MULTIPLICATIVE MIMICRY AND IMPROVEMENTS OF THE PÓLYA-VINOGRADOV INEQUALITY

by

Leo I. Goldmakher

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
2009

Doctoral Committee:

Professor Jeffrey Lagarias, Co-Chair
Professor Kannan Soundararajan, Co-Chair
Professor Donald Lewis
Professor Michael Woodroofe
Assistant Professor Djordje Milicevic

I dedicate this thesis to the giants upon whose shoulders I stand: my parents. For everything that I have, for everything that I think is important, for everything that I am – I am incredibly grateful.

ACKNOWLEDGEMENTS

First, it is a pleasure to thank my advisor Soundararajan, who managed simultaneously to give me space to work independently and to push me when I needed it. A natural teacher, he shows his students how to think about a problem, rather than just the necessary techniques. He also created a fantastic working environment for analytic number theorists at Stanford. I am happy to thank also Sound's other students: Fai Chandee, Xiannan Li, and Bob Hough. Together we explored a lot of interesting mathematics, many great restaurants, and several cities.

Two people at Stanford who were particularly supportive are Ravi Vakil and Denis Trotabas. I am deeply grateful for their encouragement, which helped me through some trying times.

I owe many thanks to Jeff Lagarias at the University of Michigan for taking on the burden of mentoring someone else's student. In addition to making many valuable suggestions during the writing process, he has helped me get through several bureaucratic tight spots, and generally was a great moral support. Karen Rhea was very understanding of my unusual research circumstances and gave me the wonderful opportunity of working with her as a course co-coordinator of the introductory calculus course, from which I gained a whole new appreciation for her job (and learned a lot about drawing pictures in LaTeX). Karen Smith and Juha Heinonen were extremely supportive at a very confusing time for me early on in graduate school; were it not for their encouragement, I would almost certainly be doing something quite different now. I am very grateful for their support.

I was fortunate at both the University of Michigan and Stanford University to work with an extremely efficient and warm support staff. In particular, I would like to thank Tara McQueen and Karen Adkins for taking great care of me.

My interest in math was born in the Math Circle, a wonderful program which I was lucky to be a part of both as a student and as a teacher. It is no exaggeration to say that I would not be in math today were it not for the creators of this program, Bob and Ellen Kaplan – thank you!

There are many other people I am very grateful to. These include my parents Viktor and Nina, who always supported and loved me unconditionally; my older brother Greg for being everything one might hope from an older brother; Steven J. Miller for being a great mentor in my last two years of undergrad; Bill Duke for being very encouraging of my work; the members of StanShakes for giving me a wonderful creative outlet and social circle outside of the math department at Stanford; and my friends, among them Dave Anderson, Ilya Baran, Jonathan Bober, Ania Dabrowski, Kimberly Johnson, Catherine Kealhofer, Aaron Kofman, Katie Lu, Nathan Pflueger, Felipe Ramirez, John Saunders, Ben Schmidt, and Liz Vivas. To all my friends, thank you for your support and understanding during this difficult time in my life - your friendship means a lot to me.

TABLE OF CONTENTS

LIST OF APPENDICES

Appendix

NOTATION

1. **Abbreviations**

- RH refers to the Riemann Hypothesis, the conjecture that $\zeta(s)$ has no zeros in the region Re $s > \frac{1}{2}$.
- GRH refers to the Generalized Riemann Hypothesis, the conjecture that $L(s, \chi)$ has no zeros in the region Re $s > \frac{1}{2}$.

2. **Indices**

- $\displaystyle\sum_{n \le N}$ should always be interpreted as $\displaystyle\sum_{n=1}^{N}$.
- When p appears as an index (e.g. $\displaystyle\sum_{p \le y}$ or $\displaystyle\prod_{p}$) this means the sum or product runs over only the primes in the given range.

3. **Sets**

- $\mu_N := \left\{ e\left(\frac{n}{N}\right) \right\}_{n \le N}$ 'the complex N^{th} roots of unity'.
- $\mathcal{S}(y) := \{n \in \mathbb{N} : p \le y$ for every prime $p \mid n\}$ 'the y-smooth numbers'
- $\mathbb{U} := \{z \in \mathbb{C} : |z| \le 1\}$ 'the complex unit disc'
- $\mathcal{F} := \{f : \mathbb{Z} \to \mathbb{U}$ such that $f(mn) = f(m)f(n)$ for all $m, n\}$.
- supp(f) denotes the support of f, i.e. the closure of the set $\{x : f(x) \ne 0\}$.

4. **Functions**

- $\mu(n)$ is the Möbius function, which is defined:

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n = p_1 p_2 \cdots p_k \text{ where the } p_i \text{ are distinct primes} \\ 0 & \text{otherwise} \end{cases}$$

One particularly nice property we will use is

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

- $e(x) := e^{2\pi i x}$
- $[x]$ denotes the largest integer n satisfying $n \leq x$ (the floor function).
- $\{x\}$ is the fractional part of x: $\{x\} = x - [x]$.
- $d(n)$ denotes the number of integer divisors of n (the divisor function).
- $\mathrm{rad}(n)$ denotes the radical of n, which is defined $\mathrm{rad}(n) = \prod_{p|n} p$.
- $\mathcal{P}(n)$ denotes the largest prime factor of n.
- $S_\chi(t)$ is the character sum $\sum_{n \leq t} \chi(n)$, where χ is a Dirichlet character.
- f_y denotes the y-smoothed version of a function f, i.e.

$$f_y(n) := \begin{cases} f(n) & \text{if } n \in \mathcal{S}(y) \\ 0 & \text{otherwise.} \end{cases}$$

5. **Relations**

We often want to compare the behavior of a complicated function f to that of a simpler function g. We will use several notations to this end:

- $f = O(g)$ means that there exists some positive constant C such that $|f| \leq Cg$ at all arguments. Here, g is assumed to be a non-negative function, and the implicit constant C is absolute, i.e. independent not only of the argument, but also of any external parameters. If C is allowed to depend on some parameter k, we write it in the subscript: $f = O_k(g)$.
- $f \ll g$ means precisely the same thing as $f = O(g)$. As above, the implicit constant should be assumed to be independent of any parameters not appearing as a subscript to the $\ll$.
- $f \asymp g$ means that both $f \ll g$ and $g \ll f$. Note that the implicit constants may be different.

CHAPTER I

Introduction

This thesis is concerned with some properties of Dirichlet characters, classical objects which have been central to number theory research for over two centuries. Despite brilliant contributions by many celebrated mathematicians, much about their behavior remains conjectural at best. In this chapter we give a brief introduction to the subject, including an attempt to motivate the theorems presented in subsequent chapters. A more careful development of the theory can be found in Appendix A.

1.1 History, pre-20th century

The use of analysis to study the integers may be traced back to the 1730's. The young Swiss mathematician Leonhard Euler was being celebrated for solving the Basel problem, which had stymied all previous efforts by the best mathematicians of the day: determine a closed form expression for $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$. Euler gave several brilliant proofs that the series sums to $\frac{\pi^2}{6}$.[1] As was often the case with his work, his proofs are not rigorous by today's standards, because they predated the development of the necessary theory by more than a century – indeed, one of his proofs of this result anticipates the fundamental Weierstrass and Hadamard factorization theorems in complex analysis.

[1] Euler later popularized the notation π for the ratio of the circumference to the diameter; at the time this paper appeared there was no standard notation for the ratio, and Euler described it in words.

Inspired by this problem, Euler studied a more general series, which Riemann later called the zeta function:

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

This series is absolutely convergent and smooth for real $s > 1$, and thus can be studied by analytic methods. In addition to evaluating $\zeta(2n)$ for all positive integers n, Euler discovered the formula

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}.$$

By taking the natural logarithm of both sides and expanding into a power series the terms $\log(1-p^{-s})$ which appear on the right side, and then taking the limit of $\zeta(s)$ as $s \to 1^+$, Euler deduced that $\sum_p \frac{1}{p}$ diverges – a surprising result since the conventional wisdom of the time held that the harmonic series was the "smallest" infinity.[2] Euler had used continuous methods (taking the limit of a function) to study discrete objects (the primes), and analytic number theory was born.

Many years later (in 1785), Euler wrote a paper in which he asserts that both

$$\sum_{p\equiv 1 \pmod 4} \frac{1}{p} \quad \text{and} \quad \sum_{p\equiv -1 \pmod 4} \frac{1}{p} \tag{1.1}$$

diverge, and that moreover any such sum should diverge, citing $\sum_{p\equiv 1 \pmod{100}} \frac{1}{p}$ as an example. He then proceeds to use Leibniz's famous alternating series[3] for $\pi/4$ to

[2]Euler writes that $\sum_p \frac{1}{p} = \log\log(+\infty)$, indicating that he understood the asymptotic behavior of $\sum_{p\leq x} \frac{1}{p}$.

[3]The formula, along with the beginnings of calculus, had been discovered centuries earlier by Madhava and his followers at the Kerala school.

calculate to several decimals the sum

$$\sum_p \frac{\chi(p)}{p} \quad \text{where} \quad \chi(n) = \begin{cases} 1 & \text{for } n \equiv 1 \pmod 4 \\ -1 & \text{for } n \equiv -1 \pmod 4 \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

It is almost certain that Euler could prove that the sums (1.1) diverge – his earlier method combined with Leibniz's series proves that (1.2) converges, and the result immediately follows – but his second assertion, that the sum of $\frac{1}{p}$ over just those primes congruent to $1 \pmod{100}$ diverges, is a significantly harder question. This type of assertion was not resolved until the brilliant work of Dirichlet, who in the late 1830's published a proof of the following:

Theorem 1.1. *If* $(a,q)=1$, *then* $\displaystyle\sum_{p\equiv a \,(mod\ q)} \frac{1}{p}$ *is a divergent series.*

Inspired by Euler, Dirichlet looked at generalizations of $\zeta(s)$,

$$L(s,\chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}. \tag{1.3}$$

To follow Euler's proof we wish to rewrite this sum as a product over the primes; a straightforward way to accomplish this is to require that χ be completely multiplicative, i.e. that $\chi(ab) = \chi(a)\chi(b)$ for all integers a, b. This condition implies that for all values of s for which $L(s,\chi)$ converges, we have an 'Euler product'

$$L(s,\chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

After taking logarithms and expanding $\log\left(1 - \dfrac{\chi(p)}{p^s}\right)$ as a power series, we obtain what we hope is a main term $\displaystyle\sum_p \frac{\chi(p)}{p^s}$. The next step is to identify primes in some arithmetic progression, whose common difference is q, say. To distinguish between different arithmetic progressions, we assign a different weight to each progression $a \pmod q$ by requiring χ to be periodic with period q. Finally, since there are only finitely many primes in any arithmetic progression $a \pmod q$ with $(a,q) > 1$, we can ignore these progressions by requiring $\chi(n) = 0$ whenever $(n,q) > 1$. Any function $\chi : \mathbb{Z} \longrightarrow \mathbb{C}$ satisfying these three properties (complete multiplicativity, periodicity with period q, $\operatorname{supp}(\chi)$ a subset of the integers relatively prime to q) is called a Dirichlet character, and denoted $\chi \pmod q$; the corresponding function $L(s,\chi)$ is called a Dirichlet L-function. Although it is not immediately obvious from this definition that any non-trivial Dirichlet characters exist, we will shortly see that they are in plentiful supply and enjoy many interesting properties. Perhaps the oldest example is the Legendre symbol $\left(\frac{a}{p}\right)$, whose study was initiated by Euler and Legendre, and was brought to the forefront of number theory by Gauss; it is a Dirichlet character $\pmod p$.

Before discussing the Dirichlet characters further, we make a simple observation: that the third property of Dirichlet characters ($\chi(n) = 0$ whenever $(n,q) > 1$) is in some sense a notational convenience. We justify this statement briefly. Call any function $f : \mathbb{Z} \longrightarrow \mathbb{C}$ which is completely multiplicative and periodic with period Q a pseudocharacter $\pmod Q$. The trivial pseudocharacter is the function which outputs 1 for all inputs, and the zero pseudocharacter outputs 0 for all inputs. Note that if f is non-zero then $f(1) = 1$, and if f is non-trivial then $f(0) = 0$.

Proposition 1.2. *Any non-zero, non-trivial pseudocharacter $\chi \pmod Q$ is a Dirichlet character $\pmod q$ for some $q \mid Q$.*

Proof. Let d be the largest factor of Q such that $\chi(d) \neq 0$ (such a d exists since $\chi(1) = 1$). If $d = 1$, then χ is a Dirichlet character $\pmod Q$, and we conclude. Thus we can assume $d > 1$. For any integers k, r we have

$$\chi(d)\,\chi\left(\frac{Q}{d}k + r\right) = \chi\left(d\left(\frac{Q}{d}k + r\right)\right) = \chi(dr) = \chi(d)\,\chi(r).$$

Since $\chi(d) \neq 0$ we conclude that χ is a pseudocharacter $\pmod{Q/d}$, and Q/d is a proper divisor of Q. Iterating this argument, we see that it must terminate, since Q only has a finite number of proper divisors. $\square$

We next describe some of the nice properties satisfied by the Dirichlet characters. The reader should note that the following is a sketchy and unintuitive development of the theory, and is meant purely as an expedient approach to the main work of this thesis. For a more natural development, see Appendix A.

Because of the three defining properties of a Dirichlet character $\chi \pmod q$, one can restrict its range and view it as a homomorphism from $(\mathbb{Z}/q\mathbb{Z})^*$ to $\mathbb{C}$. We shall abuse notation and refer to this restricted map as χ, as well.[4] A first observation is that every element of $(\mathbb{Z}/q\mathbb{Z})^*$ is a $\varphi(q)^{\text{th}}$ root of unity, whence the range of the restricted χ is contained in $\mu_{\varphi(q)}$, the complex $\varphi(q)^{\text{th}}$ roots of unity. Thus for any $\chi \pmod q$ and any integer n we have $|\chi(n)| = 0$ or 1, and it is 1 iff $(n, q) = 1$. After a few definitions, we tabulate some of the nicer features of Dirichlet characters.

Definition 1.3.

1. The *trivial* or *principal* character modulo q, denoted $\chi_0 \pmod q$, is defined by $\chi_0(n) = 1$ for all n relatively prime to q.

2. A character $\psi \pmod d$ is said to *induce* $\chi \pmod q$ if $\chi = \psi \cdot \chi_0$, where χ_0 is the principal character $\pmod q$.

3. A non-principal $\chi \pmod q$ is *primitive* if the only character inducing it is itself.

Proposition 1.4. *(See Appendix A for proofs of these statements.)*

1. *Any Dirichlet character χ satisfies $\chi(-1) = +1$ or -1. In the former case χ is said to be even; in the latter, odd.*

[4] In many ways it is more natural, though also more abstract, to develop the theory in reverse, i.e. defining a Dirichlet character to be any homomorphism from $(\mathbb{Z}/q\mathbb{Z})^*$ to $\mathbb{C}^*$ and extending it to a function from $\mathbb{Z}$ to $\mathbb{C}$. This is the approach taken in Appendix A.

2. *If $\psi \pmod{d}$ induces $\chi \pmod{q}$ then $d \mid q$.*

3. *Every nonprincipal Dirichlet character is induced by a unique primitive character.*

4. *The set of characters $\pmod{q}$ forms an abelian group under multiplication, with $\chi_0 \pmod{q}$ as the identity and $\overline{\chi}$ (the complex conjugate) as the inverse of χ. The order of the group is $\varphi(q)$.*

5. *For any $\chi \pmod{q}$ and any integer $k \geq 1$,*

$$\frac{1}{\varphi(q)} \sum_{n \leq q} \chi(n)^k = \begin{cases} 1 & \text{if } ord(\chi) \mid k \\ 0 & \text{otherwise} \end{cases}$$

6. *For any integer n,*

$$\frac{1}{\varphi(q)} \sum_{\chi \pmod{q}} \chi(n) = \begin{cases} 1 & \text{if } n \equiv 1 \pmod{q} \\ 0 & \text{otherwise} \end{cases}$$

Properties 5 (with $k = 1$) and 6 are jointly known as the orthogonality relations, and are both very useful. From 5 we deduce that for any nonprincipal $\chi \pmod{q}$ the associated character sum $S_\chi(t) := \sum_{n \leq t} \chi(n)$ is periodic with period q; since $|\chi(n)| \leq 1$ for all n, we see that $|S_\chi(t)| \leq \min(t, q)$. This is called the trivial bound on $S_\chi(t)$, and will play an important role in bounding $L(s, \chi)$. From Property 6, one sees that $\frac{1}{\varphi(q)} \sum_{\chi \pmod{q}} \overline{\chi}(a) \cdot \chi$ is an indicator function for the arithmetic progression $a \pmod{q}$.

We are now in a position to sketch Dirichlet's argument. This will be a highly non-rigorous outline; for the details, see the lovely account given in [9]. Following

Euler's approach, we find that for $s > 1$ and $\chi \pmod q$

$$\log L(s,\chi) = \sum_p \frac{\chi(p)}{p^s} + O(1)$$

whence for any a relatively prime to q we find

$$\begin{aligned}
\frac{1}{\varphi(q)} \sum_{\chi \pmod q} \overline{\chi}(a) \log L(s,\chi) &= \frac{1}{\varphi(q)} \sum_{\chi \pmod q} \overline{\chi}(a) \left(\sum_p \frac{\chi(p)}{p^s} \right) + O(1) \\
&= \sum_p \frac{1}{p^s} \left(\frac{1}{\varphi(q)} \sum_{\chi \pmod q} \overline{\chi}(a)\chi(p) \right) + O(1) \\
&= \sum_{p \equiv a \pmod q} \frac{1}{p^s} + O(1).
\end{aligned}$$

On the other hand, from the Euler product it can be seen that

$$\frac{1}{\varphi(q)} \log L(s,\chi_0) = \frac{1}{\varphi(q)} \log \zeta(s) + O(1)$$

which implies that

$$\frac{1}{\varphi(q)} \sum_{\chi \pmod q} \overline{\chi}(a) \log L(s,\chi) = \frac{1}{\varphi(q)} \log \zeta(s) + \frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0 \pmod q} \overline{\chi}(a) \log L(s,\chi) + O(1).$$

Thus if we can prove that $\log L(s,\chi)$ is bounded for all nonprincipal characters $\pmod q$ as $s \to 1^+$, we would deduce that

$$\sum_{p \equiv a \pmod q} \frac{1}{p^s} = \frac{1}{\varphi(q)} \log \zeta(s) + O(1);$$

taking $s \to 1^+$ yields Dirichlet's theorem.

It remains to prove that for any nonprincipal $\chi \pmod q$, $L(s,\chi) \asymp 1$ (where all implicit constants are allowed to depend on q). To find an upper bound we use partial summation: for $s > 1$, we have

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \int_{1^-}^{\infty} \frac{1}{t^s}\, dS_\chi(t) \tag{1.4}$$
$$= \left.\frac{S_\chi(t)}{t^s}\right|_{1^-}^{\infty} + s\int_1^{\infty} \frac{S_\chi(t)}{t^{s+1}}\, dt$$
$$= s\int_1^{\infty} \frac{S_\chi(t)}{t^{s+1}}\, dt$$

Using the trivial estimate $|S_\chi(t)| \le q$ yields that $|L(s,\chi)| \le q$ for all $s > 1$. Moreover, $\left| s\int_1^{\infty} \frac{S_\chi(t)}{t^{s+1}}\, dt \right| \le q$ for all $s > 0$; thus this 'integral representation' of $L(s,\chi)$ gives an alternative definition of the function which has the advantage of a wider range of convergence. To prove Dirichlet's theorem it now suffices to prove that $|L(s,\chi)|$ is bounded away from 0 as $s \to 1^+$. This is a much more delicate problem, and the deepest aspect of Dirichlet's proof; to accomplish it, he developed his celebrated class number formula, a fundamental result in its own right. Although we now have significantly simpler proofs that $L(1,\chi) \ne 0$, Dirichlet's difficulties in surmounting this problem foreshadowed those to come in showing that $L(s,\chi) \ne 0$ for values of s just slightly smaller than 1; this remains a major open question.

The next major development after Dirichlet was in 1859, with the appearance of Riemann's eight-page masterpiece *Über die Anzahl der Primzahlen unter einer gegebenen Größe*, a work which would define the direction number theory would take through the present day. Where Euler and Dirichlet had considered the L-functions only for real arguments, Riemann realized that allowing s to be complex led to much more powerful consequences in the study of primes. After meromorphically continuing the ζ function to the entire complex plane with only a simple pole of residue 1 at $s = 1$, Riemann proved the functional equation for $\zeta(s)$ (a relation which shows that the function has a certain symmetry with respect to the vertical line Re $s = \frac{1}{2}$). He also sketches proofs that $\zeta(s)$ admits a product representation (a remarkable insight, predating Weierstrass' factorization theorem by 20 years), that the number of zeros of $\zeta(s)$ in the 'critical strip' $0 \leq \text{Re } s \leq 1$ with imaginary part between 0 and T should be asymptotic to $\frac{T}{2\pi} \log \frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T)$, and that the number of primes less than or equal to x, denoted $\pi(x)$, should be related to the logarithmic integral $\text{li}(x)$ via an explicit formula in terms of the zeros of $\zeta(s)$. Finally, Riemann conjectured what has since become known as the Riemann Hypothesis: that all of the zeros of $\zeta(s)$ in the critical strip must lie on the line $\text{Re } s = \frac{1}{2}$. With the exception of this last conjecture, all of the other statements were justified over the course of the next 40 years, principally through the efforts of Hurwitz, von Mangoldt, Hadamard, and de la Vallée Poussin.

Analogues of all of these results hold (or are expected to hold) for all Dirichlet L-functions; in particular, the Generalized Riemann Hypothesis is the conjecture that all zeros in the critical strip of any Dirichlet L-function must lie on the line $\text{Re } s = \frac{1}{2}$. One of the major distinctions between $\zeta(s)$ and $L(s, \chi)$ for $\chi \pmod q$ nonprincipal, is that the former has a pole at $s = 1$ while the latter is holomorphic there; this follows immediately from the integral representation we derived in (1.4). Thus while $\zeta(s)$ cannot have zeros in a neighborhood of $s = 1$, a Dirichlet L-function might. Disproving the existence of these 'Landau-Siegel' zeros remains one of the largest open problems in number theory. Of course, this is a much weaker conjecture than the Generalized Riemann Hypothesis (GRH), which is also believed to be true.

The above survey is woefully incomplete, but gives some picture of the rise of analytic number theory from a curious observation of Euler's through the proof of the prime number theorem in the last few years of the 19th century. As our discussion enters the developments of the 20th century, we narrow our gaze and focus on the main subject of this thesis – character sums.

1.2 The development of the theory of character sums

Recall from (1.4) that for a given nonprincipal Dirichlet character $\chi \pmod q$, we have an integral representation of the associated L-function in terms of the character sum $S_\chi(t)$:

$$L(s,\chi) = s \int_1^\infty \frac{S_\chi(t)}{t^{s+1}}\, dt.$$

This formula shows that $L(s,\chi)$ can be extended to a holomorphic function in the region Re $s > 0$; in fact, following Riemann, one can extend $L(s,\chi)$ analytically to the entire complex plane. Analogous to Riemann's work on the ζ function, for primitive characters $\chi \pmod q$ there is a 'functional equation' relating the value of $L(1-s,\chi)$ to $L(s,\overline{\chi})$ in an explicit way (originally proved by Hurwitz in 1882, and subsequently by de la Vallée Poussin in 1896 using a different method). Since $L(s,\chi)$ is well-behaved for Re $s > 1$ – in this region the function can be described by the Dirichlet series (1.3) – the functional equation guarantees tame behavior in the region Re $s < 0$ also, leaving the critical strip $0 \le$ Re $s \le 1$ as the only region of the complex plane in which the behavior of $L(s,\chi)$ is mysterious. (Actually, again by the symmetry provided by the functional equation, it suffices to study $L(s,\chi)$ in the region $\frac{1}{2} \le$ Re $s \le 1$.) This highlights the importance of the integral representation, which provides a fairly straightforward formula in this region in terms of $S_\chi(t)$. We are thus naturally led to investigating the character sum function $S_\chi(t)$.

We mentioned above that $S_\chi(q) = 0$ and that $\chi(n)$ is a $\varphi(q)$th root of unity for every n relatively prime to q, and deduced that $S_\chi(t)$ is periodic with period q and can be bounded 'trivially' by $\min(t,q)$. From this we can already obtain an upper bound on the magnitude of $L(1,\chi)$ which, up to a constant, remains the strongest known upper bound to date:

$$|L(1,\chi)| \le \left| \int_1^q \frac{S_\chi(t)}{t^2}\ dt \right| + \left| \int_q^\infty \frac{S_\chi(t)}{t^2}\ dt \right| \le \log q + 1.$$

One consequence of GRH (due to Littlewood) is that $L(1,\chi)$ is well approximated by a 'short' Euler product: a truncated version of the regular Euler product which runs over only those primes less than $\log^2 q$. From this it immediately follows that $L(1,\chi) \asymp \log\log q$, an indicator that perhaps the trivial bound on $S_\chi(t)$ can be improved. Indeed, on the assumption of GRH one can show that[5]

$$S_\chi(t) \ll_\epsilon \sqrt{t} \cdot q^\epsilon. \tag{1.5}$$

However, even unconditionally some serious improvements of the trivial estimate can be made.

Before describing these improvements, we briefly discuss the Gauss sum, a quantity which is closely related to $S_\chi(t)$ and which will appear in many of the calculations below. The Gauss sum associated to a character $\chi \pmod q$ is

$$\tau(\chi) := \sum_{n \le q} \chi(n)\, e\Big(\frac{n}{q}\Big),$$

where $e(x) := e^{2\pi i x}$. Observe that for all integers a coprime to q, we have the identity

$$\overline{\chi}(a)\tau(\chi) = \sum_{n \le q} \chi(n)\, e\Big(\frac{an}{q}\Big). \tag{1.6}$$

It can be shown (see Appendix A) that if χ is primitive, the above relation holds for *every* a, not just those coprime to q. Taking the absolute value of each side of (1.6),

[5]Here and throughout we will use Vinogradov's useful notation $f \ll g$ interchangeably with $f = O(g)$; the implicit constant should be assumed to depend only on subscripts appearing on the $\ll$ symbol.

squaring, summing over all $a \pmod q$, expanding and simplifying, one finds that for χ primitive, $|\tau(\chi)| = \sqrt{q}$. This fact is crucial for obtaining non-trivial bounds on $S_\chi(t)$.

In 1918 George Pólya [28], then a young mathematician at ETH Zürich, realized that since $S_\chi(qt)$ is a periodic function with period 1, one can expand it in a Fourier series, say

$$S_\chi(qt) \sim \sum_{n\in\mathbb{Z}} a_n\, e(nt).$$

A straightforward calculation shows that if $\chi \pmod q$ is primitive,

$$\begin{aligned} a_0 &= -\frac{1}{q}\sum_{k\le q} k\,\chi(k) \\ a_n &= \frac{\chi(-1)\,\tau(\chi)}{2\pi i}\cdot\frac{\overline{\chi}(n)}{n} \quad \text{for } n \neq 0. \end{aligned}$$

Since $S_\chi(qt)$ is of bounded variation, its Fourier series converges to the value of the function everywhere except at points of discontinuity. With some more work (see [24]), Pólya obtained the following quantitative form of the Fourier expansion, which holds for all t: for any N,

$$S_\chi(t) = \frac{\tau(\chi)}{2\pi i}\sum_{1\le|n|\le N}\frac{\overline{\chi}(n)}{n}\left(1 - e\left(-\frac{nt}{q}\right)\right) + O\left(1 + \frac{q\log q}{N}\right). \tag{1.7}$$

Taking N of size q and estimating trivially we find

$$\begin{aligned} |S_\chi(t)| &\ll \sqrt{q}\left|\sum_{1\le|n|\le q}\frac{\overline{\chi}(n)}{n}\left(1 - e\left(-\frac{nt}{q}\right)\right)\right| + \log q \\ &\ll \sqrt{q}\,\log q. \end{aligned} \tag{1.8}$$

Very shortly after Pólya's result became known, Issai Schur [29] gave a very short proof of the above inequality; however, Schur's proof yields only an upper bound for

$S_\chi(t)$, rather than the asymptotic (1.7) produced by Pólya.

Simultaneously and entirely independently, Ivan Matveevich Vinogradov[6], a young docent at the State University of Perm in eastern Russia, discovered both of these proofs, in reverse order [31], [32]. The upper bound is now named after the two original discoverers:

Theorem 1.5 (The Pólya-Vinogradov Inequality)**.**

$$\max_t |S_\chi(t)| \ll \sqrt{q}\,\log q \qquad \textit{for every nonprincipal } \chi \,(\textit{mod } q).$$

where the implied constant is absolute.

Note that the theorem is stated for all nonprincipal $\chi \,(\mathrm{mod}\ q)$, while the sketch of the proof we gave depended on the primitivity of χ. We briefly justify this extension of the result to imprimitive characters by using a trick described by Montgomery and Vaughan in [23]. Suppose $\psi \,(\mathrm{mod}\ r)$ is the primitive character inducing $\chi \,(\mathrm{mod}\ q)$. Then $r \mid q$, and

$$\begin{aligned} \sum_{n\le t}\chi(n) &= \sum_{\substack{n\le t\\ (n,q)=1}}\psi(n) = \sum_{\substack{n\le t\\ (n,q/r)=1}}\psi(n) \\ &= \sum_{n\le t}\psi(n)\sum_{d\mid(n,q/r)}\mu(d) \\ &= \sum_{d\mid\frac{q}{r}}\mu(d)\psi(d)\sum_{m\le t/d}\psi(m). \end{aligned}$$

Applying the Pólya-Vinogradov inequality to the primitive character ψ we see that

$$|S_\chi(t)| \le d\left(\frac{q}{r}\right)\sqrt{r}\,\log r \ll \sqrt{q}\,\log q\ ;$$

here $d(n)$ is the number of divisors of n, and is well-known to be $O_\epsilon(n^\epsilon)$.

[6]Not to be confused with A. I. Vinogradov of the (Barban-)Bombieri-Vinogradov theorem.

It should be noted that for t of size q, (1.8) is very strong, comparable to the general bound (1.5) implied by GRH. On the other hand, for sums shorter than $\sqrt{q}$, the Pólya-Vinogradov bound is weaker than the trivial estimate. For many applications it is important to estimate $S_\chi(t)$ non-trivially for t as small as possible, and in a series of fundamental papers [2]-[8] Burgess did precisely this, obtaining a non-trivial bound for all $t > q^{1/4+o(1)}$. Hildebrand [19] was able 'interpolate' this result, extending the range of t on which Burgess' estimate is non-trivial to all $t > q^{1/4-o(1)}$. In general this is the strongest such result, but in certain special cases improvements are known. In 1974 H. Iwaniec (building on work of P. X. Gallagher and A. G. Postnikov) proved an important theorem, one consequence of which is a non-trivial bound for $S_\chi(t)$ in the range $t > q^\epsilon$ for primitive $\chi \pmod q$ with q sufficiently powerful (i.e. the radical of q is small). This set the stage for another fantastic theorem, published in 1990 by S.W. Graham and C. J. Ringrose, bounding $S_\chi(t)$ in terms of t, q, and the largest prime factor of q; for numbers which are suitably smooth (i.e. all of whose prime factors are small), their bound gives a non-trivial estimate for $t > q^\epsilon$. We will describe both of these results in more detail in Chapter III. Finding estimates which are non-trivial in ranges shorter than those handled by Burgess' work remains an important problem.

For $t > q^{5/8+o(1)}$ Burgess' results are weaker than the Pólya-Vinogradov inequality, which remains the strongest known bound in this range for the general character sum. Moreover, it is quite close to being a best-possible bound over the whole range $t \leq q$: applying partial summation to $\tau(\chi)$ immediately produces the lower bound

$$\max_{t\leq q} |S_\chi(t)| \geq \frac{1}{2\pi}\sqrt{q}$$

for any primitive $\chi \pmod q$. Furthermore, in 1932 Paley constructed an infinite class $\mathbb{P}$ of quadratic characters such that $\max_{t\leq q} |S_\chi(t)| \gg \sqrt{q}\log\log q$ over all $\chi \pmod q$ in $\mathbb{P}$. Thus it was an important step when in 1977 Montgomery and Vaughan [23] proved, on the assumption of GRH, that

$$S_\chi(t) \ll \sqrt{q}\,\log\log q. \tag{1.9}$$

Actually, they prove something rather more general, which will play a key role in our work in Chapters II and III; we briefly discuss their results.

Recall that to deduce the Pólya-Vinogradov inequality from Pólya's Fourier expansion (1.7), we estimated the sum $\left|\sum_{1\le|n|\le q} \frac{\overline{\chi}(n)}{n}\left(1-e\left(-\frac{nt}{q}\right)\right)\right|$ trivially by $\log q$. Of course, one expects the actual magnitude of the sum to be a good deal smaller. More generally, let $\mathcal{F}$ denote the family of all completely multiplicative functions $f : \mathbb{Z} \to \mathbb{U}$, where $\mathbb{U} = \{|z| \le 1\}$ is the complex unit disc; for any $f \in \mathcal{F}$ consider

$$\sum_{1\le|n|\le N} \frac{f(n)}{n}\, e(n\alpha). \tag{1.10}$$

As before, one can estimate the sum trivially by $\log N$, but this is expected to be far from the truth unless $f(n) \approx 1$ almost everywhere and α is very small. Indeed, for the sum to accumulate there would have to exist some constant θ such that $f(n) \approx e(\theta)\, e(-n\alpha)$ for many small n – an unlikely scenario, since f is a completely multiplicative function while $e(-n\alpha)$ is an additive one. It has been surprisingly difficult to quantify this heuristic.

In [23] Montgomery and Vaughan succeeded in obtaining (unconditionally) non-trivial cancelation in the sum $\sum_{n\le N} f(n)\, e(n\alpha)$ for an important class of real numbers α which we will discuss after stating their theorem. One of the remarkable features of their bound is its uniformity with respect to $\mathcal{F}$. The statement below does not appear explicitly in their paper, but is easily obtained from their work by partial summation – this particular formulation is due to Granville and Soundararajan and appears as Lemma 4.2 in [15]:

Theorem 1.6 (Montgomery-Vaughan)**.** *Suppose* α *is any real number and* $\frac{b}{r}$ *is any reduced fraction satisfying* $\left|\alpha - \frac{b}{r}\right| \le \frac{1}{r^2}$. *Then for any* $R \in [2, r]$, *any* $N \ge rR$, *and every* $f \in \mathcal{F}$,

$$\sum_{rR \le n \le N} \frac{f(n)}{n} e(n\alpha) \ll \log\log N + \frac{(\log R)^{3/2}}{\sqrt{R}} \log N$$

where the implicit constant is absolute.

Note that taking large values of r in the Theorem 1.6 gives strong results, so the theorem is particularly useful for those α which are approximable by fractions with large denominator. Let us be a little more precise. Recall Dirichlet's theorem of Diophantine approximation: for any $M \ge 1$ and any real number α there exists a reduced fraction $\frac{b}{r}$ such that

$$\left|\alpha - \frac{b}{r}\right| \le \frac{1}{rM} \quad \text{and} \quad 1 \le r \le M. \tag{1.11}$$

If such an approximation exists with r large (where the meaning of 'large' depends on the context), then α is said to lie in a *minor arc*; otherwise (if all rational approximations satisfying (1.11) have small denominator), α lies in a *major arc*. In this language, Theorem 1.6 improves the trivial estimate on (1.10) for α in a minor arc. This part of [23] is unconditional; to conclude the proof of (1.9) Montgomery and Vaughan demonstrated that on the assumption of GRH, $\sum_{1 \le |n| \le q} \frac{\overline{\chi}(n)}{n} e(n\alpha)$ can be well-approximated by restricting the sum to run over only smooth numbers (i.e. numbers all of whose prime factors are small). This is closely related to Littlewood's result on the size of $L(1, \chi)$ mentioned earlier, and was further explored by Granville and Soundararajan in [14].

1.3 Recent advances: the work of Granville-Soundararajan

Inspired by Montgomery and Vaughan's paper, Granville and Soundararajan developed a method of quantifying to what extent the trivial estimate of (1.10) can be improved for α in a major arc. The first step towards this is the following identity, a precursor of which can be found in equation (29) of [23], and whose proof is elementary: for $b \neq 0$ and $(b,r)=1$,

$$\sum_{1\leq |n| \leq N} \frac{\overline{\chi}(n)}{n} e\left(\frac{b}{r}n\right) = \sum_{d|r} \frac{\overline{\chi}(d)}{d} \cdot \frac{1}{\varphi(r/d)} \sum_{\substack{\psi \left(\text{mod } \frac{r}{d}\right) \\ \psi(-1)=-\chi(-1)}} \tau\left(\overline{\psi}\right) \psi(b) \left(\sum_{a \leq N/d} \frac{\overline{\chi}\psi(a)}{a} \right).$$

From this relation it is clear that if r is small, then the only situation in which $\sum_{1\leq |n|\leq N} \frac{\overline{\chi}(n)}{n} e\left(\frac{b}{r}n\right)$ can have large magnitude is if there exists some ψ a character of small conductor and parity opposite to that of χ, such that $\left(\sum_{a\leq N/d} \frac{\overline{\chi}\psi(a)}{a}\right)$ is large. One way this can happen is if $\chi(a) \approx \psi(a)$ for many small values of a. Since the values of a character are determined by its values at the primes, it suffices to check to what extent χ and ψ agree at the small primes. To accomplish this, Granville and Soundararajan introduced a pseudometric on $\mathcal{F}$, defined for $f,g \in \mathcal{F}$ and $y > 1$ by

$$\mathbb{D}(f,g;y) := \left(\sum_{p\leq y} \frac{1 - \text{Re } f(p)\overline{g}(p)}{p} \right)^{1/2}.$$

We will discuss this pseudometric extensively in Chapters II and III. With this notion in hand, Granville and Soundararajan proved the following remarkable theorems (Theorems 2.1 and 2.4 of [15]):

Theorem 1.7. *Let $\chi \,(mod\; q)$ be a primitive character. Of all primitive characters with conductor below $(\log q)^{1/3}$ let $\xi \,(mod\; m)$ denote a character for which $\mathbb{D}(\chi, \xi; q)$*

is a minimum. Then

$$\max_t |S_\chi(t)| \ll \big(1 - \chi\xi(-1)\big) \frac{\sqrt{m}}{\varphi(m)} (\sqrt{q} \log q) \exp\Big(-\frac{1}{2}\mathbb{D}(\chi, \xi; q)^2\Big) + \sqrt{q} (\log q)^{6/7}.$$

Theorem 1.8. *Assume GRH, and let $\chi \,(mod\ q)$ be a primitive character. Of all primitive characters with conductor below $(\log\log q)^{1/3}$ let $\xi \,(mod\ m)$ denote a character for which $\mathbb{D}(\chi, \xi; \log q)$ is a minimum. Then*

$$\max_t |S_\chi(t)| \ll \big(1-\chi\xi(-1)\big) \frac{\sqrt{m}}{\varphi(m)} (\sqrt{q} \log\log q) \exp\Big(-\frac{1}{2}\mathbb{D}(\chi, \xi; \log q)^2\Big) + \sqrt{q} (\log\log q)^{6/7}.$$

Thus, Granville and Soundararajan simultaneously improve Pólya-Vinogradov (unconditionally) and Montgomery-Vaughan (on GRH), *unless* there exists a character of small conductor and opposite parity to χ, which is very close to χ with respect to the Granville-Soundararajan pseudometric.

From the methods developed in [15], Granville and Soundararajan are able to deduce many interesting consequences. In particular, by finding lower bounds on the distance between a character of odd order and one of even order and small conductor, they deduce an improvement of Pólya-Vinogradov for characters of odd order. Suppose $\chi \,(\mathrm{mod}\ q)$ is a primitive character of odd order g, and set $\delta_g := 1 - \frac{g}{\pi} \sin \frac{\pi}{g}$. Theorems 1 and 4 of [15] assert that $|S_\chi(t)| \ll_g \sqrt{q} (\log Q)^{1-\frac{\delta_g}{2}+o(1)}$, where Q is q or $\log q$ depending on whether GRH is being assumed, $o(1) \to 0$ as $q \to \infty$, and the implicit constant depends only upon g. In their paper the authors remark that on GRH, one can construct an infinite class of characters of a fixed odd order g, whose sums get as large (up to a constant) as $\sqrt{q} (\log\log q)^{1-\delta_g+o(1)}$. This leads them to make the following

Conjecture 1.9. *For any primitive Dirichlet character $\chi \,(mod\ q)$ of odd order g,*

$$S_\chi(t) \ll_g \sqrt{q} (\log\log q)^{1-\delta_g+o(1)}$$

where the implicit constant depends only on g and $o(1) \to 0$ as $q \to \infty$. Moreover, this bound is best-possible.

1.4 Results of this thesis

After stating Granville and Soundararajan's identity at the start of the section, we observed that if r is small, the only way that the left side of the identity, $\sum_{1\le|n|\le N} \frac{\overline{\chi}(n)}{n} e\left(\frac{b}{r} n\right)$, can have large magnitude is if there exists some character ψ of opposite parity and small conductor such that $\sum_{a\le N} \frac{\overline{\chi}\psi(a)}{a}$ has large magnitude. One way this can occur is for χ to mimic such a ψ at many small arguments, and Granville and Soundararajan's theorems bound $S_\chi(t)$ in terms of such mimicry. However, there is another scenario in which $\sum_{a\le N} \frac{\overline{\chi}\psi(a)}{a}$ has large magnitude: if $\chi(n)$ mimics the function $\psi(n) \cdot n^{it}$ for some real t. Moreover, this is essentially characterizes the situations in which the sum can be large, as was shown by Halász in his seminal work on mean values of multiplicative functions in the late 1960's and early 1970's.

In Chapter II, we refine Granville and Soundararajan's methods by introducing Halász's results into their arguments; this will allow us to give an unconditional proof of the following:

Theorem 2.1. *Given $\chi \pmod q$ a primitive character of odd order g. Then*

$$S_\chi(t) \ll_g \sqrt{q}\,(\log q)^{1-\delta_g+o(1)}$$

where $\delta_g = 1 - \frac{g}{\pi}\sin\frac{\pi}{g}$, the implicit constant depends only on g, and $o(1) \to 0$ as $q \to \infty$.

It is expected that the methods presented in Chapter II can be adapted to prove Conjecture 1.9 conditionally on GRH, but this project is not yet completed. Also in progress is an unconditional construction à la Paley of an infinite class of cubic characters whose sums get as large as $\sqrt{q}\,(\log\log q)^{1-\delta_g+o(1)}$.

For the improvements of Pólya-Vinogradov for odd-order characters, both [15] and [11] relied on the geometry of the range of Dirichlet characters to derive lower bounds on the mimicry metric $\mathbb{D}$. In Chapter III we take a different approach to studying the metric. We first find a relation between $\mathbb{D}(\chi, 1; y)$ and the value of $L(s,\chi)$ at s slightly larger than 1; the relation allows us to translate upper bounds on $L(s,\chi)$ into lower bounds on $\mathbb{D}(\chi, 1; y)$, which in turn (via the results of [15]) lead to improvements of the Pólya-Vinogradov inequality. In Chapter III we illustrate this method by improving bounds on $S_\chi(t)$ for χ a character of smooth or powerful conductor. In addition to the theorems of Granville and Soundararajan, we will require results of Iwaniec-Gallagher-Postnikov and Graham-Ringrose on characters to smooth and powerful moduli. We prove the following two theorems:

Theorem 1. *Given $\chi\,(mod\ q)$ primitive, with q squarefree. Then*

$$|S_\chi(x)| \ll \sqrt{q}\,(\log q)\left(\frac{(\log\log\log q)^2}{\log\log q} + \frac{(\log\log\log q)^2\log\mathcal{P}(q)}{\log q}\right)^{\frac{1}{4}}$$

where $\mathcal{P}(q)$ is the largest prime factor of q and the implied constant is absolute.

Theorem 2. *Given $\chi\,(mod\ q)$ a primitive Dirichlet character with q large and*

$$rad(q) \le \exp\left((\log q)^{3/4}\right),$$

where the radical of q is defined

$$rad(q) := \prod_{p|q} p.$$

Then

$$|S_\chi(x)| \ll_\epsilon \sqrt{q}\,(\log q)^{7/8+\epsilon}.$$

To illustrate these results, we give two concrete infinite classes of characters for which we can improve the Pólya-Vinogradov inequality.

Example 1.10. Let q_k denote the product of the first k primes. Consider the set $\mathcal{G}$ of all characters (mod q_k) ranging over all positive integers k. Then for all primitive $\chi\,(\text{mod } q)$ in $\mathcal{G}$,

$$S_\chi(t) \ll \frac{\sqrt{q}\,\log q}{(\log\log q)^{1/4-o(1)}}$$

Example 1.11. Consider the set $\mathcal{G}'$ of characters (mod p^k) ranging over all primes p and integers $k \geq (\log p)^{1/3}$. Then for all primitive $\chi\,(\text{mod } q)$ in $\mathcal{G}'$,

$$S_\chi(t) \ll_\epsilon \sqrt{q}\,(\log q)^{7/8+\epsilon}$$

CHAPTER II

Odd order character sums

In the introduction to this thesis we discussed a classical bound on character sums, proved independently by Pólya and Vinogradov in 1918. Their inequality remained the strongest known upper bound for $S_\chi(t)$ in the full range $t \le q$ for almost ninety years, until the work of Granville and Soundararajan [15] showed that, at least for some infinite classes of characters, it can be improved. For easy reference, we restate the Pólya-Vinogradov theorem here: for any non-principal character $\chi \pmod q$,

$$|S_\chi(t)| \ll \sqrt{q}\log q \tag{2.1}$$

where the implicit constant is absolute (i.e. independent of χ, q, and t).

The aim of this chapter is to refine the methods of Granville and Soundararajan in [15], which (loosely speaking) we will do by expanding the set of multiplicative functions considered from just the Dirichlet characters to all functions of the form $\chi(n)\cdot n^{it}$, where χ is a Dirichlet character and t is real. This is not a new idea. Indeed, it can be said to stem from the work of Riemann, which inspired number theorists to study functions of the form $\displaystyle\sum_n \frac{\chi(n)\cdot n^{it}}{n^\sigma}$ (where σ, t are real) rather than the restricted family $\displaystyle\sum_n \frac{\chi(n)}{n^\sigma}$ considered by Dirichlet. More recently, Balog, Granville, and Soundararajan [1] used a similar aproach in their study of the distribution of values of multiplicative functions. One of the concrete advantages of studying this wider class of arithmetic functions is the possibility of using powerful results of Halász on mean values of multiplicative functions, which will allow us to improve Granville and Soundararajan's bound on odd order character sums (Theorem 1 of [15]) by a power of $\log q$; we shall prove

Theorem 2.1. *Given $\chi \pmod q$ a primitive character of odd order g. Then*

$$S_\chi(t) \ll_g \sqrt{q}\,(\log q)^{1-\delta_g+o(1)}$$

where $\delta_g = 1 - \frac{g}{\pi}\sin\frac{\pi}{g}$, the implicit constant depends only on g, and $o(1) \to 0$ as $q \to \infty$.

For example for cubic characters $\chi \pmod q$ the theorem asserts that

$$S_\chi(t) \ll \sqrt{q}\,(\log q)^{\frac{3\sqrt{3}}{2\pi}+o(1)}.$$

Note that $\frac{3\sqrt{3}}{2\pi} \approx 0.827 < 5/6$.

It is expected that the methods presented here can be adapted to give an analogous improvement of Theorem 4 of [15], thus giving a conditional proof of Conjecture 1.9 on the assumption of GRH; this would be a best possible result (again on the assumption of GRH). However, this work is still in progress.

2.1 Introduction

For $\chi \pmod q$ a primitive character, Pólya noted [28] that $S_\chi(t)$ is periodic with period q, and thus can be expanded in a formal Fourier series whose coefficients are easily computed. With some more work (see [24]), he obtained a quantitative version of the formal expansion:

$$S_\chi(t) = \frac{\tau(\chi)}{2\pi i} \sum_{1\le |n| \le N} \frac{\overline{\chi}(n)}{n} \left(1 - e\Big(-\frac{nt}{q}\Big)\right) + O\left(1 + \frac{q\log q}{N}\right) \tag{2.2}$$

where N is any positive integer and $\tau(\chi) = \sum_{n\le q} \chi(n)\, e\left(\frac{n}{q}\right)$ is the Gauss sum. Since χ is primitive, $|\tau(\chi)| = \sqrt{q}$, so it remains to understand the behavior of sums of the form

$$\sum_{1\le |n| \le N} \frac{\overline{\chi}(n)}{n} e(n\alpha). \tag{2.3}$$

This sum is trivially $\ll \log N$ by the triangle inequality, and this proves sufficient to deduce the Pólya-Vinogradov inequality.

In this chapter we will obtain a non-trivial bound on (2.3) for primitive characters $\chi \pmod q$ of odd order, improving previous results of Granville and Soundararajan [15]. Before stating our bound, observe that for a character χ of odd order, $\chi(-1) = 1$ (recall that such a character is called *even*). This implies that

$$\sum_{1\le |n| \le N} \frac{\overline{\chi}(n)}{n} = 0$$

which, upon substitution into (2.2), shows that to improve the Pólya-Vinogradov inequality for odd order characters it suffices to non-trivially bound (2.3) for nonzero α. We will prove:

Theorem 2.2. *Given $\chi \pmod q$ a primitive character of odd order g. For any real $\alpha \neq 0$,*

$$\sum_{1\le |n| \le q} \frac{\chi(n)}{n} e(n\alpha) \ll_g (\log q)^{1-\delta_g + o(1)}$$

where the implicit constant depends only on g, and $o(1) \to 0$ as $q \to \infty$.

Applying this to estimate the sum in Pólya's Fourier expansion immediately yields Theorem 2.1.

The proof of Theorem 2.2 follows very closely the ideas in Granville and Soundararajan's original paper on the subject, [15]. At the heart of their work is their insight that a character sum $S_\chi(t)$ cannot attain large magnitudes unless χ mimics the behavior of a character ξ of opposite parity and very small conductor; this idea seems to have been originally enunciated by Hildebrand [20]. Granville and Soundararajan's real success in [15] was to make this precise, by formulating a notion of distance between multiplicative functions and determining an upper bound on $S_\chi(t)$ in terms of the distance from χ to the nearest character ξ among all those of opposite parity and small conductor. In some special cases (e.g. for characters of odd order) one can find lower bounds on the distance between χ and every Dirichlet character of opposite parity and small conductor, leading to concrete improvements of the Pólya-Vinogradov inequality. The main contribution in our work is to determine a bound on $S_\chi(t)$ in terms of the distance from χ to the nearest multiplicative function of the form $\psi(n)n^{it}$, where ψ is a Dirichlet character of opposite parity and small conductor and t is real. Using this larger family of multiplicative functions allows us to employ powerful theorems of Halász on mean values of multiplicative functions, and gives a more refined version of Granville and Soundararajan's bound on $S_\chi(t)$. As will be demonstrated below (in Theorem 2.14), one can find a lower bound on the distance from a character χ of odd order to the elements of our larger class of multiplicative functions, which is comparable to the lower bound obtained by Granville and Soundararajan. Using this lower bound in our refined bound on $S_\chi(t)$ yields an improvement of Theorem 1 of [15].

We briefly mention here that the functions $f(n) = n^{-it}$ have some properties highly reminiscent of Dirichlet characters. Clearly these functions are completely multiplicative and have magnitude 1; less obvious is that they satisfy an analogue of the Pólya-Vinogradov inequality: for all $t \geq 2$,

$$\sum_{n \leq N} n^{-it} \ll \frac{N}{t} + \sqrt{t}\,\log t.$$

On the other hand, these functions are not periodic, and are therefore not Dirichlet characters. Moreover, for any fixed t bounded away from 0, the values n^{it} are dense on the complex unit circle. By contrast, the values of any Dirichlet character $\chi(n)$ form a discrete subset of $\mathbb{C}$ (namely, $\mu_g \cup \{0\}$, where g is the order of χ and μ_g denotes the g^{th} roots of unity).

As discussed in the introduction to this thesis, we prove Theorem 2.2 by handling separately the cases when α lies in a major or minor arc. More concretely, set $M = e^{\sqrt{\log q}}$. Dirichlet's theorem on Diophantine approximation implies that there exists a reduced fraction $\frac{b}{r}$ such that

$$\left|\alpha - \frac{b}{r}\right| \leq \frac{1}{rM} \quad \text{and} \quad r \leq M. \tag{2.4}$$

Lemma 2.10 below (a variant of Lemma 6.2 of [15]) will allow us to replace α by its rational approximation in the exponential sum, at the cost of possibly shortening the range of summation slightly and adding a negligible error. In our context, it asserts the existence of an $N \leq q$ such that

$$\sum_{n \leq q} \frac{\overline{\chi}(n)}{n}\, e(n\alpha) = \sum_{n \leq N} \frac{\overline{\chi}(n)}{n}\, e\left(\frac{b}{r}\,n\right) + O(\log\log q). \tag{2.5}$$

Note that N depends on α, but the implicit constant in the error term is absolute.

Following [15] we consider two cases. If there exists an approximation satisfying (2.4) with $r > (\log q)^{2\delta_g}$ we say α belongs to a minor arc; if no such approximation exists, α is said to be in a major arc. Typically, the minor arcs require more technically demanding estimates; the problem considered here is curious in that the minor arc calculations were completed long before the major arc ones.

Corollary 2.8 (a slight generalization of Lemma 4.2 of [15], which in turn was built on the work of Montgomery and Vaughan [23]) will imply that

$$\sum_{n\leq q} \frac{\overline{\chi}(n)}{n} e(n\alpha) \ll \log r + \frac{1+(\log r)^{5/2}}{\sqrt{r}} \log q + \log\log q. \tag{2.6}$$

In particular, for α in a minor arc, we conclude that the left hand side is $\ll_g (\log q)^{1-\delta_g+o(1)}$, thus proving Theorem 2.2 for such α (since if α belongs to a minor arc, so does $-\alpha$, and all implicit constants are independent of α).

For those α belonging to a major arc the calculations are slightly more involved. Roughly speaking, the argument goes like this. In [15], Granville and Soundararajan discovered a simple but enlightening formula for the exponential sum (2.3) in terms of Gauss sums and characters of small modulus and opposite parity relative to χ. In our context (see Proposition 2.12) their identity says that for b, r positive coprime integers,

$$\sum_{1\leq |n|\leq N} \frac{\overline{\chi}(n)}{n} e\left(\frac{b}{r} n\right) = \sum_{d|r} \frac{\overline{\chi}(d)}{d} \cdot \frac{1}{\varphi\left(r/d\right)} \sum_{\substack{\psi \left(\mathrm{mod}\ \frac{r}{d}\right) \\ \psi(-1)=-\chi(-1)}} \tau\left(\overline{\psi}\right)\psi(b)\left(\sum_{a\leq N/d} \frac{\overline{\chi}\psi(a)}{a}\right). \tag{2.7}$$

For α in a major arc, r is small, whence everything on the right hand side of the above formula is small with the possible exception of the sum $\sum_{a\leq N/d} \frac{\overline{\chi}\psi(a)}{a}$. One expects a fair amount of cancellation in this sum; the largest the sum can be is if $\overline{\chi}\psi(a)$ points in the same direction for many small a – in other words, if χ mimics ψ. Thus, the exponential sum on the left side of this formula can be large only if χ mimics some character ψ of small conductor and opposite parity. A good measure of the extent to which one character mimics another was formally introduced by Granville and Soundararajan in [15]:

Definition 2.3. For any positive y and any two Dirichlet characters χ and ψ, let

$$\mathbb{D}(\chi, \psi; y) := \left(\sum_{p \leq y} \frac{1 - \text{Re } \overline{\chi}\psi(p)}{p} \right)^{1/2}.$$

It is easy to verify that $\mathbb{D}$ is symmetric, that $\mathbb{D} \geq 0$, and that $\mathbb{D}(\chi, \psi; y) = 0$ implies $\chi(p) = \psi(p)$ for all $p \leq y$. Less obvious but still true is that $\mathbb{D}$ satisfies the triangle inequality: $\mathbb{D}(\chi, \psi; y) + \mathbb{D}(\psi, \xi; y) \geq \mathbb{D}(\chi, \xi; y)$ for any three characters. It should be noted that all these properties hold not just for characters, but for arbitrary completely multiplicative functions from $\mathbb{Z}$ to $\mathbb{U}$, the complex unit disc. Since this family of functions will appear quite frequently, we give it a name:

Definition 2.4. $\mathcal{F}$ shall henceforth denote the set of all completely multiplicative functions $f : \mathbb{Z} \to \mathbb{U}$.

We note that the only obstacle to $\mathbb{D}$ being a true metric is that it is possible for the distance from χ to itself to be non-zero (if $\chi(p) = 0$ for some $p \leq y$). This is not much of an obstacle – one could simply restrict the sum in the definition of $\mathbb{D}$ to run over all primes $p \leq y$ at which $\chi(p)\psi(p) \neq 0$. In any event, $\mathbb{D}$ is pseudometric, and its definition can easily be tweaked to render it a true metric.

In [15], Granville and Soundararajan prove many interesting and useful properties of this mimicry measure. In particular, they prove two fundamental theorems relating the magnitude of a character sum to the distance between the character and that character of opposite parity and small conductor which the original most closely mimics. (See Theorems 2.1 and 2.2 of [15]; also, Theorems 2.4 and 2.5 for conditional versions.) Combining these theorems with upper and lower bounds on the distance between characters has led to improvements of the Pólya-Vinogradov inequality in some special cases: Theorems 1 and 4 of [15] improved Pólya-Vinogradov for primitive characters of odd order, and Theorems 1 and 2 of [11] (reproduced in Chapter III of this thesis) improved Pólya-Vinogradov for primitive characters to smooth or powerful conductor.

Even though they were the first to formalize the mimicry measure $\mathbb{D}$ and make an in-depth study of its properties, Granville and Soundararajan were not the first to phrase results in terms of this measure. Almost four decades previously, Gábor Halász [17] had realized that the existence of the mean value of any $f \in \mathcal{F}$ is closely related to whether f mimics the function $n^{i\beta}$ for some real β. Moreover, if the mean value of f tends to 0, Halász quantified how quickly it does so in terms of the distance from $f(n)$ to the nearest function of the form $n^{i\beta}$; here, distance means precisely Granville and Soundararajan's mimicry measure $\mathbb{D}$. In [22], Montgomery refined and simplified some of Halász's arguments, leading Tenenbaum to deduce the following elegant theorem (see page 343 of [30]):

Theorem 2.5 (Halász-Montgomery-Tenenbaum)**.** *Suppose $f \in \mathcal{F}$. Then for any $x \geq 3$ and $T \geq 1$,*

$$\frac{1}{x}\sum_{n\leq x} f(n) \ll (1+\mathcal{M})e^{-\mathcal{M}} + \frac{1}{\sqrt{T}}$$

where $\mathcal{M} := \min\limits_{|\beta|\leq T} \mathbb{D}\big(f(n), n^{i\beta}; x\big)^2$ and the implicit constant is absolute.

Applying partial summation to the right hand side of the Granville-Soundararajan identity (2.7), and subsequently Halász's theorem, one can establish an upper bound on the exponential sum in terms of lower bounds on the distance between χ and the closest function of the form $\psi(n)\, n^{i\beta}$, where ψ has small conductor and opposite parity to χ. In Theorem 2.14, we find a lower bound on the distance between such functions; this generalizes Lemma 3.2 of [15].

Finally we will conclude that for α in a major arc, if $\frac{b}{r}$ approximates α, then for any N

$$\sum_{1\leq|n|\leq N} \frac{\overline{\chi}(n)}{n}\, e\left(\frac{b}{r}\, n\right) \ll_g \frac{1}{\sqrt{r}}(\log N)^{1-\delta_g+o(1)} + (\log q)^{3\delta_g}$$

where $o(1) \to 0$ as $q \to \infty$. By Lemma 2.10 mentioned above, we have that for some

$N \le q$

$$\begin{aligned}\sum_{1\le|n|\le q} \frac{\overline{\chi}(n)}{n} e(n\alpha) &= \sum_{1\le|n|\le N} \frac{\overline{\chi}(n)}{n} e\left(\frac{b}{r} n\right) + O(\log\log q)\\ &\ll \frac{1}{\sqrt{r}} (\log N)^{1-\delta_g+o(1)} + (\log q)^{3\delta_g} + \log\log q\\ &\ll (\log q)^{1-\delta_g+o(1)}\end{aligned}$$

where the last inequality follows from the fact that $3\delta_g \le 1-\delta_g$ for all $g \ge 3$. (This is easily seen from a computation: for all $g \ge 3$, $\delta_g \le \delta_3 \approx 0.173$.)

Thus for any nonzero α, we obtain the corresponding non-trivial bound on the exponential sum (2.3), which immediately proves Theorem 2.1. See section 2.5 for a more precise overview of the whole argument.

2.2 Background and tools

In this section we begin by demonstrating that those α in minor arcs can be controlled via Granville and Soundararajan's adaptation of the Montgomery-Vaughan bound (2.6). After this we prove a lemma which will allow us to pass between α and its rational approximation. Finally, we state and prove the Granville-Soundararajan identity (2.7), which is the starting point of the major arc calculation.

2.2.1 The minor arc case

Lemma 2.6. *For $y \ge 10$ and $N \ge 2$,*

$$\sum_{p\le y} \frac{1}{p^{1-\frac{1}{\log N}}} = \log\log y + O\left(\frac{\exp\left(\frac{\log y}{\log N}\right)}{1+\frac{\log y}{\log N}}\right)$$

In particular, for $N \ge \exp\left(\dfrac{\log y}{\log_3 y}\right)$ we have $\displaystyle\sum_{p\le y} \frac{1}{p^{1-\frac{1}{\log N}}} = \Big(1+o(1)\Big)\log\log y$.

Proof. Write $\displaystyle\sum_{p\le t}\frac{\log p}{p}=\log t+R(t)$. It is a classical fact (see the proof of Mertens' inequality, for example) that $R(t)\ll 1$. Applying partial summation we have

$$\begin{aligned}\sum_{p\le y}\frac{1}{p^{1-\frac{1}{\log N}}} &= \int_{2^-}^{y}\frac{t^{\frac{1}{\log N}}}{\log t}\,d\sum_{p\le t}\frac{\log p}{p}\\ &= \int_{2}^{y}\frac{dt}{t^{1-\frac{1}{\log N}}\log t}+\frac{R(t)}{\log t}\,t^{\frac{1}{\log N}}\bigg|_{2^-}^{y}+\\ &\qquad +\int_{2}^{y}\frac{R(t)}{t^{1-\frac{1}{\log N}}\log^2 t}\,dt-\frac{1}{\log N}\int_{2}^{y}\frac{R(t)}{t^{1-\frac{1}{\log N}}\log t}\,dt\end{aligned}$$

We consider these terms individually. Making the substitution $t\mapsto\dfrac{\log t}{\log N}$ gives

$$\begin{aligned}\int_{2}^{y}\frac{dt}{t^{1-\frac{1}{\log N}}\log t} &= \int_{\frac{\log 2}{\log N}}^{\frac{\log y}{\log N}} e^t\,\frac{dt}{t}\\ &= \log\log y+\int_{\frac{\log 2}{\log N}}^{\frac{\log y}{\log N}}\left(e^t-1\right)\frac{dt}{t}+O(1)\end{aligned}$$

Note that

$$\int_{0}^{x}\left(e^t-1\right)\frac{dt}{t}\le\frac{4e^x}{1+x}$$

for all positive x (the inequality is easily verified for $0\le x\le 1$, and for $x>1$ the right side grows more quickly than the left). Thus we conclude that

$$\int_{2}^{y}\frac{dt}{t^{1-\frac{1}{\log N}}\log t}=\log\log y+O\left(\frac{e^{\frac{\log y}{\log N}}}{1+\frac{\log y}{\log N}}\right). \tag{2.8}$$

From this we immediately deduce

$$\frac{1}{\log N}\int_2^y \frac{R(t)}{t^{1-\frac{1}{\log N}}\log t}\,dt \ll \frac{\log\log y}{\log N}+\frac{1}{\log N}\cdot\frac{e^{\frac{\log y}{\log N}}}{1+\frac{\log y}{\log N}}.$$

Since

$$\left(1+\frac{\log y}{\log N}\right)\frac{\log\log y}{\log N}\le \frac{\log y}{\log N}+\frac{\log\log y}{\log y}\left(\frac{\log y}{\log N}\right)^2\le \exp\left(\frac{\log y}{\log N}\right)$$

for $y\ge 10$, we have $\dfrac{\log\log y}{\log N}\le \dfrac{e^{\frac{\log y}{\log N}}}{1+\frac{\log y}{\log N}}$, whence

$$\frac{1}{\log N}\int_2^y \frac{R(t)}{t^{1-\frac{1}{\log N}}\log t}\,dt \ll \frac{e^{\frac{\log y}{\log N}}}{1+\frac{\log y}{\log N}}. \tag{2.9}$$

Using the same change of variables as above and integrating by parts we find

$$\begin{aligned}
\int_2^y \frac{R(t)}{t^{1-\frac{1}{\log N}}\log^2 t}\,dt &\ll \frac{1}{\log N}\int_{\frac{\log 2}{\log N}}^{\frac{\log y}{\log N}} e^t\,\frac{dt}{t^2}\\
&= \frac{1}{\log N}\int_{\frac{\log 2}{\log N}}^{\frac{\log y}{\log N}} e^t\,\frac{dt}{t}+O\left(1+\frac{e^{\frac{\log y}{\log N}}}{\log y}\right)\\
&= \frac{\log\log y}{\log N}+O\left(\frac{e^{\frac{\log y}{\log N}}}{\log N+\log y}+1+\frac{e^{\frac{\log y}{\log N}}}{\log y}\right)\\
&\ll \frac{e^{\frac{\log y}{\log N}}}{1+\frac{\log y}{\log N}} \qquad (2.10)
\end{aligned}$$

from our work above and elementary estimates. Finally,

$$\left.\frac{R(t)}{\log t}\,t^{\frac{1}{\log N}}\right|_{2^-}^{y} \ll \frac{e^{\frac{\log y}{\log N}}}{1+\frac{\log y}{\log N}}. \tag{2.11}$$

Combining equations (2.8) – (2.11), we deduce the lemma. $\square$

Theorem 2.7 (Montgomery-Vaughan). *Suppose $f \in \mathcal{F}$ and $|\alpha - b/r| \leq \frac{1}{r^2}$ with $(b,r) = 1$. Then for every $R \in [2, r]$ and any $N \geq Rr$ we have*

$$\sum_{Rr \leq n \leq N} \frac{f(n)}{n} e(n\alpha) \ll \log\log N + \frac{(\log R)^{3/2}}{\sqrt{R}} \log N$$

where the implicit constant is absolute.

Proof. This follows immediately from Corollary 1 of [23] by partial summation; our formulation of this theorem is lifted from Lemma 4.2 of [15]. $\square$

Montgomery and Vaughan's proof of the above theorem required both ingenuity and hard analysis, as might be expected in a minor arc estimate. With their result in hand, we can deduce the following corollary (which is closely modeled on Lemma 6.1 of [15]) without much exertion. We will use the notation $\mathcal{S}(y)$ to denote the set of all y-smooth numbers, i.e. the set of all positive integers n with the property that none of the prime factors of n exceeds y.

Corollary 2.8. *Given any $f \in \mathcal{F}$ and any nonzero α and a reduced fraction $\frac{b}{r}$ such that $\left|\alpha - \frac{b}{r}\right| \leq \frac{1}{r^2}$. Then for $x \geq 2$ and $y \geq 10$,*

$$\sum_{\substack{n \leq x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) \ll \log r + \frac{1 + (\log r)^{5/2}}{\sqrt{r}} \log y + \log\log y$$

where the implicit constant is absolute.

Remark 2.9. Taking $f = \overline{\chi}$ and $x = y = q$ yields the version (2.6) found in the introduction.

Prior to proving this, we introduce one more piece of notation. Given $f : \mathbb{Z} \to \mathbb{C}$

and any positive number y, we define the y-smoothed function f_y:

$$f_y(n) = \begin{cases} f(n) & \text{if } n \in \mathcal{S}(y) \\ 0 & \text{otherwise.} \end{cases}$$

Note that if $f \in \mathcal{F}$, then $f_y \in \mathcal{F}$ as well.

Proof. The bound is trivially true for $x \le r^2$, so we assume $x > r^2$.

If $x \le y^{\log r}$ we apply Theorem 2.7 to f_y:

$$\begin{aligned} \sum_{\substack{n \le x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) &= \sum_{n \le x} \frac{f_y(n)}{n} e(n\alpha) \\ &= \sum_{n < r^2} \frac{f_y(n)}{n} e(n\alpha) + \sum_{r^2 \le n \le x} \frac{f_y(n)}{n} e(n\alpha) \\ &\ll \log r + \frac{(\log r)^{3/2}}{\sqrt{r}} \log x + \log\log x \\ &\ll \log r + \frac{(\log r)^{5/2}}{\sqrt{r}} \log y + \log\log y \end{aligned}$$

thus proving the claim in this case.

In the case that $x \ge y^{\log r}$ we use the above to deduce:

$$\begin{aligned} \sum_{\substack{n \le x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) &= \sum_{\substack{n \le y^{\log r} \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) + \sum_{\substack{y^{\log r} < n \le x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) \\ &\ll \log r + \frac{(\log r)^{5/2}}{\sqrt{r}} \log y + \log\log y + \sum_{\substack{y^{\log r} < n \le x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) \end{aligned} \tag{2.12}$$

It remains only to bound the sum on the right hand side of (2.12). Note that

$n > y^{\log r}$ if and only if $n > r \cdot n^{1-\frac{1}{\log y}}$. Therefore,

$$\begin{aligned}
\sum_{\substack{y^{\log r} < n \le x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) &\ll \frac{1}{r} \sum_{\substack{y^{\log r} < n \le x \\ n \in \mathcal{S}(y)}} \frac{1}{n^{1-\frac{1}{\log y}}} \\
&\le \frac{1}{r} \prod_{p \le y} \left(1 - \frac{1}{p^{1-\frac{1}{\log y}}}\right)^{-1}
\end{aligned}$$

We now apply Lemma 2.6 with $N = y$ to bound the product:

$$\begin{aligned}
\log \prod_{p \le y} \left(1 - \frac{1}{p^{1-\frac{1}{\log y}}}\right)^{-1} &= \sum_{p \le y} \frac{1}{p^{1-\frac{1}{\log y}}} + O(1) \\
&= \log\log y + O(1)
\end{aligned}$$

Thus,

$$\sum_{\substack{y^{\log r} < n \le x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) \ll \frac{1}{r} \log y.$$

Combining this with (2.12) we conclude. $\square$

2.2.2 Passing from α to its rational approximation

Lemma 2.10. *Given $f \in \mathcal{F}$. Suppose y and M are large, α is a given real number, and the reduced fraction $\frac{b}{r}$ with $r \le M$ is a Dirichlet approximation to α (i.e. $\left|\alpha - \frac{b}{r}\right| \le \frac{1}{rM}$). Set $N = \min\left\{x, \frac{1}{|r\alpha - b|}\right\}$. Then for all $R \in \left[2, \frac{N}{2}\right]$,*

$$\sum_{\substack{n \le x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) = \sum_{\substack{n \le N \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e\left(\frac{b}{r} n\right) + O\left(1 + \log R + \frac{(\log R)^{3/2}}{\sqrt{R}} (\log y)^2 + \log\log y\right)$$

where the implied constant in the error term is absolute.

Moreover, if $M \ge 2(\log y)^4 \log\log y$, the error term above can be replaced by $O(\log\log y)$.

Remark 2.11. As above, taking $f = \overline{\chi}$, $x = y = q$, and $M = e^{\sqrt{\log q}}$ yields the version given in (2.5) in the introduction.

Proof. If $N = x$ then $\left|\alpha - \frac{b}{r}\right| \leq \frac{1}{rx}$ whence

$$\sum_{\substack{n \leq x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n}\left(e(n\alpha) - e\left(\frac{b}{r}n\right)\right) \ll \sum_{\substack{n \leq x \\ n \in \mathcal{S}(y)}} \frac{1}{n} \cdot n \left|\alpha - \frac{b}{r}\right| \ll 1$$

We therefore assume that $N = \frac{1}{|r\alpha - b|} < x$. Note that this immediately implies that $N \geq M$ and that

$$\left|\alpha - \frac{b}{r}\right| = \frac{1}{rN}.$$

By Dirichlet's theorem, there is a reduced fraction $\frac{b_1}{r_1}$ with $r_1 \leq 2N$ such that

$$\left|\alpha - \frac{b_1}{r_1}\right| \leq \frac{1}{2r_1 N}$$

Note that $\frac{b}{r} \neq \frac{b_1}{r_1}$, since $\left|\alpha - \frac{b_1}{r_1}\right| < \frac{1}{r_1 N}$. Thus,

$$\frac{1}{rr_1} \leq \left|\frac{b}{r} - \frac{b_1}{r_1}\right| \leq \frac{1}{2r_1 N} + \frac{1}{rN}$$

whence $r_1 \geq N - \frac{r}{2}$. Since $r \leq M \leq N$, we see that

$$\frac{N}{2} \leq r_1 \leq 2N$$

so we can trivially bound the (possibly empty) sum

$$\sum_{\substack{N < n \leq Rr_1 \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n} e(n\alpha) \ll \log \frac{Rr_1}{N} = \log R + O(1).$$

Once again applying Montgomery-Vaughan's Theorem 2.7 to f_y (which we can do since $R \leq \frac{N}{2} \leq r_1$) we see that

$$\sum_{\substack{Rr_1 < n \leq e^{(\log y)^2} \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n}\, e(n\alpha) \quad = \quad \sum_{Rr_1 < n \leq e^{(\log y)^2}} \frac{f_y(n)}{n}\, e(n\alpha)$$

$$\ll \quad \log\log y + \frac{(\log R)^{3/2}}{\sqrt{R}}\, (\log y)^2$$

Finally, similarly to the proof of the lemma above, we see that

$$\sum_{\substack{e^{(\log y)^2} < n \leq x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n}\, e(n\alpha) \ll \sum_{\substack{e^{(\log y)^2} < n \leq x \\ n \in \mathcal{S}(y)}} \frac{1}{n} \ll \frac{1}{y} \sum_{n \in \mathcal{S}(y)} \frac{1}{n^{1-\frac{1}{\log y}}} \ll 1.$$

Combining these three bounds, we deduce

$$\sum_{\substack{n \leq x \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n}\, e(n\alpha) = \sum_{\substack{n \leq N \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n}\, e(n\alpha) + O\left(1 + \log R + \frac{(\log R)^{3/2}}{\sqrt{R}}\, (\log y)^2 + \log\log y\right).$$

Just as at the start of the proof, we have

$$\sum_{\substack{n \leq N \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n}\, e(n\alpha) = \sum_{\substack{n \leq N \\ n \in \mathcal{S}(y)}} \frac{f(n)}{n}\, e\left(\frac{b}{r}\, n\right) + O(1)$$

and we conclude the proof of the first part of the theorem.

For the second claim, if $M \geq 2(\log y)^4 \log\log y$, then

$$r_1 \geq N - \frac{r}{2} \geq M - \frac{M}{2} \geq (\log y)^4 \log\log y.$$

Taking $R = (\log y)^4 \log\log y$ renders the error $O(\log\log y)$. $\square$

2.2.3 The Granville-Soundararajan Identity

For α in a major arc (i.e. it is well-approximated by $\frac{b}{r}$ with $r \leq (\log q)^{2\delta_g}$), the situation is more complicated. The following identity, due to Granville and Soundararajan, is elementary but indicates an approach to the major arcs.

Proposition 2.12. *Given a primitive Dirichlet character $\chi \,(mod\ q)$ and positive coprime integers b and r,*

$$\sum_{\substack{1\leq|n|\leq N \\ n\in\mathcal{S}(y)}} \frac{\overline{\chi}(n)}{n}\, e\left(\frac{b}{r}\, n\right) = \sum_{\substack{d|r \\ d\in\mathcal{S}(y)}} \frac{\overline{\chi}(d)}{d} \cdot \frac{1}{\varphi\,(r/d)} \sum_{\substack{\psi\,\left(mod\ \frac{r}{d}\right) \\ \psi(-1)=-\chi(-1)}} \tau\left(\overline{\psi}\right)\psi(b) \left(\sum_{\substack{a\leq N/d \\ a\in\mathcal{S}(y)}} \frac{\overline{\chi}\psi(a)}{a} \right).$$

Thus for small r, $\displaystyle\sum_{\substack{1\leq|n|\leq N \\ n\in\mathcal{S}(y)}} \frac{\overline{\chi}(n)}{n}\, e\left(\frac{b}{r}\, n\right)$ can be large only if $\displaystyle\sum_{\substack{a\leq N/d \\ a\in\mathcal{S}(y)}} \frac{\overline{\chi}\psi(a)}{a}$ is large for some ψ of small conductor and opposite parity to χ. But this sum can only be large if $\chi(a) = \psi(a)$ for many small (and y-smooth) a, i.e. if χ mimics ψ. One of Granville and Soundararajan's major contributions in [15] was to write down a pseudometric on the space of multiplicative functions, which is a useful tool in studying multiplicative mimicry. We will recall the definition and basic properties of this pseudometric after proving the proposition.

Proof of Proposition 2.12. First, observe that

$$\sum_{\substack{1\leq|n|\leq N \\ n\in\mathcal{S}(y)}} \frac{\overline{\chi}(n)}{n}\, e\left(\frac{b}{r}\, n\right) = \sum_{\substack{n\leq N \\ n\in\mathcal{S}(y)}} \frac{\overline{\chi}(n)}{n}\, e\left(\frac{b}{r}\, n\right) - \chi(-1) \sum_{\substack{n\leq N \\ n\in\mathcal{S}(y)}} \frac{\overline{\chi}(n)}{n}\, e\left(\frac{-b}{r}\, n\right). \tag{2.13}$$

We examine the first sum on the right hand side. Summing over all possible greatest common divisors d of n and r, and setting $a = n/d$ we find

$$\sum_{\substack{n\le N\\ n\in\mathcal{S}(y)}} \frac{\overline{\chi}(n)}{n}\, e\!\left(\frac{b}{r}\,n\right) = \sum_{\substack{d\mid r\\ d\in\mathcal{S}(y)}} \frac{\overline{\chi}(d)}{d} \sum_{\substack{a\le \frac{N}{d}\\ \left(a,\frac{r}{d}\right)=1\\ a\in\mathcal{S}(y)}} \frac{\overline{\chi}(a)}{a}\, e\!\left(\frac{ab}{r/d}\right). \tag{2.14}$$

Now,

$$e\!\left(\frac{ab}{r/d}\right) = \sum_{k \left(\mathrm{mod}\ \frac{r}{d}\right)} e\!\left(\frac{k}{r/d}\right)\, \delta_{ab}(k)$$

where δ_x is the indicator function of x. By orthogonality of characters, we can express the indicator function in terms of characters:

$$\delta_{ab}(k) = \frac{1}{\varphi\left(\frac{r}{d}\right)} \sum_{\psi \left(\mathrm{mod}\ \frac{r}{d}\right)} \overline{\psi}(k)\, \psi(ab)$$

whence, switching the order of summation,

$$e\!\left(\frac{ab}{r/d}\right) = \frac{1}{\varphi\left(\frac{r}{d}\right)} \sum_{\psi \left(\mathrm{mod}\ \frac{r}{d}\right)} \tau\!\left(\overline{\psi}\right)\, \psi(ab).$$

Plugging this back into (2.14) and once again switching order of summation yields

$$\sum_{\substack{n\le N\\ n\in\mathcal{S}(y)}} \frac{\overline{\chi}(n)}{n}\, e\!\left(\frac{b}{r}\,n\right) = \sum_{\substack{d\mid r\\ d\in\mathcal{S}(y)}} \frac{\overline{\chi}(d)}{d} \cdot \frac{1}{\varphi\left(r/d\right)} \sum_{\psi \left(\mathrm{mod}\ \frac{r}{d}\right)} \tau\!\left(\overline{\psi}\right) \psi(b) \left(\sum_{\substack{a\le N/d\\ a\in\mathcal{S}(y)}} \frac{\overline{\chi}\psi(a)}{a} \right).$$

Using this identity to evaluate the second sum on the right hand side of (2.13) (with b replaced by $-b$) and combining the two sums, we conclude. $\square$

2.3 A Lower Bound on $\mathbb{D}(\chi(n), \xi(n)\, n^{i\beta}; y)$

We begin by making a concrete choice of parameters in Halász's Theorem (Theorem 2.5).

Theorem 2.13. *For all $x \geq 3$, for all $f \in \mathcal{F}$,*

$$\frac{1}{x}\sum_{n\leq x} f(n) \ll \big(\log\log x\big)\, e^{-\mathcal{M}(x,f)} + \frac{1}{\log x}$$

where

$$\mathcal{M}(x,f) = \min_{|\beta|\leq \log^2 x} \mathbb{D}\big(f(n), n^{i\beta}; x\big)^2$$

and the implicit constant is absolute.

In Lemma 3.2 of [15], Granville and Soundararajan proved that for any primitive character $\chi \pmod q$ of odd order g, and any primitive character ξ of opposite parity and conductor smaller than a power of $\log y$,

$$\mathbb{D}(\chi, \xi; y)^2 \geq \big(\delta_g + o(1)\big) \log\log y. \tag{2.15}$$

Our goal in this section is to prove a generalization of this:

Theorem 2.14. *Given $y \geq 2$, $\chi \pmod q$ a primitive character of odd order g, and any odd character $\xi \pmod m$ with $m < (\log y)^A$. Then for all β satisfying $|\beta| \leq \log^2 y$,*

$$\mathbb{D}\big(\chi(n), \xi(n) n^{i\beta}; y\big)^2 \geq \big(\delta_g + o(1)\big) \log\log y$$

where $o(1) \to 0$ as $y \to \infty$ for any fixed values of g and A.

Our plan of attack is as follows. We partition the interval $[2, y]$ into many small intervals of the form $(x, (1+\delta)x]$, where δ is of size $(\log y)^{-3}$. For p in such an interval, we approximate $p^{-i\beta}$ by $x^{-i\beta}$. This reduces our problem to estimating sums of the form

$$\sum_{\ell \,(\mathrm{mod}\ k)} \sum_{\substack{x<p\le(1+\delta)x \\ \xi(p)=e\left(\frac{\ell}{k}\right)}} \frac{1}{p}\left(1-\mathrm{Re}\ \chi(p)\, e\!\left(-\frac{\ell}{k}\right) x^{-i\beta}\right).$$

Following Granville and Soundararajan's proof of (2.15), we ignore the arithmetic properties of χ and view it simply as a g^{th} root of unity. Thus, it suffices to find a lower bound on

$$\sum_{\ell \,(\mathrm{mod}\ k)} \sum_{\substack{x<p\le(1+\delta)x \\ \xi(p)=e\left(\frac{\ell}{k}\right)}} \frac{1}{p} \min_{z\in\mu_g\cup\{0\}} \left(1-\mathrm{Re}\ z\, e\!\left(-\frac{\ell}{k}\right) x^{-i\beta}\right).$$

Using Siegel-Walfisz, we will easily deduce a lower bound of

$$\sum_{\substack{x<p\le(1+\delta)x \\ \xi(p)=e\left(\frac{\ell}{k}\right)}} \frac{1}{p}$$

which is independent of ℓ; the sum

$$\sum_{\ell \,(\mathrm{mod}\ k)} \min_{z\in\mu_g\cup\{0\}} \left(1-\mathrm{Re}\ z\, e\!\left(-\frac{\ell}{k}\right) x^{-i\beta}\right)$$

can then be evaluated by arguments inspired by those of [15]. Summing over all the small intervals will yield the desired lower bound in Theorem 2.14.

Before proceeding to the calculation, we observe that we may assume that β is not too small. Indeed, for those β which are $o\left(\frac{\log\log y}{\log y}\right)$ we have

$$\begin{aligned}
\mathbb{D}\left(\chi(n), \xi(n)n^{i\beta}; y\right)^2 &= \sum_{p\leq y} \frac{1}{p}\left(1 - \operatorname{Re}\, \chi\overline{\xi}(p)e^{-i\beta\log p}\right) \\
&= \sum_{p\leq y} \frac{1}{p}\left(1 - \operatorname{Re}\, \chi\overline{\xi}(p)\left(1 + O\left(|\beta|\log p\right)\right)\right) \\
&= \mathbb{D}\left(\chi, \xi; y\right)^2 + O\left(|\beta| \sum_{p\leq y} \frac{\log p}{p}\right) \\
&= \mathbb{D}\left(\chi, \xi; y\right)^2 + o\left(\log\log y\right)
\end{aligned}$$

and thus for such β, the bound of Theorem 2.14 follows from (2.15).

2.3.1 The contribution from short intervals

Our first goal is to obtain a lower bound on the sum over a short interval

$$\sum_{x<p\leq(1+\delta)x} \frac{1}{p}\left(1 - \operatorname{Re}\, \chi\overline{\xi}(p)\, p^{-i\beta}\right) \tag{2.16}$$

where

$$\delta \asymp \frac{1}{\log^3 y}\,.$$

Note that for any prime $p \in \big(x, (1+\delta)x\big]$, we may approximate $p^{i\beta}$ by $x^{i\beta}$: we have $0 \leq \log p - \log x \leq \log\big((1+\delta)x\big) - \log x = \log(1+\delta) \leq \delta$, whence

$$\begin{aligned}
\left|p^{-i\beta} - x^{-i\beta}\right| &= \left|1 - e^{i\beta(\log p - \log x)}\right| \\
&= 2\left|\sin\left(\frac{\beta}{2}(\log p - \log x)\right)\right| \\
&\leq \left|\beta(\log p - \log x)\right| \\
&\leq \delta|\beta|.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{x<p\leq(1+\delta)x}\frac{1}{p}\left(1-\mathrm{Re}\ \chi\overline{\xi}(p)p^{-i\beta}\right) &= \sum_{x<p\leq(1+\delta)x}\frac{1}{p}\left(1-\mathrm{Re}\ \chi\overline{\xi}(p)x^{-i\beta}\right)+O\left(\delta|\beta|\sum_{x<p\leq(1+\delta)x}\frac{1}{p}\right)\\
&= \sum_{x<p\leq(1+\delta)x}\frac{1}{p}\left(1-\mathrm{Re}\ \chi\overline{\xi}(p)x^{-i\beta}\right)+O\left(\frac{\delta^2|\beta|}{\log x}\right)\\
&= \sum_{x<p\leq(1+\delta)x}\frac{1}{p}\left(1-\mathrm{Re}\ \chi\overline{\xi}(p)e(\theta_x)\right)+O\left(\frac{\delta^2\log^2 y}{\log x}\right) \qquad (2.17)
\end{aligned}$$

where $\theta_x=-\frac{\beta}{2\pi}\log x$. We bound the sum from below in terms of the orders of χ and ξ:

$$\begin{aligned}
\sum_{x<p\leq(1+\delta)x}\frac{1}{p}\left(1-\mathrm{Re}\ \chi\overline{\xi}(p)\,e(\theta_x)\right) &= \sum_{\ell\ (\mathrm{mod}\ k)}\sum_{\substack{x<p\leq(1+\delta)x\\ \xi(p)=e\left(\frac{\ell}{k}\right)}}\frac{1}{p}\left(1-\mathrm{Re}\ \chi(p)\,e\left(-\frac{\ell}{k}\right)e(\theta_x)\right)\\
&\geq \sum_{\ell\ (\mathrm{mod}\ k)}\sum_{\substack{x<p\leq(1+\delta)x\\ \xi(p)=e\left(\frac{\ell}{k}\right)}}\frac{1}{p}\min_{z\in\mu_g\cup\{0\}}\left(1-\mathrm{Re}\ z\cdot e\left(\theta_x-\frac{\ell}{k}\right)\right)
\end{aligned}$$

A straightforward application of the Siegel-Walfisz theorem will give the following:

Lemma 2.15. *Fix $\epsilon>0$, $y\geq 2$, and χ and ξ as in the statement of Theorem 2.14. Then for all $x\geq\exp\left((\log y)^{\epsilon}\right)$,*

$$\sum_{\substack{x<p\leq(1+\delta)x\\ \xi(p)=e\left(\frac{\ell}{k}\right)}}\frac{1}{p}=\frac{\delta}{k\log x}\left(1+o(1)\right)$$

where $o(1)\to 0$ as $y\to\infty$ and depends only on y and ϵ.

Note that this estimate is independent of ℓ. Thus, the following general result, combined with Lemma 2.15, will furnish a lower bound on the sum (2.16):

Lemma 2.16. *Given $g \geq 3$ odd, $k \geq 2$ even, and $\theta \in \left(-\frac{1}{2}, \frac{1}{2}\right]$. Set $k^* = \frac{k}{(g,k)}$. Then*

$$\frac{1}{k} \sum_{\ell \,(mod\ k)} \min_{z \in \mu_g \cup \{0\}} \left(1 - Re\ z \cdot e\left(\theta - \frac{\ell}{k}\right)\right) = 1 - \frac{\sin \frac{\pi}{g}}{k^* \tan \frac{\pi}{gk^*}} F_{gk^*}(-gk^*\theta) \tag{2.18}$$

where

$$F_N(\omega) = \cos \frac{2\pi\{\omega\}}{N} + \left(\tan \frac{\pi}{N}\right) \sin \frac{2\pi\{\omega\}}{N}.$$

To make sense of this lemma, we examine some properties of $F_N(\omega)$. First, since $F_N(\omega) = F_N(\{\omega\})$ we may assume that $\omega \in [0, 1)$. Second, since k^* must be even, $gk^* \geq 6$, and we can therefore assume that $N \geq 6$. Under these assumptions, one easily checks that

- $F_N(0) = 1$ and $F_N(0.5) = \frac{1}{\cos \frac{\pi}{N}}$,
- $F_N(\omega)$ is concave down everywhere on $[0, 1)$,
- On the unit interval, F_N is symmetric about $\omega = \frac{1}{2}$, and
- The average value of F_N over the unit interval is $\frac{N}{\pi} \tan \frac{\pi}{N}$.

Thus, for the 'typical' θ we expect the right side of (2.18) to be δ_g. It is appreciably larger than δ_g when $gk^*\theta$ is close to an integer, and somewhat smaller than δ_g when $gk^*\theta$ is close to a half-integer. In the context of [15], $\theta = 0$, which allowed Granville and Soundararajan to bound (2.18) from below by δ_g quite easily. Although our arguments are also not difficult, the computations are naturally somewhat more involved; we will isolate the proof in a separate subsection.

Before proving the two lemmata, we deduce from them a lower bound on (2.16).

The main term of (2.17) can be bounded from below as follows, for all $x \geq \exp\big((\log y)^\epsilon\big)$:

$$\begin{aligned}
\sum_{x<p\leq(1+\delta)x} \frac{1}{p}\big(1-\mathrm{Re}\ \chi\overline{\xi}(p)e(\theta_x)\big) &\geq \sum_{\ell \,(\mathrm{mod}\ k)} \left(\sum_{\substack{x<p\leq(1+\delta)x\\ \xi(p)=e\left(\frac{\ell}{k}\right)}} \frac{1}{p}\right) \min_{z\in\mu_g\cup\{0\}}\left(1-\mathrm{Re}\ z\cdot e\left(\theta_x-\frac{\ell}{k}\right)\right)\\
&= \frac{\delta\big(1+o(1)\big)}{\log x}\left(1-\frac{\sin\frac{\pi}{g}}{k^*\tan\frac{\pi}{gk^*}}\,F_{gk^*}(-gk^*\theta_x)\right)
\end{aligned}$$

Let

$$G(t) = 1-\frac{\sin\frac{\pi}{g}}{k^*\tan\frac{\pi}{gk^*}}\,F_{gk^*}\left(\frac{\beta gk^*}{2\pi}\,t\right).$$

Note that G is minimized at values of t for which F_{gk^*} is maximized, whence $G(t) \geq 1-\frac{\sin\frac{\pi}{g}}{k^*\sin\frac{\pi}{gk^*}}$. It is a calculus exercise to show that $1-\frac{\sin\frac{\pi}{g}}{x\sin\frac{\pi}{gx}} > 0$ for all $x \geq 2$, from which we conclude that as a function of t, $G(t)$ is bounded away from 0. This combined with our choice of δ of size $(\log y)^{-3}$ shows that we can write (2.17) in the form

$$\begin{aligned}
\sum_{x<p\leq(1+\delta)x} \frac{1}{p}\big(1-\mathrm{Re}\ \chi\overline{\xi}(p)p^{-i\beta}\big) &\geq \frac{\big(1+o(1)\big)\delta}{\log x}\,G(\log x)+O\left(\frac{\delta^2\log^2 y}{\log x}\right)\\
&= \frac{\big(1+o(1)\big)\delta}{\log x}\,G(\log x) \qquad (2.19)
\end{aligned}$$

where the $o(1)$ term in (2.19) tends to 0 as $y\to\infty$ and depends only on y, ϵ, g, and k.

We now go back and prove the two lemmata.

Proof of Lemma 2.15. A consequence of the Siegel-Walfisz Theorem says that for any fixed $\epsilon > 0$ and $A > 0$, for all $X \geq \exp(m^\epsilon)$,

$$\theta(X;m,a) := \sum_{\substack{p\leq X\\ p\equiv a\,(\mathrm{mod}\ m)}} \log p = \frac{X}{\varphi(m)}\left(1+O\left(\frac{1}{(\log X)^A}\right)\right)$$

where the constant implicit in the O-term depends only upon A and ϵ. In particular, for all $X \geq \exp\bigl((\log y)^{\epsilon}\bigr)$,

$$\theta(X; m, a) = \frac{X}{\varphi(m)}\left(1 + O_{\epsilon}\left(\frac{1}{(\log X)^{4/\epsilon}}\right)\right) \tag{2.20}$$

where the implicit constant only depends on ϵ.

To apply Siegel-Walfisz, we must first express the sum in question as a sum over primes in arithmetic progressions:

$$\sum_{\substack{x<p\leq(1+\delta)x \\ \xi(p)=e\left(\frac{\ell}{k}\right)}} \frac{1}{p} = \sum_{\substack{a \,(\mathrm{mod}\ m) \\ \xi(a)=e\left(\frac{\ell}{k}\right)}} \sum_{\substack{x<p\leq(1+\delta)x \\ p\equiv a \,(\mathrm{mod}\ m)}} \frac{1}{p}.$$

Next, we rewrite the inner sum so that we are weighing the primes by $\log p$, rather than $\frac{1}{p}$. Observe that the condition $x < p \leq (1+\delta)x$ is equivalent to $\frac{1}{1+\delta}\,p \leq x < p$, so we have

$$\frac{x\log x}{p\log p} = \frac{x}{p}\cdot\frac{\log x}{\log p} = \bigl(1+O(\delta)\bigr)\cdot\left(1 + O\left(\frac{\delta}{\log p}\right)\right) = 1 + O(\delta).$$

Thus,

$$\begin{aligned}
\sum_{\substack{x<p\leq(1+\delta)x \\ p\equiv a \,(\mathrm{mod}\ m)}} \frac{1}{p} &= \frac{1+O(\delta)}{x\log x} \sum_{\substack{x<p\leq(1+\delta)x \\ p\equiv a \,(\mathrm{mod}\ m)}} \log p \\
&= \frac{1+O(\delta)}{x\log x}\Bigl(\theta\bigl((1+\delta)x; m, a\bigr) - \theta\bigl(x; m, a\bigr)\Bigr) \\
&= \frac{\delta}{\varphi(m)\,\log x} + O\left(\frac{\delta^2}{\varphi(m)\,\log x}\right) + O_{\epsilon}\left(\frac{1}{\varphi(m)\,(\log x)^{1+4/\epsilon}}\right)
\end{aligned}$$

upon applying (2.20) to each θ term separately.

By assumption, $\log x \geq (\log y)^{\epsilon}$ and δ is of size $(\log y)^{-3}$, so

$$\sum_{\substack{x<p\leq(1+\delta)x \\ p\equiv a \,(\mathrm{mod}\ m)}} \frac{1}{p} = \frac{\delta}{\varphi(m)\,\log x}\left(1+O_{\epsilon}\left(\frac{1}{\log y}\right)\right). \tag{2.21}$$

Since this estimate is independent of a, to prove the lemma it remains only to show that

$$\sum_{\substack{a\,(\mathrm{mod}\ m) \\ \xi(a)=e\left(\frac{\ell}{k}\right)}} 1 = \frac{\varphi(m)}{k}. \tag{2.22}$$

This is well-known, but the proof is short so we include it. Clearly,

$$\sum_{\substack{a\,(\mathrm{mod}\ m) \\ \xi(a)=e\left(\frac{\ell}{k}\right)}} 1 = \sum_{\substack{a\,\in\,(\mathbb{Z}/m\mathbb{Z})^{*} \\ \xi(a)=e\left(\frac{\ell}{k}\right)}} 1$$

For brevity, denote $(\mathbb{Z}/m\mathbb{Z})^{*}$ by G. Since ξ has order k, there is some $b \in G$ such that $1, \xi(b), \xi(b)^2, \ldots, \xi(b)^{k-1}$ are all distinct; on the other hand, all these must be k^{th} roots of unity. In particular, there exists some $g \in G$ such that $\xi(g) = e\left(\frac{1}{k}\right)$.

Let H be the kernel of ξ, i.e. $H = \{a \in G : \xi(a) = 1\}$. This is a normal subgroup of G, and $g^{\ell}H = \left\{a \in G : \xi(a) = e\left(\frac{\ell}{k}\right)\right\}$. G can therefore be decomposed as a disjoint union of the k cosets $g^{\ell}H$ with $0 \leq \ell \leq k-1$. Since $\left|g^{\ell}H\right| = |H|$, (2.22) must hold. Combining this with (2.21) yields the lemma. ☐

2.3.1.1 Proof of Lemma 2.16

Recall that $g \geq 3$ is odd, $k \geq 2$ is even, and $\theta \in \left(-\frac{1}{2}, \frac{1}{2}\right]$. Let $d = (g, k)$ and set $k^{*} = \frac{k}{d}$ and $g^{*} = \frac{g}{d}$.

To prove (2.18), it suffices to show

$$\sum_{\ell \,(\mathrm{mod}\ k)} \max_{z\in\mu_g\cup\{0\}} \mathrm{Re}\ z\cdot e\left(\theta-\frac{\ell}{k}\right) = d\cdot\frac{\sin\frac{\pi}{g}}{\tan\frac{\pi}{gk^*}}\cdot F_{gk^*}(-gk^*\theta) \tag{2.23}$$

Let $\mathcal{A}_0 = \left\{e(\beta) : -\frac{1}{2g} < \beta \le \frac{1}{2g}\right\}$ and set $\mathcal{A}_n = e\left(\frac{n}{g}\right)\mathcal{A}_0$; note that the disjoint union of $\mathcal{A}_n$ as n runs over any complete set of residues of $\mathbb{Z}/g\mathbb{Z}$ is the complex unit circle. In particular, for any $\ell\in\mathbb{Z}$ there is a unique $n_\ell \in \left(-\frac{g}{2},\frac{g}{2}\right]$ such that $e\left(\theta-\frac{\ell}{k}\right)\in\mathcal{A}_{n_\ell}$. By definition, this means that $e\left(-\frac{n_\ell}{g}\right)e\left(\theta-\frac{\ell}{k}\right)\in\mathcal{A}_0$. Since for all other $n\in\left(-\frac{g}{2},\frac{g}{2}\right]$ we have $e\left(-\frac{n}{g}\right)e\left(\theta-\frac{\ell}{k}\right)\notin\mathcal{A}_0$, we deduce that

$$\begin{aligned}\max_{z\in\mu_g\cup\{0\}} \mathrm{Re}\ z\cdot e\left(\theta-\frac{\ell}{k}\right) &= \mathrm{Re}\ e\left(-\frac{n_\ell}{g}\right)e\left(\theta-\frac{\ell}{k}\right)\\ &= \mathrm{Re}\ e(\theta)\,e\left(\frac{f(\ell)}{gk}\right)\end{aligned}$$

where $f:\mathbb{Z}\to\mathbb{Z}$ is defined $f(\ell) = -(g\,\ell + k\,n_\ell)$. This allows us to rewrite the left hand side of the inequality (2.23):

$$\sum_{\ell \,(\mathrm{mod}\ k)} \max_{z\in\mu_g\cup\{0\}} \mathrm{Re}\ z\cdot e\left(\theta-\frac{\ell}{k}\right) = \mathrm{Re}\ e(\theta)\sum_{\ell \,(\mathrm{mod}\ k)} e\left(\frac{f(\ell)}{gk}\right). \tag{2.24}$$

Our aim is rewrite the sum on the right side of (2.24) in terms of geometric series.

It is not hard to see that if $\ell_1\equiv\ell_2 \,(\mathrm{mod}\ k)$ then $f(\ell_1)\equiv f(\ell_2)\,(\mathrm{mod}\ gk)$. However, more is true:

Lemma 2.17. $\ell_1\equiv\ell_2\,(mod\ k^*) \Longrightarrow f(\ell_1)\equiv f(\ell_2)\,(mod\ gk)$

Proof. Given $\ell_1\equiv\ell_2\,(\mathrm{mod}\ k^*)$. Then $k \mid g(\ell_2-\ell_1)$, since $g(\ell_2-\ell_1) = g^*k\,\frac{\ell_2-\ell_1}{k^*}$. Equivalently, there exists $m\in\mathbb{Z}$ such that $-\frac{\ell_1}{k} = -\frac{\ell_2}{k}+\frac{m}{g}$. Therefore, by the definition of n_ℓ, we find that both $e\left(\frac{m-n_{\ell_1}}{g}\right)$ and $e\left(-\frac{n_{\ell_2}}{g}\right)$ belong to the set $e\left(\frac{\ell_2}{k}-\theta\right)\mathcal{A}_0$. But this implies that $n_{\ell_1} \equiv m + n_{\ell_2} \pmod{g}$, whence

$$e\left(\frac{f(\ell_1)}{gk}\right) = e\left(\frac{f(\ell_2)}{gk}\right)$$

and we conclude. $\square$

Thus, we can restrict the sum on the right side of (2.24) to $\mathbb{Z}/k^*\mathbb{Z}$:

$$\sum_{\ell \,(\mathrm{mod}\ k)} e\left(\frac{f(\ell)}{gk}\right) = d \cdot \sum_{\ell^* \,(\mathrm{mod}\ k^*)} e\left(\frac{f(\ell^*)}{gk}\right). \tag{2.25}$$

We now prove a weaker form of Lemma 2.17, which has the advantage of a converse.

Lemma 2.18. *$\ell_1 \equiv \ell_2 \,(mod\ k^*) \Longleftrightarrow f(\ell_1) \equiv f(\ell_2) \,(mod\ k)$*

Proof.

$$\begin{aligned} f(\ell_1) \equiv f(\ell_2) \pmod{k} &\implies k \mid g(\ell_2 - \ell_1) \\ &\implies k^* \mid g^*(\ell_2 - \ell_1) \\ &\implies \ell_1 \equiv \ell_2 \pmod{k^*} \end{aligned}$$

since $(g^*, k^*) = 1$. On the other hand,

$$\ell_1 \equiv \ell_2 \pmod{k^*} \implies k \mid d(\ell_2 - \ell_1)$$

whence

$$\begin{aligned} f(\ell_1) - f(\ell_2) &= g(\ell_2 - \ell_1) + k(n_{\ell_1} - n_{\ell_2}) \\ &= g^* d(\ell_2 - \ell_1) + k(n_{\ell_1} - n_{\ell_2}) \\ &\equiv 0 \pmod{k}. \end{aligned}$$

$\square$

Proposition 2.19. *The map f restricted to $\left[-\frac{k^*}{2}+k^*\theta, \frac{k^*}{2}+k^*\theta\right) \cap \mathbb{Z}$ is an injection into $\left(-\frac{k}{2}-gk\theta, \frac{k}{2}-gk\theta\right] \cap \mathbb{Z}$.*

Proof. Injectivity follows immediately from Lemma 2.18, so it suffices to show that the image of $\left[-\frac{k^*}{2}+k^*\theta, \frac{k^*}{2}+k^*\theta\right) \cap \mathbb{Z}$ under f lands in the claimed target. In fact, we will show a slightly stronger statement. Observe that because $|\theta| \le \frac{1}{2}$, $\left[-\frac{k^*}{2}+k^*\theta, \frac{k^*}{2}+k^*\theta\right) \subseteq \left[-\frac{k}{2}+k\theta, \frac{k}{2}+k\theta\right)$; we claim that the image under f of the larger set lands inside the claimed target.

Fix any $\ell \in \left[-\frac{k}{2}+k\theta, \frac{k}{2}+k\theta\right)$; this is equivalent to requiring $\theta - \frac{\ell}{k} \in \left(-\frac{1}{2}, \frac{1}{2}\right]$. By definition of n_ℓ we have $e\left(\theta-\frac{\ell}{k}\right) \in \mathcal{A}_{n_\ell}$, from which we deduce that for some integer N, $\theta - \frac{\ell}{k} \in \left(N + \frac{2n_\ell-1}{2g}, N+\frac{2n_\ell+1}{2g}\right]$. By our restriction on ℓ, N must equal 0 (recall that $-\frac{g-1}{2} \le n_\ell \le \frac{g-1}{2}$). It follows that $f(\ell) \in \left(-\frac{k}{2}-gk\theta, \frac{k}{2}-gk\theta\right]$. $\square$

Note that $d \mid f(\ell)$ for all ℓ. Combining this fact with Proposition 2.19 we conclude that

$$\left\{ f(\ell^*) : -\frac{k^*}{2} + k^*\theta \le \ell^* < \frac{k^*}{2} + k^*\theta \right\}$$

is a set of k^* distinct multiples of d, all contained in $\left(-\frac{k}{2}-gk\theta, \frac{k}{2}-gk\theta\right]$. But this set contains *precisely* k^* multiples of d. Therefore:

$$\begin{aligned}
\sum_{\ell^* \pmod{k^*}} e\left(\frac{f(\ell^*)}{gk}\right) &= \sum_{-\frac{k^*}{2}+k^*\theta \le \ell^* < \frac{k^*}{2}+k^*\theta} e\left(\frac{f(\ell^*)}{gk}\right) \\
&= \sum_{\frac{1}{d}\left(-\frac{k}{2}-gk\theta\right) < m \le \frac{1}{d}\left(\frac{k}{2}-gk\theta\right)} e\left(\frac{md}{gk}\right) \\
&= \sum_{-\frac{k^*}{2}-gk^*\theta < m \le \frac{k^*}{2}-gk^*\theta} e\left(\frac{m}{gk^*}\right) \qquad (2.26)
\end{aligned}$$

This is a k^*-term geometric series with first term $e\left(\frac{1}{gk^*}\left[\frac{k^*}{2}-gk^*\theta\right]\right)$ and ratio $e\left(-\frac{1}{gk^*}\right)$.

Summing the series and performing standard algebraic manipulations, one finds

$$\sum_{-\frac{k^*}{2}-gk^*\theta<m\leq\frac{k^*}{2}-gk^*\theta} e\left(\frac{m}{gk^*}\right) = e\left(-\theta+\frac{1-2c}{2gk^*}\right)\frac{\sin\frac{\pi}{g}}{\sin\frac{\pi}{gk^*}}$$

where $c = \{-gk^*\theta\} \in [0,1)$. Tracing back through equations (2.24)-(2.26) and simplifying, we see that

$$\begin{aligned}\sum_{\ell \pmod k} \max_{z\in\mu_g\cup\{0\}} \mathrm{Re}\ z\cdot e\left(\theta-\frac{\ell}{k}\right) &= d\cdot\frac{\sin\frac{\pi}{g}}{\sin\frac{\pi}{gk^*}}\cdot\cos\left(\frac{\pi}{gk^*}(1-2c)\right)\\ &= d\cdot\frac{\sin\frac{\pi}{g}}{\tan\frac{\pi}{gk^*}}\cdot F_{gk^*}\big(-gk^*\theta\big)\end{aligned}$$

proving (2.23), and thus the lemma. □

2.3.2 Completing the proof of Theorem 2.14

Let $x_0 = \exp\big((\log y)^\epsilon\big)$ and set $x_r = x_0(1+\delta)^r$. Then from (2.19) we deduce

$$\begin{aligned}\mathbb{D}(\chi(n),\xi(n)\,n^{i\beta};y)^2 &= \sum_{p\leq y}\frac{1}{p}\big(1-\mathrm{Re}\ \chi\overline{\xi}(p)p^{-i\beta}\big)\\ &\geq \sum_{x_0<p\leq y}\frac{1}{p}\big(1-\mathrm{Re}\ \chi\overline{\xi}(p)p^{-i\beta}\big)\\ &\geq \sum_{\substack{r\geq 0\\ x_{r+1}\leq y}}\ \sum_{x_r<p\leq x_{r+1}}\frac{1}{p}\big(1-\mathrm{Re}\ \chi\overline{\xi}(p)p^{-i\beta}\big)\\ &\geq \sum_{\substack{r\geq 0\\ x_{r+1}\leq y}}\frac{\big(1+o(1)\big)\delta}{\log x_r}G(\log x_r)\\ &\geq \big(1+o(1)\big)\log(1+\delta)\sum_{\substack{r\geq 0\\ x_{r+1}\leq y}}\frac{G(\log x_r)}{\log x_r}\end{aligned}\tag{2.27}$$

We recognize the sum above as the left Riemann sum – with subintervals of length $\log(1+\delta)$ – for the integral $\int\limits_{\log x_0}^{\log x_m}\frac{G(t)}{t}\,dt$, where m is the integer such that $x_m \leq y <$

x_{m+1}. We recall that

$$G(t) = 1 - \frac{\sin\frac{\pi}{g}}{k^* \tan\frac{\pi}{gk^*}} F_{gk^*}\left(\frac{\beta g k^*}{2\pi} t\right)$$

where $F_N(\omega) = \cos\frac{2\pi\{\omega\}}{N} + \left(\tan\frac{\pi}{N}\right)\sin\frac{2\pi\{\omega\}}{N}$; various properties of F_N are listed on page 44.

In general, if L is a left Riemann sum of a continuously differentiable function f, then

$$\left|\int_a^b f(x)\ dx - L\right| \leq \frac{(b-a)}{2} \cdot \Delta \cdot \max_{[a,b]} |f'(x)|$$

where Δ is the size of the longest subinterval. In our case, $\frac{G(t)}{t}$ is not everywhere differentiable – it has cusps at integer values of t – but one can circumvent this issue by choosing δ so that 1 is an integer multiple of $\log(1+\delta)$, and perturbing x_0 to make $\log x_0$ an integer. We have

$$\begin{aligned}
\left|\frac{d}{dt}\left(\frac{G(t)}{t}\right)\right| &\leq \left|\frac{G'(t)}{t}\right| + \left|\frac{G(t)}{t^2}\right| \\
&\leq \frac{\frac{\sin\frac{\pi}{g}}{k^*\tan\frac{\pi}{gk^*}} F'_{gk^*}(0)}{\log x_0} + \frac{2}{(\log x_0)^2} \\
&\ll 1
\end{aligned}$$

for all $t \geq \log x_0$, whence

$$\begin{aligned}
\left|\log(1+\delta) \sum_{\substack{r\geq 0 \\ x_{r+1}\leq y}} \frac{G(\log x_r)}{\log x_r} - \int_{\log x_0}^{\log y} \frac{G(t)}{t}\ dt\right| &\ll (\log y)\cdot\log(1+\delta) + \left|\int_{\log x_m}^{\log y} \frac{G(t)}{t}\ dt\right| \\
&\ll \frac{1}{\log^2 y}
\end{aligned}$$

Therefore, continuing our calculation from where we left it in (2.27),

$$\mathbb{D}(\chi(n), \xi(n)\, n^{i\beta}; y)^2 \geq \big(1 + o(1)\big) \int\limits_{\log x_0}^{\log y} \frac{G(t)}{t}\ dt + O(1) \tag{2.28}$$

where

$$G(t) = 1 - \frac{\sin\frac{\pi}{g}}{k^* \tan\frac{\pi}{gk^*}}\, F_{gk^*}\left(\frac{\beta g k^*}{2\pi} t\right).$$

To prove Theorem 2.14 it remains only to bound the integral on the right side of (2.28) from below by $\big(\delta_g + o(1)\big) \log\log y$. Recall that $F_N(t)$ is concave down everywhere on the unit interval and symmetric about $t = \frac{1}{2}$, with minima at the endpoints of the interval. Further, $\overline{F_N}$, the mean value of F_N on the unit interval, is $\frac{N}{\pi} \tan\frac{\pi}{N}$. Observe that it suffices to prove

$$\int\limits_{a(y)}^{b(y)} \frac{1}{t} F_N(t)\ dt \leq \big(\overline{F_N} + o(1)\big) \log\log y$$

where $a(y) = \frac{N\beta}{2\pi}(\log y)^\epsilon$, $b(y) = \frac{N\beta}{2\pi}\log y$, and $\dfrac{\log\log y}{\log y} \ll |\beta| \leq \log^2 y$.

Given any $x \geq 1$ we find

$$\int\limits_{1}^{x} \frac{1}{t}\, F_N(t)\ dt = \overline{F_N} \cdot \log x + O(1),$$

by splitting the integral into unit intervals (with at most one exception) and on each interval bounding $\frac{1}{t}$ from above and below trivially. Thus if $a(y) \geq 1$, we immediately find

$$\begin{aligned} \int\limits_{a(y)}^{b(y)} \frac{1}{t}\, F_N(t)\ dt \quad &= \quad \overline{F_N} \cdot \log\frac{b(y)}{a(y)} + O(1) \\ &\leq \quad \Big(\overline{F_N} + o(1)\Big) \log\log y. \end{aligned}$$

Now we consider the case when $a(y) < 1$. We have

$$\begin{aligned}\int_{a(y)}^{b(y)} \frac{1}{t} F_N(t)\, dt &= \int_{a(y)}^{1} \frac{1}{t} F_N(t)\, dt + \int_{1}^{b(y)} \frac{1}{t} F_N(t)\, dt \\ &= \int_{1}^{\frac{1}{a(y)}} \frac{1}{t} F_N\left(\frac{1}{t}\right)\, dt + \overline{F_N} \cdot \log b(y) + O(1)\end{aligned}$$

Therefore, it suffices to show that

$$\int_{1}^{x} \frac{1}{t} F_N\left(\frac{1}{t}\right)\, dt \leq \overline{F_N} \cdot \log x + O(1). \tag{2.29}$$

For all sufficiently large x, $F_N(x) \leq \overline{F_N}$, whence

$$\frac{d}{dx}\left(\int_{1}^{x} \frac{1}{t} F_N\left(\frac{1}{t}\right)\, dt\right) \leq \frac{d}{dx}\left(\overline{F_N} \cdot \log x\right)$$

for all large x. This implies (2.29), and Theorem 2.14 is proved. $\square$

2.4 The major arc case

Given $\chi \pmod q$ of odd order g. Recall that we wish to bound $\sum_{1\leq|n|\leq N} \frac{\overline{\chi}(n)}{n} e(n\alpha)$ non-trivially. In § 2.2.1 we proved Theorem 2.2 for the case that α belongs to a minor arc, i.e. when there exists a diophantine approximation to α, $\frac{b}{r}$, with $(\log q)^{2\delta_g} < r \leq M = e^{\sqrt{\log q}}$. The goal of this section is to complete the proof of Theorem 2.2, by showing that the same bound holds for α in a major arc.

For a character $\chi \pmod q$ of odd order the Granville-Soundararajan identity

(Proposition 2.12) asserts that for positive coprime integers b and r,

$$(2.30)\quad \sum_{1\le |n|\le N} \frac{\overline{\chi}(n)}{n}\, e\left(\frac{b}{r}\,n\right) = \sum_{d|r} \frac{\overline{\chi}(d)}{d}\cdot\frac{1}{\varphi\left(r/d\right)} \sideset{}{^o}\sum_{\psi\left(\mathrm{mod}\ \frac{r}{d}\right)} \tau\left(\overline{\psi}\right)\psi(b)\left(\sum_{a\le N/d}\frac{\overline{\chi}\psi(a)}{a}\right)$$

where $\sum^o$ indicates that the sum runs over odd characters. As remarked previously, the only way the left hand side can be large is if the sum $\displaystyle\sum_{a\le N/d}\frac{\overline{\chi}\psi(a)}{a}$ is large for some primitive odd character ψ whose conductor divides r. We will deduce a nontrivial bound on this quantity by combining Halász' theorem with results of Balog, Granville, and Soundararajan on the mimicry metric.

For the rest of this section, let $T_\chi(N,r) := \displaystyle\sum_{1\le |n|\le N}\frac{\overline{\chi}(n)}{n}\, e\left(\frac{b}{r}\,n\right)$.

By triangle inequality and partial summation we have

$$\begin{aligned}
T_\chi(N,r) &\le \sum_{d|r}\frac{1}{d}\cdot\frac{1}{\varphi\left(r/d\right)}\sideset{}{^o}\sum_{\psi\left(\mathrm{mod}\ \frac{r}{d}\right)}\sqrt{\frac{r}{d}}\left|\sum_{a\le N/d}\frac{\overline{\chi}\psi(a)}{a}\right| \\
&\le \sqrt{r}\sum_{d|r}\left(\frac{1}{d}\right)^{\frac{3}{2}}\frac{1}{\varphi\left(\frac{r}{d}\right)}\sideset{}{^o}\sum_{\psi\left(\mathrm{mod}\ \frac{r}{d}\right)}\left(\left|\frac{1}{N/d}\sum_{n\le\frac{N}{d}}\overline{\chi}\psi(n)\right| + \int_1^{\frac{N}{d}}\frac{1}{t}\left|\frac{1}{t}\sum_{n\le t}\overline{\chi}\psi(n)\right|\,dt\right)
\end{aligned}$$

We have the trivial estimate $\left|\dfrac{1}{N/d}\displaystyle\sum_{n\le\frac{N}{d}}\overline{\chi}\psi(n)\right| \ll 1$, so the contribution from this term to the total is $\ll\sqrt{r}$ which in turn is $\ll(\log q)^{\delta_g}$, i.e.

$$T_\chi(N,r) \ll \sqrt{r}\sum_{d|r}\left(\frac{1}{d}\right)^{\frac{3}{2}}\frac{1}{\varphi\left(\frac{r}{d}\right)}\sideset{}{^o}\sum_{\psi\left(\mathrm{mod}\ \frac{r}{d}\right)}\int_1^{\frac{N}{d}}\frac{1}{t}\left|\frac{1}{t}\sum_{n\le t}\overline{\chi}\psi(n)\right|\,dt + O\big((\log q)^{\delta_g}\big)$$

Next, we show that the part of the integral with $1\le t\le e^r$ is not too large: making the trivial estimate $\left|\frac{1}{t}\sum_{n\le t}\overline{\chi}\psi(n)\right| \ll 1$ we deduce

$$\sqrt{r}\sum_{d|r}\left(\frac{1}{d}\right)^{\frac{3}{2}}\frac{1}{\varphi\left(\frac{r}{d}\right)}\sum_{\psi\left(\text{mod }\frac{r}{d}\right)}^{o}\int_{1}^{e^r}\frac{1}{t}\left|\frac{1}{t}\sum_{n\le t}\overline{\chi}\psi(n)\right|\,dt \ll r^{3/2}$$

whence

$$T_\chi(N,r) \ll \sqrt{r}\sum_{d|r}\left(\frac{1}{d}\right)^{\frac{3}{2}}\frac{1}{\varphi\left(\frac{r}{d}\right)}\sum_{\psi\left(\text{mod }\frac{r}{d}\right)}^{o}\int_{e^r}^{\frac{N}{d}}\frac{1}{t}\left|\frac{1}{t}\sum_{n\le t}\overline{\chi}\psi(n)\right|\,dt + (\log q)^{3\delta_g}.$$

Applying Halász's Theorem 2.13 we find

$$T_\chi(N,r) \ll \sqrt{r}\sum_{d|r}\left(\frac{1}{d}\right)^{\frac{3}{2}}\frac{1}{\varphi\left(\frac{r}{d}\right)}\sum_{\psi\left(\text{mod }\frac{r}{d}\right)}^{o}\int_{e^r}^{\frac{N}{d}}\left(\left(\log\log t\right)e^{-\mathcal{M}(t,\chi\overline{\psi})}+\frac{1}{\log t}\right)\frac{dt}{t}+(\log q)^{3\delta_g}.$$

The $\frac{1}{\log t}$ term contributes $\ll \sqrt{r}\log\log N$, which is negligible compared to the error term. So, we have

$$\begin{aligned}
T_\chi(N,r) &\ll \sqrt{r}\sum_{d|r}\left(\frac{1}{d}\right)^{\frac{3}{2}}\frac{1}{\varphi\left(\frac{r}{d}\right)}\sum_{\psi\left(\text{mod }\frac{r}{d}\right)}^{o}\int_{e^r}^{\frac{N}{d}}\left(\log\log t\right)e^{-\mathcal{M}(t,\chi\overline{\psi})}\frac{dt}{t}+(\log q)^{3\delta_g}\\
&= \sqrt{r}\sum_{d|r}\left(\frac{1}{d}\right)^{\frac{3}{2}}\int_{e^r}^{\frac{N}{d}}\left(\log\log t\right)\left(\frac{1}{\varphi\left(\frac{r}{d}\right)}\sum_{\psi\left(\text{mod }\frac{r}{d}\right)}^{o}e^{-\mathcal{M}(t,\chi\overline{\psi})}\right)\frac{dt}{t}+(\log q)^{3\delta_g}
\end{aligned}$$

From the work of Balog-Granville-Soundararajan [1], we know that for all but at most $\sqrt{\log\log t}$ of the characters $\left(\text{mod }\frac{r}{d}\right)$ we have

$$\mathcal{M}(t,\chi\overline{\psi}) \ge \left(1+o(1)\right)\log\log t.$$

Moreover, for all but at most one possible exceptional character $\left(\text{mod }\frac{r}{d}\right)$, Lemma 3.3

of [1] asserts that

$$\mathcal{M}(t,\chi\overline{\psi}) \geq \left(\frac{1}{3}+o(1)\right)\log\log t.$$

Finally, for the possible exceptional character $\xi_d \pmod{\frac{r}{d}}$, our Theorem 2.14 guarantees that

$$\mathcal{M}(t,\chi\overline{\xi_d}) \geq \big(\delta_g+o(1)\big)\log\log t.$$

Putting all this together, we see that

$$\begin{aligned}
\frac{1}{\varphi\left(\frac{r}{d}\right)}\sideset{}{^{o}}\sum_{\psi \left(\mathrm{mod}\ \frac{r}{d}\right)} e^{-\mathcal{M}(t,\chi\overline{\psi})} &\leq \frac{1}{(\log t)^{1+o(1)}}+\frac{1}{\varphi\left(\frac{r}{d}\right)}\frac{\sqrt{\log\log t}}{(\log t)^{1/3+o(1)}}+\frac{1}{\varphi\left(\frac{r}{d}\right)}\frac{1}{(\log t)^{\delta_g+o(1)}} \\
&\ll \frac{1}{(\log t)^{1+o(1)}}+\frac{1}{\varphi\left(\frac{r}{d}\right)}\frac{1}{(\log t)^{\delta_g+o(1)}} \qquad (2.31)
\end{aligned}$$

The contribution to $T_\chi(N,r)$ from the first term of (2.31) is

$$\ll \sqrt{r}\sum_{d|r}\left(\frac{1}{d}\right)^{\frac{3}{2}}\int_{e^r}^{\frac{N}{d}}\frac{1}{(\log t)^{1+o(1)}}\frac{dt}{t} \ll (\log N)^{2\delta_g}.$$

The second term of (2.31) makes a potentially larger contribution to $T_\chi(N,r)$: it adds

$$\begin{aligned}
&\ll \sqrt{r}\sum_{d|r}\left(\frac{1}{d}\right)^{\frac{3}{2}}\int_{e^r}^{\frac{N}{d}}\frac{1}{\varphi\left(\frac{r}{d}\right)}\frac{1}{(\log t)^{\delta_g+o(1)}}\frac{dt}{t} \\
&\ll_g \frac{1}{r}\sum_{d|r}\frac{d^{3/2}}{\varphi(d)}(\log N)^{1-\delta_g+o(1)} \\
&\ll_g \frac{1}{\sqrt{r}}(\log N)^{1-\delta_g+o(1)}.
\end{aligned}$$

Thus we have proved the following bound:

$$T_\chi(N,r) \ll_g \frac{1}{\sqrt{r}}(\log N)^{1-\delta_g+o(1)}+(\log q)^{3\delta_g}. \qquad (2.32)$$

2.5 Concluding the proof of Theorem 2.1

Suppose $\chi \pmod q$ is a primitive character of odd order g. From Pólya's Fourier expansion (2.2) we deduce that

$$\begin{aligned} S_\chi(t) &= -\frac{\tau(\chi)}{2\pi i}\sum_{1\le|n|\le q}\frac{\overline{\chi}(n)}{n}\,e\left(-\frac{nt}{q}\right)+O(\log q) \\ &\ll \sqrt{q}\left|\sum_{1\le|n|\le q}\frac{\overline{\chi}(n)}{n}\,e\left(-\frac{nt}{q}\right)\right|+\log q \qquad (2.33)\end{aligned}$$

We may assume that $0<\frac{t}{q}<1$. Set $\alpha=-\frac{t}{q}$ and let $M=e^{\sqrt{\log q}}$. By Dirichlet's theorem on diophantine approximation, there exists a reduced fraction $\frac{b}{r}$ with $r\le M$ such that

$$\left|\alpha-\frac{b}{r}\right|\le\frac{1}{rM}.$$

If α belongs to a minor arc, i.e. if $(\log q)^{2\delta_g}<r\le M$, then from (2.6), which is a special case of Corollary 2.8, we deduce that

$$\begin{aligned} \sum_{1\le|n|\le q}\frac{\overline{\chi}(n)}{n}\,e(n\alpha) &\ll \log r+\frac{1+(\log r)^{5/2}}{\sqrt{r}}\log q+\log\log q \\ &\ll \sqrt{\log q}+\frac{\log q}{r^{1/2-o(1)}} \\ &\ll_g (\log q)^{1-\delta_g+o(1)}. \qquad (2.34)\end{aligned}$$

Now suppose instead that α belongs to a major arc, i.e. $r\le(\log q)^{2\delta_g}$ and no minor arc approximation exists. By (2.5), which is a special case of Lemma 2.10, we conclude that there exists an $N\le q$ such that

$$\sum_{n\le q}\frac{\overline{\chi}(n)}{n}\,e(n\alpha)=\sum_{n\le N}\frac{\overline{\chi}(n)}{n}\,e\left(\frac{b}{r}\,n\right)+O(\log\log q). \qquad (2.35)$$

From (2.32) we know

$$\sum_{1\le|n|\le N} \frac{\overline{\chi}(n)}{n}\, e\!\left(\frac{b}{r}\, n\right) \ll_g \frac{1}{\sqrt{r}}(\log N)^{1-\delta_g+o(1)} + (\log q)^{3\delta_g}$$

whence from (2.35) we conclude

$$\begin{aligned} \sum_{1\le|n|\le q} \frac{\overline{\chi}(n)}{n}\, e(n\alpha) &\ll_g \frac{1}{\sqrt{r}}(\log N)^{1-\delta_g+o(1)} + (\log q)^{3\delta_g} + \log\log q \\ &\ll_g (\log q)^{1-\delta_g+o(1)}. \end{aligned} \tag{2.36}$$

Combining (2.34) and (2.36), we see that for *any* nonzero α,

$$\sum_{1\le|n|\le q} \frac{\overline{\chi}(n)}{n}\, e(n\alpha) \ll_g (\log q)^{1-\delta_g+o(1)}$$

which is Theorem 2.2. Plugging this bound into (2.33) we deduce the bound

$$S_\chi(t) \ll_g \sqrt{q}\,(\log q)^{1-\delta_g+o(1)}$$

as claimed in Theorem 2.1.

CHAPTER III

Characters to smooth moduli

3.1 Introduction

Introduced by Dirichlet to prove his celebrated theorem on primes in arithmetic progressions (see [9]), Dirichlet characters have proved to be a fundamental tool in number theory. In particular, character sums of the form

$$S_\chi(x) := \sum_{n\leq x} \chi(n)$$

(where $\chi \pmod q$ is a Dirichlet character) arise naturally in many classical problems of analytic number theory, from estimating the least quadratic nonresidue (mod p) to bounding L-functions. Recall that for any character $\chi \pmod q$, $|S_\chi(x)|$ is trivially bounded above by $\varphi(q)$. A folklore conjecture (which is a consequence of the Generalized Riemann Hypothesis) predicts that for non-principal characters the true bound should look like[1]

$$|S_\chi(x)| \ll_\epsilon \sqrt{x} \cdot q^\epsilon.$$

[1] Here and throughout we use Vinogradov's notation $f \ll g$ to mean $f = O(g)$, with variables in subscript to indicate dependence of the implicit constant.

AMS subject classification: 11L40, 11M06.

Although we are currently very far from being able to prove such a statement, there have been some significant improvements over the trivial estimate. The first such is due (independently) to Pólya and Vinogradov: they proved that

$$|S_\chi(x)| \ll \sqrt{q}\, \log q$$

(see [9], pages 135-137). Almost 60 years later, Montgomery and Vaughan [23] showed that conditionally on the Generalized Riemann Hypothesis (GRH) one can improve Pólya-Vinogradov to

$$|S_\chi(x)| \ll \sqrt{q}\, \log\log q.$$

This is a best possible result, since in 1932 Paley [26] had given an unconditional construction of an infinite class of quadratic characters for which the magnitude of the character sum could be made $\gg \sqrt{q}\, \log\log q$.

In their recent work [15], Granville and Soundararajan give a characterization of when a character sum can be large; from this they are able to deduce a number of new results, including an improvement of Pólya-Vinogradov (unconditionally) and of Montgomery-Vaughan (on GRH) for characters of small odd order. In the present paper we explore a different application of their characterization. Recall that a positive integer N is said to be *smooth* if its prime factors are all small relative to N; if in addition the product of all its prime factors is small, N is *powerful*. Building on the work of Granville-Soundararajan and using a striking estimate developed by Graham and Ringrose, we will obtain (in Section 3.5) the following improvement of Pólya-Vinogradov for characters of smooth conductor:

Theorem 1. *Given $\chi\,(mod\ q)$ a primitive character, with q squarefree. For any* *integer n, denote its largest prime factor by $\mathcal{P}(n)$. Then*

$$|S_\chi(x)| \ll \sqrt{q}\,(\log q)\left(\left(\frac{\log\log\log q}{\log\log q}\right)^{\frac{1}{2}} + \left(\frac{(\log\log\log q)^2 \log\left(\mathcal{P}(q)\, d(q)\right)}{\log q}\right)^{1/4}\right)$$

where $d(q)$ is the number of divisors of q, and the implied constant is absolute.

From the well-known upper bound $\log d(q) < \frac{\log q}{\log\log q}$ (see, for example, Ex. 1.3.3 of [25]), we immediately deduce the following weaker but more palatable bound:

Corollary. *Given $\chi\,(mod\ q)$ primitive, with q squarefree. Then*

$$|S_\chi(x)| \ll \sqrt{q}\,(\log q)\left(\frac{(\log\log\log q)^2}{\log\log q} + \frac{(\log\log\log q)^2 \log \mathcal{P}(q)}{\log q}\right)^{\frac{1}{4}}$$

where the implied constant is absolute.

For characters with powerful conductor, we can do better by appealing to work of Iwaniec [21]. We prove:

Theorem 2. *Given $\chi\,(mod\ q)$ a primitive Dirichlet character with q large and*

$$rad(q) \leq \exp\left((\log q)^{3/4}\right),$$

where the radical of q is defined

$$rad(q) := \prod_{p|q} p.$$

Then

$$|S_\chi(x)| \ll_\epsilon \sqrt{q}\,(\log q)^{7/8+\epsilon}.$$

The key ingredient in the proofs of Theorems 1 and 2 is also at the heart of [15]. In that paper, Granville and Soundararajan introduce a notion of 'distance' on the set of characters, and then show that $|S_\chi(x)|$ is large if and only if χ is close (with respect to their distance) to a primitive character of small conductor and opposite parity (ideas along these lines had been earlier approached by Hildebrand in [20], and - in the context of mean values of arithmetic functions - by Halász in [17, 18]). More precisely, given characters χ, ψ, let

$$\mathbb{D}(\chi, \psi; y) := \left(\sum_{p \leq y} \frac{1 - \text{Re } \chi(p)\overline{\psi}(p)}{p} \right)^{\frac{1}{2}}.$$

Although one can easily furnish characters $\chi \neq \psi$ and a y for which $\mathbb{D}(\chi, \psi; y) = 0$, all the other properties of a distance function are satisfied; in particular, a triangle inequality holds:

$$\mathbb{D}(\chi_1, \psi_1; y) + \mathbb{D}(\chi_2, \psi_2; y) \geq \mathbb{D}(\chi_1\chi_2, \psi_1\psi_2; y)$$

(see [16] for a more general form of this 'distance' and its role in number theory). Granville and Soundararajan's characterization of large character sums comes in the form of the following two theorems:

Theorem A ([15], Theorem 2.1)**.** *Given $\chi \,(mod\ q)$ primitive, let $\xi \,(mod\ m)$ be any primitive character of conductor less than $(\log q)^{\frac{1}{3}}$ which minimizes the quantity $\mathbb{D}(\chi, \xi; q)$. Then*

$$|S_\chi(x)| \ll (1 - \chi(-1)\xi(-1)) \frac{\sqrt{m}}{\varphi(m)} \sqrt{q} \log q \exp\left(-\frac{1}{2} \mathbb{D}(\chi, \xi; q)^2 \right) + \sqrt{q}\ (\log q)^{\frac{6}{7}}.$$

Theorem B ([15], Theorem 2.2)**.** *Given $\chi \,(mod\ q)$ a primitive character, let $\xi \,(mod\ m)$*

be any primitive character of opposite parity. Then

$$\max_x |S_\chi(x)| + \frac{\sqrt{m}}{\varphi(m)} \sqrt{q} \log\log q \gg \frac{\sqrt{m}}{\varphi(m)} \sqrt{q} \log q \exp\left(-\mathbb{D}(\chi,\xi;q)^2\right).$$

Roughly, the first theorem says that $|S_\chi(x)|$ is small (i.e. $\ll \sqrt{q}\,(\log q)^{6/7}$) unless there exists a primitive character ξ of small conductor and opposite parity, whose distance from χ is small (i.e. $\mathbb{D}(\chi,\xi;q)^2 \leq \frac{2}{7}\log\log q$); the second theorem says that if there exists a primitive character $\xi \pmod m$ of small conductor and of opposite parity, whose distance from χ is small, then $|S_\chi(x)|$ gets large. In particular, to improve Pólya-Vinogradov for a primitive character $\chi \pmod q$ it suffices (by Theorem A) to find a lower bound on the distance from χ to primitive characters of small conductor and opposite parity. For example, if one can find a positive constant δ, independent of q, for which

$$\mathbb{D}(\chi,\xi;q)^2 \geq (\delta + o(1))\log\log q \tag{3.1}$$

then Theorem A would immediately yield an improvement of Pólya-Vinogradov:

$$\max_x |S_\chi(x)| \ll \sqrt{q}\ (\log q)^{1-\frac{\delta}{2}+o(1)}.$$

As it turns out (see Lemma 3.2 of [15]), it is not too difficult to show (3.1) holds for χ a character of odd order g, with $\delta = \delta_g = 1 - \frac{g}{\pi}\sin\frac{\pi}{g}$.

Thus, to derive bounds on character sums from Theorem A, one must understand the magnitude of $\mathbb{D}(\chi,\xi;q)$; this is the problem we take up in Section 3.2. Since $\mathbb{D}(\chi,\xi;q) = \mathbb{D}(\chi\,\overline{\xi},1;q)$, we are naturally led to study lower bounds on distances of the form $\mathbb{D}(\chi,1;y)$, for χ a primitive character and y a parameter with some flexibility.

By definition,

$$\mathbb{D}(\chi, 1; y)^2 = \sum_{p \le y} \frac{1}{p} - \text{Re} \sum_{p \le y} \frac{\chi(p)}{p}.$$

The first sum on the right hand side is well-approximated by $\log\log y$ (a classical estimate due to Mertens, see pages 56-57 of [9]); we will show that the second sum is comparable to $|L(s_y, \chi)|$, where

$$s_y := 1 + \frac{1}{\log y}.$$

To be precise, in Section 3.2 we prove:

Lemma 3. *For all $y \ge 2$,*

$$\mathbb{D}(\chi, 1; y)^2 = \log \left| \frac{\log y}{L(s_y, \chi)} \right| + O(1).$$

Our problem is now reduced to finding upper bounds on $|L(s, \chi)|$ for s slightly larger than 1. This is a classical subject, and many bounds are available. Thanks to the remarkable work of Graham and Ringrose [13] on short character sums, a particularly strong upper bound on L-functions is known when the character has smooth modulus; from a slight generalization of their result we will deduce (in Section 3.3) the following:

Lemma 4. *Given a primitive character χ (mod Q), let r be any positive number such that for all $p \ge r$, $ord_p\, Q \le 1$. Let*

$$q' = q'_r := \prod_{p < r} p^{ord_p\, Q}$$

and denote by $\mathcal{P}(Q)$ the largest prime factor of Q. Then for all $y > 3$,

$$|L(s_y, \chi)| \ll \log q' + \frac{\log Q}{\log\log Q} + \sqrt{(\log Q)\big(\log \mathcal{P}(Q) + \log d(Q)\big)}$$

where the implied constant is absolute.

Using the bound $\log d(q) < \frac{\log q}{\log\log q}$ one deduces the friendlier but weaker bound

$$|L(s_y, \chi)| \ll \log q' + \frac{\log Q}{(\log\log Q)^{1/2}} + \sqrt{(\log Q)\big(\log \mathcal{P}(Q)\big)}.$$

Lemma 4 will enable us to prove Theorem 1. For the proof of Theorem 2, we need a corresponding bound for $L(s_y, \chi)$ when the conductor of χ is powerful. In Section 3.4, using a potent estimate of Iwaniec [21] we will prove:

Lemma 5. *Given χ (mod Q) a primitive Dirichlet character with Q large and*

$$rad(Q) \leq \exp\left(2\,(\log Q)^{3/4}\right).$$

Then for all $y > 3$,

$$|L(s_y, \chi)| \ll_\epsilon (\log Q)^{3/4+\epsilon}.$$

In the final section of the paper, we synthesize our results and prove Theorems 1 and 2.

Acknowledgements: I am indebted to Professor Soundararajan for suggesting the problem in the first place, for encouraging me throughout, and for making innumerable improvements to my exposition. I am also grateful to the referee for meticulously reading the manuscript and catching an important error in the proofs of Theorems 1 and 2, to Denis Trotabas and Bob Hough for some helpful discussions, and to the Stanford Mathematics Department, where the bulk of this project was completed.

3.2 The size of $\mathbb{D}(\chi, 1; y)$

How large should one expect $\mathbb{D}(\chi, 1; y)$ to be? Before proving Lemma 3 we gain intuition by exploring what can be deduced from GRH.

Proposition 3.1. *Assume GRH. For any non-principal character $\chi \pmod Q$ we have*

$$\mathbb{D}(\chi, 1; y)^2 = \log\log y + O(\log\log\log Q).$$

Proof. Since

$$\sum_{p \le y} \frac{1}{p} = \log\log y + O(1)$$

by Mertens' well-known estimate, we need only show that

$$\sum_{p \le y} \frac{\chi(p)}{p} = O\left(\log\log\log Q\right).$$

We may assume that $y > (\log Q)^6$, else the estimate is trivial. Recall that on GRH, for all $x > (\log Q)^6$ we have:

$$\theta(x, \chi) := \sum_{p \le x} \chi(p) \log p \ll \sqrt{x}\,(\log Qx)^2 \ll x^{5/6}$$

(such a bound may be deduced from the first formula appearing on page 125 of [9]). Partial summation now gives

$$\sum_{(\log Q)^6 < p \le y} \frac{\chi(p)}{p} = \int\limits_{(\log Q)^6}^{y} \frac{1}{t \log t}\, d\theta(t, \chi) \ll \frac{1}{\log Q} \ll 1$$

and the proposition follows. $\square$

We now return to unconditional results. Recall that the prime number theorem gives $\theta(x) := \sum_{p \le x} \log p \sim x$.

Proof of Lemma 3: As before, by Mertens' estimate it suffices to show that

$$\operatorname{Re} \sum_{p \le y} \frac{\chi(p)}{p} = \log |L(s_y, \chi)| + O(1) \tag{3.2}$$

where $s_y := 1 + (\log y)^{-1}$. From the Euler product we know

$$\log |L(s_y, \chi)| = \operatorname{Re} \sum_p \sum_{k=1}^{\infty} \frac{\chi(p)^k}{k\, p^{k s_y}} = \operatorname{Re} \sum_p \frac{\chi(p)}{p^{s_y}} + O(1)$$

so that (3.2) would follow from

$$\sum_{p \le y} \left(\frac{1}{p} - \frac{1}{p^{s_y}} \right) + \sum_{p > y} \frac{1}{p^{s_y}} \ll 1.$$

The first term above is

$$\sum_{p \le y} \left(\frac{1}{p} - \frac{1}{p^{s_y}} \right) = \sum_{p \le y} \frac{1 - \exp\left(-\frac{\log p}{\log y} \right)}{p} \le \frac{1}{\log y} \sum_{p \le y} \frac{\log p}{p} = \frac{1}{\log y} \int_1^y \frac{1}{t} \, d\theta(t) \ll 1$$

by partial summation and the prime number theorem. A second application of partial summation and the prime number theorem yields

$$\sum_{p > y} \frac{1}{p^{s_y}} = \int_y^{\infty} \frac{1}{t^{s_y} \log t} \, d\theta(t) \ll 1.$$

The lemma follows. $\square$

For a clearer picture of where we are heading, we work out a simple consequence of this result. Let $\chi \,(\mathrm{mod}\ q)$ and $\xi \pmod m$ be as in Theorem A. By Lemma 3,

$$\mathbb{D}(\chi, \xi; q)^2 = \mathbb{D}(\chi \overline{\xi}, 1; q)^2 = \log \left| \frac{\log q}{L(s_q, \chi \,\overline{\xi})} \right| + O(1) \,,$$

and Theorem A immediately yields:

Proposition 3.2. *Let $\chi\,(mod\ q)$ be a primitive character, and ξ a character as in Theorem A. Then*

$$|S_\chi(x)| \ll \sqrt{q}\,\sqrt{(\log q)\ \left|L(s_q,\chi\,\overline{\xi})\right|} + \sqrt{q}\,(\log q)^{6/7}.$$

Thus to improve Pólya-Vinogradov it suffices to prove

$$L(s_q,\chi\,\overline{\xi}) = o(\log q).$$

This is the problem we explore in the next two sections.

3.3 Proof of Lemma 4

We ultimately wish to bound $|L(s_q,\chi\,\overline{\xi})|$; in this section we explore the more general quantity $|L(s_y,\chi)|$, where throughout y will be assumed to be at least 3, and Q will denote the conductor of χ.

By partial summation (see (8) on page 33 of [9]),

$$L(s_y,\chi) = s_y \int_1^\infty \frac{1}{t^{s_y+1}} \left(\sum_{n\le t} \chi(n)\right)\ dt.$$

When $t > Q$, the character sum is trivially bounded by Q, so that this portion of the integral contributes an amount $\ll 1$. For $t \le T$ (a suitable parameter to be chosen later), we may bound our character sum by t, and therefore this portion of the integral contributes an amount $\ll \log T$. Thus,

$$|L(s_y,\chi)| \ll \left|\int_T^Q \frac{1}{t^2}\left(\sum_{n\le t}\chi(n)\right)\ dt\right| + 1 + \log T. \tag{3.3}$$

To bound the character sum in this range, we invoke a powerful estimate of Graham and Ringrose [13]. For technical reasons, we need a slight generalization of their theorem:

Theorem 3.3 (compare to Lemma 5.4 of [13])**.** *Given a primitive character χ (mod Q), with q' and $\mathcal{P}(Q)$ defined as in Lemma 4. Then for any $k \in \mathbb{N}$, writing $K := 2^k$, we have*

$$\left| \sum_{M<n\leq M+N} \chi(n) \right| \ll N^{1-\frac{k+3}{8K-2}}\, \mathcal{P}(Q)^{\frac{k^2+3k+4}{32K-8}}\, Q^{\frac{1}{8K-2}}\, (q')^{\frac{k+1}{4K-1}}\, d(Q)^{\frac{3k^2+11k+8}{16K-4}}\, (\log Q)^{\frac{k+3}{8K-2}}$$

where $d(Q)$ is the number of divisors of Q, and the implicit constant is absolute.

Our proof of this is a straightforward extension of the arguments given in [13]. For the sake of completeness, we write out all the necessary modifications explicitly in the appendix.

Armed with Theorem 3.3, we deduce Lemma 4 in short order. Set

$$T := \mathcal{P}(Q)^{3k}\, Q^{\frac{1}{k}}\, (q')^2\, d(Q)^{3k}\, (\log Q)^{\frac{16K}{k}}.$$

If $T \leq Q$, then for all $t \geq T$ Theorem 3.3 implies

$$\left| \sum_{n\leq t} \chi(n) \right| \ll \frac{t}{\log Q}$$

whence

$$\left| \int_T^Q \frac{1}{t^2} \left(\sum_{n\leq t} \chi(n) \right) dt \right| \ll 1.$$

From the bound (3.3) we deduce that for $T \leq Q$, $|L(s_y, \chi)| \ll \log T$. But for $T > Q$

such a bound holds trivially (irrespective of our choice of T). Therefore

$$|L(s_y,\chi)| \ll \log T \ll k\log\mathcal{P}(Q) + \frac{1}{k}\log Q + \log q' + k\log d(Q) + \frac{K}{k}\log\log Q.$$

It remains to choose k appropriately. Let

$$k' := \min\left\{\frac{1}{10}\log\log Q, \sqrt{\frac{\log Q}{\log\mathcal{P}(Q)+\log d(Q)}}\right\},$$

and set $k = [k'] + 1$. Writing $K' = 2^{k'}$ we have

$$k'\log\mathcal{P}(Q) + k'\log d(Q) \ll \sqrt{(\log Q)\big(\log\mathcal{P}(Q)+\log d(Q)\big)} \ll \frac{1}{k'}\log Q$$

and

$$\frac{K'}{k'}\log\log Q \ll (\log Q)^{\frac{\log 2}{10}}(\log\log Q) \ll (\log Q)^{\frac{1}{10}} \ll \frac{1}{k'}\log Q$$

Finally, since $K \ll K'$ and $k \asymp k'$ (i.e. $k \ll k' \ll k$) for all Q sufficiently large, we deduce:

$$|L(s_y,\chi)| \ll \log q' + \frac{1}{k}\log Q \ll \log q' + \frac{\log Q}{\log\log Q} + \sqrt{(\log Q)\big(\log\mathcal{P}(Q)+\log d(Q)\big)}.$$

The proof of Lemma 4 is now complete.

3.4 Proof of Lemma 5

Iwaniec, inspired by work of Postnikov [27] and Gallagher [10], proved the following:

Theorem 3.4 (See Lemma 6 of [21]). *Given χ (mod Q) a primitive Dirichlet char-*

acter. Then for all N, N' *satisfying* $(\text{rad } Q)^{100} < N < 9Q^2$ *and* $N < N' < 2N$*,*

$$\left| \sum_{N \le n \le N'} \chi(n) \right| < \gamma_{\scriptscriptstyle N}\, N^{1-\epsilon_N}$$

where

$$\gamma_x := \exp(C_1 z_x \log^2 C_2 z_x) \qquad \epsilon_x := \frac{1}{C_3 z_x^2 \log C_4 z_x} \qquad z_x := \frac{\log 3Q}{\log x}$$

and the C_i *are effective positive constants independent of* Q*.*

In fact, Lemma 6 of [21] is more general (bounding sums of $\chi(n)\, n^{it}$), and provides explicit choices of the constants C_i.

Proof of Lemma 5:

Recall the bound (3.3):

$$|L(s_y, \chi)| \ll \left| \int_T^Q \frac{1}{t^2} \left(\sum_{n \le t} \chi(n) \right) \, dt \right| + 1 + \log T.$$

Writing

$$\left| \sum_{n \le t} \chi(n) \right| \le \sqrt{t} + \left| \sum_{\sqrt{t} < n \le t} \chi(n) \right|,$$

partitioning the latter sum into dyadic intervals, and applying Iwaniec's result to each of these, we deduce that so long as $\sqrt{t} > (\text{rad } Q)^{100}$,

$$\left| \sum_{n \le t} \chi(n) \right| \ll (\log t)\, \gamma_t \, t^{1-\epsilon_t}$$

with $C_1 = 400$, $C_2 = 2400$, $C_3 = 4 \cdot 1800^2$, $C_4 = 7200$ in the definitions of γ_t and ϵ_t. Choosing $T = \exp((\log Q)^\alpha)$ for some $\alpha \in (0, 1)$ to be determined later, and assuming that $T > (\mathrm{rad}\ Q)^{200}$, our bound becomes

$$|L(s_y,\chi)| \ll (\log Q)^{\alpha} + \int\limits_{\exp((\log Q)^{\alpha})}^{Q} \frac{\log t}{t^2}\, \gamma_t\, t^{1-\epsilon_t}\; dt \tag{3.4}$$

Denote by $\int$ the integral in (3.4), and set $\delta_Q = \frac{\log 3}{\log Q}$. Making the substitution $z = \frac{\log 3Q}{\log t}$ and simplifying, one finds

$$\begin{aligned}
\int &= (\log^2 3Q) \int\limits_{1+\delta_Q}^{(1+\delta_Q)(\log Q)^{1-\alpha}} \frac{1}{z^3} \exp\left(C_1 z \log^2 C_2 z - \frac{\log 3Q}{C_3 z^3 \log C_4 z}\right)\, dz \\
&\ll \exp\left(2\log\log 3Q + C_1 (\log Q)^{1-\alpha}(\log\log Q)^2 - \frac{(\log Q)^{3\alpha-2}}{C_3 \log\log Q}\right) \int\limits_{1+\delta_Q}^{(1+\delta_Q)(\log Q)^{1-\alpha}} \frac{dz}{z^3} \\
&\ll 1
\end{aligned}$$

upon choosing $\alpha = \frac{3}{4} + \epsilon$. Plugging this back into (3.4) we conclude. □

It is plausible that with a more refined upper bound on the integral in (3.4) one could take a smaller value of α, thus improving the exponents in both Lemma 5 and Theorem 2.

3.5 Upper bounds on character sums

Given $\chi \pmod q$ a primitive character, recall from Proposition 3.2 the bound

$$|S_\chi(x)| \ll \sqrt{q}\sqrt{(\log q)\ \left|L(s_q, \chi\,\overline{\xi})\right|} + \sqrt{q}\,(\log q)^{6/7}$$

where $\xi \pmod m$ is the primitive character with $m < (\log q)^{1/3}$ which χ is closest to, and $s_q := 1 + \frac{1}{\log q}$.

To prove Theorems 1 and 2, we would like to apply Lemmas 4 and 5 (respectively) to derive a bound on $|L(s_q, \chi\overline{\xi})|$. An immediate difficulty is that both lemmas require the character to be primitive, which is not necessarily true of $\chi\overline{\xi}$. Instead, we will apply the lemmas to the primitive character which induces $\chi\overline{\xi}$; thus, we must understand the size of the conductor of $\chi\overline{\xi}$. This is the goal of the following simple lemma, which is surely well-known to the experts but which the author could not find in the literature. We write $[a, b]$ to denote the least common multiple of a and b, and $\mathrm{cond}(\psi)$ to denote the conductor of a character ψ.

Lemma 3.5. *For any non-principal Dirichlet characters $\chi_1 \pmod{q_1}$ and $\chi_2 \pmod{q_2}$,*

$$cond(\chi_1\chi_2) \,\Big|\, \big[cond(\chi_1), cond(\chi_2)\big]$$

Proof. First, observe that $\chi_1\chi_2$ is a character modulo $[q_1, q_2]$: one needs only check that it is completely multiplicative, periodic with period $[q_1, q_2]$, and that $\chi_1\chi_2(n) = 0$ if and only if $(n, [q_1, q_2]) > 1$. Since the conductor of a character divides its modulus, the lemma is proved in the case that both χ_1 and χ_2 are primitive.

Now suppose that χ_1 and χ_2 are not necessarily primitive; denote by $\tilde{\chi}_i \pmod{\tilde{q}_i}$ the primitive character which induces χ_i. By the argument above, we know that

$$\mathrm{cond}(\tilde{\chi}_1\tilde{\chi}_2) \,\Big|\, \big[\tilde{q}_1, \tilde{q}_2\big]. \tag{3.5}$$

Next we note that the character $\tilde{\chi}_1\tilde{\chi}_2$, while not necessarily primitive, does induce $\chi_1\chi_2$ (i.e. $\chi_1\chi_2 = \tilde{\chi}_1\tilde{\chi}_2\chi_0$ for χ_0 the trivial character modulo $[q_1, q_2]$), whence $\mathrm{cond}(\tilde{\chi}_1\tilde{\chi}_2) = \mathrm{cond}(\chi_1\chi_2)$. Plugging this into (3.5) we immediately deduce the lemma. □

Given $\chi \pmod{q}$ and $\xi \pmod{m}$ as at the start of the section, denote by $\psi \pmod{Q}$

the primitive character inducing $\chi\overline{\xi}$. Taking $\chi_1 = \chi$ and $\chi_2 = \overline{\xi}$ in Lemma 3.5, we see that $Q \mid [q, m]$; in particular, $Q \leq qm$. On the other hand, making the choice $\chi_1 = \chi\overline{\xi}$ and $\chi_2 = \xi$ yields $q \mid [Q, m]$, so $q \leq Qm$. Combining these two estimates, we conclude that

$$\frac{q}{m} \leq Q \leq qm \tag{3.6}$$

Since we will be working with both $L(s, \chi\overline{\xi})$ and $L(s, \psi)$, the following estimate will be useful:

Lemma 3.6. *Given $\chi\,(mod\ q)$ and $\xi\,(mod\ m)$ primitive characters, let $\psi\,(mod\ Q)$ be the primitive character which induces $\chi\overline{\xi}$. Then for all s with $Re(s) > 1$,*

$$\left| \frac{L(s, \chi\overline{\xi})}{L(s, \psi)} \right| \ll 1 + \log\log m.$$

Proof. For $\mathrm{Re}(s) > 1$ we have

$$\frac{L(s, \chi\overline{\xi})}{L(s, \psi)} = \prod_{\substack{p \mid [q,m] \\ p \nmid Q}} \left(1 - \frac{\psi(p)}{p^s}\right)$$

whence

$$\left| \frac{L(s, \chi\overline{\xi})}{L(s, \psi)} \right| \leq \prod_{\substack{p \mid [q,m] \\ p \nmid Q}} \left(1 + \frac{1}{p}\right).$$

From Lemma 3.5 we know $q \mid [Q, m]$; it follows that if $p \mid [q, m]$ and $p \nmid Q$ then p must divide m. Thus,

$$\prod_{\substack{p \mid [q,m] \\ p \nmid Q}} \left(1 + \frac{1}{p}\right) \leq \prod_{p \mid m} \left(1 + \frac{1}{p}\right).$$

Since

$$\log \prod_{p \mid m} \left(1 + \frac{1}{p}\right) = \sum_{p \mid m} \log\left(1 + \frac{1}{p}\right) \leq \sum_{p \mid m} \frac{1}{p},$$

to prove the lemma it suffices to show that for all m sufficiently large,

$$\sum_{p|m} \frac{1}{p} \le \log\log\log m + O(1). \tag{3.7}$$

Let $P = P(m)$ denote the largest prime such that $\prod_{p\le P} p \le m$. Then $\omega(m) \le \pi(P)$ (otherwise we would have $m \ge \mathrm{rad}(m) > \prod_{p\le P} p$, contradicting the maximality of P); therefore,

$$\sum_{p|m} \frac{1}{p} \le \sum_{p\le P} \frac{1}{p} = \log\log P + O(1).$$

Finally from the prime number theorem we know that for all m sufficiently large, $\theta(P) \ge \frac{1}{2}P$, whence $P \le 2\log m$ and the bound (3.7) follows. $\square$

With these lemmas in hand we can now prove Theorems 1 and 2 without too much difficulty.

Proof of Theorem 1: Given $\chi\,(\mathrm{mod}\ q)$ primitive with q squarefree, define the character $\xi\,(\mathrm{mod}\ m)$ as in Theorem A, and let $\psi\,(\mathrm{mod}\ Q)$ be the primitive character inducing $\chi\bar{\xi}$. Recall that we denote the largest prime factor of n by $\mathcal{P}(n)$.

From Proposition 3.2 we have

$$|S_\chi(x)| \ll \sqrt{q}\sqrt{(\log q)\ \left|L(s_q, \chi\,\bar{\xi})\right|} + \sqrt{q}\,(\log q)^{6/7}, \tag{3.8}$$

and Lemma 3.6 yields the bound

$$\left|L(s_q, \chi\,\bar{\xi})\right| \ll |L(s_q, \psi)|\ \log\log\log q. \tag{3.9}$$

Lemma 3.5 tells us that $Q \mid [q, m]$, whence for all primes $p > m$ we have

$$\mathrm{ord}_p\, Q \le \max(\mathrm{ord}_p\, q, \mathrm{ord}_p\, m) = \mathrm{ord}_p\, q \le 1$$

since q is squarefree. Therefore we may apply Lemma 4 to the character ψ, taking $y = q$ and

$$q' = \prod_{p \le m} p^{\operatorname{ord}_p Q} ;$$

this gives the bound

$$|L(s_q, \psi)| \ll \log q' + \frac{\log Q}{\log\log Q} + \sqrt{(\log Q) \log \big(\mathcal{P}(Q) d(Q)\big)}.$$

It remains only to bound the right hand side in terms of q, which we do term by term. The first term is small:

$$\begin{aligned}
\log q' &= \sum_{p \le m} (\operatorname{ord}_p Q) \log p \\
&\le \sum_{p \le m} (\operatorname{ord}_p q) \log p + \sum_{p \le m} (\operatorname{ord}_p m) \log p \\
&\le \theta(m) + \log m \\
&\ll (\log q)^{\frac{1}{3}}.
\end{aligned}$$

From (3.6) we deduce

$$\frac{\log Q}{\log\log Q} \ll \frac{\log q}{\log\log q} .$$

For the last term, Lemma 3.5 yields

$$d(Q) \le d(qm) \le d(q) d(m) \le d(q) (\log q)^{\frac{1}{3}}$$

and

$$\mathcal{P}(Q) \le \max\big(\mathcal{P}(q), \mathcal{P}(m)\big) \le \mathcal{P}(q)\mathcal{P}(m) \le \mathcal{P}(q)(\log q)^{\frac{1}{3}} ,$$

while (3.6) gives $\log Q \ll \log q$. Putting this all together, we find

$$|L(s_q, \psi)| \ll \frac{\log q}{\log\log q} + \sqrt{(\log q) \log \big(\mathcal{P}(q) d(q)\big)} ;$$

plugging this into (3.9) and (3.8) we deduce the theorem. $\square$

Proof of Theorem 2: Given $\chi \,(\mathrm{mod}\ q)$ with q large and $\mathrm{rad}(q) \leq \exp\left((\log q)^{\frac{3}{4}}\right)$, let $\xi \,(\mathrm{mod}\ m)$ be defined as in Theorem A, and let $\psi \,(\mathrm{mod}\ Q)$ denote the primitive character which induces $\chi\overline{\xi}$. We have $\mathrm{rad}(m) \leq \exp\big(\theta(m)\big)$, whence by the prime number theorem $\exists\, C > 0$ with

$$\begin{aligned} \mathrm{rad}(Q) &\leq \mathrm{rad}(q)\,\mathrm{rad}(m) \\ &\leq \exp\left((\log q)^{3/4} + C\,(\log q)^{1/3}\right) \\ &\leq \exp\left(\frac{4}{3}\,(\log q)^{3/4}\right) \end{aligned}$$

for all q sufficiently large. From (3.6) we deduce

$$\left(\frac{\log Q}{\log q}\right)^{\frac{3}{4}} \geq \left(\frac{\log \frac{q}{m}}{\log q}\right)^{\frac{3}{4}} \geq \left(1 - \frac{\log\log q}{\log q}\right) \geq \frac{2}{3}$$

for q sufficiently large, whence

$$\mathrm{rad}(Q) \leq \exp\left(2\,(\log Q)^{3/4}\right) .$$

Combining Lemma 5 with (3.9) and (3.6) we obtain

$$\left|L(s_q,\, \chi\,\overline{\xi})\right| \ll_\epsilon (\log\log\log q)(\log Q)^{3/4+\epsilon} \leq (\log\log\log q)(\log qm)^{3/4+\epsilon} \ll_\epsilon (\log q)^{3/4+\epsilon}.$$

Plugging this into Proposition 3.2 yields Theorem 2. $\square$

APPENDICES

APPENDIX A

Properties of Dirichlet characters

In the sequel, we use the convention that $(\mathbb{Z}/1\mathbb{Z})^* := \{1\}$.

Definition 1.1. Some terminology:

- A **Dirichlet character**, written $\chi \pmod q$, is any group homomorphism

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

- The **principal** or **trivial character** modulo q, denoted $\chi_0 \pmod q$, is the trivial homomorphism whose kernel is the entire group $(\mathbb{Z}/q\mathbb{Z})^*$.

- Given $\chi \pmod q$, we say $d \in \mathbb{N}$ is an **induced modulus** if the projection map

$$\begin{aligned} \pi : (\mathbb{Z}/q\mathbb{Z})^* &\longrightarrow (\mathbb{Z}/d\mathbb{Z})^* \\ x + q\mathbb{Z} &\longmapsto x + d\mathbb{Z} \end{aligned}$$

  is well-defined and χ factors through $(\mathbb{Z}/d\mathbb{Z})^*$ via the map π, i.e. there exists a function (not necessarily a homomorphism) $\chi' : (\mathbb{Z}/d\mathbb{Z})^* \longrightarrow \mathbb{C}^*$ which makes the following diagram commute:

$$\begin{tikzcd}
(\mathbb{Z}/q\mathbb{Z})^* \arrow[r, "\chi"] \arrow[d, "\pi"'] & \mathbb{C}^* \\
(\mathbb{Z}/d\mathbb{Z})^* \arrow[ur, dashed, "\chi'"'] &
\end{tikzcd} \tag{A.1}$$

- The **conductor** of $\chi \pmod q$ is its smallest induced modulus.

- $\chi \pmod q$ is **primitive** if its conductor is q.

Any Dirichlet character $\chi \pmod q$ can be extended to a completely multiplicative, periodic function

$$\begin{aligned} f_\chi : \quad \mathbb{Z} &\longrightarrow \mathbb{C} \\ n &\longmapsto \begin{cases} \chi(n + q\mathbb{Z}) & \text{if } (n,q) = 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{A.2}$$

Conversely, if f is any function from $\mathbb{Z}$ to $\mathbb{C}$ which is (1) completely multiplicative, (2) periodic with period q, and (3) non-zero iff the argument is relatively prime to q, then $f\Big|_{(\mathbb{Z}/q\mathbb{Z})^*}$ is a character $\chi \pmod q$, and $f = f_\chi$. Thus, although they are different functions, χ and f_χ carry precisely the same data. Following convention, we will abuse notation and refer to the extended function f_χ simply as χ.

If d is an induced modulus of $\chi \pmod q$, then by definition π must be well-defined, i.e.

1. if $x \in (\mathbb{Z}/q\mathbb{Z})^*$ then $\pi(x) \in (\mathbb{Z}/d\mathbb{Z})^*$; and

2. if $x \equiv y \pmod q$ and $(x,q) = 1$ then $\pi(x) = \pi(y)$.

The second condition is equivalent to requiring that $x + q \equiv x \pmod d$ whenever $(x,q) = 1$. Since $(\mathbb{Z}/q\mathbb{Z})^*$ is non-empty, we have proved:

Proposition 1.2. *If d is an induced modulus of $\chi \pmod q$, then $d|q$.*

Remark 1.3. Condition 2 implies condition 1 above: if condition 2 holds, then $d|q$, and writing $q = kd$ we have $(x, q) = 1 \Longrightarrow (x, kd) = 1 \Longrightarrow (x, d) = 1$. Thus we have actually proved that π is well-defined iff $d|q$.

Next, we wish to show that when d is an induced modulus of $\chi \pmod q$, the projection map $\pi : (\mathbb{Z}/q\mathbb{Z})^* \longrightarrow (\mathbb{Z}/d\mathbb{Z})^*$ is a surjective homomorphism.

Proposition 1.4. *Given arbitrary $n, d, q \in \mathbb{N}$ such that $(n, d) = 1$. Then $\exists N \equiv n \pmod d$ such that $(N, q) = 1$.*

Proof. For all primes p set $A_p := \{k \in \mathbb{Z} : p|n + kd\}$. We will first show that all the elements of A_p are congruent to each other $\pmod p$. An application of the Chinese Remainder Theorem will then furnish a suitable lift N of n.

Suppose $\exists m, \ell \in A_p$ such that $m \not\equiv \ell \pmod p$. Then $p|n+md$ and $p|n+\ell d$ whence $p|(m - \ell)d$. This would imply that $p|d$, whence $p|md$, whence $p|n$. But this would mean $p|(n, d)$, which is impossible since $(n, d) = 1$.

Thus as claimed, all elements of A_p are congruent to each other $\pmod p$; let a_p denote the residue class to which they all belong $\pmod p$. Since there are only a finite number of distinct prime factors of q, the Chinese Remainder Theorem yields an integer b which simultaneously satisfies $b \equiv a_p+1 \pmod p$ for all $p|q$. In particular, $b \not\equiv a_p \pmod p$ for every $p|q$, whence $b \notin A_p$ for every $p|q$. But this means q shares no prime factors with $N := n + bd$. So $(N, q) = 1$, and clearly $N \equiv n \pmod d$. $\square$

It immediately follows that the projection map π is surjective; it is also clearly a homomorphism. As a corollary, we can now prove the following basic but important result:

Proposition 1.5. *For any induced modulus d of $\chi \pmod q$, the function χ' appearing in the commutative diagram* (A.1) *is a character $\pmod d$, and is unique.*

Proof. The surjectivity of π guarantees the uniqueness of χ', so it suffices to show that for any $a, b \in (\mathbb{Z}/d\mathbb{Z})^*$, $\chi'(a)\chi'(b) = \chi'(ab)$. Given any $n \in (\mathbb{Z}/d\mathbb{Z})^*$, the surjectivity of π guarantees the existence of an $\widetilde{n} \in (\mathbb{Z}/q\mathbb{Z})^*$ such that $\pi(\widetilde{n}) = n$. If $\widetilde{n_1}, \widetilde{n_2} \in \pi^{-1}(n)$ are any two lifts of n, then from the commutativity of the diagram (A.1) and the hypothesis that χ' is a function we deduce that

$$\chi(\widetilde{n_1}) = \chi'\big(\pi(\widetilde{n_1})\big) = \chi'(n) = \chi'\big(\pi(\widetilde{n_2})\big) = \chi(\widetilde{n_2})$$

i.e. χ is independent of the choice of lift. Therefore if $\widetilde{a}, \widetilde{b} \in (\mathbb{Z}/q\mathbb{Z})^*$ are lifts of $a, b \in (\mathbb{Z}/d\mathbb{Z})^*$, we have

$$\begin{aligned}
\chi'(a)\chi'(b) &= \chi(\widetilde{a})\chi\left(\widetilde{b}\right) = \chi\left(\widetilde{a}\widetilde{b}\right) \\
&= \chi'\left(\pi\left(\widetilde{a}\widetilde{b}\right)\right) = \chi'\left(\pi(\widetilde{a})\pi\left(\widetilde{b}\right)\right) \\
&= \chi'(ab)
\end{aligned}$$

$\square$

Conversely, given a character χ' (mod d) and q any multiple of d, we know (see Remark 1.3) that the projection $\pi : (\mathbb{Z}/q\mathbb{Z})^* \longrightarrow (\mathbb{Z}/d\mathbb{Z})^*$ is a well-defined (surjective) homomorphism, whence χ' induces a unique character $\chi \pmod{q}$: $\chi = \chi' \circ \pi$. In particular, every character $\chi \pmod{q}$ is induced by precisely one primitive character.

If we now think of a character as having domain $\mathbb{Z}$, we can recast some of the properties above in a more concrete form. For example, if $\chi \pmod{q}$ induces / is induced by the character $\chi' \pmod{d}$, then $\chi = \chi'\chi_0$, where χ_0 is the trivial character mod q. This implies that $\chi \pmod{q}$ is primitive iff there is no character $\chi' \pmod{d}$, $d < q$, which agrees with χ whenever $\chi \neq 0$.

Proposition 1.6. *Given χ (mod q) and $\tilde{q} \mid q$ such that*

$$\chi(n_1) = \chi(n_2) \quad \textit{whenever } n_1 \equiv n_2 \, (\textit{mod } \tilde{q}) \textit{ and } (n_1, q) = 1 = (n_2, q). \tag{A.3}$$

Then $\exists \widetilde{\chi}$ (mod $\tilde{q}$) inducing χ.

Proof. For $n \in \mathbb{Z}$ with $(n, \tilde{q}) = 1$ we define

$$\tilde{n} = \begin{cases} n & \text{if } (n, q) = 1 \\ n + R(q, n) \cdot \tilde{q} & \text{otherwise} \end{cases}$$

where

$$R(q, n) = \frac{\text{rad } q}{\text{rad } n} = \prod_{\substack{p \mid q \\ p \nmid n}} p.$$

We make a series of observations, which we prove immediately following:

$$(n, \tilde{q}) = 1 \Longrightarrow (\tilde{n}, q) = 1 \tag{A.4}$$

$$(n, \tilde{q}) = 1 \Longrightarrow n \equiv \tilde{n} \, (\text{mod } \tilde{q}) \tag{A.5}$$

$$n_1 \equiv n_2 \, (\text{mod } \tilde{q}) \text{ and } (n_1, \tilde{q}) = 1 = (n_2, \tilde{q}) \Longrightarrow \widetilde{n_1} \equiv \widetilde{n_2} \, (\text{mod } \tilde{q}) \tag{A.6}$$

$$(n_1, \tilde{q}) = 1 = (n_2, \tilde{q}) \Longrightarrow \widetilde{n_1 n_2} \equiv \widetilde{n_1}\widetilde{n_2} \, (\text{mod } \tilde{q}) \tag{A.7}$$

We prove these in order. If $(n, q) = 1$ then (A.4) holds by definition of $\tilde{n}$, so we may assume $(n, q) \neq 1$. It suffices to show that for every prime factor ℓ of q, ℓ divides exactly one of the two numbers $\{n, \, R(q, n) \cdot \tilde{q}\}$ (because then it would not divide their sum).

- If $\ell \mid n$ then $\ell \nmid \tilde{q}$ by hypothesis and $\ell \nmid R(q, n)$ by definition.
- If $\ell \mid R(q, n)$ then $\ell \nmid n$ by definition.
- If $\ell \mid \tilde{q}$ then $\ell \nmid n$ by hypothesis.

This completes the proof of (A.4). The statement (A.5) is clear, and immediately implies both (A.6) and (A.7).

For $(n, \tilde{q}) = 1$, define $\widetilde{\chi}(n) = \chi(\tilde{n})$. Then $\widetilde{\chi}$ is a character (mod $\tilde{q}$):

- $\widetilde{\chi}(n) = \chi(\tilde{n}) \neq 0$ for $(n, \tilde{q}) = 1$, by (A.4).
- Given $n_1 \equiv n_2 \pmod{\tilde{q}}$ such that $(n_1, \tilde{q}) = 1 = (n_2, \tilde{q})$, we have

$$\begin{aligned} \widetilde{\chi}(n_1) = \chi(\widetilde{n_1}) &= \chi(\widetilde{n_2}) \qquad \text{by (A.6) and (A.3)} \\ &= \widetilde{\chi}(n_2) \end{aligned}$$

- Given $n_1, n_2 \in (\mathbb{Z}/q\mathbb{Z})^*$, we have

$$\begin{aligned} \widetilde{\chi}(n_1 n_2) = \chi(\widetilde{n_1 n_2}) &= \chi(\widetilde{n_1}\widetilde{n_2}) \qquad \text{by (A.7) and (A.3)} \\ &= \chi(\widetilde{n_1})\chi(\widetilde{n_2}) \\ &= \widetilde{\chi}(n_1)\widetilde{\chi}(n_2) \end{aligned}$$

Finally, by definition, for $(n, q) = 1$ we have $\widetilde{\chi}(n) = \chi(\tilde{n}) = \chi(n)$, whence $\widetilde{\chi}$ induces χ. □

Denote the Gauss sum for $\chi \pmod{q}$ by $\tau(\chi) = \sum_{n \leq q} \chi(n)\, e\left(\frac{n}{q}\right)$.

Proposition 1.7. *Given $\chi \pmod{q}$ primitive,*

$$\overline{\chi}(a)\tau(\chi) = \sum_{n \leq q} \chi(n)\, e\left(\frac{an}{q}\right) \qquad \forall a \in \mathbb{Z} \tag{A.8}$$

Proof. If $(a, q) = 1$, (A.8) holds for any $\chi \pmod{q}$, independent of the primitivity of χ. Thus, it suffices to prove that when $(a, q) > 1$,

$$\sum_{n \leq q} \chi(n)\, e\left(\frac{an}{q}\right) = 0.$$

Write $\frac{a}{q} = \frac{a_1}{q_1}$ with $(a_1, q_1) = 1$. Then it suffices to prove

$$\sum_{n \le q} \chi(n)\, e\left(\frac{a_1 n}{q_1}\right) = 0 \qquad \text{for } q_1 \mid q,\ q_1 \ne q, \text{ and } (a_1, q_1) = 1. \tag{A.9}$$

(A.9) holds trivially if $q_1 = 1$, so we may assume $1 < q_1 < q$. Write $q = q_1 q_2$. Then we have

$$\begin{aligned} \sum_{n \le q} \chi(n)\, e\left(\frac{a_1 n}{q_1}\right) &= \sum_{r \,(\mathrm{mod}\ q_1)} \sum_{\substack{n \le q \\ n \equiv r \,(\mathrm{mod}\ q_1)}} \chi(n)\, e\left(\frac{a_1 r}{q_1}\right) \\ &= \sum_{r \,(\mathrm{mod}\ q_1)} S(r)\, e\left(\frac{a_1 r}{q_1}\right) \end{aligned} \tag{A.10}$$

where

$$S(r) = \sum_{\substack{n \le q \\ n \equiv r \,(\mathrm{mod}\ q_1)}} \chi(n)\,.$$

To prove (A.9), and thence the proposition, it suffices to prove $S(r) = 0$ for every r. Suppose this is not the case for some r. Note that for any $m \equiv 1 \,(\mathrm{mod}\ q_1)$ relatively prime to q we have

$$\chi(m)\, S(r) = \sum_{\substack{n \le q \\ n \equiv r \,(\mathrm{mod}\ q_1)}} \chi(mn) = \sum_{\substack{k \le q \\ k \equiv r \,(\mathrm{mod}\ q_1)}} \chi(k) = S(r)$$

upon setting $k = mn$. Since we're assuming $S(r) \ne 0$, we must have

$$\chi(m) = 1 \qquad \forall\, m \equiv 1 \,(\mathrm{mod}\ q_1) \text{ such that } (m, q) = 1. \tag{A.11}$$

But then for any $n_1 \equiv n_2 \,(\mathrm{mod}\ q_1)$ with $(n_1, q) = 1 = (n_2, q)$ we would have $\chi(n_1 \overline{n_2}) = 1$ (where $\overline{n_2}$ denotes the multiplicative inverse of n_2 modulo q), whence $\chi(n_1) = \chi(n_2)$. By Proposition 1.6, χ must be induced by some $\chi_1 \,(\mathrm{mod}\ q_1)$, contradicting our hypothesis that $\chi \,(\mathrm{mod}\ q)$ is primitive. $\square$

Note that we've implicitly proved a variant of Proposition 1.6:

Corollary 1.8. *Given $\chi\,(mod\ q)$ and $\tilde{q} \mid q$ such that $\chi(n) = 1$ for all $n \equiv 1\,(mod\ \tilde{q})$ which are relatively prime to q. Then $\exists\,\widetilde{\chi}\,(mod\ \tilde{q})$ inducing χ.*

We've also proved that primitive characters behave nicely on arithmetic progressions whose constant difference is a proper divisor of q:

Corollary 1.9. *Given $\chi\,(mod\ q)$ primitive and $\tilde{q} \mid q$ such that $\tilde{q} \neq q$. Then $\forall\, r \in \mathbb{Z}$,*

$$\sum_{\substack{n \leq q \\ n \equiv r\,(mod\ \tilde{q})}} \chi(n) = 0.$$

To conclude this brief survey, we prove the orthogonality relations. Fix any integer $q \geq 2$, and set $G = (\mathbb{Z}/q\mathbb{Z})^*$ and $\widehat{G} = \{\chi\,(\text{mod}\ q)\}$. The fundamental theorem of finite abelian groups implies that $G \simeq \widehat{G}$.

Proposition 1.10.

$$\frac{1}{|G|}\sum_{n\in G}\chi(n)^k = \begin{cases} 1 & \textit{if } ord(\chi) \mid k \\ 0 & \textit{otherwise} \end{cases}$$

Proof. For any a coprime to q and for any positive integer k,

$$\chi(a)^k \sum_{n\in G}\chi(n)^k = \sum_{n\in G}\chi(an)^k = \sum_{ninG}\chi(n)^k.$$

This implies that either $\chi(a)^k = 1$ for all a coprime to q, or that $\sum_{n\leq q}\chi(n)^k = 0$; this concludes the proof. ☐

Proposition 1.11.

$$\frac{1}{|\widehat{G}|}\sum_{\chi\in\widehat{G}}\chi(n) = \begin{cases} 1 & \textit{if } n \equiv 1\,(mod\ q) \\ 0 & \textit{otherwise} \end{cases}$$

Proof.

$$\begin{aligned}
\frac{1}{|G|}\sum_{n\in G}\left|\sum_{\chi\in\widehat{G}}\chi(n)\right|^2 &= \frac{1}{|G|}\sum_{n\in G}\sum_{\chi,\psi\in\widehat{G}}\chi\overline{\psi}(n)\\
&= \sum_{\chi,\psi\in\widehat{G}}\frac{1}{|G|}\sum_{n\in G}\chi\overline{\psi}(n)\\
&= |\widehat{G}|
\end{aligned}$$

by the previous proposition, with $k=1$.

Since

$$\frac{1}{|G|}\sum_{n\in G}\left|\sum_{\chi\in\widehat{G}}\chi(1)\right|^2 = |\widehat{G}|,$$

we have

$$\frac{1}{|G|}\sum_{\substack{n\in G\\ n\neq 1}}\left|\sum_{\chi\in\widehat{G}}\chi(n)\right|^2 = 0$$

whence each term of the sum is 0, proving the proposition. $\square$

APPENDIX B

Proof of Theorem 3.3

We follow the original proof of Graham and Ringrose very closely; indeed, we will only explicitly write down those parts of their arguments which must be modified to obtain our version of the result. We refer the reader to sections 3 - 5 of [13]. Set $S := \sum_{M<n\leq M+N} \chi(n)$.

We begin by restating Lemma 3.1 of [13], but skimming off some of the unnecessary hypotheses given there:

Lemma 2.1 (Compare to Lemma 3.1 of [13])**.** *Let $k \geq 0$ be an integer, and set $K := 2^k$. Let $q_0, \ldots, q_k$ be arbitrary positive integers, and let $H_i := N/q_i$ for all i. Then*

$$|S|^{2K} \leq 8^{2K-1} \left(\max_{0\leq j\leq k} \left(N^{2K-K/J} q_j^{K/J} \right) + \frac{N^{2K-1}}{H_0 \cdots H_k} \sum_{h_0 \leq H_0} \cdots \sum_{h_k \leq H_k} |S_k(\mathbf{h})| \right) \tag{B.1}$$

where $J = 2^j$ and $S_k(\mathbf{h})$ satisfies the bound given below.

A bound on $S_k(\mathbf{h})$ is given by (3.4) of [13]:

$$|S_k(\mathbf{h})| \ll N\, Q^{-1} \, |S(Q;\, \chi,\, f_k,\, g_k,\, 0)| + \sum_{0<|s|\leq Q/2} \frac{1}{|s|} \, |S(Q;\, \chi,\, f_k,\, g_k,\, s)| \, . \tag{B.2}$$

See pages 279-280 of [13] for the definitions of f_k, g_k, and $S(Q; \chi, f_k, g_k, s)$.

Let $q := Q/q'$. We have $(q, q') = 1$, whence from Lemma 4.1 of [13] we deduce

$$S(Q; \chi, f_k, g_k, s) = S(q'; \chi', f_k, g_k, s\overline{q})S(q; \eta, f_k, g_k, s\overline{q'})$$

for some primitive characters χ' (mod q') and η (mod q), where $q\overline{q} \equiv 1$ (mod q') and $q'\overline{q'} \equiv 1$ (mod q). By construction, q is squarefree, so Lemmas 4.1-4.3 of [13] apply to give

$$\left|S(q;\, \eta,\, f_k,\, g_k,\, s\overline{q'})\right| \leq d(q)^{k+1} \left(\frac{q}{(q, Q_k)}\right)^{1/2} (q, Q_k, |s\overline{q'}|)$$

where $Q_k := \prod_{i \leq k} h_i q_i$. Combining this with the trivial estimate $|S(q'; \chi', f_k, g_k, s\overline{q})| \leq q'$ yields:

Lemma 2.2 (Compare with Lemma 4.4 of [13])**.** *Keep the notation as above. Then for any positive integers $q_1, \ldots, q_k$,*

$$|S(Q;\, \chi,\, f_k,\, g_k,\, s)| \leq q'\, d(q)^{k+1} \left(\frac{q}{(q, Q_k)}\right)^{1/2} (q, Q_k, |s\overline{q'}|).$$

We shall need the following simple lemma (versions of which appear implicitly in [13]):

Lemma 2.3. *Given q, $\overline{q'}$ be as above; let x and H be arbitrary. Then*

$$\text{1.} \quad \sum_{0<|s|\leq x} \frac{(q,\, |s\overline{q'}|)}{|s|} \ll d(q)\, \log x \tag{B.3}$$

$$\text{2.} \quad \sum_{h \leq H} (q,\, h)^{\frac{1}{2}} \leq d(q)\, H \tag{B.4}$$

Proof.

1. Since $(q, q') = 1$, we have $(q, \overline{q'}) = 1$, whence

$$\sum_{0<|s|\leq x} \frac{(q,\,|s\overline{q'}|)}{|s|} = 2 \sum_{1\leq s\leq x} \frac{(q,\,s\overline{q'})}{s} = 2 \sum_{1\leq s\leq x} \frac{(q,\,s)}{s} = 2 \sum_{n\geq 1} \frac{a_n}{n}$$

where $a_n := \#\{s \leq x : n = \frac{s}{(q,\,s)}\}$. Note that $a_n = 0$ for all $n > x$, and that

$$a_n = \#\{s \leq x : s = (q,\,s)\,n\} \leq \#\{s \leq x : s = dn,\, d \mid q\} \leq d(q).$$

Therefore

$$\sum_{0<|s|\leq x} \frac{(q,\,|s\overline{q'}|)}{|s|} \ll \sum_{n\geq 1} \frac{a_n}{n} \leq d(q) \sum_{n\leq x} \frac{1}{n} \ll d(q) \log x.$$

2. Write

$$\sum_{h\leq H} (q,\,h)^{\frac{1}{2}} = \sum_{n\geq 1} a_n \sqrt{n}$$

where $a_n := \#\{h \leq H : n = (q,\,h)\}$. It is clear that $a_n = 0$ whenever $n \nmid q$. Also, if $(q,\,h) = n$ then $n \mid h$, whence

$$a_n \leq \#\{h \leq H : n \mid h\} \leq \frac{H}{n}.$$

Therefore

$$\sum_{h\leq H} (q,\,h)^{\frac{1}{2}} = \sum_{n\geq 1} a_n \sqrt{n} \leq \sum_{n\mid q} \frac{H}{\sqrt{n}} \leq d(q)\,H.$$

□

Lemma 2.4 (Compare to Lemma 4.5 of [13]). *Keep the notation from above. For any real number $A_0 \geq 1$,*

$$|S|^{4K} \ll 8^{4K-2} \left(A\,A_0^{2K} + B\,A_0^{-2K+1}\,(q')^2 + C\,A_0^{2K-1}\,(q')^2\right)$$

where

$$\begin{aligned} A &= N^{2K} \\ B &= N^{6K-k-4}\, P^{k+1}\, Q\, d(Q)^{2k+4} \log^2 Q \\ C &= N^{2K+k+2}\, Q^{-1}\, d(Q)^{4k+4} \end{aligned}$$

and the implied constant is independent of k.

(Note that in the original paper, there is a persistent typo of writing M rather than N.)

Proof. Following the proof of Lemma 4.5 in [13] and applying (B.3) with $x = Q/2$ yields the following analogue of equation (4.5) from that paper:

$$\sum_{h_k \le H_k} \sum_{0<|s|\le Q/2} \frac{1}{|s|} \, |S(Q;\, \chi,\, f_k,\, g_k,\, s)| \ll q' \sqrt{q}\, d(q)^{k+2}\, H_k\, R_k^{-\frac{1}{2}} \log Q. \tag{B.5}$$

Setting $S_j := h_0 \cdots h_j$, one deduces the following analogue of equation (4.6) of [13]:

$$N\, Q^{-1} \sum_{h_k \le H_k} |S(Q;\, \chi,\, f_k,\, g_k,\, 0)| \le N\, q' \, \frac{\sqrt{q\, R_k}}{Q}\, d(Q)^{k+2}\, H_k \sqrt{(q,\, S_{k-1})}.$$

From (B.4) and the bound $(q,\, S_j) \le (q,\, S_{j-1})\,(q,\, h_j)$ one sees that

$$\sum_{h_0 \le H_0} \cdots \sum_{h_{k-1} \le H_{k-1}} \sqrt{(q,\, S_{k-1})} \le d(q)^k\, H_0 \cdots H_{k-1}. \tag{B.6}$$

Plugging (B.2) into (B.1) and applying (B.5) and (B.6), one obtains:

$$\begin{aligned} |S|^{2K} \ll\ & 8^{2K-1} \max_{0\le j\le k} \left(N^{2K-K/J}\, q_j^{K/J} \right) + \\ & 8^{2K-1}\, q'\, N^{2K-1}\, d(q)^{k+2}\, (\log Q) \sqrt{\frac{q}{R_k}} + \\ & 8^{2K-1}\, q'\, N^{2K}\, d(q)^{2k+2}\, \frac{\sqrt{q}}{Q} \sqrt{R_k}. \end{aligned}$$

Since $q \mid Q$, we have that $q \leq Q$ and $d(q) \leq d(Q)$. Therefore from the above we deduce the following analogue of (4.7) in [13]:

$$
\begin{aligned}
|S|^{2K} \ll \; & 8^{2K-1} \max_{0 \leq j \leq k} \left(N^{2K-K/J}\, q_j^{K/J} \right) + \\
& 8^{2K-1}\, q'\, N^{2K-1}\, d(Q)^{k+2}\, (\log Q) \sqrt{\frac{Q}{R_k}} + \\
& 8^{2K-1}\, q'\, N^{2K}\, d(Q)^{2k+2} \sqrt{\frac{R_k}{Q}}.
\end{aligned}
$$

The rest of the proof given in [13] can now be copied exactly to yield our claim. □

Chasing through the arguments in [13] gives this analogue of Lemma 5.3, which we record for reference:

Lemma 2.5 (Compare to Lemma 5.3 of [13])**.**

$$
\begin{aligned}
|S| \ll \; & N^{1-\frac{k+3}{8K-2}}\, P^{\frac{k+1}{8K-2}}\, Q^{\frac{1}{8K-2}}\, d(Q)^{\frac{k+2}{4K-1}}\, (\log Q)^{\frac{1}{4K-1}}\, (q')^{\frac{1}{4K-1}} + \\
& N^{1-\frac{1}{4K}}\, P^{\frac{k+1}{8K}}\, d(Q)^{\frac{3k+4}{4K}}\, (\log Q)^{\frac{1}{4K}}\, (q')^{\frac{1}{2K}}.
\end{aligned}
$$

Finally, we arrive at:

Proof of Theorem 3.3. Let E_k be the right hand side of the bound claimed in the statement of the theorem. The rest of the proof given in [13] now goes through almost verbatim. □

This concludes the proof of Theorem 3.3. Note that one can extend this to a bound on all non-principal characters by following the argument given directly after Lemma 5.4 in [13]; however, for our applications the narrower result suffices.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Balog, A. Granville, K. Soundararajan, *Multiplicative functions in arithmetic progressions*, to appear.

[2] D. A. Burgess, *The distribution of quadratic residues and non-residues*, Mathematika, **4** (1957), 106-112.

[3] D. A. Burgess, *On character sums and primitive roots*, Proc. Lond. Math. Soc. (3), **12** (1962), 179-192.

[4] D. A. Burgess, *On character sums and L-series*, Proc. Lond. Math. Soc. (3), **12** (1962), 193-206.

[5] D. A. Burgess, *On character sums and L-series II*, Proc. Lond. Math. Soc. (3), **13** (1963), 524-536.

[6] D. A. Burgess, *Estimation of character sums modulo a small power of a prime*, J. Lond. Math. Soc. (2), **30** (1984), *no. 3*, 385-393.

[7] D. A. Burgess, *Estimation of character sums modulo a power of a prime*, Proc. Lond. Math. Soc. (3), **52** (1986), *no. 2*, 215-235.

[8] D. A. Burgess, *The character sum estimate with* $r = 3$, J. Lond. Math. Soc. (2), **33** (1986), *no. 2*, 219-226.

[9] H. Davenport, *Multiplicative Number Theory*, 3rd ed., Graduate Texts in Mathematics, vol. 74, Springer-Verlag, New York, 2000.

[10] P. X. Gallagher, *Primes in progressions to prime-power modulus*, Invent. Math. **16** (1972), 191-201.

[11] L. Goldmakher, *Improved upper bounds on odd order character sums*, in preparation.

[12] L. Goldmakher, *Character sums to smooth moduli are small*, Canad. J. Math., *accepted for publication.*

[13] S. W. Graham and C. J. Ringrose, *Lower Bounds for least quadratic nonresidues*, Analytic number theory (Allerton Park, IL, 1989), Birkhäuser, 1990, pp. 269-309.

[14] A. Granville, K. Soundararajan, *Large Character Sums*, J. Amer. Math. Soc. **14** (2001), 365-397.

[15] A. Granville, K. Soundararajan, *Large Character Sums: Pretentious Characters and the Pólya-Vinogradov Theorem*, J. Amer. Math. Soc. **20** (2007), 357-384.

[16] A. Granville, K. Soundararajan, *Pretentious multiplicative functions and an inequality for the zeta-function*, Anatomy of integers, 191-197, CRM Proc. Lecture Notes 46, Amer. Math. Soc., Providence, RI, 2008.

[17] G. Halász, *On the distribution of additive and mean-values of multiplicative functions*, Studia Sci. Math. Hungar. **6** (1971), 211-233.

[18] G. Halász, *On the distribution of additive arithmetic functions*, Acta Arith. **27** (1975), 143-152.

[19] A. Hildebrand, *A note on Burgess' character sum estimate*, C.R. Math. Rep. Acad. Sci. Canada **8** (1986), 35-37.

[20] A. Hildebrand, *Large values of character sums*, J. Number Theory **29** (1988), 271-296.

[21] H. Iwaniec, *On zeros of Dirichlet's L series*, Invent. math. **23** (1974), 97-104.

[22] H. L. Montgomery, *A note on mean values of multiplicative functions*, Institute Mittag-Leffler, Report No. **17** (1978).

[23] H. L. Montgomery, R. C. Vaughan, *Exponential sums with multiplicative coefficients*, Invent. Math. **43** (1977), 69-82.

[24] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge University Press, 2006.

[25] M. Ram Murty, *Problems in Analytic Number Theory*, Graduate Texts in Mathematics, vol. 206, Springer-Verlag, New York, 2001, Readings in Mathematics.

[26] R. E. A. C. Paley, *A theorem on characters*, J. London Math. Soc. **7** (1932), 28-32.

[27] A. G. Postnikov, *On Dirichlet L-series with the character modulus equal to the power of a prime number*, J. Indian Math. Soc. (N.S.) **20** (1956), 217-226.

[28] G. Pólya, *Über die Verteilung der quadratischen Reste und Nichtreste*, Göttinger nachrichten (1918), 21-29.

[29] I. Schur, *Einige Bemerkungen zu der vorstehenden Arbeit des Herrn G. Pólya: Über die Verteilung der quadratischen Reste und Nichtreste*, Göttinger nachrichten (1918), 30-36.

[30] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge University Press, 1995.

[31] I. M. Vinogradov, *On an asymptotic equality in the theory of quadratic forms* (Russian), J. Phys.-Mat. ob-va Permsk Univ. **1** (1918), 18-28.

[32] I. M. Vinogradov, *Sur la distribution des résidus et des nonrésidus des puissances* (French), J. Phys.-Mat. ob-va Permsk Univ. **1** (1918), 94-98.

[33] I. M. Vinogradov, *On the distribution of quadratic residues and nonresidues* (Russian), J. Phys.-Mat. ob-va Permsk Univ. **2** (1919), 1-16.