The Ternary Goldbach Conjecture

Luke Collins

under the supervision of
Prof. Adam Harper

THE UNIVERSITY OF
WARWICK

30th March, 2020
MA469: Research Project
Mathematical Institute

EU funds
for Malta
2014-2020

MINISTRY FOR
EDUCATION AND
EMPLOYMENT

Abstract

We review Hardy–Littlewood's proof of the ternary Goldbach conjecture for sufficiently large odd numbers, which assumes the Generalised Riemann Hypothesis, then discuss Vinogradov's improvement of the minor arcs bound to prove the result unconditionally (i.e., Vinogradov's theorem), and finally explore some ideas from Helfgott's 2014 proof of the ternary Goldbach conjecture for all odd numbers larger than 7.

The research work disclosed in this publication is partially funded by the ENDEAVOUR Scholarships Scheme. The scholarship is part-financed by the European Social Fund.

Operational Programme II
European Structural and Investment Funds 2014 - 2020
*"Investing in human capital to create more opportunities and
promote the wellbeing of society"*
Scholarships part-financed by the European Social Fund
Co-financing rate: 80% EU Funds; 20% National Funds

Acknowledgements

I was able to pursue this Master's degree at Warwick thanks to the funding provided by the Maltese government's ENDEAVOUR scholarship scheme, for which I am very thankful. The scholarship is part-financed by the EU European Social Fund.

I am lucky to have had Prof. Adam Harper as a supervisor for this project. The discussions we had together were always very insightful, both on material related to the project and on analytic number theory generally, as well as career advice. This Master's course was my first real exposition to analytic number theory, and I couldn't have had a better person to guide me.

There are also other people who helped me along my journey of mathematical education, some of whom I am compelled to mention.

First of all, I would like to start by thanking my parents for their undying love and support, and for always ensuring that I received the best education, even if it meant sacrifices for them.

I would also like to thank the members of faculty at the Mathematics Department of the University of Malta, where I completed my undergraduate studies. Especially Prof. Irene Sciriha for providing me with the opportunity to carry out mathematical research at the department, as well as for supervising my undergraduate dissertation. I would also like to mention Prof. Joseph Muscat, Prof. Peter Borg and the head of department, Prof. David Buhagiar, for providing me with advice about furthering my studies. Prof. Adrian Francalanza from the Department of Computer Science also gave me excellent advice about this.

The support of my friends is always important, for that I would like to thank Stefania, Pierre, Steven and Daniele.

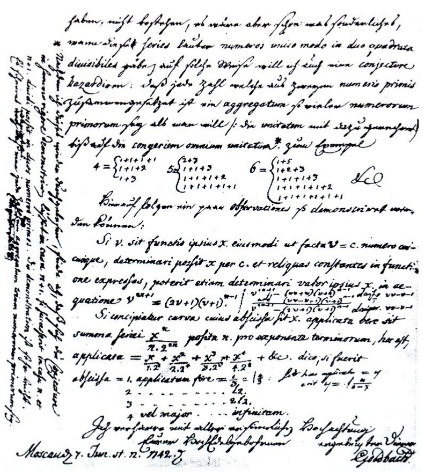
Finally, I would like to thank my A-level maths teacher, Sandro Grech from St Aloysius' College, who definitely inspired me to continue studying mathematics at university level.

Table of Contents

Introduction

On the 7th of June 1742, German mathematician Christian Goldbach conjectured in a letter to Leonhard Euler that *every even number greater than two can be expressed as a sum of two primes.* Proof of this statement, which is well known today as *Goldbach's conjecture*, has evaded mathematicians to the present day, and remains one of the oldest and best known unsolved problems in mathematics.

The so-called *ternary* Goldbach conjecture is a "weaker" analogue to the Goldbach conjecture. It was proven to be true in a preprint published on ArXiV by H. Helfgott in 2014,[10] and states the following.



haben, nicht bestehen, es wäre aber schon mal sonderlich, wenn dieselbe series lauter numeros unico modo in duo quadrata divisibiles gäbe. Auf solche Weise will ich auch eine conjecture hazardiren: daß jede Zahl welche aus zweyen numeris primis zusammengesetzt ist ein aggregatum so vieler numerorum primorum sey als man will (: die unitatem mit dazu gerechnet :) bis auf die congeriem omnium unitatum*. zum Exempel

$4 = \begin{cases} 1+1+1+1 \\ 1+1+2 \\ 1+3 \end{cases}$ $5 = \begin{cases} 2+3 \\ 1+1+3 \\ 1+1+1+2 \\ 1+1+1+1+1 \end{cases}$ $6 = \begin{cases} 1+5 \\ 1+2+3 \\ 1+1+1+3 \\ 1+1+1+1+2 \\ 1+1+1+1+1+1 \end{cases}$ &c

Hierauf folgen ein paar observationes so demonstriret werden können:

Si v. sit functio ipsius x. eiusmodi ut facta v = c. numero cuicunque, determinari possit x per c. et reliquas constantes in functione expressas, poterit etiam determinari valor ipsius x. in aequatione $v^{2x+1} = (2v+1)(v+1)^{x-1}$

Si concipiatur curva cuius abscissa sit x. applicata vero sit summa seriei $\frac{x^n}{n.2^{2n}}$ posita n. pro exponente terminorum, hoc est, applicata $= \frac{x}{1.2} + \frac{x^2}{2.2^4} + \frac{x^3}{3.2^6} + \frac{x^4}{4.2^8} + \&c.$ dico, si fuerit abscissa = 1. applicatam fore $= \frac{1}{3} = l\frac{4}{3}$

2 . . . . l 2.

3 . . . . 2 l 2.

4 vel major . . . . infinitam.

Ich verharre mit aller ersinnlichen Hochachtung
Euer Hochedelgebohren ergebenster Diener
Goldbach.

Moscau d. 7. Jun. st. n. 1742.

* Nachdem ich dieses wieder durchgelesen, finde ich, daß sich die conjecture in summo rigore demonstriren lässet in casu n+1, si successerit in casu n, et n+1 dividi possit in duos numeros primos. Die demonstration ist sehr leicht. Es scheinet wenigstens, daß eine jede Zahl die größer ist als 2 ein aggregatum trium numerorum primorum sey.

Figure 1.1: Goldbach's letter to Euler

Theorem 1.1 (Ternary Goldbach Conjecture). *Every odd integer greater than* 7 *can be written as the sum of three primes.*

It is weaker in the sense that, if the Goldbach conjecture is true, then the ternary Goldbach conjecture is automatically true. Indeed, suppose the Goldbach conjecture is true, and N is an odd number greater than 7. Then $N-3$ is even, and we can express $N-3 = p_1 + p_2$ using the Goldbach conjecture. Thus, $N = p_1+p_2+3$, and the ternary Goldbach conjecture is true.

What follows is a brief timeline of the results established that lead towards the proof of theorem 1.1. In 1923, the British mathematicians G. H. Hardy and J. E. Littlewood made use of their so-called *circle method* to establish that, assuming the Generalised Riemann Hypothesis, the ternary Goldbach conjecture is true for *sufficiently large* odd numbers.

In other words, there exists some N large enough so that, for all odd numbers greater than N, the statement is true. In 1937, Soviet mathematician I. M. Vinogradov was able to remove the dependence on the Generalised Riemann Hypothesis and proved directly, by studying exponential sums, that the ternary Goldbach conjecture was true, again for sufficiently large N.[19] Vinogradov also did not give a bound for "sufficiently large", but in 1956, a student of Vinogradov, K. Borozdkin, determined that $3^{3^{15}}$ is large enough. Even though this leaves a finite number of cases to check, it is far from feasible to do so ($3^{3^{15}}$ has 6 846 169 digits). In the 2014 preprint, Helfgott uses a tweaked version of the circle method to improve this bound to something much more reasonable: odd numbers larger 10^{27}; the remaining cases can easily be checked by computer (it takes around 5 hours on a modern Linux desktop).[10]

In this project, we shall first go over Hardy–Littlewood's proof (chapter 2), next we will see a simplified version of Vinogradov's proof which makes use of Vaughan's identity to make the argument unconditional (chapter 3), and finally we will discuss how Helfgott obtains a log-free bound for the minor arcs, which is the main novelty of his 2014 proof (chapter 4).

1.1 Notation and Preliminaries

Let us introduce some notation and get some preliminary definitions out of the way, so that in the next section, we may formulate the result more precisely.

Let $\mathscr{P}$ be a proposition. Then we denote by $\mathbb{1}_{\mathscr{P}}$ the indicator function of $\mathscr{P}$, i.e.,

$$\mathbb{1}_{\mathscr{P}} := \begin{cases} 1 & \text{if } \mathscr{P} \text{ is true} \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, if X is a set, we write $\mathbb{1}_X(x)$ for $\mathbb{1}_{x\in X}$. Without a subscript, $\mathbb{1}$ denotes the function $\mathbb{1}(x) = 1$ for all x.

We write $f(x) = O(g(x))$, or equivalently $f(x) \ll g(x)$, to denote the fact that for some $C > 0$, we have $|f(x)| \leqslant C\, g(x)$ for all x under consideration; usually for all x larger than a fixed constant. If both $f(x) \ll g(x)$ and $g(x) \ll f(x)$, then we write $f \asymp g$. For $g(x) \neq 0$, we write $f(x) = o(g(x))$ as $x \to \infty$ if for all $\epsilon > 0$, there exists $N > 0$ such that $|f(x)| \leqslant \epsilon\, g(x)$ for all $x \geqslant N$. Finally, we write $f(x) \sim g(x)$ if $f(x) - g(x) = o(g(x))$. This is stronger than the $\asymp$ relation; for instance, we have $2x \asymp x$ but $2x \nsim x$ as $x \to \infty$.

$\mathbb{R}/\mathbb{Z}$ denotes some real interval of length 1 (such as $[0, 1]$), and we define the map $e\colon \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ by $e(\theta) := e^{2\pi i\theta}$. The "circle" in the circle method is the image of $\mathbb{R}/\mathbb{Z}$ under this map.

For a ring R, we denote the group of units by $R^\times$. For instance, the group of integers modulo q which have a multiplicative inverse is denoted $(\mathbb{Z}/q\mathbb{Z})^\times$. This is just the set $\{n \in \{0, \dots, q-1\} : (n, q) = 1\}$, and we denote its cardinality by $\varphi(n)$, where φ is known as the (Euler) totient function.

For $\alpha \in \mathbb{R}$, we denote the distance to the nearest integer by $\|\alpha\|$, i.e., $\|\alpha\| = \min_{n\in\mathbb{Z}} |\alpha - n|$.

The *von Mangoldt function*, $\Lambda\colon \mathbb{N} \to \mathbb{R}$, is defined by

$$\Lambda(n) := \begin{cases} \log p & \text{if } n = p^k \text{ for some prime } p \text{ and integer } k \geqslant 1 \\ 0 & \text{otherwise,} \end{cases}$$

and the *Chebychev* Ψ *function* $\Psi\colon \mathbb{R} \to \mathbb{R}$ is defined by $\Psi(x) := \sum_{n\leqslant x} \Lambda(n)$. These functions play a significant role in analytic number theory, they often make it simpler to phrase major results. For example, the prime number theorem (PNT) $\pi(x) \sim \frac{x}{\log x}$ is equivalent to the fact that $\Psi(x) \sim x$, where $\pi(n) = \#\{p \leqslant n : p \text{ is prime}\}$.

We also make use of the *Möbius function*, denoted by $\mu(n)$, defined by

$$\mu(n) := \begin{cases} (-1)^k & \text{if } n \text{ is square-free and has } k \text{ distinct prime divisors} \\ 0 & \text{otherwise,} \end{cases}$$

where by *square-free*, we mean that there is no $d \geqslant 2$ such that $d^2 \mid n$. Note that $\mu(1) = 1$.

Finally, we introduce *Dirichlet convolution*. For two functions $f, g\colon \mathbb{N} \to \mathbb{C}$, we define their Dirichlet (or multiplicative) convolution $f * g\colon \mathbb{N} \to \mathbb{C}$ by

$$(f * g)(n) := \sum_{d\mid n} f(d)\, g(n/d) = \sum_{ab=n} f(a)\, g(b).$$

Analogously to the usual additive convolutions, we have that Dirichlet convolution is commutative, associative and distributive. To avoid confusion of Dirichlet convolution with additive convolution, we will denote the latter by $f \star g$, i.e.,

$$(f \star g)(t) := \int_{\mathbb{R}} f(\tau)\, g(t-\tau)\, d\tau.$$

The Fourier transform of a function f is denoted by $\hat{f}$, i.e., $\hat{f}(\xi) = \int_{\mathbb{R}} f(x)\, e(x\xi)\, dx$, and the inverse Fourier transform is denoted $\check{f}$, i.e., $\check{f}(\xi) = \hat{f}(-\xi)$. We can relate some of the functions we have introduced using the language of Dirichlet convolutions.

Proposition 1.2 (Dirichlet convolution properties). *Let $\mathbb{1}_{=1}\colon \mathbb{N} \to \mathbb{N}$ denote the function which is 1 at 1, and 0 everywhere else, i.e., $\mathbb{1}_{=1}(n) := \mathbb{1}_{n=1}$. Then we have the following.*

(i) $f = f * \mathbb{1}_{=1} = \mathbb{1}_{=1} * f$,

(ii) $\mathbb{1} * \mu = \mathbb{1}_{=1}$,

(iii) (Möbius inversion). For any $f, g\colon \mathbb{N} \to \mathbb{C}$, $g = f * \mathbb{1}$ if and only if $f = g * \mu$,

(iv) $\mathbb{1} * \Lambda = \log$.

The proofs are straightforward and are omitted for brevity.

1.2 Formulating the Ternary Goldbach Conjecture

Let $\mathbb{1}_P$ be the indicator for the set P of primes. Then observe that for odd N, the sum

$$t(N) := \sum_{k_1+k_2+k_3=N} \mathbb{1}_P(k_1)\,\mathbb{1}_P(k_2)\,\mathbb{1}_P(k_3)$$

over $k_i \in \mathbb{N}$, counts the number of ways in which N can be expressed as the sum of three primes. The statement of the ternary Goldbach conjecture is therefore equivalent to the statement that $t(N) > 0$ for odd $N \geqslant 7$.

It turns out that calculations are made simpler if we instead consider a weighted version of the above sum. Rather than working with $t(N)$, we instead consider

$$r(N) := \sum_{k_1+k_2+k_3=N} \Lambda(k_1)\,\Lambda(k_2)\,\Lambda(k_3),$$

which counts the number of ways N can be expressed as a sum of three prime powers ${p_1}^a + {p_2}^b + {p_3}^c$ (not just primes), with a weight of $\log(p_1)\log(p_2)\log(p_3)$ attached to each such representation. Now, define

$$S(N,\alpha) := \sum_{k\leqslant N} \Lambda(k)\,e(k\alpha).$$

Then

$$S(N,\alpha)^3 = \sum_n r_N(n)\,e(n\alpha),$$

where $r_N(n)$ is defined similarly to $r(n)$, with the added condition that each $k_i \leqslant N$. In particular, we have $r(n) = r_N(n)$ for $n \leqslant N$, and since $S(N,\alpha)^3$ is a trigonometric polynomial, we may extract $r(N)$ by taking the inverse Fourier transform, i.e.,

$$r(N) = \int_{\mathbb{R}/\mathbb{Z}} S(N,\alpha)^3\,e(-N\alpha)\,d\alpha. \tag{1.1}$$

We will see that the integrand turns out to be large when α is close to rational points with small denominators, and by estimating the contribution at these points, we manage to prove the following; from which the ternary Goldbach conjecture follows for N sufficiently large.

Theorem 1.3 (Vinogradov, 1937)**.** *Let N be an odd integer. Then*

$$r(N) \sim \tfrac{1}{2}\,\mathfrak{S}(N)\,N^2, \tag{1.2}$$

where $\mathfrak{S}(N) := \prod_{p\mid N}\left(1 - \frac{1}{(p-1)^2}\right)\prod_{p\nmid N}\left(1 + \frac{1}{(p-1)^3}\right)$.

How does it follow? Firstly, observe that since the nth prime is greater than n,

$$\mathfrak{S}(N) \leqslant \prod_p \left(1 + \frac{1}{(p-1)^3}\right) \leqslant \prod_n \left(1 + \frac{1}{n^3}\right) \leqslant \prod_n \exp\left(\frac{1}{n^3}\right) = \exp(\zeta(3)) \approx 3.32,$$

and that when N is odd,

$$\mathfrak{S}(N) = \prod_{p \mid N} \frac{(p-1)^2 - 1}{(p-1)^2} \prod_{p \nmid N} \frac{(p-1)^3 + 1}{(p-1)^3} \geqslant \prod_{n \geqslant 3} \frac{n-2}{n-1} \frac{n}{n-1} = \frac{1}{2},$$

so $\mathfrak{S}(N) \asymp 1$ for odd N, and it follows from the theorem that $r(N) > 0$ for N large enough. But this counts the representations of N as three prime powers, not just primes. Consider the set P'_N of proper prime powers less than N, i.e., $P'_N \coloneqq \{p^k \leqslant N : p \in P, k \geqslant 2\}$. Clearly each $p^k \in P'_N$ must have $p \leqslant N^{1/2}$, and the number of choices for k is $O(\log N)$, so $\#P'_N \ll \pi(N^{1/2}) \log N \ll N^{1/2}$ by PNT. Thus the contribution to $r(N)$ by representations $N = k_1 + k_2 + k_3$ with $k_1 \in P'_N$ is

$$\begin{aligned} \sum_{\substack{k_1+k_2+k_3=N \\ k_1 \in P'_N}} \Lambda(k_1)\,\Lambda(k_2)\,\Lambda(k_3) &\leqslant \log^3 N \sum_{\substack{k_1+k_2+k_3=N \\ k_1 \in P'_N}} 1 \\ &\ll (\log^3 N)\,(\#P'_N)\,N \\ &\ll N^{3/2} \log^3 N, \end{aligned}$$

and similarly for $k_2 \in P'_N$, $k_3 \in P'_N$. Thus by theorem 1.3,

$$\sum_{\substack{p_1+p_2+p_3=N \\ p_1, p_2, p_3 \in P}} \Lambda(k_1)\,\Lambda(k_2)\,\Lambda(k_3) = r(N) - O(N^{3/2} \log^3 N) \gg N^2.$$

Therefore we have obtained a lower-bound for the number of ways of writing large odd numbers as a sum of three primes.

Chapter 2

Hardy–Littlewood's Proof

In this chapter, we introduce Dirichlet characters and see how they help us study primes in progressions, then we discuss Hardy–Littlewood's conditional proof of theorem 1.3 on GRH. We mainly follow the arguments outlined in [16, pp. 18–22] for the proof, adding details for more clarity where necessary.

2.1 Dirichlet Characters, PNT in Progressions

We will first start by introducing Dirichlet characters, and subsequently use them to formulate the part of the proof which depends on GRH. I referred to [1] as well as [7] when compiling this section.

Let q be a positive whole number. A *Dirichlet character mod q* is a multiplicative function $\chi\colon (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}^\times$, i.e., we have $\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in (\mathbb{Z}/q\mathbb{Z})^\times$. Moreover, we extend Dirichlet characters to $\mathbb{Z}$ by periodicity mod q, i.e., $\chi(n+q) = \chi(n)$ for all n, and by setting $\chi(n) \coloneqq 0$ for all $n \in \mathbb{Z}$ not coprime to q. The Dirichlet character defined by $\chi(n) \coloneqq 1$ for all $n \in (\mathbb{Z}/q\mathbb{Z})^\times$ is denoted by χ_0 and called the *principal character*. (We still have $\chi_0(n) = 0$ if $(n, q) \neq 1$.)

From this definition one may deduce several facts. For instance, it is easily seen that $\chi(1) = 1$ and $\chi(0) = 0$ for all Dirichlet characters (unless $q = 1$ and $\chi = \chi_0$, in which case $\chi_0(n) = 1$ for all $n \in \mathbb{Z}$, this is called the *trivial character*). Additionally, by Euler's theorem, we see that $\chi(n)$ is always some $\varphi(q)$th of unity,[1] i.e., $\chi(n) = e(m/\varphi(q))$ for some m. Thus, $\chi(n)$ is some point on the unit circle for $(n, q) = 1$. There are $\varphi(q)$ distinct Dirichlet characters mod q. Dirichlet characters form an orthonormal basis for the set of functions $(\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$. In particular, they satisfy the orthogonality relation

$$\frac{1}{\varphi(q)} \sum_{\chi \bmod q} \chi(n)\,\bar{\chi}(a) = \mathbb{1}_{n \equiv a \bmod q}. \tag{2.1}$$

[1] *Euler's theorem.* For all $a \in \mathbb{Z}$, $a^{\varphi(q)} \equiv 1 \bmod q$.

For any Dirichlet character χ, the Dirichlet L-function is defined for $\Re(s) > 1$ by

$$L(s,\chi) := \sum_{n=0}^{\infty} \frac{\chi(n)}{n^s},$$

and can be extended to a meromorphic function on all of $\mathbb{C}$ by analytic continuation. The Generalised Riemann Hypothesis (GRH) asserts that for any Dirichlet character χ and $s \in \mathbb{C}$ with $L(s,\chi) = 0$, if s is not a negative real number, then $\Re(s) = {}^1\!/_2$. The Riemann Hypothesis (RH) is equivalent to the GRH in the case where χ is the trivial character, in which case, $L(s,\chi_0)$ equals the so-called Riemann zeta function, $\zeta(s)$.

The *Generalised Chebychev* Ψ *function* is defined by

$$\Psi(x,\chi) := \sum_{n \leqslant x} \Lambda(n)\,\chi(n).$$

This allows us to state *Dirichlet's theorem*, which is an analogue to the prime number theorem ($\Psi(x) = \Psi(x,\chi_0) \sim x$ with $q = 1$) for primes in arithmetic progressions. Dirichlet's theorem asserts that

$$\Psi(x,\chi) = \begin{cases} x + o(x) & \text{if } \chi = \chi_0 \\ \quad o(x) & \text{otherwise,} \end{cases}$$

where the implicit constants depend on q. Unfortunately the error terms here are not good enough for us to prove Vinogradov's theorem. It is at this stage (and only this stage) that we invoke GRH, to get a better bound to use in the proof.

Theorem 2.1 (PNT in arithmetic progressions, on GRH)**.** *Let χ be a Dirichlet character mod q, and assume GRH for $L(s,\chi)$. Then*

$$\Psi(x,\chi) = \begin{cases} x + O(\sqrt{x}\log^2 x + \log x \log q) & \textit{if } \chi = \chi_0 \\ \qquad O(\sqrt{x}\log^2 qx) & \textit{otherwise.} \end{cases}$$

For a derivation of this consequence of GRH, see, for example, the end of [1, §20].

Finally, consider $\Psi(x;q,a)$, defined by

$$\Psi(x;q,a) := \sum_{\substack{n \leqslant x \\ n \equiv a \bmod q}} \Lambda(n),$$

a variant of the Ψ function that sums over $n \leqslant x$ congruent to a mod q. It follows from (2.1) that the relation between $\Psi(x;q,a)$ and $\Psi(x,\chi)$ is

$$\Psi(x;q,a) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \bar{\chi}(a)\,\Psi(x,\chi),$$

and so theorem 2.1 gives, on GRH, that

$$\Psi(x;q,a) = \frac{x}{\varphi(q)} + O(\sqrt{x}\log^2 qx). \tag{2.2}$$

2.2 Gauß and Ramanujan Sums

In this section, we discuss a pair of sums which we will encounter in the proof of Vinogradov's theorem, and establish properties about them which will prove useful. The proofs of these properties are not given in [16], we provide them here.

Let χ be a Dirichlet character modulo q, and q be a positive integer. Then the *Gauß sum* $\tau(\chi)$ is the sum defined by

$$\tau(\chi) := \sum_{b \bmod q} \chi(b)\, e(b/q).$$

We will make use of the following two facts on Gauß sums.

Proposition 2.2. *Let χ be a Dirichlet character modulo q. Then*

(i) if $\chi = \chi_0$, then $\tau(\chi) = \mu(q)$,

(ii) $|\tau(\chi)| \leqslant \sqrt{q}$.

Proof. We prove (i), and omit the proof of (ii) for brevity. See [1, §9] for a proof of (ii).

$$\tau(\chi_0) = \sum_{b=1}^{q} \chi_0(b)\, e(b/q) = \sum_{\substack{b=1\\(b,q)=1}}^{q} e(b/q) = \sum_{d=1}^{q} \mu(d) \sum_{a=1}^{q/d} e(ad/q) = \mu(q)$$

since $\sum_{d|n} \mu(d) = 1$ only if $n = 1$, and by summing roots of unity. □

Next, we have the *Ramanujan sum* $c_q(n)$, defined by

$$c_q(n) := \sum_{\substack{b \bmod q\\(b,q)=1}} e(bn/q).$$

We will make use of the following.

Proposition 2.3. *Let $c_q(n)$ be the Ramanujan sum defined above. Then*

(i) $c_q(n)$ is multiplicative in q, i.e., $c_{qr}(n) = c_q(n)\, c_r(n)$,

(ii) for p prime,

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

Proof. For (i), observe that

$$\mathbb{1}(q) * c_q(n) = \sum_{d|q} c_{q/d}(n) = \sum_{d|q} \sum_{\substack{b \bmod q\\(b,q)=d}} e(bn/q) = \sum_{b \bmod q} e(bn/q) = (\mathbb{1}_{n\equiv 0 \bmod q})\, q$$

by summing roots of unity. Convolving with μ both sides, we get that

$$c_q(n) = \sum_{d\mid(q,n)} d\,\mu(q/d) \tag{2.3}$$

by proposition 1.2. We conclude that $c_q(n)$ is always an integer, and that it is multiplicative in q, since it depends only on its divisors. Now for (ii), observe that if $p \mid n$, then by (2.3), $c_p(n) = \sum_{d\mid p} d\,\mu(p/d) = \mu(p) + p\,\mu(1) = p-1 = \varphi(p)$, whereas if $p \nmid n$, $c_p(n) = \sum_{d\mid 1} d\,\mu(p/d) = \mu(p) = -1$. □

2.3 Bounds on $S(N,\alpha)$

In this section, we will make use of theorem 2.1 to obtain a bound for the sum $S(N,\alpha)$. Notice that since $|e(n\alpha)| = 1$, we have the obvious bound

$$S(N,\alpha) \leqslant \Psi(N) \ll N \tag{2.4}$$

by the PNT. Now in the introduction, we mentioned that the integrand in $r(N)$ (1.1) turns out to be large when α is close to rational points with small denominators. Since this will be important in proving theorem 1.3, we will need better estimates for $S(N,\alpha)$ which incorporate this information about α. Let us start by giving a bound for $S(x,\alpha)$ for rational values of α.

Proposition 2.4. *Let a/q be a rational number with a coprime to q. Then on GRH, we have*

$$S(x, a/q) = \frac{\mu(q)}{\varphi(q)}\,x + O(\sqrt{qx}\log^2 qx). \tag{2.5}$$

Proof. Notice that terms with $(n,q) > 1$ barely contribute to the sum, in particular,

$$S(x, a/q) = \sum_{\substack{n\leqslant x\\(n,q)=1}} \Lambda(n)\,e(an/q) + O(\log x \log q), \tag{2.6}$$

since q has $\ll \log q$ prime factors, and $p^k \mid q$ implies that $\Lambda(p)+\cdots+\Lambda(p^k) = \log p^k \ll \log x$. Now at this stage, we could split n into progressions mod q and invoke (2.2), but this introduces an error of $O(q\sqrt{x}\log^2 qx)$ which is not good enough to get (2.5). Instead, we proceed as follows. For $(an,q) = 1$, by orthogonality of Dirichlet characters, we have

$$\begin{aligned} e(an/q) &= \sum_{b \bmod q} e(b/q)\,\mathbb{1}_{b\equiv an \bmod q} \\ &= \frac{1}{\varphi(q)} \sum_{b \bmod q} \sum_{\chi \bmod q} \chi(b)\,\bar{\chi}(an)\,e(b/q) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \tau(\chi)\,\bar{\chi}(an). \end{aligned}$$

Using this in (2.6), we get

$$S(x, {}^{a}\!/_{q}) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \tau(\chi)\, \bar{\chi}(a)\, \Psi(x, \bar{\chi}) + O(\log x \log q).$$

By theorem 2.1 and proposition 2.2(ii), we see that for all non-principal characters, the contribution to the above sum is $\ll \sqrt{qx}\log^2 qx$, and for the principal character χ_0, the contribution is

$$\frac{1}{\varphi(q)}\tau(\chi_0)(x + O(\sqrt{x}\log^2 qx)) = \frac{\mu(q)}{\varphi(q)}(x + O(\sqrt{x}\log^2 qx)),$$

by proposition 2.2(i). □

Next, we use summation by parts on the result of proposition 2.4 to obtain a bound for α "close to" a rational ${}^{a}\!/_{q}$.

Corollary 2.5. *Let $\alpha = {}^{a}\!/_{q} + \beta$, where a is coprime to q and $|\beta| < {}^{1}\!/_{2}$. Then on GRH, we have*

$$S(N,\alpha) = \frac{\mu(q)}{\varphi(q)} \int_0^N e(\beta x)\, dx + O((1 + |\beta| N)\sqrt{qN}\log^2 qN).$$

Proof. We have

$$\begin{aligned} S(N,\alpha) &= \sum_{n \leqslant N} \Lambda(n)\, e({}^{an}\!/_{q})\, e(n\beta) \\ &= \int_0^N e(x\beta)\, d(S(x, {}^{a}\!/_{q})) \\ &= \int_0^N e(x\beta)\, d\bigg(\frac{\mu(q)}{\varphi(q)}\, x + \epsilon(x, {}^{a}\!/_{q})\bigg) \end{aligned}$$

by proposition 2.4, where $\epsilon(x, {}^{a}\!/_{q})$ denotes the error term. The first term of the integral gives the first term of the corollary. For the second term, integration by parts yields

$$\begin{aligned} \int_0^N e(x\beta)\, d(\epsilon(x, {}^{a}\!/_{q})) &= e(N\beta)\, \epsilon(N, {}^{a}\!/_{q}) - 2\pi i\beta \int_0^N e(x\beta)\, \epsilon(x, {}^{a}\!/_{q})\, dx \\ &\ll O(\sqrt{qN}\log^2 N) - 2\pi i\beta\, N\, O(\sqrt{qN}\log^2 qN) \\ &\ll O((1 + |\beta| N)\sqrt{qN}\log^2 qN), \end{aligned}$$

as required. □

Next, let us make precise the notion of "close to rational points with small denominators", which we alluded to in the introduction.

Definition 2.6. We shall say that a rational number ${}^a\!/\!{}_q$ with $(a,q)=1$ *approximates* the number $\alpha \in \mathbb{R}$ if

$$|\alpha - {}^a\!/\!{}_q| \leqslant {}^1\!/\!{}_{qQ},$$

where $Q \geqslant q$ is suitably large (we define "suitably large" in a bit, in terms of N).

Such rational approximations of real numbers are known as *Diophantine approximations.* It turns out to be an easy consequence of the pigeonhole principle that we can always find one for $Q \geqslant q$, see proposition A.1 in appendix A.

Now, notice that by corollary 2.5, if ${}^a\!/\!{}_q$ approximates α, then

$$\begin{aligned} S(N,\alpha) &\lll \frac{N}{\varphi(q)} + \Big(1+\frac{N}{qQ}\Big)\sqrt{qN}\log^2 N \\ &\lll \frac{N}{\varphi(q)} + \Big(\sqrt{QN}+\frac{N^{3/2}}{Q}\Big)\log^2 N, \end{aligned}$$

and by choosing $Q = N^{2/3}$, we get that

$$S(N,\alpha) \lll \frac{N}{\varphi(q)} + N^{5/6+\epsilon} \tag{2.7}$$

for any $\epsilon > 0$.

2.4 Conditional Proof of Vinogradov's Theorem

Motivated by (2.7), following Hardy–Littlewood, we shall say, for the remainder of the chapter, that α is close to a rational number with small denominator if there exists ${}^a\!/\!{}_q$ which approximates α, with $Q = N^{2/3}$ and $q \leqslant \log^{10} N$ (the latter condition is for "small denominators", the power 10 is arbitrary). If α is such a number, we say α *lies on a major arc*, and denote the set of such points by $\mathfrak{M}$. Otherwise, α lies on a *minor arc*, and the set of such points is denoted $\mathfrak{m}$. In other words, we have

$$\mathfrak{M} = \bigcup_{q\leqslant \log^{10} N}\ \bigcup_{\substack{1\leqslant a\leqslant q\\(a,q)=1}} \big({}^a\!/\!{}_q - {}^1\!/\!{}_{qQ}, {}^a\!/\!{}_q + {}^1\!/\!{}_{qQ}\big) \qquad \text{and} \qquad \mathfrak{m} = (\mathbb{R}/\mathbb{Z}) \smallsetminus \mathfrak{M}.$$

This idea of partitioning the circle $\mathbb{R}/\mathbb{Z}$ into $\mathfrak{M} \cup \mathfrak{m}$ *is* Hardy–Littlewood's circle method. The bulk of the contribution to $r(N)$ (1.1) (i.e., the main term in (1.2)) will come from integrating over the major arcs, and then we can bound the contribution on the minor arcs close to zero. Note that both $\mathfrak{M}$ and $\mathfrak{m}$ are measurable, and we are therefore justified in splitting the integral. Moreover, for N large enough, we have that $\mathfrak{M}$ is a union of *disjoint* intervals: indeed, if $q, q' \leqslant \log^{10} N$, and ${}^a\!/\!{}_q, {}^b\!/\!{}_{q'} \in \mathfrak{M}$ with ${}^a\!/\!{}_q \neq {}^b\!/\!{}_{q'}$, then

$$\big|{}^a\!/\!{}_q - {}^b\!/\!{}_{q'}\big| \geqslant {}^1\!/\!{}_{qq'} > {}^1\!/\!{}_{\log^{20} N} > {}^2\!/\!{}_{N^{2/3}}$$

for N sufficiently large, so ${}^{a}\!/_{q}$, ${}^{b}\!/_{q'}$ cannot lie in the same arc.

A historical note: the initial idea behind the circle method originates in the joint work of G. H. Hardy and S. Ramanujan from 1916, when they were studying asymptotics of integer partitions. Hardy and J. E. Littlewood later developed the method in a series of papers on Waring's problem.[5, 18]

Now, let us start by bounding the minor arc contribution to (1.1).

Proposition 2.7 (The minor arc contribution)**.** *On GRH,*

$$\Big|\int_{\mathfrak{m}} S(N,\alpha)^3\, e(-N\alpha)\, d\alpha\Big| \ll \frac{N^2}{\log^8 N}.$$

Proof. Since $\varphi(q) \gg q/\log\log q \gg q/\log q$ (see [6, ch. 18, thm. 328]), for $q > \log^{10} N$, we have $\varphi(q) \geqslant \log^9 N$, therefore by (2.7), we have $S(N,\alpha) \ll N/\log^9 N$, so

$$\begin{aligned}\Big|\int_{\mathfrak{m}} S(N,\alpha)^3\, e(-N\alpha)\, d\alpha\Big| &\leqslant \int_{\mathfrak{m}} |S(N,\alpha)|^3\, d\alpha \\ &\ll \frac{N}{\log^9 N}\int_0^1 |S(N,\alpha)|^2\, d\alpha = \frac{N}{\log^9 N}\sum_{n\leqslant N}\Lambda(n)^2 \ll \frac{N^2}{\log^8 N},\end{aligned}$$

since $\sum_{n\leqslant N}\Lambda(n)^2 \ll N\log N$. □

Notice that although we get some saving using Parseval's identity, the step of extracting the $N/\log^9 N$ term is quite costly, yet we are still able to prove the result. This attests to the fact that (2.7) is an excellent bound, resting on the strength of GRH. Next, we determine the major arc contribution, which we expect to dominate the contribution of $O(N^2/\log^8 N)$ from the minor arcs. The proof here is much more detailed than the one given in [16].

Proposition 2.8 (The major arc contribution)**.** *On GRH,*

$$\int_{\mathfrak{M}} S(N,\alpha)^3\, e(-N\alpha)\, d\alpha \sim \tfrac{1}{2}\,\mathfrak{S}(N)\, N^2.$$

Proof. We have $|\alpha - {}^{a}\!/_{q}| \leqslant {}^{1}\!/_{qQ}$ with $q \leqslant \log^{10} N$, so the major arc contribution equals

$$\sum_{q\leqslant \log^{10} N}\ \sum_{\substack{1\leqslant a\leqslant q\\ (a,q)=1}} \int_{-1/qQ}^{1/qQ} S(N, {}^{a}\!/_{q}+\beta)^3\, e(-N({}^{a}\!/_{q}+\beta))\, d\beta, \tag{2.8}$$

since intervals for different rational numbers are disjoint. Now, observe that

$$\int_0^N e(\beta x)\, dx = \frac{e(N\beta)-1}{2\pi i\beta} \ll \frac{1}{\|\beta\|},$$

and also $\int_0^N e(\beta x)\,dx \leqslant \int_0^N |e(\beta x)|\,dx = N$, so we have $\int_0^N e(\beta x)\,dx \ll \min\{N, {}^{1}\!/\!{}_{\|\beta\|}\}$. Thus by corollary 2.5, $S(N, {}^{a}\!/\!{}_{q} + \beta)^3$ equals

$$\begin{aligned}\frac{\mu(q)^3}{\varphi(q)^3}\Big(\int_0^N e(\beta x)\,dx\Big)^3 &+ O\Big(\frac{1}{\varphi(q)^2}\min\Big\{N^2, \frac{1}{\|\beta\|^2}\Big\}(1+\|\beta\|N)\sqrt{qN}\log^2 N\Big)\\ &+ O\Big(\frac{1}{\varphi(q)}\min\Big\{N, \frac{1}{\|\beta\|}\Big\}(1+\|\beta\|N)^2 qN\log^4 N\Big)\\ &+ O((1+\|\beta\|N)^3(qN)^{3/2}\log^6 N).\end{aligned}$$

Now notice that $\min\{N^2, \frac{1}{\|\beta\|^2}\}(1+\|\beta\|N) \ll N^2$ and $\min\{N, \frac{1}{\|\beta\|}\}(1+\|\beta\|N) \ll N$, and also $\|\beta\| \leqslant {}^{1}\!/\!{}_{qQ} = {}^{1}\!/\!{}_{qN^{2/3}}$. Therefore, since $q \leqslant \log^{10} N \ll N^{\epsilon}$ for all $\epsilon > 0$, the error terms become

$$O(N^2\sqrt{N^{1+\epsilon}}\,N^{\epsilon}) + O(N\,(1+N^{1/3})\,N^{1+\epsilon}) + O((1+N^{1/3})^3 N^{3/2+\epsilon}) = O(N^{5/2+\epsilon}).$$

Thus, we can write (2.8) as

$$\begin{aligned}&\sum_{q\leqslant\log^{10}N}\ \sum_{\substack{1\leqslant a\leqslant q\\(a,q)=1}}\int_{-1/qQ}^{1/qQ}\Big(\frac{\mu(q)^3}{\varphi(q)^3}\Big(\int_0^N e(\beta x)\,dx\Big)^3 + O(N^{5/2+\epsilon})\Big)\,e(-N({}^{a}\!/\!{}_{q}+\beta))\,d\beta\\ &= \sum_{q\leqslant\log^{10}N}\frac{\mu(q)^3}{\varphi(q)^3}\Big(\sum_{\substack{1\leqslant a\leqslant q\\(a,q)=1}} e(\,{-N a}/q)\Big)\Big(\int_{-1/qQ}^{1/qQ}\Big(\int_0^N e(\beta x)\,dx\Big)^3 e(-N\beta)\,d\beta\Big) + O(N^{11/6+\epsilon}),\end{aligned}\tag{2.9}$$

since $\int_{-1/qQ}^{1/qQ} O(N^{5/2+\epsilon})\,d\beta \ll N^{5/2+\epsilon}N^{-2/3} \ll N^{11/6+\epsilon}$. Let us simplify the integral first. We make the substitutions $x = Ny$ and $N\beta = \xi$, and the integral becomes

$$\begin{aligned}&N^2\int_{-N/qQ}^{N/qQ}\Big(\int_0^1 e(y\xi)\,dy\Big)^3 e(-\xi)\,d\xi\\ =&N^2\Big(\int_{-\infty}^{\infty}\Big(\int_0^1 e(y\xi)\,dy\Big)^3 e(-\xi)\,d\xi + O\Big(\frac{1}{({}^{N}\!/\!{}_{qQ})^2}\Big)\Big),\end{aligned}\tag{2.10}$$

since the integrand is $\ll {}^{1}\!/\!{}_{\xi^3}$. Notice that the integral is simply

$$\widecheck{(\widehat{\mathbb{1}}_{[0,1]})^3}(1),$$

which, by the convolution theorem, equals $(\mathbb{1}_{[0,1]}\star\mathbb{1}_{[0,1]}\star\mathbb{1}_{[0,1]})(1)$. We have $\mathbb{1}_{[0,1]}\star\mathbb{1}_{[0,1]} = \triangle_{[0,2]}$, where $\triangle_{[0,2]}$ denotes a symmetric triangle of height 1 on the interval $[0,2]$, i.e., $\triangle_{[0,2]}(x) = (1-|1-x|)\,\mathbb{1}_{[0,2]}(x)$. The second convolution evaluated at 1 is

$$(\mathbb{1}_{[0,1]}\star\triangle_{[0,2]})(1) = \int_{-\infty}^{\infty}\triangle_{[0,2]}(\tau)\,\mathbb{1}_{[0,1]}(1-\tau)\,d\tau = \int_0^1 \tau\,d\tau = \frac{1}{2},$$

and so (2.10) equals ${}^{N^2}\!/2 + O((qQ)^2) = {}^{N^2}\!/2 + O(N^{4/3+\epsilon})$, which means we can write our major arc contribution in (2.9) as

$$\frac{N^2}{2} \sum_{q \leqslant \log^{10} N} \frac{\mu(q)^3}{\varphi(q)^3} \Bigg(\sum_{\substack{1 \leqslant a \leqslant q \\ (a,q)=1}} e\big({}^{-Na}\!/q \big) \Bigg) + O(N^{11/6+\epsilon}).$$

Now we can recognise the inner sum as $c_q(-N)$. We are interested in it for square-free q, since otherwise the factor of $\mu(q)^3$ annihilates the summand. Since both $c_q(n)$ and φ are multiplicative, we have that $c_q(-N) \leqslant \varphi(q)$ by proposition 2.3. Thus we can extend the outer sum to infinity, with error term at most $\sum_{q > \log^{10} N} {}^{\mu(q)^3}\!/\varphi(q)^2 \ll {}^{1}\!/\log^{10} N$, and so our major arc contribution is asymptotic to

$$\begin{aligned} \frac{N^2}{2} \sum_{q=1}^{\infty} \frac{\mu(q)^3}{\varphi(q)^3} c_q(-N) &= \frac{N^2}{2} \prod_p \Big(1 - \frac{c_p(-N)}{\varphi(p)^3}\Big) \\ &= \frac{N^2}{2} \prod_{p \mid N} \Big(1 - \frac{1}{(p-1)^2}\Big) \prod_{p \nmid N} \Big(1 + \frac{1}{(p-1)^3}\Big), \end{aligned}$$

which completes the proof. □

Therefore we conclude that

$$\begin{aligned} r(N) &= \int_{\mathfrak{M}} S(N,\alpha)^3 \, e(-N\alpha) \, d\alpha + \int_{\mathfrak{m}} S(N,\alpha)^3 \, e(-N\alpha) \, d\alpha \\ &= \tfrac{1}{2} \, \mathfrak{S}(N) \, N^2 + O(N^2 / \log^8 N), \end{aligned}$$

which establishes theorem 1.3 on GRH.

Remark 2.9. Notice that in our proof of the major arc estimate, we made use of corollary 2.5, which depends on GRH (because of the dependency on theorem 2.1). However instead of theorem 2.1, we can adapt the argument to use the bound

$$\Psi(x, \chi) \ll x \exp(-c_B \sqrt{\log x})$$

for non-principal χ (where c_B is a constant[2]), which is a consequence of Siegel's theorem and does not depend on GRH (see [1, §22] for a derivation). Indeed, in [1, §26], the bound

$$\int_{\mathfrak{M}} S(\alpha, N)^3 \, e(-N\alpha) \, d\alpha = \tfrac{1}{2} \mathfrak{S}(N) \, N^2 + O(N^2 / \log^{B-1} N)$$

is derived this way. Thus the real work in making the proof unconditional will be in the treatment of the minor arcs.

[2] c_B depends on B, where B is such that $q \leqslant \log^B N$, in this chapter we have $B = 10$.

Chapter 3

Vinogradov's Proof

The key to Vinogradov's proof is in the careful treatment of exponential sums to bound the contribution of the minor arcs. In particular, we will be making use of the following bound.

Theorem 3.1 (Vinogradov's Bound)**.** *Let α be a real number, let ${}^a/_q$ be a rational number such that $|\alpha - {}^a/_q| \leqslant {}^1/_{q^2}$ where $(a,q) = 1$, and suppose $N \geqslant q$. Then*

$$S(N,\alpha) \ll \left(\frac{N}{\sqrt{q}} + N^{4/5} + \sqrt{Nq}\right) \log^4 N.$$

In view of this theorem, we shall adjust slightly our definitions of major and minor arcs. We shall say that the rational number ${}^a/_q$ *approximates* α if $|\alpha - {}^a/_q| \leqslant {}^1/_{q^2}$. (In other words, we are taking $Q = q$ rather than $Q = N^{2/3}$ as in the last chapter.) We shall say that α is close to a rational number if there exists ${}^a/_q$ which approximates α, with $q \leqslant \log^B N$, where $B > 0$ will be specified later. If α is such a number, we say α *lies on a major arc*, and denote the set of such points by $\mathfrak{M}$. Otherwise, α lies on a *minor arc*, and the set of such points is denoted $\mathfrak{m} = (\mathbb{R}/\mathbb{Z}) \setminus \mathfrak{M}$.

3.1 Vaughan's Identity

In [19], Vinogradov introduces a technique for estimating sums of the form $\sum_{p \leqslant N} f(p)$ where f is periodic, e.g., $f(p) = e(p\alpha)$. Following [20, ch. IX], set $P := \prod_{p \leqslant \sqrt{N}} p$. Then the sieve of Eratosthenes asserts that for $1 \leqslant n \leqslant N$, $(n,P) = 1$ if and only if $n = 1$ or n is a prime in the range $\sqrt{N} \leqslant n \leqslant N$. Thus

$$f(1) + \sum_{\sqrt{N} \leqslant p \leqslant N} f(p) = \sum_{\substack{n \leqslant N \\ (n,P)=1}} f(n) = \sum_{\substack{d|P \\ d \leqslant N}} \mu(d) \sum_{y \leqslant N/d} f(yd),$$

and we are led to bound sums of the form $\sum_{y\leqslant N/d} f(yd)$. We would like to show that these sums are small, but we cannot hope to get much cancellation when d is nearly as large as N, so Vinogradov rearranges the terms arising from $\delta N \leqslant d \leqslant N$ with $d \mid P$, but this becomes rather complicated.

In [18, ch. 3], Vaughan simplifies Vinogradov's method substantially. He starts by giving the following identity.

Proposition 3.2 (Vaughan, [18]). *For a function f, let $f_{\leqslant A}$ and $f_{>A}$ denote the functions $f_{\leqslant A}(x) := f(x)\,\mathbb{1}_{x\leqslant A}$, and $f_{>A}(x) := f(x)\,\mathbb{1}_{x>A}$ respectively. In this notation, we have*

$$\Lambda = \mu_{\leqslant X} * \log - \mu_{\leqslant X} * \Lambda_{\leqslant Y} * \mathbb{1} + \mu_{>X} * \Lambda_{>Y} * \mathbb{1} + \Lambda_{\leqslant Y}. \tag{3.1}$$

In [18], Vaughan obtains this as a consequence of another combinatorial identity which he proves directly. To make things simpler here, we have stated the result in terms of Dirichlet convolutions instead, so that we may give an easy proof by using known properties of convolutions.

Proof. The result is essentially a consequence of the fact that $\mathbb{1} * \mu = \mathbb{1}_{=1}$ (proposition 1.2). Indeed,

$$\begin{aligned}\Lambda &= \Lambda_{>Y} + \Lambda_{\leqslant Y}\\ &= \Lambda_{>Y} * (\mathbb{1} * \mu) + \Lambda_{\leqslant Y}\\ &= \Lambda_{>Y} * \mathbb{1} * (\mu_{\leqslant X} + \mu_{>X}) + \Lambda_{\leqslant Y}\\ &= (\Lambda - \Lambda_{\leqslant Y}) * \mathbb{1} * \mu_{\leqslant X} + \Lambda_{>Y} * \mathbb{1} * \mu_{>X} + \Lambda_{\leqslant Y}\\ &= (\Lambda * \mathbb{1}) * \mu_{\leqslant X} - \Lambda_{\leqslant Y} * \mathbb{1} * \mu_{\leqslant X} + \Lambda_{>Y} * \mathbb{1} * \mu_{>X} + \Lambda_{\leqslant Y}\\ &= \mu_{\leqslant X} * \log - \mu_{\leqslant X} * \Lambda_{\leqslant Y} * \mathbb{1} + \mu_{>X} * \Lambda_{>Y} * \mathbb{1} + \Lambda_{\leqslant Y},\end{aligned}$$

as required. □

The motivation behind this identity is to express Λ in terms of sums of the form $\sum_{x\mid N, x\leqslant X} a_x$ ranging over small numbers (called type I sums), and sums of the form $\sum_{xy=N, x>X, y>Y} a_x b_y$ which range over large numbers (called type II sums). Sums of type I can be dealt with by bounding the magnitude of the inner sum, whereas type II sums can be dealt with using bilinear methods (usually Cauchy–Schwarz[18]), as we shall see in the proof of Vinogradov's bound.

Applying the identity to $S(N, \alpha)$, by proposition A.2 (appendix A), we get that

$$S(N,\alpha) = \sum_{n\leqslant N} \Bigg(\sum_{\substack{xy=n\\ x\leqslant X}} \mu(x)\, \log y - \sum_{\substack{xyz=n\\ x\leqslant X, y\leqslant Y}} \mu(x)\, \Lambda(y) + \sum_{\substack{xyz=n\\ x>X, y>Y}} \mu(x)\, \Lambda(y) + \Lambda_{\leqslant Y}(n) \Bigg)\, e(\alpha n)$$

$$\begin{aligned}
&= \sum_{x\leqslant X}\sum_{y\leqslant N/x} \mu(x)\,(\log y)\,e(\alpha xy) - \sum_{x\leqslant XY}\sum_{y\leqslant N/x}\Big(\sum_{\substack{dz=x\\ d\leqslant X, z\leqslant Y}} \mu(d)\,\Lambda(z)\Big)\,e(\alpha xy)\\
&\qquad\qquad + \sum_{x>X}\sum_{Y<y\leqslant N/x}\Big(\sum_{\substack{d\mid x\\ d\leqslant Z}}\mu(d)\Big)\,\Lambda(y)\,e(\alpha xy) + S(Y,\alpha)\\
&=: S_{\mathrm{I},1} + S_{\mathrm{I},2} + S_{\mathrm{II}} + O(Y). \qquad (3.2)
\end{aligned}$$

3.2 Proving Vinogradov's Bound

Before we can prove theorem 3.1, we need the following technical result. The result is stated this way in [18], but we give a proof adapted from [1]. We gloss over the detail that the numbers ra/q are "spread out" mod 1, a proof of this can be seen in [2].

Lemma 3.3. *Suppose that A, B, α are real numbers with $A, B \geqslant 1$, and that $|\alpha - a/q| \leqslant 1/q^2$ with $(a, q) = 1$. Then*

$$\sum_{x\leqslant A} \min\Big\{\frac{AB}{x}, \frac{1}{\|\alpha x\|}\Big\} \ll AB\Big(\frac{1}{q} + \frac{1}{B} + \frac{q}{AB}\Big)\log(2Aq).$$

Proof. Let S denote the sum. Writing $x = qj + r$, we have

$$S \leqslant \sum_{0\leqslant j\leqslant A/q}\sum_{r\leqslant q} \min\Big\{\frac{AB}{qj+r}, \frac{1}{\|\alpha(qj+r)\|}\Big\}.$$

Now let $\beta := \alpha - a/q$ (notice $|\beta| \leqslant 1/q^2$). Then $\|\alpha(qj+r)\| = \|ra/q + jq\beta + r\beta\|$, and terms with $j = 0$ and $1 \leqslant r \leqslant q/2$ have $|r\beta| \leqslant 1/2q$, which implies that their contribution to the sum is

$$\ll \sum_{r\leqslant q/2} \frac{1}{\|ra/q\| - 1/2q} \leqslant \sum_{\substack{d \bmod q\\ d\not\equiv 0}} \frac{1}{\|d/a\| - 1/2q} \leqslant 2\sum_{1\leqslant d\leqslant q/2} \frac{1}{d/q - 1/2q} \ll q\sum_{n\leqslant q-1}\frac{1}{n} \ll q\log 2q.$$

Now the remaining terms contribute

$$\ll \sum_{0\leqslant j\leqslant A/q}\sum_{r\leqslant q} \min\Big\{\frac{AB}{q(j+1)}, \frac{1}{\|ra/q + jq\beta + r\beta\|}\Big\},$$

and for fixed j, the values of $ra/q + jq\beta + r\beta$ are "spread out" modulo 1 for $r = 1, \ldots, q$, thus we get a contribution

$$\sum_{0\leqslant j\leqslant A/q}\Big(\frac{AB}{q(j+1)} + \sum_{k\leqslant q/2}\frac{q}{k}\Big) \ll \frac{AB}{q}\log(2A) + \Big(\frac{A}{q}+1\Big)q\log(2q),$$

which completes the proof. □

Now we can give the proof of Vinogradov's bound. We follow [18] here again, but give much more detail, referring to [7, ch. 2].

Proof of Theorem 3.1. Recall that by (3.2), we may write $S(N,\alpha) = S_{\mathrm{I},1} + S_{\mathrm{I},2} + S_{\mathrm{II}} + O(Y)$. We start by bounding the two type I sums, $S_{\mathrm{I},1}$ and $S_{\mathrm{I},2}$. Writing $\log y = \int_1^y \frac{dt}{t}$, $S_{\mathrm{I},1}$ equals

$$\sum_{x\leqslant X}\sum_{y\leqslant N/x}\mu(x)\int_1^y \frac{e(\alpha xy)}{t}\,dt = \sum_{x\leqslant X}\mu(x)\int_1^{N/x}\sum_{t\leqslant y\leqslant N/x}\frac{e(\alpha xy)}{t}\,dt,$$

and because

$$\begin{aligned}\Big|\sum_{t\leqslant y\leqslant N/x} e(\alpha xy)\Big| &= \Big|\frac{e(\alpha x(1+\lfloor N/x\rfloor)) - e(\alpha x\lceil t\rceil)}{e(\alpha x)-1}\Big| \\ &\leqslant \frac{2}{|e(\alpha x/2) - e(-\alpha x/2)|} = \frac{1}{|\sin(\pi\alpha x)|} \leqslant \frac{1}{2\|\alpha x\|},\end{aligned}$$

together with the trivial bound $|\sum_{t\leqslant y\leqslant N/x} e(\alpha xy)| \leqslant N/x$, we get that

$$S_{\mathrm{I},1} \ll \log N \sum_{x\leqslant X}\min\Big\{\frac{N}{x}, \frac{1}{\|\alpha x\|}\Big\}.$$

Also, $\sum_{dz=x}\mu(d)\,\Lambda(z) \ll \log x$, so similarly we get that $S_{\mathrm{I},2} \ll \sum_{x\leqslant XY}\int_1^x \sum_{t\leqslant y\leqslant N/x}\frac{e(\alpha xy)}{t}\,dt$, and using the same bound for $\sum e(\alpha xy)$, that $S_{\mathrm{I},2} \ll \log N \sum_{x\leqslant XY}\min\{N/x, 1/\|\alpha x\|\}$. Putting $X = N^{2/5} = Y$ and applying lemma 3.3 with $A = XY$ and $B = N^{1/5}$, we get

$$S_{\mathrm{I},1}, S_{\mathrm{I},2} \ll (\log N)\,N\Big(\frac{1}{q} + \frac{1}{N^{1/5}} + \frac{q}{N}\Big)\log N = (\log^2 N)\Big(\frac{N}{q} + N^{4/5} + q\Big).$$

Next we treat the type II sum S_{II}, which gives the main term. Let k be the integer such that $2^k N^{4/5} \leqslant N \leqslant 2^{k+1}N^{4/5}$, and set $\mathscr{A} = \{2^\ell N^{2/5} : \ell = 1,\dots,k\}$. Then we can split S_{II} over intervals $[M, 2M]$ for $M \in \mathscr{A}$:

$$S_{\mathrm{II}} = \sum_{M\in\mathscr{A}} S(M)$$

where

$$S(M) = \sum_{M<x\leqslant 2M}\sum_{X<y\leqslant N/x}\Big(\sum_{\substack{d\mid x\\ d\leqslant X}}\mu(d)\Big)\Lambda(y)\,e(\alpha xy)$$

(taking $Z = X$). Now we apply Cauchy–Schwarz to extract the two inner sums

$$|S(M)|^2 = \Big|\sum_{M<x\leqslant 2M}\Big(\sum_{\substack{d\mid x\\ d\leqslant X}}\mu(d)\Big)\Big(\sum_{X<y\leqslant N/x}\Lambda(y)\,e(\alpha xy)\Big)\Big|^2$$

$$\begin{aligned}
&\leqslant \sum_{M<x\leqslant 2M} \Big| \sum_{\substack{d|x\\ d\leqslant X}} \mu(d) \Big|^2 \sum_{M<x\leqslant 2M} \Big| \sum_{X<y\leqslant N/x} \Lambda(y)\, e(\alpha xy) \Big|^2 \\
&\leqslant \sum_{x\leqslant 2M} \Big| \sum_{d|x} 1 \Big|^2 \sum_{M<x\leqslant 2M} \Big| \sum_{X<y\leqslant N/x} \Lambda(y)\, e(\alpha xy) \Big|^2 \\
&= \sum_{x\leqslant 2M} d(x)^2 \sum_{M<x\leqslant 2M} \Big| \sum_{X<y\leqslant N/x} \Lambda(y)\, e(\alpha xy) \Big|^2 .
\end{aligned}$$

$\sum_{x\leqslant A} d(x)^2 \ll A \log^3 2A$ (see proposition A.3 in appendix A), so we have

$$\begin{aligned}
|S(M)|^2 &\ll M \log^3 M \sum_{M<x\leqslant 2M} \sum_{X<y\leqslant N/x} \Lambda(y)\, e(\alpha xy) \sum_{X<z\leqslant N/x} \overline{\Lambda(z)\, e(\alpha xz)} \\
&\ll M \log^3 M \sum_{M<x\leqslant 2M} \sum_{X<y\leqslant N/x} \sum_{X<z\leqslant N/x} \Lambda(y)\, \Lambda(z)\, e(\alpha x(y-z)) \\
&\ll M \log^3 M \sum_{y\leqslant N/M} \Lambda(y) \sum_{z\leqslant N/M} \Lambda(z) \sum_{M<x\leqslant 2M} e(\alpha x(y-z)) \\
&\ll M\ \log^5 N \sum_{y\leqslant N/M} \sum_{z\leqslant N/M} \min\Big\{M, \frac{1}{\|\alpha(y-z)\|}\Big\},
\end{aligned}$$

then by lemma 3.3,

$$|S(M)|^2 \ll N\ \log^6 N \left(\frac{N}{q} + M + \frac{N}{M} + q\right),$$

and so

$$\begin{aligned}
S_{\text{II}} &\ll \sum_{M\in\mathscr{A}} \log^3 N \left(\frac{N}{\sqrt{q}} + \sqrt{NM} + \frac{N}{\sqrt{M}} + \sqrt{Nq}\right) \\
&\ll \left(\frac{N}{\sqrt{q}} + N^{4/5} + \sqrt{Nq}\right) \log^4 N,
\end{aligned}$$

as required. □

3.3 Deducing Vinogradov's theorem

In this section, we adapt the reasoning presented in the conclusion of [1, ch. 26].

Recall that in our conditional proof, the key to bounding the minor arcs was the bound in (2.7), which required GRH. Our new bound, although worse than (2.7), is still strong enough for us to prove theorem 1.3.

By Dirichlet's theorem on Diophantine approximation (proposition A.1), we can approximate α by a/q with $|\alpha - a/q| \leqslant 1/qQ \leqslant 1/q^2$, taking $Q = N \log^{-B} N$. We said that if

$q \leqslant \log^B N$, then $\alpha \in \mathfrak{M}$, otherwise, $\alpha \in \mathfrak{m}$. Thus for $\alpha \in \mathfrak{m}$, we have $\log^B N < q \leqslant N \log^{-B} N$, and so

$$S(N, \alpha) \ll N \log^{-B/2+4} N$$

by Vinogradov's bound (theorem 3.1). Proceeding similarly to the proof of proposition 2.7, we obtain that

$$\int_{\mathfrak{m}} S(N, \alpha)^3 \, e(-N\alpha) \, d\alpha \ll N^2 \log^{-B/2+5} N.$$

Now recall that in the last chapter, the major arcs consisted of reals which can be approximated with denominator $q \leqslant \log^{10} N$, where the integer 10 was arbitrary. But if we put $B = 10$ into the above, we get an error of N^2, which is not good enough for the proof! If we adjust the definition of minor arcs by taking any $B > 10$, then we can adapt the proof of proposition 2.8 to still get the same result, but with an error of $O(N^2/\log^B N)$ (rather than $O(N^2/\log^{10} N)$, which is what we got in the end). Thus the term $\frac{1}{2}N^2\mathfrak{S}(N)$ dominates the minor arc contribution, and the proof is complete.

Chapter 4

Helfgott's Proof

In [10], Helfgott studies the sum

$$S_\eta(x,\alpha) = \sum_n \Lambda(n)\, e(\alpha n)\, \eta(n/x), \tag{4.1}$$

where $\eta\colon \mathbb{R} \to \mathbb{R}$ is a smooth function which decays fast enough for convergence. (Usually $\eta(t) = 0$ for $t \geqslant 1$, so x here is playing a similar role to our usual upper bound N.) Using different weights η, Helfgott is able to get more control over explicit constants in his bounds in order to establish that the contribution over the major arcs wins over that of the minor arcs for $N \geqslant 10^{27}$.

For our considerations, we are interested in Helfgott's techniques for improving the bounds, but not so much in explicit constants. The main bound which Helfgott achieves is the following.

Theorem 4.1. *Let $S_\eta(x,\alpha)$ be as in (4.1), with $\eta(t) = 4\max\{\log 2 - |\log 2t|, 0\}$. Let $2\alpha = {}^a\!/\!_q + {}^\delta\!/\!_x$ with a, q coprime integers, where $|{}^\delta\!/\!_x| \leqslant {}^1\!/\!_{qQ}$, $Q = ({}^3\!/\!_4)x^{2/3}$, and $q \leqslant \sqrt[3]{x}/6$. Then*

$$|S_\eta(x,\alpha)| \lll \frac{x \log q}{\sqrt{\varphi(q)}}. \tag{4.2}$$

Comparing this with the bound in theorem 3.1, we notice that the improvement is essentially in the removal of of four logarithm factors. Throughout this chapter we will use some results from previous chapters, such as Vaughan's identity and the lemmas in chapter 3, but state them in a different way to be consistent with Helfgott's notation in [11].

4.1 Main Ideas

The general strategy of Helfgott's proof is similar to that of the unconditional proof presented in chapter 3, utilising the circle method and Vaughan's identity to obtain the bound in (4.2).

In his proof, the major arcs $\mathfrak{M}$ consist of those $\alpha \in \mathbb{R}/\mathbb{Z}$ which can be approximated (in the usual sense of definition 2.6) with $q \leqslant 300\,000$ and Q a constant times $x/300\,000$. The major arc bounds which Helfgott obtains rely on verifying GRH up to a given imaginary height T in the complex plane, i.e., showing that $L(s,\chi)$ has no zeros with $\Re(s) \neq {}^1\!/\!_2$ and $|\Im(s)| < T$. These bounds work only for $|\delta| < 4 \cdot 300\,000/q$, where $\alpha = {}^a\!/\!_q + {}^\delta\!/\!_x$.

Thus for the minor arcs $\mathfrak{m}$, we want a bound on $|S_\eta(x,\alpha)|$ which decreases as q and δ increase, where $\alpha = {}^a\!/\!_q + {}^\delta\!/\!_x$, and because of the restrictions on the major arc bounds, we cannot allow any $\log x$ factors in front of terms such as $x/\sqrt{q}$, because for large x we would have that this is worse than the trivial bound x (2.4). Helfgott remarks that a main bound proportional to $(\log^2 q/\sqrt{q})x$ was not good enough either for computing purposes, such a bound was obtained by Tao in 2014.[11, 17] We should also mention that there are better asymptotic bounds than (4.2) in the literature. Indeed, in [14], Ramaré gives the bound

$$|S(x, {}^{an}\!/\!_q)| \leqslant 13\,000 \frac{\sqrt{q}}{\varphi(q)}\, x,$$

which is the best sort of bound one expects to obtain through Vinogradov's method, but the large explicit constant, in addition to the requirement $q \leqslant x^{1/48}$, makes this bound unsuitable for keeping the explicit constants small. Particularly because we care mainly about q around $300\,000$, which is the cut-off point that Helfgott sets for major arc denominators.

The four $\log x$ factors are removed using ideas detailed in the following sections. Notice that two of the four logarithm factors can be traced back to Vaughan's identity itself (indeed, summing over the left-hand side of (3.1) gives $\sum_{n\leqslant x} \Lambda(n) = \Psi(n) \sim x$, but the sum over the right-hand side is $\lll x \log^2 x$). We closely follow chapters 1 and 3 of [11] in this exposition.

4.2 Type I Sums

Just as before, we have two type I sums, namely

$$S_{\mathrm{I},1} = \sum_{m \leqslant U} \sum_{n} \mu(m)\,(\log n)\, e(\alpha m n)\, \eta({}^{mn}\!/\!_x)$$

and

$$S_{\mathrm{I},2} = \sum_{v \leqslant V} \Lambda(v) \sum_{u \leqslant U} \sum_{n} \mu(m)\, e(\alpha v u n)\, \eta({}^{vun}\!/\!_x)$$

where $\alpha = {}^a\!/\!_q + {}^\delta\!/\!_x$, the difference here from chapter 3 is simply the smoothing η (we are also using U and V in place of X and Y in Vaughan's identity). As these are usually treated the same, here we will study the simpler sum

$$S_{\mathrm{I}} = \sum_{m \leqslant D} \sum_{n} \mu(m)\, e(\alpha m n)\, \eta({}^{mn}\!/\!_x),$$

where D is, as usual, a bound smaller than x. It suffices to study this sum because $S_{\mathrm{I},1}$ is basically of the same form (log is slowly varying so we can ignore it for small numbers), and the inner sum of $S_{\mathrm{I},2}$ is the same as S_{I} with αv in place of α, which means that for q small we can use S_{I} to estimate $S_{\mathrm{I},2}$. For q not small, we can treat $S_{\mathrm{I},2}$ as $\sum_n (\Lambda_{\leqslant V} * \mu_{\leqslant U})(n)$ and bound it trivially (by $\log n$).

Now recall that in Vinogradov's proof, we bound $|\sum_{n\leqslant N} e(\alpha n)|$ by $\min\{N, {}^1\!/\!_{\|\alpha\|}\}$, and invoke lemma 3.3. The proof of the lemma essentially involved splitting up the sum into sums of length q and bounding $\|\alpha\|$ from below. Applying this directly to our sum we get a bound of the form

$$\begin{aligned}\Big|\sum_{y<m\leqslant y+q} \mu(m) \sum_{n\leqslant N} e(\alpha mn)\Big| &\leqslant \sum_{y<m\leqslant y+q} \Big|\sum_{n\leqslant N} e(\alpha mn)\Big| \\ &\leqslant 2\min\big\{{}^x\!/\!_y, q\big\} + 2 \sum_{r\leqslant (q-1)/2} \frac{1/2}{r/q} \\ &\leqslant 2\min\big\{{}^x\!/\!_y, q\big\} + q\log q\end{aligned}$$

for all y. There are two obvious improvements to be made here, firstly we can estimating the inner sum more precisely. One can define a smoothing η and get that

$$\Big|\sum_{n\leqslant N} e(\alpha n)\, \eta({}^n\!/\!_x)\Big| \leqslant \min\Big\{x\|\eta\|_1 + \frac{\|\eta'\|_1}{2}, \frac{\|\eta'\|_1}{2|\sin\pi\alpha|}, \frac{\|\widehat{\eta''}\|_\infty}{4x\sin^2\pi\alpha}\Big\}, \tag{4.3}$$

where $\|\cdot\|_r$ denotes the L^r norm $\big(\int_{\mathbb{R}} |\cdot|^r\big)^{1/r}$.[10] This gives an improvement for large m, however the resulting term is still $\ll x/y$, resulting in a contribution of $(x\log x)/q$ to S_{I} (so still not log-free). When m is small, the terms which cause $m\alpha$ to be close to zero are those with $q \mid m$. If we exclude them, we can get a bound of the form

$$\sum_{\substack{y<m\leqslant y+q \\ q\nmid m}} \min\Big\{A, \frac{B}{|\sin\pi\alpha n|}, \frac{C}{|\sin\pi\alpha n|^2}\Big\} \ll \min\big\{Cq^2, \sqrt{AC}, Bq\max\{2, \log {}^{Cq}\!/\!_B\}\big\},$$

where m is small in the sense that $y+q \leqslant {}^Q\!/\!_2$, Q being the usual value so that $|\alpha - {}^a\!/\!_q| \leqslant {}^1\!/\!_{qQ}$ holds. There are still terms having $m \leqslant \min\{D, {}^Q\!/\!_2\}$ with $q \mid m$ and those with ${}^Q\!/\!_2 < m \leqslant D$ left. For the former, we use the Poisson summation formula ($\Sigma f = \Sigma \hat{f}$) to estimate the inner sum. Summing over m *without* applying the triangle inequality (in contrast to the above), we get the main term

$$\frac{x\,\mu(q)}{q}\,\hat{\eta}(-\delta) \sum_{\substack{a\leqslant \min\{D, Q/2\}/q \\ (a,q)=1}} \frac{\mu(a)}{a}$$

where $aq = m$. Now comes the second improvement: cancellation over μ. It can be shown that the sum over a is at most 1.[3] This does not give us back a factor of $\log x$ though. Luckily, in [15], Ramaré provides the bound

$$\Big|\sum_{\substack{a\leqslant x\\(a,q)=1}}\frac{\mu(a)}{a}\Big| \leqslant \frac{4}{5}\,\frac{q}{\varphi(q)}\,\frac{1}{\log {}^{x}\!/_{q}}$$

for $q \leqslant x$. This bound is obtained by non-elementary methods (using known properties of the Riemann ζ function). Finally, for $m > {}^{Q}\!/_{2}$, we can obtain a bound of the form

$$\sum_{y<m\leqslant y+q}\min\Big\{A,\,\frac{C}{|\sin\pi\alpha n|^2}\Big\} \ll A + q\sqrt{AC} \tag{4.4}$$

for any y, using ideas not dissimilar to the proof of lemma 3.3. The term A in the bound is proportional to $\|\eta_1\| x/y$ (from (4.3)), which then results in a multiple of $(x\log x)/q$ in S_{I}. Since m is large here, αm being close to zero no longer necessarily corresponds to terms with $m \equiv 0 \bmod q$, so we cannot extract them as before. So here, Helfgott reapplies Dirichlet's approximation theorem (proposition A.1) to obtain another approximation for α, this time taking $Q = x/|\delta q|$. If δ is very small (in other words, if the original ${}^{a}\!/_{q}$ is a very good approximation) then there will be no terms with ${}^{Q}\!/_{2} < m \leqslant D$, since Q will exceed $2D$. If this is not the case, let ${}^{a'}\!/_{q'}$ be another approximation of α taking some $Q' > Q$ in proposition A.1. Then $|{}^{a}\!/_{q} - {}^{a'}\!/_{q'}| \geqslant {}^{1}\!/_{qq'}$, which implies that $q' \geqslant \frac{\epsilon}{1+\epsilon}Q$, and so if we apply (4.4) with this new approximation instead, this effectively gets rid of A, since for the first sum over $y < m \leqslant y + q'$ with $y \geqslant {}^{Q}\!/_{2}$, the contribution is at most $x/({}^{Q}\!/_{2})$, and all other contributions of A add up to $\ll (x\log q)/q'$.

Summing everything up produces a bound with main terms

$$\frac{x\min\{1, {}^{1}\!/_{\delta^2}\}}{\varphi(q)\,\log {}^{x}\!/_{q}}, \qquad D, \qquad \text{and} \qquad q\log(\max\{{}^{D}\!/_{q}, q\}),$$

and in most cases, the main term is the first one. Notice this has the shape $x/(\varphi(q)\log x)$ for small q, and decreases rapidly as the error δ increases.

4.3 Type II Sums

The type II sum we have is

$$S_{\mathrm{II}} = \sum_{m}\sum_{n>V}(1 * \mu_{>U})(m)\,\Lambda(n)\,e(\alpha mn)\,\eta({}^{mn}\!/_{x}).$$

We assume that the smoothing η is the multiplicative convolution of two functions η_0 and η_1, i.e., $\eta(t) = (\eta_0 \circledast \eta_1)(t) := \int_0^\infty \eta_0(\tau)\,\eta_1({}^{t}\!/_{\tau})\,\frac{d\tau}{\tau}$. Analogously to the proof of theorem 3.1, we write the sum as

$$\int_V^{x/U} \sum_m \Big(\sum_{\substack{d>U \\ d|m}} \mu(d) \Big)\, \eta_0\Big(\frac{m}{x/W}\Big) \sum_{n\geqslant V} \Lambda(n)\, e(\alpha mn)\, \eta_1({}^{n}\!/\!{}_{W})\, \frac{dW}{W}$$

so that the we can apply Cauchy–Schwarz to the integrand. Indeed, the integrand is at most $\sqrt{S_1(U,W)\, S_2(U,W)}$ where

$$S_1(U,W) := \sum_{x/2W\leqslant m\leqslant x/W} \Big| \sum_{\substack{d>U \\ d|m}} \mu(d) \Big|^2, \;\; S_2(U,W) := \sum_{x/2W\leqslant m\leqslant x/W} \Big| \sum_{\max\{V, W/2\}\leqslant n\leqslant W} \Lambda(n)\, e(\alpha mn) \Big|^2.$$

Now in the last chapter, we would bound $S_1(U,W)$ by something like ${}^{x}\!/\!{}_{W} \log^3({}^{x}\!/\!{}_{W})$ using proposition A.3. What we would like here is a bound $\ll x/W$. Indeed, this is possible, using bounds on $\sum_{n\leqslant t} \mu(n)/n$, and properties of the divisor sum function $\sigma(n) = n \prod_{p|n} \big(1 + \frac{1}{p}\big)$.

To bound S_2, we make use of the large sieve. The idea is the following. Suppose $f\colon \mathbb{Z} \to \mathbb{C}$ is a function supported on an interval I of length ℓ. Parseval's identity gives us that $\int_{\mathbb{R}/\mathbb{Z}} |\hat{f}(\alpha)|^2\, d\alpha = \sum_n |f(n)|^2$. If we take a "sample" of reasonably spaced out points in $\mathbb{C}$, say $\alpha_1, \ldots, \alpha_k$ where $|\alpha_i - \alpha_j| \geqslant \beta$ for $i \neq j$, we have that

$$\sum_{i\leqslant k} |\hat{f}(\alpha_i)|^2 \leqslant (\ell + {}^{1}\!/\!{}_{\beta}) \sum_n |f(n)|^2,$$

which we can think of as a statistical equivalent of Parseval's identity. Now suppose $\alpha_1 = \alpha$, $\alpha_2 = 2\alpha$, and so on. If $\alpha = {}^{a}\!/\!{}_{q}$, then the angles $\alpha_1, \ldots, \alpha_q$ are spaced out with a distance of ${}^{1}\!/\!{}_{q}$ between any two, and $\alpha_{q+1} = \alpha_1$. So by the above, we can split S_2, which is a sum over an interval of length ${}^{x}\!/\!{}_{2W}$, into q bits of length $\lceil ({}^{x}\!/\!{}_{2Wq}) \rceil$ to get a bound of the form

$$\frac{\log W}{\log {}^{W}\!/\!{}_{2q}} \Big(\frac{x}{\varphi(q)} + \frac{qW}{\varphi(q)} \Big) W \tag{4.5}$$

after applying Montgomery's inequality.[1] Now if the error numerator $|\delta|$ is not close to zero, then α_1 and α_{q+1} are different, in particular α_{q+1} is at least $q|\delta|/x$ away from each α_i. Thus we can plug in these angles to the large sieve instead; namely $\alpha_1, \ldots, \alpha_m$ until there is overlap (α_m is the first within $< {}^{1}\!/\!{}_{q}$ of the others). Thus doing this $\lceil \ell/m \rceil \leqslant \lceil \ell/(x/|\delta|q) \rceil$ times, where $\ell = x/2W$, we get a bound of $\lceil \ell/(x/|\delta|q) \rceil (W/2 + x/|\delta|q) \sum_n |f(n)|^2$, which results in about

$$\Big(\frac{Wx}{Q} + x \Big) \log W$$

provided $\ell \geqslant x/|\delta|q$. If not, there is no overlap, and we put *all* α_i into the large sieve. The total bound obtained this way is $(W^2/4 + xW/2|\delta|q) \log W$. If ℓ is significantly than $x/|\delta|q$,

[1] $\big|\sum_n \alpha_n\big|^2 \mu(q)^2 \prod_{p|q} \frac{\omega(p)}{p-\omega(p)} \leqslant \sum_{a\leqslant q,(a,q)=1} \big|\sum_{n\leqslant N} \alpha_n e({}^{an}\!/\!{}_{q})\big|$, see [13, pp. 27–29].

then the α_i "swarm" around rationals ${}^a\!/\!_q$. If we apply Montgomery's inequality here, this has the effect of spreading out these angles, while keeping them sufficiently separated. This gives us a bound with the shape

$$\frac{\log W}{\log {}^W\!/\!_{|\delta|q}}\Big(\frac{x}{|\delta|\,\varphi(q)}+\frac{qW}{\varphi(q)}\Big)\,W,$$

which saves a factor of $|\delta|$ when compared with (4.5).

Thus we have no log factors, apart from $\log x/UV$ which comes from evaluating the integral. Thus we want to choose the parameters of Vaughan's identity to be close to x to counteract this, but at the same time we want to keep the term $D = UV$ small, since it is one of the main terms in the bound for the type I sums. Thus U and V are chosen so that

$$UV = x/\sqrt{q\max(4,|\delta|)}.$$

Combining everything, with various intricate arguments, Helfgott finally obtains the bound

$$|S_\eta(\alpha,x)| \leqslant \frac{R_{x,\delta_0 q}\log\delta_0 q+0.5}{\sqrt{\delta_0\varphi(q)}}\,x+\frac{2.5x}{\sqrt{\delta_0 q}}+\frac{2x}{\delta_0 q}\,L_{x,\delta_0 q,q}+3.36x^{5/6},$$

where $\delta_0 = \max\{2,\frac{|\delta|}{4}\}$,

$$R_{x,t} = 0.27125\log\left(1+\frac{\log 4t}{2\log\frac{9x^{1/3}}{2.004t}}\right)+0.41415,$$

and

$$L_{x,t,q} = \frac{q}{\varphi(q)}\big(\tfrac{13}{4}\log t+7.82\big)+13.66\log t+37.55.$$

The factor $R_{x,t}$ is bounded, Helfgott states that for "difficult" values of x and $\delta_0 x$ it is still less than one. We therefore have the main term in (4.2).

Chapter 5

Conclusion

Let us end by giving a retrospective bird's eye view of the proof, and compare the different minor arc bounds we had in the different versions we saw. The general strategy is to ensure that the quantity $r(N)$, which counts the number of ways of representing N as the sum of three prime powers, is positive (we saw in the introduction that proper prime powers contribute little to $r(N)$ overall). In particular, we achieve this by showing that

$$r(N) = \int_{\mathfrak{M}} S(N,\alpha)^3 \, e(-N\alpha)\, d\alpha + \int_{\mathfrak{m}} S(N,\alpha)^3 \, e(-N\alpha)\, d\alpha$$

is non-negative for large enough N, since the major arc contribution will be $\sim N^2$, and the minor arc term will be $O(N^2 \log^{-A} N)$ for some $A > 0$.

The strategy for estimating the major arc integral is the following. Using either the PNT in arithmetic progressions (on GRH) or Siegel's theorem (unconditional), we can obtain that $S(N, {}^{a}\!/_{q}) \sim (\mu(q)/\varphi(q)) \cdot N$. Then if $\mathfrak{M}(a,q)$ denotes the arc centred at ${}^{a}\!/_{q}$, we have

$$\begin{aligned}\int_{\mathfrak{M}} S(N,\alpha)^3 \, e(-N\alpha)\, d\alpha &= \sum_{\substack{q \leqslant Q \\ (a,q)=1}} \int_{\mathfrak{M}(a,q)} S(N,\alpha)^3 \, e(-N\alpha)\, d\alpha \\ &\sim N^2 \sum_{q \leqslant Q} \frac{\mu(q)}{\varphi(q)^3} \sum_{(a,q)=1} e(\,{}^{-Na}\!/_{q}) \asymp N^2.\end{aligned}$$

Loosely speaking, Helfgott's treatment of the major arcs is closer to the conditional proof than the one using Siegel's theorem, because the constants which come from Siegel's theorem are not explicit. Indeed, to prove the bounds he obtains, he uses D. Platt's finite verification method to verify GRH up to a certain imaginary height.

For the minor arcs, the strategy is the following. In both Hardy–Littlewood and Vinogradov's proof, we made use of Parseval's identity to get that

$$\int_{\mathfrak{m}} S(N,\alpha)^3 \, e(-N\alpha)\, d\alpha \leqslant \sup_{\alpha \in \mathfrak{m}} |S(N,\alpha)| \sum_{n \leqslant N} \Lambda(n)^2 \lll \sup_{\alpha \in \mathfrak{m}} |S(N,\alpha)|\, N \log N,$$

so to win over the major arcs, we needed a bound at least as good as $N/\log^{1+\epsilon} N$. In the conditional proof, the good error terms in the conditional PNT in arithmetic progressions (theorem 2.1) allowed us to deduce a bound on $S(N,\alpha)$ with main term $N/\varphi(q) \leqslant N/\log^9 N$ (since we defined the minor arcs to have $q \geqslant \log^{10} N$). In the unconditional Vinogradov proof, we used Vaughan's identity to carefully take advantage of cancellation in exponential sums, obtaining a bound with main term $(N\log^4 N)/\sqrt{q}$. Here we had $q > \log^B N$ with B unspecified, which makes the bound $\leqslant N \log^{-B/2+4} N$. Thus taking any $B > 10$ gives us what we need.

Finally, Helfgott gives a bound of the form $\leqslant N \log q/\sqrt{\varphi(q)}$ for $S(N,\alpha)$, but his approach to bounding the minor arc contribution is different to the classical one, since the presence of $\sum_{n \leqslant N} \Lambda(n)^2$ introduces a $\log N$ factor so the bound is not log-free. In his proof, the contribution over the minor arcs turns out to be proportional to $\|\eta_+\|_2{}^2\|\eta_*\|_1$, which are two smoothing weights appearing in the minor arc integral $\int_{\mathfrak{m}} |S_{\eta_+}(N,\alpha)|^2 S_{\eta_*}(N,\alpha)\, e(-N\alpha)\, d\alpha$.

5.1 The Binary Goldbach Conjecture

Having developed this proof strategy, why not try to attack the binary Goldbach conjecture in a similar way? If we let

$$r(N) = \sum_{k_1+k_2=N} \Lambda(k_1)\,\Lambda(k_2) \qquad \text{and} \qquad S(N,\alpha) = \sum_{k \leqslant N} \Lambda(k)\, e(k\alpha),$$

then just as in the ternary case, we get $r(N) = \int_{\mathbb{R}/\mathbb{Z}} S(N,\alpha)^2\, e(-N\alpha)\, d\alpha$, so we would want to define major and minor arcs so that

$$r(N) = \int_{\mathfrak{M}} S(N,\alpha)^2\, e(-N\alpha)\, d\alpha + \int_{\mathfrak{m}} S(N,\alpha)^2\, e(-N\alpha)\, d\alpha > 0.$$

(We similarly have that proper prime powers contribute a negligible amount). By analogous arguments, we can establish that $S(N, {}^a/_q) \sim (\mu(q)/\varphi(q)) \cdot N$, and that for even N, $\mathfrak{S}(N) = \sum_q \frac{\mu(q)^2}{\varphi(q)^2} \sum_{(a,q)=1} e(\,{}^{-Na}/_q) \asymp 1$. But this time, we get that

$$\int_{\mathfrak{M}} S(N,\alpha)^2\, e(-N\alpha)\, d\alpha \asymp N,$$

but for the minor arc contribution, both pointwise and square-integral estimates are too large; we have

$$\int_{\mathbb{R}/\mathbb{Z}} |S(N,\alpha)|^2\, d\alpha \leqslant N \log N \qquad \text{and} \qquad \sup_{\alpha \in \mathfrak{m}} |S(N,\alpha)|^2 \leqslant \frac{N^2}{\log^A N}$$

for any $A > 0$, which follow by Parseval and Vinogradov's bound respectively. Therefore it is simply not true that the minor arcs contribute less to $r(N)$ than the major arcs do. Thus, to prove the binary Goldbach conjecture, a fundamentally new idea is required.

Appendix A

Auxiliary Results

Proposition A.1 (Dirichlet's approximation theorem)**.** *Let α and Q be real numbers with $Q \geqslant 1$. Then there exists a rational number ${}^{a}/_{q}$ with $(a,q)=1$ and $1 \leqslant q \leqslant Q$ such that*

$$|\alpha - {}^{a}/_{q}| \leqslant {}^{1}/_{qQ}.$$

Proof. Let $\beta_t := \alpha t - \lfloor \alpha t \rfloor \in [0,1)$ for $t = 1, \dots, \lfloor Q \rfloor$. Partition $[0,1)$ into intervals

$$B_r := \left[\frac{r-1}{\lfloor Q \rfloor + 1}, \frac{r}{\lfloor Q \rfloor + 1} \right)$$

where $r = 1, \dots, \lfloor Q \rfloor + 1$. If there is a $\beta_t \in B_1$ or $B_{\lfloor Q \rfloor + 1}$ then we are done. Indeed, if $\beta_t \in B_1$, then $|\alpha - {}^{a}/_{q}| < {}^{1}/_{q(\lfloor Q \rfloor + 1)} < {}^{1}/_{qQ}$ where $a = \lfloor \alpha \rfloor$ and $q = 1$, and a similar argument applies for $B_{\lfloor Q \rfloor + 1}$. If not, then by the pigeonhole principle, there are $\beta_u, \beta_v \in B_r$ for some u, v, r with $u < v$ and $2 \leqslant r \leqslant \lfloor Q \rfloor$. Set $a := \lfloor \alpha v \rfloor - \lfloor \alpha u \rfloor$ and $q := v - u$. Then

$$|\alpha - {}^{a}/_{q}| = {}^{1}/_{q}\, |\beta_v - \beta_u| < {}^{1}/_{q(\lfloor Q \rfloor + 1)} < {}^{1}/_{qQ},$$

as required. □

Proposition A.2. *Let N, X, Y be positive integers and $f \colon \mathbb{N}^2 \to \mathbb{C}$, $g \colon \mathbb{N} \to \mathbb{C}$. Then*

(i) $$\sum_{n \leqslant N} \sum_{\substack{xy = n \\ x \leqslant X}} f(x,y)\, g(n) = \sum_{x \leqslant X} \sum_{y \leqslant {}^{N}/_{x}} f(x,y)\, g(xy),$$

(ii) $$\sum_{n \leqslant N} \sum_{\substack{xyz = n \\ x \leqslant X \\ y \leqslant Y}} f(x,y)\, g(n) = \sum_{x \leqslant XY} \sum_{y \leqslant {}^{N}/_{x}} \sum_{\substack{dz = x \\ d \leqslant X, z \leqslant Y}} f(d,z)\, g(xy),$$

(iii) $$\sum_{n \leqslant N} \sum_{\substack{xyz = n \\ x > X \\ y > Y}} f(x,y)\, g(n) = \sum_{x > X} \sum_{Y < y \leqslant {}^{N}/_{x}} \sum_{\substack{d \mid x \\ d \leqslant Z}} f(d,y)\, g(xy) \text{ for any } Z \text{ with } XYZ \geqslant N.$$

Proof. For (i), we have

$$\begin{aligned}\sum_{n\leqslant N}\sum_{\substack{xy=n\\x\leqslant X}} f(x,y)\,g(n) &= \sum_{n\leqslant N}\sum_{x\leqslant X}\mathbb{1}_{xy=n}\,f(x,y)\,g(n)\\ &= \sum_{x\leqslant X}\sum_{n\leqslant N}\mathbb{1}_{xy=n}\,f(x,y)\,g(xy)\\ &= \sum_{x\leqslant X}\sum_{y\leqslant N/x} f(x,y)\,g(xy),\end{aligned}$$

for (ii), let $r = xy$. Then

$$\begin{aligned}\sum_{n\leqslant N}\sum_{\substack{xyz=n\\x\leqslant X\\y\leqslant Y}} f(x,y)\,g(n) &= \sum_{n\leqslant N}\sum_{x\leqslant X}\sum_{y\leqslant Y}\mathbb{1}_{xyz=n}\,f(x,y)\,g(n)\\ &= \sum_{n\leqslant N}\sum_{x\leqslant X}\sum_{y\leqslant Y}\Big(\sum_{r\leqslant XY}\mathbb{1}_{r=xy}\Big)\mathbb{1}_{xyz=n}\,f(x,y)\,g(rz)\\ &= \sum_{r\leqslant XY}\sum_{n\leqslant N}\sum_{x\leqslant X}\sum_{y\leqslant Y}\mathbb{1}_{r=xy}\mathbb{1}_{xyz=n}\,f(x,y)\,g(rz)\\ &= \sum_{r\leqslant XY}\sum_{z\leqslant N/r}\sum_{x\leqslant X}\sum_{y\leqslant Y}\mathbb{1}_{r=xy}\mathbb{1}_{xyz=n}\,f(x,y)\,g(rz)\\ &= \sum_{r\leqslant XY}\sum_{z\leqslant N/r}\sum_{\substack{xy=r\\x\leqslant X\\y\leqslant Y}} f(x,y)\,g(rz),\end{aligned}$$

and for (iii), suppose $Z \geqslant N/XY$. Then if $xyz = n$ with $x > X$ and $y > Y$, we have $z \leqslant Z$, so

$$\begin{aligned}\sum_{n\leqslant N}\sum_{\substack{xyz=n\\x>X\\y>Y}} f(x,y)\,g(n) &= \sum_{n\leqslant N}\sum_{x>X}\sum_{z\leqslant Z}\mathbb{1}_{xyz=n}\mathbb{1}_{y>Y}\,f(x,y)\,g(n)\\ &= \sum_{n\leqslant N}\sum_{x>X}\sum_{z\leqslant Z}\Big(\sum_{r>X}\mathbb{1}_{r=xz}\Big)\mathbb{1}_{xyz=n}\mathbb{1}_{y>Y}\,f(x,y)\,g(ry)\\ &= \sum_{r>X}\sum_{y\leqslant N/r}\sum_{z\leqslant Z}\sum_{x>X}\mathbb{1}_{r=xz}\mathbb{1}_{xyz=n}\mathbb{1}_{y>Y}\,f(x,y)\,g(ry)\\ &= \sum_{r>X}\sum_{Y<y\leqslant N/r}\sum_{z\leqslant Z}\sum_{x>X}\mathbb{1}_{r=xz}\mathbb{1}_{xyz=n}\,f(x,y)\,g(ry)\\ &= \sum_{r>X}\sum_{Y<y\leqslant N/r}\sum_{\substack{zx=r\\z\leqslant Z}}\sum_{x>X}\mathbb{1}_{xyz=n}\,f(x,y)\,g(ry)\\ &= \sum_{r>X}\sum_{Y<y\leqslant N/r}\sum_{\substack{zx=r\\z\leqslant Z}} f(x,y)\,g(ry).\end{aligned}$$

□

Proposition A.3. *Let $d(n) = \#\{m \in \mathbb{N} : m \mid n\}$, i.e., the number of divisors of n. Then*

$$\sum_{x \leqslant n} d(x)^2 \ll n \, \log^3 n.$$

Proof.

$$\begin{aligned}
\sum_{x \leqslant n} d^2(x) = \sum_{x \leqslant n} \sum_{a \mid x} \sum_{b \mid x} 1 &= \sum_{a \leqslant n} \sum_{b \leqslant n} \sum_{k \leqslant n / \operatorname{lcm}\{a,b\}} 1 \\
&\leqslant \sum_{c \leqslant n} \sum_{k \leqslant n/c} \sum_{\ell \leqslant n/ck} \sum_{m \leqslant n/ck\ell} 1 \\
&\leqslant \sum_{c \leqslant n} \sum_{k \leqslant n/c} \sum_{\ell \leqslant n/ck} n/ck\ell \\
&= \sum_{c \leqslant n} \sum_{k \leqslant n/c} n/ck \sum_{\ell \leqslant n/ck} 1/\ell \\
&\leqslant \sum_{c \leqslant n} \sum_{k \leqslant n/c} n/ck(\log n + 1) \\
&\leqslant \sum_{c \leqslant n} n/c(\log n + 1)^2 \\
&\leqslant n(\log n + 1)^3.
\end{aligned}$$

□

Bibliography

[1] Davenport, H. *Multiplicative Number Theory*, third ed. Springer–Verlag, 1995.

[2] Gowers, T. Additive and combinatorial number theory. (Course notes, Winter 2019).

[3] Granville, A., and Ramaré, O. Explicit bounds on exponential sums and the scarcity of squarefree binomial coefficients. *Mathematika 43*, 1 (1996), 73–107.

[4] Hardy, G. H., and Littlewood, J. E. Some problems of 'partito numerorum'; iii: On the expression of a number as a sum of primes. *Acta Mathematica 44* (1923), 1–70.

[5] Hardy, G. H., and Ramanujan, S. Asymptotic formulæ in combinatory analysis. *Proceedings of the London Mathematical Society S2–17*, 1 (1918), 75–115.

[6] Hardy, G. H., and Wright, E. M. *An Introduction to the Theory of Numbers*, fourth ed. Oxford University Press, 1959.

[7] Harper, A. Analytic number theory. (Course notes, Winter 2019).

[8] Helfgott, H. A. Minor arcs for Goldbach's problem. *ArXiv math.NT* (2012), 1205.5252.

[9] Helfgott, H. A. Major arcs for Goldbach's problem. *ArXiv math.NT* (2013), 1305.2897.

[10] Helfgott, H. A. The ternary Goldbach conjecture is true. *ArXiv math.NT* (2014), 1312.7748.

[11] Helfgott, H. A. The ternary Goldbach problem (book, first edition). *ArXiv math.NT* (2015), 1501.05438.

[12] Miller, S. J., and Bighash, R. T. *An Invitation to Modern Number Theory*, first ed. Princeton University Press, 2006.

[13] Montgomery, H. L. Topics in multiplicative number theory. *Lecture Notes in Mathematics, Springer-Verlag, 227* (1971), 27–29.

[14] Ramaré, O. On Bombieri's asymptotic sieve. *Journal of Number Theory 130*, 5 (2010), 1155–1189.

[15] Ramaré, O. Explicit estimates on several summatory functions involving the moebius function. *Mathematics of Computation 84*, 293 (2015), 1359–1387.

[16] Soundararajan, K. Additive combinatorics. (Course notes, Winter 2007). Available online, last accessed: 13th December, 2019, url: `https://math.stanford.edu/~ksound/Notes.pdf`.

[17] Tao, T. Every odd number greater than 1 is the sum of at most 5 primes. *Mathematics of Computation 83*, 286 (2014), 997–1038.

[18] Vaughan, R. C. *The Hardy–Littlewood Method*, second ed. Cambridge University Press, 1997.

[19] Vinogradov, I. M. Representation of an odd number as the sum of three primes. *Doklady Akademii Nauk SSSR 15* (1937), 291–294.

[20] Vinogradov, I. M. *The Method of Trigonometrical Sums in the Theory of Numbers*, revised ed. Dover Publications, 2004.

Index