Universiteit Utrecht

Bachelor Thesis

Primes and Arithmetic progressions

Author:
Pol van Hoften
4001613

Supervisor:
Prof. dr. F. Beukers

January 27, 2015

Achknowledgements

This thesis would not have been possible without the help of several individuals who contributed to the completion of this work.

First and foremost I would like to thank prof. dr. F. Beukers for his guidance and encouraging words. Secondly I would like to thank my friends and my girlfriend for keeping me sane and happy.

Contents

1 Introduction

1.1 Arithmetic progressions

This thesis will be centred upon the following concept: An *arithmetic progression*(AP) is a sequence of numbers with constant difference between consecutive terms. For instance, consider the following arithmetic progression of length 3 with common difference 6:

$$11, 17, 23.$$

In general, an arithmetic progression of length k and common difference d is a sequence of numbers $a, a+d, a+2d, \cdots, a+(k-1)d$. It is clear that arithmetic progressions with common difference 0 aren't very interesting, so we shall call those *trivial*. The existence of non-trivial arithmetic progressions in certain subsets of the natural numbers (or integers) is quite an interesting subject. For example, how big can a set $A \subset \mathbb{N}$ be, without containing an arithmetic progression of length k? Do the primes contain arithmetic progressions of length k (and how many)? How many arithmetic progression of length k, does a 'random' set of size M contain?

1.2 History

The Dutch mathematician Van der Waerden showed in 1927 that if one colours the integers with r colours, that there is, for every $k \in \mathbb{N}$, a monochromatic arithmetic progression of length k. Erdős and Turan then conjectured that a stronger result might hold. Given a number $N \in \mathbb{N}$, let $r_k(N)$ denote the size of the biggest subset of $[N] := \{1, \cdots, N\}$ that has no k term arithmetic progressions. Erdős and Turan conjectured that $\lim_{N\to\infty} \frac{r_k(N)}{N} = 0$. In 1953, Klaus Roth proved his (second-most) famous Theorem which says that $\lim_{N\to\infty} \frac{r_3(N)}{N} = 0$ ([Rot53]). This is the first result I shall prove in this thesis. In 1969 [Sze69] Szemerédi showed that the same holds for $r_4(N)$ and in 1975 [Sze75] he proved that $\lim_{N\to\infty} \frac{r_k(N)}{N} = 0$ for all k, using a very complicated combinatorial argument.

Erdős and Turan later generalised their conjecture:

Conjecture 1.1 (Erdős-Turan)**.** *Let $A \subset \mathbb{N}$, if*

$$\sum_{n\in A} \frac{1}{n}$$

diverges, then A contains arbitrarily long arithmetic progressions.

This generalised conjecture contains as a special case $A = P$, the set of primes. The Dutch mathematician Van der Corput showed in the 1930's that the primes contain infinitely many arithmetic progressions of length three. This is the second result that I shall prove in this thesis. Finally in 2003, Green and Tao([GT08]) showed that the primes contain arbitrarily long arithmetic progressions.

Theorem 1.1 (Green-Tao)**.** *Let $k \in \mathbb{N}$, then there is an arithmetic progression of length k in the primes.*

1.3 Heuristics

Consider the set $[N]$, we can calculate how many arithmetic progressions of length 3 (3AP's) $[N]$ contains quite easily. Say we know the number of 3AP's in $[N]$, denoted by $f(N)$. We want to know the number of 3AP's in $[N+1]$. We know that the number of arithmetic progressions in $\{1, \cdots, N\}$ is the same as the number of arithmetic progressions in $\{2, \cdots, N+1\}$. So $f(N+1) = f(N) + s(N+1)$, where $s(N+1)$ is the number of arithmetic progressions in $[N+1]$ containing the number 1, since we haven't counted those yet. Using induction and the fact that $s(1) = 1$ we obtain

$$f(N) = \sum_{n=1}^{N} s(n).$$

All that remains is determining $s(n)$, which shall prove to be easy. Notice that $s(1) = 1, s(2) = 1, s(3) = 2$. For every odd number $2j+1 \leq n$, the progression $1, n+1, 2n+1$ is available, and all progressions containing 1 are of this form. This shows that $s(n) = 1 + \lfloor \frac{2n-1}{2} \rfloor$. Summing over n we obtain

$$\begin{aligned} f(N) &= \sum_{n=1}^{N} s(n) \\ &= \sum_{n=1}^{N} 1 + \lfloor \frac{2n-1}{2} \rfloor, \\ &= \sum_{n=1}^{N} \lfloor \frac{2n+1}{2} \rfloor, \\ &= \sum_{n=1}^{N} \left(\frac{2n+1}{2} + O(1) \right), \\ &= \frac{N^2}{4} + O(N). \end{aligned}$$

This result allows us to predict the number of 3AP's in subsets of $[N]$. Let $0 < \delta < 1$ and form the set A by picking δN elements uniformly at random from $[N]$. Every arithmetic progression has a chance of δ^3 to lie in A and the expected number of arithmetic progressions in A is $\frac{\delta^3 N^2}{4}$. As it turns out, a set $A \subset [N]$ of size δN will contain a 3AP if δ is not too small (we will make this precise in a bit).

If we consider P_N, the set of primes up to N, then we can also apply the above procedure (even though the primes aren't 'random') using the probabilistic model invented by Cramer [Cra35]. By the prime number theorem there are about $\frac{N}{\log N}$ primes up to N, so every number between 0 and N has a 'chance'of $\frac{1}{\log N}$ to be prime. Now consider an arithmetic progression of length three, the chance that it lies in P_N is $\frac{1}{(\log N)^3}$ so there are about $\frac{N^2}{4(\log N)^3}$ arithmetic progression in the primes up to N. This is surprisingly close to the truth, it is possible to improve this prediction by considering the fact that the common difference cannot be odd (then one the elements must be even) and other similar behaviours.

The rest of this thesis will be committed to proving the following two theorems, section 2 will be dedicated to the proof of the first theorem and section 3 will be dedicated to the proof of the second.

Theorem 1.2 (Roth(1953)). *Let $\delta > 0$ and $N \in \mathbb{N}$, let A be a subset of $[N]$. Then there exists an absolute constant c, such that if $N > \exp(\exp(\frac{c}{\delta}))$ and $|A| \geq \delta N$, then A contains at least one 3AP.*

Theorem 1.3 (Van der Corput (1939)). *Let $A > 0$ be a real number, and let $n \in \mathbb{N}$, then the number of 3AP's in the primes up to n, denoted by $R(n)$, satisfies*

$$R(n) \geq c\frac{n^2}{4(\log n)^3} + O\left(\frac{n^2}{(\log n)^A}\right),$$

where $c \geq 1$

1.4 Prerequisites and organisation

No prior knowledge is required to understand the proof of Roth's theorem in section 1.2, it is almost completely self contained. The proof of the theorem of Van der Corput in section 3 assumes a little bit more; knowledge of (multiplicative) arithmetical functions is assumed but will be included in the appendix, as well as some basic identities concerning these functions. We have referred to the classical book of Hardy and Wright [HW08] for most of the proofs, but it is not necessary to understand them to be able to follow the proof. Furthermore, we use the Theorem of Siegel and Walfiesz as a black box in the proof.

2 Roth's theorem.

The approach taken in this section has been adapted from Lyall's expository article [Lya05], the only original work done by the author is in the appendix.

2.1 Preliminaries

In this section, we will look at sets $A \subset [N] := \{1, \cdots, N\}$, with *density* $\delta > 0$, i.e., $|A| = \delta N$. We will generally fix δ and let $N = N(\delta)$ depend on δ. We will attempt to find a condition on N depending on δ or equivalently a condition on δ depending on N that guarantees the existence of an arithmetic progression of length 3 (often abbreviated by 3AP) in sets $A \subset N$ of density δ. Finding arithmetic progressions of length 3 in such a set A amounts to solving the equation $x + y = 2z$ there.

We will embed the problem in $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$ by considering solutions to this equation in $\mathbb{Z}_N$. We do lose some information about our problem, since $\mathbb{Z}_N$ contains nontrivial solutions to the above equation that are not solutions in $\mathbb{Z}$, for example $4, 8, 1$ is an arithmetic progression in $\mathbb{Z}_{11}$ but not in $\mathbb{Z}$. To make the difference clear we shall use the term 3AP $\mod N$ for solutions of the equation $x + y = 2z \mod N$. The term genuine 3AP is sometimes used for solutions in $\mathbb{Z}$, to emphasise the difference. All of this turns out to be a minor technicality, we can differentiate between 3AP's $\mod N$ and genuine 3AP's by a trick explained later.

The purpose of translating our problem to $\mathbb{Z}_N$ is the fact that we can do (finite) Fourier transformations there. We will generally assume N to be odd to guarantee the two-divisibility of $\mathbb{Z}_N$, i.e., that the map $x \mapsto 2x$ is bijective.

Define the *Fourier transform* of a function $f : \mathbb{Z}_N \to \mathbb{C}$ as

$$\hat{f}(k) =: \sum_{x=0}^{N-1} f(x) e^{-\frac{2\pi i}{N} xk},$$

where we will often write $\omega = e^{\frac{2\pi i}{N}}$ for convenience. The number $\hat{f}(k)$ will sometimes be referred to as the k-th Fourier coefficient. In the rest of this section we will prove some properties of the Fourier transform and we will recall some basic analysis. From now one all sums are, unless otherwise indicated, over $\mathbb{Z}_N$ (where N is fixed). Readers who have previously seen (finite) Fourier analysis can safely skip the proofs in this section and move on to section 2.2.

Proposition 2.1 (Properties of the Fourier transform)**.** *The Fourier transform of functions $f, g : \mathbb{Z}_N \to \mathbb{C}$ has the following properties:*

1. *The 0-th Fourier coefficient $\hat{f}(0) = \sum_x f(x)$ and $|\hat{f}(k)| \leq \sum_x |f(x)|$.*
2. *The Fourier transform commutes with convolutions*

$$\widehat{\left(\sum_y f(y)\overline{g(y-x)}\right)}(k) = \hat{f}(k)\bar{\hat{g}}(k),$$

where $\bar{c}$ is used to denote complex conjugation.

3. The orthogonality relation:

$$\frac{1}{N}\sum_k \omega^{-xk} = \begin{cases} 1 & \text{if } x \equiv 0 \mod N \\ 0 & \text{if otherwise} \end{cases} \tag{1}$$

4. Plancherel's theorem:

$$\sum_x |f(x)|^2 = \frac{1}{N}\sum_k |\hat{f}(k)|^2$$

5. The Cauchy-Schwarz inequality:

$$\sum_k |a_k b_k| \leq \left(\sum_k |a_k|^2\right)^{\frac{1}{2}} \left(\sum_k |b_k|^2\right)^{\frac{1}{2}}$$

Proof. 1. The first follows directly from the definition and the second from the fact that $|\omega^{-xk}| = 1$ and the triangle inequality..

2. Expanding the definitions gives

$$\begin{aligned} \hat{f}(k)\overline{\hat{g}(k)} &= \left(\sum_x f(x)\omega^{-xk}\right)\left(\sum_y \overline{g(y)}\omega^{yk}\right), \\ &= \sum_x \sum_y f(x)\overline{g(y)}\omega^{k(y-x)}, \\ &= \sum_x f(x) \sum_z \overline{g(x-z)}\omega^{-zk}, \qquad (z = x-y) \\ &= \sum_z \sum_x f(x)\overline{g(x-z)}\omega^{-zk}, \\ &= \widehat{\left(\sum_y f(y)\overline{g(y-x)}\right)}(k). \end{aligned}$$

This might look strange but we are considering $\sum_y f(y)(\overline{g(y-x)})$ as a function of x here, and we Fourier transform it as a function of x (in other words, the Fourier transform hat is supposed to be applied to the entire sum in the parenthesis).

3. It is clear that if $x = 0$ that $\sum_k \omega^{xk} = N$. If $x \neq 0$ we can find y such that $\omega^y \neq 1$, then

$$\omega^y \sum_k \omega^{xk} = \sum_k \omega^{xk+y},$$
$$\sum_k \omega^{xk}.$$

And so either $\omega^y = 1$ (but that's impossible by construction) or $\sum_k \omega^{xk} = 0$, which is what we wanted.

4. We can verify this directly

$$\begin{aligned}\sum_k |\widehat{f}(k)|^2 &= \sum_k \left|\sum_x f(x)\omega^{xk}\right|^2, \\ &= \sum_k \left(\sum_x f(x)\omega^{-xk}\right)\left(\sum_x \overline{f(x)}\omega^{xk}\right), \\ &= \sum_k \sum_{a-b\leq N-1} f(a)\overline{f(b)}\omega^{k(a-b)}, \\ &= \sum_{a-b\leq N-1} f(a)\overline{f(b)}\sum_k \omega^{k(a-b)},\end{aligned}$$

we now apply (1) and see

$$\begin{aligned}\sum_k |\widehat{f}(k)|^2 &= N \sum_{\substack{a-b\leq N-1 \\ ab=0 \mod N}} f(a)\overline{f(b)}, \\ &= N\sum_x |f(x)|^2.\end{aligned}$$

5. The proof can be found in any analysis text and shall not be replicated here.

□

2.2 Strategy of the proof

We have seen in the introduction that we expect a set $A \subset [N]$ chosen uniformly at random with density δ to have $\delta^3 N^2$ 3AP's. We will attempt to make this heuristic precise by defining a notion of (pseudo)-randomness for subsets of $\mathbb{Z}_N$ and then showing that sets that are sufficiently random in this way do indeed contain 3AP's $\mod N$.

Afterwards, we're going to run an argument by contradiction. We will assume that there is a set of density δ that does not have any 3AP. This set will then not be sufficiently random and we can use this fact to show that there is a (long) arithmetic progression P on which A has density $\delta + \epsilon$ for some $\epsilon > 0$. Since P is an arithmetic progression elements p of P are of the form $p = k \cdot a + b$ where a is the common difference and b smallest element of P. We can identify P with $P' = \{0, 1, \cdots, |P| - 1\}$ under the transformation $(ak + b) \mapsto k$,. It turns out this this transformation preserves arithmetic progressions. We can now form a new set $A_1 = T(A \cap P)$, which will still have 0 arithmetic progressions. Since A has density $\delta + \epsilon$ on P, our new set A_1 has density $\delta + \epsilon$ on P'.

Now since A_1 contains no arithmetic progressions, it must not be sufficiently random and so there must exist a (long) arithmetic progression P_1 on which A_1 has density $\delta + 2\epsilon$, etc.

If we iterate this argument long enough, eventually A_k has density greater than 1 on a nonzero arithmetic set P_k and this is impossible.

2.3 Defining (pseudo)-randomness.

Definition 2.1. *Let $\epsilon > 0$. We say that A is ϵ-uniform if $|\widehat{A}(k)| \leq \epsilon N$ for all $1 \leq k \leq N-1$ where*

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

is the indicator- or characteristic function of A.

We will now show that ϵ-uniform sets do indeed contain many arithmetic progressions $\bmod N$. Consider solutions to the equation $x+y-2z = 0 \mod N$, then using the orthogonality relation we can count these solutions using the sum

$$\begin{aligned}
\sum_x \sum_y \sum_z \frac{1}{N} \sum_{k=0}^{N-1} \omega^{-(x+y-2z)k} &= \sum_k \frac{1}{N} \left(\sum_x A(x)\omega^{-xk} \right) \left(\sum_x A(x)\omega^{-yk} \right) \left(\sum_x A(x)\omega^{-z(-2k)} \right), \\
&= \sum_k \frac{1}{N} \widehat{A}(k)\widehat{A}(k)\widehat{A}(-2k), \\
&= \frac{|A|^3}{N} + \frac{1}{N} \sum_{k=1}^{N-1} \widehat{A}(k)^2 \widehat{A}(-2k), \\
&= \delta^3 N^2 + \frac{1}{N} \sum_{k=1}^{N-1} \widehat{A}(k)^2 \widehat{A}(-2k).
\end{aligned}$$

Perhaps unsurprisingly, we find the $\delta^3 N^2$ that our heuristic argument predicted, we just need to get rid of the error terms. If we could show that the error is smaller than $\delta^3 N^2$ we would have more than zero arithmetic progressions mod N. This is where we use the assumption that A is ϵ-uniform to see that

$$\begin{aligned}
\frac{1}{N} \left| \sum_{k=1}^{N-1} \widehat{A}(k)^2 \widehat{A}(-2k) \right| &\leq \frac{1}{N} \max_{k \neq 0} |\widehat{A}(-2k)| \sum_{k=1}^{N-1} |\widehat{A}(k)|^2, \\
&\leq \frac{1}{N} \max_{k \neq 0} |\widehat{A}(-2k)| \sum_{k=0}^{N-1} |\widehat{A}(k)|^2, \\
&= \frac{1}{N} \max_{k \neq 0} |\widehat{A}(-2k)| N \sum_{x=0}^{N-1} |A(x)|^2, \qquad \text{(by Plancherel's Theorem)} \\
&\leq \epsilon N \delta N,
\end{aligned}$$

and this is surely smaller than $\delta^3 N^2$ for $\epsilon(\delta)$ chosen sufficiently small.

2.4 The (pseudo-)random sets

We shall first need to pass from 3AP's $\bmod N$ to genuine arithmetic progressions. This can be achieved by choosing arithmetic progressions $\bmod N$ that lie close enough to $N/2$. We claim that if $x + y = 2z \bmod N$ and $x, z \in M_A := A \cap [\frac{N}{3}, \frac{2}{3}N)$, then $x + y = 2z$ also in $\mathbb{Z}$. This is true because $y = z + |z - y| =$

$x + |x - z| < \frac{2}{3}N + \frac{N}{3}$ and $y = z + |z - y| > \frac{N}{3} + \frac{N}{3}$ so $\frac{2}{3}N < y < N$ (here we sneakily identify the equivalence class $[y] \mod N$ with its smallest positive representative $y \in \mathbb{Z}$). Using this fact we prove the following Lemma:

Lemma 2.1. *If A is ϵ-uniform for $\epsilon < \frac{\delta^2}{8}$, $N > \frac{8}{\delta^2}$ and $|M_A| \geq \frac{\delta}{4}N$, then there is at least one non-trivial 3AP in A.*

Proof. We perform the same calculation as in the previous section and let M denote the number of 3AP's in A, we only get a lower bound for M since there are 3AP's that are not in M_A.

$$\begin{aligned} M &\geq \frac{1}{N}\sum_{k=0}^{N-1} \widehat{M_A}(k)\widehat{A}(k)\widehat{M_A}(-2k) \\ &= \delta|M_A|^2 + \frac{1}{N}\sum_{k=1}^{N-1} \widehat{M_A}(k)\widehat{A}(k)\widehat{M_A}(-2k), \end{aligned}$$

using the fact that $\widehat{f}(0) = \sum f(x)$ proven in Proposition 2.1. Now we are left to deal with the error term, which can be taken caren of by using the Cauchy-Schwarz inequality and Plancherel's theorem proven in Proposition 2.1.

$$\begin{aligned} \left|\sum_{k=1}^{N-1} \widehat{M_A}(k)\widehat{A}(k)\widehat{M_A}(-2k)\right| &\leq \max_{k\neq 0}|\widehat{A}(k)| \sum_{k=1}^{N-1} |M_A(k)M_A(-2k)|, \\ &\leq \epsilon N \sum_{k=1}^{N-1} |\widehat{M_A}(k)\widehat{M_A}(-2k),| \\ &\leq \epsilon N \left(\sum_k |\widehat{M_A}(k)|^2\right)^{\frac{1}{2}} \left(\sum_k |\widehat{M_A}(-2k)|^2\right)^{\frac{1}{2}}, \\ &= \epsilon N \sum_k |\widehat{M_A}(k)|^2, \\ &= \epsilon N^2 \sum_x |M_A(x)|^2, \text{ (by Plancherel's Theorem)} \\ &= \epsilon N^2 \sum_x |M_A(x)|, \\ &\leq \epsilon N^2 |M_A|, \end{aligned}$$

using the fact that $x \mapsto -2x$ is a bijection from $\mathbb{Z}_N \to \mathbb{Z}_N$ because we assumed that N is odd. Now taking $\epsilon < \frac{\delta^2}{8}$ we see that

$$\begin{aligned} M &> \delta|M_A|^2 - \epsilon N|M_A|, \\ &\geq \frac{\delta^3}{16}N^2 - \frac{\delta^3}{32}N^2, \\ &= \frac{\delta^3}{32}N^2. \end{aligned}$$

We should not forget the existence of trivial 3 AP's , i.e., solutions $x = y = z$. It is true, however, that if we pick $N \geq \frac{8}{\delta^2}$ there always is atleast one 3AP (there are only δN trivial solutions).

$$\begin{aligned}\mathcal{N} = M - \delta N &\geq \frac{\delta^3}{32}N^2 - \delta N, \\ &\geq 2\delta N - \delta N, \\ &> 0.\end{aligned}$$

□

Now we shall turn the contrapositive of Lemma 2.1 into a Proposition, namely:

Proposition 2.2. *Let $\delta > 0$ and let $A \subset \{0, 1, \cdots, N-1\}$ with $|A| \geq \delta N$. If A contains no non-trivial 3AP's, then one of the following must hold.*

1. *$N \leq \frac{8}{\delta^2}$*
2. *The set A is not ϵ-uniform for $\epsilon = \frac{\delta^2}{8}$.*
3. *There exists an arithmetic progression P of length at least $\frac{N}{3} - 1$ such that $|A \cap P| \geq (\frac{9}{8}\delta)|P|$.*

Proof. The contrapositive of Lemma 1 gives us (i) and (ii), we have to deduce (iii) from $|M_A| < \frac{\delta}{4}N$. But this is quite easy, since $A \setminus |M_A|$ now has more than $\frac{3}{4}\delta N$ elements in $[0, \frac{N}{3}] \cup [\frac{2}{3}N, N)$ which are both arithmetic progressions of length $\frac{N}{3} - 1$. One of them has to contain at least half of the elements so

$$\max[|A \cap [0, \tfrac{N}{3})|, |A \cap [\tfrac{2}{3}N, \tfrac{N}{3})|] \geq \frac{3}{8}\delta N = \frac{9}{8}\delta(\tfrac{N}{3}),$$

which means that A has density at least $\frac{9}{8}\delta$ in an arithmetic progression of length $\frac{N}{3} - 1$. □

2.5 Sets with structure

The most important (and hardest) part of the proof will be proving that sets that are not ϵ-uniform have increased density on an arithmetic progression of large size (non pseudo-random sets have 'structure'). This will allow us to iterate Proposition 2.2 to derive a contradiction. As discussed in the proof strategy, we want to find an arithmetic progression in $\mathbb{Z}$ on which A has increased density. Since we are still working in $\mathbb{Z}_N$ we need to find a way to recognise 3AP's mod N that are also genuine 3AP's. The following definition and Lemma do just that:

Definition 2.2. *If P is an arithmetic progression in $\mathbb{Z}_N$, we say that it is* non-overlapping *if its length L and common difference D satisfy $LD < N$.*

Lemma 2.2. *Suppose that B is a non-overlapping $\mathbb{Z}_N$ progression on which A has density $\delta + \epsilon$ then there is a $\mathbb{Z}$-progression P of length at least $\epsilon\frac{|B|}{2}$ on which A has density at least $\delta + \epsilon/2$.*

Proof. We first prove that B is the union of two $\mathbb{Z}$-progressions. Label B as $b, b+D, \cdots, b+(L-1)D$. If $b+jD < b+(j+1)D \mod N$ for some j then we obtain two arithmetic progressions in $\mathbb{Z}$ namely $(b, b+D, \cdots, b+jD$ and $b+(j+1)D-N, b+(j+2)D-N, \cdots, b+(L-1)D-N$. We get from the hypothesis that there is at most one such j. If there is not such a j, we can take $P=B$ and are done.

Assuming that such a j exists we write $B = P_1 \cup P_2$. If $|P_1| \leq \frac{\epsilon|B|}{2}$ then $|P_2| > |B| - \frac{\epsilon|B|}{2} \geq \frac{\epsilon|B|}{2}$ and the density of A in P satisfies

$$\begin{aligned} |A \cap P_2| &\geq |A \cap B| - |P_1|, \\ &= (\delta+\epsilon)|B| - |P_1|, \\ &\geq (\delta + \epsilon/2)|B|, \end{aligned}$$

as required. If both P_1 and P_2 are of size greater than $\frac{\epsilon|B|}{2}$ than A has density at least $\delta+\epsilon$ on one of them. □

We will use this to find 'structure' in sets that are not pseudo-random.

Proposition 2.3. *Suppose that A is not ϵ-uniform, i.e., that there is a $t \in \mathbb{Z}_N$ such that $|\hat{A}(t)| > \epsilon N$, then there exists a non-overlapping $\mathbb{Z}_N$ progression C, with $|C| > \sqrt{N}/4$ such that*

$$|A \cap C| \geq (\epsilon/4 + \delta)|C|$$

The proof is lengthy, and requires some definitions first:

Definition 2.3. *We define the balanced function of A to be (where $\delta = |A|/N$)*

$$\begin{aligned} F_A(x) &= \begin{cases} 1-\delta & \text{if } x \in A \\ -\delta & \text{if } x \notin A \end{cases}, \\ &= A(x) - \delta. \end{aligned}$$

It is easy to see that this means that $\sum_{x \in \mathbb{Z}_N} F_A(x) = 0$, consider

$$\begin{aligned} \sum_{x \in \mathbb{Z}_N} F_A(x) &= \sum_{x \in \mathbb{Z}_N} A(x) - \delta, \\ &= \delta N - \sum_{x \in \mathbb{Z}_N} \delta, \\ &= 0. \end{aligned}$$

However, this balanced function does have the same Fourier transform, for all $k \neq 0$ we have that

$$\begin{aligned} \hat{F_A}(k) &= \sum_{x=0}^{N-1} F_A(x)\omega^{-xk}, \\ &= \sum_{x=0}^{N-1} A(x)\omega^{-xk} - \delta \sum_{x=0}^{N-1} \omega^{-xk}, \end{aligned}$$

$$= \hat{A}(k),$$

using the orthogonality relation 2.1.

Proof of Proposition 2.3. We will prove an auxiliary Lemma first:

Lemma 2.3. *For all $1 \leq t \leq N-1$ there exists a non-overlapping progression B of length at least $\sqrt{N}/4$ such that*

$$|\hat{B}(t)| > \frac{|B|}{2}$$

Proof. Take t as in the hypothesis, that is, take t such that $|\widehat{A}(t)| > \epsilon N$. Consider the square $[0, N-1] \times [0, N-1]$, partition this into $N-1$ equal squares and consider the following sequence in $(\mathbb{Z}_N)^2$.

$$\{(0,0), (1,t), \cdots, (N-1, (N-1)t)\}.$$

This sequence of points all lie in $(\mathbb{Z}_N)^2$ and since there are N points and $N-1$ squares, one of these squares must contain two points, say (k, kt) and (l, lt). Since these points are in the same square, and the square has sides less than $\sqrt{N}$, we must have that

$$\begin{aligned} &|(k-l)| \leq \sqrt{N}, \qquad \text{(we may assume that } k > l \text{, w.l.o.g.)} \\ &t(k-l) \leq \sqrt{N} \mod N. \end{aligned}$$

Now define an arithmetic progression of length $|B| = \lfloor \frac{\sqrt{N}}{\pi} \rfloor$ and common difference $d = k - l$ to be

$$\{\cdots, -d, 0, d \cdots\}.$$

Calculating $\hat{B}(t)$ we notice that

$$\begin{aligned} |\hat{B}(t) - |B|| &\leq |\sum_x B(x) \left(\omega^{-xt} - 1\right)|, \\ &\leq \sum_{x \in B} |\left(\omega^{-xt} - 1\right)|, \\ &= \sum_{k=-\frac{|B|}{2}}^{\frac{|B|}{2}} |\left(\omega^{-kdt} - 1\right)|. \end{aligned}$$

Now we use some geometry and analysis(done in appendix A.5, the argument is lengthy) to obtain

$$= |\sum_{k=-\frac{|B|}{2}}^{\frac{|B|}{2}} \left(\omega^{-kdt} - 1\right)| \leq \frac{|B|}{2},$$

which implies that $|\widehat{B}(t)| > \frac{|B|}{2}$, as claimed. □

It turns out that the arithmetic progression C that we need to prove Proposition 2.3 is B translated by a number $x \in \mathbb{Z}_N$. We will not construct this number, but we will show that it exists. If C is a translate of B then the indicator function $C(y) = B(y - x)$ and so we can calculate the cardinality of $A \cap C$, it is

$$|A \cap C| = \sum_y A(y)B(y - x).$$

If we want to show that this is bigger than $(\epsilon/4 + \delta)|C|$ we have to show that there is an x such that

$$\sum_y A(y)B(y - x) > (\epsilon/4 + \delta)|C|,$$

$$\sum_y (A(y)B(y - x)) - \delta|C| > \epsilon\frac{|C|}{4},$$

$$\sum_y (A(y) - \delta)(B(y - x) > \epsilon\frac{|C|}{4},$$

$$\sum_y F_A(y)B(y - x) > \epsilon\frac{|C|}{4},$$

where F_A is the balanced function of A. Following Lyall [Lya05] we define

$$G(x) = \sum_y F_A(y)B(y - x).$$

Applying part 1 of Proposition 2.1 we obtain

$$\sum_x |G(x)| \geq |\hat{G}(t)|,$$

where t is the number such that $\hat{A}(t) > \epsilon N$, which exists because A is not ϵ uniform by assumption. Applying part 2 of Proposition 2.1 we get

$$\begin{aligned}\sum_x |G(x)| &\geq |\hat{G}(t)|,\\ &= |\hat{F_A}(t)\hat{B}(t)|,\\ &\geq \epsilon N\tfrac{|B|}{2}.\end{aligned}$$

For the next step we first need to calculate

$$\begin{aligned}\sum_x G(x) &= \sum_x \sum_y (A(x) - \delta)B(x - y),\\ &= 0,\end{aligned}$$

since $\sum_y B(x - y)$ is a constant function of x. Using this we see that

$$\begin{aligned}\sum_x G(x) &= \sum_x \left(|G(x)| + G(x)\right),\\ &\geq \frac{\epsilon N|B|}{2},\end{aligned}$$

therefore there must be an x' such that $|G(x')| + G(x') \geq \epsilon/2|B|$ (at least one of the summands must take at least the average value). With this we can conclude that $G(x') \geq \epsilon/4|B|$, as required.

This concludes the proof of Proposition 2.3. □

2.6 The density increment argument

In section 2.2 we discussed wanting to iterate Propositions 2.3 and 2.2 by identifying the long arithmetic progressions P they give us with $\{0, 1 \cdots, |P| - 1\}$. However, we need to make sure that if we restrict this transformation to $A \cap P$ we don't lose the fact that A has no length three arithmetic progressions (as per the contradiction hypothesis). The following Lemma achieves exactly this:

Lemma 2.4. *Let P be an arithmetic progression with common difference k and let B be a subset of P containing no 3AP's. Let $T : P \to \mathbb{N}$ be the map that sends the smallest element of P to 0 and the second smallest to $1, \cdots$. Then $T(B)$ also contains no 3AP's.*

Proof. Assume that $T(B)$ contains an arithmetic progressions $a, a+d, a+2d$, and let $b = T^{-1}(A)$ (T is a bijection). Since P is an arithmetic progression $a + d$ gets mapped to $b + dk$ and $a + 2d$ gets mapped to $a + 2dk$. Notice that $b, b + dk, b + 2dk$ is once again an arithmetic progression in B, a contradiction. □

Using this Lemma we can combine Propositions 2.3 and 2.2 by assuming the worst case scenario of Proposition 2.3 (which has the shortest long arithmetic progression of the two).

Proposition 2.4. *Let $\delta > 0$ and let $N \in \mathbb{N}$, let $B \subset [N]$ such that $|B| \geq \delta N$ and $N \geq \frac{8}{\delta^2}$ and assume B contains no arithmetic progression of length three, then there is a $\mathbb{Z} - AP$ of length at least $\frac{\delta^2}{256}\sqrt{N}$ such that*

$$|B \cap P| \geq (\delta + \frac{\delta^2}{64})|P|.$$

Proof. If B contains no arithmetic progressions, then either it is not ϵ-uniform for $\epsilon \geq \frac{\delta^2}{8}$ or there exists an arithmetic progression P of length $\frac{N}{3} - 1$ such that

$$\begin{aligned} |B \cap P| &\geq (9/8\delta)|P|, \\ &\geq \delta + \frac{\delta^2}{64}. \end{aligned}$$

In the latter case we are done. In the former case we can use Proposition 2.3 to give us a non overlapping arithmetic progression $\mod N$, called P', of length at least $\sqrt{N}/4$ on which B has density at least $\delta + \epsilon/4$. By Lemma 2.5 there is a $\mathbb{Z}$-progression P of size at least $\epsilon|P'|/8$ on which B has density at least $\delta + \epsilon/8$. We get the follow lower bound for this size of P:

$$|P| \geq \epsilon/8|P'| \geq \frac{\epsilon\sqrt{N}}{32} = \frac{\delta^2\sqrt{N}}{256}$$

□

We fix $\delta > 0$ and let $N = N(\delta) \in \mathbb{N}$, let $A \subset [N]$ with density at least δ and assume A contains no 3AP's. We can now define sequences $A_1, A_2, \cdots$ and $P_0, P_1, \cdots$. We let P_0 be the AP given to us by Proposition 2.4 with $B = A$ and $N = N$. We define the rest of the sequence inductively:

If P_k and A_k are already defined, we define A_{k+1} to be $T(A_k \cap P_k)$, were T is defined as in Lemma 2.6. Then we can use Proposition 2.4 with $B = A_{k+1}$ and $N = |P_k|$ to give us an arithmetic progression P_{k+1}. Note that since T preserves the fact that A has no 3AP's, none of the A_i will have a 3AP.

When $k > \frac{64}{\delta}$, the density $\delta_k := \frac{A_k}{|P_k|}$ will be bigger than 2δ since the increase is atleast $\frac{\delta^2}{64}$ per step. Taking $\frac{64}{2\delta}$ more steps we $\delta_k > 4\delta$ since the increase is now atleast $\frac{(2\delta)^2}{64}$. To get a density $\delta_k > 2^l\delta$ we only need

$$\frac{64}{\delta}\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{l-1}}\right) \leq \frac{128}{\delta}$$

steps. So we will reach a density δ_k greater than one, and thus a contradiction before $k = \frac{128}{\delta}$. It is of course important to note that the hypothesis of Proposition 2.4 includes the fact that $N \geq \frac{8}{\delta^2}$ and since we have chosen $N = |P_k|$ we require that $|P_k| \geq \frac{8}{\delta^2}$ for every k.

We also need to know that $|P_k| > 0$ because if P_k is empty, no contradiction can be reached. The size of P_0 is at least $\frac{\delta^2\sqrt{N}}{256}$ and the size of P_k is

$$P_k \geq \frac{\delta_k^2}{256}\sqrt{P_{k-1}}$$

Solving the recursion gives

$$P_k \geq \frac{\delta^4}{256^2}N^{2^{-k}},$$

and since this is a decreasing function of k we only need to check the inequality for the biggest possible k, which is $\frac{128}{\delta}$.

$$\begin{aligned}
\frac{\delta^4}{256^2}N^{2^{-k}} &\geq 1,\\
N^{2^{-k}} &\geq \frac{256^2}{\delta^4},\\
2^{-k}\log N &\geq 4\log 16 - 4\log\delta,\\
2^{-\frac{128}{\delta}}\log N &\geq 16\log 2 - 4\log\delta,\\
\log N &\geq (16\log 2 - 4\log\delta)\,2^{-\frac{128}{\delta}}.
\end{aligned}$$

According to Lyall [Lya05], it is easy to see that

$$16\log 2 + 4\log\delta^{-1} \leq 2^{4\delta^{-1}},$$

and the author doesn't feel it is useful to fabricate a proof of this fact. Finally, applying this fact we obtain

$$\begin{aligned}
\log N &\leq 2^{132\delta^{-1}},\\
\log\log N &\leq \log 132\log(2)\delta^{-1},\\
N &\leq \exp(\exp(132\log(2)\delta^{-1}).
\end{aligned}$$

As was announced, we have just proven Theorem 2 with absolute constant(it doesn't depend on δ) $c = 132\log 2$, which we shall repeat below:

Theorem (Roth(1953))**.** *Let $\delta > 0$ and $N \in \mathbb{N}$, let A be a subset of $[N]$. Then there exists an absolute constant c, such that if $N > \exp(\exp(\frac{c}{\delta}))$ and $|A| \geq \delta N$, then A contains atleast one 3AP.*

3 Arithmetic progressions in primes.

The proof in this section is largely the same as Vaughan's proof of Vinogradov's three primes theorem in his book[Vau81], the author has also consulted Gowers's lecture notes [Gow06] and Nathanson's book [Nat96]. The three sources cited above prove a different theorem than the one we prove here, and some of the proof is original work by the author, although the result is far from new. Section 3.3 is (almost) completely adapted from the sources cited above and section 3.4 only has original work in the last subsection. Furthermore section 3.2 contains most of the original work, the subsection 3.2.3 and the subsection 3.5 are largely original, for example. The approach we take in these subsections has been suggested to the author by R.C. Vaughan and is in fact a special case of exercise 3.2 in his book [Vau81].

Let us start by repeating Theorem 1.3, as we will spend the rest of this section proving it:

Theorem (Van der Corput (1939))**.** *Let $A > 0$ be a real number, and let $n \in \mathbb{N}$, then the number of 3AP's in the primes up to n, denoted by $S(n)$, satisfies*

$$S(n) \geq c\frac{n^2}{4(\log n)^3} + O\left(\frac{n^2}{(\log n)^A}\right),$$

where $c \geq 1$.

3.1 Introduction and organisation

As we have seen in the previous section, finding three term arithmetic progression in some subset A of $[N]$ involves finding $\sum_{n=0}^{N-1} \hat{A}(n)^2\hat{A}(-2n)$. When trying to apply this principle to the primes, one quickly realises that this approach does not work because the primes are not ϵ-uniform. There are various ways to get around this and we have chosen to use the classical approach called the circle method. Instead of working with discrete Fourier analysis we work with Fourier analysis on $\mathbb{Z}$ and the circle (which we identify with $[0, 1)$).

When the circle method was invented, it concerned analysing singularities of complex generating functions (such as described in [FS09]) on the circle. It was Vinogradov who vastly simplified this approach, by looking at finite 'generating functions' of the following form.

$$\begin{aligned} f_n(z) &= \sum_{\substack{a\leq n\\ a\in A}} e^{2\pi iaz} \\ &= \sum_{\substack{a<n\\ a\in A}} e(az), \end{aligned}$$

where we write $e(x) = e^{2\pi ix}$. If we then look at $f_n(z)^2$, the coefficient of $e(mz)$ will denote the number of ways m can be written as the sum of two elements in A that are smaller than n. The constant term of $f_n(z)^2 f_n(-2z)$ will denote the number of arithmetic progressions in A consisting of elements smaller than n. The final (and trivial) observation is that

$$\int_0^1 e(bz)\mathrm{d}z = \begin{cases} 0 & \text{if } b \neq 0 \\ 1 & \text{if } b = 0 \end{cases},$$

which will serve as an analog of the orthogonality relation we had in Proposition 2.1. We will take the set of primes P for our set A and use the following function

$$f_n(\alpha) = \sum_{p \leq n} \log(p)\ e(\alpha p),$$

where the notation $p \leq n$ indicates summing over all primes below n. Essentially this means that we don't use the normal indicator function of the primes but that we weigh each prime by its logarithm, which will be needed for a technical reason. As said before the generating function for 3AP's in the primes up to n will be

$$f_n(\alpha)^2 f_n(-2\alpha).$$

The weighted number of 3AP's in the primes below n, which we shall denote by $R(n)$ is then

$$\begin{aligned} R(n) &= \int_0^1 f_n(\alpha)^2 f_n(-2\alpha) \mathrm{d}\alpha, \\ &= \int_{-1/2}^{1/2} f_n(\alpha)^2 f_n(-2\alpha) \mathrm{d}\alpha, \end{aligned}$$

and we are left with the task of calculating the integral.

3.1.1 Defining the arcs

As we have seen in the previous section we would like to integrate our generating function $f_n(\alpha)^2 f_n(-2\alpha)$ over $(0,1)$. To do this, we will divide up the interval into *major arcs*($\mathfrak{M}$) and *minor arcs*($\mathfrak{m}$), which we shall define below.

Let n be a natural number, $B > 0$ and let $P \leq (\log n)^B$ (We will choose the precise value of B later). We want to try to approximate $f_n(\alpha)$ when α is close to a rational number with a small denominator, that is, when there are a, q with $a \leq q \leq P$ such that $|\alpha - a/q| < P/n$. For practical purposes, we shall demand that our fractions a/q are reduced, i.e., that $(a,q) = 1$. Let us define

$$\begin{aligned} \mathfrak{M}(a,q) &= \{x \in (0,1) : |x - a/q| < P/n, a \leq q \leq P\}, \\ \mathfrak{M} &= \bigcup_{q \leq P} \bigcup_{\substack{a \leq q \\ (a,q)=1}} \mathfrak{M}(a,q), \\ \mathfrak{m} &= (0,1) \setminus \mathfrak{M}. \end{aligned}$$

We expect that our generating function is 'large' on the major arcs and 'small' on the minor arcs. Since our generating function is smooth, we hope to be able to estimate it on a major arc $\mathfrak{M}(a,q)$ by knowing it's value in a/q. The strategy will now be to obtain an asymptotic expression for the integral of our generating function $f_n(\alpha)$ over the major arcs, and an upper bound for the integral of our generating function over the minor arcs. We'd like to be able to write

$$R(n) = \int_0^1 f_n(\alpha)^2 f_n(-2\alpha) \mathrm{d}\alpha,$$

$$= \int_{\mathfrak{M}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha + \int_{\mathfrak{m}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha,$$

but this is only true when the major arcs are disjoint. Assume then, that there is a number γ such that γ is in two different major arcs, so then

$$|\gamma - a/q| < P/n,$$
$$|\gamma - a'/q'| < P/n.$$

By the triangle inequality this implies that $|a/q - a'/q'| \leq 2P/n$ and we know that $|a/q - a'/q'| \geq 1/qq'$ since the difference is a nonzero fraction with denominator qq'. Finally $q, q' \leq P$ implies

$$2P/n \geq |a/q - a'/q'| \geq 1/P^2,$$
$$2P^3 \geq n,$$

and this is impossible for n large enough.

Since we know that $R(n)$ is a positive, increasing function of n, we need to show that the major arcs satisfy some estimate of the form

$$\int_{\mathfrak{M}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha = cg(n) + o(g(n)),$$

and the minor arcs satisfy some estimate of the form

$$\int_{\mathfrak{m}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha = o(g(n)),$$

where $g(n)$ is some positive increasing function of n. This will then allow us to conclude that $R(n) \sim cg(n)$.

We have seen in the introduction that we expect the number of arithmetic progressions in the primes up to n to be $\frac{n^2}{4\log(n)^3}$, but we have weighted each prime p with a weight of $\log p$. It turns out we will get an asymptotic of the form $g(n) = c\frac{n^2}{4}$ where $c \geq 1$.

In section 3.2, we prove the following Proposition:

Proposition 3.1. *Let $n \in \mathbb{N}$ as above, then*

$$\int_{\mathfrak{M}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha = c\frac{n^2}{4} + O\left(\frac{n^2}{P}\right),$$

where $c \geq 1$.

In section 3.4, we prove the following Proposition:

Proposition 3.2. *Let A be a positive constant, then*

$$\left|\int_{\mathfrak{m}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha\right| \ll \frac{n^2}{(\log n)^A}.$$

To prove this we need to do some preliminary work which is done in section 3.3. We will derive Theorem 1.3 from these Propositions in Section 3.5

3.2 Major Arcs

3.2.1 Approximating $f_n(\alpha)$ on the major arcs

To prove Proposition 3.1 we will first prove a series of Lemmas. In this subsection, we prove the following Lemma

Lemma 3.1. *Let $\alpha \in \mathfrak{M}(a,q)$, then there is a constant $c > 0$ such that*

$$f_n(\alpha)^2 f_n(-2\alpha) = \frac{(-1)^{q+1}\mu(q)}{\phi(q)^3} v_n(\beta)^2 v_n(-2\beta) + \left(n^3 \exp(-c\sqrt{\log n})\right),$$

where $\beta = \alpha - a/q$.

Proof. Let us take a look at the value of f_n in a rational point, this is

$$f_n\left(\frac{a}{q}\right) = \sum_{p \leq n} (\log p) e(pa/q).$$

Now, $e\left(\frac{ap}{q}\right)$ is a periodic function with period q, so we can sort our sum into residue classes modulo q. Since both a and p are relatively prime to q, except for a small number of exceptions, we have to sum over all residue classes coprime to q.

$$f_n\left(\frac{a}{q}\right) = E + \sum_{\substack{k=1 \\ (k,q)=1}}^{q} \left(e\left(\frac{ka}{q}\right)) \sum_{\substack{p \leq n \\ p=k \mod q}} \log p \right),$$

where E an error obtained by discarding all primes that divide q. To estimate E we simply sum over all primes below q, and find

$$|E| < \sum_{p<q} \log p \ll q,$$

by Chebyshev's estimate (A.7).

To estimate

$$\sum_{\substack{p \leq n \\ p=k \mod q}} \log p$$

we need a form of the prime number theorem in arithmetic progressions, the theorem of Siegel and Walfisz. This theorem gives an estimate of the number of primes that are $k \mod q$ where $(k,q) = 1$ and q is allowed to go to infinity with X, albeit slowly. We will accept this theorem as a black box and instead refer to Chapter 11 of [MV07] for the proof.

Theorem 3.1 (Siegel-Walfisz)**.** *Let $B > O$ and $X > 1$ and let $k < q < (\log X)^B$ such that $(k, q) = 1$, then there is a constant c_1 depending only on B, such that*

$$\sum_{\substack{p \leq X \\ p = k \bmod q}} \log p = \frac{X}{\phi(q)} + O\left(X \exp(-c_1 \sqrt{\log X})\right),$$

where ϕ is the Euler-totient function as defined in Appendix A.6.

Using this theorem with $X = n$ and noticing that the error term is bigger than the error term E we previously had, we write

$$\begin{aligned} f_n\left(\frac{a}{q}\right) &= \sum_{\substack{k=1 \\ (k,q)=1}}^{q} e\left(\frac{ka}{q}\right) \frac{n}{\phi(q)} + O\left(n \exp(-c_1 \sqrt{\log n})\right), \\ &= \frac{n}{\phi(q)} \sum_{\substack{k=1 \\ (k,q)=1}}^{q} e\left(\frac{ka}{q}\right) + O\left(n \exp(-c_1 \sqrt{\log n})\right), \\ &= \frac{\mu(q)}{\phi(q)} n + O\left(n \exp(-c_1 \sqrt{\log n})\right). \end{aligned}$$

Here we used the identity $\mu(q) = \sum_{k=1,(k,q)=1}^{q} e\left(\frac{ka}{q}\right)$ which is a special case of Theorem A.3.

Let $\alpha \in \mathcal{M}(a, q)$, if $a < \frac{q}{2}$ then we can approximate -2α by $-2\frac{a}{q}$ and only lose a factor 2 in the approximation. If $a > \frac{q}{2}$, then $\alpha > 1/2$ (for n large enough) and so -2α will be bigger than 1. Since e is 1-periodic can instead consider $-2\alpha + 1$, which is well approximated by $-(2a - q)/q$.

A difficulty arises here, though, it is no longer true that $-2ap$ is always coprime to q. When q is even, they are never coprime! We deviate from our sources to tackle this difficulty. We apply theorem 272 from [HW08] which can be found in the appendix as Theorem A.3. This solves our issue but makes matters more complex as we now have to distinguish between even and odd q from now on. Continuing with the algebra we see

$$\begin{aligned} f_n\left(-\frac{2a}{q}\right) &= E + \sum_{\substack{r \leq q \\ (r,q)=1}} e\left(-\frac{2ra}{q}\right) \sum_{\substack{p \leq n \\ p \equiv r \bmod q}} \log p, \\ &= \sum_{\substack{r \leq q \\ (r,q)=1}} e\left(-\frac{2ra}{q}\right) \frac{n}{\phi(q)} + O\left(n \exp(-c_1 \sqrt{\log n})\right) \qquad \text{(Siegel Walfisz)}, \\ &= \frac{n}{\phi(q)} \sum_{\substack{r \leq q \\ (r,q)=1}} e\left(-\frac{2ra}{q}\right) + O\left(n \exp(-c_1 \sqrt{\log n})\right), \end{aligned}$$

$$= \frac{n}{\phi(q)} \frac{\mu\left(\frac{q}{(q,2)}\right)\phi(q)}{\phi\left(\frac{q}{(q,2)}\right)} + O\left(n\exp(-c_1\sqrt{\log n})\right), \qquad \text{(by Theorem 272)}$$
$$= \frac{n}{\phi(q)} g(q) + O\left(n\exp(-c_1\sqrt{\log n})\right),$$

where we define

$$g(q) := \begin{cases} \mu(q) & \text{if } q \text{ is odd}, \\ \frac{\mu(\frac{q}{2})\phi(q)}{\phi(\frac{q}{2})} & \text{if } q \text{ is even}. \end{cases}$$

We would now like to show that if α is close enough to $\frac{a}{q}$, then $f_n(\alpha)$ is also close to $f_n(\frac{a}{q})$. Define a sequence of numbers

$$c_n = \begin{cases} e\left(\frac{an}{q}\right)\log n - \frac{\mu(q)}{\phi(q)} & \text{if n is prime,} \\ -\frac{\mu(q)}{\phi(q)} & \text{if otherwise.} \end{cases}$$

In the appendix(A.1) we find the following theorem:

Theorem 3.2 (Partial Summation)**.** *Let a_n be a sequence of complex numbers, $X > 0$ and let $h : \mathbb{R} \to \mathbb{C}$ be a continuously differentiable function, then*

$$\sum_{x\le X} a_x h(x) = h(X)\sum_{x\le X} a_x - \int_1^X h'(t)\sum_{x\le t} a_x \mathrm{d}t.$$

Applying the Theorem above with $h(x) = e(\beta x)$ where $\beta = \alpha - \frac{a}{q}$, choosing $X = n$ and $a_n = c_n$ we obtain

$$\sum_{x\le n} c_x e(\beta x) = e(\beta n)\sum_{x\le n} c_x - \int_1^n 2\pi i\beta e(\beta t)\sum_{x\le t} c_x \mathrm{d}t,$$
$$\sum_{p\le n} e\left(\frac{ap}{q}\right)\log(p)e(\beta p) - \frac{\mu(q)}{\phi(q)}\sum_{x\le n} e(\beta x) = e(\beta n)\sum_{x\le n} c_x - \int_1^n 2\pi i\beta e(\beta t)\sum_{x\le t} c_x \mathrm{d}t,$$
$$f_n(\alpha) - \frac{\mu(q)}{\phi(q)}\sum_{x\le n} e(\beta x) = e(\beta n)\sum_{x\le n} c_x - \int_1^n 2\pi i\beta e(\beta t)\sum_{x\le t} c_x \mathrm{d}t.$$

Now the left side is an approximation to $f_n(\alpha)$ on $\mathfrak{M}(a,q)$ and so the right side is an error term we'd like to estimate. Write $\sum_{x\le t} c_x = E(t)$ and then

$$\left|f_n(\alpha) - \frac{\mu(q)}{\phi(q)}\sum_{x\le n} e(\beta x)\right| = \left|e(\beta n)\sum_{x\le n} c_x - \int_1^n 2\pi i\beta e(\beta t)\sum_{x\le t} c_x \mathrm{d}t.\right|,$$
$$\le |E(n)| + 2\pi n\beta \sup_{x\le n}|E(t)|.$$

We have previously seen that $|E(t)| = |f_t(\frac{a}{q}) - \frac{\mu(q)}{\phi(q)}t| = |O\left(t\exp(-c_1\sqrt{\log t})\right)|$, which is an increasing function of t so $\sup_{t\leq n}|E(t)| = |E(n)|$. This implies that

$$\begin{aligned}|E(n)| + 2\pi n \sup_{x\leq n}|E(t)| &\leq (1+2\pi n\beta)E(n),\\ &\leq (1+2\pi P)E(n), \qquad (\text{Since } \beta \leq P/n)\\ &\ll O\left(n\exp(-c\sqrt{\log n})\right),\end{aligned}$$

where c is a new constant.

For $f_n(-2\alpha)$, we do something similar, discriminating between the cases q odd and q even. Let $\alpha \in \mathfrak{M}(a,q)$, then define a sequence of numbers

$$d_n = \begin{cases} e\left(-\frac{2an}{q}\right)\log n - \frac{g(q)}{\phi(q)} & \text{if n is prime,}\\ -\frac{g(q)}{\phi(q)} & \text{if otherwise.}\end{cases}$$

Applying the partial summation formula we obtain

$$\sum_{p\leq n} e(-2\alpha p)\log p - \frac{g(q)}{\phi(q)}\sum_{x\leq n} e(-2\beta x)) = e(-2\beta n)\sum_{x\leq n} d_n + \int_1^n 4\pi i\beta e(-2\beta t)\sum_{x\leq t} d_n \mathrm{d}t,$$

$$\begin{aligned}|f_n(-2\alpha) - g(q)v_n(-2\beta)| &\ll (1+2\pi n)| - 2\alpha + \tfrac{2a}{q}|\, O\left(n\exp(-c\sqrt{\log n})\right),\\ &\ll O\left(n\exp(-c\sqrt{\log n})\right),\end{aligned}$$

where some details were omitted because of the similarity to the previous argument.

We need to find an estimate for $f_n(\alpha)^2 f_n(-2\alpha)$, multiplying our approximations for $f_n(\alpha)$ gives

$$\begin{aligned}f_n(\alpha)^2 &= \frac{\mu(q)^2}{\phi(q)^2}v_n(\beta)^2 + 2\frac{\mu(q)}{\phi(q)}v_n(\beta)O\left(n\exp(-c\sqrt{\log n}\right) + O\left(n^2\exp(-c\sqrt{\log n}\right),\\ &= \frac{\mu(q)^2}{\phi(q)^2}v_n(\beta)^2 + O\left(n^2\exp(-c\sqrt{\log n}\right),\end{aligned}$$

using that $|v_n(\beta)| \leq n$. Combining these estimates with our estimate for $f_n(-2\alpha)$ we get

$$\begin{aligned}f_n(\alpha)^2 f_n(-2\alpha) &= \frac{\mu(q)^2 g(q)}{\phi(q)^2}v_n(\beta)^2 v_n(-2\beta) + v_n(-\beta)\left(n^2\exp(-c\sqrt{\log n})\right) + \left(n^3\exp(-c\sqrt{\log n})\right),\\ &= v_n(\beta)^2 v_n(-2\beta)h(q) + \left(n^3\exp(-c\sqrt{\log n})\right),\end{aligned}$$

where

$$\begin{aligned}h(q) &= \begin{cases}\frac{\mu(q)}{\phi(q)^3} & \text{if } q \text{ is odd}\\ \frac{\mu(q)^2\mu(q/2)}{\phi(q)^2\phi(q/2)} & \text{if } q \text{ is even}\end{cases},\\ &= \begin{cases}\frac{\mu(q)}{\phi(q)^3} & \text{if } q \text{ is odd}\\ \frac{-\mu(q)}{\phi(q)^3} & \text{if } q \text{ is even}\end{cases},\end{aligned}$$

$$= (-1)^{q+1}\frac{\mu(q)}{\phi(q)^3},$$

using the fact that $\mu(q)$ and $\phi(q)$ are multiplicative. In the above we also used that $\mu(x) = 0$ if x is divisible by a square so in the above we may assume x is squarefree. This implies that 2 and $\frac{q}{2}$ are coprime. □

Putting it al; together we have proven Lemma. 3.1 which we shall repeat below

Lemma. *Let $\alpha \in \mathfrak{M}(a,q)$, then there is a constant $c > 0$ such that*

$$f_n(\alpha)^2 f_n(-2\alpha) = \frac{(-1)^{q+1}\mu(q)}{\phi(q)^3} v_n(\beta)^2 v_n(-2\beta) + \left(n^3 \exp(-c\sqrt{\log n})\right),$$

where $\beta = \alpha - a/q$.

3.2.2 Integrating over the major arcs

Using this Lemma, we can now approximate the integral of our generating function over the major arcs. Let $\mathfrak{M}(a,q)$ be a major arc, then as α runs between $a/q - P/n$ and $a/q + P/n$, the difference β runs between $-P/n$ and P/n and so

$$\begin{aligned}\int_{\mathfrak{M}(a,q)} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha &= \frac{(-1)^{q+1}\mu(q)}{\phi(q)^3}\int_{-P/n}^{P/n} v_n(\beta)^2 v_n(-2\beta) + O\left(n^3\exp(-c\sqrt{\log n})\right)\mathrm{d}\beta,\\ &= \frac{(-1)^{q+1}\mu(q)}{\phi(q)^3}\int_{-P/n}^{P/n} v_n(\beta)^2 v_n(-2\beta)\mathrm{d}\beta + O\left(Pn^2\exp(-c\sqrt{\log n})\right).\end{aligned}$$

We know that $v_n(\beta)^2 v_n(-2\beta)$ is the generating function for arithmetic progressions in $[n]$, but to apply the result from the introduction (that the number of arithmetic progressions in $\{1,\cdots,n\}$ is $n^2/4+O(n)$) we would need to integrate over a bigger interval. As it turns out, we can do so without too much of an error

Lemma 3.2.

$$\int_{-P/n}^{P/n} v(\beta)^2 v(-2\beta)\mathrm{d}\beta = \int_{-1/2}^{1/2} v(\beta)^2 v(-2\beta)\mathrm{d}\beta + O\left(\frac{n^2}{P}\right)$$

Proof. We will provide an estimate for the error term $\int_{P/n}^{1/2}$, the other ërror term follows by symmetry. On this interval we know that $\|\beta\| = \beta$, and using the inequalities $|v_n(\beta)| \leq \|\beta\|^{-1}$ and $|v_n(-2\beta)| \leq n$ we obtain

$$\left|\int_{P/n}^{1/2} v_n(\beta)^2 v_n(-2\beta)\mathrm{d}b\right| \leq n\int_{P/n}^{1/2}\frac{1}{\beta^2}\mathrm{d}\beta,$$

$$\ll \frac{n^2}{P}$$

□

The Lemma above allows us to continue where we left of. Using the result from the introduction (the number of 3AP's in $\{1, \cdots, n\} = n^2/4 + O(n)$) we calculate the integral of our generating function over all the major arcs.

$$\begin{aligned}
\int_{\mathfrak{M}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha &= \sum_{q\leq P} \sum_{\substack{a<q \\ (a,q)=1}} \int_{\mathfrak{M}(a,q)} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha, \\
&= \sum_{q\leq P} \sum_{\substack{a<q \\ (a,q)=1}} \frac{(-1)^{q+1}\mu(q)}{\phi(q)^3} \left(\int_{-P/n}^{P/n} v_n(\beta)^2 v_n(-2\beta)\mathrm{d}\beta + O\left(Pn^2 \exp(-c\sqrt{\log n}) \right) \right), \\
&= \sum_{q\leq P} \sum_{\substack{a<q \\ (a,q)=1}} \frac{(-1)^{q+1}\mu(q)}{\phi(q)^3} \left(\int_{-1/2}^{1/2} v_n(\beta)^2 v_n(-2\beta)\mathrm{d}\beta + O\left(\frac{n^2}{P} \right) \right), \qquad \text{(Lemma 3.2)} \\
&= \sum_{q\leq P} \sum_{\substack{a<q \\ (a,q)=1}} \left[\frac{(-1)^{q+1}\mu(q)}{\phi(q)^3} \left(\frac{n^2}{4} + O\left(\frac{n^2}{P} \right) \right) \right], \\
&= \sum_{q\leq P} \sum_{\substack{a<q \\ (a,q)=1}} \left[\frac{(-1)^{q+1}\mu(q)}{\phi(q)^3} \frac{n^2}{4} \right] + O\left(\frac{n^2}{P} \right), \\
&= \frac{n^2}{4} \left(\sum_{q\leq P} \frac{(-1)^{q+1}\mu(q)}{\phi(q)^2} \right) + O\left(\frac{n^2}{P} \right), \\
&= \mathfrak{S}(P) n^2/4 + O\left(\frac{n^2}{P} \right),
\end{aligned}$$

where $\mathfrak{S} = \lim_{P\to\infty} \mathfrak{S}(P)$ is called the singular series (for historical reasons). Since we will show in the next section that $\mathfrak{S}(P)$ is $O(1)$, the manipulations with the error term above are justified.

3.2.3 The Singular series

The rest of Section 3.2 will consist of original work done by the author.

Lastly, we would like to find an asymptotic for the singular series. We start by proving absolutely convergence. Theorem 327 in Hardy and Wright([HW08]) tells us that $\phi(q) \gg q^{1-\delta}$ for all $\delta > 0$, which can be found in the appendix as Theorem A.6. This means that if we use Cauchy-Schwarz we get convergence by the integral-comparison test.

Now for n (and thus P) large enough we can approximate

$$\begin{aligned}\mathfrak{S}(P) &= \sum_{q\leq P}\frac{(-1)^{q+1}\mu(q)}{\phi(q)^2},\\ &= \sum_{q=1}^{\infty}\frac{(-1)^{q+1}\mu(q)}{\phi(q)^2} - \sum_{q>P}\frac{(-1)^{q+1}\mu(q)}{\phi(q)^2},\\ &= \mathfrak{S} - \sum_{q>P}\frac{(-1)^{q+1}\mu(q)}{\phi(q)^2}.\end{aligned}$$

To estimate the error term we apply the triangle inequality and again use that $\phi(q) \gg q^{1-\delta}$ which implies

$$\begin{aligned}|\sum_{q>P}\frac{(-1)^{q+1}\mu(q)}{\phi(q)^2}| &\leq \sum_{q>P}\frac{1}{\phi(q)^2},\\ &\leq \sum_{q>P}\frac{1}{q^{2-2\delta}}\\ &\ll \int_P^{\infty}\frac{1}{t^{2-2\delta}}\mathrm{d}t = O\left(\frac{1}{P^{1-2\delta}}\right).\end{aligned}$$

We are nearly done with the error term now, we just have to notice that if we multiply the error term by $n^2/4$ we still end up in $O\left(\frac{n^2}{(\log n)^A}\right)$ for arbitrary A by choosing B correctly (remember that $P = (\log n)^B$).

At this point we would like to use the fact that we are dealing with well behaved (multiplicative) arithmetical functions, the only problem is the factor $(-1)^{q+1}$ (which is not multiplicative). We can solve this by summing over odd and even q separately.

$$\sum_{q=1}^{\infty}\frac{(-1)^{q+1}\mu(q)}{\phi(q)^2} = \sum_{q\text{ odd}}\frac{\mu(q)}{\phi(q)} - \sum_{q\text{ even}}\frac{\mu(q)}{\phi(q)},$$

Since μ is zero on numbers that are divisible by squares, we can assume q to be squarefree, which means that $q/2$ and 2 are coprime and so $\mu(q/2)\mu(2) = \mu(q)$ and $\phi(q/2)\phi(2) = \phi(q)$ for all even q, this simplifies our sum to be

$$\begin{aligned}\mathfrak{S} &= \sum_{q\text{ odd}}\frac{\mu(q)}{\phi(q)^2} - \frac{\mu(2)}{\phi(2)^2}\sum_{q\text{ even}}\frac{\mu(q/2)}{\phi(q/2)^2},\\ &= \sum_{q\text{ odd}}\frac{\mu(q)}{\phi(q)^2} + \sum_{q\text{ even}}\frac{\mu(q/2)}{\phi(q/2)^2},\\ &= \sum_{q\text{ odd}}\frac{\mu(q)}{\phi(q)^2} + \sum_{q\text{ odd}}\frac{\mu(q)}{\phi(q)^2},\\ &= 2\sum_{q\text{ odd}}\frac{\mu(q)}{\phi(q)^2},\end{aligned}$$

where we once again use that q is squarefree. We have now obtained a sum over multiplicative functions which is absolutely convergent and so we may apply the Euler factorisation:

$$\sum_{x=1}^{\infty}r(x) = \prod_p\left(1 + r(p) + r(p^2) + \cdots\right),$$

for multiplicative functions $r : \mathbb{N} \to \mathbb{R}$. This result can be found in the appendix as Theorem (A.4). Since we are summing over all odd numbers, we need to leave out all numbers divisible by two so we take the product over all primes greater than two (the reader might need to convince himself that this works). Therefore

$$\begin{aligned}
\mathfrak{S} &= 2\prod_{p>2}\left(1+\frac{\mu(p)}{\phi(p)^2}+\frac{\mu(p^2)}{\phi(p^2)^2}+\cdots\right), \\
&= 2\prod_{p>2}\left(1+\frac{\mu(p)}{\phi(p)^2}\right), \qquad (\ \mu(p^k)=0 \text{ for all } k>0\ ) \\
&= 2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right).
\end{aligned}$$

Remark 3.1. *The constant*

$$C_2 = \prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right),$$

is known in the literature as the twin-prime constant and its value has been calculated up to many decimal places and is $C_2 \approx 0.660161$ *[Wik].*

The last observation we need is that $1-\frac{1}{(p_j-1)^2} \geq 1-\frac{1}{(p_{j-1})^2}$ (with equality occuring when $j=2$), where p_i is the i-th prime. Using this and a well known identity by Euler we get

$$\begin{aligned}
\mathfrak{S} &\geq 2\prod_{p}\left(1-1/p^2\right), \\
&= \frac{2}{\zeta(2)}, \\
&= \frac{12}{\pi^2}, \\
&\approx 1.2.
\end{aligned}$$

This concludes the proof of Proposition 3.1 which we shall repeat below.

Proposition. *Let* $n \in \mathbb{N}$ *as above, then*

$$\int_{\mathfrak{M}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha = c\frac{n^2}{4} + O\left(\frac{n^2}{P}\right),$$

where $c \geq 1$.

3.3 Weyl's Inequality

3.3.1 Preliminaries

The purpose of this section is to prove an exponential sum estimate needed for the minor arcs estimates. We will largely follow the approach of Gowers [Gow] but the author also found chapter 1 of [Vau81] useful.

Lemma 3.3. *Let $0 < \alpha < 1/2$, then*

$$2\alpha < \sin \pi\alpha < \pi\alpha$$

Proof. Write $f(x) = \sin(\pi x) - 2x$, notice that $f(0) = 0 = f(1/2)$ and that $f(1/4) = \frac{\sqrt{2}}{2} - 1/4 > 0$. Suppose that $f(x) \leq 0$ in $(0, 1/2)$ then its derivative must have two zeroes in $(0, 1/2)$ (because then it takes a minimum and a maximum on $(0, 1/2)$), but $\pi \cos(\pi x) - 2$ only has one zero in $(0, 1)$, so $f(x) > 0$ for all $x \in (0, 1)$.

We consider the function $g(x) = \pi x - \sin(\pi x)$, we can Taylor-expand the sine function to get

$$g(x) = \pi x - \pi x + \frac{\pi^3 x^3}{6} \pm \frac{\pi^5 (1/2)^5}{5!} > 0$$

and so we can conclude that $\sin \pi x < \pi x$. □

Lemma 3.4. *Let $a \in (0, 1)$ and write $e(x) = \exp(2\pi i x)$, then*

$$|\sum_{n \leq X} e(\alpha n)| \leq \min[X, \frac{1}{\|\alpha\|}],$$

where $\|\alpha\| = \min_{k \in \mathbb{Z}} |\alpha - k|$.

Proof. Using the triangle equality we immediately obtain

$$\begin{aligned} |\sum_{n \leq X} e(\alpha n)| &\leq \sum_{n \leq X} |e(\alpha n)|, \\ &= X. \end{aligned}$$

For the other inequality, we notice that we are dealing with the partial sums of a geometric series and write

$$\begin{aligned} |\sum_{n \leq X} e(\alpha n)| &= |\frac{1 - e(\alpha(\lfloor X \rfloor + 1) - 1}{e(\alpha n) - 1}|, \\ &\leq \frac{2}{|e(\alpha) - 1|}, \\ &= \frac{2}{|e(\alpha/2) - e(-\alpha/2)|}, \end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{|2i(\sin 2\pi\alpha/2)|}, \\
&= \frac{1}{|\sin \pi\alpha|}, \\
&= \frac{1}{\sin(\pi\|\alpha\|}, \\
&\leq \frac{1}{2\|\alpha\|},
\end{aligned}
$$

using Lemma 3.3 and the fact that $0 < \|\alpha\| < 1/2$. □

3.3.2 A technical result

In the treatment of the minor arcs, we need an estimate of the following form twice, so we have proven it here in the current form.

Proposition 3.3 (Weyl's inequality)**.** *Let $\alpha \in (0,1)$ and $X > 1$, and let $a, q \in \mathbb{Z}$ be such that $|\alpha - \frac{a}{q}| \leq \frac{1}{q^2}$ and $(a,q) = 1$, then*

$$\sum_{x \leq X} \min\left[\frac{n}{x}, \frac{1}{\|x\alpha\|}\right] \ll (\log n)^2(q + X + \tfrac{n}{q})$$

We will choose $X = n^{2/5}$ in the next section, but we will prove the Proposition for general X here.

We will split up the result into several parts and start with the following Lemma:
Lemma 3.5. *Let $\alpha \in \mathbb{R}$ and let $a, q \in \mathbb{Z}$ such that $(a,q) = 1$ and $|\alpha - \frac{a}{q}| < q^{-2}$. Then the numbers $\alpha, 2\alpha, \cdots, \lfloor \frac{q}{2} \rfloor \alpha$ are $\frac{1}{2q}$ separated, i.e., $\|s\alpha - r\alpha\| \geq 1/2q$ for all $1 \leq s, r \leq \lfloor \frac{q}{2} \rfloor$.*

Proof. Let $1 \leq s, r \leq \lfloor \frac{q}{2} \rfloor$, then we can write

$$
\begin{aligned}
s\alpha - r\alpha &= (s-r)(a/q + \alpha - \tfrac{a}{q}) \\
&= (s-r)(a/q) + (s-r)(\alpha - \tfrac{a}{q}).
\end{aligned}
$$

Since $s - r \leq \frac{q}{2}$ we know that $(s-r)a \neq 0 \mod q$ and so $\|(s-r)a/q\| \geq 1/q$. Using the hypothesis we obtain that $(s-r)(\alpha - a/q) \leq \frac{1}{2q}$ and so

$$\|s\alpha - r\alpha\| \geq \frac{1}{2q}$$

□

Using this fact we can split up our sum into pieces of size $\frac{q}{2}$ and then sum over those pieces separately, we will shall formulate this into another Lemma:

Lemma 3.6. *Let $\theta_1, \cdots \theta_m$ be real numbers that are r^{-1} separated and such that $r > m$, and let $Q \in \mathbb{N}$ such that $Q \geq 2$. Then*

$$\sum_{i=1}^{m} \min\left[\frac{1}{\|\theta_i\|}, Q\right] \leq 6\log(Q)(Q+r).$$

Proof. We can assume without loss of generality that the θ_i are in $[-1/2, 1/2)$ and that the non-negative θ_j contribute at least half of the total sum. Suppose we order $0 < \theta_1 < \cdots < \theta_k$ then we can write

$$\begin{aligned}\sum_{i=1}^{m} \min\left[\frac{1}{\|\theta_i\|}, Q\right] &\leq 2\sum_{i=1}^{k} \min\left[\frac{1}{\|\theta_i\|}, Q\right], \\ &\leq 2\sum_{i \leq r/Q} Q + 2\sum_{r/Q < i \leq k} \frac{r}{i},\end{aligned}$$

In this last expression we just notice that $\|\theta_j\| > (j-1)/r$ because of the r^{-1} seperatedness of the θ_j and that $r/(j-1) > Q$ for $j - 1 < r/Q$. We calculate the truncated harmonic series explicitly and obtain our bound:

$$\begin{aligned}\leq 2\sum_{i \leq r/Q} Q + 2\sum_{r/Q < i < k} \frac{r}{i} &\leq 2Q(r/Q + 1) + 4r(\log k - \log r + \log Q) \\ &\leq 2(r + Q + 4r\log Q) \\ &\leq 6\log(Q)(Q+r)\end{aligned}$$

□

We now use this Lemma to prove Proposition 3.3.

Proof of Proposition 3.3. We partition X into intervals of length $\frac{q}{2}$(assume w.l.o.g. that q is even, otherwise take $2q$ since that only loses us a multiplicative constant) and take some extra terms for convenience:

$$\sum_{x \leq X} \min\left[\frac{n}{x}, \frac{1}{\|x\alpha\|}\right] \leq \sum_{i=0}^{\frac{2X}{q}} \sum_{x=\frac{iq}{2}+1}^{\frac{(i+1)q}{2}} \min\left[\frac{n}{x}, \frac{1}{\|x\alpha\|}\right].$$

Now we note that $(x\alpha)_{x=\frac{iq}{2}}^{\frac{(i+1)q}{2}}$ is $1/(2q)$ seperated for each $i \geq 1$, and also that $\frac{n}{x} < \frac{2n}{iq}$ for all $i \geq 1$. We split off the case $i = 0$

$$\sum_{x \leq X} \min\left[\frac{n}{x}, \frac{1}{\|x\alpha\|}\right] \leq \sum_{i=1}^{\frac{2X}{q}} \sum_{x=\frac{qi}{2}}^{\frac{q(i+1)}{2}} \min\left[\frac{2n}{iq}, \frac{1}{\|x\alpha\|}\right] + \sum_{x=1}^{\frac{q}{2}} \min\left[\frac{n}{x}, \frac{1}{\|x\alpha\|}\right].$$

We can estimate the inner sum of the first sum using the previous Lemma with $Q = \frac{2n}{iq}, r = 2q$, $m = \frac{q}{2}$ and take $\theta_k = \left(\frac{qi}{2} + k\right)\alpha$

$$\begin{aligned} S_i &= \sum_{x=\frac{qi}{2}+1}^{\frac{q(i+1)}{2}} \min\left[\frac{2n}{iq}, \frac{1}{\|x\alpha\|}\right], \\ &= \sum_{k=1}^{\frac{q}{2}} \min\left[\frac{2n}{iq}, \frac{1}{\|\theta_k\|}\right], \\ &\leq 6\log(\tfrac{2n}{iq})(\tfrac{2n}{iq} + 2q). \end{aligned}$$

Summing this over i we get

$$\begin{aligned} \sum_{i=1}^{\frac{2X}{q}} S_i &\leq \sum_{i=1}^{\frac{2X}{q}} 6\log(\tfrac{2n}{iq})(\tfrac{2n}{iq} + 2q), \\ &\ll \sum_{i=1}^{\frac{2X}{q}} (\frac{\log n}{i} - \log iq)(\tfrac{n}{iq} + q), \\ &= \sum_{i=1}^{\frac{2X}{q}} \frac{n}{iq}\log n + q\log n + q\log iq + \tfrac{n}{iq}\log iq \\ &\ll \tfrac{n}{q}(\log n)^2 + X\log n + q(\log n)^2, \\ &\ll (\log n)^2(\tfrac{n}{q} + X + q). \end{aligned}$$

All that remains now is to estimate the second sum

$$\begin{aligned} \sum_{x=1}^{\frac{q}{2}} \min\left[\frac{X}{x}, \frac{1}{\|x\alpha\|}\right] &\leq X\sum_{i=1}^{\frac{q}{2}} \frac{1}{x}, \\ &\ll X\log q. \end{aligned}$$

This finished the proof of Proposition 3.3 □

3.4 Minor arcs

In this section we will use the same approach as Vaughan uses in his book [Vau81] to prove Vinogradov's thee prime theorem (every sufficiently large odd number is the sum of three prime numbers). Once again both Gowers's lecture notes([Gow06]) and Nathanson's book ([Nat96]) have provided some helpful insights. The proof of Vaughan's identity in subsection 3.4.1 has been adapted from Tao ([Tao14]).

Let us remind ourselves what we are going to prove in this section:

Proposition. *Let A be a positive constant, then*

$$|\int_{\mathfrak{m}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha| \ll \frac{n^2}{(\log n)^A}.$$

We again consider the same function.

$$f_n(\alpha) = \sum_{p\leq n} \log(p) e(\alpha p),$$

to estimate the integral over the minor arcs, we will try to bound f over the minor arcs so that we can write

$$\begin{aligned}\int_{\mathfrak{m}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha &\leq \sup_{\alpha\in\mathfrak{m}} |f_n(-2\alpha)| \int_{\mathfrak{m}} |f_n(\alpha)|^2 \mathrm{d}\alpha, \\ &\leq \sup_{\alpha\in\mathfrak{m}} |f_n(-2\alpha)| \int_0^1 |f_n(\alpha)|^2 \mathrm{d}\alpha.\end{aligned}$$

Calculating $\int_0^1 |f_n(\alpha)|^2$ is straightforward, notice that

$$\int_0^1 |f_n(\alpha)|^2 \mathrm{d}\alpha = \int_0^1 f_n(\alpha)\overline{f_n}(\alpha)\mathrm{d}\alpha,$$

where $\overline{c}$ is complex conjugation. Now since complex conjugation distributes over sums we can write

$$\begin{aligned}\int_0^1 |f_n(\alpha)|^2 \mathrm{d}\alpha &= \int_0^1 \sum_{p_1,p_2\leq n} \log(p_1)\log(p_2) e((p_1-p_2)\alpha)\mathrm{d}\alpha, \\ &= \sum_{p_1,p_2\leq n} \log(p_1)\log(p_2) \int_0^1 e((p_1-p_2)\alpha)\mathrm{d}\alpha, \\ &= \sum_{p\leq n} \log^2 p,\end{aligned}$$

using the fact that $\int_0^1 e(n\alpha)\mathrm{d}\alpha = 0$ when $n \neq 0$. The next step is estimating this sum

$$\sum_{p\leq n} \log^2 p \ll n\log n,$$

by applying Chebyshev's estimate (A.7) and so

$$|\int_{\mathfrak{m}} f_n(\alpha)^2 f_n(-2\alpha)\mathrm{d}\alpha| \ll \left(\sup_{\alpha\in\mathfrak{m}} |f_n(-2\alpha)|\right) n\log n,$$

so we need to show that $\sup_{a\in\mathfrak{m}} |f_n(-2a)| \ll n(\log n)^{-A}$ for A arbitrary. Remember that we defined our major arcs

$$\mathfrak{M}(a,q) := \{x \in (0,1) : |x - a/q| < P/n, a \leq q \leq P\},$$

where $P = (\log n)^B$, we said that we would choose B later, and it turns out we can choose B depending on A to get the inequality above.

3.4.1 Vaughan's Identity

To estimate $f_n(\alpha)$, we will split the sum up into three pieces using a decomposition called Vaughan's identity. Afterwards we estimate each sum seperately.

Lemma 3.7 (Vaughan's Identity)**.** *Let $f_n(\alpha) = \sum_{p\leq n} \log(p)e(\alpha p)$ and $X = n^{2/5}$, then*

$$f_n(\alpha) = S - U - V + O(\sqrt{n}),$$

where

$$S = \sum_{x\leq X} \sum_{y\leq n/x} \mu(x)\log(y)e(\alpha xy),$$

$$U = \sum_{x\leq X^2} \sum_{y\leq n/x} c_x e(\alpha xy) \qquad \text{where } c_x = \sum_{\substack{d,r\leq X \\ dr=x}} \mu(d)\Lambda(r),$$

$$V = \sum_{\substack{x,y>X \\ xy\leq n}} \sum_{\substack{d\leq X \\ d|x}} \mu(d)\Lambda(y)e(\alpha xy),$$

where Λ is the Von-Mangoldt function as defined in the appendix. (The choice of $X = n^{2/5}$ is not used in the proof, but it is what we will use in later sections).

Proof. We will follow Tao([Tao14]) for the proof. For readers that haven't seen Dirichlet convolution before, now might be a good moment to look at it's definition in the appendix (A.4)and to convince yourself of some basic properties such as commutativity and associativity. We will also use some basic identities involving the Von-Mangoldt and Möbius functions that can also be found in the appendix.

We first prove that $f_n(\alpha) = \sum_n \Lambda(n)e(\alpha n) + O(\sqrt{n})$. Consider the difference

$$\begin{aligned}\sum_n \Lambda(n)e(\alpha n) - f_n(\alpha) &= \sum_{\substack{p^k\leq n \\ k\geq 2}} \log(p)e(\alpha n), \\ &\leq \sum_{\substack{p^k\leq n \\ k\geq 2}} \log p, \\ &\leq \sum_{p\leq\sqrt{n}} \log p = O(\sqrt{n}),\end{aligned}$$

using Chebyshev's estimate in the last step.

Now following Tao ([Tao14], proof of Lemma 4.11), we define

$$1_{\leq X}(x) = \begin{cases} 1 & \text{if } x\leq X \\ 0 & \text{if otherwise} \end{cases}, \qquad 1_{>X}(x) = \begin{cases} 1 & \text{if } x> X \\ 0 & \text{if otherwise} \end{cases},$$

$$\Lambda_{\leq X}(x) = \Lambda(x)1_{\leq X}(x), \qquad \Lambda_{>X}(x) = \Lambda(x)1_{>X}(x),$$

$$\mu_{\leq X}(x) = \mu(x)1_{\leq X}(x), \qquad \mu_{>X}(x) = \mu(x)1_{>X}(x).$$

We can now write

$$\begin{aligned}\mu \star \Lambda &= (\mu_{\leq X} + \mu_{>X}) \star (\Lambda_{\leq X} + \Lambda_{>X}), \\ &= (\mu_{\leq X} + \mu_{>X}) \star \Lambda_{\leq X} + \mu_{\leq X} \star \Lambda_{>X} + \mu_{>X} \star \Lambda_{>X}, \\ &= \mu \star \Lambda_{\leq X} + \mu_{\leq X} \star \Lambda - \mu_{\leq X} \star \Lambda_{\leq X} + \mu_{>X} \star \Lambda_{>X}.\end{aligned}$$

We know now convolve the entire expression with the identity function 1 to get rid of the μ on the left side. This is done using $\mu \star 1 = e$, the identity for convolution (Theorem A.2). We also apply $\Lambda \star 1 = \log$ and obtain

$$\begin{aligned}\Lambda &= (1 \star \mu) \star \Lambda_{\leq X} + \mu_{\leq X} \star (\Lambda \star 1) - \mu_{\leq X} \star \Lambda_{\leq X} \star 1 + \mu_{>X} \star \Lambda_{>X} \star 1, \\ &= \Lambda_{\leq X} + \mu_{\leq X} \star \log - \mu_{\leq X} \star \Lambda_{\leq X} \star 1 + \mu_{>X} \star \Lambda_{>X} \star 1\end{aligned}$$

Now we can sum $\sum_{x \leq n} \Lambda(x)\, e(\alpha x)$ using this identity and see

$$\begin{aligned}\sum_{x \leq n} \Lambda(x)\, e(\alpha x) = &\sum_{x \leq X} \Lambda(x)\, e(\alpha x) + \sum_{x \leq n} (\mu_{\leq X} \star \log)(x)\, e(\alpha x) \\ &- \sum_{x \leq n} (\mu_{\leq X} \star \Lambda_{\leq X} \star 1)(x)\, e(\alpha x) + \sum_{x \leq n} (\mu_{>X} \star \Lambda_{>X} \star 1)(x)\, e(\alpha x),\end{aligned}$$

We get rid of the first term by noticing that $\Lambda(x) \leq \log(X)$ and then using triangle inequality to estimate the first term to $X \log X = O(\sqrt{n})$. We can rewrite the second term as follows, write $y = x/d$ to see

$$\begin{aligned}\sum_{x \leq n} (\mu_{\leq X} \star \log)(x)\, e(\alpha x) &= \sum_{x \leq n} \sum_{\substack{d \mid n \\ d \leq X}} \mu(d) \log(x/d)\, e(\alpha x), \\ &= \sum_{d \leq X} \sum_{y \leq n/d} \mu(d) \log(y) e(\alpha d y) = S.\end{aligned}$$

We can do something similar for the third term ($y = x/rd$)

$$\begin{aligned}-\sum_{x \leq n} (\mu_{\leq X} \star \Lambda_{\leq X} \star 1)(x)\, e(\alpha x) &= -\sum_{x \leq n} \sum_{\substack{d \mid x \\ d \leq X}} \mu(d) \sum_{\substack{r \mid \frac{x}{d} \\ r \leq X}} \Lambda(r)\, e(\alpha x), \\ &= -\sum_{d \leq X} \sum_{r \leq X} \sum_{y \leq n/rd} \mu(d) \Lambda(r) e(\alpha y r d), \\ &= -\sum_{z \leq X^2} \sum_{y \leq n/z} e(\alpha z y) \sum_{\substack{r,d \leq X \\ rd = z}} \mu(d) \Lambda(r) = -U,\end{aligned}$$

where in the last step, we substituted $z = dr$ and sum over all possible combinations of r and d in the innermost sum.

Lastly, the fourth term:

$$\sum_{x \leq n} (\mu_{>X} \star \Lambda_{>X} \star 1)(x)\, e(\alpha x) = \sum_{x \leq n} (\Lambda_{>X} \star \mu_{>X}) \star 1)(x)\, e(\alpha x),$$

$$= \sum_{x \leq n} \sum_{\substack{d|n \\ d>X, \frac{x}{d}>X}} \mu(d)\Lambda\left(\frac{x}{d}\right) \sum_{k|\frac{x}{d}} 1,$$
$$\sum_{\substack{ydr \leq n \\ d,r>X}} \mu(d)\Lambda(r)e(\alpha ydr).$$

Substitute $k = yd$ to see

$$\sum_{\substack{ydr \leq n \\ d,r>X}} \mu(d)\Lambda(r)e(\alpha ydr) = \sum_{\substack{rk \leq n \\ r,k>X}} \Lambda(r) \sum_{\substack{d|k \\ d>X}} \mu(d),$$

If we complete the sum over d, it vanishes since $(\mu \star 1)(k) = 0$ because $k > 0$. So therefore we can instead sum $-\mu(d)$ over $d \leq X$. This implies

$$\sum_{\substack{ydr \leq n \\ d,r>X}} \mu(d)\Lambda(r)e(\alpha ydr) = -\sum_{\substack{rk \leq n \\ r,k>X}} \Lambda(r) \sum_{\substack{d|k \\ d \leq X}} \mu(d) = -V.$$

□

3.4.2 Estimating the pieces

We will now provide an estimate for S, U and V. We will use the same approach as Vaughan uses in his book [Vau81], once again both Gowers([Gow06]) and Nathanson ([Nat96]) have provided some helpful insights.

Lemma 3.8. *Let S be as above and let (a, q) be integers such that $|\alpha - a/q| < q^{-2}$, then $|S| \ll (\log n)^2(\frac{n}{q} + n^{2/5} + q)$.*

Proof. We write

$$|S| = |\sum_{x \leq X} \mu(x) \sum_{y \leq n/x} \log(y)e(\alpha xy)|,$$
$$\leq \sum_{x \leq X} |\sum_{y \leq n/x} \log(y)e(\alpha xy)|.$$

We would like to be able to pull the $\log y$ out of the absolute value by saying $\log y < \log n$ but this would be incorrect since $e(\alpha xy)$ takes both positive and negative values. We can circumvent this problem using partial summation, we use Theorem A.1 with $a_y = e(\alpha xy)$ and $h(x) = \log x$. The Theorem gives us

$$\left| \sum_{y \leq n/x} e(\alpha xy) \log x \right| = \left| \log(\tfrac{n}{x}) \sum_{y \leq n/x} e(\alpha xy) - \int_1^{n/x} \sum_{1 \leq y \leq t} e(\alpha xy) \frac{1}{t} \mathrm{d}t \right|,$$

$$\begin{aligned}
&\leq \log(\tfrac{n}{x})|\sum_{y\leq n/x} e(\alpha xy)| + \int_1^{n/x} |\sum_{1\leq y\leq t} e(\alpha xy)|\frac{1}{t}\mathrm{d}t, \qquad \text{(triangle inequality)}\\
&\leq \log(\tfrac{n}{x})\min[n/x, \|\alpha x\|^{-1}] + \int_1^{n/x} \min[t], [\|\alpha x\|^{-1}]\frac{1}{t}\mathrm{d}t, \qquad \text{(Lemma 3.4)}\\
&\leq \log(\tfrac{n}{x})\min[n/x, \|\alpha x\|^{-1}] + \log(\tfrac{n}{x}) \sup_{1\leq t\leq n/x} \left(\min[t], [\|\alpha x\|^{-1}]\right),\\
&\leq 2\log(\tfrac{n}{x})\min[n/x, \|\alpha x\|^{-1}].
\end{aligned}$$

Now we can sum over x and use Proposition 3.3 which implies

$$\begin{aligned}
|S| \leq \sum_{x\leq X} 2\log(\tfrac{n}{x})\min[n/x, \|\alpha x\|^{-1}] &\leq \log n \sum_{x\leq X} \min[n/x, \|\alpha x\|^{-1}],\\
&\leq (\log n)^3\left(\frac{n}{q} + X + q\right),\\
&= (\log n)^3\left(\frac{n}{q} + n^{2/5} + q\right).
\end{aligned}$$

□

Lemma 3.9. *Let U be as above and let (a, q) be integers such that $|\alpha - a/q| < q^{-2}$, then $|U| \ll (\log n)^2(\frac{n}{q} + n^{4/5} + q)$*

Proof. We will first estimate c_x, write

$$\begin{aligned}
|c_x| &= |\sum_{\substack{d,y\leq X\\ dy=x}} \mu(d)\Lambda(y)|,\\
&= |\sum_{\substack{d\leq X\\ d|x}} \mu\left(\frac{x}{d}\right)\Lambda(d)|,\\
&\leq |\sum_{\substack{d\leq X\\ d|x,}} \Lambda(d)|,\\
&\leq |\sum_{d|x} \Lambda(d)|,\\
&= \log(x),
\end{aligned}$$

using the identity $\sum_{d|x}\Lambda(d) = \log x$ (Appendix A.5). The rest of the estimate is straightforward using our estimate for exponential sums.

$$\begin{aligned}
|U| &= |\sum_{x\leq X^2}\sum_{y\leq n/x} c_x e(\alpha xy)|,\\
&= |\sum_{x\leq X^2} c_x \sum_{y\leq n/x} e(\alpha xy)|,
\end{aligned}$$

$$\begin{aligned}
&\leq \sum_{x\leq X^2} |c_x| \left| \sum_{y\leq n/x} e(\alpha xy) \right|, \\
&\leq \log n \sum_{x\leq X^2} |\sum_{y\leq n/x} e(\alpha xy)|, \qquad \text{(using } \log(x) \leq \log(n)\text{)}, \\
&\leq \log n \sum_{x\leq X^2} \min[n/x, \|\alpha x\|^{-1}],
\end{aligned}$$

where we used Lemma 3.4 for the last step. Applying Proposition 3.3 gives us the estimate

$$|U| \ll (\log n)^2 \left(\frac{n}{q} + n^{4/5} + q \right),$$

remembering that $X = n^{2/5}$. □

The last term will take some effort :

Lemma 3.10. *Let V be as above and let (a, q) be integers such that $|\alpha - a/q| < q^{-2}$, then $|V| \ll (\log n)^4 \left(\frac{n}{\sqrt{q}} + n^{4/5} \right)$.*

Proof. Write $t_X(x) = \sum_{d\leq X, d|x} \mu(d)$ as before and note that this allows us to write

$$V = \sum_{\substack{x>X, y>X \\ xy\leq n}} t_X(x)\Lambda(y)e(\alpha xy).$$

We divide the interval $[X, n^{3/5}]$(this is the range that we sum x over) into so called *dyadic* pieces, writing k for the first integer such that $2^{k+1} > X$. Take l to be the first integer such that $2^l > n^{3/5}$. We enlarge our sum, letting x run over $[2^k, 2^l]$. This does not, however change our sum, since if $x > n^{3/5}$ then there is no $y > X$ such that $xy \leq n$ because $n^{3/5} = n/X$. We make sure to truncate our first dyadic interval at X, so as not to gain extra terms, this is only a slight technicality and will not matter in the estimation of U_i. Using this observation we write

$$|V| = \sum_{i=k}^{j} |U_i|,$$

where we define for $i > k$

$$U_i = \sum_{2^{i-1}\leq x<2^i-1} \sum_{X<y\leq n/x} t_X(x)\Lambda(y)e(\alpha xy),$$

and for $i = k$

$$U_k = \sum_{X\leq x<2^{k+1}} \sum_{X<y\leq n/x} t_X(x)\Lambda(y)e(\alpha xy).$$

We will estimate the U_i separately and uniformly in i, and then use the fact that there are at most $O(\log n)$ terms to sum over.

We can use the Cauchy-Schwarz inequality for sums in a clever way to derive from

$$U_i = \sum_{2^{i-1}\le x<2^i-1} t_X(x) \sum_{X<y\le n/x} \Lambda(y)e(\alpha xy),$$

the inequality

$$|U_i|^2 \le \left(\sum_{2^{i-1}\le x<2^i-1} |t_X(x)|^2\right)\left(\sum_{2^{i-1}\le x<2^i-1} |\sum_{X<y\le n/x} \Lambda(y)e(\alpha xy)|^2\right).$$

To estimate the first sum we notice that

$$|T_X(x)| = |\sum_{\substack{d\le X\\ d|x}} \mu(d)| \le \sum_{\substack{d\le X\\ d|x}} 1 \le \sum_{d|x} 1 = d(x)$$

and remember that $\sum_{a\le 2^i} d(a)^2 \ll 2^i \log(n)^3$ (A.8). For the second sum we observe that $|f|^2 = f\overline{f}$ where $\overline{f}$ is complex conjugation. We can now expand the square

$$\begin{aligned}\sum_{2^{i-1}\le x<2^i-1} \left|\sum_{X<y\le n/x} \Lambda(y)e(\alpha xy)\right|^2 &= \sum_{2^{i-1}\le x<2^i-1} \left(\sum_{X<y\le n/x} \Lambda(y)e(\alpha xy)\right)\left(\sum_{X<y'\le n/x} \Lambda(y)e(-\alpha xy')\right),\\ &= \sum_{2^{i-1}\le x<2^i-1} \sum_{X<y\le n/x} \sum_{X<y'\le n/x} \Lambda(y)\Lambda(y')e(\alpha x(y-y')).\end{aligned}$$

We now want to exchange the order of summation, summing over x in the inner sum so that we can apply our exponential sum estimate. This change is quite tricky, but it is best explained as follow: both y and y' run between X and n/x, and x runs between 2^{i-1} and 2^i-1. If we want to sum over y and y' first, we need to make sure that all inequalities hold simultaneously. It is quite clear that y and y' need to run between X and 2^i-1, but to make sure that $y, y' \le n/x$ we need that $x \le n/y$ and n/y', and so $x \le \min[n/y, n/y']$.

$$\begin{aligned}\sum_{2^{i-1}\le x<2^i-1} \sum_{\substack{X<y\le n/x\\ X<y'<n/x}} \Lambda(y)\Lambda(y')e(\alpha x(y-y')) &= \sum_{X<y\le \frac{n}{2^{i-1}}} \sum_{X<y'\le \frac{n}{2^{i-1}}} \sum_{\substack{2^{-1}<x\le 2^i-1\\ x\le \min[n/y,n/y']}} \Lambda(y)\Lambda(y')e(\alpha x(y-y')),\\ &\le \sum_{X<y\le \frac{n}{2^{i-1}}} \sum_{X<y'\le \frac{n}{2^{i-1}}} \Lambda(y)\Lambda(y')\min\left[2^{i-1}, \|\alpha(y-y')\|^{-1}\right],\end{aligned}$$

because there are at most 2^{i-1} terms in the inner sum and the partial geometric series is always smaller than $\|\alpha(y-y')\|^{-1}$ by Lemma 3.4. We bound $\Lambda(y) \le \log n$ to simplify our sum:

$$|U_i|^2 \le \left(\sum_{2^{i-1}\le x<2^i-1} |t_X(x)|^2\right)\left(\sum_{2^{i-1}\le x<2^i-1} |\sum_{X<y\le n/x} \Lambda(y)e(\alpha xy)|^2\right),$$

$$\ll \left((2^i-1)\log(n)^3\right)\log(n)^2 \sum_{X<y\leq \frac{n}{2^{i-1}}} \sum_{X<y'\leq \frac{n}{2^{i-1}}} \min\left[2^{i-1}, \|\alpha(y-y')\|^{-1}\right],$$
$$\ll 2^i(\log n)^5 \sum_{X<y\leq \frac{n}{2^{i-1}}} \sum_{X<y'\leq \frac{n}{2^{i-1}}} \min\left[2^{i-1}, \|\alpha(y-y')\|^{-1}\right].$$

At this point we use another trick; every difference $z = y - y'$ occurs at most $\frac{n}{2^{i-1}}$ times because there are at most $\frac{n}{2^{-1}}$ ways to write a numbers below $\frac{n}{2^{-1}}$ as the difference of two numbers below $\frac{n}{2^{-1}}$(which is a gross overestimation). This gives

$$|U_i|^2 \ll 2^i(\log n)^5 \frac{n}{2^{i-1}} \sum_{z\leq n/2^i} \min\left[2^{i-1}, \|\alpha z\|^{-1}\right].$$

In the last step of the proof of Lemma 3.10 we proceed similarly as in the proof of Proposition 3.3. We divide $n/2^i$ into pieces of size $q/2$, assuming without loss of generality that q is even. We apply Lemma 3.6 with $r = q/2$ and $Q = 2^{i-1}$.

$$|U_i|^2 \ll n(\log n)^5 \sum_{j=1}^{\frac{n}{2^{i+1}q}} \sum_{z=\frac{jq}{2}+1}^{\frac{(j+1)q}{2}} \left[2^{i-1}, \|\alpha(z)\|^{-1}\right],$$
$$\ll n(\log n)^5 \sum_{j=1}^{\frac{n}{2^{i+1}q}} \log(2^{i-1})(2^{i-1}+\tfrac{q}{2}),$$
$$\ll n(\log n)^6 \sum_{j=1}^{\frac{n}{2^{i+1}q}} 2^{i-1}+\tfrac{q}{2},$$
$$= n(\log n)^6(\tfrac{n}{4q}+\tfrac{n}{2^{i+1}}),$$
$$\ll n(\log n)^6(\tfrac{n}{q}+\tfrac{n}{X}),$$
$$\ll n(\log n)^6(\tfrac{n}{q}+\tfrac{n}{X}),$$
$$|U_i| \ll (\log n)^3\left(\frac{n}{\sqrt{q}}+n^{4/5}\right),$$

This implies

$$|V| \leq |\sum_{i=k}^{l} U_i| \leq \log n \max_i |U_i| \ll (\log n)^4\left(\frac{n}{\sqrt{q}}+n^{4/5}\right),$$

using that $\log_2(n) \ll \log n$. This finishes the proof of Lemma 3.10 □

3.4.3 Applying the estimates

We have been estimating $f_n(\alpha)$ all this time when what we really wanted was an estimate for $f(-2\alpha)$. Luckily this will just be a small technicality: If α is in the minor arcs, then either -2α is also in the minor

arcs (at which point we are done) or it is in a major arc $\mathfrak{M}(a,q)$. Now this means $|-2\alpha - a/q| < P/n$ and so $|\alpha + a/2q| < P/2n$, if $2|a$ this fraction simplifies and we can conclude that α is in a major arc, a contradiction. Since α is not in a major arc, $2q$ must be too big, i.e., $2q > P$ and so $q > P/2$.

Using Lemma 3.7 we decompose f into pieces, which we can estimate using Lemma's 3.8, 3.9, 3.10 proven earlier. We conclude that

$$\begin{aligned} |f_n(-2\alpha)| &\ll (\log n)^4 \left(\frac{n}{\sqrt{q}} + n^{4/5}\right), \\ &\ll (\log n)^4 \left(\frac{n}{\sqrt{q}}\right), \end{aligned}$$

here q is the denominator of a reduced fraction a/q with $|-2\alpha - \frac{2a}{q}| < q^{-2}$. If -2α is in a major arc $\mathfrak{M}(a,q)$ we know that $q > P/2$. If -2α is in a minor arc we can approximate it arbitrarily well by fractions (by Dirichlet's method or continued fractions) and so we find a reduced fraction a/q such that $|-2\alpha - a/q| < P/n$. This then allows us to conclude that $q > P$ because -2α is in the minor arcs.

Clearly $q > P/2$ is the worst case scenario and so aplying this to our integral estimate at the beginning of the section gives

$$\begin{aligned} |f_n(-2\alpha)| &\ll (\log n)^4 \left(\frac{n}{\sqrt{P/2}}\right), \\ &\ll \frac{n}{(\log n)^{B/2-4}}, \\ |\int_{\mathfrak{m}} f_n(\alpha)^2 f_n(-2\alpha) \mathrm{d}\alpha| &\leq \sup_{\alpha \in \mathfrak{m}} |f_n(-2\alpha)| n \log n, \\ &\ll \frac{n^2}{(\log n)^{B/2-5}}, \end{aligned}$$

Now choosing $B = 2A + 10$ we complete the proof of Proposition 3.2, which we shall repeat below:

Proposition. *Let A be a positive constant, then*

$$|\int_{\mathfrak{m}} f_n(\alpha)^2 f_n(-2\alpha) \mathrm{d}\alpha| \ll \frac{n^2}{(\log n)^A}.$$

3.5 Conclusion

Combining Proposition 3.2 and Proposition 3.1 gives us Theorem 1.3 which we will repeat a third time here:

Theorem (Van der Corput (1939))**.** *Let $A > 0$ be a real number, and let $n \in \mathbb{N}$, then the number of 3AP's in the primes up to n, denoted by $S(n)$, satisfies*

$$S(n) \geq \mathfrak{S}\frac{n^2}{4(\log n)^3} + O\left(\frac{n^2}{(\log n)^A}\right),$$

where $\mathfrak{S} \geq 1$

Proof. We have seen from Proposition 3.2 and Proposition 3.1 that

$$\begin{aligned} R(n) &= \sum_{\substack{p_1,p_2,p_3 \leq n \\ p_1+p_2=2p_3}} \log p_1 \log p_2 \log p_3, \\ &= \int_0^1 f_n(\alpha)^2 f_n(-2\alpha) \mathrm{d}\alpha, \\ &= \mathfrak{S}\frac{n^2}{4} + O\left(\frac{n^2}{(\log n)^A}.\right), \end{aligned}$$

but we want an unweighted estimate $S(n)$, which measures the number of 3AP's in the primes below n. The following estimate is trivial when n is large enough.

$$R(n) \leq S(n)(\log n)^3$$

This concludes the proof of Theorem 1.3 [1] □

[1] It is possible to prove an upper bound for $S(N)$ which implies an asymptotic. Unfortunately the author has not been able to produce such a proof nor been able track one down.

A Selected results

A.1 Partial Summation

Theorem A.1 (Partial Summation, [Mur08] Theorem 2.1.1)**.** *Let a_n be a sequence of complex numbers and let $h : \mathbb{R} \to C$ be a continuously differentiable function, then*

$$\sum_{x \leq X} a_x h(x) = f(X) \sum_{x \leq X} a_x - \int_1^X h'(t) \sum_{x \leq t} a_x \mathrm{d}t.$$

A.2 Results from Hardy and Wright

Many results from Hardy and Wright are quoted in the text, I have decided to list them here for convenience.

Theorem A.2 ([HW08], Theorem 263)**.** *Let $\mu(n)$ be the Möbius function defined as*

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if there is a square dividing } n \\ (-1)^k & \text{if } n = p_1 \cdot p_2 \cdots p_k \end{cases},$$

then

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

Theorem A.3 ([HW08], Theorem 272)**.** *Let $c_n(m)$ be Ramanujan's sum defined as*

$$c_n(m) = \sum_{\substack{1 \leq h \leq n \\ (h,n)=1}} e\left(\frac{hm}{n}\right),$$

then

$$c_n(m) = \frac{\mu(\frac{n}{(n,m)})\phi(n)}{\phi(\frac{n}{(n,m)})}.$$

Theorem A.4 ([HW08], Theorem 286)**.** *Let $f(m)$ be a multiplicative function, that means that $f(nm) = f(n)f(m)$ when n and m are relatively prime. If $\sum_{n=1}^{\infty} f(n)$ converges absolutely then*

$$\sum_{n=1}^{\infty} f(n) = \prod_p \left(1 + f(p) + f(p^2) + \cdots\right).$$

Theorem A.5 ([HW08] Theorem 296)**.** *Let Λ be the von Mangoldt function defined as*

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \\ 0 & \text{if otherwise} \end{cases},$$

then

$$\sum_{d|n} \Lambda(d) = \log n, \qquad \sum_{d|n} \mu(d) \log(\tfrac{n}{d}) = \Lambda(n).$$

Theorem A.6 ([HW08], Theorem 327)**.** *Let $\phi(n)$ be the Euler phi function, which is defined as the number of natural numbers below n that are coprime to n and let $\delta > 0$, then*

$$\lim_{n\to\infty} \frac{\phi(n)}{n^{1-\delta}} \to \infty.$$

Theorem A.7 ([HW08], Theorem 415)**.** *The following inequalities hold*

$$\theta(x) := \sum_{p\leq x} \log p \leq 2x \log 2 = O(x),$$
$$\psi(x) := \sum_{n\leq x} \Lambda(x) = O(x).$$

A.3 Other Results

Theorem A.8 ([Gow06], Lemma 3)**.** *The following inequality holds, let $m \in \mathbb{N}$ then*

$$\sum_{x\leq m} d(x)^2 := \sum_{x\leq n} \left(\sum_{d|x} 1 \right)^2 = O(n(\log n)^3)$$

A.4 Dirichlet convolution

Definition A.1. *Let $f, g : \mathbb{N} \to \mathbb{C}$ be functions, then their Dirichlet convolution $f \star g$ is defined as*

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

This operation is commutative and associative, has the function

$$e(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if otherwise} \end{cases}$$

as identity element. We will also use the following identities proven in Theorem A.5 and Theorem A.2.

$$\mu \star 1 = e, \qquad \mu \star e = 1,$$
$$\mu \star \log = \Lambda, \qquad \Lambda \star \mu = \log.$$

A.5 A trigonometric estimate

Theorem A.9. *The following inequality holds with* $|B| = \lfloor \frac{\sqrt{N}}{\pi} \rfloor$, $d \leq \sqrt{N}$ *and* $dt \leq \sqrt{N} \mod N$*:*

$$\sum_{k=-\frac{|B|}{2}}^{\frac{|B|}{2}-1} |\left(e^{-\frac{2\pi i}{N}kdt} - 1\right)| \leq \frac{|B|}{2}$$

Proof. We are summing the distance between the point $(0,1)$ and a point on the unit circle with angle $\frac{2\pi kdt}{N}$, and since $|k| \leq \frac{\sqrt{N}}{\pi}$ and $dt \leq \sqrt{N}$ we are summing over angles between -1 and 1 radians. By symmetry we can sum twice over all positive angles to see

$$\sum_{k=-\frac{|B|}{2}+1}^{\frac{|B|}{2}-1} |\left(e^{-\frac{2\pi i}{N}kdt} - 1\right)| = 2\sum_{k=0}^{\frac{|B|}{2}-1} |\left(e^{\frac{2\pi i}{N}kdt} - 1\right)|.$$

Now we use some geometry to see that $|e^{i\alpha} - 1| < \alpha$ since α is the length of the arc going to $e^{i\alpha}$ which is longer than the path directly from 1 to $e^{i\alpha}$. This gives

$$\begin{aligned}
\sum_{k=0}^{\frac{|B|}{2}-1} |\left(e^{\frac{2\pi i}{N}kdt} - 1\right)| &\leq \sum_{k=0}^{\frac{|B|}{2}-1} \frac{2\pi kdt}{N}, \\
&= \frac{4\pi dt}{N} \sum_{k=0}^{\frac{|B|}{2}-1} k, \\
&\leq \frac{4\pi}{\sqrt{N}} \sum_{k=0}^{\frac{|B|}{2}-1} k, \\
&= \frac{\pi}{\sqrt{N}}|B|\frac{|B|-1}{2}, \\
&\leq \frac{|B|-1}{2}, \\
&\leq \frac{|B|}{2}
\end{aligned}$$

□

References

[Cra35] H. Cramér. *Prime numbers and probability.* Skand. Math. Kongr, 8:107–115, 1935.

[FS09] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics.* Cambridge University Press, Cambridge, 2009.

[Gow] T. Gowers. *Weyl's Inequality.* Chapter 3 of lecture notes available at `https://www.dpmms.cam.ac.uk/~wtg10/addnoth.notes.dvi`.

[Gow06] T. Gowers. *Vinogradov's three primes theorem.* Lecture notes available at `https://www.dpmms.cam.ac.uk/~wtg10/3primes.dvi`, 2006.

[GT08] Ben Green and Terence Tao. *The primes contain arbitrarily long arithmetic progressions.* Ann. of Math. (2), 167(2):481–547, 2008.

[HW08] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers.* Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

[Lya05] N. Lyall. *Roth's theorem, the Fourier analytic approach.* Lecture notes available at `http://www.math.uga.edu/~lyall/REU/Roth.pdf`, 2005.

[Mur08] M. Ram Murty. *Problems in analytic number theory*, volume 206 of *Graduate Texts in Mathematics.* Springer, New York, second edition, 2008. Readings in Mathematics.

[MV07] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2007.

[Nat96] Melvyn B. Nathanson. *Additive number theory*, volume 164 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1996. The classical bases.

[Rot53] K. F. Roth. *On certain sets of integers.* J. London Math. Soc., 28:104–109, 1953.

[Sze69] E. Szemerédi. *On sets of integers containing no four elements in arithmetic progression.* Acta Math. Acad. Sci. Hungar., 20:89–104, 1969.

[Sze75] E. Szemerédi. *On sets of integers containing no k elements in arithmetic progression.* Acta Arith., 27:199–245, 1975. Collection of articles in memory of Juriĭ Vladimirovič Linnik.

[Tao14] Terence Tao. *Every odd number greater than* 1 *is the sum of at most five primes.* Math. Comp., 83(286):997–1038, 2014.

[Vau81] R. C. Vaughan. *The Hardy-Littlewood method*, volume 80 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge-New York, 1981.

[Wik] Wikipedia. *Twin Primes.* Website available at `http://en.wikipedia.org/wiki/Twin_prime`.