Fourier Analytic Applications to Number Theory

by

Mariah Hamel

B.A., Colby College, 2002

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate Studies

(Department of Mathematics)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

October 5, 2004

Library Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Mariah E Hamel
Name of Author *(please print)*

04/10/2004
Date (dd/mm/yyyy)

Title of Thesis: Fourier Analytic Applications to Number Theory.

Degree: MSc Year: 2004

Department of Mathematics
The University of British Columbia
Vancouver, BC Canada

Abstract

In this paper we give expositions of Roth's theorem, Weyl's inequality and Vinogradov's three-primes theorem. In the proofs, we will frequently use exponential sums and more specifically the discrete Fourier transform. In the proof of Vinogradov's three-primes theorem we will use Hardy and Littlewood's circle method. This paper is intended to be self-contained and will hopefully be readable to someone with little background in the area.

Contents

Preface

When I first read about Vinogradov's three-primes theorem, I was overwhelmed. After examining a few sources, I came across Gowers' lecture notes [3] on the internet. His notes were much more readable and intuitive than any of the other sources I had seen. This paper originates with those notes. My goal in this paper is to provide a gentle introduction to the applications of Fourier analysis in Number Theory.

At the suggestion of my advisor, Izabella Łaba, I began studying this area of Number Theory with Roth's theorem. Besides the applications to Szemerédi's theorem [4], the proof of Roth's theorem provided me with a nice way to become more comfortable with the discrete Fourier transform. Therefore, I have included a proof of this theorem here.

Weyl's inequality is typically used in minor arcs estimates in applications of the circle method. I have included a proof of Weyl's inequality for quadratic polynomials as well as the related lemma A.0.5 which we will use in the minor arcs estimates for Vinogradov's three-primes theorem. The inclusion of Weyl's inequality provides for a fairly self contained proof of the three-primes theorem and also gives the reader important background material for other problems in this field.

In the proof of Vinogradov's three-primes theorem my primary goal was to clarify Gowers' notes and to fill in the gaps that may have resulted from the notes being written as a supplement to class lectures. I have chosen to omit constants since they originally distracted me from overlying ideas.

Throughout this paper, I have tried to combine available sources and to to take the best elements from each of them. I have tried to provide the reader with a foresight in the form of outlines of proofs, where a lemma or theorem may be used, etc. I have attempted to clarify portions of the proofs that I found difficult upon first reading. And, finally, I have used the opportunity to write this thesis to become familiar with this area of mathematics.

Acknowledgements

I would like to thank Izabella Łaba for all of her help and encouragement with this project. I would like to thank David Boyd for agreeing to be the second reader. I would also like to thank Ben Green for discussions of material relating to this thesis and for suggestions and open problems in this area.

Chapter 1

Introduction

The purpose of this paper is to provide an introduction to the application of Fourier analytic techniques in number theory. The theorems presented below combine methods from additive, combinatorial and analytic number theory. Specifically, we will prove Roth's theorem, Weyl's inequality and Vinogradov's three-primes theorem. In an effort to keep this paper fairly self-contained, we include an introductory section to the discrete Fourier transform. We also briefly describe the history of each theorem and summarize recent related results.

The theorems presented all concern the structure of subsets of the integers. The following are natural questions: Does a subset contain an arithmetic or geometric progression? Does a given subset form a base of the integers? When is the distribution of a subset of integers 'random'? What can we extrapolate about a subset from its sumset? The theorems in this paper formulate and provide answers to some of these questions.

We say that a subset of nonnegative integers, A, is a basis of finite order k if $\mathbb{N} \subset A + A + ... + A = \{a_1 + ... + a_k : a_i \in A\}$. For example, it is clear that $A = \{0, 1, 3, 5, ...\}$ is a basis of order 2. In 1770, Waring stated without proof that every natural number is the sum of at most four squares, nine cubes, nineteen fourth powers, etc. Stated precisely he claimed that $A = \{0, 1^n, 2^n, 3^n, ...\}$ is a basis of finite order for each n. The Goldbach conjecture states that every even integer greater than or equal to six is the sum of exactly two prime numbers. This general type of problem can be characterized as an attempt to show that the natural numbers can be represented as solutions to arithmetic equations over a restricted domain.

An arithmetic progression of length k is a set of the form $P = \{a, a + s, a + 2s, ..., a + (k-1)s\}$. Szemerédi's theorem proves that an arithmetic progression of arbitrary length can be found in every sufficiently dense subset of the integers. An example of a set which is not sufficiently dense, and hence does not satisfy the hypothesis of Szemerédi's theorem, is the prime numbers. Despite this, Green and Tao recently (May 2004) proved that the primes

also contain arithmetic progressions of arbitrary length. Another interesting result in this area is that any sufficiently dense subset of the integers must contain two elements which differ by a perfect square. This result was first proved by Sárkozy and separately by Furstenberg, who came by the result as a corollary to his proof of Szemerédi's theorem.

A third topic in this area is called inverse additive number theory. While we will not address this in detail the area may be of interest to the reader. Problems from this area involve the study of sumsets or difference sets. Let A be a subset of the positive integers. We define a sumset to be the set $A + A = \{a_1 + a_2 : a_1, a_2 \in A\}$ and similarly the difference set $A - A$. An important result is Freiman's theorem which states that if the sumset $A + A$ is small, then A must be contained in a generalized arithmetic progression. Freiman's theorem is related to the Balog-Szemerédi theorem which Gowers uses in his proof of Szemerédi's theorem.

Many different methods can be applied to the above problems. A good illustration is Szemerédi's theorem which can be proved using combinatorial or ergodic or analytic methods. Each approach has its own advantage, and in this paper we will focus on Fourier analytic techniques. Although this approach gives quantitative bounds (which ergodic methods do not give), we will not always provide them in an effort to keep the proofs as readable as possible.

1.1 Notation

Throughout this paper we will use the following notation:

We will be estimating many exponential sums and therefore the following will be very useful.

$$e(x) = e^{2\pi i x}$$

$$\exp(x) = e^{x}$$

At times, we will find that it is easier not to keep track of certain constants and so if f and g are functions, $g(x) > 0$ for all x and there exists a constant M such that $|f(x)| \leq Mg(x)$ for all x, then we will write

$$f(x) = O(g(x))$$

or

$$f(x) \ll g(x).$$

The standard floor and ceiling functions are defined to be

$$\lfloor \alpha \rfloor = \max\{n \in \mathbb{Z} : \alpha \geq n\}$$

and

$$\lceil \alpha \rceil = \min\{n \in \mathbb{Z} : \alpha \leq n\}.$$

We define the fractional part of the real number α to be

$$\{\alpha\} = \alpha - \lfloor \alpha \rfloor.$$

We will denote the distance from a real number α to the closest integer by

$$||\alpha|| = \min(|n - \alpha| : n \in \mathbb{Z}).$$

We will also use some functions typically used in Number Theory:
We define the Möbius function to be

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n \text{ is divisible by the square of a prime} \\ (-1)^r & \text{if } n \text{ is the product of } r \text{ distinct primes.} \end{cases}$$

The Euler ϕ-function is defined to be the number of positive integers less than n which are relatively prime to n. We can count the number of primes less than any real number x which we denote by $\pi(x)$. Finally, we define the von Mangoldt function to be

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^m \text{ for some prime } p \text{ and some } m \geq 1, \\ 0 & \text{otherwise .} \end{cases}$$

We refer the reader to appendix B for some interesting results relating to these functions.

Chapter 2

The discrete Fourier transform

In this section we will introduce the discrete Fourier transform and prove several related identities. We take f to be a function which maps the group $\mathbb{Z}_N$, the set $\{0, 1, ..., N-1\}$ under modular addition, to the complex numbers. The Fourier transform $\widetilde{f}$ of f preserves certain important properties and is often easier to study. The Fourier inversion formula will provide a way to recover the original function f.

Although this necessary background comes ahead of where we must use it, we will attempt to provide some motivation. Roth's theorem states that any subset A of the numbers $\{0, 1, ..., N-1\}$ of size δN must contain an arithmetic progression of length three for any $\delta > 0$ assuming that N is sufficiently large. To prove Roth's theorem, we will cover the two cases defined by the size of the Fourier transform of the characteristic function of the set A.

Definition 2.0.1. *Let $f : \mathbb{Z}_N \to \mathbb{C}$. Then for any $r \in \mathbf{C}$ define the* **discrete Fourier transform** *of f to be*

$$\widetilde{f}(r) = \sum_{s=0}^{N-1} f(s)e(-rs/N).$$

In the case of Roth's theorem, we define $A(x)$ to be the characteristic function of A. Then by definition 2.0.1 $\widetilde{A}(r) = \sum_{x=0}^{N-1} A(x)e(-rx/N) = \sum_{a \in A} e(-ra/N)$. Therefore, the magnitude of the Fourier transform $\widetilde{A}$ is dependent on the distribution of points $e(-ra/N)$ on the unit circle. We will see that small Fourier transforms, which means these points have a fairly even density on the entire circle, coincides with the set A being "random".

For the remainder of this section, where the limits of summation are clear we will omit the bounds.

Definition 2.0.2. *Let $f, g : \mathbb{Z}_N \to \mathbb{C}$. Then*

$$f * g(s) = \sum_t f(t)\overline{g(t-s)}$$

is defined to be the **convolution** *of f and g.*

The next identity provides a relation between the discrete Fourier transforms of f and g and their convolution.

Lemma 2.0.3. $\widetilde{(f * g)} = \widetilde{f}(r)\overline{\widetilde{g}(r)}$.

Proof: By definition 2.0.1 and definition 2.0.2 we have

$$\begin{aligned}
\widetilde{(f * g)} &= \sum_s (f * g)(s) e(-rs/N) \\
&= \sum_s (\sum_t f(t)\overline{g(t-s)}) e(-rs/N) \\
&= \sum_{s,t} f(t)\overline{g(t-s)} e(-rt/N) e(-r(t-s)/N) \\
&= \sum_{t,u} f(t) e(-rt/N) \overline{g(u) e(-ru/N)} \\
&= \widetilde{f}(r)\overline{\widetilde{g}(r)}. \ \square
\end{aligned}$$

We will use the following identity in the proof of Parseval's identity, Plancherel's formula, the Fourier inversion formula and Roth's theorem.

$$\sum_{s=0}^{N-1} e(-rs/N) = \begin{cases} N \text{ if } r = 0 \\ 0 \text{ if } r \neq 0 \end{cases} \tag{2.0.1}$$

Lemma 2.0.4. Parseval's identity.

$$\sum_r \widetilde{f}(r)\overline{\widetilde{g}(r)} = N \sum_s f(s)\overline{g(s)}.$$

Proof: Using lemma 2.0.3 and equation 2.0.1 we have

$$\begin{aligned}
\sum_r \widetilde{f}(r)\overline{\widetilde{g}(r)} &= \sum_r \sum_s (f * g)(s) e(-rs/N) \\
&= \sum_s (f * g)(s) \sum_r e(-rs/N) \\
&= N((f * g)(0)) \\
&= N \sum_s f(t)\overline{g(t)}. \ \square
\end{aligned}$$

Plancherel's formula says that the L^2 norm of the discrete Fourier transform of f is proportional to the L^2 norm of f.

Lemma 2.0.5. Plancherel's formula.

$$\sum_r |\widetilde{f}(r)|^2 = N \sum_s |f(s)|^2.$$

Proof: Using lemma 2.0.4 let $\widetilde{g(r)} = \overline{\widetilde{f}(r)}$. Then $\widetilde{f}(r)\overline{\widetilde{f}(r)} = |\widetilde{f}(r)|^2$. □
Finally, we are ready to prove the Fourier inversion formula.

Lemma 2.0.6. Fourier inversion formula.

$$f(s) = N^{-1} \sum_r \widetilde{f}(r) e(rt/N).$$

Proof:

$$\begin{aligned} N^{-1} \sum_r \widetilde{f}(r) e(rs/N) &= N^{-1} \sum_r (\sum_t f(t) e(-rt/N)) e(rs/N) \\ &= N^{-1} \sum_r \sum_t f(t) e(r(s-t)/N) \\ &= N^{-1} \sum_t f(t) \sum_r e(r(s-t)/N) \\ &= N^{-1}(f(s)N) \\ &= f(s) \end{aligned}$$

by equation 2.0.1. □

Chapter 3

Roth's Theorem

3.1 History

In this section we prove Roth's theorem on arithmetic progressions of length three. One of the first developments in the study of arithmetic progressions in sets of integers was given by van der Waerden in 1927 when he proved the following theorem:

Theorem 3.1.1. *Let k and $r \in \mathbb{N}$. Then there exist $M \in \mathbb{N}$ which depends on k and r such that if $\{1, ..., M\}$ is partitioned into any r subsets, then one of these subsets must contain an arithmetic progression of length k.*

In 1936, Erdős and Turán attempted to strengthen van der Waerden's result with their conjecture that an arithmetic progression of length k could be found in any sufficiently dense subset of the integers. In 1953, Roth made the first progress toward this conjecture by proving the case with $k = 3$. Szemerédi proved the conjecture for $k = 4$ in 1969 and generalized for all k in 1974. Later, Furstenberg was able to prove the theorem using ergodic theory. A third proof is due to Gowers in 2001 in which the methods of Roth are generalized.

Theorem 3.1.2. Szeméredi's Theorem. *Let k be a positive integer and let $\delta \geq 0$. Then there exists $N = N(k, \delta)$, such that every subset A of $\{1, ..., N\}$ such that $|A| \geq \delta N$ must contain an arithmetic progression of length k.*

Roth's proof of the case $k = 3$ also provided a quantitative bound on the size of N in relation to the density δ of the subset A. We define $N(\delta, k)$ to be how large N must be to guarantee that any subset $A \subset \{0, 1, ..., N-1\}$ with density δ contains an arithmetic progression of length k. In Roth's proof, he was able to show that $N(\delta, 3) = C\exp(\exp(c\delta^{-1}))$. Szemerédi's proof gave the bound $N(\delta, 3) = C_1 \exp(\delta^{-C_2})$. In 1999, Bourgain gave the best bound known with $N(\delta, 3) \leq C\exp(\delta^{-2}\log^2(\delta^{-1}))$.

Szemerédi's proof of progression of length k also contained a quantitative bound on $N(\delta, k)$, although it is extremely difficult to write down. Furstenberg's methods gave no bounds at all. Gowers' proof did significantly improve Szemerédi's bound. Gowers showed that $N(\delta, k) = 2^{2^{\delta^{-2^{2^{k+9}}}}}$.

The next natural question is to find a lower bound for $N(\delta, 3)$. The best bound known to date was given by Behrend [2] in 1946. In his argument he constructed a subset of $\{0, 1, ..., N-1\}$ of size $N^{1-\frac{\sqrt{2}\log 2}{\sqrt{\log N}}+\frac{\epsilon}{\sqrt{\log N}}}$ without arithmetic progression of length three. In other words, he showed that $N(\delta, 3) > c_1 \exp(c_2 \log^2(\delta^{-1}))$. His construction is based on the fact that higher dimensional spheres are convex, and so any line which passes through the sphere can intersect the sphere at most twice. There is also a result for longer arithmetic progressions due to Rankin [12]. He showed that for $k > 1$, there are subsets of $\{1, ..., N\}$ of cardinality at least $N \exp(-C(logN)^{1/(k+1)})$ that do not contain arithmetic progressions of length $1 + 2^k$.

The remainder of this chapter follows Gowers' paper "A New Proof of Szemerédi's Theorem" [4].

3.2 Roth's Theorem

Before presenting the details of Roth's theorem, it is useful to give an outline of the proof. We begin with $A \subset \{0, 1, ..., N-1\}$ with $|A| = \delta N$. If A is "random", then we will be able to show that A contains many arithmetic progressions of length three. On the other hand, if A is not "random", then we are able to choose a subset $A' \subset A$ which has higher density in some subprogression of the integers. We must then determine if A' is "random" or not, and we iterate the argument until the theorem is proved. For this argument to work, we must count the number of necessary steps to ensure that we do not eliminate too many elements for the number of steps needed. Eventually we will have a set contained in an arithmetic progression with density one and at least three elements.

We recall that for any set $A \in \mathbb{Z}_N$ we can define its **characteristic function**

$$A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases}$$

To the characteristic function, we associate the **balanced function** f of A defined to be $f(x) = A(x) - \delta$.

As a final note before we begin the details of the proof, we will need to distinguish between an arithmetic progression in the integers (an arithmetic $\mathbb{Z}$-progression) and a progression modulo N (an arithmetic $\mathbb{Z}_N$-progression).

We will say that a set A is random if the Fourier coefficients of the balanced function are small.

Definition 3.2.1. *We say that the mapping f of $\mathbb{Z}_N$ to the closed unit disk in $\mathbb{C}$ is α-uniform if $|\widetilde{f}(r)| \leq \alpha N$ for all $r \neq 0$. If f is the balanced function of $A(x)$ then we will say that A is α-**uniform** or **random**.*

We will now state and prove results which will be useful when a Fourier coefficient is large, in other words, when A is not random.

Lemma 3.2.2. *Assume N is a sufficiently large prime. Pick integers r and s such that $0 \leq r \leq N$ and $0 < s \leq N-1$. Then the set $\{0, 1, ..., N-1\}$ can be partitioned into arithmetic $\mathbb{Z}$-progressions P_j such that (1) $\sqrt{s}/2 \leq |P_j| \leq \sqrt{s}$ and (2) if $x_1, x_2 \in P_j$, then $|x_1 r - x_2 r| \leq 2s$ in $\mathbb{Z}_N$.*

Proof: Partition the interval $[0, N-1]$ into $N/2\sqrt{s}$ intervals of equal length. Consider the set $S = \{0, r, 2r, ..., \lfloor N/\sqrt{2s} \rfloor\}$ modulo N. Then, by the pigeonhole principle, there must be two elements from S, kr and lr, which lie in the same interval. Assume that $k \geq l$. Then $|kr - lr| \leq \sqrt{4s}$. Set $u = k - l$. Now we consider the set $\{0, 1, ..., N-1\}$ modulo u. Each residue class will have $\lfloor N/u \rfloor$ or $\lceil N/u \rceil$ elements. Now, $\lceil N/u \rceil \geq \lfloor N/u \rfloor \geq N/(N/\sqrt{2s}) = \sqrt{2s}$. Now will simply divide each residue class into subprogressions, P_j of sequential elements with the desired length; $\sqrt{s}/2 \leq |P_j| \leq \sqrt{s}$. Then given $x_1, x_2 \in P_j$ we have $|x_1 r - x_2 r| \leq \sqrt{s} \cdot ru \leq \sqrt{s} \cdot \sqrt{4s} = 2s$ as desired. □

Lemma 3.2.3. *Let N be a sufficiently large odd integer and let f be a function such that $f : \{0, 1, ..., N-1\} \to \{z : |z| \leq 1\}$. Assume that $|\widetilde{f}(r)| \geq \alpha N$ for some $r \neq 0$. Then the set $\{0, 1, ..., N-1\}$ can be partitioned into arithmetic $\mathbb{Z}$-progressions P_j such that $|P_j| \geq \sqrt{\alpha N/32\pi}$ and $\sum_j |\sum_{x \in P_j} f(x)| \geq \alpha N/2$.*

Proof: Set $s = \alpha N/8\pi$ and apply lemma 3.2.2 in order to partition $\{0, 1, ..., N-1\}$ into arithmetic progressions P_j such that $\sqrt{\alpha N/32\pi} \leq |P_j|$ and given x_1 and $x_2 \in P_j$ we have $|x_1 r - x_2 r| \leq 2s = \alpha N/4\pi$ in $\mathbb{Z}_N$.

By assumption and definition of the Fourier transform, we have

$$\begin{aligned}\alpha N \leq |\widetilde{f}(r)| &= |\sum_{x=1}^{N-1} f(x)e(-xr/N)| \\ &= |\sum_j \sum_{x\in P_j} f(x)e(-xr/N)| \\ &\leq \sum_j |\sum_{x\in P_j} f(x)e(-xr/N)| \\ &\leq \sum_j |\sum_{x\in P_j} f(x)| + \sum_j |P_j| \cdot \alpha/2\end{aligned}$$

To see this, we estimate the inside sum for fixed j as follows.

$$\begin{aligned}&|\sum_{x\in P_j} f(x)e(-rx/N)| \\ &\quad = |\sum_{x\in P_j} f(x)(e(-x_jr/N) + e(-xr/N) - e(-x_jr/N))| \\ &\quad \leq |\sum_{x\in P_j} f(x)e(-x_jr/N)| + |\sum_{x\in P_j} f(x)(e(-xr/N) - e(-x_jr/N)| \\ &\quad \leq |\sum_{x\in P_j} f(x)| + |P_j| \cdot \max_{x\in P_j} |f(x)| \cdot \max_{x\in P_j} |e(-xr/N) - e(-x_jr/N)| \\ &\quad \leq |\sum_{x\in P_j} f(x)| + |P_j| \cdot \alpha/2\end{aligned}$$

Here we use the fact that $|xr - x_jr| \leq \alpha N/4\pi$ by lemma 3.2.2 to show $|e(-xr/N) - e(-x_jr/N)| \leq \alpha/2$.

Therefore we have proved $\alpha N - \alpha N/2 = \alpha N/2 \leq \sum_j |\sum_{x\in P_j} f(x)|$ as desired. □

Corollary 3.2.4. *Let $A \subset \{0, 1, ..., N-1\}$ such that $|A| = \delta N$. Assume that $|\widetilde{A}(r)| \geq \alpha N$ for some $r \neq 0$ and $\alpha = \delta^2/10$. Then there exists an arithmetic $\mathbb{Z}$-progression $P = \{a, a+u, ..., a+mu\}$ of length at least $\sqrt{\delta^2 N/320\pi}$ such that $\frac{|A\cap P|}{|P|} \geq \delta + \delta^2/40$.*

Proof: Assume $r \neq 0$. Consider the balanced function $f(x) = A(x) - \delta$. We first note that $\widetilde{f}(r)$ and $\widetilde{A}(r)$ are equal since

$$\begin{aligned}\widetilde{f}(r) &= \sum_{x=0}^{N-1}(A(x)-\delta)e(-xr/N)\\ &= \sum_{x=0}^{N-1}A(x)e(-xr/N) - \delta\sum_{x=0}^{N-1}e(-xr/N)\\ &= \widetilde{A}(r)\end{aligned}$$

We apply lemma 3.2.3 to partition $\{0,1,...,N-1\}$ into arithmetic progressions P_j such that $|P_j| \geq \sqrt{\alpha N/32\pi} = \sqrt{\delta^2 N/320\pi}$ and $\sum_j |\sum_{x\in P_j} f(x)| \geq \alpha N/2 = \delta^2 N/20$. Since

$$\begin{aligned}\sum_{x=0}^{n-1} f(x) &= \sum_{x=0}^{N-1}(A(x)-\delta)\\ &= \sum_{x=0}^{N-1}A(x) - \sum_{x=0}^{N-1}\delta\\ &= |A| - \delta N\\ &= 0,\end{aligned}$$

we have

$$\begin{aligned}\sum_j |\sum_{x\in P_j} f(x)| + \sum_{x=0}^{n-1} f(x) &= \sum_j \Big(|\sum_{x\in P_j} f(x)| + \sum_{x\in P_j} f(x)\Big)\\ &\geq \delta^2 N/20\end{aligned}$$

Therefore, there must be at least one j such that $|\sum_{x\in P_j} f(x)| + \sum_{x\in P_j} f(x) \geq \delta^2|P_j|/20$. This implies that $\sum_{x\in P_j} f(x) \geq \delta^2|P_j|/40$ since $|a|+a \geq b \geq 0 \Longrightarrow a \geq b/2$. Thus, for this chosen j, we have

$$\begin{aligned}|A\cap P_j| &= \sum_{x\in P_j} A(x)\\ &= \sum_{x\in P_j}(f(x)+\delta)\\ &= \sum_{x\in P_j} f(x) + \delta|P_j|\\ &\geq \delta^2|P_j|/40 + \delta|P_j|\end{aligned}$$

as desired. □

Lemma 3.2.5. *Bertrand's Postulate: If $n \geq 1$ then there is a prime p such that $n \leq p \leq 2n$.*

We are now ready to prove Roth's theorem.

Theorem 3.2.6. *Let $\delta > 0$, let $N \geq \exp\exp(C\delta^{-1})$ (where C is an absolute constant) and let $A \subset \{1, 2, ..., N\}$ be a set of size at least δN. Then A contains an arithmetic progression of length three.*

Proof: As previously outlined, we will prove Roth's theorem by considering the distribution of the set A. We recall that we will show that if A is uniform, then we can find a progression of length three. If A is not uniform, then we are able to find a subprogression of N where the density of A is higher. We iterate this argument, until we have a density of one. Finally, we will show that N is large enough to perform the number of iterations necessary to reach a density of one and to have at least 3 elements remaining in our set.

Assume N_0 is a large postive integer and let $\delta_0 > 0$. Assume $A_0 \subset \{0, 1, ..., N_0 - 1\}$ such that $|A_0| \geq \delta_0 N_0$. We will need N to be prime and since our argument is iterative, this argument allows us to choose N prime in each step. Applying lemma 3.2.5, we choose a prime $N \in [N_0/3, 2N_0/3]$. Let $A = A_0 \cap \{0, 1, ..., N - 1\}$.

Case 0: If $|A| \leq \delta_0(1 - \delta_0/160)N$, then we have

$$\begin{aligned} |A_0 \cap \{N, ..., N_0 - 1\}| &\geq |A_0| - |A| \\ &\geq \delta_0(N_0 - (1 - \delta_0/160)N) \\ &= \delta_0((N_0 - N) + \delta_0 N/160) \\ &\geq \delta_0(1 + \delta_0/320)(N_0 - N) \end{aligned}$$

Set $\delta = \delta_0(1 - \delta_0/160)$.

Case 1: Assume $|A| = \delta N$. Let $B = A \cap [N/3, 2N/3]$ and assume $|B| \leq \delta N/5$. In this case, we have a small density in the middle third of the set $\{0, 1, ..., N - 1\}$ which must mean that A has a higher density in at least of of the other intervals $[0, N/3)$ or $[2N/3, N - 1]$. In this case, without loss of generality, $|A \cap [0, N/3)| \geq 2\delta N/5 = 6\delta/5 \cdot N/3$.

Case 2: Let $\alpha = \delta^2/10$. Assume that A is not α-uniform. Then there exists $r \neq 0$ such that $|\widetilde{A}(r)| \geq \alpha N$. Here we have satisfied the hypothesis of corollary 3.2.4 and we know that there is an arithmetic progression such that A has higher density in this progression than in our original set. Specifically, the progression will have length at least $\sqrt{\delta^2 N/320\pi}$ and $|A \cap P|/|P| \geq \delta + \delta^2/40$. This will be the basis for our iteration argument. We will consider $(A \cap P) \subset P$.

Case 3: Assume that none of the previous cases holds. In this case we are able to show that A contains an arithmetic progression of length three. By assumption, we know that $|\widetilde{A}(r)| \leq \alpha N$ for each nonzero $r \in \mathbb{C}$.

We would like to put a lower bound on the number of progressions of length three. We first note that an arithmetic progression modulo N is not necessarily an arithmetic $\mathbb{Z}$-progression (consider 10,12,1 modulo 13 for example). However, the number of arithmetic $\mathbb{Z}$-progressions $(x, y, z) \in \mathbb{Z}^3$ must be greater than or equal to the number of arithmetic progressions restricted to $(x, y, z) \in A \times B \times B$ where $B = A \cap [N/3, 2N/3]$ as defined above. The triple $(x, y, z) \in A \times B \times B$ is an arithmetic progression modulo N if and only if $x + z = 2y$ modulo N. We will now estimate the number of such triples.

$$\begin{aligned}
\sum_{(x,y,z)\in A\times B^2, x+z=2y} 1 &= N^{-1} \sum_{x\in A}\sum_{y\in B}\sum_{z\in B}\sum_{r=0}^{N-1} e(r(2y-x-z)/N) \\
&= N^{-1}\sum_{r=0}^{N-1} \widetilde{A}(r)\widetilde{B}(-2r)\widetilde{B}(r) \\
&\geq N^{-1}|A||B|^2 \\
&\quad - N^{-1}\max_{r\neq 0}|\widetilde{A}(r)|(\sum_{r=1}^{N-1}|\widetilde{B}(-2r)|^2)^{1/2}(\sum_{r=1}^{N-1}|\widetilde{B}(r)|^2)^{1/2} \\
&\geq \delta|B|^2 - \alpha|B|N.
\end{aligned}$$

Since our technique has counted trivial progressions ($x = y = z$), and there are $|B|$ such progressions, we want to show that $\delta|B|^2 - \alpha|B|N \geq |B|+1$. Now we use the fact that $|B|$ is at most N and $|B|$ is at least $\delta N/5$. Therefore, we can conclude that $N \geq (50 + \sqrt{2500 + 4\delta^3})/2\delta^3$.

In Case 3, we are able to produce an arithmetic progression of length three. Now we will begin our iteration argument for the other three cases.

First we must establish a method for finding subprogressions P such that $A \cap P$ has a higher density in P. The result of corollary 3.2.4 tells us that in Case 2, there exists an arithmetic progression P of $\{0, 1, ..., N-1\}$ of cardinality at least $\sqrt{\delta^2 N/320\pi}$ and $|A \cap P|/|P| \geq \delta + \delta^2/40$. We now apply this to our original progression $\{0, 1, ..., N_0 - 1\}$ by noting that $\delta(1+\delta/40) \geq \delta_0(1 + \delta_0/320)$. Now since N_0 is at most three times N, we know that this subprogression P must be at least of cardinality $\alpha N_0/96\pi$ and $|A_0 \cap P| \geq \delta_0(1 + \delta_0/320|P|)$. In case 0 we can take $A_0 \cap \{N+1, ..., N_0 - 1\}$ such that $|A_0 \cap \{N+1, ..., N_0 - 1\}|/|\{N, ..., N_0 - 1\}| \geq \delta_0 + \delta_o^2/320$. In case 1, we assume without loss of generality that $|A \cap [0, N/3)|/(N/3) \geq 6\delta/5$.

We now begin our iteration argument. We need to be sure of two conclusions: that it is possible to reach a density for which an arithmetic progression is guaranteed (in our case we will actually reach density one, and that our N is large enough so that we do not run out of possible subprogressions. Beginning with our first step, we start with a density of δ_0. Then in each subsequent iteration, we have the density increasing by at least $\delta_0^2/320$. Then we will reach a density of $2\delta_0$ after at most $320\delta_0^{-1}$ steps. Now, at any point where we satisfy case 3, we can stop. However, there is no way to guarantee this. Instead we will reach density one. Reaching a density of $2\delta_0$ after at most $320\delta_0^{-1}$ steps, we can then see that we will reach a density of $4\delta_0$ after at most $320(2\delta_0^{-1})$ further steps. In general, at step m where the density is δ_m, we will reach a density $2\delta_m$ after at most $320\delta_m^{-1}$ further steps. We now calculate the maximum number of steps required to reach a density of one: $320(1/\delta_0 + 1/2\delta_0 + 1/4\delta_0 + ...) = 640/\delta_0$. Now, we ask what bound we must put on N in order that this number of steps makes sense. First observe that at each step of the iteration the size of the subprogression chosen is about the square root of the progression of the previous step. Therefore after the first step we go from a progression of length N_0 to a progression of at least length $(c_1 N_0)^{1/2}$. Then after $640\delta^{-1}$ steps we will have a length of at least $cN_0^{1/(2^{640\delta_0^{-1}})}$. Now, if we are still to have a progression of length three, then we must have $cN_0^{1/(2^{640\delta_0^{-1}})} \geq 3$. This is equivalent to showing $N_0^{1/(2^{640\delta_0^{-1}})} \geq 3/c$. Taking the log of both sides, we have $\log N_0 \geq 2^{640\delta_0^{-1}} \cdot 3/c$. Therefore, we must have $N_0 \geq e^{c_1^{640\delta_0^{-1}}} \geq e^{e^{\log c_1 \cdot 640\delta_0^{-1}}}$. $\square$

3.3 Szemerédi's Theorem

Now that we have proved Roth's theorem, we will be able to discuss Gowers' generalization of the proof used to prove Szemerédi's theorem. The outline of Gowers' proof mirrors that of Roth's theorem. If a subset A of $\{0, 1, ..., N-1\}$ is 'random', then one can show that A must contain a progression of length k. Otherwise, there is a subprogression, $P \subset \{0, 1, ..., N-1\}$ such that $|A \cap P|/|P| \geq \delta$. The argument can then be iterated as in Roth's theorem.

The most obvious way to generalize would be to use α-uniformity to show that A contains an arithmetic progression of length k. Unfortunately the information given by α-uniformity does not seem to be sufficient to find progressions of length greater than three using known methods of estimation. Gowers' idea is to use a stricter definition of 'randomness' which lends itself to finding progressions of arbitrary length. However, this definition will of course complicate the iteration argument in the non-random case since fewer subsets of $\{0, 1, ..., N-1\}$ will satisfy the new definition of random.

We give here the definition of higher degree uniformity for all k, however, we will focus on the case with $k = 4$ in Szemerédi's theorem.

Definition 3.3.1. *Assume $f : \mathbb{Z}_N \to [-1, 1]$ and let $x \in \mathbb{Z}_{\mathbb{N}}$. Then we define the difference function $\Delta(f; x)$ to be*

$$\Delta(f; x)(s) = f(s)f(s - x).$$

Given a set $\{x_1, x_2, ..., x_d\}$ one can iterate the difference function. In the case $d = 2$ (which we will need for progressions of length four) this iteration is straightforward:

$$\Delta(f; x_1, x_2)(s) = \Delta(\Delta(f; x_1); x_2)(s) = f(s)f(s - x_1)f(s - x_2)f(s - x_1 - x_2) \tag{3.3.1}$$

In general we have $\Delta(f; x_1, ..., x_d) = \Delta(\Delta \cdots (\Delta(f; x_1) \cdots ; x_d)$.

Definition 3.3.2. *Let f be a function which maps $\mathbb{Z}_N$ to $[-1, 1]$. Then we say that f is α-uniform of degree d if*

$$\sum_{x_1,...,x_d} |\sum_{s \in \mathbb{Z}_n} \Delta(f; x_1, ..., x_d)(s)|^2 \leq \alpha N^{d+2}.$$

In the case where $d = 2$ we say that f is quadratically α-uniform.

Given $d = 1$ this definition means

$$\sum_x |\sum_s \Delta(f;x)|^2 = \sum_x |\sum_s f(s)f(s-x)|^2 < \alpha N^3.$$

One can see that this definition implies α-uniformity as defined for Roth's theorem since

$$\sum_x |\sum_s f(s)f(s-x)|^2 = \sum_{a-b=c-d} f(a)f(b)f(c)f(d) = N^{-1}\sum_r |\widetilde{f}(r)|^4$$

which is easy to verify and follows from Chapter 2.

We will give a more detailed outline of the proof that the set $A \subset \{0, 1, ..., N-1\}$ which is α-uniform of degree 2 and satisfies the hypotheses of Szemerédi's theorem must contain an arithmetic progression of length four. Before this outline, we would like to make a few comments regarding the non-uniform case. Here we assume that the subset $A \subset \{0, 1, ..., N-1\}$ is not quadratically α-uniform. Precisely, this definition states that there are at least αN integers k such that there exists $r \in \mathbb{Z}_N$ such that

$$|\sum_x A(x)A(x-k)e(-rx/N)| \geq \alpha N. \tag{3.3.2}$$

Denote the set of such k by B and define the function $\phi : B \to \mathbb{Z}_N$ by $\phi(k) = r$ (if more than one r satisfies equation 3.3.2 we can just pick one such r for a given k). One would then like to show that the function ϕ satisfies some sort of linear properties. To do this, one shows that there exist $\alpha^4 N^3$ "additive quadruples" $(a, b, c, d) \in B^4$ such that $a + b = c + d$ and $\phi(a) + \phi(b) = \phi(c) + \phi(d)$. One then uses Gowers' quantitative version of the Balog-Szemerédi theorem to show that there exists an long arithmetic progression modulo N on which $\phi(s) = \lambda s + \mu$ for many $s \in B$. By restricting ϕ to this large arithmetic progression, one is able to iterate the argument in a similar manner to the proof of Roth's theorem.

Now we return to the α-uniform case. As before, we denote the characteristic function of the set A by $A(x)$ and the balanced function associated with A by $f(x)$. We say that A is quadratically α-uniform if f is quadratically α-uniform.

Now we will use the characteristic function of A to count arithmetic progressions. If $A(x)A(x-y)A(x-2y)A(x-3y)$ is one for some x and y then we clearly have an arithmetic progression. We will sum over all x and y modulo

N to count the total number of progressions. Here we will assume that N is prime.

The number of $\mathbb{Z}_N$ progressions contained in A is given by

$$\sum_x \sum_y A(x)A(x-y)A(x-2y)A(x-3y)$$

$$= \sum_x \sum_y (f(x)+\delta)(f(x-y)+\delta)(f(x-2y)+\delta)(f(x-3y)+\delta)$$

$$= N^2\delta^4$$

$$+ \delta\Big(\sum_x \sum_y f(x)f(x-y)f(x-2y) \tag{3.3.3}$$

$$+ \sum_x \sum_y f(x)f(x-2y)f(x-3y) \tag{3.3.4}$$

$$+ \sum_x \sum_y f(x)f(x-y)f(x-3y) \tag{3.3.5}$$

$$+ \sum_x \sum_y f(x-y)f(x-2y)f(x-3y)\Big) \tag{3.3.6}$$

$$+ \sum_x \sum_y f(x)f(x-y)f(x-2y)f(x-3y), \tag{3.3.7}$$

where we use the fact that $\sum_x f(x) = 0$.

Using this method we count progressions modulo N as well as the trivial progressions. This is a technicality similar to the one in the proof of Roth's theorem. What we have in the above sum is the expected number of progression, $N^2\delta^4$, plus an error term which we would like to show is small. Ignoring the modulo N technicality, we will show that the magnitude of each of the terms, (3.3.3), (3.3.4), (3.3.5), (3.3.6) and (3.3.7), is small, therefore forcing a positive number of progressions. This will give a rough outline of Gowers' method to show that higher degree α-uniformity guarantees progressions of length four. This problem can be generalized for progressions of arbitrary length using the same techniques.

Lemma 3.3.3. *Assume $f : \mathbb{Z}_N \to [-1,1]$ and f is α-uniform of degree d. Then there exists a function $\beta : \mathbb{Z}_N \to [0,1]$ such that $\sum_{x\in\mathbb{Z}_N} \beta(x) = \alpha N$ and $\Delta(f;x)$ is $\beta(x)$-uniform of degree $d-1$.*

Proof: We prove this for the case when f is quadratically α-uniform. For $\Delta(f;x)$ to be $\beta(x)$-uniform of degree 1, we would like to show that

$$\sum_a \Big(\sum_s \Delta(\Delta(f;x);s)\Big)^2 = \sum_a \Big(\sum_s f(s)f(s-x)f(s-a)f(s-x-a)\Big)^2 \leq \beta(x)N^3$$

where $\sum_x \beta(x) = \alpha N$. But we know that

$$\sum_x \sum_a \Big(\sum_s f(s)f(s-x)f(s-a)f(s-x-a)\Big)^2 \leq \alpha N^4$$

since f is quadratically α-uniform. Therefore, we could take

$$\beta(x) = \begin{cases} N^{-3}\sum_a \big(\sum_s \Delta(f;x;s)\big)^2 \text{ if } x \neq 0 \\ \alpha N - N^{-3}\sum_{x\neq 0}\sum_a \big(\sum_s \Delta(f;x,s)\big)^2 \text{ otherwise,} \end{cases}$$

to prove the lemma. □

We present the following theorem exactly as it appears in Gowers paper [4]. There does not appear to be an advantage to treating the theorem separately for the case $k = 4$.

Theorem 3.3.4. *Let $k \geq 2$ and let $f_1, ..., f_k$ be functions from $\mathbb{Z}_N \to [-1,1]$ such that f_k is α-uniform of degree $k-2$. Then*

$$\Big|\sum_{x\in\mathbb{Z}_N}\sum_{y\in\mathbb{Z}_N} f_1(x)f_2(x-y)...f_k(x-(k-1)y)\Big| \leq \alpha^{1/2^{k-1}}N^2.$$

We can see that using this lemma, we will be able to bound each of the terms (3.3.3), (3.3.4), (3.3.5) and (3.3.6) using linear uniformity (ie. α-uniformity of degree 1). For example, the sum (3.3.4) can be estimated using $g(x) = f(x-y)$ to give

$$\sum_x \sum_y f(x)f(x-2y)f(x-3y) = \sum_x \sum_y f(x)g(x-y)g(x-2y),$$

which we can estimate using theorem 3.3.4. The other term, (3.3.7), we estimate directly from theorem 3.3.4 using quadratic α-uniformity.

Proof of theorem 3.3.4: We will prove this theorem by induction. For the case $k = 2$ we have

$$\left|\sum_x \sum_y f_1(x) f_2(x-y)\right| = \left|\left(\sum_x f_1(x)\right)\left(\sum_u f_2(u)\right)\right| \leq \alpha^{1/2} N^2,$$

since $|\sum_x f_1(x)|$ is trivially bounded by N and $|\sum_u f_2(u)| \leq \alpha^{1/2} N$ using α-uniformity of degree zero.

For $k > 2$ we assume that the inequality holds for $k-1$. Since we assume that f_k is α-uniform of degree $k-2$, by lemma 3.3.3 we know that there exists a function $\beta : \mathbb{Z}_N \to [0,1]$ such that $\sum_x \beta(x) = \alpha N$ and $\Delta(f;x)$ is $\beta(x)$-uniform of degree $k-3$. Now we would like to have a form to which we can apply this inductive hypothesis.

$$\begin{aligned}
&\left|\sum_{x\in\mathbb{Z}_N}\sum_{y\in\mathbb{Z}_N} f_1(x) f_2(x-y)...f_k(x-(k-1)y)\right|^2 \\
&\leq N\sum_x \left|\sum_y f_1(x) f_2(x-y)...f_k(x-(k-1)y)\right|^2 \\
&\leq N\sum_x \left|\sum_y f_2(x-y) f_3(x-2y)...f_k(x-(k-1)y)\right|^2 \\
&= N\sum_x\sum_y\sum_u f_2(x-y) f_2(x-u)...f_k(x-(k-1)y) f_k(x-(k-1)u) \\
&= N\sum_x\sum_y\sum_v f_2(x) f_2(x-v)...f_k(x-(k-2)y) f_k(x-(k-2)y-(k-1)v) \\
&= N\sum_x\sum_y\sum_v \Delta(f_2;v)(s)\Delta(f_3;2v)(x-y)...\Delta(f_k;(k-1)v)(x-(k-2)y) \\
&\leq N\sum_v \beta((k-1)v)^{1/2^{k-2}} N^2.
\end{aligned}$$

Therefore,

$$\left|\sum_{x\in\mathbb{Z}_N}\sum_{y\in\mathbb{Z}_N} f_1(x) f_2(x-y)...f_k(x-(k-1)y)\right| \leq \alpha^{1/2^{k-1}} N^2$$

using the induction hypothesis. □

This shows that each of the terms must be small as desired. We hope that this brief introduction to Gowers' proof will entice the reader to read his paper [4].

Chapter 4

Weyl's inequality

4.1 History

In this chapter we will state Weyl's inequality, prove a special case of the theorem and also give a simple application.

Although Weyl's inequality has many applications, our primary reason for including it here is for its importance in applying the circle method, which we will use to prove Vinogradov's three-primes theorem. We will briefly describe the circle method here, however we leave the details for the next chapter. Hardy and Littlewood first developed the circle method in the early 1920s. They used it to prove Waring's problem, described in the introduction. It was soon realized that their method would have many applications, and the technique was refined most notably by Vinogradov, Vaughan and Wooley. Vinogradov's refinement enabled integration over the interval $[0, 1]$ rather than the original circle of integration used by Hardy and Littlewood.

In the case of Waring's problem, we denote the number of representations of N as a sum of s positive k^{th} powers (ie $N = x_1^k + ... + x_s^k$) by $r_{k,s}(N)$. The first step in the circle method is to write $r_{k,s}(N)$ as an integral, which one then estimates. In this case, the goal is to show that the integral must be positive. In estimating the integral, one divides the bounds of integration into two disjoint sets called the major arcs and the minor arcs. One then shows that the integral over the major arcs is positive and is the main contribution to the entire estimate. The integral over the minor arcs is shown to contribute a small error term. It is in the estimate of the integral over the minor arcs that one is able to apply Weyl's inequality.

We will also use Weyl's inequality to prove that the set $S = \{\{P(n)\} : n = 1, 2, ...\}$ defined by the polynomial $P(n) = \alpha n^2 + \beta n + \gamma$ is uniformly distributed in the unit interval whenever α is irrational. Weyl's inequality is a useful tool for proving a set is uniformly distributed because Weyl's criterion gives a set of equivalent conditions for a set to be uniformly distributed.

The primary source used in this section is Montgomery's *Ten Lectures*

on the Interface Between Analytic Number Theory and Harmonic Analysis [9]. We also consulted Nathanson [10] and took examples in the applications section 4.3 from a course at the University of British Columbia taught by Izabella Łaba in the spring of 2004.

4.2 Weyl's inequality

Theorem 4.2.1. ***Weyl's inequality***
Let $P(x) = \sum_{j=0}^{k} \alpha_j x^j$ *where* $\alpha_i \in \mathbb{R}$ *for each* i *and* $|\alpha_k - a/q| \leq q^{-2}$ *for some* a, $q \in \mathbb{Z}$ *and* $(a, q) = 1$. *Then*

$$\sum_{n=1}^{N} e(P(n)) \leq C_{k,\epsilon} N^{1+\epsilon}(q^{-1} + N^{-1} + qN^{-k})^{2^{1-k}}.$$

We will prove Weyl's inequality for $P(x) = \alpha^2 + \beta x + \gamma$. The following lemma will be used for Weyl's inequality and in proving Vinogradov's three primes theorem. We will also prove that given $\alpha \notin \mathbb{Q}$ there exists a sequence $q_1, q_2, ...$ such that $\lim_{n\to\infty} q_N$ and $q_N \leq N$ such that $|\alpha - a_N/q_N| \leq q_N^{-2}$ in lemma 4.3:4 which allows us to make use of Weyl's inequality.

Lemma 4.2.2. *Let* α, $\beta \in \mathbb{R}$ *and let* M, $N \in \mathbb{N}$. *Then*

$$|\sum_{n=M}^{N} e(\alpha n + \beta)| \leq \min\{N, \frac{1}{2||\alpha||}\}.$$

Proof: The sum is trivially bounded by N since

$$|\sum_{n=M}^{N} e(\alpha n + \beta)| \leq \sum_{n=1}^{N} |e(\alpha n + \beta)| \leq N.$$

Since $e(\alpha n + \beta) = e(\alpha n)e(\beta)$ we can take $\beta = 0$. To show the other bound, we have

$$\begin{aligned}\sum_{m\le u\le n}|e(\alpha u)| &= |e(\alpha m)\sum_{0\le u\le n-m}e(\alpha u)|\\ &= |\frac{e(\alpha(n-m+1))-1}{e(\alpha)-1}|\\ &\le \frac{2}{|e(\alpha)-1|}\\ &= \frac{2}{|e(\alpha/2)-e(-\alpha/2)|}\\ &= \frac{2}{|2i\sin\pi\alpha|}\\ &= |\sin\pi\alpha|^{-1}\\ &\le (2||\alpha||)^{-1}.\Box\end{aligned}$$

Theorem 4.2.3. ***Weyl's inequality for quadratic polynomials*** *Let* $P(x)=\alpha^2+\beta x+\gamma$ *where* $\alpha\in\mathbb{R}$. *Assume* $|\alpha-a/q|\le q^{-2}$, $a,\ q\in\mathbb{Z}$ *and* $(a,q)=1$. *Define* $S=\sum_{n=1}^{N}e(P(n))$. *Then*

$$|S|\ll(N/\sqrt{q}+\sqrt{N\log q}+\sqrt{q\log q})$$

Proof: To prove Weyl's inequality, we will estimate $|S|^2$ as follows. By definition,

$$\begin{aligned}
|S|^2 &= |\sum_{n=1}^{N} e(P(n))|^2 \\
&= \sum_{n,m=1}^{N} e(P(m) - P(n)) \\
&= \sum_{n=1}^{N} \sum_{h=1-n}^{N-n} e(P(n+h) - P(n)) \\
&= \sum_{h=1-N}^{N-1} \sum_{1\leq n\leq N, 1-n\leq h\leq N-h} e(P(n+h) - P(n)) \\
&= N + \sum_{h=1}^{N-1} \sum_{n=1}^{N-h} (e(P(n+h) - P(h)) + e(P(n) - P(n+h)) \\
&= N + \sum_{h=1}^{N-1} \sum_{n=1}^{N-h} 2\mathrm{Re}(e(P(n+h) - P(n))) \\
&\leq N + 2\sum_{h=1}^{N-1} |\sum_{n=1}^{N-h} e(P(n+h) - P(n))|
\end{aligned}$$

Noticing that $P(n+h) - P(n) = 2\alpha hn + \alpha h^2 + \beta h = \alpha' n + \beta'$ where we take $\alpha' = 2\alpha h$ and $\beta' = \alpha h^2 + \beta h$, we can apply lemma 4.2.2 which gives us the bound

$$\begin{aligned}
|S|^2 &\leq N + 2\sum_{h=1}^{N-1} \min\{N-h, (2||2\alpha h||)^{-1}\} \\
&\leq N + 2\sum_{h=1}^{2N} \min\{N, ||h\alpha||^{-1}\}.
\end{aligned}$$

We will consider the case when α is rational and the case when α is not rational.

Case 1: If $\alpha = a/q$ where $(a, q) = 1$ and $a,\ q \in \mathbb{Z}$, then we have

$$\sum_{n=1}^{q} \min\{N, ||h\alpha||^{-1}\} = \sum_{h=1}^{q} \min\{N, ||h(a/q)||^{-1}\}$$

$$\begin{aligned}
&= N + \sum_{h=1}^{q-1} \min\{N, ||h(a/q)||^{-1}\} \\
&\leq N + \sum_{h=1}^{q-1} ||h(a/q)||^{-1} \\
&= N + \sum_{h=1}^{q-1} ||h/q||^{-1} \\
&\leq N + 2\sum_{h=1}^{q-1} q/h \\
&\leq N + Cq\log q
\end{aligned}$$

The fourth step in the above estimates holds since $\{a, 2a, ..., (q-1)a\}$ give a complete list of residues modulo q. We also note that the same bound works when we consider any sum $\sum_{h=kq+1}^{(k+1)q-1}$. We now partition $1, ..., 2N$ into blocks of length q. We will have less than or equal to $2N/q + 1$ such blocks. Therefore, we have

$$\begin{aligned}
|S|^2 &\leq N + (2N/q + 1)(N + Cq\log q) \\
&\ll N^2/q + N\log q + q\log q.
\end{aligned}$$

Finally, taking the square root, we have

$$|S| \ll N/\sqrt{q} + \sqrt{N\log q} + \sqrt{q\log q}.$$

Case 2: We assume $|\alpha - a/q| \leq q^{-2}$, a, $q \in \mathbb{Z}$ and $(a, q) = 1$.

As in case 1, it will be useful to split the sum we wish to estimate into separate sums. In this case, we fix a block of q integers, $M < n \leq M + q$. Write $\alpha = a/q + r$ where $r \leq 1/q^2$. Then for every u there are at most 3 choices of n in the specified interval such that $||n\alpha - u|| \leq 1/2q$. To see this, we will assume $||n\alpha - u|| \leq 1/2q$ and prove that there are at most three possibilities. Let $n = M + m$, $v = u - M\alpha$ and $1 \leq m \leq q$. Then

$$\begin{aligned}
||m\alpha - v|| &= ||M\alpha + m\alpha - v - M\alpha|| \\
&= ||(M + m)\alpha - v + v - u|| \\
&= ||n\alpha - u|| \leq 1/2q
\end{aligned}$$

by assumption. Further, $||mr|| \leq q|r| \leq 1/q$. Therefore,

$$\begin{aligned} ||ma/q - v|| &= ||ma/q + mr - mr - v|| \\ &= ||m\alpha - v - mr|| \\ &\leq ||m\alpha - v|| + ||mr|| \\ &\leq 1/2q + 1/q = 3/2q. \end{aligned}$$

Therefore, there are at most three distinct choices of m and hence at most three distinct choices of n for which $||n\alpha - u|| \leq 1/2q$.

Let $S = \{1/q, 2/q, ..., (q/2)/q\}$ Then

$$\begin{aligned} \sum_{h=M+1}^{M+q} \min(N, ||h\alpha||^{-1}) &= (2/q) \sum_{u \in S} \sum_{h=M+1}^{M=q} \min(N, ||h\alpha||^{-1}) \\ &\leq (2/q) \sum_{u \in S} (3N + \sum_{h:||h\alpha - u|| \leq 1/2q} ||h\alpha||^{-1}) \\ &\leq 3N + (2/q) \sum_{u \in S} \sum_{h:||h\alpha - u|| \leq 1/2q} ((||u|| - 1/2q))^{-1} \\ &\leq 3N + 2 \sum_{u \in S} 1/(||u|| - 1/2q) \\ &\leq 3N + 2 \sum_{m=1}^{q/2} 1/(m/q - 1/2q) \\ &\leq 3N + 2q \log q. \end{aligned} \tag{4.2.1}$$

Now using the same argument as case 1 we have

$$\sum_{h=1}^{2N} \min(N, 1/||h\alpha||) \leq 10(N^2/q + N \log q + q \log q).$$

Therefore, we have

$$|S|^2 \ll N + (N^2/q + N \log q + q \log q).$$

Taking the square root of both sides, we have the desired result. □

4.3 Applications to Uniform Distribution

Definition 4.3.1. *We say that a sequence* $a_1, a_2, ... \in [0,1]$ *is* **uniformly distributed** *if for every* $\alpha \in [0,1]$

$$\lim_{N\to\infty} \frac{1}{N}|\{n : 0 \leq a_n \leq \alpha, n \leq N\}| = \alpha.$$

We will now state and prove Weyl's criterion.

Theorem 4.3.2. ***Weyl's criterion***

The following are equivalent:

(1) The sequence $\{a_n\}_{n=1}^{\infty}$ *is uniformly distributed in the interval* $[0,1]$

(2) For each $k \in \mathbb{Z}$ *and* $k \neq 0$, $\sum_{n=1}^{N} e(-ka_n) = o(N)$

(3) If $F(x)$ *is a bounded and Riemann integrable function on* $[0,1]$, *then* $\lim_{N\to\infty} \sum_{n=1}^{N} F(a_n) = \int_0^1 F(x)dx$

Before proving Weyl's criterion, we give the immediate consequence that the set $\{\alpha n\}$ where α is irrational is uniformly distributed. By Weyl's criterion, condition (2), we can consider the exponential sum

$$\begin{aligned} |\sum_{n=1}^{N-1} e(-k(\alpha n)) &\leq \min(N, (2||k\alpha||)^{-1}) \\ &\leq 1/(2||k\alpha||) \\ &= o(N) \end{aligned}$$

since $||k\alpha|| \neq 0$ since α is irrational.

Proof of Weyl's criterion: (1) $\Rightarrow$ (3) Since we assume F Riemann integrable, we can compute the left hand side of (3) using Riemann sums. Let $\epsilon \geq 0$. Then there exists n such that if we partition $[0,1]$ into n intervals, $0 = x_0 \leq x_1 \leq ... \leq x_n = 1$ and let $M_j = \max_{x_j \leq x \leq x_{j+1}} F(x)$ we have $\sum_{j=0}^{n-1} M_j \Delta x_j - \int_0^1 F(x)dx \leq \epsilon$.

For right hand side of (3) we have

$$\begin{aligned} N^{-1}\sum_{k=1}^{N} F(a_k) &= \sum_{j=0}^{n-1} N^{-1} \sum_{l:a_l\in[x_j,x_j+1)} F(a_l) \\ &\leq \sum_{j=0}^{n-1} 1/N|\{l : a_l \in [x_j, x_{j+1})\}| \cdot M_j \\ &\leq \sum_{j=0}^{n-1} M_j \Delta x_j + \epsilon \\ &\leq \int_0^1 F(x)dx + 2\epsilon \end{aligned}$$

as $N \to \infty$ since $\lim_{N\to\infty} 1/N|\{k : a_k \in [x_j, x_{j+1})\}| = x_{j+1} - x_j = \Delta x_j$ by definition of uniformly distributed.

Using lower Riemann sums in the same way, we get $\lim_{N\to\infty} 1/N \sum_{k=1}^{N} F(a_k) \geq \int_0^1 F(x)dx - 2\epsilon$. Therefore, we can conclude $\lim_{N\to\infty} 1/N \sum_{k=1}^{N} F(a_k) = \int_0^1 F(x)dx$ as desired.

(3) $\Rightarrow$ (2) Set $F(x) = e(-kx)$ with $k \neq 0$. Then by assumption,

$$\begin{aligned} \lim_{N\to\infty} \sum_{n=1}^{N} e(-ka_n) &= \int_0^1 e(-kx)dx \\ &= -k^{-1}e(-kx)|_0^1 \\ &= 0 \end{aligned}$$

(3) $\Rightarrow$ (1) Set $F(x) = \chi_{[0,\alpha)}(x)$ where χ is the characteristic function. Then

$$\begin{aligned} \lim_{N\to\infty} |\{n \leq N : 0 \leq a_n \leq \alpha\}| &= \lim_{N\to\infty} N^{-1} \sum_{n=1}^{N} \chi_{[o,\alpha)}(a_n) \\ &= \int_0^1 \chi_{[o,\alpha)}(x)dx = \alpha \end{aligned}$$

(2) $\Rightarrow$ (3) Here we give an outline of the proof. We can approximate Riemann integrable functions with step functions, step functions with continuous functions and continuous functions by trigonometric polynomials.

Since

$$\lim_{N\to\infty} N^{-1}\sum_{n=1}^{N} e(-ka_n) = 0 = \int_0^1 e(-kx)dx$$

by assumption and

$$\lim_{N\to\infty} N^{-1}\sum_{n=1}^{N} 1 = \int_0^1 1dx,$$

the equality must hold for trigonometric polynomials and therefore, must hold for all functions by approximation. □

Now we will prove that the set defined by $\{P(n)\}$ given by the polynomial $P(x) = \alpha x^2 + \beta x + \gamma$ is uniformly distributed if α is irrational.

Lemma 4.3.3. *Given $\alpha \in \mathbb{R}$ and $N \in \mathbb{N}$, there exists $0 < q \leq N$ such that $||\alpha q|| \leq 1/N$.*

Proof: Consider the set $S = \{0, 1, \{\alpha\}, \{2\alpha\}, ..., \{(N-1)\alpha\}\}$. Partition $[0, 1]$ into intervals $[i/N, (i+1)/N]$. Then we have $N-1$ intervals and $N+1$ elements of S. Hence, by the pigeonhole principle, there must be two elements of S contained in the same interval. Assume those two elements are $\{\alpha x_1\} = |\alpha x_1 - y_1|$ and $\{\alpha x_2\} = |\alpha x_2 - y_2|$. Without loss of generality, assume that $x_1 \geq x_2$. Then $|(x_1\alpha - y_1) - (x_2\alpha - y_2)| = |(x_1 - x_2)\alpha - (y_1 - y_2)| \leq 1/N$. Thus, $||(x_1 - x_2)\alpha|| \leq 1/N$ as desired. □

Lemma 4.3.4. *If $\alpha \notin \mathbb{Q}$ then there exists $q_1, q_2, ...$ such that $\lim_{n\to\infty} q_n = \infty$, $q_N \leq N$ and $|\alpha - a_N/q_N| \leq 1/q_N^2$ for some a_N.*

Proof: Consider $\min_a |\alpha - a/q_N| = 1/q_N \min_a |\alpha q_N - a| = 1/q_N ||\alpha q_N||$. By lemma 4.3.3, for each N, there exists q_N such that $||\alpha q_N|| \leq 1/N \leq 1/q_N$. Since we require α to be irrational, $||\alpha q_N|| \neq 0$ for any q. □

Now the desired result is a simple corollary.

Corollary 4.3.5. *The set $\{\{P(n)\} : n = 1, ..., N\}$ is uniformly distributed in $[0, 1]$.*

Proof: By Weyl's criterion, it suffices to prove

$$\sum_{n=1}^{N} e(kP(n)) = o(N).$$

For each N, choose a_N and q_N such that $|\alpha - a_n/q_n| \leq 1/q_N^2$ as given by lemma 4.3.4. Then, applying Weyl's inequality to $kP(n)$, we have

$$\begin{aligned} |\sum_{n=1}^{N-1} e(kP(n))| &\leq C(N/q_N + \sqrt{N \log q_N} + \sqrt{q_N \log q_N}) \\ &= o(N). \square \end{aligned}$$

Chapter 5

Vinogradov's three-primes theorem

5.1 History, outline and setup

In 1742, Goldbach wrote a letter to Euler conjecturing two statements that would remain open problems for years to come. The first statement was what is today the Goldbach conjecture– namely that any even integer, greater than or equal to six, can be written as the sum of two odd primes. The second statement asserted that every odd integer, greater than or equal to nine, can be written as the sum of three odd primes. It is clear that proving the first statement would guarantee the second. Although the Goldbach conjecture remains open, Goldbach's second conjecture has been proven for sufficiently large N. The first progress was due to Hardy and Littlewood [5] with an application of their circle method in 1923. They proved that if N was sufficiently large, then the conjecture holds assuming the weak generalized Riemann hypothesis. In 1934, using a refinement of the circle method, Vinogradov [15] was able to remove the dependence on the generalized Riemann hypothesis. In 1946 Linnik [8] proved the theorem for large N using Riemann-Hadamard's method of L-series and contour integration. More recently, in 1997, Deshouillers, Effinger, te Riele and Zinoviev [6] proved the complete conjecture that every odd number greater than six can be written as the sum of three prime numbers where they assume the generalized Riemann hypothesis. Their proof is divided into two parts. First, they proved the theorem for $N \geq 10^{20}$. Second, a computational result (independently due to Saouter [13] and Deshouillers- te Riele) shows that the theorem must be true for primes greater than or equal to seven.

The proof of Vinogradov's three primes theorem that is given here follows the lecture notes of Gowers [3] and gives neither a bound on how large N must be nor an asymptotic formula for the number of ways which we can write a large number N as the sum of three prime numbers. We would also

like to cite here Nathanson's text [10] and Vaughan's text [14] which were very useful in writing up this material. The current bound given for the size of N is $10^{43000} \leq N$ and was proved by Chen and Wang [7]. We define

$$r(N) = \sum_{p_1+p_2+p_3=N} 1$$

to be the number of ways to write N as the sum of three primes.

Theorem 5.1.1. *If N is odd and sufficiently large, then*

$$r(N) = \mathcal{G}(N)\frac{N^2}{2(\log N)^3}(1 + O(\frac{\log\log N}{\log N}))$$

where $\mathcal{G}(N)$ is the singular series *for the three-primes theorem defined by*

$$\mathcal{G}(N) = \sum_{q=1}^{\infty} \frac{\mu(q)c_q(N)}{\phi(q)^3}$$

where

$$c_q(N) = \sum_{a=1,(a,q)=1}^{q} e(aN/q)$$

[10].

Our goal will be to show that $r(N) > 0$ for large enough N. Since we are not attempting to show how large N must be, or to give an asymptotic formula for $r(N)$ this will be adequate for our purposes. Let

$$R(N) = \sum_{p_1+p_2+p_3=n} \log p_1 \log p_2 \log p_3.$$

Then showing $R(N) > 0$ is equivalent to showing that $r(N)$ is positive. Letting

$$F(\alpha) = \sum_{p\leq N} (\log p)e(p\alpha)$$

we have

$$\begin{aligned} R(N) &= \sum_{p_1+p_2+p_3=N} \log p_1 \log p_2 \log p_3 \\ &= \sum_{p_1\le N, p_2\le N, p_3\le N} \log p_1 \log p_2 \log p_3 \int_0^1 e((p_1+p_2+p_3)\alpha)e(-N\alpha)d\alpha \\ &= \int_0^1 \sum_{p_1\le N, p_2\le N, p_3\le N} \log p_1 \log p_2 \log p_3 e((p_1+p_2+p_3)\alpha)e(-N\alpha)d\alpha \\ &= \int_0^1 F(\alpha)^3 e(-N\alpha)d\alpha. \end{aligned}$$

Our goal is now is to prove that $\int_0^1 F(\alpha)^3 e(-N\alpha)d\alpha > 0$. To do this, we must be able to estimate $F(\alpha)$, which is difficult. Therefore, we will simplify the problem by considering the set of "almost primes" which behaves similarly to the weighted primes. Since the primes have few divisors, we must make sure that integers in our new set have few divisors.

Definition 5.1.2. *Let $p_1, p_2, ..., p_k$ to be the primes less than or equal to $(\log N)^A$ where A is an absolute constant. Then let $Q = \{x \le N : p_i \nmid x, 1 \le i \le k\}$.*

We will rely on the following functions in our estimates:

$$h(\alpha) = \sum_{x\in Q} e(\alpha x)$$

and

$$h_1(\alpha) = K \sum_{x\in Q} e(\alpha x)$$

where

$$K = \prod_{i=1}^{k} (1 - p_i^{-1})^{-1}$$

where k is as in definition 5.1.2. It is clear that $\int_0^1 h(\alpha)^3 e(-N\alpha)$ counts the number of ways which we can write N as the sum of three elements in Q. Our choice of K has the identity

$$K = \prod_{i=1}^{k} (1 - 1/p_i)^{-1} = e^{\gamma} \log((\log N)^A) + O(1) \qquad (5.1.1)$$

which is Mertens' formula B.0.10. We will prove that the difference

$$|\int_0^1 F(\alpha)^3 e(-N\alpha)d\alpha - \int_0^1 h_1(\alpha)^3 e(-N\alpha)d\alpha|$$

is small in relation to the size of $\int_0^1 h_1(\alpha)^3 e(-N\alpha)$ and hence $\int_0^1 F(\alpha)^3 e(-N\alpha)d\alpha$ must be positive. The key to proving this will be to show that $F(\alpha)$ and $h_1(\alpha)$ are close to each other. Now we are ready to begin our application of the circle method. In showing that $|F(\alpha) - h_1(\alpha)|$ is small, we must consider the case when α is contained in the minor arcs and when α is contained in the major arcs. We will find that if $\alpha \in [0,1]$ is close to a rational with a small denominator, then $F(\alpha)$ is large. If not, then $F(\alpha)$ is small. We now make precise what it means for α to be close to a rational with small denominator.

We will first define the major arcs and then the minor arcs will be everything leftover. Let $B = 16$. Then given $a/q \in [0,1]$, a rational with small denominator means that $1 \leq q \leq (\log N)^B$. We also have the condition that $0 \leq a \leq q$. Now for $\alpha \in [0,1]$ to be close to such a rational means that

$$|\alpha - a/q| \leq (\log N)^B/N.$$

Therefore, we define the **major arcs**, $M(q,a)$ to be an interval of all $\alpha \in [0,1]$ such that α is close to a/q. Precisely, $M(q,a) = \{\alpha : |\alpha - a/q| \leq (\log N)^B/N\}$.

Since we will be using the major and minor arcs to estimate an integral representation of $R(N)$, we must assure that the major arcs $M(q,a)$ are disjoint in order to integrate. We will show this by contradiction. Assume that for different q_1, a_1 and q_2, a_2 we have $|a_1q_2 - a_2q_1| \geq 1$ and $M(q_1,a_1) \cap M(q_2,a_2)$ is nonempty. Take $\alpha_0 \in M(q_1,a_1) \cap M(q_2,a_2)$. Then

$$\frac{1}{(\log N)^{2B}} \leq \frac{1}{q_1q_2} \leq \frac{|a_1q_2 - a_2q_1|}{q_1q_2} = |\frac{a_1q_2}{q_1q_2} - \frac{a_2q_1}{q_1q_2}| = |\frac{a_1}{q_1} - \alpha + \alpha - \frac{a_2}{q_2}|$$

$$\leq |\frac{a_1}{q_1} - \alpha| + |\alpha - \frac{a_2}{q_2}| \leq \frac{2(\log N)^B}{N}.$$

Rearranging the inequality, we have $N \leq 2(\log N)^{3B}$. If we choose N to be large, then this inequality is certainly false. This proves our claim that the major arcs are disjoint. Explicitly we have:

$$M = \bigcup_{q=1}^{\lfloor (\log N)^B \rfloor} \bigcup_{\substack{a=0 \\ (a,q)=1}}^{q} M(q,a) \subseteq [0,1].$$

We can now define the **minor arcs** m to be the complement of M in $[0,1]$.

Using this new notation we rewrite the integral for $R(N)$:

$$R(N) = \int_M F(\alpha)^3 e(-N\alpha) d\alpha + \int_m F(\alpha)^3 e(-N\alpha) d\alpha.$$

In estimating the above integral, we will show that $F(\alpha)$ is small when α is not close to a rational with small denominator and we will estimate $F(\alpha)$ when $\alpha \in M(q,a)$ for some rational number a/q.

5.2 Minor Arcs

The goal of this section is to show that both $F(\alpha)$ and $h_1(\alpha)$ are small when $\alpha \in m$. For each of the following lemmas we use the same hypothesis as when defining the major and minor arcs. We will denote the distance to the closest integer to α by $||\alpha||$.

Now we are ready to begin our estimate. Instead of estimating $F(\alpha)$ directly, we will use the function

$$g(\alpha) = \sum_{x \leq N} \Lambda(x) e(\alpha x),$$

which is easier to estimate, and which we can show is close to $F(\alpha)$. Recall the von Mangoldt function

$$\Lambda(n) = \begin{cases} \log p \text{ if } n = p^k \text{ where p is prime and } k \geq 1 \\ 0 \text{ otherwise} \end{cases}$$

The next lemma proves that $g(\alpha)$ is in fact a good approximation for $F(\alpha)$.

Lemma 5.2.1. *For every $\alpha \in [0,1]$, $|F(\alpha) - g(\alpha)| \leq C\sqrt{N}$ where C is an absolute constant.*

Proof:

$$\begin{aligned}
|g(\alpha) - F(\alpha)| &= |\sum_{x \le N} \Lambda(x) e(\alpha x) - \sum_{p \le N} \log p e(\alpha p)| \\
&= |\sum_{p^k \le N, k \ge 1} \log p e(\alpha p^k) - \sum_{p \le N} \log p e(\alpha p)| \\
&= |\sum_{p^k \le N, k \ge 2} \log p e(\alpha p^k)| \\
&\le (\log \sqrt{N}) \sum_{p \le \sqrt{N}} 1 \\
&\le C\sqrt{N}.
\end{aligned}$$

where the last line is shown using Chebyshev's theorem B.0.8. □

We now turn to the estimation of the minor arcs. Here we will use results due to Vaughan. In the following lemma, we write $g(\alpha)$ as the sum of three terms plus an error term. We will then show that each of the terms is bounded and hence we will be able to bound $g(\alpha)$. Our goal is to show that $g(\alpha)$ is small when α is not close to a rational with small denominator. We will use two methods to show this. First, we will be able to consider the cancelation in the exponential sum as in lemma 4.2.2. Second, we can relate $g(\alpha)$ to $\sum_{d|x} \Lambda(x) = \log x$, discussed above, which should be easier to estimate. We will, in fact, use both ideas to show that $g(\alpha)$ must be small with these conditions.

Lemma 5.2.2. *Let $X = N^{2/5}$. Then $g(\alpha) = \sum_{x \le N} \Lambda(x) e(\alpha x) = S - T - U + O(N^{2/5})$ where*

$$S = \sum_{d \le X} \mu(d) \sum_{z \le N/d} \sum_{x \le N/zd} \Lambda(x) e(\alpha dxz),$$

$$T = \sum_{d \le X} \mu(d) \sum_{z \le N/d} \sum_{x \le X, x \le N/zd} \Lambda(x) e(\alpha dxz),$$

and

$$U = \sum_{X \le u \le N} \sum_{d|u, d \le X} \mu(d) \sum_{X \le x \le N/u} \Lambda(x) e(\alpha xu).$$

Our proof of lemma 5.2.2 incorporates Vaughan's identity which we will use specifically for $g(\alpha)$, but can also be proved in general for any arithmetic function of two variables. We will begin with $g(\alpha)$ and work towards a form that we hope to understand.

Proof of lemma 5.2.2: Using Chebyshev's theorem B.0.8, we can bound the sum

$$\begin{aligned}\sum_{x\leq X}\Lambda(x)e(\alpha x) &= |\sum_{p^k\leq X}\log p e(\alpha p^k)|\\ &\leq \sum_{p^k\leq X}\log p\\ &= \psi(X)\\ &= O(X)\end{aligned}$$

Therefore, by definition,

$$\begin{aligned}g(\alpha) &= \sum_{x\leq N}\Lambda(x)e(\alpha x)\\ &= \sum_{X\leq x\leq N}\Lambda(x)e(\alpha x)+\sum_{x\leq X}\Lambda(x)e(\alpha x)\\ &= \sum_{X\leq x\leq N}\Lambda(x)e(\alpha x)+O(N^{2/5})\end{aligned}$$

Recall that

$$\sum_{d|u}\mu(d)=\begin{cases}1 & \text{if } u=1,\\ 0 & \text{if } u>1.\end{cases}$$

Then

$$\begin{aligned}g(\alpha) &= \sum_{u\leq X}\sum_{d|u}\mu(d)\sum_{X\leq x\leq N/u}\Lambda(x)e(\alpha xu)+O(N^{2/5})\\ &= \sum_{u\leq N}\sum_{d|u,d\leq X}\mu(d)\sum_{X\leq x\leq N/u}\Lambda(x)e(\alpha xu)-U+O(N^{2/5})\\ &= \sum_{d\leq X}\sum_{zd\leq N}\mu(d)\sum_{X\leq x\leq N/dz}\Lambda(x)e(\alpha xzd)-U+O(N^{2/5})\\ &= \sum_{d\leq X}\mu(d)\sum_{z\leq N/d}\sum_{X\leq x\leq N/dz}\Lambda(x)e(\alpha xzd)-U+O(N^{2/5})\\ &= S-T-U+O(N^{2/5}).\end{aligned}$$

Note here that in step one, we use the fact that $d|u$ and $u \leq X$ which implies that $d \leq X$. In step two, we let $u = dz$. □

We now show that S, T and U are small under the prescribed conditions.

Lemma 5.2.3. $|S| \ll (\log N)^2(q + X + N/q)$.

Proof: By definition

$$|S| = |\sum_{d \leq X} \mu(d) \sum_{z \leq N/d} \sum_{x \leq N/zd} \Lambda(x) e(\alpha dxz)|.$$

In the next sequence of inequalities we will split the sum into pieces of the form $\sum_{u \leq n/d} \sum_{x|u} \Lambda(x) e(\alpha u)$ and then apply lemma B.0.6. Now letting $u = xz$, we have

$$\begin{aligned}
|S| &= |\sum_{d \leq X} \mu(d) \sum_{u \leq N/d} \sum_{x|u} \Lambda(x) e(\alpha du)| \\
&= |\sum_{d \leq X} \mu(d) \sum_{u \leq N/d} e(\alpha du) \sum_{x|u} \Lambda(x)| \\
&= |\sum_{d \leq X} \mu(d) \sum_{u \leq N/d} e(\alpha du) \log u| \\
&\leq |\sum_{d \leq X} 1 \sum_{u \leq N/d} e(\alpha du) \log u| \\
&= \sum_{d \leq X} |\sum_{u \leq N/d} e(\alpha du) \log u|.
\end{aligned}$$

Now we notice that

$$\begin{aligned}
|\sum_{u \leq N/d} e(\alpha d) \log u| &= |\sum_{u \leq N/d} \int_1^u e(\alpha du) dt/t| \\
&\leq |\int_1^{N/d} \sum_{t \leq u \leq N/d} e(\alpha du) dt/t| \\
&\leq \int_1^{N/d} |\sum_{t \leq u \leq N/d} e(\alpha du)| dt/t.
\end{aligned}$$

And thus, by lemma 4.2.2

$$\begin{aligned}|\sum_{u\leq N/d} e(\alpha u)\log u| &\leq \int_1^{N/d} \min\{||\alpha d||^{-1}, N/d\}dt/t \\ &\leq \log N \min\{||\alpha d||^{-1}, N/d\}\end{aligned}$$

Now apply lemma A.0.5 and we have

$$\begin{aligned}|S| &\leq \sum_{d\leq X}|\sum_{u\leq N/d} e(\alpha u)\log u| \leq \sum_{d\leq X}\log N \min\{||\alpha d||^{-1}, N/d\} \\ &\ll (\log N)^2(q+X+N/q)\end{aligned}$$

as desired. □

Lemma 5.2.4. $|T| \ll (\log N)^2(q+X^2+N/q)$

Proof:

$$\begin{aligned}|T| &= |\sum_{d\leq X}\mu(d)\sum_{z\leq N/d}\sum_{x\leq X, x\leq N/zd}\Lambda(x)e(\alpha dxz)| \\ &= |\sum_{d\leq X}\mu(d)\sum_{x\leq X, x\leq N/d}\Lambda(x)\sum_{z\leq N/dx}e(\alpha dxz)| \\ &\leq \sum_{d\leq X}\mu(d)\sum_{d\leq X}\Lambda(x)|\sum_{z\leq N/dx}e(\alpha dxz)| \\ &\leq \sum_{d\leq X}\sum_{d\leq X}\Lambda(x)|\sum_{z\leq N/dx}e(\alpha dxz)| \\ &= \sum_{y\leq X^2}\sum_{x\leq X, x|y}\Lambda(x)|\sum_{z\leq N/y}e(\alpha yz)|\end{aligned}$$

when we let $y = dx$. Now we know that $\sum_{x\leq X, x|y}\Lambda(x) \leq \log y \leq \log N$ by lemma B.0.6 and $|\sum_{z\leq N/y} e(\alpha yz)| \leq \min\{||\alpha y||^{-1}, N/y\}$ by lemma 4.2.2. Therefore, as in the previous lemma, we can apply lemma A.0.5 to achieve the desired bound. Note that here we have $X^2 = N^{4/5}$. □

Lemma 5.2.5. $|U| \ll (\log N)^4(N^{1/2}q^{1/2} + N/X^{1/2} + Nq^{-1/2})$

Proof: We know that

$$\begin{aligned}|U| &= |\sum_{X\le u\le N}\sum_{d|u,d\le X}\mu(d)\sum_{X\le x\le N/u}\Lambda(x)e(\alpha xu)|\\ &\le \sum_{X\le u\le N}|\sum_{d|u,d\le X}\mu(d)|\cdot|\sum_{X\le x\le N/u}\Lambda(x)e(\alpha xu)|.\end{aligned}$$

We will split the sum into pieces, each of which we will be able to bound. For each positive integer i, let

$$U_i = \sum_{u=2^{i-1}}^{2^i-1}\left|\sum_{d|u,d\le X}\mu(d)\right|\left|\sum_{X\le x\le N/u}\Lambda(x)e(\alpha xu)\right|.$$

The first important observation is that we will be summing over a finite number of pieces since U_i must be zero when $2^{i-1}\ge N/X$ since then for the last part of the sum we would have $X\le x\le N/(N/X)=X$. Therefore, we will have $\log N$ pieces since we are summing from $2^i\ge X$ and $2^{i-1}\le N/X$.

Now for each U_i we can apply the Cauchy-Schwarz inequality to U_i^2. Thus

$$U_i^2 \le \left(\sum_{u=2^{i-1}}^{2^i-1}\left(\left|\sum_{d|u,d\le X}\mu(d)\right|\right)^2\right)\cdot\left(\sum_{u=2^{i-1}}^{2^i-1}\left|\sum_{X\le x\le N/u}\Lambda(x)e(\alpha xu)\right|^2\right). \tag{5.2.2}$$

We now wish to bound each term of this product.

Since $\mu(d)$ is at most 1 we know that $|\sum_{d|u,d\le X}\mu(d)|\le d(u)$ where $d(u)$ denotes the number of divisors of u. Therefore, we are able to bound the first term of the product (5.2.2) as

$$\begin{aligned}\sum_{u=2^{i-1}}^{2^i-1}\left(\left|\sum_{d|u,d\le X}\mu(d)\right|\right)^2 &\le \sum_{u=2^{i-1}}^{2^i-1}\left(\sum_{d|u,d\le X}d(u)\right)^2\\ &\le \sum_{i=1}^{2^i}d(u)^2\\ &\ll (\log N)^3,\end{aligned}$$

using lemma B.0.7.

Now consider the second term in the product and expand. Thus

$$\sum_{u=2^{i-1}}^{2^i-1} \left| \sum_{X \le x \le n/u} \Lambda(x) e(\alpha x u) \right|^2$$

$$= \sum_{u=2^{i-1}}^{2^i-1} \sum_{X \le x \le N/u} \sum_{X \le y \le N/u} \Lambda(x)\Lambda(y) e(\alpha(x-y)u)$$

$$= \sum_{X \le x \le N/2^{i-1}} \sum_{X \le y \le N/2^{i-1}} \Lambda(x)\Lambda(y) \left| \sum_{2^{i-1} \le u \le 2^i, u \le \min\{N/x, N/y\}} e(\alpha(x-y)u) \right|$$

$$\le \sum_{X \le x \le N/2^{i-1}} \sum_{X \le y \le N/2^{i-1}} \Lambda(x)\Lambda(y) \min\{||\alpha(x-y)||^{-1}\}, 2^{i-1}\}$$

$$\le (\log N)^2 (N/2^i) \sum_{N/2^{i-2} \le z \le N/2^{i-1}} \min\{||\alpha z||^{-1}, 2^{i-1}\}$$

$$\le (\log N)^2 (N/2^i) \sum_{z \le N/2^{i-1}} \min\{||\alpha z||^{-1}, N/z\}$$

$$\ll N(\log N)^3 (N/q + X + q)$$

where we use lemma 4.2.2 for step three and lemma A.0.5 for step five.

Therefore, putting both estimates together and using the facts that $N/2^i \le N/X$ and $2^{i-1} \le N/X$, we have

$$\begin{aligned} U_i^2 &\ll (\log N)^3 \cdot N \cdot (\log N)^3 \cdot N \cdot (\log N)^3 (q + N/2^{i-1} + N/q) \\ &\ll N(\log N)^6 (q + N/X + N/q), \end{aligned}$$

where we use the fact that we have $N/2^i \le N/X$.

Now taking square roots of both sides, we have

$$U_i \ll (\log N)^3 (q^{1/2} N^{1/2} + N X^{-1/2} + N q^{-1/2}).$$

Finally, as mentioned above, we have at most $\log N$ possible i, hence we have

$$U \ll (\log N)^4 (q^{1/2} N^{1/2} + N X^{-1/2} + N q^{-1/2})$$

as desired. □

Lemma 5.2.6. *Let a and q be positive integers with $(a,q)=1$ and let $\alpha \in \mathbb{R}$ such that $|\alpha - a/q| \leq q^{-2}$. Then, given N sufficiently large, $F(\alpha)$ and $g(\alpha)$ are both at most $C(\log N)^4(N^{1/2}q^{1/2} + N^{4/5} + Nq^{-1/2})$.*

Proof: This result combines the results from lemma 5.2.1, lemma 5.2.2 and the lemmas bounding S, T and U. □

We will now use definition 5.1.2 and put a bound on the function $h(\alpha) = \sum_{x \in Q} e(\alpha x)$ defined in the introduction. This bound will be used later to show that distance between $F(\alpha)$ and $h_1(\alpha) = K \cdot h(\alpha)$ is small when we are considering α in a minor arc. This will be our last estimate needed for the minor arcs. We will do the same for α in a major arc, but we must estimate separately.

Lemma 5.2.7. *Let a and q be positive integers with $(a,q)=1$ and let $\alpha \in \mathbb{R}$ such that $|\alpha - a/q| \leq q^{-2}$. Then*

$$|h(\alpha)| \ll (\log N)^2(N^{1/2} + q + Nq^{-1} + N^{1-1/4A}).$$

Proof: We want to write $h(\alpha)$ in some form that we can estimate. Therefore, we notice that

$$h(\alpha) = \sum_{s=0}^{k}(-1)^s \sum_{1 \leq i_1 \leq \ldots \leq i_s \leq k} \sum_{y \leq N/p_{i_1}\ldots p_{i_s}} e(\alpha p_{i_1}\ldots p_{i_s} y). \qquad (5.2.3)$$

Since, if $x \in Q$, then $e(\alpha x)$ is added if and only if $s = 0$. On the other hand, if $x \notin Q$, then we can write $x = p_{j_1}^{\alpha_1} \cdots p_{j_r}^{\alpha_r} w$ where $w \in Q$. In this case, we add $(-1)^{|B|}e(\alpha x)$ for each subset $B \subset \{j_1, \ldots, j_r\}$ using the inclusion-exclusion formula.

We apply lemma 4.2.2 to the inner sum of (5.2.3) and note the bound

$$\sum_{y \leq N/p_{i_1}\ldots p_{i_s}} e(\alpha p_{i_1}\ldots p_{i_s} y) \leq \min\{||\alpha p_{i_1}\ldots p_{i_s}||^{-1}, N/p_{i_1}\ldots p_{i_s}\}.$$

Now we will split the outer sum of 5.2.3 into two parts where we define $t = \log N/2A\log\log N$. Then

$$|g(\alpha)| \leq |\sum_{s=0}^{t}(-1)^s \sum_{1 \leq i_1 \leq \ldots \leq i_s \leq k} \sum_{y \leq N/p_{i_1}\ldots p_{i_s}} e(\alpha p_{i_1}\ldots p_{i_s} y)| +$$

$$|\sum_{s=t+1}^{k}(-1)^s \sum_{1\le i_1\le\ldots\le i_s\le k} \sum_{y\le N/p_{i_1}\ldots p_{i_s}} e(\alpha p_{i_1}\ldots p_{i_s}y)|.$$

We estimate each term separately. For the first sum we observe that for any combination of the primes $p_1,\ldots,p_k$, we have $p_{i_1}\ldots p_{i_t} \le ((\log N)^A)^t = \sqrt{N}$. Thus

$$\begin{aligned}
&|\sum_{s=0}^{t}(-1)^s \sum_{1\le i_1\le\ldots\le i_s\le k} \sum_{y\le N/p_{i_1}\ldots p_{i_s}} e(\alpha p_{i_1}\ldots p_{i_s}y)| \\
&\le |\sum_{s=0}^{t}(-1)^s \sum_{1\le i_1\le\ldots\le i_s\le k} \min\{||\alpha p_{i_1}\ldots p_{i_s}||^{-1}, N/p_{i_1}\ldots p_{i_s}\}| \\
&\le \sum_{x\le\sqrt{N}} \min\{||\alpha x||^{-1}, N/x\} \\
&\ll (\log N)^2(N^{1/2}+q+Nq^{-1})
\end{aligned}$$

using lemma A.0.5 for the last step.

For our second estimate, we have

$$\begin{aligned}
&|\sum_{s=t+1}^{k}(-1)^s \sum_{1\le i_1\le\ldots\le i_s\le k} \sum_{y\le N/p_{i_1}\ldots p_{i_s}} e(\alpha p_{i_1}\ldots p_{i_s}y)| \\
&\le \sum_{s=t+1}^{k} \sum_{1\le i_1\le\ldots\le i_s\le k} \min\{||\alpha p_{i_1}\ldots p_{i_s}||^{-1}, N/p_{i_1}\ldots p_{i_s}\} \\
&\le \sum_{s=t+1}^{k} \sum_{1\le i_1\le\ldots\le i_s\le k} N\prod_{j=1}^{s} p_{i_j}^{-1} \\
&\le N\sum_{s=t+1}^{k}(s!)^{-1}(\sum_{m=1}^{k} p_m^{-1})^s \\
&\le CN\sum_{s=t+1}^{k}(s/e)^{-s}(\log\log(\log N)^A)^s \\
&\le CN\sum_{s=t+1}^{k}(e/s\cdot 2\log\log\log N)^s
\end{aligned}$$

where we use Mertens' theorem B.0.9 and Stirling's formula in step four. The function $(2es^{-1}\log\log\log N)^s$ is decreasing when $s \geq 2e\log\log\log N$ and so to estimate our inequality, we can substitute $s = t = \log N/2A\log\log N \geq 4e\log\log\log N$, so we can conclude,

$$\sum_{s=t+1}^{k}(2es^{-1}\log\log\log N)^s \leq k(2et^{-1}\log\log\log N)^t.$$

We want to prove that this last quantity is less than or equal to $C(N^{-1/4A})$. We know that $k = \pi((\log N)^A) \leq C\frac{(\log N)^A}{\log(\log N)^A} \leq CN^\epsilon$ for any $\epsilon \geq 0$ and N sufficiently large. Then

$$\begin{aligned}
\log(2et^{-1}\log\log\log N)^t &= t\log(2et^{-1}\log\log\log N)\\
&= \frac{\log N}{2A\log\log N}\log(Ae(\log N)^{-1+\epsilon/2})\\
&\leq \frac{\log N}{2A\log\log N}(\log(eA) - (1-\epsilon/2)\log\log N)\\
&= \frac{-2+2\epsilon}{4A}\log N\\
&\leq (-1/4A - \epsilon_0)\log N.
\end{aligned}$$

Therefore, we have

$$(2et^{-1}\log\log\log N)^t \leq CN^{-1/4A-\epsilon_0}.$$

Using this and our upper bound for k, we can conclude that

$$|\sum_{s=t+1}^{k}(-1)^s \sum_{1\leq i_1\leq\ldots\leq i_s\leq k}\ \sum_{y\leq N/p_{i_1}\ldots p_{i_s}} e(\alpha p_{i_1}\ldots p_{i_s}y)| \leq N^{-1/4A},$$

which we combine with our estimate for the first part of the sum to have our desired conclusion. □

5.3 Major Arcs

Now we move on to estimates for the major arcs. Our goal will be to estimate $F(\alpha)$ and $h(\alpha)$ when α is close to a rational with small denominator.

In the next section, along with the results we have proved for the minor arcs, we will show that $F(\alpha) - h_1(\alpha)$ is small regardless of our choice of α. Before continuing with the estimates, we will provide some motivation for the following lemmas.

Let X be an arithmetic progression of the form $\{a, a+1, ..., a+(m-1)q\}$ where a and q are relatively prime and $1 \leq a \leq n-(m-1)q$. Define a function

$$G(x) = \begin{cases} \log x - KQ(x) \text{if } x \text{ is prime} \\ -KQ(x) \text{ otherwise} \end{cases} \tag{5.3.4}$$

where $Q(x)$ is the characteristic function of Q. Then it is clear that

$$\sum_{x \leq N} G(x)e(\alpha x) = F(\alpha) - h_1(\alpha).$$

We will relate G to arithmetic progressions of the above form and therefore estimate the difference between $F(\alpha)$ and $h_1(\alpha)$ by estimating $\sum_{p \in X} \log p$ and $|X \cap Q|$. The following version of the Siegel-Walfisz theorem is taken from Nathanson [10].

Theorem 5.3.1. *(Siegel-Walfisz) If $q \geq 1$ and $(a,q) = 1$, then for any $C \geq 0$,*

$$\sum_{p \leq x, p \equiv a (mod\ q)} \log p = \frac{x}{\phi(q)} + O\left(\frac{x}{(\log x)^C}\right), \tag{5.3.5}$$

for all $x \geq 2$ and where the implied constant depends only on C and ϕ is defined to be the Euler ϕ-function.

The modular condition given to the sum of (5.3.5) provides an estimate for the arithmetic progressions X described above. Since we chose $1 \leq a \leq N-(m-1)q$ then we have

$$\sum_{p \in X} \log p = \frac{mq}{\phi(q)} + O(N/(\log N)^C). \tag{5.3.6}$$

Here we want the estimate for the major arcs and hence we take $q \leq (\log N)^B$ which allows us to conclude that the implied constant must only depend on B and C.

Lemma 5.3.2. *Let* $q \leq (\log N)^B$, *let* $X = \{a, a+q, ..., a+(m-1)q\} \subseteq \{1, ..., N\}$, $m \geq N^{1/2}$ *and let* $(a,q) = 1$. *Then*

$$|X \cap Q| = \frac{mq}{\phi(q)} \prod_{i=1}^{k} (1 - p_i^{-1}) + O(mN^{-1/4A}).$$

Proof: The proof relies on the Brun Sieve for arithmetic progresions. First note that $X \cap Q = \{x \in X : p_i \nmid x \, i = 1, ..., k\}$. Let $x \in X$ be chosen uniformly at random. Define the X_i to be the event $p_i | x$. We let $P(Y)$ be the probability of the event Y. Then

$$P(X_i) = \begin{cases} p_i^{-1} + O(m^{-1}) \text{ if } p_i \nmid q \\ 0 \text{ if } p_i | q \end{cases}$$

Then

$$\begin{aligned} |X \cap Q| &= m(1 - P(\bigcup_{i=1}^{k} X_i)) \\ &= m(1 - P(\bigcup_{i=1}^{r} X_i)), \end{aligned}$$

where $P(X_i) \neq 0$ for $1 \leq i \leq r$. In order to compute $P(\bigcup_{i=1}^{r} X_i)$ we will use the inclusion-exclusion formula. Given a set of events $X_{i_1}, ..., X_{i_s}$ with $1 \leq i_1 \leq ... \leq i_s \leq k$ we can compute the probability that a combination of the events happens as

$$P(X_{i_1} \cap ... \cap X_{i_s}) = \prod_{j=1}^{s} 1/p_{i_j} + O(m^{-1}),$$

if $p_{i_1}, ...p_{i_j} \nmid q$ and 0 otherwise.

Therefore, for any t, we have

$$\begin{aligned} 1 - P(\bigcup_{i=1}^{r} X_i) &= \sum_{s=0}^{t} (-1)^s \sum_{1 \leq i_1 \leq ... \leq i_s \leq k} \prod_{j=1}^{s} 1/p_{i_j} + O(m^{-1}) \sum_{s=1}^{t} r^s \\ &= \sum_{s=0}^{g} (-1)^s \sum_{1 \leq i_1 \leq ... \leq i_s \leq k} \prod_{j=1}^{s} 1/p_{i_j} + O(m^{-1}) \sum_{s=1}^{t} r^s \\ &+ \sum_{s=t+1}^{g} (-1)^s \sum_{1 \leq i_1 \leq ... \leq i_s \leq k} \prod_{j=1}^{s} 1/p_{i_j} + O(m^{-1}) \sum_{s=1}^{t} r^s \end{aligned}$$

For the first sum, we have

$$\begin{aligned}\sum_{s=0}^{g}(-1)^s \sum_{1\le i_1\le\ldots\le i_s\le k}\prod_{j=1}^{s}1/p_{i_j} &= \prod_{i=1}^{g}(1-1/p_i)\\ &= \prod_{i=1}^{k}(1-1/p_i)\prod_{p_i|q}(1-1/p_i)^{-1}\\ &= \prod_{i=1}^{k}(1-1/p_i)\prod_{p_i|q}(1-1/p_i)^{-1}\\ &= \frac{q}{\phi(q)}\prod_{i=1}^{k}(1-1/p_i)\end{aligned}$$

using theorem B.0.12 in the last step.

The second sum, we can estimate as in lemma 5.2.7

$$\sum_{s=t+1}^{g}(-1)^s \sum_{1\le i_1\le\ldots\le i_s\le k}\prod_{j=1}^{s}1/p_{i_j} \le (4et^{-1}\log\log\log)^t$$

where $t \ge 8e\log\log\log N$.

Putting this all together, we have

$$\begin{aligned}1-P(\cup_{i=1}^{g}X_i) &= \prod_{i=1}^{g}(1-1/p_i)+O((\log N)^{At}+(4e\log\log\log N/t)^t)\\ &= \frac{q}{\phi(q)}\prod_{j=1}^{k}(1-1/p_i)+O(N^{-1/4A})\end{aligned}$$

where we take $t = \log N/2A\log\log N$ and estimate as in lemma 5.2.7. Multiplying everything by m we get the desired result. □

Corollary 5.3.3. *Let $q \le (\log N)^A$, let $X = \{a, a+q, ..., a+(m-1)q\}$. Assume that C is any positive constant. Then*

$$K|X\cap Q| - \sum_{p\in X}\log p = O(N(\log N)^{-C}).$$

Proof: Recall that $K = \prod_{i=1}^{k}(1 - p_i^{-1})$. When $(a, q) = 1$, this follows directly from lemma 5.3.2 and from equation 5.3.6. If $(a, q) \neq 1$ then the arithmetic progression X contains at most one prime, which must be a, and the bound holds trivially. □

Lemma 5.3.4. *Let $q \leq (\log N)^A$, let $(b, q) = 1$ and assume $\alpha \in \mathbb{R}$ such that $|\alpha - b/q| \leq (\log N)^A/qN$. Define $G : \{1, 2, ..., N\} \longrightarrow \mathbb{R}$ such that $|G(x)| \leq \log N$ for every x and*

$$|\sum_{x \in X} G(x)| \ll N(\log N)^{-A}$$

for every arithmetic progression X defined as above where $B \geq 4A+2$. Then

$$|\sum_{x \leq N} G(x)e(\alpha x)| = O(N(\log N)^{-A})$$

Proof: The idea of this proof is to split

$$\sum_{x \leq N} G(x)e(\alpha x) = \sum_{i=1}^{k} \sum_{x \in X_i} G(x)e(\alpha x)$$

where the progressions X_i partition $\{1, 2, ..., N\}$ into $2N/m_0$ sets with $m \leq m_0 = N(\log N)^{-B/2}$. We can estimate the magnitude of $\sum_{x \in X} G(x)e(\alpha x)$ using the following fact and lemma 5.3.3. If $x, y \in X$ and $\beta = \alpha - b/q$ then,

$$\begin{aligned} |e(\beta x) - e(\beta y)| &= |e(\beta x)(1 - e(\beta(y - x)))| \\ &= |(1 - e(\beta(y - x))| \\ &\leq 2\pi|y - x||\beta| \\ &\leq 2\pi m(\log N)^A/N. \end{aligned}$$

Let $x_0 \in X$ be fixed. Then for each progression we have

$$\begin{aligned}
|\sum_{x\in X} G(x)e(\alpha x)| &= |\sum_{x\in X} G(x)e((\beta + b/q)x)| \\
&= |\sum_{x\in X} G(x)e(bx/q)e(\beta x)| \\
&\leq |e(ab/q)\sum_{x\in X} G(x)[e(\beta x) + e(\beta x_0) - e(\beta x_0)]| \\
&\leq |e(ab/q)\sum_{x\in X} G(x)(e(\beta x) - e(\beta x_0))| \\
&\qquad + |e(ab/q)e(\beta x_0)\sum_{x\in X} G(x)| \\
&\leq \sum_{x\in X} |G(x)(e(\beta x) - e(\beta x_0))| + |\sum_{x\in X} G(x)| \\
&\ll \sum_{x\in X} \log N \cdot 2\pi m(\log N)^A/N + N(\log N)^{-B} \\
&\ll (\log N)^{A+1} m^2 N^{-1} + N(\log N)^{-B}.
\end{aligned}$$

Since we have split the sum into $2N/m_0$ partitions, we can conclude

$$|\sum_{x\leq N} G(x)e(\alpha x)| \ll 2N/m_0((\log N)^{A+1} m^2 N^{-1} + N(\log N)^{-B})$$

which gives the desired result. □

5.4 Final calculations

We now have everything that we need to show that $F(\alpha)$ and $h_1(\alpha)$ are close. In these final calculations, we will show that $\int_0^1 h_1(\alpha)e(-N\alpha) \geq N^2/\log N$, that the difference $|\int_0^1 F(\alpha)e(-N\alpha) - \int_0^1 h_1(\alpha)e(-N\alpha)| = O(N^2(\log N)^{-3})$ and hence that $\int_0^1 F(\alpha)e(-N\alpha) \geq 0$.

Lemma 5.4.1. *Let $A = 16$. Then for every $\alpha \in \mathbb{R}$, $F(\alpha) - h_1(\alpha) = O(N(\log N)^{-4})$.*

Proof: Fix α. Then there exist b and q with $(b, q) = 1$ such that $q \leq N(\log N)^{-A}$ and $|\alpha - b/q| \leq (\log N)^A/Nq$ (in particular, $|\alpha - b/q| \leq q^{-2}$).

For the case of α in the minor arcs, we consider $q \geq (\log N)^A$. Applying lemma 5.2.6 and using the upper bound and lower bound for q we have

$$\begin{aligned} F(\alpha) &\ll (\log N)^4(N^{1/2}q^{1/2} + N^{4/5} + Nq^{-1/2}) \\ &\ll (\log N)^4(N(\log N)^{-A/2} + N^{4/5} + N(\log N)^{-A/2}) \end{aligned}$$

and for large enough N, we have

$$\ll N(\log N)^{4-A/2} = N(\log N)^{-4}.$$

We now estimate $h_1(\alpha)$ using lemma 5.2.7 and (5.1.1). We have

$$\begin{aligned} |h_1(\alpha)| &\ll K(\log N)^2(N^{1/2} + q + Nq^{-1} + N^{1-1/4A}) \\ &\ll K(\log N)^2 N(\log N)^{-A} \\ &\ll \log(\log N)N(\log N)^{2-A} \\ &= N(\log N)^{-13} \end{aligned}$$

Since we have show that both $F(\alpha)$ and $h_1(\alpha)$ are smaller than the desired bound, it is clear that their difference is bounded as desired. □

For the major arcs, we have $q \leq (\log N)^A$. Let $Q(x)$ be the characteristic function for Q. Recall that X is the progression $X = \{a, a+q, ..., a+(m-1)q\}$. Then we have

$$\begin{aligned} |\sum_{x\in X} G(x)| &= |\sum_{x\in X} KQ(x) - \sum_{p\in X} \log p| \\ &= K|X \cap Q| - \sum_{p\in X} \log p \\ &= O(N(\log N)^{-C}) \end{aligned}$$

by corollary 5.3.3. We can relate G to F and h_1 as follows:

$$\begin{aligned} \sum_{x\leq N} G(x)e(\alpha x) &= \sum_{p\leq N} (\log p - KQ(p))e(\alpha x) - K \sum_{x\leq N, x\neq p} Q(x)e(\alpha x) \\ &= \sum_{p\leq N} (\log p)e(\alpha p) - K \sum_{x\leq N} Q(x)e(\alpha x) \\ &= F(\alpha) - h_1(\alpha). \end{aligned}$$

Applying lemma 5.3.4 we have $F(\alpha) - h_1(\alpha) = O(N(\log N)^{-A})$. □

Now we are ready to prove Vinogradov's theorem. We begin by estimating the number of ways of writing an integer m as the sum of two elements of Q and we obtain an estimate for the number of ways to write m as the sum of three elements of Q as a corollary. Since we have found that Q behaves like the weighted primes, we will use these two results to prove that every large N can be written as the sum of three primes.

Lemma 5.4.2. *Let $m \in \mathbb{Z}$. Then the number of ways of writing $m = x + y$ where $x, y \in Q$ is at least*

$$m \prod_{i=1}^{k} (1 - r_i/p_i) + O(m^{-1} N^{1/2} + m N^{-1/4A})$$

where $r_i = \begin{cases} 1 \text{ if } p_i | m \\ 2 \text{ otherwise} \end{cases}$.

Proof: Choose $x \in \{1, 2, ..., m\}$. For each $1 \leq i \leq k$, let X_i be the event that $p_i | x$ or $p_i | (m - x)$. Note that if $p_i | m$ then we have $p_i | x$ if and only if $p_i | (m - x)$. If $p_i \nmid m$ then $p_i | x$ and $p_i | (m - x)$ are mutually exclusive. Therefore we have

$$P(X_i) = r_i/p_i + O(m^{-1})$$

and

$$P(X_{i_1} \cap \cdots \cap X_{i_s}) = \prod_{j=1}^{s} r_{i_j}/p_{i_j} + O(m^{-1}).$$

The proof then follows directly as in lemma 5.3.2 so we have

$$1 - P(\cup_{i=1}^{k} X_i) = \prod_{i=1}^{k} (1 - r_i/p_i) + O(m^{-1} (\log N)^{At} + (8et^{-1} \log\log\log N)^t)$$

for $t \geq 16e \log\log\log N$. Taking $t = \log N / 2A \log\log N$ the lemma follows. □

As we approach the final result, we first prove a bound on the number of ways which we can write a sufficiently large odd integer, N, as the sum of three elements of Q. Then, using this set, which shares properties with the primes, we will be able to prove Vinogradov's theorem.

Corollary 5.4.3. *Let N be a large odd integer. Then,*

$$\sum_{N=q_1+q_2+q_3, q_i \in Q} 1 \gg (N^2/16)K^{-1} \prod_{i=2}^{k}(1-2/p_i)$$
$$\gg N^2/\log N$$

Proof: If z is odd and $z \leq N/2$, then we can apply lemma 5.4.2 to $N-z$. Then

$$\sum_{N-z=x+y, x,y \in Q} 1 \geq N/4 \prod_{i=2}^{k}(1-2/p_i) + O(N^{1-1/4A}).$$

The number of possible $z \leq N/2$ and $z \in Q$ is given by lemma 5.3.2. Letting $X = \{1, 2, ..., \lfloor N/2 \rfloor\}$, we have

$$|X \cap Q| \geq \lfloor N/2 \rfloor \prod_{i=1}^{k}(1-p_i^{-1})$$
$$\geq (N/4)K^{-1}$$

Therefore,

$$\sum_{N=q_1+q_2+q_3, q_i \in Q} 1 \geq (N/4)K^{-1} \sum_{N-z=x+y, x,y \in Q} 1$$
$$\gg N^2K^{-1} \prod_{i=1}^{k}(1-2/p_i)$$
$$\gg N^2K^{-1} \prod_{i=10}^{k}(1-p_i^{-1}) \prod_{i=9}^{k}(1-p_i^{-1})$$
$$\gg N^2K^{-3}$$
$$\gg N^2(\log\log N)^{-3}$$
$$\gg N^2/\log N,$$

when N is sufficiently large, we use identity (5.1.1) and the fact that since $p_i - 1 \geq p_{i-1}$,

$$1-2/p_i = (1-1/p_i)(\frac{1-2/p_i}{1-1/p_i})$$
$$= (1-1/p_i)(\frac{p_i-2}{p_i-1})$$
$$\geq (1-1/p_i)(1-1/p_{i-1}).\square$$

Finally we are able to prove Vinogradov's three primes theorem by relating $F(\alpha)$ and $h_1(\alpha)$.

Theorem 5.4.4. *If N is a large odd integer, then N can be written as the sum of three primes.*

Proof: As we noted previously, and using lemma 5.4.3 we have

$$\begin{aligned}\int_0^1 h_1(\alpha)^3 e(-N\alpha)K^3 &\geq \int_0^1 h(\alpha)^3 e(-N\alpha)d\alpha \\ &= K^3 \sum_{N=q_1+q_2+q_3, q_i \in Q} 1 \\ &\geq N^2/\log N.\end{aligned}$$

Set $A = 4$. Recall that in lemma 5.4.1 we proved that for every real number α,

$$F(\alpha) - h_1(\alpha) = O(N(\log N)^{-4}).$$

Therefore, we are able to prove the desired result given in the introduction. Namely,

$$\begin{aligned}&\left|\int_0^1 F(\alpha)^3 e(-N\alpha)d\alpha - \int_0^1 h_1(\alpha)^3 e(-N\alpha)d\alpha\right| \\ &\qquad = \left|\int_0^1 (F(\alpha)^3 - h_1(\alpha)^3)e(-N\alpha)d\alpha\right|\end{aligned}$$

$$
\begin{aligned}
&\leq \int_0^1 |F(\alpha)^3 - h_1(\alpha)^3||e(-N\alpha)|d\alpha \\
&\leq \int_0^1 |F(\alpha) - h_1(\alpha)||F(\alpha)^2 + F(\alpha)h_1(\alpha) + h_1(\alpha)^2|d\alpha \\
&= O(N(\log N)^{-A/4}) \int_0^1 |F(\alpha) + h_1(\alpha)|^2 d\alpha \\
&= O(N(\log N)^{-A/4}) \int_0^1 |F(\alpha)|^2 + |h_1(\alpha)|^2 d\alpha \\
&= O(N(\log N)^{-A/4})(\int_0^1 |\sum_{p\leq N} \log p e(\alpha p)|^2 d\alpha + \int_0^1 |Kh(\alpha)|^2 d\alpha) \\
&= O(N(\log N)^{-A/4})(\int_0^1 \sum_{p\leq N} |\log p e(\alpha p)|^2 d\alpha + \int_0^1 |K|^2 |h(\alpha)|^2 d\alpha) \\
&= O(N(\log N)^{-A/4})(\sum_{p\leq N} (\log p)^2 + K^2|Q|) \\
&= O(N^2 \log N (\log N)^{-4})
\end{aligned}
$$

where we use the estimate from B.0.11 for $\sum_{p\leq N}(\log p)^2$. Hence

$$
\begin{aligned}
\int_0^1 F(\alpha)^3 e(-N\alpha)d\alpha &\geq N^2/\log N - O(N^2(\log N)^{-3}) \\
&> 0.
\end{aligned}
$$

Finally, we have proved Vinogradov's three-primes theorem .

Appendix A

A Minor Arcs Lemma

This lemma is in the same spirit as Weyl's inequality, and to prove it, we use elements from the proof of Weyl's inequality. We use this variation in our estimates of the minor arcs in Vinogradov's three-primes thoerem.

Lemma A.0.5. *Let a, q, $N \in \mathbb{N}$. Assume $\alpha \in [0,1]$, $(a,q) = 1$ such that $|\alpha - a/q| \leq q^{-2}$ $q \leq X$ and $X = N^{2/5}$. Then*

$$\sum_{d \leq X} \min\{||\alpha d||^{-1}, N/d\} \ll (\log 2qX)(N/q + X + q).$$

Proof: The first observation that we make is that one can write d uniquely as $d = kq + r$ where $1 \leq r \leq q$ and $0 \leq k \leq X/q$. Therefore we can bound the sum we wish to estimate as

$$\sum_{d \leq X} \min\{||\alpha d||^{-1}, N/d\} \leq \sum_{0 \leq k \leq X/q} \sum_{r=1}^{q} \min\{||\alpha(kq+r)||^{-1}, N/(kq+r)\} \tag{A.0.1}$$

We will estimate the right hand side of equation in two steps. In the first case, we assume $k = 0$ and $r \leq q/2$. Then we would like to bound $\sum_{r=1}^{\lfloor q/2 \rfloor} \min\{||\alpha r||^{-1}, N/r\}$

We have $\alpha = \frac{a}{q} + \frac{u}{q^2}$ where $-1 \leq u \leq 1$ since $|\alpha - \frac{a}{q}| \leq q^{-2}$. Thus $\alpha r = \frac{ar}{q} + \frac{ur}{q^2}$. Note that we can bound the magnitude of the second term $|\frac{ur}{q^2}| \leq \frac{r}{q^2} \leq \frac{q}{2q^2} = \frac{1}{2q}$. Also, since $ar \in \mathbb{Z}$ we can write $ar = jq + s$ for $0 \leq s < q$ and s unique for each r. Therefore, we have the following estimate

for $||\alpha r||$.

$$\begin{aligned} ||\alpha r|| &= ||\frac{ar}{q} + \frac{ur}{q^2}|| \\ &= ||j + \frac{s}{q} + \frac{ur}{q^2}|| \\ &= ||\frac{s}{q} + \frac{ur}{q^2}|| \\ &\geq ||\frac{s}{q}|| - ||\frac{ur}{q^2}|| \\ &\geq ||\frac{s}{q}|| - \frac{1}{2q}. \end{aligned}$$

Therefore, we are able to pair each number $||\alpha r||$ with a unique number $||\frac{s}{q}|| - \frac{1}{2q} = \frac{s}{q} - \frac{1}{2q}$ since $s \leq q/2$ implies that $s/q \leq 1/2$. Therefore, we have

$$\begin{aligned} \sum_{r \leq q/2} \min\{||\alpha r||^{-1}, Nr^{-1}\} &\leq \sum_{r \leq q/2} ||\alpha r||^{-1} \\ &\leq \sum_{s \leq q/2} \frac{1}{\frac{s}{q} - \frac{1}{2q}} \\ &= 2q \sum_{s \leq q/2} \frac{1}{2s - 1} \\ &\leq 2q \sum_{s \leq q/2} \frac{1}{s} \\ &\ll q \log q. \end{aligned}$$

Now we will estimate the sum for $1 \leq k$ or $k = 0$ and $q/2 \leq r \leq q$. Then we have $\frac{1}{kq+r} \leq \frac{2}{(k+1)q}$ since if $1 \leq k$ then $kq + r \geq kq \geq \frac{(k+1)q}{2}$ and otherwise $kq + r = r \geq q/2 = \frac{(k+1)q}{2}$. We would like to bound $\sum_{r=1}^{q} \min\{||\alpha(kq + r)||^{-1}, \frac{N}{(k+1)q}\}$. Although our summation runs from $r = 1$ to $r = q$, the $(kq+r)$ translates the sum and places us in the case of Weyl's inequality (4.2.1). Therefore, we have $\sum_{r=1}^{q} \min\{||\alpha(kq + r)||^{-1}, \frac{N}{(k+1)q}\} \ll \frac{N}{(k+1)q} + q \log q$.

Hence,

$$\begin{aligned}\sum_{d\le X}\min\{||\alpha d||^{-1},N/d\} &\ll q\log q+\sum_{0\le k\le X/q}\sum_{r=1}^{q}\min\{||\alpha(kq+r)||^{-1}\frac{N}{(k+1)q}\}\\ &\ll q\log q+\sum_{0\le k\le X/q}(\frac{N}{(k+1)q}+q\log q)\\ &= q\log q+\frac{N}{q}\sum_{0\le k\le X/q}\frac{1}{k+1}+\sum_{0\le k\le X/q}q\log q\\ &\ll q\log q+\frac{N}{q}\log(X/q+1)+X\log q+q\log q\\ &\ll(\log 2qX)(N/q+X+q)\end{aligned}$$

since $X/q+1\le X+q\le 2\max(q,X)\le 2qX$. □

Appendix B

Theorems from Analytic Number Theory

Here we list a few estimates that we have used in the above proofs. I have found Nathanson's "Elementary Methods in Number Theory"[11] and Apostol's "Introduction to Analytic Number Theory"[1] to be good sources for the following material.

Lemma B.0.6.

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

Proof: Write $x = p_1^{a_1} \cdots p_k^{a_k}$. Then we have the sum of $\log p_1$, a_1 times, $\log p_2$, a_2 times and so on. □

Lemma B.0.7. *Let $d(n)$ denote the number of divisors of $n \in \mathbb{N}$. Then*

$$\sum_{x \leq n} d(x)^2 \leq 2n(\log n)^3$$

Theorem B.0.8. Chebyshev: [11] *Let $\pi(x) = \sum_{p \leq x} 1$, $\vartheta(x) = \sum_{p \leq x} \log p$ and $\psi(x) = \sum_{p^k \leq x} \log p$. Then there exist positive constants A and B such that*

$$Ax \leq \vartheta(x) \leq \psi(x) \leq \pi(x) \log x \leq Bx.$$

The following two theorems are due to Mertens. The proofs can be found in Nathanson [11].

Theorem B.0.9. *There exists a constant b_1 such that*

$$\sum_{p \leq x} 1/p = \log \log x + b_1 + O(\frac{1}{\log x})$$

for $x \geq 2$.

Theorem B.0.10. Mertens' formula: *There exists a constant γ such that for $x \geq 2$,*

$$\prod_{p \leq x} (1 - p^{-1})^{-1} = e^{\gamma} \log x + O(1).$$

Here γ is Euler's constant which is defined by $\gamma = \lim_{n \to \infty} (\sum_{k=1}^{n} 1/k - \log N)$.

Lemma B.0.11.

$$\sum_{p \leq N} (\log p)^2 \ll N \log N$$

Proof: Define $\ell(x)$ to be $\log x$ if x is prime and 0 otherwise. Then

$$\begin{aligned} \sum_{p \leq N} (\log p)^2 &= \sum_{x \leq N} \ell(x) \log x \\ &= \vartheta(N) \log N - \int_1^N \frac{\vartheta(t)}{t} dt \\ &\ll N \log N, \end{aligned}$$

where we use Chebyshev's theorem B.0.8 and partial summation. □

Theorem B.0.12. *The Euler ϕ-function has the identity*

$$\phi(n) = n \prod_{p | n} (1 - p^{-1}).$$

Bibliography

[1] T. M. Apostol. *Introduction to Analytic Number Theory.* Springer, New York, 1976.

[2] F. A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci. U.S.A*, 32:331–332, 1946.

[3] T. Gowers. Vinogradov's three-primes theorem. notes available at http://www.dpmms.cam.ac.uk/ wtg10/3primes.dvi.

[4] T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001. Preprint available at http://www.dpmms.cam.ac.uk/ wtg10/papers.html.

[5] G. H. Hardy and J. E. Littlewood. Some problems of "partitio numerorum"; iii: On the expression of a number as a sum of primes. *Acta Math*, 44:1–70, 1923.

[6] H. te Riele D. Zinoviev J.-M. Deshouillers, G. Effinger. A complete Vinogradov 3-primes theorem under the Riemann hypothesis. *Electronic Research Announcements of the American Mathematical Society*, 3:99–104, 1997.

[7] Chen J.R. and Wang T.Z. On odd Goldbach problem. *Acta Math. Sinica*, 32:702–718, 1989.

[8] J. V. Linnik. A new proof of the Vinogradov-Goldbach theorem. *Mat Sbornik*, 19:3–8, 1946.

[9] H. L. Montgomery. *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis.* American Mathematical Society, Providence, Rhode Island, 1994.

[10] M. B. Nathanson. *Additive Number Theory: The Classical Bases.* Springer, New York, 1996.

[11] M. B. Nathanson. *Elementary Methods in Number Theory.* Springer, New York, 2000.

[12] R. Rankin. Sets of integers containing not more than a given number of terms in arithmetical progression. *Proc. Roy. Soc. Edinburgh, Sect. A*, 65:332–344, 1960/1961.

[13] Y. Saouter. Checking the odd Goldbach conjecture up to 10^{20}. *Math. Comp.*, 67(222):863–866, 1998.

[14] R. C. Vaughan. *The Hardy-Littlewood Method.* Cambridge University Press, Cambridge; New York; Melbourne, second edition, 1997, 1981.

[15] I. M. Vinogradov. Representation of an odd number as a sum of three primes. *Comptes Rendus (Doklady) de l'Académi des Sciences de l'URSS*, XV(3):291–294, 1937.