CHROMATIC SUMS OF SQUARES AND PRIMES

By

P. AKHILESH

MATH08200904003

Harish-Chandra Research Institute, Allahabad

A thesis submitted to the

Board of Studies in Mathematical Sciences

In partial fulfillment of requirements

for the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

October, 2015

Homi Bhabha National Institute[1]

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Mr. P. AKHILESH entitled ”Chromatic Sums of Squares and Primes” and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman - Prof. B. Ramakrishnan **Date:** 4/3/2016

Guide / Convener - Prof. D. Surya Ramana **Date:** 4 Mar 2016

Co-guide - N/A **Date:**

Examiner - Prof. Ritabrata Munshi **Date:** 4/3/16

Member 1- R. Thangadurai **Date:** 04/03/16.

RThangadurai

Member 2- **Date:**

Final approval and acceptance of this thesis is contingent upon the candidate’s submission of the final copies of the thesis to HBNI.

I/We hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 7 March 2016

Place: Allahabad

D. Surya Ramana

Guide

[1] This page is to be included only for final submission after successful completion of viva voce.

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

P. Akhilesh

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

P. Akhilesh

List of Publications arising from the thesis

Journal

1. "*A Remark on the Beurling-Selberg function*", P. Akhilesh and D.S. Ramana, *Acta Math. Hungarica*, **2013**, 139 no. 4, 354-362.

2. "*A Chromatic version of Lagrange's Four Squares Theorem*", P. Akhilesh and D.S. Ramana, *Monatshefte für Mathematik,* **2015,** 176 no. 1, 17-29.

P. Akhilesh

To

My Teachers

and

My Purankal Tharavad

Acknowledgments

I express my sincere gratitude to my advisor Prof. D.S. Ramana for his support throughout my Ph.D studies and research. I am grateful to the members of my doctoral committee, Prof. B Ramakrishnan and Prof. R Thangadurai, for their helpful comments and encouragement throughout my work on this thesis. I would also like to thank all faculty members of HRI for the courses they gave during my years of study here. My special thanks to Prof. Gyan Prakash for a number of useful discussions.

I thank all the staff of the administration of HRI for making my stay comfortable. I am especially thankful to the staff of the library, who were helpful at all times. I am thankful to the administration of the HPC cluster facility at HRI, on which several numerical calculations were carried out.

I thank all my friends at HRI for their camaraderie. I would like to especially thank Sohani, Roji, Jay, Kasi, Pradip, Pradeep, Mallesham, Pramod, Bhuwanesh, Vikas, Rajarshi, Senthil, Ritika and Mithun for their understanding and constant encouragement.

I am deeply thankful to my family and all members of my Purankal Tharavad for their love, support and sacrifices.

P. Akhilesh

Contents

List of Key Notations

$\mathbf{R}$	The set of real numbers.
$\mathbf{Z}$	The set of integers.
$\mathbf{T}$	$\mathbf{R}/\mathbf{Z}$
lS	$lS = \{s_1 + s_2 + ... + s_l \mid (s_1, s_2, ..., s_l) \in S^l\}$
$E_m(S)$	See page 2.
$\mathrm{sgn}(z)$	1 when $\mathrm{Re}(z) \geq 0$ and -1 when $\mathrm{Re}(z) < 0$
$a(t)$	$1 - \lvert t \rvert$ when $\lvert t \rvert < 1$ and 0 when $\lvert t \rvert \geq 1$
$c(t)$	$\mathrm{sgn}(t)$ when $\lvert t \rvert < 1$ and 0 when $\lvert t \rvert \geq 1$
$v_p(n)$	The largest integer d such that p^d divides n
$\pi^*(N)$	The number of primes in $(2N, 3N]$
$\omega(n)$	The number of prime divisors of an integer n
$\tau(n)$	The number of divisors of an integer n
$\hat{f}(t)$	$\int_{\mathbf{R}} f(x) e^{-2\pi ixt} dx$

SYNOPSIS

This thesis contains three chapters. The themes of the first two chapters are chromatic versions of the classical theorems of Lagrange and Vinogradov on representing natural numbers as sums of squares and primes, respectively. A key tool in these chapters is the large sieve inequality and certain analogues of it. A description of the contents of these chapters of the thesis is given in Section 1 below. The third chapter of this thesis concerns the Beurling-Selberg function. This function has a number of applications in analytic number theory, in particular to the large sieve inequality. The contents of this chapter are detailed in Section 2 below.

1 Chromatic sums of squares and primes

Lagrange's four squares theorem asserts that every natural number can be written as the sum of four squares of natural numbers. Following A. Sárközy [23], one may ask for a chromatic version of this theorem.

More precisely, let $\mathcal{S}$ denote the set of squares of the natural numbers and let $K \geq 1$ be an integer. Then we define $s(K)$ to be the smallest integer with the property that there is an integer $n(K)$ such that given any partition $\mathcal{S} = \cup_{1 \leq i \leq K} \mathcal{S}_i$ of $\mathcal{S}$ into K (disjoint) subsets $\mathcal{S}_i$ and an integer $n \geq n(K)$ there is an i, $1 \leq i \leq K$, such that n can be expressed as the sum of no more than $s(K)$ squares, all belonging to same $\mathcal{S}_i$. In intuitive terms, if the elements of $\mathcal{S}$ are each coloured by one of K colours then $s(K)$ is the smallest integer such that every sufficiently large integer can be expressed as a sum of at most $s(K)$ squares, all of the same colour.

A. Sárközy remarks on page 29 of [23] that it is easily seen that $s(K)$ is finite and then poses the problems of finding upper bounds, in terms of K, for $s(K)$ as well as the corresponding integer in the analogous question for the set of primes (see Problem 40 of [23]).

N. Hegyvári and F. Hennecart took up Sárközy's problems in [11] and obtained for $s(K)$ the bound $s(K) \ll (K \log K)^5$ (see Theorem 1, page 318 of [11]). They obtain this bound by observing first that if S is any subset of the squares in the interval $(N, 4N]$ satisfying $|S| \geq N^{\frac{1}{2}}/A$ for an integer N and a given real number $A \geq 1$, then we have the lower bound

$$|5S| \gg \frac{N}{A^5} \tag{1}$$

for the cardinality of the sum set $5S = S + S + S + S + S$, where the implied constant is absolute. This is deduced in [11] as an immediate consequence of the well-known asymptotics for the number of representations of a given integer as the sum of five squares, supplied by the circle method. Hegyvári and Hennecart then apply the inequality (1) for $A = K \log K$ within an elegant argument based on a finite analogue of Kneser's theorem, again due to A. Sárközy, to arrive at their upper bound for $s(K)$.

In the paper [1], co-authored with D.S. Ramana, we improved the upper bound for $s(K)$ given by [11] by showing that $s(K) \ll_\epsilon K^{2+\epsilon}$, for any $K \geq 1$ and $\epsilon > 0$. The first chapter of this thesis gives a full account of the proof of this improved upper bound, following [1]. Our path to this bound for $s(K)$ passes through the theorem below, which we state with the aid of the following notation.

For any subset S of the integers and any integer $k \geq 2$, $E_k(S)$ shall denote the number of tuples $(x_1, x_2, \ldots, x_{2k})$ in S^{2k} satisfying the relation

$$x_1 + x_2 + \cdots + x_k = x_{k+1} + x_{k+2} + \cdots + x_{2k} \ . \tag{2}$$

Theorem 1.1. *Let A and ϵ be real numbers with $A \geq 1$ and $\epsilon > 0$. Then there is an integer N_0, depending only on A and ϵ, such that for all $N \geq N_0$ and any subset S of the squares in the interval $[1, N]$ with $|S| \geq N^{\frac{1}{2}}/A$ we have*

$$E_6(S) \ll_\epsilon |S|^{10} A^\epsilon \ . \tag{3}$$

This theorem is proved in [1] by means of the circle method and a certain large sieve inequality for polynomial amplitudes. This inequality is employed within the application of the circle method to estimate the contribution from the major arcs. This method can be seen as an elaboration of the proof of the pruning lemma of Brüdern (see Lemma 2 of [6]). Also, it turns out that this method is, in a sense, “dual“ to an argument given by J. Bourgain [5] in a different context. It is, in particular, possible to obtain an alternate proof of Theorem 1.1, that does not rely on aforementioned large sieve inequality, by modifying Bourgain’s method. We detail all of this in the first chapter of the thesis.

Our upper bound for $s(K)$ is deduced from Theorem 1.1 by means of the following well-known application of the Cauchy-Schwarz inequality, valid for any integer $k \geq 2$

and any subset of the integers S :

$$|kS|\,E_k(S) \geq |S|^{2k}. \tag{4}$$

Here, and in agreement with the notation used earlier, kS denotes, for any integer $k \geq 1$, the sumset $S + S + \cdots + S$, with k summands. Indeed, on using the above inequality with $k = 6$ together with the conclusion of Theorem 1.1 we obtain

$$|6S| \gg \frac{N}{A^{2+\epsilon}}\ , \tag{5}$$

for any $\epsilon > 0$, where S, N and A are as in the statement of Theorem 1.1 and the implied constant depends on ϵ alone. The above inequality is our analogue of (1). Indeed, the improved upper bound for $s(K)$ is deduced in [1] from (5) applied to a suitable S and N with $A = K$ by means of the argument from [11], involving the finite version of Kneser's theorem. The first chapter of the thesis concludes with a description of this deduction.

We now turn to the contents of the second chapter of the thesis, which is on Sárközy's problem for the primes. We begin by explicitly stating this problem. Thus let $\mathcal{P}$ denote the set of prime numbers and suppose that $K \geq 1$ is an integer. Then the problem is to determine upper bounds in terms of K for the smallest integer $t(K)$ with the property that there is an integer $n(K)$ such that given any partition $\mathcal{P} = \cup_{1 \leq i \leq K} \mathcal{P}_i$ of $\mathcal{P}$ into K (disjoint) subsets $\mathcal{P}_i$ and an integer $n \geq n(K)$ there is an i, $1 \leq i \leq K$, such that n can be expressed as the sum of no more than $t(K)$ prime numbers all belonging to $\mathcal{P}_i$.

In [11], Hegyvári and Hennecart showed that $t(K) \leq 1500K^3$. This result was improved by D.S. Ramana and O. Ramaré [18], who obtained

$$t(K) \leq CK \log\log 4K, \tag{6}$$

with C an absolute constant. This upper bound for $t(K)$ is the best possible up to the value of C, on account of a lower bound for $t(K)$ provided by [11]. The bound (6) is deduced in [18] by means of the following, which we state using the notation $\pi^*(N)$ for the number of prime numbers in the interval $(\frac{N}{2}, N]$, for any integer $N \geq 1$.

Theorem 1.2. *For any integer $K \geq 1$ there is an integer $N(K)$ such that for all $N \geq N(K)$ and any subset S of the prime numbers in the interval $(\frac{N}{2}, N]$ with $|S| \geq \pi^*(N)/K$ we have*

$$E_2(S) \;\leq\; \frac{M}{\phi(M)} \frac{|S|^3}{\log\left(\frac{N}{2}\right)} \exp\left(\frac{16}{\log\log 4K}\right) \;, \tag{7}$$

where M is the product of all prime numbers not exceeding $(4 \log 4K \log\log 4K)^2$.

The proof of Theorem 1.2 in [18] comprises two key steps. The first is a (discrete) probabilistic argument that gives a bound of the expected order for the number of triples (x_1, x_2, x_3) in S^3 such that $x_1 + x_2 - x_3$ is invertible modulo U, the product of all prime numbers not exceeding $4^{11}K^2$. The second is an application of the following improved large sieve inequality for the primes.

Proposition 1.3. *Let $N \geq 100$ be an integer and u_n be a finite sequence of complex* *numbers supported on integers all of whose prime factors exceed $N^{\frac{1}{2}}$. Then for any Q satisfying $1 \leq Q \leq N^{\frac{1}{2}}$ we have*

$$\sum_{1 \leq q \leq Q} \sum_{a \bmod^* \mathrm{q}} \left| \sum_n u_n e\left(\frac{an}{q}\right) \right|^2 \;\leq\; \frac{7N \log Q}{\log N} \sum_n |u_n|^2 \,. \tag{8}$$

Note that (8) improves upon the bound supplied by the classical large sieve inequality by the factor $\log Q / \log N$. The inequality (8) is Theorem 5.3 on page 43 of [19]. It's proof depends on O. Ramaré's theory of the large sieve developed in Ramaré and Rusza [20] and in [19].

Our purpose in the second chapter of this thesis is to remark that by working with $E_3(S)$ rather than $E_2(S)$, we may substitute the use of (8) in the proof of (6) given by [18] with an application of the usual large sieve inequality and a bound for $E_2(S)$ obtained by a standard application of the circle method. More precisely, with notation as in Theorem 1.2, we show that

$$E_3(S) \;\leq\; \frac{M}{\phi(M)} \frac{|S|^5}{\log\left(\frac{N}{2}\right)} \exp\left(\frac{C_1}{\log\log 4K}\right) \;, \tag{9}$$

for an absolute constant C_1. Our proof of this bound simplifies the method of [18] to the extent that it does not rely on (8) but only on classical devices and the probabilistic argument from [18] mentioned above. We then combine (9) with (2) for $k = 3$ and the argument from [11] based on the finite version Kneser's theorem to obtain (6) with a constant C only slightly larger than the one given by [18].

2 A remark on the Beurling-Selberg function

Let $K(z)$ denote the entire function $\left(\frac{\sin \pi z}{\pi z}\right)^2$ and set $\mathrm{sgn}(z) = 1$ when $\mathrm{Re}(z) \geq 0$ and $\mathrm{sgn}(z) = -1$ when $\mathrm{Re}(z) < 0$ for any complex z. Then the Beurling-Selberg function $B(z)$ is the function defined on the complex plane by the relation

$$B(z) = 2zK(z) + \sum_{n \in \mathbf{Z}} \mathrm{sgn}(n) K(z-n)\ . \tag{10}$$

$B(z)$ defines an entire function of exponential type 2π on the complex plane that interpolates the values of $\mathrm{sgn}(z)$ at the integers.

If $I = [\alpha, \beta]$ is any compact real interval with integer end points α and β, the restriction to the real line of the complex function F_I defined by $F_I(z) = \frac{1}{2}(B(z-\alpha) + B(\beta - z))$ is an optimal majorant of the characteristic function χ_I of the interval I from the point of view of Fourier analysis on the real line. In more definite terms, among all integrable function f on the real line with Fourier transforms supported in $[-1, 1]$ and satisfying $f(x) \geq \chi_I(x)$ for all real x, the restriction of F_I to the real line deviates the least from χ_I in $L^1(\mathbf{R})$. This property of F_I makes it the ideal choice for a test function in a number of applications of the Fourier transform to number theory, in particular, to the large sieve inequality.

An entire function ϕ defined by a series $\sum_{n \geq 0} a(z+n)$ that is normally convergent on compact subsets of the complex plane may be viewed as the unique solution to the difference equation $\Delta\phi(z) = a(z)$ satisfying the condition that $\phi(x+m)$ tends to 0 as m tends to $+\infty$ for every real number x, where $\Delta\phi(z)$ denotes $\phi(z) - \phi(z+1)$ for any complex z.

In the third and final chapter of the thesis we show that this point of view provides simple proofs of the standard properties of the Beurling-Selberg function given in the literature, for example in [26], and in particular, of those properties mentioned above. This chapter is based on our paper [2] with D.S. Ramana.

CHAPTER 1

Introduction

This thesis contains three chapters in addition to this introduction. The second and third chapters of this thesis are centered around a problem of A. Sárközy that asks for chromatic versions of classical results on representing natural numbers as sums of squares and primes. Inequalities of the large sieve type play a key role in these chapters. A description of their contents is given in Section 1.1 below. Chapter 2 is based on our paper [1].

The fourth chapter of this thesis, which is a version of our paper [2], is on the Beurling-Selberg function. This function has a number of applications in analytic number theory, in particular to the classical large sieve inequality. The contents of this chapter are summarised in Section 1.2.

1.1 Sárközy's Problem

In an article [23] listing a number of unsolved problems in number theory, A. Sárközy asks for chromatic versions, in the spirit of Ramsey theory, of the classical theorems of Lagrange and Vinogradov on additive representation of natural numbers by squares and primes respectively. Indeed, as Sárközy remarks on page 26 of [23], it is not difficult to see that for each integer $K \geq 1$ there is a smallest integer — which in this thesis is denoted by $s(K)$ — such that in every colouring of the set of squares of the integers in K colours, every large enough integer can be represented as a sum of at most $s(K)$ squares, all of the same colour. Sárközy then proposes as Problem 40 of his list the question of obtaining (optimal) upper bounds for $s(K)$ and the corresponding integer $t(K)$ in the analogous question for the set of primes.

N. Hegyvári and F. Hennecart took up Sárközy's problem in [11] and obtained the bounds $s(K) \ll (K \log K)^5$ and $t(K) \leq 1500K^3$ (see Theorems 1 and 3, on pages 318 and 322 of [11]). The principal result of the second chapter of this thesis, stated as Theorem 2.1.1, improves the first of these upper bounds to $s(K) \ll_\epsilon K^{2+\epsilon}$, for any $\epsilon > 0$.

Let us describe the principle of the proof of Theorem 2.1.1, referring to Section 2.1 of Chapter 2 for a more detailed overview. For any subset S of the integers and any integer $m \geq 1$ we will write $E_m(S)$ for the number of tuples $(x_1, x_2, \ldots, x_{2m})$ in S^{2m} satisfying

$$x_1 + x_2 + x_3 + \ldots + x_m = x_{m+1} + x_{m+2} + \ldots + x_{2m} \,. \tag{1.1}$$

We begin by showing that if A and ϵ are real numbers with $A \geq 1$ and $\epsilon > 0$ then there is an integer N_0, depending only on A and ϵ, such that for all $N \geq N_0$ and any subset S of the squares in the interval $[1, N]$ with $|S| \geq N^{\frac{1}{2}}/A$ we have

$$E_6(S) \ll_\epsilon |S|^{10} A^\epsilon \, . \tag{1.2}$$

This is the conclusion of Theorem 2.1.2, which we prove in Section 2.3 by means of the circle method and a large sieve inequality for polynomial amplitudes. In effect, we use this inequality to estimate the contribution from the major arcs within our application of the circle method.

Theorem 2.1.1 is then deduced from Theorem 2.1.2 in Section 2.4 by means of a classical application of the Cauchy-Schwarz inequality followed by an appeal to an elegant argument from [11] that relies on a finite addition theorem due to Sárközy himself.

It turns out that our method for the proof of Theorem 2.1.2, in a sense, "dual" to an argument given by J. Bourgain [5] in a different context. It is, in particular, possible to obtain an alternate proof of Theorem 2.1.2 that does not rely on aforementioned large sieve inequality for polynomial amplitudes by modifying Bourgain's method. We detail this in the final section of Chapter 2.

We now turn to the third chapter of the thesis, which is on Sárközy's problem for the primes. Explicitly, the problem is to determine, for each integer $K \geq 1$, optimal upper bounds for the smallest integer $t(K)$ with the property that in any

K-colouring of the set of primes, every large enough integer is the sum of no more than $t(K)$ primes, all of the same colour. The original upper bound for $t(K)$ given by Hegyvári and Hennecart in [11] was improved by D.S. Ramana and O. Ramaré [18], who showed

$$t(K) \leq CK \log\log 4K, \tag{1.3}$$

for an absolute constant C. This upper bound is the best possible up to the value of C, on account of the lower bound for $t(K)$ given by Theorem 2, page 319 of [11]. Let us also note here that in [14], K. Matomäki gives optimal lower bounds for sum sets of positive density of the set of primes using the techniques of Green and Green-Tao. One may immediately combine her results with the method of [11] to obtain an alternate proof of (1.3), with the constant C that is half that obtained from the method [18].

With a view to discussing the contents of Chapter 3, let us briefly recount the method of [18]. The bound (1.3) is deduced in [18] by means of the theorem below, where $\pi^*(N)$, following the notation of [18] in this one instance, is the number of prime numbers in the interval $(\frac{N}{2}, N]$, for any integer $N \geq 1$.

Theorem 1.1.1. *For any integer $K \geq 1$ there is an integer $N(K)$ such that such that for all $N \geq N(K)$ and any subset S of the prime numbers in the interval $(\frac{N}{2}, N]$ with $|S| \geq \pi^*(N)/K$ we have*

$$E_2(S) \ \leq \ \frac{M}{\phi(M)} \frac{|S|^3}{\log\left(\frac{N}{2}\right)} \exp\left(\frac{16}{\log\log 4K}\right) \ , \tag{1.4}$$

where M is the product of all prime numbers not exceeding $(4\log 4K \log\log 4K)^2$.

The proof of this theorem in [18] uses a variant of a method of Ramaré and Ruzsa [20] and comprises two key steps. The first is a (discrete probabilistic) combinatorial argument that gives a bound of the expected order for the number of triples (x_1, x_2, x_3) in S^3 such that $x_1 + x_2 - x_3$ is invertible modulo U, the product of all prime numbers not exceeding $4^{11}K^2$. The second is a suitable application of the following improved large sieve inequality for the primes.

Proposition 1.1.2. *Let $N \geq 100$ be an integer and u_n be a finite sequence of complex numbers supported on integers all of whose prime factors exceed $N^{\frac{1}{2}}$. Then for any Q satisfying $1 \leq Q \leq N^{\frac{1}{2}}$ we have*

$$\sum_{1\leq q\leq Q} \sum_{a \bmod^* q} \left| \sum_n u_n e\left(\frac{an}{q}\right) \right|^2 \leq \frac{7N\log Q}{\log N} \sum_n |u_n|^2 . \tag{1.5}$$

Note that (1.5) improves upon the bound supplied by the classical large sieve inequality by the factor $\log Q/\log N$, for small enough Q. The inequality (1.5) is the conclusion of Theorem 5.3 on page 43 of [19]. Its proof depends on O. Ramaré's theory of the large sieve developed in Ramaré and Rusza [20] and in [19].

Our purpose in the Chapter 3 is to show that an upper bound for $t(K)$ of the form (1.3), for some constant C, can be obtained by a simplification of the method of [18] that does not rely any more on (1.5) but only on classical devices and the combinatorial argument from [18] mentioned above. We do this by working with $E_3(S)$ rather than $E_2(S)$. This allows us to substitute the use of (1.5) in the method of [18] with an application of the classical large sieve inequality and a simple bound for $E_2(S)$ obtained by a standard application of the circle method. As a consequence we show, with S a subset of the primes in $(2N, 3N]$ satisfying $|S| \geq \frac{\pi^*(N)}{K}$, where $\pi^*(N)$ is now the number of primes in $(2N, 3N]$, and with other notation as in the statement of Theorem 1.1.1, that

$$E_3(S) \;\leq\; \frac{M}{\phi(M)} \frac{|S|^5}{\log{(2N)}} \exp\left(\frac{C_1}{\log\log 4K}\right) \;, \tag{1.6}$$

for an absolute constant C_1. On combining (1.6) with an application of the Cauchy-Schwarz inequality and the argument from [11] based on the finite addition theorem of Sárközy we obtain (1.3) with a constant C only slightly larger than the one given by the method of [18].

1.2 The Beurling-Selberg function

Let $K(z)$ denote $\left(\frac{\sin \pi z}{\pi z}\right)^2$ and set $\mathrm{sgn}(z) = 1$ when $\mathrm{Re}(z) \geq 0$ and $\mathrm{sgn}(z) = -1$ when $\mathrm{Re}(z) < 0$ for any complex z. Then the Beurling-Selberg function $z \mapsto B(z)$ is the function defined on the complex plane by

$$B(z) = 2zK(z) + \sum_{n \in \mathbf{Z}} \mathrm{sgn}(n) K(z-n) \;. \tag{1.7}$$

It is not difficult to see that B is an entire function of exponential type 2π satisfying $B(n) = \mathrm{sgn}(n)$ for each integer n. It is only slightly harder to verify, using the classical identity $\sum_{n\in\mathbf{Z}} K(z-n) = 1$ for all complex z that the restriction to the real line of the function b defined by $b(z) = B(z) - \mathrm{sgn}(z)$ is a positive integrable function satisfying $\int_{\mathbf{R}} b(x)dx = 1$.

The characteristic function χ_I of any compact real interval $I = [\alpha, \beta]$ satisfies $\chi_I(x) = \frac{1}{2}(\mathrm{sgn}(x-\alpha) + \mathrm{sgn}(\beta - x))$, for all real x. Therefore if we define the function F_I by

$$F_I(x) = \frac{1}{2}(B(x-\alpha) + B(\beta - x)) \tag{1.8}$$

for all real x, so that we have

$$F_I(x) = \chi_I(x) + \frac{1}{2}(b(x-\alpha) + b(\beta - x)), \tag{1.9}$$

then from the properties of the function b recalled above we may conclude that F_I is an integrable function on the real line such that $\widehat{F_I}(0) = \beta - \alpha + 1$ and $F_I(x) \geq \chi_I(x)$ for all real x.

The function F_I is called the Selberg majorant of χ_I. A basic property of F_I is that its Fourier transform is supported in the interval $[-1, 1]$. This is shown in the literature either by an appeal to the Paley-Weiner theorem (page 154 of [10]) or by means of the contour integral argument underlying the proof of this theorem (pages 16 to 18 of [3]). A simple alternate proof proceeds as follows. For any complex function $\phi(z)$ let us write $\Delta\phi(z)$ for $\phi(z) - \phi(z+1)$ and $\Delta\phi$ for the function $z \mapsto \Delta\phi(z)$. Then $\Delta\mathrm{sgn}$ is, up to a constant, the characteristic function of $[-1, 0)$. Consequently, we have from (1.7) that

$$\Delta(B(x)) = 2\Delta(xK(x)) - 2K(x+1), \tag{1.10}$$

for all real x. Each term on the right hand side of (1.10) defines an integrable function on the real line with Fourier transform supported in $[-1,1]$. Indeed, the Fourier transforms of $x \mapsto K(x)$ and $x \mapsto \Delta(xK(x))$ are, respectively, the functions $t \mapsto a(t)$ and $t \mapsto -\frac{c(t)(1-e(t))}{2\pi i}$, where $a(t)$ and $c(t)$ are as defined on page v. Thus ΔB is also such a function and so is ΔF_I, from (1.8). Since for any real t we have $\widehat{F}_I(t)(1-e(t)) = \widehat{\Delta F_I}(t)$, we see on dividing by $1-e(t)$ when t is not an integer and using the continuity of the Fourier transform that $\widehat{F}_I$ is also supported in $[-1,1]$.

In Chapter 4 of this thesis, which is its final chapter, we give a full account of the fundamental properties of the Beurling-Selberg function by elaborating on the method of the preceding paragraph.

1.3 Notations

Throughout this thesis, we will use Vinogradov's well-known symbols $\ll$ and $\gg$, in addition to the O and o symbols. Any dependencies on certain parameters of the constants implicit in Vinogradov's notations will be denoted by indicating these parameters as subscripts to the symbols $\ll$ and $\gg$. When a and b are non-zero integers (a,b) denotes the g.c.d. of $|a|$ and $|b|$. If $a=0$ and $b \neq 0$ then $(a,b) = |b|$. Also, we will write $e(z)$ to denote $e^{2\pi i z}$ for any complex number z. A list of key notations used in this thesis is given on page v.

CHAPTER 2

Sárközy's Problem for the Squares

2.1 Introduction

For any integer $K \geq 1$, a colouring in K colours of a subset X of the natural numbers is partition of X into K disjoint subsets. Each subset of X in such a partition is called a colour of the colouring. We shall write $\mathcal{S}$ for the set of squares of the natural numbers and continue to use $s(K)$, for any integer $K \geq 1$, to denote the smallest integer with the property that for any colouring of $\mathcal{S}$ in K colours, every sufficiently large integer is expressible as a sum of at most $s(K)$ squares, all of the same colour.

As stated in Section 1.1, it can be shown that $s(K)$ is finite for each $K \geq 1$ and, indeed, Sárközy's problem for the squares asks for optimal upper bounds for $s(K)$ in terms of K. Our contribution towards the solution of this problem is the following theorem, obtained in the joint work [1].

Theorem 2.1.1. *For any integer $K \geq 1$ and $\epsilon > 0$ we have $s(K) \ll_{\epsilon} K^{2+\epsilon}$.*

This result improves on the bound $s(K) \ll (K \log K)^5$ supplied by Theorem 1, page 318 of N. Hegyvári and F. Hennecart [11], which heretofore was the best known upper bound for $s(K)$.

Let $M \geq 2$ be an integer and let us consider the natural colouring of the set of squares $\mathcal{S}$ induced by the congruence classes modulo M. Thus, let $c_1, c_2, \ldots, c_K$ be the quadratic residues modulo M and let us set

$$\mathcal{S}_i = \{x \in \mathcal{S} | x \equiv c_i \mod M\}. \tag{2.1}$$

Then $\mathcal{S} = \cup_{1 \leq i \leq K} \mathcal{S}_i$ is a colouring of $\mathcal{S}$ in K colours. Suppose now that n is a square-free multiple of M and that

$$n = x_1 + x_2 + \ldots + x_s, \tag{2.2}$$

with all the x_j belonging to $\mathcal{S}_i$, for some i with $1 \leq i \leq K$. Then c_i is necessarily coprime to M. For, if a prime p divides (c_i, M) then, since each x_j is a square congruent to c_i modulo M, p^2 must divide x_j for each j. Consequently, p^2 must divide n, which is absurd since n is square-free. On now reading (2.2) modulo M and recalling that n is a multiple of M we see that M divides s, since c_i is coprime to M. This means in particular that $s \geq M$. Since n can be taken to be an arbitrarily large square-free multiple of M, it follows that $s(K) \geq M$. By means of a simple argument that develops on this observation, Hegyvári and Hennecart show on page 319 of [11] that we have

$$s(K) \gg K \exp\left((\log 2 + o(1))\frac{\log K}{\log\log K}\right) \tag{2.3}$$

for all $K \geq 2$. Our upper bound for $s(K)$, given by Theorem 2.1.1, is off from this lower bound by a factor of order $K^{1+\epsilon}$. While we believe the lower bound above represents the true order of $s(K)$, we are unable to prove this at present.

The purpose of this chapter is to present a proof of Theorem 2.1.1 following [1]. Let us first summarize the method of Hegyvári and Hennecart for their upper bound for $s(K)$. They begin by remarking that if S is any subset of the squares in the interval $(N, 4N]$ satisfying $|S| \geq N^{\frac{1}{2}}/A$ for an integer N and a given real number $A \geq 1$, then we have the lower bound

$$|5S| \gg \frac{N}{A^5} \tag{2.4}$$

for the cardinality of the sum set $5S = S + S + S + S + S$. This is deduced in [11] as an immediate consequence of the well-known asymptotics for the number of representations of a given integer as the sum of five squares, provided by the circle method. Hegyvári and Hennecart then apply inequality (2.4) taking $A = K \log K$ within an elegant argument based on a finite addition theorem, also due to A. Sárközy, to arrive at their upper bound for $s(K)$.

Our proof of Theorem 2.1.1 follows the strategy of Hegyvári and Hennecart, except that in place of (2.4) we use the conclusion of theorem stated below. We recall from Section 1.1 that for any subset S of the integers $E_6(S)$ denotes the number of tuples

$(x_1, x_2, \ldots, x_{12})$ in S^{12} such that

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12}\ . \tag{2.5}$$

Theorem 2.1.2. *Let A and ϵ be real numbers with $A \geq 1$ and $\epsilon > 0$. Then there is an N_0, depending only on A and ϵ, such that for all $N \geq N_0$ and any subset S of the squares in the interval $[1, N]$ with $|S| \geq N^{\frac{1}{2}}/A$ we have*

$$E_6(S) \ll_\epsilon |S|^{10} A^\epsilon\ . \tag{2.6}$$

We prove Theorem 2.1.2 in Section 2.3 by means of the circle method and a large sieve inequality for polynomial amplitudes given by Proposition 2.2.2 of Section 2.2. We employ this inequality to estimate the contribution from the major arcs within our application of the circle method.

Theorem 2.1.1 is deduced from Theorem 2.1.2 through a classical application of the Cauchy-Schwarz inequality followed by an appeal to the argument from [11], involving the finite addition theorem of Sárközy. We provide the details of this deduction in Section 2.4. In Section 2.2 below we recall basic facts from the circle method as applied to the study of sums of squares, as also Sárközy's finite addition theorem, and present the aforementioned large sieve inequality.

2.2 Preliminaries

2.2.1 The Quadratic Weyl Bound

Let t be a real number and a, q integers with $(a,q)=1$, $q\neq 0$ and such that

$$\left|t-\frac{a}{q}\right|\leq\frac{1}{q^2}\ . \tag{2.7}$$

Then for any $P\geq 1$ we have from Theorem 1 on page 41 of [16] that

$$\left|\sum_{1\leq n\leq P}e(n^2t)\right|^2\ll\frac{P^2}{q}+P\log q+q\log q\ . \tag{2.8}$$

2.2.2 Majorisation on the Major Arcs

Let N, P and L be real numbers with $1\leq L^3\leq P\ll N^{\frac{1}{2}}$. Suppose further that t is a real number and a,q are integers with $0\leq a\leq q\leq L$ and $(a,q)=1$ satisfying

$$\left|t-\frac{a}{q}\right|\leq\frac{L}{N}. \tag{2.9}$$

Then we have that

$$\left|\sum_{1\leq n\leq P}e(n^2t)\right|\ll\frac{P}{q^{1/2}}\frac{1}{\left(1+P^2|t-\frac{a}{q}|\right)^{1/2}}\ . \tag{2.10}$$

Let us summarise the proof of this bound. Indeed, on writing $\beta=t-\frac{a}{q}$, the proof of Lemma 4.2 on page 17 of [8], in the particular case of the polynomial t^2, gives

$$\sum_{1\le n\le P} e(n^2t) = P\,\frac{S(q,a)}{q}\int_0^1 e(P^2\beta x^2)\,dx + O((|\beta|P^2+1)q)\ , \tag{2.11}$$

where $S(q,a)$ is the Gauss sum $\sum_{r\,\mathrm{mod}\,q} e(\frac{r^2a}{q})$. On using the well-known bound $|S(q,a)| \le 2q^{\frac{1}{2}}$ for the Gauss sum together with the van der Corput estimate $\int_0^1 e(\lambda x^2)dx \ll (1+|\lambda|)^{-\frac{1}{2}}$, valid for any real λ, we deduce from (2.11) that

$$\sum_{1\le n\le P} e(n^2t) \ll \frac{P}{q^{1/2}}\frac{1}{(1+P^2|\beta|)^{1/2}} + O(L^2)\ , \tag{2.12}$$

on taking account of (2.9) and the assumptions $P \ll N^{\frac{1}{2}}$, $1 \le q \le L$. Finally, since we have $P^2|\beta| \ll L$ and $\frac{P}{(Lq)^{1/2}} \ge \frac{P}{L} \ge L^2$ as well, we conclude that the first term on the right hand side of (2.12) majorises the second term, yielding (2.10).

2.2.3 A Large Sieve Inequality for Polynomial Amplitudes

Proposition 2.2.2 of this subsection is a variant, for small moduli, of the large sieve inequality for polynomial amplitudes given by Theorem 1 on page 429 of [17] and, in a special case, by O. Ramaré, Theorem 5.4, page 44 of [19]. These inequalities improve on the original such inequality given by L. Zhao, Theorem 2, page 166 of [27]. We shall prove Proposition 2.2.2 by combining the method of [17] with the "tensor power trick" as in the proof of Theorem 3, page 563 of Ramaré and Rusza [20]. See also the proof of Lemma 2, page 30 of J. Brüdern [6] and the proof of Lemma 4.28, page 307 of [5].

Let $P(T) = a_0T^k + a_1T^{k-1} + \ldots + a_k$ be a polynomial in $\mathbf{Z}[T]$ of degree $k \geq 1$. Also, for any integer $m \geq 1$, let $S(m)$ be the set of roots of $P(T)$ modulo m and let $\rho(m) = |S(m)|$ be the number of roots of $P(T)$ modulo m. Finally, let Q be any real number ≥ 1.

The proof of the large sieve inequality of Proposition 2.2.2 relies on an upper bound for $\sum_{1\leq m\leq Q} \frac{\rho(m)}{m}$ supplied by Proposition 1 on page 430 of [17]. For the convenience of the reader we reproduce this proposition, together with its proof, below.

Proposition 2.2.1. *For any integer $k \geq 1$, let $\theta(k)$ denote $k\binom{k+1}{2}$. Then for $\rho(m)$ and Q as above we have*

$$\sum_{1\leq m\leq Q} \frac{\rho(m)}{m} \ll_{a_0,k} (\log 2Q)^{\omega(a_0)+\theta(k)} , \tag{2.13}$$

where $\omega(a_0)$ is the number of prime divisors of $|a_0|$.

Note that the right hand side of (2.13) depends only on Q and the highest degree term of $P(T)$. Since $k \geq 1$, it is thus independent of the constant term of $P(T)$. This feature of the bound (2.13) is crucial to the proof of Proposition 2.2.2.

The proof of Proposition 2.2.1 depends on the following lemma.

Lemma 2.2.1. *Let a_0, k and ρ be as above and m be any integer ≥ 1. Further, let $a(m,k)$ denote the smallest integer $\geq \frac{m}{\binom{k+1}{2}}$. Then we have*

(i) $\frac{\rho(p^m)}{p^m} \leq \frac{k}{p}$ *for any prime p.*

(ii) $\frac{\rho(p^m)}{p^m} \leq \frac{(k+1)}{p^{a(m,k)}}$ *for a prime p that does not divide $|a_0|$, when $k \geq 2$.*

Proof.— Let p be a prime number and m, n be integers ≥ 1. When $m \geq n$ the image of $S(p^m)$ under the canonical surjection from $\mathbf{Z}/p^m\mathbf{Z}$ onto $\mathbf{Z}/p^n\mathbf{Z}$ is contained in $S(p^n)$. Therefore we have $\frac{\rho(p^m)}{p^m} \leq \frac{\rho(p^n)}{p^n}$ whenever $m \geq n \geq 1$. Since $\rho(p) \leq k$ it then follows that $\frac{\rho(p^m)}{p^m} \leq \frac{k}{p}$ for all $m \geq 1$, which is (i).

Let us prove (ii). We first show that any real interval of length $p^{a(m,k)}$ contains no more than $k+1$ integers x such that $P(x)$ is divisible by p^m. To verify this, it suffices to remark that if $x_1, x_2, \ldots, x_{k+1}$ are $k+1$ distinct integers such that $P(x_i)$ is divisible by p^m for each i, then we have $\sup_{(i,j)} |x_i - x_j| \geq p^{a(m,k)}$. Indeed, on recalling the well known identity for the Vandermonde determinant we have

$$a_0 \prod_{1\leq i<j\leq k+1} (x_i - x_j) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_{k+1} \\ \vdots & \vdots & & \vdots \\ x_1^{k-1} & x_2^{k-1} & \ldots & x_{k+1}^{k-1} \\ P(x_1) & P(x_2) & \ldots & P(x_{k+1}) \end{vmatrix}, \qquad (2.14)$$

after a simple manipulation of this determinant. Since the right hand side of (2.14) is divisible by p^m and p does not divide $|a_0|$, we see that p^m divides $\prod_{1\leq i<j\leq k+1}(x_i - x_j)$. Consequently,

$$\binom{k+1}{2} \sup_{i\neq j} v_p(x_i - x_j) \geq \sum_{1\leq i<j\leq k+1} v_p(x_i - x_j) \geq m\,. \qquad (2.15)$$

It follows from (2.15) that $\sup_{i\neq j} v_p(x_i - x_j) \geq a(m,k)$ and, because the x_i are distinct, that $\sup_{i,j} |x_i - x_j| \geq p^{a(m,k)}$, justifying our remark.

For each integer $m \geq 1$, the set $S(p^m)$ is in bijection with the subset of the integers x

in the interval $[0, p^m)$ such that $P(x)$ is divisible by p^m. On noting that $a(m,k) \leq m$ for all $m \geq 1$ when $k \geq 2$ and dividing the interval $[0, p^m)$ into subintervals of length $p^{a(m,k)}$, we then conclude that when p does not divide $|a_0|$ we have $\rho(p^m) \leq \frac{(k+1)p^m}{p^{a(m,k)}}$, for all $m \geq 1$ if $k \geq 2$, which gives (ii).

Proof of Proposition 2.2.1. — Since $m \mapsto \rho(m)$ is a multiplicative function, the sum $\sum_{1 \leq m \leq Q} \frac{\rho(m)}{m}$ is majorised by

$$\prod_{2 \leq p \leq Q} \left(\sum_{\substack{m \geq 0, \\ p^m \leq Q}} \frac{\rho(p^m)}{p^m} \right) \leq (2 \log 2Q)^{\omega(a_0)} \prod_{\substack{2 \leq p \leq Q, \\ (p, a_0) = 1.}} \left(\sum_{m \geq 0} \frac{\rho(p^m)}{p^m} \right), \tag{2.16}$$

where we have used the trivial bound $\rho(p^m) \leq p^m$ for the primes p dividing a_0.

We now estimate the product on the right hand side of (2.16). Thus suppose that $(p, a_0) = 1$. Then if $k \geq 2$ we have by Lemma 2.2.1 that

$$\sum_{m \geq 0} \frac{\rho(p^m)}{p^m} \leq 1 + \sum_{1 \leq m \leq \binom{k+1}{2}} \frac{k}{p} + \sum_{m > \binom{k+1}{2}} \frac{(k+1)}{p^{a(m,k)}} . \tag{2.17}$$

On dividing the sum over $m > \binom{k+1}{2}$ on the right hand side of the above relation into sums over congruence classes modulo $\binom{k+1}{2}$ and noting that $a(l + d\binom{k+1}{2}, k) = d + 1$, when d is any integer and l an integer satisfying $0 < l \leq \binom{k+1}{2}$, we deduce that

$$\sum_{m \geq 0} \frac{\rho(p^m)}{p^m} \leq 1 + \frac{\theta(k)}{p} + (k+1)\binom{k+1}{2} \sum_{d \geq 1} \frac{1}{p^{d+1}} \tag{2.18}$$

when $k \geq 2$. When $k = 1$ we have $\rho(p^m) = 1$ for all $m \geq 1$, so that the bound (2.18) remains valid in this case as well.

The proposition now follows on substituting (2.18) into the right hand side of (2.16), dropping the condition $(a_0, p) = 1$ in the resulting product and recalling that $\prod_{2\leq p\leq Q}(1+\frac{a}{p}+b\sum_{m\geq 2}\frac{1}{p^m}) \ll_{a,b} (\log 2Q)^a$, for any real numbers a and $b \geq 0$.

Hereafter we set

$$\theta(a_0, k) = \omega(a_0) + \theta(k) \tag{2.19}$$

for any polynomial $P(T) = a_0T^k + \ldots + a_k$ in $\mathbf{Z}[T]$ of degree k. Also, we write $\tau(n)$ for the number of divisors of an integer n and write $a \operatorname{mod}^* q$ to mean a is an invertible residue class modulo q.

Proposition 2.2.2. *Let $p \geq 1$ be an integer and N, Q, ℓ real numbers with $\ell > 1$ and $1 \leq Q^p \leq N$. Further, let $P(T) = a_0T^k + \ldots + a_k$ be a polynomial in $\mathbf{Z}[T]$ of degree $k \geq 1$. Then for any set X of integers in a real interval of length N and any complex numbers $c(x)$, $x \in X$, with $|c(x)| \leq 1$ for all x, we have the inequality*

$$\sum_{1\leq q\leq Q}\frac{1}{q^\ell}\sum_{a \operatorname{mod}^* q}\left|\sum_{x\in X}c(x)e\left(\frac{P(x)a}{q}\right)\right|^2 \ll_{a_0,k,p,\ell} |X|^2\left(\frac{N}{|X|}\right)^{1/p}(\log 2Q)^{\theta(a_0,k)/p}. \tag{2.20}$$

It is easily seen from the proof below that a version of proposition above for the case $\ell \leq 1$ is obtained by multiplying the right hand side of (2.20) with $Q^{1-\ell}\tau^*(Q^p)$ for such ℓ, where $\tau^*(u)$ denotes $\max_{1\leq n\leq u}\tau(n)$, for any real number $u \geq 1$.

Proof.— For any integers n and q, with $q \neq 0$, we set $\delta_q(n) = 1$ when q divides n and to be 0 otherwise, so that we have

$$\sum_{a \bmod q} e\left(\frac{na}{q}\right) = q\, \delta_q(n)\ . \tag{2.21}$$

Then on extending the summation over $\bmod^* q$ to all residue classes modulo q, for each q in $[1, Q]$, we see that the left hand side of (2.20) does not exceed

$$\sum_{1 \leq q \leq Q} \frac{1}{q^{\ell}} \sum_{a \bmod q} \left| \sum_{x \in X} c(x) e\left(\frac{P(x)a}{q}\right) \right|^2 \leq \sum_{(x,x') \in X^2} \sum_{1 \leq q \leq Q} q^{1-\ell} \delta_q(P(x) - P(x'))\ , \tag{2.22}$$

where the inequality results on opening the square on the left hand side, using (2.21) together with $|c(x)| \leq 1$, and interchanging summations.

Let $p \geq 1$ be an integer. Then an application of Hölder's inequality shows that right hand side of (2.22) does not exceed

$$|X|^{2-\frac{2}{p}} \left(\sum_{(x,x') \in X^2} \left(\sum_{1 \leq q \leq Q} q^{1-\ell} \delta_q(P(x) - P(x')) \right)^p \right)^{\frac{1}{p}}\ . \tag{2.23}$$

On expanding the summand in the sum over (x, x') in the above expression, we see that (2.23) is the same as

$$|X|^{2-\frac{2}{p}} \left(\sum_{(x,x') \in X^2} \sum_{\substack{(q_1, \ldots, q_p), \\ 1 \leq q_i \leq Q.}} \prod_{1 \leq i \leq p} {q_i}^{1-\ell} \delta_{q_i}(P(x) - P(x')) \right)^{\frac{1}{p}}\ . \tag{2.24}$$

If for a p-tuple of integers $\mathbf{q} = (q_1, q_2, \ldots, q_p)$, we write $[\mathbf{q}]$ to denote the least common multiple (lcm) of the integers $q_1, q_2, \ldots, q_p$ then we have

$$\prod_{1\leq i\leq p} \delta_{q_i}(P(x) - P(x')) \;=\; \delta_{[\mathbf{q}]}(P(x) - P(x')) \tag{2.25}$$

for any integers x and x'. Also, we have $(q_1 q_2 \ldots q_p)^{1-\ell} \leq [\mathbf{q}]^{1-\ell}$, since $\ell > 1$. Thus on writing $g(m)$, for any integer m, to denote the number of p-tuples $\mathbf{q} = (q_1, \ldots, q_p)$, with each q_i an integer in $[1, Q]$, such that $m = [\mathbf{q}]$ we see that the sum over $(q_1, \ldots, q_p)$ in (2.24) does not exceed

$$\sum_{1\leq m\leq Q^p} g(m) m^{1-l} \delta_m(P(x) - P(x')) \tag{2.26}$$

for any integers x and x'. If $m = [\mathbf{q}]$ for a tuple $\mathbf{q} = (q_1, \ldots, q_p)$, then q_i divides m for each i so that we have $g(m) \leq \tau(m)^p$, for any integer $m \geq 1$. Since $\ell > 1$ we have $\tau(m)^p \leq C(p, \ell) m^{\ell-1}$, for some real number $C(p, \ell)$ depending only on p and ℓ, by a classical bound for $\tau(m)$. We then conclude that (2.24) does not exceed

$$|X|^{2-\frac{2}{p}} \left(C(p, \ell) \sum_{x'\in X} \sum_{1\leq m\leq Q^p} \sum_{x\in X} \delta_m(P(x) - P(x')) \right)^{\frac{1}{p}}, \tag{2.27}$$

after an interchange of summations.

Given $x' \in X$ and an integer m, we have $\delta_m(P(x) - P(x')) = 1$ if and only if $x \equiv \alpha \bmod m$ where α is a root of the polynomial $P(T) - P(x')$ in $\mathbf{Z}/m\mathbf{Z}$. Therefore if $\rho(x', m)$ denotes the number of such roots then we have

$$\sum_{x'\in X}\sum_{1\leq m\leq Q^p}\sum_{x\in X}\delta_m(P(x)-P(x'))\leq 2N\sum_{x'\in X}\sum_{1\leq m\leq Q^p}\frac{\rho(x',m)}{m}\ , \tag{2.28}$$

since the number of $x\in X$ lying in a given class $\alpha \bmod m$ does not exceed $\frac{N}{m}+1\leq\frac{2N}{m}$, taking account of the hypotheses that X is contained in an interval of length N and $m\leq Q^p\leq N$. Further, we have for all x' in X the bound

$$\sum_{1\leq m\leq Q^p}\frac{\rho(x',m)}{m}\ll_{a_0,k}(\log 2Q^p)^{\theta(a_0,k)}\ \ll_{a_0,k,p}(\log 2Q)^{\theta(a_0,k)}\ , \tag{2.29}$$

from Proposition 2.2.1 and (2.19). On combining (2.28) with (2.29) we then obtain

$$\sum_{x'\in X}\sum_{1\leq m\leq Q^p}\sum_{x\in X}\delta_m(P(x)-P(x'))\ll_{a_0,k,p}N|X|(\log 2Q)^{\theta(c_0,k)}\ . \tag{2.30}$$

Substituting this bound into (2.27), which, by what we have seen, is an upper bound for the right hand side of (2.22), we conclude after an obvious rearrangement of terms that we indeed have the inequality (2.20), completing the proof of Proposition 2.2.2.

The trivial upper bound for the left hand side of (2.20) is $|X|^2\sum_{1\leq q\leq Q}q^{1-\ell}$, up to a constant. Thus Proposition 2.2.2 gives a useful bound only when $\ell<2$. Also, if $c(x)=1$ for all x in X, the term corresponding to $q=1$ on the left hand side of (2.22) reduces to $|X|^2$, which, by positivity, is therefore a lower bound for the left hand side of (2.22) in this case. Thus by taking p to be large we see that there are choices of the coefficients $c(x)$ such that (2.22) is close to optimal when $|X|$ is a "dense" subset of $[1, N]$, that is, $N/|X|$ is bounded above independent of N, and N is suitably large. This is the case, for example, in the following corollary, which gives the form in which we will use Proposition 2.2.2.

Corollary 2.2.3. *Let Q, A and ϵ be real numbers with $Q, A \geq 1$ and $\epsilon > 0$. Then for all N sufficiently large, depending only on Q and ϵ, and any subset S of the squares in $[1, N]$ satisfying $|S| \geq N^{\frac{1}{2}}/A$ we have*

$$\sum_{1\leq q\leq Q} \frac{1}{q^\ell} \sum_{a \bmod^* q} \left| \sum_{s\in S} c(s) e\left(\frac{sa}{q}\right)\right|^2 \ll_{\epsilon,\ell} |S|^2 A^\epsilon \log(2Q) \tag{2.31}$$

where $c(s)$, with s varying over S, are any complex numbers with $|c(s)| \leq 1$ and ℓ is any real number with $\ell > 1$.

Proof.— Applying Proposition 2.2.2 with $P(T) = T^2$, $X \subset [1, N^{\frac{1}{2}}]$ taken to be the set of $\sqrt{s}$, with s varying over S, and p any integer that is larger than $\frac{1}{\epsilon}$ and $\theta(1, 2)$, we see that the corollary holds when $N^{\frac{1}{2}} \geq Q^p$.

2.2.4 Sárközy's Finite Addition Theorem

Let N and k be integers ≥ 1 and A be subset of the integers in $[1, N]$ such that $|A| > \frac{N}{k} + 1$. Then there exist integers d, l and m with $1 \leq d \leq k - 1$ and $1 \leq l < 118k$ and such that $\{(m+1)d, (m+2)d, \dots, (m+N)d\}$ is contained in the sum set lA.

This is Theorem 1, page 115 of [22]. V. Lev [13], Theorem 2, page 128, gives an optimal version of this result. However, the version stated above suffices for our purpose.

2.3 Proof of Theorem 2.1.2

With S and $E_6(S)$ as in the statement of the theorem, it is evident that $E_6(S)$ does not exceed the number $E_6^*(S)$ of tuples $(y_1, y_2, x_1, \ldots, x_{10})$ satisfying the relation

$$y_1^2 + \sum_{1\le i\le 5} x_i = y_2^2 + \sum_{6\le i\le 10} x_i \tag{2.32}$$

with each x_i in S and y_1, y_2 integers in the interval $[1, N^{\frac{1}{2}}]$. Thus, on writing $\widehat{S}(t)$ to denote $\sum_{s\in S} e(st)$ and $\phi(t)$ to denote $\sum_{1\le n\le N^{\frac{1}{2}}} e(n^2 t)$ we have that

$$E_6(S) \le E_6^*(S) = \int_0^1 |\widehat{S}(t)|^{10}\, |\phi(t)|^2 \, dt \ , \tag{2.33}$$

where the equality follows by orthogonality of the functions $t \mapsto e(nt)$ on $[0, 1)$.

We shall presently apply the circle method to estimate the integral in (2.33). We begin by recording a simple remark on $E_5(S)$, which, we recall, means the number of $(x_1, x_2, \ldots, x_{10})$ in S^{10} satisfying the relation

$$x_1 + x_2 + x_3 + x_4 + x_5 = x_6 + x_7 + x_8 + x_9 + x_{10} \ . \tag{2.34}$$

We have that

$$\int_0^1 |\widehat{S}(t)|^{10}\,dt = E_5(S) \ll |S|^8 A^3 \ , \tag{2.35}$$

where the equality is again a consequence of orthogonality. As for the inequality in (2.35), if $R_5(n)$ denotes the number of representations of an integer n as a sum of five elements of S we have

$$E_5(S) = \sum_{n\geq 1} R_5^2(n) = \sum_{1\leq n\leq 5N} R_5^2(n) \ll N^{\frac{3}{2}} \sum_{n\geq 1} R_5(n) = |S|^5 N^{\frac{3}{2}} \ , \tag{2.36}$$

since we have $R_5(n) = 0$ when $n > 5N$ and $R_5(n) \leq r_5(n)$, the number of representations of n as a sum of five squares of natural numbers, and have the bound $r_5(n) \ll n^{\frac{3}{2}}$, by a standard application of the circle method. Since $|S| \geq N^{\frac{1}{2}}/A$, (2.36) implies the inequality in (2.35).

Let us now set $L = (\log N)^2$, $Q = A^5$, $M = \frac{N}{L}$ and $P = N^{\frac{1}{2}}$. Also, for integers a and q with

$$0 \leq a \leq q \leq Q \ \text{ and } \ (a,q) = 1 \ , \tag{2.37}$$

we shall call the interval $[\frac{a}{q} - \frac{1}{M}, \frac{a}{q} + \frac{1}{M})$ the major arc $\mathfrak{M}(\frac{a}{q})$. It is easily verified that distinct major arcs are in fact disjoint when N is sufficiently large, depending only on A. Also, we shall denote by $\mathfrak{M}$ the union of the family of major arcs $\mathfrak{M}(\frac{a}{q})$ and by $\mathfrak{m}$ the complement in $[0,1)$ of $\mathfrak{M}$.

2.3.1 The Minor Arc Contribution

Let us estimate the contribution to the integral in (2.33) from t in $\mathfrak{m}$. For any such real number t, we have from Dirichlet's theorem that there exists a rational number $\frac{a}{q}$ satisfying $|t - \frac{a}{q}| \leq \frac{1}{qM}$ with $(a, q) = 1$ and $1 \leq q \leq M$. Since $t \in \mathfrak{m}$ is contained in $[\frac{1}{M}, 1 - \frac{1}{M})$, it follows that $\frac{a}{q}$ is in $[0, 1]$ so that we have $0 \leq a \leq q$. Since, however, t does not belong to $\mathfrak{M}$, we must necessarily have $Q < q$ on account of the conditions (2.79) defining $\mathfrak{M}$. Since $q^2 \leq qM$, we conclude that for each t in $\mathfrak{m}$ there are integers a and $q \neq 0$ with $(a, q) = 1$ satisfying

$$\left| t - \frac{a}{q} \right| \leq \frac{1}{q^2} \quad \text{and} \quad Q < q \leq M. \tag{2.38}$$

Applying the bound given by (2.8) with $\frac{a}{q}$ as in (2.38) above, and recalling the definitions of Q, M and P, we now obtain

$$|\phi(t)|^2 \ll \frac{N}{Q} + N^{\frac{1}{2}} \log M + M \log M \ll \frac{N}{A^5} \tag{2.39}$$

for all $t \in \mathfrak{m}$, when N is sufficiently large, depending only on A. Thus, for all such N, we have

$$\int_{\mathfrak{m}} |\widehat{S}(t)|^{10} \, |\phi(t)|^2 \, dt \ll \frac{N}{A^5} \int_0^1 |\widehat{S}(t)|^{10} \, dt \; \ll |S|^{10}, \tag{2.40}$$

on using (2.35) together with $|S| \geq N^{\frac{1}{2}}/A$.

2.3.2 The Major Arc Contribution

Let us now pass to the contribution to the integral in (2.33) from t in the complement of $\mathfrak{m}$ in $[0,1)$. For each such t there are integers a and q satisfying (2.79) such that t belongs to $\mathfrak{M}(\frac{a}{q})$. Therefore, on introducing, for the sake of brevity, the notation

$$\mathcal{E}(f) = \sum_{1\le q\le Q} \sum_{a \bmod^* q} \int_{\mathfrak{M}(\frac{a}{q})} f(t)\, dt \tag{2.41}$$

for any continuous function f on $[0,1)$, we see that the contribution to the integral in (2.33) from t in $[0,1)\setminus\mathfrak{m}$ does not exceed $\mathcal{E}(|\widehat{S}(t)|^{10}|\phi(t)|^2)$.

Let p and ℓ be positive real numbers such that $\frac{1}{p}+\frac{1}{\ell}=1$. Then an application of Hölder's inequality shows that

$$\mathcal{E}(|\widehat{S}(t)|^{10}|\phi(t)|^2) \le \mathcal{E}(|\widehat{S}(t)|^{8p})^{\frac{1}{p}}\, \mathcal{E}(|\widehat{S}(t)|^{2\ell}|\phi(t)|^{2\ell})^{\frac{1}{\ell}}\ . \tag{2.42}$$

Let p be large enough so that $8p \ge 10$. Since distinct major arcs are disjoint and since $\mathfrak{M}$ is contained in $[-1,2]$, we have, on remarking that $\widehat{S}(t)$ is periodic with period 1, that

$$\mathcal{E}(|\widehat{S}(t)|^{8p}) \le \int_{-1}^{2} |\widehat{S}(t)|^{8p}\, dt = 3\int_0^1 |\widehat{S}(t)|^{8p}\, dt \ll |S|^{8p-2}A^3, \tag{2.43}$$

where we have used $|\widehat{S}(t)|^{8p} \le |S|^{8p-10}|\widehat{S}(t)|^{10}$ and (2.35). It now follows that

$$\mathcal{E}(|\widehat{S}(t)|^{8p})^{\frac{1}{p}} \ll (|S|^{8p-2}A^3)^{\frac{1}{p}} \ll |S|^8 A^{\frac{5}{p}} N^{-\frac{1}{p}}\ , \tag{2.44}$$

since $|S| \geq N^{\frac{1}{2}}/A$.

We now dispose of the second term on the right hand side of (2.42). Since we have $A^5 \leq L$ and $L^3 \leq P \ll N^{\frac{1}{2}}$ for all large enough N, depending only on A, we may apply (2.10) and obtain for all t in any given major arc $\mathfrak{M}(\frac{a}{q})$ the bound

$$|\phi(t)|^{2\ell} \ll \frac{N^\ell}{q^\ell \left(1 + N|t - \frac{a}{q}|\right)^\ell} \, . \tag{2.45}$$

Since $\ell > 1$ we have $|\widehat{S}(t)|^{2\ell} \leq |S|^{2\ell-2}|\widehat{S}(t)|^2$, for all t. Consequently, we deduce that

$$\mathcal{E}(|\widehat{S}(t)|^{2\ell}|\phi(t)|^{2\ell}) \ll |S|^{2\ell-2}N^\ell \sum_{1 \leq q \leq Q} \sum_{a \bmod^* q} \int_{\mathfrak{M}(\frac{a}{q})} \frac{|\widehat{S}(t)|^2}{q^\ell \left(1 + N|t - \frac{a}{q}|\right)^\ell} \, dt. \tag{2.46}$$

On making the change of variables $N(t - \frac{a}{q}) = \beta$ in each of the integrals in (2.46) and interchanging the resulting integral with the summations, we then see that the right hand side of (2.46) is the same as

$$|S|^{2\ell-2}N^{\ell-1} \int_{|\beta| \leq L} \frac{1}{(1+|\beta|)^\ell} \sum_{1 \leq q \leq Q} \frac{1}{q^\ell} \sum_{a \bmod^* q} \left| \sum_{s \in S} e\left(\frac{\beta s}{N}\right) e\left(\frac{sa}{q}\right) \right|^2 d\beta \, . \tag{2.47}$$

Let $\epsilon > 0$ be fixed. Then, for each β, and on recalling that $Q = A^5$, we bound the sum over q in the integrand in (2.47) by means of Corollary 2.2.3 with $c(s) = e\left(\frac{\beta s}{N}\right)$ for all s in S. This allows us to conclude that there is a $C(\epsilon)$, depending only on ϵ, such that, for all large enough N, depending only on A and ϵ, (2.47) does not exceed

$$C(\epsilon)|S|^{2\ell-2}N^{\ell-1}|S|^2A^{\frac{\epsilon}{2}}\int_{\mathbf{R}}\frac{d\beta}{(1+|\beta|)^{\ell}} \ll_{\epsilon,\ell} |S|^{2\ell}N^{\ell-1}A^{\frac{\epsilon\ell}{2}}\ , \tag{2.48}$$

since $\ell > 1$ and $A \geq 1$. Since (2.47) is itself an upper bound for $\mathcal{E}(|\widehat{S}(t)|^{2\ell}|\phi(t)|^{2\ell})$ we now obtain from (2.42) and (2.44) that

$$\mathcal{E}(|\widehat{S}(t)|^{10}|\phi(t)|^2) \ll_{\epsilon,p} |S|^8A^{\frac{5}{p}}N^{-\frac{1}{p}}(|S|^{2\ell}N^{l-1}A^{\frac{\epsilon\ell}{2}})^{\frac{1}{\ell}} \ll_{\epsilon,p} |S|^{10}A^{\frac{5}{p}+\frac{\epsilon}{2}}, \tag{2.49}$$

where we have used $\frac{1}{p}+\frac{1}{\ell}=1$. Now we choose p to be larger than both $\frac{10}{\epsilon}$ and $\frac{10}{8}$, so that we have $8p \geq 10$ and $\frac{5}{p} \leq \frac{\epsilon}{2}$, and conclude that

$$\mathcal{E}(|\widehat{S}(t)|^{10}|\phi(t)|^2) \ll_{\epsilon} |S|^{10}A^{\epsilon}, \tag{2.50}$$

when N is sufficiently large depending only on A and ϵ. Finally, on recalling (2.40) and that the left hand side of (2.50) majorises the contribution to the integral in (2.33) from $[0,1)\setminus\mathfrak{m}$, we deduce (2.3) from (2.33) thus obtaining Theorem 2.1.2.

2.4 Proof of Theorem 2.1.1

We shall now deduce Theorem 2.1.1 from Theorem 2.1.2 by a slight variant of an argument in [11]. Thus, let $\epsilon > 0$ be fixed and let $\cup_{1\leq i\leq K}\mathcal{S}_i$ be a partition of the set $\mathcal{S}$ of the squares into K subsets , for a given integer $K \geq 1$. Since Theorem 2.2.2

with $K = 1$ is certainly covered by Lagrange's theorem, we assume $K \geq 2$.

2.4.1 The Set S

Let $B \geq 4$ be a real number and $\mathcal{U}$ be the subset of integers that are coprime to every integer in $[1, B]$. Also, for any integer $N \geq 1$, let $\mathcal{U}(N)$ denote $\mathcal{U} \cap (N^{\frac{1}{2}}, 2N^{\frac{1}{2}}]$. Then the principle of inclusion and exclusion together with the effective Mertens' formula given by (3.27), page 70 of [21] shows that

$$|\mathcal{U}(N)| \geq N^{\frac{1}{2}} \prod_{p \leq B} \left(1 - \frac{1}{p}\right) - 2^B \geq \frac{N^{\frac{1}{2}}}{4 \log B} - 2^B \geq \frac{N^{\frac{1}{2}}}{8 \log B}\,, \tag{2.51}$$

when N sufficiently large, depending on B. For each such N there is an i, $1 \leq i \leq K$, such that $S_i = \mathcal{S}_i \cap (N, 4N]$ contains at least $|\mathcal{U}(N)|/K$ of the squares of the elements of $\mathcal{U}(N)$. In particular, for such an i we have that $S_i \subset [1, 4N]$ and

$$|S_i| \;\geq\; \frac{|\mathcal{U}(N)|}{K} \;\geq\; \frac{(4N)^{\frac{1}{2}}}{A}\,, \tag{2.52}$$

where $A = 16K \log B \geq 1$. When N is sufficiently large depending on ϵ, B and K, we then conclude by means of Theorem 2.1.2 that there is an i, $1 \leq i \leq K$, such that S_i satisfies $E_6(S_i) \ll_\epsilon |S_i|^{10}(K \log B)^\epsilon$. Since we have

$$E_6(S_i)\,|6S_i| \;\geq\; |S_i|^{12} \gg \frac{|S_i|^{10}\,N}{(K \log B)^2}\,, \tag{2.53}$$

where the first inequality is a classical consequence of the Cauchy-Schwarz inequality, it now follows that there exists a real number $C(\epsilon)$, depending on ϵ alone, such that for each sufficiently large N, depending on ϵ, B and K, there is an i so that S_i satisfies the inequality

$$|6S_i| \geq \frac{18N}{C(\epsilon)(K\log B)^{2+\epsilon}} > \frac{18N}{k} + 1 \ , \tag{2.54}$$

for an integer k with $k \leq 2C(\epsilon)(K\log B)^{2+\epsilon}$. Also, since S_i contains squares of elements $\mathcal{U}(N)$, there is *at least one integer* in S_i that is coprime to every integer in $[1, B]$.

2.4.2 Endgame

With S_i as above, $6S_i$ is contained in $(6N, 24N]$, we may apply Sárközy's finite addition theorem, recalled in Subsection 2.2.4 , to the set $6S_i - 6N$, which is contained in $[1, 18N]$. Taking (2.54) into account, we deduce that there are integers l, d and m with $1 \leq l < 118k$ and $d < k$ such that $6lS_i$ contains the arithmetical progression

$$\mathcal{A} = 6lN + \{(m+1)d, (m+2)d, \ldots, (m+18N)d\} \ , \tag{2.55}$$

with $18N$ terms and to the modulus d.

Let $b \geq 4 + 2\epsilon$ be a real number such that $2C(\epsilon)b^{2+\epsilon} \leq 2^{b-4-2\epsilon}$. Then on setting $B = K^b$ and recalling that $K \geq 2$, we infer from the inequalities

$$2C(\epsilon)(K\log B)^{2+\epsilon} = 2C(\epsilon)b^{2+\epsilon}(K\log K)^{2+\epsilon} \leq 2^{b-4-2\epsilon}K^{4+2\epsilon} \leq B \tag{2.56}$$

that $k \leq B$. Since we have $d < k$ for the modulus d of $\mathcal{A}$, and since S_i contains an integer that is coprime to every integer in $[1, B]$, it follows from (2.56) that, on the one hand, S_i contains an integer m coprime to d. On the other hand, since $S_i \subset (N, 4N]$, m is at most $4N$ and therefore the number of terms in $\mathcal{A}$ exceeds m. We then conclude that the arithmetical progression $\mathcal{A}$ contains a complete system of residue classes modulo m. In particular, every integer n can be written as $n = a+rm$ with a in $\mathcal{A} \subset 6lS_i$ and r an integer.

Let us write $I(N)$ to denote the interval $(2832kN, 2833kN]$ and apply the conclusion of the preceding paragraph to the integers in this interval. To this end, let us note that since $\mathcal{A}$ is contained in $6lS_i$, which is a subset of $(6lN, 24lN]$, we have $0 \leq a \leq 2832kN$ for each a in $\mathcal{A}$, on recalling that $l < 118k$. Thus if n in $I(N)$ is written as $n = a + rm$ we have $0 \leq rm \leq 2833kN$, which gives $0 \leq r \leq 2833k$ since $N < m$. Therefore n can be expressed as sum of no more than $6l + 2833k < 4000k$ terms of S_i.

In conclusion, we have verified that for each large enough N, depending on ϵ and K alone, there exists an i, with $1 \leq i \leq K$, such that every integer in the interval $I(N)$ can be written as a sum of no more than $4000k$ squares all belonging to S_i. In other words, for all sufficiently large N, every integer in the interval $I(N)$ is the sum of no more than $4000k$ squares, all of the same colour. Since

$$4000k \leq 8000C(\epsilon)(K \log B)^{2+\epsilon} \ll_\epsilon K^{2+\epsilon} \tag{2.57}$$

and since the interval $I(N)$ meets $I(N+1)$ for all sufficiently large N, so that the union of the intervals $I(N)$, over all such N, contains all sufficiently large integers, we obtain Theorem 2.1.1.

2.5 Bourgain's Method

In [5] J. Bourgain shows that for any integer $N \geq 1$ and complex numbers $a_1, a_2, \ldots, a_N$ we have

$$\int_0^1 \left| \sum_{1 \leq n \leq N} a_n e(n^2 t) \right|^p dt \ll_p N^{\frac{p}{2}-2} \|a\|_2^p \, , \tag{2.58}$$

where as usual $\|a\|_2 = (\sum_{1 \leq n \leq N} |a_i|^2)^{\frac{1}{2}}$. This follows on combining the definition of K_p given by (1.8) on page 293 of [5] with (4.1) on page 304 of [5]. Theorem 2.1.2 is a consequence of a version of the above inequality where the L^2 norm $\|a\|_2$ is replaced with the L^∞ norm of $\{a_i\}_{1 \leq i \leq N}$. Such a version may be obtained by making suitable modifications to the contents of Section 4, pages 304-307 of [5]. We show here how this may be done by providing the full details for the following special case of the required version.

Theorem 2.5.1. *For any real numbers p, A and ϵ with $p > 4$, $A \geq 1$ and $\epsilon > 0$ there is an integer N_0 such that for all $N \geq N_0$ and any subset S of the squares in the interval $[1, N]$ with $|S| \geq N^{\frac{1}{2}}/A$ we have*

$$\int_0^1 |\widehat{S}(t)|^p \, dt \ll_{\epsilon,p} |S|^{p-2} A^{\epsilon} \, . \tag{2.59}$$

Here $\widehat{S}(t)$ denotes $\sum_{s \in S} e(st)$, as before. Since the left hand side of (2.59) is $E_6(S)$

when $p = 12$, we recover Theorem 2.1.2 from the above theorem, which is in fact a much deeper assertion. Moreover, the statement of Theorem 2.5.1 is optimal with respect to p in the sense that (2.59) is false in general when $p = 4$. In fact, we have

$$\int_0^1 |\widehat{S}(t)|^4 \, dt \;=\; \int_0^1 \left| \sum_{1 \le n \le 2N} r_S(n) e(nt) \right|^2 dt \;=\; \sum_{1 \le n \le 2N} r_S^2(n) \;, \tag{2.60}$$

where $r_S(n)$ is the number of pairs $(x_1, x_2) \in S^2$ such that $x_1 + x_2 = n$, for any integer n. When S is taken to be the set of *all* squares in $[1, N]$, so that $A = 1$, the third term in (2.60) is known to be asymptotic to $CN \log N$, for some $C > 0$, as $N \to +\infty$, whereas the right hand side of (2.59) is $|S|^2 = N$. This implies that (2.59) does not hold for $p = 4$ in this case.

We shall hereafter write $\mathbf{T}$ for $\mathbf{R}/\mathbf{Z}$. Also, $\| \, \|$ and μ shall denote, respectively, the usual metric and measure on $\mathbf{T}$. Then we have

$$\int_0^1 |\widehat{S}(t)|^p \, dt = \int_{\mathbf{T}} |\widehat{S}(t)|^p \, d\mu \, . \tag{2.61}$$

Proof of Theorem 2.5.1.— Let $p > 4$ be fixed. Since $A \ge 1$, it is enough to prove (2.59) for all sufficiently small $\epsilon > 0$. We may therefore suppose that ϵ is a fixed real number satisfying $0 < \epsilon < \min(1, p - 4)$. On remarking that

$$\int_{\mathbf{T}} |\widehat{S}(t)|^p \, d\mu = p \, |S|^p \int_0^1 u^{p-1} \mu \left(t \in \mathbf{T} \mid |\widehat{S}(t)| \ge u|S| \right) \, du, \tag{2.62}$$

where μ is the usual measure on $[0, 1)$, we then see that (2.59) is a consequence of the distributional inequality

$$\mu\left(t\in\mathbf{T}\mid|\widehat{S}(t)|\geq\delta|S|\right)\ll_\epsilon\frac{A^\epsilon}{\delta^{4+\epsilon}|S|^2}\,, \tag{2.63}$$

for all δ in $(0,1)$.

To prove (2.63) we begin by observing that for all small enough δ this inequality is easily deduced from (2.60). In effect, if $r(n)$ is the number of representations of n as a sum of two squares then we have that $r_S(n)\leq r(n)\leq\tau(n)$, for any integer n. By means of a classical bound for $\tau(n)$ we then conclude that

$$\sum_{1\leq n\leq 2N}r_S^2(n)\ \leq\ \max_{1\leq n\leq 2N}\tau(n)\sum_{1\leq n\leq 2N}r_S(n)\ \leq\ |S|^2\exp\left(\frac{C\log N}{\log\log N}\right), \tag{2.64}$$

for some real number $C>0$. On the other hand, we have

$$\delta^4|S|^4\mu\left(t\in\mathbf{T}\mid|\widehat{S}(t)|\geq\delta|S|\right)\ \leq\ \int_0^1|\widehat{S}(t)|^4\,dt\,. \tag{2.65}$$

From (2.60), (2.64) and (2.65) we then conclude that

$$\mu\left(t\in\mathbf{T}\mid|\widehat{S}(t)|\geq\delta|S|\right)\ \leq\ \frac{1}{\delta^4|S|^2}\exp\left(\frac{C\log N}{\log\log N}\right), \tag{2.66}$$

which implies (2.63) when $0<\delta\leq\exp\left(-\frac{C\log N}{\epsilon\log\log N}\right)$, since $A\geq 1$. It therefore suffices to prove (2.63) supposing that

$$\exp\left(-\frac{C\log N}{\epsilon\log\log N}\right)<\delta\,. \tag{2.67}$$

The proof of (2.63) for δ satisfying (2.67) depends on the following simple yet crucial remark, which is in fact valid for all δ in $(0,1)$.

Let R be the *largest* integer such that there is a sequence of points $t_1, t_2, \dots, t_R$ in $\mathbf{T}$ satisfying the conditions

$$\|t_r - t_s\| \geq \frac{1}{N} \text{ when } r \neq s \tag{2.68}$$

and

$$|\widehat{S}(t_r)| \geq \delta |S| \text{ for each } r, \tag{2.69}$$

with $1 \leq r, s \leq R$. Then to prove (2.63) it suffices to show that for all $\ell > 2$ and $\kappa > 0$ we have the estimate

$$R \ll_{\kappa,\ell} \frac{A^{2\kappa} N^{\frac{\ell}{2}}}{\delta^{2\ell+5\kappa} |S|^{\ell}} \, . \tag{2.70}$$

To verify this remark we need only note that by the maximality of R, every t in $\mathbf{T}$ such that $|\widehat{S}(t)| \geq \delta |S|$ necessarily satisfies $\|t - t_r\| < \frac{1}{N}$ for some r. This means that

$$\mu\left(t \in \mathbf{T} \mid |\widehat{S}(t)| \geq \delta |S|\right) \leq \frac{2R}{N} \, . \tag{2.71}$$

Now (2.70) with $\kappa > 0$ and $l > 2$ satisfying $l - 2 + 2\kappa \leq \epsilon$ and $2\ell + 5\kappa \leq 4 + \epsilon$ gives

$$\frac{2R}{N} \ll_{\kappa,\ell} \frac{1}{\delta^{2\ell+5\kappa}|S|^2}\left(\frac{A^{2\kappa}|S|^2N^{\frac{\ell}{2}}}{N|S|^\ell}\right) \ll_\epsilon \frac{A^\epsilon}{\delta^{4+\epsilon}|S|^2}\ , \tag{2.72}$$

since $\left(\frac{N^{\frac{1}{2}}}{|S|}\right)^{l-2} \leq A^{l-2}$ for $l > 2$, from which and (2.71) the bound (2.63) follows.

We commence the proof of (2.70) by an application of duality. Thus let complex numbers c_r be defined by $|\widehat{S}(t_r)| = c_r\widehat{S}(t_r)$, for each r. Then on adding the inequalities (2.69) for the various r and recalling that $\widehat{S}(t_r) = \sum_{s\in S} e(st_r)$ we get

$$\sum_{s\in S}\sum_{1\leq r\leq R} c_r e(st_r) = \sum_{1\leq r\leq R}\sum_{s\in S} c_r e(st_r) = \sum_{1\leq r\leq R} |\widehat{S}(t_r)| \geq \delta|S|R\ . \tag{2.73}$$

By an application of the Cauchy-Schwarz inequality we then have that

$$\sum_{s\in S}\sum_{1\leq r\leq R} c_r e(st_r) \leq |S|^{\frac{1}{2}}\left(\sum_{s\in S}\left|\sum_{1\leq r\leq R} c_r e(st_r)\right|^2\right)^{\frac{1}{2}}. \tag{2.74}$$

On extending the sum over $s \in S$ on the right hand side of the above inequality to a sum over n^2 with n varying over $[1, P]$, where $P = N^{\frac{1}{2}}$, opening the square and using the triangle inequality together the fact that $|c_r| = 1$ for each r, we obtain

$$\sum_{s\in S}\left|\sum_{1\leq r\leq R} c_r e(st_r)\right|^2 \leq \sum_{1\leq r,s\leq R}\left|\sum_{1\leq n\leq P} e(n^2(t_r - t_s))\right|. \tag{2.75}$$

after an interchange of summations. Finally, on combining (2.73) through (2.75) we conclude that

$$\sum_{1\leq r,s\leq R}\left|\sum_{1\leq n\leq P} e(n^2(t_r-t_s))\right| \;\geq\; \delta^2 R^2 |S|\;. \tag{2.76}$$

The points t_r in Bourgain's method are the analogues of the rational numbers $\frac{a}{q}$ in the statement of Proposition 2.2.2. Bourgain's method treats sums of $e(n^2(t_r-t_s))$ for given t_r, t_s whereas our proof of Proposition 2.2.2 dealt with sums of $e(\frac{a}{q}(n^2-m^2))$ for given n, m. It is in this sense that we have said in the introduction to this chapter that our method is dual to that of Bourgain's argument. However, as in the proof of Proposition 2.2.2 the next step here is again an application of Hölder's inequality. Thus let $\ell > 2$ be fixed and p be such that $\frac{1}{l}+\frac{1}{p}=1$. Then an application of this inequality to the left hand side of (2.76) gives

$$R^{\frac{2}{p}}\left(\sum_{1\leq r,s\leq R}\left|\sum_{1\leq n\leq P} e(n^2(t_r-t_s))\right|^{\ell}\right)^{\frac{1}{\ell}} \;\geq\; \delta^2 R^2 |S|\;. \tag{2.77}$$

from which we get, after an obvious rearrangement of terms,

$$\sum_{1\leq r,s\leq R} |\phi(t_r-t_s)|^{\ell} \;\geq\; \delta^{2\ell} R^2 |S|^{\ell}\;, \tag{2.78}$$

on recalling the notation $\phi(t)=\sum_{1\leq n\leq P} e(n^2 t)$ introduced in Section 2.3.

The next step in Bourgain's method is to use a major arc - minor arc division of $\mathbf{T}$ and estimate $|\phi(t_r-t_s)|$ by means of the bounds provided in Subsections 2.2.1 and 2.2.2 depending on whether t_r-t_s lies in a majorc arc or a minor arc.

Let us set $L=(\log N)^2$, $Q=\frac{A^2}{\delta^5}$, $M=\frac{N}{L}$ and $P=N^{\frac{1}{2}}$. Then for any integers a and

q satisfying

$$0 \leq a < q \leq Q \ \text{ and } \ (a,q) = 1 \tag{2.79}$$

we call the set of $t \in \mathbf{T}$ satisfying $\|t - \frac{a}{q}\| \leq \frac{1}{M}$ the major arc $\mathfrak{M}(\frac{a}{q})$. An application of the triangle inequality on $\mathbf{T}$ together with (2.67) shows that distinct major arcs are in fact disjoint when N is sufficiently large . We denote the union of the family of major arcs $\mathfrak{M}(\frac{a}{q})$ by $\mathfrak{M}$ and by $\mathfrak{m}$ the complement in $\mathbf{T}$ of $\mathfrak{M}$.

Suppose now that $t_r - t_s$ lies in $\mathfrak{m}$ for some indices r, s. Then arguing in as at the top of page 25 we see that there is a θ in $\mathbf{R}$ such that $\theta \equiv t_r - t_s \bmod \mathbf{Z}$ and satisfying

$$|\theta - \frac{a}{q}| \leq \frac{1}{q^2} \tag{2.80}$$

with $Q < q \leq M$ and $(a,q) = 1$. Since $\phi(t_r - t_s) = \phi(\theta)$ we have on applying (2.8) to θ that

$$|\phi(t_r - t_s)|^2 \ll \frac{N}{Q} + N^{\frac{1}{2}} \log M + M \log M \ll \frac{N\delta^5}{A^2} + N^{\frac{1}{2}} \log N \ll \frac{N\delta^5}{A^2}\,, \tag{2.81}$$

on account of (2.67). Consequently, we have

$$\sum_{\substack{1 \leq r,s \leq R, \\ t_r - t_s \in \mathfrak{m}.}} |\phi(t_r - t_s)|^\ell \ \ll \ \frac{R^2 N^{\frac{\ell}{2}} \delta^{\frac{5\ell}{2}}}{A^\ell} \ \ll \delta^{\frac{5\ell}{2}} R^2 |S|^l \,, \tag{2.82}$$

since $|S| \geq \frac{N^{\frac{1}{2}}}{A}$. On comparing with (2.78) and recalling that $\mathfrak{m}$ is the complement of $\mathfrak{M}$ in $\mathbf{T}$ we then conclude that

$$\sum_{\substack{1\le r,s\le R,\\ t_r-t_s\in\mathfrak{M}.}} |\phi(t_r-t_s)|^\ell \gg \delta^{2\ell}R^2|S|^\ell \, , \tag{2.83}$$

We estimate the left hand side of the above inequality with the aid of the real valued function F defined on $\mathbf{T}$ by

$$F(t) = \frac{1}{(1+P^2|\sin(\pi t)|)^{\frac{\ell}{2}}} \, . \tag{2.84}$$

In effect, suppose that t lies in $\mathfrak{M}\left(\frac{a}{q}\right)$. Then by the estimate (2.10) we have

$$|\phi(t)|^\ell \ll \frac{P^\ell}{q^{\frac{\ell}{2}}} \frac{1}{\left(1+P^2\|t-\frac{a}{q}\|\right)^{\frac{\ell}{2}}} \ll \frac{P^\ell}{q^{\frac{\ell}{2}}} F\left(t-\frac{a}{q}\right) \, , \tag{2.85}$$

on using the classical inequality $|\sin\pi\theta| \le \pi\|\theta\|$, with $\theta = t-\frac{a}{q}$. By positivity, we therefore have for all t in $\mathfrak{M}$ that

$$|\phi(t)|^\ell \ll P^\ell \sum_{1\le q\le Q} \frac{1}{q^{\frac{\ell}{2}}} \sum_{a \bmod q} F\left(t-\frac{a}{q}\right) \, , \tag{2.86}$$

If for any t in $\mathbf{T}$ we set

$$G(t) = \sum_{1\le q\le Q} \frac{1}{q^{\frac{\ell}{2}}} \sum_{a \bmod q} F\left(t-\frac{a}{q}\right) \, , \tag{2.87}$$

we then obtain from (2.83) and positivity that

$$\sum_{1\leq r,s\leq R} G(t_r - t_s) \;\gg\; \frac{\delta^{2\ell} R^2 |S|^{\ell}}{N^{\frac{\ell}{2}}}\;, \tag{2.88}$$

since $P = N^{\frac{1}{2}}$.

Let χ be the characteristic function of the subset of t in $\mathbf{T}$ satisfying $\|t\| \leq \frac{1}{4N}$ and let

$$\psi(t) = \sum_{1\leq r\leq R} \chi(t - t_r)\,. \tag{2.89}$$

With $\check{\psi}(t) = \psi(-t)$ we shall verify that

$$\int_{\mathbf{T}} \psi * \check{\psi}(t)\, G(t) d\mu \;\geq\; \frac{2^{\frac{\ell}{2}}}{8N^2} \sum_{1\leq r,s\leq R} G(t_r - t_s). \tag{2.90}$$

To this end, let us set temporarily set $j(\theta) = (1 + P^2|\sin(\pi t)|)^{-1}$. Then we have $F(\theta) = j(\theta)^{\frac{\ell}{2}}$. Suppose now that $\|\theta - \theta'\| \leq \frac{1}{4N}$. Then we obtain

$$|j(\theta) - j(\theta')| \leq N\,\big|\,|\sin(\pi\theta)| - |\sin(\pi\theta')|\,\big|\, j(\theta) \leq j(\theta), \tag{2.91}$$

since $\big|\,|\sin(\pi\theta)| - |\sin(\pi\theta')|\,\big| \leq \pi\|\theta - \theta'\|$, by the mean value theorem, and $N \geq 1$. Consequently, $j(\theta) \leq 2j(\theta')$. This gives $F(\theta) \leq 2^{\frac{\ell}{2}} F(\theta')$ and finally $G(\theta) \leq 2^{\frac{\ell}{2}} G(\theta')$ when $\|\theta - \theta'\| \leq \frac{1}{4N}$. Applying this with $\theta = t_r - t_s$ and $\theta' = t$ we then see that for any indices r, s we have

$$\int_{\mathbf{T}} \chi(t - (t_r - t_s))\, G(t) d\mu \;\geq\; \frac{2^{\frac{\ell}{2}-1}}{N} G(t_r - t_s)\;. \tag{2.92}$$

Now we note that $\chi * \chi(t) = \frac{1}{2N} - \|t\|$ and therefore that $\chi * \chi(t) \geq \frac{1}{4N}\chi(t)$ for all t in $\mathbf{T}$. This gives

$$\int_{\mathbf{T}} \chi * \chi(t - (t_r - t_s))\, G(t) d\mu \;\geq\; \frac{2^{\frac{\ell}{2}-3}}{N^2} G(t_r - t_s)\ . \tag{2.93}$$

The inequality (2.90) now follows on summing both sides of the above relation over $1 \leq r, s \leq R$ and noting the identity

$$\psi * \check{\psi}(t) \;=\; \sum_{1 \leq r,s \leq R} \chi * \chi(t - (t_r - t_s))\ , \tag{2.94}$$

valid since $\chi(t) = \chi(-t)$. On combining (2.90) with (2.88) we then conclude that

$$\frac{2^{\frac{\ell}{2}} \delta^{2\ell} R^2 |S|^{\ell}}{N^{2+\frac{\ell}{2}}} \;\ll\; \int_{\mathbf{T}} \psi * \check{\psi}(t)\, G(t) d\mu\ . \tag{2.95}$$

To estimate the right hand side of the above relation we use a majorant for the function ψ with properties given by the following lemma.

Lemma 2.5.2. *There is a real number $C > 0$ such that for any $\frac{1}{N}$-spaced sequence $t_1, t_2, \ldots, t_r$ in $\mathbf{T}$ we have a real valued function σ on $\mathbf{T}$ satisfying the conditions*

(i) $0 \leq \sigma(t) \leq C$ for all $t \in \mathbf{T}$.

(ii) $1 \leq \sigma(t)$ if $\|t - t_r\| \leq \frac{1}{2N}$ for some r.

(iii) $\int_{\mathbf{T}} |\sigma(t)|\, d\mu \leq \frac{CR}{N}$.

(iv) $\widehat{\sigma}(k) = 0$ when $|k| \geq N$.

Proof.— Using the well-known properties of the Fejer kernel we easily verify that the requirements of the lemma are met with $C = \left(1 + \frac{1}{2}\sum_{k\geq 1}\frac{1}{k^2}\right)\frac{\pi^2}{4}$ and σ defined by

$$\sigma(t) = \frac{\pi^2}{4N^2}\sum_{1\leq r\leq R}\left(\frac{\sin \pi N(t-t_r)}{\sin \pi(t-t_r)}\right)^2. \tag{2.96}$$

□

It follows from (ii) of the lemma above and the fact that the t_r are $\frac{1}{N}$ spaced that $\sigma(t) \geq \psi(t)$ for all t in $\mathbf{T}$. We therefore obtain

$$\frac{2^{\frac{\ell}{2}}\delta^{2\ell}R^2|S|^\ell}{N^{2+\frac{\ell}{2}}} \ll \int_{\mathbf{T}} \sigma * \check{\sigma}(t)\, G(t) d\mu \;= \sum_{|k|\leq N}|\widehat{\sigma}(k)|^2\widehat{G}(k). \tag{2.97}$$

using the Parseval relation and (iv) of Lemma 2.5.2. From the definition of $G(t)$ in (2.87) we have

$$\widehat{G}(k) = \sum_{\substack{1\leq q\leq Q,\\ q|k.}} q^{1-\frac{\ell}{2}}\widehat{F}(k). \tag{2.98}$$

On noting that

$$|\widehat{F}(k)| \leq \int_{\mathbf{T}}|F(t)|d\mu = \int_{\mathbf{T}}\frac{d\mu}{1+N|\sin(\pi t)|} \ll \frac{1}{N}, \tag{2.99}$$

and since $\ell > 2$ we deduce from (2.98) that

$$|\widehat{G}(k)| \ll \frac{\tau_Q(k)}{N} \tag{2.100}$$

where $\tau_Q(k) = |\{q|k \,|\, 1\leq q\leq Q\}|$. For a given $\kappa > 0$ we have

$$\sum_{|k|\le N} |\widehat{\sigma}(k)|^2 \widehat{G}(k) \ \le\ \frac{Q}{N} \sum_{\substack{|k|\le N, \\ \tau_Q(k)>Q^\kappa.}} |\widehat{\sigma}(k)|^2 + \frac{Q^\kappa}{N} \sum_{k\in\mathbf{Z}} |\widehat{\sigma}(k)|^2 \tag{2.101}$$

where we have used the trivial bound $\tau_Q(k) \le Q$ for k such that $\tau_Q(k) > Q^\kappa$ and, for the reamining k, enlarged the summation over k to all $k \in \mathbf{Z}$. Using (iii) of Lemma 2.5.2 we have $|\widehat{\sigma}(k)| \le \frac{CR}{N}$ for all $k \in \mathbf{Z}$. Also, by the Plancherel relation we have

$$\sum_{k\in\mathbf{Z}} |\widehat{\sigma}(k)|^2 = \int_{\mathbf{T}} |\sigma(t)|^2 d\mu \ \le\ \frac{C^2 R}{N}, \tag{2.102}$$

on account of (i) and (iii) of Lemma 2.5.2. Finally, Lemma 4.8 on page 307 of [5] tells us that the number of integers k with $|k| \le N$ and $\tau_Q(k) > Q^\kappa$ does not exceed $C_{\kappa,L} N Q^{-L}$ for any $L > \kappa$. With these remarks we deduce from (2.97) and (2.101) that

$$\frac{2^{\frac{\ell}{2}} \delta^{2\ell} R^2 |S|^\ell}{N^{2+\frac{\ell}{2}}} \ \ll_{\kappa,L}\ \frac{R^2}{Q^{L-1} N^3} + \frac{Q^\kappa R}{N^2} \tag{2.103}$$

for any $\kappa > 0$ and $L > \kappa$. Since $R \le 2N$ by a trivial estimate, the second term on the right hand side of the inequality above dominates the first when $L \ge 1 + \kappa$. Consequently, we obtain

$$\frac{2^{\frac{\ell}{2}} \delta^{2\ell} R^2 |S|^\ell}{N^{2+\frac{\ell}{2}}} \ \ll_{\kappa}\ \frac{Q^\kappa R}{N^2}\ , \tag{2.104}$$

which gives (2.70) on recalling that $Q = \frac{A^2}{\delta^5}$ and rearranging terms. This completes the proof of Theorem 2.5.1.

CHAPTER 3

Sárközy's Problem for the Primes

3.1 Introduction

We write $\mathcal{P}$ for the set of primes and, as in Section 1.1, use $t(K)$, for any integer $K \geq 1$, to denote the smallest integer with the property that for any colouring of $\mathcal{P}$ in K colours, every sufficiently large integer is expressible as a sum of at most $t(K)$ primes, all of the same colour. As stated in Chapter 1, it has been shown in [18] that

Theorem 3.1.1. *We have $t(K) \leq CK \log\log 4K$.*

An example in [11] tells us that the bound for $t(K)$ given by the above therem is optimal up to the value of the constant C. Our purpose in this chapter is to give a proof of this bound, with only a slightly different constant, but by a simplified version of the method of [18]. In particular, we shall not make use of the improved large sieve inequality for the primes given by Proposition 1.1.2 but only use the classical large sieve inequality and a simple estimate obtained by means of the circle method, together with a combinatorial result from [18].

We shall deduce Theorem 3.1.1 in Section 3.4 from the theorem below and the argument from [11], this time for the primes, based on Sárközy's finite addition theorem. We recall here our notation $\pi^*(N)$ for the number of primes in the interval $(2N, 3N]$, for any integer $N \geq 1$.

Theorem 3.1.2. *For any integer $K \geq 1$ there is an integer $N(K)$ such that for all $N \geq N(K)$ and any subset S of the prime numbers in the interval $(2N, 3N]$ with $|S| \geq \pi^*(N)/K$ we have*

$$E_3(S) \;\leq\; \frac{M}{\phi(M)} \frac{|S|^5}{\log{(2N)}} \exp\left(\frac{C_1}{\log\log 4K} \right) \;, \tag{3.1}$$

where M is the product of all prime numbers not exceeding $(4 \log 4K \log\log 4K)^2$.

Let us outline the proof of Theorem 3.1.1, referring to Section 3.3 for a detailed presentation. For any $\mathbf{x} = (x_1, x_2, \ldots, x_5) \in S^5$, we write

$$l(\mathbf{x}) = x_1 + x_2 + x_3 - x_4 - x_5 \; . \tag{3.2}$$

We begin by observing that the proof of Theorem 3.1.1 amounts to the estimation of the sum

$$\sum_{\mathbf{x} \in S^5} \Lambda(l(\mathbf{x})) \; , \tag{3.3}$$

where Λ is the Van Mangoldt function. To estimate this sum we use the simple sieve identity for Λ given by the relation

$$\Lambda(n) = -\sum_{d|n} \mu(d) \log d = -\sum_{\substack{d|n,\\ d\leq L.}} \mu(d) \log d - \sum_{\substack{d|n,\\ d>L.}} \mu(d) \log d \tag{3.4}$$

for any integer $n \geq 1$ and any $L \geq 2$. With $L = N^{1/2}$, we insert (3.4) into (3.3) and obtain two sums, one over $d \leq L$ and the other over $d > L$. A standard application of Davenport's classical bound for $\sum_n \mu(n)e(nt)$ shows that the sum over $d > L$ is indeed majorised by the right hand side of (3.1). We then we use properties of the arithmetical function $\omega(q, L)$ defined for integers $q, L \geq 1$ by

$$\omega(q, L) \;=\; -\sum_{\substack{1\leq l\leq L,\\ l\equiv 0 \text{ mod } q.}} \frac{\mu(l) \log l}{l} \tag{3.5}$$

to show that the estimation of the sum over $d \leq L$ reduces to that of

$$\sum_{1\leq q\leq(\log N)^4} \frac{\mu(q)}{\phi(q)} \sum_{a \bmod^* q} \sum_{\mathbf{x}\in S^5} e\left(\frac{al(\mathbf{x})}{q}\right) . \tag{3.6}$$

Finally, we treat the sum (3.6) by means of a simple corollary to the classical large sieve inequality and a combinatorial probabilistic proposition from [18], recalled in Subsection 3.2.4 below.

3.2 Preliminaries

3.2.1 Bounds for $\omega(q, L)$

We record here two key estimates for the arithmetical function $\omega(q, L)$ defined for integers $q, L \geq 1$ by (3.5). This arithemtical function appears in a number of modern treatments of Vinogradov's three primes theorem based on the large sieve inequality. For instance, see Chapter 10 of [19].

Since $\omega(q, L) \neq 0$ only when q is a squarefree integer, which we shall assume this for the remainder of this subsection. With this assumption we have on writing $l = qk$ and rearranging terms in (3.5) that

$$\omega(q, L) \;=\; -\frac{\mu(q)}{q} \sum_{\substack{1 \leq k \leq L/q, \\ (k,q)=1.}} \frac{\mu(k) \log qk}{k} \;. \tag{3.7}$$

An application of the triangle inequality then yields, for all $q, L \geq 1$, the upper bound

$$|\omega(q, L)| \leq \;\frac{1}{q}\,(\log L)(1 + \log(L/q)) \leq \;\frac{(\log 2L)^2}{q}\;. \tag{3.8}$$

When $1 \leq q \leq L^{\frac{1}{2}}$ the preceding bound may be refined, by an application of Perron's formula to the sum on the right hand side of (3.7), to the asymptotic formula

$$\omega(q, L) \;=\; \frac{\mu(q)}{\phi(q)} + O\left(\frac{2^{\omega(q)} \log 2q}{q(\log L)^A}\right) \;, \tag{3.9}$$

for any $A \geq 1$, where the constant implicit in the O symbol depends only on A. In the error term in (3.9), $\omega(q)$ stands, of course, for the number of prime divisors of q.

3.2.2 An Application of the Large Sieve Inequality

The classical large sieve inequality is the inequality stated below. We refer, for example, to Theorem 7.28, page 178 of [12], for the optimal form and a proof.

Proposition 3.2.1. *Let $N, Q \geq 1$ be integers. Then for any complex numbers $u(n)$, with $n \in I$, an interval of length N, we have*

$$\sum_{1\leq q\leq Q}\sum_{a \bmod^* q}\left|\sum_{n} u(n) e\left(\frac{an}{q}\right)\right|^2 \leq (N+Q^2)\sum_{n}|u(n)|^2 . \qquad (3.10)$$

Here, as in the previous chapter, $a \bmod^* q$ means that a runs over all the invertible residue classes modulo q. Our substitute for Proposition 1.1.2 used in [18] is the following corollary. For a finite subset S of the integers we will write $\widehat{S}(t) = \sum_{n\in S} e(nt)$.

Corollary 3.2.2. *Let $N \geq 1$ be an integer and let S be a set of primes in the interval $(2N, 3N]$. Then for any integer Q with $1 \leq Q \leq N^{\frac{1}{2}}$ we have*

$$\sum_{1\leq q\leq Q}\sum_{a \bmod^* q}\left|\widehat{S}\left(\frac{a}{q}\right)\right|^4 \ll \frac{N^4}{(\log N)^4} . \qquad (3.11)$$

Proof.— Let $u(n)$ for any integer n be the number of pairs $(x_1, x_2) \in S^2$ such that $n = x_1 + x_2$. Since $S + S$ is contained in $(4N, 6N])$, we then have that

$$\widehat{S}(t)^2 = \sum_{4N<n\leq 6N} u(n) e(nt) . \qquad (3.12)$$

Substituting this into the left hand side of (3.11) and using (3.10) we see that

$$\sum_{1\le q\le Q}\sum_{a \bmod^* q}\left|\widehat{S}\left(\frac{a}{q}\right)\right|^4 \ll N \sum_{4N<n\le 6N} u(n)^2. \tag{3.13}$$

If, temporarily, we denote by $r(n)$ the number of pairs of prime numbers (p_1, p_2) such that $n = p_1 + p_2$, then evidently $u(n) \le r(n)$, for any integer n. The conclusion of the corollary now follows on noting that

$$N \sum_{4N<n\le 6N} u(n)^2 \le N \sum_{1\le n\le 6N} r(n)^2 \ll \frac{N^4}{(\log N)^4}\ . \tag{3.14}$$

on using the classical bound $\sum_{1\le n\le 6N} r(n)^2 \ll \frac{N^3}{(\log N)^4}$, supplied, for instance, by an application of the circle method.

3.2.3 A Simple Optimisation Principle

In this subsection and the next we recall the key inputs from [18] that we will use in our proof of Theorem 3.1.2.

Let $n \ge 1$ be an integer and let P, T be positive real numbers and suppose that the subset K of points $x = (x_1, x_2, \ldots, x_n)$ in $\mathbf{R}^n$ such that

$$\sum_{1\le i\le n} x_i = P \ \text{ and } \ 0 \le x_i \le T \ \text{ for all } i\ . \tag{3.15}$$

is not empty. Then K is a non-empty, compact and convex subset of $\mathbf{R}^n$.

Proposition 3.2.3. *Let $f(x,y) = \sum_{1\le i,j\le n} c_{ij} x_i y_j$ from $\mathbf{R}^n \times \mathbf{R}^n$ to $\mathbf{R}$ be a bilinear* *form with real coefficients c_{ij}. Then there exist extreme points x^* and y^* of the convex set K such that $f(x,x) \leq f(x^*, y^*)$ for all x in K.*

For the proof we refer to Section 2.3 of [18]. We will require a description of the extreme points of K to apply this proposition. Again, it is shown in Section 2.3 of [18] that if x is an extreme point of K then, excepting at most one, all co-ordinates of x are equal to either 0 or T. Further that, if k is the number of co-ordinates of x that are distinct from 0 then we have from (3.15) that k is determined by the inequalities $kT \geq P \geq (k-1)T$.

3.2.4 A Combinatorial Problem

Let X be the product of a finite family of finite sets $\{X_i\}_{i \in I}$ and A and B be non-empty subsets of X. Given a subset J of I, the proposition below, which is a particular case of Proposition 2.2 of Subsection 2.3 in [18], provides an upper bound for the number $\mathcal{T}_J(A,B)$ of pairs (a,b) in $A \times B$ such that $a_i \neq b_i$ for all $i \in J$, where a_i and b_i are the co-ordinates of index i of a and b respectively.

Proposition 3.2.4. *For any integer t satisfying $1 \leq t \leq \inf_{i \in J} |X_i|^{\frac{1}{2}}$ we have that*

$$\mathcal{T}_J(A,B) \leq |A||B| \exp\left(-\sum_{i \in J} \frac{1}{|X_i|}\right) \exp\left(\frac{L(A,B)}{t} + tw_J\right) . \tag{3.16}$$

where

$$L(A,B) = \log\left(\frac{|X|^2}{|A||B|}\right) \quad \textit{and} \quad w_J = \sum_{i \in J} \frac{1}{|X_i|^2} . \tag{3.17}$$

3.3 Proof of Theorem 3.1.2

The proof given here follows the method of [18] in all details, except that we will use Corollary 3.2.2 in place of Proposition 1.1.2.

Throughout this section we fix an integer $K \geq 2$. Also, N will denote a sufficiently large integer, its actual magnitude varying to meet our (finitely many) requirements.

We recall that $E_3(S)$ is the number of quadruples $(x_1, x_2, x_3, x_4, x_5, x_6)$ in S^6 such that

$$x_1 + x_2 + x_3 = x_4 + x_5 + x_6 \ . \tag{3.18}$$

In other words, $E_3(S)$ is the number of $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$ in S^5 such that $\ell(\mathbf{x})$, defined by (3.2), is in S. Therefore $E_3(S)$ does not exceed number of $\mathbf{x}$ in S^5 such that $\ell(\mathbf{x})$ is a prime number $> 2N$ or, more generally, such that $\Lambda(\ell(\mathbf{x})) > \log(2N)$. This justifies the inequality

$$E_3(S)\log(2N) \leq \sum_{\mathbf{x}\in S^5} \Lambda(\ell(\mathbf{x})) \ , \tag{3.19}$$

since $\ell(\mathbf{x}) \geq 1$ for all $\mathbf{x}$ in S^5, on recalling that S is in $(2N, 3N]$.

We use (3.4) with $L = N^{1/2}$ to estimate the sum on the right hand side of (3.19). Following usual notation, we write $\Lambda(n) = \Lambda^\sharp(n) + \Lambda^\flat(n)$, where $\Lambda^\sharp(n)$ and $\Lambda^\flat(n)$ are, respectively, the sums in (3.4) over $d \leq L$ and $d > L$, together with their signs. Substituting into (3.19) and using $r(n)$ to denote the number of $\mathbf{x}$ in S^5 such that $\ell(\mathbf{x}) = n$, for any integer n, we have

$$E_3(S)\log(2N) \ \leq \ \sum_n r(n)\Lambda^\sharp(n) + \sum_n r(n)\Lambda^\flat(n)\ . \tag{3.20}$$

Let us first estimate the second sum on the right hand side of the above relation. This we do by a standard application of Davenport's bound for the exponential sum $\sum_{1\leq d\leq u}\mu(d)e(dkt)$, given, for example, by Theorem 13.10, page 348 of [12], for a any k and t and any $u \geq 1$.

In effect, on recalling from Subsection 3.2.2 the notation $\widehat{S}(t) = \sum_{n\in S} e(nt)$ and since $r(n) = 0$ for n outside the interval $[1, 5N]$, we have

$$\sum_n r(n)\Lambda^\flat(n) \ = \ \int_0^1 \widehat{S}(t)^3\,\widehat{S}(-t)^2 \left(\sum_{1\leq n\leq 5N} \Lambda^\flat(n)e(nt)\right)\,dt\ , \tag{3.21}$$

by orthogonality of the functions $t \mapsto e(nt)$ on $[0, 1]$. From the definition of $\Lambda^\flat(n)$, and on setting $n = dk$, we obtain after an interchange of summations that

$$\sum_{1\leq n\leq 5N} \Lambda^\flat(n)e(nt) \ = \ -\sum_{1\leq k<\frac{5N}{L}}\ \sum_{L<d\leq\frac{5N}{k}} \mu(d)\log d\, e(dkt)\ . \tag{3.22}$$

For given k and t, let $T(u)$ denote the sum $\sum_{1\leq d\leq u}\mu(d)e(dkt)$ for any $u \geq 1$. Then

$$\sum_{L<d\leq\frac{5N}{k}} \mu(d)\log d\, e(dkt) = \int_L^{\frac{5N}{k}} \log u\ dT(u)\ . \tag{3.23}$$

Davenport's bound tells us that for any $A > 0$, we have $T(u) \ll u(\log u)^{-A}$ for all $u \geq 2$, uniformly in k and t. The implied constant here depends only on A. Using this bound with $A = 4$ and integrating by parts we have

$$\left| \int_L^{\frac{5N}{k}} \log u \, dT(u) \right| \leq 3 \log \left(\frac{5N}{k} \right) \max_{L \leq u \leq \frac{5N}{k}} |T(u)| \ll \frac{N \log(5N)}{k (\log L)^4} \, , \tag{3.24}$$

since $\log u$ is an increasing function for $L \leq u \leq \frac{5N}{k}$. Finally, since $L = N^{1/2}$, we conclude from (3.22) through (3.24) that, for all t in $[0, 1]$, we have

$$\sum_{1 \leq n \leq 5N} \Lambda^{\flat}(n) e(nt) \ll \frac{N}{(\log N)^3} \sum_{1 \leq n \leq 5N} \frac{1}{k} \ll \frac{N}{(\log N)^2} \, , \tag{3.25}$$

We then obtain from (3.21) and (3.25) that

$$\sum_n r(n) \Lambda^{\flat}(n) \ll \frac{N}{(\log N)^2} \int_0^1 |\widehat{S}(t)|^3 |\widehat{S}(-t)|^2 \, dt \ll \frac{N |S|^4}{(\log N)^2} \, , \tag{3.26}$$

where the last inequality follows on using the trivial bound $|\widehat{S}(t)||\widehat{S}(-t)|^2 \leq |S|^3$ together with Parseval's relation $\int_0^1 |\widehat{S}(t)|^2 \ dt = |S|$. The implied constants in (3.26) are absolute.

We now turn to the first sum on the right hand side of (3.20). The definition of $\Lambda^{\sharp}(n)$ gives

$$\sum_n r(n) \Lambda^{\sharp}(n) = - \sum_{1 \leq d \leq L} \mu(d) \log d \sum_{n \equiv 0 \bmod d} r(n) \, , \tag{3.27}$$

after an interchange of summations. Let us note that

$$\sum_{n\equiv 0 \bmod d} r(n) = \frac{1}{d}\sum_{a \bmod d}\sum_n r(n)e\left(\frac{an}{d}\right) = \frac{1}{d}\sum_{q|d}\sum_{a \bmod^* q}\sum_n r(n)e\left(\frac{an}{q}\right), \quad (3.28)$$

by orthogonality of characters on the group $\mathbf{Z}/d\mathbf{Z}$ and also that

$$\sum_n r(n)e\left(\frac{an}{q}\right) = \widehat{S}\left(\frac{a}{q}\right)^3 \widehat{S}\left(-\frac{a}{q}\right)^2 = \left|\widehat{S}\left(\frac{a}{q}\right)\right|^4 \widehat{S}\left(\frac{a}{q}\right), \quad (3.29)$$

from the definition of $r(n)$. Thus, on combining (3.29) and (3.28) with (3.27), interchanging summations and recalling the definition of the arithmetical function $\omega(q,L)$ from (3.5) we obtain

$$\sum_n r(n)\Lambda^\sharp(n) = \sum_{1\le q\le L}\omega(q,L)\sum_{a \bmod^* q}\left|\widehat{S}\left(\frac{a}{q}\right)\right|^4 \widehat{S}\left(\frac{a}{q}\right). \quad (3.30)$$

We now estimate the contribution to the sum on the right hand side of (3.30) from q with $(\log N)^4 < q \le L$. We shall verify that

$$\sum_{(\log N)^4<q\le L}|\omega(q,L)|\sum_{a \bmod^* q}\left|\widehat{S}\left(\frac{a}{q}\right)\right|^4\left|\widehat{S}\left(\frac{a}{q}\right)\right| \le \frac{2N|S|^4}{(\log N)^2}. \quad (3.31)$$

Indeed, the bound for $\omega(q,L)$ given by (3.8) and the trivial bound $|\widehat{S}(a/q)| \le |S|$ show that the absolute value of the left hand side of (3.31) does not exceed

$$\frac{(\log 2L)^2|S|^3}{(\log N)^4}\sum_{1\le q\le L}\sum_{a \bmod^* q}\left|\widehat{S}\left(\frac{a}{q}\right)\right|^2 \le \frac{(\log 2L)^2(N+L^2)|S|^4}{(\log N)^4}, \quad (3.32)$$

where we have extended the range of summation for the sum over q back to $1 \leq q \leq L$ and applied the classical large sieve inequality recalled in Subsection 3.2.2. Since $L = N^{1/2}$, we have $\log 2L \leq \log N$ for $N \geq 4$. Thus, (3.32) gives (3.31) when N is sufficiently large.

Hereafter, let T be the sum on the right hand side of (3.30) restricted to q satisfying $1 \leq q \leq (\log N)^4$ and let us apply the asymptotic formula for $\omega(q, L)$ given by (3.9) to such q. The contribution to T from the error term of this formula is easily estimated. Indeed, we have

$$\sum_{1 \leq q \leq (\log N)^4} \frac{2^{\omega(q)} \log(2q)}{q(\log L)^2} \sum_{a \bmod^* q} \left| \widehat{S}\left(\frac{a}{q}\right) \right|^4 \left| \widehat{S}\left(\frac{a}{q}\right) \right| \leq \frac{12N|S|^4}{(\log L)^2}, \tag{3.33}$$

where we have used the large sieve inequality as above together with the trivial bounds $2^{\omega(q)} \log(2q) \leq \tau(q) \log(2q) \leq 6q$, for any $q \geq 1$. Therefore, (3.9) applied with $A = 2$ gives

$$T = \sum_{1 \leq q \leq (\log N)^4} \frac{\mu(q)}{\phi(q)} \sum_{a \bmod^* q} \left| \widehat{S}\left(\frac{a}{q}\right) \right|^4 \widehat{S}\left(\frac{a}{q}\right) + O\left(\frac{N|S|^4}{(\log N)^2}\right), \tag{3.34}$$

where the constant implicit in the O symbol is absolute.

It remains to treat the sum on the right hand side of (3.34). To this end, let us set $X = K^6$ and write U to denote the product of all primes not exceeding X. Then for sufficiently large N, depending on K, we have $U \leq (\log N)^4$. We then set $T(U)$ to be the sum over q on the right hand side of (3.34) restricted to those q that divide U. Since with all other q we have either $\mu(q) = 0$ or $q > X$, the triangle inequality applied to (3.34) gives

$$|T - T(U)| \leq |S| \sum_{X<q\leq(\log N)^4} \frac{1}{\phi(q)} \sum_{a \bmod^* q} \left|\widehat{S}\left(\frac{a}{q}\right)\right|^4 + O\left(\frac{N|S|^4}{(\log N)^2}\right) . \tag{3.35}$$

We shall presently apply Corollary 3.2.2 to estimate the sum on the right hand side of (3.35). Notice that since we have chosen to estimate $E_3(S)$ rather than $E_2(S)$, we have $|\widehat{S}\left(\frac{a}{q}\right)|^4$ in the above sum in place of $|\widehat{S}\left(\frac{a}{q}\right)|^2$ that appears in [18]. Then Corollary 3.2.2 is no more applicable and one requires the deeper Proposition 1.1.2. By means of the trivial bounds $\frac{q}{\phi(q)} \leq \omega(q) + 1 \leq 2\log 2q$ we see that the sum over q on the right hand side of the above relation does not exceed

$$\frac{2\log 2X}{X} \sum_{1\leq q\leq(\log N)^4} \sum_{a \bmod^* q} \left|\widehat{S}\left(\frac{a}{q}\right)\right|^4 \ll \frac{(\log 2K)N^4}{K^6(\log N)^4} \ll \frac{(\log 2K)|S|^4}{K} , \tag{3.36}$$

where we have used the bound provided by Corollary 3.2.2 and $|S| \gg \frac{N}{K\log N}$. Thus, we obtain from (3.35) that

$$T - T(U) \ll \frac{|S|^5 \log 2K}{K^2} + \frac{|S|^5}{\log N} , \tag{3.37}$$

where again we have used $|S| \gg \frac{N}{K\log N}$ in the error term on the right hand side of (3.35). Since T stands for the sum on the right hand side of (3.30), we conclude on combining (3.26) and (3.30) with (3.37) that

$$\sum_{\mathbf{x}\in S^5} \Lambda(\ell(\mathbf{x})) - T(U) \ll \frac{|S|^5 \log 2K}{K^2} \ . \tag{3.38}$$

for all $N \geq N(K)$ depending only on K. To commence the final step of the proof of Theorem 3.1.2, which is the estimation of $T(U)$, we note that in fact

$$T(U) = \frac{U}{\phi(U)} \left|\{\mathbf{x} \in S^5 \,|\, (\ell(\mathbf{x}), U) = 1\}\right| \ . \tag{3.39}$$

Indeed, since U is a squarefree integer, for any integer $n \neq 0$ we have the identity

$$\sum_{q|U} \frac{\mu(q)}{\phi(q)} \sum_{a \,\mathrm{mod}^*\, q} e\left(\frac{an}{q}\right) \ = \ \frac{U}{\phi(U)} \sum_{q|(U,n)} \mu(q) \ , \tag{3.40}$$

by classical properties of Ramanujan sums. Since by the Möbius inversion formula the right hand side of (3.40) is $\frac{U}{\phi(U)}$ when $(n,U) = 1$ and is 0 otherwise, we have on recalling the definitions of $r(n)$ and $T(U)$ that

$$T(U) = \sum_{q|U} \frac{\mu(q)}{\phi(q)} \sum_{a \,\mathrm{mod}^*\, q} \sum_{n} r(n) e\left(\frac{an}{q}\right) \ = \ \frac{U}{\phi(U)} \sum_{(n,U)=1} r(n) \ , \tag{3.41}$$

where the last sum is nothing but the right hand side of (3.39).

For any subset Z of the integers, $\tilde{Z}$ denote its canonical image in $\mathbf{Z}/U\mathbf{Z}$. Also, we will write $z = (x_3, x_4, x_5)$ for elements of S^3 and set $\tilde{z}$ to denote the canonical image of $x_4 + x_5 - x_3$ in $\mathbf{Z}/U\mathbf{Z}$. Further, for any residue class a modulo U, let $m(a)$ be the number of elements of the set S that belong to this residue class. If D denotes $\frac{6N}{\phi(U)\log N}$, we then have that

$$\sum_{a\in\tilde{S}} m(a) = |S| \text{ and } 0 \leq m(a) \leq D \tag{3.42}$$

for all sufficiently large N, where the upper bound for $m(a)$ follows from the Prime Number Theorem for arithmetical progressions on recalling that S is a set of prime numbers in the interval $(2N, 3N]$. Finally, for any $z \in S^3$ and residue classes a, b modulo U, let us set $C_z(a,b)$ to be 1 when $a+b-\tilde{z}$ is invertible in $\mathbf{Z}/U\mathbf{Z}$ and to be 0 otherwise. Then translating (3.39) in terms of these notations we see that

$$T(U) = \frac{U}{\phi(U)} \sum_{z\in S^3} \sum_{(a,b)\in\tilde{S}^2} C_{\tilde{z}}(a,b)\, m(a)m(b)\ . \tag{3.43}$$

Let us fix a z in S^3. Then the inner sum on the right hand side of (3.43) can be bounded with the aid of the optimization principle given by Proposition 3.2.3, taking account of the conditions (3.42). From the description of extreme points in Subsection 3.2.3 we then have that

$$\sum_{(a,b)\in\tilde{S}^2} C_{\tilde{z}}(a,b)\, m(a)m(b) \leq \sum_{(a,b)\in\tilde{S}^2} C_{\tilde{z}}(a,b)\, x_a^* y_b^*\,, \tag{3.44}$$

for some x_a^* and y_b^*, with a and b varying over $\tilde{S}$, satisfying the following conditions. All the x_a^*, and similarly all the y_b^*, are either 0 or D excepting at most one, which must lie in $(0, D)$. Moreover, if A and B denote, respectively, the subsets of $\tilde{S}$ for which $x_a^* \neq 0$ and $y_b^* \neq 0$, then $|A|D \geq |S| > (|A|-1)D$ and the same inequalities hold with $|A|$ replaced by $|B|$.

By the Prime Number Theorem we have $|S| \geq \frac{N}{2K \log N}$ for sufficiently large N. Also, $\phi(U) \geq K^6$. Therefore we have $\frac{|S|}{D} \geq \frac{\phi(U)}{6K} \geq K$, since $K \geq 2$. From the preceding paragraph it follows that

$$|A|,\, |B| \;\geq\; \frac{\phi(U)}{6K} \;\geq\; K\,, \tag{3.45}$$

from which we obtain

$$D^2 \;\leq\; \frac{|S|^2}{(|A|-1)(|B|-1)} \;\leq\; \frac{|S|^2}{|A||B|} \exp\left(\frac{2}{K}\right)\,, \tag{3.46}$$

by the inequalities $(1-x)^{-1} \leq 1 + 2x \leq \exp(2x)$ when x is in $[0, 1/2]$, applied with x taken to be $\frac{1}{|A|}$ and then to be $\frac{1}{|B|}$.

Since $C_{\tilde{z}}(a,b)$ is always positive and $x_a*, y_b^* \leq D$, we have from (3.44) and (3.46) that the left hand side of (3.44) does not exceed

$$D^2 \sum_{(a,b)\in A\times B} C_{\tilde{z}}(a,b) \;\leq\; \frac{|S|^2}{|A||B|} \exp\left(\frac{2}{K}\right) \sum_{(a,b)\in A\times B} C_{\tilde{z}}(a,b)\,. \tag{3.47}$$

The sum on the right hand side in (3.47) is equal to the number of pairs (a,b) in $A \times B$ such that $a + b - \tilde{z}$ is an invertible residue class modulo U. We shall presently bound this sum by means of Proposition 3.2.4.

Let I be the set of prime numbers not exceeding K^6. The Chinese remainder theorem gives

$$\mathbf{Z}/U\mathbf{Z} = \prod_{p \in I} \mathbf{Z}/p\mathbf{Z} \; . \tag{3.48}$$

Thus $\mathbf{Z}/U\mathbf{Z}$ is a finite product of the finite sets $\mathbf{Z}/p\mathbf{Z}$. For any a in $\mathbf{Z}/U\mathbf{Z}$ and p in I, let a_p be the canonical image of a in $\mathbf{Z}/p\mathbf{Z}$. Also, we identify the sets A and B with their images in the product on the right hand side of (3.48). If we set $B' = \tilde{z} - B$, then the sum on the right hand side in (3.47) is the same as the number of pairs (a, b') in $A \times B'$ such that $a_p \neq b'_p$ for all p in I.

Let us now set $Q = 4 \log 4K \log\log 4K$. Then $2 < Q^2 < X$. If J is the subset of I comprising the p in I with $p > Q^2$, then from the conclusion of the preceding paragraph we certainly have

$$\sum_{(a,b) \in A \times B} C_{\tilde{z}}(a, b) \leq \left| \{(a, b') \in A \times B' \,|\, a_p \neq b'_p \;\text{ for all } p \in J\} \right| \; . \tag{3.49}$$

We now use Proposition 3.2.4 to estimate the right hand side of (3.49). Since $|B'| = |B|$, we have $|A|, |B'| \geq \frac{\phi(U)}{6K}$ from (3.45). By means of the trivial bound $\frac{\phi(U)}{U} \geq \frac{1}{K^6}$, we then see that

$$L(A, B') \leq 14 \log 4K \; . \tag{3.50}$$

In our case ω_J of Proposition 3.2.4 is $\sum_{p \in J} \frac{1}{p^2}$ and we have

$$\omega_J \leq \sum_{n>Q^2} \frac{1}{n^2} \leq \frac{2}{Q^2} \; . \tag{3.51}$$

Finally, if P temporarily denotes the product of the primes in J then we have

$$\exp\left(-\sum_{p\in J}\frac{1}{p}\right) \;\leq\; \exp\left(\sum_{p\in J}\frac{2}{p^2}\right)\prod_{p\in J}(1-p^{-1}) \;\leq\; \frac{\phi(P)}{P}\exp\left(\frac{4}{Q^2}\right) \; . \tag{3.52}$$

since $-\log(1-x) \leq x+2x^2$ for $0 \leq x \leq 1/2$. Therefore, we conclude from Proposition 3.2.4 that for any integer t with $1 \leq t \leq Q$ we have

$$\sum_{(a,b)\in A\times B} C_{\tilde{z}}(a,b) \;\leq\; \frac{\phi(P)}{P}|A||B|\exp\left(\frac{14\log 4K}{t}+\frac{6t}{Q^2}\right) \; . \tag{3.53}$$

Since $Q \geq 1$, there is an integer $t \geq 1$ such that $Q/2 \leq t \leq Q$. Taking t to be such an integer we see that the expression in the brackets on the right hand side of (3.53) does not exceed

$$\frac{28\log 4K}{Q}+\frac{6}{Q} \;\leq\; \frac{13}{\log\log 4K} \; . \tag{3.54}$$

On combining (3.47) and (3.53) with the bound above and recalling that the left hand side of (3.47) is an upper bound for the left hand side of (3.44), we obtain that

$$\sum_{(a,b)\in\tilde{S}^2} C_{\tilde{z}}(a,b)\, m(a)m(b) \leq \frac{\phi(P)}{P}|S|^2\exp\left(\frac{13}{\log\log 4K}+\frac{2}{K}\right) \; , \tag{3.55}$$

uniformly for all $z \in S^3$. Since M of Theorem 3.1.2 is nothing but $\frac{U}{P}$, we conclude from (3.55) and (3.43) that

$$T(U) \leq \frac{M}{\phi(M)} |S|^5 \exp\left(\frac{13}{\log\log 4K} + \frac{2}{K}\right) , \tag{3.56}$$

for all sufficiently large N. Finally, on combining (3.19), (3.38) and (3.56) we obtain

$$E(S) \;\leq\; \frac{M}{\phi(M)} \frac{|S|^5}{\log{(2N)}} \exp\left(\frac{14}{\log\log 4K} + \frac{2}{K}\right) \left(1 + \frac{C(\log 2K)}{K^2}\right). \tag{3.57}$$

for some $C > 0$. Theorem 3.1.2 now follows from (3.57) on using the inequality $1 + x \leq \exp(x)$.

3.4 Application to Sárközy's Problem

We now proceed to derive Theorem 3.1.1 from Theorem 3.1.2 and the argument from [11] based on Sárközy's finite addition theorem, recalled in Subsection 2.2.4 of Chapter 2.

Let $\mathcal{P} = \cup_{1\leq i\leq K}\mathcal{P}_i$ be a colouring of the set of primes $\mathcal{P}$ and let $N \geq 1$ be an integer. Let us set $P_i = \mathcal{P}_i \cap (2N, 3N]$ for each i. Then there is an i, $1 \leq i \leq K$, such that $|P_i| \geq \frac{\pi^*(N)}{K}$, where as before $\pi^*(N)$ is the number of primes in $(2N, 3N]$.

We apply (3.1) to $S = P_i$ for an i as above. Assuming N to be sufficiently large and with M as in Theorem 3.1.2, we then see that there is a $C > 0$ such that

$$E_3(P_i) \ \leq \ \frac{M}{\phi(M)} \frac{|S|^5}{\log{(2N)}} \exp\left(\frac{C_1}{\log\log 4K} \right) \ \leq \ \frac{C|S|^5}{2\log N} \log\log 4K \ , \qquad (3.58)$$

By a standard application of the Cauchy-Schwarz inequality we have

$$|3P_i| \, E_3(P_i) \geq |P_i|^6 \ . \qquad (3.59)$$

Since $|P_i| \geq \frac{\pi^*(N)}{K} \geq \frac{N}{2K\log N}$ for all large enough N, by the Prime Number Theorem, we conclude from from (3.58) and (3.59) that

$$|3P_i| \geq \frac{N}{CK\log\log 4K} > \frac{N}{L} + 1 \ . \qquad (3.60)$$

where $L \geq 1$ is the smallest integer such that the second inequality holds. For all sufficiently large N there exists such an integer L satisfying $L \leq 2CK\log\log 4K$.

The set $3P_i$ is contained in the interval $(6N, 9N])$. Thus if A is the set $3P_i - 6N$ then A is contained in $[1, 3N]$ and satsifies

$$|A| > \frac{3N}{k} + 1 \ . \qquad (3.61)$$

with $k = 3L$. Sárközy's finite addition theorem applied to A then tells us that there is an arithmetical progression $\mathcal{A}$ of $3N$ terms contained in $3lP_i$ and of the form

$$6lN + \{(m+1)d, (m+2)d, \ldots, (m+3N)d\},$$

where the integers l and d satisfy $1 \leq d < 3L$ and $1 \leq l \leq 354L$. Also, since $3lP_i$ is contained in $(6lN, 9lN]$, we have that $0 \leq a \leq 3186LN$ for any term a of $\mathcal{A}$.

Let p be any element of P_i. Then p is a prime number in $(2N, 3N]$ and if $N \geq 3186L$, we surely have $d < p$ from the bound on d. In particular, p is coprime to d. Since the modulus of the arithmetical progression $\mathcal{A}$ is d and since it contains $3N \geq p$ terms, it follows that $\mathcal{A}$ contains a complete system of residue classes modulo p. Therefore for every integer n there is an a in $\mathcal{A}$ and an integer b such that $n = a + bp$.

We apply the conclusion of the preceding paragraph to the integers n in the interval $I(N) = [3186LN, 3187LN]$. For all such n we have $n = a + bp$ with $0 \leq b \leq 6374L$, since $0 \leq a \leq 3186LN$ and $\frac{N}{2} < p$. On recalling that a is in $3lP_i$ and p is in P_i, we conclude that every integer in the interval $I(N)$ can be written as the sum of no more than $9560L$ elements of P_i. Since $L \leq 2CK \log\log 4K$ this means that for each sufficiently large N there is an i such that every integer in $I(N)$ is the sum of no more than $19120CK \log\log 4K$ prime numbers, all lying in P_i. Finally, we observe that for any $N \geq 3186$, the interval $I(N)$ meets $I(N+1)$. Thus as N varies over all sufficiently large integers, the union of the intervals $I(N)$ contains all sufficiently large integers and we obtain Theorem 3.1.1.

CHAPTER 4

A Remark on the Beurling-Selberg function

4.1 Introduction

Conforming to notation introduced in Section 1.2 we write $K(z)$ to denote $\left(\frac{\sin \pi z}{\pi z}\right)^2$ and set $\operatorname{sgn}(z) = 1$ when $\operatorname{Re}(z) \geq 0$ and $\operatorname{sgn}(z) = -1$ when $\operatorname{Re}(z) < 0$, for any complex number z. Also, we recall that the Beurling-Selberg function B is the complex function by the relation

$$B(z) = 2zK(z) + \sum_{n \in \mathbf{Z}} \operatorname{sgn}(n) K(z-n) \ . \tag{4.1}$$

Since $z \mapsto zK(z)$ and $z \mapsto K(z)$ are entire functions with $|K(z)| \leq \frac{e^{2\pi|\operatorname{Im}(z)|}}{\pi^2|z|^2}$ for all complex $z \neq 0$, it is easily verified that the series in (4.1) converges normally on compact subsets of the complex plane and that B is an entire function of exponential type 2π. We have $K(0) = 1$ and $K(n) = 0$ for all other integers n. It follows that $B(n) = \operatorname{sgn}(n)$ for all integers n. Thus, B is an entire function of exponential type

2π that interpolates the values of $\operatorname{sgn}(z)$ at the integers.

The function B was originally examined in unpublished work of A. Beurling and was subsequently rediscovered by A. Selberg on account of its properties described by the following pair of theorems.

Theorem 4.1.1. *Let $I = [\alpha, \beta]$ be a compact real interval with characteristic function χ_I and let $F_I(z)$ denote $\frac{1}{2}(B(z-\alpha) + B(\beta - z))$. We then have the following.*

(i) $B(x) \geq \operatorname{sgn}(x)$ for all real x with equality if and only if x is an integer. Further, we have $\int_{\mathbf{R}} B(x) - \operatorname{sgn}(x)\, dx = 1$.

(ii) F_I is an integrable function on the real line satisfying $F_I(x) \geq \chi_I(x)$ for all real x and $\int_{\mathbf{R}} F_I(x)\, dx = \int_{\mathbf{R}} \chi_I(x)\, dx + 1$.

(iii) $\widehat{F}_I$ is supported in $[-1, 1]$.

Since for any integrable function f on the real line that majorises χ_I we have $\int_{\mathbf{R}} f(x)\, dx \geq \int_{\mathbf{R}} \chi_I(x)\, dx$, the value of $\int_{\mathbf{R}} F_I(x)$ given by (ii) is close to optimal. Selberg has remarked that when the interval I has integer end-points this value is in fact optimal. This is made precise by (ii) of the following theorem.

Theorem 4.1.2. *Let $I = [\alpha, \beta]$ be a compact real interval with integer end-points. We then have the following.*

(i) For every entire function $F(z)$ of exponential type 2π satisfying the condition that $F(x) \geq \operatorname{sgn}(x)$ for all real x we have $\int_{\mathbf{R}} F(x) - \operatorname{sgn}(x)\, dx \geq 1$ with equality only if $F(z) = B(z)$ for all complex z.

(ii) For every integrable function f on the real line such that $\widehat{f}$ is supported in $[-1,1]$ and $f(x) \geq \chi_I(x)$ for all real x, we have $\int_{\mathbf{R}} f(x)\,dx \geq \int_{\mathbf{R}} \chi_I(x)\,dx + 1$.

Part (i) of the above theorem was known to Beurling several years before Selberg, according to pages 225-226 of [24], and it tells us that not only is B an extremal majorant of the function $z \mapsto \operatorname{sgn}(z)$ on the real line but also that B is the unique entire function of exponential type 2π with this property. In contrast, it has been shown by Selberg (see [24]) that F_I is never unique in being an extremal majorant of χ_I for a compact real interval I.

The Beurling-Selberg function has a number of applications, especially to analytic number theory. For a survey of many of these applications, and indeed much else, we refer the reader to the well-known article of Vaaler on this subject [26]. In certain of these applications one requires a more detailed description of the properties of B than given by Theorem 4.1.1. This is supplied by Theorem 4.1.3 below, which we state, following [26], in terms $H(z)$, the odd part of $B(z)$.

By definition we have $H(z) = \frac{1}{2}(B(z) - B(-z))$. Also, we have from that (4.1) $K(z) = \frac{1}{2}(B(z) + B(-z))$, so that $K(z)$ is the even part of $B(z)$. Therefore for all complex numbers z we have $B(z) = H(z) + K(z)$. Further, we set $a(t) = 1 - |t|$ when $|t| < 1$ and $a(t) = 0$ when $|t| \geq 1$ and also let $c(t) = \operatorname{sgn}(t)$ when $|t| < 1$ and $c(t) = 0$ when $|t| \geq 1$.

Theorem 4.1.3. *Let us define complex functions h and J by $h(z) = H(z) - \operatorname{sgn}(z)$ and $J(z) = \frac{1}{2}H'(z)$ for any complex number z. Then we have the following.*

(i) For all real $x \neq 0$ we have $h(-x) = -h(x)$. Further, $-\frac{K(x)}{x+1} \leq h(x) \leq 0$ for $x > 0$. Therefore restriction of h is an integrable function on the real line.

(ii) $\widehat{h}(t) = \frac{1}{i\pi}(c(t) + \pi a(t)\cot(\pi t) - \frac{1}{t})$ *when* $t \neq -1, 0, 1$.

(iii) J is an even function on the real line satisfying $|J(x)| \leq \frac{2}{\pi^2 x^3}$ when $x > 0$. Therefore J is integrable on the real line and, in particular, H is of bounded variation on $\mathbf{R}$.

(iv) $\widehat{J}(t) = c(t)t + \pi t a(t)\cot \pi t$ *when* $t \neq -1, 0, 1$.

(v) $\widehat{J}$ is an even positive function on the real line supported in $[-1, 1]$ satisfying $\widehat{J}(0) = 1$. Also, $\widehat{J}$ is strictly decreasing on $[0, 1]$.

Theorem 4.1.3 combines Lemma 5, Theorem 6 and Corollary 7 on pages 191 to 193 of [26] except that Lemma 5 of [26] gives $-K(x) \leq h(x) \leq 0$ for $x > 0$ in place of the bound for $h(x)$ in (i). The bound given here reflects the actual decay rate of $h(x)$ at infinity.

An entire function ϕ defined by a series $\sum_{n\geq 0} a(z+n)$ that is normally convergent on compact subsets of the complex plane may be viewed as the unique entire solution to the difference equation $\Delta\phi(z) = a(z)$ satisfying the condition that $\phi(x+m)$ tends to 0 as m tends to $+\infty$ for every real number x, where $\Delta\phi(z)$ denotes $\phi(z) - \phi(z+1)$ for any complex z. Elaborating on our discussion in Section 1.2, we shall show in this chapter that this point of view provides simple proofs of the properties of the Beurling-Selberg function given by the above theorems. This is, in essence, on account of the fact that Δsgn is, up to a constant, the characteristic function of $[-1, 0)$.

We begin with a proof of Theorem 4.1.3 in Section 4.2. While it easily possible to deduce Theorem 4.1.1 from Theorem 4.1.3, it is, nevertheless, useful to have an independent proof of the former. We provide this, from our point of view, in Section 4.3. Finally, we take up Theorem 4.1.2 in Section 4.4.

Throughout this chapter we continue write $\Delta\phi(z)$ to denote $\phi(z)-\phi(z+1)$ and $\Delta(\phi(z))^2$ for $(\phi(z)-\phi(z+1))^2$, for any complex function ϕ and complex number z. We write $\Delta\phi$ to denote the function $z \mapsto \Delta\phi(z)$. Also, we will generally use x and t to denote real numbers and z for complex numbers.

4.2 A Difference Proof of Theorem 4.1.3

The proof of Theorem 4.1.3 given below may be compared with that on pages 16-18 of [9], through the exercises on page 136 of [7], and on pages 303 to 308 of [4], among others. All of these sources follow the method of Vaaler [26], pages 191 to 193. The main simplification in our account is in the proof of parts (ii) to (iv) of the theorem.

Since $B(z)=H(z)+K(z)$, the classical identity $1=\sum_{n\in Z} K(z-n)$ for all complex z and (4.1) give

$$H(z)+K(z)-1-2zK(z)=-2\sum_{n\geq 1} K(z+n)\ . \tag{4.2}$$

On applying Δ to both sides of (4.2) and observing that $\Delta(zK(z))$ and $K(z+1)$ can be written, respectively, as $\frac{\sin^2(\pi z)}{\pi^2 z(z+1)}$ and $\frac{\sin^2(\pi z)}{\pi^2(z+1)^2}$ we obtain that

$$\Delta H(z) = 2\Delta(zK(z)) - K(z) - K(z+1) = -\frac{\sin^2(\pi z)}{\pi^2}\Delta\left(\frac{1}{z}\right)^2 , \qquad (4.3)$$

where the last expression takes its meaning for $z = 0, -1$ by continuity. Also, if $\phi(z) = \sum_{n\geq 1} K(z+n)$ then we have $\phi(z+m) = \sum_{n>m} K(z+n)$ and $\phi'(z+m) = \sum_{n>m} K'(z+n)$, for any integer m. When m tends to $+\infty$, $\phi(z+m)$ and $\phi'(z+m)$ are thus sums of tails of convergent series and therefore tend to 0. We then conclude from (4.2) that, in particular, $H(x+m)$ and $J(x+m)$ tend, respectively, to 1 and 0 as m tends to $+\infty$, for any real x. We will now show that all parts of Theorem 4.1.3 may be read off from (4.3) and these limits.

For $x \geq 0$ we have $h(x) = \sum_{n\geq 0} \Delta H(x+n)$, since $H(x+m)$ tends to 1 as m tends to $+\infty$. The bounds for $h(x)$ in (i) follow from this relation and the inequalities $-\Delta(\frac{K(x)}{x+1}) \leq \Delta H(x) \leq 0$ for $x > 0$, which result from the second equality in (4.3) on remarking that $0 \leq \Delta(\frac{1}{x})^2 \leq \Delta(\frac{1}{x^2(x+1)})$ when $x > 0$. The other assertions in (i) are evident.

We have $\widehat{\Delta h}(t) = \widehat{h}(t)(1 - e(t))$. Also, $-\frac{1}{2}\Delta\mathrm{sgn}$ is the characteristic function of $[-1, 0)$ and $\widehat{\Delta\mathrm{sgn}}(t) = -\frac{1-e(t)}{i\pi t}$ when $t \neq 0$. The Fourier transform of $a(t)$ is $K(x)$ and that of $-\frac{c(t)}{2\pi i}$ is $xK(x)$. It then follows from the first equality in (4.3) that $\Delta H(x)$ is integrable on the real line and that $\widehat{\Delta H}(t) = \frac{c(t)}{\pi i}(1 - e(t)) - a(t)(1 + e(t))$. Since $-i\frac{1+e(t)}{1-e(t)} = \cot(\pi t)$, we obtain (ii) on passing to Fourier transforms in the relation $\Delta(h(x)) = \Delta(H(x)) - \Delta(\mathrm{sgn}(x))$ and dividing throughout the result by $1 - e(t)$ when t not an integer.

Since $\Delta J(x) = \frac{1}{2}\frac{d}{dx}\Delta H(x)$, we see using the second expression for $\Delta H(x)$ in (4.3) that $\Delta J(x)$ is given, for $x > 0$, by $-\frac{\sin(2\pi x)}{2\pi^2}\Delta(\frac{1}{x})^2 + \frac{\sin^2(\pi x)}{2\pi^2}\Delta(\frac{1}{x})\Delta(\frac{1}{x^2})$. Since both terms in this expression are majorised by $\frac{1}{\pi^2}\Delta(\frac{1}{x^3})$ we conclude that $|\Delta J(x)| \leq \frac{2}{\pi^2}\Delta(\frac{1}{x^3})$ for $x > 0$. The bound for $J(x)$ in (iii) follows from this inequality combined with the relation $J(x) = \sum_{n\geq 0} \Delta J(x+n)$ by means of the triangle inequality. This last equality holds for any real x because $J(x+m)$ tends to 0 as $m \to +\infty$.

Since $\Delta H(x)$ tends to 0 as x tends to $\pm\infty$, an integration by parts shows that the Fourier transform of $\frac{d}{dx}\Delta H(x)$ is $2i\pi t\widehat{\Delta H}(t)$. Consequently, we obtain $2\widehat{J}(t)(1 - e(t)) = 2i\pi t\widehat{\Delta H}(t)$ from which the expression for $\widehat{J}(t)$ in (iv) follows on dividing by $1 - e(t)$ when t is not an integer.

The expression for $\widehat{J}(t)$ in (iv) immediately shows that $\widehat{J}$ is an even function on the real line supported in $[-1, 1]$ and that $\widehat{J}(0) = 1$. Also, we have that $\widehat{J}(t) + \widehat{J}(1 - t) = 1$ for t in $[0, 1]$. To complete the proof of (v) it thus suffices to verify that $\widehat{J}(t)$ is decreasing on $(0, 1/2)$. On this interval $\widehat{J}(t)$ is given by $t + \pi t(1-t)\cot \pi t$ and, on writing $u(t)$ for its derivative, we see that $\sin^2(\pi t)u(t) = \sin^2(\pi t) + \pi(1 - 2t)\sin \pi t \cos \pi t - \pi^2 t(1-t)$. Since the right hand side of this relation is negative on account of the inequalties $\sin \pi t < \pi t$ and $\cos \pi t < 1$, it follows that $\widehat{J}(t)$ is strictly decreasing for t in $(0, 1/2)$.

4.3 Selberg's Majorant

The proof of Theorem 4.1.1 given below follows the same method as in the previous section. Our method differs from that on pages 153 and 154 of [10], 558 and 559 of [15], 67 and 68 of [25], 307 and 308 of [4] or pages 16 to 18 of [3] mainly in the proof of (iii) of Theorem 4.1.1, for which, as we have remarked in Section 1.2, these sources either quote the Paley-Weiner Theorem or else use the well-known contour integral argument underlying the proof of this theorem.

By means of the classical identity $\sum_{n\in Z} K(z-n) = 1$ valid for all complex z we see that

$$B(z) - 1 - 2zK(z) = -2\sum_{n\in Z} K(z+n) \tag{4.4}$$

Applying Δ to both sides of this relation we obtain

$$\Delta(B(z)) = 2\Delta(zK(z)) - 2K(z+1) = \frac{2\sin^2(\pi z)}{\pi^2 z(z+1)^2} \tag{4.5}$$

for all complex z. We also have from (4.4) that $B(x+m)$ tends, respectively, to 1 and -1 as m tends to $\pm\infty$, for any real x.

As in Section 1.2, let us set $b(z) = B(z) - \operatorname{sgn}(z)$. It follows from the preceding limits that we have $b(x) = \sum_{n\geq 0} \Delta B(x+n)$ for $x \geq 0$ and similarly that we have $b(x) = -\sum_{n\geq 1} \Delta B(x-n)$ for $x < 0$. From the second equality in (4.5) we see that $\Delta B(x) \geq 0$ for $x \geq 0$ and $\Delta B(x) \leq 0$ for $x < 0$ with equality in these inequalities only when x is an integer. We then conclude that $b(x) \geq 0$ for all real x with equality only when x is an integer.

Since the even part of $B(z)$ is $K(z)$ it follows that $b(x) + b(-x) = 2K(x)$ for all real $x \neq 0$. Since $b(x) \geq 0$ for all real x and $x \mapsto K(x)$ is integrable on $\mathbf{R}$, we conclude from this relation that b is integrable on $\mathbf{R}$ as well. Also, on integrating both sides of this relation we get $\int_{\mathbf{R}} b(x)dx = \int_{\mathbf{R}} K(x)dx = 1$. This completes the proof of (i).

The characteristic function χ_I of I satisfies $\chi_I(x) = \frac{1}{2}(\mathrm{sgn}(x-\alpha) + \mathrm{sgn}(\beta - x))$ for all real x. Therefore we have $F_I(x) = \chi_I(x) + \frac{1}{2}(b(x-\alpha) + b(\beta - x))$ for all real x, from which and the properties of $b(x)$ verified above we immediately conclude (ii).

From the first equality in (4.5) we see that $x \mapsto \Delta B(x)$ is integrable on $\mathbf{R}$. Since the Fourier transform of $x \mapsto \Delta(xK(x))$ is $c(t)(1-e(t))$ and that of $x \mapsto K(x+1)$ is $a(t)e(t)$, we also see that $\widehat{\Delta B}$ is supported in $[-1,1]$. Since from the definition of F_I we have $\Delta(F_I(x)) = \frac{1}{2}(\Delta B(x-\alpha) + \Delta B(\beta - x))$, we conclude that $\widehat{\Delta F_I}$ is supported in $[-1,1]$ as well. Finally, since $\widehat{F}_I(t)(1-e(t))) = \widehat{\Delta F_I}(t))$ for all real t, we obtain (iii) on dividing by $1-e(t)$ when t is not an integer.

4.4 Extremality

In its outline, the proof of Theorem 1.2 given below is the same as that of Theorem 8 on pages 194 and 190 of [26]. Our presentation, however, bypasses a number of technical points that arise in [26]. For instance, we will not use the Plancherel-Polya Theorem. Also, the proof below explains why it is natural to view $B(z)$ as the solution of a difference equation.

We begin with the following simple remark. If $g_0(x) \leq g_1(x) \leq \ldots$ is an increasing sequence of continuous function on $[0,1)$ that converges in $\mathrm{L}^1[0,1)$ to a function $g(x)$ that is continuous on $[0,1)$ then we have

$$g_k(x) \leq g(x) \text{ for all } x \text{ in } [0,1). \tag{4.6}$$

Indeed, if x in $[0,1)$ and $\epsilon > 0$ are such that $x+\epsilon$ is also in $[0,1)$ then our hypotheses imply that $\int_x^{x+\epsilon} g_k(u)\,du \leq \int_x^{x+\epsilon} g(u)\,du$, for all $k \geq 0$. Dividing by ϵ, letting ϵ tend to 0 and noting that g_k and g are continuous at x, we get (4.6).

Let us now verify (*ii*) of Theorem 4.1.2. For any integer $k \geq 0$ let $g_k(x)$ denote $\sum_{|n|\leq k} f(x+n)$. Since $\widehat{f}(t)$ is compactly supported, the Fourier inversion formula exhibits f as a Fourier transform and thus f is continuous on the real line. Since $f(x) \geq \chi_I(x)$ for all real x, f is a positive function. Thus $\{g_k\}$ is an increasing sequence of continuous functions on $[0,1)$. Since f is integrable on $\mathbf{R}$ and since

$$\int_0^1 g_k(x)\,dx = \int_{-k}^{k} f(x)\,dx \tag{4.7}$$

for all $k \geq 1$, we have that $g_k(x)$ is a Cauchy sequence in $\mathrm{L}^1[0,1)$. If $g(x)$ is its limit, then the n-th Fourier coefficient of g is $\widehat{f}(n)$, and since $\widehat{f}(n) = 0$ for $n \neq 0$, it follows that $g(x) = \widehat{f}(0)$ for all x in $[0,1)$. Applying (4.6) we get we conclude that $\sum_{n\in\mathbf{Z}} f(x+n) \leq \widehat{f}(0)$ for all x in $[0,1)$.

Putting $x = 0$ in the preceding inequality we have $\sum_{n\in\mathbf{Z}} f(n) \leq \widehat{f}(0)$. This implies $\widehat{f}(0) \geq \beta-\alpha+1$ since we have $f(n) \geq 1$ for all integers n in I and since $\beta-\alpha+1$ is the number of integers in $I = [\alpha,\beta]$ when α and β are integers. Since $\int_{\mathbf{R}} f(x)\,dx = \widehat{f}(0)$ and $\int_{\mathbf{R}} \chi_I(x)\,dx = \beta - \alpha$, we obtain (*ii*) of Theorem 4.1.2.

Turning to (*i*) of Theorem 4.1.2, let $F(z)$ be an entire function of exponential type 2π

such that $F(x) \geq \operatorname{sgn}(x)$ for all x in $\mathbf{R}$. We shall hereafter set $f(x) = F(x) - \operatorname{sgn}(x)$. Thus f is a positive function on the real line and we may assume that $\int_{\mathbf{R}} f(x)\, dx$ is finite. Then $\Delta f(x)$ is integrable on the real line and, since $\Delta \operatorname{sgn}(x)$ is evidently integrable, so is $\Delta F(x)$. Since $\Delta F(z)$ is also an entire function of exponential type 2π, the Paley-Weiner Theorem tells us that $\widehat{\Delta F}(t)$ is a continuous function supported in $[-1, 1]$. Thus on passing to Fourier transforms in the relation $\Delta f(x) = \Delta F(x) - \Delta \operatorname{sgn}(x)$ and dividing throughout the result by $1 - e(t)$ when t is not an integer, we conclude that $\widehat{f}(t) = -\frac{1}{i\pi t}$ when $|t| \geq 1$.

For every integer n we therefore have that $\widehat{f}(n)$ is the same as n-th Fourier coefficient of the function $x \mapsto \widehat{f}(0) + 2x - 1$ in $\mathrm{L}^1[0, 1)$. Arguing as in the proof of (ii) by means of (4.6) above we then conclude that

$$\sum_{n \in \mathbf{Z}} f(x + n) \leq \widehat{f}(0) + 2x - 1 \tag{4.8}$$

for all x in $[0, 1)$. Putting $x = 0$ and noting that the left hand side of this inequality is positive, we obtain $\int_{\mathbf{R}} f(x)\, dx = \widehat{f}(0) \geq 1$.

From (4.8) we certainly have that $\sum_{n \in \mathbf{Z}} f(x + n)$ converges for every x in $[0, 1)$. Therefore $f(x + n)$ tends to 0 as n tends to $\pm\infty$, for any x in $[0, 1)$ and hence for all real x. Thus, for any real x, $F(x + m)$ tends, respectively, to 1 and -1 as m tends to $\pm\infty$. We deduce from these limits that $f(x) = \sum_{n \geq 0} \Delta F(x + n)$ when $x \geq 0$ and $f(x) = -\sum_{n \geq 1} \Delta F(x - n)$ when $x < 0$.

Suppose now that $F(z)$ is such that $\widehat{f}(0) = 1$. Then we have $\sum_{n \in \mathbf{Z}} f(n) = 0$. Since f is positive on the real line and differentiable when $x \neq 0$, it follows that $f(n) = 0$

for all integers n. This in turn implies that f has a local minimum at each integer n. Since f is differentiable when $x \neq 0$ we deduce that $f'(n) = 0$ when $n \neq 0$. We then conclude that $\Delta F(-1) = -2$, $\Delta F'(-1) = -F'(0)$ and $\Delta F'(0) = F'(0)$. Moreover, that $\Delta F(n) = \Delta F'(n) = 0$ for all other integers n.

We will now show that $\Delta F(z) = 2\Delta(zK(z)) - 2K(z+1)$ for all complex z, from which it follows that $F(z) = B(z)$. Indeed, on account of (4.5) this implies that $\Delta F(x) = \Delta B(x)$ so that $p(x) = F(x) - B(x)$ is a periodic function of period 1. Since, for any real x, $p(x+m)$ tends to 0 as m tends to $+\infty$, it follows that $p(x) = 0$ and hence that $F(z) = B(z)$.

Let us set $G(z) = \Delta F(z) - F'(0)\Delta(zK(z)) - 2K(z+1)$. Then $G(z)$ is an entire function such that $G(x)$ is the Fourier transform of a continuous function $\phi(t)$ supported in $[-1, 1]$. Thus $G'(x)$ is the Fourier transform of $-2i\pi t\phi(t)$. Consequently, the n-th Fourier coefficient of the periodisations $\sum_n \phi(t+n)$ and $-2\pi i \sum_n (t+n)\phi(t+n)$ are, respectively, $G(n)$ and $G'(n)$, both of which are 0 for each integer n. It follows that

$$\sum_n \phi(t+n) = 0 \text{ and } \sum_n (t+n)\phi(t+n) = 0 \text{ for all real } t. \tag{4.9}$$

Let t in $[-1, 1]$ be distinct from -1, 0 and 1. Then there is a unique integer $m \neq 0$ such that $t+m$ is in $[-1, 1]$. On multiplying the first of the above relations by $t+m$ and subtracting the second relation, we obtain $\phi(t) = 0$ on noting that $\phi(t+n) = 0$ when n is distinct from 0 and m. Since ϕ is a continuous function we conclude that $\phi(t) = 0$ for all real t. Consequently, $G(z) = 0$ for all complex z. In other words, $\Delta F(z) = F'(0)\Delta(zK(z)) - 2K(z+1)$ for all complex z. This is relation may also be written as $\Delta F(z) = \frac{2\sin^2(\pi z)((F'(0)-2)z+F'(0))}{\pi^2 z(z+1)^2}$ for all complex z.

Suppose that $F'(0) \neq 2$ and let x be a real number that is not an integer. If x is sufficiently large, the last equality of the preceding paragraph tells us that $\Delta F(x+n)$ and $\Delta F(-x-n)$ are distinct from 0 and have the same sign as $F'(0)-2$ for all integers $n \geq 0$. The relations $f(x) = \sum_{n\geq 0} \Delta F(x+n)$ and $f(-x) = -\sum_{n\geq 1} \Delta F(-x-n)$ then show that $f(x)$ and $f(-x)$ are distinct from 0 and have opposite signs. This is absurd since $f(x)$ is positive for all real x. Consequently, we have $F'(0) = 2$ and the proof of (i) of Theorem 4.1.2 is also complete.

Bibliography

[1] P. Akhilesh and D.S. Ramana, A chromatic version of Lagrange's four squares theorem, *Monatsh. Math.* **176** (2015), no. 1, 17-29.

[2] P. Akhilesh and D.S. Ramana, A remark on the Beurling-Selberg function, *Acta Math. Hungar.* **139** (2013), no. 4, 354-362.

[3] R.C. Baker, *Diophantine Inequalities*, London Mathematical Society Monographs **1**, Oxford University Press (1986).

[4] P.T. Bateman and H. Diamond, *Analytic Number Theory*, World Scientific Publishing Co (2007).

[5] J. Bourgain, On $\Lambda(p)$-subsets of squares, *Israel J. Math.* **67** (1989), no. 3, 291-311.

[6] J. Brüdern, A problem in additive number theory, *Math. Proc. Cambridge Philos. Soc.* **103** (1988), 27-33.

[7] H. Cohen, *Number Theory, Volume II*, Graduate Texts in Mathematics **240**, Springer-Verlag (2007).

[8] H. Davenport, T.D. Browning, *Analytic Methods for Diophantine Equations and Diophantine Inequalities*, Cambridge Mathematics Library, CUP (2005).

[9] M. Drmota and R.F. Tichy, *Sequences, Discrepancies and Applications*, Lecture Notes in Mathematics **1651**, Springer-Verlag (1997).

[10] J. Friedlander and H. Iwaniec, *Opera de Cribro*, Colloquium Publications **57**, AMS (2010).

[11] N. Hegyvári and F. Hennecart, On monochromatic sums of squares and primes, *J. Number Theory* **124** (2007), 314-324.

[12] H. Iwaniec and I.Kowalski, Analytic Number Theory, American Mathematical Society Colloquium Publications 53, A.M.S., 2004.

[13] V. Lev, Optimal representation by sumsets and subsets sums, *J. Number Theory* **62** (1997), 127-143.

[14] K. Matomäki, Sums of positive density subsets of the primes, *Acta Arith.* **159** (2013), 201-225.

[15] H.L. Montgomery, The analytic principle of the large sieve, *Bull. Amer. Math. Soc.* **84** (1978), 547-567.

[16] H.L. Montgomery, *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis*, Regional Conference Series in Mathematics **84**, CBMS (1994).

[17] G. Prakash and D.S. Ramana, The large sieve inequality for integer polynomial amplitudes, *J. Number Theory* **129** (2009), 428-433.

[18] D.S. Ramana and O. Ramaré, Additive energy of dense sets of primes and monochromatic sums, *Israel J. Math.*, **199** (2014), 955-974.

[19] O. Ramaré, *Arithmetical aspects of the large sieve inequality. With the collaboration of D. S. Ramana*, Harish-Chandra Research Institute Lecture Notes **1**, Hindustan Book Agency, New Delhi (2009).

[20] O. Ramaré and I.Z. Ruzsa, Additive properties of dense subsets of sifted sequences, *J. de théorie des nombres de Bordeaux* **13** (2001), no. 2, 559-581.

[21] B. Rosser and I. Schoenfeld, Approximate formulas for some functions of prime numbers, *Illinois J. Math.* **6** (1962), no. 1, 64-94.

[22] A. Sárközy, Finite addition theorem I, *J. Number Theory* **48** (1994), 197-218.

[23] A. Sárközy, Unsolved problems in number theory, *Period. Math. Hungar.* **42** (2001), 17-35.

[24] A. Selberg, *Collected Papers, Volume II*, Springer-Verlag (1991).

[25] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Studies in Advanced Mathematics **46**, Cambridge University Press (1995).

[26] J.D. Vaaler, Some extremal functions in Fourier analysis, *Bull. Amer. Math. Soc.* **12** (1985), 183-216.

[27] L. Zhao, Large sieve inequality with quadratic amplitudes, *Monatsh. Math.* **151** (2007), 165-173.