Analytic Number Theory, II : a second course.

TDW Spring 2021.

§1. Introduction.

The principal conclusions of a first course in analytic number theory are the Prime Number Theorem

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

and the corresponding Prime Number Theorem in arithmetic progressions, asserting that when $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a,q)=1$, then

$$\pi(x;q,a) := \sum_{\substack{p \le x \\ p \equiv a \pmod q}} 1 \sim \frac{x}{\varphi(q)\log x}.$$

Even if one makes use of the strongest unconditional error terms available, in the form

$$\psi(x) := \sum_{n \le x} \Lambda(n) = x + O\left(x \exp\left(-c\sqrt{\log x}\right)\right), \quad \text{suitable } c>0,$$

and the Siegel-Walfisz theorem

$$\psi(x;q,a) := \sum_{\substack{n \le x \\ n \equiv a \pmod q}} \Lambda(n) = \frac{x}{\varphi(q)} + O_A\left(x \exp\left(-c\sqrt{\log x}\right)\right),$$

for $q \le (\log x)^A$, any $A>0$ and $c = c_A > 0$,

one has little control of the more subtle features of the distribution of prime numbers.

Two problems serve as motivation:

Problem 1: Show that, when $1 > \theta > 0$, one has

$$\pi(x + x^\theta) - \pi(x) \sim \frac{x^\theta}{\log x} \quad (\text{primes in short intervals}).$$

Problem 2: Show that, when $1 > \theta > 0$ and $q \le x^\theta$, then

$$\pi(x;q,a) \sim \frac{x}{\varphi(q)\log x} \quad \text{whenever } (a,q)=1.$$

It is known that the asymptotic formula in Problem 1 holds when $\theta > 7/12$ (Huxley, 1972), and subject to the validity of the Riemann Hypothesis, it holds for $\theta > 1/2$. However, a pedestrian application of the approach from a first course fails to access any value of θ smaller than 1.

Likewise, in Problem 2, the Siegel - Walfisz theorem falls short of accessing parameters θ with $\theta > 0$, and even the Generalised Riemann Hypothesis would show only that the asserted asymptotic formula holds when $\theta > 1/2$.

By suitable averaging arguments, it is possible to draw conclusions in Problems 1 and 2 approaching what can be said if the Riemann Hypothesis (or GRH) were to hold. In this course we explore such averaging arguments and the associated theoretical machinery. Our first objective is the celebrated Bombieri - Vinogradov theorem.

<u>Theorem 1.1.</u> (Bombieri - Vinogradov theorem) Let $A > 0$ be fixed. Then whenever $x^{1/2}(\log x)^{-A} \le Q \le x^{1/2}$, one has

$$\sum_{1 \le q \le Q} \sup_{1 \le y \le x} \max_{\substack{1 \le a \le q \\ (a,q)=1}} \left| \psi(y; q, a) - \frac{y}{\varphi(q)} \right| \ll Q x^{1/2} (\log x)^3 .$$

Thus, as one averages over moduli q no larger than Q, one sees that almost always one has

$$\psi(x; q, a) - \frac{x}{\varphi(q)} \ll x^{\frac{1}{2}+\varepsilon}.$$

Thus, for most such q and all a with $(a,q)=1$, one has

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

a conclusion essentially as strong as can be extracted from GRH — "in average".

§2. The Large Sieve Inequality, I: trigonometric polynomials.

The large sieve was conceived as an arithmetic sieve (think: Eratosthenes sieve for prime numbers), evolved into an analytic statement concerning exponential sums, and nowadays is regarded as a mean square bound for discrete sums of oscillating arithmetic coefficients such as character sums or averages of cusp form coefficients. We begin with exponential sums (or "trigonometric polynomials").

The central object of study are exponential sums of the shape

$$T(x) = \sum_{n=M+1}^{M+N} c_n\, e(nx) \qquad (c_n \in \mathbb{C}). \quad \text{— (2.1)}$$

We consider a sequence of real points (x_r) having the property that for some $\delta > 0$, the points x_r are δ-spaced modulo 1, so that

$$\|x_r - x_s\| \geqslant \delta \quad \text{for} \quad 1 \leqslant r < s \leqslant R.$$

Here, as usual, we write $\|\theta\| := \min_{y\in\mathbb{Z}} |\theta - y|$. Our goal is to obtain a discrete analogue of the orthogonality relation

$$\int_0^1 |T(x)|^2\,dx = \int_0^1 \sum_{n_1=M+1}^{M+N} \sum_{n_2=M+1}^{M+N} c_{n_1}\overline{c_{n_2}}\, e(x(n_1-n_2))\,dx$$

$$= \sum_{M+1\le n_1,n_2\le M+N} c_{n_1}\overline{c_{n_2}} \underbrace{\int_0^1 e(x(n_1-n_2))\,dx}_{0 \text{ unless } n_1=n_2}$$

$$= \sum_{n=M+1}^{M+N} |c_n|^2 .$$

Goal: Determine the smallest value of $\Delta = \Delta(N,\delta)$ such that, uniformly in $\underline{c}$, M, R one has

$$\sum_{r=1}^{R} |T(x_r)|^2 \ll \Delta \sum_{n=M+1}^{M+N} |c_n|^2.$$

Elementary considerations One must have $\Delta \gg \max\{N, 1/\delta - 1\}$.

To see this, observe first that when $R=1$, one has

$$\sum_{r=1}^{R}|T(x_r)|^2 = |T(x_1)|^2 = \left|\sum_{n=M+1}^{M+N} c_n e(nx_1)\right|^2 = \left(\sum_{n=M+1}^{M+N} 1\right)^2 = N\sum_{n=M+1}^{M+N}|c_n|^2$$

in the particular case $c_n = e(-nx_1)$.

Thus $\Delta \geqslant N$ for certain $\underline{c}$, R.

Also, by orthogonality,

$$\int_0^1 \sum_{r=1}^{R}\left|T\left(x+\frac{r}{R}\right)\right|^2 dx = R\int_0^1 |T(x)|^2\,dx = R\sum_{n=M+1}^{M+N}|c_n|^2,$$

so by the Mean Value Theorem, there exists a value of x for which

$$\sum_{r=1}^{R}\left|T\left(x+\frac{r}{R}\right)\right|^2 \geqslant R\sum_{n=M+1}^{M+N}|c_n|^2 .$$

By choosing $R = \lfloor 1/\delta \rfloor$, the points $x_r = x + r/R$ $(1\le r\le R)$ satisfy $\|x_r - x_s\| \geqslant 1/R \geqslant \delta$ $(1\le r<s\le R)$, and thus

$$\Delta \geq R \geq 1/\delta - 1.$$

We first derive Gallagher's version of the Large Sieve Inequality, showing that up to constant factors, one cannot do much better than having Δ be as large as $\max\{N, 1/\delta - 1\}$.

<u>Theorem 2.1.</u> Let $M, N \in \mathbb{Z}$ satisfy $N \geq 1$, and define $T(x)$ as in (2.1). Suppose also that $\delta > 0$ and that the real numbers x_r $(1 \leq r \leq R)$ are δ-spaced. Then one has

$$\sum_{r=1}^{R} |T(x_r)|^2 \leq \left(\pi N + 1/\delta\right) \sum_{n=M+1}^{M+N} |c_n|^2.$$

Later, we shall see that the factor $\underset{\Delta}{\underset{\shortparallel}{\pi N + 1/\delta}}$ can be replaced by $N + 1/\delta - 1$, which is best possible.

A robust argument that may be applied in wide generality makes use of the Gallagher-Sobolev inequality.

<u>Lemma 2.2</u> (Sobolev) Suppose that a and b are real numbers with $a < b$, and $f : [a,b] \longrightarrow \mathbb{C}$ is a complex-valued function with continuous first derivative. Then one has

$$\left| f\left(\frac{a+b}{2}\right)\right| \leq \frac{1}{b-a} \int_a^b |f(x)|\, dx + \frac{1}{2} \int_a^b |f'(x)|\, dx,$$

and for any $x \in [a,b]$, one has

$$|f(x)| \leq \frac{1}{b-a} \int_a^b |f(t)|\, dt + \int_a^b |f'(t)|\, dt.$$

Proof. Suppose that $x \in [a,b]$. Then by integrating by parts, one sees that

$$\int_x^b f(u)\,du = \Big[f(u)(u-b) \Big]_x^b - \int_x^b f'(u)(u-b)\,du$$
$$= (b-x) f(x) - \int_x^b f'(u)(u-b)\,du,$$

and

$$\int_a^x f(u)\,du = \Big[f(u)(u-a) \Big]_a^x - \int_a^x f'(u)(u-a)\,du$$
$$= (x-a) f(x) - \int_a^x f'(u)(u-a)\,du.$$

By adding these relations, we obtain

$$(b-a) f(x) = \int_a^b f(u)\,du + \int_x^b f'(u)(u-b)\,du + \int_a^x f'(u)(u-a)\,du$$

$$\Rightarrow \quad (b-a)|f(x)| \le \int_a^b |f(u)|\,du + (b-x)\int_x^b |f'(u)|\,du + (x-a)\int_a^x |f'(u)|\,du.$$

Then for any $x \in [a,b]$, we see that

$$(b-a)|f(x)| \le \int_a^b |f(u)|\,du + (b-x+x-a)\int_a^b |f'(u)|\,du,$$

and when $x = \frac{a+b}{2}$, in which case $b-x = x-a = \frac{b-a}{2}$, we have

$$(b-a)|f(x)| \le \int_a^b |f(u)|\,du + \frac{b-a}{2}\left(\int_a^{(b+a)/2} |f'(u)|\,du + \int_{(b+a)/2}^b |f'(u)|\,du \right).$$

//

Proof of Theorem 2.1. Put $U(x) = T(x)\, e(-Lx)$, where

$$L = K + M + 1 \quad \text{and} \quad K = \lfloor N/2 \rfloor.$$

Then one sees that, since $N-K-1 \le K$

⑦
$$U(x) = \sum_{n=M+1}^{M+N} c_n \, e((n-L)x) = \sum_{k=-K}^{K} b_k \, e(kx),$$

$\boxed{k = n-L}$

where $b_k = c_{k+L}$ when $k+L \leq M+N$, and $b_k = 0$ for $k+L > M+N$.

Since $|U(x)| = |T(x)|$, we see that it suffices to consider
$$\sum_{r=1}^{R} |U(x_r)|^2 .$$

We apply Lemma 2.2 (the Gallagher - Sobolev inequality) to deduce that
$$\sum_{r=1}^{R} |U(x_r)|^2 \leq \frac{1}{\delta} \sum_{r=1}^{R} \int_{x_r - \delta/2}^{x_r + \delta/2} |U(x)|^2 \, dx + \sum_{r=1}^{R} \frac{1}{2} \int_{x_r - \delta/2}^{x_r + \delta/2} |2U(x) U'(x)| \, dx$$
$$\leq \frac{1}{\delta} \int_0^1 |U(x)|^2 \, dx + \int_0^1 |U(x) U'(x)| \, dx$$

note: the intervals $[x_r - \delta/2, x_r + \delta/2]$ are non-intersecting for $r = 1, 2, \ldots, R$.

By orthogonality, one has
$$\int_0^1 |U(x)|^2 \, dx = \int_0^1 \Big| \sum_{k=-K}^{K} b_k \, e(kx) \Big|^2 dx$$
$$= \sum_{k=-K}^{K} |b_k|^2 = \sum_{n=M+1}^{M+N} |c_n|^2 .$$

Similarly, by Schwarz's inequality,
$$\int_0^1 |U(x) U'(x)| \, dx \leq \Big(\int_0^1 |U(x)|^2 dx \Big)^{1/2} \Big(\int_0^1 |U'(x)|^2 dx \Big)^{1/2}$$
$$= \Big(\sum_{k=-K}^{K} |b_k|^2 \Big)^{1/2} \Big(\sum_{k=-K}^{K} |2\pi i k \, b_k|^2 \Big)^{1/2}$$
$$\leq 2\pi K \sum_{k=-K}^{K} |b_k|^2 = 2\pi K \sum_{n=M+1}^{M+N} |c_n|^2 .$$

Hence

$$\sum_{r=1}^{R} |U(x_r)|^2 \leq \left(\frac{1}{\delta} + 2\pi K \right) \sum_{n=M+1}^{M+N} |c_n|^2$$

$$\leq \left(\frac{1}{\delta} + \pi N \right) \sum_{n=M+1}^{M+N} |c_n|^2 . \; //$$

By working (much) harder, as we shall discuss in brief later, it is possible to replace the factor $1/\delta + \pi N$ by the best possible factor $1/\delta + N - 1$.

Theorem 2.3 (Montgomery and Vaughan, 1974; Selberg, 1974). Let $M, N \in \mathbb{Z}$ ~~satisfy~~ $N \geq 1$, and define $T(x)$ as in (2.1). Suppose also ~~that~~ $\delta > 0$ and ~~that~~ the real numbers x_r $(1 \leq r \leq R)$ are δ-spaced. Then one has

$$\sum_{r=1}^{R} |T(x_r)|^2 \leq \left(N + \frac{1}{\delta} - 1 \right) \sum_{n=M+1}^{M+N} |c_n|^2 .$$

Corollary 2.4. Let M and N be integers with $N \geq 1$, and define $T(x)$ as in (2.1). Then for any positive integer Q, one has

$$\sum_{q=1}^{Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |T(a/q)|^2 \leq (N + Q^2 - 1) \sum_{n=M+1}^{M+N} |c_n|^2 .$$

Proof. We apply Theorem 2.3 with the sequence $(a/q)_{\substack{1 \leq a \leq q \leq Q \\ (a,q)=1}}$ in place of (x_r).

Observe that when ~~...~~ $1 \le a_i < q_i \le Q$ satisfy $(a_i, q_i) = 1$ for $i = 1, 2$, and $a_1/q_1 \ne a_2/q_2$, ~~then~~

$$\left| \frac{a_1}{q_1} - \frac{a_2}{q_2} \right| = \left| \frac{a_1 q_2 - a_2 q_1}{q_1 q_2} \right| \geq \frac{1}{q_1 q_2} \geq \frac{1}{Q^2}.$$

Then the real points $\frac{a}{q}$ in question are δ-spaced with $\delta = 1/Q^2$, and the desired conclusion is immediate from Theorem 2.3. //

§3. The Large Sieve inequality, II: arithmetic formulation.

Suppose ~~that~~ N is a large positive integer, and we have a subset of $\{1, 2, \ldots, N\}$ ~~that~~ contains no integers x with $x \equiv 0 \pmod p$, nor with $x \equiv -2 \pmod p$, ~~for the~~ primes p with $p \le \sqrt{N+2}$. This is the case for the set of twin prime integers q with $q > \sqrt{N}$ for which both q and $q+2$ are prime. How large can this subset be? What if we omit different congruence classes modulo p? This is the subject of sieve theory.

We set things up in wider generality by counting integers with arbitrary complex weights $(c_n)_{n \in \mathbb{Z}}$. Let $M, N \in \mathbb{Z}$ satisfy $N \geq 1$, and define

$$Z = \sum_{n=M+1}^{M+N} c_n$$

and

$$Z(q,h) = \sum_{\substack{n=M+1 \\ n \equiv h \pmod q}}^{M+N} c_n .$$

<u>Goal</u>: Give an upper bound for $|Z|$.
Have in mind $c_n = \begin{cases} 1, & n \in \mathcal{S} \\ 0, & n \notin \mathcal{S} \end{cases}$ some sequence $\mathcal{S}$.

Referring back to (2.1), we have

$$T(a/q) = \sum_{h=1}^{q} \sum_{\substack{n=M+1 \\ n \equiv h \pmod q}}^{M+N} c_n e(na/q)$$

$$= \sum_{h=1}^{q} Z(q,h)\, e(ha/q).$$

Then by Parseval's identity, we have

$$q^{-1} \sum_{a=1}^{q} |T(a/q)|^2 = \sum_{h=1}^{q} |Z(q,h)|^2.$$

This relation allows us to relate the mean square $\sum_{a=1}^{q} |T(a/q)|^2$ to the variance of $Z(q,h)$, the mean value being

$$\frac{1}{q} \sum_{h=1}^{q} Z(q,h) = \frac{Z}{q}.$$

<u>Lemma 3.1.</u> One has

$$\sum_{h=1}^{q} |Z(q,h) - Z/q|^2 = \frac{1}{q} \sum_{a=1}^{q-1} |T(a/q)|^2.$$

<u>Proof.</u> One has

$$\sum_{h=1}^{q} |Z(q,h) - Z/q|^2 = \sum_{h=1}^{q} |Z(q,h)|^2 - 2\,\mathrm{Re}\left(\frac{\overline{Z}}{q} \sum_{h=1}^{q} Z(q,h)\right) + \frac{|Z|^2}{q}$$

$$= \sum_{h=1}^{q} |Z(q,h)|^2 - |Z|^2/q$$

$$= q^{-1} \sum_{a=1}^{q} |T(a/q)|^2 - |T(0)|^2/q.$$

$$= q^{-1} \sum_{a=1}^{q-1} |T(a/q)|^2. \quad //$$

The simplest consequences of Lemma 3.1 relate to the situation in which q is a prime number.

<u>Theorem 3.2.</u> Let $\mathcal{N}$ be a set of Z integers lying in the interval $[M+1, M+N]$. Put

$$Z(q,h) := \#\{n \in \mathcal{N} : n \equiv h \pmod{q}\}.$$

Then for any $Q \in \mathbb{N}$, we have

$$\sum_{p \leq Q} p \sum_{h=1}^{p} \left(Z(p,h) - Z/p \right)^2 \leq (N+Q^2) Z.$$

<u>Proof.</u> We fix $c_n = \begin{cases} 1, & n \in \mathcal{N} \\ 0, & n \notin \mathcal{N} \end{cases}$ and apply Lemma 3.1 and Corollary 2.4. Thus, we deduce that

$$\sum_{p \leq Q} \sum_{a=1}^{p-1} |T(a/p)|^2 = \sum_{p \leq Q} p \sum_{h=1}^{p} (Z(p,h) - Z/p)^2$$

$$\wedge\!\!\wedge$$

$$(N+Q^2-1) \sum_{n=M+1}^{M+N} |c_n|^2 \leq (N+Q^2) Z. \quad //$$

<u>Corollary 3.3.</u> Let $\mathcal{N} \subseteq [M+1, M+N]$ be a sets of integers. Fix a real number τ with $0 < \tau \leq 1$, and let $\mathcal{P}$ denote the set of primes p with $p \leq Q$ such that $Z(p,h) = 0$ for at least τp residue classes $h \pmod{p}$. Finally, write $P = \mathrm{card}(\mathcal{P})$. Then

$$Z \leq \frac{N+Q^2}{\tau P}.$$

<u>Proof.</u> Consider the conclusion of Theorem 3.2. When $p \in \mathcal{P}$, one has $Z(p,h) = 0$ for at least τp values of h with $1 \leq h \leq p$. Thus

$$(N+Q^2) Z \geq \sum_{p \in \mathcal{P}} p \cdot (Z/p)^2 \cdot \tau p = \tau Z^2 P$$

$$\Rightarrow \quad Z \le \frac{N+Q^2}{\tau P} \quad //$$

The utility of this estimate comes when our sets $\mathcal{N}$ under consideration avoid a large number of residue classes modulo p, hence the name large sieve. Suppose, for example, that $\mathcal{N}$ consists of the set of integers that for half the primes (say for $p \equiv 1 \pmod 4$) are congruent to squares (mod p), and for the remaining half of the primes (say for $p \equiv 3 \pmod 4$) are congruent to 1 plus a square (mod p).

Then for all primes $p \ge 3$, one has $Z(p,h) = 0$ for at least $\frac{1}{2}(p-1)$ values of h, namely those with $\left(\frac{h}{p}\right) = -1$. For $p \equiv 1 \pmod 4$, one again has $Z(p,h) = 0$ for at least $\frac{1}{2}(p-1)$ values of h. In this case one has $Z(p,h) = 0$ when $\left(\frac{h+1}{p}\right) = -1$.

Then we may take $Q = \sqrt{N}$,

$$\mathcal{P} = \{3 < p \le \underset{Q}{\underset{\shortparallel}{\sqrt{N}}} : p \text{ prime}\} \Rightarrow P \asymp \sqrt{N}/\log N,$$

$$\tau p = \tfrac{1}{2}(p-1) \Rightarrow \tau \ge 1/3.$$

Then Corollary 3.3 shows that

$$Z \le \frac{N+Q^2}{\tau P} \ll \frac{N}{\frac{1}{3}\sqrt{N}/\log N} \ll \sqrt{N}\log N.$$

A similar argument applies for $\mathcal{N} = \{n^2 : 1 \le n \le \sqrt{N}\}$, so this kind of estimate can be expected to be close to best possible.

When τ is much smaller (say $1/p$, or $2/p$, in a suitable sense), the conclusion of Corollary 3.3 is not especially efficient. If we vary τ with p so that $\tau = 1/p$ and consider the situation where $Z(p,h)=0$ for $h=0$ (only), then a modification of Corollary 3.3 shows that

$$Z \le \frac{N+Q^2}{\sum_{p\le\sqrt{N}} \frac{1}{p}} \ll \frac{N}{\log\log N}.$$

Since $\mathcal{N}$ is then the set of prime numbers not exceed N, and $\pi(N) \sim N/\log N$, one sees the relative inefficiency!

However, by considering all composite moduli instead of only prime moduli, one does better. This was achieved by Montgomery (1968).

<u>Lemma 3.4</u> (Montgomery). Consider a sequence of complex numbers $(c_n)_{M+1}^{M+N}$. For each prime number p, define

$$\mathcal{D}(p) = \{0 \le d < p : \ c_n = 0 \ \text{ whenever } n \equiv d \pmod p\}$$

and

$$\mathcal{R}(p) = \{0 \le d < p : \ c_n \ne 0 \ \text{ for some } n \equiv d \pmod p\}.$$

Write $\delta(p) = \operatorname{card}(\mathcal{D}(p))$ and note ~~that~~ $p - \delta(p) = \operatorname{card}(\mathcal{R}(p))$. Then, equipped with the notation of this section, for all squarefree integers q, one has

$$\sum_{\substack{a=1\\(a,q)=1}}^{q} |T(a/q)|^2 \ge |Z|^2 \prod_{p|q} \frac{\delta(p)}{p-\delta(p)}.$$

<u>Proof.</u> [Think of $\mathcal{D}(p)$ as the <u>deleted</u> residue classes, and $\mathcal{R}(p)$ as those that remain.]

We proceed by induction on the number of primes dividing q, beginning with the situation in which q is prime. Here we see from Lemma 3.1 that when $q=p$ is prime, one has

$$\sum_{a=1}^{p-1} |T(a/p)|^2 = p \sum_{h=1}^{p} |Z(p,h) - Z/p|^2$$

$$= p \sum_{h \in \mathcal{S}(p)} |Z(p,h) - Z/p|^2 + p \sum_{h \in \mathcal{R}(p)} |Z(p,h) - Z/p|^2$$

$$= p\left|\frac{Z}{p}\right|^2 \underbrace{|\mathcal{S}(p)|}_{\delta(p)} + p \sum_{h \in \mathcal{R}(p)} |Z(p,h) - Z/p|^2 .$$

But by Cauchy's inequality, one sees that

$$\left|\sum_{h \in \mathcal{R}(p)} \left(Z(p,h) - Z/p\right)\right|^2 \leq (p - \delta(p)) \sum_{h \in \mathcal{R}(p)} |Z(p,h) - Z/p|^2 ,$$

$$\| $$

$$|Z - (p - \delta(p)) Z/p|^2$$

$$\| $$

$$\frac{\delta(p)^2}{p^2} Z^2$$

Whence

$$\sum_{a=1}^{p-1} |T(a/p)|^2 \geq \frac{\delta(p)}{p} |Z|^2 + \frac{\delta(p)^2}{p(p-\delta(p))} |Z|^2$$

$$= |Z|^2 \frac{\delta(p)}{p-\delta(p)} .$$

This confirms the conclusion of the lemma when $q = p$ is prime. □

Suppose now that the conclusion of the lemma holds whenever q is the product of at most $t \geq 1$ primes. Let q be the product of $t+1$ primes, and write $q = q_1 q_2$ with $(q_1, q_2) = 1$, $q_1 > 1$ and $q_2 > 1$. Then each of q_1 and q_2 are the product of at most t primes, so we may assume that the conclusion of the lemma holds when q is replaced by q_i ($i = 1$ and 2). As a consequence of the Chinese Remainder Theorem, one has

$$\sum_{\substack{a=1\\(a,q)=1}}^{q} |T(a/q)|^2 = \sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1} \sum_{\substack{a_2=1\\(a_2,q_2)=1}}^{q_2} \Big| \underbrace{T\Big(\frac{a_1}{q_1}+\frac{a_2}{q_2}\Big)}_{\sum_{n=M+1}^{M+N}\left(c_n e\left(\frac{a_1}{q_1}n\right)\right) e\left(\frac{a_2}{q_2}n\right)} \Big|^2$$

$$\geq \sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1} \Big|T\Big(\frac{a_1}{q_1}\Big)\Big|^2 \prod_{p\mid q_2} \frac{\delta(p)}{p-\delta(p)}$$

$$\geq |T(0)|^2 \prod_{p\mid q_1} \frac{\delta(p)}{p-\delta(p)} \cdot \prod_{p\mid q_2} \frac{\delta(p)}{p-\delta(p)}$$

$$= |Z|^2 \prod_{p\mid q} \frac{\delta(p)}{p-\delta(p)}.$$

//

<u>Theorem 3.5.</u> Let $\mathcal{N} \subseteq [M+1, M+N]$ be a set of Z integers. For each prime number p, let $\delta(p)$ denote the number of residue classes (mod p) not represented by any elements $n \in \mathcal{N}$. Then for any $Q \in \mathbb{N}$, one has

$$Z \leq \frac{N+Q^2}{L},$$

where

$$L = \sum_{q\leq Q} \mu^2(q) \prod_{p\mid q} \frac{\delta(p)}{p-\delta(p)}.$$

<u>Proof.</u> By Lemma 3.4, one has

$$\sum_{1\leq q\leq Q} \sum_{\substack{a=1\\(a,q)=1}}^{q} |T(a,q)|^2 \geq Z^2 \sum_{1\leq q\leq Q} \mu^2(q) \prod_{p\mid q} \frac{\delta(p)}{p-\delta(p)} = Z^2 L$$

Corollary 2.4 $\wedge$

$(N+Q^2)Z$

Thus, if $Z>0$, one sees that

$$Z \leq (N+Q^2)/L.$$

//

Example. Let $\mathcal{N}$ be the set of prime numbers p lying in the interval $[1, N]$ for which $p+2$ is also prime. Then for each prime π with $\pi \leq \sqrt{N}$, one has for each $n \in \mathcal{N}$ ~~that~~

$$n \not\equiv 0 \pmod{\pi} \quad \text{and} \quad n \not\equiv \pi - 2 \pmod{\pi},$$

so we may put $\delta(\pi) = 2$ for all such primes.

Applying Theorem 3.5 with $Q = \sqrt{N}$, we find that

$$L = \sum_{1 \leq q \leq Q} \mu^2(q) \prod_{p \mid q} \frac{2}{p-2}$$

$$\gg (\log Q)^2 \gg (\log N)^2 \qquad \text{(using Qn C5 from Problem sheet I),}$$

whence

$$\text{card}(\mathcal{N}) = Z \leq (N + Q^2)/L \ll N/(\log N)^2. \quad //$$

Exercise. Show that

$$\sum_{\substack{p \text{ prime} \\ \& \\ p+2 \text{ prime}}} \frac{1}{p} < \infty .$$

---

§4. The Large Sieve Inequality, III: character sum formulation.

Our goal in this section is a mean square estimate for the character sum

$$S(\chi) := \sum_{n=M+1}^{M+N} c_n \chi(n) \qquad (c_n \in \mathbb{C}),$$

analogues to the Large Sieve Inequality in the shape of Corollary 2.4. Specifically, we will establish the following theorem.

Theorem 4.1. Let M and N be integers with $N \geq 1$. Then for any integer $Q \geq 1$, one has

$$\sum_{1\le q\le Q} \frac{q}{\varphi(q)} \sum_{\chi \pmod q}^{*} |S(\chi)|^2 \le (N+Q^2) \sum_{n=M+1}^{M+N} |c_n|^2 .$$

Here, the sum over Dirichlet characters $\chi \pmod q$ indicates (via the asterisk) that the characters are restricted to be primitive.

We establish the conclusion of Theorem 4.1 by transferring the underlying problem from the domain of multiplicative characters to that of additive characters. This entails a brief discussion (recalling some properties of) Gauss sums:

Gauss sums.

Definition 4.2. Given a Dirichlet character χ modulo q, we define the Gauss sum $\tau(\chi)$ associated with χ via the relation

$$\tau(\chi) = \sum_{a=1}^{q} \chi(a)\, e(a/q).$$

More generally, given an ~~natural number~~ integer n, one may consider the analogue of Ramanujan's sum given by

$$c_\chi(n) = \sum_{a=1}^{q} \chi(a)\, e(an/q).$$

We now explore some basic properties of these Gauss sums.

Theorem 4.3. Suppose ~~that~~ χ is a character modulo q, and $(n,q)=1$. Then

$$\chi(n)\tau(\bar\chi) = \sum_{a=1}^{q} \bar\chi(a)\, e(an/q)$$

and

$$\overline{\tau(\chi)} = \chi(-1)\,\tau(\bar\chi).$$

Proof. When $(n,q)=1$, we may make a change of variables to obtain the relation

$$\chi(n)\, c_\chi(n) = \sum_{a=1}^{q} \chi(an)\, e(an/q) = \sum_{b=1}^{q} \chi(b)\, e(b/q) = \tau(\chi),$$

whence

$$\chi(n)\,\tau(\bar\chi) = \chi(n)\cdot \bar\chi(n)\, c_{\bar\chi}(n) = c_{\bar\chi}(n) = \sum_{a=1}^{q} \bar\chi(a)\, e(an/q). \;\square$$

Also, on putting $n=-1$, ~~this last relation~~ yields

$$\chi(-1)\,\tau(\bar\chi) = \overline{\sum_{a=1}^{q} \chi(a)\, e(a/q)} = \overline{\tau(\chi)}. \;\square \;//$$

Gauss sums possess a quasi-multiplicative relation.

<u>Theorem 4.4.</u> Suppose that $(q_1,q_2)=1$, and that for $i=1$ and 2, the characters χ_i modulo q_i ~~satisfy~~ $\chi=\chi_1\chi_2$ and $q=q_1q_2$. Then one has

$$\tau(\chi) = \tau(\chi_1)\,\tau(\chi_2)\,\chi_1(q_2)\,\chi_2(q_1).$$

<u>Proof.</u> By the Chinese Remainder Theorem, we may express each residue $a \pmod{q_1q_2}$ in the shape $a_1q_2 + a_2q_1$. Thus

$$\tau(\chi) = \sum_{a=1}^{q} \chi(a)\, e(a/q) = \sum_{a_1=1}^{q_1}\sum_{a_2=1}^{q_2} \overbrace{\chi_1(a_1q_2+a_2q_1)}^{\chi(a_1q_2)}\,\overbrace{\chi_2(a_1q_2+a_2q_1)}^{\chi(a_2q_1)} \times e\left(\frac{a_1q_2+a_2q_1}{q_1q_2}\right)$$

$$= \sum_{a_1=1}^{q_1} \chi_1(a_1q_2)\, e\left(\frac{a_1}{q_1}\right) \sum_{a_2=1}^{q_2} \chi_2(a_2q_1)\, e\left(\frac{a_2}{q_2}\right)$$

$$= \chi_1(q_2)\,\chi_2(q_1) \sum_{a_1=1}^{q_1} \chi_1(a_1)\, e\left(\frac{a_1}{q_1}\right) \sum_{a_2=1}^{q_2} \chi_2(a_2)\, e\left(\frac{a_2}{q_2}\right)$$

$$= \chi_1(q_2)\,\chi_2(q_1)\,\tau(\chi_1)\,\tau(\chi_2). \;\square \;//$$

We require a good bound on $\tau(\chi)$ in order to make use of expansions of additive / multiplicative characters in terms of multiplicative / additive characters. Fortunately, we have excellent control of $|\tau(\chi)|$.

<u>Theorem 4.5.</u> Suppose that χ is a primitive character modulo q. Then one has

$$\chi(n)\,\tau(\bar\chi) = \sum_{a=1}^{q} \bar\chi(a)\, e(an/q)$$

(19)

for all integers n. Moreover, one has $|\tau(\chi)| = \sqrt{q}$.

Proof. Theorem 4.3 delivers the desired relation when $(n,q)=1$, so we may suppose ~~that~~ $(n,q)>1$, in which case $\chi(n)=0$ and we must show ~~that~~

$$\sum_{a=1}^{q} \bar{\chi}(a)\, e(an/q) = 0.$$

Put $d=(n,q)$, and ~~then write~~ $m=n/d$ and $r=q/d$. Thus

$$\sum_{a=1}^{q} \bar{\chi}(a)\, e(an/q) = \sum_{a=1}^{q} \bar{\chi}(a)\, e(am/r)$$

$$= \sum_{h=1}^{r} \Bigg(\underbrace{\sum_{\substack{a=1 \\ a\equiv h \pmod r}}^{q} \bar{\chi}(a)}_{=0} \Bigg) e(hm/r) \qquad = 0.$$

0 since $r \mid q$ and $r<q$.

(see p.134 of AANT notes).

□

Finally, we must confirm ~~that~~ $|\tau(\chi)|^2 = q$. But

$$\bar{\chi}(n)\,\tau(\chi) = \sum_{a=1}^{q} \chi(a)\, e(an/q) \qquad \text{(Theorem 4.3)},$$

so

$$\sum_{n=1}^{q} |\bar{\chi}(n)\,\tau(\chi)|^2 = \varphi(q)\, |\tau(\chi)|^2$$

$$\parallel$$

$$\sum_{n=1}^{q} \sum_{a_1=1}^{q} \chi(a_1)\, e(a_1 n/q) \sum_{a_2=1}^{q} \bar{\chi}(a_2)\, e(-a_2 n/q)$$

$$\parallel$$

$$\sum_{a_1=1}^{q} \sum_{a_2=1}^{q} \chi(a_1)\bar{\chi}(a_2) \underbrace{\sum_{n=1}^{q} e(n(a_1-a_2)/q)}_{= \begin{cases} 0, & \text{when } a_1 \not\equiv a_2 \pmod q \\ q, & \text{when } a_1 \equiv a_2 \pmod q \end{cases}}$$

$$\parallel$$

$$q \sum_{\substack{a=1 \\ (a,q)=1}}^{q} 1 = q\,\varphi(q).$$

So $|\tau(\chi)|^2 = q$. □ //

Finally (for now), we record a means of expressing character sums of multiplicative type in terms of additive characters, and vice versa.

<u>Theorem 4.6</u> (i) Suppose that χ is a primitive character modulo q. Then for any integer n,

$$\chi(n) = \frac{1}{\tau(\bar{\chi})} \sum_{a=1}^{q} \bar{\chi}(a)\, e(an/q).$$

(ii) Suppose that $(n,q)=1$. Then

$$e(n/q) = \frac{1}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\, q)} \bar{\chi}(n)\, \tau(\chi).$$

<u>Proof.</u> For part (i), observe that when χ is primitive modulo q one has $\tau(\bar{\chi}) \neq 0$, whence Theorem 4.5 shows that

$$\chi(n) = \frac{1}{\tau(\bar{\chi})} \sum_{a=1}^{q} \bar{\chi}(a)\, e(an/q). \quad \square$$

Meanwhile, by applying orthogonality of characters, one sees that when $(n,q)=1$,

$$\sum_{\chi \,(\mathrm{mod}\, q)} \bar{\chi}(n)\, \tau(\chi) = \sum_{a=1}^{q} \underbrace{\sum_{\chi \,(\mathrm{mod}\, q)} \bar{\chi}(n)\, \chi(a)}_{= \begin{cases} 0, & \text{when } n \not\equiv a \pmod q, \\ \varphi(q), & \text{when } n \equiv a \pmod q. \end{cases}} e(a/q)$$

$$= \varphi(q)\, e(n/q). \quad \square \quad //$$

We now return to the topic of the Large Sieve Inequality for character sums. Recall that $S(\chi) = \sum_{n=M+1}^{M+N} c_n \chi(n)$ and $T(\alpha) = \sum_{n=M+1}^{M+N} c_n e(\alpha)$.

Lemma 4.7. One has

$$\frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \pmod q} |S(\chi)|^2 \leq \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |T(a/q)|^2 .$$

Proof. From Theorem 4.6 (i), one has

$$\tau(\bar{\chi})S(\chi) = \tau(\bar{\chi}) \sum_{n=M+1}^{M+N} c_n \chi(n) = \sum_{a=1}^{q} \bar{\chi}(a) \sum_{n=M+1}^{M+N} c_n e(an/q)$$

$$= \sum_{a=1}^{q} \bar{\chi}(a)\, T(a/q).$$

Thus, ~~one~~ recalling ~~that~~ when χ is primitive (mod q), one has $|\tau(\chi)| = \sqrt{q}$, it follows ~~that~~

$$q \sideset{}{^*}\sum_{\chi} |S(\chi)|^2 = \sideset{}{^*}\sum_{\chi \pmod q} \Big| \sum_{a=1}^{q} \bar{\chi}(a) T(a/q) \Big|^2$$

$$\leq \sum_{\chi \pmod q} \Big| \sum_{a=1}^{q} \bar{\chi}(a) T(a/q) \Big|^2$$

$$= \varphi(q) \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |T(a/q)|^2 \quad \text{(using orthogonality of characters)}.$$

[& noting $\bar{\chi}(a) = 0$ when $(a,q) > 1$]. //

The proof of Theorem 4.1: Making use of Lemma 4.7, we deduce ~~that~~

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \pmod q} |S(\chi)|^2 \leq \sum_{q \leq Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |T(a/q)|^2$$

$$\overset{\text{Corollary 2.4}}{\leq} (N+Q^2) \sum_{n=M+1}^{M+N} |c_n|^2 . \quad //$$

§5. The Large Sieve inequality of Montgomery & Vaughan, and of Selberg.

We return to Theorem 2·3, and offer a fairly detailed sketch of the proof.

Theorem 2·3. (Montgomery & Vaughan, 1974; Selberg, 1974)

Let $M, N \in \mathbb{Z}$ satisfy $N \geq 1$, and define $T(x)$ as in (2·1). Suppose also ~~that~~ $\delta > 0$, and ~~that~~ the real numbers x_r $(1 \leq r \leq R)$ are δ-spaced. Then one has

$$\sum_{r=1}^{R} |T(x_r)|^2 \leq \left(N + \frac{1}{\delta} - 1\right) \sum_{n=M+1}^{M+N} |c_n|^2.$$

Lemma 5.1. (Montgomery and Vaughan, 1974). Suppose that $(\lambda_n)_{n=1}^{\infty}$ is a sequence of real numbers with $|\lambda_r - \lambda_s| \geq \delta > 0$ for $r \neq s$. Then for arbitrary complex sequences (x_n) and (y_n), one has

$$\left| \sum_{\substack{1 \leq m, n \leq N \\ m \neq n}} \frac{x_n y_m}{\lambda_m - \lambda_n} \right| \leq \frac{\pi}{\delta} \left(\sum_{n=1}^{N} |x_n|^2 \right)^{1/2} \left(\sum_{m=1}^{N} |y_m|^2 \right)^{1/2}.$$

Proof. By Cauchy's inequality, one has

$$\left| \sum_{\substack{1 \leq m, n \leq N \\ m \neq n}} \frac{x_n y_m}{\lambda_m - \lambda_n} \right|^2 \leq \left(\sum_{m=1}^{N} |y_m|^2 \right) \left(\sum_{m=1}^{N} \left| \sum_{\substack{1 \leq n \leq N \\ n \neq m}} \frac{x_n}{\lambda_m - \lambda_n} \right|^2 \right),$$

so it suffices to show ~~that~~

$$\sum_{m=1}^{N} \left| \sum_{\substack{1 \leq n \leq N \\ n \neq m}} \frac{x_n}{\lambda_m - \lambda_n} \right|^2 \leq \frac{\pi^2}{\delta^2} \sum_{n=1}^{N} |x_n|^2. \qquad \text{——— (5.1)}$$

Let $A = (a_{mn})_{1 \leq m, n \leq N}$ denote the matrix with elements

$$a_{mn} = \begin{cases} \dfrac{1}{\lambda_m - \lambda_n}, & \text{when } m \neq n, \\ 0, & \text{when } m = n. \end{cases}$$

(23) Note that the adjoint $A^* = (\overline{A})^T$ of A ~~satisfies~~ $A^* = -A$, so that A is skew-hermitian. Then by the theory of matrix operator norms, the extremal vector $\underline{x}$ in (5.1) is an eigenvector. Thus

$$\sum_{\substack{1 \le n \le N \\ n \ne m}} \frac{x_n}{\lambda_n - \lambda_m} = (A\underline{x})_{m\text{-th coordinate}} = \lambda x_m ,$$

for some eigenvalue λ of A. Since $-iA$ is hermitian, it follows that λ is purely imaginary, say $\lambda = i\nu$ for some $\nu \in \mathbb{R}$.

Next, expanding the left hand side of (5.1), we see that it becomes

$$\sum_{r=1}^{N} \sum_{s=1}^{N} x_r \overline{x}_s \sum_{\substack{1 \le m \le N \\ m \ne r \\ m \ne s}} \frac{1}{(\lambda_m - \lambda_r)(\lambda_m - \lambda_s)} .$$

The contribution from terms with $r = s$ is

$$\sum_{r=1}^{N} |x_r|^2 \sum_{\substack{1 \le m \le N \\ m \ne r}} \frac{1}{(\lambda_m - \lambda_r)^2} ,$$

while the contribution from those terms with $r \ne s$ is

$$\sum_{\substack{1 \le r,s \le N \\ r \ne s}} x_r \overline{x}_s \sum_{\substack{1 \le m \le N \\ m \ne r \\ m \ne s}} \underbrace{\frac{1}{(\lambda_m - \lambda_r)(\lambda_m - \lambda_s)}}_{\frac{1}{\lambda_r - \lambda_s}\left(\frac{1}{\lambda_m - \lambda_r} + \frac{1}{\lambda_s - \lambda_m}\right)}$$

$$= \sum_{\substack{1 \le r,s \le N \\ r \ne s}} \frac{x_r \overline{x}_s}{\lambda_r - \lambda_s} \left(\sum_{\substack{1 \le m \le N \\ m \ne r \\ m \ne s}} \frac{1}{\lambda_m - \lambda_r} - \sum_{\substack{1 \le m \le N \\ m \ne s \\ m \ne r}} \frac{1}{\lambda_m - \lambda_s} \right)$$

$$= \underbrace{2 \sum_{\substack{1 \le r,s \le N \\ r \ne s}} \frac{x_r \bar{x}_s}{(\lambda_r - \lambda_s)^2}}_{T_1} + \underbrace{\sum_{\substack{1 \le r,s \le N \\ r \ne s}} \frac{x_r \bar{x}_s}{\lambda_r - \lambda_s} \sum_{\substack{1 \le m \le N \\ m \ne r}} \frac{1}{\lambda_m - \lambda_r}}_{T_2}$$

include term $m = s$ in first term, & $m = r$ in second

$$- \underbrace{\sum_{\substack{1 \le r,s \le N \\ r \ne s}} \frac{x_r \bar{x}_s}{\lambda_r - \lambda_s} \sum_{\substack{1 \le m \le N \\ m \ne s}} \frac{1}{\lambda_m - \lambda_s}}_{T_3}.$$

The second term T_2 here is equal to

$$T_2 = \sum_{r=1}^{N} x_r \Big(\underbrace{\sum_{\substack{1 \le s \le N \\ s \ne r}} \frac{\bar{x}_s}{\lambda_r - \lambda_s}}_{\overline{i\nu x_r}} \Big) \Big(\sum_{\substack{1 \le m \le N \\ m \ne r}} \frac{1}{\lambda_m - \lambda_r} \Big) = -i\nu \sum_{r=1}^{N} |x_r|^2 \sum_{\substack{1 \le m \le N \\ m \ne r}} \frac{1}{\lambda_m - \lambda_r},$$

and similarly, the third term is

$$T_3 = -i\nu \sum_{s=1}^{N} |x_s|^2 \sum_{\substack{1 \le m \le N \\ m \ne s}} \frac{1}{\lambda_m - \lambda_s}.$$

Since $T_2 = T_3$, we conclude that

$$\sum_{r=1}^{N} \sum_{s=1}^{N} x_r \bar{x}_s \sum_{\substack{1 \le m \le N \\ m \ne r \\ m \ne s}} \frac{1}{(\lambda_m - \lambda_r)(\lambda_m - \lambda_s)} = \sum_{r=1}^{N} |x_r|^2 \sum_{\substack{1 \le m \le N \\ m \ne r}} \frac{1}{(\lambda_m - \lambda_r)^2} + T_1,$$

where, using the elementary inequality $|2 x_r \bar{x}_s| \le |x_r|^2 + |x_s|^2$, we have

$$T_1 \le \sum_{r=1}^{N} |x_r|^2 \sum_{\substack{1 \le s \le N \\ r \ne s}} \frac{1}{(\lambda_r - \lambda_s)^2}.$$

Hence

$$\sum_{m=1}^{N} \Big| \sum_{\substack{1\le n\le N \\ n\ne m}} \frac{x_n}{\lambda_m-\lambda_n} \Big|^2 \le 3 \sum_{n=1}^{N} |x_n|^2 \sum_{\substack{1\le m\le N \\ m\ne n}} \frac{1}{(\lambda_m-\lambda_n)^2}.$$

There is no loss of generality in supposing that $\lambda_1 < \lambda_2 < \dots$, whence $|\lambda_m - \lambda_n| \ge \delta |m-n|$, and thus

$$\sum_{\substack{1\le m\le N \\ m\ne n}} \frac{1}{(\lambda_m-\lambda_n)^2} \le \frac{1}{\delta^2} \cdot 2 \sum_{l=1}^{\infty} \frac{1}{l^2} = \frac{\pi^2}{3\delta^2}.$$

Consequently,

$$\sum_{m=1}^{N} \Big| \sum_{\substack{1\le n\le N \\ n\ne m}} \frac{x_n}{\lambda_m-\lambda_n} \Big|^2 \le 3 \cdot \frac{\pi^2}{3\delta^2} \sum_{n=1}^{N} |x_n|^2,$$

delivering (5.1) and the conclusion of the theorem. //

<u>Corollary 5.2.</u> Suppose that (α_r) is a sequence of δ-spaced points modulo 1, and that (z_r) is a complex sequence. Then one has

$$\Big| \sum_{\substack{1\le r,s<\infty \\ r\ne s}} \frac{z_r \bar{z}_s}{\sin(\pi(\alpha_r-\alpha_s))} \Big| \le \frac{1}{\delta} \sum_{r=1}^{\infty} |z_r|^2.$$

<u>Proof.</u> We begin by observing that

$$\frac{\pi}{\sin \pi w} = \lim_{K\to\infty} \sum_{k=-K}^{K} \frac{(-1)^k}{w-k} \qquad (w\in\mathbb{C}). \qquad (5.2)$$

In order to justify this assertion, recall that the Weierstrass product formula for $\sin(\pi w)$ asserts that

$$\sin \pi w = \pi w \prod_{k=1}^{\infty} \left(1-\frac{w}{k}\right)\left(1+\frac{w}{k}\right).$$

On taking logarithmic derivatives, we therefore see that

$$\pi \cot(\pi w) = \frac{1}{w} + \sum_{k=1}^{\infty}\left(\frac{1}{w-k} + \frac{1}{w+k}\right),$$

whence

$$\frac{1}{\sin(\pi w)} = \frac{1}{2}\cot\left(\frac{\pi w}{2}\right) - \frac{1}{2}\cot\left(\frac{\pi(w+1)}{2}\right)$$

$$= \frac{1}{\pi}\left(\frac{1}{w} + \sum_{k=1}^{\infty}(-1)^k\left(\frac{1}{w-k} + \frac{1}{w+k}\right)\right).$$

With the relation (5.2) in hand, we see that

$$\sum_{\substack{1\le r,s<\infty \\ r\ne s}} \frac{z_r\bar{z}_s}{\sin(\pi(\alpha_r-\alpha_s))} = \frac{1}{\pi}\lim_{K\to\infty}\sum_{\substack{1\le r,s<\infty \\ r\ne s}} z_r\bar{z}_s \sum_{k=-K}^{K}\frac{(-1)^k}{\alpha_r-\alpha_s-k}$$

$$= \frac{1}{\pi}\lim_{K\to\infty}\sum_{\substack{1\le r,s<\infty \\ r\ne s}} z_r\bar{z}_s \sum_{k=-K}^{K}\frac{(-1)^k(1-|k|/K)}{\alpha_r-\alpha_s-k}.$$

The introduction of the weight $(1-|k|/K)$ allows us to replace the sum over k by a sum over m and n, say, symmetrically as

$$\sum_{k=-K}^{K}\frac{(-1)^k(1-|k|/K)}{\alpha_r-\alpha_s-k} = \frac{1}{K}\sum_{1\le m,n\le K}\frac{(-1)^{n-m}}{\alpha_r-\alpha_s+n-m} \qquad (r\ne s).$$

When $r=s$, we may sum symmetrically to see that

$$\sum_{\substack{1\le m,n\le K \\ m\ne n}}\frac{(-1)^{n-m}}{n-m} = 0,$$

and hence

$$\sum_{\substack{1\le r,s<\infty \\ r\ne s}} z_r\bar{z}_s \sum_{k=-K}^{K}\frac{(-1)^k(1-|k|/K)}{\alpha_r-\alpha_s-k} = \frac{1}{K}\sum_{\substack{(n,r)\ne(m,s) \\ 1\le m,n\le K}}\frac{(-1)^{n-m}z_r\bar{z}_s}{(n+\alpha_r)-(m+\alpha_s)}.$$

The last expression may be rewritten in the form

$$\left|\frac{1}{K}\sum_{\substack{(n,r)\ne(m,s) \\ 1\le m,n\le K}}\frac{x_{nr}\,y_{ms}}{\lambda_{ms}-\lambda_{nr}}\right| \le \frac{\pi}{\delta}\left(\frac{1}{K}\sum_{n=1}^{K}\sum_{r}|x_{nr}|^2\right)^{1/2}\left(\frac{1}{K}\sum_{m=1}^{K}\sum_{s}|y_{ms}|^2\right)^{\frac{1}{2}}.$$

by putting $x_{nr} = (-1)^n z_r$, $y_{ms} = (-1)^{1-m}\bar{z}_s$, $\lambda_{ms} = m+\alpha_s$, $\lambda_{nr} = n+\alpha_r$.

(27)

Thus we conclude that

$$\left| \sum_{\substack{1 \le r,s < \infty \\ r \ne s}} \frac{z_r \overline{z_s}}{\sin(\pi(\alpha_r - \alpha_s))} \right| \le \frac{1}{\pi K} \cdot \frac{\pi}{\delta} \left(K \sum_r |z_r|^2 \right)^{\frac{1}{2}} \left(K \sum_s |z_s|^2 \right)^{\frac{1}{2}}$$
$$= \frac{1}{\delta} \sum_{r=1}^{\infty} |z_r|^2 . //$$

<u>Corollary 5.3.</u> For any real number θ, one has

$$\left| \sum_{\substack{1 \le r,s < \infty \\ r \ne s}} z_r \overline{z_s} \frac{\sin(2\pi\theta(\alpha_r - \alpha_s))}{\sin(\pi(\alpha_r - \alpha_s))} \right| \le \frac{1}{\delta} \sum_{r=1}^{\infty} |z_r|^2 .$$

<u>Proof.</u> One has

$$\sum_{\substack{1 \le r,s < \infty \\ r \ne s}} z_r \overline{z_s} \frac{\sin(2\pi\theta(\alpha_r - \alpha_s))}{\sin(\pi(\alpha_r - \alpha_s))} = \frac{1}{2i} \sum_{\substack{1 \le r,s < \infty \\ r \ne s}} \frac{z_r e(\theta\alpha_r) \overline{z_s e(\theta\alpha_s)}}{\sin(\pi(\alpha_r - \alpha_s))}$$
$$- \frac{1}{2i} \sum_{\substack{1 \le r,s < \infty \\ r \ne s}} \frac{z_r e(-\theta\alpha_r) \overline{z_s e(-\theta\alpha_s)}}{\sin(\pi(\alpha_r - \alpha_s))} ,$$

so from Corollary 5.2 one sees ~~that~~

$$\left| \sum_{\substack{1 \le r,s < \infty \\ r \ne s}} z_r \overline{z_s} \frac{\sin(2\pi\theta(\alpha_r - \alpha_s))}{\sin(\pi(\alpha_r - \alpha_s))} \right| \le \frac{1}{2} \left(\sum_{r=1}^{\infty} |z_r|^2 + \sum_{r=1}^{\infty} |z_r|^2 \right). //$$

<u>The proof of Theorem 2.3.</u> We first prove Theorem 2.3 with $N + 1/\delta$ in place of $N + 1/\delta - 1$. Here we make use of ~~the~~ duality principle for matrix operator norms. Thus, ~~the~~ validity of the upper bound

$$\sum_{n=M+1}^{M+N} \left| \sum_{r=1}^{R} z_r e(n\alpha_r) \right|^2 \le (N + \delta^{-1}) \sum_{r=1}^{R} |z_r|^2 ,$$

uniformly for all $z_r \in \mathbb{C}$, is equivalent to the validity of the bound

$$\sum_{r=1}^{R} \left| \sum_{n=M+1}^{M+N} c_n \, e(n x_r) \right|^2 \leq (N+\delta^{-1}) \sum_{n=M+1}^{M+N} |c_n|^2 .$$

But by squaring out, we see that

$$\sum_{n=M+1}^{M+N} \left| \sum_{r=1}^{R} z_r \, e(n x_r) \right|^2 = N \sum_{r=1}^{R} |z_r|^2 + \underbrace{\sum_{\substack{1 \leq r,s \leq R \\ r \neq s}} z_r \bar{z}_s \sum_{n=M+1} e(n(x_r - x_s))}_{T}.$$

The off-diagonal contribution T is equal to

$$\sum_{\substack{1 \leq r,s \leq R \\ r \neq s}} z_r \, e(K x_r) \, \overline{z_s \, e(K x_s)} \; \frac{\sin(\pi N(\alpha_r - \alpha_s))}{\sin(\pi(\alpha_r - \alpha_s))},$$

where $K = M + \frac{1}{2}(N+1)$. Then by Corollary 5.3, we see that

$$\sum_{n=M+1}^{M+N} \left| \sum_{r=1}^{R} z_r \, e(n x_r) \right|^2 \leq N \sum_{r=1}^{R} |z_r|^2 + \frac{1}{\delta} \sum_{r=1}^{R} |z_r|^2 .$$

$$= (N + 1/\delta) \sum_{r=1}^{R} |z_r|^2 . \quad \square$$

We now refine ~~this~~ bound by applying a device of Paul Cohen. Let K be a large ~~integer~~, and expand ~~the~~ original set of δ-spaced points x_r modulo 1 $(1 \leq r \leq R)$ to ~~the set~~ $\mathfrak{X}_K = \left\{ \frac{x_r + k}{K} : 1 \leq r \leq R, \; 1 \leq k \leq K \right\}$. Then we have shown that for the $\frac{\delta}{K}$-spaced points $y_{r,k} \in \mathfrak{X}_K$, one has

$$\sum_{n=M+1}^{M+N} c_n \, e\left(Kn \left(\frac{x_r}{K} + \frac{k}{K} \right) \right) = \sum_{n=M+1} c_n \, e(n x_r)$$

$$\| $$

$$\sum_{n'=M'+1}^{M'+N'} c'_{n'} \, e\left(n' \left(\frac{x_r}{K} + \frac{k}{K} \right) \right),$$

where $M' = MK + K - 1$, $N' = NK - K + 1$, and

$$c'_{n'} = \begin{cases} c_n, & \text{when } n' = Kn, \\ 0, & \text{otherwise}. \end{cases}$$

Then, by what we have already proved concerning the large sieve inequality, one has

$$K \sum_{r=1}^{R} \Big| \sum_{n=M+1}^{M+N} c_n e(nx_r) \Big|^2 = \sum_{k=1}^{K} \sum_{r=1}^{R} \Big| \sum_{n'=M'+1}^{M'+N'} c'_{n'} e\Big(n'\Big(\frac{x_r+k}{K}\Big)\Big)\Big|^2$$

$$\leq \left(N' + (\delta/K)^{-1}\right) \sum_{n'=M'+1}^{M'+N'} |c'_{n'}|^2$$

$$= (NK - K + 1 + K/\delta) \sum_{n=M+1}^{M+N} |c_n|^2 .$$

Hence

$$\sum_{r=1}^{R} |T(x_r)|^2 \leq \left(N - 1 + \frac{1}{K} + 1/\delta\right) \sum_{n=M+1}^{M+N} |c_n|^2 ,$$

and the conclusion of Theorem 2.3 follows on taking the limit as $K \to \infty$. //

---

§6. Estimates for arithmetic sums involving prime numbers.

We have seen certain prime number sums in the first course in analytic number theory. By applying the prime number theorem, sums of the shape

$$\sum_{1 \leq n \leq N} \Lambda(n) f(n)$$

may be estimated via Riemann–Stieltjes integration when f is reasonably smooth. When $f(n)$ is multiplicative, on the other hand, we may apply methods involving the Dirichlet series $F(s) = \sum_{n=1}^{\infty} f(n) n^{-s}$ via the logarithmic integral $\frac{F'(s)}{F(s)}$. Such ideas are effective for functions f of the form $f(n) = \chi(n) n^{-it}$. We now turn to consider functions f having neither of these properties.

A naïve idea would be to approximate sums $\sum_{p \le N} f(p)$ via the subsums $\sum_{\sqrt{N} < p \le N} f(p)$, and to detect the primality condition via an elementary sieve. Thus, writing

$$P = \prod_{p \le \sqrt{N}} p,$$

we have

$$f(1) + \sum_{\sqrt{N} < p \le N} f(p) = \sum_{\substack{1 \le n \le N \\ (n,P)=1}} f(n) = \sum_{\substack{t \mid P \\ t \le N}} \mu(t) \sum_{r \le N/t} f(rt)$$

[Note that the middle sum is equal to $\sum_{1 \le n \le N} f(n) \sum_{\substack{t \mid P \\ t \mid n}} \mu(t)$, which gives the final expression on interchanging the order of summation].

This is a double sum, which is good. But when t is close to N the inner sum is too short to be usefully estimated.

Our goal is to rearrange these sums to obtain a useful decomposition, and this was a task first accomplished by I. M. Vinogradov. We seek to write

$$S = \sum_{1 \le n \le N} f(n) \Lambda(n)$$

as a linear combination of

Type I sum: $\displaystyle\sum_{1 \le t \le T} a(t) \sum_{r \le N/t} f(tr)$

Type II sum: $\displaystyle\sum_{\substack{1 \le n \le N \\ n = mk \\ m > U,\ k > V}} b(m)\, c(k)\, f(n)$

in which $a(t), b(m), c(k)$ are suitable arithmetic functions independent of f, and T, U, V are parameters to be chosen optimally.

Vaughan's identity: In 1977, Vaughan supplied a simple yet effective version of this ~~strategy~~ of great utility. In the next section, we shall apply this in the special case that $f(n) = e(n\alpha)$, for a given $\alpha \in \mathbb{R}$, though for now we proceed in wide generality. This is tidily motivated by considering Dirichlet series. Observe ~~that~~ when $\sigma > 1$ and F, G have abscissa of convergence larger than 1,

$$-\frac{\zeta'}{\zeta}(s) = \left(-\zeta'(s) - F(s)\,\zeta(s)\right)\left(\frac{1}{\zeta(s)} - G(s)\right)$$

$$+ F(s) - \zeta(s)F(s)G(s) - \zeta'(s)G(s) \quad . \text{ --- } (6.1)$$

$$\| \qquad \sum_{n=1}^{\infty} \Lambda(n) n^{-s}$$

We apply this identity with

$$F(s) = \sum_{d \le U} \Lambda(d)\, d^{-s} \quad \text{and} \quad G(s) = \sum_{k \le V} \mu(k)\, k^{-s}$$

"approximates $-\frac{\zeta'}{\zeta}(s)$" "approximates $\frac{1}{\zeta(s)}$".

Writing $\zeta(s) = \sum_{r=1}^{\infty} r^{-s}$ and $-\zeta'(s) = \sum_{m=1}^{\infty} (\log m)\, m^{-s}$, one can interpret (6.1) as an identity on Dirichlet series coefficients:

$$\Lambda(n) = c_1(n) + c_2(n) + c_3(n) + c_4(n), \qquad \text{——— (6.2)}$$

where

$$c_1(n) = \begin{cases} \Lambda(n), & \text{when } n \le U, \\ 0, & \text{when } n > U, \end{cases}$$

$$c_2(n) = - \sum_{\substack{rdk = n \\ d \le U \\ k \le V}} \Lambda(d)\, \mu(k),$$

$$C_3(n) = \sum_{\substack{mk = n \\ k \le V}} \mu(k) \log m,$$

and $c_4(n)$ corresponds to the first complicated term. This is simplified by observing that

$$-\zeta'(s) - F(s)\zeta(s)$$

has Dirichlet coefficients

$$\log m - \sum_{\substack{d \mid m \\ d \le U}} \Lambda(d) = \sum_{\substack{d \mid m \\ d > U}} \cancel{\log}\, \Lambda(d)$$

and $\dfrac{1}{\zeta(s)} - G(s)$ has Dirichlet coefficients $\begin{cases} \mu(k), & \text{when } k > V \\ 0, & \text{when } k \le V, \end{cases}$

so that

$$c_4(n) = \sum_{\substack{mk = n \\ m > U \\ k > V}} \Big(\sum_{\substack{d \mid m \\ d > U}} \Lambda(d) \Big) \mu(k).$$

We may now write

$$S = \sum_{1 \le n \le N} f(n)\Lambda(n) = \sum_{1 \le n \le N} f(n)\big(c_1(n) + \dots + c_4(n)\big)$$

$$= S_1 + S_2 + S_3 + S_4,$$

where

$$S_i = \sum_{1 \le n \le N} f(n)\, c_i(n).$$

We now briefly consider each of these terms in turn for a general weight $f(n)$, having in mind later ~~that~~ we shall consider the special case $f(n) = e(n\alpha)$.

(i) the term S_1. Here we have

$$S_1 = \sum_{n \le U} f(n)\Lambda(n),$$

and by choosing U small enough this term may be treated trivially (it is a shortened version of the original sum).

(ii) the term S_2. Here, on writing

$$a(t) = -\sum_{\substack{dk = t \\ d \le U \\ k \le V}} \Lambda(d)\mu(k),$$

we see that $c_2(n) = \sum_{t \mid n} a(t)$, and thus

$$S_2 = \sum_{t \le UV} a(t) \sum_{r \le N/t} f(tr).$$

This is a type I sum. Moreover, since $|a(t)| \le \sum_{d \mid t} \Lambda(d) = \log t$, it follows that $S_2 \ll (\log(UV)) \sum_{t \le UV} \left| \sum_{r \le N/t} f(rt) \right|$.

(iii) the term S_3. We have

$$S_3 = \sum_{k \leq V} \mu(k) \sum_{m \leq N/k} f(km) \log m .$$

The factor $\log m$ prevents this from immediately being of type I, but this factor is at least smoothly increasing. By noting ~~that~~

$$\log m = \int_1^m \frac{dw}{w},$$

it follows that

$$S_3 = \int_1^N \sum_{k \leq V} \mu(k) \sum_{w \leq m \leq N/k} f(km) \frac{dw}{w}$$

$$\ll (\log N) \sum_{k \leq V} \max_{w \geq 1} \Big| \sum_{w \leq m \leq N/k} f(km) \Big| .$$

This is a maximal variant of a type I sum.

(iv) the term S_4. This is superficially more complicated, but on writing

$$b(m) = \sum_{\substack{d \mid m \\ d > U}} \Lambda(d),$$

we see ~~that~~

$$c_4(n) = \sum_{\substack{mk = n \\ m > U \\ k > V}} b(m) \mu(k),$$

and thus

$$S_4 = \sum_{\substack{mk \leq N \\ m > U \\ k > V}} b(m) \mu(k) f(mk) = \sum_{U < m \leq N/V} b(m) \sum_{V < k \leq N/m} \mu(k) f(mk),$$

which is a Type II sum. Notice that this sum has a bilinear structure and can be estimated in terms of sums of the shape

$$\sum_{M<m\leq 2M} x_m \sum_{k\leq N/m} y_k\, f(mk),$$

in which $x_m = b(m)$ and $y_k = \mu(k)$. For arbitrary complex sequences (x_m), (y_k), one may seek an optimal $\Delta(M) = \Delta(M, N, f)$ ~~so that~~ this bilinear sum Σ has the bound

$$|\Sigma| \leq \Delta(M) \Big(\sum_{M<m\leq 2M} |x_m|^2 \Big)^{\frac{1}{2}} \Big(\sum_{k\leq N/M} |y_k|^2 \Big)^{\frac{1}{2}}, \quad \text{———} (6\cdot 2)$$

and thus we again enter the realm of matrix operator norms. Note that in our set-up, one has

$$\sum_{M<m\leq 2M} |x_m|^2 \leq \sum_{M<m\leq 2M} (\log m)^2 \ll M\,(\log(2M))^2$$

and

$$\sum_{k\leq N/M} |y_k|^2 \leq N/M,$$

so that we have

$$|\Sigma| \ll N\,(\log(2M))^2\, \Delta(M).$$

Let us pursue the inequality $(6\cdot 2)$ a little further. Observe first that by Cauchy's inequality, one has

$$|\Sigma| \leq \Big(\sum_{M<m\leq 2M} |x_m|^2 \Big)^{\frac{1}{2}} \Big(\underbrace{\sum_{M<m\leq 2M} \Big| \sum_{k\leq N/m} y_k\, f(mk) \Big|^2}_{\Sigma_1} \Big)^{\frac{1}{2}},$$

in which

$$\Sigma_1 = \sum_{k_1\leq N/M} y_{k_1} \sum_{k_2\leq N/M} \overline{y_{k_2}} \sum_{\substack{M<m\leq 2M \\ m\leq N/k_1 \\ m\leq N/k_2}} f(mk_1)\,\overline{f(mk_2)}.$$

Since $|y_{k_1} y_{k_2}| \le \frac{1}{2}(|y_{k_1}|^2 + |y_{k_2}|^2)$, we find ~~that~~

$$\Sigma_1 \ll \sum_{k_1 \le N/M} |y_{k_1}|^2 \sum_{k_2 \le N/M} \Bigg| \sum_{\substack{M < m \le 2M \\ m \le N/k_i \ (i=1,2)}} f(mk_1)\overline{f(mk_2)} \Bigg|$$

$$\le \Bigg(\sum_{k_1 \le N/M} |y_{k_1}|^2 \Bigg) \max_{k_1 \le N/M} \sum_{k_2 \le N/M} \Bigg| \sum_{\substack{M < m \le 2M \\ m \le N/k_i \ (i=1,2)}} f(mk_1)\overline{f(mk_2)} \Bigg| .$$

We therefore deduce ~~that~~ on putting

$$\Delta(M) = \Bigg(\max_{k_1 \le N/M} \sum_{k_2 \le N/M} \Bigg| \sum_{\substack{M < m \le 2M \\ m \le N/k_i \ (i=1,2)}} f(mk_1)\overline{f(mk_2)} \Bigg| \Bigg)^{1/2}, \tag{6.3}$$

we have $|S_4| \le |\Sigma| \ll N(\log(2M))^2 \Delta(M)$.

§7. An exponential sum over prime numbers.

In this section, we apply these ideas in the case $f(n) = e(n\alpha)$, ~~thereby~~ analysing

$$\sum_{n \le X} \Lambda(n)\, e(n\alpha),$$

or, more or less equivalently,

$$\sum_{p \le X} e(p\alpha).$$

These estimates can be used to establish approximations to the conjectures of Goldbach:

(i) all large enough odd numbers are the sum of 3 primes;

(ii) almost all even numbers are the sum of 2 primes.

(These results are due essentially to I.M. Vinogradov, 1937).

Theorem 7.1. Let $\alpha \in \mathbb{R}$, and suppose that $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a,q)=1$ and $|\alpha - a/q| \leq 1/q^2$. Then

$$\sum_{1 \leq n \leq N} \Lambda(n)\, e(n\alpha) \ll N(\log N)^{5/2}\left(q^{-1} + N^{-2/5} + qN^{-1}\right)^{1/2}.$$

Proof. We apply the discussion of the previous section, whence

$$\sum_{1 \leq n \leq N} \Lambda(n)\, e(n\alpha) = S_1 + S_2 + S_3 + S_4,$$

where, for appropriate choices of the parameters U and V, one has

$$S_1 = \sum_{n \leq U} \Lambda(n)\, e(n\alpha) \ll \sum_{n \leq U} \Lambda(n) \ll U,$$

$$S_2 = \sum_{t \leq UV} a(t) \sum_{r \leq N/t} e(tr\alpha) \ll (\log(UV)) \sum_{t \leq UV} \Big| \sum_{r \leq N/t} e(rt\alpha) \Big|,$$

$$S_3 = \sum_{k \leq V} \mu(k) \sum_{m \leq N/k} (\log m)\, e(km\alpha) \ll (\log N) \sum_{k \leq V} \max_{W \geq 1} \Big| \sum_{W \leq m \leq N/k} e(km\alpha) \Big|,$$

and

$$S_4 = \sum_{U < m \leq N/V} b(m) \sum_{V < k \leq N/m} \mu(k)\, e(mk\alpha).$$

By dividing the range of summation over m into dyadic intervals (and if necessary extending the definition of $b(m)$ to be 0 outside of $U < m \leq N/V$), we see that

$$S_4 \ll (\log N) \max_{U \leq M \leq N/V} \Big| \sum_{M < m \leq 2M} b(m) \sum_{V < k \leq N/m} \mu(k)\, e(mk\alpha) \Big|$$

$$\ll (\log N) \max_{U \leq M \leq N/V} \Big(\sum_{M < m \leq 2M} |b(m)|^2 \Big)^{\frac{1}{2}} \Big(\sum_{M < m \leq 2M} \Big| \sum_{V < k \leq N/m} \mu(k)\, e(mk\alpha) \Big|^2 \Big)^{\frac{1}{2}}$$

$$\ll (\log N) \max_{U \le M \le N/V} \left(M (\log M)^2 \right)^{1/2} \left(\sum_{k \le N/M} |\mu(k)|^2 \right)^{1/2} \max_{k_1 \le N/M} S_5(M, k_1)^{\frac{1}{2}},$$

where

$$S_5 = \sum_{k_2 \le N/M} \Big| \sum_{\substack{M < m \le 2M \\ m \le N/k_i \ (i=1,2)}} e\left(m (k_2 - k_1) \alpha \right) \Big| .$$

Thus, we have

$$S_4 \ll N^{\frac{1}{2}} (\log N)^2 \max_{U \le M \le N/V} \max_{k_1 \le N/M} S_5(M, k_1)^{1/2}.$$

We may estimate the sums S_2 and S_3 by making use of standard exponential sum estimates (from AHA, for example). Thus, we have

$$\sum_{X \le x \le X+Y} e(x\theta) \ll \min \{ Y, \|\theta\|^{-1} \}. \quad \text{———} \ (7.1)$$

Also, when $\alpha \in \mathbb{R}$, and $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ ~~satisfy~~ $(a,q) = 1$ and $|\alpha - a/q| \le 1/q^2$, we have

$$\sum_{1 \le x \le X} \min \{ XY/x, \ \| \alpha x \|^{-1} \} \ll XY \left(q^{-1} + Y^{-1} + q(XY)^{-1} \right) \cdot \log (2qX). \quad \text{———} \ (7.2)$$

We shall return to these estimates shortly.

By applying (7.1) and (7.2) in turn, we find first ~~that~~

$$\sum_{t \le UV} \Big| \sum_{r \le N/t} e(rt\alpha) \Big| \ll \sum_{t \le UV} \min \{ N/t, \ \| t\alpha \|^{-1} \}$$

$$\ll N \left(q^{-1} + \frac{UV}{N} + q N^{-1} \right) \log (2qUV),$$

whence

$$S_2 \ll N \left(q^{-1} + \frac{UV}{N} + q N^{-1} \right) \left(\log (2qUV) \right)^2 . \quad \text{———} \ (7.3)$$

Similarly,

$$S_3 \ll (\log N) \sum_{k \le V} \cancel{\max_{w \ge 1}} \min \left\{ \frac{N}{k}, \ \| k\alpha \|^{-1} \right\},$$

whence

$$S_3 \ll (\log(2qNV))^2 \cdot N\left(q^{-1} + \frac{V}{N} + qN^{-1}\right). \quad \text{— (7.4)}$$

Finally, one has

$$S_5(M, k_1) \ll \sum_{k_2 \leq N/M} \min\{M, \|(k_2 - k_1)\alpha\|^{-1}\}$$

$$\ll \sum_{1 \leq k \leq N/M} \min\{N/k, \|k\alpha\|^{-1}\} + M$$

$$\ll N\left(q^{-1} + \frac{M}{N} + \frac{1}{M} + q/N\right)\log(2qN).$$

Thus

$$S_4 \ll N^{\frac{1}{2}}(\log N)^2 \Big(N\big(q^{-1} + \underset{(M\leq N/V)}{V^{-1}} + \underset{(M\geq U)}{U^{-1}} + qN^{-1}\big)\log(2qN)\Big)^{\frac{1}{2}}$$

$$\ll N(\log(2qN))^{5/2}\left(q^{-\frac{1}{2}} + V^{-1/2} + U^{-1/2} + q^{\frac{1}{2}}N^{-\frac{1}{2}}\right). \quad \text{— (7.5)}$$

Notice ~~that~~ in (7.3) – (7.5), we may suppose ~~that~~ $q \leq N$, for otherwise the conclusion of Theorem 7.1 is trivial. Then it follows ~~that~~

$$\sum_{1 \leq n \leq N} \Lambda(n)\, e(n\alpha) = S_1 + \cdots + S_4$$

$$\ll U + N(\log(2qN))^{5/2}\left(q^{-1} + U^{-1} + V^{-1} + \left(\frac{UV}{N}\right)^2 + qN^{-1}\right)^{\frac{1}{2}},$$

provided that $UV \leq N$. The right hand side here is optimised by taking $U = V$ and $U^{-1} = U^4/N^2$, which is to say $U = V = N^{2/5}$. In this way, we have

$$\sum_{1 \leq n \leq N} \Lambda(n)\, e(n\alpha) \ll N^{2/5} + N(\log N)^{5/2}\left(q^{-1} + N^{-2/5} + qN^{-1}\right)^{\frac{1}{2}},$$

and the conclusion of the theorem follows. //

Before returning to justify (7.1) and (7.2), we pause to make two remarks:

(i) Dirichlet's approximation theorem shows that whenever $\alpha \in \mathbb{R}$, there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a,q)=1$, $1 \le q \le Q$ and

$$|\alpha - a/q| \le 1/qQ \le 1/q^2 .$$

Thus the initial assumption in Theorem 7.1 is harmless.

(ii) We could attempt to apply information gained from the Prime Number Theorem in arithmetic progressions to approach Theorem 7.1. If $\beta = \alpha - a/q$ is sufficiently small, then one may apply partial summation to estimate

$$\sum_{1 \le n \le N} \Lambda(n)\, e(n\alpha) \quad \text{in terms of} \quad \sum_{1 \le n \le N} \Lambda(n)\, e(na/q).$$

But

$$\sum_{1 \le n \le N} \Lambda(n)\, e(na/q) = \sum_{r=1}^{q-1} e(ra/q) \sum_{\substack{1 \le n \le N \\ n \equiv r \,(\mathrm{mod}\, q)}} \Lambda(n) + \sum_{\substack{1 \le n \le N \\ q \mid n}} \Lambda(n).$$

Now

$$\sum_{\substack{1 \le n \le N \\ n \equiv r \,(\mathrm{mod}\, q)}} \Lambda(n) = \frac{N}{\phi(q)} + O_A\left(N \exp(-c\sqrt{\log N})\right),$$

when $(r,q)=1$ and $q \le (\log N)^A$. Thus we have the prospect of a satisfactory estimate when $q \le (\log N)^A$, but permitting q to be as large as $\sqrt{N}$, say, would require something along the lines of GRH. Thus Theorem 7.1 already encodes information concerning an averaged version of a quasi-GRH.

We now return to confirm (7.1) and (7.2). First, on noting the trivial bound $|e(\alpha x)|\le 1$, we see that $\left|\sum_{X<x\le X+Y} e(\alpha x)\right| \le Y$. Hence we may suppose ~~that~~ $\|\alpha\| \neq 0$. But then the geometric progression formula shows ~~that~~

$$\sum_{X<x\le X+Y} e(\alpha x) = \frac{e(\alpha\lfloor X+Y+1\rfloor) - e(\alpha\lfloor X+1\rfloor)}{e(\alpha)-1}$$

$$\ll |e(\alpha/2)-e(-\alpha/2)|^{-1} \ll |\sin(\pi\alpha)|^{-1}.$$

Since $2\alpha \le \sin(\pi\alpha) \le \pi\alpha$ for $0\le\alpha\le 1/2$, it follows ~~that the~~ right hand side here is $O(\|\alpha\|^{-1})$. $\square$ This confirms (7.1).

Now for (7.2). We may assume ~~that~~ $\alpha\in\mathbb{R}$, $a\in\mathbb{Z}$, $q\in\mathbb{N}$ satisfy $(a,q)=1$ and $|\alpha - a/q| \le 1/q^2$, and we seek to confirm ~~that~~

$$S := \sum_{1\le x\le X} \min\{XY/x,\ \|\alpha x\|^{-1}\} \ll XY\left(q^{-1}+Y^{-1}+q(XY)^{-1}\right) \times \log(2qX).$$

Divide the range of summation into arithmetic progressions modulo q. Thus

$$S \le \sum_{0\le j\le X/q}\ \sum_{r=1}^{q} \min\left\{\frac{XY}{qj+r},\ \|\alpha(qj+r)\|^{-1}\right\}.$$

Write $\alpha - a/q = \theta/q^2$, so ~~that~~ $|\theta|\le 1$, and note ~~that~~

$$\alpha(jq+r) = \frac{[\alpha j q^2] + \{\alpha j q^2\}}{q} + \frac{\theta r}{q^2} + \frac{ar}{q}.$$

Divide into cases:

(i) $j=0$ and $1\le r\le q/2$: Then

$$\|\alpha(qj+r)\| \ge \|ar/q\| - |\theta r|/q^2 \ge \|ar/q\| - 1/2q \ge \tfrac{1}{2}\|ar/q\|.$$

(ii) $j\neq 0$, or $j=0$ and $\tfrac{1}{2}q < r\le q$. Then $qj+r \gg q(j+1)$, and

$$\left\|\frac{[\alpha j q^2]+ar}{q}\right\| \ge \frac{3}{q},$$ unless $ar + [\alpha j q^2] \equiv b \pmod{q}$

for some b with $|b| \le 2$. For a fixed j, there are at most 5 possible values of r with $\frac{1}{2}q < r \le q$ ~~satisfying~~ this inequality, and for the remaining values

$$\| \alpha (qj+r) \| \geqslant \left\| \frac{[\alpha j q^2] + ar}{q} \right\| - \frac{2}{q} \geqslant \frac{1}{3} \left\| \frac{[\alpha j q^2] + ar}{q} \right\| .$$

Thus

$$S \ll \sum_{0 \le j \le X/q} \left(\frac{XY}{q(j+1)} + \sum_{\substack{r=1 \\ q \nmid ([\alpha j q^2] + ar)}}^{q} \left\| \frac{[\alpha j q^2]+ar}{q} \right\|^{-1} \right) + \sum_{1 \le r \le \frac{1}{2}q} \left\| \frac{ar}{q} \right\|^{-1}$$

$$\ll \frac{XY}{q} \sum_{0 \le j \le X/q} \frac{1}{j+1} + \left(\frac{X}{q} + 1 \right) \sum_{1 \le h \le \frac{1}{2}q} \frac{q}{h}$$

$$\ll \frac{XY}{q} \log \left(\frac{X}{q} + 1 \right) + (X+q) \log q$$

$$\ll XY \left(q^{-1} + Y^{-1} + q(XY)^{-1} \right) \log (2qX). \quad //$$

§8. Sums of primes : the Goldbach problems.

In 1742, Goldbach wrote to Euler conjecturing that integers should be the sum of either 2 or 3 prime numbers. The ternary Goldbach problem, to show that ~~integers~~ odd integers $n \ge 7$ should be the sum of 3 primes:

$$n = p_1 + p_2 + p_3 ,$$

was solved for large enough odd integers by Vinogradov in 1937. The binary Goldbach problem, to show that even integers $n \ge 4$ should be the sum of two primes:

$$n = p_1 + p_2 ,$$

remains open. However, an "almost-all" analogue is available.

Let

$$E(X) := \text{card}\,\{1 \le m \le X : \ 2m \text{ is \underline{not} a sum of } \cancel{\text{two}} \text{ primes}\}$$

Then it can be shown ~~that~~ $E(X) = o(X)$ for $X \to \infty$, which shows ~~that~~ almost all even integers are ~~the~~ sum of ~~two~~ primes (in ~~the~~ sense of natural density). Both results employ ~~estimates~~ from §7 as well as the Siegel-Walfisz theorem, as we now explain.

We apply a simple version of the Hardy-Littlewood (circle) method, as will be familiar to ~~those~~ who took AHA. We first investigate

$$r_3(n) := \sum_{p_1+p_2+p_3=n} (\log p_1)(\log p_2)(\log p_3).$$

Plainly, if $r_3(n) \to \infty$ as $\underset{\text{odd}}{n} \to \infty$, then all large odd integers are the sum of 3 primes.

Define

$$f(\alpha) = \sum_{p \le N} (\log p)\, e(p\alpha),$$

in which we suppose ~~that~~ $n \le N$. Then, by orthogonality, we have

$$\int_0^1 f(\alpha)^3 e(-n\alpha)\,d\alpha = \sum_{p_1, p_2, p_3 \le N} (\log p_1)(\log p_2)(\log p_3) \underbrace{\int_0^1 e(\alpha(p_1+p_2+p_3-n))\,d\alpha}_{= \begin{cases} 0, & \text{when } p_1+p_2+p_3 \ne n, \\ 1, & \text{when } p_1+p_2+p_3 = n. \end{cases}}$$

$$= r_3(n).$$

It thus remains asymptotically to calculate the integral on the left hand side, and here we are able to make use of Theorem 7.1.

Let B be a large positive number, and write $Q = (\log N)^B$, in which we fix $N = 3n$. Also, define the set of <u>major arcs</u>

$$\mathfrak{M} = \bigcup_{\substack{0 \le a \le q \le Q \\ (a,q)=1}} \mathfrak{M}(q,a),$$

where

$$\mathfrak{M}(q,a) = \{ \alpha \in [0,1) : |\alpha - \tfrac{a}{q}| \le QN^{-1} \}.$$

Notice that when $\frac{a_1}{q_1}$ and $\frac{a_2}{q_2}$ are distinct with $(a_i, q_i) = 1$ and $q_i \le Q$, one has

$$\frac{1}{q_1 q_2} \le \frac{|a_1 q_2 - a_2 q_1|}{q_1 q_2} \le \left| \frac{a_1}{q_1} - \frac{a_2}{q_2} \right|, \quad \text{whence} \quad \left| \frac{a_1}{q_1} - \frac{a_2}{q_2} \right| \gg \frac{1}{Q^2} = (\log N)^{-2B}$$

Thus $\mathfrak{M}(q_1, a_1) \cap \mathfrak{M}(q_2, a_2) = \emptyset$. Finally, we define the set of minor arcs

$$\mathfrak{m} = [0,1) \setminus \mathfrak{M}.$$

We then have

$$r_3(n) = \int_{\mathfrak{M}} f(\alpha)^3 e(-n\alpha)\, d\alpha + \int_{\mathfrak{m}} f(\alpha)^3 e(-n\alpha)\, d\alpha.$$

<u>Lemma 8.1.</u> One has

$$\sup_{\alpha \in \mathfrak{m}} |f(\alpha)| \ll N (\log N)^{(5-B)/2}.$$

<u>Proof.</u> Suppose that $\alpha \in \mathfrak{m}$. By Dirichlet's theorem on Diophantine approximation, there exists $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a,q)=1$ and $1 \le q \le N/Q$ such that $|\alpha - a/q| \le 1/(qN/Q) \le Q/N$. Since $\alpha \in \mathfrak{m}$, moreover, one <u>cannot</u> have $q \le Q$, for then $\alpha \in \mathfrak{M}(q,a) \subseteq \mathfrak{M}$. Thus

$$Q < q \le N/Q.$$

We therefore deduce from Theorem 7.1 that

$$\sum_{1 \le n \le N} \Lambda(n) e(n\alpha) \ll N (\log N)^{5/2} \left(q^{-1} + N^{-2/5} + qN^{-1} \right)^{1/2}$$

$$\ll N(\log N)^{5/2}\left(Q^{-1}+N^{-2/5}+(N/Q)N^{-1}\right)^{1/2}$$

$$\ll NQ^{-1/2}(\log N)^{5/2}.$$

But

$$\sum_{p\le N}(\log p)e(p\alpha) = \sum_{1\le n\le N}\Lambda(n)e(n\alpha) + O\Bigg(\sum_{p\le N^{1/2}}\sum_{\substack{h=2\\(n=p^h,\ h\ge 2)}}^{2\log N}\log p\Bigg)$$

$$\ll N(\log N)^{(5-B)/2} + O(N^{1/2}\log N)$$

$$\ll N(\log N)^{(5-B)/2}.$$

The desired conclusion is now immediate. //

This estimate allows us to take care of the minor arc contribution.

<u>Lemma 8.2.</u> One has

$$\int_{\mathfrak{m}} f(\alpha)^3 e(-n\alpha)\,d\alpha \ll N^2(\log N)^{(7-B)/2}.$$

<u>Proof.</u> By the triangle inequality,

$$\int_{\mathfrak{m}} f(\alpha)^3 e(-n\alpha)\,d\alpha \ll \sup_{\alpha\in\mathfrak{m}}|f(\alpha)|\int_0^1|f(\alpha)|^2\,d\alpha.$$

By orthogonality,

$$\int_0^1|f(\alpha)|^2\,d\alpha = \sum_{\substack{p_1,p_2\le N\\p_1=p_2}}(\log p_1)(\log p_2)$$

$$= \sum_{p\le N}(\log p)^2 \ll \frac{N}{\log N}\cdot(\log N)^2.$$

Thus, on applying Lemma 8.1, we see that

$$\int_{\mathfrak{m}} f(\alpha)^3 e(-n\alpha)\,d\alpha \ll (N\log N)\cdot\left(N(\log N)^{(5-B)/2}\right). \quad //$$

It remains now to asymptotically estimate the contribution of the major arcs $\int_{\mathfrak{M}} f(\alpha)^3 e(-n\alpha)\,d\alpha$.

We introduce the simple function
$$v(\beta) = \sum_{m=1}^{N} e(m\beta).$$

Lemma 8.3. There is a positive constant c with the property ~~that~~, whenever $\alpha \in \mathfrak{M}(q,a) \subseteq \mathfrak{M}$, one has

$$f(\alpha) = \frac{\mu(q)}{\varphi(q)} v(\alpha - a/q) + O(N \exp(-c\sqrt{\log N})).$$

Proof. Since $\alpha \in \mathfrak{M}(q,a) \subseteq \mathfrak{M}$, we may suppose ~~that~~ $0 \le a \le q \le Q$, $(a,q) = 1$ and $|\alpha - a/q| \le QN^{-1}$. Our first step is to investigate the situation when $\alpha = a/q$. Put

$$f(\alpha; t) = \sum_{p \le t} (\log p) e(p\alpha).$$

Then one has, for $t \le N$,

$$f(a/q; t) = \sum_{\substack{r=1 \\ (r,q)=1}}^{q} \sum_{\substack{p \le t \\ p \equiv r \,(\mathrm{mod}\, q)}} (\log p)\, e(ra/q) + O((\log N)(\log q))$$

contribution from $p|q$.

$$\cdots = \sum_{\substack{r=1 \\ (r,q)=1}}^{q} e(ra/q) \cdot \theta(t; q, r) + O((\log N)(\log q)).$$

But

$$\theta(t; q, r) = \frac{t}{\varphi(q)} + O(N \exp(-c\sqrt{\log N})).$$

(some $c > 0$)

Thus,

$$f\left(\frac{a}{q}; t\right) = \frac{t}{\varphi(q)} \underbrace{\sum_{\substack{r=1 \\ (r,q)=1}}^{q} e(ar/q)}_{\text{Ramanujan's sum (exercise sheet on ANT 1)} \; = \mu(q).} + O(N \exp(-\tfrac{1}{2}c\sqrt{\log N})).$$

$$\Rightarrow \quad f\left(\frac{a}{q}; t\right) = \frac{\mu(q)}{\varphi(q)} t + O(N \exp(-\tfrac{1}{2}c\sqrt{\log N})).$$

We now apply Riemann - Stieltjes integration to move from a/q to $\alpha = \beta + a/q$, with $|\beta| \le QN^{-1}$. Thus,

$$
\begin{aligned}
f(\beta + a/q) &= \sum_{p \le N} (\log p)\, e\left(p\tfrac{a}{q}\right) \cdot e(p\beta) \\
&= \int_{2^-}^{N} e(t\beta)\, d f\left(\tfrac{a}{q}; t\right) \\
&= \left[f\left(\tfrac{a}{q}; t\right) e(t\beta) \right]_{2^-}^{N} - \int_{2^-}^{N} f\left(\tfrac{a}{q}; t\right) 2\pi i \beta\, e(t\beta)\, dt \\
&= f\left(\tfrac{a}{q}; N\right) e(N\beta) - 2\pi i \beta \cdot \left(\frac{\mu(q)}{\varphi(q)} \int_{2^-}^{N} t\, e(t\beta)\, dt \right. \\
&\qquad\qquad \left. + O\left(\int_{2^-}^{N} N \exp\left(-\tfrac{1}{2} c \sqrt{\log N}\right) dt \right) \right).
\end{aligned}
$$

But

$$
\begin{aligned}
2\pi i \beta \int_{2^-}^{N} t\, e(t\beta)\, dt &= 2\pi i \beta \int_{2^-}^{N} \Big(\sum_{n \le t} 1 + O(1) \Big) e(t\beta)\, dt \\
&= O(N|\beta|) + \Big[e(t\beta) \sum_{n \le t} 1 \Big]_{2^-}^{N} - \int_{2^-} e(t\beta)\, d\Big(\sum_{n \le t} 1 \Big) \\
&= N e(N\beta) + O(Q) - \underbrace{\sum_{n \le N} e(n\beta)}_{v(\beta)},
\end{aligned}
$$

Whence

$$
\begin{aligned}
f(\beta + a/q) &= \left(f\left(\tfrac{a}{q}; N\right) - N \frac{\mu(q)}{\varphi(q)} \right) e(N\beta) + O\left(QN \exp\left(-\tfrac{1}{2} c \sqrt{\log N}\right) \right) \\
&\quad + \frac{\mu(q)}{\varphi(q)} v(\beta) \\
&= \frac{\mu(q)}{\varphi(q)} v(\beta) + O\left(N \exp\left(-\tfrac{1}{3} c \sqrt{\log N}\right) \right),
\end{aligned}
$$

since $Q \le (\log N)^{B} \le \exp\left(+\tfrac{1}{6} c \sqrt{\log N}\right)$. //

Lemma 8.4. One has

$$\int_{\mathfrak{M}} f(\alpha)^3 e(-n\alpha)\,d\alpha = \tfrac{1}{2} n^2 \mathfrak{S}(n) + O\left(n^2 (\log n)^{-B/2}\right),$$

where

$$\mathfrak{S}(n) = \prod_{p \mid n} \left(1 - \frac{1}{(p-1)^2}\right) \cdot \prod_{p \nmid n} \left(1 + \frac{1}{(p-1)^3}\right).$$

Proof. When $\alpha \in \mathfrak{M}(q,a) \subseteq \mathfrak{M}$, one has

$$f(\alpha)^3 - \left(\frac{\mu(q)}{\varphi(q)} v(\alpha - a/q)\right)^3 \ll N^3 \exp(-c\sqrt{\log N}),$$

whence

$$\int_{\mathfrak{M}} f(\alpha)^3 e(-n\alpha)\,d\alpha = \sum_{1 \le q \le Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} e(na/q) \int_{\mathfrak{M}(q,a)} \frac{\mu(q)}{\varphi(q)^3} v(\alpha - a/q)^3 e(-n(\alpha - a/q))\,d\alpha$$

$$+ O\left(Q^3 N^{-1} \cdot N^3 \exp(-c\sqrt{\log N})\right)$$

$$= \mathfrak{S}(n,Q) \int_{-Q/N}^{Q/N} v(\beta)^3 e(-n\beta)\,d\beta$$

$$+ O\left(N^2 \exp(-\tfrac{1}{2} c\sqrt{\log N})\right),$$

where

$$\mathfrak{S}(n,Q) = \sum_{q=1}^{Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \frac{\mu(q)}{\varphi(q)^3} e\left(-\frac{na}{q}\right).$$

One has $v(\beta) = \sum_{n=1}^{N} e(n\beta) \ll \min\{N, \|\beta\|^{-1}\}$, whence

$$\int_{-Q/N}^{Q/N} v(\beta)^3 e(-n\beta)\,d\beta = \int_{-\frac12}^{\frac12} v(\beta)^3 e(-n\beta)\,d\beta + O\left(\int_{Q/N}^{\infty} \beta^{-3}\,d\beta\right)$$

$$= \#\{n_1+n_2+n_3=n : 1\le n_i\le N\} + O(N^2/Q^2)$$

$$= \tfrac{1}{2}n^2 + O\left(N^2 (\log N)^{-2B}\right).$$

Also,

$$\mathfrak{S}(n,Q) = \sum_{q=1}^{\infty} \frac{\mu(q)}{\varphi(q)^3} \sum_{\substack{a=1\\(a,q)=1}}^{q} e\left(-\frac{na}{q}\right) + O\left(\sum_{q=Q+1}^{\infty} \frac{1}{\varphi(q)^2}\right)$$

$$= \sum_{q=1}^{\infty} \underbrace{\frac{\mu(q)}{\varphi(q)^3}\cdot\frac{\mu(q/(q,n))\varphi(q)}{\varphi(q/(q,n))}}_{\text{multiplicative}} + O\left(\frac{1}{Q^{1-\varepsilon}}\right)$$

$$= \prod_{p\mid n}\left(1-\frac{1}{(p-1)^3}\cdot\frac{p-1}{\text{[redacted]}1}\right)\prod_{p\nmid n}\left(1-\frac{1}{(p-1)^3}\cdot\left(-\frac{(p-1)}{(p-1)}\right)\right) + O(Q^{\varepsilon-1})$$

$$= \mathfrak{S}(n) + O\left((\log N)^{-B/2}\right).$$

We may ~~therefore~~ conclude ~~that~~

$$\int_{\mathfrak{M}} f(\alpha)^3 e(-n\alpha)\,d\alpha = \left(\tfrac{1}{2}n^2 + O\left(N^2(\log N)^{-2B}\right)\right) \times \left(\mathfrak{S}(n) + O\left((\log N)^{-B/2}\right)\right)$$

$$= \tfrac{1}{2}n^2\mathfrak{S}(n) + O\left(N^2(\log N)^{-B/2}\right). \quad //$$

Notice ~~that~~ when n is even, one has

$$\mathfrak{S}(n) = \left(1-\frac{1}{(2-1)^2}\right)\prod_{p>3}\cdots = 0,$$

and when n is odd,

$$\mathfrak{S}(n) \asymp 1 \quad \text{by comparison with} \quad \prod_p\left(1+\frac{8}{p^3}\right).$$

Finally, by combining Lemmata 8.2 and 8.4, we deduce that

$$r_3(n) = \int_{\mathfrak{M}} f(\alpha)^3 e(-n\alpha)\,d\alpha + \int_{\mathfrak{m}} f(\alpha)^3 e(-n\alpha)\,d\alpha$$

$$= \tfrac{1}{2}\mathfrak{S}(n)\,n^2 + O(N^2(\log N)^{-B/2}) + O(N^2(\log N)^{(7-B)/2})$$

$$= \left(\mathfrak{S}(n) + o(1)\right)\frac{n^2}{2}.$$

<u>Theorem 8.5.</u> Whenever $A>0$, one has

$$r_3(n) = \tfrac{1}{2}n^2\,\mathfrak{S}(n) + O(n^2(\log n)^{-A}),$$

where

$$\mathfrak{S}(n) = \left(\prod_{p \mid n}\left(1 - \frac{1}{(p-1)^2}\right)\right)\cdot\left(\prod_{p \nmid n}\left(1 + \frac{1}{(p-1)^3}\right)\right).$$

We next consider the binary Goldbach problem. Here we aim only for an "almost-all" results, and consider what can be said in mean square. Define

$$r_2(n) := \sum_{p_1+p_2=n} (\log p_1)(\log p_2).$$

A formal application of the method ~~that~~ we applied in considering the ternary Goldbach problem ~~suggests that~~

$$r_2(n) = \int_{\mathfrak{M}} f(\alpha)^2 e(-n\alpha)\,d\alpha + \int_{\mathfrak{m}} f(\alpha)^2 e(-n\alpha)\,d\alpha,$$

in which $\mathfrak{M}$ and $\mathfrak{m}$ are defined as before (this much is true!), wherein

$$\int_{\mathfrak{m}} f(\alpha)^2 e(-n\alpha)\,d\alpha = o(N) \qquad \text{(true on average!)}$$

and

$$\int_{\mathfrak{M}} f(\alpha)^2 e(-n\alpha)\,d\alpha = \mathfrak{S}'(n)\,n + o(n),$$

where $\mathfrak{S}'(n)$ is a suitably convergent singular series. The analysis of

$\mathfrak{S}'(n)$ takes some care owing to the delicacy of the convergence issues. We begin by proving that the minor arc contribution is small on average.

<u>Lemma 8.6.</u> One has, for each large number N,

$$\sum_{N/2 < n \leq N} \left| \int_{\mathfrak{m}} f(\alpha)^2 e(-n\alpha)\, d\alpha \right|^2 \ll N^3 (\log N)^{6-B}.$$

<u>Proof.</u> It follows from Bessel's inequality that

$$\begin{aligned}\sum_n \left| \int_{\mathfrak{m}} f(\alpha)^2 e(-n\alpha)\, d\alpha \right|^2 &\leq \int_{\mathfrak{m}} |f(\alpha)|^4\, d\alpha \\ &\ll \Big(\sup_{\alpha\in\mathfrak{m}} |f(\alpha)|\Big)^2 \int_0^1 |f(\alpha)|^2\, d\alpha \\ &\ll \left(N(\log N)^{(5-B)/2}\right)^2 N\log N \\ &\ll N^3 (\log N)^{6-B}. \quad // \qquad \text{(cf. Lemmata 8·1, 8·2).}\end{aligned}$$

Next we show that the major arc contribution is well-behaved on average to complement this conclusion, which shows that for all but at most $N(\log N)^{-B/2}$ integers n with $N/2 < n \leq N$, one has

$$\int_{\mathfrak{m}} f(\alpha)^2 e(-n\alpha)\, d\alpha \ll N(\log N)^{3-B/4}.$$

<u>Lemma 8.7.</u> One has

$$\int_{\mathfrak{M}} f(\alpha)^2 e(-n\alpha)\, d\alpha = n\mathfrak{S}'(n,Q) + O(N(\log N)^{1-B}), \quad \text{for } \tfrac{N}{2} < n \leq N,$$

where

$$\mathfrak{S}'(n,Q) = \sum_{1\leq q\leq Q}{}' \sum_{\substack{a=1\\(a,q)=1}}^{q} \frac{\mu(q)^2}{\varphi(q)^2}\, e(-an/q).$$

<u>Proof:</u> We apply Lemma 8·3 as in the proof of Lemma 8·4. Thus, when $\alpha \in \mathfrak{M}(q,a) \subseteq \mathfrak{M}$, one has

$$f(\alpha)^2 - \left(\frac{\mu(q)}{\varphi(q)} v(\alpha - a/q)\right)^2 \ll N^2 \exp(-c\sqrt{\log N}),$$

whence

$$\int_{\mathfrak{M}} f(\alpha)^2 e(-n\alpha)\,d\alpha = \sum_{1\le q\le Q} \sum_{\substack{a=1\\(a,q)=1}}^{q} e(-na/q) \int_{\mathfrak{M}(q,a)} \frac{\mu(q)^2}{\varphi(q)^2} v(\alpha-a/q)^2 e(-n(\alpha-a/q))\,d\alpha$$

$$+ \; O(Q^3N^{-1}\cdot N^2 \exp(-c\sqrt{\log N}))$$

$$= \mathfrak{S}'(n,Q) \int_{-Q/N}^{Q/N} v(\beta)^2 e(-n\beta)\,d\beta + O(N\exp(-\tfrac{1}{2}c\sqrt{\log N})).$$

One has

$$\int_{-Q/N}^{Q/N} v(\beta)^2 e(-n\beta)\,d\beta = \int_{-\frac12}^{\frac12} v(\beta)^2 e(-n\beta)\,d\beta + O\left(\int_{Q/N}^{\infty} \beta^{-2}\,d\beta\right)$$

$$= \#\{n_1+n_2=n : 1\le n_i\le N\} + O(N/Q)$$

$$= n + O(N(\log N)^{-B}).$$

Hence

$$\int_{\mathfrak{M}} f(\alpha)^2 e(-n\alpha)\,d\alpha = n\mathfrak{S}'(n,Q) + O(N\exp(-\tfrac{1}{2}c\sqrt{\log N}))$$

$$+ \; O\left(N(\log N)^{-B} \sum_{1\le q\le Q} \frac{1}{\varphi(q)}\right)$$

$$= n\mathfrak{S}'(n,Q) + O(N(\log N)^{1-B}). \quad //$$

What remains is to understand the behaviour of the singular series $\mathfrak{S}'(n,Q)$ on average over n.

<u>Lemma 8.8.</u> The singular series

$$\mathfrak{S}'(n) = \sum_{q=1}^{\infty} \frac{\mu(q)^2}{\varphi(q)^2} \sum_{\substack{a=1\\(a,q)=1}}^{q} e(-an/q)$$

converges, and one has

$$\sum_{N/2<n\le N} |\mathfrak{S}'(n,Q) - \mathfrak{S}'(n)|^2 \ll N(\log N)^{2-B}.$$

53

Proof:

When $Y > X$, one has

$$\sum_{X<q\le Y} \frac{\mu(q)^2}{\varphi(q)^2} \sum_{\substack{a=1\\(a,q)=1}}^{q} e\left(-\frac{an}{q}\right) = \sum_{X<q\le Y} \frac{\mu(q)^2}{\varphi(q)^2} \frac{\mu(q/(q,n))\,\varphi(q)}{\varphi(q/(q,n))}$$

$$= \sum_{d|n} \frac{\mu(d)^2}{\varphi(d)} \sum_{\substack{X/d<r\le Y/d\\(r,d)=1}} \frac{\mu(r)}{\varphi(r)^2}$$

$$\ll \sum_{d|n} \frac{\mu(d)^2}{\varphi(d)} \min\left\{\frac{d}{X}, 1\right\}. \qquad \boxed{\begin{array}{l} d=(q,n),\ r=q/d. \\ (r,d)>1 \Rightarrow \mu(q)=\mu(rd)=0 \end{array}}$$

$$\ll_n \frac{1}{X}.$$

Then it follows ~~that~~ $\mathfrak{S}'(n)$ converges. Moreover, one has

$$\sum_{N/2<n\le N} |\mathfrak{S}'(n,Q) - \mathfrak{S}'(n)|^2 \ll \max_{N/2<n\le N} |\mathfrak{S}'(n,Q)-\mathfrak{S}'(n)| \cdot \sum_{N/2<n\le N} |\mathfrak{S}'(n,Q)-\mathfrak{S}'(n)|,$$

and here

$$\sum_{N/2<n\le N} |\mathfrak{S}'(n,Q)-\mathfrak{S}'(n)| \ll \sum_{N/2<n\le N} \sum_{d|n} \frac{\mu(d)^2}{\varphi(d)} \min\left\{\frac{d}{Q}, 1\right\}$$

$$\ll \sum_{d\le N} \frac{\mu(d)^2}{\varphi(d)} \min\left\{\frac{d}{Q}, 1\right\} \sum_{m\le N/d} 1$$

$$\ll \frac{N}{Q} \sum_{d\le N} \frac{\mu(d)^2}{\varphi(d)} \ll N(\log N)\, Q^{-1}.$$

Also,

$$\max_{N/2<n\le N} |\mathfrak{S}'(n,Q) - \mathfrak{S}'(n)| \ll \sum_{d|n} \frac{\mu(d)^2}{\varphi(d)} \min\left\{\frac{d}{Q}, 1\right\}$$

$$\ll \cancel{Q^{-1} \sum_{\substack{d|n\\ d\le Q}} \frac{\mu(d)^2}{\varphi(d)}\cdot d} + \sum_{d\le n} \frac{\mu(d)^2}{\varphi(d)} \ll \log n.$$

Hence

$$\sum_{N/2<n\le N} |\mathfrak{S}'(n,Q) - \mathfrak{S}'(n)|^2 \ll N(\log N)^2 Q^{-1} \ll N(\log N)^{2-B}.$$

It follows from Lemma 8·6 that for all but $N(\log N)^{-B/2}$ integers n with $N/2 < n \le N$, one has

$$\int_{\mathfrak{m}} f(\alpha)^2 e(-n\alpha)\,d\alpha \ll N(\log N)^{3-B/4}.$$

Similarly, for all but $N(\log N)^{-B/2}$ integers n with $N/2 < n \le N$, it follows from Lemma 8·8 ~~that~~

$$|\mathfrak{S}'(n,Q) - \mathfrak{S}'(n)| \ll (\log N)^{1-B/4}.$$

Thus, for all but $O(N(\log N)^{-B/2})$ integers with $N/2 < n \le N$, one has

$$r_2(n) = \int_{\mathfrak{M}} f(\alpha)^2 e(-n\alpha)\,d\alpha + \int_{\mathfrak{m}} f(\alpha)^2 e(-n\alpha)\,d\alpha$$

$$= n\mathfrak{S}'(n) + O(N(\log N)^{-C}), \qquad \text{(Lemma 8·7)}.$$

where $C = \min\{\frac{B}{4}-1, \frac{B}{4}-3, B-1\}$.

But

$$\mathfrak{S}'(n) = \sum_{q=1}^{\infty} \frac{\mu(q)^2}{\varphi(q)^2} \sum_{\substack{a=1\\(a,q)=1}}^{q} e\left(-\frac{an}{q}\right) = \sum_{q=1}^{\infty} \frac{\mu(q)^2}{\varphi(q)^2} \frac{\mu(q/(q,n))}{\varphi(q/(q,n))}\varphi(q)$$

$$= \prod_{p \nmid n}\left(1 - \frac{1}{(p-1)^2}\right) \cdot \prod_{p \mid n}\left(1 + \frac{1}{p-1}\right) \gg 1.$$

Thus $r_2(n) \gg n$, with the exception of $O(N(\log N)^{-B/2})$ integers n with $N/2 < n \le N$, provided ~~that~~ B is taken sufficiently large.
Note that $\mathfrak{S}'(n) = 0$ if n is odd, in which case $2 \nmid n$.

<u>Theorem 8.9.</u> One has, for any $A > 0$,

$$E(X) := \text{card}\{1 \le m \le X : 2m \text{ is } \underline{\text{not}} \text{ the sum of two primes}\}$$

$$\ll X(\log X)^{-A}.$$

<u>Proof:</u> We have shown that $E(X) - E(X/2) \ll X(\log X)^{-B/2}$. By

Summing over dyadic intervals one sees that

$$E(X) - E(\sqrt{X}) \ll X(\log X)^{1-B/2},$$

whence $E(X) \ll \sqrt{X} + X(\log X)^{1-B/2} \ll X(\log X)^{-A}$,

provided ~~that~~ $B \geq 2A+1$. //

§9. Incomplete character sums, I: the Pólya-Vinogradov inequality.

Let χ be a character modulo q. The sum $\sum_{n=M+1}^{M+N} \chi(n)$ is called an incomplete character sum. Of course, if $N=q$ then we have a sum that is either 0 or $\phi(q)$, so incomplete character sums remain to be estimated. If the $\chi(n)$ were essentially randomly distributed amongst their allowable value set with $\chi \neq \chi_0$, then we might hope for square-root cancellation (cf. "random walk"). We are able very nearly to achieve such a conclusion for $N \asymp q$.

Theorem 9.1. (The Pólya-Vinogradov inequality, 1918). Let χ be a non-principal character modulo q. Then for any integers M and N with $N>0$, one has

$$\sum_{n=M+1}^{M+N} \chi(n) \ll \sqrt{q}\log q.$$

Proof. We begin by considering the situation in which χ is a primitive character modulo q with $q>1$. By Theorem 4.6 (i), we have

$$\chi(n) = \frac{1}{\tau(\bar{\chi})} \sum_{a=1}^{q} \bar{\chi}(a)\, e(an/q),$$

whence

$$\sum_{n=M+1}^{M+N} \chi(n) = \frac{1}{\tau(\bar{\chi})} \sum_{a=1}^{q} \bar{\chi}(a) \sum_{n=M+1}^{M+N} e(an/q).$$

But when $a/q \notin \mathbb{Z}$, we have

$$\sum_{n=M+1}^{M+N} e(an/q) = \frac{e((M+N+1)^a/q) - e((M+1)^a/q)}{e(a/q) - 1}$$

$$= e\left((2M+N+1)^a/_{2q}\right) \frac{\sin(\pi Na/q)}{\sin(\pi a/q)}.$$

Thus, since $\bar{\chi}(a) = 0$ when $q \mid a$, we see ~~that~~

$$\left|\sum_{n=M+1}^{M+N} \chi(n)\right| < \frac{1}{|\tau(\bar{\chi})|} \sum_{\substack{a=1\\(a,q)=1}}^{q-1} \frac{1}{\sin(\pi a/q)} = \frac{1}{\sqrt{q}} \sum_{\substack{a=1\\(a,q)=1}}^{q-1} \frac{1}{\sin(\pi a/q)}.$$

(Theorem 4·5)

Since $\sin(\pi a/q)$ is symmetric around $a = q/2$, we can economise slightly. When $q \equiv 2 \pmod 4$ there is no primitive character modulo q, and thus $4 \mid q$, whence $(q/2, q) > 1$. Hence

$$\left|\sum_{n=M+1}^{M+N} \chi(n)\right| < \frac{2}{\sqrt{q}} \sum_{a=1}^{(q-1)/2} \frac{1}{\sin(\pi a/q)} \leq \frac{2}{\sqrt{q}} \sum_{a=1}^{(q-1)/2} \frac{1}{2a/q}$$

$$= \sqrt{q} \sum_{a=1}^{(q-1)/2} \frac{1}{a} < \sqrt{q} \sum_{a=1}^{(q-1)/2} \log\left(\frac{1+1/2a}{1-1/2a}\right)$$

$$= \sqrt{q} \sum_{a=1}^{(q-1)/2} \left(\log(2a+1) - \log(2a-1)\right) = \sqrt{q} \log q.$$

This conclusion, valid for primitive characters, is easily extended to imprimitive non-principal characters. For suppose ~~that~~ χ is induced by χ^* modulo d, and let

$$r = \prod_{\substack{p \mid q\\ p \nmid d}} p.$$

Then one has

$$\sum_{n=M+1}^{M+N} \chi(n) = \sum_{\substack{n=M+1\\(n,r)=1}}^{M+N} \chi^*(n) = \sum_{n=M+1}^{M+N} \chi^*(n) \sum_{k \mid (n,r)} \mu(k)$$

$$= \sum_{k|r} \mu(k) \sum_{\substack{M+1 \le n \le M+N \\ k|n}} \chi^*(n)$$

$$= \sum_{k|r} \mu(k)\, \chi^*(k) \sum_{\frac{M+1}{k} \le m \le \frac{M+N}{k}} \chi^*(m) .$$

Thus $\displaystyle \left| \sum_{n=M+1}^{M+N} \chi(n) \right| \le \sum_{k|r} |\mu(k)\chi^*(k)| \sqrt{d} \log d$

$$\ll 2^{\omega(r)} \sqrt{d} \log d .$$

But $\displaystyle \frac{2^{\omega(r)}}{r^{1/2}} \le \prod_{p|r} \frac{2}{\sqrt{p}} \le \frac{2}{\sqrt{2}} \cdot \frac{2}{\sqrt{3}} = \frac{4}{\sqrt{6}} < 2$. Thus

$$\left| \sum_{n=M+1}^{M+N} \chi(n) \right| < 2\sqrt{dr} \log q \le 2\sqrt{q} \log q . \quad //$$

We shall make use of the Pólya - Vinogradov inequality later in our treatment of the Bombieri - Vinogradov theorem. The non-trivial cancellation available from Theorem 9.1 has other interesting consequences.

<u>Corollary 9.2.</u> Let χ be a non-principal character modulo p, and let n_χ denote the least positive integer having the property that $\chi(n) \ne 1$. Then one has $n_\chi \ll_\varepsilon p^{1/2\sqrt{e} + \varepsilon}$.

<u>Proof.</u> We apply an argument considered in ANT1. Suppose that $\chi(n) = 1$ for $1 \le n \le y$, and seek to choose y large enough that we contradict the cancellation exhibited in the Pólya - Vinogradov inequality. Thus, $\chi(n) = 1$ whenever n is a product of prime numbers π with $\pi \le y$. Recall that

$$\psi(x,y) = \sum_{\substack{1 \le n \le x \\ \pi \mid n \Rightarrow \pi \le y}} 1 .$$

Then, provided that $y \le x < y^2$, one has

$$\sum_{1 \le n \le x} \chi(n) = \sum_{\substack{1 \le n \le x \\ \pi \mid n \Rightarrow \pi \le y}} \chi(n) + \sum_{y < \pi \le x} \sum_{\substack{1 \le n \le x \\ \pi \mid n}} \chi(n)$$

$$= \psi(x,y) + \sum_{y < \pi \le x} \chi(\pi) \left\lfloor \frac{x}{\pi} \right\rfloor ,$$

Whence

$$\Big| \sum_{1 \le n \le x} \chi(n) \Big| \geqslant \psi(x,y) - \sum_{y < \pi \le x} \left\lfloor \frac{x}{\pi} \right\rfloor$$

$$= \lfloor x \rfloor - 2 \sum_{y < \pi \le x} \left\lfloor \frac{x}{\pi} \right\rfloor$$

$$= x \Big(1 - 2 \sum_{y < \pi \le x} \frac{1}{\pi} \Big) + O\left(\frac{x}{\log x} \right)$$

$$= x \left(1 - 2 \log\left(\frac{\log x}{\log y} \right) \right) + O\left(\frac{x}{\log x} \right).$$

Then by the Pólya-Vinogradov inequality, if we take $x = p^{1/2} (\log p)^2$, we see ~~that~~ if one were to have $\chi(n) = 1$ for $1 \le n \le y$,

$$\underset{\gg p^{1/2} \log p.}{\frac{x}{\log x}} \gg x \left(1 - 2 \log\left(\frac{\log x}{\log y} \right) \right) + O\left(\frac{x}{\log x} \right),$$

Whence

$$2 \log\left(\frac{\log x}{\log y} \right) \geq 1 + O\left(\frac{1}{\log x} \right)$$

$$\Rightarrow \quad \frac{\log x}{\log y} \geq e^{1/2} + O\left(\frac{1}{\log x} \right)$$

$$\Rightarrow \quad \log y \le e^{-1/2} \log x + O(1) .$$

But then we have $y \ll x^{4/\sqrt{e}} \ll p^{\frac{1}{2\sqrt{e}}+\varepsilon}$. Consequently, whenever $y > p^{\frac{1}{2\sqrt{e}}+2\varepsilon}$, then we must have $\chi(n) \neq 1$ for some integer n with $1 \leq n \leq y$ (and $(y,p)=1$). Hence $n_\chi \ll p^{\frac{1}{2\sqrt{e}}+2\varepsilon}$ /

Conjecture. (Vinogradov) One has $n_\chi \ll p^\varepsilon$. (any $\varepsilon > 0$).

Corollary 9.3. Let $Z(M,N;p)$ denote the number of primitive roots modulo p in the interval $[M+1, M+N]$. Then one has

$$Z(M,N;p) = \frac{\varphi(p-1)}{p} N + O(p^{1/2+\varepsilon}).$$

Proof. We prepare a detector function for primitive roots. Let $\pi_1, \dots, \pi_r$ be the distinct prime factors of $p-1$, and put

$$q = \prod_{i=1}^{r} \pi_i .$$

Then n is a primitive root modulo p if and only if

$$(\mathrm{ind}_g\, n, q) = 1,$$

for any fixed primitive root g modulo p. For $(1 \leq i \leq r)$ we put

$$\chi_i(n) = e\left(\frac{\mathrm{ind}_g\, n}{\pi_i}\right).$$

Then we have

$$\frac{1}{\pi_i} \sum_{a=1}^{\pi_i} \chi_i(n)^a = \begin{cases} 1, & \text{when } \pi_i \mid \mathrm{ind}_g\, n, \\ 0, & \text{when } \pi_i \nmid \mathrm{ind}_g\, n. \end{cases}$$

Consequently, one has

$$\prod_{i=1}^{r} \left(\chi_0(n) - \frac{1}{\pi_i} \sum_{a_i=1}^{\pi_i} \chi_i(n)^{a_i} \right) = \begin{cases} 1, & \text{when } n \text{ is primitive (mod } p), \\ 0, & \text{when } n \text{ is not primitive (mod } p). \end{cases}$$

Observe that this detector may be rewritten in the shape

$$\prod_{i=1}^{r}\left(\left(1-\frac{1}{\pi_i}\right)\chi_0(n)-\frac{1}{\pi_i}\sum_{a_i=1}^{\pi_i-1}\chi_i^{a_i}(n)\right)=\prod_{i=1}^{r}\left(\frac{\varphi(\pi_i)}{\pi_i}\chi_0(n)+\frac{\varphi(1)}{1}\cdot\frac{\mu(\pi_i)}{\pi_i}\times\sum_{\substack{\chi\\ \operatorname{ord}\chi=\pi_i}}\chi(n)\right)$$

$$=\sum_{d\mid q}\frac{\varphi(q/d)}{q/d}\cdot\frac{\mu(d)}{d}\sum_{\substack{\chi\\ \operatorname{ord}\chi=d}}\chi(n).$$

By applying this primitive root detector, we see that

$$\Sigma(M,N;p)=\frac{1}{q}\sum_{d\mid q}\varphi(q/d)\,\mu(d)\sum_{\substack{\chi\\ \operatorname{ord}\chi=d}}\ \sum_{n=M+1}^{M+N}\chi(n).$$

The main term hoe with $d=1$ is given by the principal character χ_0, which contributes

$$\frac{\varphi(q)}{q}\sum_{n=M+1}^{M+N}\chi_0(n)=\frac{\varphi(q)}{q}\left(\left(1-\frac{1}{p}\right)N+O(1)\right)$$

$$=\frac{\varphi(p-1)}{p}N+O(1).$$

Meanwhile, any character with $d>1$ is non-principal, and so the Pólya-Vinogradov inequality may be applied for such terms to show that

$$\Sigma(M,N;p)=\frac{\varphi(p-1)}{p}N+O\Bigg(\frac{1}{q}\sum_{\substack{d\mid q\\ d>1}}\varphi(q/d)|\mu(d)|\underbrace{\sum_{\substack{\chi\\ \operatorname{ord}\chi=d}}}_{\varphi(d)\text{ chars.}}p^{1/2}\log p\Bigg)$$

$$=\frac{\varphi(p-1)}{p}N+O\left(\frac{\varphi(q)}{q}p^{1/2}(\log p)\sum_{d\mid(p-1)}|\mu(d)|\right)$$

$$=\frac{\varphi(p-1)}{p}N+O\Big(\underbrace{2^{\omega(p-1)}p^{1/2}\log p}_{\ll\, p^{1/2+\varepsilon}}\Big)$$

//

The conclusion of Corollary 9.3 shows that primitive roots are equidistributed modulo p.

We can derive lower bounds to complement the Pólya–Vinogradov inequality.

<u>Theorem 9.4.</u> Suppose ~~that~~ χ is a non-principal character modulo q. Then one has

$$\max_{M,N} \left| \sum_{n=M+1}^{M+N} \chi(n) \right| \geq \frac{|\tau(\chi)|}{\pi} = \frac{\sqrt{q}}{\pi}.$$

<u>Proof.</u> One has

$$\left| \sum_{M=1}^{q} e(M/q) \sum_{n=M+1}^{M+N} \chi(n) \right| \leq \sum_{M=1}^{q} \left| \sum_{n=M+1}^{M+N} \chi(n) \right| \leq q \max_{M} \left| \sum_{n=M+1}^{M+N} \chi(n) \right|$$

$$\left| \sum_{n=1}^{N} \sum_{M=1}^{q} e(M/q)\, \chi(M+n) \right| = \Bigg| \underbrace{\sum_{n=1}^{N} e(-n/q) \sum_{M=1}^{q} \chi(M)\, e(M/q)}_{\tau(\chi)\, e\left(-\frac{(N+1)}{2q}\right) \frac{\sin(\pi N/q)}{\sin(\pi/q)}} \Bigg| .$$

If q is even, take $N = q/2$, and then

$$\frac{\sin(\pi N/q)}{\sin(\pi/q)} = \frac{1}{\sin(\pi/q)} \geq \frac{q}{\pi}.$$

If q is odd, then take $N = (q-1)/2$, in which case

$$\frac{\sin(\pi N/q)}{\sin(\pi/q)} = \frac{\cos\left(\frac{\pi}{2q}\right)}{\sin(\pi/q)} = \frac{1}{2\sin(\pi/2q)} \geq \frac{q}{\pi}.$$

In either case, therefore, we have

$$\sqrt{q} \cdot \frac{q}{\pi} \leq q \max_{M,N} \left| \sum_{n=M+1}^{M+N} \chi(n) \right|,$$

and the desired conclusion follows. //

This shows that the Pólya–Vinogradov inequality cannot be sharpened ~~too~~ much.

§ 10. Incomplete character sums, II: Burgess' inequality.

If one expects square-root cancellation amongst terms in short character sums, then one might conjecture that

$$\sum_{n=M+1}^{M+N} \chi(n) \ll_\varepsilon N^{1/2} q^\varepsilon. \qquad (\varepsilon>0).$$

The inequality of Burgess makes some progress in this direction, exhibiting cancellation in intervals shorter than $\sqrt{q}$.

We make use of the Riemann Hypothesis for curves over a finite field. This was proved by Weil (1948), but in a form due to Wolfgang Schmidt (1976). (see "Equations over finite fields. An elementary approach", Lemma 4B and Theorem 2C').

Lemma 10.1 (Weil) Let p be a prime number, and suppose that $d>1$ satisfies $d \mid (p-1)$. Let χ be a character modulo p of order d. Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial of positive degree having the property that $f(x) \equiv g(x)^d \pmod p$ identically for no $g(x) \in \mathbb{Z}[x]$.

Then one has

$$\left| \sum_{n=1}^{p} \chi(f(n)) \right| \leq (k-1)p^{1/2},$$

where k denotes the number of distinct zeros of $f(t)$ in $\overline{\mathbb{F}}_p$.

[Connects with curve $y^d = f(x)$ over $\mathbb{F}_q$].

Corollary 10.2. Let p be a prime, $d>1$, and $d \mid (p-1)$. Let χ be a character modulo p of order d. Suppose further that $k \in \mathbb{N}$, and the positive integers $e_1, \dots, e_k$ satisfy the condition that $d \nmid e_j$ for some index j with $1 \leq j \leq k$. Then provided that $c_1, \dots, c_k$ are distinct modulo p, one has

$$\left| \sum_{n=1}^{p} \chi\left((n+c_1)^{e_1} (n+c_2)^{e_2} \dots (n+c_k)^{e_k} \right) \right| \leq (k-1)p^{1/2}.$$

We apply this estimate to deine an average over a short character sum.

<u>Lemma 10.3.</u> Suppose that χ is a non-principal character modulo p, and put
$$S_{h,r} = \sum_{n=1}^{p} \Big| \sum_{m=1}^{h} \chi(m+n) \Big|^{2r}.$$
Then for each positive integer r, one has
$$S_{h,r} \ll r^{2r} \left(h^r p + h^{2r} p^{1/2} \right).$$

<u>Proof.</u> The sum $S_{h,r}$ contains a difficult to handle short inner sum with a long average over n on the outside. It therefore makes sense to expand and interchange orders of summation so as to move the long sum inside. Note that when $m+n \not\equiv 0 \pmod p$, one has (when $\operatorname{ord} \chi = d$),
$$\overline{\chi(m+n)} = (\chi(m+n))^{d-1} = \chi((m+n)^{d-1}).$$

Thus
$$\begin{aligned} S_{h,r} &= \sum_{n=1}^{p} \sum_{m_1=1}^{h} \cdots \sum_{m_{2r}=1}^{h} \chi(m_1+n)\chi(m_2+n)\cdots\chi(m_r+n)\overline{\chi(m_{r+1}+n)}\cdots\overline{\chi(m_{2r}+n)} \\ &= \sum_{\underline{m} \in [1,h]^{2r}} \sum_{n=1}^{p} \chi\left((m_1+n)\cdots(m_r+n)(m_{r+1}+n)^{d-1}\cdots(m_{2r}+n)^{d-1}\right) \\ &= \sum_{\underline{m} \in [1,h]^{2r}} \sum_{n=1}^{p} \chi(f_{\underline{m}}(n)), \qquad (10.1) \end{aligned}$$

where
$$f_{\underline{m}}(n) = \left(\prod_{i=1}^{r} (n+m_i) \right) \cdot \left(\prod_{i=r+1}^{2r} (n+m_i)^{d-1} \right).$$

We next estimate the inner sum of (10.1) by exploiting Corollary 10.2. The first problem is to take care of the possibility that there are repeated values amongst the integers m_i. Here, without loss of generality, we may suppose ~~that~~ $h \leq p$, whence $1 \leq m_i \leq p$ for each index i. Let $c_1 < c_2 < \dots < c_k$ denote the distinct values of $m_1, \dots, m_{2r}$, and put

$$a_\ell = \mathrm{card}\{j \in [1,r] : m_j = c_\ell\}, \qquad b_\ell = \mathrm{card}\{j \in [r+1,2r] : m_j = c_\ell\}.$$

We may ~~then~~ define

$$e_\ell = a_\ell + (d-1)b_\ell \qquad (1 \leq \ell \leq k),$$

and we see ~~that~~

$$f_{\underline{m}}(n) = \prod_{\ell=1}^{k} (n + c_\ell)^{e_\ell}.$$

We deduce ~~that~~

$$S_{h,r} = S^{(1)} + S^{(2)},$$

where $S^{(1)}$ denotes the contribution to (10.1) arising from those $2r$-tuples $\underline{m}$ for which there is an index ℓ with $d \nmid e_\ell$, and $S^{(2)}$ denotes the corresponding contribution in which $d \mid e_\ell$ for $1 \leq \ell \leq r$.

By Corollary 10.2, one has

$$S^{(1)} \leq \sum_{\underline{m} \in [1,h]^{2r}} (k(\underline{m}) - 1) p^{1/2} \leq h^{2r}(2rp^{1/2}).$$

In order to bound the contribution from $S^{(2)}$, note ~~that~~ if $d \mid e_\ell$ for $1 \leq \ell \leq r$, ~~then~~

$$\begin{aligned} e_1 + \dots + e_k &\geq kd \\ \| \\ a_1 + \dots + a_k + (d-1)(b_1 + \dots + b_k) &= rd. \end{aligned}$$

Thus $k \leq r$. But given $c_1, \dots, c_k \in [1,h]^k$, ~~the~~ number of choices of $m_1, \dots, m_{2r}$ with $m_\ell \in \{c_1, \dots, c_k\}$ is at most k^{2r}, and the number of choices for these integers $c_1, \dots, c_k$ is at most $\binom{h}{k}$. Then we conclude ~~that~~

$$S^{(2)} \leq \sum_{1\leq k\leq r} k^{2r} \binom{h}{k} p \ll r^{2r} h^r p,$$

By combining ~~these estimates~~, we arrive ~~at the~~ bound

$$S_{h,r} = S^{(1)} + S^{(2)} \ll r^{2r}(h^r p + h^{2r} p^{1/2}). \quad //$$

We are now equipped to prove the estimate of Burgess (1957) in the following form (the proof here is a development of one due to Friedlander and Iwaniec following the exposition of Montgomery and Vaughan).

<u>Theorem 10.4</u> (Burgess) Suppose that p is an odd prime and $r \in \mathbb{N}$. Then one has, for any non-principal character χ,

$$\sum_{n=M+1}^{M+N} \chi(n) \ll r\, N^{1-1/r}\, p^{\frac{r+1}{4r^2}} (\log p)^{\alpha_r},$$

where

$$\alpha_r = \begin{cases} 1, & \text{when } r = 1, 2, \\ \frac{3}{2r}, & \text{when } r > 2. \end{cases}$$

<u>Proof.</u> The starting point is a Vinogradov-like trick for creating averaging not at first apparent. Thus, we have for $a, b \in \mathbb{N}$,

$$\begin{aligned} S(M,N) := & \sum_{n=M+1}^{M+N} \chi(n) \\ = & \sum_{n=M+1}^{M+N} \chi(n+ab) + \sum_{n=M+1}^{M+ab} \chi(n) - \sum_{n=M+N+1}^{M+N+ab} \chi(n) \\ = & \sum_{n=M+1}^{M+N} \chi(n+ab) + S(M,ab) - S(M+N,ab). \end{aligned}$$

We would like to average over values of a, b here, and it is useful to regard the final two terms as error terms. Put

$$\tilde{S}(y) = \max_{\substack{M,N \\ y \geq N}} |S(M,N)| .$$

Then we have

$$S(M,N) = \sum_{n=M+1}^{M+N} \chi(n+ab) + 2\theta \tilde{S}(ab),$$

where $|\theta| \leq 1$, and from this it follows ~~that~~ for $A \geq 1$ and $B \geq 1$, we have

$$AB\, S(M,N) = \sum_{a=1}^{A} \sum_{b=1}^{B} \sum_{n=M+1}^{M+N} \chi(n+ab) + 2AB\theta' \tilde{S}(AB),$$

for some real number θ' with $|\theta'| \leq 1$. In particular,

$$S(M,N) = \frac{1}{AB} \sum_{a=1}^{A} \sum_{b=1}^{B} \sum_{n=M+1}^{M+N} \chi(n+ab) + 2\theta' \tilde{S}(AB). \quad \text{— (10.2)}$$

We now seek to exploit the averaging over a and b via Lemma 10.3. Some simplifications smooth our ride. Observe first ~~that~~ the conclusion of Theorem 10.4 is weaker than that provided by ~~the~~ the Pólya – Vinogradov inequality when $r = 1$, or $N > p^{5/8}$ and $r \geq 2$. In the latter case

$$N^{1-1/r} p^{\frac{r+1}{4r^2}} > p^{\theta_r}, \text{ with } \theta_r = \frac{5}{8}\left(1 - \frac{1}{r}\right) + \frac{r+1}{4r^2}$$
$$= \frac{1}{2} + \frac{r^2 - 3r + 2}{8r^2},$$

so ~~that~~ $\theta_r = 0$ when $r = 2$ and $\theta_r > \frac{1}{2}$ when $r > 2$. Also, when $r > 2$ and $N > p^{1/2}$, one sees similarly ~~that~~

$$N^{1-1/r} p^{\frac{r+1}{4r^2}} = N^{\frac{1}{2}} p^{\frac{3}{16}} \cdot N^{\frac{1}{2} - \frac{1}{r}} p^{\frac{r+1}{4r^2} - \frac{3}{16}}$$
$$> p^{\theta_r'} \cdot N^{\frac{1}{2}} p^{\frac{3}{16}},$$

where $$\theta_r' = \frac{1}{4} - \frac{1}{2r} + \frac{r+1}{4r^2} - \frac{3}{16} = \frac{1}{16} - \left(\frac{r-1}{4r^2}\right).$$

Thus, when $r > 2$ and $N > p^{1/2}$, the stated bound is worse than ~~that~~ derived from the case $r = 2$. Finally, when $N \le p^{\frac{r+1}{4r}}$, the bound is worse than trivial. We may ~~therefore~~ suppose ~~that~~ p is large,

$$r \ge 2 \quad \text{and} \quad \begin{cases} p^{\frac{r+1}{4r}} < N \le p^{5/8}, & \text{when } r = 2 \\ p^{\frac{r+1}{4r}} < N \le p^{1/2}, & \text{when } r > 2. \end{cases}$$

Now recall (10.2) and examine the main sum on the right hand side. We have

$$\sum_{a=1}^{A} \sum_{b=1}^{B} \sum_{n=M+1}^{M+N} \chi(n+ab) = \sum_{l=1}^{p} \sum_{\substack{n \in [M+1, M+N] \\ 1 \le a \le A \\ n \equiv al \pmod p}} \sum_{b=1}^{B} \chi(al + ab)$$

$$= \sum_{l=1}^{p} \sum_{\substack{n, a \\ n \equiv al \pmod p}} \chi(a) \sum_{b=1}^{B} \chi(l+b)$$

$$\le \sum_{l=1}^{p} \nu(l) \left| \sum_{b=1}^{B} \chi(l+b) \right|,$$

where

$$\nu(l) = \#\{ n \in [M+1, M+N],\ a \in [1, A] : n \equiv al \pmod p \}.$$

By Hölder's inequality, moreover, we have

$$\left(\sum_{l=1}^{p} \nu(l) \left| \sum_{b=1}^{B} \chi(l+b) \right| \right)^{2r} \le \left(\sum_{l=1}^{p} \nu(l) \right)^{2r-2} \left(\sum_{l=1}^{p} \nu(l)^2 \right) \left(\sum_{l=1}^{p} \left| \sum_{b=1}^{B} \chi(l+b) \right|^{2r} \right) \tag{10.3}$$

We suppose ~~that~~ $A < p$, so ~~that~~ $(a, p) = 1$ for $1 \le a \le A$.

Then
$$\sum_{l=1}^{p} \nu(l) = \sum_{l=1}^{p} \#\{ n \in [M+1, M+N],\ a \in [1, A] : na^{-1} \equiv l \pmod p \}$$
$$= AN.$$

Also, provided ~~that~~ $AN < \frac{1}{2}p$, we ~~see~~ ~~that~~

$$\sum_{\ell=1}^{p} \nu(\ell)^2 = \#\Big\{ n_1, n_2 \in [M+1, M+N],\ a_1, a_2 \in [1, A] : \underset{\equiv \ell}{n_1 a_1^{-1}} \equiv \underset{\equiv \ell}{n_2 a_2^{-1}} \pmod p \Big\}$$

$$= \#\{ \underline{n}, \underline{a} : \quad n_1 a_2 \equiv n_2 a_1 \pmod p \}$$

$$= \#\{ \underline{n}, \underline{a} : \quad n_1 a_2 = n_2 a_1 \}$$

$$\leq \sum_{1 \leq m \leq AN} \tau(m)^2 \ll AN (\log(AN))^3 \ll AN (\log p)^3 .$$

Finally, as a consequence of Lemma 10.3, one sees ~~that~~

$$\sum_{\ell=1}^{p} \Big| \sum_{b=1}^{B} \chi(\ell + b) \Big|^{2r} \ll r^{2r} \left(B^r p + B^{2r} p^{1/2} \right).$$

This ~~estimate~~ ~~motivates~~ us ~~to~~ ~~take~~ $B = \lfloor p^{1/(2r)} \rfloor$.

Combining ~~these~~ ~~estimates~~ ~~within~~ (10.3), we deduce ~~that~~

$$\frac{1}{AB} \sum_{a=1}^{A} \sum_{b=1}^{B} \sum_{n=M+1}^{M+N} \chi(n+ab) \ll \frac{1}{AB} (AN)^{1-\frac{1}{r}} \left(AN (\log p)^3 \right)^{\frac{1}{2r}} \left(r^{2r} B^r p \right)^{\frac{1}{2r}}$$

$$\ll_r \frac{N^{1-\frac{1}{2r}}}{A^{\frac{1}{2r}} B^{\frac{1}{2}}} p^{\frac{1}{2r}} (\log p)^{\frac{3}{2r}} . \qquad \text{—— (10.4)}$$

We take A as large as possible ~~with~~ AB smaller ~~than~~ N, so ~~that~~ ~~the~~ error term in ~~the~~ ~~estimate~~

$$S(M,N) = \frac{1}{AB} \sum_{n,a,b} \chi(n+ab) + 2\theta_1 \tilde{S}(AB)$$

is sensible. So we take $A = \lfloor \frac{1}{10} N p^{-1/(2r)} \rfloor$ and $B = \lfloor p^{1/(2r)} \rfloor$.

Then $AB \asymp \frac{N}{10}$ and (10.4) becomes

$$\frac{1}{AB} \sum_{a=1}^{A} \sum_{b=1}^{B} \sum_{n=M+1}^{M+N} \chi(n+ab) \ll r N^{1-\frac{1}{r}} p^{\frac{1}{4r} + \frac{1}{4r^2}} (\log p)^{\frac{3}{2r}} .$$

Now we refer back to (10.2), deducing that

$$\tilde{S}(N) = \max_{\substack{M, N' \\ N \geq N'}} |S(M, N')|$$

$$\leq N^{1-1/r} \lambda(p,r) + 2\tilde{S}(N/10),$$

where for a suitable constant C we have

$$\lambda(p,r) = C r p^{\frac{r+1}{4r^2}} (\log p)^{\frac{3}{2r}}.$$

This enables us to bound $\tilde{S}(N)$ iteratively. By taking $K = \frac{9}{10}\log_{10} N$, say, we deduce that

$$\tilde{S}(N) \leq N^{1-1/r} \lambda(p,r) \sum_{k=0}^{K} 2^k \left(10^{-(1-1/r)}\right)^k + 2^{K+1}\tilde{S}(N/10^{K+1})$$

$$\ll N^{1-1/r} \lambda(p,r) + \frac{N}{5^{\frac{9}{10}\log_{10} N}} = N^{1-\frac{1}{r}} \lambda(p,r) + O(\sqrt{N}).$$

Thus we conclude that

$$|S(M,N)| \ll N^{1-1/r} \lambda(p,r) \ll r N^{1-\frac{1}{r}} p^{\frac{r+1}{4r^2}} (\log p)^{\frac{3}{2r}}. \quad //$$

The utility of Burgess' estimate becomes apparent through the flexible options for the parameter r. Thus far, we have seen that there is cancellation in character sums of length slightly larger than $q^{1/2}\log q$ (Pólya–Vinogradov). Suppose $\delta > \frac{1}{4}$ and $N > p^{\delta}$. Then we have

$$\sum_{n=M+1}^{M+N} \chi(n) \ll N\left(r^r \frac{p^{\frac{r+1}{4r}}}{N} (\log p)^{\frac{3}{2}}\right)^{\frac{1}{r}},$$

so that on choosing r in such a way that

$$\frac{r+1}{4r} < \delta, \qquad (\text{this is possible, since } \delta > \tfrac{1}{4}!).$$

we are showing that the right hand side here is $o(N)$, and that there is considerable cancellation. This idea can be used to show that there exists n with $\chi(n) \neq 1$ satisfying

$$n = n_\chi \ll_\varepsilon p^{\frac{1}{4\sqrt{e}} + \varepsilon}.$$

---

§11. Primes in arithmetic progression on average.

A first course in ANT delivers estimates for $\psi(x;q,a)$ via bounds on

$$\psi(x,\chi) := \sum_{n \le x} \Lambda(n)\chi(n).$$

Here we have asymptotics for $\psi(x,\chi_0)$, and good bounds for $\psi(x,\chi)$, but only so long as $q \le (\log x)^A$, and with error terms $O(x\exp(-c\sqrt{\log x}))$. This is all owing to our weak knowledge concerning the zero-free region of $L(s,\chi)$, and the potential existence of Landau-Siegel zeros. However, by working on average (using the Large Sieve inequality), we can do much better.

Our goal in this section is the average estimate embodied in the theorem following.

Theorem 11·1. When $Q \ge 1$ and $x \ge 2$, one has

$$\sum_{1 \le q \le Q} \frac{q}{\varphi(q)} \sum_{\chi}{}^{*} \sup_{y \le x} |\psi(y,\chi)| \ll \left(x + x^{5/6}Q + x^{1/2}Q^2\right)(\log x)^3.$$

Here, we adopt the convention that $\sum_\chi^*$ denotes a sum over all primitive characters modulo q.

Note that the term with $q=1$ already contributes a term $\psi(y,\chi_0)$ with $q=1$, which is $\sim x$. Also, we expect (but cannot prove) that $\psi(y,\chi)$ is of size at most $y^{1/2}$ or thereabouts, so the contribution arising from these terms may be expect to be roughly

$$\sum_{1 \le q \le Q} q \max_{y \le x} y^{\frac{1}{2}} \asymp Q^2 x^{\frac{1}{2}}.$$

Thus, but for the term $x^{5/6}Q$, the claimed conclusion is close to the truth.

We require a maximal inequality for partial sums of series, the motivation for which comes from the analogue for Dirichlet series of partial sum estimates of Fourier series (cf. a weaker classical precursor to the Carleson - Hunt Theorem).

Consider a series $\sum_{n=1}^{N} a_n$ and the associated Dirichlet series

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

We seek a bound for

$$\max_{y \le N} \Big| \sum_{n \le y} a_n \Big|$$

amenable to our standard toolbox.

<u>Lemma 11·2.</u> Let U be a positive number with $U \geq N$. Then one has

$$\max_{y \le N} \Big| \sum_{n \le y} a_n \Big| \ll \int_{-U}^{U} |D(iu)| \min\{\log N, \tfrac{1}{|u|}\}\, du + \frac{N}{U} \sum_{n=1}^{N} |a_n|.$$

<u>Proof.</u> We follow a path not dissimilar to the treatment of a quantitative truncated form of Perron's formula. We begin by observing that when α and β are real numbers, one has

$$\int_{-U}^{U} e^{i\beta u} \underbrace{\frac{\sin(\alpha u)}{u}}_{\substack{\uparrow \\ \text{even function}}} du = \int_{-U}^{U} \cos(\beta u) \frac{\sin(\alpha u)}{u}\, du$$

$$= \int_{-U}^{U} \frac{\frac{1}{2}\big(\sin((\alpha+\beta)u) + \sin((\alpha-\beta)u)\big)}{u}\, du$$

(change of variables)

$$= \operatorname{sgn}(\alpha+\beta) \int_0^{|\alpha+\beta| U} \frac{\sin u}{u}\, du + \operatorname{sgn}(\alpha-\beta) \int_0^{|\alpha-\beta| U} \frac{\sin u}{u}\, du$$

Recall next that $\int_0^\infty \frac{\sin u}{u}\, du = \frac{\pi}{2}$, and the sine integral $si(x)$ is defined by

$$\mathrm{si}(x) := -\int_x^\infty \frac{\sin u}{u}\, du.$$

Then what we have shown is ~~that~~

$$\int_{-U}^{U} e^{i\beta u} \frac{\sin(\alpha u)}{u}\, du = \mathrm{sgn}(\alpha+\beta)\left(\frac{\pi}{2} + \mathrm{si}(|\alpha+\beta|U)\right) + \mathrm{sgn}(\alpha-\beta)\left(\frac{\pi}{2} + \mathrm{si}(|\alpha-\beta|U)\right).$$

On recalling that $\mathrm{si}(x) \ll \min\{1, 1/x\}$ for $x \geq 0$, we may relate this expression to the characteristic function χ_I of the interval $[-\alpha, \alpha]$.

Thus

$$\chi_I(\beta) = \frac{1}{\pi}\int_{-U}^{U} e^{i\beta u} \frac{\sin(\alpha u)}{u}\, du + O\left(\min\left\{1, \frac{1}{U|\alpha-\beta|}\right\} + \min\left\{1, \frac{1}{U|\alpha+\beta|}\right\}\right)$$

Having obtained a formula for the characteristic function of $[-\alpha, \alpha]$, we now apply this to deduce the partial sum ~~estimate~~ claimed in the statement of the lemma. Suppose that K is an integer with $0 \leq K < N$, and put $\alpha = \log(K + \frac{1}{2})$ and $\beta = -\log n$. Then by multiplying each side by a_n and summing over n, we obtain the relation

$$\sum_{n=1}^{K} a_n = \frac{1}{\pi}\int_{-U}^{U} \underbrace{\left(\sum_{n=1}^{N} a_n n^{-iu}\right)}_{D(iu)} \frac{\sin(\alpha u)}{u}\, du + O\left(\sum_{n=1}^{N} |a_n| \min\left\{1, \frac{1}{U\left|\log\left(\frac{n}{K+1/2}\right)\right|}\right\}\right).$$

Since $\dfrac{\sin(\alpha u)}{u} \ll \min\{|\alpha|, \frac{1}{|u|}\}$ and $\left|\log\left(\dfrac{n}{K+1/2}\right)\right| \gg \dfrac{1}{N}$, we deduce that

$$\max_{y \leq N} \left|\sum_{n \leq y} a_n\right| \ll \int_{-U}^{U} |D(iu)| \min\left\{\log N, \frac{1}{|u|}\right\} du + \frac{N}{U}\sum_{n=1}^{N} |a_n|.$$

//

Our strategy for proving Theorem 11.1 involves applying Vaughan's identity to rewrite $\psi(y,\chi)$ as a bilinear sum. The expression

$$\sup_{y \le x} |\psi(y,\chi)|$$

may then be disentangled by means of the next lemma.

Lemma 11.3. One has

$$\sum_{q \le Q} \frac{q}{\phi(q)} \sum_{\chi}^{*} \sup_{y} \Big| \sum_{\substack{1 \le m \le M \\ mn \le y}} \sum_{1 \le n \le N} a_m b_n \chi(mn) \Big|$$

$$\ll (M+Q^2)^{1/2} (N+Q^2)^{1/2} \Big(\sum_{m=1}^{M} |a_m|^2 \Big)^{1/2} \Big(\sum_{n=1}^{N} |b_n|^2 \Big)^{1/2} \log(2MN).$$

Proof: Note that in the absence of the $\sup_y$ on the lhs, ~~this~~ conclusion was obtained in one of the homework exercises — we pay the $\log(2MN)$ factor for this additional flexibility.

We begin by observing ~~that~~ Lemma 11.2 shows ~~that~~ for $T \ge MN$,

$$\sup_{y} \Big| \sum_{\substack{1 \le m \le M \\ mn \le y}} \sum_{1 \le n \le N} a_m b_n \chi(mn) \Big| \ll \int_{-T}^{T} \Big| \sum_{\substack{1 \le m \le M \\ 1 \le n \le N}} a_m b_n \chi(mn)(mn)^{-it} \Big| \min\{\log(MN), 1/|t|\} \, dt$$

$$+ \frac{MN}{T} \sum_{\substack{1 \le m \le M \\ 1 \le n \le N}} |a_m| \cdot |b_n| .$$

By Cauchy's inequality, the last term here is

$$\ll \frac{(MN)^{3/2}}{T} \Big(\sum_{m=1}^{M} |a_m|^2 \Big)^{1/2} \Big(\sum_{n=1}^{N} |b_n|^2 \Big)^{1/2} .$$

Then if we take $T = (MN)^{3/2}$, we see ~~that~~ this term is harmless. Meanwhile, since we have removed the condition $mn \le y$ from the first term, we may apply the large sieve inequality for characters to deduce that

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\left|\sum_{\substack{1\le m\le M\\1\le n\le N}} a_m b_n \chi(mn)(mn)^{-it}\right|$$

$$\le\left(\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\left|\sum_{1\le m\le M}\frac{a_m}{m^{it}}\chi(m)\right|^2\right)^{\frac12}\left(\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\left|\sum_{1\le n\le N}\frac{b_n}{n^{it}}\chi(n)\right|^2\right)^{\frac12}$$

$$\ll (M+Q^2)^{1/2}\left(\sum_{m=1}^{M}\left|\frac{a_m}{m^{it}}\right|^2\right)^{\frac12}\cdot(N+Q^2)^{1/2}\left(\sum_{n=1}^{N}\left|\frac{b_n}{n^{it}}\right|^2\right)^{\frac12}$$

$$=(M+Q^2)^{\frac12}(N+Q^2)^{\frac12}\left(\sum_{m=1}^{M}|a_m|^2\right)^{\frac12}\left(\sum_{n=1}^{N}|b_n|^2\right)^{\frac12}.$$

Hence

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\sup_{y}\left|\sum_{1\le m\le M}\sum_{\substack{1\le n\le N\\ mn\le y}} a_m b_n \chi(mn)\right|$$

$$\ll\int_{-T}^{T}\min\left\{\log(MN),\frac{1}{|t|}\right\}\cdot\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\left|\sum_{\substack{1\le m\le M\\1\le n\le N}} a_m b_n \chi(mn)(mn)^{-it}\right|dt$$

$$+\,O\left(\left(\sum_{1\le m\le M}|a_m|^2\right)^{\frac12}\left(\sum_{1\le n\le N}|b_n|^2\right)^{\frac12}\right)$$

$$\ll (M+Q^2)^{\frac12}(N+Q^2)^{\frac12}\left(\sum_{m=1}^{M}|a_m|^2\right)^{\frac12}\left(\sum_{n=1}^{N}|b_n|^2\right)^{\frac12}$$

$$\times\underbrace{\left(1+\int_{-T}^{T}\min\left\{\log(MN),\frac{1}{|t|}\right\}dt\right)}_{\ll\,\log(2MN)}.$$

(recall: $T=(MN)^{3/2}$).

The conclusion of the lemma follows at once. //

<u>The proof of Theorem 11·1:</u> Our goal is the proof of the bound:

$$\sum_{1\le q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}^{*}\sup_{y\le x}|\psi(y,\chi)| \ll (x + x^{5/6}Q + x^{1/2}Q^2)(\log x)^3. \quad \text{— (11·1)}$$

We observe first that if $Q^2 > x$, then we may apply Lemma 11·3 with $M=1$, $a_1 = 1$, $N = \lfloor x \rfloor$, $b_n = \Lambda(n)$ to show that

$$\sum_{q\le Q}\frac{q}{\phi(q)}\sum_{\chi}^{*}\sup_{y\le x}\Bigl|\underbrace{\sum_{1\le n\le y}\Lambda(n)\chi(n)}_{\psi(y,\chi)}\Bigr| \ll (1+Q^2)^{\frac12}(x+Q^2)^{\frac12}\Bigl(\sum_{1\le n\le x}\Lambda(n)^2\Bigr)^{\frac12}\log x$$

$$\ll Q^2(x\log x)^{\frac12}\log x,$$

which suffices to prove (11·1). So we may suppose henceforth that $Q \le x^{1/2}$.

Our strategy is to apply Vaughan's identity with $f(n) = \chi(n)$. Thus (see §6),

$$\underset{\displaystyle\sum_{n\le y}\chi(n)\Lambda(n)}{\psi(y,\chi)} = S_1 + S_2 + S_3 + S_4,$$

where

$$S_1(y,\chi) = \sum_{n\le U}\Lambda(n)\chi(n),$$

$$S_2(y,\chi) \ll (\log(UV))\sum_{t\le UV}\Bigl|\sum_{r\le y/t}\chi(rt)\Bigr|,$$

$$S_3(y,\chi) \ll (\log y)\sum_{k\le V}\sup_{W}\Bigl|\sum_{W\le m\le y/k}\chi(mk)\Bigr|,$$

$$S_4(y,\chi) = \sum_{U<m\le y/V} b(m)\sum_{V<k\le y/m}\mu(k)\chi(mk),$$

in which $b(m) \ll \log m$.

Trivially, one has

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}^{*}\sup_{y\le x}|S_1(y,\chi)| \ll \sum_{q\le Q} q \sum_{n\le U}\Lambda(n) \ll Q^2 U .$$

Next we treat $S_4(y,\chi)$, observing that from Lemma 11.3 one has

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}^{*}\sup_{y\le x}\Big|\sum_{\substack{U<m\le y/V\\ M<m\le 2M}} b(m)\sum_{V<k\le y/m}\mu(k)\chi(mk)\Big|$$

$$\ll (M+Q^2)^{1/2}(x/M+Q^2)^{1/2}\Big(\underbrace{\sum_{m=1}^{M}|b(m)|^2}_{O(M(\log M)^2)}\Big)^{\frac12}\Big(\underset{O(x/M)}{\sum_{k\le x/M}|\mu(k)|^2}\Big)^{\frac12}(\log x)$$

$$\ll (x + QxM^{-\frac12} + Qx^{1/2}M^{1/2} + Q^2x^{1/2})(\log x)^2 .$$

We sum this over dyadic intervals with $M=2^l$ for $U/2\le M=2^l\le x/V$, whence we deduce that

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}^{*}\sup_{y\le x}|S_4(y,\chi)| \ll (x+QxU^{-\frac12}+QxV^{-\frac12}+Q^2x^{\frac12})(\log x)^3 .$$

Next we examine $S_2(y,\chi)$, observing that

$$S_2(y,\chi) \ll \log(UV)\,(S_2^{*}(y,\chi)+S_2^{\dagger}(y,\chi)),$$

where

$$S_2^{*}(y,\chi) = \sum_{1\le t\le U}\Big|\sum_{r\le y/t}\chi(rt)\Big|$$

and

$$S_2^{\dagger}(y,\chi) = \sum_{U<t\le UV}\Big|\sum_{r\le y/t}\chi(rt)\Big| .$$

We can write

$$S_2^{\dagger}(y,\chi) = \sum_{U<t\le UV} w_t \sum_{1\le r\le y/t}\chi(rt),$$

where w_t is a unimodular function of t (related to the argument of the inner sum as a complex number). We can therefore apply the argument of the treatment of $S_4(y,\chi)$ to deduce that

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\sup_{y\le x}\left|S_2^{\dagger}(y,\chi)\right| \ll (\log x)^2 \max_{\frac{U}{2}\le M\le UV}(M+Q^2)^{1/2}\left(X/M+Q^2\right)^{1/2}$$

$$\cdot\left(\sum_{m=1}^{M}|w_m|^2\right)^{\frac12}\left(\sum_{k\le X/M}1\right)^{1/2}$$

$$\ll \left(x+QxU^{-1/2}+Qx^{1/2}(UV)^{1/2}+Q^2x^{1/2}\right)(\log x)^2.$$

Meanwhile, when $q=1$, we have

$$S_2^{*}(y,\chi)\ll\sum_{1\le t\le U}\frac{y}{t}\ll y(\log U),$$

whilst for $q>1$ we may apply the Pólya–Vinogradov theorem to deduce that

$$S_2^{*}(y,\chi)=\sum_{1\le t\le U}\left|\sum_{r\le y/t}\chi(rt)\right|$$

$$\ll\sum_{1\le t\le U}q^{1/2}\log q\ll q^{1/2}U\log q.$$

Thus we deduce that

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\sup_{y\le x}\left|S_2^{*}(y,\chi)\right|\ll x(\log U)+(\log Q)U\sum_{q\le Q}q^{3/2}$$

$$\ll\left(x+Q^{5/2}U\right)\log(Ux).$$

By combining these estimates, we see that

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\sup_{y\le x}\left|S_2(y,\chi)\right|\ll\left(x+QxU^{-1/2}+Qx^{1/2}(UV)^{1/2}+Q^{5/2}U+Q^2x^{1/2}\right)(\log x)^3.$$

The treatment of S_3 is similar to that of S_2^{*}. Thus we see that

$$S_3(y,\chi)\ll(\log y)\sum_{k\le V}\sup_{w}\Big|\underbrace{\sum_{w\le m\le y/k}\chi(mk)}_{\ll\, q^{1/2}\log q}\Big|\qquad(\chi\ne\chi_0)$$

$$\ll \quad q^{1/2} V (\log(qy))^2,$$

whence

$$\sum_{q\le Q}\frac{q}{\varphi(q)} \sum_{\chi}^{*} \sup_{y\le x} |S_3(y,\chi)| \ll (x + Q^{5/2}V)(\log(qVx))^2.$$

We may now combine the contributions from $S_1,\dots,S_4$ to obtain the upper bound

$$\sum_{q\le Q}\frac{q}{\varphi(q)} \sum_{\chi}^{*} \sup_{y\le x} |\psi(y,\chi)| \ll \Big(x + QxU^{-1/2} + QxV^{-1/2} + Q^2x^{1/2} + Qx^{1/2}(UV)^{1/2} + Q^{5/2}U + Q^{5/2}V + Q^2U\Big) \times (\log(xUV))^3.$$

The optimal estimate holds with $U=V$. We may already assume that $Q\le x^{1/2}$. If, in addition, one has $Q\ge x^{1/3}$, ~~then~~ the right hand side is optimised with $U=V=x^{2/3}/Q$ when it becomes

$$(x + QxU^{-1/2} + Q^2x^{1/2} + Qx^{1/2}U + Q^{5/2}U)(\log(xU))^3$$

$$U^{3/2} = xQ^{-3/2}$$

$$\ll (x + Q^{3/2}x^{2/3} + Q^2x^{1/2})(\log x)^3$$

$$\ll (x + Q^2x^{1/2})(\log x)^3.$$

When $1\le Q\le x^{1/3}$, meanwhile, we may take $U=V=x^{1/3}$ to obtain that the right hand side is likewise

$$\ll (x + Qx^{5/6} + Q^2x^{1/2} + Q^{5/2}x^{1/3})(\log x)^3$$

$$\ll (x + Qx^{5/6} + Q^2x^{1/2} + Q^2x^{1/6+1/3})(\log x)^3.$$

In all cases, therefore, we have

$$\sum_{q\le Q}\frac{q}{\varphi(q)} \sum_{\chi}^{*} \sup_{y\le x} |\psi(y,\chi)| \ll (x + Qx^{5/6} + Q^2x^{1/2})(\log x)^3. \quad //$$

§ 12. The Bombieri – Vinogradov Theorem.

In this section we investigate the extent to which $\psi(x;q,a)$ is close to $x/\varphi(q)$, as we average over q and take the most extreme value of a. To this end, when $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a,q)=1$, we define

$$E(x;q,a) := \psi(x;q,a) - \frac{x}{\varphi(q)},$$

$$E(x;q) = \max_{\substack{1\le a\le q\\(a,q)=1}} |E(x;q,a)|,$$

$$E^*(x;q) = \sup_{y\le x} E(y;q).$$

It transpires that one can show (and this is what the course has been leading up to thus far):

<u>Theorem 12.1</u> (The Bombieri – Vinogradov theorem) Let $A>0$ be fixed. Then one has

$$\sum_{q\le Q} E^*(x;q) \ll x^{\frac{1}{2}} Q \,(\log x)^3,$$

provided only that $x^{1/2}(\log x)^{-A} \le Q \le x^{1/2}$.

Note that for most values of q with $q \le \underbrace{x^{1/2}(\log x)^{-A}}_{Q}$, this conclusion shows that $E^*(x;q)$ is at most $x^{\frac{1}{2}}(\log x)^3$, and hence smaller than $x/\varphi(q)$. Thus, in particular, for each fixed choice of a one has $\psi(x;q,a) \sim x/\varphi(q)$ for almost all $q \le x^{\frac{1}{2}}(\log x)^{-A}$. But note here that one can allow a to vary with q, and also x in an appropriate sense.

This is roughly the same bound ~~that~~ we would obtain assuming GRH. More is expected to be true:

__Conjecture__ (Montgomery). Suppose ~~that~~ $a \in \mathbb{Z}$, $q \in \mathbb{N}$ satisfy $(a,q)=1$ and $1 \le q \le x$. Then one has

$$\psi(x;q,a) - \frac{x}{\varphi(q)} \ll x^{\varepsilon}\left(\frac{x}{q}\right)^{1/2}.$$

__Conjecture.__ (Elliott – Halberstam Hypothesis) Let $A>0$ and $\varepsilon>0$ be fixed. Then one has

$$\sum_{q\le Q} E^*(x;q) \ll x(\log x)^{-A},$$

provided only that $Q \le x^{1-\varepsilon}$.

The Elliott – Halberstam Hypothesis shows that for almost all $q \le x^{1-\varepsilon}$, one has $E^*(x;q) \ll \frac{x}{\varphi(q)}(\log x)^{-A/2}$, whence

$$\psi(x;q,a) \sim \frac{x}{\varphi(q)}.$$

__The proof of Theorem 12.1.__ We have

$$\psi(x;q,a) = \frac{1}{\varphi(q)} \sum_{\chi(\mathrm{mod}\, q)} \bar{\chi}(a)\,\psi(x;\chi).$$

Thus, on putting

$$\psi'(x,\chi) = \psi(x,\chi) - E_0(\chi)\,x,$$

where

$$E_0(\chi) = \begin{cases} 1, & \text{when } \chi=\chi_0, \\ 0, & \text{when } \chi \ne \chi_0, \end{cases}$$

we see ~~that~~

$$\psi(x;q,a) - \frac{x}{\varphi(q)} = \frac{1}{\varphi(q)} \sum_{\chi} \bar{\chi}(a)\,\psi'(x,\chi).$$

In particular, we infer ~~that~~

$$E(x;q) \le \max_{\substack{1\le a\le q\\ (a,q)=1}} \frac{1}{\varphi(q)} \left| \sum_{\chi} \bar{\chi}(a)\,\psi'(x,\chi)\right|$$

$$\leq \frac{1}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\, q)} |\psi'(x,\chi)|.$$

Finally (for now), we conclude ~~that~~

$$E^*(x;q) \leq \frac{1}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\, q)} \sup_{y \leq x} |\psi'(y,\chi)|.$$

In order to make use of the averaged estimates of the previous section, we must reduce to a consideration of primitive characters. With this in mind, we observe that when $d|q$ and the character χ modulo q is induced by the primitive character χ^* modulo d, one has

$$\psi'(y,\chi^*) - \psi'(y,\chi) = \sum_{p|q} \sum_{\substack{k \in \mathbb{N} \\ p^k \leq y}} \chi^*(p)^k \log p$$

$$\ll \sum_{p|q} \log y = \omega(q) \log y \ll (\log(qy))^2.$$

Thus

$$\sum_{q \leq Q} E^*(x;q) \ll \sum_{d \leq Q} \sum_{\substack{q \leq Q \\ d|q}} \frac{1}{\varphi(q)} \sum_{\chi^* \,(\mathrm{mod}\, d)} \left(\sup_{y \leq x} |\psi'(y,\chi^*)| + O\left((\log(qy))^2\right)\right)$$

$$= \sum_{d \leq Q} \sum_{\chi^* \,(\mathrm{mod}\, d)} \left(\sup_{y \leq x} |\psi'(y,\chi^*)| + O\left((\log(Qx))^2\right)\right) \sum_{\substack{q \leq Q \\ d|q}} \frac{1}{\varphi(q)}.$$

On writing $q = dm$ and noting that $\varphi(dm) \geq \varphi(d)\varphi(m)$, however, we see ~~that~~

$$\sum_{\substack{1 \leq q \leq Q \\ d|q}} \frac{1}{\varphi(q)} = \sum_{1 \leq m \leq Q/d} \frac{1}{\varphi(dm)} \leq \frac{1}{\varphi(d)} \sum_{1 \leq m \leq Q/d} \frac{1}{\phi(m)}$$

$$\leq \frac{1}{\varphi(d)} \prod_{p \leq Q/d} \left(1 + \frac{1}{p-1} + \frac{1}{p(p-1)} + \cdots\right)$$

$$= \frac{1}{\varphi(d)} \prod_{p \leq Q/d} \left(1 + \frac{p}{(p-1)^2}\right)$$

$$= \frac{1}{\varphi(d)} \prod_{p \le Q/d} \left(1 - \frac{1}{p}\right)^{-1} \left(1 + \frac{1}{p(p-1)}\right)$$

$$\ll \frac{1}{\varphi(d)} \log(2Q/d).$$

Hence

$$\sum_{1 \le q \le Q} E^*(x;q) \ll \sum_{1 \le d \le Q} \frac{\log(2Q/d)}{\varphi(d)} \sum_{\chi^* (\mathrm{mod}\, d)} \sup_{y \le x} |\psi'(y,\chi^*)|$$

$$+ \underbrace{O\left(\sum_{d \le Q} (\log(Qx))^2 \cdot \frac{\varphi_2(d)}{\varphi(d)} \log(2Q/d) \right)}_{= O(Q(\log(Qx))^2).}$$

$$= \sum_{1 \le q \le Q} \frac{\log(2Q/q)}{\varphi(q)} \sum_{\chi (\mathrm{mod}\, q)}^{*} \sup |\psi'(y,\chi)|$$

$$+ O(Q(\log(Qx))^2).$$

The last expression is of essentially the correct shape to apply Theorem 11·1. It is convenient to take care of the very small values of q separately, so we put $Q_1 = (\log x)^{A+1}$, and divide into dyadic intervals using a parameter U with $Q_1 \le U \le Q$. Hence, as a consequence of Theorem 11·1, we see that

$$\sum_{U < q \le 2U} \frac{\log(2Q/q)}{\varphi(q)} \sum_{\chi (\mathrm{mod}\, q)}^{*} \sup_{y \le x} |\psi'(y,\chi)|$$

$$\ll \frac{\log(4Q/U)}{U} \sum_{U < q \le 2U} \frac{q}{\varphi(q)} \sum_{\chi}^{*} \sup_{y \le x} |\psi'(y,\chi)|$$

$$\ll \left(\frac{x}{U} + x^{5/6} + x^{1/2} U\right)(\log x)^3 \cdot \log\left(\frac{4Q}{U}\right),$$

whence by summing over $U = 2^k Q_1$, we deduce ~~that~~

$$\sum_{Q_1 < q \le Q} \frac{\log(2Q/q)}{\varphi(q)} \sum_{\chi (\mathrm{mod}\, q)}^{*} \sup_{y \le x} |\psi'(y,\chi)| \ll \frac{x}{Q_1}(\log x)^4 + x^{5/6}(\log x)^5 + x^{\frac{1}{2}} Q (\log x)^3.$$

$$\ll x^{\frac{1}{2}} Q (\log x)^3 .$$

(Note here that $Q \geq x^{1/2} (\log x)^{-A}$, by supposition). We must still handle the contribution of the summands with $1 \leq q \leq Q_1$. But given a primitive character χ modulo q with $q \leq Q_1$, the Siegel - Walfisz theorem shows that

$$\sup_{y \leq x} |\psi'(y, \chi)| \ll x \exp(-c_1 \sqrt{\log x}) ,$$

for a suitable $c_1 > 0$. Thus

$$\sum_{q \leq Q_1} \frac{\log(2Q/q)}{\varphi(q)} \sum_{\chi}{}^{*} \sup_{y \leq x} |\psi'(y,\chi)| \ll x \exp(-c_1 \sqrt{\log x}) \sum_{q \leq Q_1} \log(2Q/q)$$

$$\ll (\log x)^{A+2} x \exp(-c_1 \sqrt{\log x})$$

$$\ll x \exp(-\tfrac{2}{3} c_1 \sqrt{\log x}) .$$

$$\ll x^{1/2} Q (\log x)^3 .$$

By combining this estimate with our conclusion above, we conclude that

$$\sum_{1 \leq q \leq Q} \frac{\log(2Q/q)}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sup_{y \leq x} |\psi'(y,\chi)| \ll x^{\frac{1}{2}} Q (\log x)^3 ,$$

whence

$$\sum_{1 \leq q \leq Q} E^*(x; q) \ll x^{\frac{1}{2}} Q (\log x)^3 . \quad //$$

We have established an averaged bound for $\psi(x; q, a)$. These estimates can of course be converted into similarly powerful bounds for $\pi(x; q, a)$ on average. Define

$$\tilde{E}(x; q, a) = \pi(x; q, a) - \frac{li(x)}{\varphi(q)} ,$$

$$\tilde{E}(x; q) = \max_{\substack{1 \leq a \leq q \\ (a,q)=1}} |\tilde{E}(x; q, a)| ,$$

$$\tilde{E}^*(x; q) = \sup_{y \leq x} \tilde{E}(y; q) .$$

Corollary 12.2. Let $A>0$ be fixed. Then one has

$$\sum_{1\le q\le Q} \tilde{E}^*(x;q) \ll x^{\frac12} Q (\log x)^2,$$

provided only ~~that~~ $x^{1/2}(\log x)^{-A} \le Q \le x^{1/2}$.

Proof. We have

$$\psi(y;q,a) - \theta(y;q,a) \le \overbrace{\psi(y)-\theta(y)}^{\sum_{\substack{p^k\le y\\ k\ge 2}} \log p} \ll y^{\frac12},$$

whence

$$|\theta(y;q,a) - \frac{y}{\varphi(q)}| \ll E(y;q,a) + y^{\frac12}.$$

Thus

$$\pi(y;q,a) - \frac{li(y)}{\varphi(q)} = \int_{2^-}^{y} \frac{1}{\log u}\, d\,\theta(u;q,a) - \frac{li(y)}{\varphi(q)}$$

$$= \int_{2^-}^{y} \frac{1}{\log u}\, d\left(\theta(u;q,a) - \frac{u}{\varphi(q)}\right)$$

$$= \frac{\theta(u;q,a) - \frac{u}{\varphi(q)}}{\log u}\Bigg|_{2^-}^{y} + \int_2^y \frac{\theta(u;q,a) - \frac{u}{\varphi(q)}}{u(\log u)^2}\,du.$$

Observing ~~that~~

$$\theta(u;q,a) = \sum_{\substack{p\le u\\ p\equiv a \pmod q}} \log p \ll (\log u) \sum_{\substack{m\le u\\ m\equiv a \pmod q}} 1 \ll \frac{u\log u}{q},$$

it ~~therefore~~ follows ~~that~~ when $y\ge\sqrt{x}$,

$$\left|\pi(y;q,a) - \frac{li(y)}{\varphi(q)}\right| \ll \frac{\left|\theta(y;q,a) - \frac{y}{\varphi(q)}\right|}{\log y} + \int_2^{\sqrt{x}} \frac{(u\log u)/q}{u(\log u)^2}\,du$$

$$+ \int_{\sqrt{x}}^{y} \frac{E(u;q,a) + y^{1/2}}{u(\log u)^2}\,du.$$

When $y<\sqrt{x}$ we have the same bound, ~~though~~ omitting ~~the~~ last term on ~~the~~ right hand side. Hence, when $y\ge\sqrt{x}$,

$$\tilde{E}(y;q) \ll \frac{E(y;q)+y^{1/2}}{\log y} + \frac{x^{1/2}}{q} + y^{1/2} + E^*(y;q)\int_{\sqrt{x}}^{y}\frac{du}{u(\log u)^2}$$

$$\ll \frac{E^*(x;q) + x^{\frac{1}{2}}}{\log \sqrt{x}} + x^{\frac{1}{2}} + \frac{E^*(x;q)}{\log x}.$$

Meanwhile, when $y < \sqrt{x}$, we instead ~~obtain~~

$$\tilde{E}(y;q) \ll \frac{(y\log y)/q}{\log y} + \frac{1}{q}\int_2^{\sqrt{x}} \frac{du}{\log u} \ll \frac{\sqrt{x}}{q}.$$

Then
$$\tilde{E}^*(x;q) \ll \frac{E^*(x;q)}{\log x} + x^{\frac{1}{2}},$$

whence
$$\sum_{1\le q\le Q} \tilde{E}^*(x;q) \ll \frac{\sum_{1\le q\le Q} E^*(x;q)}{\log x} + Qx^{\frac{1}{2}}$$

$$\ll x^{\frac{1}{2}}Q(\log x)^2. \quad //$$

We present an application of the Bombieri–Vinogradov theorem in order to illustrate its utility, though we leave some assertions as exercises in order to save time. We begin with two preliminary lemmata.

Lemma 12.3. Suppose that X is a large natural number, and $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ are integers with $(a,q)=1$ and $q \le X$. Then, for some absolute constant $C > 0$, one has

$$\pi(X;q,a) \le \frac{CX}{\varphi(q)\log(X/q)}.$$

Note: The Brun–Titchmarsh theorem gives $C = 2 + o(1)$.

Proof. Apply the Large Sieve inequality. We take the set of integers $\mathcal{Z}$ to be integers m with $1 \le mq+a \le X$ such that $mq+a \not\equiv 0 \pmod{\pi}$ for any prime π

with $\pi \nmid q$ and $\pi \le \underset{\substack{\| \\ Q}}{\sqrt{X/q}}$. Then by Montgomery's sieve,

$$\pi(x;q,a) \ll \frac{X/q}{L},$$

where
$$L = \sum_{\substack{n \le Q \\ (n,q)=1}} \mu^2(n) \prod_{p \mid n} \frac{1}{p-1} \gg \sum_{\substack{n \le Q \\ (n,q)=1}} \frac{\mu^2(n)}{n}$$

$$\gg \frac{\varphi(q)}{q} \log Q \qquad \text{(exercise 1!).}$$

Thus
$$\pi(x;q,a) \ll \frac{X/q}{\frac{\varphi(q)}{q} \log \sqrt{\frac{X}{q}}} \ll \frac{X}{\varphi(q)\log(X/q)}.$$

//

Lemma 12.4 One has

$$\sum_{n \le x} \frac{1}{\varphi(n)} = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \log x + O(1).$$

Proof. Observe ~~that~~
$$\frac{n}{\varphi(n)} = \sum_{d \mid n} \frac{\mu(d)^2}{\varphi(d)} = \prod_{p \mid n}\left(1 + \frac{1}{p-1}\right),$$

Whence
$$\sum_{n \le x} \frac{1}{\varphi(n)} = \sum_{n \le x} \frac{1}{n} \sum_{d \mid n} \frac{\mu(d)^2}{\varphi(d)} = \sum_{1 \le d \le x} \frac{\mu(d)^2}{d\varphi(d)} \sum_{1 \le m \le X/d} \frac{1}{m}$$

$$= \sum_{1 \le d \le x} \frac{\mu(d)^2}{d\varphi(d)} \left(\log\left(\frac{x}{d}\right) + O(1) \right)$$

$$= \left(\sum_{1 \le d \le x} \frac{\mu(d)^2}{d\varphi(d)} \right) \log x + O(1)$$

$$= \sum_{d=1}^{\infty} \frac{\mu(d)^2}{d\varphi(d)} \cdot \log x + O(1).$$

Thus
$$\sum_{n\le x}\frac{1}{\varphi(n)} = \log x\prod_p\left(1+\frac{1}{p(p-1)}\right) + O(1)$$
$$= (\log x)\prod_p\left(\frac{p^2-p+1}{p(p-1)}\right) + O(1)$$
$$= (\log x)\prod_p\left(\frac{(p^3+1)(p^3-1)}{p(p^2-1)(p^3-1)}\right) + O(1)$$
$$= (\log x)\prod_p\frac{(1-p^{-6})}{(1-p^{-3})(1-p^{-2})} + O(1)$$
$$= \frac{\zeta(2)\zeta(3)}{\zeta(6)}\log x + O(1). \quad //$$

<u>Theorem 12.5</u> One has
$$\sum_{p\le x}\tau(p-1) = \frac{\zeta(2)\zeta(3)}{\zeta(6)}x + O\left(\frac{x\log\log x}{\log x}\right).$$

<u>Proof.</u> By applying the hyperbola method, we see ~~that~~
$$\sum_{p\le x}\tau(p-1) = \sum_{p\le x}\sum_{de=p-1}1$$
$$= \sum_{1\le d\le x^{\frac12}(\log x)^{-A}}\ \sum_{\substack{p\le x\\ p\equiv 1 \pmod d}}1$$
$$+ \sum_{1\le e\le x^{\frac12}(\log x)^{-A}}\ \sum_{\substack{p\le x\\ p\equiv 1 \pmod e}}1$$
$$+ \sum_{x^{\frac12}(\log x)^{-A}\le d\le x^{\frac12}(\log x)^{A}}\ \sum_{\substack{p\le x\\ p\equiv 1 \pmod d}}1$$
$$- \sum_{\substack{1\le d\le x^{1/2}(\log x)^{+A}\\ 1\le e\le x^{1/2}(\log x)^{-A}}}\ \sum_{p-1=de}1$$
(to remove the double counting in this range)

Then

$$\sum_{p \le x} \tau(p-1) = 2 \sum_{1 \le d \le x^{1/2} (\log x)^{-A}} \pi(x; d, 1)$$

$$+ \sum_{x^{1/2}(\log x)^{-A} \le d \le x^{\frac{1}{2}}(\log x)^{A}} \pi(x; d, 1)$$

$$- \sum_{1 \le d \le x^{\frac{1}{2}}(\log x)^{A}} \pi\left(d x^{\frac{1}{2}} (\log x)^{-A}; d, 1\right).$$

By Lemma 12.3, the second term here is

$$\ll \sum_{x^{\frac{1}{2}}(\log x)^{-A} \le d \le x^{\frac{1}{2}}(\log x)^{A}} \frac{x}{\varphi(d) \log(x/d)}$$

$$\ll \frac{x}{\log x} \sum_{x^{\frac{1}{2}}(\log x)^{-A} \le d \le x^{\frac{1}{2}}(\log x)^{A}} \frac{1}{\varphi(d)}.$$

By Lemma 12.4, we see ~~that~~ ~~this~~ is

$$\ll \frac{x}{\log x}\left(\log\left(\frac{x^{\frac{1}{2}}(\log x)^{A}}{x^{\frac{1}{2}}(\log x)^{-A}}\right) + O(1)\right) \ll \frac{x \log\log x}{\log x}.$$

Similarly, the third term is

$$\ll \sum_{1 \le d \le x^{\frac{1}{2}}(\log x)^{A}} \frac{d x^{\frac{1}{2}} (\log x)^{-A}}{\varphi(d) \log(x^{\frac{1}{2}} (\log x)^{-A})}$$

$$\ll x^{\frac{1}{2}} (\log x)^{-A-1} \cdot \sum_{1 \le d \le x^{\frac{1}{2}}(\log x)^{A}} \log\log d \qquad (\text{using } \varphi(d) \gg d/\log\log d)$$

$$\ll \frac{x}{\log x} \log\log x.$$

Hence, we may conclude thus far that

$$\sum_{p \le x} \tau(p-1) = 2 \sum_{1 \le d \le x^{\frac{1}{2}} (\log x)^{-A}} \pi(x; d, 1) + O\left(\frac{x \log\log x}{\log x}\right).$$

Meanwhile, by applying the Bombieri–Vinogradov theorem, it is apparent that the main term here is equal to

$$2 \sum_{1 \le d \le x^{\frac{1}{2}}(\log x)^{-A}} \frac{li(x)}{\varphi(d)} + O\Big(\sum_{1 \le d \le x^{\frac{1}{2}}(\log x)^{-A}} \tilde{E}^*(x;d) \Big)$$

$$= 2li(x) \sum_{1 \le d \le x^{\frac{1}{2}}(\log x)^{-A}} \frac{1}{\varphi(d)} + O\big(x^{\frac{1}{2}} (x^{\frac{1}{2}} (\log x)^{-A}) \cdot (\log x)^2 \big).$$

Then, again from Lemma 12·4, this first term is equal to

$$2 li(x) \Big(\frac{\zeta(2)\zeta(3)}{\zeta(6)} \log (x^{\frac{1}{2}} (\log x)^{-A}) + O(1) \Big) + O(x (\log x)^{2-A})$$

$$= li(x) \frac{\zeta(2)\zeta(3)}{\zeta(6)} \log x + O\Big(\frac{x \log\log x}{\log x} \Big). \qquad \text{(when } A \ge 3\text{, say)}.$$

By combining these estimates, we see that

$$\sum_{p \le x} \tau(p-1) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} x + O\Big(\frac{x \log\log x}{\log x} \Big). \quad //$$

§ 13. The Barban–Davenport–Halberstam theorem.

If one is prepared to average not merely over the modulus $q \le Q$, but also over the residue class $a \pmod q$, then significantly larger values of Q may be accommodated.

Theorem 13·1. (the Barban–Davenport–Halberstam theorem). Let $A > 0$ be fixed. Then whenever $x(\log x)^{-A} \le Q \le x$, one has

$$\sum_{q \le Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \Big(\psi(x;q,a) - \frac{x}{\varphi(q)} \Big)^2 \ll Qx \log x.$$

<u>Proof.</u> Recall that
$$\psi(x;q,a) - \frac{x}{\varphi(q)} = \frac{1}{\varphi(q)} \sum_{\chi} \bar{\chi}(a)\, \psi'(x,\chi).$$
Then by the orthogonality of Dirichlet characters, it follows that
$$\sum_{\substack{a=1\\(a,q)=1}}^{q} \left(\psi(x;q,a) - \frac{x}{\varphi(q)}\right)^2 = \sum_{\substack{a=1\\(a,q)=1}}^{q} \left|\frac{1}{\varphi(q)} \sum_{\chi\,(\mathrm{mod}\,q)} \bar{\chi}(a)\, \psi'(x,\chi)\right|^2$$
$$= \frac{1}{\varphi(q)} \sum_{\chi\,(\mathrm{mod}\,q)} |\psi'(x,\chi)|^2 .$$
Suppose that the character χ^* is primitive and induces the character $\chi \pmod q$. We have seen that in such circumstances, one has
$$\psi'(x,\chi^*) - \psi'(x,\chi) \ll \log^2(qx).$$
Thus
$$\sum_{q\le Q} \sum_{\substack{a=1\\(a,q)=1}}^{q} \left(\psi(x;q,a) - \frac{x}{\varphi(q)}\right)^2 \ll \sum_{q\le Q} \frac{1}{\varphi(q)} \sum_{\chi\,(\mathrm{mod}\,q)} \left(|\psi'(x,\chi^*)|^2 + \log^4(qx)\right)$$
$$\ll \left(\sum_{d\le Q} \sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,d)} |\psi'(x,\chi)|^2\right)\left(\sum_{\substack{q\le Q\\ d\mid q}} \frac{1}{\varphi(q)}\right) + Q\log^4(Qx).$$
$$\ll \sum_{q\le Q} \frac{\log(2Q/q)}{\varphi(q)} \sideset{}{^*}\sum_{\chi} |\psi'(x,\chi)|^2 + Q\log^4(Qx).$$
We now concentrate on the first term on the right hand side.

Let R be a real number with $1 < R \le Q$. Then one has for each q with $R < q \le 2R$ that $\psi'(x,\chi) = \psi(x,\chi)$ for any $\chi \pmod q$ with χ primitive. The contribution of such values of q in the right hand side above is
$$\ll \sum_{R<q\le 2R} \frac{\log(2Q/R)}{R} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi} |\psi(x,\chi)|^2$$
$$\overset{\text{large-sieve}}{\ll} \frac{\log(2Q/R)}{R}(x+R^2) \sum_{n\le x} \Lambda(n)^2 \ll \left(x^2/R + xR\right)(\log x)\log(2Q/R).$$
Thus, on summing over dyadic intervals $(R = 2^l R_0)$ with $R_0 = (\log x)^{A+2}$,

we find ~~that~~

$$\sum_{R_0 < q \le Q} \frac{\log(2Q/q)}{\varphi(q)} \sum_{\chi}^{*} |\psi'(x,\chi)|^2 \ll \frac{x^2}{R_0}(\log x)(\log Q) + xQ\log x$$

$$\ll x^2(\log x)^{-A} + Qx\log x .$$

On the other hand, when $q \le (\log x)^{A+2}$, it follows from the Siegel - Walfisz theorem that there is a constant $c > 0$ for which

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

Hence

$$\sum_{1 \le q \le R_0} \frac{\log(2Q/q)}{\varphi(q)} \sum_{\chi}^{*} |\psi(x,\chi)|^2 \ll x^2(\log x)^{A+3}\exp(-2c\sqrt{\log x})$$

$$\ll x^2(\log x)^{-A} .$$

We therefore conclude that when $x(\log x)^{-A} \le Q \le x$, one has

$$\sum_{q \le Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left(\psi(x;q,a) - \frac{x}{\varphi(q)}\right)^2 \ll Qx\log x + x^2(\log x)^{-A}$$

$$\ll Qx\log x.$$

Notice that the conclusion of Theorem 13.1 shows that, for almost all pairs (q,a) with $1 \le q \le Q$ and $(a,q) = 1$, with $x(\log x)^{-A} \le Q \le x$, one has

$$\psi(x;q,a) - \frac{x}{\varphi(q)} \ll \left(\frac{Qx\log x}{Q^2}\right)^{\frac{1}{2}} \simeq \frac{(x\log x)^{\frac{1}{2}}}{Q^{1/2}} .$$

The right hand side here is $o(x/\varphi(q))$ whenever $Q \le x/(\log x)^{1+\delta}$, suitable $\delta > 0$.

We remark that, as was shown by Montgomery, and subsequently refined by $\underbrace{\text{Hooley - Hooley - ... - Hooley}}_{17}$ and Vaughan, one can derive ~~the~~ asymptotic formula

$$\sum_{q \le Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left(\psi(x;q,a) - \frac{x}{\varphi(q)}\right)^2 = Qx\log x + O(Qx),$$

which shows ~~that~~ there is a non-trivial secondary contribution to the asymptotics for $\psi(x;q,a)$.

§14. Further properties of zeros of L-functions: Hadamard products.

We seek to provide explicit formulae for $\psi(x)$ and $\psi(x;q,a)$ in terms of zeros of $\zeta(s)$ and of $L(s,\chi)$, and to this end we return (from ANT 1) to the distribution of zeros of these L-functions, beginning with $\zeta(s)$.

Recall that the function

$$\xi(s) = \tfrac{1}{2}s(s-1)\zeta(s)\Gamma(s/2)\pi^{-s/2}$$

is entire, and that $\xi(s) = \xi(1-s)$. Since $\xi(s)$ is entire, it has no poles, but it may have zeros. We would like to characterise the function $\frac{\xi'}{\xi}(s)$ in order to investigate $\frac{\zeta'}{\zeta}(s)$ (in order to more carefully investigate $\psi(x)$). We are therefore led to the study of what may happen when an entire function has infinitely many zeros. By analogy with the situation for polynomials, it is tempting to believe that

$$\xi(s) = c\prod_{\rho}\left(1-\frac{s}{\rho}\right),$$

where the product is taken over all (perhaps infinitely many) zeros of $\xi(s)$. Such conclusions are addressed in the next lemma.

<u>Lemma 14·1.</u> Let $f(z)$ be an entire function having a zero of order K at $z=0$, and with the property that $f(z)=0$ for $z\in\{z_k : k\in\mathbb{N}\}$. Suppose also that there exists a constant θ, with $1<\theta<2$, having the property that for R sufficiently large, one has

$$\max_{|z|\le R}|f(z)| \le \exp(R^{\theta}).$$

Then there exist $A=A(f)$ and $B=B(f)$ (complex numbers) with the property that

$$f(z) = z^K e^{A+Bz}\prod_{k=1}^{\infty}\left(1-\frac{z}{z_k}\right)e^{z/z_k} \qquad (z\in\mathbb{C}).$$

Here, the product is uniformly convergent for z in any compact set. There are more general conclusions available, but this suffices for our purposes.

Proof. We may simplify to the situation with $K=0$ by replacing, if necessary, $f(z)$ by $f(z)/z^K$. Let

$$N_f(R) = \#\{\text{zeros of } f(z) \text{ lying in } |z|\le R\}.$$

By Jensen's inequality, the number of zeros of $f(z)$ lying in $|z|\le R$ is

$$N_f(R) \le \frac{\log\big(\overbrace{\exp((2R)^\theta)}^{\max |f(z)| \text{ in } |z|\le 2R}/|f(0)|\big)}{\log(2R/R)} \le 8R^\theta. \qquad (R \text{ large}).$$

Then

$$\sum_{R<|z_k|\le 2R} |z_k|^{-2} \le 8(2R)^{\theta-2},$$

whence by summing over dyadic intervals, noting that $\theta<2$, we obtain

$$\sum_{k=1}^{\infty} |z_k|^{-2} < \infty.$$

Since $(1-z)e^z = 1 + O(|z|^2)$ uniformly for $|z|\le 1$, it follows that

$$g(z) = \prod_{k=1}^{\infty}\Big(1-\frac{z}{z_k}\Big)e^{z/z_k}$$

is uniformly convergent in compact regions, and hence represents an entire function. In particular, one sees that

$$h(z) = \frac{f(z)}{f(0)\,g(z)}$$

is a non-vanishing entire function with $h(0)=1$.

We now turn to examine the function $h(z)$ through the medium of the Borel–Carathéodory theorem, and first we determine its order of growth. With this in mind, we define

$$P_i(z) = \prod_{k \in K_i} \left(1 - \frac{z}{z_k}\right) e^{z/z_k},$$

where

$$K_1 = \{k \in \mathbb{N} : \ |z_k| \le R/2\},$$
$$K_2 = \{k \in \mathbb{N} : \ R/2 < |z_k| \le 3R\},$$
$$K_3 = \{k \in \mathbb{N} : \ |z_k| > 3R\}.$$

Then $g(z) = P_1(z) P_2(z) P_3(z)$, a product we now investigate when $R \le |z| \le 2R$.

Observe first that when $|z_k| \le R/2$, one has

$$|1 - z/z_k| \ge |z/z_k| - 1 \ge 1,$$

whence

$$|P_1(z)| \ge \prod_{k \in K_1} e^{-2R/|z_k|}.$$

Since

$$\sum_{k \in K_1} \frac{1}{|z_k|} \le \frac{R}{2} \sum_{\substack{k \in K_1 \\ |z_k| \le R/2}} \frac{1}{|z_k|^2} \ll R \cdot R^{\theta - 2} = R^{\theta - 1},$$

it follows that

$$|P_1(z)| \ge e^{-c_1 R^\theta},$$

for a suitable constant $c_1 > 0$.

Next, since

$$\operatorname{card}(K_2) \le (3R)^2 \sum_{R/2 < |z_k| \le 3R} |z_k|^{-2}$$
$$\le c_2 R^\theta,$$

for a suitable $c_2 > 0$, there is some gap between zeros (in absolute value) $\in [R, 2R]$ of size $\ge \dfrac{R}{c_2 R^\theta} = \dfrac{1}{c_2} R^{1-\theta}$. Hence, there exists a real number r with $R \le r \le 2R$ for which $|r - |z_k|| \ge 1/R^2$ for all $k \in K_2$. Fix such a choice of r and consider any complex number with $|z| = r$. We have

$$|1 - z/z_k| \ge \frac{|r - |z_k||}{|z_k|} \ge \frac{1}{3R^3} \qquad (k \in K_2),$$

and consequently

$$|P_2(z)| \geq \prod_{k\in K_2} \left(1-\frac{z}{z_k}\right)e^{z/z_k}$$

$$\geq \left(\frac{1}{3R^3}\right)^{c_2R^\theta}\cdot\left(e^{-2r/R}\right)^{c_2R^\theta} \geq e^{-c_3R^\theta\log R},$$

for a suitable $c_3>0$, and whenever $|z|=r$.

As for K_3, we observe that when $|z|\leq 2R$ one has

$$|P_3(z)| \geq \prod_{k\in K_3} e^{-c_4R^2/|z_k|^2} \geq e^{-c_5R^\theta},$$

for a suitable $c_5>0$.

By combining these estimates, we see that when R is large there exists a real number $r\in[R,2R]$ having the property that

$$|g(z)| = |P_1(z)P_2(z)P_3(z)| \geq e^{-c_6R^\theta\log R},$$

whenever $|z|=r$. Hence $|h(z)| = \dfrac{|f(z)|}{|f(0)g(z)|} \leq \dfrac{e^{c_6R^\theta\log R}e^{R^\theta}}{|f(0)|}$

$$\leq e^{c_7R^\theta\log R},$$

under the same condition. Thus, by the maximum modulus principle, we have

$$\max_{|z|\leq R}|h(z)| \leq e^{c_7R^\theta\log R}.$$

(an entire function!)

By putting $W(z) = \log h(z)$ and noting that $W(0)=0$, we see that $\mathrm{Re}(W(z)) \leq c_7R^\theta\log R$ for large R, whence by the Borel–Carathéodory theorem,

$$|W(z)| \ll R^\theta\log R.$$

Since $\theta<2$, it follows that $W(z)$ must be a polynomial of degree at most 1, so that $W(z) = A+Bz$ for some complex numbers A and B. Hence

$$f(z) = e^{A+Bz}\prod_{k=1}^{\infty}\left(1-\frac{z}{z_k}\right)e^{z/z_k}.$$

//

In order to apply Lemma 14·1 to $\xi(s)$ we require an upper bound for $|\xi(s)|$. From the bound

$$\zeta(s) \ll (1+\tau^{1-\sigma}) \min\left\{\frac{1}{|\sigma-1|}, \log\tau\right\} \qquad (\delta \le \sigma \le 2,\ |t| \ge 1)$$

(see Lemma 10·4 of ANT 1) and the estimate

$$\zeta(s) = \frac{1}{s-1} + O(1) \qquad (\delta \le \sigma \le 2 \text{ and } |t| \le 1),$$

we see that

$$\zeta(s) \ll |s|^{1/2} \quad \text{for} \quad \sigma \ge \tfrac{1}{2} \text{ and } |s| \ge 2.$$

By Stirling's formula, one has

$$\Gamma(s) = \sqrt{2\pi}\, s^{s-1/2} e^{-s} (1 + O(1/|s|)),$$

whence

$$\Gamma(s/2) \ll \exp(\tfrac{2}{3}|s| \log|s|),$$

and furthermore

$$\xi(s) = \tfrac{1}{2} s(s-1)\zeta(s)\Gamma(s/2)\pi^{-s/2}$$

$$\ll \exp(|s| \log|s|). \qquad (\sigma \ge \tfrac{1}{2} \text{ \& } |s| \ge 2).$$

Also, we have

$$\xi(s) = (s-1)\zeta(s) \cdot \underbrace{\tfrac{1}{2} s\,\Gamma(s/2)}_{\Gamma(1+s/2)} \pi^{-s/2},$$

so that since $\xi(s) = \xi(1-s)$, we have

$$\xi(0) = \xi(1) = \left(\lim_{s\to 1} (s-1)\zeta(s)\right) \Gamma(3/2)\pi^{-1/2}$$

$$= \tfrac{1}{2}.$$

Then as a corollary of Lemma 14·1, we obtain the following conclusion.

<u>Corollary 14·2.</u> There is a constant B having the property that

$$\xi(s) = \tfrac{1}{2} e^{Bs} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho},$$

for all $s \in \mathbb{C}$, where the product is taken over all the zeros ρ of $\xi(s)$.

<u>Proof:</u> We have $\max_{|z| \le R} |\xi(z)| \le \exp(R^{\theta})$, for any $\theta > 1$, and

$\xi(0) \neq 0$, so ~~there are constants A and B such that~~

$$\xi(z) = e^{A+Bz} \prod_{\rho} (1 - z/\rho) e^{z/\rho}.$$

By noting that $\frac{1}{2} = \xi(0) = e^{A}$, the desired conclusion follows. //

The precise identity of the constant B remains in question. We presently apply our knowledge of $\zeta(s)$ to give B explicitly. We pause, however, to remark that it is already evident that there are infinitely many zeros of $\xi(s)$. For if $\xi(s)$ had at most finitely many zeros, then one would have $\xi(s) \ll \exp(C|s|)$ for some $C > 0$ and s large, yet by Stirling's formula

$$\xi(\sigma) = \exp\left(\tfrac{1}{2}\sigma\log\sigma + O(\sigma)\right) \quad \text{as } \sigma \to \infty. \;\#.$$

<u>Corollary 14.3.</u> Let

$$B = -\frac{C_0}{2} - 1 + \tfrac{1}{2}\log(4\pi) = -0.0230957\ldots .$$

Then one has

$$B = -\tfrac{1}{2}\sum_{\rho}\left(\frac{1}{1-\rho} + \frac{1}{\rho}\right) = -\sum_{\rho}\mathrm{Re}\left(\frac{1}{\rho}\right),$$

where the sum is taken over all the zeros of $\xi(s)$ taken with multiplicity,

$$\frac{\xi'}{\xi}(s) = B + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right), \qquad (14.1)$$

and

$$\frac{\zeta'}{\zeta}(s) = B + \tfrac{1}{2}\log\pi - \frac{1}{s-1} - \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\tfrac{1}{2}s+1\right) + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right). \qquad (14.2)$$

<u>Proof.</u> First, by taking logarithmic derivatives in the conclusion of Corollary 14.2, we see that

$$\frac{\xi'}{\xi}(s) = B + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right),$$

confirming (14.1). Next, recalling that $\xi(s) = \frac{1}{2}s(s-1)\zeta(s)\Gamma(s/2)\pi^{-s/2}$, we find that

$$\frac{\xi'}{\xi}(s) = \frac{1}{s} + \frac{1}{s-1} + \frac{\zeta'}{\zeta}(s) + \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}\right) - \frac{1}{2}\log\pi$$

$$\underbrace{= \frac{1}{s-1} + \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{1}{2}s+1\right) + \frac{\zeta'}{\zeta}(s) - \frac{1}{2}\log\pi}_{(14\cdot 3)} \quad \left(\text{using } \Gamma\left(\tfrac{1}{2}s+1\right) = \tfrac{1}{2}s\,\Gamma\left(\tfrac{1}{2}s\right)\right).$$

Then (14·2) is immediate from (14·1).

Next, since $\xi(s) = \xi(1-s)$, we have

$$\frac{\xi'}{\xi}(s) = -\frac{\xi'}{\xi}(1-s),$$

Whence (14·1) shows that

$$\left.\begin{aligned} \frac{\xi'}{\xi}(1) &= B + \sum_\rho \left(\frac{1}{1-\rho} + \frac{1}{\rho}\right) \\ &\parallel \\ -\frac{\xi'}{\xi}(0) &= -\left(B + \sum_\rho\left(\frac{-1}{\rho} + \frac{1}{\rho}\right)\right) = -B \end{aligned}\right\} \quad B = -\frac{1}{2}\sum_\rho\left(\frac{1}{1-\rho} + \frac{1}{\rho}\right).$$

Also, from (14·3) we have

$$\frac{\xi'}{\xi}(0) = -1 + \frac{1}{2}\frac{\Gamma'}{\Gamma}(1) + \frac{\zeta'}{\zeta}(0) - \frac{1}{2}\log\pi.$$

Now $\Gamma'(1) = -C_0$, $\zeta(0) = -\frac{1}{2}$, and $\zeta(s) = \zeta(1-s)\,2^s\pi^{s-1}\Gamma(1-s)\sin\left(\frac{\pi s}{2}\right)$.

Then

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

$$\parallel$$

$$1 + C_0(s-1) + \cdots$$

(Differentiate and take $s=1$)

$$\Rightarrow \quad C_0 = 2\zeta'(0) - 2\zeta(0)\log(2\pi) + 2\zeta(0)\Gamma'(1)$$

$$\Downarrow$$

$$\zeta'(0) = -\frac{1}{2}\log(2\pi).$$

Then we deduce that

$$B = \frac{\xi'}{\xi}(0) = -1 - \frac{1}{2}C_0 + \log(2\pi) - \frac{1}{2}\log\pi$$

$$= -\frac{1}{2}C_0 - 1 + \frac{1}{2}\log(4\pi).$$

This confirms the identity given first for B. Since B is real, we see that

$$B = -\frac{1}{2}\sum_\rho\left(\mathrm{Re}\left(\frac{1}{1-\rho}\right) + \mathrm{Re}\left(\frac{1}{\rho}\right)\right).$$

The zeros ρ have real parts between 0 and 1, and hence each of these sums are absolutely convergent provided that the zeros are not too dense. We will show below that if

$$N(T) := \#\{\rho = \beta + i\gamma : \xi(\rho) = 0 \text{ and } 0 < \beta < 1,\ 0 < \gamma \leq T\},$$

then

$$N(T+1) - N(T) \ll \log(T+2). \qquad (14.4)$$

Then one sees that since $0 < Re(\rho) < 1$, one has

$$\sum_{\substack{\rho \\ T < \gamma \leq T+1}} Re\left(\frac{1}{\rho}\right) \ll \frac{\log(T+2)}{T^2} \qquad \left(\text{Note: } \frac{1}{\beta + i\gamma} = \frac{\beta - i\gamma}{\beta^2 + \gamma^2}\right),$$

and the absolute convergence follows. The corresponding argument with $1-\rho$ in place of ρ follows mutatis mutandis. But if $\xi(\rho) = 0$ then $\xi(1-\rho) = 0$, whence

$$B = -\frac{1}{2} \sum_\rho \left(\frac{1}{1-\rho} + \frac{1}{\rho}\right) = -\sum_\rho Re\left(\frac{1}{\rho}\right). \quad //$$

It remains to justify our assertion concerning $N(T)$. Here we apply Jensen's formula to $\zeta(s)$ on a disc with centre $2 + i(T + 1/2)$ and radius $R = 11/6$. If we take $r = 7/4$ in Jensen's formula, we see that the number of zeros ρ in $\frac{1}{2} \leq \beta \leq 1$, $T \leq \gamma \leq T+1$ is

$$\ll \frac{\log\left(T / \overbrace{|\zeta(2 + i(T+1/2))|}^{\gg 1}\right)}{\log\left(\frac{11/6}{7/4}\right)} \ll \log T.$$

But there are no zeros of $\Gamma(s)$ in this region, since (Weierstrass product)

$$\Gamma(s) = \frac{e^{-C_0 s}}{s} \prod_{n=1}^{\infty} \frac{e^{s/n}}{1 + s/n},$$

and hence the only zeros of $\xi(s)$ in this region are those of $\zeta(s)$. Moreover, if ρ is a zero of $\xi(s)$ then so too is $1 - \bar{\rho}$, and vice versa, so the rectangle $0 \leq \beta \leq \frac{1}{2}$, $T \leq \gamma \leq T+1$ contains

 the same number of zeros as does $\frac{1}{2} \le \beta \le 1$, $T \le \gamma \le T+1$. Hence
$$N(T+1) - N(T) \ll \log(T+2). \square$$

Note that it follows from the Schwarz reflection principle ~~that~~ $\xi(\bar{s}) = \overline{\xi(s)}$, so that for each zero ρ of $\xi(s)$, one has
$$0 = \xi(\rho) = \overline{\xi(\bar{\rho})},$$
so that $\bar{\rho}$ is a zero, whence also $1-\bar{\rho}$ is a zero, using the functional equation $\xi(s) = \xi(1-s)$.

§15. An explicit formula for $\psi(x)$.

Write
$$\psi_0(x) = \tfrac{1}{2}(\psi(x^+) + \psi(x^-)).$$
Then it follows from Perron's formula ~~that~~
$$\psi_0(x) = \lim_{T\to\infty} -\frac{1}{2\pi i}\int_{\sigma_0 - iT}^{\sigma_0 + iT} \frac{\zeta'}{\zeta}(s)\,\frac{x^s}{s}\,ds,$$

for $\sigma_0 > 1$. The formula (14·2) enables us to compute this contour integral in terms of the zeros ρ of $\xi(s)$, the pole of $\zeta(s)$ at $s=1$, and the zeros of $\Gamma(\frac{s}{2})$ at the negative even integers. Our goal in this section is the following explicit formula:
$$\psi_0(x) = x - \lim_{T\to\infty}\sum_{\substack{\rho \\ |\gamma|\le T}} \frac{x^\rho}{\rho} - \frac{\zeta'}{\zeta}(0) + \sum_{k=1}^{\infty}\frac{x^{-2k}}{2k}.$$

$$= x - \lim_{T\to\infty}\sum_{\substack{\rho \\ |\gamma|\le T}} \frac{x^\rho}{\rho} - \log(2\pi) - \frac{1}{2}\log(1 - 1/x^2).$$

(Here and throughout, $\rho = \beta + i\gamma$).

We begin with some auxiliary estimates.

Lemma 15·1 Uniformly for $-1 \le \sigma \le 2$, one has
$$\frac{\zeta'}{\zeta}(s) = -\frac{1}{s-1} + \sum_{\substack{\rho \\ |\gamma - t|\le 1}} \frac{1}{s-\rho} + O(\log\tau).$$

$(\tau = |t|+3)$.

Proof. From equation (14·2) of Corollary 14·3, one sees ~~that~~

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

$$= -\frac{1}{s-1} + \sum_{\rho}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right) - \tfrac{1}{2}\log\tau + O(1)$$

$$\left(\text{using } \frac{\Gamma'}{\Gamma}(s) = \log s + O\left(\frac{1}{|s|}\right)\right).$$

Now $\frac{\zeta'}{\zeta}(2+it) \ll 1$, and thus by considering $\frac{\zeta'}{\zeta}(\sigma+it) - \frac{\zeta'}{\zeta}(2+it)$, we obtain

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

Next, consider for a given $n \in \mathbb{N}$ ~~the zeros~~ ρ ~~satisfying~~

$$n \le |\gamma - t| \le n+1.$$

Since

$$\frac{1}{s-\rho} - \frac{1}{2+it-\rho} = \frac{2-\sigma}{(s-\rho)(2+it-\rho)} \ll \frac{1}{n^2},$$

we deduce ~~that~~

$$\sum_{\substack{\rho \\ |\gamma-t|>1}}\left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) \ll \sum_{n=1}^{\infty}\frac{N(t+n+1)-N(t+n)+N(t-n)-N(t-n-1)}{n^2}$$

$$\overset{(14\cdot4)}{\ll} \sum_{n=1}^{\infty}\frac{\log(\tau+n)}{n^2} \ll \log\tau.$$

Hence

$$\frac{\zeta'}{\zeta}(s) = -\frac{1}{s-1} + \sum_{\substack{\rho \\ |\gamma-t|\le 1}}\left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) + O(\log\tau).$$

But

$$\sum_{\substack{\rho \\ |\gamma-t|\le 1}}\frac{1}{2+it-\rho} \ll \sum_{\substack{\rho \\ |\gamma-t|\le 1}} 1 \overset{(14\cdot4)}{\ll} \log\tau,$$

and thus we conclude ~~that~~

$$\frac{\zeta'}{\zeta}(s) = -\frac{1}{s-1} + \sum_{\substack{\rho \\ |\gamma-t|\le 1}}\frac{1}{s-\rho} + O(\log\tau). \quad //$$

Lemma 15.2. Let $T \geq 2$ be a real number. Then ~~there~~ exists a real number T_1 with $T \leq T_1 \leq T+1$ for which

$$\frac{\zeta'}{\zeta}(\sigma + iT_1) \ll (\log T)^2$$

uniformly for $-1 \leq \sigma \leq 2$.

Proof. We have shown ~~that~~

$$N(T+1) - N(T) \ll \log (T+2),$$

and hence $\exists\, T_1 \in [T, T+1]$ ~~such that~~ for all ~~zeros~~ $\rho = \beta + i\gamma$ of $\zeta(s)$, one has $|T_1 - \gamma| \gg \frac{1}{\log T}$. But by Lemma 15.1 one has

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

and each summand here is $O(\log T)$ when $s = \sigma + iT_1$. Thus

$$\frac{\zeta'}{\zeta}(\sigma + iT_1) \ll (\log T) \sum_{\substack{\rho \\ |\gamma - T_1| \leq 1}} 1 + O(\log T) \ll (\log T)^2. \;//$$

There is a modicum of irritation caused by the ~~zeros~~ of $\zeta(s)$ at negative even integers, and this is resolved by the following additional estimate.

Lemma 15.3. Let

$$\mathcal{B} = \{ s \in \mathbb{C} : \ \sigma \leq -1 \ \text{ and } \ |s + 2k| \geq \tfrac{1}{4} \text{ for all } k \in \mathbb{N} \}.$$

Then, uniformly for $s \in \mathcal{B}$, one has

$$\frac{\zeta'}{\zeta}(s) \ll \log (|s| + 1).$$

Proof. Recall that

$$\zeta(s) = \zeta(1-s)\, 2^s \pi^{s-1}\, \Gamma(1-s) \sin\left(\tfrac{1}{2}\pi s\right),$$

whence

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

When $\mathrm{Re}(s) \leq -1$, one has $\frac{\zeta'}{\zeta}(1-s) \ll 1$, and hence

$$\frac{\zeta'}{\zeta}(s) \ll 1 + \left|\frac{\Gamma'}{\Gamma}(1-s)\right| + \left|\frac{e^{i\pi s}+1}{e^{i\pi s}-1}\right|$$

$$\ll 1 + \left|\frac{\Gamma'}{\Gamma}(1-s)\right| + \frac{1}{|e^{i\pi s}-1|} .$$

The last term here is absolutely bounded, since $|s+2k| \geqslant \frac{1}{4}$ for all $k \in \mathbb{N}$. On the other hand, one has

$$\frac{\Gamma'}{\Gamma}(1-s) \ll \log(|s|+1).$$

Then we conclude that

$$\frac{\zeta'}{\zeta}(s) \ll 1 + \log(|s|+1) \quad \text{for} \quad s \in \mathcal{S}.$$
$$\ll \log(|s|+1) .$$

//

~~Lemma~~ **Theorem 15.4.** Fix a real number $c > 1$, and suppose that $x \geqslant c$ and $T \geqslant 2$. Then one has

$$\psi_0(x) = x - \sum_{\substack{\rho \\ |\gamma| \leq T}} \frac{x^\rho}{\rho} - \log 2\pi - \frac{1}{2}\log\left(1 - \frac{1}{x^2}\right) + R(x,T),$$

where

$$R(x,T) \ll \log x \min\left\{1, \frac{x}{T\langle x\rangle}\right\} + \frac{x}{T}\log^2(xT),$$

in which we use the notation

$$\langle\theta\rangle := \min\{|\theta - p^h| : p \text{ prime}, h \in \mathbb{N} \ \& \ \theta \neq p^h\}.$$

Proof: Consider a real number $T \geqslant 2$. As a consequence of Lemma 15.2, there exists a real number T_1 with $T \leq T_1 \leq T+1$ such that, uniformly for $-1 \leq \sigma \leq 2$, one has

$$\frac{\zeta'}{\zeta}(\sigma + iT_1) \ll (\log T)^2.$$

Then by the explicit version of Perron's formula, we find ~~that~~ with $\sigma_0 = 1 + 1/\log x$, one has

$$\psi_0(x) = -\frac{1}{2\pi i}\int_{\sigma_0 - iT_1}^{\sigma_0 + iT_1} \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, ds + R_1,$$

where
$$R_1 \leqslant \sum_{\substack{\frac{x}{2} < n \leqslant 2x \\ n \neq x}} \Lambda(n) \min\left\{1, \frac{x}{T|x-n|}\right\} + \frac{x}{T} \underbrace{\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\sigma_0}}}_{\substack{\| \\ -\frac{\zeta'}{\zeta}(\sigma_0) \asymp \frac{1}{\sigma_0 - 1} = \log x}} .$$

Since
$$\sum_{x+1 \leq n < 2x} \Lambda(n) \min\left\{1, \frac{x}{T|x-n|}\right\} \ll \sum_{x+1 \leq n < 2x} \frac{x \log x}{T(n-x)} \ll \frac{x}{T} (\log x)^2,$$

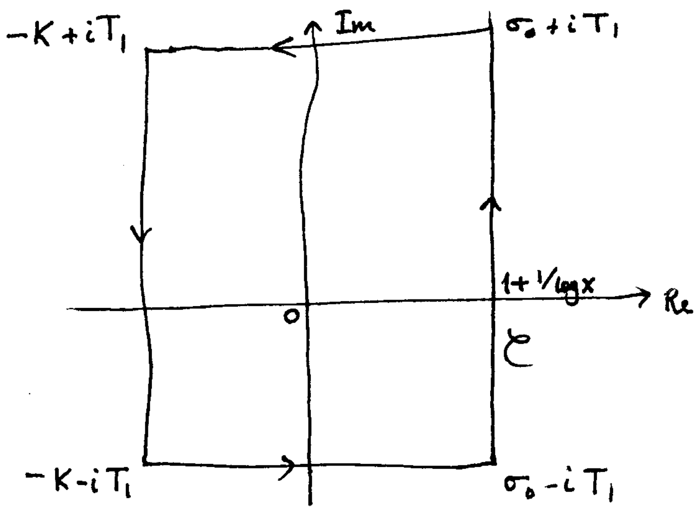
and the contribution arising analogously from ~~those terms~~ with $\frac{x}{2} < n \leq x-1$ may be treated in similar fashion. Thus

$$R_1 \ll \frac{x}{T} (\log x)^2 + \sum_{\substack{x-1 < n < x+1 \\ n \neq x}} \Lambda(n) \min\left\{1, \frac{x}{T|x-n|}\right\}$$

$$\ll \frac{x}{T} (\log x)^2 + (\log x) \min\left\{1, \frac{x}{T\langle x \rangle}\right\} .$$

We now turn to consider the main term in our formula for $\psi_0(x)$.

Let $K = 2k+1$ for some fixed $k \in \mathbb{N}$ (we're avoiding negative even integers!) Then it follows from Cauchy's integral theorem that we can evaluate $\psi_0(x)$ by examining the contour integral over the path $\mathcal{C}$. We have



$$\psi_0(x) = -\frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} \, ds + R_1 + R_2 ,$$

where
$$R_2 = \frac{1}{2\pi i} \int_{-K-iT_1}^{\sigma_0 - iT_1} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} \, ds + \frac{1}{2\pi i} \int_{-K+iT_1}^{-K-iT_1} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} \, ds + \frac{1}{2\pi i} \int_{\sigma_0+iT_1}^{-K+iT_1} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} \, ds .$$

From Corollary 14·3 we have

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

and hence

$$-\frac{1}{2\pi i}\int_{\mathcal{C}} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}\,ds = \underbrace{x}_{\substack{\text{simple pole at}\\ s=1}} - \underbrace{\sum_{\substack{\rho\\ |\gamma|<T_1}} \frac{x^{\rho}}{\rho}}_{\substack{\text{simple poles}\\ \text{at } s=\rho=\beta+i\gamma\\ \text{for } |\gamma|<T_1}} + \underbrace{\sum_{l=1}^{k}\frac{x^{-2l}}{2l}}_{\substack{\text{simple poles}\\ \text{of } \frac{\Gamma'}{\Gamma}(\frac{1}{2}s+1)\\ \text{(see below)}}} - \underbrace{\frac{\zeta'}{\zeta}(0)}_{\substack{\text{simple pole}\\ \text{at } s=0\\ \text{of } \frac{x^s}{s}}}.$$

Here, we made use of the Weierstrass product

$$\Gamma(s) = \frac{e^{-C_0 s}}{s}\prod_{n=1}^{\infty}\frac{e^{s/n}}{1+s/n}$$

to deduce that

$$\frac{\Gamma'}{\Gamma}(s) = -\frac{1}{s} - C_0 - \sum_{n=1}^{\infty}\left(\frac{1}{s+n} - \frac{1}{n}\right),$$

Whence $\frac{\Gamma'}{\Gamma}(\frac{1}{2}s+1)$ has simple poles at $s=-2l$ $(l\in\mathbb{N})$ with residues -1.

We may conclude thus far that

$$\psi_0(x) = x - \sum_{\substack{\rho\\ |\gamma|<T_1}}\frac{x^{\rho}}{\rho} - \log(2\pi) - \tfrac{1}{2}\log\left(1-\frac{1}{x^2}\right) + O\left(\frac{x}{T}(\log x)^2\right) + R_2 + (\log x)\min\left\{1, \frac{x}{T\langle x\rangle}\right\}.$$

In order to estimate R_2, we proceed directly with each of the three terms. Observe first that $|\sigma \pm iT_1| \geq T_1 \geq T$ (for $\sigma\in\mathbb{R}$), and hence Lemma 15·2 shows that

$$\int_{-1\pm iT_1}^{\sigma_0\pm iT_1}\frac{\zeta'}{\zeta}(s)\frac{x^s}{s}\,ds \ll \frac{(\log T)^2}{T}\int_{-1}^{\sigma_0} x^{\sigma}\,dx \ll \frac{x(\log T)^2}{T\log x} \ll \frac{x}{T}(\log T)^2,$$

When $\sigma\in[-K,-1]$, meanwhile, one has $|\sigma+2k\pm iT_1| \geq T_1 \geq \frac{1}{4}$ $(k\in\mathbb{N})$, and hence $\frac{\zeta'}{\zeta}(\sigma\pm iT_1) \ll \log(|\sigma\pm iT_1|+1)$ by Lemma 15·3. But under these conditions we have

$$\frac{\log|\sigma\pm iT_1|}{|\sigma\pm iT_1|} \ll \frac{\log(\sigma^2+T^2)}{\sqrt{\sigma^2+T^2}} \ll \frac{\log T}{T},$$

and hence

$$\int_{-K\pm iT_1}^{-1\pm iT_1} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}\,ds \ll \frac{\log T}{T}\int_{-\infty}^{-1} x^\sigma\,d\sigma \ll \frac{\log T}{xT\log x} \ll \frac{\log T}{T}.$$

Also, since $|-K+it| \geq K$ for $t \in [-T_1, T_1]$, and one has

$$\left|-K+it+2k\right| \geq (1+t^2)^{1/2} \geq \tfrac{1}{4} \quad \text{for } k \in \mathbb{N},$$

it follows from Lemma 15.3 ~~that~~

$$\int_{-K-iT_1}^{-K+iT_1} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}\,ds \ll \frac{\log(KT)}{K}\,x^{-K}\int_{-T_1}^{T_1} dt \ll \frac{T\log(KT)}{Kx^K}.$$

We ~~therefore~~ conclude ~~that~~

$$R_2 \ll \frac{x}{T}(\log T)^2 + \frac{T\log(KT)}{Kx^K}.$$

By ~~substituting~~ this estimate into our earlier formula for $\psi_0(x)$, we conclude ~~that~~

$$\psi_0(x) = x - \sum_{\substack{\rho \\ |\gamma|<T_1}} \frac{x^\rho}{\rho} - \log(2\pi) - \tfrac{1}{2}\log\left(1-\frac{1}{x^2}\right) + O\left(\frac{x}{T}\log^2(xT) + \frac{T\log(KT)}{Kx^K}\right)$$
$$+\ O\left((\log x)\min\left\{1, \frac{x}{T\langle x\rangle}\right\}\right).$$

If we now take the limit as $K \to \infty$, then we obtain the conclusion of the theorem save for one detail, which is ~~that~~ ~~the sum~~ over ρ is over ~~zeros~~ ~~satisfying~~ $|\gamma| < T_1$ instead of $|\gamma| < T$. However, since $N(T+1) - N(T) \ll \log(T+2)$, it follows ~~that~~

$$\sum_{\substack{\rho \\ T<|\gamma|<T_1}} \frac{x^\rho}{\rho} \ll \log(T+2)\,\frac{x}{T},$$

which is an acceptable error. //

§16. Zeros of $\zeta(s)$ and the error term in the Prime Number Theorem.

In Lemma 15.4 we have an explicit formula for $\psi(x)$. We now explore what this has to say for the size of the error term in the Prime Number Theorem. We begin with the consequences of:

The Riemann Hypothesis: For all zeros ρ of $\zeta(s)$, one has $Re(\rho) = \frac{1}{2}$.

Theorem 16.1. Assume the truth of the Riemann Hypothesis. Then for $x \geq 2$ one has

$$\psi(x) = x + O(x^{1/2}(\log x)^2),$$

$$\theta(x) = x + O(x^{1/2}(\log x)^2),$$

$$\pi(x) = li(x) + O(x^{1/2}\log x).$$

Proof. Taking $T = x$ in Lemma 15.4, we obtain

$$\psi_0(x) = x - \sum_{\substack{\rho \\ |\gamma| \leq x}} \frac{x^{\frac{1}{2}+i\gamma}}{\frac{1}{2}+i\gamma} + E(x),$$

where

$$E(x) = -\log(2\pi) - \tfrac{1}{2}\log\left(1 - \tfrac{1}{x^2}\right) + O\left((\log x)\min\left\{1, \tfrac{1}{\langle x \rangle}\right\} + (\log x)^2\right)$$

$$\ll (\log x)^2.$$

But

$$\sum_{\substack{\rho \\ |\gamma| \leq x}} \frac{x^{1/2+i\gamma}}{1/2+i\gamma} \ll x^{1/2} \sum_{\substack{\rho \\ |\gamma| \leq x}} \frac{1}{|\rho|} \ll x^{\frac{1}{2}} \sum_{1 \leq n \leq x} \frac{\log(n+2)}{n} \ll x^{\frac{1}{2}}(\log x)^2.$$

Thus

$$\psi_0(x) = x + O(x^{\frac{1}{2}}(\log x)^2). \quad \square$$

The remaining two assertions follow as exercises in elementary analytic number theory (see Problem Sheet 6). //

If one does not have available RH, but instead the weaker

information that for some real number Θ with $\frac{1}{2} < \Theta < 1$ one has
$$\Theta = \sup\{\operatorname{Re}(\rho) : \zeta(\rho) = 0\},$$
then the above argument readily yields the formula
$$\psi_0(x) = x + O(x^{\Theta}(\log x)^2)$$
(see Problem Sheet 6). Suppose in fact that for any $\varepsilon > 0$ one has
$$\psi(x) = x + O(x^{\theta+\varepsilon}),$$
for some real number $\theta > 0$. Then one sees that
$$\sum_{n \leqslant x} (\Lambda(n) - 1) = \psi(x) - x \ll x^{\theta+\varepsilon}.$$

It follows that the associated Dirichlet series
$$\sum_{n=1}^{\infty} \frac{\Lambda(n)-1}{n^s} \qquad \left(= -\frac{\zeta'(s)}{\zeta(s)} - \zeta(s) \right)$$
converges in the half plane $\sigma > \theta$ (from ANT 1), and hence one must have $\zeta(s) \neq 0$ for $\sigma > \theta$ [otherwise one would have a pole of $\frac{\zeta'(s)}{\zeta}$ in this halfplane #]. Thus one cannot have
$$\psi(x) = x + O_{\varepsilon}(x^{\theta+\varepsilon}) \quad \text{for any } \theta < \Theta. \text{ Equivalently,}$$
$$\psi(x) - x = \Omega(x^{\Theta-\varepsilon}).$$
This line of reasoning may be applied to obtain superficial refinements to asymptotic formulae for prime counting functions.

<u>Theorem 16.2.</u> Suppose that $\theta \geqslant \frac{1}{2}$ is a real number satisfying the property that, whenever $\varepsilon > 0$ one has
$$\psi(x) = x + O(x^{\theta+\varepsilon}).$$
Then one has
$$\psi(x) = x + O(x^{\theta}(\log x)^2).$$

Proof. If
$$\Theta := \sup\{\operatorname{Re}(\rho) : \zeta(\rho)=0\},$$
and $\Theta \leqslant \theta$, then the argument of Theorem 16.1 applies, and we have
$$\psi_0(x) = x + O(x^{\Theta}(\log x)^2),$$
and we are done. Hence we may suppose ~~that~~ $\Theta > \theta$, say $\Theta = \theta + \delta$ with $\delta > 0$. But ~~then~~
$$\psi(x) - x = \Omega(x^{\Theta - \varepsilon}) = \Omega(x^{\theta + \delta/2}),$$
say, and this contradicts the relation $\psi(x) = x + O(x^{\theta+\varepsilon})$. Then the former assertion does indeed hold, and we are done. //

We will return to Ω- estimates for the error term in $\psi(x)$ in due course. Our final conclusion for the present conditional on RH shows that $\psi(x) - x$ is usually $O(x^{1/2})$ (note: no powers of $\log x$!)

Theorem 16.3. Suppose that RH holds. Then for $X \geqslant 2$, one has
$$\int_X^{2X} (\psi(x) - x)^2 \, dx \ll X^2.$$

Proof. By Lemma 15.4 with $T = X$, one has for $x \in [X, 2X]$ the relation
$$\psi(x) = x - \sum_{|\gamma| \leqslant X} \frac{x^\rho}{\rho} + R(x),$$
where
$$R(x) = -\log 2\pi - \tfrac{1}{2}\log\left(1 - \tfrac{1}{x^2}\right) + O(\log^2(2X))$$
$$\ll \log^2(2X).$$
Thus
$$\int_X^{2X} R(x)^2 \, dx \ll X \log^4(2X).$$

The main action concerns the sum over zeros. We have

$$\int_X^{2X} \Big| \sum_{|\gamma|\le X} \frac{x^\rho}{\rho} \Big|^2 dx = \sum_{\substack{\gamma_1,\gamma_2 \\ |\gamma_i|\le X}} \frac{1}{\rho_1 \bar{\rho_2}} \int_X^{2X} x^{1+i(\gamma_1-\gamma_2)} dx$$

$$(\rho_i = \tfrac{1}{2} + i\gamma_i)$$

$$\ll X^2 \sum_{\substack{\gamma_1,\gamma_2 \\ |\gamma_i|\le X}} \frac{1}{|\rho_1\rho_2|\,|2+i(\gamma_1-\gamma_2)|} .$$

We now seek to show that this last sum is finite, from which it follows that

$$\int_X^{2X} (\psi(x)-x)^2\, dx \ll \int_X^{2X} R(x)^2\, dx + \int_X^{2X} \Big| \sum_{|\gamma|\le X} \frac{x^\rho}{\rho} \Big|^2 dx$$

$$\ll X \log^4(2X) + X^2 \ll X^2 . \square$$

The sum over γ_1, γ_2 may be dealt with by considering the contribution from γ_2 in different ranges. By symmetry, we may suppose that $\gamma_1 > 0$. We have

$$\sum_{\gamma_2 < -\frac{\gamma_1}{2}} \frac{1}{|\rho_1\rho_2|\;|2+i(\gamma_1-\gamma_2)|} \ll \frac{1}{|\gamma_1|} \sum_{\gamma_2 < -\frac{\gamma_1}{2}} \frac{1}{|\gamma_2|(1+|\gamma_2-\gamma_1|)}$$

$$\ll \frac{1}{|\gamma_1|} \sum_{\gamma_2 < -\frac{\gamma_1}{2}} \frac{1}{|\gamma_2|^2}$$

$$\ll \frac{1}{|\gamma_1|} \sum_{n > \frac{\gamma_1}{2}} \frac{\log n}{n^2} \ll \frac{\log \gamma_1}{\gamma_1^2} ,$$

$\uparrow$ consider $n < |\gamma_2| \le n+1$

Whence

$$\sum_{\gamma_1>0} \sum_{\gamma_2 < -\gamma_1} \frac{1}{|\rho_1\rho_2|\,|2+i(\gamma_1-\gamma_2)|} \ll \sum_{\gamma_1>0} \frac{\log \gamma_1}{\gamma_1^2}$$

$$\ll \sum_n \frac{(\log n)^2}{n^2} < \infty .$$

Similarly,

$$\sum_{|\gamma_2| \le \frac{1}{2}\gamma_1} \frac{1}{|\rho_1\rho_2|\,|2+i(\gamma_1-\gamma_2)|} \ll \frac{1}{|\gamma_1|} \sum_{0<\gamma_2\le\gamma_1} \frac{1}{|\gamma_2|\cdot|\gamma_1|}$$

$$\ll \frac{1}{|\gamma_1|^2} \sum_{1 \le n \le \gamma_1} \frac{\log n}{n} \ll \frac{(\log \gamma_1)^2}{|\gamma_1|^2}.$$

Thus

$$\sum_{\gamma_1 > 0} \sum_{|\gamma_2| \le \frac{1}{2}\gamma_1} \frac{1}{|\rho_1 \rho_2| \cdot |2 + i(\gamma_1 - \gamma_2)|} \ll \sum_{\gamma_1 > 0} \frac{(\log \gamma_1)^2}{\gamma_1^2}$$

$$\ll \sum_n \frac{(\log n)^3}{n^2} < \infty.$$

Next, we have

$$\sum_{\frac{1}{2}\gamma_1 < \gamma_2 < \frac{3}{2}\gamma_1} \frac{1}{|\rho_1 \rho_2|\,|2 + i(\gamma_1 - \gamma_2)|} \ll \frac{1}{\gamma_1^2} \sum_{\frac{1}{2}\gamma_1 < \gamma_2 \le \frac{3}{2}\gamma_1} \frac{1}{1 + |\gamma_1 - \gamma_2|}$$

$$\ll \frac{1}{\gamma_1^2} \sum_{1 \le n \le \gamma_1} \frac{\log(2n)}{n} \ll \frac{(\log \gamma_1)^2}{\gamma_1^2},$$

whence

$$\sum_{\gamma_1 > 0} \sum_{\frac{1}{2}\gamma_1 < \gamma_2 < \frac{3}{2}\gamma_1} \frac{1}{|\rho_1 \rho_2| \cdot |2 + i(\gamma_1 - \gamma_2)|} \ll \sum_{\gamma_1 > 0} \frac{(\log \gamma_1)^2}{\gamma_1^2} < \infty.$$

Finally,

$$\sum_{\gamma_2 \ge \frac{3}{2}\gamma_1} \frac{1}{|\rho_1 \rho_2| \cdot |2 + i(\gamma_1 - \gamma_2)|} \ll \frac{1}{\gamma_1} \sum_{\gamma_2 \ge \frac{3}{2}\gamma_1} \frac{1}{\gamma_2^2} \ll \frac{1}{\gamma_1} \sum_{n > \gamma_1} \frac{\log n}{n^2} \ll \frac{\log \gamma_1}{\gamma_1^2},$$

whence

$$\sum_{\gamma_1 > 0} \sum_{\gamma_2 \ge \frac{3}{2}\gamma_1} \frac{1}{|\rho_1 \rho_2| \cdot |2 + i(\gamma_1 - \gamma_2)|} \ll \sum_{\gamma_1 > 0} \frac{\log \gamma_1}{\gamma_1^2} < \infty.$$

We may now conclude, as described above, that

$$\int_X^{2X} (\psi(x) - x)^2 \, dx \ll X^2.$$

//

---

§17. Oscillation of error terms.

Thus far we have connected error terms in asymptotic formulae for $\psi(x)$, $\theta(x)$ and $\pi(x)$ with zeros of $\zeta(s)$, in particular proving Ω-results showing that $|\psi(x) - x|$ may be large.

We now consider $\Omega_\pm$ results. If $\psi(x) - x = \Omega_\pm(x^\theta)$, then there are arbitrarily large values of x for which

$$\psi(x) - x > C_+ x^\theta,$$

for a suitable constant $C_+ > 0$, and likewise arbitrarily large values of x for which

$$\psi(x) - x < -C_- x^\theta,$$

for a suitable constant $C_- > 0$. Thus we have oscillation of the error term. In preparation we consider again some properties of Dirichlet series.

Recall:

Theorem (Landau) Let $\alpha(s) = \sum_{n=1}^\infty a(n) n^{-s}$ be a Dirichlet series whose abscissa of convergence σ_c is finite. If $a_n \geq 0$ for all n, then the point σ_c is a singularity of $\alpha(s)$.

Analogously ...

Lemma 17.1. Suppose that $A(x)$ is a bounded Riemann integrable function in any finite interval $1 \leq x \leq X$, and further $A(x) \geq 0$ for all $x > X_0$. Let σ_c denote the infimum of the set of $\sigma \in \mathbb{R}$ for which

$$\int_{X_0}^\infty A(x) x^{-\sigma} \, dx < \infty.$$

Then the function

$$F(s) = \int_1^\infty A(x) x^{-s} \, dx$$

is analytic in the half-plane $\sigma > \sigma_c$, but not at the point $s = \sigma_c$.

Proof. Putting

$$F_1(s) = \int_1^{X_0} A(x) x^{-s} \, dx \quad \text{and} \quad F_2(s) = \int_{X_0}^\infty A(x) x^{-s} \, dx,$$

we see that

$$F(s) = F_1(s) + F_2(s).$$

The function $F_1(s)$ is entire. When $\sigma > \sigma_c$, we have

$$\int_{X_0}^\infty |A(x) x^{-s}| \, dx \leq \int_{X_0}^\infty A(x) x^{-\sigma} \, dx < \infty,$$

and hence $F_2(s)$ is analytic for $\sigma > \sigma_c$ (and hence $F(s) = F_1(s) + F_2(s)$ is also analytic for $\sigma > \sigma_c$).

We now consider the behaviour of $F_2(s)$ at $s = \sigma_c$. By replacing $A(x)$ by $A(x)x^{-\sigma_c}$, we may suppose that $\sigma_c = 0$. Suppose that $F_2(s)$ is analytic at $s = 0$, so $F_2(s)$ is analytic in the domain $\mathcal{D} = \{s : \sigma > 0\} \cup \{|s| < \delta\}$, for sufficiently small $\delta > 0$. We expand $F_2(s)$ as a power series at $s = 1$:

$$F_2(s) = \sum_{k=0}^{\infty} c_k (s-1)^k, \qquad (17.1)$$

where

$$c_k = \frac{F_2^{(k)}(1)}{k!} = \frac{1}{k!}\int_{x_0}^{\infty} A(x)(-\log x)^k x^{-1}\,dx.$$

The radius of convergence of the power series (17.1) is the distance from 1 to the nearest singularity of $F_2(s)$. Since $F_2(s)$ is analytic in $\mathcal{D}$, and since the nearest points not in $\mathcal{D}$ are $\pm i\delta$, we see that the radius of convergence is at least $\sqrt{1+\delta^2} = 1 + \delta'$, say.

Thus

$$F_2(s) = \sum_{k=0}^{\infty} \frac{(1-s)^k}{k!} \int_{x_0}^{\infty} A(x)(\log x)^k x^{-1}\,dx$$

for $|s-1| < 1+\delta'$. If $s < 1$, then ~~the function $A(x)(\log x)^k$~~ the terms in this infinite sum are all non-negative, and hence can be rearranged. Thus, for $-\delta' < s < 1$ one has

$$F_2(s) = \int_{x_0}^{\infty} A(x)x^{-1} \sum_{k=0}^{\infty} \frac{(1-s)^k (\log x)^k}{k!}\,dx$$

$$= \int_{x_0}^{\infty} A(x)x^{-1}\exp\left((1-s)\log x\right)dx$$

$$= \int_{x_0}^{\infty} A(x)x^{-s}\,dx.$$

Hence the function $F_2(s)$ converges at $s = -\delta'/2$, contradicting the assumption that $\sigma_c = 0$. Then $F(s)$ is not analytic at $s = \sigma_c$. //

Theorem 17.2. Let

$$\Theta = \sup \{ \mathrm{Re}(\rho) : \zeta(\rho) = 0 \}.$$

Then for every $\varepsilon > 0$, one has

$$\psi(x) - x = \Omega_{\pm} (x^{\Theta - \varepsilon}).$$

Proof. We begin by noting ~~that~~

$$-\frac{\zeta'}{\zeta}(s) = s \int_1^\infty \psi(x)\, x^{-s-1}\, dx \qquad (\sigma > 1).$$

Hence

$$\int_1^\infty (\psi(x) - x)\, x^{-s-1}\, dx = -\frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1} \qquad (\sigma > 1). \quad (17.2)$$

Suppose, if possible, ~~that~~ $\psi(x) - x < \frac{1}{2} x^{\Theta - \varepsilon}$ for all $x > X_0(\varepsilon)$.

We have

$$\int_1^\infty \big(\underbrace{x^{\Theta - \varepsilon} - \psi(x) + x}_{>0 \text{ for } x > X_0(\varepsilon).} \big)\, x^{-s-1}\, dx = \frac{1}{s - \Theta + \varepsilon} + \frac{\zeta'(s)}{s\zeta(s)} + \frac{1}{s-1}. \quad (17.3)$$

In view of (17.2), one sees ~~that~~ the assumption $\psi(x) - x < \frac{1}{2} x^{\Theta - \varepsilon}$ implies ~~that~~

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

is analytic for $\mathrm{Re}(s) > \Theta - \varepsilon$. Hence the right hand side of (17.3) is analytic for $\mathrm{Re}(s) > \Theta - \varepsilon$, with a simple pole at $s = \Theta - \varepsilon$. Then the identity (17.3) holds for $\sigma > \Theta - \varepsilon$, and both lhs and rhs are analytic in $\sigma > \Theta - \varepsilon$. But $\frac{\zeta'}{\zeta}(s)$ has poles with real part exceeding $\Theta - \varepsilon$, and so we derive a contradiction. Then we have $\psi(x) - x = \Omega_{+} (x^{\Theta - \varepsilon})$.

We may argue similarly to obtain a ~~contradiction to the~~ assumption ~~that~~ $\psi(x) - x > - \frac{1}{2} x^{\Theta - \varepsilon}$. In ~~this case~~ we see that

$$\int_1^\infty (x^{\Theta - \varepsilon} + \psi(x) - x)\, x^{-s-1} dx = \frac{1}{s - \Theta + \varepsilon} - \frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1},$$

and we again deduce that lhs and rhs are analytic in $\sigma > \Theta - \varepsilon$. #

Thus $\psi(x) - x = \Omega_-(x^{\Theta-\varepsilon})$. //

One may also establish that

$$\pi(x) - li(x) = \Omega_{\pm}(x^{\Theta-\varepsilon}),$$

though this requires a consideration of the Mellin transforms of $li(x)$ and

$$\Pi(x) := \sum_{n \le x} \frac{\Lambda(n)}{\log n},$$

namely

$$s\int_2^\infty li(x)\, x^{-s-1}\,dx = -\log(s-1) + \underset{\text{entire}}{r(s)}$$

and

$$s\int_2^\infty \Pi(x)\, x^{-s-1}\,dx = \log \zeta(s). \qquad (\sigma > 1).$$

Theorem 17.3. Suppose that Θ is $\sup\{Re(\rho) : \zeta(\rho) = 0\}$, and that there exists a ~~zero~~ ρ of $\zeta(s)$ with $Re(\rho) = \Theta$, say $\rho = \Theta + i\gamma$. Then one has

$$\limsup_{x\to\infty} \frac{\psi(x) - x}{x^\Theta} \ge \frac{1}{|\rho|},$$

and

$$\liminf_{x\to\infty} \frac{\psi(x) - x}{x^\Theta} \le -\frac{1}{|\rho|}.$$

Proof. Suppose that $\psi(x) \le x + cx^\Theta$ for all $x \ge X_0$.

Then, by Lemma 17.1, one has

$$F(s) := \int_1^\infty (cx^\Theta - \psi(x) + x)\,x^{-s-1}\,dx = \frac{c}{s-\Theta} + \frac{\zeta'(s)}{s\zeta(s)} + \frac{1}{s-1},$$

for $\sigma > \Theta$. We now engineer some positivity in the argument of the integral. We have

$$F(s) + \tfrac{1}{2}e^{i\varphi}F(s+i\gamma) + \tfrac{1}{2}e^{-i\varphi}F(s-i\gamma)$$
$$= \int_1^\infty (cx^\Theta - \psi(x) + x)(1 + \cos(\varphi - \gamma\log x))\,x^{-s-1}\,dx, \qquad (17.4)$$

for $\sigma > \Theta$. As $s = \sigma \to \Theta +$, we see that the integral on the right hand side from 1 to X_0 is uniformly bounded, while from X_0 to ∞ it is non-negative. Thus

$$\liminf_{\sigma \to \Theta +} \left(F(\sigma) + \tfrac{1}{2} e^{i\varphi} F(\sigma + i\gamma) + \tfrac{1}{2} e^{-i\varphi} F(\sigma - i\gamma) \right) > -\infty . \qquad (17.5)$$

Meanwhile, since

$$F(s) = \frac{c}{s - \Theta} + \frac{\zeta'(s)}{s\zeta(s)} + \frac{1}{s-1},$$

we see that $F(s) + \frac{1}{2} e^{i\varphi} F(s + i\gamma) + \frac{1}{2} e^{-i\varphi} F(s - i\gamma)$ has a pole at $s = \Theta$ with residue

$$c + \frac{m e^{i\varphi}}{2\rho} + \frac{m e^{-i\varphi}}{2\bar{\rho}},$$

where $m \geq 1$ is the multiplicity of ρ.

Choose φ in such a manner that $\dfrac{e^{i\varphi}}{\rho} = -\dfrac{1}{|\rho|}$. Then the above residue is equal to

$$c - \frac{m}{|\rho|} .$$

This quantity must be non-negative, for if it were negative then the lhs of (17.4) would $\to -\infty$ as $s = \sigma \to \Theta^+$, a possibility we ruled out in (17.5). Thus $c \geq \dfrac{1}{|\rho|}$, whence

$$\limsup_{x \to \infty} \frac{\psi(x) - x}{x^{\Theta}} \geq \frac{1}{|\rho|} .$$

The argument for the proof of

$$\liminf_{x \to \infty} \frac{\psi(x) - x}{x^{\Theta}} \leq -\frac{1}{|\rho|}$$

is similar, and we leave this as an easy exercise. //

<u>Corollary 17.4.</u> One has
$$\psi(x) - x = \Omega_\pm(x^{1/2}), \qquad \theta(x) - x = \Omega_-(x^{1/2}),$$
and
$$\pi(x) - \mathrm{li}(x) = \Omega_-(x^{1/2}/\log x).$$

<u>Proof.</u> We can divide into two cases according to whether the Riemann Hypothesis is true or false. If RH is false, ~~then~~ there is a zero $\rho = \beta + i\gamma$ with $\beta > \frac{1}{2}$. In such circumstances, one has
$$\Theta = \sup\{\mathrm{Re}(\rho) : \zeta(\rho) = 0\} > \tfrac{1}{2},$$
and it follows from Theorem 17.2 ~~that~~
$$\psi(x) - x = \Omega_\pm(x^{\beta - \varepsilon}) = \Omega_\pm(x^{1/2}).$$
If RH is true, meanwhile, then one has (from Theorem 17.3)
$$\limsup_{x\to\infty} \frac{\psi(x) - x}{x^{\frac{1}{2}}} \geq \frac{1}{|\rho|} \Rightarrow \psi(x) - x = \Omega_+(x^{\frac{1}{2}})$$
$$\liminf_{x\to\infty} \frac{\psi(x) - x}{x^{1/2}} \leq -\frac{1}{|\rho|} \Rightarrow \psi(x) - x = \Omega_-(x^{1/2}).$$
So $\psi(x) - x = \Omega_\pm(x^{1/2})$. $\square$

It follows from Problem sheet 6 ~~that~~
$$\theta(x) = \psi(x) - x^{1/2} + O(x^{1/3}), \quad \text{whence} \quad \theta(x) - x = \Omega_-(x^{1/2}),$$
and
$$\pi(x) - \mathrm{li}(x) = \frac{\theta(x) - x}{\log x} + O\left(\frac{x^{1/2}}{(\log x)^2}\right), \quad \text{whence} \quad \pi(x) - \mathrm{li}(x) = \Omega_-\left(\frac{x^{1/2}}{\log x}\right).$$

//

In order to go beyond the $\Omega_\pm$ - results presented in Corollary 17.4, one must start to unravel the Diophantine properties of the zeros of $\zeta(s)$. This was achieved by Littlewood in 1914. We now present a version of this argument, starting with a couple of preliminary lemmata.

<u>Lemma 17.5.</u> Suppose ~~that~~ the Riemann Hypothesis is true. Then whenever $x \geq 4$ and $\frac{1}{2x} \leq \delta \leq \frac{1}{2}$, one has

$$\frac{1}{(e^{\delta}-e^{-\delta})x}\int_{e^{-\delta}x}^{e^{\delta}x}(\psi(u)-u)\,du = -2x^{\frac12}\sum_{\gamma>0}\left(\frac{\sin(\gamma\delta)}{\gamma\delta}\right)\frac{\sin(\gamma\log x)}{\gamma}+O(x^{\frac12}).$$

Notice that as $\delta\to 0$, this expression is essentially a small average of $\psi(u)-u$ with u close to x, and is approximately

$$-2x^{\frac12}\sum_{\gamma>0}\frac{\sin(\gamma\log x)}{\gamma}.$$

We may now seek to choose x in such a manner ~~that the~~ expressions $\sin(\gamma\log x)$ conspire with each other not to cancel, and ~~thereby~~ aim to show that this expression is ~~rather~~ larger in magnitude than $x^{\frac12}$.

Proof of Lemma 17.5. We have

$$\psi(u) = u - \sum_{\substack{\rho\\ |\gamma|\le T}}\frac{u^\rho}{\rho} - \log(2\pi) - \tfrac12\log\left(1-\frac{1}{u^2}\right) + R(u,T),$$

where

$$R(u,T) \ll (\log u)\min\left\{1,\frac{u}{T\langle u\rangle}\right\} + \frac{u}{T}\log^2(uT).$$

Thus

$$\int_0^x(\psi(u)-u)\,du = -\sum_{\substack{\rho\\ |\gamma|\le T}}\int_0^x\frac{u^\rho}{\rho}\,du - x\log(2\pi) + O\left(\frac{x^2\log^2 x}{T^{1/2}}\right) + O(1),$$

for large enough values of T. Hence

$$\int_0^x(\psi(u)-u)\,du = -\sum_{\substack{\rho\\ |\gamma|\le T}}\frac{x^{\rho+1}}{\rho(\rho+1)} - x\log(2\pi) + O\left(\frac{x^2\log^2x}{T^{1/2}}\right) + O(1)$$

The series

$$\sum_\rho\frac{1}{\rho(\rho+1)}$$

is absolutely convergent and moreover

$$\sum_{|\gamma|>T}\frac{1}{|\rho(\rho+1)|} \le \sum_{n>T}\frac{\log n}{n^2} \ll \frac{\log T}{T}.$$

Thus, on taking the limit as $T\to\infty$, we deduce ~~that~~

$$\int_0^x(\psi(u)-u)\,du = -\sum_\rho\frac{x^{\rho+1}}{\rho(\rho+1)} - x\log(2\pi) + O(1).$$

Thus

$$\int_{e^{-\delta}x}^{e^{\delta}x} (\psi(u)-u)\,du = -\sum_{\rho} \frac{(e^{\delta}x)^{\rho+1}-(e^{-\delta}x)^{\rho+1}}{\rho(\rho+1)} - (e^{\delta}-e^{-\delta})x\log(2\pi) + O(1)$$

$$\Rightarrow \frac{1}{(e^{\delta}-e^{-\delta})x}\int_{e^{-\delta}x}^{e^{\delta}x} (\psi(u)-u)\,du = -\frac{\delta}{\sinh\delta}\sum_{\rho}\frac{\left(e^{\delta(\rho+1)}-e^{-\delta(\rho+1)}\right)x^{\rho}}{2\delta\,\rho(\rho+1)} + O(1).$$

Observe ~~that~~, in view of our assumption of RH, we have

$$e^{\pm\delta(\rho+1)} = e^{\pm i\gamma\delta}\,(1+O(\delta)) = e^{\pm i\gamma\delta} + O(\delta),$$

whence

$$\frac{1}{(e^{\delta}-e^{-\delta})x}\int_{e^{-\delta}x}^{e^{\delta}x} (\psi(u)-u)\,du = \frac{-\delta}{\sinh\delta}\sum_{\rho}\frac{(e^{i\gamma\delta}-e^{-i\gamma\delta})x^{\rho}}{2\delta\,\rho(\rho+1)}$$

$$+\,O\Bigg(\frac{\delta}{|\sinh\delta|}\,x^{\frac12}\underbrace{\sum_{\rho}\frac{1}{|\gamma|^2}}_{\ll \sum_n \frac{\log n}{n^2}<\infty}\Bigg) + O(1)$$

$$= -i\,x^{\frac12}\left(\frac{\delta}{\sinh\delta}\right)\sum_{\rho}\frac{\sin(\gamma\delta)}{\delta}\,\frac{x^{i\gamma}}{\rho(\rho+1)} + O(x^{\frac12}).$$

But $\dfrac{\delta}{\sinh\delta} = 1+O(\delta^2)$, and

$$\sum_{\rho}\frac{\sin(\gamma\delta)}{\delta}\,\frac{x^{i\gamma}}{\rho(\rho+1)} \ll \frac{1}{\delta}\sum_{\gamma>1/\delta}\frac{1}{\gamma^2} + \sum_{0<\gamma\le 1/\delta}\frac{\gamma\delta}{\delta}\cdot\frac{1}{\gamma^2}$$

$$\ll \frac{1}{\delta}\,\frac{\log(1/\delta)}{(1/\delta)} + \log^2(1/\delta) \ll \log^2(1/\delta).$$

Thus

$$\frac{1}{(e^{\delta}-e^{-\delta})x}\int_{e^{-\delta}x}^{e^{\delta}x} (\psi(u)-u)\,du = -i\,x^{\frac12}\sum_{\rho}\frac{\sin(\gamma\delta)}{\delta}\,\frac{x^{i\gamma}}{\rho(\rho+1)} + O(x^{\frac12}) + \underbrace{O\left(\delta^2\log^2(1/\delta)\right)}_{O(x^{1/2})}.$$

Next, since

$$\frac{1}{\rho(\rho+1)} = \frac{1}{(\frac{1}{2}+i\gamma)(\frac{3}{2}+i\gamma)} = \frac{1}{-\gamma^2+2i\gamma+\frac{3}{4}},$$

we see ~~that~~

$$\frac{1}{\rho(\rho+1)} + \frac{1}{\gamma^2} = \frac{2i\gamma+3/4}{\gamma^2(-\gamma^2+2i\gamma+3/4)} = O\left(\frac{1}{\gamma^3}\right),$$

and likewise

$$\frac{1}{\bar\rho(\bar\rho+1)} + \frac{1}{\gamma^2} = O\left(\frac{1}{|\gamma|^3}\right).$$

Thus

$$\sum_\rho \frac{\sin(\gamma\delta)}{\delta}\,\frac{x^{i\gamma}}{\rho(\rho+1)} = -\sum_\rho \frac{\sin(\gamma\delta)}{\gamma\delta}\,\frac{x^{i\gamma}}{\gamma} + O\Big(\underbrace{\sum_\rho \frac{\gamma\delta}{\delta}\cdot\frac{1}{\gamma^3}}_{\ll 1}\Big).$$

Finally,

$$-ix^{\frac{1}{2}}\sum_\rho \frac{\sin(\gamma\delta)}{\delta}\,\frac{x^{i\gamma}}{\rho(\rho+1)} = +ix^{\frac{1}{2}}\sum_{\gamma>0}\frac{\sin(\gamma\delta)}{\gamma\delta}\left(\frac{x^{i\gamma}-x^{-i\gamma}}{\gamma}\right) + O(x^{1/2})$$

$$= -2x^{\frac{1}{2}}\sum_{\gamma>0}\frac{\sin(\gamma\delta)}{\gamma\delta}\,\frac{\sin(\gamma\log x)}{\gamma} + O(x^{1/2}).$$

We have ~~therefore~~ shown ~~that~~

$$\frac{1}{(e^\delta-e^{-\delta})x}\int_{e^{-\delta}x}^{e^\delta x}(\psi(u)-u)\,du = -2x^{\frac{1}{2}}\sum_{\gamma>0}\left(\frac{\sin(\gamma\delta)}{\gamma\delta}\right)\frac{\sin(\gamma\log x)}{\gamma} + O(x^{1/2}). \quad //$$

We next recall a multidimensional generalisation of Dirichlet's theorem on Diophantine approximation.

Lemma 17·6. Suppose ~~that~~ $\theta_1,\dots,\theta_R\in\mathbb{R}$ and $Q\in\mathbb{N}$. Then there exists a positive integer q with $1\le q\le Q^R$ for which $\|q\theta_r\| < 1/Q$ $(1\le r\le R)$.

Proof. Let n be an integer with $0 \le n \le Q^R$. The R-tuple $(\{n\theta_1\}, \{n\theta_2\}, \ldots, \{n\theta_R\})$ lies in the hypercube $[0,1)^R$. We partition ~~this~~ unit hypercube up ~~into~~ Q^R sub-hypercubes each of side length $1/Q$ in the obvious manner. Thus, there are Q^R+1 R-tuples in Q^R boxes, and ~~in~~ some such sub-hypercube there must be 2 or more of ~~these~~ R-tuples, corresponding say to n_1 and $n_2 > n_1$. We then have

$$\| n_2\theta_r - n_1\theta_r \| \le | \{n_2\theta_r\} - \{n_1\theta_r\} | < 1/Q \quad (1 \le r \le R),$$

whence the conclusion of the lemma holds with $q = n_2 - n_1$, where $1 \le q \le Q$. //

We require one further result before we can prove Littlewood's oscillation theorem, though this concerns the density of the zeros of $\zeta(s)$. We will defer its proof to later.

Theorem 17.7. Independent of ~~Assume~~ the truth of the Riemann Hypothesis, one has

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

On RH, the error term is $O(\log T/\log\log T)$.

Corollary. Suppose ~~that~~ $H(T)$ is a function of T monotonically increasing with T and having the property ~~that~~ $H(T) \to \infty$ as $T \to \infty$. Then for all large T,

$$N\left(T + \frac{H(T)}{\log\log T}\right) - N(T) \gg H(T)\frac{\log T}{\log\log T}.$$

<u>Theorem 17.8</u> (Littlewood). One has
$$\psi(x) - x = \Omega_{\pm}(x^{1/2}\log\log\log x),$$
and
$$\pi(x) - \mathrm{li}(x) = \Omega_{\pm}\left(x^{1/2}\frac{\log\log\log x}{\log x}\right).$$

<u>Proof.</u> We first consider $\psi(x)-x$, noting ~~that~~ if RH is false, then $\Theta = \sup_{\rho}\{\mathrm{Re}(\rho)\} > \frac{1}{2}$, and it follows from Theorem 17.2 ~~that~~ $\psi(x) - x = \Omega_{\pm}(x^{\Theta-\varepsilon})$, which is already ~~stronger~~ than ~~the~~ assertion ~~that~~ $\psi(x)-x = \Omega_{\pm}(x^{\frac{1}{2}}\log\log\log x)$. We may ~~therefore~~ suppose ~~that~~ RH holds. We will apply ~~the~~ conclusion of Lemma 17.5 for carefully chosen x:
$$\frac{1}{(e^{\delta}-e^{-\delta})x}\int_{e^{-\delta}x}^{e^{\delta}x}(\psi(u)-u)\,du = -2x^{\frac{1}{2}}\sum_{\gamma>0}\left(\frac{\sin(\gamma\delta)}{\gamma\delta}\right)\frac{\sin(\gamma\log x)}{\gamma} + O(x^{1/2}).$$

Consider the real numbers $\frac{\gamma\log N}{2\pi}$, with $0 \le \gamma \le T$, where $T := N\log N$ and N is a large integer. The number of zeros in question here is
$$N(T) \asymp T\log T \asymp N(\log N)^2.$$
Put $R = N(T)$, and apply Lemma 17.5 to the R real numbers $\frac{\gamma\log N}{2\pi}$. We see ~~that~~ ~~there exists~~ an integer n with $1 \le n \le N^R$ for which
$$\left\|\frac{\gamma n}{2\pi}\log N\right\| < \frac{1}{N} \qquad (0<\gamma\le T). \quad \text{— (17.6)}$$

Hence $\frac{\gamma n \log N}{2\pi}$ is close to an integer for all of those γ.

Put $x = N^n e^{\pm 1/N}$ and $\delta = 1/N$ in Lemma 17.5.

We have

$$\log x = n \log N \pm \frac{1}{N},$$

whence

$$\sin(\gamma \log x) = \sin\left(\gamma n \log N \pm \frac{\gamma}{N}\right).$$

But

$$|\sin(2\pi\alpha) - \sin(2\pi\beta)| \leq 2\pi \|\alpha - \beta\|,$$

so that from (17.6) we see that

$$|\sin(\gamma \log x) \mp \sin(\gamma/N)| \leq \frac{2\pi}{N}.$$

Thus

$$\frac{1}{(e^{1/N} - e^{-1/N})x} \int_{e^{-1/N}x}^{e^{1/N}x} (\psi(u) - u)\, du = \mp 2x^{\frac{1}{2}} \sum_{0<\gamma\leq T} \frac{\sin(\gamma/N)}{\gamma/N} \cdot \frac{\sin(\gamma/N)}{\gamma}$$

$$+ E,$$

where

$$E \ll x^{1/2} + x^{\frac{1}{2}} \sum_{\gamma > T} \frac{1}{\gamma/N} \cdot \frac{1}{\gamma} + x^{\frac{1}{2}} \sum_{0<\gamma\leq T} \left| \frac{\sin(\gamma/N)}{\gamma/N} \cdot \frac{1/N}{\gamma} \right|.$$

The last error term here is

$$\ll x^{\frac{1}{2}} \sum_{\gamma\leq T} \frac{1}{\gamma^2} \ll x^{\frac{1}{2}},$$

whilst the penultimate one is

$$\ll x^{\frac{1}{2}} N \sum_{\gamma > N\log N} \frac{1}{\gamma^2} \ll x^{\frac{1}{2}} N \sum_{n > N\log N} \frac{\log n}{n^2} \ll x^{\frac{1}{2}} N \frac{\log N}{N\log N}.$$

Thus $E \ll x^{\frac{1}{2}}$. It remains to estimate

$$\sum_{0<\gamma\le T} \left(\frac{\sin(\gamma/N)}{\gamma/N} \right)^2 .$$

When $\gamma \le \frac{\pi}{2} N$, we have $\dfrac{\sin(\gamma/N)}{\gamma/N} \geq \dfrac{2}{\pi}$, and thus

$$\sum_{0<\gamma\le T} \left(\frac{\sin(\gamma/N)}{\gamma/N} \right)^2 \gg \sum_{0<\gamma\le\frac{\pi}{2}N} 1 \gg N \log N :$$

We have ~~therefore~~ shown ~~that~~

$$\frac{1}{(e^{1/N} - e^{-1/N})x} \int_{e^{-1/N}x}^{e^{1/N}x} (\psi(u) - u)\, du = \mp M + O(x^{1/2}),$$

where

$$M \gg x^{\frac{1}{2}} \frac{N \log N}{N} = x^{\frac{1}{2}} \log N .$$

We have yet to take care of the book-keeping. We have

$$x \lesssim N^{N^R} e^{1/N} \quad \text{with} \quad R = N(T) \asymp N(\log N)^2,$$

so that

$$\log\log x \ll \log(N^R \log N) \ll N(\log N)^3 .$$

Consequently, we have

$$\log N \geq (1+o(1)) \log\log\log x ,$$

and hence

$$\frac{1}{(e^{1/N} - e^{-1/N})x} \int_{e^{-1/N}x}^{e^{1/N}x} (\psi(u) - u)\, du = \Omega_\pm \left(x^{\frac{1}{2}} \log\log\log x \right).$$

The expression on the left hand side is the average of $\psi(u) - u$ over a neighbourhood of x, so we see that $\psi(y) - y = \Omega_\pm(y^{\frac{1}{2}} \log\log\log y)$

for some $y \in [e^{-1/N} x, e^{1/N} x]$.

The assertion concerning $\pi(x) - li(x)$ follows from Problem Set 6: $\pi(x) - li(x) = \frac{\theta(x) - x}{\log x} + O\left(\frac{x^{1/2}}{\log^2 x}\right)$
$\theta(x) = \psi(x) - x^{1/2} + O(x^{1/3})$.

§18. The zeros of the Riemann zeta function.

Recall that

$$N(T) := \begin{cases} \#\{\overset{\beta+i\gamma}{\overset{\parallel}{\rho}} : \zeta(\rho) = 0 \text{ and } 0 < \beta < 1,\ 0 < \gamma < T\}, & \text{when } T \neq \gamma \text{ for } \rho = \beta + i\gamma \\ (N(T^+) + N(T^-))/2, & \text{when } T = \gamma \text{ for some } \rho = \beta + i\gamma. \end{cases}$$

We briefly explore the basic properties of $N(T)$.

Theorem 18·1. When $t \in \mathbb{R}$, put

$$S(t) = \frac{1}{\pi} \arg \zeta(\tfrac{1}{2} + it).$$

Then, whenever $T > 0$, one has

$$N(T) = \frac{1}{\pi} \arg \Gamma(\tfrac{1}{4} + \tfrac{1}{2} iT) - \frac{T}{2\pi} \log \pi + S(T) + 1.$$

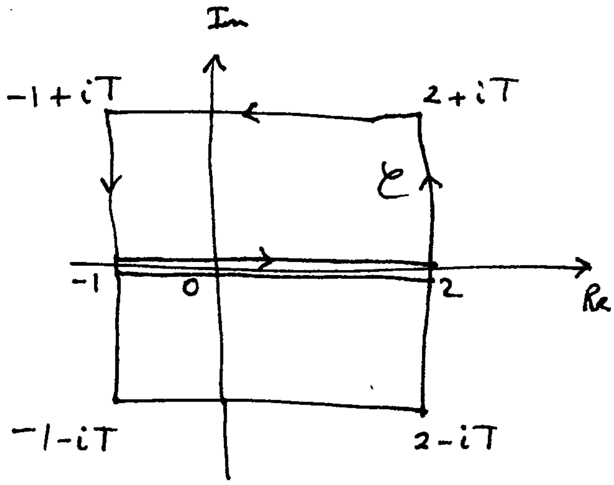
Proof. We have $N(T) = \frac{1}{2}(N(T^+) + N(T^-))$

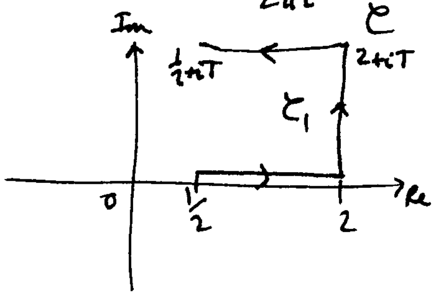
and $S(T) = \frac{1}{2}(S(T^+) + S(T^-))$,

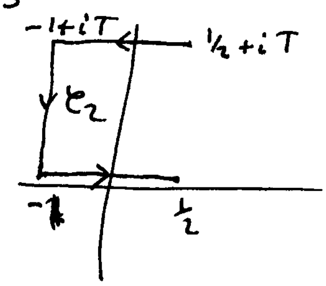
and thus we may restrict attention to the situation in which T is not the ordinate of a zero.

We may restrict attention to zeros with positive imaginary part. One has

$$N(T) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\zeta'}{\zeta}(s)\, ds.$$







We have

$$N(T) = \frac{1}{2\pi i}\int_{\mathcal{C}_1} \frac{\xi'}{\xi}(s)\,ds + \frac{1}{2\pi i}\int_{\mathcal{C}_2} \frac{\xi'}{\xi}(s)\,ds.$$

For $s \in \mathcal{C}_2$, we apply the identity arising from the functional equation

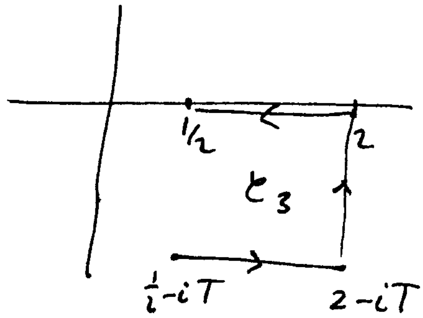
$$\frac{\xi'}{\xi}(s) = -\frac{\xi'}{\xi}(1-s),$$

whence

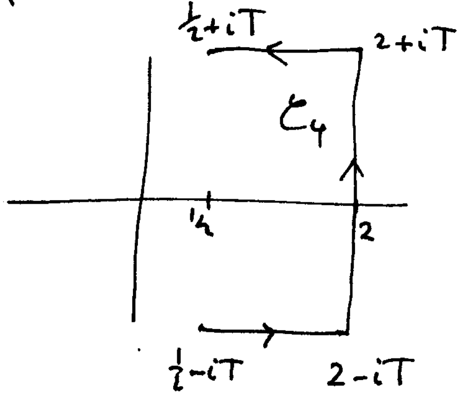
$$\int_{\mathcal{C}_2} \frac{\xi'}{\xi}(s)\,ds = -\int_{\mathcal{C}_2} \frac{\xi'}{\xi}(1-s)\,ds = \int_{\mathcal{C}_3} \frac{\xi'}{\xi}(s)\,ds.$$

Thus

$$N(T) = \frac{1}{2\pi i}\int_{\mathcal{C}_4} \frac{\xi'}{\xi}(s)\,ds,$$



where



But

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

so

$$N(T) = \frac{1}{2\pi i}\Big[\log s + \log(s-1) + \log\zeta(s) + \log\Gamma\left(\tfrac{1}{2}s\right) - \tfrac{1}{2}s\log\pi\Big]_{\frac{1}{2}-iT}^{\frac{1}{2}+iT}$$

By the Schwarz reflection principle, the real parts cancel and the imaginary parts reinforce. Thus

$$N(T) = \frac{1}{\pi}\Big(\arg\zeta\left(\tfrac{1}{2}+iT\right) + \arg\Gamma\left(\tfrac{1}{4}+\tfrac{1}{2}iT\right) - \tfrac{1}{2}T\log\pi + \arg\left(\tfrac{1}{2}+iT\right) + \arg\left(-\tfrac{1}{2}+iT\right)\Big)$$

$$= \frac{1}{\pi} \arg \Gamma\left(\tfrac{1}{4} + \tfrac{1}{2} iT\right) - \frac{T}{2\pi}\log \pi + S(T) + 1. \quad //$$

[Note: if $\arg(\tfrac{1}{2} + iT) = \theta$, then $\arg(-\tfrac{1}{2} + iT) = \pi - \theta$, so $\arg(\tfrac{1}{2}+iT) + \arg(-\tfrac{1}{2}+iT) = \pi$].

We can convert this conclusion into one almost explicit by making use of Stirling's formula:

$$\log \Gamma(s) = (s - \tfrac{1}{2})\log s - s + \tfrac{1}{2}\log(2\pi) + O(1/|s|).$$

<u>Corollary 18.2.</u> When $T \geq 2$, one has

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

<u>Proof.</u> We have $\arg(\Gamma(\tfrac{1}{4} + \tfrac{1}{2} iT)) = \operatorname{Im}(\log \Gamma(\tfrac{1}{4} + \tfrac{1}{2} iT))$, and according to Stirling's formula,

$$\operatorname{Im}(\log \Gamma(\tfrac{1}{4} + \tfrac{1}{2} iT)) = \operatorname{Im}\left((-\tfrac{1}{4} + \tfrac{1}{2} iT)\log(\tfrac{1}{4} + \tfrac{1}{2} iT) - (\tfrac{1}{4} + \tfrac{1}{2} iT)\right) + O(\tfrac{1}{T}).$$

Moreover,

$$\arg(\tfrac{1}{4} + iT/2) = \frac{\pi}{2} + O(1/T)$$

and

$$\log\left(\tfrac{1}{16} + \tfrac{1}{4}T^2\right) = 2\log\left(\frac{T}{2}\right) + O\left(\frac{1}{T^2}\right),$$

so

$$\begin{aligned}\operatorname{Im}(\log \Gamma(\tfrac{1}{4} + \tfrac{1}{2} iT)) &= -\tfrac{1}{4}\arg(\tfrac{1}{4} + \tfrac{1}{2} iT) + \frac{T}{4}\log\left(\frac{1}{16} + \frac{T^2}{4}\right) - \frac{T}{2} + O\left(\frac{1}{T}\right) \\ &= \frac{T}{2}\log\left(\frac{T}{2}\right) - \frac{T}{2} - \frac{\pi}{8} + O\left(\frac{1}{T}\right).\end{aligned}$$

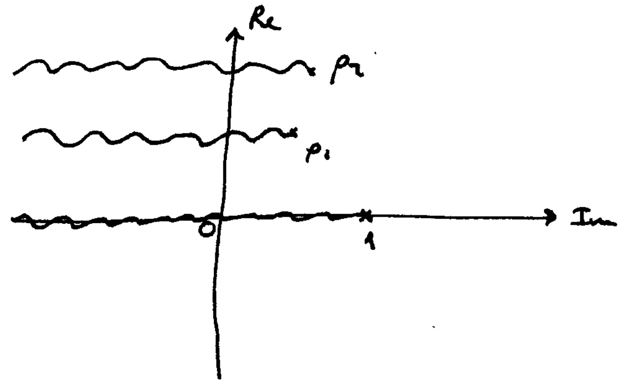
Substituting this conclusion into ~~that~~ of Theorem 18.1, we deduce ~~that~~

$$\begin{aligned}N(T) &= \frac{T}{2\pi}\log\left(\frac{T}{2}\right) - \frac{T}{2\pi} - \frac{1}{8} - \frac{T}{2\pi}\log\pi + S(T) + 1 + O\left(\frac{1}{T}\right) \\ &= \frac{T}{2\pi}\log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi} + \frac{7}{8} + S(T) + O\left(\frac{1}{T}\right). \quad //\end{aligned}$$

~~Recall that~~ we have defined

$$S(t) := \frac{1}{\pi}\arg \zeta(\tfrac{1}{2} + it).$$

A question we have avoided thus far is whether ~~this quantity~~ is well-defined. The function $\log \zeta(s)$ has a branch point at $s=1$, and also at the zeros ρ of the zeta function.



In order to obtain a single branch of $\log \zeta(s)$ (and hence to properly define $\arg \zeta(s)$), we insert branch cuts $(-\infty, 1]$ and $(-\infty + i\gamma, \beta + i\gamma]$ for each zero $\rho = \beta + i\gamma$ of $\zeta(s)$. The remaining complex set is simply connected, and we may take as the branch of $\log \zeta(s)$ that for which $\log \zeta(s) \to 0$ as $\sigma \to \infty$ (consistent with $\zeta(s) \to 1$ as $\sigma \to \infty$). We may then define, for $t \neq \gamma$ for any zero $\rho = \beta + i\gamma$ of $\zeta(s)$, $\arg \zeta(s) = \operatorname{Im} \log \zeta(s)$ by continuous variation from $\infty + it$ to $\sigma + it$, i.e.

$$\arg \zeta(s) = -\int_\sigma^\infty \operatorname{Im} \frac{\zeta'}{\zeta}(\alpha + it)\, d\alpha .$$

When $t = \gamma$ for some zero $\rho = \beta + i\gamma$ of $\zeta(s)$, then instead we put

$$\arg \zeta(s) = \tfrac{1}{2}\left(\arg \zeta(\sigma + it^+) + \arg \zeta(\sigma + it^-)\right).$$

<u>Lemma 18.3.</u> When $t \in \mathbb{R}$ and $-1 \leq \sigma < 2$, one has

$$\arg \zeta(\sigma + it) \ll \log \tau \qquad (\text{recall: } \tau = |t| + 3).$$

<u>Proof.</u> Suppose that $-1 \leq \sigma < 2$ and $t \neq \gamma$ for any zero $\rho = \beta + i\gamma$ of $\zeta(s)$. Then one has

$$\arg \zeta(\sigma + it) = \arg \zeta(2 + it) - \int_\sigma^2 \operatorname{Im} \frac{\zeta'}{\zeta}(\alpha + it)\, d\alpha .$$

But $\arg(\zeta(2+it)) \ll 1$ uniformly in t, since

$$\log \zeta(s) = \sum_{n=1}^\infty \frac{\Lambda(n)}{\log n} n^{-s} \ll 1 \qquad \text{for } \sigma > 1 .$$

Recall next from Lemma 15·1 ~~that~~

$$\frac{\zeta'}{\zeta}(s) = -\frac{1}{s-1} + \sum_{\substack{\rho \\ |\gamma - t| \le 1}} \frac{1}{s-\rho} + O(\log \tau) \qquad (-1 \le \sigma \le 2).$$

Then

$$\arg \zeta(\sigma+it) = -\sum_{|\gamma-t|\le 1} \int_\sigma^2 \operatorname{Im}\left(\frac{1}{\alpha+it-\rho}\right) d\alpha + O(\log \tau)$$

$$= \sum_{|\gamma-t|\le 1} \int_\sigma^2 \frac{t-\gamma}{(\alpha-\beta)^2+(t-\gamma)^2}\, d\alpha + O(\log\tau)$$

$$= \sum_{|\gamma-t|\le 1} \left(\tan^{-1}\left(\frac{2-\beta}{t-\gamma}\right) - \tan^{-1}\left(\frac{\sigma-\beta}{t-\gamma}\right)\right) + O(\log\tau).$$

When $t > \gamma$ the summand lies between 0 and π, whereas for $t < \gamma$, it lies between 0 and $-\pi$. In any case it is bounded. Moreover, we have

$$N(t+1) - N(t-1) \ll \log(|t|+2),$$

and hence

$$\arg\zeta(\sigma+it) \ll N(t+1) - N(t-1) + O(\log\tau)$$

$$\ll \log\tau.$$

Finally, when $t = \gamma$ for some zero $\rho = \beta + i\gamma$ of $\zeta(s)$, one has

$$\arg\zeta(\sigma+it) = \tfrac{1}{2}\left(\arg\zeta(\sigma+it^+) + \arg\zeta(\sigma+it^-)\right)$$

$$\ll \log\tau. \quad //$$

<u>Theorem 18.4.</u> When $T \ge 2$, one has

$$N(T) = \frac{T}{2\pi}\log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi} + \frac{7}{8} + O(\log T).$$

<u>Proof.</u> It follows from Lemma 18.3 that $S(T) \ll \log T$, and hence the desired conclusion is immediate from Corollary 18.2. //

§19. Bounds for $\zeta(s)$ conditional on RH.

We make use of an explicit formula to estimate $\zeta(s)$ and $\frac{\zeta'}{\zeta}(s)$, ~~thus~~ making use of smooth weights. We begin by making some observations concerning Riesz typical means. Consider a sequence $(a_n)_{n=1}^{\infty}$ with associated Dirichlet series
$$\alpha(s) = \sum_{n=1}^{\infty} a_n n^{-s}$$
and for which $A(x) := \sum_{n \le x} a_n \ll x^{\sigma_c + \varepsilon}$. When $k \in \mathbb{N}$ and $x \in \mathbb{R}$, we consider the Riesz ~~typical~~ mean
$$R_k(x) = \frac{1}{k!} \sum_{n \le x} a_n \left(\log \frac{x}{n}\right)^k.$$
Thus
$$R_k(x) = \int_0^x R_{k-1}(u)\,\frac{du}{u} \qquad \text{with} \qquad R_0(x) = A(x),$$
and hence $R_k(x) \ll x^{\theta}$ whenever $\theta > \max\{0, \sigma_c\}$.

[Notice here ~~that~~
$$\int_0^x R_{k-1}(u)\,\frac{du}{u} = \frac{1}{(k-1)!} \int_0^x \sum_{n \le u} a_n \left(\log \frac{u}{n}\right)^{k-1} \frac{du}{u}$$
$$= \frac{1}{(k-1)!} \sum_{n \le x} a_n \int_n^x \left(\log \frac{u}{n}\right)^{k-1} \frac{du}{u}$$
$$= \frac{1}{(k-1)!} \sum_{n \le x} a_n \frac{1}{k} \left[\left(\log \frac{u}{n}\right)^k\right]_n^x = \frac{1}{k!} \sum_{n \le x} a_n \left(\log \frac{x}{n}\right)^k = R_k(x).$$
]

We have
$$\alpha(s) = s \int_1^{\infty} \underset{R_0(x)}{\underset{\parallel}{A(x)}}\, x^{-s-1}\,dx,$$
and so by ~~integrating~~ by parts repeatedly we ~~obtain~~
$$\alpha(s) = s^{k+1} \int_1^{\infty} R_k(x)\, x^{-s-1}\,dx \qquad (\sigma > \max\{0, \sigma_c\}).$$
But ~~then~~
$$R_k(x) = \frac{1}{2\pi i} \int_{\sigma_0 - i\infty}^{\sigma_0 + i\infty} \alpha(s) \frac{x^s}{s^{k+1}}\,ds \qquad (x > 0 \ \& \ \sigma_0 > \max\{0, \sigma_c\})$$
(Perron's formula proof: ~~use that~~
$$\frac{1}{2\pi i} \int_{\sigma_0 - i\infty}^{\sigma_0 + i\infty} \frac{y^s}{s^{k+1}}\,ds = \operatorname{Res}\left(\frac{y^s}{s^{k+1}}\right)\Big|_{s=0} = \frac{1}{k!}(\log y)^k.)$$

From this discussion, we see that with

$$W(u) = W(x;y;u) = \begin{cases} 1, & \text{when } 1 \le u \le x; \\ 1 - \dfrac{\log(u/x)}{\log y}, & \text{when } x \le u \le xy; \\ 0, & \text{when } u \ge xy, \end{cases}$$

we have

$$(\log y) \sum_{n \le xy} W(n) \frac{\Lambda(n)}{n^s} = \frac{1}{2\pi i} \int_{\sigma_0 - i\infty}^{\sigma_0 + i\infty} -\frac{\zeta'}{\zeta}(s+w) \left(\frac{(xy)^w}{w^2} - \frac{x^w}{w^2} \right) dw$$

$$= \sum_{n \le xy} \frac{\Lambda(n)}{n^s} \left(\log \frac{xy}{n} \right) - \sum_{n \le x} \frac{\Lambda(n)}{n^s} \left(\log \frac{x}{n} \right)$$

$$= \sum_{n \le x} (\log y) \frac{\Lambda(n)}{n^s} + \sum_{x \le n \le xy} (\log y - \log(n/x)) \frac{\Lambda(n)}{n^s}.$$

Thus
$$\sum_{n \le xy} W(n) \frac{\Lambda(n)}{n^s} = -\frac{1}{2\pi i \log y} \int_{\sigma_0 - i\infty}^{\sigma_0 + i\infty} \frac{\zeta'}{\zeta}(s+w) \left(\frac{(xy)^w - x^w}{w^2} \right) dw.$$

We can move the contour to the left to show that, provided that $s \ne 1$ and $\zeta(s) \ne 0$, one has

$$\sum_{n \le xy} W(n) \frac{\Lambda(n)}{n^s} = \underset{\text{pole at } w=0}{\underset{\uparrow}{-\frac{\zeta'}{\zeta}(s)}} + \underset{\text{pole at } w = -s+1}{\frac{(xy)^{1-s} - x^{1-s}}{(1-s)^2 \log y}}$$

$$+ \left(-\underset{\text{poles at } w = -s+\rho}{\underset{\uparrow}{\sum_\rho \frac{(xy)^{\rho - s} - x^{\rho - s}}{(\rho - s)^2 \log y}}} - \underset{\text{poles at } w = -s-2k.}{\underset{\uparrow}{\sum_{k=1}^{\infty} \frac{(xy)^{-2k-s} - x^{-2k-s}}{(2k+s)^2 \log y}}} \right) \quad \text{—— (19.1)}$$

<u>Theorem 19.1.</u> Suppose that RH is true. Then one has

$$\left| \frac{\zeta'}{\zeta}(s) \right| \le \sum_{n \le (\log \tau)^2} \frac{\Lambda(n)}{n^\sigma} + O\left((\log \tau)^{2 - 2\sigma} \right),$$

uniformly for $\frac{1}{2} + 1/\log\log\tau \le \sigma \le 3/2$ and $|t| \ge 1$.

Proof. We begin by estimating the sum over ~~zeros~~ ρ in (19.1). Here, on noting that when $\sigma \geq 1/2$, one has

$$|y^{\rho-s} - 1| \leq 2, \qquad \text{(recall - assuming RH here!)}$$

and thus

$$\left| \sum_{\rho} \frac{(xy)^{\rho-s} - x^{\rho-s}}{(\rho-s)^2 \log y} \right| \ll \frac{2x^{\frac{1}{2}-\sigma}}{\log y} \sum_{\rho} \frac{1}{|s-\rho|^2}.$$

In order to estimate this sum, we make use of relation (14.2) from Corollary 14.3:

$$\frac{\zeta'}{\zeta}(s) = B + \tfrac{1}{2}\log\pi - \frac{1}{s-1} - \frac{1}{2}\cdot\frac{\Gamma'}{\Gamma}\left(\tfrac{1}{2}s+1\right) + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right),$$

where $B = -\sum_{\rho} \mathrm{Re}\left(\frac{1}{\rho}\right)$. Then

$$\sum_{\rho} \mathrm{Re}\left(\frac{1}{s-\rho}\right) = \mathrm{Re}\left(\frac{\zeta'}{\zeta}(s)\right) + \frac{1}{2}\mathrm{Re}\left(\frac{\Gamma'}{\Gamma}\left(\tfrac{1}{2}s+1\right)\right) - \tfrac{1}{2}\log\pi + \mathrm{Re}\left(\frac{1}{s-1}\right)$$

$$\Rightarrow \sum_{\rho} \frac{\sigma-\frac{1}{2}}{(\sigma-\frac{1}{2})^2+(t-\gamma)^2} = \mathrm{Re}\left(\frac{\zeta'}{\zeta}(s)\right) + \frac{1}{2}\mathrm{Re}\left(\frac{\Gamma'}{\Gamma}\left(\tfrac{1}{2}s+1\right)\right) - \tfrac{1}{2}\log\pi + \frac{\sigma-1}{(\sigma-1)^2+t^2}.$$

Since $\frac{\Gamma'}{\Gamma}(s) = \log s + O(1/|s|)$, it follows that

$$\left(\sigma-\tfrac{1}{2}\right)\sum_{\rho} \frac{1}{(\sigma-\frac{1}{2})^2+(t-\gamma)^2} = \mathrm{Re}\left(\frac{\zeta'}{\zeta}(s)\right) + \tfrac{1}{2}\log\tau + O(1).$$

Hence in (19.1) we see that there is a complex number ω with $|\omega| \leq 1$ such that

$$\frac{\zeta'}{\zeta}(s) = -\sum_{n \leq xy} w(n)\frac{\Lambda(n)}{n^s} + \omega\frac{2x^{\frac{1}{2}-\sigma}}{(\sigma-\frac{1}{2})\log y}\left|\mathrm{Re}\left(\frac{\zeta'}{\zeta}(s)\right)\right|$$
$$+ O\left(\frac{x^{1/2-\sigma}}{(\sigma-\frac{1}{2})\log y}\cdot\log\tau\right) + O\left(\frac{(xy)^{1-\sigma}}{\tau^2} + \frac{x^{1-\sigma}}{\tau^2}\right). \qquad \text{(19.2)}$$

We shall arrange that

$$\frac{2x^{1/2-\sigma}}{(\sigma-\frac{1}{2})\log y} \leq c \quad \text{for some real constant } c < 1.$$

Then we see that

$$\frac{\zeta'}{\zeta}(s) \ll \left|\sum_{n \leq xy} w(n)\frac{\Lambda(n)}{n^s}\right| + \frac{x^{\frac{1}{2}-\sigma}\log\tau}{(\sigma-\frac{1}{2})\log y} + \frac{(xy)^{1-\sigma}}{\tau^2} + \frac{x^{1-\sigma}}{\tau^2}. \qquad \text{(19.3)}$$

Put
$$y = \exp\left(\frac{1}{\sigma-\frac12}\right) \quad \text{and} \quad x = \frac{(\log\tau)^2}{y}.$$

Then we have
$$\frac{2x^{\frac12-\sigma}}{(\sigma-\frac12)\log y} = 2e\,(\log\tau)^{1-2\sigma} \leq 2e\cdot\exp\left(-2\frac{\log\log\tau}{\log\log\tau}\right) = \frac{2}{e}<1,$$
$$\left(\text{using } \sigma \geq \tfrac12 + \tfrac{1}{\log\log\tau}\right)$$

and since
$$\sum_{n\leq xy} w(n)\frac{\Lambda(n)}{n^s} \ll \sum_{n\leq(\log\tau)^2}\frac{\Lambda(n)}{n^{\frac12}} \ll \log\tau,$$

we conclude ~~that~~, uniformly for $\sigma \geq \frac12 + \frac{1}{\log\log\tau}$ and $|t|\geq 1$, one has
$$\frac{\zeta'}{\zeta}(s) \ll \log\tau.$$

Finally, by inserting ~~this upper~~ bound into (19.2), we see ~~that~~
$$\left|\frac{\zeta'}{\zeta}(s)\right| \leq \sum_{n\leq xy} w(n)\frac{\Lambda(n)}{n^\sigma} + O\Bigg(\underbrace{\frac{x^{\frac12-\sigma}\log\tau}{(\sigma-\frac12)\log y} + \frac{(xy)^{1-\sigma}}{\tau^2} + \frac{x^{1-\sigma}}{\tau^2}}_{=O((\log\tau)^{2-2\sigma})}\Bigg)$$
$$= \sum_{n\leq(\log\tau)^2}\frac{\Lambda(n)}{n^\sigma} + O\Bigg(\sum_{x\leq n\leq \underbrace{xy}_{=(\log\tau)^2}}\frac{\log(n/x)}{\log y}\frac{\Lambda(n)}{n^\sigma}\Bigg) + O((\log\tau)^{2-2\sigma}).$$

But
$$\sum_{x\leq n\leq \underbrace{xy}_{=(\log\tau)^2}}\frac{\log(n/x)}{\log y}\frac{\Lambda(n)}{n^\sigma} \leq \sum_{x\leq n\leq xy}\frac{\Lambda(n)}{n^\sigma} \ll (xy)^{-\sigma}\sum_{n\leq xy}\Lambda(n) \ll (xy)^{1-\sigma}$$
$$\ll (\log\tau)^{2-2\sigma}.$$

We ~~therefore~~ conclude ~~that~~
$$\left|\frac{\zeta'}{\zeta}(s)\right| \leq \sum_{n\leq(\log\tau)^2}\frac{\Lambda(n)}{n^\sigma} + O\left((\log\tau)^{2-2\sigma}\right). \quad //$$

<u>Corollary 19.2.</u> Suppose that RH is true. Then for $\frac{1}{2} + \frac{1}{\log\log\tau} \le \sigma \le \frac{3}{2}$ and $|t| \ge 1$, one has

$$\frac{\zeta'}{\zeta}(s) \ll \left((\log\tau)^{2-2\sigma} + 1\right) \min\left\{\frac{1}{|\sigma-1|}, \log\log\tau\right\}.$$

<u>Proof.</u> By a weak version of the Prime Number Theorem (such as Chebyshev's upper bound), we have

$$\sum_{U \le n < eU} \frac{\Lambda(n)}{n^\sigma} \ll U^{-\sigma} \sum_{U \le n < eU} \Lambda(n) \ll U^{1-\sigma},$$

whence for $\sigma \ne 1$ one has

$$\sum_{n \le (\log\tau)^2} \frac{\Lambda(n)}{n^\sigma} \ll \frac{1 + (\log\tau)^{2-2\sigma}}{|1 - e^{1-\sigma}|}$$

$$\ll \left((\log\tau)^{2-2\sigma} + 1\right) \min\left\{\frac{1}{|\sigma-1|}, 1\right\}.$$

When $|\sigma - 1| \le 1/\log\log\tau$, meanwhile, one has

$$\sum_{n \le (\log\tau)^2} \frac{\Lambda(n)}{n^\sigma} \ll \sum_{n \le (\log\tau)^2} \frac{\Lambda(n)}{n} \ll \log\log\tau,$$

and thus, uniformly in σ,

$$\sum_{n \le (\log\tau)^2} \frac{\Lambda(n)}{n^\sigma} \ll \left((\log\tau)^{2-2\sigma} + 1\right) \min\left\{\frac{1}{|\sigma-1|}, \log\log\tau\right\}.$$

By substituting this conclusion into Theorem 19.1, we conclude that

$$\frac{\zeta'}{\zeta}(s) \ll \left((\log\tau)^{2-2\sigma} + 1\right) \min\left\{\frac{1}{|\sigma-1|}, \log\log\tau\right\}. \quad //$$

Corollary 19.3. Suppose ~~that~~ RH is true. Then for $\frac{1}{2} + \frac{1}{\log\log\tau} \le \sigma \le \frac{3}{2}$ and $|t| \ge 1$, one has

$$|\log \zeta(s)| \le \sum_{n \le (\log\tau)^2} \frac{\Lambda(n)}{n^\sigma \log n} + O\left(\frac{(\log\tau)^{2-2\sigma}}{\log\log\tau}\right).$$

Proof. We have

$$\log \zeta(\sigma+it) = \log\zeta(\tfrac{3}{2}+it) - \int_\sigma^{3/2} \frac{\zeta'}{\zeta}(\alpha+it)\,d\alpha.$$

$$\Rightarrow \quad |\log \zeta(\sigma+it)| \le |\log \zeta(\tfrac{3}{2}+it)| + \int_\sigma^{3/2} \left|\frac{\zeta'}{\zeta}(\alpha+it)\right| d\alpha.$$

By Theorem 19.1, we ~~therefore~~ deduce ~~that~~

$$|\log\zeta(s)| \le |\log(\zeta(\tfrac{3}{2}+it))| + \sum_{n \le (\log\tau)^2} \Lambda(n) \int_\sigma^{3/2} n^{-\alpha}\,d\alpha + O\left(\int_\sigma^{3/2} (\log\tau)^{2-2\alpha}\,d\alpha\right)$$

We have

$$\int_\sigma^{3/2} n^{-\alpha}\,d\alpha = \frac{n^{-\sigma}-n^{-3/2}}{\log n} \quad \text{and} \quad \int_\sigma^{3/2} (\log\tau)^{2-2\alpha}\,d\alpha \ll \frac{(\log\tau)^{2-2\sigma}}{\log\log\tau}.$$

Also,

$$|\log \zeta(\tfrac{3}{2}+it)| = \left|\sum_{n=1}^\infty \frac{\Lambda(n)}{\log n} n^{-\frac{3}{2}-it}\right| \le \sum_{n=1}^\infty \frac{\Lambda(n)}{\log n} n^{-3/2}$$

$$\le \sum_{n \le (\log\tau)^2} \frac{\Lambda(n)}{\log n} n^{-\frac{3}{2}} + \sum_{n > (\log\tau)^2} \frac{\Lambda(n)}{\log n} n^{-3/2}.$$

We ~~therefore~~ conclude ~~that~~

$$|\log\zeta(s)| \le \sum_{n \le (\log\tau)^2} \frac{\Lambda(n)}{\log n} n^{-\sigma} + \sum_{n > (\log\tau)^2} \frac{\Lambda(n)}{\log n} n^{-3/2} + O\left(\frac{(\log\tau)^{2-2\sigma}}{\log\log\tau}\right).$$

The middle term on the right hand side may be estimated via Chebyshev's inequality once again, yielding the bound

$$\sum_{U<n\le eU} \frac{\Lambda(n)}{\log n} n^{-3/2} \ll \frac{U^{-3/2}}{\log U}\cdot U = \frac{U^{-1/2}}{\log U},$$

whence

$$\sum_{n>(\log\tau)^2} \frac{\Lambda(n)}{\log n} n^{-3/2} \ll \frac{(\log\tau)^{-1}}{\log\log\tau} \ll \frac{(\log\tau)^{2-2\sigma}}{\log\log\tau}.$$

Then
$$|\log\zeta(s)| \leqslant \sum_{n\le(\log\tau)^2} \frac{\Lambda(n)}{\log n} n^{-\sigma} + O\left(\frac{(\log\tau)^{2-2\sigma}}{\log\log\tau}\right). \quad //$$

We may now translate the bound of Corollary 19.3 into a uniform estimate.

<u>Corollary 19.4.</u> Suppose that RH is true. Then:

(i) for $1+1/\log\log\tau \le \sigma \le 3/2$, one has
$$|\log\zeta(s)| \le \log\left(\frac{1}{\sigma-1}\right) + O(\sigma-1);$$

(ii) for $1-1/\log\log\tau \le \sigma \le 1+1/\log\log\tau$, one has
$$|\log\zeta(s)| \le \log\log\log\tau + O(1);$$

(iii) for $1/2 + 1/\log\log\tau \le \sigma \le 1 - 1/\log\log\tau$, one has
$$|\log\zeta(s)| \le \log\left(\frac{1}{1-\sigma}\right) + O\left(\frac{(\log\tau)^{2-2\sigma}}{(1-\sigma)\log\log\tau}\right).$$

<u>Proof.</u> We begin with establishing part (i). Here we can note merely that when $1<\sigma\le 3/2$, one has
$$|\log\zeta(s)| = \left|\sum_{n=1}^{\infty}\frac{\Lambda(n)}{\log n} n^{-s}\right| \le \sum_{n=1}^{\infty}\frac{\Lambda(n)}{\log n} n^{-\sigma} = \log\zeta(\sigma)$$
$$= \log\left(\frac{1}{\sigma-1}+O(1)\right) = \log\left(\frac{1}{\sigma-1}\right)+O(\sigma-1). \quad \square$$

Next, part (ii). We may suppose that $1-1/\log\log\tau \le \sigma \le 1+1/\log\log\tau$. We put $z=(\log\tau)^2$, so that $\sigma = 1+O(1/\log z)$. Thus, from Corollary 19.3 we have

$$|\log \zeta(s)| \le \sum_{n\le z} \frac{\Lambda(n)}{n\log n} + \sum_{n\le z} \frac{\Lambda(n)}{\log n}\left(n^{-\sigma} - n^{-1}\right) + O\left(\frac{z^{1/\log z}}{\log z}\right).$$

Here, since by Mertens' ~~estimates~~ we have

$$\sum_{n\le z} \frac{\Lambda(n)}{n\log n} = \log\log z + O(1) = \log\log\log \tau + O(1),$$

and

$$n^{-\sigma} - n^{-1} = \int_1^{\sigma} n^{-\alpha}\log n \; d\alpha \ll |\sigma - 1|\, n^{-1}\log n,$$

we find ~~that~~

$$|\log\zeta(s)| \le \log\log\log\tau + O(1) + \underbrace{|\sigma-1| \sum_{n\le z}\frac{\Lambda(n)}{n}}_{\ll\, |\sigma-1|\log z \,\ll\, 1}.$$

Thus $|\log\zeta(s)| \le \log\log\log\tau + O(1)$. □

Finally, we turn to part (iii). We again put $z = (\log\tau)^2$, and note that when $\frac{1}{2} + \frac{1}{\log\log\tau} \le \sigma \le 1 - \frac{1}{\log\log\tau}$, one has by R-S ~~integration~~,

$$\sum_{n\le z} \frac{\Lambda(n)}{\log n} n^{-\sigma} = \int_{2^-}^{z} \frac{1}{u^{\sigma}\log u}\, d\psi(u)$$

$$= \int_{2^-}^{z} \frac{du}{u^{\sigma}\log u} + \frac{\psi(z) - z}{z^{\sigma}\log z} + \frac{2}{2^{\sigma}\log 2} + \int_2^z \frac{\psi(u) - u}{u^{\sigma+1}\log u}\left(\sigma + \frac{1}{\log u}\right) du$$

$$= li(z^{1-\sigma}) - li(2^{1-\sigma}) + O\left(\frac{z^{1/2}(\log z)^2}{z^{\sigma}\log z}\right) + O\left(\int_2^z \frac{u^{1/2}(\log u)^2}{u^{\sigma+1}(\log u)}\, du\right)$$

$$= O\left(\frac{z^{1-\sigma}}{(1-\sigma)\log z}\right) + \int_{2^{1-\sigma}}^{2} \frac{dv}{\log v} + O\Big(z^{\frac{1}{2}-\sigma}\log z + \underbrace{\int_2^{\infty} u^{-\sigma-\frac{1}{2}}\log u\, du}_{\ll (\sigma - \frac{1}{2})^{-2}}\Big)$$

$$= \int_{2^{1-\sigma}}^{2} \left(\frac{1}{v-1} + O(1)\right) dv + O\left(\frac{(\log\tau)^{2-2\sigma}}{(1-\sigma)\log\log\tau}\right).$$

Then by Corollary 19.3 we see that

$$|\log \zeta(s)| \leq -\log(2^{1-\sigma}-1) + O(1) + O\left(\frac{(\log \tau)^{2-2\sigma}}{(1-\sigma)\log\log \tau}\right)$$

$$= \log\left(\frac{1}{\sigma-1}\right) + O\left(\frac{(\log \tau)^{2-2\sigma}}{(1-\sigma)\log\log \tau}\right). \quad \square \quad //$$

We may now return to the zeros of $\zeta(s)$.

__Lemma 19.5.__ Suppose that RH is true. Then for $T \geq 4$, one has

$$N(T + 1/\log\log T) - N(T) \ll \frac{\log T}{\log\log T}.$$

__Proof.__ We apply Corollary 19.2 with $s = \frac{1}{2} + \frac{1}{\log\log T} + iT$. Thus

$$\frac{\zeta'}{\zeta}(s) \ll \log T,$$

and from Lemma 15.1 we see ~~that~~

$$\sum_{\substack{\rho \\ |\gamma - T| \leq 1}} \frac{1}{s-\rho} = \frac{\zeta'}{\zeta}(s) + \frac{1}{s-1} + O(\log T)$$

$$\ll \log T.$$

For each zero ρ occurring in this summation, one has $\operatorname{Re}\left(\frac{1}{s-\rho}\right) > 0$, and, moreover, for $T \leq \gamma \leq T + 1/\log\log T$ one has

$$\operatorname{Re}\left(\frac{1}{s-\rho}\right) = \operatorname{Re}\left(\frac{1}{\frac{1}{\log\log T} + i(T-\gamma)}\right) \geq \frac{1/\log\log T}{\frac{1}{(\log\log T)^2} + \frac{1}{(\log\log T)^2}}$$

$$\geq \frac{1}{2}\log\log T.$$

Hence

$$\sum_{\substack{\rho \\ T \leq \gamma \leq T + 1/\log\log T}} \left(\frac{1}{2}\log\log T\right) \ll \log T$$

$$\Rightarrow \quad N(T + 1/\log\log T) - N(T) \ll \log T/\log\log T. \quad //$$

We can apply this conclusion to refine the argument of Lemma 18.3.

Lemma 19.6. Suppose that RH is true. Then, whenever

$$|\sigma - 1/2| \le 1/\log\log\tau,$$

one has
$$\frac{\zeta'}{\zeta}(s) = \sum_{\substack{\rho \\ |\gamma - t| \le \frac{1}{\log\log\tau}}} \frac{1}{s-\rho} + O(\log\tau).$$

Proof. Put $\sigma_1 = \frac{1}{2} + \frac{1}{\log\log\tau}$, and then take $s_1 = \sigma_1 + it$. By Lemma 15.1, we have

$$\frac{\zeta'}{\zeta}(s) = \underbrace{\frac{\zeta'}{\zeta}(s_1)}_{\substack{\text{Cor. 19.2 } \Uparrow \\ \log\tau}} + \sum_{|\gamma - t| \le 1} \left(\frac{1}{s-\rho} - \frac{1}{s_1 - \rho} \right) + O(\log\tau).$$

$$= \sum_{\substack{\rho \\ |\gamma - t| \le 1/\log\log\tau}} \frac{1}{s-\rho} - \sum_{\substack{\rho \\ |\gamma - t| \le 1/\log\log\tau}} \frac{1}{s_1 - \rho}$$
$$+ \sum_{k=1}^{\infty} \sum_{\substack{|\gamma - t| \le 1 \\ \frac{k}{\log\log\tau} \le |\gamma - t| \le \frac{k+1}{\log\log\tau}}} \left(\frac{1}{s-\rho} - \frac{1}{s_1-\rho} \right) + O(\log\tau).$$

The second term here, in which (by Lemma 19.5) there are $O\left(\frac{\log\tau}{\log\log\tau}\right)$ summands, contributes

$$\ll \frac{\log\tau}{\log\log\tau} \cdot \frac{1}{1/\log\log\tau} \ll \log\tau.$$

Meanwhile, the third term, for the summand indexed by k, again by Lemma 19.5, contributes for each k:

$$\ll \frac{\log\tau}{\log\log\tau} \cdot \frac{|s - s_1|}{|s-\rho||s_1-\rho|} \ll \frac{(\log\tau)}{(\log\log\tau)} \cdot \frac{1/(\log\log\tau)}{k^2/(\log\log\tau)^2}$$

Thus the contributions from all integers k in this third term is

$$\ll \sum_{k=1}^{\infty} \frac{\log \tau}{k^2} \ll \log \tau .$$

Combining these estimates, we conclude that

$$\frac{\zeta'}{\zeta}(s) = \sum_{\substack{\rho \\ |\gamma - t| \leq 1/\log\log\tau}} \frac{1}{s-\rho} + O(\log \tau).$$

//

<u>Theorem 19.7.</u> Suppose that RH is true. Then for $\sigma \geq \frac{1}{2}$ and $|t| \geq 1$, one has

$$\arg \zeta(s) \ll \frac{\log \tau}{\log\log\tau}.$$

<u>Proof.</u> Wlog one may assume that $t \neq \gamma$ for any zero $\rho = \beta + i\gamma$ of $\zeta(s)$. Put $\sigma_1 = \frac{1}{2} + \frac{1}{\log\log\tau}$ and $s = \sigma_1 + it$. If $\sigma \geq \sigma_1$, then from Corollary 19.4 (iii) one sees that (when $\sigma \leq 1 - 1/\log\log\tau$)

$$|\arg \zeta(s)| \leq |\log \zeta(s)| \leq \log\left(\frac{1}{1-\sigma}\right) + O\left(\frac{(\log\tau)^{2-2\sigma}}{(1-\sigma)\log\log\tau}\right)$$

$$\ll \log\log\log\tau + \frac{(\log\tau)}{\log\log\tau} \ll \frac{\log\tau}{\log\log\tau},$$

whilst for $\sigma > 1 - 1/\log\log\tau$, one finds similarly from Corollary 19.4 (i), (ii) that

$$|\arg \zeta(s)| \ll \log\log\log\tau.$$

Suppose then that $\frac{1}{2} \leq \sigma \leq \sigma_1$, and note that

$$\arg \zeta(s) = \arg \zeta(s_1) - \int_{\sigma}^{\sigma_1} \mathrm{Im}\left(\frac{\zeta'}{\zeta}(\alpha + it)\right) d\alpha .$$

Since $0 \leq \sigma_1 - \sigma \leq 1/\log\log\tau$, it follows from Lemma 19.6 that

$$\arg \zeta(s) = \arg \zeta(s_1) - \sum_{|\gamma - t| \leq 1/\log\log\tau} \int_{\sigma}^{\sigma_1} \mathrm{Im}\left(\frac{1}{\alpha + it - \rho}\right) d\alpha$$

$$+ O\left(\frac{\log\tau}{\log\log\tau}\right)$$

Using Corollary 19.4 (iii) as above, we ~~thus~~ see ~~that~~

$$\arg \zeta(s) = - \sum_{|\gamma - t| \le \frac{1}{\log\log \tau}} \int_{\sigma}^{\sigma_1} \operatorname{Im}\left(\frac{1}{\alpha + it - \rho}\right) d\alpha + O\left(\frac{\log \tau}{\log\log \tau}\right)$$

$$= - \sum_{|\gamma - t| \le \frac{1}{\log\log \tau}} \underbrace{\left(\tan^{-1}\left(\frac{\sigma - 1/2}{\gamma - t}\right) - \tan^{-1}\left(\frac{\sigma_1 - 1/2}{\gamma - t}\right)\right)}_{= O(1)} + O\left(\frac{\log \tau}{\log\log \tau}\right)$$

$$\overset{\text{Lemma 19.5}}{\ll} \frac{\log \tau}{\log\log \tau}. \quad //$$

<u>Corollary 19.8.</u> One has, assuming the truth of RH,

$$S(t) \ll \frac{\log t}{\log\log t},$$

and hence

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

<u>Proof.</u> We have proved ~~that~~

$$S(T) = \frac{1}{\pi} \arg \zeta(\tfrac{1}{2} + iT) \overset{\text{Thm 19.7}}{\ll} \frac{\log T}{\log\log T}.$$

Hence, by Corollary 18.2,

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

$$= \frac{T}{2\pi} \log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi} + O\left(\frac{\log T}{\log\log T}\right). \quad //$$