RAMANUJAN-FOURIER SERIES AND APPLICATIONS

Ryan Case

A thesis submitted in
partial fulfillment of the
requirements for the Degree
of Master of Science in
Mathematics and Statistics

Queen’s University
Department of Mathematics and Statistics
August, 2013

Copyright © Ryan Case, 2013
Abstract

We look at infinite series expansions for arithmetic functions, first considered by Srinivasa Ramanujan in 1918. A basis for these expansions is investigated, for which several properties are proven. Examples of these infinite series are established using multiple techniques. Finally, we apply this theory to study the famous twin prime problem, and the problem of computing exact values of arithmetic functions.
Acknowledgements

A resounding thank you to my supervisor, Ram Murty. I learned a lot in your classes and while working with you at the black board in your office, and I appreciate the advice and guidance you’ve offered me throughout the year.

Thank you to the staff in the department of mathematics and statistics at Queen’s, especially Jennifer Read. You’ve been extremely helpful since before I even arrived at Queen’s and I’m grateful for everything you’ve done for me.

I would also like to thank Serban Belinschi and Mike Roth for serving on my examination committee and offering their thoughts and corrections on this project.
# Table of Contents

1 Introduction 2

2 Properties of Ramanujan sums 4

3 Ramanujan-Fourier series 18

4 Some examples 25

5 The twin prime conjecture and the Wiener-Khinchin formula 34

6 Computing values of arithmetic functions 44

Bibliography 48
Chapter 1

Introduction

In 1918 the Indian mathematician Srinivasa Ramanujan published a paper titled “On Certain Trigonometrical Sums and their Applications in the Theory of Numbers”[16] in which he studied sums of the form

\[ c_q(n) = \sum_{\substack{a=1 \\ \ \ \ (a,q)=1}}^{q} \cos \left( 2\pi n \frac{a}{q} \right), \]

where \( q \) and \( n \) are natural numbers and \((a, q)\) is the greatest common divisor of \( a \) and \( q \). More recently we consider equal sums, now called Ramanujan sums, of the form

\[ c_q(n) = \sum_{\substack{a=1 \\ \ \ \ (a,q)=1}}^{q} e^{2\pi in \frac{a}{q}} \]

and for cleanliness of notation we will write \( e(t) = e^{2\pi it} \) for \( t \in \mathbb{R} \).

An arithmetic function is a complex valued function defined on the set of natural numbers, the most useful of which express some number-theoretic property. Ramanujan used the sums \( c_q(n) \) as a sort of basis to represent arithmetic functions as infinite series in a way that is analogous to the Fourier series representation of a function. It is this idea that will be the focus of this paper.
We will follow the reasoning of Gadiyar and Padma and use these infinite series as a tool to study the twin prime problem and generalizations. By choosing an appropriate arithmetic function and naively applying a natural analogue of the Wiener-Khinchin formula, we make a conjecture about the asymptotic value of the number of twin primes less than any value.

Finally, we use the infinite series representation of an arithmetic function to compute its values exactly. By truncating the sum after sufficiently many summands we are assured to be close enough to the actual value of the function that rounding the result will give us the correct value.
Chapter 2

Properties of Ramanujan sums

Ramanujan sums have some useful properties that we will need to establish before we can begin our investigation of arithmetic functions.

1. The function \( c_q(n) \) is multiplicative in the argument \( q \), that is, if \( (q, q') = 1 \) then

\[
c_{qq'}(n) = c_q(n)c_{q'}(n)
\]

for every \( n \in \mathbb{N} \).

To prove this property we will need a small lemma about the residue classes of the product of coprime integers.

**Lemma 1.** Suppose \( q, q' \in \mathbb{N} \) with \( (q, q') = 1 \), then every residue class modulo \( qq' \) can be written uniquely in the form \( aq + bq' \) with \( 1 \leq a \leq q' \) and \( 1 \leq b \leq q \). Moreover, every coprime residue class modulo \( qq' \) can be written uniquely as \( aq + bq' \) with \( (a, q') = 1 \) and \( (b, q) = 1 \).

**Proof.** We first show that all residue classes modulo \( qq' \) of the form \( aq + bq' \) are
distinct. Indeed suppose not, and say
\[ a_1 q + b_1 q' \equiv a_2 q + b_2 q' \pmod{qq'}. \]
Then
\[ a_1 q + b_1 q' \equiv a_2 q + b_2 q' \pmod{q} \]
and \[ a_1 q + b_1 q' \equiv a_2 q + b_2 q' \pmod{q'} \]
so
\[ q(a_1 - a_2) \equiv 0 \pmod{q'} \]
and \[ q'(b_1 - b_2) \equiv 0 \pmod{q'} \].
Since \( q \) and \( q' \) are coprime, there exists \( q^{-1} \) modulo \( q' \) and \( q'^{-1} \) modulo \( q \), and therefore
\[ a_1 \equiv a_2 \pmod{q'} \]
and \[ b_1 \equiv b_2 \pmod{q} \].
Our hypothesis was that \( 1 \leq a_i \leq q' \) and \( 1 \leq b_i \leq q \) for \( i = 1, 2 \) so \( a_1 = a_2 \) and \( b_1 = b_2 \). It follows that the residue classes \( aq + bq' \) with \( 1 \leq a \leq q' \) and \( 1 \leq b \leq q \) are unique modulo \( qq' \), and since there are \( qq' \) ways to choose \( a \) and \( b \), each residue class modulo \( qq' \) is of this form with no repetition.

To prove the second claim, notice that if \((a, q') = d > 1\) say, then \( d|aq + bq'\) so \((aq + bq', qq') \geq d > 1\). Therefore if \((aq + bq', qq') = 1\), then \((a, q') = 1\). Similarly if \((aq + bq', qq') = 1\) we see that \((b, q) = 1\). We have shown above that each residue class modulo \( qq' \) can be represented as \( aq + bq' \), with \( 1 \leq a \leq q' \) and \( 1 \leq b \leq q \) in exactly one way, so this gives us the uniqueness condition for the representation of the coprime residue classes modulo \( qq' \).

Now we can use this lemma to prove property 1.
Proof of Property 1. If \( q \) and \( q' \) are coprime, then

\[
c_{qq'}(n) = \sum_{a=1}^{qq'} \sum_{(a,qq')=1} e\left( \frac{na}{qq'} \right)
\]

\[
= \sum_{a=1}^{q'} \sum_{b=1}^{q} e\left( \frac{na}{qq'} \right) e\left( \frac{nb}{q} \right), \text{ by lemma 1,}
\]

\[
= \sum_{a=1}^{q'} e\left( \frac{na}{q'} \right) \sum_{b=1}^{q} e\left( \frac{nb}{q} \right)
\]

\[
= c_{q'}(n)c_q(n).
\]

\( \square \)

2 The function \( c_q(n) \) is periodic in the argument \( n \) with period \( q \).

Proof. Fix \( q, n \in \mathbb{N} \) and \( k \in \mathbb{Z} \), then

\[
c_q(n + kq) = \sum_{a=1}^{q} e\left( \frac{n + kq}{q} \right)
\]

\[
= \sum_{a=1}^{q} e\left( \frac{n}{q} \right) e(k)
\]

\[
= \sum_{a=1}^{q} e\left( \frac{n}{q} \right)
\]

\[
= c_q(n).
\]

\( \square \)
3 If $\mu$ is the usual Möbius function defined as

$$\mu(n) = \begin{cases} 
(-1)^k & \text{if } n = p_1p_2 \cdots p_k \text{ for distinct primes } p_i, \\
0 & \text{otherwise}, 
\end{cases}$$

then

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

In particular, $c_q(n)$ is an integer.

The proof of property 3 will require the Möbius inversion formula which we will state as a lemma. Proof can be found in Hardy and Wright[10] theorem 266.

**Lemma 2.** Let $f$ and $g$ be two arithmetic functions such that

$$g(n) = \sum_{d|n} f(d)$$

for $n \in \mathbb{N}$. Then

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

**Proof of property 3.** Observe that the sum $c_q(n)$ is just the sum of the $n^{\text{th}}$ powers of the primitive $q^{\text{th}}$ roots of unity. Fix $\zeta_q = e^{\frac{2\pi i}{q}}$, then the powers of $\zeta_q$ are all the $q^{\text{th}}$ roots of unity. It is well known, for example in Hardy and Wright[10] theorems 65 and 66, that the set of $q^{\text{th}}$ roots of unity is exactly the union of the sets of primitive $d^{\text{th}}$ roots of unity for every $d$ that divides $q$.

Define

$$\eta_q(n) = \sum_{k=1}^{q} \zeta_q^{kn},$$

then by the above discussion we see that

$$\eta_q(n) = \sum_{d|q} c_d(n).$$
We can apply Möbius inversion to this equation with 

\[ f(q) = c_q(n) \text{ and } g(q) = \eta_q(n) \]

to see that

\[ c_q(n) = \sum_{d|q} \mu\left(\frac{q}{d}\right) \eta_d(n). \]

It is easy to see that if \( q \mid n \), then \( \zeta_q^{kn} = 1 \) for all \( k \in \mathbb{Z} \), so \( \eta_q(n) = q \). If \( q \nmid n \) then \( \zeta_q^n \neq 1 \) so we get

\[
\eta_q(n) = \sum_{k=1}^{q} \zeta_q^{kn} \\
= \sum_{k=0}^{q} \zeta_q^{kn} - 1 \\
= \frac{1 - (\zeta^n)^{q+1}}{1 - \zeta^n} - 1 \\
= \frac{1 - \zeta^{qn+n}}{1 - \zeta^n} - 1 \\
= \frac{1 - \zeta^n}{1 - \zeta^n} - 1 \\
= 0.
\]

Therefore

\[
\eta_d(n) = \begin{cases} 
  d & \text{if } d \mid n, \\
  0 & \text{if } d \nmid n,
\end{cases}
\]

and it follows that

\[
c_q(n) = \sum_{d|(q,n)} \mu\left(\frac{q}{d}\right) d.
\]

4 If \( p \) is a prime, then

\[
c_p(n) = \begin{cases} 
  p - 1 & \text{if } p \mid n, \\
  -1 & \text{if } p \nmid n,
\end{cases}
\]

\[ \square \]
and
\[ c_{p^k}(n) = \begin{cases} 
0 & \text{if } p^{k-1} \nmid n, \\
-p^{k-1} & \text{if } p^{k-1} \mid n \text{ and } p^k \nmid n, \\
p^{k-1}(p-1) & \text{if } p^k \mid n,
\end{cases} \]
for any positive integer \( k > 1 \).

**Proof.** This follows from property 3 and the definition of \( c_q(n) \). If \( p \mid n \) then
\[ c_p(n) = \mu(p) \cdot 1 + \mu(1) \cdot p = p - 1 \]
and if \( p \nmid n \) then
\[ c_p(n) = \mu(p) \cdot 1 = -1. \]

Now suppose \( k > 1 \) is an integer and \( p^k \mid n \). Then
\[ c_{p^k}(n) = \sum_{a=1}^{p^k} e \left( \frac{na}{p^k} \right) 
= \sum_{a=1 \atop (a,p^k)=1}^{p^k} 1 
= \varphi(p^k) = p^{k-1}(p-1), \]
where \( \varphi(n) \) is the Euler totient function that counts the number of coprime residue classes modulo \( n \).

If \( p^{k-1} \mid n \) but \( p^k \nmid n \), we can use the last result to see that
\[
p^{k-1}(p-1) = \sum_{d \mid (p^k,n)} \mu \left( \frac{p^k}{d} \right) d 
= \sum_{d \mid (p^{k-1},n)} \mu \left( \frac{p^k}{d} \right) d + \mu(1)p^k 
= c_{p^k}(n) + p^k.
\]
and by rearranging we get

\[ c_p^k(n) = p^{k-1}(p - 1) - p^k = -p^{k-1}. \]

Finally if \( p^{k-1} \nmid n \), then by using the same method as above can see

\[
p^{k-1}(p - 1) = \sum_{d|\,(p^k,n)} \mu\left(\frac{p^k}{d}\right) d
= \sum_{d|\,(p^{k-2},n)} \mu\left(\frac{p^k}{d}\right) d + \mu(p)p^{k-1} + \mu(1)p^k
= c_{p^k}(n) - p^{k-1} + p^k.
\]

If we rearrange this we get

\[ c_{p^k}(n) = p^{k-1}(p - 1) + p^{k-1} - p^k = 0. \]

5 An “orthogonality” relation:

Let \( N = \text{lcm}(q, r) \), then

\[
\frac{1}{N} \sum_{n=1}^{N} c_q(n)c_r(n) = \begin{cases} 
\varphi(N) & \text{if } q = r = N, \\
0 & \text{if } q \neq r
\end{cases}
\]

Furthermore,

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} c_q(n)c_r(n + h) = \begin{cases} 
c_q(h) & \text{if } q = r, \\
0 & \text{if } q \neq r.
\end{cases}
\]
Proof. We begin with the first identity. Observe that if \( a, b, q, \) and \( r \) are positive integers, \((a, q) = 1, (b, r) = 1, \) and \( q \neq r \) then \( a/q + b/r \notin \mathbb{Z} \). Indeed if it were, then
\[
a/q + b/r = n,
\]
so
\[
ar = nr - b \quad \text{and} \quad bq = nq - a
\]
and since \( nr - b, nq - a \in \mathbb{Z} \) and \((a, q) = 1, (b, r) = 1 \) we get that \( q|r \) and \( r|q \). Therefore \( q = r \), a contradiction.

Now if \( q \neq r \), then
\[
\frac{1}{N} \sum_{n=1}^{N} c_q(n)c_r(n)
\]
\[
= \frac{1}{N} \sum_{a=1}^{q} \sum_{b=1}^{r} \sum_{n=1}^{N} e \left( \frac{na}{q} \right) e \left( \frac{nb}{r} \right)
\]
\[
= \frac{1}{N} \sum_{a=1}^{q} \sum_{b=1}^{r} \sum_{n=1}^{N} e \left( n \left( \frac{a}{q} + \frac{b}{r} \right) \right)
\]
\[
= \frac{1}{N} \sum_{a=1}^{q} \sum_{b=1}^{r} \left( \frac{1 - e \left( n \left( \frac{a}{q} + \frac{b}{r} \right) \right)}{1 - e \left( n \left( \frac{a}{q} + \frac{b}{r} \right) \right)} - 1 \right)^{N+1}
\]
\[
= \frac{1}{N} \sum_{a=1}^{q} \sum_{b=1}^{r} \left( \frac{1 - e \left( n \left( \frac{a}{q} + \frac{b}{r} \right) \right)}{1 - e \left( n \left( \frac{a}{q} + \frac{b}{r} \right) \right)} - 1 \right)
\]
\[
= \frac{1}{N} \sum_{a=1}^{q} \sum_{b=1}^{r} (1 - 1)
\]
\[
= 0,
\]
since by the discussion above, \( e \left( n \left( \frac{a}{q} + \frac{b}{r} \right) \right) \neq 1.\)
If \( q = r = N \), then
\[
\sum_{n=1}^{N} e\left( n \left( \frac{a + b}{N} \right) \right) = \begin{cases} 
N & \text{if } \frac{a+b}{N} \in \mathbb{Z}, \\
0 & \text{otherwise.}
\end{cases}
\]
This is easy to see using the same argument as we have used above since this is a geometric sum.

Notice that if \( a, N \in \mathbb{N}, p \mid N \) and \( p \nmid a \) then \( p \nmid N - a \). Thus, \((a, N) = 1\) implies that \((N - a, N) = 1\). Even more, we know that given \( a \) as above, there is only one \( 0 < b < N \) such that \( a + b \equiv 0 \pmod{N} \) since \( 0 < a + b < 2N \). Therefore
\[
\# \{ a \in \mathbb{N} : a + b = N, (a, N) = 1, (b, N) = 1, b > 0 \} = \varphi(N).
\]

It follows that
\[
\frac{1}{N} \sum_{n=1}^{N} c_N(n) c_N(n) = \frac{1}{N} \sum_{n=1}^{N} \sum_{a=1}^{N} \sum_{b=1}^{N} e\left( n \left( \frac{a + b}{N} \right) \right) = \frac{1}{N} \sum_{a=1}^{N} \sum_{b=1}^{N} \sum_{n=1}^{N} e\left( n \left( \frac{a + b}{N} \right) \right) = \frac{1}{N} \varphi(N) N = \varphi(N),
\]
so the first identity is true.
For the second identity observe that
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} c_q(n)c_r(n+h) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left( \sum_{\substack{a=1 \\text{(mod } q) \\text{ and } (a,q)=1}}^{q} e\left(\frac{na}{q}\right) \right) \left( \sum_{\substack{b=1 \\text{(mod } r) \\text{ and } (b,r)=1}}^{r} e\left(\frac{(n+h)b}{r}\right) \right)
\]
\[
= \sum_{\substack{b=1 \\text{(mod } r) \\text{ and } (b,r)=1}}^{r} e\left(\frac{hb}{r}\right) \sum_{\substack{a=1 \\text{(mod } q) \\text{ and } (a,q)=1}}^{q} \left( \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e\left(\frac{n(ar+bq)}{qr}\right) \right)
\]
\[
= c_r(h) \sum_{\substack{a=1 \\text{(mod } q) \\text{ and } (a,q)=1}}^{q} \left( \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e\left(\frac{n(ar+bq)}{qr}\right) \right).
\]

If \( qr \mid ar + bq \), then \( e\left(\frac{(ar+bq)}{qr}\right) = 1 \) so
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e\left(\frac{n(a+b)}{qr}\right) = \lim_{N \to \infty} \frac{1}{N} N = 1.
\]

As we have shown previously, this occurs only when \( r = q \), and for fixed \( b \) for only one \( 0 \leq a < q \).

In the cases where \( qr \nmid ar + bq \), we know that \( e\left(\frac{(a+b)}{qr}\right) \neq 1 \) and therefore we can find some \( \epsilon(a,b) > 0 \) such that
\[
\left| 1 - e\left(\frac{(a+b)}{qr}\right) \right| > \epsilon(a,b).
\]
Then
\[
\lim_{N \to \infty} \left| \frac{1}{N} \sum_{n=1}^{N} e \left( \frac{n(a+b)}{qr} \right) \right| 
\]
\[
= \lim_{N \to \infty} \frac{1}{N} \left| \frac{1 - e \left( \frac{a+b}{qr} \right)^{N+1}}{1 - e \left( \frac{a+b}{qr} \right)} - 1 \right|
\]
\[
\leq \lim_{N \to \infty} \frac{1}{N} \left( \left| \frac{1 - e \left( \frac{a+b}{qr} \right)^{N+1}}{1 - e \left( \frac{a+b}{qr} \right)} \right| + 1 \right)
\]
\[
\leq \lim_{N \to \infty} \frac{1}{N} \left( \frac{2}{\epsilon(a,b) + 1} \right)
\]
\[
= 0,
\]
and the second identity follows.

6 Two upper bounds of the function \( c_q(n) \):

- For fixed \( q \), \( c_q(n) \leq \varphi(q) \),
- For fixed \( n \), \( c_q(n) \leq \sigma(n) \),

where \( \sigma(n) \) is the sum of the divisors of \( n \).

**Proof.** First, fix \( q \), then
\[
|c_q(n)| = \left| \sum_{a=1 \atop (a,q)=1}^{q} e \left( \frac{na}{q} \right) \right|
\]
\[
\leq \sum_{a=1 \atop (a,q)=1}^{q} \left| e \left( \frac{na}{q} \right) \right|
\]
\[
= \varphi(q).
\]
Next, for fixed $n$ we can use property 3 to get

\[
|c_q(n)| = \left| \sum_{d\mid (q,n)} \mu \left( \frac{q}{d} \right) d \right| \\
\leq \sum_{d\mid (q,n)} \left| \mu \left( \frac{q}{d} \right) d \right| \\
\leq \sum_{d\mid n} d \\
= \sigma(n).
\]

7 An incredible formula:

\[ c_q(n) = \mu \left( \frac{q}{(q,n)} \right) \frac{\varphi(q)}{\varphi \left( \frac{q}{(q,n)} \right)}. \]

Proof. For simplicity of notation define

\[ q' = \frac{q}{(q,n)}. \]

Then

\[
c_q(n) = \sum_{d\mid (q,n)} \mu \left( \frac{q}{d} \right) d \\
= \sum_{cd=(q,n)} \mu \left( \frac{q}{d} \right) d \\
= \sum_{c\mid (q,n)} \mu(cq') \frac{(q,n)}{c} \\
= (q,n) \sum_{c\mid (q,n)} \frac{\mu(cq')}{c}.
\]

Observe that

\[
\mu(cq') = \begin{cases} 
\mu(c)\mu(q') & \text{if } (c,q') = 1, \\
0 & \text{otherwise},
\end{cases}
\]
therefore
\[ c_q(n) = (q, n)\mu(q') \sum_{\substack{c \mid (q, n) \\ (c, q') = 1}} \frac{\mu(c)}{c}. \]

The sum
\[ \sum_{\substack{c \mid (q, n) \\ (c, q') = 1}} \frac{\mu(c)}{c} \]

is made up of terms that are the reciprocal of the product of a subset of the distinct primes that divide \((q, n)\), but don’t divide \(q'\). It is evident then that
\[
\sum_{\substack{c \mid (q, n) \\ (c, q') = 1}} \frac{\mu(c)}{c} = \prod_{\substack{p \mid (q, n) \\ p \nmid q'}} \left(1 + \frac{\mu(p)}{p}\right)
\]
\[
= \prod_{\substack{p \mid (q, n) \\ p \nmid q'}} \left(1 - \frac{1}{p}\right),
\]
so we have
\[
c_q(n) = (q, n)\mu(q') \prod_{\substack{p \mid (q, n) \\ p \nmid q'}} \left(1 - \frac{1}{p}\right)
\]
\[
= \mu(q')(q, n) \frac{\varphi(q)}{\varphi(q)} \prod_{\substack{p \mid (q, n) \\ p \nmid q'}} \left(1 - \frac{1}{p}\right)
\]
\[
= \mu(q')(q, n) \frac{\varphi(q)}{q \prod_{p \mid q'} \left(1 - \frac{1}{p}\right)} \prod_{\substack{p \mid (q, n) \\ p \nmid q'}} \left(1 - \frac{1}{p}\right)
\]
\[
= \mu(q') \frac{\varphi(q)}{q' \prod_{p \mid q'} \left(1 - \frac{1}{p}\right)} \prod_{\substack{p \mid (q, n) \\ p \nmid q'}} \left(1 - \frac{1}{p}\right)
\]
\[
= \frac{\mu(q') \varphi(q)}{\varphi(q')}.
\]
The properties above were not published first by Ramanujan. In 1906 Kluyver[13] published some results, including property 3 above, about these sums, but in Ramanujan’s words, “they have never been considered from the point of view which [he adopts]” in his 1918 paper. It is this point of view, and the interesting results he produces using these sums, that caused G.H. Hardy to name them after Ramanujan. The equation in property 7 seems to not have been found by Ramanujan, and was first published in 1936 by Hölder.
Chapter 3

Ramanujan-Fourier series

The properties of Ramanujan sums that we have proved now lead us to the main topic of the previously mentioned 1918 paper. In a way that is analogous to the Fourier series expansion of a function, Ramanujan used these sums as a basis for infinite series expansions for arithmetic functions.

**Definition.** Let $a : \mathbb{N} \to \mathbb{C}$ be an arithmetic function. Then a Ramanujan-Fourier series, or Ramanujan expansion, for the function $a(n)$ is an infinite series of the form

$$a(n) = \sum_{q=1}^{\infty} a_q c_q(n).$$

Using elementary methods he was able to produce infinite series expansions for many of the commonly used arithmetic functions. A typical example is

$$\frac{\sigma(n)}{n} = \frac{\pi^2}{6} \sum_{q=1}^{\infty} \frac{c_q(n)}{q^2},$$

where $\sigma(n) = \sum_{d|n} d$.

It is not true that Ramanujan expansions for all arithmetic functions can be found using elementary properties of infinite series and simple algebra. In fact the following
expansion,

\[ 0 = \sum_{q=1}^{\infty} \frac{c_q(n)}{q}, \]

is actually equivalent to the prime number theorem as we will see in chapter 4. This example also shows us that the Ramanujan expansion of a function is not unique in general, since the function that is identically zero also has the trivial expansion with all coefficients equal to zero.

Absent from Ramanujan’s paper was a formula for the general coefficient in a Ramanujan expansion. A special case of this was done later in 1930 by Carmichael[3]. In that paper Carmichael also generalized Ramanujan’s idea so that any arithmetic function with similar properties to \( c_q(n) \) can be used in the same way as a basis for an infinite series expansion for another arithmetic function.

Carmichael’s results were heavily based on the concept of the mean-value of an arithmetic function.

**Definition.** For an arithmetic function \( a(n) \), the limit

\[ M(a) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} a(n), \]

if it exists, is called the mean value of the function \( a \).

Carmichael’s discovery was a general formula for the coefficients of the Ramanujan expansion of an arithmetic function \( a(n) \) for which the mean value \( M(ac_q) \) exists for every \( q \in \mathbb{N} \). It is not true in general that the mean value \( M(ac_q) \) exists for every \( q \in \mathbb{N} \), but some progress has been made as to when we can be sure of existence.

We will require some preliminary definitions before those results can be stated. Additionally, the following results deal with two classes of arithmetic functions,
namely additive and multiplicative functions. As we have seen before a multiplicative function is an arithmetic function \( a(n) \) for which

\[
a(nm) = a(n)a(m)
\]

whenever \((n, m) = 1\). An example of this is the Ramanujan sum \( c_{q}(n) \) as was proved previously. An additive function is an arithmetic function for which

\[
a(nm) = a(n) + a(m)
\]

whenever \((n, m) = 1\). An example of an additive function is the restriction of the logarithm to the natural numbers, since

\[
\log(ab) = \log(a) + \log(b),
\]

in this case for all integers \( a \) and \( b \).

We will also define a semi-norm that will provide us with a crucial hypothesis for the upcoming results.

**Definition.** For an arithmetic function \( a(n) \) and \( 1 \leq q < \infty, \ q \in \mathbb{R}, \) define the semi-norm

\[
\|a\|_{q} = \left( \limsup_{x \to \infty} \frac{1}{x} \sum_{n \leq x} |a(n)|^{q} \right)^{1/q}
\]

and

\[
\|a\|_{\infty} = \sup \{|a(n)| : n \in \mathbb{N}\}.
\]

Note that when \( q = \infty \) we actually have a norm above.

We will associate to an arithmetic function \( a(n) \) the Dirichlet series

\[
\tilde{a}(s) = \sum_{n=1}^{\infty} \frac{a(n)}{n^{s}},
\]
absolutely convergent in the half-plane $\Re(s) > k + 1$ when the sequence $\{a_n\}_{n=1}^{\infty}$ satisfies $a_n = O(n^k)$.

For any prime $p$ define the function $a_p : \mathbb{N} \to \mathbb{C}$ by

$$a_p(n) = \begin{cases} a(n) & \text{if } n = p^k, k \in \mathbb{N} \cup \{0\}, \\ 0 & \text{otherwise.} \end{cases}$$

Also, for $l \in \mathbb{N} \cup \{0\}$, define

$$\tilde{a}_{p,l}(s) = \sum_{k=l}^{\infty} a(p^k) p^{ks}.$$

Finally we require one more definition before the results can be stated.

**Definition.** For two arithmetic functions $a$ and $b$, define the Dirichlet convolution of $a$ and $b$, denoted $a \ast b$, by

$$(a \ast b)(n) = \sum_{d|n} a(d) b \left( \frac{n}{d} \right).$$

We can now state our results that give us conditions for the existence of the coefficients of the Ramanujan expansion for certain arithmetic functions.

**Theorem.** For a multiplicative arithmetic function $a(n)$, let $\gamma = \mu \ast a$ and suppose $\|a\|_q < \infty$ for some $q > 1$. If $M(a)$ exists, and is nonzero, then

$$M(a) = \prod_p \tilde{\gamma}_p(1),$$

and $a$ has a Ramanujan expansion with coefficients

$$a_q = M(a) \prod_{p'|q} \frac{\tilde{\gamma}_{p'}(1)}{\tilde{\gamma}_p(1)},$$

where $p' \parallel n$ means that $p' \parallel n$ but $p'^{k+1} \nmid n$. 21
The formula for the mean value $M(a)$ is due to Elliott[6], which was applied by Tuttas[17] and Indlekofer[12] to arrive at the formula for the coefficients of the Ramanujan expansion. The next theorem about additive functions is due to Hildebrand and Spilker[11].

**Theorem.** For an additive arithmetic function $a(n)$, let $\gamma = \mu * a$ and suppose $\|a\|_q < \infty$ for some $q > 1$. If $M(a)$ exists, then

$$M(a) = \sum_p \gamma_p(1),$$

and $a$ has a Ramanujan expansion with coefficients

$$a_q = \frac{M(ac_q)}{\varphi(q)}.$$

These methods of computing Ramanujan expansions rely on the mean-value of the arithmetic function, but there are plenty of arithmetic functions for which the mean-value does not exist so we are unable to apply the above results.

A 2010 paper by Lucht[15] introduced a new concept that would greatly increase the number of functions for which we are able to compute a Ramanujan expansion. His idea is largely based on the fact that $c_q(n)$ is closely related to the Möbius function $\mu(q)$. His first result is useful in that not only can it help to determine the coefficients of a Ramanujan expansion, it can also be used to sum a Ramanujan expansion to find the arithmetic function it represents.

**Theorem.** Let $a(n)$ be an arithmetic function. If the series

$$\gamma(d) = d \sum_{n=1}^\infty a(dn) \mu(n)$$

converges for every $d \in \mathbb{N}$, then

$$(1 * \gamma)(n) = \sum_{q=1}^{\infty} a(q)c_q(n)$$

22
for every \( n \in \mathbb{N} \), where 1 denotes the constant function which takes the value 1 at each natural number.

An example where this theorem is applied will be given in chapter 4.

Lucht proceeds to use this result to prove an explicit formula for the coefficients of a Ramanujan expansion for certain additive arithmetic functions. Let \( \mathcal{F} \) be the unital commutative complex algebra of arithmetic functions with pointwise addition, Dirichlet convolution \(*\) and unity \( \epsilon \) defined by \( \epsilon(1) = 1 \) and \( \epsilon(n) = 0 \) for \( n > 1 \).

**Theorem.** Let \( g \in \mathcal{F} \) be additive. Suppose that \( \tilde{g}_p(1) \) converges for each prime \( p \), and that 
\[
\sum_p \tilde{g}_p(1),
\]
where the sum runs over primes \( p \), converges. Then \( g \) has a Ramanujan expansion with coefficients
\[
g_n = \begin{cases} 
-\frac{g(p^{k-1})}{p^{k-1}} + \left(1 - \frac{1}{p}\right) \tilde{g}_{p,k}(1) & \text{if } n = p^k, k \in \mathbb{N}, p \text{ prime}, \\
\sum_p g_p & \text{if } n = 1, \\
0 & \text{otherwise}.
\end{cases}
\]

Extending the mean-value results for multiplicative functions proved to be more difficult and we will require some more notation and definitions before we can state the result.

Let \( \mathcal{M} \) be the group of multiplicative arithmetic functions with Dirichlet convolution \(*\) and identity \( \epsilon \). For fixed \( k \in \mathbb{N} \cup \{0\} \) define the classes
\[
\mathcal{G}_k = \left\{ g \in \mathcal{M} : \sum_p \frac{|g(p)|^2}{p^2} \log^{2k} p < \infty, \sum_{p,v \geq 2} \frac{|g(p^v)|}{p^v} \log^k p^v < \infty \right\}
\]
\[
\mathcal{G}_k^* = \left\{ g \in \mathcal{G}_k : \tilde{g}_p(s) \neq 0 \text{ for all primes } p \text{ and } \Re(s) \geq 1 \right\}.
\]
We can show that $\mathcal{G}_k^*$ is a multiplicative group.

For a Dirichlet series
\[
\tilde{g}(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s}
\]
we denote the $j^{th}$ derivative of $\tilde{g}(s)$ by
\[
\tilde{g}^{(j)}(s) = (-1)^j \sum_{n=1}^{N} \frac{g(n)}{n^s} \log^j n
\]
for $\Re(s) \geq 1$. Finally for $g \in \mathcal{G}_k^*$ let $\gamma = \mu * g$, and $\Gamma = \gamma^{-1} = 1 * g^{-1}$. Define the class of multiplicative functions $\mathcal{R}_k$ to be all the $g \in \mathcal{G}_k^*$ such that each $\tilde{\Gamma}^{(j)}(1)$ converges for $j = 0, 1, \cdots, k$, and in addition they satisfy
\[
\tilde{\Gamma}^{(k)}(1) \neq 0 \text{ and } \tilde{\Gamma}^{(j)}(1) = 0 \text{ for } j = 0, 1, \cdots, k - 1.
\]

Using all these definitions and facts we can conclude the following result.

**Theorem.** Fix $k \in \mathbb{N} \cup \{0\}$. For every $g \in \mathcal{R}_k$, there exists a convergent Ramanujan expansion with coefficients
\[
g_n = \frac{(-1)^k}{\tilde{\Gamma}^{(k)}(1)} \log^k n \prod_{p \mid n} \frac{\tilde{\gamma}_p(1)}{\tilde{\gamma}_p(1)}
\]
where $\gamma = \mu * g$.

The results above of Lucht give us a much wider range of arithmetic functions for which we can find Ramanujan expansions.
Chapter 4

Some examples

In the same paper that introduced the infinite series expansions we have been studying, Ramanujan computed the expansions of some of the well known arithmetic functions used in number theory. He did this using some elementary properties of infinite series and simple algebraic manipulations. For the following results we follow the reasoning of Ramanujan[16].

Let $F(u, v)$ be any function of $u$ and $v$, and define

$$D(n) = \sum_{d|n} F \left( d, \frac{n}{d} \right). \quad (4.1)$$

Observe that by re-ordering the terms we also have

$$D(n) = \sum_{d|n} F \left( \frac{n}{d}, d \right). \quad (4.2)$$

Recall the function $\eta_q(n)$ defined by

$$\eta_q(n) = \sum_{k=1}^{q} \zeta_q^{kn},$$
which we determined is equal to $q$ if $q|n$ and 0 if $q \nmid n$. Then for $t \geq n$ it is easy to see that

$$D(n) = \sum_{1 \leq m \leq t} \frac{\eta_m(n)}{m} F \left( m, \frac{n}{m} \right).$$

We saw earlier that

$$\eta_q(n) = \sum_{d|n} c_q(n)$$

where $c_q(n)$ is Ramanujan’s sum. By substituting (3.4) into (3.3) we get

$$D(n) = \sum_{1 \leq m \leq t} \frac{1}{m} \sum_{q=1}^{m} c_q(n) F \left( m, \frac{n}{m} \right)$$

and by (3.2) we also know

$$D(n) = \sum_{1 \leq q \leq t} c_q(n) \sum_{1 \leq m \leq \frac{t}{q}} \frac{1}{q m} F \left( q m, \frac{n}{q m} \right).$$

Define

$$F_1(u, v) = F(u, v) \log u \text{ and } F_2(u, v) = F(u, v) \log v,$$

then

$$D(n) \log n = \sum_{d|n} F \left( d, \frac{n}{d} \right) \log \left( d \cdot \frac{n}{d} \right)$$

$$= \sum_{d|n} F \left( d, \frac{n}{d} \right) \log d + \sum_{d|n} F \left( \frac{n}{d}, d \right) \log \frac{n}{d}$$

$$= \sum_{d|n} F_1 \left( d, \frac{n}{d} \right) + \sum_{d|n} F_2 \left( \frac{n}{d}, d \right).$$

By substituting (3.5) for the first sum, and (3.6) for the second sum in (3.7), if $r, t \geq n$ we get
\[ D(n) \log n = \sum_{1 \leq q \leq r} c_q(n) \sum_{1 \leq m \leq \frac{r}{q}} \frac{\log q m}{q m} F \left( q m, \frac{n}{q m} \right) \]
\[ + \sum_{1 \leq q \leq t} c_q(n) \sum_{1 \leq m \leq \frac{t}{q}} \frac{\log q m}{q m} F \left( \frac{n}{q m}, q m \right) \quad (4.8) \]

and if in addition we have that \( F(m, n) = F(n, m) \) for all \( m, n \in \mathbb{N} \) then
\[ \frac{1}{2} D(n) \log n = \sum_{1 \leq q \leq t} c_q(n) \sum_{1 \leq m \leq \frac{t}{q}} \frac{\log q m}{q m} F \left( q m, \frac{n}{q m} \right). \quad (4.9) \]

Now we can use these general equations to prove the Ramanujan expansions for one class of arithmetic functions. Define
\[ \sigma_s(n) = \sum_{d|n} d^s \]
for \( s \geq 0 \) so that \( \sigma_1(n) \) is the familiar \( \sigma(n) \) function that gives the sum of the divisors of \( n \) and \( \sigma_0(n) = d(n) \), the number of divisors of \( n \). If we let \( F(u, v) = v^s \), then
\[ D(n) = \sum_{d|n} F \left( d, \frac{n}{d} \right) = \sum_{d|n} \left( \frac{n}{d} \right)^s = \sigma_s(n) \]
so if we substitute this into (3.6) we get
\[ \frac{\sigma_s(n)}{n^s} = \sum_{1 \leq q \leq t} c_q(n) \sum_{1 \leq m \leq \frac{t}{q}} \frac{1}{n^s \cdot q m} \left( \frac{n}{q m} \right)^s \]
\[ = \sum_{1 \leq q \leq t} c_q(n) \sum_{1 \leq m \leq \frac{t}{q}} \frac{1}{(q m)^{s+1}}. \quad (4.10) \]

Suppose now that \( s > 0 \), then
\[ \sum_{1 \leq m \leq \frac{t}{q}} \frac{1}{(q m)^{s+1}} = \sum_{m=1}^\infty \frac{1}{(q m)^{s+1}} + O \left( \frac{1}{q t^s} \right) \]
\[ = \zeta(s + 1) \frac{1}{q^{s+1}} + O \left( \frac{1}{q t^s} \right). \]
Recall property 6 from chapter 2, that \( c_q(n) \leq \sigma(n) \), so

\[
\frac{\sigma_s(n)}{n^s} = \sum_{1 \leq q \leq t} c_q(n) \sum_{1 \leq m \leq \frac{t}{q}} \frac{1}{(qm)^{s+1}}
\]

\[
= \sum_{1 \leq q \leq t} c_q(n) \left( \zeta(s + 1) \frac{1}{q^{s+1}} + O \left( \frac{1}{qt^s} \right) \right)
\]

\[
= \zeta(s + 1) \sum_{1 \leq q \leq t} \frac{c_q(n)}{q^{s+1}} + O \left( \sum_{1 \leq q \leq t} \frac{c_q(n)}{qt^s} \right)
\]

\[
= \zeta(s + 1) \sum_{1 \leq q \leq t} \frac{c_q(n)}{q^{s+1}} + O \left( \frac{1}{t^s} \sum_{1 \leq q \leq t} \frac{1}{q} \right)
\]

\[
= \zeta(s + 1) \sum_{1 \leq q \leq t} \frac{c_q(n)}{q^{s+1}} + O \left( \frac{\log t}{t^s} \right)
\]

where we used the fact that

\[
\sum_{1 \leq q \leq t} \frac{1}{q} \sim \log t.
\]

Finally if we let \( t \to \infty \) then

\[
\frac{\log t}{t^s} \to 0
\]

and we get our result

\[
\frac{\sigma_s(n)}{n^s} = \zeta(s + 1) \sum_{q=1}^{\infty} \frac{c_q(n)}{q^{s+1}} \quad (4.11)
\]

In particular we have

\[
\sigma(n) = \frac{n \pi^2}{6} \sum_{q=1}^{\infty} \frac{c_q(n)}{q^{s+1}}.
\]

Notice that

\[
\sigma_s(n) = \sum_{d|n} d^s
\]

\[= n^s \sum_{d|n} d^{-s}\]

\[= n^s \sigma_{-s}(n)\]
and so by (3.11),

$$\frac{\sigma_s(n)}{\zeta(s+1)} = \sum_{q=1}^{\infty} \frac{c_q(n)}{q^{s+1}}.$$

The finite Dirichlet series $\sigma_s(n)$ is clearly convergent and by Landau[14]

$$\lim_{s \to 0} \frac{1}{\zeta(s+1)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0.$$

From this we get the previously mentioned expansion

$$0 = \lim_{s \to 0} \frac{\sigma_s(n)}{\zeta(s+1)} = \sum_{q=1}^{\infty} \frac{c_q(n)}{q}.$$

In chapter 5 we will require the Ramanujan expansion for the function

$$\frac{\varphi(n)\Lambda(n)}{n},$$

where $\Lambda(n)$ is the von Mangoldt function defined as

$$\Lambda(n) = \begin{cases} 
\log p & \text{if } n = p^k \text{ for prime } p, k \in \mathbb{N}, \\
0 & \text{otherwise.}
\end{cases}$$

For this proof we follow the reasoning of Hardy[8]. Define

$$f(s) = \sum_{q=1}^{\infty} \frac{\mu(q)c_q(n)}{q^{s-1}\varphi(q)}$$

for $s \in \mathbb{C}$ with $\Re(s) > 1$. To analyze this we will use the Euler product identity for a multiplicative function. Suppose $a(n)$ is an arithmetic function such that $a(mn) = a(m)a(n)$ whenever $(m, n) = 1$. Then

$$\sum_{n=1}^{N} a(n) = \prod_{p} \sum_{k=0}^{\infty} a(p^k).$$
where the product runs over all primes $p$. If in addition we have that $a(p^k) = 0$ for $k \geq 2$ for all primes $p$, then (3.14) reduces to

$$
\sum_{n=1}^{N} a(n) = \prod_{p} (a(1) + a(p)) = \prod_{p} (1 + a(p)).
$$

Observe that

$$
\frac{\mu(q)c_q(n)}{q^{s-1}\varphi(q)} = \prod_{p} \left( 1 - \frac{c_p(n)}{p^{s-1}\varphi(p)} \right)
$$

is multiplicative and $\mu(p^k) = 0$ for all primes $p$ and $k \geq 2$, then

$$
f(s) = \prod_{p} \left( 1 + \frac{\mu(p)c_p(n)}{p^{s-1}\varphi(p)} \right) = \prod_{p} \left( 1 - \frac{c_p(n)}{p^{s-1}\varphi(p)} \right) = \prod_{p|n} \left( 1 - \frac{1}{p^{s-1}} \right) \prod_{p|n} \left( 1 + \frac{1}{p^{s-1}(p-1)} \right) = \prod_{p|n} \left( \frac{(p-1)(p^{s-1}-1)}{p^s - p^{s-1} + 1} \right) \prod_{p} \left( 1 + \frac{1}{p^{s-1}(p-1)} \right) = g(s)h(s),
$$

say, where we have used property 4 of $c_q(n)$, that

$$
c_p(n) = \begin{cases} 
  p - 1 & \text{if } p \mid n, \\
  -1 & \text{if } p \nmid n.
\end{cases}
$$

for primes $p$. Now suppose $n$ is a prime power, say $n = p^k$ then

$$
g(s) = \frac{(p-1)(p^{s-1}-1)}{p^s - p^{s-1} + 1}
$$

has a simple zero at $s = 1$ and

$$
g'(1) = \frac{p - 1}{p} \log p.
$$
If $n$ is not a prime power, $g(s)$ must have a zero or order greater than 1.

Consider
\[
\frac{h(s)}{\zeta(s)} = \prod_p \left( 1 + \frac{1}{p^{s-1}(p-1)} \right) \left( 1 - \frac{1}{p^s} \right)
\]
\[
= \prod_p \left( 1 + \frac{1}{p^{s-1}(p-1)} - \frac{1}{p^s} - \frac{1}{p^{2s-1}(p-1)} \right)
\]
\[
= \prod_p \left( 1 + O(p^{-\sigma-1}) \right)
\]
where $s = \sigma + it$. This product is uniformly convergent for $\sigma > 0$ and thus
\[
\frac{h(s)}{\zeta(s)}
\]
is holomorphic in the half plane $\Re(s) > 0$. Since $\zeta(s)$ has a simple pole at $s = 1$ with residue 1, so must $h(s)$.

Combining these results tell us that $f(s)$ is holomorphic at $s = 1$ and that
\[
f(1) = \frac{p-1}{p} \log p
\]
when $n$ is a prime power and $f(1) = 0$ otherwise. By using methods similar to those also used by Landau to show that
\[
\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0,
\]
the convergence of the series (3.13) can be established for $s = 1$ and so we have
\[
\sum_{q=1}^{\infty} \frac{\mu(q)c_q(n)}{\varphi(q)} = \frac{\varphi(n)\Lambda(n)}{n}.
\]

Finally, we will use a method from chapter 3 to prove the Ramanujan expansion for the divisor function $d(n)$ which counts the number of positive divisors of $n$. Recall the theorem:
Theorem. Let \( \hat{g}(n) \) be an arithmetic function. If the series

\[
\gamma(d) = d \sum_{n=1}^{\infty} \hat{g}(dn) \mu(n)
\]

converges for every \( d \in \mathbb{N} \), then

\[
(1 \ast \gamma)(n) = \sum_{q=1}^{\infty} \hat{g}(q)c_q(n)
\]

for every \( n \in \mathbb{N} \).

If we choose \( \hat{g}(n) = \frac{\log n}{n} \) then

\[
\gamma(d) = \sum_{n=1}^{\infty} \frac{\mu(n) \log(nd)}{nd} = \sum_{n=1}^{\infty} \frac{\mu(n) \log(n)}{n} + (\log d) \sum_{n=1}^{\infty} \frac{\mu(n)}{n}
\]

\[
= \sum_{n=1}^{\infty} \frac{\mu(n) \log n}{n} = -1,
\]

the last equality being another result of Landau. But

\[
(1 \ast -1)(n) = \sum_{d \mid n} 1 \cdot (-1) = -d(n),
\]

so we have

\[
-d(n) = \sum_{q=1}^{\infty} \frac{\log q}{q} c_q(n).
\]
Although this result was known to Ramanujan, note that the divisor function $d(n)$ has no mean value and thus the Ramanujan expansion could not have been found using the mean value methods.
The twin prime conjecture and the Wiener-Khinchin formula

The twin prime conjecture is a well known and long standing problem about the distribution of pairs of primes. In particular, it asks whether there exists an infinite number of primes $p$ for which $p + 2$ is also prime. The number 2 is not special in this question, and it makes just as much sense to ask if there are infinitely many primes $p$ for which $p + h$ is also prime for $h$ any even positive integer. In 1849, de Polignac[5] asked about this natural generalization which has now become known as de Polignac’s conjecture.

A stronger version of de Polignac’s conjecture was conjectured in 1922 by Hardy and Littlewood[9]. Using heuristic methods they were led to the following question.

**Conjecture.** Define the function

$$\pi_h(x) = \# \{ p \in \mathbb{N} : p \leq x, p \text{ and } p + h \text{ are prime.} \}$$
CHAPTER 5. THE TWIN PRIME CONJECTURE AND THE WIENER-KHINCHIN FORMULA

for \( x \in \mathbb{R} \) and \( h \) an even positive integer. Then

\[
\pi_h(x) \sim C_h \frac{x}{(\log x)^2}
\]

where

\[
C_h = \left( \prod_{p|h} \left( 1 + \frac{1}{p-1} \right) \prod_{p \nmid h} \left( 1 - \frac{1}{(p-1)^2} \right) \right)
\]

with the products running over primes \( p \). In particular,

\[
\pi_h(x) \to \infty \text{ as } x \to \infty
\]

We can see that the conjecture of Hardy and Littlewood is just a stronger version of de Polignac’s conjecture so a positive result would settle the twin prime conjecture and its generalizations.

The content of this chapter is another method of coming to the same conjecture without having to use the much more technical circle method that was employed by Hardy and Littlewood. Gadiyar and Padma\cite{7} gained inspiration for this line of reasoning from a theorem used primarily in physics. The Wiener-Khinchin formula is a statement that relates the autocorrelation function of a wide-sense stationary random process with the power spectrum of the process. More detailed information about the Wiener-Khinchin formula and the other notions mentioned, see \cite{4}. For our purposes, it is sufficient to consider a much more simplified version.

The Wiener-Khinchin formula tells us that if \( f(t) \) has Fourier expansion

\[
f(t) = \sum_{n \in \mathbb{Z}} f_n e^{int},
\]

then

\[
\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} f(t + \tau) \overline{f(t)} dt = \sum_{n \in \mathbb{Z}} |f_n|^2 e^{int}.
\]
To apply this to our situation we consider the following analogue. Let \( a(n) \) be an arithmetic function with Ramanujan expansion

\[
a(n) = \sum_{q=1}^{\infty} a_q c_q(n),
\]

then we would hope that

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{\infty} a(n)a(n + h) = \sum_{q=1}^{\infty} a_q^2 c_q(h).
\]

Recall from chapter 4 the Ramanujan expansion

\[
\frac{\varphi(n) \Lambda(n)}{n} = \sum_{q=1}^{\infty} \frac{\mu(q)}{\varphi(n)} c_q(n).
\]

If we assume that we may apply the Wiener-Khinchin formula to this function we get

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \frac{\varphi(n) \Lambda(n) \varphi(n + h) \Lambda(n + h)}{n \ n + h} = \sum_{q=1}^{\infty} \frac{\mu^2(q)}{\varphi^2(q)} c_q(h).
\]

An issue arises though with the application of the Wiener-Khinchin formula above. The exchange of limits happening behind the scenes of the formula cannot easily be justified, and indeed this is the only remaining gap in the proof. We can see that for an arithmetic function \( a(n) \) with Ramanujan expansion

\[
a(n) = \sum_{q=1}^{\infty} a_q c_q(n),
\]

if we ignore any issues of convergence we can prove our version of the Wiener-Khinchin
CHAPTER 5.  THE TWIN PRIME CONJECTURE AND THE WIENER-KHINCHIN FORMULA

formula:
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{\infty} a(n)a(n+h)
\]

\[
= \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \sum_{q=1}^{\infty} a_q c_q(n) \sum_{r=1}^{\infty} a_r c_r(n+h)
\]

\[
= \sum_{q=1}^{\infty} \sum_{r=1}^{\infty} a_q a_r \left( \frac{1}{N} \lim_{N \to \infty} \sum_{n=1}^{\infty} c_q(n)c_r(n+h) \right)
\]

\[
= \sum_{q=1}^{\infty} a_q^2 c_q(h)
\]

where we have used the “orthogonality” property of \(c_q(n)\) from chapter 2. Unfortunately, the Ramanujan expansion for the function

\[
\frac{\varphi(n)\Lambda(n)}{n}
\]

is not absolutely convergent and so the above computation is not valid for this situation. It is still fruitful though to continue on under the assumption that the Wiener-Khinchin formula is justified so we will do so.

We wish to analyze more closely the sum we obtained after applying the formula. Observe that

\[
\frac{\mu^2(q)}{\varphi^2(q)} c_q(n)
\]

is multiplicative since each of \(\mu(q), \varphi(q)\) and \(c_q(n)\) are multiplicative, and \(\mu(p^k) = 0\) for \(k \geq 2\) and all primes \(p\), so our function also vanishes at prime powers. Therefore we can apply the Euler product identity

\[
\sum_{q=1}^{\infty} \frac{\mu^2(q)}{\varphi^2(q)} c_q(n) = \prod_p \left( 1 + \frac{\mu^2(p)}{\varphi(p)} c_q(n) \right).
\]

Now we know that \(\mu^2(p) = 1\) and \(\varphi(p) = p - 1\) for all primes \(p\), and using property 4
from chapter 2, namely
\[ c_p(n) = \begin{cases} 
  p - 1 & \text{if } p \mid n, \\
  -1 & \text{if } p \nmid n 
\end{cases} \]
for prime \( p \), we get
\[
\prod_p \left( 1 + \frac{\mu^2(p)}{\varphi(p)} c_q(h) \right) \\
= \prod_{p \mid h} \left( 1 + \frac{p - 1}{(p - 1)^2} \right) \prod_{p \nmid h} \left( 1 + \frac{-1}{(p - 1)^2} \right) \\
= \prod_{p \mid h} \left( 1 + \frac{1}{p - 1} \right) \prod_{p \nmid h} \left( 1 - \frac{1}{(p - 1)^2} \right) = C_h,
\]
which amazingly is Hardy and Littlewood’s conjectured prime pair constant. It follows that
\[
\sum_{n=1}^{N} \frac{\varphi(n)\Lambda(n)\varphi(n+h)\Lambda(n+h)}{n(n+h)} \sim C_h N \text{ as } N \to \infty.
\]
Notice that the terms in the sum on the left are non zero exactly when \( n \) and \( n + h \) are prime powers, say \( n = p^k \) and \( n + h = q^l \). Then
\[
\frac{\varphi(n)}{n} = \frac{p^{k-1}(p - 1)}{p^k} = \frac{p - 1}{p},
\]
and
\[
\frac{\varphi(n+h)}{n+h} = \frac{q^{l-1}(q - 1)}{q^l} = \frac{q - 1}{q},
\]
and we also have
\[
\Lambda(n) = \log p \text{ and } \Lambda(n+h) = \log q.
\]
CHAPTER 5. THE TWIN PRIME CONJECTURE AND THE WIENER-KHINCHIN FORMULA

We can use these facts to simplify our estimate for $C_h N$. We compute

$$
\sum_{n \leq N} \Lambda(n) \Lambda(n + h) - \sum_{n \leq N} \frac{\varphi(n) \Lambda(n) \varphi(n + h) \Lambda(n + h)}{n + h}
$$

$$
= \sum_{p^k \leq N \atop p^k + h = q^l} \log p \log q - \sum_{p^k \leq N \atop p^k + h = q^l} \frac{(p - 1)(q - 1) \log p \log q}{pq}
$$

$$
= \sum_{p^k \leq N \atop p^k + h = q^l} \frac{\log p \log q}{p} + \sum_{p^k \leq N \atop p^k + h = q^l} \frac{\log p \log q}{q} - \sum_{p^k \leq N \atop p^k + h = q^l} \frac{\log p \log q}{pq}
$$

$$
= O(\log^2 N \log \log N),
$$

so we have reduced our estimate to

$$
\sum_{n \leq N} \Lambda(n) \Lambda(n + h) \sim C_h N.
$$

Observe now that

$$
\sum_{n \leq N} \Lambda(n) \Lambda(n + h)
$$

$$
= \sum_{p \leq N \atop p + h \text{ prime}} \log p \log(p + h) + \sum_{p^k \leq N \atop k \geq 2} \Lambda(p^k) \Lambda(p^k + h) + \sum_{p \leq N \atop p + h = q^l, l \geq 2} \log p \log(p + h)
$$

$$
\leq \sum_{p \leq N \atop p + h \text{ prime}} \log p \log(p + h) + 2 \log^2 N (N^{1/2} + N^{1/3} + \cdots + N^{1/\Theta(\log N)})
$$

$$
= \sum_{p \leq N \atop p + h \text{ prime}} \log p \log(p + h) + O(\sqrt{N} \log^3 N),
$$

thus we have

$$
\sum_{p \leq N \atop p + h \text{ prime}} \log p \log(p + h) \sim C_h N.
$$

We simplify this further by using the fact that

$$
\sum_{p \leq N} \frac{\log p}{p} \sim \log N
$$
which can be found in Apostol[1], theorem 4.9, and that

$$\log(1 + x) = x + O(x^2)$$

for $|x| < 1$.

We can see that

$$\sum_{p \leq N \text{ prime}} \log p \log (n + h) = \sum_{p \leq N \text{ prime}} \log p \left( \log \left( 1 + \frac{h}{p} \right) + \log p \right)$$

$$= \sum_{p \leq N \text{ prime}} \log^2 p + \sum_{p \leq N \text{ prime}} \log p \log \left( 1 + \frac{h}{p} \right)$$

$$= \sum_{p \leq N \text{ prime}} \log^2 p + \sum_{p \leq N \text{ prime}} \left( \frac{h}{p} + O \left( \frac{h^2}{p^2} \right) \right) \log p$$

$$= \sum_{p \leq N \text{ prime}} \log^2 p + \left( h + O \left( \frac{h^2}{p} \right) \right) \sum_{p \leq N \text{ prime}} \frac{\log p}{p}$$

$$= \sum_{p \leq N \text{ prime}} \log^2 p + O(\log N)$$

so we end up with

$$\sum_{p \leq N \text{ prime}} \log^2 p \sim C_h N.$$

To continue on we require the technique of partial summation, generally attributed to Abel. We will state the version that we will be using as a lemma.

**Lemma 3.** Fix $a \geq 1$. Let $\{c_m\}_{m=1}^t$ be a sequence of complex numbers,

$$C(x) = \sum_{a \leq m \leq x} c_m,$$

and $f(x)$ a continuously differentiable function on $[a, v]$. Then

$$\sum_{a \leq m \leq v} c_m f(m) = C(x)f(x) - \int_a^v C(x)f'(x)dx.$$
Now for each \( m \in \mathbb{N} \) put
\[
c_m = \begin{cases} 
1 & \text{if } m, m + h \text{ are prime} \\
0 & \text{otherwise} 
\end{cases}
\]
so
\[
C(x) = \sum_{1 \leq m \leq x} c_m = \pi_h(x)
\]
and
\[
f(x) = \log^2 x.
\]
Then
\[
\sum_{p \leq N \atop p + h \text{ prime}} \log^2 p = \log^2 N \pi_h(N) - 2 \int_2^N \frac{\pi_h(x) \log x}{x} \, dx.
\]
It can be shown, see for example Bateman and Diamond[2], that there exists some \( C > 0 \) such that
\[
\pi_h(x) \leq \frac{Cx}{\log^2 x}
\]
and therefore
\[
\sum_{p \leq N \atop p + h \text{ prime}} \log^2 p = \log^2 N \pi_h(N) + O\left( \int_2^N \frac{dx}{\log x} \right)
\]
\[
= \log^2 N \pi_h(N) + O\left( \frac{N}{\log N} \right).
\]
It follows that
\[
\sum_{p \leq N \atop p + h \text{ prime}} \log^2 p \sim \log^2 N \pi_h(N)
\]
so that
\[
\pi_h(N) \sim C_h \frac{N}{\log^2 N}
\]
which is precisely the same result as Hardy and Littlewood.

By assuming the justification of an interchange of limits we arrived at a conjecture about the twin prime problem. The formula that we used was

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n \leq N} a(n)a(n + h) = \sum_{q=1}^{\infty} a_q^2 c_q(h)$$

for an arithmetic function $a(n)$ with Ramanujan expansion

$$a(n) = \sum_{q=1}^{\infty} a_q c_q(n).$$

This exchange of limits is well justified in the case that the Ramanujan expansion for $a(n)$ is absolutely and uniformly convergent, but it is often the case that we only have pointwise convergence. The question we ask now is: Is there enough evidence for us to assume the validity of the exchange of limits in the arithmetic analogue of the Wiener-Khinchin formula when we don’t have absolute and uniform convergence?

The result about the twin prime problem gives us some evidence toward the validity of the formula since it matches the conjecture made by Hardy and Littlewood in 1922, but this is still strictly conjectural. Can we find any evidence to the contrary?

By using an example that we have seen before, namely that

$$0 = \sum_{q=1}^{\infty} \frac{c_q(n)}{q},$$

we can find some evidence against the validity of the arithmetic Wiener-Khinchin formula. Indeed, if we let $a(n) = 0$ for all $n \in \mathbb{N}$ and blindly apply the formula we would end up with

$$0 = \lim_{N \to \infty} \frac{1}{N} \sum_{n \leq N} a(n)a(n + h)$$

$$= \sum_{q=1}^{\infty} \frac{c_q(h)}{q^2}$$

$$= \frac{6\sigma(h)}{\pi^2 h}$$
for all $h \in \mathbb{N}$ which is obviously absurd.

From this we can conclude that we should be skeptical about the application of this formula to a Ramanujan expansion that is not absolutely and uniformly convergent.
Chapter 6

Computing values of arithmetic functions

The computation of the values of some arithmetic functions can sometimes be very difficult. For example, computing \( \sigma(n) \), the sum of the positive divisors of \( n \), may require factoring \( n \) into primes which is a difficult computational problem for large values of \( n \).

By using Ramanujan expansions of a certain class of these functions, we can compute their values exactly and possibly improve on the efficiency of their computation. The idea stems from the observation that many of the useful arithmetic functions are integer valued. Once we have computed the values up to a close enough accuracy, we can round to the nearest integer and we have the correct value of the function.

The task then becomes to approximate the tail end of the Ramanujan expansion so that we know an upper bound to how many terms we must sum up so that we have an accurate enough estimate. If for an arithmetic function \( a(n) \) with Ramanujan
expansion

\[ a(n) = \sum_{q=1}^{\infty} a_q c_q(n) \]

we write,

\[ a(n) = \sum_{q=1}^{\infty} a_q c_q(n) = \sum_{q \leq t} a_q c_q(n) + \sum_{q \geq t} a_q c_q(n) \]

then by choosing the value of \( t \) for which

\[ \left| \sum_{q \geq t} a_q c_q(n) \right| < \frac{1}{2} \]

we know that the sum up to the \( t^{th} \) term will be sufficiently close to the real value of the function that we can round to the nearest integer and have the correct result.

When we wish to compute the values of an integer valued arithmetic function \( a(n) \) for which

\[ \sum_{q=1}^{\infty} a_q \]

is convergent, we have a crude estimate that can always be used. By using the upper bound

\[ c_q(n) \leq \sigma(n) \]

we get that

\[ \sum_{q \geq t} a_q c_q(n) \leq \sigma(n) \sum_{q \geq t} a_q \]

If we choose an upper bound \( A(t) \) so that

\[ \sum_{q \geq t} a_q \leq A(t) \]
and observe the bound
\[ \sigma(n) = \sum_{d \mid n} d \]
\[ = n \sum_{d \mid n} \frac{1}{d} \]
\[ \leq n \log n, \]

we finally have
\[ \sum_{q \geq t} a_q c_q(n) \leq A(t)n \log n. \]

It remains only to rearrange the following inequality to get a bound for \( t \):
\[ A(t)n \log n < \frac{1}{2}. \]

As an example of this, we compute a value of \( t \) sufficiently large enough to give us the value of \( \sigma(n) \) itself. We have seen that
\[ \sigma(n) = \frac{\pi^2 n}{6} \sum_{q=1}^{\infty} \frac{c_q(n)}{q^2}. \]

and we will use the estimate
\[ \sum_{q \geq t} \frac{1}{q^2} \leq \frac{1}{t - 1} \]

for \( t > 1 \). Combining these we end up with
\[ \frac{\pi^2 n}{6} \sum_{q \geq t} \frac{c_q(n)}{q^2} \leq \frac{\pi^2 n^2 \log n}{6(t - 1)}. \]

We set this less than \( 1/2 \) then rearrange to find a lower bound for \( t \):
\[ \frac{\pi^2 n^2 \log n}{6(t - 1)} < \frac{1}{2} \]
\[ \frac{\pi^2 n^2 \log n}{3} + 1 < t. \]

Choosing \( t \) in this way will then force
\[ \sigma(n) = \left\lfloor \frac{\pi^2 n}{6} \sum_{q \leq t} \frac{c_q(n)}{q^2} \right\rfloor, \]
where $[x]$ is the nearest integer function for $x \in \mathbb{R}$ and $x \notin \mathbb{Z} + 1/2$. This formula is interesting enough in its own right, but it is possible that some practical application could come of it. It seems as though that $t$ grows too quickly as $n$ increases, but it is conceivable that if a large table of values for $c_q(n)$ was kept that pulling values and summing up the finite series could lead to a more efficient way to calculate $\sigma(n)$.

For a slightly different value of $t$, we could use the approximation

\[
c_q(n) \leq \sigma(n) = \sum_{d|n} d
\]

\[
\leq n \sum_{d|n} 1
\]

\[
= nd(n),
\]

where $d(n)$ is the number of divisors of $n$, so that we can choose

\[
t > \frac{\pi^2 n^2 d(n)}{3} + 1.
\]
Bibliography


