Université de Montréal

Distribution of sums of the Legendre Symbol

par

Sana Mehkari

Département de mathématiques et de statistique

Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures

en vue de l'obtention du grade de

Maître ès sciences (M.Sc.)
en mathématiques

juillet 2005

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

Distribution of sums of the Legendre Symbol

présenté par

Sana Mehkari

a été évalué par un jury composé des personnes suivantes :

Chantal David

(président-rapporteur)

Andrew Granville

(directeur de recherche)

Habiba Kadiri

(membre du jury)

Mémoire accepté le:

le 14 juillet 2005

ABSTRACT

In this thesis we study the sum of quadratic residues (modulo a prime p), of length px where $0 < x < 1$. In 1974 Hugh L. Montgomery studied the distribution of this sum for primes $p \equiv 3 \bmod 4$. More precisely, he modelled the Legendre symbol by a totally multiplicative arithmetic function and obtained estimates for the proportion of $x \in [0, 1]$ for which the sum is positive. We give a different approach to this problem by using a different model and applying techniques from analytic number theory and combinatorics. We also perform a series of numerical computations. Unlike Montgomery's results, ours assume the Generalized Riemann Hypothesis. In spite this additional condition, we are unable to improve Montgomery's result. However, the results obtained via our method suggest that Montgomery's estimate is probably the best possible. We note that our ideas can be used to determine similar results for primes $p \equiv 1 \bmod 4$.

Key Words

Analytic Number Theory, Legendre Symbols.

RÉSUMÉ

Dans ce mémoire, on étudie la somme de résidus quadratiques (modulo un premier p), de longueur px où $0 < x < 1$. En 1974, Hugh L. Montgomery a étudié la distribution de cette somme pour les premiers $p \equiv 3 \bmod 4$. Plus précisément, il a modélisé le symbole de Legendre par une fonction arithmétique qui est totalement multiplicative et a obtenu des estimations pour la proportion des $x \in [0, 1]$ pour lesquels la somme est positive. On donne une approche différente pour ce problème en utilisant un modèle différent et en appliquant des techniques de théorie des nombres analytique et combinatoire. On fait aussi une série de calculs numériques. Contrairement aux résultats de Montgomery, notre résultat suppose l'Hypothèse de Riemann Generalisée. Malgré cette condition supplémentaire, on ne peut pas améliorer le résultat de Montgomery. Cependant, les résultats obtenus par notre méthode suggérent que l'estimation de Montgomery est probablement la meilleur possible. On note que les mêmes idées peuvent être utilisées pour déterminer des resultats similaires pour les premiers $p \equiv 1 \bmod 4$.

Mots Clefs

Théorie des nombres analytique, symbole de Legendre.

ACKNOWLEDGEMENTS

I would like to begin by thanking my Director Andrew Granville, who firstly introduced me to the wonderful world of Analytic Number Theory and then with utmost patience guided me to and through this particular masters thesis. His patience, availability, numerous ideas and encouragement have made him an exceptional director and it was a truly gratifying experience to work under his supervision.

Being a project that was heavily dependent on a multitude of computations, I am obliged to extend my gratitude towards Rony Touma and Francis Forget for all their technical help and support. This thesis could never have been completed without the unselfish support of my friends Habiba Kadiri, Nathan Ng and Youness Lamzouri who were not only always willing to help me at any theoretical and technical junctures but were also most encouraging throughout. I owe my gratitude to my office mates Alexandre Girouard, Nicolas Beauchemin and Amel Kaouche for their patience in bearing with my moods and for providing me with any technical and moral support. I would also like to especially thank Amna, Fawaz, Fouzia, Madiha, Rabi, Saif, Salman and Zahra for the constant encouragement and support that they extended. My list of acknowledgements would be incomplete without the mention of the mathematics department of the Université de Montréal, which provided me the opportunity to study and produce such work. And finally, I am most grateful and indebted to my parents, family and friends who despite being far away, were always there to encourage me, support me, hear me out and make sure that I kept on going.

It is due to all these inspiring and selfless people that I was able to complete this thesis and produce the following work. Thank you all!

CONTENTS

INTRODUCTION

In the 18th century, mathematicians were interested in knowing which primes p are squares modulo a given prime q. Two famous mathematicians of the time, Leonhard Euler and Adrien-Marie Legendre observed that the answer seemed to depend only on the "reciprocal" property (see [**10**]) i.e. whether q is a square (mod p). More precisely, they conjectured that when both p and q are primes congruent to 3 (mod 4), p is a square (mod q) if and only if q is not a square (mod p), and otherwise p is a square (mod q) if and only if q is also a square (mod p). This is what became the much celebrated *Law of Quadratic Reciprocity*, which now serves as a powerful tool to attack problems of quadratic congruence. In 1785, Legendre proposed a proof for the law which depended on the then unproved assumption that any arithmetic progression $an + b$, $gcd(a, b) = 1$, $n = 1, 2, 3 \ldots$ has infinitely many primes. Later in 1796, at the age of 18, Carl Friedrich Gauss independently discovered the law and gave its first complete proof which was independent of Legendre's assumption. Legendre's additional contribution to the field came in 1798 with the introduction of a symbol to represent an integer's quadratic character with respect to a given prime p. This symbol denoted as $\left(\frac{n}{p}\right)$ equals 1 if n is a quadratic residue modulo the prime p and -1 if n is a quadratic non-residue modulo the prime p. It is now widely used and commonly referred to as the Legendre symbol. Thus, the 18th century laid the foundations and marked only the beginning of an entire area of study of quadratic residues and non-residues, questions concerning which intrigue many mathematicians even today.

Several questions have been proposed concerning the distribution of quadratic residues and non-residues, but as yet only very incomplete answers are known. By simple congruence arguments, it is easy to show that for any given odd prime p, there exist exactly $(p-1)/2$ quadratic residues $\pmod p$ and $(p-1)/2$ quadratic non-residues $\pmod p$. However, the question remains, how does one know where exactly (between 1 and $p-1$ $\pmod p$) these quadratic residues and quadratic non-residues occur. Many people are interested in finding the size of the smallest quadratic non-residue modulo a large prime p. S.W. Graham and C.J. Ringrose address this question in [**13**].

Based on the above observation of an equal number of quadratic residues and non-residues (in $[1, p-1]$) modulo p, which appear to be fairly randomly distributed, Gauss recognized a rather intriguing question involving the character of consecutive numbers. He proved the expected, that each pair of consecutive numbers RR, NN, NR, RN (where R represents a quadratic residue and N a quadratic non-residue) occurs about equally often. Many mathematicians have continued study in this area whereby H. Davenport [**7**] gave upper bounds for RRR and NNN in 1931. In fact, we know that for any pattern $RN\ldots$ of length k, the number of occurrences of that pattern is $\frac{p}{2^k} + O_k(\sqrt{p})$. Some of the other work in this area includes that of A. Brauer [**3**], R.H. Hudson [**20**], Z.H. Sun [**31**] and R. Peralta [**26**].

In trying to understand the distribution of quadratic residues and non-residues, many mathematicians have appealed to the well known Pólya-Vinogradov inequality of 1918. This result is central to analytic number theory and gives an estimate for the sum

$$\sum_{n \le x} \chi(n) \tag{0.0.1}$$

where χ is a non-principal Dirichlet character modulo q. The Legendre symbol $\left(\frac{n}{q}\right)$, being a quadratic character modulo q, then becomes a particular case of the above. It is clear that such a sum is always less than q in absolute value, however G. Pólya [**27**] and I.M. Vinogradov [**32**] proved independently that such

a character sum is $\ll \sqrt{q}\log q$. This was the first significant estimate in this area and apart from the implied constant, the inequality was not improved until very recently when in [**16**], A. Granville and K. Soundrarajan showed that the power of the log term can be improved for primitive characters of odd, bounded order. However, for all non-principal characters in general, the Pólya-Vinogradov inequality is close to best possible, for in 1918, I. Schur [**30**] proved that for all primitive characters $\chi \pmod q$,

$$\max_x \Big|\sum_{n\le x}\chi(n)\Big| > \frac{1}{2\pi}\sqrt{q}.$$

Then, in 1932 R.E.A.C. Paley [**25**] showed that

$$\max_x \Big|\sum_{n\le x}\Big(\frac{d}{n}\Big)\Big| > \frac{1}{7}\sqrt{d}\log\log d$$

for infinitely many quadratic discriminants $d > 0$. In the opposite direction, in 1977, assuming the Generalized Riemann Hypothesis (GRH), H.L. Montgomery and R.C. Vaughan [**24**] showed that

$$\sum_{n\le x}\chi(n) \ll \sqrt{q}\log\log q$$

for all non-principal characters $\pmod q$. In 1993, J.B. Friedlander and H. Iwaniec [**11**] gave a different proof of the Pólya-Vinogradov inequality.

Although the Pólya-Vinogradov inequality is close to being best possible, for many purposes it is useful to have an estimate which is sharper when x is small compared with q. This brings us to introduce some results of D.A. Burgess who has significantly contributed to the area of short character sums. In many applications, one is interested in when the character sum (0.0.1) is $o(x)$ with x substantially smaller than $q^{\frac{1}{2}+o(1)}$. In 1957, Burgess [**4**] showed

$$\Big|\sum_{n\le x}\chi(n)\Big| = o(x) \tag{0.0.2}$$

whenever $x > q^{\frac{1}{4}+o(1)}$, for any quadratic character $\pmod q$, with q prime. Via a series of papers (see [**5**] and [**6**]), Burgess later generalized his results to non-principal characters modulo k (where k is not necessarily prime) other than for

just the quadratic character. Later in 1994, J.B. Friedlander and H. Iwaniec [**12**] used a different approach to prove Burgess' result for short character sums. In 1986, A. Hilderbrand [**17**] observed that one can extrapolate Burgess' bound to the range $x > q^{\frac{1}{4}-o(1)}$. In 2001, A. Granville and K. Soundrarajan [**15**] also investigated in what range the estimate given in (0.0.2) should hold, and showed that assuming the Riemann Hypothesis for $L(s,\chi)$, it would hold if $\frac{\log x}{\log\log q} \to \infty$ as $q \to \infty$.

Burgess' result for quadratic characters modulo q (q prime) allowed him to improve Vinogradov's [**33**] estimate for the magnitude of the least (positive) quadratic non-residue $\pmod q$. Using the Pólya-Vinogradov inequality for Legendre symbols, Vinogradov had proved that this least quadratic non-residue is $O(q^{\alpha})$ (as $q \to \infty$) for any fixed $\alpha > \frac{1}{2}e^{-\frac{1}{2}}$. Burgess improved the range of α to $> \frac{1}{4}e^{-\frac{1}{2}}$.

Of the improvements done to the implied constant of the Pólya-Vinogradov inequality, the first was done by E. Landau [**22**] soon after the inequality was established. Subsequent improvements were made by P.T. Bateman and A. Hildebrand (see [**18**] and [**19**]). In 1988, Hildebrand [**19**] proved a more general estimate for the size of the character sum which, apart from yielding improved values for the constant, shows that the character sum becomes large only very rarely, and in fact only when x/q is close to a rational number with small denominator. The best known result for the size of the implied constant was given very recently by A. Granville and K. Soundrarajan [**16**].

Another interesting topic arose in the area of character sums when in 1974, H.L. Montgomery [**23**] took to studying the distribution of the sum

$$S_p(x) := \sum_{n \le px} \left(\frac{n}{p}\right) \tag{0.0.3}$$

where p is an odd prime and $\left(\frac{n}{p}\right)$ is the Legendre symbol. More specifically, he observed that this sum, when considered as a function of x, is even for primes $p \equiv 3 \pmod 4$ and for such primes, decided to study how frequently it is positive

for $x \in [0,1]$. He defined the measure

$$\alpha(p) = |\{x \in [0,1] : S_p(x) > 0\}| \quad \text{where } p \equiv 3 \pmod 4, \tag{0.0.4}$$

and proved

Theorem 0.0.1 (H.L. Montgomery). *For any $\delta > 0$ there are infinitely many primes $p \equiv 3 \bmod 4$ such that $\alpha(p) < \frac{1}{3} + \delta$. In the opposite direction, $\alpha(p) > \frac{1}{50}$ for all primes $p \equiv 3 \bmod 4$.*

Our work in this thesis is inspired by this study of Montgomery and our goal is to extend Montgomery's result to a positive proportion of primes $p \equiv 3 \bmod 4$ and perhaps to even improve the values of the bounds he obtains for $\alpha(p)$. As a result of our efforts, we prove:

Theorem 0.0.2. *Assume GRH. For a positive proportion of primes $p \equiv 3 \bmod 4$, $\alpha(p) \leq 0.746$ and of these, there is at least one prime for which $\alpha(p) \leq 0.341$. In the opposite direction, for a positive proportion of primes $p \equiv 3 \bmod 4$, $\alpha(p) > 0.285$ and of these, there is at least one prime for which $\alpha(p) > 0.631$.*

As is clear from the above theorem, our results are based on GRH (which gives rise to the assumption of a positive proportion of primes) and even then, we are unable to improve Montgomery's result. However, our approach used to arrive at Theorem 0.0.2, suggests that Montgomery's result is probably the best possible.

Our work in this thesis is geared towards proving Theorem 0.0.2. To build up some basics, in the first chapter we lay down some fundamental definitions, notation and standard theorems of analytic number theory. We then move towards our particular area of interest which concerns quadratic residues and character sums. In this respect, in Chapter 2 we introduce the central ideas concerning quadratic residues and non-residues, the Legendre symbol and formally state the Law of Quadratic Reciprocity. In the same chapter we also discuss questions concerning the distribution of the quadratic residues and non-residues and formally state the Pólya-Vinogradov Inequality, give its proof and then give D.A.Burgess'

result for the sum of quadratic characters. Subsequently, we move to a brief discussion about what the Pólya-Vinogradov Inequality and D.A.Burgess' result suggest about the possible Legendre symbol sequences. In this way, we develop a motivation for our specific interest in the distribution of the sum of Legendre symbols.

Having created a foundation for our work, next we state and prove H.L. Montgomery's 1974 result concerning the distribution of the sum of Legendre symbols $\left(\frac{\cdot}{p}\right)$, where p is a prime congruent to 3 (mod 4). A detailed proof is given in Chapter 3 where we establish his result by giving the proofs as furnished in his article, only that here we fill in all the details as well. In his work, Montgomery uses the Fourier expansion given by Pólya for the sum $S_p(x)$ which states that for primes $p \equiv 3 \bmod 4$,

$$S_p(x) \sim \frac{\sqrt{p}}{\pi}\left(L_p(1) - \sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}\right)$$

where $L_p(1) = \sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{1}{n}$. Owing to this Fourier expansion, Montgomery noticed that it was possible to bound the measure of the set of $x \in [0,1]$ for which $S_p(x)$ is positive, by the measure of the set of $x \in [0,1]$, for which the Fourier series $\sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$ is negative. Hence, he reduced the problem to studying the measure of the set of $x \in [0,1]$, for which this Fourier series is negative.

To obtain his lower bound result, Montgomery first computes the mean of the Fourier series associated to $S_p(x)$ and then then uses the Hölder's inequality to isolate the measure of $x \in [0,1]$ for which the Fourier series $\sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$ is negative. In doing this, he uses explicit values obtained for the upper bounds for the second and fourth moments of the Fourier series $\sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$ for all primes $p \equiv 3 \bmod 4$. Then, using the fact that $L_p(1)$ is positive for all primes p, a lower bound for $\alpha(p)$ directly follows. While studying Montgomery's proof, it was hard to ignore the fact that his lower bound result (which states that for all primes $p \equiv 3 \bmod 4$, the sum $S_p(x)$ is positive for at least 2 percent of the values of $x \in [0,1]$) could be improved by a simple refinement of his own proof. At the

end of Chapter 3, as a result of a minor observation, we give an improvement to this result of Montgomery, whereby we show that for all primes $p \equiv 3 \bmod 4$, the sum $S_p(x)$ is in fact positive for at least $\frac{5437}{250000}$ of the values of $x \in [0,1]$.

To obtain his upper bound result, Montgomery models the Legendre symbol by another Legendre symbol which we know explicitly i.e $\left(\frac{q}{3}\right)$. More precisely, for all primes $q \leq y$ (where y is a parameter), he modelled the Legendre symbol $\left(\frac{q}{p}\right)$ by $\left(\frac{q}{3}\right)$ for $q \neq 3$ and set $\left(\frac{q}{p}\right) = -1$ if $q = 3$. We observe that this model is close to reality since the second moments of the error is small (of size $1/y$). Montgomery then proved that given a fixed y, there exist infinitely many primes $p \equiv 3 \bmod 4$ such that for all $q \leq y$, $\left(\frac{q}{p}\right) = \left(\frac{q}{3}\right)$. We note here that we can perform numerical computations with $\left(\frac{q}{3}\right)$. Further, since his choice of function generated a specific $\{1,-1\}$ sequence, he was able to obtain an upper bound for the Fourier series corresponding to his modelling function, and in turn use this bound to obtain an upper bound for $\alpha(p)$.

Our work is reflected in the results of Theorem 0.0.2. The proof of this theorem is given in Chapter 5 and uses some preparatory results established in Chapter 4. Coming to an insight of the proof of Theorem 0.0.2 : we begin by employing the key idea of Montgomery, which was to reduce the problem to finding the corresponding measure related to the Fourier series $-\sum_{n=1}^{\infty} \lambda(n)\frac{\cos 2\pi nx}{n}$. Next, unsure if Montgomery's choice of model for the Legendre symbol ($\lambda(q) = \left(\frac{q}{3}\right)$ for all $q \leq y$) really is the best choice or not, we use a more "general" function. This "general" nature of our modelling function does not allow us to continue with the same approach as Montgomery. Hence, we employ a different approach whereby we average over a specific set of primes $p \equiv 3 \bmod 4$ and obtain bounds for $\alpha(p)$ that hold for a positive proportion of primes of this particular set. In doing so, we also appeal to combinatorial techniques and perform a series of numerical computations.

As in Montgomery's case, our modelling function λ is also totally multiplicative. We define $\lambda(q)$ to be equal to 1 or -1 for all primes $q \leq y$ (where y is a parameter) and equal to 0 for primes $q > y$. We realize that for a given y, this choice of a more "general" function would result in 2^m (where m is the number of primes $\leq y$) possibilities of $\{1, -1\}$ sequences. This in turn would involve studying 2^m $\{1, -1\}$ sequences to ultimately deduce the one giving the optimal result. (An explanation of what kind of result we are looking for follows later.) However, aware of the immense computing power predominant in today's world, we know that our vast data can be handled via a multitude of numerical computations, to finally yield a meaningful result. The numerical computations give us values for the truncated series of $-\sum_{\substack{n=1 \\ q|n \Rightarrow q \leq y}}^{\infty} \lambda(n)\frac{\cos 2\pi nx}{n}$, for all $x \in [0, 1]$, for all possible λ sequences. It is now only natural to somehow ensure some sort of proximity between the two Fourier series $-\sum_{n=1}^{\infty} \lambda(n)\frac{\cos 2\pi nx}{n}$ and $-\sum_{n=1}^{\infty} \left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$. This would allow us to deduce bounds for the measure relating to the Fourier series $-\sum_{n=1}^{\infty} \left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$, from bounds for the measure relating to the series $-\sum_{\substack{n=1 \\ q|n \Rightarrow q \leq y}}^{\infty} \lambda(n)\frac{\cos 2\pi nx}{n}$.

More specifically, the proof of Theorem 0.0.2 rests on two main theorems, each of which follow the same basic approach. For a fixed λ, y, the first theorem gives an explicit formula for a lower bound for the measure of the set of $x \in [0, 1]$ for which the Fourier series $-\sum_{n=1}^{\infty} \left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$ is negative. This explicit formula is dependent on the number of $0 \leq j \leq N-1$ for which the series $-\sum_{\substack{n=1 \\ q|n \Rightarrow q \leq y}}^{\infty} \lambda(n)\frac{\cos 2\pi n(j/N)}{n}$ is negative. The second theorem gives the analogous result for the upper bound of the concerned measure.
These theorems in turn depend on three main lemmas.

The first of these is combinatorial in nature and defines the basic approach of the main theorems. We fix λ, y, and N. N is chosen to be large enough (explicitly stated in the theorems) and such that all its prime divisors are less than

y. This condition on N simplifies the working of a certain technical detail, and in fact we can do without this condition by doing more sophisticated analysis. We consider the computed values of the Fourier series $-\sum_{\substack{n=1 \\ q|n \Rightarrow q \leq y}}^{\infty} \lambda(n)\frac{\cos 2\pi n(j/N)}{n}$ for $0 \leq j \leq N-1$, as a finite, ordered set. The combinatorial lemma states that given such a set, if there exists another set whose elements are close to the given set, then we can estimate bounds for the cardinality of the positive or negative elements of the second set. Hence, if we can show a closeness between the two Fourier series (one for the Legendre symbol and the other for λ), then using this lemma, we could obtain bounds for the measure under question, and we are done. However, a closeness between the two Fourier series is required.

This brings us to our second main lemma. Here, for a fixed λ, y we consider all those primes $p \equiv 3 \pmod 4$ for which $\left(\frac{q}{p}\right) = \lambda(q)$ for all primes $q \leq y$, and analytically evaluate an upper bound for the second moments of the difference of the two Fourier series. As is naturally expected, the second moment turns out to be rather small, and we successfully establish a closeness. However, we realize that this closeness has been established in an analytic setting while the measure is being calculated (via the combinatorial lemma) in a discrete setting, hence the need for a third main lemma.

Assuming GRH, this third lemma proves that the discrete second moments can be approximated by the analytic second moments up to an error term, which is very small for large y. In our computations however, we assume that this error term is very small for all $y \geq 29$. Based on this assumption, we calculate an upper bound for the discrete second moments and thus, successfully establish the required closeness in a discrete setting, satisfy all the requirements of the combinatorial lemma, derive the main theorems and via a series of numerical computations eventually arrive at a proof for Theorem 0.0.2.

Our numerical computations are based on the main theorems which for a given λ and y, give explicit formulas for evaluating bounds for the measure of

the set of $x \in [0,1]$ for which the Fourier series $-\sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$ is negative. Owing to the fact that for a given y, calculating bounds for all the corresponding λ sequences involves working with 2^m (where m is the number of primes $\leq y$) λ sequences, we adopt a greedy approach when doing our computations. We start with $y = 29$ (the 10^{th} prime) and based on the value of the bound obtained for each λ sequence, we employ a "pruning" process whereby we choose only those λ sequences that will serve as viable candidates for larger y. Next, we increment the value of y by the next 5 primes, and calculate the value of the bound for only those λ sequences that are obtained by appending all possible $\{1,-1\}$ sequences (of length 5) to the qualifying λ sequences of the previous round. Of these, again based on the value of the bound obtained for each λ sequence, we choose only those λ sequences that will serve as viable candidates for larger y. We continue with this "pruning" operation for each successive value of y until the number of λ sequences (at any given y) is close to exhausting the computer's memory. When calculating the upper bound, we were able to take y up to 349 and when calculating the lower bound, we were able to take y up to 281.

Interestingly, in spite the restriction on the value of y caused by the memory limitation of our computing resources, a careful study of the λ sequences that yield the final upper bound result, shows that one of these sequences matches the $\{1,-1\}$ sequence that would be generated by Montgomery's choice of modelling function a_3 (see Chapter 3). Therefore, even though our modelling function λ covered a larger range of possibilities of Legendre symbol sequences, we observe that Montgomery's specific modelling function choice was sufficient.

We note here that our final bounds for the measure $\alpha(p)$ as are stated in Theorem 0.0.2 were obtained via numerical computations based on the main theorems. Our computational results suggest that there exists a positive proportion of primes $p \equiv 3 \bmod 4$ for which the values of the upper and lower bounds are close to each other. In fact, we expect them to converge at some value, and owing to Montgomery's result we conjecture that, there exists a positive proportion of primes $p \equiv 3 \bmod 4$ for which $S_p(x)$ is negative for at least $2/3$ of the values of

$x \in [0, 1]$.

The numerical computations were done using algorithms which were coded in C++ and run on a two processor Xeon3 2.8GHz machine with 2GB RAM DDR266. These algorithms are based on theorems derived from the combinatorial lemmas and are available in Appendix A. The algorithms were iterative and so results obtained at each iteration are tabulated in Appendix B. We observe that our results can be improved by enhancing our computing environment or by a slight change in the data-types used for the variables in our coding. These suggestions are explained in more detail in the conclusion.

Finally, we observe that even though in our study we only considered the primes congruent to 3 (mod 4), the same ideas can be used to prove similar results for primes $p \equiv 1 \pmod 4$. We have via Lemma 2.6.3, that for such primes, the sum $S_p(x)$ (considered as a function of x) is an odd function. Thus, for such primes, we know that for $x \in [0, 1]$, $S_p(x)$ is as frequently positive as it is negative. Hence, it would be interesting to study its behaviour for $0 \leq x \leq \frac{1}{2}$. Since Pólya gave a Fourier expansion of $S_p(x)$ for primes $p \equiv 1 \bmod 4$ as well (see Lemma 3.1.1), our proof of Chapter 5 can be modified to also give results for such primes.

Chapter 1

NOTATION, DEFINITIONS, AND STANDARD THEOREMS

In analytic number theory there are certain standard arithmetic functions, notation and methods that are used over and over again. In this chapter we will give a brief overview of these tools for future reference.

1.1. Standard Arithmetic Functions

Here are some standard arithmetic functions used in number theory:

$$\omega(n) = |\{p : p|n, p \text{ prime}\}| = \sum_{\substack{p|n \\ p \text{ prime}}} 1$$

$$\Omega(n) = |\{p^a : p^a|n, p \text{ prime}\}|$$

$$\tau(n) = |\{d : d|n\}| = \sum_{d|n} 1$$

$$\sigma(n) = \sum_{d|n} d$$

$$\pi(x) = |\{p \leq x : p \text{ prime}\}| = \sum_{\substack{p \leq x \\ p \text{ prime}}} 1$$

$$\theta(x) = \sum_{\substack{p \leq x \\ p \text{ prime}}} \log p$$

The Mobius Function:

$$\mu(n) = \begin{cases} 0 & \text{if } p^2|n, \text{ for some prime } p, \\ (-1)^{\omega(n)} & \text{otherwise.} \end{cases}$$

The Euler ϕ Function:

$$\phi(n) = |\{m : 1 \leq m \leq n, gcd(m,n) = 1\}| = n \prod_{\substack{p|n \\ p \text{ prime}}} \left(1 - \frac{1}{p}\right)$$

Von-Mangoldt's Function:

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

1.2. Big Oh, Little Oh, and Vinogradov Notation

Given two functions $f(x)$ and $g(x)$, we say that $f(x)$ is $O(g(x))$, which is read as "$f(x)$ is big-Oh of $g(x)$" if there exists a positive number C such that for all x, $|f(x)| \leq C|g(x)|$.

If $\lim_{x\to\infty} f(x)/g(x) = 1$, then we say that $f(x)$ is asymptotic to $g(x)$, and we denote this by $f(x) \sim g(x)$.

We say that $f(x) = o(g(x))$, which is read as "$f(x)$ is little-Oh of $g(x)$" if $\lim_{x\to\infty} f(x)/g(x) = 0$.

Vinogradov's notation: $f(x) \ll g(x)$ means there exists a positive constant C such that $|f(x)| < C|g(x)|$.

1.3. Partial Summation

The method of partial summation is used frequently in analytic number theory. It is a method for estimating the sum $A(x) = \sum_{1\leq n\leq x} a_n f(n)$, if one has an estimate for $B(t) = \sum_{1\leq n\leq t} a_n$, for $t \leq x$, where a_n is some arbitrary sequence, $f(x)$ is some differentiable function and $f'(x)$ is continuous. The method works by the following formula:

$$A(x) = f(x)B(x) - \int_1^x B(t)f'(t)\, dt. \tag{1.3.1}$$

1.4. Merten's Theorem

Merten's Theorem states that for a prime, as $x \to \infty$,

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log x} \tag{1.4.1}$$

where $\gamma = 0.57721566...$ is Euler's constant. Taking the logarithm of this relation gives

$$\sum_{p \leq x} \frac{1}{p} = \log \log x + A + o(1)$$

where

$$A = \gamma + \sum_{p \text{ prime}} \left(\log \left(1 - \frac{1}{p}\right) + \frac{1}{p} \right).$$

1.5. The Euler Product

The fundamental theorem of arithmetic states that every integer $n > 1$ can be written as a product of prime factors in only one way, apart from the order of the factors. The next theorem, discovered by Euler in 1737, is sometimes called the analytic version of the fundamental theorem of arithmetic.

Theorem 1.5.1. *If f is a multiplicative arithmetic function, and $\sum_{n \geq 1} f(n)$ converges absolutely, then*

$$\sum_{n \geq 1} f(n) = \prod_{p \text{ prime}} (f(1) + f(p) + f(p^2) + ...)$$

and if, moreover, f is totally multiplicative, then

$$\sum_{n \geq 1} f(n) = \prod_{p \text{ prime}} (1 - f(p))^{-1}.$$

In each of the above cases, the product is called the *Euler product* of the series.

1.6. The Prime Number Theorem

The strong form of the Prime Number Theorem (as was independently proved by J. Hadamard and C.J. de la Vallée-Poussin in 1896) states that for some positive number C,

$$\pi(x) = \text{li}(x) + O(xe^{-C\sqrt{\log(x)}})$$

where

$$\mathrm{li}(x) = \int_2^x \frac{1}{\log t}\, dt \sim \frac{x}{\log x}.$$

In 1837 Dirichlet proved the existence of primes in an arithmetic progression. He further proved that if a, d are coprime integers and $d > 0$, the arithmetic progression $\{a, a+d, a+2d, ...\}$ contains infinitely many primes. It is now known that if $\pi(x; d, a)$ denotes the number of primes in the stated progression that do not exceed x, then for fixed coprime integers a, d with $d > 0$,

$$\pi(x; d, a) \sim \frac{1}{\phi(d)}\pi(x) \sim \frac{1}{\phi(d)}\mathrm{li}(x)$$

as $x \to \infty$ and where ϕ is the Euler function.

1.7. Dirichlet Characters

Consider the following group homomorphism:

$$\chi : (\mathbb{Z}/k\mathbb{Z})^* \longrightarrow \mathbb{C}$$

where $a \bmod k$ (with $(a, k) = 1$) is a multiplicative group of residue classes. Now consider χ as a function on $\mathbb{Z}$ by setting $\chi(n) = 0$ whenever $(n, k) \neq 1$. This makes χ a *Dirichlet character*, a function on $\mathbb{Z}$, periodic modulo k and completely multiplicative.

The *principal character* χ_0 of modulus k is that character which has the properties:

$$\chi_0(n) = \begin{cases} 1 & \text{if } (n, k) = 1, \\ 0 & \text{if } (n, k) > 1. \end{cases}$$

There are $\phi(k)$ distinct Dirichlet characters modulo k, each of which is completely multiplicative. Following are some other properties of the Dirichlet characters modulo k:

1. $\frac{1}{\phi(k)} \sum_{n=1}^{k} \chi(n) = \begin{cases} 1 & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise,} \end{cases}$
   where χ_0 is the principal character.
2. $\frac{1}{\phi(k)} \sum_{\chi \bmod k} \chi(n) = \begin{cases} 1 & \text{if } n \equiv 1 \pmod{k}, \\ 0 & \text{otherwise,} \end{cases}$
   where the summation is over all $\phi(k)$ characters.

Property 2 gives rise to the orthogonality relation for characters which says that if $(a,k)=1$, then

$$\frac{1}{\phi(k)}\sum_{\chi}\bar{\chi}(a)\chi(n)=\begin{cases}1 & \text{if } n\equiv a \bmod k,\\ 0 & \text{otherwise.}\end{cases}$$

for we have $\bar{\chi}(a)\chi(n)=\chi(n')$ and $n'\equiv 1 \bmod k$ if and only if $n\equiv a \bmod k$.

Let $\chi(n)$ be any character to the modulus k other than the principal character. If $(n,k)>1$, then $\chi(n)=0$; if $(n,k)=1$, then $\chi(n)\neq 0$, being a root of unity, and is a periodic function of n with period k. It is possible, however, that for values of n restricted by the condition $(n,k)=1$, the function $\chi(n)$ may have period less than k. If so, we say that χ is *imprimitive*, and otherwise *primitive*. Further, every non-principal character χ modulo a prime p, is a primitive character mod p.

1.8. Dirichlet L-Functions

For χ a Dirichlet character modulo k, the Dirichlet L-function is defined as follows:

$$L(s,\chi):=\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}.$$

This sum converges in the region $Re(s)>1$. Taking the Euler product, we have:

$$L(s,\chi)=\prod_{p \text{ prime}}(1-\chi(p)p^{-s})^{-1}.$$

For the case when $\chi(n)=1$ for all n, we get the Riemann zeta function which we will formally introduce next.

1.9. The Riemann Zeta Function

In his classic memoir of 1860, Riemann showed that the key to the deeper investigation of the distribution of primes lies in the study of the Riemann zeta

function $\zeta(s)$ defined as follows:

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}$$

where s is allowed to attain complex values. This sum converges absolutely for $Re(s) > 1$. As was proved by Euler, for $Re(s) > 1$, it can be expressed as an Euler Product giving

$$\zeta(s) = \prod_{p \text{ prime}} (1 - p^{-s})^{-1}.$$

The Riemann zeta function satisfies a functional equation which relates $\zeta(s)$ to $\zeta(1-s)$. From this functional equation we can deduce that $\zeta(s)$ has an analytic continuation to the whole complex plane $\mathbb{C}$ except at $s = 1$. Also, the functional equation tells us that the zeros of $\zeta(s)$ lie in the critical strip $0 \leq Re(s) \leq 1$, with the exception of the *trivial zeros* at $s = -2, -4, -6....$ Moreover, we can observe that the zeros are symmetrically arranged about the real axis, and also about the *critical line* given by $Re(s) = 1/2$. Following are some interesting facts about the theoretical applications of ζ:

1. The fact that $\zeta(s) \to \infty$ as $s \to 1$ implies the infinitude of primes.
2. The fact that $\zeta(s)$ has no zeros on the line $Re(s) = 1$ leads to the Prime Number Theorem.

In his memoir, Riemann made a remarkable conjecture, which has become a central conjecture for all of number theory:

Conjecture 1.9.1 (Riemann Hypothesis (RH)). *All the non-trivial zeros of* $\zeta(s)$ *in the critical strip* $0 \leq Re(s) \leq 1$ *lie on the line* $Re(s) = 1/2$.

Further, the Generalized Riemann Hypothesis (GRH) says that every Dirichlet L-function has its non-trivial zeros in the critical strip $0 \leq Re(s) \leq 1$ on the line $Re(s) = 1/2$.

In our work we will also refer to existing results for sums of the form

$$\zeta(q_1, q_2, \ldots, q_g) := \sum_{a_1 > a_2 > \cdots > a_g \geq 1} \frac{1}{{a_1}^{q_1}} \frac{1}{{a_2}^{q_2}} \cdots \frac{1}{{a_g}^{q_g}} \tag{1.9.1}$$

where all the a_i's and q_i's are positive integers, with $q_1 \geq 2$. Note that it is necessary that $q_1 \geq 2$ else the sum will diverge. We observe that in the case where $g = 1$, we obtain the Riemann zeta function.

1.10. NOTATION

Here is some notation that will be used during our work.

1. If e_1 is an m-dimensional vector and e_2 is an n-dimensional vector, the vector E written as $(e_1|e_2)$ is an $m+n$ dimensional vector in which we append e_2 to e_1.
2. $e(n)$ means $e^{2\pi in}$.
3. log means logarithm to the base e.

Chapter 2

QUADRATIC RESIDUES AND SOME QUESTIONS CONCERNING THEIR DISTRIBUTION

This chapter will begin with an introduction to quadratic residues, the definition of the Legendre symbol and the famous Quadratic Reciprocity Law. We will give some basic properties of the quadratic residues and the Legendre symbol since these will set the foundations of our later work.

Having established the basics, we will discuss some natural questions concerning the distribution of the quadratic residues and give a result of Gauss which relates to the character of consecutive numbers.

Finally, we will recall some results of the sums of quadratic residues. In doing so, we will state the well-known Pólya-Vinogradov Inequality and give its proof for primitive characters. We will also state Burgess' 1957 result for quadratic residues and non-residues. Then, we will introduce our specific area of interest of this thesis, which is the study of the distribution of the sum of Legendre symbols. Here we will outline a statistical notion concerning Legendre symbols, and define the sum $S_p(x)$. Finally, we will state Hugh L. Montgomery's [**23**] result of 1974 for the distribution of this sum. The proof of Montgomery's result will be given in detail in Chapter 3.

2.1. Quadratic Residues

Here we will define quadratic residues and prove that for p an odd prime, there exist exactly $(p-1)/2$ quadratic residues modulo p and $(p-1)/2$ quadratic non-residues modulo p.

Definition 2.1.1 (Quadratic Residues). *For coprime integers m, a, with m positive, we say that a is a quadratic residue* $\pmod m$ *if and only if the congruence*

$$x^2 \equiv a \pmod m \tag{2.1.1}$$

is solvable for some integer x. If the congruence is not solvable, a is said to be a quadratic non-residue $\pmod m$.

Theorem 2.1.2. *Let p be an odd prime. There exist exactly $(p-1)/2$ quadratic residues modulo p and $(p-1)/2$ quadratic non-residues modulo p.*

Proof. The quadratic residues are the nonzero numbers that are squares modulo p, so they are the numbers

$$1^2, 2^2, \ldots, (p-1)^2 \pmod p.$$

However, since we know that $(p-b) \equiv -b \mod p$, $(1 \leq b \leq p-1)$, it follows that

$$(p-b)^2 = p^2 - 2pb + b^2 \equiv b^2 \mod p.$$

Now, if we want to list all of the (nonzero) numbers that are squares modulo p, we only need to compute half of them,

$$1^2, 2^2, \ldots, \left(\frac{p-1}{2}\right)^2 \pmod p.$$

Since the same numbers are repeated in reverse order, if we square the remaining numbers, we get

$$\left(\frac{p+1}{2}\right)^2, \ldots, (p-2)^2, (p-1)^2 \pmod p.$$

Now, to show that there are exactly $(p-1)/2$ quadratic residues modulo p, we have to show that the numbers $1^2, 2^2, \ldots, \left(\frac{p-1}{2}\right)^2$ are all distinct modulo p.
Take $i, j \in \mathbb{Z}$. If $1 \leq i, j \leq \frac{p-1}{2}$ and $i^2 \equiv j^2 \bmod p$, then $p \mid (j^2 - i^2) = (j-i)(j+i)$. Therefore, either p divides $(j-i)$ or p divides $(j+i)$.

However, since $2 \leq j+i \leq p-1$, $(j+i)$ cannot be divisible by p. Thus p must divide $(j-i)$. But $|j-i| < (p-1)/2$, so we must have $i = j$ if p is to divide $(j-i)$. This shows that the numbers $1^2, 2^2, \ldots, \left(\frac{p-1}{2}\right)^2$ are all distinct modulo p, and so there are exactly $(p-1)/2$ quadratic residues modulo p. Further, since p does not divide any other integer n, $1 \leq n \leq p-1$, it trivially follows that there are also exactly $(p-1)/2$ quadratic non-residues modulo p. □

2.2. The Legendre Symbol and its Properties

The Legendre symbol is used to signify whether or not $a \not\equiv 0 \pmod{p}$ is a square $\pmod{p}$. The symbol was introduced by Adrien-Marie Legendre in his *Essai sur la theorie des nombres* in 1798 and is defined as follows:

Definition 2.2.1 (The Legendre Symbol). *For odd prime p, the Legendre symbol $\left(\frac{a}{p}\right)$ is defined as*

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } a \equiv 0 \pmod{p}, \\ 1 & \text{if } a \text{ is a quadratic residue} \pmod{p}, \\ -1 & \text{if } a \text{ is a quadratic non-residue} \pmod{p}. \end{cases} \tag{2.2.1}$$

Now we will state some of the frequently used properties of the Legendre symbol.

1. The Legendre symbol $\left(\frac{n}{p}\right)$ is a completely multiplicative function of n i.e.
$$\left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right)\left(\frac{n}{p}\right)$$
2. The Legendre symbol $\left(\frac{\cdot}{p}\right)$ is periodic with period p.
3. *Euler's criterion* states that for p an odd prime, $\left(\frac{n}{p}\right) = n^{(p-1)/2} \pmod{p}$ for all n.

Remark 2.2.2. *Owing to the first two properties and the fact that $\left(\frac{n}{p}\right)$ vanishes when $p|n$, it follows that $\left(\frac{n}{p}\right) = \chi(n)$ where χ is one of the Dirichlet characters modulo p. The Legendre symbol is called the* quadratic character mod p.

Next we state a simple consequence of Theorem 2.1.2.

Corollary 2.2.3. *The sum* $\sum_{n=1}^{p} \left(\frac{n}{p}\right) = 0$ *for any fixed* p.

PROOF.

$$\sum_{n=1}^{p} \left(\frac{n}{p}\right) = \sum_{\substack{n=1 \\ n \equiv a^2 \bmod p \\ a \in \mathbb{Z}}}^{p} \left(\frac{n}{p}\right) + \sum_{\substack{n=1 \\ n \not\equiv a^2 \bmod p \\ a \in \mathbb{Z}}}^{p} \left(\frac{n}{p}\right) = \frac{p-1}{2}(1) + \frac{p-1}{2}(-1) = 0.$$

□

Remark 2.2.4. *We note here that since the Legendre symbol* $\left(\frac{\div}{p}\right)$ *is in fact a Dirichlet character modulo* p *(Remark 2.2.2), Corollary 2.2.3 also follows directly from Property 1 of Section 1.7.*

There are two basic problems that dominate the theory of quadratic residues:

1. Given an odd prime p, determine which n are quadratic residues $\pmod p$ and which are quadratic non-residues $\pmod p$.
2. Given n, determine those primes for which n is a quadratic residue and for which it is a quadratic non-residue.

While the first of these can be solved via Euler's criterion, the second is more difficult and its solution depends on a remarkable theorem known as the *quadratic reciprocity law.* We will study this theorem in the next section. There are still many open questions regarding quadratic residues, for example, can we find an n, a quadratic non-residue $\pmod p$, in polynomial time?

2.3. THE LAW OF QUADRATIC RECIPROCITY

The theory of quadratic residues is dominated by the famous *Law of Quadratic Reciprocity* which was referred to as "the gem of higher arithmetic" by Gauss. This law was first stated in 1751 by Euler and was rediscovered by Legendre in 1785 who gave a partial proof. In 1796, at the age of 18, Gauss independently discovered the Law of Quadratic Reciprocity and gave the first proof which is available in the fourth section of his famous work, *Disquisitiones Arithemeticae.* His efforts on arriving at the proof are as follows,

> "I discovered this theorem independently in 1795 at a time when I was totally ignorant of what had been achieved in higher arithmetic, and consequently had not the slightest aid from the literature on the subject. For a whole year this theorem tormented me and absorbed my greatest efforts until at last I obtained a proof given in the fourth section of *Disquisitiones Arithemeticae*. Later I ran across three other proofs which were built on entirely different principles. One of these I have already given in the fifth section, the others, which do not compare with it in elegance, I have reserved for future publication. Although these proofs leave nothing to be desired as regards rigor, they are derived from sources much too remote, except perhaps the first, which however proceeds with laborious arguments and is overloaded with extended operations."

The law reads as follows:

Theorem 2.3.1 (Law of Quadratic Reciprocity). *If p and q are distinct odd primes, then*

$$\left(\frac{p}{q}\right) = \begin{cases} \left(\frac{q}{p}\right) & \text{if } p \equiv 1 \pmod 4 \text{ or } q \equiv 1 \pmod 4, \\ -\left(\frac{q}{p}\right) & \text{if } p \equiv q \equiv 3 \pmod 4. \end{cases} \tag{2.3.1}$$

It is often stated in the following symmetric form given by Legendre. For p and q distinct odd primes,

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{(p-1)(q-1)/4}. \tag{2.3.2}$$

For p an odd prime, the following is then a natural consequence of the Quadratic Reciprocity Law:

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1 \pmod 4, \\ -1 & \text{if } p \equiv 3 \pmod 4. \end{cases} \tag{2.3.3}$$

We also have the following relation for p an odd prime:

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{if } p \equiv \pm 1 \pmod 8, \\ -1 & \text{if } p \equiv \pm 3 \pmod 8. \end{cases}$$

2.4. Distribution of the Quadratic Residues

The distribution of the quadratic residues and non-residues has intrigued many minds for many years. By Theorem 2.1.2, we know that for any given p an odd prime, half of the numbers

$$1, 2, \ldots, p-1$$

are quadratic residues, and the other half are quadratic non-residues. A natural question that arises is: How are these quadratic resides and quadratic non-residues distributed? Studying patterns for a given p suggests that if p is a large prime, the quadratic residues and quadratic non-residues have a distribution that appears to be fairly random. And so arise various questions targeted towards testing the random character of the distribution. One such question, which was answered by Gauss, concerns the characters of consecutive numbers. That is, if n and $n+1$ are two consecutive numbers in the series $1, 2, \ldots, p-1$, how often do they have prescribed characters. The possible pair of characters for a pair of numbers are denoted as RR, RN, NR, NN, where R represents a quadratic residue and N a quadratic non-residue. Based on our observation that quadratic residues and quadratic non-residues appear to be distributed randomly, we may expect that each of the four types occur about equally often. This is in fact the case and was proved by Gauss. For a proof, see [8]. This result of Gauss has been restated in [31] in the following manner:

Theorem 2.4.1. [31] *Let p be an odd prime, $\left(\frac{\cdot}{p}\right)$ the Legendre symbol and R_p the complete set of residues modulo p. We have*

$$(RR) = |\{n : \left(\frac{n}{p}\right) = \left(\frac{n+1}{p}\right) = 1, n \in R_p\}| = \left[\frac{p-3}{4}\right]. \tag{2.4.1}$$

$$(NN) = |\{n : \left(\frac{n}{p}\right) = \left(\frac{n+1}{p}\right) = -1, n \in R_p\}| = \left[\frac{p-1}{4}\right]. \tag{2.4.2}$$

where $[.]$ is the greatest integer function.

Similarly, for p an odd prime, R_p the complete set of residues modulo p, using Gauss' proof we can show the following:

$$(NR) = |\{n : \left(\frac{n}{p}\right) = -1 \text{ and } \left(\frac{n+1}{p}\right) = 1, n \in R_p\}| = \left[\frac{p+1}{4}\right]. \tag{2.4.3}$$

$$(RN) = |\{n : \left(\frac{n}{p}\right) = 1 \text{ and } \left(\frac{n+1}{p}\right) = -1, n \in R_p\}| = \left[\frac{p-1}{4}\right]. \qquad (2.4.4)$$

This proves the assertion that for large primes, each of (RR), (NN), (RN), (NR) is about $p/4$. An important step in the proof evaluates the sum of the Legendre symbols $\left(\frac{n(n+1)}{p}\right)$. Clearly, $\left(\frac{p}{p}\right) = 0$, and this sum is given by

$$\sum_{n=0}^{p-1} \left(\frac{n(n+1)}{p}\right) = -1. \qquad (2.4.5)$$

In fact for any $(b, p) = 1$, writing $n = mb$,

$$\sum_{n=0}^{p-1} \left(\frac{n(n+b)}{p}\right) = \sum_{m=0}^{p-1} \left(\frac{mb(mb+b)}{p}\right) = \left(\frac{b^2}{p}\right) \sum_{m=0}^{p-1} \left(\frac{m(m+1)}{p}\right) = 1(-1) = -1. \qquad (2.4.6)$$

Extensive literature exists in this area. In 1931 H. Davenport [**7**] gave upper bounds for RRR and NNN. Denote by R_t and N_t the number of occurrences of t consecutive quadratic residues and non-residues, respectively. A. Brauer, in 1928 showed that for any t and large enough prime p, both R_t and N_t are greater than 0 (see [**3**]). In 1983, R.H. Hudson [**20**] showed that $RRNR > 0$ and $RNRR > 0$ for large enough primes p. In fact, we know that for any pattern $RN\ldots$ of length k, the number of occurrences of that pattern is $\frac{p}{2^k} + O_k(\sqrt{p})$. More recent work includes Z.H. Sun's paper [**31**] of 2002 which is an extension of Gauss' work to determine all those values of n (for n in a complete set of residues (mod p)) for which $\left(\frac{n}{p}\right) = \left(\frac{n+1}{p}\right)$ or $\left(\frac{n-1}{p}\right) = \left(\frac{n}{p}\right) = \left(\frac{n+1}{p}\right)$. Sun's work shows that $\left(\frac{n}{p}\right) = \left(\frac{n+1}{p}\right)$ if and only if $n \equiv \frac{(x-1)^2}{4x} \pmod{p}$ for some $x \in \mathbb{Z}$ and that $\left(\frac{n-1}{p}\right) = \left(\frac{n}{p}\right) = \left(\frac{n+1}{p}\right)$ if and only if $n \equiv \frac{(x^2+1)^2}{4x^3-4x} \pmod{p}$ for some $x \in \mathbb{Z}$. Much work has also been dedicated to the question that what would be the expected value of the least quadratic non-residue (mod p).

2.5. Lower Bound Results for the Sum of Quadratic Characters

2.5.1. A result of Pólya-Vinogradov

In this section we will recall the famous Pólya-Vinogradov inequality for χ a non-principal character $\pmod q$ and will prove it only for χ a primitive quadratic character and more precisely the Legendre Symbol $\left(\frac{\cdot}{q}\right)$.

Theorem 2.5.1 (The Pólya-Vinogradov Inequality). *Let M and N be positive integers. For χ a non-principal character $\pmod q$,*

$$\sum_{n=M+1}^{M+N} \chi(n) \ll \sqrt{q}\log q. \tag{2.5.1}$$

If we take $\chi(n) = \left(\frac{n}{q}\right)$, then we can deduce that the interval $M+1 \leq n \leq M+N$ contains $\frac{1}{2}N + O(\sqrt{q}\log q)$ quadratic residues (mod q). By considering the sum $\sum_{n\leq qx} \chi(n)$ as a function with period 1, and determining its Fourier expansion, Pólya thus deduced that for q an odd prime,

$$\left|\sum_{n=M+1}^{M+N} \chi(n)\right| < \sqrt{q}\log q. \tag{2.5.2}$$

Here, we will prove (2.5.2) for primitive characters only.

Proof. Let χ be a primitive character $\pmod q$, $q > 1$ such that $\chi(n) = \left(\frac{n}{q}\right)$. We begin by defining the Gaussian sum $\tau(\chi)$ for any character $\chi(n)$ to the modulus q,

$$\tau(\chi) := \sum_{m=1}^{q} \chi(m) e^{\frac{2\pi i m}{q}}. \tag{2.5.3}$$

If $(n,q) = 1$, $\tau(\chi) \neq 0$, then consider $m \equiv na \bmod q$ and we have

$$\chi(n)\tau(\bar{\chi}) = \sum_{m=1}^{q} \bar{\chi}(m)\chi(n) e^{\frac{2\pi i m}{q}} = \sum_{a=1}^{q} \bar{\chi}(a) e^{\frac{2\pi i n a}{q}}. \tag{2.5.4}$$

Therefore,

$$\chi(n) = \frac{1}{\tau(\bar{\chi})} \sum_{a=1}^{q} \bar{\chi}(a) e^{\frac{2\pi i n a}{q}}.$$

We will let the sum run from $-(\frac{q-1}{2}) \leq a \leq (\frac{q-1}{2})$, $a \neq 0$. We then have

$$\sum_{n=M+1}^{M+N} \chi(n) = \frac{1}{\tau(\bar{\chi})} \sum_{\substack{-(\frac{q-1}{2}) \leq a \leq (\frac{q-1}{2}) \\ a \neq 0}} \bar{\chi}(a) \sum_{n=M+1}^{M+N} e^{\frac{2\pi i n a}{q}}. \tag{2.5.5}$$

We will now evaluate the inner sum which is given by a geometric series:

$$\begin{aligned} \sum_{n=M+1}^{M+N} e^{\frac{2\pi i n a}{q}} &= \frac{e^{\frac{2\pi i a(M+1)}{q}}\left(1 - e^{\frac{2\pi i a N}{q}}\right)}{1 - e^{\frac{2\pi i a}{q}}} = \frac{(e^{\frac{2\pi i a(M+1)}{q}})e^{\frac{2\pi i a N}{2q}}\left(e^{\frac{-2\pi i a N}{2q}} - e^{\frac{2\pi i a N}{2q}}\right)}{e^{\frac{2\pi i a}{2q}}\left(e^{\frac{-2\pi i a}{2q}} - e^{\frac{2\pi i a}{2q}}\right)} \\ &= \frac{(e^{\frac{2\pi i a(M+1)}{q}})e^{\frac{2\pi i a N}{2q}}}{e^{\frac{2\pi i a}{2q}}}\left(\frac{-2i\sin(\pi N a/q)}{-2i\sin(\pi a/q)}\right) = e^{\frac{2\pi i a(M+\frac{1}{2}+\frac{N}{2})}{q}}\left(\frac{\sin(\pi N a/q)}{\sin(\pi a/q)}\right) \end{aligned}$$

Therefore

$$\left| \sum_{n=M+1}^{M+N} e^{\frac{2\pi i n a}{q}} \right| \leq \frac{1}{|\sin(\pi a/q)|}$$

and hence,

$$\left| \sum_{n=M+1}^{M+N} \chi(n) \right| = \left| \frac{1}{\tau(\bar{\chi})} \sum_{\substack{-(\frac{q-1}{2}) \leq a \leq (\frac{q-1}{2}) \\ a \neq 0}} \bar{\chi}(a) \sum_{n=M+1}^{M+N} e^{\frac{2\pi i n a}{q}} \right| \leq 2\frac{1}{\tau(\bar{\chi})} \sum_{a=1}^{(q-1)/2} \frac{1}{|\sin(\pi a/q)|}. \tag{2.5.6}$$

For convex functions $f(\alpha)$, we have

$$f(\alpha) \leq \frac{1}{\delta} \int_{\alpha - \frac{1}{2}\delta}^{\alpha + \frac{1}{2}\delta} f(\beta)\, d\beta.$$

Let $\alpha = \frac{a}{q}$ and take $f(\alpha) = (\sin \pi\alpha)^{-1}$, $\delta = \frac{1}{q}$. Then

$$\sum_{a=1}^{(q-1)/2} \frac{1}{|\sin(\pi a/q)|} = \sum_{\alpha = \frac{1}{q}}^{(q-1)/2q} f(\alpha) \leq q \sum_{\alpha = \frac{1}{q}}^{(q-1)/2q} \int_{\alpha - \frac{1}{2q}}^{\alpha + \frac{1}{2q}} (\sin \pi\beta)^{-1}\, d\beta = q \int_{\frac{1}{2q}}^{\frac{1}{2}} (\sin \pi\beta)^{-1}\, d\beta.$$

Since, for $0 < \beta < \frac{1}{2}$, $\sin \pi\beta > 2\beta$,

$$\sum_{a=1}^{(q-1)/2} \frac{1}{|\sin(\pi a/q)|} < q \int_{\frac{1}{2q}}^{\frac{1}{2}} \frac{1}{2\beta}\, d\beta = \frac{q}{2}\Big[\log \beta\Big]_{\frac{1}{2q}}^{\frac{1}{2}} = \frac{q \log q}{2}. \tag{2.5.7}$$

Further, we can show that for a primitive character χ, the Gauss sum $|\tau(\chi)| = \sqrt{q}$.

We will evaluate the Gauss sum by squaring (2.5.4) so that

$$|\tau(\bar{\chi})|^2 |\chi(n)|^2 = \sum_{a_1=1}^{q} \bar{\chi}(a_1) e^{\frac{2\pi i n a_1}{q}} \sum_{a_2=1}^{q} \bar{\chi}(a_2) e^{\frac{2\pi i n a_2}{q}}.$$

Now, sum for n over a complete set of residues $\pmod q$ which gives

$$|\tau(\bar{\chi})|^2 \sum_{n=0}^{q-1} |\chi(n)|^2 = \sum_{a_1=1}^{q} \sum_{a_2=1}^{q} \bar{\chi}(a_1)\chi(a_2) \sum_{n=0}^{q-1} e^{\frac{2\pi i n(a_1-a_2)}{q}}.$$

Since $\sum_{n=0}^{q-1} e^{\frac{2\pi i n(a_1-a_2)}{q}} = 0$ unless $a_1 = a_2$, we have

$$|\tau(\bar{\chi})|^2 \sum_{n=0}^{q-1} |\chi(n)|^2 = \sum_{a_1=1}^{q} |\chi(a_1)|^2 \sum_{n=0}^{q-1} 1 = q \sum_{a_1=1}^{q} |\chi(a_1)|^2$$

and hence $|\tau(\chi)| = \sqrt{q}$. Using this and (2.5.7) in (2.5.6) gives us (2.5.2). □

2.5.2. A result of Burgess

Burgess' contribution to this area was in the form of proving bounds for shorter character sums. In 1957 Burgess [4] stated the following result for quadratic residues and non-residues $\pmod p$.

Theorem 2.5.2 (D.A Burgess). *Let δ and ϵ be fixed positive numbers. Then, for all sufficiently large p and N, we have*

$$\left| \sum_{n=N+1}^{N+H} \left(\frac{n}{p}\right) \right| < \epsilon H$$

provided $H > p^{\frac{1}{4}+\delta}$.

This implies in particular, that the maximum number of consecutive quadratic residues or quadratic non-residues $\pmod p$ is $O(p^{\frac{1}{4}+\delta})$ for large p.

2.6. Distribution Questions concerning the Sum of Legendre Symbols

2.6.1. A Statistical Notion

The discussion about the distribution of quadratic residues in section 2.4, and in particular Gauss' result about the occurrence of consecutive quadratic residues or non-residues shows that, if the Legendre symbol is thought of as generated by a "random coin flip", then there would be $p/4$ occurrences of a given pair $(\pm 1, \pm 1)$. This would imply that the Legendre symbol is in some sense

random. However, the attractive Pólya-Vinogradov inequality says that indeed the "statistical fluctuation" of the quadratic residue/nonresidue count, starting at any initial point, is always bounded by a "variance factor" $\sqrt{p}\log p$ and hence *not* just any coin-flip sequence can be a Legendre symbol sequence. It is clear that we cannot have a Legendre symbol sequence where the first half are 1's and the second half are -1's. In fact it suggests that we cannot build up more than an $O(\sqrt{p}\log p)$ excess of one symbol over the other and Burgess' result (Theorem 2.5.2) further narrows down the maximum consecutive number of 1's or -1's to $O(p^{\frac{1}{4}+\delta})$ $(\delta > 0)$. Further, in p flips of a coin, the frequency of a given pattern of length k is $p2^{-k}$ and so for $k > c\log p$ we do not expect to see a given pattern.

2.6.2. The Sum $S_p(x)$

In view of the statistical notion of Legendre symbols as stated above, it would be of particular interest to study the distribution of the following sum

$$S_p(x) = \sum_{n \le px} \left(\frac{n}{p}\right)$$

where p is a prime and $x \in [0,1]$. This sum $S_p(x)$ has period one. In this thesis, we will study this sum for primes $p \equiv 3 \bmod 4$ and for such primes, will determine how frequently $S_p(x)$ is negative. Hence, we will prove Theorem 0.0.2 in Chapter 5. Here we will show that this sum, considered as a function of x, is even for primes $p \equiv 3 \pmod 4$ and odd for primes $p \equiv 1 \pmod 4$.

Owing to the periodicity of $S_p(x)$, we note that $S_p(x) = S_p(1+x)$ for all $x \in [0,1]$, and so

$$S_p(1+x) = \sum_{p \le n \le p(1+x)} \left(\frac{n}{p}\right).$$

In fact for any $b \in \mathbb{N}$, we have for $x \in [0,1]$:

$$S_p(x) = S_p(b+x) = \sum_{pb \le n \le p(b+x)} \left(\frac{n}{p}\right).$$

We will now define $S_p(x)$ for $x \in [-1,0]$.

Definition 2.6.1. *For $x \in [0,1]$,*

$$S_p(-x) = S_p(1-x) = \sum_{1 \le n \le p(1-x)} \left(\frac{n}{p}\right). \tag{2.6.1}$$

Extending the above definition, we now define $S_p(x)$ for all $x < 0$.

Definition 2.6.2. *For $x < 0$, if m is the smallest integer greater than or equal to x, then*

$$S_p(-x) = S_p(m-x) = \sum_{1 \le n \le p(m-x)} \left(\frac{n}{p}\right). \tag{2.6.2}$$

Lemma 2.6.3. *For $p \equiv 1 \bmod 4$, the sum $S_p(x) = \sum_{n \le px} \left(\frac{n}{p}\right)$ is an odd function and for $p \equiv 3 \bmod 4$, the sum $S_p(x) = \sum_{n \le px} \left(\frac{n}{p}\right)$ is an even function.*

PROOF. Suppose $x < 0$ and m is the smallest integer greater than or equal to x, then by Definition 2.6.2, we have

$$S_p(-x) = S_p(m-x) = \sum_{1 \le n \le p(m-x)} \left(\frac{n}{p}\right)$$

Taking $k = pm - n$ and owing to Corollary 2.2.3, we get

$$\begin{aligned} S_p(-x) &= \sum_{n \le pm} \left(\frac{n}{p}\right) - \sum_{1 \le k \le px} \left(\frac{pm-k}{p}\right) = -\sum_{1 \le k \le px} \left(\frac{-k}{p}\right) \\ &= -\left(\frac{-1}{p}\right) \sum_{1 \le k \le px} \left(\frac{k}{p}\right) = -\left(\frac{-1}{p}\right) S_p(x). \end{aligned} \tag{2.6.3}$$

Now, by the Law of Reciprocity (Theorem 2.3.1), we have

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1 \bmod 4, \\ -1 & \text{if } p \equiv 3 \bmod 4. \end{cases}$$

Hence, for $p \equiv 1 \bmod 4$, (2.6.3) gives

$$-S_p(-x) = S_p(x), \tag{2.6.4}$$

which proves that $S_p(x)$ is an odd function for $p \equiv 1 \bmod 4$.

Similarly, for $p \equiv 3 \bmod 4$, we get

$$S_p(-x) = S_p(x), \tag{2.6.5}$$

and hence prove that $S_p(x)$ is an even function for $p \equiv 3 \bmod 4$. □

2.6.3. A result of Montgomery

Our work in this thesis was inspired by Hugh L. Montgomery's 1974 article titled "Distribution Questions concerning a Character Sum" [**23**]. In this article, he studied the distribution of the sum $S_p(x)$ (as defined above) by calculating the measure of the set of $x \in [0,1]$ where $S_p(x) > 0$ for primes congruent to 3 (mod 4). Montgomery defined this measure as:

$$\alpha(p) = |\{x \in [0,1] : S_p(x) > 0\}| \quad \text{where } p \equiv 3 \pmod 4.$$

and proved Theorem 0.0.1. Montgomery achieved his upper bound for $\alpha(p)$ (which states that for any $\delta > 0$, there are infinitely many primes $p \equiv 3 \bmod 4$ such that $\alpha(p) < \frac{1}{3} + \delta$) by modelling the Legendre symbol $\left(\frac{q}{p}\right)$, (q a prime, and p a prime congruent to 3 mod 4) by a totally multiplicative arithmetic function a_n, where $a_q = \left(\frac{q}{3}\right)$ for $q \neq 3$, and $a_3 = -1$. A detailed proof of this theorem is given in Chapter 3. We will improve Montgomery's lower bound result for $\alpha(p)$ (which states that $\alpha(p) > \frac{1}{50}$ for all primes $p \equiv 3 \bmod 4$) from $\frac{1}{50}$ to $\frac{5437}{250000}$ in Chapter 3. Theorem 0.0.1 tells us that $S_p(x)$ can be negative for nearly 2/3 of the values of $x \in [0,1]$, but not for almost all values of $x \in [0,1]$. The ideas found in Montgomery's proof can also be used to show that there are infinitely many primes $p \equiv 3 \pmod 4$ for which $\alpha(p) > 1 - \delta$.

Further, his following theorem of the same article, shows that the above results would not be affected if we were to consider the set of $x \in [0,1]$ for which $S_p(x) \geq 0$ instead of $S_p(x) > 0$.

Theorem 2.6.4 (H.L. Montgomery). *Let* $\alpha(p) = \max_a |\{x \in [0,1] : S_p(x) = a\}|$. *Then* $\alpha(p)$ *tends to* 0 *as* p *tends to infinity.*

Chapter 3

PROOF OF MONTGOMERY'S THEOREM

In this chapter, we will provide a detailed proof of Theorem 0.0.1 using the ideas that were outlined by Hugh L. Montgomery in his article "Distribution Questions concerning a Character Sum" [**23**]. Montgomery proved that the sum $S_p(x) = \sum_{n \leq px} \left(\frac{n}{p}\right)$ can be negative for nearly 2/3 of the values of $x \in [0,1]$ but not for all values of $x \in [0,1]$. He showed this by modelling the Legendre symbol $\left(\frac{\cdot}{p}\right)$ by a totally multiplicative arithmetic function. He defined this function to be equal to the value of the specific Legendre symbol given by $\left(\frac{q}{3}\right)$ (q a prime) if $q \neq 3$ and equal to -1 if $q = 3$.

The last section of this chapter will give an improvement on the lower bound result obtained by Montgomery. We will improve the bound from $\frac{1}{50}$ (as was stated by Montgomery) to $\frac{5437}{250000}$. To arrive at our new result, we will essentially just refine Montgomery's proof. We recall here the following definition:

$$\alpha(p) = |\{x \in [0,1] : S_p(x) > 0\}| \quad \text{where } p \equiv 3 \pmod 4.$$

3.1. LEMMAS

In this section, we will collect together the lemmas that were used by Montgomery in his proof.

The following lemma gives Fourier expansions for the sum $S_p(x) = \sum_{n \leq px} \left(\frac{n}{p}\right)$ as were established by G. Pólya [**27**].

Lemma 3.1.1. *Let $x \notin \{\frac{m}{p} : m \in \mathbb{N}\}$. If $p \equiv 1 \bmod 4$ then $S_p(x)$ has the Fourier expansion*

$$S_p(x) = \frac{\sqrt{p}}{\pi}\left(\sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\sin 2\pi nx}{n}\right). \tag{3.1.1}$$

If $p \equiv 3 \bmod 4$ then $S_p(x)$ has the Fourier expansion

$$S_p(x) = \frac{\sqrt{p}}{\pi}\left(L_p(1) - \sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}\right) \tag{3.1.2}$$

where $L_p(1) = \sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{1}{n}$. These Fourier expansions converge uniformly (in x, for p fixed), except in neighbourhoods of jump discontinuities of $S_p(x)$ in which case the partial sums are uniformly bounded. These jump discontinuities occur when $x \in \{\frac{m}{p} : m \in \mathbb{N}\}$.

Remark 3.1.2. *We note here that the set $\{\frac{m}{p} : m \in \mathbb{N}\}$ is of measure null. Since we will be studying a measure, we can safely assume as in the above lemma that $x \notin \{\frac{m}{p} : m \in \mathbb{N}\}$.*

In view of the Fourier expansion for the case $p \equiv 3 \bmod 4$, write:

$$T_p(x) = \sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n} \tag{3.1.3}$$

so that

$$S_p(x) = \frac{\sqrt{p}}{\pi}(L_p(1) - T_p(x)).$$

The following two lemmas evaluate the second and fourth moments respectively, for the function $T_p(x)$ as introduced above.

Lemma 3.1.3.

$$\int_0^1 T_p(x)^2\, dx = \frac{\pi^2}{12}\left(1 - \frac{1}{p^2}\right). \tag{3.1.4}$$

PROOF.

$$\begin{aligned}\int_0^1 T_p(x)^2\,dx &= \int_0^1 \left(\sum_{n\geq 1}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}\right)^2 dx \\ &= \int_0^1 \left(\sum_{n\geq 1}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}\right)\left(\sum_{m\geq 1}\left(\frac{m}{p}\right)\frac{\cos 2\pi mx}{m}\right) dx \\ &= \int_0^1 \sum_{m,n\geq 1}\left(\frac{mn}{p}\right)\frac{\cos 2\pi mx}{m}\frac{\cos 2\pi nx}{n}\,dx.\end{aligned}$$

Since,

$$\int_0^1 \cos 2\pi mx\ \cos 2\pi nx\ dx = \begin{cases}\frac{1}{2} & \text{if } m = n, \\ 0 & \text{if } m \neq n.\end{cases} \tag{3.1.5}$$

we have

$$\begin{aligned}\int_0^1 T_p(x)^2\,dx &= \frac{1}{2}\sum_{\substack{m,n\geq 1\\ m=n}}\left(\frac{mn}{p}\right)\frac{1}{mn} = \frac{1}{2}\sum_{n\geq 1}\left(\frac{n^2}{p}\right)\frac{1}{n^2} \\ &= \frac{1}{2}\sum_{\substack{n\geq 1\\ p\nmid n}}\frac{1}{n^2} \qquad \text{since } \left(\frac{n^2}{p}\right) = \begin{cases}1 & \text{if } (p,n)=1, \\ 0 & \text{if } p|n.\end{cases}\end{aligned}$$

We can write this as an Euler product, giving:

$$\begin{aligned}\int_0^1 T_p(x)^2\,dx &= \frac{1}{2}\prod_{q\neq p}\left(1-\frac{1}{q^2}\right)^{-1} = \frac{1}{2}\left(1-\frac{1}{p^2}\right)\prod_q\left(1-\frac{1}{q^2}\right)^{-1} \\ &= \frac{1}{2}\zeta(2)\left(1-\frac{1}{p^2}\right) = \frac{\pi^2}{12}\left(1-\frac{1}{p^2}\right).\end{aligned}$$

□

The latter part of the following proof uses a different approach from that of Montgomery.

Lemma 3.1.4.

$$\int_0^1 T_p(x)^4\,dx < \frac{19\pi^4}{240} \tag{3.1.6}$$

PROOF.

$$\begin{aligned}\int_0^1 T_p(x)^4\,dx &= \int_0^1 \left(\sum_{n\geq 1}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}\right)^4 dx \\ &= \sum_{\substack{n_j\geq 1\\ 1\leq j\leq 4}} \frac{1}{n_1n_2n_3n_4}\left(\frac{n_1n_2n_3n_4}{p}\right)\int_0^1 \prod_{j=1}^4 (\cos 2\pi n_j x)\,dx \\ &= \frac{1}{2^4}\sum_{\substack{n_j\geq 1\\ 1\leq j\leq 4}} \frac{1}{n_1n_2n_3n_4}\left(\frac{n_1n_2n_3n_4}{p}\right)\int_0^1 \prod_{j=1}^4 \left(e^{2\pi i n_j x}+e^{-2\pi i n_j x}\right)\,dx\end{aligned}$$

Considering $\delta_j \in \{1,-1\}$, the typical term for

$$\int_0^1 \prod_{j=1}^4 \left(e^{2\pi i n_j x}+e^{-2\pi i n_j x}\right)\,dx$$

can be written as

$$\int_0^1 e^{2\pi i\left(\sum_{j=1}^4 \delta_j n_j\right)x}\,dx = \begin{cases} 1 & \text{if } \sum_{j=1}^4 \delta_j n_j = 0, \\ 0 & \text{otherwise.}\end{cases}$$

This gives

$$\int_0^1 T_p(x)^4\,dx = \frac{1}{2^4}\sum_{\delta_1,\delta_2,\delta_3,\delta_4\in\{1,-1\}}\ \sum_{\substack{n_j\geq 1\\ \sum \delta_j n_j=0}} \frac{1}{n_1n_2n_3n_4}\left(\frac{n_1n_2n_3n_4}{p}\right).$$

Change the variables by setting $m=\delta_1n_1+\delta_2n_2=-\delta_3n_3-\delta_4n_4$, $n=\delta_2n_2$ and $N=-\delta_3n_3$. Further since $p\equiv 3 \bmod 4$, we have $\frac{1}{n_1}\left(\frac{n_1}{p}\right)=\frac{1}{(-n_1)}\left(\frac{-n_1}{p}\right)$ and so,

$$\int_0^1 T_p(x)^4\,dx = \frac{1}{2^4}\sum_{m\in\mathbb{Z}}\sum_{\substack{n,N\in\mathbb{Z}\\ n,N\neq 0 \text{ or } m}} \frac{1}{(m-n)\,n\,(m-N)\,N}\left(\frac{(m-n)\,n\,(m-N)\,N}{p}\right)$$

$$= \frac{1}{2^4}\sum_{m\in\mathbb{Z}}\left(\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0 \text{ or } m}} \frac{1}{n\,(m-n)}\left(\frac{n\,(m-n)}{p}\right)\right)^2$$

$$< \frac{1}{2^4}\sum_{m\in\mathbb{Z}}\left(\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0 \text{ or } m}} \frac{1}{|n\,(m-n)\,|}\right)^2$$

$$= \frac{1}{2^4}\left(\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}} \frac{1}{|n(-n)|}\right)^2 + \frac{1}{2^4}\sum_{\substack{m\in\mathbb{Z}\\ m\neq 0}}\left(\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0,m}} \frac{1}{|n(m-n)|}\right)^2$$

$$= \frac{1}{2^4}\left(2\sum_{n\geq 1}\frac{1}{n^2}\right)^2 + \frac{1}{2^3}\sum_{m\geq 1}\left(\sum_{1\leq n<m}\frac{1}{n(m-n)} - 2\sum_{n>m}\frac{1}{n(m-n)}\right)^2$$

$$= \frac{1}{4}\zeta(2)^2 + \frac{1}{2^3}\sum_{m\geq 1}\frac{1}{m^2}\left(2\underbrace{\sum_{m<n}\left(\frac{1}{n}+\frac{1}{m-n}\right)}_{S_1} - \underbrace{\sum_{1\leq n<m}\left(\frac{1}{n}+\frac{1}{m-n}\right)}_{S_2}\right)^2 \tag{3.1.7}$$

We will now evaluate each of the inner sums, S_1 and S_2. First we evaluate S_2.

$$S_2 = \sum_{1\leq n<m}\frac{1}{n} + \sum_{1\leq n<m}\frac{1}{m-n} = 2\sum_{1\leq n\leq m-1}\frac{1}{n}. \tag{3.1.8}$$

To evaluate S_1, take M large and $r = n - m$. Then we can write

$$\sum_{m<n\leq M}\left(\frac{1}{n}+\frac{1}{m-n}\right) = \sum_{m<n\leq M}\frac{1}{n} - \sum_{1\leq r\leq M-m}\frac{1}{r} = -\sum_{1\leq r\leq m}\frac{1}{r} + \sum_{M-m\leq r\leq M}\frac{1}{r} \tag{3.1.9}$$

and when $M\to\infty$, we are left with $S_1 = -\frac{1}{m} - \sum_{1\leq r\leq m-1}\frac{1}{r}$. These results then give

$$(2S_1 - S_2)^2 = \left(-4\sum_{1\leq j\leq m-1}\frac{1}{j} - \frac{2}{m}\right)^2 \tag{3.1.10}$$

$$= 16\Big(\sum_{1\leq j\leq m-1}\frac{1}{j}\Big)^2 + \frac{16}{m}\sum_{1\leq j\leq m-1}\frac{1}{j} + \frac{4}{m^2}$$

$$= 16\Big(\sum_{1\leq i\neq j<m}\frac{1}{ij} + \sum_{1\leq i<m}\frac{1}{i^2}\Big) + \frac{16}{m}\sum_{1\leq j<m}\frac{1}{j} + \frac{4}{m^2} \tag{3.1.11}$$

and so

$$\begin{aligned}
&\frac{1}{2^3}\sum_{m\geq 1}\frac{1}{m^2}(2S_1-S_2)^2 \\
&= 2\sum_{m\geq 1}\frac{1}{m^2}\sum_{1\leq i\neq j<m}\frac{1}{ij}+2\sum_{m\geq 1}\frac{1}{m^2}\sum_{1\leq i<m}\frac{1}{i^2}+2\sum_{m\geq 1}\frac{1}{m^3}\sum_{1\leq j\leq m}\frac{1}{j}+\frac{1}{2}\sum_{m\geq 1}\frac{1}{m^4} \\
&= 4\sum_{m>j>i\geq 1}\frac{1}{m^2ij}+2\sum_{m>i\geq 1}\frac{1}{m^2i^2}+2\sum_{m>j\geq 1}\frac{1}{m^3j}+\frac{1}{2}\zeta(4). \qquad (3.1.12)
\end{aligned}$$

We observe that the first three sums in (3.1.12) are of the form $\sum_{a_1>a_2>\cdots>a_g\geq 1}\frac{1}{a_1^{q_1}}\frac{1}{a_2^{q_2}}\cdots\frac{1}{a_g^{q_g}}$. In Chapter 1, we defined such sums as $\zeta(q_1,q_2,\ldots,q_g)$ (see (1.9.1)). In [**14**] A. Granville gives formulas for evaluating $\zeta(N-1,1)$ when $N\geq 3$ and for evaluating $\zeta(N-2,1,1)$ when $N\geq 4$, and so by identities (6), (7) and (8) of page 98 of [**14**] we have

$$\begin{aligned}
\sum_{m>j>i\geq 1}\frac{1}{m^2ij} &= \zeta(2,1,1)=\zeta(4) \\
\sum_{m>i\geq 1}\frac{1}{m^2i^2} &= \zeta(2,2)=\frac{1}{2}(\zeta(2)^2-\zeta(4)) \\
\sum_{m>j\geq 1}\frac{1}{m^3j} &= \zeta(3,1)=\frac{3}{2}\zeta(4)-\frac{1}{2}\zeta(2)^2.
\end{aligned}$$

Substituting these results in (3.1.12) gives $\frac{1}{2^3}\sum_{m\geq 1}\frac{1}{m^2}(2S_1-S_2)^2=\frac{13}{2}\zeta(4)$. Since $\zeta(2)=\frac{\pi^2}{6}$ and $\zeta(4)=\frac{\pi^4}{90}$, putting this in (3.1.7) we get

$$\int_0^1 T_p(x)^4\,dx \;<\frac{\pi^4}{144}+\frac{13\pi^4}{180}=\frac{19\pi^4}{240}.$$

□

3.2. Proof of Theorem 0.0.1

We will first prove the assertion that for any $\delta>0$, there exist infinitely many primes $p\equiv 3 \bmod 4$ such that $\alpha(p)<\frac{1}{3}+\delta$, and later that $\alpha(p)>\frac{1}{50}$ for all primes $p\equiv 3 \bmod 4$.

Proof. Suppose $p\equiv 3 \bmod 4$. Let $\delta>0$ be given, and put $P=4\prod_{p\leq y}p$, where y is a parameter which will depend only on δ. Our first aim is to show that there is a reduced residue class $a \pmod P$ such that if $p\equiv a \pmod P$, then $T_p(x)\geq\frac{1}{10}$

for x in a set $\mathcal{S} \subseteq [0,1]$, where $|\mathcal{S}| > \frac{2}{3} - \delta$.

We define an arithmetic function a_n as follows: Let a_n be totally multiplicative, $a_p = \left(\frac{p}{3}\right)$ for $p \neq 3$, and $a_3 = -1$. Put

$$U(x) = \sum_{n=1}^{\infty} \frac{a_n}{n} \cos 2\pi n x.$$

Then $U(x)$ is related to $T_3(x)$, for

$$\begin{aligned} U(x) &= \sum_{k=0}^{\infty} \sum_{\substack{n=1 \\ 3^k \| n}}^{\infty} \frac{a_n}{n} \cos 2\pi n x \\ &= \sum_{k=0}^{\infty} \sum_{\substack{m=1 \\ 3 \nmid m, \\ n=3^k m}}^{\infty} \frac{(-1)^k \left(\frac{m}{3}\right)}{3^k m} \cos(2\pi 3^k m x) \\ &= \sum_{k=0}^{\infty} (-1)^k 3^{-k} T_3(3^k x). \end{aligned}$$

We use this formula to trace the behaviour of S_3 back to T_3, and hence to U. First, we note that, since

$$S_3(1/3) = \sum_{n \leq 1} \left(\frac{n}{3}\right) = \left(\frac{1}{3}\right) = 1,$$

and

$$S_3(2/3) = \sum_{n \leq 2} \left(\frac{n}{3}\right) = \left(\frac{1}{3}\right) + \left(\frac{2}{3}\right) = 1 + (-1) = 0.$$

we have:

$$S_3(x) = \begin{cases} 0 & \text{if } 0 \leq x < \frac{1}{3}, \\ 1 & \text{if } \frac{1}{3} \leq x < \frac{2}{3}, \\ 0 & \text{if } \frac{2}{3} \leq x < 1. \end{cases} \tag{3.2.1}$$

Define $\| \Theta \|$ as the distance from Θ to the nearest integer, $\| \Theta \| = \min_n | \Theta - n |$. Hence,

$$\text{for } \| x \| < \frac{1}{3},\ S_3(x) = 0, \text{ and so } T_3(x) = L_3(1) = \frac{\pi}{3\sqrt{3}}$$

and

$$\text{for } \| x \| > \frac{1}{3},\ S_3(x) = 1, \text{ and so } T_3(x) = L_3(1) - \frac{\pi}{\sqrt{3}} = \frac{-2\pi}{3\sqrt{3}}.$$

Thus for $\| x \| < \frac{1}{3}$, using that $\frac{-2\pi}{3\sqrt{3}} \leq T_3(y) \leq \frac{\pi}{3\sqrt{3}}$ for all $y \in \mathbb{R}$,

$$\begin{aligned} U(x) = \sum_{k=0}^{\infty} (-1)^k 3^{-k} T_3(3^k x) &= T_3(x) - \sum_{\substack{k \geq 1 \\ k \text{ odd}}} \frac{1}{3^k} T_3(3^k x) + \sum_{\substack{k \geq 2 \\ k \text{ even}}} \frac{1}{3^k} T_3(3^k x) \\ &\geq \frac{\pi}{3\sqrt{3}} - \sum_{\substack{k \geq 1 \\ k \text{ odd}}} \frac{1}{3^k} \frac{\pi}{3\sqrt{3}} + \sum_{\substack{k \geq 2 \\ k \text{ even}}} \frac{1}{3^k} \frac{-2\pi}{3\sqrt{3}} \\ &= \frac{\pi}{3\sqrt{3}} \left(1 - \sum_{\substack{k \geq 1 \\ k \text{ odd}}} \frac{1}{3^k} - 2 \sum_{\substack{k \geq 2 \\ k \text{ even}}} \frac{1}{3^k} \right) \\ &= \frac{\pi}{3\sqrt{3}} \left(1 - \left(\frac{1}{3} + \frac{2}{9} \right) \sum_{j \geq 0} \frac{1}{9^j} \right) \\ &= \frac{\pi}{3\sqrt{3}} \left(1 - \frac{5}{9} \left(\frac{1}{1 - 1/9} \right) \right) \\ &= \frac{\pi}{3\sqrt{3}} \left(1 - \frac{5}{8} \right) = \frac{\pi}{8\sqrt{3}} > 0.2. \end{aligned}$$

Therefore, for $\| x \| < \frac{1}{3}$,

$$U(x) > \frac{1}{5}. \tag{3.2.2}$$

Recall that by definition, $a_q = \left(\frac{q}{3}\right)$ where $q \neq 3$ and $a_3 = -1$. Now, we would like to show that there exist infinitely many primes $p \equiv 3 \bmod 4$ such that for q prime,

$$\left(\frac{q}{p}\right) = a_q = \begin{cases} \left(\frac{q}{3}\right) & \text{if } q \neq 3 \text{ and } q \leq y, \\ -1 & \text{if } q = 2, \\ -1 & \text{if } q = 3. \end{cases} \tag{3.2.3}$$

To do this, we will show that there exist infinitely many integers a such that

$$a \equiv 3 \bmod 8,$$

$$a \equiv 1 \bmod q, \text{ for } \begin{cases} q = 3, \text{ i.e. for } \left(\frac{3}{q}\right) = 1 \\ q \equiv 1 \bmod 12, \\ q \equiv 11 \bmod 12. \end{cases}$$

$$a \equiv n_q \bmod q \text{ for } \begin{cases} q \equiv 5 \bmod 12, \text{ i.e. for } \left(\frac{3}{q}\right) = -1 \\ q \equiv 7 \bmod 12. \end{cases}$$

where n_q is such that $\left(\frac{n_q}{q}\right) = -1$. The above is true by the Chinese Remainder Theorem. Now, Dirichlet's theorem on primes in an arithmetic progression implies that there exist infinitely many primes $p \equiv a \pmod{q}$. So, $\left(\frac{a}{q}\right) = \left(\frac{p}{q}\right)$. Now, by the Law of Quadratic Reciprocity,

$$\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)(-1)^{\frac{q-1}{2}}$$
$$\text{and } \left(\frac{q}{3}\right) = \left(\frac{3}{q}\right)(-1)^{\frac{q-1}{2}}.$$

So, when $\left(\frac{3}{q}\right) = 1$ then,

$$\left(\frac{a}{q}\right) = 1 \text{ since } a \equiv 1 \bmod q$$
$$\text{and } \left(\frac{p}{q}\right) = 1 \text{ since } p \equiv a \bmod q.$$

Similarly, when $\left(\frac{3}{q}\right) = -1$ then,

$$\left(\frac{a}{q}\right) = -1 \text{ since } a \equiv n_q \bmod q \text{ and } \left(\frac{n_q}{q}\right) = -1$$
$$\text{and } \left(\frac{p}{q}\right) = -1 \text{ since } p \equiv a \bmod q.$$

Putting the above together then gives $\left(\frac{q}{p}\right) = \left(\frac{q}{3}\right) = a_q$ for all primes $q \leq y$. Hence, for such p, $\left(\frac{n}{p}\right) = a_n$ for all $n \leq y$. This forces $T_p(x)$ to resemble $U(x)$,

$$\begin{aligned}
\int_0^1 (T_p(x) - U(x))^2 \, dx &= \int_0^1 \left(\sum_{n=1}^{\infty} \left(\frac{n}{p}\right) \frac{\cos 2\pi nx}{n} - \sum_{n=1}^{\infty} a_n \frac{\cos 2\pi nx}{n} \right)^2 dx \\
&= \int_0^1 \left(\sum_{n=1}^{\infty} \left(\left(\frac{n}{p}\right) - a_n \right) \frac{\cos 2\pi nx}{n} \right)^2 dx \\
&= \int_0^1 \sum_{n,m \geq 1} \left(\left(\frac{n}{p}\right) - a_n \right) \left(\left(\frac{m}{p}\right) - a_m \right) \frac{\cos 2\pi nx}{n} \frac{\cos 2\pi mx}{m} \, dx
\end{aligned}$$

And by (3.1.5), we have

$$\begin{aligned}\int_0^1 (T_p(x)-U(x))^2\,dx &= \frac{1}{2}\sum_{\substack{n,m\geq 1\\ m=n}}\left(\left(\frac{n}{p}\right)-a_n\right)\left(\left(\frac{m}{p}\right)-a_m\right)\frac{1}{mn}\\ &= \frac{1}{2}\sum_{n\geq 1}\left(\left(\frac{n}{p}\right)-a_n\right)^2\frac{1}{n^2}.\end{aligned}$$

Since $\left(\frac{n}{p}\right)=0,1$ or -1, $a_n=\left(\frac{n}{p}\right)$ for $n\leq y$ and $a_n=1$ or -1, $\left(\left(\frac{n}{p}\right)-a_n\right)^2=0,1$ or 4;

$$\int_0^1 (T_p(x)-U(x))^2\,dx \leq 2\sum_{n>y}\frac{1}{n^2}.$$

Further, since

$$\sum_{n>y}\frac{1}{n^2}\leq\int_y^\infty\frac{1}{t^2}\,dt=\frac{1}{y},$$

we get

$$\int_0^1 (T_p(x)-U(x))^2\,dx\leq\frac{2}{y}. \tag{3.2.4}$$

From (3.2.4) and the lower bound for $U(x)$ (3.2.2), we see that if $T_p(x)<\frac{1}{10}$ and $\parallel x\parallel<\frac{1}{3}$, then $\mid T_p(x)-U(x)\mid>\frac{1}{10}$. But in view of the above, the set $\mathcal{E}$ of x with this latter property satisfies

$$\mid\mathcal{E}\mid\int_0^1\frac{1}{100}\,dx<\int_0^1(T_p(x)-U(x))^2\,dx\leq\frac{2}{y}$$

and so

$$\mid\mathcal{E}\mid 10^{-2}\leq\frac{2}{y}.$$

Thus, if $y>\frac{200}{\delta}$, then

$$\frac{\mid\mathcal{E}\mid}{100}<\frac{2}{200}\delta$$

and so $\mid\mathcal{E}\mid<\delta$.

We have now established that for any $\delta>0$, there are numbers a, P, $(a,P)=1$, such that if $p\equiv a\pmod P$ then

$$S_p(x)\leq\frac{\sqrt{p}}{\pi}\left(L_p(1)-\frac{1}{10}\right)$$

on a set of measure $>\frac{2}{3}-\delta$. Thus if we can show that there are infinitely many primes $p\equiv a\pmod P$ for which $L_p(1)<\frac{1}{10}$, we will prove that $S_p(x)>0$ on a set of measure $<\frac{1}{3}+\delta$. Suppose that $p\equiv a\pmod P$, and that $\left(\frac{q}{p}\right)=-1$ for all primes q such that $y<q\leq e^y$. By quadratic reciprocity this will be the case

if $p \equiv b \pmod{Q}$, $Q = 4 \prod_{p \le e^y} p$, where b is chosen suitably. From Montgomery's article [23], we have the asymptotic average of $L_p(1)$ for primes $p \equiv b \pmod{Q}$ given by

$$\prod_{q \le y} \left(1 - \frac{a_q}{q}\right)^{-1} \prod_{y < q \le e^y} \left(1 + \frac{1}{q}\right)^{-1} \tag{3.2.5}$$

where $a_q = \left(\frac{q}{3}\right)$, for q a prime, $(q \neq 3)$ and $a_3 = -1$; which equals

$$\left(1 + \frac{1}{3}\right)^{-1} \prod_{q \le y} \left(1 - \left(\frac{q}{3}\right)\frac{1}{q}\right)^{-1} \prod_{y < q \le e^y} \left(1 + \frac{1}{q}\right)^{-1}$$

$$= \frac{3}{4} \prod_{q \le y} \left(1 - \left(\frac{q}{3}\right)\frac{1}{q}\right)^{-1} \prod_{y < q \le e^y} \left(1 + \frac{1}{q}\right)^{-1}.$$

As $y \to \infty$, the first product is $\sim L_3(1) = \frac{\pi}{3\sqrt{3}}$. We will obtain an approximation for the second product:

$$\prod_{y < q \le e^y} \left(1 + \frac{1}{q}\right)^{-1} = \frac{\prod_{q \le e^y} \left(1 + \frac{1}{q}\right)^{-1}}{\prod_{q \le y} \left(1 + \frac{1}{q}\right)^{-1}} = \frac{\prod_{q \le e^y} \left(1 - \frac{1}{q}\right)\left(1 - \frac{1}{q^2}\right)^{-1}}{\prod_{q \le y} \left(1 - \frac{1}{q}\right)\left(1 - \frac{1}{q^2}\right)^{-1}}.$$

Appealing to Merten's Theorem (1.4.1), for large y, we have

$$\prod_{y < q \le e^y} \left(1 + \frac{1}{q}\right)^{-1} \sim \frac{e^{-\gamma}/\ln e^y \prod_{q \le e^y} \left(1 - \frac{1}{q^2}\right)^{-1}}{e^{-\gamma}/\ln y \prod_{q \le y} \left(1 - \frac{1}{q^2}\right)^{-1}} = \frac{\ln y}{\ln e^y} \prod_{y < q \le e^y} \left(1 - \frac{1}{q^2}\right)^{-1}.$$

Further,

$$\log \prod_{y < q \le e^y} \left(1 - \frac{1}{q^2}\right)^{-1} = -\sum_{y < q \le e^y} \log\left(1 - \frac{1}{q^2}\right)^{-1} = \sum_{y < q \le e^y} \frac{1}{q^2} + \frac{1}{2q^4} + \dots$$

$$< \sum_{q > y} \frac{2}{q^2} < \int_y^\infty \frac{2}{t^2}\, dt = \frac{2}{y},$$

and so

$$\prod_{y < q \le e^y} \left(1 - \frac{1}{q^2}\right)^{-1} = 1 + O\left(\frac{1}{y}\right).$$

Therefore, we have

$$\prod_{y < q \le e^y} \left(1 + \frac{1}{q}\right)^{-1} \sim \frac{\log y}{y}.$$

Hence, if y is sufficiently large, $\epsilon > 0$, then an asymptotic average of $L_p(1)$ for primes $p \equiv b \pmod{Q}$ is approximately $\frac{\pi \log y}{4\sqrt{3y}} < \epsilon$. Choose $\epsilon = \frac{1}{10}$ and since we

have shown that on average, $L_p(1) < \frac{1}{10}$ for primes $p \equiv b \pmod{Q}$; it follows that $L_p(1) < \frac{1}{10}$ for infinitely many primes $p \equiv b \pmod{Q}$, and hence $S_p(x) > 0$ on a set of measure $< \frac{1}{3} + \delta$. This completes the proof of the first assertion of Theorem 0.0.1.

We now prove the second assertion of Theorem 0.0.1 which states that $\alpha(p) > \frac{1}{50}$ for all primes $p \equiv 3 \bmod 4$. We observe that for $p = 3$ (3.2.1), the assertion that $\alpha(p) > \frac{1}{50}$, is true. We assume $p \geq 7$. Applying Hölder's inequality and using the results of Lemmas 3.1.3 and 3.1.4, we have

$$\int_0^1 T_p(x)^2 \, dx \leq \left(\int_0^1 |T_p(x)| \, dx \right)^{\frac{2}{3}} \left(\int_0^1 T_p(x)^4 \, dx \right)^{\frac{1}{3}},$$

and so

$$\left(\int_0^1 |T_p(x)| \, dx \right)^2 \geq \frac{(\frac{\pi^2}{12}(1 - p^{-2}))^3}{\frac{19\pi^4}{240}} = \frac{5\pi^2}{19 \times 36} \left(1 - \frac{1}{p^2}\right)^3.$$

Therefore,

$$\int_0^1 |T_p(x)| \, dx \geq \frac{\pi}{6} \sqrt{\frac{5}{19}} \left(1 - \frac{1}{p^2}\right)^{3/2} \tag{3.2.6}$$

and since $p \geq 7$,

$$\int_0^1 |T_p(x)| \, dx \geq \frac{\pi}{6} \sqrt{\frac{5}{19}} \left(\frac{48}{49}\right)^{3/2} > 0.26 = \frac{13}{50}. \tag{3.2.7}$$

We know that

$$\int_0^1 T_p(x) \, dx = \sum_{n=1}^{\infty} \left(\frac{n}{p}\right) \frac{1}{n} \int_0^1 \cos 2\pi n x \, dx = \sum_{n=1}^{\infty} \left(\frac{n}{p}\right) \frac{1}{n} \left[\frac{\sin 2\pi n x}{2\pi n} \right]_0^1 = 0,$$

therefore,

$$\int\limits_{\substack{0<x<1 \\ x:T_p(x)<0}} -T_p(x) \, dx = \frac{1}{2} \int_0^1 (|T_p(x)| - T_p(x)) \, dx = \frac{1}{2} \int_0^1 (|T_p(x)|) \, dx,$$

and by (3.2.7),

$$\int\limits_{\substack{0<x<1 \\ x:T_p(x)<0}} -T_p(x) \, dx > \frac{13}{100}. \tag{3.2.8}$$

Then, by Cauchy's inequality,

$$\int\limits_{\substack{0<x<1\\ x:T_p(x)<0}} -T_p(x)\,dx \le \left(\int\limits_{\substack{0<x<1\\ x:T_p(x)<0}} 1\,dx\right)^{\frac{1}{2}} \left(\int_0^1 T_p(x)^2\,dx\right)^{\frac{1}{2}}$$

and using (3.2.8),

$$\frac{13}{100} < \left(\int\limits_{\substack{0<x<1\\ x:T_p(x)<0}} 1\,dx\right)^{\frac{1}{2}} \left(\int_0^1 T_p(x)^2\,dx\right)^{\frac{1}{2}}.$$

Then appealing to Lemma 3.1.3 gives:

$$\begin{aligned}\int\limits_{\substack{0<x<1\\ x:T_p(x)<0}} 1\,dx &> \frac{(13/100)^2}{\int_0^1 T_p(x)^2\,dx} = \frac{(13/100)^2}{\frac{\pi^2}{12}(1-p^{-2})} \\ &= \left(\frac{13}{100}\right)^2 \frac{12}{\pi^2}(1-p^{-2})^{-1} > \left(\frac{13}{100}\right)^2 \frac{12}{\pi^2} > \frac{1}{50}.\end{aligned}$$

Thus, we find that $T_p(x) < 0$ on a set of measure greater than $\frac{1}{50}$. However, if $T_p(x) < 0$, then $S_p(x) > \frac{\sqrt{p}}{\pi}(L_p(1)) > 0$, since $L_p(1)$ is a positive function. This then proves the assertion that $\alpha(p) > \frac{1}{50}$ for all primes $p \equiv 3 \bmod 4$. □

3.3. A Lower Bound for $\alpha(p)$

We observe that the second assertion of Montgomery's theorem can be improved, such that we obtain:

Theorem 3.3.1. *For all primes $p \equiv 3 \bmod 4$, $\alpha(p) > \frac{5437}{250000}$. For all primes $p \equiv 3 \bmod 4$, if $p \ge q$ (q a prime), then $\alpha(p) > \frac{5}{228}\left(1-\frac{1}{q^2}\right)^3$.*

Proof. From (0.0.4) and Lemma 3.1.1, we have

$$\alpha(p) = \int\limits_{\substack{0<x<1\\ x:S_p(x)>0}} 1\,dx = \int\limits_{\substack{0<x<1\\ x:L_p(1)-T_p(x)>0}} 1\,dx > \int\limits_{\substack{0<x<1\\ x:T_p(x)<0}} 1\,dx.$$

We will first prove the second assertion of Theorem 3.3.1. Thus we will show that

$$\int\limits_{\substack{0<x<1\\ x:T_p(x)<0}} 1\,dx > \frac{5}{228}\left(1-\frac{1}{q^2}\right)^3 \tag{3.3.1}$$

and since for $q \geq 821$,

$$\frac{5}{228} - \frac{5}{228}\left(1 - \frac{1}{q^2}\right)^3 < 10^{-7},$$

we can say that for $p \geq q \geq 821$

$$\int\limits_{\substack{0<x<1 \\ x:T_p(x)<0}} 1 \, dx > \frac{5}{228} - 10^{-7}. \tag{3.3.2}$$

Taking $p \geq q$, and using the same approach as Montgomery, we will prove (3.3.1). If $p \geq q$, (3.2.6) gives

$$\int_0^1 |T_p(x)| \, dx \geq \frac{\pi}{6}\sqrt{\frac{5}{19}}\left(1 - \frac{1}{q^2}\right)^{3/2}.$$

Hence, we obtain (as in (3.2.8)),

$$\int\limits_{\substack{0<x<1 \\ x:T_p(x)<0}} -T_p(x) \, dx > \frac{\pi}{12}\sqrt{\frac{5}{19}}\left(1 - \frac{1}{q^2}\right)^{3/2}.$$

Next, we apply Cauchy's inequality, and obtain the result:

$$\int\limits_{\substack{0<x<1 \\ x:T_p(x)<0}} 1 \, dx > \frac{12}{\pi^2}\left(\frac{\pi}{12}\sqrt{\frac{5}{19}}\left(1 - \frac{1}{q^2}\right)^{3/2}\right)^2 = \frac{5}{228}\left(1 - \frac{1}{q^2}\right)^3.$$

Thus, we find that if $p \geq q$, $T_p(x) < 0$ on a set of measure greater than $\frac{5}{228}\left(1 - \frac{1}{q^2}\right)^3$. Owing to the fact that $L_p(1)$ is always positive, it follows that $\alpha(p) > \frac{5}{228}\left(1 - \frac{1}{q^2}\right)^3$ for $p \geq q$. Further for $q \geq 821$, we have by (3.3.2) that $\alpha(p) > \frac{5}{228} - 10^{-7}$.

We now prove the first assertion of Theorem 3.3.1 which states that for all primes $p \equiv 3 \bmod 4$, $\alpha(p) > \frac{5437}{250000}$. We observe that for the first three primes $p \equiv 3 \bmod 4$, (i.e $p = 3, 7, 11$), $S_p(x) \geq 0$ for all values of $x \in [0, 1]$. Therefore, our assertion clearly holds true for the primes $p = 3, 7$ and 11. Taking $q = 19$ in (3.3.1), directly gives $\alpha(p) > \frac{5437}{250000}$ for all $p > 11$. □

Chapter 4

PRELIMINARY LEMMAS

In this chapter we collect together some technical lemmas that will form the basis of the proof of Theorem 0.0.2 which will be given in Chapter 5. The first three of these lemmas are simple analytical deductions of existing results of analytic number theory. The next two however follow a combinatorial approach in which we estimate the maximum and minimum possible size of a set of real numbers satisfying certain conditions.

4.1. LEMMAS

The following lemma shows that we can truncate the Fourier series (whose Fourier coefficients are given by $\left(\frac{n}{p}\right)\frac{1}{n}$), up to an error term. The error term is determined by our choice of the length of the truncated sum. This lemma uses the Pólya-Vinogradov inequality and Fourier expansions for both infinite and finite character sums as were given by Pólya [**27**] for all primitive characters.

Lemma 4.1.1. *For p a prime congruent to* 3 mod 4 *and $H \geq \sqrt{p}\log^2 p$,*

$$\sum_{n\geq 1}\left(\frac{n}{p}\right)\frac{\cos(2\pi nx)}{n} = \sum_{n\leq H}\left(\frac{n}{p}\right)\frac{\cos(2\pi nx)}{n} + O\Big(\frac{\sqrt{p}\log p}{H}\Big). \tag{4.1.1}$$

PROOF. Recall from (3.1.2) that for $p \equiv 3 \bmod 4$, $S_p(x)$ has the Fourier expansion given by:

$$S_p(x) = \frac{\sqrt{p}}{\pi}\left(L_p(1) - \sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}\right).$$

Further, from Lemma 1 of [**24**], we observe that Pólya also gave the following truncated expansion:

$$\sum_{n\leq\alpha}\chi(n)=\frac{\tau(\chi)}{2\pi i}\sum_{\substack{n=-H\\ n\neq 0}}^{H}\frac{\overline{\chi}(n)}{n}(1-e(-n\alpha/q))+O(1)+O\Big(\frac{q\log q}{H}\Big) \tag{4.1.2}$$

where χ is a primitive character $\pmod q$ and $\tau(\chi)$ is the Gauss sum.

Consider $\chi(n)=\left(\frac{n}{p}\right)$ where $p\equiv 3 \bmod 4$. We know that for $p\equiv 3 \bmod 4$, $\tau(\chi)=i\sqrt{p}$ and $\left(\frac{-n}{p}\right)=-\left(\frac{n}{p}\right)$. Rewriting (4.1.2) for $p\equiv 3 \bmod 4$ gives

$$S_p(x)=\frac{\sqrt{p}}{\pi}\left(\sum_{n\leq H}\left(\frac{n}{p}\right)\frac{1}{n}-\sum_{n\leq H}\left(\frac{n}{p}\right)\frac{\cos(2\pi nx)}{n}+O\Big(\frac{\sqrt{p}\log p}{H}\Big)+O\Big(\frac{1}{\sqrt{p}}\Big)\right). \tag{4.1.3}$$

Now, comparing (3.1.2) and (4.1.3), we have

$$\begin{aligned}&\sum_{n\geq 1}\left(\frac{n}{p}\right)\frac{1}{n}-\sum_{n\geq 1}\left(\frac{n}{p}\right)\frac{\cos(2\pi nx)}{n}\\ =\;&\sum_{n\leq H}\left(\frac{n}{p}\right)\frac{1}{n}-\sum_{n\leq H}\left(\frac{n}{p}\right)\frac{\cos(2\pi nx)}{n}+O\Big(\frac{\sqrt{p}\log p}{H}\Big)+O\Big(\frac{1}{\sqrt{p}}\Big).\end{aligned} \tag{4.1.4}$$

Note that by Stieltjes integral,

$$\sum_{n>H}\left(\frac{n}{p}\right)\frac{1}{n}=\int_H^\infty\frac{1}{t}\,d\left(\sum_{n\leq t}\left(\frac{n}{p}\right)\right) \tag{4.1.5}$$

and by the Pólya-Vinogradov Inequality, $\Big|\sum_{n\leq t}\left(\frac{n}{p}\right)\Big|\leq\sqrt{p}\log p$ and by integration by parts of (4.1.5),

$$\sum_{n>H}\left(\frac{n}{p}\right)\frac{1}{n}=O\Big(\frac{\sqrt{p}\log p}{H}\Big).$$

Using this result in (4.1.4) and taking $H\geq\sqrt{p}\log^2 p$ the result follows. □

In the next lemma we give an upper bound for an infinite sum of the reciprocals of integers with small prime factors. The corollary that follows as a result of this lemma gives a more specific result where we consider all integers with prime factors smaller than $\log^2 p$, where p is a large prime. This result will be directly applied in our later work.

Lemma 4.1.2. *For any σ, $0 < \sigma < 1$, we have*

$$\sum_{\substack{n>H \\ P(n)\leq y}} \frac{1}{n} \leq \frac{c_{y,\sigma}}{H^{1-\sigma}} \tag{4.1.6}$$

where $c_{y,\sigma} = \prod_{\substack{q \text{ prime} \\ q\leq y}} \left(1 - \frac{1}{q^{\sigma}}\right)^{-1}$.

PROOF. The result is a simple consequence of Rankin's trick, whereby

$$\sum_{\substack{n>H \\ P(n)\leq y}} \frac{1}{n} \leq \sum_{\substack{n\geq 1 \\ P(n)\leq y}} \frac{1}{n}\left(\frac{n}{H}\right)^{1-\sigma} = \frac{1}{H^{1-\sigma}} \sum_{\substack{n\geq 1 \\ P(n)\leq y}} \frac{1}{n^{\sigma}} = \frac{1}{H^{1-\sigma}} \prod_{\substack{q \text{ prime} \\ q\leq y}} \left(1-\frac{1}{q^{\sigma}}\right)^{-1} = \frac{c_{y,\sigma}}{H^{1-\sigma}}.$$

□

Corollary 4.1.3. *For $y \leq \log^2 p$, we have*

$$\sum_{\substack{n>H \\ P(n)\leq y}} \frac{1}{n} \leq \frac{p^{o(1)}}{\sqrt{H}}. \tag{4.1.7}$$

PROOF. Taking $\sigma = \frac{1}{2}$ in (4.1.6), gives

$$\sum_{\substack{n>H \\ P(n)\leq y}} \frac{1}{n} \leq \frac{c_{y,\frac{1}{2}}}{\sqrt{H}}.$$

From Lemma 4.1.2, we have that $c_{y,\frac{1}{2}} = \prod_{\substack{q \text{ prime} \\ q\leq y}} \left(1 - \frac{1}{q^{\frac{1}{2}}}\right)^{-1}$. We will evaluate $c_{y,\frac{1}{2}}$ for $y \leq \log^2 p$:

$$\log c_{y,\frac{1}{2}} = -\sum_{\substack{q \text{ prime} \\ q\leq y}} \log\left(1 - \frac{1}{q^{\frac{1}{2}}}\right) = O\Big(\sum_{\substack{q \text{ prime} \\ q\leq y}} \frac{1}{q^{\frac{1}{2}}}\Big),$$

and by partial summation and the Prime Number Theorem,

$$\sum_{\substack{q \text{ prime} \\ q\leq y}} \frac{1}{q^{\frac{1}{2}}} = \int_2^y \frac{1}{t^{\frac{1}{2}}}\, d\Big(\sum_{\substack{q \text{ prime} \\ q\leq t}} 1\Big) = \int_2^y \frac{1}{t^{\frac{1}{2}}}\, d(\pi(t)) = O(\frac{\sqrt{y}}{\log y})$$

giving $c_{y,\frac{1}{2}} = e^{O(\frac{\sqrt{y}}{\log y})}$. Hence for $y \leq \log^2 p$, $c_{y,\frac{1}{2}} = e^{O(\frac{\log p}{\log\log p})} = p^{o(1)}$. □

Let P be a finite, ordered set of real numbers such that the cardinality of the positive elements of P is known. Lemma 4.1.4 states that if there exists any

ordered set Q of the same cardinality as P and such that the elements of Q are close enough to the elements of P, then we can deduce an upper bound for the cardinality of the positive elements of Q. Corollary 4.1.5 and Lemma 4.1.6 that follow are based on exactly the same approach as this lemma (Lemma 4.1.4). While Corollary 4.1.5 gives a lower bound for the cardinality of the negative elements of the set Q, Lemma 4.1.6 gives an upper bound for the same.

Lemma 4.1.4. *Let $N \in \mathbb{N}$, $a_i \in \mathbb{R}$ where $0 \leq i \leq N-1$. Reorder the a_i's, as follows:*

$$a_{r-1} \leq \cdots \leq a_1 \leq a_0 < 0 \leq a_r \leq \cdots \leq a_{N-1}.$$

Fix $c > 0$. Let k be that integer less than r such that

$$\sum_{i=0}^{k-1} a_i^2 < cN \leq \sum_{i=0}^{k} a_i^2 \tag{4.1.8}$$

if such an integer k exists, otherwise let $k = r$. If $b_0, \ldots, b_{N-1} \in \mathbb{R}$ are such that $\frac{1}{N} \sum_{i=0}^{N-1} (a_i - b_i)^2 < c$, then

$$|\{i : b_i \geq 0\}| \leq (N - r) + k.$$

PROOF. Call $A = \{i : b_i \geq 0\}$ and $m = |A|$. We can write our set A as:

$$A = A_+ \cup A_-$$

where

$$A_+ = A \cap \{i : a_i \geq 0\} = \{i : b_i, a_i \geq 0\} \subset \{i : a_i \geq 0\}.$$

and

$$A_- = A \cap \{i : a_i < 0\} = \{i : a_i < 0 \leq b_i\}.$$

And thus,

$$A \subset \{i : a_i \geq 0\} \cup \{i : a_i < 0 \leq b_i\}. \tag{4.1.9}$$

Further, since

$$\frac{1}{N} \sum_{\substack{i:a_i<0 \\ b_i \geq 0}} (a_i - b_i)^2 \leq \frac{1}{N} \sum_{i=0}^{N-1} (a_i - b_i)^2,$$

using the hypothesis that $b_0, \ldots, b_{N-1}$ are such that $\frac{1}{N}\sum_{i=0}^{N-1}(a_i - b_i)^2 < c$, we have

$$\frac{1}{N}\sum_{\substack{i:a_i<0\\ b_i\geq 0}}(a_i - b_i)^2 < c.$$

Also,

$$(a_i - b_i)^2 = a_i^2 + b_i^2 - 2a_ib_i \geq a_i^2 \qquad \text{for } a_i < 0, b_i \geq 0$$

and so:

$$\frac{1}{N}\sum_{i\in A_-} a_i^2 = \frac{1}{N}\sum_{\substack{i:a_i<0\\ b_i\geq 0}} a_i^2 \leq \frac{1}{N}\sum_{\substack{i:a_i<0\\ b_i\geq 0}}(a_i - b_i)^2 < c. \tag{4.1.10}$$

Note from the hypothesis of this lemma that we have the a_i's ordered as follows:

$$a_{r-1} \leq a_{r-2} \leq \cdots \leq a_1 \leq a_0 < 0 \leq a_r \leq \cdots \leq a_{N-1}.$$

Suppose

$$A_- = \{0 \leq i_0 < i_1 < \cdots < i_{l-1} \leq r-1\} \text{ so that } i_t \geq t \text{ where } 0 \leq t \leq l-1.$$

The ordering of the a_i's gives:

$$|a_{r-1}| \geq |a_{r-2}| \geq \cdots \geq |a_1| \geq |a_0| > 0$$

and as $r - 1 \geq i_t \geq t \geq 0$, we have $|a_{i_t}| \geq |a_t|$. Therefore:

$$\sum_{t=0}^{l-1} a_{i_t}^2 \geq \sum_{t=0}^{l-1} a_t^2.$$

Further, we have supposed that $|A_-| = l$ and so:

$$\sum_{i\in A_-} a_i^2 = \sum_{t=0}^{l-1} a_{i_t}^2 \geq \sum_{t=0}^{l-1} a_t^2. \tag{4.1.11}$$

Owing to (4.1.10) and (4.1.11), we have:

$$\frac{1}{N}\sum_{t=0}^{l-1} a_t^2 \leq \frac{1}{N}\sum_{i\in A_-} a_i^2 < c$$

Our hypothesis tells us that either there exists a $k < r$, such that

$$\frac{1}{N}\sum_{i=0}^{k-1} a_i^2 < c \leq \frac{1}{N}\sum_{i=0}^{k} a_i^2.$$

or else we let $k = r$. Thus, it follows that

$$\frac{1}{N}\sum_{t=0}^{l-1} a_t^2 \leq \frac{1}{N}\sum_{i=0}^{k-1} a_i^2 < c$$

and so $l \leq k$. Hence $|A_-| \leq k$. From (4.1.9), we then have:

$$m = |A| \leq |\{i : a_i \geq 0\}| + |A_-| = (N - r) + |A_-| \leq (N - r) + k.$$

□

Lemma 4.1.4 immediately leads to the following corollary:

Corollary 4.1.5. *Let $N \in \mathbb{N}$, $a_i \in \mathbb{R}$ where $0 \leq i \leq N-1$. Fix $c > 0$. If $b_0, \ldots, b_{N-1} \in \mathbb{R}$ are such that $\frac{1}{N}\sum_{i=0}^{N-1}(a_i - b_i)^2 < c$, then*

$$|\{i : b_i < 0\}| \geq r - k,$$

where k and r are as defined in Lemma 4.1.4.

Lemma 4.1.6. *Let $N \in \mathbb{N}$, $a_i \in \mathbb{R}$ where $0 \leq i \leq N-1$. Reorder the a_i's, as follows:*

$$a_{N-1} \leq \cdots \leq a_{r+1} \leq a_r < 0 \leq a_0 \leq a_1 \cdots \leq a_{r-1}.$$

Fix $c > 0$. Let k be that integer less than r such that

$$\sum_{i=0}^{k-1} a_i^2 < cN \leq \sum_{i=0}^{k} a_i^2$$

if such an integer exists, otherwise let $k = r$. Then, if $b_0, \ldots, b_{N-1} \in \mathbb{R}$ are such that $\frac{1}{N}\sum_{i=0}^{N-1}(a_i - b_i)^2 < c$, then

$$|\{i : b_i < 0\}| \leq (N - r) + k.$$

PROOF. The proof for this lemma follows in exactly the same way as the proof for Lemma 4.1.4. Note that in this case $r = |\{i : a_i \geq 0\}|$. □

Chapter 5

THE DISTRIBUTION OF $S_p(x)$ FOR $p \equiv 3 \pmod 4$

5.1. INTRODUCTION

In Chapter 3, we studied in detail Hugh L. Montgomery's [**23**] 1974 result in which he used analytic techniques to show how frequently the sum $S_p(x) = \sum_{n \leq px} \left(\frac{n}{p}\right)$ is positive (where p is a prime congruent to 3 (mod 4)). Our goal in this chapter will be to study how frequently the same sum $S_p(x)$ is negative, and hence we will prove Theorem 0.0.2. In Chapter 3 we saw that Montgomery's main idea was to model the Legendre symbol by another totally multiplicative arithmetic function. Even though we will employ this same idea, our choice of function will be different. While Montgomery's modelling function was dependent on the prime 3, ours will be essentially independent, allowing us to consider all possible sequences of the Legendre symbol. In order to derive sensible results from all these possible values, we will use analytic techniques, combinatorial ideas and perform a series of numerical computations.

As is evident from Theorem 0.0.2, assuming GRH, we showed that for a positive proportion of primes $p \equiv 3 \mod 4$, $S_p(x)$ is negative for at most 71.5 percent of the values of $x \in [0, 1]$; and of these primes, there is at least one prime for which $S_p(x)$ is negative for at most 36.9 percent of the values of $x \in [0, 1]$. Also based on GRH, we further showed that for a positive proportion of primes $p \equiv 3 \mod 4$, $S_p(x)$ is negative for at least 25.4 percent of the values of $x \in [0, 1]$; and of these

primes, there is at least one prime for which $S_p(x)$ is negative for at least 65.9 percent of the values of $x \in [0,1]$. However, Montgomery's results are unconditional and imply that $S_p(x)$ can be negative for nearly 2/3 of the values of $x \in [0,1]$, but not for almost all values of $x \in [0,1]$. Thus, even though we were unable to improve this result of Montgomery's; we observe that one of the Legendre symbol sequences which correspond to our result matches the $\{1,-1\}$ sequence that would be generated by Montgomery's choice of modelling function. This suggests that Montgomery's result is in fact the best possible.

Our results were obtained via algorithms which were coded in C++, and run on a two processor Xeon3 2.8GHz machine with 2GB RAM DDR266. The results obtained can be further improved (to match Montgomery's result) by using machines with higher RAM or even by employing distributed processing.

For simplicity, we will define the following measure:

$$\mu(p) := |\{x \in [0,1] : S_p(x) < 0\}|. \tag{5.1.1}$$

Recall that in the notation of Theorem 0.0.2, $\alpha(p)$ was defined to be the measure of the set of $x \in [0,1]$ for which $S_p(x)$ is positive for primes $p \equiv 3 \pmod 4$ (see (0.0.4)). Note then that $\mu(p) = 1 - \alpha(p)$. To give a brief overview of how we will prove Theorem 0.0.2, we will begin by recalling from Lemma 3.1.1, that for primes $p \equiv 3 \pmod 4$, the Fourier expansion of $S_p(x)$ as given by Pólya gives

$$S_p(x) = \frac{\sqrt{p}}{\pi}\left(L_p(1) - \sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}\right)$$

where $L_p(1) = \sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{1}{n}$. Taking the Euler product we can prove that $L_p(1)$ is always positive for all primes p. We will see later that, owing to the above Fourier expansion of $S_p(x)$, the measure defined as $\mu(p)$, can be bounded by the measure of the set of $x \in [0,1]$ for which the Fourier Series $-\sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$ is negative. And thus, we will calculate explicit upper and lower bounds for this measure.

To do so, we will approximate this Fourier series by another series $-\sum_{\substack{n=1 \\ q|n \Rightarrow q \le y}}^{\infty}\lambda(n)\frac{\cos 2\pi nx}{n}$, where λ is a totally multiplicative arithmetic function used to model the Legendre symbol $\left(\frac{\cdot}{p}\right)$. We will define our function $\lambda(q)$ to be equal

to 1 or -1 for all primes $q \leq y$ (where y is a parameter) and equal to zero for primes $q > y$. Moreover, since λ is independent of the prime p; we will be able to compute values for the corresponding truncated series, for a fixed y, for all possible λ. These computations will be performed based on Algorithm A.1.1 and the first part of Algorithms A.1.2 and A.1.4. These algorithms are available in Appendix A.

Then, given λ and y, we will consider those primes $p \equiv 3 \pmod 4$ for which $\left(\frac{q}{p}\right) = \lambda(q)$ for all primes $q \leq y$. In this case, $-\sum\limits_{\substack{n=1 \\ q|n \Rightarrow q \leq y}}^{\infty} \lambda(n)\frac{\cos 2\pi nx}{n}$ will serve as a good model for the Fourier series $-\sum\limits_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$. Via Lemma 5.4.1, we will prove that these two series are very close to each other. Owing to the closeness of these two series, and using the numerically computed values of the truncated series of $-\sum\limits_{\substack{n=1 \\ q|n \Rightarrow q \leq y}}^{\infty} \lambda(n)\frac{\cos 2\pi nx}{n}$; we will appeal to the combinatorial lemmas of Chapter 4 to obtain bounds for the measure of $x \in [0,1]$ for which $-\sum\limits_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}$ is negative. More precisely, on the basis of these lemmas, we will formulate two main theorems, Theorems 5.2.1 and 5.2.3. While Theorem 5.2.1 will provide a method to obtain a lower bound for our measure, Theorem 5.2.3 will provide a method to obtain an upper bound for it.

These theorems will serve as a basis for formulating the latter parts of Algorithms A.1.2 and A.1.4 and the iterative Algorithms A.1.3 and A.1.5 (refer to Appendix A). These will in turn be used to perform a series of numerical computations and yield the results for the desired bounds. The explicit results obtained at each iteration have been tabulated and are available in Appendix B.

5.1.1. Definitions and Notation

Here we will familiarize ourselves with certain notation and definitions that will be central to our work in this chapter.
We will denote the Fourier series whose measure will be studied as

$$T_p(x) = -\sum_{n=1}^{\infty}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}. \tag{5.1.2}$$

Remark 5.1.1. *Note that the above definition of $T_p(x)$ is different from the one used for Montgomery's proof in Chapter 3 (see (3.1.3)).*

Then, we note here that in view of the Fourier expansion of $S_p(x)$ as given by Pólya (Lemma 3.1.1) we can write

$$S_p(x) = \frac{\sqrt{p}}{\pi}\left(L_p(1) + T_p(x)\right).$$

We put

$$\beta(p) := |\{x \in [0,1] : T_p(x) < 0\}|. \tag{5.1.3}$$

Definition 5.1.2. *For q a prime number and y a parameter, let $\lambda(q)$ be an arithmetic function that is totally multiplicative which satisfies $\lambda(q) = -1$ or 1 if $q \leq y$, and $\lambda(q) = 0$ if $q > y$.*

Remark 5.1.3. *This is the function which we will use to model the Legendre symbol (as was explained in the description that preceded the definitions).*

Definition 5.1.4. *$P(n)$ is defined as the largest prime factor of n.*

Definition 5.1.5. *Given λ and y as above, $T_{\lambda,y}(x)$ is defined as*

$$T_{\lambda,y}(x) := -\sum_{\substack{n\geq 1 \\ P(n)\leq y}} \lambda(n)\frac{\cos 2\pi n x}{n}. \tag{5.1.4}$$

Remark 5.1.6. *We will compute values for a truncated series of $T_{\lambda,y}(x)$, and use these values to approximate upper and lower bounds for $\beta(p)$.*

Having presented the above notation and definitions, the next section will now record the main theorems, Theorems 5.2.1 and 5.2.3. The proofs of these theorems will be given later in section 5.4 after we state some preparatory lemmas in section 5.3.

5.2. The Main Theorems

The main theorems assume GRH and give upper and lower bounds for $\beta(p)$. For fixed λ and y, and N such that $P(N) \leq y$ and large enough (explicitly stated in the theorem), we will compute values for a truncated series of $T_{\lambda,y}(\frac{j}{N})$ for all $0 \leq j \leq N-1$. (We note here that the first condition on N serves as a technical requirement in our work, and in fact we can do without this condition simply

by doing more sophisticated analysis). Next, we will consider primes p for which $\lambda(q) = \left(\frac{q}{p}\right)$ for all primes $q \leq y$. Under this condition, the second moments of the difference between $T_p(x)$ and $T_{\lambda,y}(x)$ will be very small. Owing to the resulting similarity between these two functions, we expect the measure of the set $x \in [0,1]$ for which $T_{\lambda,y}(x)$ is negative, to be close to the measure $\beta(p)$. Hence, we will be able to bound $\beta(p)$.

Given that we are able to analytically bound the second moments of $T_p(x) - T_{\lambda,y}(x)$; assuming GRH, we will deduce a bound for the corresponding discrete second moments of $T_p(\frac{j}{N}) - T_{\lambda,y}(\frac{j}{N})$, under the assumption that the error term is very small. In section 5.4 we will prove that this is true for a positive proportion of primes $p \equiv 3 \bmod 4$ for which $\lambda(q) = \left(\frac{q}{p}\right)$ for all primes $q \leq y$, provided y is large. However, during our computations, we will assume that the error term is very small for all $y \geq 29$. Having established a bound for the discrete second moments, we observe that if we consider our numerically computed values of the truncated series of $T_{\lambda,y}(\frac{j}{N})$ for all $0 \leq j \leq N-1$ to be the given set in the combinatorial lemmas of Chapter 4, our problem is analogous to the scenario as given by these lemmas. Thus, for a fixed λ, y a simple application of the corollary of Lemma 4.1.4 gives Theorem 5.2.1 and hence an explicit formula for a lower bound for $\beta(p)$; and applying Lemma 4.1.6 gives Theorem 5.2.3 and hence an explicit formula for an upper bound for $\beta(p)$.

We will now state the theorems. Their proofs will follow later in section 5.4 after we have established some results that will serve as the basis of these proofs.

Theorem 5.2.1. *Assume GRH. Fix λ and y and define $c_y = \frac{1}{2}\left(\frac{\pi^2}{6} - \prod_{q \leq y}\left(1 - \frac{1}{q^2}\right)^{-1}\right)$. Suppose that $\int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\, dx < c_y$ for all primes $p \equiv 3 \bmod 4$ such that $\left(\frac{q}{p}\right) = \lambda(q)$ for all primes $q \leq y$. Fix $k_1, k_2 > 0$ and suppose that*

$$\left|\frac{1}{N}\sum_{j=0}^{N-1}\left(T_p\left(\frac{j}{N}\right) - T_{\lambda,y}\left(\frac{j}{N}\right)\right)^2 - \int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\, dx\right| < \frac{k_1\sqrt{\log y}}{\sqrt{y}} + \frac{k_2}{\sqrt{N}} \tag{5.2.1}$$

holds for all N (such that $P(N) \leq y$) satisfying

$$N > \left(\frac{100k_2}{c_y - \int_0^1 (T_p(x) - T_{\lambda,y}(x))^2 \, dx} \right)^2. \tag{5.2.2}$$

Define $a_0, a_1, \cdots, a_{N-1}$ to be the set of numbers $\{T_{\lambda,y}(\frac{j}{N}) : 0 \leq j \leq N-1\}$, ordered as follows

$$a_{r-1} \leq \cdots \leq a_1 \leq a_0 < 0 \leq a_r \leq \cdots \leq a_{N-1}$$

and k to be the smallest integer such that

$$\sum_{j=0}^{k} a_j^2 \geq N\Big(c_y + \frac{k_1\sqrt{\log y}}{\sqrt{y}}\Big). \tag{5.2.3}$$

Then,

$$\beta(p) \geq \frac{r-k}{N} \tag{5.2.4}$$

for a positive proportion of primes $p \equiv 3 \bmod 4$ for which $\left(\frac{q}{p}\right) = \lambda(q)$ for all primes $q \leq y$.

Remark 5.2.2. *Note that $T_{\lambda,y}(x)$ is close to $T_p(x)$ since $c_y \to 0$ as $y \to \infty$. We observe via (3.2.4) that Montgomery obtained a similar proximity between $T_p(x)$ and the Fourier series $U(x)$.*

Theorem 5.2.3. *Assume the same hypotheses as in Theorem 5.2.1. Define $a_0, a_1, \cdots, a_{N-1}$ to be the set of numbers $\{T_{\lambda,y}(\frac{j}{N}) : 0 \leq j \leq N-1\}$, ordered as follows*

$$a_{N-1} \leq \cdots \leq a_{r+1} \leq a_r < 0 \leq a_0 \leq a_1 \cdots \leq a_{r-1}.$$

and k to be the smallest integer such that

$$\sum_{j=0}^{k} a_j^2 \geq N\Big(c_y + \frac{k_1\sqrt{\log y}}{\sqrt{y}}\Big).$$

Then,

$$\beta(p) \leq \frac{(N-r)+k}{N} \tag{5.2.5}$$

for a positive proportion of primes $p \equiv 3 \bmod 4$ for which $\left(\frac{q}{p}\right) = \lambda(q)$ for all primes $q \leq y$.

We remark here, that the final computed values obtained for the upper and lower bound of $\beta(p)$ suggest that for large y, there exists a prime $p \equiv 3 \bmod 4$ for which the upper and lower bounds for $\beta(p)$ are close to each other.

5.3. Preparatory Lemmas

This section records some lemmas that will help in establishing the hypotheses of the main theorems which were stated in the previous section. Before we state and prove these lemmas, we will introduce certain definitions.

5.3.1. Definitions

Definition 5.3.1. *For p a prime, $n \in \mathbb{Z}$, y a parameter and $P(n)$ as defined earlier,*

$$\left(\frac{n}{p}\right)_y := \begin{cases} 0 & \text{if } P(n) \leq y, \\ \left(\frac{n}{p}\right) & \text{otherwise} \end{cases}$$

where $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol.

Definition 5.3.2. *For $\left(\frac{\cdot}{p}\right)_y$ as defined above, we define $T^*_{p,y}(x)$ as:*

$$T^*_{p,y}(x) := -\sum_{n\geq 1} \left(\frac{n}{p}\right)_y \frac{\cos(2\pi n x)}{n}. \tag{5.3.1}$$

Remark 5.3.3. *For λ and y fixed, if $\lambda(q) = \left(\frac{q}{p}\right)$ for all primes $q \leq y$, then we observe that for primes $p \equiv 3 \bmod 4$,*

$$T^*_{p,y}(x) = T_p(x) - T_{\lambda,y}(x).$$

Definition 5.3.4. *Fix λ and y. Let $\mathcal{P}$ be the set of primes $p \equiv 3 \mod 4$ for which $\left(\frac{q}{p}\right) = \lambda(q)$ for all primes $q \leq y$ and for x large, $\mathcal{P}_x := \mathcal{P} \cap [x, 2x]$.*

5.3.2. Lemmas

We note that the second hypothesis of the main theorems states that for a fixed λ and y, N such that $P(N) \leq y$ and large enough, and $k_1, k_2 > 0$,

$$\left|\frac{1}{N}\sum_{j=0}^{N-1}\left(T_p\left(\frac{j}{N}\right) - T_{\lambda,y}\left(\frac{j}{N}\right)\right)^2 - \int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\, dx\right| < \frac{k_1\sqrt{\log y}}{\sqrt{y}} + \frac{k_2}{\sqrt{N}}.$$

We will prove the above in section 5.4. The proof will involve averaging over the set of primes $\mathcal{P}_x$ (Definition 5.3.4), and hence the first two lemmas of this section give results pertaining to $\mathcal{P}_x$. The first of these, Lemma 5.3.5 establishes the congruence class in which a prime p of $\mathcal{P}_x$ belongs, and applies the theorem based on GRH for the number of primes in an arithmetic progression. Then, using standard analytic number theory methods, it gives the size of the set $\mathcal{P}_x$, up to an error term. The second lemma, Lemma 5.3.6 uses the result as established in the first lemma for the congruence class in which a prime p of $\mathcal{P}_x$ belongs, and applies the orthogonality relation for characters. Next, it appeals to a result of character sums which is based on GRH, to deduce an upper bound for the average of the sum of Legendre symbols $\left(\frac{n}{p}\right)$ over the primes p in $\mathcal{P}_x$, where n has a large prime divisor which appears to an odd power.

Lemma 5.3.5. *Assume GRH. For $\mathcal{P}_x$ as in Definition 5.3.4, x large and $y \leq \frac{\log x}{5}$, we have*

$$|\mathcal{P}_x| = \frac{\pi(x)}{2^{\pi(y)}} + O(x^{\frac{3}{4}} \log x).$$

PROOF. By the definition of $\mathcal{P}$ (Definition 5.3.4), the set of primes p in $\mathcal{P}$ are such that $\lambda(q) = \left(\frac{q}{p}\right)$ for all primes $q \leq y$. It follows then by quadratic reciprocity, that $\left(\frac{p}{q}\right) = \pm 1$ depending on the value of $\lambda(q)$ and the congruence class of q mod 4. This implies that p is either a quadratic residue or a quadratic non-residue mod q. Further, since there are exactly $\frac{q-1}{2}$ congruence classes that correspond to quadratic residues mod q and exactly $\frac{q-1}{2}$ congruence classes that correspond to quadratic non-residues mod q, we can say that for each prime $q \leq y$, p lies in one of $\frac{q-1}{2}$ congruence classes mod q. Now, owing to the fact that the primes in $\mathcal{P}$ are also congruent to 3 mod 4, put $Q = 4 \prod_{3 \leq q \leq y} q$. And so by the Chinese Remainder Theorem, p lies in one of $\prod_{3 \leq q \leq y} (\frac{q-1}{2})$ congruence classes mod Q. We evaluate Q for $y \leq \frac{\log x}{5}$:

$$Q = 4 \prod_{3 \leq q \leq y} q = 2 \prod_{q \leq y} q = 2e^{\theta(y)} \leq 2x^{\frac{2}{5}}. \tag{5.3.2}$$

Now, since $Q \leq x^{\frac{1}{2}-\epsilon}$, under GRH we have that $\pi(x,Q,a) = \frac{\pi(x)}{\phi(Q)} + O(\sqrt{x}\log x)$ (see chapter 20 of [9]) and so,

$$|\mathcal{P}_x| = \left(\frac{\pi(x)}{\phi(Q)} + O(\sqrt{x}\log x)\right) \prod_{3\leq q\leq y} (\frac{q-1}{2}). \tag{5.3.3}$$

We will first evaluate the main term. Owing to the multiplicativity of the Euler ϕ function,

$$\phi(Q) = \phi(4 \prod_{3\leq q\leq y} q) = \phi(4) \prod_{3\leq q\leq y} \phi(q) = 2 \prod_{3\leq q\leq y} (q-1) \tag{5.3.4}$$

and so the main term evaluates to

$$\frac{\pi(x) \prod_{3\leq q\leq y} (q-1)}{\phi(Q) \prod_{3\leq q\leq y} 2} = \frac{\pi(x)}{\prod_{q\leq y} 2} = \frac{\pi(x)}{2^{\pi(y)}}. \tag{5.3.5}$$

Next, we will evaluate the error term. By the PNT and for $y \leq \frac{\log x}{5}$,

$$\prod_{3\leq q\leq y} (\frac{q-1}{2}) \leq e^{\theta(y)} \leq x^{\frac{1}{4}}.$$

And thus we obtain our error term as $O(x^{\frac{3}{4}}\log x)$. □

Lemma 5.3.6. *Assume GRH. Let $\mathcal{P}_x$ be as in Definition 5.3.4, x large and $y \leq \frac{\log x}{5}$. Then, if n has a prime divisor $> y$ which appears to an odd power,*

$$\frac{1}{|\mathcal{P}_x|}\left|\sum_{p\in\mathcal{P}_x} \left(\frac{n}{p}\right)\right| \ll \frac{\log(nx)\log x}{x^{\frac{1}{10}}}.$$

PROOF. Using the same arguments as were used in the proof of Lemma 5.3.5 we have in this case as well, that for p a prime in $\mathcal{P}$ (Definition 5.3.4), we can say that p lies in one of the $\prod_{3\leq q\leq y} (\frac{q-1}{2})$ congruence classes mod Q where $Q = 4 \prod_{\substack{q \text{ prime} \\ 3\leq q\leq y}} q$. Also, note that if for q prime, $q|k \Rightarrow q \leq y$ then by Definition 5.3.4, $\left(\frac{q}{p_1}\right) = \left(\frac{q}{p_2}\right) = \lambda(q)$ for all $p_1, p_2 \in \mathcal{P}_x$. Therefore, if we write $n = kn_1$, where $q|k \Rightarrow q \leq y$ and $(n_1, Q) = 1$ then

$$\sum_{p\in\mathcal{P}_x} \left(\frac{n}{p}\right) = \lambda(k) \sum_{\substack{p\in\mathcal{P} \\ x<p<2x}} \left(\frac{n_1}{p}\right) = \lambda(k) \sum_{\substack{\prod_{3\leq q\leq y} (\frac{q-1}{2}) \\ \text{choices of } b \bmod Q}} \sum_{\substack{p\equiv b \bmod Q \\ x<p<2x}} \left(\frac{n_1}{p}\right) \tag{5.3.6}$$

and by the orthogonality relation,

$$\sum_{p\in\mathcal{P}_x}\left(\frac{n}{p}\right)=\lambda(k)\sum_{\substack{\prod_{3\le q\le y}(\frac{q-1}{2})\\ \text{choices of } b \bmod Q}}\frac{1}{\phi(Q)}\sum_{\chi \bmod Q}\sum_{x<p<2x}\left(\frac{n_1}{p}\right)\chi(b^{-1})\chi(p). \quad (5.3.7)$$

Taking the absolute value then gives

$$\left|\sum_{p\in\mathcal{P}_x}\left(\frac{n}{p}\right)\right|\le\sum_{\substack{\prod_{3\le q\le y}(\frac{q-1}{2})\\ \text{choices of } b \bmod Q}}\frac{1}{\phi(Q)}\sum_{\chi \bmod Q}\left|\sum_{x<p<2x}\left(\frac{n_1}{p}\right)\chi(p)\right|. \quad (5.3.8)$$

Note that $\left(\frac{n_1}{\cdot}\right)\chi$ is a character of modulus at most $n_1Q=4n_1\prod_{3\le q\le y}q$. Further, owing to the fact that $(n_1,Q)=1$, and that n_1 has a prime divisor $>y$ which appears to an odd power, $\left(\frac{n_1}{\cdot}\right)\chi$ must be a non-principal character.

Now, by taking $T=\sqrt{x}$ in (13) on page 120 of [**9**] and using GRH, we have for ψ a non-principal character $\pmod m$ and $\Lambda(n)$ the von-Mangoldt's function,

$$\sum_{n\le x}\psi(n)\Lambda(n)\ll\sqrt{x}\log(mx)\log x.$$

By the above and partial summation, we then have

$$\sum_{p\le x}\psi(p)\ll\sqrt{x}\log(mx).$$

Applying this result to (5.3.8) gives

$$\left|\sum_{p\in\mathcal{P}_x}\left(\frac{n}{p}\right)\right|\ll\sum_{\substack{\prod_{3\le q\le y}(\frac{q-1}{2})\\ \text{choices of } b \bmod Q}}\frac{1}{\phi(Q)}\sum_{\chi \bmod Q}\sqrt{x}\log(n_1Qx). \quad (5.3.9)$$

Note that by (5.3.2) we have that $Q\ll x^{\frac{2}{5}}$. Using this in (5.3.9) gives

$$\left|\sum_{p\in\mathcal{P}_x}\left(\frac{n}{p}\right)\right|\ll\sqrt{x}\log(nx)\prod_{3\le q\le y}(\frac{q-1}{2}). \quad (5.3.10)$$

And so, since for our choice of $y\le\frac{\log x}{5}$, $\sqrt{x}\log x\ll\pi(x)/\phi(Q)$, owing to (5.3.3) we obtain

$$\frac{1}{|\mathcal{P}_x|}\left|\sum_{p\in\mathcal{P}_x}\left(\frac{n}{p}\right)\right|\ll\frac{\sqrt{x}\log(nx)\phi(Q)}{\pi(x)}. \quad (5.3.11)$$

Then, taking $y \leq \frac{\log x}{5}$ in (5.3.4) gives

$$\phi(Q) = 2 \prod_{3 \leq q \leq y} (q-1) \leq \prod_{q \leq y} q = e^{\theta(y)} \leq x^{\frac{2}{5}}.$$

Further owing to the fact that $\pi(x) \sim \frac{x}{\log x}$, the result follows. □

As per Remark 5.3.3, and Definition 5.3.4 of the set $\mathcal{P}$, for primes $p \in \mathcal{P}$, $T^*_{p,y}(t) = T_p(t) - T_{\lambda,y}(t)$ where $T^*_{p,y}(t)$ is as in Definition 5.3.2. We notice that the hypotheses of the main theorems involves the evaluation of $(T_p(t) - T_{\lambda,y}(t))^2$ (where $p \in \mathcal{P}_x$) and hence that of the square of the function $T^*_{p,y}(t)$. The following lemma gives an estimate for $T^*_{p,y}(t)^2$. In doing so, it appeals to the truncated Fourier series result of Lemma 4.1.1 and the bound for the sum of the reciprocals of integers with small prime factors as given by Corollary 4.1.3.

Lemma 5.3.7. *For $T^*_{p,y}(t)$ as in Definition 5.3.2 and $x \geq \sqrt{p}\log^2 p$ we have,*

$$T^*_{p,y}(t)^2 = \Big(\sum_{n \leq x} \Big(\frac{n}{p}\Big)_y \frac{\cos(2\pi nt)}{n}\Big)^2 + O\Big(\frac{p^{o(1)}\log x}{\sqrt{x}} + \frac{\sqrt{p}\log p \log x}{x} + \frac{p\log^2 p}{x^2}\Big). \tag{5.3.12}$$

PROOF. By the definition of $T^*_{p,y}(t)$, it follows that

$$T^*_{p,y}(t) = -\sum_{n \geq 1} \Big(\frac{n}{p}\Big) \frac{\cos(2\pi nt)}{n} + \sum_{\substack{n \geq 1 \\ P(n) \leq y}} \Big(\frac{n}{p}\Big) \frac{\cos(2\pi nt)}{n}$$

and owing to Lemma 4.1.1 and Lemma 4.1.2 we can write:

$$\begin{aligned} T^*_{p,y}(t) &= -\sum_{n \leq H} \Big(\frac{n}{p}\Big) \frac{\cos(2\pi nt)}{n} + O\Big(\frac{\sqrt{p}\log p}{H}\Big) + \sum_{\substack{n \leq H \\ P(n) \leq y}} \Big(\frac{n}{p}\Big) \frac{\cos(2\pi nt)}{n} + O\Big(\frac{c_{y,\sigma}}{H^{1-\sigma}}\Big) \\ &= -\sum_{n \leq H} \Big(\frac{n}{p}\Big)_y \frac{\cos(2\pi nt)}{n} + O\Big(\frac{c_{y,\sigma}}{H^{1-\sigma}}\Big) + O\Big(\frac{\sqrt{p}\log p}{H}\Big). \end{aligned}$$

Let $\sigma = \frac{1}{2}$, $H = x \geq \sqrt{p}\log^2 p$, then:

$$T^*_{p,y}(t) = -\sum_{n \leq x} \Big(\frac{n}{p}\Big)_y \frac{\cos(2\pi nt)}{n} + O\Big(\frac{c_{y,\frac{1}{2}}}{\sqrt{x}}\Big) + O\Big(\frac{\sqrt{p}\log p}{x}\Big)$$

and for $y \leq \log^2 p$, by Corollary 4.1.3

$$T^*_{p,y}(t) = -\sum_{n \leq x} \Big(\frac{n}{p}\Big)_y \frac{\cos(2\pi nt)}{n} + O\Big(\frac{p^{o(1)}}{\sqrt{x}}\Big) + O\Big(\frac{\sqrt{p}\log p}{x}\Big).$$

Therefore,

$$\begin{aligned} T^*_{p,y}(t)^2 &= \Big(-\sum_{n\le x}\Big(\frac{n}{p}\Big)_y \frac{\cos(2\pi nt)}{n} + O\Big(\frac{p^{o(1)}}{\sqrt{x}} + \frac{\sqrt{p}\log p}{x}\Big)\Big)^2 \\ &= \Big(\sum_{n\le x}\Big(\frac{n}{p}\Big)_y \frac{\cos(2\pi nt)}{n}\Big)^2 + O\Big(\frac{p^{o(1)}}{\sqrt{x}} + \frac{\sqrt{p}\log p}{x}\Big)^2 \\ &+ O\Big(\Big(\sum_{n\le x}\Big(\frac{n}{p}\Big)_y \frac{\cos(2\pi nt)}{n}\Big)\Big(\frac{p^{o(1)}}{\sqrt{x}} + \frac{\sqrt{p}\log p}{x}\Big)\Big). \end{aligned}$$

Now, since

$$\left|\sum_{n\le x}\Big(\frac{n}{p}\Big)_y \frac{\cos(2\pi nt)}{n}\right| \le \sum_{n\le x}\frac{1}{n} \le 1 + \int_1^x \frac{1}{u}\,du \ll \log x$$

the result follows. □

5.4. Satisfying the Hypotheses in the Main Theorems

Before giving the proofs of Theorems 5.2.1 and 5.2.3, we notice that the hypotheses of these theorems includes the following assumptions: assuming GRH, for a fixed λ and y, and $\left(\frac{q}{p}\right) = \lambda(q)$ for all $q \le y$ where $p \equiv 3 \bmod 4$,

$$\int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\,dx < c_y$$

where $c_y = \frac{1}{2}\Big(\frac{\pi^2}{6} - \prod_{q\le y}\Big(1-\frac{1}{q^2}\Big)^{-1}\Big)$ and for N large enough such that $P(N) \le y$ and $k_1, k_2 > 0$,

$$\Big|\frac{1}{N}\sum_{j=0}^{N-1}\Big(T_p\Big(\frac{j}{N}\Big) - T_{\lambda,y}\Big(\frac{j}{N}\Big)\Big)^2 - \int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\,dx\Big| < \frac{k_1\sqrt{\log y}}{\sqrt{y}} + \frac{k_2}{\sqrt{N}}.$$

Here we will first prove the above two assumptions and in turn, frame the hypotheses of Theorems 5.2.1 and 5.2.3. We will use the results of the preparatory lemmas given in Section 5.3.

5.4.1. Establishing the Hypotheses

The first hypothesis is established by Lemma 5.4.1 which uses the definitions of $T_p(x)$ and $T_{\lambda,y}(x)$ as given in (5.1.2) and Definition 5.1.5, respectively. In order to force a resemblance between the two functions $T_p(x)$ and $T_{\lambda,y}(x)$, this lemma evaluates an upper bound for the second moments of $T_p(x) - T_{\lambda,y}(x)$. The result

as is stated, is obtained by elementary integration of trigonometric functions, the definition of the Legendre symbol, and the standard form of expressing certain sums as Euler products.

Lemma 5.4.1. *Fix λ and y. For all primes $p \equiv 3 \pmod 4$ such that $\left(\frac{q}{p}\right) = \lambda(q)$ for all primes $q \leq y$, we have*

$$\int_0^1 (T_p(x) - T_{\lambda,y}(x))^2 \, dx < c_y \tag{5.4.1}$$

where $c_y = \frac{1}{2}\left(\frac{\pi^2}{6} - \prod_{q \leq y}\left(1 - \frac{1}{q^2}\right)^{-1}\right)$.

PROOF.

$$\begin{aligned}
\int_0^1 (T_p(x) - T_{\lambda,y}(x))^2 \, dx &= \int_0^1 \left(-\sum_{n\geq 1}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n} + \sum_{\substack{n\geq 1 \\ P(n)\leq y}} \lambda(n)\frac{\cos 2\pi nx}{n}\right)^2 dx \\
&= \int_0^1 \left(-\sum_{\substack{n\geq 1 \\ P(n)> y}}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}\right)^2 dx \\
&= \int_0^1 \left(\sum_{\substack{n\geq 1 \\ P(n)> y}}\left(\frac{n}{p}\right)\frac{\cos 2\pi nx}{n}\right)\left(\sum_{\substack{m\geq 1 \\ P(m)> y}}\left(\frac{m}{p}\right)\frac{\cos 2\pi mx}{m}\right) dx \\
&= \int_0^1 \sum_{\substack{m,n\geq 1 \\ P(m),\ P(n)> y}} \left(\frac{mn}{p}\right)\frac{\cos 2\pi mx}{m}\frac{\cos 2\pi nx}{n}\, dx.
\end{aligned}$$

Now,

$$\int_0^1 \cos 2\pi mx \ \cos 2\pi nx \, dx = \begin{cases} \frac{1}{2} & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases} \tag{5.4.2}$$

And so,

$$\begin{aligned}\int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\,dx &= \frac{1}{2} \sum_{\substack{m,n\geq 1 \\ m\equiv n \\ P(m),\ P(n)>y,}} \left(\frac{mn}{p}\right)\frac{1}{mn} = \frac{1}{2}\sum_{\substack{n\geq 1 \\ P(n)>y}} \left(\frac{n^2}{p}\right)\frac{1}{n^2} \\ &= \frac{1}{2} \sum_{\substack{n\geq 1 \\ p\nmid n,\ P(n)>y}} \frac{1}{n^2} \qquad \text{since } \left(\frac{n^2}{p}\right) = \begin{cases} 1 & \text{if } (p,n)=1, \\ 0 & \text{if } p|n, \end{cases} \\ &= \frac{1}{2}\left(\sum_{\substack{n\geq 1 \\ p\nmid n}} \frac{1}{n^2} - \sum_{\substack{n\geq 1 \\ p\nmid n,\ P(n)\leq y}} \frac{1}{n^2}\right).\end{aligned}$$

The above summations can then be written as Euler products, giving:

$$\int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\,dx = \frac{1}{2}\left(\prod_{q\neq p}\left(1-\frac{1}{q^2}\right)^{-1} - \prod_{\substack{q\leq y \\ q\neq p}}\left(1-\frac{1}{q^2}\right)^{-1}\right).$$

Since

$$\prod_{\substack{q\leq y \\ q\neq p}}\left(1-\frac{1}{q^2}\right)^{-1} = \begin{cases} \left(1-\frac{1}{p^2}\right)\prod_{q\leq y}\left(1-\frac{1}{q^2}\right)^{-1} & \text{if } p\leq y, \\ \prod_{q\leq y}\left(1-\frac{1}{q^2}\right)^{-1} & \text{if } p > y. \end{cases}$$

and assuming $p > y$, we consider

$$\prod_{\substack{q\leq y \\ q\neq p}}\left(1-\frac{1}{q^2}\right)^{-1} = \prod_{q\leq y}\left(1-\frac{1}{q^2}\right)^{-1},$$

which then gives:

$$\begin{aligned}\int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\,dx &= \frac{1}{2}\left(\left(1-\frac{1}{p^2}\right)\zeta(2) - \prod_{q\leq y}\left(1-\frac{1}{q^2}\right)^{-1}\right) \\ &< \frac{1}{2}\left(\frac{\pi^2}{6} - \prod_{q\leq y}\left(1-\frac{1}{q^2}\right)^{-1}\right).\end{aligned}$$

□

The second hypothesis of the main theorems assumes GRH and is established via Lemma 5.4.2 and Corollary 5.4.3. In Corollary 5.4.3 we will evaluate the explicit value for the constant k_1 (see (5.2.1)) for all $y \geq 29$. Lemma 5.4.2 uses the definition of $T^*_{p,y}(t)$ as given in Definition 5.3.2 and of the sets $\mathcal{P}$, $\mathcal{P}_x$ as in Definition 5.3.4. The proof of this lemma is essentially technical, and uses basic

trigonometric identities, standard results of analytic number theory and Cauchy-Schwarz's inequality. It appeals to Remark 5.3.3 which notes that for all primes $p \in \mathcal{P}$, $T^*_{p,y}(t) = T_p(t) - T_{\lambda,y}(t)$ and also to Lemma 5.3.7 which gives a formula for $T^*_{p,y}(t)^2$ where the main term is given by a finite sum. It also uses the results of Lemmas 5.3.5 and 5.3.6 which assume GRH and respectively, estimate the size of the set $\mathcal{P}_x$ and give a bound for the average of a sum of Legendre symbols $\left(\frac{n}{p}\right)$ over the set of primes $p \in \mathcal{P}_x$ for all n that have a large prime divisor which appears to an odd power. The basic approach will be to take an average over the set of primes $\mathcal{P}_x$ and hence approximate the value of

$$\left| \frac{1}{|\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \left(\frac{1}{N} \sum_{j=0}^{N-1} T^*_{p,y}\left(\frac{j}{N}\right)^2 - \int_0^1 T^*_{p,y}(t)^2 \, dt \right) \right|.$$

We notice that the analytic and discrete second moments of $T^*_{p,y}(t) = -\sum_{n \geq 1} \left(\frac{n}{p}\right)_y \frac{\cos(2\pi nt)}{n}$ are very close for small n. Hence, in order to approximate the above, we will split the finite sum in the formula for $T^*_{p,y}(t)^2$ into two finite sums (of which the sum over the larger n will yield the main term). We will then expand out the resulting expression and obtain six terms, each of which we will evaluate individually.

Lemma 5.4.2. *Assume GRH. Fix y and let $T^*_{p,y}(t)$ be as in Definition 5.3.2. Then, for x large, $\mathcal{P}_x$ as in Definition 5.3.4 and $c_1 > 0$, there exist at least $\frac{|\mathcal{P}_x|}{2}$ primes in $\mathcal{P}_x$, for which*

$$\left| \frac{1}{N} \sum_{j=0}^{N-1} T^*_{p,y}\left(\frac{j}{N}\right)^2 - \int_0^1 T^*_{p,y}(t)^2 \, dt \right| < c_1 \left(\frac{1}{\sqrt{N}} + \frac{\sqrt{\log y}}{\sqrt{y}} \right) \tag{5.4.3}$$

where N is such that $P(N) \leq y$. Further, for z sufficiently large, and k a small positive constant, the number of such primes $p \leq z$ is at least $ze^{\frac{-k \log z}{\log \log z}}$.

PROOF. We will show that for $c_1 > 0$,

$$\frac{1}{|\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \left| \frac{1}{N} \sum_{j=0}^{N-1} T^*_{p,y}\left(\frac{j}{N}\right)^2 - \int_0^1 T^*_{p,y}(t)^2 \, dt \right| < \frac{c_1}{2} \left(\frac{1}{\sqrt{N}} + \frac{\sqrt{\log y}}{\sqrt{y}} \right). \tag{5.4.4}$$

We claim then that there exist at least $\frac{|\mathcal{P}_x|}{2}$ primes $p \in \mathcal{P}_x$ such that (5.4.3) holds. If our claim is false, we must have

$$\Big|\frac{1}{N}\sum_{j=0}^{N-1} T_{p,y}^*\Big(\frac{j}{N}\Big)^2 - \int_0^1 T_{p,y}^*(t)^2\,dt\Big| \geq c_1\Big(\frac{1}{\sqrt{N}} + \frac{\sqrt{\log y}}{\sqrt{y}}\Big)$$

for at least $\frac{|\mathcal{P}_x|}{2}$ primes $p \in \mathcal{P}_x$. If so, we will get

$$\begin{aligned}\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\Big|\frac{1}{N}\sum_{j=0}^{N-1} T_{p,y}^*\Big(\frac{j}{N}\Big)^2 - \int_0^1 T_{p,y}^*(t)^2\,dt\Big| &\geq \frac{1}{|\mathcal{P}_x|}c_1\Big(\frac{1}{\sqrt{N}} + \frac{\sqrt{\log y}}{\sqrt{y}}\Big)\frac{|\mathcal{P}_x|}{2}\\ &= \frac{c_1}{2}\Big(\frac{1}{\sqrt{N}} + \frac{\sqrt{\log y}}{\sqrt{y}}\Big)\end{aligned}$$

which contradicts (5.4.4) and thus there must be at least $\frac{|\mathcal{P}_x|}{2}$ primes $p \in \mathcal{P}_x$ for which (5.4.3) holds. Note that by Lemma 5.3.5, $\frac{|\mathcal{P}_x|}{2} \geq \frac{\pi(x)}{2^{\pi(y)+1}} + O(x^{\frac{3}{4}}\log x)$.
To prove the second assertion of the lemma, for z sufficiently large, consider primes $\sqrt{z} \leq p \leq z$. Now, divide this interval into dyadic intervals:

$$I_j = \Big[\frac{z}{2^j}, \frac{z}{2^{j-1}}\Big] \text{ for } 1 \leq j \leq J$$

where J is the largest integer such that $\sqrt{z} \leq \frac{z}{2^J}$. By the first assertion of the lemma, we have that in each interval I_j, at least half the number of primes $p \equiv 3 \bmod 4$ for which $\left(\frac{q}{p}\right) = \lambda(q)$ for all primes $q \leq y$, satisfy (5.4.3). Therefore, the number of primes $p \leq z$ that satisfy (5.4.3), is at least

$$\begin{aligned}&\frac{1}{2}\sum_{j=1}^{J}|\{p \in I_j : p \equiv 3 \bmod 4, \left(\frac{q}{p}\right) = \lambda(q) \text{ for all primes } q \leq y\}|\\ =\ &\frac{1}{2}|\{\sqrt{z} \leq p \leq z : p \in \mathcal{P}\}|\end{aligned}$$

and appealing to Lemma 5.3.5 we then have that for $y \leq \frac{\log z}{5}$, the number of primes $p \leq z$ that satisfy (5.4.3) is at least $ze^{\frac{-k\log z}{\log\log z}}$ where k is a positive constant.

Now we will prove that (5.4.4) holds. By (5.3.12), for $p > x \geq \sqrt{p}\log^2 p$

$$\begin{aligned}
&\frac{1}{N}\sum_{j=0}^{N-1} T_{p,y}^*\left(\frac{j}{N}\right)^2 - \int_0^1 T_{p,y}^*(t)^2\,dt \\
&= \frac{1}{N}\sum_{j=0}^{N-1}\Big(\sum_{n\leq x}\left(\frac{n}{p}\right)_y \frac{\cos(\frac{2\pi nj}{N})}{n}\Big)^2 - \int_0^1\Big(\sum_{n\leq x}\left(\frac{n}{p}\right)_y \frac{\cos(2\pi nt)}{n}\Big)^2\,dt \\
&+ O\Big(\frac{p^{o(1)}\log x}{\sqrt{x}} + \frac{\sqrt{p}\log p\log x}{x} + \frac{p\log^2 p}{x^2}\Big).
\end{aligned}$$

and for $x < p < 2x$, the error term is $O(x^{-\frac{1}{2}+o(1)})$. Now, we take an average over the set of primes $\mathcal{P}_x$, which gives

$$\left|\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\Big(\frac{1}{N}\sum_{j=0}^{N-1} T_{p,y}^*\left(\frac{j}{N}\right)^2 - \int_0^1 T_{p,y}^*(t)^2\,dt\Big)\right| \leq |S| + O\Big(\frac{1}{x^{\frac{1}{2}-o(1)}}\Big)$$

where

$$S = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\Big(\frac{1}{N}\sum_{j=0}^{N-1} T_{p,y,x}^*\left(\frac{j}{N}\right)^2 - \int_0^1 T_{p,y,x}^*(t)^2\,dt\Big)$$

and

$$T_{p,y,x}^*(t) = \sum_{n\leq x}\left(\frac{n}{p}\right)_y \frac{\cos(2\pi nt)}{n}. \tag{5.4.5}$$

Let $M = \frac{N}{2}$ and $M < x$. Then write

$$T_{p,y,x}^*(t) = T_{p,y,M}^*(t) + T_{p,y,M}^{**}(t)$$

where we define

$$T_{p,y,M}^{**}(t) := \sum_{M<n\leq x}\left(\frac{n}{p}\right)_y \frac{\cos(2\pi nt)}{n}. \tag{5.4.6}$$

Note that:

$$\begin{aligned}
&\frac{1}{N}\sum_{j=0}^{N-1} T^*_{p,y,x}\left(\frac{j}{N}\right)^2 - \int_0^1 T^*_{p,y,x}(t)^2\,dt \\
=\ &\frac{1}{N}\sum_{j=0}^{N-1}\left(T^*_{p,y,x}\left(\frac{j}{N}\right)^2 - \int_{-1/2}^{1/2} T^*_{p,y,x}\left(\frac{j+\theta}{N}\right)^2 d\theta\right) \\
=\ &\frac{1}{N}\sum_{j=0}^{N-1}\left(\int_{-1/2}^{1/2}\left(T^*_{p,y,x}\left(\frac{j}{N}\right)^2 - T^*_{p,y,x}\left(\frac{j+\theta}{N}\right)^2\right) d\theta\right) \\
=\ &\frac{1}{N}\sum_{j=0}^{N-1}\left(\int_{-1/2}^{1/2}\left[\left(T^*_{p,y,M}\left(\frac{j}{N}\right) + T^{**}_{p,y,M}\left(\frac{j}{N}\right)\right)^2\right.\right. \\
-\ &\left.\left.\left(T^*_{p,y,M}\left(\frac{j+\theta}{N}\right) + T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)\right)^2\right] d\theta\right) \\
=\ &\frac{1}{N}\sum_{j=0}^{N-1}\int_{-1/2}^{1/2}\left[T^*_{p,y,M}\left(\frac{j}{N}\right)^2 + T^{**}_{p,y,M}\left(\frac{j}{N}\right)^2 - T^*_{p,y,M}\left(\frac{j+\theta}{N}\right)^2 - T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)^2\right. \\
+\ &\left. 2T^*_{p,y,M}\left(\frac{j}{N}\right)T^{**}_{p,y.M}\left(\frac{j}{N}\right) - 2T^*_{p,y,M}\left(\frac{j+\theta}{N}\right)T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)\right] d\theta
\end{aligned}$$

In view of the above, we write

$$S = S_1 + S_2 - S_3 - S_4 + 2S_5 - 2S_6 \tag{5.4.7}$$

where

$$S_1 = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\int_{-1/2}^{1/2} T^*_{p,y,M}\left(\frac{j}{N}\right)^2 d\theta \tag{5.4.8}$$

$$S_2 = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\int_{-1/2}^{1/2} T^{**}_{p,y,M}\left(\frac{j}{N}\right)^2 d\theta \tag{5.4.9}$$

$$S_3 = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\int_{-1/2}^{1/2} T^*_{p,y,M}\left(\frac{j+\theta}{N}\right)^2 d\theta \tag{5.4.10}$$

$$S_4 = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\int_{-1/2}^{1/2} T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)^2 d\theta \tag{5.4.11}$$

$$S_5 = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\int_{-1/2}^{1/2} T^{*}_{p,y,M}\left(\frac{j}{N}\right)T^{**}_{p,y,M}\left(\frac{j}{N}\right)\,d\theta \tag{5.4.12}$$

$$S_6 = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\int_{-1/2}^{1/2} T^{*}_{p,y,M}\left(\frac{j+\theta}{N}\right)T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)\,d\theta. \tag{5.4.13}$$

Evaluating each of the above terms of S individually, we will obtain:

$$S_2 \ll \frac{\log^4 x}{Nx^{\frac{1}{10}}} + \frac{1}{M} + \frac{\log y}{y} \tag{5.4.14}$$

$$S_4 \ll \frac{1}{M} \tag{5.4.15}$$

$$S_5 \ll \frac{\log^2 x}{x^{\frac{1}{20}}\sqrt{N}} + \frac{1}{\sqrt{M}} + \frac{\sqrt{\log y}}{\sqrt{y}} \tag{5.4.16}$$

$$S_6 \ll \frac{1}{\sqrt{M}} \tag{5.4.17}$$

and

$$S_1 - S_3 \leq \frac{1}{2M^2}. \tag{5.4.18}$$

The details of the above evaluations will follow later. For now, using the above results, we have

$$\begin{aligned} |S| &\leq |S_1 - S_3| + |S_2| + |S_4| + 2|S_5| + 2|S_6| \\ &\ll \frac{1}{M^2} + \frac{\log^4 x}{Nx^{\frac{1}{10}}} + \frac{1}{M} + \frac{\log y}{y} + \frac{\log^2 x}{x^{\frac{1}{20}}\sqrt{N}} + \frac{1}{\sqrt{M}} + \frac{\sqrt{\log y}}{\sqrt{y}} \\ &\ll \frac{\log^2 x}{x^{\frac{1}{20}}\sqrt{N}} + \frac{1}{\sqrt{M}} + \frac{\sqrt{\log y}}{\sqrt{y}}. \end{aligned}$$

Thus, for x large, $y \leq \frac{\log x}{5}$, $M = \frac{N}{2}$, and $c_1 > 0$, we obtain (5.4.4). We will now prove each of the results (5.4.14), (5.4.15), (5.4.16), (5.4.17), (5.4.18). Recall that $T^{*}_{p,y,M}(x)$ and $T^{**}_{p,y,M}(x)$ are as defined in (5.4.5) and (5.4.6), respectively. The details of the evaluation of S_1, S_2, S_3, S_4, S_5 and S_6 are then as follows.

Evaluating S_1 and S_3: Note that besides evaluating $S_1 - S_3$ as a single term as stated in (5.4.18), we will also evaluate S_1 and S_3 individually, since their values will be used when evaluating S_5 and S_6. We will obtain:

$$S_1 < 1, \tag{5.4.19}$$

$$S_3 < 1. \tag{5.4.20}$$

By the definition of S_1 (5.4.8) and S_3 (5.4.10), it is clear that $S_1, S_3 > 0$. Now we will give the details of the values obtained for $S_1 - S_3$, S_1 and S_3. We will write the diagonal and non-diagonal terms of S_1 and S_3 separately where the non-diagonal terms are given by:

$$S_1' = \frac{1}{|\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \frac{1}{N} \sum_{\substack{n,m \leq M \\ n \neq m}} \left(\frac{m}{p}\right)_y \left(\frac{n}{p}\right)_y \frac{1}{mn} \sum_{j=0}^{N-1} \cos\left(\frac{2\pi m j}{N}\right) \cos\left(\frac{2\pi n j}{N}\right) \tag{5.4.21}$$

$$S_3' = \frac{1}{|\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \frac{1}{N} \sum_{\substack{n,m \leq M \\ n \neq m}} \frac{\left(\frac{m}{p}\right)_y \left(\frac{n}{p}\right)_y}{mn} \sum_{j=0}^{N-1} \int_{-1/2}^{1/2} \cos\left(\frac{2\pi m(j+\theta)}{N}\right) \cos\left(\frac{2\pi n(j+\theta)}{N}\right) d\theta \tag{5.4.22}$$

and the diagonal terms are given by:

$$S_1'' = \frac{1}{|\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \frac{1}{N} \sum_{m \leq M} \left(\frac{m^2}{p}\right)_y \frac{1}{m^2} \sum_{j=0}^{N-1} \cos^2\left(\frac{2\pi m j}{N}\right) \tag{5.4.23}$$

$$S_3'' = \frac{1}{|\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \frac{1}{N} \sum_{m \leq M} \left(\frac{m^2}{p}\right)_y \frac{1}{m^2} \sum_{j=0}^{N-1} \int_{-1/2}^{1/2} \cos^2\left(\frac{2\pi m(j+\theta)}{N}\right) d\theta \tag{5.4.24}$$

where

$$S_1 = S_1' + S_1'' \tag{5.4.25}$$

and

$$S_3 = S_3' + S_3''. \tag{5.4.26}$$

We will first evaluate the non-diagonal terms of S_1 and S_3, i.e. we will evaluate S_1' and S_3', and will show that in fact, they both equal zero. To evaluate S_1', we will begin by evaluating the inner sum in (5.4.21):

$$\begin{aligned} \sum_{j=0}^{N-1} \cos\left(\frac{2\pi m j}{N}\right) \cos\left(\frac{2\pi n j}{N}\right) &= \frac{1}{2} \sum_{j=0}^{N-1} \left[\cos\left(\frac{2\pi(m+n)j}{N}\right) + \cos\left(\frac{2\pi(m-n)j}{N}\right)\right] \\ &= \frac{1}{4} \sum_{j=0}^{N-1} \left(e^{2\pi i\left(\frac{(m+n)j}{N}\right)} + e^{-2\pi i\left(\frac{(m+n)j}{N}\right)}\right) \\ &+ \frac{1}{4} \sum_{j=0}^{N-1} \left(e^{2\pi i\left(\frac{(m-n)j}{N}\right)} + e^{-2\pi i\left(\frac{(m-n)j}{N}\right)}\right). \end{aligned} \tag{5.4.27}$$

Since for $l \in \mathbb{Z}$,

$$\sum_{j=0}^{N-1} e^{\frac{2\pi i l j}{N}} = \begin{cases} N & \text{if } N|l, \\ 0 & \text{otherwise.} \end{cases}$$

we have

$$\sum_{j=0}^{N-1} \left(e^{2\pi i\left(\frac{(m\pm n)j}{N}\right)} + e^{-2\pi i\left(\frac{(m\pm n)j}{N}\right)} \right) = \begin{cases} 2N & \text{if } N|(m \pm n), \\ 0 & \text{otherwise.} \end{cases} \tag{5.4.28}$$

Then, substituting (5.4.28) in (5.4.27) and using the subsequent result in (5.4.21) gives

$$S_1' = \frac{1}{2|\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \left(\sum_{\substack{n,m \le M \\ m \neq n \\ N|(m\pm n)}} \left(\frac{m}{p}\right)_y \left(\frac{n}{p}\right)_y \frac{1}{mn} \right). \tag{5.4.29}$$

However, owing to our choice of $M = \frac{N}{2}$, for $1 \le m \neq n \le M$, we see that $N \nmid (m \pm n)$ since $0 < |m \pm n| < 2M = N$. Therefore, $\sum_{\substack{n,m \le M \\ m \neq n \\ N|(m\pm n)}} \left(\frac{m}{p}\right)_y \left(\frac{n}{p}\right)_y \frac{1}{mn} = 0$.

And hence, using this fact in (5.4.29), the non-diagonal term of S_1,

$$S_1' = 0. \tag{5.4.30}$$

Next, to evaluate S_3', we will begin by evaluating the integral in (5.4.22):

$$\int_{-1/2}^{1/2} \cos\left(\frac{2\pi m(j+\theta)}{N}\right) \cos\left(\frac{2\pi n(j+\theta)}{N}\right) d\theta$$

$$= \frac{1}{2} \int_{-1/2}^{1/2} \left[\cos\left(\frac{2\pi(m+n)(j+\theta)}{N}\right) + \cos\left(\frac{2\pi(m-n)(j+\theta)}{N}\right) \right] d\theta \tag{5.4.31}$$

and

$$\frac{1}{2} \int_{-1/2}^{1/2} \cos\left(\frac{2\pi(m\pm n)(j+\theta)}{N}\right) d\theta = \frac{1}{4} \int_{-1/2}^{1/2} \left[e^{\frac{2\pi i(m\pm n)(j+\theta)}{N}} + e^{\frac{-2\pi i(m\pm n)(j+\theta)}{N}} \right] d\theta$$

$$= \frac{1}{4} e^{\frac{2\pi i(m\pm n)j}{N}} \int_{-1/2}^{1/2} e^{\frac{2\pi i(m\pm n)\theta}{N}} d\theta + \frac{1}{4} e^{\frac{-2\pi i(m\pm n)j}{N}} \int_{-1/2}^{1/2} e^{\frac{-2\pi i(m\pm n)\theta}{N}} d\theta.$$

Since

$$\int_{-1/2}^{1/2} e^{\frac{2\pi i(m\pm n)\theta}{N}} d\theta = \int_{-1/2}^{1/2} e^{\frac{-2\pi i(m\pm n)\theta}{N}} d\theta = \frac{N}{2\pi i(m \pm n)} \left[e^{\pi i(m\pm n)/N} - e^{-\pi i(m\pm n)/N} \right]$$

$$= \frac{N}{\pi(m \pm n)} \sin\left(\frac{\pi(m \pm n)}{N}\right)$$

and $\sin\left(\frac{\pi(m\pm n)}{N}\right) = 0$ if $N|(m \pm n)$, we have that if $N \nmid (m \pm n)$, then

$$\begin{aligned}\frac{1}{2}\int_{-1/2}^{1/2} \cos\left(\frac{2\pi(m \pm n)(j+\theta)}{N}\right) d\theta &= \frac{N}{4\pi(m \pm n)} \sin\left(\frac{\pi(m \pm n)}{N}\right)[e^{\frac{2\pi i(m\pm n)j}{N}} + e^{\frac{-2\pi i(m\pm n)j}{N}}] \\ &= \frac{N}{2\pi(m \pm n)} \sin\left(\frac{\pi(m \pm n)}{N}\right) \cos\left(\frac{2\pi(m \pm n)j}{N}\right).\end{aligned}$$

Hence we have

$$\frac{1}{2}\int_{-1/2}^{1/2} \cos\left(\frac{2\pi(m \pm n)(j+\theta)}{N}\right) d\theta = \begin{cases} \frac{N}{2\pi(m\pm n)} \sin\left(\frac{\pi(m\pm n)}{N}\right) \cos\left(\frac{2\pi(m\pm n)j}{N}\right) & \text{if } N \nmid (m \pm n), \\ 0 & \text{otherwise.} \end{cases} \tag{5.4.32}$$

Then, substituting (5.4.32) in (5.4.31) and using the subsequent result in (5.4.22) gives

$$\begin{aligned} S_3' &= \frac{1}{2\pi|\mathcal{P}_x|} \sum_{p\in\mathcal{P}_x} \Bigg(\sum_{\substack{n,m\le M \\ m\neq n \\ N\nmid(m+n)}} \frac{\left(\frac{m}{p}\right)_y \left(\frac{n}{p}\right)_y}{(m+n)mn} \sin\left(\frac{\pi(m+n)}{N}\right) \sum_{j=0}^{N-1} \cos\left(\frac{2\pi(m+n)j}{N}\right) \\ &+ \sum_{\substack{n,m\le M \\ m\neq n \\ N\nmid(m-n)}} \frac{\left(\frac{m}{p}\right)_y \left(\frac{n}{p}\right)_y}{(m-n)mn} \sin\left(\frac{\pi(m-n)}{N}\right) \sum_{j=0}^{N-1} \cos\left(\frac{2\pi(m-n)j}{N}\right) \Bigg). \end{aligned} \tag{5.4.33}$$

However, by (5.4.28) we have

$$\sum_{j=0}^{N-1} \cos\left(\frac{2\pi(m \pm n)j}{N}\right) = \frac{1}{2}\sum_{j=0}^{N-1} \left(e^{\frac{2\pi i(m\pm n)j}{N}} + e^{\frac{-2\pi i(m\pm n)j}{N}}\right) = \begin{cases} N & \text{if } N|m \pm n, \\ 0 & \text{otherwise.} \end{cases} \tag{5.4.34}$$

and since $N \nmid (m \pm n)$,

$$\sum_{j=0}^{N-1} \cos\left(\frac{2\pi(m \pm n)j}{N}\right) = 0. \tag{5.4.35}$$

Therefore, using (5.4.35) in (5.4.33), we have the non-diagonal term of S_3,

$$S_3' = 0. \tag{5.4.36}$$

Having shown that the non-diagonal terms of S_1 and S_3, i.e. S_1' and S_3' (respectively) equal zero we now have by (5.4.25) and (5.4.26) that

$$S_1 = S_1'' \tag{5.4.37}$$

and

$$S_3 = S_3''. \tag{5.4.38}$$

Thus, evaluating S_1 and S_3 is equivalent to evaluating just their diagonal terms given by S_1'' and S_3'' respectively. We will first evaluate S_1''. To do so, we will begin by evaluating the inner sum in S_1'' as given in (5.4.23):

$$\frac{1}{N}\sum_{j=0}^{N-1}\cos^2\left(\frac{2\pi mj}{N}\right) = \frac{1}{2N}\sum_{j=0}^{N-1}\left[\cos\left(\frac{4\pi mj}{N}\right)+1\right] = \frac{1}{2N}\sum_{j=0}^{N-1}\cos\left(\frac{4\pi mj}{N}\right)+\frac{1}{2}.$$

By (5.4.34), we have that

$$\frac{1}{2N}\sum_{j=0}^{N-1}\cos\left(\frac{4\pi mj}{N}\right) = \begin{cases} \frac{1}{2} & \text{if } N|2m, \\ 0 & \text{otherwise.} \end{cases} \tag{5.4.39}$$

and thus

$$\frac{1}{N}\sum_{j=0}^{N-1}\cos^2\left(\frac{2\pi mj}{N}\right) = \begin{cases} 1 & \text{if } N|2m, \\ \frac{1}{2} & \text{otherwise.} \end{cases} \tag{5.4.40}$$

Using (5.4.40) in (5.4.23) we have:

$$S_1'' = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\Bigg(\sum_{\substack{m\le M\\ N|2m}}\left(\frac{m^2}{p}\right)_y\frac{1}{m^2}+\frac{1}{2}\sum_{\substack{m\le M\\ N\nmid 2m}}\left(\frac{m^2}{p}\right)_y\frac{1}{m^2}\Bigg) \le \sum_{\substack{m\le M\\ N|2m\\ P(m)>y}}\frac{1}{m^2}+\frac{1}{2}\sum_{\substack{m\le M\\ N\nmid 2m\\ P(m)>y}}\frac{1}{m^2}. \tag{5.4.41}$$

Further, since $M=\frac{N}{2}$, $N\nmid 2m$ for $m<M$, and hence we only have one term in the sum $\sum_{\substack{m\le M\\ N|2m}}\frac{1}{m^2}$ which is when $m=M$. Using this fact in (5.4.41) gives

$$S_1'' \le \frac{1}{M^2}+\frac{1}{2}\sum_{\substack{m<M\\ N\nmid 2m\\ P(m)>y}}\frac{1}{m^2} < \frac{1}{M^2}+\frac{1}{2}\sum_{m<M}\frac{1}{m^2}. \tag{5.4.42}$$

Note that for $N>10$,

$$\frac{1}{M^2} < \frac{1}{2}\left(\frac{1}{M^2}+\frac{1}{(M+1)^2}+\frac{1}{(M+2)^2}\right), \tag{5.4.43}$$

and so using this in (5.4.42) gives

$$S_1'' < \frac{1}{2}\sum_{m\ge 1}\frac{1}{m^2} = \frac{\pi^2}{12} < 1. \tag{5.4.44}$$

Then, by (5.4.44) and (5.4.37), we have $S_1 < 1$ and hence have proved (5.4.19).

We will now prove (5.4.20) which states that $S_3 < 1$. We claim that $S_1 - S_3 \geq 0$. We will actually prove this when we evaluate $S_1 - S_3$ to give $S_1 - S_3 \leq \frac{1}{2M^2}$ (see (5.4.18)). Since $\left(\frac{M}{p}\right)^2 \leq 1$, this will follow from (5.4.50) which gives $S_1 = S_3 + \frac{1}{2M^2}\left(\frac{M}{p}\right)^2$. Given this and that $S_1 < 1$, it directly follows that $S_3 < S_1 < 1$.

We will now prove that $S_1 = S_3 + \frac{1}{2M^2}\left(\frac{M}{p}\right)^2$. Note that owing to (5.4.37) and (5.4.38),

$$S_1 - S_3 = S_1'' - S_3'' \tag{5.4.45}$$

and further by (5.4.23) and (5.4.24), we have:

$$S_1 - S_3 = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\sum_{m\leq M}\left(\frac{m^2}{p}\right)_y \frac{I_{j,m}}{m^2} \tag{5.4.46}$$

where

$$I_{j,m} = \int_{-1/2}^{1/2}\left[\cos^2\left(\frac{2\pi mj}{N}\right) - \cos^2\left(\frac{2\pi m(j+\theta)}{N}\right)\right]d\theta. \tag{5.4.47}$$

We will begin by evaluating $I_{j,m}$ which is defined in (5.4.47). We obtain

$$\begin{aligned}
I_{j,m} &= \int_{-1/2}^{1/2}\left[\cos^2\left(\frac{2\pi mj}{N}\right) - \cos^2\left(\frac{2\pi m(j+\theta)}{N}\right)\right]d\theta \\
&= \frac{1}{2}\int_{-1/2}^{1/2}\left[\cos\left(\frac{4\pi mj}{N}\right) - \cos\left(\frac{4\pi m(j+\theta)}{N}\right)\right]d\theta \\
&= \frac{1}{2}\cos\left(\frac{4\pi mj}{N}\right) - \frac{N}{8\pi m}\left[\sin\left(\frac{4\pi m(j+\frac{1}{2})}{N}\right) - \sin\left(\frac{4\pi m(j-\frac{1}{2})}{N}\right)\right] \\
&= \frac{1}{2}\cos\left(\frac{4\pi mj}{N}\right) - \frac{N}{4\pi m}\left[\cos\left(\frac{4\pi mj}{N}\right)\sin\left(\frac{2\pi m}{N}\right)\right] \\
&= \frac{1}{2}\cos\left(\frac{4\pi mj}{N}\right)\left(1 - \frac{N\sin(\frac{2\pi m}{N})}{2\pi m}\right).
\end{aligned} \tag{5.4.48}$$

Using (5.4.48) in (5.4.46), we have

$$S_1 - S_3 = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\sum_{m\leq M}\left(\frac{m^2}{p}\right)_y \frac{1}{m^2}\left(1 - \frac{N\sin(\frac{2\pi m}{N})}{2\pi m}\right)\frac{1}{2N}\sum_{j=0}^{N-1}\cos\left(\frac{4\pi mj}{N}\right) \tag{5.4.49}$$

and appealing to (5.4.39) for the inner sum $\frac{1}{N}\sum_{j=0}^{N-1}\cos\left(\frac{4\pi mj}{N}\right)$, (5.4.49) gives

$$S_1 - S_3 = \frac{1}{2|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\sum_{\substack{m\leq M\\ N|2m}}\left(\frac{m^2}{p}\right)_y\frac{1}{m^2}\left(1-\frac{N\sin(\frac{2\pi m}{N})}{2\pi m}\right) = \frac{1}{2M^2}\left(\frac{M^2}{p}\right) \leq \frac{1}{2M^2} \tag{5.4.50}$$

since $M = \frac{N}{2}$ and therefore we only have a term for $m = M$.

Evaluating S_2:

$$\begin{aligned} S_2 &= \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}T^{**}_{p,y,M}\left(\frac{j}{N}\right)^2 \\ &= \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{M<n,m\leq x}\left(\frac{m}{p}\right)_y\left(\frac{n}{p}\right)_y\frac{1}{mn}\sum_{j=0}^{N-1}\cos\left(\frac{2\pi mj}{N}\right)\cos\left(\frac{2\pi nj}{N}\right) \\ &= S_2' + S_2'' \end{aligned} \tag{5.4.51}$$

where S_2' and S_2'' are the non-diagonal and diagonal terms respectively and are defined as follows:

$$S_2' = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{\substack{M<n,m\leq x\\ n\neq m}}\left(\frac{m}{p}\right)_y\left(\frac{n}{p}\right)_y\frac{1}{mn}\sum_{j=0}^{N-1}\cos\left(\frac{2\pi mj}{N}\right)\cos\left(\frac{2\pi nj}{N}\right) \tag{5.4.52}$$

$$S_2'' = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{M<m\leq x}\left(\frac{m^2}{p}\right)_y\frac{1}{m^2}\sum_{j=0}^{N-1}\cos^2\left(\frac{2\pi mj}{N}\right) \tag{5.4.53}$$

We will first evaluate the non-diagonal term S_2'. Comparing S_1' as stated in (5.4.21) and S_2' as stated in (5.4.52), we notice that the only difference in these two sums are the ranges of m and n. Hence, in the same way as we evaluated the sum over the j's in S_1', we can evaluate the sum over the j's in S_2' and thus in the same way as we obtained (5.4.29), for S_2' we have:

$$S_2' = \frac{1}{2|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\left(\sum_{\substack{M<n,m\leq x\\ m\neq n\\ N|(m\pm n)}}\left(\frac{m}{p}\right)_y\left(\frac{n}{p}\right)_y\frac{1}{mn}\right) = \frac{1}{2|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\left(\sum_{\substack{M<n,m\leq x\\ m\neq n\\ N|(m\pm n)\\ P(m),P(n)>y}}\left(\frac{mn}{p}\right)\frac{1}{mn}\right). \tag{5.4.54}$$

Now, write the inner sum as $D_1 + D_2$ where

$$D_1 = \sum_{\substack{M<n,m\leq x \\ m\neq n \\ N|(m\pm n) \\ P(m),P(n)>y \\ mn=\square r, P(r)\leq y}} \left(\frac{mn}{p}\right)\frac{1}{mn}, \tag{5.4.55}$$

$$D_2 = \sum_{\substack{M<n,m\leq x \\ m\neq n \\ N|(m\pm n) \\ P(m),P(n)>y \\ mn=\square r, P(r)>y}} \left(\frac{mn}{p}\right)\frac{1}{mn}. \tag{5.4.56}$$

To evaluate D_1, write $m = dus^2$, $n = dvt^2$ where for q prime, $q|d \Rightarrow q > y$ and $P(uv) \leq y$. And so we have

$$|D_1| < \sum_{\substack{d\leq x \\ q|d\Rightarrow q>y}} \frac{1}{d^2} \sum_{\substack{us^2,vt^2\leq x \\ N|d(us^2\pm vt^2) \\ P(uv)\leq y}} \frac{1}{(us^2)(vt^2)}. \tag{5.4.57}$$

Since N is such that for q prime, $q|N \Rightarrow q \leq y$ and d is such that for q prime, $q|d \Rightarrow q > y$, it follows that $(N, d) = 1$ and we obtain:

$$|D_1| < \sum_{\substack{d\leq x \\ q|d\Rightarrow q>y}} \frac{1}{d^2} \sum_{P(uv)\leq y} \frac{1}{uv} \sum_{s\geq 1} \frac{1}{s^2} \sum_{t\geq 1} \frac{1}{t^2} \ll \frac{1}{y\log y}\log^2 y = \frac{\log y}{y}. \tag{5.4.58}$$

Substituting the estimate obtained for D_1 in (5.4.54) gives us

$$S_2' \ll \frac{\log y}{y} + \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x} D_2, \tag{5.4.59}$$

and we can rearrange $\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x} D_2$ to give

$$\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x} D_2 \leq \sum_{\substack{M<n,m\leq x \\ m\neq n \\ N|(m\pm n) \\ P(m),P(n)>y \\ mn=\square r, P(r)>y}} \frac{1}{mn}\frac{1}{|\mathcal{P}_x|}\left|\sum_{p\in\mathcal{P}_x}\left(\frac{mn}{p}\right)\right|. \tag{5.4.60}$$

Owing to the conditions on mn, we can apply Lemma 5.3.6, which gives

$$\frac{1}{|\mathcal{P}_x|}\left|\sum_{p\in\mathcal{P}_x}\left(\frac{mn}{p}\right)\right| \ll \frac{\log(mnx)\log x}{x^{\frac{1}{10}}} \ll \frac{\log^2 x}{x^{\frac{1}{10}}} \tag{5.4.61}$$

and hence, substituting (5.4.61) in (5.4.60), gives

$$\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x} D_2 \ll \frac{\log^2 x}{x^{\frac{1}{10}}} \sum_{\substack{M<n,m\le x\\ m\neq n\\ N|(m\pm n)\\ P(m),P(n)>y\\ mn=\square r,P(r)>y}} \frac{1}{mn} < \frac{\log^2 x}{x^{\frac{1}{10}}}\Bigg(\underbrace{\sum_{\substack{M<n,m\le x\\ N|(m+n)}}\frac{1}{mn}}_{A_1} + \underbrace{\sum_{\substack{M<n,m\le x\\ N|(m-n)}}\frac{1}{mn}}_{A_2}\Bigg). \tag{5.4.62}$$

We will now evaluate A_1 and A_2. To evaluate A_1, write $m \equiv -n \equiv k \bmod N$ where $1 \le k \le N$, so that

$$A_1 = \sum_{\substack{M<n,m\le x\\ N|(m+n)}}\frac{1}{mn} = \sum_{k=1}^{N}\underbrace{\Bigg(\sum_{\substack{m\equiv k \bmod N\\ M<m\le x}}\frac{1}{m}\Bigg)}_{A_1^+}\underbrace{\Bigg(\sum_{\substack{n\equiv -k \bmod N\\ M<n\le x}}\frac{1}{n}\Bigg)}_{A_1^-}. \tag{5.4.63}$$

To evaluate A_1^+, we write $m = k + Nu$, $\frac{M-k}{N} < u \le \frac{x-k}{N}$, so that

$$A_1^+ = \sum_{\substack{m\equiv k \bmod N\\ M<m\le x}}\frac{1}{m} = \sum_{\frac{M-k}{N}<u\le\frac{x-k}{N}}\frac{1}{k+Nu} \le \int_{\frac{M-k}{N}}^{\frac{x-k}{N}}\frac{1}{Nt+k}\,dt \ll \frac{\log x}{N}. \tag{5.4.64}$$

Similarly, to evaluate A_1^-, we write $n = Nu - k$, $\frac{M+k}{N} < u \le \frac{x+k}{N}$, so that

$$A_1^- = \sum_{\substack{n\equiv -k \bmod N\\ M<n\le x}}\frac{1}{n} = \sum_{\frac{M+k}{N}<u\le\frac{x+k}{N}}\frac{1}{Nu-k} \le \int_{\frac{M+k}{N}}^{\frac{x+k}{N}}\frac{1}{Nt-k}\,dt \ll \frac{\log x}{N}. \tag{5.4.65}$$

Thus using (5.4.64) and (5.4.65) in (5.4.63), we have:

$$A_1 \ll N\Big(\frac{\log x}{N}\Big)^2 = \frac{\log^2 x}{N}. \tag{5.4.66}$$

To evaluate A_2, write $m \equiv n \equiv k \bmod N$ where $1 \le k \le N$, so that

$$A_2 = \sum_{\substack{M<n,m\le x\\ N|(m-n)}}\frac{1}{mn} = \sum_{k=1}^{N}\Bigg(\sum_{\substack{m\equiv k \bmod N\\ M<m\le x}}\frac{1}{m}\Bigg)\Bigg(\sum_{\substack{n\equiv k \bmod N\\ M<n\le x}}\frac{1}{n}\Bigg) = \sum_{k=1}^{N}\Bigg(\sum_{\substack{m\equiv k \bmod N\\ M<m\le x}}\frac{1}{m}\Bigg)^2. \tag{5.4.67}$$

We observe that $\sum_{\substack{m\equiv k \bmod N\\ M<m\le x}}\frac{1}{m} = A_1^+$ as in (5.4.63). Thus by using (5.4.64) in (5.4.67) we have

$$A_2 \ll N\Big(\frac{\log x}{N}\Big)^2 = \frac{\log^2 x}{N}. \tag{5.4.68}$$

Then, by (5.4.62), (5.4.66) and (5.4.68), we have,

$$\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x} D_2 \ll \Big(\frac{\log^2 x}{x^{\frac{1}{10}}}\Big)\Big(\frac{\log^2 x}{N}\Big) = \frac{\log^4 x}{Nx^{\frac{1}{10}}} \tag{5.4.69}$$

and using this in (5.4.59), we have the non-diagonal term of S_2,

$$S_2' \ll \frac{\log y}{y} + \frac{\log^4 x}{Nx^{\frac{1}{10}}}. \tag{5.4.70}$$

Now we will evaluate the diagonal term of S_2 which is given in (5.4.53) as

$$S_2'' = \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{M<m\le x}\left(\frac{m^2}{p}\right)_y\frac{1}{m^2}\sum_{j=0}^{N-1}\cos^2\left(\frac{2\pi mj}{N}\right).$$

Appealing to (5.4.40), the inner sum $\sum_{j=0}^{N-1}\cos^2\left(\frac{2\pi mj}{N}\right)$ takes the value $\frac{N}{2}$ if $N\nmid 2m$ and N otherwise. And so,

$$\begin{aligned} S_2'' &= \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\left(\sum_{\substack{M<m\le x\\ N\mid 2m}}\left(\frac{m^2}{p}\right)_y\frac{1}{m^2}+\frac{1}{2}\sum_{\substack{M<m\le x\\ N\nmid 2m}}\left(\frac{m^2}{p}\right)_y\frac{1}{m^2}\right)\\ &< \sum_{\substack{M<m\le x\\ N\mid 2m}}\frac{1}{m^2}+\frac{1}{2}\sum_{\substack{M<m\le x\\ N\nmid 2m}}\frac{1}{m^2}\\ &< \frac{3}{2}\sum_{M<m\le x}\frac{1}{m^2} < \frac{3}{2}\int_M^x\frac{1}{t^2}dt \ll \frac{1}{M}. \end{aligned} \tag{5.4.71}$$

Finally, replacing the values of S_2' and S_2'' (as obtained in (5.4.70) and (5.4.71) respectively) in (5.4.51) we have

$$S_2 \ll \frac{\log^4 x}{Nx^{\frac{1}{10}}} + \frac{1}{M} + \frac{\log y}{y}$$

and hence have proved (5.4.14).

Evaluating S_4:

$$\begin{aligned} S_4 &= \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\int_{-1/2}^{1/2} T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)^2 d\theta\\ &= \frac{1}{N|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\sum_{M<n,m\le x}\frac{\left(\frac{m}{p}\right)_y\left(\frac{n}{p}\right)_y}{mn}\sum_{j=0}^{N-1}\int_{-1/2}^{1/2}\cos\left(\frac{2\pi m(j+\theta)}{N}\right)\cos\left(\frac{2\pi n(j+\theta)}{N}\right)d\theta\\ &= S_4' + S_4'' \end{aligned} \tag{5.4.72}$$

where S_4' and S_4'' are the non-diagonal and diagonal terms respectively and are defined as follows:

$$S_4' = \frac{1}{N|\mathcal{P}_x|} \sum_{p\in\mathcal{P}_x} \sum_{\substack{M<n,m\leq x \\ m\neq n}} \frac{\left(\frac{m}{p}\right)_y \left(\frac{n}{p}\right)_y}{mn} \sum_{j=0}^{N-1} \int_{-1/2}^{1/2} \cos\left(\frac{2\pi m(j+\theta)}{N}\right) \cos\left(\frac{2\pi n(j+\theta)}{N}\right) d\theta \tag{5.4.73}$$

$$S_4'' = \frac{1}{N|\mathcal{P}_x|} \sum_{p\in\mathcal{P}_x} \sum_{M<m\leq x} \left(\frac{m^2}{p}\right)_y \frac{1}{m^2} \sum_{j=0}^{N-1} \int_{-1/2}^{1/2} \cos^2 \frac{2\pi m(j+\theta)}{N} d\theta. \tag{5.4.74}$$

We will first evaluate the non-diagonal term S_4'. Comparing S_3' as stated in (5.4.22) and S_4' as stated in (5.4.73), we notice that the only difference in these two sums are the ranges of m and n. Hence, evaluating the integral in S_4' in the same way as we evaluated the integral for S_3', we obtain (as we obtained (5.4.33) for S_3'):

$$\begin{aligned} S_4' \;=\; & \frac{1}{2\pi|\mathcal{P}_x|} \sum_{p\in\mathcal{P}_x} \Bigg(\sum_{\substack{M<n,m\leq x \\ m\neq n \\ N\nmid(m+n)}} \frac{\left(\frac{m}{p}\right)_y \left(\frac{n}{p}\right)_y}{(m+n)mn} \sin\left(\frac{\pi(m+n)}{N}\right) \sum_{j=0}^{N-1} \cos\left(\frac{2\pi(m+n)j}{N}\right) \\ & + \sum_{\substack{M<n,m\leq x \\ m\neq n \\ N\nmid(m-n)}} \frac{\left(\frac{m}{p}\right)_y \left(\frac{n}{p}\right)_y}{(m-n)mn} \sin\left(\frac{\pi(m-n)}{N}\right) \sum_{j=0}^{N-1} \cos\left(\frac{2\pi(m-n)j}{N}\right) \Bigg). \end{aligned} \tag{5.4.75}$$

Further owing to (5.4.34), it follows that the non-diagonal term of S_4,

$$S_4' = 0 \tag{5.4.76}$$

and hence we have

$$S_4 = S_4''. \tag{5.4.77}$$

Now we will evaluate the diagonal term S_4''. To do so, we will begin by evaluating the integral in S_4'' as given in (5.4.74):

$$\int_{-1/2}^{1/2} \cos^2\left(\frac{2\pi m(j+\theta)}{N}\right) d\theta = \frac{1}{2} \int_{-1/2}^{1/2} \cos\left(\frac{4\pi m(j+\theta)}{N}\right) d\theta + \frac{1}{2}. \tag{5.4.78}$$

By (5.4.32), we have that

$$\int_{-1/2}^{1/2} \cos\left(\frac{4\pi m(j+\theta)}{N}\right) d\theta = \begin{cases} \frac{N}{2\pi m} \sin\left(\frac{2\pi m}{N}\right) \cos\left(\frac{4\pi mj}{N}\right) & \text{if } N \nmid 2m, \\ 0 & \text{otherwise.} \end{cases} \tag{5.4.79}$$

Using (5.4.79) in (5.4.78) and the subsequent result in (5.4.74), we have:

$$\begin{aligned} S_4'' &= \frac{1}{|\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \frac{1}{N} \sum_{\substack{M < m \leq x \\ N \nmid 2m}} \left(\frac{m^2}{p}\right)_y \frac{1}{m^2} \sum_{j=0}^{N-1} \left(\frac{N}{4\pi m} \sin\left(\frac{2\pi m}{N}\right) \cos\left(\frac{4\pi mj}{N}\right) + \frac{1}{2}\right) \\ &+ \frac{1}{|\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \frac{1}{N} \sum_{\substack{M < m \leq x \\ N \mid 2m}} \left(\frac{m^2}{p}\right)_y \frac{1}{m^2} \sum_{j=0}^{N-1} \frac{1}{2} \\ &\leq \frac{1}{4\pi |\mathcal{P}_x|} \sum_{p \in \mathcal{P}_x} \sum_{\substack{M < m \leq x \\ N \nmid 2m}} \left(\frac{m^2}{p}\right)_y \sin\left(\frac{2\pi m}{N}\right) \frac{1}{m^3} \sum_{j=0}^{N-1} \cos\left(\frac{4\pi mj}{N}\right) + \frac{1}{2} \sum_{M < m \leq x} \frac{1}{m^2}. \end{aligned} \tag{5.4.80}$$

We notice that the second term in (5.4.80) evaluates as follows:

$$\frac{1}{2} \sum_{M < m \leq x} \frac{1}{m^2} < \frac{1}{2} \int_M^x \frac{1}{t^2}\, dt \ll \frac{1}{M}. \tag{5.4.81}$$

In the first term of (5.4.80), notice that we are summing over all $M < m \leq x$ for which $N \nmid 2m$. Moreover, by appealing to (5.4.34), we have that the inner sum of the first term of (5.4.80), given by $\sum_{j=0}^{N-1} \cos\left(\frac{4\pi mj}{N}\right) = 0$ when $N \nmid 2m$, and so the term in question equals zero. Hence by using this fact and substituting (5.4.81) in (5.4.80) and using this result in (5.4.77), we have

$$S_4 \ll \frac{1}{M}$$

and hence have proved (5.4.15).

Evaluating S_5 and S_6:

To evaluate S_5 and S_6, we will appeal to the results of S_1, S_2, S_3 and S_4 and use Cauchy-Schwarz's inequality. We will first evaluate S_5 by applying Cauchy's

inequality successively in the following way:

$$\begin{aligned} S_5 &= \frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}T^*_{p,y,M}\left(\frac{j}{N}\right)T^{**}_{p,y,M}\left(\frac{j}{N}\right) \\ &\leq \frac{1}{N|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\left(\sqrt{\sum_{j=0}^{N-1}T^*_{p,y,M}\left(\frac{j}{N}\right)^2}\sqrt{\sum_{j=0}^{N-1}T^{**}_{p,y,M}\left(\frac{j}{N}\right)^2}\right) \\ &\leq \sqrt{\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}T^*_{p,y,M}\left(\frac{j}{N}\right)^2}\sqrt{\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}T^{**}_{p,y,M}\left(\frac{j}{N}\right)^2} \\ &= \sqrt{S_1}\sqrt{S_2}. \end{aligned} \tag{5.4.82}$$

And owing to (5.4.19) and (5.4.14) for the values of S_1 and S_2 respectively, we have

$$S_5 \ll \sqrt{\left(\frac{\log^4 x}{Nx^{\frac{1}{10}}}+\frac{1}{M}+\frac{\log y}{y}\right)} < \frac{\log^2 x}{x^{\frac{1}{20}}\sqrt{N}}+\frac{1}{\sqrt{M}}+\frac{\sqrt{\log y}}{\sqrt{y}}$$

and hence have proved (5.4.16).

Finally we evaluate S_6. By Cauchy-Schwarz, we have

$$\begin{aligned} S_6 &= \frac{1}{N|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\sum_{j=0}^{N-1}\left(\int_{-1/2}^{1/2}T^*_{p,y,M}\left(\frac{j+\theta}{N}\right)T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)\,d\theta\right) \\ &\leq \frac{1}{N|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\sum_{j=0}^{N-1}\left(\sqrt{\int_{-1/2}^{1/2}T^*_{p,y,M}\left(\frac{j+\theta}{N}\right)^2\,d\theta}\sqrt{\int_{-1/2}^{1/2}T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)^2\,d\theta}\right) \\ &\leq \frac{1}{N|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\sqrt{\sum_{j=0}^{N-1}\int_{-1/2}^{1/2}T^*_{p,y,M}\left(\frac{j+\theta}{N}\right)^2\,d\theta}\sqrt{\sum_{j=0}^{N-1}\int_{-1/2}^{1/2}T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)^2\,d\theta} \\ &\leq \sqrt{\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\int_{-\frac{1}{2}}^{\frac{1}{2}}T^*_{p,y,M}\left(\frac{j+\theta}{N}\right)^2\,d\theta}\sqrt{\frac{1}{|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}\frac{1}{N}\sum_{j=0}^{N-1}\int_{-\frac{1}{2}}^{\frac{1}{2}}T^{**}_{p,y,M}\left(\frac{j+\theta}{N}\right)^2\,d\theta} \\ &\leq \sqrt{S_3}\sqrt{S_4}. \end{aligned}$$

And owing to (5.4.20) and (5.4.15) for the values of S_3 and S_4 respectively, we have:

$$S_6 \ll \frac{1}{\sqrt{M}}$$

and hence have proved (5.4.17). □

The following corollary evaluates an explicit value for k_1 for all $y \geq 29$ since in our computations we will consider values of $y \geq 29$.

Corollary 5.4.3. *Assume the same hypothesis as in Lemma 5.4.2. Then, for $k_1, k_2 > 0$, there exist at least $\frac{|\mathcal{P}_x|}{2}$ primes in $\mathcal{P}_x$, for which*

$$\left|\frac{1}{N}\sum_{j=0}^{N-1} T_{p,y}^*\left(\frac{j}{N}\right)^2 - \int_0^1 T_{p,y}^*(t)^2\, dt\right| < \frac{k_1\sqrt{\log y}}{\sqrt{y}} + \frac{k_2}{\sqrt{N}} \tag{5.4.83}$$

where N is such that $P(N) \leq y$. Moreover, for $y \geq 29$, $k_1 = 3.84$.

PROOF. By (5.4.3) of Lemma 5.4.2 we have that for $c_1 > 0$ and N such that $P(N) \leq y$,

$$\left|\frac{1}{N}\sum_{j=0}^{N-1} T_{p,y}^*\left(\frac{j}{N}\right)^2 - \int_0^1 T_{p,y}^*(t)^2\, dt\right| < c_1\Big(\frac{\sqrt{\log y}}{\sqrt{y}} + \frac{1}{\sqrt{N}}\Big).$$

For $k_1, k_2 > 0$ such that $c_1 = \max(k_1, k_2)$, the above follows directly from (5.4.83).

Now, to prove the second assertion of this corollary, we will explicitly evaluate the value of k_1 for $y \geq 29$. From the proof of Lemma 5.4.2, we observe that the $\frac{\sqrt{\log y}}{\sqrt{y}}$ term arises from the evaluation of the sums S_2 (5.4.14) and S_5 (5.4.16). In the case of S_2, the term depending on y comes from the non-diagonal terms of S_2, i.e from the sum $S_2' = \frac{1}{2|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x}(D_1 + D_2)$ (see (5.4.54)) where D_1 is as in (5.4.55) and D_2 is as in (5.4.56). More precisely, the sum $\frac{1}{2|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x} D_1$ evaluates as $\ll \frac{\log y}{y}$ (see (5.4.58)). In the case of S_5, (5.4.82) gives $S_5 \leq \sqrt{S_1 S_2}$ and since $S_1 < 1$ by (5.4.19), we obtain a $\frac{\sqrt{\log y}}{\sqrt{y}}$ term in the evaluation of S_5 which is basically the square root of the term obtained for S_2.

We will first evaluate the value of the implied constant in (5.4.58), and hence will study the sum $\frac{1}{2|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x} D_1$ where D_1 is as in (5.4.55). We have:

$$\frac{1}{2|\mathcal{P}_x|}\sum_{p\in\mathcal{P}_x} D_1 < \frac{1}{2}\sum_{\substack{M<n,m\leq x\\ m\neq n\\ N|(m\pm n)\\ P(m),P(n)>y\\ mn=\square r, P(r)\leq y}} \frac{1}{mn}.$$

Taking $m = dus^2$ and $n = dvt^2$ where for q prime, $q|d \Rightarrow q > y$ and $P(uv) \leq y$, and since $(N, d) = 1$ (recall that N is such that $P(N) \leq y$) we have

$$\frac{1}{2} \sum_{\substack{M<n, m\leq x \\ m\neq n \\ N|(m\pm n) \\ P(m),P(n)>y \\ mn=\square r, P(r)\leq y}} \frac{1}{mn} < \frac{1}{2} \sum_{\substack{d\leq x \\ q|d\Rightarrow q>y}} \frac{1}{d^2} \sum_{\substack{us^2, vt^2\leq x \\ N|(us^2\pm vt^2) \\ us^2\neq vt^2 \\ P(uv)\leq y}} \frac{1}{(us^2)(vt^2)}. \tag{5.4.84}$$

Now, the sum over d evaluates as follows:

$$\sum_{\substack{d\leq x \\ q|d\Rightarrow q>y}} \frac{1}{d^2} < \prod_{q>y} \left(1 - \frac{1}{q^2}\right)^{-1} - 1 < e^{\frac{1}{y\log y}} - 1. \tag{5.4.85}$$

Since $e^{\frac{1}{y\log y}} = \sum_{t\geq 0} \frac{1}{t!(y\log y)^t}$ and $\sum_{t\geq 4} \frac{1}{t!(y\log y)^t} \leq \frac{1}{(y\log y)^4} \sum_{t\geq 4} \frac{1}{t!} = \frac{e-8/3}{(y\log y)^4}$, we have

$$e^{\frac{1}{y\log y}} - 1 = \frac{1}{y\log y}\left(1 + \frac{1}{2y\log y} + \frac{1}{6(y\log y)^2} + \frac{e-8/3}{(y\log y)^3}\right)$$

and so for $y \geq 29$, the sum over d is $< \frac{1.005137}{y\log y}$.

Since $N|(ui^2 \pm vj^2)$ and $ui^2 \neq vj^2$, $\max(ui^2, vj^2) \geq \frac{N}{2}$, and we can write the inner sum in (5.4.84) as follows:

$$\sum_{\substack{us^2, vt^2\leq x \\ N|(us^2\pm vt^2) \\ us^2\neq vt^2 \\ P(uv)\leq y}} \frac{1}{(us^2)(vt^2)} \leq \sum_{1\leq i\leq N} \frac{1}{i^2} \sum_{P(u)\leq y} \frac{1}{u} \sum_{\substack{1\leq j\leq N \\ P(v)\leq y \\ vj^2>N/2}} \frac{1}{vj^2} \leq \frac{\pi^2}{6} \prod_{\substack{q\leq y \\ q \text{ prime}}} \left(1-\frac{1}{q}\right)^{-1} \sum_{\substack{1\leq j\leq N \\ P(v)\leq y \\ vj^2>N/2}} \frac{1}{vj^2}. \tag{5.4.86}$$

By identity (3.30) of page 70 of **[29]** we have for $y > 1$,

$$\prod_{\substack{q\leq y \\ q \text{ prime}}} \left(1 - \frac{1}{q}\right)^{-1} < e^{\gamma}\log y\left(1 + \frac{1}{\log^2 y}\right)$$

where γ is Euler's constant. Via simple computations, we can show that for $y \in \mathbb{Z}$ and $y \geq 24$,

$$\prod_{\substack{q\leq y \\ q \text{ prime}}} \left(1 - \frac{1}{q}\right)^{-1} < e^{\gamma}\log y\left(1 + \frac{0.822}{\log^2 y}\right)$$

and so for $y \geq 29$, we have

$$\prod_{\substack{q\leq y \\ q \text{ prime}}} \left(1 - \frac{1}{q}\right)^{-1} < 1.07249 e^{\gamma}\log y. \tag{5.4.87}$$

Next, assuming $N > 2y$, we evaluate the inner sum in (5.4.86) as follows:

$$\sum_{\substack{1\le j\le N\\ P(v)\le y\\ vj^2>N/2}} \frac{1}{vj^2} \le \sum_{\substack{1\le v\le N/2\\ P(v)\le y}} \frac{1}{v} \sum_{j>\sqrt{\frac{N}{2v}}} \frac{1}{j^2} + \frac{\pi^2}{6} \sum_{\substack{v>N/2\\ P(v)\le y}} \frac{1}{v}$$

$$\le \sqrt{\frac{2}{N}} \sum_{\substack{v\le N/2\\ P(v)\le y}} \frac{1}{\sqrt{v}} + \frac{\pi^2}{6}\Big(\sum_{v:P(v)\le y} \frac{1}{v} - \sum_{v\le y} \frac{1}{v}\Big). \tag{5.4.88}$$

By (5.4.87), for $y \ge 29$, $\sum_{v:P(v)\le y} \frac{1}{v} < 1.07249e^{\gamma}\log y$ and since $\sum_{v\le y} \frac{1}{v} > \log y + \gamma$, we get

$$\sum_{\substack{1\le j\le N\\ P(v)\le y\\ vj^2>N/2}} \frac{1}{vj^2} \le 2 + \frac{\pi^2}{6}(1.07249e^{\gamma}\log y - \log y - \gamma). \tag{5.4.89}$$

Substituting (5.4.89) and (5.4.87) in (5.4.86), and the result in (5.4.84), where the sum over $d < \frac{1.005137}{y\log y}$ we have:

$$\frac{1}{2} \sum_{\substack{M<n,m\le x\\ m\ne n\\ N\mid(m\pm n)\\ P(m),P(n)>y\\ mn=\square r,P(r)\le y}} \frac{1}{mn} < \frac{1.07799e^{\gamma}\pi^2(2-\frac{\pi^2\gamma}{6})}{12y} + \frac{1.07799e^{\gamma}\pi^4}{72}(e^{\gamma}-1)\frac{\log y}{y}$$

and so for $y \ge 29$,

$$\frac{1}{2} \sum_{\substack{M<n,m\le x\\ m\ne n\\ N\mid(m\pm n)\\ P(m),P(n)>y\\ mn=\square r,P(r)\le y}} \frac{1}{mn} < 2.4597\frac{\log y}{y}. \tag{5.4.90}$$

Recall that we take the square root of (5.4.90) to obtain the $\frac{\sqrt{\log y}}{\sqrt{y}}$ term of S_5 and further, owing to (5.4.7) we have to consider twice the resulting value. Therefore, for $y \ge 29$ we get

$$2.4597\frac{\log y}{y} + 2\sqrt{2.4597\frac{\log y}{y}} \le 3.84\sqrt{\frac{\log y}{y}} \tag{5.4.91}$$

which gives $k_1 = 3.84$. □

Remark 5.4.4. *The proof of Corollary 5.4.3 shows that we can in fact calculate an explicit value for k_1 for any given y.*

5.4.2. Proving the Main Theorems

Via the above lemmas, we have established the hypotheses of Theorems 5.2.1 and 5.2.3, and are now in a position to prove these theorems. The proofs show how we can apply our established hypotheses to obtain a bound for the discrete second moments $\frac{1}{N}\sum_{j=0}^{N-1}\left(T_p\left(\frac{j}{N}\right)-T_{\lambda,y}\left(\frac{j}{N}\right)\right)^2$. Further, as we remarked earlier, owing to the definition of λ, for fixed λ and y, we can compute values for a truncated series of $T_{\lambda,y}(\frac{j}{N})$ for $0 \leq j \leq N-1$. Therefore, with all the necessary tools in place, we can then simply map our scenario to the one set in the combinatorial lemmas and corollaries of Chapter 4. These computed values, will give a finite ordered set of elements such that the cardinality of its negative elements is known. The lemmas state that given such a set, if there exists any other ordered set of the same cardinality as the given set, such that the elements of the two sets are close enough, then we can deduce upper and lower bounds for the cardinality of the negative elements of the second set. Hence, applying these lemmas and corollaries we will subsequently obtain an upper and lower bound for $\beta(p)$, the measure of the set of $x \in [0,1]$ for which $T_p(x)$ is negative.

Both theorems have the same proof, the only difference lying in the fact that while the proof of Theorem 5.2.1 appeals to Corollary 4.1.5 to establish a formula for a lower bound for $\beta(p)$, the proof of Theorem 5.2.3 appeals to Lemma 4.1.6 to establish a formula for an upper bound for $\beta(p)$.

PROOF OF THEOREM 5.2.1. Define

$$S = \frac{1}{N}\sum_{j=0}^{N-1}\left(T_p\left(\frac{j}{N}\right)-T_{\lambda,y}\left(\frac{j}{N}\right)\right)^2$$

and

$$I = \int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\, dx.$$

By Lemma 5.4.2 and Corollary 5.4.3, we have that for any z sufficiently large, and k' a positive constant, there exist at least $ze^{\frac{-k'\log z}{\log\log z}}$ primes $p \leq z$ which are congruent to 3 (mod 4) and satisfy the condition that $\lambda(q) = \left(\frac{q}{p}\right)$ for all primes $q \leq y$, for which $|S-I| < \frac{k_1\sqrt{\log y}}{\sqrt{y}} + \frac{k_2}{\sqrt{N}}$. Therefore, by Remark 5.4.4 and given $\int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\, dx < c_y$, (5.2.1) and a lower bound for N as $\left(\frac{100k_2}{c-I}\right)^2$ in

(5.2.2), we have that for a positive proportion of such primes,

$$\begin{aligned} S \leq |S-I|+I &< \frac{k_1\sqrt{\log y}}{\sqrt{y}} + \frac{k_2}{\sqrt{N}} + I \\ &< \frac{k_1\sqrt{\log y}}{\sqrt{y}} + \frac{c_y}{100} + \frac{99I}{100} \qquad (5.4.92) \\ &< \frac{k_1\sqrt{\log y}}{\sqrt{y}} + c_y, \qquad (5.4.93) \end{aligned}$$

and (5.2.4) follows from Corollary 4.1.5. □

PROOF OF THEOREM 5.2.3. The proof of Theorem 5.2.3 follows in exactly the same way as the proof of Theorem 5.2.1. Only note that in this case, the hypothesis for the a_i's corresponds to the hypothesis for the a_i's in Lemma 4.1.6 and so once we have established that $S < \frac{k_1\sqrt{\log y}}{\sqrt{y}} + c_y$ as in the proof of Theorem 5.2.1, (5.2.5) follows from Lemma 4.1.6. □

5.5. PROOF OF THEOREM 0.0.2

We begin by recalling that by (5.1.2),

$$T_p(x) = -\sum_{n=1}^{\infty} \left(\frac{n}{p}\right) \frac{\cos 2\pi nx}{n},$$

by Definition 5.1.5,

$$T_{\lambda,y}(x) = -\sum_{\substack{n\geq 1 \\ P(n)\leq y}} \lambda(n) \frac{\cos 2\pi nx}{n},$$

by (5.1.3),

$$\beta(p) = |\{x \in [0,1] : T_p(x) < 0\}|,$$

and that for a fixed λ and y, the set $\mathcal{P}$ (Definition 5.3.4) contains those primes $p \equiv 3 \pmod 4$ for which $\lambda(q) = \left(\frac{q}{p}\right)$ for all primes $q \leq y$. Also recall that for a fixed λ and y, using the computed values of the truncated series of $T_{\lambda,y}(\frac{j}{N})$ for $0 \leq j \leq N-1$, for primes $p \in \mathcal{P}$, the main theorems, Theorem 5.2.1 and Theorem 5.2.3 respectively give an explicit formula for a lower and upper bound for $\beta(p)$. More precisely, for a fixed λ, y, (5.2.4) of Theorem 5.2.1 gives

$$\beta(p) \geq \frac{r-k}{N}$$

for a positive proportion of primes $p \in \mathcal{P}$, where N, r, k are as in Theorem 5.2.1. Similarly, for a fixed λ, y, (5.2.5) of Theorem 5.2.3 gives

$$\beta(p) \leq \frac{(N-r)+k}{N}$$

for a positive proportion of primes $p \in \mathcal{P}$, where N, r, k are as in Theorem 5.2.3. Note here that both r and k represent different quantities in the above two formulas.

Based on the above information, we will now proceed to prove Theorem 0.0.2. We notice that the statement of this theorem can be divided into two parts: the first gives a lower bound for $\mu(p)$ i.e. the measure of the set of $x \in [0,1]$ for which $S_p(x)$ is negative and the second gives an upper bound for the same. Thus we will first prove the first assertion of the theorem and then the second. Further, as was stated in the introduction of this chapter, we remarked that owing to the Fourier expansion of $S_p(x)$ as given by Pólya (see Lemma 3.1.1), we can bound the measure $\mu(p)$ by the measure $\beta(p)$ for a positive proportion of primes $p \in \mathcal{P}$. Thus, using the explicit formulas as given by Theorems 5.2.1 and 5.2.3, we will first calculate bounds for $\beta(p)$ where $p \in \mathcal{P}$. These calculations will follow from the algorithms of Appendix A, which are based on the main theorems, Theorem 5.2.1 and Theorem 5.2.3. While Algorithms A.1.2 and A.1.3 are based on Theorem 5.2.1 and give a lower bound for $\beta(p)$, Algorithms A.1.4 and A.1.5 are based on Theorem 5.2.3 and give an upper bound for $\beta(p)$. In the next few sections we will explain how (via these algorithms) we calculate a lower and an upper bound for $\beta(p)$, and using these results deduce the corresponding bound for $\mu(p)$. Having established this, the lower and upper bound results stated in Theorem 0.0.2 will immediately follow.

First, we state the following definitions:

Definition 5.5.1. *For a given y, $\Lambda_y = \{-1,1\}^n$, where $n = \pi(y)$ ($\pi(x) = \sum_{p \leq x} 1$, p prime).*

Definition 5.5.2. *Given λ as in Definition 5.1.2 and for a fixed y, $V_{\lambda,y}$ is a vector defined as:*

$$V_{\lambda,y} := (\lambda(q_1), \lambda(q_2), \ldots, \lambda(q_n)) \in \Lambda_y \tag{5.5.1}$$

where q_i is the i^{th} prime.

Remark 5.5.3. *The vector $V_{\lambda,y}$ as defined above then serves as a representation for a given λ-sequence of length n, where $n = \pi(y)$.*

5.5.1. Calculating a Lower Bound for $\beta(p)$

Here we will show how we apply Theorem 5.2.1 (which gives a lower bound for $\beta(p)$ by using Corollary 4.1.5) to determine an algorithm for obtaining a value for the lower bound for $\beta(p)$ where $p \in \mathcal{P}$. The actual algorithms written in pseudo-code, are available in Appendix A.

As is evident from Theorem 5.2.1, for a given λ, y, the algorithms will involve computing the values of the truncated series of $T_{\lambda,y}(\frac{j}{N})$ for $0 \leq j \leq N-1$. However, as the definitions of the function $T_{\lambda,y}(x)$ (Definition 5.1.5) and that of the vector $V_{\lambda,y}$ (Definition 5.5.2) suggest, the value of $T_{\lambda,y}(x)$ depends on the vector $V_{\lambda,y}$. Thus, for a fixed y, a lower bound for $\beta(p)$ will be calculated for each vector $V_{\lambda,y} \in \Lambda_y$ where $\lambda(q) = \left(\frac{q}{p}\right)$ for all $q \leq y$. This lower bound is given by $\frac{r-k}{N}$ in (5.2.4) where r, k are as in Theorem 5.2.1 and will be determined using the computed values of the truncated series of $T_{\lambda,y}(\frac{j}{N})$ for $0 \leq j \leq N-1$. For simplicity, we will denote this lower bound formula by $L(V_{\lambda,y})$. Therefore, for each vector $V_{\lambda,y}$ where $\lambda(q) = \left(\frac{q}{p}\right)$ for all $q \leq y$, and r, k as in Theorem 5.2.1 (where r and k depend on λ and y and are determined through calculation),

$$L(V_{\lambda,y}) = \frac{r-k}{N}. \tag{5.5.2}$$

The proof of Lemma 4.1.4 shows that the calculation of k (as above) depends on the value of the upper bound for the sum $S = \frac{1}{N}\sum_{j=0}^{N-1}(T_p(\frac{j}{N}) - T_{\lambda,y}(\frac{j}{N}))^2$. For computational purposes, we will assume that for $k_1 = 3.84$ (as evaluated in Corollary 5.4.3) and c_y as in Lemma 5.4.1, $\frac{y}{\log y} > \left(\frac{100k_1}{99(c_y - \int_0^1 (T_p(x) - T_{\lambda,y}(x))^2\, dx)}\right)^2$. Using this assumption in (5.4.92) of the proof of Theorem 5.2.1 gives $S < c_y$. We will then accordingly replace the condition in (5.2.3) (see the statement of Theorem 5.2.1) by $\sum_{j=0}^{k} a_j^2 \geq N c_y$ and calculate the corresponding value of k.

Now, given λ and y, first, Algorithm A.1.2 will be used to calculate values for the truncated series of $T_{\lambda,y}(\frac{j}{N})$ for $0 \leq j \leq N-1$. Using these computed values,

the algorithm will use the method as outlined in the statement of Theorem 5.2.1 to calculate the corresponding value for $L(V_{\lambda,y})$. For a fixed y, we will then use this algorithm to calculate the values of $L(V_{\lambda,y})$ for each vector $V_{\lambda,y} \in \Lambda_y$. Then, comparing the computed values of $L(V_{\lambda,y})$ for all $V_{\lambda,y} \in \Lambda_y$, a lower bound will be obtained for $\beta(p)$ (for the given y).

We note here that for a fixed λ, y, the explicit formula for a lower bound for $\beta(p)$ given by Theorem 5.2.1 holds for a positive proportion of primes $p \in \mathcal{P}$ i.e. for those primes $p \equiv 3 \bmod 4$ for which $\left(\frac{q}{p}\right) = \lambda(q)$ for all primes $q \leq y$. For such primes $p \in \mathcal{P}$, it follows that for larger values of y, $\left(\frac{q}{p}\right)$ equals $\lambda(q)$ for more primes q and so for such y, $\lambda(q)$ is a better model for $\left(\frac{q}{p}\right)$. Consequently, for larger y, $T_{\lambda,y}(x)$ will serve as a better approximation for $T_p(x)$ and as a result, for large y we will obtain better estimates for $\beta(p)$.
However, to calculate bounds for $\beta(p)$ (for a fixed y), we would like to first calculate $L(V_{\lambda,y})$ for all the vectors $V_{\lambda,y} \in \Lambda_y$. This would mean calculating $L(V_{\lambda,y})$ for 2^n ($n = \pi(y)$) vectors, and for large y, computing $L(V_{\lambda,y})$ for 2^n vectors would very quickly exceed our computing restrictions, both in terms of time and memory. Therefore, to make more efficient use of our computing resources, we will start our calculations with a small y and successively increment the value of y.

At each iteration, we will choose a set of "qualifying" vectors (this choice is explained in detail later) that will go through to the next round of computations for the new y. In this way, for each successive y we will be computing $L(V_{\lambda,y})$ for much less then $2^{\pi(y)}$ vectors and will eventually obtain an improved lower bound for $\beta(p)$. Algorithm A.1.3 will be used to perform these iterations and eventually give a final value for a lower bound for $\beta(p)$. We will give a brief explanation of the ideas employed in Algorithm A.1.3. Note that all references made are to steps of this algorithm.

We will begin by setting $y = 29$, i.e the 10^{th} prime. Then, via Algorithm A.1.2, for each vector $V_{\lambda,29} \in \Lambda_{29}$ we will compute the corresponding value of $L(V_{\lambda,29})$. Of all such values of $L(V_{\lambda,29})$ computed for the vectors $V_{\lambda,29} \in \Lambda_{29}$, we

will then pick the minimum and maximum values (step 4) so that:

$$MIN_{29} = \min_{V_{\lambda,29} \in \Lambda_{29}} (L(V_{\lambda,29})),$$

$$MAX_{29} = \max_{V_{\lambda,29} \in \Lambda_{29}} (L(V_{\lambda,29})).$$

These values imply that for $y = 29$,

$$\beta(p) \geq MIN_{29}$$

for a positive proportion of primes $p \in \mathcal{P}$ and of these primes, there is at least one prime for which

$$\beta(p) \geq MAX_{29}.$$

Based on these values of MIN_{29} and MAX_{29}, we will perform a series of "pruning operations" (steps 5 and 6). In these pruning operations we will choose only those vectors $V_{\lambda,29} \in \Lambda_{29}$ for which the value of the bound $L(V_{\lambda,29})$ is close to MIN_{29} or MAX_{29}. More precisely, we will consider parameters I_{min} and I_{max} (for more detail on the choice of these parameters, see step 5 and Remark A.2.1) and then set the following parameters:

$$B_{29,min} := MIN_{29} + I_{min},$$

$$B_{29,min} := MAX_{29} - I_{max}.$$

Next, in step 6, we will choose those vectors $V_{\lambda,29} \in \Lambda_{29}$ that qualify for one of the two sets $W_L(B_{29,min})$ and $Z_L(B_{29,max})$, which we define as follows:

$$W_L(B_{29,min}) := \{V_{\lambda,29} \in \Lambda_{29} : L(V_{\lambda,y}) < B_{29,min}\},$$

$$Z_L(B_{29,max}) := \{V_{\lambda,29} \in \Lambda_{29} : L(V_{\lambda,y}) > B_{29,max}\}.$$

For $y = 29$, we obtained $MIN_{29} = 0.205$ and $MAX_{29} = 0.560$. Following the strategy for choosing I_{min} and I_{max} as explained in Remark A.2.1; for $y = 29$, we put $I_{min} = 0.04$ and $I_{max} = 0.03$ which gave $B_{29,min} = 0.245$, $B_{29,max} = 0.53$, $|W_L(B_{29,min})| = 107$ and $|Z_L(B_{29,max})| = 27$, i.e. in all 134 vectors qualified for the next round of computations.

Having chosen the "qualifying" vectors; in preparation for the next iteration, we will increment the value of y by the next five primes, i.e. $y = 47$ (the 15^{th}

prime). For this new value of y, in step 9 we will determine new sets $\Lambda'_y \subset \Lambda_y$ and $\Lambda''_y \subset \Lambda_y$ which for $y = 47$, we define as follows:

$$\Lambda'_{47} = \{(e_1|e_2) \in \Lambda_{47} : e_1 \in W_L(B_{29,min}), e_2 \in \{1,-1\}^5\},$$

$$\Lambda''_{47} = \{(e_1|e_2) \in \Lambda_{47} : e_1 \in Z_L(B_{29,max}), e_2 \in \{1,-1\}^5\}.$$

Thus, in effect the vectors in Λ'_{47} and Λ''_{47} are obtained by appending all possible $\{1,-1\}$ sequences (of length 5) to each of the "qualifying" vectors. It is these vectors of Λ'_{47} and Λ''_{47} that will be used in the next iteration, where for each of them, we will compute $L(V_{\lambda,47})$, and consequently obtain a value for MIN_{47} (from the vectors of Λ'_{47}) and MAX_{47} (from the vectors of Λ''_{47}). As before, based on the values of MIN_{47} and MAX_{47}, the sets $W_L(B_{47,min})$ and $Z_L(B_{47,max})$ will be determined. In case the termination condition of step 7 is not satisfied, we will continue iterating in the same manner i.e. increment y by the next 5 primes, compute $L(V_{\lambda,y})$ for all the chosen vectors $V_{\lambda,y}$, determine MIN_y and MAX_y and consequently determine the sets $W_L(B_{y,min})$ and $Z_L(B_{y,max})$, until we meet the termination conditions of step 7. (For an explanation of the choice of the termination conditions refer to Remark A.2.2).

Table B.1.1 of Appendix B gives a summary of the results obtained whereby the values of MIN_y and MAX_y (see step 4 of Algorithm A.1.3) for all y for which computations were performed are tabulated.

Note that when the termination condition is satisfied, the value of MIN_y at that iteration is stored in $MINVAL$ and similarly the value of MAX_y (at the iteration at which the termination condition for it is satisfied) is stored in $MAXVAL$, and ultimately these values of $MINVAL$ and $MAXVAL$ are returned. It is these values that will give the final lower bound value for $\beta(p)$ where $p \in \mathcal{P}$. From Table B.1.1, we observe that the values of $MINVAL$ and $MAXVAL$ returned at step 7 of Algorithm A.1.3 are:

$$MINVAL = 0.254,$$

$$MAXVAL = 0.659.$$

Now, since Theorem 5.2.1 tells us that the bound obtained for $\beta(p)$ holds for a positive proportion of primes $p \in \mathcal{P}$ (Definition 5.3.4), we conclude the following:

For a positive proportion of primes $p \equiv 3 \mod 4$,

$$\beta(p) \geq 0.254 \tag{5.5.3}$$

and of these primes, there is at least one for which

$$\beta(p) \geq 0.659. \tag{5.5.4}$$

5.5.2. A Lower Bound for $\mu(p)$

In this section, we will begin by recalling our objective, which is essentially to prove the lower bound results for $\mu(p)$ as stated in Theorem 0.0.2. We will finally prove how we can reduce the question at hand to finding a lower bound for $\beta(p)$. Given $\mu(p)$ is the measure of the set of $x \in [0,1]$ for which $S_p(x)$ is negative, we recall Lemma 3.1.1 from which we deduce that this is the same as the measure of the set of $x \in [0,1]$ for which

$$L_p(1) + T_p(x) \leq 0. \tag{5.5.5}$$

Via the Law of Quadratic Reciprocity and a direct application of Merten's Theorem, the following lemma shows that we can in fact simplify (5.5.5) by showing that $L_p(1)$ is very small for a positive proportion of primes $p \in \mathcal{P}$ (Definition 5.3.4).

Lemma 5.5.4. *Fix an $\epsilon > 0$. For all λ, y, there exists a positive proportion of primes $p \equiv 3 \mod 4$ such that $\lambda(q) = \left(\frac{q}{p}\right)$ for all primes $q \leq y$ and $L_p(1) < \epsilon$.*

PROOF. Select b_q (prime) such that $\left(\frac{q}{b_q}\right) = \lambda(q)$ for all primes $q \leq y$ and $\left(\frac{q}{b_q}\right) = -1$ for all primes q such that $y < q \leq e^y$. Put $Q = 4 \prod_{p \leq e^y} p$. Then, by quadratic reciprocity and the Chinese Remainder Theorem we can determine a $b \pmod{Q}$ such that $b \equiv b_q \pmod{Q}$ and $b \equiv 3 \pmod 8$. This will give a prime $p \in \mathcal{P}$ that is congruent to $b \pmod{Q}$.

Using Montgomery's formula [**23**] for the asymptotic average of $L_p(1)$ for primes $p \equiv b \pmod{Q}$ (3.2.5), and substituting our function $\lambda(q)$ for Montgomery's function a_p, we have an asymptotic average of $L_p(1)$ for primes $p \equiv b \pmod{Q}$

given by:

$$\prod_{q\leq y}\left(1-\frac{\lambda(q)}{q}\right)^{-1}\prod_{y<q\leq e^y}\left(1+\frac{1}{q}\right)^{-1}.$$

By Definition 5.1.2, $\lambda(q)=-1$ or $+1$, and hence $\lambda(q)\leq 1$ for all primes $q\leq y$, therefore:

$$L_p(1)<\prod_{q\leq y}\left(1-\frac{1}{q}\right)^{-1}\prod_{y<q\leq e^y}\left(1+\frac{1}{q}\right)^{-1}.$$

By Merten's Theorem (1.4.1), the first product is approximately $e^{\gamma}\ln y$. We now calculate an approximate value for the second product:

$$\begin{aligned}\prod_{y<q\leq e^y}\left(1+\frac{1}{q}\right)^{-1} &= \frac{\prod_{q\leq e^y}\left(1-\frac{1}{q}\right)\prod_{q\leq e^y}\left(1-\frac{1}{q^2}\right)^{-1}}{\prod_{q\leq y}\left(1-\frac{1}{q}\right)\prod_{q\leq y}\left(1-\frac{1}{q^2}\right)^{-1}} \sim \frac{e^{-\gamma}/\ln e^y\prod_{q\leq e^y}\left(1-\frac{1}{q^2}\right)^{-1}}{e^{-\gamma}/\ln y\prod_{q\leq y}\left(1-\frac{1}{q^2}\right)^{-1}}\\ &= \frac{\ln y}{\ln e^y}\prod_{y<q\leq e^y}\left(1-\frac{1}{q^2}\right)^{-1}\sim\frac{\ln y}{y}\qquad \text{as } y\to\infty.\end{aligned}$$

Hence, if y is sufficiently large, then an asymptotic average of $L_p(1)$ for primes $p\equiv b \bmod Q$ is approximately $\frac{e^{\gamma}(\ln y)^2}{y}<\epsilon$. And so, since we have shown that on average, $L_p(1)<\epsilon$ for primes $p\equiv b \bmod Q$. It follows that $L_p(1)<\epsilon$ for a positive proportion of primes $p\equiv b \bmod Q$ and consequently $L_p(1)<\epsilon$ for a positive proportion of primes $p\in\mathcal{P}$ i.e. primes $p\equiv 3 \bmod 4$ for which $\lambda(q)=\left(\frac{q}{p}\right)$ for all primes $q\leq y$. □

By (5.5.5) and Lemma 5.5.4, we then have that there exists a positive proportion of primes $p\in\mathcal{P}$ for which

$$S_p(x)<T_p(x)+\epsilon \tag{5.5.6}$$

and hence, for such primes $p\in\mathcal{P}$,

$$\mu(p)=\int\limits_{\substack{0<x<1\\ x:S_p(x)\leq 0}}1\,dx\geq\int\limits_{\substack{0<x<1\\ x:T_p(x)<-\epsilon}}1\,dx\geq\int\limits_{\substack{0<x<1\\ x:T_p(x)<0}}1\,dx=\beta(p). \tag{5.5.7}$$

We have thus proved that there exists a positive proportion of primes $p\in\mathcal{P}$ for which we can reduce our objective of finding a lower bound for $\mu(p)$, to finding a lower bound for $\beta(p)$. We will now define this particular set of primes $\mathcal{P}'\subset\mathcal{P}$ as follows:

Definition 5.5.5. *For $\epsilon > 0$,*

$$\mathcal{P}' := \{p \in \mathcal{P} : L_p(1) < \epsilon\}$$

where $\mathcal{P}$ is as in Definition 5.3.4 and $L_p(1) = \sum_{n\geq 1} \left(\frac{n}{p}\right) \frac{1}{n}$.

Then by (5.5.7), for primes $p \in \mathcal{P}'$, we have $\mu(p) \geq \beta(p)$. And so, by (5.5.3) and (5.5.4), we can deduce the following:
For a positive proportion of primes $p \equiv 3 \bmod 4$,

$$\mu(p) \geq 0.254 \tag{5.5.8}$$

and of these primes, there is at least one for which

$$\mu(p) \geq 0.659. \tag{5.5.9}$$

Hence we have proved the first assertion of Theorem 0.0.2.

5.5.3. Calculating an Upper Bound for $\beta(p)$

The calculation of the upper bound for $\beta(p)$ $(p \in \mathcal{P})$ will follow in essentially the same way as the calculation of its lower bound. In this case we will formulate Algorithms A.1.4 and A.1.5 based on Theorem 5.2.3. (which gives an explicit formula for an upper bound for $\beta(p)$ by using Lemma 4.1.6). These algorithms are available in Appendix A. First, for a fixed y, for each vector $V_{\lambda,y} \in \Lambda_y$ where $\lambda(q) = \left(\frac{q}{p}\right)$ for all $q \leq y$ we will compute values for the truncated series of $T_{\lambda,y}(\frac{j}{N})$ for $0 \leq j \leq N-1$, and using these values, we will then calculate an upper bound for $\beta(p)$, where $p \in \mathcal{P}$. To do this, we will first determine r, k (as in Theorem 5.2.3) by using the computed values of the truncated series of $T_{\lambda,y}(x)$ for $0 \leq j \leq N-1$. For these values of r and k, Theorem 5.2.3 gives an explicit formula for an upper bound for $\beta(p)$ $(p \in \mathcal{P})$ as $\frac{(N-r)+k}{N}$ (see (5.2.5)). For simplicity, we will denote this upper bound formula by $U(V_{\lambda,y})$. Therefore, for each vector $V_{\lambda,y} \in \Lambda_y$ where $\lambda(q) = \left(\frac{q}{p}\right)$ for all $q \leq y$, and r, k as in Theorem 5.2.3, (where r and k depend on λ and y and are determined by calculation),

$$U(V_{\lambda,y}) = \frac{(N-r)+k}{N}. \tag{5.5.10}$$

For a fixed y, Algorithm A.1.4 is used to compute the values of $U(V_{\lambda,y})$ for each $V_{\lambda,y} \in \Lambda_y$. Using the values obtained for $U(V_{\lambda,y})$ for each $V_{\lambda,y} \in \Lambda_y$ for successive values of y, an explicit upper bound is obtained for $\beta(p)$ by the iterative Algorithm A.1.5. (The iterative procedure employs exactly the same methods as Algorithm A.1.3 did for computing the explicit lower bound for $\beta(p)$). Table B.1.2 of Appendix B gives a summary of the results obtained whereby the values of $MIN_y = \min_{V_{\lambda,y} \in \Lambda_y} (U(V_{\lambda,y}))$ and $MAX_y = \min_{V_{\lambda,y} \in \Lambda_y} (U(V_{\lambda,y}))$ (see step 4 of Algorithm A.1.5) for all y, for which computations were performed are tabulated.

As in Algorithm A.1.3, in Algorithm A.1.5 too, the $MINVAL$ and $MAXVAL$ values returned give the final upper bound values for $\beta(p)$. From Table B.1.2, we observe that the values of $MINVAL$ and $MAXVAL$ returned at step 7 of Algorithm A.1.5 are:

$$MINVAL = 0.369,$$

$$MAXVAL = 0.715.$$

Now, since Theorem 5.2.3 tells us that the bound obtained for $\beta(p)$ holds for a positive proportion of primes $p \in \mathcal{P}$ (Definition 5.3.4), we conclude the following: For a positive proportion of primes $p \equiv 3 \mod 4$,

$$\beta(p) \leq 0.715, \tag{5.5.11}$$

and of these primes, there is at least one for which

$$\beta(p) \leq 0.369. \tag{5.5.12}$$

5.5.4. An Upper Bound for $\mu(p)$

As we did in section 5.5.2, we first prove how we can reduce the question of finding an upper bound for $\mu(p)$ to finding an upper bound for $\beta(p)$. We again recall that Lemma 3.1.1 implies that the measure $\mu(p)$ is the same as the measure of the set $x \in [0,1]$ for which $L_p(1) + T_p(x) \leq 0$. We know that $L_p(1) = \sum_{n=1}^{\infty} \left(\frac{n}{p}\right) \frac{1}{n}$

is a positive function for all primes p, and so we have

$$\mu(p) = \int_{\substack{0<x<1 \\ x:\ S_p(x)\leq 0}} 1\,dx = \int_{\substack{0<x<1 \\ x:\ L_p(1)+T_p(x)\leq 0}} 1\,dx = \int_{\substack{0<x<1 \\ x:\ T_p(x)\leq 0, \\ |T_p(x)|>L_p(1)}} 1\,dx < \int_{\substack{0<x<1 \\ x:\ T_p(x)<0}} 1\,dx = \beta(p). \tag{5.5.13}$$

Thus, we have proved that there exists a positive proportion of primes $p \equiv 3 \bmod 4$ for which we can reduce our objective of finding an upper bound for $\mu(p)$, to finding an upper bound for $\beta(p)$. We will call this particular set of primes $\mathcal{P}''$. (It is clear that for $\mathcal{P}'$ as in Definition 5.5.5, $\mathcal{P}' \cap \mathcal{P}'' = \emptyset$). And so, by (5.5.13), for primes $p \in \mathcal{P}''$, we have $\mu(p) < \beta(p)$. Then by (5.5.11) and (5.5.12), we can deduce the following:

For a positive proportion of primes $p \equiv 3 \mod 4$,

$$\mu(p) < 0.715 \tag{5.5.14}$$

and of these primes, there is at least one for which

$$\mu(p) < 0.369. \tag{5.5.15}$$

Hence we have also proved the second assertion of Theorem 0.0.2.

Chapter 6

CONCLUSION

In Chapter 5 we mentioned that our results can be improved by using machines with higher RAM or even by employing distributed processing. Here we will outline the use of these ideas and show how certain changes to our algorithm can yield better results.

- In Remark A.2.2 we outline our choice of the termination condition. This clearly implies that if we had access to machines with greater memory, we would be able to perform more iterations and thus improve our results considerably so as to observe a convergence of values.
- The suggestion for processing our data on parallel units is in line both with the memory constraints of an alone station and also with the computation time of a single iteration. This suggestion can be put to use simply by distributing the vectors at each iteration between parallel units.
- We notice that in order to compute values of the series $-\sum_{\substack{n=1 \\ q|n \Rightarrow q \leq y}}^{\infty} \lambda(n)\frac{\cos 2\pi n x}{n}$, we need to truncate this series at some point. It is possible that if we truncate this series at a value higher than 1000 (as was our choice), we would obtain more accurate results. However, we note that this change will cost computation time.
- Similar to the previous suggestion, it is possible that we would obtain more accurate results if for our computations, the value of N (see Theorem 5.2.1) was taken to be larger than 1000 (as what we took). However, again, this larger value of N will cost considerable computation time.

- In line with the problem caused by memory constraints, we observe that storing the values of the vectors involves storing the integer values of 1 and -1. It is known that while an integer data type is of size 4 bytes, a boolean data type is of size only 1 byte. Hence, we could save on memory by storing the vectors as boolean data types, (*true* = 1 and *false* = -1), where the corresponding integer value could be considered only at the time of computation.

We invite readers to experiment with the above suggestions and in turn yield results that would prove our conjecture that H.L. Montgomery's results are in fact the best possible.

BIBLIOGRAPHY

[1] M. Abramowitz, I. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965.

[2] T. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, New York, 1976.

[3] A. Brauer, *Combinatorial methods in the distribution of k^{th} power residues*, Combinatorial Mathematics and Its Applications (R.C Bose and T.A Dowling, eds.), Univ. of North Carolina Press, Chapel Hill, (1969), 14-37.

[4] D.A. Burgess, *The distribution of quadratic residues and non-residues*, Mathematika **4**, (1957), 106-112.

[5] D.A. Burgess, *On character sums and L-series, I*, Proc. London Math Soc.(3) **123**, (1962), 193-206.

[6] D.A. Burgess, *On character sums and L-series, II*, Proc. London Math Soc.(3) **13**, (1963), 524-536.

[7] H. Davenport, *On the distribution of quadratic residues* (mod p), J. London Math Soc. **8** (1931), 49-54.

[8] H. Davenport, *The Higher Arithmetic*, Fifth edition, Cambridge University Press, London, New York, 1982, 74-76.

[9] H. Davenport, *Multiplicative Number Theory*, Third Edition, Springer-Verlag, New York, 2000.

[10] P.G.L. Dirichlet, *Lectures on Number Theory*, Amer. Math. Soc., Providence, RI, 1999.

[11] J.B Friedlander and H. Iwaniec, *Estimates for character sums*, Proc. of Amer. Math. Soc., **119** (1993), 365-372.

[12] J.B Friedlander and H. Iwaniec, *A note on character sums*, Contemp. Math J. **166** (1994), 295-299.

[13] S.W. Graham and C.J. Ringrose, *Lower Bounds for least quadratic residues*, Prog. Math **85** (1990), Birkhäuser, Boston, 269-309.

[14] A. Granville, *A Decomposition of Riemann's Zeta-Function*, London Mathematical Society Lecture Notes (Proceedings of the Kyoto Conference), vol 247, (1997), 95-101.

[15] A. Granville and K. Soundrarajan, *Large character sums*, Journal of the Amer. Math. Society, vol 14 (2001), 365-397.

[16] A. Granville and K. Soundrarajan, *Large character sums : Pretentious characters and the Pólya-Vinogradov Theorem*, (*to appear*).

[17] A. Hildebrand, *A note on Burgess' character sum estimate*, C.R. Acad. Sci. Roy. Soc. Canada **8** (1986), 35-37.

[18] A. Hildebrand, *Large values of character sums*, J. Number Theory **29** (1988), 271-296.

[19] A. Hildebrand, *On the constant in the Pólya Vinogradov Inequality* , Canad. Math. Bull. **31** (1988), 347-352.

[20] R.H. Hudson, *On the first occurrence of certain patterns of quadratic residues and non-residues* , Isreal J. Math. **44** (1983), 23-32.

[21] H. Iwaniec, E. Kowalski, *Analytic Number Theory*, American Mathematical Society, Providence, Rhode Island, 2004.

[22] E. Landau, *Abschätzungen von Charaktersummen, Einheiten und Klassenzahlen*, Göttinger Nachrichten (1918), 79-97.

[23] H.L. Montgomery, *Distribution questions concerning a character sum*, Colloquia Mathematica Societatis János Bolyai, (1974), 195-203.

[24] H.L. Montgomery, R.C Vaughan, *Exponential sums with multiplicative coefficients*, Inventiones math. **43**, (1977), 69-82.

[25] R.E.A.C Paley, *A theorem on characters*, J.London Math. Soc., **7**, (1932), 28-32.

[26] R. Peralta, *On the distribution of quadratic residues and nonresidues modulo a prime number*, Mathematics of Computation, Vol. 58, Number 197, (1992), 433-440.

[27] G. Pólya, *Über die Verteilung der quadratischen Reste und Nichtreste* , Göttinger Nachrichten, (1918), 21-29.

[28] C. Pomerance, R. Crandall, *Prime Numbers, A Computational Perspective*, Springer-Verlag, New York, 2001.

[29] J. B. Rosser, L. Schoenfeld, *Approximate Formulas for some Functions of Prime Numbers*, Illinois J. Math **6**, (1962), 64-94.

[30] I. Schur, *Einige Bemerkungen zur vorstehenden Arbeit des Herrn G. Pólya* , Göttinger Nachrichten, (1918), 30-36.

[31] Zhi-Hong. Sun, *Consecutive numbers with the same Legendre symbol*, Proceedings of the American Mathematical Society, Volume 130, Number 9, (2002), 2503-2507.

[32] I.M Vinogradov, *Über die Verteilung der quadritischen Reste und Nichtreste* , J. Soc. Phys. Math. Univ. Permi **2**,(1919), 1-14.

[33] I.M Vinogradov, *On a general theorem concerning the distribution of the residues and non-residue of powers*, Trans. American Math. Soc.,**29**, (1927), 209-217.

Appendix A

A.1. The Algorithms

The computations were performed based on the following algorithms. For a given y, Algorithm A.1.1 sets up an array using the given vector $V_{\lambda,y}$ (Definition 5.5.2) and the definition of $\lambda(n)$ (Definition 5.1.2). For the vector $V_{\lambda,y}$, given the array (as set up in Algorithm A.1.1), Algorithm A.1.2 calculates the corresponding value of the lower bound for $\beta(p)$, i.e. $L(V_{\lambda,y})$ (5.5.2), using Theorem 5.2.1. Finally, using the values of $L(V_{\lambda,y})$ for each vector $V_{\lambda,y}$ (obtained via Algorithm A.1.2), the iterative Algorithm A.1.3 determines a final lower bound for $\beta(p)$ (Definition 5.1.3). To determine a final upper bound for $\beta(p)$, we make minor changes to Algorithm A.1.2 and Algorithm A.1.3 and obtain Algorithms A.1.4 and A.1.5. Results obtained as a consequence of our computations based on the following algorithms are tabulated in Appendix B.

Algorithm A.1.1 (Setting up the array depending on the vector $V_{\lambda,y}$).
Given m, y such that y is the m^{th} prime; the vector $V_{\lambda,y} = (\lambda(q_1), \lambda(q_2), \ldots, \lambda(q_m))$ where q_i is the i^{th} prime, and using the definition of the totally multiplicative arithmetic function $\lambda(n)$, this algorithm will set up an array g of size M ($M \in \mathbb{N}$).

1. [Initializing an array for the vector $V_{\lambda,y}$]
   int L[m];
   for(int $i = 1$; $i \leq m$; $i++$) $L[i] = \lambda(q_i)$;
2. [Initializing the array g]
   int $g[M]$;
   for(int $i = 1$; $i \leq M$; $i++$) $g[i] = 1$;
3. [Setting $\lambda(n) = 0$ for all $n \leq M$ such that $P(n) > y = q_m$]

for(int $q = q_{m+1}$; $q \leq M$; $q++$)

if(isprime(q)) for(int $i = 1$; $i \leq M/q$; $i++$) $g[i*q] = 0$;

4. [Transforming the vector $V_{\lambda,y}$ of dimension m into an array of size y]

int l=0; int $V[y]$;

for(int $i = 1$; $i \leq y$; $i++$)

if(isprime(i)) {l=l+1; V[i] = l;}

5. [Completing the array g]

int $a = 1$, $q = 1$;

for(int $p = 2$; $p \leq y$; $p++$)

{

if(isprime(p))

{

$a = 1$; $q = p^a$;

while($q \leq M$)

{ for(int $i = 1$; $i \leq M/q; i++$)

$\{g[i*q] = g[i*q] * L[V[p]];\}$

$a = a + 1$; $q = p^a$; }

}

}

6. Return g;

Algorithm A.1.2 (Algorithm for Theorem 5.2.1).

Given c_y, the vector $V_{\lambda,y}$ and $N \in \mathbb{N}$, this algorithm first calculates $T_{\lambda,y}(\frac{j}{N})$ for each j, $(0 \leq j \leq N-1)$. Using these values, it then calculates a lower bound for $\beta(p)$ which is given by (5.2.4) and denoted by $L(V_{\lambda,y})$ (5.5.2). Note here that for the purpose of computation, the infinite sum over n in $T_{\lambda,y}(\frac{j}{N})$ is truncated such that $1 \leq n \leq M$ for $M \in \mathbb{N}$.

1. [Initializing]

int $r = 0$; float $S[N]$;

2. [Setting up the array for the vector $V_{\lambda,y}$]

int g[M]; //Via Algorithm A.1.1

3. [Determining the number of $j < N$ for which $T_{\lambda,y}(\frac{j}{N}) < 0$]

for (int $j = 0$; j<N ; j++)

{

float $t = 0$; float $x = j/N$;

for (int $n = 1$; $n \leq M$; $n++$)

{ $t = t + (g[n] * \cos(2\pi nx)/n)$; }

$S[j+1] = -t$;

if($S[j+1] < 0$) $r = r + 1$;

}

4. [Copying those $T_{\lambda,y}(\frac{j}{N})$ that are < 0 into a new array of size r]

float $S1[r]$, int $l = 0$;

for(int $i = 1$; $i \leq N$; $i++$)

if ($S[i] < 0$) $\{l = l + 1$; $S1[l] = S[i]$;$\}$

5. [Reordering those $T_{\lambda,y}(\frac{j}{N})$ that are < 0, in order of increasing magnitude]

float $t1$;

for(int $i = 1$; $i \leq r - 1$; $i++$)

{

for(int $j = i + 1$; $j \leq r$; $j++$)

{

if ($S1[i] < S1[j]$)

$\{t1 = S1[i]$; $S1[i] = S1[j]$; $S1[j] = t1$;$\}$

}

}

6. [Calculating the optimal k (Theorem 5.2.1)]

float $val = c_y * N$, int $k = 0$;

(a) for(int $i = 1$; $i \leq r$; $i++$) $\{S1[i] = S1[i]^2$;$\}$

(b) float $S2[r]$; $S2[1] = S1[1]$;

for(int $i = 2$; $i \leq r$; $i++$) $S2[i] = S1[i] + S2[i-1]$;

(c) float $P[r-1][2]$;

(d) for(int $i = 1$; $i \leq r - 1$; $i++$)

$\{P[i][1] = val - S2[i]$;

$P[i][2] = S2[i+1] - val$; $\}$

for(int $j = 1$; $j \leq r - 1$; $j++$)

if($P[j][1] > 0$ AND $P[j][2] \geq 0$) {k=j; break;}

7. [Calculating b, the value of $L(V_{\lambda,y})$ for the given $V_{\lambda,y}$]

float $b = 0$;

$b = (r - k)/N$;

return b;

Algorithm A.1.3 (Calculating the lower bound for $\beta(p)$). *This iterative algorithm returns a lower bound for* $\beta(p) = |\{x \in [0,1] : T_p(x) < 0\}|$, *where* $p \equiv 3 \mod 4$.

1. [Initializing]

int $m = 10$;

int $y = 29$; (the m^{th} prime)

float $MIN_y = 0$, $MAX_y = 0$;

float $omin = 0$, $omax = 0$;

int $imin = 1$, $imax = 1$;

float $MINVAL = 0$, $MAXVAL = 0$;

Set $\Lambda_{29} = \{V_{\lambda,29} \in \{1, -1\}^{10}\}$;

2. [Calculate c_y]

float $c_y = \frac{1}{2}\left(\frac{\pi^2}{6} - \prod_{\substack{q \text{ prime} \\ q \leq y}} \left(1 - \frac{1}{q^2}\right)^{-1}\right)$;

3. [Calculate $L(V_{\lambda,y})$ via Algorithm A.1.2]

If ($m == 10$)

for each $V_{\lambda,29} \in \Lambda_{29}$, calculate $L(V_{\lambda,29})$;

else

(a) if ($imin == 1$) for each $V'_{\lambda,y} \in \Lambda'_y$, calculate $L(V'_{\lambda,y})$;

(b) if ($imax == 1$) for each $V''_{\lambda,y} \in \Lambda''_y$, calculate $L(V''_{\lambda,y})$;

4. [Determine $\min_{V_{\lambda,y} \in \Lambda_y} (L(V_{\lambda,y}))$ and $\max_{V_{\lambda,y} \in \Lambda_y} (L(V_{\lambda,y}))$]

If ($m == 10$)

(a) $MIN_{29} = \min_{V_{\lambda,29} \in \Lambda_{29}} (L(V_{\lambda,29}))$;

(b) $MAX_{29} = \max_{V_{\lambda,29} \in \Lambda_{29}} (L(V_{\lambda,29}))$;

else

(a) if $(imin == 1)$ $MIN_y = \min_{V'_{\lambda,y} \in \Lambda'_y} (L(V'_{\lambda,y}))$;

(b) if $(imax == 1)$ $MAX_y = \max_{V''_{\lambda,y} \in \Lambda''_y} (L(V''_{\lambda,y}))$;

5. [Setting parameters for "pruning"]

float I_{max}, I_{min}, $B_{y,min}$, $B_{y,max}$;

//Remark A.2.1 explains the choice of I_{min} and I_{max}

(a) if $(imin == 1)$ $B_{y,min} = MIN_y + I_{min}$;

(b) if $(imax == 1)$ $B_{y,max} = MAX_y - I_{max}$;

6. [Choose the "qualifying" vectors i.e those that satisfy the "pruning" conditions]

If $(m == 10)$

(a) $W_L(B_{29,min}) = \{V_{\lambda,29} \in \Lambda_{29} : L(V_{\lambda,29}) < B_{29,min}\}$;

(b) $Z_L(B_{29,max}) = \{V_{\lambda,29} \in \Lambda_{29} : L(V_{\lambda,29}) > B_{29,max}\}$;

else

(a) if $(imin == 1)$

$W_L(B_{y,min}) = \{V'_{\lambda,y} \in \Lambda'_y : L(V'_{\lambda,y}) < B_{y,min}\}$;

(b) if $(imax == 1)$

$Z_L(B_{y,max}) = \{V''_{\lambda,y} \in \Lambda''_y : L(V''_{\lambda,y}) > B_{y,max}\}$;

7. [Termination condition]

//see Remark A.2.2 for more information about the termination condition.

if(m $>$ 10)

if $((imin == 1$ AND $|omin - MIN_y| < 10^{-3})$ OR

$(imin == 1$ AND $|W_L(B_{y,min})| > 208000))$

$imin = 0$; $MINVAL = MIN_y$

if $((imax == 1$ AND $|omax - MAX_y| < 10^{-3})$ OR

$(imax == 1$ AND $|Z_L(B_{y,max})| > 208000))$

$imax = 0$; $MAXVAL = MAX_y$

if $(imax == 0$ AND $imin == 0)$

return $MAXVAL$ and $MINVAL$;

8. [Set the new y]

$m = m + 5$;

int y_1 = the m^{th} prime;

9. [Setting up all vectors $V_{\lambda,y}$ for the next iteration]

   (a) if $(imin == 1)$

   $\Lambda'_{y_1} = \{(e_1|e_2) \in \Lambda_{y_1} : e_1 \in W_L(B_{y,min}), e_2 \in \{1,-1\}^5\}$;

   $V'_{\lambda,y_1} \in \Lambda'_{y_1}$;

   (b) if $(imax == 1)$

   $\Lambda''_{y_1} = \{(e_1|e_2) \in \Lambda_{y_1} : e_1 \in Z_L(B_{y,max}), e_2 \in \{1,-1\}^5\}$;

   $V''_{\lambda,y_1} \in \Lambda''_{y_1}$;

10. [Start next iteration]

    $omin = MIN_y$; $omax = MAX_y$;

    $y = y_1$;

    Go to step 2.

Algorithm A.1.4 (Algorithm for Theorem 5.2.3).

Given c_y, the vector $V_{\lambda,y}$ and $N \in \mathbb{N}$, this algorithm first calculates $T_{\lambda,y}(\frac{j}{N})$ for each j, $(0 \leq j \leq N-1)$. Using these values, it then calculates an upper bound for $\beta(p)$ which is given by (5.2.5) and denoted by $U(V_{\lambda,y})$ (5.5.10). Note here that for the purpose of computation, the infinite sum over n in $T_{\lambda,y}(\frac{j}{N})$ is truncated such that $1 \leq n \leq M$ for $M \in \mathbb{N}$. This algorithm follows the same steps as Algorithm A.1.2, with some changes in steps 3,4,5 and 7.

1,2. As in Algorithm A.1.2

3. [Determining the number of $j < N$ for which $T_{\lambda,y}(\frac{j}{N}) > 0$]

```
for ( int j = 0 ; j<N ; j++)
{
    float t = 0; float x = j * 1/N;
    for (int n = 1 ; n ≤ M ; n++)
        { t = t + (g[n] * cos(2πnx)/n); }
    S[j + 1] = −t;
    if(S[j + 1] > 0) r = r + 1;
}
```

4. [Copying those $T_{\lambda,y}(\frac{j}{N})$ that are > 0 into a new array of size r]

```
float S1[r], int l = 0;
for(int i = 1; i ≤ N; i++ )
```

if $(S[i] > 0)$ $\{l = l + 1;\ S1[l] = S[i];\}$

5. [Reordering those $T_{\lambda,y}(\frac{j}{N})$ that are > 0, in order of increasing magnitude]

float $t1$;

for(int $i = 1;\ i \leq r - 1;\ i++$)

{

for(int $j = i + 1;\ j \leq r;\ j++$)

{

if $(S1[i] > S1[j])$

$\{t1 = S1[i];\ S1[i] = S1[j];\ S1[j] = t1;\}$

}

}

6. As in Algorithm A.1.2

7. [Calculating b, the value of $U(V_{\lambda,y})$ for the given $V_{\lambda,y}$]

float $b = 0$;

$b = ((N - r) + k)/N$;

return b;

Algorithm A.1.5 (Calculating the upper bound for $\beta(p)$). *This iterative algorithm returns an upper bound for* $\beta(p) = |\{x \in [0,1] : T_p(x) < 0\}|$, *where* $p \equiv 3 \mod 4$.

1,2. As in Algorithm A.1.3

3. [Calculate $U(V_{\lambda,y})$ via Algorithm A.1.4]

If $(m == 10)$

for each $V_{\lambda,29} \in \Lambda_{29}$ calculate $U(V_{\lambda,29})$;

else

(a) if $(imax == 1)$ for each $V'_{\lambda,y} \in \Lambda'_y$, calculate $U(V'_{\lambda,y})$;

(b) if $(imin == 1)$ for each $V''_{\lambda,y} \in \Lambda''_y$, calculate $U(V''_{\lambda,y})$;

4. [Determine $\min\limits_{V_{\lambda,y} \in \Lambda_y}(U(V_{\lambda,y}))$ and $\max\limits_{V_{\lambda,y} \in \Lambda_y}(U(V_{\lambda,y}))$]

If $(m == 10)$

(a) $MAX_{29} = \max\limits_{V_{\lambda,29} \in \Lambda_{29}}(U(V_{\lambda,29}))$;

(b) $MIN_{29} = \min\limits_{V_{\lambda,29} \in \Lambda_{29}}(U(V_{\lambda,29}))$;

else

(a) if $(imax == 1)$ $MAX_y = \max_{V'_{\lambda,y} \in \Lambda'_y} (U(V'_{\lambda,y}))$;

(b) if $(imin == 1)$ $MIN_y = \min_{V''_{\lambda,y} \in \Lambda''_y} (U(V''_{\lambda,y}))$;

5. [Setting parameters for "pruning"]

float I_{max}, I_{min}, $B_{y,min}$, $B_{y,max}$;

//Remark A.2.1 explains the choice of I_{min} and I_{max}

(a) if $(imax == 1)$ $B_{y,max} = MAX_y - I_{max}$;

(b) if $(imin == 1)$ $B_{y,min} = MIN_y + I_{min}$;

6. [Choose the "qualifying" vectors i.e those that satisfy the "pruning" conditions]

If $(m == 10)$

(a) $W_U(B_{29,max}) = \{V_{\lambda,29} \in \Lambda_{29} : U(V_{\lambda,29}) > B_{29,max}\}$;

(b) $Z_U(B_{29,min}) = \{V_{\lambda,29} \in \Lambda_{29} : U(V_{\lambda,29}) < B_{29,min}\}$;

else

(a) if $(imax == 1)$

$W_U(B_{y,max}) = \{V'_{\lambda,y} \in \Lambda'_y : U(V'_{\lambda,y}) > B_{y,max}\}$;

(b) if $(imin == 1)$

$Z_U(B_{y,min}) = \{V''_{\lambda,y} \in \Lambda''_y : U(V''_{\lambda,y}) < B_{y,min}\}$;

7,8. As in Algorithm A.1.3

9. [Setting up all vectors $V_{\lambda,y}$ for the next iteration]

(a) if $(imax == 1)$

$\Lambda'_{y_1} = \{(e_1|e_2) \in \Lambda_{y_1} : e_1 \in W_U(B_{y,max}), e_2 \in \{1,-1\}^5\}$;

$V'_{\lambda,y_1} \in \Lambda'_{y_1}$;

(b) if $(imin == 1)$

$\Lambda''_{y_1} = \{(e_1|e_2) \in \Lambda_{y_1} : e_1 \in Z_U(B_{y,min}), e_2 \in \{1,-1\}^5\}$;

$V''_{\lambda,y_1} \in \Lambda''_{y_1}$;

10. As in Algorithm A.1.3

A.2. Remarks on the Algorithms

Here we will explain the choice of the parameters I_{min} and I_{max}, and of the termination condition as was used in the algorithms of the previous section.

Remark A.2.1 (Choice of I_{min} and I_{max}). I_{min} and I_{max} are parameters that determine the results of the pruning process. As is evident from steps 6 and 7 of Algorithm A.1.3, the choice of I_{min} and I_{max} determines the vectors that comprise the sets $W_L(B_{y,min})$ and $Z_L(B_{y,max})$ and in turn, the size of these sets. In our computations, an optimal choice for I_{min} and I_{max} was made at each iteration following some trial runs so as to ensure that this chosen value kept the number of vectors in $W_L(B_{y,min})$ and $Z_L(B_{y,max})$ to a minimum, without losing the best values for MIN_y and MAX_y obtained at the next iteration. Our choice of values for I_{min} and I_{max} ranged between 0.0015 and 0.004.

Remark A.2.2 (Termination Condition in Algorithms A.1.3 and A.1.5). The natural condition for terminating our computations would be when our values of MAX_y and MIN_y at subsequent iterations are seen to converge and stabilize at some value. The first condition of our *if* statement in step 7 of Algorithm A.1.3 employs this convergence criterion. However, during our computations we observed that before satisfying the convergence criterion, in certain cases we run into problems of space and time complexity. Due to these issues, we chose to stop further iterations when the number of "qualifying" vectors (step 6 of Algorithm A.1.3), i.e. the number of vectors to which we will append all possible $\{1, -1\}^5$ sequences for the new y, exceeded $208,000$ and hence the second condition of our *if* statement in step 7 of Algorithm A.1.3. Note that to set up the next iteration, each of these vectors that satisfy the "pruning condition", is used to create 2^5 new vectors (see step 9 of Algorithm A.1.3), and hence in effect, we stop iterating when either the convergence criterion is satisfied or when we find that number of vectors to be computed in the next iteration will exceed $6,656,000$ $(208,000 * 2^5)$. Note that we were working on machines with 2GB RAM; taking each vector to be of size at most 300 bytes, it is clear that storing $6,656,000$ vectors for a single iteration would very nearly exhaust the memory limit.

Appendix B

B.1. Tabulated Results of Computations

Take $c_y = \frac{1}{2}\left(\frac{\pi^2}{6} - \prod_{\substack{q \text{ prime} \\ q \leq y}} \left(1 - \frac{1}{q^2}\right)^{-1}\right)$. (Refer to Appendix A for details of the algorithms employed).

Applying Algorithm A.1.3, the following results were obtained:

y	c_y	$MIN_y = \min_{V'_{\lambda,y} \in \Lambda'_y} (L(V'_{\lambda,y}))$	$MAX_y = \max_{V''_{\lambda,y} \in \Lambda''_y} (L(V''_{\lambda,y}))$
29	0.0059317	0.205	0.560
47	0.0031830	0.218	0.599
71	0.0020899	0.222	0.620
97	0.0014944	0.226	0.634
113	0.0011313	0.232	0.642
149	0.0009093	0.235	0.647
173	0.0007522	0.239	0.651
197	0.0006356	0.241	0.654
229	0.0005483	0.244	0.656
257	0.0004789	0.246	0.657
281	0.0004234	0.247	0.659
313	0.0003780	0.249	–
349	0.0003414	0.251	–
379	0.0003107	0.252	–
409	0.0002844	0.254	–

Table B.1.1

Applying Algorithm A.1.5, the following results were obtained:

y	c_y	$MIN_y = \min_{V''_{\lambda,y} \in \Lambda''_y} (U(V''_{\lambda,y}))$	$MAX_y = \max_{V'_{\lambda,y} \in \Lambda'_y} (U(V'_{\lambda,y}))$
29	0.0059317	0.480	0.743
47	0.0031830	0.442	0.736
71	0.0020899	0.421	0.732
97	0.0014944	0.409	0.730
113	0.0011313	0.400	0.728
149	0.0009093	0.393	0.727
173	0.0007522	0.387	0.725
197	0.0006356	0.383	0.723
229	0.0005483	0.379	0.722
257	0.0004789	0.376	0.720
281	0.0004234	0.373	0.718
313	0.0003780	0.371	0.717
349	0.0003414	0.369	0.715

TABLE B.1.2