Vinogradov's Three Primes Theorem

Ian N Petrow Supervisor: Prof. Ben Green

2/5/08

Contents

1	Introduction		
	1.1	Major and Minor Arcs	3
	1.2	Proof of Vinogradov's Theorem	5
2	Major Arcs		
	2.1	Explicit Formula	13
		2.1.1 The Mellin Transform	14
		2.1.2 Perron's Formula and application to $\Lambda(n)$	15
	2.2	$N(T, \chi)$	21
		2.2.1 Hadamard Product for $\xi(s, \chi)$	22
		2.2.2 Several Estimates for L'/L	24
	2.3	Zero-Free Regions for $L(s, \gamma)$	26
		2.3.1 Regions for Complex γ	$\frac{-}{26}$
		2.3.2 Regions for Real γ	29
		2.3.3 Siegel Zeros	30
	2.4	Siegel's Theorem	31
	2.1	2.4.1 Goldfeld's Proof	33
	2.5	Final estimate for $\psi(x, y)$	34
	2.6	Ramanujan Sums	39
3	Mir	or Arcs	41
U	3.1	Vaughan's Identity	41
	3.2	Orthogonality of $\mu(n)$ and $e(n\alpha)$	47
	0.2	3.2.1 Some Harmonic Analysis Tools	49
4	Con	clusion	56

Chapter 1

Introduction

The Goldbach Conjecture is one of the most fascinating unsolved problems in number theory. It was (famously) first posed by Christian Goldbach in his 1742 letter to Leonard Euler. His original question was whether all integers greater than 5 are representable as a sum of three primes. The problem immediately reduces to asking whether all even numbers greater than 2 are representable as a sum of two primes.

Although it is still an open conjecture to show that all even numbers are expressible as a sum of two primes, the case for odd numbers is easier. It was shown in 1937 by I. M. Vinogradov that all sufficiently large odd integers are expressible as a sum of three primes. Vinogradov proved the three-primes theorem by analytical means, using a major arc/minor arc decomposition. The major arcs are treated by estimating the sum

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

The proof of this estimate relies on a large portion of classical theory of Dirichlet *L*-functions, and has as an immediate corollary the corresponding estimate for $\pi(x, a, q)$, the number of primes less than x equivalent to $a \mod q$ (a, q relatively prime).

The minor arcs are treated by Vinogradov's type I/type II sum decomposition. Such a decomposition provides bounds on the sum

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

which are sensitive to α . In this essay we work with the möbius function instead of the Von Mangoldt function, and apply a significant amount of harmonic analysis to achieve our results.

To prove the three-primes theorem, we work with the sum

$$r(n) = \sum_{k_1+k_2+k_3=n} \Lambda(k_1)\Lambda(k_2)\Lambda(k_3),$$

which counts representations of n as a sum of three prime powers, with a weight of $\log(p_1)\log(p_2)\log(p_3)$ attached to each such representation. The contribution to this sum deriving from proper prime powers is small compared to the contribution from primes themselves (use partial summation). If there are no such representations, r(n) = 0. For the proof, we bound r(n) away from zero for odd n. We now state

Theorem 1.0.1 (Vinogradov). For any fixed A > 0,

$$r(N) = \frac{1}{2} \mathfrak{S}(N) N^2 + O\left(\frac{N^2}{\log^A(N)}\right),$$

where

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

If N is even, the above is not useful. Indeed, assume N is even. Then one of the k_i is a power of two. If we consider the number of representations of N as a sum of a power of 2 and two integers, this gives a trivial upper bound of $r(N) = O(N \log^4(N))$. When N is even, $\mathfrak{S}(N) = 0$ in the above, and the bound given by Vinogradov is considerably worse than the trivial one.

For N odd, we have

$$\prod_{p \neq 2} \left(1 - \frac{1}{(p-1)^2} \right) = .6602..., \text{ and } \prod_p \left(1 + \frac{1}{(p-1)^3} \right) = 2.301...,$$

and so $1.320... < \mathfrak{S}(N) < 2.301...$ Hence, the N^2 term eventually dominates the second term and hence gives a lower bound for r(N).

The remainder of the introductory chapter of this paper is devoted to the proof of Vinogradov's theorem modulo 3 key lemmas. The first should really not be called a lemma, but a theorem, for it is the bound on $\psi(x, \chi)$, equivalent to the prime number theorem for arithmetic progressions. For this estimate, we follow a hybrid of Green's notes [Gr1] and Davenport's classical book [Dav]. The second is a basic fact about Ramanujan sums. The proof of these two facts forms chapter 2. The third lemma is an estimate of the exponential sum $\sum \mu(n)e(n\alpha)$ originally by Davenport, however we follow a proof of Green and Tao [GT], with some harmonic analysis from Montgomery [Mo]. The proof of this lemma makes up chapter 3. The precise statements of all of these can be found in the remainder of this chapter.

1.1 Major and Minor Arcs

We now proceed with the proof of the main theorem. To begin, consider the exponential sum,

$$S(\alpha) = \sum_{k \le N} \Lambda(n) e(k\alpha).$$

Then

$$S(\alpha)^{3} = \sum_{k_{1},k_{2},k_{3} \leq N} \Lambda(k_{1})\Lambda(k_{2})\Lambda(k_{3})e((k_{1}+k_{2}+k_{3})\alpha)$$
$$= \sum_{n} \left(\sum_{\substack{k_{1}+k_{2}+k_{3}=n\\k_{1},k_{2},k_{3} \leq N}} \Lambda(k_{1})\Lambda(k_{2})\Lambda(k_{3})\right)e(n\alpha)$$
$$= \sum_{n} r(n,N)e(n\alpha)$$

Note that r(n) = r(n, N) for $n \leq N$. Hence, we see that $S(\alpha)^3$ is a Fourier series, and thus the coefficients r(N) can be recovered

$$r(N) = \int_{\mathbb{R}/\mathbb{Z}} S(\alpha)^3 e(-N\alpha) \, d\alpha.$$
 (1.1)

Bounding this integrand from below with therefore give us a bound on r(n) and Vinogradov's theorem. The integrand of (1.1) turns out to be large when α is near a rational number with small denominator, and small when far from one. Hence, we proceed by a major arc / minor arc decomposition which is common in such problems of additive number theory.

We divide \mathbb{R}/\mathbb{Z} into subintervals \mathfrak{M} (major arcs) and \mathfrak{m} (minor arcs). The major arcs correspond to subintervals close to a rational number with small denominator, and the minor arcs their complement in \mathbb{R}/\mathbb{Z} . We make these notions precise as follows. Let $P = \log^B(N)$, and $Q = N/\log^{2B}(N)$, where B will be chosen later in terms of A (see A from Vinogradov's theorem, above). For all $q \leq P$, and all $1 \leq a \leq q$ with (a,q) = 1, define

$$\mathfrak{M}(a,q) = \{ \alpha \in \mathbb{R}/\mathbb{Z} \text{ such that } |\alpha - \frac{a}{q}| \leq \frac{1}{Q} \}.$$

And then let \mathfrak{M} be the union of all the $\mathfrak{M}(a,q)$. Let \mathfrak{m} be the complement of \mathfrak{M} in \mathbb{R}/\mathbb{Z} .

Observation 1.1.1. m is non-empty.

Proof:

Any two of the $\mathfrak{M}(a,q)$ are disjoint. Indeed, taking $a/q\neq a'/q',$ we see for large enough N that

$$\left|\frac{a}{q} - \frac{a'}{q'}\right| \ge \frac{1}{qq'} \ge \frac{1}{P^2} > \frac{2}{Q}.$$

Hence \mathfrak{M} is not all of \mathbb{R}/\mathbb{Z} .

We now carry out a separate analysis for each of \mathfrak{M} and \mathfrak{m} .

1.2 Proof of Vinogradov's Theorem

We start with an individual $\mathfrak{M}(a,q)$. Take the Gauss sum

$$\tau(\chi) = \sum_{m \in \mathbb{Z}/q\mathbb{Z}} \chi(m) e\left(\frac{m}{q}\right),$$

where χ is any character to modulus q. We then have that

$$\frac{1}{\phi(q)}\sum_{\chi\in\widehat{\mathbb{Z}/q\mathbb{Z}}}\chi(n)\tau(\overline{\chi})=0 \qquad \text{if } (n,q)>1$$

because then $\chi(n) = 0$ for all n. If (n,q) = 1, then

$$\begin{split} \frac{1}{\phi(q)} \sum_{\chi \in \widehat{\mathbb{Z}/q\mathbb{Z}}} \chi(n)\tau(\overline{\chi}) &= \frac{1}{\phi(q)} \sum_{\chi \in \widehat{\mathbb{Z}/q\mathbb{Z}}} \sum_{m \in \mathbb{Z}/q\mathbb{Z}} \chi(n)\overline{\chi}(m)e\left(\frac{m}{q}\right) \\ &= \frac{1}{\phi(q)} \sum_{\chi \in \widehat{\mathbb{Z}/q\mathbb{Z}}} \sum_{h \in \mathbb{Z}/q\mathbb{Z}} \overline{\chi}(h)e\left(\frac{nh}{q}\right) \\ &= \frac{1}{\phi(q)} \sum_{h \in \mathbb{Z}/q\mathbb{Z}} e\left(\frac{nh}{q}\right) \sum_{\chi \in \widehat{\mathbb{Z}/q\mathbb{Z}}} \overline{\chi}(h) \\ &= e\left(\frac{n}{q}\right) \end{split}$$

where in the second line we have taken $m \equiv nh \mod q$. With reference to $\mathfrak{M}(a,q)$, we now take $\alpha = a/q + \beta$. We then have

$$\begin{split} S(\alpha) &= \sum_{\substack{k \le N \\ (k,q)=1}} \Lambda(k) e(\frac{ka}{q}) e(k\beta) + O(\log^2(N)) \\ &= \frac{1}{\phi(q)} \sum_{k \le N} \Lambda(k) \sum_{\chi \in \widehat{\mathbb{Z}/q\mathbb{Z}}} \chi(ka) \tau(\overline{\chi}) e(k\beta) + O(\log^2(N)) \\ &= \frac{1}{\phi(q)} \sum_{\chi \in \widehat{\mathbb{Z}/q\mathbb{Z}}} \tau(\overline{\chi}) \chi(a) \sum_{k \le N} \chi(k) \Lambda(k) e(k\beta) + O(\log^2(N)). \end{split}$$

Taking the inner sum of the above and applying summation by parts we get

$$\sum_{k \le N} \chi(k) \Lambda(k) e(k\beta) = e(N\beta) \psi(N,\chi) - 2\pi i\beta \int_1^N e(u\beta) \psi(u,\chi) \, du.$$
(1.2)

Where

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

as in the introduction. We analyze the trivial character χ_0 separately. We now need our first

Lemma 1.2.1. Let $\chi \neq \chi_0$ be any character to modulus q. If $q \leq \log^M(x)$ for some positive constant M, then we have

$$|\psi(x,\chi)| = O(xe^{-C(M)\sqrt{\log(x)}})$$
(1.3)

for some positive constant C(M) which depends only on M.

As noted above, this "lemma" is essentially equivalent to the prime number theorem for arithmetic progressions (and correspondingly makes up a significant portion of this paper).

Combining the lemma with formula (1.2) we get

$$\sum_{k \le N} \chi(k) \Lambda(k) e(k\beta) = O((1 + |\beta|N)Ne^{-c\sqrt{\log(N)}}).$$

Now we estimate the contribution from the trivial character χ_0 . Recall that $\psi(x)$ is defined

$$\psi(x) = \sum_{n \le x} \Lambda(n)$$

as in the prime number theorem. Because $\chi_0(n) = 1$ for (n,q) = 1 and $\chi_0(n) = 0$ for (n,q) > 1, $\psi(x,\chi_0)$ resembles $\psi(x)$, precisely,

$$|\psi(x,\chi_0) - \psi(x)| \le \sum_{\substack{n \le x \\ (n,q) > 1}} \Lambda(n) = O(\log(q)\log(x)).$$

Recall that the prime number theorem gives

$$\psi(x) = x + O(xe^{-c\sqrt{\log(x)}}),$$

the proof of which is simpler and follows the same form as our lemma 1.1. Hence, setting $\psi(u, \chi_0) = \lfloor u \rfloor + R(u)$, we have

$$R(u) = O(ue^{-c\sqrt{\log(u)}}).$$
(1.4)

Similar to the origin of (1.2), we set

$$T(\beta) = \sum_{1 \le k \le N} e(k\beta),$$

and apply summation by parts to get

$$T(\beta) = e(N\beta) - 2\pi i\beta \int_1^N e(u\beta) \lfloor u \rfloor du$$

Combine this with (1.2) and then apply (1.4) in the second line below to get

$$\sum_{k \le N} \chi_0(k) \Lambda(k) e(k\beta) = T(\beta) + e(N\beta) R(N) - 2\pi i\beta \int_1^N e(u\beta) R(u) \, du$$
$$= T(\beta) + O((1+|\beta|N) N e^{-c\sqrt{\log(N)}}).$$

Therefore, the sum on the left for the trivial character only differs from the sums of the other characters by the term $T(\beta)$. We have hence reduced our sum to

$$S(\alpha) = \frac{1}{\phi(q)} \left(\tau(\chi_0) (T(\beta) + O((1 + |\beta|N)Ne^{-c\sqrt{\log(N)}})) + \sum_{\chi \neq \chi_0} \tau(\overline{\chi})\chi(a) \cdot O((1 + |\beta|N)Ne^{-c\sqrt{\log(N)}}) \right)$$
(1.5)

Two quick facts about Gauss sums will reduce this expression further.

Observation 1.2.2. $\tau(\chi_0) = \mu(q)$

Proof:

$$\tau(\chi_0) = \sum_{m=1}^q \chi_0(m) e\left(\frac{m}{q}\right)$$
$$= \sum_{\substack{m \le q \\ (m,q)=1}} e\left(\frac{m}{q}\right)$$
$$= \sum_{d=1}^q \mu(d) \sum_{1 \le m' \le q/d} e\left(\frac{m'd}{q}\right)$$
$$= \mu(q)$$

Where we have set m'd = m.

 \Box .

Observation 1.2.3. $|\tau(\chi)| \leq q^{\frac{1}{2}}$

Proof:

Recall that if χ is primitive,

$$\chi(n)\tau(\overline{\chi}) = \sum_{h\in\mathbb{Z}/q\mathbb{Z}}\overline{\chi}(h)e\left(\frac{nh}{q}\right),$$

hence

$$\sum_{n \in \mathbb{Z}/q\mathbb{Z}} |\chi(n)|^2 |\tau(\overline{\chi})|^2 = \sum_{h_1 \in \mathbb{Z}/q\mathbb{Z}} \sum_{h_2 \in \mathbb{Z}/q\mathbb{Z}} \overline{\chi}(h_1)\chi(h_2) \sum_{n \in \mathbb{Z}/q\mathbb{Z}} e\left(\frac{n(h_1 - h_2)}{q}\right).$$

Therefore

$$\phi(q)|\tau(\overline{\chi})|^2 = q \sum_{h \in \mathbb{Z}/q\mathbb{Z}} |\chi(h)|^2 = q\phi(q).$$

Hence

$$|\tau(\chi)| = q^{\frac{1}{2}}$$

if χ is primitive. In the case of χ imprimitive, one can show that $\tau(\chi) \leq q^{\frac{1}{2}}$, however, we omit this proof for sake of brevity. \Box .

These two facts applied to (1.5) yield

$$S(\alpha) = \frac{\mu(q)}{\phi(q)} T(\beta) + O((1 + |\beta|N)q^{\frac{1}{2}}Ne^{-c\sqrt{\log(N)}}).$$

Now one applies the fact that our α lies in the major arc $\mathfrak{M}(a,q)$. Using $q \leq P$, and $|\beta| \leq 1/Q$ one obtains,

$$S(\alpha) = \frac{\mu(q)}{\phi(q)}T(\beta) + O(Ne^{-c\sqrt{\log(N)}}),$$

thereby making significant improvement to the error term. We now cube and take the fourier series to find the contribution of $\mathfrak{M}(a,q)$ to the integral (1.1) we originally set out to study;

$$\int_{\mathfrak{M}(a,q)} S(\alpha)^3 e(-N\alpha) \, d\alpha = \frac{\mu(q)}{\phi(q)^3} e\left(\frac{-aN}{q}\right) \int_{-1/Q}^{1/Q} T(\beta)^3 e(-N\beta) \, d\beta + O(N^2 e^{-c'\sqrt{\log(N)}})$$

Summing over all major arcs, we have

$$\int_{\mathfrak{M}} S(\alpha)^3 e(-N\alpha) \, d\alpha = \sum_{q \le P} \frac{\mu(q)}{\phi(q)^3} \left(\sum_{\substack{a=1\\(a,q)=1}}^q e\left(\frac{aN}{q}\right) \right) \int_{-1/Q}^{1/Q} T(\beta)^3 e(-N\beta) \, d\beta + O(N^2 e^{-c''\sqrt{\log(N)}})$$
(1.6)

The inside sum

$$c_q(n) = \sum_{\substack{a=1\\(a,q)=1}}^{q} e\left(\frac{aN}{q}\right)$$

is known as Ramanujan's sum. To evaluate the first factor in (1.6), we employ our second

Lemma 1.2.4. Let $c_q(n)$ be the Ramanujan sum defined above.

- 1. The sum $c_q(n)$ is multiplicative.
- 2. If p^{α} is the highest power of p dividing n, $c_q(n)$ is given on prime powers by

$$c_{p^{\beta}}(n) = \begin{cases} \phi(p^{\beta}) & \text{if } \beta \leq \alpha \\ -p^{\alpha} & \text{if } \beta = \alpha + 1 \\ 0 & \text{if } \beta > \alpha + 1 \end{cases}$$

The proof of which is actually rather short. Nonetheless, we confine it to Chapter 2. Now observe that all of the functions in sum on the first factor of (1.6) are multiplicative, and thus we may evaluate by euler product,

$$\begin{split} \sum_{q=1}^{\infty} \frac{\mu(q)}{\phi(q)^3} c_q(N) &= \prod_p \left(1 - \frac{c_p(N)}{(p-1)^3} \right) \\ &= \prod_{p|N} \left(1 - \frac{1}{(p-1)^2} \right) \prod_{p \nmid N} \left(1 + \frac{1}{(p-1)^3} \right) \\ &= \mathfrak{S}(N). \end{split}$$

From the trivial estimate $|c_q(n)| \leq \phi(q)$, we find that

$$\sum_{q>P} \frac{\mu(q)}{\phi(q)^3} c_q(N) = O\left(\sum_{q>P} \frac{1}{\phi(q)^2}\right) = O(\log^{-B+1}(N))$$

Hence

$$\sum_{q \le P} \frac{\mu(q)}{\phi(q)^3} c_q(N) = \mathfrak{S}(N) + O(\log^{-B+1}(N)).$$
(1.7)

Now we evaluate the factor in (1.6) involving the integral. We have

$$\int_{-1/Q}^{1/Q} T(\beta)^3 e(-N\beta) \, d\beta = \int_0^1 T(\beta)^3 e(-N\beta) \, d\beta + O\left(\int_{1/Q}^{1-1/Q} |T(\beta)|^3 \, d\beta\right).$$

We observe that $T(\beta)$ is the sum of a geometric series by definition, and hence it is evaluated

$$T(\beta) = \frac{e((N+1)\beta) - e(\beta)}{e(\beta) - 1} = O(\min\left(N, \frac{1}{||\beta||_{\mathbb{R}/\mathbb{Z}}}\right)),$$

where $||\cdot||_{\mathbb{R}/\mathbb{Z}}$ is the distance of \cdot to the nearest integer. We thus have

$$\int_{-1/Q}^{1/Q} T(\beta)^3 e(-N\beta) \, d\beta = \int_0^1 T(\beta)^3 e(-N\beta) \, d\beta + O(Q^2) = \int_0^1 T(\beta)^3 e(-N\beta) \, d\beta + O\left(\frac{N^2}{\log^{4B}(N)}\right).$$

By analogous construction to the beginning of section 1.1, the integral on the right of this equation is equal to the number of ways of representing N as a sum of three arbitrary integers, $N = k_1 + k_2 + k_3$, hence

$$\int_0^1 T(\beta)^3 e(-N\beta) \, d\beta = \frac{1}{2}(N-1)(N-2) = \frac{1}{2}N^2 + O(N).$$

Therefore,

$$\int_{-1/Q}^{1/Q} T(\beta)^3 e(-N\beta) \, d\beta = \frac{1}{2}N^2 + O\left(\frac{N^2}{\log^{4B}(N)}\right). \tag{1.8}$$

Combining (1.6), (1.7) and (1.8), we get that

$$\int_{\mathfrak{M}} S(\alpha)^3 e(-N\alpha) \, d\alpha = \frac{1}{2} \mathfrak{S}(N) N^2 + O\left(\frac{N^2}{\log^{B-1}(N)}\right). \tag{1.9}$$

Vinogradov's theorem follows if we can show that the minor arcs contribute a smaller amount. We now proceed to their treatment. Take

$$\begin{aligned} \left| \int_{\mathfrak{m}} S(\alpha)^{3} e(-N\alpha) \, d\alpha \right| &\leq \left(\max_{m} |S(\alpha)| \right) \int_{\mathfrak{m}} |S(\alpha)^{2}| \, d\alpha \\ &\leq \left(\max_{m} |S(\alpha)| \right) \int_{0}^{1} |S(\alpha)^{2}| \, d\alpha. \end{aligned}$$

The integral evaluates as

$$\int_0^1 |S(\alpha)^2| \, d\alpha = \sum_{k_1 \le N} \Lambda(k_1) \sum_{k_2 \le N} \Lambda(k_2) \int_0^1 e((k_1 - k_2)\alpha) \, d\alpha$$
$$= \sum_{k \le N} \Lambda(k)^2 = O(N \log(N)).$$

We must now bound $S(\alpha)$ itself on \mathfrak{m} . We have

$$\left|\alpha - \frac{a}{q}\right| > \frac{1}{Q} = \frac{\log^{2B}(N)}{N}$$

for all $q \leq P$ if $\alpha \in \mathfrak{m}$. Hence

$$\inf_{1 \le q \le P} \left| \alpha - \frac{a}{q} \right| > \frac{\log^{2B}(N)}{N}$$
(1.10)

We now evoke our third

Lemma 1.2.5. Let $\alpha \in \mathbb{R}$, and C > 0. If for sufficiently large N, and any fixed k we have

$$\inf_{1 \le d \le 16 \log^{8(C+4)}(N)} ||d\alpha||_{\mathbb{R}/\mathbb{Z}} \ge k \frac{\log^{28(C+4)}(N)}{N},$$

then

,

$$|\frac{1}{N}\sum_{1 \le n \le N} \mu(n)e(n\alpha)| = O(\log^{-C+1}(N))$$

which essentially says that $\alpha \in \mathfrak{m}$ causes $\mu(n)$ to be asymptotically orthogonal to $e(n\alpha)$. A statement about the möbius function is easily transformed into a statement about the von Mangoldt function. We have

$$\Lambda(n) = \sum_{d|n} \log(d) \mu\left(\frac{n}{d}\right).$$

Taking $\mu(x) = 0$ if $x \notin \mathbb{Z}$, we may write

$$\Lambda(n) = \sum_{d \le n} \log(d) \mu\left(\frac{n}{d}\right).$$

Hence

$$\begin{split} S(\alpha) &= \sum_{n \leq N} \Lambda(n) e(n\alpha) &= \sum_{n \leq N} \sum_{\delta \leq n} \log(\delta) \mu\left(\frac{n}{\delta}\right) e(n\alpha) \\ &= \sum_{\delta \leq N} \log(\delta) \sum_{m \leq \frac{N}{\delta}} \mu(m) e(m\delta\alpha). \end{split}$$

Our minor arc condition (1.10) implies

$$\inf_{1 \le q \le \log^B(N)} ||q\alpha||_{\mathbb{R}/\mathbb{Z}} > \frac{\log^{4B}(N)}{N}.$$

Taking B = 9(C + 4) satisfies the conditions of the lemma. We then may evaluate

$$S(\alpha) < \sum_{\delta \le N} \log(\delta) \frac{N/\delta}{\log^{\frac{B}{9}-5}(N/\delta)}$$
$$\leq \frac{1}{\log^{\frac{B}{9}-6}(N)} \sum_{\delta \le N} \frac{N}{\delta}$$
$$= O\left(\frac{N}{\log^{\frac{B}{9}-7}(N)}\right)$$

Combining this estimate with the above, we then have

$$\int_{\mathfrak{m}} S(\alpha)^3 e(-N\alpha) \, d\alpha = O\left(\frac{N^2}{\log^{\frac{B}{9}-6}(N)}\right) \tag{1.11}$$

Upon taking B = 9(A + 6) and combining (1.9) and (1.11) we have proved Vinogradov's theorem.

In the next chapter we treat lemmas (1.2.1) and (1.2.4), which are used to bound the major arcs above. The proof of lemma (1.2.1) makes up the bulk of the chapter. We mainly follow Davenport for chapter 2. In chapter 3 we prove lemma (1.2.5) used to bound the minor arcs. The treatment here is more contemporary, following a paper of Green and Tao on the quadratic uniformity of the möbius function.

Chapter 2

Major Arcs

2.1 Explicit Formula

For the bulk of this chapter we work to derive good bounds on the sum

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

For each modulus q there are a collection of $\phi(q)$ such sums, and the prime number theorem in arithmetic progressions quickly follows from knowing good bounds on these. In the Dirichlet convolution algebra we have $\Lambda = \log *\mu$. Because Dirichlet characters χ are completely multiplicative, we also have

$$\chi(n)\Lambda(n) = \sum_{d|n} (\chi(d)\log(d)) \cdot \left(\chi\left(\frac{n}{d}\right)\mu\left(\frac{n}{d}\right)\right)$$

hence $\chi \Lambda = \chi \log * \chi \mu$. Convolution in the Dirichlet algebra becomes multiplication of the corresponding Dirichlet series,

$$\sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n^s} = \left(\sum_{i=1}^{\infty} \frac{\chi(i)\log(i)}{i^s}\right) \cdot \left(\sum_{j=1}^{\infty} \frac{\chi(j)\mu(j)}{j^s}\right)$$
$$= -\frac{L'(s,\chi)}{L(s,\chi)},$$

valid for $\operatorname{Re}(s) > 1$. From the above, we see the connection between our sum $\psi(x, \chi)$ and Dirichlet *L*-functions. To obtain a partial sum from such a series, we work with the *L*-functions and employ a variant of Perron's formula and obtain the formula

$$\psi(x,\chi) = -\sum_{|\rho| \le T} \tilde{\phi}(\rho) - \sum_{m=0}^{\infty} \tilde{\phi}(\mathfrak{a} - 2m) + O_{\phi}\left(\frac{x\log(x)\log^2(qT)}{T}\right)$$
(2.1)

Where the first sum is over all nontrivial zeros of $L(s,\chi)$, ϕ is a smooth cutoff function, and $\tilde{\phi}$ is its Mellin transform, which we define and discuss later. Note Dirichlet *L*-functions have slightly different forms for "odd" and "even" characters. The difference is realized in the above formula via our definition of

$$\mathfrak{a} = \begin{cases} 0 & \text{if } \chi(-1) = 1 \\ 1 & \text{if } \chi(-1) = -1 \end{cases}$$

Where in the first case the character is called "even" and the second it is called "odd". Equation (2.1) is called the *explicit formula* for the Dirichlet *L*-function. It is particularly important because it directly shows the duality between prime numbers (i.e. the sum $\psi(x, \chi)$) and non-trivial zeros of the *L*-function (i.e. the first sum on the right side of (2.1)). This connection is a guiding principle of analytic number theory.

From this point forward, we derive a bound on $\psi(x, \chi)$ by bounding each term in the explicit formula. In the remainder of this section we show how the explicit formula is derived and set up an approximate formula where we restrict to non-trivial zeros with imaginary part less than T. In the next two sections (2.2) and (2.3), we study $N(T, \chi)$ (that is, the number of zeros of $L(s, \chi)$ in the critical strip with imaginary part less than T), and zero-free regions of $L(s, \chi)$. These allow us the bound the first term in the above, and are essentially the most we can say about the location of the all-important critical zeros. In section (2.3) we encounter the complication of Siegel zeros - which may or may not exist - and spend the following section (2.4) bounding their contribution as best we can via Siegel's theorem. Finally, in section (2.5) we evaluate all of our Mellin transforms, and obtain our final estimate for $\psi(x, \chi)$. The brief proof of lemma (1.2.4) on Ramanujan Sums makes up section (2.6).

2.1.1 The Mellin Transform

We want to analyze the sum $\psi(x,\chi)$. However, to apply analytic techniques, it is often more appropriate to reduce to a smoothed function. Therefore, we take the function $\phi(n)$ to be any C^{∞} function on \mathbb{R} , compactly supported on a subset of $(1,\infty)$. Instead of $\psi(x,\chi)$ we work with

$$\psi_{\phi}(x,\chi) = \sum_{n} \chi(n) \Lambda(n) \phi(n).$$

We prove the explicit formula for $\psi_{\phi}(x,\chi)$, and then reduce to (2.1) later. The proof of the explicit formula works extensively with Mellin transforms, which for $s \in \mathbb{C}$ we define

$$\tilde{\phi}(s) = \int_0^\infty \phi(x) x^s \, \frac{dx}{x},$$

the Mellin transform of ϕ . Because ϕ has compact support, $\tilde{\phi}$ is analytic. The Mellin transform is analogous to the Fourier transform, except with respect to the group \mathbb{R}^+ with multiplication. We have that x^s is a character for this group, and that dx/x is an invariant measure. In like fashion to Fourier, we have the

Proposition 2.1.1 (Mellin Inversion Theorem). Let $\phi : \mathbb{R} \to \mathbb{R}$ be any C^{∞} function with compact support contained in $(1, \infty)$. Then for $\sigma \in \mathbb{R}$ we have

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

Proof:

Follows from the Fourier inversion theorem and a change of variables. \Box .

We know that in fourier analysis, smoothness of a function is manifested in the decay of its fourier transform. Similar properties hold with Mellin transforms. In fact, if ϕ were analytic, then $\tilde{\phi}$ would have exponential decay in vertical strips (i.e. in the $\pm i\infty$ direction), see Titchmarsh [Tit]. However, analytic functions which approximate the characteristic function of an interval do not exist. We must instead use the following lemma which applies to C^{∞} functions and gives a slightly weaker result.

Lemma 2.1.2 (Mellin Decay). Suppose that $\phi : \mathbb{R} \to \mathbb{R}$ is a C^{∞} function supported in the interval [X, 2X] for some X > 1. If $\sigma_1 < \sigma_2$ are real numbers, we have the estimate

$$|\tilde{\phi}(\sigma + it)| \le C_m(\sigma_1, \sigma_2)|t|^{-m} X^{\sigma-1} \sum_{j=0}^m X^j ||D^j \phi||_{L^1}$$

uniformly for $\sigma \in [\sigma_1, \sigma_2]$. For fixed m, $C_m(\sigma_1, \sigma_2)$ grows at most polynomially in $|\sigma_1| + |\sigma_2|$.

The lemma more or less states that $\tilde{\phi}$ decays faster than any polynomial in vertical strips, where the constants depend on the order of polynomial, the support of ϕ and the size of its derivatives.

Proof:

Follows from the corresponding lemma about Fourier transforms. \Box .

2.1.2 Perron's Formula and application to $\Lambda(n)$

We now apply Mellin inversion and the regularity results above to derive a smoothed version of Perron's formula.

Proposition 2.1.3 (Perron's Formula with smooth cutoff). If (b_n) is an arithmetic sequence with Dirichlet series f(s) convergent for $Re(s) > s_0$, and ϕ : $\mathbb{R} \to \mathbb{R}$ is C^{∞} and compactly supported on some subset of $(1, \infty)$, then

$$\sum_{n=1}^{\infty} b_n \phi(n) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} f(s) \tilde{\phi}(s) \, ds$$

for $\sigma > s_0$.

Proof: We have

$$\sum_{n=1}^{\infty} b_n \phi(n) = \frac{1}{2\pi i} \sum_{n=1}^{\infty} b_n \int_{\sigma-i\infty}^{\sigma+i\infty} \tilde{\phi}(s) n^{-s} ds$$
$$= \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \left(\sum_{n=1}^{\infty} \frac{b_n}{n^s}\right) \tilde{\phi}(s) ds$$

Where the interchange of the summation and integration is justified wherever $\sigma > \sigma_0$.

Applying the formula to $\psi_{\phi}(x,\chi)$ we have

$$\sum_{n} \chi(n) \Lambda(n) \phi(n) = -\frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{L'(s,\chi)}{L(s,\chi)} \tilde{\phi}(s) \, ds$$

To get the explicit formula for $\psi_{\phi}(x,\chi)$ we evaluate the above by contour integration.

Proposition 2.1.4 (Explicit Formula with smooth cutoff ϕ). If ϕ is a C^{∞} function compactly supported on a subset of $(1, \infty)$, we have

$$\psi_{\phi}(x,\chi) = -\sum_{\rho} \tilde{\phi}(\rho) - \sum_{j=1}^{\infty} \tilde{\phi}(\mathfrak{a} - 2j)$$

where \mathfrak{a} equals 0 or 1 depending on whether χ is even or odd, and the first sum is over all of the non-trivial zeros ρ of $L(s, \chi)$ counted with multiplicity.

Proof:

We would like to prove the formula for ϕ compactly supported on some subset of $(1, \infty)$. However, it suffices to consider ϕ supported on [x, 2x] for x > 1because any other ϕ supported on a different interval may be reconstructed via a smooth partition of unity as a finite sum of intervals of this type. Therefore, we will be able to use lemma (2.1.2).

To prove the proposition, we will use a standard contour integration technique, integrating around a closed square to the left of the line $\operatorname{Re}(s) = \sigma$ and then expanding the contour to $-\infty$ to pick up the residues of all the poles of $L'(s,\chi) \cdot \tilde{\phi}(s)/L(s,\chi)$. Because $L'(s,\chi)$ and $\tilde{\phi}$ have no poles, we are only concerned with the zeros of $L(s,\chi)$. All trivial poles have

$$\operatorname{Res}_{s=\mathfrak{a}-2j}\frac{L'(s,\chi)}{L(s,\chi)} = 1$$

and all poles ρ in the critical strip have residue equal to their multiplicity.

We now choose a contour $C = C_1 \cup C_2 \cup C_3 \cup C_4$, with

$$\begin{array}{rcl} C_1 &=& [\sigma_0 - iT, \sigma_0 + iT] \\ C_2 &=& [\sigma_0 + iT, -R + iT] \\ C_3 &=& [-R + iT, -R - iT] \\ C_4 &=& [-R - iT, \sigma_0 - iT] \end{array}$$

With R and T later tending to infinity. In picking such a contour, we take care to pick R and T such that C_3 avoids any of the trivial poles of L'/L (i.e. choose $R \notin \mathbb{Z}$), and that C_2 and C_4 avoid any of the critical poles. If we define

$$I_j = -\frac{1}{2\pi i} \int_{C_j} \frac{L'(s,\chi)}{L(s,\chi)} \tilde{\phi}(s) \, ds,$$

The reside theorem states

$$I_1 + I_2 + I_3 + I_4 = -\sum_{|\rho| \le T} \tilde{\phi}(\rho) - \sum_{m=0}^{R/2} \tilde{\phi}(\mathfrak{a} - 2m).$$
(2.2)

If we let $T \to \infty$,

$$I_1 \longrightarrow \psi_{\phi}(x,\chi) = \sum_{n=1}^{\infty} \chi(n) \Lambda(n) \phi(n),$$

so to complete our proof, we must bound the integrals I_1, I_2 , and I_3 . Lemma 2.1.2 provides a bound for ϕ . We have the following lemmas which bound L'/L.

Lemma 2.1.5. *If* $-1 \le \sigma \le 2$ *,*

$$\frac{L'(\sigma+iT,\chi)}{L(\sigma+iT,\chi)} = O(\log^2(qT))$$

Proof:

This proof depends on material from section (2.2.1), however our arguments are not circular. Let $\rho = \beta + i\gamma$ be a nontrivial zero of $L(s, \chi)$. Observation (2.2.3) gives that the number of zeros with $|\gamma - T| < 1$ is $O(\log(q(|T| + 2)))$. There are then $O(\log(q(|T| + 2)))$ zeros with imaginary parts within an interval of length 2. Hence there must be a gap of length $> c \log(q(|T| + 2))$ amongst them. Therefore, by varying T by a bounded amount, we have that

$$|\gamma - T| > \frac{C}{\log(qT)}$$

for all zeros. Formula (2.13) from section (2.2.1), states that for s with $-1 \leq \operatorname{Re}(s) \leq 2$ one has

$$\frac{L'(s,\chi)}{L(s,\chi)} = \sum_{\substack{\rho = \beta + i\gamma \\ |\gamma - t| < 1}} \frac{1}{s - \rho} + O(\log(q(|t| + 2))).$$

From the above choice of T, we have that each term of the sum is $O(\log(qT))$, and from observation (2.2.3) that there are $O(\log(qt))$ terms in the sum. Hence, the lemma follows.

Lemma 2.1.6. For $s \in \mathbb{C}$ with $Re(s) \leq -1$ and excluding a circle around each trivial zero of L, we have

$$\frac{L'(s,\chi)}{L(s,\chi)} = O(\log(q|s|)).$$

Proof:

We need the functional equation for Dirichlet L-functions, which we will assume without proof,

$$\frac{(i)^{\mathfrak{a}}q^{\frac{1}{2}}}{\tau(\chi)}\left(\frac{q}{\pi}\right)^{\frac{1}{2}s+\frac{1}{2}\mathfrak{a}}\Gamma(\frac{1}{2}s+\frac{1}{2}\mathfrak{a})L(s,\chi) = \left(\frac{q}{\pi}\right)^{-\frac{1}{2}s+\frac{1}{2}(\mathfrak{a}+1)}\Gamma(-\frac{1}{2}s+\frac{1}{2}(\mathfrak{a}+1))L(1-s,\overline{\chi}).$$

Employing some standard functional equation relations from complex analysis (proved by analyzing Hadamard products),

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

we obtain

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

Applying this to the functional equation, we obtain an asymmetric form

$$L(1-s,\chi) = \frac{(i)^{\mathfrak{a}}q^{\frac{1}{2}}}{\tau(\chi)} 2^{1-s} \pi^{-s} q^{s-\frac{1}{2}} \cos(\frac{\pi}{2}(s-\mathfrak{a})) \Gamma(s) L(s,\overline{\chi}).$$

Taking the logarithmic derivative,

$$\frac{L'(1-s,\chi)}{L(1-s,\chi)} = -\log(2) - \log(\pi) + \log(q) - \frac{\pi}{2}\sin(\frac{\pi}{2}(s-\mathfrak{a}))\log\left(\cos(\frac{\pi}{2}(s-\mathfrak{a}))\right) + \frac{\Gamma'(s)}{\Gamma(s)} + \frac{L'(s,\overline{\chi})}{L(s,\overline{\chi})}$$

If we avoid the poles produced by the $\log(\cos(\cdot))$, we see that this term is bounded. Furthermore we have that the Γ'/Γ term is $O(\log(|s|))$, and that the L'/L term is bounded when $\operatorname{Re}(s) \geq 2$ (see later lemma (2.1.7), we obtain the estimate in the lemma.

With the bounds given in lemmas (2.1.2), (2.1.5) and (2.1.6) we can bound I_2 , I_3 and I_4 . We begin with I_2 and I_4 . With reference to lemma (2.1.2) (Mellin Decay), we will later choose ϕ to resemble the characteristic function of an interval. Therefore, we will have $||\phi||_{L^1} = O(x)$, $||D\phi||_{L^1} = O(1)$ and $||D^2\phi||_{L^1} = O(x)$.

For I_3 we use lemma (2.1.2) (Mellin Decay) with m = 2, and lemma (2.1.6). We have

$$\begin{aligned} |I_3| &\leq \frac{1}{2\pi} \int_{-R-iT}^{-R+iT} \left| \frac{L'(s,\chi)}{L(s,\chi)} \right| \cdot |\tilde{\phi}(s)| \, ds \\ &= \int_{-T}^{T} O(\log(q|-R+it|)) \cdot C_R |t|^{-2} x^{-R-1} (||\phi||_{L^1} + x||D\phi||_{L^1} + x^2 ||D^2\phi||_{L^1}) \, dt \\ &= O_{\phi}(C_R x^{-R+2} \log(q)) \int_{-T}^{T} \frac{\log(qt)}{t^2} \, dt \\ &= O_{\phi}(C_R x^{-R+2} \log(q)). \end{aligned}$$

Lemma (2.1.2) (Mellin Decay) gives that C_R is at most a polynomial in R. Because we have that x > 1, the integral over C_3 goes to 0 as $R \to \infty$.

We now address I_2 and I_4 . We use all three of our lemmas, but take m = 1 in lemma (2.1.2) (Mellin Decay) instead.

$$\begin{split} |I_{2}| &= |I_{4}| \leq \frac{1}{2\pi} \int_{-R+iT}^{\sigma_{0}+iT} \left| \frac{L'(s,\chi)}{L(s,\chi)} \right| \cdot |\tilde{\phi}(s)| \, ds \\ &= \frac{1}{2\pi} \int_{-R}^{-1} \left| \frac{L'(\sigma+iT,\chi)}{L(\sigma+iT,\chi)} \right| \cdot |\tilde{\phi}(\sigma+iT)| \, d\sigma + \frac{1}{2\pi} \int_{-1}^{\sigma_{0}} \left| \frac{L'(\sigma+iT,\chi)}{L(\sigma+iT,\chi)} \right| \cdot |\tilde{\phi}(\sigma+iT)| \, d\sigma \\ &= \int_{-R}^{-1} + O(\log^{2}(qT)) \int_{-1}^{\sigma_{0}} C_{1}(-1,\sigma_{0})|T|^{-1}x^{\sigma-1}(||\phi||_{L^{1}} + x||D\phi||_{L^{1}}) \, d\sigma \\ &= \int_{-R}^{-1} + O_{\phi} \left(\frac{\log^{2}(qT)}{T} \right) \int_{-1}^{\sigma_{0}} x^{\sigma-1} \, d\sigma \\ &= O(\log(qR)) \int_{-R}^{-1} |\tilde{\phi}(\sigma+iT)| \, d\sigma + O_{\phi} \left(\frac{x^{\sigma_{0}-1}\log(x)\log^{2}(qT)}{T} \right) \\ &= O_{\phi}(\log(qR)) \int_{-R}^{-1} C_{1}(-R,-1)|T|^{-1}x^{\sigma} \, d\sigma + O_{\phi} \left(\frac{x^{\sigma_{0}-1}\log(x)\log^{2}(qT)}{T} \right) \\ &= O_{\phi} \left(\frac{\log(q)C'_{R}}{T} \right) \int_{-R}^{-1} x^{\sigma} \, d\sigma + O_{\phi} \left(\frac{x^{\sigma_{0}-1}\log(x)\log^{2}(qT)}{T} \right) \\ &= O_{\phi} \left(\frac{\log(q)\log(x)x^{-R}C'_{R}}{T} \right) + O_{\phi} \left(\frac{x^{\sigma_{0}-1}\log(x)\log^{2}(qT)}{T} \right). \end{split}$$

Where again C'_R is at most polynomial in R, and hence is controlled by x^{-R} . Both of these terms are seen to go to 0 as $T \to \infty$. Drawing together equation (2.2) and these last two estimates, we have

$$|I_1| = -\sum_{|\rho| \le T} \tilde{\phi}(\rho) - \sum_{m=0}^{R/2} \tilde{\phi}(\mathfrak{a} - 2m) + O_{\phi}(C_R x^{-R+1} \log(q)) + O_{\phi}\left(\frac{\log(q)\log(x)x^{-R}C_R'}{T}\right) + O_{\phi}\left(\frac{x^{\sigma_0 - 1}\log(x)\log^2(qT)}{T}\right),$$

where the first O term is as $R\to\infty$ and the second two are as $T\to\infty.$ Letting R go to infinity gives

$$|I_1| = -\sum_{|\rho| \le T} \tilde{\phi}(\rho) - \sum_{m=0}^{\infty} \tilde{\phi}(\mathfrak{a} - 2m) + O_{\phi}\left(\frac{x^{\sigma_0 - 1}\log(x)\log^2(qT)}{T}\right).$$
(2.3)

Finally, we have that

Lemma 2.1.7. If $\sigma_0 > 1$, ϕ as above,

$$|\psi_{\phi}(x,\chi) - I_1| = O_{\phi,x}\left(\frac{1}{T}\right)$$

Proof:

$$|\psi_{\phi}(x,\chi) - I_1| \le \frac{1}{\pi} \int_{\sigma_0 + iT}^{\sigma_0 + i\infty} \left| \frac{L'(s,\chi)}{L(s,\chi)} \right| \cdot |\tilde{\phi}(s)| \, ds$$

For $\sigma_0 > 1$, we have that $|L'(\sigma_0 + it)| \leq \zeta'(\sigma_0) = O(1)$. In the same range we have the absolutely convergent Euler product

$$\begin{aligned} |\log(L(s,\chi))| &= \left| \log \prod_{p} \frac{1}{1-\chi(p)p^{-s}} \right| \\ &\leq \sum_{p} |\log(1-\chi(p)p^{-s})| \\ &= \sum_{p} \left| \sum_{n=1}^{\infty} \frac{(\chi(p)p^{-s})^{n}}{n} \right| \\ &\leq \sum_{p} \sum_{n=1}^{\infty} \frac{p^{-\sigma n}}{n} \\ &\leq \sum_{p} \frac{1}{p^{-\sigma}-1} = O(1). \end{aligned}$$

So $\log(L(s,\chi))$ remains bounded independent of t, hence $1/L(s,\chi)$ also remains bounded as $t \to \pm \infty$. We know from lemma (2.1.2) (Mellin Decay) that $\tilde{\phi}$ has polynomial decay in vertical strips, hence the result. \Box . Combining lemma (2.1.7) with equation (2.3), we have

$$\psi_{\phi}(x,\chi) = |I_1| + O_{\phi,x}(1/\log(T)) = -\sum_{|\rho| \le T} \tilde{\phi}(\rho) - \sum_{m=0}^{\infty} \tilde{\phi}(\mathfrak{a}-2m) + O_{\phi}\left(\frac{x^{\sigma_0-1}\log(x)\log^2(qT)}{T}\right),$$

and letting $T \to \infty$ yields the proposition. \Box .

and letting $T \to \infty$ yields the proposition.

Now that we have the explicit formula for a smooth cutoff function ϕ in terms of its Mellin transform, we must bound its terms to achieve the lemma

$$|\psi(x,\chi)| = O(xe^{-C(M)\sqrt{\log(x)}})$$

which we originally set out to prove. We need some significant results about the nontrivial zeros of $L(s, \chi)$ before we pick ϕ and evaluate the entire expression in section (2.5).

2.2 $N(T, \chi)$

In this section we establish an asymptotic estimate for the number of zeros $N(T,\chi)$ of $L(s,\chi)$ with imaginary part |t| < T in the critical strip. Throughout we assume that χ is not the trivial character unless otherwise noted. As one might expect, we apply the argument principle from complex analysis. We pick the contour $R = R_1 \cup R_2 \cup R_3 \cup R_4$ to be

$$R_{1} = \left[\frac{5}{2} - iT, \frac{5}{2} + iT\right]$$

$$R_{2} = \left[\frac{5}{2} + iT, -\frac{3}{2} + iT\right]$$

$$R_{3} = \left[-\frac{3}{2} + iT, -\frac{3}{2} - iT\right]$$

$$R_{4} = \left[-\frac{3}{2} - iT, \frac{5}{2} - iT\right].$$

Rather than work with $L(s, \chi)$ itself, it is more appropriate to work with the variant

$$\xi(s,\chi) = \left(\frac{q}{\pi}\right)^{\frac{1}{2}s + \frac{1}{2}\mathfrak{a}} \Gamma(\frac{1}{2}s + \frac{1}{2}\mathfrak{a})L(s,\chi)$$
(2.5)

because of the simplified functional equation

$$\xi(1-s,\overline{\chi}) = \frac{i^{\mathfrak{a}}q^{\frac{1}{2}}}{\tau(\chi)}\xi(s,\chi).$$

Recall that $\tau(\chi)$ is the Gauss sum. Recall the lemma of chapter 1 gives $|\tau(\chi)| = q^{\frac{1}{2}}$ for primitive χ , hence the multiplying factor in the above equation has absolute value 1. In the definition of $\xi(s,\chi)$, the poles of Γ cancel with the trivial zeros of $L(s,\chi)$, hence, the only possible zeros are those in the critical strip. We have

$$N(T,\chi) = \frac{1}{2\pi i} \int_{R} \frac{\xi'(s,\chi)}{\xi(s,\chi)} \, ds = \frac{1}{2\pi} \left(\text{Im} \log(\xi(s,\chi)) \right)_{R}.$$
 (2.6)

Because of the symmetry in the functional equation for $\xi(s, \chi)$,

$$\operatorname{Im}\log(\xi(\sigma+it,\chi)) = \operatorname{Im}\log(\overline{\xi(1-\sigma-it,\chi)}) + c,$$

we have that the left hand side and right hand side of R have the same contribution to (2.6). We evaluate each term of $\xi(s,\chi)$ over the half-contour $H = [1/2 - iT, 3/2 - iT] \cup [3/2 - iT, 3/2 + iT] \cup [3/2 + iT, 1/2 + iT]$. We have

$$\left(\operatorname{Im}\log\left(\frac{q}{\pi}\right)^{\frac{1}{2}s+\frac{1}{2}\mathfrak{a}}\right)_{H} = T\log\left(\frac{q}{\pi}\right)$$

for the first term. Note our notation for evaluation along H. For the next term of $\xi(s, \chi)$, we evoke Stirling's formula for the gamma function

Lemma 2.2.1 (Stirling's Approximation). If Re(s) > 0,

$$\log(\Gamma(s)) = s \log(s) - s + O(\log(s)).$$

Proof:

See Whittaker and Watson, Chapter 12 [WW].

This gives us

$$\operatorname{Im}\log\Gamma\left(\frac{1}{2}s+\frac{1}{2}\mathfrak{a}\right)_{H}=T\log\left(\frac{T}{2}\right)-T+O(\log(T)).$$

It remains then to evaluate $\operatorname{Im} \log(L(s,\chi))$ along *H*. For this, we will need some basic results about $L(s,\chi)$.

2.2.1 Hadamard Product for $\xi(s, \chi)$

We know from undergraduate complex analysis that entire functions are determined (up to $e^{f(z)}$) by their zeros. This notion is made precise via the Hadamard product representation of a function. Recall that the order of growth ρ_0 of an entire function g is defined to be the infimum over all ρ such that

$$|g(z)| \le A e^{B|z|^p}.$$

Then we have

 \Box .

Lemma 2.2.2 (Hadamard Product). Suppose that f is an entire function and has growth order ρ_0 . Let k bet the integer so that $k \leq \rho_0 < k + 1$. If a_1, a_2, \ldots denote the (non-zero) zeros of f, then

$$f(z) = e^{P(z)} z^m \prod_{n=1}^{\infty} E_k\left(\frac{z}{a_n}\right),$$

where P is a polynomial of degree $\leq k$, m is the order of the zero of f at z = 0, and $E_k(z)$ are the canonical factors, defined

$$E_k(z) = (1-z)e^{z+z^2/2+\dots+z^k/k}$$

Proof:

See Stein, Complex Analysis, 2003 [St1].

We can apply this to $\xi(s, \chi)$ because it has no poles or other singularities, and is hence an entire function. First, we compute its growth order, so we need estimates on each part of factor of $\xi(s, \chi)$. By summation by parts, we have that

$$L(s,\chi) = s \int_{1}^{\infty} \frac{S(x)}{x^{s+1}} dx, \quad \text{where } S(x) = \sum_{n \le x} \chi(n),$$

valid for $\operatorname{Re}(s) > 0$. If χ is not the trivial character, we have $|S(x)| \leq q$, and

$$|L(s,\chi)| \le 2q|s| \quad \text{for } \operatorname{Re}(s) \ge 1/2.$$

Combining the above and estimates on Γ from Stirling's formula (lemma (2.2.1)), we have

$$\begin{aligned} |\xi(s,\chi)| &\leq 2q^{\frac{1}{2}\operatorname{Re}(s)+\frac{3}{2}}|s||\Gamma(\frac{1}{2}s+\frac{1}{2}\mathfrak{a})| \\ &< q^{\frac{1}{2}\operatorname{Re}(s)+\frac{3}{2}}e^{C|s|\log|s|} \end{aligned}$$

when $\operatorname{Re}(s) \geq 1/2$. The functional equation gives an identical bound when instead $\operatorname{Re}(s) \leq 1/2$. Therefore we have that the order of growth of $\xi(s, \chi)$ is 1. We have then by lemma (2.2.2) that

$$\xi(s,\chi) = e^{A+Bs} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{\frac{s}{\rho}}, \qquad (2.7)$$

where A and B depend on χ . We now use this and (2.5) to compute a formula for L'/L. Taking the logarithmic derivatives of (2.5) and (2.7), we have

$$\frac{L'(s,\chi)}{L(s,\chi)} = -\frac{1}{2}\log\frac{q}{\pi} - \frac{1}{2}\frac{\Gamma'(\frac{1}{2}s + \frac{1}{2}\mathfrak{a})}{\Gamma(\frac{1}{2}s + \frac{1}{2}\mathfrak{a})} + B(\chi) + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right).$$
 (2.8)

 \Box .

2.2.2 Several Estimates for L'/L

We take the negative real part of (2.8), and apply Stirling's approximation again (2.2.1), we have

$$\begin{aligned} -\operatorname{Re}\left(\frac{L'(s,\chi)}{L(s,\chi)}\right) &= \frac{1}{2}\log\frac{q}{\pi} + \frac{1}{2}\operatorname{Re}\left(\frac{\Gamma'(\frac{1}{2}s + \frac{1}{2}\mathfrak{a})}{\Gamma(\frac{1}{2}s + \frac{1}{2}\mathfrak{a})}\right) - \operatorname{Re}(B(\chi)) - \operatorname{Re}\left(\sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right)\right) \\ &< c(\log(q) + \log(t+2)) - \operatorname{Re}(B(\chi)) - \operatorname{Re}\left(\sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right)\right), \end{aligned}$$

where the approximation of Γ is valid in the region $t \geq 2$ and $1 \leq \sigma \leq 2$. We can further simplify this if we analyze the constant $B(\chi)$. We have from logarithmic differentiation of (2.7) and the functional equation,

$$\begin{split} B(\chi) &= \frac{\xi'(0,\chi)}{\xi(0,\chi)} &= -\frac{\xi'(1,\overline{\chi})}{\xi(1,\overline{\chi})} \\ &= -B(\overline{\chi}) - \sum_{\rho} \left(\frac{1}{1-\overline{\rho}} + \frac{1}{\overline{\rho}}\right). \end{split}$$

Because $B(\overline{\chi}) = \overline{B(\chi)}$, we have that

$$2\operatorname{Re}(B(\chi)) = \sum_{\rho} \left(\operatorname{Re}\left(\frac{1}{1-\overline{\rho}}\right) + \operatorname{Re}\left(\frac{1}{\overline{\rho}}\right) \right).$$

By the functional equation again, we may write

$$\operatorname{Re}(B(\chi)) = -\frac{1}{2} \sum_{\rho} \left(\frac{1}{\rho} + \frac{1}{\overline{\rho}}\right) = -\sum_{\rho} \operatorname{Re}\left(\frac{1}{\rho}\right).$$
(2.9)

We use (2.9) to remove $B(\chi)$ from our estimate of L'/L above, obtaining

$$-\operatorname{Re}\left(\frac{L'(s,\chi)}{L(s,\chi)}\right) < c(\log(q) + \log(t+2)) - \sum_{\rho} \operatorname{Re}\left(\frac{1}{s-\rho}\right).$$
(2.10)

This form will be quite useful in the next section when we find zero-free regions of $L(s, \chi)$ for real χ . It is also useful for the remainder of the present section. By the proof of lemma (2.1.7), we know that |L'/L| is bounded for $\operatorname{Re}(s) > 1$, thus we have that

$$\sum_{\rho} \operatorname{Re}\left(\frac{1}{s-\rho}\right) < c(\log(q) + \log(t+2)).$$

Taking $\rho = \beta + i\gamma$ and s = 2 + iT say, we compute

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

and hence find that

$$\sum_{\rho} \frac{1}{1 + (t - \gamma)^2} = O(\log(q(|t| + 2))).$$
(2.11)

The above formula (2.11) has two important consequences.

Observation 2.2.3. The number of zeros with $T - 1 < \gamma < T + 1$ is

$$O(\log(q(|t|+2)))$$

Observation 2.2.4. The sum

$$\sum_{\substack{\rho=\beta+i\gamma\\\gamma\notin(T-1,T+1)}}\frac{1}{(T-\gamma)^2} = O(\log(q(|t|+2))).$$

We now restrict to s with $-1 \leq \text{Re}(s) \leq 2$. Using the partial fraction expansion (2.8) at both s and 2 + it, and subtracting the two yields

$$\frac{L'(s,\chi)}{L(s,\chi)} = O(1) + \sum_{\rho} \left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho} \right).$$
(2.12)

In the spirit of the above two observations, we split this into the regions where the zeros ρ have either $|\gamma - t| \ge 1$ or $|\gamma - t| < 1$. For the first of these two cases we have

$$\sum_{\substack{\rho=\beta+i\gamma\\|\gamma-t|\geq 1}} \left|\frac{1}{s-\rho} - \frac{1}{s+it-\rho}\right| = \sum_{\substack{\rho=\beta+i\gamma\\|\gamma-t|\geq 1}} \frac{2-\sigma}{|(s-\rho)(2+it-\rho)|} \le \sum_{\substack{\rho=\beta+i\gamma\\|\gamma-t|\geq 1}} \frac{3}{|\gamma-t|^2} = O(\log(q(|t|+2)))$$

by observation (2.2.4). Observe that for terms with $|\gamma - t| < 1$, we have $|2 + it - \rho| \ge 1$, thus we have reduced our estimate (2.12)

$$\frac{L'(s,\chi)}{L(s,\chi)} = O(\log(q(|t|+2))) + \sum_{\substack{\rho=\beta+i\gamma\\|\gamma-t|<1}} \left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) \\
\leq O(\log(q(|t|+2))) + \sum_{\substack{\rho=\beta+i\gamma\\|\gamma-t|<1}} \frac{1}{s-\rho} + \sum_{\substack{\rho=\beta+i\gamma\\|\gamma-t|<1}} 1 \\
= O(\log(q(|t|+2))) + \sum_{\substack{\rho=\beta+i\gamma\\|\gamma-t|<1}} \frac{1}{s-\rho}$$
(2.13)

where the last line follows from observation (2.2.3).

By this point the reader may have lost track of what we were originally trying to prove. Referring back to the beginning of the section, we were looking for an appropriate bound on $\text{Im} \log(L(s,\chi))$ in the application of the argument principle for finding $N(T,\chi)$. We now have such an appropriate approximation in (2.13). Thus,

$$(\operatorname{Im}\log(L(s,\chi)))_{H} = O(1) - \int_{\frac{1}{2} + iT}^{2+iT} \operatorname{Im}\left(\frac{L'(s,\chi)}{L(s,\chi)}\right) \, ds - \int_{\frac{1}{2} - iT}^{2-iT} \operatorname{Im}\left(\frac{L'(s,\chi)}{L(s,\chi)}\right) \, ds$$

where the O(1) comes from the integral along [2 - iT, 2 + iT]. Then

$$\left(\mathrm{Im}\log(L(s,\chi))\right)_{H} = O\left(\log(q(|t|+2))\right) + \sum_{\substack{\rho = \beta + i\gamma \\ |\gamma - t| < 1}} \left(\int_{\frac{1}{2} + iT}^{2+iT} \mathrm{Im}\left(\frac{1}{s-\rho}\right) \, ds + \int_{\frac{1}{2} - iT}^{2-iT} \mathrm{Im}\left(\frac{1}{s-\rho}\right) \, ds\right) + O\left(\log(q(|t|+2))\right) + O\left$$

By application of the argument principle again to this simpler function, we have that both integrals are bounded by π . By observation (2.2.3), the number of terms in the sum is $O(\log(q(|t|+2)))$, and hence we finally have

$$(\operatorname{Im}\log(L(s,\chi)))_{H} = O(\log(q(|t|+2))).$$
(2.14)

Hence the contribution to $N(T, \chi)$ from $L(s, \chi)$ is small compared to that of $\left(\frac{q}{\pi}\right)^{s/2+\mathfrak{a}/2}$ and Γ . Drawing these results together and recalling that H is only half of the full contour R, we have that

$$N(T,\chi) = \frac{T}{\pi} \log\left(\frac{qT}{2\pi}\right) - \frac{T}{\pi} + O(\log(qT)).$$
(2.15)

2.3 Zero-Free Regions for $L(s, \chi)$

The result of the above section is one important estimate which allows us to bound the side of $\tilde{\phi}(\rho)$ in the explicit formula. However, we also need a bound on how close zeros may be to the line $\operatorname{Re}(s) = 1$ to estimate $\tilde{\phi}(\rho)$. In this section, the distinction between real and complex characters becomes significant, and we will see that the possibility of a zero on the real line close to s = 1 (a Siegel zero) cannot be ruled out, despite being able to otherwise establish a zero-free region for $L(s, \chi)$ of width $O(1/\log(t))$ elsewhere. Throughout we assume that χ is not the trivial character unless otherwise noted.

2.3.1 Regions for Complex χ

We begin with the trivial but useful inequality

$$0 \leq 2(1 + \cos(\theta))^2$$

= 3 + 4 \cos(\theta) + \cos(2\theta)

Next, if $s = \sigma + it$ and $\sigma > 1$, compute from the Dirichlet series

$$-\operatorname{Re}\left(\frac{L'(s,\chi)}{L(s,\chi)}\right) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\sigma}} \operatorname{Re}(\chi(n)e^{-it\log(n)}).$$

We have that for (n,q) = 1, the modulus $|\chi(n)e^{-it\log(n)}| = 1$, hence

$$\chi(n)e^{-it\log(n)} = e^{i\theta}$$

for some θ , and

$$\operatorname{Re}(\chi(n)e^{-it\log(n)}) = \cos(\theta).$$

Squaring or multiplying by $e^{-i\theta}$ the first of these two likewise yields

$$\operatorname{Re}(\chi^2(n)e^{-2it\log(n)}) = \cos(2\theta)$$

and

$$\operatorname{Re}(\chi_0) = \cos(0) = 1.$$

Hence, by our trivial trigonometric inequality above,

$$-3\frac{L'(\sigma,\chi_0)}{L(\sigma,\chi_0)} - 4\operatorname{Re}\left(\frac{L'(\sigma+it,\chi)}{L(\sigma+it,\chi)}\right) - \operatorname{Re}\left(\frac{L'(\sigma+2it,\chi^2)}{L(\sigma+2it,\chi^2)}\right) \ge 0.$$
(2.16)

We have $\chi^2 = \chi_0$ if and only if χ is a real character. This creates difficulties for us later and leads to the possibility of a Siegel zero. First, suppose that χ is complex and primitive. The first term in (2.16) resembles the zeta function,

$$-\frac{L'(\sigma,\chi_0)}{L(\sigma,\chi_0)} = \sum_{n=1}^{\infty} \frac{\chi_0(n)}{n^{\sigma}} \le \frac{\zeta'(\sigma)}{\zeta(\sigma)} = \frac{1}{\sigma-1} + O(1)$$

close to, but on the right of $\sigma = 1$. For the next two terms in (2.16), we appeal to the estimate (2.10) of the previous section, which states

$$-\operatorname{Re}\left(\frac{L'(s,\chi)}{L(s,\chi)}\right) < c(\log(q(t+2)) - \sum_{\rho} \operatorname{Re}\left(\frac{1}{s-\rho}\right).$$
(2.17)

The terms in the sum are

_

$$\operatorname{Re}\left(\frac{1}{s-\rho}\right) = \frac{\sigma-\beta}{|s-\rho|^2} \ge 0,$$

therefore are able to omit some or all of the terms in the sum and the estimate will remain valid. Even though we assumed that χ was primitive above, χ^2 may

not be, although this does not affect the argument as we will show with a quick estimate. If χ_1 induces χ^2 , then

$$L(s,\chi^2) = L(s,\chi_1) \prod_{p|q} (1-\chi_1(p)p^{-s}), \qquad (2.18)$$

and hence

$$\begin{aligned} \left| \frac{L'(s,\chi^2)}{L(s,\chi^2)} - \frac{L'(s,\chi_1)}{L(s,\chi_1)} \right| &\leq \sum_{p|q} \frac{\log(p)p^{-\sigma}}{1 - p^{-\sigma}} \\ &\leq \sum_{p|q} \log(p) \leq \log(q). \end{aligned}$$

This is within the bound in (2.17), so we need not concern ourselves with the possibility of χ^2 being imprimitive. Given our above remarks about the positivity of the terms in (2.17), for χ^2 we choose to omit the entire sum, hence

$$-\operatorname{Re}\left(\frac{L'(\sigma+2it,\chi^2)}{L(\sigma+2it,\chi^2)}\right) = O(\log(q(t+2))).$$

Lastly, we bound the middle term. Given some zero $\rho = \beta + i\gamma$, we take $t = \gamma$. In the sum in (2.17) we keep only the term corresponding to this zero,

$$-\operatorname{Re}\left(\frac{L'(\sigma+it,\chi)}{L(\sigma+it,\chi)}\right) < -\frac{1}{\sigma-\beta} + c\log(q(t+2)).$$

Combining the estimates for all three terms of (2.16) we see that for the zero ρ we picked above,

$$\frac{4}{\sigma-\beta} < \frac{3}{\sigma-1} + c_2 \log(q(t+2)).$$

We have chosen t but are still free to choose σ , and now take $\sigma = 1 + c_3/(\log(q(t+2)))$ for an appropriately chosen c_3 . Applying this to the above equation gives

$$\beta < 1 + \frac{c_3}{\log(q(t+2))} - \frac{4c_3}{3(\log(q(t+2))) + c_3c_4} < 1 - \frac{c_5}{\log(q(t+2))}$$

for appropriately chosen c_5 . We have this for any complex primitive character. But if χ is imprimitive, the only extra zeros arise from euler factors as in (2.18), and accordingly have $\beta = 0$. Thus we have shown that

Observation 2.3.1. There exists a positive absolute constant c_5 such that, if χ is a complex character mod q, any zero $\beta + i\gamma$ of $L(s, \chi)$ satisfies

$$\beta < 1 - \frac{c_5}{\log(q(|\gamma| + 2))}.$$
(2.19)

2.3.2 Regions for Real χ

Now we proceed to real characters, which, as is common in the theory of Dirichlet *L*-functions, are much more difficult. Our estimates for $L(\sigma, \chi_0)$ and $L(\sigma + it, \chi)$ from the previous section still hold. However, we now have that $\chi^2 = \chi_0$ so (2.17) is no longer valid. Because $L(s, \chi_0)$ differs from $\zeta(s)$ only by a number of euler factors which divide q, we have again that

$$\left|\frac{L'(s,\chi_0)}{L(s,\chi_0)} - \frac{\zeta'(s)}{\zeta(s)}\right| \le \log(q)$$

when $\operatorname{Re}(s) > 1$. We have a similar formula to (2.17) for the zeta function (we omit its proof, but it is along the same lines)

$$-\operatorname{Re}\left(\frac{\zeta'(s)}{\zeta(s)}\right) < \operatorname{Re}\frac{1}{s-1} + c_5\log(t+2)$$

Thus

$$-\operatorname{Re}\left(\frac{L'(\sigma+2it,\chi^2)}{L(\sigma+2it,\chi^2)}\right) < \operatorname{Re}\left(\frac{1}{\sigma-1+2it}\right) + c_6(\log(q(t+2))).$$

Since this is about as good as we can hope for, we use it along with our estimates of the other terms in (2.16) and obtain

$$\frac{4}{\sigma-\beta} < \frac{3}{\sigma-1} + \operatorname{Re}\left(\frac{1}{\sigma-1+2it}\right) + c_7(\log(q(t+2)))$$

where we take $t = \gamma$ and $\sigma = 1 + \delta/(\log(q(t+2)))$ as in the case for complex χ .

However, we will now need an extra supposition to make our calculations work, and we *postulate* in addition that $\gamma \geq \delta/(\log(q(t+2)))$. We obtain

$$\frac{4}{\sigma-\beta} < \log(q(|\gamma|+2)) \left(\frac{3}{\delta} + \frac{1}{5\delta} + c_7\right)$$

and

$$\beta < 1 - \left(\frac{\delta}{\log(q(|\gamma|+2))}\right) \left(\frac{4 - 5c_7\delta}{16 + 5c_7\delta}\right),$$

where our postulate was necessary to obtain the second term in the first equation. In order for the zero-free region we have just constructed to be *inside* the critical strip, we must have that the second parentheses is a strictly positive quantity. We can obtain this if let δ be sufficiently small compared to c_7 , say less than c_8 (which depends of course on previous constants). Note that if we had not assumed the extra postulate, we would not have had the 4 in this equation, and the second parentheses would have been strictly negative, giving a zero free region *outside* the critical strip, which is clearly useless. Therefore, our results are subject to the condition $\gamma \geq \delta/\log(q(|\gamma| + 2))$, or the slightly stronger condition $\gamma \geq \delta/\log(q)$. We have thus proved **Observation 2.3.2.** There exists a positive absolute constant c_8 such that, if $0 < \delta < c_8$, and χ a real nontrivial character mod q, then any zero $\rho = \beta + i\gamma$ of $L(s,\chi)$ for which

$$|\gamma| \geq \frac{\delta}{\log(q)}$$

satisfies

$$\beta < 1 - \frac{\delta}{5\log(q(|\gamma|+2))}.$$

We have not addressed imprimitive characters, however they can be dealt with by the same process as in the case of complex characters.

2.3.3 Siegel Zeros

It remains to address the region of the critical strip $|\text{Im}(s)| < \delta/\log(q)$, where additional zeros may lie. We will show that all but one of these zeros are far from s = 1, in the sense that it has $\beta < 1 - \epsilon/\log(q)$ for suitable positive ϵ . Then there exists at most one zero with $\beta \ge 1 - \epsilon/\log(q)$, hence it must be real (because otherwise we would also have it's conjugate). This possible zero is called a Siegel zero.

We assume there are two zeros with $|\text{Im}(\rho)| < \delta/\log(q)$, and show that they must be away from s = 1. We assume that these zeros are complex and come in a pair $\beta \pm i\gamma$. If one assumes there are two real zeros, the proof is similar and we omit it. Begin by returning to our favourite estimate (2.10) or (2.17),

$$-\frac{L'(\sigma,\chi)}{L(\sigma,\chi)} < c_9 \log(q) - \sum_{\rho} \frac{1}{\sigma - \rho},$$

valid for $s = \sigma > 1$. Recall that the terms in the sum are positive, so that we can drop arbitrarily many of them and retain a true statement. We drop all terms in the sum except for the two zeros whose existence we have assumed above. We now bound either side of this, obtaining

$$-\frac{1}{\sigma-1}-c_{10} < \frac{\zeta'(\sigma)}{\zeta(\sigma)} \le -\frac{L'(\sigma,\chi)}{L(\sigma,\chi)} < c_9 \log(q) - \sum_{\rho} \frac{1}{\sigma-\rho} < c_9 \log(q) - \frac{2(\sigma-\beta)}{(\sigma-\beta)^2 + \gamma^2}.$$
(2.20)

We now take $\sigma = 1 + 2\epsilon / \log(q)$, giving

$$\frac{1}{2}(\sigma - \beta) > \frac{1}{2}(\sigma - 1) = \frac{\epsilon}{\log(q)} > |\gamma|$$

Combining the last two equations and simplifying, we obtain

$$-\frac{1}{\sigma - 1} < c_{11}\log(q) - \frac{8}{5(\sigma - \beta)}$$

from which it follows

$$\beta < 1 - \frac{6 - 20c_{11}\epsilon}{5\log(q)(1 + 2c_{11}\epsilon)}$$

Hence if ϵ is chosen sufficiently small, we have for a new ϵ' positive,

$$\beta < 1 - \frac{\epsilon'}{\log(q)}.$$

Therefore we have

Observation 2.3.3. There exists a positive absolute constant c_{12} such that, if $0 < \epsilon < c_{12}$, the only possible zero of $L(s, \chi)$ for a real, nontrivial χ , and satisfying both

$$|\gamma| < \frac{\epsilon}{\log(q)}, \quad and \quad \beta > 1 - \frac{\epsilon}{\log(q)}$$

is a single simple real zero. Such a zero is called a Siegel zero.

2.4 Siegel's Theorem

We have only one more zero to locate. All other zeros were bounded appropriately far away from the line Re(s) = 1 in observations (2.3.1) and (2.3.2), however, we still need to address the possible Siegel zero. An appropriate bound of this zero away from s = 1 is given by Siegel's theorem.

Proposition 2.4.1 (Siegel's Theorem). For any $\varepsilon > 0$, there exists a positive number $c(\varepsilon)$ such that, if χ is a real primitive character mod q, then

$$L(1,\chi) > \frac{c(\varepsilon)}{q^{\varepsilon}}.$$

The result about the location of Siegel zeros which we are interested in and want to use follows from this as a corollary.

Corollary 2.4.2. For any $\varepsilon > 0$, there exists a positive number $c'(\varepsilon)$ such that, if χ is a real primitive nontrivial character mod q, then $L(s, \chi) \neq 0$ for

$$s > 1 - \frac{c'(\varepsilon)}{q^{\varepsilon}}.$$

We prove the corollary first, and then proceed to the proof of the original theorem second.

Proof (Corollary):

Suppose we have a zero β which is close to s = 1 and hence contradicts our corollary. We want to apply the mean value theorem, so we have

$$L(1,\chi) = L(1,\chi) - L(\beta,\chi) = (1-\beta)L'(c,\chi)$$
(2.21)

for some $c \in [\beta, 1]$. We have that

$$c \in \left[1 - \frac{c'(\varepsilon)}{q^{\varepsilon}}, 1\right] \subseteq \left[1 - \frac{1}{\log(q)}, 1\right]$$

for large enough q. We now derive a bound for L' in this slightly larger interval by a standard main term and tail argument. We have

$$L'(\sigma,\chi) = -\sum_{n=1}^\infty \frac{\log(n)\chi(n)}{n^\sigma}$$

for $\sigma > 0$. We take σ real and lying in the slightly larger region above. Split the sum at n = q. We first take the main term. We have that $n \leq q$ and $1 - \sigma \leq 1/\log(q)$, hence

$$1 \ge \frac{\log(n)}{\log(q)} \ge \log(n)(1-\sigma),$$

which implies

$$n^{-\sigma} = e^{-\sigma \log(n)} \le e^{1 - \log(n)} = \frac{e}{n}.$$

Hence, we have for the main term that

$$\left|\sum_{n=1}^{q} \frac{\log(n)\chi(n)}{n^{\sigma}}\right| \le e \sum_{n=1}^{q} \frac{\log(n)}{n} = O(\log^2(q)).$$

For the tail we have n > q and apply partial summation,

$$\left|\sum_{n=q+1}^{\infty} \frac{\log(n)\chi(n)}{n^{\sigma}}\right| \le \log(q)q^{-\sigma} \max_{M} \left|\sum_{m=q+1}^{M} \chi(m)\right| = O(\log(q)).$$

Hence we have that for $\sigma \in [1 - \frac{c'(\varepsilon)}{q^{\varepsilon}}, 1]$,

$$L'(\sigma, \chi) = O(\log^2(q)).$$

We use these two facts to evaluate the right side of (2.21), giving

$$L(1,\chi) = (1-\beta)L'(c,\chi) \le \frac{c'(\varepsilon)\log^2(q)}{q^{\varepsilon}}.$$

This is a contradiction with Siegel's theorem upon picking ε correctly. Hence such a zero cannot exist.

2.4.1 Goldfeld's Proof

We now proceed directly to the proof of Siegel's theorem. We follow Goldfeld's 1974 proof [Gol] which employs the Dedekind zeta function of a biquadratic field.

Proof (Siegel's Theorem):

Let K be a biquadratic extension of \mathbb{Q} . Let $f(s) = \zeta_K(s) = \zeta(s)L(s,\chi_1)L(s,\chi_2)L(s,\chi_1\chi_2)$. It has a pole at s = 1 due to the Riemann zeta function with residue $\lambda = L(1,\chi_1)L(1,\chi_2)L(1,\chi_1\chi_2)$. We pull out a quick

Lemma 2.4.3. For every $\varepsilon > 0$ there exists some $\chi_1 \mod q_1$ and a β satisfying $1 - \varepsilon < \beta < 1$ such that $f(\beta) \leq 0$ independent of what $\chi_2 \mod q_2$ is.

Proof of lemma:

We have two cases. First suppose that there are no zeros in $[1 - \varepsilon, 1]$ for any $L(s, \chi)$. Then $L(s, \chi_1)$, $L(s, \chi_2)$ and $L(s, \chi_1\chi_2)$ are all positive in this interval but $\zeta(s)$ is negative, so $f(\beta) < 0$. In the other case, such a real zero does exist for some $L(s, \chi_1)$. Let β be such a zero of this L. Then $f(\beta) = 0$ regardless of what χ_2 is.

Now we apply Perron's formula, picking

$$\phi_x(n) = \begin{cases} \frac{(n-x)^4}{24n^4} & \text{if } n \in [1,x] \\ 0 & \text{if } n \in [x,\infty). \end{cases}$$

Note that ϕ is a C^4 cutoff function instead of C^{∞} . This changes our Mellin decay properties lemma, however, does not affect the outcome of our argument, so we do not go into detail. Perron's formula gives

$$1 \ll \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} f(s+\beta) \frac{x^s}{s(s+1)(s+2)(s+3)(s+4)} ds$$
$$= \lambda \frac{x^{1-\beta}}{(1-\beta)(2-\beta)(3-\beta)(4-\beta)(5-\beta)} + \frac{f(\beta)}{4!} + O\left(\frac{(q_1q_2)^{1+\varepsilon}x^{-\beta}}{1-\beta}\right) (2-\beta)(2-\beta)(3-\beta)(4-\beta)(5-\beta)$$

Our integrand has poles at $1 - \beta$, 0, -1, -2, -3, and -4, and the first two terms in the second line represent the residues of these poles. By the lemma, we have that $f(\beta) \leq 0$, and we are free to pick x, so we choose one with $(q_1q_2)^{2+\varepsilon} = O(x)$. We now need a crude bound on $L(1,\chi)$. Recall Dirichlet's class number formula for an imaginary quadratic field,

$$h(-d) = \frac{1}{2\pi} w \sqrt{dL}(1, \chi_d),$$

where we have that w = 2, 4, or 6, and χ_d is the real character given by the Kronecker symbol mod d. Because we know that the class number is always

positive, i.e. $h(d) \ge 1$, the formula gives that $L(1, \chi) = O(d^{-1/2})$. These three facts applied to the above yield

$$1 \ll \lambda \frac{x^{1-\beta}}{1-\beta},$$

where the bound on the *L*-function given by the class number formula was used to estimate that $\lambda \gg 1/q_1q_2$. That is to say, the error term in (2.22) is controlled under the conditions we have set forth.

Using a main term and tail argument similar to (but simpler than) that in the proof of corollary (2.4.1), we also have for any nontrivial $\chi \mod q$ that $L(1,\chi) = O(\log(q))$. Hence we have also that

$$\lambda \ll L(1,\chi_2)\log(q_1)\log(q_1q_2).$$

Pulling our last two equations together, we have

$$L(1,\chi_2) > \frac{c(\chi_1,q_1,\varepsilon)}{q_2^{(2+\varepsilon)(1-\beta)}\log(q_2)}.$$

In actuality, χ_1 and q_1 only depend on ε and hence the constant above only depends on ε . If $0 < \varepsilon < 1$, then $(2 + \varepsilon)(1 - \beta) \leq 3\varepsilon$, and

$$L(1,\chi) > \frac{c(\varepsilon)}{q^{3\varepsilon}\log(q)}$$

The theorem then follows from this for sufficiently large q and upon relabeling ε .

Although Siegel's theorem does give us a nice bound on the size of the L-function at 1, the constants implied are unfortunately *ineffective*. That is, any proof of Siegel's theorem to date (including the one above) does not give a way of assigning a numerical value to the constant $c(\varepsilon)$ for a given value of ε .

2.5 Final estimate for $\psi(x,\chi)$

We how return to the ideas of section (2.1), having collected all of the estimates on zeros of L functions we need. Recall the explicit formula from section (2.1),

$$\sum_{n} \chi(n)\Lambda(n)\phi(n) = -\sum_{|\rho| \le T} \tilde{\phi}(\rho) - \sum_{m=0}^{\infty} \tilde{\phi}(\mathfrak{a} - 2m) + O_{\phi}\left(\frac{x\log(x)\log^2(qT)}{T}\right).$$

whenever ϕ is a C^{∞} function with compact support contained in $(1, \infty)$.

In light of the last section, we pull out the possible Siegel zero from our explicit formula to later treat it separately.

$$\sum_{n} \chi(n)\Lambda(n)\phi(n) = -\frac{x^{\beta_1}}{\beta_1} - \sum_{|\rho| \le T} \tilde{\phi}(\rho) - \sum_{m=0}^{\infty} \tilde{\phi}(\mathfrak{a} - 2m) + O_{\phi}\left(\frac{x\log(x)\log^2(qT)}{T}\right)$$
(2.23)

We now want to pick ϕ , calculate its Mellin transform, and apply the bounds of the previous three sections to the above. In principle, we would like to use a C^{∞} function which closely resembles the characteristic function of the interval [2, X], so that the left hand side of the above is a finite sum without a cutoff function. Because the left hand side is a discrete sum, choosing a cutoff function ϕ so that the non-constant transitions between 0 and 1 lie between integers will be sufficient for obtaining the sum

$$\sum_{n \leq x} \chi(n) \Lambda(n)$$

on the left hand side. We choose the characteristic function of the interval [2-1/(2X), X+1/2], and mollify it using Mellin convolution. The convolution is defined with respect to the group (\mathbb{R}^+, \cdot) , hence

$$(f*g)(v) = \int_0^\infty f(uv^{-1})g(u)\,\frac{du}{u}.$$

Note the analogy with the familiar additive convolution. Just as in the additive case, (f * g) has the smoothness of either f or g, whichever is more differentiable. Again in analogy with the familiar theory of fourier transforms, we have that

$$\widetilde{(f * g)}(s) = \widetilde{f}(s) \cdot \widetilde{g}(s),$$

where \tilde{f} is the Mellin transform of f. We now pick a smooth bump function $\eta_{\frac{1}{X}}(u)$ of width 1/X, centered at 1, and with $||\eta||_{L^1(\mathbb{R}^+, du/u)} = 1$. Therefore, For $u \in (0, 1 - 1/(2X)) \cup (1 + 1/(2X), \infty)$, $\eta_{\frac{1}{X}}(u) = 0$. We now pick our ϕ in the explicit formula,

$$\phi(v) = (\eta_{\frac{1}{X}} * \chi_{[2-1/(2X), X+1/2]})(v) = \int_0^\infty \eta_{\frac{1}{X}}(uv^{-1})\chi_{[2-1/(2X), X+1/2]}(u)\frac{du}{u},$$

where χ_I is the characteristic function of the interval I. Note that by the above comments ϕ is a C^{∞} function. We now check that ϕ has the properties we are interested in. Consider the multiplicative translates of $\eta_{\frac{1}{X}}$, as they appear in the integrand above. We have that

$$\operatorname{supp}_{u}(\eta_{\frac{1}{X}}(uv^{-1})) = \left[v - \frac{v}{2X}, v + \frac{v}{2X}\right].$$

Because du/u is an invariant measure for multiplication, we have that any of these multiplicative translates has $||\eta(uv^{-1})||_{L^1(\mathbb{R}^+, du/u)} = 1$. In addition, given v for which

$$\operatorname{supp}_u(\eta_{\frac{1}{X}}(uv^{-1})) \subseteq \left[2 - \frac{1}{2X}, X + \frac{1}{2}\right],$$

we have that the integral reduces to 1. However, combining the last two displayed equations shows that the set of such v is exactly [2, X]. Furthermore if

$$\operatorname{supp}_u(\eta_{\frac{1}{X}}(uv^{-1})) \cap \left[2 - \frac{1}{2X}, X + \frac{1}{2}\right] = \emptyset$$

then the above integral vanishes. The set of such v is $(0, \frac{4X-1}{2X+1}) \cup (\frac{2X^2+X}{2X-1}, \infty)$. Hence, we have that

Observation 2.5.1. ϕ as defined above is a C^{∞} function with

$$\phi(v) = \begin{cases} 0 & \text{if } v \in (0, \frac{4X-1}{2X+1}) \cup (\frac{2X^2+X}{2X-1}, \infty) \\ 1 & \text{if } v \in [2, X]. \end{cases}$$
(2.24)

and having smooth transitions elsewhere.

We now compute the Mellin Transform of ϕ , which is conveniently enough

$$\widetilde{\phi}(s) = (\eta_{\frac{1}{X}} * \widetilde{\chi_{[2-\frac{1}{2X}, X+\frac{1}{2}]}})(s) = \widetilde{\eta_{\frac{1}{X}}}(s) \cdot \widetilde{\chi_{[2-\frac{1}{2X}, X+\frac{1}{2}]}}(s)$$
$$= \widetilde{\eta_{\frac{1}{X}}}(s) \cdot \left(\frac{(X+\frac{1}{2})^s}{s} - \frac{(2-\frac{1}{2X})^s}{s}\right)$$
$$\leq c\widetilde{\eta_{\frac{1}{X}}}(s) \cdot \left(\frac{X^s}{s} - \frac{2^s}{s}\right).$$
(2.25)

To estimate the Mellin transform of our bump function, we must now appeal to our estimates of section (2.1). Recall from lemma (2.1.2) that the estimate applies for ψ supported in some interval [Y, 2Y], with Y > 1. However, our $\eta_{\frac{1}{x}}$ has very small support, and we may take Y in the statement of the lemma to be, say, 2. If we set $\beta(u) = \eta_{\frac{1}{x}}(u/3)$, then $\beta(u)$ has support contained in [2, 4], and we have that $|\tilde{\beta}(s)| = 3^{\sigma} |\tilde{\eta}_{\frac{1}{x}}(s)|$. So we may apply our Mellin decay properties lemma (2.1.2) with m = 1,

$$|\tilde{\psi}(\sigma+it)| \le C_1(\sigma_1,\sigma_2) \frac{Y^{\sigma-1}}{|t|} (||\psi||_{L^1} + Y||D\psi||_{L^1}),$$

and Y will be an absolute constant. For our bump function $\eta_{\frac{1}{x}}$ we have $||\eta||_{L^1} = O(Y)$ and $||D\eta||_{L^1} = O(1)$. Computing for our specific case we have

$$|\tilde{\eta}_{\frac{1}{x}}(s)| \le C_1(\sigma_1, \sigma_2) \frac{2^{\sigma}}{3^{\sigma}|t|}.$$

Furthermore, for the sum over the nontrivial zeros in the explicit formula, we have $0 \le \sigma \le 1$, so the $\left(\frac{2}{3}\right)^{\sigma}$ factor is bounded. Taking $\rho = \beta + i\gamma$ we have for these zeros

$$|\widetilde{\eta_{\frac{1}{X}}}(\rho)| = O(|\gamma|^{-1}).$$

Drawing these results together, we have for nontrivial zeros ρ

$$\tilde{\phi}(\rho) \le c \frac{1}{|\gamma|} \left(\frac{X^{\rho}}{\rho} - \frac{2^{\rho}}{\rho} \right).$$
(2.26)

For the trivial zeros, our Mellin transform estimates yield

$$\tilde{\phi}(-2m + \mathfrak{a}) = O(X^{-2m + \mathfrak{a}}).$$

Using our computations of the Mellin transform at the zeros of $L(s, \chi)$, we may now fully compute (2.23). We have

$$\left|\sum_{n \le x} \chi(n) \Lambda(n)\right| \ll -\frac{x^{\beta_1}}{\beta_1} - \sum_{|\rho| \le T} \frac{1}{|\gamma|} \left(\frac{x^{\rho}}{\rho} - \frac{2^{\rho}}{\rho}\right) - \sum_{m=0}^{\infty} x^{-2m+\mathfrak{a}} + O_{\phi} \left(\frac{x \log(x) \log^2(qT)}{T}\right).$$
(2.27)

We see immediately that

$$\sum_{|\rho| \le T} \frac{2^{\rho}}{|\gamma|\rho}$$

contributes only a bounded amount as $T \to \infty$, and that

$$\sum_{m=0}^{\infty} x^{-2m+\mathfrak{a}} = O(\log(x)).$$

Hence we absorb these into the error term, and have

$$\left|\sum_{n \le x} \chi(n) \Lambda(n)\right| \ll -\frac{x^{\beta_1}}{\beta_1} - \sum_{|\rho| \le T} \frac{x^{\rho}}{|\gamma|\rho} + O_{\phi}\left(\frac{x \log(x) \log^2(qT)}{T}\right) + O(\log(x)).$$
(2.28)

We are finally in a position to apply the estimates of sections (2.2), (2.3) and (2.4). Recall that from our conclusions at the end of all three subsections of (2.3), we have that all of the non-trivial zeros $\rho = \beta + i\gamma$ satisfy

$$\beta < 1 - \frac{c}{\log(qT)}$$

for some constant c. Hence for the numerator of (2.28) we have

$$|x^{\rho}| = x^{\beta} < x e^{-c \frac{\log(x)}{\log(qT)}}.$$

From section (2.2), we deduced that the number of zeros with imaginary part $|\gamma| \leq T$ was $N(T, \chi) = O(T \log(qT))$. Hence we have that

$$\sum_{|\gamma| < T} \frac{1}{|\gamma^2|} = \int_0^T t^{-2} \, dN(t,\chi) = \frac{N(T,\chi)}{T^2} + 2 \int_0^T \frac{N(t,\chi)}{t^3} \, dt \ll \frac{\log(qT)}{T} + \log(q).$$

Therefore

.

$$\sum_{|\rho| \le T} \frac{x^{\rho}}{|\gamma|\rho} = O(xe^{-c\frac{\log(x)}{\log(qT)}}).$$
(2.29)

Taking this with equation (2.28) we obtain

$$\left|\sum_{n \le x} \chi(n) \Lambda(n)\right| \ll -\frac{x^{\beta_1}}{\beta_1} + O_\phi\left(\frac{x \log(x) \log^2(qT)}{T}\right) + O(x e^{-c \frac{\log(x)}{\log(qT)}}). \quad (2.30)$$

We will apply Siegel's theorem to bound the final term above. However, first we impose a condition on q and pick T in terms of x to combine the error terms in the above. We suppose that

$$q \le \log^M(x)$$

for some positive constant M. In particular, this is stronger than

$$q \le \log^M(x) \le e^{C\sqrt{\log(x)}}.$$

We pick T similarly, setting

.

$$T = e^{C\sqrt{\log(x)}}.$$

Hence, our large error term reduces to

$$\left|\sum_{n \le x} \chi(n) \Lambda(n)\right| \ll -\frac{x^{\beta_1}}{\beta_1} + O_{\phi}(x e^{-C'\sqrt{\log(x)}}).$$
(2.31)

For the final step, we apply Siegel's theorem subject to our constraint on q. We have by the result of the corollary to Siegel's theorem in section (2.4) that for any $\varepsilon > 0$ the exists $c'(\varepsilon)$ such that

$$\beta_1 < 1 - \frac{c'(\varepsilon)}{q^{\varepsilon}}.$$

Hence

$$x^{\beta_1} < x e^{-c'(\varepsilon)\frac{\log(x)}{q^{\varepsilon}}}$$

Recalling our restriction on q from above, we take $\varepsilon=\frac{1}{2M},$ getting $q^{\varepsilon}\leq\sqrt{\log(x)}$ and

$$x^{\beta_1} < x e^{-c^{\prime\prime}(M)\sqrt{\log(x)}},$$

Therefore, we at long last have

Lemma 2.5.2 (First Main Lemma). If χ is a nontrivial character mod q and $q \leq \log^{M}(x)$ for some positive M, then we have

$$|\psi(x,\chi)| = O(xe^{-C(M)\sqrt{\log(x)}})$$
 (2.32)

where C(M) is an ineffective constant which depends only on M.

2.6 Ramanujan Sums

Two lemmas are necessary to bound the major arcs \mathfrak{M} in the proof of Vinogradov's three primes theorem. The first was a bound on $\psi(x, \chi)$ which made up the bulk of this chapter. We now give some information about Ramanujan sums, which is substantially easier than the preceding several sections.

Recall that we have defined

$$c_q(n) = \sum_{\substack{a \mod q \\ (a,q)=1}} e\left(\frac{an}{q}\right)$$

to be the Ramanujan sum. To prove the lemma, we utilize the Dirichlet convolution algebra, and write with fixed \boldsymbol{n}

$$1(q)*c_q(n) = \sum_{d|q} c_{\frac{q}{d}}(n) = \sum_{d|q} \sum_{\substack{a \mod q \\ (a,q)=d}} e\left(\frac{an}{q}\right) = \sum_{a \mod q} e\left(\frac{an}{q}\right) = \begin{cases} q & \text{if } n \equiv 0 \mod q \\ 0 & \text{else} \end{cases}$$

We have that $1*\mu=\mathrm{id}$ in the convolution algebra, so we convolve μ with either side to obtain

$$c_q(n) = \sum_{\substack{d|n\\d|q}} d \cdot \mu\left(\frac{q}{d}\right).$$
(2.33)

From this expression we can see that $c_q(n)$ only depends on q in terms of its divisors, and hence is easily seen to be multiplicative. This proves part (a) of the original lemma.

It is now a simple matter to evaluate $c_q(n)$ on prime powers. Let α be the largest power of p which divides n. Then we have that

$$c_{p^{\beta}}(n) = \sum_{i=0}^{\alpha} p^{i} \mu(p^{\beta-i}) = \begin{cases} p^{\beta} - p^{\beta-1} = \phi(p^{\beta}) & \text{if } \beta \leq \alpha \\ -p^{\alpha} & \text{if } \beta = \alpha + 1 \\ 0 & \beta > \text{if } \alpha + 1 \end{cases}$$

where the last evaluation follows because μ vanishes on proper prime powers. This proves part (b) of our lemma.

Chapter 3

Minor Arcs

In this chapter we investigate sums of the form $\sum \mu(n)e(n\alpha)$ and find that estimates for these depend on the real number α . In particular they depend on how near α is to a rational number with small denominator - i.e. whether α lies in a major or minor arc. We develop this theory with an eye towards proving the third main lemma in chapter 1. We follow Green and Tao [GT] in this exposition. To do this, we establish an alternate formula for the μ function, called Vaughan's identity. This formula allows us to decompose sums of the form $\sum \mu(n)f(n)$ into two parts (type I and type II sums), each with a distinct behaviour. Type I sums tend to be large for periodic f and type II sums large for multiplicative f. We apply bounds on each of these sums to establish the result of the first chapter.

In this chapter we adopt an "inverse approach" in proving statements, whereas in chapter 1 we apply our result in the forward direction. In the case of $f(n) = e(\alpha n)$ above, we have that if α is close to a rational number with large denominator (major arc), then one of the type I or type II sums is large. On the other hand, if α is far from such a rational number (minor arc), the both the type I and type II sums are small. The "inverse approach" means that we will instead show that a type I or type II sum being large implies α lies on a major arc. This approach is easier to apply when one deals with more devious f.

3.1 Vaughan's Identity

Lemma 3.1.1 (Vaughan's Indentity). Let U, V, N be positive integers with $UV \leq N$. Then for $n \leq N$, we have

$$\mu(n) = -\sum_{\substack{bc|n\\b\leq U,\ c\leq V}} \mu(b)\mu(c) + \sum_{\substack{bc|n\\b>U,\ c>V}} \mu(b)\mu(c)$$
(3.1)

Proof:

Observe that in the Dirichlet convolution algebra we have $1 * \mu * \mu = \mu$ Writing out the definition of the Dirichlet convolutions this is

$$\mu(n) = \sum_{bc|n} \mu(b)\mu(c).$$

We now split the summation into four ranges and label the corresponding sums $\sum_1 \dots \sum_4$,

$$\sum_{1} = \sum_{\substack{b \leq U \\ c \leq V}}$$
$$\sum_{2} = \sum_{\substack{b > U \\ c \leq V}}$$
$$\sum_{3} = \sum_{\substack{b \leq U \\ c > V}}$$
$$\sum_{4} = \sum_{\substack{b > U \\ c > V}}.$$

Because of the symmetry in b and c, we see that $\sum_2 = \sum_3$. We also see that $\sum_2 = -\sum_1$, indeed

$$\mu(n) = \sum_{bc|n} \mu(b)\mu(c) = \sum_{c|n} \mu(c) \sum_{b|\frac{n}{c}} \mu(b).$$

But the interior sum here is 0 unless n = c in which case it equals 1. Hence we have

$$\sum_{1} + \sum_{2} = \sum_{\substack{bc|n \\ c \le V}} \mu(b)\mu(c) = \sum_{\substack{c|n \\ c \le V}} \mu(c)\delta_{c,n} = 0$$

if V < n. So $-\sum_1 = \sum_2 = \sum_3$, hence

$$\mu(n) = -\sum_1 + \sum_4,$$

which implies the result.

We now apply Vaughan's identity to an expression of the form

$$\frac{1}{N}\sum_{N< n\leq 2N}\mu(n)\overline{f(n)}.$$

Using Vaughan's identity in the above one has

 \Box .

$$\frac{1}{N} \sum_{N < n \le 2N} \mu(n) \overline{f(n)} = -\frac{1}{N} \sum_{N < n \le 2N} \sum_{\substack{bc \mid n \\ b \le U, \ c \le V}} \mu(b) \mu(c) \overline{f(n)} + \frac{1}{N} \sum_{N < n \le 2N} \sum_{\substack{bc \mid n \\ b > U, \ c > V}} \mu(b) \mu(c) \overline{f(n)} \\
= -T_{\mathrm{I}} + T_{\mathrm{II}}.$$
(3.2)

Where $T_{\rm I}$ and $T_{\rm II}$ are the type I and type II sums we have referred to in the introduction to this chapter. If we take the type I sum, and make the change of variables d = bc and n = dw, we obtain

$$T_{\rm I} = \frac{1}{N} \sum_{1 \le d \le UV} a_d \sum_{N/d < w \le 2N/d} \overline{f(dw)}$$
(3.3)

where

$$a_d = \sum_{\substack{bc=d\\b\leq U,\ c\leq V}} \mu(b)\mu(c).$$

We can similarly make a change of variables for the type II sum. Taking w = b and n = dw we obtain

$$T_{\rm II} = \frac{1}{N} \sum_{V < d \le 2N/U} \sum_{\max(U, N/d) < w \le N/d} \mu(w) b_d \, \overline{f(dw)}$$
(3.4)

where

$$b_d = \sum_{\substack{c|d\\c>V}} \mu(c).$$

These sums allow us to investigate sums of the form $\sum \mu(n)f(n)$. It is easy to see that if this sum is large, then either $T_{\rm I}$ or $T_{\rm II}$ is large. The forms (3.3) and (3.4) of the type I and II sums, given by Vaughan's identity allow us to analyze their respective sizes. We simplify the expressions and make this precise in the following

Proposition 3.1.2. Let U, V, N be positive integers with $UV \leq N$, and let $f : \mathbb{N} \to \mathbb{C}$ be a function with $||f||_{L^{\infty}} = O(1)$ such that

$$\frac{1}{N} \left| \sum_{N < n \le 2N} \mu(n) \overline{f(n)} \right| \ge \delta$$

for some $\delta > 0$. Then either

• T_I is large. There exists an integer $1 \le D \le UV$ such that

$$\frac{1}{N} \left| \sum_{N/d < w \le 2N/d} f(dw) \right| \gg \frac{\delta}{\log^{5/2}(N)}$$
(3.5)

for $\gg \delta^2 D \log^{-5}(N)$ integers d such that $D < d \le 2D$.

• T_{II} is large. There exist integers D, W with $V/2 \le D \le 4N/U$ and $N/4 \le DW \le 4N$, such that

$$\frac{1}{DW} \left| \sum_{D < d, \, d' \le 2D} \left| \sum_{W < w, \, w' \le 2W} f(dw) \overline{f(dw')} f(d'w) f(d'w') \right| \gg \frac{\delta^4}{\log^{14}(N)}$$
(3.6)

Proof:

We take N to be large. We then have that either $|T_{\rm I}| \ge \delta/2$ or $|T_{\rm II}| \ge \delta/2$. We first deal with the case that the type I sum is large. With reference to the form (3.3) for $T_{\rm I}$, we have the trivial estimate

$$|a_d| \leq \sum_{\substack{bc=d\\b\leq U,\ c\leq V}} |\mu(b)\mu(c)| \leq \sum_{b|d} 1 = \tau(d).$$

Which allows us to estimate the type I sum,

$$\delta \ll |T_{\mathrm{I}}| \leq \sum_{1 \leq d \leq UV} \frac{\tau(d)}{d} \left| \frac{1}{N/d} \sum_{N/d < w \leq 2N/d} f(dw) \right|$$

Regarding the outer sum as an inner product, we may apply the Cauchy-Schwarz inequality and obtain

$$\delta^2 \ll \left(\sum_{1 \le d \le UV} \frac{1}{d} \left| \frac{1}{N/d} \sum_{N/d < w \le 2N/d} f(dw) \right|^2 \right) \cdot \left(\sum_{1 \le d \le UV} \frac{\tau^2(d)}{d} \right).$$

To compute the second sum, we observe that

$$\sum_{n=1}^{\infty} \frac{\tau^2(n)}{n^s} = \frac{\zeta^4(s)}{\zeta(2s)},$$

which by a quick application of Perron's formula yields

$$\sum_{d \le X} \frac{\tau^2(d)}{d} = O(\log^4(X)).$$

Hence

$$\sum_{1 \le d \le UV} \frac{1}{d} \left| \frac{1}{N/d} \sum_{N/d < w \le 2N/d} f(dw) \right|^2 \gg \frac{\delta^2}{\log^4(N)}.$$

Dividing the outer region of summation $1 \le d \le UV$ into dyadic blocks $D < d \le 2D$ allows us to remove the d^{-1} from the above. Applying the pigeonhole principle we obtain

$$\sum_{D \le d \le 2D} \left| \frac{1}{N/d} \sum_{N/d < w \le 2N/d} f(dw) \right|^2 \gg \frac{\delta^2 D}{\log^5(N)}$$

for some D with $1 \leq D \leq UV$. Recall that at the outset we specified $||f||_{L^{\infty}} = O(1)$, so the whole summand of the outer sum is also O(1). Therefore, not too many summands can be too small, and this sort of averaging argument shows that

$$\left|\frac{1}{N/d}\sum_{N/d < w \leq 2N/d} f(dw)\right| \gg \frac{\delta}{\log^{5/2}(N)}$$

for at least $\gg \delta^2 D \log^{-5}(N)$ values of d. This proves the proposition for type I sums.

Now we suppose that $|T_{\rm II}| \ge \delta/2$. Much like with type I sums we see that $|b_d| \le \tau(d)$, and from formula (3.4) we have

$$N\delta \ll \sum_{V < d \le 2N/U} \tau(d) \left| \sum_{N/d < w \le 2N/d} \chi_{(U,2N/d]}(w) \mu(w) f(dw) \right|.$$

Again, we think of this as in inner product and apply Cauchy-Schwarz,

$$N^2 \delta^2 \ll \left(\sum_{V < d \le 2N/U} \frac{\tau^2(d)}{d} \right) \cdot \left(\sum_{V < d \le 2N/U} d \left| \sum_{N/d < w \le 2N/d} \chi_{(U,2N/d]}(w) \mu(w) f(dw) \right|^2 \right)$$

and our estimate for $\sum \tau^2(d)d^{-1}$,

$$\frac{N^2 \delta^2}{\log^4(N)} \ll \sum_{V < d \le 2N/U} d \left| \sum_{N/d < w \le 2N/d} \chi_{(U,2N/d]}(w) \mu(w) f(dw) \right|^2.$$

Once again, we use a dyadic decomposition to remove the d. There exist integers D, W with $V/2 \le D \le 4N/U$ and $N/4 \le DW \le 4N$ such that

$$\frac{N^2 \delta^2}{D \log^5(N)} \ll \sum_{D < d \le 2D} \left| \sum_{W < w \le 2W} \chi_J(w) \mu(w) f(dw) \right|^2.$$

Where $J = (U, 2N/d] \cap (N/d, 2N/d]$. We want to remove the characteristic function χ of the interval from the above expression. For this we will need a technical

Lemma 3.1.3 (Completion of Sums). Let $I \subset \mathbb{Z}$ be a discrete interval, and $f : \mathbb{Z} \to \mathbb{C}$ be a function. Then we have

$$\sup_{J\subseteq I} |\sum_{n\in J} f(n)| \ll \log(1+|I|) \sup_{\alpha\in\mathbb{R}/\mathbb{Z}} |\sum_{n\in I} f(n)e(\alpha n)|$$

where the supremum on the left ranges over discrete sub-intervals of I. More generally, if $I' \subset \mathbb{Z}$ is another discrete interval, and $K : \mathbb{Z} \times \mathbb{Z} \to \mathbb{C}$ is a function, then we have

$$\sum_{m \in I'} |\sum_{n \in I} \chi_{J_m}(n) K(n,m)|^2 \ll \log^2(1+|I|) \sup_{\alpha \in \mathbb{R}/\mathbb{Z}} \sum_{m \in I'} |\sum_{n \in I} K(n,m) e(\alpha n)|^2$$

where for each $m \in I'$, $J_m \subset \mathbb{Z}$ is an arbitrary discrete interval.

Proof:

We apply some discrete Fourier analysis. Assume that I is nonempty and translate the interval to $I = \{1, \ldots, L\}$. We identify this interval with $\mathbb{Z}/L\mathbb{Z}$ so that we may use Fourier analysis. For any subinterval J of I we can expand

$$\sum_{n \in J} f(n) = \sum_{n \in I} \chi_J(n) f(n)$$
$$= \sum_{\xi \in \mathbb{Z}/L\mathbb{Z}} \widehat{\chi}_J(\xi) \sum_{n \in \mathbb{Z}/L\mathbb{Z}} f(n) e\left(\frac{n\xi}{L}\right)$$

Recall that the finite fourier transform is defined

$$\widehat{\chi}_J(\xi) = \frac{1}{L} \sum_{n \in \mathbb{Z}/L\mathbb{Z}} \chi_J(n) e\left(-\frac{n\xi}{L}\right)$$

hence we can crudely approximate it as a geometric series; it is then easy to sum and write a good bound for. We have

$$|\widehat{\chi}_J(\xi)| \le 4 \min\left(1, \frac{1}{L||\xi/L||_{\mathbb{R}/\mathbb{Z}}}\right).$$

We apply this bound with the triangle inequality to the above.

$$\begin{split} |\sum_{n \in J} f(n)| &\leq 4 \sum_{\xi \in \mathbb{Z}/L\mathbb{Z}} \min\left(1, \frac{1}{L||\xi/L||_{\mathbb{R}/\mathbb{Z}}}\right) \left| \sum_{n \in \mathbb{Z}/L\mathbb{Z}} f(n)e\left(\frac{n\xi}{L}\right) \right| \\ &\ll \sum_{\xi \in \mathbb{Z}/L\mathbb{Z}} \min\left(1, \frac{1}{L||\xi/L||_{\mathbb{R}/\mathbb{Z}}}\right) \sup_{\alpha \in \mathbb{R}/\mathbb{Z}} \sum_{n \in I} f(n)e(n\alpha) \\ &\ll \log(L+1) \sup_{\alpha \in \mathbb{R}/\mathbb{Z}} \sum_{n \in I} f(n)e(n\alpha). \end{split}$$

This proves the first part of the lemma. The second part of the lemma follows by essentially the same argument, but working in $l^2(\mathbb{Z})$ instead. We have

$$\begin{split} (\sum_{m\in I'}|\sum_{n\in I}\chi_{J_m}(n)K(n,m)|^2)^{\frac{1}{2}} &\ll \left(\sum_{m\in I'}(\sum_{\xi\in\mathbb{Z}/N\mathbb{Z}}\min\left(1,\frac{1}{L||\xi/L||_{\mathbb{R}/\mathbb{Z}}}\right)|\sum_{n\in\mathbb{Z}/L\mathbb{Z}}K(n,m)e\left(\frac{n\xi}{L}\right)|)^2\right)^{\frac{1}{2}} \\ &\ll \sum_{\xi\in\mathbb{Z}/N\mathbb{Z}}\min\left(1,\frac{1}{L||\xi/L||_{\mathbb{R}/\mathbb{Z}}}\right)\left(\sum_{m\in I'}|\sum_{n\in I}K(n,m)e\left(\frac{n\xi}{L}\right)|^2\right)^{\frac{1}{2}} \\ &\ll \log(L+1)\sup_{\alpha\in\mathbb{R}/\mathbb{Z}}\left(\sum_{m\in I'}|\sum_{n\in I}K(n,m)e\left(\frac{n\xi}{L}\right)|^2\right)^{\frac{1}{2}}. \end{split}$$
 Hence the second part of the lemma.
$$\Box.$$

Hence the second part of the lemma.

Returning to the original proposition we were proving, we take K(w,d) = $\mu(w)f(dw)$, and applying the second form of the lemma we obtain

$$\frac{N\delta^2}{\log^7(N)} \ll \sum_{D < d \le 2D} \left| \sum_{W < w \le 2W} \mu(w) f(dw) e(w\alpha) \right|^2$$

for some $\alpha \in \mathbb{R}/\mathbb{Z}$. We expand the right hand side as

$$\sum_{W < w, w' \le 2W} \sum_{D < d \le 2D} u(w, w') f(dw) \overline{f(dw')},$$

where $u(\cdot, \cdot)$ is a bounded function whose exact form we do not worry about. The proposition follows from the Cauchy-Schwarz inequality on the Gowers U^2 norm. \Box .

Orthogonality of $\mu(n)$ and $e(n\alpha)$ 3.2

We now apply the results of the preceding section to the special case where $f(n) = e(n\alpha)$. Throughout the notation $||\alpha||_{\mathbb{R}/\mathbb{Z}}$ is used to denote the distance of α to the nearest integer. In our results, we wish to say that $\sum \mu(n)e(n\alpha)$ is small if α is far from a rational number with small denominator (minor arc). However, in keeping with our "inverse" philosophy, we will prove instead that if the sum is large, then α must be in a major arc. These are, of course, logically the same thing. We now approach

Proposition 3.2.1. Let $\alpha \in \mathbb{R}$, let A > 0, and let N be a large integer such that

$$\left|\frac{1}{N}\sum_{N < n \le 2N} \mu(n)e(-n\alpha)\right| \ge \log^{-A}(N).$$
(3.7)

Then there exists $D, 1 \leq D \ll N^{2/3}$, such that

$$\#\{1 \le d \le 2D : ||\alpha d||_{\mathbb{R}/\mathbb{Z}} \ll \frac{D}{N} \log^{4A+14}(N)\} \gg D \log^{-4A-14}(N).$$
(3.8)

Proof:

Apply the previous proposition (3.1.2). For parameters U and V we pick $U = V = N^{\frac{1}{3}}$. The proposition says that either the type I or type II sum is large. We will show that either of these possibilities yields (3.8). Specifically, the proposition gives that one of the following holds,

• $T_{\rm I}$ is large. There exists an integer $1 \le D \le N^{\frac{2}{3}}$ such that

$$\frac{1}{N} \left| \sum_{N/d < w \le 2N/d} e(\alpha dw) \right| \gg \frac{1}{\log^{A+5/2}(N)}$$
(3.9)

for $\gg D \log^{-2A-5}(N)$ integers d such that $D < d \le 2D$.

• $T_{\rm II}$ is large. There exist integers D, W with $N^{\frac{1}{3}} \ll D \ll N^{\frac{2}{3}}$ and $N/8 \le DW \le 8N$, such that

$$\frac{1}{DW} \left| \sum_{D < d, \, d' \le 2D} \sum_{W < w, \, w' \le 2W} e(\alpha(dw - dw' - d'w + d'w')) \right| \gg \frac{1}{\log^{4A + 14}(N)}$$
(3.10)

We begin with assuming that $T_{\rm I}$ is large. Recall our bound on a geometric series from the above lemma. If I is any interval in \mathbb{Z} , then

$$\left|\sum_{n\in I} e(n\alpha)\right| \le 4\min\left(|I|, \frac{1}{||\alpha||_{\mathbb{R}/\mathbb{Z}}}\right).$$

Hence, there are $\gg D \log^{-2A-5}(N)$ values of $d, D < d \le 2D$, for which

$$||\alpha d||_{\mathbb{R}/\mathbb{Z}} \ll \frac{D}{N} \log^{A+5/2}(N),$$

which gives (3.8) with room to spare.

Now we assume instead that $T_{\rm II}$ is large. By the pigeonhole principle we can find d', w' such that

$$\frac{1}{DW} \left| \sum_{D < d \leq 2D} \left| \sum_{W < w \leq 2W} e(\alpha(dw - dw' - d'w + d'w')) \right| \gg \frac{1}{\log^{4A + 14}(N)}.$$

Now using the triangle inequality and reducing two factors to trivial we have

$$\frac{1}{DW} \sum_{D < d \le 2D} \left| \sum_{W < w \le 2W} e(\alpha w(d - d')) \right| \gg \frac{1}{\log^{4A + 14}(N)}.$$

Applying our geometric sum estimate again gives

$$\frac{1}{D} \sum_{D < d \le 2D} \min\left(1, \frac{D}{N ||\alpha(d - d')||_{\mathbb{R}/\mathbb{Z}}}\right) \gg \frac{1}{\log^{4A + 14}(N)}$$

Averaging, we have

$$\#\{D < d \le 2D : ||\alpha(d - d')||_{\mathbb{R}/\mathbb{Z}} \ll \frac{D}{N} \log^{4A + 14}(N)\} \gg D \log^{-4A - 14}(N).$$

We now make the change of variables $\tilde{d} = d - d'$, and obtain

$$\#\{-2D \le d \le 2D : ||\alpha \tilde{d}||_{\mathbb{R}/\mathbb{Z}} \ll \frac{D}{N} \log^{4A+14}(N)\} \gg D \log^{-4A-14}(N).$$

Because $D \geq N^{\frac{1}{3}}$, we may remove the case $\tilde{d} = 0$ and because $||\beta||_{\mathbb{R}/\mathbb{Z}} = ||-\beta||_{\mathbb{R}/\mathbb{Z}}$, we have symmetry in the above expression and hence the result. \Box .

We now have a proposition about the size of our sum $\sum \mu(n)e(n\alpha)$ in relation to where α is. However, the conclusion of the proposition does not make it entirely obvious under what conditions our exponential sum is small. We see that it involves a major arc / minor arc condition, for if α is close to a rational number with denominator d then $||\alpha d||_{\mathbb{R}/\mathbb{Z}}$ is small. It turns out that the size of the denominator can be made much smaller than this. However, to make this precise, we need some tools from harmonic analysis.

3.2.1 Some Harmonic Analysis Tools

In this section we develops some technical lemmas along the lines of Montgomery [Mo] and Green and Tao [GT]. We begin with some definition and state Vaaler's Lemma, from which one deduces the Erdős-Turán inequality. This inequality allow us to prove an important lemma relating recurrent linear function to the major arcs of our proof, and in turn proves the more precise version of the above proposition.

We begin by defining the saw-tooth function:

$$s(x) = \begin{cases} \{x\} - \frac{1}{2} & \text{if } x \notin \mathbb{Z} \\ 0 & \text{if } x \in \mathbb{Z}. \end{cases}$$

And recall the definition of the Fejér kernel,

$$\Delta_{K}(x) = \sum_{-K}^{K} \left(1 - \frac{|k|}{K} \right) e(kx) = \left(\frac{\sin(\pi Kx)}{\sin(\pi x)} \right)^{2}.$$
 (3.11)

Then we have

Definition 3.2.2 (Vaaler's Polynomial). Vaaler's polynomial is

$$V_{K}(x) = \frac{1}{K+1} \sum_{k=1}^{K} \left(\frac{k}{K+1} - \frac{1}{2} \right) \Delta_{K+1} \left(x - \frac{k}{K+1} \right) + \frac{1}{2\pi(K+1)} \sin(2\pi(K+1)x) - \frac{1}{2\pi} \Delta_{K+1}(x) \sin(2\pi x).$$
(3.12)

 $V_K(x)$ is a good approximation to the saw-tooth function s(x), we have

$$V_K(x) = s(x) + O\left(\min\left(1, \frac{1}{K^3 ||x||_{\mathbb{R}/\mathbb{Z}}^3}\right)\right).$$

Another possible approximation to s(x) is the

Definition 3.2.3 (Beurling polynomial). The Buerling polynomial is

$$B_K(x) = V_K(x) + \frac{1}{2(K+1)} \Delta_{K+1}(x).$$
(3.13)

The relationship between these functions is given by

Lemma 3.2.4 (Vaaler's Lemma). If $0 \le x \le 1/2$ then $s(x) \le V_K(x) \le B_K(x)$, while if $1/2 \le x \le 1$ then $V_K(x) \le s(x) \le B_K(x)$. If T(x) is a trigonometric polynomial of degree $\le K$ such that $T(x) \ge s(x)$ for all x then

$$\int_{\mathbb{R}/\mathbb{Z}} T(x) \, dx \ge \frac{1}{2(K+1)}$$

with equality if and only if $T(x) = B_K(x)$.

Proof:

The proof of Vaaler's lemma is a long series of calculations involving approximations to the derivatives of the functions involved. For full details see Montgomery [Mo].

We now make a few more definitions. Given a sequence $\{u_n\}$ of points $U_n \in \mathbb{R}/\mathbb{Z}$, and $0 \le \alpha \le \beta \le 1$, define

$$Z(N,\alpha,\beta) = \sum_{\substack{n \le N \\ \alpha \le u_n \le \beta}} 1$$
(3.14)

Then the sequence $\{u_n\}$ is uniformly distributed if

$$\lim_{N \to \infty} \frac{1}{N} Z(N, \alpha, \beta) = \beta - \alpha$$

for every choice of α, β . If we let U_N be the measure on \mathbb{R}/\mathbb{Z} that places a point mass at each point of $\{u_n\}$, then we may take the fourier transform of this measure,

$$\widehat{U}_{N}(k) = \int_{\mathbb{R}/\mathbb{Z}} e(-k\alpha) \, dU_{N}$$
$$= \sum_{n=1}^{N} e(-ku_{n}).$$
(3.15)

If we define

$$D(N, \alpha, \beta) = Z(N, \alpha, \beta) - (\beta - \alpha)N$$

we can define the *discrepancy*

$$D(N) = ||D(N, \alpha, \beta)||_{L^{\infty}(\alpha, \beta)}$$

The discrepancy, in essence, measures how close $\{u_n\}$ is to being uniformly distributed. It is easy to see that if D(N) = o(N) is equivalent to the sequence being uniformly distributed. We can apply Vaaler's lemma to study the discrepancy, however, we need one more

Definition 3.2.5 (Selberg Polynomials). The Selberg polynomials are given by

$$S_K^+(x) = \beta - \alpha + B_K(x - \beta) + B_K(\alpha - x)$$

and

$$S_K^-(x) = \beta - \alpha - B_K(\beta - x) - B_K(x - \alpha).$$

Now we are ready to state and prove

Lemma 3.2.6 (Erdős-Turán inequality). For any positive integer K,

$$|D(N,\alpha,\beta)| \le D(N) \le \frac{N}{K+1} + 3\sum_{k=1}^{K} \frac{1}{k} \left| \sum_{n=1}^{N} e(ku_n) \right|.$$
(3.16)

Proof:

We begin with Vaaler's lemma which gives that $-B_K(-x) \le s(x) \le B_K(x)$ for all x. Because

$$\chi_{\mathcal{J}}(x) = \beta - \alpha + s(x - \beta) + s(\alpha - x),$$

we see that the Selberg polynomials are trigonometric polynomials of degree at most K which approximate the characteristic function of the arc $\mathcal{J} = [\alpha, \beta] \subseteq \mathbb{R}/\mathbb{Z}$. By applying Vaaler's lemma to the definitions of the Selberg polynomials, we have that $S_K^-(x) \leq \chi_{\mathcal{J}}(x) \leq S_K^+(x)$ for all x and that

$$\int_{\mathbb{R}/\mathbb{Z}} S_K^{\pm}(x) \, dx = \beta - \alpha \pm \frac{1}{K+1}.$$

Now we approach the discrepancy. Fourier inversion and our facts concerning the Selberg polynomials give that

$$Z(N, \alpha, \beta) = \sum_{n=1}^{N} \chi_{\mathcal{J}}(u_n)$$

$$\leq \sum_{n=1}^{N} S_K^+(u_n)$$

$$= \sum_{n=1}^{N} \sum_{k=-K}^{K} \widehat{S}_K^+(k) e(ku_n)$$

$$= \sum_{k=-K}^{K} \widehat{S}_K^+(k) \widehat{U}_N(-k).$$

We have symmetry in k, and clearly by the above deductions that $\widehat{U}_N(0) = N$ and $\widehat{S}_K^+(0) = \beta - \alpha + \frac{1}{K+1}$. Hence we arrive at

$$D(N,\alpha,\beta) \le \frac{N}{K+1} + \sum_{0 < |k| \le K} \widehat{S}_K^+(k)\widehat{U}_N(-k).$$

We must now estimate the summand. We employ the estimate $|\hat{f}(k)| \leq ||f||_{L^1}$ to the function $S_K^+(x) - \chi_{\mathcal{J}}(x)$. Thus we calculate

$$|\hat{S}_{K}^{+}(k) - \hat{\chi}_{\mathcal{J}}(k)| \le ||S_{K}^{+} - \chi_{\mathcal{J}}||_{L^{1}} = \frac{1}{K+1}.$$

It is a introductory exercise to compute

$$\widehat{\chi}_{\mathcal{J}}(k) = \frac{e(k\alpha) - e(k\beta)}{2\pi i k},$$

therefore

$$|\widehat{\chi}(k)| = \left|\frac{\sin(\pi k(\beta - \alpha))}{\pi k}\right| \le \min\left(\beta - \alpha, \frac{1}{\pi |k|}\right),$$

when $k \neq 0$. Combining this with the above we have

$$|\widehat{S}_{K}^{+}(k) \leq \frac{1}{K+1} + \min\left(\beta - \alpha, \frac{1}{\pi|k|}\right).$$
 (3.17)

This allows us to evaluate our discrepancy estimate

$$D(N,\alpha,\beta) \le \frac{N}{K+1} + 2\sum_{k=1}^{K} \left(\frac{1}{K+1} + \min\left(\beta - \alpha, \frac{1}{\pi|k|}\right)\right) |\widehat{U}_N(k)|.$$

A analogous argument produces a lower bound on $D(N,\alpha,\beta)$ using $\widehat{S}^+_K(k).$ Finally, observing that

$$\frac{1}{K+1} + \frac{1}{\pi k} < \frac{3}{2k}$$

we have the estimate on the discrepancy D(N) in the lemma.

 \Box .

We will apply the Erdős-Turán inequality in the proof of another lemma, which will be more useful to us,

Lemma 3.2.7. Let $I \subset \mathbb{Z}$ be an interval, let $\alpha \in \mathbb{R}/\mathbb{Z}$, and suppose the set

$$\mathcal{L} = \{l \in I : ||lpha l||_{\mathbb{R}/\mathbb{Z}} \le \delta_1\}$$

is of size at least $\delta_2|I|$ for some $0 < \delta_1$, $\delta_2 < 1$ with $\delta_1 \leq \frac{1}{4}\delta_2$. Then

1. If
$$|I| > 1/\delta_2$$
, then $\inf_{1 \le d \le 8/\delta_2} ||\alpha d||_{\mathbb{R}/\mathbb{Z}} \le \frac{2^8}{\delta_2^2 |I|}$.
2. If $|I| > 2/\delta_2^2$, then $\inf_{1 \le d \le 16/\delta_2^2} ||\alpha d||_{\mathbb{R}/\mathbb{Z}} \le \frac{2^1 5\delta_1}{\delta_2^6 |I|}$.

Proof:

Take I to be{M + 1, ..., M + L}. Define a sequence $\{u_l\}_{l=1}^{\infty}$ as

$$u_l = \alpha(M+l) \mod 1.$$

The lower bound on the size of $\mathcal L$ in the hypotheses gives us a estimate on the discrepancy

$$D(L, -\delta_1, \delta_1) \ge \delta_2 - 2\delta_1)L \ge \frac{1}{2}\delta_2L.$$

Comparing this with the estimate from the Erdős-Turán inequality, we have that

$$\frac{1}{2}\delta_2 L \le \frac{L}{K+1} + 3\sum_{k=1}^{K} \frac{1}{k} \left| \sum_{l=1}^{L} e(ku_l) \right|$$

for any K. We pick $K = \lfloor \frac{4}{\delta_2} \rfloor$. Hence, there exists $k \leq \frac{8}{\delta_2}$ so that

$$\left|\sum_{l=1}^{L} e(ku_l)\right| \ge \frac{\delta_2^2 L}{2^6}.$$

Now, we use the estimate for a geometric sum

$$|\sum_{n\in I} e(n\alpha)| \leq 4\min\left(|I|, \frac{1}{||\alpha||_{\mathbb{R}/\mathbb{Z}}}\right)$$

from the first half of this chapter, rearrange, and part 1. of the lemma follows.

The estimate 2. is much stronger, and we use a bootstrap method to work 1. into 2. Assume $\delta_1 < \frac{\delta_2^2}{16}$ because otherwise 1. already implies the result. Let $1 \le m \le L$ (which will be chosen later), and define

$$\mathcal{L}_b = \{b+1, \dots, b+m\} \cap \mathcal{L}_b$$

By the lower bound for \mathcal{L} given in the hypotheses, we apply the pigeon-hole principle to find that there exists some b so that $|\mathcal{L}_b| \geq \frac{\delta_2 m}{2}$. Fix this b. Now we consider the set $m\mathcal{L} + \mathcal{L}_b$. By our above estimate we have that

$$|m\mathcal{L} + \mathcal{L}_b| \ge \frac{\delta_2 mL}{2},$$

and we also have that

$$m\mathcal{L} + \mathcal{L}_b \subseteq I' = \{m(M+1) + b + 1, \dots, m(M+L) + b + m\},\$$

and $|I| \leq mL$. Finally, for any $x \in m\mathcal{L} + \mathcal{L}_b$ one will have $||\alpha x||_{\mathbb{R}/\mathbb{Z}} \leq 2\delta_1 m$. We can now apply part 1. of the lemma with $I \to I'$, $\delta_1 \to 2m\delta_1$ and $\delta_2 \to \delta_2^2/2$ provided we satisfy the conditions of the lemma. If we set $m = \lfloor \delta_2^2/16\delta_1 \rfloor$, we have $m \leq \delta_2^2/16\delta_1$ and $mL > 2/\delta_2^2$, so we can apply 1. and the result follows. \Box .

Given this lemma, we now have to tools at our disposal to return to proposition (3.2.1) regarding the orthogonality of the möbius function and a linear phase giving us a major arc condition. We are prepared to make the proposition more precise with our most important

Corollary 3.2.8. Let $\alpha \in \mathbb{R}$, let A > 0 and let N be a large integer such that

$$\left|\frac{1}{N}\sum_{N < n \le 2N} \mu(n)e(-n\alpha)\right| \ge \log^{-A}(N)$$

holds. Then

1

$$\inf_{\leq d \leq 16 \log^{8(A+4)}(N)} ||\alpha d||_{\mathbb{R}/\mathbb{Z}} \ll \frac{\log^{28(A+4)}(N)}{N}$$
(3.18)

Proof:

We apply proposition (3.2.1). Therefore we have that there exists a D with $1 \leq D \leq N^{\frac{2}{3}}$ such that (3.8) holds. If D is less than $\log^{8(A+4)}(N)$, then we have he corollary directly from (3.8). If $D \geq \log^{8(A+4)}(N)$, then the argument in somewhat more difficult. Observe that if we have a linear sequence αl which has small circle norm for many l in an interval I, then α must lie on a major arc. To see this, we apply the lemma just proved in the case that $D \geq \log^{8(A+4)}(N)$, we have our interval I = [1, 2D]. We pick the parameters δ_1 and δ_2 to be

$$\delta_1 \ll \frac{D}{N} \log^{4(A+4)}(N)$$

$$\delta_2 \ll \frac{1}{\log^{4(A+4)}(N)}.$$

Then, conclusion (2) of the lemma gives the corollary.

 \Box .

From the corollary, we take the contrapositive. Extending our sum from [N, 2N] to [1, N], one picks up at most an extra factor of $\log(N)$, and obtains

Lemma 3.2.9. Let $\alpha \in \mathbb{R}$, and C > 0. If for sufficiently large N, and any fixed k we have

$$\inf_{1 \le d \le 16 \log^{8(C+4)}(N)} ||d\alpha||_{\mathbb{R}/\mathbb{Z}} \ge k \frac{\log^{28(C+4)}(N)}{N},$$

then

$$|\frac{1}{N}\sum_{1 \le n \le N} \mu(n)e(n\alpha)| = O(\log^{-C+1}(N)),$$

which was our third main lemma from the introductory chapter.

Chapter 4

Conclusion

Drawing together the results on chapters 2 and 3 and their application in chapter 1, we prove Vinogradov's theorem on sums of three primes. From the proof given, one sees two main ingredients. The first is the prime number theorem for arithmetic progressions. That is, the explicit formula given by Mellin transform, and the classical zero-free regions. Second is a theory for the treatment of exponential sums, which in this essay is given by Vaughan's identity and many analytic estimates, proving that the Möbius function is orthogonal to linear phases. This orthogonality, of course, is related to the Möbius randomness law, and hence is related to zero-free regions of L-functions and the Riemann hypothesis. All together, the proof draws on a wealth of deep and interesting mathematics.

Unfortunately, Vinogradov's result does not show that all odd integers are a sum of three primes, but rather the weaker statement that all sufficiently large odd integers are a sum of three primes. To determine exactly how large sufficiently large is one would have to investigate more carefully the constants in the $O\left(\frac{N^2}{\log^A(N)}\right)$ term. Unfortunately, Siegel's theorem (which gives an *ineffective* bound on how far a Siegel zero may be from s = 1) is invoked in the proof of Vinogradov's theorem, and hence determining the constants implicit in the O term is quite difficult.

Some progress has been made in this area. In 1956 Borodzkin showed that Vinogradov's theorem holds for all odd $n > 3^{3^{15}}$. In 1989, Chen and Wang reduced the number to 10^{43000} , which is still unfortunately beyond the range of a computer check. See Ribenboim, [Rib]. As an area for future research, one would like to have a more useful way of making constants in Siegel's theorem effective, and get these bound down to reasonable sizes.

Bibliography

- [Dav] H. Davenport, *Multiplicative Number Theory*, Springer, Berlin, Heidelberg, New York, 2000.
- [Gol] D. Goldfeld, A Simple Proof of Siegel's Theorem, Proceedings of the National Academy of Sciences, USA 71 (1974), 1055.
- [Gr1] B. J. Green, Additive Number Theory, Lecture Notes, Cambridge University, Lent Term 2007.
- [GT] B. J. Green and T. Tao, *Quadratic Uniformity of the Möbius Function*, Preprint.
- [IK] H. Iwaniec and E. Kowalski, Analytic Number Theory, American Mathematical Society Colloquium Publications, Providence, RI, USA, 2004.
- [Mo] H. L. Montgomery, Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis, Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, Providence, RI, USA, 1994.
- [Rib] P. Ribenboim, The New Book of Prime Number Records, Springer, Berlin, Heidelberg, New York, 1996.
- [St1] E.M. Stein and R. Shakarchi, *Complex Analysis*, Princeton University Press, Princeton, NJ, USA, 2003.
- [St2] E.M. Stein and R. Shakarchi, Real Analysis, Measure Theory, Integration and Hilbert Spaces, Princeton University Press, Princeton, NJ, USA, 2005.
- [Tit] E.C. Titchmarsh, Introduction to the Theory of Fourier Integrals, Oxford University Press, Oxford, UK, 1975.
- [WW] E.T. Whittaker and G.N. Watson, A Course of Modern Analysis, Fourth Edition, Cambridge University Press, Cambridge, UK, 1963.