HARMONIC ANALYSIS ON THE POSITIVE RATIONALS: MULTIPLICATIVE FUNCTIONS AND EXCEPTIONAL DIRICHLET CHARACTERS

by

JONATHAN KISH

B.A., Carleton College, 1999

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics

2013

This thesis entitled:
Harmonic Analysis on the Positive Rationals: Multiplicative Functions and Exceptional
Dirichlet Characters
written by Jonathan Kish
has been approved for the Department of Mathematics

_____

Peter D. T. A. Elliott

_____

Katherine E. Stange

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both
the content and the form meet acceptable presentation standards of scholarly work in the
above mentioned discipline.

Kish, Jonathan (Ph.D., Mathematics)

Harmonic Analysis on the Positive Rationals: Multiplicative Functions and Exceptional
    Dirichlet Characters

Thesis directed by Prof. Peter D. T. A. Elliott

An improved estimate for multiplicative functions on arithmetic progressions is demonstrated, at the expense of potentially a uniformly bounded number of sums involving such functions braided with Dirichlet characters being separated for particular attention. An introduction to new methods for classifying these characters, which we call exceptional, is offered in the Conclusion.

DEDICATION

To Laura, for persevering through the perversity,

and to Mr. Tines, for making the world a smaller place.

ACKNOWLEDGEMENTS

CONTENTS

CHAPTER

APPENDIX

CHAPTER 1

NOTATION AND DEFINITIONS

The symbols n and N will always denote positive integers, p a positive prime integer, x and y real numbers. An *arithmetic function*, generally denoted by f or g, will be a complex-valued function defined on the positive integers. An arithmetic function is *real* if it is real-valued. A *multiplicative function*, generally denoted by g, will be an arithmetic function which satisfies $g(ab) = g(a)g(b)$ whenever a and b are coprime, that is, whenever $(a, b) = 1$. A *completely multiplicative function* satisfies $g(ab) = g(a)g(b)$ for *all* pairs of positive integers a and b.

$\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}, \mathbb{N}$ the rational, real and complex number fields, the ring of (rational) integers, the set of positive (rational) integers

$m \mid n$ denotes that the integer m divides the integer n

$s = \sigma + it$ (or $i\tau$)   a complex variable, with $\sigma = \operatorname{Re}(s)$ the real part of s

(a, b) the greatest common divisor of the integers a and b

$[x]$ the largest integer not exceeding x

$[m_1, \ldots, m_n]$ the least common multiple of the integers $m_1, \ldots, m_n$

$\omega(n)$ the number of distinct prime divisors of the integer n, $\omega(1) = 0$

$\mu(n)$ Möbius' (arithmetic) function,

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^{\omega(n)} & \text{if } n > 1 \text{ is squarefree} \\ 0 & \text{otherwise} \end{cases}$$

$\Lambda(n)$ Von Mangoldt's (arithmetic) function,

$$\Lambda(n) = \begin{cases} \log p & \text{if } n \text{ is a power of a prime} \\ 0 & \text{otherwise} \end{cases}$$

$\varphi(n)$ Euler's (arithmetic) function, the order of the group of reduced residue classes modulo n

χ, χ_j Dirichlet characters to a modulus $D \geq 1$

exceptional

$|f| \leq 1$ an abbreviation for $|f(n)| \leq 1$ for all $n \in \mathbb{N}$

$f * g$ the Dirichlet convolution of the arithmetic functions f and g,

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d) = \sum_{uv=n} f(u)g(v).$$

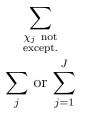|	
f^{-1}	the inverse of f with respect to Dirichlet convolution, which exists and is unique as long as $f(1) \neq 0$ (see Theorem 2.8 of [1]).
$O(\cdot), \ll$	$a(x) = O(b(x))$ and $a(x) \ll b(x)$ both denote that $|a(x)| \leq Cb(x)$, for some constant C, holds uniformly on some specified set of x-values. $O(b(x))$ denotes a function a that satisfies $a(x) \ll b(x)$
$O_\varepsilon(\cdot), \ll_\varepsilon$	both mean essentially the same thing as $O(\cdot)$ and $\ll$, except that the implied constant C depends upon ε
$\displaystyle\sum_{\chi_j}$	a sum over all $\varphi(D)$ Dirichlet characters modulo D
$\displaystyle\sum_{\chi_j \text{ except.}}$	a sum over only exceptional Dirichlet characters
$\displaystyle\sum_{\substack{\chi_j \text{ not} \\ \text{except.}}}$	a sum over all nonexceptional Dirichlet characters
$\displaystyle\sum_{j} \text{ or } \sum_{j=1}^{J}$	a sum over some (or J) Dirichlet characters, which ones being (temporally) insignificant
B_0	$\displaystyle\max_{\substack{x \leq N, |\tau| \leq T, \chi_j, \\ \sigma = 1 + (\log t)^{-1} \\ 2 \leq t \leq N}} |G(s, \chi_j)|$ (p. 36)
$E_0, \widetilde{E}_0$	see pp. 13 and 29, respectively
F_1, F_2, F_3	see p. 27

$$G(s) \qquad \sum_{n=1}^{\infty} g(n)n^{-s}, \quad \text{the Dirichlet series corresponding to } g$$

$$G(s,\chi) \qquad \sum_{n=1}^{\infty} g(n)\chi(n)n^{-s}, \quad \text{the Dirichlet series corresponding to } g\chi$$

H

$$L \qquad \sum_{D<p\leq x} \frac{1}{p} \quad \text{(p. 22)}$$

$L_1,\ L_2$

$\mathcal{L}(a)$

$$\mathcal{M}(x) = \mathcal{M}(x,g) \qquad \sum_{n\leq x} g(n)$$

$$\widetilde{\mathcal{M}}(x) = \widetilde{\mathcal{M}}(x,g) \qquad \sum_{\substack{n\leq x \\ (n,Q)=1}} g(n)$$

$$\mathcal{N}(x) = \mathcal{N}(x,g) \qquad \sum_{n\leq x} g(n)\log n$$

$$\widetilde{\mathcal{N}}(x) = \widetilde{\mathcal{N}}(x,g) \qquad \sum_{\substack{n\leq x \\ (n,Q)=1}} g(n)\log n$$

Q will generally denote a positive integer

$$Q_c \qquad \prod_{p\leq D^c} p \quad \text{(p. 28)}$$

$S,\ S_1,\ S_2$

$$\mathcal{S}_j \qquad\qquad \text{see p. 41}$$

$$\Sigma_1, \Sigma_2 \qquad\qquad \text{see p. 51 and p. 52, resp.}$$

$$Y(f,a,x) \qquad \sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, D)}} f(n) - \sum_{\chi_j \text{ except.}} \frac{\overline{\chi_j}(a)}{\varphi(D)} \sum_{n \leq x} f(n)\chi_j(n) \qquad (\text{p. 41})$$

Note that, for $\sigma > 1$, the Dirichlet series $G(s)$ and $G(s,\chi)$ define analytic functions of s and have Euler product representations in this half-plane (see Lemma 2.13 of [5, p. 95]). Moreover,

$$G'(s,\chi) = \sum_{n=1}^{\infty} g(n)\chi(n)(\log n)n^{-s},$$

which is related to $\mathcal{N}(x, g\chi_j)$. If $1 \leq u < 2$, then $\widetilde{\mathcal{N}}(u)$ (and $\mathcal{N}(u)$) are zero since for such u's there is only one term in the sum corresponding to $n = 1$, and this term is zero due to the logarithm.

I assign empty sums the value 0, whereas empty products are assigned the value 1.

Throughout this paper, the labeling 'Theorem $X.Y$' refers to the Yth Theorem in Chapter X, for instance, whereas '$(X.Y)$' refers to the Yth equation in Chapter X. On occasion I felt that a more detailed explanation was needed, but in order to not interrupt the continuity of the exposition I put the details in the Appendix. In that case X is A.

I will use Linnik's convention with constants (as opposed to Landau's thorough renumbering), in the sense that a letter representing an arbitrary constant need not have the same value at each occurrence. There will be at most finitely many changes for each constant.

CHAPTER 2

INTRODUCTION

We begin at least as far back as Euclid's *Elements* (ca. 300 BCE), wherein it is proved that there are infinitely many prime numbers. Following in the footsteps of Euler's subsequent analytic proof of the infinitude of primes, Dirichlet's 1837 Theorem on primes in arithmetic progressions, specifically that there are infinitely primes p congruent to a modulo D whenever a and D are coprime integers, is another influential milestone that we will return to in the Conclusion.

Erdős and Selberg's 1949 elementary proof of the Prime Number Theorem concerning the distribution of the primes – originally proved by Hadamard and de la Vallée Poussin in 1896 – forms an important part of the foundation of the current paper. Their work illustrated that not only was analytic continuation not essential to the proof of the Prime Number Theorem, but also that complex analytic methods (including the involvement of the Riemann zeta function) in their entirety were not necessary.

There is another result that deserves special mention: in 1911 Landau [24] showed that

$$x^{-1} \sum_{n \leq x} \mu(n) \to 0, \qquad x \to \infty,$$

is equivalent to the Prime Number Theorem, and that one can be derived from the other using elementary arguments (see Chapter 19 in [6] and the exercises in Chapter 15 of [14] for a thorough treatment of these topics). This showed conclusively that the (limiting) mean value of at least one multiplicative function is deeply significant in number theory. One might wonder if multiplicative functions on arithmetic progressions are similarly useful.

There is an enormous literature on primes in arithmetic progressions (see [25], for instance), but any of the classical methods will not apply to a *general* multiplicative function g on an arithmetic progression. It took until the middle of the twentieth century to begin a systematic study of the (limiting) mean values of general multiplicative functions in earnest. An early result is due to Delange [3], which categorizes those multiplicative functions g with $|g| \leq 1$ for which the mean value exists and is *nonzero*. Notice this result is not equivalent to the Prime Number Theorem in light of the above mean value of the Möbius function. Delange's results are powerful enough, however, to provide an alternative proof of a theorem of Erdős and Wintner concerning the limiting behavior of certain frequencies of additive arithmetic functions (although such matters will not concern us here). The case when the mean value of an arbitrary multiplicative function exists and is *zero* proved more difficult.

In the 1960's Eduard Wirsing proved results, regarding the mean values of multiplicative functions, that are as deep as the Prime Number Theorem (see [27, 28]) in the sense that the latter can be deduced from the former by elementary means. His papers have ideas in common with Selberg's aforementioned proof. The limitation of his method, however, was that his functions were essentially real-valued. This limitation was addressed in Halász's 1968 paper [20], which provides a taxonomy of complex-valued multiplicative functions g, satisfying $|g| \leq 1$, in terms of their mean values and corresponding Dirichlet series. I will use methods adopted from Wirsing and Halász, for example, that of introducing a logarithm and that of factoring G' as $G \cdot G'/G$, respectively. Using a logarithm to study primes appears in the work of Chebyshev, of course, but it seems as though Wirsing was the first to take advantage of this in a general way. See Chapter 6 in [5] for a thorough treatment of the relevant work of Delange, Wirsing and Halász. Unfortunately multiplicative functions on arithmetic progressions still needed attention.

The foregoing methods, amongst many original ones, were used by Elliott (see [8, 9, 10, 11, 12, 13, 16]) to study general multiplicative functions with values in the complex unit disc on arithmetic progressions. In particular, the denouement of this series of papers is

Theorem 1 in [16]:

Theorem 2.1 (ELLIOTT). *Let D be an integer, $2 \leq D \leq x$, $\varepsilon > 0$. Let g be a multiplicative function with values in the complex unit disc.*

There is a character $\chi_1 \,(mod\,D)$, real if g is real, such that when $0 < \gamma < 1$,

$$\sum_{\substack{n \leq y \\ n \equiv a \,(mod\,D)}} g(n) - \frac{1}{\phi(D)} \sum_{\substack{n \leq y \\ (n,D)=1}} g(n) - \frac{\overline{\chi_1(a)}}{\phi(D)} \sum_{n \leq y} g(n)\chi_1(n) \ll \frac{y}{\phi(D)} \left(\frac{\log D}{\log y} \right)^{1/4 - \varepsilon}$$

uniformly for $(a, D) = 1$, $D \leq y$, $x^\gamma \leq y \leq x$, the implied constant depending at most upon ε, γ.

An earlier version of this Theorem was used to obtain deep results on primes in arithmetic progressions (see [13, p. 202]), whereas a related, weaker version of Theorem 2.1 was used (see [15]) as the foundation of a new proof of Linnik's celebrated Theorem: that for some constant C, every reduced residue class $(mod\,D)$ contains a prime representative not exceeding D^C. The proof does not involve the use of estimates for the density of zeros of Dirichlet L-series (and hence of the Deuring-Heilbronn phenomenon), analytic continuation, or nontrivial zero-free regions of the associated L-series. In the above language, Elliott's results are at least as deep as Linnik's Theorem since the latter may be deduced from the former.

The main Theorem that I intend to prove is a modification of Theorem 2.1, namely:

Theorem 2.2. *Let D be an integer, $(\log x)^{\varepsilon_0} \leq D \leq x$ for some $\varepsilon_0 > 0$, and let α be a real number, $0 < \alpha < 1$. Let g be a multiplicative function with values in the complex unit disc. Then there are Dirichlet characters χ_j, which we call exceptional, such that when $0 < \gamma < 1$,*

$$\sum_{\substack{n \leq y \\ n \equiv a \,(mod\,D)}} g(n) - \sum_{\chi_j \ except.} \frac{\overline{\chi_j(a)}}{\varphi(D)} \sum_{n \leq y} g(n)\chi_j(n) \ll_{\alpha,\gamma} \frac{y}{D} \left(\frac{\log D}{\log y} \right)^{1-\alpha}$$

uniformly for $(a, D) = 1$, $D \leq y$, and $x^\gamma \leq y \leq x$.

The Dirichlet characters considered *exceptional* in the following account will be those that are close to g in an appropriate metric.

A typical metric, on equivalence classes of multiplicative functions, is given by

$$\rho(g, h) = \left(\sum_p p^{-\sigma} |g(p) - h(p)|^2 \right)^{1/2}$$

for a suitably chosen value of $\sigma > 1$.

For example, in Corollary 3.11 of my account a family of such metrics, parameterized by $\sigma = 1 + (\log t)^{-1}$, $2 \le t \le N$, is implicit (see p. 24). See also Chapter 12 in [14].

For simplicity of exposition I have assumed that $D > (\log x)^{\varepsilon_0}$ for some $\varepsilon_0 > 0$. Any modulus D less than this small power of a logarithm can be handled using results from [8, 11]. The main difficulty in the present circumstances is when D is "large" compared to x, that is, up to some power (less than one) of x.

The improvement from $\frac{1}{4}$ to 1 in the exponent is at the expense of having to remove potentially more characters. This is made precise in Corollary 3.11. The improvement in the denominator, from $\varphi(D)$ to D, is shown in Chapter 6.

Since the various uniformities employed are important, I have taken great care with the details. Although the general line of attack follows that of [13, 16], serious modifications are made at appropriate points to unify, simplify and extend the method(s).

Until further notice, g will denote a *completely* multiplicative function with values in the complex unit disc that is not identically zero. I will sometimes further restrict g to require that it vanish on "small" primes (those not exceeding D^c, where c is a real constant to be specified later), but that will be clear from the context. Moreover, since g is completely multiplicative now (and so determined by its values on the primes), $g(n) = 0$ for any n divisible by a prime $p \le D^c$. In this regard we may sum over n's for which $p \mid n \implies p > D^c$. Yet another way to think of this constraint is $g(n) = 0$ if $(n, Q_c) > 1$, where

$$Q_c = \prod_{p \le D^c} p.$$

This will help simplify some notation later. Note that this constraint does not alter (complete) multiplicativity, in the sense that any (completely) multiplicative g will still be (completely) multiplicative subject to this tighter restriction. It is also technically convenient to assume that $g(p) = 0$ for all $p > x$ (this is *not* the same as $g(n) = 0$ for $n > x$, of course). In a sense this is a vacuous requirement since we never sum values of g for inputs that are more than x, but this will simplify certain arguments (the proof of Corollary 3.11, for instance).

I will first prove a version of Theorem 2.2 with g a completely multiplicative function that vanishes on the primes not exceeding D^c (see (5.14)). In Chapter 6, I remove these restrictions from g and show that the upper bound obtained in the modified version of Theorem 2.2 continues to hold. This will prove the Theorem in its full generality, as stated above.

CHAPTER 3

PRELIMINARY RESULTS

I will use the following standard results, collected into a single Lemma and including some vintage Chebyshev (1851/1852) as well as a Theorem of Mertens (1874), without explicitly mentioning them. Proofs can be found in [23], Theorems 414 and 415 (p. 453), Theorems 424 and 425 (p. 462), Theorems 427 and 429 (p. 466), Theorem 62 (p. 64) and Theorem 296 (p. 333) in that order; see also "Notes" on p. 497 for historical remarks and references.

Lemma 3.1. *The following estimates hold for $Y \geq 2$:*

$$\sum_{p \leq Y} \log p \ll \sum_{d \leq Y} \Lambda(d) \ll Y,$$

$$\sum_{d \leq Y} \frac{\Lambda(d)}{d} = \sum_{p \leq Y} \frac{\log p}{p} + O(1) = \log Y + O(1),$$

$$\sum_{p \leq Y} \frac{1}{p} = \log \log Y + O(1),$$

$$\prod_{p \leq Y} \left(1 - \frac{1}{p}\right) \ll \frac{1}{\log Y}$$

Moreover, the representations

$$\varphi(n) = n \prod_{p \mid n} \left(1 - \frac{1}{p}\right) \qquad and \qquad \log n = \sum_{d \mid n} \Lambda(d)$$

hold for all $n \geq 1$.

Another useful result we shall need is the following:

Lemma 3.2 (DUALITY PRINCIPLE). *Let c_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, be mn complex numbers. Let λ be a real number. Then the inequality*

$$\sum_{i=1}^{m} \left| \sum_{j=1}^{n} c_{ij} a_j \right|^2 \leq \lambda \sum_{j=1}^{n} |a_j|^2$$

is valid for all complex numbers $a_1, \ldots, a_n$ if and only if the inequality

$$\sum_{j=1}^{n} \left| \sum_{i=1}^{m} c_{ij} b_i \right|^2 \leq \lambda \sum_{i=1}^{m} |b_i|^2$$

is valid for all complex numbers $b_1, \ldots, b_m$.

Proof of Lemma 3.2. This is Lemma 4.3 of [5, p. 150]; c.f. Lemma 3.3 of [14, p. 30] and [4]. The proof is an exercise in the use of the Cauchy-Schwarz inequality.

For nonnegative real numbers u and any sequence of complex numbers a_n, $n \geq 1$, define the function

$$A(u) = \sum_{n \leq u} a_n.$$

If $F(s)$ formally (i.e., not worrying about convergence for the moment) denotes the Dirichlet series $\sum_{n=1}^{\infty} a_n n^{-s}$ corresponding to the a_n's, then integrating by parts we obtain

$$F(s) = s \int_{1^-}^{\infty} y^{-s-1} A(y) \, dy$$

for $\sigma > 1$ and hence

$$s^{-1} F(s) = \int_{0}^{\infty} A(e^w) e^{-w\sigma} \cdot e^{-iw\tau} \, dw$$

after the change of variable $y = e^w$. It follows that $s^{-1}F(s)$ as a function of τ and $A(e^w)e^{-w\sigma}$ as a function of w are Fourier transforms. By Plancherel's Theorem we obtain the following result.

Lemma 3.3.
$$\int_{-\infty}^{\infty} \left| \frac{F(s)}{s} \right|^2 d\tau = 2\pi \int_{1}^{\infty} \frac{|A(y)|^2}{y^{2\sigma+1}} \, dy$$

provided one of the integrals exists in the appropriate L^2 sense.

Proof of Lemma 3.3. See [5, p. 228] and Lemma 10 of [13, p. 188].

Note that $s^{-1}G(s)$ and $s^{-1}G'(s, \chi)$ (see Chapter 1 for notation) are $O(|\tau|^{-1})$, as $|\tau| \to \infty$, and so belong to the Lebesgue class $L^2(\mathbb{R})$ when $\sigma > 1$. For an example in the application of Lemma 3.3, see the proof of Lemma 4.1.

Integration over τ is understood as integration along the line $\mathrm{Re}\,(s) = \sigma$.

Lemma 3.4. *Let $g(n)$ be a multiplicative function, $|g(n)| \leq 1$ for all $n \in \mathbb{N}$. For each prime p define*

$$h(p) = \sum_{k=1}^{\infty} g(p^k)p^{-ks}. \tag{3.1}$$

Then there is a representation

$$G(s) = (1 + h(2)) \exp \left(\sum_{p \geq 3} g(p)p^{-s} \right) G_1(s),$$

valid in the half-plane $\sigma > 1$. Moreover, $G_1(s)$ is analytic in the half-plane $\sigma > 1/2$, and is bounded by

$$e^{-5} \leq |G_1(s)| \leq e^5 \quad and \quad |G_1'(s)| \leq e^{11}$$

in the half-plane $\sigma \geq 1$.

Proof of Lemma 3.4. This is Lemma 6.6 of [5, p. 230]; cf. Lemma 6.2 of [26, p. 339] as well.

Lemma 3.5. *Let $r > 0$, $y = w - w(\log w)^{-r}$. Then*

$$\mathcal{N}(w) \ll w \int_2^w \frac{|\mathcal{N}(u)|}{u^2 \log u}\, du + \sum_{d \leq (\log w)^{2r}} \frac{d\Lambda(d)|g(d)|}{w - y} \int_{y/d}^{w/d} |\mathcal{M}(u)|\, du + E_0,$$

where

$$E_0 = E_0(g) = \max_{\substack{u,v \\ y \leq u \leq v \leq w}} \left| \sum_{u < n \leq v} g(n) \log n \right|$$

holds uniformly for $w \geq 2$, for all completely multiplicative functions g.

Proof of Lemma 3.5. This is Lemma 1 of [11, p. 205]. See also Lemma 19.3 in [6, p. 213], where Elliott remarks that "[t]he essential ingredient of this result ... is that it relates $\mathcal{N}(w)$ to a weighted average of itself, which may be more easy to deal with."

We now proceed with the proof. Uniformly for $y \leq u \leq w$,

$$|\mathcal{N}(w) - \mathcal{N}(u)| = \left| \sum_{n \leq w} g(n) \log n - \sum_{n \leq u} g(n) \log n \right| = \left| \sum_{u < n \leq w} g(n) \log n \right| \leq E_0 \qquad (3.2)$$

since summing over $u < n \leq w$ is among the possibilities covered in the maximum defining E_0. Therefore

$$\int_y^w \{\mathcal{N}(w) - \mathcal{N}(u)\} \, du = (w - y)\mathcal{N}(w) - \int_y^w \mathcal{N}(u) \, du$$

which implies that

$$\left| (w - y)\mathcal{N}(w) - \int_y^w \mathcal{N}(u) \, du \right| = \left| \int_y^w \{\mathcal{N}(w) - \mathcal{N}(u)\} \, du \right|$$

$$\leq \int_y^w |\mathcal{N}(w) - \mathcal{N}(u)| \, du$$

$$= (w - y)E_0,$$

by (3.2); here we use uniformity. This inequality is certainly true if $w = y$ (for then everything is zero), so assuming that $w \neq y$ (which forces $w > y$ since $w \geq y$ from the start) we obtain

$$\frac{1}{w - y} \left| (w - y)\mathcal{N}(w) - \int_y^w \mathcal{N}(u) \, du \right| \leq E_0$$

or

$$\left| \mathcal{N}(w) - \frac{1}{w - y} \int_y^w \mathcal{N}(u) \, du \right| \leq E_0$$

since $w - y > 0$, so it may be pulled through the modulus without issue. This last identity is the same as

$$\mathcal{N}(w) = \frac{1}{w - y} \int_y^w \mathcal{N}(u) \, du + O(E_0), \qquad (3.3)$$

with an implied constant actually equal to 1, this will be ultimately of no consequence. Recall that g is completely multiplicative. Using the representation $\log n = \sum_{d|n} \Lambda(d)$, we

see that

$$\mathcal{N}(u) = \sum_{n \leq u} g(n) \log n$$

$$= \sum_{n \leq u} g(n) \sum_{d \mid n} \Lambda(d)$$

$$= \sum_{d \leq u} \Lambda(d) \sum_{md \leq u} g(md)$$

$$= \sum_{d \leq u} \Lambda(d) g(d) \sum_{m \leq u/d} g(m)$$

$$= \sum_{d \leq u} \Lambda(d) g(d) \mathcal{M}(u/d).$$

Substituting this into (3.3), we see that

$$\mathcal{N}(w) = \int_y^w \sum_{d \leq u} \Lambda(d) g(d) \mathcal{M}(u/d) \, du + O(E_0). \tag{3.4}$$

The contribution to the integral in (3.4) from those terms with $d \leq (\log w)^{2r}$ does not exceed

$$\sum_{d \leq (\log w)^{2r}} \Lambda(d) |g(d)| \frac{1}{w-y} \int_y^w \left| \mathcal{M}\left(\frac{u}{d}\right) \right| \, du = \sum_{d \leq (\log w)^{2r}} \frac{d\Lambda(d) |g(d)|}{w-y} \int_{y/d}^{w/d} |\mathcal{M}(u)| \, du \tag{3.5}$$

in absolute value. Here we have used the fact that the d's no longer depend on u, justifying the interchange of the sum and the second integral in (3.5), and the change of variable $u \mapsto td$ in the integral. This gives the second term in the bound of the Lemma.

Next, over the range $1 \leq u \leq v = w(\log w)^{-2r}$, we note that

$$\sum_{y/u < p \leq w/u} p \log p \leq \frac{w}{u} \log \frac{w}{u} \sum_{y/u < p \leq w/u} 1 \ll \left(\frac{w}{u} \log \frac{w}{u}\right) \cdot \left(\frac{(w-y)/u}{\log((w-y)/u)}\right)$$

$$= \frac{w(w-y)}{u^2} \cdot \frac{\log(w/u)}{\log((w-y)/u)}, \tag{3.6}$$

where we used an old sieve estimate regarding the number of primes in an interval (see "Remarks" [14, p. 226] or Theorem 3.7 [21, p. 107]; an estimate of this kind may also be found in [22]; an adequate version may be found in Lemma 3.12 at the end of this Chapter).

For this range of u,

$$w \geq u(\log w)^{2r} \implies wu^{-1} \leq \{u^{-1}w(\log w)^{-r}\}^2 = \{u^{-1}(w - y)\}^2 \tag{3.7}$$

since, from the definition of y, $w - y = w(\log w)^{-r}$. Note that the $(\log w)^{-2r}$ is a typo in the last line of [11, p.205]. It should be as in (3.7). It follows that

$$\log(wu^{-1}) \leq 2\log\left(u^{-1}(w - y)\right),$$

and substituting this into (3.6) gives

$$\sum_{y/u < p \leq w/u} p \log p \leq \frac{w(w - y)}{u^2} \cdot \frac{2\log\left(u^{-1}(w - y)\right)}{\log\left((w - y)/u\right)}$$

$$\ll u^{-2}w(w - y). \tag{3.8}$$

Moreover,

$$\sum_{\substack{p^k \leq w/u \\ k \geq 2}} p^k \log p \ll (w/u)^{3/2} \ll u^{-2}w(w - y) \tag{3.9}$$

since

$$\sum_{\substack{p^k \leq y \\ k \geq 2}} p^k \log p \leq y \sum_{\substack{p^k \leq y \\ k \geq 2}} \log p = y\left(\sum_{\substack{p^2 \leq y \\ (k=2)}} \log p + \sum_{\substack{p^3 \leq y \\ (k=3)}} \log p + \sum_{\substack{p^4 \leq y \\ (k=4)}} \log p + \cdots\right)$$

$$= y\left(\sum_{p \leq \sqrt{y}} \log p + \sum_{p \leq \sqrt[3]{y}} \log p + \sum_{p \leq \sqrt[4]{y}} \log p + \cdots\right)$$

$$\leq y\left(\sum_{p \leq \sqrt{y}} \log p + \frac{\log y}{\log 2} \sum_{p \leq \sqrt[3]{y}} \log p\right)$$

$$\ll y\left(y^{1/2} + \frac{\log y}{\log 2} \cdot y^{1/3}\right)$$

$$\ll y^{3/2},$$

where we have used the fact that in the sum

$$\sum_{p \le \sqrt[3]{y}} \log p + \sum_{p \le \sqrt[4]{y}} \log p + \cdots$$

there are no more than $\frac{\log y}{\log 2}$ terms, none of which is larger than the first. We have also used Lemma 3.1 and the fact that $\log y \ll y^{1/6}$ to ensure that $\frac{\log y}{\log 2} \cdot y^{1/3} \ll y^{1/2}$. Note also that the condition $u \le w(\log w)^{-2r}$ implies that

$$u^{1/2} \le \frac{w^{1/2}}{(\log w)^r} \implies (w/u)^{3/2} \le \frac{w^2}{u^2(\log w)^r} \implies (w/u)^{3/2} \le u^{-2}w(w-y).$$

The contribution to the integral in (3.4) which arises from terms with $d > (\log w)^{2r}$ then satisfies

$$\left| \frac{1}{w-y} \int_y^w \sum_{(\log w)^{2r} < d \le w} \Lambda(d)g(d)\mathcal{M}(u/d)\, du \right| \le \sum_{(\log w)^{2r} < d \le w} \Lambda(d)|g(d)| \frac{1}{w-y} \int_y^w \left| \mathcal{M}\left(\frac{u}{d}\right) \right|\, du$$

$$\le \sum_{(\log w)^{2r} < d \le w} \Lambda(d) \frac{1}{w-y} \int_y^w \left| \mathcal{M}\left(\frac{u}{d}\right) \right|\, du = \sum_{(\log w)^{2r} < d \le w} d\Lambda(d) \frac{1}{w-y} \int_{y/d}^{w/d} |\mathcal{M}(t)|\, dt$$

$$\le \int_1^v |\mathcal{M}(t)| \cdot \frac{1}{w-y} \sum_{y/t < d \le w/t} d\Lambda(d)\, dt \ll \int_1^v |\mathcal{M}(t)| \cdot \frac{1}{w-y} \cdot t^{-2}w(w-y)\, dt$$

$$= w \int_1^v \frac{|\mathcal{M}(t)|}{t^2}\, dt, \tag{3.10}$$

where we used $|g| \le 1$, the substitution $t = \frac{u}{d}$, the fact that $\mathcal{M}(t)$ is zero for $t < 1$ and

$$\sum_{y/t < d \le w/t} d\Lambda(d) = \sum_{\substack{y/t < p^k \le w/t \\ k \ge 1}} p^k \log p \le \sum_{y/t < p \le w/t} p \log p + \sum_{\substack{p^k \le w/t \\ k \ge 2}} p^k \log p$$

combined with the estimates (3.8) and (3.9). Integration by parts shows that for $t \ge 2$,

$$\mathcal{M}(t) = \int_{2^-}^t \frac{1}{\log z}\, d\mathcal{N}(z) = \frac{\mathcal{N}(t)}{\log t} + \int_2^t \frac{\mathcal{N}(z)}{z(\log z)^2}\, dz$$

since $\mathcal{N}(z) = 0$ for $z < 2$, and therefore

$$\frac{|\mathcal{M}(t)|}{t^2} \le \frac{|\mathcal{N}(t)|}{t^2 \log t} + \frac{1}{t^2} \int_2^t \frac{|\mathcal{N}(z)|}{z(\log z)^2}\, dz. \tag{3.11}$$

Note also that

$$\int_1^2 \frac{|\mathcal{M}(t)|}{t^2}\,dt \leq \int_1^2 \frac{2}{t^2}\,dt = 1, \tag{3.12}$$

since there are at most two terms in the sum defining $\mathcal{M}(t)$ when $1 \leq t \leq 2$, neither of which is larger than 1 in modulus. From (3.11) and (3.12), we obtain

$$\int_1^v \frac{|\mathcal{M}(t)|}{t^2}\,dt \leq \int_1^w \frac{|\mathcal{M}(t)|}{t^2}\,dt \leq 1 + \int_2^w \frac{|\mathcal{M}(t)|}{t^2}\,dt$$

$$\ll \int_2^w \left\{ \frac{|\mathcal{N}(t)|}{t^2 \log t} + \frac{1}{t^2} \int_2^t \frac{|\mathcal{N}(z)|}{z(\log z)^2}\,dz \right\} dt$$

$$= \int_2^w \frac{|\mathcal{N}(t)|}{t^2 \log t}\,dt + \iint_{\substack{2 \leq z \leq t \\ 2 \leq t \leq w}} \frac{|\mathcal{N}(z)|}{t^2 z (\log z)^2}\,dz\,dt. \tag{3.13}$$

The double integral can be written

$$\int_2^w \frac{|\mathcal{N}(z)|}{z(\log z)^2} \left(\int_z^w \frac{dt}{t^2} \right) dz \leq \int_2^w \frac{|\mathcal{N}(z)|}{(z \log z)^2}\,dz$$

since

$$\int_z^w \frac{dt}{t^2} \leq \frac{1}{z}.$$

However,

$$\int_2^w \frac{|\mathcal{N}(z)|}{(z \log z)^2}\,dz \leq \int_2^w \frac{|\mathcal{N}(t)|}{t^2 \log t}\,dt,$$

so going back to (3.13) we see that

$$\int_1^v \frac{|\mathcal{M}(t)|}{t^2}\,dt \ll \int_2^w \frac{|\mathcal{N}(t)|}{t^2 \log t}\,dt = \int_1^w \frac{|\mathcal{N}(t)|}{t^2 \log t}\,dt$$

since $\mathcal{N}(t) = 0$ for $1 \leq t < 2$. Substituting this into (3.10), combining the resulting inequality with (3.5) and appealing to (3.4) completes the proof of Lemma 3.5.

The next inequality, of Maximal Gap Large Sieve type, will be used to control the error term E_0 in the previous Lemma amongst other things. For convenience we provide a useful Corollary.

Lemma 3.6. *Let $0 < \varepsilon < 1$. The inequality*

$$\sum_{j=1}^{J} \max_{v-u \leq H} \left| \sum_{\substack{u < n \leq v \\ (n,Q)=1}} a_n \chi_j(n) \right|^2 \ll_\varepsilon \left(H \prod_{\substack{p|Q \\ p \leq H}} \left(1 - \frac{1}{p}\right) + JH^\varepsilon D^{\frac{1}{2}} \log D \right) \sum_{n=1}^{\infty} |a_n|^2$$

where the χ_j are distinct Dirichlet characters $(\bmod\, D)$, $D \geq 2$, Q a positive integer, $H \geq 0$, holds for all square-summable complex numbers a_n.

Proof of Lemma 3.6 (c.f. Lemma 7, 9 or 11 of [13] and the proof of Lemma 3 in [11, p. 207]). This is Lemma 3 of [18]. With $0 \leq v_j - u_j \leq H$, define

$$t_j(n) = \begin{cases} \chi_j(n) & \text{if } u_j < n \leq v_j, \\ 0 & \text{otherwise,} \end{cases} \qquad j = 1, \ldots, J.$$

For any real λ_d, $d \mid Q$, constrained by $\lambda_1 = 1$, the dual form

$$S = \sum_{\substack{n \leq x \\ (n,Q)=1}} \left| \sum_{j=1}^{J} c_j t_j(n) \right|^2$$

does not exceed

$$\sum_{n \leq x} \left(\sum_{d|(n,Q)} \lambda_d \right)^2 \left| \sum_{j=1}^{J} c_j t_j(n) \right|^2 = \sum_{d_i|Q} \lambda_{d_1} \lambda_{d_2} \sum_{j,k=1}^{J} c_j \bar{c}_k \sum_{n \equiv 0 \,(\bmod\, [d_1,d_2])} t_j(n) \overline{t_k(n)}, \qquad (3.14)$$

where $[d_1, d_2]$ denotes the least common multiple of d_1 and d_2. For those terms with $j \neq k$, the innermost sum has the form

$$\chi_j \bar{\chi}_k ([d_1, d_2]) \sum_m \chi_j \bar{\chi}_k(m)$$

with the integers m over the intersection of two intervals and is, by a classical result of Pólya and Vinogradov (see for example [2, Chapter 23]), $O\left(D^{1/2} \log D\right)$ since $\chi_j \overline{\chi_k}$ is not

the principal character when $j \neq k$. The corresponding contribution to (3.14) satisfies

$$\left| \sum_{d_i|Q} \lambda_{d_1}\lambda_{d_2} \sum_{\substack{j,k=1 \\ j \neq k}}^{J} c_j \bar{c}_k \sum_{n \equiv 0 \,(\text{mod}\,[d_1,d_2])} t_j(n)\overline{t_k(n)} \right|$$

$$\leq \sum_{d_i|Q} |\lambda_{d_1}\lambda_{d_2}| \sum_{\substack{j,k=1 \\ j \neq k}}^{J} |c_j \bar{c}_k| \left| \chi_j \overline{\chi_k}([d_1,d_2]) \sum_m \chi_j \overline{\chi}_k(m) \right|$$

$$\ll D^{1/2} \log D \sum_{d_i|Q} |\lambda_{d_1}\lambda_{d_2}| \sum_{\substack{j,k=1 \\ j \neq k}}^{J} |c_j \bar{c}_k|. \qquad (3.15)$$

Now,

$$\sum_{\substack{j,k=1 \\ j \neq k}}^{J} |c_j \bar{c}_k| \leq \sum_{\substack{j,k=1 \\ j \neq k}}^{J} \frac{1}{2}(|c_j|^2 + |c_k|^2)$$

$$= \frac{1}{2} \sum_{j=1}^{J} |c_j|^2 \sum_{\substack{k=1 \\ k \neq j}}^{J} 1 + \frac{1}{2} \sum_{k=1}^{J} |c_k|^2 \sum_{\substack{j=1 \\ j \neq k}}^{J} 1$$

$$= (J-1) \sum_{j=1}^{J} |c_j|^2$$

by symmetry. Substituting this into (3.15), we see that the contribution to (3.14) arising from those terms with $j \neq k$ is

$$\ll JD^{1/2} \log D \left(\sum_{d|Q} |\lambda_d| \right)^2 \sum_{j=1}^{J} |c_j|^2. \qquad (3.16)$$

For those terms with $j = k$ we reform the square in the λ_d to gain a contribution

$$\sum_{j=1}^{J} |c_j|^2 \sum_{n \leq x} \left(\sum_{d|(n,Q)} \lambda_d \right)^2 |t_j(n)|^2. \qquad (3.17)$$

Since $|t_j(n)| \leq 1$, the innersum over n does not exceed

$$\sum_{d_i|Q} \lambda_{d_1}\lambda_{d_2} \sum_{\substack{u_j < n \leq v_j+H \\ n \equiv 0 \,(\text{mod}\,[d_1,d_2])}} 1 = H \sum_{d_i|Q} \lambda_{d_1}\lambda_{d_2}[d_1,d_2]^{-1} + O\left(\left(\sum_{d|Q} |\lambda_d| \right)^2 \right). \qquad (3.18)$$

We may follow a standard appeal to the method of Selberg (see Lemma 3.12) with $\lambda_d = 0$ if $d > H^{\varepsilon/2}$ which, in particular, gives $|\lambda_d| \leq 1$ for all remaining λ_d. As a consequence,

$$\left(\sum_{d|Q} |\lambda_d| \right)^2 \ll H^\varepsilon$$

since there are no more than $H^{\varepsilon/2}$ terms in the sum. Thus (3.16) is

$$\ll J H^\varepsilon D^{1/2} \log D \sum_{j=1}^{J} |c_j|^2. \tag{3.19}$$

The O term in (3.18) is also $\ll H^\varepsilon$, and we absorb this into preceding estimate. Moreover, as may be seen in Lemma 3.12, the function λ_d may be chosen so that the quadratic form in (3.18) satisfies

$$\sum_{\substack{d_i|Q \\ d_i \leq z}} \lambda_{d_1} \lambda_{d_2} [d_1, d_2]^{-1} \ll \prod_{\substack{p|Q \\ p \leq z}} \left(1 - \frac{1}{p} \right).$$

Utilizing this in (3.18) with $z = H^{\varepsilon/2}$, combining it with (3.19) and (3.17) and substituting everything back into (3.14) gives

$$S \ll \left(H \prod_{\substack{p|Q \\ p \leq H}} \left(1 - \frac{1}{p} \right) + J H^\varepsilon D^{1/2} \log D \right) \sum_{j=1}^{J} |c_j|^2.$$

Dualizing (see Lemma 3.2) yields the inequality of Lemma 3.6. See Remark A.5 for more detail about the sieve method used.

Corollary 3.7. *Let $0 < \varepsilon < 1$. The inequality*

$$\sum_{j=1}^{J} \left| \sum_{p \leq x} a_p \chi_j(p) \right|^2 \ll_\varepsilon \left(\frac{x}{\log x} + x^\varepsilon D^{\frac{3}{2}} \log D \right) \sum_{p \leq x} |a_p|^2$$

where the χ_j are distinct Dirichlet characters $(\bmod\, D)$, $D \geq 2$, holds for all square-summable complex numbers a_p.

Proof of Corollary 3.7. Apply Lemma 3.6 with $H = x$ and $Q = \displaystyle\prod_{p \leq x^\varepsilon} p$.

I shall employ the main Theorem from [17]. A related particular result was privately circulated at the American Mathematical Society's Mathematics Research Communities *The Pretentious View of Analytic Number Theory* meeting in Snowbird, Utah, which I attended in the summer of 2011. The following more general Theorem 3.8 has a different proof.

Theorem 3.8. *For each positive real B there is a real C such that*

$$\sum_{j=1}^{J} \max_{y \leq x} \max_{\sigma \geq 1, |t| \leq D^B} \left| \sum_{D < p \leq y} a_p \chi_j(p) p^{-s} \right|^2 \leq (4L + (J-1)C) \sum_{D < p \leq x} |a_p|^2 p^{-1},$$

with $s = \sigma + it$, $L = \sum_{D < p \leq x} p^{-1}$, uniformly for a_p in $\mathbb{C}$ and distinct Dirichlet characters χ_j (mod D), $1 \leq D \leq x$.

For convenience I reproduce a proof here. A detailed discussion of inequalities of this type can be found in the aforementioned paper and the relevant references. Theorem 3.8 (as well as Theorems 7.1 and 7.2 in the Conclusion) rely on the vital Lemma 3.9 of Elliott, which I now state, omitting the assertion concerning the principal character to which we shall not appeal, and to which we did not appeal in [17, 18].

Lemma 3.9. *Given $B > 0$,*

$$Re \sum_{w < p \leq y} \chi(p) p^{-s}$$

is bounded above in terms of B alone, uniformly for $s = \sigma + it$, $\sigma \geq 1$, $|t| \leq D^B$, $y \geq w \geq D$ and all non-principal characters $\chi \, (mod \, D)$, $D \geq 1$.

Proof of Lemma 3.9. For two proofs using complex-analytic properties of Dirichlet L-series (differing only in the degree to which analytic continuation is used), see Lemma 1 in [16] and then Lemmas 1, 14 in [13]; an elementary proof via Selberg's sieve (and which yields a better dependence upon B) is given in [15]; the case $\sigma = 1$ by continuity.

Proof Theorem 3.8. Since the sum $\sum_{D < p \leq x} |a_p| p^{-\sigma}$ approaches zero as $\sigma \to \infty$, the innermost maximum may be taken over a bounded rectangle. In view of the uniformity in y, Abel summation allows us to restrict to the case $\sigma = 1$.

For reals t_j, y_j, $|t_j| \leq D^B$, $D < y_j \leq x$, define

$$\delta_{j,p} = \begin{cases} \chi_j(p)p^{-\frac{1}{2}-it_j} & \text{if } D < p \leq y_j, \\ \\ 0 & \text{otherwise}, \end{cases}$$

$j = 1, \ldots, J$, and consider the inequality

$$\sum_{D<p\leq x} \left| \sum_{j=1}^{J} b_j \delta_{j,p} \right|^2 \leq \Delta \sum_{j=1}^{J} |b_j|^2,$$

where the b_j are for the moment real and nonnegative. The expanded sum is

$$\sum_{j=1}^{J} b_j^2 L + 2 \sum_{1\leq j<\ell\leq J} b_j b_\ell \operatorname{Re} \sum_{D<p\leq x} \chi_j \overline{\chi}_\ell(p) p^{-1-it_j+it_\ell}.$$

An appeal to Lemma 3.9 followed by an application of the Cauchy-Schwarz inequality shows that we may take $\Delta = L + (J-1)C_1$ for a certain C_1 depending at most upon B.

If now b_j is complex, we represent it as a sum

$$\max(\operatorname{Re} b_j, 0) + \min(\operatorname{Re} b_j, 0) + i \max(\operatorname{Im} b_j, 0) + i \min(\operatorname{Im} b_j, 0)$$

and correspondingly partition the innersum over j. Since the coefficients in each subsum all have the same argument, a second application of the Cauchy-Schwarz inequality allows us to conclude that with $\Delta = 4\left(L + (J-1)C_1\right)$ the above inequality holds for all complex b_j.

Dualizing (see Lemma 3.2),

$$\sum_{j=1}^{J} \left| \sum_{D<p\leq x} a_p \delta_{j,p} \right|^2 \leq 4(L + (J-1)C_1) \sum_{D<p\leq x} |a_p|^2$$

for all complex a_p. Replacing a_p by $a_p p^{-\frac{1}{2}}$ completes the proof of Theorem 3.8.

At this point it is helpful to note the following application of Theorem 3.8, which is an upper bound on $|G(s,\chi)|$, the Dirichlet series attached to the multiplicative function g braided with the Dirichlet character χ. This will aid in understanding the quantity B_0 (see p. 36), constructed as we remove the exceptional characters during the proof of Lemma 4.1.

Lemma 3.10. *Given ε, $0 < \varepsilon < 1$, the estimate*

$$\max_{\substack{x \leq N, |\tau| \leq D^B \\ \sigma = 1 + (\log t)^{-1} \\ 2 \leq t \leq N}} \left| \sum_{D^c < p \leq x} a_p \chi_j(p) p^{-s} \right| \leq \varepsilon \sum_{D^c < p \leq N} \frac{1}{p}$$

holds uniformly for all $a_p \in \mathbb{C}$ with $|a_p| \leq 1$ and all Dirichlet characters $\chi_j \, (\mathrm{mod}\, D)$, $D \geq 1$, with the possible exception of finitely many characters depending on ε.

Proof of Lemma 3.10. Suppose that, for some fixed ε, $0 < \varepsilon < 1$, and some $\chi_j \, (\mathrm{mod}\, D)$,

$$\max_{\substack{x \leq N, |\tau| \leq D^B \\ \sigma = 1 + (\log t)^{-1} \\ 2 \leq t \leq N}} \left| \sum_{D^c < p \leq x} a_p \chi_j(p) p^{-s} \right| > \varepsilon \sum_{D^c < p \leq N} \frac{1}{p} = \varepsilon L, \tag{3.20}$$

where L is now defined in terms of N instead of x. Squaring both sides of the previous inequality and summing over the distinct Dirichlet characters $\chi_j \, (\mathrm{mod}\, D)$, $j = 1, \ldots, J$, for which it holds, we obtain

$$\sum_{j=1}^{J} \max_{\substack{x \leq N, |\tau| \leq D^B \\ \sigma = 1 + (\log t)^{-1} \\ 2 \leq t \leq N}} \left| \sum_{D^c < p \leq x} a_p \chi_j(p) p^{-s} \right|^2 > \sum_{j=1}^{J} (\varepsilon L)^2 = J \varepsilon^2 L^2.$$

Appealing to Theorem 3.8, however, this would imply that

$$J \varepsilon^2 L^2 < (4L + (J-1)C)L \leq 4L^2 + CJL \qquad \text{or} \qquad J < \frac{4}{\varepsilon^2} + \frac{CJ}{\varepsilon^2 L}$$

with a constant C that depends at most upon B. Suppose that

$$J \geq \left[\frac{6}{\varepsilon^2} \right] \geq \frac{6}{\varepsilon^2} - 1 > \frac{5}{\varepsilon^2}$$

(which is valid if and only if $\frac{1}{\varepsilon^2} > 1$, or, $0 < \varepsilon < 1$). Then

$$\frac{5}{\varepsilon^2} < J < \frac{4}{\varepsilon^2} + \frac{CJ}{\varepsilon^2 L} \implies L < CJ \leq \frac{6C}{\varepsilon^2}.$$

This last inequality can certainly be falsified in light of Lemma 3.1, however, since

$$L = \log\left(\frac{\log N}{\log D^c} \right) + O(1) \tag{3.21}$$

and we are free to restrict the size of D in comparison to N. Hence there are at most $\left[\frac{6}{\varepsilon^2} \right]$ characters for which (3.20) holds, completing the proof.

Corollary 3.11. *Given any ε, $0 < \varepsilon < 1$, the estimate*

$$\max_{\substack{x \leq N, |\tau| \leq D^B \\ \sigma = 1 + (\log t)^{-1} \\ 2 \leq t \leq N}} |G(s, \chi)| \ll_{\varepsilon} \left(\frac{\log N}{\log D} \right)^{\varepsilon}$$

holds uniformly for all Dirichlet characters $\chi \,(\mathrm{mod}\, D)$ and completely multiplicative functions g with $|g| \leq 1$, with the exception of at most a finite number of characters depending on ε.

Proof of Corollary 3.11. Let χ be any Dirichlet character modulo D. Since g and χ are completely multiplicative and $|g\chi| \leq 1$, using an Euler product representation and a standard estimate for the principal value of the logarithm,

$$|\log(1 + z) - z| \leq |z|^2,$$

valid for $z \in \mathbb{C}$ with $|z| \leq \frac{1}{2}$, using $z = -\frac{g(p)\chi(p)}{p^s}$ we have

$$G(s, \chi) = \exp \left\{ \log \prod_{D^c < p \leq x} \left(1 - \frac{g(p)\chi(p)}{p^s} \right)^{-1} \right\} = \exp \left\{ -\sum_{D^c < p \leq x} \log \left(1 - \frac{g(p)\chi(p)}{p^s} \right) \right\}$$

$$= \exp \left\{ \sum_{D^c < p \leq x} \left(\frac{g(p)\chi(p)}{p^s} + O\left(\frac{1}{p^{2s}} \right) \right) \right\} = \exp \left(\sum_{D^c < p \leq x} \frac{g(p)\chi(p)}{p^s} + O(1) \right).$$

Thus

$$G(s, \chi) \ll \exp \left(\mathrm{Re} \sum_{D^c < p \leq x} \frac{g(p)\chi(p)}{p^s} \right) \tag{3.22}$$

since $|e^z| = e^{\mathrm{Re}\,(z)}$ for all $z \in \mathbb{C}$. Given an ε, $0 < \varepsilon < 1$, since $\mathrm{Re}\,(z) \leq |z|$ for any $z \in \mathbb{C}$, by Lemma 3.10, it follows that

$$\mathrm{Re} \sum_{D^c < p \leq x} \frac{g(p)\chi(p)}{p^s} \leq \left| \sum_{D^c < p \leq x} \frac{g(p)\chi(p)}{p^s} \right| \leq \varepsilon L$$

for all but possibly $O_{\varepsilon}(1)$ characters $\chi \,(\mathrm{mod}\, D)$. In light of (3.21) and (3.22),

$$G(s, \chi) \ll \exp(\varepsilon L) \ll_{\varepsilon} \left(\frac{\log N}{\log D} \right)^{\varepsilon},$$

for all but possibly $O_{\varepsilon}(1)$ characters, having used $c \geq 1$. This completes the proof. We note for the record that there is a corresponding lower bound.

A Dirichlet character modulo D will be called *exceptional* (with respect to ε, x, N) if the estimate in Corollary 3.11 fails. For practical purposes the significant parameter is ε.

For reference I close this Chapter with a sieve result that is more than adequate for our purposes.

Lemma 3.12. *Let $f(n)$ be a real-valued, nonnegative arithmetic function. Let a_n, $n = 1, \ldots, N$ be a sequence of rational integers. Let r be a positive real number, and let $p_1 < p_2 < \cdots p_s \leq r$ be rational primes. Set $Q = p_1 \cdots p_s$. If $d \mid Q$ then let*

$$\sum_{\substack{n=1 \\ a_n \equiv 0 \, (mod \, d)}} f(n) = \eta(d)X + R(n,d),$$

where X, R are real numbers, $X \geq 0$, and $\eta(d_1 d_2) = \eta(d_1)\eta(d_2)$ whenever d_1 and d_2 are coprime divisors of Q.

Assume that for each prime p, $0 \leq \eta(p) < 1$.

Let $I(N,Q)$ denote the sum

$$\sum_{\substack{n=1 \\ (a_n,Q)=1}}^{N} f(n).$$

Then the estimate

$$I(N,Q) = (1 + 2\theta_1 H)X \prod_{p|Q}(1 + \eta(p)) + 2\theta_2 \sum_{\substack{d|Q \\ d \leq z^3}} 3^{\omega(d)}|R(N,d)|$$

holds uniformly for $r \geq 2$, $\max(\log r, S) \leq \frac{1}{8}\log z$, where $|\theta_1| \leq 1$, $|\theta_2| \leq 1$, and

$$H = \exp\left(-\frac{\log z}{\log r}\left\{\log\left(\frac{\log z}{S}\right) - \log\log\left(\frac{\log z}{S}\right) - \frac{2S}{\log z}\right\}\right)$$

$$S = \sum_{p|Q}\frac{\eta(p)}{1 - \eta(p)}\log p.$$

When these conditions are satisfied there is a positive absolute constant c so that $2H \leq c < 1$.

Proof of Lemma 3.12. This is Lemma 2.1 of [5, p. 79].

CHAPTER 4

MAIN RESULTS

The main result of this Chapter is the following Lemma.

Lemma 4.1. *Let $0 < \gamma < 1$, $0 < \delta < 1$ and $2 \leq \log N \leq D^c \leq N$. If g is a completely multiplicative function that vanishes on the primes not exceeding D^c, then upon removing $O_\delta(1)$ exceptional characters modulo D, the remaining characters satisfy*

$$\sum_{\substack{\chi_j \ not \\ except.}} \max_{2 \leq y \leq t} \left| \sum_{n \leq y} g(n)\chi_j(n) \right|^2 \ll_\delta \left(\frac{t}{\log t} \right)^2 \left(\frac{\log x}{\log D} \right)^{2\delta}$$

uniformly for $D^c \leq t \leq x$ and uniformly for $N^\gamma \leq x \leq N$.

Note that for x in the range $N^\gamma \leq x \leq N$, $\frac{\log x}{\log N}$ is between two constants. This simplifies the formulation of Theorem 2.2.

The following result, in the notation of Chapter 1, prepares us for the application of harmonic analysis.

Lemma 4.2. *If $x \geq 2$, then*

$$\sum_j \max_{2 \leq w \leq x} |\widetilde{\mathcal{N}}(w, g\chi_j)|^2 \ll F_1 + F_2 + F_3 \tag{4.1}$$

where

$$F_1 = x^2 \sum_j \left(\int_2^x \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u} \, du \right)^2$$

$$F_2 = \sum_j \max_{2 \le w \le x} \left(\left(\sum_{\substack{d \le (\log w)^{2r} \\ (d,Q)=1}} \frac{\Lambda(d)}{d} \right) \cdot \left(\sum_{\substack{d \le (\log w)^{2r} \\ (d,Q)=1}} d^3 \Lambda(d) \left\{ \frac{1}{w-y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right\}^2 \right) \right)$$

$$F_3 = \sum_j \max_{\substack{u,v \\ v-u \le x(\log x)^{-r}}} \left| \sum_{\substack{u < n \le v \\ (n,Q)=1}} g(n)\chi_j(n) \log n \right|^2$$

and the summations may be taken (identically) over any collection of the characters $\chi_j \pmod D$.

Moreover, if

$$Q_c = \prod_{\substack{q \ prime \\ q \le D^c}} q, \qquad c \ge 1, \tag{4.2}$$

and $g(n) = 0$ unless $(n, Q_c) = 1$, then with $0 < \varepsilon < 1$,

$$\sum_j \max_{2 \le w \le x} |\widetilde{\mathcal{N}}(w, g\chi_j)|^2 \ll_\varepsilon x^2 \sum_j \left(\int_{D^c}^x \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u} \, du \right)^2 + \left(\frac{x}{\log D} \right)^2 \left(\log \left(\frac{\log x}{\log D} \right) \right)^2. \tag{4.3}$$

Proof of Lemma 4.2. We proceed along the lines of Lemma 8 in [13, p. 186] with modifications important to the present circumstances. For some constant C, say, and for any j with $1 \le j \le \phi(D)$, by Lemma 3.5 we have

$$|\mathcal{N}(w, g\chi_j)| \le C \left\{ w \int_1^w \frac{|\mathcal{N}(u, g\chi_j)|}{u^2 \log u} \, du + \sum_{d \le (\log w)^{2r}} \frac{d\Lambda(d)|g(d)\chi_j(d)|}{w-y} \int_{y/d}^{w/d} |\mathcal{M}(u, g\chi_j)| \, du \right.$$

$$\left. + E_0(g\chi_j) \right\}.$$

Modifying this result to take the condition $(n, Q) = 1$ into account yields

$$|\widetilde{\mathcal{N}}(w, g\chi_j)| \leq C \left\{ w \int_1^w \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u} \, du + \sum_{\substack{d \leq (\log w)^{2r} \\ (d,Q)=1}} \frac{d\Lambda(d)|g(d)\chi_j(d)|}{w - y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right. $$

$$\left. + \widetilde{E}_0(g\chi_j) \right\},$$

where we set

$$\widetilde{E}_0(g) = \max_{\substack{u,v \\ y \leq u \leq v \leq w}} \left| \sum_{\substack{u < n \leq v \\ (n,Q)=1}} g(n) \log n \right|.$$

Then $|\widetilde{\mathcal{N}}(w, g\chi_j)|^2$ does not exceed

$$\leq C^2 \left\{ w \int_1^w \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u} \, du + \sum_{\substack{d \leq (\log w)^{2r} \\ (d,Q)=1}} \frac{d\Lambda(d)|g(d)\chi_j(d)|}{w - y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du + \widetilde{E}_0(g\chi_j) \right\}^2$$

$$\leq 3C^2 \left\{ \left(w \int_1^w \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u} \, du \right)^2 + \left(\sum_{\substack{d \leq (\log w)^{2r} \\ (d,Q)=1}} \frac{d\Lambda(d)|g(d)\chi_j(d)|}{w - y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right)^2 \right.$$

$$\hfill (4.4)$$

$$\left. + (\widetilde{E}_0(g\chi_j))^2 \right\}$$

after an application of the Cauchy-Schwarz inequality. Taking the maximum over $2 \leq w \leq x$ on both sides of (4.4) and summing over the j characters, the first term on the righthand side yields F_1 since

$$\max_{2 \leq w \leq x} \left(w \int_1^w \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u} \, du \right)^2 = \left(x \int_2^x \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u} \, du \right)^2,$$

where $\widetilde{\mathcal{N}}(u, g\chi_j) = 0$ for $1 \leq u < 2$.

The middle term in (4.4) satisfies

$$
\left(\sum_{\substack{d \leq (\log w)^{2r} \\ (d,Q)=1}} \frac{d\Lambda(d)|g(d)\chi_j(d)|}{w-y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right)^2
$$

$$
= \left(\sum_{\substack{d \leq (\log w)^{2r} \\ (d,Q)=1}} \left(\frac{\Lambda(d)}{d} \right)^{1/2} \cdot \frac{(d^3\Lambda(d))^{1/2}|g(d)\chi_j(d)|}{w-y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right)^2
$$

$$
\leq \left(\sum_{\substack{d \leq (\log w)^{2r} \\ (d,Q)=1}} \left\{ \left(\frac{\Lambda(d)}{d} \right)^{1/2} \right\}^2 \right) \left(\sum_{\substack{d \leq (\log w)^{2r} \\ (d,Q)=1}} \left\{ \frac{(d^3\Lambda(d))^{1/2}|g(d)\chi_j(d)|}{w-y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right\}^2 \right)
$$

$$
\leq \left(\sum_{\substack{d \leq (\log w)^{2r} \\ (d,Q)=1}} \frac{\Lambda(d)}{d} \right) \cdot \left(\sum_{\substack{d \leq (\log w)^{2r} \\ (d,Q)=1}} d^3\Lambda(d) \left\{ \frac{1}{w-y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right\}^2 \right)
$$

the first inequality from (another application of) the Cauchy-Schwarz inequality, and the second from the fact that $|g\chi_j| \leq 1$ for any j. Taking the maximum over $2 \leq w \leq x$ and summing over j gives F_2.

As for F_3, note that the final term in (4.4) satisfies

$$
\sum_j \max_{2 \leq w \leq x} (\widetilde{E}_0(g\chi_j))^2 = \sum_j \max_{2 \leq w \leq x} \left(\max_{\substack{u,v \\ y \leq u \leq v \leq w}} \left| \sum_{\substack{u < n \leq v \\ (n,Q)=1}} g(n)\chi_j(n) \log n \right| \right)^2
$$

$$
\leq \sum_j \max_{2 \leq w \leq x} \max_{\substack{u,v \\ y \leq u \leq v \leq w}} \left| \sum_{\substack{u < n \leq v \\ (n,Q)=1}} g(n)\chi_j(n) \log n \right|^2
$$

$$
\leq \sum_j \max_{\substack{u,v \\ v-u \leq x(\log x)^{-r}}} \left| \sum_{\substack{u < n \leq v \\ (n,Q)=1}} g(n)\chi_j(n) \log n \right|^2 .
$$

Together these estimates establish (4.1).

To prove the second part of the Lemma, fix $r = 2$ and $Q = Q_c$. Then the first factor in F_2 satisfies

$$\sum_{\substack{d \leq (\log w)^4 \\ (d, Q_c) = 1}} \frac{\Lambda(d)}{d} \ll \log\left(\frac{\log w}{\log D}\right) \tag{4.5}$$

by (4.2) and Lemma A.1.

For the second factor in F_2, note that

$$\frac{1}{w-y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \leq \frac{1}{d} \cdot \max_{y/d \leq u \leq w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)|,$$

and

$$\max_{2 \leq w \leq x} \left\{ \frac{1}{w-y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right\}^2 \leq \max_{2 \leq w \leq x} \left\{ \frac{1}{d} \max_{y/d \leq u \leq w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \right\}^2$$

$$\leq \frac{1}{d^2} \max_{\beta - \alpha \leq x/d} \left| \sum_{\substack{\alpha < n \leq \beta \\ (n, Q_c) = 1}} g(n)\chi_j(n) \right|^2. \tag{4.6}$$

It follows by applying (4.5) and (4.6) that

$$F_2 \leq \sum_j \max_{2 \leq w \leq x} \left(\sum_{\substack{d \leq (\log w)^4 \\ (d, Q_c) = 1}} \frac{\Lambda(d)}{d} \right) \cdot \max_{2 \leq w \leq x} \left(\sum_{\substack{d \leq (\log w)^4 \\ (d, Q_c) = 1}} d^3 \Lambda(d) \left\{ \frac{1}{w-y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right\}^2 \right)$$

$$\ll \max_{2 \leq w \leq x} \log\left(\frac{\log w}{\log D}\right) \cdot \sum_{\substack{d \leq (\log x)^4 \\ (d, Q_c) = 1}} d^3 \Lambda(d) \cdot \sum_j \max_{2 \leq w \leq x} \left\{ \frac{1}{w-y} \int_{y/d}^{w/d} |\widetilde{\mathcal{M}}(u, g\chi_j)| \, du \right\}^2$$

$$\ll \log\left(\frac{\log x}{\log D}\right) \cdot \sum_{\substack{d \leq (\log x)^4 \\ (d, Q_c) = 1}} d^3 \Lambda(d) \cdot \frac{1}{d^2} \sum_j \max_{\beta - \alpha \leq x/d} \left| \sum_{\substack{\alpha < n \leq \beta \\ (n, Q_c) = 1}} g(n)\chi_j(n) \right|^2. \tag{4.7}$$

By Lemma 3.6, after canceling the d's the second factor above is

$$\ll_\varepsilon \sum_{\substack{d \le (\log x)^4 \\ (d,Q_c)=1}} d\Lambda(d) \left\{ \frac{x}{d} \prod_{\substack{p|Q_c \\ p \le x/d}} \left(1 - \frac{1}{p}\right) + J\left(\frac{x}{d}\right)^\varepsilon D^{1/2} \log D \right\} \sum_{\substack{n \le x/d \\ (n,Q_c)=1}} |g(n)|^2$$

$$\ll \sum_{\substack{d \le (\log x)^4 \\ (d,Q_c)=1}} d\Lambda(d) \left\{ \frac{1}{\log D} \cdot \frac{x^2}{d^2} \prod_{\substack{p|Q_c \\ p \le x/d}} \left(1 - \frac{1}{p}\right) + \frac{1}{\log D} \left(\frac{x}{d}\right)^{1+\varepsilon} D^{3/2} \log D \right\}$$

$$= \frac{x^2}{\log D} \sum_{\substack{d \le (\log x)^4 \\ (d,Q_c)=1}} \frac{\Lambda(d)}{d} \prod_{\substack{p|Q_c \\ p \le x/d}} \left(1 - \frac{1}{p}\right) + x^{1+\varepsilon} D^{3/2} \sum_{\substack{d \le (\log x)^4 \\ (d,Q_c)=1}} \Lambda(d) d^{-\varepsilon}$$

$$\ll \left(\frac{x}{\log D}\right)^2 \log\left(\frac{\log x}{\log D}\right) + x^{1+\varepsilon} D^{3/2} (\log x)^4$$

(see Remark A.2). Combining this with (4.7) we obtain

$$F_2 \ll_\varepsilon \left(\frac{x}{\log D}\right)^2 \left(\log\left(\frac{\log x}{\log D}\right)\right)^2 + x^{1+\varepsilon} D^{3/2} (\log x)^4 \log\left(\frac{\log x}{\log D}\right)$$

$$\ll \left(\frac{x}{\log D}\right)^2 \left(\log\left(\frac{\log x}{\log D}\right)\right)^2. \tag{4.8}$$

As for F_3, we use Lemma 3.6 again. With $r = 2$,

$$F_3 \ll_\varepsilon \left(x(\log x)^{-2} \prod_{\substack{p|Q_c \\ p \le x(\log x)^{-2}}} \left(1 - \frac{1}{p}\right) + J\left(x(\log x)^{-2}\right)^\varepsilon D^{1/2} \log D \right) \sum_{\substack{n \le x \\ (n,Q_c)=1}} |g(n)\chi_j(n) \log n|^2$$

$$\ll \left(x(\log x)^{-2} \prod_{p|Q_c} \left(1 - \frac{1}{p}\right) + \left(x(\log x)^{-2}\right)^\varepsilon D^{3/2} \log D \right) \cdot \frac{x(\log x)^2}{\log D}$$

$$= \left(\frac{x}{\log D}\right)^2 + x^{1+\varepsilon} (\log x)^{2(1-\varepsilon)} D^{3/2}$$

(see Remark A.3), which is certainly smaller than the bound for F_2 in (4.8). This establishes (4.3) and completes the proof of Lemma 4.2.

We proceed to the proof of the main result of this Chapter, which will be crucial in the next. In application of Lemma 4.2, bounding what corresponded to F_1 is the hardest part, to which we turn now.

Proof of Lemma 4.1 (c.f. Lemma 12 in [13, p. 189] and the 'Main Lemma' of [11, p. 209]). The sum that we wish to bound is

$$\sum_{\substack{\chi_j \text{ not} \\ \text{except.}}} \max_{2 \leq y \leq x} |\widetilde{\mathcal{M}}(y, g\chi_j)|^2$$

with $Q = Q_c$ (which we will assume for the duration of this proof unless otherwise specified). In order to achieve this bound, we introduce a logarithm and apply Lemma 4.2 to the related sum

$$\sum_{\chi_j} \max_{2 \leq y \leq x} |\widetilde{\mathcal{N}}(y, g\chi_j)|^2,$$

removing the exceptional characters at an important step and stripping out the logarithm at the end. Define

$$S(t) = \sum_{\chi_j} \max_{2 \leq y \leq t} |\widetilde{\mathcal{N}}(y, g\chi_j)|^2, \qquad 2 \leq t \leq N.$$

By Lemma 4.2, an upper bound on $S(t)$ boils down to an estimation of

$$t^2 \sum_{\chi_j} \left(\int_{D^c}^{t} \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u} \, du \right)^2.$$

Let θ be a real number, $0 < \theta < 1$ (we will fix it at an absolute value shortly), and let

$$S_1(t) = t^2 \sum_{\chi_j} \left(\int_{D^c}^{t^\theta} \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u} \, du \right)^2$$

and $S_2(t)$ the similar expression with the range of integration changed to $t^\theta < u \leq t$. Then $S(t) = S_1(t) + S_2(t)$ (recall that $g(n) = 0$ unless $(n, Q_c) = 1$, and any $n \leq D^c$ is certainly divisible by a prime not exceeding D^c. Of course, g may be zero for n's larger than D^c). We handle $S_2(t)$ first.

By the Cauchy-Schwarz inequality,

$$\left(\int_{t^\theta}^t \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^2 \log u}\, du\right)^2 = \left(\int_{t^\theta}^t \frac{1}{u^{1/2}\log u} \cdot \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^{3/2}}\, du\right)^2$$

$$\leq \left(\int_{t^\theta}^t \frac{du}{u(\log u)^2}\right)\left(\int_{t^\theta}^t \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|^2}{u^3}\, du\right)$$

for any j, thus

$$S_2(t) \leq t^2 \left(\int_{t^\theta}^t \frac{du}{u(\log u)^2}\right) \sum_{\chi_j} \int_{t^\theta}^t \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|^2}{u^3}\, du.$$

Let $\sigma = 1 + (\log t)^{-1}$. Over the range $1 \leq u \leq t$ (which includes $t^\theta < u \leq t$), $u^\sigma \leq eu$ and thus $u^{-3} \leq e^2 u^{-2\sigma - 1}$ (see Remark A.6). Hence

$$S_2(t) \leq t^2 \cdot \frac{1}{\theta \log t} \sum_{\chi_j} \int_{t^\theta}^t \frac{e^2 |\widetilde{\mathcal{N}}(u, g\chi_j)|^2}{u^{2\sigma + 1}}\, du$$

$$\leq \frac{(et)^2}{\theta \log t} \sum_{\chi_j} \int_1^\infty \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|^2}{u^{2\sigma + 1}}\, du \qquad\qquad \text{(nonnegativity)}$$

$$= \frac{(et)^2}{2\pi\theta \log t} \sum_{\chi_j} \int_{-\infty}^\infty \left|\frac{G'(s, \chi_j)}{s}\right|^2\, d\tau \qquad\qquad (4.9)$$

by Lemma 3.3, since in this setup $\widetilde{\mathcal{N}}(u, g\chi_j)$ is precisely the summatory function $A(u)$ corresponding to $G'(s, \chi_j)$ with $a_n = g(n)\chi_j(n)(\log n)$.

Let $T \geq 2$, to be specified shortly (see p. 39). Setting

$$L_1 = \frac{t^2}{\theta \log t} \sum_{\chi_j} \int_{|\tau| \leq T} \left|\frac{G'(s, \chi_j)}{s}\right|^2\, d\tau$$

and L_2 a similar expression except that the range of integration is changed to $|\tau| > T$, by (4.9) we have that $S_2(t) \ll L_1 + L_2$. We bound L_2 first.

Since $2 \le t \le N$, we have $1 < \sigma \le 3$ which in turn implies $|s|^2 \le 10$. Hence

$$\sum_j \int_{|\tau| \le 1} |G'(s, \chi_j)|^2 \, d\tau \le 10 \sum_{\chi_j} \int_{|\tau| \le 1} \left| \frac{G'(s, \chi_j)}{s} \right|^2 \, d\tau$$

$$\ll \sum_{\chi_j} \int_{-\infty}^{\infty} \left| \frac{G'(s, \chi_j)}{s} \right|^2 \, d\tau$$

$$= 2\pi \int_{D^c}^{\infty} \sum_{\chi_j} \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|^2}{u^{2\sigma+1}} \, du, \qquad (4.10)$$

the third step by Lemma 3.3. Note the appeal to the vanishing of g on small primes. By Lemma 3.6 the last integrand is (with $0 < \varepsilon < 1$, $J \le D$ and $Q = Q_c$)

$$\ll_\varepsilon \left(u \prod_{\substack{p | Q_c \\ p \le u}} \left(1 - \frac{1}{p}\right) + u^\varepsilon D^{3/2} \log D \right) \sum_{\substack{n \le u \\ p | n \implies p > D^c}} |g(n)\chi_j(n)(\log n)|^2 \cdot \frac{1}{u^{2\sigma+1}}$$

$$\ll u \prod_{p \le D^c} \left(1 - \frac{1}{p}\right) \cdot (\log u)^2 \cdot \frac{u}{\log D^c} \cdot \frac{1}{u^{2\sigma+1}}$$

$$\ll \left(\frac{\log u}{\log D} \right)^2 \cdot \frac{1}{u^{2\sigma-1}}$$

(see Remark A.7). Returning to (4.10), it follows that

$$\sum_{\chi_j} \int_{|\tau| \le 1} |G'(s, \chi_j)|^2 \, d\tau \ll_\varepsilon \frac{1}{(\log D)^2} \int_{D^c}^{\infty} \frac{(\log u)^2}{u^{2\sigma-1}} \, du$$

$$\ll \frac{(\log t)^3}{(\log D)^2} \qquad (4.11)$$

after an integration by parts.

Let λ be a real number, and let $G_\lambda(s, \chi_j)$ temporarily denote the series

$$\sum_{n=1}^{\infty} g(n) n^{-i\lambda} \chi_j(n) n^{-s}.$$

Since $g(n)n^{-i\lambda}$ is still a completely multiplicative function with modulus no larger than 1 that vanishes on the primes $p \le D^c$, (4.11) holds with the integrand $|G'_\lambda(s, \chi_j)|^2$ as well.

Moreover, from

$$\int_{|\tau-\lambda|\leq 1}|G'(s,\chi_j)|^2\,d\tau = \int_{|v|\leq 1}|G_\lambda(s,\chi_j)|^2\,dv,$$

and $|\tau - m| \leq 1 \implies |s|^{-2} \leq 4|m|^{-2}$ for $|m| > 0$, it follows that

$$\sum_j \int_{|\tau|>T}\left|\frac{G'(s,\chi_j)}{s}\right|^2 d\tau \ll \sum_{|m|>T}|m|^{-2}\sum_j \int_{|\tau-m|\leq 1}|G'(s,\chi_j)|^2\,d\tau$$

$$\ll_\varepsilon \frac{(\log t)^3}{T(\log D)^2},$$

from which

$$L_2 \ll_\varepsilon \frac{t^2}{\theta\log t}\cdot\frac{(\log t)^3}{T(\log D)^2} = \frac{1}{\theta T}\left(\frac{t\log t}{\log D}\right)^2. \tag{4.12}$$

Now, define

$$B_0 = \max_{\substack{x\leq N,|\tau|\leq T,\chi_j,\\ \sigma=1+(\log t)^{-1}\\ 2\leq t\leq N}}|G(s,\chi_j)|.$$

The maximum exists since the relevant functions are continuous on a compact set and there are only finitely many j's to consider. Then

$$L_1 = \frac{t^2}{\theta\log t}\sum_{\chi_j}\int_{|\tau|\leq T}\left|G(s,\chi_j)\cdot\frac{G'(s,\chi_j)}{sG(s,\chi_j)}\right|^2 d\tau$$

$$\leq \frac{B_0^2 t^2}{\theta\log t}\sum_{\chi_j}\int_{|\tau|\leq T}\left|\frac{G'(s,\chi_j)}{sG(s,\chi_j)}\right|^2 d\tau$$

$$\ll \frac{B_0^2 t^2}{\theta\log t}\sum_{\chi_j}\int_{-\infty}^{\infty}\left|\frac{G'(s,\chi_j)}{sG(s,\chi_j)}\right|^2 d\tau. \tag{4.13}$$

To estimate L_1 therefore amounts to bounding B_0 and the last integral. For the latter, since g is completely multiplicative and vanishes unless $(n,Q_c)=1$, we have the representation

$$-\frac{G'(s,\chi_j)}{G(s,\chi_j)} = \sum_{\substack{n=1\\(n,Q_c)=1}}^{\infty}\frac{g(n)\chi_j(n)\Lambda(n)}{n^s}.$$

By the Plancherel Lemma 3.3 again,

$$\sum_{\chi_j}\int_{-\infty}^{\infty}\left|\frac{G'(s,\chi_j)}{sG(s,\chi_j)}\right|^2 d\tau = \sum_{\chi_j}\int_{D^c}^{\infty}\frac{|A(y)|^2}{y^{2\sigma+1}}\,dy \tag{4.14}$$

where

$$A(y) = \sum_{\substack{n \leq y \\ (n, Q_c) = 1}} g(n) \chi_j(n) \Lambda(n).$$

Interchanging the sum and the integral, we obtain an integrand which by Corollary 3.7 (with $0 < \varepsilon < 1$) is

$$\ll_\varepsilon \left(\frac{y}{\log y} + y^\varepsilon D^{3/2} \log D \right) \left(\sum_{\substack{p^k \leq y \\ p > D^c}} |g(p) \chi_j(p) (\log p)|^2 \right) \cdot \frac{1}{y^{2\sigma + 1}}$$

$$\ll \frac{1}{y^{2\sigma - 1}}$$

having used that $|g\chi_j| \leq 1$ for any j and

$$\sum_{\substack{p^k \leq y \\ p > D^c}} (\log p)^2 \leq \sum_{p^k \leq y} (\log p)^2 = O(y \log y)$$

by Lemma 3.1. The integral in (4.14) is then

$$\ll \int_{D^c}^\infty y^{-2\sigma + 1} \, dy \ll \log t.$$

Substituting this into (4.13) yields

$$L_1 \ll \frac{B_0^2 t^2}{\theta}. \tag{4.15}$$

Combining (4.12) and (4.15), we obtain

$$S_2(t) \ll \frac{B_0^2 t^2}{\theta} + \frac{1}{\theta T} \left(\frac{t \log t}{\log D} \right)^2. \tag{4.16}$$

We shall see that the first term involving B_0 is the important piece, upon choosing T appropriately (see p. 39).

We treat $S_1(t)$ indirectly. For $D^c \leq u \leq x$ and $0 < \delta < 1$ fixed, define

$$H(u) = \max_{D^c \leq z \leq u} \frac{1}{z^2 (\log z)^\delta} \sum_{\chi_j} \max_{2 \leq y \leq z} |\widetilde{\mathcal{N}}(y, g\chi_j)|^2$$

$$\left(= \max_{D^c \leq z \leq u} \frac{1}{z^2 (\log z)^\delta} S(z) \right).$$

Note that $H(u)$ would be zero if $u < D^c$ and we allowed the maximum to go down to 2 (due to the stipulation on g). Moreover, H is nondecreasing by definition. By the Cauchy-Schwarz inequality,

$$S_1(t) = t^2 \sum_{\chi_j} \left(\int_{D^c}^{t^\theta} \frac{1}{u^{1/2}(\log u)^{1/2-\delta/4}} \cdot \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|}{u^{3/2}(\log u)^{1/2+\delta/4}} \, du \right)^2$$

$$\leq t^2 \left(\int_{D^c}^{t^\theta} \frac{du}{u(\log u)^{1-\delta/2}} \right) \sum_{\chi_j} \left(\int_{D^c}^{t^\theta} \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|^2}{u^3(\log u)^{1+\delta/2}} \, du \right)$$

$$= t^2 \left(\int_{D^c}^{t^\theta} \frac{du}{u(\log u)^{1-\delta/2}} \right) \left(\int_{D^c}^{t^\theta} \sum_{\chi_j} \frac{|\widetilde{\mathcal{N}}(u, g\chi_j)|^2}{u^2(\log u)^\delta} \cdot \frac{1}{u(\log u)^{1-\delta/2}} \, du \right)$$

$$\leq t^2 \left(\int_{D^c}^{t^\theta} \frac{du}{u(\log u)^{1-\delta/2}} \right) \left(\int_{D^c}^{t^\theta} \frac{H(u)}{u(\log u)^{1-\delta/2}} \, du \right)$$

$$\leq t^2 H(t^\theta) \left(\int_{D^c}^{t^\theta} \frac{du}{u(\log u)^{1-\delta/2}} \right)^2$$

$$\ll_\delta t^2(\theta \log t)^\delta H(t^\theta)$$

Combining this with (4.16) and the error from Lemma 4.2,

$$S(t) \ll_\delta t^2(\theta \log t)^\delta H(t^\theta) + \frac{B_0^2 t^2}{\theta} + \frac{1}{\theta T}\left(\frac{t \log t}{\log D}\right)^2 + \left(\frac{t}{\log D}\right)^2 \left(\log\left(\frac{\log t}{\log D}\right)\right)^2$$

for $2 \leq t \leq N$. Call the righthand side of the previous inequality $\psi(t)$ temporarily. Then

$$H(x) = \ll_\delta \max_{D^c \leq t \leq x} \frac{1}{t^2(\log t)^\delta} \psi(t)$$

$$= \max_{D^c \leq t \leq x} \left\{ \theta^\delta H(t^\theta) + \frac{B_0^2}{\theta(\log t)^\delta} + \frac{(\log t)^{2-\delta}}{\theta T(\log D)^2} + \frac{1}{(\log t)^\delta(\log D)^2} \left(\log\left(\frac{\log t}{\log D}\right)\right)^2 \right\}$$

$$= \theta^\delta H(x^\theta) + \frac{B_0^2}{c^\delta \theta(\log D)^\delta} + \frac{(\log x)^{2-\delta}}{\theta T(\log D)^2} + \frac{1}{c^\delta(\log D)^{2+\delta}} \left(\log\left(\frac{\log x}{\log D}\right)\right)^2.$$

As mentioned $H(x^\theta) \leq H(x)$ (since $0 < \theta < 1$), therefore if we fix θ at a small enough value to ensure $\theta^\delta H(x^\theta) \leq \frac{1}{2}H(x)$, transferring this term to the lefthand side of the previous

inequality yields

$$H(x) \ll_{\delta,\theta} \frac{B_0^2}{(\log D)^\delta} + \frac{(\log x)^{2-\delta}}{T(\log D)^2} + \frac{1}{(\log D)^{2+\delta}} \left(\log \left(\frac{\log x}{\log D} \right) \right)^2.$$

Now we choose T. Under the current Lemma's assumptions, $x \leq N$ and $\log N \leq D^c$. Hence $D^c \geq \log x$ which implies that $\frac{1}{D^{cB}} \leq \frac{1}{(\log x)^B}$ for any $B > 0$. If we choose $T = D^{cB}$ for $B \geq 2$, say, then assuming that enough exceptional χ_j have been removed to ensure

$$B_0 \ll_\delta \left(\frac{\log x}{\log D} \right)^{\delta/2} \tag{4.17}$$

(that this is possible is proved in Corollary 3.11), and using

$$\left(\log \left(\frac{\log x}{\log D} \right) \right) \ll_\delta \left(\frac{\log x}{\log D} \right)^{\delta/2},$$

we obtain

$$H(x) \ll_{\delta,\theta} \left(\frac{\log x}{\log D} \right)^\delta \cdot \frac{1}{(\log D)^\delta} + \frac{1}{(\log x)^\delta (\log D)^2} + \frac{1}{(\log D)^{2+\delta}} \cdot \left(\frac{\log x}{\log D} \right)^\delta$$

$$\ll \left(\frac{\log x}{\log D} \right)^\delta \frac{1}{(\log D)^\delta}.$$

From the definition of $H(x)$, it follows at once that

$$\sum_{\substack{\chi_j \text{ not} \\ \text{except.}}} \max_{2 \leq y \leq t} |\widetilde{\mathcal{N}}(y, g\chi_j)|^2 \ll_\delta t^2 (\log t)^\delta \left(\frac{\log x}{\log D} \right)^\delta \frac{1}{(\log D)^\delta}$$

$$\ll t^2 \left(\frac{\log x}{\log D} \right)^{2\delta} \tag{4.18}$$

uniformly for $D^c \leq t \leq x$ and for any fixed δ, $0 < \delta < 1$. This uniformity will be critical when we apply Lemma 4.1 in the following Chapters (see pp. 45 and 52).

The extra $\log n$ factor in the sum $\widetilde{\mathcal{N}}(y, g\chi_j)$ may be removed by appealing to Lemma 3.6. Since the Lemma holds for any collection of characters, with $H = t$, $Q = 1$ and for

$0 < \varepsilon < 1$,

$$\sum_{\substack{\chi_j \text{ not} \\ \text{except.}}} \max_{2 \leq y \leq t} \left| \sum_{n \leq y} g(n)\chi_j(n) \log(t/n) \right|^2 \ll_\varepsilon \left(t \cdot 1 + t^\varepsilon D^{3/2} \log D \right) \sum_{\substack{n \leq t \\ n > D^c}} |g(n) \log(t/n)|^2$$

$$\ll t \sum_{n \leq t} (\log(t/n))^2 \ll t^2 \qquad (4.19)$$

by Lemma A.8, possibly ignoring some cancellation. Since $\log t = \log(t/n) + \log n$, by the Cauchy-Schwarz inequality

$$(\log t)^2 \left| \sum_{n \leq y} g(n)\chi_j(n) \right|^2 = \left| \sum_{n \leq y} g(n)\chi_j(n) \log t \right|^2$$

$$\leq 2 \left\{ \left| \sum_{n \leq y} g(n)\chi_j(n) \log(t/n) \right|^2 + \left| \sum_{n \leq y} g(n)\chi_j(n) \log n \right|^2 \right\}.$$

Taking the maximum over $2 \leq y \leq t$ and summing over the nonexceptional χ_j, we obtain

$$(\log t)^2 \sum_{\substack{\chi_j \text{ not} \\ \text{except.}}} \max_{2 \leq y \leq t} \left| \sum_{n \leq y} g(n)\chi_j(n) \right|^2 \ll t^2 + t^2 \left(\frac{\log x}{\log D} \right)^{2\delta} \ll t^2 \left(\frac{\log x}{\log D} \right)^{2\delta}$$

by (4.18) and (4.19). Dividing through by $(\log t)^2$ yields

$$\sum_{\substack{\chi_j \text{ not} \\ \text{except.}}} \max_{2 \leq y \leq t} \left| \sum_{n \leq y} g(n)\chi_j(n) \right|^2 \ll_\delta \left(\frac{t}{\log t} \right)^2 \left(\frac{\log x}{\log D} \right)^{2\delta}.$$

Since $0 < \delta < 1$ was fixed but arbitrary and the number of exceptional characters removed depended on it, this completes the proof of Lemma 4.1.

CHAPTER 5

FROM L^2 TO L^∞

For an arithmetic function f, define

$$Y(f, a, x) = \sum_{\substack{n \leq x \\ n \equiv a \, (\mathrm{mod} \, D)}} f(n) - \sum_{\chi_j \text{ except.}} \frac{\overline{\chi_j}(a)}{\varphi(D)} \sum_{n \leq x} f(n) \chi_j(n).$$

Theorem 2.2 will follow from an adequate upper bound for $Y(g, a, x)$, where $(a, D) = 1$, $2 \leq D \leq x$ and g is an unrestricted multiplicative function with values in the complex unit disc. We will accomplish this by an appropriate application of Lemma 4.1, where the connection to the previous Chapter is provided by the orthogonality property of Dirichlet characters. Recalling that

$$\frac{1}{\varphi(D)} \sum_{\chi_j} \overline{\chi_j}(a) \chi_j(n) = \begin{cases} 1 & \text{if } n \equiv a \, (\mathrm{mod} \, D) \\ 0 & \text{otherwise,} \end{cases} \tag{5.1}$$

for any a that is coprime to D, our first step is the following Lemma:

Lemma 5.1. *Let* χ_j, $j \in \mathcal{J}$, *be a collection of Dirichlet characters* $(\mathrm{mod} \, D)$. *Let*

$$\mathcal{L}(a) = \sum_{\substack{n \geq 1 \\ n \equiv a \, (\mathrm{mod} \, D)}} b_n - \sum_{j \notin \mathcal{J}} \frac{\overline{\chi_j}(a)}{\varphi(D)} \mathcal{S}_j,$$

where $b_n \in \mathbb{C}$, $b_n = 0$ *for all* $n > n_0$ *(say), and*

$$\mathcal{S}_j = \sum_{n \geq 1} b_n \chi_j(n).$$

Then

$$\varphi(D) \sum_{\substack{a=1 \\ (a,D)=1}}^{D} |\mathcal{L}(a)|^2 = \sum_{j \in \mathcal{J}} |\mathcal{S}_j|^2.$$

Proof of Lemma 5.1 (c.f. [13, p. 180, line -6]). From (5.1) it follows easily that

$$\mathcal{L}(a) = \sum_{j \in \mathcal{J}} \frac{\overline{\chi_j}(a)}{\varphi(D)} \mathcal{S}_j,$$

and hence

$$\sum_{\substack{a=1 \\ (a,D)=1}}^{D} |\mathcal{L}(a)|^2 = \sum_{\substack{a=1 \\ (a,D)=1}}^{D} \frac{1}{(\varphi(D))^2} \left| \sum_{j \in \mathcal{J}} \overline{\chi_j}(a) \mathcal{S}_j \right|^2$$

$$= \sum_{\substack{a=1 \\ (a,D)=1}}^{D} \frac{1}{(\varphi(D))^2} \sum\sum_{j_1,j_2 \in \mathcal{J}} \overline{\chi_{j_1}}(a) \chi_{j_2}(a) \mathcal{S}_{j_1} \overline{\mathcal{S}_{j_2}}$$

$$= \frac{1}{(\varphi(D))^2} \sum\sum_{j_1,j_2 \in \mathcal{J}} \mathcal{S}_{j_1} \overline{\mathcal{S}_{j_2}} \underbrace{\sum_{\substack{a=1 \\ (a,D)=1}}^{D} \overline{\chi_{j_1}}(a) \chi_{j_2}(a)}_{= \begin{cases} \varphi(D) & \text{if } j_1 = j_2 \\ 0 & \text{if } j_1 \neq j_2 \end{cases}}$$

$$= \frac{1}{\varphi(D)} \sum_{j \in \mathcal{J}} |\mathcal{S}_j|^2,$$

completing the proof after an appeal to the above alternate version of the orthogonality of characters.

We continue to assume that the conditions in Lemma 4.1 are in force: that g is a completely multiplicative function that vanishes on the primes $p \leq D^c$ (and $p > x$), g satisfies $|g| \leq 1$ and δ, $0 < \delta < 1$, is a parameter that the number of exceptional characters depends upon. In the notation of Lemma 5.1, $\mathcal{L}(a) = Y(g, a, x)$, where $g(n) = b_n$, $n_0 = [x]$,

and $\mathcal{J}$ is the set of *nonexceptional* characters. Therefore

$$\varphi(D) \sum_{\substack{a=1 \\ (a,D)=1}}^{D} |Y(g,a,x)|^2 = \sum_{\substack{\chi_j \text{ not} \\ \text{except.}}} \left| \sum_{\substack{n \leq x \\ p|n \implies p>D^c}} g(n)\chi_j(n) \right|^2, \tag{5.2}$$

the righthand side of which can be bounded by Lemma 4.1. The resulting estimate can be viewed as an L^2 estimate for $Y(g,a,x)$ over *all* the reduced residue classes modulo D. Ultimately we require an L^∞ estimate (of the desired form) for

$$\sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, D)}} g(n),$$

which is a sum over a *particular* reduced residue class. We now leverage the former into the latter.

As before we will accomplish this by introducing a logarithm and removing it at the end.

Using the representation $\log n = \sum_{d|n} \Lambda(d)$,

$$Y(g\log, a, x) = \sum_{\substack{d \leq x \\ p|d \implies p>D^c}} g(d)\Lambda(d) \left(\sum_{\substack{m \leq x/d \\ md \equiv a \,(\mathrm{mod}\, D) \\ p|m \implies m>D^c}} g(m) \right.$$

$$\left. - \sum_{\chi_j \text{ except.}} \frac{\overline{\chi_j}(a)\chi_j(d)}{\varphi(D)} \sum_{\substack{m \leq x/d \\ p|m \implies p>D^c}} \chi(m)g(m) \right)$$

$$= \sum_{\substack{d \leq x \\ p|d \implies p>D^c}} g(d)\Lambda(d)Y\left(g, a\overline{d}, \frac{x}{d}\right) \tag{5.3}$$

where $d\overline{d} \equiv 1 \,(\mathrm{mod}\, D)$. Note that if $c \geq 1$, then the condition $p \mid n \implies p > D^c$ implies

that $(md, D) = 1$. By (5.3), since $|g| \leq 1$, for a c_1 to be specified shortly, we have

$$|Y(g \log, a, x)| \leq \sum_{\substack{d \leq x \\ p|d \implies p > D^c}} \Lambda(d) \left| Y\left(g, a\bar{d}, \frac{x}{d}\right) \right| \leq \sum_{D^c < d \leq x} \Lambda(d) \left| Y\left(g, a\bar{d}, \frac{x}{d}\right) \right|$$

$$= \sum_{D^c < d \leq x/D^{c_1}} \Lambda(d) \left| Y\left(g, a\bar{d}, \frac{x}{d}\right) \right| + \sum_{x/D^{c_1} < d \leq x} \Lambda(d) \left| Y\left(g, a\bar{d}, \frac{x}{d}\right) \right|, \qquad (5.4)$$

where in the second inequality we have potentially let more d's into the sum. We will handle the second sum directly, but first we use the Cauchy-Schwarz inequality and the L^2 estimate from the previous Chapter to bound the first sum. To this end we divide the interval $(D^c, x/D^{c_1}]$ into a union of adjoining subintervals of the form $(U, 2U]$, where U runs through the powers of 2 constrained by $D^c/2 \leq 2^k \leq 2x/D^{c_1}$. Furthermore, for some β with $0 < \beta < 1$, we split each interval $(U, 2U]$ into a union of adjoining subintervals of the form $(V, V + U^\beta)$, where V begins at U and the last such subinterval contains $2U$ (for each U). Then

$$\sum_{D^c < d \leq x/D^{c_1}} \Lambda(d) \left| Y\left(g, a\bar{d}, \frac{x}{d}\right) \right| \leq \sum_U \sum_V \sum_{V < d < V + U^\beta} \Lambda(d) \left| Y\left(g, a\bar{d}, \frac{x}{d}\right) \right|. \qquad (5.5)$$

Replacing $\frac{x}{d}$, in the innermost sum, by $\frac{x}{V}$ introduces an error that is

$$\ll \sum_U \sum_V \sum_{V < d < V + U^\beta} \Lambda(d) \left\{ \sum_{\substack{x/d < m \leq x/V \\ m \equiv a\bar{d} \,(\text{mod } D) \\ p|m \implies p > D^c}} 1 + \sum_{\chi_j \text{ except.}} \frac{1}{\varphi(D)} \left| \sum_{\substack{x/d < m \leq x/V \\ p|m \implies p > D^c}} g(m)\chi_j(m) \right| \right\}$$

$$\ll_\delta \sum_U \sum_V \sum_{V < d < V + U^\beta} \Lambda(d) \left\{ 1 + \frac{1}{\varphi(D)} \left(\frac{x}{V} - \frac{x}{d} \right) \right\}$$

$$\ll \frac{x}{\varphi(D)} \sum_{D^c/2 < d \leq 2x/D^{c_1}} \frac{\Lambda(d)}{d^{2-\beta}} + \sum_{D^c/2 < d \leq 2x/D^{c_1}} \Lambda(d)$$

$$\ll \frac{x}{\varphi(D)(D^c)^{1-\beta}} + \frac{x}{D^{c_1}} \qquad (5.6)$$

(see Remark A.9). The second term in (5.6) may be omitted provided $c_1 \geq 1 + (1 - \beta)c$, which we may assume.

Compared to (5.5), we are reduced to estimating the sum

$$\sum_{U}\sum_{V}\sum_{V<d<V+U^\beta} \Lambda(d)\left|Y\left(g,a\bar{d},\frac{x}{V}\right)\right|.\tag{5.7}$$

A typical innermost sum over d is

$$\sum_{\substack{b=1\\(b,D)=1}}^{D}\sum_{\substack{V<d<V+U^\beta\\d\equiv b\,(\mathrm{mod}\,D)}}\Lambda(d)\left|Y\left(g,a\bar{b},\frac{x}{V}\right)\right|$$

$$=\sum_{\substack{b=1\\(b,D)=1}}^{D}\left|Y\left(g,a\bar{b},\frac{x}{V}\right)\right|\sum_{\substack{V<d<V+U^\beta\\d\equiv b\,(\mathrm{mod}\,D)}}\Lambda(d)$$

$$\ll\frac{U^\beta}{\varphi(D)}\cdot\sum_{\substack{b=1\\(b,D)=1}}^{D}\left|Y\left(g,a\bar{b},\frac{x}{V}\right)\right|,$$

once again by an application of a sieve, provided $U^\beta\geq D^{1+\varepsilon}$ for some $\varepsilon>0$ (that may be chosen arbitrarily at the outset). We may assume this is satisfied, provided we choose c and β to satisfy $c\beta>1$.

Applying the Cauchy-Schwarz inequality, the above bound for the sum over d is

$$\leq\frac{U^\beta}{\varphi(D)}\left\{\varphi(D)\cdot\sum_{\substack{b=1\\(b,D)=1}}^{D}\left|Y\left(g,a\bar{b},\frac{x}{V}\right)\right|^2\right\}^{1/2}$$

$$=\frac{U^\beta}{\varphi(D)}\left\{\sum_{\substack{\chi_j\text{ not}\\\text{except.}}}\left|\sum_{\substack{n\leq x/V\\p|n\implies p>D^c}}g(n)\chi_j(n)\right|^2\right\}^{1/2}$$

by (5.2), where we relied on the fact that $a\bar{b}$ traverses a complete set of reduced residues as b does. By Lemma 4.1, the previous quantity is

$$\ll_\delta\frac{U^\beta}{\varphi(D)}\left\{\left(\frac{x/V}{\log(x/V)}\right)^2\left(\frac{\log x}{\log D}\right)^{2\delta}\right\}^{1/2}$$

$$=\frac{x}{\varphi(D)}\left(\frac{\log x}{\log D}\right)^\delta\cdot U^\beta\cdot\frac{1}{V\log(x/V)}.$$

The whole triple sum over U, V, d in (5.7) must therefore be

$$\ll_\delta \frac{x}{\varphi(D)} \left(\frac{\log x}{\log D}\right)^\delta \sum_U U^\beta \sum_V \frac{1}{V \log(x/V)}$$

$$\leq \frac{x}{\varphi(D)} \left(\frac{\log x}{\log D}\right)^\delta \sum_U \frac{1}{\log(x/2U)}$$

$$\ll \frac{x}{\varphi(D)} \left(\frac{\log x}{\log D}\right)^\delta \log\left(\frac{\log x}{\log D}\right) \tag{5.8}$$

after an integration by parts (see Remark A.10). Returning to (5.5), by (5.6) and (5.8) we obtain

$$\sum_{D^c < d \leq x/D^{c_1}} \Lambda(d) \left|Y\left(g, a\bar{d}, \frac{x}{d}\right)\right| = \sum_{D^c < d \leq x/D^{c_1}} \Lambda(d) \left|Y\left(g, a\bar{d}, \frac{x}{V}\right)\right| + (\text{error})$$

$$\ll_\delta \frac{x}{\varphi(D)} \left(\frac{\log x}{\log D}\right)^\delta \log\left(\frac{\log x}{\log D}\right) + \frac{x}{\varphi(D)(D^c)^{1-\beta}}. \tag{5.9}$$

Recalling (5.4), we still need to bound

$$\sum_{x/D^{c_1} < d \leq x} \Lambda(d) \left|Y\left(g, a\bar{d}, \frac{x}{d}\right)\right|, \tag{5.10}$$

but thankfully this is easier. Using the definition of Y, we argue crudely. The previous quantity is then

$$\ll \sum_{x/D^{c_1} < d \leq x} \Lambda(d) \left\{ \sum_{\substack{m \leq x/d \\ m \equiv a\bar{d} \,(\mathrm{mod}\, D) \\ p|m \implies p > D^c}} |g(m)| - \sum_{\chi_j \text{ except.}} \frac{1}{\varphi(D)} \sum_{\substack{m \leq x/d \\ p|m \implies p > D^c}} |g(m)\chi_j(m)| \right\}$$

$$\ll_\alpha \sum_{m \leq D^{c_1}} |g(m)| \sum_{\substack{d \leq x/m \\ d \equiv a\bar{m} \,(\mathrm{mod}\, D)}} \Lambda(d) + \frac{1}{\varphi(D)} \sum_{m \leq D^{c_1}} |g(m)| \sum_{d \leq x/m} \Lambda(d) \tag{5.11}$$

$$\ll \frac{x}{\varphi(D)} \sum_{m \leq D^{c_1}} \frac{|g(m)|}{m} \ll \frac{x}{\varphi(D)} \prod_{D^c < p \leq D^{c_1}} \left(1 - \frac{1}{p}\right)^{-1}$$

$$\ll \frac{x}{\varphi(D)}, \tag{5.12}$$

where we used a sieve and then Lemma 3.1 to bound the sums over d in (5.11), from left to right, respectively, and an Euler product representation in the step thereafter. That the product over primes between two powers of D in the penultimate step is $O(1)$ is considered in Remark A.5, and the sieve estimate is certainly valid if $x > D^{c_1+1+\eta}$, $\eta > 0$, since $m \leq D^{c_1}$ then implies $x/m > D^{1+\eta}$. Combining (5.12) with (5.9) and (5.4), we obtain

$$Y(g\log, a, x) \ll_\delta \frac{x}{\varphi(D)} \left(\frac{\log x}{\log D} \right)^{2\delta}, \qquad (5.13)$$

where we have used

$$\log\left(\frac{\log x}{\log D} \right) \ll_\delta \left(\frac{\log x}{\log D} \right)^\delta.$$

In the present circumstances, δ comes from the definition of H (see p. 37). We could choose $0 < \delta < \frac{1}{2}$ instead (or just replace δ by $\frac{\delta}{2}$). Without loss of generality, (5.13) then holds with δ instead of 2δ in the exponent.

To remove the logarithm, we appeal to the estimates

$$\sum_{n \leq x} \log\left(\frac{x}{n} \right) \ll x \qquad \text{and} \qquad \sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, D)}} \log\left(\frac{x}{n} \right) \ll \frac{x}{\varphi(D)}$$

from Lemma A.11. Since $\log x - \log n = \log\left(\frac{x}{n} \right)$,

$$(\log x)Y(g, a, x) - Y(g\log, a, x) =$$

$$\sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, D) \\ p|n \implies p > D^c}} g(n)\log\left(\frac{x}{n} \right) - \sum_{\chi_j \text{ except.}} \frac{\overline{\chi_j}(a)}{\varphi(D)} \sum_{\substack{n \leq x \\ p|n \implies p > D^c}} g(n)\chi_j(n)\log\left(\frac{x}{n} \right)$$

$$\ll_\alpha \sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, D) \\ p|n \implies p > D^c}} \log\left(\frac{x}{n} \right) + \frac{1}{\varphi(D)} \sum_{\substack{n \leq x \\ p|n \implies p > D^c}} \log\left(\frac{x}{n} \right)$$

$$\ll \frac{x}{\varphi(D)}$$

where we have eliminated the condition $p \mid n \implies p > D^c$ (so that the sums can only increase), used $|g| \leq 1$, $|\overline{\chi_j}(a)| \leq 1$ and the fact that there are only $O_\delta(1)$ exceptional

characters. Since this difference is less than the bound in (5.13), dividing through by $\log x$ provides the estimate

$$Y(g, a, x) \ll_\delta \frac{x}{\varphi(D) \log x} \left(\frac{\log x}{\log D} \right)^\delta. \tag{5.14}$$

The proof of Theorem 2.2 will be complete after we remove the vanishing and multiplicative conditions from g. This is the subject of the next Chapter.

CHAPTER 6

STRIPPING THE RESTRICTIONS AND IMPROVING THE ERROR

First we remove the support condition that g vanishes on the primes $p \leq D^c$ (c.f. [16, p. 26] or [13, p. 194]). In order to do so, however, we need one more Lemma providing a bound for the number of integers, not exceeding x, that lie in a residue class modulo D and have prime factors not exceeding D^c. Such a bound is well known in probabilistic number theory and may be proven by appealing to Lemmas 3.3 and 3.5 in [5, p. 127 and 132, resp.]; cf. Lemma 13 in [13, p. 192]. As an alternative, we utilize Lemma 4 in [18, p. 7], which does not require the construction of a probability space and will help improve the error. I reproduce a proof here for convenience.

Lemma 6.1. *Let $c > 0$. For any positive integer ℓ, the number of integers n in the interval $x^{1/2} < n \leq x$ made up of primes not exceeding D^c and which satisfy $n \equiv a \,(mod\, D)$ is*

$$\ll \frac{x}{D} \left(\frac{\log D}{\log x} \right)^{\ell}$$

uniformly in $D \leq x$ and a.

Proof of Lemma 6.1. Let $(a, D) = 1$, $D \leq x^{1/2}$. Suppose that the integers n are squarefree. Then their number does not exceed

$$\sum_{\substack{n \leq x \\ n \equiv a \,(mod\, D) \\ p|n \implies p \leq D^c}} \mu^2(n) \left(\frac{\log n}{\log x^{1/2}} \right)^{\ell}$$

$$= \left(\frac{1}{2}\log x\right)^{-\ell} \sum_{p_j \leq D^c} \log p_1 \cdots \log p_\ell \sum_{\substack{m \leq x[p_1,\ldots,p_\ell]^{-1} \\ m[p_1,\ldots,p_\ell] \equiv a \,(\mathrm{mod}\, D)}} 1.$$

Here $[p_1,\ldots,p_\ell] \leq D^{\ell c}$ and if $D^{\ell c} \leq x^{1/4}$, say, then the innersum is $\ll D^{-1}x[p_1,\ldots,p_\ell]^{-1}$.

The terms with the p_j distinct contribute

$$\ll \frac{x}{D(\log x)^\ell} \left(\sum_{p \leq D^c} \frac{\log p}{p}\right)^\ell \ll \frac{x}{D}\left(\frac{\log D}{\log x}\right)^\ell.$$

Those terms with two p_j equal contribute

$$\ll \frac{x}{D(\log x)^\ell} \left(\sum_{p \leq D^c} \frac{\log p}{p}\right)^{\ell-2} \cdot \sum_{p \leq D^c} \frac{(\log p)^2}{p},$$

which has a similar bound; and so on.

If $D > x^{1/2}$ or $D^{\ell c} \geq x^{1/4}$ then the desired bound is evident, since we may then ignore the condition that the prime factors of n do not exceed D.

More generally, any positive integer has a representation rm^2 with r squarefree and we may estimate the number of such integers not exceeding x which have only prime factors not exceeding D^c, and which satisfy $rm^2 \equiv a \,(\mathrm{mod}\, D)$. For $D \leq x^{1/8}$ those integers with $m \leq x^{1/4}$ are, by our above argument,

$$\ll \sum_{m \leq x^{1/4}} \frac{x}{Dm^2} \left(\frac{\log D}{\log x}\right)^\ell \ll \frac{x}{D}\left(\frac{\log D}{\log x}\right)^\ell$$

in number. Omitting side conditions, those with $m > x^{1/4}$ are

$$\ll \sum_{m > x^{1/4}} m^{-2}x \ll x^{3/4} \ll D^{-1}x^{7/8}$$

in number. Once again, the condition $D \leq x^{1/8}$ may be removed in favor of $D \leq x$.

If now $(a, D) = t$, then the integers n in the Lemma have the form tw where $w \leq x/t$, w has only prime factors not exceeding D^c and $w \equiv t^{-1}a \,(\mathrm{mod}\, t^{-1}D)$. For $D \leq x^{1/2}$ their number is

$$\ll \frac{t^{-1}x}{t^{-1}D} \left(\frac{\log t^{-1}D}{\log t^{-1}x}\right)^\ell \ll \frac{x}{D}\left(\frac{\log D}{\log x}\right)^\ell,$$

and once again the condition $D \leq x^{1/2}$ may be relaxed to $D \leq x$. This completes the proof of Lemma 6.1.

Suppose now that g is an arbitrary completely multiplicative function with $|g| \leq 1$. Define further completely multiplicative functions h and k by

$$h(p) = \begin{cases} g(p) & \text{if } p \leq D^c \\ 0 & \text{otherwise} \end{cases} \qquad k(p) = \begin{cases} g(p) & \text{if } p > D^c \\ 0 & \text{otherwise.} \end{cases}$$

Then $g = h * k$ since the Dirichlet convolution of completely multiplicative functions is (at least) multiplicative while g and the convolution agree on prime-powers. Indeed, for any $r \geq 1$,

$$(h * k)(p^r) = \sum_{i=0}^{r} h(p^i)k(p^{r-i}) = \begin{cases} 1 \cdot k(p^r) + \sum_{i=1}^{r} 0 \cdot k(p^{r-i}) = g(p^r) & \text{if } p > D^c \\ \sum_{i=0}^{r-1} h(p^i) \cdot 0 + h(p^r) \cdot 1 = g(p^r) & \text{if } p \leq D^c. \end{cases}$$

We wish to estimate $Y(g, a, x)$, and this time we use the representation $g(n) = \sum_{d|n} h(d)k(n/d)$. With $d\bar{d} \equiv 1 \pmod{D}$,

$$Y(g, a, x) = \sum_{\substack{n \leq x \\ n \equiv a \pmod{D}}} \sum_{d|n} h(d)k\left(\frac{n}{d}\right) - \sum_{\chi_j \text{ except.}} \frac{\overline{\chi_j}(a)}{\varphi(D)} \sum_{n \leq x} \sum_{d|n} h(d)k\left(\frac{n}{d}\right) \chi_j(n)$$

$$= \sum_{d \leq x} h(d) \left\{ \sum_{\substack{m \leq x/d \\ m \equiv a\bar{d} \pmod{D}}} k(m) - \sum_{\chi_j \text{ except.}} \frac{\overline{\chi_j}(a\bar{d})}{\varphi(D)} \sum_{m \leq x/d} k(m)\chi_j(m) \right\}.$$

Splitting the outermost sum into two sums, one over $d \leq x^\beta$ and the other over $x^\beta < d \leq x$, $0 < \beta < 1$ fixed, we write $Y(g, a, x) = \Sigma_1 + \Sigma_2$ where

$$\Sigma_1 = \sum_{d \leq x^\beta} h(d) \left\{ \sum_{\substack{m \leq x/d \\ m \equiv a\bar{d} \pmod{D}}} k(m) - \sum_{\chi_j \text{ except.}} \frac{\overline{\chi_j}(a\bar{d})}{\varphi(D)} \sum_{m \leq x/d} k(m)\chi_j(m) \right\}$$

and

$$\Sigma_2 = \sum_{x^\beta < d \le x} h(d) \left\{ \sum_{\substack{m \le x/d \\ m \equiv a\bar{d} \,(\mathrm{mod}\, D)}} k(m) - \sum_{\chi_j \text{ except.}} \frac{\overline{\chi_j}(a\bar{d})}{\varphi(D)} \sum_{m \le x/d} k(m)\chi_j(m) \right\}$$

$$= \sum_{m \le x^{1-\beta}} k(m) \left\{ \sum_{\substack{x^\beta < d \le x/m \\ d \equiv a\bar{m} \,(\mathrm{mod}\, D)}} h(d) - \sum_{\chi_j \text{ except.}} \frac{\overline{\chi_j}(a\bar{m})}{\varphi(D)} \sum_{x^\beta < d \le x/m} h(d)\chi_j(d) \right\}$$

upon changing the order of summation. By (5.14), since k is completely multiplicative and vanishes on the primes $p \le D^c$ by construction, the sum Σ_1 satisfies

$$\Sigma_1 = \sum_{d \le x^\beta} h(d) Y\left(k, a\bar{d}, \frac{x}{d}\right) \ll_\delta \sum_{d \le x^\beta} |h(d)| \frac{x/d}{\varphi(D)(\log(x/d))} \left(\frac{\log(x/d)}{\log D}\right)^\delta$$

$$\ll_\beta \frac{x}{\varphi(D)\log x} \left(\frac{\log x}{\log D}\right)^\delta \sum_{d \le x^\beta} \frac{|h(d)|}{d}, \tag{6.1}$$

where we have used

$$d \le x^\beta \implies \frac{1}{\log(x/d)} \ll_\beta \frac{1}{\log x}.$$

At this point a relatively simple observation will help us improve the error term. Since $(a, D) = 1$ from the beginning and $n \equiv a \,(\mathrm{mod}\, D)$, n (and all of its divisors) is coprime to D. Hence every d in the outermost sum in (6.1) is coprime to D. From the Euler product representation of h, therefore, since $|h| \le 1$ where it *doesn't* vanish (a property it inherits from g), the aforementioned sum is

$$\sum_{\substack{d \le x^\beta \\ (d,D)=1}} \frac{|h(d)|}{d} \le \prod_{\substack{p \le D^c \\ (p,D)=1}} \left(1 - \frac{1}{p}\right)^{-1} \ll \frac{\varphi(D)\log D}{D}$$

by Lemma 3.1. Combining this with (6.1) and rearranging shows that

$$\Sigma_1 \ll_\delta \frac{x}{D} \left(\frac{\log D}{\log x}\right)^{1-\delta}. \tag{6.2}$$

Similarly, using

$$\sum_{m \le x^{1-\beta}} \frac{|k(m)|}{m} \le \prod_{D^c < p \le x} \left(1 - \frac{1}{p}\right)^{-1} \ll \frac{\log x}{\log D}$$

and Lemma 6.1 with $\ell = 2$, it follows that

$$\Sigma_2 \ll_\delta \frac{x}{D} \left(\frac{\log D}{\log x} \right)^2 \cdot \frac{\log x}{\log D} = \frac{x}{D} \left(\frac{\log D}{\log x} \right),$$

which is smaller than the bound for Σ_1. Therefore

$$Y(g, a, x) \ll_\delta \frac{x}{D} \left(\frac{\log D}{\log x} \right)^{1-\delta} \tag{6.3}$$

for completely multiplicative functions g with $|g| \leq 1$ but *without* vanishing restrictions.

To remove the need for g to be completely multiplicative, I follow [13, p. 201]. For a general multiplicative function g with $|g| \leq 1$, define a completely multiplicative function g_1 by $g_1(p^r) = g(p)^r$ for all integers $r \geq 1$. Then

$$g_1^{-1}(p^r) = \mu(p^r)g_1(p^r) = \begin{cases} 1 & \text{if } r = 0 \\ -g(p) & \text{if } r = 1 \\ 0 & \text{if } r > 1 \end{cases}$$

(see Remark A.12), from which it follows easily that if we define g_2 by $g_2 = g * g_1^{-1}$, then

$$g_2(p) = 0 \quad \text{and} \quad g_2(p^r) = g(p^r) - g(p)g(p^{r-1}) \quad \text{for } r \geq 2.$$

In particular, $g = g_1 * g_2$, $g_2(p^r) \leq 2$ for all $r \geq 2$ and g_2 is multiplicative. Using an Euler product again,

$$\sum_{m \leq x} |g_2(m)| m^{-1/2} \leq \prod_{p \leq x} \left(1 + \frac{0}{p^{1/2}} + \frac{2}{p} + \frac{2}{p^{3/2}} + \cdots \right)$$

$$= \prod_{p \leq x} \left\{ 1 + \frac{2}{p} + O\left(\frac{2}{p^{3/2}} \right) \right\}$$

$$\ll \exp\left(\sum_{p \leq x} \frac{2}{p} \right)$$

$$\ll (\log x)^2, \tag{6.4}$$

upon using a geometric progression, the inequality $1 + x \leq e^x$ (valid for all $x \geq 0$, say) and Lemma 3.1. Note that, for any $M \geq 1$,

$$\sum_{M < m \leq 2M} |g_2(m)| m^{-1} \leq M^{-1/2} \sum_{M < m \leq 2M} |g_2(m)| m^{-1/2} \ll \frac{(\log M)^2}{M^{1/2}}$$

by (6.4). Using diadic intervals,

$$\sum_{m=1}^{\infty} |g_2(m)| m^{-1} = 1 + \sum_{\ell=0}^{\infty} \sum_{2^\ell < m \leq 2^{\ell+1}} |g_2(m)| m^{-1}$$

$$\ll \sum_{\ell=0}^{\infty} \frac{(\log 2^\ell)^2}{2^{\ell/2}} \ll \sum_{\ell=0}^{\infty} \frac{\ell^2}{2^{\ell/2}},$$

which certainly converges. In other words, $\sum_{m=1}^{\infty} g_2(m) m^{-1}$ converges absolutely. Since $g = g_1 * g_2$, using the representation

$$g(n) = \sum_{uv=n} g_1(u) g_2(v),$$

a calculation similar to that leading up to (5.3) allows us to write

$$Y(g, a, x) = \sum_{\substack{v \leq x \\ (v, D) = 1}} g_2(v) Y(g_1, a\bar{v}, x/v),$$

where $v\bar{v} \equiv 1 \,(\mathrm{mod}\, D)$. For any $0 < \beta < 1$, we split the preceding sum over v into one sum over $v \leq x^\beta$ and another over $x^\beta < v \leq x$. Since g_1 is completely multiplicative we may apply (6.3) to $Y(g_1, a\bar{v}, x/v)$ for $v \leq x^\beta$:

$$\sum_{\substack{v \leq x^\beta \\ (v, D) = 1}} g_2(v) Y(g_1, a\bar{v}, x/v) \ll_\delta \sum_{\substack{v \leq x^\beta \\ (v, D) = 1}} |g_2(v)| \cdot \frac{x}{vD} \left(\frac{\log D}{\log(x/v)} \right)^{1-\delta}$$

$$\ll_\beta \frac{x}{D} \left(\frac{\log D}{\log x} \right)^{1-\delta}, \tag{6.5}$$

where we have (again) used the fact that $v \leq x^\beta \implies \frac{1}{\log(x/v)} \ll_\beta \frac{1}{\log x}$ as well as the aforementioned absolute convergence of the series involving g_2.

For $x^\beta < v \le x$, since

$$\sum_{x^\beta < v \le x} |g_2(v)|v^{-1} \ll x^{-\beta/2}(\log x)^2$$

from (6.4), estimating $Y(g_1, a\overline{v}, x/v)$ crudely by a sieve as before (see p. 46) yields

$$\sum_{\substack{x^\beta < v \le x \\ (v,D)=1}} g_2(v)Y(g_1, a\overline{v}, x/v) \ll_\delta \sum_{\substack{x^\beta < v \le x \\ (v,D)=1}} |g_2(v)| \left(\frac{x}{v\varphi(D)} + 1\right)$$

$$\ll \frac{x^{1-\beta/2}(\log x)^2}{\varphi(D)}$$

$$\ll \frac{x^{1-\beta/3}}{\varphi(D)}. \tag{6.6}$$

Combining (6.5) and (6.6) establishes the estimate

$$Y(g, a, x) \ll_\delta \frac{x}{D}\left(\frac{\log D}{\log x}\right)^{1-\delta}$$

for any multiplicative function g with $|g| \le 1$ and the uniformities inherited from Lemma 4.1. Renaming δ as α completes the proof of Theorem 2.2.

CHAPTER 7

CONCLUSION

The foregoing account is meant to simplify and extend the results from [13, 16] (among others). In particular, the exponent in the error term from Theorem 2.1 was improved from $\frac{1}{4}$ to 1 (provided potentially more Dirichlet characters are excluded from the relevant sums, of course), and the error term was further improved by replacing $\varphi(D)$ by D. Theorem 2.2 will supply a major step in the proof of Linnik's Theorem and offers corresponding improvements to estimates for primes in arithmetic progressions mentioned in the Introduction. Careful reading of Theorem 2.1 reveals an important detail, however: that the single exceptional character involved is real if g is real. In the present circumstances there may be more than one exceptional character. A direction for further investigation is provided by the following two Theorems from [18], which classify the exceptional characters in terms of their support.

Theorem 7.1. *To each positive real B there is a further real c with the following property:*

If χ is a Dirichlet character $(\bmod D)$, $D \geq 2$, $8D^{-3B} \leq \delta \leq 1$, t, real, satisfies $|t| \leq D^B$ and S is a set of primes p in the interval $(D, x]$ for which

$$\sum_{p \in S} p^{-1} \left| 1 - \chi(p)p^{it} \right|^2 \leq \delta L,$$

where $L = \sum_{D < p \leq x} p^{-1}$, then

$$\text{either } \sum_{p \in S} p^{-1} \leq 4\delta^{1/3}L + c, \text{ or the order of } \chi \text{ is less than } 2\delta^{-1/3}.$$

Theorem 7.2. *Under the hypotheses of Theorem 7.1, if $r^3\delta \leq 1$ and χ has order at least r,*

then

$$\sum_{p \in S} \frac{1}{p} \leq \left(1 + (r^3\delta)^{1/2}\right) \frac{L}{r} + c,$$

for a suitable constant c.

The general result is of the following form: letting $\omega_r = \frac{1}{4}\min\left(\frac{1}{r^3}, \frac{\varepsilon^2}{r}\right)$, for a multi-plicative function g with *unconstrained* support S whose restriction to the interval $(D, x]$ satisfies $\beta \geq 1/r + \varepsilon$, $\beta = L^{-1}\sum_{p \in S} p^{-1}$, there is a representation

$$\sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, D)}} g(n) = \frac{1}{\varphi(D)}\sum_{\substack{n \leq x \\ (n, D) = 1}} g(n) + \frac{1}{\varphi(D)}\sum_{\substack{\chi \text{ has order} \\ 2, \dots, r-1}} \frac{\overline{\chi_1\chi}(a)}{\varphi(D)}\sum_{n \leq x} g(n)\chi_1\chi(n)$$

$$+ O\left(\frac{1}{\varphi(D)}\frac{x}{\log x}\prod_{\substack{p \leq x \\ p \in S}}\left(1 + \frac{1}{p}\right)\left(\frac{\log D}{\log x}\right)^{\omega_r}\log\left(\frac{\log x}{\log D}\right)\right),$$

with various attached uniformities that we do not pursue here.

Certain refinements are possible.

As a final remark, I elaborate on the perspective that informed this work, begun in the Introduction. From a modern viewpoint, Dirichlet's Theorem on primes in arithmetic progressions involves homomorphisms defined on the group $(\mathbb{Z}/D\mathbb{Z})^\times$ (the multiplicative group of reduced residue classes modulo D) into the complex unit circle (i.e. group characters in the dual group) and finite Fourier analysis. If we let $\mathbb{Q}^*$ denote the multiplicative group of positive rational numbers (this group may be denoted $\mathbb{Q}_+^\times$, see [19]), then any nonvanishing completely multiplicative function g may be viewed as a homomorphism of $\mathbb{Q}^*$ into the multiplicative group of nonzero complex numbers by defining

$$g\left(\frac{a}{b}\right) = \frac{g(a)}{g(b)}$$

for any $\frac{a}{b} \in \mathbb{Q}^*$. Since any $\frac{a}{b}$ in $\mathbb{Q}^*$ can be expressed uniquely as a ratio of integers $\frac{m}{n}$, where $(m, n) = 1$, we can similarly extend the definition of any nonvanishing (not necessarily

completely) multiplicative function to all of $\mathbb{Q}^*$. That this is possible reflects the fact that $\mathbb{Q}^*$ is a free group with the primes as generators. In this way one may view multiplicative functions that are never zero, with the aforementioned slight relaxing of the term, as group characters on $\mathbb{Q}^*$.

The dual group of $\mathbb{Q}^*$ is isomorphic to the direct product of denumerably many copies of $\mathbb{R}/\mathbb{Z}$, one for each prime. If we define Dirichlet characters to be one (instead of zero, contrary to classical practice) on the prime divisors of their modulus, then they are dense in this dual group. For more background on this perspective, and number-theoretic information that can be obtained from studying this dual group, see [7], in particular Chapters 15 and 16, as well as Chapter 12 in [14].

BIBLIOGRAPHY

[1] T. M. Apostol, *Introduction to analytic number theory*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1976.

[2] H. Davenport, *Multiplicative number theory*, third ed., Graduate Texts in Mathematics, vol. 74, Springer-Verlag, New York, 2000, Revised and with a preface by Hugh L. Montgomery.

[3] H. Delange, *Sur les fonctions arithmétiques multiplicatives*, Ann. Sci. École Norm. Sup. (3) **78** (1961), 273–304.

[4] P. D. T. A. Elliott, *On inequalities of large sieve type*, Acta Arith. **18** (1971), 405–422.

[5] _____, *Probabilistic number theory, I: Mean value theorems*, Grundlehren der Mathematischen Wissenschaften, vol. 239, Springer-Verlag, New York, 1979.

[6] _____, *Probabilistic number theory, II: Central limit theorems*, Grundlehren der Mathematischen Wissenschaften, vol. 240, Springer-Verlag, Berlin, 1980.

[7] _____, *Arithmetic functions and integer products*, Grundlehren der Mathematischen Wissenschaften, vol. 272, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1985.

[8] _____, *Multiplicative functions on arithmetic progressions*, Mathematika **34** (1987), no. 2, 199–206.

[9] _____, *Multiplicative functions on arithmetic progressions. II*, Mathematika **35** (1988), no. 1, 38–50.

[10] _____, *Multiplicative functions on arithmetic progressions. III. The large moduli*, A tribute to Paul Erdős, Cambridge Univ. Press, Cambridge, 1990, pp. 177–194.

[11] _____, *Multiplicative functions on arithmetic progressions. IV. The middle moduli*, J. London Math. Soc. (2) **41** (1990), no. 2, 201–216.

[12] _____, *Multiplicative functions on arithmetic progressions. V. Composite moduli*, J. London Math. Soc. (2) **41** (1990), no. 3, 408–424.

[13] _____, *Multiplicative functions on arithmetic progressions. VI. More middle moduli*, J. Number Theory **44** (1993), no. 2, 178–208.

[14] _____ , *Duality in analytic number theory*, Cambridge Tracts in Mathematics, vol. 122, Cambridge University Press, Cambridge, 1997.

[15] _____ , *The least prime primitive root and Linnik's theorem*, Number theory for the millennium, I (Urbana, IL, 2000), A K Peters, Natick, MA, 2002, pp. 393–418.

[16] _____ , *Multiplicative functions on arithmetic progressions. VII. Large moduli*, J. London Math. Soc. (2) **66** (2002), no. 1, 14–28.

[17] P. D. T. A. Elliott and J. Kish, *A large sieve inequality for Euler products*, arXiv:1203.0804 (2012).

[18] _____ , *Multiplicative functions and a taxonomy of Dirichlet characters*, arXiv:1208.0051 (2012).

[19] S. S. Gelbart, *Automorphic forms on adèle groups*, Princeton University Press, Princeton, N.J., 1975, Annals of Mathematics Studies, No. 83, also University of Tokyo Press, 1975.

[20] G. Halász, *Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen*, Acta Math. Acad. Sci. Hungar. **19** (1968), 365–403.

[21] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974, London Mathematical Society Monographs, No. 4.

[22] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), no. 1, 1–70.

[23] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, sixth ed., Oxford University Press, Oxford, 2008, Revised by D. R. Heath-Brown and J. H. Silverman.

[24] E. Landau, *Über die Äquivalenz zweier Hauptsätze der analytischen Zahlentheorie*, Sitzungsberichte der Kaiserlichen Akad. der Wiss. in Wien. Math. Naturw. Klasse **120** (1911), 1–16, see also *Collected Works*, Vol. 5, Thales Verlag, Essen, 1985, 46-61.

[25] K. Prachar, *Primzahlverteilung*, Springer-Verlag, Berlin, 1957.

[26] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics, vol. 46, Cambridge University Press, Cambridge, 1995, Translated from the second French edition (1995) by C. B. Thomas.

[27] E. Wirsing, *Das asymptotische Verhalten von Summen über multiplikative Funktionen*, Math. Ann. **143** (1961), 75–102.

[28] _____ , *Das asymptotische Verhalten von Summen über multiplikative Funktionen. II*, Acta Math. Acad. Sci. Hungar. **18** (1967), 411–467.

APPENDIX A

DETAILS

Lemma A.1. *The estimate*

$$\sum_{\substack{d \leq (\log w)^4 \\ (d, Q_c) = 1}} \frac{\Lambda(d)}{d} \ll \log\left(\frac{\log w}{\log D}\right), \qquad Q_c = \prod_{\substack{q \ prime \\ q \leq D^c}} q,$$

holds for all $w \geq 2$, $D \geq 2$, $c \geq 1$.

Proof of Lemma A.1. Similar to a standard estimate in Lemma 3.1,

$$\sum_{\substack{d \leq (\log w)^4 \\ (d, Q_c) = 1}} \frac{\Lambda(d)}{d} = \sum_{\substack{p \leq (\log w)^4 \\ (p, Q_c) = 1}} \frac{\log p}{p} + O(1).$$

For $(\log w)^4 < D^c$ the sum over primes is empty since Q_c is the product of *all* primes not exceeding D^c. This leaves $D^c \leq (\log w)^4$. Since $\log x \leq \sqrt{x}$ for all $x \geq 1$,

$$\log D^{1/2} \leq D^{1/4} \leq D^{c/4}$$

and hence

$$\left(\frac{1}{2}\log D\right)^4 \leq D^c$$

for any $c \geq 1$. If we enlarge the range of allowable primes, the preceding sum over p can

only increase. Thus

$$\sum_{\substack{p \le (\log w)^4 \\ (p,Q_c)=1}} \frac{\log p}{p} + O(1) = \sum_{D^c < p \le (\log w)^4} \frac{\log p}{p} + O(1) \qquad (p, Q_c) = 1 \text{ is superfluous now}$$

$$\le \sum_{\left(\frac{1}{2}\log D\right)^4 < p \le (\log w)^4} \frac{\log p}{p} + O(1)$$

$$= \log(\log w)^4 - \log\left(\frac{1}{2}\log D\right)^4 + O(1)$$

$$= 4\log\left(\frac{2\log w}{\log D}\right) + O(1)$$

$$\ll \log\left(\frac{\log w}{\log D}\right),$$

completing the proof.

Remark A.2. In the second step we used the facts that the n's in the sum defining $\widetilde{\mathcal{M}}$ in (4.6) are actually $\le u \le x/d$, $J \le \phi(D) \le D$, as well as $|g| \le 1$ and Lemma A.4 with $M = D^c$. In the last step we have used Lemma A.1,

$$\prod_{\substack{x/d < p \le x \\ p|Q_c}} \left(1 - \frac{1}{p}\right) \ll 1 \implies \prod_{\substack{p \le x/d \\ p|Q_c}} \left(1 - \frac{1}{p}\right) \ll \prod_{\substack{p \le x \\ p|Q_c}} \left(1 - \frac{1}{p}\right)$$

and

$$\prod_{\substack{p \le x \\ p|Q_c}} \left(1 - \frac{1}{p}\right) = \prod_{p|Q_c} \left(1 - \frac{1}{p}\right) = \prod_{p \le D^c} \left(1 - \frac{1}{p}\right) \ll \frac{1}{\log D}.$$

The first estimate is true since $d \le (\log x)^4$ so that $x/d \ge x^\eta$ for some $0 < \eta < 1$; the second equality is true since $p \mid Q_c \iff p \le D^c$ and $D^c \le x$ so the $p \le x$ is superfluous. We also used

$$\sum_{\substack{d \le (\log x)^4 \\ (d,Q_c)=1}} \Lambda(d)d^{-\varepsilon} \le \sum_{\substack{d \le (\log x)^4 \\ (d,Q_c)=1}} \Lambda(d) \le \sum_{d \le (\log x)^4} \Lambda(d) = (\log x)^4 + O(1)$$

by Lemma 3.1.

Remark A.3. In the first step we recall that the n's in the original sum defining F_3 were $\leq x(\log x)^{-2}$, but we let all $n \leq x$ in anyway (which may seem wasteful but it won't matter). We have also used $J \leq D$, Lemma A.4 for the extra $\log D^c$ in the trivial estimate of the sum, and

$$\prod_{\substack{x(\log x)^{-2} < p \leq x \\ p \mid Q_c}} \left(1 - \frac{1}{p}\right) \ll 1$$

as in Remark A.2.

Lemma A.4. *The number of integers not exceeding x which have no prime factor up to $M (\leq x)$ is $\ll x(\log M)^{-1}$.*

Proof of Lemma A.4. See [13, p. 187, line 1].

Remark A.5. For a proof that $|\lambda_d| \leq 1$ in the circumstances mentioned, see [21, p. 100] or [14, p. 225]. For the estimate regarding the quadratic form, see Lemma 25.7 in [14, p. 224] or Lemma 3.1 in [21, p. 102]. A further comment is appropriate, since in Lemma 3.6 $z = H^{\varepsilon/2}$, yet the result involves

$$\prod_{\substack{p \leq H \\ p \mid Q}} \left(1 - \frac{1}{p}\right).$$

Part of the problem is that we're not assuming that Q is made up of primes not exceeding z, but this is of no consequence. The upshot is that

$$\prod_{\substack{p \leq H^{\varepsilon/2} \\ p \mid Q}} \left(1 - \frac{1}{p}\right) = \prod_{\substack{p \leq H^{\varepsilon/2} \\ p \mid Q}} \left(1 - \frac{1}{p}\right) \prod_{\substack{H^{\varepsilon/2} < p \leq H \\ p \mid Q}} \left(1 - \frac{1}{p}\right) \prod_{\substack{H^{\varepsilon/2} < p \leq H \\ p \mid Q}} \left(1 - \frac{1}{p}\right)^{-1}$$

$$= \prod_{\substack{p \leq H \\ p \mid Q}} \left(1 - \frac{1}{p}\right) \prod_{\substack{H^{\varepsilon/2} < p \leq H \\ p \mid Q}} \left(1 - \frac{1}{p}\right)^{-1}$$

$$= \prod_{\substack{p \leq H \\ p \mid Q}} \left(1 - \frac{1}{p}\right) \cdot O_\varepsilon(1), \tag{A.1}$$

since

$$\log \prod_{\substack{a<p\leq a^m \\ p|Q}} \left(1-\frac{1}{p}\right)^{-1} = -\sum_{\substack{a<p\leq a^m \\ p|Q}} \log\left(1-\frac{1}{p}\right) = -\sum_{\substack{a<p\leq a^m \\ p|Q}} \left\{-\frac{1}{p}+O\left(\frac{1}{p^2}\right)\right\}$$

$$= \sum_{\substack{a<p\leq a^m \\ p|Q}} \frac{1}{p}+O(1) \leq \sum_{a<p\leq a^m} \frac{1}{p}+O(1)$$

$$= \sum_{p\leq a^m} \frac{1}{p} - \sum_{p\leq a} \frac{1}{p}+O(1) = \log\log a^m - \log\log a + O(1)$$

$$= \log\left(\frac{m\log a}{\log a}\right) + O(1) = \log m + O(1)$$

$$= O_m(1),$$

a fairly standard calculation. Using $a = H^{\varepsilon/2}$, $m = 2/\varepsilon$ and exponentiating establishes (A.1).

Remark A.6. Note that

$$u^\sigma = u^{1+(\log t)^{-1}} = u \cdot u^{(\log t)^{-1}} = u \cdot e^{(\log u)(\log t)^{-1}} \leq u \cdot e$$

since $u \leq t \iff \log u \leq \log t$. Hence

$$u^\sigma \leq eu \iff u^{2\sigma} \leq e^2 u^2 \iff u^{2\sigma+1} \leq e^2 u^3 \iff u^{-3} \leq e^2 u^{-2\sigma-1}.$$

Moreover,

$$\int_{t^\theta}^t \frac{du}{u(\log u)^2} = -\frac{1}{\log u}\Big|_{t^\theta}^t = -\left(\frac{1}{\log t} - \frac{1}{\log t^\theta}\right) = \frac{1}{\theta\log t} - \frac{1}{\log t} \leq \frac{1}{\theta\log t}$$

since $\theta < 1$.

Remark A.7. We used Lemma A.4 (with $M = D^c$) in the estimate of the sum over the n's after factoring out the $\log u$. Moreover, since $u \geq D^c$ (from the integral), the product over $p \mid Q_c$, $p \leq u$, is the product over *all* primes $p \leq D^c$. But

$$\prod_{p\leq D^c} \left(1-\frac{1}{p}\right) \ll \frac{1}{\log D^c} \ll \frac{1}{\log D}$$

for $c \geq 1$ by Lemma 3.1.

Lemma A.8. *The estimate*

$$\sum_{n \leq x} \left(\log(x/n) \right)^2 \ll x$$

holds for all $x \geq 2$.

Proof of Lemma A.8. For $x \geq 2$, using a Stieltjes integral,

$$\sum_{n \leq x} \left(\log(x/n) \right)^2 = \int_{2^-}^{x^+} \left(\log(x/y) \right)^2 \, d[y]$$

$$= \left(\log(x/y) \right)^2 [y] \Big]_{2^-}^{x^+} - \int_{2^-}^{x^+} [y] \cdot \frac{d}{dy} \left[\left(\log(x/y) \right)^2 \right] \, dy$$

$$= 0 + 2 \int_{2^-}^{x^+} \frac{[y]}{y} \log(x/y) \, dy$$

$$\ll \int_{2^-}^{x^+} \log(x/y) \, dy$$

$$= y \log(x) - y \log(y) + y \Big]_{2^-}^{x^+}$$

$$\ll x,$$

completing the proof (after incorporating any stray $\log x$ terms).

Remark A.9. Since there are $O_\delta(1)$ exceptional characters,

$$\sum_{\chi_j \text{ except.}} \frac{1}{\varphi(D)} \left| \sum_{\substack{x/d < m \leq x/V \\ p|m \implies p > D^c}} g(m) \chi_j(m) \right| \ll_\delta \frac{1}{\varphi(D)} \sum_{x/d < m \leq x/V} 1 \leq \frac{1}{\varphi(D)} \left(\frac{x}{V} - \frac{x}{d} \right),$$

The other sum in the brackets also does not exceed this bound upon removing the condition $p \mid m \implies p > D^c$. The extra 1 controls for there being only a single term in either sum. For any d with $d < V + U^\beta$, $\frac{x}{V+U^\beta} < \frac{x}{d}$. We enlarge the range on the sum over m, thus increasing it. For a fixed U and V,

$$\frac{x}{V} - \frac{x}{d} \leq \frac{x}{V} - \frac{x}{V + U^\beta} \leq \frac{xU^\beta}{V^2} \leq \frac{xU^\beta}{U^2} \leq \frac{x}{d^{2-\beta}}$$

since $d \geq V \geq U$. Now we use the fact that a convergent series of positive terms truncated at the front (say) is majorized by the term of highest index that was deleted:

$$\sum_{D^c/2 < d \leq 2x/D^{c_1}} \frac{\Lambda(d)}{d^{2-\beta}} \leq \sum_{D^c/2 < d} \frac{\Lambda(d)}{d^{2-\beta}} \leq \sum_{D^c/2 < n} \frac{\log n}{n^{2-\beta}} \ll \frac{\log(D^c/2)}{(D^c/2)^{2-\beta}} \leq \frac{(D^c/2)}{(D^c/2)^{2-\beta}} \ll \frac{1}{(D^c)^{1-\beta}}$$

since $2^{1-\beta} \leq 2$. Note also that, for $\varepsilon > 1$ arbitrary,

$$\int_{1^-}^{\infty} \frac{\log y}{y^\varepsilon} \, d[y] = \frac{\log y}{y^\varepsilon} \cdot [y] \Big]_1^{\infty} - \int_1^{\infty} y^{-\varepsilon-1}(1 - \varepsilon \log y)[y] \, dy$$

and both terms converge to a finite limit (depending on ε only).

Remark A.10. Since $U \leq V \leq 2U$ for any fixed U and any corresponding V,

$$\frac{1}{V} \leq \frac{1}{U} \qquad \text{and} \qquad \frac{1}{\log(x/V)} \leq \frac{1}{\log(x/2U)}.$$

Moreover,

$$\sum_V 1 \leq U^{1-\beta}$$

since for any fixed U there are U/U^β subintervals $(V, V + U^\beta)$ (of length U^β) necessary to cover the subinterval $(U, 2U]$. The U's then cancel. In addition, note that

$$\sum_{D^c/2 \leq 2^k \leq 2x/D^{c_1}} \frac{1}{\log(x/2^{k+1})} = \int_{y_1^-}^{y_2^+} \frac{1}{\log(x/2^{y+1})} \, d[y],$$

where

$$y_1 = \frac{\log(D^c/2)}{\log 2} \qquad \text{and} \qquad y_2 = \frac{\log(2x/D^{c_1})}{\log 2}.$$

The substitution $w = \log x - (y+1)\log 2$ (say) is useful in evaluating the integral.

Lemma A.11. *For $D \geq 1$, the estimate*

$$\sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, D)}} \log\left(\frac{x}{n}\right) \ll \frac{x}{\varphi(D)}$$

holds uniformly for $(a, D) = 1$, $x \geq 1$.

Proof of Lemma A.11. Noting that

$$\frac{d}{dy}\left[\log\left(\frac{x}{y}\right)\right] = -\frac{1}{y},$$

using a Stieltjes integral we have

$$\sum_{\substack{n\le x\\ n\equiv a\,(\mathrm{mod}\,D)}} \log\left(\frac{x}{n}\right) = \int_{1^-}^{x^+} \log\left(\frac{x}{y}\right) d\left[\sum_{\substack{n\le y\\ n\equiv a\,(\mathrm{mod}\,D)}} 1\right]$$

$$= \log\left(\frac{x}{y}\right)\cdot \sum_{\substack{n\le y\\ n\equiv a\,(\mathrm{mod}\,D)}} 1\Bigg]_{1^-}^{x^+} - \int_{1^-}^{x^+} \sum_{\substack{n\le y\\ n\equiv a\,(\mathrm{mod}\,D)}} 1\cdot \frac{d}{dy}\left[\log\left(\frac{x}{y}\right)\right]dy$$

$$= \int_1^x \frac{1}{y} \sum_{\substack{n\le y\\ n\equiv a\,(\mathrm{mod}\,D)}} 1\,dy \tag{A.2}$$

$$= \int_1^D \frac{1}{y} \sum_{\substack{n\le y\\ n\equiv a\,(\mathrm{mod}\,D)}} 1\,dy + \int_D^x \frac{1}{y} \sum_{\substack{n\le y\\ n\equiv a\,(\mathrm{mod}\,D)}} 1\,dy$$

$$\ll \int_1^D \frac{1}{y}\cdot 1\,dy + \int_D^x \frac{1}{y}\cdot \frac{y}{\varphi(D)}\,dy$$

$$= \log D + \frac{x}{\varphi(D)} - D \le \frac{x}{\varphi(D)},$$

almost completing the proof. It should be mentioned that it is possible to pick up an extra $\log x$ in step (A.2), but this is $\ll \frac{x}{\varphi(D)}$ as long as D is less than x^η, $0 < \eta < 1$, which is true in our circumstances.

Remark A.12. For if g_1 is completely multiplicative, then

$$(\mu g_1 * g_1)(n) = \sum_{d|n} \mu(d)g_1(d)g_1(n/d) = \sum_{d|n} \mu(d)g_1(n) = g_1(n)\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise} \end{cases}$$

since $g_1(1) = 1$, and this is exactly the arithmetic function that is the identity with respect to Dirichlet convolution. See Theorem 2.17 of [1, p. 36].