APPLICATION OF THE CIRCLE METHOD IN FIVE NUMBER THEORY PROBLEMS

A Dissertation
Presented to
The Academic Faculty

By

Seyyed Hamed Mousavi

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology

August 2022

APPLICATION OF THE CIRCLE METHOD IN FIVE NUMBER THEORY PROBLEMS

Thesis committee:

Dr. Ernie Croot
School of Mathematics
Georgia Institute of Technology

Dr. Ben Krause
Department of Mathematics
King's College Lonon

Dr. Michael Lacey
School of Mathematics
Georgia Institute of Technology

Dr. Larry Rolen
School of Mathematics
Vanderbilt University

Dr. Konstantin Tikhomirov
School of Mathematics
Georgia Institute of Technology

Date approved: July 15, 2022

The degree to which I can create relationships, which facilitate the growth of others as separate persons, is a measure of the growth I have achieved in myself.

Carl.R.Rogers

To my parents and my sisters

ACKNOWLEDGMENTS

I would like to thank my advisor Prof. Ernie Croot for all his help over these five years. He taught me most of my knowledge in analytic number theory, and reviewed and corrected many reports that I have written. Moreover, he frequently helped me to see the big picture in our projects. I also wish to thank Prof. Michael Lacey for his constant support and for giving me the opportunity to study discrete harmonic analysis. He was always very kind and helpful and provided a very relaxed environment for me to do research. I learned a lot of deep concepts in harmonic analysis from him. I want to express my deepest gratitude to Dr. Ben Krause, my future Postdoc supervisor, for his patience and for inviting me to participate in a seminar series on discrete harmonic analysis. I enjoyed many hours of discussing various topics on the subject with him, from basic notions to state of art problems. I learned most of my probability knowledge from Dr. Konstantin Tikhomirov, for which I am extremely thankful. I also appreciate his aid in my Postdoc application and Ph.D. defense. Furthermore, I would like to express my appreciation to Dr. Larry Rolen for his excellent class about modular forms, in which he gave a summary of important parts of his book, and also for accepting to be on my defense committee.

I am forever indebted to my parents, who have been constantly supporting me in different stages of my life. They frequently decided to sacrifice their happiness to give me mine. I am extremely lucky to have my sisters Arefeh, Tayebe, Mohaddesseh, and Foroogh, and most grateful for their emotional support and encouragement. They have been a constant light in my life, even on the most difficult days. Special thanks to my brothers-in-law, Mehdi and Hamid, for their support in these years. I am extremely grateful to have my niece Maryam, and my nephew, Mohammad Hossein, whom I love unconditionally. I never would have been able to get my Ph.D. without the support of my family.

I am very thankful to be a master's student of Dr. Javad Haghighat at Shiraz University of Technology, and Dr. Kagan Kursungoz at Sabanci University. They helped me to understand many basic notions and supported me to get a chance of studying at Georgia Tech for my Ph.D. I wish to express my gratitude to Dr. Ali Rajaee, my supervisor in Tarbiat Modares, and my co-authors Prof. Ahmad Mousavi and Dr. Ayesha Qureshi at Sabanci University. Many thanks to Farshid Pahlavan, Dr. Mohammad Hassan Shirdarreh, Dr. Babak Amini, and Dr. Afshin Amini for their aid and inspiration to participate in mathematics competitions, which played a crucial role for me to continue studying mathematics. I extremely appreciate all the support from Dr. Morag Burke, both for my postdoc applications and her consultation throughout my Ph.D. I appreciate my ten years of friendship with Dr. Hojjat Farzadfar for introducing analytic number theory to me, and for all his support these years.

Last but not the least, I would like to thank my friends Reza Dastbasteh, Farzad Padashnik, Yaghoub Rahimi, Amin Bayat Barooni, Babak Ahmadi, Mohammad Moein, Nima Rasekh, Parivash Tahani, Pourya Salehi, Zohreh Ghorbani, Yashar Kiarashinejad, Maxie Schmidt, Christina Giannitsi, Majid Mousavi, Mostafa Mousavi, and many other people who helped, encouraged and motivated me in different places of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis consists of three applications of the circle method in number theory problems. In the second chapter, we study a question of Graham. Are there infinitely many integers n for which the central binomial coefficient $\binom{2n}{n}$ is relatively prime to $105 = 3 \cdot 5 \cdot 7$? By Kummer's Theorem, this is the same as asking if there are infinitely many integers n, so that n added to itself base 3, 5, or 7, has no carries. A probabilistic heuristic of Pommerance predicts that there should be infinitely many such integers n. We establish a result of a statistical nature supporting Pommerance's heuristic. The proof consists of a Fourier analysis method, as well as geometrically bypassing an old conjecture about the primes.

In the third chapter, we discover an unexpected cancellation on the sums involving exponential functions. Applying this theorem on the first terms of the Ramanujan-Hardy-Rademacher expansion for the partition function gives us a natural proof of a "weak" pentagonal number theorem. We find several similar upper bounds for many different partition functions. Additionally, we prove another set of "weak" pentagonal number theorems for the primes, which allows us to count the number of primes in certain intervals with small error. Finally, we show an approximate solution to the Prouhet-Tarry-Escott problem using a similar technique. The core of the proofs is an involved circle method argument.

The fourth chapter of this thesis is about an endpoint scale independent ℓ^p-improving inequality for averages over the prime numbers. The primes are almost full-dimensional, hence one expects improving estimates for all $p > 1$. Those are known, and relatively easy to establish. The endpoint estimates are far more involved however, engaging for instance Siegel zeros, in the unconditional case, and the Generalized Riemann Hypothesis (GRH) in the general case. Assuming GRH, we prove the sharpest possible bound up to a constant. Unconditionally, we prove the same inequality up to a logarithmic factor. The proof is based on a circle method argument, and utilizing smooth numbers to gain additional control of Ramanujan sums.

CHAPTER 1

INTRODUCTION

In this dissertation, we use the notation $f(x) \simeq g(x)$, which means that there exists $C > 0$ such that

$$\limsup_{x\to\infty} \frac{|f(x)|}{|g(x)|} = C;$$

In particular, we show $f \sim g$ if $C = 1$. We sometimes need to use $f(x) = O(g(x))$ or $f(x) \lesssim g(x)$, which both of them mean that

$$\limsup_{x\to\infty} \frac{|f(x)|}{|g(x)|} < \infty.$$

Also we say $f(x) = o(g(x))$ if

$$\liminf_{x\to\infty} \frac{|f(x)|}{|g(x)|} = 0.$$

and $f(x) = \omega(g(x))$ if $g(x) = o(f(x))$. Finally we would like to emphasize that the constants in these notations are functions of a specific variable like y. In that case we write them like $\simeq_y$, $\lesssim_y$, $O_y(f(x))$, ...

1.1 On the Distribution of Prime Numbers

One of the major results involving the distribution of primes is the *Prime Number Theorem*, which in its simplest form states that

$$\pi(x) \sim \frac{x}{\log(x)} \tag{1.1}$$

where $\pi(x)$ is the number of primes less than or equal to x. We need the definition of the *Riemann zeta function* for the rest of the argument. Define the following function for $Re(s) > 1$:

$$\zeta(s) := \sum_{n \in \mathbb{N}} \frac{1}{n^s}$$

In order to extend the definition to the whole complex plane, we can use a specific explicit formula, which is explained later, and use analytic continuation to extend the definition up to $Re(s) \geqslant 1/2$. To further expand the definition to the whole complex plane, define

$$\xi(s) := \frac{1}{2}s(s-1)\pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})\zeta(s). \tag{1.2}$$

It is known that we have the functional equation:

$$\xi(s) = \xi(1-s),$$

we can use analytic continuation to expand the definition of ζ to the whole plane. Note that ζ has a simple pole at $s = 1$ and trivial zeroes at $\{-2n\}$ for $n \in \mathbb{N}$. The statement (1.1) is basically the same as saying the Riemann Zeta function $\zeta(s)$ does not have zero on the line $Re(s) = 1$. A stronger result gives

$$\Psi_0(x) := \sum_{n<x} \Lambda(n) + \frac{1}{2}\Lambda(x) = x + O(xe^{-c\sqrt{\log(x)}}) \text{ for a universal constant } c > 0 \tag{1.3}$$

and Λ is the *Von Mangoldt function*. To prove this result we need to show that there exists a *zero-free* region in the critical strip $0 < Re(s) < 1$. One can show that the zero-free region for $\zeta(\sigma + it)$ is

$$\sigma \geqslant 1 - \frac{c}{\log(|t| + 2)} \text{ for some } c > 0$$

Aside from this zero-free region and counting the number of zeroes of the zeta function, we also need another ingredient to prove (1.3). We can prove the following *explicit formula* for Ψ_0 in terms of ζ function:

$$\Psi_0(x) = x - \sum_{Im(\rho)<T} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log(1-x^{-2}) + O\left((\log x)\min\left(\frac{x}{T \parallel x \parallel}, 1\right) + \frac{x\log^2(xT)}{T}\right).$$

where the sum is over the nontrivial zeroes ρ of the Riemann zeta function. Related to this matter, we have the ***Riemann Hypothesis (RH)***, which assumes that all zeroes of the zeta function $\zeta(s)$ lie on the critical line $Re(s) = \frac{1}{2}$. Assuming the RH we can show that

$$\Psi(x) = x + O(x^{\frac{1}{2}}\log x).$$

In a more general case, we can define the $L-$function $L(s,\chi)$ to be

$$L(s,\chi) := \sum_{n\in\mathbb{N}} \frac{\chi(n)}{n^s} \text{ for } Re(s) > 1$$

where χ is a *Dirichlet Character*. We can use analytic continuation to extend $L(s,\chi)$ to the whole complex plane for every nonprincipal character χ. Zeroes of $L-$functions are important as they have a direct relation with the distribution of prime numbers in arithmetic progressions. A well-known result states that there are no zeroes in the region

$$\sigma \geqslant 1 - \frac{c}{\log q(1+|t|)} \text{ for some } c > 0$$

for just one exception in some cases of q. This special zero is called a *Siegel zero* and the character χ_q is called the *exceptional character*. An application of studying the zeroes of

$L(s,\chi)$ is the *Siegel–Walfisz Theorem*, which states that

$$\Psi(x;q,r) := \sum_{\substack{n<x \\ n\equiv r \pmod q}} \Lambda(n) + \frac{1}{2}\Lambda(x)1_{x\equiv r} = \frac{x}{\phi(q)} + \frac{x^{\beta}\chi(x)}{\phi(q)\beta} + O(qxe^{-c\sqrt{\log(x)}})$$

where β is the Siegel zero. Note that exceptional characters happen rarely. In fact, if q_n is the n^{th} integer with an exceptional character, then there exists $C > 0$ such that $q_{n+1} > q_n^C$. The *Generalized Riemann Hypothesis (GRH)* assumes that the non-trivial zeroes of $L(s,\chi)$ are living on the line $Re(s) = \frac{1}{2}$. Using the GRH one can show that

$$\Psi(x;q,r) = \frac{x}{\phi(q)} + O\left(x^{\frac{1}{2}}(\log qx)^2\right).$$

Obviously in this case, we do not have any Siegel zeroes. There are more generalized categories of $L-$functions, whose definitions we will not provide here.

Related to the distribution of prime numbers, an integer x is called $y-$*smooth*, if all of the prime factors of x are less than or equal to y. In other words, if $p|x$, then $p \leqslant y$. An important property of the smooth numbers that we will use in chapter 4 is their multiplicative property. That is, if x, z are $y-$smooth numbers, then xz is also a $y-$smooth number.

One of the central questions in this area is to count the number of the $y-$smooth numbers $x < N$, which is denoted by $\Psi(N,y)$ (see [1]). Note that $\Psi(N,y) \leqslant N$, and if $y > N$, then obviously $\Psi(N,y) = N$. *Dickman* in [2] showed that for any fixed $u \geqslant 1$, we have the following estimate for $\Psi(x,y)$:

$$\Psi(x,y) \sim x\rho(u) \text{ where } y = x^{1/u}$$

and $\rho(y)$ is called the *Dickman-De Brujin function* and is non zero as $x \to \infty$. It is obvious that $\rho(u) = 1$ for $0 \leqslant u \leqslant 1$ and it is shown that

$$\rho(u) = 1 - \log(u) \text{ for } 1 \leqslant u \leqslant 2$$

and

$$\rho(u) = \frac{1}{u} \int_{u-1}^{u} \rho(t) dt. \tag{1.4}$$

The *sieve methods* produce the estimate

$$\Psi(x, x^{1/u}) = x(1 - \log u) + o(x) \text{ for } 1 \leqslant u \leqslant 2.$$

Rankin proved the upper bound

$$\Psi(x, (\log x)^A) = x^{1 - \frac{1}{A} + O(1/\log\log x)}. \tag{1.5}$$

Assuming the Riemann hypothesis it is shown in [3] that

$$\Psi(x, x^{1/u}) = x\rho(u)\left(1 + O\left(\frac{\log(u+1)}{\log y}\right)\right) \text{ for } y \geqslant (\log x)^{2+\epsilon}$$

Another interesting problem in this area is to count the number of smooth numbers in short intervals. A challenge problem is to prove that

$$\Psi(x + c\sqrt{x}, x^\alpha) - \Psi(x, x^\alpha) > 0 \text{ for all } \alpha > 0 \text{ and large } x > 0 \tag{1.6}$$

Croot in [4] showed that one can get (1.6) for $\alpha = \frac{3}{14\sqrt{e}} + \epsilon$. He in fact proved the following lower bound:

$$\Psi(x + c\sqrt{x}, x^\alpha) - \Psi(x, x^\alpha) \gg \frac{\sqrt{x}}{(\log x)^{\log 4 + \epsilon}}$$

This bound has been improved in [5]. Soundarajan in [6] proved the challenge problem (1.6) assuming the RH.

1.2 Partition Theory

Theory of *Partitions* has been a subject of interest in mathematics for centuries. We denote the partition function $p(n)$ as the number of representations of the positive integer n as sum of increasing positive integers. For example $p(4) = 5$; since $4 = 1+1+1+1$ or $4 = 1+3$ or $4 = 1+1+2$ or $4 = 2+2$ or $4 = 4$ ($1+3$ and $3+1$ are considered to be the same). The generating function of the number of partitions is:

$$F(q) := \sum_{n=0}^{\infty} p(n)q^n = \prod_{n=1}^{\infty}\left(\sum_{m=0}^{\infty} q^{mn}\right) = \prod_{n=1}^{\infty}\frac{1}{1-q^n} \text{ for } |q| < 1.$$

One of the main properties of partition functions is that they satisfy the *Pentagonal Number Theorem*. Let $G_n = \frac{n(3n+1)}{2}$ be the pentagonal numbers, then

$$\sum_{G_n \leqslant x} (-1)^n p(x - G_n) = 0$$

In other words, the Pentagonal Number theorem states that

$$(q;q)_{\infty} := \prod_{n=1}^{\infty}(1-q^n) = 1 + \sum_{k=1}^{\infty}(-1)^k\left(q^{\frac{k(3k+1)}{2}} + q^{\frac{k(3k-1)}{2}}\right) \tag{1.7}$$

There are different proofs for this combinatorial property (see, for example, [7]). An exact formula for $p(n)$ is a well-known result due to *Rademacher*, *Hardy*, and *Ramanujan*.

$$p(n) = \frac{1}{\pi\sqrt{2}}\sum_{k=1}^{\infty}\left(\sqrt{k}\left(\sum_{\substack{0\leqslant h<k\\(h,k)=1}}\omega(h,k)e^{-\frac{2\pi i h n}{k}}\right)\frac{d}{dx}\left(\frac{\sinh\left(\frac{\pi}{k}\sqrt{\frac{2}{3}(x-\frac{1}{24})}\right)}{\sqrt{x-\frac{1}{24}}}\right)\Bigg|_{x=n}\right), \tag{1.8}$$

where $\omega(h,k)$ is a sum over some roots of unity.

Aside from the usual partitions, there are variants of it which have been studied extensively. For example, let $Q(n)$ be the number of partitions of n with odd parts. An exact formula for $Q(n)$ has been proved in [8]. Simply speaking, $Q(n)$ can be approximated as

$$Q(n) = A\frac{d}{dx}\left(I_0\left(\pi\sqrt{\frac{x+\frac{1}{24}}{3}}\right)\right)\Bigg|_{x=n} + O(\sqrt{Q(n)}).$$

where A is a certain constant and I_0 is the second type *Bessel function* with degree zero. More generally, let $p(n;\alpha,M)$ be the number of partitions with parts of the form $Mt \pm \alpha$, $1 \leqslant \alpha \leqslant M-1$, and $(\alpha, M) = 1$. It is shown in [9, Theorem 4] that $p(n;\alpha,M)$ has the following form:

$$p(n;\alpha,M) = \frac{\pi\csc(\frac{\pi\alpha}{M}) I_1\left(\frac{\pi\sqrt{12Mn-6\alpha^2+6M\alpha-M^2}}{3M}\right)}{\sqrt{12Mn-6\alpha^2+6M\alpha-M^2}} + O(e^{\frac{\pi\sqrt{n}}{\sqrt{3M}}}). \tag{1.9}$$

1.3 Vinogradov Mean Value Theorem

An old problem in number theory proposed by *Waring* asks for the number of ways that one can write an integer n into the sum of k^{th} powers. In other words, it asks for $r_{n,k}(x)$ the number of solutions for the following equation:

$$x = a_1^k + a_2^k + \cdots a_n^k \text{ for } a_i \in \mathbb{N} \cup \{0\}.$$

Related to this problem, one might ask for the number of solutions $J_{n,k}(N)$ for the following system of equations:

$$a_1^r + \cdots + a_n^r = b_1^r + \cdots + b_n^r \text{ where } 0 \leqslant a_i, b_i \leqslant N \text{ and } 1 \leqslant r \leqslant k \tag{1.10}$$

The main conjecture of this topic is to prove that for all $n,k \geqslant 1$ and all large N and every

$\epsilon > 0$

$$J_{n,k}(N) \lesssim_{n,k,\epsilon} N^{\epsilon}\left(N^{n} + N^{2n-\frac{1}{2}k(k+1)}\right). \tag{1.11}$$

Obviously there are around N^n trivial solutions where $\{a_i\} = \{b_i\}$. The nontrivial cases of the upper bound (1.11) occurs when $k^2 < 2n$. Using the same technique as the Waring problem, Vinogradov proved inequality (1.11) for $k^2 \lesssim n \log(n)$.

Conjecture (1.11) was proved for $k = 3$ by *Wooley* in 2014 and $k \geqslant 4$ by *Bourgain*, *Demeter*, and *Guth* in 2016. Wooley used an efficient congruence method, which helped him to apply the Vinogradov method inductively. Bourgain, Demeter and Guth used decoupling, induction on scales, and symmetries over the Fourier transform of f_k on a certain submanifold.

As an application, we view the same problem as (1.10) with a different perspective. We want to find the range of (k, n) where a nontrivial solution for (1.10) exists. Remember that a straightforward argument shows that n should be bigger than k. We call a solution for the case $k = n - 1$ a *perfect solution*. Finding a perfect solution is extremely hard, and $n = 12$ is the largest known n for a perfect solution (see [10]). We present here three categories of results. The first type of result is to find a constructive solution, that is to explicitly give a_i, b_i. To our knowledge the best possible range for k in this case is actually $O(\log n)$. In the other category, we only care about the existence of solutions, which are called non-constructive solutions. The best current range for this case is $k = O(n^{1/2})$, which is also achievable using an elementary Pigeon-hole argument. We will give this proof in chapter 3. The last category is to give a "statistical" solution, which means that $\sum a_i^r - b_i^r$ will not be zero, but very small for every $1 \leqslant r \leqslant k$.

Note that using the Vinogradov mean value theorem (1.11), the trivial upper bound becomes sharper when $n < 2k^2$ (the right hand side becomes $N^{n+\epsilon}$). So we expect to have a harder time finding a nontrivial solution. In chapter 3 we will give statistical solutions in this out-of-reach range.

1.4 Trigonometric Sums Over Primes

After studying the coefficients of *Dirichlet's series* on a certain equation based on the zeta functions we reach an inequality called *Vaughn's identity*. That is let $|f| \leqslant 1$ be an arithmetic function over integers and $UV < N$. Then

$$\sum_{n<N} f(n)\Lambda(n) \lesssim U + (\log N) \sum_{t \leqslant UV} \max_{1\leqslant w \leqslant N} \Big| \sum_{w \leqslant r \leqslant \frac{N}{t}} f(rt) \Big| + N^{\frac{1}{2}}(\log N) \max_{U\leqslant M \leqslant N/V} \max_{V \leqslant j \leqslant N/M} \left(\sum_{V<k\leqslant N/M} \left| \sum_{M<m<\min(2M, \frac{N}{j}, \frac{N}{k})} f(mj)\overline{f(mk)} \right| \right)^{\frac{1}{2}}. \tag{1.12}$$

This formula gives a trivial estimate when f is completely multiplicative.

As a result of this inequality, we can pick f to be an exponential function and arrive at the following theorem.

Theorem 1.13. *Assume that*

$$|\alpha - \frac{a}{q}| < \frac{1}{q^2}$$

Then Vinogradov proved that

$$\sum_{n<N} \Lambda(n) e\left(-n\alpha\right) \lesssim N(\log N)^3 \left(q^{-1/2} + N^{-1/5} + \frac{q}{N} \right).$$

Theorem 1.13 implies that for a large enough q, the Discrete Fourier transform of Λ is small.

1.5 Ramanujan's Sums

In this section, we mention a few properties that are related to the *Gauss sums*. We start with *Ramanujan's sum*, given by

$$\tau_q(x) = \sum_{a \in \mathbb{A}_q} e(ax/q). \tag{1.14}$$

Throughout, we denote $\mathbb{A}_q = \{a \in \mathbb{Z}/q\mathbb{Z} \ : \ (a, q) = 1\}$, so that $|\mathbb{A}_q| = \phi(q)$, the *totient function*. This lower bound on the totient function is well known: For all $0 < \epsilon < 1$, we have

$$\phi(q) \gg \frac{q}{\log\log q} \gg q^{1-\epsilon}. \tag{1.15}$$

Cancellative properties of the Ramanujan sums are very important for us, and are expressed in different ways. The first of these is

$$\tau_q(x) = \mu(q) \qquad (q, x) = 1. \tag{1.16}$$

The next cancellation property is known as *Cohen's identity*:

$$\sum_{r \in \mathbb{A}_q} \tau_q(x + r) = \mu(q)\tau_q(-x). \tag{1.17}$$

Define the Gauss sum

$$G(\chi_q, a) := \frac{1}{\phi(q)} \sum_{b \pmod q} \chi(b) e(\frac{ab}{q}).$$

We also need to use the following properties of the Gauss sums. Assume that χ is a non-principal character and $a \geqslant 1$. Also let χ^* modulo q^* be the primitive character corresponding to χ modulo q (obviously $q^*|q$). It is known that

$$G(\chi_q, a) = G(\chi, 1) \sum_{d|\gcd(a,q/q^*)} d\chi^*_{q^*}(\frac{a}{d})\mu(\frac{q}{dq^*}) \tag{1.18}$$

in particular, if $(a, q) = 1$, then $G(\chi_q, a) = \bar{\chi}_q(a)G(\chi_q, 1)$. Also we always have the inequality

$$|G(\chi_q, a)| \lesssim \begin{cases} q^{-1+\epsilon}\gcd(a,q)\sqrt{q^*} & \text{if } \chi^*_{q^*} \text{ is not principal} \\ q^{-1+\epsilon}\gcd(a,q)q^* & \text{Otherwise.} \end{cases}$$

In the case that χ_q is a real character, we have $|G(\chi_q, a)| = \sqrt{q}$ if $\gcd(a, q) = 1$. Otherwise, if $r = \gcd(q, a)$, then

$$G(\chi_q, a) = \frac{1}{\phi(q/r)}\chi^*_{q^*}(a/r)\chi^*_{q^*}(\frac{q}{rq^*})\mu(\frac{q}{rq^*})\tau(\chi^*_{q^*})\mathbf{1}_{r|q/q^*}. \tag{1.19}$$

In particular, if χ is the primitive real character, we get $G(\chi_q, a) = G(\chi_q, 1)\chi(a)1_{r=1}$.

1.6 Ergodic Theorem: Discrete Theory

After Birkhoff's theorem, Bourgain started the discrete generalized harmonic analysis field in the 1980s by proving the ergodic theorem along the square integers. Today, it is a vibrant field with several recent important results. Although there are various transference theorems to connect the discrete settings with the continuous results, note that the proofs are generally harder in the discrete cases.

Bourgain generalized the Birkhoff Ergodic Theorem for the square integers. For $f \in L^2(X)$,

$$K_N f(x) := \frac{1}{\sqrt{N}} \sum_{k<\sqrt{N}} f(T^{k^2}x) \text{ converges } \mu - \text{almost everywhere.}$$

One of his results was the fact that $\sup_N K_N$ is an ℓ^p-bounded operator, for $1 < p \leqslant \infty$. In order to prove this theorem, he created the *multifrequency* argument.

Theorem 1.20. *Let $\lambda_1, \cdots, \lambda_L \in \mathbb{R}$ be distinct points with $|\lambda_i - \lambda_s| \geqslant 2^{-j_0}$ for $i \neq s$. Then*

$$\left\| \sup_{j>j_0} \left| \sum_{\ell=1}^{L} e(\lambda_\ell x) \left(\phi_j * e(-\lambda_\ell \cdot) f\right)(x) \right| \right\|_2 \lesssim (\log L)^3 \parallel f \parallel_2$$

where ϕ_j can be certain smooth Schwartz functions.

Inequalities like in theorem 1.20 are called a *quantitative* result, while the Birkhoff theorem was a *qualitative* statement.

1.6.1 Ergodic Theorem Along the Prime Numbers

We also may study the average over sets other than polynomials. These kinds of problems are considered to be ergodic theorems with *arithmetic weights*. We can consider the set of primes, related to the Vinogradov theorem, to study

$$T_N f(x) := \frac{1}{N} \sum_{n<N} \Lambda(n) f(x-n)$$

The set of primes is "*full dimensional*", so one can see that in an appropriate sense, an ℓ^1 function is improved to an ℓ^∞ function. Our result in chapter 4 gives such an improving upper bound. We can also check the cases where $p \neq 1$, which are not the endpoints. There is the following ℓ^p-improving bound for $1 < p < 2$:

$$\parallel T_N f \parallel_{\ell^{p'}} \lesssim N^{\frac{1}{p'} - \frac{1}{p}} \parallel f \parallel_{\ell^p} \text{ where } \frac{1}{p} + \frac{1}{p'} = 1$$

We can say that these inequalities are *improving*, since we improve from ℓ^p to $\ell^{p'}$.

1.7 Schanuel's Conjecture

Let $\bar{\mathbb{Q}}$ be the set of algebraic numbers, and

$$\mathbb{L} := \{x \text{ such that } x = \log y \text{ for some } y \in \bar{\mathbb{Q}}\}.$$

There are many problems involving the set $\mathbb{L}$ in *transcendental number theory*. For example, the *Hermite–Lindemann theorem* states that any nonzero element of $\mathbb{L}$ is transcendental. Also any pair of elements $\lambda_1, \lambda_2 \in \mathbb{L}\backslash\{0\}$ which are independent over $\mathbb{Q}$ should be linearly independent over $\bar{\mathbb{Q}}$. A generalization of this result is known as the *Baker's theorem*, which states that:

Theorem 1.21. *Assume that $\lambda_1, \cdots, \lambda_r \in \mathbb{L}$ are independent over $\mathbb{Q}$, and $\beta_1, \cdots, \beta_r \in \bar{\mathbb{Q}}$. Also assume that H is the maximum of heights of β_i (For the definition of height, please see [11, 12, 13]). Then*

$$|\lambda_1\beta_1 + \cdots + \lambda_r\beta_r| > H^{-C}$$

where C is an effectively computable constant with respect to r, λ_i and maximum degree of β_i.

It immediately, gives the following result: For algebraic numbers $\alpha_1, \cdots, \alpha_r \neq 0, 1$, and for rationally independent algebraic irrational numbers $\beta_1, \cdots, \beta_r$ the number $\alpha_1^{\beta_1} \cdots \alpha_r^{\beta_r}$ is a transcendental number.

Baker's theorem concludes the question regarding the transcendence of the value at algebraic numbers of the polynomials with algebraic coefficients. A more general question, which is still a conjecture asks for transcendence of rational functions with coefficients that are algebraic numbers. In other world, we are interested in the following conjecture:

Conjecture 1 (Schanuel's Conjecture)**.** *Assume that $z_1, \cdots, z_r \in \mathbb{C}$ are linearly independent over the rational numbers. Then the field extension $F = \mathbb{Q}(z_1, \cdots, z_r, e^{z_1}, \cdots, e^{z_r})$*

has transcendence degree at least r over $\mathbb{Q}$.

As a particular case, Schanuel's conjecture states that for distinct prime numbers $q_1, \cdots, q_r$, the set

$$\{\frac{1}{\log q_1}, \cdots, \frac{1}{\log q_r}\}$$

is independent over $\mathbb{Q}$. Although this conjecture does not seem achievable at the moment, there is partial progress like in [14] and [15].

1.8 Thesis Organization

This thesis has four chapters. The first chapter is the introduction, which consists of the necessary preliminaries. In the second chapter we study the Graham problem about the $p-$divisibility of the central binomial coefficients. We show that for every $r \geqslant 1$, and all r distinct (sufficiently large) primes $p_1, ..., p_r > p_0(r, \varepsilon)$, there exist infinitely many integers n such that $\binom{2n}{n}$ is divisible by these primes to only low multiplicity. From a theorem of Kummer, an upper bound for the number of times that a prime p_j can divide $\binom{2n}{n}$ is $1 + \log n / \log p_j$; and our theorem shows that we can find integers n where for $j = 1, ..., r$, p_j divides $\binom{2n}{n}$ with multiplicity at most ε times this amount. This work is under review for publication (see [16]).

In the third chapter, we study an unexpected cancellation involving exponential sums. Following attempts at an analytic proof of the Pentagonal Number Theorem, we report on the discovery of a general principle leading to an unexpected cancellation of oscillating sums. It turns out that sums in the class we consider are much smaller than would be predicted by certain probabilistic heuristics. After stating the motivation, and our theorem, we apply it to prove several results on the Prouhet-Tarry-Escott Problem, integer partitions, and the distribution of prime numbers. We solve an approximate version of the Prouhet-Tarry-Escott Problem, and in doing so we give some evidence that a certain pigeonhole argument for solving the exact version of the Problem can be improved. In fact, our work in the approximate case exceeds the bounds one can prove using a pigeonhole argument, which seems remarkable. Also, we prove that

$$\sum_{\ell^2<n} (-1)^{\ell} p(n-\ell^2) \ \sim \ (-1)^n 2^{-3/4} n^{-1/4} \sqrt{p(n)},$$

where $p(n)$ is the usual partition function. We get a "Weak pentagonal number theorem", in which we can replace the partition function $p(n)$ with Chebyshev Ψ function. Our result is stronger than one would get using a strong form of the Prime Number Theorem and also a naive use of the Riemann Hypothesis in each interval, since the widths of the intervals are smaller than $e^{\frac{1}{2}\sqrt{x}}$, making the Riemann Hypothesis estimate "trivial". This project is also under review for publication (see [17]).

In the last chapter, we study an ergodic average along the primes. We prove sharp ℓ^p-improving for these averages, and sparse bounds for the maximal function. The inequality assuming the GRH is sharp. The proof depends upon the Circle Method, and an interpolation argument of Bourgain. This work has been published in Math Research Letter journal (see [18]). Related to this topic, we also published a similar result in [19].

CHAPTER 2

ON A CONJECTURE OF GRAHAM ON THE P-DIVISIBILITY OF CENTRAL BINOMIAL COEFFICIENTS

2.1 Introduction

In [20] and [21] it is mentioned that R. L. Graham had offered $1,000 to settle the problem of whether or not there are infinitely many integers n such that $\binom{2n}{n}$ is relatively prime to $105 = 3 \cdot 5 \cdot 7$. From the following theorem of Kummer [22] we immediately see that Graham's problem is equivalent to asking whether there are infinitely many integers $n \geqslant 1$ with the property that when we add n to itself in bases 3, 5, and 7, there are no carries.

Kummer's Theorem: For a prime p we have that the number of times that p divides $\binom{n}{m}$ equals the number of carries when adding the numbers m and $n - m$ in base-p.

In other words, are there infinitely many integers $n \geqslant 1$ such that all the base-3 digits are in $\{0, 1\}$, all the base-5 digits are in $\{0, 1, 2\}$, and all the base-7 digits are in $\{0, 1, 2, 3\}$? If so, then there are infinitely many integers n such that $\gcd\left(\binom{2n}{n}, 105\right) = 1$; and if not, then there are at most finitely many integers $n \geqslant 1$ with $\gcd\left(\binom{2n}{n}, 105\right) = 1$.

In [23], Erdős, Graham, Ruzsa, and Straus proved that for every *pair* of primes p, q, there are infinitely many integers $n \geqslant 1$ with $\gcd\left(\binom{2n}{n}, pq\right) = 1$; however, there are no such results in the literature for 3 or more primes (though, for example, there are results [24, 25] on when $\binom{2n}{n}$ is coprime to n and [26] when $\binom{2n}{n}$ is squarefree). Apart from whether one can give a proof of whether there are or aren't infinitely many n with $\gcd\left(\binom{2n}{n}, 105\right) = 1$, one can at least ask whether it's *plausible* or not that such integers $n \geqslant 1$ exist. Pomerance gave a simple heuristic for why there should exist infinitely many $n \geqslant 1$ with this property (see, for example, [27]): if we choose a random $n \in [1, x]$, the probability that all its base-3 digits are in $\{0, 1\}$ should be about $(2/3)^{\log(x)/\log 3} \approx x^{-0.37}$; the probability that all its base-5 digits are $\{0,1,2\}$ should be about $(3/5)^{\log(x)/\log 5} \approx x^{-0.32}$; and the probability that all its base-7 digits are $\{0,1,2,3\}$ should be about $(4/7)^{\log(x)/\log 7} \approx x^{-0.29}$. Assuming independence, the probability that a random $n \in [1,x]$ satisfies all three conditions is about $x^{-0.37}x^{-0.32}x^{-0.29} = x^{-0.98}$. So, we would expect there to be about $x^{0.02}$ numbers $n \in [1,x]$ with the property that $\gcd\left(\binom{2n}{n}, 105\right) = 1$, which clearly tends to infinity the larger we take x to be.

One can extend Pomerance's heuristic to any number of odd primes, making the same independence assumptions (that the events $E_1, ..., E_r$ are mutually independent, where for a randomly chosen integer $n \in [1,x]$, E_j is the event that the base-p_j digits of n are in $\{0,1,...,(p_j-1)/2\}$). When one does this, one would expect there to exist infinitely many integers $n \geqslant 1$ such that $\gcd\left(\binom{2n}{n}, p_1 \cdots p_r\right) = 1$, for distinct odd primes $p_1, ..., p_r$, provided that

$$-\sum_{j=1}^{r} \frac{\log\left(\frac{1}{2} + \frac{1}{2p_j}\right)}{\log(p_j)} \ < \ 1; \tag{2.1}$$

and that (using the Borel-Cantelli Lemma) there should be only *finitely* many such n if the $>$ is replaced with a $<$. We make no guesses about the possible case when the left-hand-side equals 1, exactly – if it is even possible.

What is interesting here is that even if we consider a slight weakening of the problem where we allow $\binom{2n}{n}$ to be divisible by the primes $p_1, ..., p_r$ to *low multiplicity*, we get the same condition (2.1) guaranteeing the existence of infinitely many such $n \geqslant 1$: in light of Kummer's theorem, the number of times that a prime p_j can divide a number n is at most about $1 + \log(n)/\log p_j$, since this is an upper bound on the number of base-p_j digits of n. If we select a random $n \in [1,x]$, the probability that all but at most k of the base-p_j digits are in $\{0,1,2,...,(p_j-1)/2\}$ is

$$\asymp \binom{[\log x/\log p_j]}{k} \left(\frac{1}{2} + \frac{1}{2p_j}\right)^{\log(x)/\log p_j - k} \left(\frac{1}{2} - \frac{1}{2p_j}\right)^{k},$$

for $k = o(\log x)$. This has size (assuming $k = o(\log x)$)

$$\left(\frac{1}{2} + \frac{1}{2p_j}\right)^{(1-o(1))\log(x)/\log p_j},$$

which, apart from the factor $1 - o(1)$ in the exponent, has the same form as the probability for the case where *every* base-p_j digit of n is in $\{0, 1, ..., (p_j - 1)/2\}$. Making the same independence assumptions as before, we thus would expect that if (2.1) holds, then there should exist infinitely many integers $n \geqslant 1$ where for $j = 1, ..., r$, p_j divides $\binom{2n}{n}$ to multiplicity at most $o(\log n)$; and, if, instead, the left-hand-side of (2.1) is > 1, we would expect there to be only *finitely* many such $n \geqslant 1$.

In this paper, we don't quite prove that (2.1) implies there are infinitely many such $n \geqslant 1$, but we do prove something in this direction:

Theorem 2.2. *Suppose $r \geqslant 1$, $\varepsilon > 0$, and let $p_1, ..., p_r \geqslant c_0(r, \varepsilon)$ be distinct primes, where $c_0(r, \varepsilon)$ is some function of r and ε (can be deduced from the proof). Then, there is a sequence $n_1, n_2, ...$ of integers n such that for all $i = 1, ..., r$,*

$$\nu_{p_i}\left(\binom{2n}{n}\right) \leqslant \frac{\varepsilon \log n}{\log p_i},$$

where $\nu_p(x)$ denotes the number of times the prime p divides x.

As we said, a trivial upper bound for $\nu_{p_i}(\binom{2n}{n})$ is $1 + (\log n)/\log p_i$, since n has at most this many base-p_i digits; so the theorem is saying that we can find infinitely many n where we are smaller than this amount by a factor ε, for all the primes $p_1, ..., p_r$, simultaneously.

As one will see, the proof is fairly technical. What would greatly simplify it is if one had that the numbers $1/\log 2, 1/\log p_1, ..., 1/\log p_r$ were linearly independent over the rationals. This is not known to be true for arbitrary sets of primes, but it would follow from the following conjecture:

Schanuel's Conjecture [28] Given any n complex numbers $z_1, ..., z_n$ that are linearly independent over the rationals, the field extension $\mathbb{Q}(z_1, ..., z_n, e^{z_1}, ..., e^{z_n})$ has transcendence degree at least n over $\mathbb{Q}$.

If Schanuel's Conjecture holds, then taking $n = r + 1$, and taking $z_1 = \log 2$, $z_2 = \log p_1$, ..., $z_{r+1} = \log p_r$, we see that $\mathbb{Q}(\log 2, \log p_1, ..., \log p_r)$ has transcendence degree $r + 1$. Now suppose we had a linear combination

$$\frac{\lambda_1}{\log 2} + \frac{\lambda_2}{\log p_1} + \cdots + \frac{\lambda_{r+1}}{\log p_r} = 0,$$

where $\lambda_1, ..., \lambda_{r+1} \in \mathbb{Q}$ and not all 0. Without loss, assume that $\lambda_1 \neq 0$. Then, the linear relation would imply that

$$\mathbb{Q}(\log 2, \log p_1, ..., \log p_r) = \mathbb{Q}(\log p_1, \log p_2, ..., \log p_r),$$

which can have transcendence degree at most r, which would be a contradiction. Thus, no such linear relations can hold.

2.2 Proof of the Main Theorem

As we said in the introduction, the central binomial coefficients in the statement of the theorem are somewhat of a distraction, in light of Kummer's Theorem. This theorem implies that if all but at most $\varepsilon(\log n)/\log p_j$ of the base-p_j digits of n are $\leqslant p_j/2 - 1$, then $v_{p_j}\left(\binom{2n}{n}\right) \leqslant \varepsilon(\log n)/\log p_j$; and establishing that there are infinitely many integers n with this property (few base-p_j digits that are $> p_j/2 - 1$) is the path we will take to prove Theorem 2.2.

In carrying out this verification, we will make use of the following theorem:

Theorem 2.3. *Suppose that $p_1, ..., p_r$ are distinct primes. For $i = 1, ..., r$, and $n \geqslant 1$, define*

$$\alpha_i(n) := p_i^{\{n(\log 2)/\log p_i\}-1} = p_i^{n(\log 2)/\log p_i - [n(\log 2)/\log p_i] - 1},$$

Define, for $H \geqslant 1$ and $i = 1, 2, ..., r$,

$$U_i(H) \ := \ \left\{ \frac{d_1}{p_i} + \frac{d_2}{p_i^2} + \cdots + \frac{d_H}{p_i^H} \ : \ 0 \ \leqslant \ d_1, ..., d_H \ \leqslant \ \frac{p_i}{3} \right\} \ + \ \left[0, \ \frac{1}{p_i^H} \right). \qquad (2.4)$$

Then, for every $\varepsilon > 0$ and some $H = H(N)$ tending to infinity slowly with N (in a sense that can be made precise by following the proof), we have that for $N \geqslant 1$ and for arbitrary sequences of real numbers $\{\beta_i(n)\}_{n=1}^{\infty}$, $i = 1, ..., r$,

$$\frac{\#\{n \leqslant N \ : \ \exists \ s \leqslant 2^{\varepsilon H} \ \forall \ j = 1, ..., r, \ \{s\alpha_j(n) + \beta_j(n)\} \in U_j(H)\}}{N} \ \geqslant \ 1 - o(1). \quad (2.5)$$

Now let us see that this theorem implies Theorem 2.2: it clearly suffices to prove that for each integer N sufficiently large, we can find an integer n satisfying

$$2^{N/2} \ < \ n \ \leqslant \ 2^N,$$

so that for all $i = 1, 2, ..., r$, all but at most $\varepsilon(\log n)/\log p_j$ of the base-p_j digits of n are $\leqslant p_j/3$.

So, let us suppose N is given. Let $f(N)$ denote the minimum possible value of the ratio on the left-hand-side of (2.5), for some choice of $H = H(N)$ tending to infinity with N, over all choices of $\{\beta_i(n)\}_{n=1}^{\infty}$, $i = 1, ..., r$. Note that

$$f(N) \ \geqslant \ 1 - o(1).$$

Next, let

$$\ell = \ell(N) \ := \ \lfloor \min((1 - f(N))^{-1/2}, H(N))^{1/2} \rfloor.$$

(It's also worth mentioning that the exponent $1/2$ here is a little arbitrary, and can be replaced with any exponent in $(0, 1)$, as far as our proof below.)

Let

$$N' := \lfloor N/\ell \rfloor.$$

We now construct the number

$$n := n_0 2^{\ell N'} + n_1 2^{\ell(N'-1)} + \cdots + n_{N'}$$

as follows: we start by letting $n_0 = 1$. Assume we have constructed $n_0, ..., n_{d-1}$. Now we show how to construct n_d: for $j = 1, ..., r$, we let

$$\beta_j(\ell(N'-d)) := \frac{n_0 2^{\ell N'} + n_1 2^{\ell(N'-1)} + \cdots + n_{d-1} 2^{\ell(N'-d+1)}}{p_j^{m_{j,d}}},$$

where for an integer h we define

$$m_{j,h} = \left[\frac{\ell(N'-h)\log 2}{\log p_j} \right] + 1.$$

(Alternatively: $m_{j,h}$ is the unique integer so that $2^{\ell(N'-h)}/p_j^{m_{j,h}}$ lies in $[1/p_j, 1)$.)

If it exists, we let $1 \leqslant n_d \leqslant 2^{\varepsilon\ell}$ be any integer where

$$\{n_d \alpha_j(\ell(N'-d)) + \beta_j(\ell(N'-d))\} \in U_j(H). \tag{2.6}$$

If no such n_d exists, just let $n_d = 0$.

In order to see that this construction works, we begin by noting that for any integer h,

$$\alpha_j(h) = p_j^{\{h(\log 2)/\log p_j\}-1} = \frac{p_j^{h(\log 2)/\log p_j}}{p_j^{[h(\log 2)/\log p_j]+1}} = \frac{2^h}{p_j^{h'}},$$

where h' is the unique integer such that this belongs to the interval $[1/p_j, 1)$.

Thus, when we go to construct n_d, we will have

$$n_d\alpha_j(\ell(N'-d)) + \beta_j(\ell(N'-d)) \;=\; \frac{n_0 2^{\ell N'} + n_1 2^{\ell(N'-1)} + \cdots + n_d 2^{\ell(N'-d)}}{p_j^{m_{j,d}}}.$$

It follows that if we write

$$n_0 2^{\ell N'} + n_1 2^{\ell(N'-1)} + \cdots + n_d 2^{\ell(N'-d)} \;=\; c_0 + c_1 p_j + c_2 p_j^2 + \cdots\,, \tag{2.7}$$

where $0 \leqslant c_i \leqslant p_j - 1$, then from (2.6) we deduce that if $n_d \neq 0$ then

$$0 \;\leqslant\; c_{m_{j,d}-1},\; c_{m_{j,d}-2},\; \ldots,\; c_{m_{j,d}-H} \;\leqslant\; \frac{p_j}{3},$$

and so in particular, since $m_{j,d} - m_{j,d+1} < \ell + 1 = o(H(N))$, we have that $0 \leqslant c_u \leqslant p_j/3$ for

$$m_{j,d+1} \;\leqslant\; u \;\leqslant\; m_{j,d}.$$

Now, if we continue adding on additional terms to (2.7),

$$n_{d+1}2^{\ell(N'-d-1)},\; n_{d+2}2^{\ell(N'-d-2)},\; \ldots \tag{2.8}$$

these will only have an effect on the terms $c_t p_j^t$ where

$$t \;\leqslant\; m_{j,d+1} + [(\log n_{d+1})/\log p_j] + 1 \;\leqslant\; m_{j,d+1} + \varepsilon\ell(\log 2)/\log p_j + 1.$$

Thus, the terms $c_u p_j^u$ where

$$m_{j,d+1} + \varepsilon\ell(\log 2)/\log p_j + 1 \;<\; u \;\leqslant\; m_{j,d}$$

in (2.7) will be unchanged, as will all the other higher-order terms with $u > m_{j,d}$.

Now we distinguish two possbilities for each $d \leqslant N'$: we let $D^\sharp$ be those d such that there *does* exist an $n_d \leqslant 2^{\varepsilon\ell}$ where (2.6) holds, and we let $D^\flat$ be those d for which it doesn't.

For each $d \in D^\sharp$ we have that for each $j = 1, ..., r$, at most $\varepsilon\ell(\log 2)/\log p_j + 1$ base-p_j digits c_u with $m_{j,d+1} \leqslant u \leqslant m_{j,d}$ are $> p_j/3$; and for each $d \in D^\flat$, in the worst cast for every $j = 1, ..., r$, all of the c_u with $m_{j,d+1} \leqslant u \leqslant m_{j,d}$ could be $> p_j/3$. Note that in this case (the case $d \in D^\flat$) there are at most $\ell + 1$ bad digits c_u with $m_{j,d+1} \leqslant u \leqslant m_{j,d}$.

Now, Theorem 2.3 implies that the number of $d \leqslant N'$ for which (2.6) doesn't hold is at most

$$N(1 - f(N)) \leqslant \frac{N}{\ell^2} = o(N').$$

All told, the total number of bad base-p_j digits that are $> p_j/3$ in this case, over all $d \in D^\sharp$, is at most

$$(\ell + 1)o(N') = o(N),$$

for every $j = 1, ..., r$. And the total number of bad base-p_j digits arising for the $d \in D^\flat$ is at most

$$N'(\varepsilon\ell(\log 2)/\log p_j + 1) < \varepsilon N,$$

for each $j = 1, ..., r$.

In total, then, for every $j = 1, ..., r$, the number of bad base-p_j digits (that are $> p_j/2$) is at most

$$\varepsilon N + o(N).$$

But since $\varepsilon > 0$ was arbitrary, it's obvious that for every $j = 1, ..., r$, the number of bad base-p_j digits is εN. This is just what we need to show in order to prove Theorem 2.2.

2.3 Proof of Theorem 2.3

In proving this theorem we will need to understand how the vectors

$$\begin{aligned}
&(\alpha_1(n),\ \alpha_2(n),\ ...,\ \alpha_r(n)) \\
&\quad = \ (p_1^{\{n(\log 2)/\log p_1\}-1},\ p_2^{\{n(\log 2)/\log p_2\}-1},\ ...,p_r^{\{n(\log 2)/\log p_r\}-1}).
\end{aligned}$$

are distributed, as we vary over $n \leqslant N$.

2.3.1 The 2-dimensional case

To better understand what is going on, we first consider the case where $r = 2$. There are two possibilities: the first possibility is that there do not exist integers n_1, n_2, n_3, with $n_1, n_2, n_3 \neq 0$, such that

$$n_1 \frac{\log 2}{\log p_1} + n_2 \frac{\log 2}{\log p_2} \ = \ n_3. \tag{2.9}$$

If this occurs, then as a consequence of

Theorem 2.10 (Multidimensional Weyl's Theorem)**.** *Suppose* $1, \vartheta_1, ..., \vartheta_r$ *are real numbers that are linearly independent over the rationals. Then, for* $\vec{\vartheta} = (\vartheta_1, ..., \vartheta_r) \in \mathbb{R}^r$*, the sequence* $\{k\vec{\vartheta}\}_{k=1}^{\infty}$ *is uniformly distributed in* $\mathbb{R}^r/\mathbb{Z}^r$*.*

(see [29, example 6.1]) we have that the vector $(n(\log 2)/\log p_1, n(\log 2)/\log p_2)$ is uniformly distributed mod 1 as we vary over $n = 1, 2, 3, ...$; and, therefore, the set $(\alpha_1(n), \alpha_2(n))$ is dense in the box $[1/p_1, 1] \times [1/p_2, 1]$.

The second possibility is that there *do* exist integers $n_1, n_2, n_3 \neq 0$ such that (2.9) holds (If n_1 were allowed to be 0, then we would have that (2.9) implies $n_2 \log 2 = n_3 \log p_2$, which can only hold if $n_2 = n_3 = 0$; and a similar thing occurs for when $n_2 = 0$ or when $n_3 = 0$; so, if one of these n_i were 0, the others would have to be as well.)

By multiplying through by -1 as needed, we can assume $n_2 > 0$; and we will assume that the p_1 and p_2 are arranged so that

$$|n_1/n_2| \leqslant 1.$$

We will show that the set

$$(\alpha_1(n),\ \alpha_2(n)) \ = \ (p_1^{\{n(\log 2)/\log p_1\}-1},\ p_2^{\{n(\log 2)/\log p_2\}-1}),\ n = 1, 2, 3, ... \tag{2.11}$$

is contained in a union of a finite set of non-linear curves. It turns out that, moreover, the set is equidistributed on these curves (when we restrict to $[1/p_1, 1] \times [1/p_2, 1]$) with respect to the right measure; though, we don't actually need the full strength of such a statement, so don't bother to prove it.

We claim that for each integer $n \geqslant 1$,

$$\left\{\frac{n \log 2}{\log p_2}\right\} \ = \ f(n) - \frac{n_1}{n_2}\left\{\frac{n \log 2}{\log p_1}\right\},$$

where $f(n) \in S$, a finite set of possibilities. To see this, we begin by rewriting (2.9) as

$$\frac{n \log 2}{\log p_2} \ = \ \frac{nn_3}{n_2} - \frac{n_1}{n_2}\frac{n \log 2}{\log p_1}. \tag{2.12}$$

Now we write

$$\frac{n \log 2}{\log p_1} \ = \ \left[\frac{n \log 2}{\log p_1}\right] + \left\{\frac{n \log 2}{\log p_1}\right\} \ = \ k(n)n_2 + a(n) + \left\{\frac{n \log 2}{\log p_1}\right\}, \tag{2.13}$$

where $k(n)$ is an integer, and $0 \leqslant a(n) \leqslant n_2 - 1$. We also write

$$n \ = \ \ell(n)n_2 + b(n),\ \text{where } 0 \leqslant b(n) \leqslant n_2 - 1,\ \text{and } \ell(n) \in \mathbb{Z}.$$

It follows, then, upon plugging this and (2.13) into (2.12), that

$$\frac{n\log 2}{\log p_2} \;=\; \ell(n)n_3 + \frac{b(n)n_3}{n_2} - k(n)n_1 - \frac{a(n)n_1}{n_2} - \frac{n_1}{n_2}\left\{\frac{n\log 2}{\log p_1}\right\}.$$

Thus, since $|n_1/n_2| \leqslant 1$,

$$\left\{\frac{n\log 2}{\log p_2}\right\} \;=\; \left\{\frac{b(n)n_3}{n_2} - \frac{a(n)n_1}{n_2}\right\} - \frac{n_1}{n_2}\left\{\frac{n\log 2}{\log p_1}\right\} + \delta,\ \text{where } \delta \in \{0,1,-1\}.$$

We would take $\delta = 0$ if the preceding terms add to a number in $[0,1)$; take $\delta = 1$ if they produce a number in $[-1,0)$; and take $\delta = -1$ if they produce a number in $[1,2]$.

It follows that we may take S to be

$$S \;=\; \left\{\left\{\frac{bn_3}{n_2} - \frac{an_1}{n_2}\right\} \;:\; a,b = 0,1,...,n_2-1\right\} \;+\; \{0,1,-1\}.$$

Thus,

$$|S| \;\leqslant\; 3n_2^2.$$

We conclude that, for $n \geqslant 1$,

$$\begin{aligned}&(p_1^{\{n\log 2/\log p_1\}-1},\ p_2^{\{n\log 2/\log p_2\}-1})\\ &\qquad = \; (p_1^{\{n\log p/\log p_1\}-1},\ c(n)p_2^{-(n_1/n_2)\{n\log p/\log p_1\}-1}),\end{aligned} \tag{2.14}$$

where $c(n) = p_2^{f(n)}$, where, recall, $f(n) \in S$.

As we vary over $n \leqslant N$, and let $N \to \infty$, all the points (2.14) lie on set of at most $|S| \leqslant 3n_2^2$ curves of the form

$$C_s \;:=\; \{(p_1^{t-1},\ c_s p_2^{-(n_1/n_2)t-1} \;:\; 0 \leqslant t < 1\},\ \text{where } c_s = p_2^s,\ \text{where } s \in S.$$

2.3.2 None of these curves are lines

Each of these curves are just dilates of one another in the second coordinate. So, to show that none are lines, it suffices to show that the curve with points

$$z(t) \ := \ (p_1^t,\ p_2^{-(n_1/n_2)t}),$$

is not a line.

To see this it suffices to prove that

$$p_1 \ \neq \ p_2^{-(n_1/n_2)},$$

which is clearly the case, since upon raising both sides to the n_2 power, if they were equal we would have

$$p_1^{n_2} \ = \ p_2^{-n_1},$$

which can't hold if p_1 and p_2 are distinct primes.

2.4 Generalizing to higher dimensions

Now suppose we have r primes $p_1, ..., p_r$, and we wish to understand the possible vectors

$$(p_1^{\{n \log 2/\log p_1\}-1},\ p_2^{\{n \log 2/\log p_2\}-1},\ ...,\ p_r^{\{n \log 2/\log p_r\}-1}), \tag{2.15}$$

given that we have relations similar to (2.9). In this case, there can be more than one such relation. We can express this set of relations as

$$\begin{aligned} a_{1,1}\frac{\log 2}{\log p_1} + a_{1,2}\frac{\log 2}{\log p_2} + \cdots + a_{1,r}\frac{\log 2}{\log p_r} \quad &= \quad a_{1,r+1} \\ &\vdots \\ a_{k,1}\frac{\log 2}{\log p_1} + a_{k,2}\frac{\log 2}{\log p_2} + \cdots + a_{k,r}\frac{\log 2}{\log p_r} \quad &= \quad a_{k,r+1}, \end{aligned}$$

where all the $a_{i,j} \in \mathbb{Q}$, where $k \leqslant r-1$, and where all these relations are linearly independent. Note that if there were $k = r$ linearly independent relations, then this would imply that all the $\log 2/\log p_i$ are rational numbers, which would imply that for each $i = 1, ..., r$, $\log 2$ and $\log p_i$ are linearly dependent over the rationals, which we know is false, as it would imply that there is an integer power of 2 that equals an integer power of p_i.

Upon applying row-reduction to these equations, and permuting the p_j's as needed, we can reduce the above system to the following one: for $j = 1, ..., k$, we have

$$\frac{\log 2}{\log p_{r-j+1}} \ = \ b_{j,1}\frac{\log 2}{\log p_1} + \cdots + b_{j,r-k}\frac{\log 2}{\log p_{r-k}} + b_{j,r+1},$$

where the $b_{j,h} \in \mathbb{Q}$. We have, also, that (recalling that the p_j's have been permuted from their original ordering)

$$1,\ \frac{\log 2}{\log p_1},\ ...,\ \frac{\log 2}{\log p_{r-k}} \text{ are independent over } \mathbb{Q}. \tag{2.16}$$

We note that this holds also in the case $k = 0$, where there are *no* linear relations as above.

Getting a common denominator, we can rewrite the above as: for $j = 1, ..., k$, we have

$$\frac{\log 2}{\log p_{r-j+1}} \ = \ \frac{m_{j,1}}{n_j}\frac{\log 2}{\log p_1} + \cdots + \frac{m_{j,r-k}}{n_j}\frac{\log 2}{\log p_{r-k}} + \frac{m_{j,r+1}}{n_j}, \tag{2.17}$$

where, for all $j = 1, ..., k$ and $h = 1, ..., r-k, r+1$, the $n_j \geqslant 1$ and the $m_{j,h}$ are integers.

Now, we claim that for $n \geqslant 1$, and $j = 1, ..., k$,

$$\left\{\frac{n\log 2}{\log p_{r-j+1}}\right\} \ = \ g_j(n) \ + \ \frac{m_{j,1}}{n_j}\left\{\frac{n\log 2}{\log p_1}\right\} + \cdots + \frac{m_{j,r-k}}{n_j}\left\{\frac{n\log 2}{\log p_{r-k}}\right\}, \tag{2.18}$$

where $g_j(n)$ takes on values in a finite set S of possibilities.

To see this, we proceed as with the 2-dimensional case: for $j = 1, ..., k$ and $h =$

$1, ..., r-k$, we define the numbers $\ell_{j,h}(n) \in \mathbb{Z}$ and $0 \leqslant a_{j,h}(n) \leqslant n_j - 1$ as follows

$$\frac{n \log 2}{\log p_h} \;=\; \ell_{j,h}(n) \cdot n_j + a_{j,h}(n) \;+\; \left\{ \frac{n \log 2}{\log p_h} \right\}.$$

Thus, from (2.17) we have that

$$\frac{n \log 2}{\log p_{r-j+1}} \;=\; \sum_{h=1}^{r-k} \ell_{j,h}(n) m_{j,h} + \frac{a_{j,h}(n) m_{j,h}}{n_j} + \frac{m_{j,h}}{n_j} \left\{ \frac{n \log 2}{\log p_h} \right\} + n b_{j,r+1}.$$

Taking the fractional part of both sides, we find that

$$\left\{ \frac{n \log 2}{\log p_{r-j+1}} \right\} \;=\; \left\{ \sum_{h=1}^{r-k} \frac{a_{j,h}(n) m_{j,h}}{n_j} \right\} + \left(\sum_{h=1}^{r-k} \frac{m_{j,h}}{n_j} \left\{ \frac{n \log 2}{\log p_h} \right\} \right) + \delta_j, \tag{2.19}$$

where δ_j is an integer chosen so as to make the right-hand-side of this equation be a real number in $[0, 1)$. Clearly, $\delta_j \in \{-\Delta, -\Delta + 1, ..., 0, ..., \Delta\}$, where

$$\Delta \;=\; 1 + r \cdot \max_{j,h} \left\lfloor \frac{|m_{j,h}|}{|n_j|} \right\rfloor.$$

Thus, if we let

$$S \;:=\; \{c_j / n_j \;:\; j = 1, ..., k, \text{ and } 0 \leqslant c_j \leqslant n_j - 1\} \;+\; \{-\Delta, -\Delta + 1, ..., 0,\ 1, ...,\ \Delta\},$$

then from (2.19) we see that

$$\left\{ \frac{n \log 2}{\log p_{r-j+1}} \right\} \;=\; g_j(n) \;+ \sum_{h=1}^{r-k} \frac{m_{j,h}}{n_j} \left\{ \frac{n \log 2}{\log p_h} \right\},$$

where $g_j(n) \in S$.

Thus, proceeding as in the 2-dimensional case, we see that the set of points (2.15) all lie on one of the following finite set of surfaces given as follows:

$$(p_1^{t_1-1},\ p_2^{t_2-1},\ ...,\ p_{r-k}^{t_{r-k}-1},\ c_1 p_{r-k+1}^{\theta_1(t_1,...,t_{r-k})-1},\ ...,\ c_k p_r^{\theta_k(t_1,...,t_{r-k})-1}), \tag{2.20}$$

where

$$\text{for } i = 1, ..., k,\ c_i \ = \ p_{r-k+i}^{s_i},\ \text{for some } s_i \in S,$$

and where

$$\theta_i(t_1, ..., t_{r-k}) \ = \ \sum_{h=1}^{r-k} \frac{m_{k-i+1,h}}{n_{k-i+1}} t_h.$$

We note that for $k = 0$ (no linear relations) the surface (2.20) just becomes

$$(p_1^{t_1-1},\ p_2^{t_2-1},\ ...,\ p_r^{t_r-1}).$$

2.5 Passing to parameterized curves

2.5.1 An illustrative example

We would like to break these surfaces up into a union of parameterized curves of the form

$$(c_1 \cdot \alpha_1^t, c_2 \cdot \alpha_2^t, ..., c_r \cdot \alpha_r^t),$$

where the α_j's are all distinct, and none of the c_j's are 0. One attempt at doing this would be to take a surface of the form (2.20), and set all but one of the t_j's to fixed values. For example, if our surface were of the form

$$(p_1^{t_1},\ p_2^{t_2},\ p_3^{t_1+t_2}),\ t_1, t_2 \in [0, 1), \tag{2.21}$$

then if we were to freeze t_1 and vary t_2, we would get a curve of the form

$$(c,\ p_2^t,\ d \cdot p_3^t).$$

Unfortunately, only two coordinates vary, not all three. However, if we parameterize differently, then we can get a full, 3-dimensional curve: let $t_1 = t$, $t_2 = t + \delta \bmod 1$, where δ is fixed and t varies in $[0, 1)$. Then, we get the parameterized curve

$$(p_1^t,\ c \cdot p_2^t,\ d \cdot p_3^{2t}), \tag{2.22}$$

where $c = p_2^{\delta}$, $d = p_3^{\delta}$. Actually, this isn't quite right, since, for example $2t > 1$ for $t > 1/2$; we need to introduce another curve to account for these possibilities. Basically, we consider all curves of (2.22) where $c \in \{p_2^{\delta}, p_2^{\delta-1}\}$ and $d \in \{p_3^{\delta}, p_3^{\delta-1}, p_3^{\delta-2}\}$. This covers all the cases; and, as we vary over all $\delta \in [0, 1)$, we get a union of curves, where this union is exactly the surface (2.21) when restricted to $(1/p_1, 1] \times (1/p_2, 1] \times (1/p_3, 1]$.

2.5.2 Applying this idea to the surface (2.20)

To attempt something similar for the surfaces (2.20), we will choose

$$t_1 = t,\ t_2 = Lt + \delta_1,\ t_3 = L^2 t + \delta_2,\ ...,\ t_{r-k} = L^{r-k-1} t + \delta_{r-k-1}, \tag{2.23}$$

where the $\delta_1, ..., \delta_{r-k-1} \in [0, 1)$, and where L is an integer chosen suitably large so that

$$\rho_i \ := \ \theta_i(1, L, L^2, ..., L^{r-k-1}) \ \neq \ 0.$$

It isn't hard to see that one can take

$$L \ \leqslant \ 2 \cdot \mathrm{lcm}(n_1, ..., n_k) \max_{i,j} |m_{i,j}|.$$

Using the parameterization (2.23), we will get that

$$\theta_i(t_1, ..., t_{r-k}) \ = \ t \cdot \theta_i(1, L, L^2, ..., L^{r-k-1}) + \theta_i(0, \delta_1, ..., \delta_{r-k-1}) \ = \ t \cdot \rho_i + \theta_i(0, \delta_1, ..., \delta_{r-k-1})$$

and applying this to (2.20), we will get curves of the form

$$(p_1^{t-1}, d_2 p_2^{Lt-1},\ \ldots,\ d_{r-k} p_{r-k}^{L^{r-k-1}t-1},\ d_{r-k+1} p_{r-k+1}^{\rho_1 t-1},\ \ldots\ d_r p_r^{\rho_k t-1}), \tag{2.24}$$

where

$$d_2 = p_2^{\delta_1},\ d_3 = p_3^{\delta_2},\ \ldots,\ d_{r-k} = p_{r-k}^{\delta_{r-k-1}},$$

and

$$d_{r-k+1} = p_{r-k+1}^{\theta_1(0,\delta_1,\ldots,\delta_{r-k-1})} c_1,\ \ldots,\ d_r = p_r^{\theta_k(0,\delta_1,\ldots,\delta_{r-k-1})} c_k,$$

where the c_i are of the form $p_{r-k+i}^{s_i}$, where $s_i \in S$.

Similar to how we dealt with (2.22) not quite covering all possible curves, we actually need to expand the set of possibilities for the d_i, given a fixed choice for $\delta_1, \ldots, \delta_{r-k}$ (that is, our curves (2.24) don't quite cover everything): we need to also include dilates by integral powers of the p_i, to handle, for example, p_2^{Lt} not always being in the range $(1/p_2, 1]$ (basically, the exponent Lt needs to be considered mod 1). Thus, in fact, we want to consider the d_i's in dilated sets

$$d_2 \in p_2^{\delta_1} D_2,\ d_3 \in P_3^{\delta_2} D_3,\ \ldots,\ d_{r-k} \in p_{r-k}^{\delta_{r-k-1}} D_{r-k}, \tag{2.25}$$

and

$$d_{r-k+1} \in p_{r-k+1}^{\theta_1(0,\delta_1,\ldots,\delta_{r-k-1})} c_1 D_{r-k+1},\ \ldots,\ d_r \in p_r^{\theta_k(0,\delta_1,\ldots,\delta_{r-k-1})} c_k D_r, \tag{2.26}$$

where

$$D_j \ = \ \{p_j^i \ : \ |i| \leqslant I\},\ j = 2, \ldots, r,$$

where I is a suitably large integer. A trivial upper bound for I would be $rL^r \max_{i,j} |m_{i,j}|$.

Of course, with such a large collection of possible curves, some may fail to intersect $(1/p_1, 1] \times \cdots \times (1/p_r, 1]$, and so will not contain any points of the form (2.15) at all. It will not cause a problem, because all we were interested in was a set of disjoint curves that *do* cover all those points, and that can be suitably discretized later to prove certain theorems.

2.5.3 An important property of the parameterized curves

When all is said and done, the curves from the previous section that we generate have the form

$$(p_1^{t-1},\ e_2 p_2^{q_2 t},\ e_3 p_3^{q_3 t},\ \ldots,\ e_r p_r^{q_r t}),$$

where the q_j's are non-zero rational numbers. An important property here is the fact that $p_1, p_2^{q_2}, \ldots, p_r^{q_r}$ are all distinct, which fulfills a goal mentioned at the beginning of section 2.5. This property holds since if two of them were equal, we would have, for example,

$$q_i \log p_i \;=\; q_j \log p_j,$$

yet we know that the $\log p_i$'s are linearly independent over $\mathbb{Q}$.

2.6 Discretized curves

Now we produce discretized versions of the curves produced in section 2.5. We begin by defining $\mathcal{C}$ to be the set of all curves produced at the end of subsection 2.5 with the property that the curve has non-empty intersection with the set

$$\Gamma \;:=\; (1/p_1, 1] \;\times\; (1/p_2, 1] \;\times\; \cdots \;\times\; (1/p_r, 1].$$

Now, any curve in $\mathcal{C}$ may be parameterized by a vector

$$(\delta_1, \ldots, \delta_{r-k-1}, c_1, \ldots, c_k, \tau_2, \ldots, \tau_r), \tag{2.27}$$

where

$$c_i = p_{r-k+i}^{s_i},\ \text{where } s_i \in S,\ i = 1, \ldots, k;\ \text{and } \tau_i \in D_i,\ i = 2, \ldots, r. \tag{2.28}$$

The $\delta_1, ..., \delta_{r-k-1}$ can take on a continuum of values in $[0, 1)$, while the values taken on by the $c_1, ..., c_k, \tau_2, ..., \tau_r$ are finite in number.

Given a prime P satisfying

$$P \ > \ (\max_{j=1,...,r} p_j)^{2H},$$

we define a family of sets $\mathcal{F}$ as follows: for each choice of numbers $0 \leqslant f_1, ..., f_{r-k-1} \leqslant P - 1$, and choice of $c_1, ..., c_k, \tau_2, ..., \tau_r$ as above, let $F(f_1, ..., f_{r-k-1}, c_1, ..., c_k, \tau_2, ..., \tau_r)$ denote the set of all points

$$(x_1, x_2, ..., x_r) \ \in \ \{0, 1..., P-1\}^r,$$

such that there exists a curve in $\mathcal{C}$ with parameter vector (2.27), incident to a point $(y_1, ..., y_r) \in \Gamma$, such that

$$(\delta_1, ..., \delta_r) \ \in \ (f_1/P, ..., f_r/P) + [0, 1/P]^r, \tag{2.29}$$

and

$$(y_1, ..., y_r) \ \in \ (x_1/P, ..., x_r/P) + [0, 1/P]^r. \tag{2.30}$$

If this set $F(f_1, ..., f_{r-k-1}, c_1, ..., c_k, \tau_2, ..., \tau_r)$ is non-empty, then we add it to the family $\mathcal{F}$; otherwise, we don't.

One can easily see that, since there are at most P^{r-k-1} choices for $f_1, ..., f_{r-k-1}$, and since there are only a bounded number possibilities for the c_j's and τ_j's,

$$|\mathcal{F}| \ \lesssim \ P^{r-k-1}.$$

Likewise, for each choice of the f_j's, there is at least one choice of the other parameters making $F(f_1, ..., f_{r-k-1}, c_1, ..., c_k, \tau_2, ..., \tau_r)$ non-empty; and so, we have that

$$P^{r-k-1} \lesssim |\mathcal{F}| \lesssim P^{r-k-1}. \tag{2.31}$$

2.7 Two propositions and the proof of Theorem 2.3

The two propositions we will need to prove Theorem 2.3 are:

Proposition 2.32. *We have that for every point $x \in \{0, 1, 2, ..., P-1\}^r$,*

$$\#\{n \leqslant N \ : \ (\alpha_1(n), ..., \alpha_r(n)) \in x/P + [0, 1/P)^r\} \lesssim N|\mathcal{F}|^{-1}P^{-1}.$$

And:

Proposition 2.33. *Suppose*

$$K(t) = (\zeta_1\theta_1^t,\ \zeta_2\theta_2^t,\ ...,\ \zeta_r\theta_r^t) \tag{2.34}$$

and suppose that

$$F \subseteq \{0, 1, ..., P-1\}^r$$

is the set of all vectors such that if $(y_1, ..., y_r) = K(t)$, for some $t \in [0, 1)$, then there exists $(x_1, ..., x_r) \in F$ such that

$$(y_1, ..., y_r) \in \frac{1}{P}(x_1, x_2, ..., x_r) + [0, 1/P)^r.$$

Now, let

$$A_1,\ A_2,\ ...,\ A_r \subseteq \mathbb{F}_P, \text{ with } |A_1|, ..., |A_r| \geqslant P^{1-\varepsilon}.$$

We claim that for all but at most $o(P)$ elements $(x_1, ..., x_r) \in F$, there exist

$$1 \leqslant n \leqslant P^{7r^3\varepsilon}, \text{ and } (\delta_1, ..., \delta_r) \in \{0, 1, ..., [P^{7r^3\varepsilon}]\}^r,$$

such that

$$n\cdot(x_1,...,x_r)-(\delta_1,...,\delta_r)\ \in\ (A_1{+}A_1{+}A_2)\times(A_2{+}A_2{+}A_2)\times\cdots\times(A_r{+}A_r{+}A_r). \quad (2.35)$$

2.7.1 Completion of the proof of Theorem 2.3

Let $\beta=(\beta_1,...,\beta_r)\in[0,1)^r$ be arbitrary. Let $\beta'=(\beta'_1,...,\beta'_r)\in\{0,...,P-1\}^r$ be defined via

$$\frac{\beta'_j}{P}\ \leqslant\ \beta_j\ <\ \frac{\beta'_j+1}{P}. \quad (2.36)$$

Thus, β' is some kind of discretized version of $P\cdot\beta$.

We will later apply Proposition 2.33 using, for $j=1,...,r$,

$$\begin{aligned} A_j\quad :=\quad & -3^{-1}\beta'_j+\left\{d_1\lceil P/p_j\rceil+d_2\lceil P/p_j^2\rceil+\cdots+d_H\lceil P/p_j^H\rceil\ :\ 0\leqslant d_1,...,d_H<p_i/10\right\}\\ & +\{0,1,...,\lceil P/p_j^H\rceil-1\}.\end{aligned}$$

(Note that 3^{-1} denotes the multiplicative inverse of 3 in $\mathbb{F}_P$.) We note that

$$|A_j|\ \gg\ (p_j/10)^H(P/p_j^H)\ =\ P/10^H.$$

This follows from the fact that all the expressions

$$d_1\lfloor P/p_j\rfloor+\cdots+d_H\lfloor P/p_j^H\rfloor+x,\ \text{where } 0\leqslant d_1,...,d_H<p_j/10,$$

are unique mod P for arbitrary x (it is easy to see this, using a similar proof as the one showing base-p_j representations are unique).

We will have that, working in $\mathbb{F}_P$,

$$\begin{aligned} A_j+A_j+A_j\quad =\quad & -\beta'_j+\{e_1\lceil P/p_j\rceil+\cdots+e_H\lceil P/p_j^H\rceil\ :\ 0\leqslant e_1,...,e_H\leqslant 3\lfloor p_j/10\rfloor\}\\ & +\{0,1,...,3\cdot\lceil P/p_j^H\rceil-3\} \end{aligned} \quad (2.37)$$

Thinking of this set as a subset of $\{0, 1, 2, ..., P-1\}$, if we divide its elements by a factor P, then we get a set of numbers contained in the set

$$-\frac{\beta_j'}{P} + \left\{ \frac{e_1}{p_j} + \frac{e_2}{p_j^2} + \cdots + \frac{e_H}{p_j^H} \ :\ 0 \leqslant e_1, ..., e_H < \frac{3p_j}{10} - 1 \right\} + \text{error} + \mathbb{Z},$$

where the error is the sum total of the errors in approximating the $(e_j/P)\lceil P/p_j^i \rceil$ by e_j/p_j^i; this error is bounded from above by $Hp_j/3P$. It is clear, then, that $P^{-1}(A_j + A_j + A_j)$ is contained in the set

$$-\beta_j + \left\{ \frac{e_1}{p_j} + \frac{e_2}{p_j^2} + \cdots + \frac{e_H}{p_j^H} \ :\ 0 \leqslant e_1, ..., e_H < \frac{3p_j}{10} - 1 \right\} + \left[0,\ \frac{1 + Hp_j/3}{P}\right] + \mathbb{Z},$$

where the β_j satisfies (2.36).

We now let

$$F \ = \ F(f_1, ..., f_{r-k-1}, c_1, ..., c_k, \tau_2, ..., \tau_r) \ \in \mathcal{F} \tag{2.38}$$

be one of the sets in $\mathcal{F}$; F is thus a discretized version of a curve of general shape (2.34), where the ζ_j's depend on the choice of $f_1, .., f_{r-k-1}, c_1, ..., c_k, \tau_2, ..., \tau_r$.

Applying Proposition 2.33 to the curve F, using ε to be some function a little slower-decaying than $(\log P)^{-1}$ as a function of P – say, take $\varepsilon = (\log P)^{-1/2}$ – and then dividing (2.35) through by a factor P (interpreting coordinates now as integers instead of elements of $\mathbb{F}_P$), we get that for all but $o(P)$ of the $(x_1, ..., x_r) \in F$, if we let $(y_1, ..., y_r) \in [0, 1)$ be such that

$$\frac{x_j}{P} \ \leqslant \ y_j \ < \ \frac{x_j + 1}{P},$$

then there exists $1 \leqslant n \leqslant P^{7r^2\varepsilon}$ such that for $j = 1, ..., r$,

$$\begin{aligned} ny_j + \beta_j \ \in \ \frac{nx_j}{P} + \frac{\beta_j'}{P} + \left[0,\ \frac{2}{P}\right] \quad &\subseteq \quad P^{-1}(A_j + A_j + A_j) + \frac{\beta_j' + \delta_j}{P} + \left[0,\ \frac{2}{P}\right] + \mathbb{Z} \\ &\subseteq \quad U_j(H) + \mathbb{Z}, \end{aligned}$$

where, recall, $U_j(H)$ is defined in (2.4). Note that in deducing this last containment we have used the fact that $\delta_j/P < P^{7r^3\varepsilon-1}$, which is much smaller than $1/p_j^H$, the width of the interval in the definition of $U_j(H)$, using $\varepsilon = (\log P)^{-1/2}$. Taking fractional parts of both sides, we get that, for all $j = 1, ..., r$,

$$\{ny_j + \beta_j\} \in U_j(H). \tag{2.39}$$

We will use the notation

$$F = F^\flat \sqcup F^\sharp,$$

where $F^\flat$ denotes the exceptional set of $x \in F$ for which we don't get (2.39) holding for every $y \in x + [0, 1/P)$; and $F^\sharp$ denotes the rest of F. Note that from what we just proved, $|F^\flat| = o(P)$, and so $|F^\sharp| = |F| - o(P)$.

We will say that an integer $n \leqslant N$ is *good* if

$$\exists\ s \leqslant 2^{\varepsilon H}\ \forall\ j = 1, ..., r,\ \{s\alpha_j(n) + \beta_j(n)\} \in U_j(H)$$

and, otherwise, we will say that it is *bad*. We have that the number of $n \leqslant N$ that are bad is at most

$$\sum_{F \in \mathcal{F}} \sum_{x \in F^\flat} \#\{n \leqslant N\ :\ (\alpha_1(n), ..., \alpha_r(n)) \in x/P + [0, 1/P)^r\}$$

Applying Proposition 2.32 and (2.31) we get that this count is

$$\lesssim \sum_{F \in \mathcal{F}} |F^\flat| N P^{-r+k} = |\mathcal{F}| \cdot o(P) \cdot NP^{-r+k} = o(N).$$

This completes the proof of Theorem 2.3.

2.8 Proof of the Proposition 2.32

Fix a point $x = (x_1, ..., x_r) \in \mathbb{F}_P^r$. We will only focus on conting the $n \leqslant N$ such that $(\alpha_1(n), ..., \alpha_{r-k}(n))$ to belong to $(x_1, ..., x_{r-k})/P + [0, 1/P)^{r-k}$. This is legal, since the proposition only claims an upper bound.

Now, since $\alpha_j(n) = p_j^{\{n \log 2/\log p_j\}-1}$, in order for this to belong to $x_j/P + [0, 1/P)$, we need that $\{n \log 2/\log p_j\}$ belongs to a certain set $I_j + \mathbb{Z}$, where I_j is an interval of width at most $1/P \log p_j$. Thus, our goal is to count the number of $n \leqslant N$ such that

$$\left(\left\{\frac{n\log 2}{\log p_1}\right\} - 1,\ \left\{\frac{n\log 2}{\log p_2}\right\} - 1,\ ...,\ \left\{\frac{n\log 2}{\log p_{r-k}}\right\} - 1\right) \ \in\ I_1 \times I_2 \times \cdots \times I_{r-k} + \mathbb{Z}^{r-k}.$$

Now, from (2.16) and Theorem 2.10 we have that the number of such $n \leqslant N$ is, asymptotically,

$$N(|I_1| \cdots |I_{r-k}| + o(1)) \ \lesssim\ NP^{-r+k} \ \lesssim\ N|\mathcal{F}|^{-1}P^{-1},$$

where the last expression follows from (2.31). Note that the implied constants for the $\lesssim$'s depend on the p_j's.

This completes the proof since the upper bound on the set of $n \leqslant N$ has the form claimed by the proposition.

2.9 Proof of Proposition 2.33

Basic setup

For this proof we will use discrete Fourier methods. Given a function $f : \mathbb{F}_P^r \to \mathbb{C}$, and a vector $(a_1, ..., a_r) \in \{0, 1, 2, ..., P-1\}^r$, we define the Fourier transform

$$\widehat{f}(a_1, ..., a_r) \ :=\ \sum_{(n_1,...,n_r)\in\{0,1,...,P-1\}^r} f(n_1, ..., n_r)e^{2\pi i(a_1,...,a_r)\cdot(n_1,...,n_r)/P}.$$

A consequence of Parseval is that

$$\sum_{0 \leqslant s_1,...,s_r \leqslant P-1} |\widehat{1}_{A_1 \times A_2 \times \cdots \times A_r}(s_1, ..., s_r)|^2 \ = \ P^r |A_1| \cdots |A_r|.$$

Thus, if Q is the set of all places $(s_1, ..., s_r)$ where

$$|\widehat{1}_{A_1 \times A_2 \times \cdots \times A_r}(s_1, ..., s_r)| \ \geqslant \ P^{r(1-3\varepsilon)},$$

then

$$|Q| \ \leqslant \ P^{-2r(1-3\varepsilon)} P^r |A_1| \cdots |A_r| \ \leqslant \ P^{6r\varepsilon}$$

Let $Q' \subseteq Q$ be all those places $(s_1, ..., s_r) \in Q$, $|s_i| < P/2$, satisfying the additional constraint that

$$|s_i| \ \leqslant \ P^{1-(7r^3-r)\varepsilon}, \ i = 1, 2, ..., r. \tag{2.40}$$

Let $N = |Q'|$, and note that

$$N \ \leqslant \ |Q| \ \leqslant \ P^{6r\varepsilon}.$$

Now we let E denote the set of all $(x_1, ..., x_r) \in F$, such that there exists $(s_1, ..., s_r) \in Q'$, $(s_1, ..., s_r) \neq (0, ..., 0)$, such that

$$\left\| \frac{(x_1, ..., x_r) \cdot (s_1, ..., s_r)}{P} \right\| \ = \ \left\| \frac{x_1 s_1 + \cdots + x_r s_r}{P} \right\| \ < \ \frac{1}{P^{(7r^3-2r)\varepsilon}}. \tag{2.41}$$

Theorem follows if we can show $|E| = o(P)$

We will show that $|E| = o(P)$. If this holds, then let us see how it implies the conclusion of the Proposition: let $L = [\log P]$,

$$U \ := \ \{0, 1, 2, ..., [P^{7r^3\varepsilon}/L]\}^r,$$

and define $g(\vec{\delta}) = g(\delta_1, ..., \delta_r)$ to be the following L-fold convolution

$$g(\vec{\delta}) \ := \ 1_U * 1_U * \cdots * 1_U(\delta_1, ..., \delta_r).$$

Now, let

$$(x_1, ..., x_r) \ \in \ F \backslash E \tag{2.42}$$

be any of the $|F| - o(P)$ vectors such that (2.41) fails to hold, for every $(s_1, ..., s_r) \in Q'$. Let

$$M \ := \ [P^{7r^3\varepsilon}],$$

and let f be the indicator function for the set

$$\{(-nx_1, -nx_2, ..., -nx_r) \ : \ 1 \leqslant n \leqslant M\}.$$

Then, we have that if

$$1_{A_1 \times \cdots \times A_r} * 1_{A_1 \times \cdots \times A_r} * 1_{A_1 \times \cdots \times A_r} * g * f(\vec{0}) \ > \ 0, \tag{2.43}$$

then there exists $1 \leqslant n \leqslant M$ and $(\delta_1, ..., \delta_r)$, so that (2.35) holds.

Expressing the left-hand-side of (2.43) in terms of Fourier transforms, one sees that it equals:

$$\begin{aligned} P^{-r} & \sum_{(s_1,...,s_r) \in \mathbb{F}_P^r} \widehat{1}_{A_1 \times \cdots \times A_r}(s_1, ..., s_r)^3 \widehat{g}(s_1, ..., s_r) \widehat{f}(s_1, ..., s_r) \\ & = \ P^{-r} \sum_{\vec{s} \in \mathbb{F}_P^r} \widehat{1}_{A_1 \times \cdots \times A_r}(\vec{s})^3 \widehat{1}_U(\vec{s})^L \widehat{f}(\vec{s}). \end{aligned} \tag{2.44}$$

We split the terms in the second sum into the term with $(s_1, ..., s_r) = (0, ..., 0)$, the terms $(s_1, ..., s_r) \in Q$, and then the remaining terms.

The contribution of the term $(s_1, ..., s_r) = (0, ..., 0)$ is

$$P^{-r} M |U|^L |A_1|^3 \cdots |A_r|^3. \tag{2.45}$$

Now suppose $(s_1, ..., s_r) \in Q \backslash Q'$. Then, for some $i = 1, ..., r$ we have that $P^{1-(7r^3-r)\varepsilon} < |s_i| < P/2$. Thus,

$$|\widehat{g}(\vec{s})| \lesssim \prod_{i=1}^{r} \min(|U|^{L/r}, \|s_i/P\|^{-L}) < |U|^{L(r-1)/r} P^{(7r^3-r)\varepsilon L} \leqslant |U|^L ((L+1)P^{-r\varepsilon})^L.$$

It follows, then, that the contribution of all such $(s_1, ..., s_r) \in Q \backslash Q'$ to the right-hand-side of (2.44) is bounded from above by

$$P^{-r} N |A_1|^3 \cdots |A_r|^3 |U|^L ((L+1)P^{-r\varepsilon})^L M,$$

which is much smaller than (2.45), on account of the $((L+1)P^{-r\varepsilon})^L$ factor, even when using the crude upper bounds: $|A_i| \leqslant P$, $i = 1, ..., r$, and $M \leqslant P^{7r^3\varepsilon}$, $N \leqslant P^{6r\varepsilon}$.

Next, we consider the contribution of all terms with $(s_1, ..., s_r) \in Q'$. Then, since $(x_1, ..., x_r)$ satisfies (2.42), and in particular that it is not E, we have that

$$\begin{aligned} |\widehat{f}(s_1, ..., s_r)| &= \left| \sum_{1 \leqslant n \leqslant M} e^{2\pi i n (x_1, ..., x_r) \cdot (s_1, ..., s_r)/P} \right| \\ &\lesssim \frac{1}{\|(x_1, ..., x_r) \cdot (s_1, ..., s_r)/P\|} \\ &\leqslant P^{(7r^3-2r)\varepsilon}. \end{aligned}$$

So, the contribution of the terms in (2.44) with $(s_1, ..., s_r) \in Q'$, $(s_1, ..., s_r) \neq (0, ..., 0)$, is, by Parseval,

$$\begin{aligned}
&\lesssim \quad P^{-r}P^{(7r^3-2r)\varepsilon}|U|^L \sum_{0\leqslant s_1,\ldots,s_r\leqslant P-1} |\widehat{1}_{A_1\times\cdots\times A_r}(s_1,\ldots,s_r)|^3 \\
&\leqslant \quad P^{-r+(7r^3-2r)\varepsilon}|U|^L|A_1|\cdots|A_r| \sum_{0\leqslant s_1,\ldots,s_r\leqslant P-1} |\widehat{1}_{A_1\times\cdots\times A_r}(s_1,\ldots,s_r)|^2 \\
&\leqslant \quad P^{(7r^3-2r)\varepsilon}|U|^L|A_1|^2\cdots|A_r|^2 \\
&\lesssim \quad P^{-(1+\varepsilon)r}M|U|^L|A_1|^3\cdots|A_r|^3,
\end{aligned}$$

which is smaller than the contribution of the term with $(s_1, \ldots, s_r) = (0, \ldots, 0)$ given in (2.45).

Finally, we consider the contribution of the remaining terms. For these terms we have

$$|\widehat{1}_{A_1\times\cdots\times A_r}(s_1,\ldots,s_r)| \ < \ P^{r(1-3\varepsilon)} \ \leqslant \ |A_1|\cdots|A_r|P^{-2r\varepsilon}.$$

Using this in those terms on the right-hand-side of (2.44), we find that, using Parseval again, they contribute at most

$$\begin{aligned}
&P^{-r-2r\varepsilon}M|U|^L|A_1|\cdots|A_r| \sum_{0\leqslant s_1,\ldots,s_r\leqslant P-1} |\widehat{1}_{A_1\times\cdots\times A_r}(s_1,\ldots,s_r)|^2 \\
&\quad \leqslant \ P^{-2r\varepsilon}M|U|^L|A_1|^2\cdots|A_r|^2 \ \leqslant \ P^{-r-r\varepsilon}M|U|^L|A_1|^3\cdots|A_r|^3,
\end{aligned}$$

which is also appreciably smaller than the contribution of the term with $(s_1, \ldots, s_r) = (0, \ldots, 0)$, as in (2.45).

Thus, there exists $1 \leqslant n \leqslant M$ and $0 \leqslant \delta_1, \ldots, \delta_r \leqslant P^{7r^3\varepsilon}$ so that

$$n(x_1,\ldots,x_r) + (\delta_1,\ldots,\delta_r) \ \in \ (3A_1)\times(3A_2)\times\cdots\times(3A_r).$$

And since this holds for $(1-o(1))|F|$ vectors $(x_1, \ldots, x_r) \in F$, the proposition is proved.

Proving $|E| = o(P)$

We begin by noting that we may assume that Q' contains at least one non-zero vector, since otherwise in the previous subsection we never need to make use of bounds on $|\widehat{f}(s_1, ..., s_r)|$, nor reference to $(x_1, ..., x_r)$ – we obtain the same bounds independent of choice of $(x_1, ..., x_r)$, which would imply that E is empty.

We note, by the pigeonhole principle, that there exist $(s_1, ..., s_r) \in Q'$, such that (2.41) holds for at least $|E|/N$ vectors $(x_1, ..., x_r) \in E$. Call this new set of vectors $E' \subseteq E$; so, we have

$$|E'| \geqslant |E|/N.$$

Let us suppose, for proof by contradiction, that

$$|E|/N > P^{1-7r\varepsilon}. \tag{2.46}$$

Note that if we establish a contradiction, then we would be forced to conclude that

$$|E| \leqslant NP^{1-7r\varepsilon} \leqslant P^{1-r\varepsilon},$$

which would imply $|E| = o(P)$, and which is just what we wanted to show.

For each $\vec{x} = (x_1, ..., x_r) \in E'$, let $t = t(\vec{x})$ be any value of t, so that if $\vec{y} = K(t)$, then

$$\vec{y} \in (x_1/P, ..., x_r/P) + [0, 1/P)^r. \tag{2.47}$$

Also, for any vector $\vec{v} \in [0, 1)^r$, let $\pi(v)$ denote the unique $\vec{x} \in \{0, ..., P-1\}^r$, so that

$$\vec{v} \in \frac{1}{P}\vec{x} + \left[0, \frac{1}{P}\right]^r.$$

Now, if we consider the set of all points in a cube

$$\vec{w} + \left[0, \frac{1}{P}\right]^r, \tag{2.48}$$

where $\vec{w}$ is some arbitrary r-dimensional vector, the function π will map that set to a set of size at most 2^r. Thus, if we let

$$T \ := \ \{t(\vec{x}) \ : \ \vec{x} \in E'\},$$

then we claim that any interval of width P^{-1} can have at most $2^r \log P$ elements of T. The reason this holds is that if we restrict t to an interval I of width at most $(P \log P)^{-1}$, then the coordinates of $K(t)$ will vary by $o(1/P)$; and so, the set $\{K(t) \ : \ t \in I\}$ will be contained in one of the cubes (2.48), which can correspond to at most 2^r vectors $\vec{x} \in \{0, 1, ..., P-1\}^r$.

By picking at most one element of T in each interval of width P^{-1}, we can pass to a subset

$$T' \ \subseteq \ T, \text{ where } |T'| > 2^{-r}|T|(\log P)^{-1} = 2^{-r}|E'|(\log P)^{-1} > P^{1-7r\varepsilon-o(1)},$$

such that every pair of elements of T' is at least $1/P$ apart.

Furthermore, we eliminate the elements of T' that are $\leqslant P^{-8r\varepsilon}$ in size. Call this new set $T'' \subseteq T'$. There can be at most $P^{1-8r\varepsilon+o(1)}$ elements in T' that are $\leqslant P^{-8r\varepsilon}$; and so,

$$|T''| \ \geqslant \ |T'| - P^{1-8r\varepsilon+o(1)} \ \geqslant \ P^{1-7r\varepsilon-o(1)}.$$

Now we index the elements or T'' as follows:

$$T'' \ := \ \{t_1, t_2, ..., t_n\},$$

where

$$t_1 \ < \ t_2 \ < \ \cdots \ < \ t_n.$$

Then, we extract disjoint subsets $T_1, ..., T_r \subseteq T''$ as follows: we let

$$T_i \ := \ \{t_j \ : \ (2i-2)n/2r < j < (2i-1)n/2r\},$$

which satisfies

$$|T_i| \ \gg \ n/r \ \gg \ |T''| \ > \ P^{1-7r\varepsilon-o(1)} \tag{2.49}$$

Let

$$d(T_i, T_j) \ := \ \min_{t \in T_i, u \in T_j} |t-u|.$$

Since the elements of T'' are spaced at least $1/P$ apart, we must have that

$$\min_{1 \leqslant i < j \leqslant r} d(T_i, T_j) \ \geqslant \ n/2rP \ > \ P^{-7r\varepsilon - o(1)}. \tag{2.50}$$

Define, also, the associated intervals

$$I_i \ := \ [t_{\lceil (2i-2)n/2r \rceil}, \ t_{\lfloor (2i-1)n/2r \rfloor}].$$

Note that if $t \in T_i$, then $t \in I_i$.

We now define $u_1, ..., u_r$ as follows: we let u_i be any element in the interval I_i such that $|h'(u)|$ is minimal, where

$$h(t) \ := \ (s_1, ..., s_r) \cdot K(t) \ = \ s_1 \zeta_1 \theta_1^t + \cdots + s_r \zeta_r \theta_r^t.$$

Note that

$$h'(t) \ := \ s_1 \zeta_1 \theta_1^t \log \theta_1 + s_2 \zeta_2 \theta_2^t \log \theta_2 + \cdots + s_r \zeta_r \theta_r^t \log \theta_r.$$

Bundling together $h'(u_1), ..., h'(u_r)$, we get the following matrix equation

$$\begin{bmatrix} \theta_1^{u_1} & \theta_2^{u_1} & \cdots & \theta_r^{u_1} \\ \theta_1^{u_2} & \theta_2^{u_2} & \cdots & \theta_r^{u_2} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_1^{u_r} & \theta_2^{u_r} & \cdots & \theta_r^{u_r} \end{bmatrix} \begin{bmatrix} s_1\zeta_1 \log\theta_1 \\ s_2\zeta_2 \log\theta_2 \\ \vdots \\ s_r\zeta_r \log\theta_r \end{bmatrix} = \begin{bmatrix} h'(u_1) \\ h'(u_2) \\ \vdots \\ h'(u_r) \end{bmatrix}. \tag{2.51}$$

Now we need the following lemma (which makes use of an idea from [30, page 99, book 2, example 1], though our proof is self-contained):

Lemma 2.52. *Let*

$$0 < x_1 < x_2 < \cdots < x_r, \text{ and } 0 < y_1 < y_2 < \cdots < y_r$$

be two sets of increasing real numbers. Define the matrix

$$A := \begin{bmatrix} x_1^{y_1} & x_2^{y_1} & \cdots & x_r^{y_1} \\ x_1^{y_2} & x_2^{y_2} & \cdots & x_r^{y_2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{y_r} & x_2^{y_r} & \cdots & x_r^{y_r} \end{bmatrix}.$$

Let

$$\sigma := \min_{\substack{(c_1,\ldots,c_r) \\ \|(c_1,\ldots,c_r)\|_2 = 1}} \|[c_1 \ \cdots \ c_r] \cdot A\|_2 = \min_{\substack{(c_1,\ldots,c_r) \\ \|(c_1,\ldots,c_r)\|_2 = 1}} \left\|[c_1 \ \cdots \ c_r] \cdot A^T\right\|_2.$$

Then,

$$\sigma \geqslant r^{-r+1/2}(x_r+1)^{-(r-1)y_r}(x_r/x_1)^{-y_1} x_1^{y_1+y_2+\cdots+y_r}\sigma_0, \tag{2.53}$$

where

$$\sigma_0 := \min_{i=1,\ldots,r-1} \prod_{1 \leqslant j < \ell \leqslant r} ((x_{i+1}/x_i)^{y_\ell/(r-1)} - (x_{i+1}/x_i)^{y_j/(r-1)}).$$

Note that when $r = 1$ this gives $\sigma \geqslant x_1^{y_1}$ (and $\sigma_0 = 1$ since the minimum is empty), and it *is easy to see that σ exactly equals $x_1^{y_1}$ in this case.*

Applying this lemma to (2.51), using $x_i = \theta_i$ and $y_i = u_i$, $i = 1, ..., r$, reordering the columns as necessary (because $\theta_1, \theta_2, ...$ may not be in increasing order), and shuffling the ordering of the coordinates of the column vector in left-hand-side of (2.51) accordingly (if you reorder the columns of the square matrix, you have to do the same for the column vector), we conclude that

$$\begin{aligned} \|(h'(u_1), ..., h'(u_r))\|_2 &\geqslant \sigma \|(s_1\zeta_1 \log\theta_1, s_2\zeta_2 \log\theta_2, ..., s_r\zeta_r \log\theta_r)\|_2 \\ &\geqslant \sigma \cdot \min_i |\zeta_i \log\theta_i| \cdot \|(s_1, ..., s_r)\|_2, \end{aligned}$$

where σ satisfies (2.53). Letting $i = 1, ..., r$ be any value where $|h'(u_i)|$ is maximal; that is,

$$|h'(u_i)| \;=\; \max_{j=1,...,r} |h'(u_j)|,$$

we will have that for every $t \in I_i$,

$$|h'(t)| \;\geqslant\; |h'(u_i)| \;\geqslant\; r^{-1/2}\|(h'(u_1), ..., h'(u_r))\|_2 \;\geqslant\; r^{-1/2}\sigma \cdot \min_j |\zeta_j \log\theta_j| \cdot \|(s_1, ..., s_r)\|_2. \tag{2.54}$$

By the Cauchy-Schwarz inequality we also have the following upper bound that holds for any $t \in I_i$:

$$|h'(t)| \;\leqslant\; r^{1/2} \max_j |\zeta_j \theta_j^t \log\theta_j| \cdot \|(s_1, ..., s_r)\|_2.$$

We wish to bound σ from below. First, note that for $\alpha > 1$ and $0 < u < t < 1$,

$$\begin{aligned} \alpha^t - \alpha^u \;=\; \alpha^u(\alpha^{t-u} - 1) &= \alpha^u(e^{(t-u)\log\alpha} - 1) \\ &> \alpha^u(t-u)\log\alpha. \end{aligned}$$

Thus, since

$$P^{-8r\varepsilon} \;<\; u_1 \;<\; \cdots \;<\; u_r \;\leqslant\; 1,$$

and since (2.50) holds, one sees that for any $i, i' = 1, ..., r$, and $\theta_i/\theta_{i'} > 1$, $j < \ell$,

$$\begin{aligned}(\theta_i/\theta_{i'})^{u_\ell/(r-1)} - (\theta_i/\theta_{i'})^{u_j/(r-1)} \quad &> \quad (\theta_i/\theta_{i'})^{u_j/(r-1)} \frac{(u_\ell - u_j)(\log \theta_i/\theta_{i'})}{r-1} \\ &> \quad \kappa P^{-7r\varepsilon - o(1)} \log \kappa,\end{aligned}$$

where

$$\kappa \;:=\; \min_{\substack{i,i',j=1,\ldots,r \\ \theta_i > \theta_{i'}}} (\theta_i/\theta_{i'})^{u_j/(r-1)}.$$

Thus, (2.53) implies

$$\sigma \;>\; P^{-7r^3\varepsilon/2 - o(1)}.$$

(The implied constants in the $o(1)$ depend on r, ε, the x_i's and y_i's; the term $o(1)$ tends to 0 as $P \to \infty$.) It follows from (2.54) that for every $t \in I_i$,

$$|h'(t)| \;>\; P^{-7r^3\varepsilon/2 - o(1)} \|(s_1, ..., s_r)\|_2. \tag{2.55}$$

In particular, this means that $h'(t) \neq 0$ for all $t \in I_i$, so that $h(t)$ is either strictly increasing on the interval I_i, or strictly decreasing on the interval I_i.

Now, for $t \in T_i$ we have that

$$|h(t)| \;=\; |(s_1, ..., s_r) \cdot K(t)| \;\leqslant\; \|(s_1, ..., s_r)\|_2 \|K(t)\|_2 \;\lesssim\; \|(s_1, ..., s_r)\|_2. \tag{2.56}$$

Applying (2.41) and (2.47), we also conclude that

$$\|h(t)\| \lesssim \; P^{-(7r^3 - 2r)\varepsilon}, \tag{2.57}$$

where $\|\cdot\|$ denotes the nearest integer function.

Now, combining (2.49) and (2.56), and applying the Pigeonhole Principle, we let T_i' be a maximal subset of T_i where the nearest integer to all the $h(t)$, $t \in T_i'$, is the same. We have that

$$|T_i'| \; > \; P^{1-7r\varepsilon-o(1)} \|(s_1, ..., s_r)\|_2^{-1}. \tag{2.58}$$

And from (2.57) we have that if z is the nearest integer to all the elements of T_i', we get that

$$\text{For every } t \in T_i', \; |h(t) - z| \; \leqslant \; P^{-(7r^3-2r)\varepsilon}. \tag{2.59}$$

However, we will see that this cannot hold, by using the Mean Value Theorem and the bound (2.55): without loss of generality, assume $h(t)$ is *increasing* in I_i (we know it is either increasing or decreasing, and it doesn't matter which). Write the set T_i' in increasing order as

$$t_1' \; < \; t_2' \; < \; \cdots \; < \; t_{n'}'.$$

Since h in increasing across this set, we have that

$$h(t_1') \; < \; h(t_2') \; < \; \cdots \; < \; h(t_{n'}').$$

Now, from the Mean Value Theorem, (2.55), the fact that the t_j''s are spaced at least $1/P$ apart, and our bound on $|T_i'|$ in (2.58), we have that

$$\begin{aligned} |h(t_1') - h(t_{n'}')| \; \gg \; |t_1' - t_{n'}'| \min_{t \in [t_1', t_{n'}']} |h'(t)| \quad &\geqslant \quad (n'/P) P^{-7r^3\varepsilon/2 - o(1)} \|(s_1, ..., s_r)\|_2 \\ &\geqslant \quad P^{-(7r^3/2+7r)\varepsilon - o(1)}. \end{aligned}$$

This is impossible, since from (2.59) we deduce from the triangle inequality that

$$|h(t_1') - h(t_{n'}')| \; \lesssim \; P^{-(7r^3-2r)\varepsilon}.$$

We conclude that (2.46) is false, and so our theorem is proved.

2.9.1 Proof of Lemma 2.52

The claim clearly holds for $r = 1$. Assume we've proved it for all $1 \leqslant r \leqslant k$. Now we prove it for $r = k+1$: So, we assume we have a matrix of that size; and assume, for proof by contradiction, that (2.53) fails to hold.

We let $(c_1, ..., c_{k+1})$ denote a vector of norm 1 such that

$$\|[c_1 \ \cdots \ c_{k+1}] \cdot A\| \ = \ \sigma.$$

Define

$$f(x) \ := \ \sum_{j=1}^{k+1} c_j x^{y_j},$$

and note that since $f(x_i)$ is the ith coordinate of $[c_1 \ \cdots \ c_{k+1}] \cdot A$, we must have

$$|f(x_i)| \ \leqslant \ \sigma, \ i = 1, ..., k+1, \tag{2.60}$$

all of which are rather small in magnitude, since we are assuming (2.53) fails to hold, making σ very small. We wish to show (for reasons explained below) that there exist $z_1, ..., z_k$, where

$$x_i \ < \ z_i \ < \ x_{i+1}, \ i = 1, 2, ..., k,$$

such that

$$|f(z_i)| \ > \ (z_i/x_i)^{y_1}|f(x_i)|, \text{ and } |f(z_i)| \ > \ |f(x_{i+1})|, \ i = 1, ..., k.$$

To see that such z_i exist, let

$$\delta \ = \ \frac{\log x_{i+1}}{\log x_i} \ - \ 1,$$

and note that

$$x_i^{1+\delta} \;=\; x_{i+1}.$$

Then, consider the numbers

$$f(x_i),\; f(x_i^{1+\delta/k}),\; f(x_i^{1+2\delta/k}),\; ...,\; f(x_i^{1+\delta}) = f(x_{i+1}).$$

Written as a row vector we have

$$[f(x_i)\; f(x_i^{1+\delta/k})\; \cdots\; f(x_{i+1})] \;=\; [c_1\; c_2\; \cdots\; c_{k+1}] \cdot V,$$

where

$$V \;:=\; \begin{bmatrix} x_i^{y_1} & (x_i^{y_1})^{1+\delta/k} & \cdots & (x_i^{y_1})^{1+\delta} \\ x_i^{y_2} & (x_i^{y_2})^{1+\delta/k} & \cdots & (x_i^{y_2})^{1+\delta} \\ \vdots & \vdots & \ddots & \vdots \\ x_i^{y_{k+1}} & (x_i^{y_{k+1}})^{1+\delta/k} & \cdots & (x_i^{y_{k+1}})^{1+\delta} \end{bmatrix}.$$

The square matrix V here is a Vandermonde (well, after dividing out by certain factors down columns), so its determinant can be explicitly computed:

$$\det(V) \;=\; x_i^{y_1+y_2+\cdots+y_{k+1}} \prod_{1\leqslant j<\ell\leqslant k+1} (x_i^{y_\ell\delta/k} - x_i^{y_j\delta/k}). \tag{2.61}$$

Letting $J = VV^T$, we then also have

$$\det(J) \;=\; \det(V)^2 \;=\; x_i^{2y_1+2y_2+\cdots+2y_{k+1}} \prod_{1\leqslant j<\ell\leqslant k+1} (x_i^{y_\ell\delta/k} - x_i^{y_j\delta/k})^2.$$

A crude upper bound on the largest eigenvalue of J can be found as follows: let μ be the maximum value of the entries of J. We note that

$$\mu \;\leqslant\; (k+1) \max_{j,\ell=1,\ldots,k+1} |x_j^{y_\ell}|^2 \;\leqslant\; (k+1)(x_{k+1}+1)^{2y_{k+1}}.$$

Then, for any vector $\vec{v} := (v_1, ..., v_{k+1})$ satisfying $\|\vec{v}\|_2 = 1$, we have that all the entries of $J\vec{v}$ can be bounded from above by

$$\mu\|\vec{v}\|_1 \ \leqslant\ \mu(k+1)\|\vec{v}\|_\infty.$$

Thus,

$$\mu(k+1) \ \leqslant\ (k+1)^2(x_{k+1}+1)^{2y_{k+1}}$$

is an upper bound for any eigenvalue for J.

Also, if $\alpha > 0$ is the smallest eigenvalue (in magnitude) of J, and $\beta > 0$ the largest eigvenvalue (in magnitude) of J, then since $\det(J)$ is the product of its eigenvalues,

$$\det(J) \ \leqslant\ \alpha\beta^k.$$

So, recalling that $\|(c_1, ..., c_{k+1})\|_2 = 1$, we have

$$\begin{aligned}
\|(f(x_i), f(x_i^{1+\delta/k}), ..., f(x_i^{1+\delta}))\|_2^2 \ &=\ (c_1, ..., c_{k+1})VV^T(c_1, ..., c_{k+1})^T \\
&\geqslant\ \alpha \\
&\geqslant\ \det(J)\cdot\beta^{-k} \\
&>\ \det(J)\cdot((k+1)^2(x_{k+1}+1)^{2y_{k+1}})^{-k} \\
&=\ \det(V)^2((k+1)^2(x_{k+1}+1)^{2y_{k+1}})^{-k}.
\end{aligned}$$

Note that the first inequality here is by the Rayleigh Principle:

$$\min_{\substack{(c_1,...,c_{k+1}) \\ \|(c_1,...,c_{k+1})\|=1}} (c_1, ..., c_{k+1})J(c_1, ..., c_{k+1})^T \ =\ \alpha.$$

Thus,

$$\max_{j=0,...,k}|f(x_i^{1+j\delta/k})| \ \geqslant\ (k+1)^{-k-1/2}(x_{k+1}+1)^{-ky_{k+1}}|\det(V)|.$$

Now, since we are operating under the assumption that (2.53) fails to hold for $r = k+1$, expressing (2.60) in terms of $\det(V)$ (and using (2.61)), we find that

$$|f(x_i)| \leqslant \sigma \leqslant (k+1)^{-k-1/2}(x_{k+1}+1)^{-ky_{k+1}}(x_{k+1}/x_1)^{-y_1}|\det(V)|.$$

Thus

$$\max_{j=0,\dots,k}|f(x_i^{1+j\delta/k})| \geqslant (x_{k+1}/x_1)^{y_1}|f(x_i)|, \text{ and } \geqslant (x_{k+1}/x_1)^{y_1}|f(x_{i+1})|.$$

We therefore have found the z_i we were looking for, since: First, for $j = 0, ..., k$ we have that

$$x_i \leqslant z_i := x_i^{1+j\delta/k} \leqslant x_{i+1},$$

where the j arising from the max above cannot be $j=0$ or $j=k$, since the max is bigger than $|f(x_i)|$ and $|f(x_{i+1})|$. And, second, we also have

$$|f(z_i)| \geqslant (x_{k+1}/x_1)^{y_1}|f(x_i)| > (z_i/x_i)^{y_1}|f(x_i)|,$$

and

$$|f(z_i)| > |f(x_{i+1})|.$$

Now, if we let

$$g(x) := x^{-y_1}f(x),$$

then note that for this choice of j (chosen by the max above),

$$|g(z_i)| = z_i^{-y_1}|f(z_i)| > x_i^{-y_1}|f(x_i)| = |g(x_i)|,$$

and, likewise,

$$|g(z_i)| \ = \ z_i^{-y_1}|f(z_i)| \ \geqslant \ z_i^{-y_1}|f(x_{i+1})| \ > \ x_{i+1}^{-y_1}|f(x_{i+1})| \ = \ |g(x_{i+1})|.$$

Thus, by Rolle's Theorem, there exists a point $w_i \in (x_i, x_{x+1})$ where the derivative

$$g'(w_i) \ = \ 0, \ i = 1, 2, ..., k. \tag{2.62}$$

But, since

$$g(w) \ = \ w^{-y_1}f(w) \ = \ \sum_{j=1}^{k+1} c_j w^{y_j - y_1},$$

we find that

$$g'(w) \ = \ \sum_{j=2}^{k+1} c_j(y_j - y_1)w^{y_j - y_1};$$

so, we have that

$$\begin{bmatrix} w_1^{y_2-y_1} & w_1^{y_3-y_1} & \cdots & w_1^{y_{k+1}-y_1} \\ w_2^{y_2-y_1} & w_2^{y_3-y_1} & \cdots & w_2^{y_{k+1}-y_1} \\ \vdots & \vdots & \ddots & \vdots \\ w_k^{y_2-y_1} & w_k^{y_3-y_1} & \cdots & w_k^{y_{k+1}-y_1} \end{bmatrix} \begin{bmatrix} c_2(y_2 - y_1) \\ c_3(y_3 - y_1) \\ \vdots \\ c_{k+1}(y_{k+1} - y_1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

However, the induction hypotheses for the case $r = k$ tells us that the square matrix on the left is non-singular (in fact, it gives a non-trivial lower bound, in magnitude, for its smallest singular value). So this is impossible.

We conclude that our assumption that (2.53) was wrong is incorrect, and so the induction is proved.

CHAPTER 3

ON A CLASS OF SUMS WITH UNEXPECTEDLY HIGH CANCELLATION, AND ITS APPLICATIONS

3.1 Introduction

Remember that the Pentagonal Number Theorem of Euler asserts that for an integer $x \geqslant 2$,

$$\sum_{G_n \leqslant x} (-1)^n p(x - G_n) = 0$$

where $G_n = \frac{n(3n-1)}{2}$ is nth pentagonal number. Various proofs of this theorem have been developed over the decades and centuries (see [7]); but we wondered whether it was possible to produce an "analytic proof", using the Ramanujan-Hardy-Rademacher formula (1.8) for $p(x)$: Considering just the first two terms in this formula, one sees that (see [31])

$$p(x) = p_2(x) + O(\sqrt{p(x)}), \quad \text{where} \quad p_2(x) = \frac{\sqrt{12}e^{\frac{\pi}{6}\sqrt{24x-1}}}{24x-1}\left(1 - \frac{6}{\pi\sqrt{24x-1}}\right).$$

So, the Pentagonal Number Theorem implies

$$\sum_{G_n \leqslant x} (-1)^n p_2(x - G_n) \lesssim \sqrt{xp(x)}; \tag{3.1}$$

In fact, one can get a better bound by using more terms in the Ramanujan-Hardy-Rademacher expression; one might call this a "Weak Pentagonal Number Theorem", which is an interesting and non-trivial bound for the size of this oscillating sum of exponential functions $(-1)^n p_2(x - G_n)$.

It is worth pointing out that this bound is much smaller than what would be expected on probabilistic grounds: if we consider a sum

$$S(X_1, X_2, \cdots) = \sum_{G_n < x} X_n p_2(x - G_n),$$

where the X_n's are independent random variables taking the values $+1$ and -1, each with probability $\frac{1}{2}$, then

$$E(S^2) = \sum_{G_n \leqslant x} p_2(x - G_n)^2.$$

So the quality of bound we would expect to prove is

$$|S| \lesssim \left(\sum_{G_n \leqslant x} p_2(x - G_n)^2 \right)^{\frac{1}{2}} \lesssim \sqrt[4]{x} p(x),$$

However, the bound (3.1) is much smaller than the RHS here.

What we have discovered is that (3.1) is just the tip of the iceberg, and that there is a very general class of sums like this that are small - much smaller than one would guess based on a probabilistic heuristic. Roughly, we will prove that

$$\sum_{f(n) \leqslant x} (-1)^n e^{c\sqrt{x - f(n)}} = \text{"small"}, \tag{3.2}$$

where f is a quadratic polynomial (with positive leading coefficient), and c is some constant. It is possible to produce a more general class of sums with a lot of cancellation; and we leave it to the reader to explore. As a consequence of this and the Ramanujan-Hardy-Rademacher expansion for $p(n)$, we will prove that

$$\sum_{l^2 < x} (-1)^l p(x - l^2) \sim \frac{e^{\pi i x} \sqrt{p(x)}}{2^{3/4} x^{1/4}}. \tag{3.3}$$

As another category of results, we will also prove a corollary of Theorem 3.8 related to prime numbers. In fact let $x > 0$ be large enough and $T = e^{0.786\sqrt{x}}$. Then

$$\sum_{0\leqslant \ell < \frac{1}{2}\sqrt{xT}} \Psi\left(\left[e^{\sqrt{x-\frac{(2\ell)^2}{T}}},\ e^{\sqrt{x-\frac{(2\ell-1)^2}{T}}}\right]\right) = \Psi(e^{\sqrt{x}})\left(\frac{1}{2} + O\left(e^{-0.196\sqrt{x}}\right)\right).$$

Finally we will develop polynomial identities that occur naturally in the Taylor expansion in (3.2). For example

$$\sum_{|\ell|\leqslant x}(4x^2-4\ell^2)^{2r} - \sum_{|\ell|<x}(4x^2-(2\ell+1)^2)^{2r} = \text{polynomial w.r.t. } x \text{ with degree } 2r-1.$$

Many of the results stated above can be deduced from the following:

Theorem 3.4. *Let $b, d \in \mathbb{R}$, $a, c > 0$; Also, let $h(x) = (\alpha x + \beta)^t$ for $\alpha, \beta, t \in \mathbb{R}$. Then*

$$\sum_{n: an^2+bn+d<x} (-1)^n e^{c\sqrt{x-(an^2+bn+d)}} h(x-(an^2+bn+d)) \lesssim e^{(w+\epsilon)c\sqrt{x}}. \tag{3.5}$$

where $w > 0$ is defined as follows. Set

$$\Delta := \sup_{r\geqslant 0} \sqrt{\sqrt{a}r\frac{\sqrt{ar^2+4}+r\sqrt{a}}{2}} - \frac{\pi r}{c}$$

Then $w = \min(1, \Delta)$.

Remark 3.6. *Obviously forcing w to be less than one is to avoid getting a trivial result, and if a, c, r are chosen in such a way that $\Delta > 1$ then this theorem becomes useless.*

Conjecture 2. *Observing the numerical results suggest that*

$$\sum_{\ell^2<x} (-1)^\ell e^{\sqrt{x-\ell^2}} = e^{o(\sqrt{x})}.$$

There is another generalization when we pick a complex c in (3.5). In this case, having an upper bound for the sum is harder, as we have both the fast growth of exponential functions and the extra oscillation coming from the imaginary exponent.

Theorem 3.7. *For large enough $x > 0$, let $T := T(x)$ be at least $\Omega(x^2)$ as $x \to \infty$. Also let $\alpha + i\beta \in \mathbb{C}$ and $0 \leqslant \alpha < 1 + \epsilon$ for a fixed $\epsilon > 0$, and $\beta < \sqrt{T}$. Then for arbitrary $\delta > 0$*

$$\sum_{l^2 < Tx} (-1)^l e^{(\alpha+i\beta)\sqrt{x - \frac{l^2}{T}}} \lesssim \sqrt{\frac{Tx}{|\beta|+1} e^{\alpha(\sqrt{\frac{2}{2+\pi^2}}+\delta)\sqrt{x}}} + \sqrt{T}.$$

Note that if $\beta = 0$ and T sufficiently large, theorem 3.7 becomes a special case of theorem 3.4 for $a = 1$, $b, d = 0$, and $c \to 0$ with a weaker result.

Even these theorems do not exhaust all the cancellation types of oscillatory sums of this form, for we can replace the square-root by a fourth- root, and then replace the quadratic polynomial with a quartic. We will not bother to develop the most general theorem possible here. Next, we prove three applications for these oscillation sums.

3.1.1 Applications to the Chebyshev Ψ function

We show that in the "Weak pentagonal number theorem" we can replace the partition function $p(n)$ with Chebyshev Ψ function.

Theorem 3.8. *Assume $\epsilon > 0$, x is large enough and $a = 1 - \sqrt{\frac{2}{2+\pi^2}}$. We have*

$$\sum_{l^2 < xe^{\frac{4a}{3}\sqrt{x}}} (-1)^l \Psi\left(e^{\sqrt{x - l^2 e^{-\frac{2a}{3}\sqrt{x}}}}\right) \lesssim e^{(1-\frac{a}{3}+\epsilon)\sqrt{x}} := e^{w\sqrt{x}}. \tag{3.9}$$

We give an argument to show a relation between the theorem and the distribution of prime numbers. A weak version of theorem can be written as

$$\Psi(e^{\sqrt{x}}) = 2 \sum_{0 < \ell < \sqrt{xT}/2} \left(\Psi\left(e^{\sqrt{x - \frac{(2\ell-1)^2}{T}}}\right) - \Psi\left(e^{\sqrt{x - \frac{(2\ell)^2}{T}}}\right)\right) + O\left(e^{(\frac{5}{6}+\epsilon)\sqrt{x}}\right) \quad \text{where } T := e^{\frac{2\sqrt{x}}{3}}.$$

Define

$$I := \bigcup_{0<\ell<\sqrt{x}T/2} \left(e^{\sqrt{x-\frac{(2\ell)^2}{T}}}, e^{\sqrt{x-\frac{(2\ell-1)^2}{T}}} \right)$$

One can see that the measure of I is almost half of the length of the interval $[0, e^{\sqrt{x}}]$. Roughly speaking theorem 3.8 states that the number of primes in I, with weight $\log(p)$, is half of the number of primes, with the same weight. This prime counting gives a stronger result than one would get using a strong form of the Prime Number Theorem and also the Riemann Hypothesis(RH), where one naively estimates the Ψ function on each of the intervals. Because the widths of the intervals are smaller than $e^{\frac{\sqrt{x}}{2}}$, making the Riemann Hypothesis estimate "trivial". However, a less naive argument can give an improvement like corollary 3.10. See table 3.1 for comparison.

PNT	Naive RH + Theorem 3.4	Our unconditional result	Our result with RH
1.41	0.91	0.79	0.47

Table 3.1: The upper bound of w in (3.9)

Corollary 3.10. *Assuming RH*

$$\sum_{l^2 < xe^{\frac{2(1+a)}{3}\sqrt{x}}} (-1)^l \Psi \left(e^{\sqrt{x - l^2 e^{-\frac{4(1+a)}{3}\sqrt{x}}}} \right) \lesssim e^{(\frac{2}{3} - \frac{a}{3} + \epsilon)\sqrt{x}} \lesssim e^{0.47\sqrt{x}}. \qquad (3.11)$$

The proof needs careful computations of a cancellation sum involving zeroes of the Riemann zeta function. In fact, we use our cancellation formula to control the low-height zeroes; The Van der Corput bound for exponential sums combined with the Montgomery Mean-value theorem to control the high-height zeroes.

Remark 3.12. *Note that numerical results up to $x < 300$ show a very smaller error term* *in comparison to* (3.11). *In particular, for example,*

$$\left|\sum_{l<2400}(-1)^l\Psi(e^{\sqrt{300-l^2/T}})\right| < 50 \quad \textit{where } T \sim 20000.$$

Remark 3.13. *A more applicable identity may be the case with fewer terms (with lower frequency) in* (3.9). *We can choose the parameters to get*

$$\sum_{l^2<xe^{2\epsilon\sqrt{x}}}(-1)^l\Psi(e^{\sqrt{x-l^2e^{-2\epsilon\sqrt{x}}}}) \lesssim x^2e^{(1-\epsilon)\sqrt{x}}.$$

This identity does not give the same level of cancellation as RH anymore but still is better than the best cancellation one can get from the current unconditional estimates for Ψ *function. Also, the advantage is that the intervals* $\left(e^{\sqrt{x-(2\ell)^2\epsilon\sqrt{x}}}, e^{\sqrt{x-(2\ell-1)^2\epsilon\sqrt{x}}}\right)$ *are not as small as what we had in* (3.9). *So it possibly is more suitable for combinatorial applications.*

3.1.2 Applications to the usual and restricted partitions

A generalization of the Pentagonal Number Theorem is the second application of the cancellation result. It is an interesting question to find the second dominant term of general,“Meinardus type” integer partitions. Our result is applicable in general if the second term of Meinardus's Theorem for an arbitrary partition function is known. But the known asymptotic formulas rely heavily on analytic properties of the parts. For many types, we see a formula like

$$\lambda(n) \sim (g(n))^q\, e^{(k(n))^\theta}\left(1-\frac{1}{(h(n))^r}\right) + O(\lambda(n)^s) \tag{3.14}$$

where $0 < s < 1$ and $\theta, r, q > 0$ and $k(n)$ is a linear polynomial and $g(n), h(n)$ are rational functions. For example for the usual partition function we have

$$g(n) = \frac{\sqrt{12}}{24n-1} \quad , \quad h(n) = \frac{\pi^2}{36}(24n-1) \quad , \quad k(n) = \frac{\pi^2}{36}(24n-1) \quad , \quad s = \theta = \frac{q}{2} = r = \frac{1}{2}$$

Assuming a partition function has form (3.14), we can conclude that for a quadratic polynomial $t(n) = an^2 + bn + d$

$$\sum_{\ell : t(\ell) < n} (-1)^l \lambda(n - t(\ell)) \lesssim \lambda^{\kappa}(n)$$

where $\kappa = \max(w, s)$ and w is defined as in Theorem 3.4, and s in (3.14). As long as $\kappa < 1$, we can get a nontrivial approximation of the Pentagonal Number Theorem. We give a few specific examples.

First, we mention a weak pentagonal number theorem for certain approximations of the partition function.

Proposition 3.15. *Let*

$$\begin{aligned}
p_1(x) &= \frac{e^{\pi\sqrt{\frac{2x}{3}}}}{4x\sqrt{3}} \\
p_2(x) &= \Big(\frac{\sqrt{12}}{24x-1} - \frac{6\sqrt{12}}{\pi(24x-1)^{\frac{3}{2}}}\Big)e^{\frac{\pi}{6}\sqrt{24x-1}} \\
p_3(x) &= \Big(\frac{\sqrt{6}e^{\pi i x}}{24x-1} - \frac{12\sqrt{6}e^{\pi i x}}{\pi(24x-1)^{\frac{3}{2}}}\Big)e^{\frac{\pi}{12}\sqrt{24x-1}}
\end{aligned}$$

be the "first" term, first "two" terms, and second "two" terms of Ramanujan-Hardy-

Rademacher formula, respectively. Then

$$\sum_{G_l<x}(-1)^l p_1(x-G_l) \lesssim p(x)^{0.14} \tag{3.16}$$

$$\sum_{G_l<x}(-1)^l (1+\frac{1}{24(x-G_l)-1}) p_2(x-G_l) \lesssim p(x)^{0.21} \tag{3.17}$$

$$\sum_{G_l<x}(-1)^l \sqrt{p_1(x-G_l)} \lesssim p(x)^{0.065} \tag{3.18}$$

$$\sum_{l^2<x} p_3(x-l^2) \lesssim p(x)^{0.065} \tag{3.19}$$

Note that equation (3.19) does not have the factor $(-1)^l$, because $\sum_h \omega(h,2)$ in equation (1.8) is $\frac{(-1)^l}{\sqrt{2}}$. So it can cancels out the other $(-1)^l$ from the weak pentagonal number theorem to eliminate the cancellation. In fact, if we put $(-1)^l$, we get the following proposition.

Proposition 3.20. *For large enough x*

$$\sum_{l^2<x}(-1)^l p_3(x-l^2) \sim \frac{e^{\pi i x}\sqrt{p(x)}}{\sqrt[4]{8x}}.$$

So if $p_4(x) = p_2(x) + p_3(x)$ is the first "four" terms in the Ramanujan-Hardy-Rademacher expression for the partition function, then we get

$$\sum_{l^2<x}(-1)^l p_4(x-l^2) \sim \frac{\sqrt{p(x)}}{\sqrt[4]{8x}}. \tag{3.21}$$

We mention another set of examples. Remember from (1.9) that $p(n;\alpha,M)$ is the number of partitions with parts of the form $Mt \pm \alpha$, $1 \leqslant \alpha \leqslant M-1$, and $(\alpha, M) = 1$. Theorem 3.4 can show a weak pentagonal number expression like

$$\sum_{am^2+bm+d<x}(-1)^m p\left(x-am^2-bm-d;\alpha,M\right) = \text{"small function w.r.t. } x,\alpha,a,M\text{"}.$$

We check two cases $M = 2$ and $M = 5$ as examples.

Corollary 3.22. *Let $q_1(n)$ be the first "two" term in the expansion of $q(n)$ in equation* (**??**). *For large n*

$$\sum_{l^2 \leqslant n} (-1)^l q_1(n - l^2) \lesssim q(n)^{0.151}$$
$$\sum_{l^2 \leqslant n} (-1)^l q(n - l^2) \lesssim \sqrt[3]{q(n)}.$$

Also for $p = 5$, see [32], there exists a constant $A > 0$ such that

$$p(n; a, 5) = \frac{B\pi \csc(\frac{\pi a}{5})}{\sqrt{(60n - A)^{\frac{3}{4}}}} e^{\frac{\pi\sqrt{60n-A}}{15}} + O(\sqrt{p(n; a, 5)}).$$

Corollary 3.23. *Let $h(n)$ be the first two term in the expansion of $p(n; a, 5)$. For large n*

$$\sum_{l^2 \leqslant n} (-1)^l h(n - l^2) \lesssim p(n; a, 5)^{0.13}$$
$$\sum_{\ell^2 \leqslant n} (-1)^\ell p(n - \ell^2; a, 5) \lesssim \sqrt{p(n; a, 5)}.$$

Note that if we generalize theorem 3.4 for third or fourth or in general nth root, we might be able to prove more expressions like the Pentagonal Number Theorem. As an example, we can prove a "pentagonal number theorem" for $p_k(n)$, which is the number of partitions of n with parts that are k^{th} powers of positive integers (see (**??**)).

3.1.3 Applications to the Prouhet-Tarry-Escott Problem

Another application of our method involves the so-called Prouhet-Tarry-Escott Problem (see chapter 1). One could consider a weakening of this problem, where the left and right hand sides of (1.10) are merely required to be "close to each other". One way to naturally view this approximation is to interpret $\{a_i\}, \{b_i\}$ as events in two discrete uniform random variables A, B both of whose moments (up to a certain level) and their moment generating functions are "close"; i.e. the probability density function of these random variables becomes almost the same. Approximating moment generating functions is an important problem in the literature - see for example [33, 34]; and what we are interested in is that the probability space is a subset of $\mathbb{Q}$. This makes the problem non-trivial.

This problem can be also viewed from another perspective that is related to the Vingradov mean value theorem (see chapter 1 or check the survey paper [35] for more information).

Let us formulate the problem as follows.

Problem 3.24. *Let $0 < c = c(N, n, k) < 1$ be the smallest constant such that there exist sequences of integers*

$$1 \leqslant a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n \leqslant N \quad \text{and} \quad 1 \leqslant b_1 \leqslant b_2 \leqslant \cdots \leqslant b_n \leqslant N$$

that do not overlap, i.e. $a_i \neq b_j$, such that for all $1 \leqslant r \leqslant k$,

$$\left| \sum_{i=1}^{n} a_i^r - b_i^r \right| \leqslant N^{cr} \tag{3.25}$$

How small can we take c to be for various ranges of k and n?

As we mentioned in chapter 1, there has been little progress in solving the original PTE problem since the 19th century. For example for an ideal solution (when $k = n - 1$) the largest known solution is for $n = 12$, see [10]. To our knowledge, the best constructive solution is perhaps for the range $k = O(\log(n))$. Using a pigeonhole argument we can do much better, and give non-constructive solutions with k as large as $k \sim c\sqrt{n}$. In section 3.5 we will briefly explain this argument, which gives one of the best known non-constructive ways to solve the problem 3.24.

Even applying the pigeonhole argument to the approximate version (Problem 3.24) we cannot make k much larger; for example, we cannot prove the existence of non-decreasing sequences a_i and b_i such that

$$|\sum_i a_i^r - b_i^r| < N^{r(1-\frac{1}{\log(r)})} \quad \text{for all} \quad 1 \leqslant r \leqslant \sqrt{n}\log^2(n).$$

In other word, one cannot guarantee that the value of c in Problem 3.24 should look like $c < 1 - 1/\sqrt{\log(N)}$ when $k > \sqrt{n}\log^2(n)$. We will see that this range for c is much, much weaker than what our construction gives. This suggests that it might be possible to beat the bounds that the pigeonhole principle gives for the exact version of the problem. In section 3.5, we will give a proof to the following theorem, which states a *constructive* solution for problem 3.24 when $M(n)$ is much bigger than $\sqrt{n}$.

Theorem 3.26. *Let $L \geqslant 1$ and $m \in \mathbb{N}$ and define $M = \lfloor (2L)^{\frac{2m}{2m+1}} \rfloor$. Define for $1 \leqslant i \leqslant n := L^m$*

$$a_i = M^{2m+1} - (2i-2)^2 \quad, b_i = M^{2m+1} - (2i-1)^2 \quad, 1 \leqslant r \leqslant k \lesssim_m \frac{M^m}{\log(M)}.$$

Then

$$\sum_{1 \leqslant i \leqslant L} a_i^r - \sum_{1 \leqslant i \leqslant L} b_i^r \lesssim r^r M^{(2m+1)\frac{r}{2}+m} \lesssim r^r a_1^{\frac{r+1}{2}}. \tag{3.27}$$

So we have two sets of around n integers less than $N := M^{2m+1}$, and they are satisfying the equation (3.25) *with $\frac{1}{2} \leqslant c < 1 - \frac{1}{4m+2}$ and $k \simeq n^{1-\frac{1}{m}}$.*

For example if we put $m = 3$ in Theorem 3.26; we get the next corollary.

Corollary 3.28. *Equation* (3.25) *has a constructive solution for $\frac{1}{2} \leqslant c < \frac{13}{14}$ and $k \sim \frac{n^{6/7}}{\log(n)}$ as follows. For $1 \leqslant i \leqslant \sqrt{N}$*

$$a_i = N - (2i-2)^2 \in \mathbb{N} \quad \textit{and} \quad b_i = N - (2i-1)^2 \in \mathbb{N}.$$

Then for all $1 \leqslant r \leqslant k \sim N^{3/7}$ *we have*

$$\sum_{1 \leqslant i \leqslant n} a_i^r - \sum_{1 \leqslant i \leqslant n} b_i^r \lesssim r^r N^{\frac{r+1}{2}} \lesssim N^{\frac{r+1}{2} + r\frac{\log(r)}{\log(N)}}. \tag{3.29}$$

Remark 3.30. *There is a conjecture in [36] stating that if* $\{a_n \geqslant 0\}, \{b_n \geqslant 0\}$ *be an ideal solution of Prouhet-Tarry-Escott and* $a_1 < b_1$*, then for all* i

$$(a_i - b_i)(a_{i+1} - b_{i+1}) < 0. \tag{3.31}$$

Although our example cannot resolve the conjecture, it shows that equation (3.31) *is not true for the solutions of Problem 3.24 for any* c*.*

Remark 3.32. *Note that we can win by a constant factor - i.e. increase* $M(n)$ *by a constant, if we pick a suitable quadratic polynomial* $q(l)$ *instead of* l^2*.*

Lastly, we investigate the problem more concretely by viewing a_i, b_i as polynomials. Then this cancellation sum can be considered as an operator in $\mathbb{Z}[x]$ which cuts the degree to half.

Theorem 3.33. *Let* $M \in \mathbb{N}$*, and define* $f_r(M) := \sum_{|\ell| < 2M} (-1)^\ell (4M^2 - \ell^2)^r$*; Then,* $f_r(M)$ *is a polynomial of degree* $r - 1$ *in* M *when* r *is even, and is a polynomial of degree* r *in* M *when* r *is odd; that is, when* r *is even,*

$$f_r(M) = c_0(r) + c_1(r)M + \cdots + c_{r-1}(r)M^{r-1},$$

where $c_0(r), ..., c_{r-1}(r)$ *are integer functions of* r *only (and not of* M*). The same general form holds for* r *odd, except that the degree here is* r*, not* $r - 1$*. Furthermore, under the assumption* $r \lesssim \frac{M}{\log(M)}$ *we have that all the coefficients have size* $O(r^{r+\epsilon})$*.*

3.2 Proof of the oscillation sums

In this section we mainly prove theorems 3.4 and 3.7. First, we mention a lemma.

Lemma 3.34. *Let $z = A + iB$ be a complex number and $q(s) = as^2 + bs + c$ be a quadratic polynomial and $x \in \mathbb{R}$. Then*

$$\textit{Re}\left(\sqrt{x - q(z)}\right) = \sqrt{\frac{1}{2}\left(\sqrt{D^2 + (2aAB + bB)^2} + D\right)} \tag{3.35}$$

$$\textit{Im}\left(\sqrt{x - q(z)}\right) = \sqrt{\frac{1}{2}\left(\sqrt{D^2 + (2aAB + bB)^2} - D\right)}. \tag{3.36}$$

where

$$D := x - (A^2 - B^2)a - bA - c.$$

Proof. We only prove (3.35). We have

$$x - q(z) = x - (A^2 - B^2)a - bA - c - i(2ABa + bB).$$

It impliess that

$$\text{Re}\left(\sqrt{x - q(z)}\right) = \sqrt[4]{D^2 + (2aAB + bB)^2}\cos\left(\frac{1}{2}\tan^{-1}\left(\frac{2aAB + bB}{D}\right)\right).$$

Noting that $\cos^2(y) = \frac{1+\cos(2y)}{2}$ and $\cos(\arctan(y)) = \frac{1}{\sqrt{1+y^2}}$

$$\text{Re}\left(\sqrt{x - q(z)}\right) = \sqrt[4]{D^2 + (2aAB + bB)^2}\sqrt{\frac{1}{2} + \frac{1}{2\sqrt{1 + \frac{(2aAB+bB)^2}{D^2}}}}.$$

Straightforward computation results in equation (3.35). □

Next, we prove Theorem 3.4.

Proof. Let $q(z) := az^2 + bz + d$ and $f(z) = \sqrt{x - q(z)}$ with branch points α_1, α_2. We choose $(-\infty, \alpha_1] \cup [\alpha_2, \infty)$ as the branch cut and let G be the interior of the square with vertices

$$\pm(\sqrt{\frac{x}{a}} - \frac{2b}{a}) \pm iu\sqrt{x},$$

where $u > 0$ will be chosen later. Note that

$$\alpha_2 > \sqrt{\frac{x}{a}} - \frac{2b}{a}$$

and we have a similar condition for α_1. Without loss of generality we assume that h is holomorphic inside G.

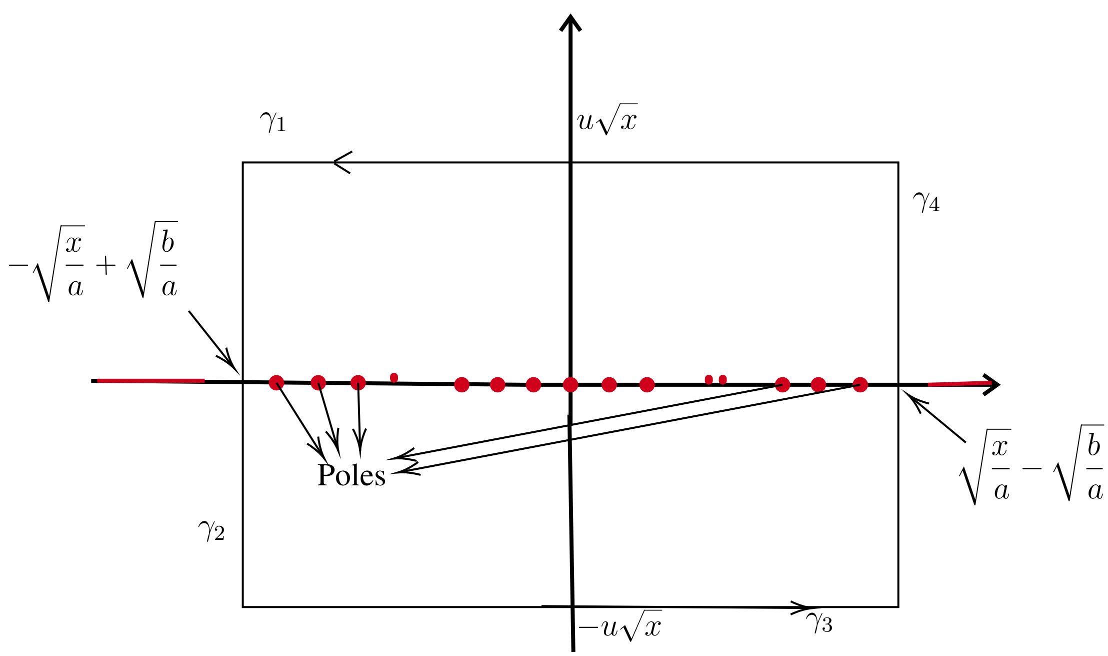

Figure 3.1: The contour γ

Let $g(z) = e^{cf(z)}$, which is analytic inside G. Define

$$H(z) = \frac{g(z)h(z)}{\sin(\pi z)}.$$

Assume that γ is the boundary of G (see figure 3.1). Using the residue theorem

$$\int_\gamma H(z)dz = 2\pi i \sum_{z_j\text{: poles}} \text{Res}(H(z))|_{z_j} = 2\pi i \sum_{q(n)<x} (-1)^n h(n) e^{c\sqrt{x-q(n)}}. \tag{3.37}$$

We wish to show that the integral in LHS has size of at most $e^{cw\sqrt{x}}$. First assume that we choose $z \in \gamma_1 \cup \gamma_3$. So $z = t \pm iu\sqrt{x}$ for $-\sqrt{\frac{x}{a}} + \frac{2b}{a} < t < \sqrt{\frac{x}{a}} - \frac{2b}{a}$. If $t = o(\sqrt{x})$, then $\sqrt{x - az^2 - bz - d} \sim \sqrt{x(1 + au^2)}$. Otherwise by lemma 3.34

$$Re(\sqrt{x - az^2 - bz - d}) \lesssim \sqrt{\frac{\sqrt{(x - at^2 + au^2x)^2 + 4a^2t^2u^2x} + x + au^2x - at^2}{2}} \tag{3.38}$$

A straightforward computation shows that the maximum of RHS of (3.38) is at $t = o(\sqrt{x})$. So

$$Re(\sqrt{x - az^2 - bx - d}) \leqslant \sqrt{x(1 + au^2)}.$$

As $c > 0$, we conclude in both cases that $e^{c\sqrt{x-az^2-bz-d}} \lesssim e^{c\sqrt{x(1+au^2)}}$. Also we have $|\sin(\pi z)| \sim \frac{1}{2}e^{\pi u\sqrt{x}}$. So we will get that for $z \in \gamma_1, \gamma_3$

$$|H(z)| \lesssim e^{c\sqrt{x(1+au^2)} - \pi u\sqrt{x}} \tag{3.39}$$

We desire to make the contribution from $z \in \gamma_1, \gamma_3$ to be approximately equal to the contribution from $z \in \gamma_2, \gamma_4$. It means that we need $c\sqrt{x(1 + au^2)} - \pi u\sqrt{x} < wc\sqrt{x}$, where w is defined in the theorem. We need to express u in terms of w. After solving this we get two cases. If $\pi^2 \neq ac^2$, then

$$\frac{-cw\pi + c\sqrt{\pi^2 - ac^2 + ac^2w^2}}{\pi^2 - ac^2} < u. \tag{3.40}$$

Otherwise, we will get

$$\frac{c(1-w^2)}{2\pi w} < u. \tag{3.41}$$

Now we compute the case $z \in \gamma_2, \gamma_4$. We have $z = \pm\sqrt{\frac{x}{a}} \mp \frac{2b}{a} + it$ and $-u\sqrt{x} < t < u\sqrt{x}$. If $t = o(\sqrt{x})$, then $\sqrt{x - q(z)} = o(\sqrt{x})$. Otherwise, using lemma 3.34

$$Re(\sqrt{x-q(z)}) \lesssim \sqrt{t\sqrt{a}\frac{\sqrt{at^2+4x}+t\sqrt{a}}{2}}.$$

Let $t = r\sqrt{x}$. We need to choose a proper α as follows.

$$\alpha = \text{argmax}_r \left(c\sqrt{r\sqrt{a}\frac{\sqrt{ar^2+4}+r\sqrt{a}}{2}} - \pi r \right) \quad , \quad 0 \leqslant r \leqslant u.$$

Also we assume that $\pm\sqrt{\frac{x}{a}} \mp \frac{2b}{a}$ is far enough from integers (otherwise we shift the legs γ_2, γ_4 slightly to avoid $Re(z)$ being near to integer). So we conclude that $|\sin(\pi r z)| > \lambda > 0$ for a fixed λ. Then we have

$$\int_{\gamma_2,\gamma_4} \frac{e^{c\sqrt{x-q(z)}}h(z)}{\sin(\pi z)} \lesssim \sqrt{x}e^{c\sqrt{x\alpha\sqrt{a}\frac{\sqrt{a\alpha^2+4}+\alpha\sqrt{a}}{2}}-\pi\alpha\sqrt{x}}h(\sqrt{x})$$

Finally in order to satisfy (3.40) and (3.41) and the fact that $u \geqslant r$, we choose

$$u = \max\left(\frac{-cw\pi + c\sqrt{\pi^2 - ac^2 + ac^2w^2}}{\pi^2 - ac^2}, \alpha\right) \quad \text{or} \quad u = \max\left(\frac{c(1-w^2)}{2\pi w}, \alpha\right).$$

where

$$w = \sqrt{\alpha\sqrt{a}\frac{\sqrt{a\alpha^2+4}+\alpha\sqrt{a}}{2}} - \frac{\pi\alpha}{c}.$$

□

In this paper, we need two versions of the Van der Corput lemma. The versions we give here are a little different than what is known in [37]. But these versions are straightforward and enough for the purpose of this paper.

Lemma 3.42. *[Simpler version] Let $F(x)$ be a second differentiable function in (a,b); also $0 < M < |F'(x)|$, and $|G(x)| < R$ for $x \in (a,b)$. Assume that $\frac{G(x)}{F'(x)}$ is a piecewise monotone function. Then*

$$\int_a^b e^{iF(x)} G(x)dx \lesssim \frac{R}{M}. \tag{3.43}$$

Lemma 3.44. *[38] Suppose that $f(x)$ is a real-valued function such that $0 < \lambda_2 \leqslant f''(x)$ for all $x \in [a,b]$, and suppose that $|f^{(3)}(x)| \leqslant \lambda_3$ and that $|f^{(4)}(x)| \leqslant \lambda_4$ throughout this interval. Put $f'(a) = \alpha$, $f'(b) = \theta$. For integers $\nu \in [\alpha - 1, \theta + 1]$ let x_ν be the root of the equation $f'(x) = \nu$. Then*

$$\sum_{a \leqslant n \leqslant b} e^{2\pi i f(n)} = e^{\frac{\pi i}{4}} \sum_{\alpha-1 \leqslant \nu \leqslant \theta+1} \frac{e^{2\pi i (f(x_\nu) - \nu x_\nu)}}{\sqrt{f''(x_\nu)}} + O\left(\log(4 + \theta - \alpha)\right) + O(\lambda_2^{-\frac{1}{2}}(\theta - \alpha + 2))$$
$$+ O\left((\lambda_3^2 \lambda_2^{-3} + \lambda_4 \lambda_2^{-2})(b-a)(\theta - \alpha + 2)\right).$$

Note that if $f''(x) < -\lambda_2 < 0$, then $e^{\frac{\pi i}{4}}$ will change $e^{-\frac{\pi i}{4}}$.

Proof of Theorem 3.7.

Let $T > 0$ and γ be the contour with vertices

$$\pm\sqrt{\eta x T} \pm iu\sqrt{x}$$

where $\eta = \frac{\pi^2}{1+\pi^2} - \epsilon$ and $0 < u$ will be determined later (see figure 3.2). Let

$$h_T(z) = \frac{e^{(\alpha+i\beta)\sqrt{x - \frac{z^2}{T}}}}{\sin(\pi z)}$$

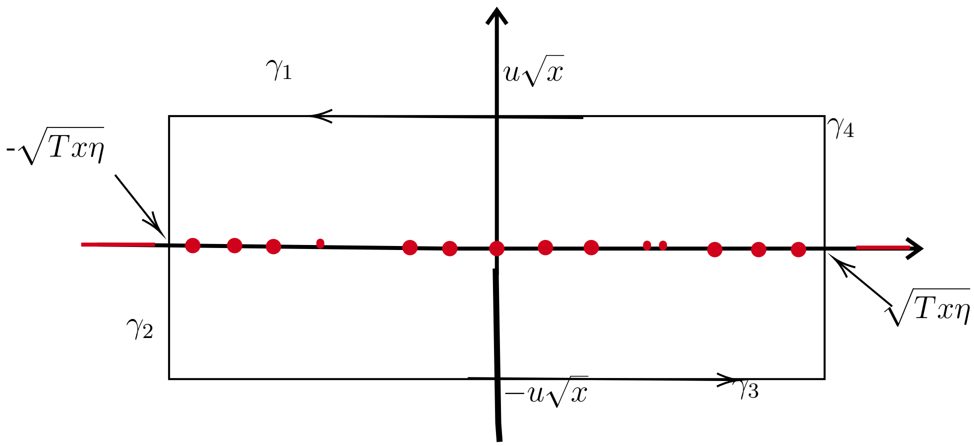

Figure 3.2: Contour γ for complex c case

We take the branch cut to be $(-\infty, -\sqrt{xT}] \cup [\sqrt{xT}, \infty)$. The Residue Theorem implies

$$\int_\gamma h_T(z)dz = 2\pi i \sum_{\ell^2 < Tx\eta} (-1)^\ell e^{(\alpha+i\beta)\sqrt{x-\frac{\ell^2}{T}}}. \tag{3.45}$$

Now we compute the case where $z \in \gamma_1, \gamma_3$. So $z = t \pm iu\sqrt{x}$ and $-\sqrt{Tx\eta} < t < \sqrt{Tx\eta}$. As $x \lesssim \sqrt{T}$ we have

$$\sqrt{x - \frac{z^2}{T}} = \sqrt{x - \frac{t^2 - u^2x \pm 2iut\sqrt{x}}{T}} \sim \sqrt{x + \frac{-t^2 \mp 2iut\sqrt{x}}{T} + O(\frac{1}{\sqrt{T}})}$$

If $t\sqrt{x} = o(\sqrt{T})$, then noting $x - \frac{t^2}{T} = x + o(\frac{1}{x})$ we conclude that

$$\sqrt{x - \frac{z^2}{T}} \sim \sqrt{x} + C(x,T) + iB(x,T), \text{where} B(x,T) = o(\frac{1}{\sqrt{T}}) \text{ and } C(x,T) = o(\frac{1}{x}). \tag{3.46}$$

Otherwise, recall that in the worst case, $t \sim \sqrt{xT}$. We use lemma 3.34 to get

$$\sqrt{x - \frac{z^2}{T}} \sim \sqrt{\frac{\sqrt{(x-\frac{t^2}{T})^2 + \frac{4u^2t^2x}{T^2}} + (x - \frac{t^2}{T})}{2}} + i\sqrt{\frac{\sqrt{(x-\frac{t^2}{T})^2 + \frac{4u^2t^2x}{T^2}} - (x - \frac{t^2}{T})}{2}} \tag{3.47}$$

Because of the range of values of t, $x - \frac{t^2}{T} \geqslant (1-\eta)x$. So

$$\sqrt{(x-\frac{t^2}{T})^2 + \frac{4u^2t^2x}{T^2}} - (x - \frac{t^2}{T}) = \frac{\frac{4u^2t^2x}{T^2}}{\sqrt{(x-\frac{t^2}{T})^2 + \frac{4u^2t^2x}{T^2}} + (x-\frac{t^2}{T})} \leqslant \frac{2u^2t^2}{(1-\eta)T^2}.$$

Therefore in all cases for t

$$Re(\sqrt{x - \frac{z^2}{T}}) \leqslant \sqrt{x} \quad \text{and} \quad Im(\sqrt{x-\frac{z^2}{T}}) \leqslant \sqrt{\frac{xu^2\eta}{(1-\eta)T}}.$$

Noting that $|\beta| \leqslant \sqrt{T}$

$$Re\left((\alpha+i\beta)\sqrt{x-\frac{z^2}{T}}\right) \leqslant \alpha\sqrt{x} + \sqrt{\frac{xu^2\eta}{(1-\eta)}}\frac{|\beta|}{\sqrt{T}} \leqslant \sqrt{x}\left(\alpha + u\sqrt{\frac{\eta}{(1-\eta)}}\right)$$

So for $z \in \gamma_1, \gamma_3$ we have

$$\frac{e^{(\alpha+i\beta)\sqrt{x-\frac{z^2}{T}}}}{\sin(\pi z)} \lesssim e^{\left(\alpha+u\sqrt{\frac{\eta}{(1-\eta)}}-\pi u\right)\sqrt{x}}. \tag{3.48}$$

We will later choose proper w, η, u such that

$$\alpha + u\sqrt{\frac{\eta}{(1-\eta)}} - \pi u + \frac{\log(T)}{2\sqrt{x}} < w. \tag{3.49}$$

Next we assume that $z \in \gamma_2, \gamma_4$. We have $z = \pm\sqrt{\eta x T} + it$ and $-u\sqrt{x} \leqslant t \leqslant u\sqrt{x}$. As $t \lesssim \sqrt{x} \lesssim \sqrt[4]{T}$, then $\frac{t\sqrt{x}}{\sqrt{T}} \lesssim 1$ and $\frac{t^2}{T} \lesssim \frac{x}{T} = o(1)$. We use lemma 3.34 to conclude that

$$\begin{aligned}\sqrt{x-\frac{z^2}{T}} &\sim \sqrt[4]{\frac{x}{T}}\left(\sqrt{\frac{\sqrt{4t^2\eta + xT(1-\eta)^2} + (1-\eta)\sqrt{xT}}{2}}\right.\\ &\qquad \left. + i\sqrt{\frac{\sqrt{4t^2\eta + xT(1-\eta)^2} - (1-\eta)\sqrt{xT}}{2}}\right)\\ &\sim \sqrt{(1-\eta)x} + it\sqrt{\frac{\eta}{T(1-\eta)}}.\end{aligned}$$

This together with the fact that $|\beta| < \sqrt{T}$ imply that

$$Re\left((\alpha+i\beta)\sqrt{x-\frac{z^2}{T}}\right) \leqslant \alpha\sqrt{(1-\eta)x} + \frac{|\beta|}{\sqrt{T}}\sqrt{\frac{\eta t^2}{1-\eta}} \leqslant \alpha\sqrt{(1-\eta)x} + |t|\sqrt{\frac{\eta}{1-\eta}}$$

Again we assume that $\sqrt{xT\eta}$ is far away from the integers; so as $\sin(\pi z) > \lambda > 0$ for some fixed λ, therefore for $z \in \gamma_2, \gamma_4$

$$\frac{e^{(\alpha+i\beta)\sqrt{x-\frac{z^2}{T}}}}{\sin(\pi z)} \lesssim e^{\alpha\sqrt{(1-\eta)x}+|t|\left(\sqrt{\frac{\eta}{1-\eta}}-\pi\right)}.$$

As $\sqrt{\frac{\eta}{1-\eta}} < \pi$, the maximum of the following function occurs at $y = 0$:

$$G(y) = \alpha\sqrt{(1-\eta)} + y\sqrt{\frac{\eta}{1-\eta}} - y\pi.$$

We conclude that

$$\begin{aligned} 2\pi i \sum_{\ell^2 < Tx\eta} (-1)^\ell e^{(\alpha+i\beta)\sqrt{x-\frac{\ell^2}{T}}} &= \int_\gamma h_T(z)dz \\ &\lesssim \sqrt{Tx}e^{(\alpha+u\sqrt{\frac{\eta}{(1-\eta)}}-\pi u)\sqrt{x}} + \sqrt{x}e^{\alpha\sqrt{(1-\eta)}\sqrt{x}}. \end{aligned} \tag{3.50}$$

A straightforward calculation shows that

$$\left|\sum_{xT\eta\leqslant\ell^2<Tx} (-1)^l e^{(\alpha+i\beta)\sqrt{x-\frac{\ell^2}{T}}}\right| \lesssim \sum_{xT\eta\leqslant\ell^2<Tx} e^{\alpha\sqrt{x-\frac{\ell^2}{T}}} \lesssim \sqrt{xT}e^{\alpha\sqrt{x(1-\eta)}}.$$

For a sharper bound, we use lemma 3.44 to control the tail. Without loss of generality assume that $\beta < 0$. We prove that for $|\beta| \lesssim \sqrt{T}$,

$$\left|\sum_{xT\eta\leqslant\ell^2<Tx-T} (-1)^\ell e^{i\beta\sqrt{x-\frac{\ell^2}{T}}}\right| \lesssim \frac{\sqrt{T}x^{3/2}}{\sqrt{|\beta|+1}} + \frac{\sqrt{T}x^{11}}{|\beta|+1} + \log x. \tag{3.51}$$

It is trivial to get the bound for $|\beta| < 1$, so we assume otherwise. Let $f(\ell) := \frac{1}{2}\ell +$

$\frac{\beta}{2\pi}\left(x-\frac{\ell^2}{T}\right)^{1/2}$. Then

$$f'(\ell)=\frac{1}{2}-\frac{\beta\ell}{2\pi T}\left(x-\frac{\ell^2}{T}\right)^{-1/2} \qquad f''(\ell)=-\frac{\beta x}{2\pi T}(x-\frac{\ell^2}{T})^{-3/2}$$
$$f^{(3)}(\ell)=-\frac{3\beta x\ell}{2\pi T^2}(x-\frac{\ell^2}{T})^{-5/2} \qquad f^{(4)}(\ell)=-\frac{3\beta x}{2\pi T^2}(x+\frac{4\ell^2}{T})(x-\frac{\ell^2}{T})^{-7/2}.$$

First we evaluate f' at the endpoints. Without loss of generality we consider $\sqrt{xT\eta}<\ell<\sqrt{xT-T}$. Assuming that $\eta<0.9$, we have

$$f'(\sqrt{xT\eta})=\frac{1}{2}-\frac{\beta\sqrt{x\eta}}{2\pi\sqrt{T}}(x-\eta x)^{-1/2}\in[0,1] \tag{3.52}$$
$$f'(\sqrt{xT-T})=\frac{1}{2}-\frac{\beta\sqrt{xT-T}}{2\pi T}<\sqrt{x}. \tag{3.53}$$

As $f'',f^{(3)},f^{(4)},f^{(5)}$ are positive, we can find $\lambda_2,\lambda_3,\lambda_4$ easily at the endpoints.

$$\lambda_2=\inf_{\sqrt{Tx\eta}<\ell<\sqrt{xT-T}} f''(\ell)=f''(\sqrt{xT\eta})=\frac{|\beta|}{2\pi T\sqrt{x}(1-\eta)^{3/2}}\gg\frac{|\beta|}{T\sqrt{x}}$$
$$\lambda_3=\sup_{\sqrt{Tx\eta}<\ell<\sqrt{xT-T}} f^{(3)}(\ell)=f^{(3)}(\sqrt{xT-T})=\frac{-3\beta x}{2\pi T^2}(xT-T)^{1/2}\lesssim\frac{|\beta|x^{3/2}}{T^{3/2}}$$
$$\lambda_4=\sup_{\sqrt{Tx\eta}<\ell<\sqrt{xT-T}} f^{(4)}(\ell)=f^{(4)}(\sqrt{xT-T})=\frac{-3\beta x}{2\pi T^2}(5x-4)\lesssim\frac{|\beta|x^2}{T^2}. \tag{3.54}$$

It implies that

$$\lambda_3^2\lambda_2^{-3}+\lambda_4\lambda_2^{-2}\lesssim\frac{x^{9/2}}{|\beta|+1}.$$

Noting that $f'(\sqrt{xT-T})-f'(\sqrt{xT\eta})=O(\sqrt{x})$ and applying lemma 3.44 implies the equation (3.51). For $xT\eta<t^2<xT-T$ we define

$$S(t):=\sum_{xT\eta\leqslant\ell^2<t^2}(-1)^\ell e^{i\beta\sqrt{x-\frac{\ell^2}{T}}}.$$

Similar to what we just did, we know that $|S(t)|\lesssim x^5t/\sqrt{|\beta|}\lesssim\sqrt{x^{11}T/|\beta|}$ for $xT\eta<$

$t^2 < xT - T$. Using Abel's summation formula we get

$$\left|\sum_{xT\eta \leqslant \ell^2 < Tx-T} (-1)^\ell e^{(\alpha+i\beta)\sqrt{x-\frac{\ell^2}{T}}}\right| \\ \lesssim |S(\sqrt{xT-T})|e^{\alpha} + |S(\sqrt{\eta xT})|e^{\alpha\sqrt{(1-\eta)x}} + \left|\int_{\sqrt{\eta xT}}^{\sqrt{xT-T}} S(t)v(t)dt\right| \tag{3.55}$$

where $v(t) := \frac{d}{dt}\exp\left(\alpha\sqrt{x-\frac{t^2}{T}}\right)$. We bound the integral in the RHS.

$$\int_{\sqrt{\eta xT}}^{\sqrt{xT-T}} S(t)v(t)dt = -\frac{\alpha}{T}\int_{\sqrt{\eta xT}}^{\sqrt{xT-T}} \frac{tS(t)e^{\alpha\sqrt{x-\frac{t^2}{T}}}}{\sqrt{x-\frac{t^2}{T}}}dt.$$

Straightforward computation gives that

$$G(t) := \frac{tS(t)e^{\alpha\sqrt{x-\frac{t^2}{T}}}}{\sqrt{x-\frac{t^2}{T}}} \lesssim \left(\frac{\sqrt{Tx^{3/2}}}{\sqrt{|\beta|+1}} + \log x + \frac{\sqrt{Tx^{11}}}{|\beta|+1}\right)e^{\alpha\sqrt{(1-\eta)x}}\sqrt{xT}$$

It implies that

$$\left|\int_{\sqrt{\eta xT}}^{\sqrt{xT-T}} S(t)v(t)dt\right| \lesssim \frac{\sqrt{xT}}{T}\left(\frac{\sqrt{Tx^{3/2}}}{\sqrt{|\beta|+1}} + \log x + \frac{\sqrt{Tx^{11}}}{|\beta|+1}\right)e^{\alpha\sqrt{(1-\eta)x}}\sqrt{xT} \\ \lesssim \left(\frac{\sqrt{Tx^{5/2}}}{\sqrt{|\beta|+1}} + \sqrt{x}\log x + \frac{\sqrt{Tx^{12}}}{|\beta|+1}\right)e^{\alpha\sqrt{(1-\eta)x}}.$$

This, (3.51), and (3.55) give

$$\left|\sum_{xT\eta \leqslant \ell^2 < Tx-T} (-1)^\ell e^{(\alpha+i\beta)\sqrt{x-\frac{\ell^2}{T}}}\right| \lesssim \left(\frac{\sqrt{Tx^{5/2}}}{\sqrt{|\beta|+1}} + \sqrt{x}\log x + \frac{\sqrt{Tx^{12}}}{|\beta|+1}\right)e^{\alpha\sqrt{(1-\eta)x}}.$$

Considering the range of β in our application, we conclude that

$$\left|\sum_{xT\eta\leqslant\ell^2<Tx}(-1)^\ell e^{(\alpha+i\beta)\sqrt{x-\frac{\ell^2}{T}}}\right|\leqslant\left|\sum_{xT\eta\leqslant\ell^2<Tx-T}(-1)^\ell e^{(\alpha+i\beta)\sqrt{x-\frac{\ell^2}{T}}}\right|+\left|\sum_{xT-T\leqslant\ell^2<Tx}(-1)^\ell e^{(\alpha+i\beta)\sqrt{x-\frac{\ell^2}{T}}}\right|$$
$$\lesssim\sqrt{\frac{x^3T}{|\beta|+1}}e^{\alpha\sqrt{(1-\eta)x}}+\sqrt{Tx}. \tag{3.56}$$

We used a trivial bound for the second sum. We want to have

$$\left|\sum_{\ell^2<Tx}(-1)^\ell e^{(\alpha+i\beta)\sqrt{x-\frac{\ell^2}{T}}}\right|\lesssim\sqrt{\frac{T}{|\beta|+1}}e^{w\sqrt{x}}.$$

Adding (3.50) and (3.56) we need to have

$$\begin{cases}\alpha\sqrt{1-\eta}\leqslant w\\ \alpha+u\sqrt{\frac{\eta}{(1-\eta)}}-\pi u+\frac{1}{2\sqrt{x}}\log(|\beta|+1)\leqslant w.\end{cases} \tag{3.57}$$

Remember that $\eta=\frac{\pi^2}{1+\pi^2}-\epsilon$. Comparing with β, if we choose u large enough then the left hand side of the second condition in (3.57) becomes negative. So

$$w=\alpha\sqrt{\frac{1}{1+\pi^2}}+\epsilon$$

from the first condition. This completes the proof. □

3.3 Proof related to prime distribution

Inspired by the proof of the Prime Number Theorem (PNT) we compute the following sum in two ways.

$$\frac{1}{2\pi i}\sum_{\ell^2<Tx}(-1)^\ell\int_{1+\epsilon-i\sqrt{T}}^{1+\epsilon+i\sqrt{T}}\frac{\zeta'(s)}{\zeta(s)}\frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}ds. \tag{3.58}$$

In this section, we assume that $T < e^{\frac{4}{3}\sqrt{x}}$.

Lemma 3.59. *For large enough x*

$$\sum_{\ell^2 < Tx} (-1)^\ell \int_{1+\epsilon - i\sqrt{T}}^{1+\epsilon+i\sqrt{T}} \frac{\zeta'(s)}{\zeta(s)} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s} ds \lesssim \sqrt{T}x e^{(\frac{1}{\sqrt{1+\pi^2}}+\epsilon)\sqrt{x}}. \tag{3.60}$$

Proof. We consider the contour γ in figure 3.3, where $\epsilon > 0$ is a very small real number and U is a very large real number far enough from any negative even integer $-2m$. Using the Residue Theorem

$$\sum_{\ell^2 < Tx} (-1)^\ell \int_\gamma \frac{\zeta'(s)}{\zeta(s)} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s} ds$$
$$= 2\pi i \sum_{\ell^2 < Tx} (-1)^\ell \Bigg(\Big(\lim_{s\to 0} e^{s\sqrt{x-\frac{\ell^2}{T}}} \big(\frac{\zeta'(s)}{\zeta(s)}\big) \Big) + e^{\sqrt{x-\frac{\ell^2}{T}}}$$
$$+ \sum_{|Im(\rho_m)| < \sqrt{T}} \frac{e^{\rho_m \sqrt{x-\frac{\ell^2}{T}}}}{\rho_m} - \sum_{1 \leqslant m \leqslant U/2} \frac{e^{-2m\sqrt{x-\frac{\ell^2}{T}}}}{2m} \Bigg), \tag{3.61}$$

where ρ_m is the m^{th} non trivial zeroes of the Rieman zeta function.

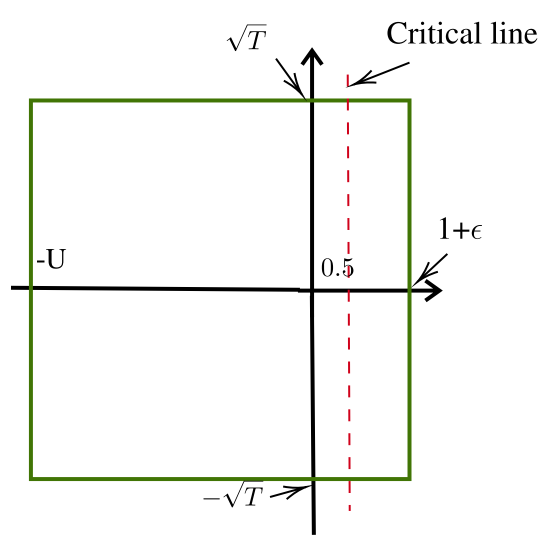

Figure 3.3: The contour γ

An easy computation shows that the first and fourth terms in the RHS sum have contribution at most $\sqrt{Tx}$. Using Theorem 3.4 (by tending $c \to 0$) the second term is bounded from above by $\sqrt{Tx}e^{\epsilon\sqrt{x}}$. We can use Theorem 3.7 to show that

$$\sum_{\ell^2<Tx}(-1)^\ell e^{\rho_m\sqrt{x-\frac{\ell^2}{T}}} \lesssim \sqrt{\frac{Tx}{|\text{Im}(\rho_m)|+1}}e^{(\frac{1}{\sqrt{1+\pi^2}}+\epsilon)\sqrt{x}}.$$

Finally, using the fact that

$$\sum_{|\text{Im}(\rho_m)|<\sqrt{T}}\frac{1}{\text{Im}(\rho_m)^{\frac{3}{2}}}$$

converges, we can conclude that the third term of RHS of (3.61) has contribution at most $\sqrt{Tx}e^{(\frac{1}{\sqrt{1+\pi^2}}+\epsilon)\sqrt{x}}$. So we have

$$\sum_{\ell^2<Tx}(-1)^\ell\int_\gamma\frac{\zeta'(s)}{\zeta(s)}\frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}ds \lesssim \sqrt{Tx}e^{(\frac{1}{\sqrt{1+\pi^2}}+\epsilon)\sqrt{x}}. \tag{3.62}$$

As $\frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}$ tends to zero for Re$(s) \to -\infty$, we can pick U large enough to have

$$\begin{aligned}\sum_{\ell^2<Tx}(-1)^\ell\int_\gamma\frac{\zeta'(s)}{\zeta(s)}\frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}ds \simeq &\sum_{\ell^2<Tx}(-1)^\ell\int_{1+\epsilon-i\sqrt{T}}^{1+\epsilon+i\sqrt{T}}\frac{\zeta'(s)}{\zeta(s)}\frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}ds\\ &+\sum_{\ell^2<Tx}(-1)^\ell\int_{-U-i\sqrt{T}}^{1+\epsilon-i\sqrt{T}}\frac{\zeta'(s)}{\zeta(s)}\frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}ds\\ &+\sum_{\ell^2<Tx}(-1)^\ell\int_{1+\epsilon+i\sqrt{T}}^{-U+i\sqrt{T}}\frac{\zeta'(s)}{\zeta(s)}\frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}ds.\end{aligned} \tag{3.63}$$

The second integral in the RHS is

$$\begin{aligned}&\left|\sum_{\ell^2<Tx}(-1)^\ell\int_{-U-i\sqrt{T}}^{1+\epsilon-i\sqrt{T}}\frac{\zeta'(s)}{\zeta(s)}\frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}ds\right|\\ &\quad=\left|\sum_{\ell^2<Tx}(-1)^\ell\int_{-U-i\sqrt{T}}^{-i\sqrt{T}}\frac{\zeta'(s)}{\zeta(s)}\frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}ds+\int_{-i\sqrt{T}}^{1+\epsilon-i\sqrt{T}}\frac{\zeta'(s)}{\zeta(s)}\frac{\sum_{\ell^2<Tx}(-1)^\ell e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s}ds\right|\\ &\quad\lesssim\sqrt{Tx}\int_{-\infty}^0\left|\frac{\zeta'(\sigma-i\sqrt{T})}{\zeta(\sigma-i\sqrt{T})}\right|\frac{e^{\sigma\sqrt{x}}}{\sqrt{T}}d\sigma+\frac{1}{\sqrt{T}}\int_0^{1+\epsilon}\left|\frac{\zeta'(s)}{\zeta(s)}\right|\sqrt[4]{Tx^2}e^{\sigma(\frac{1}{\sqrt{1+\pi^2}}+\epsilon)\sqrt{x}}d\sigma.\end{aligned}$$

Note that in the last inequality we used Theorem 3.7. We can use the fact that $\frac{\zeta'}{\zeta}(\sigma + it) = \sum_\rho \frac{1}{\sigma+it-\rho} + O(\log(t))$ to choose a proper T such that $\frac{\zeta'}{\zeta}(\sigma \pm i\sqrt{T}) \lesssim \log^2(T)$ for $-\infty \leqslant \sigma < 1+\epsilon$. So we have

$$\begin{aligned}
&\sum_{\ell^2<Tx}(-1)^\ell \int_{-U-i\sqrt{T}}^{1+\epsilon-i\sqrt{T}} \frac{\zeta'(s)}{\zeta(s)} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s} ds \\
&\quad \lesssim \sqrt{x}\log^2(T) + \frac{\sqrt{x}}{\sqrt[4]{T}} \int_0^{1+\epsilon} \left| \frac{\zeta'(\sigma - i\sqrt{T})}{\zeta(\sigma - i\sqrt{T})} \right| e^{\sigma(\frac{1}{\sqrt{1+\pi^2}}+\epsilon)\sqrt{x}} d\sigma \\
&\quad \lesssim \frac{\log^2(T)}{\sqrt[4]{T}} e^{(1+\epsilon)\sqrt{\frac{x}{1+\pi^2}}}.
\end{aligned}$$

The third integral can be similarly bounded. This, (3.62), and (3.63) give the result. □

Proof of Theorem 3.8. We compute (3.58) another way. We have

$$\begin{aligned}
\sum_{\ell^2<Tx}(-1)^\ell \int_{1+\epsilon-i\sqrt{T}}^{1+\epsilon+i\sqrt{T}} \frac{\zeta'(s)}{\zeta(s)} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s} ds &= \sum_{\ell^2<Tx}(-1)^\ell \int_{1+\epsilon-i\sqrt{T}}^{1+\epsilon+i\sqrt{T}} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s} ds \\
= \sum_{\ell^2<Tx}(-1)^\ell \left(\sum_{1\leqslant n \leqslant e^{\sqrt{x-\frac{\ell^2}{T}}}} + \sum_{e^{\sqrt{x-\frac{\ell^2}{T}}} \leqslant n} \int_{1+\epsilon-i\sqrt{T}}^{1+\epsilon+i\sqrt{T}} \frac{\Lambda(n)}{n^s} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s} ds \right) &:= A_1 + A_2
\end{aligned} \tag{3.64}$$

First we compute A_1. Again, we use the contour γ in figure 3.3 to compute the integral. Knowing $\left|\frac{e^{\sqrt{x-\frac{l^2}{T}}}}{n}\right| > 1$, we conclude that the integrand is tending to zero as $\text{Re}(s) \to -\infty$.

Considering sufficiently large U and Using the Residue Theorem give

$$\begin{aligned}
A_1 \simeq & \sum_{\ell^2<Tx}(-1)^\ell \sum_{1\leqslant n\leqslant e^{\sqrt{x-\frac{\ell^2}{T}}}} \Lambda(n)\int_\gamma \left(\frac{e^{\sqrt{x-\frac{\ell^2}{T}}}}{n}\right)^s \frac{ds}{s} \\
& + \sum_{\ell^2<Tx}(-1)^\ell \sum_{1\leqslant n\leqslant e^{\sqrt{x-\frac{\ell^2}{T}}}} \Lambda(n)\int_{-U+i\sqrt{T}}^{1+\epsilon+i\sqrt{T}} \left(\frac{e^{\sqrt{x-\frac{\ell^2}{T}}}}{n}\right)^s \frac{ds}{s} \\
& - \sum_{\ell^2<Tx}(-1)^\ell \sum_{1\leqslant n\leqslant e^{\sqrt{x-\frac{\ell^2}{T}}}} \Lambda(n)\int_{-U-i\sqrt{T}}^{1+\epsilon-i\sqrt{T}} \left(\frac{e^{\sqrt{x-\frac{\ell^2}{T}}}}{n}\right)^s \frac{ds}{s} \\
=& 2\pi i \sum_{\ell^2<Tx}(-1)^\ell \Psi\left(e^{\sqrt{x-\frac{l^2}{T}}}\right) + \sum_{\ell^2<Tx}(-1)^\ell \sum_{n\leqslant e^{\sqrt{x-\frac{\ell^2}{T}}}} \Lambda(n)\int_{-U+i\sqrt{T}}^{1+\epsilon+i\sqrt{T}} \left(\frac{e^{\sqrt{x-\frac{\ell^2}{T}}}}{n}\right)^s \frac{ds}{s} \\
& - \sum_{\ell^2<Tx}(-1)^\ell \sum_{n\leqslant e^{\sqrt{x-\frac{\ell^2}{T}}}} \Lambda(n)\int_{-U-i\sqrt{T}}^{1+\epsilon-i\sqrt{T}} \left(\frac{e^{\sqrt{x-\frac{\ell^2}{T}}}}{n}\right)^s \frac{ds}{s}.
\end{aligned} \tag{3.65}$$

We bound the integrals in RHS. Define

$$y_n(V) := \Lambda(n)\int_{-U\pm iV}^{1+\epsilon\pm iV} \sum_{\ell^2<Tx-T\log^2(n)} (-1)^\ell \left(\frac{e^{\sqrt{x-\frac{\ell^2}{T}}}}{n}\right)^s \frac{ds}{s}$$

Inspired by a mean value method by Montgomery in [39] there exists $\sqrt{T} < V < 2\sqrt{T}$

such that

$$
\begin{aligned}
\sum_{n\leqslant e^{\sqrt{x}}}|y_n(V)| &= \sum_{n\leqslant e^{\sqrt{x}}}\left(\overline{y_n(V)}y_n(V)\right)^{1/2}\\
&= \sum_{n\leqslant e^{\sqrt{x}}}\left(\Lambda(n)^2\left|\int_{-U\pm iV}^{1+\epsilon\pm iV}\sum_{\ell^2<Tx-T\log^2(n)}(-1)^\ell\left(\frac{e^{\sqrt{x-\frac{\ell^2}{T}}}}{n}\right)^s\frac{ds}{s}\right|^2\right)^{1/2}\\
&\lesssim \sum_{n\leqslant e^{\sqrt{x}}}\Lambda(n)\left(\frac{1}{\sqrt{T}}\int_{\sqrt{T}}^{2\sqrt{T}}\int_{-U}^{1+\epsilon}\left|\sum_{\ell^2<Tx-T\log^2(n)}(-1)^\ell\left(\frac{e^{\sqrt{x-\frac{\ell^2}{T}}}}{n}\right)^{\sigma+it}\right|^2\frac{d\sigma}{\sigma^2+t^2}dt\right)^{1/2}\\
&\lesssim \sum_{n\leqslant e^{\sqrt{x}}}\frac{\Lambda(n)}{\sqrt[4]{T}}\left(\int_{-U}^{1+\epsilon}\sum_{\ell_1^2<\ell_2^2<T(x-\log^2(n))}\frac{e^{\sigma(\sqrt{x-\frac{\ell_1^2}{T}}+\sqrt{x-\frac{\ell_2^2}{T}})}}{n^{2\sigma}}\left|\int_{\sqrt{T}}^{2\sqrt{T}}e^{it(\sqrt{x-\frac{\ell_1^2}{T}}-\sqrt{x-\frac{\ell_2^2}{T}})}\frac{dt}{\sigma^2+t^2}\right|d\sigma\right)^{\frac{1}{2}}\\
&\quad+\sum_{n\leqslant e^{\sqrt{x}}}\frac{\Lambda(n)}{\sqrt[4]{T}}\left(\sum_{\ell^2<T(x-\log^2(n))}\int_{-U}^{1+\epsilon}\frac{e^{2\sigma\sqrt{x-\frac{\ell^2}{T}}}}{n^{2\sigma}}\int_{\sqrt{T}}^{2\sqrt{T}}\frac{dt}{\sigma^2+t^2}d\sigma\right)^{\frac{1}{2}}. \qquad (3.66)
\end{aligned}
$$

We use Lemma 3.42 for $G(t) := \frac{1}{\sigma^2+t^2}$ and $F(t) := t(\sqrt{x-\frac{\ell_1^2}{T}}-\sqrt{x-\frac{\ell_2^2}{T}})$ (i.e. $F'(t) \geqslant \frac{\ell_2^2-\ell_1^2}{2T\sqrt{x}}$) for the off-diagonal terms in the last expression of RHS in (3.66). Note that we could get the same result without using the lemma, but this way is more straightforward. Then

$$
\begin{aligned}
\sum_{n\leqslant e^{\sqrt{x}}}|y_n(V)| &\lesssim \frac{\sqrt[4]{x}}{\sqrt[4]{T}}\sum_{n\leqslant e^{\sqrt{x}}}\Lambda(n)\left(\sum_{\ell_1^2<\ell_2^2<T(x-\log^2(n))}\int_{-U}^{1+\epsilon}\frac{e^{\sigma(\sqrt{x-\frac{\ell_1^2}{T}}+\sqrt{x-\frac{\ell_2^2}{T}})}}{n^{2\sigma}(\ell_2^2-\ell_1^2)}d\sigma\right)^{\frac{1}{2}}\\
&\quad+\frac{1}{\sqrt{T}}\sum_{n\leqslant e^{\sqrt{x}}}\Lambda(n)\left(\sum_{\ell^2<T(x-\log^2(n))}\frac{e^{2(1+\epsilon)\sqrt{x-\frac{\ell^2}{T}}}}{n^{2(1+\epsilon)}}\right)^{\frac{1}{2}}\\
&\lesssim \left(\sum_{m<\sqrt{Tx}}\frac{\tau(m)}{m}\right)^{\frac{1}{2}}\frac{\sqrt[4]{x}e^{(1+\epsilon)\sqrt{x}}}{\sqrt[4]{T}}\sum_{n\leqslant e^{\sqrt{x}}}\frac{\Lambda(n)}{n^{1+\epsilon}}+\frac{\sqrt[4]{x}e^{(1+\epsilon)\sqrt{x}}}{\sqrt[4]{T}}\sum_{n\leqslant e^{\sqrt{x}}}\frac{\Lambda(n)}{n^{1+\epsilon}}
\end{aligned}
$$

where $\tau(m)$ is the number of divisors of m. So there exists $\sqrt{T} < V < 2\sqrt{T}$ we have

$$\sum_{n\leqslant e^{\sqrt{x}}} \Lambda(n) \int_{-U\pm iV}^{1+\epsilon\pm iV} \sum_{\ell^2<T(x-\log^2(n))} (-1)^\ell \left(\frac{e^{\sqrt{x-\frac{\ell^2}{T}}}}{n}\right)^s \frac{ds}{s} \lessapprox \sum_{n\leqslant e^{\sqrt{x}}} |y_n(V)| \lessapprox \frac{x^{\frac{1}{4}} e^{(1+\epsilon)\sqrt{x}} \log T}{\sqrt[4]{T}}.$$

This and (3.65) imply that

$$A_1 = 2\pi i \sum_{l^2<Tx} (-1)^l \Psi\left(e^{\sqrt{x-\frac{l^2}{T}}}\right) + O\left(\frac{x^{\frac{1}{4}} e^{(1+\epsilon)\sqrt{x}}}{\sqrt[4]{T}}\right). \tag{3.67}$$

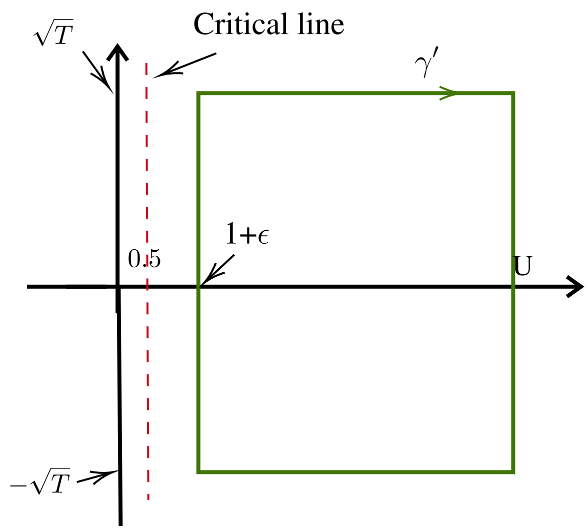

Figure 3.4: The contour γ'

Next we compute A_2. We consider contour γ' in Figure 3.4. As $\frac{e^{s\sqrt{x-\frac{l^2}{T}}}}{sn^s}$ does not have poles inside γ', choosing large enough U and using the Cauchy's integral theorem give

$$A_2 \simeq \sum_{\ell^2<Tx} (-1)^l \left(\sum_{e^{\sqrt{x-\frac{\ell^2}{T}}}\leqslant n} \int_{1+\epsilon+i\sqrt{T}}^{U+i\sqrt{T}} \frac{\Lambda(n)}{n^s} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s} ds - \sum_{e^{\sqrt{x-\frac{\ell^2}{T}}}\leqslant n} \int_{1+\epsilon-i\sqrt{T}}^{U-i\sqrt{T}} \frac{\Lambda(n)}{n^s} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{s} ds \right).$$

Similar to y_n, we define z_n as follows:

$$z_n(V) := \Lambda(n) \sum_{T(x-\log^2(n))<\ell^2<Tx} (-1)^\ell \int_{1+\epsilon\pm iV}^{U\pm iV} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{sn^s} ds$$

In this case, we will have

$$\begin{aligned}\sum_n |z_n(V)| = &\sum_{n<e^{\sqrt{x}}} \Lambda(n) \left| \sum_{T(x-\log^2(n))<\ell^2<Tx} (-1)^\ell \int_{1+\epsilon\pm iV}^{U\pm iV} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{sn^s} ds \right| \\ &+ \sum_{e^{\sqrt{x}}<n} \Lambda(n) \left| \sum_{\ell^2<Tx} (-1)^\ell \int_{1+\epsilon\pm iV}^{U\pm iV} \frac{e^{s\sqrt{x-\frac{\ell^2}{T}}}}{sn^s} ds \right|. \end{aligned} \tag{3.68}$$

As these cases are similar, we only compute the bound for the case $e^{\sqrt{x}} < n$. There exists $\sqrt{T} < V < 2\sqrt{T}$ such that

$$\begin{aligned}\sum_{e^{\sqrt{x}}<n} |z_n(V)| &\lesssim \frac{1}{\sqrt[4]{T}} \sum_{e^{\sqrt{x}}<n} \Lambda(n) \left(\int_{\sqrt{T}}^{2\sqrt{T}} \int_{1+\epsilon}^{U} \left| \sum_{\ell^2<Tx} (-1)^\ell e^{s\sqrt{x-\frac{\ell^2}{T}}} \right|^2 \frac{d\sigma}{n^{2\sigma}(\sigma^2+t^2)} dt \right)^{\frac{1}{2}} \\ &\lesssim \frac{1}{\sqrt[4]{T}} \sum_{e^{\sqrt{x}}<n} \Lambda(n) \left(\int_{1+\epsilon}^{U} \frac{1}{n^{2\sigma}} \sum_{\ell_1^2<\ell_2^2<Tx} e^{\sigma(\sqrt{x-\frac{\ell_1^2}{T}}+\sqrt{x-\frac{\ell_2^2}{T}})} \int_{\sqrt{T}}^{2\sqrt{T}} \frac{e^{it(\sqrt{x-\frac{\ell_1^2}{T}}-\sqrt{x-\frac{\ell_2^2}{T}})}}{(\sigma^2+t^2)} dt d\sigma \right)^{\frac{1}{2}} \\ &+ \frac{1}{\sqrt[4]{T}} \sum_{e^{\sqrt{x}}<n} \Lambda(n) \left(\int_{\sqrt{T}}^{2\sqrt{T}} \int_{1+\epsilon}^{U} \left| \sum_{\ell^2<Tx} (-1)^\ell e^{s\sqrt{x-\frac{\ell^2}{T}}} \right|^2 \frac{d\sigma}{n^{2\sigma}(\sigma^2+t^2)} dt \right)^{\frac{1}{2}} \\ &\lesssim \frac{1}{\sqrt[4]{T}} \sum_{e^{\sqrt{x}}<n} \Lambda(n) \left(\int_{1+\epsilon}^{U} \frac{1}{n^{2\sigma}} \sum_{\ell_1^2<\ell_2^2<Tx} e^{\sigma(\sqrt{x-\frac{\ell_1^2}{T}}+\sqrt{x-\frac{\ell_2^2}{T}})} \int_{\sqrt{T}}^{2\sqrt{T}} \frac{e^{it(\sqrt{x-\frac{\ell_1^2}{T}}-\sqrt{x-\frac{\ell_2^2}{T}})}}{(\sigma^2+t^2)} dt d\sigma \right)^{\frac{1}{2}} \\ &\quad + \frac{1}{\sqrt[4]{T}} \sum_{e^{\sqrt{x}}<n} \Lambda(n) \left(\int_{\sqrt{T}}^{2\sqrt{T}} \sum_{\ell^2<Tx} \int_{1+\epsilon}^{U} \frac{e^{2\sigma\sqrt{x-\frac{\ell^2}{T}}}}{n^{2\sigma}} \frac{d\sigma}{(\sigma^2+t^2)} dt \right)^{\frac{1}{2}}. \end{aligned} \tag{3.69}$$

Let $F(t) = t(\sqrt{x-\frac{\ell_1^2}{T}} - \sqrt{x-\frac{\ell_2^2}{T}})$ and $G(t) = \frac{1}{(\sigma^2+t^2)}$. Then we conclude that $|F'(t)| \gg$

$\frac{\ell_2^2-\ell_1^2}{T\sqrt{x}}$ and $|G(t)| \lesssim \frac{1}{(\sigma^2+T)}$. Using lemma 3.42

$$\begin{aligned}\sum_{e^{\sqrt{x}}<n} |z_n(V)| &\lesssim \frac{\sqrt[4]{x}}{\sqrt[4]{T}} \sum_{e^{\sqrt{x}}<n} \Lambda(n) \left(\sum_{\ell_1^2<\ell_2^2<Tx} \frac{1}{\ell_2^2-\ell_1^2} \int_{1+\epsilon}^{U} \frac{e^{\sigma(\sqrt{x-\frac{\ell_1^2}{T}}+\sqrt{x-\frac{\ell_2^2}{T}})}}{n^{2\sigma}} d\sigma \right)^{\frac{1}{2}} \\ &\quad + \frac{\sqrt[4]{x}}{\sqrt[4]{T}} \sum_{e^{\sqrt{x}}<n} \Lambda(n) \frac{e^{(1+\epsilon)\sqrt{x}}}{n^{1+\epsilon}} \\ &\lesssim \frac{\sqrt[4]{x}}{\sqrt[4]{T}} \sum_{e^{\sqrt{x}}<n} \Lambda(n) \frac{e^{(1+\epsilon)\sqrt{x}}}{n^{1+\epsilon}} \left(\sum_{m<Tx} \frac{\tau(m)}{m} \right)^{\frac{1}{2}} \\ &\quad + \frac{\sqrt[4]{x}}{\sqrt[4]{T}} \sum_{e^{\sqrt{x}}<n} \Lambda(n) \frac{e^{(1+\epsilon)\sqrt{x}}}{n^{1+\epsilon}} \lesssim \frac{\sqrt[4]{x}\log(T) e^{(1+\epsilon)\sqrt{x}}}{\sqrt[4]{T}}.\end{aligned}$$

So there exists $\sqrt{T} < V < 2\sqrt{T}$ such that

$$A_2 \lesssim \sum_{e^{\sqrt{x}}\leqslant n} |z_n(V)| \lesssim \frac{\sqrt[4]{x}\log(T) e^{(1+\epsilon)\sqrt{x}}}{\sqrt[4]{T}}. \tag{3.70}$$

Putting (3.67) and (3.70) into (3.64) and comparing it with (3.60) gives

$$\sum_{\ell^2<Tx} (-1)^l \Psi\left(e^{\sqrt{x-\frac{\ell^2}{T}}}\right) \lesssim \sqrt{Tx}\, e^{(1+\epsilon)\sqrt{\frac{x}{1+\pi^2}}} + \frac{x^{\frac{3}{4}}}{\sqrt[4]{T}} e^{(1+\epsilon)\sqrt{x}}.$$

Taking $T = e^{\frac{4(1+\epsilon)}{3}\sqrt{x}(1-\sqrt{\frac{1}{1+\pi^2}})}$ gives the desired result. ∎

3.4 Proof related to the pentagonal number theorem.

We start this section by proving the weak pentagonal number theorem for truncation of the usual partition function.

We start with the proof of proposition 3.15.

Proof. For (3.16) we only need to put $c = \pi\sqrt{\frac{2}{3}}$, $a = \frac{3}{2}$, $b = -\frac{1}{2}$, $d = 0$ in theorem 3.4; for equation (3.18), pick $c = \frac{\pi}{\sqrt{6}}$, $a = \frac{3}{2}$, $b = -\frac{1}{2}$, $d = 0$ and use Theorem 3.4; and for equation

(3.19) we need to pick $c = \frac{\pi}{\sqrt{6}}$ and $a = 1$, and $b = d = 0$.

We prove equation (3.17). Let $f(z) = \pi\sqrt{\frac{24(x-\frac{z(3z-1)}{2})-1}{36}}$ and $b^2 = \frac{1}{12}$. We choose the branch cut $(-\infty, \alpha_1] \cup [\alpha_2, \infty)$. Then let G be the interior of the square with vertices (see figure 3.1)

$$\pm\sqrt{\frac{2x}{3}} \mp 1 \pm ib\sqrt{x}, \tag{3.71}$$

Define

$$h(z) := \frac{e^{f(z)}}{\sin^3(\pi z)} = \frac{e^{\pi\sqrt{\frac{24(x-\frac{z(3z-1)}{2})-1}{36}}}}{\sin^3(\pi z)}. \tag{3.72}$$

Using the residue theorem

$$\int_\gamma h(z)dz = 2\pi i \cdot \sum_{z_i\text{: poles}} \text{Res}(h(z_i)) \tag{3.73}$$

We compute the residues of $h(z)$. We know that for z near to $\ell \in \mathbb{Z}$ we have

$$\frac{1}{(\sin(\pi z))^3} = \frac{(-1)^\ell}{\pi^3(z-\ell)^3} + \frac{(-1)^\ell}{2\pi(z-\ell)} + \cdots$$
$$e^{f(z)} = e^{f(\ell)} + f'(\ell)e^{f(\ell)}(z-\ell) + \frac{(f'(\ell))^2 + f''(\ell)}{2}e^{f(\ell)}(z-\ell)^2 + \cdots$$

Also

$$(f(z))' = \pi(1-6z)\left(24(x-\frac{z(3z-1)}{2})-1\right)^{-1/2}$$
$$(f(z))'' = -144\pi x\left(24(x-\frac{z(3z-1)}{2})-1\right)^{-3/2} \tag{3.74}$$

So we have

$$\begin{aligned}
\text{Res } h(z)|_{z=\ell} &= (-1)^\ell e^{f(\ell)}\left(\frac{1}{2\pi}+\frac{(f'(\ell))^2+f''(\ell)}{2\pi^3}\right)\\
&= \frac{(-1)^\ell e^{f(\ell)}}{2\pi(24(x-G_\ell)-1)}\left(24(x-G_\ell)-1+(1-6\ell)^2-\frac{144x}{\pi\sqrt{24(x-G_\ell)-1}}\right)\\
&= \frac{(-1)^\ell e^{f(\ell)}}{2\pi(24(x-G_\ell)-1)}\left(24x-\frac{144x}{\pi\sqrt{24(x-G_\ell)-1}}\right)
\end{aligned}$$

It implies that

$$\begin{aligned}
\int_\gamma h(z)dz &= 24ix\sum_{G_l<x}(-1)^l\frac{e^{\frac{\pi}{6}\sqrt{24(x-G_l)-1}}}{(24(x-G_l)-1)}\left(1-\frac{6}{\pi\sqrt{24(x-G_l)-1}}\right)\\
&= \frac{24ix}{\sqrt{12}}\sum_{G_l<x}(-1)^l p_2(x-G_l). \qquad (3.75)
\end{aligned}$$

We bound the integral. First assume that we choose $z \in \gamma_1 \cup \gamma_3$. So $z = t \pm ib\sqrt{x}$ for $-\sqrt{\frac{2x}{3}}+1<t<\sqrt{\frac{2x}{3}}-1$. For large enough x we have

$$\frac{24(x-\frac{z(3z-1)}{2})-1}{36} \sim \frac{2}{3}x-t^2+b^2x\mp 2ibt\sqrt{x}.$$

Similar to the proof of theorem 3.4 for $z \in \gamma_1, \gamma_3$

$$e^{\pi\sqrt{\frac{24\left(x-\frac{z(3z-1)}{2}\right)-1}{36}}} \leqslant e^{\pi\sqrt{\frac{2x}{3}+b^2x}}.$$

Also $|\sin^3(\pi z)| \sim \frac{1}{8}e^{3\pi b\sqrt{x}}$, and considering $b^2=\frac{1}{12}$ we get that

$$\pi\left(\sqrt{\frac{2}{3}+b^2}-3b\right)=0. \qquad (3.76)$$

So the contribution of the horizontal legs is at most $o(x)$. Now we compute the case $z \in$

γ_2, γ_4. We have $z = \pm\sqrt{\frac{2}{3}x} \mp 1 + it$ and $-b\sqrt{x} < t < b\sqrt{x}$. We have

$$\frac{24(x - \frac{z(3z-1)}{2}) - 1}{36} = \frac{2}{3}x - \frac{2}{3} \times \frac{2x \mp 2it\sqrt{6x} - 3t^2 \mp 7it - (2\sqrt{6} + \sqrt{\frac{2}{3}})\sqrt{x} + 4}{2} - \frac{2}{3}$$

If $t = o(\sqrt{x})$, then $\frac{24(x - \frac{z(3z-1)}{2}) - 1}{36} = o(x)$. Otherwise, since $\sqrt{x}, t$ are negligible in comparison to x, t^2

$$\frac{24(x - \frac{z(3z-1)}{2}) - 1}{36} \sim t^2 \mp it\sqrt{\frac{8x}{3}}.$$

Using lemma 3.34 we have

$$Re\left(\sqrt{\frac{24(x - \frac{z(3z-1)}{2}) - 1}{36}}\right) \leqslant \sqrt{\frac{t}{2\sqrt{3}}\left(\sqrt{3t^2 + 8x} + t\sqrt{3}\right)} \tag{3.77}$$

Hence in any case we get

$$\left|\frac{e^{\pi\sqrt{\frac{24(x - \frac{z(3z-1)}{2}) - 1}{36}}}}{(\sin(\pi z))^3}\right| \lesssim e^{\pi\sqrt{\frac{t}{2\sqrt{3}}\left(\sqrt{3t^2+8} + t\sqrt{3}\right)} - 3\pi t}. \tag{3.78}$$

Maximizing for t in the RHS of (3.78), we get that the integral in the LHS of (3.73) can be at most $e^{0.21\sqrt{x}}$. It completes the proof. □

Proof of Proposition 3.20. We have

$$\sum_{\ell^2<x}(-1)^\ell p_3(x-\ell^2) = \sqrt{6}e^{\pi ix}\sum_{\ell^2<x}\left(\frac{1}{24(x-\ell^2)-1}-\frac{12}{\pi(24(x-\ell^2)-1)^{\frac{3}{2}}}\right)e^{\frac{\pi}{12}\sqrt{24(x-\ell^2)-1}}$$

$$= \sqrt{6}e^{\frac{\pi}{12}\sqrt{24x-1}+\pi ix}\sum_{\ell^2<\frac{x}{4}} e^{\frac{-2\pi\ell^2}{\sqrt{24x-1}(\sqrt{1-\frac{24\ell^2}{24x-1}}+1)}}\left(\frac{1}{24(x-\ell^2)-1}+O\left(\frac{1}{\sqrt{x^3}}\right)\right)$$

$$\sim \frac{e^{\frac{\pi}{12}\sqrt{24x-1}+\pi ix}}{4x\sqrt{6}}\sum_{\ell^2<\sqrt{x}\ln x} e^{-\ell^2/2\sigma^2},$$

where

$$\sigma^2 = \frac{\sqrt{6x}}{\pi}.$$

The last expression in the above can be approximated as follows

$$\begin{aligned}
\frac{1}{4x\sqrt{6}}e^{\frac{\pi}{12}\sqrt{24x-1}+\pi ix}\sum_{\ell^2<\sqrt{x}\ln x} e^{-\frac{\ell^2}{2\sigma^2}} &\sim \frac{1}{4x\sqrt{6}}e^{\frac{\pi}{12}\sqrt{24x-1}+\pi ix}\int_{-\infty}^{\infty} e^{-t^2/2\sigma^2}dt \\
&= \frac{\sigma\sqrt{2\pi}}{4x\sqrt{6}}e^{\frac{\pi}{12}\sqrt{24x-1}+\pi ix} \\
&\sim \frac{e^{\pi ix}}{2^{3/4}x^{1/4}}\sqrt{p(x)}.
\end{aligned}$$

□

We need the next lemma.

Lemma 3.79. *With the same notation as theorem 3.4*

$$\sum_{n:an^2+bn+d<x}(-1)^n I_\alpha\left(c\sqrt{x-an^2+bn+d}\right)h(n) = O\left(e^{cw\sqrt{x}}\right). \qquad (3.80)$$

where I_α is the Bessel function.

Before we mention the proof note that for fixed α and large enough x

$$I_\alpha(x) \sim \frac{e^x}{\sqrt{2\pi x}}\left(1+O\left(\frac{1}{x}\right)\right).$$

Proof. Since the proof is very similar to proof of 3.4, we skip the details. Let $H(z) = \frac{h(z)I_\alpha(\sqrt{x-z^2})}{\sin(\pi z)}$ and $q(n) = an^2 + bn + d$ and assume the contour γ in 3.1. Then

$$\sum_{\ell:q(\ell)<x} (-1)^\ell I_\alpha\left(c\sqrt{x-q(\ell)}\right)h(\ell) = \int_\gamma H(z)dz$$

For $z \in \gamma_1, \gamma_3$

$$|H(z)| \lesssim \frac{I_\alpha\left(c\sqrt{x(1+au^2)}\right)}{e^{\pi u\sqrt{x}}} \lesssim e^{c\sqrt{x(1+au^2)}-\pi u\sqrt{x}}.$$

Also for $z \in \gamma_2, \gamma_4$

$$|H(z)| \lesssim I_\alpha\left(c\sqrt{\frac{ux}{2}\left(\sqrt{au^2+4}+u\sqrt{a}\right)}\right) \lesssim \sqrt{x}e^{c\sqrt{x\sqrt{a}\alpha\frac{\sqrt{a\alpha^2+4}+\alpha\sqrt{a}}{2}}-\pi\alpha\sqrt{x}}.$$

with the same notation as in proof of theorem 3.4. As the bound of argument of Bessel function is the same as exponents in the proof of theorem 3.4 we get the same bound. □

Proof of corollaries 3.22 and 3.23. For corollary 3.22 pick $a = 1$, and $c = \sqrt{\frac{2\pi^2}{3}}$ in the Lemma 3.79. For corollary 3.23 pick $c = \frac{2\pi}{\sqrt{15}}$ and $a = 1$ in Lemma 3.79. □

3.5 Proof related to the Prouhet-Tarry-Escott problem

It is worth establishing a "baseline result" related to problem 3.24 for N large, relative to k, n, that we get easily from a Pigeonhole Argument: consider all vectors $(x, x^2, \cdots, x^k)$ with $1 \leqslant x \leqslant N$. The sum of n of these lie in a box of volume $n^k N^{k(k+1)/2}$; and if two such sums belong to the same box with dimensions $N^c \times N^{2c} \times \cdots \times N^{kc}$, then they give a solution to (3.24) for all $1 \leqslant i \leqslant k$. The number of non over-lapping $N^c \times \cdots \times N^{kc}$ boxes that fit inside our volume $n^k N^{k(k+1)/2}$ is at most $n^k N^{(1-c)k(k+1)/2}$; and with a little work one can see that the large box can be covered with approximately (up to a constant factor) this many smaller boxes. If this (the number of smaller boxes in a covering) is smaller than the number of sets of n vectors $(x, x^2, \cdots, x^k)$ that produce our vector sum (this count is at least $\frac{N^n}{n!}$ for N large enough relative to n) then we get a "collision", that is a pair of sequences $a_1, \cdots, a_n$ and $b_1, \cdots, b_n$ leading to a solution to 3.24 for all $1 \leqslant i \leqslant k$. In other words, we get such a solution when

$$n^k N^{\frac{(1-c)k(k+1)}{2}} < \frac{N^n}{n!}.$$

For N large, then, we get that there is a solution so long as

$$c > 1 - \frac{2n}{k(k+1)}. \tag{3.81}$$

When k is a little smaller than $\sqrt{2n}$, note that the RHS is negative, implying that we can take $c = 0$ (since it must be non-negative).

Curiously, when k is only a little bigger than $\sqrt{n}$ (say, $\sqrt{n}\log(n)$), then this pigeonhole argument only gives us pairs of sequences with c near to 1. Basically, then, we don't get a much better result for the weakening than we do for the original Prouhet-Tarry-Escott Problem, if we insist on finding solutions with $c < \frac{1}{2}$, say.

We prove a lemma before introducing a set of solutions for the weak Prouhet-Terry-Escott problem (problem 3.24).

Lemma 3.82. *For large x, let $k \lesssim \frac{\sqrt{x}}{\log(xT)}$ and $T := T(x) = o(x)$. Then for every $1 \leqslant r \leqslant k$ there exists $c > 0$ such that*

$$\sum_{\ell^2 < xT} (-1)^\ell \left(xT - \ell^2\right)^{\frac{r}{2}} \lesssim \sqrt{x}(Tx)^{\frac{r}{4}}(Ar)^{r/2} \tag{3.83}$$

Remark 3.84. *Note that the proof becomes easier if we just choose r to be even. But we propose a more general case here.*

Proof. Let $u = o(\sqrt{x})$, to be determined later. Define

$$f_r(z) = \frac{(xT - z^2)^{\frac{r}{2}}}{\sin(\pi z)}.$$

Let γ be the contour in Figure 3.5. Using the residue Theorem

$$\int_\gamma f_r(z)dz = 2\pi i \sum_{\ell^2 < xT} (-1)^\ell \left(Tx - \ell^2\right)^{\frac{r}{2}}.$$

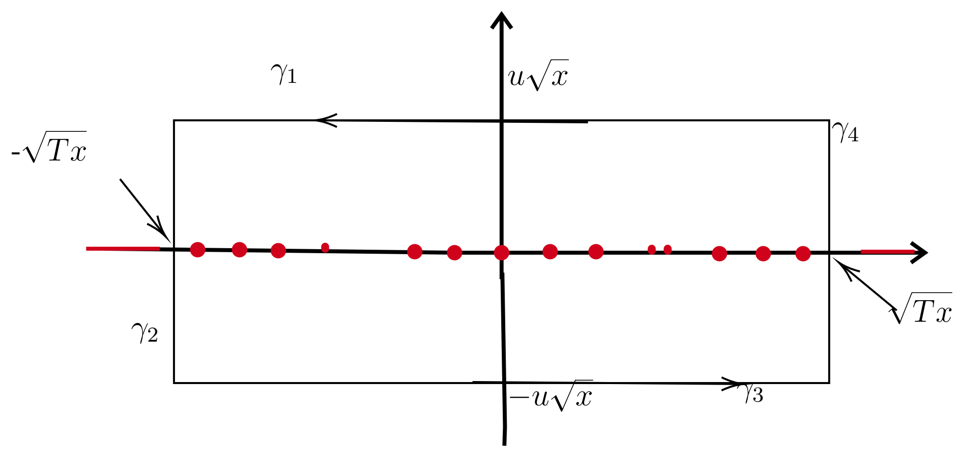

Figure 3.5: The contour γ

Let $z \in \gamma_1, \gamma_3$. So $z = t \pm iu\sqrt{x}$ and $-\sqrt{xT} < t < \sqrt{xT}$. Then

$$\left|xT - z^2\right|^2 = \left(xT + u^2x - t^2\right)^2 + 4t^2u^2x.$$

Note that u is a constant as x tends to infinity and $T = o(x)$. With simple computation we can conclude that the RHS is maximaized at $t = 0$, so on γ_1, γ_3 we have

$$|f_r(z)| \lesssim \left|xT + u^2x\right|^{\frac{r}{4}} e^{-\pi u\sqrt{x}} \sim (xT)^{\frac{r}{4}} e^{-\pi u\sqrt{x}}. \tag{3.85}$$

By assumption $r \leqslant k \lesssim \sqrt{x}/\log(xT)$, so we can pick u to be large enough so as the contribution of horizental legs become small. For γ_2, γ_4 we have $z = \pm\sqrt{xT} + it$ and

$-u\sqrt{x} < t < u\sqrt{x}$. We can show that

$$\left|xT - z^2\right|^2 = t^4 + 4t^2 xT.$$

So we need to maximize the RHS of the following expression for $t \leqslant u\sqrt{x}$

$$|f_r(z)| \lesssim (t^4 + 4t^2 xT)^{\frac{r}{4}} e^{-\pi t}$$

Simple computation shows that it happens when $t \sim Cr$ for some $C > 0$. Hence, there exist $A > 0$ such that

$$|f_r(z)| \lesssim A^r (xT)^{\frac{r}{4}} \left(r^2 + \frac{r^4}{xT}\right)^{\frac{r}{4}} \lesssim (Ar)^{r/2} (xT)^{\frac{r}{4}}.$$

This completes the proof. $\square$

Remark 3.86. *We could increase the height of vertical lines of figure 3.5 to x^α, $\alpha > \frac{1}{2}$, to make it possible for k to become bigger - say $k \gg x^\alpha$. This in turn results in larger $k = M(n)$ and larger error term.*

Proof of Theorem 3.26. Let M be a large number.

$$x_i = M^{2m+b} - (2i-2)^2 \qquad y_i = M^{2m+b} - (2i-1)^2$$

Then $\max(x_i^r, y_i^r) \sim M^{2m+b}$. Lemma 3.82 concludes that for $x = M^{2m}$ and $T = M^b$ and $1 \leqslant r \leqslant k$

$$\sum_i x_i^r - \sum_i y_i^r \lesssim (r)^{\frac{r}{2}+\epsilon} M^{(2m+b)\frac{r}{4}+m}.$$

If we pick $k \leqslant \frac{u\pi M^m}{12m^2 \log(M)}$ and $b = 1$, then the result follows. $\square$

Proof of Theorem 3.33. We first show that $f_r(M)$ is a polynomial in M – that is,

$$f_r(M) = c_0(r) + c_1(r)M + \cdots + c_d(r)M^d,$$

where d is yet to be determined. This follows upon applying the binomial theorem to the terms in the definition of $f_r(M)$, together with the fact that $\sum_{|\ell|<2M}(-1)^\ell \ell^k$ is a polynomial in M. The coefficients are obviously integers and we also can show the coefficients as sums involving Bernouli numbers. Note that the degree d of that polynomial doesn't depend on M.

Let's assume that r is even. We now leverage this fact to prove that $d = r - 1$. To do this, note that it suffices to prove that $|f_r(M)| = o_r(M^r)$, and $|f_r(M)| \gg_r M^{r-1}$. To put that another way: fix r, and then we show that

$$\lim_{M\to\infty} \frac{\log(|f_r(M)|)}{\log(M)} = r - 1.$$

Write $f_r(M)$ as the contour integral

$$\frac{1}{2\pi i}\int_\gamma f(z)dz := \frac{1}{2\pi i}\int_\gamma \frac{(4M^2 - z^2)^r}{\sin(\pi z)}dz,$$

where γ is in figure **??**. Note that because f has a removable singularity at $z = \pm 2M$, it is possible to compute the contribution of the integral in these vertical legs.

Now, one easily sees that the contribution of γ_1, γ_3 is negligible, and at least for M large relative to r the main contribution will come from the part of the contour near the real axis. These two parts of the contour can be parametrized as $z = 2M + it$ and $z = -2M + it$, $|t| \leqslant 2M$. So, for M large relative to r we will have that the integral is

$$\sim \frac{1}{\pi}\int_{-2M}^{2M} \frac{(-4Mit + t^2)^r}{\sin(\pi it)}dt = \frac{1}{\pi}\int_{-2M}^{2M} \frac{(-4Mit)^r}{\sin(\pi it)}dt + \frac{1}{\pi}\int_{-2M}^{2M} \frac{r(-4Mit)^{r-1}t^2}{\sin(\pi it)}dt + O(M^{r-2})$$
$$\sim 0 + \frac{r}{\pi}(-4Mi)^{r-1}\int_{-\infty}^{\infty} \frac{t^{r+1}}{\sin(\pi it)}dt \sim cM^{r-1},$$

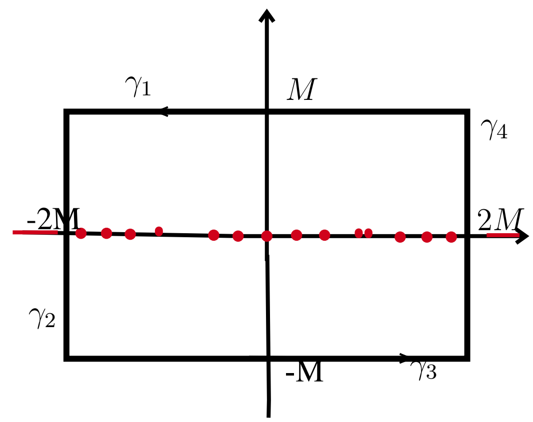

Figure 3.6: The contour γ

for a constant c that depends only on r. Note that the first term of RHS is zero by symmetry. This means that $f_r(M)$ is of degree $r-1$. Also we bound the size of $f_r(M)$ from above in the range $r \lesssim \frac{M}{\log(M)}$.

$$\int_\gamma f(z)dz \sim cM^{r-1} \lesssim e^{r(\log(r)+\log\log(r))}M^{r-1}.$$

□

CHAPTER 4

ENDPOINT ℓ^R IMPROVING ESTIMATES FOR PRIME AVERAGES

4.1 Introduction

We consider discrete averages over the prime integers. The averages are weighted by the von Mangoldt function.

$$A_N f(x) = \frac{1}{N} \sum_{1 \leqslant n \leqslant N} f(x-n)\Lambda(n) \tag{4.1}$$

Our interest is in *scale free* ℓ^r improving estimates for these averages. The question presents itself in different forms.

For an interval I in the integers and function $f : I \to \mathbb{C}$, set

$$\langle f \rangle_{I,r} = \Big[|I|^{-1} \sum_{x \in I} |f(x)|^r \Big]^{1/r}. \tag{4.2}$$

If $r = 1$, we will suppress the index in the notation. And, set $\operatorname{Log} x = 1 + |\log x|$, for $x > 0$.

The kind of estimate we are interested in takes the the following form, in the simplest instance. What is the 'smallest' function $\psi : [0,1] \to [1, \infty)$ so that for all integers N and indicator functions $f, g : I \to \{0, 1\}$, there holds

$$N^{-1} \langle A_N f, g \rangle \leqslant \langle f \rangle_I \langle g \rangle_I \psi(\langle f \rangle_I \langle g \rangle_I).$$

That is, the right hand side is independent of N, making it scale-free. We specified that f, g be indicator functions as that is sometimes the sharp form of the inequality. Of course it is interesting for arbitrary functions, but the bound above is not homogeneous, so not the most natural estimate in that case.

The points of interest in these two results arises from, on the one hand, the distinguished role of the prime integers. And, on the other, endpoint results are significant interest in Harmonic Analysis, as the techniques which apply are the sharpest possible. In this instance, the sharp methods depend very much on the prime numbers.

For the primes, we expect that the Riemann Hypothesis to be relevant. We state unconditional results, and those that depend upon the Generalized Riemann Hypothesis (GRH). Remember that according to GRH all zeroes in the critical strip $0 < Re(s) < 1$ of an arbitrary $L-$function $L(f,s)$ are on the critical line $Re(s) = \frac{1}{2}$. Under GRH, the primes are equitably distributed mod q, with very good error bounds. Namely,

$$\psi(x,q,a) = \sum_{\substack{n<x \\ n\equiv a \pmod q}} \Lambda(n) = \frac{x}{\phi(q)} + O(x^{\frac{1}{2}}\log^2(q)). \tag{4.3}$$

Theorem 4.4. *There is a constant C so that this holds. For integers $N > 30$, and interval I of length N, the following inequality holds for all functions $f = \mathbf{1}_F$ and $g = \mathbf{1}_G$ with $F, G \subset I$*

$$N^{-1}\langle A_N f, g\rangle \leqslant C\langle f\rangle_I\langle g\rangle_I \times \begin{cases} \operatorname{Log}(\langle f\rangle_I\langle g\rangle_I) & \text{assuming GRH} \\ (\operatorname{Log}(\langle f\rangle_I\langle g\rangle_I))^2 & \end{cases} \tag{4.5}$$

The inequality assuming GRH is sharp, as can be seen by taking f to be the indicator of the primes, and $g = \mathbf{1}_0$. It is also desirable to have a form of the inequality above that holds for the maximal function

$$A^*f = \sup_N |A_N f|.$$

Our second main theorem is sparse bound for A^*. The definition of a sparse bound is postponed to Definition 4.52. Remarkably, the inequality takes the same general form, although we consider a substantially larger operator.

Theorem 4.6. *For functions $f = \mathbf{1}_F$ and $g = \mathbf{1}_G$, for finite sets $F, G \subset \mathbb{Z}$, there is a sparse collection of intervals $\mathcal{S}$ so that we have*

$$\langle A^* f, g\rangle \lesssim \sum_{I \in \mathcal{S}} \langle f\rangle_I \langle g\rangle_I (\operatorname{Log}\langle f\rangle_I \langle g\rangle_I)^t |I|, \tag{4.7}$$

where we can take $t = 1$ under GRH, and otherwise we take $t = 2$.

The sparse bound is very strong, implying weighted inequalities for the maximal operator A^*. These inequalities could be further quantified, but we do not detail those consequences, as they are essentially known. See [40].

This subject is an outgrowth of Bourgain's fundamental work on arithmetic ergodic theorems [41, 42]. These inequalities proved therein focused on the diagonal case, principally ℓ^p to ℓ^p estimates for maximal functions. Bourgain's work has been very influential, with a very rich and sophisticated theory devoted to the diagonal estimates. We point to [43, 44], and very recently [45, 46]. The subject is very rich, and the reader should consult the references in these papers.

Shortly after Bourgain's first results, Wierdl [47] studied the primes, and the simpler form of the Circle method in that case allowed him to prove diagonal inequalities for all $p > 1$, which was a novel result at that time. The result was revisited by Mirek and Trojan [48]. The approach of this paper differs in some important aspects from the one in [49]. (The low/high decomposition is dramatically different, to point to the single largest difference.)

The subject of sparse bounds originated in harmonic analysis, with a detailed set of applications in the survey [50], with a wide set of references therein. The paper [51] initiated the study of sparse bounds in the discrete setting. While the result in that paper of an 'ϵ improvement' nature, for averages it turns out there are very good results available, as was first established for the discrete sphere in [52, 53]. There is a rich theory here, with a range of inequalities for the Magyar-Stein-Wainger [54] maximal function in [55]. Nearly sharp results for certain polynomial averages are established in [56, 57], and a surprisingly good estimate for arbitrary polynomials is in [58]. The latter result plays an interesting role in the innovative result of Krause, Mirek and Tao [59].

The ℓ^p improving property for the primes was investigated in [60], but not at the endpoint. That paper result established the first weighted estimates for the averages for the prime numbers. This paper establishes the sharp results, under GRH. Mirek [61] addresses the diagonal case for Piatetski-Shapiro primes. It would be interesting to obtain ℓ^p improving estimates in this case.

Our proof uses the Circle Method to approximate the Fourier multiplier, following Bourgain [41]. In the unconditional case, we use Page's Theorem, which leads to the appearance of exceptional characters in the Circle method. Under GRH, there are no exceptional characters, and one can identify, as is well known, a very good approximation to the multiplier.

The Fourier multiplier is decomposed at the end of §4.3 in such a way to fit an interpolation argument of Bourgain [62], also see [63]. We call it the High/Low Frequency method. To acheive the endpoint results, this decomposition has to be carefully phrased. There are two additional features of this decomposition we found necessary to add in. First, certain difficulties associated with Ramanujan sums are addressed by making a significant change to a Low Frequency term. The sum defining the Low Frequency term (4.28) is over all Q-smooth square free denominators. Here, the integer Q can vary widely, as small as 1 and as large as $N^{1/10}$, say. (The largest Q-smooth square denominator will be of the order of e^Q.) Second, in the unconditional case, the exceptional characters are grouped into their own term. As it turns out, they can be viewed as part of the Low Frequency term. The properties we need for the High/Low method are detailed in §4.4. The following sections are applications of those properties.

4.2 Notation

We write $A \lesssim B$ if there is a constant C so that $A \leqslant CB$. In such instances, the exact nature of the constant is not important.

Let $\mathcal{F}$ denote the Fourier transform on $\mathbb{R}$, defined for by

$$\mathcal{F}f(\xi) = \int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\,dx, \qquad f \in L^1(\mathbb{R}).$$

The Fourier transform on $\mathbb{Z}$ is denoted by $\widehat{f}$, defined by

$$\widehat{f}(\xi) = \sum_{n\in\mathbb{Z}} f(n)e^{-2\pi i n\xi}, \qquad f \in \ell^1(\mathbb{Z}).$$

Throughout this chapter, we denote $A_q = \{a \in \mathbb{Z}/q\mathbb{Z} \ : \ (a,q) = 1\}$, so that $|A_q| = \phi(q)$, the totient function. We have

$$\frac{q}{\operatorname{Log}\operatorname{Log} q} \lesssim \phi(q) \leqslant q - 1. \tag{4.8}$$

It is known that for non-principal characters χ, we have $|G(\chi, a)| < q^{-\frac{1}{2}}$, see [64, Chapter 3]. As for the principal character, if χ is identity, then we get Ramanujan's sum

$$c_q(n) := \phi(q)G(\mathbf{1}_{A_q}, a) = \sum_{r\in A_q} e\Big(\frac{ra}{q}\Big). \tag{4.9}$$

Let χ_q denote the exceptional character. It is a non-trivial quadratic Dirichlet character modulo q, that is χ_q takes values $-1, 0, 1$, and takes the value -1 at least once. We also know that χ_q is primitive, namely that its period is q. As a matter of convenience, if q does not have an exceptional character, we will set $\chi_q \equiv 0$, and $\beta_q = 1$. These properties are important to Lemma 4.45.

Page's Theorem uses the exceptional characters to give an approximation to the prime counting function. Counting primes in an arithmetic progression of modulus q, we have

$$\psi(N; q, r) - \frac{N}{\phi(q)} + \frac{\chi_q(x)}{\phi(q)} \beta_q^{-1} x^{\beta_q} \lesssim N e^{c\sqrt{\log N}}. \tag{4.10}$$

4.3 Approximations of the Kernel

Denote the kernel of A_N with the same symbol, so that $A_N(x) = N^{-1} \sum_{n \leqslant N} \Lambda(n) \delta_n(x)$. It follows that

$$\widehat{A_N}(\xi) = \frac{1}{N} \sum_{n \leqslant N} \Lambda(n) e^{-2\pi n \xi}.$$

The core of the paper is the approximation to $\widehat{A_N}(\xi)$, and its further properties, detailed in the next section.

Set

$$M_N^\beta = \frac{1}{N\beta} \sum_{n \leqslant N} [n^\beta - (n-1)^\beta] \delta_n, \qquad \tfrac{1}{2} < \beta \leqslant 1. \tag{4.11}$$

We write $M_N = M_N^1$ when $\beta = 1$, which is the standard average. For $\beta < 1$, these are not averaging operators. They are the operators associated to the exceptional characters. The Fourier transforms are straight forward to estimate.

Proposition 4.12. *We have the estimates*

$$|\widehat{M_N}(\xi)| \lesssim \min\{1, (N|\xi|)^{-1}\}, \tag{4.13}$$

$$|\widehat{M_N^\beta}(\xi)| \lesssim (N|\xi|)^{-1}, \tag{4.14}$$

$$|\widehat{M_N^\beta}(\xi) - \beta^{-1} N^{\beta-1}| \lesssim N^\beta |\xi|. \tag{4.15}$$

For integers q and $a \in A_q$,

$$\widehat{L_N^{a,q}}(\xi) = G(\mathbf{1}_{A_q}, a) \widehat{M_N}(\xi) - G(\chi_q, a) \widehat{M_N^{\beta_q}}(\xi) \tag{4.16}$$

We state the approximation to the kernel at rational point, with small denominator.

Lemma 4.17. *Assume that* $|\xi - \frac{a}{q}| \leqslant N^{-1}Q$ *for some* $1 \leqslant a \leqslant q \leqslant Q$ *and* $\gcd(a,q) = 1$. *Then*

$$\widehat{A_N}(\xi) = \widehat{L_N^{a,q}}(\xi - \tfrac{a}{q}) + \begin{cases} O(QN^{-\frac{1}{2}+\epsilon}), & \textit{Assuming GRH} \\ O(Qe^{-c\sqrt{n}}), & \textit{Otherwise} \end{cases} \tag{4.18}$$

Proof. We proceed under GRH, and return to the unconditional case at the end of the argument. The key point is that we have the approximation (4.3) for $\psi(N;q,r)$. Set $\alpha := \xi - \frac{a}{q}$. Using Abel summation, we can write

$$N\widehat{M_N}(\alpha) = Ne(\alpha N) - \sqrt{N}e(\alpha\sqrt{N}) - 2\pi i\alpha \int_{\sqrt{N}}^{N} e^{t\alpha}\, dt + O(\sqrt{N}).$$

Turning to the primes, we separate out the sum below according to residue classes mod q. Since $\xi = \frac{a}{q} + \alpha$,

$$\begin{aligned} \sum_{\ell \leqslant N} e(\xi\ell)\Lambda(\ell) &= \sum_{\substack{0 \leqslant r \leqslant q \\ \gcd(r,q)=1}} \sum_{\substack{\ell \leqslant N \\ \ell \equiv r \mod q}} e(\xi\ell)\Lambda(\ell) \\ &= \sum_{r \in A_q} e\Big(\frac{ra}{q}\Big) \sum_{\substack{\ell \leqslant N \\ \ell \equiv r \mod q}} e(\alpha\ell)\Lambda(\ell). \end{aligned}$$

Examine the inner sum. Using Abel's summation formula, and the notation ψ for prime counting function, we have

$$\begin{aligned} \sum_{\substack{\ell \leqslant N \\ \ell \equiv r \mod q}} e(\alpha\ell)\Lambda(\ell) = {} & \psi(N;q,r)e(\alpha N) - \psi(\sqrt{N};q,r)e(\alpha\sqrt{N}) \\ & - 2\pi i\alpha \int_{\sqrt{N}}^{N} \psi(t;q,r)e(\alpha t)dt + O(\sqrt{N}). \end{aligned}$$

At this point we can use the Generalized Riemann Hypothesis. From (4.3), it follows that

$$\sum_{\substack{\ell \leqslant N \\ \ell \equiv r \mod q}} e(\alpha\ell)\Lambda(\ell) - \frac{N}{\phi(q)}\widehat{M_N}(\alpha) = (\psi(N;q,r) - \frac{N}{\phi(q)}e(\alpha N))e(\alpha N)$$
$$- 2\pi i\alpha \int_{\sqrt{N}}^{N} e(t\alpha)(\psi(t;q,r) - t)\,dt + O(\sqrt{N})$$
$$\lesssim N^{\frac{1}{2}+\epsilon} + \frac{Q}{N}\int_{\sqrt{N}}^{N} t^{\frac{1}{2}+\epsilon} dt + O(N^{\frac{1}{2}+\epsilon})$$
$$\lesssim QN^{\frac{1}{2}+\epsilon}.$$

The proof without GRH uses Page's Theorem (4.10) in place of (4.3). We omit the details.

□

The previous Lemma approximates $\widehat{A_N}(\xi)$ near a rational point. We extend this approximation to the entire circle. This is done with these definitions.

$$\widehat{V_{s,n}}(\xi) = \sum_{a/q \in \mathcal{R}_s} G(\mathbf{1}_{A_q}, a)\widehat{M_N}(\xi - a/q)\eta_s(\xi - a/q), \tag{4.19}$$

$$\widehat{W_{s,n}}(\xi) = \sum_{a/q \in \mathcal{R}_s} G(\chi_q, a)\widehat{M_N^{\beta_q}}(\xi - a/q)\eta_s(\xi - a/q), \tag{4.20}$$

$$\mathcal{R}_s = \{a/q \ : \ a \in A_q, \ 2^s \leqslant q < 2^{s+1}\}, \tag{4.21}$$

and $\mathcal{R}_0 = \{0\}$. Further $\mathbf{1}_{[-1/4,1/4]} \leqslant \eta \leqslant \mathbf{1}_{[-1/2,1/2]}$, and $\eta_s(\xi) = \eta(4^s\xi)$. In (4.27), recall that if q is not exceptional, we have $\chi_q = 0$. Otherwise, χ_q is the associated exceptional Dirichlet character. Given an integer $N = 2^n$, set

$$\tilde{N} = \begin{cases} e^{c\sqrt{n}/4} & \text{where } c \text{ is as in (4.18)} \\ N^{1/5} & \text{under GRH} \end{cases} \tag{4.22}$$

Lemma 4.23. *Let* $N = 2^n$*. Write* $A_N = B_N + \mathrm{Err}_N$*, where*

$$B_N = \sum_{s\,:\,2^s < (\tilde{N})^{1/400}} V_{s,n} - W_{s,n}. \tag{4.24}$$

Then, we have $\|\mathrm{Err}_N f\|_{\ell^2} \lesssim (\tilde{N})^{-1/1000} \|f\|_{\ell^2}$*.*

Proof. We estimate the ℓ^2 norm by Plancherel's Theorem. That is, we bound

$$\|\widehat{A_N} - \widehat{B_N}\|_{L^\infty(\mathbb{T})} \lesssim (\tilde{N})^{-1/1000}.$$

Fix $\xi \in \mathbb{T}$, where we will estimate the L^∞ norm above. By Dirichlet's Theorem, there are relatively prime integers a, q with $0 \leqslant a < q \leqslant (\tilde{N})^{1/5}$ with

$$|\xi - a/q| < \frac{1}{q^2}.$$

The argument now splits into cases, depending upon the size of q.

Assume that $(\tilde{N})^{1/400} < q \leqslant (\tilde{N})^{1/5}$. This is a situation for which the classical Vinogradov inequality [65]*Chapter 9 was designed. That estimate is however is not enough for our purposes. Instead we use [64, Chapter 9] for the estimate below.

$$|\widehat{A_N}(\xi)| \lesssim (q^{-1/2} + (q/N)^{1/2} + N^{-1/5}) \log^3 N \lesssim (\tilde{N})^{-1/1000}.$$

So, in this case we should also see that $\widehat{B_N}(\xi)$ satisfies the same bound. The function $\widehat{B_N}$ is a sum over $\widehat{V_{s,n}}$ and $\widehat{W_{s,n}}$. The argument for both is the same. Suppose that $\widehat{V_{s,n}}(\xi) \neq 0$. The supporting intervals for $\eta_s(\xi - a/q)$ for $a/q \in \mathcal{R}_s$ are pairwise disjoint. We must have $|\xi - a_0/q_0| < 2^{-2s}$ for some $a_0/q_0 \in \mathcal{R}_s$, where $2^s < (\tilde{N})^{1/400}$. Then,

$$|\xi - a_0/q_0| \geqslant |a_0/q_0 - a/q| - |\xi - a/q| \geqslant (qq_0)^{-1} - q^{-2} \geqslant q_0^{-4}.$$

But then by the decay estimate (4.13), we have

$$|G(\mathbf{1}_{A_q}, a_0)\widehat{M_N}(\xi - a_0/q_0)| \lesssim (Nq_0^{-4})^{-1} \lesssim N^{-1}(\tilde{N})^{1/100}$$

This estimate is summed over $s \leqslant (\tilde{N})^{1/400}$ to conclude this case.

Proceed under the assumption that $q \leqslant N_0 = (\tilde{N})^{1/400}$. From Lemma 4.17, the inequality (4.18) holds.

$$\widehat{A_N}(\xi) = \widehat{L_N^{a,q}}(\xi - \tfrac{a}{q}) + O(N_0^{-1/2})$$

The Big O term is as is claimed, so we verify that $\widehat{B_N}(\xi) - \widehat{L_N^{a,q}}(\xi - \frac{a}{q}) \lesssim N_0^{-1/2}$.

The analysis depends upon how close ξ is to a/q. Suppose that $|\xi - a/q| < \frac{1}{4}N_0^{-2}$. Then a/q is the unique rational b/r with $(b, r) = 1$ and $0 \leqslant b < r \leqslant N_0$ that meets this criteria. That means that

$$\widehat{B_N}(\xi) = \widehat{L_N^{a,q}}(\xi - a/q)\eta_s(\xi - a/q)$$

where in the last term on the right, $2^s \leqslant q < 2^{s+1}$. By definition $\eta_s(\xi - a/q) = \eta(4^s(\xi - a/q))$, which equals one by assumption on ξ. That completes this case.

Continuing, suppose that there is no a/q with $|\xi - a/q| < N_0^{-2}$. The point is that we have the decay estimates (4.13) and (4.14) which imply

$$|\widehat{M_N}(\xi - a/q)| + |\widehat{M_N^{\beta}}(\xi - a/q)| \lesssim [N(\xi - a/q)]^{-1} \lesssim \frac{N_0^2}{N} \lesssim N^{-3/5}.$$

But then, from the definition (4.16), we have

$$|\widehat{L_N^{a,q}}(\xi - \tfrac{a}{q})| \lesssim N^{-1/5}.$$

And as well, trivially bounding Gauss sums by 1, we have

$$|\widehat{B_N}(\xi)| \lesssim \frac{n^{3/5}}{N} \lesssim N^{-1/5},$$

by just summing over all $a/q \in \mathcal{R}_s$, with $s < (\tilde{N})^{1/400}$. That completes the proof.

□

The discussion to this point is of a standard nature. We state here a decomposition of the operator B_N defined in (4.24). It encodes our High/Low/Exceptional decomposition, and requires some care to phrase, in order to prove our endpoint type results for the prime averages. It depends upon a supplementary parameter Q. This parameter Q will play two roles, controlling the size and smoothness of denominators. Recall that an integer q is Q-*smooth* if all of its prime factors are less than Q. Let $\mathbb{S}_Q$ be the collection of square-free Q-smooth integers.

$$\widehat{V_{s,n}^{Q,\mathrm{lo}}}(\xi) = \sum_{\substack{a/q \in \mathcal{R}_s \\ q \in \mathbb{S}_Q}} G(\mathbf{1}_{A_q}, a)\widehat{M_N}(\xi - a/q)\eta_s(\xi - a/q), \tag{4.25}$$

$$\widehat{V_{s,n}^{Q,\mathrm{hi}}}(\xi) = \sum_{\substack{a/q \in \mathcal{R}_s \\ q \notin \mathbb{S}_Q}} G(\mathbf{1}_{A_q}, a)\widehat{M_N}(\xi - a/q)\eta_s(\xi - a/q), \tag{4.26}$$

$$\widehat{W_{s,n}}(\xi) = \sum_{a/q \in \mathcal{R}_s} G(\chi_q, a)\widehat{M_N^{\beta_q}}(\xi - a/q)\eta_s(\xi - a/q), \tag{4.27}$$

Define

$$\mathrm{Lo}_{Q,N} = \sum_s V_{s,n}^{Q,\mathrm{lo}}, \tag{4.28}$$

$$\mathrm{Hi}_{Q,N} = \sum_{s \,:\, Q \leqslant 2^s \leqslant (\tilde{N})^{1/400}} V_{s,n}^{Q,\mathrm{hi}} - W_{s,n} \tag{4.29}$$

$$\mathrm{Ex}_{Q,N} = \sum_{s \,:\, 2^s \leqslant Q} W_{s,n} \tag{4.30}$$

Concerning these definitions, in the Low term (4.28), there is no restriction on s, but the sum only depends upon the finite number of square-free Q-smooth numbers in $\mathbb{S}_Q$. (Due to (4.42), the non-square free integers will not contribute to the sum.) The largest integer in $\mathbb{S}_Q$ will be about e^Q, and the value of Q can be as big as $\tilde{N}$. In the High term (4.29), there are two parts associated with the principal and exceptional characters. For the principal characters, we exclude the square free Q-smooth denominators which are both larger than Q and less than $(\tilde{N})^{1/400}$. These are included in the Low term. We include all the denominators for the exceptional characters. In the Exceptional term (4.30), we just impose the restriction on the size of the denominator to be not more than Q. This will be part of the Low term.

The sum of these three terms well approximates B_N.

Proposition 4.31. *Let* $1 \leqslant Q \leqslant \tilde{N}$. *We have the estimate*

$$\|\mathrm{Err}'_{Q,N} f\|_{\ell^2} \lesssim (\tilde{N})^{-1/2} \|f\|_{\ell^2}, \tag{4.32}$$

We have the estimate

$$\|\mathrm{Err}'_{Q,N} f\|_{\ell^2} \lesssim (\tilde{N})^{-1/2} \|f\|_{\ell^2}, \tag{4.33}$$

where

$$\mathrm{Err}'_N = \mathrm{Lo}_{Q,N} + \mathrm{Hi}_{Q,N} + \mathrm{Ex}_N + \mathrm{Err}_N - B_N. \tag{4.34}$$

Proof. From (4.24), we see that

$$\widehat{\mathrm{Err}'_N}(\xi) = \sum_{s \,:\, 2^s > (\tilde{N})^{1/400}} \widehat{V_{s,n}^{Q,\mathrm{lo}}}(\xi)$$

Recalling the definition of $V_{s,n}^{Q,\mathrm{lo}}$ from (4.25), it is straight forward to estimate this last sum in $L^\infty(\mathbb{T})$, using the Gauss sum estimate $G(\mathbf{1}_{A_q}, a) \lesssim \frac{\operatorname{Log}\operatorname{Log} q}{q}$. □

4.4 Properties of the High, Low and Exceptional Terms

The further properties of the High, Low and Exceptional terms are given here, in that order.

4.4.1 The High Terms

We have the ℓ^2 estimates for the fixed scale, and for the supremum over large scales, for the High term defined in (4.29). Note that the supremum is larger by a logarithmic factor.

Lemma 4.35. *We have the inequalities*

$$\|\mathrm{Hi}_{Q,N}\|_{\ell^2\to\ell^2} \lesssim \frac{\log\log Q}{Q}, \tag{4.36}$$

$$\|\sup_{N>Q^2} |\mathrm{Hi}_{Q,N}\, f|\|_2 \lesssim \frac{\log\log Q \cdot \log Q}{Q} \|f\|_{\ell^2}. \tag{4.37}$$

We comment that the insertion of the Q smooth property into the definition of $V_{s,n}^{Q,\mathrm{hi}}$ in (4.26) is immaterial to this argument.

Proof. Below, we assume that there are no exceptional characters, as a matter of convenience as the exceptional characters are treated in exactly the same manner. For the inequality (4.36), we have from the definition of the High term in (4.29), and (4.26),

$$\begin{aligned}
\|\mathrm{Hi}_{Q,N}\|_{\ell^2\to\ell^2} &= \|\widehat{\mathrm{Hi}_{Q,N}}\|_{L^\infty(\mathbb{T})} \\
&= \Big\| \sum_{s\,:\,Q\leqslant 2^s\leqslant \tilde N} \widehat{V_{s,n}^{Q,\mathrm{hi}}} \Big\|_{L^\infty(\mathbb{T})} \\
&\leqslant \sum_{s\,:\,Q\leqslant 2^s\leqslant \tilde N} \|\widehat{V_{s,n}^{Q,\mathrm{hi}}}\|_{L^\infty(\mathbb{T})} \\
&\leqslant \sum_{s\,:\,Q\leqslant 2^s\leqslant \tilde N} \max_{2^s\leqslant q<2^{s+1}} \max_{a\in A_q} |G(\mathbf{1}_{A_q}, a)| \\
&\lesssim \sum_{s\,:\,Q\leqslant 2^s\leqslant \tilde N} \max_{2^s\leqslant q<2^{s+1}} \frac{1}{\phi(q)} \\
&\lesssim \sum_{s\,:\,Q\leqslant 2^s} \log s \cdot 2^{-s} \lesssim \frac{\log\log Q}{Q}.
\end{aligned}$$

The first line is Plancherel, and the subsequent lines depend upon definitions, and the fact that the functions below are disjointly supported.

$$\{\eta_s(\cdot - a/q) \ : \ 2^s \leqslant q < 2^{s+1}, \ a \in A_q\}.$$

Last of all, we use a well known lower bound $\phi(q) \gg q/\log\log q$.

For the maximal inequality (4.37), we have an additional logarithmic term. This is direct consequence of the Bourgain multi-frequency inequality, stated in Lemma 4.38. We then have

$$\begin{aligned}
\| \sup_{N>Q^2} |\mathrm{Hi}_{Q,N} f| \|_{\ell^2} &\leqslant \sum_{s\,:\,Q\leqslant 2^s} \Big\| \sup_{N>Q^2} |V^{Q,\mathrm{hi}}_{s,n} f| \Big\|_{\ell^2} \\
&\lesssim \sum_{s\,:\,Q\leqslant 2^s} s \cdot \max_{2^s\leqslant q<2^{s+1}} \frac{1}{\phi(q)} \cdot \|f\|_{\ell^2} \lesssim \frac{\log Q \cdot \log\log Q}{Q} \|f\|_{\ell^2}.
\end{aligned}$$

□

Lemma 4.38. *Let $\theta_1, \dots, \theta_J$ be points in $\mathbb{T}$ with $\min_{j\neq k}|\theta_j - \theta_k| > 2^{-2s_0+2}$. We have the inequality*

$$\Big\| \sup_{N>4^{s_0}} \Big| \sum_{j=1}^{J} \mathcal{F}^{-1}\Big(\widehat{f} \sum_{j=1}^{J} \tilde{M}_N(\cdot - \theta_j)\eta_{s_0}(\cdot - a/q)\Big) \Big| \Big\|_{\ell^2} \lesssim \log J \cdot \|f\|_{\ell^2}.$$

This is one of the main results of [42]. It is stated therein with a higher power of $\log J$. But it is well known that the inequality holds with a single power of $\log J$. This is discussed in detail in [60].

4.4.2 The Low Terms

From the Low terms defined in (4.28), the property is

Lemma 4.39. *For a functions f, g supported on interval I of length $N = 2^n$, we have*

$$N^{-1}\langle \mathrm{Lo}_{Q,N} * f, g\rangle \lesssim \log Q \cdot \langle f\rangle_I \langle g\rangle_I. \tag{4.40}$$

The following Möbius Lemma is well known.

Lemma 4.41. *For each q, we have*

$$\sum_{a\in A_q} G(\mathbf{1}_{A_q}, a)\mathcal{F}^{-1}(\widehat{M}_N \cdot \eta_s(\cdot - a/q))(x) = \frac{\mu(q)}{\phi(q)} c_q(-x). \tag{4.42}$$

Proof. Compute

$$\sum_{a\in A_q} G(\mathbf{1}_{A_q}, a)\mathcal{F}^{-1}(\widehat{M}_N \cdot \eta_s(\cdot - a/q))(x) = M_N * \mathcal{F}^{-1}\eta_s(x) \sum_{a\in A_q} G(\mathbf{1}_{A_q}, a) e(ax/q).$$

We focus on the last sum above, namely

$$\begin{aligned} S_q(x) &= \sum_{a\in A_q} G(\mathbf{1}_q, a) e(xa/q) \\ &= \frac{1}{\phi(q)} \sum_{r\in A_q} \sum_{a\in A_q} e(a(r+x)/q) \\ &= \frac{1}{\phi(q)} \sum_{r\in A_q} c_q(r+x) = \frac{\mu(q)}{\phi(q)} c_q(-x). \end{aligned} \tag{4.43}$$

The last line uses Cohen's identity. □

The two steps of inserting of the property of being Q smooth in (4.25), as well as dropping an restriction on s in (4.28), were made for this proof.

Proof of Lemma 4.39. By (4.42), the kernel of the operator $\mathrm{Lo}_{Q,N}$ is

$$\begin{aligned} \mathrm{Lo}_{Q,N}(x) &= M_N * \mathcal{F}^{-1}\eta_s(x) \cdot S(-x), \\ \text{where} \quad S(x) &= \sum_{q\in \mathbb{S}_Q} \frac{\mu(q)}{\phi(q)} c_q(x). \end{aligned} \tag{4.44}$$

We establish a pointwise bound $\|S\|_{\ell^\infty} \lesssim \log Q$, which proves the Lemma.

Assume $x \neq 0$. We exploit the multiplicative properties of the summands, as well as the fact that if prime p divides x, we have $\frac{\mu_p(x)}{\phi(p)} c_q(x) = \mu_p(x)$. Let $\mathcal{Q}_1$ be the primes $p < Q$ such that $(p, x) = 1$, and set $\mathcal{Q}_2$ to be the primes less than Q which are not in $\mathcal{Q}_1$.

The multiplicative aspect of the sums allows us to write

$$\frac{\mu(q)}{\phi(q)} c_q(-x) = \frac{\mu(q_1)}{\phi(q_1)} c_{q_1}(-x) \cdot \mu(q_2)$$

where $q = q_1 q_2$, and all prime factors of q_j are in $\mathcal{Q}_j$. If $\mathcal{Q}_j$ is empty, set $q_j = 1$. Thus, $S(x) = S_1(x) S_2(x)$, where the two terms are associated with $\mathcal{Q}_1$ and $\mathcal{Q}_2$ respectively. We have

$$\begin{aligned} S_1(x) &= \sum_{q \text{ is } \mathcal{Q}_1 \text{ smooth}} \frac{\mu(q)}{\phi(q)} c_q(-x) \\ &= \prod_{p \in \mathcal{Q}_1} 1 + \frac{\mu(p) c_p(-x)}{\phi(p)} \\ &= \prod_{p \in \mathcal{Q}_1} 1 + \frac{1}{p-1} = A_x. \end{aligned}$$

This is so, since $\mu(p) c_p(x) = 1$. It is a straight forward consequence of the Prime Number Theorem that $A_x \lesssim \log Q$. Here, and below, we say that q is $\mathcal{Q}$ smooth if all the prime factors of q are in the set of primes $\mathcal{Q}$.

The second term is as below, where $d = |\mathcal{Q}_2|$. Here, in the definition (4.28), there is no restriction on s, hence all the smooth square free numbers are included. If $\mathcal{Q}_2 = \emptyset$, then

$S_2(x) = 1$, otherwise

$$\begin{aligned} S_2(x) &= \sum_{q \text{ is } \mathcal{Q}_2 \text{ smooth}} \mu(q) \\ &= \sum_{j=1}^{d} \binom{d}{j} (-1)^j \\ &= -1 + \sum_{j=0}^{d} \binom{d}{j} (-1)^j = -1. \end{aligned}$$

If $x = 0$, then $S(0) = S_2(x) = -1$. That completes the proof.

□

4.4.3 The Exceptional Term

The Exceptional terms are always of a smaller order than the Low terms.

Lemma 4.45. *Let χ be an exceptional character modulo q. For $x \in \mathbb{Z}$,*

$$\Big| \sum_{a \in A_q} G(\chi, a) e(xa/q) \Big| = \frac{q}{\phi(q)} \tag{4.46}$$

provided $(x, q) = 1$, otherwise the sum is zero.

Proof. It is also known that exceptional characters are primitive - see [64, Theorem 5.27]. So the sum is zero if $(x, q) > 1$. We use the common notation

$$\tau(\chi, x) = \sum_{a \in A_q} \chi(a) e(ax/q)$$

which is $\phi(q) G(\chi, x)$. Assuming $(x, q) = 1$,

$$\tau(\chi, a) = \tau(\chi, 1).$$

This leads immediately to

$$\sum_{a\in A_q} \tau(\chi,a)e(\frac{ax}{q}) = \tau(\chi,1)\sum_{a\in A_q}\chi(a)e(-\frac{ax}{q})$$
$$= \frac{\tau(\chi)\overline{\tau(\chi,x)}}{\phi(q)} = \frac{|\tau(\chi)|^2\overline{\chi(x)}}{\phi(q)}.$$

It is known that $|\tau(\chi)|^2 = q$ for primitive characters. And the exceptional character is quadratic, so this completes the proof. □

Lemma 4.47. *For a function f supported on interval I of length $N = 2^n$, we have*

$$\langle \mathrm{Ex}_{Q,N} * f\rangle_\infty \lesssim (\log\log Q)^2 \cdot \langle f\rangle_I. \tag{4.48}$$

The term on the left is defined in (4.30).

Proof. Following the argument from Lemma 4.39, we have

$$\mathrm{Ex}_{Q,N}(x) = \sum_{q<Q}\sum_{a\in A_q} G(\chi_q,a)e(xa/q)\cdot M_N^{\beta_v} * \mathcal{F}^{-1}\eta_{s_q}(x).$$

Above, $2^{s_q} \leqslant q < 2^{s_q+1}$. The interior sum above is estimated in (4.46). Using the lower bound on the totient function in (4.8), we have

$$\mathrm{Ex}_{Q,N}(x)f \lesssim \log\log Q \cdot \langle f\rangle_I \sum_{\substack{q<Q\\ q \text{ exceptional}}} 1.$$

We know that the exceptional q grow at the rate of a double exponential, that is for q_v being the vth exceptional q, we have $q_v \gg C^{C^v}$, for some $C > 1$. It follows that the sum above is at most $\log\log Q$. □

4.5 Proofs of the Fixed Scale and Sparse Bounds

Proof of Theorem 4.4. Let $N = 2^n$, and recall that $f = \mathbf{1}_F$ and $g = \mathbf{1}_G$ where $F, G \subset I$, and interval of length N.

Let us address the case in which we do not assume GRH. We always have the estimate

$$N^{-1}\langle A_N f, g\rangle \lesssim n \cdot \langle f\rangle_I \langle g\rangle_I. \tag{4.49}$$

Hence, if we have $\langle f\rangle_I \langle g\rangle_I \lesssim e^{-c\sqrt{n}/100}$, the inequality with a squared log follows.

We assume that $e^{-c\sqrt{n}} \lesssim \langle f\rangle_I \langle g\rangle_I$, and then prove a better estimate. We turn to the Low/High/Exceptional decomposition in (4.28)—(4.30), for a choice of integer Q that we will specify. We have

$$A_N = \mathrm{Lo}_{Q,N} + \mathrm{Hi}_{Q,N} - \mathrm{Ex}_{Q,N} + \mathrm{Err}_N + \mathrm{Err}'_N \tag{4.50}$$

These terms are defined (4.28), (4.29), (4.30), (4.24) and (4.34) df respectively.

For the 'High' term we have by (4.36),

$$N^{-1}|\langle \mathrm{Hi}_{Q,N}\, f, g\rangle| \lesssim \frac{\log\log Q}{Q} \langle f\rangle_{I,2} \langle g\rangle_{I,2}$$

The same inequality holds for both $\mathrm{Err}_{Q,N}\, f$ and $\mathrm{Err}'_{Q,N}\, f$ by Lemma 4.23 and Proposition 4.31.

Concering the Low term, by (4.40), we have

$$N^{-1}|\langle \mathrm{Lo}_{Q,N}\, f, g\rangle| \lesssim \log Q \langle f\rangle_I \langle g\rangle_I$$

The Exceptional term satisfies the same estimate by (4.48).

Combining estimates, choose Q to minimize the right hand side, namely

$$N^{-1}\langle A_N f, g\rangle \lesssim \frac{\log\log Q}{Q}\big[\langle f\rangle_I \langle g\rangle_I\big]^{1/2} + \log Q \cdot \langle f\rangle_I \langle g\rangle_I. \tag{4.51}$$

This value of Q is

$$Q\frac{\log Q}{\log\log Q} \simeq \big[\langle f\rangle_I \langle g\rangle_I\big]^{-1/2}.$$

Since $e^{-c\sqrt{n}} \lesssim \langle f\rangle_I \langle g\rangle_I$, this is an allowed choice of Q. And, then, we prove the desired inequality, but only need a single power of logarithm.

Assuming GRH, from (4.49), we see that the inequality to prove is always true provided $\langle f\rangle_I \langle g\rangle_I < cN^{-1/4}$. Assuming this inequality fails, we follow the same line of reasoning above that leads to (4.51). That value of Q will be at most $N^{1/4}$, so the proof will complete, to show the bound with a single power of the logarithmic term.

□

Turning to the sparse bounds, let us begin with the definitions.

Definition 4.52. *A collection of intervals $\mathcal{S}$ is called* sparse *if to each interval $I \in \mathcal{S}$, there is a set $E_I \subset I$ so that $4|E_I| \geqslant |I|$ and the collection $\{E_I \ :\ I \in \mathcal{S}\}$ are pairwise disjoint. All intervals will be finite sets of consecutive integers in $\mathbb{Z}$.*

The form of the sparse bound in Theorem 4.6 strongly suggests that one use a recursive method of proof. (Which is indeed the common method.) To formalize it, we start with the notion of a *linearized* maximal function. Namely, to bound the maximal function $A^* f$, it suffices to bound $A_{\tau(x)} f(x)$, where $\tau \ :\ \mathbb{Z} \to \{2^n \ :\ n \in \mathbb{N}\}$ is a function, taken to realize the supremum. The supremum in the definition of $A^* f$ is always attained if f is finitely supported.

Definition 4.53. *Let I_0 an interval, and let f be supported on $3I_0$. A map $\tau \ :\ I_0 \to$*

$\{1, 2, 4, \ldots, |I_0|\}$ *is said to be* admissible *if*

$$\sup_{N \geqslant \tau(x)} M_N f(x) \leqslant 10\langle f \rangle_{3I_0,1}.$$

That is, τ is admissible if at all locations x, the averages of f over scales larger than $\tau(x)$ are controlled by the global average of f.

Lemma 4.54. *Let f and τ be as in Definition 4.53. Further assume that f and g are indicator functions, with g supported on I_0. Then, we have*

$$|I_0|^{-1}\langle A_\tau f, g\rangle \lesssim \langle f\rangle_{I_0,1}\langle g\rangle_{I_0,1} \cdot (\mathrm{Log}\langle f\rangle_{3I_0,1}\langle g\rangle_{I_0,1})^t, \tag{4.55}$$

where $t = 1$ assuming RH, and $t = 2$ otherwise.

Proof. We restrict τ to take values $1, 2, 4, \ldots, 2^t, \ldots,$. Let $|I_0| = N_0 = 2^{n_0}$. We always have the inequalities

$$|I_0|^{-1}\langle A_\tau f, g\rangle \lesssim n_0\langle f\rangle_{I_0,1}\langle g\rangle_{I_0,1}$$
$$|I_0|^{-1}\langle \mathbf{1}_{\tau<T} A_\tau f, g\rangle \lesssim (\log T)\langle f\rangle_{I_0,1}\langle g\rangle_{I_0,1}.$$

The top line follows from admissibility.

We begin by not assuming GRH. Then, the conclusion of the Lemma is immediate if we have $(\mathrm{Log}\langle f\rangle_{I_0,1}\langle g\rangle_{I_0,1})^2 \gg n_0$. It is also immediate if $\log\tau \lesssim (\mathrm{Log}\langle f\rangle_{I_0,1}\langle g\rangle_{I_0,1})^2$. We proceed assuming

$$p_0^2 = C(\mathrm{Log}\langle f\rangle_{I_0,1}\langle g\rangle_{I_0,1})^2 \leqslant c_0 \min\{n_0, \log\tau\}, \tag{4.56}$$

where $0 < c_0 < 1$ is sufficiently small.

We use the definitions in (4.28)—(4.30) for a value of $Q < e^{c\sqrt{n_0}}$ that we will specify. We address the High, Low, Exceptional and both Error terms, as in (4.50). First, the Error terms. The error terms come in the form of Err_N from Lemma 4.23 and Err'_N from (4.33). Both are similar. Concerning the second error term, from the estimate (4.34) and (4.56), we have by a straight forward square function argument,

$$\begin{aligned}
\|\operatorname{Err}_{Q,\tau} f\|_2^2 &\leqslant \sum_{n\,:\,p_0^2\leqslant n\leqslant n_0} \|\operatorname{Err}_{Q,2^n} f\|_{\ell^2}^2 \\
&\lesssim \|f\|_{\ell^2}^2 \sum_{n\,:\,p_0^2\leqslant n\leqslant n_0} e^{-c\sqrt{n}} \\
&\lesssim \|f\|_{\ell^2}^2 \cdot p_0^2 e^{-cp_0} \lesssim \|f\|_{\ell^2}^2 \cdot \langle f\rangle_{3I_0,1}\langle g\rangle_{I_0,1}.
\end{aligned}$$

This provided C in (4.56) is large enough. This is a much smaller estimate than we need. The second error term in Proposition 4.31 is addressed by the same square function argument.

For the High term, apply (4.37) to see that

$$\| \sup_{N>Q^2} |\operatorname{Hi}_{Q,N} f| \|_2 \lesssim \frac{\log Q\cdot \log\log Q}{Q} \|f\|_{\ell^2}. \tag{4.57}$$

For the Low term the definition of admissibility and (4.40) that

$$|I_0|^{-1} |\langle \operatorname{Lo}_{Q,\tau(x)} f(x), g\rangle \lesssim (\log Q)\langle f\rangle_I \langle g\rangle_I.$$

The Exceptional term also satisfies this bound.

We conclude that

$$|I_0|^{-1}\langle A_\tau f, g\rangle \lesssim \frac{\log Q\cdot \log\log Q}{Q}\langle f\rangle_{I,2}\langle g\rangle_{I,2} + \log Q\cdot \langle f\rangle_I\langle g\rangle_I.$$

This is optimized by taking Q so that

$$\frac{Q}{\log\log Q} \simeq \big[\langle f\rangle_I \langle g\rangle_I\big]^{-1/2}.$$

And this will be an allowed value of Q since (4.56) holds. Again, the resulting estimate is better by power of the logarithmic term than what is claimed.

Under RH, the proof is very similar, but a wider range of Q's are allowed. In particular, only a single power of logarithm is needed.

$\square$

REFERENCES

[1] A. Granville, "Smooth numbers: Computational number theory and beyond," *Algorithmic number theory: lattices, number fields, curves and cryptography*, vol. 44, pp. 267–323, 2008.

[2] K. Dickman, "On the frequency of numbers containing prime factors of a certain relative magnitude," *Arkiv for matematik, astronomi och fysik*, vol. 22, no. 10, A–10, 1930.

[3] A. Hildebrand, "Integers free of large prime factors and the riemann hypothesis," *Mathematika*, vol. 31, no. 2, pp. 258–271, 1984.

[4] E. Croot, "Smooth numbers in short intervals," *International Journal of Number Theory*, vol. 3, no. 01, pp. 159–169, 2007.

[5] K. Matomäki, "Another note on smooth numbers in short intervals," *International Journal of Number Theory*, vol. 12, no. 02, pp. 323–340, 2016.

[6] K. Soundararajan, "Smooth numbers in short intervals," *arXiv preprint arXiv:1009.1591*, 2010.

[7] B. C. Berndt, *Number theory in the spirit of Ramanujan*. American Mathematical Soc., 2006, vol. 34.

[8] L.-k. Hua, "On the number of partitions of a number into unequal parts," *Transactions of the American Mathematical Society*, vol. 51, no. 1, pp. 194–201, 1942.

[9] S. Iseki, "A partition function with some congruence condition," *American Journal of Mathematics*, vol. 81, no. 4, pp. 939–961, 1959.

[10] S. Raghavendran and V. Narayanan, "The prouhet tarry escott problem: A review," *Mathematics*, vol. 7, no. 3, p. 227, 2019.

[11] A. Baker, "On the class number of imaginary quadratic fields," *Bulletin of the American Mathematical Society*, vol. 77, no. 5, pp. 678–684, 1971.

[12] ——, "Linear forms in the logarithms of algebraic numbers (ii)," *Mathematika*, vol. 14, no. 1, pp. 102–107, 1967.

[13] ——, "Linear forms in the logarithms of algebraic numbers (iii)," *Mathematika*, vol. 14, no. 2, pp. 220–228, 1967.

[14] J. Kirby and B. Zilber, “The uniform schanuel conjecture over the real numbers,” *Bulletin of the London Mathematical Society*, vol. 38, no. 4, pp. 568–570, 2006.

[15] C. Bertolin, “Périodes de 1-motifs et transcendance,” *Journal of Number Theory*, vol. 97, no. 2, pp. 204–221, 2002.

[16] E. Croot, H. Mousavi, and M. Schmidt, “On a conjecture of graham on the p-divisibility of central binomial coefficients,” *arXiv preprint arXiv:2201.11274*, 2022.

[17] E. Croot and H. Mousavi, “On a class of sums with unexpectedly high cancellation, and its applications,” *arXiv preprint arXiv:1909.12470*, 2019.

[18] M. T. Lacey, H. Mousavi, and Y. Rahimi, “Endpoint ℓ^r improving estimates for prime averages,” *arXiv preprint arXiv:2101.10401*, 2021.

[19] C. Giannitsi, M. T. Lacey, H. Mousavi, and Y. Rahimi, “Improving and maximal inequalities for primes in progressions,” *Banach Journal of Mathematical Analysis*, vol. 16, no. 3, pp. 1–25, 2022.

[20] D. Berend and J. E. Harmse, “On some arithmetical properties of middle binomial coefficients,” *Acta Arithmetica*, vol. 84, no. 1, pp. 31–41, 1998.

[21] N. J. Sloane, “The on-line encyclopedia of integer sequences, published electronically at http://www. research. att. com/ njas/sequences,” *Sequence A000372: Dedekind numbers: number of monotone Boolean functions of n variables or number of antichains of subsets of an n-set*, 2009.

[22] E. E. Kummer, “Über die ergänzungssätze zu den allgemeinen reciprocitätsgesetzen.,” 1852.

[23] P. Erdös, R. Graham, I. Ruzsa, and E. Straus, “On the prime factors of $\binom{2n}{n}$,” *Mathematics of Computation*, pp. 83–92, 1975.

[24] C. Sanna, “Central binomial coefficients divisible by or coprime to their indices,” *International Journal of Number Theory*, vol. 14, no. 04, pp. 1135–1141, 2018.

[25] K. Ford and S. Konyagin, “Divisibility of the central binomial coefficient,” *Transactions of the American Mathematical Society*, vol. 374, no. 2, pp. 923–953, 2021.

[26] A. Granville and O. Ramaré, “Explicit bounds on exponential sums and the scarcity of squarefree binomial coefficients,” *Mathematika*, vol. 43, no. 1, pp. 73–107, 1996.

[27] C. Pomerance, “Divisors of the middle binomial coefficient,” *The American Mathematical Monthly*, vol. 122, no. 7, pp. 636–644, 2015.

[28] S. Lang, *Introduction to transcendental numbers*. Addison-Wesley Publishing Company, 1966.

[29] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*. Courier Corporation, 2012.

[30] F. R. Gantmacher, "The theory of matrices, chelsea pub," *Co., New York*, vol. 2, 1959.

[31] S. DeSalvo and I. Pak, "Log-concavity of the partition function," *The Ramanujan Journal*, vol. 38, no. 1, pp. 61–73, 2015.

[32] J. Lehner, "A partition function connected with the modulus five," *Duke Mathematical Journal*, vol. 8, no. 4, pp. 631–655, 1941.

[33] R. W. Butler and A. T. Wood, "Saddlepoint approximation for moment generating functions of truncated random variables," *The Annals of Statistics*, vol. 32, no. 6, pp. 2712–2730, 2004.

[34] J.-h. Xu and J.-h. Zhao, "Moment generating functions of random variables and asymptotic behaviour for generalized feller operators," *Journal of Computational Mathematics*, pp. 173–182, 2000.

[35] L. B. Pierce, "The vinogradov mean value theorem [after wooley, and bourgain, demeter and guth]," *arXiv preprint arXiv:1707.00119*, 2017.

[36] *Unsolved problems and conjectures*, http://eslpower.org/unsolve.htm, Accessed: 2010-09-30.

[37] *254a, notes 5: Bounding exponential sums and the zeta function*, https://terrytao.wordpress.com/2015/02/07/254a-notes-5-bounding-exponential-sums-and-the-zeta-function/, Accessed: 2010-09-30.

[38] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, 84. American Mathematical Soc., 1994.

[39] H. L. Montgomery and R. C. Vaughan, "The large sieve," *Mathematika*, vol. 20, no. 2, pp. 119–134, 1973.

[40] D. Frey and Z. Nieraeth, "Weak and strong type $A_1 - A_\infty$ estimates for sparsely dominated operators," *The Journal of Geometric Analysis*, vol. 29, no. 1, pp. 247–282, 2019.

[41] J. Bourgain, "On the pointwise ergodic theorem on L^p for arithmetic sets," *Israel J. Math.*, vol. 61, no. 1, pp. 73–84, 1988.

[42] J. Bourgain, "Pointwise ergodic theorems for arithmetic sets," *Inst. Hautes Études Sci. Publ. Math.*, no. 69, pp. 5–45, 1989, With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein.

[43] A. D. Ionescu and S. Wainger, "L^p boundedness of discrete singular Radon transforms," *J. Amer. Math. Soc.*, vol. 19, no. 2, pp. 357–383, 2006.

[44] M. Mirek, "Square function estimates for discrete Radon transforms," *Anal. PDE*, vol. 11, no. 3, pp. 583–608, 2018.

[45] T. Tao, "The Ionescu–Wainger multiplier theorem and the adeles," *arXiv e-prints*, arXiv:2008.05066, arXiv:2008.05066, Aug. 2020. arXiv: 2008.05066 `[math.CA]`.

[46] L. B. Pierce, "On superorthogonality," *arXiv e-prints*, arXiv:2007.10249, arXiv:2007.10249, Jul. 2020. arXiv: 2007.10249 `[math.CA]`.

[47] M. Wierdl, "Pointwise ergodic theorem along the prime numbers," *Israel J. Math.*, vol. 64, no. 3, 315–336 (1989), 1988.

[48] M. Mirek and B. Trojan, "Cotlar's ergodic theorem along the prime numbers," *J. Fourier Anal. Appl.*, vol. 21, no. 4, pp. 822–848, 2015.

[49] B. Trojan, "Endpoint estimates for the maximal function over prime numbers," *Journal of Fourier Analysis and Applications*, vol. 25, no. 6, pp. 3123–3153, 2019.

[50] M. C. Pereyra, "Dyadic harmonic analysis and weighted inequalities: The sparse revolution," in *New Trends in Applied Harmonic Analysis, Volume 2*, Springer, 2019, pp. 159–239.

[51] A. Culiuc, R. Kesler, and M. T. Lacey, "Sparse bounds for the discrete cubic Hilbert transform," *Anal. PDE*, vol. 12, no. 5, pp. 1259–1272, 2019.

[52] R. Kesler and M. T. Lacey, "ℓ^p-Improving Inequalities for Discrete Spherical Averages," *Anal. Math.*, vol. 46, no. 1, pp. 85–95, 2020.

[53] K. Hughes, "ℓ^p-improving for discrete spherical averages," *Ann. H. Lebesgue*, vol. 3, pp. 959–980, 2020.

[54] A. Magyar, E. M. Stein, and S. Wainger, "Discrete analogues in harmonic analysis: Spherical averages," *Ann. of Math. (2)*, vol. 155, no. 1, pp. 189–208, 2002.

[55] R. Kesler, M. T. Lacey, and D. Mena, "Sparse bounds for the discrete spherical maximal functions," *Pure Appl. Anal.*, vol. 2, no. 1, pp. 75–92, 2020.

[56] R. Han, M. T. Lacey, and F. Yang, “Averages along the Square Integers: ℓ^p improving and Sparse Inequalities,” *arXiv e-prints*, arXiv:1907.05734, arXiv:1907.05734, Jul. 2019. arXiv: 1907.05734 [math.CA].

[57] S. Dasu, C. Demeter, and B. Langowski, “Sharp l^p-improving estimates for the discrete paraboloid,” *arXiv e-prints*, arXiv:2002.11758, arXiv:2002.11758, Feb. 2020. arXiv: 2002.11758 [math.CA].

[58] R. Han, V. Kovač, M. T. Lacey, J. Madrid, and F. Yang, “Improving estimates for discrete polynomial averages,” *J. Fourier Anal. Appl.*, vol. 26, no. 3, Paper No. 42, 11, 2020.

[59] B. Krause, M. Mirek, and T. Tao, “Pointwise ergodic theorems for non-conventional bilinear polynomial averages,” *arXiv e-prints*, arXiv:2008.00857, arXiv:2008.00857, Aug. 2020. arXiv: 2008.00857 [math.DS].

[60] R. Han, B. Krause, M. T. Lacey, and F. Yang, “Averages along the primes: Improving and sparse bounds,” *Concr. Oper.*, vol. 7, no. 1, pp. 45–54, 2020.

[61] M. Mirek, “Roth’s theorem in the Piatetski-Shapiro primes,” *Rev. Mat. Iberoam.*, vol. 31, no. 2, pp. 617–656, 2015.

[62] J. Bourgain, “Estimations de certaines fonctions maximales,” *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 301, no. 10, pp. 499–502, 1985.

[63] A. D. Ionescu, “An endpoint estimate for the discrete spherical maximal function,” *Proc. Amer. Math. Soc.*, vol. 132, no. 5, pp. 1411–1417, 2004.

[64] H. Iwaniec, “E. kowalski analytic number theory,” in *American Mathematical Society Colloquium Publications*, vol. 53, 2004.

[65] I. M. Vinogradov, *Elements of number theory*. Dover Publications, Inc., New York, 1954, pp. viii+227, Translated by S. Kravetz.