k}(\sqrt{x})}\left(\underline{a}+\left[-\frac{1}{2},\frac{1}{2}\right]^k\right)\subset B_{\mathbb{R}^k}\left(\sqrt{x}+d\right)\,.$$

Now let γ_k be the volume of the unit ball in k dimensions. Then $\mathrm{vol}\left(B_{R^k}(\sqrt{x}+O(1))\right)=\gamma_k\left(\sqrt{x}+O(1)\right)^k=\gamma_k x^{\frac{k}{2}}+O\left(x^{\frac{k-1}{2}}\right)$.

COROLLARY 21 (Gauss). *We have*

$$\#\left(\mathbb{Z}^k\cap B_{\mathbb{R}^k}(\sqrt{x})\right)=\gamma_k x^{\frac{k}{2}}+O\left(x^{\frac{k-1}{2}}\right)\,.$$

Note that the error term has a natural interpretation as the volume of the sphere.

Consider first the case $k=2$, where the size of the error term is known as the *Gauss Circle Problem.*

THEOREM 22 (Hardy 1915). *Write* $\#\left(\mathbb{Z}^2 \cap B_{\mathbb{R}^2}(\sqrt{x})\right) = \pi x + E(x)$. *Then* $E(x) \gg x^{1/4}\log^{1/4} x$ *infinitely often.*

CONJECTURE 23 (Hardy). $E(x) \ll_\varepsilon x^{\frac{1}{4}+\varepsilon}$.

We later give Voronoi's bound $E(x) \ll_\varepsilon x^{\frac{1}{3}+\varepsilon}$ (see section XX). The world record is

THEOREM 24 (Huxley 2003). $E(x) \ll_\varepsilon x^{\frac{131}{208}+\varepsilon}$.

When $k \geq 4$ the situation is easier, because $r_4(n)$ is a nicer function.

THEOREM 25 (Jacobi).

$$r_4(n) = 8(2+(-1)^n)\sum_{\substack{d|n \\ d\ odd}} d\,.$$

COROLLARY 26. $\sum_{n\leq x} r_4(n) = \frac{\pi^2}{2}x^2 + O(x\log x)$.

PROOF. By the usual method

$$\begin{aligned}
\sum_{n\leq x} r_4(n) &= \sum_{\substack{n=md\leq x \\ d \text{ odd}}} 8(2+(-1)^n)d \\
&= 8\sum_{m\leq x}(2+(-1)^m)\sum_{\substack{d\leq \frac{x}{m} \\ d \text{ odd}}} d \\
&= 8\sum_{m\leq x}(2+(-1)^m)\left(\frac{1}{2}\cdot\frac{1}{2}\cdot\left(\frac{x}{m}\right)^2 + O\left(\frac{x}{m}\right)\right) \\
&= 2x^2\sum_{m\leq x}\frac{2+(-1)^m}{m^2} + O\left(x\sum_{m\leq x}\frac{1}{m}\right) \\
&= 2x^2\left(\zeta(2)+\frac{1}{2}\zeta(2)\right) + O(x\log x) \\
&= 3\zeta(2)x^2 + O(x\log x) \\
&= \frac{\pi^2}{2}x^2 + O(x\log x)\,.
\end{aligned}$$

□

Note that Gauss's argument would have given the error term $O(x^{3/2})$.

EXERCISE 27. Improve for $k \geq 5$ the error term to $O(x^{\frac{k}{2}-1})$ using the result for $k=4$.

1.2.3. Dirichlet hyperbola method ("divisor switching"). The calculatution above of the average of $\tau(n)$ is inefficient, since the estimate $\left[\frac{x}{d}\right] = \frac{x}{d} + O(1)$ is bad for large d. However (Dirichlet) every $n \leq x$ has a divisor smaller than $\sqrt{x}$. Thus

$$\sum_{n\leq x} \tau(n) = 2\sum_{d\leq\sqrt{x}} \left[\frac{x}{d}\right] - \left[\sqrt{x}\right]^2$$

(error coming from cases where both divisors are $\leq x$, including square n). Thus

$$\begin{aligned}
\sum_{n\leq x} \tau(n) &= 2\sum_{d\leq\sqrt{x}} \frac{x}{d} - x + O(\sqrt{x}) \\
&= 2x\sum_{d\leq\sqrt{x}} \frac{1}{d} - x + O(\sqrt{x}) \\
&= x\left(2\log\sqrt{x} + 2\gamma + O\left(\frac{1}{\sqrt{x}}\right) - 1\right) + O(\sqrt{x}) \\
&= x\log x + (2\gamma - 1)x + O(\sqrt{x}).
\end{aligned}$$

We conlcude thatn

$$\frac{1}{x}\sum_{n\leq x} \tau(n) = \log x + (2\gamma - 1) + O(x^{-1/2}).$$

EXERCISE 28. Prove by the hyperbola method that $\frac{1}{x}\sum_{n\leq x}\tau_k(n) = P_k(\log x) + O\left(x^{1-\frac{1}{k}}\right)$ where P_k is a polynomial of degree k.

EXERCISE 29. Let $k \geq 4$. Writing $r_k(n) = \sum_{x_1,\ldots,x_{k-4}} r_4\left(n - \sum_{i=1}^{k} x_k^2\right)$ and changing the order of summation, show that

$$\sum_{n\leq x} r_k(n) = \frac{(\pi x)^{k/2}}{\Gamma\left(\frac{k}{2}+1\right)} + O\left(x^{\frac{k}{2}-1}\log x\right)$$

Note that the same formula with error term $O\left(x^{\frac{k-1}{2}}\right)$ follows from a volume argument as in the circle method.

1.3. Elementary prime estimates

1.3.1. Cramer's model (in practice discussed in Lecture 1). Let $A \subset [2,x]$ be chosen as follows: each $2 \leq n \leq x$ independently declares itself "prime" with probability $\frac{1}{\log n}$. Then

$$\mathbb{E}|A| = \sum_{n\leq x} \frac{1}{\log n} \approx \int_2^x \frac{\mathrm{d}t}{\log t}.$$

DEFINITION 30. $\mathrm{Li}(x) = \int_2^x \frac{\mathrm{d}t}{\log t}$.

CONJECTURE 31 (Gauss). $\pi(x) \stackrel{def}{=} |P \cap [0,x]| \sim \mathrm{Li}(x) \sim \frac{x}{\log x}$.

Similarly we find

$$\mathbb{E}\sum_{n\in A} \log n = \sum_{n\leq x} 1 \approx x. \tag{1.3.1}$$

$$\mathbb{E}\sum_{n\in A}\frac{1}{n}=\sum_{n\leq x}\frac{1}{n\log n}\approx\int_2^x\frac{\mathrm{d}t}{t\log t}=\log\log x+O(1)\,. \tag{1.3.2}$$

$$\mathbb{E}\sum_{n\leq x}A(n)A(n+2)\approx\sum_{n\leq x}\frac{1}{\log^2 n}\approx\int_2^x\frac{\mathrm{d}t}{\log^2 t}\sim\frac{x}{\log^2 x}\,. \tag{1.3.3}$$

While these look similar, for the true set of primes (1.3.2) is easy (we are about to prove it), (1.3.1) is hard (one of the highlights of the course) and (1.3.3) is open:

CONJECTURE 32 (Hardy–Littlewood twin primes conjecture). *$\sum_{n\leq x}P(n)P(n+2)\sim 2C_2\int_2^x\frac{\mathrm{d}t}{\log^2 t}$ where $C_2=\prod_p\frac{p(p-2)}{(p-1)^2}$.*

REMARK 33. Numerical estimates show our model to be somewhat off. The reason is that primality is not independent. For example, if n is prime then $n+1$ is not. A better model is to fix a small parameter z (say $z\approx C\log\log x$), take the primes up to z as known, and exclude from A any n divisible by a small prime.

CONJECTURE 34 (Generalized Hardy–Littlewood). *See Green–Tao.*

1.3.2. Chebychev's estimate (Lecture 4, 13/1/2014).

- Idea: dyadic decomposition

Let $n<p\leq 2n$. Then $p\|\binom{2n}{n}$ since p divides $(2n)!$ once and $n!$ not at all. Given x set $n=\left\lfloor\frac{x}{2}\right\rfloor$. Then

$$\begin{aligned}\sum_{\frac{x}{2}<p\leq x}\log p &\leq \sum_{n<p\leq 2n}\log p+\log x\\ &\leq \log\binom{2n}{n}+\log x\leq\log(4^n)+\log x\\ &\leq x\log 2+\log x\,.\end{aligned}$$

Setting $\theta(x)=\sum_{p\leq x}\log p$ we find

$$\theta(x)\leq\theta\left(\frac{x}{2}\right)+x\log 2+\log x$$

so

$$\begin{aligned}\theta(x) &\leq x\log 2\sum_{j=0}^{\log_2 x}\frac{1}{2^j}+\sum_{j=0}^{\log_2 x}\log x\\ &\leq (2\log 2)x+\log^2 x\\ &= O(x)\,.\end{aligned}$$

- Idea: there are very few prime powers

Now set $\psi(x)=\sum_{n\leq x}\Lambda(n)$. Then $\psi(x)=\theta(x)+\theta(x^{1/2})+\theta(x^{1/3})+\cdots=O(x+x^{1/2}+x^{1/3}+\cdots)=O(x)$ as well.

REMARK 35. Can also get a lower bound $\theta(x)\geq cx$ from this method, by noting that primes $\frac{2}{3}n<p<n$ don't divide $\binom{2n}{n}$ at all, and bounding the number of times primes $\sqrt{n}<p<\frac{2}{3}n$ can divide. Note that $\binom{2n}{n}\geq\frac{4^n}{2n+1}$ since it's the largest of $2n+1$ summands.

1.3.3. Mertens's formula (Lecture 4, continued). Note that

$$\sum_{d|n} \Lambda(d) = \sum_{p^j|n} \log p = \sum_{p^e \| n} e \log p = \log\left(\prod_{p^e\|n} p^e\right) = \log n\,.$$

Thus

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

Now $\sum_{d\le x} \Lambda(d) = \psi(d) = O(x)$ and

$$\begin{aligned}
\sum_{n\le x} \log n &= \int_1^x \log t \,\mathrm{d}t + O(\log x) \\
&= x\log x - x + O(\log x)\,.
\end{aligned}$$

Dividing by x we thus find

$$\sum_{d\le x} \frac{\Lambda(d)}{d} = \log x + O(1)\,.$$

Using the principle of "very few prime powers" it also follows that

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

We are now ready to prove

THEOREM 36 (Mertens). *There is a constant C such that $\sum_{p\le x} \frac{1}{p} = \log\log x + C + O(\frac{1}{\log x})$.*

PROOF. Let $S_n = \sum_{p \le n} \frac{\log p}{p}$. Then

$$\begin{aligned}
\sum_{p\le x} \frac{1}{p} &= \sum_{n\le x} \frac{1}{\log n}\left(S_n - S_{n-1}\right) \\
&= \sum_{n\le x} S_n \left(\frac{1}{\log n} - \frac{1}{\log(n+1)}\right) + C + O(\frac{1}{\log x}) \\
&= \sum_{n\le x} \left(1 - \frac{\log n}{\log(n+1)}\right) + O\left(\sum_{n\le x} \frac{1}{\log n} - \frac{1}{\log(n+1)}\right) + C + O(\frac{1}{\log x}) \\
&= \sum_{n\le x} \frac{\log(n+1) - \log n}{\log(n+1)} + C + O(\frac{1}{\log x}) \\
&= \sum_{n\le x} \frac{\frac{1}{n} + O(\frac{1}{n^2})}{\log(n+1)} + C + O(\frac{1}{\log x}) \\
&= \sum_{n\le x} \frac{1}{n\log n} + C + O(\frac{1}{\log x}) \\
&= \int_2^x \frac{\mathrm{d}t}{t\log t} + C + O(\frac{1}{\log x}) \\
&= \log\log x + C + O(\frac{1}{\log x}).
\end{aligned}$$

$\square$

REMARK 37. Can express as a Riemann–Stieltjes integral and integrate by parts instead.

1.3.4. The number of prime divisors (Lecture 5, 15/1/2014).

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

Thus

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

We now compute the standard deviation

$$\begin{aligned}
\frac{1}{x}\sum_{n\le x}(\omega(n)-\log\log x)^2 &= \frac{1}{x}\sum_{n\le x}(\omega(n))^2-\frac{2}{x}\sum_{n\le x}\omega(n)\log\log x+(\log\log x)^2\\
&= \frac{1}{x}\sum_{p_1,p_2\le x}\sum_{p_1,p_2|n\le x}1-2\log\log x\,(\log\log x+O(1))+(\log\log x)^2\\
&= \frac{1}{x}\left(\sum_{p\le x}\left[\frac{x}{p}\right]+\sum_{p_1\ne p_2\le x}\left[\frac{x}{p_1p_2}\right]-\sum_{p\le x}\left[\frac{x}{p^2}\right]\right)-(\log\log x)^2+O\left(\log\log x\right).\\
&\le \sum_{p\le x}\frac{1}{p}+\left(\sum_{p\le x}\frac{1}{p}\right)^2-(\log\log x)^2+O(\log\log x)\\
&= \log\log x+C+O(\frac{1}{\log x})+(\log\log x)^2+2C\log\log x+C^2+O\left(\frac{\log\log x}{\log x}\right)-(\log\mathrm{lo}\\
&= O(\log\log x).
\end{aligned}$$

(Theorem of Turan–Kubilius).

We can, in fact, determine the constant here. We have

$$\frac{1}{x}\sum_{n\le x}(\omega(n)-\log\log x)^2=\frac{1}{x}\left(\sum_{p\le x}\left[\frac{x}{p}\right]+\sum_{p_1\ne p_2\le x}\left[\frac{x}{p_1p_2}\right]-\sum_{p\le x}\left[\frac{x}{p^2}\right]\right)-\frac{2\log\log x}{x}\sum_{p\le x}\left[\frac{x}{p}\right]+(\log\log x)^2$$

Now $\frac{1}{x}\sum_{p\le x}\left[\frac{x}{p}\right]=\log\log x+O(1)$ (our main term),

$$\begin{aligned}
\frac{1}{x}\sum_{p_1\ne p_2}\left[\frac{x}{p_1p_2}\right] &= \frac{2}{x}\sum_{p_1\le\sqrt{x}}\sum_{p_2\le\frac{x}{p_1}}\left[\frac{x}{p_1p_2}\right]-\frac{1}{x}\sum_{p_1,p_2\le\sqrt{x}}\left[\frac{x}{p_1p_2}\right]\\
&= 2\sum_{p_1\le\sqrt{x}}\sum_{p_2\le\frac{x}{p_1}}\frac{1}{p_1p_2}+O(\frac{\sqrt{x}}{x})-\sum_{p_1,p_2\le\sqrt{x}}\frac{1}{p_1p_2}+O\left(\frac{\sqrt{x}^2}{x}\right)\\
&= 2\sum_{p_1\le\sqrt{x}}\frac{1}{p_1}\left(\log\log\frac{x}{p_1}+C+O(\frac{1}{\log(x/p_1)})\right)-\left(\log\log\sqrt{x}+C+O(\frac{1}{\log\sqrt{x}})\right)+O(1)\\
&= 2\sum_{p_1\le\sqrt{x}}\frac{1}{p_1}\left(\log\log x+\log\left(1-\frac{\log p_1}{\log x}\right)+C+O(\frac{1}{\log(x/p_1)})\right)
\end{aligned}$$

, $\frac{1}{x}\sum_{p\le x}\left[\frac{x}{p^2}\right]=O(1)$.

COROLLARY 38 (Hardy–Ramanujan). *Most $n\le x$ have about $\log\log n$ prime divisors.*

PROOF. By the triangle inequality in ℓ^2,

$$\left(\frac{1}{x}\sum_{n\le x}(\omega(n)-\log\log n)^2\right)^{1/2}\le\left(\frac{1}{x}\sum_{n\le x}(\omega(n)-\log\log x)^2\right)^{1/2}+\left(\frac{1}{x}\sum_{n\le x}(\log\log x-\log\log n)^2\right)^{1/2}.$$

Now for $n \geq \sqrt{x}$, $\log n \geq \frac{1}{2}\log x$ and $\log\log n \geq \log\log x - \log 2$. It follows that $\frac{1}{x}\sum_{n\leq x}(\log\log x - \log\log n)^2 \leq O(1) + \frac{(\log\log x)^2}{\sqrt{x}} = O(1)$ and, squaring, that

$$\frac{1}{x}\sum_{n\leq x}(\omega(n) - \log\log n)^2 = O(\log\log x)$$

as well. Now if $|\omega(n) - \log\log n| \geq (\log\log n)^{3/2}$ $\square$

THEOREM 39 (Erdős–Kac). *Fix a,b. Then*

$$\frac{1}{x}\#\left\{n \leq x \mid a \leq \frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} \leq b\right\} \xrightarrow[x\to\infty]{} \frac{1}{2\pi}\int_a^b e^{-t^2/2}\,\mathrm{d}t$$

Math 539: Problem Set 1 (due 29/1/2014)

1. (The standard divisor bound)
   (a) Let $f(n)$ be multiplicative, and suppose that $f \to 0$ along prime powers – that is, for every $\varepsilon > 0$ there is N such that if $p^m > N$ then $|f(p^m)| \le \varepsilon$. Show that $\lim_{n\to\infty} f(n) = 0$.
   (b) Show that for all $\varepsilon > 0$, $d(n) = O(n^\varepsilon)$.

2. Establish the following identities in the ring of formal Dirichlet series
   (a) Let $d_k(n) = \sum_{\prod_{i=1}^k a_i = n} 1$ be the generalized divisor functions, counting factorizations of n into k parts (so $d_2(n) = d(n)$ is the usual divisor function). Show $\sum_n d_k(n) n^{-s} = (\zeta(s))^k$ and that $d_k * d_l = d_{k+l}$.
   (b) Define $d_{1/2}(n)$. Calculate $d_{1/2}(p)$, $d_{1/2}(12)$.
   (c) Let $\sigma_\alpha(n) = \sum_{d|n} d^\alpha$ be the generalized sum-of-divisors function. Show that $\sum_{n=1}^\infty \sigma_\alpha(n) n^{-s} = \zeta(s)\zeta(s-\alpha)$.
   (d) Show that $\sum_{n\ge 1} d(n^2) n^{-s} = \frac{\zeta(s)^3}{\zeta(2s)}$.

3. Let $D_\varphi(s) = \sum_{n\ge1} \varphi(n) n^{-s}$.
   (a) Represent the series in terms of $\zeta(s)$ formally.
   (b) Show the series converges absolutely for $\Re(s) > 2$.
   (c) Show that the series failes to converge at $s = 2$.

4. Recall the function $\mathrm{Li}(x) = \int_2^x \frac{\mathrm{d}t}{\log t}$.
   (a) ("Asymptotic expansion") Show that for fixed K, $\mathrm{Li}(x) = \sum_{k=1}^K (k-1)! \frac{x}{\log^k x} + O_K\left(\frac{x}{\log^{K+1} x}\right)$.
   (b) (Asymptotic expansions are not series expansions) Show that $\sum_{k=1}^\infty (k-1)! \frac{x}{\log^k x}$ diverges.
   (c) Use summation by parts to estimate $\pi(x) = \sum_{p\le x} 1$ using the known asymptotics for $\sum_{p\le x} \frac{1}{p}$. Can you show $\pi(x) \ll \mathrm{Li}(x)$? $\pi(x) \gg \mathrm{Li}(x)$? That $\frac{\pi(x)}{\mathrm{Li}(x)} = 1 + o(1)$?

   NOTATION $f = \Theta(g)$ means $f = O(g)$ and $g = O(f)$, that is $0 \le \frac{1}{C} f(x) \le g(x) \le C f(x)$ for all large enough x.
   (d) Deduce $\pi(x) = \Theta(\mathrm{Li}(x))$ from Chebychev's estimate $\psi(x) = \Theta(x)$.

5. (Chebychev's lower bound) Let $\theta(x) = \sum_{p\le x} \log p$. We will find an explicit $c > 0$ such that $\theta(x) \ge cx$ for $x \ge 2$.
   (a) Let $v_p(n)$ denote the number of times p divides n. Show that $v_p(n!) = \sum_{k=1}^\infty \left\lfloor \frac{n}{p^k} \right\rfloor$.
   (b) Show that if $n < p \le 2n$ then $v_p\left(\binom{2n}{n}\right) = 1$.
   (c) (main saving) Show that if $\frac{2}{3}n < p \le n$ then $v_p\left(\binom{2n}{n}\right) = 0$ unless $n = p = 2$.
   (d) Show that if $\sqrt{2n} < p \le n$ then $v_p\left(\binom{2n}{n}\right) \le 1$.
   (e) Show that $v_p\left(\binom{2n}{n}\right) \le \log_p 2n$.
   (f) Show that $\log\binom{2n}{n} - (\theta(2n) - \theta(n)) \le \theta\left(\frac{2n}{3}\right) + 2\sqrt{2n}\log(2n)$.
   (g) Find a constant $c > 0$ such that $\theta(x) \ge cx$ for $2 \le x \le 4$ and such that if $\theta(x) \ge cx$ for all $2 \le x \le X$ then $\theta(x) \ge cx$ for $X < x \le 2X$.

6. Notation: $f(x) = o(g(x))$ ("little oh") if $\lim_{x\to\infty} \frac{|f(x)|}{g(x)} = 0$.
   (a) Show $\prod_p \left(1 - \frac{1}{p}\right) e^{\frac{1}{p}}$ converges.
   (b) Show that $\prod_{p\le z} \left(1 - \frac{1}{p}\right) = \frac{C(1+o(1))}{\log z}$.

7. Let $\sigma > 0$
   (a) Show that $\prod_{p\le x}(1 + p^{-\sigma}) \le \exp\left(O(x^{1-\sigma}/\log x)\right)$.
   (b) Let $a_p \in \mathbb{C}$ satisfy $\left|a_p\right| \le p^{-\sigma}$. Show that $f(n) = \prod_{p|n}(1+a_p) \le \exp\left(O\left((\log^{1-\sigma} n)(\log\log n)^{-1}\right)\right)$.
   (c) Show that $\sum_{n\le x} f(n) = cx + O(x^{1-\sigma})$ where $c = \prod_p \left(1 + \frac{a_p}{p}\right)$.

8. Let $\mathcal{A}_n$ denote a set of representative for the isomorphism classes of abelian groups of order n, $A_n = \#\mathcal{A}_n$ the number of isomorphism classes.
   (a) Show that $\sum_{n\ge 1} A_n n^{-s} = \prod_{k=1}^{\infty} \zeta(ks)$ in the ring of formal Dirichlet series.
   (b) Show that $\sum_{n\le x} A_n = cx + O\left(x^{1/2}\right)$ where $c = \prod_{k=2}^{\infty} \zeta(k)$.

9. Let $A \subset \mathbb{N}$. We define its lower and upper *natural densities* by
$$\underline{d} = \liminf_{N\to\infty} \frac{1}{N} \sum_{n\le N} A(n)\,, \qquad \bar{d} = \limsup_{N\to\infty} \frac{1}{N} \sum_{n\le N} A(n)\,.$$
   If the two are equal the limit exists and we call it the *natural density* of A. Similarly, the lower and upper *logarithmic densities* are
$$\underline{\delta} = \liminf_{N\to\infty} \frac{1}{\log N} \sum_{n\le N} \frac{A(n)}{n}\,, \qquad \bar{\delta} = \limsup_{N\to\infty} \frac{1}{\log N} \sum_{n\le N} \frac{A(n)}{n}\,,$$
   and a set has a *logarithmic density* if it lower and upper densities agree.
   (a) Show that $0 \le \underline{d} \le \underline{\delta} \le \bar{\delta} \le \bar{d} \le 1$ for all A. Conclude that if A has natural density it has logarithmic density and the two are equal.
   (b) Let A be the set of integers whose most significant digit is 4. Compute the lower and upper natural and logarithmic densities of A. Does it have natural density? Logarithmic density?

Hint for 4(a): Repeatedly integrate by parts, and for the error estimate effectuate the mantra "log is a constant function" by breaking the interval of integration in two.

Supplementary problems

A. Give the ring of formal Dirichlet series the *ultrametric* topology. In other words, say that $D_n(s) \to D(s)$ if each coefficient eventually stabilizes. This allows us to define products of Dirichlet series.

   (a) Let $G(T) = \sum_{n=0}^{\infty} a_n T^n \in \mathbb{C}[[T]]$ be a formal power series, and let f be an arithmetic function with $f(1) = 0$. Show that $G(D_f) = \sum_{n=0}^{\infty} a_n \left(D_f(s)\right)^n$ is a convergent series in the ring of formal Dirichlet series.

   (b) Let $D(s) = \sum_{n\geq 1} b_n n^{-s}$ be a formal Dirichlet series with $b_1 = 1$. Realize $\log D(s)$ as a formal Dirichlet series without constant term, and show that $\exp \log D(s) = D(s)$.

   (c) If f is multiplicative then $\sum_{n\geq 1} f(n)n^{-s} = \prod_p \left(\sum_{m=0}^{\infty} f(p^m)p^{-ms}\right)$, in the sense that the product on the right converges in the ultrametric topology.

   (d) (The complex topology is different) Give a multiplicative function f and a number s such that $\prod_p \left(1 + \left|\sum_{m=1}^{\infty} f(p^m)p^{-ms}\right|\right)$ converges but $\sum_{n\geq 1} f(n)n^{-s}$ does not.

B. For each $z \in \mathbb{C}$ define a multiplicative function $d_z(n)$ giving a natural generalization of d_k, and satisfying $d_z * d_w = d_{z+w}$. Evaluate $d_z(p)$, $d_z(p^k)$, $d_z(360)$.

[Note for the future: problem on sum-free sets]

CHAPTER 2

Fourier analysis

NOTATION 40. For $z \in \mathbb{C}$ set $e(z) = \exp(2\pi i z)$.

2.1. The Fourier transform on $\mathbb{Z}/N\mathbb{Z}$

2.1.1. Basics.

DEFINITION 41. For $f\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ set $\mathbb{E}_x f(x) = \frac{1}{N}\sum_{x \bmod N} f(x)$. Set $e_N(x) = e\left(\frac{x}{N}\right)$. Set $\psi_k(x) = e_N(kx)$. Note that N is implicit and that kx is well-defined mod N.

LEMMA 42. *$\{\psi_k\}_{k\in\mathbb{Z}/N\mathbb{Z}}$ is a complete orthonormal system in $L^2(\mathbb{Z}/N\mathbb{Z})$ (wrt the probability measure).*

COROLLARY 43 (Fourier analysis mod N). *Set $\hat{f}(k) = \langle \psi_k, f\rangle = \mathbb{E}_x \psi_{-k}(x) f(x)$. Then*

(1) (continuity in L^1) $\left\|\hat{f}\right\|_\infty \le \|f\|_1$.
(2) (Fourier inversion) $f(x) = \sum_{k(N)} \hat{f}(k)\psi_k(x)$.
(3) (Parseval formula) $\frac{1}{N}\sum_{x(N)} |f(x)|^2 = \sum_{k(N)} \left|\hat{f}(k)\right|^2$.
(4) (Expansion of δ distribution) $\frac{1}{N}\sum_{k(N)} \psi_{-k}(x)\psi_k(y) = \delta_{x,y}$. Equivalently,

$$\frac{1}{N}\sum_{k(N)} e_N\left(k(x-y)\right) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}.$$

DEFINITION 44. Let $f, g \in L^2(\mathbb{Z}/N\mathbb{Z})$ we define their *convolution* to be

$$\begin{aligned} (f*g)(x) &= \frac{1}{N}\sum_{a+b=x} f(a)g(b) \\ &= \mathbb{E}_y f(y) g(x-y). \end{aligned}$$

LEMMA 45. $\widehat{f*g}(k) = \hat{f}(k)\hat{g}(k)$.

PROOF. $\mathbb{E}_x (f*g)(x) e_{-k}(x) = \mathbb{E}_{x,y} f(y) g(x-y)\psi_{-k}(y)\psi_{-k}(x-y) = \mathbb{E}_{x,z} f(y) g(z)\psi_{-k}(y)\psi_{-k}(z)$. □

2.1.2. Application: Roth's Theorem.

PROBLEM 46. Let $A \subset \mathbb{Z}/N\mathbb{Z}$ be large enough. Must A contain a 3-AP, that is a solution to $x + z = 2y$?

Let $\alpha = \frac{\#A}{N} = \|A\|_1$ be the *density* of A. Here's an easy combinatorial argument.

LEMMA 47. *Suppose $\alpha > \frac{1}{2}$. Then A contains $\Theta(N^2)$ 3-APs.*

PROOF. For $x \in A$ consider the sets $\{d \mid x+d \in A\}$, $\{d \mid x-d \in A\}$ (basically shifts of A). Each has density $\alpha > \frac{1}{2}$ and hence their intersection has density$> 2\alpha - 1$. It follows that x is the middle element of $\Theta(N)$ 3-APs. $\square$

We count 3-APs using a Fourier expansion instead. Set

$$\Lambda_3(f_1, f_2, f_3) = \frac{1}{N^2} \sum_{x,d} f(x-d) f(x) f(x+d)\,,$$

so that $\Lambda_3(A,A,A)$ is the (normalized) number of 3-APs in A, including degenerate ones. Then

$$\begin{aligned}
\Lambda_3(f_1, f_2, f_3) &= \frac{1}{N^2} \sum_{x,d} \sum_{k_1,k_2,k_3} e_N\left(k_1(x-d) + k_2 x + k_3(x+d)\right) \hat{f}_1(k_1) \hat{f}_2(k_2) \hat{f}_3(k_3) \\
&= \frac{1}{N^2} \sum_{k_1,k_2,k_3} \hat{f}_1(k_1) \hat{f}_2(k_2) \hat{f}_3(k_3) \sum_{x,d} e_N\left((k_1+k_2+k_3)x + (k_3-k_1)d\right) \\
&= \frac{1}{N} \sum_{k_1,k_2} \hat{f}_1(k_1) \hat{f}_2(k_2) \hat{f}_3(k_1) \sum_{x} e_N\left((2k_1+k_2)x\right) \\
&= \sum_k \hat{f}_1(k) \hat{f}_2(-2k) \hat{f}_3(k)\,.
\end{aligned}$$

In particular, let $f_1 = f_2 = f_3 = A$ be the characteristic functions of A. We then have

$$\begin{aligned}
\Lambda_3(A,A,A) &= \sum_k \hat{A}(k)^2 \hat{A}(-2k) \\
&= \alpha^3 + \sum_{k \neq 0} \hat{A}(k)^2 \hat{A}(-2k)
\end{aligned}$$

Naural to let $f_A(x) = A(x) - \alpha$ be the *balanced function,* which has $\hat{f}_A(k) = \begin{cases} \hat{A}(k) & k \neq 0 \\ 0 & k = 0 \end{cases}$. Then

$$\Lambda_3(A,A,A) = \Lambda_3(\alpha,\alpha,\alpha) + \Lambda_3(f_A, f_A, f_A) = \alpha^3 + \Lambda_3(f_A, f_A, f_A)$$

since in each of the other 6 terms some argument has $\hat{f}$ supported away from zero, and some argument has $\hat{f}$ supported at zero. We conclude that:

$$\begin{aligned}
\left|\Lambda_3(A,A,A) - \alpha^3\right| &= \left|\Lambda_3(f_A, f_A, f_A)\right| \\
&\leq \sum_k \left|\hat{f}_A(k)\right|^2 \left\|\hat{f}_A\right\|_\infty \\
&= \left(\frac{1}{N} \sum_x (A(x) - \alpha)^2\right) \left\|\hat{f}_A\right\|_\infty \\
&= \left(\frac{1}{N} \sum_x \left(A(x) - 2\alpha A(x) + \alpha^2\right)\right) \left\|\hat{f}_A\right\|_\infty \\
&= \alpha(1-\alpha) \left\|\hat{f}_A\right\|_\infty\,.
\end{aligned}$$

COROLLARY 48 (Base case). *Suppose $\alpha > \frac{1}{2}$. Then A contains $\Theta_\alpha(N^2)$ 3-APs.*

PROOF. $\left\|\hat{f}_A\right\|_\infty \leq \|f_A\|_1 = \frac{1}{N}\left(\#A(1-\alpha)+(N-\#A)\alpha\right) = 2\alpha(1-\alpha)$. Thus

$$\begin{aligned}\frac{1}{N^2}\sum_{x,d} A(x)A(x+d)A(x-d) &\geq \alpha^3 - 2\alpha^2(1-\alpha)^2 \\ &\geq \alpha^3 - \frac{1}{8}.\end{aligned}$$

□

Idea: If $\hat{f}_A(k)$ is large for some k, then $\hat{f}_A$ correlates strongly with the function $e_N(kx)$, which is constant on relatively lengthy APs. This forces $\hat{f}_A$ to be relatively constant along such progressions, showing that the restriction of A to such a progression has somewhat larger density, at which point one can give an argument by induction.

THEOREM 49 (Roth 1953). *For all $\alpha > 0$ there is $N_0 = N_0(\alpha)$ such that if $N > N_0$ and $A \subset \{1,\ldots,N\}$ has density at least α then A has $\Theta_\alpha(N^2)$ 3-APs.*

PROOF. By downward induction on α ("density increment method"). Specifically, we show that for any $\alpha > 0$ if the theorem is true for $\alpha + \frac{\alpha^2}{10}$ is it true for α as well. Applying this to the infimum of the α for which the Theorem holds shows the infimum is 0.

Let $A \subset [N]$ have density α. In order to deal with "wraparound" issues embed A in $\mathbb{Z}/M\mathbb{Z}$ where $M = 2N+1$ and let

$$f_A(x) = \begin{cases} A(x)-\alpha & 0 \leq x < N \\ 0 & N \leq x < M. \end{cases}$$

and

$$1_N(x) = \begin{cases} 1 & 0 \leq x < N \\ 0 & N \leq x < M \end{cases}$$

so that $A = f_A + \alpha 1_A$ as functions in $\mathbb{Z}/M\mathbb{Z}$. Repeating the calculation above we find

$$\Lambda_3(A,A,A) = \alpha^3\Lambda_3(1_N,1_N,1_N) + \text{seven terms}.$$

Here, $\Lambda_3(1_N,1_N,1_N) = \sum$ can be computed exactly, and each of the other error terms has the form $\Lambda_3(f_1,f_2,f_3)$ where each f_i is either f_A or the balanced version of $\alpha 1_N$ (since $\hat{f}_A(0) = 0$). Now by C-S and Parseval,

$$|\Lambda_3(f_1,f_2,f_3)| = \left|\sum_{k(M)} \hat{f}_1(k)\hat{f}_2(-2k)\hat{f}_3(k)\right| \leq \left\|\hat{f}_i\right\|_\infty \left\|f_j\right\|_2 \|f_k\|_2$$

for any permutation (i,j,k) of $(1,2,3)$. Now $\|f_A\|_2 = \left(\frac{1}{M}\left(\alpha N(1-\alpha)^2 + (1-\alpha)N\alpha^2\right)\right)^{1/2} = \left(\frac{N}{2N+1}\right)^{1/2}(\alpha(1-\alpha))^{1/2}$ and

$$\left\|\alpha 1_N - \alpha\frac{N}{M}\right\|_2 = \alpha\left(\frac{1}{M}\left(N\left(\frac{N+1}{M}\right)^2 + (N+1)\left(\frac{N}{M}\right)^2\right)\right)^{1/2} = \alpha\frac{(N(N+1))^{1/2}}{M} \leq \frac{\alpha}{2}.$$

It follows that each of the seven terms is bounded above by one of $\left\|\hat{f}_A\right\|_\infty \frac{\alpha^2}{4}$ or $\left\|\hat{f}_A\right\|_\infty \frac{\alpha^{3/2}}{2\sqrt{2}}$ or $\left\|\hat{f}_A\right\| \frac{\alpha}{2}$. Setting

We now repeat the calculation above.

Setting $\varepsilon = \frac{\alpha^2}{2}$ we divide in two cases:

(1) ("quasi-randomness") If $\left\|\hat{f}_A\right\|_\infty \le \varepsilon$ then we have shown:

$$\Lambda_3(A,A,A) \ge \alpha^3 - \alpha(1-\alpha)\frac{\alpha^2}{2} > \frac{\alpha^3}{2}\,.$$

(2) ("structured case") Suppose $\left|\hat{A}(k)\right| \ge \varepsilon$ for some $k \ne 0$. We construct a longish AP $P \subset \mathbb{Z}/N\mathbb{Z}$ on which $A \cap P$ has larger density, and then apply the induction hypothesis to $A \cap P$, noting that any 3-AP in $A \cap P$ is an AP in A.

(a) There is $1 \le r \le \sqrt{N}$ such that kr has a representative of magnitude at most $\sqrt{N}$ (if not then there are $1 \le r_1 < r_2 \le \sqrt{N}$ such that kr_1, kr_2 have distance at most $\sqrt{N}$, and take $r = r_2 - r_1$).

(b) Let $L = \left\lceil \frac{\varepsilon\sqrt{N}}{4\pi} \right\rceil$,and let $P = r[L] = \{jr\}_{j=0}^{L-1}$. The point is that e_k is roughly constant on any progression $b+P$: since kr has a representative of magnitude at most $\sqrt{N}$,

$$|e_k(b+jr) - e_k(b)| = |e_k(jr) - 1| = 2\left|\sin\left(\pi\frac{kr}{N}j\right)\right| \le 2\pi\frac{1}{\sqrt{N}}L \le \frac{\varepsilon}{2} + \frac{2\pi}{\sqrt{N}}$$

(c) We now compute $\hat{f}_A$ by averaging over all translates of P:

$$\begin{aligned}
\varepsilon &\le \left|\hat{f}_A(k)\right| \\
&= \left|\frac{1}{N}\sum_b \frac{1}{L}\sum_{y\in P} f_A(b+y)e_{-k}(b+y)\right| \\
&\le \left|\frac{1}{N}\sum_b e_{-k}(b)\frac{1}{L}\sum_{y\in P} f_A(b+y)\right| + \frac{1}{N}\sum_b \frac{1}{L}\sum_{y\in P} |f_A(b+y)|\,|e_{-k}(b+y) - e_{-k}(b)| \\
&\le \left|\frac{1}{N}\sum_b e_{-k}(b)\frac{1}{L}\sum_{y\in P} f_A(b+y)\right| + \frac{\varepsilon}{2} + \frac{2\pi}{\sqrt{N}}\,,
\end{aligned}$$

that is

$$|\mathbb{E}_b e_{-k}(b)\mathbb{E}_{x\in b+P} f_A(x)| \ge \frac{\varepsilon}{2} - \frac{2\pi}{\sqrt{N}}\,.$$

(d) (Endgame) Let $e(-\theta)$ be the phase of the term in the paranthesis. Then we have found

$$\mathbb{E}_b e_{-k}(b)e(\theta)\mathbb{E}_{x\in b+P} f_A(x) \ge \frac{\varepsilon}{2} - \frac{2\pi}{\sqrt{N}}\,.$$

Since f_A averages to zero, this can also be written as

$$\mathbb{E}_b\left(e_{-k}(b)e(\theta) + 1\right)\mathbb{E}_{x\in b+P} f_A(x) \ge \frac{\varepsilon}{2} - \frac{2\pi}{\sqrt{N}}\,.$$

The real parts of $(e_{-k}(b)e(\theta)+1)$ are in $[0,2]$. Get b such that

$$\mathbb{E}_{x\in b+P} f_A(x) \ge \frac{\varepsilon}{4} - \frac{2\pi}{\sqrt{N}}\,.$$

Therefore, for N large enough, the restriction of A to $b+P$ has density at least $\alpha + \frac{\varepsilon}{5} = \alpha + \frac{\alpha^2}{10}$, and P itself is long.

$\square$

REMARK 50 (Corner cases). (1) Degenerate triples.
(2) The claim in $\mathbb{Z}$ and wraparound.

In fact, we have shown:

THEOREM 51. *Let $A \subset \{1, \dots, N\}$ have density $\gg \frac{1}{\log\log N}$. Then A has a 3-AP.*

The best result to date is

THEOREM 52 (Bourgain [**2**]). *Let $A \subset \{1, \dots, N\}$ have density $\gg \sqrt{\frac{\log\log N}{\log N}}$. Then A has a 3-AP.*

Compare also

THEOREM 53 (Sárközy, Furstenberg). *Let $A \subset \{1 \dots, N\}$ have density $\gg (\log\log N)^{-2/5}$. Then there are distinct $a, a' \in A$ such that $a - a'$ is a perfect square.*

2.1.3. Remarks: additive number theory.

- Szemeredi's Theorem and higher-order Fourier analysis.
- Corners Theorem
- Sum-product; Bourgain–Katz–Tao.

2.2. Dirichlet characters and the Fourier transform on $(\mathbb{Z}/N\mathbb{Z})^\times$

2.2.1. The Ramanujan sum.

DEFINITION 54. The *Ramanujan sum* is $c_N(k) = \sum'_{a(N)} e_N(ka)$, that is the fourier transform of the characteristic function of $(\mathbb{Z}/N\mathbb{Z})^\times$.

PROPOSITION 55. *$\sum_{d|n} c_d(k) = \begin{cases} n & n|k \\ 0 & n \nmid k \end{cases}$, so that $c_n(k) = \sum_{d|(k,n)} d\mu\left(\frac{n}{d}\right)$, and in particular $c_n(1) = \mu(n)$.*

PROOF. We sum:

$$\sum_{d|n} c_d(k) = \sum_{d|n} c_{n/d}(k) = \sum_{d|n} \sum_{\substack{a(n) \\ (a,n)=d}} e_n(ka) = \sum_{a(n)} e_n(ka) = \begin{cases} n & n|k \\ 0 & n \nmid k \end{cases}.$$

Now apply Möbius inversion. $\square$

COROLLARY 56. $\mathbb{1}_{(x,N)=1} = \frac{1}{N}\sum_{k(N)}\sum_{d|(k,N)} d\mu\left(\frac{N}{d}\right) e_N(kxxa) = \sum_{k(N)}\left(\sum_{d|(k,N)} \frac{d}{N}\mu\left(\frac{N}{d}\right)\right) e_N(kx)$

2.2.2. Basics, arithmetic progressions.

Construction. For each $a \in (\mathbb{Z}/N\mathbb{Z})^\times$, let $M_a \in U\left(L^2\left((\mathbb{Z}/N\mathbb{Z})^\times\right)\right)$ be multiplication by A. Clearly $M_a M_b = M_{ab}$ so this is a commuting family of unitary operators, hence jointly diagonalizable. Let $a \mapsto \chi(a)$ be an eigenvalue system. The multiplicative relation above gives $\chi(a)\chi(b) = \chi(ab)$ so $\chi\colon (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ is a group homomorphism

Every associated eigenvector f satisfies $f(a) = (M_a f)(1) = \chi(a) f(1)$, so the eigenspace is 1-dimensional and spanned by χ. Conversely, every $\chi \in \mathrm{Hom}\left((\mathbb{Z}/N\mathbb{Z})^\times, \mathbb{C}\right)$ lies in an eigenspace, and we see $\mathrm{Hom}\left((\mathbb{Z}/N\mathbb{Z})^\times, \mathbb{C}\right)$ is an orthonormal basis of $L^2\left((\mathbb{Z}/N\mathbb{Z})^\times\right)$ (prob measure).

DEFINITION 57. A *Dirichlet character* (of modulus N) a group homomorphism $(\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$, or equivalently its pullback to $\mathbb{Z}$: a multiplicative map $\chi\colon \mathbb{Z} \to \mathbb{C}$ such that $\chi(n) = 0$ iff $(n,N) \neq 1$.

Primitive characters. Note that if χ is a Dirichlet character mod N and $N|N'$ then we can obtain a Dirichlet character mod N' by setting $\chi'(n) = \begin{cases} \chi(n) & (n,N') = 1 \\ 0 & (n,N') = 1 \end{cases}$. If $N' > N$ we say that χ' is *imprimitive*. If χ is not imprimitive we say it is *primitive*. Given a Dirichlet character χ and $q \geq 1$ say q is a *period* of χ if whenever a,b are prime to q and $a \equiv b\,(d)$ we have $\chi(a) = \chi(b)$.

LEMMA 58 (The conductor). *Let $q(\chi)$ be the minimal period of χ.*

(1) $q(\chi)$ divides N.

(2) Let q be a period. Then there is a unique character χ' mod q which agrees with χ on n prime to Nq.

(3) The minimal period divides all periods.

PROOF. (1) If q is a period then so is $xq + yN$ for all x,y. (2) Let n be prime to q. For any j such that $n + jq$ is prime to N (for example, j can be the produt of the primes dividing N but not n) set $\chi'(n) = \chi(n + jq)$, noting that the RHS is independent of the choice of j since q is a period. This is clearly multiplicative and uniquely defined. (3) Follows from (1),(2). □

DEFINITION 59. We call q_0 the *conductor* of q and denote it $q(\chi)$.

Values. For fixed $a \in (\mathbb{Z}/N\mathbb{Z})^\times$ we will consider the possible values $\chi(a)$ as χ ranges over $\widehat{(\mathbb{Z}/N\mathbb{Z})}^\times$. For this let r be the multiplicative order of p mod N. Then for each χ, $\chi(a)$ must be a root of unity of order dividing r. The set $\{\chi(a)\}_{\chi \in \widehat{(\mathbb{Z}/N\mathbb{Z})}^\times}$ is a finite group of roots of unity, hence cyclic (a finite subgroup of a field), say of order $s|r$. It follows that $\chi(a^s) = 1$ for all $\chi \in \widehat{(\mathbb{Z}/N\mathbb{Z})}^\times$. Let δ_1 be the characteristic function of $1 \in (\mathbb{Z}/N\mathbb{Z})^\times$. Then $\frac{1}{\varphi(N)} \sum_\chi \chi(a) = \sum_\chi \langle \delta_1, \chi \rangle \chi(a) = \delta_1(a)$. In particular, if $\chi(a^s) = 1$ for all χ then $a^s = 1$ and hence $s = r$. It follows that the set of values $\{\chi(a)\}$ is exactly the set of roots of unity of order r. Finally, let χ be such that $\chi(a) = \zeta_r$ is a primitive root of unity of order r. Then multiplication by χ^j gives a bijection between $\{\chi \mid \chi(a) = \zeta_r^u\}$, $\left\{\chi \mid \chi(a) = \zeta_r^{u+j}\right\}$ so all these sets must have the same size. We have shown:

PROPOSITION 60 (Existence of characters). *Let $a \in (\mathbb{Z}/N\mathbb{Z})^\times$ have order r. Then for each root of unity $\zeta \in \mu_r$ there are $\frac{\varphi(N)}{r}$ Dirichlet characters χ mod N such that $\chi(a) = \zeta$.*

2.2.3. L-functions and Dirichlet's Theorem on primes in arithmetic progressions. We now reprise the analysis of Section **??**.

DEFINITION 61. For a Dirichlet character χ let $L(s;\chi)$ be the Dirichlet series $\sum_{n\geq 1} \chi(n) n^{-s}$.

LEMMA 62. *$L(s;\chi)$ converges absolutely in $\Re(s) > 1$, where it has the Euler product $L(s;\chi) = \prod_p (1 - \chi(p) p^{-s})^{-1}$.*

EXAMPLE 63. Let χ_0 be the principal character mod q. Then $L(s;\chi_0) = \left[\prod_{p|q} (1 - p^{-s})\right] \zeta(s)$. In particular, $L(s;\chi_0)$ continues to $\Re(s) > 0$ and has a pole at $s = 1$.

LEMMA 64. *Let χ be a non-principal character. Then $L(s;\chi)$ converges in $\Re(s)>0$.*

PROOF. We have $\sum_{a(qN)}\chi(a)=\sum'_{a(q)}\chi(a)=\varphi(q)\langle\chi_0,\chi\rangle=0$ so the series $\sum_n\chi(n)$ is bounded. For $\sigma>0$ $\{n^{-\sigma}\}_{n=1}^{\infty}$ converges monotonically to zero so by Dirichlet's criterion the series $\sum_{n\geq1}\chi(n)n^{-\sigma}$ converges. $\square$

PROPOSITION 65. *Let χ be non-principal. Then $L(1;\chi)\neq 0$.*

PROOF. Consider the Dedekind zetafunction of $K=\mathbb{Q}(\zeta_q)$,

$$\zeta_K(s)=\prod_{\chi}L(s;\chi)\,.$$

The Euler factor at $p\nmid N$ is

$$\prod_{\chi}\frac{1}{(1-\chi(p)p^{-s})^{-1}}\,.$$

Suppose that p has order r mod N. By Proposition 60, this proudct is exactly

$$\left[\prod_{\zeta\in\mu_r}\left(1-\zeta p^{-s}\right)\right]^{-\varphi(N)/r}=\left(1-p^{-rs}\right)^{-\varphi(N)/r}$$

since $\prod_{\zeta\in\mu_r}(1-\zeta X)=1-X^r$ (the two polynomials have degree r, agree at the $r+1$ points $\mu_r\cup\{0\}$). It follows that $\zeta_K(s)$ is a Dirichlet series with non-negative coefficients, and in particular that $\zeta_K(\sigma)\geq1$ for $\sigma>1$. Suppose $\chi\neq\bar{\chi}$. Then if $L(1;\chi)=0$ then also $L(1;\bar{\chi})=0$ and so the product of the two zeroes will cancel the pole of $L(s;\chi_0)$ at $s=1$, a contradiction

The real case requires more work, and we give three proofs.

(1) [**3**, pp. 33-34] Suppose χ is real and $L(1;\chi)=0$. Consider the auxiliary Dirichlet series $\psi(s)=\frac{L(s;\chi_0)L(s;\chi)}{L(2s;\chi_0)}$, which converges absolutely in $\Re(s)>1$, is meromorphic in $\Re(s)>0$, and is regular for $\Re(s)>\frac{1}{2}$ (the numerator is regular at $s=1$ and the denominator is non-vanishing in $\Re(s)>\frac{1}{2}$). Its Euler product (convergent in $\Re(s)>1$) is

$$\prod_{p\nmid q}\frac{\left(1-p^{-2s}\right)}{(1-p^{-s})(1-\chi(p)p^{-s})}=\prod_{p\nmid q}\frac{(1+p^{-s})}{(1-\chi(p)p^{-s})}=\prod_{\chi(p)=1}\frac{1+p^{-s}}{1-p^{-s}}\,.$$

In particular, $\psi(s)=\sum_{n\geq1}a_n n^{-s}$ for some positive coefficients a_n. Now consider its Taylor expansion about $s=2$, which has radius of convergence at least $\frac{3}{2}$. Differentiating m times we see $\psi^{(m)}(2)=(-1)^m\sum_{n\geq1}a_n(\log n)^m n^{-2}$ so there are $b_m\geq0$ such that

$$\psi(s)=\sum_{m\geq0}(-1)^m b_m(s-2)^m=\sum_{m\geq0}b_m(2-s)^m\,.$$

Now any $\frac{1}{2}<\sigma<2$ is in the domain of convergence and since $(2-\sigma)>0$ we have $\psi(\sigma)\geq b_0=\psi(2)>1$. But $\psi(\frac{1}{2})=0$ due to the pole of the denominator there.

(2) By Landau's Theorem, the domain of convergence ends with a singularity on the real axis. If $L(1;\chi)=0$ for some χ then this will cancel the simple pole of $L(s;\chi_0)$ there, so that

$\zeta_K(s)$ will be regular at $s = 1$. Since $\zeta(s), L(s;\chi)$ are regular on $(0,1)$ it would follow that $\zeta_K(s)$ converges in $\Re(s) > 0$. But for real $\sigma > 0$,

$$\zeta_K(s) = \prod_{p\nmid N} \left(1 - p^{-r\sigma}\right)^{-\varphi(N)/r} \geq \prod_{p\nmid N} \left(1 - p^{-\varphi(N)\sigma}\right)^{-1} = \sum_{(n,N)=1} n^{-\varphi(N)\sigma},$$

which diverges for $\sigma = \frac{1}{\varphi(N)}$ by comparison with the harmonic series.

(3) Replacing χ with its primitive counterpart changes only finitely many Euler factors in $L(s;\chi)$ and doesn't affect vanishing at $s = 1$. Now if $\chi^2 = 1$ then $\chi(n) = \chi_d(n) = \left(\frac{d}{n}\right)$ (Kronecker symbol) for some quadratic discriminant d, and we have

THEOREM 66 (Dirichlet's class number formula 1839; Conj. Jacobi 1832). *For $d < 0$, $L(1;\chi_d) = \frac{2\pi h(d)}{w|d|^{1/2}} > 0$. For $d > 0$, $L(1;\chi_d) = \frac{h(d)\log\varepsilon}{d^{1/2}} > 0$ where w is the number of roots of unity in $\mathbb{Q}(\sqrt{d})$ (usually $w = 2$), $h(d)$ is the number of equivalence classes of binary quadratic forms of discriminant d and ε is a fundamental unit of norm* 1.

□

THEOREM 67 (Dirichlet 1837 [**4**]). *Let $(a,N) = 1$. Then there are infinitely many primes p such that $p \equiv a\,(N)$.*

PROOF. For each character χ mod N and s with $\Re(s) > 1$ consider $\log L(s;\chi) = \sum_p \sum_{m=1}^{\infty} \chi(p)^m p^{-ms}$ (note that $|\chi(p)p^{-s}| < 1$ so we may use the Taylor expansion for $\log(1-\chi(p)p^{-s})$. Since $\sum_p \sum_{m\geq 2} p^{-m} \leq \sum_{n\geq 2}\sum_{m\geq 2} n^{-m} = \sum_{n\geq 2} \frac{1}{1-\frac{1}{n}} n^{-2} = \sum_{n\geq 2} \frac{1}{n(n-1)} = \frac{1}{2}$ we see that for $\Re(s) > 1$ we have

$$\log L(s;\chi) = \sum_p \chi(p)p^{-s} + O(1).$$

Now $\frac{1}{\varphi(N)}\sum_\chi \bar{\chi}(a)\chi(n) = \sum_\chi \langle\delta_a,\chi\rangle\chi = \delta_a$. Thus

$$\sum_{p\equiv a\,(N)} p^{-s} = \sum_p \delta_a(p)p^{-s} = \frac{1}{\varphi(N)}\sum_\chi \bar{\chi}(a)\log L(s;\chi) + O(1).$$

Now let $s \to 1^+$ through real values. For non-principal χ we have $\log L(s;\chi) \to \log L(1;\chi)$ which is finite by the Proposition. For the principal character, $\log L(s;\chi_0) \to \infty$ since $L(s;\chi_0) = \prod_{p|N}(1-p^{-s})\,\zeta(s)$. It follows that the RHS diverges as $s \to 1^+$. By the MCT we conclude that

$$\sum_{p\equiv a\,(N)} \frac{1}{p} = \infty.$$

In particular, there are infinitely many such primes. □

REMARK 68. In fact, our proof shows

$$\sum_{\substack{p\equiv a\,(N)\\ p\leq x}} p^{-1} = \frac{1}{\varphi(N)}\log\log x + O(1).$$

Moreover, it is natural to believe that the primes are evenly distributed between the residue classes. We will prove a quantitative version, but note the theory of "prime number races".

2.2.4. Additive transform of multiplicative characters: Gauss's sum. For future reference we evaluate

$$\hat{\chi}(k) = \sum_{a(q)} \chi(a) e(-\frac{ka}{q}) .$$

Math 539: Problem Set 2 (due 10/3/2014)

Dirichlet Characters

0. List all Dirichlet characters mod 15 and mod 16. Determine which are primitive.

1. Fix $q > 1$.
   (a) Let χ be a non-principal Dirichlet character mod q. Show that $\sum_p \frac{\chi(p)}{p}$ converges.
   (b) Let $(a,q) = 1$. Show that $\sum_{p \equiv a(q), p \leq x} \frac{1}{p} = \frac{1}{\varphi(q)} \log\log x + O(1)$
   (*c) Improve the error term to $C + O\left(\frac{1}{\log x}\right)$.

Dirichlet Series

2. (Convergence of Dirichlet series) Let $D(s) = \sum_{n \geq 1} a_n n^{-s}$ be a formal Dirichlet series. We will study the convergence of this series as s varies in $\mathbb{C}$.
   (a) Suppose that $D(s)$ converges absolutely at some $s_0 = \sigma_0 + it$. Show that $D(s)$ converges uniformly absolutely in the closed half-plane $\Re(s) = \sigma \geq \sigma_0$.
   (b) Conclude that there is an *abcissa of absolute convergence* $\sigma_{\mathrm{ac}} \in [-\infty, +\infty]$ such that one of the following holds: (1) ($\sigma_{\mathrm{ac}} = \infty$) $D(s)$ does not converge absolutely for any $s \in \mathbb{C}$; (2) ($\sigma_{\mathrm{ac}} \in (-\infty, +\infty)$)$D(s)$ converges absolutely exactly in the half-plane $\sigma > \sigma_{\mathrm{ac}}$ or $\sigma \geq \sigma_{\mathrm{ac}}$; (3) ($\sigma_{\mathrm{ac}} = -\infty$) $D(s)$ converges absolutely in $\mathbb{C}$. In cases (2),(3) the convergence is uniform in any half-plane whose closure is a proper subset of the domain of convergence.
   (c) Suppose that $D(s)$ converges at some s_0. Show that $D(s)$ converges in the open half-plane $\sigma > \sigma_0$, locally uniformly in every half-plane of the form $\sigma \geq \sigma_1 > \sigma_0$, and that $D(s)$ converges absolutely in the half-plane $\sigma > \sigma_0 + 1$.
   (d) Conclude that there is an *absicssa of convergence* $\sigma_{\mathrm{c}} \in [-\infty, \infty]$ such that on of the following holds: (1) ($\sigma_{\mathrm{c}} = \infty$) $D(s)$ does not converge for any $s \in \mathbb{C}$; (2) ($\sigma_{\mathrm{c}} \in (-\infty, +\infty)$)$D(s)$ converges in the open half-plane $\sigma > \sigma_{\mathrm{c}}$ and diverges in the open half-plane $\sigma < \sigma_{\mathrm{c}}$; the convergence is locally uniform in any half-plane $\sigma \geq \sigma_1 > \sigma_{\mathrm{c}}$ (3) ($\sigma_{\mathrm{ac}} = -\infty$) $D(s)$ converges absolutely in $\mathbb{C}$. In cases (2) the convergence is uniform in any half-plane. Furthermore, σ_{c} and σ_{ac} are either both $-\infty$, both $+\infty$, or both finite, and in the latter case $\sigma_{\mathrm{c}} \leq \sigma_{\mathrm{ac}} \leq \sigma_{\mathrm{c}} + 1$.

3. Let $D(s)$ have abcissa of absolute convergence σ_{ac}.
   (a) Suppose $\sigma_{\mathrm{ac}} \geq 0$. Show that $\sum_{n \leq x} |a_n| \ll_\varepsilon x^{\sigma_{\mathrm{ac}} + \varepsilon}$.
   (b) Suppose $\sigma_{\mathrm{ac}} < 1$. Show that $\sum_{n > x} |a_n| n^{-1} \ll_\varepsilon x^{\sigma_{\mathrm{ac}} + \varepsilon}$

4. (Convergence of sums and products) Let $D_1(s) = \sum_{n \geq 1} a_n n^{-s}$ and $D_2(s) = \sum_{n \geq 1} b_n n^{-s}$, and let $(D_1 + D_2)(s) = \sum_{n \geq 1} (a_n + b_n) n^{-s}$, $(D_1 \cdot D_2)(s) = \sum_{n \geq 1} c_n n^{-s}$ where $c = a * b$ is the Dirichlet convolution.
   (a) Show that the domain of absolute convergence of $D_1 + D_2$ and $D_1 D_2$ is at least the intersection of the domains of absolute convergence of D_1, D_2.
   (**b) (Mertens) Suppose that D_1, D_2 have abcissa of convergence σ_{c}. Show that $D_1 D_2$ has abcissa of convergence at most $\sigma_c + \frac{1}{2}$.

5. (Uniqueness of Dirichlet series) Suppose that $D(s)=\sum_{n\geq 1}a_n n^{-s}$ converges somewhere
   (a) Suppose that $a_n=0$ if $n<N$ and $a_N\neq 0$. Show that $\lim_{\Re(s)\to\infty}N^s D(s)=a_N$.
   (b) Suppose that $D_2(s)=\sum_{n\geq 1}b_n n^{-s}$ also converges somewhere, and that $D(s_k)=D_2(s_k)$ for $\{s_k\}$ in the common domain of convergence such that $\lim_{k\to\infty}\Re(s_k)=\infty$. Show that $a_n=b_n$ for all n.

6. (Landau's Theorem; proof due to K. Kedlaya) Let $D(s)=\sum_{n\geq 1}a_n n^{-s}$ have non-negative coefficients.
   (a) Show that $\sigma_{\mathrm{c}}=\sigma_{\mathrm{ac}}$ for this series.
   (b) Suppose that $D(s)$ extends to a holomorphic function in a small ball $|s-\sigma_{\mathrm{c}}|<\varepsilon$. Show that if $s<\sigma_c<\sigma$ and s,σ are close enough to σ_{c} then s is in the domain of convergence of the Taylor expansion of D at σ.
   (c) Using that $D^{(k)}(\sigma)=\sum_{n=1}^{\infty}a_n(-\log n)^k n^{-\sigma}$, write $D(s)$ as the sum of a two-variable series with positive terms.
   (d) Changing the order of summation, show that $D(s)$ converges at s, a contradiction to the definition of σ_{c}.
   (e) Obtain *Landau's Theorem*: if $D(s)$ has positive coefficients, has abcissa of convergence σ_{c}, and agrees with a holomorphic function in some punctured neighbourhood of σ_{c} then the singularity at $s=\sigma_{\mathrm{c}}$ is not removable.

Fourier Analysis

7. (Basics of Fourier series)
   (a) Let $D_N(x)=\sum_{|k|\leq N}e(kx)$ be the Dirichlet kernel. Show that $\int_0^1|D_N(x)|\,\mathrm{d}x\gg\log N$.
   (b) Let $F_N(x)=\sum_{|k|<N}\left(1-\frac{|k|}{N}\right)e(kx)$ be the Fejér kernel. Show that for $\delta\leq|x|\leq\frac{1}{2}$, we have $|F_N(x)|\leq\frac{1}{N\sin^2(\pi\delta)}$ so that for $f\in L^1(\mathbb{R}/\mathbb{Z})$,
   $$\lim_{N\to\infty}\int_{\delta\leq|x|\leq\frac{1}{2}}|f(x)|\,|F_N(x)|\,\mathrm{d}x=0\,.$$
   (c) In class we showed that "smoothness implies decay": if $f\in C^r(\mathbb{R}/\mathbb{Z})$ then for $k\neq 0$, $\left|\hat{f}(k)\right|\ll_r\|f\|_{C^r}|k|^{-r}$. Show the following partial converse: if $\left|\hat{f}(k)\right|=O(k^{-r-\varepsilon})$ then $\sum_{k\in\mathbb{Z}}\hat{f}(k)e(kx)\in C^{r-1}(\mathbb{R}/\mathbb{Z})$.

8. (The Basel problem) Let $f(x)$ be the $\mathbb{Z}$-periodic function on $\mathbb{R}$ such that $f(x)=x^2$for $|x|\leq\frac{1}{2}$.
   (a) Find $\hat{f}(k)$ for $k\in\mathbb{Z}$.
   (b) Show that $\zeta(2)=\frac{\pi^2}{6}$.
   (c) Apply *Parseval's identity* $\|f\|_{L^2(\mathbb{R}/\mathbb{Z})}=\left\|\hat{f}\right\|_{L^2(\mathbb{Z})}$ to evaluate $\zeta(4)$.

9. Let $\varphi\in\mathcal{S}(\mathbb{R})$.
   (a) Let $c\in L^2(\mathbb{Z}/q\mathbb{Z})$. Show that $\sum_{n\in\mathbb{Z}}c(n)\varphi(n)=\sum_{k\in\mathbb{Z}}\hat{c}(-k)\hat{\varphi}(k/q)$.
   (b) Let χbe a primitive Dirichlet character mod q. Show that
   $$\sum_{n\in\mathbb{Z}}\chi(n)\varphi(n)=\frac{G(\chi)}{q}\sum_{k\in\mathbb{Z}}\bar{\chi}(k)\hat{\varphi}\left(\frac{k}{q}\right).$$

Application: Weyl differencing and equidistribution on the circle

10. (Equidistribution) Let X be a compact space, μ a fixed probability measure on X (thought of as the "uniform" measure). We say that a sequence of probability measures $\{\mu_n\}_{n=1}^{\infty}$ is *equidistributed* if it converges to μ in the weak-* sense, that is if for every $f \in C(X)$, $\lim_{n\to\infty} \mu_n(f) = \mu(f)$ (equivalently, if for every open set $U \subset X$, $\mu_n(U) \to \mu(U)$).
    (a) Show that it is enough to check convergence on a set $B \subset C(X)$ such that $\mathrm{Span}_{\mathbb{C}}(B)$ is dense in $C(X)$.
    (b) (Weyl criterion) We will concentrate on the case $X = \mathbb{R}/\mathbb{Z}$, $\mu =$ Lebesgue. Show that in that case it is enough to check whether $\int_0^1 e(kx)\,\mathrm{d}\mu_n(x) \xrightarrow[n\to\infty]{} 0$ for each non-zero $k \in \mathbb{Z}$. (Hint: Stone–Weierstrass)
    DEF We say that a sequence $\{x_n\}_{n=1}^{\infty} \subset X$ is equidistributed (w.r.t. μ) if the sequence $\left\{\frac{1}{n}\sum_{k=1}^{n} \delta_{x_k}\right\}_{k=1}^{\infty}$ is equidistributed, that is if for every open set U the proportion of $1 \le k \le n$ such that $x_k \in U$ converges to $\mu(U)$, the proportion of the mass of X carried by μ.
    (c) Let α be irrational. Show directly that the sequence fractional parts $\{n\alpha \mod 1\}_{n=1}^{\infty}$ is dense in $[0,1]$.
    (d) Let α be irrational. Show that the sequence of fractional parts $\{n\alpha \mod 1\}_{n=1}^{\infty}$ is equidistributed in $[0,1]$.
    (e) Returning to the setting of parts (a),(b). suppose that $\mathrm{supp}(\mu) = X$. Show that every equidistributed sequence is dense.

2.3. The Fourier transform on $\mathbb{R}/\mathbb{Z}$ and the Poisson summation formula

2.3.1. Fourier series (Lecture 13, 3/2/2014).

2.3.1.1. *L^2 theory.*

- $\{e(kx)\}_{k\in\mathbb{Z}} \subset C^\infty(\mathbb{R}/\mathbb{Z}) \subset C(\mathbb{R}/\mathbb{Z}) \subset L^2(\mathbb{R}/\mathbb{Z})$ is a set of characters, hence an orthonormal system in $L^2(\mathbb{R}/\mathbb{Z})$ (prob measure). The unital algebra they span is closed under complex conjugation and separates the points, hence is dense in $C(\mathbb{R}/\mathbb{Z})$. This is dense in $L^2(\mathbb{R}/\mathbb{Z})$ so $\{e(kx)\}_{k\in\mathbb{Z}}$ is a complete orthonormal system. Set
$$\hat{f}(k) = \langle e(kx), f\rangle_{L^2(\mathbb{R}/\mathbb{Z})} = \int_{\mathbb{R}/\mathbb{Z}} f(x)e(-kx)\,\mathrm{d}x\,.$$
- Then for $f \in L^2$ we have $f = \sum_{k\in\mathbb{Z}} \hat{f}(k)e(kx)$ (congergence in L^2). This must converge almost everywhere, but at no specific point.
- We have *Parseval's identity* $\|f\|_{L^2}^2 = \sum_{k\in\mathbb{Z}} |\langle e_k, f\rangle|^2 = \|\hat{f}\|_{L^2(\mathbb{Z})}$.
- As usual we note $\|\hat{f}\|_{L^\infty(\mathbb{Z})} \le \|f\|_{L^1(\mathbb{R}/\mathbb{Z})}$.

We are interested in *pointwise* convergence of the Fourier expansion. We divide this in two parts.

2.3.1.2. *Smoothness $\Rightarrow$ decay.* Suppose $f \in C^1$. Then integrating by parts shows that for $k \neq 0$,
$$\hat{f}(k) = \frac{1}{2\pi ik}\hat{f}'(k)\,.$$
By induction, this means that for $k \neq 0$ and $d \geq 0$,
$$\left|\hat{f}(k)\right| \le \frac{\|f\|_{C^k}}{(2\pi)^d}|k|^{-d}\,.$$

COROLLARY 69. *For $f \in C^2$, the series $\sum_{k\in\mathbb{Z}} \hat{f}(k)e(kx)$ converges uniformly absolutely.*

2.3.1.3. *Convergence to f.*

DEFINITION 70. Set $(s_n f)(x) = \sum_{|k|\le n} \hat{f}(k)e(kx)$ and $(\sigma_N f)(x) = \frac{1}{N}\sum_{|n|<N}(s_n f)(x) = \sum_{|k|<N}\left(1-\frac{|k|}{N}\right)\hat{f}(k)e($

The second sum is *smoother*, so we expect it to be better behaved.

LEMMA 71. *$s_n f = D_n * f$, $\sigma_N f = F_N * f$ where*
$$D_n(x) = \sum_{|k|\le N} e(kx) = \frac{\sin\left(2\pi\left(N+\frac{1}{2}\right)x\right)}{\sin(\pi x)}$$
("Dirichlet kernel"*) and*
$$F_N(x) = \frac{1}{N}\sum_{n<N} D_n(x) = \frac{1}{N}\left(\frac{\sin(\pi Nx)}{\sin(\pi x)}\right)^2$$
("Fejér kernel"*).*

Both kernels satisfy $\int_{\mathbb{R}/\mathbb{Z}} D_n(x)\,\mathrm{d}x = \int_{\mathbb{R}/\mathbb{Z}} F_N(x)\,\mathrm{d}x = 1$. Moreover, $F_N(x) \ge 0$ for all x.

PROOF. Calculation. □

THEOREM 72 (Fejér). *Suppose $f \in L^1(\mathbb{R}/\mathbb{Z})$ is continuous at x. Then $\lim_{N\to\infty}(\sigma_N f)(x) = x$. In particular, if $f \in L^1(\mathbb{R}/\mathbb{Z})$ and $\lim_{n\to\infty} s_n f(x)$ exists, it equals $f(x)$.*

PROOF. We have

$$\begin{aligned}\sigma_N f(x) - f(x) &= \int_{\mathbb{R}/\mathbb{Z}} F_N(y) f(x+y)\,\mathrm{d}y - \int_{\mathbb{R}\mathbb{Z}} F_N(y) f(x)\,\mathrm{d}y \\ &= \int_{\mathbb{R}/\mathbb{Z}} F_N(y)\left(f(x+y) - f(x)\right)\mathrm{d}y\,.\end{aligned}$$

Given $\varepsilon > 0$ let $0 < \delta \leq \frac{1}{2}$ be such that $|f(x+y) - f(x)| \leq \varepsilon$ if $|y| \leq \delta$. Then

$$|\sigma_N f(x) - f(x)| \leq \varepsilon \int_{|y|\leq\delta} F_N(y)\,\mathrm{d}y + C_N \int_{\delta\leq|y|\leq\frac{1}{2}} \left(|f(x+y)| + |f(x)|\right)\mathrm{d}y$$

where $C_N(\delta) = \max\left\{F_N(y) \mid \delta \leq |y| \leq \frac{1}{2}\right\}$. Since $\int_{|y|\leq\delta} F_N(y)\,\mathrm{d}y \leq \int_{\mathbb{R}/\mathbb{Z}} F_N(y)\,\mathrm{d}y = 1$ we see that

$$|\sigma_N f(x) - f(x)| \leq \varepsilon + \left(\|f\|_{L^1} + (1-2\delta)\,|f(x)|\right) C_N(\delta)\,.$$

Since $C_N(\delta) = O_\delta(N^{-1})$ (PS2), the claim follows. □

REMARK 73. The Dirichlet kernel takes negative values. Since $\|D_n\|_{L^1} \gg \log n$, the proof would not have worked with it.

In fact, Fejér's theorem can be strengthned to

THEOREM 74 (Fejér). *Suppose $f \in L^1(\mathbb{R}/\mathbb{Z})$ has $\lim_{x\to x_0^\pm} f(x) = L_\pm$. Then $\lim_{N\to\infty} (\sigma_N f)(x) = \frac{L_+ + L_-}{2}$. In particular, if $s_n f(x)$ converges it converges to that limit.*

REMARK 75. Suppose f has a jump discontinuity at x_0 and is otherwise smooth. Then $s_N f(x) \to f(x)$ for all $x \neq 0$ (pointwise), but this convergence is not uniform: for fixed N, $s_N f(x)$ has a "spike" of height about $L_+ + c\,(L_+ - L_-)$ at a point $x_N = x_0 + \frac{1}{2N}$, an similarly $\lim_{N\to\infty} s_N f(x_0 - \frac{1}{2N}) = L_- + c\,(L_- - L_+)$.

2.3.2. The Poisson Summation formula (Lecture 14, 5/2/2014).

LEMMA 76. *Let $\varphi \in \mathcal{S}(\mathbb{R})$. Then $\Phi(x) = \sum_{n\in\mathbb{Z}} \varphi(x+n) \in C^\infty(\mathbb{R}/\mathbb{Z})$.*

By our Fourier inversion theorem, this means that

$$\Phi(x) = \sum_{k\in\mathbb{Z}} \hat{\Phi}(k) e(kx) \tag{2.3.1}$$

where

$$\begin{aligned}\hat{\Phi}(k) &= \int_{\mathbb{R}/\mathbb{Z}} \Phi(x) e(-kx)\,\mathrm{d}x = \int_0^1 \left(\sum_{n\in\mathbb{Z}} \varphi(n+x)\right) e(-kx)\,\mathrm{d}x \\ &= \sum_{n\in\mathbb{Z}} \int_n^{n+1} \varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}} \varphi(x) e(-kx)\,\mathrm{d}x\,.\end{aligned}$$

DEFINITION 77. For $f \in L^1(\mathbb{R})$ set $\hat{f}(k) = \int_{\mathbb{R}} f(x) e(-kx)\,\mathrm{d}x$.

PROPOSITION 78 (Poisson sum). *Let $\varphi \in \mathcal{S}(\mathbb{R})$. Then*

$$\sum_{n\in\mathbb{Z}} \varphi(n) = \sum_{k\in\mathbb{Z}} \hat{\varphi}(k)$$

PROOF. Set $x = 0$ in (2.3.1). □

2.4. Application: Pólya–Vinogradov

2.4.1. The meaning of "Smooth cutoff".

LEMMA 79 (Cutoff at). *Let* $\varphi \in \mathcal{S}(\mathbb{R})$*, and let* $X \geq 1$*. Then* $\sum_{n\in\mathbb{Z}} \left|\varphi\left(\frac{n}{X}\right)\right| = O_{\varphi}(X)$. *In particular, for any bounded* $f\colon \mathbb{Z} \to \mathbb{C}$ *we have*

$$\sum_{n\in\mathbb{Z}} f(n)\varphi\left(\frac{n}{X}\right) = O_{\varphi,\|f\|_\infty}(X)\,.$$

PROOF. Fix $T > 1$. Then there is $C = C(\varphi, T)$ such that for $|x| \geq 1$, $|\varphi(x)| \leq Cx^{-T}$ and hence

$$\begin{aligned}
\sum_{|n|>X} \left|\varphi\left(\frac{n}{X}\right)\right| &\leq 2C \sum_{|n|>X} \left(\frac{n}{X}\right)^{-T} \\
&\leq 2C\left(1 + \int_X^\infty \left(\frac{x}{X}\right)^{-T} \mathrm{d}x\right) \\
&= 2C\left(1 + X^T\left[\frac{x^{1-T}}{1-T}\right]_X^\infty\right) \\
&\leq 2C\left(X + \frac{X}{T-1}\right).
\end{aligned}$$

Also,

$$\sum_{|n|\leq X} \left|\varphi\left(\frac{n}{X}\right)\right| \leq (2X+1)\,\|\varphi\|_\infty = O_\varphi(X)\,.$$

□

We also need the "dual" version

LEMMA 80. *Let* $\varphi \in \mathcal{S}(\mathbb{R})$ *and let* $X \geq 1$*. Then for any* $T > 1$*,* $\sum_{|n|\geq 1} |\varphi(nX)| = O_{\varphi,T}\left(X^{-T}\right)$*. In particular, for any bounded* $f\colon \mathbb{Z} \to \mathbb{C}$ *we have*

$$\sum_{n\in\mathbb{Z}} f(n)\varphi\left(\frac{n}{X}\right) = f(0)\varphi(0) + O_{\|f\|_\infty,\varphi,T}\left(X^{-T}\right)\,.$$

PROOF. Let C be such that $|\varphi(x)| \leq Cx^{-T}$ for $|x| \geq 1$. Then

$$\sum_{|n|\geq 1} |\varphi(nX)| \leq 2C\sum_{n=1}^{\infty}(nX)^{-T} = \frac{2C\zeta(T)}{X^T}\,.$$

□

2.4.2. Smooth version, applications.

THEOREM 81 (Polya–Vinogradov). *Let* χ *be primitive mod* $q > 1$ *and let* $\varphi \in \mathcal{S}(\mathbb{R})$*. Then* $\sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n-M}{N}\right) \ll_\varphi \sqrt{q}$*.*

PROOF. Several stages.

(1) The sum is long: We have the trivial bound

$$\sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n-M}{N}\right) \leq \sum_{n\in\mathbb{Z}} \left|\varphi\left(\frac{n}{N}\right)\right| = O_\varphi(N)\,.$$

In particular, the claim is trivial unless $N \gg \sqrt{q}$, which we assume from now on.

(2) Gauss sum: Since χ is primitive we have $\chi(n) = \frac{\tau(\chi)}{q}\sum_{k(q)}\bar{\chi}(k)e_q(kn)$ so

$$\sum_{n\in\mathbb{Z}}\chi(n)\varphi\left(\frac{n-M}{N}\right) = \frac{\tau(\chi)}{q}\sum_{n\in\mathbb{Z}}\sum_{k(q)}\bar{\chi}(k)e_q(kn)\varphi\left(\frac{n-M}{N}\right).$$

(3) Poisson sum: Let $f(x) = \varphi\left(\frac{x-M}{N}\right)e\left(\frac{kx}{q}\right)$. Then $\hat{f}(\xi) = Ne\left(-M(\xi-\frac{k}{q})\right)\hat{\varphi}\left(N(\xi-\frac{k}{q})\right)$ and hence

$$\sum_{n\in\mathbb{Z}}e_q(kn)\varphi\left(\frac{n-M}{N}\right) = N\sum_{n\in\mathbb{Z}}e\left(-M(n-\frac{k}{q})\right)\hat{\varphi}\left(N(n-\frac{k}{q})\right)$$

and

$$\sum_{n\in\mathbb{Z}}\chi(n)\varphi\left(\frac{n-M}{N}\right) = \frac{N\tau(\chi)}{q}\sum_{|k|\le\frac{q}{2}}\bar{\chi}(k)\sum_{n\in\mathbb{Z}}e\left(-M(n-\frac{k}{q})\right)\hat{\varphi}\left(N(n-\frac{k}{q})\right).$$

(4) Rapid decay of $\hat{\varphi}$: This will shorten our dual sum. We choose k so that $\left|\frac{k}{q}\right| \le \frac{1}{2}$, at which point the proof of Lemma 80 still applies to the inner sum, so

$$= \frac{N\tau(\chi)}{q}\sum_{|k|\le\frac{q}{2}}\bar{\chi}(k)e\left(\frac{Mk}{q}\right)\hat{\varphi}\left(\frac{Nk}{q}\right) + \frac{N\tau(\chi)}{q}\cdot qO_{\varphi,T}\left(N^{-T}\right).$$

The remaining sum is certainly at most $\sum_{|k|\ge1}\left|\hat{\varphi}\left(\frac{Nk}{q}\right)\right|$.

(a) If $N \le q$ we apply Lemma 79 to get the bound

$$\sum_{n\in\mathbb{Z}}\chi(n)\varphi\left(\frac{n-M}{N}\right) = \frac{N\tau(\chi)}{q}O\left(\frac{q}{N}\right) + \tau(\chi)O(q^{-\frac{T-1}{2}}) = O(q^{1/2}).$$

(b) If $N \ge q$ we may apply Lemma 80 get for any T,

$$\begin{aligned}\sum_{n\in\mathbb{Z}}\chi(n)\varphi\left(\frac{n-M}{N}\right) &= \tau(\chi)O\left(\left(\frac{q}{N}\right)^T\right) + \tau(\chi)O(q^{-\frac{T-1}{2}}) \\ &= O(q^{1/2}).\end{aligned}$$

□

REMARK 82. Note that the precise choice of φ and the precise values for T are immaterial. This can be extended to non-primitive χ.

COROLLARY 83. *Let χ be primitive, of conductor $q > 1$. Let n be minimal such that $\chi(n) \ne 1$ (perhaps $\chi(n) = 0$. Then $n = O(q^{1/2})$.*

PROOF. Let φ be supported on $[-\varepsilon, 1+\varepsilon]$, valued in $[0,1]$, and satisfying $\varphi \equiv 1$ on $[0,1]$. Suppose that $\chi(n) = 1$ if $|n| \le N$ is prime to q. Then on the one hand

$$\sum_{n\in\mathbb{Z}}\chi(n)\varphi(\frac{n}{N}) = O(q^{1/2})$$

and on the other hand

$$\sum_{n\in\mathbb{Z}}\chi(n)\varphi\left(\frac{n}{N}\right) \ge N - 2\varepsilon N = (1-2\varepsilon)N.$$

It follows that

$$N = O(q^{1/2}).$$

$\square$

This can be improved, noting that if $\chi(n) = 1$ up to some y then the bias toward 1s continues much farther.

DEFINITION 84. Call $n \in \mathbb{Z}$ *y-smooth* if every prime divisor of n is at most y. Let $\psi(x;y)$ denote the number of y-smooth numbers up to x.

In those terms ,if $\chi(n) = 1$ for $n \le y$ then also $\chi(n) = 1$ for all y-smooth n, with the same φ as before,

$$\sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n}{x}\right) \ge \psi(x;y) - 2\varepsilon x - \sum_{y<p\le q} \frac{x}{p}.$$

Now suppose that $\sqrt{x} < y < x$. Then (since no integer up to x is divisible by *two* primes $> \sqrt{x}$) we have

$$\psi(x;y) = [x] - \sum_{y<p\le x} \left[\frac{x}{p}\right] \ge x - 1 - \sum_{y<p\le x} \frac{x}{p}.$$

It follows that if $\chi(n) = 1$ up to y, and if $y < x < y^2$ then for any $\varepsilon > 0$

$$\begin{aligned}
O_\varepsilon(q^{1/2}) &\ge (1-2\varepsilon)x - 1 - 2x \sum_{y<p\le x} \frac{1}{p} \\
&= (1-2\varepsilon)x - 1 - 2x\left[\log\log x + C + O\left(\frac{1}{\log x}\right) - \log\log y - C + O\left(\frac{1}{\log y}\right)\right] \\
&= x(1 - 2\varepsilon - 2\log\frac{\log x}{\log y}) + O\left(\frac{x}{\log x}\right).
\end{aligned}$$

Given $\delta > 0$ suppose $y = x^{\frac{1}{\sqrt{e}}(1+\delta)}$. Then we have $\frac{\log x}{\log y} = \frac{1}{\frac{1}{\sqrt{e}}(1+\delta)} = \sqrt{e}(1+\delta)^{-1}$ and hence

$$O\left(\frac{x}{\log x}\right) + 2\left(1 - 2\varepsilon - 1 + 2\log\left(1+\delta\right)\right)x = O(q^{1/2}).$$

Now given a small $\delta > 0$ choose $\varepsilon < \log(1+\delta)$. Then the LHS is $\Omega(x)$, and hence

$$x = O(q^{1/2})$$

and

$$y \ll_\varepsilon q^{\frac{1}{2\sqrt{e}}+\varepsilon}.$$

Theorem 81 is essentially best possible.

PROPOSITION 85. *There are N, M such that* $\sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n-M}{N}\right) \gg$

PROOF. Consider

$$\left|\sum_{M(q)} e_q(-M) \sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n-M}{N}\right)\right| \le q \max_M \left|\sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n-M}{N}\right)\right|.$$

Applying Poisson sum as before we have

$$\sum_{M(q)} e_q(-M) \sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n-M}{N}\right) = \frac{N\tau(\chi)}{q} \sum_{k(q)} \bar{\chi}(k) \sum_{n\in\mathbb{Z}} \hat{\varphi}\left(N(n-\frac{k}{q})\right) \sum_{M(q)} e_q(-M) e\left(-M(n-\frac{k}{q})\right).$$

Now since M is integral, $e(-Mn) = 0$ and $\sum_{M(q)} e_q(M(k-1)) = \begin{cases} q & k=1 \\ 0 & k\neq 1 \end{cases}$ so

$$= \frac{N\tau(\chi)}{q} q \sum_{n\in\mathbb{Z}} \hat{\varphi}\left(N(n-\frac{1}{q})\right).$$

Using Lemma 80 we see that

$$\sqrt{q}\frac{N}{q}\left(\left|\hat{\varphi}(\frac{N}{q})\right| + \text{small}\right) \leq \max_M \left|\sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n-M}{N}\right)\right|.$$

Now $\hat{\varphi}(x)$ is an analytic function, and in particular is non-vanishing on $[0,1]$. Letting $N = qx$ where $\hat{\varphi}(x) \neq 0$ gives

$$\max_M \left|\sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n-M}{N}\right)\right| \geq \sqrt{q}x\,|\hat{\varphi}(x)| - \text{small}.$$

Taking $N = [qx]$ will be find since $\varphi'(x)$ is bounded. □

2.4.3. Sharp cutoff. For the following, see [**3**] , [**7**, §9.4], or [**6**, Ch. 12] which gives a good constant also covers the smooth case.

(1) For χ primitive, $\left|\sum_{n=M+1}^{n=M+N} \chi(n)\right| \leq q\log q$, and for χ non-principal $\left|\sum_{n=M+1}^{n=M+N} \chi(n)\right| \leq 6\sqrt{q}\log q$.
   (a) On GRH (Montgomery–Vaughan Inv Math 43; simpler proof by Granville–Soundararajan JAMS 20 2007) $\left|\sum_{n=M+1}^{n=M+N} \chi(n)\right| \ll \sqrt{q}\log\log q$.
   (b) For all q there are N,M such that $\left|\sum_{n=M+1}^{n=M+N} \chi(n)\right| \geq \frac{\sqrt{q}}{\pi}$.
   (c) (Paley) There is $c>0$ such that for infinitely many quadratic discriminants d,

$$\max_{M,N} \left|\sum_{n=M+1}^{n=M+N} \chi_d(n)\right| > c\sqrt{d}\log\log d.$$

   (d) Oor bounds are trivial for $N \ll \sqrt{q}$. It is believed that $\left|\sum_{n=M+1}^{n=M+N} \chi(n)\right| \ll_\varepsilon N^{\frac{1}{2}} q^\varepsilon$.

2.4.4. Connection to Dirichlet L-functions [see Goldmakher's Thesis]. Set $S_\chi(t) = \sum_{n\in\mathbb{Z}} \chi(n)\varphi\left(\frac{n}{t}\right)$ (with $\varphi \in \mathcal{S}(\mathbb{R})$ having the same perity as χ) Then for $\Re(s) > 1$,

$$\begin{aligned} \int_0^\infty S_\chi(t) t^{-s}\frac{\mathrm{d}t}{t} &= 2\sum_{n=1}^\infty \chi(n) \int_0^\infty \varphi\left(\frac{n}{t}\right) t^{-s}\frac{\mathrm{d}t}{t} \\ &= \left(2\int_0^\infty \varphi(t)t^s\frac{dt}{t}\right)\left(\sum_{n=1}^\infty \chi(n)n^{-s}\right), \end{aligned}$$

and the manipulation is justified by the absolute convergence. Now $\int_0^\infty \varphi(t)t^s\frac{dt}{t}$ is holomorphic for $\Re(s) > 0$. Our bound $S_\chi(t) \ll_\varphi \sqrt{q}$ shows that the LHS converges absolutely for $\Re(s) > 0$, and on the RHS the same is true for the Mellin transform $\left(2\int_0^\infty \varphi(t)t^s\frac{dt}{t}\right)$. It follows that $L(s;\chi)$ extends meromorphically to $\Re(s) > 0$. In fact, the extension is holomorphic, since by varying φ we can ensure the denominator in the following expression is non-vanishing at any specific points

$$L(s;\chi) = \frac{\int_0^\infty S_\chi(t) t^{-s} \frac{\mathrm{d}t}{t}}{2\int_0^\infty \varphi(t) t^s \frac{dt}{t}}\,.$$

EXAMPLE 86. $L(1;\chi) \ll \log q$.

PROOF. $\int_0^\infty S_\chi(t) t^{-1} \frac{\mathrm{d}t}{t} = \int_0^q S_\chi(t) t^{-2}\,\mathrm{d}t + \int_q^\infty S_\chi(t) t^{-2}\,\mathrm{d}t \ll \int_0^q t^{-1}\,\mathrm{d}t + \sqrt{q}\int_q^\infty t^{-2}\,\mathrm{d}t = O(\log q + 1)$. □

EXERCISE 87 (Convexity bound). $L(\frac{1}{2};\chi) \ll q^{1/4}$.

CONJECTURE 88 (ELH). $L\left(\frac{1}{2};\chi\right) \ll_\varepsilon q^\varepsilon$.

THEOREM 89 (Burgess). $L\left(\frac{1}{2};\chi\right) \ll_\varepsilon q^{\frac{1}{4}-\frac{1}{16}+\varepsilon}$.

2.5. The Fourier transform on $\mathbb{R}^n$

LEMMA 90. *Assuming all integrals converge,*

$$\varphi\widehat{(ax+b)}(k) = \int_{\mathbb{R}} \varphi(ax+b) e(-kx)\,\mathrm{d}x = \frac{e\left(\frac{kb}{a}\right)}{a}\hat{\varphi}\left(\frac{k}{a}\right),$$

integration by parts shows (smoothness therefore decay)

$$\hat{\varphi}(k) = \frac{1}{(2\pi i k)^d}\widehat{\varphi^{(d)}}(k)\,.$$

and differentiation under the integral sign gives (decay therefore smoothness)

$$\hat{\varphi}^{(r)}(k) = (-2\pi i)^r \widehat{x^r\varphi(x)}\,.$$

COROLLARY 91. *Let $\varphi \in \mathcal{S}(\mathbb{R})$. Then $\hat{\varphi} \in \mathcal{S}(\mathbb{R})$.*

THEOREM 92 (Fourier inversion formula). *Let $\varphi \in \mathcal{S}(\mathbb{R})$.*

$$\varphi(x) = \int_{\mathbb{R}} \hat{\varphi}(k) e(kx)\,\mathrm{d}k\,.$$

PROOF. We have

$$\sum_{n\in\mathbb{Z}} \varphi(Tn+x) = \frac{1}{T}\sum_{k\in\mathbb{Z}} \hat{\varphi}\left(\frac{k}{T}\right) e\left(\frac{k}{T}x\right).$$

Letting $T \to \infty$, the LHS converges to $\varphi(x)$, the RHS to $\int_{\mathbb{R}} \hat{\varphi}(k) e(kx)\,\mathrm{d}k$. □

LEMMA 93 (Fourier inversion). *Let $f \in L^1(\mathbb{R})$ and suppose that $S = \int_{-\infty}^{+\infty} \hat{f}(t) e^{2\pi i t x}\,\mathrm{d}x$ converges as a symmetric improper integral for some t. Suppose that f is continuous at x. Then $S = f(x)$.*

PROOF. Let $\varphi(u) \in C_c^\infty(\mathbb{R})$ be odd. Setting $S_u(x) = \int_{-u}^{u} \hat{f}(t) e^{2\pi i t x}\,\mathrm{d}t = (D_u * f)(x)$, we consider the average

$$\frac{1}{T}\int_0^\infty \varphi\left(\frac{u}{T}\right) S_u(x)\,\mathrm{d}u = \int_0^\infty \varphi(u) S_{Tu}(x)\,\mathrm{d}u \xrightarrow[T\to\infty]{} \left(\int_0^\infty \varphi(u)\,\mathrm{d}u\right) S\,.$$

On the other hand,

$$\frac{1}{T}\int_0^\infty \varphi(u/T)S_u(x)\,\mathrm{d}u = \frac{1}{T}\int_{u=0}^{u=\infty}\mathrm{d}u\varphi(u/T)\int_{t=-u}^{t=u}\mathrm{d}te^{2\pi itx}\int_{y=-\infty}^{y=+\infty}\mathrm{d}ye^{-2\pi ity}f(y)\,\mathrm{d}y$$

converges absolutely so we may change the order of integration and obtain

$$\begin{aligned}
\frac{1}{T}\int_0^T \varphi(u/T)S_u(x)\,\mathrm{d}u &= \frac{1}{T}\int_{u=0}^{u=\infty}\int_{y=-\infty}^{y=+\infty}\mathrm{d}u\,\mathrm{d}y\varphi(u/T)f(y)\int_{t=-u}^{t=u}\mathrm{d}te^{2\pi it(x-y)} \\
&= \frac{1}{T}\int_{y=-\infty}^{y=+\infty}\mathrm{d}yf(y)\int_{u=0}^{u=\infty}\frac{e^{2\pi i(x-y)u}-e^{-2\pi i(x-y)u}}{2\pi i(x-y)}\,\mathrm{d}u\varphi(u/T) \\
&= -\frac{1}{2\pi i}\int_{y=-\infty}^{y=+\infty}\mathrm{d}yf(y)\frac{\hat{\varphi}\left(T(x-y)\right)}{x-y} \\
&= -\frac{1}{2\pi i}\int_{\mathbb{R}}\mathrm{d}yf(x+y)\frac{\hat{\varphi}(Ty)}{y}\,.
\end{aligned}$$

Now $\frac{\hat{\varphi}(y)}{2\pi iy}$ is the Fourier transform of $\int_{-\infty}^u \varphi(t)\,\mathrm{d}t$. In particular, $\int_{\mathbb{R}}\frac{\hat{\varphi}(Ty)}{2\pi iy}\,\mathrm{d}y = \int_{\mathbb{R}}\frac{\hat{\varphi}(y)}{2\pi iy}\,\mathrm{d}y = \int_{-\infty}^0\varphi(u)\,\mathrm{d}u = -\int_0^\infty\varphi(u)\,\mathrm{d}u$. It follows that

$$\left(\int_0^\infty \varphi(u)\,\mathrm{d}u\right)S - \left(\int_0^\infty \varphi(u)\,\mathrm{d}u\right)f(x) = \frac{1}{2\pi i}\int_{\mathbb{R}}\mathrm{d}y\left[f(x)-f(x+y)\right]\frac{\hat{\varphi}(Ty)}{y}\,.$$

Choosing δ such that $|f(x+y)-f(x)|\leq\varepsilon$ for $|y|\leq\delta$ and setting $C_T = \sup\{|\hat{\varphi}(y)| \mid |y|\geq T\}$ we get:

$$\left|\int_0^\infty \varphi(u)\,\mathrm{d}u\right||S-f(x)| \leq \frac{\varepsilon}{2\pi}\int_{\mathbb{R}}\left|\frac{\hat{\varphi}(Ty)}{y}\right|\mathrm{d}y + \frac{1}{2\pi}|f(x)|\int_{|y|\geq T\delta}\left|\frac{\hat{\varphi}(y)}{y}\right|\mathrm{d}y + \frac{1}{2\pi\delta}C_{T\delta}\left\|f\right\|_{L^1}$$

and, taking $T\to\infty$ and using Riemann–Lebesuge we get

$$\left|\int_0^\infty \varphi(u)\,\mathrm{d}u\right||S-f(x)| \leq \frac{\varepsilon}{2\pi}\left\|\frac{\hat{\varphi}(y)}{y}\right\|_{L^1(\mathbb{R})}\,.$$

□

COROLLARY 94 (Fourier inversion formula). *Suppose that $f,\hat{f}\in L^1(\mathbb{R})$ and in addition $f\in C(\mathbb{R})$. Then $f(x)=\int_{\mathbb{R}}\hat{f}(k)e(kx)$ for all x.*

2.5.1. Application: (Weakened) Voronoi bound for the Gauss circle problem. Let $\varphi\in C_c^\infty(\mathbb{R}^2)$ be smooth, non-negative, radial, integrating to one, supported in the unit disc. Let $f = \varphi\left(\frac{\cdot}{Y}\right)*\mathbb{1}_{B(X)}$. Then for $|x|>X+Y$, $f(x)=0$ and for $|x|<X-Y$, $f(x)=1$. It follows that

$$\left|\sum_{n\in\mathbb{Z}^2}f(n) - \sum_{|n|\leq\sqrt{X}}1\right| \leq \sum_{\sqrt{X}-Y\leq m\leq\sqrt{X}+Y} r_2(m) \ll YX^{\varepsilon}\,.$$

On the other hand,

$$\sum_{n\in\mathbb{Z}^2}f(n) = \sum_{k\in\mathbb{Z}^2}\hat{f}(k) = X\sum_{k\in\mathbb{Z}^2}\hat{\varphi}\left(Yk\right)\hat{\chi}_D(\sqrt{X}k)\,.$$

Since $\hat{\varphi}(0) = 1$, $\hat{\chi}_D(0) = \pi$, $k = 0$ gives the main term πX. Thus with $\psi(|k|) = \hat{\varphi}(k)$ and $\hat{\chi}_D(k) = \frac{1}{|k|} J_1(2\pi |k|)$, we have

$$\sum_{n \in \mathbb{Z}^2} f(n) = \pi X + X \sum_{m=1}^{\infty} r_2(m) \psi\left(Y\sqrt{m}\right) \frac{J_1\left(2\pi\sqrt{mX}\right)}{\sqrt{mX}}.$$

Thus:

$$\sum_{m \leq X} r_2(m) = \pi X + O(YX^{\varepsilon}) + \sqrt{X} \sum_{m=1}^{\infty} r_2(m) m^{-1/2} \psi\left(Y\sqrt{m}\right) J_1\left(2\pi\sqrt{mX}\right).$$

We now break the sum in two. For m small we use $J_1(x) = O(x^{-1/2})$ and $|\psi(x)| \leq 1$ to get

$$\left|\sqrt{X} \sum_{m \leq T} r_2(m) m^{-1/2} \psi\left(Y\sqrt{m}\right) J_1\left(2\pi\sqrt{mX}\right)\right| \ll X^{1/4} \sum_{m \leq T} m^{\varepsilon - 3/4} \ll X^{1/4} T^{\varepsilon + 1/4}.$$

For m large, we add the decay of ψ to get

$$\left|\sqrt{X} \sum_{m > T} r_2(m) m^{-1/2} \psi\left(Y\sqrt{m}\right) J_1\left(2\pi\sqrt{mX}\right)\right| \ll X^{1/4} \sum_{m > T} m^{\varepsilon - \frac{3}{4}} \left(Y\sqrt{m}\right)^{-k} \ll X^{1/4} Y^{-k} T^{\varepsilon + \frac{1}{4} - \frac{k}{2}}.$$

We thus get

$$\sum_{m \leq X} r_2(m) = \pi X + O\left(YX^{\varepsilon} + X^{1/4} T^{\varepsilon + \frac{1}{4}} + X^{1/4} Y^{-k} T^{\varepsilon + \frac{1}{4} - \frac{k}{2}}\right).$$

Trying $Y = X^{\alpha}$, $T = X^{\beta}$ the error term is

$$X^{\alpha} + X^{\frac{\beta+1}{4}} \left(1 + X^{-(\alpha + \beta/2)k}\right)$$

CHAPTER 3

The Prime Number Theorem

We'd like to estimate $\sum_{n\leq x} a_n$. We first try estimating $\sum_n a_n \varphi\left(\frac{n}{X}\right)$. We saw how to derive estimate on this via an additive Fourier expansion of φ (Poisson sum). We will now make a *multiplicative* Fourier expansion. We first investigate the associated transform.

3.1. Preliminaries

3.1.1. The Mellin Transform.

DEFINITION 95. For a reasonable φ defined on $\mathbb{R}_{>0}$ set

$$\tilde{\varphi}(s) = \int_0^\infty \varphi(x) x^s \frac{\mathrm{d}x}{x}$$

can call this the *Mellin transform* of φ.

This is the Fourier transform on the locally compact abelian group $\mathbb{R}^\times_{>0}$, isomorphic to $\mathbb{R}^+$ via the logarithm map. We thus get:

THEOREM 96. *Suppose φ decays rapidly enough. Then $\tilde{\varphi}$ extends to a meromorphic function, and in any vertical strip where the integrals converge absolutely, we have*

$$\varphi(x) = \frac{1}{2\pi i}\int_{(c)} \tilde{\varphi}(s) x^{-s}\,\mathrm{d}s\,.$$

3.1.2. Zetafunction counting.

3.1.2.1. *Setup and motivation.* For this fix a smoth cutoff $\varphi \in C_c^\infty(\mathbb{R})$. We then have for c large enough that

$$\begin{aligned}\sum_{n=1}^\infty a_n \varphi\left(\frac{n}{X}\right) &= \frac{1}{2\pi i}\sum_{n=1}^\infty a_n \int_{c-i\infty}^{c+i\infty} \tilde{\varphi}(s)\left(\frac{n}{X}\right)^{-s} \mathrm{d}s \\ &= \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \tilde{\varphi}(s) X^s D(s)\,\mathrm{d}s\,,\end{aligned}$$

assuming the integral and the series converge absolutely. Here $D(s)$ is the multiplicative generating series

$$D(s) = \sum_{n=1}^\infty a_n n^{-s}\,.$$

We need c large enough so that the series converges absolutely, and small enough to be in the strip of definition of $\tilde{\varphi}$.

COROLLARY 97. *When everything converges absolutely, we have*

$$\left|\sum_{n=1}^\infty a_n \varphi\left(\frac{n}{X}\right)\right| \ll \left(\int_{(c)} |\tilde{\varphi}(s) D(s)|\,\mathrm{d}\,|s|\right) X^c\,.$$

In particular,

$$\left|\sum_{n=1}^{\infty} a_n \varphi\left(\frac{n}{X}\right)\right| \ll_\varepsilon X^{\sigma_{ac}+\varepsilon}.$$

We would now like to shift the contour of integration as far to the left as possible, depending on the domain of holomorphy of $D(s)$ and $\tilde{\varphi}(s)$. This would have the effect of making the X^s term smaller. Along the way we pick up contributions of the form $X^\rho \operatorname{Res}_{s=\rho} \tilde{\varphi}(s)D(s)$ where ρ ranges over poles of $D(s)$,and $\frac{1}{s}$.

We are therefore motivated to investiate analytical continuation of $D(s)$ as far to the left as possible.

Why use a smooth cutoff? Suppose we took $\varphi(x) = \mathbb{1}_{[0,1]}$. Then $\tilde{\varphi}(s) = \frac{1}{s}$, and the integral $\int D(s)X^s \frac{\mathrm{d}s}{s}$ may only converge conditionally.

3.1.2.2. *Multiplicative smoothing.* Define $(f \star g)(x) = \int_{y=0}^{\infty} f(y)g\left(\frac{x}{y}\right)\frac{\mathrm{d}y}{y}$ (Multiplicative convolution). Then in the region of absolute convergence,

$$\begin{aligned}\widetilde{(f\star g)}(s) &= \int_0^\infty x^s \frac{\mathrm{d}x}{x}\int_{y=0}^\infty f(y)g\left(\frac{x}{y}\right)\frac{\mathrm{d}y}{y} \\ &= \int_{y=0}^\infty y^s f(y)\frac{\mathrm{d}y}{y}\int_{x=0}^\infty \left(\frac{x}{y}\right)^s g\left(\frac{x}{y}\right)\frac{\mathrm{d}x}{x} = \tilde{f}(s)\tilde{g}(s).\end{aligned}$$

Let $\psi_\delta(x) = \frac{1}{\delta}\rho\left(\frac{\log x}{\delta}\right)$ for some positive symmetric test function $\rho \in C_c^\infty(\mathbb{R})$ integrating to 1 and supported in $[-1,1]$. Let $\varphi_\delta = \mathbb{1}_{[0,1]} \star \psi_\delta$. Then

$$\tilde{\psi}_\delta(s) = \int_0^\infty \rho\left(\frac{\log x}{\delta}\right)x^s\frac{\mathrm{d}x}{\delta x} \overset{x=y^\delta}{=} \int_0^\infty \rho(\log y)y^{\delta s}\frac{\mathrm{d}y}{y} = \tilde{\psi}(\delta s),$$

so

$$\tilde{\varphi}_\delta(s) = \frac{1}{s}\tilde{\psi}(\delta s).$$

Counting with $0 \le \varphi_\delta \le 1$ we find

$$\sum_{n=1}^\infty a_n\varphi_\delta\left(\frac{n}{X}\right) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\tilde{\psi}(\delta s)X^s D(s)\frac{\mathrm{d}s}{s},$$

where $\tilde{\psi}(s)$ is entire and decays vertically. Note that for δs small, $\tilde{\psi}(\delta s) = \tilde{\psi}(0) + O(\delta s) = 1 + O(\delta s)$. Note that $\varphi_\delta(x) = 1$ for $x \le e^{-\delta}$, $\psi_\delta(x) = 0$ for $x \ge e^\delta$, and that $0 \le \psi_\delta(x) \le 1$ in between. Thus

$$\sum_{n=1}^\infty a_n\varphi_\delta\left(\frac{n}{X}\right) - \sum_{n\le X} a_n \le \sum_{e^{-\delta}X\le n\le e^\delta X}|a_n|.$$

If a_n are bounded, for example, this is $O(\delta X)$, and if the first pole is at the expected point $s=1$ then using the Taylor explansion we get

$$\sum_{n\le X} a_n = \tilde{\psi}(\delta)\operatorname{Res}_{s=1}D(s)X + O(\delta X) + \text{error} = \operatorname{Res}_{s=1}D(s)X + O(\delta X) + \text{error}$$

Now say error comes from integrating on $\Re(s) = \sigma$, that $D(s)$ is bounded on that line. Get error like $X^\sigma\int_{\Re(s)=\delta\sigma}|\tilde{\psi}(s)|\frac{|\mathrm{d}s|}{|s|}$. This is very similar to the constant $\int_{\Re(s)=0}\tilde{\psi}(s)\frac{ds}{s}$ except for part near zero, so we get error of size $\delta^{-1}X^{\sigma}$ from the $\frac{1}{s}$. Thus we have roughly

$$\sum_{n\le X} a_n = \mathrm{Res}_{s=1} D(s)X + O(\delta X) + O(\delta^{-1}X^{\sigma})$$

so the error is $O(X^{\frac{\sigma+1}{2}})$. Not as good as what we should get (X^{σ}) but requires less work.

3.1.3. Convergence of Dirichlet series. Given an arithmetical function$\{a_n\}_{n=1}^{\infty} \subset \mathbb{C}$ get the Dirichlet series $D(s) = \sum_{n\ge 1} a_n n^{-s}$ (generating function for multiplicative convlution). Let $R \subset \mathbb{C}$be its domain of convergence, R_{a} be its domain of absolute convergence.

LEMMA 98. *R is non-empty iff $|a_n| = O(n^T)$ for some T.*

PROPOSITION 99 (Domain of convergence). *Fix $s_0 = \sigma_0 + it_0 \in \mathbb{C}$*

(1) Suppose $D(s)$ converges absolutely at s_0. Then it converges absolutely in the half-plane $\{\Re(s) \ge \sigma_0\}$, uniformly absolutely in any half-plane $\{\Re(s) > \sigma_0 + \varepsilon\}$.
(2) Suppose $D(s)$ converges at s_0. Then it converges in the half-plane $\{\Re(s) > \sigma_0\}$, uniformly in any half-plane $\{\Re(s) > \sigma_0 + \varepsilon\}$. Furthermore, it converges absolutely in $\Re(s) > \sigma_0 + 1$.

COROLLARY 100. *Suppose R is non-empty. Then the interiors of R and R_a are half-planes $\{\sigma > \sigma_c\} \supset \{\sigma > \sigma_{ac}\}$.*

DEFINITION 101. $\sigma_{\mathrm{c}}, \sigma_{\mathrm{ac}}$ are called the *abcissas* of *convergence* and *absolute convergence*, respectively.

EXAMPLE 102. The abcissa of convergence and absolute convergence of $\zeta(s) = \sum_{n\ge 1} n^{-s}$ is clearly $\sigma_{\mathrm{c}} = 1$. The function blows up there by the MCT since $\sum_{n\ge 1} n^{-1} = \infty$. But $\zeta(s) = \prod_p (1-p^{-s})^{-1}$ and each individual factor is regular at $s = 1$ (the poles are at $2\pi i \log p\mathbb{Z}$). We conclude that there are infinitely many primes.

We can show a little more by elementary means. Let $D(s) = \sum_{n=1}^{\infty} (-1)^n n^{-s}$. This converges for $\sigma > 0$ by Dirichlet's criterion, hence for $\Re(s) > 0$. For $\Re(s) > 1$ we have $D(s) + \zeta(s) = 2\sum_{k=1}^{\infty} (2k)^{-s} = 2\cdot 2^{-s}\zeta(s)$. It follows that $\zeta(s) = -\frac{D(s)}{1-2^{1-s}}$ on $\Re(s) > 1$, showing that $\zeta(s)$ continues meromorphically to $\Re(s) > 0$. At $s = 1$ $\frac{1}{1-2^{1-s}}$ has a simple pole wtih residue $-\frac{1}{\log 2}$, and $D(1) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\log 2 \ne 0$ so $\zeta(s)$ has a simple pole at $s = 1$ with residue 1. We will later see that $\zeta(s)$ is regular at $1 + it$, $t \ne 0$ so the other singularities of $\frac{D(s)}{1-2^{1-s}}$ are removable.

3.2. Counting primes with the Riemann zetafunction

After Gauss it is natural to count primes with the weight $\log p$. Riemann pointed out it is better to count with with von Mangolt function. We note that:

Consider the *logarithmic derivative*

$$\begin{aligned} -\frac{\zeta'(s)}{\zeta(s)} &= \sum_p \frac{\mathrm{d}}{\mathrm{d}s} \log\left(1-p^{-s}\right) = \sum_p \log p \frac{p^{-s}}{1-p^{-s}} = \sum_p \sum_{m=1}^{\infty} \frac{\log p}{p^{ms}} \\ &= \sum_{n\ge 1} \Lambda(n) n^{-s}. \end{aligned}$$

The latter series converges absolutely for $\Re(s) > 1$. Thus, for $c > 1$,

$$\sum_{n=1}^{\infty} \Lambda(n)\varphi\left(\frac{n}{X}\right) = -\frac{1}{2\pi i}\int_{(c)} \frac{\zeta'(s)}{\zeta(s)}\tilde{\varphi}(s)X^s\,\mathrm{d}s\,.$$

We have already seen that $\zeta(s)$ continues meromorphically to $\Re(s) > 0$, with a unique pole at $s = 1$. Recall, however, that the logarithmic derivative has a pole at every zero and pole of the original function, with residue equal to the order. Thus, shifting formally to some $c' < 1$, and assuming there are no zeroes on the line $\Re(s) = c'$ itself, we formally:

$$\sum_{n=1}^{\infty} \Lambda(n)\varphi\left(\frac{n}{X}\right) = \tilde{\varphi}(1)X - \sum_{\zeta(\rho)=0} \tilde{\varphi}(\rho)X^{\rho} - \frac{1}{2\pi i}\int_{(c')} \frac{\zeta'(s)}{\zeta(s)}\tilde{\varphi}(s)X^s\,\mathrm{d}s\,.$$

The first term is the desires main term, conjectured by Gauss. Assuming $\frac{\zeta'(s)}{\zeta(s)}$ does not grow too fast, the last term is clearly an error term. The problem is with the term in the middle – we have no idea where the zeros are, or how many there are. If $\Re(\rho)$ is close to one (perhaps equal to one) or if they are very dense , these "error terms" could overwhelm the main term. In the next parts we first analytically continue $\zeta(s)$ to all of C, allowing us to take c' to $-\infty$. We then establish enough about the zeroes to prove the Riemann's formula above. We then improve our control on the zeroes (obtaining the "zero free region") to obtain the Prime Number Theorem.

3.2.1. Analytical continuation of the Riemann zetafunction. For even $\varphi \in \mathcal{S}(\mathbb{R})$ set $\varphi(r\mathbb{Z}) = \sum_{n\in\mathbb{Z}} \varphi(rn) - \varphi(0)$. This decays faster than any polynomial at infinity, and grows at most like r^{-1} as $r \to 0$. It follows that the Mellin transform

$$Z(\varphi;s) = \int_0^{\infty} \varphi(r\mathbb{Z})\,r^s\frac{\mathrm{d}r}{r}$$

converges absolutely for $\Re(s) > 1$. In that domain we may exchange summation and integration to get

$$Z(\varphi;s) = 2\zeta(s)\tilde{\varphi}(s)\,.$$

Since $\tilde{\varphi}(s)$ can be chosen entire (say if φ is ompactly supported away from 0),to meromorphically continue $\zeta(s)$ it is enough to continue $Z(\varphi;s)$.

PROPOSITION 103. *$Z(\varphi;s)$ extends (AC) to a meromorphic function in $\mathbb{C}$, (BVS) bounded in vertical strips, satisfying (FE) the* functional equation

$$Z(\varphi;s) = Z(\hat{\varphi};1-s)$$

and with simple poles at $s = 0,1$ where the residues are $-\varphi(0)$, $\hat{\varphi}(0)$ respectively.

PROOF. Calculation, using Poisson sum:

$$\begin{aligned}
Z(\varphi;s) &= \int_0^\infty \varphi(r\mathbb{Z})r^s\frac{\mathrm{d}r}{r} \\
&\int_1^\infty \varphi(r\mathbb{Z})r^s\frac{\mathrm{d}r}{r}+\int_0^1\left[\sum_{n\in\mathbb{Z}}\varphi(rn)\right]r^s\frac{\mathrm{d}r}{r}-\varphi(0)\int_0^1 r^s\frac{\mathrm{d}r}{r} \\
&\int_1^\infty \varphi(r\mathbb{Z})r^s\frac{\mathrm{d}r}{r}-\frac{\varphi(0)}{s}+\int_0^1\left[\sum_{n\in\mathbb{Z}}\hat{\varphi}(r^{-1}n)\right]r^s\frac{\mathrm{d}r}{r} \\
&\int_1^\infty \varphi(r\mathbb{Z})r^s\frac{\mathrm{d}r}{r}-\frac{\varphi(0)}{s}+\int_1^\infty\left[\sum_{n\in\mathbb{Z}}\hat{\varphi}(rn)\right]r^{1-s}\frac{\mathrm{d}r}{r} \\
&\int_1^\infty \varphi(r\mathbb{Z})r^s\frac{\mathrm{d}r}{r}-\frac{\varphi(0)}{s}+\int_1^\infty \hat{\varphi}(r\mathbb{Z})r^{1-s}\frac{\mathrm{d}r}{r}-\frac{\hat{\varphi}(0)}{1-s}.
\end{aligned}$$

□

Now let $\varphi(x)=e^{-\pi x^2}$. Then $\hat{\varphi}=\varphi$ and $\tilde{\varphi}(s)=\int_0^\infty e^{-\pi x^2}x^s\frac{\mathrm{d}x}{x}=\int_0^\infty e^{-t}\left(\frac{t}{\pi}\right)^{s/2}\frac{1}{2}\frac{\mathrm{d}t}{t}=\frac{1}{2}\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)$.

DEFINITION 104. $\Gamma_{\mathbb{R}}(s)=\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)$.

COROLLARY 105. *Let $\xi(s)=\Gamma_{\mathbb{R}}(s)\zeta(s)$. Then $\xi(s)$ has AC, BVS, the FE*

$$\xi(s)=\xi(1-s)$$

and with poles at $s=0$ (residue -1) and at $s=1$ (residue 1). Moreover, $\zeta(k)=0$ for $k\in-2\mathbb{Z}_{\geq 1}$.

THEOREM 106. *$\zeta(s)$ itself is polynomially bounded in vertical strips.*

PROOF. For $\Re(s)\geq\sigma>1$, $\zeta(s)$ is uniformly bounded by absolute convergence. By the functional equation,

$$\zeta(1-s)=\frac{\Gamma_{\mathbb{R}}(s)}{\Gamma_{\mathbb{R}}(1-s)}\zeta(s).$$

Stirling's approximation shows:

$$\begin{aligned}
\Re\log\Gamma_{\mathbb{R}}(\sigma+it) &= -\frac{s}{2}\log\pi+\frac{s-1}{2}\log\frac{s}{2}-\frac{s}{2}+\frac{1}{2}\log(2\pi)+O\left(\frac{1}{t}\right) \\
&= C(\sigma)+\left(\frac{\sigma-1}{2}\right)\Re\log s-\frac{t}{2}\Im\log\frac{\sigma+it}{2}+O\left(\frac{1}{t}\right) \\
&= C(\sigma)+\left(\frac{\sigma-1}{2}\right)\log\left(t\sqrt{1+\frac{\sigma^2}{t^2}}\right)-\frac{t}{2}\arccos\left(\frac{\sigma}{t}\cdot\frac{1}{\sqrt{1+(\sigma/t)^2}}\right)+O(\frac{1}{t}) \\
&= C(\sigma)+\frac{\sigma-1}{2}\log t-\frac{t}{2}\left[\frac{\pi}{2}-\frac{\sigma}{t}+O(\frac{1}{t^2})\right]+O(\frac{1}{t}) \\
&= C(\sigma)+\frac{\sigma-1}{2}\log t-\frac{\pi t}{4}+O\left(\frac{1}{t}\right).
\end{aligned}$$

In other words,

$$\Gamma_{\mathbb{R}}(\sigma+it)=C(\sigma)\,|t|^{\frac{\sigma-1}{2}}e^{-\frac{\pi}{4}|t|}\left(1+O\left(\frac{1}{t}\right)\right).$$

Note that the exponential decay term is independent of σ. Thus for $s=\sigma+it$ with $\sigma>1$ we have

$$\begin{aligned}\zeta(1-s) &= \frac{C(\sigma)}{C(1-\sigma)}|t|^{\frac{\sigma-1}{2}-\frac{(1-\sigma)-1}{2}}\left(1+O\left(\frac{1}{t}\right)\right)\zeta(s)\\ &\ll |t|^{\sigma-\frac{1}{2}}.\end{aligned}$$

Finally, we apply Phragmen–Lindelöf. □

COROLLARY 107. *$\xi(s)$ is of order* 1.

PROOF. By FE enough to check for $\sigma\geq\frac{1}{2}$. There $\Gamma_{\mathbb{R}}(s)$ satisfies the bound by Stirling, and $\zeta(s)$ grows at most polynoimally as we saw above. □

3.2.2. Functions of finite order.

DEFINITION 108. Call an entire function f of *order*$\leq\alpha$ if $|f(z)|\ll_{\varepsilon}\exp\left(|z|^{\alpha+\varepsilon}\right)$. The *order* is the least α for which this holds.

Call a meromorphic function of *order*$\leq\alpha$ if it is of the form $\frac{f}{g}$ where f,g are entire of order$\leq$ α.

LEMMA 109. *The set of entire functions of order$\leq\alpha$ (or finite order) is an algebra; the corresponding sets of meromorphic functions are divison algebras.*

LEMMA 110 (Jensen's formula). *Let f be holomorphic in $|z|<R$, continuous in the closed ball, and non-vanishing on the circle and at the origin. Then*

$$\frac{1}{2\pi}\int_0^{2\pi}\log\left|f\left(Re^{i\theta}\right)\right|\mathrm{d}\theta=\log|f(0)|+\sum_{f(z_k)=0}\log\frac{R}{|z_k|}$$

where the sum is over the zeroes in the ball, counted with multiplicity.

PROOF. Write $f(z)=g(z)\prod_{k=1}^{n}(z-z_k)$ with g non-vanishing. Then the formula holds for g since $\log|g(z)|$ is harmonic, and for $z-z_k$ by direct calculation. □

COROLLARY 111. *Let f have order$\leq\alpha$, and let $\{z_k\}_{k=1}^{\infty}$ enumerate its zeroes. Then $\sum_{k=1}^{\infty}|z_k|^{-\beta}$ converges for any $\beta>\alpha$.*

PROOF. $\log\left|f(Re^{i\theta})\right|<R^{\alpha+\varepsilon}$ for R large enough. By the maximum principle also $|f(0)|\leq R^{\alpha+\varepsilon}$ and hence

$$\sum_{\substack{f(z_k)=0\\R/2\leq|z_k|<R}}\log 2\leq\sum_{\substack{f(z_k)=0\\0<|z_k|<R}}\log\frac{R}{|z_k|}\leq 2R^{\alpha+\varepsilon}$$

so the number of zeroes of f of magnitude between $\frac{R}{2}$ and R is at most $\frac{2R^{\alpha+\varepsilon}}{\log 2}$. Thus, ignoring the finite contributions small radii,

$$\begin{aligned}\sum_{k=1}^{\infty}|z_k|^{-\beta} &\le C+\sum_{n=N}^{\infty}\sum_{2^n\le|z_k|<2^{n+1}}|z_k|^{-\beta}\le C+\sum_{n=N}^{\infty}2^{-\beta n}\cdot 2\cdot\left(2^{n+1}\right)^{\alpha+\varepsilon}\\ &\ll \sum_{n=N}^{\infty}2^{-(\beta-\alpha-\varepsilon)n}.\end{aligned}$$

Now choosing ε small enough so that $\alpha+\varepsilon<\beta$ solves the problem. □

THEOREM 112 (Hadamard factorization). *Let f have order $\le\alpha$, with zeros $\{z_k\}_{k=1}^{\infty}$ excepting possibly zero. Then for some polynomial g of degree $\le\alpha$,*

$$f(z)=e^{g(z)}z^e\prod_{k=1}^{\infty}\left(1-\frac{z}{z_k}\right)\exp\left\{\sum_{1\le m\le\alpha}\frac{1}{m}\left(\frac{z}{z_k}\right)^m\right\}.$$

COROLLARY 113. *We have the product representation*

$$s(s-1)\xi(s)=e^{Bs}\prod_{\xi(\rho)=0}\left(1-\frac{s}{\rho}\right)e^{s/\rho}, \tag{3.2.1}$$

where ρ runs over the zeroes of $\xi(s)$, which all occur in the critical strip.

PROOF. Applying the theorem gives this except the initial exponential is e^{A+Bs}. $s(s-1)\xi(s)\xrightarrow[s\to 1]{}\mathrm{Res}_{s=1}\xi(s)=1$ so by the FE the function has the value 1 at 0, and $e^A=1$. □

3.2.3. Counting zeroes. Taking the logarithmic derivative of (3.2.1) gives:

$$-\frac{\zeta'}{\zeta}(s)=\frac{1}{2}\frac{\Gamma'}{\Gamma}(\frac{s}{2})-\frac{1}{2}\log\pi-B+\frac{1}{s}-\frac{1}{1-s}-\sum_{\rho}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right). \tag{3.2.2}$$

It will be useful to recall the Stirling's approximation for the digamma function:

$$F(s)\overset{\text{def}}{=}\frac{\Gamma'(s)}{\Gamma(s)}=\log s-\frac{1}{2s}+O_\delta\left(|s|^{-2}\right),$$

valid in any cone $|\arg(s)|\le\pi-\delta$ (for proof see PS0).

LEMMA 114. *Let $\rho=\beta+i\gamma$ run through the zeroes. Then as $T\to\infty$,*

$$\sum_{\rho}\frac{1}{4+(T-\gamma)^2}=O(\log T).$$

PROOF. Setting $s=2+iT$ in (3.2.2), we have $F\left(\frac{s}{2}\right)=O(\log T)$ by Stirling's formula, so

$$-\Re\frac{\zeta'}{\zeta}(s)=O(\log T)-\Re\sum_{\rho}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right).$$

Next, $-\frac{\zeta'}{\zeta}(s)=\sum_{n\ge 1}\Lambda(n)n^{-s}$ is uniformly bounded in any halfplane $\sigma\ge 1+\varepsilon$, so we get

$$\Re\sum_{\rho}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right)=O(\log T).$$

Finally, $\Re\frac{1}{\rho} = \frac{\beta}{|\rho|^2} > 0$ and $\Re\frac{1}{s-\rho} = \frac{2-\beta}{|s-\rho|^2} > 0$ ($0 \le \beta \le 1$), so each term in the series is positive. Specifically,

$$\Re\frac{1}{s-\rho} = \frac{2-\beta}{(2-\beta)^2+(T-\gamma)^2} \ge \frac{1}{4+(T-\gamma)^2},$$

and the claim follows. $\square$

COROLLARY 115. *$N(T+1) - N(T-1) = O(\log T)$, and $\sum_{|\gamma-T|>1} \frac{1}{(T-\gamma)^2} = O(\log T)$.*

Next, let T be large and let $-1 \le \sigma \le 2$. Subtracting (3.2.2) evaluated at $s = \sigma + iT, 2+iT$ we get

$$\frac{\zeta'(s)}{\zeta(s)} = O(1) + \sum_\rho \left(\frac{1}{s-\rho} - \frac{1}{2+iT-\rho}\right),$$

since $F\left(\frac{\sigma+iT}{2}\right) - F\left(\frac{2+iT}{2}\right) = O(1)$. Now for ρ with $\gamma \notin (T-1, T+1)$, we have

$$\left|\frac{1}{s-\rho} - \frac{1}{2+iT-\rho}\right| \le \frac{2-\sigma}{|s-\rho|\,|2+iT-\rho|} \le \frac{3}{|\gamma-T|^2}$$

and for ρ with $\gamma \in (t-1, t+1)$ we have $\left|\frac{1}{2+iT-\rho}\right| \le \frac{1}{|2-\beta|} \le 1$. We have shown:

LEMMA 116. *Let $s = \sigma + iT$, $\sigma \in [-1,2]$. Then*

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{\gamma\in(T-1,T+1)} \frac{1}{s-\rho} + O(\log T).$$

COROLLARY 117. *For each $T > 2$ there exists $T' \in [T, T+1]$ such that for $s = \sigma + iT'$, $\sigma \in [-1,2]$ we have*

$$\frac{\zeta'(s)}{\zeta(s)} = O(\log^2 T').$$

PROOF. There are $O(\log T)$ zeroes with $\gamma \in [T, T+1]$. In particular, there is a gap of length $O(\frac{1}{\log T})$ there, and we can choose T' in the middle of the gap. Then $|\gamma - T'| \gg \frac{1}{\log T}$ for all zeroes of the zetafunction, so that

$$\left|\frac{\zeta'}{\zeta}(\sigma+iT')\right| \ll \left(N(T'+1) - N(T'-1)\right) O(\log T) + O(\log T') = O(\log^2 T').$$

$\square$

THEOREM 118. *$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T)$.*

PROOF. Suppose T is not the ordinate of any zero, and let R be the rectangle $[-1,2]\times[-T,T]$. We need to calculate the real number

$$2N(T) - 2 = \frac{1}{2\pi i}\oint_{\partial R} \frac{\xi'(s)}{\xi(s)}\,\mathrm{d}s.$$

Since $\overline{\xi(\bar{s})} = \xi(s)$ and by the functional equation $\xi(1-s) = \xi(s)$, it is enough to consider the quarter-rectangle $2 \to 2+iT \to \frac{1}{2}+iT$. Recall that $\xi(s) = \pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)$. The argument of $\pi^{-s/2}$ changes exactly by $-\frac{1}{2}T\log\pi$. The argument of $\Gamma\left(\frac{s}{2}\right)$ changes by $\Im\log\Gamma\left(\frac{1}{4}+\frac{1}{2}iT\right) = \frac{T}{2}\log\left(\frac{T}{2}\right) - \frac{\pi}{8} - \frac{T}{2} + O(T^{-1})$. It remains to estimate the change $S(T)$ in $\arg\zeta(s)$. Since $\Re(\zeta(2+$

$it)) \geq 1 - \sum_{n=2}^{\infty} \frac{1}{n^2} > 0$, the change of argument in $[2, 2+iT]$ is at most π. On $\left[\frac{1}{2}+iT, 2+iT\right]$ Lemma 116 gives:

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{\gamma \in (t-1, t+1)} (\log(s-\rho))' + O(\log T)\,.$$

Now the change of the argument of each $s-\rho$ on the interval is at most $\frac{\pi}{2}$, so the total change in the argument of $\zeta(s)$ is $O(\log T)$. In summary, we have:

$$2\frac{1}{4}2\pi N(T) = \frac{T}{2}\log\left(\frac{T}{2}\right) - \frac{T}{2}\log\pi - \frac{T}{2} + O(\log T)\,.$$

□

REMARK 119. Note that the "main" term came from the argument of $\Gamma_{\mathbb{R}}(s)$, the "error term" from the argument of $\zeta(s)$, despite the zeroes being those of $\zeta(s)$. The reason is the functional equation, which is symmetrical only for $\xi(s)$. In the left half of the rectangular path, the argument of $\zeta(s)$ will change considerably (note that $\zeta(1-s) \neq \zeta(s)$!).

The functional equation connects the zeroes $\rho, 1-\rho$ and hence zeroes with opposite imaginary parts, showing that indeed R contains $2N(T)$ zeroes. The real-on-the real axis relation $\overline{\xi(\bar{s})} = \xi(s)$ shows that the zeroes are symmetric about the critical line. In particular, a zero slightly off the line must have a "partner" on the other side, and so a numerical countour integral argument can *prove* that a suspected simple zero is *exactly* on the line rather than off it. Of course, a double zero would be indistinguishable from two off-the-line zeroes, but no such double zero has ever been found, and conjecturally they don't exist.

REMARK 120. A more precise version of the Theorem is

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} + S(T) + O(T^{-1})\,.$$

- This is easy to prove (just keep track of the constant term in the Stirling approximation and of the contribution of the two poles).
- To see that this is significant note that (Littlewood)

$$\int_0^T S(t)\,\mathrm{d}t = O(\log T)\,,$$

  showing massive cancellation. It is clear that the term $\frac{7}{8}$ is significant when averaging the rest of the formula.
- This is important in numerical calculation of the zeroes: suppose we missed a zero between $[0, T - O(\log T)]$. Then $\int_0^T N(t)\,\mathrm{d}t$ will be small by $O(\log T)$. But the RHS can be calculated to that precision.

Now let $N_0(T)$ denote the number of zeroes $\frac{1}{2}+i\gamma$, $0 \leq \gamma \leq T$. Hardy–Littlewood shows that $N_0(T) \gg T$. This was improved:

THEOREM 121. *Let* $\kappa = \liminf_{T\to\infty} \frac{N_0(T)}{N(T)}$, κ^* *similar for simple zeroes. Then:*

(1) (Selberg 1942) $\kappa > 0$.
(2) (Levinson 1974) $\kappa > 34.74\%$
(3) (Heath–Brown 1979) $\kappa^* > 34.74\%$
(4) (Conrey 1989) $\kappa > 40.88\%$, $\kappa^* > 40.13\%$

(5) (Bui–Conrey–Young 2012) $\kappa \geq 41.05\%$, $\kappa^* \geq 40.58\%$
(6) (Feng 2012) $\kappa \geq 41.28\%$

THEOREM 122. *(Zero density estimate)*

3.2.4. A smooth cutoff. Let $\eta \in C_c^\infty(\mathbb{R})$ be positive, supported in $[-1,1]$ and such that $\int_{\mathbb{R}} \eta = 1$. For $H > 0$ set $\eta_H(x) = H\eta\,(H\log x)$, and let $\varphi_H = \eta_H * \mathbb{1}_{[0,1]}$ (multiplicative convolution), a smooth function on $\mathbb{R}_{>0}^\times$. Then $\tilde{\varphi}_H(s) = \tilde{\eta}_H(s)\tilde{\mathbb{1}}_{[0,1]}(s)$. The second integral is $\frac{1}{s}$ so that

$$\tilde{\varphi}_H(s) = \frac{1}{s}\tilde{\eta}_H(s)\,.$$

LEMMA 123. *We have $\tilde{\eta}_H(0) = 1$, and*

$$\tilde{\eta}_H(s) = \int_{\mathbb{R}} \eta(u)\exp\left\{\frac{us}{H}\right\}\mathrm{d}u\,,$$

and in particular the estimates:

(1) For $\frac{|s|}{H}$ bounded, $\tilde{\eta}_H(s) = 1 + O(\frac{|s|}{H})$.
(2) In the region $|\Re(s)| \leq A$ we have for $k \geq 0$ that

$$|\tilde{\eta}_H(s)| \ll_{\eta,k} \exp\left\{\frac{A}{H}\right\}\frac{H^k}{|s|^k}\,.$$

PROOF. Setting $x = e^{\frac{u}{H}}$ in the Mellin transform gives:

$$\tilde{\eta}_H(s) = \int_0^\infty H\eta\,(H\log x)\,x^s\frac{\mathrm{d}x}{x} = \int_{-\infty}^\infty \eta(u)\exp\left\{\frac{us}{H}\right\}\mathrm{d}u\,.$$

Integrating by parts k times we get

$$\tilde{\eta}_H(s) = (-1)^k\left(\frac{H}{s}\right)^k\int_{-\infty}^\infty \eta^{(k)}(u)\exp\left\{\frac{us}{H}\right\}\mathrm{d}u\,,$$

and taking absolute values we get

$$\begin{aligned}|\tilde{\eta}_H(s)| &\leq \frac{H^k}{|s|^k}\int_{-\infty}^{+\infty}\left|\eta^{(k)}(u)\right|\exp\left\{\frac{u\Re(s)}{H}\right\}\mathrm{d}u \\ &\leq \left\|\eta^{(k)}\right\|_{L^1}\exp\left\{\frac{A}{H}\right\}\frac{H^k}{|s|^k}\,.\end{aligned}$$

□

COROLLARY 124. *$\tilde{\varphi}_H(s)$ extends to a meromorphic function in $\mathbb{C}$ with a unique pole at $s = 0$ with residue 1, satisfying the esimate*

$$\tilde{\varphi}_H(s) \ll \exp\left\{\frac{A}{H}\right\}\frac{H^k}{|s|^{k+1}}$$

in $|\Re s| \leq A$. In particular, $\tilde{\varphi}_H$ decays rapidly in vertical strips away from its pole at $s = 0$, and the Mellin inversion formula applies to it on any vertical line to the right of this pole.

3.2.5. The explicit formula.

LEMMA 125. *For $\sigma \le -1$ we have $\frac{\zeta'(s)}{\zeta(s)} = O(\log|s|)$.*

PROOF. By the duplication formula,

$$\zeta(1-s) = 2^{1-s}\pi^{-s}\cos\left(\frac{\pi s}{2}\right)\Gamma(s)\zeta(s)$$

and hence

$$\frac{\zeta'(1-s)}{\zeta(1-s)} = -\frac{1}{2}\pi\tan\left(\frac{\pi s}{2}\right) + \frac{\Gamma'(s)}{\Gamma(s)} + \frac{\zeta'(s)}{\zeta(s)}\,.$$

Now if $\sigma \ge 2$ the last term is $O(1)$, the second term is $O(\log|s|) = O(\log|1-s|)$ and if $1-s$ is away from the trivial zeroes, then the first term is $O(1)$ as well. □

PROPOSITION 126. *Let $U \ge 1$ not be an even integer. Then*

$$\sum_{n\le x}\Lambda(n) + O(\frac{x\log x}{H}) = x - \sum_{\rho}\tilde{\eta}_H(\rho)\frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} + \sum_{2m<U}\tilde{\eta}_H(-2m)\frac{x^{-2m}}{2m} - \frac{1}{2\pi i}\int_{(-U)}\frac{\zeta'(s)}{\zeta(s)}\tilde{\eta}_H(s)x^s\frac{\mathrm{d}s}{s}\,.$$

PROOF. In Section 3.1.2 we obtained the formula:

$$\sum_{n=1}^{\infty}\Lambda(n)\varphi_H\left(\frac{n}{x}\right) = -\frac{1}{2\pi i}\int_{(2)}\frac{\zeta'(s)}{\zeta(s)}\tilde{\varphi}_H(s)x^s\,\mathrm{d}s\,.$$

On the left-hand-side, $\varphi_H\left(\frac{n}{x}\right) = 1$ if $n \le x$, $\tilde{\varphi}_H\left(\frac{n}{x}\right) = 0$ if $n \ge xe^{1/H} = x + O(\frac{x}{H})$ and for $x \le n \le x + O(\frac{x}{H})$ we have $\Lambda(n)\varphi_H(\frac{n}{x}) \le \log x$, so

$$\text{LHS} = \sum_{n\le x}\Lambda(n) + O\left(\frac{x\log x}{H}\right)\,.$$

On the RHS we would like to shift the contour to the line $(-U)$. For this, let T not be the height of a zero and let $R_T = [-U,2]\times[-T,T]$. By the Residuum Theorem,

$$\frac{1}{2\pi i}\oint_{\partial R_T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\tilde{\varphi}_H(s)x^s\,\mathrm{d}s = \tilde{\varphi}_H(1)x - \sum_{|\gamma|<T}\tilde{\varphi}_H(\rho)x^\rho - \frac{\zeta'(0)}{\zeta(0)} - \sum_{2m<U}\tilde{\varphi}_H(-2m)x^{-2m}\,.$$

Thus

$$\begin{aligned}-\frac{1}{2\pi i}\int_{2-iT}^{2+iT}\frac{\zeta'(s)}{\zeta(s)}\tilde{\varphi}_H(s)x^s\,\mathrm{d}s \;&=\; x - \sum_{|\gamma|<T}\tilde{\eta}_H(\rho)\frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} + \sum_{2m<U}\tilde{\eta}_H(-2m)\frac{x^{-2m}}{2m} - \frac{1}{2\pi i}\int_{(-U)}\frac{\zeta'(s)}{\zeta(s)}\tilde{\varphi}_H(s)x^s\,\mathrm{d}s \\ &\quad +R_1(T) + R_2(U,T)\end{aligned}$$

where $R_1(T)$ represents the integral over $[-1,2]\times\{\pm T\}$, $R_2(U,T)$ the integral over $[-U,-1]\times\{\pm T\}$. Let T be on of the heights guaranteed by Corollary 117. Then

$$\begin{aligned}R_1(T) \;&\ll\; \int_{-1+iT}^{2+iT}\left[\log^2 T\right]\left[\exp\left\{\frac{2}{H}\right\}\frac{H^k}{T^{k+1}}\right]\left[x^2\right]\mathrm{d}s \\ &\ll\; x^2H^k\exp\left\{\frac{2}{H}\right\}\frac{\log^2 T}{T^{k+1}}\,.\end{aligned}$$

For the rest of the integration we use the bound $\left|\frac{\zeta'}{\zeta}(s)\right| \ll \log|s|$ of Lemma 125 to get

$$
\begin{aligned}
R_2(U,T) &\ll \int_{-U+iT}^{-1+iT} [\log|s|] \left[\exp\left\{\frac{U}{H}\right\} \frac{H^k}{|s|^{k+1}}\right] \left[x^{-1}\right] \mathrm{d}s \\
&\ll H^k \exp\left\{\frac{U}{H}\right\} \frac{U}{x} \frac{\log U + \log T}{|T|^{k+1}}.
\end{aligned}
$$

Now letting $T \to \infty$ we see $R_1(T), R_2(U,T) \to 0$. The superpolynomial decay of $\tilde{\varphi}_H(s)$ along vertical lines shows that the vertical integrals converge to the intergrals alone (2), $(-U)$ respectively. $\square$

REMARK 127. By the FE, $-\frac{\zeta'(0)}{\zeta(0)} = \log(2\pi)$.

COROLLARY 128 (von Mangoldt's explicit formula). *Interpreting the sum over the zeroes symmetrically, we have*

$$
\sum_{n \le x} \Lambda(n) = x - \sum_{\rho} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log(1 - x^{-2}).
$$

PROOF. We take $H = U \to \infty$ in the proposition. The LHS is fine. The last term on the RHS reads:

$$
\begin{aligned}
-\frac{1}{2\pi i}\int_{(-U)} \frac{\zeta'(s)}{\zeta(s)} \tilde{\eta}_H(s) x^s \frac{\mathrm{d}s}{s} &\ll \int_{-\infty}^{+\infty} \log|H+iT| \exp\left\{\frac{U}{H}\right\} \frac{H^k}{|H+iT|^{k+1}} x^{-H} \,\mathrm{d}T \\
&\ll x^{-H} \int_0^{+\infty} \frac{\log H + \log|1+iT|}{|1+iT|^k} \,\mathrm{d}T \\
&\ll x^{-H} \log H.
\end{aligned}
$$

We need to show:

$$
\lim_{H\to\infty} \sum_{\rho} \tilde{\eta}_H(\rho) \frac{x^\rho}{\rho} = \lim_{T\to\infty} \sum_{|\gamma|<T} \frac{x^\rho}{\rho}
$$

and

$$
\lim_{H\to\infty} \sum_{2m<H} \tilde{\eta}_H(-2m) \frac{x^{-2m}}{2m} = \frac{1}{2}\sum_{m=1}^{\infty} \frac{x^{-2m}}{m} = -\frac{1}{2}\log\left(1 - x^{-2}\right).
$$

For the second claim, we have

$$
\sum_{\rho} \tilde{\eta}_H(\rho) \frac{x^\rho}{\rho} = \sum_{\rho} \tilde{\eta}_1\left(\frac{\rho}{H}\right) \frac{x^\rho}{\rho}
$$

For the third claim, if $2m < H$ we have $\left|\tilde{\eta}_H(-2m)\frac{x^{-2m}}{2m}\right| \le \exp\left\{\left(\frac{2m}{H}\right)\right\} \frac{x^{-2m}}{2m} \le e\frac{x^{-2m}}{2m}$ and we are done by the bounded convergence theorem. $\square$

PROPOSITION 129. *Let $\beta(T)$ be such that if $|\gamma| < T$ then $\beta \le \beta(T)$. We then have*

$$
|\psi(x) - x| \ll \log^2 T \cdot x^{\beta(T)} + \frac{x\log x}{H} + \frac{xH\log T}{T}.
$$

PROOF. In the previous proposition take $U = 1$. Then the U integral reads

$$-\frac{1}{2\pi i}\int_{(-U)}\frac{\zeta'(s)}{\zeta(s)}\tilde{\eta}_H(s)x^s\frac{\mathrm{d}s}{s} \ll x^{-1}H\exp\left\{\frac{1}{H}\right\}\int_{-\infty}^{+\infty}\frac{\log|1+it|}{|1+it|^2}\,\mathrm{d}t\,.$$

Since the zero density is about $\log t$ at height t, and since $\tilde{\eta}_H(\rho) \ll 1$, we can bound $\sum_{|\gamma|<T}\tilde{\eta}_H(\rho)\frac{x^\rho}{\rho}$ by $x^{\beta(T)}\int_1^T\frac{\log t}{t}\,\mathrm{d}t = \log^2 T\cdot x^{\beta_{\max}}$. Similarly, $\left|\sum_{|\gamma|>T}\tilde{\eta}_H(\rho)\frac{x^\rho}{\rho}\right|$ is bounded by

$$x\int_T^\infty\frac{H}{t}\frac{\log t\,\mathrm{d}t}{t} \ll \frac{xH\log T}{T}\,.$$

□

THEOREM 130 (Prime Number Theorem). *Suppose, further, that every zero have $\beta \leq 1-\frac{c}{\log\gamma}$. Then*

$$|\psi(x)-x| \ll x\exp\left\{-c'\sqrt{\log x}\right\}\,.$$

On RH we have

$$\psi(x) = x + O(\sqrt{x}\log x)\,.$$

PROOF. On RH we have

$$\sqrt{x}\log^2 T + \frac{x\log x}{H} + \frac{xH\log T}{T}$$

can take $H = \sqrt{x}$, $T = x^2$. □

With zero-free region get bound

$$x\log^2 T\exp\left\{-c\frac{\log x}{\log T}\right\} + \frac{x\log x}{H} + \frac{xH\log T}{T}$$

and taking $T = \exp\left\{c_1(\log x)^{1/2}\right\}$, $H = \exp\left\{c_2\sqrt{\log x}\right\}$ with $c_1 > c_2$ works.

3.2.5.1. *Proof from Iwaniec–Kowalski.* We have

$$\sum_n \Lambda(n)\varphi_H\left(\frac{n}{x}\right) = \frac{1}{2\pi i}\int_{(2)}\left(-\frac{\zeta'}{\zeta}(s)\right)\tilde{\eta}_H(s)x^s\frac{\mathrm{d}s}{s}\,.$$

Shift to the contour $1-\sigma = \frac{c}{\log(|t|+2)}$. We pick up the pole, but not zeroes, and get

$$\sum_{n\leq x}\Lambda(n) + O\left(\frac{x\log x}{H}\right) \ll x\int_0^\infty\left(\log\left(|t|+2\right)\right)\tilde{\eta}_H\left(1-\frac{c}{\log\left(|t|+2\right)}+it\right)x^{-c/\log(|t|+2)}\frac{\mathrm{d}t}{|t|+2}$$

3.2.6. The zero-free region.

LEMMA 131 (Hadamard / de la Vallée Poussin; argument due to Mertens). *$\zeta(1+it) \neq 0$ if $t \neq 0$.*

PROOF. For $s = \sigma + it$, $\sigma > 1$ have $\log\zeta(s) = \sum_p\sum_{m=1}^\infty m^{-1}p^{-m\sigma}p^{-mit}$ so that

$$\Re\log\zeta(s) = \sum_p\sum_{m=1}^\infty m^{-1}p^{-m\sigma}\cos\left(t\log p^m\right)\,.$$

Using $2\left(1+\cos\theta\right)^2 = 3+4\cos\theta+\cos 2\theta \geq 0$, get

$$3\log\zeta(\sigma) + 4\Re\log\zeta(\sigma+it) + \Re\log\zeta(\sigma+2it) \geq 0\,,$$

that is

$$\zeta^3(\sigma)\left|\zeta^4(\sigma+it)\zeta(\sigma+2it)\right| \geq 1\,.$$

Letting $\sigma \to 1$, suppose $\zeta(\sigma+it)=0$. Then $\zeta(\sigma+2it)$ must be a pole, lest $\zeta^3(s)\zeta^4(s+it)\zeta(\sigma+2it)$ vanish there. □

THEOREM 132. *If* $\zeta(\beta+i\gamma)=0$ *then* $\beta < 1-\frac{c}{\log\gamma}$.

PROOF. The same identity shows

$$-3\frac{\zeta'}{\zeta}(\sigma)-4\Re\frac{\zeta'}{\zeta}(\sigma+it)-\Re\frac{\zeta'}{\zeta}(\sigma+2it) \geq 0\,.$$

Now $-\frac{\zeta'}{\zeta}(\sigma) \leq \frac{1}{\sigma-1}+A$ (pole!), and we know

$$-\Re\frac{\zeta'}{\zeta}(s) < A\log t-\sum_\rho \Re\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right)$$

where each summand is positive. In particular, $-\Re\frac{\zeta'}{\zeta}(\sigma+2it) \leq A\log t$. Setting $t=\gamma$ we have $s-\rho=\sigma-\beta$ so

$$-\Re\frac{\zeta'}{\zeta}(\sigma+i\gamma) \leq A\log t-\frac{\sigma-\beta}{|s-\rho|^2} = A\log t-\frac{1}{\sigma-\beta}\,.$$

It follows that

$$\frac{3}{\sigma-1}+3A+4A\log t-\frac{4}{\sigma-\beta}+A\log t \geq 0$$

so

$$\frac{4}{\sigma-\beta} \leq A\log\gamma+\frac{3}{\sigma-1}\,.$$

Take $\sigma = 1+\frac{c}{A\log\gamma}$. Then $\frac{4}{\sigma-\beta} \leq (1+\frac{3}{c})A\log\gamma$ so $1+\frac{3c}{A\log\gamma}-\beta \leq \frac{4}{(1+\frac{3}{c})A\log\gamma}$ so

$$1-\beta \leq \left(\frac{4}{1+\frac{3}{c}}-c\right)\frac{1}{A\log\gamma} = \frac{c(1-c)}{(c+3)A}\frac{1}{\log\gamma} \ll \frac{1}{\log\gamma}$$

if $0<c<1$. □

Problem Set 3

Analysis

1. (Mellin inversion)
   (a) Let $\phi(x)$ be as bounded on $(0,\infty)$, decaying at infinity faster than $x^{-\alpha}$ for some $\alpha > 0$). For $0 < \sigma < \alpha$ Show that $\tilde{\phi}(\sigma + it) = \hat{f}_\sigma(-t)$ where $f_\sigma(u) = 2\pi\phi\left(e^{2\pi u}\right)e^{2\pi\sigma u}$.

   RMK By Fourier theory, the rapid decay of f_σ shows that $t \mapsto \tilde{\phi}(\sigma + it)$ is smooth, but we already know $s \mapsto \tilde{\phi}(s)$ is holomorphic.

   (b) Suppose, in addition, that $\hat{f}_\sigma \in L^1(\mathbb{R})$, and obtain the *Mellin inversion formula*: at every $x > 0$ where ϕ is continuous, and for every $0 < \sigma < \alpha$, we have

$$\phi(x) = \frac{1}{2\pi i}\int_{\sigma - i\infty}^{\sigma + i\infty} \tilde{\phi}(s)x^s\,\mathrm{d}s\,.$$

   (c) Show that if ϕ', ϕ'' are bounded and decay polynomially at infinity then the additional hypothesis of (b) is satisfied.

   (d) Now suppose that ϕ is smooth on $[0,\infty)$ and that ϕ and all its derivatives are bounded there and decay polynomially at infinity (for example, $\phi \in C_c^\infty(0,\infty)$). Show that the meromorphic function $\tilde{\phi}$ (see PS0 problem B) *decays in vertical strips*: for every $a < b$ and $k \geq 0$ we have $\left|\tilde{\phi}(\sigma + it)\right| \leq C_{a,b}\,|t|^{-k}$ for $\sigma \in [a,b]$, $|t| \geq 1$ (if there are no poles in $[a,b] \cap \mathbb{Z}_{\leq 0}$ then we can work in the full strip $[a,b] \times \mathbb{R}$ with the bound $C_{a,b}\,(1+|t|)^{-k}$.

2. (Theorems of Phragmen–Lindelöf and Hadamard) Let $F(s)$ be holomorphic in the open strip $\{a < \sigma < b\}$, continuous in the closed strip $\{a \leq \sigma \leq b\}$.
   (a) Suppose that F grows at most polynomially in the closed strip, and that $|F(a+it)| \leq M, |F(b+it)| \leq M$. Show that $F(s) \leq M$ in the interior of the strip as well (hint: apply the maximum modulus principle to $F(s)e^{\varepsilon s^2}$).

   (b) Extend (a) to the case of F of finite order ($F(s) = O\left(e^{|s|^k}\right)$ in the strip).

   (c) (Convexity principle I) Suppose that F is of finite order in the strip, and that $|F(a+it)| \leq M_a$, $|F(b+it)| \leq M_b$. Show that if $\sigma = (1-x)a + xb$ then $|F(\sigma+it)| \leq M_a^{1-x}M_b^x$.

   (d) (Convexity principle II) Suppose that F is of finite order in the strip, and that $F(a+it) = O\left((1+|t|)^A\right)$, $F(b+it) = O\left((1+|t|)^B\right)$. Show that $F(\sigma+it) = O\left((1+|t|)^{(1-x)A+xB}\right)$.

The Riemann zetafunction

3. (Continuing the Riemann Zeta function)
   (a) (Trick) Show that the Mellin transform of $\frac{e^{-x}}{1-e^{-x}}$ agrees with $\Gamma(s)\zeta(s)$ for $\Re(s)$ large enough, and conclude that $\Gamma(s)\zeta(s)$ extends to a meromorphic function in $\mathbb{C}$ with at most simple poles at $\mathbb{Z}_{\leq 0}$, and with $\mathrm{Res}_{s=1}\left(\Gamma(s)\zeta(s)\right) = 1$.

   (b) Conclude that $\zeta(s)$ extends to a meromorphic function in $\mathbb{C}$ with a unique simple pole at $s = 1$, where $\mathrm{Res}_{s=1}\,\zeta(s) = 1$.

(c) Define the *Bernoulli numbers* by $\frac{s}{e^s-1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} s^k$. Show that $\zeta(1-k) = -\frac{B_k}{k}$ for $k \geq 1$ (in particular, find $\zeta(0)$ and $\zeta(-1)$) and conlcude that $\zeta(1-2k) = 0$ if $k \geq 1$ (the so-called "trivial zeroes").

4. In this problem we apply 2(d) to obtain bounds on the Riemann zetafunction and on Dirichlet L-functions.
   (a) ("Convexity bound") Show that $\zeta\left(\frac{1}{2}+it\right) = O_\varepsilon(t^{1/4+\varepsilon})$ as $t \to \infty$.
   (b) Generalize this to $L(\frac{1}{2}+it;\chi) = O_\varepsilon\left(q(s)^{1/4}\right)$ where $q(s) = q(|s+a|+3)$ is the analytic conductor.
   RMK The *Lindelöf Hypothesis* is the conjecture that $\zeta\left(\frac{1}{2}+it\right) = O_\varepsilon(t^\varepsilon)$. It is a consequences of the Riemann Hypothesis.

Some counting

3. (Averaging $2^{\omega(n)}$).
   (a) Show that $\frac{\zeta^2(s)}{\zeta(2s)} = \sum_{n=1}^{\infty} 2^{\omega(n)} n^{-s}$.
   (b) Show that for some A, B with $A > 0$ we have $\sum_{n \leq x} 2^{\omega(n)} = Ax\log x + Bx + O_\varepsilon\left(x^{\frac{1}{2}+\varepsilon}\right)$.
   (c) Let X be a Gaussian random variable with expectation $\log\log n$
   RMK Note that $2^{\omega(n)} \approx A\log n$ means $\omega(n) \approx \frac{\log\log n}{\log 2}$ which many standard deviations away from the typical value $\log\log n$. In other words, the sum is dominated by outliers.
   (c) Show that $\sum_{n \geq x} \frac{2^{\omega(n)}}{n} \geq C\log^2 x$ for some $C > 0$.
   (d) Find an asymptotic formula for $\frac{1}{x}\sum_{n \leq x} d_3(n)$.
   (e) Show that $\frac{1}{x}\sum_{n \leq x} d_k(n) = P_k(x) + O_\varepsilon(x^{-1/k+\varepsilon})$

4. Call $n \in \mathbb{Z}_{\geq 1}$*power-full* if $p|n \to p^2|n$. Let $\mathcal{F}$ be the set of power-full numbers.
   (a) Show that $\sum_{n \in \mathcal{F}} n^{-s} = \frac{\zeta(2s)\zeta(3s)}{\zeta(6s)}$ formally, and as holomorphic functions in $\sigma > \frac{1}{2}$.
   (b) Show that $\#\{n \in \mathcal{F} \mid n \leq x\} = Ax^{1/2} + Bx^{1/3} + O(x^{1/5})$, and express the constants A, B in terms of values of $\zeta(s)$.

5. Asymptotics for $\sigma(n)$.
6. MV sec 2.1 ex 10 (p 41)
7. Supposing $\psi(x) = x + O\left(x\exp\left(-c\sqrt{\log x}\right)\right)$, show that $\pi(x) = \mathrm{Li}(x) + O\left(x\exp\left(-c\sqrt{\log x}\right)\right)$ for the same c.
8. Let p_n denote the nth prime number. Show that $\frac{p_n}{n} = \log n + \log\log n - 1 + O\left(\frac{\log\log n}{\log n}\right)$.
9. Show $\sum_{n \leq x} \phi(n) = \frac{3}{\pi^2}x^2 + O\left(x^2\exp\left(-c\sqrt{logx}\right)\right)$.

3.3. The Prime Number Theorem in Arithmetic Progressions

Follow same scheme, using Dirichlet L-functions

(1) AC, FE, BVS
(2) ...

New features:

(1) The *conductor* q, and the *analytic conductor* $q(s) = q \cdot (|s+a|+3)$.
(2) The *root number* w.

3.3.1. Analytic continuation. From now until Section xxx we fix a primitive Dirichlet character χ mod $q > 1$. We have the Dirichlet series

$$L(s;\chi) = \sum_{n \geq 1} \chi(n) n^{-s}$$

convergent in $\Re(s) > 0$, absolutely in $\Re(s) > 1$. Define $a \in \{0,1\}$ by $\chi(-1) = (-1)^a$.

For $\varphi \in \mathcal{S}(\mathbb{R})$ of the same parity as χ ($\varphi(-x) = \chi(-1)\varphi(x)$) set

$$F(\chi;\varphi;r) = \sum_{n \in \mathbb{Z}} \chi(n)\varphi(rn)\,.$$

LEMMA 133 (Properties of F). *(1) $F(r) = F(\chi;\varphi;r)$ decays rapidly as $r \to \infty$.*
(2) $F(\chi;\varphi;r) = \frac{G(\chi)}{rq} F(\bar{\chi};\hat{\varphi};\frac{1}{rq})$.
(3) $F(r) \to 0$ rapidly as $r \to 0$.

PROOF. $|F(r)| \leq \sum_{n \neq 0} |\varphi(rn)|$. The second claim is Poisson sum (see PS2), and the third follows from the second. $\square$

DEFINITION 134. Let

$$Z(\chi;\varphi;s) = \int_0^\infty F(\chi;\varphi;r) r^s \frac{\mathrm{d}r}{r}\,.$$

This converges absolutely for $\Re(s) > 0$. For $\Re(s) > 1$ we can change the order of summation and integration and get:

$$Z(\chi;\varphi;s) = 2\sum_{n=1}^{\infty} \chi(n) n^{-s} \tilde{\varphi}(s) = 2L(s;\chi)\tilde{\varphi}(s)\,.$$

We now break the integral in two:

$$\begin{aligned} Z(\chi;\varphi;s) &= \int_{\sqrt{q}}^\infty F(\chi;\varphi;r) r^s \frac{\mathrm{d}r}{r} + \int_0^{\sqrt{q}} F(\chi;\varphi;r) r^s \frac{\mathrm{d}r}{r} \\ &= \int_{\sqrt{q}}^\infty F(\chi;\varphi;r) r^s \frac{\mathrm{d}r}{r} + \frac{G(\chi)}{q} \int_0^{\sqrt{q}} F(\bar{\chi};\hat{\varphi};\frac{1}{rq}) r^{s-1} \frac{\mathrm{d}r}{r} \\ &= \int_{\sqrt{q}}^\infty F(\chi;\varphi;r) r^s \frac{\mathrm{d}r}{r} + \frac{G(\chi)}{\sqrt{q}} q^{\frac{1}{2}-s} \int_{\sqrt{q}}^\infty F(\bar{\chi};\hat{\varphi};r) r^{1-s} \frac{\mathrm{d}r}{r}\,. \end{aligned}$$

We have shown:

$$q^{s/2} Z(\chi;\varphi;s) = q^{s/2} \int_{\sqrt{q}}^\infty F(\chi;\varphi;r) r^s \frac{\mathrm{d}r}{r} + \frac{G(\chi)}{\sqrt{q}} q^{\frac{1-s}{2}} \int_{\sqrt{q}}^\infty F(\bar{\chi};\hat{\varphi};r) r^{1-s} \frac{\mathrm{d}r}{r}\,.$$

COROLLARY 135. *$Z(\chi;\varphi;s)$ extends to an entire function.*

Next, note that $\left|\frac{G(\chi)}{\sqrt{q}}\right| = 1$ and $F(\chi;\hat{\varphi};r) = \chi(-1)F(\chi;\varphi;r)$. Thus:

$$q^{\frac{1-s}{2}}Z(\bar{\chi};\hat{\varphi};1-s) = q^{\frac{1-s}{2}}\int_{\sqrt{q}}^{\infty}F(\bar{\chi};\hat{\varphi};r)r^{1-s}\frac{\mathrm{d}r}{r} + \frac{\chi(-1)G(\bar{\chi})}{\sqrt{q}}q^{\frac{s}{2}}\int_{\sqrt{q}}^{\infty}F(\chi;\varphi;r)r^{s}\frac{\mathrm{d}r}{r},$$

and applying $G(\chi)G(\bar{\chi}) = q\chi(-1)$ gives:

$$q^{s/2}Z(\chi;\varphi;s) = \frac{G(\chi)}{\sqrt{q}}q^{\frac{1-s}{2}}Z(\bar{\chi};\hat{\varphi};1-s).$$

COROLLARY 136 (Non-symmetric FE).

$$L(s;\chi) = G(\chi)q^{-s}\frac{\tilde{\hat{\varphi}}(1-s)}{\tilde{\varphi}(s)}L(1-s;\bar{\chi}).$$

We now make a specific choice: $\varphi_a(x) = x^a e^{-\pi x^2}$. For $a=0$ we have $\hat{\varphi}_a(k) = \varphi_a(k)$. For $a=1$ we have $\hat{\varphi}_a(k) = -i\varphi_a(k)$. Also, $\tilde{\varphi}_a(s) = \tilde{\varphi}_0(s+a) = \Gamma_{\mathbb{R}}(s+a)$ is nowhere zero and $\tilde{\hat{\varphi}}_a(s) = (-i)^a\tilde{\varphi}_a(s)$. We conclude:

THEOREM 137. *Let $\Lambda(s;\chi) = q^{s/2}\Gamma_{\mathbb{R}}(s+a)L(s;\chi)$. Then $\Lambda(s;\chi)$ extends to an entire function, and satisfies the functional equation*

$$\Lambda(s;\chi) = w\Lambda(1-s;\bar{\chi})$$

with the root number $w = w(\chi) = \frac{G(\chi)(-i)^a}{\sqrt{q}}$. *Since $q^{s/2}\Gamma_{\mathbb{R}}(s+a)$ is nowhere zero, $L(s;\chi)$ extends to an entire function. This has "trivial" zeroes at $\{a-2n \mid n \geq 1\}$.*

Note that by the absolute convergence of the Euler product, $L(s;\chi)$ hence $\Lambda(s;\chi)$ has no zeroes in $\Re(s) > 1$ hence in $\Re(s) < 0$.

3.3.2. The Hadamard product. In the right half-plane $\Re(s) > \varepsilon$ we have $L(s;\chi)$ bounded, and $q^{s/2}, \Gamma_{\mathbb{R}}(s+a)$ of order 1 (Stirling). Applying the FE we see that $\Lambda(s;\chi)$ is of order 1, and therefore has the expansion

$$\Lambda(s;\chi) = e^{A+B(\chi)s}\prod_{\rho}\left(1-\frac{s}{\rho}\right)e^{s/\rho}.$$

Taking the logarithmic derivative, we find:

$$\frac{\Lambda'}{\Lambda}(s;\chi) = B(\chi) + \sum_{\rho}\left[\frac{1}{s-\rho}+\frac{1}{\rho}\right].$$

Now $\Re B(\chi)$ will contribute to the

$$B(\chi) = \frac{\Lambda'}{\Lambda}(0;\chi) = -\frac{\Lambda'}{\Lambda}(1;\bar{\chi}) = -B(\bar{\chi}) - \sum_{\rho}\left[\frac{1}{1-\bar{\rho}}+\frac{1}{\bar{\rho}}\right].$$

Thus

$$2\Re B(\chi) = -\sum_{\rho}\left[\frac{1}{\rho}+\frac{1}{\bar{\rho}}\right] = -2\sum_{\rho}\Re\frac{1}{\rho} < 0.$$

Finally, we note that

$$-\frac{L'(s;\chi)}{L(s;\chi)} = \frac{1}{2}\log\frac{q}{\pi} + \frac{1}{2}\frac{\Gamma'(\frac{s+a}{2})}{\Gamma(\frac{s+a}{2})} - B(\chi) - \sum_{\rho}\left[\frac{1}{s-\rho} + \frac{1}{\rho}\right]. \tag{3.3.1}$$

and that $\frac{1}{2}\log\frac{q}{\pi} + \frac{1}{2}\frac{\Gamma'(\frac{s+a}{2})}{\Gamma(\frac{s+a}{2})} \approx \log q(s)$.

3.3.3. The zero-free region. Note that

$$-\Re\frac{L'(s;\chi)}{L(s;\chi)} = O(\log q(s)) - \Re B(\chi) - \Re\sum_{\rho}\left[\frac{1}{s-\rho} + \frac{1}{\rho}\right] = O(\log q(s)) - \sum_{\rho}\frac{\sigma-\beta}{|\sigma-\beta|^2 + |\gamma-t|^2}.$$

In particular, if $\sigma > 1$ then for any single zero ρ,

$$-\Re\frac{L'(s;\chi)}{L(s;\chi)} \le O(\log q(s)) - \frac{\sigma-\beta}{|\sigma-\beta|^2 + |\gamma-t|^2}$$

From the Euler product we have for $s = \sigma + it$ with $\sigma > 1$ that

$$-\frac{L'(s;\chi)}{L(s;\chi)} = \sum_n \chi(n)\Lambda(n)n^{-s} = \sum_n \frac{\Lambda(n)}{n^{\sigma}}\left(\chi(n)n^{-it}\right).$$

Applying Mertens's identity again we get:

$$-3\frac{L'}{L}(\sigma,\chi_0) - 4\Re\frac{L'}{L}(\sigma+it,\chi) - \Re\frac{L'}{L}(\sigma+2it,\chi^2) \ge 0. \tag{3.3.2}$$

Note that χ_0 isn't and χ^2 need not be primitive, and that we may have $\chi^2 = \chi_0$. We first note that if ψ is a Dirichlet character mod q, ψ_1 its primitive counterpart then at $\sigma > 1$ their logarithmic derivatives differ by at most

$$\sum_{p|q}\frac{\log p p^{-\sigma}}{1-p^{-\sigma}} \le \sum_{p|q}\log p \le \log q.$$

It follows that our estimate

$$-\Re\frac{L'}{L}(s;\psi_1) \le \Re\frac{\delta_{\psi_1}}{s-1} + O(\log q_1(s))$$

also gives

$$-\Re\frac{L'}{L}(s;\psi_1) \le \Re\frac{\delta_{\psi}}{s-1} + O(\log q(s)).$$

Applying this in (3.3.2) gives with $s = \sigma + i\gamma$, for the zero $\rho = \beta + i\gamma$ gives:

$$3\Re\frac{1}{s-1} - \frac{4}{\sigma-\beta} + \Re\frac{\delta_{\chi^2}}{\sigma+2it-1} + O(\mathscr{L}) \ge 0$$

with $\mathscr{L} = \log q(\gamma)$. Thus:

$$\frac{4}{\sigma-\beta} \le \frac{3}{\sigma-1} + \Re\frac{\delta_{\chi^2}}{\sigma+2it-1} + C\mathscr{L}.$$

Case 1. If χ^2 is non-principal ("complex"), take $\sigma = 1 + \frac{\delta}{\mathscr{L}}$ and get

$$1 - \beta + \frac{\delta}{\mathscr{L}} \geq \frac{4}{\frac{3}{\delta} + C} \frac{1}{\mathscr{L}}$$

so

$$1 - \beta \geq \left(\frac{4\delta}{3 + C\delta} - \delta \right) \frac{1}{\mathscr{L}} \gg \frac{1}{\mathscr{L}}.$$

Case 2. If $\chi^2 = \chi_0$, suppose $\gamma \geq \frac{\delta}{\mathscr{L}}$ and $\sigma = 1 + \frac{\delta}{\mathscr{L}}$. Then $\Re \frac{1}{\sigma + 2it - 1} = \frac{\sigma - 1}{|\sigma - 1|^2 + 4t^2} \leq \frac{\mathscr{L}}{5\delta}$. Then

$$\frac{4}{\sigma - \beta} \leq \frac{3\mathscr{L}}{\delta} + \frac{\mathscr{L}}{5\delta} + C\mathscr{L}$$

so

$$\beta < 1 - \frac{4 - 5C\delta}{16 + 5C\delta} \frac{\delta}{\mathscr{L}}.$$

In other words, we have our zero-free region for $\gamma > \frac{\delta}{\log q}$.

Now suppose χ is real. We need to study small zeroes. For this recall

$$-\frac{L'}{L}(\sigma;\chi) = O(\log q) - \sum_{\rho} \frac{1}{\sigma - \rho}.$$

Now $-\frac{L'}{L}(\sigma;\chi) \geq -\sum_n \Lambda(n) n^{-\sigma} = \frac{\zeta'}{\zeta}(\sigma) = -\frac{1}{\sigma - 1} - O(1)$. Suppose two complex zeroes. Then

$$-\frac{1}{\sigma - 1} - O(1) \leq O(\log q) + \frac{2(\sigma - \beta)}{(\sigma - \beta)^2 + \gamma^2}.$$

For $\sigma = 1 + \frac{2\delta}{\log q}$ get $|\gamma| < \frac{1}{2}(\sigma - \beta)$ so

$$-\frac{1}{\sigma - 1} = O(\log q) - \frac{8}{5(\sigma - \beta)}$$

and if δ is small enough get $\beta < 1 - \frac{\delta}{\log q}$. Same if two real zeroes.

THEOREM 138. *There exists C such that if $0 < \delta < C$ the only possible zero for $L(s;\chi)$ with $|\gamma| < \frac{\delta}{\log q}$ and $\beta > 1 - \frac{\delta}{\log q}$ is a single real zero, and this only if χ is real. In any case all zeroes with $|\gamma| \geq \frac{\delta}{\log q}$ satisfy $1 - \beta \gg \frac{1}{\log q(\gamma)}$.*

REMARK 139. Note that if χ is imprimitive, coming from primitive χ_1 then $L(s;\chi)$ and $L(s;\chi_1)$ have same zeroes except for zeroes of Euler factors $(1 - \chi(p)p^{-s})$ for $p|q$, and these are all on the line $\Re(s) = 0$, and we still obtain the conclusion of the Theorem.

REMARK 140. (Landau) Let χ_1, χ_2 be two quadratic characters. Then the Euler product $\zeta(s)L(s;\chi_1)L(s;\chi_2)L(s;\chi_1\chi_2)$ has positive coefficients. From this can deduce that Siegel zeroes are rare: at most one character mod q can have then, and the sequence of moduli supporting such characters must satisfy $q_{n+1} \geq q_n^2$.

3.3.4. Counting zeroes. We return to the formula

$$-\Re\frac{L'(s;\chi)}{L(s;\chi)} = O(\log q(s)) - \sum_\rho \frac{\sigma-\beta}{|\sigma-\beta|^2+|\gamma-t|^2}\,.$$

Applying this with $\sigma=2$, where $\left|\frac{L'(s;\chi)}{L(s;\chi)}\right| \le \left|\frac{\zeta'(2)}{\zeta(2)}\right|$, and using $\sigma-\beta\ge 1$, we get:

$$\#\{\gamma \mid |\gamma-T|\le 1\} = O(\log qT) \tag{3.3.3}$$

and

$$\sum_{|\gamma-T|>1} \frac{1}{(T-\gamma)^2} = O(\log qT)\,. \tag{3.3.4}$$

LEMMA 141. *Let $s=\sigma+iT$, $\sigma\in[-1,2]$. Then*

$$\frac{L'}{L}(s;\chi) = \sum_{\gamma\in(T-1,T+1)} \frac{1}{s-\rho} + O(\log qT)\,.$$

PROOF. Subtracting (3.3.1) evaluated at $s=\sigma+iT, 2+iT$ we get

$$\frac{L'}{L}(s;\chi) = O(1) + \sum_\rho \left(\frac{1}{s-\rho} - \frac{1}{2+iT-\rho}\right),$$

since $F\left(\frac{\sigma+iT}{2}\right) - F\left(\frac{2+iT}{2}\right) = O(1)$. Now for ρ with $\gamma\notin(T-1,T+1)$, we have

$$\left|\frac{1}{s-\rho} - \frac{1}{2+iT-\rho}\right| \le \frac{2-\sigma}{|s-\rho|\,|2+iT-\rho|} \le \frac{3}{|\gamma-T|^2}$$

and for ρ with $\gamma\in(t-1,t+1)$ we have $\left|\frac{1}{2+iT-\rho}\right| \le \frac{1}{|2-\beta|} \le 1$. □

COROLLARY 142. *For each $T>2$ there exists $T'\in[T,T+1]$ such that for $s=\sigma+iT'$, $\sigma\in[-1,2]$ we have*

$$\frac{L'}{L}(s;\chi) = O(\log^2 qT')\,.$$

PROOF. Same as Corollary 117. □

DEFINITION 143. $N_\chi(T)$ counts zeroes of $L(s;\chi)$ up to height T.

THEOREM 144. $N_\chi(T) = \frac{T}{2\pi}\log\frac{qT}{2\pi} - \frac{T}{2\pi} + O(\log qT)$.

PROOF. Suppose T is not the ordinate of any zero, and let R be the rectangle $[-1,2]\times[-T,T]$. We need to calculate the real number

$$2N_\chi(T) = \frac{1}{2\pi i}\oint_{\partial R} \frac{\Lambda'}{\Lambda}(s;\chi)\,\mathrm{d}s\,.$$

Since $\overline{\Lambda(\bar{s};\chi)} = \Lambda(s;\bar\chi)$ and by the functional equation $\Lambda(1-s;\chi) = w(\bar\chi)\Lambda(s;\bar\chi)$, it is enough to consider the quarter-rectangle $2\to 2+iT\to \frac{1}{2}+iT$. Recall that $\Lambda(s;\chi) = q^{s/2}\pi^{-s/2}\Gamma\left(\frac{s+a}{2}\right)L(s;\chi)$. The argument of $\left(\frac{q}{\pi}\right)^{s/2}$ changes exactly by $\frac{1}{2}T\log\frac{q}{\pi}$. The argument of $\Gamma\left(\frac{s+a}{2}\right)$ changes by $\Im\log\Gamma\left(\frac{1+2a}{4}+\frac{1}{2}iT\right) = \frac{T}{2}\log\left(\frac{T}{2}\right) - \frac{T}{2} - \frac{\pi}{8} + \frac{\pi a}{4} + O(T^{-1})$. It remains to estimate the change $S(T)$

in $\arg\zeta(s)$. Since $\Re(L(2+it;\chi)) \geq 1-\sum_{n=2}^{\infty}\frac{1}{n^2} > 0$, the change of argument in $[2, 2+iT]$ is at most π. On $\left[\frac{1}{2}+iT, 2+iT\right]$ Lemma 141 gives:

$$\frac{L'}{L}(s;\chi) = \sum_{\gamma\in(t-1,t+1)} (\log(s-\rho))' + O(\log qT)\,.$$

Now the change of the argument of each $s-\rho$ on the interval is at most $\frac{\pi}{2}$, so the total change in the argument of $\zeta(s)$ is $O(\log qT)$. In summary, we have:

$$2\frac{1}{4}2\pi N_\chi(T) = \frac{T}{2}\log\left(\frac{T}{2}\right) + \frac{T}{2}\log q - \frac{T}{2}\log\pi - \frac{T}{2} + O(\log qT)\,.$$

□

Now get

$$N_\chi(T) = \frac{T}{2\pi}\log\left(\frac{qT}{2\pi}\right) - \frac{T}{2\pi} + O(\log qT)\,.$$

3.3.5. The explicit formula for $L(s;\chi)$.

LEMMA 145. *For $\sigma \leq -1$ we have $\frac{L'}{L}(s;\chi) = O(\log q|s|)$.*

PROOF. By the duplication formula,

$$L(1-s;\chi) = w(\chi)2^{1-s}\pi^{-s}q^{s-\frac{1}{2}}\cos\left(\frac{\pi(s-a)}{2}\right)\Gamma(s)L(s;\bar\chi)$$

and hence

$$\frac{L'}{L}(1-s;\chi) = \log q - \frac{1}{2}\pi\tan\left(\frac{\pi(s+a)}{2}\right) + \frac{\Gamma'(s)}{\Gamma(s)} + \frac{L'}{L}(s;\bar\chi)\,.$$

Now if $\sigma \geq 2$ the last term is $O(1)$, the third term is $O(\log|s|) = O(\log|1-s|)$ and if $1-s$ is away from the trivial zeroes, then the second term is $O(1)$ as well. □

PROPOSITION 146. *Let $U \geq 1$ not be an even integer. Then*

$$\sum_{n\leq x}\chi(n)\Lambda(n) + O(\frac{x\log x}{H}) = -\sum_\rho \tilde\eta_H(\rho)\frac{x^\rho}{\rho} + (1-a)\log x + b(\chi) - \frac{1}{2\pi i}\int_{(-1)}\frac{L'}{L}(s;\chi)\tilde\eta_H(s)x^s\frac{\mathrm{d}s}{s}\,,$$

where $b(\chi)$ is the zeroes order term in the Laurent expansion of $-\frac{L'}{L}(s;\chi)$ at $s=0$.

PROOF. More-or-less as before: we have

$$\sum_{n\leq x}\chi(n)\Lambda(n) + O(\frac{x\log x}{H}) = \sum_n \chi(n)\Lambda(n)\varphi_H\left(\frac{n}{x}\right) = \frac{1}{2\pi i}\int_{(2)}\frac{L'}{L}(s;\chi)\tilde\eta_H(s)x^s\frac{\mathrm{d}s}{s}\,.$$

We now shift the contour to (-1), acquiring contributions from the poles of $\frac{1}{s}\frac{L'}{L}(s;\chi)$. These occur at the zeroes of $L(s;\chi)$ (which itself has no poles), accounting for the terms $\tilde\eta_H(\rho)\frac{x^\rho}{\rho}$, and at $s=0$. To understand the contribution of $s=0$ we go back to the logarithmic derivative (3.3.1). If $a=1$ this is regular at $s=0$ and $b(\chi) = -\frac{L'}{L}(0;\chi)$ (note that $\tilde\eta_H(0)=1$). If $a=0$, however, $\Gamma(\frac{s}{2})$ has pole at $s=0$, and so does its logarithmic derivative, at which point the integrand has a double pole. In that case near $s=0$,

$$-\frac{L'}{L}(s;\chi) = \frac{1}{s} + b(\chi) + O(s); \qquad \frac{x^s}{s} = \frac{1}{s} + \log x + O(s)\,.$$

Now,
$$\frac{\mathrm{d}}{\mathrm{d}s}\tilde{\eta}_H(s)=\frac{\mathrm{d}}{\mathrm{d}s}\int_{-\infty}^{+\infty}\eta(u)\exp\left\{\frac{us}{H}\right\}\mathrm{d}u=\frac{1}{H}\int_{-\infty}^{+\infty}u\eta(u)\exp\left\{\frac{us}{H}\right\}\mathrm{d}u\,.$$
In partiular, choosing since η symmetric we see that $\tilde{\eta}_H'(0)=0$ so that $\tilde{\eta}_H(s)=1+O(s^2)$ and the residue of the integrand is $\log x+b(\chi)$. $\square$

We need to estimate $b(\chi)$.

LEMMA 147. *We have*
$$b(\chi)=O(\log q)+\sum_{|\gamma|<1}\frac{1}{\rho}\,.$$

PROOF. Subtract (3.3.1) at $s,2$ and use $-\frac{L'}{L}(2;\chi)=O(1)$ and $\frac{1}{2}\digamma\left(\frac{s+a}{2}\right)=\frac{1-a}{s}+O(1)$ ($O(1)$ absolute) to get
$$-\frac{L'}{L}(s;\chi)=\frac{1-a}{s}+O(1)-\sum_{\rho}\left(\frac{1}{s-\rho}-\frac{1}{2-\rho}\right),$$
with the $O(1)$ absolute. Now $\frac{1}{\rho}+\frac{1}{2-\rho}=\frac{2}{\rho(2-\rho)}$. In particular, $\left|\sum_{|\gamma|\geq 1}\frac{1}{\rho}+\frac{1}{2-\rho}\right|\ll\sum_{|\gamma|>1}\frac{1}{|\gamma|^2}\ll\log q$ by (3.3.4). If $|\gamma|<1$ then $\left|\frac{1}{2-\rho}\right|=O(1)$ so $\sum_{|\gamma|<1}\frac{1}{2-\rho}=O(\log q)$ by (3.3.3). $\square$

PROPOSITION 148. *Let $\beta(T)$ be such that if $|\gamma|<T$ then $\beta\leq\beta(T)$, except possibly for the single real zero β_0. We then have*
$$\psi(x;\chi)=\sum_{n\leq x}\chi(n)\Lambda(n)\ll\psi(x;\chi)\ll-\frac{x^{\beta_0}}{\beta_0}+\left[x^{1/4}\log x+\log^2 qT\cdot x^{\beta(T)}+\frac{x\log x}{H}+\frac{xH\log(qT)}{T}+\frac{H\log q}{x}\right].$$

PROOF. The integral in the last Proposition satisfies:
$$-\frac{1}{2\pi i}\int_{(-1)}\frac{L'}{L}(s;\chi)\tilde{\eta}_H(s)x^s\frac{\mathrm{d}s}{s}\ll x^{-1}H\exp\left\{\frac{1}{H}\right\}\int_{-\infty}^{+\infty}\frac{\log\left(q|1+it|\right)}{|1+it|^2}\,\mathrm{d}t\,.$$

Since the zero density is about $\log qt$ at height t, and since $\tilde{\eta}_H(\rho)\ll 1$, we can bound $\sum_{1<|\gamma|<T}\tilde{\eta}_H(\rho)\frac{x^\rho}{\rho}$ by
$$x^{\beta(T)}\int_1^T\frac{\log qt}{t}\,\mathrm{d}t\leq x^{\beta(T)}\int_1^{qT}\frac{\log t}{t}\,\mathrm{d}t\ll\log^2(qT)x^{\beta(T)}\,.$$
Similarly, $\left|\sum_{|\gamma|>T}\tilde{\eta}_H(\rho)\frac{x^\rho}{\rho}\right|$ is bounded by
$$x\int_T^\infty\frac{H}{t}\frac{\log(qt)\,\mathrm{d}t}{t}\ll\frac{xH\log(qT)}{T}\,.$$

In summary so far, we have
$$\psi(x;\chi)=\sum_{|\gamma|<1}\left(\frac{1}{\rho}-\frac{x^\rho}{\rho}\right)+(1-a)\log x+O(\log q)+O(\log^2 qT\cdot x^{\beta(T)}+\frac{x\log x}{H}+\frac{xH\log(qT)}{T}+\frac{H\log q}{x}).$$

Now zeroes with $\beta<1-\frac{c}{\log q}$ also have $\beta>\frac{c}{\log q}$ by the functional equation, except (if χ is real) for a single pair of zeroes $\beta_0,1-\beta_0$, where we have $\beta_0>\frac{3}{4}$ since can take c small and $q\geq 3$. Thus $\frac{1}{\beta_0}$ is $O(1)$. The sum over $\frac{x^\rho}{\rho}$ is $O(\log^2 q)x^{\beta(T)}$, so absorbed in the existing error terms. Also, $\frac{1-x^{1-\beta_0}}{1-\beta_0}=x^\sigma\log x$ for $0<\sigma<1-\beta_0<\frac{1}{4}$ and we get the claim. $\square$

THEOREM 149. *For* $\log q \ll (\log x)^{1/2}$

$$\psi(x;\chi) = -\frac{x^{\beta_0}}{\beta_0} + O(x\exp\left\{-c'\sqrt{\log x}\right\}).$$

On RH we have for $q \le x$ *that*

$$\psi(x;\chi) \ll \sqrt{x}\log^2 x.$$

PROOF. On RH we have the bound

$$x^{1/4}\log x + \sqrt{x}\log^2 qT + \frac{x\log x}{H} + \frac{xH\log(qT)}{T} + \frac{H\log q}{x}.$$

Take $T = \sqrt{x}$, $H = x$. Then for $x \ge q$, the error term is

With zero-free region get bound

$$x\log^2 qT\exp\left\{-c\frac{\log x}{\log qT}\right\} + \frac{x\log x}{H} + \frac{xH\log qT}{T} + \frac{H\log q}{x}.$$

Taking $T = \exp\left\{c_1(\log x)^{1/2}\right\}$, $H = \exp\left\{c_2(\log x)^{1/2}\right\}$ with $c_1 > c_2$ works if $\log q \le C(\log x)^{1/2}$.

Finally, we note that if χ is a (possibly non-primitive) character mod q, with primitive associate χ_1 mod q_1. Then

$$\psi(x;\chi_1) - \psi(x;\chi) = \sum_{\substack{p^m \le x \\ p\mid q \\ p\nmid q_1}} \chi(p^m)\log p \ll \sum_{p\mid q}\log p \sum_{p^m \le x} 1 \ll \log q\log x,$$

which can be absorbed in our error terms in either case. Thus we may apply the theorem for non-primitive characters as well. □

3.3.6. The PNT in APs. Averaging Theorem (149) over the group of characters, we find:

$$\sum_{x \ge p^m \equiv a(q)} \log p = \frac{1}{\varphi(q)}\sum_{n\le x}\bar{\chi}(a)\psi(x;\chi) = \frac{x}{\varphi(q)} + \text{error}$$

where on the RH the error is $O(x^{1/2}\log^2 x)$ for $q \le x$, and $-\frac{\chi(a)x^{-\beta_0}}{\varphi(q)\beta_0} + O\left(x\exp\left\{-C\sqrt{\log x}\right\}\right)$ unconditionally, if $\log q \ll (\log x)^{1/2}$.

LEMMA 150. $\beta_0 \le 1 - \frac{c}{q^{1/2}\log^2 q}$.

PROOF. Using $h(d) \ge 1$ in the class number formula we get $L(1;\chi) \gg q^{-1/2}$. Now for $1 - \frac{c}{\log q} \le \sigma \le 1$ we have $n^{-\sigma} \le \frac{1}{n}\exp\left\{\frac{c\log n}{\log q}\right\}$. Thus

$$\sum_{n=1}^{q}\frac{\chi(n)\log n}{n^\sigma} \ll \sum_{n=1}^{q}\frac{\log n}{n} \ll \log^2 q.$$

Also, partial summation gives

$$\sum_{q+1}^{\infty}\frac{\chi(n)\log n}{n^\sigma} \ll \sum_{q+1}^{\infty}|S(n)|\left[\frac{\log n}{n^\sigma} - \frac{\log(n+1)}{(n+1)^\sigma}\right] \ll q^{1/2}\frac{\log q}{q^\sigma} \ll \log q.$$

It follows that $L'(\sigma;\chi) \ll \log^2 q$ for $\beta \le \sigma \le 1$. Then

$$q^{-1/2} \ll L(1;\chi) = L(1;\chi) - L(\beta;\chi) \ll (1-\beta)\log^2 q\,.$$

$\square$

THEOREM 151 (PNT I). *For $q \ll (\log x)^{1-\delta}$ we have*

$$\sum_{x \ge p \equiv a(q)} \log p = \frac{x}{\varphi(q)} + O\left(x\exp\left\{-C\sqrt{\log x}\right\}\right)\,.$$

PROOF. $x^{\beta_0 - 1} \le \exp\left\{-\frac{\log x}{q^{1/2}\log q}\right\}$. $\square$

COROLLARY 152. *The first prime in an AP occurs before* $\exp\left(q^{1+\delta}\right)$.

Note that RH predicts $q^{2+\delta}$ and probably $q^{1+\delta}$ is enough.

THEOREM 153 (Siegel 1935). $L(1;\chi) \ge C(\varepsilon)q^{-\varepsilon}$ *for some ineffective constant.*

COROLLARY 154. *Any exceptional zero has $\beta \le 1 - \frac{c(\varepsilon)}{q^{\varepsilon}}$, and the error term holds for $q \ll (\log x)^A$ for A arbitrarily large. The first prime in an AP occurs before* $\exp(q^{\varepsilon})$.

3.3.7. Statement of Bombieri–Vinogradov.

DEFINITION 155. Let $\psi(x;q,a) = \sum_{x \ge n \equiv a\,(q)} \Lambda(n)$.

We expect this to be about $\frac{x}{\varphi(q)} + O(\sqrt{x}\log^2 x)$.

THEOREM 156 (Bombieri, Vinogradov 1965 [**1, 8**]). *Given $A > 0$ and for $x^{1/2}(\log x)^{-A} \le Q \le x^{1/2}$ we have*

$$\frac{1}{Q}\sum_{q \le Q} \max_{(a,q)=1} \max_{y \le x} \left|\psi(y;q,a) - \frac{y}{\varphi(q)}\right| \ll x^{1/2}(\log x)^5\,.$$

CONJECTURE 157 (Elliott–Halberstam). *Can take $Q \le x^{\theta}$ for $\theta < 1$.*

THEOREM 158 (Zhang 2013). *Can take $Q \le x^{\theta}$ for some $\theta > \frac{1}{2}$ if restrict q to be sufficiently smooth.*