# Topics in multiplicative number theory




Jared Duker Lichtman

Balliol College

University of Oxford

Supervisor: Prof. James Maynard

A thesis submitted for the degree of

*Doctor of Philosophy*

July 2023

# Abstract

In this thesis we use a blend of analytic and combinatorial techniques to address a variety of problems about primes and related multiplicative structures:

We introduce a modified linear sieve, and obtain new upper bound for twin primes, refining a 2004 result of Wu. This represents the largest improvement on the problem since 1986.

We prove the equidistribution of primes $p$ in arithmetic progressions to moduli of size $p^{0.5313}$, and consequently obtain the infinitude of shifted primes $p-1$ without prime factors above $p^{0.2844}$, refining a 1998 result of Baker and Harman.

We introduce a new hybrid conjecture of the famous conjectures of Hardy–Littlewood and Chowla on correlations of the Möbius and von Mangoldt functions. We prove several 'on average' results on the problem (in part jointly with J. Teräväinen), extending work of Matomäki, Radziwiłł, and Tao.

We study primitive sets. The *Erdős primitive set conjecture*, posed in 1986, asserts that the set of primes is maximal among all primitive sets, in a precise sense. We explore variations of this problem in multiple directions, where many open questions remain. We disprove a natural generalization of the conjecture, due to Banks and Martin in 2013, by showing the $k$-almost primes are minimal when $k = 6$, in the same sense of Erdős. We also show a translated analogoue of the conjecture is false already for translates $h \approx 1.04$, highlighting the subtlety of the original problem.

Finally, this thesis culminates with a proof of the Erdős primitive set conjecture.

Dedicated to Jonina Duker

**תייבל בת רחל לאה**

**ושלמה יהושע**

# Acknowledgements

For me it is deeply meaningful to be part of such a remarkable research group in analytic number theory, during my time here at Oxford.

I have immense gratitude to James Maynard for his supervision, insight, and many fruitful discussions over the years. I would also like to thank Ben Green, Jon Keating, and Joni Teräväinen for their mentorship roles throughout my degree, as well as Carl Pomerance for his unwavering and enduring support.

Beyond explicit teachings, my experience has been greatly enriched by constant contact among members of our group, absorbing tacit knowledge in the intellectual atmosphere. In this regard, I thank my academic siblings Oliver McGrath, Julia Stadlmann, Alex Pascadi, and Cédric Pilatte, along with the wisdom I've gleaned from postdoctoral members Thomas Bloom, Ofir Gorodetsky, Lasse Grimmelt, Jori Merikoski, Sarah Peluse, and Kyle Pratt.

I would like to give thanks to my collaborators over the years, including Jonathan Bayless, Tim Browning, Tsz Ho Chan, Sary Drappeau, Paul Kinlaw, and Mo Dick Wong.

Finally, I thank my parents Jonina Duker and Alan D. Lichtman, who at great sacrifice, have acted and continue to act in my best interest.

My doctorate was supported by the Clarendon Scholarship (from Oxford University Press), and Marvin Bower Scholarship (from Balliol College). I also received funding from the Oxford Mathematical Institute, G-Research, and the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (grant agreement No. 851318).

# Introduction

A whole number is *prime* if its only divisors are one and itself. This simple definition becomes of foundational importance through the Fundamental Theorem of Arithmetic, which states that every whole number greater than one factors uniquely into prime numbers (up to permutation).

The prime numbers have fascinated mathematicians since Antiquity, where Euclid of Alexandria proved the first significant theorem, showing there are infinitely many primes. There remain many basic questions about prime numbers which have gone unsolved for decades (and sometimes millenia!). Over time, mathematicians have developed an array of tools to investigate these questions, which have proved fruitful for gaining a more structured understanding of the mathematical landscape.

In this thesis we use a blend of analytic and combinatorial techniques to address a number of such basic problems about primes, as well as related multiplicative structures. The thesis consists of twelve chapters, which may be thematically grouped in three:

I. In Chapters 1–4 we study *primes* with prescribed patterns. This includes twin primes, tuples of (simultaneous) primes, primes in arithmetic progressions, and shifted primes without large prime factors.

II. In Chapters 5–8 we study *$k$-almost primes*: numbers with exactly $k$ prime factors, counting multiplicity. We focus in particular on understanding aspects of their distributions as $k$ varies.

III. In Chapters 9–12 we study *primitive sets*: sets of numbers, no member of which divides another. This is a large class of sets, which includes our objects of study in the prior chapters, i.e. the primes and $k$-almost primes.

**Chapter 1:** *A modification of the linear sieve, and the count of twin primes*, J. D. Lichtman, Algebra & Number Theory, to appear.

We introduce a modification of the linear sieve whose weights satisfy strong factorization properties, and consequently equidistribute primes up to size $x$ in arithmetic progressions to moduli up to $x^{10/17}$. This surpasses the level of distribution $x^{4/7}$ with the linear sieve weights from well-known 1986 work of Bombieri, Friedlander, and Iwaniec, and which was recently extended to $x^{7/12}$ by Maynard. As an application, we obtain a new upper bound on the count of twin primes, refining a 2004 result of Wu by about 3%. This represents the largest improvement on the problem since 1986.

**Chapter 2:** *Primes in arithmetic progressions to large moduli, and shifted primes without large prime factors*,
J. D. Lichtman, Mathematische Annalen, to appear.

We prove the infinitude of shifted primes $p-1$ without prime factors above $p^{0.2844}$. This refines $p^{0.2961}$ from Baker and Harman in 1998. Consequently, we obtain an improved lower bound on the distribution of Carmichael numbers.

Our main technical result is a new mean value theorem for primes in arithmetic progressions to large moduli. Namely, we estimate primes of size $x$ with quadrilinear forms of moduli up to $x^{17/32}$. This extends moduli beyond $x^{11/21}$, recently obtained by Maynard, improving $x^{29/56}$ from well-known 1986 work of Bombieri, Friedlander, and Iwaniec.

**Chapter 3:** *Averages of the Möbius function on shifted primes*,
J. D. Lichtman, Quarterly Journal of Mathematics (2021), 1–29.

It is a folklore conjecture that the Möbius function exhibits cancellation on shifted primes; that is, $\sum_{p \leq X} \mu(p+h) = o(\pi(X))$ as $X \to \infty$ for any fixed shift $h > 0$. This appears in print at least since Hildebrand in 1989. We prove the conjecture on average for shifts $h \leq H$, provided $\log H / \log \log X \to \infty$. We also obtain results for shifts of prime $k$-tuples, and for higher correlations of Möbius with von Mangoldt and divisor functions. Our argument combines sieve methods with a refinement of Matomäki, Radziwiłł, and Tao's work on an averaged form of Chowla's conjecture.

**Chapter 4:** *On the Hardy–Littlewood–Chowla conjecture on average*, J. D. Lichtman, J. Teräväinen, Forum of Mathematics, Sigma, 10 (2022), 17 pp.

There has been recent interest in a hybrid form of the celebrated conjectures of Hardy–Littlewood and of Chowla. We prove that for any $k, \ell \geq 1$ and distinct integers $h_1, \dots, h_{k-1}, a_1, \dots, a_\ell$, we have

$$\sum_{n \leq X} \mu(n+h_1) \cdots \mu(n+h_k)\Lambda(n+a_1) \cdots \Lambda(n+a_\ell) = o(X)$$

for all except $o(H)$ values of $h_1 \leq H$, so long as $H \geq (\log X)^{\ell - o(1)}$. This improves on the range $H \geq (\log X)^{\psi(X)}$, $\psi(X) \to \infty$ obtained in previous work of the first author. Our results also generalize from the Möbius function $\mu$ to an arbitrary (non-pretentious) multiplicative function $f$, provided $H \geq (\log X)^{\ell}$. This improves on the range $H \geq \exp((\log X)^{5/8+o(1)})$, implied by Fourier uniformity bounds of Matomäki, Radziwiłł, Tao, Ziegler, and the second author.

**Chapter 5:** *Mertens' prime product formula, dissected*, J. D. Lichtman, Integers, 21A (2021), Ron Graham Memorial Volume, #A17, 15 pp.

In 1874, Mertens famously proved an asymptotic formula for the product $p/(p-1)$ over all primes $p$ up to $x$. On the other hand, one may expand Mertens' prime product into series of over numbers $n$ with only small prime factors. It is natural to restrict such series to numbers $n$ with a fixed number $k$ of prime factors. In this article, we obtain asymptotic formulae for these series for each $k$, which together dissect Mertens' original formula. The proof is by elementary methods of a combinatorial flavor.

**Chapter 6:** *Higher Mertens constants for almost primes*, J. Bayless, P. Kinlaw, J. D. Lichtman, Journal of Number Theory, 234 (2022), 448–475.

For $k \geq 1$, a $k$-almost prime is a positive integer with exactly $k$ prime factors, counted with multiplicity. In this article we give elementary proofs of precise asymptotics for the reciprocal sum of $k$-almost primes. Our results match the strength of those of classical analytic methods. We also study the limiting behavior of the constants appearing in these estimates, which may be viewed as higher analogues of the Mertens constant $\beta =$

0.2614... Further, in the case $k = 2$ of semiprimes we give yet finer-scale and explicit estimates, as well as a conjecture.

**Chapter 7:** *Almost primes and the Banks–Martin conjecture*,
J. D. Lichtman, Journal of Number Theory, 211 (2020), 513–529.

It has been known since Erdős that the sum of $1/(n \log n)$ over numbers $n$ with exactly $k$ prime factors (with repetition) is bounded as $k$ varies. We prove that as $k$ tends to infinity, this sum tends to 1. Banks and Martin have conjectured that these sums decrease monotonically in $k$, and in earlier papers this has been shown to hold for $k$ up to 3. However, we show that the conjecture is false in general, and in fact a global minimum occurs at $k = 6$.

**Chapter 8:** *On Erdős sums of almost primes*,
O. Gorodetsky, J. D. Lichtman, M. D. Wong, submitted.

In 1935, Erdős proved the sums $f_k = \sum_n 1/(n \log n)$, over integers $n$ with exactly $k$ prime factors, are uniformly bounded, and in 1993 Zhang proved $f_k$ is maximized by the prime sum $f_1 = \sum_p 1/(p \log p)$. According to a 2013 conjecture of Banks and Martin, the sums $f_k$ are predicted to decrease monotonically in $k$. In this article, we show the sums restricted to odd integers are indeed monotonically decreasing in $k$, sufficiently large. By contrast, contrary to the conjecture we prove the sums $f_k$ increase monotonically in $k$, sufficiently large.

Our main result gives an asymptotic for $f_k$ which identifies the (negative) secondary term, i.e. $f_k = 1 - (a + o(1))k^2/2^k$ for an explicit constant $a = .0656\cdots$. This is proven by a refined method combining real and complex analysis, while the classical results of Sathe and Selberg on products of $k$ primes imply the weaker estimate $f_k = 1 + O_\varepsilon(k^{\varepsilon - 1/2})$. We also give an alternate, probability-theoretic argument related to the Dickman distribution. Here the proof reduces to showing a new sequence of integrals converges to $e^{-\gamma}$, which may be of independent interest.

**Chapter 9:** *A generalization of primitive sets and a conjecture of Erdős*, T. H. Chan, J. D. Lichtman, C. Pomerance, Discrete Analysis 2020:16, 13 pp.

A set of integers greater than 1 is *primitive* if no element divides another. Erdős proved in 1935 that the sum of $1/(n \log n)$ for $n$ running over a primitive set $A$ is universally bounded over all choices for $A$. In 1986 he asked if this universal bound is attained by the set of prime numbers. We answer the Erdős question in the affirmative for 2-primitive sets. Here a set is *2-primitive* if no element divides the product of 2 other elements.

**Chapter 10:** *On the critical exponent for k-primitive sets*, T. H. Chan, J. D. Lichtman, C. Pomerance, Combinatorica (2021), 19 pp.

A set of positive integers is primitive (or 1-primitive) if no member divides another. Erdős proved in 1935 that the weighted sum $\sum 1/(n \log n)$ for $n$ ranging over a primitive set $A$ is universally bounded over all choices for $A$. In 1986 he asked if this universal bound is attained by the set of prime numbers. One source of difficulty in this conjecture is that $\sum n^{-\lambda}$ over a primitive set is maximized by the primes if and only if $\lambda$ is at least the critical exponent $\tau_1 \approx 1.14$.

A set is $k$-primitive if no member divides any product of up to $k$ other distinct members. One may similarly consider the critical exponent $\tau_k$ for which the primes are maximal among $k$-primitive sets. In recent work the authors showed that $\tau_2 < 0.8$, which directly implies the Erdős conjecture for 2-primitive sets. In this article we study the limiting behavior of the critical exponent, proving that $\tau_k$ tends to zero as $k \to \infty$.

**Chapter 11:** *Translated sums of primitive sets*,
J. D. Lichtman, Comptes Rendus. Mathématique, 360 (2022), 409–414.

The Erdős primitive set conjecture states that the sum $f(A) = \sum_{a \in A} \frac{1}{a \log a}$, ranging over any primitive set $A$ of positive integers, is maximized by the set of prime numbers. Recently Laib, Derbal, and Mechik proved that the translated Erdős conjecture for the sum $f(A, h) = \sum_{a \in A} \frac{1}{a(\log a + h)}$ is false starting at $h = 81$, by comparison with semiprimes. In this note we prove that such falsehood occurs already at $h = 1.04\cdots$, and show

this translate is best possible for semiprimes. We also obtain results for translated sums of $k$-almost primes with larger $k$.

**Chapter 12:** *A proof of the Erdős primitive set conjecture*,
J. D. Lichtman, Forum of Mathematics, Pi, 11 (2023) 22 pp.

A set of integers greater than 1 is primitive if no member in the set divides another. Erdős proved in 1935 that the series $f(A) = \sum_{a \in A} 1/(a \log a)$ is uniformly bounded over all choices of primitive sets $A$. In 1986 he asked if this bound is attained for the set of prime numbers. In this article we answer in the affirmative.

As further applications of the method, we make progress towards a question of Erdős, Sárközy, and Szemerédi from 1968. We also refine the classical Davenport–Erdős theorem on infinite divisibility chains, and extend a result of Erdős, Sárközy, and Szemerédi from 1966.

# Contents

**1 A modification of the linear sieve, and the count of twin primes 18**
1.1 Introduction . . . 19
1.1.1 Application to twin primes . . . 21
1.1.2 Outline of main ideas in Theorem 1.1.1 . . . 22
1.2 Technical setup and results . . . 23
1.2.1 Factorization of weights and their level of distribution . . . 23
1.2.2 Sieve theory setup and bounds . . . 26
1.3 Programmably factorable support . . . 28
1.3.1 Refined factorization of $\mathcal{D}^{\text{well}}$ . . . 34
1.4 Modification of the linear sieve . . . 36
1.4.1 Sieve function computation . . . 40
1.5 Factorable remainder, after Iwaniec . . . 41
1.5.1 Variable level of distribution for the linear sieve weights . . . 46
1.5.2 Equidistribution for products of primes . . . 47
1.6 Upper bound for twin primes . . . 48
1.6.1 Lemmas . . . 49
1.6.2 Computations . . . 51
1.6.3 Completing the proof of Theorem 1.1.2 . . . 57

**2 Primes in arithmetic progressions to large moduli, and shifted primes without large prime factors 60**
2.1 Introduction . . . 61
2.1.1 Primes in arithmetic progressions to large moduli . . . 62
2.2 Proof outline . . . 63
2.3 Shifted primes without large prime factors . . . 65
2.4 Main propositions . . . 69
2.5 Preliminary Lemmas . . . 72
2.6 Exponential sum estimates . . . 73

2.7 Zhang-style estimates . . . 76
2.8 Proof of Proposition 2.4.1 (type II estimate) . . . 82
2.9 Proof of Proposition 2.4.3 (4 prime factors) . . . 84
2.10 Proof of Proposition 2.4.2 (sieve asymptotics) . . . 84
2.11 Proof of Proposition 2.4.4 (3 prime factors) . . . 88

**3 Averages of the Möbius function on shifted primes 90**
3.1 Introduction . . . 91
3.1.1 Higher correlations . . . 92
3.1.2 Correlations with divisor functions . . . 93
3.1.3 Beyond Möbius . . . 94
3.1.4 Overview of the proof of Theorem 3.1.1 . . . 94
3.2 Initial reductions . . . 96
3.3 Key Fourier estimate . . . 101
3.3.1 Minor arc . . . 102
3.3.2 Major arc . . . 104
3.4 Preparatory lemmas . . . 106
3.5 Mean value of multiplicative functions . . . 110
3.5.1 Mean value of Dirichlet polynomials . . . 111
3.6 Average Chowla-type correlations . . . 114
3.6.1 Deduction of results . . . 116
3.6.2 Almost all shifts . . . 117
3.7 Non-pretentious multiplicative functions . . . 118

**4 On the Hardy–Littlewood–Chowla conjecture on average 121**
4.1 Introduction . . . 122
4.1.1 Other multiplicative functions . . . 124
4.1.2 Outline of the proof . . . 125
4.2 Auxiliary results . . . 126
4.2.1 A short exponential sum estimate . . . 126
4.2.2 Upper-bounding correlations of primes . . . 128
4.2.3 Correlations of primes and integers having typical factorization 133
4.3 A Fourier-analytic argument . . . 135
4.4 Proofs of the main theorems . . . 138

**5 Mertens' prime product formula, dissected** **141**
5.1 Introduction . . . 142
5.1.1 Uniform estimates via complex analysis . . . 143
5.1.2 The coefficients $c_k$ . . . 145
5.2 Elementary combinatorial proof for $k$ fixed . . . 146
5.3 Combinatorial proof of asymptotics for coefficients $c_k$ . . . 148
5.4 Analytic proof for $k$ in uniform range . . . 153

**6 Higher Mertens constants for almost primes** **155**
6.1 Introduction . . . 156
6.1.1 Finer-scale estimates for semiprimes. . . . 158
6.2 Outline and perspectives on the argument . . . 161
6.2.1 Outline of Theorem 6.1.1 . . . 163
6.3 The Proof of Theorem 6.2.3 . . . 164
6.4 Proof of Proposition 6.3.0.3 . . . 167
6.5 The Proof of Theorem 6.1.1 . . . 169
6.6 Proof of Equivalence of Theorems 6.2.1 and 6.2.2 . . . 171
6.7 The Proof of Theorem 6.1.1 . . . 172
6.8 The Proofs of Theorems 6.1.2 and 6.1.3 . . . 176
6.9 A new proof of Theorem 6.1.4 . . . 180

**7 Almost primes and the Banks–Martin conjecture** **182**
7.1 Introduction . . . 183
7.2 Statement of results . . . 184
7.3 Zeta function method for small $k$ . . . 185
7.4 Asymptotic behavior for large $k$ . . . 188
7.5 Further progress via zeta functions for large $k$ . . . 191
7.5.1 A global minimum via zeta functions . . . 195
7.5.2 Remark on integration . . . 196

**8 On Erdős sums of almost primes** **198**
8.1 Introduction . . . 199
8.1.1 A sequence of integrals . . . 201
8.1.2 Proof of monotonicity results . . . 201
8.1.3 Proof method: permutations . . . 202
8.2 Sifted sums . . . 203
8.2.1 Preparation . . . 203

8.2.2 Proof of Theorem 8.1.3 . . . 205
8.2.3 Proof of Lemma 8.2.2 . . . 207
8.2.4 Proof of Lemma 8.2.3 . . . 208
8.3 Proof of Theorem 8.1.2 . . . 211
8.3.1 Proof of Lemma 8.3.2 . . . 212
8.3.2 Proof of Lemma 8.3.1 . . . 212
8.3.3 Proof of Lemma 8.3.3 . . . 214
8.4 Probability-theoretic argument . . . 217
8.4.1 Preliminary lemmas . . . 217
8.4.2 Poof of Proposition 8.4.1 . . . 220
8.5 A sequence of integrals . . . 223
8.5.1 Probabilistic setup . . . 223
8.5.2 Proof of Lemma 8.4.7 . . . 226
8.5.3 A second proof that $I_n$ tends to $e^{-\gamma}$ . . . 229

**9 A generalization of primitive sets and a conjecture of Erdős 231**
9.1 Introduction . . . 232
9.2 Combinatorial Lemmas . . . 234
9.3 Proof of Theorem 9.1.2 . . . 236

**10 On the critical exponent for $k$-primitive sets 243**
10.1 Introduction . . . 244
10.1.1 Generalizations . . . 246
10.1.2 $\infty$-primitive sets . . . 247
10.2 Preliminary lemmas . . . 248
10.3 Theorem for $k$-primitive sets . . . 250
10.4 Theorem for lcm $k$-primitive sets . . . 255
10.5 Theorem for strongly $k$-primitive sets . . . 257

**11 Translated sums of primitive sets 260**
11.1 Introduction . . . 261
11.2 Proof of Theorem 1 . . . 261
11.2.1 Computations . . . 264
11.3 Proof of Theorem 11.1.2 . . . 265

**12 A proof of the Erdős primitive set conjecture** **268**
12.1 Introduction . . . . . 269
12.1.1 Proof outline of Theorem 12.1.2 . . . . . 272
12.1.2 L-primitive sets . . . . . 273
12.1.3 Density and divisibility chains . . . . . 274
12.2 Preliminaries on L-primitive sets . . . . . 275
12.3 Primitive sets . . . . . 278
12.4 Deduction of Theorems 12.1.2, 12.1.3, 12.1.5 . . . . . 280
12.4.1 Proof of Theorem 12.1.5 . . . . . 284
12.5 L-primitive sets revisited . . . . . 285
12.5.1 Upper density . . . . . 285
12.5.2 The Erdős L-primitive set conjecture . . . . . 286
12.6 Deduction of Theorems 12.1.8, 12.1.10 . . . . . 289
12.7 Closing remarks . . . . . 293

**A Mathematica code** **295**

**Bibliography** **296**

# Notation

Let $\Omega(n)$ denote the number of prime factors of $n$, counted with repetition. We will use the Vinogradov $\ll$ and $\gg$ asymptotic notation, and the big oh $O(\cdot)$ and $o(\cdot)$ asymptotic notation. $f \asymp g$ will denote the conditions $f \ll g$ and $g \ll f$ both hold. We will also use the standard notation $f \sim g$ to denote $f = (1 + o(1))g$, as well as the (non-standard) notation $f \lesssim g$ and $f \gtrsim g$ to denote $f \leq (1 + o(1))g$ and $f \geq (1 + o(1))g$, respectively. Dependence on a parameter will be denoted by a subscript.

Let $\mathcal{P}$ denote the set of prime numbers. A variable denoted $p$ will always be assumed a prime, and $\pi(x) := \#\{p \leq x\}$ denotes the count of primes up to $x$. We use $\phi$ to denote the Euler totient function, $e(x) := e^{2\pi i x}$ the complex exponential, $\tau_k(n)$ the $k$-fold divisor function, $\mu(n)$ the Möbius function. We let $p^-(n)$, $P^+(n)$ denote the smallest and largest prime factors of $n$, respectively. Also let $P_j^+(n)$ denote the $j$th largest prime factor of $n$, so $P_1^+(n) = P^+(n)$.

Let $\hat{f}$ denote the Fourier transform of $f$, i.e. $\hat{f}(\xi) = \int_{-\infty}^{\infty} f(t)e(-\xi t)dt$. We use $\mathbf{1}$ to denote the indicator function of a statement. For example,

$$\mathbf{1}_{n \equiv a \pmod q} = \begin{cases} 1 & \text{if } n \equiv a \pmod q, \\ 0 & \text{else.} \end{cases}$$

We will use $(a, b)$ to denote $\gcd(a, b)$ when it does not conflict with notation for ordered pairs. For $(n, q) = 1$, we will use $\overline{n}$ to denote the inverse of the integer $n$ modulo $q$; the modulus will be clear from the context. For example, we may write $e(a\overline{n}/q)$ - here $\overline{n}$ is interpreted as the integer $m \in \{0, \dots, q-1\}$ such that $mn \equiv 1 \pmod q$. Occasionally we will also use $\overline{\lambda}$ to denote complex conjugation; the distinction of the usage should be clear from the context. For a complex sequence $\alpha_{n_1,\dots,n_k}$, $\|\alpha\|_2$ will denote the $\ell^2$ norm $\|\alpha\|_2 = (\sum_{n_1,\dots,n_k} |\alpha_{n_1,\dots,n_k}|^2)^{1/2}$.

Summations assumed to be over all positive integers unless noted otherwise. In a summation, the notation $n \sim N$ denotes the conditions $N < n \leq 2N$.

Throughout this thesis, we apply partial summation [80, §1.5] (sometimes called Abel summation) and often use the following standard notation.

**Partial summation.** Let $f : \mathbb{R} \to \mathbb{C}$ be a continuously differentiable function. For a sequence $a_n \in \mathbb{C}$, denote $A(x) = \sum_{n \leq x} a_n$. Then we have

$$\sum_{y < n \leq x} a_n f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)\mathrm{d}t.$$

## Notation in Chapter 2

In Chapter 2 specifically, we also adopt the following conventions:

We will let $z_0 := x^{1/(\log\log x)^3}$ and $y_0 := x^{1/\log\log x}$ two parameters depending on $x$, which we will think of as a large quantity. We will let $\psi_0 : \mathbb{R} \to \mathbb{R}$ denote a fixed smooth function supported on $[1/2, 5/2]$ which is identically equal to 1 on the interval $[1, 2]$ and satisfies the derivative bounds $\|\psi_0^{(j)}\|_\infty \ll (4^j j!)^2$ for all $j \geq 0$. See [14, Page 368, Corollary to Lemma 9] for the construction of such a function $\psi_0$.

**Definition** (Siegel-Walfisz condition)**.** We say that a complex sequence $\alpha_n$ satisfies the **Siegel-Walfisz condition** if for every $d \geq 1$, $q \geq 1$ and $(a, q) = 1$ and every $A > 1$ we have

$$\Big| \sum_{\substack{n \sim N \\ n \equiv a \pmod q \\ (n,d)=1}} \alpha_n - \frac{1}{\phi(q)} \sum_{\substack{n \sim N \\ (n,dq)=1}} \alpha_n \Big| \ll_{A,B_0} \frac{N\tau(d)^{B_0}}{(\log N)^A}. \tag{1}$$

We note that $\alpha_n$ certainly satisfies the Siegel-Walfisz condition if $\alpha_n = 1$, if $\alpha_n = \mu(n)$, or if $\alpha_n$ is the indicator function of the primes.

# Chapter 1

# A modification of the linear sieve, and the count of twin primes

## 1.1 Introduction

Given a finite set $\mathcal{A}$ of positive integers, sieve methods offer a broad framework for estimating the number of elements in $\mathcal{A}$ all whose prime factors exceed $z$, denoted $S(\mathcal{A}, z)$, in terms of the approximate density $g_{\mathcal{A}}(d) = g(d)$ of multiples of $d$ in $\mathcal{A}$, denoted $\mathcal{A}_d$. Note one often expects $S(\mathcal{A}, z) \approx |\mathcal{A}| \prod_{p<z} \big(1 - g(p)\big)$. Combinatorial sieves may be viewed as refinements of the basic inclusion-exclusion principle, and are described by a sequence of weights $\lambda(d) \in \{-1, 0, 1\}$ supported on integers up to some level $D \geq 1$. We refer the reader to Opera de Cribro [62] for a more thorough introduction to the subject.

In particular, the upper bound weights $\lambda^+(d)$ for the linear sieve satisfy

$$S(\mathcal{A}, z) \;\leq\; |\mathcal{A}| \prod_{p<z} \big(1 - g(p)\big) \Big(F\big(\tfrac{\log D}{\log z}\big) + o(1)\Big) \;+ \sum_{\substack{d \leq D \\ p|d \Rightarrow p<z}} \lambda^+(d) \big(|\mathcal{A}_d| - |\mathcal{A}| g(d)\big) \tag{1.1}$$

as $D \to \infty$, provided $g = g_{\mathcal{A}}$ satisfies some mild conditions. Here the function $F : \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 1}$ is defined by a delay-differential equation, as in (1.10). For sets $\mathcal{A}$ sufficiently equidistributed in arithmetic progressions the second sum over $d \leq D$ in (1.1) contributes negligibly, in which case the main term is $S(\mathcal{A}, z) \lesssim |\mathcal{A}| \prod_{p<z} \big(1 - g(p)\big) F(s)$, where $z = D^{1/s}$. In fact, $F(s) \to 1$ as $s \to \infty$ so the main term confirms the naïve expectation in this case. Moreover, $F$ is optimal in the sense that the bound (1.1) is attained sharply for a particular set $\mathcal{A}$.

We introduce this sieve theory setup more formally in §1.2.2 below, and define the sieve weights $\lambda^+$ explicitly in §1.3. See [62, §12] for further details on the linear sieve ($\beta = 2$), as well as [62, §11] for its generalization to the $\beta$-sieve.

The linear sieve is powerful when combined with equidistribution estimates which make the final sum in (1.1) small. For example, the Bombieri–Vinogradov Theorem shows that for every $\varepsilon, A > 0$, letting $Q = x^{\frac{1}{2}-\varepsilon}$ we have

$$\sum_{q \leq Q} \sup_{(a,q)=1} \Big|\pi(x; q, a) - \frac{\pi(x)}{\phi(q)}\Big| \;\ll_{\varepsilon, A}\; \frac{x}{(\log x)^A}. \tag{1.2}$$

So by taking $D = Q$, (1.1) can give a good upper bound when the set $\mathcal{A}$ is related to the primes, such as when $\mathcal{A} = \{p + 2 : p \leq x\}$, in which case (1.1) gives an upper bound for the count of twin primes.

The estimate (1.2) may be viewed as an assertion of the Generalized Riemann Hypothesis on average over moduli up to $Q = x^{\frac{1}{2}-\varepsilon}$. It remains an important open

problem to extend the range to $Q = x^{\frac{1}{2}+\delta}$ for some fixed $\delta > 0$. Indeed, Elliott and Halberstam [39] conjectured such an extension up to $Q = x^{1-\varepsilon}$ for any $\varepsilon > 0$.

In some contexts it suffices to relax the setup in (1.2) in order to raise the level of distribution. In particular, in the case of a fixed residue class $a \in \mathbb{Z}$, and the absolute values replaced by well-factorable weights $\lambda(q)$ (c.f. Def. 1.2.1), the celebrated result of Bombieri–Friedlander–Iwaniec [13] raised the level up to $Q = x^{\frac{4}{7}-\varepsilon}$,

$$\sum_{\substack{q \leq Q \\ (q,a)=1}} \lambda(q) \Big( \pi(x; q, a) - \frac{\pi(x)}{\phi(q)} \Big) \ \ll_{a,A,\varepsilon} \ \frac{x}{(\log x)^A}. \tag{1.3}$$

While the linear sieve weights are not themselves well-factorable, Iwaniec [79] constructed a well-factorable variant $\widetilde{\lambda}^+$ of the weights $\lambda^+$ (and so (1.3) holds with $\lambda = \widetilde{\lambda}^+$), which are only slightly altered from $\lambda^+$ so that $\widetilde{\lambda}^+$ enjoys an analogous linear sieve bound as in (1.1), notably with an identical form of the main term,

$$S(\mathcal{A}, z) \ \leq \ |\mathcal{A}| \prod_{p<z} \big(1 - g(p)\big) \Big( F\big(\tfrac{\log D}{\log z}\big) + o(1) \Big) \ + \sum_{\substack{d \leq D \\ p|d \Rightarrow p<z}} \widetilde{\lambda}^+(d) \big( |\mathcal{A}_d| - |\mathcal{A}| g(d) \big). \tag{1.4}$$

The bound (1.3) stood for several decades, but quite recently Maynard [102] managed to extend the level in (1.3) further to $Q = x^{\frac{7}{12}-\varepsilon}$ in the case of the weights $\lambda = \widetilde{\lambda}^+$. Given the currently available equidistribution estimates for primes, we note the level $x^{\frac{7}{12}}$ is a natural barrier for these weights.

In this chapter, we modify the technical construction of the linear sieve weights to avoid this barrier, and thereby produce new sieve weights that induce stronger equidistribution estimates for primes.

**Theorem 1.1.1.** *Let $D = x^{\frac{10}{17}-\varepsilon}$. There exists a sequence $\widetilde{\lambda}^*(d) \in \{-1, 0, 1\}$ satisfying:*

(1) Equidistribution for primes: *for any fixed $a \in \mathbb{Z}$, $A, \varepsilon > 0$, we have*

$$\sum_{\substack{d \leq D \\ (d,a)=1}} \widetilde{\lambda}^*(d) \Big( \pi(x; d, a) - \frac{\pi(x)}{\phi(d)} \Big) \ \ll_{a,A,\varepsilon} \ \frac{x}{(\log x)^A}.$$

(2) Sieve upper bound: *for $s \geq 1$, $z = D^{1/s}$, we have*

$$S(\mathcal{A}, z) \ \leq \ |\mathcal{A}| \prod_{p<z} \big(1 - g(p)\big) \big( F^*(s) + o(1) \big) \ + \sum_{\substack{d \leq D \\ p|d \Rightarrow p<z}} \widetilde{\lambda}^*(d) \big( |\mathcal{A}_d| - |\mathcal{A}| g(d) \big),$$

*where $F^*(s) \leq 1.000081\, F(s)$ when $1 \leq s \leq 3$, for the linear sieve function $F$ as in (1.10).*

The key feature of Theorem 1.1.1 is to obtain equidistribution up to level $x^{\frac{10}{17}}$ at the cost of only a tiny loss in the main term. See Theorem 1.2.10 and Proposition 1.5.4 for full technical statements and additional variations that may be of independent interest.

### 1.1.1 Application to twin primes

We expect that Theorem 1.1.1 should give numerous improvements to sieve bounds related to the primes. As proof of concept in this direction, we give a new upper bound for the count of twin primes up to $x$, denoted $\pi_2(x)$. Recall Hardy and Littlewood [72] conjectured the asymptotic formula

$$\pi_2(x) \sim \frac{2x}{(\log x)^2} \prod_{p>2} \frac{1-2/p}{(1-1/p)^2} =: \Pi(x). \tag{1.5}$$

**Theorem 1.1.2.** *As $x$ tends to infinity, we have*

$$\pi_2(x) \lesssim 3.29956\,\Pi(x).$$

Theorem 1.1.2 gives a 2.94% refinement from the previous record bound of Wu [141]. For reference, this gives the largest percentage improvement since the work of Bombieri, Friedlander, and Iwaniec [13]. See below for a chronology of the known upper bounds on $\pi_2(x)/\Pi(x)$. Also see Siebert [132], Riesel–Vaughan [122, Lemma 5] for numerically explicit forms of Selberg's bound [129].

| Year | Author(s) | $\pi_2(x)/\Pi(x) \lesssim$ |
|---|---|---|
| 1947 | Selberg [129] | 8 |
| 1964 | Pan [111] | 6 |
| 1966 | Bombieri–Davenport [12] | 4 |
| 1978 | Chen [24] | 3.9171 |
| 1983 | Fouvry–Iwaniec [59] | $3.7777\cdots = 34/9$ |
| 1984 | Fouvry [56] | $3.7647\cdots = 64/17$ |
| 1986 | Bombieri–Friedlander–Iwaniec [13] | 3.5 |
| 1986 | Fouvry–Grupp [58] | 3.454 |
| 1990 | Wu [140] | 3.418 |
| 2003 | Cai–Lu [17] | 3.406 |
| 2004 | Wu [141] | 3.39951 |

The main ingredients for these results come from applying sieve bounds to the set $\mathcal{A} = \{p+2 : p \le x\}$, and using equidistribution of primes in arithmetic progressions to handle remainder terms. Bombieri–Davenport obtained $\pi_2(x)/\Pi(x) \lesssim 4$ as a consequence of the Bombieri–Vinogradov theorem (1.2) and a standard sieve upper

bound of level $x^{\frac{1}{2}-\varepsilon}$. More generally, if one proves level of distribution $x^{\theta-\varepsilon}$ then one immediately obtains $\pi_2(x)/\Pi(x) \lesssim 2/\theta$. Bombieri–Friedlander–Iwaniec proved $\pi_2(x)/\Pi(x) \lesssim 7/2$ by the well-factorable variant (1.3) level of distribution $x^{\frac{4}{7}-\varepsilon}$, together with the linear sieve with well-factorable remainder (1.4).

The other key ingredient to subsequent improvements is the *switching principle*, introduced in Chen's celebrated result that there are infinitely many primes $p$ such that $p+2$ has at most two prime factors [23]. The basic insight is to use a weighted sieve inequality to split the problem into multiple cases, apply sieve bounds to $\mathcal{A} = \{p+2 : p \leq x\}$ in certain cases, and then reinterpret the remaining cases as new sieving problems for switched sets $\mathcal{B} = \{m-2 \leq x\}$ where the numbers $m$ are constructed from $\mathcal{A}$ (as prescribed depending on the case).

### 1.1.2 Outline of main ideas in Theorem 1.1.1

Maynard's new equidistribution results show equidistribution of the primes with sieve weights $\lambda(d)$ provided $d = p_1 \cdots p_r$ is restricted to suitably well-factorable integers. Unfortunately, the original linear sieve weights only partially satisfy these well-factorable conditions. In particular for $\eta > 0$, when looking at the linear sieve of level $x^{\frac{7}{12}+\eta}$, some integers $d$ in its support do not satisfy the conditions, which means that $x^{\frac{7}{12}}$ is the limit for the linear sieve given our current equidistribution technology. Nevertheless, the key observation here is that only a few exceptional $d$ fail to satisfy these conditions. Moreover up to level $x^{\frac{10}{17}}$, i.e. $\eta < 1/204$, the anatomy of exceptional $d$ may be precisely characterized in terms of $\eta$ (given specifically as $\mathcal{P}_4$, $\mathcal{P}_6$ in (1.17)). In particular, as $\eta > 0$ grows the family of exceptional integers contribute $O(\eta^5)$ to the sieve bound. However, we note this characterization breaks down when $\eta \geq 1/204$, and the contribution becomes considerably larger and more complicated.

As such we carefully revise the construction of the linear sieve, altering a few particular inclusion-exclusion steps in order to avoid the exceptional integers $d$ with bad factorizations. Once these terms no longer contribute to the sieve, this produces a worse and more complicated main term, but since there are only a very small number of such terms the resulting loss is small. And since these modified weights now satisfy stronger factorization properties in their support, we can now leverage the full strength of Maynard's equidistribution results.

## 1.2 Technical setup and results

### 1.2.1 Factorization of weights and their level of distribution

**Definition 1.2.1** (well-factorable)**.** Let $Q \in \mathbb{R}_{\geq 1}$. A sequence $\lambda(q)$ is **well-factorable of level** $Q$, if for every factorization $Q = Q_1Q_2$ into $Q_1, Q_2 \in \mathbb{R}_{\geq 1}$, there exist sequences $\gamma_1, \gamma_2$ such that

1. $|\gamma_1(q_1)|, |\gamma_2(q_2)| \leq 1$ for all $q_1, q_2 \in \mathbb{N}$,
2. $\gamma_i(q) = 0$ if $q \notin [1, Q_i]$ for $i = 1, 2$,
3. We have $\lambda = \gamma_1 * \gamma_2$, i.e.,

$$\lambda(q) = \sum_{q=q_1q_2} \gamma_1(q_1)\gamma_2(q_2).$$

In [13, Theorem 10], Bombieri–Friedlander–Iwaniec established level of distribution $x^{\frac{4}{7}-\varepsilon}$ with well-factorable weights.

**Theorem 1.2.2** (Bombieri–Friedlander–Iwaniec [13])**.** *Fix any $a \in \mathbb{Z}$ and let $A, \varepsilon > 0$. For any well-factorable sequence $\lambda$ of level $Q \leq x^{\frac{4}{7}-\varepsilon}$, we have*

$$\sum_{\substack{q \leq Q \\ (q,a)=1}} \lambda(q) \left(\pi(x; q, a) - \frac{\pi(x)}{\phi(q)}\right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

Maynard [102] considered a natural strengthening of well-factorable sequences.

**Definition 1.2.3** (triply well-factorable)**.** Let $Q \in \mathbb{R}_{\geq 1}$. A sequence $\lambda(q)$ is **triply well-factorable of level** $Q$, if for every factorization $Q = Q_1Q_2Q_3$ into $Q_1, Q_2, Q_3 \in \mathbb{R}_{\geq 1}$, there exist sequences $\gamma_1, \gamma_2, \gamma_3$ such that

1. $|\gamma_1(q_1)|, |\gamma_2(q_2)|, |\gamma_3(q_3)| \leq 1$ for all $q_1, q_2, q_3 \in \mathbb{N}$,
2. $\gamma_i(q) = 0$ if $q \notin [1, Q_i]$ for $i = 1, 2, 3$,
3. We have $\lambda = \gamma_1 * \gamma_2 * \gamma_3$, i.e.,

$$\lambda(q) = \sum_{q=q_1q_2q_3} \gamma_1(q_1)\gamma_2(q_2)\gamma_3(q_3).$$

The definitions of well-factorable and triply well-factorable sequences are quite natural and relatively simple from a conceptual standpoint. In [102, Theorem 1.1], Maynard obtains powerful equidistribution results for triply well-factorability that are beyond the scope of well-factorability. Unfortunately, triply well-factorability is too restrictive a condition for us to produce Theorem 1.1.1. As such we are forced to identify the precise mechanism that enables Maynard's equidistribution results, and extract the following technical definition that is implicit in [102].[1]

**Definition 1.2.4** (programmably factorable)**.** Let $0 < \delta < 10^{-5}$. For $x \in \mathbb{R}_{>1}$, a sequence $\lambda(q)$ is **programmably factorable of level $Q$ (relative to $x$, $\delta$)**, if for every $N \in [x^{2\delta}, x^{\frac{1}{3}+\delta/2}]$ there exists a factorization $Q = Q_1Q_2Q_3$ with $Q_1, Q_2, Q_3 \in \mathbb{R}_{\geq 1}$, satisfying the system

$$\begin{aligned} Q_1 &\leq Nx^{-\delta}, \\ N^2Q_2Q_3^2 &\leq x^{1-\delta}, \\ N^2Q_1Q_2^4Q_3^3 &\leq x^{2-\delta}, \\ NQ_1Q_2^5Q_3^2 &\leq x^{2-\delta}. \end{aligned} \tag{1.6}$$

And for every such factorization $Q = Q_1Q_2Q_3$ there exist sequences $\gamma_1, \gamma_2, \gamma_3$ such that

1. $|\gamma_1(q_1)|, |\gamma_2(q_2)|, |\gamma_3(q_3)| \leq 1$ for all $q_1, q_2, q_3 \in \mathbb{N}$,
2. $\gamma_i(q) = 0$ if $q \notin [1, Q_i]$ for $i = 1, 2, 3$,
3. We have $\lambda = \gamma_1 * \gamma_2 * \gamma_3$, i.e.,

$$\lambda(q) = \sum_{q=q_1q_2q_3} \gamma_1(q_1)\gamma_2(q_2)\gamma_3(q_3).$$

Programmable factorability is the key technical definition in this chapter. It is named in allusion to the linear programming-type system of inequalities (1.6) that the factors satisfy. The diagram below displays the various implications among the definitions.

[1]Indeed, the definition of programmably factorable in the special case $Q_3 = 1$ gives the implicit condition (which is implied by well-factorable) that enables Bombieri–Friedlander–Iwaniec to get equidistribution (1.3). Also see Lemma 5 in [58].

$\lambda$ is **triply well-factorable** of level $Q$. $\implies$ $\lambda$ is **well-factorable** of level $Q$.

$\Downarrow$

$\lambda$ is **programmably factorable** of level $Q$ (relative to $x$, $\delta$).

In the key result [102, Theorem 1.1], Maynard extended the level of distribution up to $Q < x^{\frac{3}{5}}$ for programmably factorable weights. Note that level $x^{\frac{3}{5}}$ is the natural barrier for (1.6) to admit a solution.

**Theorem 1.2.5** (Maynard [102]). *Fix any $a \in \mathbb{Z}$ and let $A, \varepsilon > 0$. For any programmably factorable sequence $\lambda$ of level $Q \leq x^{\frac{3}{5}-\varepsilon}$ (relative to $x$, $\varepsilon/50$), we have*

$$\sum_{\substack{q \leq Q \\ (q,a)=1}} \lambda(q) \left( \pi(x; q, a) - \frac{\pi(x)}{\phi(q)} \right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

*Remark* 1.2.6. [102, Theorem 1.1] was stated for triply-factorable sequences, but its proof in fact gives the result for programmably factorable sequences.

Note the weights $\widetilde{\lambda}^+$ are composed of well-factorable—but not neccessarily programmably factorable—sequences of given level $D$. Nevertheless, Maynard showed the upper bound weights $\widetilde{\lambda}^+$ of sieve level $D = x^{\frac{7}{12}-\varepsilon}$ are programmably factorable of level $D \leq Q = x^{\frac{3}{5}-\varepsilon}$ (relative to $x$, $\varepsilon/50$). By Theorem 1.2.5 this gives [102, Theorem 1.2] below.

**Corollary 1.2.6.1** (Maynard [102]). *For any fixed $a \in \mathbb{Z}$ and $A, \varepsilon > 0$, the weights $\widetilde{\lambda}^+$ from (1.9) of sieve level $D = x^{\frac{7}{12}-\varepsilon}$ satisfy*

$$\sum_{\substack{d \leq D \\ (d,a)=1}} \widetilde{\lambda}^+(d) \left( \pi(x; d, a) - \frac{\pi(x)}{\phi(d)} \right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

Later, in Proposition 1.5.4, we shall obtain technical improvements of Corollary 1.2.6.1 for Iwaniec's weights $\widetilde{\lambda}^{\pm}$ (both upper and lower), in special cases where equidistribution is restricted to moduli which are smooth, or otherwise amenable to programmable factorization.

We may summarize the definitions and results of the section up to this point as follows:

$\lambda$ is **triply well-factorable** of level $Q$.
- equidistributed for $Q < x^{\frac{3}{5}}$.
- can take $\lambda = \widetilde{\lambda}^+$ for $D \leq Q^{\frac{2}{3}}$.

$\lambda$ is **well-factorable** of level $Q$.
- equidistributed for $Q < x^{\frac{4}{7}}$.
- can take $\lambda = \widetilde{\lambda}^+$ for $D \leq Q$.

$\lambda$ is **programmably factorable** of level $Q$ (relative to $x$, $\delta$).
- equidistributed for $Q < x^{\frac{3}{5}}$.
- can take $\lambda = \widetilde{\lambda}^+$ for $D \leq Q < x^{\frac{7}{12}}$.

For each type of sequence, we have outlined their corresponding levels of distribution, and the levels at which the type is satisfied by the upper bound weights for the linear sieve. Observe well-factorability is flexible enough to accommodate the linear sieve to any level, but has weaker equidistribution. On the other hand, triple well-factorability has stronger equidistribution, but is too rigid to accommodate the linear sieve (at nontrivial levels). Finally, programmable factorability also has strong equidistribution in addition to (nontrivially) accommodating the linear sieve, though at the cost of conceptual technicality.

### 1.2.2 Sieve theory setup and bounds

We recall the standard sieve-theoretic notation. Given a finite set $\mathcal{A} \subset \mathbb{N}$, set of primes $\mathcal{P}$, and a threshold $z > 0$, we define $\mathcal{A}_d = \{n \in \mathcal{A} : d \mid n\}$ and remainder $r_{\mathcal{A}}$ via

$$|\mathcal{A}_d| = g(d)|\mathcal{A}| + r_{\mathcal{A}}(d),$$

where $g$ is a multiplicative function, with $0 \leq g(p) < 1$ for $p \in \mathcal{P}$ (we assume $g(p) = 0$ if $p \notin \mathcal{P}$). Also define $P(z) = \prod_{p<z, p\in\mathcal{P}} p$ and $V(z) = \prod_{p|P(z)} \big(1 - g(p)\big)$. The central object of interest is the *sifted sum*

$$S(\mathcal{A}, z) = S(\mathcal{A}, \mathcal{P}, z) = \sum_{n\in\mathcal{A}} \mathbf{1}_{(n,P(z))=1}. \tag{1.7}$$

Later for our application of interest, we will set $g(d) = 1/\phi(d)$. For now, it suffices for us to assume for all $2 \le w \le z$,

$$\frac{V(w)}{V(z)} = \prod_{\substack{w \le p < z \\ p \in \mathcal{P}}} \big(1 - g(p)\big) \;=\; \frac{\log z}{\log w}\Big(1 + O\Big(\frac{1}{\log w}\Big)\Big). \tag{1.8}$$

*Remark* 1.2.7. The proof of the upper bound for the standard linear sieve only requires a one-sided inequality for $V(w)/V(z)$, whereas our modification requires the above two-sided condition (1.8).

The basic result which we shall adapt is the linear sieve with well-factorable remainder, as in [62, Theorem 12.20].

**Theorem 1.2.8** (Iwaniec [62])**.** *Let $\varepsilon > 0$ and $D > 1$ be sufficiently small and large, respectively. Then for $s \ge 1$ and $z = D^{1/s}$, we have*

$$\begin{aligned} S(\mathcal{A}, z) \;&\le\; |\mathcal{A}|V(z)\big(F(s) + O(\varepsilon)\big) \;+\; \sum_{d \mid P(z)} \widetilde{\lambda}^{+}(d)\, r_{\mathcal{A}}(d), \\ S(\mathcal{A}, z) \;&\ge\; |\mathcal{A}|V(z)\big(f(s) + O(\varepsilon)\big) \;-\; \sum_{d \mid P(z)} \widetilde{\lambda}^{-}(d)\, r_{\mathcal{A}}(d), \end{aligned}$$

*where the implied constant only depends that of* (1.8)*. Here the weights $\widetilde{\lambda}^{\pm}$ are*

$$\widetilde{\lambda}^{\pm}(d) = \sum_{j \le \exp(\varepsilon^{-3})} \lambda_j^{\pm}(d) \tag{1.9}$$

*for some well-factorable sequences $\lambda_j^{\pm}$ of level $D$. The functions $F, f : \mathbb{R}^+ \to \mathbb{R}$ satisfy the system of delay-differential equations*

$$\begin{aligned} sF(s) &= 2e^{\gamma} \qquad [s \le 3] & \big(sF(s)\big)' &= f(s-1), \\ sf(s) &= 0 \qquad [s \le 2] & \big(sf(s)\big)' &= F(s-1). \end{aligned} \tag{1.10}$$

*Remark* 1.2.9. See Iwaniec [79, Theorem 1] for an alternate formulation and proof, which gives sharper quantitative bounds than $O(\varepsilon)$. However, it is more technical than necessary for our purposes.

The main result of this chapter is the following modification of the linear sieve with programmably factorable remainder.

**Theorem 1.2.10.** *Let $\mathcal{A}$ be a finite set of positive integers with density function $g(d)$ satisfying* (1.8)*, and $F(s)$ the function defined by the system* (1.10)*. Let $\varepsilon > 0$ and*

$x > 1$ *be sufficiently small and large, respectively. Then for* $\eta \geq 0$, $D = x^{\frac{7}{12}+\eta}$, $s \geq 1$, *and* $z = D^{1/s}$, *we have*

$$S(\mathcal{A}, z) \ \leq \ |\mathcal{A}|V(z)\big(F^*(s) + O(\varepsilon)\big) \ + \ \sum_{d|P(z)} \widetilde{\lambda}^*(d)\, r_{\mathcal{A}}(d),$$

*where the implied constant only depends that of* (1.8). *Here the weights* $\widetilde{\lambda}^*$ *are*

$$\widetilde{\lambda}^*(d) = \sum_{j \leq \exp(\varepsilon^{-3})} \lambda_j^*(d) \tag{1.11}$$

*for some programmably factorable sequences* $\lambda_j^*$ *of level* $D$ *(relative to* $x$, $\varepsilon/50$*). For* $\eta < \frac{1}{204}$ *we have* $F^*(s) = F(s) + O(\eta^5)$, *and* $F^*(s) \leq 1.000081\, F(s)$ *for* $1 \leq s \leq 3$, $\eta = \frac{1}{204}$.

Note Theorem 1.2.5, applied to each $\lambda = \lambda_j^*$ above, immediately implies the following.

**Corollary 1.2.10.1.** *Given any fixed* $a \in \mathbb{Z}$ *and* $A$. *For* $\eta < \frac{1}{204}$, *the weights* $\widetilde{\lambda}^*$ *as in* (1.11) *of level* $D = x^{\frac{7}{12}+\eta}$ *satisfy*

$$\sum_{\substack{d \leq D \\ (d,a)=1}} \widetilde{\lambda}^*(d) \left(\pi(x; d, a) - \frac{\pi(x)}{\phi(d)}\right) \ \ll_{a,A,\eta} \ \frac{x}{(\log x)^A}.$$

## 1.3 Programmably factorable support

The upper and lower bound weights $\lambda^{\pm}$ for the linear sieve of level $D$ are defined by

$$\lambda^{\pm}(d) = \mathbf{1}_{d \in \mathcal{D}^{\pm}}\, \mu(d),$$

where $\mathcal{D}^{\pm} = \mathcal{D}^{\pm}(D)$ are the standard support sets

$$\mathcal{D}^{+} = \{p_1 \cdots p_r \ : \ D^{1/2} \geq p_1 \geq \cdots \geq p_r, \text{ and } p_1 \cdots p_{l-1} p_l^3 \leq D \ \text{ for each odd } l \leq r\}, \tag{1.12}$$

$$\mathcal{D}^{-} = \{p_1 \cdots p_r \ : \ D^{1/2} \geq p_1 \geq \cdots \geq p_r, \text{ and } p_1 \cdots p_{l-1} p_l^3 \leq D \ \text{ for each even } l \leq r\}.$$

Observe that both sets satisfy the containment $\mathcal{D}^{\pm}(D) \subset \mathcal{D}^{\text{well}}(D)$, where

$$\mathcal{D}^{\text{well}} = \{p_1 \cdots p_r \ : \ D^{1/2} \geq p_1 \geq \cdots \geq p_r, \text{ and } p_1 \cdots p_{l-1} p_l^2 \leq D \ \text{ for each } l \leq r\}. \tag{1.13}$$

We shall return to this observation later in the section.

In [102], Maynard deduces Corollary 1.2.6.1 for $\widetilde{\lambda}^+$ from the general Theorem 1.2.5 by means of the following key result [102, Proposition 9.1] (along with a construction of Iwaniec we shall address in later sections), which programmably factorizes elements of the support $\mathcal{D}^+$.

**Proposition 1.3.1** (Maynard [102]). *Let $0 < \delta < 10^{-3}$ and let $D = x^{\frac{7}{12}-50\delta}$, $N \in [x^{2\delta}, x^{\frac{1}{3}+\delta/2}]$. Then every $d \in \mathcal{D}^+(D)$ has a factorization $d = d_1d_2d_3$ such that $d_1 \leq Nx^{-\delta}$ and*

$$\begin{aligned} N^2d_2d_3^2 &\leq x^{1-\delta}, \\ N^2d_1d_2^4d_3^3 &\leq x^{2-\delta}, \\ Nd_1d_2^5d_3^2 &\leq x^{2-\delta}. \end{aligned} \tag{1.14}$$

*Remark* 1.3.2. The level $x^{\frac{7}{12}}$ is sharp in this construction. Indeed, roughly speaking the linear sieve weights are not programmably factorable of level $D = x^{\frac{7}{12}+\eta}$ for any $\eta > 0$, because the support set contains obstructing (families of) elements $d \in \mathcal{D}^+(D)$ of the form $d = p_1 \cdots p_r$ where $p_1 \approx \cdots \approx p_6 \approx D^{\frac{1}{7}}$, or where $p_1 \approx p_2 \approx D^{\frac{2}{7}}$ and $p_3 \approx p_4 \approx D^{\frac{1}{7}}$. These obstructions suggest how we should restrict the support set in order to increase the level.

For $\eta > 0$, level $D = x^{\frac{7}{12}+\eta}$, we define the modified weights $\lambda^* = \lambda^*_\eta$,

$$\lambda^*(d) = \mathbf{1}_{d\in\mathcal{D}^*}\,\mu(d), \tag{1.15}$$

for the support set $\mathcal{D}^*$,

$$\mathcal{D}^* \;=\; \mathcal{D}^+(x^{\frac{7}{12}}) \cup \{p_1 \cdots p_r \in \mathcal{D}^+(x^{\frac{7}{12}+\eta}) : p_1 \cdots p_i \notin \mathcal{P}_i, i \leq r, i \in \{4, 6\}\}. \tag{1.16}$$

Here $\mathcal{P}_4$ and $\mathcal{P}_6 = \mathcal{P}_{6,1} \cup \mathcal{P}_{6,2}$ are exceptional subsets of $\mathcal{D}^+(x^{\frac{7}{12}+\eta})$, given by

$$\begin{aligned} \mathcal{P}_4 &= \{p_1 \cdots p_4 \;:\; p_1 < x^{\frac{1}{6}+2\eta} \text{ and } p_2p_4 > x^{\frac{1}{4}-3\eta}\}, \\ \mathcal{P}_{6,1} &= \{p_1 \cdots p_6 \;:\; p_1p_2 < x^{\frac{1}{6}+2\eta} \text{ and } p_2p_3p_4 > x^{\frac{1}{4}-3\eta} \text{ and } p_6 > x^{\frac{1}{12}-5\eta}\}, \\ \mathcal{P}_{6,2} &= \{p_1 \cdots p_6 \;:\; p_1,\, p_2p_3 < x^{\frac{1}{6}+2\eta} \text{ and } p_1p_4,\, p_2p_3p_4 > x^{\frac{1}{4}-3\eta} \text{ and } p_6 > x^{\frac{1}{12}-5\eta}\}. \end{aligned} \tag{1.17}$$

The modified support set $\mathcal{D}^* = \mathcal{D}^*_\eta$ is understood to depend on $\eta > 0$ (as do $\mathcal{P}_4$, $\mathcal{P}_6$), but we will suppress this for notational convenience.

In this section, we establish a programmable factorization of the elements of the support $\mathcal{D}^*$ provided $D < x^{\frac{10}{17}}$, i.e. $\eta < \frac{1}{204}$. This will serve as the key technical input for the proof of Theorem 1.2.10.

**Proposition 1.3.3** (Factorization of elements of $\mathcal{D}^*$). *Let $0 < \delta < 10^{-5}$, and take $0 < \eta < \frac{1}{204} - 3\delta$ and $N \in [x^\delta, x^{\frac{1}{3}-\delta/2}]$. If $d \in \mathcal{D}^*$ for $D = x^{\frac{7}{12}+\eta-50\delta}$, then we may factor $d = d_1d_2d_3$ such that $d_1 \leq Nx^{-\delta}$ and*

$$\begin{aligned} N^2\; d_2d_3^2 &\leq x^{1-\delta}, \\ N^2d_1d_2^4d_3^3 &\leq x^{2-\delta}, \\ Nd_1d_2^5d_3^2 &\leq x^{2-\delta}. \end{aligned} \tag{1.18}$$

On the first attempt working through technicalities, we encourage the reader to set $\delta = 0$ in order to better view the key features.

Before proving the proposition, we need some lemmas. The first gives a general-purpose criterion to factor an integer $d$.

**Lemma 1.3.4.** *Let $D = x^{\frac{7}{12}+\eta-50\delta}$ for $-\frac{1}{84} < \eta < \frac{1}{60}$. A factorization $d = d_1d_2d_3$ satisfies* (1.18)*, provided $d_1, d_2, d_3 \geq 1$ satisfy*

$$\begin{aligned} d_2 &\in [x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}], \\ d_1 \leq Nx^{-\delta}, \quad &\text{and} \quad d_3 \leq D/Nd_2. \end{aligned} \tag{1.19}$$

*Proof.* By (1.19), $Nd_3 \leq D/d_2$ and so

$$\begin{aligned} N^2\, d_2d_3^2 &\leq D^2/d_2 \leq x^{2(\frac{7}{12}+\eta-50\delta)-(\frac{1}{6}+2\eta)} = x^{1-\delta}, \\ N^2d_1d_2^4d_3^3 &\leq D^3d_2 \leq x^{3(\frac{7}{12}+\eta-50\delta)+(\frac{1}{4}-3\eta)} = x^{2-\delta}, \\ Nd_1d_2^5d_3^2 &\leq D^2d_2^3 \leq x^{2(\frac{7}{12}+\eta-50\delta)+3(\frac{1}{4}-3\eta)} = x^{\frac{23}{12}-7\eta-100\delta} < x^{2-\delta}, \end{aligned}$$

using $\eta \in (-\frac{1}{84}, \frac{1}{60})$. This gives (1.18). □

The above criterion implies factorizations in the following special cases.

**Lemma 1.3.5.** *Let $D = x^{\frac{7}{12}+\eta-50\delta}$ for $\eta < \frac{1}{60}$. For $r \geq 4$, let $x^{\frac{1}{6}+2\eta} > p_1 \geq \cdots \geq p_r$ be primes for which $d = p_1 \cdots p_r \in \mathcal{D}^+(D)$. Suppose $d_2$ is one of the subproducts in $\{p_1p_4, p_2p_3, p_2p_4, p_2p_3p_4\}$. Then $d$ has a factorization $d = d_1d_2d_3$ satisfying* (1.18)*, provided*

$$d_2 \in [x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}].$$

*Proof.* Let $C = D/Nd_2$ and note either $p_1 \leq N$ or $p_1 \leq C$, since

$$p_1^2 \leq D^{\frac{2}{3}} = x^{\frac{2}{3}(\frac{7}{12}+\eta-50\delta)} < x^{\frac{1}{3}+4\eta-50\delta} \leq D/d_2 = NC.$$

We proceed by induction on $r \geq 4$. As the base case $r = 4$, by Lemma 1.3.4 it suffices for each $b$ to factor $p_1 \cdots p_4/d_2 = d_1d_3$ for $d_1 \leq N, d_3 \leq C$. Indeed, this holds when $d_2 = p_2p_3p_4$ since $p_1^2 \leq AC$, and similarly:

- If $d_2 = p_2p_4$ then $p_3^2 \leq D/p_1p_2p_3 \leq NC/p_1$ implies $p_1p_3 = d_1d_3$ for some $d_1 \leq N$, $d_3 \leq C$.
- If $d_2 = p_1p_4$ then $p_3^2 \leq D/p_1p_2p_3 \leq NC/p_2$ implies $p_2p_3 = d_1d_3$ for some $d_1 \leq N$, $d_3 \leq C$.

- If $d_2 = p_2p_3$ then $p_4^2 \le D/p_1p_2p_3 \le NC/p_1$ implies $p_1p_4 = d_1d_3$ for some $d_1 \le N$, $d_3 \le C$.

Now for $r \ge 5$, we inductively assume a factorization $p_1 \cdots p_{r-1} = d_1d_2d_3$ with $d_1 \le N$, $d_3 \le C$. Then $p_r^2 \le D/p_1 \cdots p_{r-1} = NC/(ac)$ so either $d_1p_r \le N$ or $d_3p_r \le C$, extending the factorization. Hence Lemma 1.3.4 applies again, and completes the proof. □

Finally, if the primes dividing $d$ are small enough, we may use the greedy algorithm to factor $d$ as follows.

**Lemma 1.3.6.** *Let $D = x^{\frac{7}{12}+\eta-50\delta}$ for $\eta < \frac{1}{60}$. For $r \ge 4$, let $x^{\frac{1}{6}+2\eta} > p_1 \ge \cdots \ge p_r$ be primes for which $d = p_1 \cdots p_r \in \mathcal{D}^+(D)$, and $p_6 < x^{\frac{1}{12}-5\eta}$ if $r \ge 6$. Then $d$ has a factorization $d = abc$ satisfying* (1.18)*, provided there is a factorization $p_1p_2p_3p_4 = d_1d_2d_3$ satisfying*

$$d_1 \le N, \quad d_3 \le x^{1-2\delta}/DN, \quad d_2 \le D^2/x = x^{\frac{1}{6}+2\eta}. \tag{1.20}$$

*Proof.* Let $D_1 = N$, $D_2 = D^2/x$, $D_3 = x^{1-2\delta}/(DN)$, so that $d_i \le D_i$ by assumption.

We now greedily append primes to $d_i$ while preserving $d_i \le D_i$ for all $i$, i.e. where at the $j$th step we replace $d_i \mapsto d_ip_j$ (for one of $i = 1, 2, 3$) provided $d_ip_j \le D_i$. Starting from $j = 5$, we stop either when we have exhausted all primes (i.e. $j = r$), or $d_ip_j > D_i$ for each $i = 1, 2, 3$. In the former case, we have the desired $d_1d_2d_3 = d = p_1 \cdots p_r$ and $d_i \le D_i$ so we easily get

$$\begin{aligned} D_1 &= A, \\ A^2D_2D_3^2 &= x^{1-4\delta}, \\ A^2D_1D_2^4D_3^3 &= D^5x^{-1-3\delta} \le x^{5\cdot\frac{3}{5}-1-256\delta} < x^{2-3\delta}, \\ AD_1D_2^5D_3^2 &= D^8x^{-3-2\delta} \le x^{8\cdot\frac{3}{5}-3-404\delta} < x^{2-2\delta}. \end{aligned}$$

Thus $d_1d_2d_3 = d = p_1 \cdots p_r$ gives the desired factorisation.

In the latter case, there exists a terminal index $j < r$ for which $d_ip_j > D_i$ for all $i = 1, 2, 3$. Note if $j$ is odd, then $d_ip_j \le D_i$ for some $i$, since

$$p_j^3 \le \frac{D}{p_1 \cdots p_{j-1}} = \frac{D_1D_2D_3}{d_1d_2d_3}.$$

So the terminal $j$ is even with $j \ge 6$. By assumption $p_j \le p_6 \le x^{\frac{1}{12}-5\eta}$ is smaller than the width of the interval $[x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}]$. And since $d_2 \le D_2 = x^{\frac{1}{6}+2\eta} < d_2p_j$, we deduce $e_2 := d_2p_j$ lies in the interval $e_2 \in [x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}]$.

Thus letting $E_3 := D_2D_3/e_2$, for each $l > j$ in turn we shall greedily append the prime $p_l$ onto either $d_1$ or $d_3$ while preserving $d_1 < D_1$ and $d_3 < D_3$. Indeed, for all $l > j$,

$$p_l^2 \le \frac{D}{p_1 \cdots p_{l-1}} = \frac{D_1D_2D_3}{d_1d_2d_3p_j \cdots p_{l-1}} \le \frac{D_1E_3}{d_1d_3p_{j+1} \cdots p_{l-1}},$$

so there is a factorization $e_1e_3 = d_1d_3p_{j+1} \cdots p_l$ with $e_1 \le D_1 = N$ and $e_3 \le E_3 = D/(Ae_2x^{2\delta})$. Hence the result now follows by Lemma 1.3.4 for the factorization $e_1e_2e_3 = d_1d_2d_3p_j \cdots p_l = p_1 \cdots p_l$. □

*Proof of Proposition 1.3.3.* We shall consider 3 cases, depending on the sizes of $p_1$ and $p_2p_3$ compared to the endpoints of the key interval $[x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}]$.

**CASE 1:** $p_1 \ge D^2/x = x^{\frac{1}{6}+2\eta}$**.**

Let $d_2 := p_1$, $C := D/Nd_2$. Note $C = D/Nd_2 \ge D^{\frac{2}{3}}/N \ge 1$.

Next $D \ge p_1^3 \ge p_1p_2^2$ implies $p_2^2 \le D/p_1 = NC$, so either $p_2 \le N$ or $p_2 \le C$. Similarly, since $p_1 \cdots p_{j-1}p_j^2 \le D$ for all $j \le r$, we get $p_j^2 \le \frac{AC}{p_2 \cdots p_{j-1}}$ for $3 \le j \le r$. As such, we may factor $p_2 \cdots p_r = d_1d_3$ for $d_1 \le N, d_3 \le C$. Hence by Lemma 1.3.4 $p_1 \cdots p_r = d_1d_2d_3$ satisfies (1.18).

In the remaining cases, we assume $p_1 < x^{\frac{1}{6}+2\eta}$. By Lemma 1.3.5, it remains to consider $p_2p_3 > x^{\frac{1}{4}-3\eta}$ or $p_2p_3 < x^{\frac{1}{6}+2\eta}$. Note

$$p_1p_3 \le p_1^{\frac{1}{3}}(p_1p_2p_3^3)^{\frac{1}{3}} \le (x^{\frac{1}{6}+2\eta})^{\frac{1}{3}}D^{\frac{1}{3}} = x^{\frac{1}{3}(\frac{1}{6}+\frac{7}{12}+3\eta-50\eta)} < x^{\frac{1}{4}+\eta-16\delta}. \tag{1.21}$$

**CASE 2:** $p_2p_3 > x^{\frac{1}{4}-3\eta}$ **and** $p_1 < x^{\frac{1}{6}+2\eta}$**.**

The proof follows by Lemma 1.3.5 if $p_2p_4 \in [x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}]$. Thus by definition of $\mathcal{P}_4$, in this case we may assume

$$p_2p_4 < x^{\frac{1}{6}+2\eta}. \tag{1.22}$$

Hence we have $p_4 < x^{12\eta+50\delta}$, since

$$p_2 > p_2(p_1p_2p_3^3)/D > (p_2p_3)^3/D > x^{3(\frac{1}{4}-3\eta)}/D = x^{\frac{1}{6}-10\eta+50\delta}. \tag{1.23}$$

If $p_1p_4 > x^{\frac{1}{6}+2\eta}$, then the proof follows by Lemma 1.3.5 where $d_2 = p_1p_4$ is $< x^{(\frac{1}{6}+2\eta)+12\eta+50\delta} < x^{\frac{1}{4}-3\eta}$, since $\eta < \frac{1}{204} - 3\delta$.

Else $p_1p_4 < x^{\frac{1}{6}+2\eta}$. We shall apply Lemma 1.3.6 with $d_2 = p_1p_4$.

If either $N$ or $x^{1-2\delta}/DN$ is greater than $x^{\frac{1}{4}+\eta-16\delta} \ge p_2p_3$, by (1.21), then Lemma 1.3.6 completes the proof with $(d_1, d_3) = (p_2p_3, 1)$ or $(1, p_2p_3)$, respectively. Otherwise,

$N, x^{1-2\delta}/DN \in [x^{\frac{1}{6}-2\eta-64\delta}, x^{\frac{1}{4}+\eta-16\delta}]$, since $x/D = x^{\frac{5}{12}-\eta+50\delta}$. But then, using $\eta < \frac{1}{108}$,

$$\begin{aligned} \max(N, x^{1-2\delta}/DN) &\geq (x^{1-2\delta}/D)^{\frac{1}{2}} = x^{\frac{5/2}{12}-\eta/2+24\delta} > x^{\frac{1}{6}+2\eta} > p_2, \\ \min(N, x^{1-2\delta}/DN) &\geq x^{\frac{1}{6}-2\eta-64\delta} > x^{\frac{1}{8}+\eta/2-8\delta} \geq (p_1p_3)^{1/2} \geq p_3, \end{aligned} \tag{1.24}$$

by (1.21), which suffices again for Lemma 1.3.6. Note $p_6 < x^{12\eta+50\delta} < x^{\frac{1}{12}-5\eta}$ when $r \geq 6$, using $\eta < \frac{1}{204} - 3\delta$.

**CASE 3: $p_2p_3 < x^{\frac{1}{6}+2\eta}$ and $p_1 < x^{\frac{1}{6}+2\eta}$.**

By Lemma 1.3.5, it suffices to consider either $p_1p_4 < x^{\frac{1}{6}+2\eta}$ or $p_1p_4 > x^{\frac{1}{4}-3\eta}$.

**Subcase 3.1: $p_1p_4 < x^{\frac{1}{6}+2\eta}$.**

Suppose we can show $p_6 < x^{\frac{1}{12}-5\eta}$ (when $r \geq 6$). Then since $x^{1-2\delta}/D > D^4/x^2$, either $N$ or $x^{1-2\delta}/DN$ is greater than $D^2/x$. Thus Lemma 1.3.6 will complete the proof, with $(d_1, d_2, d_3) = (p_1p_4, p_2p_3, 1)$ or $(1, p_2p_3, p_1p_4)$.

If $p_1p_3 > x^{\frac{1}{4}-3\eta}$ then in this subcase

$$x^{\frac{1}{12}+\eta} > (p_2p_3)^{\frac{1}{2}} \geq p_3 = p_4\frac{p_1p_3}{p_1p_4} > p_4x^{\frac{1}{12}-5\eta},$$

so $p_4 < x^{6\eta}$. Hence $p_6 \leq p_4 < x^{\frac{1}{12}-5\eta}$ since $\eta < \frac{1}{108}$, which completes the proof.

Else $p_1p_3 < x^{\frac{1}{4}-3\eta}$. By Lemma 1.3.5, it suffices $p_1p_3 < x^{\frac{1}{6}+2\eta}$. Then we see $p_3 > x^{\frac{1}{12}-5\eta}$ implies $p_1 < x^{\frac{1}{12}+7\eta}$. If further $p_1p_2 > x^{\frac{1}{4}-3\eta}$, then similarly

$$x^{\frac{1}{12}+7\eta} > p_1 \geq p_2 = p_4\frac{p_1p_2}{p_1p_4} > p_4x^{\frac{1}{12}-5\eta},$$

so $p_3 < x^{12\eta}$. Hence $p_6 \leq p_3 < x^{\frac{1}{12}-5\eta}$ since $\eta < \frac{1}{204}$, which completes the proof.

Else $p_1p_2 < x^{\frac{1}{4}-3\eta}$. By Lemma 1.3.5, we may assume $p_1p_2 < x^{\frac{1}{6}+2\eta}$.

Similarly, suppose $p_2p_3p_4 < x^{\frac{1}{4}-3\eta}$. By Lemma 1.3.5 we may assume $p_2p_3p_4 < x^{\frac{1}{6}+2\eta}$, and so

$$p_6 \leq (p_2p_3p_4)^{\frac{1}{3}} \leq x^{\frac{(2/3)}{12}+(2/3)\eta} \leq x^{\frac{1}{12}-5\eta},$$

using $\eta < \frac{1}{204}$, which completes the proof.

Thus we may assume $p_2p_3p_4 > x^{\frac{1}{4}-3\eta}$. But unless $p_6 < x^{\frac{1}{12}-5\eta}$, this subcase will contradict the definition of $\mathcal{P}_{6,1}$ in (1.17), hence completing the proof.

**Subcase 3.2: $p_1p_4 > x^{\frac{1}{4}-3\eta}$.**

If $d_2 = p_2p_3p_4 < x^{\frac{1}{6}+2\eta}$, then Lemma 1.3.6 completes the proof with $(d_1, d_3) = (p_1, 1)$ or $(1, p_1)$, since

$$p_6 \le (p_2p_3p_4)^{\frac{1}{3}} \le x^{\frac{1}{3}(\frac{1}{6}+2\eta)} \le x^{\frac{1}{12}-5\eta}$$

for $\eta < \frac{1}{204}$. And if $p_2p_3p_4 \in [x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}]$ the proof follows by Lemma 1.3.5.

Else $p_2p_3p_4 > x^{\frac{1}{4}-3\eta}$. Note $p_4 < x^{\frac{1}{12}+\eta}$ and $p_1 = p_1p_4/p_4 > x^{\frac{1}{6}-4\eta}$ and $p_2p_3p_4 < x^{\frac{1}{4}+3\eta}$. Also note we may factor $p_1p_4 = d_1d_3$ for $d_1 \le N$, $d_3 \le x^{1-2\delta}/DN$ (Indeed this follows if $N$ or $x^{1-2\delta}/DN$ exceeds $x^{\frac{1}{4}+3\eta} \ge p_1p_4$. Else $N, x^{1-2\delta}/DN \in [x^{\frac{1}{6}-4\eta-2\delta}, x^{\frac{1}{4}+3\eta}]$, which also works similarly as with (1.24), since $p_4 < x^{\frac{1}{12}+\eta} < x^{\frac{1}{6}-4\eta-2\delta}$ and $p_1 < x^{\frac{1}{6}+2\eta} < x^{\frac{5}{12}-\eta-\delta}$ by $\eta < \frac{1}{60}$)

If further $p_6 > x^{\frac{1}{12}-5\eta}$, then this subcase contradicts the definition of $\mathcal{P}_{6,2}$ in (1.17). Hence we have $p_6 \le x^{\frac{1}{12}-5\eta}$, and so by the above paragraph Lemma 1.3.6 completes the proof with $d_2 = p_2p_3$.

Combining all cases completes the proof of Proposition 1.3.3. □

### 1.3.1 Refined factorization of $\mathcal{D}^{\text{well}}$

Proposition 1.3.3 (programmably) factorizes each $d \in \mathcal{D}^* \subset \mathcal{D}^+$, and forms the key step to prove the weights $\lambda^*$ are programmably factorable. With applications in mind to twin primes, we shall similarly (programmably) factorize certain subsets of the well-factorable support $\mathcal{D}^{\text{well}}$, as in (1.13).

In the following result, we factorize $d \in \mathcal{D}^{\text{well}}(D)$ for variable level $D \in (x^{\frac{4}{7}}, x^{\frac{3}{5}})$, depending on the anatomy of $d$. As $\mathcal{D}^{\pm} \subset \mathcal{D}^{\text{well}}$, this has implications to both upper and lower bounds for the standard linear sieve.

**Proposition 1.3.7.** *Let $\mathcal{D}^{\text{well}}(D)$ as in (1.13) for $D = x^{\frac{7}{12}+\eta-50\delta}$ and $-\frac{1}{84} < \eta < \frac{1}{60} - 30\delta$. Let $x^{\frac{1}{4}-3\eta} \ge p_1 \ge \cdots \ge p_r$ be primes for which $d = p_1 \cdots p_r \in \mathcal{D}^{\text{well}}(D)$. Then $d$ has factorization $d = abc$ satisfying (1.18) if $p_3 \le x^{\frac{1}{12}-5\eta}$, or if*

$$d_2 \in [x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}] \qquad \text{with } \ d_2 \mid p_1p_2p_3, d_2 \neq p_3.$$

*Proof.* For $i = 1, 2, 3$, suppose $d_2 = p_1 \cdots p_i$ lies $[x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}]$, and let $C = D/Nd_2$. Since $p_1 \cdots p_{j-1}p_j^2 \le D$ for all $i < j \le r$, we get $p_j^2 \le \frac{NC}{p_{i+1} \cdots p_{j-1}}$ for $i < j \le r$. As such, we may factor $p_{i+1} \cdots p_r = d_1d_3$ for $d_1 \le N, d_3 \le C$. Hence by Lemma 1.3.4 $p_1 \cdots p_r = d_1d_2d_3$ satisfies (1.18).

Else, by assumption $p_1 < x^{\frac{1}{4}-3\eta}$ so we may assume further $p_1 < x^{\frac{1}{6}+2\eta}$. In particular this gives $p_1^2 \le D/d_2$. For the remaining $d_2 \mid p_1p_2p_3$:

- If $d_2 = p_2p_3$ then $p_1^2 \le D/d_2$ implies $p_1 \le N$ or $p_1 \le D/Nd_2$.
- If $d_2 = p_1p_3$ then $p_2^2 \le D/d_2$ implies $p_2 \le N$ or $p_2 \le D/Nd_2$.
- If $d_2 = p_2$ then $p_3^2 \le D/p_1d_2$ implies a factorization $p_1p_3 = d_1d_3$ for $d_1 \le N, d_3 \le D/Nd_2$.

For each $d_2$ above, we factored $p_1p_2p_3 = d_1d_2d_3$ for $d_1 \le N, c \le D/Nd_2$. Since $p_1 \cdots p_{j-1}p_j^2 \le D$ for all $j \le r$, by induction we may factor $p_1 \cdots p_r = d_1d_2d_3$ for $d_1 \le N, d_3 \le D/Nd_2$. By Lemma 1.3.4 $p_1 \cdots p_r = d_1d_2d_3$ satisfies (1.18).

Finally, suppose $p_3 \le x^{\frac{1}{12}-5\eta}$ is less than the width of the interval $[x^{\frac{1}{6}+2\eta}, x^{\frac{1}{4}-3\eta}]$. Since $p_1 < x^{\frac{1}{6}+2\eta}$, we have $p_1p_3 < x^{\frac{1}{4}-3\eta}$ so by the above argument we may assume $d_2 := p_1p_3 < x^{\frac{1}{6}+2\eta}$. Then $p_2^3 \le p_1p_2^2 \le D$ implies

$$p_2^2 \le x^{\frac{2}{3}(\frac{7}{12}+\eta)} < x^{\frac{5}{12}-\eta+48\delta} = \frac{x^{1-2\delta}}{D},$$

since $\eta < \frac{1}{60} - 30\delta$. Thus $p_2 \le A$ or $p_2 \le x/DA$, so there is a factorization $p_1p_2p_3 = d_1d_2d_3$ satisfying (1.20). Hence the same greedy argument as in Lemma 1.3.6 completes the proof, with $p_3$ playing the role of $p_6$. □

Taking the maximum valid $\eta$ as above, we may re-express the above factorization of level $x^\theta$, $\theta = \frac{7}{12} + \eta$, as follows. Note the maximum $\theta$ for which $t \in [\frac{1}{6} + 2\eta, \frac{1}{4} - 3\eta]$ is given by

$$\theta(t) = \begin{cases} \frac{2-t}{3} & \text{if } t > \frac{1}{5}, \\ \frac{1+t}{2} & \text{if } t \le \frac{1}{5}. \end{cases} \tag{1.25}$$

Similarly the maximum $\theta = \frac{7}{12} + \eta$ for which $t \le \frac{1}{12} - 5\eta$ is $(3-t)/5$.

**Corollary 1.3.7.1.** *Let $p_1 \ge \cdots \ge p_r$ be primes and write $p_i = x^{t_i}$. If $d = p_1 \cdots p_r \in \mathcal{D}^{\text{well}}(x^{\theta-50\delta})$, then there is a factorization $d = d_1d_2d_3$ satisfying* (1.18) *provided*

$$\theta \le \theta(t_1),$$

*for $\theta(t)$ as in* (1.25)*. Moreover if $t_1 \le \frac{1}{5}$, then it suffices that*

$$\theta \le \theta(t_1, t_2, t_3) := \max\Big\{ \frac{3-t_3}{5}, \theta(t_1), \theta(t_2), \theta(t_1+t_2+t_3), \theta(t_1+t_2), \theta(t_1+t_3), \theta(t_2+t_3) \Big\}. \tag{1.26}$$

## 1.4 Modification of the linear sieve

In this section we shall bound the modified linear sieve, analogous to the bounds for the linear sieve (sometimes called the Jurkat–Richert theorem). This bound will form the basis for our final result the next section, in which we modify the construction of Iwaniec's weights.

**Proposition 1.4.1.** *Let $\varepsilon > 0$ be sufficiently small. For $\eta \leq \frac{1}{204}$, the modified weights $\lambda^*$ as in (1.15) of level $D = x^{\frac{7}{12}+\eta-\varepsilon}$ satisfy*

$$S(\mathcal{A}, z) \ \leq \ |\mathcal{A}|V(z)\Big(F^*(\tfrac{\log D}{\log z}) + o(1)\Big) + \sum_{d|P(z)} \lambda^*(d)\, r_{\mathcal{A}}(d),$$

*where $F^* = F^*_\eta$ is a function satisfying $F^*(s) = F(s) + O(\eta^5)$ for $F$ as in (1.10).*

*Remark* 1.4.2. It suffices for our purposes to obtain qualitative error $o(1)$ in the factor accompanying $F^*$. Though as with the Jurkat–Richert theorem, with greater care one should obtain a quantitative refinement, e.g. $O((\log D)^{-1/6})$. See (12.4)–(12.8) in [62].

We now adapt the proof. Let $D = x^{\frac{7}{12}+\eta}$ and $D_0 = x^{\frac{7}{12}}$. For $n \geq 1$, primes $p_1 \geq \cdots \geq p_n$, if $p_1 \cdots p_n \notin \mathcal{D}^+(D)$ then there exists a minimal index $l \leq n$ such that $p_1 \cdots p_l \notin \mathcal{D}^+(D)$. By definition such minimal $l$ is odd (explicitly, this occurs when $p_1 \cdots p_{l-1} p_l^3 > D$ but $p_1 \cdots p_{m-1} p_m^3 \leq D$ for all odd $m < l$.) Similarly, if $p_1 \cdots p_n \notin \mathcal{D}^*$ there exists a minimal index $l \leq n$ such that $p_1 \cdots p_l \notin \mathcal{D}^*$, which is also odd.

Indeed, to show this let $l \leq n$ be the minimal index such that $p_1 \cdots p_l \notin \mathcal{D}^*$. If $(p_1, \ldots, p_j) \notin \mathcal{P}_j$ for all $j \leq l$, $j \in \{4, 6\}$, then clearly $l > j$ must be odd, as with $\mathcal{D}^+(D)$. On the other hand, if $(p_1, \ldots, p_j) \in \mathcal{P}_j$ for some $j \leq l$, $j \in \{4, 6\}$, a priori one might expect $l$ could be even. However, the key point in this case is that $p_1 \cdots p_j \in \mathcal{D}^+(D_0) \subset \mathcal{D}^*$ (since $p_1 \cdots p_j \approx D^{\frac{6}{7}}$ by definition of $\mathcal{P}_j$). Thus $l > j$ is the minimal index such that $p_1 \cdots p_l \notin \mathcal{D}^+(D_0)$, and hence must be odd as claimed.

Using this minimal index, we show the following lemma.

**Lemma 1.4.3.** *Let $h$ be a multiplicative function with $0 \leq h(p) \leq 1$ for all primes $p$. Then we have*

$$\prod_{p|n}(1 - h(p)) \ \leq \ \sum_{d|n} \lambda^*(d) h(d).$$

*Proof.* Note if $h = 1$ identically, we interpret the product as $\mathbf{1}_{n=1}$. Now by definition,

$$\sum_{d|n} \lambda^*(d)h(d) - \prod_{p|n}(1 - h(p)) \ = \ \sum_{\substack{d|n \\ d \in \mathcal{D}^*}} \mu(d)h(d) - \sum_{d|n} \mu(d)h(d) \ = \ \sum_{\substack{d|n \\ d \notin \mathcal{D}^*}} \mu(d)h(d).$$

Then splitting up $d \notin \mathcal{D}^*$ by its minimal index,

$$\sum_{\substack{d|n \\ d\notin\mathcal{D}^*}} \mu(d)h(d) = -\sum_{\text{odd } l} \sum_{\substack{p_l<\cdots<p_1<z \\ p_1\cdots p_{l-1}\in\mathcal{D}^* \\ p_1\cdots p_l\notin\mathcal{D}^*}} h(p_1\cdots p_l) \sum_{\substack{p_1\cdots p_l b|n \\ b|P(p_l)}} \mu(b)h(b) \;\le\; 0,$$

since $h \ge 0$ and the inner sum on $b$ factors as $\prod_{p|(P(p_l),n)}(1-h(p)) \ge 0$, since $h(p) \le 1$. □

By Lemma 1.4.3 with $h(d) = 1$, we have

$$\mathbf{1}_{n=1} \;\le\; \sum_{d|n} \lambda^*(d), \tag{1.27}$$

in which case we obtain

$$\begin{aligned} S(\mathcal{A},\mathcal{P},z) = \sum_{n\in A} \mathbf{1}_{(n,P(z))=1} \;&\le\; \sum_{n\in A} \sum_{d|(n,P(z))} \lambda^*(d) = \sum_{d|P(z)} \lambda^*(d)|A_d| \\ &= X \sum_{d|P(z)} \lambda^*(d)\, g(d) + \sum_{d|P(z)} \lambda^*(d)\, r_{\mathcal{A}}(d) \;=:\; XV^*(D,z) + R^*_{\mathcal{A}}(D,z). \end{aligned} \tag{1.28}$$

Following Lemma 1.4.3 with $h = g$, we have the identity

$$V^*(D,z) := \sum_{\substack{d|n \\ d\notin\mathcal{D}^*}} \mu(d)g(d) = V(z) + \sum_{\text{odd } n} \sum_{\substack{p_n<\cdots<p_1<z \\ p_1\cdots p_{n-1}\in\mathcal{D}^* \\ p_1\cdots p_n\notin\mathcal{D}^*}} g(p_1\cdots p_n)V(p_n), \tag{1.29}$$

and similarly

$$V^+(D,z) = V(z) + \sum_{\text{odd } n} \sum_{\substack{p_n<\cdots<p_1<z \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D) \\ p_1\cdots p_n\notin\mathcal{D}^+(D)}} g(p_1\cdots p_n)V(p_n) \;=:\; V(z) + \sum_{\text{odd } n} V_n(D,z). \tag{1.30}$$

Then the difference of $V^*$ and $V^+$ is

$$V^*(D,z) - V^+(D,z) \;=\; \sum_{\text{odd } n} \sum_{p_n<\cdots<p_1<z} g(p_1\cdots p_n)V(p_n)\,\boldsymbol{\Delta}, \tag{1.31}$$

where $\boldsymbol{\Delta}$ is the difference of indicator functions,

$$\boldsymbol{\Delta} := \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^* \\ p_1\cdots p_{n-1}\in\mathcal{D}^*}} - \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D)}} = \mathbf{1}_{\substack{p_1\cdots p_n\in\mathcal{D}^+(D)\setminus\mathcal{D}^* \\ p_1\cdots p_{n-1}\in\mathcal{D}^*}} - \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D)\setminus\mathcal{D}^*}},$$

recalling $\mathcal{D}^* \subset \mathcal{D}^+(D)$. Note if a point is $(p_1, .., p_6) \in \mathcal{P}_6$ then its projection is $(p_1, .., p_4) \notin \mathcal{P}_4$. So by definitions of $\mathcal{D}^*$, $\mathcal{D}^+(D)$ from (1.16), (1.12), for odd $n$ we have the identities,

$$\mathbf{1}_{\substack{p_1\cdots p_n\in\mathcal{D}^+(D)\setminus\mathcal{D}^* \\ p_1\cdots p_{n-1}\in\mathcal{D}^*}} = \sum_{\substack{j\in\{4,6\}\\ j<n}} \mathbf{1}_{(p_1,..,p_j)\in\mathcal{P}_j} \cdot \mathbf{1}_{\substack{p_1\cdots p_n\in\mathcal{D}^+(D)\setminus\mathcal{D}^+(D_0), \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D_0)}} \tag{1.32}$$

$$\mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D)\setminus\mathcal{D}^*}} = \sum_{\substack{j\in\{4,6\}\\ j<n}} \mathbf{1}_{(p_1,..,p_j)\in\mathcal{P}_j} \cdot \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D)\setminus\mathcal{D}^+(D_0)}}. \tag{1.33}$$

We may plug (1.32) and (1.33) into $\mathbf{\Delta}$. In addition, we strategically add and subtract the indicator function of $\{p_1\cdots p_n \notin \mathcal{D}^+(D),\ p_1\cdots p_{n-1} \in \mathcal{D}^+(D_0)\}$, which together give

$$\begin{aligned}
\mathbf{\Delta} &= \sum_{\substack{j\in\{4,6\}\\ j<n}} \mathbf{1}_{(p_1,..,p_j)\in\mathcal{P}_j} \cdot \left( \mathbf{1}_{\substack{p_1\cdots p_n\in\mathcal{D}^+(D)\setminus\mathcal{D}^+(D_0) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D_0)}} - \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D)\setminus\mathcal{D}^+(D_0)}} \right) \\
&= \sum_{\substack{j\in\{4,6\}\\ j<n}} \mathbf{1}_{(p_1,..,p_j)\in\mathcal{P}_j} \cdot \left( \mathbf{1}_{\substack{p_1\cdots p_n\in\mathcal{D}^+(D)\setminus\mathcal{D}^+(D_0) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D_0)}} + \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D_0)}} \right. \\
&\qquad\qquad \left. - \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D)\setminus\mathcal{D}^+(D_0)}} - \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D_0)}} \right) \\
&= \sum_{\substack{j\in\{4,6\}\\ j<n}} \mathbf{1}_{(p_1,..,p_j)\in\mathcal{P}_j} \cdot \left( \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D_0) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D_0)}} - \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D)}} \right).
\end{aligned}$$

Thus plugging $\mathbf{\Delta}$ back into (1.31) gives

$$\begin{aligned}
V^*(D,z) - V^+(D,z) &= \sum_{j\in\{4,6\}} \sum_{\substack{p_j<\cdots<p_1<z \\ (p_1,..,p_j)\in\mathcal{P}_j}} \sum_{\text{odd } n>j} \times \\
&\qquad \times \sum_{p_n<\cdots p_{j+1}<p_j} g(p_1\cdots p_n) V(p_n) \left( \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D_0) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D_0)}} - \mathbf{1}_{\substack{p_1\cdots p_n\notin\mathcal{D}^+(D) \\ p_1\cdots p_{n-1}\in\mathcal{D}^+(D)}} \right).
\end{aligned}$$

Recalling the definition of $V_n(D,z)$ in (1.30), since $g$ is multiplicative we have

$$\begin{aligned}
&V^*(D,z) - V^+(D,z) \\
&\quad = \sum_{j\in\{4,6\}} \sum_{(p_1,..,p_j)\in\mathcal{P}_j} g(p_1\cdots p_j) \sum_{\text{odd } n>j} \Big( V_{n-j}\big(\tfrac{D_0}{p_1\cdots p_j}, p_j\big) - V_{n-j}\big(\tfrac{D}{p_1\cdots p_j}, p_j\big) \Big) \\
&\quad = \sum_{j\in\{4,6\}} \sum_{(p_1,..,p_j)\in\mathcal{P}_j} g(p_1\cdots p_j) \Big( V^+\big(\tfrac{D_0}{p_1\cdots p_j}, p_j\big) - V^+\big(\tfrac{D}{p_1\cdots p_j}, p_j\big) \Big). \qquad (1.34)
\end{aligned}$$

as $V^+(D,z) - V^+(D',z) = \sum_{\text{odd } n} [V_n(D,z) - V_n(D',z)]$.

Now assuming the two-sided condition (1.8) for $g$, the proof of [62, Theorem 11.12] (c.f. (12.4)–(12.8)) gives asymptotic equality,

$$V^+(D,z) \;=\; V(z)\Big\{F(\tfrac{\log D}{\log z}) \;+\; O\big((\log D)^{-1/6}\big)\Big\} \qquad (z \le D), \tag{1.35}$$

so that (1.34) becomes

$$\begin{aligned} V^*(D,z) \;=\; & V(z)\Big\{F(\tfrac{\log D}{\log z}) \;+\; O\big((\log D)^{-\frac{1}{6}}\big)\Big\} \\ & + \sum_{j\in\{4,6\}} \sum_{(p_1,..,p_j)\in\mathcal{P}_j} g(p_1\cdots p_j)V(p_j)\times \\ & \qquad \times \Big\{F(\tfrac{\log D_0/p_1\cdots p_j}{\log p_j}) - F(\tfrac{\log D/p_1\cdots p_j}{\log p_j}) \;+\; O\big(\log\big(\tfrac{D}{p_1\cdots p_j}\big)^{-\frac{1}{6}}\big)\Big\} \end{aligned} \tag{1.36}$$

By partial summation and the prime number theorem, for each $j$ we have,

$$\begin{aligned} & \sum_{(p_1,..,p_j)\in\mathcal{P}_j} g(p_1\cdots p_j)V(p_j)F(\tfrac{\log D_0/p_1\cdots p_j}{\log p_j}) \\ & \qquad = (\tfrac{7}{12}+\eta)\int_{(x_1,..,x_j)\in P_j} \frac{\mathrm{d}x_1\cdots\mathrm{d}x_j}{x_1\cdots x_{j-1}x_j^2} F\Big(\tfrac{\frac{7}{12}-x_1-\cdots x_j}{x_j}\Big) + O\big((\log D)^{-\frac{1}{6}}\big). \end{aligned}$$

Here $P_j$ is the polytope in Euclidean space $\mathbb{R}^j$ corresponding to $\mathcal{P}_j$, as below.

Hence from (1.36), we obtain

$$V^*(D,z) = V(z)\Big\{F^*(\tfrac{\log D}{\log z}) \;+\; O\big((\log D)^{-\frac{1}{6}}\big)\Big\} \qquad (z \le D), \tag{1.37}$$

where the function $F^*$ satisfies

$$\begin{aligned} & sF^*(s) \;-\; sF(s) \;= \\ & \qquad (\tfrac{7}{12}+\eta)\cdot \sum_{j\in\{4,6\}} \int_{(x_1,..,x_j)\in P_j} \frac{\mathrm{d}x_1\cdots\mathrm{d}x_j}{x_1\cdots x_{j-1}x_j^2} \Big[F\Big(\tfrac{\frac{7}{12}-x_1-\cdots x_j}{x_j}\Big) - F\Big(\tfrac{\frac{7}{12}+\eta-x_1-\cdots x_j}{x_j}\Big)\Big]. \end{aligned} \tag{1.38}$$

Namely, $P_4 \subset \mathbb{R}^4$ is given by

$$P_4 = \{(x_1,...,x_4) \in \mathrm{D}^+(\tfrac{7}{12}+\eta) \;:\; x_1 < \tfrac{1}{6}+2\eta \;\text{ and }\; x_2+x_4 > \tfrac{1}{4}-3\eta\},$$

and $P_6 = P_{6,1} \cup P_{6,2} \subset \mathbb{R}^6$ is given by

$$\begin{aligned} P_{6,1} = \{(x_1,...,x_6) \in \mathrm{D}^+(\tfrac{7}{12}+\eta) \;:\; & x_1+x_2 < \tfrac{1}{6}+2\eta \;\text{ and }\; x_6 > \tfrac{1}{12}-5\eta \\ & \text{and }\; x_2+x_3+x_4 > \tfrac{1}{4}-3\eta\}, \\ P_{6,2} = \{(x_1,...,x_6) \in \mathrm{D}^+(\tfrac{7}{12}+\eta) \;:\; & x_1,\; x_2+x_3 < \tfrac{1}{6}+2\eta \;\text{ and }\; x_6 > \tfrac{1}{12}-5\eta \\ & \text{and }\; x_1+x_4,\; x_2+x_3+x_4 > \tfrac{1}{4}-3\eta\}, \end{aligned} \tag{1.39}$$

Similarly, $\mathrm{D}^+$ is the set in Euclidean space corresponding to $\mathcal{D}^+$, namely,

$$\begin{aligned} \mathrm{D}^+(\tau) = \{(x_1,\ldots,x_r) \;: & x_1 > \cdots > x_r > 0, \\ & \text{and } x_1+\cdots x_{l-1}+3x_l < \tau \;\text{ for each odd } l \le r\}. \end{aligned}$$

Hence Proposition 1.4.1 follows.

### 1.4.1 Sieve function computation

We now compute $F^*$ in terms of $F$.

**Proposition 1.4.4.** *Let $\eta = \frac{1}{204}$. Then for $1 \le s \le 3$, we have*

$$F^*(s) \le 1.000081\, F(s). \tag{1.40}$$

*Proof.* From (1.38) we have

$$sF^*(s) = sF(s) + \left(\tfrac{7}{12} + \eta\right) \cdot 2e^\gamma \eta (J_4 + J_6),$$

for integrals $J_j$, $j \in \{4, 6\}$,

$$\begin{aligned} J_j &:= \frac{1}{2e^\gamma \eta} \int_{(x_1,..,x_j)\in P_j} \frac{\mathrm{d}x_1 \cdots \mathrm{d}x_j}{x_1 \cdots x_{j-1} x_j^2} \left[ F\Big(\tfrac{\frac{7}{12} - x_1 - \cdots x_j}{x_j}\Big) - F\Big(\tfrac{\frac{7}{12} + \eta - x_1 - \cdots x_j}{x_j}\Big)\right] \\ &= \int_{(x_1,..,x_j)\in P_j} \frac{\mathrm{d}x_1 \cdots \mathrm{d}x_j}{x_1 \cdots x_j} \Big[ (\tfrac{7}{12} - x_1 - \cdots x_j)(\tfrac{7}{12} + \eta - x_1 - \cdots x_j)\Big]^{-1}, \end{aligned}$$

since $sF(s) = 2e^\gamma$ for $s \in [1, 3]$. In particular $|P_j| = O(\eta^j)$ implies $J_j = O(\eta^{j+1})$, and so $F^*(s) = F(s) + O(\eta^5)$.

For $\eta = \frac{1}{204}$, we use Mathematica[2] to compute that

$$J_4 \le 0.016896. \tag{1.41}$$

Next we bound $J_6$. For $(x_1, .., x_6) \in P_6$ we have $x_4 < \frac{1}{2}(x_2 + x_3) < \frac{1}{12} + \eta$ and $\frac{7}{12} + \eta - x_1 - \cdots x_6 > x_5$ so

$$J_6 \le \int_{\overline{P_6}} \frac{\mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3}{x_1 x_2 x_3} \int_{\frac{1}{12} - 5\eta < x_6 < x_5 < x_4 < \frac{1}{12} + \eta} \frac{\mathrm{d}x_4 \mathrm{d}x_5 \mathrm{d}x_6}{x_4 x_5^2 x_6 (x_5 - \eta)},$$

where $\overline{P_6} = \overline{P_{6,1}} \cup \overline{P_{6,2}}$ is the (3-dimensional) projection of $P_6$, given explicitly by

$$\begin{aligned} \overline{P_{6,1}} = \{(x_1, x_2, x_3) \in \mathrm{D}^+(\tfrac{7}{12} + \eta) \ :\ & x_1 + x_2 < \tfrac{1}{6} + 2\eta \ \text{ and } \ x_3 > \tfrac{1}{12} - 5\eta \\ & \text{and } \ x_2 + 2x_3 > \tfrac{1}{4} - 3\eta\}, \\ \overline{P_{6,2}} = \{(x_1, x_2, x_3) \in \mathrm{D}^+(\tfrac{7}{12} + \eta) \ :\ & x_1,\ x_2 + x_3 < \tfrac{1}{6} + 2\eta \ \text{ and } \ x_3 > \tfrac{1}{12} - 5\eta \\ & \text{and } \ x_1 + x_3,\ x_2 + 2x_3 > \tfrac{1}{4} - 3\eta\}. \end{aligned}$$

For $\eta = \frac{1}{204}$, we compute $J_6 \le (J_{6,1} + J_{6,2}) J_{6,0}$ where

$$\begin{aligned} J_{6,0} &= \int_{\frac{1}{12} - 5\eta < x_6 < x_5 < x_4 < \frac{1}{12} + \eta} \frac{\mathrm{d}x_4 \mathrm{d}x_5 \mathrm{d}x_6}{x_4 x_5^2 x_6 (x_5 - \eta)} \le 2.33838, \\ J_{6,1} &= \int_{\overline{P_{6,1}}} \frac{\mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3}{x_1 x_2 x_3} \le 0.000806853, \\ J_{6,2} &= \int_{\overline{P_{6,2}}} \frac{\mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3}{x_1 x_2 x_3} \le 0.00397946. \end{aligned}$$

[2]The Mathematica package and code are available at `arxiv.org/abs/2109.02851`

Hence combining with (1.41), for $s \in [1,3]$ we conclude

$$\frac{F^*(s)}{F(s)} \;\le\; 1 + (\tfrac{7}{12}+\eta)\cdot\eta\big(J_4 + (J_{6,1}+J_{6,2})J_{6,0}\big) \;\le\; 1.000081. \qquad \square$$

## 1.5 Factorable remainder, after Iwaniec

In Theorem 1.2.8, Iwaniec constructed a well-factorable variant $\widetilde{\lambda}^{\pm}$ of the weights $\lambda^{\pm}$ from the (Jurkat–Richert) linear sieve. In this section, we prove Theorem 1.2.10 for the programmably factorable variant $\widetilde{\lambda}^*$ by adapting Iwaniec's construction, similarly building on the Jurkat–Richert type Proposition 1.4.1 that we obtained in the previous section. We shall also prove a technical variation on this result, with a variable level depending on anatomy of the moduli.

To set up the construction, we first adapt [62, Proposition 12.18]. Denote $P(z,u) = P(z)/P(u) = \prod_{u<p\le z} p$.

**Proposition 1.5.1.** *Let $\eta > 0$, and $D = x^{(\frac{7}{12}+\eta)/(1+\varepsilon+\tau)}$ for $\varepsilon > 0$ sufficiently small. Let $\mathbf{D}_r^*$ be defined by (1.49). Let $\lambda^{(r)}$ be the standard (upper and lower, for $r$ odd and even, resp.) weights for the linear sieve of level $D^\varepsilon$. Then for $u = D^{\varepsilon^2}$, $\tau = \varepsilon^9$,*

$$\begin{aligned} S(\mathcal{A},z) \;&\le\; |\mathcal{A}|V(z)\big\{F^*(s)+O(\varepsilon^5)\big\} \\ &\quad + \sum_{0\le r\le \varepsilon^{-2}} \sum_{(D_1,..,D_r)\in\mathbf{D}_r^*} \frac{(-1)^r}{\gamma(D_1,\dots,D_r)} \sum_{\substack{p_1\cdots p_r \mid P(z,u)\\ D_j<p_j\le D_j^{1+\tau}}} \sum_{\substack{b\mid P(u)\\ b\le D^\varepsilon}} \lambda^{(r)}(b)\, r_{\mathcal{A}}(bp_1\cdots p_r). \end{aligned} \tag{1.42}$$

*Proof.* First we write

$$S(\mathcal{A},z) \;\le\; S^*(\mathcal{A},z) - \sum_{\text{odd } n\le N} S_n(\mathcal{A},z) \tag{1.43}$$

for any $N \ge 1$, where

$$S^*(\mathcal{A},z) := \sum_{0\le r\le\varepsilon^{-2}} (-1)^r \sum_{\substack{u\le p_r<\cdots p_1<z\\ p_1\cdots p_r\in\mathcal{D}^*}} |\mathcal{A}_{p_1\cdots p_r}|, \quad S_n(\mathcal{A},z) := \sum_{\substack{p_n<\cdots<p_1<z\\ p_1\cdots p_{n-1}\in\mathcal{D}^*\\ p_1\cdots p_n\notin\mathcal{D}^*}} S(\mathcal{A}_{p_1\cdots p_n}, p_n).$$

We apply the inequality (1.43), not for $\mathcal{A} = (a_n)$ itself but rather for the subsequence $\tilde{\mathcal{A}} = (a_n \mathbf{1}_{(n,P(u))=1})$. Here we take $u = D^{\varepsilon^2}$, and then return to $\mathcal{A}$ by means of the Fundamental Lemma.

Let $z = D^{1/s}$ with $2 \le s \le \varepsilon^{-1}$. Since $z > u$, the only change to the above bound (1.43) when passing to $\tilde{\mathcal{A}}$ is the term $S^*(\tilde{\mathcal{A}},z)$, provided that $N$ is not too large in

terms of $\varepsilon$. Specifically, we require the lower bound for $p_n$ (by induction, the linear sieve conditions imply $p_1 \cdots p_m < D^{1-2^{-m}}$)

$$p_n \geq (D/p_1 \cdots p_{n-1})^{1/3} \geq D^{2^{1-N}/3} \tag{1.44}$$

to be larger than $u = D^{\varepsilon^2}$, which certainly holds provided

$$N \leq \frac{1}{2} \log \frac{1}{\varepsilon}. \tag{1.45}$$

Now it remains to evaluate $S^*(\tilde{\mathcal{A}}, z)$,

$$S^*(\tilde{\mathcal{A}}, z) = \sum_{0 \leq r \leq \varepsilon^{-2}} (-1)^r \sum_{\substack{u \leq p_r < \cdots p_1 < z \\ p_1 \cdots p_r \in \mathcal{D}^*}} |\tilde{\mathcal{A}}_{p_1 \cdots p_r}|. \tag{1.46}$$

For each $r$, we break the range in the inner sum into boxes. Namely, we let $D_1, ..., D_r$ run over numbers of form

$$D^{\varepsilon^2 (1+\tau)^j}, \quad j = 0, 1, 2, \ldots \tag{1.47}$$

with $\tau = \varepsilon^9$. We denote by $\mathbf{D}_r^+ = \mathbf{D}_r^+(D)$ the set of $r$-tuples $(D_1, ..., D_r)$ with $D_r \leq \cdots \leq D_1 \leq \sqrt{D}$, such that

$$\mathbf{D}_r^+ = \begin{cases} \{(D_1, ..., D_r) : D_1 \cdots D_{m-1} D_m^3 < D \quad \text{for all odd } m \leq r\} & \text{if } r \text{ even}, \\ \{(D_1, ..., D_r) : D_1 \cdots D_{m-1} D_m^3 < D^{1/(1+\tau)} \quad \text{for all odd } m \leq r\} & \text{if } r \text{ odd}. \end{cases}$$

We note, for $\varepsilon > 0$ sufficiently small, the cardinalities of the $\mathbf{D}_r^+$ are bounded by

$$\sum_{0 \leq r \leq \varepsilon^{-2}} |\mathbf{D}_r^+| \ \leq \ \exp(\varepsilon^{-3}). \tag{1.48}$$

Hereafter let $D = x^{(\frac{7}{12}+\eta)/(1+\tau+\varepsilon)}$, and define

$$\mathbf{D}_r^* = \{(D_1, ..., D_r) \in \mathbf{D}_r^+(D) : (D_1, ..., D_i) \notin \mathbf{P}_{i,r} \text{ for } i \leq r, i \in \{4, 6\}\}. \tag{1.49}$$

where $\mathbf{P}_{4,r}, \mathbf{P}_{6,r}$ are ($\tau$-enlarged, for even $r$) analogues of the polytopes $\mathcal{P}_4$, $\mathcal{P}_6$ in (1.17), e.g.

$$\mathbf{P}_{4,r} = \begin{cases} \{(D_1, \ldots, D_4) \ : \ D_1^{1+\tau} < x^{\frac{1}{6}+2\eta} \ \text{ and } \ (D_2 D_4)^{1/(1+\tau)} > x^{\frac{1}{4}-3\eta}\} & \text{if } r \text{ even}, \\ \{(D_1, \ldots, D_4) \ : \ D_1 < x^{\frac{1}{6}+2\eta} \ \text{ and } \ D_2 D_4 > x^{\frac{1}{4}-3\eta}\} & \text{if } r \text{ odd}. \end{cases}$$

Observe each integer $p_1 \cdots p_r$ has a unique vector $(D_1, \ldots, D_r)$ such that $p_i \in (D_i, D_i^{1+\tau}]$ for all $i \leq r$, inducing a map $\nu : \mathbb{N} \to \bigcup_r \mathbf{D}_r^+$. As a convention $\nu(1) = ()$ is the empty vector. By construction, for even $r$ if $p_1 \cdots p_4 \notin \mathcal{P}_4$ then $\nu(p_1 \cdots p_4) \notin \mathbf{P}_{4,r}$,

and if $(D_1, \dots, D_4) \notin \mathbf{P}_{4,r}$ then $\nu^{-1}(D_1, \dots, D_4) \cap \mathcal{P}_4 = \emptyset$ for odd $r$. Continuing this argument, we deduce

$$\begin{aligned} p_1 \cdots p_r \in \mathcal{D}^* \quad &\Longrightarrow \quad \nu(p_1 \cdots p_r) \in \mathbf{D}_r^* \qquad \text{if } r \text{ even}, \\ (D_1, \dots, D_r) \in \mathbf{D}_r^* \quad &\Longrightarrow \quad \nu^{-1}(D_1, \dots, D_r) \subset \mathcal{D}^* \qquad \text{if } r \text{ odd}. \end{aligned} \tag{1.50}$$

Without loss, we may restrict $\mathbf{D}_r^*$ to vectors with nonempty preimage in $\mathcal{D}^*$. Hence by construction, (1.46) becomes[3]

$$S^*(\tilde{\mathcal{A}}, z) \le \sum_{0 \le r \le \varepsilon^{-2}} \sum_{(D_1, .., D_r) \in \mathbf{D}_r^*} \frac{(-1)^r}{\gamma(D_1, \dots, D_r)} \sum_{\substack{p_1 \cdots p_r | P(z) \\ D_j < p_j \le D_j^{1+\tau}}} |\tilde{\mathcal{A}}_{p_1 \cdots p_r}|, \tag{1.51}$$

where $\gamma(D_1, \dots, D_r) = k_1! \cdots k_\ell!$ for the corresponding multiplicities $k_i \ge 1$ (i.e. we have $r = k_1 + \cdots + k_\ell$ and $D_1 = \cdots = D_{k_1} < D_{k_1+1} = \cdots = D_{k_2} < \cdots = D_r$.). Note the term $r = 0$ corresponds to $|\tilde{\mathcal{A}}|$ with $p_1 = \cdots = p_r = 1$.

Now by the Fundamental Lemma [62, Theorem 6.9], we have upper (and lower) bounds

$$\begin{aligned} |\tilde{\mathcal{A}}_{p_1 \cdots p_r}| &= S(\mathcal{A}_{p_1 \cdots p_r}, u) \\ &\le \; g(p_1 \cdots p_r) |\mathcal{A}| V(u) \{1 + O(e^{-1/\varepsilon})\} \; + \; \sum_{b \le D^\varepsilon} \lambda^{(r)}(b)\, r_{\mathcal{A}}(b p_1 \cdots p_r), \end{aligned}$$

(with $\le$ replaced by $\ge$ for the lower bound) where $\lambda^{(r)}$ is the upper (lower) bound $\beta$-sieve of level $D^\varepsilon$ when $r$ is even (odd). For further details on the Fundamental Lemma and $\beta$-sieve, we refer the reader to [62, §6 & §11].

Plugging back into (1.51), we get

$$\begin{aligned} S(\mathcal{A}, z) \; \le S^*(\tilde{\mathcal{A}}, z) \le \sum_{0 \le r \le \varepsilon^{-2}} \sum_{(D_1, .., D_r) \in \mathbf{D}_r^*} \frac{(-1)^r}{\gamma(D_1, \dots, D_r)} \sum_{\substack{p_1 \cdots p_r | P(z) \\ D_j < p_j \le D_j^{1+\tau}}} \times \\ \times \Big\{ g(p_1 \cdots p_r) |\mathcal{A}| V(u) \{1 + O(e^{-1/\varepsilon})\} \; + \; \sum_{b \le D^\varepsilon} \lambda^{(r)}(b)\, r_{\mathcal{A}}(b p_1 \cdots p_r) \Big\}. \end{aligned} \tag{1.52}$$

We now compare the main term above to that of the modified linear sieve, as in (1.29)–(1.37) from the proof of Proposition 1.4.1, namely,

$$V^*(D, z) := \sum_{d | P(z)/P(u)} \lambda^*(d) g(d) \; = \; V(z) \big\{ F^*(s) + o(1) \big\}.$$

[3]Indeed, we have reverse engineered the definition of $\mathbf{D}_r^*$ just so that (1.51) holds.

The difference between these main terms is accounted for by those $d$ with two close prime factors, within a ratio $D^\tau$, and those $d$ near the boundary. The former contribution is

$$\sum_{\substack{d|P(z)/P(u)\\ p\le p'<pD^\tau\\ pp'|d}} g(d) \le \sum_{\substack{u\le p<z\\ p\le p'<pD^\tau}} g(pp')\cdot \prod_{u\le p<z}(1+g(p)),$$

and the latter contribution is

$$\sum_r \sum_{\substack{u<p_r<\cdots p_1<z\\ D^{1/(1+\tau)}<p_1\cdots p_m^3<D}} g(p_1\cdots p_r), \tag{1.53}$$

where $m$ is the first index ($m\le r$) for which this occurs. Both contributions may be shown to be $O(\varepsilon^5)$, see [62, pp. 254-255]. Hence the Proposition follows. □

*Remark* 1.5.2. We make a minor technical point. Namely, at an admissible cost $O(\varepsilon^5)$ we may assume Proposition 1.42 holds, where $\mathbf{D}_r^*$ is further restricted to vectors $(D_1,\dots,D_r)$ satisfying

$$\nu^{-1}(D_1^{\frac{1}{1+\tau}},\dots,D_r^{\frac{1}{1+\tau}})\subset \mathcal{D}^*, \tag{1.54}$$

regardless of parity of $r$. To show this, note by definitions of $\mathbf{P}_{4,r},\mathbf{P}_{6,r}$ the integers $p_1\cdots p_r$ with $p_j\in[D_j^{1/(1+\tau)},D_j]$, $j\le r$, that lie outside $\mathcal{D}^*$ must satisfy

$$x^{\frac16+2\eta}/p_1 \quad\text{or}\quad x^{\frac{1}{12}-5\eta}/p_6 \;\in\; [x^{-2\tau},x^{2\tau}].$$

Then for $B_1=x^{\frac16+2\eta+2\tau}$, $B_6=x^{\frac{1}{12}-5\eta+2\tau}$, we have

$$\sum_{B_i\ge p_i\ge \max(B/x^{4\tau},u)} g(p_i) \;\ll\; \log\frac{\log B_i}{\log\max(B_i/x^{4\tau},u)} \ll \frac{\log x^{4\tau}}{\log u}\ll\frac{\tau}{\varepsilon^2},$$

and so the contribution of such integers to the main term of (1.52) is

$$\ll\sum_r \sum_{\substack{u<p_r<\cdots p_1<z\\ B_i\ge p_i>\max(B_i/x^{4\tau},u),\ i\in\{1,6\}}} g(p_1\cdots p_r)\ll\frac{\tau}{\varepsilon^2}\prod_{u<p<z}\big(1+g(p)\big)\ll\frac{\tau}{\varepsilon^2}\frac{\log z}{\log u}\ll\varepsilon^5.$$

Hence (1.54) follows.

We now proceed to Theorem 1.2.10. For each vector $(D_1,..,D_r)\in\mathbf{D}_r$ we define the weight $\lambda_{(D_1,..,D_r)}$ supported on $d$ in $\nu^{-1}(D_1,..,D_r)$, namely,

$$\lambda_{(D_1,..,D_r)}(d) \;:=\; \mathbf{1}^{d=p_1\cdots p_r}_{D_j<p_j\le D_j^{1+\tau}\ \forall j}. \tag{1.55}$$

Next, we may decompose an integer $d$ into its $u$-smooth and rough components, $d = b(p_1 \cdots p_r)$. Recall for $b \mid P(u)$ we have $\lambda^{(r)}(b) = \mu(b)$ if $b \in \mathcal{D}^{(r)}$, and $\lambda^{(r)}(b) = 0$ else. Thus the convolution $\lambda^{(r)}_{(D_1,..,D_r)} := \lambda_{(D_1,..,D_r)} * \lambda^{(r)}$ equals

$$\lambda^{(r)}_{(D_1,..,D_r)}(d) \;=\; \begin{cases} \mu(b) & \text{if } d = bp_1 \cdots p_r \;\text{ for }\; b \in \mathcal{D}^{(r)}(D^{\varepsilon}),\; b \mid P(D^{\varepsilon^2}), \\ & \qquad\qquad \text{and }\; D_j < p_j \leq \; D_j^{1+\tau}\; \forall j \leq r. \\ 0 & \text{else.} \end{cases} \tag{1.56}$$

Hence the remainder in (1.42) equals

$$\sum_{0 \leq r \leq \varepsilon^{-2}} \sum_{(D_1,..,D_r) \in \mathbf{D}^*_r} \frac{(-1)^r}{\gamma(D_1, \ldots, D_r)} \sum_{d \mid P(z)} \lambda^{(r)}_{(D_1,..,D_r)}(d)\, r_{\mathcal{A}}(d) \;=\; \sum_{d \mid P(z)} \widetilde{\lambda}^*(d)\, r_{\mathcal{A}}(d),$$

for the weights

$$\widetilde{\lambda}^* = \sum_{0 \leq r \leq \varepsilon^{-2}} \sum_{(D_1,..,D_r) \in \mathbf{D}^*_r} \frac{(-1)^r}{\gamma(D_1, \ldots, D_r)} \lambda^{(r)}_{(D_1,..,D_r)}. \tag{1.57}$$

Recalling the cardinality of $\mathbf{D}^*_r \subset \mathbf{D}^+_r$ from (1.48), it suffices to show the weights $\lambda^{(r)}_{(D_1,..,D_r)}$ are programmably factorable for each vector in $\mathbf{D}^*_r$. To this we have the following.

**Lemma 1.5.3.** *For an integer $d$ denote the vector $\nu(d) = (D_1, \ldots, D_r)$ from* (1.50)*. If $d$ has a factorization as in* (1.18) *at level $D$, then the corresponding weights $\lambda_{\nu(d)}$ and $\lambda^{(r)}_{\nu(d)}$, as in* (1.55) *and* (1.56)*, resp., are programmably factorable sequences of levels $D^{1+\tau}$ and $D^{1+\tau+\varepsilon} = x^{\frac{7}{12}+\eta}$, resp. (relative to $x$, $\varepsilon/50$).*

*Proof.* By assumption for each $N \in [1, x^{1/3}]$, there is a factorization $d = d_1 d_2 d_3$ satisfying the system (1.18). For $j = 1, 2, 3$, write $d_j = \prod_{i \in I_j} p_i$ for the induced partition of indices $\{1, ..., r\} = I_1 \cup I_2 \cup I_3$. Thus letting $Q_j = \prod_{i \in I_j} D_i$, the factorization $D_1 \cdots D_r = Q_1 Q_2 Q_3$ satisfies (1.6), since $D_i < p_i$.

Further, writing the corresponding subvectors $(D_i)_{i \in I_j}$ for $j = 1, 2, 3$, the weights $\lambda_{(D_i)_{i \in I_j}}$ are 1-bounded, supported on $[1, Q_j^{1+\tau}]$, and give the desired triple convolution,

$$\lambda_{(D_1,...,D_r)} = \lambda_{(D_i)_{i \in I_1}} * \lambda_{(D_i)_{i \in I_2}} * \lambda_{(D_i)_{i \in I_3}}.$$

Hence $\lambda_{(D_1,...,D_r)}$ is programmably factorable of level $D^{1+\tau}$ as claimed. Similarly $\lambda^{(r)}_{(D_1,...,D_r)} = \lambda_{(D_i)_{i \in I_1}} * \lambda_{(D_i)_{i \in I_2}} * \lambda^{(r)}_{(D_3)_{i \in I_1}}$ is programmably factorable of level $D^{1+\tau+\varepsilon}$. □

Now for each vector $(D_1, \ldots, D_r) \in \mathbf{D}^*_r$, by (1.54) there exists $d = p_1 \cdots p_r \in \mathcal{D}^*$ for some primes $p_i \in (D_i^{1/(1+\tau)}, D_i]$. Then for all $N \in [1, x^{1/3}]$ Proposition 1.3.3 gives a factorization of $d$ as in (1.18), and so by Lemma 1.5.3 $\lambda^{(r)}_{(D_1,...,D_r)}$ is programmably factorable of level $x^{\frac{7}{12}+\eta}$.

This completes the proof of Theorem 1.2.10.

### 1.5.1 Variable level of distribution for the linear sieve weights

We now return to Iwaniec's well-factorable weights $\widetilde{\lambda}^{\pm}$ for the (upper and lower) linear sieve, given explicitly from (1.56) as the weighted sum,

$$\widetilde{\lambda}^{\pm} = \sum_{0\le r\le \varepsilon^{-2}} \sum_{(D_1,..,D_r)\in \mathbf{D}_r^{\pm}} \frac{(-1)^r}{\gamma(D_1,\dots,D_r)}\, \lambda^{(r)}_{(D_1,..,D_r)}. \tag{1.58}$$

We introduce the analogous set of well-factorable vectors $\mathbf{D}_r^{\text{well}} = \mathbf{D}_r^{\text{well}}(D)$,

$$\mathbf{D}_r^{\text{well}} = \{(D_1,...,D_r) : D_1\cdots D_{m-1}D_m^2 < D \quad \text{for all } m\le r\}. \tag{1.59}$$

Note $\mathbf{D}_r^{\pm}\subset \mathbf{D}_r^{\text{well}}$, having dropped parity conditions on the indices $m\le r$.

We have the following technical variation on Theorem 1.2.10 for the original linear sieve.

**Proposition 1.5.4.** *Let $(D_1,...,D_r)\in \mathbf{D}_r^{\text{well}}(D)$ and write $D=x^{\theta}$, $D_i = x^{t_i}$ for $i\le r$. If $\theta\le \theta(t_1)-\varepsilon$ as in (1.25), then*

$$\sum_{\substack{b=p_1\cdots p_r\\ D_i<p_i\le D_i^{1+\tau}}} \sum_{\substack{d=bc\le x^{\theta}\\ c|P(p_r)\\ (d,a)=1}} \widetilde{\lambda}^{\pm}(d)\left(\pi(x;d,a)-\frac{\pi(x)}{\phi(d)}\right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}. \tag{1.60}$$

*Moreover if $t_1\le \frac15$ and $r\ge 3$, then (1.60) holds provided that $\theta\le \theta(t_1,t_2,t_3)-\varepsilon$ as in (1.26).*

*If $t_1\le \frac15$ and $r\le 2$, then provided $\theta\le \frac{3-u}{5}-\varepsilon$,*

$$\sum_{\substack{b=p_1\cdots p_r\\ D_i<p_i\le D_i^{1+\tau}}} \sum_{\substack{d=bc\le x^{\theta}\\ c|P(x^u)\\ (d,a)=1}} \widetilde{\lambda}^{\pm}(d)\left(\pi(x;d,a)-\frac{\pi(x)}{\phi(d)}\right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

*In particular for $r=0$ (i.e. the empty vector), $\theta\le \frac{3-u}{5}-\varepsilon$, this simplifies as*

$$\sum_{\substack{d\le x^{\theta}\\ d|P(x^u)\\ (d,a)=1}} \widetilde{\lambda}^{\pm}(d)\left(\pi(x;d,a)-\frac{\pi(x)}{\phi(d)}\right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

*Proof.* Given $(D_1,...,D_r)$, take an integer $b=p_1\cdots p_r$ with $D_i<p_i\le D_i^{1+\tau}$. Then for all multiples $d$ of $b$ with $d/b \mid P(p_r)$, the weight $\lambda^{(s)}_{(D_1',..,D_s')}(d)$ vanishes unless the vector $(D_1',..,D_s')$ extends $(D_1,...,D_r)$. That is, $D_i'=D_i$ for all $i\le r$. Conversely,

given such a vector $(D'_1, .., D'_s)$ we have $\lambda^{(s)}_{(D'_1,..,D'_s)}(d') = 0$ unless $d' = p_1 \cdots p_s$ with $D_i < p_i \leq D_i^{1+\tau}$, $i \leq r$. So by definition of $\widetilde{\lambda}^{\pm}$ as in (1.58), we have

$$\sum_{\substack{b=p_1\cdots p_r \\ D_i<p_i\leq D_i^{1+\tau}}} \sum_{\substack{d=bc\leq x^\theta \\ c|P(p_r) \\ (d,a)=1}} \widetilde{\lambda}^{\pm}(d) = \sum_{r\leq s\leq \varepsilon^{-2}} \sum_{\substack{(D'_1,\dots,D'_s)\in \mathbf{D}_s^{\pm} \\ D'_i=D_i,\, i\leq r}} \frac{(-1)^s}{\gamma(D'_1,\dots,D'_s)} \sum_{\substack{d\leq x^\theta \\ (d,a)=1}} \lambda^{(s)}_{(D'_1,..,D'_s)}(d). \tag{1.61}$$

Here we have extended (by zero) the inner sum to all $d \leq x^\theta$, $(d, a) = 1$.

Next, take such a vector $(D'_1, ..., D'_s) \in \mathbf{D}_s^{\pm}(x^\theta)$ with $D'_i = D_i$ for $i \leq r$. Each integer $d = p_1 \cdots p_s$ with $D'_i < p_i \leq (D'_i)^{1+\tau}$ lies in $d \in \mathcal{D}^{\text{well}}(x^{\theta+\tau})$. In particular $p_1 \leq D_1^{1+\tau} \leq x^{t_1+\tau}$ so by Corollary 1.3.7.1, $d$ has a factorization as in (1.18) at level $x^{\theta(t_1+\tau)}$. Since $\theta(t)$ is continuous (in fact, piecewise linear), for $\tau > 0$ sufficiently small $\theta(t_1 + \tau) \geq \theta(t_1) - \varepsilon \geq \theta$. Thus by Lemma 1.5.3 the weights $\lambda^{(s)}_{(D_1,..,D_s)}$ are programmably factorable sequences of level $x^\theta$. Hence for each such vector, by Theorem 1.2.5 we have

$$\sum_{\substack{d\leq x^\theta \\ (d,a)=1}} \lambda^{(s)}_{(D'_1,..,D'_s)}(d) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}. \tag{1.62}$$

Plugging (1.62) back into (1.61) gives the bound (1.60), as claimed.

Moreover, if $t_1 \leq \frac{1}{5}$ and $r \geq 3$ then proceeding as in the above paragraph, by Corollary 1.3.7.1 $d$ will factor as in (1.18) to level $x^{\theta(t_1+\tau,t_2+\tau,\tau_3+\tau)}$. Again $\theta(t, u, v)$ is continuous, so for $\tau > 0$ sufficiently small $\theta(t_1+\tau, t_2+\tau, \tau_3+\tau) \geq \theta(t_1, t_2, t_3)-\varepsilon \geq \theta$. Hence (1.60) also follows in this case.

Similarly if $t_1 \leq \frac{1}{5}$ and $r \leq 2$, proceeding as above with $d = p_1 \cdots p_s$, the assumption $d/b \mid P(x^u)$ implies $s \leq 2$ or $p_3 \leq D_3^{1+\tau} \leq x^{u+\tau}$. Thus by Corollary 1.3.7.1 $d$ will factor as in (1.18) to level $x^{(3-u-\tau)/5} \geq x^\theta$. Hence (1.60) follows in this case as well. □

## 1.5.2 Equidistribution for products of primes

We use an extension of Theorem 1.2.5 to products of $k$ primes. This is the analogue in the programmably factorable setting of Lemma 7 [140] extending Theorem 1.2.2 of Bombieri–Friedlander–Iwaniec.

**Proposition 1.5.5.** *Let $\varepsilon > 0$ and $\lambda$ be a programmably factorable sequence of level $D \leq x^{\frac{3}{5}-\varepsilon}$ (relative to $x$, $\varepsilon/50$). Take real numbers $\varepsilon_1, \dots, \varepsilon_k \geq \varepsilon$ such that $\sum_{i\leq k} \varepsilon_i =$*

1. *Then for any fixed integer $a \in \mathbb{Z}$, $A, B > 0$, letting $\Delta = 1 + (\log x)^{-B}$ we have*

$$\sum_{\substack{d \le D \\ (d,a)=1}} \lambda(d) \Bigg( \sum_{\substack{p_1 \cdots p_k \equiv a \,(\mathrm{mod}\, d) \\ x^{\varepsilon_i}/\Delta < p_i \le x^{\varepsilon_i} \,\forall i \le k}} 1 \; - \; \frac{1}{\phi(d)} \sum_{\substack{(p_1 \cdots p_k, d)=1 \\ x^{\varepsilon_i}/\Delta < p_i \le x^{\varepsilon_i} \,\forall i \le k}} 1 \Bigg) \ll_{a,\varepsilon,A,B} \frac{x}{(\log x)^A}. \tag{1.63}$$

*Proof.* This follows by the same proof method as in Theorem 1.2.5 (i.e. Maynard's [102, Theorem 1.1]). Indeed, Maynard just uses the Heath–Brown identity to decompose the indicator function of primes into Type I/II sums. A similar decomposition holds for products of $k$ primes, after which we may apply the same Type I/II estimates in Propositions 5.1 and 5.2 of [102]. □

In addition, by replacing Theorem 1.2.5 with Proposition 1.5.5 in the proof, we obtain analogues of Proposition 1.5.4 and 1.5.5.1 for the linear sieve weights $\lambda = \widetilde{\lambda}^{\pm}$ in the case of products of $k$ primes.

**Corollary 1.5.5.1.** *Let $(D_1, ..., D_r) \in \mathbf{D}_r^{\mathrm{well}}(D)$ and write $D = x^{\theta}$, $D_i = x^{t_i}$ for $i \le r$. Let $\varepsilon > 0$ and real numbers $\varepsilon_1, \ldots, \varepsilon_k \ge \varepsilon$ such that $\sum_{i \le k} \varepsilon_i = 1$. Fix an integer $a \in \mathbb{Z}$, $A, B > 0$, and let $\Delta = 1 + (\log x)^{-B}$. If $\theta \le \theta(t_1) - \varepsilon$ as in* (1.25),

$$\sum_{\substack{b = p_1' \cdots p_r' \\ D_i < p_i' \le D_i^{1+\tau}}} \sum_{\substack{d = bc \le x^{\theta} \\ c \mid P(p_r') \\ (d,a)=1}} \widetilde{\lambda}^{\pm}(d) \Bigg( \sum_{\substack{p_1 \cdots p_k \equiv a \,(\mathrm{mod}\, d) \\ x^{\varepsilon_i}/\Delta < p_i \le x^{\varepsilon_i} \,\forall i \le k}} 1 \; - \; \frac{1}{\phi(d)} \sum_{\substack{(p_1 \cdots p_k, d)=1 \\ x^{\varepsilon_i}/\Delta < p_i \le x^{\varepsilon_i} \,\forall i \le k}} 1 \Bigg) \tag{1.64}$$
$$\ll_{a,\varepsilon,A,B} \frac{x}{(\log x)^A}.$$

*Moreover if $t_1 \le \frac{1}{5}$, $r \ge 3$, then* (1.64) *holds provided that $\theta \le \theta(t_1, t_2, t_3) - \varepsilon$ as in* (1.26).

*In addition, if $r \le 2$, $u \le t_r$, $t_1 \le \frac{1}{5}$, and $\theta \le \frac{3-u}{5} - \varepsilon$, then*

$$\sum_{\substack{b = p_1' \cdots p_r' \\ D_i < p_i' \le D_i^{1+\tau}}} \sum_{\substack{d = bc \le x^{\theta} \\ c \mid P(x^u) \\ (d,a)=1}} \widetilde{\lambda}^{\pm}(d) \Bigg( \sum_{\substack{p_1 \cdots p_k \equiv a \,(\mathrm{mod}\, d) \\ x^{\varepsilon_i}/\Delta < p_i \le x^{\varepsilon_i} \,\forall i \le k}} 1 \; - \; \frac{1}{\phi(d)} \sum_{\substack{(p_1 \cdots p_k, d)=1 \\ x^{\varepsilon_i}/\Delta < p_i \le x^{\varepsilon_i} \,\forall i \le k}} 1 \Bigg) \tag{1.65}$$
$$\ll_{a,\varepsilon,A,B} \frac{x}{(\log x)^A}.$$

## 1.6 Upper bound for twin primes

Now we shall apply the modified sieve to the set of twin primes

$$\mathcal{A} := \{p + 2 : p \le x\}.$$

In this case the sieve notation specializes as $\mathcal{P} = \{p > 2\}$ and $g(d) = 1/\phi(d)$ for odd $d$, so that $V(z) = \prod_{2<p<z}(1 - 1/\phi(p))$. Recall by Mertens theorem $V(z) \sim \mathfrak{S}_2/e^\gamma \log z$.

We begin in the spirit of Fouvry–Grupp [58], and apply a weighted sieve inequality. To each non-switched term, we apply the Buchstab identity in order to lower the sieve threshhold down to $z = x^\epsilon$ for some tiny $\epsilon > 0$. By Proposition 1.5.4, such smooth sums will satisfy level of distribution $\frac{3-\epsilon}{5}$. Combined with variations on a theme, which identify programmably-factorable weights in certain cases, the consequent increase in level will be sufficient to obtain the bound in Theorem 1.1.2. For a final bit of savings, we use refinements obtained by Wu's iteration method [141].

Before moving on, we note that sieve methods and the switching principle has also yielded progress on the Goldbach problem. Indeed, the upper bound in Wu [141, Theorem 3] for twin primes is obtained by using the same formulae as in [141, Theorem 1] for the Goldbach problem, except for altering the level from $\frac{1}{2}$ to $\frac{4}{7}$ (this amounts to replacing factors of 4 with 7/2 in a few instances). Importantly, the quantitative upper bounds for twin primes are much stronger than those of the Goldbach problem. This is because the latter relies on level $\frac{1}{2}$ from Bombieri–Vinogradov for a growing residue $a = N$, while the former may appeal to level of distribution $\frac{4}{7}$ from Bombieri–Friedlander–Iwaniec for the fixed residue $a = 2$ (and now the subsequent improvements of Maynard). As such our methods have nothing new to say for the Goldbach problem.

### 1.6.1 Lemmas

We begin with a standard lemma for $x^{1/u}$-rough numbers in terms of the Buchstab function $\omega(u) = \big(f(u) + F(u)\big)/(2e^\gamma)$ for linear sieve functions $f$, $F$ as in Theorem 1.2.8. Alternatively, $\omega$ is directly defined via

$$\begin{aligned} \omega(u) &= \frac{1}{u} && \text{for} \quad 1 \le u \le 2, \\ \big(u\,\omega(u)\big)' &= \omega(u-1) && \text{for} \quad 2 \le u. \end{aligned}$$

**Lemma 1.6.1.** *Let $x \ge 2$ and $y = x^{1/u}$. Then we have*

$$\sum_{\substack{n \le x \\ p|n \Rightarrow p \ge y}} 1 = \omega(u)\,\frac{x}{\log y} + O\Big(\frac{x}{(\log y)^2}\Big).$$

*Proof.* This is [140, Lemma 12]. □

The argument of Wu makes essential use of weighted sieve inequalities, as in Lemmas 4.1 and 4.2 [141]. We shall employ the latter inequality in the special case $d = 1$, $\sigma = 1$.

**Lemma 1.6.2.** *For $3/10 \geq \rho \geq \tau_3 > \tau_2 > \tau_1 \geq \rho' \geq 1/20$, we have*

$$5S(\mathcal{A}, x^\rho) \lesssim \sum_{1 \leq n \leq 21} \Gamma_n,$$

*where*

$$\Gamma_1 := 4S(\mathcal{A}, x^{\rho'}) + S(\mathcal{A}, x^{\tau_1}),$$

$$\Gamma_2 := -\sum_{x^{\rho'} \leq p < x^\rho} S(\mathcal{A}_p, x^{\rho'}),$$

$$\Gamma_3 := -\sum_{x^{\rho'} \leq p < x^{\tau_2}} S(\mathcal{A}_p, x^{\rho'}),$$

$$\Gamma_4 := -\sum_{x^{\rho'} \leq p < x^{\tau_3}} S(\mathcal{A}_p, x^{\rho'}),$$

$$\Gamma_5 := \sum\sum_{x^{\rho'} \leq p_1 < p_2 < x^{\tau_2}} S(\mathcal{A}_{p_1p_2}, x^{\rho'}),$$

$$\Gamma_6 := \sum\sum_{\substack{x^{\rho'} \leq p_1 < x^{\tau_1} \\ x^{\tau_2} \leq p_2 < x^{\tau_3}}} S(\mathcal{A}_{p_1p_2}, x^{\rho'}),$$

$$\Gamma_7 := \sum\sum_{x^{\rho'} \leq p_1 < p_2 < x^{\tau_1}} S(\mathcal{A}_{p_1p_2}, p_1),$$

$$\Gamma_8 := \sum\sum_{x^{\rho'} \leq p_1 < x^{\tau_1} \leq p_2 < x^{\tau_2}} S(\mathcal{A}_{p_1p_2}, p_1),$$

$$\Gamma_9 := \sum\sum\sum_{x^{\tau_1} \leq p_1 < p_2 < p_3 < x^{\tau_3}} S(\mathcal{A}_{p_1p_2p_3}, p_2),$$

$$\Gamma_{10} := \sum\sum\sum_{x^{\tau_1} \leq p_1 < p_2 < x^{\tau_2} \leq p_3 < x^\rho} S(\mathcal{A}_{p_1p_2p_3}, p_2),$$

$$\Gamma_{11} := \sum\sum\sum_{x^{\tau_1} \leq p_1 < x^{\tau_2} \leq p_2 < p_3 < x^{\tau_3}} S(\mathcal{A}_{p_1p_2p_3}, p_2),$$

$$\Gamma_{12} := \sum\sum\sum_{\substack{x^{\rho'} \leq p_1 < p_2 < x^{\tau_1} \\ x^{\tau_3} \leq p_3 < x^\rho}} S(\mathcal{A}_{p_1p_2p_3}, p_2),$$

$$\Gamma_{13} := \sum\sum\sum_{x^{\rho'} \leq p_1 < x^{\tau_1} \leq p_2 < x^{\tau_2} \leq p_3 < x^\rho} S(\mathcal{A}_{p_1p_2p_3}, p_2),$$

$$\Gamma_{14} := \sum\sum\sum_{\substack{x^{\rho'} \leq p_1 < x^{\tau_1} \\ x^{\tau_2} \leq p_2 < p_3 < x^\rho}} S(\mathcal{A}_{p_1p_2p_3}, p_2),$$

$$\Gamma_{15} := \sum\sum\sum_{x^{\tau_1} \leq p_1 < x^{\tau_2} \leq p_2 < x^{\tau_3} \leq p_3 < x^\rho} S(\mathcal{A}_{p_1p_2p_3}, p_2),$$

$$\Gamma_{16} := \sum\sum\sum\sum_{x^{\tau_2} \leq p_1 < p_2 < p_3 < p_4 < x^{\tau_3}} S(\mathcal{A}_{p_1p_2p_3p_4}, p_3),$$

$$\Gamma_{17} := \sum\sum\sum\sum_{x^{\tau_2} \leq p_1 < p_2 < p_3 < x^{\tau_3} \leq p_4 < x^\rho} S(\mathcal{A}_{p_1p_2p_3p_4}, p_3),$$

$$\Gamma_{18} := \sum\sum\sum\sum_{x^{\tau_2} \leq p_1 < p_2 < x^{\tau_3} \leq p_3 < p_4 < x^\rho} S(\mathcal{A}_{p_1p_2p_3p_4}, p_3),$$

$$\Gamma_{19} := \sum\sum\sum\sum\sum_{\substack{x^{\tau_1} \leq p_1 x^{\tau_2} \\ x^{\tau_3} \leq p_2 < p_3 < p_4 < x^\rho}} S(\mathcal{A}_{p_1p_2p_3p_4}, p_3),$$

$$\Gamma_{20} := \sum\sum\sum\sum\sum_{x^{\tau_2} \leq p_1 < x^{\tau_3} \leq p_2 < p_3 < p_4 < p_5 < x^\rho} S(\mathcal{A}_{p_1p_2p_3p_4p_5}, p_4),$$

$$\Gamma_{21} := \sum\sum\sum\sum\sum\sum_{x^{\tau_3} \leq p_1 < p_2 < p_3 < p_4 < p_5 < p_6 < x^\rho} S(\mathcal{A}_{p_1p_2p_3p_4p_5p_6}, p_5).$$

*Proof.* This is [141, Lemma 4.2] from Wu with $d = 1$, $\sigma = 1$. Here we simplify Wu's notation slightly, using $(\underline{d}^{1/s}, \underline{d}^{1/\kappa_3}, \underline{d}^{1/\kappa_2}, \underline{d}^{1/\kappa_1}, \underline{d}^{1/s'}) = (x^\rho, x^{\tau_3}, x^{\tau_2}, x^{\tau_1}, x^{\rho'})$. The basic proof idea is to iterate the Buchstab identity and to strategically neglect some terms by positivity. □

### 1.6.2 Computations

Given $0.1 \le \rho' \le \tau_1 < 0.2 \le \tau_2 < \tau_3 \le \rho \le 0.3.$, we define integrals $I_n = I_n(\rho, \rho', \tau_1, \tau_2, \tau_3)$ by

$$\begin{aligned}
I_n &= \int_{\mathbb{D}_n} \omega\Big(\frac{1-t-u-v}{u}\Big)\frac{\mathrm{d}t\mathrm{d}u\mathrm{d}v}{tu^2v} && (9 \le n \le 15),\\
I_n &= \int_{\mathbb{D}_n} \omega\Big(\frac{1-t-u-v-w}{v}\Big)\frac{\mathrm{d}t\mathrm{d}u\mathrm{d}v\mathrm{d}w}{tuv^2w} && (16 \le n \le 19),\\
I_{20} &= \int_{\mathbb{D}_{20}} \omega\Big(\frac{1-t-u-v-w-x}{w}\Big)\frac{\mathrm{d}t\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}x}{tuvw^2x},\\
I_{21} &= \int_{\mathbb{D}_{21}} \omega\Big(\frac{1-t-u-v-w-x-y}{x}\Big)\frac{\mathrm{d}t\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}x\mathrm{d}y}{tuvwx^2y},
\end{aligned} \tag{1.66}$$

where $\omega$ is the Buchstab function, and where the domains $\mathbb{D}_n$ are

$$\begin{aligned}
\mathbb{D}_9 &= \{(t,u,v) : \tau_1 < t < u < v < \tau_3\},\\
\mathbb{D}_{10} &= \{(t,u,v) : \tau_1 < t < u < \tau_2 < v < \rho\},\\
\mathbb{D}_{11} &= \{(t,u,v) : \tau_1 < t < \tau_2 < u < v < \tau_3\},\\
\mathbb{D}_{12} &= \{(t,u,v) : \rho' < t < u < \tau_1,\ \tau_3 < v < \rho\},\\
\mathbb{D}_{13} &= \{(t,u,v) : \rho' < t < \tau_1 < u < \tau_2 < v < \rho\},\\
\mathbb{D}_{14} &= \{(t,u,v) : \rho' < t < \tau_1,\ \tau_2 < u < v < \rho\},\\
\mathbb{D}_{15} &= \{(t,u,v) : \tau_1 < t < \tau_2 < u < \tau_3 < v < \rho\},\\
\mathbb{D}_{16} &= \{(t,u,v,w) : \tau_2 < t < u < v < w < \tau_3\},\\
\mathbb{D}_{17} &= \{(t,u,v,w) : \tau_2 < t < u < v < \tau_3 < w < \rho\},\\
\mathbb{D}_{18} &= \{(t,u,v,w) : \tau_2 < t < u < \tau_3 < v < w < \rho\},\\
\mathbb{D}_{19} &= \{(t,u,v,w) : \tau_1 < t < \tau_2,\ \tau_3 < u < v < w < \rho\},\\
\mathbb{D}_{20} &= \{(t,u,v,w,x) : \tau_2 < t < \tau_3 < u < v < w < x < \rho\},\\
\mathbb{D}_{21} &= \{(t,u,v,w,x,y) : \tau_3 < t < u < v < w < x < y < \rho\}.
\end{aligned}$$

Recall the definitions (1.25) and (1.26),

$$\theta(t) = \begin{cases} \frac{2-t}{3} & \text{if } t > \frac{1}{5}, \\ \frac{1+t}{2} & \text{if } t \le \frac{1}{5}. \end{cases}$$

and

$$\theta(t,u,v) := \max\Big\{ \frac{3-v}{5},\, \theta(t),\, \theta(u),\, \theta(t+u+v),\\ \theta(t+u),\, \theta(t+v),\, \theta(u+v)\Big\}.$$

We also define

$$\begin{aligned} G_1 &= 4G(\rho') + G(\tau_1), & G_3 &= G_0 + \overline{G}(\tau_2), \\ G_2 &= G_0 + \overline{G}(\rho), & G_4 &= G_0 + \overline{G}(\tau_3), \end{aligned} \tag{1.67}$$

where for $c \le 1/5$,

$$\begin{aligned} G(c) = \frac{1}{\epsilon} F\big(\theta_\epsilon/\epsilon\big) - \frac{1}{\epsilon}\int_\epsilon^c \frac{\mathrm{d}t}{t} f\big((\theta_\epsilon - t)/\epsilon\big) + \frac{1}{\epsilon}\int_\epsilon^c \int_\epsilon^t \frac{\mathrm{d}t\mathrm{d}u}{tu} F\big((\theta_\epsilon - t - u)/\epsilon\big) \\ - \int_\epsilon^c \int_\epsilon^t \int_\epsilon^u \frac{\mathrm{d}t\mathrm{d}u\mathrm{d}v}{tuv^2} f\big((\theta(t,u,v) - t - u - v)/v\big), \end{aligned} \tag{1.68}$$

and for $c > 1/5$,

$$\overline{G}(c) = -\frac{1}{\epsilon}\int_{1/5}^c \frac{\mathrm{d}t}{t} f\big((\theta(t) - t)/\epsilon\big) + \int_{1/5}^c \int_\epsilon^{\rho'} \frac{\mathrm{d}t\mathrm{d}u}{tu^2} F\big((\theta(t) - t - u)/u\big) \tag{1.69}$$

as well as

$$\begin{aligned} G_0 = -\frac{1}{\epsilon}\int_{\rho'}^{1/5} \frac{\mathrm{d}t}{t} f\big((\theta_\epsilon - t)/\epsilon\big) + \frac{1}{\epsilon}\int_{\rho'}^{1/5} \int_\epsilon^{\rho'} \frac{\mathrm{d}t\mathrm{d}u}{tu} F\big((\theta_\epsilon - t - u)/\epsilon\big) \\ - \int_{\rho'}^{1/5} \int_\epsilon^{\rho'} \int_\epsilon^u \frac{\mathrm{d}t\mathrm{d}u\mathrm{d}v}{tuv^2} f\big((\theta(t,u,v) - t - u - v)/v\big). \end{aligned} \tag{1.70}$$

We similarly let

$$\begin{aligned} G_5 &= \frac{1}{\epsilon}\int_{\rho'}^{1/5} \int_{\rho'}^t \frac{\mathrm{d}t\mathrm{d}u}{tu} F\big((\theta_\epsilon - t - u)/\epsilon\big) \;+\; \frac{1}{\rho'}\int_{1/5}^{\tau_2} \int_{\rho'}^t \frac{\mathrm{d}t\mathrm{d}u}{tu} F\big((\theta(t) - t - u)/\rho'\big) \\ &\quad - \int_{\rho'}^{1/5} \int_{\rho'}^t \int_\epsilon^{\rho'} \frac{\mathrm{d}t\mathrm{d}u\mathrm{d}v}{tuv^2} f\big((\theta(t,u,v) - t - u - v)/v\big), \\ G_6 &= \frac{1}{\rho'}\int_{\tau_2}^{\tau_3} \int_{\rho'}^{\tau_1} \frac{\mathrm{d}t\mathrm{d}u}{tu} F\big((\theta(t) - t - u)/\rho'\big), \end{aligned} \tag{1.71}$$

$$\begin{aligned} G_7 &= \frac{1}{\epsilon}\int_{\rho'}^{\tau_1} \int_{\rho'}^t \frac{\mathrm{d}t\mathrm{d}u}{tu} F\big((\theta_\epsilon - t - u)/\epsilon\big) \\ &\quad - \int_{\rho'}^{\tau_1} \int_{\rho'}^t \int_\epsilon^u \frac{\mathrm{d}t\mathrm{d}u\mathrm{d}v}{tuv^2} f\big((\theta(t,u,v) - t - u - v)/v\big), \\ G_8 &= \frac{1}{\epsilon}\int_{\tau_1}^{1/5} \int_{\rho'}^{\tau_1} \frac{\mathrm{d}t\mathrm{d}u}{tu} F\big((\theta_\epsilon - t - u)/\epsilon\big) \;+ \int_{1/5}^{\tau_2} \int_{\rho'}^{\tau_1} \frac{\mathrm{d}t\mathrm{d}u}{tu^2} F\big((\theta(t) - t - u)/u\big) \\ &\quad - \int_{\tau_1}^{1/5} \int_{\rho'}^{\tau_1} \int_\epsilon^u \frac{\mathrm{d}t\mathrm{d}u\mathrm{d}v}{tuv^2} f\big((\theta(t,u,v) - t - u - v)/v\big). \end{aligned}$$

Recall the sieve functions $F, f$ satisfy $F(s) = 2e^\gamma/s$ for $s \in [1,3]$, $f(s) = 2e^\gamma \log(s-1)/s$ for $s \in [2,4]$ and $F(s) = 2e^\gamma/s \cdot [1 + \int_2^{s-1} f(t)\mathrm{d}t]$ for all $s \geq 1$.

The main bound is the following.

**Proposition 1.6.3.** *Let $0 < \epsilon \leq 0.1 \leq \rho' \leq \tau_1 < 0.2 \leq \tau_2 < \tau_3 \leq \rho \leq 0.3$. and $\theta_\epsilon = \frac{3-\epsilon}{5}$. Then for $I_n$, $G_n$, and $G(c)$ as in* (1.66),(1.67),(1.71)*, and* (1.68)*, we have*

$$S(\mathcal{A}, x^\rho) \lesssim \frac{\Pi(x)}{5e^\gamma}\Big(\sum_{n=1}^{8} G_n + G(\tfrac{1}{5})\sum_{n=9}^{21} I_n\Big). \tag{1.72}$$

*Proof.* We first bound $S(\mathcal{A}, x^c)$ for $c \in [\epsilon, 1/5]$. By the Buchstab identity,

$$S(\mathcal{A}, x^c) = S(\mathcal{A}, x^\epsilon) - \sum_{x^\epsilon \leq p < x^c} S(\mathcal{A}_p, p).$$

Iterating twice more, we obtain

$$\begin{aligned} S(\mathcal{A}, x^c) &= S(\mathcal{A}, x^\epsilon) - \sum_{x^\epsilon \leq p_1 < x^c} S(\mathcal{A}_{p_1}, x^\epsilon) \\ &\quad + \sum_{x^\epsilon \leq p_2 < p_1 < x^c} S(\mathcal{A}_{p_1p_2}, x^\epsilon) - \sum_{x^\epsilon \leq p_3 < p_2 < p_1 < x^c} S(\mathcal{A}_{p_1p_2p_3}, p_3). \end{aligned} \tag{1.73}$$

To each term $S(\mathcal{A}_d, x^\epsilon)$ above, we apply the linear sieve of level $x^{\theta_\epsilon}$ for $\theta_\epsilon = \frac{3-\epsilon}{5}$, as in Theorem 1.2.8. And to each term $S(\mathcal{A}_{p_1p_2p_3}, p_3)$, we apply the linear sieve of level $x^\theta$ for $\theta = \theta(t_1, t_2, t_3)$, where $p_i = x^{t_i}$.

To handle the corresponding error terms, note for primes $x^\epsilon \leq p_2 < p_1 < x^c \leq x^{1/5}$ and $d \in \{1, p_1, p_1p_2\}$, the prime factors of $q/d$ above are bounded by $x^\epsilon$ so that the sets $\mathcal{A}_q$ are equidistributed to level $\theta_\epsilon$. Hence for each $x^\epsilon \leq p_2 < p_1 < x^{1/5}$, by Proposition 1.5.4 with $u = \epsilon$, $b = 1$, $p_1$, $p_1p_2$, we have

$$\sum_{\substack{q \leq x^{\theta_\epsilon} \\ q | P(x^\epsilon)}} \widetilde{\lambda}^+(q)\Big(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\phi(q)}\Big) = \sum_{\substack{q \leq x^{\theta_\epsilon} \\ q | P(x^\epsilon)}} \widetilde{\lambda}^+(q)\Big(\pi(x; q, -2) - \frac{\pi(x)}{\phi(q)}\Big) \ll_A \frac{x}{(\log x)^A},$$

and

$$\sum_{p_1} \sum_{\substack{q = p_1 m \leq x^{\theta_\epsilon} \\ m | P(x^\epsilon)}} \widetilde{\lambda}^-(q)\Big(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\phi(q)}\Big) \ll_A \frac{x}{(\log x)^A},$$

and

$$\sum_{p_2, p_1} \sum_{\substack{q = p_1p_2m \leq x^{\theta_\epsilon} \\ m | P(x^\epsilon)}} \widetilde{\lambda}^+(q)\Big(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\phi(q)}\Big) \ll_A \frac{x}{(\log x)^A}.$$

In addition for each $p_3 < p_2 < p_1 < x^{1/5}$, $p_i = x^{t_i}$, letting $\theta = \theta(t_1, t_2, t_3)$, by Proposition 1.5.4 with $b = p_1p_2p_3$,

$$\sum_{p_3,p_2,p_1} \sum_{\substack{q=p_1p_2p_3m\leq x^\theta \\ m|P(p_3)}} \widetilde{\lambda}^-(q)\Big(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\phi(q)}\Big) \ll_A \frac{x}{(\log x)^A}.$$

Thus for $c \leq \frac{1}{5}$, the linear sieve bounds give

$$S(\mathcal{A}, x^\epsilon) \lesssim |\mathcal{A}|V(x^\epsilon)F\big(\tfrac{\theta_\epsilon}{\epsilon}\big), \tag{1.74}$$

and

$$\sum_{x^\epsilon \leq p_1 < x^c} S(\mathcal{A}_{p_1}, x^\epsilon) \gtrsim \sum_{x^\epsilon \leq p_1 < x^c} |\mathcal{A}|g(p_1)V(x^\epsilon)f\big(\tfrac{\theta_\epsilon - t_1}{\epsilon}\big), \tag{1.75}$$

and

$$\sum_{x^\epsilon \leq p_2 < p_1 < x^c} S(\mathcal{A}_{p_1p_2}, x^\epsilon) \lesssim \sum_{x^\epsilon \leq p_2 < p_1 < x^c} |\mathcal{A}|g(p_1p_2)V(x^\epsilon)F\big(\tfrac{\theta_\epsilon - t_1 - t_2}{\epsilon}\big), \tag{1.76}$$

and

$$\sum_{x^\epsilon \leq p_3 < p_2 < p_1 < x^c} S(\mathcal{A}_{p_1p_2p_3}, p_3) \gtrsim \sum_{x^\epsilon \leq p_3 < p_2 < p_1 < x^c} |\mathcal{A}|g(p_1p_2p_3)V(p_3)f\big(\tfrac{\theta - t_1 - t_2 - t_3}{t_3}\big). \tag{1.77}$$

Hence by (1.74), (1.75), (1.76), (1.77), we observe that (1.73) becomes

$$\begin{aligned} S(\mathcal{A}, x^c) \lesssim & -|\mathcal{A}| \sum_{x^\epsilon \leq p_3 < p_2 < p_1 < x^c} g(p_1p_2p_3)V(p_3)f\big(\tfrac{\theta - t_1 - t_2 - t_3}{\epsilon}\big) \qquad (1.78) \\ & + |\mathcal{A}|V(x^\epsilon)\bigg(F\big(\tfrac{\theta_\epsilon}{\epsilon}\big) - \sum_{x^\epsilon \leq p_1 < x^c} g(p_1)f\big(\tfrac{\theta_\epsilon - t_1}{\epsilon}\big) + \sum_{x^\epsilon \leq p_2 < p_1 < x^c} g(p_1p_2)F\big(\tfrac{\theta_\epsilon - t_1 - t_2}{\epsilon}\big)\bigg). \end{aligned}$$

Recall $V(z) \sim \mathfrak{S}_2/e^\gamma \log z$ by Mertens theorem. Thus by partial summation and the prime number theorem, we obtain

$$S(\mathcal{A}, x^c) \lesssim \frac{\Pi(x)}{e^\gamma}G(c), \tag{1.79}$$

for $G(c)$ as in (1.68). Hence for $c = \rho', \tau_1$ we have $c \in [\epsilon, 1/5]$, so we bound $\Gamma_1$ as

$$\Gamma_1 = 4S(\mathcal{A}, x^{\rho'}) + S(\mathcal{A}, x^{\tau_1}) \lesssim \frac{\Pi(x)}{e^\gamma}\Big(4G(\rho') + G(\tau_1)\Big) = \frac{\Pi(x)}{e^\gamma}G_1.$$

Now consider $c, c' \in [1/5, 2/7]$. We shall apply the linear sieve of level $\theta(t_1)$, as in (1.25). In general, for $p_1 = x^{t_1}$ and $\theta(t_1)$ as in (1.25), Proposition 1.5.4 gives

$$\sum_{p_1} \sum_{\substack{p_1|q,\ q\le x^{\theta(t_1)} \\ q/p_1|P(p_1)}} \widetilde{\lambda}^-(q)\left(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\phi(q)}\right) \ll_A \frac{x}{(\log x)^A},$$

so that for $c, c' \in [1/5, 2/7]$, the linear sieve of level $\theta(t_1)$ gives

$$\sum_{x^{c'}\le p_1<x^c} S(\mathcal{A}_{p_1}, p_1) \gtrsim \sum_{x^{c'}\le p_1<x^c} |\mathcal{A}| g(p_1) V(p_1) f\big(\tfrac{\theta(t_1)-t_1}{t_1}\big). \tag{1.80}$$

Thus by partial summation and the prime number theorem,

$$\Gamma_2 = \sum_{x^{\rho'}\le p<x^{\rho}} S(\mathcal{A}_p, x^{\rho'}) \lesssim \frac{\Pi(x)}{e^\gamma}\big(G + \overline{G}(\rho)\big) = \frac{\Pi(x)}{e^\gamma} G_2.$$

Similarly, we obtain

$$\Gamma_n \lesssim \frac{\Pi(x)}{e^\gamma} G_n \qquad \text{for } 1 \le n \le 8. \tag{1.81}$$

Finally, for the remaining $\Gamma_n$, we apply the switching principle. Namely, for $\Gamma_9$ we have

$$\Gamma_9 = \sum\sum\sum_{x^{\tau_1}\le p_1<p_2<p_3<x^{\tau_3}} S(\mathcal{A}_{p_1p_2p_3}, p_2) = S(\mathcal{B}, x^{1/2}) + O(x^{1/2}) \tag{1.82}$$

for the set

$$\begin{aligned}\mathcal{B} = \{p_1p_2p_3m - 2 \le x : {} & x^{\tau_1} \le p_1 < p_2 < p_3 < x^{\tau_3}, \\ & p' \mid m \Rightarrow p' \ge p_2 \text{ or } p' = p_1\}.\end{aligned} \tag{1.83}$$

Note since $p_2, p_1 > x^{\tau_1} > x^{\cdot 1}$, each $m$ above has at most 7 prime factors.

Now by a standard subdivision argument, $\mathcal{B}$ is similarly equidstributed in arithmetic progressions as is $\mathcal{A}$. Indeed, the basic idea is to partition $\mathcal{B} = \bigcup_{r\le 7} \mathcal{B}^{(r)}$, where $\mathcal{B}^{(r)}$ is the subset corresponding to integers $m$ with $r$ prime factors. Then we cover the prime tuples $(p_1, \dots, p_r)$ into hypercubes of the form

$$\big[\Delta^{l_1}, \Delta^{l_1+1}\big) \times \cdots \times \big[\Delta^{l_r}, \Delta^{l_r+1}\big),$$

for $\Delta = 1 + (\log x)^{-B}$ with $B > 0$ sufficiently large, and apply Corollary 1.5.5.1 with $u = \epsilon$, $b = 1$. This gives

$$\sum_{\substack{q\le x^{\theta_\epsilon} \\ q|P(x^\epsilon)}} \widetilde{\lambda}^+(q)\left(|\mathcal{B}_q| - \frac{|\mathcal{B}|}{\phi(q)}\right) \ll_A \frac{x}{(\log x)^A}.$$

Similarly for each $x^{\varepsilon} < p_3' < p_2' < p_1' < x^{1/5}$, by Corollary 1.5.5.1 with $u = \epsilon$ and $b = p_1'$, $p_1'p_2'$, we have

$$\sum_{p_1'} \sum_{\substack{q=p_1'm' \le x^{\theta_\epsilon} \\ m'|P(x^\epsilon)}} \widetilde{\lambda}^-(q) \Big( |\mathcal{B}_q| - \frac{|\mathcal{B}|}{\phi(q)} \Big) \ll_A \frac{x}{(\log x)^A},$$

and

$$\sum_{p_2',p_1'} \sum_{\substack{q=p_1'p_2'm' \le x^{\theta_\epsilon} \\ m'|P(x^\epsilon)}} \widetilde{\lambda}^+(q) \Big( |\mathcal{B}_q| - \frac{|\mathcal{B}|}{\phi(q)} \Big) \ll_A \frac{x}{(\log x)^A}.$$

In addition for each $p_3' < p_2' < p_1' < x^{1/5}$, $p_i' = x^{t_i}$, letting $\theta = \theta(t_1, t_2, t_3)$, by Corollary 1.5.5.1 with $b = p_1'p_2'p_3'$, we have

$$\sum_{p_3',p_2',p_1'} \sum_{\substack{q=p_1'p_2'p_3'm' \le x^{\theta} \\ m'|P(p_3')}} \widetilde{\lambda}^-(q) \Big( |\mathcal{B}_q| - \frac{|\mathcal{B}|}{\phi(q)} \Big) \ll_A \frac{x}{(\log x)^A}.$$

Iterating the Buchstab identity, we have

$$\begin{aligned} S(\mathcal{B}, x^{1/2}) \le S(\mathcal{B}, x^{1/5}) \ &= \ S(\mathcal{B}, x^\epsilon) - \sum_{x^\epsilon \le p_1' < x^{1/5}} S(\mathcal{B}_{p_1'}, x^\epsilon) \\ &\quad + \sum_{x^\epsilon \le p_2' < p_1' < x^{1/5}} S(\mathcal{B}_{p_1'p_2'}, x^\epsilon) - \sum_{x^\epsilon \le p_3' < p_2' < p_1' < x^{1/5}} S(\mathcal{B}_{p_1'p_2'p_3'}, p_3'), \end{aligned}$$

and hence by the linear sieve bounds we obtain

$$\Gamma_9 \lesssim S(\mathcal{B}, x^{1/2}) \le S(\mathcal{B}, x^{1/5}) \ \lesssim \ e^{-\gamma} \frac{G(\frac{1}{5})}{\log x} \mathfrak{S}_2 |\mathcal{B}|. \tag{1.84}$$

Now to compute $|\mathcal{B}|$, we may drop the condition "or $p' = p_1$" in (1.83) at admissible cost, so that by Lemma 1.6.1 we have

$$|\mathcal{B}| \ \sim \ \frac{x}{\log x} \int_{\tau_1 < t_1 < t_2 < t_3 < \tau_2} \frac{\mathrm{d}t_1 \mathrm{d}t_2 \mathrm{d}t_3}{t_1 t_2^2 t_3} \omega\Big( \frac{1 - t_1 - t_2 - t_3}{t_2} \Big) = \frac{x}{\log x} \cdot I_9.$$

Thus we have $\Gamma_9 \lesssim e^{-\gamma} G(\frac{1}{5}) \Pi(x) I_9$. Similarly, we obtain

$$\Gamma_n \ \lesssim \ G(\tfrac{1}{5}) \frac{\Pi(x)}{e^\gamma} \, I_n \qquad \text{(for } 9 \le n \le 21\text{)}. \tag{1.85}$$

Hence plugging (1.81), (1.85) into Lemma 1.6.2 completes the proof. □

Let

$$\begin{aligned} \rho &= 0.27195, & \tau_3 &= 0.24589, \\ \rho' &= 0.12313, & \tau_2 &= 0.20867, \\ \epsilon &= 0.002, & \tau_1 &= 0.16288. \end{aligned} \tag{1.86}$$

For such choices of parameters, we compute the following integrals from Proposition 1.6.3,

$$\sum_{9\le n\le 21} I_n \ \le\ 0.174404, \qquad \sum_{1\le n\le 8} G_n \ \le\ 28.34581, \qquad G(\tfrac{1}{5}) \ \le\ 5.99237.$$

Thus by Proposition 1.6.3, we obtain the bound

$$\pi_2(x) \ \lesssim\ S(\mathcal{A}, x^\rho) \ \lesssim\ \frac{\Pi(x)}{5e^\gamma}\Big(\sum_{n=1}^{8} G_n \ +\ G(\tfrac{1}{5})\sum_{n=9}^{21} I_n\Big) \ \lesssim\ 3.30042\,\Pi(x). \tag{1.87}$$

We also record the individual bounds (see table at end of chapter).

### 1.6.3 Completing the proof of Theorem 1.1.2

We shall refine our argument in certain cases for which Wu's iteration method [141] applies directly without modification. As such we have chosen simplicity over full optimization.

In the lemma below, we consider the cases of sets $\mathcal{A}_{p_1p_2}$, where $p_1, p_2$ lie in a prescribed range, and such that for all multiples $bp_1p_2$, the sets $\mathcal{A}_{bp_1p_2}$ are equidistributed to level $x^\theta$ (we also need level $x^\theta$ for corresponding switched sets $\mathcal{B}$ of integers $mbp_1p_2 - 2$).

**Lemma 1.6.4.** *Let $\theta \in [1/2, 1]$, $s \in [2, 3]$, and $\mathcal{A} = \{p + 2 : p \le x\}$. There is a function $H_\theta(s)$, monotonically increasing in $\theta$ for fixed $s$ and decreasing in $s$ for fixed $\theta$, such that the following holds: For each $(D_1, D_2) \in \mathbf{D}_2^{\mathrm{well}}(x^\theta)$, we have*

$$\sum_{\substack{D_1<p_1<D_1^{1+\tau}\\ D_2<p_2<D_2^{1+\tau}}} S(\mathcal{A}_{p_1p_2}, z) \ \lesssim\ \frac{\Pi(x)}{e^\gamma} \sum_{\substack{D_1<p_1<D_1^{1+\tau}\\ D_2<p_2<D_2^{1+\tau}}} \frac{\log x}{\phi(p_1p_2)\log z}\big(F(s) - \frac{2e^\gamma}{s}H_\theta(s)\big), \tag{1.88}$$

*where $z^s = x^\theta/p_1p_2$, provided (1.63) holds for $\lambda = \widetilde{\lambda}^+$ at level $D = x^\theta$ with $(x^{\varepsilon_1}, x^{\varepsilon_2}) = (D_1, D_2)$, and provided for all vectors $(D_1, \dots, D_r) \in \mathbf{D}_r^{\mathrm{well}}(x^\theta)$ extending $(D_1, D_2)$,*

$$\sum_{\substack{b=p_1\cdots p_r\\ D_i<p_i\le D_i^{1+\tau}}} \sum_{\substack{b\mid d,\ d\le x^\theta\\ (d,a)=1}} \widetilde{\lambda}^{\pm}(d)\left(\pi(x; d, a) - \frac{\pi(x)}{\phi(d)}\right) \ \ll_{a,\varepsilon,A}\ \frac{x}{(\log x)^A}.$$

*Proof.* Wu iterates the weighted sieve inequality (Lemmas 4.1 and 4.2 [141]), on the subset of (non-switched) terms whose sieving parameter $s$ lies in the interval $s \in [2,3]$. This yields a recurrence relation for a function $H_\theta(s)$, which encodes the percent savings over the (normalized) linear sieve $sF(s)/(2e^\gamma)$.

Starting from a term $S(\mathcal{A}_{p_1p_2}, z)$, each successive iteration of the weighted sieve inequality is composed of terms of the form $S(\mathcal{A}_{bp_1p_2}, z')$ for some multiple $bp_1p_2$ corresponding to some vector extending $(D_1, D_2)$. By assumption, all such sets are equidistributed to level $x^\theta$, when weighted by the upper/lower linear sieve. Similarly, the switched sets are also equidistributed to level $x^\theta$ (here we only need the upper bound weights $\widetilde{\lambda}^+$). Finally, the savings function $H_\theta(s)$ inherits the stated monotonicity properties by construction of the iteration. □

The function $H_\theta$ depends on the known level of distribution $x^\theta$ (i.e. Wu used $\theta = \frac{1}{2}$ for Goldbach, and $\theta = \frac{4}{7}$ for twin primes). For parameters as in Tables 1 and 2 [141, pp.30–32],

$$H_{1/2}(t) \geq \begin{cases} 0.0223939 & \text{if } 2.0 < t \leq 2.2, \\ 0.0217196 & \text{if } 2.2 < t \leq 2.3, \\ 0.0202876 & \text{if } 2.3 < t \leq 2.4, \\ 0.0181433 & \text{if } 2.4 < t \leq 2.5, \\ 0.0158644 & \text{if } 2.5 < t \leq 2.6, \\ 0.0129923 & \text{if } 2.6 < t \leq 2.7, \\ 0.0100686 & \text{if } 2.7 < t \leq 2.8, \\ 0.0078162 & \text{if } 2.8 < t \leq 2.9, \\ 0.0072943 & \text{if } 2.9 < t \leq 3.0, \\ 0 & \text{else,} \end{cases}$$

and

$$H_{4/7}(t) \geq \begin{cases} 0.0287118 & \text{if } 2.0 \leq t \leq 2.1, \\ 0.0280509 & \text{if } 2.1 < t \leq 2.2, \\ 0.0264697 & \text{if } 2.2 < t \leq 2.3, \\ 0.0241936 & \text{if } 2.3 < t \leq 2.4, \\ 0.0214619 & \text{if } 2.4 < t \leq 2.5, \\ 0.0183875 & \text{if } 2.5 < t \leq 2.6, \\ 0.0149960 & \text{if } 2.6 < t \leq 2.7, \\ 0.0117724 & \text{if } 2.7 < t \leq 2.8, \\ 0.0094724 & \text{if } 2.8 < t \leq 2.9, \\ 0.0090024 & \text{if } 2.9 < t \leq 3.0, \\ 0 & \text{else.} \end{cases}$$

As such, in [141, Theorem 3] Wu obtained $\pi_2(x)/\Pi(x) \lesssim (7/2)(1 - H_{4/7}(2.1)) \leq 3.39951$.

To complete our proof of Theorem 1.1.2 we apply Lemma 1.6.4, now valid up to level $x^{\frac{7}{12}}$ by Corollary 1.2.6.1 and Proposition 1.5.5 for $\widetilde{\lambda}^+$. Note when the largest integration variable is $t \geq \frac{1}{5}$, we have $\theta(t) = \frac{2-t}{3} \leq \frac{7}{12}$ iff $t \geq \frac{1}{4}$. A key feature we use to satisfy the conditions of Lemma 1.6.4 is that the level $x^{\theta(t_1)}$ persists, since the largest prime $p_1 = x^{t_1}$ is preserved through successive iterations.

Thus in practice, Lemma 1.6.4 simply amounts to modifying the integral in $G_2$ by substituting $F(s) - \frac{2e^\gamma}{s}H_\theta(s)$ in for $F(s)$, $s = (\theta(t) - t - u)/u$, when $t \geq \frac{1}{4}$ (the only parameter $\geq \frac{1}{4}$ is $\rho$, so we only refine $G_2$). Denote this as $G_2^{\mathrm{Wu}}$. For ease we also use $H_\theta(s) \geq H_{4/7}(s)$, by monotonicity in $\theta$. Doing so, with the same parameter choices (1.86), we obtain $G_2^{\mathrm{Wu}} \leq -5.598667$ and hence

$$\pi_2(x) \lesssim \frac{\Pi(x)}{5e^\gamma}\Big(G_2^{\mathrm{Wu}} + \sum_{1\leq n\neq 2\leq 8} G_n + G(\tfrac{1}{5}) \sum_{9\leq n\leq 21} I_n\Big) \lesssim 3.299552\,\Pi(x). \quad (1.89)$$

This completes the proof of Theorem 1.1.2.

For slight numerical gains, one may compute $H_\theta(s)$ when $\theta \in [\frac{4}{7}, \frac{7}{12}]$, by tweaking the formulae in [141]. More substantially, Wu defined a lower bound savings $h_\theta(s)$, for a substitution of $f(s)$ by $f(s) + \frac{2e^\gamma}{s}h_\theta(s)$. But in practice, to compute $h$ would require derivations (analogous to $H$) of as yet undetermined formulae. We leave these to the reader.

| $n$ | $G_n$ | $n$ | $I_n$ | $n$ | $I_n$ |
|---|---|---|---|---|---|
| 1 | 39.00163 | 9 | 0.0332157 | 17 | 0.000315 |
| 2 | −5.591009 | 10 | 0.0228322 | 18 | 0.000269 |
| 3 | −3.986553 | 11 | 0.0092564 | 19 | 0.000164 |
| 4 | −5.060499 | 12 | 0.0150101 | 20 | $\leq 2.70 \cdot 10^{-6}$ |
| 5 | 1.864133 | 13 | 0.0547244 | 21 | $\leq 5.50 \cdot 10^{-9}$ |
| 6 | 0.741181 | 14 | 0.0260202 | | |
| 7 | 0.453663 | 15 | 0.0124636 | | |
| 8 | 0.923736 | 16 | 0.0001314 | | |

# Chapter 2

# Primes in arithmetic progressions to large moduli, and shifted primes without large prime factors

## 2.1 Introduction

Following an old conjecture of Erdős [41, 44], we expect there are infinitely many primes $p$ with $P^+(p-a) \leq p^{\varepsilon}$, for any $\varepsilon > 0$. We prove the infinitude of primes $p$ with $P^+(p-a) \leq p^{0.2844}$, from the following quantitative result.

**Theorem 2.1.1.** *For fixed nonzero $a \in \mathbb{Z}$ and $\beta > 15/32\sqrt{e} = 0.2843\cdots$, there exists $C \geq 1$ such that*

$$\sum_{\substack{x<p\leq 2x \\ P^+(p-a)\leq x^{\beta}}} 1 \gg \frac{x}{(\log x)^C}. \tag{2.1}$$

The exponent 0.2844 in Theorem 2.1.1 gives a roughly 4% refinement over the previous record exponent 0.2961 of Baker and Harman [4]. The table below gives a chronology of the known lower bounds on $\beta$.

| Year | Author(s) | $\beta >$ |
|---|---|---|
| 1998 | Baker–Harman [4] | 0.2961 |
| 1989 | Friedlander [61] | $0.3032\cdots = 1/2\sqrt{e}$ |
| 1986 | Fouvry–Grupp [58] | $0.3174\cdots = 3/7e^{.3}$ |
| 1983 | Balog [5] | 0.35 |
| 1980 | Pomerance [116] | $0.4490\cdots = 625/512e$ |
| 1979 | Wooldridge [139] | $0.8284\cdots = 2(\sqrt{2}-1)$ |
| 1969 | Goldfeld [63] | $0.5 = 1/2$ |
| 1935 | Erdős [41] | $1-\delta$ for some $\delta > 0$ |

In recent decades, this problem has attracted increased attention, due in part to applications to cryptography (see [28, 67, 117] for further discussion). Moreover, this problem is of independent interest as it sheds light on the subtle interplay between addition and multiplication. Indeed, long before connections to cryptography, in 1935 Erdős [41] showed the existence of some $\delta > 0$ for which infinitely many primes $p$ satisfy $P^+(p-1) \leq p^{1-\delta}$.

Theorem 2.1.1 implies an improved lower bound on the distribution of Carmichael numbers. Recall a composite number $n$ is a Carmichael number if $n$ is a pseudoprime to every base $b$, that is, $b^{n-1} \equiv 1 \pmod{n}$ for all coprime $(b, n) = 1$.

**Corollary 2.1.1.1.** *There are at least $x^{0.3389}$ Carmichael numbers up to $x$, sufficiently large.*

The infinitude of Carmichael numbers we first proven in the landmark result of Alford–Granville–Pomerance [2]. Moreover, their argument gave the quantitative

lower bound $x^{\frac{5}{12}(1-\beta)}$, for $\beta > 0$ satisfying (2.1). The current record is $x^{0.4736(1-\beta)}$ due to Harman [74]. Corollary 2.1.1.1 then follows by combining this with Theorem 2.1.1. Note the exponent $0.4736(1 - 15/32\sqrt{e}) = 0.33895\cdots$ in Corollary 2.1.1.1 improves on $0.4736(1 - 0.2961) = 0.33336\cdots$ from Harman [74].

We also consider the values of the Euler $\phi$-function, about which there are many celebrated questions. For example, Carmchael's conjecture [18] asserts that every number $n \in \mathbb{N}$ that occurs $n = \phi(m)$ as a value of $\phi$, does so for at least two arguments $m_1, m_2 \in \mathbb{N}$; Sierpínski's conjecture (now Ford's theorem [55]) asserts that for each $k \geq 2$, there is a number $n \in \mathbb{N}$ occuring as a value $n = \phi(m_i)$ for exactly $k$ arguments $m_1, \dots, m_k \in \mathbb{N}$. Using Theorem 2.1.1, we obtain improved bounds on the infinitude of popular values of the Euler $\phi$-function.

**Corollary 2.1.1.2.** *Denote by $n_1 < n_2 < \cdots$ the integers $n \in \mathbb{N}$ for which $n = \phi(m)$ admits at least $n^{0.7156}$ solutions $m \in \mathbb{N}$. Then the sequence $(n_i)$ is infinite, and satisfies*

$$\lim_{i\to\infty} \frac{\log n_{i+1}}{\log n_i} = 1.$$

The well-established method of Erdős and Pomerance [44, 116], yields a lower bound $m^{1-\beta}$, for $\beta > 0$ satisfying (2.1). Thus Corollary 2.1.1.2 follows directly by Theorem 2.1.1. Note the exponent $1 - 15/32\sqrt{e} = 0.7156\cdots$ in Corollary 2.1.1.2 improves on $1 - 0.2961 = 0.7039$ from Harman [74].

Finally, note if Erdős' conjecture of $\beta = \varepsilon$ holds, then the equation $m = \phi(n)$ has at least $m^{1-\varepsilon}$ solutions. This bound would be sharp up to a factor $m^{\varepsilon}$, since $m = \phi(n) \gg n/\log\log n$ bounds any solution $n$ by $n \ll m \log\log m$.

### 2.1.1 Primes in arithmetic progressions to large moduli

As the main technical result of the chapter, we establish a new estimate for primes in arithmetic progressions with quadrilinear forms of moduli. Let $\tau(q)$ denote the divisor function, and $\pi(x; q, a)$ the count of primes up to $x$ congruent to $a \pmod q$.

**Theorem 2.1.2.** *Fix nonzero $a \in \mathbb{Z}$. Let $\varepsilon > 0$ and let $Q, R, S$ satisfy*

$$QR \leq x^{1/2+\varepsilon}, \qquad QS^2 \leq x^{1/2-2\varepsilon} \qquad S^2 \leq R \leq x^{1/32-\varepsilon}. \tag{2.2}$$

*Let $\lambda_q, \nu_q, \eta_q, \mu_q$ be complex sequences with $|\lambda_q|, |\nu_q|, |\eta_q|, |\mu_q| \leq \tau(q)^{B_0}$. Then for every $A > 0$ we have*

$$\sum_{\substack{q\leq Q}}\sum_{r\leq R}\sum_{s\leq S}\sum_{t\leq S}\lambda_q\nu_r\eta_s\mu_t\Big(\pi(x; qrst, a) - \frac{\pi(x)}{\phi(qrst)}\Big) \ll_{a,\varepsilon,A,B_0} \frac{x}{(\log x)^A}.$$
$$(qrst,a)=1$$

Theorem 2.1.2 may handle quadrilinear forms of moduli up to $x^{17/32-\varepsilon}$, for the choices $(Q, R, S) = (x^{15/32+2\varepsilon}, x^{1/32-\varepsilon}, x^{1/64-2\varepsilon})$. Specifically, for the application to our main Theorem 2.1.1 we use

$$\sum_{q\leq x^{15/32+2\varepsilon}} \sum_{\substack{r\leq x^{1/32-\varepsilon} \\ (qrst,a)=1}} \sum_{s,\, t\leq x^{1/64-2\varepsilon}} \lambda_q \nu_r \eta_s \mu_t \Big(\pi(x; qrst, a) - \frac{\pi(x)}{\phi(qrst)}\Big) \ll_{a,\varepsilon,A,B_0} \frac{x}{(\log x)^A}. \tag{2.3}$$

Previously, the highest level for these applications had involved bilinear forms. In their celebrated 1986 work, Bombieri, Friedlander, and Iwaniec [13, Theorem 8] handled bilinear forms of moduli up to level $x^{29/56-\varepsilon}$, showing

$$\sum_{\substack{q\leq Q \\ (qr,a)=1}} \sum_{r\leq R} \lambda_q \nu_r \Big(\pi(x; qr, a) - \frac{\pi(x)}{\phi(qr)}\Big) \ll_{a,\varepsilon,A} \frac{x}{(\log x)^A} \tag{2.4}$$

with $Q < x^{1/3}$, $R < x^{1/5}$, $Q^5R^2 < x^2$, and $QR < x^{29/56-\varepsilon}$. This bilinear estimate was only recently extended to moduli up to level $x^{11/21-\varepsilon}$ by Maynard [101, Theorem 1.1], deducing (2.4) with $(Q, R) = (x^{10/21}, x^{1/21-\varepsilon})$ (even with weights replaced by absolute values) Note $29/56 = 0.5178\cdots$, $11/21 = 0.5238\cdots$, and $17/32 = 0.5313\cdots$.

In context, Theorem 2.1.2 may be viewed as an interpolating result between [101, Theorem 1.1] and [102, Theorem 1.1]. For example, at a given level of distribution, absolute values are the most 'rigid' weights, since an equidistribution estimate using absolute values implies a corresponding estimate using any other choice of weights. Conversely, triply well-factorable weights should be viewed as very 'flexible', and our choice of quadrilinear weights are of an intermediate nature. As a trade-off, the more rigid the weights the stronger the estimate is at a given level, but the more flexible the weights the higher one can usually take the level of distribution.

As such, the size of moduli $x^{17/32-\epsilon}$ in Theorem 2.1.2 exceeds $x^{11/21-\epsilon}$ in [101, Theorem 1.1], while not to the full extent of $x^{3/5-\epsilon}$ in [102, Theorem 1.1]. The balance of 'rigidity' and 'flexibitiy' in the weights in Theorem 2.1.2 is molded to our application to Theorem 2.1.1. Whereas, the shifted primes problem appears too rigid for [102, Theorem 1.1] to be applicable.

## 2.2 Proof outline

In this section we outline the proofs of Theorems 2.1.1 and 2.1.2.

Following the previous method of Baker–Harman [4], we restrict our attention to special factorizations $p - a = lp_0mn$ with $l \sim x^{2\theta-1}$ smooth, $p_0$ prime, and $p_0m$, $n \sim$

$x^{1-\theta}$. We aim to take $\theta > 1/2$ as large as possible. An inclusion-exclusion argument then reduces the problem to estimating primes in arithmetic progressions of the form

$$\sum_l \sum_{p_0} \sum_m \Big(\pi(x; a, lp_0m) - \frac{\pi(x)}{\phi(lp_0m)}\Big). \tag{2.5}$$

The desired bound $x/(\log x)^A$ for (2.5) above follows from [101, Theorem 1.1] when $\theta = 11/21$, by grouping $p_0$,$m$ together and inserting absolute values (this would already give some improvement over Baker–Harman). We can hope to do better by exploiting the fact we don't need absolute values. That is, the argument for [101, Theorem 1.1] was limited by $\theta \leq 11/21$ only for very specific terms in the decomposition of primes. Namely, when one factor in the decomposition is of size $x^{1/7}$, the 'Fouvry-style' estimate [101, Proposition 12.1] break down. When the coefficients factor as $c_{q,r} = \lambda_q \nu_r$, we may use a stronger result of Iwaniec–Pomykala [81] in this critical case.

In addition, we strengthen the 'Zhang-style' estimate [101, Proposition 8.2] in the case of trilinear forms of moduli. Lastly, to combine these 'Fouvry-style' and 'Zhang-style' trilinear estimates in a compatible manner, we take a common refinement of the corresponding systems of conditions. This leads to the single system (2.2) for the final quadrilinear estimate. When combined with other estimates of Maynard [101], this establishes the main Theorem 2.1.2 with $\theta = 17/32$.

Roughly speaking, the Zhang-style argument handles sums of the form

$$\sum_q \sum_r c_{q,r} \sum_n \alpha_n \sum_m \beta_m \Big(\mathbf{1}_{nm \equiv a(qr)} - \frac{1}{\phi(qr)}\Big).$$

We apply Cauchy-Schwarz in the $q, m$ variables (smoothing the $m$-summation), use Poisson summation in the $m$ variable, and simplify the exponential sums to roughly give

$$\sum_q \sum_{r_1, r_2} c_{q,r_1} c_{q,r_2} \sum_{n_1, n_2} \alpha_{n_1} \alpha_{n_2} \sum_{h < QR^2N/x} e\Big(ah\Big(\frac{\overline{n_1 r_2}}{qr_1} + \frac{\overline{n_2 r_1 q}}{r_2}\Big)\Big).$$

Then we want to apply Cauchy-Schwarz in the $q, n_1, n_2$ variables to eliminate the unknown coefficients $\alpha_n$. Then applying Poisson summation in the $n_1, n_2$ variables (inserting a smooth majorant) leads to sums of Kloosterman sums, which may be handled by the Weil bound.

In our case of trilinear forms of moduli, we may restrict to those $r$ which factor and consider the coefficients $c_{q,rs} = \lambda_q \nu_r \eta_s$. In this situation, we may also apply Cauchy-Schwarz in in the $s_1, s_2$ variables. This new situation has the significant

benefit of fewer off-diagonal terms, and the remaining diagonal terms have less sparse summations in auxiliary variables. This allows for improved estimates in this critical case, and leads to $\theta = 17/32$ when combined with the other estimates.

*Remark* 2.2.1. The previous argument of Baker–Harman [4] proceeds by estimating primes in arithmetic progressions of the form

$$\sum_{l} \sum_{p_0} \sum_{m} \pi(x; a, lp_0m) \tag{2.6}$$

in two regimes, using the following ingredients:

(a) a mean value theorem of Bombieri–Friedlander–Iwaniec [13, Theorem 9] to get asymptotics for (2.6) in the range $x^\beta < p_0 < x^{1/3}$.

(b) a Harman's sieve argument to get weaker bounds for (2.6) in the range $x^{1/3} \leq p_0 \leq 2x^{1-\theta}$, by focusing on the most amenable parts of the decomposition of the primes.

Now using Theorem 2.1.2 we upgrade (a) above, which already achieves a superior result for shifted primes without appealing to part (b). In principle, one could obtain some further improvements using a similar argument as in (b). In the interest of clarity of presentation we do not pursue this.

## 2.3 Shifted primes without large prime factors

In this section we deduce Theorem 2.1.1 from Theorem 2.1.2.

Let $\theta = 17/32 - 4\varepsilon$ and take $\beta > (1-\theta)/\sqrt{e}$. Define the integer $H = \lceil (2\theta - 1)\frac{10}{\varepsilon} \rceil$. In particular $l_0 := x^{\frac{2\theta-1}{H}}$ lies in the interval $l_0 \in [x^{.09\varepsilon}, x^{.1\varepsilon}]$ for $\varepsilon > 0$ sufficiently small. Define the set

$$\mathcal{G} = \{l_1 \cdots l_H : l_i \sim l_0, (l_i, a) = 1\},$$

so that $l \asymp x^{2\theta-1} = x^{1/16-8\varepsilon}$ for all $l \in \mathcal{G}$. Consider

$$\mathcal{N} = \big\{(p, l, m, n) \ : \ p - a = lmn, \ l \in \mathcal{G}, \ m, n \sim x^{1-\theta}, \ (m, a) = 1, \ p \sim x\big\}, \tag{2.7}$$

$$\mathcal{N}' = \big\{(p, l, m, n) \in \mathcal{N} \ : \ P^+(p - a) \leq x^\beta\big\}. \tag{2.8}$$

Observe that for Theorem 2.1.1, it suffices to prove

$$|\mathcal{N}'| \gg \frac{x}{\log x}. \tag{2.9}$$

Indeed, letting $\Gamma(p) = |\{(l,m,n) \,:\, (p,l,m,n) \in \mathcal{N}'\}|$, we have $|\mathcal{N}'| = \sum_{p\sim x} \Gamma(p)$. By Cauchy-Schwarz,

$$|\mathcal{N}'|^2 = \Big(\sum_{p\sim x} \Gamma(p)\Big)^2 \le \sum_{p\sim x} \Gamma(p)^2 \cdot \sum_{\substack{p\sim x\\ \Gamma(p)>0}} 1.$$

Also $\Gamma(p) \le \tau(p-a)^{B_0}$ so by a divisor bound $\sum_{p\sim x} \Gamma(p)^2 \ll |\mathcal{N}'|(\log x)^B$, and thus

$$\sum_{\substack{p\sim x\\ P^+(p-a)\le x^\beta}} 1 \ge \sum_{\substack{p\sim x\\ \Gamma(p)>0}} 1 \gg \frac{|\mathcal{N}'|}{(\log x)^B} \gg \frac{x}{(\log x)^{B+1}}$$

using (2.9). This completes the proof of Theorem 2.1.1. Thus it suffices to show (2.9).

Next define

$$\mathcal{N}_1 = \{(p,l,p_0m,n) \in \mathcal{N} \,:\, x^\beta < p_0 \le 2x^{1-\theta}\}$$
$$\mathcal{N}_2 = \{(p,l,m,p_0n) \in \mathcal{N} \,:\, x^\beta < p_0 \le 2x^{1-\theta}\}$$

and note by symmetry $|\mathcal{N}_1| = |\mathcal{N}_2|$. We have

$$|\mathcal{N}'| \ge |\mathcal{N}| - |\mathcal{N}_1| - |\mathcal{N}_2| = |\mathcal{N}| - 2|\mathcal{N}_1|. \tag{2.10}$$

Observe

$$|\mathcal{N}_1| = \sum_{l\in\mathcal{G}} \sum_{\substack{m\sim x^{1-\theta}/p_0\\ (m,a)=1}} \sum_{x^\beta<p_0\le 2x^{1-\theta}} \sum_{\substack{p\sim x\\ p\equiv a \pmod{lp_0m}}} 1. \tag{2.11}$$

We shall use Theorem 2.1.2 to prove that

$$|\mathcal{N}_1| = (1+O(\varepsilon))N_1, \tag{2.12}$$

where

$$N_1 = \sum_{l\in\mathcal{G}} \sum_{\substack{m\sim x^{1-\theta}/p_0\\ (m,a)=1}} \sum_{x^\beta<p_0\le 2x^{1-\theta}} \sum_{\substack{p\sim x\\ (p,a)=1}} \frac{1}{\phi(lmp_0)}. \tag{2.13}$$

The above sum on $m$ may be handled by an elementary argument,

$$\sum_{\substack{m\sim M\\ (m,a)=1}} \frac{\phi(l)}{\phi(lm)} = G_l \log 2 + O\big(\tau(a)(\log x)/M\big),$$

where $G_l = \frac{\phi(a)}{a}\prod_{p\nmid al}\Big(1+\frac{1}{p(p-1)}\Big)$. Note $G_l \gg_a 1$. Hence we obtain

$$N_1 = \big(1+O(1/\log x)\big)\,\pi(x)\log 2\sum_{l\in\mathcal{G}}\frac{G_l}{\phi(l)}\sum_{x^\beta<p_0\le 2x^{1-\theta}}\frac{1}{p_0}. \tag{2.14}$$

By an analogous (simpler) argument, we have

$$|\mathcal{N}| = \sum_{l\in\mathcal{G}}\sum_{\substack{m\sim x^{1-\theta}\\(m,a)=1}}\sum_{\substack{p\sim x\\ p\equiv a\ (lm)}} 1 = \big(1+O(1/\log x)\big)\,\pi(x)\log 2\sum_{l\in\mathcal{G}}\frac{G_l}{\phi(l)}. \tag{2.15}$$

Indeed, this is [4, (2.6)] which follows already by [13, Theorem 9]. In particular both $|\mathcal{N}|, N_1$ are of size $\gg x/\log x$. From (2.14), (2.15) we obtain

$$N_1 = \big(1+O(1/\log x)\big)|\mathcal{N}|\sum_{x^\beta<p_0\le 2x^{1-\theta}}\frac{1}{p_0} \;=\; (1+o(1))|\mathcal{N}|\cdot\log(\tfrac{1-\theta}{\beta}). \tag{2.16}$$

Hence provided (2.12) holds, i.e. $|\mathcal{N}_1| = (1+O(\varepsilon))N_1$, plugging (2.16) back into (2.10) gives

$$|\mathcal{N}'| \ge |\mathcal{N}| - 2|\mathcal{N}_1| \;=\; (1+O(\varepsilon))|\mathcal{N}|\big(1-2\log(\tfrac{1-\theta}{\beta})\big) \gg |\mathcal{N}| \gg \frac{x}{\log x}, \tag{2.17}$$

using $\beta > (1-\theta)/\sqrt{e}$. This gives (2.9) as desired. Thus it remains to show (2.12).

To prove (2.12), we subtract (2.13) from (2.11), giving

$$|\mathcal{N}_1| - N_1 = \sum_{l\in\mathcal{G}}\sum_{\substack{m\sim x^{1-\theta}/p_0\\(m,a)=1}}\sum_{x^\beta<p_0\le 2x^{1-\theta}}\Delta(lmp_0), \tag{2.18}$$

where

$$\Delta(k) := \sum_{\substack{p\sim x\\(p,a)=1}}\Big(\mathbf{1}_{p\equiv a\ (\mathrm{mod}\ k)} - \frac{1}{\phi(k)}\Big).$$

Note by a standard sieve upper bound (c.f. (7.2) in [103]), all but $O(\varepsilon M)$ integers $m\sim M$ factor as $m = m_1p_1$, for a prime $p_1\in[x^{5\varepsilon}, x^{.01}]$. Indeed,

$$\sum_{\substack{m\sim M\\ \nexists p_1\mid m, p_1\in[x^{5\varepsilon},x^{.01}]}} 1 \ll \sum_{r<x^{5\varepsilon}}\sum_{\substack{m'\sim M/r\\ P^-(m')>x^{.01}}} \ll \frac{M}{\log x}\sum_{r<x^{5\varepsilon}}\frac{1}{r} \ll \varepsilon M.$$

Thus by the Brun-Titchmarsh theorem, those exceptional $m$ (without such a prime factor) contribute to (2.18) at most

$$\sum_{l\in\mathcal{G}}\sum_{\substack{x^\beta<p_0\le 2x^{1-\theta}\\ m\sim x^{1-\theta}/p_0\\ \nexists p_1\mid m, p_1\in[x^{5\varepsilon},x^{.01}]}}\Delta(lmp_0) \ll \sum_{l\in\mathcal{G}}\sum_{\substack{x^\beta<p_0\le 2x^{1-\theta}\\ m\sim x^{1-\theta}/p_0\\ \nexists p_1\mid m, p_1\in[x^{5\varepsilon},x^{.01}]}}\frac{\pi(x)}{\phi(lmp_0)} \ll \varepsilon\pi(x).$$

So letting $\omega^*(m) = \#\{p_1 \in [x^{5\varepsilon}, x^{.01}] : p_1 \mid m\}$, we may rewrite (2.18) as

$$|\mathcal{N}_1| - N_1 + O\big(\varepsilon\pi(x)\big) = \sum_{l\in\mathcal{G}} \sum_{\substack{m\sim M\\(m,a)=1\\ \exists p_1 | m,\, p_1\in[x^{5\varepsilon},x^{.01}]}} \sum_{x^\beta<p_0\le 2x^{1-\theta}} \Delta(lmp_0)$$
$$= \sum_{l\in\mathcal{G}} \sum_{x^\beta<p_0\le 2x^{1-\theta}} \sum_{p_1\in[x^{5\varepsilon},x^{.01}]} \sum_{\substack{m_1\sim x^{1-\theta}/p_0p_1\\(m_1,a)=1}} \frac{\Delta(p_0p_1m_1l)}{\omega^*(m_1)+1}. \quad (2.19)$$

Next, we split into short intervals $p_1 \in [P_1, P_1(1+\varepsilon)]$, and replace the sum over $m_1 \sim x^{1-\theta}/p_0p_1$ with a sum over $m_1 \sim x^{1-\theta}/p_0P_1$. Note the discrepancy between these sums consist of those $m_1$ satisfying either $m_1 \sim x^{1-\theta}/p_0p_1$ or $m_1 \sim x^{1-\theta}/p_0P_1$, but not both. Since $|p_1 - P_1| \le \varepsilon P_1$, the contribution of such $m_1$ to (2.19) is $O(\varepsilon\pi(x))$ by Brun–Titchmarsh. Hence (2.19) becomes

$$|\mathcal{N}_1| - N_1 + O\big(\varepsilon\pi(x)\big) = \quad (2.20)$$
$$\sum_{P_1=(1+\varepsilon)^j\in[x^{3\varepsilon},x^{.01}]} \sum_{l\in\mathcal{G}} \sum_{x^\beta<p_0\le 2x^{1-\theta}} \sum_{p_1\in[P_1,(1+\varepsilon)P_1]} \sum_{\substack{m_1\sim x^{1-\theta}/p_0P_1\\(m_1,a)=1}} \frac{\Delta(p_0p_1m_1l)}{\omega^*(m_1)+1}.$$

Observe we have now separated the sums over variables $p_1$ and $m_1$.

By definition, each $l \in \mathcal{G}$ factors as $l = l_1\cdots l_H$ for $l_i \in [x^{.09\varepsilon}, x^{.1\varepsilon}]$. In particular for any real $\lambda \in [1, l]$, there exists an index $I \le H$ such that $\lambda/l_1\cdots l_I \in [x^{.09\varepsilon}, x^{.1\varepsilon}]$. Thus for each $P_1 \in [x^{5\varepsilon}, x^{.01}]$ there exist indices $I, J \le H$ such that

$$s := l_1\cdots l_I, \qquad s \in x^{1/64}\cdot[x^{-2.1\varepsilon}, x^{-2\varepsilon}], \quad (2.21)$$
$$t := l_{I+1}\cdots l_{2I}, \qquad t \in x^{1/64}\cdot[x^{-2.1\varepsilon}, x^{-2\varepsilon}], \quad (2.22)$$
$$q_1 := l_{2I+1}\cdots l_J, \qquad q_1 \in P_1\cdot[x^{-2.1\varepsilon}, x^{-2\varepsilon}], \quad (2.23)$$

And since $l = l_1\cdots l_H \asymp x^{2\theta-1} = x^{1/16-8\varepsilon}$, we see that

$$r_1 := l_{J+1}\cdots l_H, \qquad r_1 = \frac{l}{q_1st} \in x^{1/32}/P_1\cdot[x^{-2\varepsilon}, x^{-1.7\varepsilon}]. \quad (2.24)$$

That is, for each $P_1$, we obtain a factorization of $l$ as $l = q_1r_1st$, where the factors lie in the specified ranges above. Note $P_1 \in [x^{5\varepsilon}, x^{.01}]$ implies $q_1 \ge x^{2\varepsilon}$.

Recalling $p_1 \sim P_1$, we see (2.24) implies

$$r := p_1r_1, \qquad r \in x^{1/32}\cdot[x^{-2\varepsilon}, x^{-1.7\varepsilon}]. \quad (2.25)$$

Similarly, recalling $p_0m_1P_1 \sim x^{1-\theta} = x^{\frac{15}{32}+4\varepsilon}$, we see (2.23) implies

$$q := p_0m_1q_1, \qquad q \in x^{15/32}\cdot[x^{1.9\varepsilon}, x^{2\varepsilon}]. \quad (2.26)$$

Hence we obtain a factorization of the modulus $p_0p_1m_1l \sim x^\theta = x^{\frac{17}{32}-4\varepsilon}$ in (2.20) as

$$p_0p_1m_1l = p_0p_1m_1(q_1r_1st) = qrst,$$

with $s$, $t$ as in (2.21)–(2.22) and $r$, $q$ as in (2.25)–(2.26).

In particular $s, t \le x^{1/64-2\varepsilon}$, $r \le x^{1/32-\varepsilon}$, and $q \le x^{15/32+2\varepsilon}$. Thus (2.20) becomes

$$\begin{aligned} |\mathcal{N}_1| - N_1 \ll\ & \varepsilon\pi(x) + \\ & + \frac{\log x}{\log(1+\varepsilon)} \sup_{P_1\in[x^{5\varepsilon},x^{.01}]} \sum_{q\le x^{15/32+2\varepsilon}} \sum_{\substack{r\le x^{1/32-\varepsilon}\\(qrst,a)=1}} \sum_{s,\,t\le x^{1/64-2\varepsilon}} \lambda_q\nu_r\eta_s\mu_t\Delta(qrst), \end{aligned} \tag{2.27}$$

for the choice of coefficients $\lambda_q$, $\nu_r$, $\eta_s$, $\mu_t$,

$$\eta_s = \sum_{\substack{l_i\sim l_0\\ s=l_1\cdots l_I}} 1, \qquad \mu_t = \sum_{\substack{l_i\sim l_0\\ t=l_{I+1}\cdots l_{2I}}} 1,$$

$$\nu_r = \sum_{\substack{p_1\in[P_1,(1+\varepsilon)P_1]\\ r=p_1l_{J+1}\cdots l_H}} \sum_{l_i\sim l_0} 1, \qquad \lambda_q = \sum_{x^\beta<p_0\le 2x^{1-\theta}} \sum_{\substack{m_1\sim x^{1-\theta}/p_0P_1\\ q=p_0m_1l_{2I+1}\cdots l_J}} \sum_{l_i\sim l_0} \frac{1}{\omega^*(m_1)+1}.$$

Now for any $P_1 \in [x^{5\varepsilon}, x^{.01}]$, we apply the key mean value theorem in Theorem 2.1.2, specifically (2.3). Thus (2.27) becomes

$$|\mathcal{N}_1| - N_1 \ \ll\ \frac{\varepsilon x}{\log x} \ +\ \sup_{P_1\in[x^{5\varepsilon},x^{.01}]} \frac{x}{(\log x)^{A-1}} \ \ll\ \varepsilon N_1. \tag{2.28}$$

This gives (2.12) as claimed, and hence completes the proof of Theorem 2.1.1.

## 2.4 Main propositions

In this section we prove Theorem 2.1.2 assuming three new technical propositions, which we will then establish over the rest of the chapter. We do this by applying a sieve decomposition to break the count of primes in arithmetic progressions into counts of integers with particular prime factorizations, which can then be estimated using the relevant proposition. The sieve decomposition is based on ideas based on Harman's sieve (see [73]), but we could have used the Heath-Brown identity and some combinatorial lemmas as an alternative.

Define $S_n$ and $S_d(z)$ (depending on integers $a, q_0, q_1, q_2$ satisfying $(a, q_0q_1q_2) = 1$ which we suppress from the notation for convenience) for integers $n, d$ and a real $z$ by

$$\begin{aligned} S_n &:= \mathbf{1}_{n\equiv a \pmod{q_0q_1q_2}} - \frac{1}{\phi(q_0q_1q_2)}\mathbf{1}_{(n,q_0q_1q_2)=1}, \\ S_d(z) &:= \sum_{\substack{n\sim x/d\\ P^-(n)>z}} S_{dn}. \end{aligned}$$

With this notation, we are able to state our key propositions. The first result is a refinement of [101, Proposition 7.1]. We prove this in Section 2.8.

**Proposition 2.4.1** (Type II estimate)**.** *Let $A > 0$ and let $Q_1, Q_1, Q_2$ satisfy*

$$Q_0^2 Q_2 < x^{1/2-\epsilon}, \qquad Q_0 Q_1 Q_2^2 < x^{1-\epsilon}, \qquad Q_0^{10} Q_1^{12} Q_2^7 < x^{4-\epsilon}. \tag{2.29}$$

*Let $P_1, \dots, P_J \geq x^{1/7+10\epsilon}$ be such that $P_1 \cdots P_J \asymp x$ and*

$$x^{3/7+\epsilon} \leq \prod_{j \in \mathcal{J}} P_j \leq x^{4/7-\epsilon}$$

*for some subset $\mathcal{J} \subseteq \{1, \dots, J\}$. Let $\eta_q, \lambda_q, \nu_q$ be complex sequences with $|\eta_q|, |\lambda_q|, |\nu_q| \leq \tau(q)^{B_0}$. Then we have*

$$\sum_{\substack{q_0 \sim Q_0 \\ (q_0 q_1 q_2, a) = 1}} \sum_{q_1 \sim Q_1} \sum_{q_2 \sim Q_2} \eta_{q_0} \lambda_{q_1} \nu_{q_2} \sum_{\substack{p_1, \dots, p_J \\ p_i \sim P_i \, \forall i}}^{*} S_{p_1 \cdots p_J}(p_J) \ll_{A, B_0} \frac{x}{(\log x)^A}.$$

*Here $\sum^*$ indicates that the summation is restricted by $O(1)$ inequalities of the form $p_1^{\alpha_1} \cdots p_J^{\alpha_J} \leq B$. The implied constant may depend on all such exponents $\alpha_i$, but none of the quantities $B$.*

The second result is a refinement of [101, Proposition 7.2]. We prove this in Section 2.10.

**Proposition 2.4.2** (Sieve asymptotics)**.** *Let $A > 0$. Let $x^{3/7+\epsilon} \geq P_1 \geq \cdots \geq P_r \geq x^{1/7+10\epsilon}$ be such that $P_1 \cdots P_r \leq x^{2/3}$ and such that either $r = 1$ or $P_r \leq x^{1/4+\epsilon}$. Let $Q_1, Q_1, Q_2$ satisfy*

$$Q_0^2 Q_2 < x^{1/2-\epsilon}, \qquad Q_0 Q_1 Q_2^2 < x^{1-\epsilon}, \qquad Q_0^{10} Q_1^{12} Q_2^7 < x^{4-\epsilon}.$$

*Let $\eta_q, \lambda_q, \nu_q$ be complex sequences with $|\eta_q|, |\lambda_q|, |\nu_q| \leq \tau(q)^{B_0}$. Then we have*

$$\sum_{\substack{q_0 \sim Q_0 \\ (q_0 q_1 q_2, a) = 1}} \sum_{q_1 \sim Q_1} \sum_{q_2 \sim Q_2} \eta_{q_0} \lambda_{q_1} \nu_{q_2} \sum_{\substack{p_1, \dots, p_r \\ p_i \sim P_i \, \forall i}}^{*} S_{p_1 \cdots p_r}(x^{1/7+10\epsilon}) \ll_{A, B_0} \frac{x}{(\log x)^A}.$$

*Here $\sum^*$ means that the summation is restricted to $O(1)$ inequalities of the form $p_1^{\alpha_1} \cdots p_r^{\alpha_r} \leq B$ for some constants $\alpha_1, \dots \alpha_r$. The implied constant may depend on all such exponents $\alpha_i$, but none of the quantities $B$.*

*Moreover, we also have the related estimate*

$$\sum_{\substack{q_0 \sim Q_0 \\ (q_0 q_1 q_2, a) = 1}} \sum_{q_1 \sim Q_1} \sum_{q_2 \sim Q_2} \eta_{q_0} \lambda_{q_1} \nu_{q_2} S_1(x^{1/7+10\epsilon}) \ll_{A, B_0} \frac{x}{(\log x)^A}.$$

The third result is a refinement of [101, Proposition 7.3]. We prove this in Section 2.9.

**Proposition 2.4.3** (Numbers with 4 or more prime factors)**.** *Let $A > 0$. Let $J \geq 4$ and $P_1 \geq \cdots \geq P_J \geq x^{1/7+10\epsilon}$ with $P_1 \cdots P_J \asymp x$. Let $Q_1, Q_1, Q_2$ satisfy*

$$Q_0^2 Q_2 < x^{1/2-\epsilon}, \qquad Q_0 Q_1 Q_2^2 < x^{1-\epsilon}, \qquad Q_0^{10} Q_1^{12} Q_2^7 < x^{4-\epsilon}.$$

*Let $\eta_q, \lambda_q, \nu_q$ be complex sequences with $|\eta_q|, |\lambda_q|, |\nu_q| \leq \tau(q)^{B_0}$. Then we have*

$$\sum_{\substack{q_0 \sim Q_0 \\ (q_0 q_1 q_2, a) = 1}} \sum_{q_1 \sim Q_1} \sum_{q_2 \sim Q_2} \eta_{q_0} \lambda_{q_1} \nu_{q_2} \sideset{}{^*}\sum_{\substack{p_1, \dots, p_J \\ p_i \sim P_i \forall i}} S_{p_1 \cdots p_J} \ll_{A, B_0} \frac{x}{(\log x)^A}.$$

*Here $\sum^*$ indicates that the summation is restricted by $O(1)$ inequalities of the form $p_1^{\alpha_1} \cdots p_J^{\alpha_J} \leq B$. The implied constant may depend on all such exponents $\alpha_i$, but none of the quantities $B$.*

The final result is a refinement of [101, Proposition 7.4]. We prove this in Section 2.11.

**Proposition 2.4.4** (Numbers with three prime factors)**.** *Let $A > 0$ and let $P_1, P_2, P_3 \in [x^{1/4}, x^{3/7+\epsilon}]$ with $P_1 P_2 P_3 \asymp x$. Let $Q_1, Q_1, Q_2$ satisfy*

$$Q_0^2 Q_2 < x^{1/2-\epsilon}, \qquad Q_0 Q_1 Q_2^2 < x^{1-\epsilon}, \qquad Q_0^{10} Q_1^{12} Q_2^7 < x^{4-\epsilon}.$$

*Let $\eta_q, \lambda_q, \nu_q$ be complex sequences with $|\eta_q|, |\lambda_q|, |\nu_q| \leq \tau(q)^{B_0}$. Then we have*

$$\sum_{\substack{q_0 \sim Q_0 \\ (q_0 q_1 q_2, a) = 1}} \sum_{q_1 \sim Q_1} \sum_{q_2 \sim Q_2} \eta_{q_0} \lambda_{q_1} \nu_{q_2} \sideset{}{^*}\sum_{\substack{p_1, p_2, p_3 \\ p_i \sim P_i \forall i}} S_{p_1 p_2 p_3} \ll_{A, B_0} \frac{x}{(\log x)^A}.$$

*Here $\sum^*$ means that the summation is restricted to $O(1)$ inequalities of the form $p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} \leq B$ for some constants $\alpha_1, \alpha_2, \alpha_3$. The implied constant may depend on all such exponents $\alpha_i$, but none of the quantities $B$.*

*Proof of Theorem 2.1.2 assuming Propositions 2.4.1, 2.4.2, 2.4.3 and 2.4.4.* This follows just as in the proof of [101, Theorem 1.1], except for $Q_1$ replaced by $Q_0 Q_1$, and with trilinear weights $\eta_{q_0} \lambda_{q_1} \nu_{q_2}$ instead of absolute values. In this case, the Type II estimate and sieve asymptotics in Propositions 2.4.1, 2.4.2, 2.4.4 replace that of [101, Propositions 7.1, 7.2, 7.4], respectively. □

## 2.5 Preliminary Lemmas

In this section, we collect statements of some preliminary lemmas, which will be of use moving forward.

**Lemma 2.5.1** (Divisor function bounds)**.** *Let $|b| < x - y$ and $y \geq qx^{\epsilon}$. Then for any $C \geq 1$, we have*

$$\sum_{\substack{x-y \leq n \leq x \\ n \equiv a \pmod{q}}} \tau(n)^C \tau(n-b)^C \ll \frac{y}{q} (\tau(q) \log x)^{O_C(1)}.$$

*Proof.* This follows from Shiu's Theorem [131], and is given in [14, Lemma 12]. □

**Lemma 2.5.2** (Small sets contribute negligibly)**.** *Let $\delta > 0$, $Q \leq x^{1-\epsilon}$ and let $\mathcal{A} \subseteq [x, 2x]$. Then we have*

$$\sum_{q \sim Q} \tau(q) \Big| \sum_{\substack{n \in \mathcal{A} \\ n \equiv a \pmod{q}}} 1 - \frac{1}{\phi(q)} \sum_{\substack{n \in \mathcal{A} \\ (n,q)=1}} 1 \Big| \ll x^{\delta} \# \mathcal{A}^{1-\delta} (\log x)^{O_\delta(1)}.$$

*Proof.* See [101, Lemma 8.9]. □

**Lemma 2.5.3** (Separation of variables from inequalities)**.** *Let $Q_1 Q_2 \leq x^{1-\epsilon}$. Let $N_1, \dots, N_r \geq z_0$ satisfy $N_1 \cdots N_r \asymp x$. Let $\alpha_{n_1,\dots,n_r}$ be a complex sequence with $|\alpha_{n_1,\dots,n_r}| \leq (\tau(n_1) \cdots \tau(n_r))^{B_0}$ for some $B_0 \geq 1$. Then, for any choice of $A > 0$ there is a constant $C = C(A, B_0, r)$ and intervals $\mathcal{I}_1, \dots, \mathcal{I}_r$ with $\mathcal{I}_j \subseteq [P_j, 2P_j]$ of length $\leq P_j (\log x)^{-C}$ such that*

$$\sum_{q_1 \sim Q_1} \sum_{\substack{q_2 \sim Q_2 \\ (q_1 q_2, a)=1}} \Big| \sideset{}{^*}\sum_{\substack{n_1,\dots,n_r \\ n_i \sim N_i \forall i}} \alpha_{n_1,\dots,n_r} S_{n_1 \cdots n_r} \Big|$$
$$\ll_{r,A,B_0} \frac{x}{(\log x)^A} + (\log x)^{rC} \sum_{q_1 \sim Q_1} \sum_{\substack{q_2 \sim Q_2 \\ (q_1 q_2, a)=1}} \Big| \sum_{\substack{n_1,\dots,n_r \\ n_i \in \mathcal{I}_i \forall i}} \alpha_{n_1,\dots,n_r} S_{n_1 \cdots n_r} \Big|.$$

*Here $\sum^*$ means that the summation is restricted to $O(1)$ inequalities of the form $n_1^{\alpha_1} \cdots n_r^{\alpha_r} \leq B$ for some constants $\alpha_1, \dots \alpha_r$ and some quantity $B$. The implied constant may depend on all such exponents $\alpha_i$, but none of the quantities $B$.*

*Proof.* See [101, Lemma 8.10]. □

**Lemma 2.5.4.** *Let $C, B > 0$ be constants and let $\alpha_n$ be a sequence satisfing the Siegel-Walfisz condition (1), supported on $n \leq 2x$ with $P^-(n) \geq z_0 = x^{1/(\log\log x)^3}$ and satisfying $|\alpha_n| \leq \tau(n)^B$. Then $\mathbf{1}_{\tau(n) \leq (\log x)^C} \alpha_n$ also satisfies the Siegel-Walfisz condition.*

*Proof.* See [101, Lemma 13.7]. □

**Lemma 2.5.5** (Most moduli have small $z_0$-smooth part)**.** *Let $Q < x^{1-\epsilon}$. Let $\gamma_b, c_q$ be complex sequences with $|\gamma_b|, |c_b| \leq \tau(n)^{B_0}$ and recall $z_0 := x^{1/(\log\log x)^3}$ and $y_0 := x^{1/\log\log x}$. Let $\mathrm{sm}(n; z)$ denote the $z$-smooth part of $n$. (i.e. $\mathrm{sm}(n; z) = \prod_{p \leq z} p^{\nu_p(n)}$). Then for every $A > 0$ we have that*

$$\sum_{\substack{q \sim Q \\ \mathrm{sm}(q; z_0) \geq y_0}} c_q \sum_{b \leq x} \gamma_b \Big( \mathbf{1}_{b \equiv a \pmod q} - \frac{\mathbf{1}_{(b,q)=1}}{\phi(q)} \Big) \ll_{A, B_0} \frac{x}{(\log x)^A}.$$

*Proof.* See [101, Lemma 13.10]. □

**Lemma 2.5.6** (Splitting into coprime sets)**.** *Let $\mathcal{N} \subseteq \mathbb{Z}^2_{>0}$ be a set of pairs $(a, b)$ satisfying:*

1. *$a, b \leq x^{O(1)}$,*
2. *$\gcd(a, b) = 1$,*
3. *The number of prime factors of $a$ and of $b$ is $\ll (\log\log x)^3$.*

*Then there is a partition $\mathcal{N} = \mathcal{N}_1 \sqcup \mathcal{N}_2 \sqcup \cdots \sqcup \mathcal{N}_J$ into $J$ disjoint subsets with*

$$J \ll \exp\Big( O(\log\log x)^4 \Big),$$

*such that if $(a, b)$ and $(a', b')$ are in the same set $\mathcal{N}_j$, then $\gcd(a, b') = \gcd(a', b) = 1$.*

*Proof.* This follows immediately from [56, Lemme 6]. Also see [101, Lemma 13.2] □

## 2.6 Exponential sum estimates

**Lemma 2.6.1** (Weil bound for Kloosterman sums)**.** *Let $S(m, n; c)$ be the standard Kloosterman sum*

$$S(m, n; c) := \sum_{\substack{b \pmod c \\ (b,c)=1}} e\Big( \frac{mb + n\bar{b}}{c} \Big).$$

*Then we have that*

$$S(m, n; c) \ll \tau(c) c^{1/2} \gcd(m, n, c)^{1/2}.$$

*Proof.* This is [80, Corollary 11.12]. Also see [101, Lemma 13.3] □

**Lemma 2.6.2** (Completion of inverses). *Let $C > 0$ and $f : \mathbb{R} \to \mathbb{R}$ be a smooth function which is supported on $[-10, 10]$ and satisfies $\|f^{(j)}\|_\infty \ll_j (\log x)^{jC}$ for all $j \geq 0$. Let $(d, q) = 1$. Then we have for any $H \geq x^\epsilon dq/N$*

$$\sum_{\substack{(n,q)=1\\ n\equiv n_0 \pmod d}} f\Big(\frac{n}{N}\Big)e\Big(\frac{b\overline{n}}{q}\Big) = \frac{N\hat{f}(0)}{dq}\sum_{(c,q)=1} e\Big(\frac{bc}{q}\Big) + \frac{N}{dq}\sum_{1\leq|h|\leq H}\hat{f}\Big(\frac{hN}{dq}\Big)e\Big(\frac{n_0\overline{q}h}{d}\Big)\sum_{\substack{c \pmod q\\ (c,q)=1}} e\Big(\frac{b\overline{dc}+hc}{q}\Big) + O_C(x^{-100}).$$

*Moreover, if $\|f^{(j)}\|_\infty \ll ((j+1)\log x)^{jC}$ then we have the same result for any $H \geq (\log x)^{2C+1}dq/N$.*

*Proof.* See [101, Lemma 13.5] □

**Proposition 2.6.3** (Reduction to exponential sums). *Let $\alpha_n, \beta_m, \gamma_{q,d}, \lambda_{q,d,r}$ be complex sequences with $|\alpha_n|, |\beta_n| \leq \tau(n)^{B_0}$ and $|\gamma_{q,d}| \leq \tau(qd)^{B_0}$ and $|\lambda_{q,d,r}| \leq \tau(qdr)^{B_0}$. Let $\alpha_n$ and $\lambda_{q,d,r}$ be supported on integers with $P^-(n) \geq z_0$ and $P^-(r) \geq z_0$, and let $\alpha_n$ satisfy the Siegel-Walfisz condition* (1). *Let*

$$\mathcal{S} := \sum_{\substack{d\sim D\\ (d,a)=1}}\sum_{\substack{q\sim Q\\ (q,a)=1}}\sum_{\substack{r\sim R\\ (r,a)=1}} \lambda_{q,d,r}\gamma_{q,d}\sum_{m\sim M}\beta_m\sum_{n\sim N}\alpha_n\Big(\mathbf{1}_{mn\equiv a \pmod{qrd}} - \frac{\mathbf{1}_{(mn,qrd)=1}}{\phi(qrd)}\Big).$$

*Let $A > 0$, $1 \leq E \leq x$ and $C = C(A, B_0)$ be sufficiently large in terms of $A, B_0$, and let $N, M$ satisfy*

$$N > QDE(\log x)^C, \qquad M > (\log x)^C.$$

*Then we have*

$$|\mathcal{S}| \ll_{A,B_0} \frac{x}{(\log x)^A} + MD^{1/2}Q^{1/2}(\log x)^{O_{B_0}(1)}\Big(|\mathcal{E}_1|^{1/2} + |\mathcal{E}_2|^{1/2}\Big),$$

*where*

$$\mathcal{E}_1 := \sum_{e\sim E}\mu^2(e)\sum_{\substack{q\\(q,a)=1}}\sum_{\substack{d\sim D\\(d,a)=1}}\sum_{\substack{r_1,r_2\sim R\\(r_1r_2,a)=1}}\psi_0\Big(\frac{q}{Q}\Big)\frac{\lambda_{q,d,r_1}\overline{\lambda_{q,d,r_2}}}{\phi(qder_2)qdr_1}$$
$$\times\sum_{\substack{n_1,n_2\sim N\\(n_1,qder_1)=1\\(n_2,qder_2)=1}}\alpha_{n_1}\overline{\alpha_{n_2}}\sum_{1\le|h|\le H_1}\hat{\psi}_0\Big(\frac{hM}{qdr_1}\Big)e\Big(\frac{ah\overline{n_1}}{qdr_1}\Big),$$

$$\mathcal{E}_2 := \sum_{e\sim E}\mu^2(e)\sum_{\substack{q\\(q,a)=1}}\psi_0\Big(\frac{q}{Q}\Big)\sum_{\substack{d\sim D\\(d,a)=1}}\sum_{\substack{r_1,r_2\sim R\\(r_1,ar_2)=1\\(r_2,aqdr_1)=1}}\frac{\lambda_{q,d,r_1}\overline{\lambda_{q,d,r_2}}}{qdr_1r_2}$$
$$\times\sum_{\substack{n_1,n_2\sim N\\n_1\equiv n_2\pmod{qde}\\(n_1,n_2eqdr_1)=1\\(n_2,n_1eqdr_2)=1\\|n_1-n_2|\ge N/(\log x)^C}}\alpha_{n_1}\overline{\alpha_{n_2}}\sum_{1\le|h|\le H_2}\hat{\psi}_0\Big(\frac{hM}{qdr_1r_2}\Big)e\Big(\frac{ah\overline{n_1r_2}}{qdr_1}+\frac{ah\overline{n_2qdr_1}}{r_2}\Big),$$

$$H_1 := \frac{QDR}{M}\log^5 x,$$
$$H_2 := \frac{QDR^2}{M}\log^5 x.$$

*Proof.* This is [101, Theorem 14.4]. □

**Lemma 2.6.4** (Simplification of exponential sum)**.** *Let $N, M, Q, R, S \le x$ with $NM \asymp x$ and*

$$QR < x^{2/3}, \tag{2.30}$$
$$QR^2 < Mx^{1-2\epsilon}. \tag{2.31}$$

*Let $\lambda_{q,r}$ and $\alpha_n$ be complex sequences supported on $P^-(n), P^-(r) \ge z_0$ with $|\lambda_{q,r}| \le \tau(qr)^{B_0}$ and $|\alpha_n| \le \tau(n)^{B_0}$. Let $H := \frac{QR^2}{M}\log^5 x$ and let*

$$\mathcal{E} := \sum_{(q,a)=1}\psi_0\Big(\frac{q}{Q}\Big)\sum_{\substack{r_1,r_2\sim R\\(r_1,ar_2)=1\\(r_2,aqr_2)=1}}\frac{\lambda_{q,r_1}\overline{\lambda_{q,r_2}}}{qr_1r_2}\sum_{\substack{n_1,n_2\sim N\\n_1\equiv n_2\pmod{q}\\(n_1,n_2qr_1)=1\\(n_2,n_1qr_2)=1\\|n_1-n_2|\ge N/(\log x)^C}}\alpha_{n_1}\overline{\alpha_{n_2}}$$
$$\times\sum_{1\le|h|\le H}\hat{\psi}_0\Big(\frac{hM}{qr_1r_2}\Big)e\Big(\frac{ah\overline{n_1r_2}}{qr_1}+\frac{ah\overline{n_2qr_1}}{r_2}\Big).$$

*Then we have (uniformly in $C$)*

$$\mathcal{E} \ll_{B_0} \exp((\log\log x)^5)\sup_{\substack{H'\le H\\Q'\le 2Q\\R_1,R_2\le 2R}}|\mathcal{E}'| + \frac{N^2}{Qx^\epsilon},$$

*where*

$$\mathcal{E}' = \sum_{\substack{Q\le q\le Q'\\(q,a)=1}} \sum_{\substack{R\le r_1\le R_1\\R\le r_2\le R_2\\(r_1ar_2)=1\\(r_2,aqr_1)=1}} \frac{\lambda_{q,r_1}\overline{\lambda_{q,r_2}}}{qr_1r_2} \sum_{\substack{n_1,n_2\sim N\\n_1\equiv n_2 \pmod q\\(n_1,qr_1n_2)=1\\(n_2,qr_2n_1)=1\\(n_1r_2,n_2)\in\mathcal{N}\\|n_1-n_2|\ge N/(\log x)^C}} \alpha_{n_1}\overline{\alpha_{n_2}} \sum_{1\le|h|\le H'} e\Big(\frac{ah\overline{n_2qr_1}(n_1-n_2)}{n_1r_2}\Big),$$

*and $\mathcal{N}$ is a set with the property that if $(a,b)\in\mathcal{N}$ and $(a',b')\in\mathcal{N}$ then we have* $\gcd(a,b')=\gcd(a',b)=1$.

*Proof.* This is [101, Theorem 14.5]. □

**Lemma 2.6.5** (Deshouillers–Iwaniec)**.** *Let $b_{n,r,s}$ be a 1-bounded sequence and parameters $R,S,N,D,C\ll x^{O(1)}$. Let $g(c,d)=g_0(c/C,d/D)$ where $g_0$ is a smooth function supported on $[1/2,5/2]\times[1/2,5/2]$. Then we have*

$$\sum_{r\sim R}\sum_{\substack{s\sim S\\(r,s)=1}}\sum_{n\sim N} b_{n,r,s}\sum_{d\sim D}\sum_{\substack{c\sim C\\(rd,sc)=1}} g(c,d)e\Big(\frac{n\overline{dr}}{cs}\Big) \ll_{g_0} x^{\epsilon}\Big(\sum_{r\sim R}\sum_{s\sim S}\sum_{n\sim N}|b_{n,r,s}|^2\Big)^{1/2}\mathcal{J}.$$

*where*

$$\mathcal{J}^2 = CS(RS+N)(C+DR)+C^2DS\sqrt{(RS+N)R}+D^2NR.$$

*Proof.* This is [101, Theorem 15.1]. Also see [33, Theorem 12] (correcting a minor typo in the last term of $\mathcal{J}^2$ which is written as $D^2NR/S$). □

## 2.7 Zhang-style estimates

In this section we establish a new Zhang-style exponential sum estimate.

**Lemma 2.7.1** (Zhang exponential sum estimate)**.** *Let $Q,R,S,M,N$ satisfy $NM\asymp x$ and*

$$Q^7R^{12}S^{10}<x^{4-18\epsilon},\qquad Q<N<\frac{x^{1-5\epsilon}}{QS^2},$$

*and let $H\ll QNR^2S^2/x^{1-\epsilon}$. Let $\lambda_{q,s}$, $\nu_r$ and $\alpha_n$ be 1-bounded complex sequences supported on $r,s$ with $P^-(rs)\ge z_0$. Let*

$$\mathcal{Z} := \sum_{\substack{q\sim Q\\(q,a)=1}}\sum_{s_1,s_2\sim S}\sum_{\substack{r_1,r_2\sim R\\(r_1s_1,ar_2s_2)=1\\(r_2s_2,aqr_1s_1)=1}}\sum_{\substack{n_1,n_2\sim N\\n_1\equiv n_2 \pmod q\\(n_1,qr_1s_1n_2)=1\\(n_2,qr_2s_2n_1)=1\\|n_1-n_2|\ge N/(\log x)^C}} \frac{\alpha_{n_1}\overline{\alpha_{n_2}}\lambda_{q,s_1}\nu_{r_1}\overline{\lambda_{q,s_2}\nu_{r_2}}}{qr_1r_2s_1s_2}$$
$$\times\sum_{1\le|h|\le H}\hat{\psi}_0\Big(\frac{hM}{qr_1r_2s_1s_2}\Big)e\Big(ah\Big(\frac{\overline{n_1r_2s_2}}{qr_1s_1}+\frac{\overline{n_2qr_1s_1}}{r_2s_2}\Big)\Big),$$

*Then we have*

$$\mathcal{Z} \ll \frac{N^2}{Qx^\epsilon}.$$

*Proof.* We assume throughout that $H \ll QR^2S^2N/x^{1-\epsilon}$ and that $Q \ll N$, and deduce the other conditions are sufficient to give the result.

Since we only consider $r_1, r_2, s_1, s_2$ with $P^-(r_1r_2s_1s_2) \geq z_0$, we see $r_1, r_2, s_1, s_2$ have at most $(\log\log x)^3$ prime factors. Therefore, by Lemma 2.5.6, there are at most $O(\exp(\log\log x)^5))$ different sets $\mathcal{N}_1, \mathcal{N}_2, \dots$ which cover all possible pairs $(r_1s_1, r_2s_2)$, and such that if $(r_1s_1, r_2s_2), (r_1's_1, r_2's_2) \in \mathcal{N}_j$ then $\gcd(r_1s_1, r_2's_2) = \gcd(r_1's_1, r_2s_2) = 1$. Taking the worst such set $\mathcal{N}$, we see that

$$\mathcal{Z} \ll \exp((\log\log x)^5)\Big| \sum_{\substack{q\sim Q\\(q,a)=1}} \sum_{s_1,s_2\sim S} \sum_{\substack{r_1,r_2\sim R\\(r_1s_1,ar_2s_2)=1\\(r_2s_2,aqr_1s_1)=1\\(r_1s_1,r_2s_2)\in\mathcal{N}}} \sum_{\substack{n_1,n_2\sim N\\n_1\equiv n_2 \pmod q\\(n_1,qr_1s_1n_2)=1\\(n_2,qr_2s_2n_1)=1\\|n_1-n_2|\geq N/(\log x)^C}} \frac{\alpha_{n_1}\overline{\alpha_{n_2}}\lambda_{q,s_1}\nu_{r_1}\overline{\lambda_{q,s_2}\nu_{r_2}}}{qr_1r_2s_1s_2}$$
$$\times \sum_{1\leq|h|\leq H} \hat{\psi}_0\Big(\frac{hM}{qr_1r_2s_1s_2}\Big) e\Big(ah\Big(\frac{\overline{n_1r_2s_2}}{qr_1s_1} + \frac{\overline{n_2qr_1s_1}}{r_2s_2}\Big)\Big)\Big|.$$

We apply Cauchy-Schwarz in the $n_1, n_2, s_1, s_2, q$ variables to eliminate the $\alpha_n, \lambda_{q,s}$ coefficients and insert a smooth majorant for the $n_1$ and $n_2$ summations. This gives (using $Q \ll N$)

$$\mathcal{Z}^2 \ll \exp(2(\log\log x)^5)\Big(\sum_{q\sim Q}\sum_{s_1,s_2\sim S}\sum_{\substack{n_1,n_2\sim N\\n_1\equiv n_2 \pmod q}} \frac{1}{(qs_1s_2)^2}\Big)|\mathcal{Z}_2| \ll \frac{x^\epsilon N^2}{Q^2S^2}|\mathcal{Z}_2|,$$

where

$$\mathcal{Z}_2 := \sum_{\substack{q\sim Q\\(q,a)=1}} \sum_{s_1,s_2\sim S} \sum_{\substack{n_1,n_2\\(n_1n_2,q)=1}} \psi_0\Big(\frac{n_1}{N}\Big)\psi_0\Big(\frac{n_2}{N}\Big)$$
$$\times \Big| \sum_{\substack{r_1,r_2\sim R\\(r_1s_1,ar_2s_2n_1)=1\\(r_2s_2,aqr_1s_1n_2)=1\\(r_1s_1,r_2s_2)\in\mathcal{N}}} \frac{\nu_{r_1}\overline{\nu_{r_2}}}{r_1r_2} \sum_{1\leq|h|\leq H} \hat{\psi}_0\Big(\frac{hM}{qr_1r_2s_1s_2}\Big) e\Big(ah\Big(\frac{\overline{n_1r_2s_2}}{qr_1s_1} + \frac{\overline{n_2qr_1s_1}}{r_2s_2}\Big)\Big)\Big|^2$$
$$\leq \frac{1}{R^4} \sum_{q\sim Q}\sum_{s_1,s_2\sim S} \sum_{\substack{r_1,r_1',r_2,r_2'\sim R\\(qr_1r_1',r_2r_2')=1}} \sum_{1\leq|h|,|h'|\leq H}$$
$$\times \Big| \sum_{\substack{n_1,n_2\\n_1\equiv n_2 \pmod q\\(n_1,qr_1r_1's_1)=1\\(n_2,r_2r_2's_2)=1}} \psi_0\Big(\frac{n_1}{N}\Big)\psi_0\Big(\frac{n_2}{N}\Big) e\Big(\frac{c_1\overline{n_1}}{qr_1r_1's_1} + \frac{c_2\overline{n_2}}{r_2r_2's_2}\Big)\Big| =: \frac{1}{R^4}|\mathcal{Z}_3|$$

where $c_1 \pmod{qr_1r_1's_1}$ and $c_2 \pmod{r_2r_2's_2}$ are given by

$$c_1 = a(hr_1'r_2' - h'r_1r_2)\overline{r_2r_2's_2},$$
$$c_2 = a(hr_1'r_2' - h'r_1r_2)\overline{qr_1r_1's_1}.$$

(Here we used the fact that $(r_1s_1, r_2s_2), (r_1's_1, r_2's_2) \in \mathcal{N}$ to conclude $(r_1s_1, r_2's_2) = (r_1's_1, r_2s_2) = 1$.) In order to establish the desired bound $\mathcal{Z} \ll N^2/(x^\epsilon Q)$, it suffices to show $\mathcal{Z}_2 \ll N^2S^2/x^{3\epsilon}$, and so it suffices to prove

$$\mathcal{Z}_3 \ll \frac{N^2S^2R^4}{x^{3\epsilon}}. \tag{2.32}$$

We separate the diagonal terms $\mathcal{Z}_=$ with $hr_1'r_2' = h'r_1r_2$ and the off-diagonal terms $\mathcal{Z}_\neq$ with $hr_1'r_2' \neq h'r_1r_2$.

$$\mathcal{Z}_3 \ll \mathcal{Z}_= + \mathcal{Z}_\neq. \tag{2.33}$$

We first consider the diagonal terms $\mathcal{Z}_=$. Given a choice of $h, r_1', r_2'$ there are $x^{o(1)}$ choices of $h', r_1, r_2$ by the divisor bound. Thus, estimating the remaining sums trivially we have (using $Q \ll N$ and $H \ll NQ(RS)^2/x^{1-\epsilon}$)

$$\mathcal{Z}_= \ll x^{o(1)}Q(RS)^2HN\Big(\frac{N}{Q}+1\Big) \ll \frac{N^3Q(RS)^4}{x^{1-2\epsilon}}. \tag{2.34}$$

Now we consider the off-diagonal terms $\mathcal{Z}_\neq$. By Lemma 2.6.2, we have that

$$\sum_{\substack{n_2 \equiv n_1 \pmod q \\ (n_2, r_2r_2's_2)=1}} \psi_0\Big(\frac{n_2}{N}\Big)e\Big(\frac{c_2\overline{n_2}}{r_2r_2's_2}\Big)$$
$$= \frac{N}{qr_2r_2's_2} \sum_{|\ell_2| \leq x^\epsilon QSR^2/N} \hat{\psi}_0\Big(\frac{\ell_2 N}{qr_2r_2's_2}\Big) S(c_2, \ell_2\overline{q}; r_2r_2's_2)e\Big(\frac{\ell_2 n_1 \overline{r_2r_2's_2}}{q}\Big) + O(x^{-100}).$$

here $S(m,n;c)$ is the standard Kloosterman sum. By Lemma 2.6.2 again, we have that

$$\sum_{(n_1, qr_1r_1's_1)=1} \psi_0\Big(\frac{n_1}{N}\Big)e\Big(\frac{c_1\overline{n_2} + \ell_2 r_1r_1's_1n_1\overline{r_2r_2's_2}}{qr_1r_1's_1}\Big) \; + \; O(x^{-100})$$
$$= \frac{N}{qr_1r_1's_1} \sum_{|\ell_1| \leq x^\epsilon QSR^2/N} \hat{\psi}_0\Big(\frac{\ell_1 N}{qr_1r_1's_1}\Big) S(c_1, \ell_1\overline{q}; r_1r_1's_1) S(c_1, \ell_1\overline{r_1r_1's_1} + \ell_2\overline{r_2r_2's_2}; q).$$

Thus, we see that $\mathcal{Z}_3$ is a sum of Kloosterman sums. By the standard Kloosterman sum bound of Lemma 2.6.1 $S(m,n;c) \ll \tau(c)c^{1/2}(m,n,c)^{1/2} \ll c^{1/2+o(1)}(m,c)^{1/2}$, the

inner sum has the bound

$$\sum_{\substack{n_1,n_2\\ n_1\equiv n_2 \pmod q\\ (n_1,qr_1r_1's_1)=1\\ (n_2,r_2r_2's_2)=1}} \psi_0\Big(\frac{n_1}{N}\Big)\psi_0\Big(\frac{n_2}{N}\Big)e\Big(\frac{c_1\overline{n_1}}{qr_1r_1's_1}+\frac{c_2\overline{n_2}}{r_2r_2's_2}\Big)$$

$$\ll \frac{x^{o(1)}N^2}{Q^2S^2R^4}\sum_{\substack{|\ell_1|\le x^\epsilon QSR^2/N\\ |\ell_2|\le x^\epsilon QSR^2/N}} (QR^4S^2)^{1/2}\cdot(c_2,r_2r_2's_2)^{1/2}(c_1,qr_1r_1's_1)^{1/2}$$

$$\ll x^{3\epsilon}Q^{1/2}R^2S\cdot(hr_1'r_2'-h'r_1r_2,qr_1r_1'r_2r_2's_1s_2)^{1/2}$$

Substituting this into our expression for $\mathcal{Z}_{\neq}$ gives

$$\begin{aligned}\mathcal{Z}_{\neq} &\ll x^{3\epsilon}Q^{1/2}R^2S\sum_{s_1,s_2\sim S}\sum_{r_1,r_1'\sim R}\sum_{r_2,r_2'\sim R}\sum_{\substack{1\le|h|,|h'|\le H\\ hr_1'r_2'\neq h'r_1r_2}}\sum_{q\sim Q}(hr_1'r_2'-h'r_1r_2,qr_1r_1'r_2r_2's_1s_2)^{1/2}\\ &\ll x^{3\epsilon}Q^{1/2}R^2S\,(x^\epsilon S^2R^4H^2Q) \;=\; x^{4\epsilon}Q^{3/2}R^6S^3H^2\\ &\ll \frac{N^2Q^{7/2}R^{10}S^7}{x^{2-6\epsilon}}.\end{aligned} \tag{2.35}$$

Substituting (2.34) and (2.35) into (2.33) then gives

$$\mathcal{Z}_3 \ll \frac{N^3QR^4S^4}{x^{1-2\epsilon}}+\frac{N^2Q^{7/2}R^{10}S^7}{x^{2-6\epsilon}}.$$

This gives the desired bound (2.32) provided we have

$$N < \frac{x^{1-5\epsilon}}{QS^2}, \tag{2.36}$$

$$Q^7R^{12}S^{10} < x^{4-18\epsilon}. \tag{2.37}$$

This gives the result. □

**Lemma 2.7.2** (Second exponential sum estimate)**.** *Let $Q,R,M,N \le x^{O(1)}$ satisfy $NM \asymp x$ and*

$$NQ < x^{1-4\epsilon}, \qquad NQ^{5/2}R^3 < x^{2-4\epsilon}, \qquad N^2QR < x^{2-4\epsilon}.$$

*Let $\alpha_n$, $\lambda_{q,r}$ be 1-bounded complex sequences, $H_1 = (QR\log^5 x)/M$ and*

$$\widetilde{\mathcal{Z}} := \sum_{\substack{q\sim Q\\ (q,a)=1}}\sum_{\substack{r_1,r_2\sim R\\ (r_1r_2,a)=1}}\frac{\lambda_{q,r_1}\overline{\lambda_{q,r_2}}}{\phi(qr_2)qr_1}\sum_{\substack{n_1,n_2\sim N\\ (n_1,qr_1)=1\\ (n_2,qr_2)=1}}\alpha_{n_1}\overline{\alpha_{n_2}}\sum_{1\le|h|\le H_1}\hat{\psi}_0\Big(\frac{hM}{qr_1}\Big)e\Big(\frac{ah\overline{n_1}}{qr_1}\Big).$$

*Then we have*

$$\widetilde{\mathcal{Z}} \ll \frac{N^2}{Qx^\epsilon}.$$

*Proof.* See [101, Lemma 17.2] □

We are now able to establish the following Zhang-style estimate. This is a refinement of [101, Proposition 8.2].

**Proposition 2.7.3** (Zhang-style estimate)**.** *Let $A > 0$. Let $N, M, Q_0, Q_1, Q_2 \geq 1$ with $NM \asymp x$ be such that*

$$Q_0^{10} Q_1^{12} Q_2^7 < x^{4-20\epsilon}, \qquad x^{\epsilon} Q_2 < N < \frac{x^{1-6\epsilon}}{Q_0^2 Q_2}.$$

*Let $\beta_m, \alpha_n$ be complex sequences such that $|\alpha_n|, |\beta_n| \leq \tau(n)^{B_0}$ and such that $\alpha_n$ satisfies the Siegel-Walfisz condition* (1) *and $\alpha_n$ is supported on $n$ with all prime factors bigger than $z_0 = x^{1/(\log\log x)^3}$. Let $\eta_q, \lambda_q, \nu_q$ be 1-bounded complex sequences*

$$\Delta(q) := \sum_{m \sim M} \sum_{n \sim N} \alpha_n \beta_m \Big( \mathbf{1}_{mn \equiv a \pmod q} - \frac{\mathbf{1}_{(mn,q)=1}}{\phi(q)} \Big).$$

*Then we have*

$$\sum_{\substack{q_0 \sim Q_0 \\ (q_0 q_1 q_2, a) = 1}} \sum_{q_1 \sim Q_1} \sum_{q_2 \sim Q_2} \eta_{q_0} \lambda_{q_1} \nu_{q_2} \Delta(q_0 q_1 q_2) \ll_{A, B_0} \frac{x}{(\log x)^A}.$$

*Proof.* First we note that by Lemma 2.5.1 the set of $n, m$ with $\max(|\alpha_n|, |\beta_m|) \geq (\log x)^C$ has size $\ll x (\log x)^{O_{B_0}(1) - C}$, so by Lemma 2.5.2 these terms contribute negligibly if $C = C(A, B_0)$ is large enough. Thus, by dividing through by $(\log x)^{2C}$ and considering $A + 2C$ in place of $A$, it suffices to show the result when all the sequences are 1-bounded. ($\alpha_n$ still satisfies (1) by Lemma 2.5.4.)

We factor $q_0 = d_0 s$, $q_1 = d_1 q$, $q_2 = d_2 r$ into parts with large and small prime factors. By putting these in dyadic intervals, we see that it suffices to show for every $A > 0$ and every choice of $D_0 S \asymp Q_0$, $D_1 Q \asymp Q_1$, $D_2 R \asymp Q_2$ that

$$\sum_{\substack{s \sim S \\ P^-(qrs) > z_0 \\ (qrs, a) = 1}} \sum_{r \sim R} \sum_{q \sim Q} \sum_{\substack{d_0 \sim D_0 \\ z_0 \geq P^+(d_0 d_1 d_2)}} \sum_{d_1 \sim D_1} \sum_{d_2 \sim D_2} \eta_{s d_0} \nu_{r d_1} \lambda_{q d_2} \Delta(qrs d_0 d_1 d_2) \ll_A \frac{x}{(\log x)^A}.$$

By Lemma 2.5.5 we have the result unless $D_0, D_1, D_2 \leq y_0 = x^{1/\log\log x}$, may assume that $S = Q_0 x^{-o(1)}$, $R = Q_1 x^{-o(1)}$, $Q = Q_1 x^{-o(1)}$,. We let $d = d_0 d_1 d_2$, and extend the summation over $d_0, d_1, d_2$ to only have the constraint $d \leq y_0^3$ and then insert some divisor-bounded coefficients $c_d$ to absorb the conditions $z_0 \geq P^+(d), d \sim D$. Also we

modify the coefficients $\lambda'_q = \lambda_q \mathbf{1}_{P^-(q)>z_0}$, and similarly for $\nu'_r, \eta'_s$. Thus it suffices to show that

$$\sum_{\substack{d\le y_0^3\\(d,a)=1}} \sum_{s\sim S} \sum_{r\sim R} \sum_{q\sim Q} \eta'_s\nu'_r\lambda'_q c_d \Delta(qrsd) \ll_A \frac{x}{(\log x)^A}.$$

If we let

$$\lambda_{b_1,b_2,b_3} = \mathbf{1}_{b_1=1} \sum_{qrs=b_3} \eta'_s\nu'_r\lambda'_q c_{b_2}$$

then we see that we have a sum of the type considered in Proposition 2.6.3 (taking '$R$' to be $QR$, '$Q$' to be $D$ and '$E$' to be 1). By the assumptions of the proposition, we have that $NQRS \le NQ_0Q_1Q_2 < x^{-o(1)}$, so we have $H_1 = (QDRS\log^5 x)/M < 1$ and so the sum $\mathcal{E}_1$ of Proposition 2.6.3 vanishes. Therefore, by Proposition 2.6.3, it suffices to show that

$$\mathcal{E}_2 \ll \frac{N^2}{Qx^\epsilon},$$

where $H_2 = (QR^2S^2\log^5 x)/M$ and where

$$\mathcal{E}_2 := \sum_{(q,a)=1} \psi_0\Big(\frac{q}{Q}\Big) \sum_{\substack{r_1,r_2\sim R\\(r_1s_1,ar_2s_2)=1\\(r_2s_2,aqr_1s_1)=1}} \sum_{s_1,s_2\sim S} \frac{c_{q,r_1,s_1}\overline{c_{q,r_2,s_2}}}{qr_1r_2s_1s_2} \sum_{\substack{n_1,n_2\sim N\\n_1\equiv n_2 \pmod q\\(n_1,n_2qr_1)=1\\(n_2,n_1qr_2s_2)=1\\|n_1-n_2|\ge N/(\log x)^C}} \alpha_{n_1}\overline{\alpha_{n_2}}$$
$$\times \sum_{1\le|h|\le H_2} \hat{\psi}_0\Big(\frac{hM}{qr_1r_2s_1s_2}\Big) e\Big(\frac{ah\overline{n_1r_2s_2}}{qr_1s_1} + \frac{ah\overline{n_2qr_1s_1}}{r_2s_2}\Big).$$

Now absorbing the $\psi_0(q/Q)$ factors into the coefficients $c_{q,r,s}$, we see these are precisely the sums $\widetilde{\mathcal{Z}}$ and $\mathcal{Z}$ considered in Lemma 2.7.1 and Lemma 2.7.2. Thus, these lemmas give the result provided we have

$$Q^7R^{12}S^{10} < x^{4-18\epsilon}, \qquad Q < N < \frac{x^{1-5\epsilon}}{QS^2},$$
$$NQ < x^{1-4\epsilon}, \qquad NQ^{5/2}(RS)^3 < x^{2-4\epsilon}, \qquad N^2Q(RS) < x^{2-4\epsilon}.$$

We see that the first two conditions imply the final three (since we may assume $QRS \ge x^{1/2-\epsilon}$ or else the result follows from the Bombieri-Vinogradov theorem). Recalling that $Q = Q_2x^{o(1)}$, $R = Q_1x^{-o(1)}$, $S = Q_0x^{-o(1)}$ then gives the result. □

## 2.8 Proof of Proposition 2.4.1 (type II estimate)

In this section we prove Proposition 2.4.1 using the new Zhang-style estimate from the prior section, via the type II estimates below. Recall the notation

$$S_n := \mathbf{1}_{n\equiv a \pmod{q_0q_1q_2}} - \frac{1}{\phi(q_0q_1q_2)}\mathbf{1}_{(n,q_0q_1q_2)=1},$$
$$S_d(z) := \sum_{\substack{n\sim x/d\\ P^-(n)>z}} S_{dn}.$$

**Lemma 2.8.1** (Type II estimate away from $x^{1/2}$)**.** *Let $A>0$ and $Q_1Q_2 \le x^{127/224-\epsilon}$, and let $P_1,\dots,P_J \ge x^{1/7+10\epsilon}$ be such that $P_1\cdots P_J \asymp x$ and*

$$x^{\epsilon}Q_1Q_2 < \prod_{j\in\mathcal{J}} P_j < x^{4/7-\epsilon}$$

*for some subset $\mathcal{J}\subseteq\{1,\dots,J\}$. Then we have*

$$\sum_{q_1\sim Q_1}\sum_{\substack{q_2\sim Q_2\\ (q_1q_2,a)=1}}\Big|\sum_{\substack{p_1,\dots,p_J\\ p_i\sim P_i\forall i}}^{*} S_{p_1\cdots p_J}\Big| \ll_A \frac{x}{(\log x)^A}.$$

*Here $\sum^*$ indicates that the summation is restricted by $O(1)$ inequalities of the form $p_1^{\alpha_1}\cdots p_J^{\alpha_J}\le B$. The implied constant may depend on all such exponents $\alpha_i$, but none of the quantities $B$.*

*Proof.* This is [101, Propositon 8.12]. □

**Lemma 2.8.2** (Type II estimate near $x^{1/2}$)**.** *Let $A>0$ and let $Q_1,Q_1,Q_2$ satisfy*

$$Q_0^2Q_2 < x^{1/2-20\epsilon}, \qquad Q_0Q_1Q_2^2 < x^{1-20\epsilon}, \qquad Q_0^{10}Q_1^{12}Q_2^7 < x^{4-20\epsilon}.$$

*Let $P_1,\dots,P_J \ge x^{1/7+10\epsilon}$ be such that $P_1\cdots P_J\asymp x$ and*

$$\frac{x^{1-\epsilon}}{Q_0Q_1Q_2} \le \prod_{j\in\mathcal{J}} P_j \le x^{\epsilon}Q_0Q_1Q_2,$$

*for some subset $\mathcal{J}\subseteq\{1,\dots,J\}$.*

*Let $\eta_q,\lambda_q,\nu_q$ be complex sequences with $|\eta_q|,|\lambda_q|,|\nu_q|\le\tau(q)^{B_0}$. Then we have*

$$\sum_{q_0\sim Q_0}\sum_{\substack{q_1\sim Q_1\\ (q_0q_1q_2,a)=1}}\sum_{q_2\sim Q_2}\eta_{q_0}\lambda_{q_1}\nu_{q_2}\sum_{\substack{p_1,\dots,p_J\\ p_i\sim P_i\,\forall i}}^{*} S_{p_1\cdots p_J} \ll_{A,B_0} \frac{x}{(\log x)^A}.$$

*Here $\sum^*$ indicates that the summation is restricted by $O(1)$ inequalities of the form $p_1^{\alpha_1}\cdots p_J^{\alpha_J}\le B$. The implied constant may depend on all such exponents $\alpha_i$, but none of the quantities $B$.*

*Proof.* This follows quickly from Proposition 2.7.3. Indeed, by Lemma 2.5.3, it suffices to show that

$$\sum_{q_0\sim Q_0}\sum_{q_1\sim Q_1}\sum_{\substack{q_2\sim Q_2\\(q_1q_2,a)=1}}\eta_{q_0}\lambda_{q_1}\nu_{q_2}\sum_{\substack{p_1,\dots,p_J\\p_i\in\mathcal{I}_i\forall i}}S_{p_1\cdots p_J}\ll_B\frac{x}{(\log x)^B}.$$

for every $B>0$ and every choice of intervals $\mathcal{I}_1,\dots,\mathcal{I}_J$ with $\mathcal{I}_j\subset[P_j,2P_j]$. By reordering the indices, we may assume $\mathcal{J}=\{1,\dots,k\}$. We now take $N\asymp\prod_{j=1}^k P_j$, and $M\asymp\prod_{j=k+1}^J P_j$ and

$$\alpha_n:=\sum_{\substack{n=p_1\cdots p_k\\p_i\in\mathcal{I}_i}}1,\qquad\beta_m:=\sum_{\substack{m=p_{k+1}\cdots p_J\\p_i\in\mathcal{I}_i}}1,$$

Thus since $Q_0^{10}Q_1^{12}Q_2^7<x^{4-20\epsilon}$, we see that Proposition 2.7.3 gives the result in the range

$$Q_2x^\epsilon<N<\frac{x^{1-7\epsilon}}{Q_0^2Q_2}$$

Importantly, we also have the result in the mirrored range $Q_0^2Q_2x^{7\epsilon}<N<x^{1-7\epsilon}/Q_2$, by swapping the roles of $N,M$. And since $Q_0^2Q_2<x^{1/2-20\epsilon}$, together these cover the symmetric range

$$Q_2x^{7\epsilon}<N<\frac{x^{1-7\epsilon}}{Q_2}.$$

Finally since $Q_0Q_1Q_2^2<x^{1-20\epsilon}$, this covers the desired range

$$\frac{x^{1-\epsilon}}{Q_0Q_1Q_2}<N\le x^\epsilon Q_0Q_1Q_2.$$ □

*Proof of Proposition 2.4.1.* We note that $S_{p_1\cdots p_J}(p_J)$ is a weighted sum over integers $p_1\cdots p_Jn\sim x$ with $P^-(p_1\dots p_Jn)\ge x^{1/7+10\epsilon}$, and so with at most 6 prime factors. Expanding this into separate terms according to the exact number of prime factors, it suffices to show

$$\sum_{q_0\sim Q_0}\sum_{q_1\sim Q_1}\sum_{\substack{q_2\sim Q_2\\(q_0q_1q_2,a)=1}}\eta_{q_0}\lambda_{q_1}\nu_{q_2}\sum_{\substack{p_1,\dots,p_J\\p_i\sim P_i\forall i}}^*S_{p_1\cdots p_J}\ll_A\frac{x}{(\log x)^A}.$$

But this now follows from Lemma 2.8.1 with $(Q_0\cdot Q_1,\,Q_2)$ and Lemma 2.8.2 with $(Q_0,\,Q_1,\,Q_2)$, noting that

$$\begin{aligned}Q_0Q_1Q_2&=(Q_0^2Q_2)^{1/16}(Q_0Q_1Q_2^2)^{1/4}(Q_0^{10}Q_1^{12}Q_2^7)^{1/16}\\&<(x^{1/2})^{1/16}(x)^{1/4}(x^4)^{1/16}=x^{17/32}<x^{127/224-\epsilon}.\end{aligned}$$ □

## 2.9 Proof of Proposition 2.4.3 (4 prime factors)

In this section we prove Proposition 2.4.3. We recall the following estimate for triple convolutions.

**Proposition 2.9.1** (Estimate for triple convolutions). *Let $A, B_0 > 0$, $KLM \asymp x$, $\min(K, L, M) > x^{\epsilon}$, $a \neq 0$ and $x^{7/10-\epsilon} > Q > x^{1/2}(\log x)^{-A}$. Let $L, K$ satisfy*

$$\begin{aligned} Qx^{\epsilon} &< KL, \\ QK &< x^{1-2\epsilon}, \\ KL &< \frac{x^{153/224-10\epsilon}}{Q^{1/7}}, \\ KL^4 &< \frac{x^{57/32-10\epsilon}}{Q}. \end{aligned}$$

*Let $\eta_k, \lambda_\ell, \beta_m$ be complex sequences such that $|\eta_n|, |\lambda_n|, |\beta_n| \leq \tau(n)^{B_0}$ and such that $\eta_k$ satisfies the Siegel-Walfisz condition (1), and such that $\eta_k, \lambda_\ell$ be supported on integers with all prime factors bigger than $z_0$. Let*

$$\Delta_{\mathcal{B}}(q) := \sum_{k\sim K}\sum_{\ell\sim L}\sum_{m\sim M}\eta_k\lambda_\ell\alpha_m\Big(\mathbf{1}_{k\ell m\equiv a \pmod q} - \frac{\mathbf{1}_{(k\ell m,q)=1}}{\phi(q)}\Big).$$

*Then we have*

$$\sum_{\substack{q\sim Q\\(q,a)=1}} |\Delta_{\mathcal{B}}(q)| \ll_{A,B_0} \frac{x}{(\log x)^A}.$$

*Proof.* This is [101, Proposition 8.3]. □

*Proof of Proposition 2.4.3 from Proposition 2.4.1 and Proposition 2.9.1.* This follows just as in the proof of [101, Proposition 7.3], except for $Q_1$ replaced by $Q_0Q_1$, and trilinear weights $\eta_{q_0}\lambda_{q_1}\nu_{q_2}$ instead of absolute values. □

## 2.10 Proof of Proposition 2.4.2 (sieve asymptotics)

In this section we prove Proposition 2.4.2 using the following Fouvry-style and small divisor estimates.

**Proposition 2.10.1** (Fouvry-style estimate). *Let $A > 0$ and $C = C(A)$ be sufficiently large in terms of $A$. Assume that $N, M, Q_1, Q_2$ satisfy $NM \asymp x$ and*

$$x^{\epsilon}Q_1 < N, \tag{2.38}$$

$$N^6Q_1^4Q_2^8 < x^{5-\epsilon}, \tag{2.39}$$

$$Q_1Q_2^2 < x^{1-\epsilon}N. \tag{2.40}$$

*Let $\beta_m, \alpha_n$ be complex sequences such that $|\alpha_n|, |\beta_n| \leq \tau(n)^{B_0}$ and such that $\alpha_n$ satisfies the Siegel-Walfisz condition* (1) *and $\alpha_n$ is supported on $n$ with all prime factors bigger than $z_0 := x^{1/(\log\log x)^3}$. Let*

$$\Delta(q) := \sum_{m\sim M}\sum_{n\sim N}\alpha_n\beta_m\Big(\mathbf{1}_{mn\equiv a \pmod q} - \frac{\mathbf{1}_{(mn,q)=1}}{\phi(q)}\Big).$$

*Then we have*

$$\sum_{\substack{q_1\sim Q_1\\(q_1q_2,a)=1}}\sum_{q_2\sim Q_2}\lambda_{q_1}\nu_{q_2}\Delta(q_1q_2) \ll_{A,B_0} \frac{x}{(\log x)^A}.$$

*Proof.* This appears as [81, Proposition p.243], noting the condition therein $x^\epsilon(Q_1 + Q_1Q_2^2/x) < N < x^{5/6-\epsilon}/(Q_1Q_2^2)^{2/3}$ is equivalent to the one stated above (also see [101, Proposition 12.1]). □

**Proposition 2.10.2** (Small divisor estimate)**.** *Let $A > 0$ and $C = C(A)$ be sufficiently large in terms of $A$. Assume that $N, M, Q_1, Q_2$ satisfy $NM \asymp x$ and*

$$N^6Q_1^7Q_2^8 < x^{4-13\epsilon},$$
$$Q_1^2Q_2 < x^{1-7\epsilon}N.$$

*Let $\beta_m, \alpha_n$ be complex sequences such that $|\alpha_n|, |\beta_n| \leq \tau(n)^{B_0}$ and such that $\alpha_n$ satisfies the Siegel-Walfisz condition* (1) *and $\alpha_n$ is supported on $n$ with all prime factors bigger than $z_0 := x^{1/(\log\log x)^3}$. Let*

$$\Delta(q) := \sum_{m\sim M}\sum_{n\sim N}\alpha_n\beta_m\Big(\mathbf{1}_{mn\equiv a \pmod q} - \frac{\mathbf{1}_{(mn,q)=1}}{\phi(q)}\Big).$$

*Let $\lambda_q$ and $\nu_q$ be complex sequences supported on $P^-(q) \geq z_0$ with $|\lambda_q|, |\nu_q| \leq \tau(q)^{B_0}$. Then we have*

$$\sum_{\substack{q_1\sim Q_1\\(q_1q_2,a)=1}}\sum_{q_2\sim Q_2}|\Delta(q_1q_2)| \ll_{A,B_0} \frac{x}{(\log x)^A}.$$

*Proof.* This is [101, Proposition 12.2]. □

By combining the two results above, we now prove a small factor type II estimate for convolutions. This refines [101, Lemma 12.3].

**Lemma 2.10.3** (Small Factor Type II estimate for convolutions)**.** *Let $Q_1, Q_2$ satisfy*

$$Q_1Q_2 < x^{4/7-50\epsilon}, \qquad x^{4/7} < Q_1Q_2^2 < x^{1-100\epsilon}. \tag{2.41}$$

*Let $N$, $M$ be such that $NM \asymp x$ and*

$$x^{\varepsilon} < N < x^{1/7+10\epsilon}.$$

*Let $\beta_m, \alpha_n$ be complex sequences such that $|\alpha_n|, |\beta_n| \leq \tau(n)^{B_0}$ and such that $\alpha_n$ satisfies the Siegel-Walfisz condition* (1) *and $\alpha_n$ is supported on $n$ with all prime factors bigger than $z_0 := x^{1/(\log\log x)^3}$. Let*

$$\Delta(q) := \sum_{m\sim M}\sum_{n\sim N}\alpha_n\beta_m\Big(\mathbf{1}_{mn\equiv a \pmod q} - \frac{\mathbf{1}_{(mn,q)=1}}{\phi(q)}\Big).$$

*Let $\lambda_q$ and $\nu_q$ be complex sequences supported on $P^-(q) \geq z_0$ with $|\lambda_q|, |\nu_q| \leq \tau(q)^{B_0}$. Then we have*

$$\sum_{q_1\sim Q_1}\sum_{\substack{q_2\sim Q_2\\(q_1q_2,a)=1}}\lambda_{q_1}\nu_{q_2}\Delta(q_1q_2) \ll_{A,B_0} \frac{x}{(\log x)^A}.$$

*Proof of Lemma 2.10.3 from Proposition 2.10.1.* We may assume $Q_1Q_2 \geq x^{1/2-\epsilon}$, for otherwise the result is immediate from the Bombieri-Vinogradov Theorem. Proposition 2.10.2 (with $Q_1$ and $Q_2$ swapped) gives the result when

$$\frac{Q_1Q_2^2}{x^{1-7\epsilon}} < N < \frac{x^{2/3-3\varepsilon}}{Q_1^{8/6}Q_2^{7/6}}. \tag{2.42}$$

Note this lower bound implies $N > x^{\epsilon} \geq Q_1Q_2^2/x^{1-7\epsilon}$ by assumption. Also Proposition 2.10.1 gives the result when

$$\max\Big(\frac{Q_1Q_2^2}{x^{1-4\epsilon}}, Q_1x^{\epsilon}\Big) < N < \frac{x^{5/6-5\epsilon}}{(Q_1Q_2^2)^{2/3}}. \tag{2.43}$$

We see the lower bound max is attained by $Q_1x^{\epsilon}$. Thus the ranges in (2.43) and (2.42) overlap when

$$Q_1x^{\epsilon} > \frac{x^{2/3-3\varepsilon}}{Q_1^{8/6}Q_2^{7/6}}.$$

This is follows since $Q_1Q_2^2 > x^{4/7}$. Hence the result holds when $N > x^{\varepsilon}$ and $N$ is less than the min upper bound in (2.43). Moreover, this bound implies $N < x^{1/7+10\epsilon}$ as desired, provided

$$Q_1Q_2^2 < x^{29/28-50\varepsilon}.$$

This in turn follows from $Q_1Q_2^2 < x$, which completes the proof. □

Using Lemma 2.10.3, we deduce the following consequence. This refines [101, Proposition 101.1].

**Proposition 2.10.4** (Consequence of small factor type II estimate)**.** *Let $Q_1, Q_2$ satisfy*

$$Q_1Q_2 < x^{4/7-50\epsilon}, \qquad x^{4/7} < Q_1Q_2^2 < x^{1-100\epsilon}.$$

*Let $\alpha_d, \beta_e, \gamma_m$ complex sequences with $|\alpha_n|, |\beta_n|, |\gamma_n| \leq \tau(n)^{B_0}$ and such that $\gamma_m$ satisfies the Siegel-Walfisz condition* (1)*. Assume that*

$$D, E, P \in [x^\epsilon, x^{1/7+10\epsilon}]$$

*and let $MNDEP \asymp x$. Let $\lambda_q$ and $\nu_q$ be complex sequences supported on $P^-(q) \geq z_0$ with $|\lambda_q|, |\nu_q| \leq \tau(q)^{B_0}$. Then we have for every $A > 0$*

$$\sum_{q_1\sim Q_1} \sum_{\substack{q_2\sim Q_2 \\ (q_1q_2,a)=1}} \lambda_{q_1}\nu_{q_2} \sum_{d\sim D}\sum_{e\sim E}\sum_{p\sim P} \alpha_d\beta_e \sum_{m\sim M} \gamma_m \sum_{\substack{n\sim N \\ P^-(d),P^-(n)\geq p}}^{*} S_{nmpde} \ll_{A,B_0} \frac{x}{(\log x)^A}.$$

*Here $\sum^*$ indicates that the summation is restricted by $O(1)$ inequalities of the form $p^{\alpha_1}d^{\alpha_2}e^{\alpha_3}m^{\alpha_4}n^{\alpha_5} \leq B$. The implied constant may depend on all such exponents $\alpha_i$, but none of the quantities $B$.*

*Proof of Proposition 2.10.4 from Proposition 2.10.1.* This follows as in the proof of [101, Proposition 10.1], except with bilinear weights $\lambda_{q_1}\nu_{q_2}$ instead of absolute values. In this case, the small factor type II convolution estimate in Lemma 2.10.3 replaces that of [101, Lemma 12.3], and consequently the range $x^\epsilon < N < x^{1/7+\epsilon}$ in Lemma 2.10.3 leads to the result under the assumption $D, E, P \in [x^\epsilon, x^{1/7+10\epsilon}]$. □

We are now in a position to prove the sieve asymptotics in Proposition 2.4.2.

*Proof of Proposition 2.4.2 from Propositions 2.4.1 and 2.10.4.* The proof follows as in the proof of [101, Proposition 7.2], except for $Q_1$ replaced by $Q_0Q_1$, and trilinear weights $\eta_{q_0}\lambda_{q_1}\nu_{q_2}$ instead of absolute values. In this case, the small factor type II estimate in Proposition 2.11.2 replaces that of [101, Proposition 10.1], and consequently we use the cutoff $y_1 := x^\epsilon$ instead of $y_1 := x^{1-\epsilon}/(Q_1Q_2)^{15/8}$. □

## 2.11 Proof of Proposition 2.4.4 (3 prime factors)

In this section we prove Proposition 2.4.4. We begin by recalling the following triple divisor function estimate.

**Lemma 2.11.1.** *Let $x^{2\epsilon} \le N_1 \le N_2 \le N_3$ and $x^\epsilon \le M$ and $Q_1, Q_2 \ge 1$ be such that $Q_1Q_2 \le x^{1-\epsilon}$, $N_1N_2N_3M \asymp x$ and*

$$\frac{MQ_1^{5/2}Q_2^3}{x^{1-15\epsilon}} \le N_3 \le \frac{x^{2-15\epsilon}}{Q_1^3Q_2^2M}.$$

*Let $\alpha_m$ be a 1-bounded complex sequence, $\mathcal{I}_j \subseteq [N_j, 2N_j]$ an interval and*

$$\Delta_{\mathcal{K}}(q) := \sum_{m\sim M} \alpha_m \sum_{\substack{n_1\in\mathcal{I}_1\\ n_2\in\mathcal{I}_2\\ n_3\in\mathcal{I}_3\\ P^-(n_1n_2n_3)\ge z_0}} \Big(\mathbf{1}_{mn_1n_2n_3\equiv a \pmod q} - \frac{\mathbf{1}_{(mn_1n_2n_3,q)=1}}{\phi(q)}\Big).$$

*Then for every $A > 0$ we have*

$$\sum_{\substack{q_1\sim Q_1\\ (q_1,a)=1}} \sum_{\substack{q_2\sim Q_2\\ (q_2,a)=1}} \Big|\Delta_{\mathcal{K}}(q_1q_2)\Big| \ll_A \frac{x}{(\log x)^A}.$$

*Proof.* This is [101, Lemma 20.7]. □

We now establish a variant of the triple divisor estimate in [101, Proposition 11.1], under the weaker constraint $M = x^\varepsilon$

**Proposition 2.11.2** (Estimate for triple divisor function)**.** *Let $A > 0$. Let $Q, R$ satisfy*

$$Q^3R^2 < x^{11/7-30\epsilon}, \qquad QR < x^{8/15-30\epsilon}.$$

*Let $x^{3/7} \ge N_3 \ge N_2 \ge N_1 \ge x^\epsilon =: M$ satisfy $N_1N_2N_3M \asymp x$. Let $|\alpha_m| \le \tau(m)^{B_0}$ be a complex sequence, let $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$ be intervals with $\mathcal{I}_j \subseteq [N_j, 2N_j]$, and let*

$$\Delta_{\mathcal{K}}(q) := \sum_{m\sim M} \alpha_m \sum_{\substack{n_1\in\mathcal{I}_1\\ n_2\in\mathcal{I}_2\\ n_3\in\mathcal{I}_3\\ P^-(n_1n_2n_3)\ge z_0}} \Big(\mathbf{1}_{mn_1n_2n_3\equiv a \pmod q} - \frac{\mathbf{1}_{(mn_1n_2n_3,q)=1}}{\phi(q)}\Big).$$

*Then we have*

$$\sum_{\substack{q\sim Q\\ (q,a)=1}} \sum_{\substack{r\sim R\\ (r,a)=1}} |\Delta_{\mathcal{K}}(qr)| \ll_A \frac{x}{(\log x)^A}.$$

*Proof.* First we note that by Lemma 2.5.1 the set of $m$ with $|\alpha_m| \geq (\log x)^C$ has size $\ll x(\log x)^{O_{B_0}(1)-C}$, so by Lemma 2.5.2 these terms contribute negligibly if $C = C(A, B_0)$ is large enough. Thus, by dividing through by $(\log x)^C$ and considering $A + C$ in place of $A$, it suffices to show the result when $|\alpha_m| \leq 1$.

Since $N_1N_2N_3 \asymp x^{1-\epsilon}$ and $N_3 \geq N_2 \geq N_1$ we have $N_3 \gg x^{1/3-\epsilon}$. We first apply Lemma 2.11.1 with $M = x^\epsilon$ and the trivial factorization $Q_1 = QR$, $Q_2 = 1$. This gives the result provided

$$\frac{Q^{5/2}R^{5/2}}{x^{1-12\epsilon}} < N_3 < \frac{x^{2-16\epsilon}}{Q^3R^3}. \tag{2.44}$$

We now apply Lemma 2.11.1 the factorization $Q_1 = Q$, $Q_2 = R$, which gives the result provided

$$\frac{Q^{5/2}R^3}{x^{1-14\epsilon}} < N_3 < \frac{x^{2-16\epsilon}}{Q^3R^2}. \tag{2.45}$$

These ranges (2.44) and (2.45) overlap, provided

$$\frac{Q^{5/2}R^3}{x^{1-14\epsilon}} < \frac{x^{2-16\epsilon}}{Q^3R^3},$$

equivalently $Q^9R^8 < x^{6-30\epsilon}$. This follows easily since $(QR)^{9/6} < x^{4/5} < x^{1-30\epsilon}$ by assumption $QR < x^{8/15-30\epsilon}$. Thus the result holds in the combined range

$$\frac{Q^{5/2}R^{5/2}}{x^{1-12\epsilon}} < N_3 < \frac{x^{2-16\epsilon}}{Q^3R^2}. \tag{2.46}$$

This covers the stated range $x^{1/3-\epsilon} < N_3 < x^{3/7+\epsilon}$, since by assumption

$$Q^3R^2 < x^{11/7-30\epsilon}, \qquad QR < x^{8/15-30\epsilon}.$$

This gives the result. □

*Proof of Proposition 2.4.4 assuming Propositions 2.4.3 and 2.10.4.* This follows just as in the proof of [101, Proposition 7.4], except for $Q_1$ replaced by $Q_0Q_1$, and with trilinear weights $\eta_{q_0}\lambda_{q_1}\nu_{q_2}$ instead of absolute values. In this case, the triple divisor function estimate in Proposition 2.11.2 replaces that of [101, Proposition 11.1], and consequently we use the cutoff $y_1 := x^\epsilon$ instead of $y_1 := x^{1-\epsilon}/(Q_1Q_2)^{15/8}$. □

# Chapter 3

# Averages of the Möbius function on shifted primes

## 3.1 Introduction

Many central results in number theory may be formulated in terms of averages of the Möbius function $\mu$. Notably, the prime number theorem is equivalent to the statement $\sum_{n\le X}\mu(n) = o(X)$, and $\sum_{n\le X}\mu(n) = O(X^\theta)$ for all $\theta > \frac{1}{2}$ is equivalent to the Riemann hypothesis.

Clearly $\mu(p) = -1$ gives $\sum_{p\le X}\mu(p) = -\pi(X)$, but less is known about the Möbius function on shifted primes. It is a folklore conjecture that $\sum_{p\le X}\mu(p+h) = o(\pi(X))$ for any fixed shift $h > 0$. This appeared in print at least since Hildebrand [77, p.212], as well as Sarnak [128, Problem 5.2] and in Murty–Vatwani [107, (1.2)]). We answer an averaged version of this conjecture with quantitative bounds.

**Theorem 3.1.1.** *Suppose* $(\log X)^{300} < H < X$, *and write* $H = (\log X)^{\psi(X)}$. *Then*

$$\sum_{h\le H}\Big|\sum_{p\le X}\mu(p+h)\Big| \ll_\delta \frac{H\pi(X)}{\min\{\psi(X),(\log X)^{1/3-\delta}\}}. \tag{3.1}$$

An immediate consequence is that $\sum_{p<X}\mu(p+h)$ exhibits cancellation for all but $o(H)$ values of $h \le H = (\log X)^{\psi(X)}$, provided $\psi(X)\to\infty$. In Chapter 4, we shall focus specifically on achieving such qualitative cancellation with $H$ as small as possible.

*Remark* 3.1.2. The weaker qualitative cancellation (3.1) in the longer regime $H = X^\theta$ may be obtained more directly, using a recent Fourier uniformity result of Matomäki–Radziwiłł–Tao [99]. See Theorem 3.1.6 for details.

Theorem 3.1.1 is an illustrative example within a broad class of correlations that may be handled by the methods in this chapter, see Theorem 3.6.2 for the full technical result. Below we highlight an example of a general correlation of Möbius $\mu$ with an arbitrary function $G:\mathbb{N}\to\mathbb{C}$ (satisfying a mild $L^2$ condition).

**Theorem 3.1.3.** *For* $A > 5$, *suppose* $(\log X)^{300} < H < X$ *and write* $H = (\log X)^{\psi(X)}$. *If* $G:\mathbb{N}\to\mathbb{C}$ *satisfies* $\sum_{n\le X}|G(n)|^2 \ll X(\log X)^{A/20}$, *then we have*

$$\sum_{h\le H}\Big|\sum_{n\le X}\mu(n+h)G(n)\Big| \ll_{A,\delta} \frac{HX}{\min\{\psi(X),(\log X)^{1/3-\delta}\}}.$$

### 3.1.1 Higher correlations

The influential conjectures of Chowla [26] and Hardy–Littlewood [72] assert that for any fixed tuple $\mathcal{H} = \{h_1, .., h_k\}$ of distinct integers,

$$\begin{aligned}\sum_{n\leq X} \mu(n+h_1)\cdots\mu(n+h_k) &= o(X),\\ \sum_{n\leq X} \Lambda(n+h_1)\cdots\Lambda(n+h_k) &= \mathfrak{S}(\mathcal{H})X + o(X),\end{aligned}$$

for the singular series $\mathfrak{S}(\mathcal{H}) = \prod_p \frac{(1-\nu_p/p)}{(1-1/p)^k}$, where $\nu_p = \#\{h_1, .., h_k (\mathrm{mod}\ p)\}$. Both conjectures remain open for any $k \geq 2$.

We establish an average result for Hardy–Littlewood–Chowla correlations.

**Theorem 3.1.4.** *Suppose $(\log X)^{300} < H < X$, and write $H = (\log X)^{\psi(X)}$. Then for any $\delta > 0$, $m, k \geq 1$, and fixed tuple $\mathcal{A} = \{a_1, .., a_k\}$ of disinct integers, we have*

$$\sum_{h_1,..,h_m\leq H} \Big| \sum_{n\leq X} \prod_{j=1}^{m} \mu(n+h_j) \prod_{i=1}^{k} \Lambda(n+a_i) \Big| \ll_{\delta,m,\mathcal{A}} \frac{XH^m}{\min\{\psi(X)^m, (\log X)^{m/3-\delta}\}}. \tag{3.2}$$

It is worth emphasizing particular aspects of this result. First, (3.2) holds for an arbitrary fixed prime $k$-tuple. We must average over at least $m \geq 1$ copies of $\mu$ in order to obtain cancellation. Notably, the cancellation becomes quantitatively stronger for larger $m$, e.g. error savings $(\log X)^{m/3-\delta}$. For the case $m = 0, k = 2$, Matomäki–Radziwiłł–Tao [97] handled binary correlations $\sum_{n\leq X} \Lambda(n)\Lambda(n+h)$ on average with error savings $(\log X)^A$, though in the much larger regime $H \geq X^{8/33+\varepsilon}$.

In particular, the Chowla conjecture holds on average along the subsequence of primes.

**Corollary 3.1.4.1.** *Suppose $H < X$ and $\log H/\log_2 X \to \infty$ as $X \to \infty$. Then for any $m \geq 1$,*

$$\sum_{h_1,..,h_m\leq H} \Big| \sum_{p\leq X} \mu(p+h_1)\cdots\mu(p+h_m) \Big| = o_m\big(\pi(X)H^m\big). \tag{3.3}$$

Moreover, using Markov's inequality we may obtain qualitative cancellation for almost all shifts, with arbitrary log factor savings in the exceptional set.

**Corollary 3.1.4.2.** *Suppose $H < X$ and $\log H/\log_2 X \to \infty$ as $X \to \infty$. Then for any $A > 0$,*

$$\sum_{p\leq X} \mu(p+h_1)\cdots\mu(p+h_m) = o_m(\pi(X)),$$

*for all except $O_A(H^m(\log X)^{-A})$ shifts $(h_1, .., h_m) \in [1, H]^m$.*

These results build on earlier work of Matomäki–Radziwiłł–Tao [96], who established an average form of Chowla's conjecture,

$$\sum_{h_1,..,h_m\leq H}\Big|\sum_{n\leq X}\mu(n+h_1)\cdots\mu(n+h_m)\Big| \;=\; o_m(XH^m), \tag{3.4}$$

for any $H = H(X) \to \infty$ arbitrarily slowly. Whereas, our results require the faster growth $H = (\log X)^{\psi(X)}$ with $\psi(X) \to \infty$ arbitrarily slowly.

### 3.1.2 Correlations with divisor functions

Consider fixed integers $a \geq 1$, $k \geq l \geq 2$. The well studied correlation of two divisor functions $\tau_k, \tau_l$ is predicted to satisfy

$$\sum_{n\leq X}\tau_k(n+h)\tau_l(n) \;=\; C_{k,l,h}\cdot\big(X \;+\; o(X)\big)(\log X)^{k-l-2},$$

for a certain (explicit) constant $C_{k,l,h} > 0$. Recently, Matomäki–Radziwiłł–Tao [98] have shown the following averaged result, in the regime $H \geq (\log X)^{10000k\log k}$,

$$\sum_{h\leq H}\Big|\sum_{n\leq X}\tau_k(n+h)\tau_l(n) \;-\; C_{k,l,h}\cdot X(\log X)^{k+l-2}\Big| \;=\; o_k(HX(\log X)^{k+l-2}).$$

For higher correlations of divisor functions with Möbius, we obtain the following.

**Theorem 3.1.5.** *For any $j \geq 1$, $k_1,..,k_j \geq 2$, let $k = \sum_{i=1}^j k_i$ and take any fixed tuple $\mathcal{A} = \{a_1,..,a_j\}$ of distinct integers. If $H < X$ and $\log H/\log_2 X \to \infty$, then*

$$\sum_{h\leq H}\Big|\sum_{n\leq X}\mu(n+h)\prod_{i=1}^{j}\tau_{k_i}(n+a_i)\Big| \;=\; o_{k,\mathcal{A}}\big(HX(\log X)^{k-j}\big).$$

Again, we emphasize the need to average over the shift $h$ that inputs to Möbius $\mu(n+h)$, while $a_i$ may be fixed arbitrarily.

The main number-theoretic input is the classical Vinogradov–Korobov zero-free region

$$\Big\{\sigma+it \;:\; 1-\sigma < \frac{c}{\max\big\{\log q,\, \log(|t|+3)^{2/3}\log\log(|t|+3)^{1/3}\big\}}\Big\} \tag{3.5}$$

for $L(s,\chi)$, where $\chi$ is a Dirichlet character of modulus $q$, see [80, Corollary 8.28]. We shall apply this in the range of moduli $q \leq (\log X)^A$.

### 3.1.3 Beyond Möbius

We also consider general multiplicative functions $f : \mathbb{N} \to \mathbb{C}$, which do not pretend to be a character $f(n) \approx n^{it}\chi(n)$ for some $\chi \pmod q$. More precisely, we follow Granville and Soundararajan [68] and define the pretentious distance

$$\mathbb{D}(f, g; X) = \Big( \sum_{p \le X} \frac{1 - \Re(f(p)\overline{g(p)})}{p} \Big)^{1/2},$$

and the related quantity

$$M(f; X, Q) = \inf_{\substack{|t| \le X \\ \chi \,(q),\, q \le Q}} \mathbb{D}\big(f,\, n \mapsto n^{it}\chi(n); X\big)^2. \tag{3.6}$$

We may apply recent work of Matomäki–Radziwiłł–Tao [99] on Fourier uniformity, in order to more directly obtain (qualitative) cancellation for averages of non-pretentious multiplicative functions over shifted primes.

**Theorem 3.1.6.** *Given $\theta \in (0,1)$ let $H = X^\theta$. Given a multiplicative function $f : \mathbb{N} \to \mathbb{C}$ with $|f| \le 1$. There exists $\rho \in (0, \frac{1}{8})$ such that, if $M(f; X^2/H^{2-\rho}, Q) \to \infty$ as $X \to \infty$ for each fixed $Q > 1$, then*

$$\sum_{h \le H} \Big| \sum_{p \le X} f(p+h) \Big| \;=\; o_{\theta,\rho}\big(H\pi(X)\big).$$

In particular, $f = \mu$ does not pretend to be a Dirichlet character, a fact equivalent to the prime number theorem in arithmetic progressions. Indeed,

$$M(\mu; X, Q) \ge \inf_{\substack{|t| \le X \\ \chi \,(q),\, q \le Q}} \sum_{e^{(\log X)^{2/3+\varepsilon}} \le p \le X} \frac{1 + \Re\chi(p)p^{it}}{p} \;\ge\; \Big(\frac{1}{3} - \varepsilon\Big) \log\log X + O(1), \tag{3.7}$$

where the latter inequality follows from the zero-free region (3.5), see [105, §9.5].

### 3.1.4 Overview of the proof of Theorem 3.1.1

We now indicate the general form of the proof. We pursue a variation on the approach of Matomäki–Radziwiłł–Tao [96]. Namely, we first restrict (3.1) to 'typical' terms $\mu(n)$ for $n = p + h \in \mathcal{S}$ that have prime factors lying in certain prescribed intervals $[P_1, Q_1], [P_2, Q_2]$. The terms with $n \notin \mathcal{S}$ are sparse, and thus may be shown to contribute negligibly by standard sieve estimates. (For higher correlations, one may also use sieve estimates, along with work of Henriot [76] to handle a general class of functions with 'moderate growth' that are 'amenable to sieves.')

Once reduced to numbers with 'typical factorization,' we decouple the short interval correlation between Möbius and the indicator for the primes, using a Fourier identity and applying Cauchy–Schwarz (Lemma 3.2.1). This essentially yields a bound of $\pi(X) \ll X/\log X$ times a Fourier-type integral for $\mu$,

$$\sup_{\alpha} \int_0^X \Big| \sum_{\substack{x \leq n \leq x+H \\ n \in \mathcal{S}}} \mu(n) e(n\alpha) \Big| \mathrm{d}x.$$

This decoupling step is a gambit. It has the advantage of only needing to consider $\mu$ on its own, but loses a factor of $\log X$ from the density of the primes. To make this gambit worthwhile, we must recover over a factor of $\log X$ savings in the above Fourier integral for $\mu$. However, Matomäki–Radziwiłł–Tao [96, Theorem 2.3] bound the above integral with roughly $(\log X)^{\frac{1}{500}}$ savings (though their bound holds for any non-pretentious multiplicative function $g$.) Therefore we must refine the argument in the special case of $g = \mu$ to win back over a full factor of $\log X$. We note this task is impossible unless $H$ is larger than a power of log.

We accomplish this task in the 'key Fourier estimate' (Theorem 3.2.2), which bounds the above integral with $(\log X)^A$ savings for any $A > 0$ (though $\mathcal{S}$ will implicitly depend on $A$). As with [96], this bound is proven by reducing to the analogous estimate with the completely multiplicative Liouville function $\lambda$, and splitting up $\alpha \in [0, 1]$ into major and minor arcs.

The main technical innovation here comes from the major arcs (Proposition 3.5.1), essentially saving a factor $(\log X)^A$ in the mean values of 'typical' Dirichlet polynomials of the form

$$\sum_{\substack{X \leq n \leq 2X \\ n \in \mathcal{S}}} \frac{\lambda(n)\chi(n)}{n^s}$$

for a character $\chi$ of modulus $q \leq (\log X)^A$ in the Siegel–Walfisz range. This refines the seminal work of Matomäki–Radziwiłł [93], who obtained a fractional power of log savings for the corresponding mean values. However, Matomäki–Radziwiłł's results apply to the general setting of (non-pretentious) multiplicative functions and appeal to Halász's theorem, which offers small savings. By contrast, our specialization to the Möbius function affords us the full strength of Vinogradov–Korobov estimates (Lemma 3.4.5).

The Matomäki–Radziwiłł method saves roughly a fractional power of $P_1$ in the Dirichlet mean value when $Q_1 \approx H$. So in order to recover from our initial gambit, we are prompted to choose $P_1 = (\log X)^C$ for some large $C > 0$. Then by a standard

sieve bound the size of $\overline{\mathcal{S}}$ is morally $O(\frac{\log P_1}{\log Q_1}) = O_C(\frac{\log\log X}{\log H})$. This highlights the need for our assumption $\log H/\log\log X \to \infty$.

We remark that the Matomäki–Radziwiłł method requires two intervals $[P_1, Q_1]$, $[P_2, Q_2]$ (that define $\mathcal{S}$) in order to handle 'typical' Dirichlet polynomials in the regime $H = (\log X)^{\psi(X)}$ for $\psi(X) \to \infty$. Note in general [93] the slower $H \to \infty$ the more intervals we require (though by a neat short argument [95], only one interval is needed in the regime $H = X^\theta$ for $\theta > 0$).

# Notation

We recall the key arithmetic functions. The Möbius function $\mu$ is defined multiplicatively from primes $p$ by $\mu(p) = -1$ and $\mu(p^k) = 0$ for $k \geq 2$. Similarly the Liouville function $\lambda$ is defined *completely* multiplicatively by $\lambda(p) = -1$. The von Mangoldt function is given by $\Lambda(n) = \log p$ on prime powers $n = p^k$, and zero otherwise. For $l \geq 2$ the $l$th divisor function is $\tau_l(n) = \sum_{n=n_1\cdots n_l} 1$

We use standard asymptotic notation: $X \ll Y$ and $X = O(Y)$ both mean $|X| \leq CY$ for some some absolute constant $C$, and $X \asymp Y$ means $X \ll Y \ll X$. If $x$ is a parameter tending to infinity, $X = o(Y)$ means that $|X| \leq c(x)Y$ for some quantity $c(x)$ that tends to zero as $x \to \infty$. Let $\log_k X = \log_{k-1}(\log X)$ denote the $k$th-iterated logarithm.

Unless otherwise specified, all sums range over the integers, except for sums over the variable $p$ (or $p_1$, $p_2$,..) which are understood to be over the set of primes $\mathbb{P}$. Let $e(x) := e^{2\pi i x}$.

We use $\mathbf{1}_S$ to denote the indicator of a predicate $S$, so $\mathbf{1}_S = 1$ if $S$ is true and $\mathbf{1}_S = 0$ if $S$ is false. When $\mathcal{S}$ is a set, we write $\mathbf{1}_{\mathcal{S}}(n) = \mathbf{1}_{n\in\mathcal{S}}$ as the indicator function of $\mathcal{S}$. Also let $\mathbf{1}_{\mathcal{S}} f$ denote the function $n \mapsto \mathbf{1}_{\mathcal{S}}(n) f(n)$.

## 3.2 Initial reductions

In this section, we shall make some initial reductions along the lines of Matomäki–Radziwiłł–Tao [96, Theorem 2.3]. We shall restrict our attention to numbers $n = p+h$ with prime factors in prescribed intervals $[P_1, Q_1], [P_1, Q_1]$ (as defined in (3.11)). The exceptional $n$ are rare and thus contribute negligibly, as shown by a standard sieve bound. From here, we shall decouple the correlation between $\mu$ and the indicator for the primes, and reduce the problem to the key Fourier estimate for $\mu$, as in Theorem 3.2.2. For technical convenience, we shall further reduce to $\lambda$.

To this, we begin with a general Fourier-type bound to decouple correlations of arbitrary functions.

**Lemma 3.2.1** (Fourier bound)**.** *Given $f, g : \mathbb{N} \to \mathbb{C}$, let $F(X) := \sum_{n\le X} |f(n)|^2$. Then*

$$\sum_{|h|\le H} \Big| \sum_{n\le X} f(n)\, g(n+h) \Big|^2 \;\ll\; F(X+2H) \cdot \sup_{\alpha} \int_0^X \Big| \sum_{x\le n\le x+2H} g(n) e(n\alpha) \Big| \mathrm{d}x. \quad (3.8)$$

*Proof.* First, the lefthand side of (3.8) is

$$\sum_{|h|\le H} \Big| \sum_{n\le X} f(n)\, g(n+h) \Big|^2 \ll\; H^{-2} \sum_{|h|\le 2H} (2H-|h|)^2 \Big| \sum_{n\le X} f(n)\, g(n+h) \Big|^2 \;=:\; H^{-2}\, \Sigma. \quad (3.9)$$

Expanding the square in $\Sigma$ and letting $h = m - n = m' - n'$, we have

$$\begin{aligned} \Sigma &= \sum_{|h|\le 2H} (2H-|h|)^2 \sum_{n,n'\le X} f(n)\overline{f}(n') g(n+h)\overline{g}(n'+h) \\ &= \sum_{n,n'\le X} \sum_{m,m'} f(n)\overline{f}(n') g(m)\overline{g}(m') \mathbf{1}_{m-n=m'-n'} \\ &\qquad\qquad \times \Big( \int_0^X \mathbf{1}_{x\le n, m\le x+2H} \mathrm{d}x \Big) \Big( \int_0^X \mathbf{1}_{x'\le n', m'\le x'+2H} \mathrm{d}x' \Big). \end{aligned}$$

Then orthogonality $\mathbf{1}_{m-n=m'-n'} = \int_0^1 e((m-n-m'+n')\alpha)\mathrm{d}\alpha$ gives

$$\begin{aligned} \Sigma &= \int_0^1 \Big( \int_0^X \sum_{x\le n, m\le x+2H} f(n) g(m) e\big((m-n)\alpha\big) \mathrm{d}x \Big) \\ &\qquad \cdot \Big( \int_0^X \sum_{x'\le n', m'\le x'+2H} \overline{f}(n')\overline{g}(m') e\big((n'-m')\alpha\big) \mathrm{d}x' \Big) \mathrm{d}\alpha \\ &= \int_0^1 \Big| \int_0^X \sum_{x\le n, m\le x+2H} f(n) g(m) e\big((m-n)\alpha\big) \mathrm{d}x \Big|^2 \mathrm{d}\alpha. \end{aligned}$$

Using Cauchy–Schwarz, we bound $\Sigma$ as

$$\begin{aligned} \Sigma \;&\le\; \int_0^1 \int_0^X \Big| \sum_{x\le m\le x+2H} g(m) e(m\alpha) \Big|^2 \mathrm{d}x \cdot \int_0^X \Big| \sum_{y\le n\le y+2H} \overline{f}(n) e(n\alpha) \Big|^2 \mathrm{d}y \mathrm{d}\alpha \\ &\ll\; H \Big( \sup_{\alpha} \int_0^X \Big| \sum_{x\le m\le x+2H} g(m) e(m\alpha) \Big| \mathrm{d}x \Big) \int_0^1 \int_0^X \Big| \sum_{y\le n\le y+2H} \overline{f}(n) e(n\alpha) \Big|^2 \mathrm{d}y \mathrm{d}\alpha. \end{aligned} \quad (3.10)$$

Using $\int_0^1 e(n\alpha)\mathrm{d}\alpha = \mathbf{1}_{n=0}$ again, the second integral in (3.10) is

$$\begin{aligned}\int_0^1 \int_0^X \Big| \sum_{y\le n\le y+2H} \overline{f}(n)e(n\alpha)\Big|^2 \mathrm{d}y\mathrm{d}\alpha \\ &= \int_0^X \sum_{y\le n,n'\le y+2H} \overline{f}(n)f(n') \int_0^1 e\big((n-n')\alpha\big)\mathrm{d}\alpha\mathrm{d}y \\ &= \int_0^X \sum_{y\le n\le y+2H} |f(n)|^2\mathrm{d}y = \sum_{n\le X+2H} |f(n)|^2 \int_{n-2H}^n \mathrm{d}y \ll HF(X+2H).\end{aligned}$$

Hence plugging the bound (3.10) for $H^{-2}\Sigma$ back into (3.9) gives the result. □

Next we consider numbers with 'typical factorization.'

For $A, \delta \ge 0$, define $\psi$ via $H = (\log X)^{\psi(X)}$ and consider the intervals

$$\begin{aligned}[P_1, Q_1] &= [(\log X)^{33A}, (\log X)^{\psi(X)-4A}], \\ [P_2, Q_2] &= [\exp\big((\log X)^{2/3+\delta/2}\big), \exp\big((\log X)^{1-\delta/2}\big)],\end{aligned} \tag{3.11}$$

and define the 'typical factorization' set

$$\mathcal{S} \;=\; \mathcal{S}(X,H,A,\delta) := \{n \le X : \exists \text{ prime factors } p_1, p_2 \mid n \text{ with } p_j \in [P_j, Q_j]\}. \tag{3.12}$$

Using the Fourier bound, we shall reduce Theorem 3.1.1 to the following.

**Theorem 3.2.2** (Key Fourier estimate for $\mu$)**.** *Given any $A > 5$, $\delta > 0$, let $\mathcal{S} = \mathcal{S}(X,H,A,\delta)$ as in* (3.12)*. If $(\log X)^{40A} < H < X$, then*

$$\sup_\alpha \int_0^X \Big| \sum_{\substack{x\le n\le x+H \\ n\in\mathcal{S}}} \mu(n)e(n\alpha)\Big|\mathrm{d}x \;\ll_{A,\delta}\; \frac{HX}{(\log X)^{A/5}}.$$

*Proof of Theorem 3.1.1 from Theorem 3.2.2.* By a standard sieve upper bound [62, Theorem 7.4], for each $h \le H$, $j = 1, 2$ we have

$$\#\{\, p \le X \;:\; q \nmid p+h \;\forall q \in [P_j, Q_j]\} \;\ll\; \pi(X)\frac{\log P_j}{\log Q_j}\frac{h}{\phi(h)}. \tag{3.13}$$

Thus, recalling the choice of $[P_j, Q_j]$ in (3.11), the terms $p+h \notin \mathcal{S}$ trivially contribute to (3.1)

$$\begin{aligned}\sum_{h\le H}\Big|\sum_{\substack{p\le X \\ p+h\notin\mathcal{S}}}\mu(p+h)\Big| \;&\le \sum_{1\le j\le 2}\sum_{h\le H}\sum_{\substack{p\le X \\ q\nmid p+h\;\forall q\in[P_j,Q_j]}} 1 \\ &\ll_\delta\; \pi(X)\Big(\frac{A}{\psi(X)} + (\log X)^{\delta-1/3}\Big)\sum_{h\le H}\frac{h}{\phi(h)} \\ &\ll_{A,\delta}\; \frac{H\pi(X)}{\min\{\psi(X), (\log X)^{1/3-\delta}\}}.\end{aligned} \tag{3.14}$$

On the other hand for $p+h \in \mathcal{S}$, Lemma 3.2.1 with $f(n) = \mathbf{1}_{\mathbb{P}}(n)$, $g(n) = \mathbf{1}_{\mathcal{S}}\mu(n)$ gives

$$\sum_{h\le H}\Big|\sum_{\substack{p\le X\\ p+h\in\mathcal{S}}}\mu(p+h)\Big|^2 \ll \pi(X+2H)\cdot\sup_{\alpha}\int_0^X\Big|\sum_{\substack{x\le n\le x+2H\\ n\in\mathcal{S}}}\mu(n)e(n\alpha)\Big|\mathrm{d}x$$
$$\ll_{A,\delta} \frac{HX^2}{(\log X)^{A/5+1}},$$

assuming Theorem 3.2.2. Thus by Cauchy–Schwarz we obtain

$$\sum_{h\le H}\Big|\sum_{\substack{p\le X\\ p+h\in\mathcal{S}}}\mu(p+h)\Big| \ll \Big(H\sum_{h\le H}\Big|\sum_{\substack{p\le X\\ p+h\in\mathcal{S}}}\mu(p+h)\Big|^2\Big)^{1/2}$$
$$\ll_{A,\delta} \frac{HX}{(\log X)^{A/10+1/2}} \ll \frac{H\pi(X)}{(\log X)^{1/2}}. \tag{3.15}$$

choosing $A = 10$. Hence (3.14) and (3.15) give Theorem 3.1.1. □

Let $W = (\log X)^A$. Recall Theorem 3.2.2 asserts that $M_H(X;X) \ll_{A,\delta} XH/W^{1/5}$ for

$$M_H(X;Y) := \sup_{\alpha}\int_0^X\Big|\sum_{\substack{x\le n\le x+H\\ n\in\mathcal{S}(Y,A,\delta)}}\mu(n)e(n\alpha)\Big|\mathrm{d}x.$$

We first note, that, for technical convenience, it suffices to establish $M_{H_0}(2X;X) \ll XH_0/W^{1/5}$ with $H_0 := \min\{H, \exp((\log X)^{2/3})\}$. Indeed, if $H \in [H_0, X]$ then by the triangle inequality

$$M_H(X;X) \le \sum_{k\le\lceil H/H_0\rceil} M_{H_0}(X+kH_0;X)$$
$$\ll \sum_{k\le\lceil H/H_0\rceil}\frac{(X+kH_0)H_0}{W^{1/5}} \ll \frac{H}{H_0}\cdot\frac{XH_0}{W^{1/5}} = \frac{XH}{W^{1/5}}$$

as desired. Hence we may assume $H \le \exp((\log X)^{2/3})$ hereafter. This reduction is not strictly necessary, but will simplify the argument. For example, in this case $Q_1 < P_2$ so the intervals $[P_j, Q_j]$ are disjoint.

Consider the 'refined typical factorization' sets $\mathcal{S}_d = \{n/d : d \mid n \in \mathcal{S}\}$ for $d < P_1$, that is,

$$\mathcal{S}_d \;=\; \mathcal{S}_d(X,H,A,\delta) = \{m \le X/d : \exists \text{ prime factors } p_1, p_2 \mid m \text{ with } p_j \in [P_j, Q_j]\}. \tag{3.16}$$

So far we have reduced Theorem 3.1.1 to Theorem 3.2.2 for $\mu$. We now reduce further to the analogous estimate for its completely multiplicative counterpart $\lambda$.

**Proposition 3.2.3** (Key Fourier estimate for $\lambda$)**.** *Given any $A > 5$, $\delta > 0$, $H = (\log X)^{\psi(X)}$ with $40A \le \psi(X) \le (\log X)^{2/3}$. For $d \le W = (\log X)^A$ and $\mathcal{S}_d = \mathcal{S}_d(X, H, A, \delta)$ as in* (3.16)*, we have*

$$\sup_{\alpha} \int_0^{2X} \Big| \sum_{\substack{x \le nd \le x+H \\ n \in \mathcal{S}_d}} \lambda(n) e(\alpha n) \Big| \mathrm{d}x \ \ll_{A,\delta} \ \frac{HX}{d^{3/4} W^{1/5}}.$$

*Proof of Theorem 3.2.2 from Proposition 3.2.3.* By Möbius inversion, we have $\mu = \lambda * h$ for $h = \mu * (\mu\lambda)$, where $*$ denotes Dirichlet convolution. That is, $h(d^2) = \mu(d)$ for squarefree $d$, and zero otherwise. Thus we may write

$$\sum_{x \le n \le x+H} \mathbf{1}_{\mathcal{S}}(n) \mu(n) e(n\alpha) = \sum_{d \ge 1} h(d) \sum_{x \le md \le x+H} \mathbf{1}_{\mathcal{S}}(md) \lambda(m) e(md\alpha),$$

and so the triangle inequality gives

$$\begin{aligned} &\int_0^{2X} \Big| \sum_{x \le n \le x+H} \mathbf{1}_{\mathcal{S}}(n) \mu(n) e(n\alpha) \Big| \mathrm{d}x \\ &\qquad \le \sum_{d \ge 1} |h(d)| \int_0^{2X} \Big| \sum_{x \le md \le x+H} \mathbf{1}_{\mathcal{S}}(md) \lambda(m) e(md\alpha) \Big| \mathrm{d}x. \end{aligned} \tag{3.17}$$

Note, using the trivial bound and swapping the order of summation and integration, the contribution of $d > W$ to (3.17) is

$$\ll \sum_{W < d \le 2X} |h(d)| \sum_{md \le X+H} H \ \ll \ \frac{HX}{W^{1/4}} \sum_{d \ge 1} \frac{|h(d)|}{d^{3/4}} \ \ll \ \frac{HX}{W^{1/4}} \sum_{d \ge 1} d^{-3/2} \ \ll \ \frac{HX}{W^{1/4}}, \tag{3.18}$$

since $|h| \le 1$ is supported on squares.

On the other hand the contribution of $d \le W$ to (3.17) is

$$\le \sum_{d \ge 1} |h(d)| \int_0^{2X} \Big| \sum_{x \le md \le x+H} \mathbf{1}_{\mathcal{S}_d}(m) \lambda(m) e(md\alpha) \Big| \mathrm{d}x \ll \frac{HX}{W^{1/5}} \sum_{d \le W} \frac{|h(d)|}{d^{3/4}} \ \ll \ \frac{HX}{W^{1/5}} \tag{3.19}$$

assuming Proposition 3.2.3, and noting $\mathbf{1}_{\mathcal{S}}(md) = \mathbf{1}_{\mathcal{S}_d}(m)$ since $d \le W < P_1$. Together (3.18) and (3.19) give Theorem 3.2.2. □

*Remark* 3.2.4. In [96, Theorem 2.3], Matomäki–Radziwiłł–Tao bound the key Fourier integral for $\lambda$ with roughly $(\log X)^{\frac{1}{500}}$ savings over the trivial bound $HX$. Indeed their bound holds in general for any 1-bounded, non-pretentious multiplicative function $g$. However, the main point here is that the specific choice of $g = \lambda$ (and by extension, $\mu$) will allow us to extract a savings of any higher power of $\log x$.

## 3.3 Key Fourier estimate

In this section, we establish Proposition 3.2.3 by the circle method, following the argument in [96, Proposition 2.4].

Take $\alpha \in [0,1]$. By Dirichlet's approximation theorem there exists $\frac{a}{q} \in \mathbb{Q}$ with $(a,q)=1$ and $1 \le q \le Q_1$ for which

$$\Big|\alpha - \frac{a}{q}\Big| \ \le \ \frac{1}{qQ_1}.$$

So we may split $[0,1]$ into major arcs $\mathfrak{M}$ and minor arcs $\mathfrak{m}$, according to the size of denominator $q$ compared to $W$,

$$\mathfrak{M} = \bigcup_{q \le W} \mathfrak{M}(q) \quad \text{and} \quad \mathfrak{m} = [0,1] \setminus \mathfrak{M},$$

where $\mathfrak{M}(q) = \bigcup_{(a,q)=1}\{\alpha : |\alpha - a/q| \le 1/qQ_1\}$. Recall the definitions (3.11), (3.16),

$$\begin{aligned}
[P_1, Q_1] &= [(\log X)^{33A}, (\log X)^{\psi(X)-4A}],\\
[P_2, Q_2] &= [\exp\big((\log X)^{2/3+\delta/2}\big), \exp\big((\log X)^{1-\delta/2}\big)],\\
\mathcal{S}_d(X,H,A,\delta) &= \{m \le X/d : \exists p_1, p_2 \mid m \text{ with } p_j \in [P_j, Q_j]\}.
\end{aligned}$$

We shall obtain Proposition 3.2.3 from the following results.

**Proposition 3.3.1** (Key minor arc estimate)**.** *Given any $A > 5$, $H = (\log X)^{\psi(X)}$ with $40A \le \psi(X) \le (\log X)^{2/3}$, let $d \le W = (\log X)^A$ and $\mathcal{S}_d = \mathcal{S}_d(X, A, 0)$ as in* (3.16)*. Then for any completely multiplicative $g : \mathbb{N} \to \mathbb{C}$ with $|g| \le 1$, we have*

$$\sup_{\alpha \in \mathfrak{M}} \int_0^{2X} \Big| \sum_{\substack{x \le nd \le x+H \\ n \in \mathcal{S}_d}} g(n)\, e(n\alpha) \Big| \mathrm{d}x \ \ll_A \ \frac{HX}{d^{3/4} W^{1/5}}.$$

**Proposition 3.3.2** (Key major arc estimate for $\lambda$)**.** *Given any $A > 5$, $\delta > 0$, $H = (\log X)^{\psi(X)}$ with $40A \le \psi(X) \le (\log X)^{2/3}$, let $d \le W = (\log X)^A$ and $\mathcal{S}_d = \mathcal{S}_d(X, H, A, \delta)$ as in* (3.16)*. Then we have*

$$\sup_{\alpha \in \mathfrak{M}} \int_0^{2X} \Big| \sum_{\substack{x \le nd \le x+H \\ n \in \mathcal{S}_d}} \lambda(n)\, e(n\alpha) \Big| \mathrm{d}x \ \ll_{A,\delta} \ \frac{HX}{dW}.$$

We remark that the bounds in the minor arc hold for any bounded multiplicative function, whereas in the major arc the specific choice of $\lambda$ is needed.

### 3.3.1 Minor arc

In this subsection, we prove Proposition 3.3.1. Recall for $\alpha \in \mathfrak{M}$ in the minor arc, $|\alpha - a/q| < W^4/qH$ with $q \in [W, H/W^4]$. It suffices to show

$$I_{\mathfrak{m}} := \int_{\mathbb{R}} \theta(x) \sum_{x \le nd \le x+H} \mathbf{1}_{\mathcal{S}_d}(n) g(n)\, e(n\alpha) \mathrm{d}x \;\ll\; HX\big(\tfrac{\log\log X}{dW}\big)^{1/2} \psi(X), \tag{3.20}$$

uniformly for any $\alpha \in \mathfrak{m}$ and measurable $\theta : [0, 2X] \to \mathbb{C}$ with $|\theta(x)| \le 1$. Letting $\mathcal{P} = \{p : P_1 \le p \le Q_1\}$, by definition each $n \in \mathcal{S}_d$ has a prime factor in $\mathcal{P}$, so we use a variant of the Ramaré identity (see [62, §17.3]),

$$\mathbf{1}_{\mathcal{S}_d}(n) = \sum_{\substack{p \in \mathcal{P} \\ n = mp}} \frac{\mathbf{1}_{\mathcal{S}_d^{(1)}}(mp)}{\#\{q \in \mathcal{P} : q \mid m\} + \mathbf{1}_{p \nmid m}}, \tag{3.21}$$

where $\mathcal{S}_d^{(1)} = \{m \le X/d : \exists p \mid m,\, p \in [P_2, Q_2]\}$. As $g$ is completely multiplicative, we obtain

$$I_{\mathfrak{m}} = \sum_{p \in \mathcal{P}} \sum_{m} \frac{\mathbf{1}_{\mathcal{S}_d^{(1)}}(mp) g(m) g(p) e(mp\alpha)}{\#\{q \in \mathcal{P} : q \mid m\} + \mathbf{1}_{p \nmid m}} \int_{\mathbb{R}} \theta(x) \mathbf{1}_{x \le mpd \le x+H} \mathrm{d}x.$$

Next we split $\mathcal{P}$ into dyadic intervals $[P, 2P]$. It suffices to show for each $P \in [P_1, Q_1]$,

$$\sum_{\substack{p \in \mathcal{P} \\ P \le p \le 2P}} \sum_{m} \frac{\mathbf{1}_{\mathcal{S}_d^{(1)}}(mp) g(m) g(p) e(mp\alpha)}{\#\{q \in \mathcal{P} : q \mid m\} + \mathbf{1}_{p \nmid m}} \int_{\mathbb{R}} \theta(x) \mathbf{1}_{x \le mpd \le x+H} \mathrm{d}x \;\ll\; \frac{HX}{\log P} \big(\tfrac{\log\log X}{dW}\big)^{1/2}, \tag{3.22}$$

since then (3.20) will follow by (3.11) and the triangle inequality, using

$$\sum_{\substack{P_1 \ll P \ll Q_1 \\ P = 2^j}} \frac{1}{\log P} \ll \sum_{\log P_1 \ll j \ll \log Q_1} \frac{1}{j} \ll \log \frac{\log Q_1}{\log P_1} = \log \frac{\psi(X) - 4A}{33A} \ll_A \psi(X).$$

Fix $P$. We may replace $\mathbf{1}_{p \nmid m}$ with 1 in (3.22) at a cost of $O(HX/dP)$. Indeed, since the integral is $\int_{\mathbb{R}} \theta(x) \mathbf{1}_{x \le mpd \le x+H} \mathrm{d}x \ll H$, and $\mathbf{1}_{\mathcal{S}_d^{(1)}}(mp) = 0$ unless $m \le X/dP$, the cost of such substitution is

$$\ll \sum_{\substack{p \in \mathcal{P} \\ P \le p \le 2P}} \sum_{\substack{m \le X/dP \\ p \mid m}} H \;\ll\; P \frac{X}{dP^2} H = \frac{HX}{dP}.$$

Now the left hand side of (3.22) becomes

$$\begin{aligned}
&\sum_{m\in\mathcal{S}_d^{(1)}} \frac{g(m)}{\#\{q\in\mathcal{P}: q\mid m\}+1} \sum_{\substack{p\in\mathcal{P}\\ P\le p\le 2P}} g(p)e(mp\alpha)\mathbf{1}_{mpd\le X}\int_{\mathbb{R}}\theta(x)\mathbf{1}_{x\le mpd\le x+H}\mathrm{d}x\\
&\ll \sum_{m\le X/dP}\Big|\sum_{\substack{p\in\mathcal{P}\\ P\le p\le 2P}} g(p)e(mp\alpha)\mathbf{1}_{mpd\le X}\int_{\mathbb{R}}\theta(x)\mathbf{1}_{x\le mpd\le x+H}\mathrm{d}x\Big|\\
&\ll \Big(\frac{X}{dP}\Big)^{\frac12}\Big(\sum_{m\le X/dP}\Big|\sum_{\substack{p\in\mathcal{P}\\ P\le p\le 2P}} g(p)e(mp\alpha)\mathbf{1}_{mpd\le X}\int_{\mathbb{R}}\theta(x)\mathbf{1}_{x\le mpd\le x+H}\mathrm{d}x\Big|^2\Big)^{\frac12},
\end{aligned}$$

by the trivial bound and Cauchy–Schwarz. Hence for (3.22) it suffices to show

$$\sum_{m\le X/dP}\Big|\sum_{\substack{p\in\mathcal{P}\\ P\le p\le 2P}} g(p)e(mp\alpha)\mathbf{1}_{mpd\le X}\int_{\mathbb{R}}\theta(x)\mathbf{1}_{x\le mpd\le x+H}\mathrm{d}x\Big|^2 \ll \frac{H^2PX}{W}\frac{\log\log P}{(\log P)^2}. \tag{3.23}$$

We expand the left hand side of (3.23) and sum the resulting geometric series on $m$,

$$\begin{aligned}
&\sum_{p_1,p_2\in\mathcal{P}\cap[P,2P]}\int_{\mathbb{R}^2} g(p_1)\overline{g(p_2)}\theta(x_1)\overline{\theta(x_2)} \sum_{\substack{m\le X/dp_i\,\forall i\le 2\\ x_i\le mdp_i\le x_i+H}} e\big(m(p_1-p_2)\alpha\big)\mathrm{d}x_1\mathrm{d}x_2\\
&\ll HX\sum_{p_1,p_2\le 2P}\min\left(\frac{H}{dP},\frac{1}{\|(p_1-p_2)\alpha\|}\right),
\end{aligned}$$

since for given $d, p_1, p_2$, there are $O(X)$ choices for $x_1$ and $O(H)$ subsequent choices for $x_2$ since $x_2 = x_1(p_2/p_1) + O(H)$. Here $\|z\|$ denotes the distance of $z\in\mathbb{R}$ to the nearest integer.

Thus (3.23) reduces to showing

$$\sum_{p_1,p_2\le 2P}\min\left(\frac{H}{P},\frac{1}{\|(p_1-p_2)\alpha\|}\right) \ll \frac{HP}{W}\frac{\log\log P}{(\log P)^2}. \tag{3.24}$$

The difference of primes is $p_1 - p_2 \ll P$. Conversely, any integer $n \ll P$ may be written as $n = p_1 - p_2$ for $p_1, p_2 \le 2P$ in $\ll \frac{n}{\phi(n)}P(\log P)^{-2} \ll P\frac{\log\log P}{(\log P)^2}$ ways by a standard upper bound sieve, see [62, Proposition 6.22]. Hence for (3.24) it suffices to obtain

$$\sum_{1\le n\ll P}\min\left(\frac{H}{n},\frac{1}{\|n\alpha\|}\right) \ll \frac{H}{W} \qquad (\alpha\in\mathfrak{M}). \tag{3.25}$$

To show this, we use the following standard lemma, see [80, p.346].

**Lemma 3.3.3** (Vinogradov's lemma)**.** *Let $H, P > 1$, and take $\alpha \in [0,1]$ satisfying $|\alpha - a/q| \le 1/q^2$ for some $(a,q) = 1$. Then*

$$\sum_{1 \le n \le P} \min\left(\frac{H}{n}, \frac{1}{\|n\alpha\|}\right) \ll \frac{H}{q} + \frac{H}{P} + (P+q)\log q.$$

Thus by Lemma 3.3.3, we have the bound $H/q + H/P + (P+q)\log q \ll H/W$ since $q \in [W, H/W^4]$ and $P \in [P_1, Q_1] = [W^{24}, H/W^4]$. This gives (3.25) as desired, and completes the proof in the minor arc case.

### 3.3.2 Major arc

In this subsection, we prove the key major arc estimate assuming the following mean value result for the (twisted) Liouville function.

**Proposition 3.3.4.** *Given $A > 5$, $\delta > 0$, $H = (\log X)^{\psi(X)}$ with $40A \le \psi(X) \le (\log X)^{2/3}$, let $q \le W = (\log X)^A$, $d < W^{33}$, $\chi \pmod q$, $h \in [H/W^5, H]$, and $\mathcal{S}_d = \mathcal{S}_d(X, H, A, \delta)$ as in (3.16). Then for all $Y \in [X/W^7, 2X]$, we have*

$$J_{d,h,q}(Y;\chi) := \int_Y^{2Y} \left|\frac{1}{h} \sum_{\substack{x \le m \le x+h \\ m \in \mathcal{S}_d}} \lambda(m)\chi(m)\right|^2 \mathrm{d}x \ll_{A,\delta} \frac{Y}{W^{10}}. \tag{3.26}$$

*Proof of Proposition 3.3.2 from Proposition 3.3.4.* To obtain the key major arc estimate we shall prove the stronger bound,

$$I_{\mathfrak{m}} := \sup_{\alpha \in \mathfrak{m}} \int_0^{2X} \left| \sum_{x \le nd \le x+H} \mathbf{1}_{\mathcal{S}_d}(n)\lambda(n)\, e(n\alpha)\right| \mathrm{d}x \ll \frac{HX}{dW}. \tag{3.27}$$

In the major arc recall $\alpha = \frac{a}{q} + \theta$ with $q \le W$ and $|\theta| \le \frac{W^4}{qH}$. By partial summation with $a_n = \mathbf{1}_{>x/d}(n)\mathbf{1}_{\mathcal{S}_d}(n)\lambda(n)\, e(na/q)$, and $A(t) = \sum_{n \le t} a_n$, we have

$$\begin{aligned}
&\sum_{x \le nd \le x+H} \mathbf{1}_{\mathcal{S}_d}(n)\lambda(n)\, e(n\alpha) \\
&= e(\tfrac{x+H}{d}\theta)A(\tfrac{x+H}{d}) - e(\tfrac{x}{d}\theta)A(\tfrac{x}{d}) - 2\pi i\theta \int_{x/d}^{(x+H)/d} e(t\theta)A(t)\mathrm{d}t \\
&\ll \left| \sum_{x \le nd \le x+H} \mathbf{1}_{\mathcal{S}_d}(n)\lambda(n)\, e(an/q)\right| + |\theta| \int_0^{H/d} \left| \sum_{x/d \le n \le x/d+h} \mathbf{1}_{\mathcal{S}_d}(n)\lambda(n)\, e(na/q)\right| \mathrm{d}h.
\end{aligned}$$

Thus taking the maximizing $h$ and integrating over $x \in [0, 2X]$, we obtain

$$I_{\mathfrak{m}} \ll I_{H/d} + |\theta|\frac{H}{d} \sup_{h \le H/d} I_h \ll I_{H/d} + \frac{W^4}{qd} \sup_{h \le H/d} I_h, \tag{3.28}$$

where

$$I_h := \int_0^{2X} \Big| \sum_{x/d \le n \le x/d+h} \mathbf{1}_{\mathcal{S}_d}(n)\lambda(n)\, e(an/q)\Big| \mathrm{d}x. \tag{3.29}$$

Then splitting into residues $b \pmod q$ gives

$$\begin{aligned} I_h &\le \sum_{b\,(q)} |e(ab/q)| \int_0^{2X} \Big| \sum_{\substack{x/d \le n \le x/d+h \\ n \equiv b\,(q)}} \mathbf{1}_{\mathcal{S}_d}(n)\lambda(n)\Big| \mathrm{d}x \\ &= \sum_{b\,(q)} \int_0^{2X} \Big| \sum_{\substack{x/d \le n \le x/d+h \\ n \equiv b\,(q)}} \mathbf{1}_{\mathcal{S}_d}(n)\lambda(n)\Big| \mathrm{d}x. \end{aligned}$$

Now suppose we have the bound

$$I_h \ll \frac{qhX}{W^5} \qquad \text{for} \qquad h \in [qH/W^5, H/d]. \tag{3.30}$$

Then, combining with the trivial bound $I_h \le hX$ when $h \le qH/W^5$, (3.28) becomes

$$\begin{aligned} I_{\mathfrak{m}} &\ll I_{H/d} + \frac{W^4}{qd}\Big( \sup_{qH/W^5 \le h \le H/d} I_h + \sup_{h \le qH/W^5} hX\Big) \\ &\ll \frac{qHX}{dW^5} + \frac{W^4}{qd}\Big(\frac{qHX}{dW^5} + \frac{qHX}{W^5}\Big) \ll \frac{HX}{dW}, \end{aligned}$$

for $q \le W$ in the major arc. Hence it suffices to show (3.30).

Now to bound $I_h$, we extract the gcd. Let $c := (b, q)$ so that $c \mid n$, and we let $b' = b/c$, $q' = q/c$, $h' = h/c$, $m = n/c$. Thus since $\lambda$ is completely multiplicative, we have

$$\begin{aligned} I_h &\le \sum_{c|q} |\lambda(c)| \sideset{}{^*}\sum_{b'\,(q')} \int_0^{2X} \Big| \sum_{\substack{x/cd \le m \le x/cd+h/c \\ m \equiv b'\,(q')}} \mathbf{1}_{\mathcal{S}_d}(cm)\lambda(m)\Big| \mathrm{d}x \\ &\le \sum_{c|q} cd \sideset{}{^*}\sum_{b'\,(q')} \int_0^{2X/cd} \Big| \sum_{\substack{y \le m \le y+h' \\ m \equiv b'\,(q')}} \mathbf{1}_{\mathcal{S}_{cd}}(m)\lambda(m)\Big| \mathrm{d}y, \end{aligned}$$

using the substitution $y = x/cd$, and noting $\mathbf{1}_{\mathcal{S}_d}(cm) = \mathbf{1}_{\mathcal{S}_{cd}}(m)$ since $c \le q < P_1$. Then recalling orthogonality of characters $\phi(q')\mathbf{1}_{m \equiv b'\,(q')} = \sum_{\chi\,(q')} \overline{\chi(b')}\chi(m)$, we obtain

$$\begin{aligned} I_h &\le \sum_{c|q} cd \sideset{}{^*}\sum_{b'\,(q')} \frac{1}{\phi(q')} \sum_{\chi\,(q')} |\overline{\chi(b')}| \int_0^{2X/cd} \Big| \sum_{y \le m \le y+h'} \mathbf{1}_{\mathcal{S}_{cd}}(m)\lambda(m)\chi(m)\Big| \mathrm{d}y \\ &\le \sum_{c|q} cd \sum_{\chi\,(q')} \int_0^{2X/cd} \Big| \sum_{y \le m \le y+h'} \mathbf{1}_{\mathcal{S}_{cd}}(m)\lambda(m)\chi(m)\Big| \mathrm{d}y. \end{aligned} \tag{3.31}$$

We may discard the contribution to (3.31) of the integral over $y \le X/dW^5$, since $h' = h/c$ and $q \le W$ imply an admissible cost

$$\ll \sum_{c|q} cd\phi(q') \frac{X}{dW^5} h' \ \le \ \frac{hX}{W^5} \sum_{q'|q} \phi(q') \ = \ \frac{qhX}{W^5}.$$

For the remaining $y \in [\frac{X}{dW^5}, \frac{X}{cd}]$ in (3.31), we split into dyadic intervals so that

$$I_h \ \le \sum_{c|q} cd \sum_{\chi\,(q')} \sum_{\substack{Y=2^j \\ \frac{X}{2dW^5} \le Y \le \frac{2X}{cd}}} \int_Y^{2Y} \Big| \sum_{y \le m \le y+h'} \mathbf{1}_{\mathcal{S}_{cd}}(m)\lambda(m)\chi(m) \Big| \mathrm{d}y \ + \ O\Big(\frac{qHX}{W^5}\Big). \tag{3.32}$$

By assumption, Proposition 3.3.4 implies $J_{cd,h',q'}(Y;\chi) \ll Y/W^{10}$, so Cauchy–Schwarz gives

$$\int_Y^{2Y} \Big| \sum_{y \le m \le y+h'} \mathbf{1}_{\mathcal{S}_{cd}}(m)\lambda(m)\chi(m) \Big| \mathrm{d}y \le h' \sqrt{Y \cdot J_{cd,h',q'}(Y;\chi)} \ \ll_A \ \frac{Yh'}{W^5}.$$

So plugging back into (3.32), we obtain

$$I_h \ \ll \ \sum_{c|q} cd\phi(q') \sum_{\substack{Y=2^j \\ \frac{X}{2W^4} \le Y \le \frac{2X}{cd}}} \frac{Yh'}{W^5} \ = \ \frac{hd}{W^5} \sum_{q'|q} \phi(q') \sum_{\substack{Y=2^j \\ \frac{X}{2dW^5} < Y < \frac{2X}{cd}}} Y \ \ll \ \frac{qhX}{W^5}.$$

This gives (3.30) as desired. □

## 3.4 Preparatory lemmas

We collect some standard lemmas on Dirichlet polynomials.

The first is the integral mean value theorem [80, Theorem 9.1].

**Lemma 3.4.1** (mean value)**.** *For $D(s) = \sum_{n \le N} a_n n^{-s}$, we have*

$$\int_{-T}^{T} |D(it)|^2 \mathrm{d}t = (T + O(N)) \sum_{n \le N} |a_n|^2.$$

One may discretize the mean value theorem by replacing the integral over $[-T, T]$ with a sum over a well-spaced set $\mathcal{W} \subset [-T, T]$.

**Definition 3.4.2.** A set $\mathcal{W} \subset \mathbb{R}$ is *well-spaced* if $|w - w'| \ge 1$ for all $w, w' \in \mathcal{W}$.

Next is the Halász-Montgomery inequality [80, Theorem 9.6], which offers an improvement to the (discretized) mean value theorem when the well-spaced set is 'sparse.'

**Lemma 3.4.3** (Halász-Montgomery)**.** *Given $D(s) = \sum_{n\le N} a_n n^{-s}$ and a well-spaced set $\mathcal{W} \subset [-T, T]$. Then*

$$\sum_{t\in\mathcal{W}} |D(it)|^2 \ \ll\ (N + |\mathcal{W}|\sqrt{T})(\log 2T) \sum_{n\le N} |a_n|^2.$$

We also need a bound on the size of well-spaced sets $\mathcal{W}$ in terms of the values of prime Dirichlet polynomials on $1 + i\mathcal{W}$.

**Lemma 3.4.4.** *Let $T, L, U \ge 2$. For a sequence $a_p \in \mathbb{C}$ indexed by primes, with $|a_p| \le 1$, and define the prime polynomial*

$$P(s) = \sum_{L\le p\le 2L} \frac{a_p}{p^s}.$$

*Suppose a well-spaced set $\mathcal{W} \subset [-T, T]$ satisfies $|P(1 + it)| \ge 1/U$ for all $t \in \mathcal{W}$. Then*

$$|\mathcal{W}| \ \ll\ U^2\, T^{2(\log U + \log\log T)/\log L}.$$

*Proof.* See [93, Lemma 8]. □

**Lemma 3.4.5.** *Given $A, K > 0$, $\theta > \frac{2}{3}$, and a Dirichlet character $\chi$ mod $q \le (\log X)^A$. Assume $\exp((\log X)^\theta) \le P \le Q \le X$, and let $P(s, \chi) = \sum_{P\le p\le Q} \chi(p)p^{-s}$. Then for any $|t| \le X$,*

$$|P(1 + it, \chi)| \ \ll_{A,K,\theta}\ \frac{\log X}{1 + |t|}\ +\ (\log X)^{-K}.$$

*Proof.* This follows as with [95, Lemma 2], except that the Vinogradov–Korobov zero-free region for $\zeta(s)$ is replaced by that of $L(s, \chi)$. □

We also use a Parseval-type bound. This shows that the average of a multiplicative function in almost all short intervals can be approximated by its average on a long interval, provided the mean square of the corresponding Dirichlet polynomial is small.

**Lemma 3.4.6** (Parseval bound)**.** *Given $T_0 \in [(\log X)^{1/15}, X^{1/4}]$, and take a sequence $(a_m)_{m=1}^\infty$ with $|a_m| \le 1$. Assume $1 \le h_1 \le h_2 \le X/T_0^3$. For $x \in [X, 2X]$, define*

$$S_{h_j}(x) = \sum_{x\le m\le x+h_j} a_m, \qquad \text{and} \qquad A(s) = \sum_{X\le m\le 4X} \frac{a_m}{m^s}.$$

*Then*

$$\frac{1}{X}\int_X^{2X} \left|\tfrac{1}{h_1}S_{h_1}(x) - \tfrac{1}{h_2}S_{h_2}(x)\right|^2 \mathrm{d}x \ \ll\ \frac{1}{T_0}\ +\ \int_{T_0}^{X/h_1} |A(1 + it)|^2 \mathrm{d}t + \max_{T\ge X/h_1} \frac{X/h_1}{T} \int_T^{2T} |A(1 + it)|^2 \mathrm{d}t.$$

*Proof.* This follows as in [93, Lemma 14] with $(\log X)^{1/15}$ replaced by general $T_0$. □

We have a general mean value of products, via the Ramaré identity [93, Lemma 12].

**Lemma 3.4.7.** *For $V, P, Q \geq 1$, denote $\mathcal{P} = [P, Q] \cap \mathbb{P}$. Let $a_m, b_m, c_p$ be bounded sequences for which $a_{mq} = b_m c_q$ when $q \nmid m$ and $q \in \mathcal{P}$. Let*

$$Q_{v,V}(s) = \sum_{\substack{q \in \mathcal{P} \\ e^{v/V} \leq q \leq e^{(v+1)/V}}} \frac{c_q}{q^s},$$

$$R_{v,V}(s) = \sum_{Xe^{-v/V} \leq m \leq 2Xe^{-v/V}} \frac{b_m}{m^s} \cdot \frac{1}{\#\{p \mid m : p \in \mathcal{P}\} + 1},$$

*and take a measurable set $\mathcal{T} \subset [-T, T]$. Then for $\mathcal{I} = [\lfloor V \log P \rfloor, V \log Q] \cap \mathbb{Z}$, we have*

$$\int_{\mathcal{T}} \Big| \sum_{X \leq n \leq 2X} \frac{a_n}{n^{1+it}} \Big|^2 \mathrm{d}t \ \ll \ V \log(\tfrac{Q}{P}) \sum_{v \in \mathcal{I}} \int_{\mathcal{T}} |Q_{v,V}(1+it)\, R_{v,V}(1+it)|^2 \mathrm{d}t$$
$$+ \ \Big(\frac{T}{X} + 1\Big)\Big(\frac{1}{V} + \frac{1}{P} + \sum_{\substack{X \leq n \leq 2X \\ p \nmid n \forall p \in \mathcal{P}}} \frac{|a_n|^2}{n}\Big)$$

In the next result we employ the Fundamental Lemma of the sieve, along with the Siegel–Walfisz theorem.

**Lemma 3.4.8.** *Given $A, K > 0$, $q \leq (\log x)^A$, Dirichlet character $\chi \pmod q$, and let $\mathcal{D} = \prod_{p \in \mathcal{P}} p$ for any set of primes $\mathcal{P} \subset (q, x^{1/\log\log x})$. Then*

$$\sum_{\substack{m \leq x \\ (m, \mathcal{D}) = 1}} \lambda(m)\chi(m) \ \ll_{A,K} \ \frac{x}{(\log x)^K}.$$

*Proof.* First partition the sum on $m$ by the values of $\lambda(m), \chi(m)$,

$$S_0 := \sum_{m \leq x} \mathbf{1}_{(m,\mathcal{D})=1} \lambda(m)\chi(m) = \sum_{b \,(q), \nu \in \{\pm 1\}} \nu\chi(b) \sum_{m \in \mathcal{A}^{(b,\nu)}} \mathbf{1}_{(m,\mathcal{D})=1} \tag{3.33}$$

for the set $\mathcal{A}^{(b,\nu)} = \{m \leq x : m \equiv b \,(q), \lambda(m) = \nu\}$.

Now it suffices to prove

$$\sum_{m \in \mathcal{A}^{(b,\nu)}} \mathbf{1}_{(m,\mathcal{D})=1} \ = \ \frac{x}{2q} \prod_{p \mid \mathcal{D}} \Big(1 - \frac{1}{p}\Big) \ + \ O_{A,K}\big(x(\log x)^{-K}\big) \tag{3.34}$$

uniformly in $b, \nu$, from which it will follow

$$S_0 = \frac{x}{2qd} \prod_{p|\mathcal{D}} \Big(1 - \frac{1}{p}\Big) \sum_{b\,(q), \nu \in \{\pm 1\}} \nu\chi(b) \;+\; O\big(x(\log x)^{-K}\big) \;\ll\; \frac{x}{(\log x)^K},$$

by pairing up terms $\nu = \pm 1$. This will give the lemma.

Now to show (3.34), write $\mathcal{A} = \mathcal{A}^{(b,\nu)}$. For $d \mid \mathcal{D}$ the set of multiples $\mathcal{A}_d = \{m \in \mathcal{A} : d \mid m\}$ has size

$$|\mathcal{A}_d| = \sum_{\substack{m \le x \\ d|m,\, m \equiv b\,(q)}} \mathbf{1}_{\lambda(m) = \nu} = \sum_{\substack{n \le x/d \\ nd \equiv b\,(q)}} \frac{\nu\lambda(nd) + 1}{2} = \frac{\nu\lambda(d)}{2} \sum_{\substack{n \le x/d \\ n \equiv bd^{-1}\,(q)}} \lambda(n) \;+\; \frac{x}{2qd} + O(q)$$

noting $(q, d) = 1 = (q, \mathcal{D})$. Moreover $\max_{c\,(q)} \big|\sum_{n \equiv c\,(q)}^{n \le y} \lambda(n)\big| \ll_{A,K} y(\log y)^{-2K}$ by Siegel–Walfisz, which is valid by the assumption $q \le (\log X)^A$. Thus

$$|\mathcal{A}_d| = \frac{x}{2qd} + R_d, \qquad \text{where} \quad |R_d| \ll \frac{x}{d} \log(x/d)^{-2K} + q. \tag{3.35}$$

Now for any $D > 1$ the indicator $\mathbf{1}_{(m,\mathcal{D})=1}$ is bounded between $\sum_{d<D,\, d|(m,\mathcal{D})} \lambda^{\pm}(d)$, for the standard linear sieve weights $\{\lambda^{\pm}(d)\}_{d<D}$, as in Section 1.3 of Chapter 1. Also see [62, Lemma 6.11]. Thus the desired sum in (3.34) is bounded in between

$$\sum_{\substack{d|\mathcal{D} \\ d<D}} \lambda^{-}(d)|\mathcal{A}_d| \;\le \sum_{m \in \mathcal{A}} \mathbf{1}_{(m,\mathcal{D})=1} \;\le\; \sum_{\substack{d|\mathcal{D} \\ d<D}} \lambda^{+}(d)|\mathcal{A}_d|. \tag{3.36}$$

Note by (3.35), the upper and lower bounds are given by

$$\sum_{\substack{d|\mathcal{D} \\ d<D}} \lambda^{\pm}(d)|\mathcal{A}_d| = \frac{x}{2q} \sum_{\substack{d|\mathcal{D} \\ d<D}} \frac{\lambda^{\pm}(d)}{d} \;+\; \sum_{\substack{d|\mathcal{D} \\ d<D}} \lambda^{\pm}(d) R_d.$$

Let $z = x^{1/\log_2 x}$ so that $\mathcal{P} \subset (q, z)$. Then choosing $D = z^s$ for $s = 2K \log_2 x / \log_3 x$, the above error is $\ll \sum_{d<D} |R_d| \ll x(\log x)^{-K}$ by (3.35). And by the Fundamental Lemma [62, Lemma 6.11], the main term is

$$\sum_{\substack{d|\mathcal{D} \\ d<D}} \frac{\lambda^{\pm}(d)}{d} = (1 + O(s^{-s})) \prod_{p|\mathcal{D}} \Big(1 - \frac{1}{p}\Big).$$

Hence (3.34) follows as claimed, noting $s^{-s} \ll (\log x)^{-K}$. □

## 3.5 Mean value of multiplicative functions

In this section we prove Proposition 3.3.4 based on the following mean value theorem for Dirichlet polynomials with typical factorization. This refines Matomäki–Radziwiłł [93, Proposition 12] in the case of $g = \lambda\chi$, by leveraging Vinogradov–Korobov type bounds.

**Proposition 3.5.1.** *Given any $A > 5$, $\delta > 0$, denote $B = 11A$, and write $W = (\log X)^A$, $H = (\log X)^{\psi(X)}$ with $40A \le \psi(X) \le (\log X)^{2/3}$. Take $q \le W$, $d < W^{33}$, a Dirichlet character $\chi \pmod q$, and let $\mathcal{S}_d = \mathcal{S}_d(X, H, A, \delta)$ as in (3.16) (as such $[P_1, Q_1] = [W^{33}, H/W^4]$). For any $Y \in [X^{1/2}, X^2]$, define*

$$G(s) = \sum_{\substack{Y \le n \le 2Y \\ n \in \mathcal{S}_d}} \frac{\lambda(n)\chi(n)}{n^s}.$$

*Then for any $T \in [Y^{1/2}, Y]$, we have*

$$\int_{(\log X)^{2B}}^{T} |G(1+it)|^2 \mathrm{d}t \ \ll_{A,\delta} \ \Big(\frac{Q_1 T}{Y} \ + \ 1\Big)(\log X)^{-B}. \tag{3.37}$$

*Proof of Proposition 3.3.4 from Proposition 3.5.1.* We shall prove

$$J := \int_{Y}^{2Y} \Big|\tfrac{1}{h_1} S_{h_1}(x)\Big|^2 \mathrm{d}x \ \ll_{A,\delta} \ \frac{Y}{(\log X)^{10A}} \tag{3.38}$$

for $Y \in [X/W^7, 2X]$, $h_1 = h \in [qH/W^5, H]$ and the sum

$$S_l(x) := \sum_{\substack{x \le m \le x+l \\ m \in \mathcal{S}_d}} \lambda(m)\chi(m).$$

First we claim $S_x(0) \ll x(\log x)^{-K}$ for all $K > 0$. To this, recall from (3.16) that each $m \in \mathcal{S}_d$ has prime factors $p_1, p_2 \mid m$ with $p_j \in [P_j, Q_j]$. So by inclusion-exclusion, the indicator of $\mathcal{S}_d$ is

$$\begin{aligned} \mathbf{1}_{\mathcal{S}_d}(m) &= (1 - \mathbf{1}_{(m,\mathcal{D}_1)=1})(1 - \mathbf{1}_{(m,\mathcal{D}_2)=1}) \\ &= \mathbf{1}_{(m,\mathcal{D}_0)=1} - \mathbf{1}_{(m,\mathcal{D}_1)=1} - \mathbf{1}_{(m,\mathcal{D}_2)=1} + \mathbf{1}_{(m,\mathcal{D}_3)=1}, \end{aligned}$$

where $\mathcal{D}_j = \prod_{p \in \mathcal{P}_j} p$ for the sets of primes $\mathcal{P}_0 = \emptyset$, $\mathcal{P}_1 = [P_1, Q_1]$, $\mathcal{P}_2 = [P_2, Q_2]$, $\mathcal{P}_3 = \mathcal{P}_1 \cup \mathcal{P}_2$. Hence applying Lemma 3.4.8 to each $\mathcal{D}_j$ gives

$$S_x(0) \le \sum_{j=0}^{3} \Big| \sum_{m \le x} \mathbf{1}_{(m,\mathcal{D}_j)=1} \lambda(m)\chi(m) \Big| \ \ll_{A,K} \ x(\log x)^{-K}. \tag{3.39}$$

In particular, letting $B = 11A$ we have

$$S_{h_2}(x) = S_{x+h_2}(0) - S_x(0) \ll_A \frac{x}{(\log x)^{7B}} \ll_A \frac{h_2}{(\log x)^B},$$

where $h_2 \asymp x(\log x)^{-6B}$, and so

$$J = \frac{1}{Y}\int_Y^{2Y} \left|\tfrac{1}{h_1}S_{h_1}(x)\right|^2 \mathrm{d}x \ll (\log X)^{-B} + \frac{1}{Y}\int_Y^{2Y} \left|\tfrac{1}{h_1}S_{h_1}(x) - \tfrac{1}{h_2}S_{h_2}(x)\right|^2 \mathrm{d}x. \tag{3.40}$$

Then Lemma 3.4.6 (Parseval) with $T_0 = (\log X)^{2B}$ and $h_2 = Y/T_0^3 = Y/(\log X)^{6B}$ gives

$$J \ll (\log X)^{-B} + \int_{T_0}^{Y/h_1} |G(1+it)|^2 \mathrm{d}t + \max_{T \geq Y/h_1} \frac{Y/h_1}{T}\int_T^{2T} |G(1+it)|^2 \mathrm{d}t. \tag{3.41}$$

Now for the latter integral over $[T, 2T]$, we apply the Lemma 3.4.1 (mean value) if $T \geq X/2$, and apply Proposition 3.5.1 if $T \in [Y/h_1, X/2]$. Doing so, (3.41) becomes

$$\begin{aligned} J &\ll \big(\tfrac{Y/h_1}{Y/Q_1} + 1\big)(\log X)^{-B} + \max_{T \geq X/2} \frac{Y/h_1}{T}(T/X + 1) \\ &\qquad + \max_{Y/h_1 \leq T \leq X/2} \frac{Y/h_1}{T}\big(\tfrac{Q_1 T}{Y} + 1\big)(\log X)^{-B} \\ &\ll (\frac{Q_1}{h_1} + 1)(\log X)^{-B} + \frac{Y}{h_1 X} \ll W(\log X)^{-B} = (\log X)^{A-B}. \end{aligned} \tag{3.42}$$

Here we used $Y \leq 2X$, $Q_1 = H/W^4$, and $h_1 \geq H/W^5 (\geq W^{35})$. Hence recalling $B = 11A$ gives Proposition 3.3.4 as claimed. □

### 3.5.1 Mean value of Dirichlet polynomials

In this subsection, we prove Proposition 3.5.1. Recall the definitions (3.11), (3.16),

$$\begin{aligned} [P_1, Q_1] &= [(\log X)^{33A}, (\log X)^{\psi(X)-4A}], \\ [P_2, Q_2] &= [\exp\big((\log X)^{2/3+\delta/2}\big), \exp\big((\log X)^{1-\delta/2}\big)], \\ \mathcal{S}_d(X, H, A, \delta) &= \{m \leq X/d : \exists p_1, p_2 \mid m \text{ with } p_j \in [P_j, Q_j]\}. \end{aligned}$$

Let $B = 11A$, and $\alpha = 1/5$. Let $V = P_1^{1/3} = (\log X)^B$ and define the prime polynomial

$$Q_{v,j}(s) := \sum_{\substack{P_j \leq p \leq Q_j \\ e^{v/V} \leq p \leq e^{(v+1)/V}}} \frac{\lambda(p)\chi(p)}{p^s}. \tag{3.43}$$

Note $Q_{v,j}(s) \neq 0$ only if $v \in \mathcal{I}_j := \{v \in \mathbb{Z} : P_j \leq e^{v/V} \leq Q_j\} = [\lfloor V \log P_j \rfloor, V \log Q_j]$.

We decompose $[T_0, T] = \mathcal{T}_1 \cup \mathcal{T}_2$ as a disjoint union, where $\mathcal{T}_2 = [0,1] \setminus \mathcal{T}_1$ and

$$\mathcal{T}_1 = \{t : \ |Q_{v,1}(1+it)| \leq e^{-\alpha v/V} \quad \forall v \in \mathcal{I}_1\}. \tag{3.44}$$

For $j = 1, 2$ denote by $\mathcal{S}_d^{(j)}$ the integers containing a prime factor in the interval $[P_i, Q_i]$ with $i \neq j$ and possibly, but not necessarily, with $i = j$. That is,

$$\mathcal{S}_d^{(j)} = \{m \leq X/d : \exists p \mid m \text{ with } p \in [P_i, Q_i] \text{ for } i \neq j\}.$$

Also define the polynomial

$$R_{v,j}(s) = \sum_{\substack{Ye^{-v/V} \leq m \leq 2Ye^{-v/V} \\ m \in \mathcal{S}_d^{(j)}}} \frac{\lambda(m)\chi(m)}{m^s} \cdot \frac{1}{\#\{p \mid m : P_j \leq p \leq Q_j\} + 1}. \tag{3.45}$$

Now Lemma 3.4.7 (Ramaré) applies with $V = V, P = P_j, Q = Q_j$, and $a_m = \lambda\,\chi\,\mathbf{1}_{\mathcal{S}}(m)$, $c_p = \lambda\,\chi(p)$, $b_m = \lambda\,\chi\,\mathbf{1}_{\mathcal{S}_j}(m)$, giving

$$\int_{\mathcal{T}_j} |G(1+it)|^2 \mathrm{d}t \ \ll \ V \log Q_j \sum_{v \in \mathcal{I}_j} \int_{\mathcal{T}_j} |Q_{v,j}(1+it)\, R_{v,j}(1+it)|^2 \mathrm{d}t \\ + \ \frac{1}{V} + \frac{1}{P_j} \ + \sum_{\substack{Y \leq n \leq 2Y \\ p \nmid n \forall p \in [P_j, Q_j]}} \frac{\mathbf{1}_{\mathcal{S}_d}(n)}{n}.$$

We crucially note the latter sum vanishes since each $n \in \mathcal{S}_d$ has a prime factor $p \in [P_j, Q_j]$. Summing over $j = 1, 2$, the second and third terms above contribute

$$\ll \sum_{1 \leq j \leq 2} \Big(\frac{1}{V} + \frac{1}{P_j}\Big) \ll \frac{1}{V} = (\log X)^{-B}.$$

Hence the desired integral is

$$\int_{T_0}^{T} |G(1+it)|^2 \mathrm{d}t = \int_{\mathcal{T}_1 \cup \mathcal{T}_2} |G(1+it)|^2 \mathrm{d}t \ \ll \ E_1 + E_2 + (\log X)^{-B}, \tag{3.46}$$

where

$$E_j = V \log Q_j \sum_{v \in \mathcal{I}_j} \int_{\mathcal{T}_j} |Q_{v,j}(1+it)\, R_{v,j}(1+it)|^2 \mathrm{d}t. \tag{3.47}$$

Hence it suffices to bound $E_1, E_2 \ll (Q_1 T/Y + 1)(\log X)^{-B}$.

**Bound for $E_1$:** By definition of $t \in \mathcal{T}_1$, we have $|Q_{v,1}(1+it)| \leq e^{-\alpha v/V}$ for all $v \in \mathcal{I}_1$, so

$$E_1 \ \ll V \log Q_1 \sum_{v \in \mathcal{I}_1} e^{-2\alpha v/V} \int_{\mathcal{T}_1} |R_{v,1}(1+it)|^2 \mathrm{d}t \ \ll V \log Q_1 \sum_{v \in \mathcal{I}_1} e^{-2\alpha v/V} \Big(\frac{T}{Y/e^{v/V}} + 1\Big)$$

by Lemma 3.4.1 (mean value). Summing the resulting geometric series gives

$$\begin{aligned} E_1 &\ll V\log Q_1 \frac{P_1^{-2\alpha}}{1-e^{-2\alpha/V}}\Big(\frac{Q_1T}{Y}+1\Big) \\ &\ll (\log X)P_1^{-2\alpha}\Big(\frac{Q_1T}{Y}+1\Big) \ll (\log X)^{-B}\Big(\frac{Q_1T}{Y}+1\Big), \end{aligned} \tag{3.48}$$

noting $V/(1-e^{-2\alpha/V}) = O(1)$ and $P_1^{-2\alpha} = (\log X)^{-6B/5}$.

**Bound for $E_2$:** We choose the maximizing $v \in \mathcal{I}_2$ for $E_2$. Thus since $|\mathcal{I}_2| < V\log Q_2$,

$$\begin{aligned} E_2 &= V\log Q_2 \sum_{v\in\mathcal{I}_2}\int_{\mathcal{T}_2} |Q_{v,2}\cdot R_{v,2}(1+it)|^2\mathrm{d}t \ll (V\log Q_2)^2\int_{\mathcal{T}_2}|Q_{v,2}\cdot R_{v,2}(1+it)|^2\mathrm{d}t \\ &\le (V\log Q_2)^2\sum_n \sup_{t_n\in[n,n+1]\cap\mathcal{T}_2}|Q_{v,2}\cdot R_{v,2}(1+it_n)|^2 \\ &\le 2(V\log Q_2)^2\sum_{t\in\mathcal{W}}|Q_{v,2}\cdot R_{v,2}(1+it)|^2, \end{aligned}$$

for a well-spaced set $\mathcal{W}\subset\mathcal{T}_2$. For instance, one may take $\mathcal{W}$ as the even or odd integers in $\mathcal{T}_2$ (choose the parity that gives a larger contribution). It happens that $\mathcal{W}$ is easier to analyze than $\mathcal{T}_2$ itself.

Now is the critical step for the choice $g=\lambda\chi$ and $\log P_2 = (\log X)^{2/3+\delta}$: by Lemma 3.4.5 (Vinogradov–Korobov), we have for all $t\in[T_0,T]$

$$|Q_{v,2}(1+it)| \ll_{\delta,A} \frac{\log X}{1+T_0} + (\log X)^{-B} \ll (\log X)^{-B}, \tag{3.49}$$

for $T_0 = (\log X)^{2B}$ and $B = 11A$. So by Lemma 3.4.3 (Halász-Montgomery), we have

$$\begin{aligned} E_2 &\ll (V\log Q_2)^2(\log X)^{2-4B}\sum_{t\in\mathcal{W}}|R_{v,2}(1+it)|^2 \\ &\ll (V\log Q_2)^2(\log X)^{3-4B}(Ye^{-v/V}+|\mathcal{W}|\sqrt{T})\,\frac{e^{v/V}}{Y} \\ &\ll (\log X)^{-B}(1+|\mathcal{W}|\frac{\sqrt{T}Q_2}{Y}), \end{aligned}$$

recalling $\log T \asymp \log X$, $V=(\log X)^B$, and $e^{v/V}\le Q_2$.

Thus it suffices to bound $|\mathcal{W}|$. We shall obtain

$$E_2 \ll (\log X)^{-B}\Big(1+\frac{T}{Y}\Big) \ll_{A,\delta} (\log X)^{-B} \tag{3.50}$$

provided we show $|\mathcal{W}| \ll T^{1/2}/Q_2$. To prove this, by definition of $\mathcal{T}_2\supset\mathcal{W}$, we first partition $\mathcal{W} = \bigcup_{u\in\mathcal{I}_1}\mathcal{W}^{(u)}$ where

$$|Q_{u,1}(1+it)| > e^{-u\alpha/V} \qquad \text{for all} \quad t\in\mathcal{W}^{(u)}.$$

Hence for each $u \in \mathcal{I}_1$, we may apply Lemma 3.4.4 to the prime polynomial $Q_{u,1}$ with $U = e^{u\alpha/V}$ and $L = e^{u/V}$, so that

$$|\mathcal{W}| = \sum_{u\in\mathcal{I}_1} |\mathcal{W}^{(u)}| \ll \sum_{u\in\mathcal{I}_1} U^2 T^{2\frac{\log U+\log\log T}{\log L}} \ \ll\ |\mathcal{I}_1| U^2 T^{2\alpha+\frac{2\log\log T}{\log L}} \ \ll\ T^{2/5+2/5A+o(1)}, \tag{3.51}$$

since $|\mathcal{I}_1| < V\log Q_1 \ll T^{o(1)}$, $U^2 \le Q_1^{2\alpha} \ll T^{o(1)}$, $\log L \ge \log P_1 \ge 5A\log\log T$, by recalling $[P_1, Q_1] = [(\log X)^{33A}, (\log X)^{\psi(X)-4A}]$ and $T \in [X^{1/4}, X^2]$.

Hence $A > 5$ gives $|\mathcal{W}| \ll T^{1/2}/Q_2$, and completes the proof of Proposition 3.5.1.

## 3.6 Average Chowla-type correlations

In this section, we establish the results for higher correlations stated in the introduction.

We first exhibit quantitative cancellation among a broad class of correlations with a 'typical' factor $\mathbf{1}_{\mathcal{S}}\mu$. We use a standard 'van der Corput' argument and then apply the key Fourier estimate.

**Lemma 3.6.1.** *Let $A > 5$, $\delta > 0$, $(\log X)^{40A} < H < X$, and $\mathcal{S} = \mathcal{S}(X, H, A, \delta)$ as in (3.12). Suppose $G : \mathbb{N} \to \mathbb{C}$ satisfies $\sum_{n\le X} |G(n)|^2 \ll X(\log X)^{A/20}$. Then*

$$\sum_{h\le H} \Big| \sum_{n\le X} \mathbf{1}_{\mathcal{S}}\mu(n+h)G(n)\Big| \ \ll_{A,\delta}\ \frac{HX}{(\log X)^{A/40}}.$$

*Proof.* Let $g = \mathbf{1}_{\mathcal{S}}\mu$. By Cauchy–Schwarz it suffices to show

$$\frac{HX^2}{(\log X)^{A/20}} \ \gg\ \sum_{h\le H} \Big| \sum_{n\le X} g(n+h)G(n)\Big|^2 \ =\ \sum_{n,n'\le X} G(n)\overline{G}(n') \sum_{h\le H} g(n+h)g(n'+h).$$

Using Cauchy–Schwarz again, the right hand side above is bounded by

$$\sum_{n\le X} |G(n)|^2 \cdot \Big( \sum_{n,n'} \Big| \sum_{h\le H} g(n+h)g(n'+h)\Big|^2\Big)^{\frac12}.$$

recalling $g$ is supported on $[1, X]$. By assumption $\sum_{n\le X} |G(n)|^2 \ll X(\log X)^{A/20}$, so it suffices to prove

$$\frac{H^2X^2}{(\log X)^{A/5}} \ \gg\ \sum_{n,n'} \Big| \sum_{h\le H} g(n+h)g(n'+h)\Big|^2 = \sum_{|h|<H} (\lfloor H\rfloor - |h|)\Big| \sum_{n} g(n)g(n+h)\Big|^2.$$

But this indeed holds: since $g = \mathbf{1}_{\mathcal{S}}\mu$, we apply Lemma 3.2.1 (Fourier bound) with $f = g = \mathbf{1}_{\mathcal{S}}\mu$. Thus the trivial bound $F(X) \ll X$ and Theorem 3.2.2 (Key Fourier estimate) give

$$\sum_{|h|\le H}\Big|\sum_{n\le X}\mathbf{1}_{\mathcal{S}}\mu(n)\,\mathbf{1}_{\mathcal{S}}\mu(n+h)\Big|^2$$
$$\ll F(X+2H)\cdot\sup_{\alpha}\int_0^X\Big|\sum_{x\le n\le x+2H}\mathbf{1}_{\mathcal{S}}\mu(n)e(n\alpha)\Big|\mathrm{d}x \ \ll\ \frac{HX^2}{(\log X)^{A/5}}. \qquad \square$$

*Proof of Theorem 3.1.3.* This follows immediately from Lemma 3.6.1, recalling (3.14),

$$\sum_{h\le H}\Big|\sum_{\substack{n\le X\\ n+h\notin\mathcal{S}}}\mu(n+h)\Big| \ \le\ \sum_{1\le j\le 2}\sum_{h\le H}\sum_{\substack{n\le X\\ q\nmid n+h\ \forall q\in[P_j,Q_j]}} 1 \ll_{A,\delta}\ \frac{HX}{\min\{\psi(X),(\log X)^{1/3-\delta}\}}. \qquad \square$$

We now prove the main technical result of the chapter, which handles averaged Chowla-type correlations with $m \ge 1$ copies of the Möbius function $\mu$ and with any function $G : \mathbb{N} \to \mathbb{C}$ of 'moderate growth' which is 'amenable to sieves.'

**Theorem 3.6.2.** *Let $A > 5$, $\delta > 0$, $(\log X)^{40A} < H < X$, and $\mathcal{S} = \mathcal{S}(X,H,A,\delta)$ as in* (3.12). *Suppose $G : \mathbb{N} \to \mathbb{C}$ satisfies $\sum_{n\le X}|G(n)|^2 \ll X(\log X)^{A/20}$. Then*

$$\sum_{h_1,..,h_m\le H}\Big|\sum_{n\le X}G(n)\prod_{j=1}^m\mu(n+h_j)\Big| \tag{3.52}$$
$$\ll_{A,\delta}\ \sum_{h_1,..,h_m\le H}\sum_{n\le X}|G(n)|\prod_{j=1}^m\mathbf{1}_{\overline{\mathcal{S}}}(n+h_j)\ +\ \frac{mXH^m}{(\log X)^{A/40}}.$$

*Proof.* We observe from Lemma 3.6.1 that any correlation with a factor $\mathbf{1}_{\mathcal{S}}\mu$ exhibits strong cancellation. So we split up $\mu = \mathbf{1}_{\overline{\mathcal{S}}}\mu + \mathbf{1}_{\mathcal{S}}\mu$ until each term has a factor $\mathbf{1}_{\mathcal{S}}\mu$, except for one term with $m$ factors of $\mathbf{1}_{\overline{\mathcal{S}}}\mu$. Thus the product in (3.52) becomes

$$\prod_{j=1}^m\mu(n+h_j) = \prod_{j=1}^m\mathbf{1}_{\overline{\mathcal{S}}}\mu(n+h_j)\ +\ \sum_{i=1}^m\mathbf{1}_{\mathcal{S}}\mu(n+h_i)\prod_{1\le j<i}\mathbf{1}_{\overline{\mathcal{S}}}\mu(n+h_j)\prod_{i<j\le m}\mu(n+h_j).$$

Hence we bound the left hand side of (3.52) by $\Sigma_1 + \Sigma_2$, where

$$\Sigma_1 = \sum_{h_1,..,h_m\le H}|G(n)|\sum_{n\le X}\prod_{j=1}^m\mathbf{1}_{\overline{\mathcal{S}}}(n+h_j), \tag{3.53}$$

$$\Sigma_2 = \sum_{i=1}^m\sum_{h_1,..,h_m\le H}\Big|\sum_{n\le X}\mathbf{1}_{\mathcal{S}}\mu(n+h_i)G_i(n)\Big|, \tag{3.54}$$

where $G_i(n) = G(n)\prod_{1\le j<i}\mathbf{1}_{\overline{\mathcal{S}}}\mu(n+h_j)\prod_{i<j\le m}\mu(n+h_j)$. In particular $|G_i(n)| \le |G(n)|$.

Thus Lemma 3.6.1 applies to each $\mathbf{1}_{\mathcal{S}}(n+h)G_i(n)$, so that $\Sigma_2$ is bounded by

$$\Sigma_2 \ll_A \frac{mH^mX}{W^{1/40}}. \qquad \square$$

### 3.6.1 Deduction of results

For convenience, denote $\psi_\delta(X) = \min\{\psi(X), (\log X)^{1/3-\delta}\}$.

*Proof of Theorem 3.1.4.* Let $G(n) = \prod_{j=1}^k \Lambda(n+a_j)$ for the tuples $\mathcal{A} = \{a_1, .., a_k\}$, $\mathcal{H} = \{h_1, \dots, h_m\}$. Then $\sum_{n\le X}|G(n)|^2 \ll X(\log X)^k$, and by a standard sieve upper bound,

$$\sum_{n\le X} G(n)\prod_{j=1}^m \mathbf{1}_{\overline{\mathcal{S}}}(n+h_j) \tag{3.55}$$
$$\ll_{m,\mathcal{A}} X\mathfrak{S}(\mathcal{A}\cup\mathcal{H})\Big(\prod_{p\in[P_1,Q_1]} + \prod_{p\in[P_2,Q_2]}\Big)\Big(1-\frac{1}{p}\Big)^{m-l} \ll_{m,\delta} \frac{X\mathfrak{S}(\mathcal{A}\cup\mathcal{H})}{\psi_\delta(X)^\ell},$$

for $\ell = |\mathcal{H}\setminus\mathcal{A}|$, and using Mertens' product theorem [104]. Thus we have

$$\Sigma := \sum_{h_1,..,h_m\le H}\Big|\sum_{n\le X}\prod_{j=1}^k\Lambda(n+a_j)\prod_{j=1}^m\mathbf{1}_{\overline{\mathcal{S}}}(n+h_j)\Big| \tag{3.56}$$
$$\le \sum_{h_1,..,h_m\le H}\Big|\sum_{n\le X}G(n)\prod_{j=1}^m\mathbf{1}_{\overline{\mathcal{S}}}(n+h_j)\Big| \ll_{m,\delta}\sum_{\ell=0}^m\sum_{\substack{h_1,..,h_m\le H\\ |\mathcal{H}\setminus\mathcal{A}|=\ell}}\frac{X\mathfrak{S}(\mathcal{A}\cup\mathcal{H})}{\psi_\delta(X)^\ell}.$$

Next, by Pintz [114, Theorem 1] we have

$$\sum_{h\le H}\mathfrak{S}(\mathcal{A}\cup\{h\}) \sim_k H\mathfrak{S}(\mathcal{A}),$$

and so by induction we have

$$\sum_{\substack{h_1,..,h_\ell\le H\\ h_i \text{ distinct}}}\mathfrak{S}(\mathcal{A}\cup\{h_1,\dots,h_\ell\}) \sim_{\ell,k} H^\ell\mathfrak{S}(\mathcal{A}). \tag{3.57}$$

For $\mathcal{H}' = \mathcal{H}\setminus\mathcal{A}$, plugging (3.57) into (3.56) gives

$$\Sigma \ll_{m,\delta}\sum_{\ell=0}^m \ell!\sum_{\substack{h_1,..,h_\ell\le H\\ h_i \text{ distinct}}}\frac{X\mathfrak{S}(\mathcal{A}\cup\mathcal{H}')}{\psi_\delta(X)^\ell}$$
$$\ll_{m,k,\delta}\sum_{\ell=0}^m\frac{XH^\ell}{\psi_\delta(X)^\ell}\mathfrak{S}(\mathcal{A}) \ll_{m,k,\delta}\frac{XH^m}{\psi_\delta(X)^m}\mathfrak{S}(\mathcal{A}),$$

since $H > (\log X)^{1/3} > \psi_\delta(X)$. In particular $\Sigma \ll_{m,\delta,\mathcal{A}} X(H/\psi_\delta(X))^m$.

Hence by Theorem 3.6.2 with $A = 20(m+k)$,

$$\sum_{h_1,..,h_m \le H} \Big| \sum_{n \le X} \prod_{j=1}^{k} \Lambda(n+a_j) \prod_{j=1}^{m} \mu(n+h_j) \Big| \tag{3.58}$$
$$\ll_{m,\delta,\mathcal{A}} \ \Sigma \ + \ \frac{XH^m}{(\log X)^{A/40}} \ \ll_{m,\delta,\mathcal{A}} \ \frac{XH^m}{\psi_\delta(X)^m}. \qquad \square$$

*Proof of Theorem 3.1.5.* Let $G(n) = \prod_{i=1}^{j} \tau_{k_i}(n+a_i)$ for the tuple $\mathcal{A} = \{a_1, .., a_j\}$ and recall $k = \sum_{i=1}^{j} k_i$. By Henriot [76, Theorem 3] (see Lemma 4.2.9), we may obtain

$$\sum_{h \le H} \sum_{n \le X} \mathbf{1}_{\overline{\mathcal{S}}}(n+h) \prod_{i=1}^{j} \tau_{k_i}(n+a_i) \ \ll_{\mathcal{A}} \ \frac{HX}{(\log X)^{j+1}} \sum_{n \le \sqrt{X}} \frac{\mathbf{1}_{\overline{\mathcal{S}}}(n)}{n} \prod_{i=1}^{j} \sum_{n \le \sqrt{X}} \frac{\tau_{k_i}(n)}{n}.$$

Observe $\sum_{n \le \sqrt{X}} \tau_{k_i}(n)/n \ll (\sum_{n \le \sqrt{X}} 1/n)^{k_i} \ll (\log X)^{k_i}$, and so by Mertens' product theorem [104],

$$\sum_{n \le \sqrt{X}} \frac{\mathbf{1}_{\overline{\mathcal{S}}}(n)}{n} \ \ll \ \log X \Big( \prod_{p \in [P_1, Q_1]} + \prod_{p \in [P_2, Q_2]} \Big) \Big(1 - \frac{1}{p}\Big) \ \ll_\delta \ \frac{\log X}{\psi_\delta(X)}.$$

Thus since $\sum_{n \le X} |G(n)|^2 \ll X(\log X)^k$, Theorem 3.6.2 with $A = 20k$ gives

$$\sum_{h \le H} \Big| \sum_{n \le X} \mu(n+h) \prod_{i=1}^{j} \tau_{k_i}(n+a_i) \Big| \ \ll_{\delta,\mathcal{A}} \ \frac{HX}{\psi_\delta(X)} (\log X)^{k-j}. \qquad \square$$

### 3.6.2 Almost all shifts

Corollary 3.1.4.2 follows from the following result by the triangle inequality for $g_j = \mu$.

**Theorem 3.6.3.** *Suppose $H < X$ and $\log H / \log_2 X \to \infty$ as $X \to \infty$. Let $g_1 = \mu$ and take any $g_j : \mathbb{N} \to \mathbb{C}$ with $|g_j| \le 1$ for $1 < j \le k$. Then for any fixed shifts $h_2, ..., h_k \le H$, $K > 0$ we have*

$$\sum_{p \le X} \prod_{j=1}^{k} g_j(p+h_j) \ = \ o(\pi(X)), \tag{3.59}$$

*for all except $O_K(H(\log X)^{-K})$ shifts $h_1 \le H$.*

*Proof.* Given $\varepsilon > 0$ and fixed shifts $h_2, .., h_k \le H$, we aim to show $|\mathcal{E}| \ll_\varepsilon H(\log X)^{-K}$ for the exceptional set

$$\mathcal{E} = \Big\{ h \le H : \Big| \sum_{p \le X} \mu(p+h) \prod_{j=2}^{k} g_j(p+h_j) \Big| > 2\varepsilon \pi(X) \Big\}. \tag{3.60}$$

To this, by Markov's inequality we have

$$\begin{aligned}|\mathcal{E}|(\varepsilon\pi(X)) &\ll \sum_{h\in\mathcal{E}}\Big|\sum_{p\le X}\mu(p+h)\prod_{j=2}^{k}g_j(p+h_j)\Big| \\ &\le \sum_{h\in\mathcal{E}}\Big|\sum_{p\le X}\mathbf{1}_{\overline{\mathcal{S}}}(p+h)\Big| \;+\; \sum_{h\le H}\Big|\sum_{p\le X}\mathbf{1}_{\mathcal{S}}\mu(p+h)\prod_{j=2}^{k}g_j(p+h_j)\Big| \\ &\ll_A \quad \frac{\pi(X)}{\psi(X)}\sum_{h\in\mathcal{E}}\prod_{\substack{p\mid h\\ p>P_1}}(1+\tfrac{1}{p}) \quad + \quad \frac{H\pi(X)}{(\log X)^{A/40}},\end{aligned}$$

using Lemma 3.6.1 when $p+h\in\mathcal{S}$, and a standard sieve upper bound [62, Theorem 7.4] when $p+h\notin\mathcal{S}$. Here $\mathcal{S}=\mathcal{S}(X,H,A,\delta)$ as in (3.12) with $A=80K$ and $\delta=1/10$, say.

Observe for any $h\le H=(\log X)^{\psi(X)}$, we see the above product is at most $\prod_{P_1<p\le z}(1+\frac{1}{p})\ll\frac{\log z}{\log P_1}$ where $z=P_1+\psi(X)\log_2 X$. Recalling $P_1=(\log X)^{33A}$ this gives

$$\frac{\pi(X)}{\psi(X)}\sum_{h\in\mathcal{E}}\prod_{\substack{p\mid h\\ p>P_1}}(1+\tfrac{1}{p}) \;=\; o\big(|\mathcal{E}|\,\pi(X)\big).$$

Hence we conclude $|\mathcal{E}|\ll\frac{1}{\varepsilon}H(\log X)^{-K}$ as desired. □

## 3.7 Non-pretentious multiplicative functions

In this section we prove Theorem 3.1.6, which we restate below.

**Theorem 3.1.6** *Let $H=X^\theta$ for $\theta\in(0,1)$, and take a multiplicative function $f:\mathbb{N}\to\mathbb{C}$ with $|f|\le 1$. There exists $\rho\in(0,\frac{1}{8})$ such that, if $M(f;X^2/H^{2-\rho},Q)\to\infty$ as $X\to\infty$ for each fixed $Q$, then*

$$\sum_{h\le H}\Big|\sum_{p\le X}f(p+h)\Big| \;=\; o_{\theta,\rho}\big(H\pi(X)\big).$$

*Proof.* Consider the exponential sum $F_x(\alpha)=\sum_{x\le m\le x+2H}f(m)e(m\alpha)$. The hypotheses of our theorem are made in order to satisfy [99, Theorem 1.4], which in this case gives

$$\int_0^X \sup_\alpha|F_x(\alpha)|\mathrm{d}x \;=\; o_{\theta,\rho}(HX). \tag{3.61}$$

We critically note the supremum is *inside* the integral.

Now on to the proof, it suffices to show $S_f = o(HX)$ where

$$S_f := \sum_{h\le H}\Big|\sum_{n\le X}\Lambda(n)f(n+h)\Big| \ll \frac{1}{H}\sum_{h\le 2H}(2H-h)\Big|\sum_{n\le X}\Lambda(n)f(n+h)\Big|.$$

For each sum $z_h = \sum_{n\le X}\Lambda(n)f(n+h)$, denote 1-bounded coefficients $c(h)$ given by the relation $|z_h| = c(h)z_h$ so that

$$\begin{aligned}
S_f &\ll \frac{1}{H}\sum_{h\le 2H}(2H-h)c(h)\sum_{n\le X}\Lambda(n)f(n+h)\\
&= \frac{1}{H}\sum_{h\le 2H}c(h)\sum_{n\le X}\Lambda(n)\sum_{m\le X+2H}f(m)\mathbf{1}_{m=n+h}\cdot\int_0^X \mathbf{1}_{x\le n,m\le x+2H}\mathrm{d}x\\
&= \frac{1}{H}\int_0^X\int_0^1\sum_{h\le 2H}c(h)e(h\alpha)\sum_{x\le n,m\le x+2H}\Lambda(n)f(m)e\big((n-m)\alpha\big)\mathrm{d}\alpha\mathrm{d}x,
\end{aligned}$$

by orthogonality $\int_0^1 e(n\alpha)\mathrm{d}\alpha = \mathbf{1}_{n=0}$. That is, we have the following triple convolution

$$S_f \ll \frac{1}{H}\int_0^X\int_0^1 C_0(\alpha)L_x(-\alpha)F_x(\alpha)\mathrm{d}\alpha\mathrm{d}x, \tag{3.62}$$

denoting the sums $C_0(\alpha) = \sum_{h\le 2H}c(h)e(h\alpha)$ and $L_x(\alpha) = \sum_{x\le n\le x+2H}\Lambda(n)e(n\alpha)$.

We shall split the inner integral on $\alpha$ according to the size of $L_x$. Specifically, for each $x$ let $\mathcal{T}_x = \{\alpha\in[0,1] : |L_x(\alpha)|\ge \delta H\}$. Then by Markov's inequality, $\mathcal{T}_x$ has measure

$$\int_{\mathcal{T}_x}\mathrm{d}\alpha \le \frac{1}{(\delta H)^4}\int_{\mathcal{T}_x}|L_x(\alpha)|^4\mathrm{d}\alpha \ll \frac{1}{\delta^4 H}, \tag{3.63}$$

since the Fourier identity implies

$$\begin{aligned}
\int_0^1|L_x(\alpha)|^4\mathrm{d}\alpha &= \sum_{x\le n_1,n_2,n_3,n_4\le x+2H}\Lambda(n_1)\Lambda(n_2)\Lambda(n_3)\Lambda(n_4)\mathbf{1}_{n_1+n_2=n_3+n_4}\\
&\ll (\log X)^4\sum_{\substack{x\le p_1,p_2,p_3,p_4\le x+2H\\ p_1+p_2=p_3+p_4}}1 \ll_\theta H^3,
\end{aligned}$$

by a standard sieve upper bound [62, Theorem 7.4]. Thus plugging (3.63) into (3.62), we obtain

$$\begin{aligned}
S_f &\ll \frac{1}{H}\int_0^X\int_{[0,1]\setminus\mathcal{T}_x}C_0(\alpha)L_x(-\alpha)F_x(\alpha)\mathrm{d}\alpha\mathrm{d}x\\
&\quad+\frac{1}{\delta^4H^2}\int_0^X\sup_{\alpha\in\mathcal{T}_x}|C_0(\alpha)L_x(-\alpha)F_x(\alpha)|\mathrm{d}x.
\end{aligned}$$

Denote the two integrals above by $I_1$ and $I_2$. Observe $I_2 \ll_\theta \delta^{-4}\int_0^X \sup_\alpha |F_x(\alpha)|\mathrm{d}x$, using $|C_0(\alpha)| \le H$ trivially and $|L_x(\alpha)| \ll_\theta H$ by the Brun–Titchmarsh theorem. Then by definition of $\mathcal{T}_x$, Cauchy–Schwarz implies

$$\begin{aligned} I_1 &\le \delta \int_0^X \int_{[0,1]\setminus\mathcal{T}_x} |C_0(\alpha)F_x(\alpha)|\mathrm{d}\alpha\mathrm{d}x \\ &\le \delta \int_0^X \left(\int_0^1 |C_0(\alpha)|^2\mathrm{d}\alpha \cdot \int_0^1 |F_x(\alpha)|^2\mathrm{d}\alpha\right)^{1/2}\mathrm{d}x \ll \delta HX, \end{aligned}$$

by Parseval's identity applied to $C_0$ and $F_x$. Thus combining bounds for $I_1, I_2$ gives

$$S_f \ll_\theta \delta HX + \delta^{-4}\int_0^X \sup_\alpha |F_x(\alpha)|\mathrm{d}x. \tag{3.64}$$

Hence taking $\delta \to 0$, the Fourier uniformity bound (3.61) gives $S_f = o_{\theta,\rho}(HX)$ as claimed. □

# Chapter 4

# On the Hardy–Littlewood–Chowla conjecture on average

## 4.1 Introduction

We continue our investigations from Chapter 3, now with our focus on acheiving only qualititative cancellation, and particular attention payed to the higher correlations of Mobius $\mu$ and von Mangoldt $\Lambda$ functions.

It is natural to consider the following hybrid conjecture which, as in [135], we call the Hardy–Littlewood–Chowla conjecture.

**Conjecture 4.1.1** (Hardy–Littlewood–Chowla)**.** Let $k, \ell \geq 0$, and let $h_1, \dots, h_k$, $a_1, \dots, a_\ell$ be distinct integers. Then we have[1]

$$\sum_{n \leq X} \mu(n+h_1) \cdots \mu(n+h_k)\Lambda(n+a_1) \cdots \Lambda(n+a_\ell) \;=\; (\mathfrak{S} + o(1))X. \tag{4.1}$$

Here $\mathfrak{S} = \prod_p \frac{(1-\nu_p/p)}{(1-1/p)^k}$, where $\nu_p = \#\{a_1, .., a_\ell \pmod p\}$ if $k = 0$, and $\mathfrak{S} = 0$ if $k > 0$.

The 'pure' cases $k = 0$ and $\ell = 0$ specialize to the original conjectures of Hardy–Littlewood and of Chowla, respectively. We remark that under the hypothetical assumption of infinitely many Siegel zeros, Conjecture 4.1.1 was recently verified for $\ell \leq 2$ by Tao and Teräväinen [135] (generalizing works of Heath-Brown [75] and Chinis [25]). In the current chapter, we will be concerned with unconditional results.

Our main result is an averaged form of Conjecture 4.1.1 for the genuinely 'hybrid' cases $k, \ell \geq 1$.

**Theorem 4.1.2** (Hardy–Littlewood–Chowla on average)**.** *Let $\varepsilon > 0$ and $k, \ell \geq 1$ be fixed, and let $h_2, \dots, h_k$, $a_1, \dots, a_\ell$ be fixed and distinct positive integers.*

*(i) Let $(\log X)^{\ell+\varepsilon} \leq H \leq \exp((\log X)^a)$ with $a = a(\varepsilon, \ell) > 0$ small enough. Then*

$$\sum_{h_1 \leq H} \left| \sum_{n \leq X} \mu(n+h_1) \cdots \mu(n+h_k)\Lambda(n+a_1) \cdots \Lambda(n+a_\ell) \right| \;\ll\; HX \, \frac{\log\log H}{\log H}. \tag{4.2}$$

*(ii) Let $C \geq 1$ be fixed, and let $(\log X)^{\ell+\varepsilon} \leq H \leq (\log X)^C$. There exists an absolute constant $c > 0$ such that for any $\delta$ with $10^4 \ell \varepsilon^{-1} (\log\log H)/(\log H) \leq \delta \leq c/C$, we have*

$$\left| \sum_{n \leq X} \mu(n+h_1) \cdots \mu(n+h_k)\Lambda(n+a_1) \cdots \Lambda(n+a_\ell) \right| \leq \delta X$$

*for all except $\ll H^{1-c\delta\varepsilon/\ell}$ values of $h_1 \leq H$.*

[1]Here we use the convention that an empty product equals to 1.

In particular, (4.1) holds for all but $o(H)$ values of $h_1 \leq H$ (and the exceptional set can be taken to be nearly power-saving).

As is clear from the proof, Theorem 4.1.2 holds equally well with the Liouville function in place of the Möbius function.

*Remark* 4.1.3. It is likely that with some extra effort (in particular, refining Proposition 4.2.1 for $\mu$ in the spirit of Theorem 3.2.2) Theorem 4.1.2(ii) could be extended to the regime $(\log X)^{\ell+\varepsilon} \leq H \leq X$. However, our main interest here is in taking $H$ as small as possible.

Earlier, Matomäki, Radziwiłł and Tao [96] proved the case $\ell = 0$ of Conjecture 4.1.1 for almost all $h_1 \leq H$, with $H = \psi(X)$ tending to infinity arbitrarily slowly. In [97], they considered the case $k = 0$, $\ell = 2$ and proved the conjecture for almost all $h_1 \leq H$, with $H \geq X^{8/33+\varepsilon}$. In Chapter 3, the author obtained cancellation in the range $H \geq (\log X)^{\psi(X)}$, implied by the stronger quantitative bound

$$\sum_{h_1,\dots,h_k \leq H} \left| \sum_{n \leq X} \mu(n+h_1)\cdots\mu(n+h_k)\Lambda(n+a_1)\cdots\Lambda(n+a_\ell) \right| \ll \frac{H^k X}{\min\left\{\psi(X)^k, (\log X)^{k/3-o(1)}\right\}}.$$

The simplest hybrid case $k = \ell = 1$ reduces to the Möbius function on shifted primes. It is a folklore conjecture and a well-known model case for the parity problem in sieve theory that

$$\sum_{p \leq X} \mu(p+h) \;=\; o(\pi(X))$$

for any fixed shift $h \neq 0$. This has appeared in print (for $h = 1$) at least since Hildebrand [77]; see also Pintz [113] and Murty–Vatwani [107, (1.2)]. Theorem 4.1.2 directly implies an averaged form of this conjecture for $H \geq (\log X)^{1+\varepsilon}$.

**Corollary 4.1.3.1.** *Let $\varepsilon > 0$. Then, for $X \geq H \geq (\log X)^{1+\varepsilon}$, we have*

$$\sum_{p \leq X} \mu(p+h) = o(\pi(X)) \tag{4.3}$$

*for all except $o(H)$ values of $h \leq H$.*

This improves on the work of the author in Chapter 3, which established Corollary 4.1.3.1 when $H \geq (\log X)^{\psi(X)}$ for any function $\psi(X)$ tending to infinity with $X$.

### 4.1.1 Other multiplicative functions

We also consider a variant of Conjecture 4.1.1 where the occurrences of the Möbius function are replaced with other 1-bounded multiplicative functions $f_i : \mathbb{N} \to \mathbb{C}$, with $f_1$ not pretending to be a twisted character $\chi(n)n^{it}$ for any $\chi \pmod q$ and $t \in \mathbb{R}$. Here, following Granville and Soundararajan [68], pretentiousness is measured by the pretentious distance

$$\mathbb{D}(f, g; X) = \Big( \sum_{p \le X} \frac{1 - \Re(f(p)\overline{g(p)})}{p} \Big)^{1/2}$$

and the related quantity

$$M(f; X, Q) = \inf_{\substack{|t| \le X \\ \chi \,(q),\, q \le Q}} \mathbb{D}\big(f,\, n \mapsto n^{it}\chi(n); X\big)^2. \tag{4.4}$$

The following generalization of Conjecture 4.1.1 is closely related to Elliott's conjecture [38] on correlations of multiplicative functions (in fact, the case $\ell = 0$ is precisely Elliott's conjecture), so we shall call it the Hardy–Littlewood–Elliott conjecture.

**Conjecture 4.1.4** (Hardy–Littlewood–Elliott)**.** Let $h_1, \dots, h_k, a_1, \dots, a_\ell$ be distinct integers. Fix 1-bounded multiplicative functions $f_1, \dots, f_k$. Suppose that $f_1$ is non-pretentious in the sense that

$$M(f_1; X, Q) \xrightarrow{X \to \infty} \infty$$

for any $Q \ge 1$. Then we have

$$\sum_{n \le X} f_1(n + h_1) \cdots f_k(n + h_k)\Lambda(n + a_1) \cdots \Lambda(n + a_\ell) = o(X). \tag{4.5}$$

The case $\ell = 0$ of this is Elliott's conjecture (in the slightly corrected formulation of [134, Conjecture 1.3]).

We extend our results to non-pretentious multiplicative functions in the regime $H \ge (\log X)^{\ell+\varepsilon}$.

**Theorem 4.1.5** (Hardy–Littlewood–Elliott on average)**.** *Let $\varepsilon > 0$, $k, \ell \ge 1$ be fixed, and let $h_2, \dots, h_k, a_1, \dots, a_\ell$ be fixed and distinct positive integers. Let $(\log X)^{\ell+\varepsilon} \le H \le \exp((\log X)^{1/1000})$. Let $f_1, \dots, f_k$ be 1-bounded multiplicative functions. Then we have*

$$\sum_{h_1 \le H} \Big| \sum_{n \le X} f_1(n + h_1) \cdots f_k(n + h_k)\Lambda(n + a_1) \cdots \Lambda(n + a_\ell) \Big|$$
$$\ll HX \left( \exp\big(-M(f_1; X, \min\{(\log X)^{1/125}, (\log H)^5\}\big)/(10000\ell\varepsilon^{-1})\big) + \frac{\log\log H}{\log H} \right).$$

*In particular, if* $(\log X)^{\ell+\varepsilon} \le H \le \exp((\log X)^{1/1000})$ *and*

$$M(f_1; X, \min\{(\log X)^{1/125}, (\log H)^5\}) \xrightarrow{X\to\infty} \infty,$$

*then* (4.5) *holds for all except* $o(H)$ *values of* $h_1 \le H$.

*Remark* 4.1.6. Recall from (3.7) that $M(\mu; X, (\log X)^A) \ge (1/3 - \epsilon)\log\log X + O(1)$ by the Vinogradov–Korobov zero-free region. Thus, for $f_1 = \mu$ in the range $\exp((\log X)^a) \ge H \ge (\log X)^{\ell+\varepsilon}$, Theorem 4.1.5 simplifies to

$$\sum_{h_1\le H}\left|\sum_{n\le X}\mu(n+h_1)\cdots f_k(n+h_k)\Lambda(n+a_1)\cdots\Lambda(n+a_\ell)\right| \ll HX\frac{\log\log H}{\log H}.$$

Hence Theorem 4.1.2(i) is a special case of Theorem 4.1.5. The rest of the chapter is therefore devoted to the proofs of Theorem 4.1.5 and Theorem 4.1.2(ii).

Theorem 4.1.5 improves on the range $H \ge \exp((\log X)^{5/8+\varepsilon})$ which follows (under a slightly different pretentiousness hypothesis) from Fourier uniformity bounds of Matomäki, Radziwiłł, Tao, Ziegler and Teräväinen [100, Theorem 1.8]. See Theorem 3.1.6 for the details of this implication[2]. We also note that the case $\ell = 0$ of Conjecture 4.1.4 was proven on average by Matomäki–Radziwiłł–Tao [96] in the regime $H = \psi(X) \to \infty$.

### 4.1.2 Outline of the proof

The proof method of Theorem 4.1.2 (and Theorem 4.1.5) is somewhat different from that of Chapter 3 (or [96]). Both proofs begin with Fourier-analytic identities, namely Lemma 3.2.1 and Proposition 4.3.2 below. The latter allows one to study mixed 2-point correlations on average, for some functions $f$ and $g$ which satisfy certain technical hypotheses. Crucially, in Proposition 4.3.2 having qualitative cancellation for the exponential sum of $f$ on average over short intervals $[x, x+H]$ suffices, whereas in Lemma 3.2.1 one needs to save a factor comparable to $((1/X)\sum_{n\le X}|g(n)|^2)^{-1/2}$ (but no additional hypotheses were required of $f, g$). For intervals of length $H = (\log X)^{O(1)}$, the bound in [96] for the exponential sum of a multiplicative $f$ on average over short intervals $[x, x+H]$ saves less than $(\log X)^{-o(1)}$, which is problematic if the mean value of $|g|^2$ is a power of $\log X$, so the weakening of the Fourier assumption on $f$ in Proposition 4.3.2 is important for us.

[2]Theorem 3.1.6 relies on [99, Theorem 1.4] as the Fourier uniformity input, and hence produces a result for $H = X^\theta$. Inserting instead [100, Theorem 1.12] into the argument produces a result for $H = \exp((\log X)^{5/8+\varepsilon})$.

One ultimately applies Proposition 4.3.2 for the choices $f = \mu$ (say) and $g(n) = \mu(n+h_2)\cdots\mu(n+h_k)\Lambda(n+a_1)\cdots\Lambda(n+a_\ell)$, in which case verifying the technical hypotheses may be reduced to the following problems:

(I) cancellation in short exponential sums of $\mu$ on average,

(II) moment estimates for short exponential sums over primes, i.e. for 'typical' $x$,

$$\int_0^1 \Big| \sum_{x\le n\le x+H} \Lambda(n)e(\alpha n)\Big|^{2m} \mathrm{d}\alpha \ \ll\ H^{2m-1}.$$

Here (I) is deduced from the work of Matomäki–Radziwiłł–Tao [96], with special care paid to the shape of the error terms in order to obtain a power-saving error term in Theorem 4.1.2(ii). This is carried out in Proposition 4.2.1.

For (II), the intervals involved are too short for sieve methods to provide us with a strong enough pointwise bound on these moments. However, on average over $x$, in Proposition 4.2.6 we show that the moments above are of the expected order of magnitude by applying upper bounds for prime tuples coming from Selberg's sieve, and by performing an analysis of the resulting singular series.

The method presented is rather flexible, and could be applied (with some additional effort) to other correlation problems as well, such as the sums

$$\sum_{n\le X} \mu(n+h)\Lambda(P(n)) \quad \text{or} \quad \sum_{n\le X} \mu(n+h)\mathbf{1}_{\exists a,b\in\mathbb{Z}\,:\, n=a^2+b^2}$$

on average over $h$ (with $P(Y)\in\mathbb{Z}[Y]$). We leave the details of such generalizations to the interested reader.

## 4.2 Auxiliary results

### 4.2.1 A short exponential sum estimate

As in [96], we introduce a set $\mathcal{S}_{P_1,Q_1,X}$ (depending on parameters $10 < P_1 < Q_1 \le X$) consisting of the "typical" numbers having prime factors from certain long ranges. More precisely, $\mathcal{S}_{P_1,Q_1,X}$ is the set of positive integers having a prime factor in each of the ranges $[P_j, Q_j]$, $j = 1,\dots,J$, where for $j > 1$ we define

$$P_j = \exp(j^{4j}(\log Q_1)^{j-1}(\log P_1)),$$
$$Q_j = \exp(j^{4j+2}(\log Q_1)^j),$$

and $J$ is the largest index such that $Q_J \le \exp(\sqrt{\frac{1}{2}\log X})$. If $\mathcal{S}^c_{P_1,Q_1,X}$ denotes the complement of $\mathcal{S}_{P_1,Q_1,X}$ in $\mathbb{N}\cap[1,X]$, then the fundamental lemma of sieve theory [62, Lemma 6.17] tells us that

$$\left|\mathcal{S}^c_{P_1,Q_1,X}\right| \ll \sum_{j\le J}\frac{\log P_j}{\log Q_j}X \ll \frac{\log P_1}{\log Q_1}X. \tag{4.6}$$

We then have the following exponential sum estimate (with the supremum outside) for short sums of multiplicative functions, which follows from [96, Theorem 2.3].

**Proposition 4.2.1.** *Let* $3 \le H \le \exp((\log X)^{1/1000})$ *and*

$$\frac{\log\log H}{\log H} \le \ \delta \ \le \frac{1}{2000}.$$

*Let* $f$ *be a 1-bounded multiplicative function. Then there exist* $10 < P_1 < Q_1 \le X$ *such that*

$$\frac{\log P_1}{\log Q_1} \ll \delta + \exp(-M(f;X,Q)/2000) \tag{4.7}$$

*and such that for* $\mathcal{S} = \mathcal{S}_{P_1,Q_1,X}$ *we have*

$$\sup_\alpha \int_0^X \Big|\sum_{x\le n\le x+H} f(n)\mathbf{1}_{\mathcal{S}}(n)e(n\alpha)\Big|\mathrm{d}x \ll \Big(\exp(-M(f;X,Q)/2000) + H^{-\delta}\Big)HX, \tag{4.8}$$

*where* $Q = \min\{(\log X)^{1/125}, (\log H)^5\}$.

*Remark* 4.2.2. Applying $M(\mu;X,(\log X)^A) \ge (1/3-\epsilon)\log\log X + O(1)$ from (3.7), and assuming that $H \le (\log X)^C$, the bound (4.8) in the case of $f=\mu$ simplifies to

$$\sup_\alpha \int_0^X \Big|\sum_{x\le n\le x+H} \mu(n)\mathbf{1}_{\mathcal{S}}(n)e(n\alpha)\Big|\mathrm{d}x \ll H^{1-\delta}X \tag{4.9}$$

for some $\delta = \delta(C) > 0$. Moreover, (4.7) then simplifies to $(\log P_1)/(\log Q_1) \ll \delta$.

*Proof.* Without loss of generality, we may assume that $H \ge H_0$ for a large enough constant $H_0$. Denote the integral in (4.8) by $J_f$. Suppose first that

$$(\log H)^6 \le \exp(M(f;X,Q)/300).$$

Let

$$W = \min\{H^{6\delta}, \exp(M(f;X,Q)/300)\},$$

and let $P_1 = W^{200}$, $Q_1 = H/W^3$ be as in [96, Theorem 2.3]. Note that $W \ll (\log X)^{1/150}$ by the fact that $M(f;X,Q) \le 2\log\log X + O(1)$. Note also that $\frac{\log P_1}{\log Q_1} \ll \frac{\log W}{\log H} \ll \delta$, and that all the conditions for $W$ in [96, Theorem 2.3] are satisfied (that is, $W \ge (\log H)^5$ and $W \le \min\{H^{1/250}, (\log X)^{1/125}\}$ and $W \le \exp(M(f;X,Q)/3)$), so we obtain

$$J_f \ll \frac{(\log H)^{1/4}(\log\log H)}{W^{1/4}} HX \ll (H^{-\delta} + \exp(-M(f;X,Q)/2000))HX.$$

Suppose then that

$$\exp(M(f;X,Q)/300) := (\log H_1)^6 < (\log H)^6.$$

Let

$$W = (\log H_1)^6 = \exp(M(f;X,Q)/300), \tag{4.10}$$

and let $P_1 = W^{200}$, $Q_1 = H/W^3$ as before. We may assume $H_1 \ge 1000$, since otherwise the claim is trivial. Now, by noting that $\frac{\log P_1}{\log Q_1} \ll \exp(-M(f;X,Q)/2000)$ and splitting the sum on $H$ in (4.8) into shorter sums of length in $[H_1/4, H_1/2]$, it suffices to prove that

$$\sup_{\alpha} \int_0^X \Big| \sum_{x \le n \le x+H'} f(n)\mathbf{1}_{\mathcal{S}}(n) e(n\alpha) \Big| \mathrm{d}x \ll \exp(-M(f;X,Q)/2000) H'X,$$

uniformly for $H' \in [H_1/4, H_1/2]$. Since the conditions for $W$ in [96, Theorem 2.3] are satisfied, we can again apply that theorem to obtain the desired bound (noting that $\frac{(\log H_1)^{1/4}(\log\log H_1)}{W^{1/4}} \ll \exp(-M(f;X,Q)/2000)$ in this situation). This completes the proof. □

### 4.2.2 Upper-bounding correlations of primes

The goal of this subsection is to prove Proposition 4.2.6 below, which gives optimal bounds for even moments of the short exponential sums associated with correlations of the von Mangoldt function. We first need a few lemmas on tuples of primes and averages of singular series.

Recall for a tuple $\mathcal{H}$ of integers, $\nu_p(\mathcal{H}) = |\{h \pmod p : h \in \mathcal{H}\}|$. A well-known application of Selberg's sieve upper bounds the number of $k$-tuples of primes, for fixed $k$.

**Lemma 4.2.3.** *Let $k \geq 1$ be fixed, let $X \geq X_0(k)$, and suppose that $h_1, \dots, h_k \in [-X, X]$ are distinct integers. Denote $\mathcal{H} = \{h_1, \dots, h_k\}$. Then we have*

$$\big|\{n \leq X \ : \ n + h_1, \dots, n + h_k \in \mathbb{P}\}\big| \ \leq \ k! \cdot 2^k \, \frac{\mathfrak{S}(\mathcal{H}) X}{(\log X)^k} \Big(1 + O\Big(\frac{\log\log X}{\log X}\Big)\Big),$$

*where the singular series is given by*

$$\mathfrak{S}(\mathcal{H}) := \prod_p \Big(1 - \frac{\nu_p(\mathcal{H})}{p}\Big)\Big(1 - \frac{1}{p}\Big)^{-k}.$$

*Proof.* This is [62, Theorem 7.16]. □

We have a trivial upper bound for the values of the singular series, which we will need in what follows.

**Lemma 4.2.4.** *Let $k \geq 1$, and let $\mathcal{H} = \{h_1, \dots, h_k\}$ be a tuple of distinct integers. Then*

$$\mathfrak{S}(\mathcal{H}) \ll_k \prod_{1 \leq i < j \leq k} \frac{|h_i - h_j|}{\varphi(|h_i - h_j|)}.$$

*Proof.* We have

$$\begin{aligned} \mathfrak{S}(\mathcal{H}) = \prod_p \Big(1 - \frac{\nu_p(\mathcal{H})}{p}\Big)\Big(1 - \frac{1}{p}\Big)^{-k} &\ll_k \prod_p \Big(1 - \frac{1}{p}\Big)^{\nu_p(\mathcal{H}) - k} \\ &\ll \prod_{1 \leq i < j \leq k} \prod_{p \mid h_i - h_j} \Big(1 - \frac{1}{p}\Big)^{-1} \ll \prod_{1 \leq i < j \leq k} \frac{|h_i - h_j|}{\varphi(|h_i - h_j|)}, \end{aligned}$$

since $k - \nu_p(\mathcal{H}) \leq |\{(i, j) \ : \ 1 \leq i < j \leq k, \, p \mid h_i - h_j\}|$. □

We will also need an upper bound for the autocorrelations of $h/\varphi(h)$.

**Lemma 4.2.5.** *Let $H \geq 1$ and let $C, k, r \geq 1$ be fixed. Then we have*[3]

$$\sum_{h \leq H} \prod_{i=1}^{r} \left(\frac{|L_i(h)|}{\varphi(|L_i(h)|)}\right)^k \ll_{r,k,C} H$$

*uniformly for any linear functions $L_i(h) = a_i h + b_i$ with $a_i, b_i \in [-C, C] \cap \mathbb{Z}$.*

[3] In what follows, we interpret $0/\varphi(0)$ as 0, say.

*Proof.* First note that for $H' \geq 1$ and any integer $m \geq 1$, we have

$$\begin{aligned}\sum_{h\leq H'}\Big(\frac{h}{\varphi(h)}\Big)^m &= \sum_{h\leq H'}\Big(\sum_{d|h}\frac{1}{d}\Big)^m = \sum_{d_1,\dots,d_m\geq 1}\frac{1}{d_1\cdots d_m}\sum_{\substack{h\leq H'\\ d_i|h\ \forall i\leq m}} 1\\ &\ll H'\sum_{d_1,\dots,d_m\geq 1}\frac{1}{d_1\cdots d_m[d_1,\dots,d_m]}\\ &\ll H'\Big(\sum_{d\geq 1}\frac{1}{d^{1+1/m}}\Big)^m \ll_m H'. \end{aligned} \tag{4.11}$$

since $[d_1,\dots,d_m] \geq \max_{i\leq m} d_i \geq (d_1\cdots d_m)^{1/m}$. Thus by Hölder's inequality and (4.11) with $H' = (C+1)H$, $m = rk$, we have

$$\begin{aligned}\sum_{h\leq H}\prod_{i\leq r}\Big(\frac{|L_i(h)|}{\varphi(|L_i(h)|)}\Big)^k &\leq \prod_{i\leq r}\Big(\sum_{h\leq H}\Big(\frac{|L_i(h)|}{\varphi(|L_i(h)|)}\Big)^{rk}\Big)^{1/r}\\ &\leq \prod_{i\leq r}\Big(\sum_{h\leq (C+1)H}\Big(\frac{h}{\varphi(h)}\Big)^{rk}\Big)^{1/r} \ll_{r,k,C} H,\end{aligned}$$

using the fact that $L_i(\cdot)$ takes any integer value at most once. □

The previous lemmas lead to the following corollary.

**Corollary 4.2.5.1.** *Let $k \geq m \geq 1$ be fixed. Then*

$$\sum_{\mathbf{h}\in[0,H]^m}\sup_{L_1,\dots,L_k}\mathfrak{S}(\{L_1(\mathbf{h}),\dots,L_k(\mathbf{h})\}) \ll_{k,A} H^m,$$

*where the supremum ranges over affine-linear forms $L_j(h_1,\dots,h_m) = a_j + \sum_{i=1}^m a_{i,j}h_i$ with integer coefficients and with $|a_{i,j}|, |a_j| \leq A$.*

*Proof.* Denote by $\Sigma$ the sum of interest on left hand side. Since there are $O_{k,A}(1)$ many choices of affine-linear forms $L_1,\dots,L_k$ with $A$-bounded coefficients, it suffices to take the supremum outside the sum in $\Sigma$. Then for each $L_1,\dots,L_k$, by Lemma 4.2.4 we bound $\Sigma$ as

$$\begin{aligned}\Sigma &\ll_k \sum_{\mathbf{h}\in[0,H]^m}\prod_{1\leq j<j'\leq k}\frac{|L_j(\mathbf{h})-L_{j'}(\mathbf{h})|}{\varphi(|L_j(\mathbf{h})-L_{j'}(\mathbf{h})|)}\\ &= \sum_{\mathbf{h}\in[0,H]^{m-1}}\sum_{h\leq H}\prod_{1\leq j<j'\leq k}\frac{|L_{j,j',\mathbf{h}}(h)|}{\varphi(|L_{j,j',\mathbf{h}}(h)|)} \ll \sum_{\mathbf{h}\in[0,H]^{m-1}} H \ll H^m,\end{aligned}$$

since to each $\mathbf{h}\in[0,H]^{m-1}$, we can apply Lemma 4.2.5 with $r = \binom{k}{2}$, $C = mA$, and linear functions $L_{j,j',\mathbf{h}}(h) := L_j(\mathbf{h},h) - L_{j'}(\mathbf{h},h)$. □

Using Corollary 4.2.5.1, we can prove the following bound for high moments of short exponential sums associated with the correlations of the von Mangoldt function. This estimate will be needed in Section 4.4.

**Proposition 4.2.6.** *Let $\ell \geq 1$, $k \geq 2$ be fixed, and let $a_1, \dots, a_\ell$ be distinct fixed integers. For $X \geq H \geq 2$ and $|g(n)| \leq 1$, we have*

$$\begin{aligned} M_{2k} &:= \int_0^X \int_0^1 \Big| \sum_{x \leq n \leq x+H} g(n) e(\alpha n) \prod_{j=1}^{\ell} \Lambda(n+a_j) \Big|^{2k} \mathrm{d}\alpha \mathrm{d}x \\ &\ll X(H^{2k-1} + H^k (\log X)^{k\ell} + H(\log X)^{(2k-1)\ell}). \end{aligned}$$

*In particular if $H \geq (\log X)^{\ell k/(k-1)}$ then*

$$M_{2k} \ll XH^{2k-1}. \tag{4.12}$$

*Remark* 4.2.7. The bound above for $M_{2k}$ is optimal up to a constant factor, assuming the Hardy–Littlewood prime tuples conjecture. Indeed, one can show that $M_{2k} \gg XH^{2k-1}$ in the case $g(n) \equiv 1$, assuming the Hardy–Littlewood conjecture (this is done by considering the contribution to (4.13) below coming from those $(n_1, \dots, n_{2k})$ with $H > |n_i - n_j| > \max_i |a_i|$ for all $i \neq j$). Similarly, one can show $M_{2k} \gg XH^k (\log X)^{k\ell}$ (by considering the contribution to (4.13) below coming from those $(n_1, \dots, n_{2k})$ with $n_{k+i} = n_i$ for all $i \leq k$ and $H > |n_i - n_j| > \max_i |a_i|$). Lastly, one can show $M_{2k} \gg XH(\log X)^{(2k-1)\ell}$ (by considering the contribution to (4.13) below coming from those $(n_1, \dots, n_{2k})$ with $n_1 = \cdots = n_{2k}$).

*Proof.* Expanding out the definition of $M_{2k}$ using orthogonality, we obtain

$$M_{2k} = \int_0^X \sum_{\substack{x \leq n_1, \dots, n_{2k} \leq x+H \\ n_1 + \cdots + n_k = n_{k+1} + \cdots + n_{2k}}} g(n_1) \cdots g(n_k) \overline{g(n_{k+1}) \cdots g(n_{2k})} \prod_{\substack{j \leq \ell \\ i \leq 2k}} \Lambda(n_i + a_j) \mathrm{d}x, \tag{4.13}$$

and so $|g(n)| \leq 1$ implies

$$M_{2k} \leq \sum_{\substack{n_1, \dots, n_{2k} \leq X+H \\ n_1 + \cdots + n_k = n_{k+1} + \cdots + n_{2k}}} \prod_{\substack{j \leq \ell \\ i \leq 2k}} \Lambda(n_i + a_j) \cdot \int_0^X \mathbf{1}_{x \leq n_1 \leq x+H} \cdots \mathbf{1}_{x \leq n_{2k} \leq x+H} \mathrm{d}x. \tag{4.14}$$

For each $n_1, \dots, n_{2k} \leq X + H$ denote $M = \max_{1 \leq i,j \leq 2k} |n_i - n_j|$, so that

$$\int_0^X \mathbf{1}_{x \leq n_1 \leq x+H} \cdots \mathbf{1}_{x \leq n_{2k} \leq x+H} \mathrm{d}x = (H - M) \mathbf{1}_{0 \leq M \leq H} \ \leq \ H \mathbf{1}_{0 \leq M < H},$$

and if we let $n = n_{2k}$, $h_i = n_i - n$ (so $h_{2k} = 0$), then (4.14) becomes

$$\begin{aligned} M_{2k} \le\ & H \sum_{\substack{n_1,\dots,n_{2k} \le X+H \\ n_1+\cdots+n_k = n_{k+1}+\cdots+n_{2k} \\ \max_{i,j}|n_i - n_j| < H}} \prod_{\substack{j\le \ell \\ i \le 2k}} \Lambda(n_i + a_j) \\ \le\ & H \sum_{\substack{-H \le h_1,\dots,h_{2k-1} \le H \\ h_1+\cdots+h_k = h_{k+1}+\cdots+h_{2k-1}}} \sum_{n \le X+H} \prod_{\substack{1\le j\le \ell \\ 0 \le i \le 2k-1}} \Lambda(n + h_i + a_j). \end{aligned}$$

Write $\mathcal{A} = \{a_1, \dots, a_\ell\}$ and $\mathcal{H} = \{0, h_1, \dots, h_{2k-1}\}$ (note $h_{2k-1} = h_1 + \cdots + h_{k-1} - h_k - \cdots - h_{2k-2}$). Let $\mathcal{A} + \mathcal{H}$ denote the sumset, and observe that $|\mathcal{A} + \mathcal{H}| \in [\ell, 2k\ell]$. In what follows, for convenience of notation we denote $h_0 = 0$.

Now we partition the tuples $\mathcal{H}$ according to $|\mathcal{A}+\mathcal{H}|$, namely, for $m \in [0, (2k-1)\ell]$ denote

$$\mathrm{H}_m = \big\{(h_1, \dots, h_{2k-2}) \ :\ |h_i| \le H,\ |\mathcal{A} + \mathcal{H}| = 2k\ell - m\big\},$$

so that by Lemma 4.2.3,

$$M_{2k} \ll XH \sum_{m=0}^{(2k-1)\ell} (\log X)^m \sum_{(h_1,\dots,h_{2k-2}) \in \mathrm{H}_m} \mathfrak{S}\big(\mathcal{A} + \mathcal{H}\big). \tag{4.15}$$

Let $A = \max_{a\in\mathcal{A}} |a|$. We claim that

$$\big|\mathcal{I}\big| \ \ge\ \lceil m/\ell \rceil, \qquad \text{where} \quad \mathcal{I} := \big\{0 \le i \le 2k-1 :\ \exists i' > i :\ h_i - h_{i'} \in [-A, A]\big\}.$$

To show this, first note that all the sums $a_j + h_i$, $j \le \ell$, $i \notin \mathcal{I}$ are distinct, since if there was a coincidence $a + h_i = a' + h_{i'}$ for some $a, a' \in \mathcal{A}$, $i' > i$, then $h_i - h_{i'} = a' - a \in [-A, A]$ and so $i \in \mathcal{I}$. Thus we must have

$$2k\ell - m = |\mathcal{A} + \mathcal{H}| \ge (2k - |\mathcal{I}|)\ell,$$

so $|\mathcal{I}| \ge m/\ell$, as claimed. Now, for $I := |\mathcal{I}|$, we have a system of $I + 1$ linear inequalities constraining the vector $(h_1, \dots, h_{2k-1}) \in [-H, H]^{2k-1}$:

$$\begin{cases} h_1 + \cdots + h_{k-1} - h_k - \cdots - h_{2k-2} - h_{2k-1} = 0 \\ |h_i - h_{\sigma(i)}| \le A, \quad i \in \mathcal{I}, \end{cases}$$

where $h_0 = 0$, and $\sigma(i) \in [i+1, 2k-1]$ are some integers.

Let $I' = I + 1$ if $I \le k - 1$, and let $I' = I$ if $I \ge k$. Then by basic linear algebra the set of $I + 1$ linear forms above contains a subset of $I'$ linearly independent forms, so there exists a vector $\mathbf{u} = (u_1, \dots, u_{2k-1-I'}) \in [-H, H]^{2k-1-I'}$ (depending on $h_i$)

and linear forms $L_i : \mathbb{Z}^{2k-2-I'} \to \mathbb{Z}$ with bounded coefficients (and with $L_0 \equiv 0$) such that denoting $L_{i,j}(\mathbf{u}) = a_j + L_i(\mathbf{u})$ we have

$$\begin{aligned} \mathcal{H} &= \{L_0(\mathbf{u}) + O(1), \dots, L_{2k-1}(\mathbf{u}) + O(1)\} \\ \mathcal{A} + \mathcal{H} &= \{L_{i,j}(\mathbf{u}) + O(1) : 0 \le i \le 2k-1, j \le \ell\}. \end{aligned} \tag{4.16}$$

Thus by Corollary 4.2.5.1 and the fact that $I' \ge \lceil m/\ell \rceil + 1_{m \le (k-1)\ell}$, for each $m$ we have

$$\begin{aligned} \sum_{(h_1,\dots,h_{2k-2}) \in \mathrm{H}_m} \mathfrak{S}\big(\mathcal{A}+\mathcal{H}\big) &\le \sum_{\mathbf{u} \in [-H,H]^{2k-2-\lceil m/\ell \rceil + \mathbf{1}_{m>(k-1)\ell}}} \mathfrak{S}\big(\big(L_{i,j}(\mathbf{u}) + r\big)_{\substack{0 \le i \le 2k-1 \\ j \le \ell, r = O(1)}}\big) \\ &\ll H^{2k-2-\lceil m/\ell \rceil + \mathbf{1}_{m>(k-1)\ell}}. \end{aligned} \tag{4.17}$$

Hence the bound in (4.15) becomes

$$\begin{aligned} M_{2k} &\ll XH \sum_{m=0}^{(2k-1)\ell} (\log X)^m H^{2k-2-\lceil m/\ell \rceil + \mathbf{1}_{m>(k-1)\ell}} \\ &\ll X(H^{2k-1} + H^k (\log X)^{k\ell} + H(\log X)^{(2k-1)\ell}). \end{aligned}$$

Finally, the assumption $H \ge (\log X)^{\ell k/(k-1)}$ implies $M_{2k} \ll XH^{2k-1}$ as claimed. □

### 4.2.3 Correlations of primes and integers having typical factorization

In this subsection, we will prove Lemma 4.2.10, which upper bounds the correlations of the von Mangoldt function and the indicator of numbers having prime factors in prescribed intervals.

We first consider a class of multiplicative functions with 'moderate growth'.

**Definition 4.2.8.** Denote the set $\mathcal{M}$ of multiplicative $f : \mathbb{N} \to \mathbb{R}_{\ge 0}$ with

(i) $f(p^\nu) \le 2^\nu$ for all primes $p$, $\nu \ge 1$.

(ii) for all $\varepsilon > 0$ there exists $B = B(\varepsilon)$ such that $f(n) \le Bn^\varepsilon$ for all $n \ge 1$.

Now we give a special case of a general bound of Henriot [76] for the class $\mathcal{M}$. Henriot's result refines earlier work of Nair–Tenenbaum [108], and importantly, it is uniform in the discriminant.

**Lemma 4.2.9.** *Given $k \geq 1$ and a tuple $\mathcal{H} = \{h_1, \dots, h_k\} \subset [1, X]$, denote $\nu_p(\mathcal{H}) = |\{h \,(\mathrm{mod}\ p) : h \in \mathcal{H}\}|$. Then for any multiplicative functions $f_j \in \mathcal{M}$,*

$$\sum_{\sqrt{X} \leq n \leq X} \prod_{j=1}^{k} f_j(n + h_j) \ \ll_k \Delta_D \, X \prod_{p \leq \sqrt{X}} \Big(1 - \frac{\nu_p(\mathcal{H})}{p}\Big) \prod_{j=1}^{k} \sum_{n \leq \sqrt{X}} \frac{f_j(n)}{n} \tag{4.18}$$

*where $D = D(\mathcal{H}) = \prod_{i<j}(h_j - h_i)^2$, and*

$$\Delta_D = \prod_{p \mid D} \Bigg(1 \ + \sum_{\substack{0 \leq \nu_1, \dots, \nu_k \leq 1 \\ (\nu_1, \dots, \nu_k) \neq (0, \dots, 0)}} \frac{|\{n \,(\mathrm{mod}\ p^2) \ : \ p^{\nu_j} \| n + h_j \ \forall j\}|}{p^2} \prod_{j=1}^{k} f_j(p^{\nu_j}) \Bigg).$$

*In particular, if $|f(p)| \leq 1$ for all $p$, we have $\Delta_D \leq \prod_{p \mid D}(1 + k/p)$.*

*Proof.* This is [76, Theorem 3] in the special case of $x = \sqrt{X}$, $y = X$, $\delta = 1/(2k)$, with linear polynomials $Q_j(n) = n + h_j$, $Q(n) = \prod_{j=1}^{k} Q_j(n)$, and the multiplicative function $F(n_1, \dots, n_k) = \prod_{j=1}^{k} f_j(n_j)$. Note that the discriminant of the polynomial $Q$ is $\prod_{i<j}(h_i - h_j)^2 = D$ and the sum of coefficients is $\ll \prod_{j=1}^{k} h_j \ll X^k$.

Finally, if $|f(p)| \leq 1$ for all $p$, then

$$\Delta_D = \prod_{p \mid D} \Bigg(1 \ + \sum_{\substack{0 \leq \nu_1, \dots, \nu_k \leq 1 \\ (\nu_1, \dots, \nu_k) \neq (0, \dots, 0)}} \frac{|\{n \,(p^2) \ : \ p^{\nu_j} \| n + h_j \ \forall j\}|}{p^2} \Bigg) \leq \prod_{p \mid D} \Big(1 + \frac{k}{p}\Big),$$

since the count $|\{n \,(p^2) \ : \ p^{\nu_j} \| n + h_j \ \forall j\}|$ equals zero, unless $\nu_j = 1$ for a unique index $j \in [1, k]$ (and $\nu_i = 0$ for all $i \neq j$) in which case it equals $p$. □

We remark that the bound (4.18) is of the correct order of magnitude when the functions $f_j$ are not too small, e.g. $f_j(n) \geq \eta^{\Omega(n)}$ for some $\eta > 0$; see [76, Theorem 6].

As mentioned, we will need in Section 4.4 a simple upper bound for the correlations of the primes and integers with prescribed factorization patterns.

**Lemma 4.2.10.** *Let $\ell \geq 1$ be fixed and let $\mathcal{A} = \{a_1, \dots, a_\ell\}$ be distinct. Let $X \geq 2$ and $1 \leq h \leq X$ with $h \neq a_j$ for all $1 \leq j \leq \ell$. Let $\mathcal{I} \subset [1, X]$, and let $\mathcal{N}$ be the set of positive integers having no prime factors from $\mathcal{I}$. Then we have*

$$\sum_{n \leq X} \mathbf{1}_{\mathcal{N}}(n + h) \Lambda(n + a_1) \cdots \Lambda(n + a_\ell) \ll_{\mathcal{A}} X \prod_{j=1}^{\ell} \Bigg(\frac{|h - a_j|}{\varphi(|h - a_j|)}\Bigg)^{\ell+1} \prod_{p \in \mathcal{I}} \Big(1 - \frac{1}{p}\Big). \tag{4.19}$$

*Proof.* Let $P(z) = \prod_{p<z} p$. We may assume that the summation in (4.19) runs over $\sqrt{X} < n \leq X$, since the contribution of $n \leq \sqrt{X}$ is negligible. Therefore, the proof of (4.19) has been reduced to showing

$$\sum_{n\leq X} \mathbf{1}_{\mathcal{N}}(n+h) \prod_{j=1}^{\ell} \mathbf{1}_{(n+a_j, P(\sqrt{X})=1} \ll_{\mathcal{A}} X(\log X)^{-\ell} \prod_{j=1}^{\ell} \left( \frac{|h-a_j|}{\varphi(|h-a_j|)} \right)^{\ell+1} \prod_{p\in\mathcal{I}} \left(1-\frac{1}{p}\right). \tag{4.20}$$

Note that all the indicator functions above are multiplicative.

To this, denote $D = \prod_{1\leq i<j\leq \ell}(a_i - a_j) \prod_{j\leq \ell}(h-a_j)$. By Lemma 4.2.9, we have

$$\begin{aligned} &\sum_{n\leq X} \mathbf{1}_{\mathcal{N}}(n+h) \prod_{j=1}^{\ell} \mathbf{1}_{(n+a_j, P(\sqrt{X}))=1} \\ &\qquad \ll_{\mathcal{A}} \ \Delta_D X \prod_{p\leq X} \left(1-\frac{\ell+1}{p}\right) \sum_{n_0, n_1, \dots, n_\ell \leq X} \frac{\mathbf{1}_{\mathcal{N}}(n_0) \prod_{j=1}^{\ell} \mathbf{1}_{(n_j, P(\sqrt{X}))=1}}{n_0 n_1 \cdots n_\ell}, \end{aligned} \tag{4.21}$$

where $\Delta_D$ is crudely bounded as

$$\Delta_D \leq \prod_{p|D} \left(1+\frac{\ell+1}{p}\right) \ll_A \prod_{j=1}^{\ell} \left( \frac{|h-a_j|}{\varphi(|h-a_j|)} \right)^{\ell+1}.$$

Now (4.20) follows by noting that by Euler products and Mertens' theorem [104],

$$\begin{aligned} \sum_{n\leq X} \frac{\mathbf{1}_{\mathcal{N}}(n)}{n} &\leq \prod_{\substack{p\leq X \\ p\notin\mathcal{I}}} \left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right) = \prod_{p\leq X} \left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right) \prod_{p\in\mathcal{I}} \left(1-\frac{1}{p}\right) \\ &\ll (\log X)^{-\ell} \prod_{p\leq X} \left(1-\frac{\ell+1}{p}\right)^{-1} \prod_{p\in\mathcal{I}} \left(1-\frac{1}{p}\right), \end{aligned}$$

as well as

$$\sum_{n\leq X} \frac{\mathbf{1}_{(n,P(\sqrt{X}))=1}}{n} = 1 + \sum_{\sqrt{X}<p\leq X} \frac{1}{p} \asymp 1. \qquad \square$$

## 4.3 A Fourier-analytic argument

Our main theorems will be a consequence of the propositions in the previous section and Proposition 4.3.2 below on correlations on average under suitable hypotheses. We begin by formulating the necessary hypotheses.

In the rest of the chapter, given a function $f : \mathbb{N} \to \mathbb{C}$, let $E_f$ be a convenient upper bound for the average of $f$, so that

$$\frac{1}{X}\sum_{n\le X}|f(n)| \le E_f(X).$$

For example, we may simply take $E_f = O(1)$ if $f$ is bounded, or if $f = \Lambda$.

**Definition 4.3.1.** Let $X, H, C, p > 1$ and $\eta > 0$, and let $f : \mathbb{N} \to \mathbb{C}$. We say that $f$ satisfies hypothesis $\mathsf{H}_1(X, H, C, p, E_f)$ if $\sum_{n\le X}|f(n)| \le X\,E_f(X)$ and

$$\int_0^X \int_0^1 \Big|\sum_{x\le n\le x+2H} f(n)e(\alpha n)\Big|^p \,\mathrm{d}\alpha\,\mathrm{d}x \le CH^{p-1}XE_f(X)^p.$$

We say that $f$ satisfies hypothesis $\mathsf{H}_2(X, H, \eta)$ if

$$\sup_\alpha \int_0^X \Big|\sum_{x\le n\le x+2H} f(n)e(\alpha n)\Big|^2 \,\mathrm{d}x \le \eta H^2 XE_f(X)^2.$$

**Proposition 4.3.2.** *Let $X \ge H \ge 2$. Let $C \ge 1$, $p \ge 2$ and $\eta > 0$. Let $f : \mathbb{N} \to \mathbb{C}$ be any function supported on $[0, X]$ and satisfying $\mathsf{H}_2(X, H, \eta)$, and let $g : \mathbb{N} \to \mathbb{C}$ be any function satisfying $\mathsf{H}_1(X, H, C, p, E_g)$. Then we have*

$$\sum_{h\le H}\Big|\sum_{n\le X} f(n+h)g(n)\Big| \ll (\eta C)^{1/p}HXE_{|f|+|f|^2}(X)E_g(X).$$

Note that if $|f(n)| \le 1$, we can simply take $E_{|f|+|f|^2}(X) = O(E_{|f|}(X))$.

*Proof.* Let us denote

$$S_{f,g} := \sum_{h\le H}\Big|\sum_{n\le X} f(n+h)g(n)\Big|.$$

Then the task is to show that if $f$ satisfies $\mathsf{H}_2(X, H, \eta)$ and $g$ satisfies $\mathsf{H}_1(X, H, C, p, E_g)$, then

$$S_{f,g} \ll \eta^{1/p}C^{1/p}HXE_{|f|+|f|^2}(X)E_g(X).$$

For proving this, we first note that

$$S_{f,g} \le \frac{1}{H}\sum_{h\le 2H}(2H-h)\Big|\sum_{n\le X} f(n+h)g(n)\Big|.$$

We introduce unimodular coefficients $c(h)$ to denote the phase of $\sum_{n\le X} f(n+h)g(n)$, so that

$$\begin{aligned}|S_{f,g}| &\le \frac{1}{H}\sum_{h\le 2H}(2H-h)c(h)\sum_{n\le X} f(n+h)g(n)\\ &= \frac{1}{H}\sum_{h\le 2H} c(h)\sum_{n\le X+2H} f(n)\sum_{m\le X} g(m)\mathbf{1}_{n=m+h}\cdot\int_0^X \mathbf{1}_{x\le n,m\le x+2H}\mathrm{d}x\\ &= \frac{1}{H}\int_0^X\int_0^1 \sum_{h\le 2H} c(h)e(h\alpha)\sum_{x\le n,m\le x+2H} f(n)g(m)e\big((m-n)\alpha\big)\mathrm{d}\alpha\mathrm{d}x,\end{aligned}$$

where we used the orthogonality relation $\mathbf{1}_{n=0}=\int_0^1 e(n\alpha)\mathrm{d}\alpha$. Therefore, we have the upper bound

$$|S_{f,g}| \ \le \frac{1}{H}\left|\int_0^X\int_0^1 C_0(\alpha)F_x(-\alpha)G_x(\alpha)\mathrm{d}\alpha\mathrm{d}x\right|, \tag{4.22}$$

where in the triple convolution integral the three exponential sums are

$$\begin{aligned}C_0(\alpha) &:= \sum_{h\le 2H} c(h)e(h\alpha),\\ F_x(\alpha) &:= \sum_{x\le n\le x+2H} f(n)e(n\alpha),\\ G_x(\alpha) &:= \sum_{x\le m\le x+2H} g(m)e(m\alpha).\end{aligned}$$

By Fubini's theorem, the triangle inequality and Cauchy–Schwarz, from (4.22) we deduce that

$$|S_{f,g}| \le \frac{1}{H}\int_0^1 |C_0(\alpha)|\left(\int_0^X |F_x(-\alpha)|^2\mathrm{d}x\right)^{\frac12}\left(\int_0^X |G_x(\alpha)|^2\mathrm{d}x\right)^{\frac12}\mathrm{d}\alpha.$$

Let $q$ satisfy $1/p+1/q=1/2$. By taking supremum of the integral with $F_x(-\alpha)$ to the power $1/p$ and applying Hölder's inequality with exponents $(2,q,p)$, we obtain

$$\begin{aligned}|S_{f,g}| &\le \frac{1}{H}\sup_\alpha\left(\int_0^X |F_x(-\alpha)|^2\mathrm{d}x\right)^{\frac1p}\left(\int_0^1 |C_0(\alpha)|^2\mathrm{d}\alpha\right)^{\frac12}\\ &\qquad\cdot\left(\int_0^1\int_0^X |F_x(-\alpha)|^2\mathrm{d}x\mathrm{d}\alpha\right)^{\frac1q}\left(\int_0^1\left(\int_0^X |G_x(\alpha)|^2\mathrm{d}x\right)^{\frac p2}\mathrm{d}\alpha\right)^{\frac1p}\\ &=: \frac{1}{H}I_0^{1/p}I_1^{1/2}I_2^{1/q}I_3^{1/p},\end{aligned}$$

say. Since $f$ satisfies $\mathsf{H}_2(X,H,\eta)$, we have

$$I_0 \le \eta H^2 X E_f(X)^2.$$

By Parseval, we have

$$I_1 \le 2H, \quad I_2 \le 2H \sum_{n \le X+2H} |f(n)|^2 \le 2HXE_{|f|^2}(X).$$

Lastly, using Hölder's inequality and the fact that $g$ satisfies $\mathsf{H}_1(X, H, C, p, E_g)$, we have

$$I_3 \le X^{p/2-1} \int_0^1 \int_0^X |G_x(\alpha)|^p \mathrm{d}x \mathrm{d}\alpha \le CH^{p-1}X^{p/2}E_g(X)^p.$$

Combining the bounds, we see that

$$\begin{aligned} |S_{f,g}| &\ll H^{-1}\big(\eta H^2 X E_f(X)^2\big)^{1/p} H^{1/2}\big(HXE_{|f|^2}(X)^2\big)^{1/q}\big(CH^{p-1}X^{p/2}E_g(X)^p\big)^{1/p} \\ &\ll (\eta C)^{1/p} HXE_{|f|+|f|^2}(X)E_g(X), \end{aligned}$$

as desired. □

## 4.4 Proofs of the main theorems

We are now ready to prove Theorem 4.1.5. As already discussed in Remark 4.1.6, Theorem 4.1.2(i) concerning the Möbius case follows immediately as a special case. As we shall see, also Theorem 4.1.2(ii) follows with essentially the same proof.

*Proof of Theorem 4.1.5.* Fix $\varepsilon \in (0,1)$ and $k, \ell \ge 1$, and let $2 \le (\log X)^{\ell+\varepsilon} \le H \le \exp((\log X)^{1/1000})$. Let

$$\delta = 10^4 \, \frac{\ell}{\varepsilon} \frac{\log\log H}{\log H}.$$

Set

$$f(n) = f_1(n)\mathbf{1}_{\mathcal{S}}(n)\mathbf{1}_{[0,X]}(n), \quad \text{and} \quad g(n) = \prod_{i=2}^{k} f_i(n+h_i) \prod_{j=1}^{\ell} \Lambda(n+a_j),$$

where $\mathcal{S}$ is the set of positive integers having a prime factor in each of the intervals $[P_i, Q_i]$, with $P_i, Q_i$ as in Proposition 4.2.1 and $\delta$ as above. Then by Lemma 4.2.10 (and the fact that $\mathcal{S}^c \subset \bigcup_{j \le J} \mathcal{S}_j^c$, where $\mathcal{S}_j^c$ is the set of integers having no prime factors in $[P_j, Q_j]$) for all $X \ge 2$ we have

$$\begin{aligned} \Big| \sum_{\substack{n \le X \\ n+h \notin \mathcal{S}}} f_1(n+h)g(n) \Big| &\le \sum_{n \le X} \mathbf{1}_{\mathcal{S}^c}(n+h) \prod_{j=1}^{\ell} \Lambda(n+a_j) \\ &\ll \sum_{j \le J} X \prod_{p \in [P_j, Q_j]} \Big(1 - \frac{1}{p}\Big) \prod_{j=1}^{\ell} \Big( \frac{|h-a_j|}{\varphi(|h-a_j|)} \Big)^{2^{\ell+1}}. \end{aligned} \tag{4.23}$$

By (4.6) we have

$$\sum_{j\le J}\prod_{p\in[P_j,Q_j]}\left(1-\frac{1}{p}\right)\ll\sum_{j\le J}\frac{\log P_j}{\log Q_j}\ll\frac{\log P_1}{\log Q_1}\ll\delta+\exp(-M(f;X,Q)/2000)=:\delta'.$$

Also by Lemma 4.2.5, we have

$$\sum_{h\le H}\prod_{j=1}^{\ell}\left(\frac{|h-a_j|}{\varphi(|h-a_j|)}\right)^{2^{\ell+1}}\ll H,$$

and so plugging back into (4.23) gives

$$\sum_{h\le H}\Big|\sum_{\substack{n\le X\\ n+h\notin\mathcal{S}}}f_1(n+h)g(n)\Big|\ \ll\ \delta' XH. \tag{4.24}$$

Let

$$p:=2+2\lceil\ell/\varepsilon\rceil.$$

In view of (4.24), to prove Theorem 4.1.5 it suffices to show that

$$\sum_{h\le H}\Big|\sum_{n\le X}f(n+h)g(n)\Big|\ll\left(\frac{\log\log H}{\log H}+\exp(-M(f;X,Q)/(2000p))\right)HX, \tag{4.25}$$

since we can crudely estimate $2000p<10000\ell/\varepsilon$. Note that by Lemma 4.2.3, we can take

$$E_g(X)\asymp 1,$$

and we can trivially take $E_f(X), E_{|f|+|f|^2}(X)\asymp 1$, so it suffices to prove (4.25) with an extra factor of $E_{|f|+|f|^2}(X)E_g(X)$ on the right-hand side.

By Proposition 4.2.6 (and the fact that $p$ is even and satisfies $\ell+\varepsilon>\ell p/(p-2)$), there exists a constant $B=B_{\ell,\varepsilon}>0$ for which

$$\int_0^1\int_0^X\Big|\sum_{x\le n\le x+2H}e(\alpha n)\prod_{i=2}^{k}f_i(n+h_i)\prod_{j=1}^{\ell}\Lambda(n+a_j)\Big|^p\,\mathrm{d}x\mathrm{d}\alpha\ \le\ BXH^{p-1}.$$

Therefore, $g$ satisfies hypothesis $\mathsf{H}_1(X,H,B',p,E_g)$ for some $B'\ll 1$.

We also note that by Proposition 4.2.1 $f$ satisfies $\mathsf{H}_2(X,H,\eta)$ with

$$\eta\ll H^{-\delta}+\exp(-M(f;X,Q)/2000).$$

Now, applying Proposition 4.3.2 with the choices of $f, g$ above, we obtain

$$\sum_{h\le H}\Big|\sum_{n\le X} f_1\mathbf{1}_{\mathcal{S}}(n+h)\prod_{i=2}^{k} f_i(n+h_i)\prod_{j=1}^{\ell}\Lambda(n+a_j)\Big| \\ \ll \eta^{1/p}HX \ \ll \ (H^{-\delta/p}+\exp(-M(f;X,Q)/(2000p)))HX. \tag{4.26}$$

Thus since $\delta = 10^4\ell\varepsilon^{-1}(\log\log H)/(\log H)$, this gives (4.25), as desired. The proof of Theorem 4.1.5 (and hence of Theorem 4.1.2(i)) is now complete □

*Proof of Theorem 4.1.2(ii).* Let $\delta$ be as in the theorem, so that in particular $\delta \le c/C$ with $c > 0$ be small enough. Take $c = 1/10000$. Applying (4.26) and Markov's inequality we see that

$$\Big|\sum_{n\le X}\mu\mathbf{1}_{\mathcal{S}}(n+h)\prod_{i=2}^{k}\mu(n+h_i)\prod_{j=1}^{\ell}\Lambda(n+a_j)\Big| \\ \ll (H^{-\delta/(2p)}+H^{\delta/(2p)}\exp(-M(\mu;X,Q)/(2000p)))HX$$

for all but $\ll H^{1-\delta/(2p)}$ integers $h \le H$. By (3.7) we have $M(\mu;X,Q) \ge \frac{1}{4}\log\log X + O(1)$, so we see that

$$H^{\delta/(2p)}\exp(-M(\mu;X,Q)/(2000p)) \ll H^{-\delta/(4p)} \ll \delta,$$

and the claim follows. □

# Chapter 5

# Mertens' prime product formula, dissected

## 5.1 Introduction

We begin with the Euler-Mascheroni constant $\gamma = 0.57721\cdots$, defined as the limit of the difference between the harmonic series up to $x$ and $\log x$. The ubiquitous constant $\gamma$ crops up in many contexts, notably, in the 3rd of three results from a celebrated paper of Mertens [104] on the distribution of prime numbers.

As notation, throughout we write $f(x) = O(g(x))$ and $f(x) \ll g(x)$ to mean $|f(x)/g(x)|$ is bounded, while $f(x) \sim g(x)$ means $\lim_{x\to\infty} f(x)/g(x) = 1$. Also, let $\log_2 x = \log\log x$, and let $p$ denote a prime number.

**Theorem 5.1.1** (Mertens, 1874)**.** *There exists a constant $\beta > 0$ for which*

$$\sum_{p\le x} \frac{\log p}{p} = \log x + O(1), \qquad \sum_{p\le x} \frac{1}{p} = \log_2 x + \beta + O\Big(\frac{1}{\log x}\Big) \tag{5.1}$$

$$\prod_{p\le x} \Big(1 - \frac{1}{p}\Big)^{-1} \sim e^{\gamma} \log x. \tag{5.2}$$

Here $\beta = 0.26149\cdots$ is Mertens' constant, which is known to satisfy

$$\beta - \gamma = \sum_{p} \big(\log(1 - \tfrac{1}{p}) + \tfrac{1}{p}\big) = -\sum_{p}\sum_{j\ge 2} \frac{p^{-j}}{j} = -\sum_{j\ge 2} \frac{P(j)}{j}, \tag{5.3}$$

where $P(s) = \sum_p p^{-s}$ denotes the prime zeta function, for $s > 1$. See for instance Theorem 2.7 in [106, p.50].

Now by expanding Mertens' prime product in (5.2), we have

$$\prod_{p\le x} \Big(1 - \frac{1}{p}\Big)^{-1} = \prod_{p\le x} \Big(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\Big) = \sum_{P^+(n)\le x} \frac{1}{n} \tag{5.4}$$

where $P^+(n)$ denotes the largest prime factor of $n$.

Consider "dissecting" the sum in (5.4) according to the number of prime factors of $n$ with multiplicity, denoted $\Omega(n)$. Our main result is an asymptotic formula for this dissected sum.

**Theorem 5.1.2.** *For each fixed $k \ge 1$, we have*

$$\sum_{\substack{\Omega(n)=k\\ P^+(n)\le x}} \frac{1}{n} = \sum_{j=0}^{k} \frac{c_{k-j}}{j!} \big(\log_2 x + \beta\big)^j + O_k\Big(\frac{(\log_2 x)^{k-1}}{\log x}\Big) \tag{5.5}$$

*where the sequence $(c_k)_{k=0}^{\infty}$ is recursively defined by $c_0 = 1$ and*

$$c_k = \frac{1}{k} \sum_{j=2}^{k} c_{k-j}\, P(j). \tag{5.6}$$

Theorem 5.1.2 may be viewed as a "dissection" of Mertens' prime product formula. Indeed, as shown later in (5.18), the main term $e^\gamma \log x$ in (5.2) may be expressed as the series over all $k \geq 1$ of the the main terms in (5.5) (i.e. the sum over $j \leq k$).[1]

We note this terminology was introduced by Pollack [115], who dissected a classical mean value theorem of Hall and Tenenbaum.

### 5.1.1 Uniform estimates via complex analysis

Classically, the analogous series to (5.5) has been studied, replacing the condition $P^+(n) \leq x$ with the more common $n \leq x$.

Mertens' 1st theorem implies, by induction on each fixed $k \geq 1$,

$$\sum_{\substack{\Omega(n)=k \\ n \leq x}} \frac{1}{n} \sim \frac{(\log_2 x)^k}{k!} \tag{5.7}$$

as $x \to \infty$, see [106, p.228]. Note (5.7) historically attributed to Landau [85]. This is another example of dissection, as the sum over all $k$ of each side gives $\sum_{n \leq x} \frac{1}{n}$ and $\log x$, respectively. We also note that the asymptotic (5.7) also holds with $\Omega(n)$ replaced by $\omega(n)$, the number of distinct prime factors of $n$.

However, (5.7) only holds for fixed $k$. The celebrated theorem of Sathe and Selberg implies the following uniform estimate for $k$ less than $2 \log_2 x$.

**Theorem 5.1.3** (Sathe–Selberg). *Define* $\nu(z) = \frac{1}{\Gamma(z+1)} \prod_p (1 - \frac{z}{p})^{-1}(1 - \frac{1}{p})^z$, *and let* $r = k/\log_2 x$. *For any* $\varepsilon > 0$, *as* $x \to \infty$ *we have uniformly for* $r \leq 2 - \varepsilon$,

$$\sum_{\substack{\Omega(n)=k \\ n \leq x}} \frac{1}{n} \sim \nu(r) \frac{(\log_2 x)^k}{k!}. \tag{5.8}$$

To see this, [106, Theorem 7.19] or [137, Theorem 6.5] gives an asymptotic in the stated range

$$\sum_{\substack{\Omega(n)=k \\ n \leq x}} 1 = \nu(r) \frac{x}{\log x} \frac{(\log_2 x)^{k-1}}{(k-1)!} \Big(1 + O_\varepsilon\big(\tfrac{k}{\log_2 x}\big)\Big).$$

Then (5.8) follows by partial summation, combined with e.g. the Erdős-Sarközy upper bound $O(k^4 2^{-k} x \log x)$ uniformly for all $x, k \geq 1$, see [45].

[1] Here "dissection" is meant to highlight the formal compatibly of main terms. Whereas the estimate (5.5) itself does not necessarily hold uniformly over all $k \geq 1$.

*Remark* 5.1.4. As $\nu(r) = 1$ only when $r = 0, 1$, Landau's estimate (5.7) holds if and only if $k = o(\log_2 x)$ or $k = (1 + o(1)) \log_2 x$.

*Remark* 5.1.5. [137, Theorem 6.4] gives an analogous result with $\Omega(n)$ replaced by $\omega(n)$, by substituting the function $\nu(z)$ above with $\lambda(z) = \frac{1}{\Gamma(z+1)} \prod_p (1 + \frac{z}{p-1})(1 - \frac{1}{p})^z$.

*Remark* 5.1.6. The Sathe–Selberg theorem is proved through contour integration in the complex plane. Recently, Popa [118, 119] and Tenenbaum [136] have obtained results by similar analytic methods, for a generalized series that replaces the conditions $\Omega(n) = k$ and $n \leq x$ by the condition $p_1 \cdots p_k \leq x$ over $k$ independent prime variables. Or equivalently, they weight $n$ by the number of its ordered prime factorizations.

The uniformity coming from sophisticated analytic tools exemplifies the larger tension within mathematics, between proving the strongest results and using the simplest arguments. Of particular interest historically is the case $k = 1$, i.e. the Prime Number Theorem. Hadamard and de la Vallée Poussin initially gave proofs in 1896 using complex analysis, and for decades many believed it impossible to prove by elementary means. It came as a great shock when Selberg and Erdős did so in 1948. For an intriguing historical account, see Spencer and Graham [66].

As such, we emphasize that in Theorem 5.1.2, our particular conditions $\Omega(n) = k, P^+(n) \leq x$ in (5.5) are directly amenable to elementary methods when $k$ is fixed. Nevertheless, applying analytic tools to (5.5) do lend the advantage of uniformity in $k < (2 - \varepsilon) \log_2 x$.

**Theorem 5.1.7.** *Let $r = k/\log_2 x$. For any $\varepsilon > 0$, as $x \to \infty$ we have uniformly for $r \leq 2 - \varepsilon$,*

$$\sum_{\substack{\Omega(n)=k \\ P^+(n)\leq x}} \frac{1}{n} \sim \nu(r) e^{r\gamma} \Gamma(r+1) \frac{(\log_2 x)^k}{k!}. \tag{5.9}$$

Hence by comparision with the Sathe–Selberg theorem, we obtain the following elegant relation between sums over $P^+(n) \leq x$ with those over $n \leq x$.

**Corollary 5.1.7.1.** *Let $r = k/\log_2 x$. For any $\varepsilon > 0$, as $x \to \infty$ we have uniformly for $r \leq 2 - \varepsilon$,*

$$\sum_{\substack{\Omega(n)=k \\ P^+(n)\leq x}} \frac{1}{n} \sim e^{r\gamma} \Gamma(r+1) \sum_{\substack{\Omega(n)=k \\ n\leq x}} \frac{1}{n}. \tag{5.10}$$

*Remark* 5.1.8. One may prove an analogous result for $\omega(n)$, with the same factor $e^{r\gamma} \Gamma(r+1)$.

Note the factor $e^{r\gamma}\,\Gamma(r+1) = 1$ if and only if $r = 0$. Hence Corollary 5.1.7.1 implies

$$\sum_{\substack{\Omega(n)=k \\ P^+(n)\le x}} \frac{1}{n} \sim \sum_{\substack{\Omega(n)=k \\ n\le x}} \frac{1}{n} \tag{5.11}$$

if and only if $k$ is in the uniform range $k = o(\log_2 x)$. This is an example of friable regularity, in the following sense. Recall an integer $n$ with $P^+(n) \le x$ is called $x$-smooth or $x$-friable.

**Definition 5.1.9.** A sequence $(a_n)_{n\in\mathbb{N}}$ is *friably regular* [2] if $\sum_{n\le x} a_n \sim \sum_{P^+(n)\le x} a_n$ as $x \to \infty$.

For example, the friable regularity of $(\mu(n)/n)_{n\in\mathbb{N}}$ is equivalent to the prime number theorem. We also extend the definition to families of sequences.

**Definition 5.1.10.** A one-parameter family $(a_{n,x})_{n\in\mathbb{N}}$, indexed by $x \in \mathbb{R}$, is *friably regular* if $\sum_{n\le x} a_{n,x} \sim \sum_{P^+(n)\le x} a_{n,x}$ as $x \to \infty$.

In particular, Corollary 5.1.7.1 implies the family $(\mathbf{1}_{\Omega(n)=k}/n)_{n\in\mathbb{N}}$, indexed by $k = k(x)$, is friably regular if and only if $k = o(\log_2 x)$.

### 5.1.2 The coefficients $c_k$

Finally, we emphasize an important feature of the combinatorial approach in Theorem 5.1.2. The recursion in (5.6) enables rapid computation of the coefficients $c_k$ to high precision, the first few displayed below.

| $k$ | $c_k$ | $k$ | $c_k$ |
|---|---|---|---|
| $0, 1$ | $1, 0$ | 6 | $0.0108213\cdots$ |
| 2 | $0.226123\cdots$ | 7 | $0.0054110\cdots$ |
| 3 | $0.058254\cdots$ | 8 | $0.0027375\cdots$ |
| 4 | $0.044814\cdots$ | 9 | $0.0013752\cdots$ |
| 5 | $0.020323\cdots$ | 10 | $0.0006903\cdots$ |

At first glance, one might not expect the coefficients $c_k$ arising from (5.6) to exhibit any particular structure. However, the combinatorial approach shows $c_k$ to satisfy exponentially precise asymptotics.

**Theorem 5.1.11.** *The coefficients satisfy* $c_k = \eta\, 2^{-k} + O(3^{-k})$. *Here the constant* $\eta$ *is given by* $\eta = e^{-1}\prod_{p>2}(1-\frac{2}{p})^{-1}e^{-2/p} = 0.71206\cdots$.

[2] This extends the notion of friable regularity as in [35],[60], from equality of limits of convergent series to asymptotic equality of (possibly non-convergent) partial sums.

## 5.2 Elementary combinatorial proof for $k$ fixed

In this section we prove Theorem 5.1.2. For $x, s > 0$ define the (truncated) zeta functions

$$P_k(s,x) = \sum_{\substack{\Omega(n)=k \\ P^+(n)\le x}} n^{-s}, \qquad P(s,x) = P_1(s,x) = \sum_{p\le x} p^{-s}.$$

We first express $P_k(s,x)$ in terms of $P(s,x)$.

**Proposition 5.2.1.** *For each $k \ge 1$ and any $x, s > 0$ we have the identity*

$$P_k(s,x) = \sum_{n_1+2n_2+\cdots=k} \prod_{j\ge 1} \frac{1}{n_j!}\big(P(js,x)/j\big)^{n_j} \tag{5.12}$$

*where the sum ranges over all partitions of $k$.*

*Proof.* For any $x, s > 0$ we have a formal power series identity in $z$,

$$\sum_{k\ge 0} P_k(s,x) z^k = \sum_{P^+(n)\le x} \frac{z^{\Omega(n)}}{n^s} = \prod_{p\le x} \Big(1 + \frac{z}{p^s} + \frac{z^2}{p^{2s}} + \cdots\Big) = \prod_{p\le x} \Big(1 - \frac{z}{p^s}\Big)^{-1}$$

since the function $n \mapsto z^{\Omega(n)}/n^s$ is completely multiplicative. Thus expanding Taylor series,

$$\begin{aligned}\sum_{k\ge 0} P_k(s,x) z^k &= \exp\Big(-\sum_{p\le x} \log(1 - zp^{-s})\Big) = \exp\Big(\sum_{p\le x}\sum_{j\ge 1} \frac{(zp^{-s})^j}{j}\Big) \\ &= \exp\Big(\sum_{j\ge 1} \frac{P(js,x)}{j} z^j\Big) = \prod_{j\ge 1} \exp\Big(\frac{P(js,x)}{j} z^j\Big) \\ &= \prod_{j\ge 1}\sum_{n_j\ge 0} \frac{1}{n_j!}\Big(\frac{P(js,x)}{j} z^j\Big)^{n_j} = \sum_{k\ge 0} z^k \sum_{n_1+2n_2\cdots=k} \prod_{j\ge 1} \frac{1}{n_j!}\big(P(js,x)/j\big)^{n_j}.\end{aligned} \tag{5.13}$$

Now (5.12) follows by comparing the coefficients of $z^k$. □

*Remark* 5.2.2. This proposition generalizes [86, Proposition 3.1].

Next, the recursion (5.6) for $c_k$ leads to the explicit formula,

$$c_k = \sum_{2n_2+3n_3\cdots=k} \prod_{j\ge 2} \frac{1}{n_j!}\big(P(j)/j\big)^{n_j} \tag{5.14}$$

by the following lemma, for the choices $A_1 = 0$ and $A_j = P(j)$ when $j \ge 2$.

**Lemma 5.2.3.** *Given any sequence $(A_k)_{k=1}^{\infty}$, the sequence $(b_k)_{k=0}^{\infty}$ is given recursively by $b_0 = 1$ and $b_k = \frac{1}{k}\sum_{j=1}^{k} b_{k-j}A_j$, if and only if $(b_k)_{k=0}^{\infty}$ is given explicitly as*

$$b_k = \sum_{n_1+2n_2+\cdots=k} \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!}.$$

Note the (unique) partition of $k = 0$ has $n_j = 0$ for all $j$, so $b_0 = \prod_{j\geq 1}(A_j/j)^0/0! = 1$.

*Proof.* We prove the forward direction by induction on $k$ (the reverse direction is similar). For $k = 1$, we have $b_1 = b_0A_1 = A_1$. Then assuming the claim for each $r < k$,

$$\begin{aligned} kb_k &= \sum_{r=1}^{k} b_{k-r}A_r = \sum_{r=1}^{k} A_r \sum_{n_1+\cdots=k-r} \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!} \\ &= \sum_{r=1}^{k} \sum_{n_1+\cdots=k-r} \frac{A_r^{n_r+1}}{r^{n_r}n_r!} \prod_{j\neq r} \frac{(A_j/j)^{n_j}}{n_j!} = \sum_{r=1}^{k} \sum_{\substack{n_1+\cdots=k \\ n_r\geq 1}} rn_r \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!} \\ &= \sum_{n_1+\cdots=k} \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!} \sum_{\substack{1\leq r\leq k \\ n_r\geq 1}} rn_r = k \sum_{n_1+\cdots=k} \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!} \end{aligned}$$

In the last step, we dropped the condition $n_r \geq 1$ (since $rn_r = 0$ for $n_r = 0$) which gives $\sum_{r=1}^{k} rn_r = k$. Dividing by $k$ completes the induction. □

Now equipped with Proposition 5.2.1 and formula (5.14) for $c_k$, we now prove Theorem 5.1.2.

*Proof of Theorem 5.1.2.* For $P(j,x)$ with $j \geq 2$, we trivially bound $\sum_{p>x} p^{-j}$ by $x^{2-j}\sum_{n>x} n^{-2} = O(x^{1-j})$, which gives

$$P(j,x) := \sum_{p\leq x} p^{-j} = P(j) - \sum_{p>x} p^{-j} \;=\; P(j) + O(x^{1-j}) \qquad \text{for } j \geq 2.$$

Thus plugging into the identity for $P_k(1,x)$, Proposition 5.2.1 with $s = 1$ gives

$$P_k(1,x) \;=\; \sum_{n_1+2n_2+\cdots=k} \frac{P(1,x)^{n_1}}{n_1!} \prod_{j\geq 2} \frac{1}{n_j!}\Big(\frac{P(j)+O(x^{1-j})}{j}\Big)^{n_j}$$

For any partition of $k$, the binomial theorem implies $\prod_{j\geq 2} \frac{1}{n_j!}\Big([P(j)+O(x^{1-j})]/j\Big)^{n_j}$ equals $\prod_{j\geq 2} \frac{1}{n_j!}(P(j)/j)^{n_j}$ at negligible cost $O_k(1/x)$. Thus

$$P_k(1,x) = \sum_{n_1+2n_2+\cdots=k} \frac{P(1,x)^{n_1}}{n_1!} \prod_{j\geq 2} \frac{\big(P(j)/j\big)^{n_j}}{n_j!} \;+\; O_k(\tfrac{P(1,x)^k}{x}). \tag{5.15}$$

Then for $P(1,x)$, we recall Mertens' 2nd theorem

$$P(1,x) := \sum_{p\le x}\frac{1}{p} \;=\; \log_2 x + \beta + E(x), \qquad \text{with } E(x) = O\big(\tfrac{1}{\log x}\big) \tag{5.16}$$

so plugging in above gives

$$\begin{aligned} P_k(1,x) \;&=\; \sum_{n_1+2n_2+\cdots=k} \frac{1}{n_1!}\Big(\log_2 x+\beta+E(x)\Big)^{n_1}\prod_{j\ge 2}\frac{\big(P(j)/j\big)^{n_j}}{n_j!} \;+\; O_k\big(\tfrac{(\log_2 x)^k}{x}\big) \\ &=\; \sum_{n_1=0}^{k}\frac{(\log_2 x+\beta)^{n_1}}{n_1!}\sum_{2n_2+\cdots=k-n_1}\prod_{j\ge 2}\frac{\big(P(j)/j\big)^{n_j}}{n_j!} \;+\; O_k\big(E(x)\,(\log_2 x)^{k-1}\big) \end{aligned} \tag{5.17}$$

again by the binomial theorem. Here we used $\prod_{j\ge 2}\frac{1}{n_j!}(P(j)/j)^{n_j} = O_k(1)$.

Now recalling (5.14) and $E(x) = O(1/\log x)$ gives Theorem 5.1.2. □

From here, we may "dissect" Mertens' 3rd theorem. Indeed by (5.4),

$$\prod_{p\le x}\big(1-\tfrac{1}{p}\big)^{-1} = \sum_{P^+(n)\le x}\frac{1}{n} = \sum_{k\ge 0}P_k(1,x)$$

and using the asymptotic formula for $P_k(1,x)$ from Theorem 5.1.2,

$$\sum_{k\ge 0}\sum_{j=0}^{k}\frac{c_{k-j}}{j!}(\log_2 x+\beta)^j = \sum_{j\ge 0}\frac{(\log_2 x+\beta)^j}{j!}\sum_{k\ge j}c_{k-j} = e^{\beta}\log x\sum_{m\ge 0}c_m \;=\; e^{\gamma}\log x, \tag{5.18}$$

as desired, provided $\sum_{m\ge 0}c_m = e^{\gamma-\beta}$. This follows in turn by (5.14),

$$\begin{aligned} \sum_{m\ge 0}c_m &= \sum_{m\ge 0}\sum_{2n_2+3n_3\cdots=m}\prod_{j\ge 2}\frac{1}{n_j!}\big(P(j)/j\big)^{n_j} = \prod_{j\ge 2}\sum_{n_j\ge 0}\frac{1}{n_j!}\big(P(j)/j\big)^{n_j} \\ &= \prod_{j\ge 2}\exp\big(P(j)/j\big) = \exp\Big(\sum_{j\ge 2}\frac{P(j)}{j}\Big) \;=\; e^{\gamma-\beta} \end{aligned}$$

recalling (5.3). This shows the claim.

## 5.3 Combinatorial proof of asymptotics for coefficients $c_k$

In this section we prove a strengthening of Theorem 5.1.11. To this, we first rephrase the recursion (5.6) for $c_k$.

Let $A_1 = 0$ and $A_k = \sum_p p^{-k}$ for $k \geq 2$. Then $c_k$ is recursively defined by $c_0 = 1$ and

$$kc_k = \sum_{j=1}^{k} c_{k-j} A_j. \tag{5.19}$$

Consider the following induced sequences $A_{k,q}, c_{k,q}$ for each prime $q$: let $A_{k,2} = A_k$, $c_{k,2} = c_k$; and if $p$ is the prime preceding $q > 2$, let

$$A_{k,q} = A_{k,p} - p^{-k} \qquad \text{for } k \geq 1, \tag{5.20}$$

$$c_{k,q} = c_{k,p} - p^{-1} c_{k-1,p} \qquad \text{for } k \geq 1, \qquad \text{and} \quad c_{0,q} = c_0. \tag{5.21}$$

Explicitly we have

$$A_{k,q} = \sum_{r \geq q} r^{-k} \qquad \text{for } k \geq 2, \qquad \text{and} \quad A_{1,q} = -\sum_{p<q} p^{-1}. \tag{5.22}$$

**Lemma 5.3.1.** *For each prime $q$ and $k \geq 0$, we have the recursion*

$$kc_{k,q} = \sum_{j=1}^{k} c_{k-j,q} A_{j,q}. \tag{5.23}$$

*Proof.* We proceed by induction on the prime $q$. The base case $q = 2$ holds by (5.19).

Now assume (5.23) for $p < q$. The difference of recursions (5.23) for $c_{k,p}$ and $p^{-1} \cdot c_{p,k-1}$ is

$$\begin{aligned}
&kc_{k,p} - (k-1)p^{-1}c_{k-1,p} \\
&\qquad = \sum_{j=1}^{k-1} (c_{k-j,p} - p^{-1}c_{k-j-1,p}) A_{j,p} \;+\; c_{0,p} A_{k,p} \\
&\qquad = \sum_{j=1}^{k-1} (c_{k-j,p} - p^{-1}c_{k-j-1,p})(p^{-j} + A_{j,q}) \;+\; c_{0,p}(p^{-k} + A_{k,q}) \\
&\qquad = \sum_{j=1}^{k-1} (p^{-j}c_{k-j,p} - p^{-j-1}c_{k-j-1,p}) + c_{0,p}p^{-k} \;+\; \sum_{j=1}^{k-1} c_{k-j,q} A_{j,q} + c_{0,q} A_{k,q} \\
&\qquad = p^{-1}c_{k-1,p} \;+\; \sum_{j=1}^{k} c_{k-j,q} A_{j,q}
\end{aligned}$$

using (5.20), (5.21) and telescoping series. Subtracting $p^{-1}c_{k-1,p}$ gives

$$kc_{k,q} = k(c_{k,p} - p^{-1}c_{k-1,p}) = \sum_{j=1}^{k} c_{k-j,q} A_{j,q}. \qquad \square$$

Note Lemmas 5.2.3 and 5.3.1 together imply

$$c_{k,q} = \sum_{n_1+2n_2\cdots=k} \prod_{j\geq 1} \frac{(A_{j,q}/j)^{n_j}}{n_j!} \tag{5.24}$$

for each prime $q$, $k \geq 1$.

Now with the recursion in hand, we bound the induced sequence $c_{k,q}$.

**Lemma 5.3.2.** *For each prime $q$, we have $c_{k,q} \ll_q q^{-k}$ as $k \to \infty$.*

*Proof.* Fix $q$ and let $m_k = \max_{j\leq k} q^j|c_{j,q}|$. We shall prove $m_k \ll_q 1$, and it suffices to show this along a subsequence, since $m_k$ is itself a non-decreasing sequence. Namely, we consider the indices $k$ for which $m_k = q^k|c_{k,q}|$.

Recalling (5.22), we have for all $n \geq 1$

$$\sum_{1\leq j\leq n} q^j\, A_{j,q} = qA_{1,q} + (n-1) + \sum_{2\leq j\leq n} \sum_{r>q} (q/r)^j = n + O_q(1),$$

by summing the geometric series, and so the recursion (5.23) gives

$$\begin{aligned} kq^k\,|c_{k,q}| = \Big|\sum_{j=1}^{k} q^{k-j}\, c_{k-j,q}\cdot q^j\, A_{j,q}\Big| \;&\leq\; m_{k/2}\Big|\sum_{k/2<j\leq k} q^j\, A_{j,q}\Big| \;+\; m_k\Big|\sum_{1\leq j\leq k/2} q^j\, A_{j,q}\Big| \\ &=\; m_{k/2}\big(k/2+O(1)\big) \;+\; m_k\big(k/2+O(1)\big). \end{aligned}$$

And by our choice of $k$, we have $m_k = q^k|c_{k,q}|$ and so

$$m_k \leq m_{k/2}\big(1+O(1/k)\big). \tag{5.25}$$

Hence by induction on $k$, we conclude

$$m_k \ll m_1 \prod_{2^i\leq k} (1+O(2^{-i})) \ll \exp\sum_{2^i\leq k} O(2^{-i}) \ll 1.$$

□

Since Lemma 5.3.2 holds for every prime $q$ and the sequences $c_{k,q}$ are defined inductively on primes, Lemma 5.3.2 is self-improving. Indeed, for each pair of consecutive primes $p < q$,

$$c_{k,p} - p^{-1}c_{k-1,p} = c_{k,q} = O(q^{-k}).$$

In other words, multiplying above by $p^k$ the modified sequence $c'_{k,p} := c_{k,p}p^k$ satisfies $c'_{k,p} - c'_{k-1,p} = O((p/q)^k)$, so $(c'_{k,p})_{k\geq 1}$ is a Cauchy sequence for each prime $p$. Hence the limit

$$\eta_p := \lim_{k\to\infty} c'_{k,p} = \lim_{k\to\infty} c_{k,p}p^k$$

exists with $c'_{k,p} = \eta_p + O((p/q)^k)$. That is,

$$c_{k,p} = \eta_p\, p^{-k} + O_p(q^{-k}) \qquad \text{for each prime } p. \tag{5.26}$$

To summarize, we expanded the definition of $c_{k,q}$ and used a zeroth order expansion for each prime (Lemma 5.3.2) to prove a first order expansion for every prime simultaneously.

Continuing in this way, we obtain a $h$th order expansion for $c_{k,q}$ by induction on the order $h \geq 1$, at each step proving the respective expansion for every prime simultaneously.

**Proposition 5.3.3.** *For any $h \geq 1$, we have*

$$c_{k,p_n} \;=\; \sum_{l=0}^{h-1} \eta^{(n)}_{p_{n+l}}\, p_{n+l}^{-k} \;+\; O_{n,h}(p_{n+h}^{-k}), \qquad \textit{where} \quad \eta^{(n)}_{p_{n+l}} = \eta_{p_{n+l}} \Big/ \prod_{i=0}^{l-1} \big(1 - \tfrac{p_{n+l}}{p_{n+i}}\big) \tag{5.27}$$

*for all $n$ as $k \to \infty$. Here $p_n$ denotes the $n$th prime, and $\eta_p = \lim_{k\to\infty} c_{k,p} p^k$ as in* (5.26).

*Proof.* We proceed by induction on $h$. The base $h = 1$ holds for all $n$ by (5.26), since $\eta^{(n)}_{p_n} = \eta_{p_n}$. Now assume (5.27) holds with $h$ for every $n$, and write $c_{k,p_n} = \sum_{l=0}^{h} \eta^{(n)}_{p_{n+l}}\, p_{n+l}^{-k} + E_{k,n,h}$. By assumption $E_{k,n,h} \ll p_{n+h}^{-k}$, and we aim to show $E_{k,n,h} \ll p_{n+1+h}^{-k}$.

By (5.21) and the induction hypothesis (5.27) for $c_{k,p_{n+1}}$,

$$c_{k,p_{n+1}} = c_{k,p_n} - p_n^{-1} c_{k-1,p_n}$$

$$\sum_{l=0}^{h-1} \eta^{(n+1)}_{p_{n+1+l}}\, p_{n+1+l}^{-k} \;+\; O(p_{n+1+h}^{-k}) = \sum_{l=0}^{h} \eta^{(n)}_{p_{n+l}} (1 - \tfrac{p_{n+l}}{p_n}) p_{n+l}^{-k} \;+\; (E_{k,n,h} - p_n^{-1} E_{k-1,n,h})$$

Note by definition $\eta^{(n+1)}_{p_{n+l}} = \eta^{(n)}_{p_{n+l}}(1 - \frac{p_{n+l}}{p_n})$, and so the above simplifies as

$$O(p_{n+1+h}^{-k}) = E_{k,n,h} - p_n^{-1} E_{k-1,n,h}.$$

Thus, similarly as with (5.26), the modified sequence $E'_{k,n,h} := E_{k,n,h} p_n^k$ converges as $k \to \infty$ to some limit $\ell_{n,h}$, with $E'_{k,n,h} = \ell_{n,h} + O((p_n/p_{n+1+h})^k)$. That is,

$$E_{k,n,h} = \ell_{n,h}\, p_n^{-k} + O(p_{n+1+h}^{-k})$$

On the other hand, $E_{k,n,h} \ll p_{n+h}^{-k}$ forces $\ell_{n,h} = 0$. Hence $E_{k,n,h} = O(p_{n+1+h}^{-k})$ as desired. □

Next, we determine the expansion coefficients $\eta_p$ from (5.26).

**Proposition 5.3.4.** *For any prime $p$, the coefficient $\eta_p = \lim_{k\to\infty} c_{k,p}p^k$ equals*

$$\eta_p = e^{-\sum_{q\le p} p/q} \prod_{q>p} (1-\tfrac{p}{q})^{-1} e^{-p/q}. \tag{5.28}$$

*Proof.* Consider the generating function $C_p(z) = \sum_{k\ge 0} c_{k,p} z^k$. On one hand, the explicit formula (5.24) for $c_{k,p}$ implies

$$\begin{aligned} C_p(z) = \sum_{k\ge 0} c_{k,p} z^k &= \sum_{k\ge 0} z^k \sum_{n_1+2n_2\cdots=k} \prod_{j\ge 1} \frac{\big(A_{j,p}/j\big)^{n_j}}{n_j!} = \prod_{j\ge 1} \sum_{n_j\ge 0} \frac{1}{n_j!}\big(A_{j,p}z^j/j\big)^{n_j} \\ &= \prod_{j\ge 1} \exp\big(A_{j,p}z^j/j\big) = \exp\big(\sum_{j\ge 1} A_{j,p}z^j/j\big). \end{aligned}$$

Then recalling $A_{j,p} = \sum_{q\ge p} q^{-j}$ for $j \ge 2$,

$$\begin{aligned} C_p(z) &= \exp\big(zA_{1,p} + \sum_{q\ge p}\sum_{j\ge 2} (z/q)^j/j\big) = e^{zA_{1,p}} \exp\big(-\sum_{q\ge p} [\log(1-z/q) + z/q]\big) \\ &= e^{zA_{1,p}} \prod_{q\ge p} (1-z/q)^{-1} e^{-z/q}. \end{aligned} \tag{5.29}$$

On the other hand, by the expansion (5.26) for $c_k$ we have

$$C_p(z) = \sum_{k\ge 0} c_k z^k = \eta_p \sum_{k\ge 0} (z/p)^k \ + \ O_p\Big(\sum_{k\ge 0} (z/q)^k\Big) = \frac{\eta_p}{1-z/p} \ + \ \frac{O_p(1)}{1-z/q}. \tag{5.30}$$

since $A_{1,p} = -\sum_{q<p} q^{-1}$. So comparing $C_p(z)$ from (5.29) and (5.30) at the pole $z = p$,

$$\eta_p = \lim_{z\to p} C_p(z)(1-z/p) = e^{pA_{1,p}-1} \prod_{q>p} (1-p/q)^{-1} e^{-p/q}.$$

Hence the result follows since $A_{1,p} = -\sum_{q<p} q^{-1}$. □

Finally, we obtain an expansion for the original sequence $c_{k,p_1} = c_k$ to arbitrary order, which gives a considerable refinement of Theorem 5.1.11.

**Theorem 5.3.5.** *For each prime $q$,*

$$c_k = \sum_{p<q} \alpha_p\, p^{-k} \ + \ O_q(q^{-k})$$

*where $\alpha_p := e^{-1}\prod_{q\ne p}(1-\frac{p}{q})^{-1}e^{-p/q}$. In particular $c_k = \alpha_2\, 2^{-k} + O(3^{-k})$.*

*Proof.* Setting $n = 1$ in Proposition 5.3.3, the sequence $c_{k,p_1} = c_k$ satisfies

$$c_k = \sum_{p<q} \eta_p^{(1)}\, p^{-k} \;+\; O_q(q^{-k})$$

where, by definition of $\eta_p^{(1)}$ in (5.27), Proposition 5.3.4 gives

$$\eta_p^{(1)} := \eta_p / \prod_{q<p}(1-\tfrac{p}{q}) = e^{-1}\prod_{q\neq p}(1-\tfrac{p}{q})^{-1}e^{-p/q} = \alpha_p.$$

□

## 5.4 Analytic proof for $k$ in uniform range

We prove Theorem 5.1.7 which quantitative error, which we state below.

**Theorem 5.4.1.** *Let $r = k/\log_2 x$ and define $\eta(z) = e^{\gamma z}\prod_p(1-\frac{1}{p})^z(1-\frac{z}{p})^{-1}$. For any $\varepsilon > 0$, as $x \to \infty$ we have uniformly for $r \le 2-\varepsilon$,*

$$\sum_{\substack{\Omega(n)=k\\ P^+(n)\le x}} \frac{1}{n} \;=\; \eta(r)\;\frac{(\log_2 x)^k}{k!}\;\Big(1+O_\varepsilon\big(\tfrac{k}{(\log_2 x)^2}\big)\Big).$$

*Proof.* By Cauchy's residue formula, we have for any $r < 2$,

$$P_k(1,x) = \frac{1}{2\pi i}\int_{|z|=r} f_x(z)\frac{dz}{z^{k+1}}, \tag{5.31}$$

where $f_x$ is given by the power series

$$\begin{aligned} f_x(z) = \sum_{k\ge 0} P_k(1,x)z^k &= \sum_{P^+(n)\le x}\frac{z^{\Omega(n)}}{n}\\ &= \prod_{p\le x}\Big(1+\frac{z}{p}+\frac{z^2}{p^2}+\cdots\Big) = \prod_{p\le x}\Big(1-\frac{z}{p}\Big)^{-1}\\ &= (1+O(E(x)))e^{z\gamma}(\log x)^z\prod_{p\le x}\Big(1-\frac{z}{p}\Big)^{-1}\Big(1-\frac{1}{p}\Big)^z\\ &= \;(1+O(E(x)))\eta(z)(\log x)^z, \end{aligned}$$

as $\prod_{p\le x}(1-\frac{1}{p})^{-1} = (1+E(x))e^\gamma\log x$ by Merten's 2nd theorem in quantitative form.[3]

Hence (5.31) becomes

$$P_k(1,x) = \frac{1+O(E(x))}{2\pi i}\int_{|z|=r}\eta(z)(\log x)^z\frac{dz}{z^{k+1}}. \tag{5.32}$$

[3]This also follows from the prime number theorem.

The desired main term in Theorem 5.1.7 is given by evaluating $\eta(z)$ at $z = r$, namely

$$\frac{\eta(r)}{2\pi i}\int_{|z|=r}(\log x)^z\frac{dz}{z^{k+1}} = \eta(r)\frac{(\log_2 x)^k}{k!}. \tag{5.33}$$

For the error we follow the argument in [106, p.233], which we provide for completeness. Recall $E(x) \ll 1/\log x$. For $|z| = r = k/\log_2 x$, integration by parts gives

$$\frac{1}{2\pi i}\int_{|z|=r}(z-r)(\log x)^z\frac{dz}{z^{k+1}} = \frac{(\log_2 x)^{k-1}}{(k-1)!} - \frac{r(\log_2 x)^k}{k!} \;=\; 0, \tag{5.34}$$

$$\text{and} \qquad \eta(z) - \eta(r) - \eta'(r)(z-r) = \int_r^z (z-w)\eta''(w)\,dw \ll |z-r|^2. \tag{5.35}$$

Thus subtracting (5.33) from (5.32), the error is

$$\begin{aligned}
\eta(r)\frac{(\log_2 x)^k}{k!} \;-\; P_k(1,x) \;&\ll \int_{|z|=r}[\eta(r)-\eta(z)](\log x)^z\frac{dz}{z^{k+1}} \\
&\overset{(5.34)}{=} \int_{|z|=r}[\eta(r)-\eta(z)-\eta'(r)(z-r)](\log x)^z\frac{dz}{z^{k+1}} \\
&\overset{(5.35)}{\ll} \int_{|z|=r}|z-r|^2(\log x)^z\frac{dz}{z^{k+1}} \\
&\ll r^{2-k}\int_{-1/2}^{1/2}(\sin\pi\theta)^2 e^{k\cos(2\pi\theta)}\,d\theta \\
&\ll r^{2-k}e^k\int_0^\infty \theta^2 e^{-8k\theta^2}d\theta \ll r^{2-k}e^k k^{-3/2} \\
&= (\log_2 x)^{k-2}(e/k)^k k^{1/2} \;\ll\; k(\log_2 x)^{k-2}/k!
\end{aligned}$$

Here we used $|\sin x| \le x$, $\cos(2\pi\theta) \le 1 - 8\theta^2$ for $|\theta| \le 1/2$, and Stirling's formula. $\square$

# Chapter 6

# Higher Mertens constants for almost primes

## 6.1 Introduction

For a positive integer $n$ denote by $\Omega(n)$ the number of prime factors of $n$, counted with multiplicity. For $k \geq 1$, a $k$-almost prime is a positive integer $n$ with $\Omega(n) = k$.

In this chapter we consider the reciprocal sum of $k$-almost primes,

$$\mathcal{R}_k(x) = \sum_{\substack{\Omega(n)=k \\ n \leq x}} \frac{1}{n} = \sum_{\substack{p_1 \cdots p_k \leq x \\ p_1 \leq \ldots \leq p_k}} \frac{1}{p_1 \cdots p_k}. \tag{6.1}$$

In estimating this sum, we generalize Mertens' second theorem [104],

$$\sum_{p \leq x} \frac{1}{p} = \log_2 x + \beta + O\left(\frac{1}{\log x}\right),$$

where $\beta = 0.2614\ldots$ is the Mertens (or Meissel–Mertens) constant, which satisfies

$$\beta = \gamma + \sum_{p} \left(\frac{1}{p} + \log\left(1 - \frac{1}{p}\right)\right). \tag{6.2}$$

Here $\gamma = 0.5772\ldots$ is the Euler–Mascheroni constant.

Throughout the chapter, we denote $p$ as a prime variable and $\log_2 x = \log(\log x)$ as the iterated natural logarithm. In addition, $\Gamma$ is the Euler gamma function and $\zeta$ is the Riemann zeta function.

Our main result is the following.

**Theorem 6.1.1.** *For $x \geq 3$, we have*

$$\mathcal{R}_k(x) = \sum_{j=0}^{k} \frac{\nu_{k-j}}{j!} (\log_2 x)^j + O_k\left(\frac{(\log_2 x)^{k-1}}{\log x}\right), \tag{6.3}$$

*where $\nu_j := \nu^{(j)}(0)/j!$ are the Taylor coefficients of*

$$\nu(z) = \frac{1}{\Gamma(z+1)} \prod_{p} \left(1 - \frac{z}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^{z}. \tag{6.4}$$

The proof of Theorem 6.1.1 is completely elementary, see Section 6.2 for further remarks.

For comparision, classical analytic estimates imply the slightly cruder result

$$\mathcal{R}_k(x) = \sum_{j=1}^{k} \frac{\nu_{k-j}}{j!} (\log_2 x)^j + C_k + O_k\left(\frac{(\log_2 x)^k}{\log x}\right) \tag{6.5}$$

for some (a priori unspecified) constant $C_k$.[1] Indeed, the Sathe-Selberg theorem [137, Theorem II.6.5] states that, for each $0 < \delta < 1$, the counting function $\mathcal{N}_k(x)$ of $k$-almost primes is

$$\mathcal{N}_k(x) = \frac{x}{\log x} \sum_{j=0}^{k-1} \frac{\nu_{k-1-j}}{j!} (\log_2 x)^j + O_\delta \left( \frac{x(\log_2 x)^k}{k!(\log x)^2} \right) \tag{6.6}$$

uniformly for $x \geq 3$ and $k \leq (2-\delta)\log_2 x$. Then by partial summation, (6.6) implies (6.5) with constant coefficient

$$C_k = \int_{2^k}^{\infty} \frac{\mathcal{E}_k(t)}{t^2} \mathrm{d}t - \sum_{j=1}^{k} \frac{\nu_{k-j}}{j!} (\log_2 2^k)^j, \tag{6.7}$$

where $\mathcal{E}_k(t)$ is the error term in (6.6). Similarly as with $\nu_1 = \beta$ in the original proof of Mertens' second theorem, the added value of Theorem 6.1.1 lies in the identification of $C_k = \nu_k$.

Moreover, though not immediately apparent from the statement itself, our proof of Theorem 6.1.1 directly implies the following concise bounds, which may be of independent interest.

**Corollary 6.1.1.1.** *For each $k \geq 2$ there exists $x_0(k) > 0$ such that*

$$\frac{1}{k!} (\log_2 x)^k \ < \ \mathcal{R}_k(x) \ < \ \frac{1}{k!} (\log_2 x + \beta)^k$$

*for all $x > x_0(k)$, where $\beta$ is the Mertens constant.*

The lower bound of Corollary 6.1.1.1 holds when $k = 1$ by Mertens' second theorem, in fact for all $x > 1$ (see for instance [124]), while the upper bound fails to hold when $k = 1$. By a result of G. Robin [123], the error term in Mertens' second theorem changes sign infinitely often. As a consequence of Theorem 6.1.3 below, we find that the optimal value of $x_0(2)$ is $\exp(\exp(\sqrt{113/90} - \beta)) = 10.5998\ldots$. (See Corollary 6.1.1.2 below.)

We note that Theorem 6.1.1 generalizes Mertens' second theorem, which is the special case $k = 1$. In particular, the constant coefficients $\nu_k$ may be viewed as higher analogues of the Mertens constant $\nu_1 = \beta$.

[1] The exponent $k$ in the error term of (6.5) corrects the stated exponent $k-1$ in the published version.

We establish a recurrence for $\nu_k$ (see Proposition 6.7.0.1) which enables rapid computation of $\nu_k$ to high precision. We display results for $k \le 20$ below.

| $k$ | $\nu_k$ | $k$ | $\nu_k$ |
|---|---|---|---|
| 1 | $\beta = 2.61497 \cdot 10^{-1}$ | 11 | $1.49365 \cdot 10^{-4}$ |
| 2 | $-5.62153 \cdot 10^{-1}$ | 12 | $4.99174 \cdot 10^{-5}$ |
| 3 | $3.05978 \cdot 10^{-1}$ | 13 | $1.82657 \cdot 10^{-5}$ |
| 4 | $2.62973 \cdot 10^{-2}$ | 14 | $1.30241 \cdot 10^{-5}$ |
| 5 | $-6.44501 \cdot 10^{-2}$ | 15 | $5.52779 \cdot 10^{-6}$ |
| 6 | $3.64064 \cdot 10^{-2}$ | 16 | $2.90194 \cdot 10^{-6}$ |
| 7 | $-4.70865 \cdot 10^{-3}$ | 17 | $1.45075 \cdot 10^{-6}$ |
| 8 | $-4.33984 \cdot 10^{-4}$ | 18 | $7.19861 \cdot 10^{-7}$ |
| 9 | $1.5085 \cdot 10^{-3}$ | 19 | $3.61606 \cdot 10^{-7}$ |
| 10 | $-1.83548 \cdot 10^{-4}$ | 20 | $1.80517 \cdot 10^{-7}$ |

Though initially appearing rather erratic, the table eventually suggests $\nu_k \approx 0.189 \cdot 2^{-k}$.

In light of (or perhaps despite) this numerical evidence, we show the higher Mertens constants $\nu_k$ satisfy the following precise asymptotics.

**Theorem 6.1.1.** *We have* $\nu_k = \beta_2\, 2^{-k} + O(3^{-k})$, *where* $\beta_2 = \frac{1}{8}\prod_{p>2} \frac{(1-1/p)^2}{1-2/p} \approx 0.1893475$.

*Moreover for any prime $q$, we have*

$$\nu_k = \sum_{p<q} \beta_p\, p^{-k} \;+\; O_q(q^{-k}), \tag{6.8}$$

*where $\beta_2, \beta_3, \beta_5, \ldots$ are constants given by*

$$\beta_p := \frac{(1-1/p)^p}{p!} \prod_{p' \neq p} \Big(1 - \frac{p}{p'}\Big)^{-1} \Big(1 - \frac{1}{p'}\Big)^{p}. \tag{6.9}$$

The proof of Theorem 6.1.1 is in Section 6.7, also by elementary combinatorial methods. For an outline of the arguments see Section 6.2.

### 6.1.1 Finer-scale estimates for semiprimes.

Let $k = 2$. We establish a finer estimate for $\mathcal{R}_2(x)$. Let $E(x) = \sum_{p \le x} 1/p - (\log_2 x + \beta)$ denote the error term in Mertens' second theorem.

**Theorem 6.1.2.** *For any $N \ge 0$ we have*

$$\mathcal{R}_2(x) = \frac{1}{2}(\log_2 x + \beta)^2 + \frac{P(2) - \zeta(2)}{2} + \sum_{1 \le j \le N} \frac{\alpha_j}{\log^j x} + O_N\left((\log x)^{-N-1}\right),$$

*where $\alpha_j$ are constants given by*

$$\alpha_j := \frac{\log^j 2}{j}\left(\frac{1}{j} - \log_2 2 - \beta\right) + \int_2^\infty \frac{E(t)\log^{j-1} t}{t}\mathrm{d}t. \tag{6.10}$$

As we shall prove in Lemma 6.8.4, the constants $\alpha_j$ can be viewed as higher analogues of the constant

$$\alpha_1 = \lim_{x\to\infty}\left(\log x - \sum_{p\le x}\frac{\log p}{p}\right) = \gamma + \sum_{m\ge 2}\sum_p \frac{\log p}{p^m} = 1.332582\ldots$$

arising from a strong form of Mertens' first theorem. See for instance [37, Theorem 5.7]. We also determine an explicit bound for $\mathcal{R}_2(x)$.

**Theorem 6.1.3.** *For all $x > 1$ we have*

$$\left|\mathcal{R}_2(x) \;-\; \frac{1}{2}(\log_2 x + \beta)^2 - \frac{P(2)-\zeta(2)}{2} - \frac{\alpha_1}{\log x}\right| \;<\; (\log x)^{-3/2}.$$

As a consequence we have the following bounds which make Corollary 6.1.1.1 explicit for the case $k = 2$.

**Corollary 6.1.1.2.** *We have*

$$\frac{1}{2}(\log_2 x)^2 \;<\; \mathcal{R}_2(x) \;<\; \frac{1}{2}(\log_2 x + \beta)^2,$$

*where the lower bound holds for all $x \ge 4$ and the upper bound holds for all $x \ge 11$.*

We will prove Theorems 6.1.2 and 6.1.3 in Section 6.8. We also consider the sum $\mathcal{R}_2^*(x)$ restricted to squarefree semiprimes (i.e. products of 2 distinct primes).

**Corollary 6.1.1.3.** *For any $N \ge 0$, $\mathcal{R}_2^*(x)$ satisfies the estimate in Theorem 6.1.2 with $P(2)/2$ replaced by $-P(2)/2$. Moreover, we have*

$$\left|\mathcal{R}_2^*(x) \;-\; \frac{1}{2}(\log_2 x + \beta)^2 + \frac{P(2)+\zeta(2)}{2} - \frac{\alpha_1}{\log x}\right| \;<\; (\log x)^{-3/2},$$

*where the upper bound on $\mathcal{R}_2^*(x)$ holds for all $x \ge 227$ and the lower bound holds for all $x > 1$.*

The corollary follows from the bounds established in the proof of Theorem 6.1.3 together with Lemma 6.4.1 below.

We learned that constants related to $\alpha_j$ from (6.10) appear in a very recent preprint [29] of Crisan and Erban who derive an asymptotic expansion for the counting function $\mathcal{N}_2(x)$ of semiprimes and compute the values of $B_j := -j\alpha_j$ to high

precision. By extending their table of $B_j$ for $1 \leq j \leq 10$, we are led to conjecture that

$$\frac{\alpha_{j+1}}{\alpha_j} \sim \frac{2j^2}{j+1}. \tag{6.11}$$

Indeed, equation (3.5) in [29] yields the following numerical values for $\alpha_j$ and the ratio $(\alpha_{j+1}/\alpha_j)/(2j^2/(j+1))$ for $11 \leq j \leq 20$, computed in Pari/GP.

| $j$ | $\alpha_j$ | $(\alpha_{j+1}/\alpha_j)/(2j^2/(j+1))$ |
|---|---|---|
| 11 | $3.4791 \cdot 10^8$ | 0.98998 |
| 12 | $6.9638 \cdot 10^9$ | 0.99253 |
| 13 | $1.5342 \cdot 10^{11}$ | 0.99443 |
| 14 | $3.6886 \cdot 10^{12}$ | 0.99585 |
| 15 | $9.6096 \cdot 10^{13}$ | 0.99691 |
| 16 | $2.6965 \cdot 10^{15}$ | 0.99769 |
| 17 | $8.1071 \cdot 10^{16}$ | 0.99828 |
| 18 | $2.5999 \cdot 10^{18}$ | 0.99871 |
| 19 | $8.8587 \cdot 10^{19}$ | 0.99904 |
| 20 | $3.1957 \cdot 10^{21}$ | 0.99928 |

*Remark* 6.1.2. In private communication, Ofir Gorodetsky has given a sketch proof that

$$\alpha_j \sim (j-1)! \frac{2^{j-1}}{j},$$

which immediately implies (6.11). The basic approach is to use the representation for $\alpha_j$ obtained in Lemma 6.8.4, and apply the von Mangoldt explicit formula (with some care). Full details will be provided in forthcoming work.

We note that the work of Crisan and Erban [29] may suggest an alternate approach to Theorem 6.1.2. Conversely, Theorem 6.1.2 can be used to give a new proof of Crisan and Erban's asymptotic expansion of $\mathcal{N}_2(x)$.

**Theorem 6.1.4** (Theorem 2.5 [29])**.** *For any $N \geq 0$ we have*

$$\mathcal{N}_2(x) = \frac{x}{\log x} \sum_{n=0}^{N-1} n! \frac{\log_2 x + D_n}{\log^n x} + O_N\left(\frac{x \log_2 x}{\log^{N+1} x}\right),$$

*for constants $D_n = \sum_{j=0}^{n} B_j/j! - H_n$, where $H_n$ is the $n$-th harmonic number, $B_0 = \beta$, and $B_j = -j\alpha_j$, $j \geq 1$, for $\alpha_j$ as in* (6.10)*.*

The new proof of Theorem 6.1.4 is provided in Section 6.9. It relies on a strong form of the prime number theorem, which implies the estimate

$$E(x) \ll_A (\log x)^{-A} \text{ for all } A > 0.$$

We will make use of this estimate throughout the chapter. We remark that it follows from a version of the prime number theorem established by elementary methods. See for instance [34, Sec. 8].

## 6.2 Outline and perspectives on the argument

Our proof for $\mathcal{R}_k(x)$ proceeds by building on recent work of Tenenbaum [136] and Qi–Hu [120] for the related sum

$$\mathcal{S}_k(x) := \sum_{p_1 \cdots p_k \le x} \frac{1}{p_1 \cdots p_k} = \sum_{\substack{\Omega(n)=k \\ n \le x}} \frac{\mathrm{f}(n)}{n}, \tag{6.12}$$

where $\mathrm{f}(n)$ denotes the number of ordered prime factorizations of $n$. Recall (6.1) for comparision.

**Theorem 6.2.1** (Theorem 1 [136]). *For $x \ge 3$,*

$$\mathcal{S}_k(x) = S_k(\log_2 x) + O_k\left(\frac{(\log_2 x)^{k-1}}{\log x}\right), \tag{6.13}$$

*where $S_k(X) = \sum_{0 \le j \le k} \lambda_{j,k} X^j$ is a monic polynomial of degree $k$ and*

$$\lambda_{j,k} = \sum_{m=0}^{k-j} \binom{k}{m, j, k-m-j} (\beta - \gamma)^{k-m-j} \left(\frac{1}{\Gamma}\right)^{(m)} (1). \tag{6.14}$$

Tenenbaum's analytic proof employs the Selberg–Delange method. However, Qi and Hu [120] have recently obtained an elementary proof via Dirichlet's hyperbola method.

**Theorem 6.2.2** (Theorem 1.3 [120]). *For $x \ge 3$,*

$$\mathcal{S}_k(x) = \sum_{j=0}^{k} \binom{k}{j} a_{k-j} (\log_2 x + \beta)^j + O_k\left(\frac{(\log_2 x)^{k-1}}{\log x}\right), \tag{6.15}$$

*where the coefficients $a_j$ are defined recursively by $a_0 = 1, a_1 = 0$, and*

$$a_j = \sum_{i=1}^{j-1} (-1)^i \binom{j-1}{i} i! \zeta(i+1) a_{j-1-i}.$$

The results of Tenenbaum and Qi–Hu generalize those of Popa [118, 119] in the cases $k = 2$ and $k = 3$ (as well as Mertens when $k = 1$). Though it is not immediately

clear that the expansions in Theorems 6.2.1 and 6.2.2 are equivalent, we will show this in Section 6.6.

In the sum $\mathcal{S}_k(x)$, terms can appear up to $k!$ times due to permutation of the prime factors. We obtain the following estimate for $\mathcal{R}_k(x)$, which counts each term precisely once.

**Theorem 6.2.3.** *For $x \geq 3$, we have*

$$\mathcal{R}_k(x) = R_k(\log_2 x + \beta) + O_k\left(\frac{(\log_2 x)^{k-1}}{\log x}\right),$$

*where $R_k$ is the polynomial of degree $k$ defined by*

$$R_k(X) = \sum_{n_1+2n_2+\ldots+kn_k=k} \sum_{i=0}^{n_1} \frac{a_{i,n_1}}{n_1!} \prod_{j=2}^{k} \left(\frac{P(j)}{j}\right)^{n_j} \frac{X^j}{n_j!},$$

*where the sum ranges over all integer partitions of $k$, $P(j) = \sum_p p^{-j}$ is the prime zeta function, and the coefficients $a_{i,n}$ are given by*

$$a_{i,n} = \sum_{m=0}^{n-i} \binom{n}{m, i, n-m-i} (-\gamma)^{n-m-i} \left(\frac{1}{\Gamma}\right)^{(m)}(1).$$

When combined with the Weierstrass product formula (see (6.29)), Theorem 6.2.3 yields the following special cases:

$$\begin{aligned}
\mathcal{R}_2(x) &= \frac{1}{2}(\log_2 x + \beta)^2 + \frac{P(2) - \zeta(2)}{2} + E_2(x), \\
\mathcal{R}_3(x) &= \frac{1}{6}(\log_2 x + \beta)^3 + \frac{P(2) - \zeta(2)}{2}(\log_2 x + \beta) + \frac{P(3) + \zeta(3)}{3} + E_3(x), \\
\mathcal{R}_4(x) &= \frac{1}{24}(\log_2 x + \beta)^4 + \frac{P(2) - \zeta(2)}{4}(\log_2 x + \beta)^2 + E_4(x) \\
&\quad + \frac{P(3) + \zeta(3)}{3}(\log_2 x + \beta) + \frac{P(4)}{4} + \frac{\zeta(4)}{16} + \frac{P(2)^2}{8} - \frac{P(2)\zeta(2)}{4}.
\end{aligned}$$

Here $\zeta$ is the Riemann zeta function and

$$E_k(x) := \mathcal{R}_k(x) - R_k(\log_2 x + \beta) \ll_k (\log_2 x)^{k-1}/\log x.$$

We observe that since $P(2) - \zeta(2) < 0$, Theorem 6.2.3 implies Corollary 6.1.1.1.

Recall Theorem 6.1.1 states $\mathcal{R}_k(x) = V_k(\log_2 x) + O_k((\log_2 x)^{k-1}/\log x)$ for the polynomial

$$V_k(X) := \sum_{j=0}^{k} \frac{\nu_{k-j}}{j!} X^j. \tag{6.16}$$

In Sections 6.3-6.5, we shall establish Theorem 6.2.3 from Tenenbaum's result by an elementary combinatorial argument. By a similar argument we prove the polynomial identity

$$R_k(X+\beta) = V_k(X) \tag{6.17}$$

which gives the equivalence of Theorems 6.2.3 and 6.1.1. This allows us to obtain Theorem 6.1.1, thus implying that $C_k = \nu_k$. Hence, when combined with the hyperbola method in Theorem 6.2.2, the proof of Theorems 6.1.1 and 6.2.3 are completely elementary in nature.

As the Sathe–Selberg theorem (6.6) is proven by contour integration, Theorem 6.1.1 represents a success of elementary methods at matching the strength of results with those of more sophisticated analytic tools. One may wonder if similar elementary methods could be used to establish the Sathe–Selberg theorem itself.

### 6.2.1 Outline of Theorem 6.1.1

We note that the proof methods in Theorem 6.2.3 to handle $\mathcal{R}_k(x)$ are motivated by those of [88, Theorem 2] for the smooth variant

$$\mathcal{P}_k(x) := \sum_{\substack{\Omega(n)=k \\ P^+(n)\le x}} \frac{1}{n}, \tag{6.18}$$

where $P^+(n)$ is the largest prime factor of $n$. As noted, the proof of Theorem 6.2.3 proceeds by passing from $\mathcal{R}_k(x)$ to $\mathcal{S}_k(x)$. Whereas, the analogous method for $\mathcal{P}_k(x)$ may be applied directly, and is quite natural in the smooth setting.

Theorems 1.2 and 3.5, as well as (2.3), in [88] give the following.

**Theorem 6.2.4** ([88]). *For each fixed $k \ge 1$, we have*

$$\mathcal{P}_k(x) \;=\; \sum_{j=0}^{k} \frac{d_{k-j}}{j!}\big(\log_2 x + \beta\big)^j \;+\; O_k\Big(\frac{(\log_2 x)^{k-1}}{\log x}\Big),$$

*where the sequence $(d_k)_{k=0}^{\infty}$ is recursively defined by $d_0 = 1$ and $d_k = \frac{1}{k}\sum_{j=2}^{k} d_{k-j}\,P(j)$. Explicitly, we have*

$$d_k = \sum_{2n_2+3n_3+\cdots=k} \prod_{j\ge 2} \frac{(P(j)/j)^{n_j}}{n_j!},$$

*where the sum ranges over partitions of $k$ without singletons.*

*Moreover, for each prime $q$,*

$$d_k = \sum_{p<q} \delta_p \, p^{-k} \ + \ O_q(q^{-k}), \tag{6.19}$$

*where* $\delta_p := e^{-1} \prod_{q \neq p} (1 - \frac{p}{q})^{-1} e^{-p/q}$. *In particular* $d_k = \delta_2 \, 2^{-k} + O(3^{-k})$.

In Section 6.7, we utilize Theorem 6.2.4 in order to establish Theorem 6.1.1.

Finally, in Theorem 6.1.1, it is straightforward to verify that if such an expansion (6.8) exists, then the coefficients are determined uniquely by (6.9). Indeed, this is seen by comparing the residue from the poles of $\nu(z)$ at $z = p$ with that of

$$\begin{aligned} \nu(z) = \sum_{k \geq 0} \nu_k z^k \ &= \ \sum_{p<q} \beta_p \sum_{k \geq 0} (z/p)^k \ + \ O_q(\sum_{k \geq 0} (z/q)^k) \\ &= \sum_{p<q} \frac{\beta_p}{1 - z/p} \ + \ \frac{O_q(1)}{1 - z/q}. \end{aligned}$$

Hence the core of the proof in Theorem 6.1.1 lies in the existence of the expansion (6.8).

This situation is somewhat reminiscent of the developments leading to the prime number theorem (PNT), i.e. $\pi(x) \sim x/\log x$. Indeed, Mertens' second theorem did not directly imply PNT. However, if the expansion $\pi(x) \sim cx/\log x$ were assumed to exist for some constant $c > 0$, then Mertens' theorem would force $c = 1$ (this result is historically attributed to Chebyshev). As such the core of the proof of PNT was to establish existence of such an expansion.

## 6.3 The Proof of Theorem 6.2.3

To prove Theorem 6.2.3 it suffices to establish the following result, which allows us to apply Tenenbaum's estimate (6.13).

**Proposition 6.3.0.1.** *For $x \geq 3$ we have*

$$\mathcal{R}_k(x) = \sum_{n_1 + 2n_2 + \ldots + kn_k = k} \frac{\mathcal{S}_{n_1}(x)}{n_1!} \cdot \prod_{j=2}^{k} \frac{(P(j)/j)^{n_j}}{n_j!} + O_k\left( \frac{(\log_2 x)^{k-1}}{\log x} \right).$$

This result may be viewed as an analogue of [86, Proposition 3.1] and [88, Proposition 2.1]. These latter results concern "smooth" (or "friable") sums over $P^+(n) \leq x$, in which setting admits concise combinatorial proofs. However, the technical difficulty for Proposition 6.3.0.1 arises, since the usual sums over $n \leq x$ in $\mathcal{R}_k(x)$ require greater care in order to facilitate the combinatorics.

We will establish Proposition 6.3.0.1 after stating and proving Propositions 6.3.0.2 and 6.3.0.3 below.

**Proposition 6.3.0.2.** *We have*

$$\mathcal{R}_k(x) = \sum_{n_1+2n_2+\ldots+kn_k=k} \left( \prod_{j=1}^{k} \frac{j^{-n_j}}{n_j!} \right) \sum_{m \leq x} \frac{f(m)}{m},$$

*where* $f(m) = f_{(n_1,\ldots,n_k)}(m) = \#\{\text{orderedtuples}(p_{1j}, .., p_{n_j j})_{j \leq k} : m = \prod_{j=1}^{k} \prod_{i=1}^{n_j} p_{ij}^{j}\}$.

*Proof.* We apply Perron's formula to the characteristic function of the $k$-almost primes, along with the inclusion-exclusion formulas for the almost prime zeta function $P_k(z) = \sum_{\Omega(n)=k} n^{-z}$ (in contrast to $P(z)^k$ as in Tenenbaum's proof [136]) to obtain for any $c > 0$ and $x \in \mathbb{R}^+ \setminus \mathbb{N}$,

$$\begin{aligned}
\mathcal{R}_k(x) = \sum_{\substack{\Omega(n)=k \\ n \leq x}} \frac{1}{n} &= \frac{1}{2\pi i} \int_{c+i\mathbb{R}} P_k(z+1) \frac{x^z}{z} \mathrm{d}z \\
&= \frac{1}{2\pi i} \int_{c+i\mathbb{R}} \sum_{n_1+2n_2+\ldots+kn_k=k} \prod_{j=1}^{k} \frac{1}{n_j!} \left( \frac{P(j(z+1))}{j} \right)^{n_j} \frac{x^z}{z} \mathrm{d}z \\
&= \sum_{n_1+2n_2+\ldots+kn_k=k} \frac{1}{2\pi i} \left( \prod_{j=1}^{k} \frac{j^{-n_j}}{n_j!} \right) \int_{c+i\mathbb{R}} \prod_{j=1}^{k} P(j(z+1))^{n_j} \frac{x^z}{z} \mathrm{d}z,
\end{aligned}$$

where the equality of the first and second lines is a direct application of [86, Proposition 3.1]. Applying Perron's formula again,

$$\mathcal{R}_k(x) = \sum_{n_1+2n_2+\ldots+kn_k=k} \left( \prod_{j=1}^{k} \frac{j^{-n_j}}{n_j!} \right) \sum_{m \leq x} \frac{f(m)}{m}, \tag{6.20}$$

where $f$ satisfies

$$\sum_{m=1}^{\infty} \frac{f(m)}{m^s} = \prod_{j=1}^{k} P(js)^{n_j} = \prod_{j=1}^{k} \prod_{i=1}^{n_j} \sum_{p_{ij}} p_{ij}^{-js} = \sum_{A} \prod_{q \in A} q^{-s}$$

as $A$ ranges over all choices of multi-sets $A = \{p_{ij}^{j} : j \leq k, i \leq n_j\}$. Thus $f(m) = \sum_{m=\prod_{j=1}^{k} \prod_{i=1}^{n_j} p_{ij}^{j}} 1$, by the uniqueness of Dirichlet series coefficients. Hence

$$\sum_{m \leq x} \frac{f(m)}{m} = \sum_{\prod_{j=1}^{k} \prod_{i=1}^{n_j} p_{ij}^{j} \leq x} \prod_{j=1}^{k} \prod_{i=1}^{n_j} p_{ij}^{-j}. \tag{6.21}$$

Substituting this back into (6.20) completes the proof. □

We note that Proposition 6.3.0.2 is completely elementary and combinatorial in nature. However, it is efficiently deduced with the help of Perron's formula.

To complete the proof of Proposition 6.3.0.1 we show that (at an admissible error) we may replace the condition $\prod_{j=1}^{k}(p_{1j}\cdots p_{n_jj})^j \leq x$ in the proof of Proposition 6.3.0.2 with the weaker condition $p_{11}\cdots p_{n_11} \leq x$, to obtain

$$\sum_{m\leq x}\frac{f(m)}{m} \approx \Big(\sum_{p_{11}\cdots p_{n_11}\leq x}\prod_{i=1}^{n_1}p_{i1}^{-1}\Big)\prod_{j=2}^{k}\prod_{i=1}^{n_j}\sum_{p_{ij}}p_{ij}^{-j} = \mathcal{S}_{n_1}(x)\prod_{j=2}^{k}P(j)^{n_j}.$$

We shall do so using the following proposition, whose proof we give in the next section.

**Proposition 6.3.0.3.** *Given $\ell \geq 0, n \geq 0$, and $a_i \geq 2$ for $i = 1,\dots,\ell$, let $k = n + \sum_{i\leq\ell} a_i$. Then*

$$\sum_{m=q_1\cdots q_n\prod_{i\leq\ell}p_i^{a_i}\leq x}\frac{1}{m} = \mathcal{S}_n(x)\prod_{i\leq\ell}P(a_i) + O_k\left(\frac{(\log_2 x)^{n-1}}{\log x}\right),$$

*where the sum $\sum_{m\leq x}$ ranges over $n+\ell$ independent prime variables, $q_1,\dots q_n, p_1\dots p_\ell$, whose values are not necessarily distinct.*

*Proof of Proposition 6.3.0.1 from Propositions 6.3.0.2 and 6.3.0.3.* From (6.21),

$$\begin{aligned}\sum_{m\leq x}\frac{f(m)}{m} &= \sum_{m=p_{11}\cdots p_{n_11}\prod_{j=2}^{k}\prod_{i=1}^{n_j}p_{ij}^{j}\leq x}\frac{1}{m}\\ &= \mathcal{S}_{n_1}(x)\prod_{j=2}^{k}P(j)^{n_j} + O_k\left(\frac{(\log_2 x)^{n_1-1}}{\log x}\right) \qquad (6.22)\end{aligned}$$

by Proposition 6.3.0.3 for the choices of parameters

$$n = n_1,\quad \ell = 2n_2+\cdots+kn_k,\quad \{a_1,\dots,a_\ell\} = \bigcup_{j=2}^{k}\{j\}^{n_j}$$
$$q_1,\dots q_n = p_{11},\dots p_{n_11},\quad \prod_{i=1}^{\ell}p_i^{a_i} = \prod_{j=2}^{k}\prod_{i=1}^{n_j}p_{ij}^{j}.$$

Substituting (6.22) into Proposition 6.3.0.2 gives

$$\begin{aligned}\mathcal{R}_k(x) &= \sum_{n_1+2n_2+\dots+kn_k=k}\left(\prod_{j=1}^{k}\frac{j^{-n_j}}{n_j!}\right)\left(\mathcal{S}_{n_1}(x)\prod_{j=2}^{k}P(j)^{n_j} + O_k\left(\frac{(\log_2 x)^{n_1-1}}{\log x}\right)\right)\\ &= \sum_{n_1+2n_2+\dots+kn_k=k}\frac{\mathcal{S}_{n_1}(x)}{n_1!}\cdot\prod_{j=2}^{k}\frac{(P(j)/j)^{n_j}}{n_j!} + O_k\left(\frac{(\log_2 x)^{k-1}}{\log x}\right).\end{aligned}$$

This completes the proof of Proposition 6.3.0.1, and therefore Theorem 6.2.3. □

## 6.4 Proof of Proposition 6.3.0.3

We will use a lemma of Nguyen and Pomerance, see [109, Lemma 2.7].

**Lemma 6.4.1** (Nguyen and Pomerance)**.** *For all $x > 1$, we have*

$$\sum_{p>x} \frac{1}{p^2} < \frac{1}{x \log x}.$$

It follows that $\sum_{p>x} p^{-a} < (x^{a-1} \log x)^{-1}$ for any $x > 1$ and $a \geq 2$. Aside from the numerically explicit bounds in Theorem 6.1.3 and Corollaries 6.1.1.2–6.1.1.3, we only use an upper bound of the form $O((x^{a-1} \log x)^{-1})$, which is a consequence of Chebyshev's estimates.

*Proof of Proposition 6.3.0.3.* We proceed by induction on $\ell$. Note the claim holds for $\ell = 0$ by definition of $\mathcal{S}_n(x)$. The claim also holds for $\ell = 1$ and $n = 0$ by Lemma 6.4.1, with the convention that $\mathcal{S}_0(x) = 1$. So assume $\ell = 1$, and take any $n \geq 1$, and $a \geq 2$.

Let $\mathcal{E}$ denote the multi-set of numbers of the form $q_1 \cdots q_n$. Then

$$\begin{aligned} \sum_{\substack{p^a m \leq x \\ m \in \mathcal{E}}} \frac{1}{p^a m} &= \sum_{p \leq (x/2^n)^{1/a}} \frac{1}{p^a} \sum_{\substack{m \leq x/p^a \\ m \in \mathcal{E}}} \frac{1}{m} = \sum_{p \leq (x/2^n)^{1/a}} \frac{1}{p^a} \mathcal{S}_n(x/p^a) \\ &= \sum_{p \leq (x/2^n)^{1/a}} \left( \frac{1}{p^a} S_n\left(\log_2(x/p^a)\right) + O_n\left( \frac{1}{p^a} g_n(x/p^a) \right) \right) \end{aligned} \tag{6.23}$$

by (6.13) from Tenenbaum, where $g_n(x) := (\log_2 x)^{n-1}/\log x$. (Note the condition $t \geq 3$ in (6.13). Indeed $x/p^a \geq 3$ if $n \geq 2$, while if $n = 1$, we appeal directly to Mertens' theorem.) Splitting the sum at $x^{1/(a+1)}$, we bound the error in (6.23) as

$$\sum_{p \leq (x/2^n)^{1/a}} \frac{g_n(x/p^a)}{p^a} \leq (\log_2 x)^{n-1} \Big( \frac{a+1}{a \log x} \sum_{p \leq x^{1/(a+1)}} p^{-a} + \sum_{x^{1/(a+1)} < p} \frac{p^{-a}}{n \log 2} \Big) \ll_{a,n} g_n(x) \tag{6.24}$$

using Lemma 6.4.1 on the right sum above.

Let $s_a(t) := \sum_{p \le t} p^{-a}$. By partial summation, the main term in (6.23) is

$$\begin{aligned}
&\sum_{p \le (x/2^n)^{1/a}} \frac{1}{p^a} S_n\left(\log_2(x/p^a)\right) \\
&\qquad = s_a\left((x/2^n)^{1/a}\right) S_n(\log_2 2^n) + \int_2^{(x/2^n)^{1/a}} \frac{a s_a(t) S_n'\left(\log_2(x/t^a)\right)}{t \log(x/t^a)} \mathrm{d}t \\
&\qquad = P(a) S_n(\log_2 2^n) + P(a) \int_2^{(x/2^n)^{1/a}} \frac{a S_n'\left(\log_2(x/t^a)\right)}{t \log(x/t^a)} \mathrm{d}t \\
&\qquad\quad + O_{a,n}\left( x^{1/a-1} \int_2^{(x/2^n)^{1/a}} \frac{a \left|S_n'\left(\log_2(x/t^a)\right)\right|}{t^a \log t \log(x/t^a)} \mathrm{d}t \right)
\end{aligned}$$

using Lemma 6.4.1. We bound the latter integral in the error as $\ll_{a,n} g_n(x)$ by splitting the interval at $y = x^{1/(a+1)}$. (If $t > y$ then $\log t > \log x/(a+1)$, and if $t < y$ then $\log(x/t^a) > \log x/(a+1)$.) The main term integral may be evaluated exactly, and so the main term is

$$\begin{aligned}
&\sum_{p \le (x/2^n)^{1/a}} \frac{1}{p^a} S_n\left(\log_2(x/p^a)\right) \\
&\qquad = P(a) S_n(\log_2 2^n) - P(a) \left[ S_n\left(\log_2(x/t^a)\right) \right]_2^{(x/2^n)^{1/a}} + O_{a,n}(g_n(x)) \\
&\qquad = P(a) S_n\left(\log_2(x/2^a)\right) + O_{a,n}(g_n(x)) \\
&\qquad = P(a) S_n\left(\log_2 x\right) + O_{a,n}(g_n(x)) \\
&\qquad = P(a) \mathcal{S}_n(x) + O_{a,n}(g_n(x)), \qquad (6.25)
\end{aligned}$$

by (6.13). Here we used $\log_2(x/2^a) = \log_2 x + O_a(1/\log x)$ and $\deg(S_n) = n$, so the error term is distributed over less than $n$ factors of $\log_2 x$. Since the implied constant depends only on $a$ and $n$, it can be taken to depend only on the value of $a + n = k$. Plugging back (6.25) and (6.24) into (6.23), we obtain

$$\sum_{\substack{p^a m \le x \\ m \in \mathcal{E}}} \frac{1}{p^a m} = P(a) \mathcal{S}_n(x) + O_k(g_n(x)). \qquad (6.26)$$

This completes the base case $\ell = 1$, for any $n \ge 0$ and $a \ge 2$.

We now turn to the induction step. Assume that

$$\sum_{v = q_1 \cdots q_n \prod_{i \le \ell} p_i^{a_i} \le x} \frac{1}{v} = \mathcal{S}_n(x) \prod_{i \le \ell} P(a_i) + O_k(g_n(x)),$$

where $k = \Omega(v)$. Let $\mathcal{E}$ be the multi-set of numbers of the form $q_1 \cdots q_n \prod_{i \le \ell} p_i^{a_i}$. Then, for any $a \ge 2$, we have

$$\sum_{\substack{p^a v \le x \\ v \in \mathcal{E}}} \frac{1}{p^a v} = \sum_{p \le (x/2^k)^{1/a}} \frac{1}{p^a} \sum_{\substack{v \le x/p^a \\ v \in \mathcal{E}}} \frac{1}{v}.$$

Thus by the induction hypothesis,

$$\sum_{\substack{p^a v \le x \\ v \in \mathcal{E}}} \frac{1}{p^a v} = \sum_{p \le (x/2^k)^{1/a}} \frac{1}{p^a} \prod_{i \le \ell} P(a_i) \left( \mathcal{S}_n(x/p^a) + O_k\left(g_n(x/p^a)\right)\right)$$

$$= \prod_{i \le \ell} P(a_i) \cdot \sum_{p \le (x/2^k)^{1/a}} \frac{1}{p^a} S_n\left(\log_2(x/p^a)\right) + O_{k,\ell}\left( \sum_{p \le (x/2^k)^{1/a}} \frac{g_n(x/p^a)}{p^a} \right) \tag{6.27}$$

by (6.13). As with (6.24), the error term above is $\ll_{k,\ell} g_n(x)$. As with (6.25), we have

$$\sum_{p \le (x/2^k)^{1/a}} \frac{1}{p^a} S_n\left(\log_2(x/p^a)\right) = P(a) S_n\left(\log_2 x\right) + O_{a,k,n}(g_n(x)).$$

Finally, we note that the implied constant can be taken to depend only on $a+k$, since for any fixed $a$ and $k$, there are only finitely many possibilities for $\ell$ and $n$. Therefore,

$$\sum_{\substack{p^a v \le x \\ v \in \mathcal{E}}} \frac{1}{p^a v} = P(a) \prod_{i \le \ell} P(a_i) \cdot \mathcal{S}_n(x) + O_k(g_n(x)).$$

This completes the proof of Proposition 6.3.0.3. □

## 6.5 The Proof of Theorem 6.1.1

We now prove Theorem 6.1.1. Recall that by Theorem 6.2.3, it suffices to show that $R_k(X+\beta) = V_k(X)$ for each $k$, where $V_k$ is defined as in (6.16). Recall formulas (6.2) and (6.4) for the Mertens constant $\beta$ and the function $\nu(z)$. Also recall the prime zeta function $P(s) = \sum_p p^{-s}$.

**Lemma 6.5.1.** *Define $c_1 = c_1^* = \beta$, $c_j = P(j) - (-1)^j \zeta(j)$ and $c_j^* = (-1)^{j+1}(P(j) + \zeta(j))$, $j \ge 2$. Then for $|z| < 1$, we have the expansions*

$$\nu(z) = \exp\Big( \sum_{j \ge 1} \frac{c_j z^j}{j} \Big) \quad \text{and} \quad \nu^*(z) = \exp\Big( \sum_{j \ge 1} \frac{c_j^* z^j}{j} \Big). \tag{6.28}$$

*Proof.* We prove the expansion for $\nu(z)$. A similar argument can be used to prove the expansion for $\nu^*(z)$. The Weierstrass product formula [137, Theorem II.0.6] yields a Taylor expansion, for $|z| < 1$,

$$\log \Gamma(z+1) = -\gamma z + \sum_{j \ge 2} \zeta(j) \frac{(-z)^j}{j}$$

and so

$$\frac{1}{\Gamma(z+1)} = \exp\left(\gamma z - \sum_{j\geq 2}\zeta(j)\frac{(-z)^j}{j}\right). \tag{6.29}$$

Next, we have

$$\begin{aligned}\prod_p \Big(1-\frac{1}{p}\Big)^z\Big(1-\frac{z}{p}\Big)^{-1} &= \exp\Big(\sum_p z\log(1-\tfrac{1}{p}) - \log(1-\tfrac{z}{p})\Big)\\ &= \exp\left(z\sum_p\Big(\frac{1}{p}+\log(1-\tfrac{1}{p})\Big) + \sum_p\sum_{j\geq 2}\frac{(z/p)^j}{j}\right)\\ &= \exp\left((\beta-\gamma)z + \sum_{j\geq 2}P(j)\frac{z^j}{j}\right).\end{aligned}$$

Combining with (6.4) gives the result. □

From the first assertion of Lemma 6.5.1, we may Taylor expand $\nu$ as

$$\nu(z) = \prod_{j\geq 1}\exp\left(c_j z^j/j\right) = \prod_{j\geq 1}\sum_{n_j\geq 0}\frac{\left(c_j z^j/j\right)^{n_j}}{n_j!} = \sum_{k\geq 0} z^k \sum_{n_1+2n_2+\cdots=k}\prod_{j\geq 1}\frac{(c_j/j)^{n_j}}{n_j!}$$

from which we see that

$$\nu_k := \frac{\nu^{(k)}(0)}{k!} = \sum_{n_1+2n_2+\cdots=k}\prod_{j\geq 1}\frac{(c_j/j)^{n_j}}{n_j!}. \tag{6.30}$$

Note that $\nu_0 = 1, \nu_1 = \beta$, and recall that $V_k(X) = \sum_{j=0}^k \nu_{k-j}X^j/j!$.

On the other hand, with $R_k$ defined as in Theorem 6.2.3, we have

$$R_k(X+\beta) = \sum_{n_1+2n_2+\cdots=k}\prod_{j\geq 2}\frac{(P(j)/j)^{n_j}}{n_j!}\sum_{i=0}^{n_1}\frac{\lambda_{i,n_1}X^i}{n_1!},$$

$$\text{where}\quad \lambda_{i,n} = \frac{n!}{i!}\sum_{m=0}^{n-i}\frac{(\beta-\gamma)^{n-m-i}}{(n-m-i)!}\frac{(1/\Gamma)^{(m)}(1)}{m!}$$

is defined as in (6.14), so that

$$\begin{aligned}R_k(X+\beta) &= \sum_{i=0}^k\frac{X^i}{i!}\sum_{n_1+2n_2+\cdots=k}\prod_{j\geq 2}\frac{(P(j)/j)^{n_j}}{n_j!}\sum_{m=0}^{n_1-i}\frac{(\beta-\gamma)^{n_1-m-i}}{(n_1-m-i)!}\frac{(1/\Gamma)^{(m)}(1)}{m!}\\ &= \sum_{i=0}^k\frac{X^i}{i!}\sum_{n_1+2n_2+\cdots=k-i}\prod_{j\geq 2}\frac{(P(j)/j)^{n_j}}{n_j!}\sum_{m=0}^{n_1}\frac{(\beta-\gamma)^{n_1-m}}{(n_1-m)!}\frac{(1/\Gamma)^{(m)}(1)}{m!}\\ &= \sum_{i=0}^k\frac{X^i}{i!}\sum_{n_1+2n_2+\cdots=k-i}\frac{G^{(n_1)}(0)}{n_1!}\prod_{j\geq 2}\frac{(P(j)/j)^{n_j}}{n_j!} =: \sum_{i=0}^k\mu_{k-i}\frac{X^i}{i!}\end{aligned}$$

via $n_1 \mapsto n_1 - i$ (note $n_1 \geq i$, otherwise the inner sum on $m$ vanishes), where by the product rule

$$\sum_{m=0}^{n_1} \frac{(\beta-\gamma)^{n_1-m}}{(n_1-m)!} \frac{(1/\Gamma)^{(m)}(1)}{m!} = \frac{1}{n_1!} \frac{d^{n_1}}{dz^{n_1}} \Big[ \frac{e^{(\beta-\gamma)z}}{\Gamma(z+1)} \Big]_{z=0} = \frac{G^{(n_1)}(0)}{n_1!},$$

for $G(z) = e^{(\beta-\gamma)z}/\Gamma(z+1)$. Also by the product rule,

$$\begin{aligned} \mu_k :&= \sum_{n_1=0}^{k} \frac{G^{(n_1)}(0)}{n_1!} \sum_{2n_2+\cdots=k-n_1} \prod_{j\geq 2} \frac{(P(j)/j)^{n_j}}{n_j!} \\ &= \sum_{n_1+m=k} \frac{G^{(n_1)}(0)}{n_1!} \frac{1}{m!} \frac{d^m}{dz^m} \Big[ \exp\Big( \sum_{j\geq 2} P(j) \frac{z^j}{j} \Big) \Big]_{z=0} \\ &= \frac{1}{k!} \frac{d^k}{dz^k} \Big[ G(z) \exp\Big( \sum_{j\geq 2} P(j) \frac{z^j}{j} \Big) \Big]_{z=0} = \frac{\nu^{(k)}(0)}{k!} = \nu_k, \end{aligned}$$

recalling the expansion (6.28). This completes the proof of Theorem 6.1.1.

## 6.6 Proof of Equivalence of Theorems 6.2.1 and 6.2.2

In this section we provide a direct proof that the coefficients appearing in Tenenbaum's formula (6.13) and the formula (6.15) are equal. First, we let $b_m = a_m/m!$ so that

$$b_m = \frac{1}{m} \sum_{i=1}^{m-1} (-1)^i \zeta(i+1) b_{m-1-i} = \frac{1}{m} \sum_{j=2}^{m} (-1)^{j-1} \zeta(j) b_{m-j}.$$

This recursive identity for $b_m$ implies that, by [88, Lemma 2.2], $b_m$ is given explicitly as

$$b_m = \sum_{2n_2+3n_3\cdots=m} \prod_{j\geq 2} \frac{((-1)^{j-1}\zeta(j)/j)^{n_j}}{n_j!} = \frac{1}{m!} \frac{\mathrm{d}^m}{\mathrm{d}z^m} \Big[ \exp\Big( -\sum_{j\geq 2} \zeta(j)(-z)^j/j \Big) \Big]_{z=0}$$

and so the Weierstrass product formula in (6.29) gives

$$a_m = m!\, b_m = \frac{\mathrm{d}^m}{\mathrm{d}z^m} \Big[ \frac{e^{-\gamma z}}{\Gamma(z+1)} \Big]_{z=0} = \sum_{i=0}^{m} \binom{m}{i} (-\gamma)^{m-i} \Big( \frac{1}{\Gamma} \Big)^{(i)}(1). \tag{6.31}$$

Thus by the binomial theorem, the main term in (6.15) is

$$\begin{aligned}
\mathcal{S}_k(x) + \text{err.} &= \sum_{m=0}^{k} \binom{k}{m} a_m (\log_2 x + \beta)^{k-m} \\
&= \sum_{m=0}^{k} \binom{k}{m} a_m \sum_{j=0}^{k-m} \binom{k-m}{j} \beta^{k-m-j} (\log_2 x)^j \\
&= \sum_{m=0}^{k} \binom{k}{m} \sum_{i=0}^{m} \binom{m}{i} (-\gamma)^{m-i} \Big(\frac{1}{\Gamma}\Big)^{(i)}(1) \sum_{j=0}^{k-m} \binom{k-m}{j} \beta^{k-m-j} (\log_2 x)^j \\
&= \sum_{j=0}^{k} (\log_2 x)^j \sum_{i=0}^{k-j} \Big(\frac{1}{\Gamma}\Big)^{(i)}(1) \sum_{m=i}^{k-j} \frac{k!\, \beta^{k-m-j} (-\gamma)^{m-i}}{i!(m-i)!j!(k-m-j)!} \\
&= \sum_{j=0}^{k} (\log_2 x)^j \sum_{i=0}^{k-j} \Big(\frac{1}{\Gamma}\Big)^{(i)}(1) \frac{k!}{i!j!(k-j-i)!} (\beta - \gamma)^{k-j-i}
\end{aligned}$$

Hence we obtain Tenenbaum's main term in (6.13),

$$\mathcal{S}_k(x) + \text{err.} = \sum_{j=0}^{k} (\log_2 x)^j \sum_{i=0}^{k-j} \binom{k}{i, j, k-j-i} (\beta - \gamma)^{k-j-i} \Big(\frac{1}{\Gamma}\Big)^{(i)}(1) \;=\; S_k(\log_2 x).$$

This completes the proof.

## 6.7 The Proof of Theorem 6.1.1

We first establish the following recurrence relations for the sequences $(\nu_k)_{k=0}^{\infty}$ and $(\nu_k^*)_{k=0}^{\infty}$. These recurrences are analogous to those of $(d_k)_{k=0}^{\infty}$ in the smooth setting of Theorem 6.2.4 above, as well as Proposition 3.1 from [86].

**Proposition 6.7.0.1.** *Define sequences $c_1 = c_1^* = \beta$, $c_j = P(j) - (-1)^j \zeta(j)$ and $c_j^* = (-1)^{j+1}(P(j) + \zeta(j))$, $j \geq 2$. Then the sequences $(\nu_k)_{k=0}^{\infty}$ and $(\nu_k^*)_{k=0}^{\infty}$ are given recursively by $\nu_0 = \nu_0^* = 1$,*

$$\nu_k = \frac{1}{k} \sum_{j=1}^{k} \nu_{k-j} c_j \quad \textit{and} \quad \nu_k^* = \frac{1}{k} \sum_{j=1}^{k} \nu_{k-j}^* c_j^*.$$

*Proof.* We prove the relation for $\nu_k$ and note that a similar argument gives that of $\nu_k^*$. We proceed by induction on $k \geq 1$. For the base case $k = 1$, we have $\nu_1 = \beta = \nu_0 c_1$.

Now assume the claim for all $1 \le r < k$. By the explicit formula (6.30) for $\nu_k$ we have

$$\sum_{r=1}^{k} \nu_{k-r} c_r = \sum_{r=1}^{k} c_r \sum_{n_1+\cdots=k-r} \prod_{j\ge 1} \frac{(c_j/j)^{n_j}}{n_j!} = \sum_{r=1}^{k} \sum_{n_1+\cdots=k-r} \frac{c_r^{n_r+1}}{r^{n_r}\, n_r!} \prod_{\substack{j\ge 1 \\ j\ne r}} \frac{(c_j/j)^{n_j}}{n_j!}$$

$$= \sum_{r=1}^{k} \sum_{\substack{n_1+\cdots=k \\ n_r\ge 1}} r n_r \prod_{j\ge 1} \frac{(c_j/j)^{n_j}}{n_j!} = \sum_{n_1+\cdots=k} \prod_{j\ge 1} \frac{(c_j/j)^{n_j}}{n_j!} \sum_{\substack{1\le r\le k \\ n_r\ge 1}} r n_r = k\nu_k.$$

In the last step, we dropped the condition $n_r \ge 1$ (since $rn_r = 0$ for $n_r = 0$) which gives $\sum_{r=1}^{k} r n_r = k$. This completes the proof. □

Note this recursion enables rapid computation of $\nu_k, \nu_k^*$ to high precision. We show results for $k \le 10$ below.

| $k$ | $\nu_k$ | $\nu_k^*$ |
|---|---|---|
| 0 | 1 | 1 |
| 1 | $\beta = 2.61497 \cdot 10^{-1}$ | $\beta = 2.61497 \cdot 10^{-1}$ |
| 2 | $-5.62153 \cdot 10^{-1}$ | $-1.01440 \cdot 10^{0}$ |
| 3 | $3.05978 \cdot 10^{-1}$ | $1.87717 \cdot 10^{-1}$ |
| 4 | $2.62973 \cdot 10^{-2}$ | $3.44297 \cdot 10^{-1}$ |
| 5 | $-6.44501 \cdot 10^{-2}$ | $-1.86153 \cdot 10^{-1}$ |
| 6 | $3.64064 \cdot 10^{-2}$ | $-1.50297 \cdot 10^{-2}$ |
| 7 | $-4.70865 \cdot 10^{-3}$ | $4.29836 \cdot 10^{-2}$ |
| 8 | $-4.33984 \cdot 10^{-4}$ | $-1.30388 \cdot 10^{-2}$ |
| 9 | $1.5085 \cdot 10^{-3}$ | $-1.57532 \cdot 10^{-3}$ |
| 10 | $-1.83548 \cdot 10^{-4}$ | $2.17630 \cdot 10^{-3}$ |

Now we prove the main theorem of the section.

*Proof of Theorem 6.1.1.* Note that $\nu(z) = G(z)C(z)$, where

$$G(z) := \frac{1}{\Gamma(z+1)} \prod_p \left(1 - \frac{1}{p}\right)^z e^{z/p} = \frac{e^{(\beta-\gamma)z}}{\Gamma(z+1)},$$

$$C(z) := \prod_p \left(1 - \frac{z}{p}\right)^{-1} e^{-z/p} = \exp\Big(\sum_{j\ge 2} P(j)\frac{z^j}{j}\Big).$$

By the product rule, we thus have

$$\nu_k = \sum_{n=0}^{k} \frac{G^{(n)}(0)}{n!} d_{k-n}, \tag{6.32}$$

where

$$d_k := \frac{1}{k!} C^{(k)}(0) = \sum_{2n_2+3n_3+\cdots=k} \prod_{j\ge 2} \frac{(P(j)/j)^{n_j}}{n_j!}.$$

Now we appeal to Theorem 6.2.4, which gives

$$d_k = \sum_{p<q} \delta_p \, p^{-k} \ + \ O_q(q^{-k})$$

for any prime $q$, where

$$\delta_p := e^{-1} \prod_{p' \neq p} \Big(1 - \frac{p}{p'}\Big)^{-1} e^{-p/p'} = \lim_{z \to p} (1 - z/p) C(z). \tag{6.33}$$

Hence (6.32) becomes

$$\nu_k = \sum_{p<q} \delta_p \sum_{n=0}^{k} \frac{G^{(n)}(0)}{n!} \, p^{n-k} \ + \ O_q\Big( \sum_{n=0}^{k} \frac{G^{(n)}(0)}{n!} q^{n-k} \Big). \tag{6.34}$$

Since $G$ is entire, we have by estimate (6.34) that

$$\begin{aligned} \nu_k &= \sum_{p<q} \delta_p p^{-k} \left( G(p) - \sum_{n>k} \frac{G^{(n)}(0)}{n!} p^n \right) + O_q(G(q) q^{-k}) \\ &= \sum_{p<q} \beta_p p^{-k} \ + \ O_q(q^{-k}), \end{aligned}$$

where the last equality holds by Lemmas 6.7.1 and 6.7.2 below. □

**Lemma 6.7.1.** *We have $\delta_p G(p) = \beta_p$ for each prime $p$.*

*Proof.* By (6.9), (6.33), and the definition of $G$, we have

$$\begin{aligned} \frac{\delta_p G(p)}{\beta_p} &= e^{(\beta-\gamma)p-1} \left(1 - \frac{1}{p}\right)^{-p} \prod_{p' \neq p} \left(1 - \frac{1}{p'}\right)^{-p} e^{-p/p'} = e^{(\beta-\gamma)p} \prod_{p'} \left(1 - \frac{1}{p'}\right)^{-p} e^{-p/p'} \\ &= e^{(\beta-\gamma)p} \exp\left( -p \sum_{p'} \left( \log\left(1 - \frac{1}{p'}\right) + \frac{1}{p'} \right) \right) = 1. \end{aligned}$$

□

**Lemma 6.7.2.** *For any $m \geq 2$, we have $G^{(k)}(0)/k! \ll_m m^{-k}$ as $k \to \infty$.*

*Proof.* It suffices to prove $G^{(k)}(0)/k! \ll_m (k+1)^m m^{-k}$ for any $m \geq 2$, in which case taking $m' = m + 1$ gives the result. Now to show this, given a fixed $m$ we have

$$\begin{aligned} G(z) &= \exp(\beta z - \sum_{j \geq 2} \zeta(j)(-z)^j/j) \\ &= e^{\beta z} \exp(\sum_{1 \leq n < m} [\log(1 + z/n) - z/n] - \sum_{j \geq 2} \bar{\zeta}(j)(-z)^j/j) = G_0(z) G_1(z) \end{aligned} \tag{6.35}$$

letting $G_0(z) = e^{\beta z}\prod_{n<m}(1+z/n)e^{-z/n}$ and $G_1(z) = \exp(-\sum_{j\geq 2}\bar{\zeta}(j)(-z)^j/j)$, where $\bar{\zeta}(j) := \sum_{n\geq m} n^{-j}$. Note that $\bar{\zeta}(2) = \sum_{n\geq m} n^{-2} \leq m^{-2} + \int_m^\infty t^{-2}\mathrm{d}t = (m+1)/m^2$, so by induction $\bar{\zeta}(j) \leq (m+1)/m^j$ for all $j \geq 2$. Thus

$$G_1^{(k)}(0) = (-1)^k k! \sum_{2n_2+\cdots=k} \prod_{j\geq 2} \frac{(-\bar{\zeta}(j)/j)^{n_j}}{n_j!},$$
$$|G_1^{(k)}(0)| \leq k! \sum_{2n_2+\cdots=k} \prod_{j\geq 2} \frac{((m+1)/jm^j)^{n_j}}{n_j!} = \widetilde{G}_1^{(k)}(0)$$

for

$$\widetilde{G}_1(z) := \exp((m+1)\sum_{j\geq 2}(z/m)^j/j) = \frac{e^{-(m+1)(z/m)}}{(1-z/m)^{m+1}}.$$

Note the derivatives

$$\frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k}\Big[\frac{1}{(1-z/m)^{m+1}}\Big]_{z=0} = \frac{m^{-k}}{m!}\prod_{1\leq j\leq m}(k+j) \leq \frac{(k+1)^m}{m^k},$$
$$\frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k}\big[e^{-(m+1)(z/m)}\big]_{z=0} = \frac{1}{k!}\big(-(m+1)/m\big)^k \leq \frac{(1+1/m)^k}{k!}$$

and since $\sum_{k\in K} u^k/k! \leq e^{|u|}$ for any $u \in \mathbb{R}$, any set $K \subset \mathbb{N}$, by the product rule we have

$$\frac{\widetilde{G}_1^{(k)}(0)}{k!} \leq \sum_{a+b=k} \frac{(a+1)^m}{m^a}\frac{(1+1/m)^b}{b!} = m^{-k}\sum_{b=0}^{k}(k+1-b)^m\frac{(m+1)^b}{b!}$$
$$\leq e^{m+1}\frac{(k+1)^m}{m^k} \ll_m \frac{(k+1)^m}{m^k}.$$

We also have

$$G_0(z) = e^{\beta z}\prod_{j<m}(1+z/j)e^{-z/j} = \sum_{k\geq 0}\frac{z^k}{k!}(\beta - \sum_{j<m}1/j)^k\prod_{j<m}(1+z/j).$$

In particular $G_0^{(k)}(0) \ll_m |\beta - \sum_{j<m}1/j|^k$. Thus by the product rule for $G = G_0 \cdot G_1$ we have

$$\frac{G^{(k)}(0)}{k!} = \sum_{b+c=k}\frac{G_0^{(b)}(0)}{b!}\frac{G_1^{(c)}(0)}{c!},$$
$$\frac{|G^{(k)}(0)|}{k!} \leq \sum_{b+c=k}\frac{|G_0^{(b)}(0)|}{b!}\frac{\widetilde{G}_1^{(c)}(0)}{c!} \ll_m \sum_{b+c=k}\frac{|\beta-\sum_{j<m}1/j|^b}{b!}\frac{(c+1)^m}{m^c}$$
$$\ll_m \frac{(k+1)^m}{m^k}.$$

This gives the claim as desired. □

## 6.8 The Proofs of Theorems 6.1.2 and 6.1.3

Recall that $E(x) := \sum_{p\le x} 1/p - (\log_2 x + \beta)$. Note $E(t) \ll_A (\log t)^{-A}$ for all $A > 0$, which implies $\int_2^\infty |E(t)|(\log t)^j/t\mathrm{d}t$ converges for all $j$. So we may define the constants

$$\gamma_j = \int_2^\infty E(t)(\log t)^{j-1}\frac{\mathrm{d}t}{t}. \tag{6.36}$$

In particular, we have the relation $\alpha_j = (\log 2)^j(1/j - \log_2 2 - \beta)/j + \gamma_j$.

We cite some useful lemmas.

**Lemma 6.8.1** (Rosser & Schoenfeld [124, Theorems 5 & 20])**.** *We have*

$$-1/(2\log^2 x) < E(x) < 1/\log^2 x, \quad (x > 1), \tag{6.37}$$

$$0 < E(x), \quad (1 < x \le 10^8). \tag{6.38}$$

**Lemma 6.8.2** (Dusart [36, Theorem 6.10],[37, Theorem 5.6])**.** *We have*

$$|E(x)| \le 1/(10\log^2 x) + 4/(15\log^3 x), \quad (x \ge 10372), \tag{6.39}$$

$$|E(x)| \le 0.2/\log^3 x, \quad (x \ge 2278383).^2 \tag{6.40}$$

For ease of notation, we define

$$T(x) := \mathcal{R}_1(x) = \sum_{p\le x}\frac{1}{p}.$$

**Lemma 6.8.3.** *For all $N \ge 0$, and $\gamma_j$ as in* (6.36)*, we have*

$$\int_2^{\sqrt{x}} \frac{E(t)}{t\log(x/t)}\mathrm{d}t = \sum_{j=1}^{N}\frac{\gamma_j}{\log^j x} + O_N\left((\log x)^{-N-1}\right).$$

*Proof.* Note for $t \le \sqrt{x}$, we have the geometric series

$$\sum_{j=0}^{N-1}\left(\frac{\log t}{\log x}\right)^j = \frac{1-(\log t/\log x)^N}{1-(\log t/\log x)} = \frac{\log x}{\log(x/t)}(1 - O(2^{-N})).$$

[2]We were recently made aware of an issue in [37] which affects (6.40). This has been resolved in a paper [16] of Broadbent et al. which provides the necessary bounds on Chebyshev's function $\theta(x)$ to guarantee the validity of (6.40).

So recalling $E(t) \ll_A (\log t)^{-A}$ for all $A > 0$, we may interchange sum and integral to obtain

$$\begin{aligned}\int_2^{\sqrt{x}} \frac{E(t)}{t\log(x/t)}\mathrm{d}t &= \sum_{j=0}^{N-1} \frac{1+O(2^{-N})}{(\log x)^{j+1}} \int_2^{\sqrt{x}} (\log t)^j E(t)\frac{\mathrm{d}t}{t} \\ &= \sum_{j=1}^{N} \frac{\gamma_j}{(\log x)^j} + O_N((\log x)^{-N-1})\end{aligned}$$

since for all $j \le N$, letting $A = 2(N+1)$,

$$\int_{\sqrt{x}}^{\infty} (\log t)^j E(t)\frac{\mathrm{d}t}{t} \ll_N (\log x)^{-N-1}.$$

This completes the proof. □

We give the following limit characterization of the constants $\alpha_j$.

**Lemma 6.8.4.** *For all $N \ge j \ge 1$, and $\alpha_j$ as in* (6.10)*, we have*

$$\alpha_j = \frac{1}{j}\left(\frac{\log^j x}{j} - \sum_{p\le x}\frac{\log^j p}{p}\right) + O_N((\log x)^{-N}).$$

*Proof.* By partial summation and Mertens' second theorem, we have

$$\begin{aligned}\sum_{p\le x}\frac{\log^j p}{p} &= T(x)\log^j x - j\int_2^x T(t)\log^{j-1} t\frac{\mathrm{d}t}{t} \\ &= (\log_2 x + \beta + E(x))\log^j x - j\int_2^x (\log_2 t + \beta + E(t))\log^{j-1} t\frac{\mathrm{d}t}{t} \\ &= E(x)\log^j x + \log^j x/j + \log^j 2(\log_2 2 + \beta - 1/j) - j\int_2^x E(t)\log^{j-1} t\frac{\mathrm{d}t}{t}.\end{aligned}$$

Hence by definition of $\alpha_j$,

$$\frac{1}{j}\left(\frac{\log^j x}{j} - \sum_{p\le x}\frac{\log^j p}{p}\right) = \alpha_j - E(x)\log^j x/j - \int_x^{\infty} E(t)\log^{j-1} t\frac{\mathrm{d}t}{t}.$$

Recalling $E(x) \ll_N (\log x)^{-2N}$ completes the proof. □

**Lemma 6.8.5.** *We have*

$$\mathcal{R}_2(x) = \sum_{p\le\sqrt{x}}\frac{1}{p}T(x/p) - \frac{1}{2}T(\sqrt{x})^2 + \frac{1}{2}\sum_{p\le\sqrt{x}}\frac{1}{p^2}.$$

*Proof.* Note that $\mathbb{N}_2 \cap [1,x] = \{pq \le x : p < q\} \cup \{p^2 : p \le \sqrt{x}\}$. Also, we have $pq \le x, p < q$ if and only if $p \le \sqrt{x}$ and $p < q \le x/p$. Therefore,

$$\begin{aligned}\mathcal{R}_2(x) &= \sum_{p\le\sqrt{x}} \sum_{p<q\le\frac{x}{p}} \frac{1}{pq} + \sum_{p\le\sqrt{x}} \frac{1}{p^2} \\ &= \sum_{p\le\sqrt{x}} \frac{1}{p}\left(T\left(x/p\right) - T(p)\right) + \sum_{p\le\sqrt{x}} \frac{1}{p^2}.\end{aligned}$$

Thus it suffices to note that by the multinomial theorem, we have

$$2\sum_{p\le\sqrt{x}} \frac{T(p)}{p} = \sum_{p\le\sqrt{x}} \frac{1}{p^2} + \left(\sum_{p\le\sqrt{x}} \frac{1}{p}\right)^2. \qquad \square$$

*Proof of Theorem 6.1.2.* By Lemma 6.8.5, $\mathcal{R}_2(x) = A_1(x) + A_2(x) + A_3(x)$, where

$$\begin{aligned}A_1(x) &:= \sum_{p\le\sqrt{x}} \frac{1}{p} T\left(x/p\right) = \sum_{p\le\sqrt{x}} \frac{\log_2(x/p) + \beta}{p} + \sum_{p\le\sqrt{x}} \frac{E\left(x/p\right)}{p}, \\ A_2(x) &:= -\frac{1}{2} T(\sqrt{x})^2 = -\frac{1}{2}(\log_2\sqrt{x} + \beta + E(\sqrt{x}))^2, \\ A_3(x) &:= \frac{1}{2}\sum_{p\le\sqrt{x}} p^{-2}.\end{aligned}$$

Similarly, $\mathcal{R}_2^*(x)$ is given by negating the last term above. We next write $A_1(x) = B_1(x) + B_2(x) + B_3(x)$, where

$$B_1(x) := \sum_{p\le\sqrt{x}} \frac{\log_2(x/p)}{p}, \quad B_2(x) := \sum_{p\le\sqrt{x}} \frac{\beta}{p}, \quad B_3(x) := \sum_{p\le\sqrt{x}} \frac{E(x/p)}{p}.$$

We have $B_2(x) = \beta \cdot T(\sqrt{x}) = \beta(\log_2\sqrt{x} + \beta + E(\sqrt{x}))$. By partial summation,

$$\begin{aligned}B_1(x) &= T(\sqrt{x})\log_2\sqrt{x} + \int_2^{\sqrt{x}} \frac{T(t)}{t\log(x/t)}\mathrm{d}t \\ &= (\log_2\sqrt{x} + \beta + E(\sqrt{x}))\log_2\sqrt{x} \ + \int_{\log 2}^{\log\sqrt{x}} \frac{\log u + \beta}{\log x - u}\mathrm{d}u + \int_2^{\sqrt{x}} \frac{E(t)}{t\log(x/t)}\mathrm{d}t.\end{aligned}$$

Denoting the dilogarithm $\mathrm{Li}_2(z) = \sum_{j\ge1} z^j/j^2$, the integral on the left equals

$$\begin{aligned}\int_{\log 2}^{\log\sqrt{x}} \frac{\log u + \beta}{\log x - u}\mathrm{d}u &= -\Big[(\beta + \log u)\log(1 - \tfrac{u}{\log x}) + \mathrm{Li}_2(\tfrac{u}{\log x})\Big]_{\log 2}^{\frac{1}{2}\log x} \\ &= (\beta + \log_2 2)\log(1 - \tfrac{\log 2}{\log x}) + \mathrm{Li}_2(\tfrac{\log 2}{\log x}) \\ &\quad + (\beta + \log_2\sqrt{x})\log 2 - \mathrm{Li}_2(\tfrac{1}{2}).\end{aligned}$$

After simplifying and noting $\mathrm{Li}_2(1/2) = \zeta(2)/2 - (\log 2)^2/2$, we have

$$\mathcal{R}_2(x) = \frac{1}{2}(\log_2 x + \beta)^2 + A_3(x) - \frac{\zeta(2)}{2} + (\beta + \log_2 2)\log\left(1 - \tfrac{\log 2}{\log x}\right) + \mathrm{Li}_2\left(\tfrac{\log 2}{\log x}\right) - \frac{1}{2}E(\sqrt{x})^2 + \sum_{p \le \sqrt{x}} \frac{E(x/p)}{p} + \int_2^{\sqrt{x}} \frac{E(t)}{t\log(x/t)}\mathrm{d}t. \tag{6.41}$$

By Lemma 6.4.1, $A_3(x) = P(2)/2 + O(x^{-1/2})$. And $E(t) \ll_A 1/\log^A t$ implies

$$E(\sqrt{x})^2 \text{ and } \sum_{p \le \sqrt{x}} \frac{E(x/p)}{p} \ll_N \frac{1}{\log^{N+1} x}$$

using Mertens' second theorem. Thus by Lemma 6.8.3,

$$\begin{aligned}\mathcal{R}_2(x) &= \frac{1}{2}(\log_2 x + \beta)^2 + \frac{P(2) - \zeta(2)}{2} + \sum_{j=1}^{N} \frac{\gamma_j}{\log^j x} \\ &\quad + (\beta + \log_2 2)\log\left(1 - \tfrac{\log 2}{\log x}\right) + \mathrm{Li}_2\left(\tfrac{\log 2}{\log x}\right) \ + \ O_N\left(\frac{1}{\log^{N+1} x}\right) \\ &= \frac{1}{2}(\log_2 x + \beta)^2 + \frac{P(2) - \zeta(2)}{2} + \sum_{j=1}^{N} \frac{\gamma_j}{\log^j x} \\ &\quad + \sum_{j=1}^{N} \frac{\log^j 2}{j\log^j x}\Big(\frac{1}{j} - \log_2 2 - \beta\Big) \ + \ O_N\left(\frac{1}{\log^{N+1} x}\right)\end{aligned}$$

using $\mathrm{Li}_2(z) = \sum_{j\ge1} z^j/j^2$ and $\log(1-z) = -\sum_{j\ge1} z^j/j$ with $z = \log 2/\log x$.

Hence recalling $\alpha_j = (\log 2)^j(1/j - \log_2 - \beta)/j + \gamma_j$ completes the proof. □

*Proof of Theorem 6.1.3.* We verify the bound for all $x \le x_0 := 10^9$ by computer. Let $x > x_0$. We bound the terms in (6.41). By (6.39) we have $-0.0003/\log^2 x < -E(\sqrt{x})^2 \le 0$. By Lemma 6.8.3 we have

$$\begin{aligned}\int_2^{\sqrt{x}} \frac{E(t)}{t\log(x/t)}\mathrm{d}t &= \frac{\gamma_1}{\log x} - \frac{1}{\log x}\int_{\sqrt{x}}^{\infty} \frac{E(t)}{t}\mathrm{d}t + \frac{1}{\log x}\int_2^{\sqrt{x}} \frac{E(t)}{t}\frac{\log t}{\log(x/t)}\mathrm{d}t \\ &= \gamma_1/\log x + I_1 + I_2,\end{aligned}$$

say. To bound $I_1$ above (resp. below) we use (6.38) and (6.40) (resp. (6.39)), obtaining $-0.2515/\log^2 x < I_1 < 0.0194/\log^2 x$. To bound $I_2$ above (resp. below) we use (6.37) (resp. (6.38) and (6.40)), obtaining

$$-\frac{0.0218}{\log^2 x} < I_2 < \frac{\log_2(x/2) - \log_2 2}{\log^2 x}.$$

By Lemma 6.8.2 and following the method in [7, Theorem 5.1], we have

$$-\frac{0.2161}{\log^2 x} < \sum_{p\le\sqrt{x}} \frac{E\left(x/p\right)}{p} < \sum_{p\le\sqrt{x}} \frac{0.1258}{p\log^2(x/p)} \le \frac{0.1258\log_2 x + 0.1593}{\log^2 x}.$$

Next, by Lemma 6.4.1 we have

$$-\frac{0.0007}{\log^2 x} < -\frac{1}{2}\sum_{p>\sqrt{x}} p^{-2} < 0.$$

We find by bounding series expansions that

$$-\frac{0.2458\beta}{\log^2 x} < \beta\log\left(1-\frac{\log 2}{\log x}\right) + \frac{\beta\log 2}{\log x} < -\frac{0.2402\beta}{\log^2 x},$$

$$\frac{0.1201}{\log^2 x} < \mathrm{Li}_2\left(\frac{\log 2}{\log x}\right) - \frac{\log 2}{\log x} < \frac{0.1221}{\log^2 x},$$

and

$$\frac{0.0818}{\log^2 x} < (\log_2 2)\log\left(1-\frac{\log 2}{\log x}\right) + \frac{\log_2 2\log 2}{\log x} < \frac{0.0962}{\log^2 x},$$

where we note that $-\log_2 2 > 0$ for the last inequality above. Combining all bounds, we complete the proof of Theorem 6.1.3. □

## 6.9 A new proof of Theorem 6.1.4

We provide an alternate proof of Crisan and Erban's asymptotic expansion for $\mathcal{N}_2(x)$ using our refined estimate for $\mathcal{R}_2(x)$ given in Theorem 6.1.2. The proof relies on a strong form of the prime number theorem, i.e. $E(x) \ll_A (\log x)^{-A}$ for all $A > 0$.

*Proof of Theorem 6.1.4.* By partial summation,

$$\mathcal{N}_2(x) = x\mathcal{R}_2(x) - \int_4^x \mathcal{R}_2(t)\mathrm{d}t.$$

So by Theorem 6.1.2, at admissible error $\epsilon(x) \ll x\log_2 x/\log^{N+1} x$, we have

$$\begin{aligned}
&\mathcal{N}_2(x) + \epsilon(x) \qquad (6.42)\\
&= \frac{x}{2}(\log_2 x+\beta)^2 + x\nu_2 + \sum_{j=1}^{N}\frac{x\alpha_j}{\log^j x} - \int_4^x\left(\frac{1}{2}(\log_2 t+\beta)^2 + \nu_2 + \sum_{j=1}^{N}\frac{\alpha_j}{\log^j t}\right)\mathrm{d}t\\
&= \frac{x}{2}(\log_2 x)^2 + \beta x\log_2 x - \int_4^x\left(\frac{1}{2}(\log_2 t)^2 + \beta\log_2 t\right)\mathrm{d}t\\
&\qquad + \sum_{j=1}^{N}\alpha_j\Big(\frac{x}{\log^j x} - \int_4^x\frac{\mathrm{d}t}{\log^j t}\Big)\\
&= (\log_2 x+\beta)\mathrm{li}(x) - \int_4^x\frac{\mathrm{li}(t)}{t\log t}\mathrm{d}t - \sum_{j=1}^{N-1}\frac{j\alpha_j}{j!}\sum_{n=j}^{N-1}\frac{n!\,x}{\log^{n+1} x}
\end{aligned}$$

using the expansions

$$\begin{aligned}
\int_4^x \log_2 t \mathrm{d}t &= x\log_2 x - \mathrm{li}(x) + O(1),\\
\int_4^x (\log_2 t)^2 \mathrm{d}t &= x(\log_2 x)^2 - 2\mathrm{li}(x)\log_2 x + O(1) + 2\int_4^x \frac{\mathrm{li}(t)}{t\log t}\mathrm{d}t,\\
\int_4^x \frac{\mathrm{d}t}{\log^{j+1} t} &= \sum_{n=j}^{N-1} \frac{n!}{j!}\frac{x}{\log^{n+1} x} + O_N\left(\frac{x}{\log^{N+1} x}\right) \qquad (j \geq 0).
\end{aligned}$$

Note $j = 0$ gives expansion for the logarithmic integral $\mathrm{li}(x) = \int_2^x \mathrm{d}t/\log t$, so that (6.42) becomes

$$\begin{aligned}
&(\log_2 x + \beta)\sum_{n=0}^{N-1} \frac{n!\,x}{\log^{n+1} x} - \sum_{n=0}^{N-1} n! \int_4^x \frac{\mathrm{d}t}{\log^{n+2} t} - \sum_{j=1}^{N-1} \frac{j\alpha_j}{j!}\sum_{n=j}^{N-1} \frac{n!\,x}{\log^{n+1} x}\\
&= (\log_2 x + \beta)\sum_{n=0}^{N-1} \frac{n!\,x}{\log^{n+1} x} - \sum_{n=0}^{N-1} \frac{1}{n+1}\sum_{m=n+1}^{N-1} \frac{m!\,x}{\log^{m+1} x} - \sum_{j=1}^{N-1} \frac{j\alpha_j}{j!}\sum_{n=j}^{N-1} \frac{n!\,x}{\log^{n+1} x}.
\end{aligned}$$

Hence we conclude

$$\mathcal{N}_2(x) + \epsilon(x) = \sum_{n=0}^{N-1} \frac{n!\,x}{\log^{n+1} x}\Big(\log_2 x + \beta - \sum_{j=1}^{n}\Big(\frac{j\alpha_j}{j!} + \frac{1}{j}\Big)\Big).$$

By definition of $D_n$, this completes the proof of Theorem 6.1.4. □

# Chapter 7

# Almost primes and the Banks–Martin conjecture

## 7.1 Introduction

For each $k \geq 1$, let $\mathcal{P}^{(k)} = \{n : \Omega(n) = k\}$ be the set of numbers with $k$ prime factors with multiplicity, so-called $k$-almost primes. The sets $\mathcal{P}^{(k)}$ are the prototypical examples of *primitive* sets of natural numbers $> 1$, i.e., no member of the set divides any other. In 2013, Banks and Martin [6] conjectured that $f(\mathcal{P}^{(k)}) > f(\mathcal{P}^{(k+1)})$ for all $k \geq 1$, that is,

$$\sum_{n\in\mathcal{P}} \frac{1}{n\log n} > \sum_{n\in\mathcal{P}^{(2)}} \frac{1}{n\log n} > \sum_{n\in\mathcal{P}^{(3)}} \frac{1}{n\log n} > \cdots.$$

Their conjecture exemplifies the general view that $f(A)$ is sensitive to the prime factorizations of $n \in A$. Indeed, Banks and Martin [6] showed that, for any sufficiently small set of primes $\mathcal{Q}$ (e.g., if $\sum_{p\in\mathcal{Q}} 1/p < 1.74$), the analogous statement holds,

$$f\big(\mathcal{P}^{(k)}(\mathcal{Q})\big) > f\big(\mathcal{P}^{(k+1)}(\mathcal{Q})\big) \qquad \text{for all } k \geq 1. \tag{7.1}$$

Here $A(\mathcal{Q})$ denotes the numbers in $A$ composed only of prime factors in $\mathcal{Q}$. In fact they showed $f\big(\mathcal{P}^{(k)}(\mathcal{Q})\big) \geq f(A(\mathcal{Q}))$ for all primitive sets $A$ with $\Omega(n) \geq k$ for each $n \in A$.

Zhang [143] proved that $f(\mathcal{P}) > f(\mathcal{P}^{(k)})$ for each $k \geq 2$. Pomerance and the author [90] proved that $f(\mathcal{P}^{(k)}) \gg 1$. Bayless, Kinlaw, and Klyve [7] recently showed that $f(\mathcal{P}^{(2)}) > f(\mathcal{P}^{(3)})$, providing bounds on $f(\mathcal{P}^{(k)}), f(\mathcal{P}_*^{(k)})$ for small $k$, see 7.1. Here $\mathcal{P}_*^{(k)}$ is the set of *squarefree* $k$-almost primes.

Figure 7.1: Bounds on $f(\mathcal{P}^{(k)}), f(\mathcal{P}_*^{(k)})$ from Bayless et al. [7]

| $k$ | $f(\mathcal{P}^{(k)})$ | $k$ | $f(\mathcal{P}_*^{(k)})$ |
|---|---|---|---|
| 2 | (1.1416, 1.1484) | 2 | (0.8877, 0.8945) |
| 3 | (0.65708, 1.0841) | 3 | (0.36003, 0.7678) |
| 4 | (0.40713, 1.1891) | 4 | (0.15118, 0.8527) |

Their approach is to directly compute the series up to $10^{12}$, and then obtain explicit inequalities for the counting functions of $\mathcal{P}^{(k)}, \mathcal{P}_*^{(k)}$. By partial summation, these translate into bounds for $f(\mathcal{P}^{(k)}), f(\mathcal{P}_*^{(k)})$. As evidenced by the table, this approach becomes exceedingly difficult as $k$ grows. This is in part due to the fact that the series for $f(\mathcal{P}^{(k)}), f(\mathcal{P}_*^{(k)})$ converge quite slowly. For example, as we shall see, the partial sum up to $10^{12}$ makes up less than half of $f(\mathcal{P}^{(4)})$.

Even in the case $k = 1$ for primes, the series for $f(\mathcal{P})$ converges slowly. Nevertheless, Cohen [27] was able to compute to a remarkable degree of precision,

$$f(\mathcal{P}) \;=\; 1.63661632335126086856965800392186367118159707613129\cdots. \tag{7.2}$$

His basic idea is to write

$$f(\mathcal{P}) = \sum_p \frac{1}{p \log p} = \int_1^\infty P(s)\, ds, \tag{7.3}$$

where $P(s) = \sum_p p^{-s}$ is the prime zeta function. Since $\log \zeta(s) = \sum_{m \geq 1} P(ms)/m$, by Möbius inversion one has the rapidly converging series $P(s) = \sum_{m \geq 1} \frac{\mu(m)}{m} \log \zeta(ms)$, and in turn one may use well-known rapid computation of $\zeta(s)$.

## 7.2 Statement of results

We generate the following data in Figure 7.2, refining Figure 7.1 from [7]. Following Cohen, we write $f(\mathcal{P}^{(k)})$ as an integral in terms of the corresponding zeta function, as in (7.3). From here, the key ingredient is a set of recurrence relations for such zeta functions, see Proposition 7.3.1. Also see Figure 7.2 below, and Appendix A for the relevant Mathematica code.

Figure 7.2: Computation of $f(\mathcal{P}^{(k)}), f(\mathcal{P}_*^{(k)})$ to 20 digits.

| $k$ | $f(\mathcal{P}^{(k)})$ | $k$ | $f(\mathcal{P}_*^{(k)})$ |
|---|---|---|---|
| 2 | 1.1448165734059179915 | 2 | 0.8909254794763183321 |
| 3 | 1.0308351017932175719 | 3 | 0.7131238005098902554 |
| 4 | 0.9973421485952523597 | 4 | 0.6528129098554062569 |
| 5 | 0.9888821921300755349 | 5 | 0.6284306642973934048 |
| 6 | 0.9887534530145096063 | 6 | 0.6176406880308143497 |
| 7 | 0.9910205950027380022 | 7 | 0.6126252367925047050 |
| 8 | 0.9935373386530404095 | 8 | 0.6102275665474058560 |
| 9 | 0.9956203792390954090 | 9 | 0.6090620642567092069 |
| 10 | 0.9971495172651382446 | 10 | 0.6084897027941833669 |

In particular, we observe that the sequence $\{f(\mathcal{P}^{(k)})\}_k$ decreases for $k \leq 6$ but then increases thereafter, in particular $f(\mathcal{P}^{(6)}) < f(\mathcal{P}^{(7)})$, contrary to the conjecture of Banks–Martin.

We also prove the following theorems, which confirm some of the trends observed in the data.

**Theorem 7.2.1.** *For all positive integers $k \neq 6$, we have $f(\mathcal{P}^{(6)}) < f(\mathcal{P}^{(k)})$.*

**Theorem 7.2.2.** *In the limit as $k \to \infty$, we have*

$$f(\mathcal{P}^{(k)}) \;=\; \sum_{\Omega(n)=k} \frac{1}{n \log n} \;\sim\; 1 \quad \textit{and} \quad f(\mathcal{P}_*^{(k)}) \;=\; \sum_{\Omega(n)=k} \frac{\mu(n)^2}{n \log n} \;\sim\; \frac{6}{\pi^2}. \tag{7.4}$$

We shall prove a quantitative form of Theorem 7.2.2 in Section 7.4, namely, $f(\mathcal{P}^{(k)}) = 1 + O_\epsilon(k^{\epsilon - 1/2})$ for any $\epsilon > 0$. The proof uses partial summation and a delicate application of the Sathe-Selberg theorem to $N_k(x) := \#\{n \leq x : \Omega(n) = k\}$ in the critical range $x \approx e^{e^k}$.

**Theorem 7.2.3** (Sathe-Selberg [130])**.** *For $k \leq 1.99 \log\log x$,*

$$N_{k+1}(x) \;=\; G\Big(\frac{k}{\log\log x}\Big) \frac{x}{\log x} \frac{(\log\log x)^k}{k!} \bigg(1 \;+\; O\Big(\frac{k}{(\log\log x)^2}\Big)\bigg),$$

*where $G(z) = \frac{1}{\Gamma(1+z)} \prod_p (1 - z/p)^{-1}(1 - 1/p)^z$.*

This approach is sufficiently versatile to generalize to arithmetic progressions. However, as we shall see throughout the chapter, the zeta function method offers a much more precise understanding of $f(\mathcal{P}^{(k)})$. By this method, we shall prove Theorem 7.2.1 as well as obtain an **exponentially decaying** upper bound of $1 + O(k\, 2^{-k/2})$, which significantly refines the results by classical methods. Further, we expect zeta functions may lead to lower bounds of comparable strength, though the problem here is more subtle.

In the next section, we describe the zeta function method and show how to generate the data in Figure 7.2 for small $k$.

## 7.3 Zeta function method for small $k$

Consider the prime and $k$-almost prime zeta functions

$$P(s) = \sum_p \frac{1}{p^s} \qquad P_k(s) = \sum_{\Omega(n)=k} \frac{1}{n^s} \qquad P_k^*(s) = \sum_{\Omega(n)=k} \frac{\mu(n)^2}{n^s}$$

for $k \geq 1$. Note $P_1 = P_1^* = P$, and let $P_0(s) = 1$. Our interest in these zeta functions arises from the identity

$$f(\mathcal{P}^{(k)}) = \sum_{\Omega(n)=k} \frac{1}{n \log n} = \sum_{\Omega(n)=k} \int_1^\infty \frac{ds}{n^s} = \int_1^\infty P_k(s)\, ds \tag{7.5}$$

Here $f(\mathcal{P}^{(k)})$ is finite, and so by Tonelli's theorem the interchange of sum and integral in (7.5) is justified because all terms are positive. Similarly we have $f(\mathcal{P}_*^{(k)}) = \int_1^\infty P_k^*(s)\,ds$.

In the following proposition, we express $P_k$ explicitly in terms of $P$ and derive a handy recursion formula.[1]

**Proposition 7.3.1.** *We have*

$$P_k(s) = \sum_{n_1+2n_2+\cdots+kn_k=k} \prod_{j=1}^{k} \frac{1}{n_j!}\big(P(js)/j\big)^{n_j} \tag{7.6}$$

*where the sum ranges over all partitions of $k$. $P_k$ satisfies the recurrence*

$$P_k(s) = \frac{1}{k}\sum_{j=1}^{k} P(js)P_{k-j}(s). \tag{7.7}$$

*Similarly, $P_k^*$ satisfies*

$$P_k^*(s) \;=\; (-1)^k \sum_{n_1+2n_2+\cdots=k} \prod_{j=1}^{k} \frac{1}{n_j!}\big(-P(js)/j\big)^{n_j} \;=\; \frac{1}{k}\sum_{j=1}^{k}(-1)^{j+1}P(js)P_{k-j}^*(s). \tag{7.8}$$

*Proof.* For (7.6), we have the following identity for $\zeta(s)$,

$$\sum_{k\ge 0} P_k(s) = \zeta(s) = \prod_p (1-p^{-s})^{-1} = \exp\Big(\sum_{j\ge 1}\frac{P(js)}{j}\Big).$$

Similarly, we have a formal power series identity in $z$,

$$\begin{aligned}\sum_{k\ge 0} P_k(s)z^k &= \prod_p \big(1-zp^{-s}\big)^{-1} = \exp\Big(\sum_{j\ge 1}\frac{P(js)}{j}z^j\Big) = \prod_{j\ge 1}\exp\Big(\frac{P(js)}{j}z^j\Big)\\ &= \prod_{j\ge 1}\sum_{n_j\ge 0}\frac{1}{n_j!}\Big(\frac{P(js)}{j}z^j\Big)^{n_j} = \sum_{k\ge 0} z^k \sum_{n_1+2n_2+\cdots=k}\prod_{j\ge 1}\frac{1}{n_j!}\big(P(js)/j\big)^{n_j}.\end{aligned} \tag{7.9}$$

Now (7.6) follows by comparing the coefficients of $z^k$.

To show (7.7), each term $n^{-s}$ appearing in $\sum_{j=1}^k P(js)P_{k-j}(s)$ has $\Omega(n) = k$. Conversely, for each $\Omega(n) = k$ write $n = \prod_p p^{e_p}$ with $\sum_p e_p = k$. Then for $1 \le j \le k$, the term $n^{-s}$ appears $\#\{p : e_p \ge j\}$ times in $P(js)P_{k-j}(s)$, thus appearing

$$\sum_{j=1}^{k} \#\{p : e_p \ge j\} = \sum_p e_p = k$$

[1]This was stated with incomplete proof in a preprint [92]. We provide a full proof by a simpler method.

times in $\sum_{j=1}^{k} P(js)P_{k-j}(s)$. Hence (7.7) follows.

We proceed similarly in the squarefree case. Letting $P_k^{\mu}(s) = \sum_{\Omega(n)=k} \mu(n)/n^s = (-1)^k P_k^*(s)$,

$$\sum_{k\geq 0} P_k^{\mu}(s) z^k = \prod_p (1 - zp^{-s}) = \exp\Big( -\sum_{j\geq 1} \frac{P(js)}{j} z^j \Big)$$
$$= \sum_{k\geq 0} z^k \sum_{n_1+2n_2+\cdots=k} \prod_{j\geq 1} \frac{\big(-P(js)/j\big)^{n_j}}{n_j!}$$

by expanding exp as before. Now the the partition formula in (7.8) follows by comparing the coefficients of $z^k$, and using $P_k^{\mu}(s) = (-1)^k P_k^*(s)$. Finally, as before, by counting the number of appearances of $\mu(n)n^{-s}$ on each side (even when $\mu(n) = 0$), we get

$$P_k^{\mu}(s) \;=\; \frac{1}{k} \sum_{j=1}^{k} -P(js) \cdot P_{k-j}^{\mu}(s).$$

Using $P_k^{\mu}(s) = (-1)^k P_k^*(s)$ this reduces to (7.8) as claimed. □

For instance, the first few $P_k$ are given by

$$\begin{aligned}
2! \cdot P_2(s) &= P(s)^2 + P(2s) \\
3! \cdot P_3(s) &= P(s)^3 + 3P(2s)P(s) + 2P(3s) \\
4! \cdot P_4(s) &= P(s)^4 + 6P(2s)P(s)^2 + 3P(2s)^2 + 8P(3s)P(s) + 6P(4s) \\
5! \cdot P_5(s) &= P(s)^5 + 10P(s)^3P(2s) + 15P(s)P(2s)^2 + 20P(s)^2P(3s) \\
&\quad + 20P(2s)P(3s) + 30P(s)P(4s) + 24P(5s).
\end{aligned}$$

With the above expressions for $P_k, P_k^*$ in terms of $P$, and using the built-in function PrimeZetaP, a few lines of code in Mathematica computes $f(\mathcal{P}^{(k)}), f(\mathcal{P}_*^{(k)})$ to high precision, generating the data in Figure 7.2. See Appendix A.

The computation was also independently verified to 20 digits on Pari/GP, courtesy of Paul Kinlaw. Pari/GP does not have $P(s)$ built-in, so it was computed using a variant of the identity $P(s) = \sum_{m\geq 1} \frac{\mu(m)}{m} \log \zeta(ms)$, namely,

$$P(s) \;=\; \sum_{p\leq A} p^{-s} \;+\; \sum_{m\geq 1} \frac{\mu(m)}{m} \log \zeta_A(ms) \tag{7.10}$$

for a suitable choice of $A$, where $\zeta_A(s) = \zeta(s) \prod_{p\leq A}(1 - p^{-s})$.

The data in Figure 7.2 suggest that $f(\mathcal{P}^{(k)})$ tends to 1 and $f(\mathcal{P}_*^{(k)})$ tends to $6/\pi^2 \approx .607\cdots$ as $k$ grows. With a bit more patience, we may calculate these differences for $k \le 20$.

Figure 7.3: Further computations of $1 - f(\mathcal{P}^{(k)})$ and $f(\mathcal{P}_*^{(k)}) - 6/\pi^2$

| $k$ | $1 - f(\mathcal{P}^{(k)})$ | $k$ | $f(\mathcal{P}_*^{(k)}) - 6/\pi^2$ |
|---|---|---|---|
| 10 | $2.85\cdots \times 10^{-3}$ | 10 | $5.62\cdots \times 10^{-4}$ |
| 11 | $1.80\cdots \times 10^{-3}$ | 11 | $2.79\cdots \times 10^{-4}$ |
| 12 | $1.11\cdots \times 10^{-3}$ | 12 | $1.39\cdots \times 10^{-4}$ |
| 13 | $6.74\cdots \times 10^{-4}$ | 13 | $6.95\cdots \times 10^{-5}$ |
| 14 | $4.02\cdots \times 10^{-4}$ | 14 | $3.46\cdots \times 10^{-5}$ |
| 15 | $2.37\cdots \times 10^{-4}$ | 15 | $1.73\cdots \times 10^{-5}$ |
| 16 | $1.38\cdots \times 10^{-4}$ | 16 | $8.65\cdots \times 10^{-6}$ |
| 17 | $7.96\cdots \times 10^{-5}$ | 17 | $4.32\cdots \times 10^{-6}$ |
| 18 | $4.55\cdots \times 10^{-5}$ | 18 | $2.16\cdots \times 10^{-6}$ |
| 19 | $2.58\cdots \times 10^{-5}$ | 19 | $1.08\cdots \times 10^{-6}$ |
| 20 | $1.45\cdots \times 10^{-5}$ | 20 | $5.40\cdots \times 10^{-7}$ |

## 7.4 Asymptotic behavior for large $k$

We confirm the limits that the data suggest with the following theorem.

**Theorem 7.4.1.** *For any $\epsilon > 0$,*

$$f(\mathcal{P}^{(k)}) \;=\; 1 \;+\; O\big(k^{-1/2+\epsilon}\big) \quad \textit{and} \quad f(\mathcal{P}_*^{(k)}) \;=\; \frac{6}{\pi^2} \;+\; O\big(k^{-1/2+\epsilon}\big). \tag{7.11}$$

*Proof.* Fix $k$ large and let $N_k(x) = \#\{n \le x : \Omega(n) = k\}$. First, by partial summation

$$\begin{aligned} f(\mathcal{P}^{(k+1)}) &= \sum_{\Omega(n)=k+1} f(n) = -\int_{2^{k+1}}^{\infty} N_{k+1}(t) f'(t)\, dt \\ &= \int_{2^{k+1}}^{\infty} \frac{N_{k+1}(t)}{t^2 \log t}\Big(1 + \frac{1}{\log t}\Big)\, dt =: I. \end{aligned}$$

The Sathe–Selberg theorem [130] implies that for $r = 1.99$ and $k \le r \log\log x$ (i.e., $x \ge e^{e^{k/r}}$)

$$N_{k+1}(x) \;=\; G\Big(\frac{k}{\log\log x}\Big)\frac{x}{\log x}\frac{(\log\log x)^k}{k!}\bigg(1 \;+\; O\Big(\frac{k}{(\log\log x)^2}\Big)\bigg), \tag{7.12}$$

where $G(z) = \frac{1}{\Gamma(1+z)}\prod_p (1-z/p)^{-1}(1-1/p)^z$. As such, we split up the integral $I = I_1 + I_2$ at $t = e^{e^{k/r}}$. For $I_1$, we use the universal bound of Erdős-Sárközy [45],

$$N_{k+1}(t) \ \ll \ \frac{k^4}{2^k}\, t\log t \qquad \text{for all } t,k \geq 1. \tag{7.13}$$

Using $r = 1.99$, we have $1/r - \log 2 < -1/6$ so that

$$I_1 := \int_{2^{k+1}}^{e^{e^{k/r}}} \frac{N_{k+1}(t)}{t^2\log t}\, dt \ \ll \ \frac{k^4}{2^k}\int_{2^{k+1}}^{e^{e^{k/r}}} \frac{dt}{t} \ < \ \frac{k^4}{2^k}e^{k/r} \ \ll \ e^{-k/6}. \tag{7.14}$$

For the bulk of the integral, $I_2$, we apply Sathe–Selberg to get

$$\begin{aligned} I_2 := \int_{e^{e^{k/r}}}^{\infty} \frac{N_{k+1}(t)}{t^2\log t}\, dt &= \frac{1}{k!}\int_{e^{e^{k/r}}}^{\infty} G\Big(\frac{k}{\log\log x}\Big)\frac{(\log\log t)^k}{t(\log t)^2}\bigg(1 \ + \ O\Big(\frac{k}{(\log\log t)^2}\Big)\bigg)\, dt \\ &= \frac{1}{k!}\int_{k/r}^{\infty} G(k/y)y^k e^{-y}\big(1+O(k/y^2)\big)\, dy. \end{aligned}$$

Since $G(z) \ll 1$ for $z \leq r$, the error in $I_2$ is bounded by

$$\frac{k}{k!}\int_{k/r}^{\infty} G(k/y)y^{k-2}e^{-y}\, dy \ll \frac{\Gamma(k-1)}{(k-1)!} = \frac{1}{k-1}$$

so that

$$I_2 \ = \ \frac{1}{k!}\int_{k/r}^{\infty} G(k/y)y^k e^{-y}\, dy \ + \ O(1/k). \tag{7.15}$$

The bulk of $I_2$ lies in the range $|y-k| < k^\delta$, for $\delta = 1/2+\epsilon$, so we split accordingly $I_2 = J + J'$. Note $y^k e^{-k}$ is increasing for $y$ up to $k$, and decreasing thereafter. So letting $z = y-k$,

$$\frac{y^k e^{-y}}{k!} < \exp(k\log y - k\log k - z) = \exp\big(k\log(1+z/k) - z\big) < e^{-z^2/2k} \tag{7.16}$$

using the lower bound $k! > k^k e^{-k}$. Thus

$$\frac{y^k e^{-y}}{k!} < e^{-\frac{1}{2}k^{2\delta-1}} \qquad \text{for} \quad |y-k| > k^\delta. \tag{7.17}$$

and so, using $G(z) \ll 1$ again, the integral $J'$ is bounded by

$$J' := \frac{1}{k!}\bigg(\int_{k/r}^{k-k^\delta} + \int_{k+k^\delta}^{\infty}\bigg)G(k/y)y^k e^{-y}\, dy \ \ll \ e^{-k^\epsilon}. \tag{7.18}$$

For $|y-k| < k^\delta$, we have $G(k/y) = 1 \ + \ O\big(k^{\delta-1}\big)$ [106, p. 236] so that

$$\begin{aligned} J := \frac{1}{k!}\int_{k-k^\delta}^{k+k^\delta} G(k/y)y^k e^{-y}\, dy &= \frac{1+O(k^{\delta-1})}{k!}\int_{k-k^\delta}^{k+k^\delta} y^k e^{-y}\, dy \\ &= \ \frac{1+O(k^{\delta-1})}{k!}\Gamma(k+1) \ = \ 1+O(k^{\delta-1}). \end{aligned} \tag{7.19}$$

Combining altogether, we obtain $f(\mathcal{P}^{(k)}) = I_1 + J + J' = 1 + O(k^{-1/2+\epsilon})$.

Similarly, the squarefree version of the Sathe–Selberg theorem [106, p. 237, ex. 4] states

$$N^*_{k+1}(x) \;=\; G^*\Big(\frac{k}{\log\log x}\Big)\frac{x}{\log x}\frac{(\log\log x)^k}{k!}\bigg(1 \;+\; O\Big(\frac{k}{(\log\log x)^2}\Big)\bigg) \tag{7.20}$$

where $G^*(z) = \frac{1}{\Gamma(1+z)}\prod_p(1+z/p)(1-1/p)^z$. And since $G^*(1) = \prod_p(1-p^{-2}) = 6/\pi^2$, by an analogous argument we obtain $f(\mathcal{P}^{(k)}_*) = 6/\pi^2 + O\big(k^{-1/2+\epsilon}\big)$. □

Further, by similar arguments one may show for any choice of integers $e_p \geq 0$ for each prime $p$,

$$\sum_{\substack{\Omega(n)=k\\ p^{e_p}\nmid n\ \forall e_p>0}}\frac{1}{n\log n} \;\sim\; \prod_{e_p>0}\big(1-p^{-e_p}\big). \tag{7.21}$$

In particular, we deduce that the evens and odds asymptotically contribute equally $1/2$ to $f(\mathcal{P}^{(k)}) \sim 1$. Whereas in the squarefree case, the evens and odds contribute $2/\pi^2, 4/\pi^2$, respectively, to $f(\mathcal{P}^{(k)}_*) \sim 6/\pi^2$.

Generalizing in another direction, we may consider the contribution to $f(\mathcal{P}^{(k)})$ from an arbitrary arithmetic progression.

**Corollary 7.4.1.1.** *For any fixed integers $0 \leq a < q$, as $k \to \infty$*

$$\sum_{\substack{\Omega(n)=k\\ n\equiv a\ (q)}}\frac{1}{n\log n} \;=\; \frac{1+o_q(1)}{q}. \tag{7.22}$$

*Proof.* Consider the counting function $N_k(x;q,a) := \#\{n \leq x : \Omega(n) = k, n \equiv a\ (q)\}$. First, if $(a,q) = 1$, then Theorem 2 in Spiro [133] gives[2]

$$N_{k+1}(x;q,a) \;=\; \frac{1}{\phi(q)}G_q\Big(\frac{k}{\log\log x}\Big)\frac{x}{\log x}\frac{(\log\log x)^k}{k!}\bigg(1 \;+\; O_q\Big(\frac{k}{(\log\log x)^2}\Big)\bigg) \tag{7.23}$$

for $k \leq 1.99\log\log x$, where

$$G_q(z) := \frac{(\phi(q)/q)^z}{\Gamma(1+z)}\prod_{p\nmid q}(1-z/p)^{-1}(1-1/p)^z = G(z)\prod_{p\mid q}(1-z/p).$$

And since $G_q(1) = \phi(q)/q$, by a similar argument to Theorem 7.4.1 we obtain (7.22) (with error term $O_q(k^{-1/2+\epsilon})$).

[2]Equation (7.23) may also be derived from earlier work of Delange [32].

When $d = (a, q) > 1$, letting $i = \Omega(d)$, by the above argument we have

$$\sum_{\substack{\Omega(m)=k-i \\ m \equiv a/d \ (q/d)}} \frac{1}{m \log m} \sim \frac{1}{q/d}.$$

Thus (7.22) follows by setting $n = md$ and noting $\log md \sim \log m$. □

The fact that every progression $a \pmod q$ contributes $1/q$ to $f(\mathcal{P}^{(k)}) \sim 1$ is especially remarkable in view of (7.23), since $N_k(x; q, a) \sim N_k(x)/q$ is not always true. However, it does indeed hold in the critical range of $\log\log x \sim k$.

One may also obtain an analogous result in the squarefree case. Namely, under the same conditions as Corollary 7.4.1.1, if $d = (a, q)$ is squarefree then

$$\sum_{\substack{\Omega(n)=k \\ n \equiv a \ (q)}} \frac{\mu(n)^2}{n \log n} \ \sim \ \frac{1}{q} \prod_{p \nmid d} (1 - p^{-2}) \prod_{p \mid d} (1 - p^{-1}) = \frac{6/\pi^2}{q} \frac{d}{\sigma(d)}. \tag{7.24}$$

In the next section, we turn to the question of optimal error bounds in Theorem 7.4.1.

## 7.5 Further progress via zeta functions for large $k$

As with the results of Bayless et al. [7], Theorem 7.4.1 is proven by using partial summation and knowledge of the counting function for $\mathcal{P}^{(k)}$. And given the success of the zeta function approach for small $k$, one might hope that the approach would yield results that beat the stated bound of $O(k^{-1/2+\epsilon})$. Inspecting in Figure 7.3, the ratio of consecutive entries of $1 - f(\mathcal{P}^{(k)})$ appears to converge to $1/2$, suggesting that the true error term in Theorem 7.4.1 may be $O(2^{-k})$. Such a bound remains out of reach for the moment, but we obtain related partial progress in this direction, culminating with the **exponentially-decaying** upper bound $f(\mathcal{P}^{(k)}) \le 1 + O(k\, 2^{-k/2})$.

**Theorem 7.5.1.** *We have*

$$\frac{1}{k!} \int_1^\infty [\log \zeta(s)]^k \, ds \ = \ 1 \ + \ O(2^{-k}). \tag{7.25}$$

*Proof.* Fix $k$ sufficiently large. The result will follow from the following three claims,

$$\int_1^2 \log(\tfrac{1}{s-1})^k \, ds \ = \ k!, \qquad 0 < \int_2^\infty [\log \zeta(s)]^k \, ds \ < \ 2^{-k}, \tag{7.26}$$

and

$$0 \;<\; J_k \;\ll\; 2^{-k} \qquad \text{where} \quad J_k := \frac{1}{k!}\int_1^2 [\log\zeta(s)]^k - \log(\tfrac{1}{s-1})^k \, ds. \tag{7.27}$$

First, we have

$$\int_1^2 \log(\tfrac{1}{s-1})^k \, ds = \int_0^1 (-\log s)^k \, ds = \int_0^\infty u^k e^{-u} \, du = k!\,. \tag{7.28}$$

Second, we note $0 < \log\zeta(s) < 2^{1-s}$ for $s \geq 2$, so

$$\int_2^\infty \log\zeta(s)^k \, ds < \int_1^\infty 2^{-ks} \, ds = \frac{2^{-k}}{k\log 2} < 2^{-k}. \tag{7.29}$$

Third, the series expansion of $\log\zeta$ at $s = 1$ is

$$\log\zeta(s) = \log\big(\tfrac{1}{s-1}\big) + \gamma(s-1) + O\big((s-1)^2\big).$$

It will suffice to use $0 < \log\zeta(s) + \log(s-1) < \frac{3}{5}(s-1)$ for $s \in [1,2]$. Expanding the binomial gives

$$[\log\zeta(s)]^k < \Big(\tfrac{3}{5}(s-1) - \log(s-1)\Big)^k = \sum_{j=0}^k \binom{k}{j} [\tfrac{3}{5}(s-1)]^j [-\log(s-1)]^{k-j}$$

so that

$$J_k := \frac{1}{k!}\int_1^2 [\log\zeta(s)]^k - [-\log(s-1)]^k \, ds < \sum_{j=1}^k \frac{(3/5)^j}{j!(k-j)!} \int_0^1 s^j [-\log s]^{k-j} \, ds. \tag{7.30}$$

Note we have

$$\int_0^1 s^j [-\log s]^i \, ds = \int_0^\infty u^i \, e^{-(j+1)u} \, du = \frac{i!}{(j+1)^{i+1}}. \tag{7.31}$$

Hence using $i = k - j$, we bound $J_k$ by the hypergeometric series

$$J_k < \sum_{j=1}^k \frac{(3/5)^j}{j!}(j+1)^{-(k-j+1)} \leq \tfrac{3}{5}\cdot 2^{-k}\Big(1 + \sum_{j=2}^k (3/5)^{j-1}\Big) \ll 2^{-k}, \tag{7.32}$$

since the ratio of consecutive terms is $\frac{3}{5}\big(\frac{j+1}{j+2}\big)^{k-j} \leq \frac{3}{5}$ for $j \geq 2$. □

Recall the prime zeta function $P(s) = \sum_p p^{-s}$. Consider

$$h(s) := \log\zeta(s) - P(s) = \sum_{m\geq 2}\sum_p \frac{1}{mp^{ms}} \;=\; h(1) + h'(1)(s-1) + O\big((s-1)^2\big)$$

where the Taylor series coefficients are given by

$$c := h(1) = \sum_{m\ge 2}\sum_{p}\frac{1}{mp^m} = .315718\cdots \qquad \text{and}$$
$$h'(1) = -\sum_{m\ge 2}\sum_{p}\frac{\log p}{p^{ms}}\Big|_{s=1} = -\sum_{p}\frac{\log p}{p(p-1)} = -.7505\cdots .$$

Hence from the expansion $P(s)+h(s) = \log\zeta(s) = \log(\frac{1}{s-1}) + \gamma(s-1) + O(s-1)^2$, we obtain

$$P(s) = \log\big(\tfrac{1}{s-1}\big) - c + (\gamma - h'(1))(s-1) + O\big((s-1)^2\big). \tag{7.33}$$

In particular, we have $0 < P(s) - \log\big(\frac{\alpha}{s-1}\big) < 1.4(s-1)$ for $s\in[1,2]$, where

$$\alpha := e^{-c} = \exp\Big(-\sum_{m\ge 2}\frac{P(m)}{m}\Big) = \prod_{m\ge 2}\zeta(m)^{\mu(m)/m} = .729264\cdots . \tag{7.34}$$

Proceeding as in Theorem 7.5.1, we obtain

**Theorem 7.5.2.** *For $\alpha$ in* (7.34)*, we have*

$$\frac{1}{k!}\int_1^\infty P(s)^k\,ds \;=\; \alpha \;+\; O(2^{-k}). \tag{7.35}$$

*Proof.* The proof follows from the following three claims,

$$\int_1^2 \log(\tfrac{\alpha}{s-1})^k\,ds \;=\; \alpha\,k!, \qquad 0 < \int_2^\infty P(s)^k\,ds < 2^{-k}, \tag{7.36}$$

and

$$0 < I_k \;\ll\; 2^{-k} \qquad \text{for} \quad I_k := \frac{1}{k!}\int_1^2 P(s)^k - \log(\tfrac{\alpha}{s-1})^k\,ds. \tag{7.37}$$

The first two hold, as in Theorem 7.5.1, recalling $P(s) < \log\zeta(s)$. For the third, as in Theorem 7.5.1, using $0 < P(s) - \log\big(\frac{\alpha}{s-1}\big) < \frac{7}{5}(s-1)$ and expanding the binomial gives

$$I_k < \sum_{j=1}^{k}\frac{(7/5)^j}{j!}(j+1)^{-(k-j+1)} \;< \tfrac{7}{5}\cdot 2^{-k} + \frac{(7/5)^2}{2!}3^{1-k}k \;\ll\; 2^{-k}.$$

Here we used the fact that the terms in the series decrease for $j$ up to around $k/\log k$ and increase thereafter. Thus the series is bounded by the first term, plus $k$ times the max of the terms $j = 2, k$. □

These results culminate with the following exponentially-decaying upper bound for $f(\mathcal{P}^{(k)})$.

**Theorem 7.5.3.** *We have*

$$f(\mathcal{P}^{(k)}) = \int_1^\infty P_k(s)\,ds \;\le\; 1 \;+\; O(k\,2^{-k/2}) \tag{7.38}$$

*Proof.* We first apply the combinatorial identity (7.6),

$$f(\mathcal{P}^{(k)}) = \int_1^\infty P_k(s)\,ds \le \sum_{n_1+2n_2+\cdots=k} \prod_{j\ge2} \frac{\big(P(j)/j\big)^{n_j}}{n_j!} \int_1^\infty \frac{P(s)^{n_1}}{n_1!}\,ds$$

using $P(js) \le P(j)$ for $j \ge 2, s \ge 1$. Then by Theorem 7.5.2,

$$f(\mathcal{P}^{(k)}) \;\le \sum_{n_1+2n_2+\cdots=k} \prod_{j\ge2} \frac{\big(P(j)/j\big)^{n_j}}{n_j!}\big(\alpha + O(2^{-n_1})\big) = \sum_{n_1=0}^{k} Q(k-n_1)\big(\alpha + O(2^{-n_1})\big)$$

where $Q(m) := \sum_{2n_2+3n_3+\cdots=m} \prod_{j\ge2} \frac{1}{n_j!}\big(P(j)/j\big)^{n_j}$. Then by definition of $\alpha, c$ from (7.34),

$$\frac{1}{\alpha} = e^c = \exp\Big(\sum_{j\ge2} P(j)/j\Big) = \sum_{m\ge0} \sum_{2n_2+3n_3+\cdots=m} \prod_{j\ge2} \frac{\big(P(j)/j\big)^{n_j}}{n_j!} = \sum_{m\ge0} Q(m) \tag{7.39}$$

By Lemma 7.5.4 below, we have $\sum_{m\ge y} Q(m) \ll y\,2^{-y}$. Hence using $y = k/2$,

$$\begin{aligned} f(\mathcal{P}^{(k)}) \le \sum_{m=0}^{k} Q(m)\big(\alpha + O(2^{m-k})\big) &\le \big(\alpha + O(2^{y-k})\big) \sum_{0\le m\le y} Q(m) + O(y\,2^{-y}) \\ &\le 1 \;+\; O(k\,2^{-k/2}) \end{aligned}$$

by (7.39). This completes the proof. □

**Lemma 7.5.4.** *We have*

$$Q(k) = \sum_{2n_2\cdots=m} \prod_{j\ge2} \frac{1}{n_j!}\big(P(j)/j\big)^{n_j} \;\ll\; k\,2^{-k}. \tag{7.40}$$

*Proof.* First $P(2) < 2\cdot 2^{-2}$ implies $P(j) < 2\cdot 2^{-j}$ for $j \ge 2$, and so

$$Q(k) < \sum_{2n_2+3n_3\cdots=k} \prod_{j\ge2} \frac{1}{n_j!}\big(2\cdot 2^{-j}/j\big)^{n_j} =: \widetilde{Q}(k).$$

Then we have a power series identity

$$\begin{aligned}\sum_{k\geq 0}\widetilde{Q}(k)z^k &= \sum_{k\geq 0}\sum_{2n_2+3n_3\cdots=k}\prod_{j\geq 2}\frac{1}{n_j!}\big(2\cdot(z/2)^j/j\big)^{n_j} = \prod_{j\geq 2}\sum_{n_j\geq 0}\frac{1}{n_j!}\big(2\cdot(z/2)^j/j\big)^{n_j}\\ &= \prod_{j\geq 2}\exp\Big(2\cdot(z/2)^j/j\Big) = \exp\Big(2\sum_{j\geq 2}\frac{(z/2)^j}{j}\Big) = \exp\Big(z-2\log(1-z/2)\Big)\\ &= e^z\cdot(1-z/2)^{-2} = \sum_{n\geq 0}\frac{z^n}{n!}\sum_{m\geq 0}(m+1)(z/2)^m.\end{aligned}$$

Hence equating coefficients of $z^k$, we deduce

$$\widetilde{Q}(k) = \sum_{n+m=k}\frac{1}{n!}\frac{m+1}{2^m} = 2^{-k}\sum_{n=0}^{k}(k-n+1)\frac{2^n}{n!} < k2^{-k}e^2 \ll k2^{-k}. \qquad \square$$

While a corresponding lower bound to $f(\mathcal{P}^{(k)})$ as in Theorem 7.5.3 remains out of reach, the zeta function method does produce strong enough lower bounds to show $f(\mathcal{P}^{(k)})$ is minimal when $k = 6$.

### 7.5.1 A global minimum via zeta functions

In view of Proposition 7.3.1, the integral $(1/k!)\int_1^\infty P(s)^k\,ds$ in Theorem 7.5.2 is a lower bound for $f(\mathcal{P}^{(k)}) = \int_1^\infty P_k(s)\,ds$, and constitutes the first of the terms in the identity (7.6), one per partition of $k$. Note the terms of partitions built from small parts contribute the most. So by incorporating the terms for the partitions $k = 1\cdot(k-j)+j$ and $k = 1\cdot(k-j-2)+2+j$ for $j \leq 6$, we may obtain a sufficiently tight lower bound on $f(\mathcal{P}^{(k)})$ to conclude the following.

**Theorem 7.5.5.** *We have $f(\mathcal{P}^{(6)}) < f(\mathcal{P}^{(k)})$ for all positive integers $k \neq 6$.*

*Proof.* We have already verified the claim directly for $k \leq 20$, see Figures 7.2,7.3. Thus it suffices to assume $k > 20$. By Proposition 7.3.1,

$$\begin{aligned}f(\mathcal{P}^{(k)}) = \int_1^\infty P_k(s)\,ds \;>\; &\frac{1}{k!}\int_1^2 P(s)^k\,ds + \sum_{j=2}^{6}\frac{\int_1^2 P(s)^{k-j}P(js)\,ds}{j(k-j)!}\\ &+\;\frac{\int_1^2 P(s)^{k-4}P(2s)^2\,ds}{2!2^2(k-4)!} + \sum_{j=3}^{6}\frac{\int_1^2 P(s)^{k-j-2}P(2s)P(js)\,ds}{2j(k-j-2)!}.\end{aligned} \tag{7.41}$$

For every $k \geq 1$, we have

$$\int_1^\infty P(s)^k\,ds > \int_0^1 \log\left(\tfrac{\alpha}{s}\right)^k ds = \alpha\,k!\,. \tag{7.42}$$

Using the first order Taylor series $P(js) > P(j) + P'(j)(s-1)$ for $j \geq 2$,

$$\begin{aligned}\int_1^2 P(s)^{k-j}P(js)\,ds \;&>\; P(j)\int_0^1 \log\left(\tfrac{\alpha}{s}\right)^{k-j} ds + P'(j)\int_0^1 \log\left(\tfrac{\alpha}{s}\right)^{k-j} s\,ds \\ &= \alpha(k-j)!\Big(P(j) + \frac{\alpha P'(j)}{2^{k-j}}\Big)\end{aligned}$$

and

$$\begin{aligned}\int_1^2 P(s)^{k-j-2}P(2s)P(js)\,ds \;&>\; (P'(2)P(j) + P(2)P'(j))\int_0^1 \log\left(\tfrac{\alpha}{s}\right)^{k-j} s\,ds \\ &+\; P(2)P(j)\int_0^1 \log\left(\tfrac{\alpha}{s}\right)^{k-j-2} ds \;+\; P'(2)P'(j)\int_0^1 \log\left(\tfrac{\alpha}{s}\right)^{k-j-2} s^2\,ds \\ &=\; \alpha(k-j-2)!\Big(P(2)P(j) + \frac{\alpha^2 P'(2)P'(j)}{3^{k-j-1}} + \\ &\qquad + \frac{\alpha}{2^{k-j-1}}[P'(2)P(j) + P(2)P'(j)]\Big).\end{aligned}$$

Hence plugging back into (7.41),

$$\begin{aligned}f(\mathcal{P}^{(k)}) > \alpha\Bigg[1 + \sum_{j=2}^{6}\frac{P(j) + \alpha P'(j)/2^{k-j}}{j} \;+\; \frac{1}{8}\Big(P(2)^2 + \frac{\alpha P(2)P'(2)}{2^{k-4}} + \frac{\alpha^2 P'(2)^2}{3^{k-3}}\Big) \\ +\; \sum_{j=3}^{6}\frac{1}{2j}\left(P(2)P(j) + \frac{\alpha}{2^{k-j-1}}[P'(2)P(j) + P(2)P'(j)] + \frac{\alpha^2 P'(2)P'(j)}{3^{k-j-1}}\right)\Bigg] =: \beta_k.\end{aligned} \tag{7.43}$$

Finally, since the lower bound $\beta_k$ is clearly increasing in $k$, for $k > 20$ we obtain[3]

$$f(\mathcal{P}^{(k)}) > \beta_k > \beta_{20} > 0.99 > f(\mathcal{P}^{(6)}). \qquad \square$$

### 7.5.2 Remark on integration

The integral $f(\mathcal{P}^{(k)}) = \int_1^\infty P_k(s)\,ds$ is computed numerically with high confidence. Nevertheless, in this section, we supplement the data by rigorously bounding the integral $f(\mathcal{P}^{(k)}) = \int_1^\infty P_k(s)\,ds$ at the tail and near the singularity $s = 1$. For $k \geq 2$, the tail $s \geq 10$ contributes

$$\int_{10}^\infty P_k(s)\,ds \leq 2\int_{10}^\infty 2^{-ks}\,ds = \frac{2^{1-10k}}{k\log 2} < 2^{-10k}$$

Next for $\epsilon > 0$,

$$\int_1^{1+\epsilon} \log(\tfrac{1}{s-1})^k\,ds = \int_{\log(1/\epsilon)}^\infty u^k e^{-u}\,du$$

[3]In Mathematica, we compute $\beta_{20} = 0.991049\cdots$.

And since $P_k(s) \le \log(\frac{1}{s-1})^k$ for $s \in (1, 1+e^{-2k})$ by Lemma 7.5.6 below, letting $\epsilon_k = e^{-4k}$, we bound

$$\int_1^{1+\epsilon_k} P_k(s)\, ds < \frac{1}{k!}\int_{4k}^{\infty} u^k e^{-u}\, du < \int_{4k}^{\infty} e^{-(u-k)^2/2k}\, du < \int_{3k}^{\infty} e^{-3ku/2}\, du \ll e^{-4k^2}$$

by (7.16), which rapidly decays.

**Lemma 7.5.6.** *For all $k \ge 1$, $s \in (1, 1+e^{-2k})$, we have*

$$P_k(s) \;\le\; \frac{1}{k!}\log(\tfrac{1}{s-1})^k. \tag{7.44}$$

*Proof.* We proceed by induction on $k$. For $k=1$, $P(s) \le \log(\frac{1}{s-1})$ holds for all $s > 1$. Now assuming the claim holds for all $j < k$, the recursion (7.7) gives

$$P_k(s) = \frac{1}{k}\sum_{j=1}^{k} P(js)P_{k-j}(s) \le \frac{1}{k}\sum_{j=1}^{k} \log\left(\tfrac{1}{js-1}\right)\frac{1}{(k-j)!}\log\left(\tfrac{1}{s-1}\right)^{k-j} \le \frac{1}{k!}\log\left(\tfrac{1}{s-1}\right)^k$$

which completes the induction, provided $\log\left(\frac{1}{js-1}\right) \le \frac{(k-j)!}{k!}\log\left(\frac{1}{s-1}\right)^j$. This is equality for $j=1$, and for $j \ge 2$ it suffices to prove

$$\log\left(\frac{1}{(j+1)s-1}\right) \le \frac{1}{k-j}\log\left(\tfrac{1}{s-1}\right)\log\left(\frac{1}{js-1}\right) \qquad 2 \le j < k, s \in (1, 1+e^{-2k}). \tag{7.45}$$

Letting $z = js-1$, we seek $s$ so that $\frac{1}{k-j}\log\left(\frac{1}{s-1}\right)$ is at least

$$\frac{\log(z+s)}{\log z} = 1 + \log(1+\tfrac{s}{z}) \le 1 + \frac{s}{z} = 1 + \frac{1}{j-1/s} \le 2$$

using $j \ge 2, s \ge 1$. Hence (7.45) holds since $s - 1 < e^{-2k}$. □

# Chapter 8

# On Erdős sums of almost primes

## 8.1 Introduction

In this chapter, we denote in shorthand, for $k \geq 1$,

$$f_k = \sum_{\Omega(n)=k} \frac{1}{n \log n}.$$

Here $f_k$ is the *Erdős sum* of $k$-almost primes, also called the $k^{\text{th}}$ *fingerprint number*. In 1935 Erdős [40] showed $f_k = O(1)$ uniformly[1] and in 1993 Zhang [143] proved the primes have maximal Erdős sum, that is, $f_k \leq f_1$ for all $k$. This had given initial evidence towards the Erdős primitive set conjecture, which we address later in Chapter 12.

In 2013, Banks and Martin [6] conjectured a vast generalization of the Erdős primitive set conjecture (see [6] and Chapter 12 for details and further discussion). In particular, they conjectured that the sums $f_k$ decrease monotonically in $k$. Denote the sum $f_{k,y}$ restricting $f_k$ to integers without prime factors $\leq y$, that is,

$$f_{k,y} = \sum_{\substack{\Omega(n)=k \\ p|n \Rightarrow p>y}} \frac{1}{n \log n}.$$

Banks and Martin further conjectured that $f_{k,y}$ decrease monotonically for any $y \geq 1$.

We prove that their conjecture holds for $y \geq 2$ and $k$ sufficiently large. By contrast, we prove $f_k = f_{k,1}$ increases monotonically in $k$, sufficiently large.

**Theorem 8.1.1.** *Let $y \geq 2$. For $k$ sufficiently large, we have $f_{k-1} < f_k$ and $f_{k-1,y} > f_{k,y}$.*

When $y \geq 2$, we believe that $f_{k-1,y} > f_{k,y}$ should hold for all $k > 1$, in accordance with Banks–Martin [6]. When $y = 1$, we believe $f_{k-1} < f_k$ for all $k > 6$. These inequalities have been verified numerically up to $k \leq 20$ [86].

The classical Sathe–Selberg theorem gives asymptotics for the counting function of $k$-almost primes, and implies $f_k$ converges to 1 with square-root error. That is $f_k = 1 + O(k^{\varepsilon - 1/2})$, see [86, Theorem 4.1]. We give an exponential refinement of this result, which identifies the (negative) secondary term, of order $O(k^2/2^k)$.

**Theorem 8.1.2.** *We have*

$$f_k = 1 - \frac{\log 2}{4} 2^{-k} \big( dk^2 + O\big(k \log(k+1)\big)\big)$$

[1] Indeed, his result bounded Erdős sums uniformly over any primitive set.

*uniformly for $k \geq 1$, where*

$$d := \frac{1}{4} \prod_{p>2} \left(1 - \frac{2}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^{2} = 0.37869\cdots. \tag{8.1}$$

To motivate the proof of Theorem 8.1.2, we first handle the sifted Erdős sums $f_{k,y}$, whose (positive) secondary term is of order $O_y(2^{-k})$ when $y \geq 2$, and thereby converge more rapidly.

**Theorem 8.1.3.** *Let $y \geq 2$ be an integer. We have*

$$f_{k,y} = \prod_{p \leq y} \left(1 - \frac{1}{p}\right) + c_y d_y / 2^k + O_y(k^3/3^k)$$

*uniformly for $k \geq 1$, where*

$$c_y = \gamma + \sum_{p \leq y} \frac{\log p}{p-1} - \sum_{p>y} \frac{\log p}{(p-1)(p-2)}, \tag{8.2}$$

$$d_y = d \prod_{2<p\leq y} \left(1 - \frac{2}{p}\right), \tag{8.3}$$

*for $d$ as in* (8.1)*, and $\gamma = 0.5772\cdots$ is the Euler–Mascheroni constant.*

In particular, from Theorems 8.1.2 and 8.1.3 we see the even and odd terms in $f_k$ contribute $\frac{1}{2} - O(k^2/2^k)$ and $\frac{1}{2} + O(2^{-k})$, respectively. As will be shown in the proofs, this discrepancy in the size (and sign) of the secondary terms ultimately come from the different behavior of

$$\lim_{s \to 1^+} (s-1)^{-z} \sum_{p|n \Rightarrow p>y} \frac{z^{\Omega(n)}}{n^s}$$

when $y = 1$ and $y \geq 2$. Namely, the singularity closest to 0 when $y = 1$ is at $z = 2$ while for $y \geq 2$ it is more distant. This provides a clean answer to the hitherto unexplained numerical observations up to $k \leq 20$.

*Remark* 8.1.4. Our methods can also handle $\sum_{\omega(n)=k,\, p|n \Rightarrow p>y} 1/(n \log n)$, where $\omega(n) = \sum_{p|n} 1$. For this problem, the analysis does not have a discrepancy between $y = 1$ and $y \geq 2$.

*Remark* 8.1.5. Hankel contours over the complex plane are used in the Selberg–Delange method, concerning asymptotics for sums $\sum_{n \leq x} z^{\Omega(n)}$. Saddle point analysis is used in the Sathe–Selberg theorem, which extracts information on the counting function of $k$-almost primes from these sums. These devices are powerful but lead

to poor savings. By contrast, our proofs contain neither saddle point analysis nor Hankel contours.

Moreover, the Selberg–Delange method assumes a zero-free region for $\zeta(s)$, as well as a bound on $\log \zeta(s)$ to the left of $\Re(s) = 1$. By contrast, our proofs make no complex-analytic assumptions about $\zeta(s)$ whatsoever. We only use (very basic) singularity analysis. Specifically, we use the Taylor approximation $\zeta(t)(t-1) = 1 + \gamma(t-1) + O((t-1)^2)$ for real $t \in (1,2)$.

### 8.1.1 A sequence of integrals

We also provide an alternate, probability-theoretic argument, which gives (weaker) exponential error $f_k = 1 + O(k/2^{k/4})$, but which has potentially much wider applicability to other primitive sets beyond $k$-almost primes. We leave further development of this perspective to future work.

The basic idea is to first show an asymptotic relation $f_k \sim e^{\gamma} I_{\lfloor k/4 \rfloor}$ for a certain sequence of iterated integrals $I_k$. Specifically, let $I_0 = 1, I_1 = \log 2$, and for $k \geq 2$ let

$$I_k = \int_{[0,1]^k} \frac{\mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_k}{1 + x_1(1 + x_2(\cdots(1 + x_k)\cdots))}. \tag{8.4}$$

We prove that $I_k$ converges to $e^{-\gamma}$, which may be of independent interest. Here $\gamma = 0.5772\cdots$ is the Euler-Mascheroni constant.

**Theorem 8.1.6.** *We have $I_k = e^{-\gamma} + O(2^{-k})$.*

Our approach to Theorem 8.1.6 relies on a probabilistic reformulation of the problem, which turns out to be related to Dickman-type distributions [112]. Such distributions have appeared not only in arithmetic problems but also in various contexts such as logarithmic combinatorial structures [3, Theorem 4.6] and computer science [78], but we shall not assume any prior knowledge of it. Indeed, based on techniques of recursive distributional equations, we are able to provide a self-contained proof of a finer version of Theorem 8.1.6.

### 8.1.2 Proof of monotonicity results

Here we quickly deduce Theorem 8.1.1 assuming Theorems 8.1.2 and 8.1.3: Indeed, for large $k$ we have

$$\begin{aligned} f_k - f_{k-1} &= \frac{\log 2}{4} d\Big((k-1)^2/2^{k-1} - k^2/2^k\Big) \; + \; o(k^2/2^k) \\ &= \frac{\log 2}{4} dk^2/2^k + o(k^2/2^k) \; > \; 0. \end{aligned}$$

by Theorem 8.1.2. Similarly, by Theorem 8.1.3 we have

$$\begin{aligned} f_{k-1,y} - f_{k,y} &= c_y d_y \Big( (k-1)^2/2^{k-1} - k^2/2^k \Big) \ + \ o(k^2/2^k) \\ &= c_y d_y k^2/2^k + o(k^2/2^k) \ > \ 0. \end{aligned}$$

Here we used $c_y > 0$ for $y \geq 2$. Indeed, $c_y$ from (8.2) is clearly increasing in $y \geq 2$, so $c_y \geq c_2 > 0$ since

$$\sum_{p>2} \frac{\log p}{(p-1)(p-2)} < .9 < 1.2 < \log 2 + \gamma.$$

This completes the proof of Theorem 8.1.1.

### 8.1.3 Proof method: permutations

As a short illustration of the proof method, we consider a permutation analogue of the sums $f_k$. This result also may be of independent interest. Namely, define $f_{k,\pi}$ by

$$f_{k,\pi} := \sum_{m \geq 1} \frac{a_{m,k}}{m}$$

where $a_{m,k}$ is the probability that a permutation, chosen uniformly at random from $S_m$, has exactly $k$ disjoint cycles.

*Remark* 8.1.7. To see the analogy with the original sum, given a permutation $\pi \in S_m$ let $C(\pi)$ be the number of disjoint cycles in $\pi$, $d(\pi)$ be $m$ and $|\pi|$ be $|S_m| = m!$. The sum $f_{k,\pi}$ may be expressed as

$$f_{k,\pi} = \sum_{\substack{C(\pi)=k \\ \pi \in \cup_{m \geq 1} S_m}} \frac{1}{|\pi| d(\pi)}.$$

The weight $1/(|\pi| d(\pi))$ is analogous to $1/(n \log n)$.

Recall the Riemann zeta function $\zeta(s) = \sum_{n \geq 1} n^{-s}$.

**Proposition 8.1.8.** *For any $k \geq 1$ we have*

$$f_{k,\pi} = \zeta(k+1) = 1 + 2^{-k-1} + O(3^{-k}).$$

*Proof.* Consider the exponential generating function of $\sum_{\pi \in S_n} z^{C(\pi)}$:

$$F(u,z) := \sum_{m \geq 0} \frac{1}{m!} \left( \sum_{\pi \in S_m} z^{C(\pi)} \right) u^m = 1 + \sum_{k \geq 1} z^k \sum_{m \geq 1} a_{m,k} u^m.$$

The exponential formula for permutations shows

$$F(u,z) = \exp\left(\sum_{n\geq 1} z\frac{u^n}{n}\right) = \exp\left(-z\log(1-u)\right). \tag{8.5}$$

The Taylor series for $\exp(-z\log(1-u))$ in $z$ is

$$\exp\left(-z\log(1-u)\right) = \sum_{k\geq 0}\frac{z^k}{k!}\left(-\log(1-u)\right)^k,$$

so extracting the coefficient of $z^k$ in $F(u,z)$ yields, by (8.5),

$$\sum_{m\geq 1} a_{m,k}u^m = \frac{\left(-\log(1-u)\right)^k}{k!}. \tag{8.6}$$

Since $1/m = \int_0^1 u^{m-1}\,\mathrm{d}u$, we can integrate (8.6) over $u\in[0,1]$ to obtain

$$f_{k,\pi} = \sum_{m\geq 1}\frac{a_{m,k}}{m} = \int_0^1 \frac{\left(-\log(1-u)\right)^k}{k!}\frac{\mathrm{d}u}{u} = \int_0^\infty \frac{x^k}{\Gamma(k+1)}\frac{\mathrm{d}x}{e^x-1} = \zeta(k+1).$$

This completes the proof. Here in the third equality we performed the change of variables $u = 1-e^{-x}$, and in the fourth we applied Riemann's famous integral representation for $\zeta(s)$, with $s = k+1$,

$$\Gamma(s)\zeta(s) = \int_0^\infty \frac{x^{s-1}}{e^x-1}\mathrm{d}x.$$

□

*Remark* 8.1.9. In the same way we shall relate $f_k$ to the sequence of integrals $I_k$ in Theorem 8.1.6, one can relate $f_{k,\pi}$ to $I_k$ as well.

## 8.2 Sifted sums

### 8.2.1 Preparation

For any $y\geq 1$ let $y_1$ be the smallest prime greater than $y$. The function

$$F(s,z) := \sum_{n\geq 1}\frac{z^{\Omega(n)}}{n^s} = \prod_p\left(1-\frac{z}{p^s}\right)^{-1} \tag{8.7}$$

converges absolutely for $s>1$ and $|z|<2$. In this domain it has a power series representation in $z$ and it defines a smooth function in $s$.[2] To see this, observe that

[2]In this paper we will only need first and second derivatives with respect to $s$.

for $s \geq 1+\varepsilon$ and $|z| \leq 2^{-\varepsilon}$ the $p$th term is

$$\exp\left(\sum_{i\geq 1} \frac{z^i}{ip^{si}}\right) = \exp\left(O_\varepsilon(|z|p^{-s})\right)$$

and so the product is a uniform limit of power series in $z$ which are smooth in $s$. We define

$$G(s,z) := F(s,z)(s-1)^z$$

and explain how to extend the range of definition of $G$. We write it as

$$G(s,z) = F(s,z)\zeta(s)^{-z}(\zeta(s)(s-1))^z. \tag{8.8}$$

It is well known that $\lim_{s\to 1^+}\zeta(s)(s-1) = 1$ and that extending $\zeta(s)(s-1)$ to $s=1$ it is a smooth function in $s \geq 1$. Hence, $(\zeta(s)(s-1))^z = \exp(z\log(\zeta(s)(s-1))$ is smooth in $s \geq 1$ and has a power series representation in $z$ with infinite radius of convergence:

$$(\zeta(s)(s-1))^z = \sum_{i\geq 0}\frac{z^i}{i!}(\log(\zeta(s)(s-1)))^i.$$

We have the following Euler product for $F(s,z)\zeta(s)^{-z}$:

$$F(s,z)\zeta(s)^{-z} = \prod_p \left(1-\frac{z}{p^s}\right)^{-1}\left(1-\frac{1}{p^s}\right)^z. \tag{8.9}$$

For every prime $q$, we can use this product to define $F(s,z)\zeta(s)^{-z}$ as a product of the rational function $\prod_{p\leq q}(1-z/p^s)^{-1}$ (which has simple poles in $z$ at $z=p^s$ for every prime $p \leq q$, and is defined for $s \geq 1$) with a function that has a power series representation in $z$ with radius of convergence $q_1$ and is defined for $s \geq 1$. This is because if $|z| < q_1^{1-\varepsilon}$ and $p > q$ then the $p$th term in the right-hand side of (8.9) equals

$$\exp\left(\sum_{i=2}^{\infty}\frac{z^i - z}{ip^{si}}\right) = \exp O_\varepsilon\left((|z|^2+|z|)p^{-2s}\right),$$

and the product over $p > q$ converges absolutely and uniformly.

In summary, for every prime $q$, $\prod_{p\leq q}(1-z/p^s)G(s,z)$, as well as its derivatives in $s$, is smooth in $s \geq 1$ and has a power series expansion in $z$ with radius $q_1$. In complex analytic language, we proved that for $s \geq 1$ $G(s,z)$ has a meromorphic continuation to $z \in \mathbb{C}$ with simple poles at $z = p^s$ for every prime $p$. In particular, for $s=1$,

$$G(1,z) = \prod_p \left(1-\frac{z}{p}\right)^{-1}\left(1-\frac{1}{p}\right)^z.$$

We also introduce

$$F_y(s,z) := \sum_{\substack{n\geq 1\\ p|n\Rightarrow p>y}} \frac{z^{\Omega(n)}}{n^s} = \prod_{p>y}\Big(1-\frac{z}{p^s}\Big)^{-1} = \prod_{p\leq y}\Big(1-\frac{z}{p^s}\Big)\,F(s,z), \tag{8.10}$$

and

$$\begin{aligned} G_y(s,z) &:= \prod_{p\leq y}\Big(1-\frac{z}{p^s}\Big)\,G(s,z)\\ &= \prod_{p\leq y}\Big(1-\frac{1}{p^s}\Big)^z \prod_{p>y}\Big(1-\frac{z}{p^s}\Big)^{-1}\Big(1-\frac{1}{p^s}\Big)^z(\zeta(s)(s-1))^z. \end{aligned} \tag{8.11}$$

For any smooth function $H(s,z)$, we shall denote by $H^{(a,b)}$ the mixed partial derivative

$$H^{(a,b)} = \frac{\partial^{a+b}}{\partial s^a \partial z^b} H.$$

For every $j \geq 0$ and fixed $s \geq 1$, $G_y^{(j,0)}(s,z)$ has a meromorphic continuation to $\mathbb{C}$ with simple poles at $z = p^s$ for every prime $p > y$.

Taking the logarithmic derivative of (8.11) with respect to $s$ gives

$$\begin{aligned} \frac{G_y^{(1,0)}}{G_y}(s,z) = \frac{\partial}{\partial s}\big[\log G_y(s,z)\big] = \frac{\partial}{\partial s}\Big[\sum_{p\leq y} z\log\Big(1-\frac{1}{p^s}\Big)\\ + \sum_{p\geq y_1}\Big(z\log\Big(1-\frac{1}{p^s}\Big) - \log\Big(1-\frac{z}{p^s}\Big)\Big) + z\log\big(\zeta(s)(s-1)\big)\Big] \end{aligned}$$

so that

$$G_y^{(1,0)}(s,z) = zG_y(s,z)\left(\sum_{p\leq y}\frac{\log p}{p^s-1} - \sum_{p\geq y_1}\frac{(z-1)\log p}{(p^s-1)(p^s-z)} + (\log(\zeta(s)(s-1)))'\right). \tag{8.12}$$

*Remark* 8.2.1. These generating functions and their values have natural connections to related questions about $k$-almost primes [8,88]. In particular $G_2(1,2) = \frac{1}{4}\prod_{p>2}(1-\frac{2}{p})^{-1}(1-\frac{1}{p})^2 = d$ as in (8.1) equals $2\beta_2$, from the main term of [8, Theorem 1.2].

### 8.2.2 Proof of Theorem 8.1.3

Let $y \geq 1$ be an integer, and let $y_1$ be the smallest prime greater than $y$. We have the integral representation

$$f_{k,y} = \sum_{\substack{\Omega(n)=k\\ p|n\Rightarrow p>y}} \frac{1}{n\log n} = \int_1^\infty \sum_{\substack{\Omega(n)=k\\ p|n\Rightarrow p>y}} n^{-s}\mathrm{d}s. \tag{8.13}$$

The smallest number $n$ with $\Omega(n) = k$ and all prime factors greater than $y$ is $y_1^k$, so the contribution of $s \geq 2$ to (8.13) is

$$\int_2^\infty \sum_{\substack{\Omega(n)=k \\ p|n \Rightarrow p>y}} n^{-s} \mathrm{d}s = \sum_{\substack{\Omega(n)=k \\ p|n \Rightarrow p>y}} \frac{1}{n^2 \log n} \leq \sum_{n \geq y_1^k} \frac{1}{n^2 \log n} \ll \frac{1}{k} \int_{y_1^k - 1}^\infty \frac{dt}{t^2} \ll \frac{1}{k y_1^k}.$$

Thus we have

$$f_{k,y} = I_{k,y} + O\left(\frac{1}{k y_1^k}\right), \tag{8.14}$$

where $I_{k,y}$ is the corresponding integral over $s \in [1, 2]$, namely,

$$I_{k,y} := \int_1^2 \sum_{\substack{\Omega(n)=k \\ p|n \Rightarrow p>y}} n^{-s} \mathrm{d}s = \int_1^2 \frac{1}{k!} F_y^{(0,k)}(s, 0) \mathrm{d}s. \tag{8.15}$$

Here we term-wise differentiated $F_y(s, z)$ in (8.10) with respect to $z$, for $s > 1$.

Next from the Taylor series

$$(s-1)^{-z} = \exp\left(-z \log(s-1)\right) = \sum_{i \geq 0} \frac{\left(-\log(s-1)\right)^i}{i!} z^i, \tag{8.16}$$

we apply the product rule to $F_y(s, z) = (s-1)^{-z} G_y(s, z)$, giving

$$\frac{1}{k!} F_y^{(0,k)}(s, 0) = \sum_{i=0}^k \frac{\left(-\log(s-1)\right)^{k-i}}{(k-i)!} \frac{1}{i!} G_y^{(0,i)}(s, 0).$$

Thus (8.15) becomes

$$I_{k,y} = \sum_{i=0}^k \int_1^2 \frac{\left(-\log(s-1)\right)^{k-i}}{(k-i)!} \frac{1}{i!} G_y^{(0,i)}(s, 0) \mathrm{d}s. \tag{8.17}$$

Now we introduce a similar integral $I'_{k,y}$, given by evaluating $G^{(0,i)}(s, 0)$ in the integrand at $s = 1$, namely,

$$I'_{k,y} := \sum_{i=0}^k \frac{1}{i!} G_y^{(0,i)}(1, 0) \int_1^2 \frac{\left(-\log(s-1)\right)^{k-i}}{(k-i)!} \mathrm{d}s. \tag{8.18}$$

To handle $I'_{k,y}$, we substitute $s = 1 + e^{-t}$ and obtain, for any $j \geq 0$,

$$\int_1^2 \frac{\left(-\log(s-1)\right)^j}{j!} \mathrm{d}s = \int_0^\infty \frac{e^{-t} t^j}{j!} \mathrm{d}t = \frac{\Gamma(j+1)}{j!} = 1, \tag{8.19}$$

as we are evaluating the Gamma function at $j + 1$. Hence (8.18) simplifies as

$$I'_{k,y} = \sum_{i=0}^k \frac{1}{i!} G_y^{(0,i)}(1, 0). \tag{8.20}$$

In the upcoming subsections, we shall estimate $I_{k,y}$ by means of the following lemmas for $I'_{k,y}$ and $I_{k,y} - I'_{k,y}$.

**Lemma 8.2.2.** *Let $y \geq 1$ be an integer. We have $I'_{k,y} = G_y(1,1) + O_y(y_1^{-k})$.*

**Lemma 8.2.3.** *Let $y \geq 2$ be an integer. We have $I_{k,y} = I'_{k,y} + G_y^{(1,0)}(1,2)/2^{k+1} + O_y(k^3/3^k)$.*

*Proof of Theorem 8.1.3 assuming Lemmas 8.2.2 and 8.2.3.* Recalling (8.14), we have

$$\begin{aligned} f_{k,y} &= I_{k,y} + O(1/ky_1^k) \\ &= I'_{k,y} + G_y^{(1,0)}(1,2)/2^{k+1} + O_y(k^3/3^k) \\ &= G_y(1,1) + G_y^{(1,0)}(1,2)/2^{k+1} + O_y(k^3/3^k). \end{aligned} \tag{8.21}$$

To compute the constants above, we first note $G_y(1,1) = \prod_{p \leq y}(1-1/p)$. Next for $y \geq 2$, by (8.11) and (8.12) with $(s,z) = (1,2)$ we have

$$\begin{aligned} G_y(1,2) &= \prod_{p \leq y} \left(1 - \frac{1}{p}\right)^2 \prod_{p \geq y_1} \left(1 - \frac{2}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^2 = d_y, \\ G_y^{(1,0)}(1,2) &= 2G_y(1,2) \left( \sum_{p \leq y} \frac{\log p}{p-1} - \sum_{p \geq y_1} \frac{\log p}{(p-1)(p-2)} + \gamma \right) = 2d_y c_y. \end{aligned} \tag{8.22}$$

Here we used $(\log(\zeta(s)(s-1)))'|_{s=1} = \gamma$. Hence plugging (8.22) and $G_y(1,1) = \prod_{p \leq y}(1-1/p)$ back into (8.21), we conclude

$$f_{k,y} = \prod_{p \leq y} \left(1 - \frac{1}{p}\right) + c_y d_y / 2^k + O_y(k^3/3^k). \tag{8.23}$$

□

### 8.2.3 Proof of Lemma 8.2.2

We use the notation $[z^n]A(z) = \frac{1}{n!}(\frac{d^n}{dz^n}A)(0)$ to denote the coefficient of $z^n$ in $A$, where $A$ is a function with Taylor series representation at $z = 0$. From the representation of $I'_{k,y}$ in (8.20), we have for all $k \geq 1$,

$$I'_{k,y} = [z^k] \frac{G_y(1,z)}{1-z} = \frac{1}{2\pi i} \int_{|z|=1/2} \frac{G_y(1,z)}{1-z} \frac{\mathrm{d}z}{z^{k+1}},$$

using Cauchy's integral formula. Here the integral ranges over a circle centered around $z = 0$, oriented counterclockwise, with radius $1/2$.

The function $G_y(1,z)$ has simple poles at $z = p$ for every prime $p > y$; these are its only poles. The rational function $1/(1-z)$ has a simple pole at $z = 1$. So recalling the smallest prime $y_1 > y$, the only poles of $G_y(1,z)/(1-z)$ in the range

$1/2 < |z| < y_1 + 1/2 =: R$ occur at $z = 1$ and $z = y_1$. Thus by Cauchy's residue theorem,

$$I'_{k,y} = \frac{1}{2\pi i}\int_{|z|=R} \frac{G_y(1,z)}{1-z}\frac{\mathrm{d}z}{z^{k+1}} + G_y(1,1) - \frac{\lim_{z\to y_1}(z-y_1)\,G_y(1,z)}{(1-y_1)y_1^{k+1}}. \tag{8.24}$$

Note $\lim_{z\to y_1} G_y(1,z)(z-y_1) \ll_y 1$. Then we claim that $|G_y(1,z)| \ll_y 1$ in the integrand of (8.24), from which we conclude

$$I'_{k,y} = \int_{|z|=R} \frac{O_y(1)}{R-1}\frac{\mathrm{d}z}{R^{k+1}} + G_y(1,1) - \frac{O_y(1)}{(1-y_1)y_1^{k+1}} \;=\; G_y(1,1) + O_y(y_1^{-k}). \tag{8.25}$$

Hence the proof follows from the claim $|G_y(1,z)| \ll_y 1$ for $|z| = R$.

To show this claim, note that if $p > 2|z| = 2R$ then

$$\left|\left(1-\frac{1}{p}\right)^z\left(1-\frac{z}{p}\right)^{-1}\right| = \left|\exp\Big(\sum_{i\geq 2}\frac{z^i - z}{ip^i}\Big)\right| \leq \exp\Big(\sum_{i\geq 2}\frac{|z|^i}{p^i}\Big) \leq \exp\Big(\frac{2R^2}{p^2}\Big).$$

Hence, recalling (8.11) with $s = 1$ we obtain

$$\begin{aligned}\max_{|z|=R}|G_y(1,z)| &= \max_{|z|=R}\left|\prod_{p\leq y}\left(1-\frac{1}{p}\right)^z\prod_{p\geq y_1}\left(1-\frac{1}{p}\right)^z\left(1-\frac{z}{p}\right)^{-1}\right| \\ &\leq \prod_{p\leq y}\left(1-\frac{1}{p}\right)^{-R}\prod_{p>2R}\exp\Big(\frac{2R^2}{p^2}\Big)\times \\ &\qquad \times \max_{|z|=R}\prod_{y_1\leq p\leq 2R}\left|\left(1-\frac{1}{p}\right)^z\left(1-\frac{z}{p}\right)^{-1}\right| \;\ll_y\; 1. \end{aligned}\tag{8.26}$$

This proves the claim, and hence Lemma 8.2.2 follows.

### 8.2.4 Proof of Lemma 8.2.3

By Taylor expanding at $s = 1$, we have uniformly for $s \in [1,2]$

$$G_y^{(0,i)}(s,0) = G_y^{(0,i)}(1,0) + (s-1)b_i + O\big((s-1)^2 c_i\big),$$

for coefficients

$$b_i := G_y^{(1,i)}(1,0) \qquad \text{and} \qquad c_i := \max_{s'\in[1,2]}\left|G_y^{(2,i)}(s',0)\right|. \tag{8.27}$$

Thus subtracting (8.17) from (8.18), we have

$$\begin{aligned}\left|I_{k,y} - I'_{k,y}\right| &= \left|\sum_{i=0}^{k}\int_1^2 \frac{(-\log(s-1))^{k-i}}{(k-i)!}\frac{1}{i!}\left(G_y^{(0,i)}(s,0) - G_y^{(0,i)}(1,0)\right)\mathrm{d}s\right| \\ &= \sum_{i=0}^{k}\int_0^1 \frac{(-\log s)^{k-i}}{(k-i)!i!}\left(sb_i + O(s^2c_i)\right)\mathrm{d}s.\end{aligned}\tag{8.28}$$

Substituting $s = e^{-t}$ shows that

$$\sum_{i=0}^{k} \frac{b_i}{i!(k-i)!} \int_0^1 (-\log s)^{k-i} s \mathrm{d}s = \sum_{i=0}^{k} \frac{b_i}{i!(k-i)!} \int_0^\infty t^{k-i} e^{-2t} \mathrm{d}t$$
$$= \sum_{i=0}^{k} \frac{b_i 2^{i-k-1}}{i!(k-i)!} \int_0^\infty v^{k-i} e^{-v} \mathrm{d}v \;=\; \sum_{i=0}^{k} \frac{b_i}{i!} 2^{i-k-1} \tag{8.29}$$

and similarly

$$\sum_{i=0}^{k} \frac{c_i}{i!(k-i)!} \int_0^1 (-\log s)^{k-i} s^2 \mathrm{d}s = \sum_{i=0}^{k} \frac{c_i}{i!} 3^{i-k-1}. \tag{8.30}$$

Plugging (8.29) and (8.30) back into (8.28) gives

$$I_{k,y} - I'_{k,y} = \sum_{i=0}^{k} \Big( \frac{b_i}{i!} 2^{i-k-1} + O\Big( \frac{c_i}{i!} 3^{i-k-1} \Big) \Big). \tag{8.31}$$

So proceeding as in the proof of Lemma 8.2.2 (as in (8.24)), by Cauchy's integral formula and residue theorem,

$$\sum_{i=0}^{k} \frac{b_i}{i!} 2^{i-k-1} = 2^{-1}[z^k] \frac{G_y^{(1,0)}(1,z)}{1-\frac{z}{2}} = \frac{1}{4\pi i} \int_{|z|=1/2} \frac{G_y^{(1,0)}(1,z)}{1-\frac{z}{2}} \frac{\mathrm{d}z}{z^{k+1}}$$
$$= \frac{1}{2\pi i} \int_{|z|=R} \frac{G_y^{(1,0)}(1,z)}{2-z} \frac{\mathrm{d}z}{z^{k+1}} + G_y^{(1,0)}(1,2) 2^{-k-1}$$
$$- \frac{\lim_{z \to y_1} (z-y_1)\, G_y^{(1,0)}(1,z)}{(2-y_1) y_1^{k+1}} = G_y^{(1,0)}(1,2) 2^{-k-1} + O_y(y_1^{-k}). \tag{8.32}$$

Finally, we claim $c_i \ll_y i!(i+1)^2/y_1^i$, in which case $\sum_{i=0}^k \frac{c_i}{i!} 3^{i-k-1} \ll \sum_{i=0}^k y_1^{-i}(i+1)^2 3^{i-k-1} \ll \sum_{i=0}^k (i+1)^2\, 3^{-k} \ll_y k^3/3^k$ since $y_1 \geq 3$ (as $y \geq 2$). Thus combined with (8.32), we conclude

$$I_{k,y} - I'_{k,y} = G_y^{(1,0)}(1,2)/2^{k+1} + O_y(k^3/3^k). \tag{8.33}$$

Hence to complete the proof of Lemma 8.2.3, it suffices to show $c_i \ll_y i!(i+1)^2/y_1^i$, which by definition means that uniformly for $s \in [1,2]$,

$$[z^i] G_y^{(2,0)}(s,z) \ll_y (i+1)^2/y_1^i. \tag{8.34}$$

To this end, recall $G_y^{(1,0)}(s,z) = z G_y(s,z) c(s)$ by (8.12), where

$$c(s) = c_y(s,z) := \sum_{p \leq y} \frac{\log p}{p^s - 1} - \sum_{p \geq y_1} \frac{(z-1)\log p}{(p^s-1)(p^s - z)} + (\log(\zeta(s)(s-1)))'.$$

So differentiating again

$$\begin{aligned} G_y^{(2,0)}(s,z) &= zG_y^{(1,0)}(s,z)c(s) + zG_y(s,z)c'(s) \\ &= G_y(s,z)\big(z^2c(s)^2 + zc'(s)\big), \end{aligned} \tag{8.35}$$

noting

$$c'(s) = \sum_{p\geq y_1} \frac{(z-1)(\log p)^2(2p^{2s}-(z+1)p^s)}{(p^s-1)^2(p^s-z)^2} - \sum_{p\leq y} \frac{p^s(\log p)^2}{(p^s-1)^2} + \big(\log(\zeta(s)(s-1))\big)''.$$

For fixed $s \in [1,2]$, note $c$ and $c'$ are meromorphic in $\mathbb{C}$ whose only poles are $z = p^s$ for each $p \geq y_1$. Thus

$$\begin{aligned} G_y(s,z) &= \left(1-\frac{z}{y_1^s}\right)^{-1} H_{y,1}(s,z), \\ z^2c(s)^2 + zc'(s) &= \left(1-\frac{z}{y_1^s}\right)^{-2} H_{y,2}(s,z), \end{aligned}$$

for functions $H_{y,1}, H_{y,2}$, whose smallest pole is at $z = y_2^s$ where $y_2$ is the smallest prime larger than $y_1$.

Letting $H_y := H_{y,1}H_{y,2}$, we see (8.35) becomes

$$G_y^{(2,0)}(s,z) = \left(1-\frac{z}{y_1^s}\right)^{-3} H_y(s,z). \tag{8.36}$$

Note $H_y$ has no poles inside $|z| \leq y_2^{s-\varepsilon}$ so by compactness $\max_{s\in[1,2]} |H_y(s,z)| \ll_{y,\varepsilon} 1$ uniformly for $|z| \leq y_2^{s-\varepsilon}$. Thus by Cauchy's integral formula,

$$[z^i]H_y(s,z) = \frac{1}{2\pi i}\int_{|z|=y_2^{s-\varepsilon}} \frac{H_y(s,z)}{z^{i+1}}\mathrm{d}z \ll y_2^{-i(s-\varepsilon)} \max_{|z|=y_2^{s-\varepsilon}} |H_y(s,z)| \ll_{y,\varepsilon} y_2^{-i(1-\varepsilon)}$$

uniformly for $s \in [1,2]$. By the binomial theorem,

$$[z^i]\left(1-\frac{z}{y_1^s}\right)^{-3} = y_1^{-is}\binom{i+2}{2} \ll (i+1)^2/y_1^i$$

uniformly for $s \in [1,2]$ and $i \geq 0$. Hence by the product rule, from (8.36) we conclude

$$[z^i]G_y^{(2,0)}(s,z) = \sum_{i_1+i_2=i} [z^{i_1}]\left(1-\frac{z}{y_1^s}\right)^{-3} [z^{i_2}]H_y(s,z) \ll_y (i+1)^2/y_1^i.$$

This gives (8.34) as desired, which completes the proof.

*Remark* 8.2.4. By a similar proof as of (8.34) above, for any $y \geq 2$, $m \geq 0$,

$$[z^i]G_y^{(m,0)}(s,z) \ll_{m,y} (i+1)^m/y_1^i \tag{8.37}$$

uniformly for $s \in [1,2]$ and $i \geq 0$.

## 8.3 Proof of Theorem 8.1.2

Recall $f_k = f_{k,1}$, $F = F_1$ and $G = G_1$. By (8.14) with $y = 1$, we have

$$f_k = I_k + O\left(\frac{1}{k2^k}\right)$$

where

$$I_k = \int_1^2 \frac{1}{k!} F^{(0,k)}(s,0)\mathrm{d}s.$$

We apply the product rule to

$$F(s,z) = (s-1)^{-z}G(s,z) = (s-1)^{-z}\left(1-\frac{z}{2^s}\right)^{-1} G_2(s,z),$$

giving

$$\frac{1}{k!}F^{(0,k)}(s,0) = \sum_{i+j+l=k} \frac{(-\log(s-1))^l}{l!} 2^{-js} \frac{1}{i!} G_2^{(0,i)}(s,0),$$

using the power series in (8.16), as well as $\left(1-\frac{z}{2^s}\right)^{-1} = \sum_{j\geq 0} 2^{-js}z^j$. Thus we have

$$I_k = \sum_{i+j+l=k} \int_1^2 \frac{(-\log(s-1))^l}{l!} 2^{-js} \frac{1}{i!} G_2^{(0,i)}(s,0)\mathrm{d}s. \tag{8.38}$$

Now we introduce a similar integral $I_k'$, given by evaluating $G_2^{(0,i)}(s,0)$ in the integrand at $s = 1$, namely,

$$I_k' = \sum_{i+j+l=k} \int_1^2 \frac{(-\log(s-1))^l}{l!} 2^{-js} \frac{1}{i!} G_2^{(0,i)}(1,0)\mathrm{d}s. \tag{8.39}$$

Hence to establish Theorem 8.1.2, it suffices to prove the following lemmas for $I_k'$ and $I_k - I_k'$.

**Lemma 8.3.1.** *We have* $I_k' = 1 - \frac{\log 2}{4} 2^{-k}\big(dk^2 + O(k\log(k+1))\big)$.

**Lemma 8.3.2.** *We have* $I_k = I_k' + O(k/2^k)$.

These are the $y = 1$ analogues of Lemmas 8.2.2 and 8.2.3.

### 8.3.1 Proof of Lemma 8.3.2

By the mean value theorem we have, uniformly for $s \in [1,2]$,

$$\left|G_2^{(0,i)}(s,0) - G_2^{(0,i)}(1,0)\right| \leq (s-1)b_i$$

for coefficients

$$b_i := \max_{s' \in [1,2]} \left|G_2^{(1,i)}(s',0)\right|.$$

Thus subtracting (8.38) from (8.39), we find

$$\begin{aligned} \left|I_k - I_k'\right| &= \left| \sum_{i+j+l=k} \int_1^2 \frac{(-\log(s-1))^l}{l!} 2^{-js} \frac{1}{i!} \left(G_2^{(0,i)}(s,0) - G_2^{(0,i)}(1,0)\right) \mathrm{d}s \right| \\ &\leq \sum_{i+j+l=k} \int_1^2 \frac{(-\log(s-1))^l}{l!} 2^{-js} (s-1) \frac{b_i}{i!} \mathrm{d}s. \end{aligned} \tag{8.40}$$

By (8.37), we have uniformly for $t \in [1,2]$,

$$\frac{1}{i!} G_2^{(1,i)}(t,0) = [z^i] G_2^{(1,0)}(t,z) \ll (i+1)3^{-i}. \tag{8.41}$$

Hence $\frac{b_i}{i!} \ll (i+1)3^{-i}$, so that (8.40) becomes

$$\begin{aligned} I_k - I_k' &\ll \sum_{i+j+l=k} \int_0^\infty \frac{t^l}{l!} 2^{-j(1+e^{-t})} e^{-2t} (i+1) 3^{-i} \mathrm{d}t \\ &\leq \sum_{i+j+l=k} (i+1) 3^{-i} 2^{-j} \int_0^\infty \frac{t^l}{l!} e^{-2t} \mathrm{d}t \\ &= \sum_{i+j+l=k} (i+1) 3^{-i} 2^{-j-l} \int_0^\infty \frac{u^l}{l!} e^{-u} \mathrm{d}u \;=\; \sum_{i+j+l=k} (i+1) 3^{-i} 2^{-j-l} \end{aligned}$$

where the last equality follows from substituting $u/2$ for $t$ and recalling the integral form (8.19) of the Gamma function. Hence we conclude

$$I_k - I_k' \ll \sum_{i \leq k} (i+1) 3^{-i} 2^{i-k} \sum_{j \leq k-i} 1 \leq k 2^{-k} \sum_{i \leq k} (i+1)(2/3)^i \ll k 2^{-k}.$$

### 8.3.2 Proof of Lemma 8.3.1

Recall

$$I_k' = \sum_{i=0}^k \frac{1}{i!} G_2^{(0,i)}(1,0) \sum_{j+l=k-i} \int_1^2 \frac{(-\log(s-1))^l}{l!} 2^{-js} \mathrm{d}s.$$

We will prove in the next subsection that

**Lemma 8.3.3.** *For $k \geq 1$, we have*

$$\sum_{j+l=k} \int_1^2 \frac{(-\log(s-1))^l}{l!} 2^{-sj} \mathrm{d}s = 2 - \frac{\log 2}{4} 2^{-k} \Big( k^2 + O\big((k+1)\log(k+2)\big) \Big).$$

*Remark* 8.3.4. The saving is $k/\log k$ and it appears sharp. We find it to be an unusual saving.

Using Lemma 8.3.3 we simplify the inner sum in (8.39) and find

$$\begin{aligned} I'_k &= \sum_{i=0}^{k} \frac{1}{i!} G_2^{(0,i)}(1,0) \bigg( 2 - \frac{\log 2}{4} 2^{-k} \Big( (k-i)^2 + O\big((k+1-i)\log(k+2-i)\big) \Big) \bigg) \\ &= 2I'_{k,2} - \frac{\log 2}{4} 2^{-k} B \end{aligned} \tag{8.42}$$

where

$$B = \sum_{i=0}^{k} \frac{2^i}{i!} \Big( (k-i)^2 + O\big((k-i+1)\log(k-i+2)\big) \Big) G_2^{(0,i)}(1,0). \tag{8.43}$$

Lemma 8.2.2 with $y = 2$ yields

$$I'_{k,2} := \sum_{i=0}^{k} \frac{1}{i!} G_2^{(0,i)}(1,0) = G_2(1,1) + O(3^{-k}) = \tfrac{1}{2} + O(3^{-k}).$$

Similarly as in the proof of Lemma 8.2.2, Cauchy's integral formula and residue theorem imply

$$\begin{aligned} B' &:= \sum_{i=0}^{k} \frac{2^i}{i!} G_2^{(0,i)}(1,0) = [z^k] \frac{G_2(1,2z)}{1-z} \\ &= \frac{1}{2\pi i} \int_{|z|=1/4} \frac{G_2(1,2z)}{1-z} \frac{\mathrm{d}z}{z^{k+1}} \\ &= \frac{1}{2\pi i} \int_{|z|=2} \frac{G_2(1,2z)}{1-z} \frac{\mathrm{d}z}{z^{k+1}} + G_2(1,2) - \frac{\lim_{z \to 3/2}(z-3/2)\, G_y(1,2z)}{(1-3/2)(3/2)^{k+1}} \\ &= d + O((3/2)^{-k}). \end{aligned} \tag{8.44}$$

Here $G_2(1,2) = \frac{1}{4} \prod_{p>2} (1-\frac{2}{p})^{-1} (1-\frac{1}{p})^2 = d$. We also note $G_2(1,z)$ is meromorphic in $|z| < 4$ with simple pole at $z = 3$ (so $z = 3/2$ is the smallest pole of $G_2(1,2z)$). In particular,

$$\frac{1}{i!} G_2^{(0,i)}(1,0) = [z^i] G_2(1,z) \ll 3^{-i}. \tag{8.45}$$

Thus combining (8.43), (8.44) and (8.45), we obtain

$$\begin{aligned} B - (d + O(2/3)^k)k^2 &= B - k^2 B' \\ &= \sum_{i=0}^{k} \frac{2^i}{i!} G_2^{(0,i)}(1,0)((k-i)^2 + O((k-i+1)\log(k-i+2)) - k^2) \\ &\ll \sum_{i=0}^{k} (2/3)^i (ik + O(k\log(k+1))) \ll k\log(k+1). \end{aligned}$$

Hence $B = dk^2 + O(k\log(k+1))$, so plugging back into (8.42) we conclude

$$I'_k = 1 - \frac{\log 2}{4} 2^{-k} \big( dk^2 + O(k\log(k+1)) \big). \tag{8.46}$$

### 8.3.3 Proof of Lemma 8.3.3

We may suppose $k \geq 2$. Multiplying through by $2^k$, we aim to prove

$$A_k = 2^{k+1} - \frac{\log 2}{4} k^2 + O(k\log k), \tag{8.47}$$

for

$$A_k := 2^k \sum_{i+j=k} \int_1^2 \frac{(-\log(s-1))^i}{i!} 2^{-js} \mathrm{d}s = \sum_{i=0}^{k} \frac{2^i}{i!} J(i)$$

where, substituting $s = 1 + e^{-u}$,

$$J(i) := \int_1^2 (-\log(s-1))^i \, 2^{-j(s-1)} \mathrm{d}s = \int_0^\infty u^i e^{-u} 2^{-(k-i)e^{-u}} \mathrm{d}u.$$

Note the trivial bound $J(i) \leq i!$, using $2^{-(k-i)e^{-u}} \leq 1$.

In order to conclude (8.47), it suffices to prove the following two estimates, using $T = 15\log k$,

$$\sum_{0 \leq i \leq T} \frac{2^i}{i!} J(i) = \sum_{0 \leq i \leq T} 2^i + O(kT) \qquad \text{for} \quad k \geq T \geq \log k, \tag{8.48}$$

$$\sum_{T < i \leq k} \frac{2^i}{i!} J(i) = 2^{k+1} - \frac{\log 2}{4} k^2 - \sum_{0 \leq i \leq T} 2^i + O(kT) \qquad \text{for} \quad k \geq T \geq 15\log k. \tag{8.49}$$

We first prove (8.48). For $T \geq \log k$, the contribution of $e^u \geq k$ to $J(i)$, i.e.

$u \geq \log k$, is handled by the Taylor approximation $2^{-(k-i)e^{-u}} = 1 - O(ke^{-u})$. Thus

$$\begin{aligned} J(i) &= \int_0^{\log k} u^i e^{-u} 2^{-(k-i)e^{-u}} \mathrm{d}u + \int_{\log k}^{\infty} u^i e^{-u}(1 + O(ke^{-u}))\mathrm{d}u \\ &= \int_0^{\log k} u^i e^{-u} 2^{-(k-i)e^{-u}} \mathrm{d}u + \int_{\log k}^{\infty} u^i e^{-u}\mathrm{d}u + O(k2^{-i}) \int_0^{\infty} v^i e^{-v}\mathrm{d}v \\ &= \int_0^{\log k} u^i e^{-u} 2^{-(k-i)e^{-u}} \mathrm{d}u + \int_0^{\infty} u^i e^{-u}\mathrm{d}u - \int_0^{\log k} u^i e^{-u}\mathrm{d}u + O(k2^{-i}i!) \\ &= i! + \int_0^{\log k} u^i e^{-u}(2^{-(k-i)e^{-u}} - 1)\mathrm{d}u + O(k2^{-i}i!). \end{aligned}$$

Summing over $i \leq T$, we obtain

$$\sum_{0 \leq i \leq T} \frac{2^i}{i!} J(i) = \sum_{0 \leq i \leq T} 2^i - \sum_{0 \leq i \leq T} \frac{2^i}{i!} \int_0^{\log k} u^i e^{-u}(1 - 2^{-(k-i)e^{-u}})\mathrm{d}u + O(kT). \tag{8.50}$$

Thus to conclude (8.48) from (8.50) above, it remains show

$$\sum_{0 \leq i \leq T} \frac{2^i}{i!} \int_0^{\log k} u^i e^{-u}(1 - 2^{-(k-i)e^{-u}})\mathrm{d}u = O(k \log k). \tag{8.51}$$

For $i \leq \log k$, we have

$$\sum_{0 \leq i \leq \log k} \frac{2^i}{i!} \int_0^{\infty} u^i e^{-u}\mathrm{d}u \leq \sum_{0 \leq i \leq \log k} 2^i = O(2^{\log k}) = O(k) \tag{8.52}$$

since $2 < e$. Next recall the function $u \mapsto u^i e^{-u}$ is increasing for $u \leq i$. So for $u \leq \log k < i$, the integrand is maximized at $u = \log k$, giving

$$\begin{aligned} \sum_{\log k < i \leq T} \frac{2^i}{i!} \int_0^{\log k} u^i e^{-u}\mathrm{d}u &\leq \sum_{\log k < i \leq T} \frac{2^i}{i!} \log k (\log k)^i e^{-\log k} \\ &\leq \frac{\log k}{k} \sum_{i=0}^{\infty} \frac{(2 \log k)^i}{i!} = k \log k. \end{aligned} \tag{8.53}$$

Combining (8.52) and (8.53) gives (8.51) as desired. This completes the proof of (8.48).

Now to prove (8.49), we begin by handling the contribution of small $u$ to the integral $J(i)$. Since the function $u \mapsto u^i e^{-u}$ increases for $u \leq i$, for fixed $a \in (0, 1)$, the contribution of $u \leq ai$ to $J(i)$ is at most

$$\int_0^{ai} u^i e^{-u} 2^{-(k-i)e^{-u}} \mathrm{d}u \leq \int_0^{ai} u^i e^{-u}\mathrm{d}u \leq (ai)(ai)^i e^{-ai} \ll i!\sqrt{i}(ea/e^a)^i$$

by Stirling's approximation. Thus when e.g. $a = 0.21$ is small enough to satisfy $e^a > 2ea$, we have

$$\int_0^{ai} u^i e^{-u} 2^{-(k-i)e^{-u}} \mathrm{d}u \ll_a i! 2^{-i}. \tag{8.54}$$

Now let $c = 0.07$ (so that $a = 3c$ will satisfy $e^a > 2ea$). For $T < i \le k$, if $u \ge ci$ we have

$$(k-i)e^{-u} \le ke^{-ci} \le ke^{-cT} \le k^{1-15c} \ll k^{-.05}$$

since $T \ge 15 \log k$. In particular, we may use the 2nd order Taylor approximation:

$$2^{-(k-i)e^{-u}} = 1 - \log 2 (k-i)e^{-u} + O(k^2 e^{-2u}).$$

Thus by (8.54) with $ci < ai$, we have for $T < i \le k$,

$$\begin{aligned} J(i) &= \int_{ci}^{\infty} u^i e^{-u} 2^{-(k-i)e^{-u}} \mathrm{d}u + O(i! 2^{-i}) \\ &= \int_{ci}^{\infty} u^i e^{-u} \left(1 - \log 2 \, (k-i)e^{-u} + O\left(k^2 e^{-2u}\right)\right) \mathrm{d}u + O(i! 2^{-i}) \\ &= A_{i,1} - \log 2 (k-i) A_{i,2} + O(k^2 A_{i,3}) + O(i! 2^{-i}) \end{aligned} \tag{8.55}$$

where, for $j = 1, 2, 3$,

$$A_{i,j} := \int_{ci}^{\infty} u^i e^{-ju} \mathrm{d}u = j^{-i-1} \int_{cij}^{\infty} v^i e^{-v} \mathrm{d}v = j^{-i-1} \, i! \, (1 + O(2^{-i})). \tag{8.56}$$

Here we used (8.54) with $cij \le 3ci = ai$. So plugging back into (8.55) and dividing through by $i!$,

$$\begin{aligned} \frac{J(i)}{i!} &= (1 + O(2^{-i})) - \log 2 (k-i) 2^{-i-1} (1 + O(2^{-i})) + O(k^2 3^{-i-1} + 2^{-i}) \\ &= 1 - \log 2 (k-i) 2^{-i-1} + O(k^2 3^{-i} + 2^{-i}). \end{aligned} \tag{8.57}$$

Summing over $i \in (T, k]$, we conclude

$$\begin{aligned} \sum_{T<i\le k} \frac{2^i}{i!} J(i) &= \sum_{T<i\le k} \left(2^i - \log 2 (k-i) 2^{-1} + O(k^2 (2/3)^i + 1)\right) \\ &= 2^{k+1} - \sum_{i \le T} 2^i - \frac{\log 2}{2} \sum_{T<i\le k} (k-i) \; + \; O(k^2 (2/3)^T + k) \\ &= 2^{k+1} - \sum_{i \le T} 2^i - \frac{\log 2}{2} \Big(\frac{k^2}{2} + O(Tk)\Big). \end{aligned}$$

Here $k^2 (2/3)^T \le k^{2 + 15 \log(2/3)} < k^{-4}$ since $T \ge 15 \log k$. This gives (8.49) as desired, and hence completes the proof of Lemma 8.3.3.

## 8.4 Probability-theoretic argument

In this section, we give an alternative probabilistic interpretation of Erdős sums, showing

**Proposition 8.4.1.** *We have* $f_k = 1 + O(k/2^{k/4})$.

In view of Theorem 8.1.2 we haven't tried to optimize the exponent $2^{k/4}$.

We first begin with some preliminaries.

### 8.4.1 Preliminary lemmas

For an integer $a \geq 1$, let $P_1(a)$ and $p(a)$ denote the largest and smallest prime factors of $a$, respectively (here $P_1(1) := 1$). Define the set of L-multiples $\mathrm{L}_a$,[3]

$$\mathrm{L}_a := \{ba \in \mathbb{N} : p \mid b \Rightarrow p \geq P_1(a)\}.$$

Note the density of this set is $\mathrm{d}(\mathrm{L}_a) = \frac{1}{a}\prod_{p<P_1(a)}(1-1/p)$.

**Lemma 8.4.2.** *For any* $a \in \mathbb{N}$*, we have*

$$\sum_{p \geq P_1(a)} \mathrm{d}(\mathrm{L}_{ap}) = \mathrm{d}(\mathrm{L}_a).$$

*Proof.* Let $y > 1$. Consider the set of integers without prime factors $p < y$, and partition it according to the smallest prime factor $q \geq y$. This gives the disjoint union,

$$\big\{b \in \mathbb{N} : p \mid b \Rightarrow p \geq y\big\} \;=\; \bigcup_{q \geq y} \big\{bq \in \mathbb{N} : p \mid b \Rightarrow p \geq q\big\}.$$

Taking the density of both sides,

$$\prod_{p<y}\Big(1-\frac{1}{p}\Big) = \sum_{q \geq y}\frac{1}{q}\prod_{p<q}\Big(1-\frac{1}{p}\Big). \tag{8.58}$$

Now choosing $y = P_1(a)$, we divide by $a$ to obtain $\mathrm{d}(\mathrm{L}_a) = \sum_{q \geq P_1(a)} \mathrm{d}(\mathrm{L}_{aq})$ as desired. □

From Lemma 8.4.2, a simple induction argument on $j \geq 1$ gives

$$\sum_{\substack{\Omega(b)=j \\ p(b) \geq P_1(a)}} \mathrm{d}(\mathrm{L}_{ab}) = \mathrm{d}(\mathrm{L}_a). \tag{8.59}$$

In particular when $a = 1$, for any $j \geq 1$ we have $\sum_{\Omega(b)=j} \mathrm{d}(\mathrm{L}_b) = 1$.

[3] L for lexicographic

**Lemma 8.4.3.** *Let $a \in \mathbb{Z}_{>1}$. We have uniformly for $0 < v < 1$,*

$$\sum_{q \geq P_1(a)^{\frac{1}{v}}} \mathrm{d}(\mathrm{L}_{aq}) = v\,\mathrm{d}(\mathrm{L}_a)\Big(1 + O\Big(\tfrac{1}{\log P_1(a)}\Big)\Big). \tag{8.60}$$

*Proof.* Take $0 < v < 1$. We first recall Mertens' product theorem states that

$$\prod_{p<x}\Big(1-\frac{1}{p}\Big) = \frac{e^{-\gamma}}{\log x}\Big(1+O\Big(\frac{1}{\log x}\Big)\Big).$$

In particular, for $x = P_1(a) \geq 2$,

$$\begin{aligned}\prod_{P_1(a)\leq p<P_1(a)^{\frac{1}{v}}}\Big(1-\frac{1}{p}\Big) &= \prod_{p<P_1(a)}\Big(1-\frac{1}{p}\Big)^{-1}\prod_{p<P_1(a)^{\frac{1}{v}}}\Big(1-\frac{1}{p}\Big)\\ &= \frac{\log P_1(a)}{\log P_1(a)^{\frac{1}{v}}}\Big(1+O\Big(\frac{1}{\log P_1(a)}\Big)\Big) \;=\; v\Big(1+O(1/\log P_1(a))\Big).\end{aligned} \tag{8.61}$$

So by (8.61) and the identity (8.58) with $y = P_1(n)^{\frac{1}{v}}$,

$$\sum_{q\geq P_1(a)^{\frac{1}{v}}}\frac{1}{q}\prod_{p<q}\Big(1-\frac{1}{p}\Big) = \prod_{p<P_1(a)^{\frac{1}{v}}}\Big(1-\frac{1}{p}\Big) = v\Big(1+O(1/\log P_1(a))\Big)\prod_{p<P_1(a)}\Big(1-\frac{1}{p}\Big).$$

Dividing through by $a$ completes the proof. □

**Lemma 8.4.4.** *For $k \geq 1$, let $c_1 \geq \cdots \geq c_k \geq 0$. If $d_1, D_1, E_1, \ldots, d_k, D_k, E_k \geq 0$ satisfy $E_i \leq \sum_{j=1}^{i} d_j \leq D_i$ for all $1 \leq i \leq k$ (and let $d_0 = E_0 = D_0 = 0$), then we have*

$$\sum_{i=1}^{k} c_i(E_i - E_{i-1}) \;\leq\; \sum_{i=1}^{k} c_i d_i \;\leq\; \sum_{i=1}^{k} c_i(D_i - D_{i-1}).$$

*Proof.* By rearranging sums,

$$\sum_{i=1}^{k} c_i d_i = \sum_{i=1}^{k} c_i\Big(\sum_{j=1}^{i} d_j - \sum_{j=0}^{i-1} d_j\Big) = \sum_{i=1}^{k-1}(c_i - c_{i+1})\sum_{j=1}^{i} d_j \;+\; c_k\sum_{i=1}^{k} d_i. \tag{8.62}$$

Since $c_i - c_{i+1} \geq 0$ and $\sum_{j\leq i} d_j \leq D_i$, from (8.62) we obtain

$$\sum_{i=1}^{k} c_i d_i \leq \sum_{i=1}^{k-1}(c_i - c_{i+1})D_i \;+\; c_k D_k = \sum_{i=1}^{k} c_i(D_i - D_{i-1}).$$

Similarly, since $\sum_{j\leq i} d_j \geq E_i \geq 0$, from (8.62) we obtain

$$\sum_{i=1}^{k} c_i d_i \geq \sum_{i=1}^{k-1}(c_i - c_{i+1})E_i \;+\; c_k E_k = \sum_{i=1}^{k} c_i(E_i - E_{i-1}).$$

□

To handle the contribution of smooth numbers, we use a simple bound of Erdős and Sárközy [45, Lemma 2], whose proof we provide for completeness.

**Lemma 8.4.5** (Erdős–Sárközy)**.** *For any $k \geq 1$, $y > 1$, we have*

$$\sum_{\substack{\Omega(n)=k \\ P_1(n)<e^y}} \frac{1}{n} \ll y^2\, k/2^k.$$

*Proof.* Observe that $2^k$ times our given sum is bounded by the following Euler product,

$$\begin{aligned}\sum_{\substack{\Omega(n)=k \\ P_1(n)<e^y}} \frac{2^k}{n} &\leq \prod_{p<e^y} \Big(1 + \frac{2}{p} + \cdots + \frac{2^k}{p^k}\Big) \\ &\leq (k+1) \prod_{2<p<e^y} \Big(1 - \frac{2}{p}\Big)^{-1} \ll ky^2\end{aligned}$$

by Mertens' product theorem. Dividing by $2^k$ completes the proof. □

Let $P_j(n)$ denote the $j$th largest prime of $n$, with multiplicity, so that $n = P_1(n) \cdots P_k(n)$.

**Corollary 8.4.5.1.** *For any $1 \leq j \leq k$ and $y > 1$, we have*

$$\sum_{\substack{\Omega(n)=k \\ P_{j+1}(n)<e^y}} \frac{1}{n \log n} \ll y^2\, k\, 2^{j-k}.$$

*Proof.* First, for each $n$ with $\Omega(n) = k$, factor $n = ab$ with $\Omega(b) = j$ and $p(b) \geq P_1(a)$. Thus by (8.59) with $j$, we have

$$\sum_{\substack{\Omega(n)=k \\ P_{j+1}(n)<e^y}} \mathrm{d}(\mathrm{L}_n) = \sum_{\substack{\Omega(a)=k-j \\ P_1(a)<e^y}} \sum_{\substack{\Omega(b)=j \\ p(b)\geq P_1(a)}} \mathrm{d}(\mathrm{L}_{ab}) = \sum_{\substack{\Omega(a)=k-j \\ P_1(a)<e^y}} \mathrm{d}(\mathrm{L}_a).$$

On each side of the above identity, we apply the simple bounds $1/(n \log n) \ll \mathrm{d}(\mathrm{L}_n)$ and $\mathrm{d}(\mathrm{L}_a) \ll 1/a$. This gives

$$\sum_{\substack{\Omega(n)=k \\ P_{j+1}(n)<e^y}} \frac{1}{n \log n} \ll \sum_{\substack{\Omega(a)=k-j \\ P_1(a)<e^y}} \frac{1}{a} \ll y^2\, k\, 2^{j-k}$$

by Lemma 8.4.5 with $k$ replaced by $k - j$. □

### 8.4.2 Poof of Proposition 8.4.1

Let $k \geq 1$ be sufficiently large. We shall choose $y = 2^j$ for $j = \lfloor k/4 \rfloor$, and $N = 4^k$. Let $f_k'$ denote the sum $f_k$ restricted by $P_{j+1}(n) \geq e^y$. Thus by Corollary 8.4.5.1,

$$f_k = \sum_{\Omega(n)=k} \frac{1}{n \log n} = f_k' + O(y^2 2^{j-k}) = f_k' + O(k/2^{k/4}) \tag{8.63}$$

where, by Mertens's product theorem,

$$f_k' := \sum_{\substack{\Omega(n)=k \\ P_{j+1}(n) \geq e^y}} \frac{1}{n \log n} = \sum_{\substack{\Omega(n)=k \\ P_{j+1}(n) \geq e^y}} \big(e^\gamma + \tfrac{O(1)}{\log P_1(n)}\big) \frac{\log P_1(n)}{\log n} \mathrm{d}(\mathrm{L}_n). \tag{8.64}$$

Next, we rewrite the identity $n = P_1(n) \cdots P_k(n)$ as

$$\frac{\log n}{\log P_1(n)} = 1 + \tfrac{\log P_2(n)}{\log P_1(n)} \Big(1 + \cdots \Big(1 + \tfrac{\log P_{j+1}(n)}{\log P_j(n)} \Big(1 + \tfrac{\log \left(P_{j+2}(n) \cdots P_k(n)\right)}{\log P_{j+1}(n)}\Big)\Big) \cdots \Big).$$

Taking the reciprocal of identity above gives

$$\frac{\log P_1(n)}{\log n} = u_{j+1}\Big(\tfrac{\log P_2(n)}{\log P_1(n)}, \ldots, \tfrac{\log P_{j+1}(n)}{\log P_j(n)}, \tfrac{\log(P_{j+2}(n) \cdots P_k(n))}{\log P_{j+1}(n)}\Big),$$

for the functions $u_k \colon \mathbb{R}^k \to \mathbb{R}$,

$$u_k(x_1, \ldots, x_k) := \frac{1}{1 + x_1(1 + x_2(\cdots(1 + x_k)\cdots))}. \tag{8.65}$$

In particular, from $P_1(n) \geq \ldots \geq P_k(n)$ we infer the inequalities

$$\frac{\log P_1(n)}{\log n} \leq u_j\Big(\tfrac{\log P_2(n)}{\log P_1(n)}, \ldots, \tfrac{\log P_{j+1}(n)}{\log P_j(n)}\Big), \tag{8.66}$$

$$\geq u_j\Big(\tfrac{\log P_2(n)}{\log P_1(n)}, \ldots, \tfrac{\log P_{j+1}(n)}{\log P_j(n)}(k-j)\Big). \tag{8.67}$$

By (8.66), we see (8.64) implies

$$f_k' \leq \sum_{\substack{\Omega(a)=k-j \\ P_1(a) \geq e^y}} \big(e^\gamma + \tfrac{O(1)}{\log P_1(a)}\big) \sum_{P_1(a) \leq p_j \leq \cdots \leq p_1} u_j\Big(\tfrac{\log p_2}{\log p_1}, \ldots, \tfrac{\log P_1(a)}{\log p_j}\Big) \mathrm{d}(\mathrm{L}_{a p_j \cdots p_1}). \tag{8.68}$$

**Lemma 8.4.6.** *For any $a \in \mathbb{Z}_{>1}$, there is an absolute constant $C > 1$ such that*

$$\sum_{P_1(a) \leq p_j \leq \cdots \leq p_1} u_j\Big(\tfrac{\log p_2}{\log p_1}, \ldots, \tfrac{\log P_1(a)}{\log p_j}\Big) \mathrm{d}(\mathrm{L}_{a p_j \cdots p_1})$$

$$\leq \frac{\big(1 + \frac{C}{\log P_1(a)}\big)^j}{N^j} \sum_{i_1, \ldots, i_j = 1}^{N} u_j\Big(\tfrac{i_1 - 1}{N}, \ldots, \tfrac{i_j - 1}{N}\Big) \mathrm{d}(\mathrm{L}_a).$$

*Proof.* For each $1 \le r \le j$, it suffices to show that

$$\begin{aligned} &N^{1-r} \sum_{i_1,\dots i_{r-1}=1}^{N} \sum_{P_1(a)\le p_j\le\cdots\le p_{r+1}\le p_r} u_j\Big(\tfrac{i_1-1}{N},\dots,\tfrac{i_{r-1}-1}{N},\tfrac{\log p_{r+1}}{\log p_r},\dots,\tfrac{\log P_1(a)}{\log p_j}\Big)\,\mathrm{d}(\mathrm{L}_{ap_j\cdots p_r}) \\ &\le \Big(1+\tfrac{C}{\log P_1(a)}\Big)N^{-r} \qquad (8.69) \\ &\quad\times \sum_{i_1,\dots,i_r=1}^{N} \sum_{P_1(a)\le p_j\le\cdots\le p_{r+1}} u_j\Big(\tfrac{i_1-1}{N},\dots,\tfrac{i_r-1}{N},\tfrac{\log p_{r+2}}{\log p_{r+1}},\dots,\tfrac{\log P_1(a)}{\log p_j}\Big)\,\mathrm{d}(\mathrm{L}_{ap_j\cdots p_{r+1}}). \end{aligned}$$

Indeed, iterating (8.69) (with each $r = 1, 2, \dots, j$ in turn) completes the proof of the lemma.

Now to show (8.69), fix indices $i_1, \dots i_{r-1} \le N$ and primes $p_j \le \cdots \le p_{r+1}$ ($p_j \ge P_1(a)$). Define $c_{i_r}$, $d_{i_r}$ by

$$\begin{aligned} c_{i_r} &:= u_j\Big(\tfrac{i_1-1}{N},\dots,\tfrac{i_r-1}{N},\tfrac{\log p_{r+2}}{\log p_{r+1}},\dots,\tfrac{\log P_1(a)}{\log p_j}\Big) \\ d_{i_r} &:= \sum_{p_r\in[p_{r+1}^{N/i_r},p_{r+1}^{N/(i_r-1)})} \mathrm{d}(\mathrm{L}_{ap_j\cdots p_r}). \end{aligned}$$

(For $i_r = 1$, the range of $p_r$ in $d_{i_r}$ is to be interpreted as $[p_{r+1}^N, \infty)$.) Note for any $u \le N$, by Lemma 8.4.3

$$\sum_{i_r=1}^{u} d_{i_r} = \sum_{p_r\ge p_{r+1}^{N/u}} \mathrm{d}(\mathrm{L}_{ap_j\cdots p_r}) \le \frac{u}{N}\mathrm{d}(\mathrm{L}_{ap_j\cdots p_{r+1}})\Big(1+\tfrac{C}{\log P_1(a)}\Big) =: D_u.$$

In particular $D_u - D_{u-1} = \frac{1}{N}\mathrm{d}(\mathrm{L}_{ap_j\cdots p_{r+1}})(1 + C/\log P_1(a))$. So splitting up the sum over $p_r$ and applying Lemma 8.4.4, we have

$$\begin{aligned} &\sum_{p_r\ge p_{r+1}} u_j\Big(\tfrac{i_1-1}{N},\dots,\tfrac{i_{r-1}-1}{N},\tfrac{\log p_{r+1}}{\log p_r},\dots,\tfrac{\log P_1(a)}{\log p_j}\Big)\,\mathrm{d}(\mathrm{L}_{ap_j\cdots p_r}) \\ &\qquad\le \sum_{i_r=1}^{N} c_{i_r} \sum_{p_r\in[p_{r+1}^{N/i_r},p_{r+1}^{N/(i_r-1)})} \mathrm{d}(\mathrm{L}_{ap_j\cdots p_r}) \\ &\qquad= \sum_{i_r=1}^{N} c_{i_r}d_{i_r} \le \sum_{i_r=1}^{N} c_{i_r}(D_{i_r} - D_{i_r-1}) \le \Big(1+\tfrac{C}{\log P_1(a)}\Big)\frac{1}{N}\sum_{i_r=1}^{N} c_{i_r}\mathrm{d}(\mathrm{L}_{ap_j\cdots p_{r+1}}) \\ &\qquad= \Big(1+\tfrac{C}{\log P_1(a)}\Big)\frac{1}{N}\sum_{i_r=1}^{N} u_j\Big(\tfrac{i_1-1}{N},\dots,\tfrac{i_r-1}{N},\tfrac{\log p_{r+2}}{\log p_{r+1}},\dots,\tfrac{\log P_1(a)}{\log p_j}\Big)\,\mathrm{d}(\mathrm{L}_{ap_j\cdots p_{r+1}}). \end{aligned} \qquad (8.70)$$

Summing (8.70) over $i_1, \dots i_r \le N$ and $p_j \le \cdots \le p_{r+1}$, we obtain (8.69) as desired. □

Plugging Lemma 8.4.6 into (8.68) we obtain

$$f_k' \le e^\gamma \sum_{\substack{\Omega(a)=k-j\\ P_1(a)\ge e^y}} \mathrm{d}(\mathrm{L}_a)\,\frac{\big(1+C/y\big)^{j+1}}{N^j} \sum_{i_1,\dots,i_j=1}^{N} u_j\Big(\tfrac{i_1-1}{N},\dots,\tfrac{i_j-1}{N}\Big). \tag{8.71}$$

Note by Corollary 8.4.5.1 and (8.59) with $a=1$,

$$\sum_{\substack{\Omega(a)=k-j\\ P_1(a)\ge e^y}} \mathrm{d}(\mathrm{L}_a) = \sum_{\Omega(a)=k-j} \mathrm{d}(\mathrm{L}_a) - O(y^2 2^{j-k}) = 1 - O(2^{-k/4}),$$

recalling the definitions $y=2^j$, $j=\lfloor k/4\rfloor$. Thus

$$f_k' \le e^\gamma\,\frac{\big(1+O(2^{-k/4})\big)^{j+1}}{N^j} \sum_{i_1,\dots,i_j=1}^{N} u_j\Big(\tfrac{i_1-1}{N},\dots,\tfrac{i_j-1}{N}\Big). \tag{8.72}$$

By an analogous argument (using the lower bound in (8.67), and settin $E_i = \frac{i}{N}\,\mathrm{d}(\mathrm{L}_{p_2\cdots p_j a})(1-C/\log P_1(a))$ in Lemma 8.4.4 instead of $D_i$), we may obtain a similar lower bound

$$f_k' \ge e^\gamma \frac{\big(1-O(2^{-k/4})\big)^{j+2}}{N^j} \sum_{i_1,\dots,i_j=1}^{N} u_j\Big(\tfrac{i_1}{N},\dots,\tfrac{i_j}{N}(k-j)\Big). \tag{8.73}$$

Now for a sequence $(c_j)_j$, define the integral

$$I_j(c_j) := \int_{[0,1]^j} u_j(x_1,\dots,x_{j-1},c_jx_j)\,\mathrm{d}x_1\cdots\mathrm{d}x_j.$$

Observe that the sum in (8.72) is the Upper Riemann sum for the integral $I_j(1)$, noting that $u_j:[0,1]^j\to[0,1]$ is decreasing in each component. And since the Upper and Lower Riemann sums (which squeeze $I_j(1)$) overlap in $(N-1)^j$ points, their difference is $\ll (N^j-(N-1)^j)/N^j = 1-(1-1/N)^j \ll j/N$. In particular the sum in (8.72) equals $I_j(1)+O(j/N)$.

Similarly (8.72) is the Lower Riemann sum for $I_j(k-j)$. Thus we obtain

$$N^{-j}\sum_{i_1,\dots,i_j=1}^{N} u_j\Big(\tfrac{i_1-1}{N},\dots,\tfrac{i_j-1}{N}\Big) = I_j(1)+O(j/N),$$

$$N^{-j}\sum_{i_1,\dots,i_j=1}^{N} u_j\Big(\tfrac{i_1}{N},\dots,\tfrac{i_j}{N}(k-j)\Big) = I_j(k-j)+O(j/N).$$

Recalling $N=4^k$ and $j=\lfloor k/4\rfloor$, we see (8.72) and (8.73) become

$$f_k' \le (e^\gamma + O(k/2^{k/4}))\,I_j(1), \tag{8.74}$$

$$f_k' \ge (e^\gamma - O(k/2^{k/4}))\,I_j(k-j). \tag{8.75}$$

In the next section, we shall establish the following quantitative result.

**Theorem 8.4.7.** *Let $(c_j)_j$ be any non-negative sequence. Then $I_j(c_j)$, as in* (8.77)*, satisfies*

$$I_j(c_j) = e^{-\gamma} + O\left(2^{-j}\left(1 + c_j\right)\right).$$

*Remark* 8.4.8. The qualitative result that $I_n(c_n) = e^{-\gamma} + o(1)$ may be established in the wider regime where $\limsup_{n\to\infty} \frac{1}{n} \log c_n < 1$ (see Lemma 8.5.3 for more precise statement) but the bound above is sufficient for the purpose of this article.

In particular, Theorem 8.4.7 gives

$$I_j(1) = e^{-\gamma} + O(2^{-j}),$$
$$I_j(k-j) = e^{-\gamma} + O(k/2^j).$$

Thus plugging into (8.74) and (8.75) we obtain $f'_k = 1 + O(k/2^{k/4})$. Hence by (8.63) we conclude

$$f_k = f'_k + O(k/2^{k/4}) = 1 + O(k/2^{k/4}). \tag{8.76}$$

This completes the proof of Proposition 8.4.1.

## 8.5 A sequence of integrals

In this section, we prove Theorem 8.4.7. This implies Theorem 8.1.6 for $I_j(1)$. Recalling $u_j$ in (8.65), we defined the following sequence of integrals, for a sequence $(c_j)_j$,

$$I_j = I_j(c_j) := \int_{[0,1]^j} \frac{\mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_j}{1 + x_1(1 + x_2(\cdots(1 + x_{j-1}(1 + c_j x_j))\cdots))}. \tag{8.77}$$

### 8.5.1 Probabilistic setup

In the following, all random variables are assumed to live in a common reference probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

**Lemma 8.5.1.** *Let $U, U_1, U_2, \ldots$ be i.i.d.* Uniform$[0,1]$ *random variables, and define*

$$F_n(x) := 1 + U_n x \qquad \forall x \in \mathbb{R}, \quad n \in \mathbb{N},$$

*as well as the sequence of iterated maps $S_0(x) := x$ and*

$$S_n(x) := F_1 \circ F_2 \circ \cdots \circ F_n(x) \qquad \forall n \in \mathbb{N}. \tag{8.78}$$

*Then for any sequence of random variables $(V_n)_n$ satisfying $\theta^{-n}|V_n| \xrightarrow[n\to\infty]{a.s.} 0$ for some $\theta \in (1, e)$, we have*

$$\lim_{n\to\infty} S_n(V_n) = S_\infty \qquad \textit{almost surely} \tag{8.79}$$

*for some non-trivial random variable $S_\infty \in (0, \infty)$. In particular, $S_n(1)$ converges to $S_\infty$ almost surely.*

*Proof.* Note by definition,

$$S_n(1) = 1 + \sum_{j=1}^{n} \prod_{k=1}^{j} U_k$$

and it is immediate that $S_{n+1}(1) \geq S_n(1) \geq 1$ for all $n \in \mathbb{N}$. Therefore, by monotone convergence we have almost surely

$$S_n(1) \xrightarrow{n\to\infty} S_\infty := 1 + \sum_{j=1}^{\infty} \prod_{k=1}^{j} U_k \geq 1. \tag{8.80}$$

Moreover, we have $\mathbb{P}(S_\infty < \infty) = 1$ because the expectation $\mathbb{E}[S_\infty] = 2$ is finite. Indeed,

$$\mathbb{E}[S_\infty] = 1 + \sum_{j=1}^{\infty} \mathbb{E}\left[\prod_{k=1}^{j} U_k\right] = 1 + \sum_{j=1}^{\infty} \mathbb{E}[U]^j = \sum_{j=0}^{\infty} 2^{-j} = 2. \tag{8.81}$$

Now suppose $\theta \in (1, e)$, and $(V_n)_n$ is some sequence of random variables such that $\theta^{-n} V_n$ converges almost surely to 0 as $n \to \infty$. Since each of the functions $F_n$ is linear with Lipschitz constant $\|F_n\|_{\text{Lip}} = U_n$, we have

$$\begin{aligned} |S_n(V_n) - S_n(1)| &= |F_1 \circ F_2 \circ \cdots \circ F_n(V_n) - F_1 \circ F_2 \circ \cdots \circ F_n(1)| \\ &= \left[\prod_{j=1}^{n} \|F_j\|_{\text{Lip}}\right] |V_n - 1| = \left[\prod_{j=1}^{n} U_n\right] |V_n - 1| \\ &\leq \exp\left(\sum_{j=1}^{n} \log U_j\right) (1 + |V_n|). \end{aligned} \tag{8.82}$$

Since $\mathbb{E}[\log U_i] = \int_0^1 \log u \, \mathrm{d}u = -1$, the strong law of large numbers gives

$$\frac{1}{n} \sum_{j=1}^{n} \log U_j \xrightarrow[n\to\infty]{a.s.} -1.$$

In particular, if we choose $\varepsilon \in (0, 1 - \log\theta)$, then almost surely there exists some (random) $n_0 = n_0(\varepsilon) \in \mathbb{N}$ such that

$$\frac{1}{n}\sum_{j=1}^{n} \log U_j \leq -1 + \varepsilon \qquad \text{for all } n \geq n_0.$$

Substituting this into (8.82), we obtain

$$|S_n(V_n) - S_n(1)| \leq (\theta/e^{1-\varepsilon})^n \left[\theta^{-n}(1 + |V_n|)\right] = O\left(\theta^{-n}(1 + |V_n|)\right).$$

Thus by assumption of $V_n$ we see $S_n(V_n) - S_n(1)$ converges to 0 almost surely. And since $S_n(1)$ converges to $S_\infty$ almost surely, by (8.80), we conclude $S_n(V_n)$ also converges to $S_\infty$ almost surely. □

**Corollary 8.5.1.1.** *For the integral $I_n$ as in* (8.77)*, we have*

$$I_n = \mathbb{E}\left[\frac{1}{S_{n-1}(1 + c_n U_n)}\right].$$

*In particular, for any sequence $(c_n)_n$ satisfying $0 \leq c_n = o(\theta^n)$ for some $\theta \in (1, e)$,*

$$\lim_{n\to\infty} I_n = \mathbb{E}[S_\infty^{-1}] =: I_\infty.$$

*Proof.* The probablistic representation of the iterated integral $I_n$ follows immediately by construction. If we let $V_{n-1} := 1 + c_n U_n$, then $\theta^{-n} V_n$ converges to 0 almost surely as $n \to \infty$, and by Lemma 8.5.1 we also obtain that $S_{n-1}(V_{n-1})$ converges almost surely to $S_\infty$. Since

$$\frac{1}{S_{n-1}(V_{n-1})} \leq \frac{1}{S_{n-1}(1)} \leq 1,$$

we conclude by dominated convergence that

$$\lim_{n\to\infty} I_n = \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{S_{n-1}(V_{n-1})}\right] = \mathbb{E}[S_\infty^{-1}].$$

□

The next step is a simple but crucial characterisation of the distribution of $S_\infty$.

**Lemma 8.5.2.** *The law of $S_\infty$ is the unique proper solution*[4] *to the distributional equality*

$$S_\infty \overset{(d)}{=} 1 + U S_\infty, \tag{8.83}$$

*where on the right-hand side, the random variable $U \sim \text{Uniform}[0, 1]$ is independent of $S_\infty$.*

[4] i.e. distributions that satisfy $\mathbb{P}(|S_\infty| < \infty) = 1$.

*Proof.* Since we have ruled out distributions that charge infinity with non-trivial mass, it suffices to verify the following two claims:

1. $S_\infty := \lim_{n\to\infty} S_n(x)$ *satisfies* (8.83).

   This is straightforward by a quick reordering of the underlying i.i.d. random variables/iterated maps. Indeed,

   $$\begin{aligned} S_n(x) = F_1 \circ \cdots \circ F_n(x) &\stackrel{(d)}{=} F_n \circ F_1 \circ F_2 \circ \cdots \circ F_{n-1}(x) \\ &= 1 + U_n S_{n-1}(x) \xrightarrow[n\to\infty]{(d)} 1 + U S_\infty. \end{aligned}$$

2. *The only proper solution to* (8.83) *is given by the law of* $S_\infty$ *in* (8.80).

   To establish this claim, let $X$ be a random variable the law of which is a proper solution to (8.83). Using the notation of Lemma 8.5.1, for any $n \in \mathbb{N}$ we have

   $$X \stackrel{(d)}{=} F_1(X) \stackrel{(d)}{=} \ldots \stackrel{(d)}{=} F_1 \circ F_2 \circ \cdots \circ F_n(X) = S_n(X).$$

   Since the distribution of $X$ is proper, and in particular $2^{-n}|X| \xrightarrow[n\to\infty]{a.s.} 0$, Lemma 8.5.1 with $V_n = X$ implies that $S_n(X) \xrightarrow[n\to\infty]{a.s.} S_\infty$. In other words,

   $$X \stackrel{(d)}{=} S_\infty = 1 + \sum_{j=1}^{\infty} \prod_{k=1}^{j} U_k,$$

   which concludes the proof.

□

## 8.5.2 Proof of Lemma 8.4.7

The recursive distributional equation (8.83) is a very convenient tool that helps us control the rate of convergence of $S_n(\cdot)$ and extract information about the statistical behaviour of $S_\infty$ at the same time. We first explain how to estimate the difference between $I_n(c_n)$ and its limit $I_\infty$.

**Lemma 8.5.3.** *For any non-negative sequence* $(c_n)_n$*, we have*

$$|I_n(c_n) - I_\infty| \le 2^{1-n} + \mathbb{E}\left[\min\left(1, c_n \prod_{j\le n} U_j\right)\right] \;\le\; 2^{-n}(2 + c_n). \tag{8.84}$$

*Proof.* Recall that $V_{n-1} := 1 + c_n U_n$ and

$$I_n(c_n) = \mathbb{E}\left[\frac{1}{S_{n-1}(V_{n-1})}\right].$$

On the other hand, if we introduce a new random variable $T \stackrel{(d)}{=} S_\infty$ that is independent of all of the $U_j$'s, we see that

$$S_{n-1}(T) \stackrel{(d)}{=} S_\infty \qquad \text{and hence} \qquad I_\infty = \mathbb{E}[S_\infty^{-1}] = \mathbb{E}\left[\frac{1}{S_{n-1}(T)}\right]$$

by the distributional fixed point equation (8.83).

Using the fact that $S_{n-1}(\cdot)$ is linear with Lipschitz constant $\|S_{n-1}\|_{\mathrm{Lip}} = \prod_{j\le n-1} U_j$, we have

$$\begin{aligned} |I_n(c_n) - I_\infty| &= \left|\mathbb{E}\left[\frac{1}{S_{n-1}(V_{n-1})} - \frac{1}{S_{n-1}(T)}\right]\right| \\ &\le \mathbb{E}\left[\frac{\|S_{n-1}\|_{\mathrm{Lip}}\,|V_{n-1} - T|}{S_{n-1}(V_{n-1})S_{n-1}(T)}\right] \\ &\le \mathbb{E}\left[\frac{\|S_{n-1}\|_{\mathrm{Lip}}\,|T-1|}{S_{n-1}(T)}\right] + \mathbb{E}\left[\frac{\|S_{n-1}\|_{\mathrm{Lip}}\,|V_{n-1}-1|}{S_{n-1}(V_{n-1})}\right]. \end{aligned} \tag{8.85}$$

Since $\mathbb{E}[U_j] = 1/2$ and $\mathbb{E}[T] = \mathbb{E}[S_\infty] = 2$ by (8.81), independence implies the first term in the right-hand side of (8.85) equals

$$\mathbb{E}\left[\frac{\|S_{n-1}\|_{\mathrm{Lip}}\,|T-1|}{S_{n-1}(T)}\right] = \mathbb{E}\left[\left(\prod_{j\le n-1} U_j\right)|T-1|\right] = \mathbb{E}[T-1]\prod_{j\le n-1}\mathbb{E}\left[U_j\right] = 2^{1-n}.$$

Next, observe that $V_{n-1} - 1 = c_n U_n$ and $S_{n-1}(x) \ge 1$ for any $x \ge 0$. This means

$$\frac{\|S_{n-1}\|_{\mathrm{Lip}}\,|V_{n-1}-1|}{S_{n-1}(V_{n-1})} \le c_n \prod_{j\le n} U_j.$$

On the other hand, $S_{n-1}(V_{n-1}) \ge 1 + V_{n-1}\|S_{n-1}\|_{\mathrm{Lip}} \ge \|S_{n-1}\|_{\mathrm{Lip}}\,|V_{n-1}-1|$, which leads to a slightly improved bound

$$\frac{\|S_{n-1}\|_{\mathrm{Lip}}\,|V_{n-1}-1|}{S_{n-1}(V_{n-1})} \le \min\left(1, c_n\prod_{j\le n} U_j\right).$$

Taking expectation both sides and plugging this back into (8.85) yields the first inequality in (8.84), and the second inequality in (8.84) follows from $\mathbb{E}[c_n\prod_{j\le n} U_j] = 2^{-n}c_n$. □

It remains to show that the value of $I_\infty$ equals $e^{-\gamma}$. This will be achieved using the recursive distributional equation (8.83) with the help of Laplace transform $\phi(t) := \mathbb{E}[e^{-tS_\infty}]$, which is intrinsically related to our problem because

$$I_\infty = \mathbb{E}[S_\infty^{-1}] = \mathbb{E}\left[\int_0^\infty e^{-tS_\infty} \mathrm{d}t\right] = \int_0^\infty \phi(t)\,\mathrm{d}t \tag{8.86}$$

by Fubini's theorem. Let us first highlight that:

**Lemma 8.5.4.** *The Laplace transform $\phi(t) := \mathbb{E}[e^{-tS_\infty}]$ satisfies*

$$te^t\phi(t) = \int_0^t \phi(v)\,\mathrm{d}v, \qquad t \geq 0. \tag{8.87}$$

*In particular,*

$$I_\infty = \lim_{t\to\infty} te^t\phi(t). \tag{8.88}$$

*Proof.* From the recursive distribution equation (8.83), we have

$$\begin{aligned}\phi(t) := \mathbb{E}[e^{-tS_\infty}] &= \mathbb{E}[e^{-t(1+US_\infty)}] \\ &= e^{-t}\int_0^1 \mathbb{E}[e^{-tuS_\infty}]\mathrm{d}u = \frac{e^{-t}}{t}\int_0^t \mathbb{E}[e^{-vS_\infty}]\mathrm{d}v.\end{aligned}$$

Hence $te^t\phi(t) = \int_0^t \phi(v)\mathrm{d}v$, as claimed. In particular $\lim_{t\to\infty} te^t\phi(t) = \int_0^\infty \phi(v)\mathrm{d}v = I_\infty$. □

*Proof of Lemma 8.4.7.* Differentiating the equality (8.87) yields $te^t\phi'(t) + (te^t)'\phi(t) = \phi(t)$, which may be rewritten as

$$\frac{\phi'(t)}{\phi(t)} = \frac{1-(te^t)'}{te^t} = \frac{e^{-t}-1}{t} - 1.$$

Since $\phi(0) = 1$, we then obtain

$$\begin{aligned}\log\phi(x) = \log\phi(x) - \log\phi(0) &= \int_0^x \frac{\mathrm{d}}{\mathrm{d}u}\Big[\log\phi(u)\Big]\mathrm{d}u = \int_0^x \frac{\phi'}{\phi}(u)\mathrm{d}u \\ &= \int_0^x \Big(\frac{e^{-u}-1}{u} - 1\Big)\mathrm{d}u = \int_0^x \big(e^{-u}-1\big)\,\frac{\mathrm{d}u}{u} - x \\ &= \big[(e^{-u}-1)\log u\big]_0^x + \int_0^x e^{-u}\log u\,\mathrm{d}u - x.\end{aligned}$$

From Euler's identity for $\gamma$ [82, Eq. (2.2.8)],

$$\gamma = -\int_0^\infty e^{-u}\log u\,\mathrm{d}u,$$

we see as $x \to \infty$,

$$\log \phi(x) = -\log x - x - \gamma + o(1).$$

Substituting this into (8.88), we obtain

$$I_\infty = \lim_{x\to\infty} x e^x \phi(x) = e^{-\gamma}.$$

Combining this with Lemma 8.5.3, we conclude that $I_n = e^{-\gamma} + O(2^{-n}(1+c_n))$. $\square$

### 8.5.3 A second proof that $I_n$ tends to $e^{-\gamma}$

We provide yet another proof that $I_\infty = \lim_{n\to\infty} I_n$ equals $e^{-\gamma}$ below, using the fact that the integral of the Dickman function equals $e^\gamma$.

To this, first observe that

$$e^x \phi(x) = \mathbb{E}\left[e^{-x\tilde{S}_\infty}\right], \qquad \tilde{S}_\infty := S_\infty - 1.$$

In order to obtain the desired result, we will need to prove as $x \to \infty$

$$\mathbb{E}\left[e^{-x\tilde{S}_\infty}\right] \sim \frac{e^{-\gamma}}{x} \qquad \text{or equivalently} \qquad \mathbb{P}\left(\tilde{S}_\infty \le \frac{1}{x}\right) \sim \frac{e^{-\gamma}}{x},$$

where the equivalence follows from a standard Tauberian argument (see e.g. [54, Theorem XIII.5.3]).

Let us recall (8.83), from which we see that $\tilde{S}_\infty$ is the unique (proper) distributional solution to

$$\tilde{S}_\infty \overset{(d)}{=} U(1+\tilde{S}_\infty)$$

for an independent Uniform$[0,1]$ variable $U$. For any $s \ge 0$, we have

$$\tilde{F}(s) := \mathbb{P}\left(\tilde{S}_\infty \le s\right) = \mathbb{P}\left(U(1+\tilde{S}_\infty) \le s\right) = \int_0^1 \tilde{F}\big(\frac{s}{u} - 1\big)\mathrm{d}u = s\int_{s-1}^\infty \frac{\tilde{F}(v)\,\mathrm{d}v}{(1+v)^2}.$$

Specialising this to $s \in [0,1]$, we see that $\tilde{F}(s) = cs$ where $c = \int_0^\infty \tilde{F}(v)\frac{\mathrm{d}v}{(1+v)^2}$. As for $s > 1$ the above equality leads to

$$\left[\tilde{F}(s)/s\right]' = -\tilde{F}(s-1)/s^2 \qquad \text{or} \qquad s\tilde{F}'(s) - \tilde{F}(s) + \tilde{F}(s-1) = 0.$$

If we now express everything in terms of the probability density function $\tilde{f}(s) := \tilde{F}'(s)$ of the random variable $\tilde{S}_\infty$, then we arrive at the following differential-difference equation:

$$\begin{cases} \tilde{f}(s) = c, & s \in [0,1], \\ s\tilde{f}'(s) + \tilde{f}(s-1) = 0, & s \ge 1. \end{cases} \tag{8.89}$$

This is the initial value problem satisfied by the Dickman function $\rho(\cdot)$, except with a different initialisation on $[0,1]$. The linearity of the differential equation on $s \geq 1$ implies $\tilde{f} \propto \rho$, whereas the fact that $\tilde{f}(s)$ must integrate to 1 implies

$$\tilde{f}(s) = e^{-\gamma}\rho(s), \qquad s \geq 0,$$

since we have (c.f. [106, p. 212])

$$e^{\gamma} = \int_0^{\infty} \rho(u)\,\mathrm{d}u.$$

As $\rho(0) = 1$, we can conclude that

$$I_{\infty} = \lim_{x\to\infty} x\mathbb{P}(\tilde{S}_{\infty} \leq x^{-1}) = \tilde{f}(0) = e^{-\gamma}.$$

□

*Remark* 8.5.5. Observe $\tilde{S}_{\infty}$ has probability density function given by the Dickman distribution $\tilde{f}(s) = e^{-\gamma}\rho(s)$. As such, we may evaluate $I_{\infty}$ without the Tauberian argument by considering directly

$$\begin{aligned} I_{\infty} = \mathbb{E}\left[\frac{1}{1+\tilde{S}_{\infty}}\right] &= \int_0^{\infty} \frac{\tilde{f}(s)}{1+s}\,\mathrm{d}s \\ &= -\int_0^{\infty} \tilde{f}'(s+1)\,\mathrm{d}s = \widetilde{f}(1) = e^{-\gamma} \end{aligned}$$

where the last line follows from (8.89).

# Chapter 9

# A generalization of primitive sets and a conjecture of Erdős

## 9.1 Introduction

We recall a set of integers greater than 1 is called primitive if no element divides any other. Similarly, We say that a set $A$ of integers greater than 1 with $|A| \geq k+1$ is *$k$-primitive* if no element divides the product of $k$ distinct other elements. Note 1-primitive means primitive, and that $k$-primitive implies $j$-primitive for all $k \geq j \geq 1$. As the central example, the set of primes $\mathcal{P}$ is $k$-primitive for any $k \geq 1$.

In 1938, Erdős [42] first studied the maximal cardinality of 2-primitive sets in an interval. Chan, Győri, and Sárközy [20] extended it to all $k$ and it was subsequently improved in [19] and [110]. In this Chapter, we prove the Erdős conjecture for 2-primitive sets (and hence $k$-primitive for all $k \geq 2$).

**Theorem 9.1.1.** *For any 2-primitive set $A$, we have $f(A) \leq f(\mathcal{P})$.*

Later in Chapter 12, we shall address Erdős' conjecture in the broader context of 1-primitive sets. In the special case of 2-primitive sets we may actually prove a much stronger result, in Theorem 9.1.2 below, which does not hold for 1-primitive sets.

We emphasize that the primes are not optimal among primitive sets with respect to logarithmic density. Indeed, Erdős, Sárközy and Szemerédi [49] obtained the best possible upper bound

$$\sum_{\substack{n\in A\\ n\leq x}} \frac{1}{n} \leq \Big(\frac{1}{\sqrt{2\pi}} + o(1)\Big)\frac{\log x}{\sqrt{\log\log x}}$$

for any primitive set $A$, while Erdős [43] showed that

$$\sum_{\substack{n\in A'\\ n\leq x}} \frac{1}{n} \geq \Big(\frac{1}{\sqrt{2\pi}} + o(1)\Big)\frac{\log x}{\sqrt{\log\log x}}$$

where $A'$ is the set of positive integers $a \leq x$ with $\Omega(a) = [\log\log x]$. (Here, $\Omega(a)$ is the number of prime factors of $a$, counted with multiplicity.) By contrast, the primes satisfy

$$\sum_{p\leq x} \frac{1}{p} = \log\log x + O(1).$$

Nevertheless, one may wonder if the primes still maximize the logarithmic density among 2-primitive sets. Indeed, we prove this to be the case.

As a quick note, by cancelling primes on each side, in Theorem 9.1.1 it suffices to prove $f(A) \leq f(\mathcal{P}(A))$, for $\mathcal{P}(A) = \{p : p \mid a \in A\}$.

**Proposition 9.1.0.1.** *For all $x \geq 2$ and any 2-primitive set $A$,*

$$\sum_{\substack{n \in A \\ n \leq x}} \frac{1}{n} \leq \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p}$$

We use this to deduce Theorem 9.1.1.

*Proof of Theorem 9.1.1 given Proposition 9.1.0.1.* For any 2-primitive set $A$, we have

$$F(x) := \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p} - \sum_{\substack{n \in A \\ n \leq x}} \frac{1}{a} \geq 0$$

for all $x \geq 2$ by Proposition 9.1.0.1. Then by partial summation,

$$\sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p \log p} - \sum_{\substack{n \in A \\ n \leq x}} \frac{1}{n \log n} = \frac{F(x)}{\log x} + \int_{2^-}^{x} \frac{F(u)}{u \log^2 u} du \geq 0.$$

Hence taking $x \to \infty$ gives $f(\mathcal{P}(A)) \geq f(A)$ as desired. □

In light of Proposition 9.1.0.1, it is natural to ask if there exists an exponent $\lambda < 1$ for which

$$\sum_{\substack{n \in A \\ n \leq x}} \frac{1}{n^\lambda} \leq \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p^\lambda} \tag{9.1}$$

holds for all 2-primitive $A$, $x \geq 2$. Banks and Martin [6] settled the question in the setting of 1-primitive sets, proving (9.1) holds for all primitive $A$ if and only if

$$\lambda \geq \tau_1 := 1.1403659 \cdots,$$

where $t = \tau_1$ is the unique real solution to the equation

$$\sum_{\mathcal{P}} p^{-t} = 1 + \left(1 - \sum_{\mathcal{P}} p^{-2t}\right)^{1/2}.$$

The fact that $\tau_1$ is markedly larger than 1 gives some indication as to why the full Erdős conjecture remains open.

In the setting of 2-primitive sets, we extend the range of valid exponents $\lambda$.

**Theorem 9.1.2.** *For any $\lambda \geq 0.7983$, $x \geq 2$, and any 2-primitive set $A$,*

$$\sum_{\substack{n \in A \\ n \leq x}} \frac{1}{n^\lambda} \leq \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} \frac{1}{p^\lambda} \tag{9.2}$$

We remark it suffices to verify Theorem 9.1.2 with $\lambda = 0.7983$. Indeed, suppose that $F_\lambda(x) \geq 0$ for all $x \geq 2$, where

$$F_t(x) = \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} p^{-t} - \sum_{\substack{n \in A \\ n \leq x}} n^{-t}.$$

Then, by partial summation, for any $t > \lambda$,

$$F_t(x) = x^{\lambda - t} F_\lambda(x) + (t - \lambda) \int_2^x u^{\lambda - t - 1} F_\lambda(u)\, du \geq 0.$$

Hence we may define the **critical exponent** $\tau_2$ for 2-primitive sets, as the infimum over all $\lambda$ for which (9.2) holds.

We also note that Theorem 9.1.2 with $\lambda = 1$ gives us Proposition 9.1.0.1. However, Theorem 9.1.2 does not hold for every positive value of $\lambda$. Indeed, in [42], Erdős showed that there is a 2-primitive set $A$ in $[1, x]$ of cardinality $\pi(x) - \pi(x^{1/3}) + cx^{2/3}/(\log x)^2$. It consists of primes in $(x^{1/3}, x]$ and a subset of $\{p_1 p_2 p_3 : \text{ primes } p_i \leq x^{1/3}\}$ where the triples $\{p_1, p_2, p_3\}$ form a Steiner triple system. Thus, by the prime number theorem,

$$\sum_{a \in A} \frac{1}{a^\lambda} \geq \sum_{x^{1/3} < p \leq x} \frac{1}{p^\lambda} + \frac{cx^{2/3}}{(\log x)^2} \frac{1}{x^\lambda} > \sum_{p \leq x} \frac{1}{p^\lambda}$$

when $\lambda < 0.5$ and $x$ is sufficiently large. Hence the above argument and Theorem 9.1.2 together imply that the critical exponent lies in the interval

$$0.5 \leq \tau_2 \leq 0.7983. \tag{9.3}$$

## 9.2 Combinatorial Lemmas

Before proving Theorem 9.1.2, we need lemmas in counting the maximal number of elements in a $k$-primitive set.

We first recall the following famous result due to Erdős and Szekeres [52], whose proof we provide for completeness.

**Lemma 9.2.1** (Erdős–Szekeres)**.** *A sequence of $(r-1)(s-1)+1$ real numbers has either a monotonic nondecreasing subsequence of length $r$ or a monotonic nonincreasing subsequence of length $s$.*

*Proof.* Say the sequence is $a_1, a_2, \ldots, a_n$, where $n = (r-1)(s-1)+1$. For each $a_i$ consider the ordered pair $(b_i, c_i)$, where $b_i$ is the length of the longest nondecreasing

subsequence ending at $a_i$ and $c_i$ is the length of the longest nonincreasing subsequence ending at $a_i$. Then no two pairs $(b_i, c_i)$ and $(b_j, c_j)$ can be equal, so for at least one choice of $i$ we have $b_i \geq r$ or $c_i \geq s$. □

We next bound the size of a $k$-primitive set based on the number of prime factors used to generate its elements.

**Lemma 9.2.2.** *For $k \geq 2$, suppose $A$ is a $k$-primitive set and $T \subset A$ with $|\mathcal{P}(T)| = n$. If $n \leq k$, then $|T| \leq n$. If $n = k+1$, then $|T| \leq k+2$. Further, for $k = 2$, $n = 4$ we have $|T| \leq 19$.*

*Proof.* Let $\mathcal{P}(T) = \{q_1, \ldots, q_n\}$ and write each $t = \prod_i q_i^{e_i} \in T$ as an exponent vector $\vec{v} = (e_1, \ldots, e_n)$. Define the notation $\vec{v} \geq 0$ if $e_i \geq 0$ for all $i$, and define $\vec{v} \leq \vec{w}$ if $\vec{w} - \vec{v} \geq 0$. Suppose that $|T| \geq n$. Let $\vec{v}_1$ be such that $e_1$ is maximal. Then let $\vec{v}_2$ be such that $e_2$ is maximal among the remaining vectors, and similarly define $\vec{v}_3, \ldots, \vec{v}_n$. Thus, the chosen vectors are distinct.

Case $n \leq k$: If $|T| \geq n+1$ then $T$ has some vector $\vec{v} \neq \vec{v}_i$ for all $i$. But then $\vec{v} \leq \vec{v}_1 + \cdots + \vec{v}_n$. This implies that $T$, and hence $A$, is not $n$-primitive, and since $n \leq k$, it implies that $A$ is not $k$-primitive, a contradiction. Hence we cannot have $|T| \geq k+1$ when $n \leq k$.

Case $n = k+1$: If $|T| \geq n+2$ then $T$ has vectors $\vec{w}_1 \neq \vec{w}_2$ with $\vec{w}_j \notin \{\vec{v}_1, \ldots \vec{v}_n\}$ for $j = 1, 2$. Write $\vec{w}_j = (f_1^{(j)}, \ldots, f_n^{(j)})$. By the pigeonhole principle, we may assume

$$f_i^{(1)} \leq f_i^{(2)}$$

for at least $n/2$ values of $i$, say $i = 1, \ldots, \lceil n/2 \rceil$. Thus we deduce

$$\vec{w}_1 \leq \vec{w}_2 + \vec{v}_{\lceil n/2 \rceil + 1} + \cdots + \vec{v}_n$$

contradicting $T$ as $k$-primitive, since $1 + \lfloor n/2 \rfloor = 1 + \lfloor (k+1)/2 \rfloor \leq k$.

Now say $k = 2$, $n = 4$. Suppose there are 20 members in $T$ with corresponding vectors

$$\vec{w}_i := (e_{i,1}, e_{i,2}, e_{i,3}, e_{i_4}) \text{ for } 1 \leq i \leq 20.$$

Since $A$ is 2-primitive, so is $T$. Without loss of generality, say $\vec{w}_{18}$ has maximal first coordinate, $\vec{w}_{19} \neq \vec{w}_{18}$ has maximal second coordinate among the remaining 19 vectors, and $\vec{w}_{20} \neq \vec{w}_{18}, \vec{w}_{19}$ has maximal third coordinate among the remaining 18 vectors. Arrange the remaining 17 vectors in ascending order of their first coordinate (i.e., $e_{1,1} \leq e_{2,1} \leq \ldots \leq e_{17,1}$). By Lemma 9.2.1, there is a monotonic sequence of

length 5 among the $e_{i,2}$'s. Without loss of generality, say $e_{1,2}$, $e_{2,2}$, $e_{3,2}$, $e_{4,2}$, $e_{5,2}$ form such a sequence.

Case 1: $e_{1,2} \le e_{2,2} \le e_{3,2} \le e_{4,2} \le e_{5,2}$. Consider the numbers $e_{i,3}$ for $i = 1, \dots, 5$. By Lemma 9.2.1, there is a monotonic sequence of length 3 among the $e_{i,3}$'s, without loss of generality, say it is $e_{1,3}, e_{2,3}, e_{3,3}$. If $e_{1,3} \le e_{2,3} \le e_{3,3}$, this forces $e_{2,4} > e_{1,4} + e_{3,4}$ for otherwise $\vec{w}_2 \le \vec{w}_1 + \vec{w}_3$, contradicting $T$ being 2-primitive. But this implies that $\vec{w}_1 \le \vec{w}_2$ which contradicts $T$ being primitive. Hence, we must have $e_{1,3} \ge e_{2,3} \ge e_{3,3}$. Again, this forces $e_{2,4} > e_{1,4} + e_{3,4}$, which in turn implies that $\vec{w}_1 \le \vec{w}_2 + \vec{w}_{20}$, again a contradiction.

Case 2: $e_{1,2} \ge e_{2,2} \ge e_{3,2} \ge e_{4,2} \ge e_{5,2}$. By Lemma 9.2.1, there is a monotonic sequence of length 3 among the $e_{i,3}$'s, without loss of generality, say it is $e_{1,3}, e_{2,3}, e_{3,3}$. If $e_{1,3} \le e_{2,3} \le e_{3,3}$, then again this forces $e_{2,4} > e_{1,4} + e_{3,4}$. But then $\vec{w}_1 \le \vec{w}_2 + \vec{w}_{19}$. Hence, we must have $e_{1,3} \ge e_{2,3} \ge e_{3,3}$. This forces $e_{2,4} > e_{1,4} + e_{3,4}$. But then $\vec{w}_3 \le \vec{w}_2 + \vec{w}_{18}$, again a contradiction.

Therefore, there can be at most 19 members in $T$. □

*Remark* 9.2.1. We will not need it here, but by similar methods one can prove that if $T$ is a 2-primitive set of positive integers with $|\mathcal{P}(T)| = n \ge 3$, then $|T| \le 9^{2^{n-3}}$.

## 9.3 Proof of Theorem 9.1.2

Let $A \subset (1, x]$ be a 2-primitive set. Let $\lambda$ be a parameter with $0.79 \le \lambda < 1$, to be defined later. First, we partition $A$ into primes $S$ and composites $T$. Note $S$ and $\mathcal{P}(T)$ are disjoint since $A$ is primitive. For a prime $p$, define

$$T_p := \{t \in T : p \mid t\}$$

If some prime $p \in \mathcal{P}(T)$ satisfies

$$\sum_{t \in T_p} \frac{1}{t^\lambda} \le \frac{1}{p^\lambda}, \tag{9.4}$$

then we replace the members of $T_p$ with the prime $p$ (i.e., redefine $A = (T \backslash T_p) \cup \{p\}$). This would make $\sum_{T_p} t^{-\lambda}$ at least as big while keeping $A$ 2-primitive. Repeat the process with each prime $p \in \mathcal{P}(T)$ until no such prime satisfies (9.4). If $T = \emptyset$ after doing this, then $A = S$ consists of primes so Proposition 9.1.0.1 follows. Otherwise $T \ne \emptyset$, so we may assume

$$\sum_{t \in T_p} \frac{1}{t^\lambda} > \frac{1}{p^\lambda} \qquad \text{for all} \quad p \in \mathcal{P}(T). \tag{9.5}$$

Consider the set

$$D := \{t/p : t \in T,\ p \mid t\} \tag{9.6}$$

We record some useful properties of $T$ and $D$.

**Lemma 9.3.1.** *Let $T$ be a 2-primitive set for which (9.5) holds and let $D$ be as in (9.6).*

*(i) For each $p \in \mathcal{P}(T)$, $T_p$ has at least 3 elements.*

*(ii) The map sending ordered pairs $(t, p)$ with $t \in T$ and $p \mid t$ to $t/p \in D$ is injective.*

*(iii) Each $t \in T$ has at least 3 prime factors.*

*(iv) $D$ is a primitive set of composite numbers.*

*Proof.* (i) Note that (9.5) implies that the members of $T$ are composite. In fact, for $p \in \mathcal{P}(T)$, (9.5) implies that

$$\sum_{t \in \mathcal{P}(T)} \frac{1}{(t/p)^\lambda} > 1 > 2^{-0.79} + 3^{-0.79},$$

so (i) holds.

(ii) If not, then $t_1/p_1 = t_2/p_2$ for some $t_1, t_2$, $p_1 \mid t_1$, and $p_2 \mid t_2$. If $t_1 \neq t_2$, by (i) there exists some $p_1 k \in T_{p_1}$ other than $t_1, t_2$. But then $t_1 = (t_1/p_1)p_1 = (t_2/p_2)p_1 \mid t_2(p_1 k)$, which contradicts $T$ as 2-primitive. Hence $t_1 = t_2$, which forces $p_1 = p_2$.

(iii) If not, say $t = pq$. Since $T_p, T_q$ each have at least 3 elements, there are some $pm$ and $qn$ other than $t \in T$. But then, $t = pq \mid (pm)(qn)$ which contradicts $T$ as 2-primitive.

(iv) If not, then $(t/p) \mid (t_1/p_1)$ for some $t, t_1 \in T$, $p \mid t$, $p_1 \mid t_1$, and $t/p \neq t_1/p_1$. If $p_1 = p$, then $t \mid t_1$ which contradicts $T$ as primitive. And if $p_1 \neq p$, then there is some $pl \in T_p$ other than $t$ and $t_1$. This implies $t \mid t_1 \cdot pl$, and since $t \neq t_1$ (otherwise $p = p_1$), we have a contradiction to $T$ being 2-primitive. Thus $D$ is primitive, and also composite by (iii). □

For Theorem 9.1.2, we must show

$$\sum_{t \in T} \frac{1}{t^\lambda} - \sum_{p \in \mathcal{P}(T)} \frac{1}{p^\lambda} < 0. \tag{9.7}$$

Suppose $\mathcal{P}(T)$ consists of primes $q_1 < q_2 < \cdots < q_r$. Let $2 = p_1 < p_2 < \cdots < p_r$ be the first $r$ primes in $\mathcal{P}$. We are going to modify the set $T$ by the following process.

First, if each $q_i = p_i$, we let $T$ stand as it is. Otherwise, let $i$ be the smallest index such that $q_i > p_i$. Then $q_j = p_j$ for all $j < i$ and we have $p_i \nmid t$ for all $t \in T$. Then replace each $t \in T_{q_i}$ with $p_i/q_i \cdot t$. This keeps $T$ as 2-primitive, and by (9.5),

$$0 \ < \ \sum_{t\in T_{q_i}} \frac{1}{t^\lambda} - \frac{1}{q_i^\lambda} \ < \ \frac{q_i^\lambda}{p_i^\lambda}\left(\sum_{t\in T_{q_i}} \frac{1}{t^\lambda} - \frac{1}{q_i^\lambda}\right) = \sum_{t\in T_{q_i}} \frac{1}{(p_i/q_i \cdot t)^\lambda} - \frac{1}{p_i^\lambda}.$$

So replacing each $t \in T_{q_i}$ with $p_i/q_i \cdot t$ preserves (9.5). We repeat this process for each $i$ with $q_i > p_i$ and in the end we have $\mathcal{P}(T) = \{p_1, p_2, \dots, p_r\}$. By showing (9.7) for this $T$ it would follow that (9.5) fails for some $p_i$, and this contradiction would prove the theorem.

As just noted, we may assume that $\mathcal{P}(T)$ consists of the primes up to some $Y$, i.e., $\mathcal{P}(T) = \mathcal{P} \cap (1, Y]$, so (9.7) becomes

$$\sum_{t\in T} \frac{1}{t^\lambda} - \sum_{p\le Y} \frac{1}{p^\lambda} < 0. \tag{9.8}$$

For a parameter $0 < \theta < 1$ to be chosen later, we define $\lambda$ as

$$\lambda = \tau(1-\theta), \text{ where } \tau = 1.140366. \tag{9.9}$$

First consider those $t \in T$ with greatest prime factor $P^+(t) \ge t^\theta$. Then $t^{1-\theta} \ge t/P^+(t)$ and so $t^{-\lambda} \le (t/P^+(t))^{-\lambda/(1-\theta)} = (t/P^+(t))^{-\tau}$. Hence

$$\sum_{\substack{t\in T\\ P^+(t)\ge t^\theta}} t^{-\lambda} \le \sum_{\substack{t\in T\\ P^+(t)\ge t^\theta}} \Big(\frac{t}{P^+(t)}\Big)^{-\tau} \le \sum_{p\le Y} p^{-\tau} \tag{9.10}$$

by (9.1), since $\{t/P^+(t) : t \in T\} \subset D$ is primitive by part (iii) of Lemma 9.3.1.

For a positive integer $t$, we consider the following unique factorization

$$t = m(t)M(t)$$

into positive integers $m(t) \le M(t)$ whose ratio $M(t)/m(t)$ is minimal. Let

$$\mathcal{M}(T) = \{m(t) : t \in T\} \cup \{M(t) : t \in T\}.$$

We need two lemmas.

**Lemma 9.3.2.** *For any 2-primitive set $T$, consider the graph on the integers with edges $\{t, m(t)\}$ and $\{t, M(t)\}$ for $t \in T$, where if $m(t) = M(t)$, there is just one edge containing $t$. This graph contains a matching from $T$ into $\mathcal{M}(T)$.*

*Proof.* Let $t \in T$. If $m(t) \notin \{m(t'), M(t')\}$ for all other $t' \in T$, we can match $t$ with $m(t)$. So assume $m(t) \in \{m(t'), M(t')\}$ for some other $t' \in T$. Then $M(t) \notin \{m(t''), M(t'')\}$ for all $t'' \in T$ with $t'' \neq t, t'$, since otherwise $t \mid t't''$, contradicting $T$ being 2-primitive.

If $m(t) < M(t)$, we can match $t$ with $M(t)$. Otherwise, we have $t = m(t)^2$, and then $m(t') < M(t')$. Let $m' = t'/m(t)$. We would like to match $t'$ with $m'$ instead of $m(t)$. Suppose this is blocked by some $t''$ different from $t'$ (and necessarily different from $t$) with $m' \in \{m(t''), M(t'')\}$. But then $t' \mid tt''$, a violation of 2-primitivity. Thus, the matching can be completed. □

**Lemma 9.3.3.** *Suppose $0 < \theta < 1/3$ and that $T$ is 2-primitive with $P^+(t) < t^\theta$ for each $t \in T$. Let $N(z) = |T \cap [2, z]|$. Then, with $q$ running over primes in the interval $I := [z^{(1+\theta)/4}, z^{(1+\theta)/2})$, we have*

$$N(z) < z^{(1+\theta)/2} - \sum_{q \in I} \Big\lfloor \frac{z^{(1+\theta)/2}}{q} \Big\rfloor$$

*Proof.* By Lemma 9.3.2, it suffices to bound $|\mathcal{M}(T \cap [2, z])|$. We first show that $\mathcal{M}(T \cap [2, z]) \subset [1, z^{(1+\theta)/2})$. Let $t \in T$ with $t \leq z$. Say $t = q_1 q_2 ... q_r$ where the primes $q_i$ are written in nondecreasing order. Let $d = q_1 q_2 ... q_i$ be maximal with $d \leq t^{(1-\theta)/2}$. Then $d' = dq_{i+1}$ satisfies $t^{(1-\theta)/2} < d' < t^{(1+\theta)/2}$. Also, $d'' = t/d'$ satisfies the same double inequality. Thus,

$$t^{(1-\theta)/2} < m(t) \leq M(t) < t^{(1+\theta)/2} \leq z^{(1+\theta)/2}.$$

We further note that the members $m$ of $\mathcal{M}(T \cap [2, z])$ satisfy $P^+(m) < z^\theta$, since $m$ divides some member of $T \cap [2, z]$ and every $t$ in that set has $P^+(t) < z^\theta$. In particular, $m$ is not divisible by any prime $q \geq z^\theta$. Note that if $\theta < 1/3$, then $\theta < (1 + \theta)/4$. So, $m$ is not divisible by any prime in the interval $I$. Since no integer below $z^{(1+\theta)/2}$ is divisible by 2 primes from $I$, the lemma follows. □

Set

$$T^p = \{t \in T : P^+(t) = p\},$$

so that $T^p \subset T_p$. We have the following variant of Lemma 9.3.3.

**Lemma 9.3.4.** *For any 2-primitive set $T$ and prime $p$, let $N_p(z)$ denote the number of members $t$ of $T^p$ with $t \leq z$. With $q$ running over the primes in $I_p := (\max\{p, z^{1/4}\}, z^{1/2})$, we have*

$$N_p(z) \leq z^{1/2} - \sum_{q \in I_p} \Big\lfloor \frac{z^{1/2}}{q} \Big\rfloor.$$

*Proof.* Note that if $T$ is 2-primitive, so too is $T^p/p = \{t/p : t \in T^p\}$. Thus, we may apply Lemma 9.3.2 to obtain a matching from $T^p/p$ into $\mathcal{M}(T^p/p)$. The prime factors of each element $t/p \in T^p/p$ are at most $p$, so following the proof of Lemma 9.3.3, we have $m(t/p), M(t/p) \in [t^{1/2}/p, t^{1/2})$. The lemma then follows in the same way as Lemma 9.3.3. □

**Lemma 9.3.5.** *For $x \geq 2$ we have*

$$\sum_{\substack{x^{1/2}<q<x \\ q \text{ prime}}} \Big\lfloor \frac{x}{q} \Big\rfloor \geq \Big( \log 2 - \frac{1.25}{\log x} - \frac{2.5}{(\log x)^2} \Big) x.$$

*Proof.* First suppose that $x \geq 286^2$. We have the sum at most

$$\sum_{x^{1/2}<q<x} \frac{x}{q} - \pi(x).$$

From [124, (3.7)], we have that $\pi(x) < 1.25x/\log x$ and from [124, (3.18)] that

$$\sum_{q<x} \frac{1}{q} > \log\log x + B - \frac{1}{2(\log x)^2},$$

where $B$ is the Mertens constant. Further, from [124, (3.17)],

$$\sum_{q \leq x^{1/2}} \frac{1}{q} < \log\log x^{1/2} + B - \frac{1}{2(\log x^{1/2})^2} = \log\log x - \log 2 + B - \frac{2}{(\log x)^2}.$$

This proves the lemma in the range $x \geq 286^2$ and direct calculation shows that it holds in the wider range $x \geq 2$. □

We shall find it useful to use the following asymptotically weaker estimates in small cases. The proof follows by checking values of $x \leq 3213$ after which Lemma 9.3.5 is stronger.

**Corollary 9.3.0.1.** *For $x \geq 185$, we have $\sum_{q \in (x^{1/2}, x]} \lfloor x/q \rfloor > 0.5x$. For $x \geq 67$, we have $\sum_{q \in (x^{1/2}, x]} \lfloor x/q \rfloor > 0.45x$.*

Let

$$\theta = 0.3, \quad \lambda = 0.798257.$$

Set $\nu = 1/\theta = 10/3$. For each prime $p$, let

$$S_p = \sum_{\substack{t \in T \\ P^+(t) = p < t^\theta}} \frac{1}{t^\lambda}.$$

We are going to estimate $S_p$ for various small primes $p$. If $t \in T$, $P^+(t) < t^\theta$, then $t \le p^\nu$ implies that $P^+(t) < (p^\nu)^\theta = p$. So, by Lemma 9.2.2, $T$ has at most one member below $3^\nu$, at most 2 members below $5^\nu$, at most 4 members below $7^\nu$, and at most 19 members below $11^\nu$. Since members $t$ of $T$ with $P^+(t) < t^\theta$ have at least $\lceil \nu \rceil = 4$ prime factors, we have

$$\begin{aligned} S_2 &\le \frac{1}{2^{4\lambda}} < 0.109347, \\ S_2 + S_3 &< 0.109347 + \frac{2-1}{3^{\nu\lambda}} < 0.163106, \\ S_2 + S_3 + S_5 &< 0.163106 + \frac{4-2}{5^{\nu\lambda}} < 0.190722, \\ S_2 + S_3 + S_5 + S_7 &< 0.190722 + \frac{19-4}{7^{\nu\lambda}} < 0.275330. \end{aligned} \tag{9.11}$$

Computing $\sum_{p \le Y}(1/p^\lambda - 1/p^\tau)$ directly for $Y = 2, 3, 5, 7$ gives lower bounds

$$0.121399,\ 0.251741,\ 0.368904,\ 0.471733,$$

respectively. Thus $\sum_{p \le Y} S_p < \sum_{p \le Y}(1/p^\lambda - 1/p^\tau)$, so by (9.10) we see Theorem 4 holds when $Y = 2, 3, 5, 7$, respectively.

We now consider primes $p$ with $11 \le p \le 37$. By partial summation,

$$S_p = \int_{p^\nu}^{\infty} \frac{\lambda}{z^{1+\lambda}} N_p(z)\, dz. \tag{9.12}$$

We use Lemmas 9.3.4 and 9.3.5 to get the upper estimates for $N_p(z)$:

$$\begin{aligned} N_p(z) &\le \lfloor z^{1/2} \rfloor - \sum_{p < q \le z^{1/2}} \Big\lfloor \frac{z^{1/2}}{q} \Big\rfloor, \text{ when } p > z^{1/4}, \\ N_p(z) &\le z^{1/2}\Big(1 - \log 2 + \frac{2.5}{\log z} + \frac{10}{(\log z)^2}\Big), \text{ when } p \le z^{1/4}. \end{aligned} \tag{9.13}$$

In the first range, we bound the contribution to $S_p$ by summing over intervals $[m^2, (m+1)^2]$ getting

$$\begin{aligned} \int_{p^\nu}^{p^4} \frac{\lambda}{z^{1+\lambda}} N_p(z)\, dz \le & \Big(\frac{1}{p^{\nu\lambda}} - \frac{1}{\lceil p^{\nu/2} \rceil^{2\lambda}}\Big)\Big(\lfloor p^{\nu/2} \rfloor - \sum_{p < q \le p^{\nu/2}} \lfloor \frac{\lfloor p^{\nu/2} \rfloor}{q} \rfloor\Big) \\ & + \sum_{\lfloor p^{\nu/2} \rfloor < m < p^2} \Big(\frac{1}{m^{2\lambda}} - \frac{1}{(m+1)^{2\lambda}}\Big)\Big(m - \sum_{p < q \le m} \lfloor \frac{m}{q} \rfloor\Big). \end{aligned}$$

For the second range, when $z \ge p^4$, we numerically integrate (9.12) with (9.13) substituted in for $N_p(z)$ when $z \ge 3213^2$, and for smaller values of $z$ we use Corollary 9.3.0.1 to bound the estimate. From these, numerical integration gives the following.

| $p$ | $S_p$ | $\sum_{q\le p} S_q$ | $\sum_{q\le p}(q^{-\lambda}-q^{-\tau})$ |
|---|---|---|---|
| 11 | 0.13259 | 0.40792 | 0.55427 |
| 13 | 0.11241 | 0.52033 | 0.62966 |
| 17 | 0.08382 | 0.60415 | 0.69432 |
| 19 | 0.07601 | 0.68016 | 0.75484 |
| 23 | 0.06194 | 0.74210 | 0.80868 |
| 29 | 0.04757 | 0.78967 | 0.85521 |
| 31 | 0.04501 | 0.83468 | 0.89978 |
| 37 | 0.03680 | 0.87148 | 0.93950 |

Note that the entries in the second and third columns are upper bounds and the entries in the fourth column are lower bounds. The first entry in the third column is found by adding $S_{11}$ to the estimate in (9.11). Since the entries in the fourth column exceed the entries in the third column, we have the theorem for $Y \le 37$.

Now assume that $Y \ge 41$. We have via partial summation that

$$\sum_{P^+(t)<t^\theta} \frac{\mathbf{1}_{t\in T}}{t^\lambda} = \sum_{p\le 7} S_p + \sum_{11\le p\le 23} \int_{p^\nu}^{29^\nu} \frac{\lambda}{z^{1+\lambda}} N_p(z)\,dz + \int_{29^\nu}^{\infty} \frac{\lambda}{z^{1+\lambda}} N(z)\,dz.$$

From (9.11) the $S_p$ terms contribute at most 0.27533. Using Lemmas 9.3.3, 9.3.4, and 9.3.5, and Corollary 9.3.5, we obtain

$$\begin{aligned}\sum_{P^+(t)<t^\theta} \frac{\mathbf{1}_{t\in T}}{t^\lambda} \;<\; & 0.27533 + 0.08455 + 0.06576 + 0.03756 \\ & + 0.02953 + 0.01487 + 0.45614 \;=\; 0.96374.\end{aligned}$$

We also note that $\sum_{p\le 41}\left(\frac{1}{p^\lambda}-\frac{1}{p^\tau}\right) > 0.97661$. Since this estimate exceeds the prior one, this gives the theorem with $\lambda = 0.798257$.

# Chapter 10

# On the critical exponent for $k$-primitive sets

## 10.1 Introduction

A set $A \subset \mathbb{Z}_{>1}$ is *primitive* if no member of $A$ divides another. Erdős [40] showed that for any primitive set $A$,

$$\sum_{n\in A} \frac{1}{n\log n} \;<\; \infty.$$

In fact, his proof bounded the sum uniformly over all primitive sets $A$. Further, in 1986 he asked [46, Conjecture 2.1] if the maximizer is the set of primes $A = \mathcal{P}$. This is now referred to as the Erdős conjecture for primitive sets:

$$\textit{For primitive } A\textit{, we have } \sum_{n\in A} \frac{1}{n\log n} \;\le\; \sum_{p\in \mathcal{P}} \frac{1}{p\log p} \;=\; 1.6366\cdots, \tag{10.1}$$

The current record bound is $\sum_{n\in A} 1/(n\log n) < e^{\gamma} = 1.781\cdots$ due to Pomerance and the author [90]. Here $\gamma$ is the Euler–Mascheroni constant.

A potential approach towards the Erdős conjecture is via integration. Namely, we have

$$\sum_{n\in A} \frac{1}{n\log n} \;=\; \int_1^{\infty} \Big(\sum_{n\in A} \frac{1}{n^{\lambda}}\Big)\, d\lambda,$$

and one might hope the integrand above is dominated by $\sum_p p^{-\lambda}$ for all $\lambda > 1$. Note by a simple argument (see Lemma 10.2.1), if this inequality holds for an exponent $\lambda$, then it will continue to hold for all larger exponents $\lambda' > \lambda$.

However, the primes are not maximal among primitive sets with respect to logarithmic density (i.e., $\lambda = 1$). Indeed, by Erdős [43] and Erdős, Sárközy, and Szemerédi [49],

$$\sup_{\text{primitive } A} \sum_{\substack{n\in A\\ n\le x}} \frac{1}{n} \;=\; \Big(\frac{1}{\sqrt{2\pi}} + o(1)\Big)\frac{\log x}{\sqrt{\log\log x}},$$

where the maximizer is the set of positive integers with $\lfloor \log\log x \rfloor$ prime factors (with multiplicity). By contrast, the primes satisfy

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

Later, Banks and Martin [6] obtained the full characterization that

$$\sum_{\substack{n\in A\\ n\le x}} n^{-\lambda} \;\le\; \sum_{p\le x} p^{-\lambda}, \tag{10.2}$$

for all primitive $A$, $x > 1$, if and only if $\lambda \geq \tau_1 := 1.1403\ldots$, where $\tau = \tau_1$ is the unique real solution to the equation

$$\sum_{p\in\mathbb{P}} p^{-\tau} \;=\; 1 \;+\; \Big(1 - \sum_{p\in\mathbb{P}} p^{-2\tau}\Big)^{1/2}. \tag{10.3}$$

As such we call $\tau_1$ the *critical exponent* for primitive sets.

One may define a hierarchy of primitivity as follows. A 1-primitive set is primitive, and inductively for $k > 1$, a $(k-1)$-primitive set is $k$-primitive if no member divides the product of $k$ distinct other members. That is, a set $A \subset \mathbb{Z}_{>1}$ is *$k$-primitive* if no member of $A$ divides any product of $j$ distinct other members, for any $1 \leq j \leq k$.[1] Note that if (10.2) holds for all $\lambda > \tau$, then it holds for $\lambda = \tau$. Thus, one may similarly consider the critical exponent $\tau_k$ for which (10.2) holds for all $k$-primitive sets if and only if $\lambda \geq \tau_k$. Note that $\tau_j \geq \tau_k$ for $1 \leq j \leq k$.

Recently, Chan, Pomerance, and the author [21] proved $\tau_2 \leq 0.7983$. In particular $\tau_2 < 1$, thereby establishing the Erdős conjecture in the case of 2-primitive sets.

**Theorem 10.1.1** ([21])**.** *For $\lambda \geq 0.7983$, we have*

$$\sum_{\substack{n\in A\\ n\leq x}} n^{-\lambda} \;\leq\; \sum_{p\leq x} p^{-\lambda}$$

*for all 2-primitive sets $A$ and $x \geq 2$. In particular, any 2-primitive set $A$ satisfies*

$$\sum_{n\in A} \frac{1}{n\log n} \;\leq\; \sum_{p} \frac{1}{p\log p}.$$

In 1938, Erdős [42] first studied the maximal cardinality of 2-primitive sets (i.e., $\lambda = 0$). He used Steiner triple systems, though he didn't name them as such. Using more elaborate combinatorial ideas, Chan, Győri, and Sárközy [20] extended the Erdős results to all $k \geq 2$, also see [19] and [110]. Namely, there is an absolute constant $c > 0$ such that

$$\frac{1}{8k^2}\frac{x^{\frac{2}{k+1}}}{(\log x)^2} \;\leq\; \sup_{k\text{-primitive } A} \sum_{\substack{n\in A\\ n\leq x}} 1 \;-\; \sum_{p\leq x} 1 \;\leq\; ck^2\frac{x^{\frac{2}{k+1}}}{(\log x)^2}, \tag{10.4}$$

[1] In [20], $A$ is called $k$-primitive if no member of $A$ divides any product of $k$ distinct other members. These definitions only differ when $|A| \leq k$, and do not affect the critical exponent $\tau_k$, by Lemma 10.2.4 below.

for $x$ sufficiently large. Here the lower bound is attained by some set $A''$ consisting of the primes in $(x^{1/(k+1)}, x]$ and a size $x^{2/(k+1)}/8(k\log x)^2$ subset of products of $k+1$ primes in $(1, x^{1/(k+1)}]$. In particular, the lower bound in (10.4) implies

$$\sum_{\substack{n\in A''\\ n\le x}} n^{-\lambda} \ge \sum_{x^{1/(k+1)}<p\le x} p^{-\lambda} + \frac{1}{x^\lambda}\frac{x^{2/(k+1)}}{8(k\log x)^2} > \sum_{x^{1/(k+1)}<p\le x} p^{-\lambda} + \sum_{p\le x^{1/(k+1)}} p^{-\lambda},$$

when $\lambda < 1/k$ and $x$ is sufficiently large. Hence we quickly deduce $\tau_k \ge 1/k$.

Thus combining with Theorem 10.1.1, the critical exponent for 2-primitive sets lies in the interval

$$\tau_2 \ \in \ [0.5,\, 0.7983]. \tag{10.5}$$

It is an open question to determine the exact value of $\tau_2$, and perhaps characterize $\tau_2$ as a solution to some functional equation, as with (10.2) for $\tau_1$.

In light of this, it is natural to ask about the behavior of the decreasing sequence $\tau_1 \ge \tau_2 \ge \tau_3 \ge \cdots$, in particular whether $\tau_k$ tends to zero as $k \to \infty$. The main result of this chapter is to answer in the affirmative. Using some of the ideas in Chapter 9 we prove the following quantitative result.

**Theorem 10.1.2.** *Let $p_k$ denote the $k$th prime number. For any $k \ge 1$ and $\lambda \ge 1.5/\log p_k$, we have*

$$\sum_{\substack{n\le x\\ n\in A}} n^{-\lambda} \ \le \ \sum_{p\le x} p^{-\lambda} \tag{10.6}$$

*for all $k$-primitive sets $A$ and $x \ge 2$.*

Thus, for $k \ge 1$,

$$\frac{1}{k} \ \le \ \tau_k \ \le \ \frac{1.5}{\log p_k}. \tag{10.7}$$

Clearly the upper and lower bounds differ substantially, and we offer it as an problem to narrow this gap.

### 10.1.1 Generalizations

Upon closer inspection of the proofs in [19], [20], we observe the lower bound for (10.4) holds under a stronger notion of $k$-primitivity, namely, one forbids a member from dividing the product of $k$ other members, *not necessarily distinct*. Similarly, the

upper bound in (10.4) holds even if one relaxes to only forbid a member from dividing the *least common multiple* (lcm) of $k$ other members.

Hence this naturally suggests the following generalizations. We say a set $A \subset \mathbb{Z}_{>1}$ is "strongly $k$-primitive" if no member divides the product of $k$ other members which are not necessarily distinct. Any strongly $k$-primitive set is $k$-primitive, but not vice versa. For example, $A = \{4, 5, 6\}$ is 2-primitive but not strongly 2-primitive. In the other direction, we say a set $A \subset \mathbb{Z}_{>1}$ is "lcm $k$-primitive" if no member divides the lcm of $k$ other members. Here, every $k$-primitive set is lcm $k$-primitive, but not vice versa. An example is $A = \{4, 6, 10\}$ which is lcm 2-primitive, but not 2-primitive.

One can ask for critical exponents in the strong case and in the lcm case. Denote the former by $\tau_k^{(\mathrm{s})}$ and the latter by $\tau_k^{(\mathrm{lcm})}$. By the above comments, for each $k \geq 2$ we have

$$\frac{1}{k} \ \leq \ \tau_k^{(\mathrm{s})} \ \leq \ \tau_k \ \leq \ \tau_k^{(\mathrm{lcm})}.$$

From these definitions, two natural questions arise: Is there a better upper bound for $\tau_k^{(\mathrm{s})}$ than that afforded by Theorem 10.1.2? Is there an upper bound for $\tau_k^{(\mathrm{lcm})}$ that is $o(1)$ as $k \to \infty$?

We make progress on these two questions by proving the following two theorems.

**Theorem 10.1.3.** *For any $k \geq 1$, $\tau_k^{(\mathrm{lcm})} \leq 1.7/\log p_k$. In addition, $\tau_2^{(\mathrm{lcm})} \leq 1$, so the Erdős conjecture is true for* lcm 2*-primitive sets.*

For the $\tau_k^{(\mathrm{s})}$ case we prove a considerably stronger inequality.

**Theorem 10.1.4.** *For $k \geq 2$ we have $\tau_k^{(\mathrm{s})} \leq (3 \log k)/k$.*

Thus,

$$\frac{1}{k} \ \leq \ \tau_k^{(\mathrm{s})} \ \leq \ \frac{3 \log k}{k}$$

for all $k \geq 2$. It would be nice to so sharpen the inequalities for $\tau_k$ and $\tau_k^{(\mathrm{lcm})}$.

### 10.1.2 $\infty$-primitive sets

Finally, we offer a natural interpretation of the result $\tau_k \to 0$. We say a set $A$ is *$\infty$-primitive* if it is $k$-primitive for every $k \geq 1$. For example, the set of primes forms an $\infty$-primitive set, as does any set of pairwise coprime integers. However, the set $\{6, 10\}$ indicates that coprimality is not neccessary.

We prove that the primes are maximal among all $\infty$-primitive sets for the full range of exponents $\lambda \geq 0$.

**Corollary 10.1.0.1.** *For all $x \geq 2$, we have*

$$\sup_{\infty\text{-primitive } A} \sum_{\substack{n \in A \\ n \leq x}} 1 \ \leq \ \sum_{p \leq x} 1.$$

*Proof.* We proceed by contradiction. Suppose there exist $x \geq 2$ and an $\infty$-primitive set $T$ such that

$$\sum_{\substack{n \in T \\ n \leq x}} 1 \ > \ \sum_{p \leq x} 1.$$

Consider the sequence $a_n = \mathbf{1}_{n \leq x}(\mathbf{1}_{n \in T} - \mathbf{1}_{n \in \mathcal{P}})$ and define the Dirichlet series $F(t) = \sum_n a_n n^{-t}$. By assumption $F(0) > 0$ so by continuity of $F$ there exists $t > 0$ sufficiently small for which $F(t) > 0$. That is,

$$\sum_{\substack{n \in T \\ n \leq x}} n^{-t} \ > \ \sum_{p \leq x} p^{-t}. \tag{10.8}$$

Since $\tau_k \to 0$ by Theorem 10.1.2, we have $\tau_j < t$ for $j$ sufficiently large. But since $T$ is $\infty$-primitive, hence $j$-primitive, we see (10.8) contradicts the definition of $\tau_j$. Thus the corollary follows. □

## 10.2 Preliminary lemmas

**Lemma 10.2.1.** *Take sets $A, B \subset \mathbb{R}_{>1}$. Suppose $\lambda \geq 0$ satisfies $I_\lambda(x) \geq 0$ for all $x > 1$, where*

$$I_\lambda(x) := \sum_{\substack{a \in A \\ a \leq x}} a^{-\lambda} \ - \ \sum_{\substack{b \in B \\ b \leq x}} b^{-\lambda}.$$

*Then $I_{\lambda'}(x) \geq 0$ for all $\lambda' \geq \lambda$, $x > 1$.*

*Proof.* By partial summation,

$$I_{\lambda'}(x) \ = \ x^{\lambda - \lambda'} I_\lambda(x) \ + \ (\lambda' - \lambda) \int_1^x u^{\lambda - \lambda' - 1} I_\lambda(u) \, du.$$

Hence if $I_{\lambda_k}(x) \geq 0$ for all $x > 1$, it then follows $I_{\lambda'}(x) \geq 0$ for all $\lambda' \geq \lambda$ as claimed. □

**Lemma 10.2.2.** *Let*

$$\lambda_1 = 1.2, \quad \lambda_2 = 0.8, \quad \text{and} \ \ \lambda_k = 2.625 \prod_{i=1}^{k} \Big(1 - \frac{1}{p_i}\Big) \ \ \text{for} \ \ k \ge 3.$$

*Then*

$$\lambda_k > \frac{1.45}{\log p_k} \ \ \text{for} \ \ k \ge 62, \qquad \lambda_k < \frac{1.5}{\log p_k} \ \ \text{for} \ \ k \ge 1.$$

*In addition, let*

$$\mu_1 = 8/7 \quad \text{and} \ \ \mu_k = 3 \prod_{i=1}^{k} \Big(1 - \frac{1}{p_i}\Big) \ \ \text{for} \ \ k \ge 2.$$

*Then*

$$\mu_k > \frac{1.65}{\log p_k} \ \ \text{for} \ \ k \ge 47, \qquad \mu_k < \frac{1.7}{\log p_k} \ \ \text{for} \ \ k \ge 1.$$

*Proof.* One can verify the lemma for $p_k \le 2{,}000$ by direct computation. For larger $p_k$ we use (3.25) of Rosser and Schoenfeld [124] with the Euler–Mascheroni constant $\gamma = 0.57721...$, getting

$$\lambda_k \ge \frac{2.625 e^{-\gamma}}{\log p_k}\Big(1 - \frac{1}{2\log^2 p_k}\Big) \ge \frac{2.625 e^{-0.57722}}{\log p_k}\Big(1 - \frac{1}{2\log^2 2{,}000}\Big) \ge \frac{1.45}{\log p_k},$$

which gives the lower bound for $\lambda_k$. The lower bound for $\mu_k$ follows in the same way. For the upper bound, by (3.26) of Rosser and Schoenfeld [124] we have

$$\lambda_k < \frac{2.625 e^{-0.57721}}{\log p_k}\Big(1 + \frac{1}{2\log^2 2{,}000}\Big) < \frac{1.5}{\log p_k}.$$

Again, the upper bound for $\mu_k$ follows in the same way. This completes the proof. □

**Lemma 10.2.3.** *For* $0 < \lambda < 1$ *and* $x \ge 41$,

$$x^{1-\lambda}\Big(1 - \frac{1}{\log x}\Big) \le \sum_{p \le x} \frac{\log p}{p^\lambda} \le \frac{1.01624}{1-\lambda} x^{1-\lambda}.$$

*Proof.* By partial summation,

$$\sum_{p \le x} \frac{\log p}{p^\lambda} = \int_{2^-}^{x} \frac{d\theta(u)}{u^\lambda} = \frac{\theta(x)}{x^\lambda} + \lambda \int_2^x \frac{\theta(u)}{u^{\lambda+1}} du$$

where $\theta(x) = \sum_{p \le x} \log p$. The lemma follows from (3.16) and (3.32) in Rosser and Schoenfeld

$$x\Big(1 - \frac{1}{\log x}\Big) < \theta(x) \ \ \text{for} \ \ x \ge 41$$

and

$$\theta(x) < 1.01624x \ \ \text{for} \ \ x > 0. \qquad \square$$

For a set $A$ of integers, let $\mathcal{P}(A)$ denote the set of primes that divide some member of $A$.

**Lemma 10.2.4.** *Let $A$ be an* lcm *$k$-primitive set with $k \geq 2$. If $|\mathcal{P}(A)| \leq k$, then $|A| \leq |\mathcal{P}(A)|$ and for all $\lambda \geq 0$,*

$$\sum_{n \in A} n^{-\lambda} \ \leq \ \sum_{p \in \mathcal{P}(A)} p^{-\lambda}.$$

*Also, if $k < |\mathcal{P}(A)| < 2k$, then $|A| \leq |\mathcal{P}(A)| + 1$.*

*Proof.* Let $v_p(n)$ denote the exponent on $p$ in the prime factorization of $n$, so that $p^{v_p(n)} \parallel n$. For each $p \in \mathcal{P}(A)$ let $n_p$ be the element $n \in A$ with $v_p(n)$ maximal (breaking ties arbitrarily), and let $A^* = \{n_p : p \in \mathcal{P}(A)\}$. Thus $|A^*| \leq |\mathcal{P}(A)|$.

Suppose $|\mathcal{P}(A)| \leq k$. Then any $n \in A \setminus A^*$ would satisfy $n \mid \operatorname{lcm}(A^*)$, contradicting $A$ as lcm $k$-primitive. Thus, $A^* = A$ and $|A| \leq |\mathcal{P}(A)| \leq k$. Next, $|A| \leq k$ implies each $n \in A$ has $n \nmid \operatorname{lcm}(A \setminus \{n\})$. Thus, each $n \in A$ has a prime factor $p$ with $v_p(n) > v_p(m)$ for all $m \in A \setminus \{n\}$, so the map, call it $f$, where $f(n) = p \mid n$ is injective on $A$. Hence we conclude

$$\sum_{n \in A} n^{-\lambda} \ \leq \ \sum_{n \in A} f(n)^{-\lambda} \ \leq \ \sum_{p \in \mathcal{P}(A)} p^{-\lambda}.$$

Also, suppose $N = |\mathcal{P}(A)|$, $k < N < 2k$, and there exist distinct $n, n' \in A \setminus A^*$. Without loss, the subset $P = \{p \in \mathcal{P}(A) : v_p(n) \geq v_p(n')\}$ contains at least half of the primes in $\mathcal{P}(A)$, i.e., $|P| \geq \lceil \frac{N}{2} \rceil$. Hence

$$n' \ \mid \ \operatorname{lcm}(\{n_p : p \notin P\} \cup \{n\}),$$

which is an lcm of $1 + N - \lceil N/2 \rceil$ elements. Clearly this number is at most $k$, thus contradicting $A$ as lcm $k$-primitive. This implies $|A| \leq N + 1$. □

## 10.3 Theorem for $k$-primitive sets

In this section we prove Theorem 10.1.2. Recall the numbers $\lambda_k$ in Lemma 10.2.2. By that lemma it suffices to prove the following theorem.

**Theorem 10.3.1.** *Let $A$ be a $k$-primitive set. For each $k \geq 1$ we have*

$$\sum_{\substack{a \in A \\ a \leq x}} a^{-\lambda_k} \ \leq \ \sum_{\substack{p \in \mathcal{P}(A) \\ p \leq x}} p^{-\lambda_k}. \tag{10.9}$$

*for any $x > 1$.*

Since $\lambda_1 = 1.2, \lambda_2 = 0.8$, the theorem holds for $k = 1, 2$, so we may assume that $k \geq 3$ and that the theorem holds for $(k-1)$-primitive sets.

We partition $A$ into primes $S$ and composites $T$. Note by primitivity, the primes in $S$ and $\mathcal{P}(T)$ are disjoint. We thus may cancel the contribution of $p \in S$ from both sides of (10.9) and so reduce Theorem 10.3.1 to the case $A = T$ where every member is composite.

For a prime $p$, let $T_p = \{t \in T : p \mid t\}$. We may assume that

$$\sum_{t \in T_p} t^{-\lambda} \ > \ p^{-\lambda} \quad \text{for all} \quad p \in \mathcal{P}(T), \tag{10.10}$$

since if this fails for some $p$, the theorem for $T \setminus T_p$ implies the theorem for $T$. An immediate consequence is that

$$|T_p| \geq 2 \quad \text{for all} \quad p \in \mathcal{P}(T). \tag{10.11}$$

Further, it suffices to assume that $\mathcal{P}(T)$ consists of an initial list of primes, say

$$\mathcal{P}(T) = \mathcal{P} \cap (1, Y] \quad \text{for some} \quad Y \geq 2. \tag{10.12}$$

Indeed, if not, suppose $q$ is the smallest prime outside $\mathcal{P}(T)$, and let $p \in \mathcal{P}(T)$ be the smallest prime with $p > q$. Then by (10.10),

$$0 \ < \ (p/q)^{\lambda} \Big( \sum_{t \in T_p} t^{-\lambda} \ - \ p^{-\lambda} \Big) \ \leq \ \sum_{t' \in T'_q} (t')^{-\lambda} \ - \ q^{-\lambda},$$

where $T'$ is the ($k$-primitive) image of $T$ under the automorphism of $\mathbb{N}$ induced by swapping $q \leftrightarrow p$. Hence the proof for $T$ will follow from that of $T'$.

For an integer $t > 1$ let $Q(t)$ denote the largest prime power factor of $t$, which is possibly a prime to the first power. We first handle those $t \in T$ with $Q(t) < t^{\theta}$ for an appropriate choice of $\theta$.

**Lemma 10.3.1.** *Let $k \geq 2$ and let $0 < \theta \leq 1/k$. Suppose $T$ is* lcm *$k$-primitive with $Q(t) < t^{\theta}$ for each $t \in T$. Let $z \geq 2$, and let $N(z)$ be the number of members of $T$ up to $z$. Then*

$$N(z) \ \leq \ z^{\frac{1}{k} + \theta}.$$

*Proof.* If $t \leq z^{1/k}$, let $m_1(t) = t$. Now suppose that $t > z^{1/k}$ and decompose $t = q_1 q_2 \cdots q_r$ into its prime powers $q_1 > \cdots > q_r$. By assumption, $q_1 < t^{\theta}$. Consider $q_1 \cdots q_j \leq z^{1/k}$ with $j$ maximal. Then $m_1(t) := q_1 \cdots q_{j+1}$ lies in the interval $(z^{1/k}, z^{1/k+\theta}]$. In this way we may split $t$ into $l_t \leq k$ pairwise coprime factors

$$t = q_1 q_2 \cdots q_r = m_1(t) \cdots m_{l_t}(t) \tag{10.13}$$

with each $m_i(t) \le z^{1/k+\theta}$.

Now observe each $t \in T$ has some factor $m_i(t)$ which is distinct from all other factors $m_j(s)$, $s \in T \setminus \{t\}$. Indeed, if not, then each factor of $t$ has $m_i(t) = m_{j_i}(t_i)$ for some $t_i \in T \setminus \{t\}$ (not neccessarily distinct). And since the factors $m_i(t)$ are pairwise coprime,

$$t = m_1(t) \cdots m_{l_t}(t) \mid \operatorname{lcm}[t_1, \dots, t_{l_t}],$$

contradicting $T$ as lcm $k$-primitive.

Hence we have a one-to-one map $g : T \to \mathbb{N}$ via $g(t) = m_i(t)$. And since $m_i(t) \le z^{\frac{1}{k}+\theta}$, we conclude $|T| = |g(T)| \le z^{\frac{1}{k}+\theta}$. □

We now fix a choice for $\theta = \theta_k$. Let

$$\theta_k = \frac{1}{p_k} \text{ for } k \neq 3 \text{ and } \theta_3 = \frac{1}{8}.$$

Further, let $\nu_k = 1/\theta_k$, so that

$$\nu_k = p_k \text{ for } k \neq 3 \text{ and } \nu_3 = 8.$$

With these choices we have

$$\lambda_k = 2.4 \prod_{j \le k} (1 - \theta_j).$$

Note that if $Q(t) < t^{\theta_k}$, then $t$ must have at least $\nu_k + 1$ distinct prime factors. Let $P^+(t)$ denote the largest prime dividing $t$, so that

$$p_{\nu_k+1} \le P^+(t) \le Q(t) < t^{\theta_k} \text{ which implies } t > p_{\nu_k+1}^{\nu_k}.$$

Thus, with $\theta = \theta_k$, $\nu = \nu_k$, and $\lambda > \frac{1}{k} + \theta$,

$$\sum_{\substack{t \in T \\ Q(t) < t^\theta}} \frac{1}{t^\lambda} = \int_{p_{\nu+1}^\nu}^\infty \frac{\lambda}{z^{1+\lambda}} N(z)\, dz \le \frac{\lambda}{\lambda - \frac{1}{k} - \theta} p_{\nu+1}^{-\nu(\lambda - \frac{1}{k} - \theta)}, \tag{10.14}$$

by partial summation and Lemma 10.3.1.

**Lemma 10.3.2.** *Let $k \ge 2$ and let $T$ be an* lcm *$k$-primitive set of composite numbers. Decompose $T = T' \cup T''$, where $t \in T''$ if there exists another $s \in T$ with $Q(t) \mid s$; else $t \in T'$. Define the map $f : T \to \mathbb{N}$ via*

$$f(t) = \begin{cases} Q(t) & t \in T' \\ t/Q(t) & t \in T'' \end{cases}$$

*Then $f$ is one-to-one and $f(T)$ is an* lcm *$(k-1)$-primitive set. Further, the members of $f(T')$ are pairwise coprime proper prime powers.*

*Proof.* First, the map $f$ is one-to-one. Indeed, suppose $f(t) = f(t')$ for some $t, t' \in T$. If $t \in T'$ then $Q(t) \nmid t'$, in particular $f(t) = Q(t) \neq f(t') \in \{Q(t'), t'/Q(t')\}$. Similarly, if $t \in T''$ then $Q(t) \mid s$ for some $s \in T \setminus \{t\}$. Thus $1 = \gcd(Q(t), t/Q(t)) = \gcd(Q(t), t'/Q(t'))$ implies

$$t = Q(t) \cdot \frac{t}{Q(t)} = Q(t) \cdot \frac{t'}{Q(t')} \;\Big|\; \operatorname{lcm}[s, t'].$$

Thus lcm 2-primitivity of $T$ forces $t = t'$. Hence $f$ is indeed one-to-one.

Next suppose $f(T)$ is not lcm $(k-1)$-primitive. Then there exist $t \in T$ and $t_1, .., t_{k-1} \in T \setminus \{t\}$ such that

$$f(t) \;\Big|\; \operatorname{lcm}[f(t_1), \ldots, f(t_{k-1})].$$

If $t \in T'$ then $f(t) = Q(t)$ is a prime power, so by above $Q(t) \mid f(t_i)$ for some index $i$. Thus $Q(t) \mid t_i \in T \setminus \{t\}$, which contradicts $t \in T'$.

Similarly if $t \in T''$, then $Q(t) \mid s$ for some $s \in T \setminus \{t\}$, and so $1 = \gcd(Q(t), t/Q(t))$ gives

$$t = Q(t) \cdot \frac{t}{Q(t)} = Q(t)f(t) \;\Big|\; \operatorname{lcm}[s, t_1, \ldots, t_{k-1}]$$

contradicting $T$ as lcm $k$-primitive. Hence $f(T)$ is indeed lcm $(k-1)$-primitive. That the members of $f(T')$ are pairwise coprime follows from $f(T')$ being a primitive set of prime powers. That the members of $f(T')$ are proper prime powers follows from the fact that if $Q(t)$ is prime, then by (10.11), $T_{Q(t)}$ has at least 2 elements, and so $t \in T''$. □

Let $T_\theta = \{t \in T : Q(t) \geq t^\theta\}$. We apply Lemma 10.3.2 to $T = T_\theta$. Thus, by the induction hypothesis on the lcm $(k-1)$-primitive set $f(T_\theta)$, for $\lambda' := \lambda_{k-1} = \frac{\lambda_k}{1-\theta}$ we have

$$\sum_{t \in T_\theta} f(t)^{-\lambda'} \;=\; \sum_{t \in T'} Q(t)^{-\lambda'} + \sum_{t \in T''} \big(t/Q(t)\big)^{-\lambda'} \;=\; \sum_{d \in f(T_\theta)} d^{-\lambda'} \;\leq\; \sum_{p \leq Y} p^{-\lambda'}.$$

Now if $Q(t) \geq t^\theta$, then $t/Q(t) \leq t^{(1-\theta)}$ so that $t^{-\lambda} \leq (t/Q(t))^{-\lambda/(1-\theta)} = (t/Q(t))^{-\lambda'}$. Thus by the above,

$$\begin{aligned}\sum_{t \in T_\theta} t^{-\lambda} = \sum_{t \in T'} t^{-\lambda} + \sum_{t \in T''} t^{-\lambda} &\leq \sum_{t \in T'} Q(t)^{-\lambda} + \sum_{t \in T''} \big(t/Q(t)\big)^{-\lambda/(1-\theta)} \\ &\leq \sum_{t \in T'} \big(Q(t)^{-\lambda} - Q(t)^{-\lambda'}\big) \;+\; \sum_{p \leq Y} p^{-\lambda'}.\end{aligned}$$

Thus,

$$\sum_{t\in T_\theta} t^{-\lambda} - \sum_{p\le Y} p^{-\lambda} < \sum_{p\le Y} \big((p^{-2\lambda} - p^{-2\lambda'}) - (p^{-\lambda} - p^{-\lambda'})\big) =: S(Y), \tag{10.15}$$

using that $f(T')$ is a set of pairwise coprime proper prime powers and $\mathcal{P}(T) \subset [1, Y]$. Note that from Lemma 10.2.4 we may assume that $Y \ge p_k$.

*Claim 1:* The sequence $S(p_j)$ for $j \ge k$ is decreasing, so if $S(p_k) < 0$, then $S(Y) < 0$ for all $Y \ge p_k$.

Indeed, the terms in $S(Y)$ are of the form $h(y,z) = y - z - (y^2 - z^2)$, where $y = p^{-\lambda'}$ and $z = p^{-\lambda}$. Note that $h(y,z) = (y-z)(1-(y+z))$ and we have $0 < y < z$. Further, $p^{-\lambda} \le \frac{1}{3}$ for $p \ge p_k$ and $k \ge 3$, which follows from Lemma 10.2.2 and a short calculation. Thus, for $p \ge p_k$, the terms in $S(Y)$ are negative, establishing Claim 1.

*Claim 2:* For $k \ge 3$ we have $S(p_k) < 0$ and for $k \ge 200$ we have $S(p_k) < -0.015/\log p_k$.

We verify this directly for $3 \le k \le 199$, so assume now that $k \ge 200$. Let $F(\lambda) = \sum_{p\le p_k}(p^{-2\lambda} - p^{-\lambda})$ so that $S(p_k) = F(\lambda) - F(\lambda')$. By the mean value theorem, there exists some $\xi \in (\lambda, \lambda')$ with

$$\begin{aligned} F(\lambda) - F(\lambda') &= (\lambda - \lambda')F'(\xi) = (\lambda - \lambda') \sum_{p\le p_k} \big(p^{-\xi}\log p - 2p^{-2\xi}\log p\big) \\ &= -\theta\lambda' \sum_{p\le p_k} \big(p^{-\xi} - 2p^{-2\xi}\big)\log p \;<\; -\theta\lambda' \sum_{p\le p_k} \big(p^{-\lambda'} - 2p^{-2\lambda}\big)\log p. \end{aligned}$$

Recall that $\theta = \theta_k$, $\lambda = \lambda_k$, and $\lambda' = \lambda_{k-1}$. Using Lemma 10.2.3, we thus have

$$\begin{aligned} S(p_k) = F(\lambda) - F(\lambda') &< -\theta\lambda'\Big(p_k^{1-\lambda'}\big(1 - \frac{1}{\log p_k}\big) - \frac{2.03248}{1-2\lambda}p_k^{1-2\lambda}\Big) \\ &= -\lambda' p_k^{-\lambda}\Big(p_k^{\lambda-\lambda'}\Big(1 - \frac{1}{\log p_k}\Big) - \frac{2.03248}{1-2\lambda}p_k^{-\lambda}\Big). \end{aligned}$$

We use $1 - 2\lambda > 0.587$, $p_k^{\lambda-\lambda'} > 1 - 1/p_k$, and $e^{-1.5} < p_k^{-\lambda} < e^{-1.45}$, which follows from Lemma 10.2.2, to get

$$S(p_k) < -\frac{0.015}{\log p_k}, \qquad \text{for } \; k \ge 200, \tag{10.16}$$

completing the proof of Claim 2.

By (10.14) and (10.15),

$$I_\lambda = \sum_{\substack{t\in T\\ Q(t)<t^\theta}} t^{-\lambda} + \sum_{\substack{t\in T\\ Q(t)\ge t^\theta}} t^{-\lambda} \;-\; \sum_{p\le Y} p^{-\lambda} < \; \frac{\lambda}{\lambda - \frac{1}{k} - \theta}\, p_{\nu+1}^{-\nu(\lambda - \frac{1}{k} - \theta)} + S(Y). \tag{10.17}$$

Note though that if $Y < p_{\nu+1}$, then the first term does not appear, so Claims 1 and 2 prove that $I_\lambda < 0$. So, assume that $Y = p_{\nu+1}$ in (10.14). We check numerically that $I_\lambda < 0$ for $3 \le k \le 199$.

It remains to show that $I_\lambda < 0$ for $k \ge 200$. Note that if $k \ge 200$, then

$$\lambda - \frac{1}{k} - \theta > \frac{1.4}{\log p_k}, \quad \frac{\lambda}{\lambda - \frac{1}{k} - \theta} < 1.05,$$

using Lemma 10.2.3. Thus,

$$\frac{\lambda}{\lambda - \frac{1}{k} - \theta}\, p_{\nu+1}^{-\nu(\lambda - \frac{1}{k} - \theta)} \;<\; 1.05 p_{p_k+1}^{-1.4p_k/\log p_k} \;<\; 1.05 p_k^{-1.4p_k/\log p_k} \;=\; 1.05 e^{-1.4p_k}.$$

As a function of $p_k$ this expression is much smaller than $0.015/\log p_k$, in fact, this is so for $p_k \ge 5$. Thus, (10.16) shows that $I_\lambda < 0$ for $k \ge 200$. This completes the proof.

## 10.4 Theorem for lcm $k$-primitive sets

In this section we prove Theorem 10.1.3. The proof largely follows from the proof for $k$-primitive sets in the previous section. In fact, the only difference is that we start the induction at $k = 2$ rather than $k = 3$. By Lemma 10.2.2 it suffices to prove the following theorem.

**Theorem 10.4.1.** *Recall the numbers $\mu_k$ in Lemma 10.2.2. Let $A$ be an* lcm *$k$-primitive set. For each $k \ge 1$ we have*

$$\sum_{\substack{a \in A \\ a \le x}} a^{-\mu_k} \;\le\; \sum_{\substack{p \in \mathcal{P}(A) \\ p \le x}} p^{-\mu_k}. \tag{10.18}$$

*for any $x > 1$.*

First note that since $\tau_1 < 8/7 = \mu_1$, the theorem holds at $k = 1$, so we may assume that $k \ge 2$ and the theorem holds for lcm $(k-1)$-primitive sets.

Next note that the various reductions we made in Section 10.3 hold here, as well as Lemmas 10.3.1 and 10.3.2. Here we have

$$\theta_k = 1/p_k \;\text{ for }\; k \ne 2, \quad \theta_2 = 1/8,$$

so that for all $k \ge 1$,

$$\mu_k = \frac{16}{7} \prod_{j \le k} (1 - \theta_k).$$

Let $\nu_k = 1/\theta_k$, so that $\nu_k = p_k$ for $k \neq 2$ and $\nu_2 = 8$. With these new values, we continue to have (10.14) recorded anew as follows:

$$\sum_{\substack{t\in T\\ Q(t)<t^\theta}} \frac{1}{t^\mu} \;=\; \int_{p_{\nu+1}^\nu}^\infty \frac{\mu}{z^{1+\mu}} N(z)\,dz \;\le\; \frac{\mu}{\mu-\frac{1}{k}-\theta}\, p_{\nu+1}^{-\nu(\mu-\frac{1}{k}-\theta)}, \tag{10.19}$$

where $\mu = \mu_k$, $\theta = \theta_k$, $\nu = \nu_k$.

We have the analogue of (10.15), where $\lambda$ is replaced with $\mu = \mu_k$ and $\lambda'$ is replaced with $\mu' = \mu_{k-1}$. In addition, we continue to have Claim 1, checking that $p^{-\mu} \le \frac{1}{3}$ for $p \ge p_k$.

However, Claim 2 needs to be verified. As before, we check that $S(p_k) < 0$ for $2 \le k \le 199$. Following the argument for $k \ge 200$, we have $1 - 2\mu > 0.528$, $p_k^{\mu-\mu'} > 1 - 1/p_k$, and $e^{-1.7} < p_k^{-\lambda} < e^{-1.65}$, again following from Lemma 10.2.3. Thus,

$$\begin{aligned} S(p_k) &< -\frac{1.65e^{-1.7}}{\log p_k}\Big(\Big(1-\frac{1}{p_k}\Big)\Big(1-\frac{1}{\log p_k}\Big) - \frac{2.03248}{0.528}e^{-1.65}\Big) \\ &< -\frac{0.035}{\log p_k}, \qquad \text{for } \; k \ge 200. \end{aligned}$$

This is somewhat stronger than Claim 2.

We have the analogue of (10.17):

$$I_\mu \;<\; \frac{\mu}{\mu-\frac{1}{k}-\theta}\, p_{\nu+1}^{-\nu(\mu-\frac{1}{k}-\theta)} + S(Y), \tag{10.20}$$

where the first term does not occur if $Y < p_{\nu+1}$. Our goal is to show that $I_\mu < 0$. Thus, by Claims 1 and 2, we may assume that $Y = p_{\nu+1}$. We then check numerically that the bound in (10.20) is negative for $2 \le k \le 199$.

To show that $I_\mu < 0$ for $k \ge 200$, note that

$$\mu - \frac{1}{k} - \theta > \frac{1.6}{\log p_k}, \quad \frac{\mu}{\mu-\frac{1}{k}-\theta} < 1.05$$

in analogy to what we had before. Thus,

$$\frac{\mu}{\mu-\frac{1}{k}-\theta}\, p_{\nu+1}^{-\nu(\mu-\frac{1}{k}-\theta)} < 1.05e^{-1.6p_k},$$

which is again smaller than $0.015/\log p_k$. Hence $I_\mu < 0$ for $k \ge 200$, which completes the proof.

## 10.5 Theorem for strongly $k$-primitive sets

In this section we prove Theorem 10.1.4.

As in Section 10.3 we may assume that $A = T$ consists of composite numbers, for each $p \in \mathcal{P}(T)$ we have $|T_p| \geq 2$, and $\mathcal{P}(T)$ consists of all of the primes up to some point $Y$. Note that since $\tau_k^{(\mathrm{s})} \leq \tau_k$ for all $k$, Theorem 10.1.4 follows from Theorem 10.3.1 when $k \leq 38$. Thus, in the sequel, we assume that $k \geq 39$ and that the theorem holds for $k-1$.

For $k \geq 39$, let

$$\lambda = \lambda_k = \frac{3\log k}{k}, \quad \theta = \theta_k = 1 - \frac{\lambda_k}{\lambda_{k-1}}.$$

A simple calculation shows that

$$\nu = \nu_k := \frac{1}{\theta_k} > \frac{k\log(k-1)}{\log(k-1)-1} > k.$$

Recall that $P^+(t)$ denotes the largest prime factor of $t$. Let

$$T_0 = \{t \in T : P^+(t) < t^\theta\}.$$

We now prove a version of Lemma 10.3.1 dealing with $T_0$.

**Lemma 10.5.1.** *Let $N_0(z)$ denote the number of $t \in T_0$ with $t \leq z$. Then*

$$N_0(z) \ \leq \ z^{\frac{1}{k}+\theta}.$$

*Proof.* Let $t \in T_0$, $t \leq z$. If $t \leq z^{1/k}$, let $m_1(t) = t$. Otherwise, say the prime factorization of $t$ is $p_1p_2\cdots p_r$, where $p_1 \geq p_2 \geq \cdots \geq p_r$. Let $j$ be minimal with $p_1\cdots p_j > z^{1/k}$. Since all of these primes are at most $t^\theta \leq z^\theta$, we have $m_1(t) := p_1\cdots p_j \leq z^{\frac{1}{k}+\theta}$. Continuing in this fashion we obtain a factorization

$$t = p_1p_2\cdots p_r = m_1(t)m_2(t)\cdots m_{l_t}(t), \quad l_t \leq k, \ \text{ each } m_i(t) \leq z^{\frac{1}{k}+\theta}.$$

We claim that each $t$ has at least one factor $m_i(t)$ that does not appear in the analogous factorization for any other $t' \in T_0$. Indeed, if each $m_i(t) = m_{j_i}(t'_i)$ for some $t'_i \in T_0 \setminus \{t\}$ with $j_i \leq l_{t'_i}$, then $t \mid t'_1t'_2\cdots t'_{l_t}$, contradicting $T_0$ as strongly $k$-primitive. By mapping $t$ to such a unique factor $m_i(t)$ we obtain a one-to-one function from $T_0$ to the integers in $(1, z^{\frac{1}{k}+\theta}]$, so proving the lemma. □

Because of the change in the definition of $N(z)$ we do not have (10.14). Instead, we argue as follows. Note that every member of $T_0$ has at least $\lceil\nu\rceil$ prime factors, counted with multiplicity. Thus, the least element of $T_0$ is at least $2^\nu$. In addition, the

second smallest member of $T_0$ must be at least $3^\nu$. Indeed, if there are two members smaller than this, then $P^+(t) < t^\theta$ implies they are both powers of 2, and hence $T_0$ is not primitive. More generally, using Lemma 10.2.4, $T_0$ has at most $j$ members smaller than $p_{j+1}^{\nu_k}$ for each $j \le k$. Thus,

$$\begin{aligned}\sum_{t\in T_0}\frac{1}{t^\lambda} &< \sum_{j\le k}\frac{1}{p_j^{\nu\lambda}} + \int_{p_{k+1}^\nu}^\infty \frac{\lambda}{z^{1+\lambda}}N_0(z)\,dz \\ &< \sum_{j\le k}\frac{1}{p_j^{\nu\lambda}} + \frac{\lambda}{\lambda-\frac{1}{k}-\theta}\,p_{k+1}^{-\nu(\lambda-\frac{1}{k}-\theta)} \end{aligned} \tag{10.21}$$

by partial summation and Lemma 10.5.1.

In the next lemma we give a variant of Lemma 10.3.2 in a more general setting.

**Lemma 10.5.2.** *Let $k \ge 2$ and let $T$ be an arbitrary strongly $k$-primitive set of composite numbers such that for each prime $p \in \mathcal{P}(T)$, $|T_p| \ge 2$. Then the map $f : T \to \mathbb{N}$ given by $f(t) = t/P^+(t)$ is one-to-one and $f(T)$ is $(k-1)$-primitive.*

*Proof.* Suppose $t, t' \in T$, $t \ne t'$, and $f(t) = f(t')$. Since $|T_{P^+(t)}| \ge 2$, there is some $s \in T \setminus \{t\}$ with $P^+(t) \mid s$. Then

$$t = P^+(t)\cdot\frac{t}{P^+(t)} = P^+(t)\cdot\frac{t'}{P^+(t')} \;\Big|\; st',$$

contradicting $T$ as strongly 2-primitive. Thus, $f$ is one-to-one.

Next, suppose that $f(T)$ is not strongly $(k-1)$-primitive, so that there are $t, t_1, \dots, t_{k-1}$ in $T$ with $t \notin \{t_1, \dots, t_{k-1}\}$ and

$$f(t) \mid f(t_1)\cdots f(t_{k-1}).$$

With $P^+(t) \mid s$, $s \ne t$ as above, we have $t \mid s\cdot t_1\cdots t_{k-1}$, contradicting $T$ as strongly $k$-primitive. Thus, $f(T)$ is strongly $(k-1)$-primitive, and the proof is complete. □

Let $T_\theta = T \setminus T_0 = \{t \in T : P^+(t) \ge t^\theta\}$. We apply Lemma 10.5.2 to $T$, and so restricting the injection $f$ to $T_\theta$, we have $f(T_\theta)$ as a $(k-1)$-primitive set. Further, every $t \in T_\theta$ has $f(t) \le t^{1-\theta}$. Thus, $t^{-\lambda} \le (t/P^+(t))^{-\lambda/(1-\theta)}$ and by the induction hypothesis on the $(k-1)$-primitive set $f(T_\theta)$,

$$\sum_{t\in T_\theta} t^{-\lambda} \;\le\; \sum_{t\in T_\theta}\big(t/P^+(t)\big)^{-\lambda'} \;=\; \sum_{d\in f(T_\theta')} d^{-\lambda'} \;\le\; \sum_{p\le Y} p^{-\lambda'}$$

for $\lambda' := \lambda_{k-1} = \frac{\lambda_k}{1-\theta}$.

By way of (10.21), this allows us to replace (10.17) with

$$I_\lambda = \sum_{t\in T} t^{-\lambda} - \sum_{p\le Y} p^{-\lambda} < \sum_{p\le p_k} p^{-\nu\lambda} + \frac{\lambda}{\lambda - \frac{1}{k} - \theta}\, p_{k+1}^{-\nu(\lambda-\frac{1}{k}-\theta)} + \sum_{p\le Y} \big(p^{-\lambda'} - p^{-\lambda}\big),$$

with the goal as before to show that $I_\lambda < 0$.

By the mean value theorem, there is some $\xi \in (\lambda, \lambda')$ with

$$\sum_{p\le Y} \big(p^{-\lambda'} - p^{-\lambda}\big) = -(\lambda' - \lambda) \sum_{p\le Y} \frac{\log p}{p^{\xi}} < -\lambda'\theta \sum_{p\le Y} \frac{\log p}{p^{\lambda'}}.$$

Since by Lemma 10.2.4 we may assume that $Y \ge p_{k+1}$, it suffices, by Lemma 10.2.3, for us to show that

$$\sum_{p\le p_k} p^{-\nu\lambda} + \frac{\lambda}{\lambda - \frac{1}{k} - \theta}\, p_{k+1}^{-\nu(\lambda-\frac{1}{k}-\theta)} < \lambda'\theta p_{k+1}^{1-\lambda'}\Big(1 - \frac{1}{\log p_{k+1}}\Big). \tag{10.22}$$

Now $\nu\lambda > 3\log k$, so that

$$\sum_{p\le p_k} p^{-\nu\lambda} < 2^{-3\log k} + (k-1)3^{-3\log k} < k^{-2} + k\cdot k^{-3} = 2k^{-2}.$$

Using $k \ge 39$ we see that $\nu(\lambda - \frac{1}{k} - \theta) > 3\log k - 2$ and $\lambda/(\lambda - \frac{1}{k} - \theta) < 1.23$, so that

$$\frac{\lambda}{\lambda - \frac{1}{k} - \theta}\, p_{k+1}^{-\nu(\lambda-\frac{1}{k}-\theta)} < 1.23 p_{k+1}^{-(3\log k - 2)} < k^{-2}.$$

So the left side of (10.22) is less than $3k^{-2}$. We now get a lower bound for the right side. Using $k \ge 39$, we have $\lambda'\theta > 2(\log k)/k^2$ and $p_{k+1}^{\lambda'} < 4.4$. Thus,

$$\lambda'\theta p_{k+1}^{1-\lambda}\Big(1 - \frac{1}{\log p_{k+1}}\Big) > 0.79\frac{2\log k}{k^2} p_{k+1}/4.4 > \frac{0.36 p_{k+1}\log k}{k^2} > \frac{0.36\log^2 k}{k},$$

using that $p_{k+1} > p_k > k\log k$. We do indeed have $3/k^2 < 0.36(\log^2 k)/k$ when $k \ge 39$, so we have (10.22), and the theorem.

# Chapter 11

# Translated sums of primitive sets

## 11.1 Introduction

One may consider a translated analogue of the Erdős primitive set conjecture: namely, whether the maximum of

$$f(A,h) = \sum_{a\in A} \frac{1}{a(\log a + h)}$$

over primitive sets $A$ attained by the set of primes $A = \mathcal{P}^{(1)}$. Recently Laib, Derbal, and Mechik [84] proved that this translated conjecture is false, by showing $f(\mathcal{P}^{(1)},h) < f(\mathcal{P}^{(2)},h)$ for all $h \geq 81$. Their proof method is direct, by studying partial sum truncations of $f(A,h)$. (Laib [83] very recently announced a bound $h \geq 60$, as a refinement of the same method.)

By different methods, we extend the range of such falsehood down to $h > 1.04\cdots$, and show this translate is best possible for semiprimes.

**Theorem 11.1.1.** *Let $P(s) = \sum_p p^{-s}$ denote the prime zeta function. We have $f(\mathcal{P}^{(1)},h) < f(\mathcal{P}^{(2)},h)$ if and only if $h > h_2$, where $t = h_2 = 1.04\cdots$ is the unique real solution to*

$$\int_1^\infty \Big[P(s) - \tfrac{1}{2}\big(P(s)^2 + P(2s)\big)\Big] e^{(1-s)t}\mathrm{d}s \ = \ 0.$$

*Moreover $f(\mathcal{P}^{(1)},h) < f(\mathcal{P}^{(k)},h)$ for all $k$ sufficiently large, if and only if $h > h_\infty$, where $t = h_\infty = 0.803\cdots$ is the unique real solution to*

$$\int_1^\infty P(s) e^{(1-s)t}\mathrm{d}s \ = \ 1.$$

In view of our later work in Chapter 12, this highlights how the original conjecture of Erdős is quite sensitive to perturbations. Moreover, for the same value $h = h_2 = 1.04\cdots$ we show the primes *minimize* $f(\mathcal{P}^{(k)},h)$, which may be viewed as the inverse analogue of Zhang's maximization result.

**Theorem 11.1.2.** *For $h_2 = 1.04\cdots$, we have $f(\mathcal{P}^{(1)},h_2) < f(\mathcal{P}^{(k)},h_2)$ for all $k > 1$.*

## 11.2 Proof of Theorem 1

We introduce the zeta function for $k$-almost primes $P_k(s) = \sum_{n\in\mathcal{P}^{(k)}} n^{-s}$, whose relevance to us arises from the identity,

$$\begin{aligned} f(\mathcal{P}^{(k)},h) &= \sum_{n\in\mathcal{P}^{(k)}} \frac{1}{n(\log n + h)} = \sum_{n\in\mathcal{P}^{(k)}} \frac{1}{n\log(ne^h)} \\ &= \sum_{n\in\mathcal{P}^{(k)}} e^h \int_1^\infty (ne^h)^{-s}\mathrm{d}s = \int_1^\infty P_k(s) e^{(1-s)h}\mathrm{d}s. \end{aligned} \tag{11.1}$$

Here the interchange of sum and integral holds by Tonelli's theorem, since $f(\mathcal{P}^{(k)}, h) \leq f(\mathcal{P}^{(k)})$ converges uniformly after Erdős. The significance of the identity (11.1) was first observed when $k = 1$, $h = 0$ by H. Cohen [27, p.6], who rapidly computed $f(\mathcal{P}^{(1)}) = 1.636616\cdots$ to 50 digits accuracy. By comparison, the direct approach by partial sums $\sum_{p\leq x} 1/(p\log p)$ converge too slowly, i.e. $O(1/\log x)$. Similarly for $k > 1$, we shall see (11.1) leads to sharper results.

Note one has $P_1(s) = P(s)$ and $P_2(s) = \frac{1}{2}P(s)^2 + \frac{1}{2}P(2s)$, as well as

$$P_3(s) = \frac{1}{6}P(s)^3 + \frac{1}{2}P(s)P(2s) + \frac{1}{3}P(3s),$$
$$P_4(s) = \frac{1}{24}P(s)^4 + \frac{1}{4}P(s)^2P(2s) + \frac{1}{8}P(2s)^2 + \frac{1}{3}P(s)P(3s) + \frac{1}{4}P(4s).$$

In general for $k \geq 1$, Proposition 5.2.1 in Chapter 7 gives an explicit formula for $P_k$ in terms of $P$,

$$P_k(s) = \sum_{n_1+2n_2+\cdots=k} \prod_{j\geq 1} \frac{1}{n_j!}\big(P(js)/j\big)^{n_j}. \tag{11.2}$$

Here the above sum ranges over all partitions of $k$. Also see Proposition 2.1 in [88]. Recall (11.2) is equivalent to the recursion,

$$P_k(s) = \frac{1}{k}\sum_{j=1}^{k} P_{k-j}(s)P(js).$$

We may also rapidly compute $P_k'$ using recursion relations.

**Lemma 11.2.1.** *For $k \geq 1$ let $P_k(s) = \sum_{\Omega(n)=k} n^{-s}$. Then we have*

$$P_k'(s) = \sum_{j=1}^{k} P_{k-j}(s)P'(js). \tag{11.3}$$

*Proof.* By differentiating (11.2), we have

$$\begin{aligned} P_k'(s) &= \sum_{n_1+2n_2+\cdots=k} \sum_{i\leq k} P'(is)\frac{(P(is)/i)^{n_i-1}}{(n_i-1)!}\prod_{j\neq i}\frac{1}{n_j!}\big(P(js)/j\big)^{n_j} \\ &= \sum_{i\leq k} P'(is) \sum_{n_1+2n_2+\cdots=k-i} \prod_{j\geq 1}\frac{1}{n_j!}\big(P(js)/j\big)^{n_j} = \sum_{i\leq k} P'(is)P_{k-i}(s). \end{aligned}$$

□

As observed empirically in [6], the Dirichlet series $P_2(s) - P(s) = \frac{1}{2}[P(s)^2 + P(2s)] - P(s)$ has a unique root at $s = \sigma = 1.1403\cdots$, through which it passes from positive to negative. We prove this more generally for $k \leq 20$.

**Lemma 11.2.2.** *For $2 \le k \le 20$, the Dirichlet series $P_k(s) - P(s)$ has a unique root at $s = \sigma_k > 1$, through which it passes from positive to negative.*

*Proof.* For each $k \ge 1$, $P_k(s)$ is monotonically decreasing in $s > 1$. As such there is a unique $s_k > 1$ for which $P(s)$ passes through $(k!)^{1/(k-1)}$ from above. Using the main term in (11.2), $P_k(s) > P(s)^k/k!$ which is larger than $P(s)$ iff $P(s)^{k-1} > k!$ iff $s < s_k$ by definition. That is,

$$P_k(s) > P(s) > 0 \qquad \text{for} \quad s \in (1, s_k). \tag{11.4}$$

Also there is a unique $s_k' > 1$ for which $P_{k-1}(s)$ passes through 1 from above. Using the first term in the recursion (11.3), $-P_k'(s) > -P'(s)P_{k-1}(s)$ which is larger than $-P'(s) > 0$ iff $P_{k-1}(s) > 1$ iff $s < s_k'$ by definition. That is,

$$P_k'(s) < P'(s) < 0 \qquad \text{for} \quad s \in (1, s_k'). \tag{11.5}$$

For $c < 2^k$, $P_k(s)c^s$ is monotonically decreasing in $s > 1$, and so is $P_k(s)/(2^{-s} + 3^{-s}) = \{1/[P_k(s)2^s] + 1/[P_k(s)3^s]\}^{-1}$. Thus there is a unique $t_k > 1$ for which $P_k(s)/(2^{-s}+3^{-s})$ passes through 1 from above. Now by definition $P(s)/(2^{-s}+3^{-s}) > 1 = P_k(t_k)/(2^{-t_k} + 3^{-t_k})$, which is larger than $P_k(s)/(2^{-s} + 3^{-s})$ iff $s > t_k$ by monotonicity. That is,

$$0 < P_k(s) < P(s) \qquad \text{for} \quad s \in (t_k, \infty). \tag{11.6}$$

In summary $P_k - P$ has at most one root in $(1, s_k')$, and no roots in $(1, s_k) \cup (t_k, \infty)$. We directly compute that $s_k < t_k < s_k'$ for $2 \le k \le 20$, and so $P_k(s) - P(s)$ has a unique root $\sigma_k \in (s_k, t_k)$ as claimed. □

We deduce the following corollary, which gives (the first part of) Theorem 11.1.1 when $k = 2$.

**Corollary 11.2.2.1.** *For $2 \le k \le 20$, $f(\mathcal{P}^{(k)}, h) - f(\mathcal{P}^{(1)}, h)$ has a unique root at $h_k > 0$, through which it passes from negative to positive.*

*Proof.* For $h \ge 0$, recall $f(\mathcal{P}^{(k)}, h) = \int_1^\infty P_k(s)e^{(1-s)h}\mathrm{d}s$ by (11.1). Now by Lemma 11.2.2, $P_k(s) - P(s)$ passes from positive to negative at $s = \sigma_k > 1$. Thus

$$\begin{aligned}[f(\mathcal{P}^{(k)}, h) - f(\mathcal{P}^{(1)}, h)]e^{(\sigma_k-1)h} &= \int_1^\infty [P_k(s) - P(s)]e^{(\sigma_k-s)h}\mathrm{d}s \\ &= \int_1^{\sigma_k} [P_k(s) - P(s)]e^{(\sigma_k-s)h}\mathrm{d}s - \int_{\sigma_k}^\infty [P(s) - P_k(s)]e^{(\sigma_k-s)h}\mathrm{d}s\end{aligned}$$

is difference of two integrals with positive integrands, which are mononotically increasing and decreasing in $h \geq 0$, respectively. Hence the difference is mononotically increasing in $h \geq 0$. And $f(\mathcal{P}^{(k)}, 0) - f(\mathcal{P}^{(1)}, 0) < 0$ by Zhang [143], so the result follows. □

For the second part of Theorem 11.1.1, note for any fixed $h \geq 0$ we have $\log n + h \sim \log n$ for $n \in \mathcal{P}^{(k)}$ as $k \to \infty$, since $n \geq 2^k$. Thus $f(\mathcal{P}^{(k)}, h) \sim f(\mathcal{P}^{(k)}, 0)$ as $k \to \infty$. Hence by Theorem 2.2 5.2.1 in Chapter 7,

$$\lim_{k\to\infty} f(\mathcal{P}^{(k)}, h) = \lim_{k\to\infty} f(\mathcal{P}^{(k)}) = 1.$$

Note $1 - f(\mathcal{P}^{(1)}, h)$ passes from negative to positive at a unique root $h_\infty = 0.803524\cdots$. So for each $h > h_\infty$, we have $f(\mathcal{P}^{(k)}, h) - f(\mathcal{P}^{(1)}, h) > 0$ for $k$ sufficiently large (and similarly for the converse). This completes the proof of Theorem 11.1.1.

### 11.2.1 Computations

For $k \leq 20$, we compute the unique roots $\sigma_k$ and $h_k$ of $P_k(s) - P(s)$ and $f(\mathcal{P}^{(k)}, h) - f(\mathcal{P}^{(1)}, h)$, respectively, as well as verify that the auxiliary parameters (as defined in Lemma 11.2.2) satisfy $s_k < \sigma_k < t_k < s'_k$. We similarly compute the root of $1 - f(\mathcal{P}^{(1)}, h)$ as $h_\infty = 0.803524\cdots$.

The data are displayed in the table below, obtained using Mathematica (for technical convenience, we first compute $y_k = \log h_k$). See Appendix A.

| $k$ | $s_k$ | $t_k$ | $s'_k$ | $\sigma_k$ | $h_k$ |
|---|---|---|---|---|---|
| 2 | 1.11313 | 1.40678 | 1.39943 | 1.14037 | 1.04466 |
| 3 | 1.06861 | 1.23367 | 1.25922 | 1.09224 | 0.98213 |
| 4 | 1.04306 | 1.15231 | 1.17696 | 1.06206 | 0.93018 |
| 5 | 1.02761 | 1.104 | 1.12386 | 1.04231 | 0.89038 |
| 6 | 1.01795 | 1.07259 | 1.08784 | 1.02907 | 0.86146 |
| 7 | 1.01179 | 1.05125 | 1.06272 | 1.02007 | 0.84126 |
| 8 | 1.00779 | 1.0364 | 1.04493 | 1.0139 | 0.8276 |
| 9 | 1.00518 | 1.02594 | 1.03223 | 1.00964 | 0.8186 |
| 10 | 1.00346 | 1.0185 | 1.02312 | 1.0067 | 0.8128 |
| 11 | 1.00231 | 1.0132 | 1.01658 | 1.00466 | 0.80915 |
| 12 | 1.00155 | 1.00942 | 1.01187 | 1.00325 | 0.80689 |
| 13 | 1.00105 | 1.00672 | 1.00849 | 1.00226 | 0.80551 |
| 14 | 1.0007 | 1.00479 | 1.00607 | 1.00158 | 0.8047 |
| 15 | 1.00048 | 1.00341 | 1.00433 | 1.0011 | 0.8042 |
| 16 | 1.00032 | 1.00243 | 1.00309 | 1.00077 | 0.80391 |
| 17 | 1.00022 | 1.00173 | 1.0022 | 1.00053 | 0.80374 |
| 18 | 1.00015 | 1.00123 | 1.00157 | 1.00037 | 0.80365 |
| 19 | 1.0001 | 1.00087 | 1.00112 | 1.00026 | 0.80359 |
| 20 | 1.00007 | 1.00062 | 1.00079 | 1.00018 | 0.80356 |

We believe a unique of root $h_k$ exists as in Corollary 11.2.2.1 for all $k > 20$ as well. This would enable a strengthening of Theorem 11.1.2 to $f(\mathcal{P}^{(k)}, h) > f(\mathcal{P}^{(1)}, h)$ for all values of $k > 1$, $h \geq h_2$ (so far we only establish this for $\{1 < k \leq 20,\ h \geq h_2\}$ or $\{k > 1,\ h = h_2\}$). Uniqueness of $h_k$ would follow if $\sigma_k$ is unique, as in Lemma 11.2.2, for all $k$. In turn it would suffice to show $t_k < s'_k$ for all $k$ (note $s_k < t_k$ holds automatically by (11.4),(11.6)), though it is not clear how to establish such an inequality in general.

Moreover, it appears both $h_k$ and $\sigma_k$ are monotonically decreasing in $k$. This may be related to some empirical trends for $f(\mathcal{P}^{(k)})$ from Chapter 7.

## 11.3 Proof of Theorem 11.1.2

We have already verified the claim directly for $k \leq 20$, since in this case $h_k \leq h_2 = 1.04\cdots$.

For $k > 20$, the proof strategy is similar to that of Theorem 7.5.5 in Chapter 7. That is, the integral $f(\mathcal{P}^{(k)}, h) = \int_1^\infty P_k(s) e^{(1-s)h} \mathrm{d}s$ has its mass concentrated near 1 as $k \to \infty$, so it suffices to truncate the integration to $[1, 1.01]$ say, as a lower bound.

Thus by (11.1),

$$f(\mathcal{P}^{(k)}, h_2) = \int_1^\infty P_k(s) e^{(1-s)h_2} \mathrm{d}s > e^{-.01h_2} \int_1^{1.01} P_k(s) \mathrm{d}s. \tag{11.7}$$

Next, we may lower bound $P_k(s)$ by $P(s)^k/k!$, which constitutes the first of the terms in the identity (11.2), one per partition of $k$. Note the terms of partitions built from small parts contribute the most mass. So by also including the terms for the partitions $k = 1 \cdot (k-j) + j$ and $k = 1 \cdot (k-j-2) + 2 + j$ for $j \le 6$, we shall obtain a sufficiently tight lower bound to deduce the result. Indeed, we have

$$\begin{aligned} \int_1^{1.01} P_k(s) \mathrm{d}s \ &> \ \frac{1}{k!} \int_1^{1.01} P(s)^k \mathrm{d}s + \sum_{j=2}^{6} \frac{\int_1^{1.01} P(s)^{k-j} P(js) \mathrm{d}s}{j(k-j)!} \\ &+ \ \frac{\int_1^{1.01} P(s)^{k-4} P(2s)^2 \mathrm{d}s}{2!2^2(k-4)!} + \sum_{j=3}^{6} \frac{\int_1^{1.01} P(s)^{k-j-2} P(2s) P(js) \mathrm{d}s}{2j(k-j-2)!}. \end{aligned} \tag{11.8}$$

From (7.33) in Chapter 7, we have

$$0 < P(s) - \log(\tfrac{\alpha}{s-1}) < \tfrac{7}{5}(s-1), \qquad \text{for } s \in [1,2], \tag{11.9}$$

where $\alpha = \exp(-\sum_{m \ge 2} P(m)/m) = 0.7292\cdots$. Thus we have for every $k \ge 1$,

$$\int_1^{1.01} P(s)^k \mathrm{d}s > \int_0^{0.01} \log\left(\tfrac{\alpha}{s}\right)^k \mathrm{d}s = \alpha \int_{\log(100\alpha)}^\infty u^k e^{-u} \, du > 0.729 \, k!, \tag{11.10}$$

Also note

$$\int_0^1 s \log\left(\tfrac{\alpha}{s}\right)^k \mathrm{d}s = \alpha^2 \int_0^\infty u^k e^{-2u} \, du = \frac{\alpha^2}{2^{k+1}} k!.$$

Using the first order Taylor approximation $P(js) > P(j) + P'(j)(s-1)$ for $j \ge 2$,

$$\begin{aligned} \int_1^{1.01} P(s)^{k-j} P(js) \mathrm{d}s \ &> \ P(j) \int_0^{0.01} \log\left(\tfrac{\alpha}{s}\right)^{k-j} \mathrm{d}s + P'(j) \int_0^1 s \log\left(\tfrac{\alpha}{s}\right)^{k-j} \mathrm{d}s \\ &> \ 0.729(k-j)! \Big( P(j) + \frac{\alpha P'(j)}{2^{k-j}} \Big) \end{aligned}$$

by (11.10). Similarly,

$$\begin{aligned} &\int_1^{1.01} P(s)^{k-j-2} P(2s) P(js) \mathrm{d}s \\ &\quad > P(2)P(j) \int_0^{.01} \log\left(\tfrac{\alpha}{s}\right)^{k-j-2} \mathrm{d}s \ + \ [P'(2)P(j) + P(2)P'(j)] \int_0^1 s \log\left(\tfrac{\alpha}{s}\right)^{k-j-2} \mathrm{d}s \\ &\quad > 0.729(k-j-2)! \Big( P(2)P(j) + \frac{\alpha}{2^{k-j-1}} [P'(2)P(j) + P(2)P'(j)] \Big). \end{aligned}$$

Hence plugging back into (11.7),

$$f(\mathcal{P}^{(k)}, h_2)e^{0.01h_2} > \int_1^{1.01} P(s)^k \mathrm{d}s \ > \ \ell_k, \tag{11.11}$$

for the explicit lower bound

$$\begin{aligned}\ell_k \ := \ 0.729\Bigg[1 + \sum_{j=2}^{6} \frac{P(j) + \alpha P'(j)/2^{k-j}}{j} \ + \ \frac{1}{8}\Big(P(2)^2 + \frac{\alpha P(2)P'(2)}{2^{k-4}}\Big) \\ + \ \sum_{j=3}^{6} \frac{1}{2j}\bigg(P(2)P(j) + \frac{\alpha}{2^{k-j-1}}[P'(2)P(j) + P(2)P'(j)]\bigg)\Bigg].\end{aligned}$$

Note $\ell_k$ is clearly increasing in $k$ (recall $P'(s) < 0$). Hence for $k > 20$ we have

$$f(\mathcal{P}^{(k)}, h_2) > e^{-0.01h_2}\ell_k > e^{-0.01h_2}\ell_{20} > 0.98 > 0.91 > f(\mathcal{P}^{(1)}, h_2). \tag{11.12}$$

Here we compute $\ell_{20} = 0.99069\cdots$ and $f(\mathcal{P}^{(1)}, h_2) = 0.908599\cdots$. This completes the proof.

# Chapter 12

# A proof of the Erdős primitive set conjecture

## 12.1 Introduction

A set of integers $A \subset \mathbb{Z}_{>1}$ is *primitive* if no member in $A$ divides another. For example, the integers in a dyadic interval $(x, 2x]$ form a primitive set. Similarly the set of primes is primitive, along with the set $\mathbb{N}_k$ of numbers with exactly $k$ prime factors (with multiplicity), for each $k \geq 1$. Another well-known example is the set of perfect numbers.[1]

The study of primitive sets emerged in the 1930s as a generalization of one special problem. A classical theorem of Davenport asserts that the set of abundant numbers has a positive asymptotic density. This was originally proved by sophisticated analytic methods, but Erdős soon found an elementary proof by using primitive abundant numbers.[2] The proof ideas led people to introduce the abstract definition of primitive sets and study them for their own sake. See Hall [70] or Halberstam–Roth [69, §5] for detailed introductions to the subject.

There are a number of interesting and sometimes unexpected theorems about primitive sets. For instance, in 1934 Besicovitch [10] showed that the upper asymptotic density of a primitive set can be arbitrarily close to $1/2$, whereas in 1935 Behrend [9] and Erdős [40] proved the lower asymptotic density is always 0. In fact, Erdős proved the stronger result that

$$f(A) := \sum_{a \in A} \frac{1}{a \log a} \ < \ \infty,$$

uniformly over all primitive sets $A$. Later in 1986 Erdős [46, Conjecture 2.1] famously asked if the maximum is attained by the primes $\mathcal{P}$.

**Conjecture 12.1.1** (Erdős primitive set conjecture)**.** For any primitive set $A$, $f(A) \leq f(\mathcal{P})$.

The prime sum is $f(\mathcal{P}) = \sum_p 1/(p \log p) = 1.6366\cdots$ after computations of Cohen [27]. In 1993, Erdős and Zhang [53] proved the bound $f(A) < 1.84$ for all primitive $A$. Recently in 2019, Pomerance and the author [90] improved the bound to $f(A) < e^\gamma = 1.781\cdots$, where $\gamma$ is the Euler-Mascheroni constant. Note the tail of the series for $f(\mathcal{P})$ converges quite slowly $O(1/\log x)$, and moreover there are sets $A \subset [x, \infty)$ for which $f(A) \sim 1$ as $x \to \infty$ (in this connection see Conjecture 12.1.4 below). As

[1] Since Ancient Greece, a number $n$ is classified as 'perfect,' 'abundant,' or 'deficient,' depending on whether the sum of its proper divisors equals $n$, is greater than $n$, or is less than $n$, respectively.

[2] More precisely, 'primitive non-deficient numbers'

such, Conjecture 12.1.1 is not susceptible to direct attack by computing partial sums up to $x$.

One potential strategy to approach Conjecture 12.1.1 is via integration. Namely,

$$f(A) = \sum_{a\in A} \frac{1}{a\log a} = \sum_{a\in A} \int_1^\infty a^{-t}\mathrm{d}t = \int_1^\infty f_t(A)\mathrm{d}t,$$

letting $f_t(A) = \sum_{a\in A} a^{-t}$. So Conjecture 1 would follow if for any $t > 1$, primitive set $A$,

$$f_t(A) \ \leq \ f_t(\mathcal{P}). \tag{12.1}$$

However, it was shown in [6] that (12.1) holds if and only if

$$t \geq \tau := 1.1403\cdots,$$

where $t = \tau$ is the unique real solution to the equation

$$\sum_p p^{-t} = 1 + \Big(1 - \sum_p p^{-2t}\Big)^{1/2}.$$

The fact that $\tau$ is markedly larger than 1 gives some indication as to why the Erdős primitive set conjecture has remained open.

Similar analysis actually enables a disproof of a natural analogue of Conjecture 12.1.1 for the translated sum $f(A,h) = \sum_{a\in A} 1/a(\log a+h)$, in that there are primitive $A$ for which $f(A,h) > f(\mathcal{P},h)$ once $h \geq 81$ [84], [83]. This was refined down to just $h \geq 1.04$ in Chapter 11, and highlights the sensitivity to perturbation of the original conjecture (when $h = 0$).

A separate strategy for the problem is to split up $A$ according to the smallest prime factor. That is, for each prime $p$ let

$$A_p = \{n \in A : n \text{ has least prime factor } p\}.$$

As in [90], we say $p$ is *Erdős strong* if the singleton set $\{p\}$ maximizes $f(A)$ among all primitive sets $A$ all of whose elements have least prime factor $p$. That is, $f(A_p) \leq f(\{p\}) =: f(p)$ for all primitive $A$. Conjecture 12.1.1 would follow if every prime is Erdős strong, since then $f(A) = \sum_p f(A_p) \leq f(\mathcal{P})$.

By a short argument in [90] (also see Lemma 12.2.2), a sufficient condition for a prime $p$ to be Erdős strong is that

$$e^\gamma \prod_{q<p} \Big(1 - \frac{1}{q}\Big) \ \leq \ \frac{1}{\log p}. \tag{12.2}$$

Here $q$ runs over primes. Note the two sides of this inequality are asymptotically equal by Mertens' prime product theorem. By direct computation, (12.2) is satisfied by the first $10^8$ odd primes, but fails for $p = 2$ since $\log 2 > e^{-\gamma}$.

Moreover 99.999973% of primes[3] satisfy (12.2), assuming the Riemann Hypothesis and the Linear Independence hypothesis[4] [91]. This result is intimately related to the celebrated work of Rubinstein and Sarnak [125] on the prime number race between $\pi(x)$ and $\mathrm{li}(x)$. On the Riemann Hypothesis alone (12.2) fails for a positive proportion of primes $p$ (in log density), and even unconditionally (12.2) is known to fail for infinitely many primes $p$. This perhaps suggests Conjecture 12.1.1 might be false, or at least beyond the reach of unconditional tools.

In this chapter we establish Conjecture 12.1.1.

**Theorem 12.1.2.** *For any primitive set $A$, we have $f(A) \le f(\mathcal{P})$.*

Moreover, we show that every odd prime is Erdős strong.

**Theorem 12.1.3.** *For any primitive set $A$ and any prime $p > 2$, we have $f(A_p) \le f(p)$.*

It remains an open question whether $p = 2$ is Erdős strong.

Another question related to Conjecture 12.1.1, in 1968 Erdős, Sárközy, and Szemerédi posed the following [51, eq. (11)].

**Conjecture 12.1.4** (Erdős–Sárközy–Szemerédi)**.** We have

$$\lim_{x\to\infty} \sup_{\substack{A\subset[x,\infty)\\ A \text{ primitive}}} f(A) \le 1.$$

This also appears in [126, p. 244] as Problem 2.2, and in [127, p. 224] as Problem 2.

Not much has been proven in this direction until very recently. Recall the set $\mathbb{N}_k$ of numbers with exactly $k$ prime factors (with multiplicity) lies in $[2^k, \infty)$. Pomerance and the author [90] proved $f(\mathbb{N}_k) \gg 1$, and in Chapter 7 it was shown $f(\mathbb{N}_k) \sim 1$ as $k \to \infty$. This means that if Conjecture 12.1.4 holds, then the limit must attain an *equality* of 1. We note Theorem 7.4.1 in Chapter 7 gives for all $\varepsilon > 0$,

$$f(\mathbb{N}_k) = 1 + O_\varepsilon(k^{\varepsilon-1/2}). \tag{12.3}$$

[3]More precisely, the set of such primes has discrete, logarithmic density equal to $0.99999973\cdots$ within $\mathcal{P}$.

[4]Namely, the sequence of values $\gamma_n > 0$ such that $\zeta(\frac{1}{2} + i\gamma_n) = 0$ is linearly independent over $\mathbb{Q}$.

Moreover, computations up to $k = 20$ suggest the true rate of decay may be exponential $O(2^{-k})$, see Chapter 7.

The methods in this chapter enable the following progress towards Conjecture 12.1.4.

**Theorem 12.1.5.** *We have*

$$\lim_{x\to\infty} \sup_{\substack{A\subset[x,\infty)\\ A \text{ primitive}}} f(A) \ \le \ e^{\gamma}\frac{\pi}{4} \approx 1.399.$$

## Notation

Let $p^-(a), P^+(a)$ denote the smallest and largest prime factors of $a \in \mathbb{Z}_{>1}$, respectively, and denote $a^* = a/P^+(a)$. Let $\mathbb{N}_k = \{n : \Omega(n) = k\}$. Define $f(a) = 1/(a\log a)$ and $f(A) = \sum_{a\in A} f(a)$ for $A \subset \mathbb{Z}_{>1}$. Let $\mathcal{P}$ be the set of prime numbers, whose elements we denote by $p$ and $q$, unless otherwise stated. Also $p^k \| n$ means $p^k \mid n$ and $p^{k+1} \nmid n$.

### 12.1.1 Proof outline of Theorem 12.1.2

The proof is a refinement of the argument of [90]. The key new idea is to exploit the fact that $A$ cannot contain too many elements $a$ with $P^+(a)$ just slightly less than $a$. This improves the critical case in the argument of [90], and ultimately leads to an improvement by a factor of $\pi/4$ from a contribution from each $a \in A$ which is not prime. Since $e^{\gamma}\pi/4 < f(\mathcal{P})$, this ultimately means that $f(A)$ is maximized when all elements are prime. (Additional care is needed for small numbers, using explicit bounds.)

Let us recall the rough argument of [90] (suppressing details for primes and small numbers). By Mertens' product theorem,

$$f(A) = \sum_{a\in A}\frac{1}{a\log a} < \sum_{a\in A}\frac{1}{a\log P^+(a)} \approx e^{\gamma}\sum_{a\in A}\frac{1}{a}\prod_{p<P^+(a)}\Big(1-\frac{1}{p}\Big). \tag{12.4}$$

But $a^{-1}\prod_{p<P^+(a)}(1-p^{-1})$ is the natural density of $\mathrm{L}_a = \{ba \ : \ p \mid b \Rightarrow p \ge P^+(a)\}$, and these sets turn out to be disjoint by primitivity of $A$ (Lemma 12.2.1). So the sum of densities in (12.4) is trivially at most 1, leading to the bound $f(A) < e^{\gamma}$ for primitive $A$. This is inspired by the original 1935 argument of Erdős [40].

There is a loss in the above argument when bounding $a$ by $P^+(a)$, and this loss is largest when $a$ is far from prime. We can save an additional factor of $\log P^+(a)/\log a$ for any individual $a \in A$, and this would be a significant improvement in the case

$P^+(a)^2 < a$, say. Therefore the critical case to handle is when $a \in A$ is composite with $P^+(a)$ close to $a$ in size. The key new ingredient (Proposition 12.3.3) shows that if $P^+(a)^{1+v} > a$ uniformly for all $a \in A$ (so the savings factor is $\log P^+(a)/\log a > 1/(1+v)$), then we can bound the sum of densities in (12.4) by $\sqrt{v}$. This refines the trivial bound of 1 in the range $0 < v < 1$, and quantifies the earlier statement that $A$ contains few elements $a$ with $P^+(a)$ slightly less than $a$. As the savings $1/(1+v)$ improves with $v$, the worst-case scenario is when the subset of $a \in A$ with $P^+(a)^{1+v} \approx a$ contributes about $\frac{\mathrm{d}}{\mathrm{d}v}[\sqrt{v}] = 1/2\sqrt{v}$ to the sum of densities in (12.4). Combining these ingredients ultimately leads to a savings of $\int_0^1 \mathrm{d}v/2\sqrt{v}(1+v) = \pi/4$, as desired.

Lastly, the key Proposition 12.3.3 relies on the following observation (Lemma 12.3.1): not only are the sets $\mathrm{L}_a$ disjoint, but so too are $\mathrm{L}_{ac}$ for many choices of integers $c$ (in fact all choices of $c$ with prime factors between $P_2^+(a)$ and $P_2^+(a)^{1/\sqrt{v}}$). Thus the sum of densities of these $\mathrm{L}_{ac}$ must be at most 1. But these sets $\mathrm{L}_{ac}$ are self-similar to the $\mathrm{L}_a$, and so the sum of their densities is roughly $1/\sqrt{v}$ times that of the $\mathrm{L}_a$, giving the desired bound $\sqrt{v}$.

### 12.1.2 L-primitive sets

As outlined above, the subset of multiples of each $a \in A$,

$$\mathrm{L}_a := \big\{ba \in \mathbb{N} \,:\, p \mid b \implies p \geq P^+(a)\big\}, \tag{12.5}$$

arises naturally in our proof. As such we shall introduce 'L' refinements of our common notions (here L alludes to 'lexicographic'). Specifically, if $n \in \mathrm{L}_a$ we say $n$ is an L-*multiple* of $a$, and $a$ is an L-*divisor* of $n$. Most importantly, we introduce the following key definition.

**Definition 12.1.6.** A set $A \subset \mathbb{Z}_{>1}$ is L-*primitive* if $a' \notin \mathrm{L}_a$ for all distinct $a, a' \in A$.

That is, $A$ is L-primitive if no member of $A$ is an L-multiple of another. In particular this definition is weaker than primitive.

One may apply the basic argument as in (12.4) more generally for L-primitive sets $A$, leading to the same bound $f(A) < e^\gamma$ (again ignoring small numbers). Moreover, L-primitive sets play a central role in the proof of Theorem 12.1.2. However, it turns out the bound $e^\gamma$ is essentially best possible for L-primitive sets (see Proposition 12.5.3), which is markedly larger than $f(\mathcal{P})$. This further highlights the subtlety of Conjecture 12.1.1.

### 12.1.3 Density and divisibility chains

Recall the natural (asymptotic) density $\mathrm{d}(S) = \lim_{x\to\infty} |S \cap [1,x]|/x$ of a set $S \subset \mathbb{N}$. We also consider log density $\delta(S)$ and log log density $\Delta(S)$, given by

$$\delta(S) = \lim_{x\to\infty} \frac{1}{\log x} \sum_{n\in S, n\le x} \frac{1}{n}, \qquad \text{and} \qquad \Delta(S) = \lim_{x\to\infty} \frac{1}{\log\log x} \sum_{n\in S, 1<n\le x} \frac{1}{n\log n}, \tag{12.6}$$

provided these limits exist. Recall the corresponding upper densities $\overline{\mathrm{d}}(S)$, $\overline{\delta}(S)$, $\overline{\Delta}(S)$ always exist, by replacing $\lim_{x\to\infty}$ with $\limsup_{x\to\infty}$ (and similarly $\liminf_{x\to\infty}$ for lower densities).

Taking an abstract view, a primitive set is an antichain for the partial ordering of integers by divisibility. As such this naturally leads to the dual notion of a chain in this context. Namely, an infinite sequence of integers $1 < d_1 < d_2 < \cdots$ is a *divisibility chain* if $d_j \mid d_{j+1}$ for all $j \ge 1$. A classical 1937 theorem of Davenport and Erdős [30] asserts that if set $A \subset \mathbb{N}$ has upper log density $\overline{\delta}(A) > 0$, then it contains an infinite divisibility chain $D \subset A$.

Analogously, we introduce the following refinement.

**Definition 12.1.7.** An infinite sequence of integers $1 < d_1 < d_2 < \cdots$ is an L-*divisibility chain* if $d_{j+1} \in \mathrm{L}_{d_j}$ for all $j \ge 1$.

That is, $d_{j+1}$ is an L-multiple of $d_j$ for all $j \ge 1$. In particular this definition is stronger than a (mere) divisibility chain.

We refine the Davenport–Erdős theorem to L-divisibility chains.

**Theorem 12.1.8.** *If a set $A \subset \mathbb{N}$ has upper* log *density $\overline{\delta}(A) > 0$, then $A$ contains an infinite* L*-divisibility chain.*

In 1966 Erdős, Sárközy, and Szemerédi [48, Theorem 1] quantified the Davenport–Erdős theorem by showing such a divisibility chain $D$ which grows at least as fast as $\limsup_{y\to\infty} \sum_{d\in D, d\le y} 1/\sqrt{\log\log y} > 0$, and proved such growth rate is best possible.

They also studied the analogous question for upper log log density, which they write "seems more interesting to us." Namely, in [48, Theorem 2] they established the following quantitative result.

**Theorem 12.1.9** (Erdős–Sárközy–Szemerédi)**.** *If $A \subset \mathbb{N}$ has upper* log log *density $\overline{\Delta}(A) > 0$, then there is an infinite divisibility chain $D \subset A$ of growth*

$$\limsup_{y\to\infty} \sum_{\substack{d\in D\\ d\le y}} \frac{1}{\log\log y} \;\ge\; \frac{\overline{\Delta}(A)}{e^{\gamma}}. \tag{12.7}$$

Analogously, we quantify Theorem 12.1.8 in the case of log log density, thereby refining Theorem 12.1.9 of Erdős–Sárközy–Szemerédi to L-divisibility chains.

**Theorem 12.1.10.** *If $A \subset \mathbb{N}$ has upper* log log *density $\overline{\Delta}(A) > 0$, then there is an infinite* L*-divisibility chain $D \subset A$ of growth*

$$\limsup_{y\to\infty} \sum_{\substack{d\in D\\ d\le y}} \frac{1}{\log\log y} \;\ge\; \frac{\overline{\Delta}(A)}{e^{\gamma}}.$$

In view of Proposition 12.5.3 we believe that the lower bound $\overline{\Delta}(A)/e^{\gamma}$ above is best possible for L-divisibility chains, though we are unable to settle this. Notably, this contrasts the situation in Theorem 12.1.9, as Erdős–Sárközy–Szemerédi conjectured $\overline{\Delta}(A)/e^{\gamma}$ in (12.7) might be improved to $\overline{\Delta}(A)$, which would be best possible for divisibility chains, if true [48, eq. (5)].

## 12.2 Preliminaries on L-primitive sets

Recall the set of L-multiples $\mathrm{L}_a := \{ba \in \mathbb{N} : p \mid b \Rightarrow p \ge P^+(a)\}$ from (12.5). In particular $a \in \mathrm{L}_a$ for $b = 1$, and $p^-(b) \ge P^+(a)$ for $b > 1$. For $A \subset \mathbb{N}$ define $\mathrm{L}_A := \bigcup_{a\in A} \mathrm{L}_a$. Also let $A_a = A \cap \mathrm{L}_a$ so that $\mathbb{N}_a = \mathrm{L}_a$ and $A_q = \{a \in A : p^-(a) = q\}$ for prime $q$.[5]

Observe that $a \in \mathrm{L}_{a'}$ if and only if $\mathrm{L}_a \subset \mathrm{L}_{a'}$, as well as the following trichotomy.

**Lemma 12.2.1.** *For any integers $a, a' > 1$, if $\mathrm{L}_a \cap \mathrm{L}_{a'} \neq \emptyset$ then $a \in \mathrm{L}_{a'}$ or $a' \in \mathrm{L}_a$. Thus $\mathrm{L}_a \cap \mathrm{L}_{a'} = \emptyset$ or $\mathrm{L}_a \subset \mathrm{L}_{a'}$ or $\mathrm{L}_a \supset \mathrm{L}_{a'}$.*

*Proof.* Suppose $ba = b'a' \in \mathrm{L}_a \cap \mathrm{L}_{a'}$. If $b = 1$ or $b' = 1$ then $a \in \mathrm{L}_{a'}$ or $a' \in \mathrm{L}_a$. Otherwise $b, b' > 1$, so $P^+(a) \le p^-(b)$ and $P^+(a') \le p^-(b')$ imply $b \mid b'$ or $b' \mid b$. Thus $a' = a(b/b') \in \mathrm{L}_a$ or $a = a'(b'/b) \in \mathrm{L}_{a'}$ as well. □

As such we see $A$ is L-primitive if and only if the sets $\{\mathrm{L}_a\}_{a\in A}$ are pairwise disjoint.

**Corollary 12.2.1.1.** *If $A$ is an* L*-primitive set, then $\mathrm{L}_a$ and $\mathrm{L}_{a'}$ are disjoint for distinct $a, a' \in A$.*

Recall $\mathrm{L}_a$ has natural density $\mathrm{d}(\mathrm{L}_a) = \frac{1}{a}\prod_{p<P^+(a)}(1-\frac{1}{p})$. And by Mertens' product theorem $\prod_{p<x}(1-\frac{1}{p}) \sim 1/e^{\gamma}\log x$, where $\gamma = .57721\cdots$ is the Euler-Mascheroni constant. By a show argument below, we relate $f(A)$ to density of L-multiples. This is essentially based on Erdős [40] (also see [50, Lemma 1], [90, Proposition 2.1]).

[5]Note the notation for $A_q$ differs slightly from what is used in [53], [90]

**Lemma 12.2.2.** *For an* L*-primitive set* $A$ *and an integer* $1 < n \notin A$*, we have* $f(A_n) < e^\gamma \, \mathrm{d}(\mathrm{L}_n)$.

*Proof.* We may assume $A = A_n$ is finite, since $f(A) = \lim_{x\to\infty} f(A \cap [1,x])$. As $n \notin A$ all elements of $A$ are composite. Also $A$ is L-primitive so $\mathrm{d}(\mathrm{L}_A) = \sum_{a\in A} \mathrm{d}(\mathrm{L}_a)$ by Corollary 12.2.1.1. Next, Theorem 7 in [124] implies $\prod_{p<x} \frac{p}{p-1} < e^\gamma \log(2x)$ for all $x > 1$. Thus for any composite integer $a > 1$, we have $a > 2P^+(a)$ so that

$$f(a) = \frac{1}{a \log a} \;\le\; \frac{1}{a \log 2P^+(a)} \;<\; \frac{e^\gamma}{a} \prod_{p<P^+(a)} \Big(1 - \frac{1}{p}\Big) \;=\; e^\gamma \, \mathrm{d}(\mathrm{L}_a).$$

Hence $f(A) = \sum_{a\in A} f(a) < e^\gamma \, \mathrm{d}(\mathrm{L}_A) \le e^\gamma \, \mathrm{d}(\mathrm{L}_n)$ since $A \subset \mathrm{L}_n$. □

We shall also need a technical refinement of Lemma 12.2.2. For this, we rewrite Mertens' product theorem as $\mu_x \sim 1$, where we denote

$$\mu_x := e^\gamma \log x \prod_{p<x} \Big(1 - \frac{1}{p}\Big). \tag{12.8}$$

In particular, for a prime $q$ we have

$$f(q) = \frac{1}{q \log q} = \frac{1}{q} \frac{e^\gamma}{\mu_q} \prod_{p<q} \Big(1 - \frac{1}{p}\Big) = \frac{e^\gamma}{\mu_q} \mathrm{d}(\mathrm{L}_q). \tag{12.9}$$

We have the following explicit bounds for $\mu_x$, which critically are monotonic. We give upper bounds which hold on real $x \in \mathbb{R}$, but for lower bounds it turns out it suffices to restrict to the subsequence of primes $q \in \mathcal{P}$.

**Lemma 12.2.3** (Monotonic bounds)**.** *For* $q \in \mathcal{P}$ *and* $x \in \mathbb{R}$*, define*

$$m_q := \inf_{\substack{p\ge q \\ p\in\mathcal{P}}} \mu_p, \quad \textit{and} \quad M_x := \sup_{\substack{y\ge x \\ y\in\mathbb{R}}} \mu_y.$$

*Then we have*

$$m_q \ge \begin{cases} \mu_7 = 0.9242\cdots & q \le 7 \\ \mu_{19} = 0.9467\cdots & 7 < q \le 300\,, \\ 1 - \frac{1}{2(\log q)^2} & q > 300. \end{cases} \qquad M_x \le \begin{cases} \mu_2 = 1.235\cdots & x \le 2 \\ 1 + \frac{1}{2\log(2\cdot 10^9)^2} & 2 < x \le 2\cdot 10^9 \\ 1 + \frac{1}{2(\log x)^2} & x > 2\cdot 10^9. \end{cases}$$

*Proof.* First, Rosser-Schoenfeld [124, Theorem 7] implies the product over primes $p < x$ is bounded in between

$$\begin{aligned} e^\gamma \log x \prod_{p<x} \Big(1 - \frac{1}{p}\Big) \;&\le\; 1 + \frac{1}{2(\log x)^2} \qquad \text{for} \quad x > 1, \\ &\ge\; 1 - \frac{1}{2(\log x)^2} \qquad \text{for} \quad x > 285. \end{aligned}$$

Note $\mu_x$ is increasing on $x \in (p, p']$ for consecutive primes $p, p'$. So the upper bound follows by computing $\mu_p < 1$ for the first $10^8$ odd primes $p$ (note $p_{10^8} \geq 2 \cdot 10^9$). Hence $\mu_x < 1$ for real $2 < x \leq 2 \cdot 10^9$. Below we display $\mu_q$ for the first few primes $q$, rounded to 4 significant digits.

| $q$ | $\mu_q$ | $q$ | $\mu_q$ | $q$ | $\mu_q$ | $q$ | $\mu_q$ | $q$ | $\mu_q$ |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 1.235 | 31 | **0.9660** | 73 | 0.9766 | 127 | 0.9902 | 179 | 0.9909 |
| 3 | 0.9784 | 37 | 0.9831 | 79 | 0.9809 | 131 | 0.9887 | 181 | 0.9874 |
| 5 | 0.9555 | 41 | 0.9836 | 83 | 0.9795 | 137 | 0.9902 | 191 | 0.9921 |
| 7 | **0.9242** | 43 | 0.9720 | 89 | 0.9829 | 139 | 0.9858 | 193 | 0.9889 |
| 11 | 0.9762 | 47 | **0.9718** | 97 | 0.9906 | 149 | 0.9925 | 197 | 0.9876 |
| 13 | 0.9492 | 53 | 0.9808 | 101 | 0.9890 | 151 | 0.9885 | 199 | **0.9844** |
| 17 | 0.9679 | 59 | 0.9883 | 103 | 0.9834 | 157 | 0.9896 | | |
| 19 | **0.9467** | 61 | 0.9795 | 107 | 0.9818 | 163 | 0.9906 | | |
| 23 | **0.9551** | 67 | 0.9854 | 109 | 0.9765 | 167 | 0.9892 | | |
| 29 | 0.9811 | 71 | 0.9841 | 113 | **0.9749** | 173 | 0.9900 | | |

The lower bound follows by identifying the primes $q$ for which $\mu_q = \inf_{p \geq q} \mu_p$ (in bold above), and then computing $\mu_{199} < \mu_p$ for $199 < p \leq 300$, as well as checking $\mu_{199} < 0.9846 < 1 - \frac{1}{2(\log x)^2}$ for $x > 300$. (In practice we shall only need $\mu_q$ for $q = 7, 19$.) □

We may now prove a technical refinement of Lemma 12.2.2 using $\mu_q$.

**Lemma 12.2.4.** *Let $A$ be an* L*-primitive set. Take $v \geq 0$, an integer $n \notin A$, and denote $q = P^+(n)$. If $P^+(a)^{1+v} \leq a$ for all $a \in A_n$, then*

$$f(A_n) = \sum_{a \in A_n} \frac{1}{a \log a} \ \leq \ \frac{e^\gamma}{m_q} \frac{\mathrm{d}(\mathrm{L}_{A_n})}{1+v}.$$

*Proof.* We may assume $A = A_n$ is finite, since $f(A) = \lim_{x \to \infty} f(A \cap [1, x])$. As $n \notin A$ all elements of $A$ are composite. Also $A$ is L-primitive so $\mathrm{d}(\mathrm{L}_A) = \sum_{a \in A} \mathrm{d}(\mathrm{L}_a)$ by Corollary 12.2.1.1. Moreover $(1+v) \log P^+(a) \leq \log a$ for all $a \in A$. Thus by definition of $\mu_{P^+(a)}$ in (12.8),

$$\frac{1}{a \log a} \ \leq \ \frac{1}{1+v} \frac{1}{a \log P^+(a)} = \frac{e^\gamma}{\mu_{P^+(a)}} \frac{1}{(1+v)a} \prod_{p < P^+(a)} \Big(1 - \frac{1}{p}\Big) = \frac{e^\gamma}{\mu_{P^+(a)}} \frac{\mathrm{d}(\mathrm{L}_a)}{1+v}.$$

By monotonicity $\mu_{P^+(a)} \geq m_{P^+(a)} \geq m_q$ for $a \in A \subset \mathrm{L}_n$. Hence we conclude

$$f(A) = \sum_{a \in A} \frac{1}{a \log a} \ \leq \ \frac{e^\gamma}{m_q} \frac{1}{1+v} \sum_{a \in A} \mathrm{d}(\mathrm{L}_a) = \frac{e^\gamma}{m_q} \frac{\mathrm{d}(\mathrm{L}_A)}{1+v}.$$

□

## 12.3 Primitive sets

Given $v \in (0,1)$, we shall be interested in elements $a \in A$ for which $P^+(a)^{1+v} > a$, and their multiples $ac$, where $c \in C_a^v$ for

$$C_a^v := \big\{c \in \mathbb{N} : p \mid c \implies p \in [P^+(a^*), P^+(a^*)^{1/\sqrt{v}})\big\}. \tag{12.10}$$

Note $c = 1 \in C_a^v$. Recall $a^* = a/P^+(a)$, so $P^+(a^*)$ is the second largest prime of $a$. Also if $1 < c \in C_a^v$ then $P^+(c) \le P^+(a^*)^{1/\sqrt{v}}$ is markedly smaller than $P^+(a) \ge P^+(a^*)^{1/v}$.

The following key lemma provides an upgrade to Corollary 12.2.1.1 in the case when $A$ is primitive, not just L-primitive. Namely, the $\mathrm{L}_{ac}$ are disjoint, and so the larger set $\{ac : a \in A, c \in C_a^v\}$ is L-primitive.

**Lemma 12.3.1.** *Let $A$ be a primitive set of composite numbers, and take $v \in (0,1)$. If $P^+(a)^{1+v} > a$ for all $a \in A$, then the collection of sets $\mathrm{L}_{ac}$, ranging over $a \in A, c \in C_a^v$, are pairwise disjoint.*

*Proof.* Suppose $\mathrm{L}_{ac} \cap \mathrm{L}_{a'c'} \neq \emptyset$ for some $a, a' \in A$ and $c \in C_a^v$, $c' \in C_{a'}^v$. Without loss, by Lemma 12.2.1 we may assume $ac \in \mathrm{L}_{a'c'}$. Note if $c = 1$ then $a \in \mathrm{L}_{a'c'}$ implies $a' \mid a'c' \mid a$, which forces $a = a'$ and $c' = 1$ by primitivity of $A$. So assuming $(a,c) \neq (a',c')$ we deduce $c > 1$.

We factor $ac = p_1 \cdots p_k$ into primes $p_1 \ge \cdots \ge p_k$, so $ac \in \mathrm{L}_{a'c'}$ implies $a'c' = p_j \cdots p_k$ for some index $1 < j < k$. Since $P^+(a) > P^+(c), p^-(c) \ge P^+(a^*)$ we also have $a^* = p_i \cdots p_k$ for some $2 < i \le k$. If $i \le j$ then $a'c' \mid a^*$ so $a' \mid a$, contradicting $A$ as primitive. Hence $i > j$ so $a^* \mid a'c'$. Write $da^* = a'c'$ where $d = p_j \cdots p_{i-1}$, and note $P^+(d) = p_j = P^+(a')$. By definition of $1 < c \in C_a^v$, we have

$$p_j = P^+(d) \le P^+(c) < P^+(a^*)^{1/\sqrt{v}}. \tag{12.11}$$

Recall $P^+(a')^v > (a')^* \ge P^+((a')^*)$ for $a' \in A$. Now consider cases $c' > 1$ and $c' = 1$. When $1 < c' \in C_{a'}^v$, we have $P^+(c') = p_{j+1} \ge p_i = P^+(a^*)$. Thus

$$p_j = P^+(a') > P^+((a')^*)^{1/v} > P^+(c')^{1/\sqrt{v}} \ge P^+(a^*)^{1/\sqrt{v}}. \tag{12.12}$$

But (12.12) contradicts (12.11), so $\mathrm{L}_{ac}$ and $\mathrm{L}_{a'c'}$ are disjoint.

Similarly when $c' = 1$, we have $P^+((a')^*) = p_{j+1} \ge p_i = P^+(a^*)$ and so

$$p_j = P^+(a') > P^+((a')^*)^{1/v} \ge P^+(a^*)^{1/v}.$$

This also contradicts (12.11) (indeed $v < \sqrt{v}$). Hence $\mathrm{L}_{ac}$ and $\mathrm{L}_{a'}$ are disjoint in both cases. □

*Remark* 12.3.2. The exponent $1/\sqrt{v}$ in the definition of $C_a^v$ in (12.10) is chosen as large as possible, constrained by the final steps (12.11), (12.12) above. If one established a larger exponent in Lemma 12.3.1, this would improve the final savings factor $\int_0^1 \frac{\mathrm{d}}{\mathrm{d}v}[v^{1/2}]\frac{\mathrm{d}v}{1+v} = \pi/4$.

In the following proposition, we use Lemma 12.3.1 in order to bound the density of $\mathrm{L}_{A_n}$ by essentially a savings factor $\sqrt{v}$ from the trivial bound $\mathrm{d}(\mathrm{L}_n)$, when $P^+(a)^{1+v} > a$ for all $a \in A_n$.

**Proposition 12.3.3.** *Let $A$ be a finite primitive set. Take $v \in (0,1)$, an integer $n > 1$ with $n \notin A$, and denote $q = P^+(n)$. If $P^+(a)^{1+v} > a$ for all $a \in A_n$ then*

$$\mathrm{d}(\mathrm{L}_{A_n}) \ \leq \ \sqrt{v}\, r_q\, \mathrm{d}(\mathrm{L}_n) \tag{12.13}$$

*for the ratio $r_q := M_q/m_q$ when $q \geq 3$, and $r_2 := r_3$.*

*Proof.* Without loss assume $A = A_n$. Then $ac \in \mathrm{L}_n$ for all $a \in A$, $c \in C_a^v$ (recall $p^-(ac) = p^-(a)$), and so $\mathrm{L}_{ac} \subset \mathrm{L}_n$. Note the condition $P^+(a)^{1+v} > a$ is equivalent to $P^+(a)^v > a^*$, and $v < 1$ implies $P^+(a) \nmid a^*$. By Lemma 12.3.1, we have the following (finite) disjoint union,

$$\mathrm{L}_n \ \supset \ \bigcup_{a\in A}\bigcup_{c\in C_a^v} \mathrm{L}_{ac}. \tag{12.14}$$

Thus taking the density of (12.14), we obtain

$$\mathrm{d}(\mathrm{L}_n) \geq \mathrm{d}\Big(\bigcup_{a\in A}\bigcup_{c\in C_a^v} \mathrm{L}_{ac}\Big) = \sum_{a\in A}\sum_{c\in C_a^v} \mathrm{d}(\mathrm{L}_{ac}) = \sum_{a\in A} \mathrm{d}(\mathrm{L}_a) \sum_{c\in C_a} \frac{1}{c}, \tag{12.15}$$

noting $P^+(a) > P^+(c)$ for $1 < c \in C_a^v$, so $\mathrm{L}_{ac} = \{bac : p^-(b) \geq P^+(a)\} = c \cdot \mathrm{L}_a$. Then by definitions of $C_a^v$ and $\mu_q$ in (12.10) and (12.8),

$$\begin{aligned}\sum_{c\in C_a^v} \frac{1}{c} \ &= \prod_{p\in[P^+(a^*),P^+(a^*)^{1/\sqrt{v}})} \Big(1-\frac{1}{p}\Big)^{-1} = \prod_{p<P^+(a^*)^{1/\sqrt{v}}} \Big(1-\frac{1}{p}\Big)^{-1} \prod_{p<P^+(a^*)} \Big(1-\frac{1}{p}\Big) \\ &= \frac{\log P^+(a^*)^{1/\sqrt{v}}}{\mu_{P^+(a^*)^{1/\sqrt{v}}}} \frac{\mu_{P^+(a^*)}}{\log P^+(a^*)} = \frac{\mu_{P^+(a^*)}}{\mu_{P^+(a^*)^{1/\sqrt{v}}}} \frac{1}{\sqrt{v}}. \end{aligned} \tag{12.16}$$

When $q \geq 3$, we use $\mu_{P^+(a^*)}/\mu_{P^+(a^*)^{1/\sqrt{v}}} \geq m_q/M_q = 1/r_q$, which follows by monotonicity of $m_q, M_q$ in Lemma 12.2.3, and that $P^+(a^*), q \in \mathcal{P}$. Hence plugging (12.16) back into (12.15),

$$\mathrm{d}(\mathrm{L}_n) \geq \frac{1}{\sqrt{v}\, r_q} \sum_{a\in A} \mathrm{d}(\mathrm{L}_a) = \frac{1}{\sqrt{v}\, r_q}\, \mathrm{d}(\mathrm{L}_A)$$

as desired.

The result similarly holds when $q = 2$: if $P^+(a^*) \geq 3$ then $\mu_{P^+(a^*)}/\mu_{P^+(a^*)^{1/\sqrt{v}}} \geq m_3/M_3 = 1/r_3$ as before. And if $P^+(a^*) = 2$ then $\mu_2/\mu_{2^{1/\sqrt{v}}} \geq 1$ also suffices. □

## 12.4 Deduction of Theorems 12.1.2, 12.1.3, 12.1.5

We now apply our analysis of the density of L-multiples to our original sum of interest $f(A) = \sum_{a\in A} \frac{1}{a\log a}$. First we need a simple lemma on bounding certain monotonic sequences.

**Lemma 12.4.1.** *For $k \geq 1$, let $c_0 \geq c_1 \geq \cdots \geq c_k \geq 0$ and $0 = D_0 \leq D_1 \leq \cdots \leq D_k$. If $d_1, \ldots, d_k \geq 0$ satisfy $\sum_{j\leq i} d_j \leq D_i$ for all $i \leq k$, then we have*

$$\sum_{i\leq k} c_i d_i \ \leq \ \sum_{i\leq k} c_i(D_i - D_{i-1}).$$

*Proof.* By rearranging sums,

$$\sum_{i\leq k} c_i d_i = \sum_{i\leq k} c_i\Big(\sum_{j\leq i} d_j - \sum_{j\leq i-1} d_j\Big) = \sum_{i\leq k-1} (c_i - c_{i+1}) \sum_{j\leq i} d_j \ + \ c_k \sum_{i\leq k} d_i.$$

Since $c_i \geq c_{i+1}$ and $\sum_{j\leq i} d_j \leq D_i$, we conclude

$$\sum_{i\leq k} c_i d_i \leq \sum_{i\leq k-1} (c_i - c_{i+1}) D_i \ + \ c_k D_k = \sum_{i\leq k} c_i(D_i - D_{i-1}).$$

□

To motivate the remainder of the proof, we offer a probabilistic interpretation of Proposition 12.3.3: for $v \geq 0$, consider $D(v) := \sup_A \mathrm{d}(\mathrm{L}_{A_n})/\mathrm{d}(\mathrm{L}_n)$, ranging over primitive sets $A$ such that $P^+(a)^{1+v} > a$ for all $a \in A$. Note $D(v)$ may be viewed as a 'cumulative distribution function', since $D(0) = 0$ and $D(v) \to 1$ as $v \to \infty$. Now Proposition 12.3.3 essentially bounds $D(v)$ by $\sqrt{v}$. Using the corresponding bound $1/2\sqrt{v}$ for the 'probability density function', we establish quantitative bounds below.

**Proposition 12.4.2.** *For any primitive set $A$, and any integer $n \notin A$ with $q = P^+(n) \geq 3$,*

$$f(A_n) \ \leq \ \frac{\pi}{4}\frac{M_q}{m_q^2}\, e^{\gamma}\mathrm{d}(\mathrm{L}_n).$$

*Proof.* Without loss, we may assume $A = A_n$ is finite, since $f(A) = \lim_{x\to\infty} f(A \cap [1, x])$. Also $n \notin A$ implies all elements of $A$ are composite.

Take $k \geq 1$ and any sequence $0 = v_0 < v_1 < \cdots < v_k = 1$ and partition the set $A = \bigcup_{0\leq i\leq k} A_{(i)}$, where $A_{(k)} = \{a \in A : P^+(a)^2 \leq a\}$ and for $0 \leq i \leq k$,

$$A_{(i)} = \{a \in A : P^+(a)^{1+v_i} \leq a < P^+(a)^{1+v_{i+1}}\}.$$

Then applying Lemma 12.2.4 to each $A_{(i)}$,

$$f(A) = \sum_{0\le i\le k} f(A_{(i)}) \ \le\ \frac{e^\gamma}{m_q} \sum_{0\le i\le k} \frac{\mathrm{d}(\mathrm{L}_{A_{(i)}})}{1+v_i}. \tag{12.17}$$

Note since $A$ is primitive, $\{\mathrm{L}_{A_{(i)}}\}_{i\le k}$ are pairwise disjoint. Also for each $j<k$, the first $j$ components are $\bigcup_{0\le i\le j} A_{(i)} = \{a\in A : a < P^+(a)^{1+v_{j+1}}\} =: A^{(j)}$, so by Proposition 12.3.3 they have density

$$\sum_{0\le i\le j} \mathrm{d}(\mathrm{L}_{A_{(i)}}) = \mathrm{d}(\mathrm{L}_{A^{(j)}}) \le \sqrt{v_{j+1}}\, r_q\, \mathrm{d}(\mathrm{L}_n).$$

Also for $j=k$ we have $\sum_{0\le i\le k} \mathrm{d}(\mathrm{L}_{A_{(i)}}) = \mathrm{d}(\mathrm{L}_A) \le \mathrm{d}(\mathrm{L}_n)$, which is trivially less than $r_q\mathrm{d}(\mathrm{L}_n)$. Let $c_i = \frac{1}{1+v_i}$, $d_i = \mathrm{d}(\mathrm{L}_{A_{(i)}})$, $D_i = \sqrt{v_{i+1}}\, r_q\, \mathrm{d}(\mathrm{L}_n)$ (here we let $v_{k+1}=v_k$ so that $D_k - D_{k-1} = 0$). Thus by Lemma 12.4.1 we have

$$\sum_{0\le i\le k} \frac{\mathrm{d}(\mathrm{L}_{A_{(i)}})}{1+v_i} = \sum_{0\le i\le k} c_i d_i \ \le\ \sum_{0\le i\le k} c_i(D_i - D_{i-1}) = r_q\, \mathrm{d}(\mathrm{L}_n) \sum_{0\le i\le k} \frac{\sqrt{v_{i+1}} - \sqrt{v_i}}{1+v_i}.$$

Hence the weighted sum in (12.17) is bounded by

$$f(A) \le \frac{r_q}{m_q}\, e^\gamma\, \mathrm{d}(\mathrm{L}_n) \sum_{1\le i\le k} \frac{\sqrt{v_i} - \sqrt{v_{i-1}}}{1+v_{i-1}}. \tag{12.18}$$

As (12.18) holds for any partition $0 = v_0 < v_1 < \cdots < v_k = 1$, we may set $v_i = \frac{i}{k}$ and obtain the corresponding integral,

$$\lim_{k\to\infty} \sum_{1\le i\le k} \frac{\sqrt{v_i} - \sqrt{v_{i-1}}}{1+v_{i-1}} = \lim_{k\to\infty} \sum_{1\le i\le k} \int_{v_{i-1}}^{v_i} \frac{\mathrm{d}}{\mathrm{d}v}\Big[\sqrt{v}\Big] \frac{\mathrm{d}v}{1+v_{i-1}} = \int_0^1 \frac{\mathrm{d}v}{2\sqrt{v}(1+v)} = \frac{\pi}{4}.$$

Hence we conclude $f(A) \le \frac{\pi}{4}\frac{r_q}{m_q}\, e^\gamma\, \mathrm{d}(\mathrm{L}_n)$ as desired. □

We illustrate the value of these bounds by deducing Theorem 12.1.3 in quantitative form.

**Corollary 12.4.2.1.** *Let $A$ be a primitive set, and take an odd prime $p$. If $p\notin A$ then we have $f(A_p) < .901\, f(p)$, and moreover $f(A_p) \le (\frac{\pi}{4} + o(1))f(p)$ as $p\to\infty$. In addition, if $p>23$ and $2p\notin A$ then $f(A_{2p}) < f(2p)$.*

*Proof.* For an odd prime $q$ define $b_q := \frac{\pi}{4}\frac{M_q}{m_q^2}\mu_q$. Then Proposition 12.4.2 shows that if $n\notin A$ we have

$$f(A_n) \le \frac{\pi}{4}\frac{M_q}{m_q^2} e^\gamma \mathrm{d}(\mathrm{L}_n) = \frac{q}{n}\, b_q f(q)$$

with $q = P^+(n) \geq 3$, recalling $\mathrm{d}(\mathrm{L}_n) = \frac{q}{n}\mathrm{d}(\mathrm{L}_q)$ and (12.9). In particular for $n = q, 2q$ we have $f(A_q) \leq b_q f(q)$ and $f(A_{2q}) \leq \frac{1}{2}b_q f(q)$. Note $\mu_q, m_q, M_q \sim 1$ implies $b_q \sim \frac{\pi}{4}$ as claimed. Also the first few values of $b_q$ are displayed below.

| $q$ | $b_q$ | $q$ | $b_q$ |
|---|---|---|---|
| 3 | 0.9006 | 23 | 0.8232 |
| 5 | 0.8795 | 29 | 0.8266 |
| 7 | 0.8507 | 31 | 0.8139 |
| 11 | 0.8564 | 37 | 0.8184 |
| 13 | 0.8327 | 41 | 0.8189 |
| 17 | 0.8491 | 43 | 0.8092 |
| 19 | 0.8305 | 47 | 0.8090 |

Observe for $q > 7$, we have

$$f(A_q) \leq \frac{\pi}{4}\Big(\frac{M_q}{m_{11}}\Big)^2 f(q) \leq \frac{\pi}{4}\Big(\frac{1 + 1/2\log(2\cdot 10^9)^2}{\mu_{19}}\Big)^2 f(q) < .879 f(q).$$

In particular, with the table, we see $f(A_q) < .901 f(q)$ for all $q > 2$ as claimed.

Finally, we note $f(A_{2q}) < f(2q)$ whenever $b_q < \frac{\log q}{\log(2q)}$. The result then follows since

$$b_q = \frac{\pi}{4}\frac{M_q}{m_q^2}\mu_q \ > \ \frac{\log q}{\log(2q)} \qquad \text{iff} \quad q \leq 23. \tag{12.19}$$

Indeed, this may be checked directly for $q < 47$. And for $q \geq 47$ we observe that $\log q/\log(2q) \geq \log 47/\log 94 \geq .847$ exceeds $b_q \leq \frac{\pi}{4}(M_{2\cdot 10^9}/m_{47})^2 \leq .834$. □

Importantly $b_q < 1$ for all odd $q$, which means every odd prime is Erdős strong. However it remains an open question whether $q = 2$ is Erdős strong. Now if $2 \in A$ we immediately deduce $f(A) \leq f(\mathcal{P}) = 1.6366\cdots$. Thus to complete the proof of Theorem 12.1.2, it suffices to assume $2 \notin A$.

We achieve this in the result below. The argument is somewhat similar in spirit to that of Theorem 1.1 and Lemma 2.4 in [90].

**Theorem 12.4.3.** *For any primitive set $A$ with $2 \notin A$, we have $f(A) < 1.60$.*

*Proof.* As $2 \notin A$, denote by $K \geq 2$ the exponent for which $2^K \in A$. Note $K$ is unique by primitivity (Also if $2^k \notin A$ for all $k$ let $K = \infty$, in which case let $f(2^K) = 0$). Partition $A$ into sets $A^0 = \{a \in A : 2 \nmid a\}$ and $A^k = \{a \in A : 2^k \| a\}$ for $k \geq 1$, and let $B^k = \{a/2^k : a \in A^k\}$. We have

$$\begin{aligned} f(A) &= f(2^K) + \sum_{p\in A} f(p) + \sum_{p\notin A} f(A_p) \\ &\leq f(2^K) + \sum_{\substack{p>2\\ p\in A}} f(p) + \sum_{\substack{p>2\\ p\notin A}} b_p f(p) + \sum_{\substack{p>2\\ p\notin A}}\sum_{k=1}^{K-1} f((A^k)_{2^k p}), \end{aligned} \tag{12.20}$$

since $f((A^0)_p) \le b_p f(p)$ if $p \notin A$ by Proposition 12.4.2. More generally, if $2^k p \notin A$ then

$$f((A^k)_{2^k p}) \le 2^{-k} f((B^k)_p) \le 2^{-k} b_p f(p).$$

By comparison if $2^k p \in A$ then $f((A^k)_{2^k p}) = f(2^k p) \le 2^{-k} f(p) \frac{\log p}{\log(2p)}$.

Observe that either $2^k p \notin A$ for all $k \ge 1$, or $2^J p \in A$ for a (unique) $J = J_p \in [1, K)$, in which case $(A^k)_{2^k p} = \emptyset$ for all $k > J$ by primitivity. Thus by (12.19), it suffices to assume $2^k p \notin A$ for all $k \ge 1$ when $p \le 23$, and $2^J p \in A$ for some $J \in [1, K)$ when $p > 23$, so

$$\begin{aligned}\sum_{\substack{p>2\\ p\notin A}} \sum_{k=1}^{K-1} f((A^k)_{2^k p}) &\le (1-2^{1-K}) \sum_{\substack{2<p\le 23\\ p\notin A}} b_p f(p) + \sum_{\substack{p>23\\ p\notin A}} f(p)\Big((1-2^{1-J})b_p + 2^{-J}\frac{\log p}{\log(2p)}\Big)\\ &\le (1-2^{1-K}) \sum_{\substack{2<p\le 23\\ p\notin A}} b_p f(p) + \sum_{\substack{p>23\\ p\notin A}} b_p f(p),\end{aligned}$$

since $2b_p > 1 > \frac{\log p}{\log(2p)}$ for all $p > 2$. Moreover $(2-2^{1-K})b_p \ge (2-1/2)\frac{\pi}{4} > 1.1$, so (12.20) becomes

$$\begin{aligned} f(A) &\le f(2^K) + \sum_{\substack{p>2\\ p\in A}} f(p) + (2-2^{1-K}) \sum_{\substack{2<p\le 23\\ p\notin A}} b_p f(p) + 2\sum_{\substack{p>23\\ p\notin A}} b_p f(p)\\ &\le f(2^K) + (2-2^{1-K}) \sum_{2<p\le 23} b_p f(p) + 2\sum_{p>23} b_p f(p)\\ &=: f(2^K) + (2-2^{1-K})C_1 + 2C_2. \end{aligned} \tag{12.21}$$

Now we compute the constants $C_1, C_2$. First, let $M = M_{2\cdot 10^9} = 1.001\cdots$. Recalling $\mu_p f(p) = e^\gamma \mathrm{d}(\mathrm{L}_p)$,

$$C_2 := \sum_{p>23} b_p f(p) = \frac{\pi}{4} e^\gamma \sum_{p>23} \frac{M_p}{m_p^2} \mathrm{d}(\mathrm{L}_p) \le \frac{\pi}{4} \frac{M e^\gamma}{\mu_{23}^2} \prod_{p\le 23} \left(1-\tfrac{1}{p}\right) = 0.251135\cdots, \tag{12.22}$$

since $\sum_{p>q} \mathrm{d}(\mathrm{L}_p) = \prod_{p\le q}(1-\frac{1}{p})$. Similarly we have

$$C_1 := \sum_{2<p\le 23} b_p f(p) = \sum_{2<p\le 23} \frac{\pi}{4}\frac{M}{m_p^2} e^\gamma \mathrm{d}(\mathrm{L}_p) = \frac{\pi}{4} M e^\gamma \cdot 0.39012\cdots = 0.5463\cdots \tag{12.23}$$

Here we computed

$$\sum_{2<p\le 23} \frac{1}{m_p^2}\mathrm{d}(\mathrm{L}_p) = \frac{1}{\mu_7^2}\sum_{2<p\le 7} \mathrm{d}(\mathrm{L}_p) + \frac{1}{\mu_{19}^2}\sum_{7<p\le 19} \mathrm{d}(\mathrm{L}_p) + \frac{1}{\mu_{23}^2}\mathrm{d}(\mathrm{L}_{23}) = 0.390126\cdots,$$

using $\sum_{q<p\le q'} \mathrm{d}(\mathrm{L}_p) = \prod_{p\le q}(1-\frac{1}{p}) - \prod_{p\le q'}(1-\frac{1}{p})$.

Hence plugging (12.22) and (12.23) back into (12.21),

$$\begin{aligned} f(A) &\le f(2^K) + (2-2^{1-K})C_1 + 2C_2 \\ &\le 2^{-K}\Big(\frac{1}{\log 4} - 2C_1\Big) + 2(C_1+C_2) \le 2(C_1+C_2) \le 1.595. \end{aligned} \tag{12.24}$$

Here we used $2C_1 > .722 > 1/\log 4$. This completes the proof. □

*Remark* 12.4.4. A similar argument as in Theorem 12.4.3 shows $f(A_2) < C_1 + C_2 < 0.80$ when $2 \notin A$. We leave this to the interested reader. Note this bound improves on $f(A_2) < e^\gamma/2 \approx 0.89$ from [90, Proposition 2.1], but unfortunately still exceeds $f(2) \approx 0.72$.

### 12.4.1 Proof of Theorem 12.1.5

Take $\varepsilon > 0$. We shall introduce large parameters $y = y_\varepsilon$, $k = k_{\varepsilon,y}$, and $x = x_{\varepsilon,k}$.

By Lemma 12.2.2, we have $f(A_n) \le e^\gamma \mathrm{d}(\mathrm{L}_n)$ for any integer $n \notin A$, $n > 1$, and when $y = y_\varepsilon \in \mathbb{R}$ is sufficiently large by Proposition 12.4.2 we have the sharper bound

$$f(A_n) \;\le\; (\frac{\pi}{4}e^\gamma + \varepsilon)\mathrm{d}(\mathrm{L}_n) \qquad \text{provided} \quad P^+(n) > y. \tag{12.25}$$

Next for $k = k_\varepsilon = k_{\varepsilon,y} \in \mathbb{N}$ sufficiently large we have the crude bound

$$\sum_{\substack{n\in\mathbb{N}_k \\ P^+(n)\le y}} \mathrm{d}(\mathrm{L}_n) \le \sum_{\substack{n\ge 2^k \\ P^+(n)\le y}} \frac{1}{n} < \varepsilon. \tag{12.26}$$

Indeed, using Rankin's trick,

$$R(u) := \sum_{\substack{n\le u \\ P^+(n)\le y}} 1 \le u^{1/2} \sum_{\substack{n\le u \\ P^+(n)\le y}} n^{-1/2} \le u^{1/2} \prod_{p\le y}(1-p^{-1/2})^{-1} \ll_y u^{1/2}, \tag{12.27}$$

so by partial summation,

$$\sum_{\substack{n>u \\ P^+(n)\le y}} \frac{1}{n} = -\frac{R(u)}{u} + \int_u^\infty R(t)t^{-2}\mathrm{d}t \ll_y u^{-1/2} + \int_u^\infty t^{-3/2}\mathrm{d}t \ll_y u^{-1/2}.$$

Then setting $u = 2^k - 1$ above, we obtain (12.26) for $k = k_{\varepsilon,y}$ sufficiently large.

Finally, since $f(\mathbb{N}_j) < 2$ crudely for all $j$ there exists $x = x_{\varepsilon,k} \in \mathbb{R}$ sufficiently large so that $f\big(\bigcup_{j\le k} \mathbb{N}_j \cap [x,\infty)\big) < \varepsilon$.

Now take a primitive set $A \subset [x,\infty)$, and consider the partition $A = A' \cup \bigcup_{n\in\mathbb{N}_k\setminus A} A_n$, where $A'$ consists of elements $a \in A$ with at most $k$ prime factors, and

each other element $a \in A$ (with at least $k+1$ prime factors) then lies in $A_n = A \cap \mathrm{L}_n$, where $n \notin A$ is the product of the smallest $k$ primes of $a$. Hence we conclude

$$\begin{aligned} f(A) &= f(A') + \sum_{n \in \mathbb{N}_k \setminus A} f(A_n) \\ &\le f\big( \bigcup_{j \le k} \mathbb{N}_j \cap [x, \infty) \big) + \sum_{\substack{n \in \mathbb{N}_k \setminus A \\ P^+(n) \le y}} f(A_n) + \sum_{\substack{n \in \mathbb{N}_k \setminus A \\ P^+(n) > y}} f(A_n) \\ &\le \varepsilon + e^\gamma \sum_{\substack{n \in \mathbb{N}_k \\ P^+(n) \le y}} \mathrm{d}(\mathrm{L}_n) + (\frac{\pi}{4} e^\gamma + \varepsilon) \sum_{\substack{n \in \mathbb{N}_k \\ P^+(n) > y}} \mathrm{d}(\mathrm{L}_n) \ \le \ \varepsilon + e^\gamma \varepsilon + (\frac{\pi}{4} e^\gamma + \varepsilon), \end{aligned}$$

by (12.25), (12.26), and noting $\sum_{n \in \mathbb{N}_k, P^+(n) > y} \mathrm{d}(\mathrm{L}_n) \le 1$. Hence letting $\varepsilon \to 0$ completes the proof of Theorem 12.1.5.

## 12.5 L-primitive sets revisited

### 12.5.1 Upper density

As mentioned in the introduction, one of the striking early results in the study of primitive sets was due to Besicovitch [10], who showed

$$\sup_{A \text{ primitive}} \overline{\mathrm{d}}(A) \ = \ \frac{1}{2}.$$

This came as quite a surprise at the time, in particular disproving a conjecture of Davenport. We shall extend this phenomenon further to L-primitive sets, in Proposition 12.5.2.

To proceed, we recall a result of Erdős [47], which bounds the density of the set of multiples of an interval. Also see Hall–Tenenbaum [71, Theorem 21] for quantitatively stronger results. Denote the set of (all) multiples of $A \subset \mathbb{N}$ as $\mathrm{M}_A = \{na : n \in \mathbb{N}, a \in A\}$.

**Proposition 12.5.1** (Erdős, 1936)**.** *Let $\varepsilon(x)$ be any function with $\varepsilon(x) \to 0$ as $x \to \infty$. Then the upper density of $\mathrm{M}_{(x^{1-\varepsilon(x)}, x]}$ tends to zero as $x \to \infty$.*

We prove a Besicovitch-type result for L-primitive sets, notably with full upper density.

**Proposition 12.5.2.** *We have $\sup_A \overline{\mathrm{d}}(A) = 1$ over L-primitive sets $A$.*

*Proof.* Take $h \in \mathbb{Z}_{>1}$, $\varepsilon > 0$ and let $S = \{n \in \mathbb{N} : P^+(n) \le h\}$ be the set of $h$-smooth numbers. For a sequence of indices $k_1, k_2, \dots$ to be determined, define intervals $I_i = (h^{k_i-1}, h^{k_i}]$. Let $S_i := I_i \setminus S$ and note for $a \in S_i$ and $n \in \mathrm{L}_a$ we have $n \ge P^+(a)a > h^{k_i}$, so $n \notin I_i \supset S_i$. In particular $a' \notin \mathrm{L}_a$ for distinct $a, a' \in S_i$, so each set $S_i$ is L-primitive. Now define the L-primitive set

$$A = \bigcup_{j\ge 1} S_j \setminus \bigcup_{1\le i<j} \mathrm{M}_{I_i}. \tag{12.28}$$

Recall $|S \cap [1,x]| \ll_h \sqrt{x}$ by (12.27). For each fixed $h > 1$, by Proposition 12.5.1 we see $\overline{\mathrm{d}}(\mathrm{M}_{(x/h,x]}) \to 0$ as $x \to \infty$. So for $k_i$ large enough, we may assume $\overline{\mathrm{d}}(\mathrm{M}_{I_i}) < \varepsilon/2^i$.

For each $i$ the set of multiples $\mathrm{M}_{I_i}$ is a periodic set with period (dividing) $(h^{k_i})!$. So assuming $k_{i+1} \ge (h^{k_i})!$ the relative density of $\mathrm{M}_{I_i}$ inside $I_{i+1}$ is at most $2\overline{\mathrm{d}}(\mathrm{M}_{I_i})$. Hence

$$\begin{aligned} |A \cap [1, h^{k_j}]| &\ge |I_j| - \big|S \cap [1, h^{k_j}]\big| - 2h^{k_j} \sum_{1\le i<j} \overline{\mathrm{d}}(\mathrm{M}_{I_i}) \\ &\ge (h^{k_j} - h^{k_j-1}) - O_h\big(h^{k_j/2}\big) - 2\varepsilon h^{k_j} \sum_{i\ge 1} 2^{-i}. \end{aligned}$$

Thus dividing by $x = h^{k_j}$ we see $\overline{\mathrm{d}}(A) = \limsup_{x\to\infty} |A \cap [1,x]|/x \ge 1 - 1/h - 2\varepsilon$. Taking $h \to \infty$ and $\varepsilon \to 0$ completes the proof. □

### 12.5.2 The Erdős L-primitive set conjecture

Sets of L-multiples play a central role in our proof of Theorem 12.1.2, as the mathematical structures arising from a probabilistic interpretation of (12.4),[6] and implicit in the original 1935 argument of Erdős [40].[7] As such it is natural to pose the L-primitive analogue of Conjecture 12.1.1, namely that $f(A) \le f(\mathcal{P})$ for all L-primitive sets $A$.

However, this conjecture turns out to be false.

**Proposition 12.5.3.** *We have*

$$\sup_{A \text{ L-primitive}} f(A) = \sum_p \max\{f(p), e^\gamma \mathrm{d}(\mathrm{L}_p)\}, \quad \textit{and} \quad \lim_{x\to\infty} \sup_{\substack{A\subset[x,\infty)\\ A \text{ L-primitive}}} f(A) = e^\gamma. \tag{12.29}$$

[6] A variant in [50], a set $A$ 'possesses property I' if there is no solution to $a' = ba$ for $a, a' \in A$ with $p^-(b) > P^+(a)$. This is similar to $A$ as L-primitive, but the latter imposes the inclusive inequality $p^-(b) \ge P^+(a)$, which arises naturally from a probabilistic viewpoint. This inclusivity leads to key structural properties, notably the trichotomy in Lemma 12.2.1.

[7] The author was also recently shown 'prefix-free sets' in [1], which coincides with L-primitive for sets of squarefree numbers.

Note the prime sum in (12.29) above is at least (and well-approximated by) $f(\mathcal{P}) - f(2) + e^\gamma/2 \approx 1.805$. In particular it exceeds $f(\mathcal{P}) \approx 1.636$. As such Proposition 12.5.3, along with Conjecture 12.7.1 and related work in the literature, highlights how the Erdős primitive set conjecture is quite fragile under certain seemingly natural directions of generalization.

We now proceed to set up the proof of Proposition 12.5.3. First, the trichotomy in Lemma 12.2.1 leads to the following.

**Lemma 12.5.4.** *Every set $S \subset \mathbb{N}$ has a unique* L*-primitive subset $\langle S \rangle$ with $\mathrm{L}_{\langle S \rangle} = \mathrm{L}_S$. In particular $\langle S \rangle = S$ if $S$ is* L*-primitive.*

*Proof.* For any $s_1, s_2 \in S$, by Lemma 12.2.1 either $\mathrm{L}_{s_1} \cap \mathrm{L}_{s_2} = \emptyset$ or $\mathrm{L}_{s_1} \subset \mathrm{L}_{s_2}$ (or vice versa). Thus each $s \in S$ has a (unique) smallest L-divisor $s' \in S$, inducing a map $S \to S : s \mapsto s'$. We define $\langle S \rangle$ as the image of this map. Explicitly this is

$$\langle S \rangle := \{s \in S : s \notin \mathrm{L}_t \ \forall\, t < s, t \in S\}. \tag{12.30}$$

By minimality $\mathrm{L}_{s_1} \cap \mathrm{L}_{s_2} = \emptyset$ for all $s_1, s_2 \in \langle S \rangle$, so $\langle S \rangle$ is L-primitive. Moreover $\mathrm{L}_S = \bigcup_{s \in S} \mathrm{L}_s = \bigcup_{s' \in \langle S \rangle} \mathrm{L}_{s'} = \mathrm{L}_{\langle S \rangle}$, where the latter union over $\langle S \rangle$ is disjoint by L-primitivity. This completes the proof. □

Next take $v > 0$, $n \in \mathbb{Z}_{>1}$, and consider the set $D_v(n)$ of prime divisors of $n$ whose induced L-divisor is not smooth, i.e.

$$D_v(n) = \Big\{p \mid n \ : \prod_{q^e \| n, q < p} q^e \ \le \ p^v\Big\}. \tag{12.31}$$

We cite the following result of Bovey, based on earlier work of Erdős [71, §1.2].

**Proposition 12.5.5** (Bovey, 1977)**.** *For each $v > 0$ there is a set $N_v \subset \mathbb{N}$ of full density with*

$$\frac{|D_v(n)|}{\log\log n} \ \to \ e^{-\gamma} \int_0^v \rho(x)\mathrm{d}x \tag{12.32}$$

*as $n \to \infty$ on $N_v$. Here $\rho$ is the Dickman–de Bruijn function.*

*Remark* 12.5.6. In probability, the righthand side of (12.32) is called the Dickman distribution.

In particular, $|D_v(n)| \gg_v \log\log n$ for all $n \in N_v$. Now we may define a map $\beta : N_u \to \mathbb{N}$ sending $n$ to its L-divisor $\beta(n) = p\prod_{q^e \| n, q<p} q^e$, for the largest prime $p \in D_v(n)$.

Define the L-primitive generating set $B(v) := \langle \beta(N_v) \rangle$ as in Lemma 12.5.4. By construction $\mathrm{L}_{B(v)} = \mathrm{L}_{\beta(N_v)} \supset N_u$ has full density. Also, by definition of $\beta, D_v$,

$$B(v) \subset \beta(N_v) \subset \{n \in \mathbb{N} : n \leq P^+(n)^{1+v}\}. \tag{12.33}$$

We are now prepared to establish a local version of Proposition 12.5.3.

**Proposition 12.5.7.** *For each prime $q$, we have*

$$\lim_{y\to\infty} \sup_{\substack{A\subset[y,\infty)\\ \text{L-primitive } A\not\ni q}} f(A_q) \;=\; \sup_{\text{L-primitive } A\not\ni q} f(A_q) \;=\; e^{\gamma}\mathrm{d}(\mathrm{L}_q).$$

*Proof.* By Lemma 12.2.2 we have $f(A_p) < e^{\gamma}\mathrm{d}(\mathrm{L}_p)$ for all L-primitive $A$ not containing $p$. It now suffices to provide L-primitive sets $B \subset [y,\infty)$ with $f(B_q) \to e^{\gamma}\mathrm{d}(\mathrm{L}_q)$ as $y \to \infty$.

Fix $v > 0$. The L-primitive set $B(v)$ in (12.33) satisfies

$$f(B(v)_q) = \sum_{b\in B(v)_q} \frac{1}{b\log b} \;\geq\; \frac{1}{1+v} \sum_{b\in B(v)_q} \frac{1}{b\log P^+(b)}. \tag{12.34}$$

Next, for $x > e^{e^{e^y}}$ we may assume $N_v \subset [x,\infty)$ and retain full density. Observe then $B(v) \subset [y,\infty)$ is our candidate L-primitive set. Indeed, for each $n \in N_v$ by construction $\beta(n)$ is divisible by all primes $q \in D_v(n)$, so $\beta(n)$ is composite with $\beta(n) \geq |D_v(n)| \gg_v \log\log n \geq \log\log\log x > y$ for each $b \in B(v)$, for $y$ sufficiently large. And note Mertens' product theorem gives

$$\mathrm{d}(\mathrm{L}_b) = \frac{1}{b}\prod_{p<P^+(b)}\Big(1-\frac{1}{p}\Big) = \frac{e^{-\gamma}+o_y(1)}{b\log P^+(b)}.$$

Plugging back into (12.34), we obtain

$$f(B(v)_q) \geq \frac{e^{\gamma}+o_y(1)}{1+v} \sum_{b\in B(v)_q} \mathrm{d}(\mathrm{L}_b).$$

Recall $\mathrm{L}_{B(v)} = \mathrm{L}_{\beta(N_v)} \supset N_v$ has full density, which implies $(\mathrm{L}_{B(v)})_q = \mathrm{L}_{B(v)_q}$ has full relative density $\mathrm{d}(\mathrm{L}_{B(v)_q}) = \mathrm{d}(\mathrm{L}_q)$. Hence by Corollary 12.2.1.1 this latter sum is

$$\sum_{b\in B(v)_q} \mathrm{d}(\mathrm{L}_b) = \mathrm{d}(\mathrm{L}_{B(v)_q}) = \mathrm{d}(\mathrm{L}_q).$$

Thus taking $y \to \infty$ and $v \to 0$ gives $f(B(v)_q) \to e^{\gamma}\mathrm{d}(\mathrm{L}_q)$ as desired. □

*Proof of Proposition 12.5.3.* Take L-primitive $A \subset [x,\infty)$, so $p \geq x$ for all $p \in A$. Then by Lemma 12.2.2,

$$\begin{aligned} f(A) = \sum_p f(A_p) &= \sum_{\substack{p<x \\ \text{or } p\notin A}} f(A_p) + \sum_{p\in A, p\geq x} f(p) \\ &\leq e^\gamma \sum_{\substack{p<x \\ \text{or } p\notin A}} \mathrm{d}(\mathrm{L}_p) + \sum_{p\in A, p\geq x} f(p) \\ &\leq e^\gamma \sum_p \mathrm{d}(\mathrm{L}_p) + \big(e^\gamma + o_x(1)\big) \sum_{p\geq x} \mathrm{d}(\mathrm{L}_p) \ \leq \ e^\gamma + o_x(1) \end{aligned}$$

by Mertens' theorem, and noting $\sum_p \mathrm{d}(\mathrm{L}_p) = 1$. Thus $\lim_x \sup_{A\subset[x,\infty)} f(A) \leq e^\gamma$. Equality in the limsup holds for the choice of $B = \bigcup_q B(v)_q$ and taking $v \to 0$ as in Proposition 12.5.7. Observe such $B$ inherits L-primitivity from the $B(v)_q$. Note in general, a union $B = \bigcup_q B_q$ is L-primitive if each $B_q$ is L-primitive. (By contrast $B = \bigcup_q B_q$ is not necessarily primitive even if each $B_q$ is primitive, e.g. $B = \{3, 6\}$.)

Next, consider the primes $\mathcal{Q} = \{q : f(q) > e^\gamma \mathrm{d}(\mathrm{L}_q)\} = \{q : 1/\log q > e^\gamma \prod_{p<q}(1 - \frac{1}{p})\}$. By Lemma 12.2.2 $f(A_q) < e^\gamma \mathrm{d}(\mathrm{L}_q)$ when $q \notin A$, so in general we have $f(A_q) < \max\{f(q), e^\gamma \mathrm{d}(\mathrm{L}_q)\}$ for all L-primitive $A$ and all primes $q$. Hence

$$f(A) = \sum_q f(A_q) < \sum_q \max\{f(q), e^\gamma \mathrm{d}(\mathrm{L}_q)\} = f(\mathcal{Q}) + e^\gamma\big(1 - \mathrm{d}(\mathrm{L}_\mathcal{Q})\big).$$

This bound is attained for the choice of $B' = \mathcal{Q} \cup \bigcup_{q\notin\mathcal{Q}} B(v)_q$, and taking $v \to 0$ as in Proposition 12.5.7. Again $B'$ inherits L-primitivity from the $B(v)_q$, as desired. □

## 12.6 Deduction of Theorems 12.1.8, 12.1.10

Our study of sets of L-multiples leads to Theorem 12.1.8, refining Davenport–Erdős. This in turn enables the proof of Theorem 12.1.10, by a modification of the argument in [48, Theorem 2], with greater care given to the constants involved.

To proceed we first establish some lemmas.

**Lemma 12.6.1.** *For any* L*-primitive* $A \subset \mathbb{N}$*, we have* $\underline{\mathrm{d}}(\mathrm{L}_A) \geq \sum_{a\in A} \mathrm{d}(\mathrm{L}_a)$*. Moreover if* $\sum_{a\in A} 1/a < \infty$ *then the natural density* $\mathrm{d}(\mathrm{L}_A)$ *exists and equals* $\sum_{a\in A} \mathrm{d}(\mathrm{L}_a)$*.*

*Proof.* For each $a \in A$ we have $\underline{\mathrm{d}}(\mathrm{L}_a) = \mathrm{d}(\mathrm{L}_a)$. So taking the lower density of the finite (disjoint) union $\bigcup_{a\in A, a\leq x} \mathrm{L}_a \subset \mathrm{L}_A$, we have $\sum_{a\in A, a\leq x} \mathrm{d}(\mathrm{L}_a) \leq \underline{\mathrm{d}}(\mathrm{L}_A)$ for all $x > 1$. Thus $\sum_{a\in A} \mathrm{d}(\mathrm{L}_a) \leq \underline{\mathrm{d}}(\mathrm{L}_A)$. Moreover if $\sum_{a\in A} 1/a < \infty$, then for all $y > 1$

$$\frac{1}{x} \sum_{\substack{n\leq x \\ n\in \mathrm{L}_{A\cap(y,\infty)}}} 1 \leq \frac{1}{x} \sum_{a\in A, a>y} \left\lfloor \frac{x}{a} \right\rfloor \leq \sum_{a\in A, a>y} \frac{1}{a} = o_y(1).$$

Thus $\overline{\mathrm{d}}(\mathrm{L}_{A\cap(y,\infty)}) \to 0$ as $y \to \infty$, and so combining with $\mathrm{d}(\mathrm{L}_{A\cap[1,y]}) = \sum_{a\in A, a\le y} \mathrm{d}(\mathrm{L}_a)$ completes the proof. □

The following lemma shows that sets of L-multiples have a log density, refining Davenport–Erdős' elementary proof for sets of (all) multiples [31].

**Lemma 12.6.2.** *For any* L*-primitive* $A \subset \mathbb{N}$*, the* log *density* $\delta(\mathrm{L}_A)$ *exists and equals* $\sum_{a\in A} \mathrm{d}(\mathrm{L}_a)$.

*Proof.* In general $\underline{\mathrm{d}}(S) \le \underline{\delta}(S) \le \overline{\delta}(S) \le \overline{\mathrm{d}}(S)$ for any $S \subset \mathbb{N}$. So for $S = \mathrm{L}_A$, by Lemma 12.6.1 it suffices to show

$$\sum_{a\in A} \mathrm{d}(\mathrm{L}_a) \ \ge\ \overline{\delta}(\mathrm{L}_A). \tag{12.35}$$

To this, for $y > 1$ let $A^y = \{a \in A : P^+(a) \le y\}$ and $L^y = \{n \in \mathrm{L}_A : P^+(n) \le y\}$. Note $L^y \subset \mathrm{L}_{A^y}$. Also $\sum_{a\in A^y} \frac{1}{a} \le \prod_{p\le y}(1-\frac{1}{p})^{-1} = O_y(1)$, so by Lemma 12.6.1 $\mathrm{d}(\mathrm{L}_{A^y})$ exists and equals $\sum_{a\in A^y} \mathrm{d}(\mathrm{L}_a)$ for all $y > 1$. In particular $\mathrm{d}(\mathrm{L}_{A^y}) \to \sum_{a\in A} \mathrm{d}(\mathrm{L}_a)$ as $y \to \infty$.

Now observe each $n \in \mathrm{L}_{A^y}$ is a L-multiple of a unique $a \in A^y$, so for $x \ge y > 1$ we have

$$\begin{aligned}\sum_{n\in L^x\cap \mathrm{L}_{A^y}} \frac{1}{n} &= \sum_{a\in A^y} \frac{1}{a} \prod_{P^+(a)\le p\le x} (1-\tfrac{1}{p})^{-1} = \sum_{a\in A^y} \frac{1}{a} \prod_{p<P^+(a)} (1-\tfrac{1}{p}) \prod_{p\le x} (1-\tfrac{1}{p})^{-1} \\ &= \mathrm{d}(\mathrm{L}_{A^y}) \prod_{p\le x} (1-\tfrac{1}{p})^{-1}. \end{aligned} \tag{12.36}$$

In particular for $x = y$ we have $\sum_{n\in L^x} \frac{1}{n} = \mathrm{d}(\mathrm{L}_{A^x}) \prod_{p\le x}(1-\frac{1}{p})^{-1}$.

Then for all $x \ge y > 1$, by (12.36) and Mertens' theorem

$$\begin{aligned}\sum_{n\in L^x\setminus \mathrm{L}_{A^y}} \frac{1}{n} &= \sum_{n\in L^x} \frac{1}{n} - \sum_{n\in L^x\cap \mathrm{L}_{A^y}} \frac{1}{n} \\ &= \big(\mathrm{d}(\mathrm{L}_{A^x}) - \mathrm{d}(\mathrm{L}_{A^y})\big) \prod_{p\le x} (1-\tfrac{1}{p})^{-1} \ \ll\ (\log x)\big(\mathrm{d}(\mathrm{L}_{A^x}) - \mathrm{d}(\mathrm{L}_{A^y})\big). \end{aligned} \tag{12.37}$$

Recall the natural density $\mathrm{d}(\mathrm{L}_{A^y})$ exists, in which case equals the log density $\delta(\mathrm{L}_{A^y})$. Hence by (12.37), for each $y > 1$ the upper log density is

$$\begin{aligned}\overline{\delta}(\mathrm{L}_A) \ &=\ \limsup_{x\to\infty} \frac{1}{\log x} \sum_{\substack{n\le x\\ n\in \mathrm{L}_A}} \frac{1}{n} \le \lim_{x\to\infty} \frac{1}{\log x} \sum_{\substack{n\le x\\ n\in \mathrm{L}_{A^y}}} \frac{1}{n} \ +\ \limsup_{x\to\infty} \frac{1}{\log x} \sum_{n\in L^x\setminus \mathrm{L}_{A^y}} \frac{1}{n} \\ &= \delta(\mathrm{L}_{A^y}) \ +\ \lim_{x\to\infty} O\big(\mathrm{d}(\mathrm{L}_{A^x}) - \mathrm{d}(\mathrm{L}_{A^y})\big) \\ &= \mathrm{d}(\mathrm{L}_{A^y}) \ +\ O\Big(\sum_{a\in A} \mathrm{d}(\mathrm{L}_a) - \mathrm{d}(\mathrm{L}_{A^y})\Big). \end{aligned} \tag{12.38}$$

Hence $\mathrm{d}(\mathrm{L}_{A^y}) \to \sum_{a\in A} \mathrm{d}(\mathrm{L}_a)$ as $y \to \infty$ implies $\overline{\delta}(\mathrm{L}_A) \le \sum_{a\in A} \mathrm{d}(\mathrm{L}_a)$, giving (12.35). □

**Theorem 12.1.8.** *If $\overline{\delta}(A) > 0$, then $A$ contains an infinite* L*-divisibility chain.*

*Proof.* We claim all such $A \subset \mathbb{N}$ contain an element $a \in A$ such that $A \cap \mathrm{L}_a$ has positive upper log density. (In other words, if $\overline{\delta}(A) > 0$ then there exists an element $a \in A$ such that $\overline{\delta}(A \cap \mathrm{L}_a) > 0$.)

Assume this claim holds. Letting $A^1 = A$, $a_1 = a$, and for $i \ge 1$ suppose $\overline{\delta}(A^i) > 0$. By the claim there exists $a_i \in A^i$ such that $A^{i+1} := A^i \cap \mathrm{L}_{a_i}$ has positive upper log density. Hence by induction, we obtain an L-divisibility chain $a_1, a_2, \cdots$, as desired.

Thus it remains to establish the above claim. For sake of contradiction, suppose $A \cap \mathrm{L}_a$ has zero log density for all $a \in A$. Next, for the L-primitive generating set $B = \langle A\rangle$ by Lemma 12.6.1 $\delta(\mathrm{L}_B) = \sum_{b\in B} \mathrm{d}(\mathrm{L}_b)$ exists. Then for $z > 1$ large enough we have $\delta(\mathrm{L}_{B\cap(z,\infty)}) = \sum_{b\in B, b>z} \mathrm{d}(\mathrm{L}_b) < \overline{\delta}(A)$. Now by assumption $\overline{\delta}(A \cap \mathrm{L}_b) = 0$ for all $b \le z$, $b \in B$, and so

$$\overline{\delta}(A) = \overline{\delta}(A \cap \mathrm{L}_{B\cap(z,\infty)}) \ \le \ \delta(\mathrm{L}_{B\cap(z,\infty)}) < \overline{\delta}(A),$$

a contradiction. Hence there exists $a \in A$ such that $A \cap \mathrm{L}_a$ has positive upper log density. □

**Theorem 12.1.10.** *If $\overline{\Delta}(A) > 0$, then there is an infinite* L*-divisibility chain $D \subset A$ of growth*

$$\limsup_{y\to\infty} \sum_{\substack{d\in D\\ d\le y}} \frac{1}{\log\log y} \ \ge \ \frac{\overline{\Delta}(A)}{e^\gamma}.$$

*Proof.* Take $\varepsilon > 0$. Without loss, we may suppose $A \subset [x_\varepsilon, \infty)$ for $x_\varepsilon$ sufficiently large, so that by Proposition 12.5.3 $f(A') \le e^\gamma + \varepsilon$ for all L-primitive subsets $A' \subset A$.

By definition of upper log log density $\Delta := \overline{\Delta}(A) > 0$, there exists an unbounded sequence $(x_j)_{j=0}^\infty \subset \mathbb{R}$ such that for all $j \ge 0$,

$$f(A \cap [1, x_j]) = \sum_{\substack{a\in A\\ a\le x_j}} \frac{1}{a\log a} > (\Delta - \varepsilon)\log\log x_j. \tag{12.39}$$

Recall the L-primitive generating set $\langle S\rangle = \{s \in S : s \notin \mathrm{L}_t \ \forall t < s, t \in S\}$ of a set $S \subset \mathbb{N}$ from Lemma 12.5.4. We partition $A = \bigcup_{i\ge 0} A^i$ into a disjoint collection of L-primitive subsets, where $A^0 = \langle A\rangle$ and inductively $A^l = \langle A \backslash \bigcup_{i<l} A^i\rangle$. By construction

each $a = a_l \in A^l$ has a (finite) chain of L-divisors $a_i \in A^i$ with $\mathrm{L}_{a_0} \supset \cdots \supset \mathrm{L}_{a_l} = \mathrm{L}_a$. Also note $f(A^i) \le e^\gamma + \varepsilon$ by assumption, so in particular $A^i$ has zero log log density. Hence (12.39) implies each $A^i$ in $A = \bigcup_{i\ge 0} A^i$ is non-empty. Next, define the subset $B = \bigcup_{j\ge 0} B_j$ for

$$B_j \;:=\; A \cap [1, x_j] \setminus \bigcup_{1\le i<r_j} A^i, \qquad \text{where} \quad r_j := \frac{\Delta - 2\varepsilon}{e^\gamma + \varepsilon} \log\log x_j.$$

Note the sets $B_j$ are pairwise disjoint: Indeed, since $A = \bigcup_{i\ge 0} A^i$, for each $j$ we have $A \cap [1, x_j] \subset \bigcup_{i<s_j} A^i$ for some finite $s_j$, as determined by $x_j$. Then since $(x_j)_j$ is unbounded, (passing to a subsequence) we have $r_{j+1} > s_j$ and so $A \cap [1, x_j] \subset \bigcup_{i<r_{j+1}} A^i$. Thus $B_j = A \cap [1, x_j] \setminus \bigcup_{i<r_j} A^i \;\subset\; \bigcup_{r_j\le i<r_{j+1}} A^i$ inherits disjointness from the $A^i$, as claimed.

Since $B = \bigcup_{j\ge 0} B_j$ forms a disjoint union, for each $b \in B$ there is a unique index $J(b)$ such that $b \in B_{J(b)}$, that is,

$$b \;\le\; x_{J(b)} \quad \text{and} \quad b \notin \bigcup_{i<r_{J(b)}} A^i. \tag{12.40}$$

In addition, $B$ has positive upper log log density, since by definitions of $B$, $r_j$, and (12.39),

$$\begin{aligned} f(B \cap [1, x_j]) \;\ge\; f(B_j) \;&\ge\; f(A \cap [1, x_j]) \;-\; \sum_{i<r_j} f(A^i) \\ &>\; (\Delta - \varepsilon) \log\log x_j - r_j(e^\gamma + \varepsilon) \;=\; \varepsilon \log\log x_j. \end{aligned}$$

In particular $B$ has positive upper log density, so by Theorem 12.1.8 there exists an *infinite* L–divisibility chain $D \subset B$. Since $D := (d_k)_{k=0}^\infty$ is unbounded, (by passing to a subchain) we may assume $J(d_k) < J(d_{k+1})$ for all $k \ge 0$. Recall each $a \in A^i$ is at the end of an L-divisibility chain of length $i$. As $b \in B_{J(b)}$ and $B_j$ is contained in $\bigcup_{r_j\le i<r_{j+1}} A^i$, we infer each $d \in D \subset B$ is at the end of an L-divisibility chain of length (at least) $r_{J(d)}$. Write it as $c_0^{(k)} \mid c_1^{(k)} \mid \cdots \mid c_{r_{J(d_k)}}^{(k)} = d_k$, with

$$\mathrm{L}_{c_0^{(k)}} \supset \cdots \supset \mathrm{L}_{d_k}.$$

Now let $i_k$ be the least index such that $c_{i_k}^{(k)} > d_{k-1}$ and define

$$C := \{d_{k-1} < c_i^{(k)} \le d_k \;:\; k, i \ge 0\} \;=\; \bigcup_{k\ge 0} \{c_i^{(k)} \;:\; i \in [i_k, r_{J(d_k)}]\}.$$

We may assume $(r_j)_j$ grows fast enough, so that $\lfloor \varepsilon\, r_{J(d_k)} \rfloor > d_{k-1}$. Then the trivial bound $c_i^{(k)} > i$ implies $c^{(k)}_{\lfloor \varepsilon\, r_{J(d_k)} \rfloor} > d_{k-1}$, and so $\lfloor \varepsilon\, r_{J(d_k)} \rfloor \geq i_k$. Thus

$$\big|C \cap [1, x_{j(d_k)}]\big| \;\geq\; \big|C \cap [d_{k-1}, d_k]\big| \;\geq\; (1-\varepsilon) r_{J(d_k)} = (1-\varepsilon)\frac{\Delta - 2\varepsilon}{e^\gamma + \varepsilon} \log\log x_{j(d_k)}. \tag{12.41}$$

Hence taking $\varepsilon \to 0$ in (12.41) above gives $\limsup_{x\to\infty} \sum_{c\in C, c\leq x} 1/\log\log x \geq \Delta/e^\gamma$ as desired.

Finally note $C$ forms an infinite L-divisibility chain: for each $k$ we have $c_j^{(k)} \in \mathrm{L}_{c_i^{(k)}}$ for all $i_k \leq i < j$, in particular $d_k \in \mathrm{L}_{c_i^{(k)}}$. Also $d_k \in \mathrm{L}_{d_{k-1}}$ since $D$ is an L-divisibility chain, so there exist factorizations

$$d_k = g c_i^{(k)} = h d_{k-1},$$

with $p^-(g) \geq P^+(c_i^{(k)})$ and $p^-(h) \geq P^+(d_{k-1})$. As $c_i^{(k)} > d_{k-1}$, we deduce $c_i^{(k)} \in \mathrm{L}_{d_{k-1}}$. Thus the $k$th and $(k-1)$th pieces of $C$ are linked together. Hence $C$ is indeed an L-divisibility chain. □

## 12.7 Closing remarks

In this discussion, we attempt to sample just a few of the multitude of open questions that have quickly arisen in connection with the Erdős primitive set conjecture. We have already described a few in the introduction, including Conjecture 12.1.4, as well as whether $p = 2$ is Erdős strong. We also note recent work has studied variants of the problem in function fields $\mathbb{F}_q[x]$, see [64], [65]. In addition, it would be interesting to further extend the classical study of sets of (all) multiples and of primitive sets, e.g. see Hall [70] or Halberstam–Roth [69, §5], to sets of L-multiples and L-primitive sets.

We conclude with a related question of Banks and Martin, which offers a potential unified framework to view the results described in this chapter. For $k \geq 1$, recall $\mathbb{N}_k = \{n : \Omega(n) = k\}$, in particular $\mathbb{N}_1 = \mathcal{P}$. In 1993, Zhang [143] proved $f(\mathbb{N}_k) < f(\mathcal{P})$ for each $k > 1$. Later Bayless, Kinlaw, and Klyve [7] showed that $f(\mathbb{N}_2) > f(\mathbb{N}_3)$. Banks and Martin [6] predicted $f(\mathbb{N}_k) < f(\mathbb{N}_{k-1})$ for each $k > 1$. In fact, they posed a vast generalization to Conjecture 12.1.1.

**Conjecture 12.7.1** (odd Banks–Martin)**.** Let $k \geq 1$ and suppose $A$ is a primitive set with $\Omega(n) \geq k$ for all $n \in A$. Then for any set of odd primes $\mathcal{Q}$, we have

$$f(A(\mathcal{Q})) \;\leq\; f\big(\mathbb{N}_k(\mathcal{Q})\big). \tag{12.42}$$

Here $A(\mathcal{Q})$ denotes the set of members of $A$ composed of primes in $\mathcal{Q}$.

Banks and Martin managed to show (12.42) in the special case when the set of primes $\mathcal{Q}$ is quite sparse, namely $\sum_{p\in\mathcal{Q}} 1/p < 1.74$ (even when $2 \in \mathcal{Q}$). We note the original formulation of Conjecture 12.7.1 included the cases $2 \in \mathcal{Q}$, but this turns out to be false. Indeed, when $\mathcal{Q} = \mathcal{P}$ it was shown $f(\mathbb{N}_k) > f(\mathbb{N}_6)$ for each $k \neq 6$, as in Chapter 7. In fact, numerical evidence suggests that in fact the reverse holds $f(\mathbb{N}_k) > f(\mathbb{N}_{k-1})$ for $k > 6$. Nevertheless for $\mathcal{Q} = \mathcal{P} \setminus \{2\}$, the desired inequality $f(\mathbb{N}_k(\mathcal{Q})) < f(\mathbb{N}_{k-1}(\mathcal{Q}))$ holds up to at least $k = 20$.

Observe that Theorem 12.1.3 implies Conjecture 12.7.1 in the special case $k = 1$. Indeed, if $p \notin \mathcal{Q}$ then $A(\mathcal{Q})_p = \emptyset$, so we deduce $f(A(\mathcal{Q})) = \sum_{p\in\mathcal{Q}} f(A(\mathcal{Q})_p) \leq \sum_{p\in\mathcal{Q}} f(p) = f(\mathcal{Q})$. Moreover, if true, Conjecture 12.7.1 implies Conjecture 12.1.4 of Erdős–Sárközy–Szemerédi. This follows by an argument similar to Theorem 12.1.5, and using $f(\mathbb{N}_k(\mathcal{Q})) \to 1/2$ as $k \to \infty$ when $\mathcal{Q} = \mathcal{P} \setminus \{2\}$, see Corollary 7.4.1.1. We leave this to the interested reader.

# Appendix A

# Mathematica code

Chapter 1: *A modification of the linear sieve, and the count of twin primes*

The Mathematica code used, as well as the package 'SieveFunctions.m' by William F. Galway (2009), are available at `arxiv.org/abs/2109.02851`

Chapter 7: *Almost primes and the Banks–Martin conjecture*

We compute $f(\mathcal{P}^{(k)}) = \int_1^\infty P_k(s)\mathrm{d}s$ using zeta functions $P_k(s) = \sum_{\Omega(n)=k} n^{-s}$ for $s > 1$ and $k \leq 20$. These in turn are computed by the recursion relation in Proposition 5.2.1, $P_k(s) = \frac{1}{k}\sum_{j=1}^k P(js)P_{k-j}(s)$.

```
P[k_Integer,s_]:= If[k==1,PrimeZetaP[s],
    Expand[(Sum[P[1,j*s]*P[k-j,s],j,1,k-1]+P[1,k*s])/k]]
Do[Print[k," ",NIntegrate[P[k,s],s,1,Infinity, WorkingPrecision->30,
    AccuracyGoal -> 13,PrecisionGoal -> 13]], k, 20]
```

Chapter 11: *Translated sums of primitive sets*

We compute the constants $s_k$, $t_k$, $s'_k$, $\sigma_k$, $h_k$, for $k \leq 20$, as follows:

```
P[k_Integer,s_]:= If[k==1,PrimeZetaP[s],
    Expand[(Sum[P[1,j*s]*P[k-j,s],j,1,k-1]+P[1,k*s])/k]]
FindRoot[P[1, s] == (k!)^(1/(k - 1)), s, 1 + 1/k^3]
FindRoot[P[k, t]/(2^(-t) + 3^(-t)) == 1, t, 1 + 1/k^3]
FindRoot[P[k - 1, s1] == 1, s1, 1 + 1/k^3]
FindRoot[P[k, sigma] == P[1, sigma], sigma, 1 + 1/k^3]]
FindRoot[NIntegrate[(P[1, s] - P[k, s])/y^s,s,1,Infinity,
    WorkingPrecision->30,
    AccuracyGoal->13, PrecisionGoal->13], y, 1 + 1/k^3]
```

# Bibliography


[1] R. Ahlswede, L.H. Khachatrian, A. Sárközy, *On prefix-free and suffix-free sequences of integers*, Numbers, Information and Complexity, (2000), 1–16.

[2] W. R. Alford, A. Granville, C. Pomerance. *There are infinitely many Carmichael numbers*, Ann. of Math., **139** (1994), 703–722.

[3] R. Arratia, A. D. Barbour, and S. Tavaré, *Logarithmic combinatorial structures: a probabilistic approach*, EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich, 2003.

[4] R. C. Baker, G. Harman, *Shifted primes without large prime factors*, Acta Arith. **83** (1998), 331–360.

[5] A. Balog, $p + a$ *without large prime factors*, Sém. Théorie des Nombres Bordeaux, (1983-84), exposé 31.

[6] W. Banks, G. Martin, *Optimal primitive sets with restricted primes*, Integers **13** (2013), #A69, 10 pp.

[7] J. Bayless, P. Kinlaw, D. Klyve, *Sums over primitive sets with a fixed number of prime factors*, Math. Comp. (2019).

[8] J. Bayless, P. Kinlaw, and J. D. Lichtman, *Higher Mertens constants for almost primes*, J. Number Theory **234** (2022), 448–475.

[9] F. Behrend, *On sequences of numbers not divisible by one another*, J.London Math. Soc. **10**(1935), 42–44.

[10] A. S. Besicovitch, *On the density of certain sequences of integers*, Math. Ann. **110** (1934), 336–341.

[11] E. Bombieri, *On the large sieve*, Mathematika, **12** (1965), 201–225.

[12] E. Bombieri, H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. Ser. A **239** (1966), 1–18.

[13] E. Bombieri, J. Friedlander, H. Iwaniec, *Primes in arithmetic progressions to large moduli.*, Acta Math., **156** (1986),203–251.

[14] E. Bombieri, J. Friedlander, H. Iwaniec, *Primes in arithmetic progressions to large moduli. II*, Math. Ann., 277(3): (1987) 361–393.

[15] E. Bombieri, J. Friedlander, H. Iwaniec, *Primes in arithmetic progressions to large moduli. III*, J. Amer. Math. Soc., **2** (1989),215–224.

[16] S. Broadbent, et al. *Sharper bounds for the Chebyshev function* $\theta(x)$, Math. Comp. (2021).

[17] Y.C. Cai, M.G. Lu, *On the upper bound for* $\pi_2(x)$, Acta Arith. **110** (2003), 275–298.

[18] R. D. Carmichael, *On Euler's* $\phi$*-function*, Bull. A.M.S. **13** (1907), 241–243.

[19] T.H. Chan, On sets of integers, none of which divides the product of $k$ others, *European J. Comb.* **32** (2011), 443–447.

[20] T.H. Chan, E. Győri, A. Sárközy, On a problem of Erdős on integers, none of which divides the product of $k$ others, *European J. Comb.* **31** (2010), 260–269.

[21] T. H. Chan, J. D. Lichtman, C. Pomerance, *A generalization of primitive sets and a conjecture of Erdős*, Discrete Analysis 2020:16, 13 pp.

[22] T. H. Chan, J. D. Lichtman, C. Pomerance, *On the critical exponent for k-primitive sets*, , Combinatorica (2021), 19 pp.

[23] J.R. Chen, *On the representation of a larger even integer as the sum of a prime and the product of at most two primes*, Sci. Sinica **16** (1973), 157–176.

[24] J.R. Chen, *On the Goldbach's problem and the sieve methods*, Sci. Sinica **21** (1978), 701–739.

[25] J. Chinis, *Siegel Zeros and Sarnak's Conjecture*, (2021) `arXiv:2105.14653`

[26] S. Chowla. *The Riemann hypothesis and Hilbert's tenth problem.* Mathematics and Its Applications, Gordon and Breach Science Publishers, (1965).

[27] H. Cohen, *High precision computation of Hardy–Littlewood constants*, preprint (1991), `https://www.math.u-bordeaux.fr/~hecohen/`.

[28] R. Crandall, C. Pomerance, *Prime numbers: a computational perspective*, Springer-Verlag, New York, 2001.

[29] D. Crisan, R. Erban, *On the counting function of semiprimes*, preprint, (2020) `arXiv:2006.16491`

[30] H. Davenport, P. Erdős, *On sequences of positive integers*, Acta Arith. **2** (1937), 147–151.

[31] H. Davenport, P. Erdős, *On sequences of positive integers*, J. Indian Math. Soc. **15** (1951), 19–24.

[32] H. Delange, *Sur des formules de Atle Selberg*, Acta Arith. **19** (1971), 105–146.

[33] J.-M. Deshouillers, H. Iwaniec, *Kloosterman sums and Fourier coefficients of cusp forms*, Invent. Math., **70** (1982), 219–288.

[34] H. G. Diamond, *Elementary methods in the study of the distribution of prime numbers*, Bull. Amer. Math. Soc. 7.3 (1982): 553-589.

[35] R. J. Duffin, *Representation of Fourier integrals as sums, III.* Proc. Amer. Math. Soc. **8** (1957), 272–277.

[36] P. Dusart, *Estimates of some functions over primes without RH*, preprint, (2010) `arXiv:1002.0442`

[37] P. Dusart, *Explicit estimates of some functions over primes*, Ramanujan J. **45** (2018), 227–251.

[38] P. D. T. A. Elliott, *On the correlation of multiplicative functions*, Notas Soc. Mat. Chile, **11** (1992), 1–11.

[39] P. D. T. A. Elliott, H. Halberstam, *A conjecture in prime number theory*, Symposia Mathematica, Vol. IV, Academic Press, London, (1970), 59–72.

[40] P. Erdős, *Note on sequences of integers no one of which is divisible by any other*, J. London Math. Soc. **10** (1935), 126–128.

[41] P. Erdős, *On the normal number of prime factors of $p-1$ and some related problems concerning Euler's $\varphi$-function*, Quart. J. Math., **6** (1935), 205–213.

[42] P. Erdős, On sequences of integers no one of which divides the product of two others and on some related problems, *Tomsk. Gos. Univ. Ucen. Zap.* **2** (1938), 74–82.

[43] P. Erdős, On the integers having exactly $k$ prime factors, *Annals of Math.* **49** (1948), 53–66.

[44] P. Erdős, *On pseudoprimes and Carmichael numbers*, Publ. Math. Debrecen, **4** (1956), 201–206.

[45] P. Erdős, A. Sárközy, *On the number of prime factors of integers*, Acta Sci. Math. **42** (1980), 237–246.

[46] P. Erdős, *Problèmes et résultats en théorie des nombres*, Conférence de Paul Erdős à l'Université de Limoges, le 21 Octobre 1986 (rédigé par F. Morain), 22 pp. `http://www.lix.polytechnique.fr/Labo/Francois.Morain/Introuvables/Morain88.pdf`

[47] P. Erdős, *A Generalization of a Theorem of Besicovitch*, J. London Math. Soc. **11** (1936), 92–98.

[48] P. Erdős, A. Sárközy, E. Szemerédi, *On divisibility properties of sequences of integers*, Studia Sci. Math. Hungar., **1** (1966), 431–435.

[49] P. Erdős, A. Sárközy, E. Szemerédi, On an extremal problem concerning primitive sequences, J. London Math. Soc. **42** (1967), 484-488.

[50] P. Erdős, A. Sárközy, E. Szemerédi, *On the solvability of some equations in dense sequences of integers*, Soviet Math. Dokl., **8** (1967), 1160–1164.

[51] P. Erdős, A. Sárközy, E. Szemerédi, *On divisibility properties of sequences of integers*, Colloq. Math. Soc. János Bolyai, **2** (1968), 35–49.

[52] P. Erdős, G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* **2** (1935), 463–470.

[53] P. Erdős, Z. Zhang, *Upper bound of* $\sum 1/(a_i \log a_i)$ *for primitive sequences*, Proc. Amer. Math. Soc. **117** (1993), 891–895.

[54] W. Feller, *An introduction to probability theory and its applications. Vol. II*, second ed., John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[55] K. Ford, *The number of solutions of* $\phi(x) = m$, Annals of Math., **150** (1999), 283–312.

[56] E. Fouvry, *Autour du théorème de Bombieri-Vinogradov*, Acta Math., **152** (1984), 219–244.

[57] E. Fouvry, *Autour du théorème de Bombieri-Vinogradov. II* Ann. Sci. École Norm. Sup. (4), **20** (1987) 617–640.

[58] E. Fouvry, F. Grupp, *On the switching principle in sieve theory*, J. reine angew. Math. **370** (1986), 101–125.

[59] E. Fouvry, H. Iwaniec. *Primes in arithmetic progressions*, Acta Arith., **42** (1983), 197–218.

[60] E. Fouvry, G. Tenenbaum, *Entiers sans grand facteur premierè en progressions arithmétiques*, Proc. London Math. Soc. **63** (1991), 449–494.

[61] J.B. Friedlander, *Shifted primes without large prime factors*, *Number Theory and Applications*, Kluwer, Berlin, 393–401, 1989.

[62] J. Friedlander, H. Iwaniec, *Opera de Cribro* Amer. Math. Soc. Colloquium Publications, **57** (2010).

[63] M. Goldfeld, *On the number of primes p for which* $p+a$ *has a large prime factor*, Mathematika, **16** (1969), 23–27.

[64] A. Gómez-Colunga, C. Kavaler, N. McNew, M. Zhu, *On the size of primitive sets in function fields*, Finite Fields and their Applications, **64** (2020), 101658.

[65] A. Gómez-Colunga, C. Kavaler, N. McNew, M. Zhu, *On the Erdős primitive set conjecture in function fields*, J. Number Theory **219** (2021), 412–444.

[66] R. Graham, J. Spencer, *The elementary proof of the prime number theorem*, Math. Intelligencer **31** (2009), 18–23.

[67] A. Granville. *Smooth numbers: Computational number theory and beyond.* Algorithmic Number Theory (eds. J. Buhler and P. Stevenhagen), MSRI Publications 44, 2008.

[68] A. Granville, K. Soundararajan. *Large character sums: pretentious characters and the Pólya-Vinogradov theorem.* J. Amer. Math. Soc., **20** (2007), 357–384.

[69] H. Halberstam, K. F. Roth, *Sequences*, Oxford University Press (1966)

[70] R. R. Hall, *Sets of multiples*, Cambridge Tracts in Mathematics, **118**, Cambridge University Press (1996)

[71] R. R. Hall, G. Tenenbaum, *Divisors*, Cambridge Tracts in Mathematics, **90**, Cambridge University Press (1988)

[72] G. H. Hardy, J. E. Littlewood, *Some Problems of 'Partitio Numerorum.' III. On the Expression of a Number as a Sum of Primes*, Acta Math. **44** (1923), 1–70.

[73] G. Harman, *Prime-detecting sieves*, volume 33 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2007.

[74] G. Harman, *Watt's mean value theorem and Carmichael numbers*, Int. J. Number Theory, **4** (2008), 241–248.

[75] D. R. Heath-Brown, *Prime twins and Siegel zeros*, Proc. London Math. Soc., **47** (1983), 193–224.

[76] K. Henriot, *Nair-Tenenbaum bounds uniform with respect to the discriminant*, Math. Proc. Cambridge Philos. Soc., **152** (2012), 405–424.

[77] A. Hildebrand, *Additive and multiplicative functions on shifted primes*, Proc. London Math. Soc., bf 59, (1989), 209–232.

[78] H.-K. Hwang and T.-H. Tsai, *Quickselect and the Dickman function*, Combin. Probab. Comput. **11** (2002), 353–371.

[79] H. Iwaniec. *A new form of the error term in the linear sieve*, Acta Arith., **37** (1980), 307–320.

[80] H. Iwaniec, E. Kowalski. *Analytic number theory*, Amer. Math. Soc. Colloquium Publications, **53** (2004).

[81] H. Iwaniec, J. Pomykala, *Sums and differences of quartic norms*, Mathematika, **40** (1993), 233–245.

[82] J. C. Lagarias, *Euler's constant: Euler's work and modern developments*, Bull. Amer. Math. Soc. (N.S.) **50** (2013), 527–628.

[83] I. Laib , *Note on translated sum on primitive sequences*, Notes on Number Theory and Discrete Mathematics, **27** (2021), 39–43.

[84] I. Laib , A. Derbal , R. Mechik, *Somme translatée sur des suites primitives et la conjecture d'Erdös*, C. R. Acad. Sci. Paris, Ser. I 357 (2019), 413–417.

[85] E. Landau. *Handbuch der Lehre von der Verteilung der Primzahlen.* Erster Band. Leipzig, Berlin: B. G. Teubner, (1909).

[86] J. D. Lichtman, *Almost primes and the Banks–Martin conjecture*, J. Number Theory, **211** (2020), 513–529.

[87] J. D. Lichtman, *Averages of the Möbius function on shifted primes*, Q. J. Math. **73** (2022), 729—757.

[88] J. D. Lichtman, *Mertens' prime product formula, dissected*, Integers **21A** (2021), Ron Graham Memorial Volume, #A17, 15 pp.

[89] J. D. Lichtman, *Translated sums of primitive sets*, Comptes Rendus. Mathématique, 360 (2022), 409–414.

[90] J. D. Lichtman, C. Pomerance, *The Erdős conjecture for primitive sets*, Proc. Amer. Math. Soc. Ser. B **6** (2019), 1-–14.

[91] J. D. Lichtman, G. Martin, C. Pomerance, *Primes in prime number races*, Proc. Amer. Math. Soc. **147** (2019), 3743–3757.

[92] R. J. Mathar, *Series of reciprocal powers of k-almost primes*, preprint (2009), `arXiv:0803.0900`.

[93] K. Matomäki, M. Radziwiłł, *Multiplicative functions in short intervals*, Ann. of Math. **183** (2016), 1015–1056.

[94] K. Matomäki, M. Radziwiłł, *Multiplicative functions in short intervals II.*, (2020) `arXiv:2007.04290`

[95] K. Matomäki, M. Radziwiłł, *A note on the Liouville function in short intervals*, (2015) `arXiv:1502.02374`

[96] K. Matomäki, M. Radziwiłł, T. Tao, *An averaged form of Chowla's conjecture*, Algebra Number Theory, **9** (2015), 2167–2196.

[97] K. Matomäki, M. Radziwiłł, T. Tao, *Correlations of the von Mangoldt and higher divisor functions I. long shift ranges*, Proc. Lond. Math. Soc., **118** (2019), 284–350.

[98] K. Matomäki, M. Radziwiłł, T. Tao, *Correlations of the von Mangoldt and higher divisor functions II. divisor correlations in short ranges*, Math. Ann. **374** (2019), 793–840.

[99] K. Matomäki, M. Radziwiłł, T. Tao, *Fourier uniformity of bounded multiplicative functions in short intervals on average*, Invent. Math., **220** (2020), 1–58.

[100] K. Matomäki, M. Radziwiłł, T. Tao, J. Teräväinen, T. Ziegler, *Higher uniformity of bounded multiplicative functions in short intervals on average*, Ann. of Math., to appear.

[101] J. Maynard, *Primes in arithmetic progressions to large moduli I: fixed residue classes*, Mem. Amer. Math. Soc., to appear.

[102] J. Maynard, *Primes in arithmetic progressions to large moduli II: well-factorable estimates*, Mem. Amer. Math. Soc., to appear.

[103] J. Maynard, *Primes in arithmetic progressions to large moduli III: Uniform residue classes*, Mem. Amer. Math. Soc., to appear.

[104] F. Mertens, *Ein Beitrag zur analytischen Zahlentheorie*, J. reine angew. Math. **78** (1874), 46–62.

[105] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994.

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

[107] M. Ram Murty, A. Vatwani, *Twin primes and the parity problem*, J. Number Theory **180** (2017) 643,–659.

[108] M. Nair, G. Tenenbaum, *Short sums of certain arithmetic functions*, Acta Math. **180** (1998), 119–144.

[109] H. Nguyen, C. Pomerance, *The reciprocal sum of the amicable numbers*, Math. Comp. **88** (2019), 1503–1526.

[110] P.P. Pach, Cs. Sándor, Multiplicative bases and an Erdős problem, *Combinatorica*, **38** (2018), no. 5, 1175–1203.

[111] C. D. Pan, *A new application of the Yu. V. Linnik large sieve method*, Chinese Math. Acta **5** (1964), 642–652.

[112] M. D. Penrose, A. R. Wade, *Random minimal directed spanning trees and Dickman-type distributions*, Adv. in Appl. Probab. **36** (2004), 691–714.

[113] J. Pintz, *An approximation to the twin prime conjecture and the parity phenomenon*, Indag. Math., **26** (2015), 883–896.

[114] J. Pintz, *On the singular series in the prime k-tuple conjecture*, (2010) `arXiv:1004.1084`

[115] P. Pollack, *A generalization of the Hardy-Ramanujan inequality and applications*, J. Number Theory, **210** (2020), 171–182.

[116] C. Pomerance, *Popular values of Euler's function*, Mathematika, **27** (1980), 84–89.

[117] C. Pomerance, I.E. Shparlinski, *Smooth orders and cryptographic applications*, Proc. ANTS-V, Sydney, Australia, *Springer Lecture Notes in Computer Science* **2369** (2002), 338–348.

[118] D. Popa, *A double Mertens type evaluation*, J. Math. Anal. Appl. **409** (2014), 1159–1163.

[119] D. Popa, *A triple Mertens evaluation*, J. Math. Anal. Appl. **444** (2016), 464–474.

[120] T. Qi, S. Hu, *A multiple Mertens evaluation*, preprint (2019) `arXiv:1909.10930`

[121] K. Ramachandra, *Some problems of analytic number theory*, Acta Arith., **31** (1976), 313–324.

[122] H. Riesel, R.C. Vaughan, *On sums of primes*, Ark. Mat. **21** (1983), 45–74.

[123] G. Robin, *Sur l'ordre maximum de la fonction somme des diviseurs.* Seminaire Delange-Pisot-Poitou, Theorie des nombres 1981-1982, Progress in Mathematics **38** (1983), 233–244.

[124] J.B. Rosser, L. Schoenfeld, Approximate formulas for some functions of prime numbers, *Illinois J. Math.* **6** (1962), 64–94.

[125] M. Rubinstein, P. Sarnak, *Chebyshev's bias*, Experiment. Math. **3** (1994), 173–197.

[126] A. Sárközy, *On divisibility properties of sequences of integers*, 241–250, in *The Mathematics of Paul Erdős I* R. L. Graham, J. Nesetril (eds.), Springer-Verlag Berlin Heidelberg (1997).

[127] A. Sárközy, *On divisibility properties of sequences of integers*, 221–232, in *The Mathematics of Paul Erdős I* (2nd edition) R. L. Graham, J. Nesetril (eds.), Springer-Verlag Berlin Heidelberg (2013).

[128] American Institure of Mathematics: Problem Lists *Sarnak's conjecture*, available at: `http://aimpl.org/sarnakconjecture`

[129] A. Selberg, *On elementary methods in prime number theory and their limitations*, in: 11 Skand. Mat. kongr., Trondheim 1949, 13–22.

[130] A. Selberg, *Note on a paper of L. G. Sathe*, J. Indian Math. Soc., **18** (1954), 83–87.

[131] P. Shiu. *A Brun-Titchmarsh theorem for multiplicative functions*, J. Reine Angew. Math., 313 (1980), 161–170.

[132] H. Siebert, *Montgomery's weighted sieve for dimension two*, Monatsh. Math. **82** (1976), 327–336.

[133] C. A. Spiro, *Extensions of some formulae of A. Selberg*, Internat. J. Math. & Math. Sci., **8** (1985), 283–302.

[134] T. Tao, J. Teräväinen, *The structure of correlations of multiplicative functions at almost all scales, with applications to the Chowla and Elliott conjectures*, Algebra Number Theory **13** (2019), 2103–2150.

[135] T. Tao, J. Teräväinen, *The Hardy–Littlewood–Chowla conjecture in the presence of a Siegel zero*, J. London Math. Soc., to appear.

[136] G. Tenenbaum. *Generalized Mertens sums*, in Analytic number theory, modular forms and q-hypergeometric series – in honor of Krishna Alladi's 60th birthday, G. Andrews and F. Garvan, eds., Springer Proc. Math. Stat. **221**, Springer, Cham, (2017), 477–495.

[137] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Graduate Studies in Mathematics **163**, Amer. Math. Soc., 2015.

[138] A. I. Vinogradov, *The density hypothesis for Dirichet L-series*, Izv. Akad. Nauk SSSR Ser. Mat., **29** (1965), 903–934.

[139] K. R. Wooldridge, *Values taken many times by Euler's phi-function*, Proc. Amer. Math. Soc, **76** (1979) 229–234.

[140] J. Wu, *Sur la suite des nombres premiers jumeaux*, Acta Arith. **55** (1990), 365–394.

[141] J. Wu, *Chen's double sieve, Goldbach's conjecture and the twin prime problem*, Acta Arith. **114** (2004), 215–273.

[142] Y. Zhang, *Bounded gaps between primes,* Ann. of Math. (2), **179** (2014), 1121–1174.

[143] Z. Zhang, *On a problem of Erdős concerning primitive sequences*, Math. Comp. **60** (1993), 827–834.