The Circle Method

Steven J. Miller and
Ramin Takloo-Bighash

July 14, 2004

Contents

Chapter 1

Introduction to the Circle Method

The Circle Method is a beautiful idea for investigating many problems in additive number theory. It originated in investigations by Hardy and Ramanujan ([HR], 1918) on the partition function $P(n)$. We start our study of the Circle Method in §1.1 by reviewing the basic properties of $P(n)$ via generating functions, and then exploring generating functions of a variety of problems. In §1.2 we state the main ideas of the Circle Method, and then in §1.3 we sketch its applications to writing numbers as the sums of primes. We then perform the detailed analysis, handling most of the technicalities, for Germain primes in Chapter 2.

Our goal is to describe the key features of the Circle Method *without* handling all of the technical complications that arise in its use; we refer the reader to the excellent books [EE, Na] for complete details. We highlight the main ideas and needed ingredients for its application, and describe the types of problems it either solves or predicts the answer.

1.1 Origins

In this section we study various problems of additive number theory that motivated the development of the Circle Method. For example, consider the problem of writing n as a sum of s perfect k-powers. If $k = 1$, we have seen a combinatorial solution (see §**??** and Lemma **??**): the number of ways of writing n as a sum of s non-negative integers is $\binom{n+s-1}{s-1}$. Unfortunately, this argument does not generalize to higher k (it is easy to partition a set into s subsets; it is not clear how to partition it into s subsets where the number of elements in each subset is a perfect square). There is another method, an analytical approach, which solves the $k = 1$ case and can be generalized.

For $|x| < 1$, define the **generating function**

$$f(x) \ = \ \sum_{m=0}^{\infty} x^m \ = \ \frac{1}{1-x}. \tag{1.1}$$

Let $r_{1,s}(n)$ denote the number of solutions to $m_1 + \cdots + m_s = n$ where each m_i is a non-negative integer. We claim

$$f(x)^s \;=\; \left(\sum_{m_1=0}^{\infty} x^{m_1}\right)\cdots\left(\sum_{m_s=0}^{\infty} x^{m_s}\right) \;=\; \sum_{n=0}^{\infty} r_{1,s}(n)x^n. \tag{1.2}$$

This follows by expanding the product in (1.1). We have terms such as $x^{m_1}\cdots x^{m_s}$, which is $x^{m_1+\cdots+m_s} = x^n$ for some n. Assuming everything converges, when we expand the product we obtain x^n many times, once for each choice of $m_1, \ldots, m_s$ that adds to n. Thus the coefficient of x^n in the expansion is $r_{1,s}(m)$. On the other hand, we have

$$f(x)^s \;=\; \left(\frac{1}{1-x}\right)^s \;=\; \frac{1}{(s-1)!}\frac{d^{s-1}}{dx^{s-1}}\frac{1}{1-x}. \tag{1.3}$$

Substituting the geometric series expansion for $\frac{1}{1-x}$ gives

$$f(x)^s \;=\; \frac{1}{(s-1)!}\frac{d^{s-1}}{dx^{s-1}}\sum_{n=0}^{\infty} x^n \;=\; \sum_{n=0}^{\infty}\binom{n+s-1}{s-1}x^n, \tag{1.4}$$

which yields $r_{1,s}(n) = \binom{n+s-1}{s-1}$. It is this second method of proof that we generalize. Below we describe a variety of problems and show how to find their generating functions. In most cases, exact formulas such as (1.3) are unavailable; we develop sufficient machinery to analyze the generating functions in a more general setting.

Exercise 1.1.1. *Justify the arguments above. Show all series converge, and prove* (1.3) *and* (1.4)*.*

1.1.1 Partitions

We describe several problems where we can identify the generating functions. For $n \in \mathbb{N}$, $P(n)$ is the **partition function**, the number of ways of writing n as a sum of positive integers where we do not distinguish re-orderings. For example, if $n = 4$ then

$$\begin{aligned} 4 &= 4 \\ &= 3+1 \\ &= 2+2 \\ &= 2+1+1 \\ &= 1+1+1+1, \end{aligned} \tag{1.5}$$

and $P(4) = 5$. Note we do not count both $3+1$ and $1+3$. If we add the requirement that no two parts can be equal, there are only two ways to partition 4: 4 and $3+1$.

Proposition 1.1.2 (Euler). *We have as an identity of formal power series*

$$F(x) \;=\; \frac{1}{(1-x)(1-x^2)(1-x^3)\cdots} \;=\; 1+\sum_{n=1}^{\infty} P(n)x^n. \tag{1.6}$$

An **identity of formal power series** means that, without worrying about convergence, the two sides have the same coefficients of x^n for all n. For this example, if we use the geometric series expansion on each $\frac{1}{(1-x^k)}$ and then collect terms with the same power of x, we would have the series on the right; however, we do not know that the series on the right is finite for any x.

Exercise 1.1.3. *Prove the above proposition. Do the product or series converge for any $x > 0$?* Hint: *the combinatorial bounds from §**??** might be a useful starting point.*

$F(x)$ is called the generating function of the partition function. If $f(n)$ is an arithmetic function (see Chapter **??**), we can associate a generating function to f through a power series:

$$F_f(x) \;=\; 1+\sum_{n=1}^{\infty} f(n)x^n. \tag{1.7}$$

Exercise 1.1.4. *1. Fix $m \in \mathbb{N}$. For each n, let $p_m(n)$ be the number of partitions of n into numbers less than or equal to the given number m. Show that*

$$\frac{1}{(1-x)(1-x^2)\cdots(1-x^m)} \;=\; 1+\sum_{n=1}^{\infty} p_m(n)x^n. \tag{1.8}$$

Does the series converge for any $x > 0$?

2. Show that

$$(1+x)(1+x^2)(1+x^3)\cdots \;=\; 1+\sum_{n=1}^{\infty} q(n)x^n. \tag{1.9}$$

where $q(n)$ is the number of partitions of n into non-equal parts. Does this series converge for any $x > 0$?

3. Give similar interpretations for

$$\frac{1}{(1-x)(1-x^3)(1-x^5)\cdots} \tag{1.10}$$

and

$$(1+x^2)(1+x^4)(1+x^6)\cdots. \tag{1.11}$$

Do these products converge for any $x > 0$?

One can use generating functions to obtain interesting properties of the partition functions:

Proposition 1.1.5. *Let $n \in \mathbb{N}$. The number of partitions of n into unequal parts is equal to the number of partitions of n into odd numbers.*

Exercise 1.1.6. *Prove Proposition 1.1.5.* Hint:

$$(1+x)(1+x^2)(1+x^3)\cdots = \frac{1-x^2}{1-x}\frac{1-x^4}{1-x^2}\frac{1-x^6}{1-x^3}\cdots \tag{1.12}$$

$$= \frac{1}{(1-x)(1-x^3)(1-x^5)\cdots}. \tag{1.13}$$

For more examples of this nature, see Chapter XIX of [HW].

So far, we have studied power series expansions where the coefficients are related to the function we want to study. We now consider more quantitative questions. Is there a simple formula for $P(n)$? How rapidly does $P(n)$ grow as $n \to \infty$? Using the Circle Method, Hardy and Ramanujan showed that

$$P(n) \sim \frac{e^{\pi\sqrt{2n/3}}}{4n\sqrt{3}}. \tag{1.14}$$

We prove similar results for other additive problems.

1.1.2 Waring's Problem

It is useful to think of the partition problem in §1.1.1 as the study of the number of ways that a given number n can be written as a sum

$$\sum_i n_i^k \tag{1.15}$$

for $k = 1$, with the number of terms ranging from n (when each $n_i = 1$) to 1 (when $n_1 = n$). See also §**??**. We can now formulate the following question:

Question 1.1.7. *Let $k \in \mathbb{N}$. Let $P_k(n)$ be the number of ways that n can be written as the sum of perfect k^{th} powers. Can one calculate $P_k(n)$?*

It is clear that for all n, $P_n(n)$ is non-zero as n can be written as the sum of n ones. There is a striking difference between this case and the problem of $P(n)$ in §1.1.1. The difference is that if n is a natural number and $m < n$, then one can easily write a partition of n into m numbers. For higher powers, however, this is false; in fact not true even for $k = 2$. For example, 3 cannot be written as the sum of two squares. Hence we ask the following questions:

Question 1.1.8 (Waring's Problem). *Let $k \in \mathbb{N}$. What is the smallest number s such that every natural number can be written as the sum of at most s perfect k^{th} powers? Does such an s exist? If s exists, how does s depend on k?*

These questions can easily be translated to questions involving appropriate generating functions, as we now explain. For Question 1.1.7, we easily see that

$$1 + \sum_{n=1}^{\infty} P_k(n)x^n \;=\; \frac{1}{(1-x^{1^k})(1-x^{2^k})(1-x^{3^k})\cdots}; \tag{1.16}$$

however, this expansion is only useful if we can use it to calculate the $P_k(n)$s. For Question 1.1.8, consider the auxiliary function

$$Q_k(x) \;=\; \sum_{n=0}^{\infty} x^{n^k}. \tag{1.17}$$

As an identity of formal power series, we have

$$Q_k(x)^s \;=\; 1 + \sum_{n=1}^{\infty} a(n;k,s)x^n, \tag{1.18}$$

where $a(n;k,s)$ is the number of ways to write n as a sum of exactly s perfect k^{th} powers.

Exercise 1.1.9. *Prove* (1.18).

Remark 1.1.10 (Important). *So far, all we have done is to use generating functions to find an equivalent formulation for the original problem. We must find a good way to determine $a(n;k,s)$.*

If we could show that given a k there exists an s such that for all n, $a(n;k,s) \neq 0$, then we would have proved every number is the sum of s perfect k^{th} powers. The smallest such s, if it exists, is usually denoted by $g(k)$. In 1770 Waring stated without proof that every natural number is the sum of at most nine positive perfect cubes, also the sum of at most 19 perfect fourth powers, and so on. It was already known that every number is a sum of at most four squares. It is usually assumed that Waring believed that for all k, $g(k)$ exists. Hilbert [Hil] proved Waring's conjecture in 1909, though his method yielded poor bounds for the true value of $g(k)$.

Exercise 1.1.11. *Show that no number of the form $4k+3$ can be the sum of two squares. Show that no number of the form $4^a(8k+7)$ is the sum of three squares. This exercise shows that we cannot write all sufficiently large numbers as the sum of three squares.*

Exercise 1.1.12. *Let $n_k = 2^k\left[\left(\frac{3}{2}\right)^k\right] - 1$. How many perfect k^{th} powers are needed to represent n_k as a sum of k^{th} powers? Conclude that $g(k) \geq 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2$. This gives $g(2) \geq 4$, $g(3) \geq 9$, $g(4) \geq 19$, ...*

Exercise 1.1.13. *Using density arguments, we can often prove certain problems have no solutions. Show there are not enough perfect squares to write any (large) number as the sum of two squares. Use this method to determine a lower bound for how many perfect k^{th} powers are needed for each k.*

Let us concentrate on $g(2) = 4$. As we now know that infinitely many numbers cannot be the sum of three squares, we need to show that every natural number can be written as a sum of four squares. There are many proofs of this important fact, the first of which is due to Lagrange (though it is believed that Diophantus was familiar with the theorem). One proof uses geometric considerations based on Minkowski's theorem (see [Ste]). We refer the reader to Chapter XX of [HW] for three interesting proofs of the theorem, as well as [Na]. We are particularly interested in the proof in §20.11 and §20.12 of [HW] which uses generating functions. We set

$$\theta(x) \;=\; \sum_{m=-\infty}^{\infty} x^{m^2}. \tag{1.19}$$

If $r(n)$ is defined by

$$\theta(x)^4 \;=\; 1 + \sum_{n=1}^{\infty} r(n) x^n, \tag{1.20}$$

then $r(n)$ is equal to the number of representations of n as the sum of four squares:

$$r(n) \;=\; \#\{(m_1, m_2, m_3, m_4) : m_i \in \mathbb{Z},\; n = m_1^2 + m_2^2 + m_3^2 + m_4^2\}. \tag{1.21}$$

Here the m_is are integers, and different permutations of the m_is are counted as distinct. One can show that

$$\theta(x)^4 \;=\; 1 + 8 \sum_{n=1}^{\infty} c_n x^n \tag{1.22}$$

where

$$c_n \;=\; \sum_{m|n, 4\nmid m} m. \tag{1.23}$$

Thus, $r(n) = 8c(n)$. As $c_n > 0$, this implies that every integer is the sum of four squares.

Exercise 1.1.14. ***Give exercises sketching proof of claim.***

The above is a common feature of such proofs: we show the *existence* of at least one solution by showing there are many. We proved n is the sum of four squares by actually finding out how many different ways n is the sum of four squares. In our investigations of other problems, we will argue similarly.

1.1.3 Goldbach's conjecture

Previously we considered the question of determining the smallest number of perfect k^{th} powers needed to represent all natural numbers as a sum of k^{th} powers. One can consider the analogous question for other sets of numbers. Namely, given a set A, is there a number s_A such that every natural number can be written as a sum of at most s_A elements of A? A set of natural arithmetic interest is the set P of all prime numbers. Goldbach, in a letter to Euler (June 7, 1742), conjectured that every integer is the sum of three primes. Euler reformulated this conjecture to every even integer is the sum of two primes.

Exercise 1.1.15. *Prove that if every integer is the sum of at most three primes, then every even number must be the sum of at most two primes. Conversely, show if every even integer is the sum of at most two primes, every integer is the sum of at most three primes.*

To date, Goldbach's conjecture has been verified for all even numbers up to $2 \cdot 10^{16}$ (see [Ol]). There are deep unconditional results in the direction of Goldbach's conjecture:

1. Shnirel'man proved $s_P < \infty$. The proof is based on an ingenious density argument (see [Na], Chapter 7).

2. Estermann [Est1] proved that almost every even number is the sum of two primes.

3. Vinogradov showed every large enough odd number is the sum of three primes. We discuss the proof of Vinogradov's theorem later. Vinogradov proved his theorem in [Vin1, Vin2], where he reformulated the Circle Method from the language of complex analysis to that of Fourier series. [ChWa] has shown that sufficiently large may be taken to be $e^{e^{11.503}}$.

4. Chen proved every even number is the sum of a prime and a number that is at most the product of two primes. Chen's theorem is based on a sieve argument (see [Na], Chapters 9 and 10).

In the next section we describe the key ideas of the Circle Method. This will allow us to approximate quantities such as $a(n; k, s)$ (see (1.18)). We return to generating function approaches to Goldbach's conjecture in §1.3.

1.2 The Circle Method

We explain the key features of the Circle Method. We reinterpret some of the problems discussed in §1.1 in this new language.

1.2.1 Problems

The Circle Method was devised to deal with additive problems of the following nature:

Problem 1.2.1. *Given some subset $A \subset \mathbb{N}$ and a positive integer s, what natural numbers can be written as a sum of s elements of A, and in how many ways? Explicitly, what is*

$$\{a_1 + \cdots + a_s : a_i \in A\} \cap \mathbb{N}. \tag{1.24}$$

More generally, one has

Problem 1.2.2. *Fix a collection of subsets $A_1, \ldots, A_s \subset \mathbb{N}$ and study*

$$\{a_1 + \cdots + a_s : a_i \in A_i\} \cap \mathbb{N}. \tag{1.25}$$

We give several problems where the Circle Method is useful. We confine ourselves to two common choices for A. The first choice is P, the set of primes: $P = \{2, 3, 5, 7, 11, \ldots\}$. We denote elements of P by p. The second choice is K, the set of k^{th} powers of non-negative integers; $K = \{0, 1, 2^k, 3^k, 4^k, \ldots\}$. We denote elements of K by n^k.

1. Consider $A = P$ and $s = 2$. Thus we are investigating

$$\{p_1 + p_2 : p_i \text{ prime}\} \cap \mathbb{N}. \tag{1.26}$$

   This is Goldbach's conjecture for even numbers.

2. Again let $A = P$ but now let $s = 3$. Thus we are investigating

$$\{p_1 + p_2 + p_3 : p_i \text{ prime}\} \cap \mathbb{N}. \tag{1.27}$$

   Vinogradov's theorem asserts that every large enough odd number is included in the intersection.

3. Let $A = K$ and fix a positive integer s. We are studying

$$\{n_1^k + \cdots + n_s^k : n_i \in \mathbb{N}\} \cap \mathbb{N}. \tag{1.28}$$

   This is Waring's problem.

4. Let $-P = \{-2, -3, -5, \ldots\}$. If we consider $P - P$, we have

$$\{p_1 - p_2\} \cap \mathbb{N}. \tag{1.29}$$

   This tells us which numbers are the differences between primes. A related question is to study how many pairs (p_1, p_2) satisfy $p_1 - p_2 = n$. If we take $n = 2$, p_1 and p_2 are called **twin primes**.

In the following paragraphs we sketch the main ideas of the Circle Method, first without worrying about convergence issues, then highlighting where the technicalities lie. In Chapter 2 we work through all but one of these technicalities for a specific problem; the remaining technicality for this problem has resisted analysis to this day. We have chosen to describe an open problem rather than a problem where all the difficulties can be handled for several reasons. The first is that to handle these technicalities for one of the standard problems would take us too far afield, and there are several excellent expositions for those desiring complete detail (see [Da2, EE, Na]). Further, there are numerous open problems where the Circle Method provides powerful heuristics that agree with experimental investigations; after working through the problem in Chapter 2 the reader will have no trouble deriving such estimates for additional problems.

1.2.2 Setup

Let us consider Problem 1.2.1. As before, we consider a generating function

$$F_A(x) \ = \ \sum_{a\in A} x^a. \tag{1.30}$$

Next, we write

$$F_A(x)^s \ = \ \sum_{n=1}^{\infty} r(n;s,A)x^n. \tag{1.31}$$

Exercise 1.2.3. *Prove $r(n;s,A)$ is the number of ways of writing n as a sum of s elements of A.*

An equivalent formulation of Problem 1.2.1 is the following:

Problem 1.2.4. *Determine $r(n;s,A)$.*

In order to extract individual coefficients from a power series we have the following standard fact from complex analysis:

Proposition 1.2.5. *1. Let γ be the unit circle oriented counter-clockwise. Then*

$$\frac{1}{2\pi i}\int_\gamma z^n\,dz \ = \ \begin{cases} 1 & \text{if } n=-1; \\ 0 & \text{otherwise.} \end{cases} \tag{1.32}$$

2. Let $P(z)=\sum_{k=0}^{\infty} a_k z^k$ be a power series with radius of convergence larger than one. Then

$$\frac{1}{2\pi i}\int_\gamma P(z)z^{-n-1}\,dz \ = \ a_n. \tag{1.33}$$

See §**??** for a sketch of the proof, or any book on complex analysis (for example, [Al, La5]). Consequently, ignoring convergence problems yields

$$r(n;s,A) \;=\; \frac{1}{2\pi i}\int_\gamma F_A(z)^s z^{-n-1}\,dz. \tag{1.34}$$

Definition 1.2.6 ($e(x)$**).** *We set*

$$e(x) \;=\; e^{2\pi i x}. \tag{1.35}$$

Exercise 1.2.7. *Let* $m,n\in\mathbb{Z}$*. Prove*

$$\int_0^1 e(nx)e(-mx)dx \;=\; \begin{cases} 1 & \text{if } n=m; \\ 0 & \text{otherwise.}\end{cases} \tag{1.36}$$

An alternative, but equivalent, formulation is to consider a different generating function for A:

$$f_A(x) \;=\; \sum_{a\in A} e(ax). \tag{1.37}$$

Again, ignoring convergence problems,

$$\int_0^1 f_A(x)^s e(-nx)dx \;=\; r(n;s,A). \tag{1.38}$$

If we can evaluate the above integral, not only will we know which n can be written as the sum of s elements of A, but we will know in how many ways.

Exercise 1.2.8. *Using exercise 1.2.7, prove* (1.38)*.*

1.2.3 Convergence Issues

The additive problem considered in Problem 1.2.1 is interesting only if A is infinite; otherwise, we can just enumerate $a_1+\cdots+a_s$ in a finite number of steps. If A is infinite, the defining sum for the generating function $f_A(x)$ need not converge, or may not have a large enough radius of convergence. For each N, define

$$A_N \;=\; \{a\in A : a\le N\} \;=\; A\cap\{0,1,\dots,N\}. \tag{1.39}$$

Note the A_Ns are an increasing sequence of subsets

$$A_N \;\subset\; A_{N+1}, \tag{1.40}$$

and

$$\lim_{N\to\infty} A_N \;=\; A. \tag{1.41}$$

For each N, we consider the truncated generating function attached to A_N:

$$f_N(x) \;=\; \sum_{a\in A_N} e(ax). \tag{1.42}$$

As $f_N(x)$ is a finite sum, all the convergence issues vanish. A similar argument as before yields

$$f_N(x)^s \quad = \quad \sum_{n\le sN} r_N(n;s,A)e(nx), \tag{1.43}$$

except now we have $r_N(n;s,A)$, which is the number of ways of writing n as the sum of s elements of A with each element at most N. If $n \le N$, then $r_N(n;s,A) = r(n;s,A)$, the number of ways of writing n as the sum of s elements of A; note $f_N(x)^s$ is the generating function for the sum of s elements (at most N) of A

For example, if $A = P$ (the set of primes), $N = 10$ and $s = 2$, then $A_{10} = P_{10} = \{2,3,5,7\}$. An easy calculation gives $r_{10}(8;2,P) = r(8;2,P) = 2$. However, $r_{10}(14;2,P) = 1$ (from $7+7$) but $r(14;2,P) = 3$ (from $7+7$, $3+11$, and $11+3$).

We have shown the following, which is the key re-formulation of these additive problems:

Lemma 1.2.9. *If $n \le N$ then*

$$r(n;s,A) \;=\; r_N(n;s,A) \;=\; \int_0^1 f_N(x)^s e(-nx)dx. \tag{1.44}$$

However, having an integral expression for $r_N(n;s,A)$ is not enough; we must be able to *evaluate* the integral (either exactly, or at least bound it away from zero). Note $f_N(x)$ has $|A_N|$ terms, each term of absolute value 1. In many problems, for most $x \in [0,1]$ the size of $f_N(x)$ is about $\sqrt{|A_N|}$, while for special $x \in [0,1]$ one has $f_N(x)$ is of size $|A_N|$. The main contribution to the integral is expected to come from x where $f_N(x)$ is large, and often this integration can be performed. If we can show that the contribution of the remaining x is smaller, we will have bounded $r_N(n;s,A)$ away from zero.

1.2.4 Major and Minor arcs

The difficultly is evaluating the integral in Lemma 1.2.9. Many successful applications of the Circle Method proceed in the following manner:

1. Given a set A, we construct a generating function $f_N(x)$ for A_N. As $f_N(x)$ is a sum of complex exponentials of size 1, we expect there will often be significant cancellation. See the comments after Theorem **??** for other examples of similar cancellation in number theory.

2. Split $[0,1]$ into two disjoint pieces, called the **Major arcs** $\mathcal{M}$ and the **Minor arcs** m. Then
$$r(m;s,A) \;=\; r_N(m;s,A) \;=\; \int_{\mathcal{M}} f_N^s(x)e(-mx)dx + \int_{\mathrm{m}} f_N^s(x)e(-mx)dx. \tag{1.45}$$
The construction of $\mathcal{M}$ and m depend on N and the problem being studied.

3. On the Major arcs $\mathcal{M}$ we find a function which, up to lower order terms, agrees with $f_N^s(x)$ and is easily integrated. We then perform the integration, and are left with a contribution over the Major arcs which is bounded away from zero and is large.

4. One shows that as $N \to \infty$, the Minor arcs' contribution is of lower order than the Major arcs' contribution. This implies that for n large, $r_N(n;s,A) > 0$, which proves that large n can be represented as a sum of s elements of A.

The last is the most difficult step. It is often highly non-trivial to obtain the required cancellation over the Minor arcs. For the problems mentioned, we are able to obtain the needed cancellation for $A = P$ and $s = 3$ (every large odd number is the sum of three primes), but not $A = P$ and $s = 2$; we give some heuristics in §1.3.7 as to why $s = 2$ is so much harder than $s = 3$. For $A = K$ (the set of k^{th} powers of integers), we can obtain the desired cancellation for $s = s(k)$ sufficiently large. Hardy and Littlewood proved we may take $s(k) = 2^k + 1$. Wooley and others have improved this result; however, in general we expect the result to hold for smaller s than the best results to date.

1.2.5 Historical Remark

We briefly comment on the nomenclature: we have been talking about the Circle Method and arcs, yet there are no circles anywhere in sight! Let us consider an example. Recall from Proposition 1.1.2 that the generating function for the partition problem is
$$F(x) \;=\; \frac{1}{(1-x)(1-x^2)(1-x^3)\cdots} \;=\; 1 + \sum_{n=1}^{\infty} P(n)x^n. \tag{1.46}$$
By (1.34), and ignoring convergence issues, we need to consider
$$P(n) \;=\; \frac{1}{2\pi i}\int_{\gamma} F(z)z^{-n-1}\,dz. \tag{1.47}$$
The integrand is not defined at any point of the form $e(\frac{a}{q})$. The idea is to consider a small arc around each point $e(\frac{a}{q})$. This is where $|F(z)|$ is large. At least intuitively one expects that the integral of $F(z)$ along these arcs should be the major part of the integral. Thus, we break the unit circle into two disjoint pieces, the Major arcs (where we expect the generating function to be large), and the Minor arcs (where we expect the function to be small). While many problems proceed through generating functions that are sums of exponentials, as well as integrating over $[0,1]$ instead of a cirlce, we keep the original terminology.

1.2.6 Needed Number Theory Results

In our applications of the Circle Method, we need several results concerning prime numbers. These will be used to analyze the size of the generating function on the Major arcs. As we have seen in §**??** and §**??**, it is often easier to weight primes by $\log p$ in sums, and then remove these weights through partial summation. We use the following statements freely (see, for example, [Da2] for proofs). We constantly use partial summation; the reader is advised to review the material in §**??**.

Theorem 1.2.10 (Prime Number Theorem). *Let $\pi(x)$ denote the number of primes at most x. Then there is a constant $c < 1$ such that*

$$\sum_{p \le x} \log p \;=\; x \;+\; O\left(x \exp(-c\sqrt{\log x})\right). \tag{1.48}$$

Equivalently, by partial summation we have

$$\pi(x) \;=\; \sum_{p \le x} 1 \;=\; Li(x) \;+\; O\left(x \exp(-\frac{c}{2}\sqrt{\log x})\right), \tag{1.49}$$

*where $Li(x)$ is the **logarithmic integral**, which for any fixed positive integer k has the Taylor expansion*

$$Li(x) \;=\; \int_2^x \frac{dt}{\log t} \;=\; \frac{x}{\log x} + \frac{1!x}{\log^2 x} + \cdots + \frac{(k-1)!x}{\log^k x} + O\left(\frac{x}{\log^{k+1} x}\right). \tag{1.50}$$

The above is the original version of the Prime Number Theorem. The error term has been strengthened by Korobov and Vinogradov to $O\left(x \exp(-c_\theta \sqrt[\theta]{\log x})\right)$ for any $\theta < \frac{3}{5}$. All we will need is

$$\pi(x) \;=\; \frac{x}{\log x} + o\left(\frac{x}{\log x}\right), \quad \sum_{p \le x} \log p \;=\; x + o(x). \tag{1.51}$$

Exercise 1.2.11. *Using partial summation, deduce a good estimate for $\pi(x)$ from* (1.48).

Exercise 1.2.12. *Prove* (1.50).

Theorem 1.2.13 (Siegel-Walfisz). *Let $C, B > 0$ and let a and q be relatively prime. Then*

$$\sum_{\substack{p \le x \\ p \equiv a(q)}} \log p \;=\; \frac{x}{\phi(q)} + O\left(\frac{x}{\log^C x}\right) \tag{1.52}$$

for $q \le \log^B x$, and the constant above does not depend on x, q or a (i.e., it only depends on C and B).

One may interpret the Siegel-Walfisz Theorem as saying each residue class has, to first order, the same number of primes. Explicitly, for a fixed q there are $\phi(q)$ numbers a relatively prime to q. Up to lower order terms each residue class has $\frac{\pi(x)}{\phi(q)}$ primes (see §**??**, §**??**). *Note the main term is larger than the error term if we choose C sufficiently large*. If we were to take q as large as x^δ for some $\delta > 0$, then the error term would exceed the main term; we want to apply this theorem when q is much smaller than x. The choice of the Major arcs is crucially influenced by the error term in the Siegel-Walfsiz Theorem.

1.3 Goldbach's conjecture revisited

While we discuss the complications from estimating the integral over the Minor arcs below, we do not give details on actually bounding these integrals; the interested reader should consult [Da2, EE, Est2, Na]. It is our intention to only *introduce* the reader to the broad brush strokes of this elegant theory.

Unfortunately, such an approach means that at the end of the day, we have not solved the original problem. We have chosen this approach for several reasons. While the technical details can be formidable, for many problems these details are beautifully presented in the above (and many other) sources. Further, there are many applications of the Circle Method where the needed estimates on the Minor arcs are not known, even assuming powerful conjectures such as the Generalized Riemann Hypothesis. In these cases, while it is often reasonable to assume that the contribution from the Major arcs is the main term, one cannot prove such statements. Thus, the techniques we develop are sufficient to allow the reader to *predict* the answer for a variety of open problems; these answers can often be tested numerically.

For these reasons, we describe the ideas of the Circle Method for Goldbach's problem: What are the Major and Minor arcs? Why do we obtain the necessary cancellation when $s = 3$ but not when $s = 2$? These examples are well known in the literature, and we content ourselves with a very brief introduction. In Chapter 2 we give a very thorough treatment of another Circle Method problem, Germain primes, which has applications to cryptography. The techniques for this problem suffice to estimate the Major arc contributions in many other problems (for example, how many twin primes are there less than x).

We do not always explicitly compute the error terms below, often confining ourselves to writing the main term and remarking the correction terms are smaller. As an exercise, the reader is encouraged to keep track of these errors.

1.3.1 Setup

The Circle Method begins with a choice of a generating function specific to the problem. For analytical reasons (see remark **??** and §1.2.6), it is often convenient to analyze the weighted generating function

$$F_N(x) \ = \ \sum_{p \le N} \log p \cdot e(px) \tag{1.53}$$

instead of $f_N(x)$, and pass to the unweighted function by partial summation. One could worked only with $f_N(x)$ (see [Est2], Chapter 3); however, we prefer to use $F_N(x)$ as the weights are easily removed and simplify several formulas. Working analogously as before, to write m as a sum of s primes leads us to

$$R_{N,s}(m) \;=\; \int_0^1 F_N^s(x)e(-mx)dx, \tag{1.54}$$

where now

$$R_{N,s}(m) \;=\; \sum_{\substack{p_1+\cdots+p_s=m\\ p_i\le N}} \log p_1 \cdots \log p_s. \tag{1.55}$$

Exercise 1.3.1. *Relate $R_{N,s}(m)$ and $r_N(m; s, P)$. For details, see §2.7.*

Thus, if we can show $R_{N,s}(m)$ is positive for N and m sufficiently large, then $r(m; s, P)$ is also positive.

1.3.2 Average Value of $|F_N(x)|^2$

We use the little-Oh notation (see definition **??**). Thus, $N + o(N)$ means the answer is N plus lower order terms. Recall

$$F_N(x) \;=\; \sum_{p\le N} \log p \cdot e(px) \tag{1.56}$$

Lemma 1.3.2. $|F_N(x)| \le N + o(N)$.

Proof. By the Prime Number Theorem, (1.48), we have

$$|F_N(x)| \quad = \quad \left|\sum_{p\le N} \log p \cdot e(px)\right| \;\le\; \sum_{p\le N} \log p \;=\; N + o(N). \tag{1.57}$$

□

Lemma 1.3.3. $F_N(0) = F_N(1) = N + o(N)$, *and* $F_N(\frac{1}{2}) = -N + o(N)$.

Proof. $F_N(0)$ and $F_N(1)$ are immediate, as $e(p \cdot 1) = 1$ for all p. For $F_N(\frac{1}{2})$, note

$$e\left(p \cdot \frac{1}{2}\right) \;=\; e^{\pi ip} \;=\; \begin{cases} -1 & \text{if } p \text{ is odd} \\ +1 & \text{if } p \text{ is even} \end{cases} \tag{1.58}$$

As there is only one even prime,

$$F_N\left(\frac{1}{2}\right) \quad = \quad \log 2 \;-\; \sum_{3\le p\le N} \log p, \tag{1.59}$$

and the argument proceeds as before. □

Exercise 1.3.4. *How large are $F_N(\frac{1}{4})$ and $F_N(\frac{3}{4})$? How big can $o(N)$ be in Lemma 1.3.3?*

Thus $F_N(x)$ is occasionally as large as N; in §1.3.5 we describe the x where $F_N(x)$ is large. We can, however, show that the average square of $F_N(x)$ is significantly smaller:

Lemma 1.3.5. *The average value of $|F_N(x)|^2$ is $N \log N + o(N \log N)$.*

Proof. The following trivial observation will be extremely useful in our arguments. Let $g(x)$ be a complex-valued function, and let $\overline{g}(x)$ be its complex conjugate. Then $|g(x)|^2 = g(x)\overline{g}(x)$. In our case, as $\overline{F_N(x)} = F_N(-x)$ we have

$$\begin{aligned}\int_0^1 |F_N(x)|^2 dx &= \int_0^1 F_N(x)F_N(-x)dx \\ &= \int_0^1 \sum_{p\le N} \log p \cdot e(px) \sum_{q\le N} \log q \cdot e(-qx)dx \\ &= \sum_{p\le N}\sum_{q\le N} \log p \log q \int_0^1 e\left((p-q)x\right) dx. \end{aligned} \tag{1.60}$$

By exercise 1.2.7, the integral is 1 if $p = q$ and 0 otherwise. Therefore the only pairs (p, q) that contribute are when $p = q$, and we have

$$\int_0^1 |F_N(x)|^2 dx = \sum_{p\le N} \log^2 p. \tag{1.61}$$

Using partial summation (see exercise 1.3.9), we can show

$$\sum_{p\le N} \log^2 p \;=\; N \log N \;+\; o(N \log N). \tag{1.62}$$

Thus

$$\int_0^1 |F_N(x)|^2 dx \;=\; N \log N \;+\; o(N \log N). \tag{1.63}$$

□

Remark 1.3.6. *The above argument is extremely common. The absolute value function is not easy to work with; however, $g(x)\overline{g(x)}$ is very tractable (see also §**??**). In many problems, it is a lot easier to study $\int |g(x)|^2$ than $\int |g(x)|$ or $\int |g(x)|^3$.*

Remark 1.3.7 (Philosophy of Square-root Cancellation). *The average value of $|F_N(x)|^2$ is about $N \log N$, significantly smaller than the maximum possible value of N^2. Thus, we have almost square-root cancellation on average. In general, if one adds a "random" set of N numbers of absolute value 1, the sum could be as large as N, but often is at most of size $\sqrt{N}$. For more details and examples, see §?? and §??.*

Exercise 1.3.8. *Investigate the size of $\sum_{x=0}^{p-1} e^{2\pi i x^2/p}$ for p prime.* Hint: *rewrite the sum as two sums by using the Legendre symbol (see §??).*

Exercise 1.3.9. *Using the Prime Number Theorem and Partial Summation, prove*

$$\sum_{p \le N} \log^2 p \;=\; N \log N \;+\; o(N \log N). \tag{1.64}$$

1.3.3 Large Values of $F_N(x)$

For a fixed B, let $Q = \log^B N$. Fix a $q \le Q$ and an $a \le q$ with a and q relatively prime. We evaluate $F_N\left(\frac{a}{q}\right)$. While on average $F_N(x)$ is of size $\sqrt{N \log N}$, for x near such $\frac{a}{q}$ we shall see that $F_N(x)$ is large.

$$F_N\left(\frac{a}{q}\right) \;=\; \sum_{p \le N} \log p \cdot e\left(p\frac{a}{q}\right). \tag{1.65}$$

The summands on the right hand side depend weakly on p. Specifically, the exponential terms only depend on $p \bmod q$, which allows us to rewrite $F_N\left(\frac{a}{q}\right)$ as a sum over congruence classes:

$$\begin{aligned} F_N\left(\frac{a}{q}\right) &= \sum_{r=1}^{q} \sum_{\substack{p \equiv r(q) \\ p \le N}} \log p \cdot e\left(\frac{ap}{q}\right) \\ &= \sum_{r=1}^{q} \sum_{\substack{p \equiv r(q) \\ p \le N}} \log p \cdot e\left(\frac{ar}{q}\right) \\ &= \sum_{r=1}^{q} e\left(\frac{ar}{q}\right) \sum_{\substack{p \equiv r(q) \\ p \le N}} \log p. \end{aligned} \tag{1.66}$$

We use the Siegel-Walfisz Theorem to evaluate the sum over $p \equiv r \bmod q$. We first remark that we may assume r and q are relatively prime (see exercise 1.3.10). Briefly, if $p \equiv r \bmod q$, this means $p = \alpha q + r$ for some $\alpha \in \mathbb{N}$. If r and q have a common factor, there can be at most one prime p (namely r) such

that $p \equiv r \bmod q$, and this can easily be shown to give a negligible contribution. For any $C > 0$, by the Siegel-Walfisz Theorem

$$\sum_{\substack{p \equiv r(q) \\ p \le N}} \log p \;=\; \frac{N}{\phi(q)} \;+\; O\left(\frac{N}{\log^C N}\right). \tag{1.67}$$

As $\phi(q)$ is at most q which is at most $\log^B N$, we see that if we take $C > B$ then the main term is significantly greater than the error term. Note the Siegel-Walfisz Theorem would be useless if q were large, say $q \approx N^\delta$. Then the main term would be like $N^{1-\delta}$, which would be smaller than the error term. Thus we find

$$\begin{aligned} F_N\left(\frac{a}{q}\right) &= \sum_{\substack{r=1 \\ (r,q)=1}}^{q} e\left(\frac{ar}{q}\right) \frac{N}{\phi(q)} \;+\; O\left(\frac{qN}{\log^C N}\right) \\ &= \frac{N}{\phi(q)} \sum_{\substack{r=1 \\ (r,q)=1}}^{q} e\left(\frac{ar}{q}\right) \;+\; O\left(\frac{N}{\log^{C-B} N}\right). \end{aligned} \tag{1.68}$$

If the sum over r in (1.68) is not too small, then $F_N\left(\frac{a}{q}\right)$ is "approximately" of size $\frac{N}{q}$, with an error of size $\frac{N}{\log^{C-B} N}$. If $C > 2B$, the main term is significantly larger than the error term, and $F_N\left(\frac{a}{q}\right)$ is large.

The Siegel-Walfisz Theorem is our main tools for evaluating the necessary prime sums, and it is useful only when the error term is less than the main term. Our investigations of the (potential) size of $F_N(x)$ lead us to the proper definitions for the Major and Minor arcs in §1.3.4.

Exercise 1.3.10. *Show the terms with r and q not relatively prime in* (1.66) *contribute lower order terms.*

1.3.4 Definition of the Major and Minor Arcs

We split $[0, 1]$ into two disjoint parts, the Major and the Minor arcs. As $|F_N(x)|^2$ is of size $N \log N$ on average, there is significant cancellation in $F_N(x)$ most of the time. The Major arcs will be a union of very small intervals centered at rationals with small denominator relative to N. Near these rationals we can approximate $F_N(x)$ very well, and $F_N(x)$ will be large (of size N). The Minor arcs will be the rest of $[0, 1]$, and here we expect $F_N(x)$ to be significantly smaller than N. Obtaining such cancellation in the series expansion is *not* easy – this is the hardest part of the problem. In many cases we are unable to prove the integral over the Minor arcs is smaller than the contribution from the Major arcs, though we often believe this is the case, and numerical investigations support such claims.

Major Arcs

The choice of the Major arcs depend on the problem being investigated. In problems where the Siegel-Walfisz Theorem is used, the results from §1.3.3 suggest the following choice. Let $B > 0$, and let $Q = \log^B N \ll N$. For each $q \in \{1, 2, \dots, Q\}$ and $a \in \{1, 2, \dots, q\}$ with a and q relatively prime, consider the set

$$\mathcal{M}_{a,q} = \left\{ x \in [0,1] : \left| x - \frac{a}{q} \right| < \frac{Q}{N} \right\}. \tag{1.69}$$

We also add in one interval centered at either 0 or 1, i.e., the interval (or wrapped-around interval)

$$\left[0, \frac{Q}{N}\right) \cup \left(1 - \frac{Q}{N}, 1\right]. \tag{1.70}$$

Exercise 1.3.11. *Show that if N is large then the Major arcs $\mathcal{M}_{a,q}$ are disjoint for $q \le Q$ and $a \le q$, a and q relatively prime.*

We define the Major arcs to be the union of the arcs $\mathcal{M}_{a,q}$:

$$\mathcal{M} = \bigcup_{q=1}^{Q} \bigcup_{\substack{a=1 \\ (a,q)=1}}^{q} \mathcal{M}_{a,q}, \tag{1.71}$$

where (a, q) is the greatest common divisor of a and q.

Remark 1.3.12. *As the Major arcs depend on N and B, we should write $\mathcal{M}_{a,q}(N, B)$ and $\mathcal{M}(N, B)$; however, for notational convenience these subscripts are often suppressed.*

Exercise 1.3.13. *Show $|\mathcal{M}| \le \frac{2Q^3}{N}$. As $Q = \log^B N$, this implies $|\mathcal{M}| \to 0$ as $N \to \infty$. Thus in the limit most of $[0, 1]$ is contained in the Minor arcs; the choice of terminology reflects where $F_N(x)$ is large, and not which subset of $[0, 1]$ is larger.*

Note that the above choice for the Major arcs has two advantages. First, recall that we required the denominator q to be small relative to N: $q \le Q = \log^B N$. Once a denominator is small for some N, we can apply the Siegel-Walfisz Theorem and we can evaluate $F_N(\frac{a}{q})$ well (see §1.3.3). Second, each Major arc $\mathcal{M}_{a,q}$ has length $\frac{2Q}{N} = \frac{2\log^B N}{N}$; as these intervals are small, we expect $F_N(x) \approx F_N(\frac{a}{q})$. It should be possible to estimate the integral over $\mathcal{M}_{a,q}$. Thus, for a fixed $\frac{a}{q}$, the size of the arc about it tends to zero as N tends to infinity, but $F_N(x)$ becomes better and better understood in a smaller windows about $\frac{a}{q}$.

Exercise 1.3.14. *For large N, find a good asymptotic formula for $|\mathcal{M}|$.*

Exercise 1.3.15. *For a fixed B, how large must N be for the Major arcs to be disjoint?*

Minor Arcs

The Minor arcs, m, are whatever is *not* in the Major arcs. Thus,

$$\mathrm{m} \ = \ [0,1] - \mathcal{M}. \tag{1.72}$$

Clearly, as $N \to \infty$ almost all of $[0,1]$ is in the Minor arcs. The hope is that by staying away from rationals with small denominator, we will be able to obtain significant cancellation in $F_N(x)$.

1.3.5 The Major Arcs and the Singular Series

We are trying to write m as a sum of s primes. Let us consider the case $m = N$ and $s = 3$. We have shown the (weighted) answer is given by

$$\int_0^1 F_N(x)^3 e(-Nx)dx \ = \sum_{\substack{p_1,p_2,p_3 \le N \\ p_1+p_2+p_3=N}} \log p_1 \log p_2 \log p_3; \tag{1.73}$$

the weights can easily be removed by partial summation. We merely sketch what happens now; we handle a Major arc calculation in full detail in Chapter 2.

First one shows that for $x \in \mathcal{M}_{a,q}$, $F_N(x)$ is very close to $F_N\left(\frac{a}{q}\right)$. While one could calculate the Taylor Series expansion (see §**??**), in practice it is technically easier to find a function which is non-constant and agrees with $F_N(x)$ at $x = \frac{a}{q}$. As the Major arcs are disjoint for large N,

$$\int_{\mathcal{M}} F_N(x)^3 e(-Nx)dx \ = \ \sum_{q=1}^{Q} \sum_{\substack{a=1 \\ (a,q)=1}} \int_{\mathcal{M}_{a,q}} F_N(x)^3 e(-Nx)dx. \tag{1.74}$$

For heuristic purposes, we approximate $F_N(x)^3 e(-Nx)$ by $F_N\left(\frac{a}{q}\right)^3 e\left(-N\frac{a}{q}\right)$. After reading Chapter 2 the reader is encouraged to do these calculations correctly. Therefore

$$\int_{\mathcal{M}} F_N(x)^3 e(-Nx)dx \ \approx \ \int_{\mathcal{M}} F_N\left(\frac{a}{q}\right)^3 e\left(-N\frac{a}{q}\right) dx \ = \ F_N\left(\frac{a}{q}\right)^3 e\left(-N\frac{a}{q}\right) \cdot \frac{2Q}{N}. \tag{1.75}$$

In (1.68) we used the Siegel-Walfisz Theorem to evaluate $F_N\left(\frac{a}{q}\right)$. Again, for heuristic purposes we

suppress the lower order error terms, and find that the contribution from the Major arcs is

$$\begin{aligned}\sum_{\substack{p_1,p_2,p_3\le N\\ p_1+p_2+p_3=N}} \log p_1 \log p_2 \log p_3 &= \frac{2Q^3}{N}\sum_{q=1}^{Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\left(\frac{N}{\phi(q)}\sum_{\substack{r=1\\(r,q)=1}}^{q} e\left(\frac{ar}{q}\right)\right)^3 e\left(\frac{-Na}{q}\right)\\ &= \left[2Q^3\sum_{q=1}^{Q}\frac{1}{\phi(q)^3}\sum_{\substack{a=1\\(a,q)=1}}^{q}\left(\sum_{\substack{r=1\\(r,q)=1}}^{q} e\left(\frac{ar}{q}\right)\right)^3 e\left(\frac{-Na}{q}\right)\right]N^2 \end{aligned} \tag{1.76}$$

To complete the proof, we need to show that what multiplies N^2 is positive and not too small. If N^2 were multiplied by $\frac{1}{N^3}$, for example, the main term from the Major arcs would be of size $\frac{1}{N}$, which could easily be cancelled by the contribution from the Minor arcs. An elementary analysis often bounds the factor away from 0 and infinity.

Note that, up to factors of $\log N$ (which are important!), the contribution from the Major arcs is of size N^2. A more careful analysis, where we do not just replace $f_N(x)^3 e(-Nx)$ with $f_N(\frac{a}{q})^3 e\left(-N\frac{a}{q}\right)$ on $\mathcal{M}_{a,q}$, would show that the Major arcs contribute

$$\mathfrak{S}(N)\frac{N^2}{2} \ + \ o(N^2), \tag{1.77}$$

with

$$\mathfrak{S}(N) \ = \ \sum_{q=1}^{\infty}\frac{1}{\phi(q)^3}\sum_{\substack{a=1\\(a,q)=1}}^{q}\left(\sum_{\substack{r=1\\(r,q)=1}}^{q} e\left(\frac{ar}{q}\right)\right)^3 e\left(-N\frac{a}{q}\right). \tag{1.78}$$

$\mathfrak{S}(N)$ is called the **Singular Series**; in all Circle Method investigations, the contribution from the Major arcs is given by such a series. The singular series for Germain primes will be discussed in detail in Chapter 2; for complete details on the singular series for sums of three primes, the interested reader should see [EE, Na]. If we set

$$c_p(N) \ = \ \begin{cases} p-1 & \text{if } p|N \\ 0 & \text{otherwise} \end{cases} \tag{1.79}$$

then one can show

$$\mathfrak{S}(N) \ = \ \prod_p \left(1 - \frac{c_p(N)}{\phi(p)^3}\right). \tag{1.80}$$

The product expansion is a much more useful expression for the factor multiplying N^2 than the series expansion. If N is *odd*, there exist constants c_1 and c_2 such that

$$0 \ < \ c_1 \ < \ \mathfrak{S}(N) \ < \ c_2 \ < \ \infty. \tag{1.81}$$

This allows us to conclude the Major arcs' contribution is of order N^2.

We do not go into great detail concerning the arithmetic properties of $\mathfrak{S}(N)$, and content ourselves with an important observation. If $\mathfrak{S}(N) > c_1$ for all N, then the main term will be greater than the error term for N sufficiently large. Can $\mathfrak{S}(N)$ ever vanish?

Consider N even. Then $c_2(N) = 1$, $\phi(2) = 1$, and the factor in $\mathfrak{S}(N)$ corresponding to $p = 2$ vanishes! Thus, for even N, the main term from the Circle Method is zero. In hindsight, this is not surprising. Assume an even $N > 6$ can be written as the sum of three primes. Exactly one of the primes must be even (if all or exactly one were odd, then N would be odd; if all were even, N would be 6). Therefore, if the Circle Method tells us that we can write an even N as the sum of three primes, we could immediately conclude that $N - 2$ is the sum of two primes.

The Singular Series "knows" about the difficulty of Goldbach. For many Circle Method problems, one is able to write the main term from the Major arcs (up to computable constants and factors of $\log N$) as $\mathfrak{S}(N)N^a$, with $\mathfrak{S}(N)$ a product over primes. The factors at each prime often encode information about obstructions to solving the original problem. For more on obstructions, see §**??**.

Exercise 1.3.16. *For N odd, show there exist positive constants c_1, c_2 (independent of N) such that $0 < c_1 < \mathfrak{S}(N) < c_2 < \infty$.*

Exercise 1.3.17. *In the spirit of exercise 1.1.13, we sketch a heuristic for the expected average value of the number of ways of writing n as a sum of k primes. Consider $n \in [N, 2N]$ for N large. Count the number of k-tuples of primes with $p_1 + \cdots + p_k \in [N, 2N]$. As there are approximately N even numbers in the interval, deduce the average number of representations for such n. What if we instead considered short intervals, such as $n \in [N, N^{1-\delta}]$ for some $\delta > 0$?*

Exercise 1.3.18. *Prove* (1.78) *implies the product representation in* (1.80). Hint: *many of the sums of arithmetic functions arise in the Germain prime investigations; see §2.6.*

Remark 1.3.19 (Goldbach). *If instead we investigate writing even numbers as the sum of two primes, we would integrate $F_N(x)^2$ and obtain a new singular series, say $\widetilde{\mathfrak{S}}(N)$. The Major arcs would then contribute $\widetilde{\mathfrak{S}}(N)N$.*

1.3.6 Contribution from the Minor Arcs

We bound the contribution from the Minor arcs to $R_{N,s}(N)$:

$$\begin{aligned}\left|\int_{\mathrm{m}} F_N(x)^3 e(-Nx)dx\right| &\leq \int_{\mathrm{m}} |F_N(x)|^3 dx \\ &\leq \Big(\max_{x\in\mathrm{m}} |F_N(x)|\Big)\int_{\mathrm{m}} |F_N(x)|^2 dx \\ &\leq \Big(\max_{x\in\mathrm{m}} |F_N(x)|\Big)\int_0^1 F_N(x)F_N(-x)dx \\ &\leq \Big(\max_{x\in\mathrm{m}} |F_N(x)|\Big) N\log N. \end{aligned} \tag{1.82}$$

As the Minor arcs are most of the unit interval, replacing $\int_{\mathrm{m}}$ with $\int_0^1$ does not introduce much of an over-estimation. *In order for the Circle Method to succeed, we need a non-trivial, good bound for*

$$\max_{x\in\mathrm{m}} |F_N(x)| \tag{1.83}$$

This is where most of the difficulty arises, showing that there is significant cancellation in $F_N(x)$ if we stay away from rationals with small denominator. We need an estimate such as

$$\max_{x\in\mathrm{m}} |F_N(x)| \ \leq\ \frac{N}{\log^{1+\epsilon} N}, \tag{1.84}$$

or even

$$\max_{x\in\mathrm{m}} |F_N(x)| \ \ll\ o\left(\frac{N}{\log N}\right). \tag{1.85}$$

Relative to the average size of $|F_N(x)|^2$, which is $N\log N$, this is significantly larger. Unfortunately, as we have inserted absolute values, it is not enough to bound $|F_N(x)|$ on average – we need to obtain a good bound uniformly in x. We know such a bound cannot be true for all $x \in [0,1]$, because $F_N(x)$ is large on the Major arcs! The hope is that if x is not near a rational with small denominator, we will obtain moderate cancellation. While this is reasonable to expect, it is not easy to prove; the interested reader should see [EE, Na]. Following Vinogradov [Vin1, Vin2] one shows

$$\max_{x\in\mathrm{m}} |F_N(x)| \ \ll\ \frac{N}{\log^{D} N}, \tag{1.86}$$

which allows one to deduce any large odd number is the sum of three primes. While (1.63) gives us significantly better cancellation on average, telling us that $|F_N(x)|^2$ is usually of size N, bounds such as (1.86) are the best we can do if we require the bound to hold for *all* $x \in \mathrm{m}$.

Exercise 1.3.20. *Using the definition of the Minor arcs, bound*

$$\left| \int_0^1 |F_N(x)^2| dx \ - \ \int_{\mathrm{m}} |F_N(x)^2| dx \right|. \tag{1.87}$$

Show, therefore, that there is little harm in extending the integral of $|F_N(x)^2|$ *to all of* $[0,1]$*. In general, there is very little loss of information in integrating* $|F_N(x)|^{2^k}$*.*

1.3.7 Why Goldbach's Conjecture is Hard

We give some arguments which indicate the difficulty of applying the Circle Method to Goldbach's conjecture. To investigate $R_{N,s}(N)$, the number of ways of writing N as the sum of s primes, we considered the generating function

$$F_N(x) \ = \ \sum_{p \leq N} \log p \cdot e\left(px\right), \tag{1.88}$$

which led to

$$R_{N,s}(N) \ = \ \int_0^1 F_N(x)^s e(-Nx) dx. \tag{1.89}$$

Remember that the average size of $|F_N(x)|^2$ is $N \log N$.

We have seen that, up to logarithms, the contribution from the Major arcs is of size N^2 for $s = 3$. Similar arguments show that the Major arcs contribute on the order of N^{s-1} for sums of s primes. We now investigate why the Circle Method works for $s = 3$ but fail for $s = 2$.

When $s = 3$, we can bound the Minor arcs contribution by

$$\begin{aligned} \left| \int_{\mathrm{m}} F_N(x)^3 e(-Nx) dx \right| \ &\leq \ \max_{x \in \mathrm{m}} |F_N(x)| \int_0^1 |F_N(x)|^2 dx \\ &\leq \ \max_{x \in \mathrm{m}} |F_N(x)| \cdot N \log N. \end{aligned} \tag{1.90}$$

As the Major arcs contribute $\mathfrak{S}(N)N^2$, one needs only a small savings on the Minor arcs; Vinogradov's bound

$$\max_{x \in \mathrm{m}} |F_N(x)| \ \ll \ \frac{N}{\log^D N}. \tag{1.91}$$

suffices. What goes wrong when $s = 2$? The Major arcs' contribution is now expected to be of size N. How should we estimate the contribution from the Minor arcs? We have $F_N(x)^2 e(-Nx)$. The simplest estimate to try is to just insert absolute values, which gives

$$\left| \int_{\mathrm{m}} F_N(x)^2 e(-Nx) dx \right| \ \leq \ \int_0^1 |F_N(x)|^2 dx \ = \ N. \tag{1.92}$$

Note, unfortunately, that this is the same size as the expected contribution from the Major arcs!

We could try pulling a $\max_{x\in\mathfrak{m}}|F_N(x)|$ outside the integral, and hope to get a good saving (pulling out $|F_N(x)|^2$ clearly cannot work as the maximum of this is at least $N\log N$). The problem is this leaves us with $\int_{\mathfrak{m}}|F_N(x)|dx$. As $F_N(x)$ on average is of size $\sqrt{N\log N}$ (this is not quite right: we have only shown $|F_N(x)|^2$ on average is $N\log N$; however, let us ignore this complication and see what bound we obtain), replacing $|F_N(x)|$ in the integral with its average value leads us to

$$\left|\int_{\mathfrak{m}} F_N(x)^2 e(-Nx)dx\right| \;\leq\; \max_{x\in\mathfrak{m}}|F_N(x)|\cdot\sqrt{N\log N}. \tag{1.93}$$

As the Major arcs' contribution is of size N, we would need

$$\max_{x\in\mathfrak{m}}|F_N(x)| \;\ll\; o\left(\sqrt{\frac{N}{\log N}}\right). \tag{1.94}$$

There is no chance of such cancellation; this is better than square-root cancellation, and contradicts the average value of $|F_N(x)|^2$ from (1.63).

Another approach is to use the Cauchy-Schwarz Inequality (see Lemma **??**):

$$\int_0^1 |f(x)g(x)|dx \;\leq\; \left(\int_0^1 |f(x)|^2dx\right)^{\frac{1}{2}}\cdot\left(\int_0^1 |g(x)|^2dx\right)^{\frac{1}{2}}. \tag{1.95}$$

Thus

$$\begin{aligned}\left|\int_{\mathfrak{m}} F_N(x)^2 e(-mx)dx\right| &\leq \max_{x\in\mathfrak{m}}|F_N(x)|\int_0^1 |F_N(x)|dx\\ &\leq \max_{x\in\mathfrak{m}}|F_N(x)|\left(\int_0^1 |F_N(x)|^2dx\right)^{\frac{1}{2}}\cdot\left(\int_0^1 1^2dx\right)^{\frac{1}{2}}\\ &\leq \max_{x\in\mathfrak{m}}|F_N(x)|\cdot(N\log N)^{\frac{1}{2}}\cdot 1.\end{aligned} \tag{1.96}$$

Unfortunately, this is the same bound as (1.93), which was too large.

Remark 1.3.21. *Even though it failed, it was a good idea to use the Cauchy-Schwartz inequality. The reason is we are integrating over a finite interval; thus $\int_0^1 1^2dx$ is harmless; if the size of the interval depended on N (or was all of $\mathbb{R}$), applying Cauchy-Schwartz might be a mistake.*

While the above sketch shows the Circle Method is not, at present, powerful enough to handle the Minor arcs' contribution, all is not lost. The quantity we *need* to bound is

$$\left|\int_{\mathfrak{m}} F_N(x)^2 e(-mx)dx\right|. \tag{1.97}$$

However, we have instead been studying

$$\int_{\mathfrak{m}} |F_N(x)|^2 dx \tag{1.98}$$

and

$$\max_{x\in\mathfrak{m}} |F_N(x)| \int_0^1 |F_N(x)| dx. \tag{1.99}$$

We are ignoring the probable oscillation and cancellation in the integral $\int_{\mathfrak{m}} F_N(x)^2 e(-mx)dx$. It is this expected cancellation that would lead to the Minor arcs contributing significantly less than the Major arcs.

However, showing there is cancellation in the above integral is very difficult. It is a lot easier to work with absolute values. Further, just because we cannot prove that the Minor Arc contribution is small, does not mean the Circle Method is not useful. Numerical simulations confirm, for many problems, that the Minor arcs do not contribute for many N. For example, for $N \le 10^9$, the observed values are in excellent agreement with the Major arc predictions (see [Ci, Sch, Weir]).**ADD DATA??**

Chapter 2

Circle Method: Heuristics for Germain Primes

We apply the Circle Method to investigate Germain primes. As current techniques are unable to adequately bound the Minor arc contributions, we concentrate on the Major arcs, where we perform the calculations in great detail The methods of this chapter immediately generalize to other standard problems, such as investigating twin primes or prime tuples.

We have chosen to describe the Circle Method for Germain primes as this problem highlights many of the complications that arise in applications. Unlike the previous investigations of writing N as a sum of s primes, our generating function $F_N(x)$ is the product of two different generating functions. To approximate $F_N(x)$ on the Major arc $\mathcal{M}_{a,q}$, we could try to Taylor expand; however, the derivative is not easy to analyze or integrate. Instead we construct a new function which is easy to integrate on $[-\frac{1}{2}, \frac{1}{2}]$, has most of its mass concentrated near $\frac{a}{q}$, and is a good approximation to $F_N(x)$ on $\mathcal{M}_{a,q}$. To show the last claim requires multiple applications of partial summation. For numerical investigations of the Minor arcs, as well as spacing properties of Germain primes, see [Weir].

In §2.1 and §2.2 we define Germain primes, the generating function $F_N(x)$, and the Major and Minor arcs. In §2.3 we estimate $F_N(x)$ and find an easily integrable function $u(x)$ which should be close to $F_N(x)$ on the Major arcs. We prove $u(x)$ is a good approximation to $F_N(x)$ in §2.4; this is a technical section and can easily be skimmed on a first reading. We then determine the contribution from the Major arcs by performing the integration in §2.5 and then analyzing the singular series in §2.6. Finally, in §2.7 we remove the $\log p$ weights and then conclude with some exercises and open problems.

2.1 Germain Primes

Consider an odd prime p. Clearly $p-1$ cannot be prime, as it is even; however, $\frac{p-1}{2}$ could be prime, and sometimes is as $p = 5, 7$ and 11 show.

Definition 2.1.1 (Germain Prime). *A prime p is a Germain prime (or p and $\frac{p-1}{2}$ are a Germain prime*

pair) if both p and $\frac{p-1}{2}$ are prime. An alternate definition is to have p and $2p+1$ both prime.

Germain primes have many wonderful properties. Around 1825, Sophie Germain proved that if p is a Germain prime, then the first case of Fermat's Last Theorem, which states the only integer solutions of $x^p + y^p = z^p$ have $p|xyz$, is true for exponent p. For more on Fermat's Last Theorem, see §**??**. As another application, recent advances in cryptography are known to run faster if there are many Germain primes (see [AgKaSa]).

Germain primes are just one example of the following type of problem: Given relatively prime positive integers a and b, for $p \le N$ how often are p and $ap+b$ prime? Or, more generally, how often are $p, a_1p+b_1, \dots, a_kp+b_k$ prime? One well known example is the famous Twin Prime Conjecture, which states that there are infinitely many primes p such that $p+2$ is also prime. It is not known if this is true. Unlike the sum of the reciprocals of the primes, which diverges, Brun has shown that the sum of the reciprocal of the twin primes converges (see [Na]). Therefore, if there are infinitely many twin primes, there are in some sense fewer twin primes than primes. Explicitly, Brun proved that there exists an N_0 such that for all $N > N_0$, the number of twin primes less than N is at most $\frac{100N}{\log^2 N}$. This should be compared to the number of primes less than N, which is of size $\frac{N}{\log N}$. Using the Circle Method, Hardy and Littlewood were led to conjectures on the number of such primes, and their Major arc calculations agree beautifully with numerical investigations.

We have chosen to go through the calculation of the number of Germain primes less than N rather than twin primes as these other problems are well documented in literature ([Da2, EE, Est2, Na]). Note the Germain problem is slightly different from the original formulation of the Circle Method. Here, we are investigating how often $p_1 - 2p_2 = 1$, with $p_2 < p_1 < N$. Let

$$\begin{aligned} A_1 &= \{p : p \text{ prime}\} \\ A_2 &= \{-2p : p \text{ prime}\}. \end{aligned} \tag{2.1}$$

To construct generating functions that converge, we consider the truncated sets

$$\begin{aligned} A_{1N} &= \{p : p \text{ prime}, p \le N\} \\ A_{2N} &= \{-2p : p \text{ prime}, p \le N\}. \end{aligned} \tag{2.2}$$

We are interested in

$$\{(a_1, a_2) : a_1 + a_2 = 1, a_i \in A_{iN}\}. \tag{2.3}$$

In the original applications of the Circle Method, we were just interested in whether or not a number m was in $A_N + \cdots + A_N$. To show m could be written as the sum of elements $a_i \in A_N$, we counted the number of ways to write it as such a sum, and showed it was positive.

For Germain primes and related problems, we are no longer interested in determining all numbers that can be written as the sum $a_1 + a_2$. We only want to find pairs with $a_1 + a_2 = 1$. The common feature

with our previous investigations is showing how many ways certain numbers can be written as the sum of elements in A_{iN}. Such knowledge gives estimates for the number of Germain primes at most N.

For any $N \geq 5$ we know $1 \in A_{1N} + A_{2N}$ as 5 is a Germain prime. Note the number of ways of writing 1 as $a_1 + a_2$ with $a_i \in A_{iN}$ is the number of Germain primes at most N; similar to before, we need to compute for this problem $r(1; A_{1N}, A_{2N})$, with the obvious notation.

Exercise 2.1.2. *Looking at tables of primes less than* 100*, do you think there will be more Germain primes or twin primes in the limit? What if you study primes up to* 10^4*? Up to* 10^8*? What percent of primes less than* 100 *(*$10^4, 10^8$*) are Germain primes? Twin primes? How many primes less than* N *(for* N *large) do you expect to be Germain primes? Twin primes?*

Exercise 2.1.3. *By the prime number theorem, for primes near* x *the average spacing between primes is* $\log x$*. One can interpret this as the probability a number near* x *is prime is* $\frac{1}{\log x}$*. We flip a biased coin with probability* $\frac{1}{\log x}$ *of being a prime,* $1 - \frac{1}{\log x}$ *of being composite; this is called the* ***Cramér model****. Using such a model predict how many Germain primes and twin primes are less than* N*.*

Remark 2.1.4 (Remark on the previous exercise). *The Cramér model of the previous exercise cannot be correct – knowledge that* p *is prime gives some information about the potential primality of nearby numbers. One needs to correct the model to account for congruence information. See §***??** *and [Rub1].*

2.2 Preliminaries

We use the Circle Method to calculate the contribution from the Major arcs for the Germain problem, namely, how many primes $p \leq N$ there are such that $\frac{p-1}{2}$ is also prime. As pointed out earlier, for this problem the Minor Arc calculations cannot be determined with sufficient accuracy to be shown to be smaller than the Major arc contributions. We will, however, do the Major arc calculations in complete detail. Let

$$\begin{aligned} e(x) &= e^{2\pi i x} \\ \lambda(n) &= \begin{cases} \log p & \text{if } n = p \text{ is prime;} \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{2.4}$$

We constantly use the integral version of partial summation (§**??**) and the Siegel-Walfisz Theorem (Theorem 1.2.13). We have introduced the arithmetic function $\lambda(n)$ for notational convenience. In applying partial summation, we will have sums over integers, but our generating function is defined as a sum over primes; $\lambda(n)$ is a convenient notation which allows us to write the sum over primes as a sum over integers.

2.2.1 Germain Integral

Define

$$\begin{aligned} F_{1N}(x) &= \sum_{p_1 \le N} \log p_1 \cdot e(p_1 x) \\ F_{2N}(x) &= \sum_{p_2 \le N} \log p_2 \cdot e(-2p_2 x) \\ F_N(x) &= \sum_{p_1 \le N} \sum_{p_2 \le N} \log p_1 \log p_2 \cdot e\left((p_1 - 2p_2)x\right) \;=\; F_{1N}(x) F_{2N}(x). \end{aligned} \tag{2.5}$$

$F_N(x)$ is the generating function for the Germain primes. As $F_N(x)$ is periodic with period 1, we can integrate either over $[0,1]$ or $[-\frac{1}{2}, \frac{1}{2}]$. We choose the latter because the main contribution to the integral is from x near 0, although both choices obviously yield the same result. Letting $r(1; A_{1N}, A_{2N})$ denote the weighted number of Germain primes, we have

$$\begin{aligned} r(1; A_{1N}, A_{2N}) &= \int_{-\frac{1}{2}}^{\frac{1}{2}} F_N(x) e(-x) dx \\ &= \sum_{p_1 \le N} \sum_{p_2 \le N} \log p_1 \log p_2 \int_{-\frac{1}{2}}^{\frac{1}{2}} e\left((p_1 - 2p_2 - 1)x\right) dx. \end{aligned} \tag{2.6}$$

By exercise 1.2.7

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} e\left((p_1 - 2p_2 - 1)x\right) dx \;=\; \begin{cases} 1 & \text{if } p_1 - 2p_2 - 1 = 0 \\ 0 & \text{if } p_1 - 2p_2 - 1 \neq 0 \end{cases} \tag{2.7}$$

In (2.6) we have a contribution of $\log p_1 \log p_2$ if p_1 and $p_2 = \frac{p_1 - 1}{2}$ are both prime and 0 otherwise. Thus

$$r(1; A_{1N}, A_{2N}) \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} F_N(x) e(-x) dx \;=\; \sum_{\substack{p_1 \le N \\ p_1, p_2 = \frac{p_1 - 1}{2} \text{ prime}}} \log p_1 \log p_2. \tag{2.8}$$

The above is a weighted counting of Germain primes. We have introduced these weights to facilitate applying the Siegel-Walfisz formula; it is easy to pass from bounds for $r(1; A_{1N}, A_{2N})$ to bounds for the number of Germain primes (see §2.7).

Remark 2.2.1. *Using the λ-function from* (2.4)*, we can rewrite the generating function as a sum over pairs of integers instead of pairs of primes:*

$$F_N(x) \;=\; \sum_{m_1=1}^{N} \sum_{m_2=1}^{N} \lambda(m_1)\lambda(m_2) \cdot e((m_1 - 2m_2)x). \tag{2.9}$$

Of course, the two functions are the same; sometimes it is more convenient to use one notation over the other. When we apply partial summation, it is convenient if our terms are defined for all integers, and not just at primes.

Exercise 2.2.2. *Determine (or at least bound) the average values of $F_{1N}(x)$, $F_{2N}(x)$ and $F_N(x)$.* Hint: *for $F_N(x)$, use the Cauchy-Schwartz inequality.*

2.2.2 The Major and Minor Arcs

Let B, D be positive integers with $D > 2B$. Set $Q = \log^D N$. Define the Major arc $\mathcal{M}_{a,q}$ for each pair (a, q) with a and q relatively prime and $1 \le q \le \log^B$ by

$$\mathcal{M}_{a,q} \;=\; \left\{ x \in \left(-\frac{1}{2},\, \frac{1}{2}\right) :\; \left| x - \frac{a}{q} \right| \;<\; \frac{Q}{N} \right\} \tag{2.10}$$

if $\frac{a}{q} \neq \frac{1}{2}$ and

$$\mathcal{M}_{1,2} \;=\; \left[-\frac{1}{2},\, -\frac{1}{2} + \frac{Q}{N} \right) \bigcup \left(\frac{1}{2} - \frac{Q}{N},\, \frac{1}{2} \right]. \tag{2.11}$$

Remember, as our generating function is periodic with period 1, we can work on either $[0, 1]$ or $[-\frac{1}{2}, \frac{1}{2}]$. As the Major arcs depend on N and D, we should write $\mathcal{M}_{a,q}(N, D)$ and $\mathcal{M}(N, D)$; however, for notational convenience these subscripts are often suppressed. Note we are giving ourselves a little extra flexibility by having $q \le \log^B N$ and each $\mathcal{M}_{a,q}$ of size $\frac{\log^D N}{N}$. We see in §2.5 why we need to have $D > 2B$.

By definition, the Minor arcs m are whatever is not in the Major arcs. Thus the Major arcs are the subset of $[-\frac{1}{2}, \frac{1}{2}]$ near rationals with small denominators, and the Minor arcs are what is left. Here near and small are relative to N. Then

$$r(1; A_{1N}, A_{2N}) \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} F_N(x) e(-x) dx \;=\; \int_{\mathcal{M}} F_N(x) e(-x) dx \;+\; \int_{\mathrm{m}} F_N(x) e(-x) dx. \tag{2.12}$$

We will calculate the contribution to $r(1; A_{1N}, A_{2N})$ from the Major arcs, and then in §2.7 we remove the $\log p_i$ weights.

We chose the above definition for the Major arcs because our main tool for evaluating $F_N(x)$ is the Siegel-Walfisz formula (Theorem 1.2.13), which states that given any $B, C > 0$, if $q \le \log^B N$ and $(r, q) = 1$ then

$$\sum_{\substack{p \le N \\ p \equiv r(q)}} \log p \;=\; \frac{N}{\phi(q)} + O\left(\frac{N}{\log^C N} \right). \tag{2.13}$$

For C very large, the error term leads to small, manageable errors on the Major arcs. For a more detailed explanation of this choice for the Major arcs, see §1.3.3 and §1.3.4.

We show the Major arcs contribute, up to lower order terms, $T_2 N$, where T_2 is a constant independent of N. By choosing B, C and D sufficiently large we can ensure that the errors from the Major arc calculations are less than the main term from the Major arcs. Of course, we have absolutely no control over what happens on the Minor arcs. Similar to Chapter 1 (see §**??**), up to powers of $\log N$ we have $F_N(x)$ on average is of size N, but is of size N^2 on the Major arcs. As there is a lot of oscillation in the generating function $F_N(x)$, for generic $x \in [-\frac{1}{2}, \frac{1}{2}]$ we expect a lot of cancellation in the size of $F_N(x)$. Unfortunately, we are unable to prove that this oscillation yields the Minor arcs contributing less than the Major arcs.

We highlight the upcoming calculations. On the Major arcs $\mathcal{M}_{a,q}$, we find a function u of size N^2 such that the error from u to F_N on $\mathcal{M}_{a,q}$ is much smaller than N^2, say N^2 divided by a large power of $\log N$. When we integrate u over the Major arcs, we find the main term is of size N (because up to powers of $\log N$ the Major arcs are of size $\frac{1}{N}$), and we succeed if we can show the errors in the approximations are much smaller than N, say N divided by a large power of $\log N$. Numerical simulations for x up to 10^9 and higher support the conjecture that the Minor arcs do not contribute for the Germain problem. Explicitly, the observed number of Germain prime pairs in this range agrees with the prediction from the Major arcs (see [Weir]). We content ourselves with calculating the contribution from the Major arcs,

$$\int_{\mathcal{M}} F_N(x) e(-x) dx. \tag{2.14}$$

2.3 $F_N(x)$ and $u(x)$

After determining F_N on $\mathcal{M}_{a,q}$, we describe an easily integrable function which is close to F_N on $\mathcal{M}_{a,q}$. We calculate $F_N\left(\frac{a}{q}\right)$ for $q \le \log^B N$. Define the **Ramanujan sum** $c_q(a)$ by

$$c_q(a) \;=\; \sum_{\substack{r=1 \\ (r,q)=1}}^{q} e\left(r\frac{a}{q}\right). \tag{2.15}$$

As usual, we evaluate $F_N\left(\frac{a}{q}\right)$ with the Siegel-Walfisz Theorem (Theorem 1.2.13). We restrict below to $(r_i, q) = 1$ because if $(r_i, q) > 1$, there is at most one prime $p_i \equiv r_i \bmod q$, and one prime will give a

negligible contribution as $N \to \infty$. See also §1.3.5. We have

$$\begin{aligned} F_N\left(\frac{a}{q}\right) &= \sum_{p_1 \le N} \log p_1 \cdot e\left(p_1\frac{a}{q}\right) \sum_{p_2 \le N} \log p_2 \cdot e\left(-2p_2\frac{a}{q}\right) \\ &= \sum_{r_1=1}^{q} \sum_{\substack{p_1 \le N \\ p_1 \equiv r_1(q)}} \log p_1 \cdot e\left(p_1\frac{a}{q}\right) \sum_{r_2=1}^{q} \sum_{\substack{p_2 \le N \\ p_2 \equiv r_1(q)}} \log p_2 \cdot e\left(-2p_2\frac{a}{q}\right) \\ &= \sum_{r_1=1}^{q} e\left(r_1\frac{a}{q}\right) \sum_{r_2=1}^{q} e\left(r_2\frac{-2a}{q}\right) \sum_{\substack{p_1 \le N \\ p_1 \equiv r_1(q)}} \log p_1 \sum_{\substack{p_2 \le N \\ p_2 \equiv r_2(q)}} \log p_2 \\ &= \sum_{\substack{r_1=1 \\ (r_1,q)=1}}^{q} e\left(r_1\frac{a}{q}\right) \sum_{\substack{r_2=1 \\ (r_2,q)=1}}^{q} e\left(r_2\frac{-2a}{q}\right) \left[\frac{N}{\phi(q)} + O\left(\frac{N}{\log^C N}\right)\right]^2 \\ &= \frac{c_q(a)c_q(-2a)}{\phi(q)^2}N^2 + O\left(\frac{N^2}{\log^{C-B} q}\right). \end{aligned} \tag{2.16}$$

Exercise 2.3.1. *Show that the contribution from one prime may safely be absorbed by the error term.*

Let

$$u(x) = \sum_{m_1=1}^{N} \sum_{m_2=1}^{N} e\left((m_1 - 2m_2)x\right). \tag{2.17}$$

As $u(0) = N^2$, it is natural to compare $F_N(x)$ on the Major arcs to

$$\frac{c_q(a)c_q(-2a)}{\phi(q)^2}\, u\left(x - \frac{a}{q}\right), \tag{2.18}$$

as these two functions agree at $x = \frac{a}{q}$. The function $u(x)$ is a lot easier to analyze than $F_N(x)$. We show for $x \in \mathcal{M}_{a,q}$ that there is negligible error in replacing $F_N(x)$ with $\frac{c_q(a)c_q(-2a)}{\phi(q)^2}u(x - \frac{a}{q})$. We then integrate over $\mathcal{M}_{a,q}$, and then sum over all Major arcs. We describe in great detail in Remark 2.5.9 why it is natural to consider $u(x)$.

2.4 Approximating $F_N(x)$ on the Major arcs

In this technical section we apply partial summation multiple times to show u is a good approximation to F_N on the Major arcs $\mathcal{M}_{a,q}$. Define

$$C_q(a) = \frac{c_q(a)c_q(-2a)}{\phi(q)^2}. \tag{2.19}$$

We show

Theorem 2.4.1. *For $\alpha \in \mathcal{M}_{a,q}$,*

$$F_N(\alpha) \;=\; C_q(a)u\left(\alpha - \frac{a}{q}\right) + O\left(\frac{N^2}{\log^{C-2D} N}\right). \tag{2.20}$$

For $\alpha \in \mathcal{M}_{a,q}$, we write α as $\beta + \frac{a}{q}$, $\beta \in \left[-\frac{Q}{N}, \frac{Q}{N}\right]$. Remember $Q = \log^D N$ and $q \le \log^B N$. Note $F_N(x)$ is approximately $C_q(a)N^2$ for x near $\frac{a}{q}$, and from our definitions of F_N, u and $C_q(a)$, (2.20) is immediate for $\alpha = \frac{a}{q}$. The reader interested in the main ideas of the Circle Method may skip to §2.5, where we integrate $u(x)$ over the Major arcs. The rest of this section is devoted to rigorously showing that $|F_N(x) - C_q(a)u(x - \frac{a}{q})|$ is small.

The calculation below is a straightforward application of partial summation. The difficulty is that we must apply partial summation twice. Each application yields two terms, a boundary term and an integral term. We will have four pieces to analyze. The problem is to estimate the difference

$$S_{a,q}(\alpha) \;=\; F_N(\alpha) - C_q(a)u\left(\alpha - \frac{a}{q}\right) \;=\; F_N\left(\beta + \frac{a}{q}\right) - C_q(a)u(\beta). \tag{2.21}$$

Recall that $q \le \log^B N$ and $F_N(\frac{a}{q}) = C_q(a)u(0)$ is of size $\frac{N^2}{\phi(q)^2}$. To prove Theorem 2.4.1 we must show that $|S_{a,q}(\alpha)| \le \frac{N^2}{\log^{C-2D} N}$. As mentioned in Remark 2.2.1, it is easier to apply partial summation if we use the λ-formulation of the generating function F_N because now both F_N and u will be sums over $m_1, m_2 \le N$. Thus

$$\begin{aligned}
S_{a,q}(\alpha) &= \sum_{m_1,m_2\le N} \lambda(m_1)\lambda(m_2)e\left((m_1-2m_2)\beta\right) \;-\; C_q(a)\sum_{m_1,m_2\le N} e\left((m_1-2m_2)\beta\right) \\
&= \sum_{m_1,m_2\le N} \left[\lambda(m_1)\lambda(m_2)e\left((m_1-2m_2)\frac{a}{q}\right) - C_q(a)\right] e\left((m_1-2m_2)\beta\right) \\
&= \sum_{m_1\le N}\left[\sum_{m_2\le N}\left[\lambda(m_1)\lambda(m_2)e\left((m_1-2m_2)\frac{a}{q}\right) - C_q(a)\right] e(-2m_2\beta)\right] e(m_1\beta) \\
&= \sum_{m_1\le N}\left[\sum_{m_2\le N} a_{m_2}(m_1,N)b_{m_2}(m_1,N)\right] e(m_1\beta) \\
&= \sum_{m_1\le N} S_{a,q}(\alpha;m_1)e(m_1\beta),
\end{aligned} \tag{2.22}$$

where

$$\begin{aligned} a_{m_2}(m_1,N) &= \lambda(m_1)\lambda(m_2)e\left((m_1-2m_2)\frac{a}{q}\right)-C_q(a) \\ b_{m_2}(m_1,N) &= e(-2m_2\beta) \\ S_{a,q}(\alpha;m_1) &= \sum_{m_2\le N} a_{m_2}(m_1,N)b_{m_2}(m_1,N). \end{aligned} \tag{2.23}$$

We have written $S_{a,q}(\alpha)$ as above to illuminate the application of partial summation. We hold m_1 fixed and then use partial summation on the m_2-sum. This generates two terms, a boundary and an integral term. We then apply partial summation to the m_1-sum. The difficulty is not in evaluating the sums, but rather in the necessary careful book-keeping required.

Recall the integral version of partial summation (Lemma **??**) states

$$\sum_{m=1}^{N} a_m b(m) \;=\; A(N)b(N)-\int_1^N A(u)b'(u)du, \tag{2.24}$$

where b is a differentiable function and $A(u)=\sum_{m\le u}a_m$. We apply this to $a_{m_2}(m_1,N)$ and $b_{m_2}(m_1,N)$. As $b_{m_2}=b(m_2)=e(-2\beta m_2)=e^{-4\pi i\beta m_2}$, $b'(m_2)=-4\pi i\beta e(-2\beta m_2)$.

Applying the integral version of partial summation to the m_2-sum gives

$$\begin{aligned} S_{a,q}(\alpha;m_1) &= \sum_{m_2\le N}\left[\lambda(m_1)\lambda(m_2)e\left((m_1-2m_2)\frac{a}{q}\right)-C_q(a)\right]e(-2m_2\beta) \\ &= \sum_{m_2\le N} a_{m_2}(m_1,N)b_{m_2}(m_1,N) \\ &= \left[\sum_{m_2\le N} a_{m_2}(m_1,N)\right]e(-2N\beta)+4\pi i\beta\int_{u=1}^{N}\left[\sum_{m_2\le u} a_{m_2}(m_1,N)\right]e(-u\beta)du. \end{aligned} \tag{2.25}$$

The first term is called the boundary term, the second the integral term. We substitute these into (2.22) and find

$$\begin{aligned} S_{a,q}(\alpha) &= \sum_{m_1\le N}\left[\left[\sum_{m_2\le N} a_{m_2}(m_1,N)\right]e(-2N\beta)\right]e(m_1\beta) \\ &\quad+\sum_{m_1\le N}\left[4\pi i\beta\int_{u=1}^{N}\left[\sum_{m_2\le u} a_{m_2}(m_1,N)\right]e(-u\beta)du\right]e(m_1\beta) \\ &= S_{a,q}(\alpha;\text{Boundary})+S_{a,q}(\alpha;\text{Integral}). \end{aligned} \tag{2.26}$$

The proof of Theorem 2.4.1 is completed by showing $S_{a,q}(\alpha;\text{Boundary})$ and $S_{a,q}(\alpha;\text{Integral})$ are small. This is done in Lemmas 2.4.3 and 2.4.8 by straightforward partial summation.

Remark 2.4.2. *The factor of $4\pi i\beta$ in* (2.26) *is from differentiating $b(m_2)$. Remember $\alpha = \frac{a}{q} + \beta$ is in the Major arc $\mathcal{M}_{a,q} = \left[\frac{a}{q} - \frac{Q}{N}, \frac{a}{q} + \frac{Q}{N}\right]$. Thus, $|\beta| \le \frac{Q}{N} = \frac{\log^D N}{N}$. Even though the integral in* (2.26) *is over a range of length N, it is multiplied by β, which is small. If β was not present, this term would yield a contribution greater than the expected main term.*

2.4.1 Boundary Term

We first deal with the boundary term from the first partial summation on m_2, $S_{a,q}(\alpha;\text{Boundary})$.

Lemma 2.4.3.

$$S_{a,q}(\alpha;\textit{Boundary}) \;=\; O\left(\frac{N^2}{\log^{C-D} N}\right). \tag{2.27}$$

Proof. Recall that

$$\begin{aligned} S_{a,q}(\alpha;\text{Boundary}) &= \sum_{m_1\le N}\left[\left[\sum_{m_2\le N} a_{m_2}(m_1,N)\right]e(-2N\beta)\right]e(m_1\beta) \\ &= e(-2N\beta)\sum_{m_1\le N}\left[\sum_{m_2\le N} a_{m_2}(m_1,N)\right]e(m_1\beta). \end{aligned} \tag{2.28}$$

As $|(e(-2N\beta)| = 1$, we can ignore it in the bounds below. We again apply the integral version of partial summation with

$$\begin{aligned} a_{m_1} &= \sum_{m_2\le N} a_{m_2}(m_1,N) \;=\; \sum_{m_2\le N}\left[\lambda(m_1)\lambda(m_2)e\left((m_1-2m_2)\frac{a}{q}\right) - C_q(a)\right] \\ b_{m_1} &= e(m_1\beta). \end{aligned} \tag{2.29}$$

We find

$$\begin{aligned} e(2N\beta)S_{a,q}(\alpha;\text{Boundary}) &= \sum_{m_1\le N}\left[\sum_{m_2\le N} a_{m_2}(m_1,N)\right]e(N\beta) \\ &\quad - 2\pi i\beta\int_{t=0}^{N}\sum_{m_1\le t}\left[\sum_{m_2\le N} a_{m_2}(m_1,N)\right]e(t\beta)dt. \end{aligned} \tag{2.30}$$

To prove Lemma 2.4.3, it suffices to bound the two terms in (2.30), which we do in Lemmas 2.4.4 and 2.4.5. □

Lemma 2.4.4.

$$\sum_{m_1\le N}\left[\sum_{m_2\le N} a_{m_2}(m_1,N)\right]e(N\beta) \;=\; O\left(\frac{N^2}{\log^C N}\right) \tag{2.31}$$

.

Proof. As $|e(N\beta)|=1$, this factor is harmless, and the m_1,m_2-sums are bounded by the Siegel-Walfisz Theorem.

$$\begin{aligned}
\sum_{m_1\le N}\sum_{m_2\le N} a_{m_2}(m_1,N) &= \sum_{m_1\le N}\sum_{m_2\le N}\left[\lambda(m_1)\lambda(m_2)e\left((m_1-2m_2)\frac{a}{q}\right)-C_q(a)\right]\\
&= \left[\sum_{m_1\le N}\lambda(m_1)e\left(m_1\frac{a}{q}\right)\right]\left[\sum_{m_2\le N}\lambda(m_2)e\left(-m_2\frac{a}{q}\right)\right]-C_q(a)N^2\\
&= \left[\frac{c_q(a)N}{\phi(q)}+O\left(\frac{N}{\log^C N}\right)\right]\cdot\left[\frac{c_q(-2a)N}{\phi(q)}+O\left(\frac{N}{\log^C N}\right)\right]-C_q(a)N^2\\
&= O\left(\frac{N^2}{\log^C N}\right) \qquad (2.32)
\end{aligned}$$

as $C_q(a)=\frac{c_q(a)c_q(-2a)}{\phi(q)^2}$ and $|c_q(b)|\le\phi(q)$. □

Lemma 2.4.5.

$$2\pi i\beta\int_{t=0}^{N}\sum_{m_1\le t}\left[\sum_{m_2\le N} a_{m_2}(m_1,N)\right]e(t\beta)dt \;=\; O\left(\frac{N^2}{\log^{C-D} N}\right). \tag{2.33}$$

Proof. Note $|\beta|\le\frac{Q}{N}=\frac{\log^D N}{N}$, and $C_q(a)=\frac{c_q(a)}{\phi(q)^2}\frac{c_q(-2a)}{\phi(q)^2}$. For $t\le\sqrt{N}$, we trivially bound the m_2-sum by $2N$. Thus these t contribute at most

$$|\beta|\int_{t=0}^{\sqrt{N}}\sum_{m_1\le t}2Ndt \;=\; |\beta|N^2 \;\le\; N\log^D N. \tag{2.34}$$

An identical application of Siegel-Walfisz as in the proof of Lemma 2.4.4 yields for $t\ge\sqrt{N}$,

$$\begin{aligned}
\sum_{m_1\le t}\sum_{m_2\le N} a_{m_2}(m_1,N) &= \left[\frac{c_q(a)t}{\phi(q)}+O\left(\frac{t}{\log^C N}\right)\right]\cdot\left[\frac{c_q(-2a)N}{\phi(q)}+O\left(\frac{N}{\log^C N}\right)\right]-C_q(a)tN\\
&= O\left(\frac{tN}{\log^C N}\right). \qquad (2.35)
\end{aligned}$$

Therefore

$$|\beta| \int_{t=\sqrt{N}}^{N} \left| \sum_{m_1 \leq t} \sum_{m_2 \leq N} a_{m_2}(m_1, N) \right| dt \ = \ O\left(\frac{N^3 \beta}{\log^C N} \right) \ = \ O\left(\frac{N^2}{\log^{C-D} N} \right) \tag{2.36}$$

□

Remark 2.4.6. *Note, of course, that the contribution is only negligible while $|\beta| \leq \frac{Q}{N}$. We see a natural reason to take the Major arcs small in length.*

Remark 2.4.7. *The above argument illustrate a very common technique. Namely, if $t \in [0, N]$ and N is large, the interval $[0, \sqrt{N}]$ has negligible relative length. It is often useful to break the problem into two such regions, as different bounds are often available when t is large and small. For our problem, the Siegel-Walfisz formula requires that $q \leq \log^B t$; this condition fails if t is small compared to q. For small t, the bounds may not be as good; however, the length of such an interval is so small that weak bounds suffice. See also the example in §**??**.*

2.4.2 Integral Term

We now deal with the integral term from the first partial summation on m_2, $S_{a,q}(\alpha; \text{Integral})$.

Lemma 2.4.8.

$$S_{a,q}(\alpha; \textit{Integral}) \ = \ O\left(\frac{N^2}{\log^{C-2D} N} \right). \tag{2.37}$$

Proof. Recall

$$S_{a,q}(\alpha; \text{Integral}) \ = \ 4\pi i \beta \sum_{m_1 \leq N} \left[\int_{u=1}^{N} \left[\sum_{m_2 \leq u} a_{m_2}(m_1, N) \right] e(-u\beta) du \right] e(m_1 \beta) \tag{2.38}$$

where

$$a_{m_2}(m_1, N) \ = \ \lambda(m_1)\lambda(m_2) e\left((m_1 - 2m_2)\frac{a}{q} \right) - C_q(a). \tag{2.39}$$

We apply the integral version of partial summation, with

$$\begin{aligned} a_{m_1} &= \int_{u=1}^{N} \left[\sum_{m_2 \leq u} a_{m_2}(m_1, N) \right] e(-u\beta) du \\ b_{m_1} &= e(m_1 \beta). \end{aligned} \tag{2.40}$$

We find

$$\begin{aligned} S_{a,q}(\alpha;\text{Integral}) &= 4\pi i\beta\left[\sum_{m_1\le N}\int_{u=1}^{N}\sum_{m_2\le u}a_{m_2}(m_1,N)e(-u\beta)du\right]e(N\beta) \\ &\quad + 8\pi\beta^2\int_{t=1}^{N}\left[\sum_{m_1\le t}\int_{u=1}^{N}\sum_{m_2\le u}a_{m_2}(m_1,N)e(-u\beta)du\right]e(m_1t)dt. \quad (2.41)\end{aligned}$$

The factor of $8\pi\beta^2 = -(4\pi i\beta)\cdot(2\pi i\beta)$ and comes from the derivative of $e(m_1\beta)$. Arguing in a similar manner as in §2.4.1, in Lemmas 2.4.9 and 2.4.10 we show the two terms in (2.41) are small, which will complete the proof. □

Lemma 2.4.9.

$$4\pi i\beta\left[\sum_{m_1\le N}\int_{u=1}^{N}\sum_{m_2\le u}a_{m_2}(m_1,N)e(-u\beta)du\right]e(N\beta) \;=\; O\left(\frac{N^2}{\log^{C-D}N}\right). \quad (2.42)$$

Proof. Arguing along the lines of Lemma 2.4.5, one shows the contribution from $u\le\sqrt{N}$ is bounded by $N\log^D N$. For $u\ge\sqrt{N}$ we apply the Siegel-Walfisz formula as in Lemma 2.4.5, giving a contribution bounded by

$$\begin{aligned} & 4|\beta|\int_{u=\sqrt{N}}^{N}\left(\left[\frac{c_q(a)u}{\phi(q)}+O\left(\frac{u}{\log^C N}\right)\right]\cdot\left[\frac{c_q(-2a)N}{\phi(q)}+O\left(\frac{N}{\log^C N}\right)\right]-C_q(a)uN\right)du \\ \ll\;& |\beta|\int_{u=\sqrt{N}}^{N}\frac{uN}{\log^C N}du \\ \ll\;& \frac{N^3|\beta|}{\log^C N}. \quad (2.43)\end{aligned}$$

As $|\beta|\le\frac{\log^B N}{N}$, the above is $O\left(\frac{N^2}{\log^{C-D}N}\right)$. □

Lemma 2.4.10.

$$8\pi\beta^2\int_{t=1}^{N}\left[\sum_{m_1\le t}\int_{u=1}^{N}\sum_{m_2\le u}a_{m_2}(m_1,N)e(-u\beta)du\right]e(m_1t)dt \;=\; O\left(\frac{N^2}{\log^{C-2D}N}\right). \quad (2.44)$$

Proof. The proof proceeds along the same lines as the previous lemmas. Arguing as in Lemma 2.4.5, one shows that the contribution when $t\le\sqrt{N}$ or $u\le\sqrt{N}$ is $O\left(\frac{N}{\log^{C-2D}N}\right)$. We then apply the Siegel-Walfisz Theorem as before, and find the contribution when $t,u\ge\sqrt{N}$ is

$$\ll\; 8\beta^2\int_{t=\sqrt{N}}^{N}\int_{u=\sqrt{N}}^{N}\frac{ut}{\log^C N}dudt \;\ll\; \frac{N^4\beta^2}{\log^C N}. \quad (2.45)$$

As $|\beta| \leq \frac{\log^D N}{N}$, the above is $O\left(\frac{N^2}{\log^{C-2D} N}\right)$. □

This completes the proof of Theorem 2.4.1.

2.5 Integrals over the Major Arcs

We first compute the integral of $u(x)e(-x)$ over the Major arcs and then use Theorem 2.4.1 to deduce the corresponding integral of $F_N(x)e(-x)$.

2.5.1 Integrals of $u(x)$

By Theorem 2.4.1 we know for $x \in \mathcal{M}_{a,q}$ that

$$\left|F_N(x) - C_q(a)u\left(x - \frac{a}{q}\right)\right| \ll O\left(\frac{N^2}{\log^{C-2D} N}\right). \tag{2.46}$$

We now evaluate the integral of $u(x - \frac{a}{q})e(-x)$ over $\mathcal{M}_{a,q}$; by Theorem 2.4.1 we then obtain the integral of $F_N(x)e(-x)$ over $\mathcal{M}_{a,q}$. Remember (see (2.17)) that

$$u(x) = \sum_{m_1,m_2 \leq N} e\left((m_1 - 2m_2)x\right). \tag{2.47}$$

Theorem 2.5.1.

$$\int_{\mathcal{M}_{a,q}} u\left(\alpha - \frac{a}{q}\right) \cdot e(-\alpha)d\alpha = e\left(-\frac{a}{q}\right)\frac{N}{2} + O\left(\frac{N}{\log^D N}\right). \tag{2.48}$$

Theorem 2.5.1 will follow from a string of lemmas on various integrals of u. We first determine the integral of u over all of $[-\frac{1}{2}, \frac{1}{2}]$, and then show that the integral of $u(x)$ is small if $|x| > \frac{Q}{N}$.

Lemma 2.5.2.

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} u(x)e(-x)dx = \frac{N}{2} + O(1). \tag{2.49}$$

Proof.

$$\begin{aligned}\int_{-\frac{1}{2}}^{\frac{1}{2}} u(x)e(-x)dx &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{m_1 \leq N}\sum_{m_2 \leq N} e\left((m_1 - 2m_2)x\right) \cdot e(-x)dx \\ &= \sum_{m_1 \leq N}\sum_{m_2 \leq N} \int_{-\frac{1}{2}}^{\frac{1}{2}} e\left((m_1 - 2m_2 - 1)x\right)dx. \end{aligned} \tag{2.50}$$

By exercise 1.2.7 the integral is 1 if $m_1 - 2m_2 - 1 = 0$ and 0 otherwise. For $m_1, m_2 \in \{1, \dots, N\}$, there are $\left[\frac{N}{2}\right] = \frac{N}{2} + O(1)$ solutions to $m_1 - 2m_2 - 1 = 0$, which completes the proof. □

Define

$$I_1 \;=\; \left[-\frac{1}{2} + \frac{Q}{N}, -\frac{Q}{N}\right], \quad I_2 \;=\; \left[\frac{Q}{N}, \frac{1}{2} - \frac{Q}{N}\right]. \tag{2.51}$$

The following bound is crucial in our investigations.

Lemma 2.5.3. *For $x \in I_1$ or I_2, $\frac{1}{1-e(ax)} \ll \frac{1}{x}$ for $a \in \{1, -2\}$.*

Exercise 2.5.4. *Prove Lemma 2.5.3.*

Lemma 2.5.5.

$$\int_{x \in I_1 \cup I_2} u(x)e(-x)dx \;=\; O\left(\frac{N}{\log^D N}\right). \tag{2.52}$$

Proof. We have

$$\begin{aligned}
\int_{I_i} u(x)e(-x)dx &= \int_{I_i} \sum_{m_1, m_2 \le N} e\left((m_1 - 2m_2 - 1)x\right) dx \\
&= \int_{I_i} \sum_{m_1 \le N} e(m_1 x) \sum_{m_2 \le N} e(-2m_2 x) \cdot e(-x)dx \\
&= \int_{I_i} \left[\frac{e(x) - e((N+1)x)}{1 - e(x)}\right] \left[\frac{e(-2x) - e(-2(N+1)x)}{1 - e(-2x)}\right] e(-x)dx
\end{aligned} \tag{2.53}$$

because these are geometric series. By Lemma 2.5.3 we have

$$\int_{I_i} u(x)e(-x)dx \;\ll\; \int_{I_i} \frac{2}{x}\frac{2}{x}dx \;\ll\; \frac{N}{Q} \;=\; \frac{N}{\log^D N}, \tag{2.54}$$

which completes the proof of Lemma 2.5.5. □

Remark 2.5.6. *It is because the error term in Lemma 2.5.5 is $O\left(\frac{N}{\log^D N}\right)$ that we must take $D > 2B$.*

Lemma 2.5.7.

$$\int_{x=\frac{1}{2}-\frac{Q}{N}}^{\frac{1}{2}+\frac{Q}{N}} u(x)e(-x)dx \;=\; O\left(\log^D N\right). \tag{2.55}$$

Proof. The argument is similar to the proof of Lemma 2.5.5. The difference is we use the geometric series formula only for the m_1-sum, which is $\frac{e(x)-e((N+1)x)}{1-e(x)} \ll \frac{1}{x}$. As x is near $\frac{1}{2}$, the m_1-sum is $O(1)$. There are N terms in the m_2-sum. As each term is at most 1, we may bound the m_2-sum by N. Thus, the integrand is $O(N)$. We integrate over a region of length $\frac{2Q}{N}$ and see that the integral is $O(Q) = O(\log^D N)$ for N large. □

Note in the above proof we could not use the geometric series for both m-sums, as near $x = \frac{1}{2}$ the second sum is quite large. Fortunately, we still have significant cancellation in the first sum, and we are integrating over a small region. The situation is different in the following lemma. There, *both* m-sums are large. Not surprisingly, this is where most of the mass of u is concentrated.

Lemma 2.5.8.

$$\int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(x)e(-x)dx = \frac{N}{2} + O\left(\frac{N}{\log^D N}\right). \tag{2.56}$$

Proof. This is immediate from Lemma 2.5.2 (which shows the integral of u over $[-\frac{1}{2}, \frac{1}{2}]$ is $\frac{N}{2} + O(1)$) and Lemmas 2.5.5 and 2.5.7 (which show the integral of u over $|x| > \frac{Q}{N}$ is small). □

It is now trivial to prove Theorem 2.5.1.

Proof of Theorem 2.5.1. We have

$$\begin{aligned}
\int_{\mathcal{M}_{a,q}} u\left(\alpha - \frac{a}{q}\right) \cdot e(-\alpha)d\alpha &= \int_{\frac{a}{q}-\frac{Q}{N}}^{\frac{a}{q}+\frac{Q}{N}} u\left(\alpha - \frac{a}{q}\right) \cdot e(-\alpha)d\alpha \\
&= \int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta) \cdot e\left(-\frac{a}{q} - \beta\right) d\beta \\
&= e\left(-\frac{a}{q}\right) \int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta)e(-\beta)d\beta \\
&= e\left(-\frac{a}{q}\right) \frac{N}{2} + O\left(\frac{N}{\log^D N}\right).
\end{aligned} \tag{2.57}$$

□

Note there are two factors in Theorem 2.5.1. The first, $e\left(-\frac{a}{q}\right)$, is an arithmetical factor which depends on which Major arc $\mathcal{M}_{a,q}$ we are in. The second factor is universal, and is the size of the contribution.

Remark 2.5.9. *We remark once more on the utility of finding a function $u(x)$ to approximate $F_N(x)$, as opposed to a Taylor series expansion. We found a function that is easy to integrate and by straightforward applications of partial summation is close to our generating function. Further, most of the mass of $u(x)$ is concentrated in a neighborhood of size $\frac{2Q}{N}$ about 0. Hence integrating u (or its translates) over a Major arc is approximately the same as integrating u over the entire interval. While there are a few points where we need to be careful in analyzing the behavior of u, the slight complications are worth the effort because of how easy it is to work with $u(x)$. For this problem, it was $x = 0$ giving the main contribution, and $x = \pm\frac{1}{2}$ was a potential trouble point which turned out to give a small contribution. The reason we need to check $x = \pm\frac{1}{2}$ is due to the definition of Germain primes, namely the **2** in $F_{2N}(x) = \sum_{p_2 \le N} e(-\mathbf{2}p_2x)$. Because of this 2, when x is near $\frac{1}{2}$, $F_{2N}(x)$ is near N.*

2.5.2 Integrals of $F_N(x)$

An immediate consequence of Theorem 2.5.1 is

Theorem 2.5.10.

$$\int_{\mathcal{M}_{a,q}} F_N(x)e(-x)dx \;=\; C_q(a)e\left(-\frac{a}{q}\right)\frac{N}{2} + O\left(\frac{N}{\log^D N}\right) + O\left(\frac{N}{\log^{C-3D} N}\right) \tag{2.58}$$

Exercise 2.5.11. *Prove Theorem 2.5.10.*

From Theorem 2.5.10 we immediately obtain the integral of $F_N(x)e(-x)$ over the Major arcs $\mathcal{M}$:

Theorem 2.5.12.

$$\begin{aligned}\int_{\mathcal{M}} F_N(x)e(-x)dx &= \sum_{q=1}^{\log^B N}\sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a)e\left(-\frac{a}{q}\right)\frac{N}{2} \;+\; O\left(\frac{N}{\log^{D-2B} N} + \frac{N}{\log^{C-3D-2B} N}\right)\\ &= \mathfrak{S}_N\frac{N}{2} + O\left(\frac{N}{\log^{D-2B} N} + \frac{N}{\log^{C-3D-2B} N}\right),\end{aligned} \tag{2.59}$$

where

$$\mathfrak{S}_N \;=\; \sum_{q=1}^{\log^B N}\sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a)e\left(-\frac{a}{q}\right) \tag{2.60}$$

is the truncated singular series for the Germain primes.

Proof. As

$$\mathcal{M} \;=\; \bigcup_{q=1}^{\log^B N} \bigcup_{\substack{a=1\\(a,q)=1}}^{q} \mathcal{M}_{a,q}, \tag{2.61}$$

the number of Major arcs $\mathcal{M}_{a,q}$ is bounded by $\log^{2B} N$. In summing over the Major arcs, the error terms in Theorem 2.5.10 are multiplied by at most $\log^{2B} N$, and the claim now follows. □

We will show the main term in Theorem 2.5.12 is of size N; thus we need to take $D > 2B$ and $C > 3D + 2B$. We study $\mathfrak{S}_N$ in §2.6, and remove the $\log p_i$ weights in §2.7.

2.6 Major Arcs and the Singular Series

If we can show that there exists a constant $c_0 > 0$ (independent of N) such that

$$\mathfrak{S}_N \;>\; c_0, \tag{2.62}$$

then for $D > 2B$ and $C > 3D + 2B$ by Theorem 2.5.12 the contribution from the Major arcs is positive and of size $\mathfrak{S}_N \frac{N}{2}$ for N sufficiently large. Recall

$$\begin{aligned} c_q(a) &= \sum_{\substack{r=1\\(r,q)=1}}^{q} e\left(r\frac{a}{q}\right) \\ C_q(a) &= \frac{c_q(a)c_q(-2a)}{\phi(q)^2}. \end{aligned} \tag{2.63}$$

Substituting

$$\rho_q \;=\; \sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a) e\left(-\frac{a}{q}\right), \tag{2.64}$$

into the series expansion of $\mathfrak{S}_N$ in (2.60), we find that

$$\mathfrak{S}_N \;=\; \sum_{q=1}^{\log^B N} \rho_q. \tag{2.65}$$

The singular series for the Germain primes is

$$\mathfrak{S} \;=\; \sum_{q=1}^{\infty} \rho_q. \tag{2.66}$$

We show $\mathfrak{S}$ is given by a multiplicative product and is positive in Theorem 2.6.18, and in Theorem 2.6.20 we show $|\mathfrak{S} - \mathfrak{S}_N| = O\left(\frac{1}{\log^{(1-2\epsilon)B} N}\right)$ for any $\epsilon > 0$. This will complete our evaluation of the contribution from the Major arcs.

Many of the arithmetical functions we investigate below were studied in Chapter **??**. Recall a function f is multiplicative if $f(mn) = f(m)f(n)$ for m, n relatively prime, and completely multiplicative if $f(mn) = f(m)f(n)$; see definition **??**. The reader should consult Chapter **??** as necessary.

2.6.1 Properties of Arithmetic Functions

We follow the presentation of [Na] (Chapter 8 and Appendix A), where many of the same functions arise from studying a related Circle Method problem. Below we determine simple formulas for the arithmetic functions we have encountered, which then allows us to prove our claims about $\mathfrak{S}_N$ and $\mathfrak{S}$ (see §2.6.2).

Lemma 2.6.1. *If $(q, q') = 1$ then we can write the congruence classes relatively prime to qq' as $rq' + r'q$, with $1 \le r \le q$, $1 \le r' \le q'$ and $(r, q) = (r', q') = 1$.*

Exercise 2.6.2. *Prove Lemma 2.6.1.*

Lemma 2.6.3. *$c_q(a)$ is multiplicative.*

Proof. Using Lemma 2.6.1 we have

$$\begin{aligned} c_q(a)c_{q'}(a) &= \sum_{\substack{r=1\\(r,q)=1}}^{q} e\left(r\frac{a}{q}\right) \sum_{\substack{r'=1\\(r',q')=1}}^{q'} e\left(r'\frac{a}{q'}\right) \\ &= \sum_{\substack{r=1\\(r,q)=1}}^{q} \sum_{\substack{r'=1\\(r',q')=1}}^{q'} e\left(\frac{(rq'+r'q)a}{qq'}\right) \\ &= \sum_{\substack{\widetilde{r}=1\\(\widetilde{r},qq')=1}}^{qq'} e\left(\widetilde{r}\frac{a}{q}\right) = c_{qq'}(a). \end{aligned} \tag{2.67}$$

□

We will soon determine $c_q(a)$ for $(a, q) = 1$. We first state some needed results.

Lemma 2.6.4. *Show that*

$$h_d(a) = \sum_{r=1}^{d} e\left(r\frac{a}{d}\right) = \begin{cases} d & \text{if } d|a; \\ 0 & \text{otherwise.} \end{cases} \tag{2.68}$$

Exercise 2.6.5. *Prove the above lemma.*

Recall the Möbius function (see §**??**):

$$\mu(d) \;=\; \begin{cases} (-1)^r & \text{if } d \text{ is the product of } r \text{ distinct primes;} \\ 0 & \text{otherwise.} \end{cases} \tag{2.69}$$

By Lemma **??**,

$$\sum_{d|(r,q)} \mu(d) \;=\; \begin{cases} 1 & \text{if } (r,q)=1; \\ 0 & \text{otherwise.} \end{cases} \tag{2.70}$$

Lemma 2.6.6. *If* $(a,q)=1$ *then* $c_q(a)=\mu(q)$.

Proof.

$$c_q(a) \quad = \quad \sum_{\substack{r=1 \\ (r,q)=1}}^{q} e\left(r\frac{a}{q}\right) \;=\; \sum_{r=1}^{q} e\left(r\frac{a}{q}\right) \sum_{d|(r,q)} \mu(d), \tag{2.71}$$

where we used (2.69) to expand the sum from $(r,q)=1$ to all $r \bmod q$. Further

$$c_q(a) \quad = \quad \sum_{d|q} \mu(d) \sum_{\substack{r=1 \\ d|r}}^{q} e\left(r\frac{a}{q}\right). \tag{2.72}$$

This is because $d|(r,q)$ implies $d|r$ and $d|q$, which allows us to rewrite the conditions above. We change variables and replace r with ℓ; as r ranges from 1 to q through values divisible by ℓ, ℓ ranges from 1 to $\frac{q}{d}$. We will use Lemma 2.6.4 to evaluate this sum. Therefore

$$\begin{aligned} c_q(a) \quad &= \quad \sum_{d|q} \mu(d) \sum_{\ell=1}^{q/d} e\left(\ell\frac{a}{q/d}\right) \\ &= \quad \sum_{d|q} \mu(d) h_{q/d}(a) \\ &= \quad \sum_{d|q} \mu\left(\frac{q}{d}\right) h_d(a) \\ &= \quad \sum_{\substack{d|q \\ d|a}} \mu\left(\frac{q}{d}\right) d \\ &= \quad \sum_{d|(a,q)} \mu\left(\frac{q}{d}\right) d. \end{aligned} \tag{2.73}$$

If $(a,q)=1$ then the only term above is $d=1$, which yields $c_q(a)=\mu(q)$. □

Corollary 2.6.7. *If $q = p^k$, $k \geq 2$ and $(a, q) = 1$, then $c_q(a) = 0$.*

We have shown $c_{qq'}(a) = c_q(a)c_{q'}(a)$ if $(q, q') = 1$. Recall that the Euler ϕ-function, $\phi(q)$, is the number of numbers less than q which are relatively prime to q and is a multiplicative function (see §**??** for more details). We now have

Lemma 2.6.8. *$C_q(a)$ is multiplicative in q.*

Proof. Assume $(q, q') = 1$. We have

$$\begin{aligned} C_{qq'}(a) &= \frac{c_{qq'}(a)c_{qq'}(-2a)}{\phi(qq')^2} \\ &= \frac{c_q(a)c_{q'}(a)c_q(-2a)c_{q'}(-2a)}{\phi(q)^2\phi(q')^2} \\ &= \frac{c_q(a)c_q(-2a)}{\phi(q)^2} \cdot \frac{c_{q'}(a)c_{q'}(-2a)}{\phi(q')^2} \\ &= C_q(a)C_{q'}(a). \end{aligned} \tag{2.74}$$

□

We now prove ρ_q is multiplicative. We first prove a needed lemma.

Lemma 2.6.9. *If $(q_1, q_2) = 1$, $C_{q_1}(a_1q_2) = C_{q_1}(a_1)$.*

Proof. As $C_{q_1}(a_1q_2) = \frac{c_{q_1}(a_1q_2)c_{q_1}(-2a_1q_2)}{\phi(q_1)}$, we see it suffices to show $c_{q_1}(a_1q_2) = c_{q_1}(a_1)$ and $c_{q_1}(-2a_1q_2) = c_{q_1}(-2a_1)$. As the proofs are similar, we only prove the first statement. From the definition of $c_q(a)$, (2.63), we have

$$\begin{aligned} c_{q_1}(a_1q_2) &= \sum_{\substack{r_1=1 \\ (r_1,q_1)=1}}^{q_1} e\left(r_1\frac{a_1q_2}{q_1}\right) \\ &= \sum_{\substack{r_1=1 \\ (r_1,q_1)=1}}^{q_1} e\left(r_1q_2\frac{a_1}{q_1}\right) \\ &= \sum_{\substack{r=1 \\ (r,q_1)=1}}^{q_1} e\left(r\frac{a_1}{q_1}\right) = c_{q_1}(a), \end{aligned} \tag{2.75}$$

because $(q_1, q_2) = 1$ implies that as r_1 goes through all residue classes that are relatively prime to q_1, so too does $r = r_1q_2$. □

Lemma 2.6.10. *ρ_q is multiplicative.*

Proof. Recall

$$\rho_q \;=\; \sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a) e\left(-\frac{a}{q}\right). \tag{2.76}$$

Assume $(q_1, q_2) = 1$. By Lemma 2.6.1 we can write the congruence classes relatively prime to $q_1 q_2$ as $a_1 q_2 + a_2 q_1$, with $1 \le a_1 \le q_1$, $1 \le a_2 \le q_2$ and $(a_1, q_1) = (a_2, q_2) = 1$. Then

$$\begin{aligned}
\rho_{q_1 q_2} &= \sum_{\substack{a=1\\(a,q_1q_2)=1}}^{q_1q_2} C_{q_1q_2}(a) e\left(-\frac{a}{q_1q_2}\right)\\
&= \sum_{\substack{a=1\\(a,q_1q_2)=1}}^{q_1q_2} C_{q_1}(a) C_{q_2}(a) e\left(-\frac{a}{q_1q_2}\right)\\
&= \sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1} \sum_{\substack{a_2=1\\(a_2,q_2)=1}}^{q_2} C_{q_1}(a_1q_2 + a_2q_1) C_{q_2}(a_1q_2 + a_2q_1) e\left(-\frac{a_1q_2 + a_2q_1}{q_1q_2}\right).
\end{aligned} \tag{2.77}$$

A straightforward calculation shows $C_{q_1}(a_1q_2 + a_2q_1) = C_{q_1}(a_1q_2)$ and $C_{q_2}(a_1q_2 + a_2q_1) = C_{q_2}(a_2q_1)$, which implies

$$\begin{aligned}
\rho_{q_1q_2} &= \sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1} \sum_{\substack{a_2=1\\(a_2,q_2)=1}}^{q_2} C_{q_1}(a_1q_2) C_{q_2}(a_2q_1) e\left(-\frac{a_1q_2 + a_2q_1}{q_1q_2}\right)\\
&= \left[\sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1} C_{q_1}(a_1q_2) e\left(-\frac{a_1}{q_1}\right)\right]\left[\sum_{\substack{a_2=1\\(a_2,q_2)=1}}^{q_2} C_{q_2}(a_2q_1) e\left(-\frac{a_2}{q_2}\right)\right]\\
&= \left[\sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1} C_{q_1}(a_1) e\left(-\frac{a_1}{q_1}\right)\right]\left[\sum_{\substack{a_2=1\\(a_2,q_2)=1}}^{q_2} C_{q_2}(a_2) e\left(-\frac{a_2}{q_2}\right)\right]\\
&= \rho_{q_1} \cdot \rho_{q_2},
\end{aligned} \tag{2.78}$$

where we used Lemma 2.6.9 to replace $C_{q_1}(a_1q_2)$ with $C_{q_1}(a_1)$, and similarly for $C_{q_2}(a_2q_1)$. Thus, ρ_q is multiplicative. □

Exercise 2.6.11. *Prove $C_{q_1}(a_1q_2 + a_2q_1) = C_{q_1}(a_1q_2)$ and $C_{q_2}(a_1q_2 + a_2q_1) = C_{q_2}(a_2q_1)$.*

We now determine ρ_q.

Lemma 2.6.12. *If $k \geq 2$ and p is prime then $\rho_{p^k} = 0$.*

Proof. This follows immediately from $C_{p^k}(a) = 0$ (see Corollary 2.6.7 and the definition of $C_q(a)$). □

Lemma 2.6.13. *If $p > 2$ is prime then $\rho_p = -\frac{1}{(p-1)^2}$.*

Proof.

$$\begin{aligned} \rho_p &= \sum_{\substack{a=1 \\ (a,p)=1}}^{p} C_p(a) e\left(-\frac{a}{p}\right) \\ &= \sum_{a=1}^{p-1} \frac{c_p(a)c_p(-2a)}{\phi(p)^2} e\left(-\frac{a}{p}\right). \end{aligned} \tag{2.79}$$

For $p > 2$, $(a, p) = 1$ implies $(-2a, p) = 1$ as well. By Lemma (2.6.6), $c_p(a) = c_p(-2a) = \mu(p)$. As $\mu(p)^2 = 1$ and $\phi(p) = p - 1$ we have

$$\begin{aligned} \rho_p &= \sum_{a=1}^{p-1} \frac{1}{(p-1)^2} e\left(-\frac{a}{p}\right) \\ &= \frac{1}{(p-1)^2}\left[-e\left(-\frac{0}{p}\right) + \sum_{a=0}^{p-1} e\left(-\frac{a}{p}\right)\right] \\ &= -\frac{1}{(p-1)^2}. \end{aligned} \tag{2.80}$$

□

Lemma 2.6.14. $\rho_2 = 1$.

Proof.

$$\begin{aligned} \rho_2 &= \sum_{\substack{a=1 \\ (a,2)=1}}^{2} C_2(a) e\left(-\frac{a}{2}\right) \\ &= C_2(1) e\left(-\frac{1}{2}\right) \\ &= \frac{c_2(1)c_2(-2)}{\phi(2)^2} \cdot e^{-\pi i} \\ &= \frac{e^{\pi i} e^{-2\pi i}}{1^2} \cdot e^{-\pi i} \;=\; 1, \end{aligned} \tag{2.81}$$

where we have used $c_2(1) = e^{\pi i}$ and $c_2(-2) = e^{-2\pi i}$. □

Exercise 2.6.15. *Prove $c_2(1) = e^{\pi i}$ and $c_2(-2) = e^{-2\pi i}$.*

2.6.2 Determination of $\mathfrak{S}_N$ and $\mathfrak{S}$

We use the results from §2.6.1 to study $\mathfrak{S}_N$ and $\mathfrak{S}$, which from (2.65) and (2.66) are

$$\mathfrak{S}_N \;=\; \sum_{q=1}^{\log^B N} \rho_q, \quad \mathfrak{S} \;=\; \sum_{q=1}^{\infty} \rho_q. \tag{2.82}$$

We show that $|\mathfrak{S} - \mathfrak{S}(N)|$ is small by first determining $\mathfrak{S}$ (Theorem 2.6.18) and then estimating the difference (Theorem 2.6.20).

Exercise 2.6.16. *Let h_q be any multiplicative sequence (with whatever growth conditions are necessary to ensure the convergence of all sums below). Prove*

$$\sum_{q=1}^{\infty} h_q \;=\; \prod_{p \text{ prime}} \left(1 + \sum_{k=1}^{\infty} h_{p^k}\right). \tag{2.83}$$

Determine what growth conditions ensure convergence.

Definition 2.6.17 (Twin Prime Constant).

$$T_2 \;=\; \prod_{p>2} \left[1 - \frac{1}{(p-1)^2}\right] \;\approx\; .6601618158 \tag{2.84}$$

is the twin prime constant. Using the Circle Method, Hardy and Littlewood were led to the conjecture that the number of twin primes at most x is given by

$$\pi_2(x) \;=\; 2T_2 \frac{x}{\log^2 x} + o\left(\frac{x}{\log^2 x}\right). \tag{2.85}$$

The techniques of this chapter suffice to determine the contribution from the Major arcs to this problem as well; however, again the needed bounds on the Minor arcs are unknown.

Theorem 2.6.18. *$\mathfrak{S}$ has a product representation and satisfies*

$$\mathfrak{S} \;=\; 2T_2. \tag{2.86}$$

Proof. By exercise 2.6.16 we have

$$\begin{aligned} \mathfrak{S} &= \sum_{q=1}^{\infty} \rho_q \\ &= \prod_{p \text{ prime}} \left(1 + \sum_{k=1}^{\infty} \rho_{p^k}\right) \\ &= \prod_{p \text{ prime}} (1 + \rho_p) \end{aligned} \tag{2.87}$$

because $\rho_{p^k} = 0$ for $k \geq 2$ and p prime by Lemma 2.6.12. The product is easily shown to converge (see exercise 2.6.19). By Lemmas 2.6.13 and 2.6.14, $\rho_2 = 1$ and $\rho_p = -\frac{1}{(p-1)}$ for $p > 2$ prime. Therefore

$$\begin{aligned} \mathfrak{S} &= \prod_{p} (1 + \rho_p) \\ &= (1 + \rho_2) \prod_{p>2} (1 + \rho_p) \\ &= 2 \prod_{p>2} \left[1 - \frac{1}{(p-1)^2}\right] \\ &= 2T_2. \end{aligned} \tag{2.88}$$

□

We need to estimate $|\mathfrak{S} - \mathfrak{S}_N|$. As ρ_q is multiplicative and zero if $q = p^k$ ($k \geq 2$), we need only look at sums of ρ_p. As $\rho_p = -\frac{1}{(p-1)^2}$, it follows that the difference between $\mathfrak{S}$ and $\mathfrak{S}_N$ tends to zero as $N \to \infty$.

Exercise 2.6.19. *Show the product in* (2.87) *converges.* Hint: *take the logarithm, and Taylor expand.*

Theorem 2.6.20. *For any $\epsilon > 0$ and B, N such that $\log^B N > 2$,*

$$|\mathfrak{S} - \mathfrak{S}_N| \ll O\left(\frac{1}{\log^{(1-2\epsilon)B} N}\right). \tag{2.89}$$

Proof.

$$\mathfrak{S} - \mathfrak{S}_N = \sum_{q=\log^B N}^{\infty} \rho_q. \tag{2.90}$$

By Lemma 2.6.12, $\rho_{p^k} = 0$ for $k \geq 2$ and p prime. By Lemma 2.6.13, $\rho_p = -\frac{1}{(p-1)^2}$ if $p > 2$ is prime. By Lemma 2.6.10, ρ_q is multiplicative. Therefore for any $\epsilon > 0$

$$|\rho_q| \ \leq \ \frac{1}{\phi(q)^2} \ \ll \ \frac{1}{q^{2-2\epsilon}}; \tag{2.91}$$

if q is not square-free this is immediate, and for q square-free we note $\phi(p) = p - 1$ and $\phi(q) \gg q^{1-\epsilon}$. Hence

$$|\mathfrak{S} - \mathfrak{S}_N| \quad \ll \quad \sum_{q=\log^B N}^{\infty} \frac{1}{q^{2-2\epsilon}} \ = \ O\left(\frac{1}{\log^{(1-2\epsilon)B} N}\right). \tag{2.92}$$

□

Exercise 2.6.21. *For q square-free, prove that for any $\epsilon > 0$, $\phi(q) \gg q^{1-\epsilon}$.*

Combining the results above, we have finally determined the contribution from the Major arcs:

Theorem 2.6.22. *Let $D > 2B$, $C > 3D + 2B$, $\epsilon > 0$ and $\log^B N > 2$. Then*

$$\int_{\mathcal{M}} F_N(x)e(-x)dx \ = \ \mathfrak{S}\frac{N}{2} + O\left(\frac{N}{\log^{(1-2\epsilon)} BN} + \frac{N}{\log^{D-2B} N} + \frac{N}{\log^{C-3D-2B} N}\right), \tag{2.93}$$

where $\mathfrak{S}$ is twice the twin prime constant T_2.

In the binary and ternary Goldbach problems, to see if N could be written as the sum of two or three primes, we evaluated the Singular Series at N (see §**??**). Thus, even after taking limits, we still evaluated the Singular Series at multiple points, as we were trying to see *which* integers can be written as a sum of two or three primes, and the answer told us how many ways this was possible. Here, we really have $\mathfrak{S}(1)$; knowing how large this is tells us information about what percent of primes are Germain primes (see §2.7). As things stand, it does not make sense to evaluate this Singular Series at additional points. However, if we were interested in a more general problem, such as p, $\frac{p-b}{2}$ are both prime, b odd, this would lead to $p_1 - 2p_2 = b$. We would replace $e(-x)$ in (2.6) with $e(-bx)$. Working in such generality would lead to a Singular Series depending on b. More generally, we could consider prime pairs of the form $p, \frac{ap+b}{c}$. If we take $a = c$ and $b = 2ck$, we have the special case of prime pairs, and the Singular Series will depend on the factorization of $2k$ (see [HL3, HL4]).

Exercise 2.6.23. *Redo the calculations of this chapter for one of the problems described above or in §2.1.*

2.7 Number of Germain Primes and Weighted Sums

We now remove the $\log p_i$ weights in our counting function. By Theorem 2.6.22, we know the contribution from the Major arcs. If we assume the minor arcs contribute $o(N)$ then we would have

$$\sum_{\substack{p\le N\\ p,\frac{p-1}{2}\text{ prime}}} \log p\cdot\log\frac{p-1}{2} \;=\; \mathfrak{S}\frac{N}{2}+o(N) \;=\; T_2N+o(N). \tag{2.94}$$

We can pass from this weighted sum to a count of the number of Germain prime pairs $\left(\frac{p-1}{2},p\right)$ with $p\le N$. Again we follow [Na], Chapter 8; for more on weighted sums, see §**??**. Define

$$\begin{aligned}\pi_G(N) &= \sum_{\substack{p\le N\\ p,\frac{p-1}{2}\text{ prime}}} 1\\ G(N) &= \sum_{\substack{p\le N\\ p,\frac{p-1}{2}\text{ prime}}} \log p\cdot\log\frac{p-1}{2}.\end{aligned} \tag{2.95}$$

Theorem 2.7.1.

$$\frac{G(N)}{\log^2 N} \;\le\; \pi_G(N) \;\le\; \frac{G(N)}{\log^2 N}+O\left(N\frac{\log\log N}{\log N}\right). \tag{2.96}$$

Proof. In (2.95), $\log p\log\frac{p-1}{2}<\log^2 N$. Thus $G(N)\le\log^2 N\cdot\pi_G(N)$, proving the first inequality in (2.96).

The other inequality is more involved, and illustrates a common technique in analytic number theory. As there clearly are less Germain primes than primes, for any $\delta>0$

$$\pi_G(N^{1-\delta}) \;=\; \sum_{\substack{p\le N^{1-\delta}\\ p,\frac{p-1}{2}\text{ prime}}} 1 \;\le\; \pi(N^{1-\delta}) \;=\; \frac{N^{1-\delta}}{\log N^{1-\delta}} \;\ll\; \frac{N^{1-\delta}}{\log N}. \tag{2.97}$$

We now obtain a good upper bound for $\pi_G(N)$. If $p\ge N^{1-\delta}$, then

$$\begin{aligned}\log\frac{p-1}{2} &= \log p+\log\left(1-\frac{1}{2p}\right)\\ &\ge (1-\delta)\log N+O\left(\frac{1}{p}\right)\\ &= (1-\delta)\log N+O\left(\frac{1}{N^{1-\delta}}\right).\end{aligned} \tag{2.98}$$

In the arguments below, the error from (2.98) is negligible and is smaller than the other errors we encounter. We therefore suppress this error for convenience. Thus, up to lower order terms,

$$\begin{aligned} G(N) &\geq \sum_{\substack{p \geq N^{1-\delta} \\ p, \frac{p-1}{2} \text{ prime}}} \log p \cdot \log \frac{p-1}{2} \\ &= (1-\delta)^2 \log^2 N \sum_{\substack{p \geq N^{1-\delta} \\ p, \frac{p-1}{2} \text{ prime}}} 1 \\ &= (1-\delta)^2 \log^2 N \left(\pi_G(N) - \pi_G(N^{1-\delta})\right) \\ &\geq (1-\delta)^2 \log^2 N \cdot \pi_G(N) + O\left((1-\delta)^2 \log^2 N \cdot \frac{N^{1-\delta}}{\log N}\right). \end{aligned} \tag{2.99}$$

Therefore

$$\begin{aligned} \log^2 N \cdot \pi_G(N) &\leq (1-\delta)^{-2} \cdot G(N) + O\left(\log^2 N \cdot \frac{N^{1-\delta}}{\log N}\right) \\ 0 \leq \log^2 N \cdot \pi_G(N) - G(N) &\leq \left[(1-\delta)^{-2} - 1\right] G(N) + O\left(\log N \cdot N^{1-\delta}\right). \end{aligned} \tag{2.100}$$

If $0 < \delta < \frac{1}{2}$, then $(1-\delta)^{-2} - 1 \ll \delta$. We thus have

$$0 \leq \log^2 N \cdot \pi_G(N) - G(N) \ll N\left[\delta + O\left(\frac{\log N}{N^\delta}\right)\right]. \tag{2.101}$$

Choose $\delta = \frac{2 \log\log N}{\log N}$. Then we get

$$0 \leq \log^2 N \cdot \pi_G(N) - G(N) \leq O\left(N \frac{\log\log N}{\log N}\right), \tag{2.102}$$

which completes the proof. □

Combining our results of this section, we have proved

Theorem 2.7.2. *Assuming there is no contribution to the main term from the Minor arcs, up to lower order terms we have*

$$\pi_G(N) = \frac{T_2 N}{\log^2 N}, \tag{2.103}$$

where T_2 is the twin prime constant (see definition 2.6.17).

Remark 2.7.3 (Important). *It is a common technique in analytic number theory to choose an auxiliary parameter such as δ. Note how crucial it was in the proof for δ to depend (albeit very weakly) on N. Whenever one makes such approximations, it is good to get a feel for how much information is lost in the estimation. For $\delta = \frac{2\log\log N}{\log N}$ we have*

$$N^{1-\delta} \;=\; N \cdot N^{-2\log\log N/\log N} \;=\; N \cdot e^{-2\log\log N} \;=\; \frac{N}{\log^2 N}. \tag{2.104}$$

Hence there is little cost in ignoring the Germain primes less than $N^{1-\delta}$. Our final answer is of size $\frac{N}{\log^2 N}$. As $N^{1-\delta} = \frac{N}{\log^2 N}$ and there are $O\left(\frac{N}{\log^3 N}\right)$ primes less than $N^{1-\delta}$, there are at most $O\left(\frac{N}{\log^3 N}\right)$ Germain primes at most $N^{1-\delta}$.

Exercise 2.7.4. *Let $\Lambda(n)$ be the von Magnoldt function (see §??). Prove*

$$\sum_{n\le x}\Lambda(n) \;=\; \sum_{p\le x}\log p + O(x^{\frac{1}{2}}\log x). \tag{2.105}$$

As $\sum_{p\le x}$ is of size x, there is negligible loss in ignoring prime powers.

2.8 Exercises

The following problems are known questions (either on the Circle Method, or needed results to prove some of these claims).

Exercise 2.8.1. *Prove $\forall\epsilon > 0$, $q^{1-\epsilon} \ll \phi(q) \ll q$.*

Exercise 2.8.2. *Let*

$$c_q(N) \;=\; \sum_{\substack{a=1\\(a,q)=1}}^{q} e^{2\pi i N a/q}.$$

Prove $c_q(N)$ is multiplicative. Further, show

$$c_q(N) \;=\; \begin{cases} p-1 & \text{if } p|N \\ -1 & \text{otherwise.} \end{cases}$$

Exercise 2.8.3. *Prove $\mu(q)c_q(N)/\phi(q)^3$ is multiplicative.*

Exercise 2.8.4. *Using the above exercises and the methods of this chapter, calculate the contribution from the Major arcs to writing any integer N as the sum of three primes. Deduce for writing numbers as the sum of three primes that*

$$\begin{aligned}\mathfrak{S}_3(N) &= \sum_{q=1}^{\infty} \frac{\mu(q)c_q(N)}{\phi(q)^3} \\ &= \prod_p \left(1 + \sum_{j=1}^{\infty} \frac{\mu(p^j)c_{p^j}(N)}{\phi(p^j)^3}\right) \\ &= \prod_p \left(1 + \frac{1}{(p-1)^3}\right) \prod_{p|N} \left(1 - \frac{1}{p^2 - 3p + 3}\right).\end{aligned}$$

Note $\mathfrak{S}_3(N) = 0$ if N is even; thus the Circle Method "knows" that Goldbach is hard. We call $\mathfrak{S}(N)$ the ***Singular Series****.*

Exercise 2.8.5. *Let $\mathfrak{S}_{3,Q}(N)$ be the first Q terms of $\mathfrak{S}_3(N)$. Bound $\mathfrak{S}_3(N) - \mathfrak{S}_{3,Q}(N)$. Show for N odd there exist constants c_1, c_2 such that $0 < c_1 < \mathfrak{S}_3(N) < c_2 < \infty$.*

Exercise 2.8.6. *Assume every large integer is the sum of three primes. Prove every large even integer is the sum of two primes. Conversely, show if every large even integer is the sum of two primes, every large integer is the sum of three primes.*

Exercise 2.8.7 (Non-Trivial). *Calculate the Singular Series $\mathfrak{S}_2(N)$ and $\mathfrak{S}_{2,Q}(N)$ for the Goldbach problem (even numbers as the sum of two primes), and $\mathfrak{S}_{W,k,s}(N)$ and $\mathfrak{S}_{W,k,s,Q}(N)$ for Waring's problem (writing numbers as the sum of s perfect k-powers).* ***Warning:*** *$\mathfrak{S}_2(N) - \mathfrak{S}_{2,Q}(N)$* ***cannot be shown to be small for all even N in the Goldbach problem.*** *Do $\mathfrak{S}_{2,Q}(N)$ and $\mathfrak{S}_2(N)$ vanish for N odd?*

2.9 Research Projects

One can use the Circle Method to predict the number of primes (or prime tuples) with given properties, and then investigate these claims numerically; see, for example, [Law2, Sch, Weir] (for additional Circle Method investigations, see [Ci]). After counting the number of such primes (or prime tuples), the next natural question is to investigate the spacings between adjacent elements (see Chapter **??**).

Research Project 2.9.1. *For many questions in number theory, the Cramér model (see §**??** and Exercise 2.1.2) leads to good heuristics and predictions; recently, however, [MS] have shown that this model is inconsistent with certain simple numerical investigations of primes, and in fact the Random Matrix Theory model of the zeros of the Riemann Zeta function and the Circle Method give a prediction which agrees*

beautifully with experiments. There are many additional interesting sequences of primes to investigate and see which model is correct. Candidates include primes in arithmetic progression, twin primes, generalized twin primes (fix an integer k, look for primes such that p and p+2k are prime), prime tuples (fix integers k_1 through k_r such that $p, p+2k_1, \dots, p+2k_r$ are all prime), Germain primes, and so on. A natural project is to investigate the statistics from [MS] for these other sequences of primes, using the Circle Method and the Cramér model to predict two answers, and then see which agrees with numerics. ***ADD REF to papers from Brent***

Research Project 2.9.2. *In many successful applications of the Circle Method, good bounds are proved for the generating function on the Minor arcs. From these bounds it is then shown that the Minor arcs' contribution is significantly smaller than that from the Major arcs. However, to* prove *that the Major arcs are the main term does not require one to obtain good cancellation at every point in the Minor arcs; all that is required is that the* integral *is small.*

For problems such as Goldbach's conjecture or Germain primes, the needed estimates on the Minor arcs are conjectured to hold; by counting the number of solutions, we see that the integral over the Minor arcs is small (at least up to about 10^9). A good investigation is to numerically calculate the generating function at various points on the Minor arcs for several of these problems, and see how often large values are obtained. See [Law2] and ***REF TO CJM****. Warning: calculations of this nature are very difficult. The Major arcs are defined as intervals of size $\frac{2\log^D N}{N}$ about rationals with denominators at most $\log^B N$. For example, if $D = 10$ than $\log^D N > N$ until N is about 3.4×10^{15}, and there will not be any Minor arcs! For $N \approx 10^{15}$, there are too many primes to compute the generating function in a reasonable amount of time. Without resorting to supercomputers, one must assume that we may take B small for such numerical investigations.*

Bibliography

[Acz] A. Aczel, *Fermat's Last Theorem: Unlocking the Secret of an Ancient Mathematical Problem*, Four Walls Eight Windows, 1996.

[AKS] R. Adler, M. Keane, and M. Smorodinsky, *A construction of a normal number for the continued fraction transformation*, J. Number Theory **13** (1981), no. 1, 95–105.

[AgKaSa] M. Agrawal, N. Kayal and N. Saxena, *Primes is in P*, to appear.

[Al] Ahlfors, *Complex Analysis*, McGraw-Hill, Inc., New York, 3rd edition, 1979.

[AG] M. Aigner and G. M. Ziegler, *Proofs from THE BOOK*, Springer-Verlag Berlin Heidelberg, 1998, 199 pages, ISBN 3-540-63698-6.

[AGP] W. R. Alford, A. Granville, A. and C. Pomerance, *There are Infinitely Many Carmichael Numbers*, Ann. Math. 139, 703-722, 1994.

[AB] U. Andrews IV and J. Blatz, *Distribution of digits in the continued fraction representations of seventh degree algebraic irrationals*, Junior Thesis, Princeton University, Fall 2002,

[Ap] R. Apéry,*Irrationalité de $\zeta(3)$ et $\zeta(3)$*, Astérisque 61, 1979, 11-13.

[Apo] T. Apostol, *Introduction to Analytic Number Theory*.

[ALM] S. Arms, A. Lozano-Robledo and S. J. Miller, *Constructing One-Parameter Families of Elliptic Curves over $\mathbb{Q}(T)$ with Moderate Rank*, preprint.

[Art] M. Artin, *Algebra*, Prentice Hall.

[Ay] Ayoub, *Introduction to the analytic theory of numbers*

[Bai] Z. Bai, *Methodologies in spectral analysis of large-dimensional random matrices, a review*, Statist. Sinica **9** (1999), no. 3, 611-677.

[BR] K. Ball and T. Rivoal, *Irrationalité d'une infinité valeurs de la fonction zeta aux entiers impairs*, Invent. Math. 146, 2001, 193-207.

[BL] P. Baxandall and H. Liebeck, *Vector Calculus*, Clarendon Press, 1986.

[Be] R. Beals, *Notes on Fourier Series*, Lecture Notes, Yale University, 1994.

[Bec] M. Beceanu, *Period of the Continued Fraction of* $\sqrt{n}$, Junior Thesis, Princeton University, 2003.

[BBH] A. Berger, Leonid A. Bunimovich and T. Hill, *One-dimensional dynamical systems and Benford's Law*, accepted for publication in Transactions of the American Mathematical Society.

[BD] P. Bickel and K. Doksum, *Mathematical statistics: basic ideas and selected topics*, Holden-Day, 1977.

[Bl1] P. Bleher, *The energy level spacing for two harmonic oscillators with golden mean ratio of frequencies*, J. Stat. Phys. **61**, 1990, 869Ú876.

[Bl2] P. Bleher, *The energy level spacing for two harmonic oscillators with generic ratio of frequencies*, J. Stat. Phys. **63**, 1991, 261Ú283.

[Bol] B. Bollobás, *Random Graphs*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2001.

[BoLa] E. Bombieri and J. Lagarias, *Complements to Li's criterion for the Riemann hypothesis*, J. Number Theory **77**, no. 2, 1999, 274–287.

[Bon] D. Boneh, *Twenty years of attacks on the RSA cryptosystem*, Notices of the American Mathematical Society, **46**, 2, 1999, 203–213.

[BS] Z. Borevich and I. Shafarevich, *Number Theory*, Academic Press.

[BK] A. Boutet de Monvel and A. Khorunzhy, *Some Elementary Results around the Wigner Semicircle Law*, http://lthiwww.epfl.ch/∼leveque/Matrices/boutet_khorunzhy.pdf

[BFFMPW] T. Brody, J. Flores, J. French, P. Mello, A. Pandey, S. Wong, *Random-matrix physics: spectrum and strength fluctuations*, Rev. Mod. Phys. vol. **53**, no. 3, July 1981, $385 - 479$.

[BDJ] W. Bryc, A. Dembo, T. Jiang, *Spectral Measure of Large Randm Hankel, Markov and Toeplitz Matrices*, preprint.

[Br] A. Bryuno, *Continued frations of some algebraic numbers*, U.S.S.R. Comput. Math. and Math. Phys. **4** (1972), 1-15.

[CGI] G. Casati, I. Guarneri, and F. M. Izrailev, *Statistical Properties of the Quasi-Energy Spectrum of a Simple Integrable System*, Phys. Lett. A 124 (1987), $263 - 266$.

[Car] L. Carleson, *On the convergence and growth of partial sums of Fourier series*, Acta Math. **116**, 1966, 135Ű157.

[Ca] J. W. S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge University Press, London 1957.

[ChWa] J. R. Chen and T. Z. Wang, *On the Goldbach Problem*, Acta Math. Sinica 32, $702 - 718$, 1989.

[Ci] J. Cisneros, *Waring's Problem*, Junior Thesis, Princeton University, Spring 2001,

[CW] J. Coates and A. Wiles, *On the conjecture of Birch and Swinnterton-Dyer*, Invent. Math. **39**, 1977, $43 - 67$.

[CB] S. Chatterjee and A. Bose, *A New Method for Bounding Rates of Convergence of Empirical Spectral Distributions*,

[Coh] P. Cohen, *The Independence of the Continuum Hypothesis*, Proc. Nat. Acad. Sci. U. S. A, **50** 1963, 1143-1148 and **51** 1964, , 105-110.

[Con] J. B. Conrey, *L-functions and Random Matrices*,

[CFKRS] B. Conrey, D. Farmer, P. Keating, M. Rubinstein and N. Snaith, *Integral Moments of L-Functions*, http://arxiv.org/pdf/math.NT/0206018

[Dav] R. Daveney, *An Introduction to Chaotic Dynamical Systems*, Perseus Books, 2nd edition, 2003.

[Da1] H. Davenport, *The Higher Arithmetic*, Cambridge University Press.

[Da2] H. Davenport, *Multiplicative Number Theory, 2nd edition*, Graduate Texts in Mathematics **74**, Springer-Verlag, New York, 1980, revised by H. Montgomery.

[Da3] H. Davenport, *On the distribution of quadratic residues (mod p)*, Jour. London Math. Soc. **6** (1931), 49-54.

[Da4] H. Davenport, *On character sums in finite fields*, Acta Math. **71** (1939), 99-121.

[DSV] G. Davidoff, P. Sarnak and A. Valette, *Elementary Number Theory, Group Theory, and Ramanujan Graphs*, London Mathematical Society, Student Texts **55**, Cambridge University Press, 2003.

[Dia] P. Diaconis, *Patterns in eigenvalues: the* 70^{th} *Josiah Williard Gibbs lecture*, Bulletin of the American Mathematical Society, **40**, 2, 2003, 155-178.

[Di] T. Dimofte, *Rational Shifts of Linearly Periodic Continued Fractions*, Junior Thesis, Princeton University, 2003.

[Du] R. Durrett, *Probability: Theory and Examples*, Duxbury Press, second edition, 1996.

[Dy1] F. Dyson, *Statistical theory of the energy levels of complex systems: I, II, III*,. J. Mathematical Phys., **3**, 1962,140Ű156, 157-165, 166-175.

[Dy2] F. Dyson, *The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics*, J. Mathematical Phys., **3**, 1962, 1199Ű1215.

[Ed] H. M. Edwards, *Riemann's Zeta Function*, Academic Press, Inc., 1974.

[EE] W. J. Ellison and F. Ellison, *Prime Numbers*, John Wiley & Sons, New York, 1985.

[Est1] T. Estermann, *On Goldbach's Problem: Proof that Almost All Even Positive Integers are Sums of Two Primes*, Proc. London Math. Soc. Ser. 2 **44**, 1938, 307-314.

[Est2] T. Estermann, *Introduction to Modern Prime Number Theory*, Cambridge University Press, 1961.

[Fef] C. Fefferman, *Pointwise convergence of Fourier series*, Ann. of Math. (2) 98, 1973, 551–571.

[Fe] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. II. Second edition. John Wiley & Sons, Inc., New York-London-Sydney 1971.

[Fi] D. Fishman, *Closed Form Continued Fraction Expansions of Special Quadratic Irrationals*, Junior Thesis, Princeton University, 2003.

[Fol] G. Folland, *Real Analysis : Modern Techniques and Their Applications*, Pure and Applied Mathematics, Wiley-Interscience, second edition, 1999.

[FSV] P. J. Forrester, N. C. Snaith and J. J. M. Verbaarschot, *Developments in Random Matrix Theory*, **refer to special edition**

[Gau] M. Gaudin, *Sur la loi limite de lŠespacement des valeurs propres dŠune matrice aléatoire*, Nucl. Phys. **25**, 1961, 447-458.

[Gel] A. O. Gelfond, *Transcendental and Algebraic Numbers*, 1960.

[Gl] A. Gliga, *On continued fractions of the square root of prime numbers*, Junior Thesis, Princeton University, 2003.

[Gö] K. Gödel, *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*, Dover Publications, Inc.

[Gol1] D. GOldfeld, *The class number of quadratic fields and the conjectures of Birch and Swinnerton-Dyer*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. 3, 4, 1976, 624-663.

[Gol2] D. Goldfeld, *The Elementary proof of the Prime Number Theorem, An Historical Perspective*, Number Theory, New York Seminar 2003, Edited by D. and G. Chudnovsky, M. Nathanson, Springer, 2004, pp. 179-192.

[GKP] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A foundation for computer science*, Addison-Wesley Publishing Company, 1988.

[GT] A. Granville and T. Tucker, *It's as easy as* abc, Notices of the AMS, volume 49, number 10 (November 2002).

[GZ] B. Gross and D. Zagier, *Heegner points and derivatives of* L*-series*, Invent. Math. **84** 1986, no. 2, 225-320.

[HM] C. Hammond and S. J. Miller, *Eigenvalue spacing distribution for the ensemble of real symmetric Toeplitz matrices*, preprint.

[HL1] G. H. Hardy and J. E. Littlewood, *A new solution of Waring's Problem*, Q. J. Math. 48: $272 - 293$, 1919.

[HL2] G. H. Hardy and J. E. Littlewood, *Some problems of "Partitio Numerorum". A new solution of Waring's problem*, Göttingen Nach., $33 - 54$, 1920.

[HL3] 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.

[HL4] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum.' IV. Further researches in Waring's problem*, Math. Z., 23 1925, 1–37.

[HW] G. H. Hardy and E. Wright, *An Introduction to the Theory of Numbers*, fifth edition, Oxford Science Publications, Clarendon Press, Oxford, 1995.

[HR] G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatorial analysis*, Proc. London Math. Soc. 17: $75 - 115$, 1918.

[HuRu] C. Hughes and Z. Rudnick, *Mock Gaussian behaviour for linear statistics of classical compact groups*, J. Phys. A **36**, (2003) 2919–2932.

[Hei] H. Heillbronn, *On the average length of a class of finite continued fractions*, *Number Theory and Analysis* (A collection of papers in honor of E. Landau, VEB Deutscher Verlag, Berline 1968.

[Hej] D. Hejhal, *On the triple correlation of zeros of the zeta function*, Internat. Math. Res. Notices 1994, no. 7, $294 - 302$.

[Hil] D. Hilbert, *Beweis für die Darstellbarkeit der ganzen zahlen durch eine feste Anzahl n^{ter} Potenzen (Waringsches Problem)*, Mat. Annalen **67**, 281-300, 1909.

[Hi1] T. Hill, *The first-digit phenomenon*, American Scientists **86**, 1996, 358-363.

[Hi2] T. Hill, *A statistical derivation of the significant-digit law*, Statistical Science **10**, 1996, 354-363.

[HS] M. Hindry and J. Silverman, *Diophantine geometry: An introduction*, Graduate Texts in Mathematics, vol. 201, Springer, New York, 2000.

[HJ] K. Hrbacek and T. Jech, *Introduction to Set Theory*, Pure and Applied Mathematics, Marcel Dekker, Inc., 1984.

[Hua] Hua Loo Keng, *Introduction to Number Theory*, Springer-Verlag, 1982.

[IR] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, Springer-Verlag, Graduate Texts in Mathematics 84, 1990.

[Iw] H. Iwaniec, *Topics in Classical Automorphic Forms*, American Mathematical Society, Graduate Studies in Mathematics, vol. **17**, Providence, 1997.

[ILS] H. Iwaniec, W. Luo and P. Sarnak, *Low lying zeros of families of L-functions*, Inst. Hautes Études Sci. Publ. Math. **91**, 2000, $55 - 131$.

[JMRR] D. Jakobson, S. D. Miller, I. Rivin and Z. Rudnick, *Eigenvalue spacings for regular graphs*, Emerging applications of number theory (Minneapolis, MN, 1996), $317 - 327$.

[Je] R. Jeffrey, *Formal Logic: Its Scope and Limits*, McGraw Hill.

[Ka] S. Kapnick, *Continued fraction of cubed roots of primes*, Junior Thesis, Princeton University, Fall 2002,

[KS1] N. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues and Monodromy*, AMS Colloquium Publications **45**, AMS, Providence, 1999.

[KS2] N. Katz and P. Sarnak, *Zeros of zeta functions and symmetries*, Bull. AMS **36**, 1999, $1-26$.

[KeSn] J. P. Keating and N. C. Snaith, *Random matrices and L-functions*,

[Kel] D. Kelley, *Introduction to Probability*, Macmillian Publishing Company, 1994.

[Kh] A. Y. Khinchin, *Continued Fractions*, Third Edition, The University of Chicago Press, Chicago 1964.

[Kn] A. Knapp, *Elliptic Curves*, Princeton University Press, Princeton, 1992.

[Knu] D. Knuth, *The Art of Computer Programming, Vol. 2*, Addison-Wesley, second edition, 1981.

[Kob1] N. Koblitz, *Why study equations over finitess fields?*, Math. Mag. **55** (1982), no. 3, 144-149.

[Kob2] N. Koblitz, *Elliptic curve cryptosystems*, Math. Comp. **48** (1987), no. 177, 203-209.

[Kob3] N. Koblitz, *A survey of number theory and cryptography*, Number theory, Trends Math., Birkhäuser, Basel, 2000, 217–239.

[Ko] V. Kolyvagin, *On the Mordell-Weil group and the Shafarevich-Tate group of modular elliptic curves*, Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990), Math. Soc. Japan, Tokyo, 1991, $429-436$.

[KonMi] A. Kontorovich and S. J. Miller, *Poisson Summation, Benford's Law and values of L-functions*, preprint.

[KonSi] A. Kontorovich and Ya. G. Sinai, *Structure theorem for (d, g, h)-maps*, Bull. Braz. Math. Soc. (N.S.) 33 (2002), no. 2, 213–224.

[Kor] A. Korselt, *Probléme chinois*, L'intermédiaire math. 6, 143-143, 1899.

[Kua] F. Kuan, *Digit distribution in the continued fraction of $\zeta(n)$*, Junior Thesis, Princeton University, Fall 2002,

[KN] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, John Wiley & Sons, New York 1974.

[Ku] R. Kuzmin, *Ob odnoi zadache Gaussa*, Doklady akad. nauk, ser A, 1928, $375-380$.

[La1] S. Lang, *Introduction to Diophantine Approximations*, Addison-Wesley, Reading, 1966.

[La2] S. Lang, *Undergraduate Algebra*, Springer-Verlag, second edition, 1986.

[La3] S. Lang, *Calculus of Several Variables*, Springer-Verlag, 1987.

[La4] S. Lang, *Undergraduate Analysis*, Springer-Verlag; second edition, 1997.

[La5] S. Lang, *Complex Analysis*, Springer-Verlag, Graduate Texts in Mathematics, Vol. 103, 1999.

[LT] S. Lang and H. Trotter, *Continued fractions for some algebraic numbers*, J. Reine Angew. Math. **255**, 1972, $112 - 134$.

[LF] R. Larson and B. Farber, *Elementary Statistics: Picturing the World*, Prentice Hall, Inc., 2003.

[LP] R. Laubenbacher and D. Pengelley, *Gauss, Eisenstein, and the "third" proof of the quadratic reciprocity theorem: Ein kleines Schauspiel*, Math. Intelligencer 16 (1994), no. 2, 67-72.

[Law1] J. Law, *Kuzmin's Theorem on Algebraic Numbers*, Junior Thesis, Princeton University, Fall 2002,

[Law2] J. Law, *The Circle Method on the Binary Goldbach Conjecture*, Junior Thesis, Princeton University, Spring 2003,

[Leh] R. Lehman, *First order spacigns of Random Matrix eigenvalues*, Junior Thesis, Princeton University, 2001.

[Le] P. Lévy, *Sur les lois de probabilité dont dependent les quotients complets et incomplets d'une fraction continue*, Bull. Soc. Math., **57**, 1929, $178 - 194$.

[Li] R. Lipshitz, *Numerical Results concerning the distribution of* $\{n^2\alpha\}$, Junior Thesis, Princeton University, Spring 2000.

[Liu] Y. Liu, *Statistical Behavior of the Eigenvalues of Random Matrices*, Junior Thesis, Princeton University, ,

[Mah] K. Mahler, *Arithmetische Eigenschaften einer Klasse von Dezimalbrüchen*, Akad. Wetensch. Amsterdam Proc. = Indag. Math. 40, 421-428.

[Mai] K. Mainzer, *Natural Numbers, Integers, and Rational Numbers*, Numbers, pp. 9-26.

[Mar] J. Marklof, *Almost modular functions and the distribution of* n^2x *modulo one*, Int. Math. Res. Not. 2003, no. 39, 2131–2151.

[Maz1] B. Mazur, *Modular curves and the Eisenstein ideal*, IHES Publ. Math. 47 (1977), 33-186.

[Maz2] B. Mazur, *Rational isogenies of prime degree (with an appendix by D. Goldfeld)*, Invent. Math. 44 (1978), no. 2, 129–162.

[Maz3] B. Mazur, *Number Theory as Gadfly*, Amer. Math. Monthly, **98** (1991), 593-610.

[McK] B. McKay, *The expected eigenvalue distribution of a large regular graph*, Linear Algebra Appl. 40 (1981), $203 - 216$.

[Meh1] M. Mehta, *On the statistical properties of level spacings in nuclear spectra*, Nucl. Phys. 18, 1960, 395-.

[Meh2] M. Mehta, *Random Matrices, 2nd edition*, Academic Press Inc., Boston, 1991.

[Mic1] M. Michelini, *Independence of the digits of Continued Fractions*, Junior Thesis, Princeton University, Fall 2002,

[Mic2] M. Michelini, *Kuzmin's Extraordinaty Zero Measure Set*, Senior Thesis, Princeton University, May 2004.

[Mi] N. Miller, *Various tendencies of non-Poissonian distributions along subsequences of certain transcendental numbers*, Junior Thesis, Princeton University, Fall 2002, **GIVE WEB LINK**.

[Mill] S. D. Miller, *A simpler way to show $\zeta(3)$ is irrational*, http://www.math.rutgers.edu/$\sim$ sdmiller/simplerzeta3.pdf.

[Mil1] S. J. Miller, 1- *and* 2*-Level Densities for Families of Elliptic Curves: Evidence for the Underlying Group Symmetries*, P.H.D. Thesis, Princeton University, 2002, http://www.math.princeton.edu/$\sim$sjmiller/thesis/thesis.pdf.

[Mil2] S. J. Miller, 1- *and* 2*-Level Densities for Families of Elliptic Curves: Evidence for the Underlying Group Symmetries*, Compositio Mathematica, to appear.

[Mon] H. Montgomery, *The pair correlation of zeros of the zeta function*, Analytic Number Theory, Proc. Sympos. Pure Math. **24**, Amer. Math. Soc., Providence, 1973, $181 - 193$.

[MoMc] D. Moore and G. McCabe, *Introduction to the practice of statistics*, W. H. Freeman and Co., 2003.

[MS] H. Montgomery and K. Soundararajan, *Beyond Pair Correlation*,

[Moz1] C. J. Mozzochi, *The Fermat Diary*, American Mathematical Society, 2000.

[Moz2] C. J. Mozzochi, *The Fermat Proof*, Trafford Publishing, 2004.

[Mu1] R. Murty, *Primes in certain arithmetic progressions*, Journal of the Madras University, (1988) 161-169.

[Mu2] R. Murty, *Problems in Analytic Number Theory*, Graduate Texts in Mathematics, Vol. 206, Springer-Verlag, 2001.

[MM] M. R. Murty and V. K. Murty, *Non-vanishing of L-Functions and Applications*, Birkhäuser, Progress in Mathematics, Vol 157, 1997.

[Na] M. Nathanson, *Additive Number Theory: The Classical Bases*, Springer-Verlag, Graduate Texts in Mathematics, 1996.‘

[NT] J. von Neumann and B. Tuckerman, *Continued fraction expansion of* $2^{1/3}$, Math. Tables Aids Comput. **9** (1955), 23-24.

[NZM] I. Niven, H. Zuckerman, and H. Montgomery, *An Introduction to the Theory of Numbers*, John Wiley & Sons, Inc., fifth edition, 1991.

[Od1] A. Odlyzko, *On the distribution of spacings between zeros of the zeta function*, Math. Comp. **48**, 1987, no. 177, $273 - 308$.

[Od2] A. Odlyzko, *The* 10^{22}*-nd zero of the Riemann zeta function*, Proc. Conference on Dynamical, Spectral and Arithmetic Zeta-Functions, M. van Frankenhuysen and M. L. Lapidus, eds., Amer. Math. Soc., Contemporary Math. series, 2001, http://www.research.att.com/∼amo/doc/zeta.html

[Ok] T. Okano, *A note on the transcendental continued fractions*, Tokyo J of Math, 1987. **I will find the exact reference later**

[Ol] T. Oliveira e Silva, *Verification of the Goldbach Conjecture Up to* $6 \cdot 10^{16}$, NMBRTHRY@listserv.nodak.edu mailing list, Oct. 3, 2003, http://listserv.nodak.edu/scripts/wa.exe?A2=ind0310&L=nmbrthry&P=168 and http://www.ieeta.pt/∼tos/goldbach.html.

[Ols] L. Olsen, *Extremely non-normal continued fractions*, Acta Arith. 108 (2003), no. 2, 191-202.

[vdP] A. van der Poorten, *Notes on Fermat's Last Theorem*, John Wiley & Sons.

[Po] C. Porter (editor), Statistical Theories of Spectra: Fluctuations, Academic Press, 1965.

[QS] R. Qian and D. Steinhauer, *Eigenvalues of weighted random graphs*, Junior Thesis, Princeton University, Fall 2003.

[Re] F. Reif, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill.

[RDM] R. D. Richtmyer, M. Devaney and N. Metropolis, *Continued fraction of algebraic numbers* Numer. Math. **4** (1962) 68-84.

[Ri] G. F. B. Riemann, *Über die Anzahl der Primzahlen unter einer gegebenen Grösse*. Monatsber. Königl. Preuss. Akad. Wiss. Berlin, 671-680, Nov. 1859. **I think there is a translation of this in Edward's book**

[RSA] R. Rivest, A. Shamir, A. and L. Adleman, *A Method for Obtaining Digital Signatures and Public Key Cryptosystems*, Comm. ACM 21, 120-126, 1978.

[Ro] K. Roth, *Rational approximations to algebraic numbers*, Mathematika 2, 1955, $1 - 20$.

[Rub1] M. Rubinstein, *A simple heuristic proof of Hardy and Littlewood's conjecture B*, Amer. Math. Monthly 100, 1993, no. 5, 456-460.

[Rub2] M. Rubinstein, *Low-lying zeros of L-functions and random matrix theory*, Duke Math. J. **109** (2001), no. 1, 147–181.

[RubSa] M. Rubinstein and P. Sarnak, *Chebyshev's bias*, Experiment. Math. 3 (1994), no. 3, 173-197.

[Rud] W. Rudin, *Principles of Mathematical Analysis*, third edition, International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, 1976.

[RS] Z. Rudnick and P. Sarnak, *Zeros of principal L-functions and random matrix theory*, Duke Journal of Math. **81**, 1996, $269 - 322$.

[RS2] Z. Rudnick and P. Sarnak, *The pair correlation function of fractional parts of polynomials*, Comm. Math. Phys. 194 (1998), no. 1, 61–70.

[RSZ] Z. Rudnick, P. Sarnak, and A. Zaharescu, *The Distribution of Spacings Between the Fractional Parts of $n^2\alpha$*, Invent. Math. 145 (2001), no. 1, $37 - 57$.

[Sar] P. Sarnak *Some applications of modular forms*, Cambridge Trusts in Math. **99**, Cambridge University Press, 1990.

[Sch] D. Schmidt, *Prime Spacing and the Hardy-Littlewood Conjecture B*, Junior Thesis, Princeton University, Spring 2001,

[Se] J. P. Serre, *A Course in Arithmetic*, Springer-Verlag.

[SS] E. Stein and R. Shakarchi, *Fourier Analysis: An Introduction*, Princeton University Press, 2003.

[Sil] J. Silverman, *The Arithmetic of Elliptic Curves*, Graduate Texts in Mathematics **106**, Springer-Verlag, Berlin - New York, 1986.

[ST] J. Silverman and J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag, 1992.

[Sk] Skewes. J. London Math. Soc. 8, 277-283, 1933.

[Ste] I. Stewart, *Algebraic Number Theory*,

[St] Strang, *Linear Algebra and Its Applications*, International Thomson Publishing, 3rd edition.

[Sz] P. Szüsz, *On the length of continued fractions representing a rational number with given denominator*, Acta Arithmetica XXXVII, 1980, 55-59.

[Ta] C. Taylor, *The Gamma function and Kuzmin's Theorem*, Junior Thesis, Princeton University, Fall 2002,

[TW] R. Taylor and A. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. Math. **141**, 1995, $553 - 572$.

[TrWi] C. Tracy and H. Widom, *Correlation functions, cluster functions, and spacing distributions for random matrices*, J. Statist. Phys., **92** (5-6), 1998, 809Ű835.

[Te] Tenenbaum, *Introduction to analytic and probabilistic number theory*

[Ti] E. C. Titchmarsh, *The Theory of the Riemann Zeta-function*, Revised by D. R. Heath-Brown, Oxford 1986.

[Vin1] I. Vinogradov, *Representation of an odd number as the sum of three primes*, Doklady Akad. Nauk SSSR, $15(6 - 7)$: $291 - 294$, 1937.

[Vin2] I. Vinogradov, *Some theorems concerning the theory of primes*, Mat. Sbornik, $2(44)$: $179 - 195$, 1937.

[Vo] A. Voros, *A sharpening of Li's criterion for the Riemann Hypothesis*, http://arxiv.org/pdf/math.NT/0404213.

[VG] W. Voxman and R. Goetschel, Jr., *Advanced Calculus*, Mercer Dekker Inc., 1981.

[Wa] L. Washington, *Elliptic curves: Number theory and cryptography*, Chapman & Hall/CRC, 2003.

[Wei] A. Weil, *Numbers of Solutions of Equations in Finite Fields*, Bull. Amer. Math. Soc. **14**(1949), 497-508.

[Weir] B. Weir, *The local behavior of Germain primes*, Undergraduate Mathematics Laboratory report, Courant Institute, NYU, http://www.math.nyu.edu/Courses/V63.0393/projects/germainprimes/germain.htm.

[We] E. Weisstein, *MathWorld–A Wolfram Web Resource*, http://mathworld.wolfram.com/

[Wh] E. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies: With an Introduction to the Problem of Three Bodies*, Dover, 1944.

[WW] E. Whittaker and G. Watson, *A Course of Modern Analysis*, Cambridge University Press; 4th edition, 1996.

[Wig1] E. Wigner, *On the statistical distribution of the widths and spacings of nuclear resonance levels*, Proc. Cambridge Philo. Soc. **47**, 1951, $790 - 798$.

[Wig2] E. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Ann. of Math. **2**, 62, 1955, 548Ű564.

[Wig3] E. Wigner, *Statistical Properties of real symmetric matrices*, Canadian Mathematical Congress Proceedings, University of Toronto Press, Toronto, 1957, 174-184.

[Wig4] E. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions. II*, Ann. of Math. **2**, 65, 1957, 203Ű207.

[Wig5] E. Wigner, *On the distribution of the roots of certain symmetric matrices. Ann. of Math*, **2**, 67, 1958, 325Ű327.

[Wi] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. Math **141**, 1995, $443 - 551$.

[Wir] E. Wirsing, *On the Theorem of Gauss-Kuzmin-Lévy and a Frobenius-Type Theorem for Function Spaces*, Acta Arith. **24**, 1974, 507-528.

[Wom] N. C. Wormald, *Models of random regular graphs*,

[Wo] T. Wooley, *Large improvements in Waring's problem*, Ann. Math., 135, $131 - 164$.

[Za] I. Zakharevich, *A Generalization of Wigner's Law*, preprint.

[Zu] W. Zudilin, *One of the Numbers* $\zeta(5), \zeta(7), \zeta(9), \zeta(11)$ *Is Irrational*, Uspekhi Mat. Nauk 56, 2001, 149-150.

[Zy] A. Zygmund, *Trigonometrical Series, Volume I and II*, Cambridge University Press, 1968.

Index