Comparison of the circle method applied to the Goldbach Problem and the Restricted Digit Problem

by

Mariken Weijs

a bachelor thesis submitted to the Department of Mathematics
at Utrecht University for the degree of

Bachelor in Mathematics

Supervisor: Damaris Schindler

June 15, 2017

Utrecht University

Abstract

The Hardy-Littlewood circle method is a widely used tool in the field of analytic number theory. James Maynard [21] uses it to prove that there are infinitely many primes without a certain fixed digit in their decimal expansion. His application however is slightly different from the original approach. In this thesis the parallels and differences are discussed between the original circle method applied to the Ternary Goldbach Problem and the modified circle method applied to the Restricted Digit Problem. It is quite interesting that we can solve the Restricted Digit Problem, which is a binary problem, with the circle method. After all the Binary version of the Golbach Problem can not be solved with it.

Contents

Chapter 1

Introduction

Since this thesis is a comparison between two applications of the circle method, we look into both the problems it is applied to and then give an overview of the application of the circle method to them. First the Ternary Goldbach Problem is discussed and afterwards we follow the same steps but then for the Restricted Digit Problem. We expect this makes it easier for the reader to compare both applications of the circle method.

In Chapter 2 you can find an introduction to the Goldbach Problem. Both the Binary and the Ternary version of the problem are posed there. Furthermore the developments throughout history are discussed.
In the successive chapter the circle method is explained applied to the Ternery Goldbach Problem. We can distinguish three steps in the circle method. The first is to find an expression for a sequence we are interested in. This step is discussed in detail in section 3.1. The next step is to define major and minor arcs, which we perform in section 3.2. The last step is to find estimates for both the major and minor arcs in section 3.3. We discuss an outline of the proof for sufficiently large n, but skip a lot of details and parts of the proof that are too involved. We discuss the minor arcs a bit more in detail than the major arcs because the minor arcs is the part that makes the difference between the Binary and Ternary Goldbach Problem. After all, for the Ternary Goldbach Problem we can find a usefull estimate, but for the Binary Goldbach Problem nobody managed to do this untill now.
Chapter 4 is dedicated to the Restriced Digit Problem. The problem is introduced and the set $\mathcal{A}_k$ is defined, which plays a crucial role in the proof Maynard provides. Just as with the Goldbach Problem, some earlier works from other mathematicians is discussed to give some context. In the last chapter we give an overview of the application of the circle method. We look into the similarities between this application and the application in Chapter 3. This is however not very detailed, since the proof is quite complex.

1.1 Notation and assumptions

Throughout this thesis we assume X to be an integral power of 10 larger than 1 always when we encounter X. The reason to assume this, is that the numbers in between the integral powers of ten lead to a lot of technical difficulties that are irrelevant. The only exception is figure 4.1, where X takes al integral values between 0 and 70.
Furthermore, we choose the constant $k \in \{0, 1, \ldots, 9\}$ to be fixed.
When ϵ is used, this denotes a very small positive constant.

Throughout this thesis, the notation $e(x)$ is used for $e^{2\pi i x}$.
With $\mathbb{P} = \{2, 3, 5, 7, ...\}$ we denote the set of prime numbers.
Since the greatest common divisor is used multiple times, we write $(a, b) = c$ as shorthand notation for $\gcd(a, b) = c$.
For functions f and g, where g takes non-negative real values, we use $f \ll g$ to indicate that there exist constants $c \in \mathbb{R}_{>0}$ and $x_0 \in \mathbb{R}$, such that $|f(x)| < c \cdot g(x)$ for all $x > x_0$.
Analogously, we use $f \gg g$ to indicate that there exist constants $c \in \mathbb{R}_{>0}$ and $x_0 \in \mathbb{R}$, such that $|f(x)| > c \cdot g(x)$ for all $x > x_0$.
For two real valued functions f and g, we use the notation $f(x) = O(g(x))$ to state that there exist constants $c \in \mathbb{R}_{>0}$ and $x_0 \in \mathbb{R}$ such that $|f(x)| < c \cdot |g(x)|$ for all $x > x_0$.
We also use the stronger notion $f(x) = o(g(x))$. With this we indicate that for all $\epsilon > 0$ there exists an $x_0 \in \mathbb{R}$ such that $|f(x)| \leq \epsilon \cdot |g(x)|$ for all $x > x_0$.
For all these notations for asymptotic behaviour, a subscript η indicates that the constant c depends on a constant η.
We write ${}^{a}\log b = r$ if $b^r = a$. When $\log x$ is used without specification of the base, it is the natural logarithm, so ${}^{e}\log x$.

Chapter 2

The Goldbach Problem

In this chapter we take a look at an introduction to the Goldbach Problem and give an overview of the most significant results on it so far. For this, [19], Chapter 5 served as a basis. Many of the referenced works are mentioned there, though other relevant works are added that are not discussed in the book.

Back in 1742 Goldbach already posed the following conjecture.

Conjecture 1 (Weak Goldbach Conjecture)
For all odd $N > 5$ there exist primes p_1, p_2, p_3 such that $N = p_1 + p_2 + p_3$.

This became known as the Weak Goldbach Conjecture or the Ternary Goldbach Problem. Euler thought this conjecture could even be formulated stronger. He posed the stronger Conjecture 2.

Conjecture 2 (Strong Goldbach Conjecture)
For all even $N > 4$ there exist primes p_1, p_2 such that $N = p_1 + p_2$.

This conjecture is now called the Strong Goldbach Conjecture or the Binary Goldbach Problem. For if Conjecture 2 would be proven, Conjecture 1 follows from it. After all, since Conjecture 1 states that every even number $N > 4$ is the sum of two primes, adding 3 to it gives all odd numbers $M > 7$ written as the sum of three primes. We already know that Conjecture 1 holds for $N = 7$, since $7 = 2 + 2 + 3$. So the Weak Goldbach Conjecture follows.

For centuries both conjectures remained unproven. In 1922 Hardy and Littlewood [15] published a paper with their work on the problem so far. They managed to prove Conjecture 1, for sufficiently large odd numbers N, under condition of a generalization of the Riemann Hypothesis. They do not give a specific N_0 such that the Ternary Goldbach Problem holds for all $N > N_0$.
We recall that the Riemann Hypothesis states that the only positive zeros of the Riemann ζ-function can be found at complex numbers with real part $\frac{1}{2}$. The Riemann ζ-function is one particular Dirichlet L-function and it turns out that the Riemann Hypothesis can be generalized for all Dirichlet L-functions. This is the generalization Hardy and

Littlewood assumed.
Hardy and Littlewood used the so called circle method for their proof. Hardy and Ramanujan [18] invented the original circle method, which is explained in general in Chapter 3.

It took until 1937 to prove Conjecture 1 without assumption of the Generalized Riemann Hypothesis. This proof still only holds for sufficiently large odd values of N. Back then, it was Vinogradov [25] who proved it. Just ast Hardy and Littlewood, he did not give an explicit bound for "sufficiently large". For his result Vinogradov modified the method used by Hardy and Littlewood.

Although there was some progress in the years in between, only in 2012 Terence Tao proved that every odd number greater than 1 is the sum of at most five primes. Shortly afterwards the Weak Goldbach Conjecture was proven without lower bound and without assumption of the Generalized Riemann Hypothesis by Harold Helfgott [12], [13] en [14]. He used much of the work of Tao [23] for this. Helfgott accomplished his results by lowering the lower bound as much as possible. Thereafter he used an efficient computational method to check all the remaining cases. For the computational method, he used [17], which we wrote togeter with David Platt.

Conjecture 2 on the other hand is still not proven. This problem is now often called the Goldbach Problem. So when the "Goldbach Problem" is mentioned further on, it refers to Conjecture 2. In 1937/1938 van der Corput [7], Tchudakov [24] and Estermann [10] proved independently that Conjecture 2 holds for almost all even numbers N. If we define $P(N)$ as $\#\{n \leq N : n \text{ even }, n = p_1 + p_2 \text{ with } (p_1, p_2) \in \mathbb{P}\}$, then "almost all" means that the limit $\lim_{n\to\infty} \frac{P(N)}{N} = 1$.

Later, Chen [5] showed that every sufficiently large even number can be written as the sum of a prime and the product of at most two primes. For this prove he used seive methods. Just as Vinogradov, he did not give an explicit bound for "sufficiently large". In 2015 however, Yamada [27] did. He proved that every even number greater than $e^{e^{36}} \approx 1,7 \cdot 10^{1872344071119348}$ can be written as the sum of two primes or a prime and a product of two primes.

Chapter 3

Circle method applied to the Ternary Goldbach Problem

Because the aim of this thesis is to compare two applications of the circle method, we take a closer look at this method. We now explain the circle method applied to the Ternary Goldbach Problem stated in Conjecture 1.

We use the circle method to prove asymptotic behaviour of a sequence. We define

$$A_n = \left\{(p_1, p_2, p_3) \in \mathbb{P}^3 : p_1 + p_2 + p_3 = n\right\}. \tag{3.1}$$

Now we define the sequence

$$a_n = \#A_n, \tag{3.2}$$

which is the sequence we want to study. So a_n is the amount of possible ways to write n as the sum of three primes. The order is relevant, so for example let $a \neq b \neq c$, then (a, b, c), with $a+b+c = n$ and (a, c, b), with $a+c+b = n$ count as two different elements of A_n. If we are able to show that $a_n \geq 1$ for all odd $n > 5$, we know that every number n can be written as the sum of three primes. This proves the Ternary Goldbach Problem. To determine asymptotic behaviour of a_n, we carry out the following three steps.

1. Find an expression as Fourier integral to describe a_n
2. Define major and minor arcs
3. Estimate the major and minor arcs

Combination of the estimates we found, gives us information about the asymptotic behaviour for a_n if n goes to infinity.

3.1 Find an expression to describe a_n

In step 1 there is a difference in the approach of Hardy and Littlewood on the one hand and Vinogradov's on the other hand. In the next subsection, 3.1.1, we follow Hardy and Littlewood's original line of attack, for which [6] and [26] are used as a basis for the understanding of the circle method. In that subsection we discuss the Ternary Goldbach Problem however, instead of Waring's Problem. In subsection 3.1.2 we thereafter explain how Vinogradov approached the problem, using [3] and [4]. Vinogradov's approach is commonly used nowadays. This line of attack is however still referred to as "The Hardy-Littlewood circle method".

3.1.1 Hardy and Littlewood's approach

We are looking for an expression to describe a_n (3.2) as a Fourier integral. With that in mind, we take a look at the generating function for $\mathbb{P}$,

$$G(y) = \sum_{p \in \mathbb{P}} y^p = y^2 + y^3 + y^5 + y^7 + \ldots.$$

Before we use this series to formulate an expression for a_n, we want to make sure it is convergent. We know that for the geometric series

$$S(y) = \sum_{n=0}^{\infty} y^n = \lim_{n \to \infty} \frac{1 - y^n}{1 - y} = \frac{1}{1 - y}$$

holds if $|y| < 1$. In other words, the radius of convergence of $S(y)$ is 1. Since $\mathbb{P} \subsetneq \mathbb{N}$, we know that $G(y) \leq S(y)$ for all y, with $|y| < 1$. Thus we know that $G(y)$ converges on the unit disk $D_1 = \{y \in \mathbb{C} : |y| < 1\}$, just as $S(y)$. If $y \geq 1$ every term is bigger than or equal to 1 and therefore $G(y)$ is not convergent anymore. So we can conclude that the radius of convergence is 1. Next, we take the third power of $G(y)$. This gives

$$\begin{aligned}(G(y))^3 &= \left(\sum_{p_1 \in \mathbb{P}} y^{p_1}\right) \cdot \left(\sum_{p_2 \in \mathbb{P}} y^{p_2}\right) \cdot \left(\sum_{p_3 \in \mathbb{P}} y^{p_3}\right) \\ &= y^2y^2y^2 + y^3y^2y^2 + y^2y^3y^2 + y^2y^2y^3 + y^3y^3y^2 + \ldots \\ &= y^6 + 3y^7 + \ldots \\ &= \sum_{j=6}^{\infty} a_j y^j,\end{aligned} \tag{3.3}$$

where a_j is again the number of ways to write j as the sum of three primes. We know that the product of two convergent power series with radius of convergence $r > 0$ is again a convergent power series with radius of convergence r. Using this, we can conclude that the radius of convergence of $(G(y))^3$ is 1. Because of this, the singularities lie on the unit circle. Since we are looking for asymptotic behaviour of a_n when n goes to infinity, we need only to examine cases where $(G(y))^3$ is convergent. We thus choose y to be an element of the unit disk D_1. Since we are composing an expression for a_n, we are only interested in primes smaller than n. We can therefore use n as an upper bound for the summation in $G(y)$. We can thus choose $N > n$ such that it functions as upper bound for the single summation in $(G(y))^3$. Let us now define $\tilde{G}(y)$ as $(G(y))^3$ with N as upper bound for its summation. Let $b_k = a_{k+n+1}$. For $n > 0$, we now define

$$F(y) = \sum_{k=6-n-1}^{N-n-1} b_k y^k = \tilde{G}(y) \cdot y^{-n-1}$$
$$= \sum_{j=6}^{N} a_j y^{j-n-1}.$$

Since $b_k = a_{k+n+1}$, choosing $k = -1$ gives us $b_{-1} = a_{-1+n+1} = a_n$. Because we now have a finite summation, it converges outside the unit circle as well. The multiplication with y^{-n-1} creates however a singularity in 0. We therefore know that $F(y)$ converges on $D_1 \backslash \{0\}$.
Since $F(y)$ is a power series centered at 0, b_{-1} is the residue of F about 0, we know how to compute it using Cauchy's Residue Theorem, stated below.

Theorem 1 (Cauchy's Residue Theorem) *([20], Chapter IV, p. 173, Theorem 1.1)*
Let z_0 be an isolated singularity of f, and let C be a small circle oriented counterclockwise, centered at z_0, such that f is holomorphic on C and its interior, except possibly at z_0. Then

$$\int_C f(z)dz = 2\pi i Res_{z_0} f$$

We choose the counter clockwise curve along the unit circle as C, which is a small counter clockwise circle centered at 0. We have an expression for $F(y)$ as a power series with a positive radius of convergence and finitely many terms of negative order. So it is an analytic function on $D_1 \backslash \{0\}$ and therefore holomorphic on $D_1 \backslash \{0\}$.
In our case we hence get

$$\int_C F(y)dy = 2\pi i \cdot \mathrm{Res}_0 F = 2\pi i \cdot b_{-1}.$$

Using the notation $e(x) = e^{2\pi i x}$, we compute

$$\begin{aligned} b_{-1} &= \frac{1}{2\pi i} \int_C \tilde{G}(y) \cdot y^{-n-1} dy \\ &= \int_0^1 \tilde{G}(e^{2\pi i t})(e^{2\pi i t})^{-n-1} \cdot e^{2\pi i t} dt \\ &= \int_0^1 \tilde{G}(e^{2\pi i t})(e^{-2\pi i t n}) dt \\ &= \int_0^1 \tilde{G}(e(t)) e(-nt) dt \end{aligned}$$

Since $b_{-1} = a_n$, we now have a formula to compute a_n. Filling in $\tilde{G}$, this becomes

$$a_n = \int_0^1 \sum_{j=6}^{N} a_j \cdot e(jt) \cdot e(-nt)dt. \tag{3.4}$$

This may seem a bit weird, since we try to find an expression for a_n and use $\sum_{j=6}^{N} a_j e(jt)$ for it. It is however way easier to find an estimate of the sum over many coefficients a_j then finding an estimate of one specific a_j. After all, looking for combinations of three primes p_1, p_2, p_3, such that $p_1 + p_2 + p_3 = j$ is way more difficult than looking for combinations of three primes that add up to any number up to N. The latter corresponds to finding all combinations of three primes and stop if their sum gets larger than N. Using the steps of (3.3) in reversed order, we can rewrite (3.4) to

$$\begin{aligned} a_n &= \int_0^1 \left(\sum_{\substack{p_1 \in \mathbb{P} \\ p_1 < n}} e(p_1 t) \right) \cdot \left(\sum_{\substack{p_2 \in \mathbb{P} \\ p_2 < n}} e(p_2 t) \right) \cdot \left(\sum_{\substack{p_3 \in \mathbb{P} \\ p_3 < n}} e(p_3 t) \right) \cdot e(-nt)dt \\ &= \int_0^1 (g(t))^3 \cdot e(-nt)dt, \end{aligned} \tag{3.5}$$

where

$$g(t) = \sum_{\substack{p \in \mathbb{P} \\ p < n}} e(pt).$$

3.1.2 Vinogradov's approach

A perhaps more intuitive way to approach this problem is the line of attack Vinogradov followed. To describe the behaviour of a_n, we would like to have an indicator function for A_n. That is a function that gives 1 if we fill in an element of the set A_n defined in (3.1) and 0 otherwise. Then we can sum over this formula and thus get the number of possible ways to write n as the sum of three primes. Using again the notation $e(x) = e^{2\pi i x}$, the function

$$\int_0^1 e(mt)dt, \qquad \text{with } m \in \mathbb{N} \tag{3.6}$$

displays such behaviour. For

$$\begin{aligned} \int_0^1 e(mt)dt &= \begin{cases} [t]_0^1 = 1 - 0 & \text{if } m = 0 \\ \left[\frac{1}{2\pi i m} \cdot e^{2\pi i m t}\right]_0^1 = \frac{1}{2\pi i m} \cdot (1 - 1) & \text{if } m \neq 0 \end{cases} \\ &= \begin{cases} 1 & \text{if } m = 0 \\ 0 & \text{if } m \neq 0 \end{cases} \end{aligned}$$

holds. To make this formula of any use for our problem we substitute $p_1 + p_2 + p_3 - n$, with $p_1 1, p_2, p_3 \in \mathbb{P}$ for m in (3.6). This gives us

$$\int_0^1 e\left((p_1 + p_2 + p_3 - n)t\right) dt = \begin{cases} 1 & \text{if } p_1 + p_2 + p_3 = n \\ 0 & \text{if } p_1 + p_2 + p_3 \neq n \end{cases}$$
$$= \begin{cases} 1 & \text{if } (p_1, p_2, p_3) \in A_n \\ 0 & \text{otherwise.} \end{cases}$$

Since we are looking for combinations of three primes that add up to n, we only need to take a look at primes smaller than n. We can now formulate the following formula for a_n.

$$\begin{aligned} a_n &= \sum_{\substack{p_1 \in \mathbb{P} \\ p_1 < n}} \sum_{\substack{p_2 \in \mathbb{P} \\ p_2 < n}} \sum_{\substack{p_3 \in \mathbb{P} \\ p_3 < n}} \int_0^1 e\left((p_1 + p_2 + p_3 - n)t\right) dt \\ &= \int_0^1 \sum_{\substack{p_1 \in \mathbb{P} \\ p_1 < n}} \sum_{\substack{p_2 \in \mathbb{P} \\ p_2 < n}} \sum_{\substack{p_3 \in \mathbb{P} \\ p_3 < n}} e\left((p_1 + p_2 + p_3 - n)t\right) dt \\ &= \int_0^1 \sum_{\substack{p_1 \in \mathbb{P} \\ p_1 < n}} \sum_{\substack{p_2 \in \mathbb{P} \\ p_2 < n}} \sum_{\substack{p_3 \in \mathbb{P} \\ p_3 < n}} e(p_1 t) \cdot e(p_2 t) \cdot e(p_3 t) \cdot e(-nt) dt \\ &= \int_0^1 \left(\sum_{\substack{p_1 \in \mathbb{P} \\ p_1 < n}} e(p_1 t) \right) \cdot \left(\sum_{\substack{p_2 \in \mathbb{P} \\ p_2 < n}} e(p_2 t) \right) \cdot \left(\sum_{\substack{p_3 \in \mathbb{P} \\ p_3 < n}} e(p_3 t) \right) \cdot e(-nt) dt \\ &= \int_0^1 (g(t))^3 \cdot e(-nt) dt, \end{aligned} \tag{3.7}$$

where

$$g(t) = \sum_{\substack{p \in \mathbb{P} \\ p < n}} e(pt).$$

Note that we now found the same formula as (3.5). Because Vinogradov's approach is most consistent with Maynards line of attack, we use this for the comparison.

3.2 Define major and minor arcs

In the last section we found an expression for a_n, but if we compute this integral directly we do no get a convenient answer. Therefore we partition the unit circle in major and minor arcs and sum over them. The major arcs are small arcs where $(g(t))^3 \cdot e(-nt)$ is large. The minor arcs contain the bulk of the circle where $(g(t))^3 \cdot e(-nt)$ takes small values. We therefore want to find a bound for the minor arcs which can be seen as an error term. The major arcs yield the main term for a_n.

In this section we take a look at the major and minor arcs for the Ternary Goldbach Problem. For both this section and the next, [16], Chapter 3 and [4] are combined to function as a basis. For the Ternary Goldbach Problem it turns out to be convenient if we set

$$f(t) = \sum_{\substack{p\in\mathbb{P}\\ p<n}} (\log p) e(pt),$$

instead of

$$g(t) = \sum_{\substack{p\in\mathbb{P}\\ p<n}} e(pt),$$

where $\log p$ is the natural logarithm. This makes it possible to bring in the Von Mangoldt function that we need to estimate the major arcs. We now define $R(n)$ to be what we get if we fill in this new f in (3.7). This gives us

$$\begin{aligned}
R(n) &= \int_0^1 (f(t))^3 e(-nt)dt \\
&= \int_0^1 \sum_{\substack{p_1\in\mathbb{P}\\ p_1<n}} \sum_{\substack{p_2\in\mathbb{P}\\ p_2<n}} \sum_{\substack{p_3\in\mathbb{P}\\ p_3<n}} (\log p_1)(\log p_2)(\log p_3) e((p_1+p_2+p_3-n)t)dt \\
&= \sum_{\substack{p_1\in\mathbb{P}\\ p_1<n}} \sum_{\substack{p_2\in\mathbb{P}\\ p_2<n}} \sum_{\substack{p_3\in\mathbb{P}\\ p_3<n}} (\log p_1)(\log p_2)(\log p_3) \int_0^1 e((p_1+p_2+p_3-n)t)dt \\
&= \sum_{\substack{p_1\in\mathbb{P}\\ p_1<n}} \sum_{\substack{p_2\in\mathbb{P}\\ p_2<n}} \sum_{\substack{p_3\in\mathbb{P}\\ p_3<n \\ p_1+p_2+p_3=n}} (\log p_1)(\log p_2)(\log p_3) \\
&= \sum_{p_1\in\mathbb{P}} \sum_{p_2\in\mathbb{P}} \sum_{\substack{p_3\in\mathbb{P}\\ p_1+p_2+p_3=n}} (\log p_1)(\log p_2)(\log p_3).
\end{aligned} \tag{3.8}$$

So $R(n)$ as aboveal to a_n, which we can write as

$$a_n = \sum_{p_1\in\mathbb{P}} \sum_{p_2\in\mathbb{P}} \sum_{\substack{p_3\in\mathbb{P}\\ p_1+p_2+p_3=n}} 1,$$

but it is a weighted count of it. Hence if we prove that $R(n) > 0$ for all sufficiently large odd $n \in \mathbb{N}$, the Weak Goldbach Conjecture follows.

Now we make the major arcs specific. Let $B \in \mathbb{R}_{>0}$ be a large positive constant and write $P = (\log n)^B$ for sufficiently large n. Assuming n to be large is no problem since

we want to find asymptotic behaviour for a_n if n goes to infinity.
We now define for all $a, q \in \mathbb{N}$ such that $0 < a \leq q \leq P$ and $(a, q) = 1$;

$$\begin{aligned}\mathfrak{M}_{a,q} &= \left\{ t \in \mathbb{R} : \left| t - \frac{a}{q} \right| \leq \frac{P}{n} \right\} \\ &= \left[\frac{a}{q} - \frac{P}{n}, \frac{a}{q} + \frac{P}{n} \right].\end{aligned}$$

These sets are small intervals around the fractions $\frac{a}{q}$. Set $\mathcal{U} = \left(\frac{P}{n}, 1 + \frac{P}{n}\right]$ which is an interval of length 1. Since we are integrating along the unit circle in (3.8), we can shift the interval of integration from $(0, 1]$ to $\mathcal{U}$. To make this possible, we want all of the intervals $\mathfrak{M}_{a,q}$ to lie inside $\mathcal{U}$. This is why we check the lower bound of the interval around the smallest fraction $\frac{a}{q}$ and the upper bound of the interval around the biggest fraction $\frac{a}{q}$.

The smallest possible value of $\frac{a}{q}$ is $\frac{1}{P}$. The lower bound of the corresponding interval is given by

$$\frac{1}{P} - \frac{P}{n} \quad = \quad \frac{n - P^2}{Pn} \quad = \quad \frac{\frac{n}{P} - P}{n} \quad = \quad \frac{\frac{n}{(\log n)^B} - (\log n)^B}{n},$$

which is larger than $\frac{P}{n}$ for n large enough. After all $\frac{n}{(\log n)^B} - (\log n)^B > (\log n)^B$ for n large enough. Therefore we know for sure that the interval lies inside $\mathcal{U}$.
The biggest possible value of $\frac{a}{q}$ is 1. The upper bound of the corresponding interval is $1 + \frac{P}{n}$, which is the upper bound of $\mathcal{U}$. Once again the interval lies in $\mathcal{U}$. This ensures us that there are no major arcs cut off when we integrate over $\mathcal{U}$.

Now we want to make sure that none of the intervals overlap. Let two different midpoints $\frac{a}{q} \neq \frac{a'}{q'}$ be given such that $a, q, a', q' \in \mathbb{N}$, $0 \leq a \leq q \leq P$, $(a, q) = 1$, $0 \leq a' \leq q' \leq P$ and $(a', q') = 1$.
So $aq' - a'q \neq 0$ and therefore $|aq' - a'q| \geq 1$, since a, q, a', q' are all integers. Now we can conclude

$$\left| \frac{a}{q} - \frac{a'}{q'} \right| = \left| \frac{aq' - a'q}{qq'} \right| \geq \frac{1}{qq'} \geq \frac{1}{P^2} = \frac{1}{(\log n)^{2B}},$$

which is larger than $2\frac{P}{n}$ for n large enough. After all, we know that $\frac{1}{(\log n)^{2B}} > 2\frac{(\log n)^B}{n}$ for n large enough. Therefore we know that the midpoints $\frac{a}{q}$ and $\frac{a'}{q'}$ are so far apart that their intervals cannot intersect since the bound of each interval is $\frac{P}{n}$ away from the midpoint. So the intervals are pairwise disjoint.

We define the set of major arcs by

$$\mathfrak{M} = \bigcup_{q \leq P} \bigcup_{\substack{0 \leq a \leq q \\ (a,q)=1}} \mathfrak{M}_{a,q}.$$

We now define the set of minor arcs as $\mathfrak{m} = \mathcal{U} \backslash \mathfrak{M}$.
Using these definitions for $\mathfrak{M}$ and $\mathfrak{m}$, we know that $\mathfrak{M} \cup \mathfrak{m} = \mathcal{U}$. Now we shift the interval of integration from $(0, 1]$ to $\mathcal{U}$. We can hence write

$$\begin{aligned} R(n) &= \int_0^1 (f(t))^3 \cdot e(-nt)dt \\ &= \int_{\frac{P}{n}}^{1+\frac{P}{n}} (f(t))^3 \cdot e(-nt)dt \\ &= \int_{\mathfrak{M}} (f(t))^3 \cdot e(-nt)dt + \int_{\mathfrak{m}} (f(t))^3 \cdot e(-nt)dt. \end{aligned} \tag{3.9}$$

In the next section we give a suitable value of B for which we can find convenient estimates.

3.3 Estimate the major and minor arcs

In this section we start with an investigation of the minor arcs. In subsection 3.3.1 we give a bound for them. In subsection 3.3.2 the major arcs are discussed less extensive. We just state the outcome without much further explanation.

3.3.1 Minor arcs

To find a good estimate for them, we start with an estimation for $\int_0^1 |f(t)|^2 dt$. Thereafter we find a bound for the absolute value of f on the minor arcs. Combining these two results, we obtain an estimation for $\int_{\mathfrak{m}} |f|^3 dt$, which gives us an estimate for the minor arcs.

Since $f(t) = \sum_{\substack{p \in \mathbb{P} \\ p \leq n}} (\log p) e(pt)$, we know that $\overline{f(t)} = f(-t)$. We hence write

$$\begin{aligned} \int_0^1 |f(t)|^2 \, dt &= \int_0^1 f(t) \cdot \overline{f(t)} dt \\ &= \int_0^1 \left(\sum_{\substack{p_1 \in \mathbb{P} \\ p_1 < n}} (\log p_1) e(p_1 t) \right) \cdot \left(\sum_{\substack{p_2 \in \mathbb{P} \\ p_2 < n}} (\log p_2) e(-p_2 t) \right) dt \\ &= \int_0^1 \sum_{\substack{p_1 \in \mathbb{P} \\ p_1 < n}} \sum_{\substack{p_2 \in \mathbb{P} \\ p_2 < n}} (\log p_1)(\log p_2) e((p_1 - p_2)t) dt \\ &= \sum_{\substack{p_1 \in \mathbb{P} \\ p_1 < n}} \sum_{\substack{p_2 \in \mathbb{P} \\ p_2 < n}} (\log p_1)(\log p_2) \int_0^1 e((p_1 - p_2)t) dt. \end{aligned}$$

Since $\int_0^1 e((p_1 - p_2)t)dt = 1$ if $p_1 = p_2$ and 0 if $p_1 \neq p_2$, we get

$$\int_0^1 |f(t)|^2 \, dt = \sum_{\substack{p \in \mathbb{P} \\ p<n}} (\log x)^2.$$

To give a more specific bound, we need the Prime Number Theorem. Let $\pi(x) = \#\{p \in \mathbb{P} : p \leq x\}$.

Theorem 2 (Prime Number Theorem) *([20], Chapter XVI, p. 449, Theorem 2.5)* $\pi(x) \sim \frac{x}{\log x}$.

Here $\pi(x) \sim \frac{x}{\log x}$ means that $\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\log x}} = 1$, so in particular we know that $\pi(x) \ll \frac{x}{\log x}$.

It hence follows that $\sum_{\substack{p \in \mathbb{P} \\ p<n}} 1 \ll \frac{n}{\log n}$, so there exists a $C \in \mathbb{R}$ such that $\left|\sum_{\substack{p \in \mathbb{P} \\ p<n}} 1\right| < C \cdot \frac{n}{\log n}$. Using this in combination with the fact that $\log p$ is a strictly increasing function, we can conclude that

$$\sum_{\substack{p \in \mathbb{P} \\ p<n}} (\log p)^2 = \left|\sum_{\substack{p \in \mathbb{P} \\ p<n}} (\log p)^2\right| \leq (\log n)^2 \left|\sum_{\substack{p \in \mathbb{P} \\ p<n}} 1\right| < (\log n)^2 \cdot C \cdot \frac{n}{\log n} = C \cdot n \log n.$$

Using the previous results, we can state that

$$\int_0^1 |f(t)|^2 \, dt \ll n \log n. \tag{3.10}$$

This is the first result, which holds for the whole interval $\mathcal{U}$. This is the main contribution to bound of the minor arcs. The most important step in estimating the minor arcs however, is the following theorem.

Theorem 3 *([16], Chapter 3, p. 27, Theorem 3.1)*
Suppose that (a,q) = 1, $q \leq n$ and $|t - \frac{a}{q}| \leq \frac{1}{q^2}$. Then

$$f(t) \ll (\log n)^4 (nq^{-\frac{1}{2}} + n^{\frac{4}{5}} + n^{\frac{1}{2}} q^{\frac{1}{2}}).$$

We do not prove this theorem here. To apply this theorem, we also need the Dirichlet Lemma, which is stated below.

Lemma 1 (Dirichlet Lemma) *([16], p. 9, lemma 2.1)*
Let t denote a real number. Then for each real number $X \geq 1$ there exists a rational number $\frac{a}{q}$ with $(a,q) = 1$, $1 \leq q \leq X$ and

$$\left|t - \frac{a}{q}\right| \leq \frac{1}{qX}.$$

We now choose $Q = \frac{n}{(\log n)^B} > 1$. From the Dirichlet Lemma we know that there exists a fraction $\frac{a}{q}$ for every $t \in \mathcal{U}$ such that $(a, q) = 1$, $1 \leq q \leq Q$ and

$$\left| t - \frac{a}{q} \right| \leq \frac{1}{qQ}. \tag{3.11}$$

Since $(a, q) = 1$ and $q \leq Q = \frac{n}{(\log n)^B} < n$, the first two requirements of Theorem 3 are met. Since $q \leq Q$, we also know that $\frac{1}{qQ} < \frac{1}{q^2}$. Combining this with (3.11), gives us the third requirement of the theorem. Therefore we can now conclude that

$$f(t) \ll (\log n)^4 (nq^{-\frac{1}{2}} + n^{\frac{4}{5}} + n^{\frac{1}{2}} q^{\frac{1}{2}}). \tag{3.12}$$

Until now we did not use the fact that we are only looking at the minor arcs, but we use that right now. Let us take a closer look at (3.11).

$$\left| t - \frac{a}{q} \right| \leq \frac{1}{qQ} = \frac{1}{\frac{qn}{(\log n)^B}} = \frac{(\log n)^B}{qn} = \frac{1}{q} \cdot \frac{P}{n} \leq \frac{P}{n}.$$

This looks like the definition for the major arcs. There is a difference however. For the major arcs we require that $1 \leq a \leq q \leq P$. Since we know $t \in \mathcal{U}$, it follows that $1 \leq a \leq q$, but not that $q \leq P$. Because we are looking at the minor arcs, we want to make sure that $t \notin \mathfrak{M}$. The only way to ensure this, is to require that $q > P = (\log n)^B$. This way $\frac{a}{q}$ can not be the middle point fraction of any major arc.
Since $q > (\log n)^B$, we know that $q^{\frac{1}{2}} > (\log n)^{\frac{B}{2}}$ and therefore

$$q^{-\frac{1}{2}} = \frac{1}{q^{\frac{1}{2}}} < \frac{1}{(\log n)^{\frac{B}{2}}} = (\log n)^{-\frac{B}{2}}.$$

We also know that $q \leq Q = \frac{n}{(\log n)^B}$, so $q^{\frac{1}{2}} \leq \left(\frac{n}{(\log n)^B}\right)^{\frac{1}{2}} = \frac{n^{\frac{1}{2}}}{(\log n)^{\frac{B}{2}}}$ and therefore

$$n^{-\frac{1}{2}} q^{\frac{1}{2}} \leq \frac{n^{-\frac{1}{2}} n^{\frac{1}{2}}}{(\log n)^{\frac{B}{2}}} = (\log n)^{-\frac{B}{2}}.$$

The third observation we could make, is that $n^{-\frac{1}{5}} < (\log n)^{-\frac{B}{2}}$ for n large enough. Substituting these three results in (3.12), gives us

$$\begin{aligned}
f(t) &\ll (\log n)^4 (nq^{-\frac{1}{2}} + n^{\frac{4}{5}} + n^{\frac{1}{2}} q^{\frac{1}{2}}) \\
&\ll n(\log n)^4 (q^{-\frac{1}{2}} + n^{-\frac{1}{5}} + n^{-\frac{1}{2}} q^{\frac{1}{2}}) \\
&\ll n(\log n)^4 (3(\log n)^{-\frac{B}{2}}) \\
&\ll n(\log n)^{4-\frac{B}{2}}.
\end{aligned}$$

Since f is positive, we have $\sup_{t\in\mathfrak{m}} |f(t)| \ll n(\log n)^{4-\frac{B}{2}}$.
Combining this result with (3.10) and assuming $B \geq 2A + 10$ for some positive constant A, gives us

$$\begin{aligned}
\left|\int_{\mathfrak{m}} (f(t))^3 e(-nt)dt\right| &\leq \int_{\mathfrak{m}} |f(t)|^3 dt \\
&\ll \sup_{t\in\mathfrak{m}} |f(t)| \cdot \int_{\mathfrak{m}} |f(t)|^2 dt \\
&\ll n(\log n)^{4-\frac{B}{2}} dt \cdot \int_{\mathfrak{m}} |f(t)|^2 \\
&\ll n(\log n)^{4-\frac{2A+10}{2}} \int_{\mathfrak{m}} |f(t)|^2 dt \\
&\ll n(\log n)^{-A-1} \cdot n \log n \\
&\ll n^2(\log n)^{-A}.
\end{aligned}$$

To conclude everything we did for the minor arcs, we can formulate the following theorem.

Theorem 4 *([16], Chapter 3, p. 30, Theorem 3.2)*
Suppose that A is a positive constant [...], then

$$\int_{\mathfrak{m}} |f(t)|^3 dt \ll n^2(\log n)^{-A}.$$

We can hence write

$$\left|\int_{\mathfrak{m}} (f(t))^3 e(-nt)dt\right| \leq \int_{\mathfrak{m}} |f(t)|^3 dt = O(n^2(\log n)^{-A}). \tag{3.13}$$

3.3.2 Major arcs

Now we take a quick look at the major arcs. We start with defining for $n \geq 2$

$$\mathfrak{S}(n) = \left(\prod_{p\nmid n} \left(1 + (p-1)^{-3}\right)\right) \cdot \left(\prod_{p|n} \left(1 - (p-1)^{-2}\right)\right), \tag{3.14}$$

where p denotes the prime divisors of n. If n is an even number, one of its prime factors is 2. This gives the term $(1 - (2-1)^{-2}) = 0$ in the second product. So $\mathfrak{S}(n) = 0$ for even n. None of the terms becomes 0 if we have an odd number n, and it turns out that $\mathfrak{S}(n) \gg 1$ for odd n. We want to prove that $a_n \geq 1$, so if we use the formula above in an expression for a_n it is a problem if its value is zero. Since we are only interested in odd n however, (3.14) does not become zero, so there is no problem.
After a quite involving proof, it turns out that the following theorem holds.

Theorem 5 *([16], Chapter 3, p. 32, Theorem 3.3)*
Suppose that A is a positive constant[...]. Then

$$\int_{\mathfrak{M}} f(t)^3 e(-tn)dt = \frac{1}{2}n^2\mathfrak{S}(n) + O(n^2(\log n)^{-A})$$

where $\mathfrak{S}$ satisfies (3.14).

Combining this with (3.9) and (3.13), we obtain

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

which leads to the next theorem.

Theorem 6 *([16], Chapter 3, p. 33, Theorem 3.4)*
Suppose that A is a positive constant and $R(n)$ satisfies (3.8). Then

$$R(n) = \frac{1}{2}n^2\mathfrak{S}(n) + O(n^2(\log n)^{-A})$$

where $\mathfrak{S}$ satisfies (3.14).

Since $\mathfrak{S}(n) \gg 1$ for odd values of n, we know that $R(n) > 0$ for all odd values of n bigger than a certain n_0. So we can conclude Conjecture 1 for n large enough.

Chapter 4

The Restricted Digit Problem

In this chapter the problem is introduced that James Maynard [21] solved. In his paper Maynard proves the following theorem.

Theorem 7 (Restricted Digit Problem)
Let $k \in \{0, ..., 9\}$. Then there are infinitely many primes without k in their decimal expansion.

Although it is not a generally accepted name and Maynard does not use it, we will call this the Restricted Digit Problem from now on.
We now define

$$\tilde{\mathcal{A}}_k = \left\{ \sum_{i \geq 0} n_i 10^i : n_i \in \{0, ..., 9\} \backslash \{k\} \right\} \backslash \{\infty\} \text{ with } k \in \{0, ..., 9\}, \tag{4.1}$$

which is a slightly different notation than Maynard uses. Theorem 7 reduces to solving the equation $p = a$, for $a \in \tilde{\mathcal{A}}_k$ and $p \in \mathbb{P}$.

As we have seen in Chapter 2, the Binary Goldbach Problem is an important unsolved Problem in number theory. In the Binary Goldbach Problem we look for sums of two primes such that $n = p_1 + p_2$, which makes it a binary problem. Similarly we are looking for numbers for which $0 = p - a$ holds in the Restricted Digit Problem. So this is a binary problem as well. We would therefore expect both problems to be in some sense comparable. Untill now however, 2 seems impossible to solve, but Maynard did solve 7. That makes it quite an interesting result. The similarities are worked out further in Chapter 5, where we formulate the integral used by the circle method for the Restricted Digit Problem. First we take a closer look at $\tilde{\mathcal{A}}_k$.

4.1 Density of $\tilde{\mathcal{A}}_k$

In the last section we saw that both the Binary Goldbach problem and the Restricted Digit Problem are binary problems. In the Goldbach Problem, both parameters are elements of $\mathbb{P}$, though in the Restricted Digit Problem, one of the parameters is an element

of $\mathbb{P}$ and one of $\tilde{\mathcal{A}}_k$.
Theorem 7 can be proved due to our understanding of $\tilde{\mathcal{A}}_k$. Hence we now study the density of $\tilde{\mathcal{A}}_k$ and compare it to the density of $\mathbb{P}$. Using (4.1), we plot $n_{k,X} = \#\{a \in \tilde{\mathcal{A}}_k : a < X\}$. the resulting graph for $k = 2$ can be seen in figure 4.1.

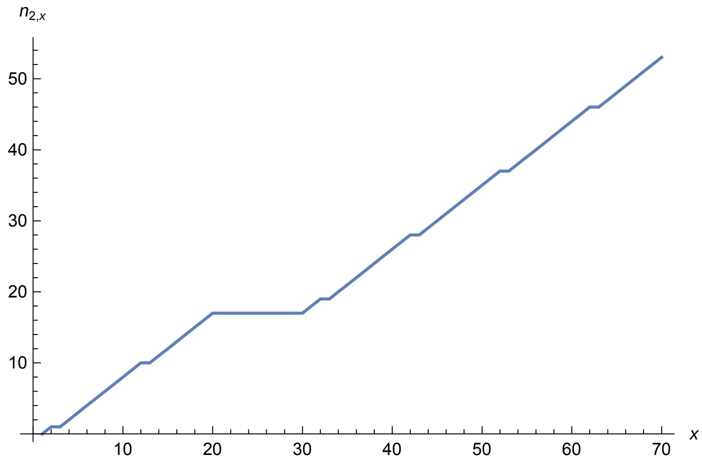

Figure 4.1: Amount of numbers smaller than X without 2 in their decimal expansion ($n_{2,X}$).

We can make similar graphs for the other values of k. This gives us an idea of the density of $\mathcal{A}_k$.
We know that there are only 9^m possible numbers with m digits if each digit is an element of $\{0, 1, ..., 9\}\backslash\{k\}$. In other words $\#\{a \in \tilde{\mathcal{A}}_k : a < X\} = 9^m$ if $X = 10^m$ for some $m \in \mathbb{N}$. That is the reason we will further assume that X is an integral power of 10. We choose such an $X \in \mathbb{N}$ and define

$$\mathcal{A}_k(X) = \left\{ \sum_{1 \geq 0} n_i 10^i < X : n_i \in \{0, ..., 9\}\backslash\{k\} \right\} \text{ with } k \in \{0, ..., 9\}. \tag{4.2}$$

Note the upper bound in which this definition differs from the definition of $\tilde{\mathcal{A}}_k$ we used before. We can now rewrite 9^m as

$$10^{^{10}\log 9^m} = (10^m)^{^{10}\log 9} = X^{^{10}\log 9}.$$

So

$$\#\mathcal{A}_k(X) = O(X^{^{10}\log 9})$$

for all values of X. Now define $c = {}^{10}\log(10/9) \approx 0,046 > 0$. Since

$$X^{^{10}\log 9} = X^{1-(1-^{10}\log 9)} = X^{1-(^{10}\log 10 - ^{10}\log 9)} = X^{1-^{10}\log(10/9)},$$

it follows that $\#\mathcal{A}_k(X)$ is $O(X^{1-c})$. We introduce the following notion.

Definition 1
A subset A of $\mathbb{N}$ is called **sparse** *in $\mathbb{N}$ if*

$$\liminf_{X\to\infty} \frac{\#\{a \in A\colon a < X\}}{X} = 0.$$

Since

$$\lim_{X\to\infty} \frac{X^{1-c}}{X}$$

exists and equals 0, we now know that

$$\liminf_{X\to\infty} \frac{X^{1-c}}{X}$$

equals 0 and therefore $\tilde{\mathcal{A}}_k$ is a sparse subset of $\mathbb{N}$. As we would like to compare $\tilde{\mathcal{A}}_k$ with $\mathbb{P}$, we now check if $\mathbb{P}$ is a sparse subset of $\mathbb{N}$ as well. We can use the Prime Number Theorem (2) to see that.

Again $\lim_{X\to\infty} \frac{\frac{X}{\log X}}{X} = \lim_{X\to\infty} \frac{X}{X\log X}$ exists and equals 0, so the $\liminf_{X\to\infty} \frac{\frac{X}{\log X}}{X}$ equals 0 as well. It thus follows that $\mathbb{P}$ also is a sparse subset of $\mathbb{N}$, but $\tilde{\mathcal{A}}_k$ is much sparser. In general, it is more difficult to solve a problem if it is about a sparser subset. The reason this is not the case for the Restricted Digit Problem, is the unusually nice Fourier structure of $\tilde{\mathcal{A}}_k$. This explains why the circle method can be successfully applied to the Restricted Digit Problem and not to the Goldbach Problem. We discuss this nice Fourier structure a bit further in section 5.3.

4.2 Previous studies of sets related to $\mathcal{A}_k(X)$

Several mathematicians have been looking at the structure of $\mathcal{A}_k(X)$ and related sets before. In this section we recall some of their results.

For example Erdős, Maudit and Sárközy [9] studied the distribution in residue classes of integers in base $g \geq 2$ not exceeding x. They denote by $S(n)$ the sum of the digits of n in basis g and define the set

$$\mathscr{U}_{(m,r)}(N) = \{n \in \mathbb{N}\colon n \leq N, S(n) \equiv r (\text{mod } m)\} \tag{4.3}$$

for $N, m \in \mathbb{N}$ and $r \in \mathbb{Z}$.
Gelfond [11] proved the following theorem about (4.3).

Theorem 8 *([9], Section 1, RU 1)*
If $m \in N$ is fixed with

$$(m, g-1) = 1,$$

then for all $r \in \mathbb{Z}$ and all "small" $q \in \mathbb{N}$ the set $\mathscr{U}_{(m,r)}(N)$ is well-distributed in the residue classes modulo q.

Since $\left|\mathscr{U}_{(m,r)}\right| \gg N$, the set $\mathscr{U}_{(m,r)}(N)$ has a positive density and is therefore not sparse. Maudit and Sárközy [22] wrote a paper in which they prove Theorem 9 for the sparser set defined below. Let

$$\mathscr{V}_k(N) = \{n \in \mathbb{N} \colon n \leq N, S(n) = k\} \tag{4.4}$$

for $k \in \mathbb{N}$ and $0 \leq k \leq (g-1)({}^g \log N + 1)$.
The analogue of Theorem 8 for (4.4) turns out to be the following.

Theorem 9 *([9], Section 1, RV 1)*
If $g \in \mathbb{N}$, $g \geq 2$, $k \in \mathbb{N}$, $0 \leq k \leq (g-1)({}^g \log N + 1)$,

$$\min\Big(k, (g-1)^k \log N - k\Big) \to \infty,$$

$m \in \mathbb{N}$ and m is "small", then $\mathscr{V}_k(N)$ is well-distributed in the residue classes modulo m.

Erdős, Maudit and Sárközy derive a similar theorem for an even sparser set. They introduce the set $\mathscr{D}$ of integers in base g without one or more digits in their g-ary expansion. To construct this, let $g \in \mathbb{N}$, $g \geq 3$ and $t \in \mathbb{N}$ such that $2 \leq t \leq g-1$. Now we can define

$$\mathscr{D} \subset \{0, 1, ..., g-1\} \text{ such that } 0 \in \mathscr{D} \text{ and } |\mathscr{D}| = t,$$

which has cardinality $N^{1-\epsilon}$.

The sets they study are different form $\mathcal{A}_k(X)$, but they are sets containing numbers with some restrictions on their digits. When studying the Restricted Digit Problem we make use of the property that both $\mathbb{P}$ and $\mathcal{A}_k(X)$ are well-distributed in the residue classes modulo $q \in \{0, 1, ..., 9\}$.

Another related work is [1] from Banks, Conflitti and Shparlinski. They found upper bounds for multiplicative character sums and exponential sums over sets of integers in base $g \geq 2$ with various restrictions on their digits. We shall not use character sums, but exponential sums are used frequently. The methods they use have similarities with the methods used by Maynard.
Banks and Shparlinski [2] wrote another paper together in which they derive asymptotic formulas for some arithmetic properties of numbers with restricted digits. This is close to what Maynard does.

In a later paper, Drmota and Maudit [8] define a set $\mathcal{N}$ as a set of non-negative integers with relations between their digits in base q. To understand their results we need to introduce the following notion.

Definition 2
Let $\{x_n\}_{n\in\mathbb{N}}$, with $x_n \in \mathbb{R}$ for all n. Then $\{x_n\}_{n\in\mathbb{N}}$ is **uniformly distributed modulo 1** *if for every $s, t \in \mathbb{R}$ with $0 \leqslant< t < 1$, we have that*

$$\lim_{N\to\infty} \frac{1}{N} \# \{j : 0 \leq j \leq n-1, \{x_j\} \in [s,t]\} = t - s,$$

where $\{x_j\} = x_j - \lfloor x_j \rfloor x_j$.

Drmota and Maudit now prove that $(\alpha n)_{n \in \mathcal{N}}$ is uniformly distributed modulo 1 for all $\alpha \in \mathbb{C} \backslash \mathbb{Q}$.

Chapter 5

Circle method applied to the Restricted Digit Problem

To apply the circle method we need to formulate a suitable sequence. Since we want to prove that there are infinitely many primes with a missing digit, we would like to formulate a sequence that gives us the amount of primes with a missing digit up to a fixed number X. Let $\mathcal{A}_k(X)$ be defined as in (4.2). Now we introduce $d_X = \#\{(p,a) \in \mathbb{P} \times \mathcal{A}_k(X) \colon p = a\}$, which we can rewrite as $d_X = \#\{p \in \mathcal{A}_k(X) : p \in \mathbb{P}\}$. So d_X is the number of primes in $\mathcal{A}_k(X)$ or equivalently the number of primes in $\tilde{\mathcal{A}_k}(X)$ smaller than X. So d_X depends on a fixed k. If we can show that this sequence goes to infinity for every $k \in \{0, 1, ..., 9\}$, Theorem 7 follows.

5.1 Find an expression to describe d_X

We can construct the generating function of d_X in the same way as we did before in section 3.1.2 for the Ternary Goldbach Problem.

$$\begin{aligned}
d_X &= \sum_{a \in \mathcal{A}_k(X)} \sum_{\substack{p \in \mathbb{P} \\ p < X}} \int_0^1 e((p-a)t)dt \\
&= \int_0^1 \sum_{a \in \mathcal{A}_k(X)} \sum_{\substack{p \in \mathbb{P} \\ p < X}} e(pt) \cdot e(-at)dt \\
&= \int_0^1 \left(\sum_{a \in \mathcal{A}_k(X)} e(-at) \right) \cdot \left(\sum_{\substack{p \in \mathbb{P} \\ p < X}} e(pt) \right) dt \qquad (5.1)
\end{aligned}$$

Now we introduce some notation. Let

$$\mathbf{1}_{\mathcal{A}_k(X)}(a) = \begin{cases} 1 & \text{if } a \in \mathcal{A}_k(X) \\ 0 & \text{otherwise.} \end{cases}$$

Since (3.6) has these properties, it could be used as **1**. Using aforementioned definition we define

$$S_{\mathcal{A}_k(X)}(\theta) = \sum_{a \in \mathbb{N}} \mathbf{1}_{\mathcal{A}_k(X)}(a) e(a\theta).$$

Analogously we define

$$S_{\mathbb{P}}(\theta) = \sum_{\substack{p \in \mathbb{N} \\ p < X}} \mathbf{1}_{\mathbb{P}}(p) e(p\theta).$$

If we plug in these definitions in (5.1), we get

$$d_X = \int_0^1 S_{\mathcal{A}_k(X)}(-t) S_{\mathbb{P}}(t) dt.$$

It turns out later that it is easier to work with a sum instead of an integral. Therefore we rewrite this integral for d_X. We divide the interval $[0, 1)$ in X steps of size $\frac{1}{X}$. Now we estimate the integral in (5.1) by the following sum.

$$\begin{aligned} d_X &\approx \sum_{0 \le t < 1} S_{\mathcal{A}_k(X)}(-t) \, S_{\mathbb{P}}(t) \cdot \frac{1}{X} \\ &= \frac{1}{X} \sum_{0 \le t < X} S_{\mathcal{A}_k(X)}(-t) \, S_{\mathbb{P}}(t) \cdot \frac{1}{X} \\ &= \frac{1}{X} \sum_{0 \le t < X} S_{\mathcal{A}_k(X)} \left(\frac{-t}{X} \right) S_{\mathbb{P}} \left(\frac{t}{X} \right) \end{aligned} \tag{5.2}$$

This is the equivalence of the integral 3.7 for the Goldbach Problem. We see that it is the product of two functions $S_{\mathcal{A}}$ and $S_{\mathbb{P}}$ since it is a binary problem.

Note that (5.2) is an approximation of d_X. It turns out however that this formula is equal to d_X. To see this we use a different approach to find an expression for d_X. For this, we study the summation

$$\frac{1}{X} \sum_{0 \le t < X} e \left(\frac{(p-a)t}{X} \right).$$

First we look at the case that $X | (p-a)$. Then we know that $\frac{(p-a)t}{X}$ is an integer and therefore $e\left(\frac{(p-a)t}{X}\right) = 0$ for all t and thus the summation equals 0. Now we look at the case that $X \nmid (p-a)$, so in particular $p - a \in \mathbb{Z} \backslash \{0\}$. We know that

$$\operatorname{Im}\left(e^{\frac{2\pi i (p-a)t}{X}}\right) = \sin\left(\frac{2\pi(p-a)t}{X}\right). \tag{5.3}$$

Since $\sin\phi = -\sin(-\phi) = -\sin(2\pi m - \phi)$ for $m \in \mathbb{Z}$, we can rewrite (5.3) to

$$\begin{aligned}
\mathrm{Im}\left(e^{\frac{2\pi i(p-a)t}{X}}\right) &= -\sin\left(-2\pi t\frac{p-a}{X}\right) \\
&= -\sin\left(2\pi(p-a) - 2\pi t\frac{p-a}{X}\right) \\
&= -\sin\left(2\pi\frac{p-a}{X}(X-t)\right) \\
&= -\mathrm{Im}\left(e^{\frac{2\pi i(p-a)(X-t)}{X}}\right)
\end{aligned}$$

Now follows that

$$\sum_{0<t<\frac{X}{2}} \mathrm{Im}\left(e^{\frac{2\pi i(p-a)t}{X}}\right) = -\sum_{0<t<\frac{X}{2}} \mathrm{Im}\left(e^{\frac{2\pi i(p-a)(X-t)}{X}}\right) = -\sum_{\frac{X}{2}<t<X} \mathrm{Im}\left(e^{\frac{2\pi i(p-a)t}{X}}\right).$$

Note that $e\left(\frac{(p-a)t}{X}\right) = 0$ for $t = 0$ and $t = \frac{X}{2}$ if X is even. Therefore we can now conclude that

$$\sum_{0\le t<X} \mathrm{Im}\left(e^{\frac{2\pi i(p-a)t}{X}}\right) = 0.$$

We also know that

$$\mathrm{Re}\left(e^{\frac{2\pi i(p-a)t}{X}}\right) = \cos\left(\frac{2\pi(p-a)t}{X}\right) = \sin\left(\frac{2\pi(p-a)t}{X} + \frac{\pi}{2}\right).$$

Now we can conclude in similar way similar tot that of the imaginary part that

$$\sum_{0\le t<X} \mathrm{Re}\left(e^{\frac{2\pi i(p-a)t}{X}}\right) = 0.$$

Together this gives us

$$\frac{1}{X}\sum_{0\le t<X} e\left(\frac{(p-a)t}{X}\right) = 0$$

if $X \nmid (p-a)$. Therefore we can now conclude that

$$\frac{1}{X}\sum_{0\le t<X} e\left(\frac{(p-a)t}{X}\right) = \begin{cases} 1 & \text{if } X|(p-a) \\ 0 & \text{if } X\nmid(p-a) \end{cases}.$$

Since we are looking for primes in $\mathcal{A}_k(X)$, we only look at values for a and p such that $a, p < X$. Therefore X only divides $(p-a)$ if $p - a = 0$. So

$$\frac{1}{X}\sum_{0\le t<X} e\left(\frac{(p-a)t}{X}\right) = \begin{cases} 1 & \text{if } p-a=0 \\ 0 & \text{if } p-a\neq 0. \end{cases}$$

In other words, this formula gives us a 1 if we have a prime in $\mathcal{A}_k(X)$ and a 0 otherwise. If we now sum over $a \in \mathcal{A}_k(X)$ and over $p \in \mathbb{P}$ with $p < X$, we get the number of primes in $\mathcal{A}_k(X)$. This is exactly what we defined d_X to be. We hence get

$$\begin{aligned}
d_X &= \sum_{a \in \mathcal{A}_k(X)} \sum_{\substack{p \in \mathbb{P} \\ p<X}} \frac{1}{X} \sum_{0 \le t < X} e\left(\frac{(p-a)t}{X}\right) \\
&= \frac{1}{X} \sum_{a \in \mathcal{A}_k(X)} \sum_{\substack{p \in \mathbb{P} \\ p<X}} \sum_{0 \le t < X} e\left(\frac{-at}{X}\right) \cdot e\left(\frac{pt}{X}\right) \\
&= \frac{1}{X} \sum_{0 \le t < X} \left(\sum_{a \in \mathcal{A}_k(X)} e\left(\frac{-at}{X}\right)\right) \cdot \left(\sum_{\substack{p \in \mathbb{P} \\ p<X}} e\left(\frac{pt}{X}\right)\right) \\
&= \frac{1}{X} \sum_{0 \le t < X} S_{\mathcal{A}_k(X)}\left(\frac{-t}{X}\right) S_{\mathbb{P}}\left(\frac{t}{X}\right).
\end{aligned}$$

which is exactly equal to (5.2), but no longer an approximation.

5.2 Define major and minor arcs

For the Ternary Goldbach Problem we defined a constant B. Similarly we define a constant $C \in \mathbb{R}_{>0}$ such that $P = (\log X)^C$, for sufficiently large X. Now let $\mathcal{V} = \left(\frac{P}{X}, 1 + \frac{P}{X}\right]$. We can again define for al $a, q \in \mathbb{N}$ such that $0 \le a \le q \le P$ and $(a,q) = 1$;

$$\mathfrak{M}_{a,q} = \left\{ t \in \mathcal{V} : \left| t - \frac{a}{q} \right| \le \frac{P}{X} \right\}$$

and let

$$\mathfrak{M} = \bigcup_{q \le P} \bigcup_{\substack{0 \le a \le q \\ (a,q)=1}} \mathfrak{M}_{a,q} \tag{5.4}$$

denote the set of major arcs. James Maynard does this in a slightly different manner since we work with a sum from 0 to X for the Restricted Digit Problem instead of an integral from 0 to 1. He defines the major arcs ([21], p. 17) as

$$\begin{aligned}
\mathfrak{M} &= \left\{ 0 \le s < X : \frac{s}{X} = \frac{a}{q} + O\left(\frac{(\log X)^C}{X}\right) \text{ for some } q \ll (\log X)^C \right\} \\
&= \bigcup_{q \ll P} \left\{ 0 \le s < X : \frac{s}{X} = \frac{a}{q} + O\left(\frac{P}{X}\right) \right\}. \qquad (5.5)
\end{aligned}$$

We know that $\frac{P}{X} > 0$, so $\left|\frac{P}{X}\right| = \frac{P}{X}$. Let $A \in \mathbb{R}_{>0}$, $t = \frac{s}{X}$ and thus $s = tX$. Then we can rewrite (5.5) as

$$\begin{aligned}
\mathfrak{M} &= \bigcup_{q \ll P} \left\{ 0 \leq t < 1 : \frac{s}{X} - \frac{a}{q} = D, \text{ with } |D| < A \cdot \left|\frac{P}{X}\right| \right\} \\
&= \bigcup_{q \ll P} \bigcup_{\substack{0 \leq a \leq q \\ (a,q)=1}} \left\{ 0 \leq t < 1 : \left| t - \frac{a}{q} \right| < A \cdot \frac{P}{X} \right\}.
\end{aligned}$$

Now we see this is almost the same as (5.4). The only differences are that $q \ll P$ instead of $q \leq P$, that $t \in [0, 1)$ instead of $t \in \mathcal{V}$ and that $\left|t - \frac{a}{q}\right| < A \cdot \frac{P}{X}$ instead of $\leq \frac{P}{X}$. The reason that Maynard uses a different definition for $\mathfrak{M}$, is that he defined a sum from 0 to X to compute an estimation for d_X instead of an integral from 0 to 1. Therefore we use (5.5) as definition for the major arcs. With this we can define the minor arcs by $\mathfrak{m} = [0, X) \backslash \mathfrak{M}$. This ensures that $\mathfrak{M} \cup \mathfrak{m} = [0, X)$, which is the interval over which the sum is taken.

Using (5.2), we can now write

$$\begin{aligned}
d_X &= \frac{1}{X} \sum_{0 \leq t < X} S_{\mathcal{A}_k(X)}\left(\frac{-t}{X}\right) S_{\mathbb{P}}\left(\frac{t}{X}\right) \\
&= \frac{1}{X} \left(\sum_{\mathfrak{m}} S_{\mathcal{A}_k(X)}\left(\frac{-t}{X}\right) S_{\mathbb{P}}\left(\frac{t}{X}\right) + \sum_{\mathfrak{M}} S_{\mathcal{A}_k(X)}\left(\frac{-t}{X}\right) S_{\mathbb{P}}\left(\frac{t}{X}\right) \right).
\end{aligned}$$

5.3 Fourier Estimates of $S_{\mathcal{A}_k(X)}(\theta)$

As said before, the reason that we can solve the Restricted Digit Problem even though it is a binary problem, is the exceptionally well controllable Fourier structure of $\mathcal{A}_k(X)$. We derive this Fourier structure in this section.

Before we estimate the behaviour of (5.2) on the major and minor arcs, we normalize $S_{\mathcal{A}_k(X)}(\theta)$. We hence define

$$F_X(t) = X^{-10 \log 9} \left| \sum_{n \in \mathbb{N}} \mathbf{1}_{\mathcal{A}_k(X)}(n) e(nt) \right|, \tag{5.6}$$

which is nothing but $\left|S_{\mathcal{A}_k(X)}(t)\right|$ with a factor $X-^{10} \log 9$ to normalize it. Recall that $X-^{10} \log 9 = \#\mathcal{A}_k(X)$.

Since we assumed X to be an integral power of 10, $X = 10^m$ for some $m \in \mathbb{N}$. Now we can rewrite (5.6) as follows,

$$\begin{aligned}
F_X(t) &= (10^m)^{10 \log \frac{1}{9}} \left| \sum_{n \in \mathcal{A}_k(X)} e(nt) \right| \\
&= 10^{10 \log \frac{1}{9^m}} \left| \sum_{n \in \mathcal{A}_k(X)} e(nt) \right| \\
&= \frac{1}{9^m} \left| \sum_{n \in \mathcal{A}_k(X)} e(nt) \right|.
\end{aligned}$$

Since $n \in \mathcal{A}_k(X)$ and so $n < X$, we know that $n = \sum_{j=0}^{m-1} n_j 10^j$. Filling in this definition in F_X gives us

$$\begin{aligned}
F_X(t) &= \frac{1}{9^m} \left| \sum_{n \in \mathcal{A}_k(X)} e\left(\sum_{j=0}^{m-1} n_j 10^j t \right) \right| \\
&= \frac{1}{9^m} \left| \sum_{n_0, \ldots, n_{m-1} \in \{0, \ldots, 9\} \setminus \{k\}} e\left(\sum_{j=0}^{m-1} n_j 10^j t \right) \right| \\
&= \frac{1}{9^m} \left| \sum_{n_0, \ldots, n_{m-1} \in \{0, \ldots, 9\} \setminus \{k\}} e^{2\pi i \cdot \sum_{j=0}^{m-1} n_j 10^j t} \right| \\
&= \frac{1}{9^m} \left| \sum_{n_0, \ldots, n_{m-1} \in \{0, \ldots, 9\} \setminus \{k\}} \prod_{j=0}^{m-1} e\left(n_j 10^j t \right) \right| \\
&= \frac{1}{9^m} \left| \sum_{n_0, \ldots, n_{m-1} \in \{0, \ldots, 9\} \setminus \{k\}} e\left(n_0 10^0 t \right) \cdot e\left(n_1 10^1 t \right) \cdot \ldots \cdot e\left(n_{m-1} 10^{m-1} t \right) \right| \\
&= \prod_{j=0}^{m-1} \left| \frac{1}{9} \sum_{n_j \in \{0, \ldots, 9\} \setminus \{k\}} e\left(n_j \cdot 10^j t \right) \right|.
\end{aligned}$$

We know that $\sum_{k=0}^{n-1} ar^k = a\frac{r^n-1}{r-1}$ for $r \neq 1$. Using this, we now write

$$\begin{aligned}
F_X(t) &= \prod_{j=0}^{m-1} \left| \left(\sum_{n_j=0}^{9} \frac{1}{9} \left(e^{2\pi i 10^j t} \right)^{n_j} \right) - \frac{1}{9} e^{2\pi i k 10^j t} \right| \\
&= \prod_{j=0}^{m-1} \left| \frac{1}{9} \frac{\left(e^{2\pi i 10^j t} \right)^{10} - 1}{e^{2\pi i 10^j t} - 1} - \frac{1}{9} e^{2\pi i k 10^j t} \right| \\
&= \prod_{j=0}^{m-1} \left| \frac{e(10^{j+1}t) - 1}{9\left(e(10^j t) - 1\right)} - \frac{1}{9} e(k \cdot 10^j t) \right|
\end{aligned}$$

Note that this is a multiplicative formula, which can be used to find an estimate for the minor arcs. Although it is quite involved, it is with most problems possible to estimate the major arcs in a straight forward manner. The minor arcs are the difficult part. That is the part where the circle method gets stuck when applied to the Binary Goldbach Problem. In the case of the Restricted Digit Problem, this multiplicative Fourier structure of $\mathcal{A}_k(X)$ is the key ingredient for the bound on the minor arcs.

5.4 Estimate the major and minor arcs

In this section we give some important ingredients for the estimation of the major and minor arcs for the Resticted Digit Problem.

5.4.1 Minor arcs

Maynard divides the minor arcs into two categories, the "Generic minor arcs" and the "Exceptional minor arcs". To give the estimate for the minor arcs, we need to introduce a few notions. Let $0 < \eta < 1$ and let $l_1, l \in \mathbb{N}_{>0}$ such that $l_1 \leq l \ll \eta^{-1}$. Now we define $\mathcal{R} \subseteq [\eta, 1]^l$ to be a convex polytope in $\mathbb{R}^l$ independent of X, such that

$$\mathbf{e} \in \mathcal{R} \Rightarrow \sum_{i=1}^{l_1} e_i \in \left[\frac{9}{25} + \epsilon, \frac{17}{40} - \epsilon \right]. \tag{5.7}$$

We also define the following function.

$$\Lambda_{\mathcal{R}}(n) = \sum_{\substack{p_1 \cdot \ldots \cdot p_l \\ \left(\frac{\log p_1}{\log X}, \ldots, \frac{\log p_l}{\log X} \right) \in \mathcal{R}}} \prod_{i=1}^{l} \log p_i.$$

Using this notion we now define $S_{\mathcal{R}}(\theta) = \sum_{n<X} \Lambda_{\mathcal{R}}(n) e(n\theta)$.

It turns out that the following theorem holds.

Theorem 10 *([21], Section 8, p. 21, Lemma 8.2)*
Let
$$\mathcal{E} = \left\{ 1 \le t < X : F_X\left(\frac{t}{X}\right) \ge \frac{1}{x^{\frac{23}{80}}} \right\}$$
[be the set of exceptional minor arcs]. Then $\mathcal{E} \ll X^{\frac{23}{40}-\epsilon}$*,*
$$\sum_{t\in\mathcal{E}} F_X\left(\frac{t}{X}\right) \ll X^{\frac{23}{80}-\epsilon},$$
and
$$\frac{1}{X}\sum_{\substack{t<X\\ t\notin\mathcal{E}}} \left| F_X\left(\frac{t}{X}\right) S_{\mathcal{R}}\left(\frac{-t}{X}\right)\right| \ll_\eta \frac{1}{X^\eta}.$$

Since F_X is a normalized version of $S_{\mathcal{A}_k(X)}$, this is not the final result we need. The denormalized and more general version can be found in the next Theorem.

Theorem 11 (Generic minor arcs) *([21], Section 6, p. 17, Proposition 6.3)*
Let $l \ll_\eta 1$ *and* $\mathcal{R}[...]$ *[as defined in (5.7)]. There is some exceptional set* $\mathcal{E} \subseteq [1, X]$ *with*
$$\#\mathcal{E} \ll X^{\frac{23}{40}},$$
such that
$$\frac{1}{X}\sum_{\substack{t<X\\ t\notin\mathcal{E}}} \left| S_{\mathcal{A}_k(X)}\left(\frac{t}{X}\right) S_{\mathcal{R}}\left(\frac{-t}{X}\right)\right| \ll_\eta \frac{\#\mathcal{A}_k(X)}{X^\epsilon}.$$

For the Exceptional minor arcs we again need to introduce some notation. Let $\delta = (\log\log X)^{-1}$, and let $\mathcal{R} = \mathcal{R}(a_1, ..., a_{l-1})$ be given by
$$\mathcal{R} = \left\{ \mathbf{e} \in \mathbb{R}^l : e_i \in [a_i, a_i + \delta^2] \text{ for } 1 \le i \le l-1, \sum_{j=1}^{l} e_i \in [1-\delta, 1] \right\},$$
for some constants $a_1, ..., a_{l-1}$ satisfying $\min_i(a_i) \ge \frac{\eta}{2}$ and $\sum_{i=1}^{l-1} < 1 - \frac{\eta}{2}$ and $l \ll_\eta 1$. Let $\mathcal{M} = \mathcal{M}(C)$ denote the major arcs defined in (5.5). Using this, the following theorem gives the estimate for the Exceptional minor arcs.

Theorem 12 (Exceptional minor arcs) *([21], Section 6, p. 17, Proposition 6.4)*
Let η*,* l*,* $\mathcal{R} = \mathcal{R}(a_1, ..., a_{l-1})$ *and* $\mathcal{M}(C)$ *be as given [before][...]. Let* $a_1, ..., a_{l-1}$ *in the definition of* $\mathcal{R}$ *satisfy* $\sum_{i=1}^{l_1} a_i \in [\frac{9}{25} + \epsilon, \frac{17}{40} - \epsilon]$ *for some* $l_1 < l$*, and let* $C = C(A, \eta)$ *in the definition of* $\mathcal{M}$ *be sufficiently large in terms of* A *and* η*. Let* $\mathcal{E} \subseteq [1, X]$ *be any set such that* $\#\mathcal{E} \ll X^{\frac{23}{40}}$*. Then we have*
$$\frac{1}{X}\sum_{\substack{t\in\mathcal{E}\\ t\notin\mathcal{M}}} S_{\mathcal{A}_k(X)}\left(\frac{t}{X}\right) S_{\mathcal{R}}\left(\frac{-t}{X}\right) \ll_\eta \frac{\#\mathcal{A}_k(X)}{(\log X)^A}.$$

In particular this holds for the the set of minor arcs $\mathcal{E}$ we defined in Theorem 10.

5.4.2 Major arcs

For the Major arcs, Maynard finds the theorem stated below.

Theorem 13 (Major arcs) *([21], Section 6, p. 17, Proposition 6.2)*
Let [η, l, $\mathcal{R} = \mathcal{R}(a_1, ..., a_{l-1})$ and $\mathcal{M}$ be given as before.] [...] Then

$$\frac{1}{X} \sum_{\substack{0 \le t < X \\ t \in \mathcal{M}}} S_{\mathcal{A}_k(X)}\left(\frac{t}{X}\right) S_{\mathcal{R}}\left(\frac{-t}{X}\right) = \kappa_2 \frac{\#\mathcal{A}_k(X)}{X^A} \sum_{n<X} \Lambda_{\mathcal{R}}(n) + O_{C,\eta}\left(\frac{\#\mathcal{A}_k(X)}{(\log x)^C}\right).$$

Here κ_2 is the constant given [...] [by

$$\kappa_2 = \begin{cases} \frac{10(\phi(10)-1)}{9\phi(10)} & \textit{if } (10, a_0) = 1 \\ \frac{10}{9} & \textit{otherwise.]} \end{cases}$$

Combination of this theorem with Theorem 11 and 12 leads in the end to

$$d_X = \#\{p \in \mathcal{A}_k(X)\} \ge (c + o(1)) \frac{\#A}{\log X}$$

for some constant c. Maynard uses numerical integration on the base of a Markov process to make sure that $c > 0$. With that Theorem 7 follows.

Chapter 6

Conclusion

In the previous chapters we compared the circle method applied to the Goldbach Problem and to the Restricted Digit Problem.
We know that the Weak Goldbach Conjecture leads to a ternary problem, since we look for sums of three primes p_1, p_2, p_3 such that $n = p_1 + p_2 + p_3$. This leads to expression (3.7) when we apply the circle method. This expression contains the product of three functions, namely $(g(t))^3$.
In the Binary Goldbach Problem we are looking for sums of two primes p_1, p_2 such that $n = p_1 + p_2$. Therefore the expression for a_n turns out to contain the product of two functions, $g(t)^2$.
Theorem 7 is about a binary problem as well. After all we are looking for numbers for which $0 = p - a$ holds. We would thus expect to find an expression containing the product of two functions. This is indeed true, since we get

$$d_X = \int_0^1 \left(\sum_{a \in \mathcal{A}_k(X)} e(-at) \right) \cdot \left(\sum_{\substack{p \in \mathbb{P} \\ p < X}} e(pt) \right) dt \tag{6.1}$$

in (5.1). We see that this is the integral of the product of $g(t)$ with another function. Now let

$$h(t) = \sum_{a \in \mathcal{A}_k(X)} e(-at).$$

This is the function that represents the requirement that one of the parameters a is an element of $\mathcal{A}_k(X)$. So we have the product of g with h in (6.1) instead of the g^2 we have for the Binary Goldbach Problem. This is what creates the difference between the Binary Goldbach Problem and the Restricted Digit Problem.

In Chapter 3 we have seen that it is possible to find a good estimate for $R(n)$ on both the major and minor arcs for the Ternary Goldbach Problem. In a similar way we can find a bound for the Binary Goldbach Problem on the major arcs. On the minor arcs however we cannot use a similar strategy to the one used for the Ternary Goldbach Problem. Since the Binary Goldbach Problem and the Restricted Digit Problem are both binary problems, we would expect them to be comparable in some way. In Chapter 5 we have seen that it is possible to find a convenient bound for the minor arcs in the Restricted Digit Problem nonetheless. This is quite interesting, since $\mathcal{A}_k(X)$ is a much sparser set than $\mathbb{P}$. In general it is more difficult to solve a problem involving sparser sets than to solve a problem involving less sparse sets, unless the sparser set has much more structure. This is the case with $\mathcal{A}_k(X)$. After all its Fourier structure yields a nice multiplicative formula, which we can use to give a bound for d_X on the minor arcs. This is an important reason why Theorem 7 is proven and Conjecture 2 not.

To use this nice multiplicative Fourier structure of $\mathcal{A}_k(X)$, we need to use a sum instead of an integral to compute an estimation for d_X. This leads to slightly different major arcs. Another difference is that we need to define two kinds of minor arcs to solve the Restricted Digit Problem, the Generic minor arcs and the exceptional minor arcs. Altogether this leads to a quite technical and extensive proof. To make it work, Maynard uses the Harman's seive, Large seive and a Markov process, which are all tools in analytic number theory.

Bibliography

[1] W.D. Banks, A. Conflitti and I.E. Shparlinski, *Character sums over integers with restricted g-ary digits*, Illinois J. Math. **46(3)** (2002), 819-836

[2] W.D. Banks an I.E. Shparlinski, *Arithmetic properties of numbers with restricted digits*, Acta Arithmetica **112(3)** (2004), 313-332

[3] W.W.L. Chen *Chapter 1 Waring's Problem*, (1997, 2013), available on: `https://rutherglen.science.mq.edu.au/wchen/lnhlmfolder/hlm01.pdf`, consulted on 29 March 2017

[4] *Chapter 2 Goldbach's Problem*, (1997, 2013) available on: `https://rutherglen.science.mq.edu.au/wchen/lnhlmfolder/hlm02.pdf`, consulted on 30 March 2017

[5] J.R. Chen, *On the representation of a larger even integer as the sum of a prime and the product of at most two primes*, Sci. Sinica. **16** (1973), 157-176.

[6] W.Cole, *The Evolution of the Circle Method in Additive Prime Number Theory*, available on `http://www.maa.org/sites/default/files/images/upload_library/46/HOMSIGMAA/1-evo-prime-wcole.pdf` consulted on 30 March 2017

[7] J.G. van der Corput, *Sur l'hypothèse de Goldbach pour presque tous les nombres premiers*, Acta Arithmetica **2** (1937), 266-290.

[8] M. Drmota and C. Maudit, *Weyl sums over integers with affine digit restrictions*, J. Number Theory **130(11)** (2010), 2404-2407

[9] P. Erdős, c. Maudit and A. Sárközy, *On Arithmetic Properties of Integers with Missing Digits I: Distribution in residue classes*, J. Number Theory **70(2)** (1998), 99-120

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

[11] A. O. Gelfond, *Sur les nombres qui ont des propriétés additives et multiplicatives donneées*, Acta Arithmetica **13** (1968), 259-265.

[12] H.A. Helfgott, *Minor arcs for Goldbach's theorem*, arXiv:1205.5252v4 [math.NT] (2013)

[13] H.A. Helfgott, *Major arcs for Goldbach's theorem*, arXiv:1305.2897v4 [math.NT] (2014)

[14] H.A. Helfgott, *The Ternary Goldbach conjecture is true*, arXiv:1312.7748v2 [math.NT] (2013)

[15] 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** (1922), 1-70.

[16] R.C. Vaughan, *The Hardy-Littlewood method, second edition*, Cambridge University Press, Cambridge Tracts in Mathematics. **125** (1997)

[17] H.A. Helfgott and D.J. Platt, *Numerical verification of the ternary goldbach conjecture up to* $8.875 \cot 10^{30}$, arXiv:1305.3062v2 [math.NT] (2014)

[18] G.H.Hardy and S.Ramanujan, *Asymptotic Formulae in Combinatory Analysis*, Proc. London Math. Z. **17** (1918), 75-115.

[19] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, American Mathematical Society, Colloquium Publications. **v. 53** (2004)

[20] S. Lang, *Complex Analysis*, Springer, Graduate Texts in Mathematics**103** (2000)

[21] J. Maynard, *Primes with restricted digits*, arXiv:1604.01041 [math.NT] (2016)

[22] C. Mauduit and A. Sárközy, *On the arithmetic structure of the integers whose sum of digits is fixed*, Acta Arithmetica **81** (1997), 145-173.

[23] T. Tao, *Every odd number greater than 1 is the sum of at most five primes*, arXiv:1201.6656 [math.NT] (2012)

[24] N.G. Tchudakov, *Sur le problème de Goldbach*, C.R. (Dokl.) Acad. Sci. URSS, n. Ser. **17** (1937), 335-338.

[25] I.M. Vinogradov, *Representation of an odd number as a sum of three primes*, Dokl. Akad. Nauk SSSR **15** (1937), 291-294.

[26] E. Warner, *An easy introduction to the circle method.* available on `https://web.stanford.edu/~ebwarner/CircleMethod.pdf` consulted on 29 March 2017

[27] T. Yamada, *Explicit Chen's Theorem*, arXiv:1511.03409v2 [math.NT] (2015)