Carl Pomerance · Michael Th. Rassias
Editors

Analytic Number Theory

In Honor of Helmut Maier's 60th Birthday

Springer

Analytic Number Theory

Professor Helmut Maier

Carl Pomerance • Michael Th. Rassias
Editors

Analytic Number Theory

In Honor of Helmut Maier's 60th Birthday

Springer

Editors
Carl Pomerance
Department of Mathematics
Dartmouth College
Hanover, NH, USA

Michael Th. Rassias
Department of Mathematics
ETH-Zürich
Zürich, Switzerland

Department of Mathematics
Princeton University
Princeton, NJ, USA

ISBN 978-3-319-22239-4 ISBN 978-3-319-22240-0 (eBook)
DOI 10.1007/978-3-319-22240-0

Library of Congress Control Number: 2015947785

Mathematics Subject Classification (2010): 11-Axx, 11-Bxx, 11-Lxx, 11-Mxx, 11-Nxx, 11-Pxx, 11-Yxx, 26-Dxx, 40-XX, 41-XX

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

This book presents a series of research and expository articles under the broad banner of Analytic Number Theory. Written by eminent mathematicians from the international mathematical community, the chapters in this book are dedicated to the 60th birthday of Professor Helmut Maier. His own research has been groundbreaking and deeply influential, especially in the problem of gaps between consecutive primes.

Helmut Maier was born in Geislingen/Steige, Germany, on October 17, 1953. He graduated with a Diploma in Mathematics from the University of Ulm in 1976, under the supervision of Professor H.E. Richert. After a period as a scientific employee at the University of Ulm, he started his doctoral studies at the University of Minnesota in 1979. He received his Ph.D. in 1981 with the thesis entitled *Some results on prime numbers based on the application of sieve theory*, under the supervision of Professor J. Ian Richards. The thesis was an extension of his paper *Chains of large gaps between consecutive primes*, which appeared in Advances in Mathematics, **39** (1981), 257–269. In this paper, Maier applied for the first time what is now called *Maier's Matrix Method*. This method later on led him and other mathematicians to the discovery of unexpected irregularities in the distribution of prime numbers. The Matrix Method was also applied to other questions like irreducible polynomials and consecutive primes in the same residue class modulo an integer.

After postdoctoral positions at the University of Michigan and the Institute for Advanced Study, Princeton, Maier then took a permanent position at the University of Georgia. During this period, he showed that the usual formulation of the Cramér model for the distribution of primes is wrong, a completely unexpected result. With C. Pomerance, he did research on the values of Euler's φ-function and large gaps between primes. One theme in this period was Professor Maier's work on the size of the coefficients of cyclotomic polynomials, done partially in collaboration with S. Konyagin and E. Wirsing. He also collaborated with H.L. Montgomery on the size of the sum of the Möbius function under the assumption of the Riemann Hypothesis. Maier and G. Tenenbaum in joint work investigated the sequence of divisors of integers, solving the famous *propinquity* problem of Erdős.

Since 1993, Maier is a Professor at the University of Ulm, Germany. Lately, M.Th. Rassias and Maier worked on problems related to the Estermann zeta function and the Nyman–Beurling criterion on the Riemann Hypothesis involving the distribution of certain cotangent sums. In their latest work, they studied prime avoidance of perfect powers of prime numbers.

Other collaborators of Professor Maier include P. Erdős, J. Friedlander, A. Granville, D. Haase, A.J. Hildebrand, M.L. Lapidus, J.W. Neuberger, A. Sankaranarayanan, A. Sárközy, and C.L. Stewart.

We are grateful to all of the mathematicians who participated in this publication.

We also wish to express our thanks to Springer for undertaking the publication of this book.

Hanover, NH, USA — Carl Pomerance
Princeton, NJ, USA — Michael Th. Rassias

Contents

CM-Points on Straight Lines

Bill Allombert, Yuri Bilu, and Amalia Pizarro-Madariaga

To Helmut Maier on his 60th anniversary

Abstract We prove that, with "obvious" exceptions, a CM-point $(j(\tau_1), j(\tau_2))$ cannot belong to a straight line in $\mathbb{C}^2$ defined over $\mathbb{Q}$. This generalizes a result of Kühne, who proved this for the line $x_1 + x_2 = 1$.

1 Introduction

In this article τ with or without indices denotes a quadratic[1] complex number with $\mathrm{Im}\tau > 0$ and j denotes the j-invariant.

In 1998 André [1] proved that a non-special irreducible plane curve in $\mathbb{C}^2$ may have only finitely many CM-points. Here a *plane curve* is a curve defined by an irreducible polynomial equation $F(x_1, x_2) = 0$, where F is a polynomial with complex coefficients. A *CM-point* in $\mathbb{C}^2$ is a point of the form $(j(\tau_1), j(\tau_2))$ with quadratic τ_1, τ_2. *Special curves* are the curves of the following types:

[1] "Quadratic" here and below means "of degree 2 over $\mathbb{Q}$ ".

B. Allombert
Université de Bordeaux, IMB, UMR 5251, F-33400 Talence, France

CNRS, F-33400 Talence, France

INRIA, F-33400 Talence, France
e-mail: Bill.Allombert@math.u-bordeaux.fr

Y. Bilu (✉)
Institut de Mathématiques de Bordeaux, Université de Bordeaux & CNRS, 351 cours de la Libération, 33405 Talence, France
e-mail: yuri@math.u-bordeaux1.fr

A. Pizarro-Madariaga
Instituto de Matemáticas, Universidad de Valparaíso, Valparaíso, Chile
e-mail: amalia.pizarro@uv.cl

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_1

- "vertical lines" $x_1 = j(\tau_1)$;
- "horizontal lines" $x_2 = j(\tau_2)$;
- *modular curves* $Y_0(N)$, realized as the plane curves $\Phi_N(x_1, x_2) = 0$, where Φ_N is the modular polynomial of level N.

Clearly, each special curve contains infinitely many CM-points, and André proved that special curves are characterized by this property.

André's result was the first non-trivial contribution to the celebrated André-Oort conjecture on the special subvarieties of Shimura varieties; see [11] and the references therein.

Independently of André the same result was also obtained by Edixhoven [7], but Edixhoven had to assume the Generalized Riemann Hypothesis for certain L-series to be true.

Further proof followed; we mention specially the remarkable argument of Pila [10]. It is based on an idea of Pila and Zannier [12] and readily extends to higher dimensions [11].

The arguments mentioned above were non-effective, because they used the Siegel-Brauer lower bound for the class number. Breuer [5] gave an effective proof, but it depended on GRH.

Recently Kühne [8, 9] and, independently, Bilu et al. [3] found unconditional effective proofs of André's theorem. Besides giving general results, both articles [9] and [3] treat also some particular curves, showing they have no CM-points at all. For instance, Kühne [9, Theorem 5] proves the following.

Theorem 1.1. *The straight line $x_1 + x_2 = 1$ has no CM-points.*

(The same result was also independently obtained in an earlier version of [3], but did not appear in the final version.)

A similar result for the curve $x_1 x_2 = 1$ was obtained in [3].

One can ask about CM-points on general straight lines defined over $\mathbb{Q}$; that is, defined by an equation

$$A_1 x_1 + A_2 x_2 + B = 0, \tag{1}$$

where $A_1, A_2, B \in \mathbb{Q}$. One has to exclude from consideration the *special straight lines*: $x_1 = j(\tau_1)$, $x_2 = j(\tau_2)$ and $x_1 = x_2$, the latter being nothing else than the modular curve $Y_0(1)$ (the modular polynomial Φ_1 is $x_1 - x_2$). According to the theorem of André, these are the only straight lines containing infinitely many CM-points.

In the present paper we obtain a rather vast generalization of Theorem 1.1.

Theorem 1.2. *Let $(j(\tau_1), j(\tau_2))$ be a CM-point belonging to a non-special straight line defined over $\mathbb{Q}$. Then we have one of the following options. Either*

$$j(\tau_1), j(\tau_2) \in \mathbb{Q}, \tag{2}$$

or

$$j(\tau_1) \neq j(\tau_2), \quad \mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2)), \quad [\mathbb{Q}(j(\tau_1)) : \mathbb{Q}] = [\mathbb{Q}(j(\tau_2)) : \mathbb{Q}] = 2. \tag{3}$$

Remark 1.3. 1. Recall that $[\mathbb{Q}(j(\tau) : \mathbb{Q}] = h(\mathscr{O}_\tau)$, the class number of the "complex multiplication order" $\mathscr{O}_\tau = \mathrm{End}\langle\tau, 1\rangle$, where $\langle\tau, 1\rangle$ is the lattice generated by τ and 1. All orders of class number 1 and 2 are well known, which means that points satisfying (2) or (3) can be easily listed. In fact, there are 169 CM-points satisfying (2) and, up to $\mathbb{Q}$-conjugacy, 217 CM-points satisfying (3); see Remark 5.3 for the details.

2. Our result is best possible because any point satisfying (2) or (3) does belong to a non-special straight line defined over $\mathbb{Q}$.
3. Kühne remarks on page 5 of his article [9] that his Theorem 4 allows one, in principle, to list all possible CM-points belonging to non-special straight lines over $\mathbb{Q}$, but the implied calculation does not seem to be feasible.
4. Bajolet [2] produced a software package for finding all CM-points on a given straight line. He illustrated its efficiency by proving that no straight line (1) with non-zero $A_1, A_2, B \in \mathbb{Z}$ satisfying $|A_1|, |A_2|, |B| \leq 10$ passes through a CM-point. This work is now formally obsolete because of our Theorem 1.2, but a similar method can be used in more general situations, where our theorem no longer applies.
5. CM-points (x_1, x_2) satisfying $x_1 x_2 \in \mathbb{Q}^\times$ are completely classified in [4]; this generalizes the above-mentioned result from [3] about the curve $x_1 x_2 = 1$.

In Sects. 2 and 3 we recall basic facts about imaginary quadratic orders, class groups, ring class fields and complex multiplications.

In Sect. 4 we investigate the field equality $\mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2))$. In particular, in Corollary 4.2 we determine all cases of such equality when $\mathbb{Q}(\tau_1) \neq \mathbb{Q}(\tau_2)$. This might be of independent interest.

After all these preparations, we prove Theorem 1.2 in Sect. 5.

2 Imaginary Quadratic Orders

In this section we recall basic facts about imaginary quadratic fields and their orders, and recall a famous result of Weinberger about class groups annihilated by 2.

2.1 Class Groups

Let K be an imaginary quadratic field and $\mathscr{O}$ an order in K of discriminant $\Delta = Df^2$, where D is the discriminant of the field K (often called the *fundamental discriminant*) and f is the conductor of $\mathscr{O}$, defined from $\mathscr{O} = \mathbb{Z} + f\mathscr{O}_K$. We denote

by $\mathrm{Cl}(\mathscr{O})$ the class group of $\mathscr{O}$ (the group of invertible fractional ideals modulo the invertible principal fractional ideals). As usual, we set $h(\mathscr{O}) = |\mathrm{Cl}(\mathscr{O})|$. Since $\mathscr{O}$ is uniquely determined by its discriminant Δ, we may also write $\mathrm{Cl}(\Delta)$, $h(\Delta)$, etc. In particular, $\mathrm{Cl}(D) = \mathrm{Cl}(\mathscr{O}_K)$ is the class group of the field K, and $h(D)$ is the class number of K.

There is a canonical exact sequence

$$1 \to \mathrm{Cl}_0(\Delta) \to \mathrm{Cl}(\Delta) \to \mathrm{Cl}(D) \to 1, \tag{4}$$

where the kernel $\mathrm{Cl}_0(\Delta)$ will be described below. This implies, in particular, that $h(D) \mid h(\Delta)$.

The structure of the group $\mathrm{Cl}_0(\Delta)$ is described, for instance, in [6], Sect. 7.D and Exercise 7.30. We briefly reproduce this description here. We will assume, with a slight abuse of notation, that $\mathbb{Z}/f\mathbb{Z}$ is a subring of $\mathscr{O}_K/f\mathscr{O}_K$. Then we have another canonical exact sequence

$$1 \to (\mathbb{Z}/f\mathbb{Z})^\times(\mathscr{O}_K^\times)_f \hookrightarrow (\mathscr{O}_K/f\mathscr{O}_K)^\times \to \mathrm{Cl}_0(\Delta) \to 1, \tag{5}$$

where $(\mathscr{O}_K^\times)_f$ is the image of the multiplicative group $\mathscr{O}_K^\times$ in $(\mathscr{O}_K/f\mathscr{O}_K)$.

The group $(\mathbb{Z}/f\mathbb{Z})^\times(\mathscr{O}_K^\times)_f$ is "not much bigger" than $(\mathbb{Z}/f\mathbb{Z})^\times$. Precisely,

$$[(\mathbb{Z}/f\mathbb{Z})^\times(\mathscr{O}_K^\times)_f : (\mathbb{Z}/f\mathbb{Z})^\times] = [\mathscr{O}_K^\times : \mathscr{O}^\times] = \begin{cases} 2 & \text{if } D = -4, f > 1, \\ 3 & \text{if } D = -3, f > 1, \\ 1 & \text{otherwise.} \end{cases}$$

An easy consequence is the following formula for $h(\Delta)$:

$$h(\Delta) = \frac{fh(D)}{[\mathscr{O}_K^\times : \mathscr{O}^\times]} \prod_{p \mid f} \left(1 - \left(\frac{D}{p}\right) p^{-1}\right), \tag{6}$$

where $(D/\cdot)$ is the Kronecker symbol.

2.2 Orders with Class Groups Annihilated by 2

In this subsection we recall the famous result of Weinberger about imaginary quadratic orders whose class group is annihilated by 2. For a multiplicatively written abelian group G we denote by G^2 its subgroup of squares: $G^2 = \{g^2 : g \in G\}$.

The group $\mathrm{Cl}(\Delta)/\mathrm{Cl}(\Delta)^2$ is usually called the *genus group* of Δ. It is known to be isomorphic to $(\mathbb{Z}/2\mathbb{Z})^\mu$, where $\mu = \mu(\Delta) \in \{\omega(\Delta) - 1, \omega(\Delta)\}$ and $\omega(\cdot)$ denote the number of distinct prime divisors. We may also remark that $\mu(\Delta) = \omega(\Delta) - 1$ when $\Delta = D$ (and $f = 1$).

Already Euler studied discriminants Δ with the property

$$|\mathrm{Cl}(\Delta)^2| = 1, \tag{7}$$

or, equivalently, $\mathrm{Cl}(\Delta) \cong (\mathbb{Z}/2\mathbb{Z})^\mu$. (Of course, he used a different terminology.) Chowla proved that the set of such Δ is finite. Using a deep result of Tatuzawa [13] about Siegel's zero, Weinberger [16] improved on this, by showing that *field discriminants* D with this property are bounded explicitly with at most one exception.

To state Weinberger's result precisely, denote by D' the square-free part of D:

$$D' = \begin{cases} D, & D \equiv 1 \bmod 4, \\ D/4, & D \equiv 0 \bmod 4. \end{cases}$$

Proposition 2.1 (Weinberger [16], Theorem 1). *There exists a negative integer D^* such that for any discriminant D of an imaginary quadratic field the property $|\mathrm{Cl}(D)^2| = 1$ implies $|D'| \le 5460$ or $D = D^*$.*

Using existing methods [15], it is easy to determine the full list of D with $|\mathrm{Cl}(D)^2| = 1$ and $|D'| \le 5674$. Since the group $\mathrm{Cl}(D)$ is a quotient of the group $\mathrm{Cl}(Df^2)$, if the latter has property (7) then the former does. Also, for every given D it is easy to find all possible f such that $|\mathrm{Cl}(Df^2)^2| = 1$, using the description of the group $\mathrm{Cl}(Df^2)$ given in Sect. 2.1. Hence the couples (D, f) for which $|\mathrm{Cl}(Df^2)^2| = 1$ and $|D'| \le 5674$ can be easily listed as well. This list is widely available in the literature since long ago; we reproduce it in Table 1.

It follows that Weinberger's result has the following consequence.

Corollary 2.2. *There exists a negative integer D^* such that $|\mathrm{Cl}(Df^2)^2| = 1$ implies that either $\Delta = Df^2$ appears in Table 1 or $D = D^*$.*

Table 1 Known Δ with $|\mathrm{Cl}(\Delta)^2| = 1$

$h(\Delta) = 1$	$-3, -3\cdot 2^2, -3\cdot 3^2, -4, -4\cdot 2^2, -7, -7\cdot 2^2, -8, -11, -19, -43, -67, -163$
$h(\Delta) = 2$	$-3\cdot 4^2, -3\cdot 5^2, -3\cdot 7^2, -4\cdot 3^2, -4\cdot 4^2, -4\cdot 5^2, -7\cdot 4^2, -8\cdot 2^2, -8\cdot 3^2, -11\cdot 3^2,$ $-15, -15\cdot 2^2, -20, -24, -35, -40, -51, -52, -88, -91, -115, -123, -148,$ $-187, -232, -235, -267, -403, -427$
$h(\Delta) \ge 4$	$-3\cdot 8^2, -7\cdot 8^2, -8\cdot 6^2, -15\cdot 4^2, -15\cdot 8^2, -20\cdot 3^2, -24\cdot 2^2, -35\cdot 3^2, -40\cdot 2^2,$ $-84, -88\cdot 2^2, -120, -120\cdot 2^2, -132, -168, -168\cdot 2^2, -195, -228, -232\cdot 2^2,$ $-280, -280\cdot 2^2, -312, -312\cdot 2^2, -340, -372, -408, -408\cdot 2^2, -420, -435,$ $-483, -520, -520\cdot 2^2, -532, -555, -595, -627, -660, -708, -715,$ $-760, -760\cdot 2^2, -795, -840, -840\cdot 2^2, -1012, -1092, -1155, -1320, -1320\cdot 2^2,$ $-1380, -1428, -1435, -1540, -1848, -1848\cdot 2^2, -1995, -3003, -3315, -5460$

Remark 2.3. Class numbers of discriminants from Table 1 are at most 16, and the results of [15] imply that Table 1 contains all Δ with $|\mathrm{Cl}(\Delta)^2| = 1$ and $h(\Delta) \leq 64$. Hence if Δ satisfies $|\mathrm{Cl}(\Delta)^2| = 1$ but does not appear in Table 1 then we must have $h(\Delta) \geq 128$.

In particular, the first two lines of Table 1 give full lists of negative quadratic discriminants Δ with $h(\Delta) = 1$ and 2.

3 Ring Class Fields and Complex Multiplication

Let K be an imaginary quadratic field, and $\mathcal{O}$ an order in K of discriminant $\Delta = Df^2$. One associates with $\mathcal{O}$ an abelian extension of K with Galois group $\mathrm{Cl}(\mathcal{O})$, called the *ring class field* of $\mathcal{O}$. We will denote it by $\mathrm{RiCF}(\mathcal{O})$, or $\mathrm{RiCF}(\Delta)$, or $\mathrm{RiCF}(K,f)$. The canonical isomorphism $\mathrm{Cl}(\mathcal{O}) \to \mathrm{Gal}(\mathrm{RiCF}(\mathcal{O})/K)$ is called the *Artin map*. For the details see, for instance, [6, Sect. 9].

The correspondence $\mathcal{O} \leftrightarrow \mathrm{RiCF}(\mathcal{O})$ is functorial in the following sense: if $\mathcal{O}'$ is a sub-order of $\mathcal{O}$, then $\mathrm{RiCF}(\mathcal{O}') \subset \mathrm{RiCF}(\mathcal{O})$, and we have the commutative diagram

$$\begin{array}{ccc} \mathrm{Cl}(\mathcal{O}) & \to & \mathrm{Gal}(\mathrm{RiCF}(\mathcal{O})/K) \\ \downarrow & & \downarrow \\ \mathrm{Cl}(\mathcal{O}') & \to & \mathrm{Gal}(\mathrm{RiCF}(\mathcal{O}')/K) \end{array}$$

where the horizontal arrows denote Artin maps and the vertical arrows are the natural maps of the class groups and the Galois groups. It follows that the Galois group of $\mathrm{RiCF}(\mathcal{O})$ over $\mathrm{RiCF}(\mathcal{O}')$ is isomorphic to the kernel of $\mathrm{Cl}(\mathcal{O}) \to \mathrm{Cl}(\mathcal{O}')$. In particular, $\mathrm{Gal}(\mathrm{RiCF}(\mathcal{O})/\mathrm{RiCF}(\mathcal{O}_K))$ is $\mathrm{Cl}_0(\Delta)$, the group introduced in Sect. 2.1. (One may notice that $\mathrm{RiCF}(\mathcal{O}_K)$ is nothing else but the *Hilbert class field* of K.)

3.1 Compositum of Ring Class Fields

In this subsection it will be more convenient to use the "conductor notation" $\mathrm{RiCF}(K,f)$.

As we have seen above, if $f_1 \mid f$ then $\mathrm{RiCF}(K,f_1) \subset \mathrm{RiCF}(K,f)$. It follows that the compositum of two ring class fields $\mathrm{RiCF}(K,f_1)$ and $\mathrm{RiCF}(K,f_2)$ is a subfield of $\mathrm{RiCF}(K,f)$, where $f = \mathrm{LCM}(f_1,f_2)$. It turns out that this compositum is "almost always" equal to $\mathrm{RiCF}(K,f)$, but there are some exceptions. Here is the precise statement. It is certainly known, but we did not find it in the available literature.

Proposition 3.1. *Let K be an imaginary quadratic field of discriminant D and f_1, f_2 positive integers. Set $f = \mathrm{LCM}(f_1,f_2)$. Then we have the following.*

1. *If* $D \neq -3, -4$ *then* $\mathrm{RiCF}(K,f_1)\mathrm{RiCF}(K,f_2) = \mathrm{RiCF}(K,f)$.
2. *Assume that* $D \in \{-3 - 4\}$. *Then* $\mathrm{RiCF}(K,f_1)\mathrm{RiCF}(K,f_2) = \mathrm{RiCF}(K,f)$ *either when one of* f_1, f_2 *is* 1 *or when* $\gcd(f_1,f_2) > 1$. *On the contrary, when* $f_1, f_2 > 1$ *and* $\gcd(f_1,f_2) = 1$, *the compositum* $\mathrm{RiCF}(K,f_1)\mathrm{RiCF}(K,f_2)$ *is a subfield of* $\mathrm{RiCF}(K,f)$ *of degree* 2 *for* $D = -4$ *and of degree* 3 *for* $D = -3$.

Proof (A Sketch). To simplify the notation, we set

$$L_0 = \mathrm{RiCF}(K,1), \qquad L_1 = \mathrm{RiCF}(K,f_1), \qquad L_2 = \mathrm{RiCF}(K,f_2),$$
$$L = \mathrm{RiCF}(K,f), \qquad L' = L_1 L_2.$$

The mutual position of these fields is illustrated here:

$$\begin{array}{ccccc} & & L_1 & & \\ & \nearrow & \downarrow & \searrow & \\ L_0 & \longrightarrow & L' & \longrightarrow & L; \\ & \searrow & \uparrow & \nearrow & \\ & & L_2 & & \end{array}$$

We want to determine the degree $[L : L']$.

We have $\mathrm{Gal}(L/L_0) = \mathrm{Cl}_0(Df^2)$. By (5), this implies

$$\mathrm{Gal}(L/L_0) = (\mathscr{O}_K/f\mathscr{O}_K)^\times/(\mathbb{Z}/f\mathbb{Z})^\times(\mathscr{O}_K^\times)_f,$$

Similarly,

$$\mathrm{Gal}(L_i/L_0) = (\mathscr{O}_K/f_i\mathscr{O}_K)^\times/(\mathbb{Z}/f_i\mathbb{Z})^\times(\mathscr{O}_K^\times)_{f_i} \qquad (i = 1, 2).$$

The Galois group $\mathrm{Gal}(L/L_i)$ is the kernel of the natural map

$$(\mathscr{O}_K/f\mathscr{O}_K)^\times/(\mathbb{Z}/f\mathbb{Z})^\times(\mathscr{O}_K^\times)_f \xrightarrow{\pi_i} (\mathscr{O}_K/f_i\mathscr{O}_K)^\times/(\mathbb{Z}/f_i\mathbb{Z})^\times(\mathscr{O}_K^\times)_{f_i}.$$

Hence $\mathrm{Gal}(L/L')$ is the common kernel of the maps π_1 and π_2. It follows that $\mathrm{Gal}(L/L') = G/(\mathbb{Z}/f\mathbb{Z})^\times(\mathscr{O}_K^\times)_f$, where G is the subgroup of $(\mathscr{O}_K/f\mathscr{O}_K)^\times$ consisting of $x \in (\mathscr{O}_K/f\mathscr{O}_K)^\times$ satisfying

$$x \in (\mathbb{Z}/f\mathbb{Z})(\mathscr{O}_K^\times)_f \bmod f_i \qquad (i = 1, 2). \tag{8}$$

In particular, $[L : L'] = [G : (\mathbb{Z}/f\mathbb{Z})^\times(\mathscr{O}_K^\times)_f]$.

If $D \neq -3, -4$ then $\mathscr{O}_K^\times = \{\pm 1\}$, which implies that

$$G = (\mathbb{Z}/f\mathbb{Z})^\times(\mathscr{O}_K^\times)_f = (\mathbb{Z}/f\mathbb{Z})^\times$$

and $L = L'$.

Now assume that $D = -4$. Then $\mathscr{O}_K^\times = \{\pm 1, \pm\sqrt{-1}\}$. When $f_1 = 1$ or $f_2 = 1$ the statement is trivial, so we may assume that $f_1, f_2 > 1$. Condition (8) can be rewritten as

$$x \in \mathbb{Z}/f\mathbb{Z} \cup (\mathbb{Z}/f\mathbb{Z})\sqrt{-1} \cup \big((\mathbb{Z}/f\mathbb{Z})f_1 + (\mathbb{Z}/f\mathbb{Z})f_2\sqrt{-1}\big) \cup \big((\mathbb{Z}/f\mathbb{Z})f_2 + (\mathbb{Z}/f\mathbb{Z})f_1\sqrt{-1}\big). \quad (9)$$

If $\gcd(f_1, f_2) > 1$, then the last two sets in (9) have no common elements with $(\mathscr{O}_K/f\mathscr{O}_K)^\times$. We obtain

$$G = \big(\mathbb{Z}/f\mathbb{Z} \cup (\mathbb{Z}/f\mathbb{Z})\sqrt{-1}\big) \cap (\mathscr{O}_K/f\mathscr{O}_K)^\times = (\mathbb{Z}/f\mathbb{Z})(\mathscr{O}_K^\times)_f,$$

and $L = L'$.

If $\gcd(f_1, f_2) = 1$, then each of the last two sets in (9) has elements belonging to $(\mathscr{O}_K/f\mathscr{O}_K)^\times$ but not to $(\mathbb{Z}/f\mathbb{Z})(\mathscr{O}_K^\times)_f$; for instance, $f_1 + f_2\sqrt{-1}$ and $f_2 + f_1\sqrt{-1}$, respectively (here we use the assumption $f_1, f_2 > 1$). Hence $[G : (\mathbb{Z}/f\mathbb{Z})(\mathscr{O}_K^\times)_f] > 1$. On the other hand, if x and y belong to the last two sets in (9), then $xy \in \mathbb{Z}/f\mathbb{Z}$ if they belong to the same set, and $xy \in (\mathbb{Z}/f\mathbb{Z})\sqrt{-1}$ if they belong to distinct sets. This shows that $[G : (\mathbb{Z}/f\mathbb{Z})(\mathscr{O}_K^\times)_f] = 2$, and hence $[L : L'] = 2$. This completes the proof in the case $D = -4$.

The case $D = -3$ is treated similarly. We omit the details. □

3.2 Complex Multiplication

Ring class fields are closely related to the Complex Multiplication. Let $\tau \in K$ with $\mathrm{Im}\tau > 0$ be such that $\mathscr{O} = \mathrm{End}\langle\tau, 1\rangle$ (where $\langle\tau, 1\rangle$ is the lattice generated by τ and 1); one says that $\mathscr{O}$ is the *complex multiplication order* of the lattice $\langle\tau, 1\rangle$.

The "Main Theorem of Complex Multiplication" asserts that $j(\tau)$ is an algebraic integer generating over K the ring class field $\mathrm{RiCF}(\mathscr{O})$. In particular, $[K(j(\tau)) : K] = h(\mathscr{O})$. In fact, one has more:

$$[K(j(\tau)) : K] = [\mathbb{Q}(j(\tau)) : \mathbb{Q}] = h(\mathscr{O}). \quad (10)$$

The proofs can be found in many sources; see, for instance, [6, Sect. 11].

Since $\mathrm{RiCF}(\Delta) = K(j(\tau))$ is a Galois extension of K, it contains, by (10), all the $\mathbb{Q}$-conjugates of $j(\tau)$. It follows that $K(j(\tau))$ is Galois over $\mathbb{Q}$; in particular, the Galois group $\mathrm{Gal}(K/\mathbb{Q})$ acts on $\mathrm{Gal}(K(j(\tau))/K)$.

Proposition 3.2. *The Galois group* $\mathrm{Gal}(K/\mathbb{Q})$ *acts on* $\mathrm{Gal}(K(j(\tau))/K)$ *"dihedrally": if* ι *is the non-trivial element of* $\mathrm{Gal}(K/\mathbb{Q})$*, then we have* $\sigma^\iota = \sigma^{-1}$ *for any* $\sigma \in \mathrm{Gal}(K(j(\tau))/K)$.

For a proof see, for instance, [6, Lemma 9.3].

Corollary 3.3. *The following properties are equivalent.*

1. *The field* $K(j(\tau))$ *is abelian over* $\mathbb{Q}$.
2. *The Galois group* $\mathrm{Gal}(K(j(\tau))/K)$ *is annihilated by* 2 *(that is, isomorphic to* $\mathbb{Z}/2\mathbb{Z} \times \cdots \times \mathbb{Z}/2\mathbb{Z}$*).*
3. *The Galois group* $\mathrm{Gal}(K(j(\tau))/\mathbb{Q})$ *is annihilated by* 2.
4. *The field* $\mathbb{Q}(j(\tau))$ *is Galois over* $\mathbb{Q}$.
5. *The field* $\mathbb{Q}(j(\tau))$ *is abelian over* $\mathbb{Q}$.

Proof. The implications 1⇒2⇒3 follow from Proposition 3.2. The implication 3⇒4 is trivial. To see the implication 4⇒5, just observe that (10) implies the isomorphism $\mathrm{Gal}(K(j(\tau))/K) \cong \mathrm{Gal}(\mathbb{Q}(j(\tau))/\mathbb{Q})$. Finally, the implication 5⇒1 is again trivial. □

3.3 The Conjugates of $j(\tau)$

Let τ and $\mathscr{O}$ be as in Sect. 3.2, and let Δ be the discriminant of $\mathscr{O}$. As we already mentioned in the beginning of Sect. 3.2, $j(\tau)$ is an algebraic integer of degree $h(\Delta)$. It is well known that the $\mathbb{Q}$-conjugates of $j(\tau)$ can be described explicitly. Below we briefly recall this description.

Denote by $T = T_\Delta$ the set of triples of integers (a, b, c) such that

$$\gcd(a, b, c) = 1, \quad \Delta = b^2 - 4ac,$$
$$\text{either} \quad -a < b \le a < c \quad \text{or} \quad 0 \le b \le a = c$$

Proposition 3.4. *All* $\mathbb{Q}$*-conjugates of* $j(\tau)$ *are given by*

$$j\left(\frac{-b + \sqrt{\Delta}}{2a}\right), \qquad (a, b, c) \in T_\Delta. \tag{11}$$

In particular, $h(\Delta) = |T_\Delta|$.

For a proof, see, for instance, [6, Theorem 7.7].

The following observation will be crucial: in the set T_Δ there exists exactly one triple (a, b, c) with $a = 1$. This triple can be given explicitly: it is

$$\left(1, r_4(\Delta), \frac{r_4(\Delta) - \Delta}{4}\right),$$

where $r_4(\Delta) \in \{0, 1\}$ is defined by $\Delta \equiv r_4(\Delta) \mod 4$. The corresponding number $j(\tau)$, where $\tau = (-r_4(\Delta) + \sqrt{\Delta})/2$, will be called *the dominant j-value* of discriminant Δ. It is important for us that it is much larger in absolute value than all its conjugates.

Lemma 3.5. *Let $j(\tau)$ be the dominant j-value of discriminant Δ, with $|\Delta| \geq 11$, and let $j(\tau') \neq j(\tau)$ be conjugate to $j(\tau)$ over $\mathbb{Q}$. Then $|j(\tau')| \leq 0.1|j(\tau)|$.*

Proof. Recall the inequality $\big||j(z)| - |q_z^{-1}|\big| \leq 2079$, where $q_z = e^{2\pi i z}$, for z belonging to the standard fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$ on the Poincaré plane [3, Lemma 1]. We may assume that $\tau = (-r_4(\Delta) + \sqrt{\Delta})/2$ and $\tau' = (-b + \sqrt{\Delta})/2a$ with $a \geq 2$. Hence $|q_\tau| = e^{\pi\sqrt{|\Delta|}} \geq e^{\pi\sqrt{11}} > 33506$ and $|q_{\tau'}| \leq |q_\tau|^{1/2}$. We obtain

$$\frac{|j(\tau')|}{|j(\tau)|} \leq \frac{|q_\tau|^{1/2} + 2079}{|q_\tau| - 2079} \leq \frac{33506^{1/2} + 2079}{33506 - 2079} < 0.1,$$

as wanted. □

The minimal polynomial of $j(\tau)$ over $\mathbb{Z}$ is called the *Hilbert class polynomial*[2] of discriminant Δ; it indeed depends only on Δ because its roots are the numbers (11). We will denote it $H_\Delta(x)$.

4 Comparing Two CM-Fields

In this section we study the field equality $\mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2))$. We distinguish two cases: $\mathbb{Q}(\tau_1) \neq \mathbb{Q}(\tau_2)$, when we obtain the complete list of all possibilities, and $\mathbb{Q}(\tau_1) = \mathbb{Q}(\tau_2)$, where we will see that Δ_1 and Δ_2 are "almost the same." Here for $i = 1, 2$ we denote by Δ_i the discriminant of the "complex multiplication order" $\mathcal{O}_i = \mathrm{End}\langle\tau_i, 1\rangle$, and write $\Delta_i = D_i f_i^2$ with the obvious meaning of D_i and f_i.

4.1 The Case $\mathbb{Q}(\tau_1) \neq \mathbb{Q}(\tau_2)$

In this subsection we investigate the case when $\mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2))$, but the fields $\mathbb{Q}(\tau_1)$ and $\mathbb{Q}(\tau_2)$ are distinct. It turns out that this is a very strong condition, which leads to a completely explicit characterization of all possible cases.

Theorem 4.1. *Let τ_1 and τ_2 be quadratic numbers such that $\mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2))$, but $\mathbb{Q}(\tau_1) \neq \mathbb{Q}(\tau_2)$. Then both Δ_1 and Δ_2 appear in Table 1.*

Proof. Denote by L the field $\mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2))$. If L is a Galois extension of $\mathbb{Q}$, then the group $\mathrm{Cl}(\Delta_1) = \mathrm{Cl}(\Delta_2)$ is annihilated by 2 by Corollary 3.3, and we can use Corollary 2.2. Since $D_1 \neq D_2$, at least one of the two discriminants D_1 and D_2

[2] Or, sometimes, *ring class polynomial*, to indicate that its root generates not the Hilbert class field, but the more general ring class field.

is distinct from D^*; say, $D_1 \neq D^*$. Then Δ_1 is in Table 1. Since $h(\Delta_1) = h(\Delta_2)$, Remark 2.3 implies that Δ_2 is in Table 1 as well. (This argument goes back to Kühne [9, Sect. 6].)

Now assume that

$$L \text{ is not Galois over } \mathbb{Q}. \tag{12}$$

We will show that this leads to a contradiction. Denote by M the Galois closure of L over $\mathbb{Q}$; then, $M = \mathbb{Q}(\tau_1, j(\tau_1)) = \mathbb{Q}(\tau_2, j(\tau_2))$. Define the Galois groups

$$G = \mathrm{Gal}(M/\mathbb{Q}), \qquad \tilde{N} = \mathrm{Gal}(M/\mathbb{Q}(\tau_1, \tau_2)),$$
$$N_i = \mathrm{Gal}(M/\mathbb{Q}(\tau_i)) = \mathrm{Cl}(\Delta_i) \quad (i = 1, 2),$$

so that $\tilde{N} = N_1 \cap N_2$ and $[N_1 : \tilde{N}] = [N_2 : \tilde{N}] = 2$.

We claim the following:

$$\text{the group } \tilde{N} \text{ is annihilated by 2.} \tag{13}$$

(One may mention that this is a special case of an observation made independently by Edixhoven [7] and André [1].)

Indeed, let ι_1 be an element of G acting non-trivially on τ_1 but trivially on τ_2. Then for any $\sigma \in N_1$ we have $\sigma^{\iota_1} = \sigma^{-1}$ by Proposition 3.2. On the other hand, for any $\sigma \in N_2$ we have $\sigma^{\iota_1} = \sigma$. It follows that $\sigma = \sigma^{-1}$ for $\sigma \in \tilde{N}$, proving (13).

Thus, each of the groups N_1 and N_2 has a subgroup of index 2 isomorphic to $(\mathbb{Z}/2\mathbb{Z})^\mu$ for some integer μ. Hence each of N_1 and N_2 is isomorphic either to $(\mathbb{Z}/2\mathbb{Z})^{\mu+1}$ or to $\mathbb{Z}/4\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})^{\mu-1}$.

If, say, $N_1 \cong (\mathbb{Z}/2\mathbb{Z})^{\mu+1}$, then L is Galois over $\mathbb{Q}$ by Corollary 3.3, contradicting (12). Therefore

$$N_1 \cong N_2 \cong \mathbb{Z}/4\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})^{\mu-1}.$$

Let $\iota \in G$ be the complex conjugation. Since i extends the non-trivial element of $\mathrm{Gal}(\mathbb{Q}(\tau_1)/\mathbb{Q})$, the group $H = \{1, \iota\}$ acts on N_1 dihedrally: $\sigma^\iota = \sigma^{-1}$ for $\sigma \in N_1$, and we have $G = N_1 \rtimes H$. Let N_1' and N_1'' be subgroups of N_1 isomorphic to $\mathbb{Z}/4\mathbb{Z}$ and $(\mathbb{Z}/2\mathbb{Z})^{\mu-1}$, respectively, such that $N_1 = N_1' \times N_1''$. Then H commutes with N_1'' and acts dihedrally on N_1'. Hence

$$G = N_1 \rtimes H = (N_1' \rtimes H) \times N_1'' \cong D_8 \times (\mathbb{Z}/2\mathbb{Z})^{\mu-1},$$

where D_{2n} denotes the dihedral group of $2n$ elements.

Now observe that $D_8 \times (\mathbb{Z}/2\mathbb{Z})^{\mu-1}$ has only one subgroup isomorphic to the group $\mathbb{Z}/4\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})^{\mu-1}$; this follows, for instance, from the fact that both groups have exactly 2^μ elements of order 4. Hence $N_1 = N_2$, which implies the equality $\mathbb{Q}(\tau_1) = \mathbb{Q}(\tau_2)$, a contradiction. □

Now one can go further and, inspecting all possible pairs of fields, produce the full list of number fields presented as $\mathbb{Q}(j(\tau_1))$ and $\mathbb{Q}(j(\tau_2))$ with $\mathbb{Q}(\tau_1) \neq \mathbb{Q}(\tau_2)$.

Corollary 4.2. *Let L be a number field with the following property: there exist quadratic τ_1 and τ_2 such that $L = \mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2))$ but $\mathbb{Q}(\tau_1) \neq \mathbb{Q}(\tau_2)$. Then L is one of the fields in Table 2.*

Table 2 Fields presented as $\mathbb{Q}(j(\tau_1))$ and $\mathbb{Q}(j(\tau_2))$ with $\mathbb{Q}(\tau_1) \neq \mathbb{Q}(\tau_2)$

Field L	$[L:\mathbb{Q}]$	Δ	$\mathrm{Cl}(\Delta)$
$\mathbb{Q}$	1	$-3,-4,-7,-8,-11,-12,-16,-19,-27,-28,-43,$ $-67,-163$	trivial
$\mathbb{Q}(\sqrt{2})$	2	$-24,-32,-64,-88$	$\mathbb{Z}/2\mathbb{Z}$
$\mathbb{Q}(\sqrt{3})$	2	$-36,-48$	$\mathbb{Z}/2\mathbb{Z}$
$\mathbb{Q}(\sqrt{5})$	2	$-15,-20,-35,-40,-60,-75,-100,-115,-235$	$\mathbb{Z}/2\mathbb{Z}$
$\mathbb{Q}(\sqrt{13})$	2	$-52,-91,-403$	$\mathbb{Z}/2\mathbb{Z}$
$\mathbb{Q}(\sqrt{17})$	2	$-51,-187$	$\mathbb{Z}/2\mathbb{Z}$
$\mathbb{Q}(\sqrt{2},\sqrt{3})$	4	$-96,-192,-288$	$(\mathbb{Z}/2\mathbb{Z})^2$
$\mathbb{Q}(\sqrt{3},\sqrt{5})$	4	$-180,-240$	$(\mathbb{Z}/2\mathbb{Z})^2$
$\mathbb{Q}(\sqrt{5},\sqrt{13})$	4	$-195,-520,-715$	$(\mathbb{Z}/2\mathbb{Z})^2$
$\mathbb{Q}(\sqrt{2},\sqrt{5})$	4	$-120,-160,-280,-760$	$(\mathbb{Z}/2\mathbb{Z})^2$
$\mathbb{Q}(\sqrt{5},\sqrt{17})$	4	$-340,-595$	$(\mathbb{Z}/2\mathbb{Z})^2$
$\mathbb{Q}(\sqrt{2},\sqrt{3},\sqrt{5})$	8	$-480,-960$	$(\mathbb{Z}/2\mathbb{Z})^3$

Explanations:

1. the third column contains the full list of discriminants Δ of CM-orders $\mathrm{End}\langle\tau,1\rangle$ such that $L = \mathbb{Q}(j(\tau))$;
2. the fourth column gives the structure of the class group $\mathrm{Cl}(\Delta)$ for any such Δ.

Proof. This is just a calculation using PARI. □

4.2 The Case $\mathbb{Q}(\tau_1) = \mathbb{Q}(\tau_2)$

Now assume that $\mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2))$ and $\mathbb{Q}(\tau_1) = \mathbb{Q}(\tau_2)$. Denote by D the discriminant of the number field $\mathbb{Q}(\tau_1) = \mathbb{Q}(\tau_2)$ and write $\Delta_i = f_i^2 D$ for $i = 1, 2$.

Proposition 4.3. *Assume that $\mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2))$ and $\mathbb{Q}(\tau_1) = \mathbb{Q}(\tau_2)$. Then either*

$$f_1/f_2 \in \{1, 2, 1/2\} \tag{14}$$

or $D = -3$ and $f_1, f_2 \in \{1, 2, 3\}$ (in which cases $\mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2)) = \mathbb{Q}$).

Proof. Put $f = \mathrm{LCM}(f_1, f_2)$. When $D \neq -3, -4$, Proposition 3.1 implies that

$$h(f^2 D) = [\mathbb{Q}(\sqrt{D}, j(\tau_1), j(\tau_2)) : \mathbb{Q}(\sqrt{D})] \tag{15}$$

Since $j(\tau_1)$ and $j(\tau_2)$ generate the same field, we obtain

$$h(f_1^2 D) = h(f_2^2 D) = h(f^2 D). \tag{16}$$

Using (6) and (16), we obtain

$$\frac{f}{f_1} \prod_{\substack{p \mid f \\ p \nmid f_1}} \left(1 - \left(\frac{D}{p}\right) p^{-1}\right) = 1,$$

which implies that $f/f_1 \in \{1, 2\}$. Similarly, $f/f_2 \in \{1, 2\}$. Hence we have (14).

Now assume that $D \in \{-3, -4\}$. If $\gcd(f_1, f_2) > 1$, then we again have (15), and the same argument proves (14).

If, say, $f_1 = 1$ then either $D = -4$ and $f_2 \in \{1, 2\}$, in which case we again have (14), or $D = -3$ and $f_2 \in \{1, 2, 3\}$.

Finally, assume that $f_1, f_2 > 1$ and $\gcd(f_1, f_2) = 1$. Then $f = f_1 f_2$ and Proposition 3.1 implies that

$$h(f_1^2 D) = h(f_2^2 D) = \ell^{-1} h(f^2 D), \tag{17}$$

where $\ell = 2$ for $D = -4$ and $\ell = 3$ for $D = -3$. Using (6) and (17), we obtain

$$f_i \prod_{p \mid f_i} \left(1 - \left(\frac{D}{p}\right) p^{-1}\right) = \ell \qquad (i = 1, 2),$$

and a quick inspection shows that in this case $D = -3$ and $\{f_1, f_2\} = \{2, 3\}$. □

5 Proof of Theorem 1.2

We assume that $P = (j(\tau_1), j(\tau_2))$ belongs to a non-special straight line ℓ defined over $\mathbb{Q}$, and show that it satisfies either (2) or (3). We define $\Delta_i = D_i f_i^2$ as in the beginning of Sect. 4.

Let $A_1 x_1 + A_2 x_2 + B = 0$ be the equation of ℓ. Since ℓ is not special, we have $A_1 A_2 \neq 0$, which implies, in particular, that

$$\mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2)). \tag{18}$$

We set

$$L = \mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2)), \qquad h = h(\Delta_1) = h(\Delta_2) = [L : \mathbb{Q}].$$

If $j(\tau_1), j(\tau_2) \in \mathbb{Q}$ (that is, $h = 1$), we are done. From now on assume that

$$j(\tau_1), j(\tau_2) \notin \mathbb{Q}.$$

If $j(\tau_1) = j(\tau_2)$, then ℓ is the special line $x_1 = x_2$, because it passes through the points $(j(\tau_1), j(\tau_1))$ and through all its conjugates over $\mathbb{Q}$. Hence

$$j(\tau_1) \neq j(\tau_2). \tag{19}$$

If $h = 2$, then we have (3). From now on assume that

$$h \geq 3, \tag{20}$$

and, in particular,

$$|\Delta_1|, |\Delta_2| \geq 23. \tag{21}$$

We will show that this leads to a contradiction.

Remark 5.1. Before proceeding with the proof, remark that, for a given pair of distinct discriminants Δ_1 and Δ_2 it is easy to verify whether there exists a point $(j(\tau_1), j(\tau_2))$ on a non-special straight line defined over $\mathbb{Q}$, such that Δ_i is the discriminant of the CM-order $\mathrm{End}\langle \tau_i, 1\rangle$. Call two polynomials $f(x), g(x) \in \mathbb{Q}[x]$ *similar* if there exist $\alpha, \beta, \lambda \in \mathbb{Q}$ with $\alpha\lambda \neq 0$ such that $f(\alpha x + \beta) = \lambda g(x)$. Now, a point $(j(\tau_1), j(\tau_2))$ as above exists if and only if the class polynomials H_{Δ_1} and H_{Δ_2} (see end of Sect. 3.3) are similar. This can be easily verified using, for instance, the PARI package.

5.1 Both Coordinates are Dominant

It turns out that we may assume, without a loss of generality, that both $j(\tau_1)$ and $j(\tau_2)$ are the *dominant j-values* of corresponding discriminants, as defined in Sect. 3.3.

Lemma 5.2. *Assume that ℓ is a non-special straight line containing a CM-point $P = (j(\tau_1), j(\tau_2))$ satisfying (18)–(20). Then ℓ contains a CM-point $P' = (j(\tau_1'), j(\tau_2'))$, conjugate to P over $\mathbb{Q}$ and such that $j(\tau_i')$ is the dominant j-value of discriminant Δ_i for $i = 1, 2$.*

Proof. Since ℓ is defined over $\mathbb{Q}$, all $\mathbb{Q}$-conjugates of P belong to ℓ as well. Replacing P by a $\mathbb{Q}$-conjugate point, we may assume that $j(\tau_1)$ is the dominant j-value for the discriminant Δ_1. If $j(\tau_2)$ is the dominant value for Δ_2 we are done; so assume it is not, and show that this leads to a contradiction.

Since both $j(\tau_1)$ and $j(\tau_2)$ generate the same field of degree h over $\mathbb{Q}$, the Galois orbit of P (over $\mathbb{Q}$) has exactly h elements; moreover, each conjugate of $j(\tau_1)$ occurs exactly once as the first coordinate of a point in the orbit, and each conjugate of $j(\tau_2)$ occurs exactly once as the second coordinate.

It follows that there is a conjugate point P^σ such that the second coordinate $j(\tau_2)^\sigma$ is the dominant j-value for Δ_2; then, its first coordinate $j(\tau_1)^\sigma$ is not dominant for Δ_1 because $P^\sigma \neq P$. Since $h \geq 3$, there exists yet another point $P^{\sigma'}$ with both coordinates not dominant for the respective discriminants.

All three points P, P^σ, and $P^{\sigma'}$ belong to ℓ. Hence

$$\begin{vmatrix} 1 & j(\tau_1) & j(\tau_2) \\ 1 & j(\tau_1)^\sigma & j(\tau_2)^\sigma \\ 1 & j(\tau_1)^{\sigma'} & j(\tau_2)^{\sigma'} \end{vmatrix} = 0.$$

The determinant above is a sum of 6 terms: the "dominant term" $j(\tau_1)j(\tau_2)^\sigma$ and 5 other terms. Each of the other terms is at most $0.1|j(\tau_1)j(\tau_2)^\sigma|$ in absolute value: this follows from Lemma 3.5, which applies here due to (21). Hence the determinant cannot vanish, a contradiction. □

5.2 *Completing the Proof*

After this preparation, we are ready to complete the proof of Theorem 1.2. Thus, let $P = (j(\tau_1), j(\tau_2))$ belong to a straight line ℓ defined over $\mathbb{Q}$. We assume that (18)–(20) are satisfied, and we may further assume that $j(\tau_1)$, $j(\tau_2)$ are the dominant j-values of Δ_1, Δ_2, respectively.

If $\mathbb{Q}(\tau_1) \neq \mathbb{Q}(\tau_2)$, then Corollary 4.2 applies, and all possible L, Δ_1, and Δ_2 can be found in Table 2. In particular, we have only 6 possible fields L and 15 possible couples Δ_1, Δ_2. All of the latter are ruled out by verifying (using PARI) that the corresponding Hilbert class polynomials $H_{\Delta_1}(x)$ and $H_{\Delta_2}(x)$ are not similar, as explained in Remark 5.1.

If $\mathbb{Q}(\tau_1) = \mathbb{Q}(\tau_2)$, then Proposition 4.3 applies, and we have $f_1/f_2 \in \{1, 2, 1/2\}$. Since both $j(\tau_1)$ and $j(\tau_2)$ are dominant, the case $f_1 = f_2$ is impossible: there is only one dominant j-value for every given discriminant, and we have $j(\tau_1) \neq j(\tau_2)$ by (19). Thus, $f_1/f_2 \in \{2, 1/2\}$.

Assume, for instance, that $f_2 = 2f_1$. Write $\Delta_2 = \Delta$, so that $\Delta_1 = 4\Delta$. Since both $j(\tau_1)$ and $j(\tau_2)$ are dominant, we may choose

$$\tau_1 = \frac{-r_4(4\Delta) + \sqrt{4\Delta}}{2} = \sqrt{\Delta}, \qquad \tau_2 = \frac{-r_4(\Delta) + \sqrt{\Delta}}{2}$$

It follows that $\tau_2 = \frac{1}{2}\gamma(\tau_1)$, where

$$\gamma = \begin{pmatrix} 1 & -r_4(\Delta) \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Hence the point $P = (j(\tau_1), j(\tau_2))$ belongs to the modular curve $Y_0(2)$ realized as the plane curve $\Phi_2(x_1, x_2) = 0$, where

$$\begin{aligned}\Phi_2(x_1, x_2) = &- x_1^2x_2^2 + x_1^3 + x_2^3 + 1488x_1^2x_2 + 1488x_1x_2^2 + 40773375x_1x_2 \\ &- 162000x_1^2 - 162000x_2^2 + 8748000000x_1 + 8748000000x_2 \\ &- 157464000000000\end{aligned}$$

is the modular polynomial of level 2. Since $\deg \Phi_2 = 4$ and P belongs to a straight line over $\mathbb{Q}$, the coordinates of P generate a field of degree at most 4 over $\mathbb{Q}$. Thus, $3 \leq h \leq 4$.

Looking into existing class number tables (or using PARI) one finds that there exist only 5 negative discriminants Δ such that $h(\Delta) = h(4\Delta) \in \{3, 4\}$:

$$\begin{aligned}h = 3: &\quad -23, -31; \\ h = 4: &\quad -7 \cdot 3^2, -39, -55.\end{aligned}$$

Verifying that the polynomials H_Δ and $H_{4\Delta}$ for these values of Δ are not similar is an easy calculation with PARI. □

Remark 5.3. We conclude the article with some computational remarks.

1. As indicated in the introduction, it is very easy to list all CM-points satisfying (2) or (3); call them *rational* and *quadratic*, respectively.

   There exist exactly 13 discriminants Δ with $h(\Delta) = 1$, the first line of Table 1 lists them all. Hence there exist 169 rational CM-points, and listing them explicitly is plainly straightforward.

   As for the quadratic CM-points, there are two kinds of them: points with

   $$\Delta_1 = \Delta_2 = \Delta, \quad h(\Delta) = 2, \quad j(\tau_1) \text{ and } j(\tau_2) \text{ are conjugate over } \mathbb{Q}, \tag{22}$$

   and points with

   $$\Delta_1 \neq \Delta_2, \quad h(\Delta_1) = h(\Delta_2) = 2, \quad \mathbb{Q}(j(\tau_1)) = \mathbb{Q}(j(\tau_2)). \tag{23}$$

   There exist exactly 29 negative quadratic discriminants Δ with $h(\Delta) = 2$, see the second line of Table 1 for the complete list. Hence there exist 29, up to conjugacy, quadratic CM-points satisfying (22).

   The quadratic CM-points satisfying (23) can be extracted from the "quadratic" part of Table 2. Indeed, if $\mathbb{Q}(\tau_1) \neq \mathbb{Q}(\tau_2)$, then Δ_1, Δ_2 are in Table 2 by Corollary 4.2. And if $\mathbb{Q}(\tau_1) = \mathbb{Q}(\tau_2)$, then $\Delta_1/\Delta_2 \in \{4, 1/4\}$ by Proposition 4.3. Inspecting the list of the 29 discriminants with $h(\Delta) = 2$, we find that the only possibility is $\{\Delta_1, \Delta_2\} = \{-15, -60\}$. Both these values appear in Table 2 in the line corresponding to the field $\mathbb{Q}(\sqrt{5})$.

Looking into Table 2 we find that there exist

$$4(4-1)+2(2-1)+9(9-1)+3(3-1)+2(2-1)=94$$

(ordered) pairs (Δ_1, Δ_2) as in (23). Each pair gives rise to two, up to conjugacy, points satisfying (23). So, up to conjugacy, there are 188 points satisfying (23), and $188+29=217$ quadratic CM-points altogether. Again, listing them explicitly is a straightforward computation.

2. Thomas Scanlon (private communication) asked whether there exists a non-special straight line over $\mathbb{C}$ passing through more than 2 CM-points. Since

$$\det\begin{bmatrix} 1728 & -884736000 \\ 287496 & -147197952000 \end{bmatrix} = 0,$$

the points $(0,0)$, $(1728, 287496)$ and $(-884736000, -147197952000)$ belong to the same straight line, and so do the points $(0,0)$, $(1728, -884736000)$ and $(287496, -147197952000)$. Notice that

$$j\left(\frac{-1+\sqrt{-3}}{2}\right)=0, \quad j(\sqrt{-1})=1728, \quad j(2\sqrt{-1})=287496,$$

$$j\left(\frac{-1+\sqrt{-43}}{2}\right)=-884736000, \quad j\left(\frac{-1+\sqrt{-67}}{2}\right)=-147197952000.$$

We verified that (up to switching the variables x_1, x_2) these are the only such lines defined over $\mathbb{Q}$. Precisely:

- no 3 rational CM-points, with the exceptions indicated above, lie on the same non-special line;
- no line passing through conjugate quadratic CM-points contains a rational CM-point;
- lines defined by pairs of conjugate quadratic CM-points are all pairwise distinct.

The verification is a quick calculation with PARI.

It is not clear to us whether the examples above (and absence of other examples) are just accidental, or admit some conceptual explanations.

3. All the computations for this paper take about 20 s.

Acknowledgements We thank Lars Kühne whose marvelous article [9] was our principal source of inspiration. We also thank Karim Belabas, Henri Cohen, Andreas Enge, and Jürg Kramer for useful conversations, and the referee for the encouraging report and many helpful comments.

Yuri Bilu was supported by the *Agence National de la Recherche* project "Hamot" (ANR 2010 BLAN-0115-01). Amalia Pizarro-Madariaga was supported by the ALGANT scholarship program.

Our calculations were performed using the PARI/GP package [14].

References

1. Y. André, Finitude des couples d'invariants modulaires singuliers sur une courbe algébrique plane non modulaire. J. Reine Angew. Math. **505**, 203–208 (1998)
2. A. Bajolet, Finding singular moduli on a complex line. Int. J. Number Theory **10**, 763–777 (2014)
3. Yu. Bilu, D. Masser, U. Zannier, An effective "Theorem of André" for CM-points on a plane curve. Math. Proc. Camb. Philos. Soc. **154**, 145–152 (2013)
4. Yu. Bilu, F. Luca, A. Pizarro-Madariaga, Rational products of singular moduli. Preprint (2014); arXiv:1410.1806; to appear in J. Number. Th.
5. F. Breuer, Heights of CM points on complex affine curves. Ramanujan J. **5**, 311–317 (2001)
6. D.A. Cox, *Primes of the Form* $x^2 + ny^2$ (Wiley, New York, 1989)
7. B. Edixhoven, Special points on the product of two modular curves. Compos. Math. **114**, 315–328 (1998)
8. L. Kühne, An effective result of André-Oort type. Ann. Math. (2) **176**, 651–671 (2012)
9. L. Kühne, An effective result of André-Oort type II. Acta Arith. **161**, 1–19 (2013)
10. J. Pila, Rational points of definable sets and results of André-Oort-Manin-Mumford type. Int. Math. Res. Not. **2009**, 2476–2507 (2009)
11. J. Pila, O-minimality and the André-Oort conjecture for $\mathbb{C}^n$. Ann. Math. (2) **173**, 1779–1840 (2011)
12. J. Pila, U. Zannier, Rational points in periodic analytic sets and the Manin-Mumford conjecture. Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **19**, 149–162 (2008)
13. T. Tatuzawa, On a theorem of Siegel. Jpn. J. Math. **21** (1951), 163–178 (1952)
14. The PARI Group, PARI/GP version 2.7.1 (2014), Bordeaux; available from http://pari.math.u-bordeaux.fr/
15. M. Watkins, Class numbers of imaginary quadratic fields. Math. Comput. **73**, 907–938 (2004)
16. P.J. Weinberger, Exponents of the class group of the complex quadratic fields. Acta Arith. **22**, 117–123 (1973)

Maass Waveforms and Low-Lying Zeros

Levent Alpoge, Nadine Amersi, Geoffrey Iyer, Oleg Lazarev, Steven J. Miller, and Liyang Zhang

To Professor Helmut Maier on his 60th birthday

Abstract The Katz–Sarnak Density Conjecture states that the behavior of zeros of a family of L-functions near the central point (as the conductors tend to zero) agrees with the behavior of eigenvalues near 1 of a classical compact group (as the matrix size tends to infinity). Using the Petersson formula, Iwaniec, Luo, and Sarnak proved that the behavior of zeros near the central point of holomorphic cusp forms agrees with the behavior of eigenvalues of orthogonal matrices for suitably restricted test functions ϕ. We prove similar results for families of cuspidal Maass forms, the other natural family of $\mathrm{GL}_2/\mathbb{Q}$ L-functions. For suitable weight functions on the space of Maass forms, the limiting behavior agrees with the expected orthogonal group. We prove this for $\mathrm{supp}(\hat{\phi}) \subseteq (-3/2, 3/2)$ when the level N tends to infinity through the square-free numbers; if the level is fixed the support decreases to being contained in $(-1, 1)$, though we still uniquely specify the symmetry type by computing the 2-level density.

L. Alpoge
Department of Mathematics, Princeton University, Princeton, NJ 08544, USA
e-mail: lalpoge@math.princeton.edu

N. Amersi
Department of Mathematics, University College London, London WC1E 6BT, UK
e-mail: n.amersi@ucl.ac.uk

G. Iyer
Department of Mathematics, UCLA, Los Angeles, CA 90095, USA
e-mail: geoff.iyer@gmail.com

O. Lazarev
Department of Mathematics, Stanford University, Stanford, CA 94305, USA
e-mail: olazarev@stanford.edu

S.J. Miller (✉)
Department of Mathematics & Statistics, Williams College, Williamstown, MA 01267, USA
e-mail: sjm1@williams.edu; Steven.Miller.MC.96@aya.yale.edu

L. Zhang
Department of Mathematics, Yale University, New Haven, CT 06520, USA
e-mail: zhangliyangmath@gmail.com

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_2

1 Introduction

In this section we set the stage for our results by quickly reviewing previous work on zeros of L-functions, leading up to n-level correlations, densities, and the conjectured correspondence with random matrix ensembles. As this is a vast field and the readership of this book is likely to have diverse backgrounds and interests, we discuss in some detail the history of the subject in order to put the present problems in context. We concentrate on some of the key theorems and statistics, and refer the reader to the extensive literature for more information. After this quick tour we describe the Katz–Sarnak conjectures for the behavior of low-lying zeros, and then in Sect. 2 we state our new results for families of Maass forms (the reader familiar with this field can skip this section and go straight to Sect. 2). The analysis proceeds by using the Kuznetsov trace formula to convert sums over zeros to exponential sums over the primes. Similar sums have been extensively studied by Maier in many papers over the years (see, for example, [12, 36–40]); it is a pleasure to dedicate this chapter to him on the occasion of his 60th birthday.

1.1 Zeros of L-Functions

The Riemann zeta function $\zeta(s)$ is defined for $\mathrm{Re}(s) > 1$ by

$$\zeta(s) \;:=\; \sum_{n=1}^{\infty} \frac{1}{n^s} \;=\; \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1}; \tag{1}$$

the Euler product expansion is equivalent to the Fundamental Theorem of Arithmetic on the unique factorization of integers into prime powers. Much can be gleaned in this regime. For example, looking at the limit as $s \to 1$ from above shows the sum of the reciprocals of the primes diverges (and with just a little work one gets $\sum_{p<x} 1/p \sim \log\log x$), and hence there are infinitely many primes. The true utility of this function, however, doesn't surface until we consider its meromorphic continuation $\xi(s)$ to the entire complex plane, where

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

The product expansion shows $\xi(s)$ has no zeros for $\mathrm{Re}(s) > 1$, and from the functional equation the only zeros for $\mathrm{Re}(s) < 0$ are at the negative even integers. The remaining zeros all have real part between 0 and 1; the Riemann Hypothesis [52] is the statement that these zeros all have real part 1/2.

Ever since Riemann's classic paper, researchers have exploited the connections between zeros of $\zeta(s)$ (and later other L-functions) to arithmetically important problems to translate information about the zeros to results in number theory.

For example, it can be shown that $\zeta(s)$ is never zero on the line $\mathrm{Re}(s) = 1$. This implies the Prime Number Theorem: the number of primes at most x, $\pi(x)$, is $\mathrm{Li}(x)$ plus a lower order term, where

$$\mathrm{Li}(x) := \int_2^x \frac{dt}{\log t}; \tag{3}$$

excellent references for this and the subsequent results are [7, 26].

Similar results about primes in arithmetic progressions modulo m follow from analogous results about the distribution of zeros of Dirichlet L-functions $L(s, \chi) := \sum_n \chi(n)/n^s$, where χ ranges over all primitive characters modulo m. It is worth noting that to study primes congruent to a modulo m it is not enough to study *one* specific Dirichlet L-function, but rather we need to understand the entire family coming from all characters of modulus m in order to invoke orthogonality relations to extract information about our progression from averages of $\chi(m)$; this notion of family will be very important in our work later.

After determining main terms, it is natural to ask about the form of the lower order terms. While the Riemann Hypothesis (RH) implies that $\pi(x) = \mathrm{Li}(x) + O(x^{1/2} \log x)$, neither it nor its generalization to other L-functions (GRH) is powerful enough to explain how the distribution of primes modulo m varies with the residue class, as these fluctuations are roughly the size of the errors from GRH. Chebyshev observed that there appeared to be more primes congruent to 3 modulo 4 than to 1 modulo 4. We now have an excellent theory (see [55]) that explains this phenomenon. A key ingredient is the Grand Simplicity Hypothesis, which asserts that the zeros of these L-functions are linearly independent over the rationals.

Assuming RH, the non-trivial zeros of $\zeta(s)$ all have real part equal to 1/2, and may thus be ordered on the line. It therefore makes sense to talk about spacings between adjacent zeros $\varrho_j = 1/2 + i\gamma_j$, or better yet spacings between adjacent normalized zeros (where we have normalized so that the average spacing is 1). Recent work has shown powerful connections between these gaps and important arithmetic quantities. For example, we can obtain excellent bounds on the size of the class groups of imaginary quadratic fields through knowing the existence of L-functions with multiple zeros at the central point [18, 19], or knowing that a positive percentage of gaps between normalized zeros of the Riemann zeta function are at least a certain fixed fraction of the average spacing [6].

The central theme in the above examples is that the more information we know about the zeros of the L-functions, the more we can say about arithmetically important questions. We started with just knowledge of the zeros on the line $\mathrm{Re}(s) = 1$, and then extended to GRH and all non-trivial zeros having real part 1/2, and then went beyond that to the distribution of the zeros on that critical line. In this chapter we expand on this last theme, and explore the distribution of zeros of L-functions on the critical line.

1.2 n-Level Correlations and Random Matrix Theory

While zeros of L-functions is a rich subject with an extensive history, our story on the number theory side begins in the 1970s with Montgomery's [47] work on the pair correlation of the zeros of $\zeta(s)$. Given an increasing sequence of numbers $\{\alpha_j\}$ and $B \subset \mathbb{R}^{n-1}$ a compact box, the n-level correlation $R_n(B)$ is defined by

$$R_n(B) := \lim_{N\to\infty} \frac{\#\left\{(\alpha_{j_1} - \alpha_{j_2}, \ \ldots, \ \alpha_{j_{n-1}} - \alpha_{j_n}) \in B, \ \ j_i \leq N\right\}}{N}, \tag{4}$$

where the indices above are distinct. Instead of using a box (which is equivalent to a sharp cut-off) it's often technically easier to consider a similar version with a smooth test function (see [56]).

While knowing *all* the n-level correlations allows one to determine all the neighbor spacings (see, for example, [41]), computing these for arbitrary B (or for any admissible test function) is well beyond current technology. There are, however, many important partial results. The first is the referred-to one of Montgomery [47], who showed that for suitable test functions the 2-level density agrees with the 2-level density of eigenvalues of the Gaussian Unitary Ensemble (GUE). There are many ways to view this ensemble of matrices. The easiest is that these are Hermitian matrices whose upper triangular entries are independently drawn from Gaussians (as the diagonal must be real, we draw from a different Gaussian for these entries than we do for the non-diagonal ones). An alternative definition, which explains the use of the word unitary, deals with the equality of the probability of choosing a matrix and its conjugation by a unitary matrix; note this is equivalent to saying the probability of a matrix is independent of the base used to write it down. From a physical point of view these matrices represent the Hamiltonian of a system. What matters are their eigenvalues, which correspond to the energy levels. While the entries of the matrix change depending on the basis used to write it down, the eigenvalues do not, which leads us to the unitary invariance condition. These and other matrix families had been extensively studied by Dyson, Mehta, and Wigner among many others; see [5, 14, 15, 21, 41, 46] and the multitude of references therein for more on the history and development of the subject.

This suggested a powerful connection between number theory and random matrix theory, which was further supported by Odlyzko's numerical investigations [48, 49] showing agreement between the spacings of zeros of $\zeta(s)$ and the eigenvalues of the GUE. Subsequent work by Hejhal [22] on the triple correlation of $\zeta(s)$ and Rudnick and Sarnak [56] on the n-level correlations for all automorphic cuspidal L-functions, again under suitable restrictions, provided additional support for the conjectured agreement between the limiting behavior of zeros of L-functions (as we move up the critical line) and eigenvalues of $N \times N$ matrices (as $N \to \infty$).

These results indicated a remarkable universality in behavior; while there are many random matrix ensembles, it appeared that only one was needed for number theory. A little thought, however, shows that this might not be the full story.

The reason is that the n-level correlations are insensitive to the behavior of finitely many zeros. In other words, we can remove finitely many zeros and not change $R_n(B)$. This is particularly troublesome, as there are many problems where only a few zeros matter. For example, the Birch and Swinnerton-Dyer conjecture [3, 4] states that the order of vanishing of the L-function associated with an elliptic curve equals the rank of its Mordell-Weil group; thus in studies on this problem we *only* care about what is happening at the central point, and not at all about what is happening far away on the critical line.

Later studies by Katz and Sarnak [28, 29] confirmed that more care is needed. The n-level correlations of the zeros of L-functions agree not only with those from the GUE, but also with those coming from the classical compact groups. The advantage of the latter is that the probability distribution a matrix is chosen is derived from the Haar measure on the group, which is a canonical choice. This is in sharp contrast to the definition of the GUE, where we fix a probability distribution and choose independent entries from it, which begs the question why one distribution was chosen over another (for the GUE, the answer is that the Gaussian is forced upon us by our assumption of the probability being invariant under unitary transformations of the basis). They proved that as $N \to \infty$ the n-level correlations of the eigenvalues are the same for all the classical compact groups (unitary, symplectic, and orthogonal, split or not split by sign). Thus one could just as easily say that the zeros of $\zeta(s)$ behave like the eigenvalues of orthogonal matrices instead of the GUE.

This led Katz and Sarnak to introduce a new statistic that is both able to distinguish the different classical compact groups and which depends on the behavior of eigenvalues near 1. We briefly describe the comparisons between number theory and random matrix theory. If we assume the Riemann hypothesis then the non-trivial zeros have real part 1/2 and we may write them as $\varrho_j = 1/2 + i\gamma_j$ for γ_j real. On the random matrix theory side, the classical compact groups are unitary matrices, and we can therefore write their eigenvalues as $e^{i\theta_k}$ with θ_k real. From intuition gleaned from earlier results, as well as function field analogues, Katz and Sarnak were led to conjecture that in the appropriate limits the behavior of zeros near 1/2 agree with the behavior of eigenvalues near 1 (more generally, one can also compare values of L-functions and characteristic polynomials of matrices).

1.3 n-Level Densities and the Katz–Sarnak Philosophy

Unfortunately, it is not possible to compare just the zeros of one L-function near the central point to the eigenvalues of one matrix. As in many problems in analytic number theory, we need to be able to execute some type of averaging and take some kind of limit in order to isolate out a main term and make progress. For the n-level correlations (or, equivalently, for Odlyzko's work on spacings between adjacent zeros), one L-function provides infinitely many zeros, and the average spacing between critical zeros at height T is on the order of $1/\log T$. Thus if we go high up,

we are essentially averaging over the zeros of that L-function, and can isolate out a universal, main term behavior. If instead we concentrate on the low-lying zeros, those near the central point, the situation is very different; due to the symmetry of the zeros about the central point we may restrict to studying the zeros in the upper half plane (this is why, in the definitions below, we use even test functions). To each L-function $L(s,f)$ we can associate a quantity, called the analytic conductor c_f, such that the first few zeros are of height $1/\log c_f$. If we rescale so that these zeros are of mean spacing one, then given any constant C there are essentially a finite number (depending basically just on C) that are at most C.

In order to make progress we need to collect a large number of L-functions which should behave similarly and are naturally connected. We call such a collection a *family* of L-functions. The definition of what is a family is still a work in progress (see [9] among others), but most natural collections of L-functions are. Examples include families of Dirichlet characters (either all of a given conductor, all whose conductor is in a given range, say $[N, 2N]$, or just quadratic characters whose conductor is in a range), cuspidal newforms (and very important subsets, one or two parameter families of elliptic curves), symmetric powers of cusp forms, and so on. Collections that are not families would include arbitrary subsets, for example, cusp forms whose third Fourier coefficient is 2 modulo 5, or cusp forms whose first zero above the central point is at least twice the average. Typically as the conductors (or range) grows we have more and more L-functions in the family. The Katz–Sarnak philosophy is that if we take averages of statistics of zeros over the family then in the limit it will converge and agree with the corresponding statistic for the scaling limit of a classical compact group as the matrix size tends to infinity.

The main statistic we study in this paper is their n-level density. For convenience of exposition we assume the Generalized Riemann Hypothesis for $L(s,f)$ (and thus all the zeros are of the form $1/2 + i\gamma_{j;f}$ with $\gamma_{j;f}$ real), though the statistic below makes sense even if GRH fails. Let $\phi(x) = \prod_{j=1}^{n} \phi_j(x_j)$ where each ϕ_j is an even Schwartz function such that the Fourier transforms

$$\widehat{\phi_j}(y) := \int_{-\infty}^{\infty} \phi_j(x) e^{-2\pi i x y} dx \tag{5}$$

are compactly supported. The n-level density for f with test function ϕ is

$$D_n(f,\phi) = \sum_{\substack{j_1,\ldots,j_n \\ j_\ell \neq j_m}} \phi_1\left(L_f \gamma_{j_1;f}\right) \cdots \phi_n\left(L_f \gamma_{j_n;f}\right), \tag{6}$$

where L_f is a scaling parameter which is frequently related to the conductor. The idea is to average over similar f, and use the explicit formula to relate this sum over zeros to a sum over the Fourier coefficients of the L-functions. See, for example, [27], the seminal paper in the subject and the first to explore these questions, and see [56] for a nice derivation of the explicit formula for general automorphic forms.

The subject is significantly harder if the conductors vary in the family, as then we cannot just pass the averaging over the forms through the test function to the Fourier coefficients of the L-functions. If we only care about the 1-level density, then we may rescale the zeros by using the average log-conductor instead of the log-conductor; as this is the primary object of study below we shall rescale all the L-functions in a family by the same quantity, which we denote R, and we emphasize this fact by writing $D_n(f,\phi,R)$. For more on these technical issues, see [42, 43], which studies families of elliptic curves where the variation in conductors must be treated. There it is shown that if the conductors vary within a family then this global renormalization leads to problems, and in the 2-level computations terms emerge where we cannot just pass the averaging through the test function. In some problems, however, it is important to compute the n-level densities. One application is to obtain significantly better bounds on the order of vanishing at the central point (see [23]). Another is to distinguish orthogonal candidates if the 1-level density can only be computed for small support; we will elaborate on this below.

Given a family $\mathscr{F} = \cup_N \mathscr{F}_N$ of L-functions with conductors tending to infinity, the n-level density $D_n(\mathscr{F},\phi,R;w)$ with test function ϕ, scaling R and a non-negative weight function w is defined by

$$D_n(\mathscr{F},\phi,R;w) := \lim_{N\to\infty} \frac{\sum_{f\in\mathscr{F}_N} w(f) D_n(f,\phi,R;w)}{\sum_{f\in\mathscr{F}_N} w(f)}. \tag{7}$$

The advantage of this statistic is that for a fixed f, individual zeros of $L(s,f)$ can contribute, with most of the contribution coming from the zeros near the central point due to the rapid decay of the test functions. Further, as we are averaging over similar forms there is a hope that there is a nice limiting behavior. In our applications later we will have weights $w_T(t) = w(t/T)$ with T a parameter tending to infinity, but we suppress the subscript T as it is always understood.

Katz and Sarnak [28, 29] proved that the n-level density is different for each classical compact group, and found nice determinant expansions for them. Set $K(y) := \frac{\sin \pi y}{\pi y}$ and $K_\epsilon(x,y) := K(x-y) + \epsilon K(x+y)$ for $\epsilon = 0, \pm 1$. They proved that if G_N is either the family of $N \times N$ unitary, symplectic or orthogonal families (split or not split by sign), the n-level density for the eigenvalues converges as $N \to \infty$ to

$$\begin{aligned} &\int \cdots \int \phi(x_1,\ldots,x_n) W_{n,G}(x_1,\ldots,x_n) dx_1 \cdots dx_n \\ &\quad = \int \cdots \int \hat{\phi}(y_1,\ldots,y_n) \hat{W}_{n,G}(y_1,\ldots,y_n) dy_1 \cdots dy_n, \end{aligned} \tag{8}$$

where

$$\begin{aligned} W_{m,\,\mathrm{SO(even)}}(x) &= \det(K_1(x_i,x_j))_{i,j\le m} \\ W_{m,\,\mathrm{SO(odd)}}(x) &= \det(K_{-1}(x_i,x_j))_{i,j\le m} + \sum_{k=1}^{m} \delta(x_k) \det(K_{-1}(x_i,x_j))_{i,j\ne k} \end{aligned}$$

$$\begin{aligned} W_{m,\mathrm{O}}(x) &= \frac{1}{2}W_{m,\ \mathrm{SO(even)}}(x) + \frac{1}{2}W_{m,\ \mathrm{SO(odd)}}(x) \\ W_{m,\mathrm{U}}(x) &= \det(K_0(x_i,x_j))_{i,j\le m} \\ W_{m,\mathrm{Sp}}(x) &= \det(K_{-1}(x_i,x_j))_{i,j\le m}. \end{aligned} \tag{9}$$

While these densities are all different, for the 1-level density with test functions whose Fourier transforms are supported in $(-1,1)$ the three orthogonal flavors cannot be distinguished from each other, though they can be distinguished from the unitary and symplectic densities. Explicitly, the Fourier Transforms for the 1-level densities are

$$\begin{aligned} \hat{W}_{1,\ \mathrm{SO(even)}}(u) &= \delta_0(u) + \frac{1}{2}\eta(u) \\ \hat{W}_{1,\mathrm{O}}(u) &= \delta_0(u) + \frac{1}{2} \\ \hat{W}_{1,\ \mathrm{SO(odd)}}(u) &= \delta_0(u) - \frac{1}{2}\eta(u) + 1 \\ \hat{W}_{1,\mathrm{Sp}}(u) &= \delta_0(u) - \frac{1}{2}\eta(u) \\ \hat{W}_{1,\mathrm{U}}(u) &= \delta_0(u), \end{aligned} \tag{10}$$

where $\eta(u)$ is 1, $1/2$, and 0 for $|u|$ less than 1, equal to 1, and greater than 1, respectively, and δ_0 is the standard Dirac Delta functional. Note that the first three densities agree for $|u| < 1$ and split (i.e., become distinguishable) for $|u| \ge 1$. Thus in order to uniquely specify a symmetry type among the three orthogonal candidates, one either needs to obtain results for support exceeding $(-1,1)$, or compute the 2-level density, as that is different for the three orthogonal groups for arbitrarily small support [42, 43].

The Katz–Sarnak Density Conjecture states that the behavior of zeros near the central point in a family of L-functions (as the conductors tend to infinity) agrees with the behavior of eigenvalues near 1 of a classical compact group (as the matrix size tends to infinity). For suitable test functions, this has been verified in many families, including Dirichlet characters, elliptic curves, cuspidal newforms, symmetric powers of GL(2) L-functions, and certain families of GL(4) and GL(6) L-functions; see, for example, [8, 9, 11, 13, 16, 17, 20, 23, 24, 27, 29, 32, 43, 45, 50, 51, 53, 54, 59, 60]. This correspondence between zeros and eigenvalues allows us, at least conjecturally, to assign a definite symmetry type to each family of L-functions (see [9, 57] for more on identifying the symmetry type of a family).

For this work, the most important families studied to date are holomorphic cusp forms. Using the Petersson formula (and a delicate analysis of the exponential sums arising from the Bessel–Kloosterman term), Iwaniec et al. [27] proved that the limiting behavior of the zeros near the central point of holomorphic cusp forms agrees with that of the eigenvalues of orthogonal matrices for suitably restricted test functions. In this chapter we look at the other $\mathrm{GL}_2/\mathbb{Q}$ family of L-functions, Maass waveforms.

2 Statement of Main Results

We first describe the needed normalizations and notation for our families of Maass forms, and then conclude by stating our new results and sketching the arguments. The beginning of the proofs is similar to that in all families studied to date: one uses the explicit formula to convert sums over zeros to sums over the Fourier coefficients of the forms. The difficulty is averaging over the family. In order to obtain support beyond $(-1, 1)$, we have to handle some very delicate exponential sums; these arise from the Bessel–Kloosterman term in the Kuznetsov trace formula. To facilitate applying it, we spend a lot of time choosing tractable weights. This is similar to previous work on cuspidal newforms where the harmonic weights were used to simplify the application of the Petersson trace formula. It is possible to remove these weights, and this is done in [27]. For some applications it is important to have unweighted families, in order to talk about the percentage of forms that vanish at the central point to a given order (see [23]); for our purposes, we are primarily interested in obtaining large enough support to uniquely determine the symmetry type, and thus choose our weight functions accordingly.

2.1 Normalizations and Notation

We quickly recall the basic properties of Maass forms (see [25, 26, 34, 35] for details), and then review the 1-level density from the last section with an emphasis on the important aspects for the subsequent computations. We use the standard conventions. Specifically, by $A \ll B$ we mean $|A| \leq c|B|$ for a positive constant c. Similarly, $A \asymp B$ means $A \ll B$ and $A \gg B$. We set $e(x) := \exp(2\pi i x)$, and define the Fourier transform of f by

$$\hat{f}(\xi) := \int_{\mathbb{R}} f(x)e(-x\xi)dx. \tag{11}$$

In the rest of the paper u always denotes a Hecke–Maass cusp form on $\Gamma_0(N)$ with N square-free. Thus u is an eigenfunction of the Laplacian with eigenvalue $\lambda_u =: (\frac{1}{2} + it_u)(\frac{1}{2} - it_u)$, and it is either even or odd with respect to the involution $z \mapsto -1/z$; if u is even we set $\epsilon = 0$, otherwise we take $\epsilon = 1$. Selberg's 3/16 ths theorem implies that we may take $t_u \geq 0$ or $t_u \in [0, \frac{1}{4}]i$. Next we Fourier expand u as follows:

$$u(z) = y^{1/2} \sum_{n \neq 0} a_n(u) K_{s-1/2}(2\pi|n|y)e(ny). \tag{12}$$

Let

$$\lambda_n(u) := \frac{a_n(u)}{\cosh(\pi t_u)^{1/2}}, \tag{13}$$

where the $K_\alpha(z)$ are the K-Bessel functions. We normalize u so that $\lambda_1(u) = 1$.

The L-function associated with u is

$$L(s,u) := \sum_{n\geq 1} \lambda_n n^{-s}. \tag{14}$$

By results from Rankin–Selberg theory the L-function is absolutely convergent in the right half-plane $\Re(s) > 1$ (one could also use the work of Kim and Sarnak [30, 31] to obtain absolutely convergent in the right half-plane $\Re(s) > 71/64$, which suffices for our purposes). These L-functions analytically continue to entire functions of the complex plane, satisfying the functional equation

$$\Lambda(s,u) = (-1)^\epsilon \Lambda(1-s,u), \tag{15}$$

with

$$\Lambda(s,u) := \pi^{-s}\Gamma\left(\frac{s+\epsilon+it}{2}\right)\Gamma\left(\frac{s+\epsilon-it}{2}\right)L(s,u). \tag{16}$$

Factoring

$$1-\lambda_p X + X^2 =: (1-\alpha_p X)(1-\beta_p X) \tag{17}$$

at each prime (the α_p, β_p are the Satake parameters at p not dividing the level N; if p divides the level, then either α_p or β_p is zero), we get an Euler product

$$L(s,u) = \prod_p (1-\alpha_p p^{-s})^{-1}(1-\beta_p p^{-s})^{-1}, \tag{18}$$

which again converges for $\Re(s)$ sufficiently large.

For the remainder of the paper $\mathscr{B}_N$ denotes an orthogonal basis of Maass cusp forms on $\Gamma_0(N)$, all normalized so that $\lambda_1 = 1$; thus, $\mathscr{B}_N$ is *not* orthonormal under the Petersson inner product on the space. Note we do not take a basis of *newforms*—that is, the delicate sieving out of oldforms as in [27] is not done. Using Weyl's law, however, one can control the sieving (see [25]).

We use the notation $\mathrm{Avg}(A; w)$ to denote the average of A over $\mathscr{B}_N$ with each element $u \in \mathscr{B}_N$ given weight $w(u)$. That is,

$$\mathrm{Avg}(A; w) := \frac{\sum_{u\in\mathscr{B}_N} A(u)w(u)}{\sum_{u\in\mathscr{B}_N} w(u)}. \tag{19}$$

Our main statistic for studying the low-lying zeros (i.e., the zeros near the central point) is the 1-level density; we quickly summarize the needed definitions and facts from Sect. 1.3. Let ϕ be an even Schwartz function such that the Fourier transform $\hat{\phi}$ of ϕ has compact support; that is,

$$\hat{\phi}(y) \ = \ \int_{-\infty}^{\infty} \phi(x) e^{-2\pi i x y} dx \tag{20}$$

and there is an $\eta < \infty$ such that $\hat{\phi}(y) = 0$ for y outside $(-\eta, \eta)$.

The 1-level density of the zeros of $L(s, u)$ is

$$D_1(u, \phi, R) \ = \ \sum_{\varrho} \phi\left(\frac{\log R}{2\pi}\gamma\right), \tag{21}$$

where $\varrho = 1/2 + i\gamma$ are the non-trivial zeros of $L(s, u)$, and $\log R$ is a rescaling parameter related to the average log-conductor in the weighted family, whose choice is forced upon us by (46). Under GRH all γ are real and the zeros can be ordered; while GRH gives a nice interpretation to the 1-level density, it is not needed for our purposes. As ϕ is a Schwartz function, most of the contribution comes from the zeros near the central point $s = 1/2$. The different classical compact groups (unitary, symplectic, and orthogonal) have distinguishable 1-level densities for arbitrarily small support; however, the 1-level densities for the even and odd orthogonal matrix ensembles are equal for test functions whose Fourier transforms are supported in $(-1, 1)$. There are two solutions to this issue. One possibility is to perform a more detailed analysis and "extend support." The other is to study the 2-level density, which Miller [42, 43] showed distinguishes the orthogonal ensembles for arbitrarily small support. For some of the families studied below we are able to calculate the support beyond $(-1, 1)$, and we may thus determine which of the orthogonal groups should be the symmetry group; for the other families our support is too limited and we instead study the 2-level density.

2.2 Main Results

Similar to how the harmonic weights facilitate applications of the Petersson formula to average the Fourier coefficients of cuspidal newforms (see, for instance, [27, 44]), we introduce nice, even weight functions to smooth the sum over the Maass forms. As we will see below, some type of weighting is necessary in order to restrict to conductors of comparable size. While our choice does not include the characteristic function of $[T, 2T]$, we are able to localize for the most part to conductors near T, and are able to exploit smoothness properties of the weight function in applications of the Kuznetsov trace formula. Further, in problems such as these the primary goal is to have as large support as possible for the Fourier transform of the test function

that hits the zeros. For more on these issues see [2], where Alpoge and Miller impose even more restrictions on the weight functions, which allows them to increase the support.

We consider the averaged one-level density weighted by two different weight functions of "nice" analytic properties. Let $\hat{H} \in C^\infty\left(\left(-\frac{1}{4}, \frac{1}{4}\right)\right)$ be an even smooth bump function of compact support on the real line (thus it should be close to 1 in some region and decay very quickly to zero outside there), and let H be its Fourier transform. We may of course (by applying this construction to a square root—recall that the support of a convolution is easily controlled) take $H \geq 0$. We may also take H to have an order K zero at 0. Let

$$H_T(r) \ := \ H\left(\frac{r}{T}\right). \tag{22}$$

This is essentially supported in a band of length $\asymp T$ about $\pm T$.

Next, in the same way, let $\hat{h} \in C^\infty\left(\left(-\frac{1}{4}, \frac{1}{4}\right)\right)$ be even. We also require h to have an order at least 8 zero at 0. Note that, by the same process as above, we may take $h(x) \geq 0$ for all $x \in \mathbb{R}$ and also (by Schwarz reflection) $h(ix) \geq 0$ for all $x \in \mathbb{R}$. Let T be a positive odd integer. We let

$$h_T(r) \ := \ \frac{\frac{r}{T} h\left(\frac{ir}{T}\right)}{\sinh\left(\frac{\pi r}{T}\right)}. \tag{23}$$

This is the same test function used in [2], and is essentially supported in a band of length $\asymp T$ about $\pm T$.

By trivially bounding the Fourier integral we observe that

$$H(x+iy),\ h(x+iy) \ \ll \ \exp\left(\frac{\pi|y|}{2}\right). \tag{24}$$

Hence

$$H_T(ir) \ \ll \ \exp\left(\frac{\pi|r|}{2T}\right), \tag{25}$$

and, using $\sinh(x) \gg e^{|x|}$, we find

$$h_T(r) \ \ll \ \exp\left(-\frac{\pi|r|}{4T}\right). \tag{26}$$

These will both be useful in what follows.

In one-level calculations we will take an even Schwartz function ϕ such that $\hat{\phi}$ is supported inside $[-\eta, \eta]$. We suppress the dependence of constants on h, H, ϕ, and η (as these are all fixed), but not T or the level N since one or both of these will be tending to infinity.

We weight each element $u \in \mathscr{B}_N$ by either $H_T(t_u)/||u||^2$ or $h_T(t_u)/||u||^2$, where

$$||u||^2 \;=\; ||u||^2_{\Gamma_0(N)\backslash\mathfrak{h}} \;=\; \frac{1}{[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)]} \int_{\Gamma_0(N)\backslash\mathfrak{h}} u(z) \frac{dxdy}{y^2} \tag{27}$$

is the L^2-norm of u on the modular curve $Y_0(N)$, and, as before, $\lambda_u = \frac{1}{4} + t_u^2$ is the Laplace eigenvalue of u. Recall that

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] \;=:\; \nu(N) \;=\; \prod_{p|N}(p+1). \tag{28}$$

The averaged weighted one-level density may thus be written (we will see in (46) that $R \asymp T^2 N$ is forced)

$$D_1(\mathscr{B}_N, \phi, R; w) \;:=\; \mathrm{Avg}\left(D_1(u, \phi, R); w(t_u)/||u||^2\right), \tag{29}$$

where $w(t_u)$ is either $H_T(t_u)$ or $h_T(t_u)$.

The main question is to determine the behavior of $D_1(\mathscr{B}_N, \phi, R; w)$ as either the level N or the weight parameter T tends to infinity; specifically, one is generally interested in the corresponding symmetry group. There are now several works [9, 28, 29, 57] which suggest ways to determine the symmetry group. For our family, they suggest the following conjecture.

Conjecture 2.1. Let h_T be as in (23), ϕ an even Schwartz function with $\hat{\phi}$ of compact support, and $R \asymp T^2 N$. Then

$$\lim_{R\to\infty} D_1\left(\mathscr{B}_N, \phi, R; w\right) \;=\; \int_{\mathbb{R}} \phi(t) W_{1,\mathrm{O}}(t) dt, \tag{30}$$

where $W_{1,\mathrm{O}} := 1 + \frac{1}{2}\delta_0$. In other words, the symmetry group associated with the family of Maass cusp forms of level N is orthogonal.

In [2] the above conjecture is shown for $N = 1, w = h_T$, with the extra restrictions that h has $2K \geq 8$ zeros at the origin and $\mathrm{supp}(\hat{\phi}) \subseteq (-2 + \frac{2}{2K+1}, 2 - \frac{2}{2K+1})$. Our first result here is in the case where the level N tends to infinity (remember that N must be square-free), T and K are fixed, and $w = w_T$ equals either h_T or H_T.

Theorem 2.2. *Fix T and K and let $R \asymp N$ with N square-free. Let H be an even, non-negative function with $K \geq 8$ zeros at 0 and Fourier transform $\hat{H} \in C^\infty\left(\left(-\frac{1}{4}, \frac{1}{4}\right)\right)$, and let h be an even function with 8 zeros at 0 and $\hat{h} \in C^\infty\left(\left(-\frac{1}{4}, \frac{1}{4}\right)\right)$. Let the weights $w = w_T$ be either H_T or h_T, where these are the functions given by (22) and (23), respectively. Let ϕ be an even Schwartz function with* $\mathrm{supp}(\hat{\phi}) \subseteq (-\frac{3}{2}, \frac{3}{2})$. *Then*

$$\lim_{\substack{N\to\infty \\ N \text{ square-free}}} D_1(\mathscr{B}_N, \phi, R; w_T) = \int_{\mathbb{R}} \phi(t) W_{1,\mathrm{O}}(t) dt. \tag{31}$$

Notice that the support in Theorem 2.2 exceeds $(-1, 1)$, and thus we have uniquely specified which orthogonal group is the symmetry group of the family.

Next we investigate the case where N is fixed and T tends to infinity through odd values. For ease of exposition we take $N = 1$.

Theorem 2.3. *Let h be an even function with 8 zeros at 0 and $\hat{h} \in C^\infty\left(\left(-\frac{1}{4}, \frac{1}{4}\right)\right)$, and define h_T as in (23). Let ϕ be an even Schwartz function with $\mathrm{supp}(\hat{\phi}) \subseteq (-1, 1)$, and take $R \asymp T^2$ (here $N = 1$). Then*

$$\lim_{\substack{T\to\infty \\ T \text{ odd}}} D_1(\mathscr{B}_1, \phi, R; h_T) = \int_{\mathbb{R}} \phi(t) W_{1,\mathrm{O}}(t) dt. \tag{32}$$

We also get a similar, though slightly weaker, result for the weight function $w = H_T$ if we allow $K = \mathrm{ord}_{z=0} H(z)$ to vary. For the argument given we invoke the work of [27] twice, since we reduce the Bessel–Kloosterman term of the Kuznetsov trace formula to sum of Kloosterman terms arising in the Petersson trace formula. Since [27] use GRH (specifically, for Dirichlet L-functions and L-functions associated with symmetric squares of holomorphic cusp forms), we must, too.

Theorem 2.4. *Assume GRH for Dirichlet L-functions and symmetric squares of holomorphic cusp forms of level 1. Let H be an even, non-negative function with K zeros at 0 and Fourier transform $\hat{H} \in C^\infty\left(\left(-\frac{1}{4}, \frac{1}{4}\right)\right)$, and let the weights $w = w_T$ be H_T, which is given by (22). Let ϕ be an even Schwartz function with $\mathrm{supp}(\hat{\phi}) \subseteq (-1 + \frac{1}{5+2K}, 1 - \frac{1}{5+2K})$. Take $R \asymp T^2$ (here $N = 1$). Then*

$$\lim_{T\to\infty} D_1(\mathscr{B}_1, \phi, R; H_T) = \int_{\mathbb{R}} \phi(t) W_{1,\mathrm{O}}(t) dt. \tag{33}$$

Notice the support in Theorems 2.3 and 2.4 is too small to uniquely determine which orthogonal symmetry is present (this is because the one-level densities of the orthogonal flavors all agree inside $(-1, 1)$). At the cost of more technical arguments, Alpoge and Miller [2] are able to extend the support beyond $(-1, 1)$ when weighting by h_T, thereby determining the symmetry group to be orthogonal. In this work we instead compute the 2-level density, which provides a second proof that the symmetry type of the family of Maass cusp forms on $\mathrm{SL}_2(\mathbb{Z})$ is orthogonal. The 2-level density is defined in (36). As any support for the 2-level density suffices to uniquely determine the symmetry group, we do not worry about obtaining optimal results.

Theorem 2.5. *Let w_T equal h_T or H_T, $R \asymp T^2$ (here $N = 1$), and let*

$$\mathscr{N}(-1) := \frac{1}{\sum_{u\in\mathscr{B}_1} \frac{w_T(t_u)}{||u_u||^2}} \sum_{\substack{u\in\mathscr{B}_1 \\ (-1)^\epsilon = -1}} \frac{w_T(t_u)}{||u||^2} \tag{34}$$

be the weighted percentage of Maass forms in $\mathscr{B}_1$ with odd functional equation. Write

$$D_2(\mathscr{B}_1, \phi_1, \phi_2, R; w_T) := \text{Avg}\left(D_2(u, \phi_1, \phi_2, R); w_T(t_u)/||u||^2\right), \tag{35}$$

with

$$\begin{aligned} D_2(u, \phi_1, \phi_2, R) &:= \sum_{i \neq \pm j} \phi_1\left(\frac{\log R}{2\pi}\gamma_i\right)\phi_2\left(\frac{\log R}{2\pi}\gamma_j\right) \\ &= D_1(u, \phi_1, R)D_1(u, \phi_2, R) \\ &\quad - 2D_1(u, \phi_1\phi_2, R) \\ &\quad + \delta_{\epsilon,1}\phi_1(0)\phi_2(0), \end{aligned} \tag{36}$$

the average 2-level density of the weighted family of level 1 Maass cusp forms, and the 2-level density of $u \in \mathscr{B}_1$, respectively. Let $f * g$ denote the convolution of f and g. Then, for $\delta \ll 1$ and $\widehat{\phi_1}, \widehat{\phi_2}$ even Schwartz functions supported in $(-\delta, \delta)$,

$$\begin{aligned} \lim_{T\to\infty} D_2(\mathscr{B}_1, \phi_1, \phi_2, R; w_T) &= \left(\frac{\phi_1(0)}{2} + \widehat{\phi_1}(0)\right)\left(\frac{\phi_2(0)}{2} + \widehat{\phi_2}(0)\right) \\ &\quad + 2\int_{\mathbb{R}} |x|\widehat{\phi_1}(x)\widehat{\phi_2}(x)dx \\ &\quad - (1 - \mathscr{N}(-1))\phi_1(0)\phi_2(0) - 2(\widehat{\phi_1} * \widehat{\phi_2})(0), \end{aligned} \tag{37}$$

agreeing with the 2-level density of the scaling limit of an orthogonal ensemble with proportions of $\mathscr{N}(-1)$ SO(odd) matrices and $1 - \mathscr{N}(-1)$ SO(even) matrices.

A similar result holds for $\mathscr{B}_N$—all the calculations will be standard given our work on the one-level densities.

2.3 Outline of Arguments

By a routine application of the explicit formula we immediately reduce the problem to studying averages of Hecke eigenvalues over the space of Maass cusp forms of level N. For this we apply the Kuznetsov trace formula, as found in [33]. We are quickly reduced to studying a term of shape

$$\nu(N)\sum_{c\geq 1}\frac{S(m, 1; cN)}{cN}\int_{\mathbb{R}} J_{2ir}\left(\frac{4\pi\sqrt{m}}{cN}\right)\frac{rw_T(r)}{\cosh(\pi r)}dr. \tag{38}$$

In all cases the idea is to move the contour from $\mathbb{R}$ to $\mathbb{R} - iY$ with $Y \to \infty$. The properties of the weights h_T or H_T ensure that the integral along the moving line vanishes in the limit, so all that is left in place of the integral is the sum over poles, of shape (up to negligible error in the case of h_T, which also has poles of its own)

$$\sum_{k\geq 0}(-1)^k J_{2k+1}\left(\frac{4\pi\sqrt{m}}{cN}\right)(2k+1)w_T\left(\frac{2k+1}{2}i\right). \tag{39}$$

Now the $N \to \infty$ limit is very easy to take, as all the Bessel functions involved have zeros at 0. So the term does not contribute (the total mass is of order $T^2\nu(N)$, canceling the $\nu(N)$ out in front). With some care we arrive at Theorem 2.2.

If instead we take $N = 1$ and $w_T = H_T$, then, by standard bounds on Bessel functions, $J_{2k+1}(\frac{4\pi\sqrt{m}}{c})$ is very small for k larger than $\asymp \frac{\sqrt{m}}{c}$. For us $\sqrt{m}$ will always be bounded in size by something that is $\asymp T^\eta$. Thus for k smaller than this range, the Bessel term is still controlled but not too small. It is the term

$$w_T\left(\frac{2k+1}{2}i\right) \asymp H\left(\frac{2k+1}{2T}i\right) \ll \left(\frac{k}{T}\right)^K \tag{40}$$

that is small. Upon taking

$$K \gg \frac{1}{1-\eta} \tag{41}$$

it is in fact small enough to bound trivially. For slightly larger support we instead appeal to the bounds of [27] on sums of Kloosterman sums, which are derived from assuming GRH for Dirichlet L-functions. This gives Theorem 2.4. In fact, the expression we get is exactly a weighted sum of terms appearing in [27] from the Kloosterman terms of Petersson formulas. It would be interesting to find a conceptual explanation for this.

The proof of Theorem 2.3 is a simplified version of the argument given in [2], except considerably shortened—instead of delicate analysis of exponential sums, we just use Euler–Maclaurin summation. As one would expect our support is thus smaller than that in [2], but the argument and main ideas are significantly easier to see.

The proof of Theorem 2.5 follows from the previous results and another application of the Kuznetsov formula, this time to the inner product of T_{p^ℓ} with $T_{q^{\ell'}}$ with p, q primes.

3 Preliminaries for the Proofs

In this section we compute and analyze some expansions and resulting expressions that are useful in the proofs of our main theorems. We start in Sect. 3.1 by using the explicit formula to relate the sum over zeros to sums over the Hecke eigenvalues of

the associated cusp forms. The weights and normalizations are chosen to facilitate applying the Kuznetsov trace formula to these sums, which we do. After trivially handling several of the resulting terms, in Sect. 3.2 we analyze the Bessel function integral that arises. We then use these results in Sect. 4 to prove the stated theorems.

3.1 Calculating the Averaged One-Level Density

We first quickly review the computation of the explicit formula; see [27, 56] for details. Let $u \in \mathscr{B}_N$, and for an even Schwartz test function ϕ set

$$\Phi(s) \ := \ \phi\left(\frac{\left(s-\frac{1}{2}\right)\log R}{2\pi i}\right). \tag{42}$$

Consider

$$\int_{\sigma=\frac{3}{2}} \Phi(s)\frac{\Lambda'}{\Lambda}(s,u)ds. \tag{43}$$

By moving the integration to $\sigma = -\frac{1}{2}$ and applying the functional equation, we find that

$$2\int_{\sigma=\frac{3}{2}} \frac{\Lambda'}{\Lambda}(s,u)\Phi(s)ds \ = \ D_1(u,\phi) \tag{44}$$

(use the rapid decay of ϕ along horizontal lines and Phragmen–Lindelöf to justify the shift). After expanding the logarithmic derivative out in the usual way, applying the Kim–Sarnak bound, and noticing that $\lambda_{p^2} = \lambda_p^2 - \chi_0(p)$ for χ_0 the principal character modulo N, this equality simplifies to

$$\begin{aligned} D_1(u,\phi) \ &= \ \frac{\phi(0)}{2} + \hat{\phi}(0)\left(\frac{\log N + \log(1+t_u^2)}{\log R}\right) + O\left(\frac{\log\log R + \log\log N}{\log R}\right) \\ &+ \ 2\sum_{\ell=1}^{2}\sum_{p}\frac{\lambda_{p^\ell}\log p}{p^{\frac{\ell}{2}}\log R}\hat{\phi}\left(\frac{\ell\log p}{\log R}\right). \end{aligned} \tag{45}$$

Thus, if w_T is essentially supported on $\asymp T$ (as are h_T and H_T), the averaged one-level density is (since $||u|| \asymp 1$ under our normalizations, by Smith [58])

$$\begin{aligned} D_1(\mathscr{B}_N,\phi,R;w_T) \ &= \ \frac{\phi(0)}{2} + \hat{\phi}(0)\left(\frac{\log(T^2N)}{\log R}\right) + O\left(\frac{\log\log R + \log\log N}{\log R}\right) \\ &+ 2\sum_{\ell=1}^{2}\sum_{p}\frac{\log p}{p^{\frac{\ell}{2}}\log R}\hat{\phi}\left(\frac{\ell\log p}{\log R}\right)\mathrm{Avg}(\lambda_{p^\ell};w_T). \end{aligned} \tag{46}$$

Notice the above computation tells us the correct scaling to use is $R \asymp T^2N$.

The difficulty is in determining the averages over Hecke eigenvalues. For this we use the Kuznetsov formula for $\mathscr{B}_N$. Let w_T equal h_T or H_T.

Theorem 3.3 (Kuznetsov Trace Formula (See [33], p. 86)). *Let $m \in \mathbb{Z}^+$. Then*

$$\begin{aligned}\sum_{u\in\mathscr{B}_N}\frac{\lambda_m(u)}{||u||^2}w_T(t_u) &= \frac{\delta_{m,1}\nu(N)}{\pi^2}\int_{\mathbb{R}} rw_T(t)\tanh(\pi r)dr \\ &\quad -\frac{1}{\pi}\sum_{(i_p)_{p|N}\in\{0,1\}^{\omega(N)}}\int_{\mathbb{R}}\frac{\tilde{\sigma}_{ir}(m,(i_p))\overline{\tilde{\sigma}_{ir}(1,(i_p))}m^{ir}w_T(r)}{||(i_p)||^2|\zeta(1+2ir)|^2}dr \\ &\quad +\frac{2i}{\pi}\frac{\nu(N)}{N}\sum_{c\geq 1}\frac{S(m,1;Nc)}{c}\int_{\mathbb{R}}J_{2ir}\left(\frac{4\pi\sqrt{m}}{Nc}\right)\frac{rw_T(r)}{\cosh(\pi r)}dr,\end{aligned} \tag{47}$$

where S is the usual Kloosterman sum, $(i_p)_{p|N}$ runs through all $0 \leq i_p \leq 1$ with p ranging over the prime factors of N,

$$\begin{aligned}&\tilde{\sigma}_{ir}(a,(i_p)) := \\ &\left(\prod_{p|N}p^{i_p}\right)^{-1-2ir}\sum_{d|a}\frac{\chi_0(d \bmod \prod_{p|N}p^{1-i_p})}{d^{2ir}}\sum_{f\in\left(\mathbb{Z}/\prod_{p|N}p^{i_p}\mathbb{Z}\right)^\times}e\left(\frac{af}{d\prod_{p|N}p^{i_p}}\right),\end{aligned} \tag{48}$$

J is the usual Bessel function, and

$$||(i_p)||^2 = \prod_{p|N}\frac{p^{1-i_p}}{1+p} = \frac{N}{\nu(N)\prod_{p|N}p^{i_p}}. \tag{49}$$

In our applications we always have $(m,N) = 1$ since we only take $m = 1, p$, or p^2, which means that the contribution from the principal character in the definition of $\tilde{\sigma}_{ir}$ may be ignored. Also the inner sum in $\tilde{\sigma}_{ir}(a,(i_p))$ is of the form

$$\sum_{\xi\in(\mathbb{Z}/n\mathbb{Z})^\times}e\left(\frac{\xi}{n}\right) = \mu(n) \ll 1. \tag{50}$$

Hence, bounding trivially and noting that our m have at most three divisors, we find

$$\tilde{\sigma}_{ir}(a,(i_p)) \ll \left(\prod_{p|N}p^{i_p}\right)^{-1}. \tag{51}$$

Also, by work of de la Vallée Poussin on the prime number theorem, $\zeta(1+2ir) \gg \log(2+|r|)^{-1}$. Hence the second term in (47), the Eisenstein contribution, is

$$\begin{aligned} &\ll \frac{\nu(N)}{N} \sum_{(i_p)_{p|N}} \frac{1}{\prod_{p|N} p^{i_p}} \int_{\mathbb{R}} w_T(r) \log(2+|r|) dr \\ &\ll \frac{\nu(N) T \log T}{N} \prod_{p|N} \left(1+\frac{1}{p}\right) \\ &= \left(\frac{\nu(N)}{N}\right)^2 T \log T. \end{aligned} \tag{52}$$

In our applications we will always divide these expressions by the corresponding expression with $m = 1$, which gives the total mass of the family ("the denominator" in the sequel). We will see that it is of order $\asymp T^2 \nu(N)$ (see Corollary 3.4). Hence, since it will be divided by something of order $\asymp T^2 \nu(N)$, the Eisenstein contribution is thus negligible for N or T large.

Note that the diagonal term (that is, the first term of (47), with $m = 1$) is

$$\nu(N) \int_{\mathbb{R}} r w_T(r) \tanh(\pi r) dr \asymp T^2 \nu(N). \tag{53}$$

Hence to show the claim about the total mass it suffices to bound the last term of (47) in the case of $m = 1$.

We have therefore reduced the computation of the weighted 1-level density to understanding the "Bessel–Kloosterman" terms. We isolate this result below.

Lemma 3.2. *If $m = 1, p$ or p^2 is coprime to N and w_T equals h_T or H_T, then*

$$\begin{aligned} \sum_{u \in \mathcal{B}_N} \frac{\lambda_m(u)}{||u||^2} w_T(t_u) &= \delta_{m,1} \cdot \left(\asymp T^2 N\right) \\ &\quad + O\left(\left(\frac{\nu(N)}{N}\right)^2 T \log T\right) \\ &\quad + \frac{2i}{\pi} \frac{\nu(N)}{N} \sum_{c \geq 1} \frac{S(m,1;Nc)}{c} \int_{\mathbb{R}} J_{2ir}\left(\frac{4\pi\sqrt{m}}{Nc}\right) \frac{r w_T(r)}{\cosh(\pi r)} dr, \end{aligned} \tag{54}$$

where $\delta_{m,1} \cdot \left(\asymp T^2 N\right)$ is the product of a term on the order of $T^2 N$ with Kronecker's delta.

3.2 Handling the Bessel Integral

As in [2], the technical heart of the analysis of the Kuznetsov formula is the following claim, which relies on the analytic properties of h_T and H_T; see Appendix 1 for a proof.

Proposition 3.3. *Let T be an odd integer. Let $X \leq T$. Let w_T equal h_T or H_T, where these are the weight functions from Theorems 2.2 through 2.5. Then*

$$\int_{\mathbb{R}} J_{2ir}(X)\frac{rw_T(r)}{\cosh(\pi r)}dr = c_1\sum_{k\geq 0}(-1)^k J_{2k+1}(X)(2k+1)w_T\left(\left(k+\frac{1}{2}\right)i\right)$$
$$\left[+c_2T^2\sum_{k\geq 1}(-1)^k J_{2kT}(X)k^2h(k)\right] \tag{55}$$
$$= c_1\sum_{k\geq 0}(-1)^k J_{2k+1}(X)(2k+1)w_T\left(\left(k+\frac{1}{2}\right)i\right)$$
$$\left[+O\left(Xe^{-c_3T}\right)\right], \tag{56}$$

where c_1, c_2, and c_3 are constants independent of X and T, and the terms in brackets are included if and only if $w_T = h_T$.

Since Bessel functions of integer order are much better-studied objects than those of purely imaginary order, this is a useful reduction. The calculation also realizes the Kloosterman term in the Kuznetsov formula as a sort of average (though the "weight function" is growing exponentially in the case of H_T) of Kloosterman terms arising in Petersson formulas over all even weights.

For what follows we quickly note the following corollary, which determines the size of the denominator (total mass) mentioned above.

Corollary 3.4. *Let w_T equal h_T or H_T (as above). Then*

$$\sum_{u\in\mathcal{B}_N}\frac{w_T(t_u)}{||u||^2} \asymp T^2\nu(N). \tag{57}$$

Proof. It suffices to show that

$$\frac{\nu(N)}{N}\sum_{c\geq 1}\frac{S(1,1;Nc)}{c}\sum_{k\geq 0}(-1)^k J_{2k+1}\left(\frac{4\pi}{cN}\right)(2k+1)w_T\left(\left(k+\frac{1}{2}\right)i\right) \ll \frac{\nu(N)T}{N}. \tag{58}$$

To see this, we bound trivially by using

$$J_k(x) \ll \frac{(x/2)^k}{k!}, \tag{59}$$

and

$$\sin\left(\frac{2k+1}{2T}\pi\right) \gg T^{-1} \tag{60}$$

in the case of $w_T = h_T$.

4 Proofs of the Main Theorems

Using the results from the previous section, we can now prove our main theorems. All arguments begin with the following reductions. We first use (46) to reduce the determination of the 1-level density to that of sums of the weighted averages of λ_p and λ_{p^2}. We then use the Kuznetsov trace formula (Theorem 3.3) to analyze these sums. By Lemma 3.2 we are reduced to bounding the contribution from the Bessel–Kloosterman term, to which we apply Proposition 3.3 to analyze these exponential sums. We now turn to the details of each of these cases.

4.1 Proof of Theorem 2.2

It suffices to study

$$\mathscr{S} := \sum_{\ell=1}^{2}\sum_{p}\frac{2\log p}{p^{\frac{\ell}{2}}\log R}\hat{\phi}\left(\frac{\ell\log p}{\log R}\right)\frac{\nu(N)}{N}\sum_{c\geq 1}\frac{S(p^\ell,1;Nc)}{c}$$
$$\cdot\sum_{k\geq 0}(-1)^k J_{2k+1}\left(\frac{4\pi p^{\frac{\ell}{2}}}{cN}\right)(2k+1)w_T\left(\frac{2k+1}{2}i\right), \tag{61}$$

and bound $\mathscr{S}$ by something growing strictly slower than $\nu(N)$. This is because we get to divide this term by the total mass, which by Corollary 3.4 is of the order $T^2\nu(N)$. As T is fixed, we are dividing by a quantity on the order of $\nu(N)$.

Bounding trivially, we find

$$\begin{aligned}\mathscr{S} &\ll \frac{\nu(N)}{N}\sum_{c\geq 1}\frac{1}{c}\sum_{k\geq 0}\frac{(2\pi)^{2k+1}w_T\left(\frac{2k+1}{2}i\right)}{(2k)!}\sum_{\ell=1}^{2}\sum_{p^{\ell/2}\leq R^{\eta/2}}\frac{2\log p}{\log R}(cN)^{\frac{1}{2}+\epsilon-2k-1}p^{k\ell}\\ &\ll \frac{\nu(N)}{N^{\frac{3}{2}-\epsilon}}\sum_{k\geq 0}\frac{(2\pi)^{2k}}{(2k)!}w_T\left(\frac{2k+1}{2}i\right)\frac{R^{\eta(k+1)}}{N^{2k}\log R}\\ &\ll \nu(N)\frac{N^{\eta-\frac{3}{2}+\epsilon}T^{2\eta}}{\log N+\log T}e^{O_T(T^{2\eta}N^{\eta-2})}.\end{aligned} \tag{62}$$

As T is fixed, the above is negligible for $\eta < 3/2$, which completes the proof. □

4.2 *Proof of Theorem 2.3*

It suffices to study

$$\mathscr{S} := \sum_{\ell=1}^{2} \sum_{p} \frac{\log p}{p^{\frac{\ell}{2}} \log T} \hat{\phi}\left(\frac{\ell \log p}{2 \log T}\right) \sum_{c \geq 1} \frac{S(p^{\ell}, 1; c)}{c} \cdot \sum_{k \geq 0} (-1)^k J_{2k+1}\left(\frac{4\pi p^{\frac{\ell}{2}}}{c}\right) (2k+1) h_T\left(\frac{2k+1}{2} i\right) \tag{63}$$

and bound $\mathscr{S}$ by something growing strictly slower than T^2. This is because N is fixed, so by Corollary 3.4 the denominator that occurs in the weighted averages is on the order of T^2.

In [2] (see their (43)–(46)—for the convenience of the reader, this argument is reproduced in (Appendix 2), it is proved that

$$\begin{aligned} &\sum_{k \geq 0} (-1)^k J_{2k+1}\left(\frac{4\pi p^{\frac{\ell}{2}}}{c}\right) (2k+1) h_T\left(\frac{2k+1}{2} i\right) \\ &\quad = c_4 T \sum_{|\alpha| < \frac{T}{2}} e\left(Y \sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{\tilde{h}}\left(\frac{\pi Y}{T} \cos\left(\frac{\pi\alpha}{T}\right)\right) + O(Y) \\ &\quad =: c_4 T S_h(Y) + O(Y), \end{aligned} \tag{64}$$

where $\tilde{\tilde{h}}(x) := x^2 h(x)$ and $Y := 2p^{\ell/2}/c$. We apply the Euler–Maclaurin summation formula to the first term, yielding

$$\begin{aligned} S_h(Y) \;=\; &\int_{-\frac{T}{2}}^{\frac{T}{2}} e\left(Y \sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{\tilde{h}}\left(\frac{\pi Y}{T} \cos\left(\frac{\pi\alpha}{T}\right)\right) d\alpha \\ &+ \sum_{k=2}^{M} \frac{B_k}{k!} \left(e\left(Y \sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{\tilde{h}}\left(\frac{\pi Y}{T} \cos\left(\frac{\pi\alpha}{T}\right)\right)\right)^{(k)} \Bigg|_{-\frac{T}{2}}^{\frac{T}{2}} \\ &+ O\left(\int_{-\frac{T}{2}}^{\frac{T}{2}} \left|\left(e\left(Y \sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{\tilde{h}}\left(\frac{\pi Y}{T} \cos\left(\frac{\pi\alpha}{T}\right)\right)\right)^{(M)}\right| d\alpha\right). \end{aligned} \tag{65}$$

In differentiating the expression

$$e\left(Y \sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{\tilde{h}}\left(\frac{\pi Y}{T} \cos\left(\frac{\pi\alpha}{T}\right)\right) \tag{66}$$

k times, the worst case is when we differentiate the exponential every single time and pick up a factor of $\left(\frac{Y}{T}\right)^k$; otherwise, we gain at least one factor of T (remember that Y should be thought of as order T^η). Hence we may bound the error term by

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} \left| \left(e\left(Y \sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{\tilde{h}}\left(\frac{\pi Y}{T}\cos\left(\frac{\pi\alpha}{T}\right)\right)\right)^{(M)} \right| d\alpha \ \ll \ \left(\frac{Y}{T}\right)^M T. \tag{67}$$

Taking $M \geq 1 + \frac{1}{1-\eta}$, the error term is thus $O(Y/T)$.

Next, by the same analysis, in the second term of (65) we either differentiate the exponential every single time, or we gain a factor of T from differentiating $\tilde{h}$ or one of the $\cos(\pi\alpha/T)$'s produced from differentiating the exponential. Thus, since we differentiate at least twice, all but one term in the k-fold derivative is bounded by Y/T^2. The last remaining term, obtained by differentiating the exponential k times, vanishes because $\tilde{\tilde{h}}$ has a zero at 0.

Hence it remains to bound the first term of (65). This we do by integrating by parts, via

$$\int e^{\phi(x)} f(x) dx \ = \ -\frac{1}{2\pi i} \int e^{\phi(x)} \left(\frac{f(x)}{\phi'(x)}\right)' dx. \tag{68}$$

We get

$$\begin{aligned} &\int_{-\frac{T}{2}}^{\frac{T}{2}} e\left(Y \sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{\tilde{h}}\left(\frac{\pi Y}{T}\cos\left(\frac{\pi\alpha}{T}\right)\right) d\alpha \\ &\quad = \ c_5 \frac{Y}{T^2} \int_{-\frac{T}{2}}^{\frac{T}{2}} e\left(Y \sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{h}'\left(\frac{\pi Y}{T}\cos\left(\frac{\pi\alpha}{T}\right)\right) \sin\left(\frac{\pi\alpha}{t}\right) d\alpha \\ &\quad \ll \ \frac{Y}{T}. \end{aligned} \tag{69}$$

Hence we obtain the bound

$$S_h(Y) \ \ll \ \frac{Y}{T}. \tag{70}$$

Thus

$$\sum_{k\geq 0} (-1)^k J_{2k+1}\left(\frac{4\pi p^{\ell/2}}{c}\right)(2k+1) h_T\left(\frac{2k+1}{2} i\right) \ \ll \ \frac{p^{\ell/2}}{c}. \tag{71}$$

That is, this tells us that

$$\begin{aligned}
&\sum_{\ell=1}^{2}\sum_{p}\frac{\log p}{p^{\ell/2}\log T}\hat{\phi}\left(\frac{\ell\log p}{2\log T}\right)\sum_{c\geq 1}\frac{S(p^{\ell},1;c)}{c}\\
&\qquad\cdot\left[\sum_{k\geq 0}(-1)^{k}J_{2k+1}\left(\frac{4\pi p^{\ell/2}}{c}\right)(2k+1)h_{T}\left(\frac{2k+1}{2}i\right)\right]\\
&\ll\sum_{\ell=1}^{2}\sum_{p}\frac{\log p}{p^{\ell/2}\log T}\left|\hat{\phi}\left(\frac{\ell\log p}{2\log T}\right)\right|\sum_{c\geq 1}c^{-\frac{1}{2}+\epsilon}\left[\frac{p^{\ell/2}}{c}\right]\\
&\ll\sum_{\ell=1}^{2}\sum_{p^{\ell/2}\leq T^{\eta}}\frac{\log p}{\log T}\\
&\ll\frac{T^{2\eta}}{\log T}. \qquad (72)
\end{aligned}$$

For $\eta < 1$ this is negligible upon division by the total mass (which is of order T^2), completing the proof. □

4.3 *Proof of Theorem 2.4*

Before proceeding it bears repeating that the same limiting support as K gets large (namely, $(-1,1)$) can be achieved by just trivially bounding as above (that is, without exploiting cancelation in sums of Kloosterman sums), but we present here an argument connecting the Kuznetsov formula to the Petersson formula as studied in [27] instead.

It suffices to study

$$\begin{aligned}
\mathscr{S}\ :=\ &\sum_{\ell=1}^{2}\sum_{p}\frac{\log p}{p^{\ell/2}\log T}\hat{\phi}\left(\frac{\ell\log p}{2\log T}\right)\sum_{c\geq 1}\frac{S(p^{\ell},1;c)}{c}\\
&\cdot\sum_{k\geq 0}(-1)^{k}J_{2k+1}\left(\frac{4\pi p^{\frac{\ell}{2}}}{c}\right)(2k+1)H_{T}\left(\frac{2k+1}{2}i\right) \qquad (73)
\end{aligned}$$

and bound $\mathscr{S}$ by something growing strictly slower than T^2. Let

$$Q_{k}^{*}(m;c)\ :=\ 2\pi i^{k}\sum_{p}S(p,1;c)J_{k-1}\left(\frac{4\pi m\sqrt{p}}{c}\right)\hat{\phi}\left(\frac{\log p}{\log R}\right)\frac{2\log p}{\sqrt{p}\log R}, \qquad (74)$$

exactly as in [27]. Then this simplifies to (dropping overall constants)

$$\sum_{c\geq 1} c^{-1} \sum_{k\geq 0} (2k+1) H\left(\frac{2k+1}{2T} i\right) Q^*_{2k+2}(1;c) \tag{75}$$

if we ignore the $\ell = 2$ term, which is insignificant by, e.g., GRH for symmetric square L-functions on $\mathrm{GL}_3/\mathbb{Q}$ (as in [27]).

In Sects. 6 and 7 (see (7.1)) of [27] they prove

Theorem 4.1. *Assume GRH for all Dirichlet L-functions. Then*

$$Q^*_k(m;c) \ll \tilde{\gamma}_k(z) m T^\eta k^\epsilon (\log(2c))^{-2} \tag{76}$$

where

$$\tilde{\gamma}_k(z) := \begin{cases} 2^{-k} & k \geq 3z \\ k^{-1/2} & \textit{otherwise} \end{cases} \tag{77}$$

and

$$z := \frac{4\pi T^\eta}{c}. \tag{78}$$

Hence the sum over c now converges with no problem, and we may ignore it. What remains is

$$\begin{aligned} &T^\eta \sum_{k\geq 0} (2k+1) H\left(\frac{2k+1}{2T} i\right) \tilde{\gamma}_k(z) k^\epsilon \\ &\quad = T^\eta \left(\sum_{0\leq k \ll T^\eta} k^{\epsilon+\frac{1}{2}} H\left(\frac{2k+1}{2T} i\right) + \sum_{T^\eta \ll k} \frac{k^{1+\epsilon} H\left(\frac{2k+1}{2T} i\right)}{2^k} \right). \end{aligned} \tag{79}$$

The second term in parentheses poses no problem. For the first term, using $H(x) \ll x^K$ for $0 \leq x \leq 1$, we see that the above is bounded by

$$T^\eta \sum_{0\leq k \ll T^\eta} k^{\epsilon+\frac{1}{2}+K} T^{-K} \ll T^{\frac{5}{2}\eta-(1-\eta)K}. \tag{80}$$

Thus we need

$$\eta < \frac{2+K}{\frac{5}{2}+K} = 1 - \frac{1}{5+2K}. \tag{81}$$

Again, by taking K even larger we could have just trivially bounded throughout and not invoked [27] or GRH, but the connection noted above may be of independent interest.

4.4 *Proof of Theorem 2.5*

By definition

$$D_2(u,\phi_1,\phi_2,R) = D_1(u,\phi_1,R)D_1(u,\phi_2,R) - 2D_1(u,\phi_1\phi_2,R) + \delta_{\epsilon,1}\phi_1(0)\phi_2(0). \quad (82)$$

Averaging, we see that for $\widehat{\phi_1}$ and $\widehat{\phi_2}$ of sufficiently small support (actually $\left(-\frac{1}{2},\frac{1}{2}\right)$ would work fine), given our results above on one-level densities, up to negligible error

$$D_2(\mathscr{B}_1,\phi_1,\phi_2,R;w_T) = \mathrm{Avg}\left(D_1(u,\phi_1,R)D_1(u,\phi_2,R);\frac{w_T(t_u)}{||u||^2}\right) - (1-\mathscr{N}(-1))\phi_1(0)\phi_2(0) - 2\widehat{\phi_1}*\widehat{\phi_2}(0) + o(1). \quad (83)$$

Now

$$D_1(u,\phi_1,R)D_1(u,\phi_2,R) = \left(\widehat{\phi_1}(0)\frac{\log(1+t_u^2)}{\log R} + \frac{\phi_1(0)}{2} - \sum_{\ell=1}^{2}\sum_p \frac{2\log p}{p^{\frac{\ell}{2}}\log R}\lambda_{p^\ell}\widehat{\phi_1}\left(\frac{\ell\log p}{\log R}\right)\right) \cdot \left(\widehat{\phi_2}(0)\frac{\log(1+t_u^2)}{\log R} + \frac{\phi_2(0)}{2} - \sum_{\ell=1}^{2}\sum_p \frac{2\log p}{p^{\frac{\ell}{2}}\log R}\lambda_{p^\ell}\widehat{\phi_2}\left(\frac{\ell\log p}{\log R}\right)\right). \quad (84)$$

Since w_T is supported essentially around $t_u \asymp T$, the $\log(1+t_u^2)$ terms are all, up to negligible error, approximately $\log R$ (again we invoke the bound of [58] on the L^2-norms occurring in the denominator). Also by an application of the Kuznetsov formula, now with the inner product of two Hecke operators (namely T_{p^ℓ} and $T_{q^{\ell'}}$ for p,q primes and $0 \le \ell,\ell' \le 2$—this uses the results on the Bessel–Kloosterman term established above), we see that the resulting average is, up to negligible error,

$$\mathrm{Avg}\left(D_1(u,\phi_1,R)D_1(u,\phi_2,R);\frac{w_T(t_u)}{||u||^2}\right) = \left(\widehat{\phi_1}(0)+\frac{\phi_1(0)}{2}\right)\left(\widehat{\phi_2}(0)+\frac{\phi_2(0)}{2}\right) + \sum_{\ell=1}^{2}\sum_p \frac{4\log^2 p}{p^\ell\log^2 R}\widehat{\phi_1}\left(\frac{\ell\log p}{\log R}\right)\widehat{\phi_2}\left(\frac{\ell\log p}{\log R}\right). \quad (85)$$

That is, only the diagonal terms $p^\ell = q^{\ell'}$ matter. Now partial summation (and the prime number theorem, as usual) finishes the calculation.

Appendix 1: Contour Integration

We prove Proposition 3.3 below. We restate it for the reader's convenience.

Proposition 5.1. *Let T be an odd integer and $X \leq T$. Let w_T equal h_T or H_T, where these are the weight functions from Theorems 2.2 to 2.5. Then*

$$\int_{\mathbb{R}} J_{2ir}(X)\frac{rw_T(r)}{\cosh(\pi r)}dr = c_1\sum_{k\geq 0}(-1)^k J_{2k+1}(X)(2k+1)w_T\left(\left(k+\frac{1}{2}\right)i\right)$$
$$\left[+c_2T^2\sum_{k\geq 1}(-1)^k J_{2kT}(X)k^2h(k)\right] \tag{86}$$
$$= c_1\sum_{k\geq 0}(-1)^k J_{2k+1}(X)(2k+1)w_T\left(\left(k+\frac{1}{2}\right)i\right)$$
$$\left[+O\left(Xe^{-c_3T}\right)\right], \tag{87}$$

where c_1, c_2, and c_3 are constants independent of X and T, and the terms in brackets are included if and only if $w_T = h_T$.

Proof. In the proof below bracketed terms are present if and only if $w_T = h_T$.

Recall that

$$J_\alpha(2x) = \sum_{m\geq 0}\frac{(-1)^m x^{2m+\alpha}}{m!\Gamma(m+\alpha+1)}. \tag{88}$$

By Stirling's formula, $\Gamma(m+2ir+1)\cosh(\pi r) \gg |m+2ir+1|^{m+\frac{1}{2}}e^{-m}$. Hence by the Lebesgue Dominated Convergence Theorem (remember that $w_T(z)$ is of rapid decay as $|\Re z| \to \infty$) we may switch sum and integral to get

$$\int_{\mathbb{R}} J_{2ir}(X)\frac{rw_T(r)}{\cosh(\pi r)}dr$$
$$= \sum_{m\geq 0}\frac{(-1)^m x^{2m}}{m!}\int_{\mathbb{R}}\frac{x^{2ir}rw_T(r)}{\Gamma(m+1+2ir)\cosh(\pi r)}dr, \tag{89}$$

where $X =: 2x$.

Now we move the line of integration down to $\mathbb{R} - iR$, $R \notin \mathbb{Z} + \frac{1}{2}$ (and $\notin T\mathbb{Z}$ if $w_T = h_T$). To do this, we note the estimate (for $A \gg 1$)

$$\int_{\pm A\to\pm A-iR}\frac{x^{2ir}rw_T(r)}{\Gamma(m+1+2ir)\cosh(\pi r)}dr \ll \int_{\pm A\to\pm A-iR} e^m x^{2R}|r||m+2ir+1|^{-m-\frac{1}{2}}|w_T(r)|dr \ll_{T,R} A^{-2}, \tag{90}$$

where again we have used the rapid decay of w_T along horizontal lines, and $B\to B-iR$ denotes the vertical line from $B\in\mathbb{C}$ to $B-iR\in\mathbb{C}$. (Rapid decay also ensures the integral along $\mathbb{R}-iR$ converges absolutely.)

Note that the integrand

$$\frac{x^{2ir}rw_T(r)}{\Gamma(m+1+2ir)\cosh(\pi r)} \tag{91}$$

has poles below the real axis precisely at $r\in -i(\mathbb{N}+\frac{1}{2})=\{-\frac{1}{2}i,-\frac{3}{2}i,\ldots\}$, and, if $w_T=h_T$, poles also at $r\in -iT\mathbb{Z}^+$. The residue of the pole at $r=-\frac{2k+1}{2}i$ $(k\geq 0)$ is, up to an overall constant independent of k,

$$\frac{x^{2k+1}}{\Gamma(m+1+(2k+1))}(-1)^k(2k+1)w_T\left(\frac{2k+1}{2}i\right). \tag{92}$$

If $w_T=h_T$, the residue of the pole at $r=-ikT$ $(k\geq 1)$ is, up to another overall constant independent of k,

$$\frac{x^{2kT}}{\Gamma(m+1+(2kT))}(-1)^k k^2T^2\frac{h(k)}{\cos(\pi kT)} = \frac{x^{2kT}}{\Gamma(m+1+(2kT))}k^2T^2h(k). \tag{93}$$

Hence the sum of (89) becomes

$$\begin{aligned}\sum_{m\geq 0}\frac{(-1)^m x^{2m}}{m!}\int_{\mathbb{R}}\frac{x^{2ir}rw_T(r)}{\Gamma(m+1+2ir)\cosh(\pi r)}dr \\ = c_6\sum_{m\geq 0}\frac{(-1)^m x^{2m}}{m!}\Bigg(\sum_{0\leq k\ll R}\frac{x^{2k+1}}{\Gamma(m+1+(2k+1))}(-1)^k(2k+1)w_T \\ \left(\frac{2k+1}{2}i\right)+\int_{\mathbb{R}-iR}\frac{x^{2ir}rw_T(r)}{\Gamma(m+1+2ir)\cosh(\pi r)}dr \\ \left[+c_7\sum_{m\geq 0}\frac{(-1)^m x^{2m}}{m!}\sum_{0\leq k\ll R}\frac{x^{2kT}}{\Gamma(m+1+(2kT))}(-1)^k k^2T^2h(k)\right]\Bigg).\end{aligned} \tag{94}$$

Now we take $R \to \infty$. Note that

$$\int_{\mathbb{R}-iR} \frac{x^{2ir} r w_T(r)}{\Gamma(m+1+2ir)\cosh(\pi r)} dr \ll \int_{\mathbb{R}-iR} x^{2R}|r||w_T(r)||m+2ir+1|^{-m-\frac{1}{2}-2R} dr$$
$$\ll_m x^{2R} e^{\frac{\pi R}{2T}} R^{-m-\frac{1}{2}-2R}, \tag{95}$$

again by Stirling and rapid decay of w_T on horizontal lines (that is, $w_T(x+iy) \ll (1+x)^{-4} e^{\frac{\pi y}{2T}}$ since both h and H have all their derivatives supported in $\left(-\frac{1}{4}, \frac{1}{4}\right)$). This of course vanishes as $R \to \infty$.

Hence we see that

$$\int_{\mathbb{R}} J_{2ir}(X) \frac{r w_T(r)}{\cosh(\pi r)} dr$$
$$= c_8 \sum_{m\geq 0} \frac{(-1)^m x^{2m}}{m!} \sum_{k\geq 0} \frac{x^{2k+1}}{\Gamma(m+1+(2k+1))} (-1)^k (2k+1) w_T\left(\frac{2k+1}{2} i\right)$$
$$\left[+ c_9 \sum_{m\geq 0} \frac{(-1)^m x^{2m}}{m!} \sum_{k\geq 0} \frac{x^{2kT}}{\Gamma(m+1+(2kT))} k^2 T^2 h(k) \right]. \tag{96}$$

Switching sums (via the exponential bounds on w_T along the imaginary axis) and applying $J_n(X) = \sum_{m\geq 0} \frac{(-1)^m x^{2m+n}}{m!(m+n)!}$ gives us the claimed calculation. For the bound on the bracketed term, use

$$J_n(ny) \ll \frac{y^n e^{n\sqrt{1-y^2}}}{(1+\sqrt{1-y^2})^n} \tag{97}$$

(see [1], p. 362) and bound trivially.

Appendix 2: An Exponential Sum Identity

The following proposition and proof are also used in [2].

Proposition 6.1. *Suppose $X \leq T$. Then*

$$S_J(X) := T \sum_{k\geq 0} (-1)^k J_{2k+1}(X) \frac{\tilde{\tilde{h}}\left(\frac{2k+1}{2T}\right)}{\sin\left(\frac{2k+1}{2T}\pi\right)}$$
$$= c_{10} T \sum_{|\alpha|<\frac{T}{2}} e\left(Y \sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{\tilde{g}}\left(\frac{\pi Y}{T} \cos\left(\frac{\pi\alpha}{T}\right)\right) + O(Y), \tag{98}$$

where c_{10} is some constant, $g(x) := \mathrm{sgn}(x) h(x)$, $X =: 2\pi Y$, and for any f set $\tilde{f}(x) := xf(x)$.

Proof. Observe that $k \mapsto \sin(\pi k/2)$ is supported only on the odd integers, and maps $2k+1$ to $(-1)^k$. Hence, rewriting gives

$$S_J(X) \;=\; T \sum_{\substack{k\geq 0\\ k\notin 2T\mathbb{Z}}} J_k(X)\tilde{\tilde{h}}\left(\frac{k}{2T}\right)\frac{\sin\left(\frac{\pi k}{2}\right)}{\sin\left(\frac{\pi k}{2T}\right)}. \tag{99}$$

Since

$$\frac{\sin\left(\frac{\pi k}{2}\right)}{\sin\left(\frac{\pi k}{2T}\right)} \;=\; \frac{e^{\frac{\pi ik}{2}}-e^{-\frac{\pi ik}{2}}}{e^{\frac{\pi ik}{2T}}-e^{-\frac{\pi ik}{2T}}} \;=\; \sum_{\alpha=-\left(\frac{T-1}{2}\right)}^{\frac{T-1}{2}} e^{\frac{\pi ik\alpha}{T}} \tag{100}$$

when k is not a multiple of $2T$, we find that

$$S_J(X) \;=\; T \sum_{|\alpha|<\frac{T}{2}} \sum_{\substack{k\geq 0\\ k\notin 2T\mathbb{Z}}} e\left(\frac{k\alpha}{2T}\right) J_k(X)\tilde{\tilde{h}}\left(\frac{k}{2T}\right). \tag{101}$$

Observe that, since the sum over α is invariant under $\alpha \mapsto -\alpha$ (and it is non-zero only for k odd!), we may extend the sum over k to the entirety of $\mathbb{Z}$ at the cost of a factor of 2 and of replacing h by

$$g(x) \;:=\; \mathrm{sgn}(x)h(x). \tag{102}$$

(This is because of the identity $J_{-k}(x) = (-1)^k J_k(x)$ and the fact that we need only consider odd k.) Note that g is as differentiable as h has zeros at 0, less one. That is to say, $\hat{g}$ decays like the reciprocal of a degree $\mathrm{ord}_{z=0}\, h(z) - 1$ polynomial at ∞. This will be crucial in what follows.

Next, we add back on the $2T\mathbb{Z}$ terms and obtain

$$\begin{aligned}\frac{1}{2}S_J(X) \;&=\; T \sum_{|\alpha|<\frac{T}{2}} \sum_{k\in\mathbb{Z}} e\left(\frac{k\alpha}{2T}\right) J_k(X)\tilde{\tilde{g}}\left(\frac{k}{2T}\right) - T^2 \sum_{k\in\mathbb{Z}} J_{2kT}(X)k^2h(k)\\ &=\; T \sum_{|\alpha|<\frac{T}{2}} \sum_{k\in\mathbb{Z}} e\left(\frac{k\alpha}{2T}\right) J_k(X)\tilde{\tilde{g}}\left(\frac{k}{2T}\right) + O\left(Xe^{-c_{11}T}\right)\\ &=:\; V_J(X) + O\left(Xe^{-c_{11}T}\right),\end{aligned} \tag{103}$$

where we have bounded the term $T^2\sum_{k\in\mathbb{Z}} J_{2kT}(X)k^2h(k)$ trivially via $J_n(2x) \ll x^n/n!$.

Now we move to apply Poisson summation. Write $X =: 2\pi Y$. We apply the integral formula (for $k \in \mathbb{Z}$)

$$J_k(2\pi x) \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} e\left(kt - x\sin(2\pi t)\right) dt \tag{104}$$

and interchange the sum and integral (via rapid decay of g) to get that

$$V_J(X) = T \sum_{|\alpha|<\frac{T}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\sum_{k\in\mathbb{Z}} e\left(\frac{k\alpha}{2T} + kt\right) \tilde{\tilde{g}}\left(\frac{k}{2T}\right)\right) e\left(-Y\sin(2\pi t)\right) dt. \tag{105}$$

By Poisson summation, (105) is just (interchanging the sum and integral once more)

$$\begin{aligned} V_J(X) &= T^2 \sum_{|\alpha|<\frac{T}{2}} \sum_{k\in\mathbb{Z}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{g}''\left(2T(t-k)+\alpha\right) e\left(-Y\sin(2\pi t)\right) dt \\ &= c_{12} T \sum_{|\alpha|<\frac{T}{2}} \int_{-\infty}^{\infty} \hat{g}''(t) e\left(Y\sin\left(\frac{\pi t}{T} + \frac{\pi\alpha}{T}\right)\right) dt \\ &=: c_{12} W_g(X). \end{aligned} \tag{106}$$

As

$$\sin\left(\frac{\pi t}{T} + \frac{\pi\alpha}{T}\right) = \sin\left(\frac{\pi\alpha}{T}\right) + \frac{\pi t}{T}\cos\left(\frac{\pi\alpha}{T}\right) - \frac{\pi^2 t^2}{2T^2}\sin\left(\frac{\pi\alpha}{T}\right) + O\left(\frac{t^3}{T^3}\right), \tag{107}$$

we see that

$$\begin{aligned} &W_g(X) \\ &= c_{13} T \sum_{|\alpha|<\frac{T}{2}} e\left(Y\sin\left(\frac{\pi\alpha}{T}\right)\right) \int_{-\infty}^{\infty} \hat{g}''(t) e\left(\frac{\pi Y t}{T}\cos\left(\frac{\pi\alpha}{T}\right)\right) dt \\ &\quad - c_{14}\frac{\pi^2 Y}{T} \sum_{|\alpha|<\frac{T}{2}} e\left(Y\sin\left(\frac{\pi\alpha}{T}\right)\right) \sin\left(\frac{\pi\alpha}{T}\right) \int_{-\infty}^{\infty} t^2 \hat{g}''(t) e\left(\frac{\pi Y t}{T}\cos\left(\frac{\pi\alpha}{T}\right)\right) dt \\ &\quad + O\left(\frac{Y}{T} + \frac{Y^2}{T^2}\right) \end{aligned} \tag{108}$$

$$\begin{aligned} &= c_{15} T \sum_{|\alpha|<\frac{T}{2}} e\left(Y\sin\left(\frac{\pi\alpha}{T}\right)\right) \tilde{\tilde{g}}\left(\frac{\pi Y}{T}\cos\left(\frac{\pi\alpha}{T}\right)\right) \\ &\quad - c_{16}\frac{Y}{T} \sum_{|\alpha|<\frac{T}{2}} e\left(Y\sin\left(\frac{\pi\alpha}{T}\right)\right) \sin\left(\frac{\pi\alpha}{T}\right) \tilde{\tilde{g}}''\left(\frac{\pi Y}{T}\cos\left(\frac{\pi\alpha}{T}\right)\right) \\ &\quad + O\left(\frac{Y}{T} + \frac{Y^2}{T^2}\right). \end{aligned} \tag{109}$$

Now bound the second term trivially to get the claim.

Appendix 3: 2-Level Calculations

The purpose of this appendix is to provide additional details to the 2-level computation of Sect. 4.4. Letting

$$\begin{aligned} S_1(u_j, \phi_i) &:= \sum_p \frac{2\lambda_p(u_j)\log p}{p^{1/2}\log R}\hat{\phi}_i\left(\frac{\log p}{\log R}\right) \\ S_2(u_j, \phi_i) &:= \sum_p \frac{2\lambda_{p^2}(u_j)\log p}{p\log R}\hat{\phi}_i\left(\frac{2\log p}{\log R}\right) \end{aligned} \tag{110}$$

and summing over the family, a standard calculation reduces the determination of the 2-level density to understanding

$$\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}\prod_{i=1}^{2}\left[\widehat{\phi_i}(0)\frac{\log(1+t_j^2)}{\log R} - S_1^{c(i)}(u_j,\phi_i) - S_2^{c(i)}(u_j,\phi_i) + O\left(\frac{\log\log R}{\log R}\right)\right]^2 \tag{111}$$

(the other terms are straightforward consequences of the combinatorial bookkeeping from the inclusion–exclusion argument), where $c(1)$ is the identity map and $c(2)$ denotes complex conjugation.

We can move the factor $O(\log\log R/\log R)$ outside the product at the cost of an error of the same size outside all the summations. To see this, since $O(\log\log R/\log R)$ is independent of u_j these terms are readily bounded by applying the Cauchy-Schwarz inequality. Letting $\mathscr{S}$ represent any of the factors in the product over i in (111), the product involving this is $O(\log\log R/\log R)$:

$$\begin{aligned} &\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}\mathscr{S}\cdot O\left(\frac{\log\log R}{\log R}\right) \\ &\ll \left[\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}O\left(\left(\frac{\log\log R}{\log R}\right)^2\right)\right]^{1/2}\cdot\left[\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}|\mathscr{S}|^2\right]^{1/2} \\ &\ll O\left(\frac{\log\log R}{\log R}\right)\cdot\left(\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}|\mathscr{S}|^2\right)^{1/2}. \end{aligned} \tag{112}$$

We now analyze the four possibilities for the sum involving $|\mathscr{S}|^2$. If $\mathscr{S}$ is either $\widehat{\phi}_i(0)\frac{\log(1+t_j^2)}{\log R}$ or $O\left(\frac{\log\log R}{\log R}\right)$, this sum is trivially $O(1)$, and thus the entire expression is $O\left(\frac{\log\log R}{\log R}\right)$. We are left with the non-trivial cases of $\mathscr{S} = S_1$ or $\mathscr{S} = S_2$. To handle these cases, we rely on results that we will soon prove in lemmas below: for sufficiently small support $|S_1|^2 = O(1)$ and $|S_2|^2 = o(1)$. Note that we use these lemmas for $\phi = \phi_1\phi_1$ instead of the usual $\phi = \phi_1\phi_2$, but this does not affect the proofs, and thus we may move the $O\left(\frac{\log\log R}{\log R}\right)$ factor outside the product.

By symmetry, it suffices to analyze the following terms to determine the 2-level density:

$$\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}\widehat{\phi}_1(0)\widehat{\phi}_2(0)\left(\frac{\log(1+t_j^2)}{\log R}\right)^2$$

$$\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}\widehat{\phi}_1(0)\frac{\log(1+t_j^2)}{\log R}S_k(u_j,\phi_2), \quad k\in\{1,2\}$$

$$\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}S_k(u_j,\phi_1)\overline{S_\ell(u_j,\phi_2)} \quad k,\ell\in\{1,2\}. \tag{113}$$

We now analyze these terms. For small support, a similar analysis as performed earlier in the paper shows that the first term in (113) satisfies

$$\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}\widehat{\phi}_1(0)\widehat{\phi}_2(0)\left(\frac{\log(1+t_j^2)}{\log R}\right)^2 = \widehat{\phi}_1(0)\widehat{\phi}_2(0) + O\left(\frac{1}{\log\log R}\right). \tag{114}$$

We next handle the terms where we have exactly one S-factor.

Lemma 7.1. *For sufficiently small support, the $\widehat{\phi}_1(0)S_1(u_j,\phi_2)$ and $\widehat{\phi}_1(0)S_2(u_j,\phi_2)$ terms are $O(\log\log R/\log R)$, and thus do not contribute.*

Proof. The proof is almost identical to the application of the Kuznetsov trace formula to prove similar results for the 1-level density, the only change being that now we have the modified weight function $h_T(t_j)\log(1+t_j^2)$.

We now turn to the more interesting terms in (113). We first handle the second term.

Lemma 7.2. *For sufficiently small support,*

$$\frac{1}{\sum_j \frac{h_T(t_j)}{\|u_j\|^2}}\sum_j \frac{h_T(t_j)}{\|u_j\|^2}S_1(u_j,\phi_1)\overline{S_1(u_j,\phi_2)} = 2\int_{-\infty}^{\infty}|z|\hat{\phi}_1(z)\hat{\phi}_2(z)dz + O\left(\frac{\log\log R}{\log R}\right). \tag{115}$$

Proof.

$$\sum_j \frac{h_T(t_j)}{\|u_j\|^2} S_1(u_j,\phi_1)\overline{S_1(u_j,\phi_2)}$$

$$= 4\sum_j \frac{h_T(t_j)}{\|u_j\|^2} \sum_{p_1,p_2} \frac{\lambda_{p_1}(u_j)\overline{\lambda_{p_2}(u_j)}}{p_1^{1/2}p_2^{1/2}} \frac{\log p_1 \log p_2}{\log^2 R} \hat{\phi}_1\left(\frac{\log p_1}{\log R}\right)\hat{\phi}_2\left(\frac{\log p_2}{\log R}\right)$$

$$= 4\sum_{p_1,p_2} \frac{1}{p_1^{1/2}p_2^{1/2}} \frac{\log p_1 \log p_2}{\log^2 R} \hat{\phi}_1\left(\frac{\log p_1}{\log R}\right)\hat{\phi}_2\left(\frac{\log p_2}{\log R}\right) \sum_j \frac{h_T(t_j)}{\|u_j\|^2} \lambda_{p_1}(u_j)\overline{\lambda_{p_2}(u_j)}. \tag{116}$$

As before, we apply the Kuznetsov formula to the inner sum. Since the formula has a δ_{p_1,p_2}, we need to split this sum into the case when $p_1 = p_2$ and the case when $p_1 \neq p_2$. For small support one easily finds the case $p_1 \neq p_2$ does not contribute; however, the case $p_1 = p_2$ does contribute. Arguing as in the 1-level calculations, after applying the Kuznetsov formula we are left with

$$4\sum_p \frac{1}{p}\frac{\log^2 p}{\log^2 R}\hat{\phi}_1\left(\frac{\log p}{\log R}\right)\hat{\phi}_2\left(\frac{\log p}{\log R}\right) = 2\int_{-\infty}^{\infty} |z|\hat{\phi}_1(z)\hat{\phi}_2(z)dz + O\left(\frac{\log\log R}{\log R}\right). \tag{117}$$

(the equality follows from partial summation and the Prime Number Theorem, see [42] for a proof).

Lemma 7.3. *For small support, the contribution from the* $S_k(u_j,\phi_1)\overline{S_\ell(u_j,\phi_2)}$ *terms is* $O(\log\log R/\log R)$ *if* $(k,\ell) = (1,2)$ *or* $(2,2)$.

Proof. The proof is similar to the previous lemma, following again by applications of the Kuznetsov trace formula. The support is slightly larger as the power of the primes in the denominator is larger.

Acknowledgements We thank Eduardo Dueñez, Gergely Harcos, Andrew Knightly, Stephen D. Miller, and Peter Sarnak for helpful conversations in an earlier version. This work was done at the SMALL REU at Williams College, funded by NSF GRANT DMS0850577 and Williams College; it is a pleasure to thank them for their support. The second named author was also partially supported by the Mathematics Department of University College London, and the fifth named author was partially supported by NSF grants DMS0970067 and DMS1265673.

References

1. M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th printing (Dover, New York, 1972)
2. L. Alpoge, S.J. Miller, *Low Lying Zeros of Maass Form L-Functions*. Int Math Res Notices (2014), 24 pages, doi:10.1093/imrn/rnu012
3. B. Birch, H.P.F. Swinnerton-Dyer, Notes on elliptic curves. I. J. Reine Angew. Math. **212**, 7–25 (1963)
4. B. Birch, H.P.F. Swinnerton-Dyer, Notes on elliptic curves. II. J. Reine Angew. Math. **218**, 79–108 (1965)
5. J.B. Conrey, *L*-Functions and random matrices, in *Mathematics Unlimited: 2001 and Beyond* (Springer, Berlin, 2001), pp. 331–352
6. J.B. Conrey, H. Iwaniec, Spacing of zeros of Hecke L-functions and the class number problem. Acta Arith. **103**(3), 259–312 (2002)
7. H. Davenport, *Multiplicative Number Theory (Graduate Texts in Mathematics)*, vol. 74, 2nd edn. (Springer, New York, 1980). Revised by H. Montgomery
8. E. Dueñez, S.J. Miller, The low lying zeros of a GL(4) and a GL(6) family of *L*-functions. Compos. Math. **142**(6), 1403–1425 (2006)
9. E. Dueñez, S.J. Miller, The effect of convolving families of *L*-functions on the underlying group symmetries. Proc. Lond. Math. Soc. (2009). doi:10.1112/plms/pdp018
10. H.M. Edwards, *Riemann's Zeta Function* (Academic, New York, 1974)
11. A. Entin, E. Roditty-Gershon, Z. Rudnick, Low-lying zeros of quadratic Dirichlet *L*-functions, hyper-elliptic curves and random matrix theory. Geom. Funct. Anal. **23**(4), 1230–1261 (2013)
12. P. Erdős, H. Maier, A. Sárközy, On the distribution of the number of prime factors of sums $a + b$. Trans. Am. Math. Soc. **302**(1), 269–280 (1987)
13. D. Fiorilli, S.J. Miller, Surpassing the ratios conjecture in the 1-level density of Dirichlet *L*-functions. Algebra & Number Theory **9**(1), 13–52 (2015)
14. F.W.K. Firk, S.J. Miller, Nuclei, primes and the random matrix connection. Symmetry **1**, 64–105 (2009). doi:10.3390/sym1010064
15. P. Forrester, *Log-Gases and Random Matrices*. London Mathematical Society Monographs, vol. 34 (Princeton University Press, Princeton, 2010)
16. E. Fouvry, H. Iwaniec, Low-lying zeros of dihedral *L*-functions. Duke Math. J. **116**(2), 189–217 (2003)
17. P. Gao, *N*-level density of the low-lying zeros of quadratic Dirichlet *L*-functions. Ph.D Thesis, University of Michigan, 2005
18. D. Goldfeld, The class number of quadratic fields and the conjectures of Birch and Swinnerton-Dyer. Ann. Scuola Norm. Sup. Pisa **3**(4), 623–663 (1976)
19. B. Gross, D. Zagier, Heegner points and derivatives of *L*-series. Invent. Math. **84**, 225–320 (1986)
20. A. Güloğlu, Low-lying zeros of symmetric power *L*-functions. Int. Math. Res. Not. **9**, 517–550 (2005)
21. B. Hayes, The spectrum of Riemannium. Am. Sci. **91**(4), 296–300 (2003)
22. D. Hejhal, On the triple correlation of zeros of the zeta function. Int. Math. Res. Not. **7**, 294–302 (1994)
23. C. Hughes, S.J. Miller, Low-lying zeros of *L*-functions with orthogonal symmtry. Duke Math. J. **136**(1), 115–172 (2007)
24. C. Hughes, Z. Rudnick, Linear statistics of low-lying zeros of *L*-functions. Q. J. Math. Oxf. **54**, 309–333 (2003)
25. H. Iwaniec, *Introduction to the Spectral Theory of Automorphic Forms* (Biblioteca de la Revista Matemática Iberoamericana, Madrid, 1995)
26. H. Iwaniec, E. Kowalski, *Analytic Number Theory*, vol. 53 (AMS Colloquium Publications, Providence, 2004)

27. H. Iwaniec, W. Luo, P. Sarnak, Low lying zeros of families of L-functions. Inst. Hautes Études Sci. Publ. Math. **91**, 55–131 (2000)
28. N. Katz, P. Sarnak, *Random Matrices, Frobenius Eigenvalues and Monodromy*, vol. 45 (AMS Colloquium Publications, Providence, 1999)
29. N. Katz, P. Sarnak, Zeros of zeta functions and symmetries. Bull. Am. Math. Soc. **36**, 1–26 (1999)
30. H. Kim, Functoriality for the exterior square of GL_2 and the symmetric fourth of GL_2. J. Am. Math. Soc. **16**(1), 139–183 (2003)
31. H. Kim, P. Sarnak, Appendix: Refined estimates towards the Ramanujan and Selberg conjectures, Appendix to "H. Kim, Functoriality for the exterior square of GL_2 and the symmetric fourth of GL_2. J. Am. Math. Soc. **16**(1), 139–183 (2003)"
32. J. Levinson, S.J. Miller, The n-level density of zeros of quadratic Dirichlet L-functions. Acta Arith. **161**, 145–182 (2013)
33. C. Li, A. Knightly, Kuznetsov's trace formula and the Hecke eigenvalues of Maass forms. Mem. Am. Math. Soc. **24**(1055), 132 (fourth of 4 numbers) (2013)
34. J. Liu, *Lectures on Maass Forms* (Postech, 2007), http://www.prime.sdu.edu.cn/lectures/LiuMaassforms.pdf. Accessed 25–27 March 2007
35. J. Liu, Y. Ye, Petersson and Kuznetsov trace formulas, in *Lie Groups and Automorphic Forms*, ed. by L. Ji, J.-S. Li, H.W. Xu, S.-T. Yau, AMS/IP Studies in Advanced Mathematics, vol. 37 (American Mathematical Society, Providence, 2006), pp. 147–168
36. H. Maier, *On Exponential Sums with Multiplicative Coefficients*, in Analytic Number Theory (Cambridge University Press, Cambridge 2009), pp. 315–323
37. H. Maier, C. Pomerance, Unusually large gaps between consecutive primes. Trans. Am. Math. Soc. **322**(1), 201–237 (1990)
38. H. Maier, A. Sankaranarayanan, On an exponential sum involving the Mbius function. Hardy-Ramanujan J. **28**, 10–29 (2005)
39. H. Maier, A. Sankaranarayanan, Exponential sums over primes in residue classes. Int. J. Number Theory **6**(4), 905–918 (2010)
40. H. Maier, G. Tenenbaum, On the set of divisors of an integer. Invent. Math. **76**(1), 121–128 (1984)
41. M. Mehta, *Random Matrices*, 2nd edn. (Academic, Boston, 1991)
42. S.J. Miller, 1- and 2-level densities for families of elliptic curves: evidence for the underlying group symmetries. Ph.D. Thesis, Princeton University, 2002
43. S.J. Miller, 1- and 2-level densities for families of elliptic curves: evidence for the underlying group symmetries. Compos. Math. **140**, 952–992 (2004)
44. S.J. Miller, D. Montague, An orthogonal test of the L-functions ratios conjecture, II. Acta Arith. **146**, 53–90 (2011)
45. S.J. Miller, R. Peckner, Low-lying zeros of number field L-functions. J. Number Theory **132**, 2866–2891 (2012)
46. S.J. Miller, R. Takloo-Bighash, *An Invitation to Modern Number Theory* (Princeton University Press, Princeton, 2006)
47. H. Montgomery, The pair correlation of zeros of the zeta function, in *Analytic Number Theory*. Proceedings of Symposia in Pure Mathematics, vol. 24 (American Mathematical Society, Providence, 1973), pp. 181–193
48. A. Odlyzko, On the distribution of spacings between zeros of the zeta function. Math. Comput. **48**(177), 273–308 (1987)
49. A. Odlyzko, The 10^{22}-nd zero of the Riemann zeta function, in *Proceedings of the Conference on Dynamical, Spectral and Arithmetic Zeta-Functions*. American Mathematical Society, Contemporary Mathematics series, 2001, ed. by M. van Frankenhuysen, M.L. Lapidus. http://www.research.att.com/$\sim$amo/doc/zeta.html
50. A.E. Özlük, C. Snyder, Small zeros of quadratic L-functions. Bull. Aust. Math. Soc. **47**(2), 307–319 (1993)
51. G. Ricotta, E. Royer, Statistics for low-lying zeros of symmetric power L-functions in the level aspect. Forum Math. **23**, 969–1028 (2011, preprint)

52. G.F.B. Riemann, Über die Anzahl der Primzahlen unter einer gegebenen Grösse, Monatsber. Königl. Preuss. Akad. Wiss. Berlin, Nov. 671–680 (1859) (see "H.M. Edwards, *Riemann's Zeta Function* (Academic, New York, 1974)" for an English translation)
53. E. Royer, Petits zéros de fonctions L de formes modulaires. Acta Arith. **99**(2), 147–172 (2001)
54. M. Rubinstein, Low-lying zeros of L-functions and random matrix theory. Duke Math. J. **109**, 147–181 (2001)
55. M. Rubinstein, P. Sarnak, Chebyshev's bias. Exp. Math. **3**(3), 173–197 (1994)
56. Z. Rudnick, P. Sarnak, Zeros of principal L-functions and random matrix theory. Duke Math. J. **81**, 269–322 (1996)
57. S.W. Shin, N. Templier, Sato-Tate theorem for families and low-lying zeros of automorphic L-functions. Invent. Math. preprint. http://arxiv.org/pdf/1208.1945v2
58. R.A. Smith, The L^2-norm of Maass wave functions. Proc. Am. Math. Soc. **82**(2), 179–182 (1981)
59. A. Yang, *Low-Lying Zeros of Dedekind Zeta Functions Attached to Cubic Number Fields*, preprint
60. M. Young, Low-lying zeros of families of elliptic curves. J. Am. Math. Soc. **19**(1), 205–250 (2006)

Théorème de Jordan Friable

Régis de la Bretèche et Gérald Tenenbaum

Pour Helmut Maier, qui compte
par plaisir et partage sans compter.

Abstract Extending a previous result, we show that, for the friable summation method, the Fourier series of any normalized function F with bounded variation on the unidimensional torus converges pointwise to F while avoiding the Gibbs phenomenon. We also prove that the convergence is uniform when F is continuous and provide an effective bound for the rate when F satisfies a uniform Lipschitz condition.

Keywords Friable integers • Friable summation • Summation methods • Fourier series • Gibbs phenomenon • Functions of bounded variation • Jordan's theorem • Integers free of large prime factors

2010 Mathematics Subject Classification.
Primary: 11N25, 42A24, secondary: 42A20.

1 Introduction et énoncé des résultats

Soit $VB^*(\mathbb{T})$ la classe des fonctions F qui sont 1-périodiques, à variation bornée sur $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, et normalisées par $F(\vartheta) = \frac{1}{2}\{F(\vartheta+) + F(\vartheta-)\}$ aux points de discontinuité. Nous nous proposons ici de généraliser aux fonctions de $VB^*(\mathbb{T})$ le théorème 5.1 de [2] relatif à la convergence friable des séries de Fourier et

R. de la Bretèche
Institut de Mathématiques de Jussieu-Paris Rive Gauche, UMR 7586, Université Paris Diderot-Paris 7, case 7012, Bâtiment Sophie Germain, 75205 Paris Cedex 13, France
e-mail: regis.de-la-breteche@imj-prg.fr

G. Tenenbaum (✉)
Institut Élie Cartan, Université de Lorraine, BP 70239,
54506 Vandœuvre-lès-Nancy Cedex, France
e-mail: gerald.tenenbaum@univ-lorraine.fr

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_3

à l'absence de phénomène de Gibbs (voir par exemple, [9], vol. I, § II.9, ou [6], chapitre 17) pour ce procédé de sommation.

Désignons par $P(n)$ le plus grand facteur premier d'un entier naturel $n \geqslant 1$, avec la convention que $P(1) = 1$. Nous notons $a_n(F)$, $b_n(F)$ les coefficients de Fourier d'une fonction F de $L^1(\mathbb{T})$, et, pour toute $y \geqslant 2$, nous définissons la suite des sommes partielles friables de la série de Fourier de F par

$$F(\vartheta; y) := a_0(F) + \sum_{P(n)\leqslant y} \big\{a_n(F)\cos(2\pi n\vartheta) + b_n(F)\sin(2\pi n\vartheta)\big\} \tag{1}$$

pour toute valeur de ϑ où cela possède un sens. C'est en particulier le cas si $F \in VB^*(\mathbb{T})$ puisque $|a_n(F)| + |b_n(F)| \ll 1/n$, ce qui implique l'absolue convergence de la série (1) pour chaque $y \geqslant 2$.

L'étude de la convergence de $F(\vartheta; y)$ lorsque y tend vers l'infini s'inscrit dans le cadre plus général de celle de la convergence friable des séries. Formellement introduit dans [3] puis dans [4], ce procédé de sommation[1] consistant à définir la somme comme limite, lorsque le paramètre de friabilité tend vers l'infini, des sous-séries restreintes aux entiers friables, a été étudié dans [1] puis de manière systématique dans [2]. Au théorème 5.1 de [2], nous avons établi la convergence simple de $F(\vartheta; y)$ vers $F(\vartheta)$ pour toute fonction F de $VB^*(\mathbb{T})$ sous réserve que la dérivée presque partout de F soit dans $\cup_{\alpha>1} L^\alpha(\mathbb{T})$. Le but de la présente note consiste à relâcher cette dernière condition.

Il est relativement facile d'exhiber des exemples de fonctions discontinues F appartenant à $VB^*(\mathbb{T}) \smallsetminus \cup_{\alpha>1} L^\alpha(\mathbb{T})$: si $f \in L^1(\mathbb{T}) \smallsetminus \cup_{\alpha>1} L^\alpha(\mathbb{T})$, on peut choisir

$$F(\vartheta) := \int_0^\vartheta f(t)\,\mathrm{d}t + B(\vartheta).$$

Un exemple de fonction f admissible est fourni par

$$f(\vartheta) = \sum_{n\geqslant 2} \frac{\cos(2\pi n\vartheta)}{\log n}. \tag{2}$$

En effet, un théorème classique concernant les séries trigonométriques $\sum_{n\in\mathbb{Z}} a_n\,\mathrm{e}^{2\pi i n\vartheta}$ dont la suite des coefficients est paire et vérifie identiquement $a_{n-1}+a_{n+1} \geqslant 2a_n$ (voir [5], théorème I.4.1, ou [9], vol. I, théorème V.1.5) implique[2] $f \in L^1(\mathbb{T})$ alors que, pour tout $\alpha > 1$, posant $r := \min(2, \alpha)$, $s := r/(r-1)$, l'hypothèse $f \in L^\alpha(\mathbb{T})$ impliquerait $f \in L^r(\mathbb{T})$ et donc (cf., par exemple, [5], théorème I.4.7)

[1] Désigné dans ces travaux sous le nom de P-convergence ou P-sommabilité.

[2] On peut aussi établir directement via la formule d'Euler–Maclaurin que, pour $0 < |\vartheta| \leqslant \frac{1}{2}$, on a $f(\vartheta) \ll 1/\{\vartheta(\log\vartheta)^2\}$.

$$\sum_{n\in\mathbb{Z}} |\hat{f}(n)|^s < \infty,$$

une condition manifestement en défaut pour la série (2).

Nous désignons par ϱ la fonction de Dickman,[3] posons $\|v\| := \min_{n\in\mathbb{Z}} |v-n|$ $(v \in \mathbb{R})$ et notons

$$B(\vartheta; y) = -\sum_{P(n)\leqslant y} \frac{\sin(2\pi n\vartheta)}{\pi n} \tag{3}$$

la somme partielle friable d'ordre y de la série de Fourier de la fonction de Bernoulli

$$B(\vartheta) = \sum_{m\geqslant 1} \frac{\sin(2\pi m\vartheta)}{\pi m} = \begin{cases} \langle\vartheta\rangle - \frac{1}{2} & \text{si } \vartheta \notin \mathbb{Z}, \\ 0 & \text{si } \vartheta \in \mathbb{Z}, \end{cases} \tag{4}$$

où $\langle\vartheta\rangle$ désigne la partie fractionnaire du nombre réel ϑ.

Théorème 1.1. *Soit $F \in VB^*(\mathbb{T})$. On a, uniformément pour $\vartheta \in \mathbb{R}$ et $y \geqslant 2$,*

$$F(\vartheta; y) - F(\vartheta) = \int_{\mathbb{T}} \left\{ \varrho\left(\frac{\log(1/\|v\|)}{\log y}\right) - 1 \right\} B(v)\, \mathrm{d}F(\vartheta - v) + O\Big(\frac{1}{\log y}\Big). \tag{5}$$

En particulier, on a

$$\lim_{y\to\infty} F(\vartheta; y) = F(\vartheta) \qquad (\vartheta \in \mathbb{T}), \tag{6}$$

$$\lim_{y\to\infty} \sup_{\vartheta\in\mathbb{T}} F(\vartheta; y) = \sup_{\vartheta\in\mathbb{T}} F(\vartheta). \tag{7}$$

De plus, si F est continue, alors $F(\cdot; y)$ tend uniformément vers F.

Remarque. Bien que l'on ait (7), la convergence de $F(\vartheta; y)$ vers $F(\vartheta)$ n'est en général pas uniforme: il est facile de voir que la série $F(\vartheta; y)$ converge uniformément et donc que, pour chaque $y \geqslant 2$ fixé, $F(\vartheta; y)$ dépend continûment de ϑ. Il est cependant à noter que, contrairement à la situation classique, la convergence demeure exploitable au voisinage des discontinuités : alors que l'on sait, classiquement, que la somme partielle d'ordre y de la série (4) n'approche pas le membre de gauche pour, par exemple, $\vartheta = 1/y$, on déduit de (5) que

$$B(1/y; y) - B(1/y) \ll 1/\log y \qquad (y \geqslant 2).$$

[3] Voir par exemple [7], chapitre III.5. Rappelons que ϱ est la solution continue de l'équation différentielle aux différences $u\varrho'(u) + \varrho(u-1) = 0$ avec la condition initiale $\varrho(u) = 1$ $(0 \leqslant u \leqslant 1)$.

Sous une hypothèse supplémentaire de régularité pour F, nous pouvons préciser la vitesse de convergence de $F(\cdot;y)$ vers F.

Théorème 1.2. *Soit $\alpha > 0$. Si $F \in VB^*(\mathbb{T})$ est uniformément lipschitzienne d'exposant $\alpha > 0$, on a*

$$\max_{\vartheta\in\mathbb{T}} |F(\vartheta;y) - F(\vartheta)| \ll \frac{1}{y^{\alpha/2}} \qquad (y \geqslant 2). \tag{8}$$

2 Preuve du Théorème 1.1

Pour toute fonction F de $VB^*(\mathbb{T})$ et tout entier $n \geqslant 1$, on a

$$\begin{aligned} a_n(F)\cos(2\pi n\vartheta) + b_n(F)\sin(2\pi n\vartheta) &= 2\int_{\mathbb{T}} \cos(2\pi n(\vartheta - v))F(v)\,\mathrm{d}v \\ &= -\int_{\mathbb{T}} \frac{\sin(2\pi n(\vartheta - v))}{\pi n}\,\mathrm{d}F(v) = -\int_{\mathbb{T}} \frac{\sin(2\pi nv)}{\pi n}\,\mathrm{d}F(\vartheta - v). \end{aligned}$$

Cette quantité étant trivialement $\ll 1/n$, on peut sommer pour $P(n) \leqslant y$. On en déduit que l'on a, pour tout $y \geqslant 2$,

$$F(\vartheta;y) = a_0(F) + \int_{\mathbb{T}} B(v;y)\,\mathrm{d}F(\vartheta - v). \tag{9}$$

Par ailleurs, une simple intégration par parties fournit, dès que $F \in VB^*(\mathbb{T})$,

$$F(\vartheta) = a_0(F) + \int_{\mathbb{T}} B(v)\,\mathrm{d}F(\vartheta - v). \tag{10}$$

En effectuant la différence de (9) et (10), nous obtenons

$$F(\vartheta;y) - F(\vartheta) = \int_{\mathbb{T}} \nabla_1(v;y)\,\mathrm{d}F(\vartheta - v), \tag{11}$$

où l'on a posé

$$\nabla_1(\vartheta;y) := B(\vartheta;y) - B(\vartheta). \tag{12}$$

Or, nous avons établi dans [2] que l'on a, uniformément pour $\vartheta \in \mathbb{R}$,

$$\nabla_1(\vartheta;y) = \{\varrho(u^*_{\vartheta,y}) - 1\}B(\vartheta) + O\Big(\frac{1}{\log y}\Big), \tag{13}$$

où nous avons posé

$$u^*_{\vartheta,y} := \frac{\log(1/\|\vartheta\|)}{\log y} \qquad (\vartheta \in \mathbb{R} \smallsetminus \mathbb{Z})$$

et convenu que $u^*_{\vartheta,y} = \infty$ si $\vartheta \in \mathbb{Z}$. Il s'ensuit que

$$F(\vartheta; y) - F(\vartheta) = \int_{\mathbb{T}} \{\varrho(u^*_{v,y}) - 1\} B(v)\, \mathrm{d}F(\vartheta - v) + O_F(1/\log y). \tag{14}$$

Comme $\{\varrho(u^*_{v,y}) - 1\}B(v)$ tend simplement vers 0 sur $\mathbb{T}$, la relation (5) implique clairement (6), en vertu du théorème de la convergence dominée.

Montrons maintenant (7). Il résulte immédiatement de (6) que

$$\limsup_{y\to\infty} \sup_{\vartheta} F(\vartheta; y) \geqslant \sup_{\vartheta} F(\vartheta).$$

Il suffit donc de prouver l'inégalité inverse. Nous pouvons supposer sans perte de généralité que $a_0(F) = 0$. Posons

$$F_y(\vartheta) := \int_{\mathbb{T}} \varrho(u^*_{v,y}) B(v)\, \mathrm{d}F(\vartheta - v),$$

de sorte que, par (10) et (14), on a

$$\sup_{\vartheta} |F(\vartheta; y) - F_y(\vartheta)| \ll 1/\log y \qquad (y \to \infty).$$

Comme l'application $v \mapsto \varrho(u^*_{v,y})B(v)$ est continue sur $\mathbb{T}$, $F_y(\vartheta)$ est, pour chaque y, une fonction continue de ϑ. Elle atteint donc son maximum en un point ϑ_y. Quitte à extraire une sous-suite, nous pouvons supposer que

$$\limsup_{y\to\infty} \sup_{\vartheta} F(\vartheta; y) = \lim_{y\to\infty} F_y(\vartheta_y)$$

et que ϑ_y tend vers une limite $\vartheta_0 \in \mathbb{T}$. Quitte à changer $F(\vartheta)$ en $F(\vartheta - \vartheta_0)$, nous pouvons encore effectuer l'hypothèse que $\vartheta_0 = 0$.

Soit $\delta_F := F(0+) - F(0-)$ le saut de F en 0. Posons

$$H(\vartheta) := 1 + \lfloor \vartheta \rfloor - \tfrac{1}{2}\mathbf{1}_{\mathbb{Z}}(\vartheta)$$

de sorte que $\widetilde{F} := F - \delta_F H$ est continue sur $\mathbb{Z}$. Pour $\vartheta \in \mathbb{T}$, nous avons

$$F_y(\vartheta) = \int_{\mathbb{T}} \varrho(u^*_{v,y}) B(v)\, \mathrm{d}\widetilde{F}(\vartheta - v) + \delta_F \varrho(u^*_{\vartheta,y}) B(\vartheta). \tag{15}$$

Comme $\varrho(u_{v,y}^*)B(v)$ tend simplement vers $B(v)$ sur $\mathbb{T}$ avec une convergence uniforme sur tout compact de $\mathbb{T}$ ne contenant pas 0, et comme $\widetilde{F}$ est continue en 0, on a

$$\lim_{y\to\infty}\int_{\mathbb{T}}\varrho(u_{v,y}^*)B(v)\,\mathrm{d}\widetilde{F}(\vartheta_y-v)=\int_{\mathbb{T}}B(v)\,\mathrm{d}\widetilde{F}(-v).$$

Il suit

$$F_y(\vartheta_y)\leqslant\int_{\mathbb{T}}B(v)\,\mathrm{d}\widetilde{F}(-v)+\tfrac12|\delta_F|+o(1).$$

Or, en faisant tendre y vers l'infini dans (15), nous obtenons successivement

$$F(0\pm)=\int_{\mathbb{T}}B(v)\,\mathrm{d}\widetilde{F}(-v)\mp\tfrac12\delta_F,\qquad \max\{F(0+),F(0-)\}=\int_{\mathbb{T}}B(v)\,\mathrm{d}\widetilde{F}(-v)+\tfrac12|\delta_F|.$$

On en déduit que

$$F_y(\vartheta_y)\leqslant\max\{F(0+),F(0-)\}+o(1)\leqslant\sup_{\vartheta}F(\vartheta)+o(1).$$

Cela achève la preuve de (7).

Lorsque F est continue, et puisque $\varrho(v)=1$ pour $0\leqslant v\leqslant 1$, la relation (5) implique

$$|F(\vartheta;y)-F(\vartheta)|\leqslant\int_{\vartheta-1/y}^{\vartheta+1/y}|\,\mathrm{d}F|(t)+O\Big(\frac{1}{\log y}\Big),$$

où la constante implicite ne dépend que de F. Comme F est continue, et donc uniformément continue en vertu de la compacité du tore, il en va de même de sa variation. La dernière intégrale tend donc vers 0 uniformément en ϑ.

3 Preuve du Théorème 1.2

Sous l'hypothèse que F est à variation bornée sur $\mathbb{T}$ et satisfait une condition de Lipschitz uniforme

$$F(\vartheta+h)-F(\vartheta)\ll|h|^\alpha,\tag{16}$$

où $\alpha>0$ et la constante implicite est indépendante de $\vartheta\in\mathbb{T}$, un théorème de Zygmund ([9], vol. I, th. VI.3.6) implique que la série de Fourier de F, i.e.

$$F(\vartheta)=\sum_{n\in\mathbb{Z}}c_n(F)\,\mathrm{e}(n\vartheta),\tag{17}$$

avec la notation familière $\mathrm{e}(t):=\mathrm{e}^{2\pi it}$ $(t\in\mathbb{R})$, est absolument convergente.

En fait, la démonstration de Zygmund fournit, sous la seule hypothèse que F est à variation bornée, la majoration

$$\sum_{2^{\nu-1}<|n|\leqslant 2^{\nu}} |c_n(F)| \leqslant \tfrac{1}{2}\sqrt{V_F\omega_F(1/2^{\nu+1})} \qquad (\nu \in \mathbb{N}^*), \tag{18}$$

où V_F désigne la variation totale de F sur $\mathbb{T}$ et $h \mapsto \omega_F(h)$ son module de continuité.

Pour la commodité du lecteur, nous reproduisons ici la courte preuve de [9]. Soient $\nu \geqslant 1$, $N := 2^{\nu+1}$. Pour tout $\vartheta \in \mathbb{T}$, nous avons

$$\sum_{1\leqslant k\leqslant N} \Big|F\Big(\vartheta + \frac{k}{N}\Big) - F\Big(\vartheta + \frac{k-1}{N}\Big)\Big|^2 \leqslant V_F\omega_F\Big(\frac{1}{N}\Big).$$

En intégrant sur $\mathbb{T}$, nous obtenons

$$4N \sum_{n\in\mathbb{Z}} |c_n(F)|^2 \sin^2\Big(\frac{\pi n}{N}\Big) \leqslant V_F\omega_F\Big(\frac{1}{N}\Big).$$

Nous en déduisons (18) en observant que, pour $N/4 < |n| \leqslant N/2$, nous avons

$$\sqrt{2}/2 \leqslant |\sin(\pi n/N)| \leqslant 1$$

et en appliquant l'inégalité de Cauchy–Schwarz.

Sous la condition (16), nous déduisons immédiatement de (18) que

$$\sum_{|n|>y} |c_n(F)| \ll_F \frac{1}{y^{\alpha/2}} \qquad (y \geqslant 1). \tag{19}$$

Posons

$$F_N(\vartheta) := \sum_{|n|\leqslant N} c_n(F)\,\mathrm{e}(n\vartheta), \qquad (N \geqslant 1,\ \vartheta \in \mathbb{T},\ y \geqslant 2).$$
$$F_N(\vartheta; y) := \sum_{\substack{|n|\leqslant N\\ P(|n|)\leqslant y}} c_n(F)\,\mathrm{e}(n\vartheta)$$

Pour tous $\vartheta \in \mathbb{T}$, $N \geqslant 1$, nous avons, par (11),

$$F_N(\vartheta; y) - F_N(\vartheta) = -\sum_{\substack{|n|\leqslant N\\ P(|n|)>y}} c_n(F)\,\mathrm{e}(n\vartheta) \ll \frac{1}{y^{\alpha/2}} \tag{20}$$

où la majoration découle de (19), la constante implicite étant donc indépendante de N.

Comme $F \in VB^*(\mathbb{T})$, on a $c_n(F) \ll 1/n$ pour tout $n \in \mathbb{Z}^*$, donc $F_N(\vartheta; y)$ tend simplement vers $F(\vartheta; y)$ lorsque N tend vers l'infini. D'après le théorème de Jordan, il en va de même de la convergence de $F_N(\vartheta)$ vers $F(\vartheta)$. Nous obtenons donc (8) en faisant tendre N vers l'infini dans (20).

Bibliographie

1. R. de la Bretèche, *P*-régularité de sommes d'exponentielles. Mathematika **45**, 145–175 (1998)
2. R. de la Bretèche, G. Tenenbaum, Séries trigonométriques à coefficients arithmétiques. J. Anal. Math. **92**, 1–79 (2004)
3. R.J. Duffin, Representation of Fourier integrals as sums, III. Proc. Am. Math. Soc. **8**, 272–277 (1957)
4. É. Fouvry, G. Tenenbaum, Entiers sans grand facteur premier en progressions arithmétiques. Proc. Lond. Math. Soc. **63**(3), 449–494 (1991)
5. Y. Katznelson, *An Introduction to Harmonic Analysis*, 2ème édn. (Dover, New York, 1976)
6. T.W. Körner, *Fourier Analysis*, 2nd edn. (Cambridge University Press, Cambridge, 1989), xii+591 pp.
7. G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, 3ème édn. (coll. Échelles, Belin, 2008), 592 pp.
8. G. Tenenbaum, J. Wu, *Exercices corriges de théorie analytique et probabiliste des nombres* (coll. Échelles, Belin, 2014), 347 pp.
9. A. Zygmund, *Trigonometric Series*. Cambridge Mathematical Library, vol. I, II, 3ème édn. (Cambridge University Press, Cambridge, 2002), vol. I: xiv+383 pp.; vol. II: viii+364 pp.

On Conjectures of T. Ordowski and Z.W. Sun Concerning Primes and Quadratic Forms

Christian Elsholtz and Glyn Harman

On the occasion of Helmut Maier's 60th birthday. With admiration for his beautiful results on the distribution of primes

Abstract We discuss recent conjectures of T. Ordowski and Z.W. Sun on limits of certain coordinate-wise defined functions of primes in $\mathbb{Q}(\sqrt{-1})$ and $\mathbb{Q}(\sqrt{-3})$. Let $p \equiv 1 \bmod 4$ be a prime and let $p = a_p^2 + b_p^2$ be the unique representation with positive integers $a_p > b_p$. Then the following holds:

$$\lim_{N\to\infty} \frac{\sum_{p\le N, p\equiv 1 \bmod 4} a_p^k}{\sum_{p\le N, p\equiv 1 \bmod 4} b_p^k} = \frac{\int_0^{\pi/4} \cos^k(x)\,dx}{\int_0^{\pi/4} \sin^k(x)\,dx}.$$

For $k = 1$ this proves, but for $k = 2$ this disproves the conjectures in question. We shall also generalise the result to cover all positive definite, primitive, binary quadratic forms. In addition we will discuss the case of indefinite forms and prove a result that covers many cases in this instance.

2010 Mathematics Subject Classification:
Primary:
11N05 Distribution of primes
Secondary:
11A41 primes
11R44 Distribution of prime ideals
11E25 Sums of squares and representations by other particular quadratic forms

C. Elsholtz (✉)
Institut für Mathematik A, Technische Universität Graz, Steyrergasse 30/II, 8010 Graz, Austria
e-mail: elsholtz@math.tugraz.at

G. Harman
Department of Mathematics, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK
e-mail: g.harman@rhul.ac.uk

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_4

1 Primes and Quadratic Forms

Tomasz Ordowski (see Sun [12], Sect. 6) conjectured:

Conjecture 1.1. Let $p \equiv 1 \bmod 4$ be a prime and let $p = a_p^2 + b_p^2$ be the unique representation with positive integers $a_p > b_p$.

$$\text{(a)} \qquad \lim_{N\to\infty} \frac{\sum_{p\le N,\, p\equiv 1 \bmod 4} a_p}{\sum_{p\le N,\, p\equiv 1 \bmod 4} b_p} = 1+\sqrt{2},$$

$$\text{(b)} \qquad \lim_{N\to\infty} \frac{\sum_{p\le N,\, p\equiv 1 \bmod 4} a_p^2}{\sum_{p\le N,\, p\equiv 1 \bmod 4} b_p^2} = \frac{9}{2}.$$

Z.W. Sun [12] stated a number of related conjectures for other quadratic forms, including the following:

Conjecture 1.2 (Sun). Let $p \equiv 1 \bmod 3$ be a prime and let $p = x_p^2 + x_p y_p + y_p^2$ be the unique representation with positive integers $x_p > y_p$.

$$\text{(a)} \qquad \lim_{N\to\infty} \frac{\sum_{p\le N,\, p\equiv 1 \bmod 3} x_p}{\sum_{p\le N,\, p\equiv 1 \bmod 3} y_p} = 1+\sqrt{3},$$

$$\text{(b)} \qquad \lim_{N\to\infty} \frac{\sum_{p\le N,\, p\equiv 1 \bmod 3} x_p^2}{\sum_{p\le N,\, p\equiv 1 \bmod 1} y_p^2} = \frac{52}{9}.$$

This conjecture can also be found in the comments on sequence A218585 in [10]. Sun further remarked that numerically

$$\lim_{N\to\infty} \frac{\sum_{p\le N,\, p\equiv 1 \bmod 3} x_p^3}{\sum_{p\le N,\, p\equiv 1 \bmod 1} y_p^3} \approx 11.15 \quad \text{and} \quad \lim_{N\to\infty} \frac{\sum_{p\le N,\, p\equiv 1 \bmod 3} x_p^4}{\sum_{p\le N,\, p\equiv 1 \bmod 1} y_p^4} \approx 20.6.$$

In this paper we will show that these limits exist and how one can evaluate them. Cases (a) of both conjectures turn out to be true, but cases (b) of both conjectures are wrong. It seems to us that the authors guessed the values based on some experimental data, but without theoretical justification. For the two quadratic forms above we will determine the nature and value of the sums for k-th moments. Throughout the text k denotes a positive integer.

Theorem 1.1. *Let $p \equiv 1 \bmod 4$ be a prime and let $p = a_p^2 + b_p^2$ be the unique representation with positive integers $a_p > b_p$. Then*

$$I_k := \lim_{N\to\infty} \frac{\sum_{p\le N, p\equiv 1 \bmod 4} a_p^k}{\sum_{p\le N, p\equiv 1 \bmod 4} b_p^k} = \frac{\int_0^{\pi/4} \cos^k(x)\,dx}{\int_0^{\pi/4} \sin^k(x)\,dx}.$$

For $k = 1$ this value is indeed $1+\sqrt{2}$. For $k = 2$ the value is $\frac{\pi+2}{\pi-2}$ which is about 4.50388. Here is a small table of exact and numerical values.

k	Exact value	Approx. numerical value
1	$1+\sqrt{2}$	2.41421
2	$\dfrac{\pi+2}{\pi-2}$	4.50388
3	$\frac{5}{7}(5+4\sqrt{2})$	7.61204
4	$\dfrac{3\pi+8}{3\pi-8}$	12.2298
5	$(43\sqrt{2})/(64-43\sqrt{2}) = \frac{43}{199}(43+32\sqrt{2})$	19.0701
6	$\dfrac{15\pi+44}{15\pi-44}$	29.1700
7	$(177\sqrt{2})/(256-177\sqrt{2}) = \frac{177}{1439}(177+128\sqrt{2})$	44.0371
8	$\dfrac{21\pi+64}{21\pi-64}$	65.8612

Remark 1.1. The method of proof can be easily adapted to prove a slightly more general result: Let $0 \le C_1 < C_2$ be nonnegative constants. Let $p \equiv 1 \bmod 4$ be a prime and let $p = a_p^2 + b_p^2$ be a representation with positive integers $C_1\, a_p < b_p < C_2\, a_p$. Then (counting representations with multiplicity if $C_1 < 1 < C_2$)

$$I_k(C_1, C_2) := \lim_{N\to\infty} \frac{\sum_{p\le N, p\equiv 1 \bmod 4} a_p^k}{\sum_{p\le N, p\equiv 1 \bmod 4} b_p^k} = \frac{\int_{\arctan C_1}^{\arctan C_2} \cos^k(x)\, dx}{\int_{\arctan C_1}^{\arctan C_2} \sin^k(x)\, dx}.$$

The special case $C_1 = 0, C_2 = 1$ immediately gives Theorem 1.1.

Remark 1.2. It is possible to investigate the number theoretic properties of the integers in the related sequences. For example, the coefficients of π in the expressions

$$\int_0^{\pi/4} \cos^k(x)\, dx = \begin{cases} \frac{2+\pi}{8} & k=2 \\ \frac{8+3\pi}{32} & k=4 \\ \frac{11}{48}+\frac{10\pi}{128} & k=6 \\ \frac{5}{24}+\frac{35\pi}{512} & k=8 \end{cases}$$

etc. lead to the sequence $1, 3, 10, 35, 126, \ldots$ which is well studied, with a lot of further comments and connections stated in the Online encyclopedia of integer sequences, sequence, A001700 [10].

On the other hand, for $k = 2\ell$ we find that $I_k = (A_\ell \pi + B_\ell)/(A_\ell \pi - B_\ell)$ where A_ℓ is the sequence

$$1, 3, 15, 21, 315, 3465, 45045, 15015, 765765, 14549535 \ldots$$

which is, at the time of writing, not in the OEIS. However, the sequence is very closely related to sequence A025547, least common multiple of $\{1, 3, 5, \ldots, 2n-1\}$: $1, 3, 15, 105, 315, 3465, 45045, 45045, 765765, 14549535, 14549535, \ldots$. The relation to this sequence is quite natural by the recursive nature of the integrals.

(Expanding the fractions would enlarge 21 to 105 and 15015 to 45045.) We do not follow any of these paths further.

We now study two properties of the values of I_k:

Corollary 1.1. *Let* $I_k = \lim_{N\to\infty} \frac{\sum_{p\le N, p\equiv 1 \bmod 4} a_p^k}{\sum_{p\le N, p\equiv 1 \bmod 4} b_p^k}$, *then* $I_k \in \mathbb{Q}(\sqrt{2})$, *with* I_k *irrational, when* k *is odd, and* $I_k \in \mathbb{Q}(\pi)$, *with* I_k *transcendental, when* $k \ge 2$ *is even.*

The following result states the asymptotic growth of I_k:

Theorem 1.2. *As* k *tends to infinity, we have the following estimate:*

$$I_k \sim \left(\frac{\pi k}{2}\right)^{\frac{1}{2}} 2^{k/2}. \tag{1}$$

We now come to Conjecture 1.2. A consequence of our main theorem below (Theorem 1.4) is the following which we state as a result in its own right.

Theorem 1.3. *Let* $p \equiv 1 \bmod 3$ *be a prime and let* $p = x_p^2 + x_p y_p + y_p^2$ *be the unique representation with positive integers* $x_p > y_p$. *Then*

$$J_k = \lim_{N\to\infty} \frac{\sum_{p\le N, p\equiv 1 \bmod 3} x_p^k}{\sum_{p\le N, p\equiv 1 \bmod 3} y_p^k} = \frac{U(k)}{V(k)},$$

where

$$U(k) = \int_0^{\pi/6} (\sqrt{3}\cos x - \sin x)^k \, dx$$

and

$$V(k) = 2^k \int_0^{\pi/6} \sin^k x \, dx.$$

In particular this gives

k	Exact value of J_k	Approx. numerical value
1	$1+\sqrt{3}$	2.73205
2	$\frac{2\pi}{2\pi - 3\sqrt{3}}$	5.78012
3	$\frac{1}{13}(67+45\sqrt{3})$	11.1494
4	$\frac{2(\sqrt{3}-2\pi)}{7\sqrt{3}-4\pi}$	20.5927
5	$\frac{7}{709}(1837+1113\sqrt{3})$	37.1698
6	$\frac{9\sqrt{3}-10\pi}{18\sqrt{3}-10\pi}$	66.2204

Again, part (a) of the conjecture is correct, while (b) is not.

Theorems 1.1 and 1.3 above, and many others, are consequences of the following general result. Before stating this we need to make some comments on the uniqueness of representations. In the examples above the representations are unique, but it is easy to find forms which give two representations. For Theorem 1.3 if we had considered instead $Q(x,y) = x^2 - xy + y^2$ so that $Q(x,y) = Q(x,x-y)$ and thus $0 < y < x$ is equivalent to $0 < x - y < x$. We thus obtain two representations for every representable prime. Similarly, if instead of $x^2 + y^2$ we considered $(x-y)^2 + y^2 = x^2 - 2xy + 2y^2$. We note that $x^2 - 3xy + 3y^2$ gives four different representations with $0 < y < x$ of primes $p \equiv 1 \bmod 6$. The worst case forms are $ax^2 + bxy + cy^2$ with $4ac - b^2 = 3$ and b large and negative. For these we have 5 or 6 representations. For example, with $Q(x,y) = x^2 - 7xy + 13y^2$ we have

$$109 = Q(8,5) = Q(16,7) = Q(27,5) = Q(33,7) = Q(41,12) = Q(43,12)$$

(6 representations),

$$103 = Q(17,2) = Q(25,9) = Q(35,11) = Q(38,9) = Q(42,11)$$

(5 representations).

We therefore need to state our general theorem carefully to take this into account. The reason for the numbers of solutions occurring will become apparent in the proof of the theorem.

Theorem 1.4. *Let $Q(x,y) = ax^2 + bxy + cy^2$ be a positive definite, primitive (i.e. $\gcd(a,b,c) = 1$), binary quadratic form with integer coefficients. Write $D = 4ac - b^2$, so $D > 0$ as the form is positive definite. Let $\delta = \sqrt{D}$, and*

$$\beta = \begin{cases} \frac{1}{2}\pi & \text{if } a + 2b = 0, \\ \arctan(\delta/(b+2a)) & \text{otherwise,} \end{cases}$$

with $\arctan x \in [0,\pi]$. Then, for primes p represented by $Q(x,y)$ with $0 < y < x$ we let x_p, y_p denote a solution to $0 < y_p < x_p$, $Q(x_p,y_p) = p$. For all other primes we write $x_p = y_p = 0$. When there is more than one pair x_p, y_p we assume all pairs are counted in the expressions that follow. We define x_n, y_n similarly for any positive integer n. We then have

$$R_k = \lim_{N\to\infty} \frac{\sum_{p\le N} x_p^k}{\sum_{p\le N} y_p^k} = \lim_{N\to\infty} \frac{\sum_{n\le N} x_n^k}{\sum_{n\le N} y_n^k} = \frac{S(k)}{T(k)},$$

where

$$S(k) = \int_0^\beta (\delta\cos\theta - b\sin\theta)^k \, d\theta \,,$$

and

$$T(k) = (2a)^k \int_0^\beta \sin^k \theta \, d\theta \, .$$

We give one further corollary to illustrate the general result.

Corollary 1.2. *Let $n \in \mathbb{Z}, n \geq 2$. Then, in the notation of Theorem 1.4 with $Q(x, y) = x^2 + ny^2$, we have*

$$R_1 = 1 + \sqrt{n+1},$$

$$R_2 = \frac{n(\sqrt{n} + (1+n)\arctan\sqrt{n})}{-\sqrt{n} + (1+n)\arctan\sqrt{n}},$$

$$R_3 = \frac{(3+2n)(2+3n+2(1+n)^{3/2})}{3+4n}.$$

This follows by taking $a = 1, b = 0, c = n$ in Theorem 1.4 which gives $\delta = 2\sqrt{n}, \beta = \arctan(\sqrt{n})$. It can also be observed that if $n+1$ is a square, then some of these limits are indeed rational numbers.

Z.W. Sun also made a conjecture about the form $u_p^2 + 3u_p v_p + v_p^2$ for prime $p \equiv \pm 1 \bmod 5$:

Conjecture 1.3 (Sun). Let $p \equiv \pm 1 \bmod 5$ be a prime and let $p = u_p^2 + 3u_p v_p + v_p^2$ be the unique representation with positive integers $u_p > v_p$.

$$\lim_{N\to\infty} \frac{\sum_{p\leq N, p\equiv \pm 1 \bmod 5} u_p}{\sum_{p\leq N, p\equiv \pm 1 \bmod 5} v_p} = 1 + \sqrt{5}.$$

He also remarked that it seems that

$$\lim_{N\to\infty} \frac{\sum_{p\leq N, p\equiv \pm 1 \bmod 5} u_p^2}{\sum_{p\leq N, p\equiv \pm 1 \bmod 5} v_p^2} \approx 8.185.$$

While Theorem 1.4 is only valid for positive definite forms we remark that a formal application of the integrals, with $a = 1, b = 3, D = \sqrt{-5}$ leads to $S(1) = -3 + \sqrt{5}$ and $T(1) = 2 - \sqrt{5}$ which predicts indeed

$$\lim_{N\to\infty} \frac{\sum_{p\leq N, p\equiv \pm 1 \bmod 5} u_p}{\sum_{p\leq N, p\equiv \pm 1 \bmod 5} v_p} = 1 + \sqrt{5}.$$

Moreover, for $k = 2$ we find $S(2) = -i(\sqrt{5} - 2\,\mathrm{artanh}(\frac{1}{\sqrt{5}}))$ and $T(2) = -\frac{i}{2}(\sqrt{5} - 4\,\mathrm{artanh}(\frac{1}{\sqrt{5}}))$, which predicts

$$\lim_{N\to\infty} \frac{\sum_{p\leq N, p\equiv \pm 1 \bmod 5} u_p^2}{\sum_{p\leq N, p\equiv \pm 1 \bmod 5} v_p^2} = \frac{2\sqrt{5} - 4\,\mathrm{artanh}(\frac{1}{\sqrt{5}})}{\sqrt{5} - 4\,\mathrm{artanh}(\frac{1}{\sqrt{5}})} \approx 8.18483.$$

Now there are several potential difficulties when looking at the problem for indefinite forms. For example, in the most general case there is not the control of variable size in terms of prime size that we have in the positive definite case. At first sight there could be issues working in fields with infinitely many units, in particular how we count the number of representations. Moreover, the results we need from Coleman's work [1, 2] are not explicitly stated for indefinite forms, although careful study shows that he does indeed prove the theorem we require. We thus are able to prove the following result which covers the above conjecture and many other cases.

Theorem 1.5. *Let $Q(x, y) = ax^2 + bxy + cy^2$ be an indefinite, primitive, binary quadratic form with integer coefficients. Write $D = b^2 - 4ac$, which we assume is not a perfect square. Let $\delta = \sqrt{D}$ and we assume that $0 < \delta/(b + 2a) < 1$, for if this condition fails there will be infinitely many representations of a prime by $Q(x, y)$ with $0 < y < x$. Put $\kappa = \operatorname{artanh}(\delta/(b + 2a))$ (so this is well defined by the previous condition). Then, for primes p represented by $Q(x, y)$ with $0 < y < x$ we define x_p, y_p by $0 < y_p < x_p, Q(x_p, y_p) = p$. When there is more than one pair x_p, y_p we assume all pairs are counted in the expressions that follow. For all other primes we write $x_p = y_p = 0$. We define x_n, y_n similarly for any positive integer n. We then have*

$$R_k = \lim_{N\to\infty} \frac{\sum_{p\le N} x_p^k}{\sum_{p\le N} y_p^k} = \lim_{N\to\infty} \frac{\sum_{n\le N} x_n^k}{\sum_{n\le N} y_n^k} = \frac{U(k)}{V(k)},$$

where

$$U(k) = \int_0^\kappa (\delta \cosh\theta - b \sinh\theta)^k \, d\theta \,,$$

and

$$V(k) = (2a)^k \int_0^\kappa \sinh^k \theta \, d\theta \,.$$

Remark 1.3. We will indicate in the proof how one gets multiple representations of primes and why representations are unique for some forms like $u_p^2 + 3u_p v_p + v_p^2$.

We give a corollary working out these integrals for a particular family:

Corollary 1.3. *Let $n \in \mathbb{Z}, b \ge 3$. Then, in the notation of Theorem 1.5 with $Q(x, y) = x^2 + bxy + y^2$, we have*

$$R_1 = 1 + \sqrt{b + 2}.$$

$$R_2 = \frac{(b-1)\sqrt{b^2-4} - 4 \operatorname{artanh} \sqrt{(b-2)/(b+2)}}{\sqrt{b^2-4} - 4 \operatorname{artanh} \sqrt{(b-2)/(b+2)}}.$$

2 Proofs: The Simplest Case

We begin by quoting a simple consequence of Coleman's work for Gaussian primes which we will use to prove Theorem 1.1 directly and which motivated the whole investigation. The reader will see that the proof of Theorem 1.4 is a straightforward generalisation of this, though the necessary terminology may at first make it look more complicated.

Lemma 2.1 (Coleman, Theorem 2.1 of [1]). *Let* $0 \le \varphi_1 \le \varphi_2 \le 2\pi$, *and* $0 \le y \le x$. *We can define* $S = S(x, y, \varphi_1, \varphi_2) = \{\mathbf{z} \in \mathbb{Z}[i] : x - y < |\mathbf{z}|^2 \le x, \varphi_1 \le \varphi \le \varphi_2\}$, *where* $\varphi = \arg(\mathbf{z}/|\mathbf{z}|)^4$. *Let* $P = P(x, y, \varphi_1, \varphi_2) = \{\mathbf{p} \in S : |\mathbf{p}|^2 = p,\ prime\}$.

Let $\varepsilon > 0$ *be given. We have the asymptotic result,*

$$\sum_{\mathbf{p} \in P(x,y,\varphi_1,\varphi_2)} 1 = \frac{(\varphi_2 - \varphi_1)y}{2\pi \log x}\left(1 + O\left(\frac{1}{\log x}\right)\right),$$

for $\varphi_2 - \varphi_1 > x^{-5/24+\varepsilon}$, $y > x^{19/24+\varepsilon}$, $x > x_\varepsilon$.

Let us first outline the idea: for asymptotically evaluating $\sum_{p \le N, p \equiv 1 \bmod 4} a_p^k$ it suffices to dissect the sector with radius N from $0 < \varphi < \pi/4$ into polar boxes. Coleman's result says that one can dissect this into fine (but not too fine) boxes, so that the number of primes, corresponding to $p = a^2 + b^2$, is asymptotically the right number, with some error of smaller order. We can therefore replace summation by integration with negligible error. The same is true for the sum in the denominator. Hence, we get:

$$\lim_{N \to \infty} \frac{\sum_{p \le N, p \equiv 1 \bmod 4} a_p^k}{\sum_{p \le N, p \equiv 1 \bmod 4} b_p^k} = \frac{\int_0^{\pi/4} \cos^k(x)\, dx}{\int_0^{\pi/4} \sin^k(x)\, dx}.$$

The reader should note that this is exactly the same relation we would get if p ran over all numbers representable as the sum of two squares.

Proof (Proof of Theorem 1.1). We now describe this in more detail. We write $T = 2[\log_2 N], \nu = \sqrt{N}$. Let us first concentrate on Gaussian primes with modulus in the interval $(\frac{\nu}{2}, \nu]$, we later sum up over intervals of type $(\frac{\nu}{2^i}, \frac{\nu}{2^{i-1}}]$. For $1 \le s \le 3T/4, 1 \le t \le T/2$ we then define polar boxes

$$\begin{aligned} B_{s,t} &= \{\mathbf{a} = re^{i\varphi} : 1 - s/T < (r/\nu)^2 \le 1 - (s-1)/T, (t-1)/T < 2\varphi/\pi \le t/T\} \\ &= \left\{\mathbf{a} = re^{i\varphi} : \sqrt{N}\sqrt{1 - s/T} < r \le \sqrt{N}\sqrt{1 - (s-1)/T}, \frac{\pi(t-1)}{2T} < \varphi \le \frac{\pi t}{2T}\right\}. \end{aligned}$$

We note that a polar box $\{(r,\theta) : R_2^2 < r \le R_1^2, \theta' \le \theta \le \theta' + \phi\}$ has area $\frac{1}{2}(R_1^2 - R_2^2)\phi$. It follows (something we will need later when we convert sums to integrals) that the box $B_{s,t}$ has area $N\pi/(4T^2)$. Write $\eta = (\log N)^{-1}$. Now $\log M = \log N + O(1)$ for $N/4 < M \le N$. Hence, by Lemma 2.1 (note though that the corresponding polar box in that lemma has an angle four times that of $B_{s,t}$) for each pair s, t we have

$$\sum_{\mathbf{p}\in B_{s,t}} 1 = \frac{N}{T^2 \log N}\left(1 + O(\eta)\right).$$

In each polar box $\cos\varphi = \cos(t\pi/2T) + O(\eta)$ and similarly for $\sin\varphi$. Also $r = \nu(1 - s/T)^{\frac{1}{2}} + O(\nu\eta)$. Hence

$$\begin{aligned}
\sum_{\frac{N}{4}<p\le N, p\equiv 1 \bmod 4} a_p^k &= \sum_{s,t}\sum_{\mathbf{p}\in B_{s,t}} a_p^k \\
&= (1 + O(\eta)) \sum_{s,t} \frac{N}{T^2 \log N}\left(\nu(1 - s/T)^{\frac{1}{2}}\cos(t\pi/2T) + O(\eta\nu)\right)^k \\
&= \frac{4}{\pi}(\eta + O(\eta^2)) \sum_{s,t} \int\int_{re^{ix}\in B_{s,t}} (r\cos x + O(\nu\eta))^k r\, dr\, dx \\
&= \frac{4}{\pi}(\eta + O(\eta^2)) \int_{\nu/2}^{\nu} r^{k+1}\, dr \int_0^{\pi/4} \cos^k x\, dx.
\end{aligned}$$

We remark that it is easy to check the case $k = 0$ of the above which must give the number of primes $\equiv 1 \bmod 4$ between $N/4$ and N. This is $\frac{3}{8}N\eta(1 + O(\eta))$ and equals the final line of the above display by an elementary calculation.

Similarly,

$$\sum_{\frac{N}{4}<p\le N, p\equiv 1 \bmod 4} b_p^k = \frac{4}{\pi}(\eta + O(\eta^2)) \int_{\nu/2}^{\nu} r^{k+1}\, dr \int_0^{\pi/4} \sin^k x\, dx.$$

Adding up over the intervals $(\frac{\nu}{2^i}, \frac{\nu}{2^{i-1}}]$, and cancelling common factors one finds that

$$\lim_{N\to\infty} \frac{\sum_{p\le N, p\equiv 1 \bmod 4} a_p^k}{\sum_{p\le N, p\equiv 1 \bmod 4} b_p^k} = \frac{\int_0^{\pi/4} \cos^k(x)\, dx}{\int_0^{\pi/4} \sin^k(x)\, dx}.$$

Proof (Proof of Corollary 1.1). We now evaluate these integrals, for small exponents, and determine the arithmetic nature of the values. It is well known that

$$\int_0^{\pi/4} \sin^n(x)\,dx = -\frac{\sin^{n-1} x\cos x}{n}\Big|_0^{\pi/4} + \frac{n-1}{n}\int_0^{\pi/4} \sin^{n-2} x\,dx$$
$$= -\frac{1}{n2^{n/2}} + \frac{n-1}{n}\int_0^{\pi/4} \sin^{n-2} x\,dx,$$

$$\int_0^{\pi/4} \cos^n(x)\,dx = \frac{\cos^{n-1} x\sin x}{n}\Big|_0^{\pi/4} + \frac{n-1}{n}\int_0^{\pi/4} \cos^{n-2} x\,dx$$
$$= \frac{1}{n2^{n/2}} + \frac{n-1}{n}\int_0^{\pi/4} \cos^{n-2} x\,dx.$$

We note for the initial values $n = 0$ and $n = 1$ that

$$\int_0^{\pi/4} \sin^0(x)\,dx = \frac{\pi}{4},$$
$$\int_0^{\pi/4} \sin^1(x)\,dx = 1 - \frac{1}{\sqrt{2}},$$
$$\int_0^{\pi/4} \cos^0(x)\,dx = \frac{\pi}{4},$$
$$\int_0^{\pi/4} \cos^1(x)\,dx = \frac{1}{\sqrt{2}}.$$

From this it is easy to calculate for small k the limit by hand. For example, let $k = 5$.

$$\int_0^{\pi/4} \sin^3(x)\,dx = -\frac{1}{3\cdot 2^{3/2}} + \frac{2}{3}(1 - \frac{1}{\sqrt{2}}) = \frac{1}{12}(8 - 5\sqrt{2}).$$

$$\int_0^{\pi/4} \sin^5(x)\,dx = -\frac{1}{5\cdot 2^{5/2}} + \frac{4}{5}\cdot\frac{1}{12}(8 - 5\sqrt{2}) = \frac{1}{120}(64 - 43\sqrt{2}).$$

$$\int_0^{\pi/4} \cos^3(x)\,dx = \frac{1}{3\cdot 2^{3/2}} + \frac{2}{3}\frac{1}{\sqrt{2}} = \frac{5\sqrt{2}}{12}.$$

$$\int_0^{\pi/4} \cos^5(x)\,dx = \frac{1}{5\cdot 2^{5/2}} + \frac{4}{5}\cdot\frac{5\sqrt{2}}{12} = \frac{43\sqrt{2}}{120}.$$

Hence

$$I_5 = \lim_{N\to\infty} \frac{\sum_{p\le N, p\equiv 1 \bmod 4} a_p^5}{\sum_{p\le N, p\equiv 1 \bmod 4} b_p^5} = \frac{\frac{43\sqrt{2}}{120}}{\frac{1}{120}(64-43\sqrt{2})} = \frac{43\sqrt{2}}{64-43\sqrt{2}}.$$

It is clear from the recursion formulae above, that for odd k, $\int_0^{\pi/4} \cos^k x\,dx = r_1\sqrt{2}$, and $\int_0^{\pi/4} \sin^k x\,dx = r_2 + r_3\sqrt{2}$, where $r_1, r_2, r_3 \in \mathbb{Q}\backslash\{0\}$. It follows that $I_k \in \mathbb{Q}(\sqrt{2}) \setminus \mathbb{Q}$. Similarly, as remarked above, when $k = 2\ell$ we find that $I_k = (A_\ell\pi + B_\ell)/(A_\ell\pi - B_\ell)$ with integers $A_\ell, B_\ell > 0$, so the value of $I(k)$ is a rational expression of π, where the π can never cancel, hence $I(k)$ is transcendental, when $k \ge 2$ is even.

Proof (Proof of Theorem 1.2). We note that

$$\int_0^{\pi/4} \cos^n(x)\,dx = \int_0^{\pi/2} \cos^n(x)\,dx + O\left(2^{-n/2}\right).$$

For n even, say $n = 2m$, we have by the recursion formula

$$\begin{aligned}\int_0^{\pi/2} \cos^n(x)\,dx &= \frac{\pi}{2}\frac{(2m-1)(2m-3)\ldots 3}{2m(2m-2)\ldots 2}\\ &= \frac{\pi}{2}\frac{(2m)!}{(2^m m!)^2}\\ &\sim \frac{\pi}{2}\frac{(2m/e)^{2m}\sqrt{4\pi m}}{2^{2m}(m/e)^{2m}2\pi m}\\ &\sim \left(\frac{\pi}{2n}\right)^{\frac{1}{2}}.\end{aligned}$$

In the above we have used Stirling's formula to obtain asymptotic formulae for $m!$ and $(2m)!$. Similarly, for odd n we have

$$\int_0^{\pi/2} \cos^n(x)\,dx \sim \left(\frac{\pi(1+n)}{2n^2}\right)^{\frac{1}{2}} \sim \left(\frac{\pi}{2n}\right)^{\frac{1}{2}}.$$

On the other hand, we note that

$$\sin(\pi/4 - x) = \frac{1}{\sqrt{2}}\left(1 - x + O(x^2)\right).$$

Hence the natural logarithm of the function $e^{kx}2^{k/2}(\sin(\pi/4-x)^k$ is

$$\tfrac{k}{2}\log 2 + kx + k\log(\sin(\pi/4 - x)) = O(kx^2).$$

From this we easily obtain

$$(\sin(\pi/4 - x))^k = 2^{-k/2} e^{-kx} \left(1 + O\left(kx^2\right)\right).$$

Hence, writing $\lambda = \pi/4 - k^{-2/3}$, $\mu = k^{1/3}$, we have

$$\begin{aligned}\int_0^{\pi/4} \sin^k(x)\,dx &= \int_0^{\lambda} \sin^k(x)\,dx + \int_{\lambda}^{\pi/4} \sin^k(x)\,dx \\ &= O\left(2^{-k/2} e^{-\mu}\right) + \frac{2^{-k/2}}{k}\left(1 + O\left(\mu^{-1}\right)\right).\end{aligned}$$

Combining the above results gives (1) as desired.

3 Proofs: The General Positive Definite Case

A number of authors, including Landau, Hecke [3, 4], Rademacher [11], Kubilius [8, 9], Kalnin′ [6] and Coleman [1, 2], established results with regard to equidistribution of prime ideals in certain regions. We shall here need to work with the distribution of prime ideals and the 1-1 correspondence that exists between these ideals and the representation of their norms by positive definite quadratic forms. A further complication arises that the discriminant of the quadratic form may not be a fundamental discriminant (those given in the next paragraph) and this forces us to use a more general result.

We must now define the notation needed to state Coleman's theorem in its full generality and we will quote this more or less verbatim from [1]. Let $\mathbb{Q}(\sqrt{\Delta})$ be the imaginary quadratic number field with discriminant Δ or 4Δ depending on whether or not $\Delta \equiv 1 \bmod 4$, respectively, and Δ is a negative square-free integer. We use Gothic letters $\mathfrak{a}, \mathfrak{f}$ to denote ideals and $\mathfrak{p}$ will represent a prime ideal. We write $N(\mathfrak{a})$ for the norm of $\mathfrak{a}$. Given a non-zero ideal $\mathfrak{f}$, let $g = g(\mathfrak{f})$ be the number of units ϵ such that $\epsilon \equiv 1 \bmod \mathfrak{f}$. We write K for an ideal class $\bmod \mathfrak{f}$ and (ξ) for the principal ideal generated by an algebraic integer ξ. For each such class we assume there has been chosen and fixed an ideal $\mathfrak{a}_0 \in K^{-1}$. Then given $\mathfrak{a} \in K$ we can define $\xi_{\mathfrak{a}} \in \mathfrak{a}_0$ by $(\xi_{\mathfrak{a}}) = \mathfrak{a}\mathfrak{a}_0$ and $\xi_{\mathfrak{a}} \equiv 1 \bmod \mathfrak{f}$. This algebraic integer is unique up to multiplication by the units $\epsilon \equiv 1 \bmod \mathfrak{f}$. We write

$$\lambda(\xi_{\mathfrak{a}}) = \left(\frac{\xi_{\mathfrak{a}}}{|\xi_{\mathfrak{a}}|}\right)^g.$$

By the definition of g, $\arg(\lambda(\xi_{\mathfrak{a}}))$ is unique $\bmod 2\pi$.

Lemma 3.2 (Coleman, Theorem 2.1 of [1]). *Given* $0 \le \varphi_1 \le \varphi_2 \le 2\pi$, $0 \le y \le x$ *and an ideal class* $K \bmod \mathfrak{f}$. *We define* $S = S(x, y, \varphi_1, \varphi_2, K) = \{\mathfrak{a} \in K : x - y < N(\mathfrak{a}) \le x, \varphi_1 \le \arg(\lambda(\xi_{\mathfrak{a}})) \le \varphi_2\}$. *Let* $P = P(x, y, \varphi_1, \varphi_2, K) = \{\mathfrak{p} \in S : N(\mathfrak{p}) = p,\ prime\}$.

Let $\varepsilon > 0$ be given. We have the asymptotic result,

$$\sum_{\mathfrak{p} \in P(x,y,\varphi_1,\varphi_2,K)} 1 = \frac{(\varphi_2 - \varphi_1)y}{2\pi h(\mathfrak{f}) \log x} \left(1 + O\left(\frac{1}{\log x}\right)\right),$$

for $\varphi_2 - \varphi_1 > x^{-5/24+\varepsilon}$, $y > x^{19/24+\varepsilon}, x > x_\varepsilon$. Here $h(\mathfrak{f})$ is the order of the abelian group of ideal classes mod $\mathfrak{f}$.

Proof (Proof of Theorem 1.4). The proof that

$$\lim_{N\to\infty} \frac{\sum_{n\le N} x_n^k}{\sum_{p\le N} y_n^k} = \frac{S(k)}{T(k)},$$

follows by applying a simple change of variable to map the region

$$0 < y < x, \quad Q(x, y) < N$$

onto the sector

$$0 < r < N^{\frac{1}{2}}, \quad 0 < \theta < \beta.$$

This change of variable will underlie the proof of the more difficult case for primes

$$\lim_{N\to\infty} \frac{\sum_{p\le N} x_p^k}{\sum_{p\le N} y_p^k} = \frac{S(k)}{T(k)},$$

so we shall concentrate on establishing this.

Let $\Delta = -D$, and we suppose initially that this is a fundamental discriminant with $\Delta < -4$. Hence there are just the two units $-1, 1$ in $\mathbb{Q}(\sqrt{\Delta})$. We follow Coleman's argument based on [7] to obtain the 1-1 correspondence between prime ideals and points at which $Q(m, n)$ is prime. In Lemma 2.1 we take

$$\mathfrak{a}_0 = \left[a, \tfrac{1}{2}(b - i\delta)\right].$$

The 1-1 correspondence is then

$$N(\mathfrak{p}) = p = Q(m, n)$$

with

$$\xi_{\mathfrak{p}} = ma + n\left(\frac{b - i\delta}{2}\right).$$

We must remark at this stage that this correspondence is 1-1 between prime ideals and m, n and not necessarily between p and m, n. This is why we may get two

representations if two distinct prime ideals with norm p fall within the sector we are about to describe (which can be larger than the first quadrant if $b < -2a$).

We note that $|\xi_{\mathfrak{p}}|^2 = ap$. So, if we write

$$\phi = \arg(\xi_{\mathfrak{p}}), \quad r = \sqrt{p},$$

we have

$$m = \frac{r}{\sqrt{a\delta}} \left(\delta \cos\phi + b \sin\phi\right), \qquad n = -2a \frac{r}{\sqrt{a\delta}} \sin\phi\,.$$

Now writing $\theta = -\phi$ the condition $0 < n < m$ translates to $0 < \theta < \beta$ with

$$m = \frac{r}{\sqrt{a\delta}} \left(\delta \cos\theta - b \sin\theta\right), \qquad n = 2a \frac{r}{\sqrt{a\delta}} \sin\theta\,.$$

We divide the region $0 < r < N^{\frac{1}{2}}, 0 < \theta < \beta$ into polar boxes as in the proof of Theorem 1.1 and apply Lemma 3.2 to the corresponding regions. As in the proof of Theorem 1.1 we can convert the sums to integrals with smaller order errors. When we divide one sum by the other the integrals over r cancel, as do various constants and the $1/\log N$ factor, leaving just $S(k)/T(k)$ as claimed.

Now if $\Delta = -3$ or -4 we have 6 or 4 units, respectively. We recall that $\xi_{\mathfrak{p}}$ is only unique up to multiplication by units. In the above argument we can have up to two distinct prime ideals and $\xi_{\mathfrak{p}}$ could be multiplied by 3 or 2 units leading to different values m, n and still remain within the sector under consideration. This leads to the multiple representations. Of course, for Theorem 1.1 $\kappa = \pi/4$ constrains $\xi_{\mathfrak{p}}$ to one value, and similarly for Theorem 1.3 $\kappa = \pi/6$, leading to one value. For the form $x^2 - 7xy + 13y^2$ we have $\Delta = -3$, $\kappa = \arctan(-\sqrt{3}/5) = 2.808\ldots$. We note that $5\pi/6 < \kappa < \pi$ and so there will be either 5 or 6 representations depending on whether or not all 6 possible values for ξ lie in the sector.

Now suppose that $D = -f^2\Delta$ with Δ a fundamental discriminant and continue working in $\mathbb{Q}(\sqrt{\Delta})$. If we repeat the above argument, then we would require

$$\xi_{\mathfrak{p}} = ma + n\left(\frac{b - if\delta}{2}\right).$$

We therefore need to restrict ourselves to counting those prime ideals which lead to such $\xi_{\mathfrak{p}}$. This corresponds to restricting the prime ideals to a union of ideal classes mod $\mathfrak{f}$ for a suitable $\mathfrak{f}$. We give the simplest case by way of illustration. Let $Q(x, y) = x^2 + 4y^2$, so $D = 16 = f^2(-\Delta)$ with $f = 2$. Let $\mathfrak{f} = (2)$. Then there are just two ideal classes coprime to $\mathfrak{f}$: $\{(u + vi) : v \text{ even}, u \text{ odd}\}$, $\{(u + vi) : u \text{ even}, v \text{ odd}\}$. Counting only prime ideals in the class with v even will then give $p = x^2 + 4y^2$ as required.

4 The Indefinite Case

We first consider how to describe the geometry in the real case. The natural embedding from $\mathbb{Q}(\delta)$ into $\mathbb{R}^2$ is $a + b\delta \longrightarrow (a, b\delta)$. We write ζ' for the algebraic conjugate of $\zeta \in \mathbb{Q}(\delta)$. We shall discover that the polar boxes of the imaginary case give way to hyperbolic boxes in this new situation. We then note that there is an analogous correspondence between prime ideals and points at which $Q(m, n)$ is prime. However we now need to be much more careful about the number of points (m, n) corresponding to each prime ideal. We need only deal with the fundamental discriminant case as the adaptation to the general case follows as previously. Now we write

$$\begin{aligned}\xi_{\mathfrak{p}} &= ma + n\left(\frac{b-\delta}{2}\right)\\ &= u + v\delta \quad \text{where } u^2 - Dv^2 = r^2 a\\ &= \sqrt{a}r(\cosh\phi + \sinh\phi) \quad \text{for a uniquely defined value } \phi.\end{aligned}$$

Note that this gives a 1-1 correspondence between $\xi_{\mathfrak{p}}$ and (m, n). Following the argument of the previous case, if we write $\theta = -\phi$ then $0 < m < n$ becomes

$$m = \frac{r}{\sqrt{a}\delta}\left(\delta\cosh\theta - b\sinh\theta\right), \qquad n = 2a\frac{r}{\sqrt{a}\delta}\sinh\theta\,,$$

with $0 < \theta < \kappa$. We thus have a hyperbolic box

$$\{(r\cosh\theta, r\sinh\theta) : 0 \le r \le N^{\frac{1}{2}}, 0 < \theta < \kappa\}\,,$$

which we dissect into small hyperbolic boxes in a corresponding manner to our earlier discussion. In particular, where $\xi_{\mathfrak{a}}$ occurs we define $t(\xi_{\mathfrak{a}})$ to be the unique value t such that

$$\xi_{\mathfrak{a}} = \sqrt{N(\mathfrak{a})}(\cosh t + \sinh t) = \sqrt{N(\mathfrak{a})}e^t\,.$$

Now $\xi_{\mathfrak{p}}$ is only unique up to multiplication by units. Let ϵ_0 be the fundamental unit of $\mathbb{Q}(\delta)$. Multiplying $\xi_{\mathfrak{p}}$ by an even power of ϵ_0 gives another candidate for $\xi_{\mathfrak{p}}$ but shifts ϕ by $2\log\epsilon_0$. Hence the maximum number of additional representations of p from this ideal will be strictly less than

$$\frac{\kappa}{2\log\epsilon_0}\,. \tag{2}$$

However, there is the ideal containing the algebraic conjugate of $\xi_{\mathfrak{p}}$, that is $\xi'_{\mathfrak{p}}$, to consider. Possibly multiplying $\xi'_{\mathfrak{p}}$ by an even power of ϵ_0 will give a number in the required range. So, writing $t = t(\xi_{\mathfrak{p}})$, we would need $0 < -t < \kappa$ and $0 <$

$t + 2h \log \epsilon_0 < \kappa$. In the case of $x^2 + 3xy + y^2$ the value of (2) is exactly $1/2$ and so the representation is unique. On the other hand, for $x^2 + xy - y^2$ the value is 1. There can therefore be no more representations from the original ideal, but if $0 < -t < \kappa$ then we obtain $0 < t + 2\log \epsilon_0 < \kappa$ giving exactly one other representation. Of course, it is easy to see there will be two representations for this form as the transformation $(x, y) \longrightarrow (x, x - y)$ leaves the expression unchanged. If $0 < \delta/(2a + b) < 1$ fails, then the only restriction on θ is $\theta > 0$ and so we obtain infinitely many solutions by multiplying $\xi_{\mathfrak{p}}$ by any positive even power of ϵ_0. That is why we cannot prove a result of the required form in this case.

We must now show that Coleman's work in [2] supplies us with the required formula for prime ideals in hyperbolic boxes as above. There Theorem 1.2 is his main theorem. We remark that Hensley [5, §5] also gives an explicit account for how to count prime ideals in hyperbolic boxes in the case of real quadratic fields. We must first describe Hecke characters in $\mathbb{Q}(\delta)$, and we quote Hecke's original definition almost verbatim from [4]. Given an integer ϱ of $\mathbb{Q}(\delta)$ we define

$$\lambda(\varrho) = \exp\left(\frac{i\pi}{\log \epsilon_0} \log \left|\frac{\varrho}{\varrho'}\right|\right).$$

This is clearly a multiplicative function of ϱ which takes the value 1 at all units (as they are powers of ϵ_0, and dividing a unit by its conjugate gives an even power of ϵ_0). The Hecke characters are then all integer powers of $\lambda(\cdot)$. If $N(\varrho) = r$, and $\varrho = r(\cosh t + \sinh t)$, then

$$\varrho' = r(\cosh t - \sinh t), \quad \frac{\varrho}{\varrho'} = e^{2t}.$$

We can therefore use the Hecke characters to pick out the condition $\sigma < t < \sigma + \tau$. Since we can investigate the size of the norm using $N(\varrho)^s (s \in \mathbb{C})$ just as in the imaginary case, we obtain the following result from [2, Theorem 2]. We continue to use the terminology stated before Lemma 3.2.

Lemma 4.3. *Let $\varepsilon > 0$ be given. Given $0 \le \varphi_0 < \varphi_0 + \tau \le 1 < x$, $x^{-1/5+\varepsilon} \le \tau < 1$ and an ideal class K* mod $\mathfrak{f}$. *We define $S = S(x, \varphi_0, \tau, K) = \{\mathfrak{a} \in K : x(1 - \tau) < N(\mathfrak{a}) \le x, \varphi_0 \le t(\xi_{\mathfrak{a}}) \le \varphi_0 + \tau\}$. Let $P = P(x, \varphi_0, \tau, K) = \{\mathfrak{p} \in S : N(\mathfrak{p}) = p,$ prime$\}$. We have the asymptotic result,*

$$\sum_{\mathfrak{p} \in P(x,\varphi_0,\tau,K)} 1 = \frac{x\tau^2}{h(\mathfrak{f}) \log x} \left(1 + O\left(\frac{1}{\log x}\right)\right).$$

Here $h(\mathfrak{f})$ is the order of the abelian group of ideal classes mod $\mathfrak{f}$.

This supplies us with precisely the formula we need for counting prime ideals in a hyperbolic box and the proof can then be easily completed.

Acknowledgements The first named author was partially supported by the Austrian Science Fund (FWF): W1230.

The authors thank the referee for helpful suggestions.

References

1. M.D. Coleman, The distribution of points at which binary quadratic forms are prime. Proc. Lond. Math. Soc. **61**(3), 433–456 (1990)
2. M.D. Coleman, The distribution of points at which norm-forms are prime. J. Number Theory **41**(3), 359–378 (1992)
3. E. Hecke, Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen. I. Math. Z. **1**, 357–376 (1918)
4. E. Hecke, Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen. II. Math. Z. **6**, 11–51 (1920)
5. D. Hensley, An asymptotic inequality concerning primes in contours for the case of quadratic number fields. Acta Arith. **28**, 69–79 (1975)
6. I.M. Kalnin′, On the primes of an imaginary quadratic field which are located in sectors (Russian). Latvijas PSR Zinātņu Akadēmijas Vēstis. Fizikas un Tehnisko Zinātņu Sērija. Izvestiya Akademii Nauk Latviĭskoĭ SSR. Seriya Fizicheskikh i Tekhnicheskikh Nauk **1965**(2), 83–92 (1965)
7. S. Knapowski, On a theorem of Hecke. J. Number Theory **1**, 235–251 (1969)
8. I. Kubilyus, The distribution of Gaussian primes in sectors and contours (Russian). Leningrad. Gos. Univ. Uc. Zap. Ser. Mat. Nauk **137**(19), 40–52 (1950)
9. I.P. Kubilyus, On some problems of the geometry of prime numbers (Russian). Mat. Sb. N.S. **31**(73), 507–542 (1952)
10. OEIS Foundation Inc., The on-line encyclopedia of integer sequences (2011), http://oeis.org/
11. H. Rademacher, Primzahlen reell-quadratischer Zahlkörper in Winkelräumen. Math. Ann. **111**, 209–228 (1935)
12. Z.W. Sun, Conjectures involving primes and quadratic forms, arXiv:1211.1588, version 26, 18 March 2013 (2013)

Large Gaps Between Consecutive Prime Numbers Containing Perfect Powers

Kevin Ford, D.R. Heath-Brown, and Sergei Konyagin

For Helmut Maier in celebration of his 60th birthday

Abstract For any positive integer k, we show that infinitely often, perfect kth powers appear inside very long gaps between consecutive prime numbers, that is, gaps of size

$$c_k \frac{\log p \log_2 p \log_4 p}{(\log_3 p)^2},$$

where p is the smaller of the two primes.

1 Introduction

In 1938, Rankin [11] proved that the maximal gap, $G(x)$, between primes $\leqslant x$, satisfies[1]

$$G(x) \geqslant (c + o(1)) \frac{\log x \log_2 x \log_4 x}{(\log_3 x)^2}, \tag{1}$$

[1] As usual in the subject, $\log_2 x = \log\log x$, $\log_3 x = \log\log\log x$, and so on.

K. Ford
Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 West Green Street, Urbana, IL 61801, USA
e-mail: ford@math.uiuc.edu

D.R. Heath-Brown
Mathematical Institute, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK
e-mail: rhb@maths.ox.ac.uk

S. Konyagin (✉)
Steklov Mathematical Institute, 8 Gubkin Street, Moscow 119991, Russia
e-mail: konyagin@mi.ras.ru

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_5

with $c = \frac{1}{3}$. The following six decades witnessed several improvements of the constant c; we highlight out only a few of these. First, Rankin's own improvement [12] $c = e^{\gamma}$ in 1963 represented the limit of what could be achieved by inserting into Rankin's original 1938 argument best possible bounds on counts of "smooth" numbers. This record stood for a long time until Maier and Pomerance [8] introduced new ideas to improve the constant to $c = 1.31256e^{\gamma}$ in 1989; these were refined by Pintz [10], who obtained $c = 2e^{\gamma}$ in 1997. Very recently, the first and third authors together with Green and Tao [2] have shown that c can be taken arbitrarily large. Independently, this was also proven by Maynard [9].

Rankin's lower bound (1) is probably very far from the truth. Based on a probabilistic model of primes, Cramér [1] conjectured that

$$\limsup_{X\to\infty} \frac{G(X)}{\log^2 X} = 1,$$

and Granville [3], using a refinement of Cramér's model, has conjectured that the lim sup above is in fact at least $2e^{-\gamma} = 1.1229\ldots$. Cramér's model also predicts that the normalized prime gaps $\frac{p_{n+1}-p_n}{\log p_n}$ should have exponential distribution, that is, $p_{n+1} - p_n \geqslant C \log p_n$ for about $e^{-C}\pi(X)$ primes $\leqslant X$.

Our aim in this paper is to study whether or not long prime gaps, say of the size of the right-hand side of the inequality in (1), occur when we impose that an integer of a specified type lies inside the interval. To be precise, we say that a number m is "prime avoiding with constant c" if $m + u$ is composite for all integers u satisfying

$$|u| \leqslant c\frac{\log m \log_2 m \log_4 m}{(\log_3 m)^2}.$$

Here we will be concerned with prime avoiding perfect powers.

Theorem 1.1. *For any positive integer k, there are a constant $c = c(k) > 0$ and infinitely many perfect kth powers which are prime-avoiding with constant c.*

It seems possible that the methods of Green, Ford, Konyagin and Tao, or of Maynard, might be adapted to handle slightly longer intervals containing a kth power but no primes. However we leave this possibility aside for the time being.

2 Sieve Estimates

Throughout, constants implied by the Landau O-symbol and Vinogradov $\ll$-symbol are absolute unless otherwise indicated, e.g. by a subscript such as $\ll_u$. The symbols p and q will always denote prime numbers. Denote by $P^+(n)$ the largest prime factor of a positive integer n, and by $P^-(n)$ the smallest prime factor of n.

We need several standard lemmas from sieve theory, the distribution of "smooth" numbers, and the distribution of primes in arithmetic progressions.

Lemma 2.1. *For large x and $z \leqslant x^{\log_3 x/(10\log_2 x)}$, we have*

$$|\{n \leqslant x : P^+(n) \leqslant z\}| \ll \frac{x}{\log^5 x}.$$

Proof. This follows from standard counts of smooth numbers. Lemma 1 of Rankin [11] also suffices. □

Lemma 2.2. *Let $\mathscr{R}$ denote any set of primes and let $a \in \{-1, 1\}$. Then*

$$|\{p \leqslant x : p \not\equiv a \pmod r \ (\forall r \in \mathscr{R})\}| \ll \frac{x}{\log x} \prod_{\substack{p \in \mathscr{R} \\ p \leqslant x}} \left(1 - \frac{1}{p}\right).$$

Proof. Standard sieve methods [4]. □

Finally, we require a bound of "large sieve" type for averages of quadratic character sums.

Lemma 2.3. *For any set $\mathscr{P}$ of primes in $[2, x]$, and for any $\varepsilon > 0$,*

$$\sum_{\substack{m \leqslant x \\ m \text{ odd}}} \mu^2(m) \Big| \sum_{p \in \mathscr{P}} \left(\frac{p}{m}\right) \Big|^2 \ll_\varepsilon x^{2+\varepsilon}.$$

Proof. This follows immediately from Theorem 1 of [5]. □

3 *k*th Power Residues and Prime Ideals

One of our principal tools is the following estimate for an average of counts of solutions of a certain kth power congruence.

Lemma 3.1. *Let k be a positive integer. For any non-zero integer u and any prime p write*

$$\varrho_{u,k}(p) = \varrho_u(p) = |\{n \pmod p : n^k + u \equiv 0 \pmod p\}|.$$

Then for any fixed $\varepsilon > 0$ and $x \geqslant 2$ we have

$$\prod_{x < p \leqslant y} \left(1 - \frac{\varrho_u(p)}{p}\right) \ll_{k,\varepsilon} |u|^\varepsilon \frac{\log x}{\log y}.$$

The proof of this is based on the Prime Ideal Theorem, and we begin by giving a formal statement of an appropriate form of the latter.

Lemma 3.2. *There is an effectively computable absolute constant $c > 0$ with the following property. Let K be an algebraic number field of degree n_K, and write d_K for the absolute value of the discriminant of K. Let β_0 be the largest simple real zero of $\zeta_K(s)$ in the interval $[\frac{1}{2}, 1]$ if any such exists. Then*

$$|\pi_K(x) - \mathrm{Li}(x)| \leqslant \mathrm{Li}(x^{\beta_0}) + c^{-1} x \exp\{-cn_K^{-1/2} \log^{1/2} x\}$$

for $x \geqslant \exp\{10 n_K \log^2 d_K\}$, where as usual, $\pi_K(x)$ denotes the number of prime ideals of K with norm at most x, and $\mathrm{Li}(x) = \int_2^x dt/\log t$. We omit the first summand on the right-hand side if β_0 does not exist.

This follows from Theorem 1.3 of Lagarias and Odlyzko [7], on choosing $L = K$ in their notation. The reader should note that the counting function $\pi_C(x, L/K)$ of [7] excludes ramified primes, but the number of these is $O(n_K \log d_K)$, which is majorized by $x \exp\{-cn_K^{-1/2} \log^{1/2} x\}$.

In order to handle the term involving the possible simple real zero β_0 we use the following result of Heilbronn [6, Theorem 1].

Lemma 3.3. *A simple real zero of $\zeta_K(s)$ must be a zero of $\zeta_k(s)$ for some quadratic subfield k of K.*

It follows that β_0 is a zero for some quadratic Dirichlet L-function $L(s, \chi)$, with a character χ of conductor dividing d_K. Thus Siegel's Theorem shows that $1 - \beta_0 \geqslant c(\varepsilon) d_K^{-\varepsilon}$, for any fixed $\varepsilon > 0$, with an ineffective constant $c(\varepsilon) > 0$. We then deduce that

$$\mathrm{Li}(x^{\beta_0}) \ll x^{\beta_0} \leqslant x \exp\{-c(\varepsilon) d_K^{-\varepsilon} \log x\} \leqslant x \exp\{-c(\varepsilon) n_K^{-1/2} \log^{1/2} x\}$$

if $n_K \log x \geqslant d_K^{2\varepsilon}$. We therefore obtain the following version of the Prime Ideal Theorem.

Lemma 3.4. *For any $\eta > 0$ there is an ineffective constant $C(\eta) > 0$ with the following property. Let K be an algebraic number field of degree n_K, and write d_K for the absolute value of the discriminant of K. Then*

$$\pi_K(x) = \mathrm{Li}(x) + O(x \exp\{-C(\eta) n_K^{-1/2} \log^{1/2} x\})$$

for

$$x \geqslant \exp\left\{\max\left(10 n_K \log^2 d_K\,,\, n_K^{-1} d_K^{\eta}\right)\right\}.$$

Proof (Proof of Lemma 3.1). Since it may happen that $-u$ is a perfect power we begin by taking a to be the largest divisor of k for which $-u$ is a perfect ath power. Then if $-u = v^a$ and $k = ab$ we see firstly that the polynomial $X^b - v$ is irreducible

over the rationals, and secondly that $n^b \equiv v \pmod{p}$ implies $n^k + u \equiv 0 \pmod{p}$, whence

$$\varrho_{u,k}(p) \geqslant \varrho_{-v,b}(p). \tag{2}$$

We will apply Lemma 3.4 to the field $K = \mathbb{Q}(\theta)$, where θ is a root of $X^b - v$. Thus K has degree $b \leqslant k$. Moreover its discriminant will be a divisor of

$$D := \mathrm{Disc}(1, \theta, \theta^2, \ldots, \theta^{b-1}) = (-1)^{b-1} b^b v^{b-1}.$$

We now set

$$x_0 = C(k, \eta) \exp(|D|^\eta).$$

If we choose the constant $C(k, \eta)$ sufficiently large, then whenever $x \geqslant x_0$ we will have

$$x \exp\{-C(\eta) b^{-1/2} \log^{1/2} x\} \leqslant x \log^{-2} x$$

and

$$x \geqslant \exp\left\{\max\left(10b(\log|D|)^2\,,\, b^{-1}|D|^\eta\right)\right\}.$$

It therefore follows from Lemma 3.4 that

$$\pi_K(x) = \mathrm{Li}(x) + O_{k,\eta}\left(x \log^{-2} x\right)$$

for $x \geqslant x_0$.

We now write $\nu_K(p)$ for the number of first degree prime ideals of K lying above p. Then

$$\pi_K(x) = \sum_{p \leqslant x} \nu_K(p) + O_k\Big(\sum_{p^e \leqslant x,\, e \geqslant 2} 1\Big).$$

Moreover, by Dedekind's Theorem we will have $\varrho_{-v,b}(p) = \nu_K(p)$ whenever $p \nmid D$. In the remaining case in which $p \mid D$ we have $\varrho_{-v,b}(p) \leqslant b \leqslant k$ and $\nu_K(p) \leqslant b \leqslant k$. It therefore follows that

$$\pi_K(x) = \sum_{p \leqslant x} \varrho_{-v,b}(p) + O_k(x^{1/2}) + O_k(\log|D|),$$

so that

$$\sum_{p \leqslant x} \varrho_{-v,b}(p) = \mathrm{Li}(x) + O_{k,\eta}(x \log^{-2} x)$$

when $x \geqslant x_0$.

We now observe that

$$\prod_{x<p\leqslant y}\left(1-\frac{\varrho_u(p)}{p}\right)\leqslant \exp\left\{-\sum_{x<p\leqslant y}\frac{\varrho_u(p)}{p}\right\}$$
$$\leqslant \exp\left\{-\sum_{x<p\leqslant y}\frac{\varrho_{-v,b}(p)}{p}\right\}$$

by (2). Assuming that $y\geqslant x_0$ we may then use summation by parts to calculate that

$$\begin{aligned}\sum_{x<p\leqslant y}\frac{\varrho_{-v,b}(p)}{p} &\geqslant \sum_{\max(x,x_0)<p\leqslant y}\frac{\varrho_{-v,b}(p)}{p}\\ &= \log\log y-\log\log\big(\max(x,x_0)\big)+O_{k,\eta}(1)\\ &\geqslant \log\log y-\log\log x-\log\log x_0+O_{k,\eta}(1)\\ &= \log\log y-\log\log x-\eta\log|D|+O_{k,\eta}(1)\\ &\geqslant \log\log y-\log\log x-\eta k\log|u|+O_{k,\eta}(1).\end{aligned}$$

We therefore have

$$\prod_{x<p\leqslant y}\left(1-\frac{\varrho_u(p)}{p}\right)\ll_{k,\eta}|u|^{k\eta}\frac{\log x}{\log y}$$

when $y\geqslant x_0$. Of course this estimate is trivial when $y\leqslant x_0$ since one then has $\log y\ll_{k,\eta}|D|^{\eta}\ll_{k,\eta}|u|^{k\eta}$. The lemma then follows. □

4 Main Argument

Fix a positive integer k. Let x be a large number, sufficiently large depending on k, let c_1 and c_2 be two positive constants depending on k to be chosen later, and put

$$N=\prod_{p\leqslant x}p,\qquad z=x^{c_1\log_3 x/\log_2 x},\qquad y=\frac{c_2x\log x\log_3 x}{(\log_2 x)^2}.$$

In the rest of the paper we will prove the following lemma.

Lemma 4.1. *There is a number $m\leqslant 2N$ such that m^k+u is composite for $|u|\leqslant y$.*

Theorem 1.1 will follow upon observing that $m^k\leqslant e^{kx+o(x)}$ as $x\to\infty$ and consequently that

$$y \gg_k \frac{\log(m^k)\log_2(m^k)\log_4(m^k)}{(\log_3(m^k))^2}.$$

We will select m by choosing residue classes for m modulo p for primes $p \leqslant x$. Let

$$\mathscr{P}_1 = \{p : p \leqslant \log x \text{ or } z < p \leqslant x/4\}, \qquad \mathscr{P}_2 = \{p : \log x < p \leqslant z\}.$$

We first choose

$$\begin{aligned} m &\equiv 0 \pmod{p} \quad (p \in \mathscr{P}_1), \\ m &\equiv 1 \pmod{p} \quad (p \in \mathscr{P}_2). \end{aligned} \tag{3}$$

Observe that $p|(m^k + u)$ if $p|u$ for some $p \in \mathscr{P}_1$. Because $y < (x/4)\log x$, any remaining value of u is thus either composed only of primes in $\mathscr{P}_2$ (in particular, u is z-smooth), including $|u| = 1$, or $|u|$ is a prime larger than $x/4$. For any u in the latter category such that $p|(u+1)$ for some $p \in \mathscr{P}_2$, $p|(m^k+u)$. Let U denote the set of exceptional values of u, that is, the set of $u \in [-y, y]$ not divisible by any prime in $\mathscr{P}_1$, and such that if $|u|$ is prime then $p \nmid (u+1)$ for all $p \in \mathscr{P}_2$. By Lemmas 2.1 and 2.2, if c_1 is sufficiently small, then

$$|U| \ll \frac{y}{\log^5 x} + \frac{y}{\log x} \prod_{p \in \mathscr{P}_2} \left(1 - \frac{1}{p}\right) \ll \frac{y \log_2 x}{\log x \log z} = \frac{c_2}{c_1} \frac{x}{\log x}.$$

Choosing c_2 appropriately, we can ensure that $|U| \leqslant \delta x/\log x$, where $\delta > 0$ depends on k (δ will be chosen later).

The remaining steps depend on whether k is odd or even. If k is odd, the construction is very easy. For each $u \in U$, associate with u a different prime $p_u \in (x/4, x]$ such that $(p_u - 1, k) = 1$ (e.g., one can take $p_u \equiv 2 \pmod{k}$ if $k \geqslant 3$). Then every residue modulo p_u is a kth power residue, and we take m in the residue class modulo p_u such that

$$m^k \equiv -u \pmod{p_u} \quad (u \in U). \tag{4}$$

By the prime number theorem for arithmetic progressions, the number of available primes is at least

$$x/(2\phi(k)\log x) \geqslant |U|$$

if δ is small enough. With this construction, $p_u|(m^k + u)$ for every $u \in U$. Therefore, $m^k + u$ is divisible by a prime $\leqslant x$ for every $|u| \leqslant y$. Furthermore, (3) and (4) together imply that m is defined modulo a number N', where $N'|N$. Therefore, there is an admissible value of m satisfying $N < m \leqslant 2N$. The prime number theorem implies that $N = e^{x+o(x)}$, thus $m^k - y > x$. Consequently, $m^k + u$ is composite for $|u| \leqslant y$.

Now suppose that k is even. There do not exist primes for which every residue modulo p is a kth power residue. However, we maximize the density of kth power residues by choosing primes p such that $(p-1,k)=2$, e.g. taking $p \equiv 3 \pmod{2k}$. For such primes p, every quadratic residue is a kth power residue. Let

$$\mathscr{P}_3 = \{x/4 < p \leqslant x/2 : p \equiv 3 \pmod{2k}\}.$$

By the prime number theorem for arithmetic progressions, $|\mathscr{P}_3| \geqslant x/(5\phi(2k)\log x)$. We aim to associate numbers $u \in U$ with distinct primes $p_u \in \mathscr{P}_3$ such that $\left(\frac{-u}{p_u}\right) = 1$. This ensures that the congruence $m^k + u \equiv 0 \pmod{p_u}$ has a solution. We, however, may not be able to find such p for every $u \in U$, but can find appropriate primes for most u. Let

$$U' = \left\{ u \in U : \left(\frac{-u}{p}\right) = 1 \text{ for at most } \frac{\delta x}{\log x} \text{ primes } p \in \mathscr{P}_3 \right\}.$$

The numbers $u \in U \backslash U'$ may be paired with different primes $p_u \in \mathscr{P}_3$ such that $\left(\frac{-u}{p_u}\right) = 1$. We then may take m such that

$$m^k \equiv -u \pmod{p_u} \quad (u \in U \backslash U'). \tag{5}$$

Next we will show that $|U'|$ is small. Write

$$S = \sum_{u \in U} \Big| \sum_{p \in \mathscr{P}_3} \left(\frac{-u}{p}\right) \Big|^2.$$

Each u may be written uniquely in the form $u = su_1^2 u_2$, where $s = \pm 1$, $u_2 > 0$ and u_2 is squarefree. By quadratic reciprocity,

$$\left(\frac{-u}{p}\right) = (-s)\left(\frac{u_2}{p}\right) = (-s)(-1)^{\frac{u_2-1}{2}}\left(\frac{p}{u_2}\right),$$

since $p \equiv 3 \pmod 4$. Given u_2, there are at most $\sqrt{y/u_2} \leqslant \sqrt{y}$ choices for u_1. Hence, using Lemma 2.3,

$$\begin{aligned} S &= \sum_{u \in U} \Big| \sum_{p \in \mathscr{P}_3} \left(\frac{p}{u_2}\right) \Big|^2 \\ &\leqslant \sum_{u_2 \leqslant y} 2y^{1/2} \Big| \sum_{p \in \mathscr{P}_3} \left(\frac{p}{u_2}\right) \Big|^2 \\ &\ll_\varepsilon x^{5/2+\varepsilon}. \end{aligned}$$

Now let $\delta = \frac{1}{15\phi(2k)}$, so that $\delta x/\log x \leqslant \frac{1}{3}|\mathscr{P}_3|$. If $u \in U'$, then clearly

$$\Big|\sum_{p\in\mathscr{P}_3}\left(\frac{-u}{p}\right)\Big| \geqslant \frac{1}{3}|\mathscr{P}_3| \geqslant \delta\frac{x}{\log x}.$$

It follows that $|S| \gg |U'|(x/\log x)^2$, and consequently that

$$|U'| \ll_\varepsilon x^{1/2+2\varepsilon}. \tag{6}$$

Let $A \mod M$ denote the set of numbers m satisfying the congruence conditions (3) and (5), where $0 \leqslant A < M$. Thus, if $m \equiv A \pmod M$ and $u \notin U'$, then $m^k + u$ is divisible by a prime $\leqslant x/2$. Let

$$K = \prod_{x/2<p\leqslant x} p.$$

We'll take $m = Mj + A$, where $1 \leqslant j \leqslant K$, and aim to show that there exists a value of j so that $(mj + A)^k + u$ is composite for every $u \in U'$. By sieve methods (see [4]),

$$\begin{aligned}\sum_{j=1}^{K}\#\{u \in U' : (Mj + A)^k + u \text{ prime}\} &= \sum_{u\in U'}\#\{1 \leqslant j \leqslant K : (Mj + A)^k + u \text{ prime}\}\\ &\ll \sum_{u\in U'} K\prod_{y<q\leqslant\sqrt{K}}\left(1 - \frac{\varrho_u(q)}{q}\right).\end{aligned}$$

By Lemma 3.1, the above product is

$$\ll_{k,\varepsilon} u^{\varepsilon/2}\frac{\log y}{\log K} \ll_{k,\varepsilon} u^{\varepsilon/2}\frac{\log x}{x}.$$

Combined with our estimate (6) for the size of $|U'|$, we find that

$$\sum_{1\leqslant j\leqslant K}\#\{u \in U' : (Mj + A)^k + u \text{ prime}\} \ll_{k,\varepsilon} \frac{K}{x^{1/2-4\varepsilon}}.$$

It follows that the left-hand side above is zero for some j. That is, $(Mj + A)^k + u$ is composite for every $u \in U'$. Therefore, $(Mj + A)^k + u$ is composite for every u satisfying $|u| \leqslant y$. Finally, we note that $Mj + A \leqslant 2N$, and the proof of Lemma 4.1 is complete.

Remark 4.1. For odd k the constant $c(k)$ in Theorem 1.1 is effective. For even k it is ineffective due to the use of Siegel's theorem in the proof of Lemma 3.1.

Acknowledgements Research of the third author was partially performed while he was visiting the University of Illinois at Urbana-Champaign. Research of the first author was partially performed while visiting the University of Oxford. Research of the first and third authors was also carried out in part at the University of Chicago, and they are thankful to Prof. Wilhelm Schlag for hosting these visits.

The first author was supported by NSF grant DMS-1201442. The research of the second author was supported by EPSRC grant EP/K021132X/1. The research of the third author was partially supported by Russian Foundation for Basic Research, Grant 14-01-00332, and by Program Supporting Leading Scientific Schools, Grant Nsh-3082.2014.1.

References

1. H. Cramér, On the order of magnitude of the difference between consecutive prime numbers. Acta Arith. **2**, 396–403 (1936)
2. K. Ford, B. Green, S. Konyagin, T. Tao, Large gaps between consecutive prime numbers. arXiv:1408.4505
3. A. Granville, Harald Cramér and the distribution of prime numbers. Scand. Actuar. J. **1**, 12–28 (1995)
4. H. Halberstam, H.-E. Richert, *Sieve Methods* (Academic, London, 1974)
5. D.R. Heath-Brown, A mean value estimate for real character sums. Acta Arith. **72**, 235–275 (1995)
6. H. Heilbronn, On real zeros of Dedekind ζ-functions. Can. J. Math. **25**, 870–873 (1973)
7. J.C. Lagarias, A.M. Odlyzko, Effective versions of the Chebotarev density theorem, in *Algebraic Number Fields* (Proceedings of the Symposia, University Durham, Durham, 1975) (Academic, London, 1977), pp. 409–464
8. H. Maier, C. Pomerance, Unusually large gaps between consecutive primes. Trans. Am. Math. Soc. **322**(1), 201–237 (1990)
9. J. Maynard, Large gaps between primes. arXiv:1408.5110
10. J. Pintz, Very large gaps between consecutive primes. J. Number Theory **63**(2), 286–301 (1997)
11. R. Rankin, The difference between consecutive prime numbers. J. Lond. Math. Soc. **13**, 242–247 (1938)
12. R.A. Rankin, The difference between consecutive prime numbers. V. Proc. Edinb. Math. Soc. **13**(2), 331–332 (1962/1963)

On the Parity of the Number of Small Divisors of n

Kevin Ford, Florian Luca, Carl Pomerance, and Jeffrey Shallit

To Professor Helmut Maier on his 60th birthday

Abstract For a positive integer j we look at the parity of the number of divisors of n that are at most j, proving that for large j, the count is even for most values of n.

Keywords Number of divisors

Mathematics Subject Classification: 11N25, 20K01

1 Introduction

Let $\tau(n)$ denote the number of positive divisors of the positive integer n. It is easy to see that $\tau(n)$ is odd if and only if n is a square, so in the sense of asymptotic density, $\tau(n)$ is almost always even. In this note we consider the function $\tau_j(n) = \#\{d \mid n : d \leqslant j\}$, the number of positive divisors of n that are at most j. Here j is a positive integer. Can we say that $\tau_j(n)$ is usually even? Evidently not. This is patently false for $j = 1$, and it is false for all odd numbers n when $j \leqslant 2$. Here's another trivial case. Say n is not a square and $n/2 \leqslant j < n$. Then $\tau_j(n)$ is odd. In fact, if we list out

K. Ford
Department of Mathematics, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA
e-mail: ford@math.uiuc.edu

F. Luca (✉)
School of Mathematics, University of the Witwatersrand, Private Bag X3, Wits 2050, South Africa
e-mail: florian.luca@wits.ac.za

C. Pomerance
Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA
e-mail: carl.pomerance@dartmouth.edu

J. Shallit
School of Computer Science, University of Waterloo, Waterloo, ON, Canada N2L 3G1
e-mail: shallit@cs.uwaterloo.ca

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_6

all of the divisors of n: $1 = d_1 < d_2 < \cdots < d_{\tau(n)} = n$ and choose j at random in $[1, n]$, when n is not a square, more than half of the time $\tau_j(n)$ will be odd, since the top interval $[n/2, n)$ takes up half of the available values of j.

We are interested in the range $2 \leqslant j \leqslant \sqrt{n}$, showing that $\tau_j(n)$ tends to be even here. Let

$$\delta = 1 - \frac{1 + \log\log 2}{\log 2} = 0.08607\ldots .$$

Theorem 1.1. *Let $N_j(x)$ denote the number of integers $n \leqslant x$ with $\tau_j(n)$ odd. Uniformly for $2 \leqslant j \leqslant \sqrt{x}$,*

$$N_j(x) = O\left(\frac{x}{(\log j)^{\delta/(1+\delta)}(\log\log(2j))^{1.5/(1+\delta)}}\right).$$

The theorem implies that when j is large and fixed, $\tau_j(n)$ is usually even as n varies.

It is interesting to look at this problem numerically. For a fixed number j, whether $\tau_j(n)$ is even or odd depends solely on the value of $\gcd(n, L_j)$, where L_j is the least common multiple of the integers in $[1, j]$. That is, $\tau_j(n) = \tau_j(\gcd(n, L_j))$. Thus, the set of integers n with $\tau_j(n)$ odd is a union of residue classes modulo L_j, so the asymptotic density of the set of such n exists; it is $N_j(L_j)/L_j$.

j	L_j	$N_j(L_j)$	$N_j(L_j)/L_j$
1	1	1	1
2	2	1	0.5
3	6	3	0.5
4	12	7	0.5833333333
5, 6	60	33	0.55
7	420	225	0.5357142857
8	840	405	0.4821428571
9	2520	1305	0.5178571429
10	2520	1235	0.4900793651
11, 12	27,720	13,635	0.4918831169
13	360,360	177,705	0.4931318681
14	360,360	170,775	0.4739010989
15	360,360	170,181	0.4722527473
16	720,720	359,073	0.4982142857
17	12,252,240	6,106,815	0.4984243697
18	12,252,240	5,919,705	0.4831528765
19	232,792,560	112,887,225	0.4849262580
20	232,792,560	109,706,355	0.4712622903
21	232,792,560	110,362,725	0.4740818392
22	232,792,560	107,787,735	0.4630205321
23, 24	5,354,228,880	2,496,334,995	0.4662361380
25	26,771,144,400	12,782,443,905	0.4774709558
26	26,771,144,400	12,538,223,775	0.4683484422
27	80,313,433,200	37,368,330,615	0.4652812005
28	80,313,433,200	36,653,106,105	0.4563757848

Our theorem implies that the right column approaches the limit 0 as $j \to \infty$ slightly faster, at the least, than $(\log j)^{-\delta/(1+\delta)}$.

In the following table we consider some larger values of j but only via some statistical experiments to approximate the density $N_j(L_j)/L_j$. The experiments involved taking the first 10^4 numbers following the kth prime, for $k = 10^5, 2 \times 10^5, \ldots, 6 \times 10^5$. The numbers in the table are actual counts of the number of odd values of $\tau_j(n)$ among the 10^4 values of n. The numbers weakly suggest that $N_j(L_j)/L_j$ decays to 0 like $(\log j)^{-\theta}$ where θ is slightly above $1/2$. However, this too is misleading. Indeed, we will show below in Theorem 2.3 that $N_j(L_j)/L_j$ decays more slowly than about $1/(\log j)^{\delta}$. We do not resolve the issue of the "correct" exponent on $\log j$, but we do give a suggested plan for proving it is asymptotically δ.

j	10^5	2×10^5	3×10^5	4×10^5	5×10^5	6×10^5
100	4131	4121	4077	4099	4123	4109
200	4061	4107	4174	4181	4231	4050
300	3800	3850	3954	3980	4002	3969
400	3630	3703	3800	3744	3877	3875
500	3466	3587	3673	3710	3793	3772
600	3351	3512	3526	3594	3722	3682
700	3294	3435	3502	3543	3627	3593
800	3213	3301	3431	3475	3577	3574
900	2822	3245	3337	3411	3522	3477
1000	2358	3197	3248	3334	3459	3439

Throughout this note, the constants implied the by Landau symbol O and by the Vinogradov symbols $\ll$ and $\gg$ are absolute. We also use the notation $A \asymp B$ if $A \ll B \ll A$. We also write $a \| b$ for positive integers a, b if $a \mid b$ and $\gcd(a, b/a) = 1$.

2 Proof of Theorem 1.1

We begin with a criterion for $\tau_j(n)$ to be even.

Lemma 2.1. *Let $j \geqslant 2$. Suppose that there is a prime $p \| n$ and n has no divisor from the interval $(j, pj]$. Then $\tau_j(n)$ is even.*

Proof. Suppose the hypotheses hold. Let

$$A = \{d \mid n : d \leqslant j, p \nmid d\}, \quad B = \{d \mid n : d \leqslant j, p \mid d\},$$

so that $\{d \mid n : d \leqslant j\}$ is the disjoint union of A and B. It suffices to show that $\#A = \#B$. For each $d \in A$ consider pd. Then $pd \mid n$ and $pd \leqslant pj$. But by our hypothesis, we must then have $pd \leqslant j$, so that $pd \in B$. Thus, $\#A \leqslant \#B$. Now take $d \in B$. We may write $d = pd'$, where $d' \mid n$, $p \nmid d'$, and $d' \leqslant j/p \leqslant j$. Thus, $d' \in A$, which shows that $\#B \leqslant \#A$. So, $\#A = \#B$, completing the proof.

For real numbers $x \geqslant z \geqslant y \geqslant 1$, let $H(x, y, z)$ denote the number of integers $n \leqslant x$ which have a divisor in the interval $(y, z]$. From the main theorem in Ford [3], we have the following result.

Lemma 2.2. *Suppose that $y \geqslant 2$ and $2y \leqslant z \leqslant y^2 \leqslant x$. Let $u = \log z/\log y - 1$, so that $z = y^{1+u}$. Then,*

$$H(x, y, z) \asymp xu^{\delta}(\log(2/u))^{-3/2}.$$

Note that [3] states this result for $y \geqslant 100$ and $x \geqslant 100{,}000$, but by adjusting the implicit constants, the result can be seen to hold in the larger range asserted.

We now proceed to the proof of the theorem. Let $2 \leqslant k \leqslant j$ be a parameter to be chosen shortly. First note that the number of $n \leqslant x$ for which there is no prime $p \leqslant k$ with $p \| n$ is $O(x/\log k)$. Indeed this follows from sieve methods, in particular [4, Theorem 2.2]. Let $u = \log k/\log j$ and $z = kj = j^{1+u}$. By Lemma 2.2, the number of $n \leqslant x$ which have a divisor in $(j, z]$ is $O(xu^{\delta}(\log(2/u))^{-3/2})$. Let us equate these two O-estimates so as to fix our parameter k:

$$\frac{x}{\log k} = xu^{\delta}(\log(2/u))^{-3/2} = x\left(\frac{\log k}{\log j}\right)^{\delta}\left(\log\left(\frac{2\log j}{\log k}\right)\right)^{-3/2}.$$

After a small calculation this leads to a reasonable choice for k being

$$k = \exp\left((\log j)^{\delta/(1+\delta)}(\log\log(2j))^{1.5/(1+\delta)}\right).$$

With this choice of k we have that the number of $n \leqslant x$ for which it is not the case that both

- there is a prime $p \leqslant k$ with $p \| n$,
- n is free of divisors from the interval $(j, kj]$,

is $O(x/\log k)$. By Lemma 2.1, if both of these conditions hold, then $\tau_j(n)$ is even. With the choice of k given just above, this completes the proof.

It is clear that any integer n which has no prime factors in $[1, j]$ also has $\tau_j(n) = 1$; that is, $\tau_j(n)$ is odd. Thus,

$$N_j(x) \gg \frac{x}{\log j},$$

using [4, Theorem 2.5].

This "trivial" lower bound can be improved using ideas similar to those used to prove Theorem 1.1.

Theorem 2.3. *Let $c > 0$ be arbitrarily small. Uniformly for $j \leqslant x^{1/2-c}$ and x sufficiently large (in terms of c), we have*

$$N_j(x) \gg \frac{x}{(\log j)^{\delta}(\log\log(2j))^{3/2}}.$$

Proof. It follows from [3, Theorem 4] that for $j \leqslant x^{1/2-c}$, the number of integers $n \leqslant x$ with exactly 1 divisor in $(j/2, j]$ is of magnitude $x(\log j)^{-\delta}(\log\log(2j)^{-3/2}$. Further, from the comments in the first paragraph of Sect. 1.3 in [3], the same is true if we ask in addition that n is odd. For such an odd number n, it follows by an argument akin to that of Lemma 2.1 that $\tau_j(2n)$ is odd. The claimed lower bound follows.

To close the gap between this lower bound and Theorem 1.1, the following strategy might be tried. It follows from Lemma 2.1 and its proof that if $\tau_j(n)$ is odd and if p is the least prime with $p\|n$, then n has a divisor in $(j, pj]$ that is divisible by p. (It is also possible that n has no prime factor p with $p\|n$, but such numbers are negligible.) Let $N_j(x, p)$ denote the number of integers $n \leqslant x$ such that (i) p is the least prime with $p\|n$ and (ii) n has a divisor in $(j, pj]$ divisible by p. Again following the thoughts in the first paragraph of Sect. 1.3 of [3], it may be possible to show that for each $p \leqslant \exp((\log j)^{1/2})$, $N_j(x, p)$ is uniformly bounded above by a constant times

$$\frac{1}{p\log p}\,\frac{x(\log p)^{\delta}}{(\log j)^{\delta}(\log\log(2j))^{3/2}}. \tag{1}$$

That is, a factor $p\log p$ is introduced in the denominator due to the condition that p is the least prime with $p\|n$. Summing this estimate for $p \leqslant \exp((\log j)^{1/2})$ yields the estimate $O(x(\log j)^{-\delta}(\log\log(2j))^{-3/2})$, with larger values of p being trivially negligible. Thus, we would have a match with the improved lower bound, at least for $j \leqslant x^{1/2-c}$.

The estimate (1) would follow if one could show that the number $N'_j(x, p)$ of integers $m \leqslant x/p$ having a divisor in $(j/p, j]$ and such that if $q\|m$ then $q > p$, is at most a constant times the expression in (1) for $p \leqslant \exp((\log j)^{1/2})$. Multiplying such a number m by p would cover all those n counted by $N_j(x, p)$, that is, $N_j(x, p) \leqslant N'_j(x, p)$. It would seem that upper bounding $N'_j(x, p)$ in this way is eminently provable using the ideas in [3], since integers m with such restrictions on their small prime divisors seem less likely to have a divisor in a given interval than integers in general.

3 A Corollary

We saw at the start that if j is randomly chosen in $[1, n]$, it is more likely than not that $\tau_j(n)$ is odd, despite our theorem. This is because of the huge weight of the interval $[n/2, n)$. To equalize things, we might take a harmonic measure. For $y \in \mathbb{R}$, $y \geqslant 1$, let $\tau_y(n) = \tau_{\lfloor y\rfloor}(n)$. Let

$$S(n) = \{y \in [1, n] \,:\, \tau_y(n) \text{ is odd}\}, \quad f(n) = \frac{1}{\log n}\int_{S(n)} \frac{dy}{y}.$$

Then we always have $0 \leqslant f(n) \leqslant 1$. Further, if n is prime, then $f(n) = 1$, while if $n = 2p$ where p is prime, then $f(n) \to 0$ as $p \to \infty$. We can ask what is the normal value of the statistic $f(n)$. The following corollary of Theorem 1.1 addresses this question.

Corollary 3.1. *There is a set of integers $\mathscr{A}$ of asymptotic density* 1*, such that if $n \to \infty$ with $n \in \mathscr{A}$, then $f(n) \to 0$.*

Proof. Since we always have $f(n) \in [0, 1]$, the assertion of the corollary is equivalent to

$$\sum_{n \leqslant x} f(n) = o(x), \quad x \to \infty,$$

and this in turn is equivalent to

$$\sum_{n \in (\frac{1}{2}x, x]} f(n) = o(x), \quad x \to \infty,$$

which is equivalent to

$$\sum_{n \in (\frac{1}{2}x, x]} \int_{S(n)} \frac{dy}{y} = o(x \log x), \quad x \to \infty. \tag{2}$$

For $n \in (\frac{1}{2}x, x]$, consider its divisors $1 = d_1 < d_2 < \cdots < d_{\tau(n)} = n$, so that

$$\int_{S(n)} \frac{dy}{y} = \sum_{\substack{i < \tau(n) \\ i \text{ odd}}} \log\left(\frac{d_{i+1}}{d_i}\right).$$

The interval (d_i, d_{i+1}) has the companion interval $(n/d_{i+1}, n/d_i)$, which is the same as $(d_{\tau(n)-i}, d_{\tau(n)-i+1})$. Further if n is not a square, i is odd if and only if $\tau(n) - i$ is odd. Thus, for n not a square,

$$\int_{S(n)} \frac{dy}{y} = 2 \int_{S(n) \cap [1, \sqrt{n}]} \frac{dy}{y}.$$

Since the squares are negligible, to prove (2), it now suffices to prove that

$$\sum_{n \in (\frac{1}{2}x, x]} \int_{S(n) \cap [1, \sqrt{n}]} \frac{dy}{y} = o(x \log x), \quad x \to \infty. \tag{3}$$

This sum is equal to

$$\sum_{n\in(\frac{1}{2}x,x]}\sum_{\substack{j\leqslant\sqrt{n}\\ \tau_j(n)\ \text{odd}}}\log\frac{j+1}{j},$$

except for a possible error of $o(1)$ as $x\to\infty$ caused by $j=\lfloor\sqrt{n}\rfloor$. Ignoring this triviality, the sum in (3) is now equal to

$$\sum_{j\leqslant\sqrt{x}}\log\frac{j+1}{j}\sum_{\substack{n\in(\frac{1}{2}x,x]\\ j\leqslant\sqrt{n}\\ \tau_j(n)\ \text{odd}}}1\ll x\sum_{j\leqslant\sqrt{x}}\frac{1}{j(\log(2j))^{\delta/(1+\delta)}}\ll x(\log x)^{1/(1+\delta)},$$

using $\log((j+1)/j)<1/j$ and Theorem 1.1. Since $1/(1+\delta)<1$, it follows that (3) holds, and as we have seen, this is sufficient for the corollary. This completes the proof.

4 Final Thoughts

Though the situation is much simpler in this note, the idea behind our Lemma 2.1 was inspired by the argument in Maier [5].

One might ask about other residue classes for $\tau_j(n)$. Our proof can show that for each fixed positive integer k, the set of numbers n such that $k\mid\tau(n)$ and $k\nmid\tau_j(n)$ has asymptotic density $o(1)$ as $j\to\infty$. For k not a power of 2, it might be interesting to investigate the density of those numbers n where $k\nmid\tau(n)$ and also $k\nmid\tau_j(n)$.

It is interesting to see several connections of this note to work of A.S. Besicovitch. First, Ford's theorem, cited in Lemma 2.2, is the latest chapter in a long story that began with work of Besicovitch [1] in 1934, when he showed that $\lim_{x\to\infty}H(x,y,2y)/x$ has liminf 0 as $y\to\infty$. And second, this note was motivated originally by looking for examples for sequences that perhaps violated a result known as the Besicovitch pseudometric, see [2]. In particular, it was thought if the densities $\lim_{x\to\infty}N_j(x)/x$ did not approach 0, then this would be a violation. It is interesting that we used a descendant of the 1934 Besicovitch result to show that these densites do approach 0 and so there is no counterexample.

Acknowledgements Research of Kevin Ford was partially supported by the grant DMS-1201442 from the National Science Foundation. Research of Florian Luca was done when he visited the Mathematics Department of Dartmouth College in Spring of 2014. He thanks this department for their hospitality. Part of this work was done when Carl Pomerance was visiting the National Center for Theoretical Sciences in Taiwan. He thanks the center for their hospitality and support.

References

1. A.S. Besicovitch, On the density of certain sequences of integers. Math. Ann. **110**, 336–341 (1934)
2. F. Blanchard, E. Formenti, P. Kurka, Cellular automata in the Cantor, Besicovitch, and Weyl topological spaces, Complex Syst. **11**, 107–123 (1997)
3. K.B. Ford, The distribution of integers with a divisor in a given interval, Ann. Math. (2) **168**, 367–433 (2008)
4. H. Halberstam, H.-E. Richert, *Sieve Methods* (Academic Press, London, 1974)
5. H. Maier, in *Analytic Number Theory, Proceedings of a Conference in Honor of Heini Halberstam*, ed. by B. Berndt, H. Diamond, A.J. Hildebrand. The size of the the coefficients of cyclotomic polynomials, vol. 2 (Birkhäuser, Boston, 1996), pp. 633–639

Counting Primes in Arithmetic Progressions

John B. Friedlander

To Helmut Maier on the occasion of his 60th birthday

These notes represent an expanded version of a lecture delivered at the Urbana meeting of June 2014 in memory of Paul and Felice Bateman and of Heini Halberstam, and, in modified form, at the October 2014 workshop at the Royal Swedish Academy of Sciences, Stockholm, on the occasion of the presentation to Yitang Zhang of the 2014 Rolf Schock Prize in Mathematics for his ground-breaking work on bounded gaps between primes.

Both Zhang's work [31] and the earlier work of Goldston–Pintz–Yildirim [19], which had inspired it, kindled novel ideas on the use of sieve methods to produce prime numbers in interesting integer sequences and sparked renewed interest, if any reminder were needed, in the central part played in such developments by the distribution of primes in arithmetic progressions.

We take advantage of this renewed activity to survey here some of the few things we know (180 years after Dirichlet!) about that distribution and some of the things we expect to be true. Inter alia, we highlight a key role which has been played in parts of this story, based on ideas of Helmut Maier. It is a great pleasure to offer this work as a tribute to Helmut on the occasion of his 60th birthday.

1 Primes in an Individual Progression

Dirichlet [8] set the subject on its feet by proving the fundamental result:

Theorem 1.1. *Let q be a positive integer and a an integer relatively prime to q. Then there are infinitely many prime numbers p with*

$$p \equiv a \pmod{q} .$$

J.B. Friedlander (✉)
Department of Mathematics, University of Toronto, Toronto, ON, Canada M5S 2E4
e-mail: frdlndr@math.toronto.edu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_7

Shortly after the proof of the prime number theorem, de la Vallée–Poussin combined his techniques with those of Dirichlet to give an asymptotic formula for the number of primes up to a given point, in an arithmetic progression. Throughout, we shall find it convenient, following Tchebyshev, to count the primes with logarithmic weight. Thus we define, for $(a, q) = 1$,[1]

$$\theta(x; q, a) = \sum_{\substack{n \le x \\ n \equiv a (\bmod q)}} \theta(n),$$

where

$$\theta(n) = \begin{cases} \log p & \text{if } n = p \text{ is prime,} \\ 0 & \text{otherwise,} \end{cases}$$

and the expected error in the asymptotic formula,

$$E(x; q, a) = \theta(x; q, a) - \frac{x}{\varphi(q)},$$

where, as usual, $\varphi(q)$ is Euler's function, the number of reduced residue classes modulo q. As is to be expected, analogous statements apply throughout the paper to the similar prime counting functions $\psi(x; q, a)$ and $\pi(x; q, a)$.

De la Vallée–Poussin gave the asymptotic formula

$$\theta(x; q, a) \sim \frac{x}{\varphi(q)}, \tag{1}$$

with an error bound (actually, quite a bit stronger than)

$$E(x; q, a) \ll_{A,q} x\mathscr{L}^{-A}$$

for arbitrary A, where here and throughout, we abbreviate $\mathscr{L} = \log x$ and the subscript indicates that the constant implied in $\ll$ is allowed to depend only on A and q.

The dependence on A is unimportant but, for numerous applications, it is vital to obtain results which make explicit the dependence on q and, even more crucially, to show that the asymptotic formula (1) holds already with q not terribly small compared to x, and the larger the better. The best we know today is still the Siegel–Walfisz theorem [30, Hilfsatz 3],

$$E(x; q, a) \ll_A x\mathscr{L}^{-A}, \tag{2}$$

[1] This co-primality restriction is assumed throughout, though not always explicitly mentioned.

which looks the same as the previous bound until you notice that the dependence on q has been removed. This gives the asymptotic formula as long as $q \to \infty$ no faster than a fixed power of $\log x$. Although of fundamental value in many applications, this still leaves q too small for the result to be of use in a great many others.

However, we expect much more to be true. It follows from the generalized Riemann hypothesis (GRH), assumed for those Dirichlet L-functions formed with the characters of modulus q, that

$$E(x; q, a) \ll x^{\frac{1}{2}} \mathscr{L}^2 ,$$

which implies the asymptotic formula in the range, say $q < x^{\frac{1}{2}} \mathscr{L}^{-3}$.

Being more optimistic, by a conjecture of H. Montgomery which is widely believed, we expect

$$E(x; q, a) \ll_\varepsilon \sqrt{\frac{x}{q}}\, x^\varepsilon ,$$

which would give the asymptotic in the range $q < x^{1-\delta}$ for arbitrarily small positive δ.

On the other hand, one can go too far with such statements. Montgomery [24, 25] actually suggested the conjecture in the stronger form

$$E(x; q, a) \ll_\varepsilon \left(\frac{x}{q}\right)^{\frac{1}{2}+\varepsilon} \mathscr{L} ,$$

which would imply the asymptotic formula, for example, in the range $q < x\mathscr{L}^{-3}$.

This, however, cannot hold, even if -3 is replaced by any other number. Thus, as a special case of the results in [15], one knows that the asymptotic formula

$$\theta(x; q, 2) \sim \frac{x}{\varphi(q)}$$

fails for many odd q in the range $q < Q$ with, just for example, $Q = x\mathscr{L}^{-70{,}000{,}000}$. This result, as well as earlier ones [14, 17] in which the residue class a was required to grow with x, depends on a simple but beautiful idea [22] of H. Maier. We give a brief description of this in Sect. 4.

2 Approaching Twin Primes with the Sieve

Perhaps there is no chance to guess who might have been the first to make the twin prime conjecture:
There exist infinitely many primes p such that $p - 2$ is also prime.

It is natural to try to attack this problem using sieve methods. As a first try, consider the sequence of numbers

$$a_n = \begin{cases} 1 & \text{if } n = m(m-2) \text{ for some } m, \\ 0 & \text{otherwise.} \end{cases}$$

We would like to study the sum

$$S(x) = \sum_{n \leqslant x} a_n X_n$$

where

$$X_n = \begin{cases} 1 & \text{if } (n, P) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

with $P = \prod_{p<z} p$, hopefully with say $z \geqslant 2x^{\frac{1}{4}}$, which would imply that both m and $m-2$ are primes.

If we try the ancient sieve of Eratosthenes, we write

$$S(x) = \sum_n a_n \sum_{q|n} \mu(q) \tag{3}$$

with μ being the Möbius function. Following an interchange of the order of summation and an approximation of the resulting inner sum, our estimations founder on the fact that the Möbius function has too large a support, so that the variable q can range right up to x and, as a result, the errors in the approximation accumulate beyond control.

Brun got the idea to replace the Möbius function $\mu(q)$ by another function λ_q, similar in many respects but having support restricted, say to $q < Q$. This lost for him the possibility of obtaining an asymptotic formula for the sum $S(x)$, but the choice of possible sequences λ_q was sufficiently rich that he could find candidates which gave useful upper (and for z not too large, lower) bounds.

Proceeding, as with (3), we have

$$\sum_n a_n \sum_{q|n} \lambda_q = \sum_{q<Q} \lambda_q \sum_{\substack{n \leqslant x \\ n \equiv 0 (\bmod q)}} a_n \, .$$

Here, the inner sum counts integers m for which

$$m(m-2) \equiv 0 (\mathrm{mod} q) \, . \tag{4}$$

We can count these quite accurately for fairly large q (the larger, the better). As a result Brun [5] obtained, in particular, the following result.

Theorem 2.2. *There exists an integer K such that, for infinitely many m, each of m and $m-2$ has no more than K prime factors, multiplicity included.*

Indeed, Brun's method produces many such m; see, for example, Chap. 6 of [18] for some explicit bounds.

Now, let's consider a second approach. We define a new sequence a_n by

$$a_n = \begin{cases} \log p & \text{if } n = p-2 \text{ for some prime } p \leqslant x, \\ 0 & \text{otherwise.} \end{cases}$$

This has a major advantage in that we are beginning with the knowledge that at least one of the two integers is a prime. Proceeding as before, we have

$$\sum_n a_n \sum_{q|n} \lambda_q = \sum_{q<Q} \lambda_q \sum_{\substack{p \leqslant x \\ p \equiv 2 (\bmod q)}} \log p \,. \tag{5}$$

Now, we can see a major disadvantage to this approach. It is far more difficult to estimate the inner sum which, in our earlier notation, is none other than $\theta(x; q, 2)$.

There is yet a third possibility for the use of sieve methods in approaching the twin prime goal. Instead of insisting on integers differing by two and attempting to minimize the number of their prime factors, one may begin with more general clusters of integers potentially containing two primes and then seek to minimize the diameter of those that are found to succeed. Until the work [19], such an idea produced results which seemed too far from the goal to be a viable competitor in the approach to twin primes. That has now greatly changed. As with the second arrangement, this newer work rests heavily on the distribution of primes in arithmetic progressions and, in particular, on bounds of the type we are about to discuss.

3 Primes in an Arithmetic Progression, on Average

If, as before, we approximate $\theta(x; q, 2)$ by $x/\varphi(q)$ and if we assume that the "sieve weights" satisfy $|\lambda_q| \leqslant 1$ for all q, the effect of which turns out to be harmless, then we see that the error in the inner sum in (5) satisfies the bound

$$\text{error} \ll \sum_{q<Q} |E(x; q, 2)| \,. \tag{6}$$

Now, we are back to the problem of counting primes in arithmetic progressions to large moduli. But we also see that, for the present purpose, we don't need to know about primes in individual progressions, just on average over the moduli $q < Q$. We still want to be able to take the range of the moduli, in other words Q, as large as possible.

Actually, rather than (6), we really need only to bound the sum $\sum_{q<Q} \lambda_q E(x;q,2)$ which could be smaller, since, like $\mu(q)$, the λ_q can be expected to change sign quite a bit. For either of these sums, it would be enough to get a bound $\ll x\mathscr{L}^{-2}$ for this application and a bound $\ll_A x\mathscr{L}^{-A}$ would handle almost any application to problems which have been found to be approachable to date.

The first bound of this type is due to Renyi [27] who showed, with particular reference to the twin prime problem, that

$$\sum_{q\in\mathscr{Q}} |E(x;q,2)| \ll_\varepsilon x\mathscr{L}^{-3} \tag{7}$$

for a certain complicated looking but usable subset $\mathscr{Q}$ (related to the support of Brun sieve weights) of the integers $\{1 \leqslant q \leqslant Q\}$ with $Q = x^{\frac{1}{3}-\varepsilon}$. As a result, he became the first to derive the following theorem.

Theorem 3.3. *There exists an integer K such that, for infinitely many primes p, $p-2$ has no more than K prime factors, multiplicity included.*

Some fifteen years later, Barban [1], using a theorem of Linnik, improved Renyi's bounds of type (7) in a number of ways, in particular, extending the level to size $Q = x^{\frac{3}{8}-\varepsilon}$. Combining this with work of B. Levin, he thus achieved $K = 4$ as an acceptable choice. In 1965, Buchstab [6] used this result of Barban in more advantageous fashion to prove Renyi's theorem with the value of $K = 3$. Finally, J.-R. Chen announced in 1965, with detailed proofs [7] appearing in 1973, the lowering of the admissible value to $K = 2$.

By the time of Chen's work, he could take advantage of the following fundamental result, which in particular improves the Renyi and Barban exponents.

Bombieri-Vinogradov theorem. For arbitrary A, the bound

$$\sum_{q<Q} \max_{(a,q)=1} |E(x;q,a)| \ll_{A,\varepsilon} x\mathscr{L}^{-A} \tag{8}$$

holds with $Q = x^{\frac{1}{2}-\varepsilon}$.

In fact, Bombieri [2] gave this in the stronger form with $Q = x^{\frac{1}{2}}\mathscr{L}^{-B}$ with some $B = B(A)$.

The B–V theorem has been the source of a multitude of applications. Note that it confirms in a certain average sense the "level of distribution" $x^{\frac{1}{2}-\varepsilon}$ which, as mentioned in Sect. 1, would follow from the GRH. The theorem also quickly gave rise to conjectures corresponding to (and in fact pre-dating) the stronger conjectures of Montgomery.

Elliott-Halberstam conjecture. The bound (8) holds with $Q = x^{1-\varepsilon}$.

Actually, here too the bound was originally conjectured to hold with a level of the type $Q = x\mathscr{L}^{-B}$; see [14] or, better still, the original work [9].

In any case, the latter version was disproved in [14, 17], again based on Maier's idea, whereas the former has remained widely believed and has proved to be important, spawning a large number of consequences. One of the most striking is its implication in [19] that $\liminf\{p_{n+1} - p_n\} \leqslant 16$ (since improved to 12 by Maynard [23]).

The fact that these conjectures of Montgomery and of Elliott and Halberstam seem reasonable for moduli of size $x^{1-\varepsilon}$, but are demonstrably wrong for moduli of size $x\mathscr{L}^{-A}$, leads naturally to the following question. Quite apart from being able to prove it,

What do we think might be the truth?

The authors in [17] made the following

Hypothesis. Let $\varepsilon > 0$. For all $q > q_0(\varepsilon)$, $(a, q) = 1$, and $x > q(\log q)^{1+\varepsilon}$, we have

$$E(x; q, a) \ll \left(\frac{x}{q}\right)^{1-(1-\varepsilon)\delta(x,q)}, \tag{9}$$

where

$$\delta(x, q) = \min\left\{\frac{1}{2}, \frac{\log\left(\log \frac{x}{q} / \log_2 x\right)}{\log_2 x}\right\}. \tag{10}$$

Here, $\log_2$ stands for log log.

An analogous hypothesis is made about the sum $\sum_{q<Q} |E(x; q, a)|$.

The expression for δ in (10) is a little complicated to digest so we illustrate it in two important special cases:

(I) One expects the asymptotic formula $\theta(x; q, a) \sim x/\varphi(q)$ to hold so long as $q < x\mathscr{L}^{-A(x)}$ for any $A(x) \to \infty$ as $x \to \infty$.

(II) We expect the asymptotic formula to hold with square-root cancellation in the error, that is $E(x; q, a) \ll (x/q)^{\frac{1}{2}+\varepsilon}$, provided that $q < x \exp(-\sqrt{\mathscr{L}})$.

In [17] it is shown that the above hypothesis is false if one replaces $1-\varepsilon$ by $1+\varepsilon$ in (9). Perhaps it is dangerous to assume a result is best possible simply because one does not know how to improve it!

4 Maier's Idea

In the context of studying primes in arithmetic progressions, the Maier approach to irregularity results amounts to a consideration of the double sum

$$\sum_r \sum_s \theta(rP + sq) \tag{11}$$

as r runs through the integers in a dyadic interval $R/2 < r \leqslant R$ and s runs through the integers $1 \leqslant s \leqslant S$. Here, P is a parameter to be chosen and, for given r, the inner sum counts, with our usual logarithmic weights, the primes in a progression modulo q. Then summing over r, we obtain a counting over many progressions modulo q. Now, if for example q is a prime not dividing P or any r, then all of the residue classes modulo q are reduced classes. If that prime is large and if we assume a conjectured result for primes in progressions to such large moduli, then we can evaluate this total counting.

Next, consider what happens if we interchange the order of summation in (11), getting

$$\sum_s \sum_r \theta(rP + sq) \, . \tag{12}$$

Here, the inner sum is over primes in (a dyadic segment of) an arithmetic progression to modulus P. If the modulus P is not too large, then we know how to evaluate that sum and then to sum that over s. But the latter sum will involve a count, as s ranges, of the number of reduced residue classes modulo P. Now, however, we choose our parameter P (of convenient size, essentially $RP = Sq$ and) of the form

$$P = \prod_{p<z} p \, . \tag{13}$$

The number of reduced classes modulo P has, depending on the size of z relative to S, a well-known oscillatory behaviour, governed by the "Buchstab" function. By making a choice of the parameter z so that the number of classes is relatively large (respectively small) compared to its mean value, we produce a rigourous evaluation of (12) which gives a count too large (respectively too small) in comparison with the count conjectured for the same double sum in the form (11).

This explanation might lead one to ask what happens if the moduli q are much smaller, why doesn't the method still work? Actually, it does work but the results it then contradicts are estimations so accurate that nobody would have believed them anyway and indeed some can be more directly seen to be false.

5 Bilinear Forms and the BFI Conjectures

Using elementary ideas of Tchebyshev [28], it is possible to replace sums over prime numbers by more flexible bilinear forms. Begun in the work of I. Vinogradov [29] on the ternary Goldbach problem and then fashioned into a versatile tool by Linnik [21] in his development of the dispersion method, these ideas have since engendered quite a variety of useful elementary identities for the von-Mangoldt function (see Chap. 17 of [18] for a representative sampling).

What these identities have in common is that they replace $\theta(k)$ (or $\Lambda(k)$ or $\mu(k)$) by the sum of a few weighted divisor functions

$$\sum_{mn=k} \alpha_m \beta_n$$

so that we can make a transformation of sums

$$\sum_q |E(x;q,a)| \quad \to \quad \text{sums of type} \quad \sum_q |\Delta_{\alpha*\beta}(x;q,a)| \,,$$

where $*$ denotes Dirichlet convolution and

$$\Delta_f(x;q,a) = \sum_{\substack{n\leqslant x \\ n\equiv a(\bmod q)}} f(n) - \frac{1}{\varphi(q)} \sum_{\substack{n\leqslant x \\ (n,q)=1}} f(n) \,.$$

In practice, the sequences α_m, β_n can be taken to be bounded by some power of the divisor function $\tau(n)$ and, after a further subdivision, supported in boxes $M/2 < m \leqslant M$, $N/2 < n \leqslant N$. The only arithmetic information required is that one of them, say β_n, satisfy a Siegel–Walfisz axiom, that is

$$\Delta_\beta(N;q,a) \ll_A ||\beta|| \sqrt{N} (\log N)^{-A} \,, \tag{14}$$

where $||\dots||$ is the ℓ_2-norm.

One main advantage of a decomposition of this type is that it breaks a sum over primes into pieces which might be estimated by different methods, no one of which is suitable for all of them. As such, the method of bilinear forms has quite general capabilities. Herein, we are specifically interested in its use in proving bounds of B–V type. In order to prove such a theorem with $Q = x^{\eta-\varepsilon}$ for some $\eta > \frac{1}{2}$, it is sufficient to prove the same result for all relevant bilinear forms.

Thus, in [3] a number of such conjectures were proposed:

Conjecture I. Let $\varepsilon > 0$. Let α_m, β_n be sequences as above with $M = x^{1-\zeta}$, $N = x^\zeta$ where $\varepsilon < \zeta < 1 - \varepsilon$. Then, for arbitrary A we have

$$\sum_{\substack{q<Q \\ (q,a)=1}} \Delta_{\alpha*\beta}(x;q,a) \ll_{a,A,\varepsilon} ||\alpha||||\beta|| \sqrt{MN} \mathscr{L}^{-A} \tag{15}$$

with $Q = x\mathscr{L}^{-B}$ and some $B = B(A)$.

Actually, this is a special case of the conjecture, which was phrased a little more generally, in terms of ℓ_p-norms.

Some up-dating remarks about this are in order. Although Conjecture I has not yet been disproven, following the work described in the previous section, it would be

advisable to postulate the safer level $Q = x^{1-\varepsilon}$. Once one does this, it now becomes reasonable to also postulate the independence of the implied constant on a (which, at the time, was irrelevant to the work in [3]). This gives

Conjecture II. In the notation of the previous conjecture, we have

$$\sum_{q<Q} \max_{(a,q)=1} |\Delta_{\alpha*\beta}(x;q,a)| \ll_{A,\varepsilon} ||\alpha||||\beta||\sqrt{MN}\mathscr{L}^{-A}\,, \tag{16}$$

with $Q = x^{1-\varepsilon}$.

Also in [3], a still weaker conjecture was proposed.

Conjecture III. Same as Conjecture I, but with $Q = x^{\frac{3}{4}-\varepsilon}$.

Here, the point was that this weaker conjecture (after a use of Cauchy's inequality, which set this lower limit) had become something that was actually provable, modulo assumption of the expected bounds for exponential sums over certain rational functions in several variables.

It should be mentioned that, in very recent work, D. Polymath [26] has revived the use of conjectures very similar to Conjectures I and II, to which they refer as Generalized Elliott–Halberstam conjectures, and from which they derive the exciting consequence that $\liminf\{p_{n+1} - p_n\} \leqslant 6$. Thus, sharpening the earlier conditional bounds of GPY and of Maynard, they show that, not only do such conjectures, as has been known for quite a while, imply the Elliott–Halberstam conjecture, but apparently they can imply important consequences stronger than those implied by the Elliott–Halberstam conjecture on its own.

6 B–V Type Theorems with Special Weights

We have seen that, in applications, it is frequently unnecessary to have bounds for the sum $\sum_q |E(x;q,a)|$ but rather for weighted sums

$$\sum_{\substack{q<Q\\(q,a)=1}} \lambda_q E(x;q,a)\,, \tag{17}$$

where the coefficients λ_q may be more flexible than those given by the absolute value signs.

For example, if we have $\lambda = \gamma * \delta$ with γ_r supported on $R/2 < r \leqslant R$ and δ_s supported on $S/2 < s \leqslant S$ with $RS = Q$ (or if it can be written as a sum of not too many of these convolutions), then we are in a position to estimate along the lines of the previous section.

Already in the work of Renyi, the relevant sum was over a subset $\mathscr{Q}$ of moduli. Following the proof of the Bombieri–Vinogradov theorem, the search for levels of Q beyond the square-root barrier returned the focus to the use of flexible weighted sums, in particular the well-factorable weights of Iwaniec [20], of great importance for their use in applications of the linear sieve. For these weights, following an initial breakthrough (beyond the Riemann Hypothesis) in [13], the level $Q = x^{\frac{4}{7}-\varepsilon}$ was reached in [3]. For the problem at hand, this permits an upper bound for the number of twin primes up to a given point, one which sharpens by a factor $7/8$, the bounds which, following Selberg's work, had been achievable, at first only on assumption of GRH, and later unconditionally, after the availability of the Bombieri–Vinogradov theorem.

The Iwaniec well-factorable weights are a little complicated to be briefly described. We give here three further examples of bounds for sums (17) having weights of greater or lesser flexibility.

Example I. Although it is the most recent, we mention as our first example that of Yitang Zhang [31], one which is extremely flexible. We take, for z not too large, the very smooth numbers

$$\mathscr{Q} = \{q < Q; p|q \Rightarrow p < z\}$$

and define

$$\lambda_q = \begin{cases} \text{sgn}\left(E(x;q,a)\right) & \text{if } q \in \mathscr{Q} \\ 0 & \text{otherwise.} \end{cases}$$

For this choice of λ_q, Zhang was able to prove the bound

$$\sum_{q<Q} \lambda_q E(x;q,a) = \sum_{q\in\mathscr{Q}} |E(x;q,a)| \ll_A x\mathscr{L}^{-A},$$

uniformly in $a = a(q)$ of special arithmetic type and with $Q = x^{\frac{1}{2}+\eta}$ for a specific positive η. Because of this uniformity in a (any such η would have sufficed), he was able to derive a spectacular consequence, the first unconditional proof that $\liminf\{p_{n+1} - p_n\} < \infty$.

In the next example, we shall take a to be fixed. In other words, we allow the implied constants to depend on the residue class a.

Example II. It is natural to ask what level Q one can reach if one asks, in contrast to the previous example, for very inflexible weights. If no flexibility at all is allowed, such as with absolute values, then we still cannot take a fixed exponent $\eta > 0$ although, if a dependence on a is allowed, then [4] comes arbitrarily close.

As an example having some, but minimal, flexibility, we have the following special case of Theorem 8 in [3].

Theorem 6.4. *Let $a \neq 0$, $A > 0$, $\varepsilon > 0$, $\theta = \frac{1}{2} + \frac{1}{56} - \varepsilon$. Then,*

$$\sum_{p_1 < x^{\frac{8}{13}\theta}} \sum_{p_2 < x^{\frac{5}{13}\theta}} E(x; p_1 p_2, a) \ll_{a,A,\varepsilon} x\mathscr{L}^{-A}$$

where, as usual, p_1, p_2 run over primes.

Actually, the result holds uniformly in a up to a small power of x, depending on ε.

Example III. Our final example is the simplest that one could think of; we take λ_q to be identically equal to 1. This example, which has application to the Titchmarsh divisor problem, was independently studied in [3, 12].

What is of particular interest here is that we can successfully treat the sum $\sum_{q<Q} E(x; q, a)$ with $Q = x^{1-\varepsilon}$, due to the Dirichlet switching principle. In other words, if we consider the equation

$$p - a = qr\,, \quad p \leqslant x\,,$$

then the roles of q and r are symmetric and at least one of the two is $\leqslant x^{\frac{1}{2}}$. As a result, as soon as you can push a little beyond the Bombieri–Vinogradov level, you get to go quite a bit further.

The fact that we can succeed with such large q means that we can rigourously investigate the limit in the size of the level Q for which a bound

$$\sum_{\substack{q<Q \\ (q,a)=1}} E(x; q, a) \ll_A x\mathscr{L}^{-A} \tag{18}$$

can hold. We find that the answer depends on the residue class a, not so much its size but rather its arithmetic structure. Thus, we have the following result of [16] (which heavily uses [4]).

Theorem 6.5. *Fix $A > 1$, $M > 1$, $\varepsilon > 0$. For some B depending on A and M and for all large x, the bound (18) holds for every $Q < x\mathscr{L}^{-B}$ and every $0 < |a| < x$ satisfying*

$$\sum_{p|a} 1 < M \log_2 x$$

but, for every B and all large x, the bound (18) fails for some $Q < x\mathscr{L}^{-B}$ and some a, $0 < |a| < x$, satisfying

$$\sum_{p|a} 1 < (\log_2 x)^{\frac{6}{5}+\varepsilon}\,.$$

If the Riemann Hypothesis is assumed then, in the last exponent, $\frac{6}{5}$ can be replaced by 1.

Thus, under RH, the result is tight apart from the epsilon. By taking an $M > 1$, we see that it is the first case that applies to a typical residue class a.

Finally, it should be mentioned that, in connection with this last example, there are interesting further developments in recent work [10, 11] of Fiorilli.

Acknowledgements The author's research is partially supported by a University Professor Grant from the University of Toronto and by the Natural Sciences and Engineering Research Council of Canada through Research Grant A5123.

References

1. M.B. Barban, The "density" of the zeros of Dirichlet L-series and the problem of the sum of primes and "near primes". Mat. Sb. (N.S.) **61**, 418–425 (1963)
2. E. Bombieri, On the large sieve. Mathematika **12**, 201–225 (1965)
3. E. Bombieri, J.B. Friedlander, H. Iwaniec, Primes in arithmetic progressions to large moduli. Acta Math. **156**, 203–251 (1986)
4. E. Bombieri, J. B. Friedlander and H. Iwaniec, Primes in arithmetic progressions to large moduli III. J. Amer. Math. Soc. **2**, 215–224 (1989)
5. V. Brun, Über das goldbachsche Gesetz und die Anzahl der Primzahlpaare. Arch. Math. Naturvid. **B34**(8), 1–19 (1915)
6. A.A. Buchstab, New results in the investigation of the Goldbach-Euler problem and the problem of prime pairs. Dokl. Akad. Nauk SSSR **162**, 735–738 (1965). (Soviet Math. Dokl. **6**, 729–732 (1965))
7. 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**, 157–176 (1973)
8. G.L. Dirichlet, Beweiss des Satzes das jede arithmetische Progression deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält, Werke pp. 313–342, Kön. Preuss. Akad. der Wissen., Berlin, 1889. Reprinted Chelsea, New York, 1969
9. P.D.T.A. Elliott, H. Halberstam, A conjecture in prime number theory, in *Symposia Mathematica*, vol. **IV** (INDAM, Rome, 1968/69) (Academic Press, London, 1970), pp. 59–72
10. D. Fiorilli, Residue classes containing an unexpected number of primes. Duke Math. J. **161**, 2923–2943 (2012)
11. D. Fiorilli, The influence of the first term of an arithmetic progression. Proc. Lond. Math. Soc. **106**, 819–858 (2013)
12. E. Fouvry, Sur le problème des diviseurs de Titchmarsh. J. Reine Angew. Math. **357**, 51–76 (1985)
13. E. Fouvry, H. Iwaniec, Primes in arithmetic progressions. Acta Arith. **42**, 197–218 (1983)
14. J.B. Friedlander, A. Granville, Limitations to the equi-distribution of primes I. Ann. Math. (2) **129**, 363–382 (1989)
15. J.B. Friedlander, A. Granville, Limitations to the equi-distribution of primes III. Compos. Math. **81**, 19–32 (1992)
16. J.B. Friedlander, A. Granville, Relevance of the residue class to the abundance of primes, in *Proceedings of the Amalfi Conference (Maiori, 1989)* (Università di Salerno, Salerno, 1992), pp. 95–103
17. J.B. Friedlander, A. Granville, A. Hildebrand, H. Maier, Oscillation theorems for primes in arithmetic progressions and for sifting functions. J. Amer. Math. Soc. **4**, 25–86 (1991)
18. J.B. Friedlander, H. Iwaniec, Opera de cribro, in *American Mathematical Society Colloquium Publications*, vol. 57 (American Mathematical Society, Providence, RI, 2010)
19. D.A. Goldston, J. Pintz, C.Y. Yildirim, Primes in tuples I. Ann. Math. **170**(2), 819–862 (2009)

20. H. Iwaniec, A new form of the error term in the linear sieve. Acta Arith. **37**, 307–320 (1980)
21. Yu.V. Linnik, *The Dispersion Method in Binary Additive Problems* (American Mathematical Society, Providence, RI, 1963)
22. H. Maier, Primes in short intervals. Michigan Math. J. **32**, 221–225 (1985)
23. J. Maynard, Small gaps between primes. Ann. Math. **181**(2), 383–413 (2015)
24. H. Montgomery, in *Topics in Multiplicative Number Theory*. Lecture Notes in Mathematics, vol. 227 (Springer, Berlin/New York, 1971)
25. H. Montgomery, Problems concerning prime numbers. in *Mathematical Developments Arising from the Hilbert Problems, Northern Illinois University, De Kalb Illinois, 1974*. Proceedings of Symposia in Pure Mathematics, vol. XXVIII, pp. 307–310 (American Mathematical Society, Providence, RI, 1976)
26. D.H.J. Polymath, Variants of the Selberg sieve, and bounded intervals containing many primes. Res. Math. Sci. (2014) (to appear)
27. A. Renyi, On the representation of an even number as the sum of a single prime and single almost-prime number. Izvestiya Akad. Nauk SSSR. Ser. Mat. **12** (1948) 57–78. (Amer. Math. Soc. Transl. **19**(2), 299–321 (1962)
28. P.L. Tchebyshev, Sur la function qui détermine la totalité des nombres prémiers inferieurs à une limite donnée. J. Math. Pures et Appl. **17**, 366–390 (1852)
29. I.M. Vinogradov, *Selected Works* (Springer, Berlin, 1985)
30. A. Walfisz, Zur additiven Zahlentheorie II. Math. Zeit. **40**, 592–607 (1936)
31. Y. Zhang, Bounded gaps between primes. Ann. Math. (2) **179**, 1121–1174 (2014)

Limit Points of the Sequence of Normalized Differences Between Consecutive Prime Numbers

Daniel A. Goldston and Andrew H. Ledoan

Dedicated to Helmut Maier on the occasion of his 60th birthday

Abstract Let p_n denote the nth prime number and let $d_n = p_{n+1} - p_n$ denote the nth difference in the sequence of prime numbers. Erdős and Ricci independently proved that the set of limit points of $\frac{d_n}{\log p_n}$, the normalized differences between consecutive prime numbers, forms a set of positive Lebesgue measure. Hildebrand and Maier answered a question of Erdős and proved that the Lebesgue measure of the set of limit points of $\frac{d_n}{\log p_n}$ in the interval $[0, T]$ is $\gg T$ as $T \to \infty$. Currently, the only specific limit points known are 0 and ∞. In this note, we use the method of Erdős to obtain specific intervals within which a positive Lebesgue measure of limit points exist. For example, the intervals $\left[\frac{1}{8}, 2\right]$ and $\left[\frac{1}{40}, 1\right]$ both have a positive Lebesgue measure of limit points.

Keywords Hardy–Littlewood prime k-tuple conjecture • Limit points of normalized differences between consecutive prime numbers • Singular series

1 Introduction

In this note, we are concerned with the limit points of the sequence of normalized differences between consecutive prime numbers. Let p_n denote the nth prime number and let $d_n = p_{n+1} - p_n$ denote the nth difference in the sequence of prime numbers. By the prime number theorem, the average size of the differences d_n is $\log p_n$, and therefore it is interesting to examine the distribution of the limit points

D.A. Goldston (✉)
Department of Mathematics, San José State University, 315 MacQuarrie Hall, One Washington Square, San José, CA 95192-0103, USA
e-mail: daniel.goldston@sjsu.edu

A.H. Ledoan
Department of Mathematics, University of Tennessee at Chattanooga, 415 EMCS Building (Mail Stop 6956), 615 McCallie Avenue, Chattanooga, TN 37403-2598, USA
e-mail: andrew-ledoan@utc.edu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_8

of the normalized differences $\frac{d_n}{\log p_n}$. Since $\log p_n \sim \log n$ as $n \to \infty$, one can equally well consider the sequence $\frac{d_n}{\log n}$ which has the same set of limit points.[1]

It is difficult to find specific limit points for the sequence of normalized prime differences. In 1931 Westzynthius [12] proved that ∞ is a limit point, and it was only in 2005 that Goldston et al. [6] proved that 0 is also a limit point. We expect that every non-negative real number is a limit point, and this is a consequence of the Hardy–Littlewood prime tuple conjectures.

We are concerned here with a result of Erdős [5] and Ricci [10] who, in 1954, independently proved that the set of limit points of $\frac{d_n}{\log p_n}$ has positive Lebesgue measure. As we will see, the limit points obtained by this method are not large. For example, no finite limit point larger than 1 can be proven to exist by this method, and the largest Lebesgue measure for the set of limit points that one can ever hope to prove by this method is certainly less than 1. It was, therefore, a great breakthrough when in 1988 Hildebrand and Maier [8] proved, among other more general results, that the Lebesgue measure of the set of limit points of $\frac{d_n}{\log p_n}$ in the interval $[0, T]$ is $\gg T$ as $T \to \infty$. Their result implies that the sequence $\frac{d_n}{\log p_n}$ has arbitrarily large finite limit points, which more than settled a question of Erdős to prove that there is at least one finite limit point greater than 1.

In this note, we use the method of Erdős to obtain specific intervals in which there must be a positive Lebesgue measure of limit points. A representative example of our results is the following Corollary of our Theorem.

Corollary. *The intervals* $\left[\frac{1}{8}, 2\right]$ *and* $\left[\frac{1}{40}, 1\right]$ *each contain a positive Lesbegue measure of limit points of the sequence* $\frac{d_n}{\log p_n}$.

2 Erdős's Method

The only information required concerning prime numbers in Erdős's method is the prime number theorem and a sieve upper bound for pairs of prime numbers. The prime number theorem states that

$$\pi(x) := \sum_{p_n \le x} 1 \sim \frac{x}{\log x}, \quad \text{as } x \to \infty.$$

We will use the following two relationships that follow immediately from this:

$$\sum_{x < p_n \le 2x} 1 \sim \frac{x}{\log x}, \quad \text{as } x \to \infty, \tag{1}$$

[1] Erdős proved his result with this latter sequence, while Ricci used the former sequence.

and

$$\sum_{x<p_n\le 2x} d_n \sim x, \quad \text{as } x \to \infty. \tag{2}$$

Let p denote a prime number. The singular series $\mathfrak{S}(k)$ is defined for all integers $k \neq 0$ by

$$\mathfrak{S}(k) = \begin{cases} 2C_2 \prod\limits_{\substack{p|k \\ p>2}} \left(\dfrac{p-1}{p-2}\right), & \text{if } k \text{ is even;} \\ 0, & \text{if } k \text{ is odd;} \end{cases} \tag{3}$$

where

$$C_2 = \prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right) = 0.66016\ldots. \tag{4}$$

Then the upper bound estimate on prime pairs that we need is

$$\sum_{\substack{x<p\le 2x \\ p+k \text{ prime}}} 1 \le (\mathscr{C} + \epsilon)\mathfrak{S}(k)\frac{x}{(\log x)^2}, \tag{5}$$

which holds for any arbitrarily small positive constant ϵ and all sufficiently large x. Here, $\mathscr{C}$ is a constant.

The value $\mathscr{C} = 4$ in (5) is due to Bombieri and Davenport [2] (See also [7]). If $k \ll (\log x)^B$ for any given B, Bombieri et al. [3] obtained $\mathscr{C} = \frac{7}{2}$. Their result is only stated for the twin primes, but the above generalization holds. The best current results are due to Wu [13], who obtained the constant $\mathscr{C} = 3.91045$ uniformly for all k, and $\mathscr{C} = 3.3996$ if $k \ll (\log x)^B$ for any given B.

We will obtain our results from estimates for the consecutive prime gap counting function for prime numbers in the interval $(x, 2x]$ defined by

$$G(M) = G(x, M) = \sum_{\substack{x<p_n\le 2x \\ \frac{d_n}{\log p_n} \le M}} 1.$$

Further, we define

$$G^*(M) := \liminf_{x\to\infty} \left(\left(\frac{x}{\log x}\right)^{-1} G(x, M)\right).$$

From (5), we prove the following bound on normalized consecutive prime differences in an interval.

Lemma 2.1. *For fixed numbers $0 \le a < b$, and $\mathscr{C}$ a real number for which (5) holds, we have that*

$$G(b) - G(a) = \sum_{\substack{x < p_n \le 2x \\ a < \frac{d_n}{\log p_n} \le b}} 1 \le (\mathscr{C} + \epsilon)(b-a)\frac{x}{\log x} \tag{6}$$

holds for any arbitrarily small positive constant ϵ and all sufficiently large x.

Using Lemma 2.1 and a standard analysis argument, we obtain the following lower bound on the measure of the set of limit points.

Lemma 2.2 (Erdős [5]). *Let*

$$S(M) = \left\{ \frac{d_n}{\log p_n} : \frac{d_n}{\log p_n} \le M \right\}$$

and let

$$S^*(M) = \{c : c \text{ is a limit point of } S(M)\}.$$

Then we have

$$m(S^*(M)) \ge \frac{1}{\mathscr{C}} G^*(M),$$

where $m(E)$ denotes the Lebesgue measure of the set E.

We now need to find lower bounds for $G^*(M)$. Erdős found two methods for doing this. The simple method he used in [5] obtains a non-trivial lower bound for $M > 1$, while the method in [4] obtains a non-trivial bound for $M > 1 - \frac{1}{2\mathscr{C}}$.

Lemma 2.3. *Fix a real number $M > 1$. Then we have*

$$G^*(M) \ge 1 - \frac{1}{M}. \tag{7}$$

If $M > 1 - \frac{1}{2\mathscr{C}}$, then we have

$$G^*(M) \ge 1 - \frac{\mathscr{C} + \frac{1}{2}}{\mathscr{C} M + 1}. \tag{8}$$

The first lower bound (7) is better than (8) when $M > 2$. Combining Lemmas 2.2 and 2.3, we obtain the following Theorem.

Theorem. *Let a be any real number for which $a > \mathscr{C}$. Then each of the intervals*

$$\left[\frac{1}{a}\left(1 - \frac{1}{M}\right), M \right], \quad \text{for any } M \ge 2, \tag{9}$$

and

$$\left[\frac{1}{a}\left(1-\frac{a+\frac{1}{2}}{aM+1}\right),M\right], \quad \textit{for any } 1-\frac{1}{2\mathscr{C}}<M<2, \tag{10}$$

contains a positive Lesbegue measure of limit points of the sequence $\frac{d_n}{\log p_n}$.

Proof of the Theorem. Combining Lemma 2.2 and (7), we have

$$\begin{aligned} m\left(S^*(M)\cap\left[\frac{1}{a}\left(1-\frac{1}{M}\right),M\right]\right) &\geq m\left(S^*(M)\right)-m\left(\left[0,\frac{1}{a}\left(1-\frac{1}{M}\right)\right]\right)\\ &\geq \left(\frac{1}{\mathscr{C}}-\frac{1}{a}\right)\left(1-\frac{1}{M}\right)>0, \end{aligned}$$

which proves that (9) contains a positive Lebesgue measure of limit points. For (10), since

$$\frac{u+\frac{1}{2}}{uM+1}$$

is an increasing function of u for $u>0$ when $M<2$, we have for any $a>\mathscr{C}$ and $1-\frac{1}{2\mathscr{C}}<M<2$ that

$$G^*(M)\geq 1-\frac{a+\frac{1}{2}}{aM+1}.$$

The same argument just used for (9) with this lower bound for G^* shows (10) has a positive Lebesgue measure of limit points. □

The Corollary mentioned in the previous section is obtained by taking the value $a=4$ which is valid since we can use a value of $\mathscr{C}<4$ and take $M=2$ in (9) and $M=1$ in (10).

3 Proof of the Lemmas

In this section, the letter ϵ will denote a small positive number which may change each time it occurs.

Proof of Lemma 2.1. We make use of the sieve bound (5) together with the singular series average

$$T(N):=\sum_{k\leq N}\mathfrak{S}(k)\sim N, \quad \text{as } N\to\infty. \tag{11}$$

(See [2].) Let p' denote, here and henceforth, a prime number that is not necessarily adjacent to the prime number p. Then we have

$$\begin{aligned}\sum_{\substack{x<p_n\le 2x\\ a<\frac{d_n}{\log p_n}\le b}} 1 &\le \sum_{\substack{x<p\le 2x\\ a\log x<p'-p\le b\log 2x}} 1\\ &= \sum_{a\log x<k\le b\log 2x}\ \sum_{\substack{x<p\le 2x\\ p+k\ \text{prime}}} 1\\ &\le (\mathscr{C}+\epsilon)\frac{x}{(\log x)^2}\sum_{a\log x<k\le b\log 2x}\mathfrak{S}(k)\\ &\le (\mathscr{C}+\epsilon)(b-a)\frac{x}{\log x}.\end{aligned}$$

This completes the proof of Lemma 2.1. □

Proof of Lemma 2.2. Since $S^*(M)$ is a set of limit points in the interval $[0, M]$, it is a closed and bounded set of real numbers, and is therefore compact. We construct a sequence of open covers of $S^*(M)$, the kth one being defined by

$$\mathscr{O}_k = \left\{ I_c(k) = \left(c-\frac{1}{k}, c+\frac{1}{k}\right) : c\in S^*(M)\right\}.$$

Letting

$$Q_k = \bigcup_{c\in S^*(M)} I_c(k),$$

we have that

$$S^*(M) = \bigcap_{k=1}^{\infty} Q_k.$$

Since $\{Q_k\}_{k=1}^{\infty}$ is an infinite decreasing sequence of measurable sets, i.e., a sequence with $Q_{k+1}\subset Q_k$ for each k, and since $m(Q_1)\le M+2$, we have

$$m(S^*(M)) = m\left(\bigcap_{k=1}^{\infty} Q_k\right) = \lim_{k\to\infty} m(Q_k),$$

by a standard property of Lebesgue measure. (See [11], Proposition 14, pp. 62–63.)

Thus, given any $\epsilon>0$, we can find a sufficiently large number k_0 such that

$$m(S^*(M)) \ge m(Q_{k_0}) - \epsilon. \tag{12}$$

By the Heine–Borel theorem, the open covering $\mathscr{O}_{k_0}$ of $S^*(M)$ has a finite subcovering $\{I_{c_1}(k_0), \ldots, I_{c_r}(k_0)\}$ with $c_i \in S^*$, $1 \le i \le r$, and

$$S^*(M) \subset \bigcup_{i=1}^{r} I_{c_i}(k_0) \subset Q_{k_0}.$$

The open intervals $I_{c_i}(k_0)$, with $i = 1, \ldots, r$, may not be disjoint from each other. However, since the union of two overlapping open intervals is itself an open interval containing both of the original intervals, we may combine overlapping open intervals to obtain a smaller number of disjoint open intervals with the same union as the original intervals. As a result, we have obtained a finite collection $\{I_1, \ldots, I_s\}$, with $s \le r$, of disjoint open intervals such that

$$S^*(M) \subset \bigcup_{i=1}^{s} I_i \subset Q_{k_0}.$$

Letting $\ell(I)$ denote the length of the interval I, we have by (12) that

$$m(S^*(M)) \ge m\left(\bigcup_{i=1}^{s} I_i\right) - \epsilon = \sum_{i=1}^{s} \ell(I_i) - \epsilon.$$

By Lemma 2.1,

$$\ell(I_i) \ge \left((\mathscr{C} + \epsilon)\frac{x}{\log x}\right)^{-1} \sum_{\substack{x < p_n \le 2x \\ \frac{d_n}{\log p_n} \in I_i}} 1,$$

and since for all sufficiently large n each number $\frac{d_n}{\log p_n}$ lies in an interval I_i for some i, we have

$$m(S^*(M)) \ge \left((\mathscr{C} + \epsilon)\frac{x}{\log x}\right)^{-1} G(x, M) - \epsilon,$$

for x sufficiently large. Since the left-hand side of this inequality does not depend on x, we take the lim inf of the right-hand side and obtain Lemma 2.2. □

Proof of Lemma 2.3. Since

$$\begin{aligned} G(M) = \sum_{\substack{x < p_n \le 2x \\ \frac{d_n}{\log p_n} \le M}} 1 &\ge \sum_{\substack{x < p_n \le 2x \\ d_n \le M \log x}} \left(1 - \frac{d_n}{M \log x}\right) \\ &\ge \frac{1}{M \log x} \sum_{\substack{x < p_n \le 2x \\ d_n \le M \log x}} (M \log x - d_n) =: H(M), \end{aligned} \tag{13}$$

by (1) and (2) we have

$$\begin{aligned} H(M) &\geq \frac{1}{M\log x}\sum_{x<p_n\leq 2x}(M\log x - d_n) \\ &= \sum_{x<p_n\leq 2x} 1 - \frac{1}{M\log x}\sum_{x<p_n\leq 2x} d_n \qquad (14) \\ &\geq \left(1-\frac{1}{M}-\epsilon\right)\frac{x}{\log x}. \end{aligned}$$

Hence

$$\left(\frac{x}{\log x}\right)^{-1} G(M) \geq \left(1-\frac{1}{M}-\epsilon\right),$$

for x sufficiently large, and (7) follows.

We can improve on this argument when M is close to 1 by using the sieve bound (5) which prevents the lengths of all of the consecutive prime differences from being concentrated in a very small interval.

Letting $\lambda > 1$ be a fixed number that will be chosen later in terms of M and $\mathscr{C}$, we have by (14) that

$$H(\lambda) \geq \left(1-\frac{1}{\lambda}-\epsilon\right)\frac{x}{\log x},$$

for x sufficiently large. Now taking $M < \lambda$, we have

$$\begin{aligned} H(\lambda) &= \frac{1}{\lambda\log x}\left(\sum_{\substack{x<p_n\leq 2x\\ d_n\leq M\log x}}(\lambda\log x - d_n) + \sum_{\substack{x<p_n\leq 2x\\ M\log x<d_n\leq\lambda\log x}}(\lambda\log x - d_n)\right) \\ &\leq G(M) + \sum_{\substack{x<p\leq 2x\\ M\log x<p'-p\leq\lambda\log x}}\left(1-\frac{p'-p}{\lambda\log x}\right) \\ &= G(M) + H_1(\lambda, M). \end{aligned}$$

We will show below that

$$H_1(\lambda, M) \leq (\mathscr{C}+\epsilon)\frac{(\lambda-M)^2}{2\lambda}\left(\frac{x}{\log x}\right),$$

and therefore on combining these last three equations we have

$$\left(\frac{x}{\log x}\right)^{-1} G(M) \geq 1 - \frac{1}{\lambda} - \epsilon - (\mathscr{C} + \epsilon)\frac{(\lambda - M)^2}{2\lambda}$$
$$\geq \frac{1}{\lambda}\left(\lambda - 1 - \frac{\mathscr{C}}{2}(\lambda - M)^2\right) - \epsilon.$$

The choice $\lambda = M + \frac{1}{\mathscr{C}}$ maximizes the expression in the parentheses, and with this choice we obtain

$$\left(\frac{x}{\log x}\right)^{-1} G(M) \geq 1 - \frac{\mathscr{C} + \frac{1}{2}}{\mathscr{C} M + 1} - \epsilon,$$

and (8) follows.

To prove the bound for H_1, we see that the same calculation used in the proof of Lemma 2.1 gives

$$H_1(\lambda, M) \leq (\mathscr{C} + \epsilon)\frac{x}{(\log x)^2} \sum_{M \log x < k \leq \lambda \log x} \left(1 - \frac{k}{\lambda \log x}\right) \mathfrak{S}(k)$$
$$= (\mathscr{C} + \epsilon) T_1(\lambda \log x, M \log x)\frac{x}{(\log x)^2},$$

where

$$T_1(W, V) = \sum_{V < k \leq W} \left(1 - \frac{k}{W}\right) \mathfrak{S}(k),$$

and $V = V(x)$, $W = W(x)$, and $W \to \infty$ as $x \to \infty$. By partial summation, we have using (11) that

$$T_1(W, V) = \int_V^W \left(1 - \frac{u}{W}\right) dT(u)$$
$$= \frac{1}{W}\int_V^W (W - u)\, du + o(W)$$
$$= \frac{1}{2W}(W - V)^2 + o(W),$$

and the bound for H_1 follows. This finishes the proof of Lemma 2.3. □

4 Further Results and Comments

Recently, Pintz [9] proved that there is an ineffective small positive constant c for which the interval $[0, c]$ consists entirely of the limit points of $\frac{d_n}{\log n}$. His proof makes use of Zhang's recent breakthrough [14] on the existence of bounded differences between prime numbers. Banks et al. [1] have used the even more recent methods of Maynard and Tao to prove that for any set of real numbers $0 \leq \lambda_1 < \lambda_2 < \ldots < \lambda_k$ at least one of the differences $\lambda_j - \lambda_i$, with $1 \leq i < j \leq k$, is a limit point of $\frac{d_n}{\log p_n}$ provided $k \geq 9$. This implies far stronger results than anything in this paper. In particular, they obtain that at least 12.5 % of all the positive real numbers are limit points of the sequence of normalized consecutive prime differences. With regard to Erdős's question of limit points greater than 1, they obtain that such limit points must exist in the interval $[1, 8 + \eta]$, for any $\eta > 0$.

It has been pointed out to us by Julian Ziegler Hunts that one can generalize the results of the current paper to examine limit points of $\frac{p_{n+r}-p_n}{\log p_n}$, in which case the intervals obtained should be unattainable by the method of Banks et al. when r is large enough.

Acknowledgements The first author received support from the National Science Foundation Grant DMS-1104434.

References

1. W.D. Banks, T. Freiberg, J. Maynard, On limit points of the sequence of normalized prime gaps. Submitted for publication (2014). [Available at http://arxiv.org/pdf/1404.5094v2.pdf]
2. E. Bombieri, H. Davenport, Small differences between prime numbers. Proc. Roy. Soc. Ser. A **293**, 1–18 (1966)
3. E. Bombieri, E., J.B. Friedlander, H. Iwaniec, Primes in arithmetic progressions to large moduli. Acta Math. **156**(3–4), 203–251 (1986)
4. P. Erdős, The difference of consecutive primes. Duke Math. J. **6**, 438–441 (1940)
5. P. Erdős, Some problems on the distribution of prime numbers. Teoria dei numeri, Mathematical Congress Varenna, 1–8 (1954). [Reprinted as pp. 79–88 in *Teoria dei numeri*, ed. by G. Ricci. C.I.M.E. Summer Schools, vol. 5 (Springer, Berlin/Heidelberg, 2011)]
6. D.A. Goldston, J. Pintz, C.Y. Yıldırım, Primes in tuples, I. Ann. Math. (2) **170**(2), 819–862 (2009)
7. H. Halberstam, H.-E. Richert, in *Sieve Methods*. London Mathematical Society Monographs, vol. 4 (Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London/New York, 1974)
8. A. Hildebrand, H. Maier, Gaps between prime numbers. Proc. Am. Math. Soc. **104**(1), 1–9 (1988)
9. J. Pintz, Polignac numbers, conjectures of Erdős on gaps between primes, arithmetic progressions in primes, and the bounded gap conjecture. Submitted for publication (2013). [Available at http://arxiv.org/pdf/1305.6289.pdf]
10. G. Ricci, Sull'andamento della differenza di numeri primi consecutivi. Riv. Mat. Univ. Parma Vol. (1) **5**, 3–54 (1954)
11. H.L. Royden, *Real Analysis*, 3rd edn. (Macmillan, New York, 1988)

12. E.J. Westzynthius, Üeber die Verteilung der Zahlen, die zu den *n* ersten Primzahlen teilerfremd sind. Comm. Phys. Math. Soc. Sci. Fenn., Helsingfors, Vol. (5) **25**, 1–37 (1931)
13. J. Wu, Chen's double sieve, Goldbach's conjecture and the twin prime problem. Acta Arith. **114**, 215–273 (2004)
14. Y. Zhang, Bounded gaps between primes. Ann. Math. (3) **179**, 1121–1174 (2014)

Spirals of the Zeta Function I

Steven M. Gonek and Hugh L. Montgomery

To Professor Helmut Maier on his 60th birthday

Abstract Assuming RH, it is shown that the curve $\zeta(1/2 + it)$ spirals in the clockwise direction for all sufficiently large t, in the sense that its curvature is negative.

1 Introduction

It is well known that $\arg \zeta(1/2 + it)$ increases for $0 \le t \le t_1 = 6.289836$, and that it is decreasing for $t \ge t_1$, apart from jump discontinuities of height $m\pi$ at t if there are precisely $m > 0$ zeros $\beta + i\gamma$ (counting multiplicity) of $\zeta(s)$ with $\gamma = t$. This behavior depends on the Hadamard product and functional equation of the zeta function, but is independent of its Euler product, and of the Riemann Hypothesis (RH). It may be similarly noted that the curve $\zeta(1/2 + it)$ turns in the counterclockwise sense (i.e., $\kappa(t) > 0$) for $0 \le t \le t_0 = 2.7564883864$, and it seems that the curve turns in the clockwise direction ($\kappa(t) < 0$) for $t \ge t_0$. We confirm this, assuming RH (Fig. 1).

Theorem 1.1. *Assume RH, and let $\kappa(t)$ denote the curvature of the parameterized curve $\zeta(1/2 + it)$. There is a $t_2 > 0$ such that the curve turns in the clockwise direction for all $t > t_2$, which is to say that $\kappa(t) < 0$ for all $t > t_2$.*

It might still be the case that $\kappa(t) < 0$ for all large t even if RH is false. Whether or not depends a great deal on the positions of zeros off the critical line. It is therefore instructive to examine the issue for a function that has a functional equation similar to that of the zeta function, but has no Euler product. Titchmarsh [4, Sect. 10.25] defined a function $f(s)$ to be a linear combination of two quartic L-functions modulo

S.M. Gonek (✉)
Department of Mathematics, University of Rochester, Rochester, NY 14627, USA
e-mail: gonek@math.rochester.edu

H.L. Montgomery
Department of Mathematics, University of Michigan, Ann Arbor, MI 48109–1043, USA
e-mail: hlm@umich.edu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_9

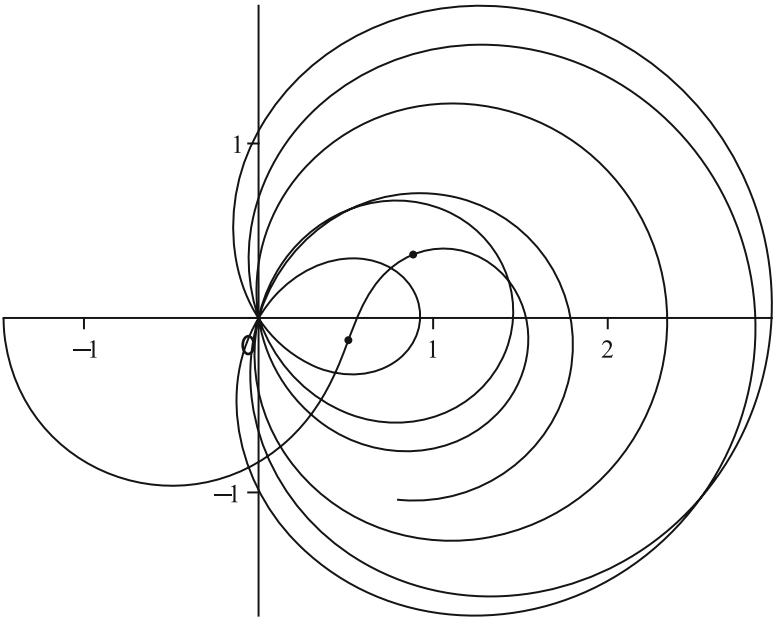

Fig. 1 The curve $\zeta(1/2+it)$ for $0 \le t \le 40$, with $\zeta(1/2+it_0)$ and $\zeta(1/2+it_1)$

5, and showed that it has zeros with real part > 1, although it satisfies a functional equation. As was first discovered by Spira [3], the first zero of $f(s)$ off the critical line is at $\varrho = 0.1914828 + i85.69934849$, together with $1-\overline{\varrho}$, $\overline{\varrho}$, and $1-\varrho$. Levinson and Montgomery [1] showed that under suitable circumstances zeros $\beta + i\gamma$ of $\zeta(s)$ with $0 < \beta < 1/2$ are in one-to-one correspondence with zeros $\beta_1 + i\gamma_1$ of $\zeta'(s)$ with $0 < \beta_1 < 1/2$. The same reasoning applies to $f(s)$, so $f'(s)$ has exactly one zero $\beta_1 + i\gamma_1$ in the rectangle $-1 \le \beta_1 < 1/2$, $2 \le \gamma_1 \le 100$, namely $\varrho_1 = 0.4053715 + i85.70513603$. Let $\kappa_f(t)$ denote the curvature of the curve $f(1/2+it)$. Then $\kappa_f(t)$ is generally negative, but this zero ϱ_1 just to the left of the critical line causes $\kappa_f(t)$ to be positive in the interval $[85.559071, 85.849586]$. In particular, $\kappa_f(\gamma_1) = 9.856$. The curve $f(1/2+it)$ for t near this height is plotted in Fig. 2. Numerical studies reveal that $f(s)$ has exactly 252 zeros with $\beta < 1/2$ and $0 < \gamma \le 5000$. Each one is accompanied by a nearby zero ϱ_1 of $f'(s)$, and $\kappa_f(\gamma_1) > 0$ for all but 10 of these γ_1. The first example of a zero ϱ with $\beta < 1/2$ for which the curvature remains negative occurs when $\varrho = 0.1022498 + i2442.530292$ with associated $\varrho_1 = 0.2664654 + i2442.535402$ for which $\kappa_f(\gamma_1) = -0.232836$. Moreover, $\kappa_f(t) < 0$ for $2425.05 \le t \le 2452.43$. Assuming that the one-to-one association of ϱ with ϱ_1 can be continued, let $N_{f-}(T)$ denote the number of zeros ϱ of $f(s)$ for which $\beta < 1/2$, $0 < \gamma \le T$, and $\kappa_f(\gamma_1) < 0$. Similarly, let $N_{f+}(T)$ denote the number of zeros ϱ of $f(s)$ for which $\beta < 1/2$, $0 < \gamma \le T$, and $\kappa_f(\gamma_1) > 0$. Not only do we conjecture that both $N_{f-}(T)$ and $N_{f+}(T)$ tend to infinity with T, but also that $N_{f-}(T) \gg T$ and

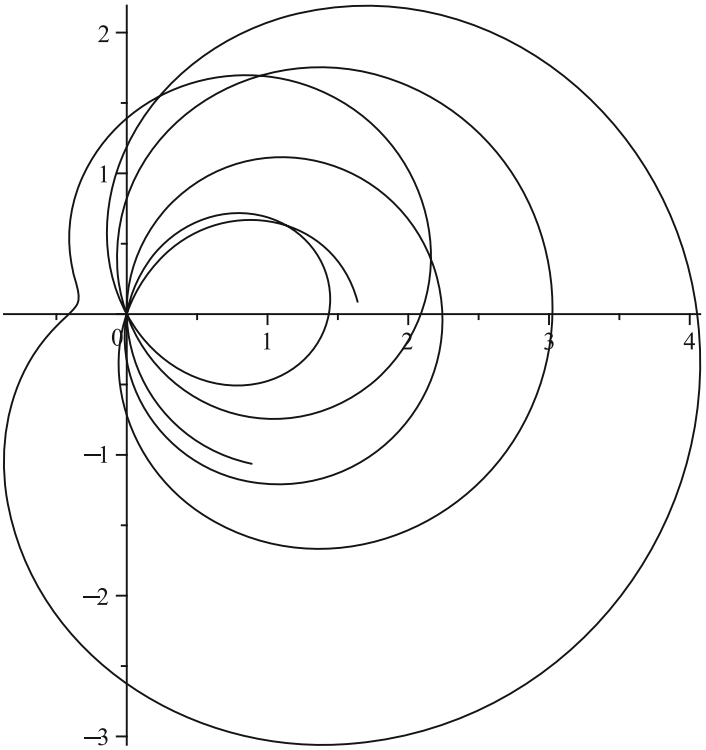

Fig. 2 The curve $f(1/2 + it)$ for $80 \le t \le 90$

$N_{f+}(T) \gg T$ for all large T. Indeed, we find it plausible that $N_{f-}(T)/T \to \infty$ and $N_{f+}(T)/T \to \infty$ as $T \to \infty$. However, we offer no speculation as to the relative asymptotic sizes of $N_{f-}(T)$ and $N_{f+}(T)$.

2 Proof of the Theorem

We show first that the curvature of the curve $\zeta(1/2 + it)$ at height t in the complex plane is

$$\kappa(t) = \frac{\Re \frac{\zeta''}{\zeta'}(1/2 + it)}{|\zeta'(1/2 + it)|} \,. \tag{1}$$

To this end we write the tangent vector $i\zeta'(1/2+it)$ in polar coordinates, so that $i\zeta'(1/2+it) = Re^{i\alpha}$. Here

$$\alpha = \frac{\pi}{2} + \arg\zeta'(1/2+it) = \frac{\pi}{2} + \Im\log\zeta'(1/2+it)\,.$$

By the Cauchy–Riemann equations it follows that

$$\frac{d\alpha}{dt} = \Re\frac{\zeta''}{\zeta'}(1/2+it)\,. \tag{2}$$

With s denoting arc-length, we have

$$\frac{ds}{dt} = R = |\zeta'(1/2+it)|\,. \tag{3}$$

Since

$$\kappa(t) = \frac{d\alpha}{ds} = \frac{\dfrac{d\alpha}{dt}}{\dfrac{ds}{dt}}, \tag{4}$$

the relation (1) follows from (2) and (3).

Levinson and Montgomery [1], assuming RH, showed that $\Re\frac{\zeta'}{\zeta}(s) < 0$ for s in various regions, and thus gave a new proof of Speiser's theorem [2] to the effect that $\zeta'(s) \neq 0$ for $0 < \sigma < 1/2$ if and only if RH is true. In the same way that it is natural to express $\frac{\zeta'}{\zeta}(s)$ in terms of a sum over zeros of $\zeta(s)$, one may write $\frac{\zeta''}{\zeta'}(s)$ in terms of a sum over zeros of $\zeta'(s)$. In this way, Yıldırım [5] showed, assuming RH, that $\zeta''(s) \neq 0$ for $0 < \sigma < 1/2$. In particular, Yıldırım showed, assuming RH, that

$$\Re\frac{\zeta''}{\zeta'}(1/2+it) < 0 \tag{5}$$

for $t \geq 100$. Since it is an easy matter (using, e.g., Maple or Mathematica) to plot $\kappa(t)$ for $2 \leq t \leq 100$, it follows that $\kappa(t) < 0$ for $t > t_0$ where t_0 is defined to be the root of $\Re\frac{\zeta''}{\zeta'}(1/2+it) = 0$ near 2.75.

Acknowledgements The authors supported in part by NSF grants DMS-1200582 and DMS-063529.

References

1. N. Levinson, H.L. Montgomery, Zeros of the derivatives of the Riemann zeta-function. Acta Math. **133**, 49–65 (1974)
2. A. Speiser, Geometrisches zur Riemannschen Zetafunktion. Math. Ann. **110**, 514–521 (1934)
3. R. Spira, Some zeros of the Titchmarsh counterexample. Math. Comp. **63**, 747–748 (1994)
4. E.C. Titchmarsh, *The Theory of the Riemann Zeta-function*, 2nd edn. (Oxford University Press, Oxford, 1986)
5. C. Yıldırım, A note on $\zeta''(s)$ and $\zeta'''(s)$. Proc. Am. Math. Soc. **124**, 2311–2314 (1996)

Best Possible Densities of Dickson m-Tuples, as a Consequence of Zhang–Maynard–Tao

Andrew Granville, Daniel M. Kane, Dimitris Koukoulopoulos, and Robert J. Lemke Oliver

Abstract We determine for what proportion of integers h one now knows that there are infinitely many prime pairs $p,\ p+h$ as a consequence of the Zhang–Maynard–Tao theorem. We consider the natural generalization of this to k-tuples of integers, and we determine the limit of what can be deduced assuming only the Zhang–Maynard–Tao theorem.

1 Introduction and Statement of Results

The twin prime conjecture states that there are infinitely many pairs of integers $(n, n+2)$ which are simultaneously primes. More generally, Hardy and Littlewood conjectured that each entry of the k-tuple $(n+h_1, \cdots, n+h_k)$ should be prime infinitely often, unless there is a trivial reason why this cannot happen. This "trivial reason" is about divisibility by small primes p. For example, the triplet $(n+2, n+4, n+6)$ can never all be prime if $n>1$ because at least one of them must be a multiple of 3. So we call a k-tuple *admissible* if, for each prime p, the reductions of the numbers $h_1, \ldots, h_k$ modulo p do not cover all of $\mathbb{Z}/p\mathbb{Z}$. With this definition in hand, the Hardy–Littlewood conjecture states that if $(h_1, \ldots, h_k)$ is an admissible k-tuple, then there are infinitely many integers n for which the numbers $n+h_1, \ldots, n+h_k$ are all prime.

Even though we are still far from proving the full Hardy–Littlewood conjecture, there has been remarkable progress made towards it recently. Firstly, in May 2013,

A. Granville • D. Koukoulopoulos
Département de mathématiques et de statistique, Université de Montréal, CP 6128 succ. Centre-Ville, Montréal, QC, Canada H3C 3J7
e-mail: andrew@dms.umontreal.ca; koukoulo@dms.umontreal.ca

D.M. Kane
Department of Mathematics, University of California-San Diego, 9500 Gilman Drive #0404, La Jolla, CA 92093, USA
e-mail: dakane@math.ucsd.edu

R.J. Lemke Oliver (✉)
Department of Mathematics, Stanford University, Building 380, Stanford, CA 94305, USA
e-mail: rjlo@stanford.edu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_10

Yitang Zhang [6] made headlines by proving that there are bounded gaps between primes, and specifically that

$$\liminf_{n\to\infty} p_{n+1} - p_n < 70{,}000{,}000,$$

where p_n denotes the n-th prime. Then, in November 2013, James Maynard [3] and Terence Tao independently showed, using somewhat different techniques, that 70,000,000 can be replaced by 600, and, as it stands right now, the best bound known is 246, due to the Polymath project [5]. Even more impressively, they proved that for any integer $m \geq 1$ there is an integer $k = k_m$ such that if $(h_1, \dots, h_k)$ is an admissible k-tuple, then there are infinitely many integers n for which at least m of the numbers $n + h_1, \dots, n + h_k$ are prime. Obviously, $k_m \geq m$, and in [5] it was shown that one can take $k_2 = 50$ and $k_m \ll e^{3.82m}$.

We call a k-tuple of integers $(h_1, \dots, h_k)$ a *Dickson k-tuple* if there are infinitely many integers n for which $n+h_1, \dots, n+h_k$ are each prime. The Hardy–Littlewood conjecture is equivalent to the statement that "*all admissible k-tuples of integers are Dickson k-tuples*", and the Maynard–Tao theorem implies that "*every admissible k_m-tuple of integers contains a Dickson m-tuple*". In particular, the Maynard–Tao theorem implies that Dickson m-tuples exist, yet no explicit example of a Dickson m-tuple is known! Nevertheless, a simple counting argument (which also appeared in [2]) yields the following attractive consequence of the Maynard–Tao theorem.

Corollary 1.1. *A positive proportion of m-tuples of integers are Dickson m-tuples.*

Proof. Let $k = k_m$, so that $m \leq k$, and define $R = \prod_{p\leq k} p$ and $x = NR$ for some (very large) integer N. We let

$$\mathscr{N} = \{n \in (-x, x] : (n, R) = 1\},$$

so that $|\mathscr{N}| = 2x\phi(R)/R$. Any subset of k elements of $\mathscr{N}$ is admissible, since it does not contain any integer $\equiv 0 \pmod p$ for any prime $p \leq k$. There are $\binom{|\mathscr{N}|}{k}$ such k-tuples. Each contains a Dickson m-tuple by the Maynard–Tao theorem.

Now suppose that there are $T(x)$ Dickson m-tuples within $\mathscr{N}$. Any such m-tuple is a subset of exactly $\binom{|\mathscr{N}|-m}{k-m}$ of the k-subsets of $\mathscr{N}$, and hence

$$T(x) \cdot \binom{|\mathscr{N}| - m}{k - m} \geq \binom{|\mathscr{N}|}{k},$$

and therefore

$$T(x) \geq \binom{|\mathscr{N}|}{k} \Big/ \binom{|\mathscr{N}| - m}{k - m} \geq (|\mathscr{N}|/k)^m = \left(\frac{\phi(R)}{kR}\right)^m \cdot (2x)^m,$$

as desired. □

The main goal of this paper is to provide better lower bounds on the proportion of m-tuples that are Dickson m-tuples. If $\Delta(m)$ is the proportion of such m-tuples, then Corollary 1.1 implies that $\Delta(m) > 0$. We are interested in determining the best possible lower bound on $\Delta(m)$ assuming only the results of Zhang, Maynard, and Tao and treating them as "black boxes": If $\mathscr{D}_m$ is the set of Dickson m-tuples, then there is an integer $k = k_m$ for which:

- $\mathscr{D}_m$ is translation-invariant, that is to say, if $(h_1, \dots, h_m) \in \mathscr{D}_m$ and $t \in \mathbb{Z}$, then $(h_1 + t, \dots, h_m + t) \in \mathscr{D}_m$;
- $\mathscr{D}_m$ is permutation-invariant, that is to say, if $(h_1, \dots, h_m) \in \mathscr{D}_m$ and $\sigma \in S_m$, then $(h_{\sigma(1)}, \dots, h_{\sigma(m)}) \in \mathscr{D}_m$;
- for any admissible k-tuple, $(x_1, \dots, x_k)$, there exist distinct $h_1, \dots, h_m \in \{x_1, \dots, x_k\}$ such that $(h_1, \dots, h_m) \in \mathscr{D}_m$.

We call any set $A \subset \mathbb{Z}^m$ that has the above properties (m, k)-*plausible*. Then we define

$$\delta(m,k) = \min\left\{ \liminf_{N\to\infty} \frac{|A \cap [-N,N]^m|}{(2N)^m} : A \subset \mathbb{Z}^m,\ A \text{ is } (m,k)-\text{plausible} \right\}; \tag{1}$$

and, for any $m \geq 1$, we have that

$$\Delta(m) \geq \delta(m, k_m)$$

by the Zhang–Maynard–Tao theorem. Our main result is the following.

Theorem 1.1. *For $k \geq m \geq 1$, we have, uniformly*

$$\delta(m,k) = \frac{(\log 2m)^{O(m)}}{(\frac{k}{m}\log\log\frac{3k}{m})^{m-1}}.$$

In the Maynard–Tao Theorem we know that one can obtain $k_m \leq e^{cm}$ for some constant $c > 0$. The best value known for c is a little smaller than 3.82, and Tao (http://terrytao.wordpress.com/, Public communication; see also [2]) showed that the Maynard–Tao technique cannot be (directly) used to obtain a constant smaller than 2. So we deduce that

$$\Delta(m) \geq \delta(m, k_m) \geq e^{-(c+o(1))m^2} \gg e^{-4m^2},$$

and that this can, at most, be improved to $\gg e^{-2m^2}$ using the currently available methods.

A *de Polignac number* is an integer h for which there are infinitely many pairs of primes p and $p + h$ (the twin prime conjecture is the special case $h = 2$). A positive proportion of integers are de Polignac numbers, as follows from both results above, but we wish to determine the best explicit lower bounds possible. In Sect. 2, using quite elementary arguments, we will show that

$$\Delta(2) \geq \delta(2,k_2) \geq \frac{1}{49}\prod_{p\leq 50}\left(1-\frac{1}{p}\right) \approx 0.002830695767\ldots > \frac{1}{354}, \tag{2}$$

and that this cannot be improved dramatically without improving the value of $k_2 = 50$. We remark that this bound is surprisingly good in light of the fact that we can only deduce that there is some $h \leq 246$ for which there are infinitely many prime pairs $p, p+h$.

Finally, we note that it would perhaps be more natural to consider the proportion $\Delta^{\mathrm{ad}}(m)$ of *admissible* m-tuples that are Dickson m-tuples, rather than the proportion of all m-tuples. To this end, in Sect. 4, we determine the proportion $\varrho_{\mathrm{ad}}(m)$ of all m-tuples that are admissible. Surprisingly, this question has seemingly not been addressed in the literature, and we prove that

$$\varrho_{\mathrm{ad}}(m) = \frac{e^{o(m)}}{(e^{\gamma}\log m)^m} \tag{3}$$

as $m \to \infty$.

2 The Density of de Polignac Numbers

To prove a lower bound on the density of de Polignac numbers, we consider admissible sets B of integers that do not contain a Dickson pair. If the prime k-tuplets conjecture is true, then necessarily $|B| = 1$, while Zhang's theorem implies that $|B| \leq k_2 - 1$. Because of this upper bound, there must be maximal such sets, in that, for any $t \notin B$ such that $B \cup \{t\}$ is admissible, $B \cup \{t\}$ contains a Dickson pair. This condition implies that $t - B$ contains a de Polignac number, and we will obtain a lower bound on $\delta(2,k_2)$ by varying t. To this end, for an admissible set B, define $\eta(B)$ to be the minimal lower density of sets A with the property that $t - B$ contains an element of A whenever $B \cup \{t\}$ is admissible. Moreover, for any integer ℓ, set $\eta(\ell)$ to be the infimum of $\eta(B)$ as B runs over admissible sets of size ℓ. We thus have that

$$\Delta(2) \geq \delta(2,k_2) \geq \min_{1\leq \ell \leq k_2-1} \eta(\ell),$$

and we will prove the following.

Proposition 2.1. *For any integer ℓ, we have that $\eta(\ell) \sim e^{-\gamma}/\ell\log\ell$, and explicitly that*

$$\frac{1}{\ell}\prod_{p\leq \ell+1}\left(1-\frac{1}{p}\right) \leq \eta(\ell) \leq \frac{1}{\ell-y}\prod_{p\leq y}\left(1-\frac{1}{p}\right)$$

for any positive integer $y \leq \ell - 1$.

For the particular application to $\Delta(2)$, we take $\ell \in \{1, \dots, 49\}$ in Proposition 2.1 to find that

$$\delta(2, k_2) \geq \min_{1 \leq \ell \leq 49} \eta(\ell) > 0.002830695767 > \frac{1}{354},$$

while taking $y = 13$ yields that $\eta(49) < 0.005328005328 < \frac{1}{187}$, so that, using our techniques, we can hope for an improvement in our lower bound for $\delta(2, k_2)$ by a factor of at most about two. However, we can do better: the explicit upper bound given in Lemma 3.1 below shows that $\delta(2, 50) \leq \frac{1}{210} < 0.0048$.

Proof of Proposition 2.1. We begin with the lower bound for $\eta(\ell)$. Suppose that B is admissible of size ℓ, so that for each prime $p \leq \ell + 1$ there exists a residue class $n_p \pmod p$ such that $p \nmid n_p + b$ for each $b \in B$. If $t \not\equiv -n_p \pmod p$ for all $p \leq \ell + 1$, then $B \cup \{t\}$ is admissible and so contains a Dickson pair, and, as remarked above, t must be one of that pair, else B would have contained a Dickson pair. If x is large, then the number of such integers $t \leq x$ is

$$\sim \prod_{p \leq \ell + 1} \left(1 - \frac{1}{p}\right) x$$

and we know that, for some $b \in B$, $t - b$ is a de Polignac number. Given $t \in \mathbb{Z}$, an integer can be written in at most ℓ ways as $t - b$ for $b \in B$, since $|B| = \ell$. Hence,

$$\eta(B) \geq \frac{1}{\ell} \prod_{p \leq \ell + 1} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\ell \log \ell}$$

as $\ell \to \infty$.

We now turn to the upper bound. We will construct sets A and B with the properties that B is admissible and that if t is such that $B \cup \{t\}$ is admissible, then there exists $b = b_t \in B$ such that $t - b_t \in A$.

Let $y \in \{1, \dots, \ell - 1\}$ and set $v = \ell - y$. Moreover, define $r = \prod_{p \leq y} p$ and $m = \prod_{p \leq \ell} p$, and select h large so that $q := hv + 1$ is a prime $> \ell$. We define A to be the set of all integers a for which $(a + 1, r) = 1$ and $a \equiv 0, 1, 2, \dots, \text{or } h - 1 \pmod q$. The density $\delta(A)$ of A is

$$\frac{h}{q} \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \leq \frac{1}{\ell - y} \prod_{p \leq y} \left(1 - \frac{1}{p}\right).$$

Our set B will be constructed as the union of two sets, B_1 and B_2. For B_1, we take $B_1 := \{b_0, b_1, \dots, b_v\}$ where each b_i is chosen to satisfy $b_i \equiv ih \pmod q$ and $b_i \equiv 1 \pmod m$, and we choose B_2 to be a set of $y - 1$ integers covering the residue classes $2, \dots, p - 1 \pmod p$ for each prime $p \leq y$. We note that $B := B_1 \cup B_2$ has ℓ elements and is admissible: it occupies the congruence classes $1, \dots, p - 1 \pmod p$ for all $p \leq y$ and covers at most y classes modulo p for $p > y$.

If t is such that $B\cup\{t\}$ is admissible, then $t \not\equiv 0 \pmod p$ for all primes $p \le y$. We can write $t \equiv ih+j \pmod q$ for some i and j satisfying $0 \le i \le v$ and $0 \le j \le h-1$. Consider $t-b_i$. This is $\equiv j \pmod q$ and $\not\equiv -1 \pmod p$ for all primes $p \le y$. Hence $t - b_i \in A$, so that $\eta(\ell) \ge \delta(A)$, and the result follows. □

Remark 2.1. Pintz [4] showed that there is an ineffective constant C such that every interval of length C contains a de Polignac number. The proof of the lower bound above furnishes a different proof of this result: the values of t that appear are periodic modulo $r := \prod_{p\le k_2} p$, so, given a maximal set B, we obtain Pintz's result with $C = r + \max_{b\in B} b - \min_{b\in B} b$.

3 Proof of Theorem 1.1

We begin with a strong version of the upper bound of Theorem 1.1.

Lemma 3.1. *If $k \ge m \ge 2$ with $k/m \to \infty$, then*

$$\delta(m,k) \le \left(\frac{e^{-\gamma}+o(1)}{\frac{k}{m-1}\log\log\frac{k}{m-1}}\right)^{m-1}.$$

Proof. Pick q as large as possible so that $(m-1)\phi(q) < k$. We claim that

$$A = \{(x_1,\dots,x_m) : x_1 \equiv x_2 \equiv \dots \equiv x_m \mod q\}$$

is (m,k)-plausible. This is because any admissible set with k elements may take at most $\phi(q)$ distinct values modulo q, so by the pigeonhole principle, at least m of these values must be congruent modulo q, giving an m-tuple contained in A. On the other hand, q is of size $\ge (e^\gamma + o(1))\frac{k}{m-1}\log\log\frac{k}{m-1}$ as $k/m \to \infty$, and A has density $1/q^{m-1}$. This completes the proof. □

The proof of the lower bound is somewhat more involved. A key ingredient is the Lovász Local Lemma:

Lovász Local Lemma. *Suppose that $E_1,\dots,E_n$ are events, each of which occurs with probability $\le p$ and depends on no more than d of the others. If $d \le 1/(4p)$, then the probability that no E_j occurs is at least e^{-2pn}.*

The input for the Lovász Local Lemma will come from the following technical result.

Lemma 3.2. *Let $k \ge m \ge 1$. Consider a translation and permutation invariant set $A \subset ((-2x,2x]\cap\mathbb{Z})^m$, and set $\mathcal{N} = \{n \in (-x,x]\cap\mathbb{Z} : (n, \prod_{p\le k} p) = 1\}$. If*

$$\frac{|A\cap\mathcal{N}^m|}{|\mathcal{N}^m|} > \frac{1}{8m\binom{k-1}{m-1}},$$

and x is large enough in terms of m and k, then

$$|A| \geq \frac{(\log 3m)^{O(m)}}{(\frac{k}{m}\log\log\frac{k}{m})^{m-1}} \cdot x^m.$$

Before we prove Lemma 3.2, we use it to deduce the lower bound in Theorem 1.1.

Proof of the lower bound in Theorem 1.1. Fix a large positive x and suppose that $A \subset [-2x, 2x]^m$ is an (m, k)-plausible set, yet

$$|A| \leq (\log 2m)^{-cm} x^m \Big/ \left(\frac{k}{m}\log\log\frac{k}{m}\right)^{m-1}$$

for some $c > 0$. If c is large enough, then Lemma 3.2 implies that

$$\frac{|A \cap \mathscr{N}^m|}{|\mathscr{N}^m|} \leq \frac{1}{8m\binom{k-1}{m-1}}, \tag{4}$$

where $\mathscr{N} := \{n \in (-x, x] \cap \mathbb{Z} : (n, \prod_{p\leq k} p) = 1\}$. Now select integers $n_1, \dots, n_k$ uniformly at random from $\mathscr{N}$. We claim that there is a positive probability that the n_i are distinct and that there is no m-element subset of $\{n_1, \dots, n_k\}$ in A, which implies that A is not (m, k)-plausible. We use the Lovász Local Lemma, in which we wish to avoid the events $n_i = n_j$ for $1 \leq i < j \leq k$ and $(x_1, \dots, x_m) \in A$ for any choice of m elements $x_1, \dots, x_m$ among the integers $n_1, \dots, n_k$. There are $\binom{k}{2} + \binom{k}{m}$ such events. Each event depends on no more than m of the n_i, and so depends on no more than $m\binom{k-1}{1} + m\binom{k-1}{m-1} \leq 2m\binom{k-1}{m-1}$ other events. If $x_1, \dots, x_m$ are m out of the k random variables, then the probability that $(x_1, \dots, x_m) \in A$ is $\leq 1/(8m\binom{k-1}{m-1})$ by (4). We finally note that the probability that $n_i = n_j$ is $1/|\mathscr{N}| \sim R/(2x\phi(R))$ as $x \to \infty$, which is certainly $\leq 1/(8m\binom{k-1}{m-1})$ for x sufficiently large.

The Lovász Local Lemma now implies that there exist distinct elements $n_1, \dots, n_k \in \mathscr{N}$ for which $(x_1, \dots, x_m) \notin A$ for all subsets $\{x_1, \dots, x_m\}$ of $\{n_1, \dots, n_k\}$. (In fact, the Lovász Local Lemma implies that this is true for a proportion $\geq e^{-k/2m^2}$ of the k-subsets of $\mathscr{N}$.) Hence, A is not an (m, k)-plausible set, a contradiction. □

To prove Lemma 3.2, we first need another result. Given an m-tuple $\mathbf{h} = (h_1, \dots, h_m)$, we denote by $n_p(\mathbf{h})$ the number of congruence classes mod p covered by $h_1, \dots, h_m$. Note that $1 \leq n_p(\mathbf{h}) \leq \min\{p, m\}$ and that both upper and lower bounds are easily obtained for some m-tuple $\mathbf{h}$.

Lemma 3.3. *Let $k \geq m \geq 1$. There is an absolute constant $c > 0$ such that if $\mathbf{h}$ is a randomly selected m-tuple from $\mathscr{N} = \{n \in (-x, x] \cap \mathbb{Z} : (n, \prod_{p\leq k} p) = 1\}$ and x is large enough in terms of m and k, then the probability that*

$$\prod_{p\leq k}\left(1 - \frac{n_p(\boldsymbol{h})}{p}\right)\left(1 - \frac{1}{p}\right)^{-m} > (\log 3m)^{cm}(\log\log k)^{m-1}$$

is

$$\le \frac{1}{16m\binom{k-1}{m-1}}.$$

Proof. The result is trivial if $k \le m^2$, so we will assume that $k > m^2$. Similarly, we may assume that k is large enough. Moreover, note that there are $\ll_k |\mathscr{N}|^{m-1} = o_{x\to\infty}(|\mathscr{N}|^m)$ m-tuples $(h_1, \dots, h_m)$ with non-distinct elements. So it suffices to show that if we randomly select a subset $B = \{b_1, \dots, b_m\}$ of distinct elements of $\mathscr{N}$, then the probability that

$$\prod_{p\le k}\left(1-\frac{n_p(B)}{p}\right)\left(1-\frac{1}{p}\right)^{-m} > (\log 3m)^{cm}(\log\log k)^{m-1}$$

is

$$\le \frac{1}{17m\binom{k-1}{m-1}}.$$

By Mertens' Theorem, the contribution of the small primes is

$$\prod_{p\le m^2}\left(1-\frac{n_p(B)}{p}\right)\left(1-\frac{1}{p}\right)^{-m} \le \prod_{p\le m^2}\left(1-\frac{1}{p}\right)^{-m} \le (\log 3m)^{c_1 m}$$

for some absolute constant $c_1 > 0$. If $c \ge 2c_1$ and we set

$$f(B) = \prod_{m^2<p\le k}\left(1-\frac{n_p(B)}{p}\right)\left(1-\frac{1}{p}\right)^{-m},$$

then we are left to show that

$$\mathbb{P}\Big(f(B) > (\log 3m)^{cm/2}(\log\log k)^{m-1}\Big) \le \frac{1}{17m\binom{k-1}{m-1}}.$$

Taking logarithms, we have

$$\begin{aligned}\log f(B) &\le \sum_{m^2<p\le k}\left(\frac{m-n_p}{p}+O\left(\frac{m^2}{p^2}\right)\right) \le \sum_{m^2<p\le k}\frac{m-n_p}{p}+O\left(\frac{1}{\log m}\right)\\ &\le (m-1)\sum_{m^2<p\le k}\frac{X_p(B)}{p}+O\left(\frac{1}{\log m}\right),\end{aligned}$$

where $X_p(B) = 1$ if $n_p(B) < m$, and $X_p(B) = 0$ if $n_p(B) = m$. Therefore, if we set $X(B) = \sum_{m^2<p\le k} X_p(B)/p$ and we assume that c is large enough, then it suffices to show that

$$\mathbb{P}\left(X(B) > \frac{c}{3}\log\log(3m) + \log\log\log k\right) \le \frac{1}{17m\binom{k-1}{m-1}}.$$

In turn, the above inequality is a consequence of the weaker estimate

$$\mathbb{P}\left(X(B) > \log\log(\max\{m^2, m\log k\}) + c'\right) \le \frac{1}{17m\binom{k-1}{m-1}}, \tag{5}$$

where c' is a sufficiently large constant.

The X_p can be viewed as independent random variables as we run over all possible sets B. As in the birthday paradox, the probability that $X_p = 0$ is

$$\left(1-\frac{1}{p-1}\right)\left(1-\frac{2}{p-1}\right)\cdots\left(1-\frac{m-1}{p-1}\right) = \exp\left(-\frac{m^2+O(m)}{2p}\right)$$
$$= 1 + O\left(\frac{m^2}{p}\right).$$

For any r, if $p \le r$, then we trivially have that $\mathbb{E}[e^{rX_p/p}] \le e^{r/p}$, and otherwise

$$\mathbb{E}[e^{rX_p/p}] = \mathbb{P}(X_p = 0) + \mathbb{P}(X_p = 1)e^{r/p} = 1 + \mathbb{P}(X_p = 1)(e^{r/p} - 1)$$
$$\le \exp\left(O\left(\frac{m^2 r}{p^2}\right)\right).$$

Therefore, for any values of s and r, we have that

$$e^{rs}\mathbb{P}(X \ge s) \le \mathbb{E}[e^{rX}] = \prod_{p\le k}\mathbb{E}[e^{rX_p/p}] \le \prod_{p\le r} e^{r/p}\prod_{p\ge r} e^{O(m^2 r/p^2)}$$
$$\le \exp\left(r\log\log r + O(r) + O(m^2/\log r)\right).$$

Thus, if $r \ge m^2$, then setting $s = \log\log r + c'$ for c' sufficiently large, we find that

$$\mathbb{P}\left(X \ge \log\log r + c'\right) \le e^{-r}.$$

Substituting $r = \max\{m^2, m\log k\}$ establishes (5) for k large enough, thus completing the proof of the lemma. □

Proof of Lemma 3.2. We partition the elements of $[-2x, 2x]^m$ into *translation classes*, putting two elements in the same class if and only if they differ by $(\ell, \ell, \ldots, \ell)$ for some integer ℓ. Each translation class T intersecting $[-x, x]^m$

contains at least $2x$ elements of $[-2x, 2x]^m$ (and at most $6x$). The main idea of the proof is that if we can find at least U elements of $A \cap \mathscr{N}^m$ whose translation class has at most M elements inside $\mathscr{N}^m$, then $|A| \geq 2x \cdot U/M$.

Note that $n_p(\boldsymbol{h})$ is fixed over all $\boldsymbol{h} \in T$, so we denote it by $n_p(T)$. The number of integers ℓ for which $\mathbf{h} + (\ell, \ell, \dots, \ell) \in \mathscr{N}^m$ is $\leq 3x \prod_{p \leq k} \left(1 - \frac{n_p(T)}{p}\right)$ for x large enough. If we set $R = \prod_{p \leq k} p$, then Lemma 3.3 implies that the proportion of elements of $\mathscr{N}^m$ for which

$$\prod_{p \leq k} \left(1 - \frac{n_p(B)}{p}\right) \leq \left(\frac{\phi(R)}{R}\right)^m (\log 3m)^{cm} (\log\log k)^{m-1} \tag{6}$$

is $\leq 1/16m\binom{k-1}{m-1}$, provided that x is large enough, with c being an absolute constant. Since A contains at least $|\mathscr{N}^m|/8m\binom{k-1}{m-1}$ elements in $\mathscr{N}^m$ by assumption, we find that A contains at least $|\mathscr{N}^m|/16m\binom{k-1}{m-1}$ for which (6) holds. Therefore, we conclude that

$$\begin{aligned} |A| &\geq \frac{|\mathscr{N}^m|}{24m\binom{k-1}{m-1}} \Big/ \left\{ \left(\frac{\phi(R)}{R}\right)^m (\log 3m)^{cm} (\log\log k)^{m-1} \right\} \\ &> \frac{x^m (\log 3m)^{O(m)}}{(\frac{k}{m} \log\log \frac{k}{m})^{m-1}} \end{aligned}$$

for x large enough, and the result follows. □

4 The Number Of Admissible k-Tuples

Our goal in this section is to show relation (3). Given a prime p, we say that an m-tuple is *admissible mod p* if its elements do not occupy all of the residue classes mod p, so an m-tuple is admissible if and only if it is admissible mod p for every prime p. By the pigeonhole principle, any set of m integers is admissible mod p if $p > m$, so to test for admissibility we need only work with the primes $p \leq m$. This implies, using the Chinese Remainder Theorem, that

$$\varrho_{\mathrm{ad}}(m) = \prod_{p \leq m} \varrho_{\mathrm{ad}}(m, p),$$

where $\varrho_{\mathrm{ad}}(m, p)$ denotes the proportion of m-tuples that are admissible mod p.

If $m/\log m < p \leq m$, then we note the trivial bounds $(1-1/p)^m \leq \varrho_{\mathrm{ad}}(m, p) \leq 1$, with the lower bound coming from counting m-tuples whose elements are not 0 mod p. Therefore

$$1 \geq \prod_{m/\log m < p \leq m} \varrho_{\mathrm{ad}}(m, p) \geq \prod_{m/\log m < p \leq m} \left(1 - \frac{1}{p}\right)^m = e^{O(\frac{m \log\log m}{\log m})}.$$

It remains to compute the contribution of primes $p \le m/\log m$. It is not difficult to determine an exact expression for $\varrho_{\mathrm{ad}}(m,p)$ using an inclusion–exclusion argument: the probability that the elements of an m-tuple $\boldsymbol{h}$ belong to a given subset of $p-1$ residue classes is $(1-\frac{1}{p})^m$. There are $\binom{p}{1}$ choices of the $p-1$ residue classes. If the elements of $\boldsymbol{h}$ belong to exactly $p-2$ residue classes mod p, then $\boldsymbol{h}$ was just counted twice and so we need to subtract the probability of this happening. That probability is $(1-\frac{2}{p})^k$, and there are $\binom{p}{2}$ choices of the $p-2$ residue classes. Continuing in the way, we find that

$$\begin{aligned}\varrho_{\mathrm{ad}}(m,p) &= \sum_{j=1}^{m-1}\binom{p}{j}(-1)^{j-1}\left(1-\frac{j}{p}\right)^m \\ &= p\left(1-\frac{1}{p}\right)^m - \binom{p}{2}\left(1-\frac{2}{p}\right)^m + \dots\end{aligned}$$

We note that ratio of two consecutive summands in absolute value is

$$\begin{aligned}\frac{\binom{p}{j}\left(1-\frac{j}{p}\right)^m}{\binom{p}{j+1}\left(1-\frac{j+1}{p}\right)^m} &= \left(1+\frac{1}{p-j-1}\right)^m\frac{j+1}{p-j} \ge \left(1+\frac{1}{p-2}\right)^m\frac{1}{p-1} \\ &\ge 2\exp\left\{\frac{m}{p-1}-\log(p-1)\right\} \ge 2\log m\end{aligned}$$

for all $p \le m/\log m$. Therefore, we deduce that

$$\varrho_{\mathrm{ad}}(m,p) = p\left(1-\frac{1}{p}\right)^m\left(1+O\left(\frac{1}{\log m}\right)\right),$$

which implies that

$$\varrho_{\mathrm{ad}}(m) = e^{O(\frac{m\log\log m}{\log m})}\prod_{p\le m/\log m}p\left(1-\frac{1}{p}\right)^m = \frac{e^{O(\frac{m\log\log m}{\log m})}}{(e^\gamma\log m)^m},$$

which proves (a quantitative version of) relation (3).

5 A Better Density in the Continuous Case

Analogous to the discrete question considered here, one can also ask about the continuous version, i.e. the set of limit points $\mathscr{L}$ of the set of values of $(p_{n+1}-p_n)/\log p_n$, where p_n is the nth prime. One can deduce from a uniform version of Zhang's Theorem that for any $0 \le \beta_1 \le \beta_2 \le \dots \le \beta_k$ with $k = k_2$ there exists

$1 \le i < j \le k$ such that $\beta_j - \beta_i \in \mathscr{L}$; in fact, this was done by Banks, Freiberg, and Maynard [1]. By a small modification of the argument used to prove Corollary 1.1, one can then show that $\mathscr{L} \cap [0, T]$ has Lebesgue measure $\gtrsim T/(k-1)$. Somewhat remarkably, Banks, Freiberg, and Maynard were able to go beyond this by showing that, given a sufficiently large k-tuple partitioned into 9 equal parts, at least two of these parts must simultaneously represent a prime. From this, they deduce that $\mathscr{L} \cap [0, T]$ has Lebesgue measure $\gtrsim T/8$.

Acknowledgements The first and third authors are supported by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada. The second and fourth authors were supported by NSF Mathematical Sciences Postdoctoral Research Fellowships.

References

1. W.D. Banks, T. Freiberg, J. Maynard, On limit points of the sequence of normalized prime gaps. Preprint. Available at http://arxiv.org/abs/1404.5094
2. A. Granville, Primes in intervals of bounded length. Bull. Am. Math. Soc. MathSciNet, MR3312631
3. J. Maynard, Small gaps between primes. Ann. Math. MathSciNet, MR3272929
4. J. Pintz, Polignac numbers, conjectures of Erdős on gaps between primes, arithmetic progressions in primes, and the bounded gap conjecture. Preprint. Available at http://arxiv.org/abs/1305.6289
5. D.H.J. Polymath, Variants of the Selberg sieve, and bounded intervals containing many primes. Res. Math. Sci. MathSciNet, MR3373710
6. Y. Zhang, Bounded gaps between primes. Ann. Math. **179**(3), 1121–1174 (2014)

A Note on Helson's Conjecture on Moments of Random Multiplicative Functions

Adam J. Harper, Ashkan Nikeghbali, and Maksym Radziwiłł

To Professor Helmut Maier on the occasion of his 60th birthday

1 Introduction

In this note we are interested in cancellations in sums of multiplicative functions. It is well known that

$$M(x) := \sum_{n \le x} \mu(n) = O(x^{1/2+\varepsilon})$$

is equivalent to the Riemann Hypothesis. On the other hand, it is also a classical result that $M(x) > x^{1/2-\varepsilon}$ for a sequence of arbitrarily large x. It is in fact conjectured that

$$\overline{\lim_{x\to\infty}} \frac{M(x)}{\sqrt{x}(\log\log\log x)^{\frac{5}{4}}} = \pm B$$

for some constant $B > 0$ (see [21]).

Wintner [24] initiated the study of what happens for a generic multiplicative function which is as likely to be 1 or -1 on the primes. Consider $f(p)$, a sequence of independent random variables taking values ± 1 with probability $1/2$ each

A.J. Harper
Jesus College, Cambridge CB5 8BL, England
e-mail: A.J.Harper@dpmms.cam.ac.uk

A. Nikeghbali (✉)
University of Zürich, Institute of Mathematics, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
e-mail: ashkan.nikeghbali@math.uzh.ch

M. Radziwiłł
Department of Mathematics, Rutgers University, Hill Center for the Mathematical Sciences, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA
e-mail: maksym.radziwill@gmail.com

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_11

(i.e., Rademacher random variables), and define a multiplicative function supported on squarefree integers n by

$$f(n) := \prod_{p|n} f(p).$$

We shall refer to such a function as a *Rademacher random multiplicative function*. By the three series theorem, the Euler product $F(s) := \prod_p (1 + f(p)p^{-s})$ converges almost surely for $\Re s > \frac{1}{2}$. From this Wintner deduced that

$$\sum_{n \le x} f(n) \ll x^{1/2+\varepsilon} \text{ almost surely (a.s.)}$$

Since then the problem of the behavior of $\sum_{n \le x} f(n)$ has attracted considerable attention [7, 11–13, 17, 18]. A closely related model is to let $f(p)$ be uniformly distributed on the complex unit circle (i.e., Steinhaus random variables), and then define $f(n) := \prod_{p^\alpha || n} f(p)^\alpha$ for all n. We shall refer to such a function as a *Steinhaus random multiplicative function*.

Very recently mean values of random multiplicative functions arose in connection with harmonic analysis. In his last paper Helson [16] considered the question of generalizing Nehari's theorem to the infinite polydisk. He noticed that the generalization could be disproved if one could show that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \Big| \sum_{n \le N} n^{-it} \Big| dt = o(\sqrt{N}). \tag{1}$$

Using Bohr's identification, we have

$$\left(\mathbb{E} \left| \sum_{n \le N} f(n) \right|^{2q} \right)^{1/2q} = \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T \left| \sum_{n \le N} n^{-it} \right|^{2q} dt \right)^{1/2q} \tag{2}$$

for all $2q > 0$, and with $f(n)$ a Steinhaus random multiplicative function. Therefore (1) is equivalent to

$$\mathbb{E} \Big| \sum_{n \le N} f(n) \Big| = o(\sqrt{N}), \tag{3}$$

with $f(n)$ a Steinhaus random multiplicative function. Helson justified his belief in (1) by observing that $N(it) := \sum_{n \le N} n^{-it}$ is the multiplicative analogue of the classical Dirichlet kernel $D(\theta) := \sum_{|n| \le N} e^{2\pi i n\theta}$. Since $\|D\|_1 = o(\|D\|_2)$ Helson conjectured that the same phenomenon should happen for the "multiplicative analogue" $N(it)$. Another reason one might believe the large cancellation in (1) to be possible is that on the $\frac{1}{2}$-line one has

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T\Big|\sum_{n\le N}\frac{1}{n^{1/2+it}}\Big|dt \ll \log^{1/4+o(1)} N,$$

as follows from the work of Bondarenko et al. [3]. This bound is stronger than one would expect assuming only squareroot cancellation, which would suggest a size more like $\log^{1/2} N$.

Recently Ortegà-Cerda and Seip [22] proved that Nehari's theorem doesn't extend to the infinite polydisk. However the problem of establishing (1) remained. There are now reasons to believe that (1) is false. In a recent paper Bondarenko and Seip [2] showed that the first absolute moment is at least $\sqrt{N}(\log N)^{-\delta+o(1)}$ for some small $\delta < 1$. Our primary goal in this note is to improve further on the lower bounds for (2). Our results also work for Rademacher random multiplicative functions.

Theorem 1.1. *Let $f(n)$ be a Rademacher or Steinhaus random multiplicative function. Then,*

$$\mathbb{E}\left|\sum_{n\le N} f(n)\right| \gg \sqrt{N}(\log\log N)^{-3+o(1)}$$

as $N\to\infty$.

The main input in the proof of Theorem 1.1 is the work [12] of the first named author on lower bounds for sums of random multiplicative functions. Using Hölder's inequality, we can extend the result of Theorem 1.1 to L^q norms.

Theorem 1.2. *Let $f(n)$ be a Rademacher or Steinhaus random multiplicative function and let $0 \le q \le 1$. Then,*

$$\mathbb{E}\left|\sum_{n\le N} f(n)\right|^{2q} \gg N^q(\log\log N)^{-6+o(1)}.$$

Theorems 1.1 and 1.2 suggest it is rather unlikely that Helson's conjecture is true. See Conjecture 1.1, below.

In addition to the above results, we establish an asymptotic estimate for the $2k$-th moment when k is a positive integer.

Theorem 1.3. *Let $k\in\mathbb{N}$. Suppose that $f(n)$ is a Steinhaus random multiplicative function. Then, as $N\to\infty$,*

$$\mathbb{E}\left|\sum_{n\le N} f(n)\right|^{2k} \sim \binom{2k-2}{k-1} k^{-(k-1)}\cdot c_k\gamma_k\cdot N^k\cdot(\log N)^{(k-1)^2},$$

where γ_k is the volume of Birkhoff polytope $\mathscr{B}_k$, defined as the $(k-1)^2$ dimensional volume of the set of $(u_{i,j})\in\mathbb{R}_+^{k^2}$ such that

$$\text{for each } i \le k : \sum_{1 \le j \le k} u_{i,j} = 1$$

$$\text{and for each } j \le k : \sum_{1 \le i \le k} u_{i,j} = 1,$$

and

$$c_k = \prod_p \Big(1 - \frac{1}{p}\Big)^{k^2} \cdot \Big(1 + \sum_{\alpha \ge 1} \frac{\binom{\alpha+k-1}{k-1}^2}{p^\alpha}\Big).$$

Note that $\mathscr{B}_k$ is a $(k-1)^2$ dimensional object embedded in a k^2 dimensional space. The $(k-1)^2$ dimensional volume of $\mathscr{B}_k$ is equal (see, e.g., Sect. 2 of Chan and Robbins [6]) to k^{k-1} times the full-dimensional volume of the set of $(u_{i,j})_{i,j \le k-1} \in \mathbb{R}^{(k-1)^2}$ such that, for all $i, j \le k-1$,

$$\sum_{j \le k-1} u_{i,j} \le 1 \text{ and } \sum_{i \le k-1} u_{i,j} \le 1 \text{ and } \sum_{i,j \le k-1} u_{i,j} \ge k-2.$$

The latter is how the volume of $\mathscr{B}_k$ will actually arise in our calculations.

It is worth pointing out that finding a closed formula for the volume of the Birkhoff polytope $\mathscr{B}_k$ is a notorious open question and would be of interest in enumerative combinatorics, statistics, and computational geometry (see [23]). There are evaluations of $\mathrm{Vol}(\mathscr{B}_k)$ for small values of k (see [1, 6]),

$$\mathrm{Vol}(\mathscr{B}_3) = \frac{3 \cdot 3^2}{2^2!}, \ \mathrm{Vol}(\mathscr{B}_4) = \frac{352 \cdot 4^3}{3^2!}, \ \mathrm{Vol}(\mathscr{B}_5) = \frac{4718075 \cdot 5^4}{4^2!}, \ \ldots$$

and an asymptotic formula is known to hold [5]

$$\mathrm{Vol}(\mathscr{B}_k) \sim \sqrt{2\pi} e^{1/3} \cdot \frac{k^{-(k-1)^2} e^{k^2}}{(2\pi)^k}, \ k \to \infty.$$

In addition the asymptotic behavior of the Euler product c_k is known (see [9, Proposition]),

$$\log c_k = -k^2 \log(2e^\gamma \log k) + o(k^2)$$

where γ is the Euler–Mascheroni constant.

We note that Conrey and Gamburd [8] compute the even integer moments

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \Big| \sum_{n \le N} n^{-1/2-it} \Big|^{2k} dt$$

on the $\frac{1}{2}$-line, and unsurprisingly the answer is extremely similar to Theorem 1.3 (in particular an Euler product and a volume related to the Birkhoff polytope again appear). Conrey and Gamburd discuss the connection between their result and the moments of certain truncated characteristic polynomials of random matrices. In general, it seems reasonable to say that the arithmetic factor c_k reflects local counting modulo different primes in the moment computation, whereas the geometric factor γ_k reflects global counting of tuples $n_1, \dots, n_k$ and $m_1, \dots, m_k$ subject to the truncation $n_i, m_i \leq N$.

We deduce Theorem 1.3 from a general result of La Bretèche [4] on mean values of multiplicative functions in several variables. Theorem 1.3 has also been obtained independently by Granville and Soundararajan (unpublished), and also very recently (and independently) by Heap and Lindqvist [15]. Additionally, Theorem 1.3 sheds light on the conjectural behavior of moments of the theta functions,

$$\theta(x, \chi) = \sum_{n \geq 1} \chi(n) e^{-\pi n^2 x/p}$$

with $p \geq 3$ a prime, and χ an even Dirichlet character modulo p. The rapidly decaying factor $e^{-\pi n^2 x/p}$ essentially restricts the sum to those n less than about $\sqrt{p}$ (if $x = 1$, say), and the average behavior of $\chi(n)$ with $n \ll p^{1/2}$ is similar to that of a Steinhaus random multiplicative function. Therefore Theorem 1.3 leads to the conjecture that

$$\frac{1}{p} \sum_{\substack{\chi \bmod p \\ \chi \text{ even}}} |\theta(1, \chi)|^{2k} \sim C_k p^{k/2} (\log p)^{(k-1)^2} \quad \text{as } p \to \infty.$$

In unpublished recent work the same conjecture was stated by Marc Munsch on the basis of his lower bound for moments of $\theta(1, \chi)$. Louboutin and Munsch [19] prove the conjecture for $k = 1$ and $k = 2$.

Combining Theorems 1.2 and 1.3 suggests the following "counter-conjecture" to Helson's claim (1).

Conjecture 1.1. If $f(n)$ is a Steinhaus random multiplicative function, then we have as $N \to \infty$,

$$\mathbb{E}\left|\sum_{n \leq N} f(n)\right|^{2q} \sim \begin{cases} C(q) N^q, & \text{for } 0 \leq q \leq 1 \\ C(q) N^q (\log N)^{(q-1)^2}, & \text{for } 1 \leq q. \end{cases}$$

Conjecture 1.1 suggests a possible line of attack on the problem of showing that for a positive proportion of even characters χ modulo p, we have $\theta(1, \chi) \neq 0$. This would be based on comparing the first and second *absolute* moments, i.e.

$$\sum_{\substack{\chi \bmod p \\ \chi \text{ even}}} |\theta(1, \chi)| \quad \text{and} \quad \sum_{\substack{\chi \bmod p \\ \chi \text{ even}}} |\theta(1, \chi)|^2.$$

We emphasize that we do not have a lot of evidence towards Conjecture 1.1 when $q \notin \mathbb{N}$, and perhaps especially when $0 < q < 1$, and it is conceivable the behavior could be more complicated. However this is the simplest possible conjecture respecting the information that we now have. In addition for $q > 1$ it perhaps seems unlikely that the distribution of the tails of $\sum_{n<N} f(n)$ (in a large deviation regime) fluctuates so significantly that it would affect the exponent $(q-1)^2$ of the logarithm when q goes from an integer to a fractional exponent. We also note that if we could obtain the order of magnitude for the $2q$-th moment suggested by Conjecture 1.1 for $q = \frac{1}{2}$, then since we know it trivially for $q = 1$ a simple argument using Hölder's inequality (as in the proof of Theorem 1.2, below) would establish the order of magnitude suggested by the Conjecture for all $0 \le q \le 1$.

Finally, following a question from the referee, we noticed that we can extend Theorem 1.3 to the Rademacher case. We omit the simple cases of $k = 1, 2$ in the theorem below, since both are different from the case $k \ge 3$.

Theorem 1.4. *Let $f(n)$ be a Rademacher random multiplicative function. Then, for $k \ge 3$ an integer, as $N \to \infty$,*

$$\mathbb{E}\Big(\sum_{n\le N} f(n)\Big)^k \sim C_k \cdot N^{k/2} (\log N)^{k(k-3)/2}$$

with $C_k > 0$ constant.

Similarly as in Theorem 1.3 the constant C_k splits into an arithmetic and geometric factor. The interested reader should have no trouble working out the details. Theorem 1.4 has also been obtained independently by Heap and Lindqvist [15].

At first glance it may seem strange that all the moments here (including the odd ones) are non-trivially large, but that is because in the Rademacher case there is no distinction between a term and its complex conjugate (and similarly if one calculated an expression like $\mathbb{E}\left|\sum_{n\le N} f(n)\right|^{2k}\left(\sum_{n\le N} f(n)\right)$ in the Steinhaus case, this would be non-trivially large provided $k \ge 1$). Note also that the moments are rather larger in the Rademacher case than the Steinhaus case, again because everything is real valued and so the terms exhibit less cancellation.

2 Lower Bounds for the First Moment

In this section we shall first prove the following result.

Proposition 2.1. *Let $f(n)$ be a Rademacher random multiplicative function. There exist arbitrarily large values of x for which*

$$\mathbb{E}\left|\sum_{n\le x} f(n)\right| \ge \frac{\sqrt{x}}{(\log\log x)^{3+o(1)}}.$$

The same is true if $f(n)$ is a Steinhaus random multiplicative function.

The above proposition is actually a fairly straightforward deduction from the work of Harper [12]. However, it is a bit unsatisfactory because it only gives a lower bound along some special sequence of x values. With more work we can correct this defect, as in the following theorem announced in the Introduction:

Theorem 2.1. *Let $f(n)$ be a Rademacher random multiplicative function. Then for* all *large x we have*

$$\mathbb{E}\left|\sum_{n\le x} f(n)\right| \ge \frac{\sqrt{x}}{(\log\log x)^{3+o(1)}}.$$

The same is true if $f(n)$ is a Steinhaus random multiplicative function.

The proof of Proposition 2.1 has two ingredients. The first is the observation, essentially due to Halász [11], that one can almost surely lower bound an average of $\left|\sum_{n\le x} f(n)\right|$ in terms of the behavior of $f(n)$ on primes only: more specifically, in the Rademacher case we almost surely have that, for any $y \ge 2$,

$$\int_1^\infty \frac{\left|\sum_{n\le z} f(n)\right|}{z^{3/2+1/\log y}}\,dz \gg \sup_{t\ge 1} \exp\Big(\sum_p \frac{f(p)\cos(t\log p)}{p^{1/2+1/\log y}} - \log t - \log\log(t+2)/2\Big).$$

Here the implicit constant in the $\gg$ notation is absolute. The reader should note that the presence of the supremum over t will be very significant here, since at any fixed t the expected size of the right-hand side would be too small to produce a useful result (about $\log^{1/4} y$, rather than about $\log y$ which is what we need).

The second ingredient is a strong lower bound for the expected size of the right-hand side, which we deduce from the work of Harper [12]. We quote the relevant statements from Harper's work as a lemma now, for ease of reference later.

Lemma 2.1 (See § 6.3 of [12]). *If $(f(p))_{p\ prime}$ are independent Rademacher random variables, then with probability $1-o(1)$ as $x\to\infty$ we have*

$$\sup_{1\le t\le 2(\log\log x)^2} \sum_p \frac{f(p)\cos(t\log p)}{p^{1/2+1/\log x}} \ge \log\log x - \log\log\log x - O((\log\log\log x)^{3/4}).$$

If $(f(p))_{p\ prime}$ are independent Steinhaus random variables, then with probability $1-o(1)$ as $x\to\infty$ we have

$$\sup_{1\le t\le 2(\log\log x)^2} \sum_p \Big(\frac{\Re(f(p)p^{-it})}{p^{1/2+1/\log x}} + \frac{1}{2}\frac{\Re(f(p)^2p^{-2it})}{p^{1+2/\log x}}\Big) \ge \log\log x - \log\log\log x - O((\log\log\log x)^{3/4}).$$

The first statement here is proved in the last paragraph in § 6.3 of [12] (noting that the quantity y there is $\log^8 x$). The second statement can be proved by straightforward adaptation of that argument, the point being that the expectation and covariance structure of these random sums in the Steinhaus case are the same, up to negligible error terms, as in the Rademacher case, so the same arguments can be applied. (See the preprint [14] for an explicit treatment of some very similar Steinhaus random sums.) The argument in [12] is quite involved, but the basic aim is to show that, for the purpose of taking the supremum, the sums $\sup_{1\le t\le 2(\log\log x)^2}\sum_p \frac{f(p)\cos(t\log p)}{p^{1/2+1/\log x}}$ behave somewhat independently at values of t that are separated by $\gg 1/\log x$, so one has something like the supremum over $\log x$ independent samples.

To prove Theorem 1.1 we introduce a third ingredient, namely we show that $\mathbb{E}\left|\sum_{n\le x} f(n)\right|$ may itself be lower bounded in terms of an integral average of $\mathbb{E}\left|\sum_{n\le z} f(n)\right|$, as follows:

Proposition 2.2. *Let $f(n)$ be a Rademacher random multiplicative function. For any large x we have*

$$\mathbb{E}\left|\sum_{n\le x} f(n)\right| \gg \frac{\sqrt{x}}{\log x}\int_1^{\sqrt{x}}\left(\frac{\mathbb{E}|\sum_{n\le z} f(n)|}{\sqrt{z}}\right)\frac{dz}{z}.$$

The same is true if $f(n)$ is a Steinhaus random multiplicative function.

This uses the multiplicativity of $f(n)$ in an essential way (as does the proof of Proposition 2.1, of course).

Theorem 1.1 then follows quickly by combining Proposition 2.2 with the proof of Proposition 2.1.

As the reader will see, the proof of Proposition 2.2 is based on a "physical space" decomposition of the sum $\sum_{n\le x} f(n)$, which is somewhat related to the martingale arguments of Harper [13]. This is unlike the other arguments above, which work by establishing a connection between the integral average of $\sum_{n\le x} f(n)$ and its Dirichlet series $\sum_n f(n)/n^s$ (on the "Fourier space" side).

2.1 Proof of Proposition 2.1

The proof of Proposition 2.1 is slightly cleaner in the Rademacher case, because then $f(p)^2 \equiv 1$ for all primes p. So we shall give the proof in that case first, and afterwards explain the small changes that arise in the Steinhaus case.

We know from work of Wintner [24] that almost surely $\sum_{n\le x} f(n) = O_\epsilon(x^{1/2+\epsilon})$. Consequently, by partial summation the Dirichlet series $F(s) := \sum_n f(n)/n^s$ is almost surely convergent in the half plane $\Re(s) > 1/2$, and then by term by term integration it satisfies

$$F(s) = s\int_1^\infty \frac{\sum_{n\le z} f(n)}{z^{s+1}}dz, \qquad \Re(s) > 1/2.$$

In particular, $F(s)$ is almost surely a holomorphic function on the half plane $\Re(s) > 1/2$.

On the other hand, since $f(n)$ is multiplicative we have for any $\Re(s) > 1$ that, in the Rademacher case,

$$\begin{aligned} F(s) &= \prod_p \Big(1 + \frac{f(p)}{p^s}\Big) = \exp\Big(\sum_p \log\Big(1 + \frac{f(p)}{p^s}\Big)\Big) \\ &= \exp\Big(\sum_p \frac{f(p)}{p^s} - \frac{1}{2}\sum_p \frac{f(p)^2}{p^{2s}} + \sum_{k\ge 3}\frac{(-1)^{k+1}}{k}\sum_p \frac{f(p)^k}{p^{ks}}\Big). \end{aligned}$$

Therefore in the Rademacher case we have

$$s\int_1^\infty \frac{\sum_{n\le z} f(n)}{z^{s+1}}dz = \exp\Big(\sum_p \frac{f(p)}{p^s} - \frac{1}{2}\sum_p \frac{f(p)^2}{p^{2s}} + \sum_{k\ge 3}\frac{(-1)^{k+1}}{k}\sum_p \frac{f(p)^k}{p^{ks}}\Big)$$

at least when $\Re(s) > 1$, since both sides are equal to $F(s)$. But all the sums involving p^{2s} and p^{ks} are clearly absolutely convergent whenever $\Re(s) > 1/2$, and therefore define holomorphic functions there. In addition, for any fixed s with $\Re(s) > 1/2$ the series $\sum_p \frac{f(p)}{p^s}$ is a sum of independent random variables, and Kolmogorov's Three Series Theorem implies it converges almost surely. Since a Dirichlet series is a holomorphic function strictly to the right of its abscissa of converge, we find that almost surely $\sum_p \frac{f(p)}{p^s}$ is a holomorphic function on the half plane $\Re(s) > 1/2$, and so almost surely we have, for all $\Re s > \frac{1}{2}$,

$$s\int_1^\infty \frac{\sum_{n\le z} f(n)}{z^{s+1}}dz = \exp\Big(\sum_p \frac{f(p)}{p^s} - \frac{1}{2}\sum_p \frac{f(p)^2}{p^{2s}} + \sum_{k\ge 3}\frac{(-1)^{k+1}}{k}\sum_p \frac{f(p)^k}{p^{ks}}\Big).$$

Next, if we write $s = \sigma + it$ and take absolute values on both sides then we find that, almost surely,

$$\begin{aligned} &|s|\int_1^\infty \frac{\big|\sum_{n\le z} f(n)\big|}{z^{\sigma+1}}dz \\ &\quad\ge \exp\Big(\Re\Big(\sum_p \frac{f(p)}{p^s} - \frac{1}{2}\sum_p \frac{f(p)^2}{p^{2s}} + \sum_{k\ge 3}\frac{(-1)^{k+1}}{k}\sum_p \frac{f(p)^k}{p^{ks}}\Big)\Big) \\ &\quad= \exp\Big(\sum_p \frac{\Re(f(p)p^{-it})}{p^\sigma} - \frac{1}{2}\sum_p \frac{\Re(f(p)^2p^{-2it})}{p^{2\sigma}} + O(1)\Big), \qquad \forall \sigma > 1/2. \end{aligned}$$

If we take $\sigma = 1/2 + 1/\log y$ for a parameter $y \geq 2$, and we note that then $|s| \asymp |t|$ provided $t \geq 1$ (say), we have almost surely that for all $y \geq 2$,

$$\int_1^\infty \frac{\left|\sum_{n\leq z} f(n)\right|}{z^{3/2+1/\log y}} dz \gg \sup_{t\geq 1} \exp\Big(\sum_p \frac{\Re(f(p)p^{-it})}{p^{1/2+1/\log y}} - \frac{1}{2}\sum_p \frac{\Re(f(p)^2 p^{-2it})}{p^{1+2/\log y}} - \log t\Big).$$

In the Rademacher case the first sum over p is $\sum_p \frac{f(p)\cos(t\log p)}{p^{1/2+1/\log y}}$, and (since $f(p)^2 = 1$) the second sum over p is $\Re\sum_p \frac{1}{p^{1+2/\log y+2it}} = \Re\log\zeta(1 + 2/\log y + 2it) + O(1)$, where ζ denotes the Riemann zeta function. Standard estimates (see, e.g., Theorem 6.7 of Montgomery and Vaughan [20]) imply that $|\log\zeta(1 + 2/\log y + 2it)| \leq \log\log(t + 2) + O(1)$ for $t \geq 1$, so we have almost surely that for all $y \geq 2$,

$$\int_1^\infty \frac{\left|\sum_{n\leq z} f(n)\right|}{z^{3/2+1/\log y}} dz \gg \sup_{t\geq 1} \exp\Big(\sum_p \frac{f(p)\cos(t\log p)}{p^{1/2+1/\log y}} - \log t - \log\log(t+2)/2\Big). \tag{4}$$

(The above argument and inequality (4) are essentially due to Halász [11], and are also related to the arguments of Wintner [24]. The only small difference is that Halász restricted to $1 \leq t \leq 2$. See Appendix A of Harper [12] for a presentation similar to the above.)

Now to prove Proposition 2.1, note that for any large parameters x and $x_0 < x_1$ we have

$$\sup_{x_0<z<x_1} \frac{\mathbb{E}|\sum_{n\leq z} f(n)|}{\sqrt{z}} \geq \frac{1}{\log x}\int_{x_0}^{x_1} \frac{\mathbb{E}\left|\sum_{n\leq z} f(n)\right|}{z^{3/2+1/\log x}} dz,$$

since $\int_{x_0}^{x_1} \frac{dz}{z^{1+1/\log x}} \leq \int_1^\infty \frac{dz}{z^{1+1/\log x}} = \log x$. Then by Cauchy–Schwarz we always have $\mathbb{E}|\sum_{n\leq z} f(n)| \leq \sqrt{z}$, so

$$\begin{aligned}\int_{x_0}^{x_1} \frac{\mathbb{E}\left|\sum_{n\leq z} f(n)\right|}{z^{3/2+1/\log x}} dz &\geq \int_1^\infty \frac{\mathbb{E}\left|\sum_{n\leq z} f(n)\right|}{z^{3/2+1/\log x}} dz \\ &\quad - \int_1^{x_0} \frac{dz}{z^{1+1/\log x}} - \int_{x_1}^\infty \frac{dz}{z^{1+1/\log x}} \\ &\geq \int_1^\infty \frac{\mathbb{E}\left|\sum_{n\leq z} f(n)\right|}{z^{3/2+1/\log x}} dz - \log x_0 - \frac{\log x}{x_1^{1/\log x}}.\end{aligned}$$

In particular, if we choose $x_0 = e^{\sqrt{\log x}}$ and $x_1 = e^{(\log x)\log\log x}$, say, then we have

$$\sup_{x_0<z<x_1} \frac{\mathbb{E}|\sum_{n\le z} f(n)|}{\sqrt{z}} \ge \frac{1}{\log x}\int_1^\infty \frac{\mathbb{E}\left|\sum_{n\le z} f(n)\right|}{z^{3/2+1/\log x}}dz - \frac{2}{\sqrt{\log x}}. \tag{5}$$

Finally, in the Rademacher case Lemma 2.1 implies that, with probability $1-o(1)$ as $x \to \infty$,

$$\sup_{1\le t\le 2(\log\log x)^2} \sum_p \frac{f(p)\cos(t\log p)}{p^{1/2+1/\log x}} \ge \log\log x - \log\log\log x - O((\log\log\log x)^{3/4}).$$

This implies that with probability $1 - o(1)$ one has

$$\sup_{t\ge 1} \exp\Big(\sum_p \frac{f(p)\cos(t\log p)}{p^{1/2+1/\log x}} - \log t - \log\log(t+2)/2\Big) \ge \frac{\log x}{(\log\log x)^{3+o(1)}},$$

and then by the Halász type lower bound inequality (4) we deduce

$$\int_1^\infty \frac{\mathbb{E}\left|\sum_{n\le z} f(n)\right|}{z^{3/2+1/\log x}}dz \ge \frac{\log x}{(\log\log x)^{3+o(1)}}. \tag{6}$$

Proposition 2.1 follows in the Rademacher case by combining this with (5).

In the Steinhaus case the initial argument of Wintner [24] still works, so the first change that is needed in the preceding argument comes in the expression for the Euler product $F(s)$, which for $\Re(s) > 1$ is now

$$F(s) = \prod_p \left(1 + \sum_{j=1}^\infty \frac{f(p)^j}{p^{js}}\right) = \exp\Big(-\sum_p \log\left(1 - \frac{f(p)}{p^s}\right)\Big)$$
$$= \exp\Big(\sum_p \frac{f(p)}{p^s} + \frac{1}{2}\sum_p \frac{f(p)^2}{p^{2s}} + \sum_{k\ge 3}\frac{1}{k}\sum_p \frac{f(p)^k}{p^{ks}}\Big).$$

Notice this is the same as we had in the Rademacher case, except now there are no alternating minus signs in the final exponential. The argument using the Three Series Theorem, etc. then continues as in the Rademacher case to yield that, almost surely,

$$s\int_1^\infty \frac{\sum_{n\le z} f(n)}{z^{s+1}}dz = \exp\Big(\sum_p \frac{f(p)}{p^s} + \frac{1}{2}\sum_p \frac{f(p)^2}{p^{2s}} + \sum_{k\ge 3}\frac{1}{k}\sum_p \frac{f(p)^k}{p^{ks}}\Big)$$
$$\forall\ \Re(s) > 1/2.$$

Putting $s = 1/2 + 1/\log y + it$ and taking absolute values on both sides, we deduce that almost surely,

$$\int_1^\infty \frac{\left|\sum_{n\le z} f(n)\right|}{z^{3/2+1/\log y}}dz \gg \sup_{t\ge 1} \exp\Big(\sum_p \Big(\frac{\Re(f(p)p^{-it})}{p^{1/2+1/\log y}} + \frac{1}{2}\frac{\Re(f(p)^2p^{-2it})}{p^{1+2/\log y}}\Big) - \log t\Big)$$
$$\forall\ y \ge 2. \tag{7}$$

Since we don't now have $f(p)^2 \equiv 1$, we cannot remove the contribution of the prime squares using estimates for the zeta function. However, by the Steinhaus case of Lemma 2.1 we still have that, with probability $1 - o(1)$ as $x \to \infty$,

$$\sup_{1\le t\le 2(\log\log x)^2} \sum_p \Big(\frac{\Re(f(p)p^{-it})}{p^{1/2+1/\log x}} + \frac{1}{2}\frac{\Re(f(p)^2p^{-2it})}{p^{1+2/\log x}}\Big) \ge \log\log x - \log\log\log x$$
$$- O((\log\log\log x)^{3/4}),$$

and therefore with probability $1 - o(1)$ we have

$$\sup_{t\ge 1} \exp\Big(\sum_p \Big(\frac{\Re(f(p)p^{-it})}{p^{1/2+1/\log x}} + \frac{1}{2}\frac{\Re(f(p)^2p^{-2it})}{p^{1+2/\log x}}\Big) - \log t\Big) \ge \frac{\log x}{(\log\log x)^{3+o(1)}}.$$

Combining this estimate with (7) and (5) then proves Proposition 2.1 in the Steinhaus case.

2.2 *Proofs of Theorem 1.1 and Proposition 2.2*

Proof (Proof of Theorem 1.1, Assuming Proposition 2.2). In view of Proposition 2.2, it will suffice to prove that for all large x we have

$$\int_1^{\sqrt{x}} \Big(\frac{\mathbb{E}|\sum_{n\le z} f(n)|}{\sqrt{z}}\Big)\frac{dz}{z} \ge \frac{\log x}{(\log\log x)^{3+o(1)}}.$$

However, for any large parameter y we have

$$\int_1^{\sqrt{x}} \Big(\frac{\mathbb{E}|\sum_{n\le z} f(n)|}{\sqrt{z}}\Big)\frac{dz}{z} \ge \int_1^{\sqrt{x}} \frac{\mathbb{E}|\sum_{n\le z} f(n)|}{z^{3/2+1/\log y}}dz$$
$$\ge \frac{\log y}{(\log\log y)^{3+o(1)}} - \int_{\sqrt{x}}^\infty \frac{\mathbb{E}|\sum_{n\le z} f(n)|}{z^{3/2+1/\log y}}dz,$$

in view of the lower bound $\int_1^\infty \frac{\mathbb{E}|\sum_{n\le z} f(n)|}{z^{3/2+1/\log y}} dz \ge \frac{\log y}{(\log\log y)^{3+o(1)}}$ obtained in (6). By Cauchy–Schwarz we always have $\mathbb{E}|\sum_{n\le z} f(n)| \le \sqrt{z}$, so the subtracted term here is at most

$$\int_{\sqrt{x}}^\infty \frac{dz}{z^{1+1/\log y}} = \frac{\log y}{(\sqrt{x})^{1/\log y}}.$$

If we choose $\log y$ somewhat smaller than $\log x$, say $\log y = (\log x)/(100 \log\log \log x)$, we deduce that

$$\int_1^{\sqrt{x}} \left(\frac{\mathbb{E}|\sum_{n\le z} f(n)|}{\sqrt{z}}\right) \frac{dz}{z} \ge \frac{\log x}{(\log\log x)^{3+o(1)}} - \frac{\log x}{(\log\log x)^{50}} = \frac{\log x}{(\log\log x)^{3+o(1)}},$$

as required.

Proof (Proof of Proposition 2.2). The first part of the proof again differs slightly depending on whether we are in the Rademacher or the Steinhaus case. We will first work in the Rademacher case and then explain the small changes needed in the other situation.

Let $A_t := \sum_{n\le t} f(n)$. If we let $P(n)$ denote the largest prime factor of n, we have

$$\sum_{n\le x} f(n) = \sum_{p\le x} \sum_{n\le x, P(n)=p} f(n) = \sum_{p\le x} f(p) \sum_{m\le x/p, P(m)<p} f(m),$$

since f is multiplicative. Here the inequality $P(m) < p$ in the final sum is strict because f is supported on squarefree numbers. Notice here that if $p > \sqrt{x}$ then $x/p < \sqrt{x} < p$, so we automatically have $P(m) < p$ in the inner sums over m. Thus we can rewrite things slightly as

$$\begin{aligned}\sum_{n\le x} f(n) &= \sum_{\sqrt{x}<p\le x} f(p) \sum_{m\le x/p} f(m) + \sum_{p\le\sqrt{x}} f(p) \sum_{m\le x/p, P(m)<p} f(m)\\ &=: \sum_{\sqrt{x}<p\le x} f(p) A_{x/p} + B_x,\end{aligned}$$

say. Notice also that the random variables $A_{x/p}$ and B_x are independent of the $f(p)$ for $\sqrt{x} < p \le x$.

We shall introduce a penultimate piece of notation, by defining the random variable

$$C_x := \sum_{\sqrt{x}<p\le x} f(p) A_{x/p}.$$

Finally, let ϵ be a Rademacher random variable that is independent of everything else.

Now since the $(f(p))_{\sqrt{x}<p\le x}$ are symmetric random variables independent of B_x and the $A_{x/p}$, it follows that

$$\sum_{n\le x} f(n) = \sum_{\sqrt{x}<p\le x} f(p)A_{x/p} + B_x \stackrel{d}{=} \epsilon \sum_{\sqrt{x}<p\le x} f(p)A_{x/p} + B_x,$$

where $\stackrel{d}{=}$ denotes equality in distribution. Then if we *condition on the values of* B_x, C_x, we find the conditional expectation

$$\mathbb{E}\left(\left|\epsilon \sum_{\sqrt{x}<p\le x} f(p)A_{x/p} + B_x\right| \Big| B_x, C_x\right)$$
$$= (1/2)|C_x + B_x| + (1/2)|-C_x + B_x| \ge |C_x|,$$

by the triangle inequality. Now if we average over values of B_x, C_x, and use the Tower Property of conditional expectations (the fact that the expectation of a conditional expectation is the unconditional expectation), we obtain

$$\mathbb{E}\left|\sum_{n\le x} f(n)\right| = \mathbb{E}\left|\epsilon \sum_{\sqrt{x}<p\le x} f(p)A_{x/p} + B_x\right| \ge \mathbb{E}|C_x|.$$

On recalling the definitions of C_x and $A_{x/p}$, we see we have proved the following:

Lemma 2.2. *For all large x we have*

$$\mathbb{E}\left|\sum_{n\le x} f(n)\right| \ge \mathbb{E}\left|\sum_{\sqrt{x}<p\le x} f(p) \sum_{m\le x/p} f(m)\right|.$$

(In the Steinhaus case one has a weak inequality $P(m) \le p$ in the definition of B_x, since f is totally multiplicative, but this makes no difference to the argument just given. Instead of choosing ϵ to be a Rademacher random variable one can choose ϵ to be uniformly distributed on the unit circle, and then one obtains exactly the same conclusion in Lemma 2.2.)

Since the $f(p)$ are Rademacher or Steinhaus random variables independent of the "coefficients" $\sum_{m\le x/p} f(m) = A_{x/p}$, an application of Khintchine's inequality (see, e.g., Gut's textbook [10]) yields that

$$\mathbb{E}\left|\sum_{\sqrt{x}<p\le x} f(p)\sum_{m\le x/p} f(m)\right| \gg \mathbb{E}\sqrt{\sum_{\sqrt{x}<p\le x}\left|\sum_{m\le x/p} f(m)\right|^2}.$$

It would be nice if we could find a way to exploit this (sharp) bound with the squares still in place on the inside, but to prove Proposition 2.2 we shall trade them away in order to remove the intractable squareroot. Thus by the Cauchy–Schwarz inequality and the fact that $\sum_{\sqrt{x}<p\le x} 1/p = \log 2 + o(1)$ we have

$$\begin{aligned}\sum_{\sqrt{x}<p\le x}\sqrt{\frac{1}{p}}\left|\sum_{m\le x/p} f(m)\right| &\le \sqrt{\sum_{\sqrt{x}<p\le x}\frac{1}{p}}\sqrt{\sum_{\sqrt{x}<p\le x}\left|\sum_{m\le x/p} f(m)\right|^2}\\ &\ll \sqrt{\sum_{\sqrt{x}<p\le x}\left|\sum_{m\le x/p} f(m)\right|^2}.\end{aligned}$$

Combining this with the above, we deduce:

Lemma 2.3. *For all large x we have*

$$\mathbb{E}\left|\sum_{n\le x} f(n)\right| \gg \sum_{\sqrt{x}<p\le x}\frac{1}{\sqrt{p}}\mathbb{E}\left|\sum_{m\le x/p} f(m)\right| \ge \frac{1}{\log x}\sum_{\sqrt{x}<p\le x}\frac{\log p}{\sqrt{p}}\cdot\mathbb{E}\left|\sum_{m\le x/p} f(m)\right|.$$

We have now almost finished the proof of Proposition 2.2. If we have two primes $z \le p \le p' \le z + z/\log^{1000} x$ for some $\sqrt{x} < z \le x$, then

$$\begin{aligned}\left|\mathbb{E}\left|\sum_{m\le x/p} f(m)\right| - \mathbb{E}\left|\sum_{m\le x/p'} f(m)\right|\right| &\le \mathbb{E}\left|\sum_{x/p'<m\le x/p} f(m)\right|\\ &\ll \sqrt{x(\frac{1}{p}-\frac{1}{p'})+1} \ll \sqrt{\frac{x}{p\log^{1000} x}}+1,\end{aligned}$$

by the Cauchy–Schwarz inequality and orthogonality of the $f(m)$. And we see

$$\begin{aligned}&\frac{1}{\log x}\sum_{\sqrt{x}<p\le x}\frac{\log p}{\sqrt{p}}\left(\sqrt{\frac{x}{p\log^{1000} x}}+1\right)\\ &\ll \frac{\sqrt{x}}{\log^{500} x}+\frac{1}{\log x}\sum_{\sqrt{x}<p\le x}\frac{\log p}{\sqrt{p}} \ll \frac{\sqrt{x}}{\log x},\end{aligned}$$

which will make a negligible contribution in Proposition 2.2, so in Lemma 2.3 we may replace each term $\mathbb{E}\left|\sum_{m\le x/p} f(m)\right|$ by an averaged version

$$\frac{\log^{1000} x}{p}\int_p^{p(1+1/\log^{1000} x)} \mathbb{E}\left|\sum_{m\le x/t} f(m)\right| dt.$$

Since we know that primes are well distributed in intervals of relative length $1+1/\log^{1000} x$ (with density 1 when weighted by $\log p$) we can rewrite Lemma 2.3 as

$$\begin{aligned}\mathbb{E}\left|\sum_{n\le x} f(n)\right| &\gg \frac{1}{\log x}\sum_{\sqrt{x}<p\le x} \log p \frac{\log^{1000} x}{p}\int_p^{p(1+1/\log^{1000} x)} \mathbb{E}\left|\sum_{m\le x/t} f(m)\right| \frac{dt}{\sqrt{t}}\\ &\gg \frac{1}{\log x}\int_{\sqrt{x}}^{x} \mathbb{E}\left|\sum_{m\le x/t} f(m)\right| \frac{dt}{\sqrt{t}}.\end{aligned}$$

Proposition 2.2 now follows by making the substitution $z = x/t$ in the integral.

3 Lower Bounds for small Moments: Proof of Theorem 1.2

The proof is a very simple argument using the Cauchy–Schwarz inequality and Hölder's inequality.

Indeed, for any $0 \le q \le 1$ we have

$$\begin{aligned}\mathbb{E}\Big|\sum_{n\le N} f(n)\Big| &\le \mathbb{E}\Big[\Big|\sum_{n\le N} f(n)\Big|^{2q}\Big]^{1/2}\cdot \mathbb{E}\Big[\Big|\sum_{n\le N} f(n)\Big|^{2-2q}\Big]^{1/2}\\ &\le \mathbb{E}\Big[\Big|\sum_{n\le N} f(n)\Big|^{2q}\Big]^{1/2}\cdot \mathbb{E}\Big[\Big|\sum_{n\le N} f(n)\Big|^{2}\Big]^{(1-q)/2}.\end{aligned}$$

Since $\mathbb{E}|\sum_{n\le N} f(n)|^2 \le N$ and $\mathbb{E}|\sum_{n\le N} f(n)| \ge \sqrt{N}/(\log\log N)^{3+o(1)}$, by rearranging we obtain the lower bound

$$\mathbb{E}\Big[\Big|\sum_{n\le N} f(n)\Big|^{2q}\Big] \ge N^q (\log\log N)^{-6+o(1)}.$$

4 Asymptotics for Even Moments: Proof of Theorem 1.3

Note that

$$\mathbb{E}\Big|\sum_{n\le X} f(n)\Big|^{2k} = \sum_{\substack{n_1,\dots,n_k\le X\\ m_1,\dots,m_k\le X}} \mathbb{E}[f(n_1)\dots f(n_k)\overline{f(m_1)\dots f(m_k)}]$$
$$= \sum_{\substack{n_1,\dots,n_k\le X\\ m_1,\dots,m_k\le X\\ n_1\dots n_k=m_1\dots m_k}} 1. \tag{8}$$

Now

$$g(n_1,\dots,n_k,m_1,\dots,m_k) = \mathbf{1}_{n_1\dots n_k=m_1\dots m_k}$$

is a multiplicative function of several variables[1] and our problem reduces to understanding the mean value of

$$\sum_{\substack{n_1,\dots,n_k\le X\\ m_1,\dots,m_k\le X}} g(n_1,\dots,n_k,m_1,\dots,m_k).$$

We notice that the associated multiple Dirichlet series

$$\sum_{\substack{n_1,\dots,n_k\\ m_1,\dots,m_k}} \frac{g(n_1,\dots,n_k,m_1,\dots,m_k)}{n_1^{s_1}\dots n_k^{s_k} m_1^{w_1}\dots m_k^{w_k}}$$
$$= \sum_n \sum_{n_1n_2\dots n_k=n} \frac{1}{n_1^{s_1}\dots n_k^{s_k}} \sum_{m_1m_2\dots m_k=n} \frac{1}{m_1^{w_1}\dots m_k^{w_k}}$$

is absolutely convergent for $\Re s_i, \Re w_i > \frac{1}{2}$ and moreover it factors as

$$H(s_1,\dots,s_k,w_1,\dots,w_k)\prod_{i=1}^{k}\prod_{j=1}^{k}\zeta(s_i+w_j)$$

[1] In other words

$$g(n_1,\dots,n_k,m_1,\dots,m_k)g(u_1,\dots,u_k,v_1,\dots,v_k) = g(n_1u_1,\dots,n_ku_k,m_1v_1,\dots,m_kv_k)$$

for any natural numbers n_i, m_i and u_i, v_i whose least common multiples are coprime.

with $H(s_1, \dots, s_k, w_1, \dots, w_k)$ absolutely convergent in the region $\Re s_i, \Re w_i > \frac{1}{4}$. In addition a direct check shows that

$$H(\tfrac{1}{2}, \dots, \tfrac{1}{2}) = \prod_p \Big(1 - \frac{1}{p}\Big)^{k^2} \cdot \Big(1 + \frac{k^2}{p} + \sum_{\alpha \geq 2} \frac{\binom{a+k-1}{k-1}^2}{p^\alpha}\Big) > 0.$$

Therefore the main result of La Bretèche [4] is applicable with the k^2 linear forms $\ell^{(i,j)}(s_1, \dots, s_k, w_1, \dots, w_k) := s_i + w_j$ with $1 \leq i, j \leq k$. We note that the rank of the collection of linear forms $\ell^{(i,j)}$ (inside the space of all $\mathbb{C}$-linear forms on $\mathbb{C}^{2k}$) is $2k - 1$. Therefore it follows from La Bretèche's result that (8) is equal to

$$(1 + o(1))C_k X^k (\log X)^{k^2 - (2k-1)}.$$

Using Théorème 2 in La Bretéche's work allows us to recover the precise value of C_k. Indeed, according to Théorème 2 in [4] we get that (8) is equal to

$$(1 + o(1))H(\tfrac{1}{2}, \dots, \tfrac{1}{2})\mathrm{Vol}(A_k(X))$$

where $A_k(X)$ is a subset of $[1, \infty)^{k^2}$ corresponding to tuples $(a_{i,j}) \in [1, \infty)^{k^2}$ with $1 \leq i, j \leq k$ such that

$$\text{for each } j \leq k : \prod_{1 \leq i \leq k} a_{i,j} \leq X$$

$$\text{and for each } i \leq k : \prod_{1 \leq j \leq k} a_{i,j} \leq X$$

Therefore it remains to understand the asymptotic behavior of

$$\mathrm{Vol}(A_k(X))$$

as $X \to \infty$. Surprisingly, this is somewhat involved, and the rest of the proof is devoted to that.

Proposition 4.3. *Let $k \geq 2$ be fixed. Then,*

$$Vol(A_k(X)) \sim \binom{2k-2}{k-1} k^{-(k-1)} \cdot Vol(\mathscr{B}_k) \cdot X^k \cdot (\log X)^{(k-1)^2}$$

where $Vol(\mathscr{B}_k)$ corresponds to the $(k-1)^2$ dimensional volume of the Birkhoff polytope $\mathscr{B}_k \subset \mathbb{R}^{k^2}$.

The proof of the Proposition depends on the following Lemma.

Lemma 4.4. *Let $n \geq 1$ be fixed. Then as $X \to \infty$ we have*

$$\iint_{\substack{0\leq x_1,\dots,x_n\leq \log X\\ 0\leq y_1,\dots,y_n\leq \log X}} \exp\Big(\min(x_1+\dots+x_n, y_1+\dots+y_n)\Big)dx_1\dots dy_n \sim \binom{2n}{n}X^n.$$

Proof. Making the substitutions $v_i = \log X - x_i$ and $w_i = \log X - y_i$ in Lemma 4.4, we see the integral there is the same as

$$X^n \iint_{\substack{0\leq v_1,\dots,v_n\leq \log X\\ 0\leq w_1,\dots,w_n\leq \log X}} \exp\Big(-\max(v_1+\dots+v_n, w_1+\dots+w_n)\Big)dv_1\dots dw_n.$$

Here we can extend all the ranges of integration up to positive infinity, at the cost of a multiplicative error term $1+o(1)$. Then by symmetry

$$\iint_{\substack{0\leq v_1,\dots,v_n\\ 0\leq w_1,\dots,w_n}} \exp\Big(-\max(v_1+\dots+v_n, w_1+\dots+w_n)\Big)dv_1\dots dw_n$$
$$= 2\int_{0\leq v_1,\dots,v_n} \exp\Big(-(v_1+\dots+v_n)\Big)\int_{w_1+\dots+w_n\leq v_1+\dots+v_n} dv_1\dots dw_n,$$

and making the further substitution $v = v_1+\dots+v_n$ in the integral, we see the above is

$$= 2\int_0^\infty e^{-v}\left(\int_{v_1+\dots+v_{n-1}\leq v} dv_1\dots dv_{n-1}\right)\left(\int_{w_1+\dots+w_n\leq v} dw_1\dots dw_n\right)dv$$
$$= 2\int_0^\infty e^{-v}v^{2n-1}\left(\int_{v_1+\dots+v_{n-1}\leq 1} dv_1\dots dv_{n-1}\right)\left(\int_{w_1+\dots+w_n\leq 1} dw_1\dots dw_n\right)dv.$$

Here the two integrals in brackets are simply the volume of the standard $n-1$ simplex and the standard n simplex, which are well known to be $1/(n-1)!$ and $1/n!$, respectively. Therefore the above integral is equal to

$$\frac{2}{(n-1)!n!}\int_0^\infty e^{-v}v^{2n-1}dv = 2\binom{2n-1}{n} = \binom{2n}{n}.$$

We conclude that the integral in the statement of Lemma 4.4 is equal to (as $X\to\infty$),

$$(1+o(1))\binom{2n}{n}X^n$$

as claimed.

We are now ready to prove the Proposition, and thus finish the proof of Theorem 1.3.

Proof (Proof of Proposition 4.3). Notice first that, if we set $u_{i,j} = \log a_{i,j}$, and if we write

$$c_j = \sum_{1\le i\le k} u_{i,j} \text{ and } r_i = \sum_{1\le j\le k} u_{i,j}$$

for all $i, j \le k$, then we find

$$\mathrm{Vol}(A_k(X)) = \int_{(u_{i,j})_{1\le i,j\le k}\subseteq[0,\infty)^{k^2}: c_j, r_i\le \log X\ \forall i,j\le k} \exp\left(\sum_{i,j\le k} u_{i,j}\right) du_{1,1}\dots du_{k,k}.$$

To prove the proposition we shall obtain upper and lower bounds for the integral on the right that are asymptotically equal.

For convenience of writing, we start by introducing a little more notation. Let $S_{k-1} := \sum_{i,j\le k-1} u_{i,j}$. Let also $\mathscr{U}_{k,\varepsilon}(X)$ be the set of $u_{i,j}$ with $i, j \le k-1$ for which

$$\sum_{i\le k-1} u_{i,j} \le \log X \text{ and } \sum_{j\le k-1} u_{i,j} \le \log X \text{ and } \sum_{i,j\le k-1} u_{i,j} > (k-2-\varepsilon)\log X.$$

Considering the vector $\mathbf{u}$ of $u_{i,j}$ with $i, j \le k-1$ as fixed, let $\mathscr{T}_{C,k}(\mathbf{u}, X)$ be the set of those $u_{k,i}$ with $i \le k-1$ for which

$$c_j \le \log X \text{ for all } j \le k-1$$

Finally, again considering the $u_{i,j}$ with $i, j \le k-1$ as fixed let $\mathscr{T}_{R,k}(\mathbf{u}, X)$ be the set of those $u_{j,k}$ with $j \le k-1$ for which

$$r_i \le \log X \text{ for all } i \le k-1.$$

We set $\epsilon = 1/\sqrt{\log X}$, say. First seeking an upper bound, we note that if we have $S_{k-1} \le (k-2-\epsilon)\log X$ then $S_k \le (k-\epsilon)\log X$, and therefore the part of the integral where $S_{k-1} \le (k-2-\epsilon)\log X$ contributes at most

$$X^{k-\epsilon} \cdot \int_{(u_{i,j})_{1\le i,j\le k}\subseteq[0,\infty)^{k^2}: c_j, r_i\le \log X\ \forall i,j\le k} 1 du_{1,1}\dots du_{k,k} \le X^{k-\epsilon}\log^{k^2} X.$$

This is asymptotically negligible (for any fixed k) by our choice of ϵ. Meanwhile, the part of the integral where $S_{k-1} > (k-2-\epsilon)\log X$ is equal to

$$\int_{\mathscr{U}_{k,\varepsilon}(X)} \exp(S_{k-1}) \int_{\mathscr{T}_{C,k}(\mathbf{u},X)} \exp\left(u_{k,1}+\ldots+u_{k,k-1}\right)\times \tag{9}$$

$$\times \int_{\mathscr{T}_{R,k}(\mathbf{u},X)} \exp\left(u_{1,k}+\ldots+u_{k-1,k}\right) \int_{u_{k,k}: c_k, r_k\le \log X} \exp(u_{k,k})\, du_{1,1}\dots du_{k,k}.$$

Here the innermost integral is over those

$$0 \le u_{k,k} \le \log X - \max(u_{k,1} + \ldots + u_{k,k-1}, u_{1,k} + \ldots + u_{k-1,k}),$$

assuming the upper range of integration is at least zero. Therefore the innermost integral is certainly bounded above (extending the lower limit to negative infinity, and then performing the integration) by

$$X \exp\Big(- \max(u_{k,1} + \ldots + u_{k,k-1}, u_{1,k} + \ldots + u_{k-1,k})\Big).$$

Substituting this in, it follows that (9) is less than

$$X \int_{\mathscr{U}_{k,\varepsilon}(X)} \int_{\mathscr{T}_{C,k}(\mathbf{u},X)} \int_{\mathscr{T}_{R,k}(\mathbf{u},X)} \exp\Big(\min(\sum_{1\le j\le k-1} c_j, \sum_{1\le i\le k-1} r_i)\Big) \prod_{(i,j)\ne(k,k)} du_{i,j}. \tag{10}$$

At this point we change variables, letting $r_1, \ldots, r_{k-1}$ and $c_1, \ldots, c_{k-1}$ run through the interval $[0, \log X]$ so that $u_{i,k} = r_i - \sum_{1\le j\le k-1} u_{i,j}$ and $u_{k,j} = c_j - \sum_{1\le i\le k-1} u_{i,j}$. Since $u_{i,k} \ge 0$ and $u_{k,j} \ge 0$ this change of variable implies the additional condition that for all $i, j \le k-1$,

$$\sum_{j\le k-1} u_{i,j} \le r_i \text{ and } \sum_{i\le k-1} u_{i,j} \le c_j \tag{11}$$

The Jacobian of this linear change of variable is equal to 1 since the linear transformation taking the $(u_{i,j})$ with $(i,j) \ne (k,k)$ into $(r_\ell, c_\ell, u_{i,j})$ with $i, j, \ell \le k-1$ is upper triangular with only 1's on the diagonal.

Given $\mathbf{r} = (r_1, \ldots, r_{k-1})$ and $\mathbf{c} = (c_1, \ldots, c_{k-1})$ we let $\tilde{\mathscr{U}}_{k,\varepsilon}(\mathbf{r}, \mathbf{c}, X)$ be the set of $u_{i,j}$ with $i, j \le k-1$ satisfying the conditions (11) and the standing condition that

$$\sum_{i,j\le k-1} u_{i,j} \ge (k - 2 - \varepsilon) \log X, \tag{12}$$

and we let $\tilde{\mathscr{T}}_k(X)$ be the set of $0 \le r_1, \ldots, r_{k-1} \le \log X$ and $0 \le c_1, \ldots, c_{k-1} \le \log X$. Then (10) can be re-written as

$$X \int_{\tilde{\mathscr{T}}_k(X)} \exp\Big(\min(\sum_{1\le j\le k-1} c_j, \sum_{1\le i\le k-1} r_i)\Big) \mathrm{Vol}\Big(\tilde{\mathscr{U}}_{k,\varepsilon}(\mathbf{r}, \mathbf{c}, X)\Big) \prod_{i\le k-1} dc_i\, dr_i \tag{13}$$

Since $r_i, c_i \le \log X$ for all $i \le k-1$, we have

$$\begin{aligned}\mathrm{Vol}(\tilde{\mathscr{U}}_{k,\varepsilon}(\mathbf{r}, \mathbf{c}, X)) &\le \mathrm{Vol}(\tilde{\mathscr{U}}_{k,\varepsilon}(\log \mathbf{X}, \log \mathbf{X}, \mathbf{X})) \\ &= (\log X)^{(k-1)^2} \cdot \mathrm{Vol}(\tilde{\mathscr{U}}_{k,\varepsilon}(\mathbf{1}, \mathbf{1}, e)) \sim (\log X)^{(k-1)^2} \cdot \mathrm{Vol}(\tilde{\mathscr{U}}_{k,0}(\mathbf{1}, \mathbf{1}, e))\end{aligned}$$

as $X \to \infty$, where $\log \mathbf{X} := (\log X, \dots, \log X)$ and $\mathbf{1} := (1, \dots, 1)$, and where we recall for the final asymptotic that $\epsilon = 1/\sqrt{\log X}$. As already mentioned in the introduction $\mathrm{Vol}(\tilde{\mathscr{U}}_{k,0}(\mathbf{1}, \mathbf{1}, e)) = k^{-(k-1)}\mathrm{Vol}(\mathscr{B}_k)$ where $\mathscr{B}_k$ is the Birkhoff polytope. It follows that (10) is

$$\leq (1 + o(1))X(\log X)^{(k-1)^2} \cdot k^{-(k-1)}\mathrm{Vol}(\mathscr{B}_k)$$
$$\times \int_{\tilde{\mathscr{T}}_k(X)} \exp\Big(\min\Big(\sum_{1\leq j\leq k-1} c_j, \sum_{1\leq i\leq k-1} r_i\Big)\Big) \prod_{i\leq k-1} dc_i dr_i$$

and by Lemma 4.4 this is less than or equal to

$$(1 + o(1))X(\log X)^{(k-1)^2} \cdot k^{-(k-1)}\mathrm{Vol}(\mathscr{B}_k) \cdot X^{k-1}\binom{2k-2}{k-1}$$

thus finishing the proof of the upper bound.

For the lower bound we restrict attention, as we may (due to positivity), to the part of the integral where $S_{k-1} > (k-2+\varepsilon)\log X$ and each r_i, c_i is $\geq (1-\varepsilon)\log X$. The point of the former condition is that if it is satisfied then

$$u_{1,k} + u_{2,k} + \ldots + u_{k-1,k} \leq (k-1)\log X - S_{k-1} \leq (1-\varepsilon)\log X$$

and similarly $u_{k,1} + u_{k,2} + \ldots + u_{k,k-1} \leq (1-\varepsilon)\log X$, and therefore

$$\log X - \max\Big(u_{k,1} + \ldots + u_{k,k-1}, u_{1,k} + \ldots + u_{k-1,k}\Big) > \varepsilon \log X = \sqrt{\log X} \to \infty.$$

Therefore arguing as above the innermost integral over $u_{k,k}$ in (9) contributes

$$(1 + o(1))X \exp\Big(-\max(u_{k,1} + \ldots + u_{k,k-1}, u_{1,k} + \ldots + u_{k-1,k})\Big).$$

Proceeding as before we thus arrive to (13) but with the additional condition that $(1-\varepsilon)\log X < r_i, c_i < \log X$ (and with the condition that $\sum_{i,j\leq k-1} u_{i,j} \geq (k-2-\varepsilon)\log X$ replaced by the condition that $\sum_{i,j\leq k-1} u_{i,j} \geq (k-2+\varepsilon)\log X$). It follows that on this set of r_i and c_i we have

$$\mathrm{Vol}(\tilde{\mathscr{U}}_{k,\varepsilon}(\mathbf{r}, \mathbf{c}, X)) > \mathrm{Vol}(\tilde{\mathscr{U}}_{k,\varepsilon}((\mathbf{1}-\boldsymbol{\varepsilon})\log \mathbf{X}, (\mathbf{1}-\boldsymbol{\varepsilon})\log \mathbf{X}, X))$$
$$= (1 + o(1))(\log X)^{(k-1)^2} \cdot k^{-(k-1)}\mathrm{Vol}(\mathscr{B}_k)$$

Therefore we obtained the following lower bound

$$(1 + o(1))X(\log X)^{(k-1)^2} \cdot k^{-(k-1)}\mathrm{Vol}(\mathscr{B}_k)$$
$$\times \iint_{\tilde{\mathscr{T}}_{k,\varepsilon}(X)} \exp\Big(\min\Big(\sum_{1\leq j\leq k-1} c_j, \sum_{1\leq i\leq k-1} r_i\Big)\Big) \prod_{i\leq k-1} dc_i\, dr_i$$

where $\tilde{\mathscr{T}}_{k,\varepsilon}(X)$ is the set of r_i, c_i satisfying $(1-\varepsilon)\log X < r_i, c_i \le \log X$ for all $i \le k-1$. Note that the condition $(1-\varepsilon)\log X < r_i, c_i$ can be dropped. Indeed the contribution to the integral of any tuple of $(r_1, \dots, r_{k-1})$ or $(c_1, \dots, c_{k-1})$ where at least one of the c_i, r_i is $\le (1-\varepsilon)\log X$ is $\le X^{k-1-\varepsilon}$ and therefore negligible. Thus we can extend the integration to all of $c_i, r_i \le \log X$. Because of this Lemma 4.4 is applicable and we have therefore obtained the lower bound

$$\ge (1+o(1))\binom{2k-2}{k-1} \cdot k^{-(k-1)}\mathrm{Vol}(\mathscr{B}_k)X^k \cdot (\log X)^{(k-1)^2}$$

as claimed. Since we have obtained asymptotically matching upper and lower bounds the proof of the proposition is finished.

5 Proof of Theorem 1.4

In the Rademacher case we have, letting $\square$ denote a generic square,

$$\begin{aligned}\mathbb{E}\Big(\sum_{n\le X} f(n)\Big)^k &= \sum_{n_1,\dots,n_k\le X} \mathbb{E}[f(n_1)\dots f(n_k)] \\ &= \sum_{\substack{n_1,\dots,n_k\le X \\ n_1\dots n_k=\square}} \mu^2(n_1)\dots\mu^2(n_k)\end{aligned}$$

Let $g(n_1, \dots, n_k)$ be a multiplicative function of several variables, supported on square-free n_i, and such that $g = 1$ when $n_1 \dots n_k = \square$ and $g = 0$ otherwise. Then we find that the Dirichlet series

$$\sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \frac{g(n_1,\dots,n_k)}{n_1^{s_1}\dots n_k^{s_k}}$$

is equal to

$$\prod_p \Big(1 + \sum_{\substack{0\le\alpha_1,\dots,\alpha_k\le 1 \\ \alpha_1+\dots+\alpha_k\equiv 0 \mod 2}} \frac{1}{p^{\alpha_1 s_1+\dots+\alpha_k s_k}}\Big).$$

This factors as

$$H(s_1,\dots,s_k)\prod_{1\le i<j\le k}\zeta(s_i+s_j)$$

with

$$H(\tfrac{1}{2},\dots,\tfrac{1}{2}) = \prod_p \Big(1-\frac{1}{p}\Big)^{k(k-1)/2}\Big(1+\sum_{1\le j\le k/2}\frac{\binom{k}{2j}}{p^j}\Big)$$

The main result of La Bretèche is applicable with $\binom{k}{2}$ linear forms $\ell^{(i,j)}(s_1,s_2,\dots,s_k) = s_i + s_j$ defined for $1 \le i < j \le k$. The rank of these linear forms is equal to k for $k \ge 3$ (for $k = 2$ the rank is equal to 1 since there is only one form in that case). Therefore applying La Bretèche's result it follows that the moment is asymptotically

$$(1+o(1))C_k X^{k/2}(\log X)^{\binom{k}{2}-k}.$$

In order to determine the constant C_k one could use Théorème 2 of La Bretèche, to conclude that the moment is asymptotically

$$(1+o(1))H(\tfrac{1}{2},\dots,\tfrac{1}{2})\mathrm{Vol}(B(X))$$

where $B(X)$ is the set of $(u_{i,j})_{i<j} \in \mathbb{R}^{k(k-1)/2}$ such that

$$\text{for all } 1 \le i \le k: \ \prod_{j<i} u_{j,i}\prod_{i<j} u_{i,j} \le X$$

and then proceed in a manner similar to Theorem 1.3. However we leave this computation to the interested reader.

Acknowledgements We are grateful to the referee for a careful reading of the paper and for asking several questions which led to Theorem 1.4 and stronger results in Theorem 1.3.

The first author is supported by a research fellowship at Jesus College, Cambridge.

References

1. M. Beck, D. Pixton, The Erhart polynomial of the Birkhoff polytope. Discrete Comput. Geom. **30**(4), 623–637 (2003)
2. A. Bondarenko, K. Seip, Helson's problem for sums of a random multiplicative function. Preprint available online at http://www.arxiv.org/abs/1411.6388
3. A. Bondarenko, W. Heap, K. Seip, An inequality of Hardy–Littlewood type for Dirichlet polynomials. J. Number Theory **150**, 191–205 (2015)
4. R. de la Bretèche, Estimation de sommes multiples de fonctions arithmétiques. Compos. Math. **128**(3), 261–298 (2001)
5. E.R. Canfield, B.D. McKay, The asymptotic volume of the Birkhoff polytope. Online J. Anal. Comb. (4), 4 pp. (2009)
6. C.S. Chan, D.P. Robbins, On the volume of the polytope of doubly stochastic matrices. Exp. Math. **8**(3), 291–300 (1999)

7. S. Chatterjee, K. Soundararajan, Random multiplicative functions in short intervals. Int. Math. Res. Not. **3**, 479–492 (2012)
8. B. Conrey, A. Gamburd, Pseudomoments of the Riemann zeta function and pseudomagic squares. J. Number Theory **117**(2), 263–278 (2006)
9. B. Conrey, S. Gonek, High moments of the Riemann zeta-function. Duke Math. J. **107**(3), 577–604 (2001)
10. A. Gut, *Probability: A Graduate Course*, 2nd edn. Springer Texts in Statistics (Springer, Berlin, 2013)
11. G. Halász, On random multiplicative functions, in *Hubert Delange Colloquium, (Orsay, 1982)*. Publications Mathématiques d'Orsay, vol. 83 (University of Paris XI, Orsay, 1983), pp. 74–96
12. A.J. Harper, Bounds on the suprema of Gaussian processes, and omega results for the sum of a random multiplicative function. Ann. Appl. Probab. **23**(2), 584–616 (2013)
13. A.J. Harper. On the limit distributions of some sums of a random multiplicative function. J. Reine Angew. Math. **678**, 95–124 (2013)
14. A.J. Harper, A note on the maximum of the Riemann zeta function, and log-correlated random variables. Preprint available online at http://www.arxiv.org/abs/1304.0677
15. W. Heap, S. Lindqvist, Moments of random multiplicative functions and truncated characteristic polynomials http://www.arxiv.org/abs/1505.03378
16. H. Helson, Hankel forms. Stud. Math. **198**(1), 79–84 (2010)
17. B. Hough, Summation of a random multiplicative function on numbers having few prime factors. Math. Proc. Camb. Philos. Soc. **150**, 193–214 (2011)
18. Y.-K. Lau, G. Tenenbaum, J. Wu, On mean values of random multiplicative functions. Proc. Am. Math. Soc. **141**, 409–420 (2013)
19. S.R. Louboutin, M. Munsch, The second and fourth moments of theta functions at their central point. J. Number Theory **133**(4), 1186–1193 (2013)
20. H.L. Montgomery, R.C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, 1st edn. (Cambridge University Press, Cambridge, 2007)
21. N. Ng, The distribution of the summatory function of the Möbius function. Proc. Lond. Math. Soc. **89**(3), 361–389 (2004)
22. J. Ortegà-Cerda, K. Seip, A lower bound in Nehari's theorem on the polydisc. J. Anal. Math. **118**(1), 339–342 (2012)
23. I. Pak, Four questions on Birkhoff polytope. Ann. Comb. **4**, 83–90 (2000)
24. A. Wintner, Random factorizations and Riemann's hypothesis. Duke Math. J. **11**, 267–275 (1944)

Large Values of the Zeta-Function on the Critical Line

Aleksandar Ivić

Dedicated to Professor Helmut Maier on the occasion of his 60th birthday

Abstract This is primarily an overview article (Lecture given during the conference "Number Theory and its Applications Workshop" in Xi'an (China), October 23–28, 2014.) dealing with the large values of $|\zeta(\frac{1}{2}+it)|$. This approach allows one to obtain upper bounds for moments (mean values) of $|\zeta(\frac{1}{2}+it)|$, which is one of the fundamental problems of the theory of the Riemann zeta-function. A sketch of the upper bound for the 12th moment of D.R. Heath-Brown (Q J Math (Oxford) 29:443–462, 1978) is presented, together with some recent results of the author. They include upper bounds obtained by the use of large values of $E^*(T)$, a function closely related to the classical function $E(T)$, the error term in the mean square formula for $|\zeta(\frac{1}{2}+it)|$. A new large values result involving $E_k(T)$, the general error-term function in the formula for the $2k$-th moment of $|\zeta(\frac{1}{2}+it)|$, is also given.

Keywords Riemann zeta-function • Moments • Large values

AMS Mathematics Subject Classification (2010) 11M06

1 The Basic Properties of $\zeta(s)$

1.1 The Definitions

The Riemann zeta-function is defined by

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p} (1-p^{-s})^{-1} \tag{1}$$

A. Ivić (✉)
Serbian Academy of Science and Arts, Knez Mihailova 35, 11000 Beograd, Serbia
e-mail: aleksandar.ivic@rgf.bg.ac.rs; aivic_2000@yahoo.com

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_12

for $\Re s > 1$, where p denotes primes. For other values of the complex variable $s = \sigma + it\,(\sigma, t \in \mathbb{R})$ it is defined by analytic continuation. It is regular for $s \in \mathbb{C}$, except at $s = 1$ where it has a simple pole with residue 1. The product representation in (1) shows that $\zeta(s)$ does not vanish for $\sigma > 1$. The *Laurent expansion of* $\zeta(s)$ at $s = 1$ reads

$$\zeta(s) = \frac{1}{s-1} + \gamma_0 + \gamma_1(s-1) + \gamma_2(s-1)^2 + \cdots,$$

where the so-called *Stieltjes constants* γ_k are given by

$$\gamma_k = \frac{(-1)^k}{k!} \lim_{N\to\infty}\left(\sum_{m\leqslant N} \frac{\log^k m}{m} - \frac{\log^{k+1} N}{k+1}\right) \qquad (k = 0, 1, 2, \ldots),$$

and $\gamma = \gamma_0 = -\Gamma'(1) = 0.5772157\ldots$ is the *Euler constant*. For more properties of $\zeta(s)$ the reader should consult the monographs of Titchmarsh [47], Edwards [9], and the author [20].

A notable feature of $\zeta(s)$, whose analogues are true for many *Dirichlet series* (or L-functions) of the form

$$F(s) = \sum_{n=1}^{\infty} a_F(n) n^{-s} \qquad (\Re s > 1), \tag{2}$$

is the functional equation, proved first by Riemann [43] in 1859. In a symmetric form it says, that for $s \in \mathbb{C}$,

$$\pi^{-s/2}\zeta(s)\Gamma(\tfrac{1}{2}s) = \pi^{-(1-s)/2}\zeta(1-s)\Gamma(\tfrac{1}{2}(1-s)). \tag{3}$$

Alternatively we can write (3) as

$$\zeta(s) \;=\; \chi(s)\zeta(1-s),$$

where

$$\chi(s) := \frac{\Gamma(\frac{1}{2}(1-s))}{\Gamma(\frac{1}{2}s)}\pi^{s-1/2},$$

and $\Gamma(s)$ is the familiar gamma-function. Using the classical *Stirling formula* for $\Gamma(s)$, we have

$$\chi(s) = \left(\frac{2\pi}{t}\right)^{\sigma+it-1/2} e^{i(t+\pi/4)}\left(1 + O\left(\frac{1}{t}\right)\right) \qquad (s = \sigma + it,\ t \geqslant 2).$$

1.2 The Selberg Class

The best known generalization of $\zeta(s)$ is the *Selberg class* $\mathscr{S}$ of Dirichlet series (see the overview papers of Bombieri [2] and Kaczorowski [36]), defined axiomatically by Selberg [45] as follows:

(a) The series in (2) converges absolutely for $\sigma = \Re s > 1$ with $a_F(1) = 1$, and the coefficients $a_F(n)$ satisfy the bound $a_F(n) \ll_\varepsilon n^\varepsilon$ for every $\varepsilon > 0$.
(b) There is a natural number m_F such that $(s-1)^{m_F} F(s)$, which is of a finite order, admits analytic continuation which is regular in $\mathbb{C}$.
(c) There is a function

$$\Phi_F(s) := Q_F^s \Gamma_F(s) F(s) \qquad (Q_F > 0)$$

with

$$\Gamma_F(s) := \prod_{j=1}^{r_F} \Gamma\Big(\lambda_j(F)s + \mu_j(F)\Big) \quad (\lambda_j(F) > 0,\ \Re\mu_j(F) \geqslant 0).$$

Then the functional equation

$$\Phi_F(s) = \omega_F \overline{\Phi}_F(1-s) \qquad (|\omega_F| = 1) \tag{4}$$

holds for $s \in \mathbb{C}$, where for any function $f(s)$ we define $\bar{f}(s) = \overline{f(\bar{s})}$.
(d) We have

$$\log F(s) = \sum_{n=2}^{\infty} b_F(n)\Lambda(n)/(n^s \log n),$$

where $b_F(n) \ll n^\theta$ for some $0 \leqslant \theta < \frac{1}{2}$, and $\Lambda(n)$ is the *von Mangoldt function*, namely $\Lambda(n) = \log p$ if $n = p^\alpha$ $(\alpha \in \mathbb{N})$ and $\Lambda(n) = 0$ otherwise.

The quantity

$$d_F := 2\sum_{j=1}^{r_F} \lambda_j(F)$$

plays an important rôle in the theory of $\mathscr{S}$. It is called *the degree of F* in $\mathscr{S}$, it does not depend on the form of the functional equation (4), and it is conjectured that it is always a non-negative integer. Broadly speaking, anything that holds for $\zeta(s)$ can be to some extent generalized to elements of $\mathscr{S}$.

Notation. Owing to the nature of this text, absolute consistency in notation could not be attained, although whenever possible standard notation is used.

By $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$ we denote the set of natural numbers, integers, real and complex numbers, respectively. The symbol ε will denote arbitrarily small positive numbers, not necessarily the same ones at each occurrence. The symbols $f(x) = O\big(g(x)\big)$ and $f(x) \ll g(x)$ both mean that $|f(x)| \leqslant Cg(x)$ for some constants $C > 0$ and $x \geqslant x_0 > 0$. By $f(x) \ll_{a,b,\dots} g(x)$ we mean that the constant implied by the $\ll$-symbol depends on $a, b, \dots$, while $a \asymp b$ means that both $a \ll b$ and $b \ll a$ hold.

2 Moments on the 1/2-Line

2.1 Basic Concepts

In zeta-function theory, of particular interest are the *moments* (or *mean values*)

$$I_k(T) := \int_0^T |\zeta(\tfrac{1}{2} + it)|^{2k}\,\mathrm{d}t.$$

Although the integral certainly makes sense if $\Re k > 0$, usually one takes $k \in \mathbb{N}$ and supposes that k is fixed. Namely in this case one can use the obvious identity $|z|^2 = z\cdot\bar{z}$, and if one has a "good" expression for $z = \zeta^k(\frac{1}{2}+it)$, then by multiplying the expressions for z and $\bar{z}$ one can tackle $I_k(T)$, at least in principle. There are two monographs dedicated solely to the subject of mean values of $\zeta(s)$: the author's [18] and Ramachandra's [42]; see also the author's recent overview paper [29] and Joyner [33].

The expression

$$I_k(T+H) - I_k(T) = \int_T^{T+H} |\zeta(\tfrac{1}{2} + it)|^{2k}\,\mathrm{d}t$$

denotes zeta-moments in "short intervals" if $1 \leqslant H < T$ and $H = o(T)$ as $T \to \infty$. It is also an important object of study.

We have (Balasubramanian–Ramachandra; see [42])

$$|\zeta(\tfrac{1}{2} + it)|^k \ll_k \log t \left(\int_{t-c}^{t+c} |\zeta(\tfrac{1}{2} + ir)|^k\,\mathrm{d}r + 1\right) \qquad (k \in \mathbb{N},\ c > 0). \tag{5}$$

This bound allows one to transform *discrete zeta sums* into sums of integrals over short intervals. In this process one loses only an (unimportant) logarithmic factor.

Conversely, one trivially can transform integrals of continuous functions $f(x)$ into discrete sums by means of the elementary inequality

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leqslant (b-a)\max_{x\in[a,b]} |f(x)| \qquad (a < b).$$

2.2 Lower Bounds for $\zeta(\frac{1}{2}+it)$ in Short Intervals

Ramachandra [42] showed unconditionally that one has

$$\int_T^{T+H} |\zeta(\tfrac{1}{2}+it)|^{2k}\,\mathrm{d}t \;\gg_k\; H(\log H)^{k^2} \qquad (k\in\mathbb{N},\ \log\log T \ll H \leqslant T). \tag{6}$$

He used complex integration and the fact that, for $w = u+iv$ $(u, v \in \mathbb{R};\ |u| < \pi/2)$,

$$|\exp(-\cos w)| = \left|\exp\Big(-\tfrac{1}{2}(e^{iw}+e^{-iw})\Big)\right|$$
$$= \left|\exp\Big(-\tfrac{1}{2}(e^{iu}e^{-v}+e^{-iu}e^{v})\Big)\right| = \exp(-\cos u\cdot\cosh v).$$

The above kernel $\exp(-\cos w)$, which decreases like a *second order exponential* in $|v|$, sets the limit to the lower bound for H in (6) (as a multiple of $\log\log T$). According to W.K. Hayman, this is essentially the best such kernel function, but it is an open problem whether, by some method different from Ramachandra's, one can improve the range for H in (6).

Recently Radziwill and Soundararajan [41] showed that (unconditionally)

$$\int_0^T |\zeta(\tfrac{1}{2}+it)|^{2k}\,\mathrm{d}t \geqslant e^{-30k^4}T(\log T)^{k^2}$$

holds for any real $k > 1$ and $T \geqslant T_0$. This bound not only holds for any $k > 1$, but it is explicit and at the same time continuous in k, although the constant e^{-30k^4} is certainly not the best one possible.

2.3 Upper Bounds for $\zeta(\frac{1}{2}+it)$ in Short Intervals Under the RH

Upper bounds for $I_k(T)$ are much more difficult to obtain than lower bounds. In 2010 Soundararajan [46] showed that, under the Riemann Hypothesis (RH, that all complex zeros of $\zeta(s)$ have real parts 1/2),

$$\int_0^T |\zeta(\tfrac{1}{2}+it)|^{2k}\,\mathrm{d}t \;\ll_\varepsilon\; T(\log T)^{k^2+\varepsilon}, \tag{7}$$

which is valid for any fixed $k > 0$ and any given $\varepsilon > 0$. This result, when $H = T$, comes very close to Ramachandra's unconditional lower bound (6).

The author [25] obtained an improvement and generalization of (7), using Soundararajan's method: let $H = T^\theta$ where $0 < \theta \leqslant 1$ is a fixed number, and let k be a fixed positive number. Then, again under the RH, we have

$$\int\limits_{T-H}^{T+H} |\zeta(\tfrac{1}{2}+it)|^{2k}\,\mathrm{d}t \ \ll\ H(\log T)^{k^2\left(1+O(1/\log_3 T)\right)},$$

where

$$\log_3 T := \log\log\log T.$$

Later in 2013 Harper [11] obtained the conjectured upper bound (also under the RH) for any fixed $k \geqslant 0$, namely

$$\int_0^T |\zeta(\tfrac{1}{2}+it)|^{2k}\,\mathrm{d}t \ \ll\ T(\log T)^{k^2}. \tag{8}$$

Harper's method can be carried over to obtain the corresponding improvement for $I_k(T+H)-I_k(T)$ as well. Essentially this result is best possible, since for $H=T$ Ramachandra's lower bound (6) (which is unconditional) for $I_k(T)$ is of the same form as the right-hand side of (8).

3 The Conjectural Formula for $I_k(T)$

It is generally conjectured that

$$I_k(T) = Tp_{k^2}(\log T) + E_k(T), \tag{9}$$

where $p_{k^2}(y)$ is an (explicit) polynomial in y whose coefficients depend on k. The function $E_k(T)$ is thought of as an *error term function*, in the sense that one expects, for fixed $k \in \mathbb{N}$,

$$E_k(T) \ = \ o(T) \qquad (T\to\infty).$$

Unfortunately, so far an asymptotic formula of the above shape is known, despite many efforts, to hold only in the cases when $k=1$ or $k=2$. The case $k=1$ (namely, the function $E_1(T)\equiv E(T)$) is extensively discussed in [18, 20], while for $k=2$ (the function $E_2(T)$) see [30, 31, 40].

There is, however, a *conjectural formula* for $I_k(T)$ which of a similar shape as (9). The formula in question is due to Conrey, Farmer, Keating, Rubinstein and Snaith [7, 8, 37]. It is of the form

$$I_k(T) = \int_0^T |\zeta(\tfrac{1}{2}+it)|^{2k}\,\mathrm{d}t = (1+o(1))\int_0^T P_k\left(\log\Big(\frac{t}{2\pi}\Big)\right)\,\mathrm{d}t \quad (T\to\infty). \tag{10}$$

In (10) $P_k(x)$ is the polynomial of degree k^2, all of whose coefficients depend on the (fixed) integer $k \geqslant 1$, given explicitly by the $2k$-fold residue

$$P_k(x) = \frac{(-1)^k}{k!^2} \frac{1}{(2\pi i)^{2k}} \int_{|z_1|=\varepsilon_1} \cdots \int_{|z_{2k}|=\varepsilon_{2k}} \times$$
$$\times \frac{G(z_1, \ldots, z_{2k}) \Delta^2(z_1, \ldots, z_{2k})}{\prod_{i=1}^{2k} z_i^{2k}} \exp\Big\{\frac{1}{2} x \sum_{i=1}^{k} (z_i - z_{i+k})\Big\} \, \mathrm{d}z_1 \ldots \mathrm{d}z_{2k},$$

where the ε_i's are small positive numbers. We have

$$\Delta(z_1, \ldots, z_m) = \prod_{1 \leqslant i < j \leqslant m} (z_j - z_i) = |z_i^{j-1}|_{m \times m},$$

which is the *Vandermonde determinant*,

$$G(z_1, \ldots, z_{2k}) = A_k(z_1, \ldots, z_{2k}) \prod_{i=1}^{k} \prod_{j=1}^{k} \zeta(1 + z_i - z_{j+k}),$$

and finally A_k is the *Euler product* (as usual, $\mathrm{e}(\theta) := \exp(2\pi i \theta)$)

$$A_k(z_1, \ldots, z_{2k}) = \prod_{p} \prod_{i,j=1}^{k} (1 - p^{-1-z_i+z_{k+j}}) \times$$
$$\times \int_0^1 \prod_{j=1}^{k} \Big(1 - \mathrm{e}(\theta) p^{-1/2-z_j}\Big)^{-1} \Big(1 - \mathrm{e}(-\theta) p^{-1/2+z_{k+j}}\Big)^{-1} \mathrm{d}\theta.$$

The authors actually conjecture, in all cases, an error term in (10) of the order $O_{k,\varepsilon}(T^{1/2+\varepsilon})$, which in the general case this author finds to be too optimistic. In fact their paper [8] brings forth a conjecture (via characteristic polynomials from *Random Matrix Theory*; see, e.g., the work [39] of M.L. Mehta on random matrices) for the complete main term in the asymptotic formulas for a wide class of L-functions. The coefficients of $P_k(x)$ are given explicitly when $2 \leqslant k \leqslant 7$, and numerically computed moments are compared to the values obtained for the main term by the moments conjecture. We mention that further numerical calculations involving $P_k(x)$ were carried out by Hiary and Odlyzko [13], and Rubinstein and Yamagishi [44]. In all cases when an asymptotic formula for the moment was rigorously proved, the main term in question coincided with the expression predicted by the authors, which renders the conjectural formulas quite important. Heretofore there have been several conjectures for particular L-functions, to mention here just the works of Keating and Snaith [37], and Conrey and Gonek [6] on the conjectural formula for $I_k(T)$.

4 The Large Values Technique for Zeta-Sums

4.1 Introduction

There are several approaches to the study of (unconditional) upper bounds for $I_k(T)$ (or more generally $I_k(T+H)-I_k(T)$). We single out the following:

1. Large values techniques through *Dirichlet polynomials* of the type $\sum_{n\asymp N} f(n)n^{-s}$ and things like the *Halász–Montgomery inequality* (see, e.g., the Appendix of [20]).
2. Spectral theory/Kloostermania. This is the approach developed by N.V. Kuznetsov, Deshouillers–Iwaniec, R. Bruggeman, Y. Motohashi (see his monograph [40]), and others. It leads to good results on $I_2(T)$, namely produces a sharp formula for the (weighted) fourth moment of $|\zeta(\frac{1}{2}+it)|$, from which results on $E_2(T)$ can be derived.
3. The *general binary additive divisor problem*, involving exponential sums with $d_k(n)d_k(n+h)$, where the *shift parameter* h is not fixed, and the *general divisor function* $d_k(n)$ is generated by $\zeta^k(s)$. For the case $k=3$, interesting in the case of the estimation of $I_3(T)$, see the author's work [19].
4. Connecting bounds for $I_k(T)$ to bounds involving $E(T)\equiv E_1(T)$ and the related function $E^*(T)$. This function will be defined in Sect. 7 and used in bounding $I_k(T)$ in Sect. 8.

The purpose of this paper is to focus on the approaches 1. and 4. In particular, a sketch of the proof of Heath-Brown's classical estimate (11) (see [12]) will be given, and this result (up to a log-factor) will be reproved in Sect. 9.

4.2 The Large Values Technique

Suppose t_r $(r=1,\dots,R)$ is a set of points lying in $[T,T+H]$ $(1\ll H\leqslant T)$ such that

$$|\zeta(\tfrac{1}{2}+it_r)|\geqslant V,\ |t_r-t_s|\geqslant G\ (r\neq s,\ 1\ll G\leqslant H;\ r,s=1,\dots,R).$$

The points t_r are then said to be *well-spaced*.

A good upper bound for R will lead to a good upper bound for

$$\sum_{r\leqslant R}|\zeta(\tfrac{1}{2}+it_r)|^k,$$

hence also to a good upper bound for

$$\int_T^{T+H} |\zeta(\tfrac{1}{2}+it)|^k \,\mathrm{d}t,$$

where $k > 0$ is a suitable constant.

This is the underlying idea of the large values technique, which obviously can be used for the estimation of moments of other L-functions, in particular to those belonging to the Selberg class $\mathscr{S}$.

5 Atkinson's Formula for the Mean Square

The formula of Atkinson (1916–2002) [1] for the mean square of $|\zeta(\frac{1}{2}+it)|$ (1949) is a fundamental result in zeta-function theory. It was used later in 1978 by Heath-Brown [12], who employed a large values technique, to prove the 12th moment estimate

$$\int_0^T |\zeta(\tfrac{1}{2}+it)|^{12} \,\mathrm{d}t \;\ll\; T^2(\log T)^{17}. \tag{11}$$

This result is even today essentially the strongest upper bound for $I_k(T)$ when $k \geqslant 6$.

Let

$$E(T) \;:=\; \int_0^T |\zeta(\tfrac{1}{2}+it)|^2 \,\mathrm{d}t - T\left(\log\Big(\frac{T}{2\pi}\Big) + 2\gamma - 1\right).$$

Here $d(n)$ is the number of divisors of n, and $\gamma = -\Gamma'(1) = 0.577215\ldots$ is Euler's constant. The function $E(T)$ is a fundamental function in zeta-function theory, as $I_1(T) = \int_0^T |\zeta(\frac{1}{2}+it)|^2 \,\mathrm{d}t$ is the fundamental moment of $|\zeta(\frac{1}{2}+it)|$. Atkinson's explicit formula for $E(T)$ is

$$E(T) = \Sigma_1(T) + \Sigma_2(T) + O(\log^2 T),$$

$$\Sigma_1(T) := 2^{1/2}(T/(2\pi))^{1/4} \sum_{n\leqslant N} (-1)^n d(n) n^{-3/4} e(T,n) \cos(f(T,n)),$$

$$\Sigma_2(T) := -2 \sum_{n\leqslant N'} \frac{d(n)}{n^{1/2}(\log T/(2\pi n))} \cos\left(T\log\Big(\frac{T}{2\pi n}\Big) - T + \frac{1}{4}\pi\right),$$

$$\begin{aligned} f(T,n) &= 2T\operatorname{ar\,sinh}\left(\sqrt{\pi n/(2T)}\right) + \sqrt{2\pi nT + \pi^2 n^2} - \tfrac{1}{4}\pi \\ &= -\tfrac{1}{4}\pi + 2\sqrt{2\pi nT} + \\ &\quad + \tfrac{1}{6}\sqrt{2\pi^3}\, n^{3/2} T^{-1/2} + a_5 n^{5/2} T^{-3/2} + a_7 n^{7/2} T^{-5/2} + \ldots, \end{aligned}$$

$$e(T,n) = (1+\pi n/(2T))^{-1/4}\Big\{(2T/\pi n)^{1/2}\mathrm{ar\,sinh}\left(\sqrt{\pi n/(2T)}\right)\Big\}^{-1}$$
$$= 1 + O(n/T) \qquad (1 \leqslant n < T),$$

$$\mathrm{ar\,sinh}\, x = \log\left(x + \sqrt{1+x^2}\right),\ N' = \frac{T}{2\pi} + \frac{N}{2} - \sqrt{\frac{N^2}{4} + \frac{NT}{2\pi}}.$$

The sum $\sum_2(T)$ in Atkinson's formula is analogous to the sum approximating $\zeta^2(s)$ (essentially sums of $d(n)n^{-s}$ of length $\asymp |t|$; see, e.g., Chap. 4 of [20]), and by averaging processes (e.g., Gaussian weights $\exp(-(T/G)^2)$) its contribution in many instances may be made fairly small. Its strength lies in the facts that it is explicit, contains two sums with elementary functions and the error term is small.

6 The 12th Moment Estimate

6.1 Technical Preparation

We sketch now the proof of (11), namely the following

Theorem 6.1. *We have*

$$\int_0^T |\zeta(\tfrac{1}{2}+it)|^{12}\,\mathrm{d}t \ \ll\ T^2(\log T)^{17}.$$

First we have, for $T^\varepsilon \leqslant G \leqslant T^{1/2-\varepsilon}, L := \log T$,

$$\int_{T-G}^{T+G} |\zeta(\tfrac{1}{2}+it)|^2\,\mathrm{d}t \leqslant e\int_{-GL}^{GL} |\zeta(\tfrac{1}{2}+iT+it)|^2\exp(-t^2G^{-2})\,\mathrm{d}t.$$

The presence of the exponential factor makes it possible, by using Atkinson's formula and the classical integral

$$\int_{-\infty}^{\infty}\exp(Ax-Bx^2)\,\mathrm{d}x = (\pi/B)^{1/2}\exp(A^2/(4B)) \qquad (\Re B > 0),$$

to obtain the following

Lemma 6.1.

$$\int_{T-G}^{T+G} |\zeta(\tfrac{1}{2}+it)|^2\,\mathrm{d}t \ll G\log T$$
$$+ G\sum_K (TK)^{-1/4}\Big(|S(K)| + K^{-1}\int_0^K |S(x)|\,\mathrm{d}x\Big)e^{-G^2K/T}.$$

Here

$$S(x) \equiv S(x, K, T) := \sum_{K \leqslant n \leqslant K+x} (-1)^n d(n) \exp(if(T, n)),$$

$$f(T, n) = 2T \operatorname{ar\,sinh}\left(\sqrt{\pi n/(2T)}\right) + \sqrt{2\pi nT + \pi^2 n^2} - \tfrac{1}{4}\pi,$$

and summation is over $K = 2^k$ such that $T^{1/3} \leqslant K \leqslant N$, where for $\delta > 0$ fixed

$$N := B^2/(T/(2\pi) - B), \quad B := T(2\pi G)^{-1} \log^{1+\delta} T.$$

6.2 Reduction to Exponential Sums

Very often problems in analytic number theory are ultimately reduced to the *estimation of exponential sums*, namely *classical van der Corput sums, sums over primes, sums over Hecke series, Kloosterman sums*, etc.

By means of Lemma 6.1 we shall prove first (see [20])

Theorem 6.2. *Let* (κ, λ) *be any exponent pair with* $\kappa > 0$ *and let* $t_1 < \cdots < t_R$ *satisfy*

$$|t_r| \leqslant T, \ |\zeta(\tfrac{1}{2} + it_r)| \geqslant V > 0 \ (r \leqslant T), \ |t_r - t_s| \geqslant 1 \ (r \neq s).$$

Then

$$R \ll TV^{-6} \log^8 T + T^{(\kappa+\lambda)/\kappa} V^{-2(1+2\kappa+2\lambda)} (\log T)^{3+6\kappa+4)/\kappa}.$$

Remark 6.1. Recall that (κ, λ) is an (one-dimensional) *exponent pair* if

$$\sum_{N < n \leqslant N' \leqslant 2N} \exp\bigl(if(n)\bigr) \ \ll \ N^{\kappa} F^{\lambda} \quad (0 < F < 1),$$

if $f(x)$ $(\in \mathbb{R})$ is sufficiently many times differentiable in $[N, 2N]$ $(N \geqslant 2)$, and essentially

$$f^{(k)}(x) \ \asymp_k \ FN^{1-k} \qquad (k = 1, 2, \ldots).$$

For a detail account on exponent pairs the reader is referred to the book of Graham and Kolesnik [10].

Thus trivially $(0, 1)$ is an exponent pair, other common ones are

$$(1/2, 1/2), \ (1/6, 4/6), \ (2/7, 4/7), \ (1/14, 11/14)$$

etc. Taking $(\kappa, \lambda) = (1/2, 1/2)$ in Theorem 6.2 we obtain

$$R \ll T^2 V^{-12} \log^{16} T,$$

and Theorem 6.1 easily follows.

Theorem 6.2 follows from

Lemma 6.2. *Let $\mathscr{A}$ be a set of numbers $\{t_r\}$ such that $T/2 \leqslant t_r \leqslant T$, $\log^2 T \leqslant G \leqslant |t_r - t_s| \leqslant J$ for $r \neq s$. With*

$$S(x) \equiv S(x, K, T) := \sum_{K \leqslant n \leqslant K+x} (-1)^n d(n) \exp(if(T, n)),$$

$$f(T, n) = 2T \operatorname{arsinh}\left(\sqrt{\pi n/(2T)}\right) + \sqrt{2\pi nT + \pi^2 n^2} - \tfrac{1}{4}\pi,$$

we have

$$\sum_{t_r \in \mathscr{A}} |S(x, K, t_r)| \ll \Big\{(K + K^{3/4} T^{1/4} G^{-1/2} \log^{1/2} T |\mathscr{A}|^{1/2} + J^{\kappa/2} T^{-\kappa/4} |\mathscr{A}| K^{(2\lambda-\kappa+2)/4}\Big\} \log^{3/2} T.$$

The result is obtained by the use of the Halász–Montgomery inequality (see, e.g., the Appendix of [20])

$$\sum_{r \leqslant R} |(\xi, \phi_r)| \leqslant ||\xi|| \left(\sum_{r,s \leqslant R} |(\phi_r, \phi_s)|\right)^{1/2}. \tag{12}$$

For Dirichlet polynomials (12) is used in the following context: If $\mathbf{a} = \{a_n\}_{n=1}^{\infty}$ and $\mathbf{b} = \{b_n\}_{n=1}^{\infty}$ are two (vector) sequences in $\mathbb{C}$, then their *standard inner product* is

$$(\mathbf{a}, \mathbf{b}) := \sum_{n=1}^{\infty} a_n \bar{b}_n. \tag{13}$$

This is used to remove the coefficients $(-1)^n d(n)$ from the sums of $S(x, K, t_r)$ and get ordinary exponential sums. Hence, with $\xi = \sum_{K < n \leqslant K+x} (-1)^n d(n)$, we have

$$||\xi||^2 = (\xi, \xi) = \sum_{K < n \leqslant K+x} d^2(n) \ll K \log^3 T$$

and by (12) and (13) the left-hand side of the sum in Lemma 6.2 is

$$\ll K^{1/2} \log^{3/2} T \left\{\sum_{t_r, t_s \in \mathscr{A}} \left|\sum_{K < n \leqslant 2K} \exp\Big(if(t_t, n) - if(t_s, n)\Big)\right|\right\}^{1/2}.$$

The terms for $r = s$ are trivially $O(K)$, and the terms $r \neq s$ are estimated by the theory of exponent pairs. Lemma 6.2 follows.

The term TV^{-6} in Theorem 6.2, which would lead to the yet unknown sixth moment

$$\int_0^T |\zeta(\tfrac{1}{2} + it)|^6 \, \mathrm{d}t \ll T \log^C T,$$

comes from the term with $|\mathscr{A}|^{1/2}$ in Lemma 6.2.

The other term is small if J is suitably chosen, namely it holds for R_0, the number of t_r's lying in an interval of length $\leqslant J$. Thus the obtained bound for R has to be multiplied by $R_0(1 + T/J)$, and this gives the other term in Theorem 6.2. Up to the choice of exponent pairs, this is the limit of the method. Note that there are new (non-classical) exponent pairs, obtained by Huxley [14–16] and Huxley and Ivić [17], for example. Huxley's work is built on the Bombieri–Iwaniec method for the estimation of exponential sums; see [3, 4]. There is a way to avoid the use of the Halász–Montgomery inequality in proving (11) and work only with the simpler Cauchy–Schwarz inequality. For this, see Chap. 8 of the author's monograph [20].

7 Large Values of $|\zeta(\frac{1}{2} + it)|$ via the Divisor Problem and $E^*(T)$

7.1 The Function $E^*(T)$

Already Atkinson's formula provides the analogy between $E(T)$ and the error term in the classical *Dirichlet divisor problem*, namely the estimation of the function

$$\Delta(x) := \sum_{n \leqslant x} d(n) - x(\log x + 2\gamma - 1). \tag{14}$$

We have

$$\begin{aligned} \Delta(x) = \frac{1}{\pi\sqrt{2}} x^{\frac{1}{4}} \sum_{n \leqslant N} d(n) n^{-\frac{3}{4}} \cos(4\pi\sqrt{nx} - \tfrac{1}{4}\pi) \\ + O_\varepsilon(x^{\frac{1}{2}+\varepsilon} N^{-\frac{1}{2}}) \quad (1 \ll N \ll x), \end{aligned} \tag{15}$$

which is the classical truncated *Voronoï formula* for $\Delta(x)$ of 1904 (see Voronoï [48] and Chap. 3 of [20]).

A better analogy is to consider, instead of $\Delta(x)$ (see Jutila [34, 35]), the function

$$\begin{aligned} \Delta^*(x) :&= -\Delta(x) + 2\Delta(2x) - \frac{1}{2}\Delta(4x) \\ &= \frac{1}{2} \sum_{n \leqslant 4x} (-1)^n d(n) - x(\log x + 2\gamma - 1), \end{aligned}$$

because Atkinson's formula has a factor $(-1)^n$ in it, and the formula for $\Delta(x)$ does not. Then it makes sense to investigate the function

$$E^*(t) \ := \ E(t) - 2\pi\Delta^*\bigl(\frac{t}{2\pi}\bigr).$$

Namely, we have the explicit formula

$$\begin{aligned}\Delta^*(x) = \frac{1}{\pi\sqrt{2}}x^{\frac{1}{4}}\sum_{n\leqslant N}(-1)^n d(n)n^{-\frac{3}{4}}\cos(4\pi\sqrt{nx} - \tfrac{1}{4}\pi)\\ + O_\varepsilon(x^{\frac{1}{2}+\varepsilon}N^{-\frac{1}{2}}) \quad (1 \ll N \ll x),\end{aligned} \tag{16}$$

and without $(-1)^n$ the formula (16) is exactly Voronoï truncated formula (14) for $\Delta(x)$.

If $N = T^{1/3-\varepsilon}$, then the first N terms in the truncated Voronoï formula for $\Delta^*(x)$ and $\sum_1(T)$ in Atkinson's formula are asymptotic to one another. This is the analogy between $\Delta^*(x)$ and $E(T)$.

7.2 Results on $E^*(T)$

Jutila [35] proved:

$$\int_T^{T+H}(E^*(t))^2\,\mathrm{d}t \ \ll_\varepsilon \ HT^{1/3}\log^3 T + T^{1+\varepsilon} \tag{17}$$

for $1 \ll H = H(T) \leqslant T$.The significance of (17) is that, in view of

$$\int_0^T(\Delta^*(t))^2\,\mathrm{d}t \ \sim \ AT^{3/2}, \quad \int_0^T E^2(t)\,\mathrm{d}t \ \sim \ BT^{3/2},$$

for $A, B > 0, T \to \infty$, the function $E^*(t)$ is in the *mean square sense* of a lower order of magnitude than either $\Delta^*(t)$ or $E(t)$. The author [26] in 2012 complemented (17) with the lower bound

$$\int_T^{T+H}(E^*(t))^2\,\mathrm{d}t \gg HT^{1/3}\log^3 T \tag{18}$$

for $T^{2/3+\varepsilon} \leqslant H = H(T) \leqslant T$.

Remark 7.2. The bounds (17) and (18) show the true order of the integral. Problem: What is the lower bound for $H = H(T)$ for which (18) holds? The author [23] proved in 2007 the asymptotic formula

$$\int_0^T (E^*(t))^2\,\mathrm{d}t \;=\; T^{4/3}P_3(\log T) + O_\varepsilon(T^{7/6+\varepsilon}),$$

where $P_3(x)$ is a cubic polynomial in x with positive leading coefficient, and all the coefficients may be written down explicitly. It is reasonable to conjecture that the above error term is $O_\varepsilon(T^{1+\varepsilon})$ and at the same time $\Omega(T)$, but this is certainly a difficult problem. However the above formula does imply neither the upper nor the lower bound for $(E^*(t))^2$ in short intervals. The author also proved [21, 22]

$$\int_0^T \bigl|E^*(t)\bigr|^3\,\mathrm{d}t \;\ll_\varepsilon\; T^{3/2+\varepsilon} \tag{19}$$

and that

$$\int_0^T \bigl|E^*(t)\bigr|^5\,\mathrm{d}t \;\ll_\varepsilon\; T^{2+\varepsilon}. \tag{20}$$

Hence by the Cauchy–Schwarz inequality for integrals it follows that

$$\int_0^T \bigl|E^*(t)\bigr|^4\,\mathrm{d}t \;\ll_\varepsilon\; T^{7/4+\varepsilon}. \tag{21}$$

An open problem is to investigate higher moments of $E^*(t)$. In particular, it would be interesting to show that the integral of $(E^*(t))^3$ is of a lower order of magnitude than that of $|E^*(t)|^3$.

Theorem 7.3. *Suppose that the bound*

$$\int_0^T (E^*(t))^3\,\mathrm{d}t \;\ll_\varepsilon\; T^{3/2-\eta+\varepsilon}$$

holds for some constant η such that $0 < \eta < 1/6$. Then for any $\varepsilon > 0$ there exist constants $T_0(\varepsilon), A > 0$ such that, for $T \geqslant T_0(\varepsilon)$, every interval $[T,\, T + T^{1-\eta+\varepsilon}]$ contains points T_1, T_2 for which

$$E^*(T_1) > AT_1^{1/6}\log^{3/2} T_1, \qquad E^*(T_2) < -AT_2^{1/6}\log^{3/2} T_2.$$

Remark 7.3. This is a recent result of the author [25]. Note that, for $A \geqslant 2$, one has

$$\int_0^T |E^*(t)|^A\,\mathrm{d}t \;\gg_A\; T^{1+A/6}(\log T)^{3A/2}.$$

Remark 7.4. We have

$$E(T) = 2\pi\Delta^*\Bigl(\frac{T}{2\pi}\Bigr) + \Omega(T^{1/6}\log^{3/2} T). \tag{22}$$

A plausible conjecture is that

$$E(T) = 2\pi\Delta^*\Big(\frac{T}{2\pi}\Big) + O_\varepsilon(T^{1/4+\varepsilon}). \tag{23}$$

This is compatible with the classical conjectures (implied neither by the Lindelöf Hypothesis (that $\zeta(\frac{1}{2} + it) \ll_\varepsilon |t|^\varepsilon$) or Riemann Hypothesis!)

$$E(T) \ll_\varepsilon T^{1/4+\varepsilon}, \qquad \Delta^*(x) \ll_\varepsilon x^{1/4+\varepsilon}, \tag{24}$$

but perhaps, in view of (22), the true exponent in (23) is even $1/6 + \varepsilon$.

Remark 7.5. Lau and Tsang [38] showed that $\Delta(x) \ll_\varepsilon x^{a+\varepsilon}$ implies $\Delta^*(x) \ll_\varepsilon x^{a+\varepsilon}$ and conversely, $\Delta^*(x) \ll_\varepsilon x^{b+\varepsilon}$ gives $\Delta(x) \ll_\varepsilon x^{b+\varepsilon}$. In the other direction, it is known that $a < 1/4$ and $b < 1/4$ cannot hold.

8 Connections of $E(T)$ and $E^*(T)$ with the Moments of $|\zeta(\frac{1}{2} + it)|$

8.1 Large Values Estimates

There are several ways to connect power moments of $|\zeta(\frac{1}{2} + it)|$ with the moments of $E(T)$ and/or $E^*(T)$. The author [25] proved

Theorem 8.3. *Let $t_1, \dots, t_R$ be points in $[T,\ 2T]$ which satisfy*

$$|\zeta(\tfrac{1}{2} + it_r)| \geqslant V \geqslant T^\varepsilon$$

and $|t_r - t_s| \geqslant 1$ for $r, s \leqslant R$ and $r \neq s$. Then we have, for $L = \log T$, $G = A(V/L)^2$ with a suitable constant $A > 0$, and $k \in \mathbb{N}$ fixed,

$$R \ll V^{-2-2k}L^{2+2k}\int\limits_{T/3}^{3T}\Big\{|E(t+2G) - E(t-2G)|^k$$

$$+ |E(t+\tfrac{1}{2}G) - E(t-\tfrac{1}{2}G)|^k\Big\}\,\mathrm{d}t.$$

Remark 8.6. The presence of the factors

$$|E(t+2G) - E(t-2G)|^k,\ |E(t+\tfrac{1}{2}G) - E(t-\tfrac{1}{2}G)|^k$$

makes one feel that it is possible to exploit the "closeness" of the arguments. Indeed, Theorem 8.3 is valid with E^* in place of E.

Corollary 8.1. *Suppose that the integral on the right-hand side in Theorem 8.3 is bounded by* $O_\varepsilon(T^{1+\alpha+\varepsilon}G^\beta)$ *for some real constants* $\alpha = \alpha(k)$ (> 0) *and* $\beta = \beta(k) \leqslant k-1$, *and* $T^\varepsilon \leqslant G = G(T) \ll T^{1/3}$. *Then we have*

$$\int_0^T |\zeta(\tfrac{1}{2}+it)|^{2+2k-2\beta}\,\mathrm{d}t \ \ll_\varepsilon \ T^{1+\alpha+\varepsilon}.$$

We also have the following result (see [24]):

Theorem 8.4. *For* $T^{3/8} \ll G = G(T) \ll T^{1/2}$ *we have*

$$\int_T^{2T} \Big(E(t+G) - E(t-G)\Big)^4\,\mathrm{d}t \ \ll_\varepsilon \ T^{1+\varepsilon}G^2,$$

and the same result holds with Δ *instead of* E.

Corollary 8.2. *As shown in [24], if the above bound holds for* $T^\varepsilon \leqslant G \leqslant T^{1/2}$, *then the weak 8th moment*

$$\int_0^T |\zeta(\tfrac{1}{2}+it)|^8\,\mathrm{d}t \ \ll_\varepsilon \ T^{1+\varepsilon}$$

holds.

Remark 8.7. The above bound holds for $E^*(T)$ in place of $E(T)$ in view of the fourth moment bound for $E^*(T)$.

A recent result of the author [28] is the following:

Theorem 8.5. *Let* $t_1, \dots, t_R$ *be points in* $[T,\ 2T]$ *which satisfy*

$$|\zeta(\tfrac{1}{2}+it_r)| \geqslant V \geqslant T^\varepsilon$$

and $|t_r - t_s| \geqslant 1$ *for* $r, s \leqslant R$ *and* $r \neq s$. *Then we have*

$$R \ll V^{-2-2k}\log T \int_{T/3}^{3T} |E^*(t)|^k\,\mathrm{d}t.$$

From the results on moments of $E^*(T)$ [see (17)–(19)] we obtain, with $k = 3, 4$,

Corollary 8.3.

$$\int_0^T |\zeta(\tfrac{1}{2}+it)|^8\,\mathrm{d}t \ll_\varepsilon T^{3/2+\varepsilon}, \qquad \int_0^T |\zeta(\tfrac{1}{2}+it)|^{10}\,\mathrm{d}t \ll_\varepsilon T^{7/4+\varepsilon}.$$

We also have, with $k = 5$,

Corollary 8.4.

$$\int_0^T |\zeta(\tfrac{1}{2}+it)|^{12}\,\mathrm{d}t \ll_\varepsilon T^{2+\varepsilon}.$$

The bounds in Corollaries 8.3 and 8.4 are, up to "ε", the sharpest ones known. D.R. Heath-Brown [see (11)], obtained $T^2\log^{17}T$ as the bound for the last integral. Corollary 8.4 thus provides an alternate proof of Heath-Brown's result.

Theorem 8.3 on large values follows by using ($C>0, L=\log T, t\asymp T$)

$$\int_{t-G}^{t+G} |\zeta(\tfrac{1}{2}+iu)|^2\,\mathrm{d}u \leqslant \frac{C}{G^2}\int_{-GL}^{GL} xE^*(t+x)e^{-(x/G)^2}\,\mathrm{d}x + CGL^2 \tag{25}$$

and the fact that

$$\Delta^*(x+H)-\Delta^*(x) \ll_\varepsilon Hx^\varepsilon \quad (1\ll H\leqslant x),$$

which is a consequence of the elementary bound

$$d(n) \ll_\varepsilon n^\varepsilon$$

and the defining relation

$$\Delta^*(x) \;=\; \tfrac{1}{2}\sum_{n\leqslant 4x}(-1)^n d(n) - x(\log x + 2\gamma - 1).$$

Conjecture 8.1 (M. Jutila). We have

$$\Delta^*(x+H)-\Delta^*(x) \ll_\varepsilon \sqrt{H}x^\varepsilon \quad (x^\varepsilon \ll H\leqslant x).$$

This conjecture is motivated by mean value results for $\Delta(x+H)-\Delta(x)$ and $\Delta^*(x+H)-\Delta^*(x)$ (see [25]). Recently Wenguang Zhai and the author [32] proposed a similar

Conjecture 8.2. Suppose that

$$\log T \leqslant U \leqslant T^{1/2}/10,\ T^{1/2}\ll H\ll T,\ HU\gg T^{1+\varepsilon}.$$

Then the estimate

$$\int_T^{T+H} \max_{0\leqslant u\leqslant U}\bigl|\Delta(x+u)-\Delta(x)\bigr|^2\,\mathrm{d}x \ll HU\log^c T$$

holds for some absolute constant $c\geqslant 0$.

This is a very strong conjecture, since it implies

Corollary 8.5. *If Conjecture 8.2 holds, then* $\Delta(x) \ll_\varepsilon x^{1/4+\varepsilon}$, *and (by Remark 7.5) also* $\Delta^*(x) \ll_\varepsilon x^{1/4+\varepsilon}$.

From (25) we have the following

Corollary 8.6. *If* $E^*(T) \ll_\varepsilon T^{\varrho+\varepsilon}$ $(0 < \varrho < 1/3)$, *then*

$$\zeta(\tfrac{1}{2} + it) \ \ll_\varepsilon \ |t|^{\varrho/2+\varepsilon}.$$

Thus if the conjectural $E^*(T) \ll_\varepsilon T^{1/6+\varepsilon}$ holds, then we have (the very strong bound)

$$\zeta(\tfrac{1}{2} + it) \ \ll_\varepsilon \ |t|^{1/12+\varepsilon}.$$

Let us mention that the best unconditional exponent for $|\zeta(\frac{1}{2} + it)|$ is $32/205 = 0.15609\ldots$, which is a result of Huxley [16] (very recently Bourgain [5] in 2014 improved Huxley's exponent to $53/342 = 0.15497\ldots$). The yet unproved Lindelöf hypothesis states that $\zeta(\frac{1}{2} + it) \ \ll_\varepsilon \ |t|^\varepsilon$. It follows from the RH (see, e.g., Chap. 1 of [20]), but the converse is unknown (and not very likely!).

9 A New Large Values Result Involving the Function $E_k(T)$

Let $E_k(T)$ be the error-term function defined by (9). We present a new result which generalizes Theorem 8.5. This is

Theorem 9.6. *Let* $E_k(T)$ *be is as in (9). Then, for a fixed* m $(\geqslant 1)$ *and* $L = \log T$, *we have*

$$\int_T^{2T} |\zeta(\tfrac{1}{2} + it)|^{2k(m+1)}\,\mathrm{d}t \ll_\varepsilon T^{1+\varepsilon} + L^{4m+k^2+1}\int_{T/2}^{2T} |E_k(t)|^m\,\mathrm{d}t. \tag{26}$$

Moreover, when $k = 1$, *(26) remains valid with* $E_1(t) = E(T)$ *replaced by* $E^*(t)$.

Theorem 9.6 is a consequence of the following large values estimate. Let $t_1, \ldots, t_R$ be points in $[T,\ 2T]$ which satisfy

$$|\zeta(\tfrac{1}{2} + it_r)| \geqslant V \geqslant T^\varepsilon \tag{27}$$

and $|t_r - t_s| \geqslant 1$ for $r, s \leqslant R$ and $r \neq s$. Then we have

$$R \ \ll \ V^{-2k(m+1)}L^{4m+k^2}\int_{T/3}^{3T} |E_k(t)|^m\,\mathrm{d}t. \tag{28}$$

Namely we divide the integral on the left-hand side of (26) into $O(L)$ integrals of the form

$$\int_{T,V\leqslant|\zeta(\frac12+it)|\leqslant 2V}^{2T} |\zeta(\tfrac12+it)|^{2k(m+1)}\,\mathrm{d}t \ \ll\ R_V V^{2k(m+1)},$$

where R_V is the number of points $t_{r,V}$ each of which satisfies (27) (with $t_r \equiv t_{r,V}$) and $|t_{r,V} - t_{s,V}| \geqslant 1$ for $r_V, s_V \leqslant R_V$. The points $t_{r,V}$ are chosen in such a way that $|\zeta(\frac12 + it)|$ attains its maximum value in suitable subintervals of length not exceeding unity, lying in $[T,\ 2T]$. Applying (28) to bound R_V one obtains (26), since the contribution of V which are $\leqslant T^{\varepsilon}$ is trivially $O(T^{1+\varepsilon})$.

To prove (26) we start from (5) to obtain

$$\sum_{r\leqslant R} |\zeta(\tfrac12 + it_r)|^{2k} \ \ll\ L\left(\sum_{r\leqslant R}\int_{t_r-\frac12}^{t_r+\frac12} |\zeta(\tfrac12+it)|^{2k}\,\mathrm{d}t + 1\right). \tag{29}$$

The interval $[T,\ 2T]$ containing the points $\{t_r\}$ is covered with those intervals of length $4GL$ ($T^{\varepsilon} \leqslant G \leqslant T^{1/3}$) (except the last which may be shorter) which have a non-empty intersection with at least one interval of the form $[t_r - \frac12, t_r + \frac12]$. The centers of such intervals are denoted by $t'_{r'}$, and we suppose that there are R' ($\leqslant R$) such intervals. Moreover, by considering separately the points $\{t'_{r'}\}$ with even and odd indices r', we are led to the estimation of the sum

$$\sum_{r'\leqslant R'}\int_{t'_{r'}-2GL}^{t'_{r'}+2GL} |\zeta(\tfrac12+it)|^{2k}\,\mathrm{d}t, \tag{30}$$

where $|t'_{r'} - t'_{s'}| \geqslant 4GL$ if $r' \neq s'$. Now we recall (9) and set $Q_{k^2} := p_{k^2} + p'_{k^2}$. Since ($H = 2GL$)

$$\begin{aligned}\int_{t_r-H}^{t_r+H} |\zeta(\tfrac12+it)|^{2k}\,\mathrm{d}t &\leqslant e\int_{-\infty}^{\infty} |\zeta(\tfrac12+it_r+ix)|^{2k} e^{-(x/H)^2}\,\mathrm{d}x\\ &= e\int_{-\infty}^{\infty}\left[(t_r+x)Q_{k^2}(t_r+x) + E'_k(t_r+x)\right]e^{-(x/H)^2}\,\mathrm{d}x,\end{aligned}$$

it follows that the expression in (30) is

$$\begin{aligned}&\ll R'G\log^{k^2} T + \sum_{r'\leqslant R'} G^{-2}\int_{-\infty}^{\infty} xE_k(t'_r+x)e^{-(x/H)^2}\,\mathrm{d}x\\ &\ll R'G\log^{k^2} T + \frac{L^2}{G}\sum_{r'\leqslant R'}\int_{-2GL}^{2GL} |E_k(t'_r+x)|\,\mathrm{d}x\\ &= R'G\log^{k^2} T + \frac{L^2}{G}\sum_{r'\leqslant R'}\int_{t'_{r'}-2GL}^{t'_{r'}+2GL} |E_k(t)|\,\mathrm{d}t.\end{aligned}$$

Since the intervals $(t'_{r'}-2GL, t'_{r'}+2GL)$ are disjoint we obtain, on applying Hölder's inequality ($m \geqslant 1$ is fixed), that the last sum above does not exceed

$$\left(\sum_{r'\leqslant R'} 1\right)^{1-1/m}\left\{\sum_{r'\leqslant R'}\left(\int_{t'_{r'}-2GL}^{t'_{r'}+2GL}|E_k(t)|\,\mathrm{d}t\right)^m\right\}^{1/m}$$

$$\ll (R')^{1-1/m}\left\{\sum_{r'\leqslant R'}\int_{t'_{r'}-2GL}^{t'_{r'}+2GL}|E_k(t)|^m\,\mathrm{d}t\left(\int_{t'_{r'}-2GL}^{t'_{r'}+2GL}1\,\mathrm{d}t\right)^{m-1}\right\}^{1/m}$$

$$\ll (R')^{1-1/m}\left(\int_{T/2}^{5T/2}|E_k(t)|^m\,\mathrm{d}t\right)^{1/m}(GL)^{1-1/m}.$$

If (27) holds, then the sum in (29) is $\gg RV^{2k}$. Therefore we obtain

$$RV^{2k} \ll R'GL^{k^2+1} + \frac{L^3}{G}(R')^{1-1/m}\left(\int_{T/2}^{5T/2}|E_k(t)|^m\,\mathrm{d}t\right)^{1/m}(GL)^{1-1/m}.$$

Now choose

$$G = cV^{2k}L^{-k^2-1} \tag{31}$$

with a suitable $c > 0$, so that $R'GL^{k^2+1}$ is dominated by RV^{2k}. Simplifying the above bound and taking into account that $R' \leqslant R$, we have

$$R \ll L^{3m}G^{-m}V^{-2km}(GL)^{m-1}\int_{T/2}^{5T/2}|E_k(t)|^m\,\mathrm{d}t. \tag{32}$$

With G given by (31), it is seen that (32) simplifies to

$$R \ll L^{4m+k^2}V^{-2k(m+1)}\int_{T/2}^{5T/2}|E_k(t)|^m\,\mathrm{d}t,$$

and this is (28). The proof in the general case is complete. The case $k = 1$ follows on using (25) and the above arguments.

Besides results obtained from the case $k = 1$ (Corollaries 8.3 and 8.4), we may consider the case $k = 2$. It was proved by Motohashi and the author [30, 31] and Motohashi [40] that

$$\int_0^T|E_2(t)|\,\mathrm{d}t \ll T^{3/2}, \quad \int_0^T|E_2(t)|^2\,\mathrm{d}t \ll T^2\log^{22}T.$$

Thus with $k = 2$ and $m = 1$, $m = 2$, we obtain

$$\int_0^T |\zeta(\tfrac{1}{2} + it)|^8 \,\mathrm{d}t \ll T^{3/2}\log^9 T, \quad \int_0^T |\zeta(\tfrac{1}{2} + it)|^{12} \,\mathrm{d}t \ll T^2\log^{35} T,$$

which are (up to log-factors) the sharpest known results.

The present knowledge about the function $E_k(T)$ does not make it possible to exploit the moments of $E_k(T)$ when $k \geqslant 3$. However, there are some heuristic reasons to believe (see, e.g., [29]) that $E_3(T) \ll_\varepsilon T^{3/4+\varepsilon}$ holds. Already this pointwise bound would lead to

$$\int_0^T |\zeta(\tfrac{1}{2} + it)|^{12} \,\mathrm{d}t \ll_\varepsilon T^{7/4+\varepsilon},$$

hence to a substantial improvement over Heath-Brown's classical bound (11). In fact, any bound of the form $E_3(T) \ll_\varepsilon T^{\alpha+\varepsilon}$ with $\alpha < 1$ would improve (11).

References

1. F.V. Atkinson, The mean value of the Riemann zeta-function. Acta Math. **81**, 353–376 (1949)
2. E. Bombieri, The classical theory of zeta and *L*-functions. Milan J. Math. **78**, 11–59 (2010)
3. E. Bombieri, H. Iwaniec, On the order of $\zeta(\frac{1}{2} + it)$. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **13**, 449–472 (1986)
4. E. Bombieri, H. Iwaniec, Some mean value theorems for exponential sums. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **13**, 473–486 (1986)
5. J. Bourgain, Decoupling, exponential sums and the Riemann zeta-function. Preprint available at `arXiv:1408.5794` (to appear)
6. J.B. Conrey, S.M. Gonek, High moments of the Riemann zeta-function. Duke Math. J. **107**, 577–604 (2001)
7. J.B. Conrey, D.W. Farmer, J.P. Keating, M.O. Rubinstein, N.C. Snaith, Integral moments of *L*-functions. Proc. Lond. Math. Soc. (3) **91**, 33–104 (2005)
8. J.B. Conrey, D.W. Farmer, J.P. Keating, M.O. Rubinstein, N.C. Snaith, Lower order terms in the full moment conjecture for the Riemann zeta function. J. Number Theory **128**, 1516–1554 (2008)
9. H.M. Edwards, *Riemann's Zeta-Function* (Academic, New York, London, 1974)
10. S.W. Graham, G. Kolesnik, *Van der Corput's Method for Exponential Sums*. London Mathematical Society Lecture Note Series, vol.126 (Cambridge University Press, Cambridge, 1991)
11. A.J. Harper, Sharp conditional bounds for moments of the Riemann zeta-function. Preprint available at `arXiv:1305.4618`
12. D.R. Heath-Brown, The twelfth power moment of the Riemann zeta-function. Q. J. Math. (Oxford) **29**, 443–462 (1978)
13. G.A. Hiary, A.M. Odlyzko, The zeta function on the critical line: numerical evidence for moments and random matrix theory models. Math. Comput. **81**, 1723–1752 (2012)
14. M.N. Huxley, Integer points, exponential sums and the Riemann zeta function, in *Number Theory for the Millennium II (Urbana, IL, 2000)*, ed. by M.A. Bennett (A. K. Peters, Natick, MA, 2002), pp. 275–290
15. M.N. Huxley, Exponential sums and lattice points III. Proc. Lond. Math. Soc. (3) **87**(3), 591–609 (2003)

16. M.N. Huxley, Exponential sums and the Riemann zeta function V. Proc. Lond. Math. Soc. (3) **90**, 1–41 (2005)
17. M.N. Huxley, A. Ivić, Subconvexity for the Riemann zeta-function and the divisor problem. Bulletin CXXXIV de l'Académie Serbe des Sciences et des Arts - 2007, Classe des Sciences mathématiques et naturelles, Sciences mathématiques No. 32, pp. 13–32
18. A. Ivić, *Mean Values of the Riemann Zeta-Function, LN's*, vol. 82 (Tata Institute of Fundamental Research, Bombay, 1991) (distr. by Springer, Berlin etc.)
19. A. Ivić, On the ternary additive divisor problem and the sixth moment of the zeta-function, in *Sieve Methods, Exponential Sums, and Their Applications in Number Theory*, ed. by G.R.H. Greaves, G. Harman, M.N. Huxley (Cambridge University Press, Cambridge, 1996), pp. 205–243
20. A. Ivić, *The Riemann Zeta-Function* (Wiley, New York, 1985); (Reissue: Dover, Mineola, New York, 2003)
21. A. Ivić, On the Riemann zeta function and the divisor problem. Cen. Eur. J. Math. **2**(4) , 1–15 (2004); and II. Cen. Eur. J. Math. **3**(2), 203–214 (2005)
22. A. Ivić, On the Riemann zeta-function and the divisor problem IV. Unif. Distrib. Theory **1**, 125–135 (2006)
23. A. Ivić, On the mean square of the zeta-function and the divisor problem. Ann. Acad. Sci. Fenn. Math. **23**, 1–9 (2007)
24. A. Ivić, On the divisor function and the Riemann zeta-function in short intervals. Ramanujan J. **19**(2), 207–224 (2009)
25. A. Ivić, On the moments of the Riemann zeta-function in short intervals. Hardy-Ramanujan J. **32**, 4–23 (2009)
26. A. Ivić, On some mean square estimates for the zeta-function in short intervals. Ann. Univ. Sci. Bp. Sect. Comput. **40**, 321–335 (2013)
27. A. Ivić, On the moments of the function $E^*(t)$. Turk. J. Anal. Number Theory **2**(3), 102–109 (2014). doi:10.12691/tjant-2-3-9. http://www.pubs.scieoub.com/tjant/2/3/9
28. A. Ivić, On some mean value results for the zeta-function in short intervals. Acta Arith. **162**(2), 141–158 (2014)
29. A. Ivić, The mean value of the Riemann zeta-function on the critical line, in *Analytic Number Theory, Approximation Theory, and Special Functions*, ed. by G.V. Milovanović, M.T. Rassias (Springer, New York, 2014), pp. 3–68
30. A. Ivić, Y. Motohashi, The mean square of the error term for the fourth moment of the zeta-function. Proc. Lond. Math. Soc. (3) **69**, 309–329 (1994)
31. A. Ivić, Y. Motohashi, On the fourth power moment of the Riemann zeta-function. J. Number Theory **51**, 16–45 (1995)
32. A. Ivić, W. Zhai, On a hybrid fourth moment involving the Riemann zeta-function, in *Topics in Mathematical Analysis and Applications*, ed. by T. Rassias, L. Tóth. Springer Optimization and Its Applications, vol. 94 (Springer, Berlin, 2014)
33. D. Joyner, *Distribution Theorems of L-Functions*. Pitman Research Notes in Mathematics Series, vol. 142 (Longman Scientific & Technical, Harlow, Essex, 1986), 247 pp.
34. M. Jutila, On a formula of Atkinson. Topics in classical number theory, Colloq. Budapest 1981, Vol. I. Colloq. Math. Soc. János Bolyai **34**, 807–823 (1984)
35. M. Jutila, Riemann's zeta-function and the divisor problem. Arkiv Mat. **21**, 75–96 (1983); and II. Arkiv Mat. **31**, 61–70 (1993)
36. J. Kaczorowski, Axiomatic theory of *L*-functions: the Selberg class, in *Analytic Number Theory*, ed. by A. Perelli, C. Viola (Springer, Berlin/Heidelberg, 2006), pp. 133–209
37. J.P. Keating, N.C. Snaith, Random matrix theory and $\zeta(\frac{1}{2} + it)$. Commun. Math. Phys. **214**, 57–89 (2000)
38. Y.-K. Lau, K.-M. Tsang, Omega result for the mean square of the Riemann zeta function. Manuscr. Math. **117**, 373–381 (2005)
39. M.L. Mehta, *Random Matrices*, 3rd edn. Pure and Applied Mathematics, vol. 142 (Elsevier/Academic, Amsterdam, 2004)

40. Y. Motohashi, *Spectral Theory of the Riemann Zeta-Function* (Cambridge University Press, Cambridge, 1997)
41. M. Radziwill, K. Soundararajan, Continuous lower bounds for moments of zeta and L-functions. Mathematika **59**, 119–128 (2013)
42. K. Ramachandra, *On the Mean-Value and Omega-Theorems for the Riemann Zeta-Function, LN's*, vol. 85 (Tata Institute of Fundamental Research, Bombay, 1995) (distr. by Springer, Berlin etc.)
43. B. Riemann, Über die Anzahl der Primzahlen unter einer gegebener Grösse. Monatsh. Preuss. Akad. Wiss., 671–680 (1859)
44. M.O. Rubinstein, S. Yamagishi, Computing the moment polynomials of the zeta function. Math. Comput. **84**, 425–454 (2015)
45. A. Selberg, *Selected Papers*, vol. I (Springer, Berlin, 1989); and vol. II (Springer, Berlin, 1991)
46. K. Soundarajan, Moments of the Riemann zeta function. Ann. Math. **170**, 981–993 (2010)
47. E.C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd edn. (Oxford University Press, Oxford, 1986)
48. G.F. Voronoï, Sur une fonction transcendante et ses applications à la sommation de quelques séries. Ann. École Normale **21**(3), 207–268 (1904); and Ann. École Normale **21**(3), 459–534 (1904)

A Note on Bessel Twists of L-Functions

J. Kaczorowski and A. Perelli

Dedicated to Professor Helmut Maier for his 60th birthday

Abstract We show that certain twists of Bessel type of a given L-function have meromorphic continuation over $\mathbb{C}$.

Keywords L-functions • Selberg class • Twists • Bessel functions

2010 Mathematics Subject Classification: 11 M 41

1 Definitions and Statements

The purpose of this note is to obtain the meromorphic continuation of certain twists of Bessel type of a given L-function.

By an L-function we mean a member of the extended Selberg class $\mathcal{S}^\sharp$, namely the class of Dirichlet series with continuation over $\mathbb{C}$ with at most a pole at $s = 1$, and satisfying a general functional equation of Riemann type. Almost every L-function from number theory and automorphic representation theory belongs, at least conjecturally, to $\mathcal{S}^\sharp$. See Selberg [17], Conrey-Ghosh [2], Ram Murty [11], and our survey papers [4, 6, 14, 15] and [16] for definitions and basic results on the Selberg class theory. In particular, we assume that the reader is familiar with the notions of degree and conductor of $F \in \mathcal{S}^\sharp$, respectively denoted by d and q.

J. Kaczorowski
Faculty of Mathematics and Computer Science, A.Mickiewicz University, 61-614 Poznań, Poland

Institute of Mathematics of the Polish Academy of Sciences, 00-956 Warsaw, Poland
e-mail: kjerzy@amu.edu.pl

A. Perelli (✉)
Dipartimento di Matematica, Università di Genova, via Dodecaneso 35, 16146 Genova, Italy
e-mail: perelli@dima.unige.it

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_13

By a twist of Bessel type, *Bessel twist* for short, we mean a twist of the following form. Let $B(\xi)$ be a complex valued function of the real variable $\xi \geq 1$ with the property that as $\xi \to +\infty$

$$B(\xi) \sim \sum_{k\in\mathbb{Z}} c_k \frac{e(-f_k(\xi,\boldsymbol{\alpha}))}{\xi^{\nu_k}}, \qquad 0 \leq \nu_0 < \nu_{\pm 1} < \nu_{\pm 2} < \cdots \to \infty, \tag{1}$$

where $\sim$ means that for any positive integer K, $B(\xi)$ equals the sum of the terms with $|k| \leq K$ plus $O_K(\xi^{-\min(\nu_{K+1}, \nu_{-K-1})})$. Moreover, $e(x) = e^{2\pi i x}$, $c_k \in \mathbb{C}$ and the functions $f_k(\xi,\boldsymbol{\alpha})$ are of type

$$f_k(\xi,\boldsymbol{\alpha}) = \sum_{j=0}^{N_k} \alpha_j \xi^{\kappa_j} \tag{2}$$

with $\alpha_j = \alpha_j(k) \neq 0$, $\kappa_j = \kappa_j(k)$ and $0 < \kappa_{N_k} < \kappa_{N_k-1} < \cdots < \kappa_0$. Clearly, $B(\xi)$ is bounded for $\xi \geq 1$. The twist of $F \in \mathcal{S}^\sharp$ by $B(\xi)$ is defined for $\sigma > 1$ by the absolutely convergent Dirichlet series

$$F(s;B) = \sum_{n=1}^{\infty} \frac{a(n)B(n)}{n^s},$$

where $a(n)$ are the coefficients of $F(s)$. The function $F(s;B)$ is called a Bessel twist since the Bessel functions are typical examples of functions $B(\xi)$; see below. In particular, as a very special case of the Theorem below we obtain that the twist

$$\sum_{n=1}^{\infty} \frac{a(n)J_\nu(2\pi\alpha n^\kappa)}{n^s}, \tag{3}$$

where $J_\nu(z)$ denotes the familiar Bessel J-function and $0 \neq \alpha \in \mathbb{R}$, has meromorphic continuation over $\mathbb{C}$ if $\kappa \leq 1/d$, d being the degree of $F(s)$. Moreover, it is entire whenever $\kappa < 1/d$.

Theorem. *Let $F \in \mathcal{S}^\sharp$ with $d > 0$ and $B(\xi)$ be as above with all $\kappa_0 \leq 1/d$. Then $F(s;B)$ is meromorphic over $\mathbb{C}$. If all $f_k(\xi,\boldsymbol{\alpha})$ have a term with exponent $< 1/d$, then $F(s;B)$ is entire.*

It is clear from the proof that we have some control on the poles of $F(s;B)$, as well as more precise information about its analytic character in certain cases. Moreover, we can also get meromorphic (or entire) continuation of $F(s;B)$ in certain cases with $\kappa_0 > 1/d$.

We finally remark that Noda [12, 13] obtained, by a different method, the meromorphic continuation of the twists of the Riemann zeta function by certain Bessel and hypergeometric functions, respectively. These results are now special cases of our Theorem, see the examples below; however, Noda obtained also several representations for such twists.

2 Proof

The Theorem is a consequence of the properties of the nonlinear twists obtained in our papers [5, 7] and [9]; we first summarize the main features of such results. Given $F \in \mathcal{S}^\sharp$ with $d > 0$ and $f(\xi, \boldsymbol{\alpha})$ as in (2) with $\kappa_0 \leq 1/d$, the nonlinear twist

$$F(s;f) = \sum_{n=1}^{\infty} \frac{a(n)}{n^s} e(-f(n, \boldsymbol{\alpha}))$$

is entire if $f(\xi, \boldsymbol{\alpha})$ has at least one exponent $< 1/d$. If instead $f(\xi, \boldsymbol{\alpha})$ is the standard twist $\alpha\xi^{1/d}$, $0 \neq \alpha \in \mathbb{R}$, then $F(s;f)$ is still entire if $|\alpha| \notin \mathrm{Spec}(F)$, while it has meromorphic continuation to $\mathbb{C}$ with at most simple poles at the points

$$s_k = \frac{d+1}{2d} - \frac{k}{d} - i\theta_F \qquad\qquad k = 0, 1, \ldots$$

and nonvanishing residue at s_0 if $|\alpha| \in \mathrm{Spec}(F)$. The spectrum $\mathrm{Spec}(F)$ is defined as

$$\mathrm{Spec}(F) = \{\alpha > 0 : a(n_\alpha) \neq 0\} \text{ where } n_\alpha = qd^{-d}\alpha^d \text{ and } a(n_\alpha) = 0 \text{ if } n_\alpha \notin \mathbb{N},$$

while the real number θ_F is the internal shift of $F(s)$ (see, e.g., [9]). Moreover, when $\kappa_0 \leq 1/d$ the function $F(s;f)$ is of order ≤ 1 whenever entire. Further, $F(s;f)$ has polynomial growth on vertical strips when $f(\xi, \boldsymbol{\alpha})$ coincides with the standard twist $\alpha\xi^{1/d}$, and satisfies the weaker bound $F(s;f) \ll \exp(|t|^\delta)$ with some $\delta < 1$ in the other cases.

The Theorem follows from the above results. Indeed, thanks to (1), for any given positive integer K we may write

$$F(s;B) = \sum_{|k| \leq K} c_k F(s;f_k) + R_K(s) = M_K(s) + R_K(s),$$

say. But $M_K(s)$ is meromorphic over $\mathbb{C}$ (or entire) by the above results, and $R_K(s)$ is absolutely convergent for $\sigma > -\min(\nu_{K+1}, \nu_{-K-1}) + 1$; the result follows since such a minimum becomes arbitrarily large as K increases.

The above results about $F(s;f)$ show that in principle we may gain some control on the poles of $F(s;B)$. In addition, we have some control on the growth of $F(s;B)$ on vertical strips. Moreover, under suitable information about the order of growth of the constants involved in (1), we can also deduce that $F(s;B)$ is of order ≤ 1 when it is entire. We finally remark that in [8] and [10] we obtained the analytic properties of $F(s;f)$ in certain cases with $\kappa_0 > 1/d$, hence corresponding result can be obtained for $F(s;B)$.

3 Examples

As we already remarked, the Bessel functions are typical examples of functions $B(\xi)$. Indeed, by (3) and (4) of Sect. 7.13 of the Bateman Project [3], the Bessel J- and Y-functions satisfy, respectively,

$$J_\nu(z) = (\pi z/2)^{-1/2}\{\cos(z-\nu\pi/2-\pi/4)\sum_{k=0}^{K-1}\frac{c_k}{z^{2k}} - \sin(z-\nu\pi/2-\pi/4)\sum_{k=0}^{K-1}\frac{d_k}{z^{2k+1}}\} + O(|z|^{-2K})$$

and

$$Y_\nu(z) = (\pi z/2)^{-1/2}\{\sin(z-\nu\pi/2-\pi/4)\sum_{k=0}^{K-1}\frac{c_k}{z^{2k}} + \cos(z-\nu\pi/2-\pi/4)\sum_{k=0}^{K-1}\frac{d_k}{z^{2k+1}}\} + O(|z|^{-2K})$$

as $|z| \to \infty$ in $-\pi < \arg z < \pi$, where c_k and d_k are certain explicit coefficients. Writing $\sin z$ and $\cos z$ in terms of $e^{\pm iz}$ and choosing $z = 2\pi f(\xi, \boldsymbol{\alpha})$, by the Theorem we obtain the continuation over $\mathbb{C}$ as meromorphic or entire functions of the Bessel twists

$$\sum_{n=1}^{\infty}\frac{a(n)J_\nu(2\pi f(n,\boldsymbol{\alpha}))}{n^s} \quad \text{and} \quad \sum_{n=1}^{\infty}\frac{a(n)Y_\nu(2\pi f(n,\boldsymbol{\alpha}))}{n^s},$$

provided $F \in \mathcal{S}^\sharp$ has degree $d > 0$ and $f(\xi, \boldsymbol{\alpha})$ has $\kappa_0 \le 1/d$. In particular, we may choose $f(\xi, \boldsymbol{\alpha}) = \alpha\xi^{\kappa_0}$ with $\alpha \in \mathbb{R}$ and $\kappa_0 \le 1/d$.

We can apply the Theorem in many other similar situations, for instance with arbitrary Fox hypergeometric functions $H(z)$ in place of the above $J_\nu(z)$ and $Y_\nu(z)$. In fact, Braaksma [1] shows that any such function admits an asymptotic expansion, for $z = \xi \to \infty$, as a finite sum of terms of the form

$$e^{\alpha\xi^\kappa}\sum_{\nu=0}^{\infty}\frac{c_\nu}{\xi^{\nu_\nu}} \qquad \alpha \in \mathbb{C}.$$

Given $F \in S^\sharp$, the twist $F(s, H)$ converges somewhere if and only if $\Re\alpha \le 0$ in all such terms. If $\Re\alpha < 0$, the corresponding term of the twist converges everywhere and hence is entire. Otherwise, if $\Re\alpha = 0$ our Theorem applies, securing meromorphic continuation whenever the exponent κ does not exceed $1/d$.

We conclude with few details on the above remark about the order of $F(s;B)$. In general, suppose that we have a function $B(\xi)$ with expansion of type

$$B(\xi) = \sum_{|\nu_k| \le K} c_k \frac{e(-f_k(\xi,\alpha))}{\xi^{\nu_k}} + R_K(\xi),$$

where

$$|\nu_k| \gg k^{\delta}, \qquad c_k \ll A^{k^C}, \qquad R_K(\xi) \ll A^{K^C} \xi^{-K}$$

for certain positive constants δ, A and C and with a finite number of different exponents $f_k(\xi, \boldsymbol{\alpha})$. Then a computation shows that the corresponding twist $F(s;B)$ is of a finite order whenever entire. The assumption about the finiteness of the number of exponents can be relaxed, but instead we would need some uniform bounds for the involved coefficients $\boldsymbol{\alpha}$. This can be justified by means of the estimates in [9], see in particular Theorem 3 there. In particular, by the known bounds for the Bessel J-function we can prove that the twist in (3) is of order ≤ 1 whenever entire.

Acknowledgements This research was partially supported by the MIUR grant PRIN2010-11 *Arithmetic Algebraic Geometry and Number Theory* and by grant 2013/11/B/ST1/ 02799 *Analytic Methods in Arithmetic* of the National Science Centre.

References

1. B.L.J. Braaksma, Asymptotic expansions and analytic continuations for a class of Barnes-integrals. Compos. Math. **15**, 239–341 (1964)
2. J.B. Conrey, A. Ghosh, On the Selberg class of Dirichlet series: small degrees. Duke Math. J. **72**, 673–693 (1993)
3. A. Erdélyi, W. Magnus, F. Oberhettinger, F.G. Tricomi, *Higher Transcendental Functions*, vol. 2 (McGraw-Hill, New York, 1953)
4. J. Kaczorowski, Axiomatic theory of L-functions: the Selberg class, in *Analytic Number Theory, C.I.M.E. Summer School, Cetraro (Italy) 2002*, ed. by A. Perelli, C. Viola. Lecture Notes in Mathematics, vol. 1891 (Springer, Berlin/Heidelberg, 2006), pp. 133–209
5. J. Kaczorowski, A. Perelli, On the structure of the Selberg class, I: $0 \le d \le 1$. Acta Math. **182**, 207–241 (1999)
6. J. Kaczorowski, A. Perelli, The Selberg class: a survey, in *Number Theory in Progress, Proc. Conf. in Honor of A.Schinzel*, ed. by K. Győry et al. (de Gruyter, Berlin, 1999), pp. 953–992
7. J. Kaczorowski, A. Perelli, On the structure of the Selberg class, VI: non-linear twists. Acta Arithmetica **116**, 315–341 (2005)
8. J. Kaczorowski, A. Perelli, On the structure of the Selberg class, VII: $1 < d < 2$. Ann. Math. **173**, 1397–1441 (2011)
9. J. Kaczorowski, A. Perelli, Twists and resonance of L-functions, I. To appear in J. European Math. Soc.
10. J. Kaczorowski, A. Perelli, Twists and resonance of L-functions, II. Preprint (2014)
11. M.R. Murty, Selberg's conjectures and Artin L-functions. Bull. Am. Math. Soc. **31**, 1–14 (1994)
12. T. Noda, On the functional properties of Bessel zeta-functions. Submittted

13. T. Noda, On the functional properties of the confluent hypergeometric zeta-function. To appear in Ramanujan J
14. A. Perelli, A survey of the Selberg class of L-functions, part II. Riv. Mat. Univ. Parma **3***(7), 83–118 (2004)
15. A. Perelli, A survey of the Selberg class of L-functions, part I. Milan J. Math. **73**, 19–52 (2005)
16. A. Perelli, Non-linear twists of L-functions: a survey. Milan J. Math. **78**, 117–134 (2010)
17. A. Selberg, Old and new conjectures and results about a class of Dirichlet series, in *Proc. Amalfi Conf. Analytic Number Theory*, ed. by E. Bombieri et al. (Università di Salerno, Salerno, 1992), pp. 367–385; *Collected Papers*, vol. 2 (Springer, Berlin, 1991), pp. 47–63

The Sound of Fractal Strings and the Riemann Hypothesis

Michel L. Lapidus

Abstract We give an overview of the intimate connections between natural direct and inverse spectral problems for fractal strings, on the one hand, and the Riemann zeta function and the Riemann hypothesis, on the other hand (in joint works of the author with Carl Pomerance and Helmut Maier, respectively). We also briefly discuss closely related developments, including the theory of (fractal) complex dimensions (by the author and many of his collaborators, including especially Machiel van Frankenhuijsen), quantized number theory and the spectral operator (jointly with Hafedh Herichi), and some other works of the author (and several of his collaborators).

Keywords Riemann zeta function • Riemann hypothesis (RH) • Quantization • Quantized number theory • Fractal strings • Geometry and spectra • Direct and inverse spectral problems for fractal strings • Minkowski dimension • Minkowski measurability • Complex dimensions • Weyl–Berry conjecture • Fractal drums • Infinitesimal shift • Spectral operator • Invertibility • Quantized Dirichlet series and Euler product • Universality • Phase transitions • Symmetric and asymmetric criteria for RH

1 Riemann Zeros and Spectra of Fractal Strings: An Informal Introduction

Unlike an ordinary (Sturm–Liouville) vibrating string, which consists of a single interval (of length ℓ, say), a fractal string consists of infinitely many intervals (of lengths $\ell_1, \ell_2, \cdots, \ell_j, \cdots$, with $\ell_j \downarrow 0$ as $j \to \infty$), vibrating independently of each other. Hence, the (eigenvalue or frequency) spectrum of a fractal string consists of the union of the spectra (counting multiplicities) of each of the countably many ordinary strings of which it is composed.

M.L. Lapidus (✉)
Department of Mathematics, University of California, 900 University Avenue, Riverside, CA 92521-0135, USA
e-mail: lapidus@math.ucr.edu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_14

A fractal string (or, equivalently, its boundary, viewed as a compact subset of the real line $\mathbb{R}$) always has (fractal) Minkowski dimension D between 0 and 1, the most extreme case $D = 0$ and $D = 1$ being referred to (following [74]) as the least and most fractal case, respectively, while the case when $D = 1/2$ is referred to (also as in [74]) as the *midfractal case.* The latter case will play a key role throughout this paper.

By listening to a fractal string, one can detect whether or not one of its complex dimensions coincides with a nontrivial Riemann zero (that is, with a zero of the Riemann zeta function which is located in the critical strip $0 < \mathrm{Re}(s) < 1$). Indeed, it turns out that the Riemann zeta function $\zeta = \zeta(s)$ mediates between the geometry and the spectrum of a fractal string:

$$\zeta_\nu(s) = \zeta(s) \cdot \zeta_{\mathscr{L}}(s), \tag{1}$$

where $\zeta_{\nu,\mathscr{L}}(s) = \zeta_\nu(s) := \sum_{k=1}^{\infty} f_k^{-s}$ *is the spectral zeta function* of the fractal string $\mathscr{L}$, with $\{f_k\}_{k=1}^{\infty}$ denoting the sequence of (suitably normalized) frequencies of $\mathscr{L}$, written in nondecreasing order according to their multiplicities, and $\zeta_{\mathscr{L}}(s) := \sum_{j=1}^{\infty} \ell_j^s$ is the *geometric zeta function* of $\mathscr{L}$. This relation (discovered in [75, 76]) has played an important role in fractal string theory, as it enables one to understand why some of the complex dimensions of $\mathscr{L}$ (i.e., the poles of the meromorphic continuation of $\zeta_{\mathscr{L}}$) may be "canceled" by (nontrivial) zeros of ζ. Accordingly, certain oscillations which are present in the intrinsic geometry of the fractal string are no longer "visible" (or rather, "audible") in the spectrum of $\mathscr{L}$.

More poetically, it is shown in [86, 87] that "One can hear the shape of a fractal string of dimension $D \neq 1/2$" (in the sense of a certain inverse spectral problem, denoted by $(\mathrm{ISP})_D$ and to be specified in Sect. 6 below, and not in the original sense of Kac [65]) if and only if the Riemann hypothesis is true. Moreover, one cannot hear it for all of the fractal strings of dimension $1/2$ (because $\zeta = \zeta(s)$ has zeros on the critical line $\mathrm{Re}(s) = 1/2$). Hence, in the present approach, the truth of the Riemann hypothesis is equivalent to the existence of a (mathematical) phase transition at $D = 1/2$ and at no other dimension D in the "critical interval" $(0, 1)$. (See Theorems 6.15 and 6.17 below.) By "hearing the shape of a fractal string" here, we mean that the inverse spectral problem $(\mathrm{ISP})_D$ in Sect. 6 below has an affirmative answer for any fractal string of dimension D.

Phrased a little more precisely: one can hear the shape of a fractal string of a given (Minkowski) dimension $D \in (0, 1)$ if and only if the Riemann zeta function $\zeta(s)$ does not have any zero on the vertical line $\mathrm{Re}(s) = D$. (See Theorem 6.14.) Furthermore, in general, one cannot hear the shape of a fractal string in the midfractal case where $D = 1/2$. (See Corollary 6.16.) Consequently, one can hear the shape of a fractal string in every possible dimension $D \in (0, 1)$ (other than $1/2$) if and only if the Riemann hypothesis [150] is true; that is, if and only if

$$\zeta(s) = 0, \quad 0 < \mathrm{Re}(s) < 1 \Rightarrow \mathrm{Re}(s) = \frac{1}{2}. \tag{2}$$

(See Theorem 6.15). These results have been established by Lapidus and Maier in [87] (announced in 1991 in [86]) building on the author's earlier work [74] (see also [75, 76]) on a partial resolution of the Weyl–Berry conjecture in any dimension [6, 7]) as well as on the ensuing work of Lapidus and Pomerance [94] (announced in 1990 in [93]) on a resolution of the one-dimensional (modified) Weyl–Berry conjecture (of [74]) and its unexpected connections with the Riemann zeta function. (See Conjecture 5.9 and Theorem 5.11 in Sect. 5.)

Later on, these results of [87] (which made use at the heuristic level of the intuition of complex dimensions) were reinterpreted (by Lapidus and van Frankenhuijsen in [99, 100]) in terms of the then rigorously defined notion of complex dimension, as well as extended to a large class of Dirichlet series and integrals, including all of the arithmetic zeta functions for which the generalized Riemann hypothesis is expected to hold. (See also [101, Chap. 9].) Moreover, a method similar to the one used in [87], but now relying in part on the explicit formulas established in [99–101] (and direct computations as well as on inverse spectral problems), was used to show that the Riemann zeta function, along with a large class of Dirichlet series and integrals (including most of the arithmetic zeta functions [131, 132, 157] and [101, Appendix A] or [79, Appendices B, C and E], other than the zeta functions of varieties over finite fields, for which the result clearly does not hold) cannot have an infinite vertical arithmetic progression of zeros. (See [101, Chap. 11] for this result and several extensions concerning the density of the zeros.)

Unknown at the time to Lapidus and van Frankenhuijsen of [99] (and of earlier papers on this and related subjects), this latter result about the zeros in arithmetic progression, in the special case of ζ, was already obtained by Putnam in [138, 139] by a completely different method, which could not be generalized to this significantly broader setting. This turned out to be quite beneficial to the general theory of complex dimensions as it led us in [99–101] to significantly improve and refine the authors' original pointwise and distributional explicit formulas. (See, e.g., [101, Chap. 5].)

In the rest of this paper, we will be more specific and explain what type of inverse spectral problem is involved here. (See, especially, Sect. 6.) First of all, we will need to precisely define (in Sect. 2) what is a fractal string as well as its Minkowski dimension and content.

We will then recall (in Sect. 3) results from [93, 94] providing a characterization of the notions of Minkowski measurability and nondegeneracy, which will play a key role in Sect. 5 and part of Sect. 6 (as well as serve as a motivation for some aspects of Sect. 7.4). In Sect. 4, we will discuss Weyl's asymptotic formula and conjecture [180, 181] for the spectral asymptotics of drums with smooth (or sufficiently "regular") boundary, as well as the Weyl–Berry conjecture [6, 7] for drums with fractal boundary (or "fractal drums") and its partial resolution obtained in [74]. In Sect. 5, we will present the resolution of the modified Weyl–Berry (MWB) conjecture [74] for fractal strings obtained in [93, 94], thereby establishing a precise connection between the corresponding *direct spectral problem* and the Riemann zeta function $\zeta = \zeta(s)$ in the critical interval $0 < s < 1$. In Sect. 6, we will

introduce the aforementioned *inverse spectral problem* $(\mathrm{ISP})_D$, for each $D \in (0,1)$, and then precisely state the results of [86, 87] connecting it with the presence of zeros of the Riemann zeta function in the critical strip $0 < \mathrm{Re}(s) < 1$ (the so-called critical or nontrivial zeros), and thereby, with the Riemann hypothesis.

Finally, in Sect. 7, we will discuss a variety of topics, closely connected to (or motivated in part by) the above developments. The subjects to be discussed include the mathematical theory of complex dimensions of fractal strings developed in [99–101] (see Sect. 7.2), its higher-dimensional counterpart recently developed in [108–113, 115, 116] (see Sect. 7.3), as well as aspects of "quantized number theory" (see Sect. 7.4) developed in [53] (and [49–52] along with [80]). In particular, in the latter subsection, we will see that the aforementioned inverse spectral problem $(\mathrm{ISP})_D$ can be rigorously reinterpreted in terms of the invertibility (or "quasi-invertibility") of the "spectral operator" $\mathfrak{a} = \zeta(\partial)$, with the "infinitesimal shift" (of the real line) ∂ now playing the role of the usual complex variable s in the definition of the quantum (or operator-valued) analog of the classic Riemann zeta function $\zeta = \zeta(s)$.

We close this introduction by providing several relevant references. For general references concerning the theory of the Riemann zeta function and related aspects of analytic number theory, we mention, for example, [29, 58, 59, 66, 79, 101, 131–133, 165, 174] along with the relevant references therein. For fractal string theory and the associated theory of complex dimensions, along with their applications to a variety of subjects, including fractal geometry, spectral geometry, number theory, and dynamical systems, we refer to [101], along with [30, 31, 44, 47, 49–53, 72–88, 90–97, 103–116, 119, 130, 134, 135, 140, 144, 172, 173, 183, 184] and the relevant references therein. In particular, Chap. 13 of [101] provides an exposition of a number of recent extensions and applications of the theory, including to fractal sprays (higher-dimensional analogs of fractal strings [95]) and self-similar systems ([101, §13.1], based on [91, 92, 104, 105, 134, 135]), p-adic (or nonarchimedean) geometry ([101, §13.2], based on [83–85, 107, 114]), multifractals ([101, §13.3], based on [30, 96, 103]), random fractal strings ([101, §13.4], based on [44]), as well as fractal membranes and the Riemann (or modular) flow on the moduli space of fractal membranes ([101, §13.5], based on the book [79] and on [88]). As a general rule, we will give specific references to the most recent monograph [101] rather than to the earlier research monographs [99, 100].

2 Fractal Strings and Minkowski Dimension

A (nontrivial) fractal string $\mathscr{L}$ is a bounded open set $\Omega \subset \mathbb{R}$ which is not a finite union of intervals. Hence, Ω is an infinite countable disjoint union of (bounded) intervals I_j, of lengths $\ell_j = |I_j|$, for $j = 1, 2, \cdots$: $\Omega = \cup_{j=1}^{\infty} I_j$. Since $|\Omega| = \sum_{j=1}^{\infty} \ell_j < \infty$, we may assume without loss of generality that $(\ell_j)_{j=1}^{\infty}$ is nonincreasing and so $\ell_j \downarrow 0$ as $j \to \infty$. Furthermore, for our purposes, we may identify a fractal string with its associated sequence of lengths (or scales), $\mathscr{L} :=$

$(\ell_j)_{j=1}^{\infty}$, written as above (counting multiplicities). Indeed, all of the geometric notions we will work with, such as $V(\varepsilon)$, the Minkowski dimension and content, as well as the geometric zeta function and the complex dimensions, depend only on $\mathscr{L} = (\ell_j)_{j=1}^{\infty}$ and not on the particular geometric realization of $\mathscr{L}$ as a bounded open subset Ω of $\mathbb{R}$.

Given $\varepsilon > 0$, the (inner) *ε-neighborhood* (or *inner tube*) of $\partial\Omega$ (or of $\mathscr{L}$) is given by

$$\Omega_\varepsilon := \{x \in \Omega : d(x, \partial\Omega) < \varepsilon\}, \tag{3}$$

where $\partial\Omega$ denotes the boundary of Ω (a compact subset of $\mathbb{R}$) and $d(\cdot, \partial\Omega)$ denotes the (Euclidean) distance to $\partial\Omega$.

In the sequel, we will let

$$V(\varepsilon) = V_{\mathscr{L}}(\varepsilon) := |\Omega_\varepsilon|, \tag{4}$$

the volume (really, the length or one-dimensional Lebesgue measure) of Ω_ε. As was mentioned above, it can be shown (cf. [93, 94]) that $V(\varepsilon)$ depends only on $\mathscr{L}$ (and not on the particular geometric realization Ω of $\mathscr{L}$). Then, given $d \geq 0$, the *d-dimensional upper Minkowski content* of $\mathscr{L}$ (or of $\partial\Omega$) is given by

$$\mathscr{M}_d^* = \mathscr{M}_d^*(\mathscr{L}) := \limsup_{\varepsilon \to 0^+} \frac{V(\varepsilon)}{\varepsilon^{1-d}}, \tag{5}$$

and similarly for the d-dimensional lower Minkowski content $\mathscr{M}_{*,d} = \mathscr{M}_{*,d}(\mathscr{L})$, except for the upper limit replaced by a lower limit.

The *Minkowski dimension* of $\mathscr{L}$ (or of its boundary $\partial\Omega$),[1] denoted by $D = D_{\mathscr{L}}$, is defined by

$$\begin{aligned} D &:= \inf\{d \geq 0 : \mathscr{M}_d^* < \infty\} \\ &= \sup\{d \geq 0 : \mathscr{M}_d^* = +\infty\}. \end{aligned} \tag{6}$$

We note that since $|\Omega| < \infty$, we always have $0 \leq D \leq 1$. Furthermore, from the physical point of view, D will play the role of a *critical exponent* (or *critical parameter*): it is the unique real number D such that $\mathscr{M}_d^* = +\infty$ for $d < D$ and $\mathscr{M}_d^* = 0$ for $d > D$.

The *upper* (resp. *lower*) *Minkowski content* of $\mathscr{L}$ is defined by $\mathscr{M}^* := \mathscr{M}_D^*$ (resp., $\mathscr{M}_* := \mathscr{M}_{*,D}$). We always have $0 \leq \mathscr{M}_* \leq \mathscr{M}^* \leq \infty$. If $0 < \mathscr{M}_* (\leq) \mathscr{M}^* < \infty$, then $\mathscr{L}$ is said to be *Minkowski nondegenerate*. If, in addition, $\mathscr{M}_* = \mathscr{M}^*$ (i.e., if the upper limit in Eq. (5) is a true limit in $(0, +\infty)$, with $d := D$),

[1] This is really the upper Minkowski dimension of $\partial\Omega$, relative to Ω, but we will not stress this point in this paper (except perhaps in Sect. 7.3).

then we denote by $\mathscr{M}$ this common value, called the *Minkowsi content* of $\mathscr{L}$, and the fractal string $\mathscr{L}$ (or its boundary $\partial\Omega$) is said to be *Minkowski measurable.* So that $\mathscr{M} = \lim_{\varepsilon\to 0+} V(\varepsilon)/\varepsilon^{1-D}$ and $0 < \mathscr{M} < \infty$.

We close this section by stating a theorem that shows the intimate connections between the geometric zeta function $\zeta_{\mathscr{L}}$ of a fractal string $\mathscr{L}$ [introduced in Sect. 1 above, just after Eq. (1)], and the Minkowski dimension of $\mathscr{L}$. (Recall that $\zeta_{\mathscr{L}}(s) = \sum_{j=1}^{\infty} \ell_j^s$, for all $s \in \mathbb{C}$ with $\mathrm{Re}(s)$ sufficiently large.) This result was first observed by Lapidus in [75, 76], using earlier work of Besicovich and Taylor in [8], and has since then been given several direct proofs in [101, Theorem 1.10] and in [101, Theorem 13.111]; see also [114]. It also follows from general results in [110, 116].

Theorem 2.1 ([75, 76]). *Let $\mathscr{L}$ be a fractal string. Then, the abscissa of convergence of $\zeta_{\mathscr{L}}$ coincides with the Minkowski dimension of $\mathscr{L}$: $\sigma = D$.*

Recall that the *abscissa of convergence* of $\zeta_{\mathscr{L}}$ is given by

$$\sigma := \inf\left\{\varrho \in \mathbb{R} : \sum_{j=1}^{\infty} \ell_j^{\varrho} < \infty\right\}. \tag{7}$$

It follows from this definition and from known results about Dirichlet series with positive coefficients (see, e.g., [165, §VI.2.2 and §VI.2.3]) that $\zeta_{\mathscr{L}} = \zeta_{\mathscr{L}}(s)$ is holomorphic for $\mathrm{Re}(s) > D$ and that the open right half-plane $\{\mathrm{Re}(s) > D\}$ is the largest right half-plane (of the form $\{\mathrm{Re}(s) > \alpha\}$, for some $\alpha \in \mathbb{R} \cup \{\pm\infty\}$) to which $\zeta_{\mathscr{L}}$ can be holomorphically continued as well as the largest such half-plane in which the Dirichlet series $\sum_{j=1}^{\infty} \ell_j^s$ is absolutely convergent (and hence, convergent). Therefore, $\zeta_{\mathscr{L}}$ does not have any pole in $\{\mathrm{Re}(s) > D\}$. (We refer, e.g., to [165, §V1.2 and §V1.3] or to [45] for an introduction to the theory of Dirichlet series.)

Furthermore, note that $s = D$ is always a singularity of $\zeta_{\mathscr{L}}$ (i.e., $\zeta_{\mathscr{L}}(s) \to +\infty$ as $s \to D^+$, $s \in \mathbb{R}$) and therefore, if $\zeta_{\mathscr{L}}$ can be meromorphically continued to an open (connected) neighborhood of D, then D is a pole of $\zeta_{\mathscr{L}}$. Theorem 2.1 above (according to which $D = \sigma$), along with this last observation, is one of the original justifications for calling the poles of a meromorphic continuation of $\zeta_{\mathscr{L}}$ (to an open connected neighborhood of $\{\mathrm{Re}(s) > D\}$) the (visible) complex dimensions of the fractal string $\mathscr{L}$. (See [99–101] and, for a brief introduction, see Sect. 7.2 below.)

Here and thereafter, $\{\mathrm{Re}(s) > \alpha\}$ stands for $\{s \in \mathbb{C} : \mathrm{Re}(s) > \alpha\}$. By convention, for $\alpha = -\infty$ or $+\infty$, respectively, it coincides with $\mathbb{C}$ or the empty set $\emptyset$. Similarly, if $\alpha \in \mathbb{R}$, $\{\mathrm{Re}(s) = \alpha\}$ denotes the vertical line $\{s \in \mathbb{C} : \mathrm{Re}(s) = \alpha\}$.

3 Characterization of Minkowski Measurability and Nondegeneracy

We recall here some of the joint results of Lapidus and Pomerance obtained in [93, 94]. Further results from that work will be discussed in Sect. 5.

Theorem 3.2 (Characterization of Minkowski Measurability [94]). *Let $\mathscr{L} = (\ell_j)_{j=1}^{\infty}$ be a fractal string of Minkowski dimension $D \in (0,1)$. Then, $\mathscr{L}$ is Minkowski measurable if and only if*

$$\ell_j \sim Lj^{-1/D} \text{ as } j \to \infty, \tag{8}$$

for some constant $L \in (0, +\infty)$. In that case, the Minkowski content of $\mathscr{L}$ is given by

$$\mathscr{M} = \frac{2^{1-D}}{1-D} L^D. \tag{9}$$

Equation (8) precisely means that the limit of $\ell_j \cdot j^{1/D}$ exists in $(0, +\infty)$ and is equal to L. Furthermore, note that (8) is equivalent to

$$N_{\mathscr{L}}(x) \sim Mx^D \text{ as } x \to +\infty, \tag{10}$$

with $M := L^D$ and where $N_{\mathscr{L}}$, the *geometric counting function* of $\mathscr{L}$, is defined for all $x > 0$ by

$$N_{\mathscr{L}}(x) = \#(\{j \geq 1 : \ell_j^{-1} \leq x\}). \tag{11}$$

Here and thereafter, $\#B$ denotes the number of elements of a finite set B. Moreover, Eq. (10) precisely means that $x^{-D} N_{\mathscr{L}}(x) \to M$ as $x \to +\infty$, for some $M \in (0, +\infty)$.

Although we will not use this result explicitly, it is helpful to also give the counterpart of Theorem 3.2 for Minkowski nondegeneracy. Let α_* (resp., α^*) be the lower (resp., upper) limit of $\ell_j \cdot j^{1/D}$ as $j \to \infty$; so that we always have $0 \leq \alpha_* \leq \alpha^* \leq \infty$.

Theorem 3.3 (Characterization of Minkowski Nondegeneracy [94]). *Let $\mathscr{L}$ be a fractal string of dimension $D \in (0,1)$. Then, $\mathscr{L}$ is Minkowski nondegenerate (i.e., $0 < \mathscr{M}_*(\leq)\mathscr{M}^* < \infty$) if and only if*

$$0 < \alpha_*(\leq)\alpha^* < \infty. \tag{12}$$

Concretely, Eq. (12) means that there exist positive constants c, $C \geq 1$ such that $c^{-1}j^{-1/D} \leq \ell_j \leq cj^{-1/D}$ for all $j \geq 1$ or, equivalently, $C^{-1}x^D \leq N_{\mathscr{L}}(x) \leq Cx^D$ for all $x > 0$.

Let $N_\nu = N_{\nu,\mathscr{L}}$ denote the *spectral* (i.e., *frequency*) *counting function* of $\mathscr{L}$. Hence, $N_\nu(x) := \#(\{f \in \sigma(\mathscr{L}) : f \leq x\})$ for all $x > 0$, where $\sigma(\mathscr{L}) = \{n \cdot \ell_j^{-1} : n \geq 1, j \geq 1\}$ denotes the (*frequency*) *spectrum* of $\mathscr{L}$. (In essence, the *frequencies* of $\mathscr{L}$ are, up to a multiplicative normalizing positive constant, equal to the square roots of the eigenvalues of $\mathscr{L}$; that is, of the eigenvalues of the Dirichlet Laplacian $-\Delta = -d^2/dy^2$ on any geometric representation Ω of $\mathscr{L}$ by a bounded open subset of $\mathbb{R}$.)

It is also shown in [94] that Eq. (12) (and hence also, the Minkowski nondegeneracy of $\mathscr{L}$, according to Theorem 3.3) is equivalent to

$$0 < \delta_*(\leq)\delta^* < \infty, \tag{13}$$

where δ_* (resp., δ^*) denotes the lower (resp., upper) limit of $\varphi_\nu(x)/x^D$ as $x \to +\infty$, and $\varphi_\nu(x)$ is the "asymptotic second term" for $N_\nu(x)$; namely, $\varphi_\nu(x) := W(x) - N_\nu(x)$, with $W(x) := |\Omega|x$ being the Weyl (or leading) term to be discussed in Sects. 4.1, 5, and 6 below. (In general, one always has $\varphi_\nu(x) \geq 0$ for all $x > 0$ and thus $0 \leq \delta_* \leq \delta^* \leq \infty$.)

Equation (13) means that there exist $c_1 \geq 1$ such that $c_1^{-1}x^D \leq \varphi_\nu(x) \leq c_1 x^D$ for all $x > 0$; in other words, the error estimates of [74], to be discussed further on in Theorem 4.8 of Sect. 4.1, are sharp in this case. More specifically, when $N = 1$ and $D \in (0, 1)$, each of the equivalent conditions (12) and (13) characterizes the sharpness of the remainder estimates of [74], recalled in Eq. (19) (see Theorem 4.8(i) below).

It is natural to wonder whether the counterpart of Theorem 3.3 and of its complement for N_ν holds for one-sided (rather than two-sided estimates). The answer may be somewhat surprising to the reader. In short, it is positive for the upper estimates but negative for the lower estimates. More specifically, still in [94], it is shown that

$$\mathscr{M}^* < \infty \Leftrightarrow \ell_j = O(j^{-1/D}) \text{ as } j \to \infty \text{ (i.e., } \alpha^* < \infty) \Leftrightarrow N_{\mathscr{L}}(x) = O(x^D) \tag{14}$$

$$\text{as } x \to +\infty \Leftrightarrow \varphi_\nu(x) = O(x^D) \text{ as } x \to +\infty, \text{ (i.e., } \delta^* < \infty),$$

and, more generally, that the exact counterpart of (14) holds for any $d > 0$, provided D and $\mathscr{M}^* = \mathscr{M}_D^*$ are replaced by d and $\mathscr{M}_d^*$, respectively. This provides, in particular, a converse (in the present one-dimensional case) to the error estimate obtained in [74] for the asymptotic second term of $N_\nu(x)$ to be discussed next; see part (i) of Theorem 4.8 below, specialized to the case where $N = 1$ and $D \in (0, 1)$.

Moreover, it is shown in [94] by means of an explicit counterexample that in the analog of (14) for $\mathscr{M}_*$, the implications in one direction holds, but not in the other direction. For example, it is *not* true, in general, that $\mathscr{M}_* > 0 \Leftrightarrow \alpha_* > 0 \Leftrightarrow \delta_* > 0$.

Here, in Eq. (14), as well as in the sequel, given $f : [0, +\infty) \to \mathbb{R}$ and $g : [0, +\infty) \to [0, +\infty)$, one writes that $f(x) = O(g(x))$ to mean that there exists a positive constant c_2 such that $|f(x)| \leq c_2 g(x)$, for all x sufficiently large. (For the type of functions or sequences we will work with, we may assume that this inequality holds for all $x > 0$.) We use the same classic Landau notation for sequences instead of for functions of a continuous variable.

In closing this section, we note that under mild assumptions on $\mathscr{L}$ (about the growth of a suitable meromorphic continuation of its geometric zeta function $\zeta_{\mathscr{L}}$), it has since then been shown that within the theory of complex dimensions developed in [99–101], the characterization of Minkowski measurability obtained in [94] (Theorem 3.2 above) can be supplemented as follows (see [101], Theorem 8.15]), under appropriate hypotheses.

Theorem 3.4 (Characterization of Minkowski Measurability Revisited [94, 101]). *Let $\mathscr{L}$ be a fractal string of Minkowski dimension D. Then, under suitable conditions on $\zeta_{\mathscr{L}}$ (specified in [101, Sect. 8.3]), the following statements are equivalent:*

(i) *$\mathscr{L}$ is Minkowski measurable.*
(ii) *Condition (8), or equivalently (10), holds.*
(iii) *D is the only complex dimension of $\mathscr{L}$ with real part D, and it is simple.*

Moreover, if any of these equivalent conditions is satisfied, then the Minkowski content of $\mathscr{L}$ is given by

$$\mathscr{M} = 2^{1-D}\frac{M}{1-D} = 2^{1-D}\frac{res(\zeta_{\mathscr{L}}, D)}{D(1-D)}, \tag{15}$$

where $res(\zeta_{\mathscr{L}}, D)$ denotes the residue of $\zeta_{\mathscr{L}}(s)$ at $s = D$, $M := L^D$ and L and M are given as in (8) and (10), respectively.

Recall that the complex dimensions of $\mathscr{L}$ are the poles of a (necessarily unique) meromorphic continuation of $\zeta_{\mathscr{L}}$ (to a connected open neighborhood of $\{\mathrm{Re}(s) > D\}$). According to a result of [75, 76] (see Theorem 2.1 above), there are no complex dimensions with real parts $> D$; see the discussion following Eq. (7) above.

Remark 3.5. When $\mathscr{L}$ is a self-similar string (in the sense of [101, Chap. 2] i.e., the boundary of $\mathscr{L}$ is a self-similar set), it is shown in [101, Theorems 2.16, 8.23 and 8.36] that no hypothesis on $\zeta_{\mathscr{L}}$ is needed in the counterpart of Theorem 3.4, and that either of the equivalent statements (i), (ii), or (iii) of Theorem 3.4 is true if and only if $\mathscr{L}$ is nonlattice; i.e., iff the subgroup of $\mathbb{R}$ generated by the logarithms of its distinct scaling ratios is not of the form $\varrho\mathbb{Z}$, for some $\varrho > 0$. If $\mathscr{L}$ is lattice, then it is Minkowski nondegenerate but is not Minkowski measurable.

Finally, we mention that the original proof of Theorem 3.2 given in [93, 94] was analytical and combinatorial in nature. Several parts of the proof have since been established in a different manner by Falconer in [31], using some techniques from the theory of dynamical systems, and more recently (and perhaps most concisely), by Rataj and Winter in [144] using techniques from geometric measure theory (in particular, from [143, 169] and the relevant references therein).

Remark 3.6. We also note that the notion of Minkowski dimension was introduced (for noninteger values of the dimension) by Bouligand in the late 1920s in [16]. The notion of (normalized) Minkowski content was introduced by Feder in [33], while that of Minkowski measurability was apparently first used by Stachó in [169].

The Minkowski dimension is also often called "box dimension" (see, e.g., [32, 97, 101, 121, 122, 175] or the applied literature on fractal dimensions), entropy dimension, or capacity dimension, for example. Contrary to the Hausdorff dimension, it is not σ-stable [that is, it is usually not true that $D(\cup_{k=1}^{\infty} A_k) = \sup_{k\geq 1} D(A_k)$],

where $D(A)$ denotes the upper Minkowski dimension of A; see, e.g., [32, 122, 175].[2] Furthermore, unlike the Hausdorff measure (which is a true positive Borel measure, in the usual mathematical sense of the term [21, 32, 35, 97, 122]), the Minkowski content (when it exists) or more generally, the upper Minkowski content, is not a measure (it is only finitely sub-additive). In fact, as is pointed out in [74] (see also [75, 76]), and somewhat paradoxically, this is precisely because it does not have all of these desirable mathematical properties that the Minkowski dimension is important in the study of aspects of harmonic analysis as well as of spectral and fractal geometry, including the study of the vibrations and spectra of fractal drums or "drums with fractal boundary" ([74–76], see also Sect. 4 below) and, in particular, of fractal strings (i.e., one-dimensional drums with fractal boundary). See, e.g., [74, Examples 5.1 and 5.1′].

4 The Weyl–Berry Conjecture for Fractal Drums

In this section, we first briefly recall Weyl's classic formula for the leading spectral asymptotics of ordinary (smooth or piecewise smooth) drums, as well as corresponding extensions and sharp remainder estimates (from [74]), valid for general fractal drums and providing a partial resolution of the Weyl–Berry conjecture [6, 7] in any dimension; see Sect. 4.1. (In Sect. 5, we will specialize the situation to the one-dimensional case, which is the main focus of this paper.) In the latter part of this section, we will comment on various aspects of Weyl's formula for fractal drums, both in the case of drums with fractal boundary (which is of most interest here) and in the related case of drums with fractal membrane (corresponding to Laplacians and Dirac operators on fractals themselves); see Sect. 4.2. We will also conclude this section by briefly mentioning some of the physical and technological applications of these mathematical results about fractal drums and of the original (or modified) Weyl–Berry conjecture.

4.1 Weyl's Asymptotic Formula with Sharp Error Term for Fractal Drums

Hermann Weyl's classic asymptotic formula ([180, 181]; see also, e.g., [23, 146]) for the frequency (or spectral) counting function $N_\nu = N_\nu(x)$ of an ordinary (N-dimensional) drum can be stated as follows:

$$N_\nu(x) = \mathscr{C}_N |\Omega|_N x^N + o(x^N) \tag{16}$$

[2] A simple counterexample is provided by $A := \{1/k : k \geq 1\}$ and $A_k := \{\frac{1}{k}\}$ for each $k \geq 1$, viewed as subsets of $\mathbb{R}$; then, $D(A) = 1/2$, whereas $D(A_k) = 0$ for all $k \geq 1$ and hence, $\sup_{k\geq 1} D(A_k) = 0$.

as $x \to +\infty$. Here, $|\Omega| := |\Omega|_N$ denotes the volume of the bounded open set $\Omega \subset \mathbb{R}^N$ (i.e., the N-dimensional Lebesgue measure of Ω). Furthermore, given $x > 0, N_\nu(x)$ is the number of (suitably normalized) frequencies f of the drum not exceeding x, and $\mathscr{C}_N$ is an explicitly known positive constant which can be expressed in terms of the volume of the unit ball of $\mathbb{R}^N$ [and therefore, in terms of appropriate values of the gamma function $\Gamma = \Gamma(s)$]. In the sequel, we denote the *leading term* in (16) by

$$W(x) := \mathscr{C}_n |\Omega|_N x^N \tag{17}$$

and call it the *Weyl term*. In addition, mathematically, the spectrum of the (N-dimensional) "drum" is interpreted as the spectrum of the Dirichlet Laplacian $-\Delta$ on a given nonempty bounded open set $\Omega \subset \mathbb{R}^N (N \geq 1)$, with boundary $\partial\Omega$. Furthermore, the Dirichlet boundary conditions are interpreted variationally (or in the distributional sense); see, e.g., [17, 120] or [74, §2].

Since the Laplacian is a second order (self-adjoint, positive) linear operator, the aforementioned (normalized) frequencies are (up to a multiplicative constant) equal to the square roots of the eigenvalues. In the present situation, the spectrum is discrete and hence, we can order these frequencies in nondecreasing order (and according to their multiplicities) as follows:

$$0 < f_1 \leq f_2 \leq \cdots \leq f_n \leq \cdots, \text{ with } f_n \to +\infty \text{ as } n \to \infty. \tag{18}$$

Remark 4.7. Physically, $W(x)$ can be interpreted as a volume in phase space. More specifically, $W(x)$ is proportional to

$$|\{(x, \xi) \in \Omega \times \mathbb{R}^N : |\xi|^2 \leq x^2\}|_{2N} = |B(0,1)|_N \, |\Omega|_N \, x^N,$$

where $\mathbb{R}^N \times \mathbb{R}^N \approx \mathbb{R}^{2N}$ is the "phase space" (the space of positions and velocities or equivalently, momenta, of the classical particle) and for $\varrho > 0, B(0, \varrho)$ denotes the ball of center the origin and radius ϱ in $\mathbb{R}^N$.

In [6, 7], extending to the fractal case a classic conjecture of the mathematician Weyl [180, 181] in the "regular" (or "smooth" case),[3] the physicist Michael Berry has conjectured that (16) should be completed to obtain an asymptotic second term for the spectral counting function $N_\nu(x)$, of the form $-C_{N,H}\mathscr{H}_H(\partial\Omega)x^H =: S(x)$, where H is the Hausdorff dimension of the boundary $\partial\Omega$ ($H \in [N-1, N]$), $\mathscr{H}_H(\partial\Omega)$ is the H-dimensional Hausdorff measure of $\partial\Omega$ (a well-known fractal generalization to noninteger dimensions of the notion of "volume"; see, e.g., [32, 33, 97, 122, 175]), and $C_{N,H}$ is a positive constant independent of Ω and depending only on N and H

[3] See, e.g., [55–57, 60–62, 71, 124, 125, 162–164] (along with the relevant references therein and in [101, §12.5 and Appendix B]) for results (in the smooth case) concerning the Weyl conjecture about the asymptotic second term of the spectral counting function $N_\nu(x)$.

(as well as expressed in terms of the gamma function, by analogy with the known results in integer dimensions for simple examples such as N-dimensional cubes).

Unfortunately, it turns out that Berry's conjecture (called the Weyl–Berry conjecture in [74] and in the literature since then), although very stimulating, is not correct, as was first noted by Brossard and Carmona by means of an explicit counterexample in [18] and then, explained from a mathematical point of view (and illustrated by a family of even simpler counterexamples) in [74]. (See, in particular, [74, Examples 5.1 and 5.1′].) Furthermore, H should be replaced by D, the (inner) Minkowski dimension of $\partial\Omega$, and $\mathscr{H}_H(\partial\Omega)$ might reasonably be replaced by $\mathscr{M}_D(\partial\Omega)$, the (inner) Minkowski content of $\partial\Omega$. Finally, as was shown in [75, 76, 94, 95], even the expected constant $C_{N,H}$ (when it exists) does not simply take the form of $C_{N,D}$ (where D is the Minkowski dimension of $\partial\Omega$) but whatever replaces the factor of proportionality in the counterpart of $S(x)$ should merely be expressed in terms of the residue at $s = D$ of the meromorphic continuation of the corresponding spectral zeta function $\zeta_\nu(s) = \sum_{n=1}^{\infty} f_n^{-s}$, where $(f_n)_{n=1}^{\infty}$ is the sequence of frequencies of the drum, as given in Eq. (18) above. In fact, in Sect. 5, we will see that when $N = 1$ (the special case of fractal strings instead of higher-dimensional fractal drums) and when $\partial\Omega$ (or, equivalently, the fractal string) is Minkowski measurable (with Minkowski content denoted by $\mathscr{M}$), then $C_{1,D}$ is proportional to the positive number $-\zeta(D)$, where $\zeta = \zeta(s)$ is the classic Riemann zeta function and $D \in (0, 1)$. (See Theorem 5.11 below, where in light of (23), $C_{1,D}$ is of the form $c_D\mathscr{M}$, with c_D given by Eq. (24).)

On the positive side, the following partial resolution of the Weyl–Berry conjecture was obtained by Lapidus in [74] (recall that the Weyl term $W(x)$ is given by (17) above and is therefore proportional to x^N):

Theorem 4.8 (Sharp Error Estimates [74]). *Let Ω be any (nonempty) bounded open subset of $\mathbb{R}^N$. Recall that we always have $D \in [N-1, N]$, where $D = D(\partial\Omega)$ is the (inner) Minkowski dimension of $\partial\Omega$; see comment (b) in Sect. 4.2 just below.*

(i) Then, in the "fractal case" where $D \in (N-1, N]$, we have

$$N_\nu(x) = W(x) + O(x^D) \text{ as } x \to +\infty, \tag{19}$$

provided $\mathscr{M}_D^*(\partial\Omega) < \infty$, where the (pointwise) remainder estimate $O(x^D)$ is sharp for every $D \in (N-1, N)$. Furthermore, if $\mathscr{M}_D^*(\partial\Omega) = +\infty$, then D should be replaced by $D + \varepsilon$, for any arbitrarily small $\varepsilon > 0$.

(ii) In the "nonfractal case" where $D = N - 1$ (which is the case, for example, if the boundary $\partial\Omega$ is piecewise smooth or, more generally, locally Lipschitz), then exactly the same result as in part (i) holds, except for the fact that the error term now takes the form $O(x^D \log x)$ as $x \to +\infty$ (with $D := N - 1$).

When $\mathscr{M}_D^*(\partial\Omega) = +\infty$, either in case (i) or (ii) of Theorem 4.8, one should try to use the later extension of this theorem obtained in [47] and expressed in terms of generalized (upper) Minkowski contents, relative to suitable gauge functions; see comment (g) in Sect. 4.2 just below.

Note that in the most fractal case when $D = N$, the remainder estimate (19) still holds but is clearly uninteresting because then, the error term is of the same order as the leading term $W(x)$ given by Eq. (17).

4.2 Further Comments, Extensions, and Applications

We next comment on various aspects and extensions (in related contexts) of Theorem 4.8 and of the Weyl–Berry conjecture. Just as in Sect. 4.1, our discussion is not meant to be exhaustive in any way but simply aims at providing various pointers and references where the interested reader can find a lot more detailed information.

(a) The nonfractal case where $D := N - 1$ (see part (ii) of Theorem 4.8 just above) was essentially already obtained by Métivier in [128] (see also [126, 127]), in a slightly less precise form and using a different terminology (not explicitly involving the notions of Minkowski dimension and content).

(b) In Theorem 4.8, the (inner) Minkowski dimension D and (inner) upper Minkowski content $\mathscr{M}_D^* := \mathscr{M}_D^*(\partial\Omega)$ are defined exactly as in the one-dimensional case where $N = 1$ (see Eqs. (6) and (5), respectively), except that on the right-hand side of (5) (with $d := D$), $V(\varepsilon)/\varepsilon^{1-D}$ should be replaced by $V(\varepsilon)/\varepsilon^{N-D}$. Since $\Omega \subset \mathbb{R}^N$ is bounded, it has finite volume and therefore it is easy to check that $D \leq N$. Furthermore, since $\partial\Omega$ is the boundary of a (nonempty) bounded open set, its topological dimension is equal to $N - 1$ and hence, $D \geq N - 1$. Consequently, as was observed in [74], we always have $N - 1 \leq D \leq N$. (For the special case of fractal strings, we have $N = 1$ and we thus recover the fact that $0 \leq D \leq 1$; see the statement following Eq. (6) in Sect. 2.)

In [74], the case where $D = N - 1$ is called the *least* or *nonfractal case*, the case where $D = N$ is called the *most fractal case*, while the case where $D = N - \frac{1}{2}$ (i.e., the *codimension* $N - D = \{D\}$ is equal to $1/2$, where $\{D\} \in [0, 1)$ is the fractional part of D), is referred to as the *midfractal case.* It turns out that each of these cases (where the codimension $N - D$ takes the value $0, 1$ or $1/2$, respectively)[4] plays an important role in the proof of Theorem 4.8 (and its generalizations) given in [74] (as well as in related spectral or geometric results obtained in [74, Examples 5.1 and 5.1′, along with Appendix C]).

(c) To show that the remainder estimates are sharp (in case (i) of Theorem 4.8 where $D \in (N - 1, N)$ and $\mathscr{M}_D^*(\partial\Omega) < \infty$), a simple one-family of examples is constructed in [74, Example 5.1 ($N = 1$) and Example 5.1′ ($N \geq 1$)]. When $N = 1$, it is of the form $\Omega_a := \cup_{j=1}^{\infty}((j+1)^{-a}, j^{-a})$, where $a > 0$ is arbitrary, so

[4] Strictly speaking, when $D = N - 1, N - D = 1$ is not equal to $\{1\}$.

that $\partial\Omega_a = \{0\} \cup \{j^{-a} : j \geq 1\}$. It follows that for every $a > 0$, $H = 0$ (since $\partial\Omega_a$ is countable), whereas $D = (a+1)^{-1}$, $\partial\Omega_a$ is Minkowski measurable (as defined in Sect. 2 above and shown in [74]) with Minkowski content $\mathscr{M} := \mathscr{M}_D(\partial\Omega)$ equal to $2^{1-D}a^D/(1-D)$; see [74, Appendix C]. This was the first explicit example of a fractal string and served as a motivation for the formulation of Theorem 3.2 above (the characterization of Minkowski measurability for fractal strings, obtained in [94]) and the MWB conjecture, stated in [74] and to be discussed in Sect. 5 below (see Conjecture 5.9). Note that the midfractal case where $D = 1/2$ corresponds to $a = 1$, while the most ($D = 1$) and least ($D = 0$) fractal cases correspond, respectively, to the limits $a \to 0^+$ and $a \to +\infty$. Furthermore, the symmetry $a \leftrightarrow 1/a$ exchanges D and $1 - D$.

Finally, when $N \geq 2$, exactly the same statements as above are true for the N-dimensional analog (a fractal comb) of Ω_a, defined by $\Omega_{a,N} := \Omega_a \times [0,1]^{N-1}$. We then have that $H = N - 1, D = N - 1 + (a+1)^{-1}$, $\partial\Omega_{a,N}$ is Minkowski measurable and $\mathscr{M}_D(\partial\Omega_{\alpha,N})$ has the same value as above. Furthermore, the remainder estimate (19) of part (i) of Theorem 4.8 is still sharp for every $a > 0$ [and hence, for every $D \in (N-1, N)$]; see [74, Example 5.1′]. Actually, it follows from the later results (from [94]) recalled in Sect. 5 below that for every $a > 0$,

$$N_\nu(x) = W(x) - c_{N,D}\,\mathscr{M}_D(\partial\Omega)x^D + o(x^D) \text{ as } x \to +\infty, \tag{20}$$

where $\Omega = \Omega_{a,N}$, $W(x)$ is given by (17), and $c_{N,D}$ is explicitly known.

(d) Under suitable hypotheses, Theorem 4.8 has an exact counterpart for the Neumann Laplacian (instead of the Dirichlet Laplacian), with the boundary conditions interpreted variationally (as in [17, 120] or [74, §2]) as well as for higher order, positive self-adjoint elliptic operators with (possibly) variable coefficients and with Dirichlet, Neumann, or mixed boundary conditions; see [74], Theorem 2.1 and its corollaries. For the Neumann Laplacian, the error estimate (19) holds, for instance, for the classic Koch snowflake domain and, more generally, for all planar quasidics [123, 136] (e.g., for the simply connected planar domains bounded by the Julia sets of the quadratic maps $z \mapsto z^2 + c$, where the complex parameter c is sufficiently small).

(e) In [18], for the Dirichlet Laplacian and using probabilistic methods, Brossard and Carmona have obtained error estimates (as $t \to 0^+$) for the partition function $Z_\nu = Z_\nu(t) := Tr(e^{t\Delta})$ (the trace of the heat semigroup), of interest in quantum statistical mechanics, probability theory, harmonic analysis, and spectral theory. These estimates are also expressed in terms of the (inner) Minkowski dimension D. We should note, however, that even though part (i) of Theorem 4.8 implies those error estimates, the converse is not true, in general. Indeed, typically, beyond the leading term (for which a classic Tauberian theorem can be used in order to show the equivalence of the corresponding results for $N_\nu(x)$ and $Z_\nu(t)$ (see, e.g., [65, 168]), pointwise asymptotic results for $N_\nu(x)$ are considerably more difficult to obtain than for $Z_\nu(t)$.

(f) The Weyl–Berry conjecture for drums with fractal boundary (and its modifications in [74]) has since been studied in a number of different contexts, both analytically and probabilistically. We refer to [101, §12.5] for a number of references on the subject, including [18, 34, 38, 39, 44, 47, 74–76, 86, 87, 93–95, 99–101, 109, 116, 129, 176]. We also refer [40, 55–57, 60–62, 71, 74–76, 124, 125, 162–164] and [101, Appendix B] for related references (including [40] in the case of partition functions) on the spectral asymptotics of smooth (as opposed to fractal) drums. A general introduction to Weyl's asymptotic formula and its analytic or probabilistic proof (for smooth or piecewise smooth boundaries, for example) can be found in [23, 40, 65, 146, 168].

Finally, we mention that the original Weyl–Berry conjecture was also formulated for "drums with fractal membrane" (as opposed to "drum with fractal boundary"). An appropriate modification of the conjecture was established in that setting by Kigami and Lapidus in [69] (building on a conjecture of [76]) for a large class of self-similar fractal drums corresponding to Laplacians *on* fractals (the so-called "finitely ramified" or p.c.f. self-similar sets), such as the Sierpinski gasket, the pentagasket, and certain fractal trees. In particular, in [69] is established a suitable analog of Weyl's classic asymptotic formula, but now for the leading spectral asymptotics of Laplacians on (self-similar) fractals (see, e.g., [68]) rather than on bounded open sets with fractal boundary. The resulting semi-classical formula, along with its geometric interpretation, is further explored and extended from the point of view of nonsmooth geometric analysis and Connes' noncommutative geometry [22], in [70, 77, 78] and (in a somewhat different setting and using suitably constructed Dirac operators and intrinsic geodesic metrics on the fractals under consideration), in [20, 98].

Beside [69, 76], see also, e.g., [101, §12.5.2] for many other references on (or related to) this subject, including [20, 26, 36, 42, 43, 70, 72, 73, 77, 78, 98, 141, 142, 153–155, 166, 172, 173].

(g) Theorem 4.8 above (from [74]), along with the results from [94] stated in Sect. 3 above and Sect. 5 below, and a part of the results from [87] discussed in Sect. 6 below have been extended by He and Lapidus in the research memoir [47] to the more general situation where the (now *generalized*) Minkowski content is no longer defined in terms of a strict power law, but in terms of a large class of "gauge functions" involving, for example, expressions of the form $x^D \log x$, $x^D \log\log x, \cdots$ or $x^D / \log\log\log x$, etc. This is of interest in the applications to fractal geometry, spectral geometry, harmonic analysis, probability theory, stochastic processes, and mathematical or theoretical physics.

We close this section by briefly commenting on the physical relevance of Weyl's asymptotic formula (see [65] for an interesting overview), Theorem 4.8 above (from [74]) and the Weyl–Berry conjecture which partially motivated it. The Weyl–Berry conjecture, along with its various modifications and extensions, has potential or actual physical and engineering applications to condensed matter physics, quantum mechanics, quantum chemistry, quantum chaos, acoustics, diffusions, and wave

propagation in fractal or random media (i.e., on or off fractals), geophysics, radar and cell phone technology (fractal antennas), as well as the making of computer microchips. (See, e.g., [6, 7, 20, 65, 70, 72–79, 89, 98, 102, 141, 142, 156, 159] and the relevant references therein.)

We will next focus on the one-dimensional case (i.e., $N = 1$), as in the rest of this paper (with the exception of the present section), and therefore on the connections between this subject and aspects of number theory, particularly the Riemann zeta function and the Riemann hypothesis, in Sects. 5 and 6, respectively.

5 Direct Spectral Problems for Fractal Strings and the Riemann Zeta Function in the Critical Interval

Having discussed in Sect. 4 the Weyl classical asymptotic formula (with error term) and the closely related Weyl–Berry conjecture, we may now specialize the situation to the one-dimensional case (that of fractal strings, corresponding to $N = 1$ in Sect. 4.1) and discuss the key result of [94] establishing the (one-dimensional) "MWB conjecture" for fractal strings; see Theorem 5.11 below. This brings to the fore direct connections between the spectra of fractal strings and the Riemann zeta function $\zeta = \zeta(s)$, in the case of the "critical interval" $0 < s < 1$. (See also Eq. (1) above for a related general formula.) This connection will be further explored in Sect. 6 (based on [87]), in relation with the critical strip $0 < \mathrm{Re}(s) < 1$, the Riemann hypothesis and the converse of Theorem 5.11; that is, an inverse (rather than a direct) spectral problem for fractal strings; see, in particular, Theorems 6.14 and 6.15 (from [87]) in Sect. 6 below.

Let Ω be an arbitrary (nontrivial) fractal string (i.e., a bounded open subset of $\mathbb{R}$), of Minkowski dimension $D \in (0, 1)$ and associated sequence of lengths $\mathscr{L} = (\ell_j)_{j=1}^{\infty}$, written in nonincreasing order (according to multiplicities) and such that $\ell_j \downarrow 0$ as $j \to \infty$; see Sect. 2. In the sequel, we will refer to such a fractal string as Ω or $\mathscr{L}$, interchangeably.

Recall from our discussion in Sects. 1 and 4.1 that a one-dimensional fractal drum (i.e., drum with fractal boundary) is nothing but a fractal string. Furthermore, recall from Sect. 2 that the (Minkowski) dimension of a fractal string always satisfies $0 \leq D \leq 1$. In the sequel, we exclude the extreme cases where $D = 0$ and $D = 1$ (the least or nonfractal case and the most fractal case, respectively) and therefore assume that D belongs to the *critical interval* $(0, 1) : 0 < D < 1$.

Letting $N = 1$ in the expression (17) for the Weyl term [i.e., the leading term in Weyl's asymptotic formula (16)] $W = W(x)$, we obtain (given our current normalization for the frequencies)

$$W(x) = |\Omega|_1 \, x, \text{ for } x > 0. \tag{21}$$

So that, according to Weyl's classic formula (16), we have

$$N_\nu(x) = W(x) + R(x), \tag{22}$$

where $W(x)$ is given by (21) and $R(x) = o(x)$ as $x \to +\infty$. Furthermore, provided $\mathscr{M}^* < \infty$ (i.e., $\mathscr{L}$ has finite upper Minkowski content), then according to part (i) of Theorem 4.8 above (from [74]), the error term $R(x)$ in (22) can be estimated as follows: $R(x) = O(x^D)$ as $x \to +\infty$. Moreover, if $\mathscr{L}$ is Minkowski nondegenerate (i.e., $0 < \mathscr{M}_* < \mathscr{M}^* < \infty$), we even have that $R(x)$ is truly of the order of x^D as $x \to +\infty$ (according to a result of [94]) briefly discussed in Sect. 2 above; and conversely (still by Lapidus and Pomerance [94]), if $R(x)$ is exactly of the order of x^D [resp., if $R(x) = O(x^D)$], then $\mathscr{L}$ is Minkowski nondegenerate (resp., $\mathscr{M}^* < \infty$). (See Theorem 3.3 and Eq. (14).)

The question is now to know when we really have an asymptotic second term, proportional to x^D, in Weyl's asymptotic formula (22). We will sometimes refer to such a term as a "*monotonic second term*." In [74], the following conjecture (called the MWB conjecture, for short) was stated.

Conjecture 5.9 (MWB Conjecture [74]). If $\mathscr{L}$ is a Minkowski measurable fractal string of dimension $D \in (0, 1)$, then

$$N_\nu(x) = W(x) - c_D\, \mathscr{M} x^D + o(x^D), \text{ as } x \to +\infty, \tag{23}$$

where $\mathscr{M}$ denotes the Minkowski content of $\mathscr{L}$ (i.e., of its boundary $\partial\Omega$) and c_D is a positive constant depending only on D.

Remark 5.10.

(a) In higher dimensions (i.e., when $N \geq 2$ and $\Omega \subset \mathbb{R}^N$ is a bounded open set), the counterpart of the MWB conjecture is not true, in general; see [34] and, especially, [95] for various counterexamples. Moreover, one does not know whether it is true for a simply connected domain in the plane (i.e., when $N = 2$). It is not the object of the present paper to discuss this issue further, although it is of significant interest and is also very intricate.

(b) When $N = 1$, the Minkowski measurability condition is necessary, in general, for the spectral counting function N_ν to admit a *monotonic* asymptotic second term, proportional to x^D, as $x \to +\infty$. For example, for the Cantor string $\mathscr{L} = \mathscr{L}_{CS}$ (defined by $\Omega_{CS} := [0, 1] \setminus C$, the complement in $[0, 1]$ of the classic ternary Cantor set C), N_ν admits an *oscillatory* (asymptotic) second term. More specifically, for the Cantor string, the asymptotic second term is of the form $-x^D\, G(\log_3 x)$, where G is a 1-periodic function which is bounded away from zero and infinity; see [94] and, especially, [101, Eq. (16.57) and §10.2.1]. This fact was first proved in [93, 94] (by a direct computation) and does not contradict the MWB conjecture since (as was also first proved in [93, 94]), the Cantor string is *not* Minkowski measurable (but is Minkowski nondegenerate). Moreover, these issues were investigated in great detail in [99–101], by using the theory of complex dimensions and associated explicit formulas developed in those monographs. (See, e.g., [101], Chap. 6, §8.4 and Chap. 10.)

It is shown, for example, in [101, Chap. 6, §8.4] (building in part on conjectures and results in [76]) that a self-similar string is Minkowski measurable

if and only it is nonlattice (i.e., the logarithms of its distinct scaling ratios are rationally independent). Moreover, it follows from *loc. cit.* that a lattice self-similar string (e.g., the Cantor string) is never Minkowski measurable (but is Minkowski nondegenerate) and that its spectral counting function $N_\nu(x)$ always has an asymptotic second term which is *oscillatory* and of the order of x^D (in fact, it is of the form $x^D\, G(\log x)$, where G is a nonconstant periodic function on $\mathbb{R}$ which is bounded away from zero and infinity).

More generally, we will see in Theorem 6.14 of Sect. 6 (based on [87]) that if $\zeta = \zeta(s)$ does not have any zeros on the vertical line $\{\mathrm{Re}(s) = D\}$, then the existence of an asymptotic second term proportional to x^D for the spectral counting function $N_\nu(x)$ implies that $\mathscr{L}$ is Minkowski measurable (i.e., the hypothesis of Conjecture 5.9 is also necessary).

The following theorem (due to Lapidus and Pomerance in [94], and announced in [93]) gives the precise asymptotic second term of the spectral (or frequency) counting function of a Minkowski measurable fractal string, thereby resolving in the affirmative the MWB conjecture for fractal strings and yielding the specific value of the positive constant c_D appearing in Eq. (23) above.

Theorem 5.11 (Resolution of the MWB Conjecture for Fractal Strings [94]). *The MWB conjecture for fractal strings* (*Conjecture 5.9 above*) *is true for every* $D \in (0,1)$. *More specifically, if* $\mathscr{L}$ *is a Minkowski measurable fractal string of dimension* $D \in (0,1)$, *then its frequency counting function* $N_\nu = N_\nu(x)$ *admits a monotonic second term, proportional to* x^D, *of the exact same form as in Eq.* (23). *Furthermore, the positive constant occurring in* (23) *is given by*

$$c_D = (1-D)\, 2^{-(1-D)}(-\zeta(D)). \tag{24}$$

Note that $c_D > 0$ because $\zeta = \zeta(s)$ is strictly negative in the critical interval $0 < s < 1$ (see, e.g., [174]): $-\zeta(D) > 0$ since $D \in (0,1)$. Furthermore, in Eq. (24), the constant c_D is proportional to the positive number $-\zeta(D)$, and hence, to the value at $s = D$ of $\zeta = \zeta(s)$ in the *critical interval* $0 < s < 1$.

We next briefly comment on the structure of the proof of Theorem 5.11 given in [94]. It relies on two different theorems. Namely, the geometric characterization of Minkowski measurability (Theorem 3.2 above, from [94]) along with the following theorem (also from [94]), which is of an analytic number theoretic nature and whose motivation will be explained after its statement. (In the sequel, given $y \in \mathbb{R}$, $[y]$ stands for the *integer part* of y and $\{y\} := y - [y] \in [0,1)$ denotes the *fractional part* of y.)

Theorem 5.12 (The Sound of Minkowski Measurable Fractal Strings [94]). *Let* $\mathscr{L}$ *be a Minkowski measurable fractal string of dimension* $D \in (0,1)$. *Equivalently, according to Theorem 3.2* (*see Eq.* (8) *in Sect. 3*), *let* $\mathscr{L} = (\ell_j)_{j=1}^\infty$ *be a nonincreasing sequence of positive numbers satisfying* (*for some positive constant* $L > 0$)

$$\ell_j \sim L j^{-1/D} \text{ as } j \to \infty. \tag{25}$$

Then

$$\sum_{j=1}^{\infty}[\ell_j x] = \left(\sum_{j=1}^{\infty}\ell_j\right)x + \zeta(D)L^D x^D + o(x^D) \text{ as } x \to +\infty, \tag{26}$$

where ζ denotes the classic Riemann zeta function. Equivalently, Eq. (26) *can be stated as follows*:

$$\sum_{j=1}^{\infty}\{\ell_j\, x\} \sim -\zeta(D)\, L^D x^D \text{ as } x \to +\infty.$$

We note that in order to state Theorem 5.12, one does not need to assume that $\mathscr{L}$ is Minkowski measurable and therefore, to rely on Theorem 3.2. Instead, one can simply assume that (25) [or, equivalently, (10)] holds for some $D \in (0,1)$. Theorem 3.2 is then used when deducing Theorem 5.11 from Theorem 5.12.

Given a fractal string $\Omega = \cup_{j=1}^{\infty} I_j$ (as in Sect. 2), its spectral counting function $N_\nu =: N_\nu(\Omega,\cdot)$ satisfies $N_\nu(\Omega,\cdot) = \sum_{j=1}^{\infty} N_\nu(I_j,\cdot)$, with $N_\nu(I_j,x) = [\ell_j x]$ for each $j \geq 1$ because I_j is an interval of length ℓ_j. It follows that $N_\nu(x) = \sum_{j=1}^{\infty}[\ell_j x]$, which is the left-hand side of (26). In light of the first expression obtained for the Minkowski content $\mathscr{M}$ in Eq. (15) (and the line following it), one then easily deduces Theorem 5.11 (and hence, the MWB conjecture for fractal strings, Conjecture 5.9) from Theorem 5.12. Note that $|\Omega|_1 = \sum_{j=1}^{\infty}\ell_j$, so that the leading term of $N_\nu(x) = \sum_{j=1}^{\infty}[\ell_j\, x]$ in (26) coincides with the Weyl term $W(x)$ [given by (21)], as it should. Furthermore, as was alluded just above, by using the expression of $\mathscr{M}$ given by the first equality in (15) and eliminating L^D in the asymptotic second term of (26), one establishes both the existence and the value of the positive constant c_D (explicitly given by Eq. (24) of Theorem 5.11).

Physically, the aforementioned relation, $N_\nu(\Omega,\cdot) = \sum_{j=1}^{\infty} N_\nu(I_j,\cdot)$, follows from the fact that the intervals I_j comprising the fractal string Ω (or, more poetically, the strings of the fractal harp) are vibrating independently of one another. Mathematically, it follows from the variational formulation of the underlying eigenvalue problem; see, e.g., [74] and the relevant references therein (including [120, 146]).

6 Inverse Spectral Problems for Fractal Strings and the Riemann Hypothesis

It is natural to wonder whether the converse of Theorem 5.11 (or essentially equivalently, of the MWB conjecture for fractal strings, Conjecture 5.9) is true. Roughly speaking, this means that if there are no oscillations of order D in the spectrum of $\mathscr{L}$, then there are no oscillations of order D in the geometry of $\mathscr{L}$. Rephrased: If $N_\nu(x)$ has an asymptotic second term, proportional to x^D, is it true that $\mathscr{L}$ is Minkowski measurable?

More precisely, given $D \in (0,1)$, the inverse spectral problem $(\mathrm{ISP})_D$ under consideration can be stated as follows:

$(\mathrm{ISP})_D$ *If $\mathscr{L}$ is a fractal string of dimension D such that*

$$N_\nu(x) = W(x) - \mathscr{C}\, x^D + o(x^D) \text{ as } x \to +\infty, \tag{27}$$

for some nonzero (real) constant $\mathscr{C}$, then is it true that $\mathscr{L}$ is Minkowski measurable (or, equivalently, in light of Theorem 3.2 and the comment following it, that $N_{\mathscr{L}}(x) \sim M\, x^D$ as $x \to +\infty$, for some positive constant M, where $N_{\mathscr{L}} = N_{\mathscr{L}}(x)$ denotes the geometric counting function of $\mathscr{L}$)?

The above problem, $(\mathrm{ISP})_D$, is called an *inverse spectral problem* since given some spectral information about the fractal string $\mathscr{L}$ (namely, the existence for $N_\nu(x)$ of a monotonic asymptotic second term, proportional to x^D), one asks whether one can recover some geometric information about $\mathscr{L}$ (namely, that $\mathscr{L}$ is Minkowski measurable). Similarly, the MWB conjecture from [74] (Conjecture 5.9 above) and its resolution given in Theorem 5.11 above (from [94]) fall naturally within the class of *direct spectral problems*.

Remark 6.13.

(a) In the statement of $(\mathrm{ISP})_D$, it is not necessary to assume that $\mathscr{L}$ is of Minkowski dimension D since a result from [94] (recalled in Theorem 3.3 above and the comment following it) shows that (27) implies that $\mathscr{L}$ is Minkowski nondegenerate (i.e., $0 < \mathscr{M}_* \leq \mathscr{M}^* < \infty$) and hence, has Minkowski dimension D.

(b) Originally, in [86, 87], the error term $o(x^D)$ in the counterpart of Eq. (27), was assumed to be slightly smaller, namely, $O(x^D/\log^{1+\delta} x)$ as $x \to +\infty$, for some $\delta > 0$. However, an unpublished work of Hilberdink (2000, Written personal communication to the author, unpublished), using an improved version of the Wiener–Ikehara Tauberian theorem (see, e.g., [137]) used in [86, 87], shows that the more general estimate $o(x^D)$ suffices. (Actually, in [86, 87], the better error estimate is only used for the sufficiency part of the proof of Theorem 6.14, that is, the part requiring the use of a Tauberian theorem.)

(c) For Dirichlet boundary conditions (which are assumed here in $(\mathrm{ISP})_D$), one must necessarily have $\mathscr{C} > 0$ in Eq. (27) because then, we have that $N_\nu(x) \leq W(x)$ for all $x > 0$, which implies that $\mathscr{C} \geq 0$. Note that by hypothesis, $\mathscr{C} \neq 0$.

In [87] (announced in [86]), Lapidus and Maier have shown that this family of inverse spectral problems $(\mathrm{ISP})_D$, for D ranging through the critical interval $(0,1)$, is intimately related to the presence of zeros of $\zeta = \zeta(s)$ in the *critical strip* $0 < \mathrm{Re}(s) < 1$ (i.e., to the *critical zeros* of ζ), and thereby, to the Riemann hypothesis. More specifically, we have the following two theorems (the second one really being a corollary of the first one).

Theorem 6.14 (Characterization of the Nonvanishing of ζ Along Vertical Lines in the Critical Strip [87]). *Fix $D \in (0,1)$. Then, the inverse spectral problem*

$(ISP)_D$ is true (that is, has an affirmative answer for every fractal string of dimension D satisfying the assumptions of $(ISP)_D$) if and only if the Riemann zeta function $\zeta = \zeta(s)$ does not have any zeros along the vertical line $\{Re(s) = D\}$; i.e., if and only the "D-partial Riemann hypothesis" is true.

Theorem 6.15 (Spectral Reformulation of the Riemann Hypothesis [87]). *The inverse spectral problem $(ISP)_D$ is true for every value of D in $(0, 1)$ other than in the midfractal case when $D = 1/2$ [or equivalently, for every $D \in (0, 1/2)$] if and only if the Riemann hypothesis is true.*

Theorem 6.15 follows at once from Theorem 6.14 since the Riemann hypothesis states that $\zeta(s) = 0$ with $0 < \mathrm{Re}(s) < 1$ implies that $\mathrm{Re}(s) = 1/2$ (i.e., that s belongs to the critical line $\{\mathrm{Re}(s) = 1/2\}$). The fact that in Theorem 6.15, D can be equivalently assumed to be in $(0, 1/2)$, $(1/2, 1)$ or in $(0, 1)\backslash\{1/2\}$, follows from the functional equation satisfied by $\zeta = \zeta(s)$, according to which

$$\xi(s) = \xi(1 - s), \text{ for all } s \in \mathbb{C}, \tag{28}$$

and hence,

$$\zeta(s) = 0 \Leftrightarrow \zeta(1 - s) = 0, \text{ for } 0 < \mathrm{Re}(s) < 1. \tag{29}$$

Here, $\xi = \xi(s)$ denotes the *completed Riemann zeta function* (or the *global zeta function* of $\mathbb{Q}$, the field of rational numbers), defined by $\xi(s) := \pi^{-s/2}\Gamma(s/2)\zeta(s)$. For the standard properties of ζ and of ξ, we refer, e.g., to [29, 58, 66, 133, 174].

The next result follows immediately from Theorem 6.14 and the well-known fact according to which ζ has a zero (and even infinitely many zeros, according to Hardy's theorem [29, 174], even though it is not needed here) along the critical line $\{\mathrm{Re}(s) = 1/2\}$.

Corollary 6.16 ([87]). *The inverse spectral problem $(ISP)_D$ is not true in the midfractal case when $D = 1/2$.*

The following interpretation of Theorem 6.15 and Corollary 6.16 has first been proposed in [75, 76] and has since been pursued in a different, but closely related context, in [49–53] and in [80], as will be briefly discussed towards the end of Sect. 7.4 below.

Theorem 6.17 (The Riemann Hypothesis as a Mathematical Phase Transition [75, 76]). *The Riemann hypothesis is true if and only if $(ISP)_D$, the inverse spectral problem for fractal strings, fails to be true (i.e., fails to have an affirmative answer) only in the midfractal case when $D = 1/2$.*

In [75, 76], Lapidus also wondered (in an open problem) whether the mathematical phase transition conditionally (i.e., under RH, and in fact, if and only if RH is true) occurring at $D = 1/2$ could be understood (in a suitable model) as a true physical phase transition, of the type occurring in the theory of critical phenomena in statistical physics or quantum field theory. He also asked whether the (then)

intuitive notion of fractal "complex dimension," underlying the proof of the key part of Theorem 6.14 (as will be explained next), could not be understood as a "complexified space-time dimension," as in Wilson's theory of phase transitions and critical phenomena [182].

We now briefly comment on the proof of Theorem 6.14, the central result of [87], from which Theorem 6.15 and Corollary 6.16 (as well as Theorem 6.17) follow, given the known properties of the Riemann zeta function. First, note that Theorem 6.14 consists of two different theorems (and was stated as such in [86, 87]). The part of Theorem 6.14 corresponding to a sufficient condition (for $(\text{ISP})_D$ to be true) is established by using the Wiener–Ikehara Tauberian theorem (see, e.g., [137] or [87] for the statement of this theorem). That is, assuming that [for a given $D \in (0,1)$], $\zeta(s) \neq 0$ for all $s \in \mathbb{C}$ on the vertical line $\{\text{Re}(s) = D\}$, one uses the aforementioned Tauberian theorem in order to show that the inverse problem $(\text{ISP})_D$ has an affirmative answer (for all fractal strings of dimension D). On the other hand, in order to establish the converse, namely, the fact that the condition that $\zeta(s) \neq 0$ for all $s \in \mathbb{C}$ on the vertical line $\{\text{Re}(s) = D\}$ is necessary for the inverse spectral problem $(\text{ISP})_D$ to be true (i.e., to have an affirmative answer for all fractal strings of dimension D), one uses in [86, 87] in a key (but rigorous) manner the intuition (at the time) of complex dimension and its intimate connections with asymptotic oscillatory behavior (both in the geometry and the spectrum), along with Theorem 3.2 (from [94]).[5]

More specifically, we reason by contraposition. Fix $D \in (0,1)$ and assume that $\zeta(\omega) = 0$, for some $\omega \in \mathbb{C}$ such that $Re(\omega) = D$. Then, due to the basic symmetry of ζ (namely, $\zeta(\overline{s}) = \overline{\zeta(s)}$, where the bar indicates that we are taking the complex conjugate of the given complex number), we also have $\zeta(\overline{\omega}) = 0$. Let us write $\omega = D + i\tau$, with $\tau \in \mathbb{R}$; so that $\overline{\omega} = D - i\tau$. Without loss of generality, we may assume that $\tau > 0$. (Clearly, $\tau \neq 0$ since $\zeta(D) < 0$ because $0 < D < 1$.)

Next, for $x > 0$, let

$$U(x) := x^D + \beta(x^{\omega} + x^{\overline{\omega}}) = x^D(1 + 2\beta\cos(\tau \log x)), \tag{30}$$

for some coefficient $\beta > 0$ sufficiently small, and let $V(x) := [U(x)]$, the integer part of $U(x)$. It is easy to check that for all β small enough, $U(x) > 0$ and U is (strictly) increasing on $(0, +\infty)$; furthermore, the range of U is all of $(0, +\infty)$. Hence, for such values of β, given any integer $j \geq 1$, we can uniquely define $\ell_j > 0$ such that $U(\ell_j) = j$ (and thus, $V(\ell_j) = j$). In this manner, we define a fractal string $\mathscr{L} = (\ell_j)_{j=1}^{\infty}$ with geometric counting function $N_{\mathscr{L}}$ coinciding with $V : N_{\mathscr{L}} = V$. In light of Eq. (30), $N_{\mathscr{L}} = V = [U]$ has sinusoidal (and hence, nontrivial periodic) oscillations, caused by the "complex dimensions" ω and $\overline{\omega}$. Therefore, $N_{\mathscr{L}}(x)$ cannot be asymptotic to x^D; equivalently, we do not have

[5] It can now also be systematically understood in terms of the generalized explicit formulas of [101, Chap. 5]; see [101, Chap. 9]. The resulting theorems and assumptions, however, are somewhat different.

$N_{\mathscr{L}}(x) \sim M\,x^D$ as $x \to +\infty$, for some $M > 0$. (Note that here, according to (30), we would have to have $M := 1$.) It then follows from Theorem 3.2 (the characterization of Minkowski measurability, from [94]) and the comment following it that $\mathscr{L}$ is *not* Minkowski measurable. Moreover, and using the fact that $\zeta(\omega) = \zeta(\overline{\omega}) = 0$, via a direct computation[6] it is shown in [87] that Eq. (27) holds for the fractal string $\mathscr{L}$, for some (explicitly known) positive constant $\mathscr{C}$. Consequently, the hypothesis (27) of $(\mathrm{ISP})_D$ is satisfied but its conclusion (namely, the Minkowski measurability of $\mathscr{L}$) is not; i.e., the inverse spectral problem $(\mathrm{ISP})_D$ cannot be true for this value of $D \in (0, 1)$ because it fails to hold for this particular fractal string $\mathscr{L}$. This proves that if $(\mathrm{ISP})_D$ is true for some $D \in (0, 1)$, we must have $\zeta(s) \neq 0$ for all $s \in \mathbb{C}$ such that $\mathrm{Re}(s) = D$, as desired.

We point out that in the above construction, the geometric oscillations caused by the nonreal complex dimensions ω and $\overline{\omega}$ of $\mathscr{L}$ remain [as is obvious from (30)], but due to the fact that by construction, $\zeta(\omega) = \zeta(\overline{\omega}) = 0$, the spectral oscillations [in the asymptotic second term of $N_\nu(x)$] disappear. Therefore, we see the subtle interplay between complex dimensions, geometric and spectral oscillations, as well as the critical zeros of ζ. (Clearly, both ω and $\overline{\omega}$ belong to the open critical strip $0 < \mathrm{Re}(s) < 1$ since we have that $\mathrm{Re}(\omega) = \mathrm{Re}(\overline{\omega}) = D$ and $D \in (0, 1)$.)

Finally, we note that by using the theory of complex dimensions developed in [99–101], it can be shown that the above fractal $\mathscr{L}$ has exactly three complex dimensions (each of which has a real part equal to D and multiplicity one). Namely, the set $\mathscr{D}_{\mathscr{L}}$ of complex dimensions of $\mathscr{L}$ is given by $\mathscr{D}_{\mathscr{L}} = \{D, \omega, \overline{\omega}\}$. Moreover, observe that $x^{\omega} = x^D\,x^{i\tau}$ and $x^{\omega} = x^D\,x^{-i\tau}$; so that the real part (resp., the imaginary part) of the complex dimension ω (or $\overline{\omega}$) determines the *amplitude* (resp., the *frequency*) of the corresponding geometric or spectral oscillations (viewed multiplicatively). This statement is now fully corroborated (for any fractal string and its complex dimensions) by the rigorous theory of complex dimensions of fractal strings and the associated (generalized) explicit formulas developed in [99–101].

7 Epilogue: Later Developments and Research Directions

In this epilogue, by necessity of concision, we very briefly discuss further developments closely connected to (or partly motivated by) the results discussed in the main body of this paper as well as by related results and conjectures in [74–76]. These topics include a geometric interpretation of the critical strip for the Riemann zeta function $\zeta = \zeta(s)$ (see Sect. 7.1), the theory of complex dimensions of fractal strings (see Sect. 7.2), fractal zeta functions and a higher-dimensional theory of complex dimensions (valid for arbitrary bounded subsets of Euclidean spaces and relative fractal drums (RFDs), see Sect. 7.3), quantized number theory and the

[6] This is now proved in [99–101] by using the (generalized) explicit formulas from [99–101] and Eq. (1).

spectral operator, along with a functional analytic reformulation of the results of [87] discussed in Sect. 6, as well as a different framework (developed in [49–53]) and a new asymmetric reformulation of the Riemann hypothesis recently obtained by Lapidus in [80] (see Sect. 7.4).

7.1 *Fractal Strings, $\zeta = \zeta(s)$, and a Geometric Interpretation of the Critical Strip*

The one-dimensional situation (i.e., the case of fractal strings) is ideally suited to the Riemann zeta function $\zeta = \zeta(s)$ in the (closed) critical strip $0 \leq \mathrm{Re}(s) \leq 1$, as we have seen in Sect. 5 and, especially, Sect. 6 above. This is due in part to the product formula (1), $\zeta_\nu(s) = \zeta(s) \cdot \zeta_{\mathscr{L}}(s)$, connecting the geometric zeta function $\zeta_{\mathscr{L}}$ of a fractal string $\mathscr{L}$ to its spectral zeta function ζ_ν,[7] combined with the fact that fractal strings always have a (Minkowski) dimension between 0 and 1 : $0 \leq D \leq 1$.

These facts, along with the results of [93, 94] and some of the results and conjectures of [74–76], have led to the following geometric interpretation of the (closed) critical strip $0 \leq \mathrm{Re}(s) \leq 1$, using the terminology introduced in [74] (in any dimension): The least (resp., most) fractal case when $D = 0$ (resp., $D = 1$), corresponds to the left- (resp., right-) hand side of the critical strip, that is, to the vertical line $\{\mathrm{Re}(s) = 0\}$ (resp., $\{\mathrm{Re}(s) = 1\}$. Furthermore, the midfractal case when $D = 1/2$ corresponds to the critical line, namely, the vertical line $\{\mathrm{Re}(s) = 1/2\}$ where (according to the Riemann hypothesis) all of the nontrivial (or critical) zeros of $\zeta = \zeta(s)$ are supposed to be located. This geometric picture of the critical strip has later been corroborated by the work in [86, 87] described in Sect. 6 above. Its complete and rigorous justification has then been provided by the theory of complex dimensions for fractal strings (that is, in the one-dimensional situation) developed in [99–101]. (See, in particular, [101, Chaps. 9 and 11].) We will next briefly describe (in Sect. 7.2) a few aspects of the latter theory.

7.2 *Complex Dimensions of Fractal Strings and Oscillatory Phenomena*

The theory of complex dimensions of fractal strings, developed by Lapidus and van Frankenhuijsen in the research monographs [99–101] (and several corresponding articles), aimed originally at obtaining a much more accurate understanding of the oscillatory phenomena which are intrinsic to fractals (in their geometries and their

[7] We note that the product formula (1) for the spectral zeta functions of fractal strings has since been extended in various ways in the setting of Laplacians on certain self-similar fractals; see [26, 72, 73, 172, 173].

spectra as well as in the underlying dynamics). This is accomplished via explicit formulas (generalizing Riemann's original 1858 explicit formula and its extensions to various aspects of number theory and arithmetic geometry, see [24, 29, 58, 131–133, 157, 174] along with [101, §5.1.2 and §5.6]) expressed in terms of the underlying complex dimensions; see [101, Chap. 5]. The latter complex dimensions are defined as the poles of a suitable zeta function, typically a geometric, spectral, dynamical, or arithmetic zeta function. (See [101].)

In the case of fractal strings, which is the main focus of the theory developed in [99–101], these explicit formulas can be applied, for example, to the geometric counting function $N_{\mathscr{L}}(x)$, the spectral counting function $N_{\nu}(x)$, the volume of tubular neighborhoods $V(\varepsilon)$ (giving rise to a "fractal tube formula"), or to the counting function of the number of primitive geodesics of an underlying dynamical system. (See [101, Chaps. 6–11].) In the special case of self-similar strings, the resulting fractal tube formulas give very precise information, in part due to the knowledge of the periodic (or, in general, the quasiperiodic) structure of the complex dimensions (see [101], Chaps. 2–3 and Sect. 8.4).

For a fractal string $\mathscr{L}$, under mild growth assumptions on the associated geometric zeta function $\zeta_{\mathscr{L}}$ (see [101, §5.1]) and assuming that all the complex dimensions (i.e., the poles of $\zeta_{\mathscr{L}}$) are simple, the *fractal tube formula* for $V(\varepsilon) = V_{\mathscr{L}}(\varepsilon)$ takes the following form (see [101, Chap. 8]):

$$V(\varepsilon) = \sum_{\omega \in \mathscr{D}} \alpha_\omega \frac{(2\varepsilon)^{1-\omega}}{\omega(1-\omega)} + R(\varepsilon), \tag{31}$$

where $\mathscr{D} = \mathscr{D}_{\mathscr{L}}$ denotes the set of (visible) complex dimensions of $\mathscr{L}$ and $\alpha_\omega := res(\zeta_{\mathscr{L}}, \omega)$ for each $\omega \in \mathscr{D}$. Furthermore, the error term $R(\varepsilon)$ is precisely estimated in [101]. In the important special case of self-similar strings (which includes the Cantor string discussed just below), one can take $R(\varepsilon) \equiv 0$ and therefore obtain an *exact* fractal tube formula (which holds pointwise); see [101, §8.4].

Remark 7.18.

(a) The explicit formula is valid for multiple poles as well but must then take a different form, in general. Namely, we then have (see [101, Theorems 8.1 and 8.7])

$$V(\varepsilon) = \sum_{\omega \in \mathscr{D}} res\left(\frac{\zeta_{\mathscr{L}}(s)\,(2\varepsilon)^{1-s}}{s(1-s)}, \omega\right) + R(\varepsilon). \tag{32}$$

(b) The fractal tube formulas (31) and (32) hold pointwise or distributionally, depending on the growth assumptions made about $\zeta_{\mathscr{L}}$; see [101], Sect. 8.1, especially, Theorems 8.1 and 8.7 and their corollaries. Also, in the so-called strongly languid case (in the sense of [101, Definition 5.3]), we can let $R(\varepsilon) \equiv 0$ and therefore obtain an *exact* formula. This is the case, for example, for all self-similar strings; see [101, §8.4].

(c) For the first higher-dimensional analog of (31) and (32), see [99–101] for specific examples of fractal sprays (in the sense of [95]), and for a fairly general class of fractal sprays and self-similar tilings (or, less generally, sets), see [91, 92, 104, 105]; see also [134, 135] and, for a direct approach in the case of the Koch snowflake curve [90]. Within the general higher-dimensional theory of complex dimensions developed in [108–113, 115, 116], the precise counterpart of (31) and (32) is provided in [108, 113] and [116, Chap. 5]; see Sect. 7.2 for a brief discussion.

Let us now specialize the above discussion to the Cantor string $\mathscr{L} = CS$, viewed geometrically as the bounded open set $\Omega \subset \mathbb{R}$, defined as the complement of the classic ternary Cantor set in $[0, 1]$; hence, $\partial\Omega$ is the Cantor set. Then $\mathscr{L} = CS = (\ell_j)_{j=1}^{\infty}$, with $\ell_1 = 1/3, \ell_2 = \ell_3 = 1/9, \ell_4 = \ell_5 = \ell_6 = \ell_7 = 1/27, \cdots$. Alternatively, the lengths of the Cantor string (or harp) are the numbers 3^{-n-1} repeated with multiplicity 2^n, for $n = 0, 1, 2, \cdots$. Then

$$\zeta_{CS}(s) = \frac{3^{-s}}{1 - 2 \cdot 3^{-s}}, \text{ for all } s \in \mathbb{C}; \tag{33}$$

so that the set $\mathscr{D}_{CS}$ of complex dimensions (here, the complex solutions of the equation $1 - 2 \cdot 3^{-s} = 0$) is given by

$$\mathscr{D}_{CS} = \{D + in\mathbf{p} : n \in \mathbb{Z}\}, \tag{34}$$

where $D := \log_3 2 = \log 2 / \log 3$ (the Minkowski dimension of CS or, equivalently, of the ternary Cantor set) and $\mathbf{p} := 2\pi / \log 3$ (the *oscillatory period* of CS). Then (31) (with $R(x) \equiv 0$) becomes

$$\begin{aligned} V_{CS}(\varepsilon) &= \frac{1}{2 \log 3} \sum_{n \in \mathbb{Z}} \frac{(2\varepsilon)^{1-D-in\mathbf{p}}}{(D + in\mathbf{p})(1 - D - in\mathbf{p})} - 2\varepsilon \\ &= \varepsilon^{1-D} \, G(\log \varepsilon^{-1}) - 2\varepsilon, \end{aligned} \tag{35}$$

where G is a periodic function which is bounded away from zero and infinity. One then recovers the result from [93, 94] according to which CS (and hence also, the Cantor set) is Minkowski nondegenerate but is not Minkowski measurable. (Actually, the values of $\mathscr{M}_*$ and $\mathscr{M}^*$ are computed explicitly in [94] as well as, in a more general context, in [101, Chap. 10].)

Moreover, the geometric counting function of CS is given by

$$N_{CS}(x) = \frac{1}{2 \log 3} \sum_{n \in \mathbb{Z}} \frac{x^{D+in\mathbf{p}}}{D + in\mathbf{p}} - 1 \tag{36}$$

while the corresponding frequency counting function is given by

$$N_{v,CS}(x) = x + \frac{1}{2\log 3}\sum_{n\in\mathbb{Z}} \zeta(D + in\mathbf{p})\frac{x^{D+in\mathbf{p}}}{D + in\mathbf{p}} + O(1) \text{ as } x \to +\infty. \tag{37}$$

(See [101], Sect. 1.1.1, Sect. 1.2.2 and Chap. 6.)

In (31) and (35)–(37), we see the intuitive (and actual) meaning of the complex dimensions. Namely, the real (resp., imaginary) parts correspond to the *amplitudes* (resp., *frequencies*) of the oscillations [in the spaces of scales, for (31) and (35)–(36), and in frequency space, for (37)].

Remark 7.19 (Fractality and Complex Dimensions). The notion of fractality is notoriously difficult to define. Mandelbrot [121] has proposed to define a fractal as a geometric object whose Hausdorff dimension is strictly greater than its topological dimension. There is an obvious problem with this definition (Mandelbrot was aware of it, as is stated in his book, [121, p. 82]). Namely, the Cantor curve (or "devil's staircase") is not fractal in this sense (since its Hausdorff, Minkowski, and topological dimensions are all equal to 1); however, as Mandelbrot states in [121], everyone would agree that the Cantor curve should be called "fractal". This issue has long preoccupied the present author.

There is, however, a satisfactory way to resolve this apparent paradox as well as many other related and unrelated issues. In the theory of complex dimensions developed in [99–101], an object is called "fractal" if it has at least one *nonreal* complex dimension (with positive real part). (See [101, §12.1].) Accordingly, the Cantor curve (CC) is, indeed, fractal because its set of complex dimensions is given by $\mathscr{D}_{CC} = \mathscr{D}_{CS} \cup \{1\}$, where $\mathscr{D}_{CS}$ is the set of complex dimensions of the Cantor set (or, equivalently, the Cantor string) given by (34). (Hence, $\mathscr{D}_{CC}$ has infinitely many nonreal complex dimensions with positive real part.) Furthermore, (nontrivial) self-similar strings (and, more generally, self-similar geometries) are fractal.[8] In order to accommodate random (and, in particular, stochastically self-similar fractals), as in [44] and the relevant references therein, the above definition of fractality has been extended to allow for a "fractal" A to be such that its associated zeta function ζ_A has a natural boundary along a suitable curve of the complex plane, called a "screen" in [101, §5.1] (and hence, such that ζ_A cannot be meromorphically extended beyond that curve). In [108–113, 115, 116], such a compact subset of $\mathbb{R}^N$ is called a "hyperfractal." In [110, 112, 116] are even constructed "maximal hyperfractals"; that is, compact subsets A of $\mathbb{R}^N$ (where $N \geq 1$ is arbitrary) such that ζ_A has a nonremovable singularity at every point of the "critical line" $\{\text{Re}(s) = D\}$, where D is the Minkowski dimension of A.

Remark 7.20 (Noncommutative Riemann Flow of Zeros and the Riemann Hypothesis [79]). In the long term, the theory of complex dimensions aims at unifying many aspects of fractal and arithmetic geometries. Many concrete and more abstract

[8] An example of a trivial self-similar set is an interval of $\mathbb{R}$ or a cube in $\mathbb{R}^N (N \geq 2)$. In such cases, all of the complex dimensions are easily seen to be real; see [101, 112, 113] and [116].

examples are provided in [101]. This direction is pursued in many different directions in the author's book, *In Search of the Riemann Zeros* [79], where is introduced the notion of fractal membrane (i.e., quantized fractal string) and the associated moduli spaces of fractal membranes (as well as of fractal strings). In [79], a still conjectural (noncommutative) flow on the moduli space of fractal membranes and correspondingly, a flow of zeta functions (or partition functions) and a flow of zeros are used in an essential manner in order to provide a possible new interpretation of (and approach to) the Riemann hypothesis.

Accordingly, the flow of fractal membranes can be viewed as some kind of noncommutative Ricci flow which is transforming (as "time" tends to infinity or physically, as the absolute "temperature" tends to zero) generalized quasicrystals into (self-dual) pure crystals. Correspondingly, the zeta functions are becoming increasingly symmetric (that is, satisfy a true functional equation, in the limit), while the zeros (of these zeta functions) are pushed under this flow (viewed as a flow on the Riemann sphere) onto the Equator, which naturally represents the critical line in this context. Conjecturally, this would also provide a proof of the generalized Riemann hypothesis (GRH), valid for all of the number theoretic zeta functions for which the analog of the Riemann hypothesis is expected to be true. (See, e.g., [131, 132, 157] and [101, Appendix A] or [79, Appendices B, C, and E].) Very concisely, this is the interpretation of (and approach to) the Riemann hypothesis proposed in [79]. Needless to say, it poses formidable mathematical and physical challenges but may nevertheless stimulate further investigations of the aforementioned (noncommutative) Riemann (or "modular") flow, whether it be viewed as a (noncommutative) Ricci-type flow or, physically, as a renormalization flow (not unlike the one which presumably describes the time evolution of our universe).

7.3 *Fractal Zeta Functions of Arbitrary Compact Sets and Higher-Dimensional Theory of Complex Dimensions*

In a forthcoming book by Lapidus, Radunović and Žubrinić, entitled *Fractal Zeta Functions and Fractal Drums* [116] (see also the series of research and survey papers [108–113, 115]), is developed a higher-dimensional theory of complex dimensions, valid for any bounded subset of N-dimensional Euclidean space $\mathbb{R}^N$, with $N \geq 1$ arbitrary. Distance and tube zeta functions are defined in that general setting. Let D be the (upper) Minkowski dimension of a given (nonempty) bounded subset A of $\mathbb{R}^N$. Then it is shown in [110, 116] that the *distance zeta function* of A, denoted by $\zeta_A = \zeta_A(s)$ and defined (for all $s \in \mathbb{C}$ with $\mathrm{Re}(s)$ sufficiently large and for a given $\delta > 0$) by

$$\zeta_A(s) := \int_{A_\delta} d(x, A)^{s-N} dx \tag{38}$$

is holomorphic for $\mathrm{Re}(s) > D$.

Remark 7.21.

(a) This new type of fractal zeta function, ζ_A, was introduced by the author in 2009 in order to be able to extend to any higher dimension N the theory of complex dimensions of fractal strings developed in [99, 100] (and now also in [101]). The case of fractal strings corresponds to $N = 1$; in that case, the precise relationship between $\zeta_{\mathscr{L}}$ and ζ_A, with $A := \partial\mathscr{L}$, is explained in [110, 116].

(b) For the simplicity of the discussion, we assume here that $|A| = 0$ (i.e., A is a Lebesgue null set), which is the case of most fractals of interest. We refer to [116] for a discussion of the general case.

A first key result of the theory is that the abscissa of convergence of this Dirichlet-type integral (viewed as a Lebesgue integral) or, equivalently, the *abscissa of* (*absolute*) *convergence* of ζ_A, coincides with D:

$$\sigma = D, \tag{39}$$

the (upper) Minkowski (or box) dimension of A. Therefore, $\{\mathrm{Re}(s) > D\}$ is the largest open right half-plane (of the form $\{\mathrm{Re}(s) > \alpha\}$, for some $\alpha \in \mathbb{R} \cup \{\pm\infty\}$), on which the Lebesgue integral appearing on the right-hand side of (38) is convergent (i.e., absolutely convergent). This result is the higher-dimensional counterpart of Theorem 2.1 above (first noted in [75, 76] in the case of fractal strings, i.e., when $N = 1$). The proof of Eq. (39) makes use of an interesting integral estimate obtained in [46] in order to study the singularities of solutions of certain linear partial differential equations. (See also [183].)

Moreover, under mild assumptions, namely, if $D < N$ (recall that we always have $0 \leq D \leq N$), $\mathscr{M}_*^D > 0$ and the Minkowski dimension of A exists (i.e., the lower and upper Minkowski dimensions of A coincide), then $\{\mathrm{Re}(s) > D\}$ is also the maximal open right-half plane to which $\zeta_A = \zeta_A(s)$ can be holomorphically continued; i.e., D coincides with $D_{hol}(\zeta_A)$, the *abscissa of holomorphic continuation* of ζ_A:

$$D = \sigma = D_{hol}(\zeta_A), \tag{40}$$

where, as above, $D = D(A)$ and $\sigma = D_{abs}(\zeta_A)$.

An entirely analogous theorem holds for the *tube zeta function* of A, denoted by $\zeta_A = \tilde{\zeta}_A(s)$ and defined by

$$\tilde{\zeta}_A(s) = \int_0^\delta t^{s-N} |A_t|_N \frac{dt}{t}, \tag{41}$$

for all $s \in \mathbb{C}$ such that $\mathrm{Re}(s) > D$.

In fact, it can be shown that the two functions, ζ_A and $(N - s)\ \tilde{\zeta}_A$, differ by an entire function, from which it follows that (for $D < N$), the two fractal zeta functions ζ_A and $\tilde{\zeta}_A$ have exactly the same qualitative properties. In particular, given a domain $U \subseteq \mathbb{C}$ containing $\{\mathrm{Re}(s) > D\}$, the common half-plane of (absolute) convergence of ζ_A and $\tilde{\zeta}_A$, ζ_A can be meromorphically extended to U if and only if

it is the case for $\tilde{\zeta}_A$. Furthermore, ζ_A and $\tilde{\zeta}_A$ have exactly the same poles, in U (with the same multiplicities). These poles are called the (visible) *complex dimensions* of A. Moreover, the residues of ζ_A and $\tilde{\zeta}_A$ at a simple pole (or, more generally, the principal parts at a multiple pole) are related in a very simple way. For example, if D is simple, then we have

$$res(\zeta_A, D) = (N - D)\, res(\tilde{\zeta}_A, D).$$

Finally, if ζ_A (or, equivalently, $\tilde{\zeta}_A$) can be meromorphically continued to an open and connected neighborhood of D, then D is a simple pole of ζ_A and $\tilde{\zeta}_A$ (i.e., it is a complex dimension of A)[9] and the residue of $\tilde{\zeta}_A$ at $s = D$ is squeezed between the lower and upper Minkowski contents of A:

$$\mathscr{M}_* \leq res(\tilde{\zeta}_A; D) \leq \mathscr{M}^*. \tag{43}$$

If, in addition, A is Minkowski measurable (with Minkowski content denoted by $\mathscr{M}$, as before), then

$$res(\tilde{\zeta}_A; D) = \mathscr{M}. \tag{44}$$

Both formulas (43) and (44) are valid even if $D = N$, although the justification of this statement requires a specific argument; see [116].

We should note that, in light of the aforementioned functional equation connecting ζ_A and $\tilde{\zeta}_A$, the choice of δ is unimportant in the definition of ζ_A and $\tilde{\zeta}_A$ (in (38) and (41), respectively). Furthermore, the residues of ζ_A and $\tilde{\zeta}_A$ at (simple) complex dimensions are independent of δ.

In the special case when $N = 1$ (that is, for fractal strings), it is shown (still in [110, 115, 116]) that the distance zeta function ζ_A and the geometric zeta function $\zeta_{\mathscr{L}}$ are closely related (here, $A := \partial\Omega$ and $\mathscr{L}$ is the fractal string associated with the bounded open set $\Omega \subset \mathbb{R}$). In fact, given any subdomain U of $\{\mathrm{Re}(s) > 0\}$ (or, more generally, of $\mathbb{C}\backslash\{0\}$), $\zeta_{\mathscr{L}}$ has a meromorphic continuation to U if and only if $\zeta_A = \zeta_{\partial\Omega}$ does. Consequently, $\zeta_{\mathscr{L}}$ and ζ_A have the same poles (with the same multiplicities) in U; i.e., they have the same visible complex dimensions. This is true independently of the geometric realization Ω of the fractal $\mathscr{L} = (\ell_j)_{j=1}^{\infty}$ as a bounded open subset of $\mathbb{R}$. In particular, we recover Theorem 2.1.

A variety of results are provided in [108–113, 115, 116], guaranteeing the existence of a suitable meromorphic extension of ζ_A (and of $\tilde{\zeta}_A$) beyond the half-plane of convergence $\{\mathrm{Re}(s) > D\}$. Moreover, many examples of computation

[9] Since ζ_A and $\tilde{\zeta}_A$ are holomorphic in the open half-plane $\{\mathrm{Re}(s) > D\}$, it then follows that

$$D = \max\{\mathrm{Re}(s) : s \in \mathscr{D}\}, \tag{42}$$

where $\mathscr{D}$ is the set of (visible) complex dimensions of A in any domain U containing this neighborhood of $\{D\}$.

of the complex dimensions of a variety of fractals (when $N = 1, 2$ or $N \geq 3$) are provided throughout [108–113, 115, 116]. These fractals include (one- or two-parameter) families of generalized Cantor sets, the Sierpinski gasket and carpet, a relative fractal drum counterpart of the Menger sponge (a three-dimensional analog of the Sierpinski carpet, see, e.g., [121]), families of spirals, fractal curves such as the Cantor curve (or devil's staircase), which is examined from several points of view, as well as geometric chirps and curves with cusps. The theory is extended to a new class of objects, called RFDs, which allow a much greater flexibility in many situations and generalize both the usual fractal drums (i.e., drums with fractal boundary, in the sense, for example, of [74–76], as discussed in Sect. 4 above) and the class of bounded subsets of Euclidean space $\mathbb{R}^N$. It is noteworthy that in the case of RFDs, the relative Minkowski dimension can be negative; it can even take the value $-\infty$, as is shown in [109, 112, 116], where geometric explanations are provided for this "dimension drop" phenomenon.

In short, a *RFD* is a pair (A, Ω), with A an arbitrary (possibly unbounded) subset of $\mathbb{R}^N$ and Ω an open subset of $\mathbb{R}^N$ with finite volume (i.e., $|\Omega|_N < \infty$) and such that $\Omega \subseteq A_\delta$, for some $\delta > 0$. The special case of an ordinary fractal drum (or "drum with fractal boundary," in the sense of [74–76]) corresponds to the case where $A = \partial\Omega$ and Ω is as above.

Furthermore, the main results of [74] (see Sect. 4.1 above, especially, Theorem 4.8) are used in an essential manner in order to show that the spectral zeta function of a fractal drum (that is, of the Dirichlet Laplacian on a bounded open subset of $\mathbb{R}^N$, with $N \geq 1$) can be meromorphically extended to (at least) the open half-plane $\{\text{Re}(s) > D\}$, where (as before) D is the (upper) Minkowski dimension of the boundary $\partial\Omega$ of the drum. [This was already observed in [76], using [74] and a well-known Abelian-type theorem (in the sense of [168]), the converse of a Tauberian theorem. It can also be deduced from the main result of [74] by a direct argument, based on the holomorphicity of an integral depending analytically on a parameter.] Moreover, using the construction of a maximal hyperfractal[10] carried out in [109, 110, 116] (and alluded to towards the end of Remark 7.19 above), it is shown that the half-plane $\{\text{Re}(s) > D\}$ is optimal, in general; i.e., it cannot usually be larger, from the point of view of the meromorphic continuation of the spectral zeta function. We note that under suitable assumptions, the Dirichlet Laplacian can be replaced by a higher order positive, self-adjoint elliptic operator, with possibly variable coefficients and with Dirichlet, Neumann or mixed boundary conditions; see [74, 109, 116].

Extensions of the results of [108–113, 115, 116] to unbounded sets, in particular, are provided in [140].

[10]Recall that in the terminology of [108–113, 115, 116], a "maximal hyperfractal" is a compact subset A of $\mathbb{R}^N$ such that ζ_A has a nonremovable singularity at every point of the "critical line" $\{\text{Re}(s) = D\}$; in addition to Remark 7.19, see Remark 7.22 below.

Finally, we note that ζ_A and $\tilde{\zeta}_A$ remain unchanged if we replace the bounded set $A \subset \mathbb{R}^N$ by its closure $\overline{A}$. As a result, we could have assumed throughout that A was a compact subset of $\mathbb{R}^N$. Moreover, we point out that fractal tube formulas, significantly extending to compact subsets of $\mathbb{R}^N$ the corresponding (pointwise and distributional) tube formulas obtained in [101, §8.1–8.3] for fractal strings (as well as the later tube formulas obtained for fractal sprays in [91, 92, 104, 105]), are established in [108, 113] and [116, Chap. 5], without any assumptions of self-similarity and in every dimension $N \geq 1$.

Remark 7.22. We close this discussion by mentioning the fact that the higher-dimensional theory of fractal zeta functions and the associated complex dimensions now enables us to extend to the general (higher-dimensional) setting the definition of fractality introduced in [99–101] (see, especially, [101, §12.1 and §12.2]). Accordingly, a geometric object is said to be "fractal" if it has at least one *nonreal* complex dimension (with a positive real part). In particular, one can now precisely talk about the fractality of any bounded subset A (or, more generally, RFD) in $\mathbb{R}^N$, with $N \geq 1$ arbitrary. (Compare with Remark 7.19 above.)

As an example, the Cantor curve (i.e., the "devil's staircase" in the terminology of [121]) is fractal according to this definition, whereas it is not fractal according to Mandelbrot's original definition in [121] (because its Hausdorff, box and topological dimensions coincide and are equal to one). (Recall from Remark 7.19 that in [121], a geometric object is said to be "fractal" if its Hausdorff dimension is strictly greater than its topological dimension.) One can also check that most of the classic "fractals" (for example, the ternary Cantor set and its generalizations, as well as the Sierpinski gasket and carpet and their higher-dimensional counterparts) are indeed fractal in this new sense. (See [109, 112, 113, 116] for these and many other examples.)

A technical challenge remains to prove this same result for all (suitable) self-similar sets (as was done when $N = 1$ for self-similar strings in [99–101] and when $N \geq 2$ for a large class of self-similar sprays or tilings in [91, 92, 104, 105]),[11] as well as for the classic types of fractals occurring in complex dynamics (Julia sets, the Mandelbrot set and their generalizations; see, e.g., [118, 121, 167]) and in conformal dynamics and/or hyperbolic geometry (for instance, limit sets of Fuchsian groups and Kleinian groups; see, e.g., [5]), possibly by using other gauge functions than the usual ones based on the standard power laws, as was originally done in [47] and further studied in [108–113, 115, 116] (from the point of view of the new higher-dimensional theory of complex dimensions). It may be difficult to do so, but there is no doubt in the author's mind that all of the classic (deterministic) fractals will eventually be found to be "fractal" in the above sense or "hyperfractal," in a sense to be explained next.

[11] *Added note*: In the case of self-similar sprays, the results of [108, 112, 113, 116] now enable one to recover and significantly extend the results of [91, 92, 104, 105].

Due in part to the work in [44] on random fractal strings and their complex dimensions, the notion of fractality was extended as follows (first in [100] and [101, §13.4.3] and then, in any dimension, in [108–113, 115, 116]). A geometric object is said to be "fractal" if it has at least one nonreal complex dimension with a positive real part (as above) or else if it is "hyperfractal"; i.e., if the associated fractal zeta function has a natural boundary along some suitable contour in $\mathbb{C}$ (a "screen," in the sense of [101, §5.3]). We note that the term "hyperfractal" was introduced in [108–113, 115, 116], in this context.

Therefore, a "hyperfractal" is such that the associated fractal zeta function cannot be meromorphically extended beyond a certain "screen." Furthermore, in [109–112, 115, 116] is also introduced the notion of "maximal hyperfractal," according to which the corresponding fractal zeta function has a nonremovable singularity at every point of the critical line of convergence $\{\mathrm{Re}(s) = D\}$, where D is the Minkowski dimension of $A \subset \mathbb{R}^N$ (or its relative counterpart, in the case of an RFD). It is then shown in ([109–112] or [116, Chap. 4]) using, in particular, countably many suitable fractal strings assembled in an appropriate way, along with Baker's theorem from transcendental number theory [4], that maximally hyperfractal strings ($N = 1$), as well as maximally hyperfractal compact subsets and RFDs of $\mathbb{R}^N$ (for every $N \geq 1$) can be explicitly constructed. This construction is completely deterministic. The author conjectures (building on [44] and [101, §13.4.3]) that for large classes of random fractals, maximal hyperfractality is an almost sure property.

In closing this remark, we mention that the above definition of fractality cannot just be applied to standard geometric objects embedded in Euclidean spaces (or, more generally, in appropriate metric measure spaces) but is also applicable, in principle, to 'virtual' geometries, spectral geometries, dynamical systems, algebraic and noncommutative geometries (not necessarily consisting of ordinary points; see, e.g., [19, 22]), as well as arithmetic geometries. In fact (along with the proper notion of zeta function and the associated complex dimensions), it should be used as a unifying tool between these apparently vastly different domains of mathematics. This long-term goal has been one of the central motivations of the author (and his collaborators) in [20, 30, 47, 74–85, 87, 88, 94, 98–101, 104, 116] (particularly, in [75, 76, 78, 80–82], as well as in the books [79, 99–101]).

For instance, the author conjectures that there exists a natural fractal-like geometric object whose complex dimensions are precisely the critical (i.e., nontrivial) zeros of the Riemann zeta function. Furthermore, the essential "shape" of this object should be understood in terms of a yet to be constructed cohomology theory associated with the underlying complex dimensions (and the pole of $\zeta = \zeta(s)$ at $s = 1$). In particular, connections with Deninger's work (and conjectures) in [24, 25] arise naturally in this context. (See, especially, [79, 81, 82], along with [101, §12.3 and §12.4].) Significant progress along those lines has recently been made by the author in [81, 82].

7.4 *Quantized Number Theory, Spectral Operators and the Riemann Hypothesis*

Formula (1), $\zeta_\nu(s) = \zeta(s) \cdot \zeta_{\mathscr{L}}(s)$, which connects the spectral zeta function $\zeta_\nu = \zeta_{\nu,\mathscr{L}}$ and the geometric zeta function $\zeta_{\mathscr{L}}$ of a fractal string $\mathscr{L}$ via the Riemann zeta function ζ, has the following counterpart, in terms of the spectral and geometric counting functions of $\mathscr{L}$ (see [101, Theorem 1.2]):

$$N_\nu(x) = \sum_{n=1}^{\infty} N_{\mathscr{L}}\left(\frac{x}{n}\right), \tag{45}$$

for any $x > 0$. Note that for a fixed $x > 0$, only finitely many terms are nonzero on the right-hand side of (45). However, the number of these terms tends to $+\infty$ as $x \to +\infty$.

The (*heuristic*) *spectral operator*, at the level of the counting functions, is then given by the map

$$g := N_{\mathscr{L}} \mapsto N_\nu(g) := \sum_{n=1}^{\infty} g\left(\frac{\cdot}{n}\right); \tag{46}$$

that is, $N_\nu(g)(x) := \sum_{n=1}^{\infty} g(x/n)$, for all $x > 0$. It can therefore be thought of as the map sending the geometry [represented by $N_{\mathscr{L}}$, given by Eq. (11)] of a fractal string $\mathscr{L}$ onto the spectrum (represented by $N_\nu = N_{\nu,\mathscr{L}}$) of a fractal string $\mathscr{L}$. Here, as in the discussion preceding Eq. (13),

$$N_\nu(x) = \#(\{f \in \sigma(\mathscr{L}) : f \le x\}), \tag{47}$$

for any $x > 0$, where $\sigma(\mathscr{L}) = \{n \cdot \ell_j^{-1} : n \ge 1, j \ge 1\}$ is the (frequency) spectrum of $\mathscr{L} = (\ell_j)_{j=1}^{\infty}$.

Remark 7.23. At an even more fundamental level (that of the "density of geometric states" $\eta := \sum_{j=1}^{\infty} \delta_{\{\ell_j^{-1}\}}$ and the "density of spectral states" $\nu := \sum_{f \in \sigma(\mathscr{L})} \delta_{\{f\}}$, see [101, §6.3.1]), the spectral operator can be viewed as the map sending the geometry (represented by η) of a fractal string $\mathscr{L} = (\ell_j)_{j=1}^{\infty}$ onto its spectrum (represented by ν) :

$$\eta := \sum_{j=1}^{\infty} \delta_{\{\ell_j^{-1}\}} \mapsto \nu = \nu(\eta) = \sum_{j,n=1}^{\infty} \delta_{\{n \cdot \ell_j^{-1}\}}, \tag{48}$$

where $\delta_{\{y\}}$ denotes the Dirac point mass at $y > 0$ and the "generalized fractal strings" η and ν are viewed as positive (local) measures (or as tempered distributions) on $(0, +\infty)$; see [101, §6.3.2]. We also refer to [101, §6.3.1] for explicit formulas expressing η and ν in terms of the underlying complex dimensions of $\mathscr{L}$, along

with [101, Chap. 4] for the notion of *generalized fractal string*, based on the notion of local measure from [27, 63, 64] and [116, Appendix A].

Viewed additively (that is, after having made the change of variable $x = e^t$, $t = \log x$, with $x > 0$ and $t \in \mathbb{R}$), the (heuristic) spectral operator, denoted by $\mathfrak{a}$, becomes the following map (where $f = f(t)$ is a suitable function of the new variable $t \in \mathbb{R}$):

$$f(t) \mapsto \mathfrak{a}(f)(t) := \sum_{n=1}^{\infty} f(t - \log n). \tag{49}$$

As before, let $\mathscr{P}$ denote the set of all prime numbers. Then, given any $p \in \mathscr{P}$, the associated (*local*) *Euler factor* $\mathfrak{a}_p$ is given by the operator

$$f(t) \mapsto \mathfrak{a}_p(f) = \sum_{m=0}^{\infty} f(t - m \log p). \tag{50}$$

The spectral operator $\mathfrak{a}$ and its (operator-valued) Euler factors $\mathfrak{a}_p$ (with $p \in \mathscr{P}$) are connected via the following (operator-valued) *Euler product* representation of $\mathfrak{a}$:

$$\mathfrak{a}(f)(t) = \left(\prod_{p \in \mathscr{P}} \mathfrak{a}_p(f) \right)(t), \tag{51}$$

with $\mathfrak{a}_p = (1 - p^{-\partial})^{-1}$. Here, we have let $\partial = d/dt$ denote the differentiation operator (also called the *infinitesimal shift* of the real line). We have (for any $h \in \mathbb{R}$)

$$(e^{-h\partial})(f)(t) = f(t - h); \tag{52}$$

in particular, we have

$$(n^{-h})(f)(t) = f(t - \log n), \tag{53}$$

so that we should also expect to be able to obtain the following (operator-valued) *Dirichlet series* representation of $\mathfrak{a}$:

$$\mathfrak{a}(f)(t) = \left(\sum_{n=1}^{\infty} n^{-\partial} \right)(f)(t). \tag{54}$$

All of these semi-heuristic formulas and definitions are given without proper mathematical justification in [101, §6.3.2] (and, originally, in [100, §6.3.2] where the heuristic spectral operator $\mathfrak{a}$ was first introduced). However, in [49–53], Herichi and Lapidus have developed a rigorous functional analytic framework within which (under suitable assumptions) all of these "definitions" and formulas are properly justified and, among many other results, the reformulation of the Riemann hypothesis obtained in [87] (see Theorem 6.15 above) can be restated in operator

theoretic terms; namely, as the "quasi-invertibility" of the spectral operator in all possible dimensions, except in the midfractal case $1/2$.

We refer the interested reader to the forthcoming book [53], titled *Quantized Number Theory, Fractal Strings and the Riemann Hypothesis*, for a complete exposition of the theory. We limit ourselves here to a brief exposition and a few additional comments.

For each fixed $c \in \mathbb{R}$, let $\mathbb{H}_c = L^2(\mathbb{R}, e^{-2ct}dt)$,[12] and let $\partial = d/dt$ be the unbounded (and densely defined) operator acting on the Hilbert space $\mathbb{H}_c$ via $\partial(f) = f'$, the distributional (or else, the pointwise almost everywhere defined) derivative of f, for all f in the domain $D(\partial)$ of ∂:

$$D(\partial) := \{f \in C_{abs}(\mathbb{R}) \cap \mathbb{H}_c : f' \in \mathbb{H}_c\}, \tag{56}$$

where $C_{abs}(\mathbb{R})$ is the space of (locally) absolutely continuous functions on $\mathbb{R}$ (see, e.g., [21, 35, 151]). Then, the *infinitesimal shift* $\partial = \partial_c$ is shown in [49, 50, 53] to be an unbounded normal operator on $\mathbb{H}_c$ (that is, $\partial^*\partial = \partial\partial^*$, where ∂^* denotes the adjoint of the closed unbounded operator ∂; see, e.g., [67] or [152]), with spectrum $\sigma(\partial) = \{\mathrm{Re}(s) = c\}$. Furthermore, the associated group $\{e^{-h\partial}\}h \in \mathbb{R}$ is shown to be the *shift group* given (for all $f \in \mathbb{H}_c$) by (52) while $n^{-\partial}$ is therefore given by (53).

Moreover, for every $c > 1$, the quantized (i.e., operator-valued) Euler product and Dirichlet series representations (51) and (54) of $\mathfrak{a} = \mathfrak{a}_c$ hold, in the following strong sense: $\mathfrak{a} = \sum_{n=1}^{\infty} n^{-\partial}$ and $\mathfrak{a} = \Pi_{p\in\mathscr{P}}(1 - p^{-\partial})^{-1}$, where both the series and the infinite product converge in $\mathscr{B}(\mathbb{H}_c)$, the Banach algebra of bounded linear operators on the Hilbert space $\mathbb{H}_c$. (See [52] or [53, Chap. 7].)

In addition, for all f in a suitable dense subspace of $D(\partial)$ (and hence, of $\mathbb{H}_c$), the *spectral operator* $\mathfrak{a} = \mathfrak{a}_c$, now rigorously defined (for every $c \in \mathbb{R}$) by the formula $\mathfrak{a} := \zeta(\partial)$, is given (for $c > 0$) by an appropriate operator-valued version of the classic analytic continuation of the Riemann zeta function $\zeta = \zeta(s)$ to the open half-plane $\{\mathrm{Re}(s) > 0\}$. Namely, formally, we have

$$\mathfrak{a} = \zeta(\partial) = \frac{1}{\partial - 1} + \sum_{n=1}^{\infty} \int_n^{n+1} (n^{-\partial} - t^{-\partial})dt. \tag{57}$$

(Compare, for example, with [165, §VI.3] or the proof of Theorem 5.12 above given in [94].) On this dense subspace, it also coincides with a quantized (and convergent) version of the Euler product and the Dirichlet series.

[12] Hence, $\mathbb{H}_c$ is the complex Hilbert space of (Lebesgue measurable, complex-valued) square integrable functions with respect to the absolutely continuous measure $e^{-2ct}dt$, and is equipped with the norm

$$||f||_c := \left(\int_{\mathbb{R}} |f(t)|^2 \, e^{-2ct} \, dt\right)^{1/2} < \infty \tag{55}$$

and the associated inner product, denoted by $(\cdot, \cdot)_c$.

Similarly, the *global spectral operator* $\mathscr{A} := \xi(\partial)$, where, as before, $\xi(s) := \pi^{-s/2}\Gamma(s/2)\,\zeta(s)$ is the *global* (or *completed*) *Riemann zeta function*, is shown to be given by an operator-valued (or quantized) version of the standard analytic continuation of $\xi = \xi(s)$ to the entire complex plane (see, e.g., [174, §2.6] or [79, §2.4]). Again, we refer the interested reader to [52] or [53, Chap. 7] for the precise statements and many related results.

We note that the spectral operator $\mathfrak{a} = \mathfrak{a}_c$, now defined for any $c \in \mathbb{R}$ by $\mathfrak{a} := \zeta(\partial)$, is obtained via the functional calculus for unbounded normal operators (see [152]); so that, formally, one substitutes the infinitesimal shift ∂ for the complex variable s in the usual definition of the Riemann zeta function (or, rather, in its meromorphic continuation to all of $\mathbb{C}$, still denoted by $\zeta = \zeta(s)$, for $s \in \mathbb{C}$).

As it turns out, according to the celebrated spectral theorem (for unbounded normal operators, see [152]), what matters are the values of $\zeta = \zeta(s)$ along the spectrum $\sigma(\partial)$ of ∂; that is, according to a key result of [49, 53], along

$$\sigma(\partial) = c\ell(\zeta(\{s \in \mathbb{C} : \mathrm{Re}(s) = c\}), \tag{58}$$

where $c\ell(G)$ denotes the closure of $G \subseteq \mathbb{C}$ in $\mathbb{C}$.

We point out that for $c = 1$, the unique (and simple) pole of ζ, we must assume that $s \neq 1$ on the right-hand side of (58). Alternatively, one views the meromorphic function ζ as a continuous function with values in the Riemann sphere $\tilde{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$, equipped with the standard chordal metric. (We set $\zeta(1) = \infty$.) Then, even for $c = 1$, the extended spectrum $\tilde{\sigma}(\mathfrak{a})$ of $\mathfrak{a}$ is given by the right-hand side of (58), *but without the closure*, where $\tilde{\sigma}(\mathfrak{a}) := \sigma(\mathfrak{a}) \cup \{\infty\}$ if $\mathfrak{a}$ is unbounded and $\tilde{\sigma}(\mathfrak{a}) := \sigma(\mathfrak{a})$ if $\mathfrak{a}$ is bounded; that is, according to [49–53], if $c \leq 1$ or if $c > 1$, respectively.

One therefore sees how the properties of the Riemann zeta function $\zeta = \zeta(s)$ (for example, its range along the vertical line $\{\mathrm{Re}(s) = 1\}$ or the existence of a pole at $s = 1$) are reflected in the properties of the spectral operator $\mathfrak{a} = \zeta(\partial)$ (for example, whether or not $\mathfrak{a}$ is bounded or invertible or, equivalently, whether or not its spectrum $\sigma(\mathfrak{a})$ is compact or does not contain the origin, respectively).

Formula (58) for $\sigma(\mathfrak{a})$ follows from a suitable version of the *spectral mapping theorem* (SMT) for unbounded normal operators (see [53, Appendix E]) according to which $\sigma(\zeta(\partial)) = c\ell(\zeta(\sigma(\partial))\backslash\{1\})$ or, equivalently, $\tilde{\sigma}(\zeta(\partial)) = \zeta(\sigma(\partial))$ (where in the latter formula, ζ is viewed as a continuous $\tilde{\mathbb{C}}$-valued function, as explained above). Note that for $c \neq 1$, we have that $1 \notin \sigma(\partial)$, so that $\zeta = \zeta(s)$ is holomorphic and hence, continuous along $\sigma(\partial) = \{\mathrm{Re}(s) = c\}$, whereas for $c = 1$, $\zeta = \zeta(s)$ is meromorphic but has a singularity (more specifically, a simple pole) at $s = 1 \in \sigma(\partial)$. Consequently, we can use the continuous (resp., meromorphic) version of SMT (stated and proved in [53, Appendix E]) in order to deduce that when $c \neq 1$ (resp., $c = 1$), the identity (58) (resp., the counterpart of (58) when $c = 1$) holds.

Next, given $T > 0$, one defines (still via the functional calculus for unbounded normal operators) a *truncated infinitesimal shift* $\partial^{(T)} = \partial_c^{(T)}$ (say, $\partial^{(T)} := \varphi^{(T)}(\partial)$, where $\varphi^{(T)}$ is a suitable cut-off function), so that (again by the aforementioned SMT), $\sigma(\partial^{(T)}) = [c - iT, c + iT]$, where $c \in \mathbb{R}$ and $i := \sqrt{-1}$. (We refer to [49, 50, 53] for the precise definitions; see also Remark 7.24 just below.)

Then, letting $\mathfrak{a}^{(T)} := \zeta(\partial^{(T)})$, we obtain the *truncated spectral operator* $\mathfrak{a}^{(T)} = \mathfrak{a}_c^{(T)}$. We say that the spectral operator $\mathfrak{a}$ is *quasi-invertible* if each of its truncations $\mathfrak{a}^{(T)}$ is invertible (in the usual sense, for bounded operators), for every $T > 0$. (Note that the normal operator $\partial^{(T)}$ is bounded since its spectrum is compact.)

Remark 7.24. For $c \neq 1$, the cut-off function $\varphi^{(T)} : \sigma(\partial) = \{\mathrm{Re}(s) = c\} \to \mathbb{C}$ is assumed to satisfy the following two conditions: (i) $\varphi^{(T)}$ is continuous and, in addition, (ii) $c\ell(\varphi^{(T)}(\{\mathrm{Re}(s) = c\})) = [c - iT, c + iT]$. For $c = 1$, we replace (i) by (i') : $\varphi^{(T)}$ has a (necessarily unique) meromorphic continuation to a connected open neighborhood of $\{\mathrm{Re}(s) = 1\}$. (Then, condition (i') and (ii) imply that (i) also holds.) Consequently, when $c \neq 1$ (resp., c =1), one can use the continuous (resp., meromorphic) version of the SMT (see [53, Appendix E]) in order to deduce that

$$\begin{aligned} \sigma(\partial^{(T)}) &= \sigma(\varphi^{(T)}(\partial)) \\ &= c\ell(\varphi^{(T)}(\{\mathrm{Re}(s) = c\})) = [c - iT, c + iT], \end{aligned} \tag{59}$$

and especially, that

$$\begin{aligned} \sigma(\mathfrak{a}^{(T)}) &= \sigma(\zeta(\partial^{(T)})) \\ &= c\ell(\zeta(\sigma(\partial^{(T)}))) = c\ell(\zeta([c - iT, c + iT])), \end{aligned} \tag{60}$$

where for $c = 1$, one should remove $\{1\}$ from $\sigma(\partial^{(T)}) = [c - iT, c + iT]$ in the last two equalities of (60), while when $c \neq 1$, (60) takes the following simpler form (since then, ζ is continuous along the vertical line segment $[c - iT, c + iT]$):

$$\sigma(\mathfrak{a}^{(T)}) = \zeta([c - iT, c + iT]). \tag{61}$$

When $c = 1$, (61) can be replaced by the following identity, which is equivalent to (60) interpreted as above (since $\zeta(1) = \{\infty\}$ and $\zeta(s) \neq \infty$ for all $s \in \mathbb{C}, s \neq 1$) :

$$\tilde{\sigma}(\mathfrak{a}^{(T)}) := \sigma(\mathfrak{a}^{(T)}) \cup \{\infty\} = \zeta([1 - iT, 1 + iT]), \tag{62}$$

where (as in an earlier comment) the meromorphic function ζ should be viewed as a $\tilde{\mathbb{C}}$-valued continuous function in the right-hand side of (62) and where $\tilde{\sigma}(\mathfrak{a}^{(T)})$ denotes the extended spectrum of $\mathfrak{a}^{(T)}$.

The full strength of the definition given in the following remark will be needed when we discuss Theorem 7.29 below. This definition and the accompanying property is also used in the statement and the proof of Theorem 7.26 (and hence, of Theorem 7.27 as well).

Remark 7.25. A possibly unbounded linear operator $L : D(L) \subseteq H \to H$ on a Hilbert space H, with domain $D(L)$, is said to be *invertible* if it is a bijection from $D(L)$ onto H and if its inverse, L^{-1}, is bounded. (If L is closed, an assumption which is satisfied by all of the operators considered in this section, then L^{-1} is

automatically bounded, by the closed graph theorem [35, 67, 151].) Furthermore, essentially by definition of the spectrum, L is invertible if and only if $0 \notin \sigma(L)$. (See, e.g., [67, 147, 152, 158].)

We can now state the counterparts of Theorems 6.14 and 6.15 (the key results from [87] discussed in Sect. 6 above) in this context. (See Remark 7.28 below.)

Theorem 7.26 (Analog of Theorem 6.14 [50, 53]). *Given any $c \in \mathbb{R}$, the spectral operator $\mathfrak{a} = \mathfrak{a}_c$ is quasi-invertible if and only if $\zeta = \zeta(s)$ does not have any zeros along the vertical line $\{Re(s) = c\}$; i.e., if and only the "c-partial Riemann hypothesis" is true.*

Hence, much as in Corollary 6.16, $\mathfrak{a}_{1/2}$ is *not* quasi-invertible (since ζ has at least one zero along the critical line $\{\mathrm{Re}(s) = 1/2\}$).[13] Furthermore, $\mathfrak{a}_c$ is quasi-invertible for every $c > 1$. Actually, for $c > 1$, $\mathfrak{a} = \mathfrak{a}_c$ is invertible (which implies that it is quasi-invertible) and its inverse is given by $\mathfrak{a}^{-1} = \prod_{p \in \mathscr{P}}(1 - p^{-\partial})$, with the convergence of the infinite product holding in $\mathscr{B}(\mathbb{H}_c)$; see [52, 53].

Just as Theorem 6.15 is a consequence of Theorem 6.14, the following key result follows from Theorem 7.26.

Theorem 7.27 (Analog of Theorem 6.15 [50, 53]). *The spectral operator is quasi-invertible for every value of c in $(0, 1)$ other than in the midfractal case when $c = 1/2$ (or, equivalently, for every $c \in (0, 1/2)$) if and only if the Riemann hypothesis is true.*

Remark 7.28. Recall from Sect. 6 that Theorems 6.14 and 6.15 are expressed in terms of the solvability of the inverse spectral problem $(\mathrm{ISP})_D$ for all fractal strings of Minkowski dimension D. Here, the role played by the dimension D is now played by the parameter c, while the solvability of $(\mathrm{ISP})_D$ is now replaced by the quasi-invertibility of $\mathfrak{a}_c$. In a precise sense, c is the analog of D in the present context; this analogy is based in part on some of the results of [94] which are briefly discussed in Sect. 3 above (see Theorem 3.3 and the comments following it), provided one thinks of the functions $f = f(t)$ as geometric counting functions of fractal strings (modulo the change of variable $x = e^t$).

We leave it to the interested reader to state (in the present context) the counterpart of Theorem 6.17 (about a mathematical phase transition at $c = 1/2$).

Like Theorem 6.15, Theorem 7.27 is a *symmetric criterion for the Riemann hypothesis* (RH). Indeed, in light of the functional equation (28) for ζ connecting $\zeta(s)$ and $\zeta(1-s)$, we can replace the open interval $(0, 1/2)$ by the symmetric interval $(1/2, 1)$, with respect to $1/2$. (See also Eq. (29) above.) In contrast, the author has recently discovered the following *asymmetric criterion for* RH, as we now explain. (See [80].)

[13] In fact, ζ has infinitely many zeros along the critical line, according to Hardy's theorem (see [29, 174]) but this is irrelevant here. We refer to [50, 53] for a version of Theorem 7.26 for which this well-known fact actually matters.

Let $\mathfrak{b} = \mathfrak{b}_c$ be the nonnegative self-adjoint operator defined by $\mathfrak{b} := \mathfrak{a}\mathfrak{a}^* = \mathfrak{a}^*\mathfrak{a}$. Note that $\mathfrak{b}$ is a nonnegative self-adjoint operator because $\mathfrak{a}$ is normal (see, e.g., [67]) and, like $\mathfrak{a}$ itself, is unbounded for all $c \in (0,1)$, while it is bounded for all $c > 1$ (see [49, 50, 53]). Then, $\mathfrak{b}$ is invertible (in the sense of possibly unbounded operators, see Remark 7.25 above) if and only if $\mathfrak{b}$ is bounded away from zero (i.e., if and only if there exists $\gamma > 0$ such that $(\mathfrak{b}f, f)_c \geq \gamma ||f||_c^2$, for all $f \in D(\mathfrak{b})$). Furthermore, $\mathfrak{b}$ is invertible if and only if $\mathfrak{a}$ is invertible, which is the case if and only $0 \notin \sigma(\mathfrak{a})$ or, equivalently, $0 \notin \sigma(\mathfrak{b})$.

Theorem 7.29 (An Asymmetric Reformulation of the Riemann Hypothesis [80]). *The following statements are equivalent*:

(i) The Riemann hypothesis is true.
(ii) The spectral operator $\mathfrak{a}$ *is invertible for all* $c \in (0, 1/2)$.
(iii) The self-adjoint operator $\mathfrak{b}$ *is invertible for all* $c \in (0, 1/2)$.
(iv) For each $c \in (0, 1/2)$, $\mathfrak{b}$ *is bounded away from zero.*

We know that in part (iv) of Theorem 7.29, the implicit lower bound $\gamma = \gamma_c$ may vary with c and even tend to 0 as $c \to (1/2)^-$ or $c \to 0^+$.

A priori, the phase transition observed at $c = 1/2$ in Theorem 7.29 just above is of a very different nature from the one observed in Theorem 6.15 (based on [87], see also Theorem 6.17) or in Theorem 7.27 (based on [50, 53], see the comment following Remark 7.28). Indeed, under RH (and, in fact, if and only RH holds, by Theorem 7.29), we have that, for all $c \in (0, 1/2)$, $\sigma(\mathfrak{a})$ is a closed unbounded subset of $\mathbb{C}$ not containing 0 and hence, $\mathfrak{a}$ is invertible. This follows from the expression (58) obtained in [49, 50, 53] for $\sigma(\mathfrak{a})$, combined with a conditional result (by Garunkštis and Steuding; see the proof of Lemma 4 and Proposition 5 in [37]) about the non-universality of ζ in the left critical strip $\{0 < \mathrm{Re}(s) < 1/2\}$.

On the other hand, for any $c \in (1/2, 1)$, we have that $\sigma(\mathfrak{a}) = \mathbb{C}$ and hence, that $\mathfrak{a}$ is not invertible. This follows again from (58), but now combined with the Bohr–Courant theorem [11] according to which for every $c \in (1/2, 1)$, the range of ζ along any vertical line $\{\mathrm{Re}(s) = c\}$ is dense in $\mathbb{C}$. This latter fact is a consequence of the universality of ζ in the right critical strip $\{1/2 < \mathrm{Re}(s) < 1\}$, as was first observed and established by Voronin in [177, 178].

Consequently, the (unconditional) universality of the Riemann zeta function ζ in the right critical strip $\{1/2 < \mathrm{Re}(s) < 1\}$ and the (conditional) non-universality of ζ in the left critical strip $\{0 < \mathrm{Re}(s) < 1/2\}$ are absolutely crucial in order to understand the mathematical phase transition occurring in the midfractal case when $c = 1/2$, according to Theorem 7.29, if and only RH holds.

We stress that the universality of ζ, initially discovered by Voronin in [178] in the mid-1970s after he extended the Bohr–Courant theorem to jets of ζ (consisting of ζ and its derivatives), in [177], has since been generalized in a number of ways; see, e.g., the books [66, 117, 171], along with [2, 3, 51, 53, 148, 149, 170] and the many relevant references therein.

Roughly speaking, the universality theorem, as generalized in [2, 3, 148], states that given any compact subset K of the right critical strip $\{1/2 < \mathrm{Re}(s) < 1\}$

with connected complement in $\mathbb{C}$, every $\mathbb{C}$-valued continuous function on K which is holomorphic and nowhere vanishing on the interior of K can be uniformly approximated on K by vertical translates of ζ.

The universality theorem has been extended to a large class of L-functions for which the GRH is expected to hold; see, e.g., [171] for an exposition, along with [170] and the relevant appendices of [51, 53]. Accordingly, Theorem 7.29 can be extended to many L-functions. Similarly, but for different and simpler reasons, after a suitable modification in the definition of the truncated infinitesimal shift $\partial^{(T)}$, Theorems 7.26 and 7.27 (based on [49–51, 53]) can be extended to essentially all of the L-functions (or arithmetic zeta functions) for which GRH is expected to hold (independently of whether or not the analog of the universality theorem holds for those zeta functions); see, e.g., [131, 132, 157], [101, Appendix A] and [79, Appendices B,C,E].

Moreover, we point out that a counterpart in the present context [that is, for the spectral operator $\mathfrak{a} = \zeta(\partial)$] of the universality of ζ is provided in [51, 53]. Interestingly, the corresponding "quantized universality" then involves the family of truncated spectral operators $\{\mathfrak{a}^{(T)} = \zeta(\partial^{(T)})\}_{T>0}$; that is, the complex variable s is not replaced by the operator ∂, as one might naively expect, but by the family of truncated infinitesimal shifts $\{\partial^{(T)}\}_{T>0}$ discussed earlier in this section; see [51] and [53].

In addition, we note that the (conditional) phase transition at $c = 1/2$ associated with Theorem 7.29 (and also with Theorem 6.15, from [86, 87], which preceded the work in [14, 15]; see also its interpretation given in Theorem 6.17, from [75, 76]) is of a very different nature from the one studied by Bost and Connes in [14, 15]. (See also [22, §V.11].) Indeed, the latter phase transition has nothing to do with the universality or with the critical zeros of ζ but instead, is merely connected with the pole of $\zeta(s)$ at $s = 1$.[14] However, it is natural to wonder whether the conditional phase transition occurring at $c = 1/2$ in Theorem 7.29 (if and only if RH holds) can be interpreted physically and mathematically (as in [14, 15]) as some kind of "symmetry breaking" associated with a suitable physical model (in quantum statistical physics or quantum field theory, for example) as well as corresponding to a change in the nature of an appropriate symmetry group yet to be attached to the present situation.

Remark 7.30. It is also natural to wonder what is the inverse of the spectral operator, when $0 < c < 1/2$ and assuming that RH holds. Naturally, according to the functional calculus, we must have $\mathfrak{a}^{-1} = (1/\zeta)(\partial)$, even if $\mathfrak{a}^{-1}$ is not bounded (i.e., even if $D(\mathfrak{a}^{-1}) \neq \mathbb{H}_c$ or equivalently, if $R(\mathfrak{a}) \neq \mathbb{H}_c$, where $R(\mathfrak{a})$ denotes the range of $\mathfrak{a}$). We conjecture that under RH, we have for all $c \in (0, 1/2)$ and all $f \in D(\mathfrak{a}^{-1}) = \mathbb{H}_c$, $\mathfrak{a}^{-1}(f) = \sum_{n=1}^{\infty} \mu(n) n^{-\partial}(f)$, so that $\mathfrak{a}^{-1}(f)(t) = \sum_{n=1}^{\infty} \mu(n) f(t - \log n)$, both

[14] A similar phase transition (or "symmetry breaking") is observed at $c = 1$, as is discussed in [49–51, 53], since $\mathfrak{a}$ is bounded and invertible for $c > 1$, unbounded and not invertible for $1/2 \leq c \leq 1$, while, likewise, $\sigma(\mathfrak{a})$ is a compact subset of $\mathbb{C}$ not containing the origin for $c > 1$ and $\sigma(\mathfrak{a}) = \mathbb{C}$ is unbounded and contains the origin if $1/2 < c < 1$.

for a.e. $t \in \mathbb{R}$ and as an identity between functions in $\mathbb{H}_c$. (Here, $\mu = \mu(n)$ denotes the classic Möbius function, given by $\mu(n) = 1$ or -1, respectively, if $n \geq 2$ is a square-free integer that is a product of an even or odd number of distinct primes, respectively, and $\mu(n) = 0$ otherwise; see, e.g., [29] or [174].) Observe that if that were the case, then "quantized Dirichlet series" would behave very differently from ordinary (complex-valued) Dirichlet series. In particular, the quantized Riemann zeta function $\mathfrak{a} = \zeta(\partial)$ would be rather different from its classic counterpart, the ordinary Riemann zeta function, $\zeta = \zeta(s)$ (where $s \in \mathbb{C}$). Indeed, as is well known, even if the Riemann hypothesis is true, the series $\sum_{n=1}^{\infty} \mu(n)\, n^{-s}$ cannot converge for any value of $s := s_0$ with $0 < Re(s_0) < 1/2$. Otherwise, it would follow from the standard properties of (numerical) Dirichlet series and the Möbius inversion formula (see, e.g., [29]) that the sum of the series $\sum_{n=1}^{\infty} \mu(n) n^{-s}$ would have to be holomorphic and to coincide with $(1/\zeta)(s)$ in the half-plane $\{\mathrm{Re}(s) > \mathrm{Re}(s_0)\}$, which is, of course, impossible because the meromorphic function $1/\zeta$ must have a pole at every zero of ζ along the critical line $\{\mathrm{Re}(s) = 1/2\}$ (of which there are infinitely many).

We close this discussion by a final comment related to the operator $\mathfrak{b}$ (appearing in the statement of parts (iii) and (iv) of Theorem 7.29) and to the possible origin of the (conditional) phase transition occurring in the midfractal case $c = 1/2$.

Remark 7.31. The nonnegative self-adjoint operator $\mathfrak{b} = \mathfrak{a}\mathfrak{a}^* = \mathfrak{a}^*\mathfrak{a}$ is given by the following formula:

$$\mathfrak{b}(f)(t) = \zeta(2c) \sum_{(k,n)=1} f\left(t - \log \frac{k}{n}\right) n^{-2c}, \tag{63}$$

where the sum is taken over all pairs of integers $k, n \geq 1$ without common factor. The appearance of the terms n^{-2c} (for $n \geq 1$) and of the factor $\zeta(2c)$ are very interesting features of this formula (which first appeared in [100, 101, §6.3.2] without formal justification). A priori, the above formula was derived (in [52, 53], motivated in part by the new result of [80] described in Theorem 7.29 above) by assuming that $c > 1$. However, the series $\sum_{n=1}^{\infty} n^{-2c}$ is convergent for $c > 1/2$ and the term $\zeta(2c)$ is singular only at $c = 1/2$, due to the pole of $\zeta(s)$ at $s = 1$. This may shed new light on the origin of the phase transition occurring at $c = 1/2$ in the statement of Theorem 7.29.

Given the reformulation of the Riemann hypothesis provided above in terms of the invertibility of the spectral operator $\mathfrak{b} = \mathfrak{b}_c$, where $\mathfrak{b} = \mathfrak{a}^*\mathfrak{a} = \mathfrak{a}\mathfrak{a}^*$ (see the equivalence of (i), (iii), and (iv) in Theorem 7.29), it is tempting to extend formula (63) to a suitable class $\mathscr{C}$ of "test functions" $f = f(t)$ (technically, a "core" for the operator $\mathfrak{b}$, that is, a dense subspace of $D(\mathfrak{b})$ in the Hilbert graph norm $|||f|||_c := (||f||_c^2 + ||bf||_c^2)^{1/2}$) so that it becomes valid for all $c \in (0, 1/2)$.

More specifically, for each $c \in (0, 1/2)$, we would like to show that there exists a constant $\gamma = \gamma_c > 0$ (which may depend on the parameter c) such that

$$||bf||_c \geq \gamma ||f||_c, \tag{64}$$

for all $f \in \mathscr{C}$. (Compare with part (iv) of Theorem 7.29 above as well as with the comments preceding the statement of Theorem 7.29.)

In light of the new asymmetric criterion for RH obtained in [80] (see Theorem 7.29 and, especially, the equivalence of (i) and (iv) in Theorem 7.29), this conjectured inequality (64) would imply (in fact, would be equivalent to) the Riemann hypothesis. Thus far, however, the author has only been able to verify it for a certain class of test functions, unfortunately not yet large enough to fulfill the required conditions.

A last comment is in order.

Remark 7.32. In two works in preparation [81, 82], the author explores some of the applications of this new formalism ("quantized number theory" from [49–53, 80]) to potentially reformulating the Weyl conjectures (for curves and higher-dimensional varieties over finite fields or, equivalently, for function fields),[15] as well as for constructing generalized Polya–Hilbert operators[16] with spectra the Riemann zeros or, in fact, the zeros (and the poles) of general L-functions. For this purpose, Lapidus has introduced a related but different formalism, in [81, 82] associated with harmonic analysis and operator theory in weighted Bergman spaces of analytic functions (see, e.g., [1, 48]).[17] On the one hand, this new functional analytic framework offers greater flexibility and ease of use, since it only involves bounded operators, albeit of an unusual nature. On the other hand, the framework (from [49–53, 80]) discussed in this section presents the significant advantage, in particular, of naturally parametrizing the critical strip $0 < \mathrm{Re}(s) < 1$ by means of the dimension parameter $c \in (0, 1)$. Only future research on both approaches will help us to eventually determine whether one formalism should be preferred to the other one, or as is more likely to happen, whether both formalisms should be used in order to further develop different aspects of quantized number theory.

Acknowledgements This research was partially supported by the US National Science Foundation (NSF) under the grants DMS-0707524 and DMS-1107750 (as well as by many earlier NSF grants since the mid-1980s). Part of this work was completed during several stays of the author as a visiting professor at the Institut des Hautes Etudes Scientifiques (IHES) in Paris/Bures-sur-Yvette, France.

[15] See, e.g., [79, Appendix B] and the references therein.

[16] See, e.g., [79, Chaps. 4 and 5], along with the relevant references therein.

[17] Some of this work may eventually become joint work with one of the author's current Ph.D. students, Tim Cobler.

References

1. A. Atzmon, B. Brive, Surjectivity and invariant subspaces of differential operators on weighted Bergman spaces of entire functions, in *Bergman Spaces and Related Topics in Complex Analysis*, ed. by A. Borichev, H. Hedenmalm, K. Zhu. Contemporary Mathematics, vol. 404 (American Mathematical Society, Providence, RI, 2006), pp. 27–39
2. B. Bagchi, The statistical behaviour and universality properties of the Riemann zeta-function and other allied Dirichlet series. Ph.D. thesis, Indian Statistical Institute, Calcutta, 1981
3. B. Bagchi, A joint universality theorem for Dirichlet L-functions. Math. Z. **181**, 319–334 (1982)
4. A. Baker, *Transcendental Number Theory* (Cambridge University Press, Cambridge, 1975)
5. T. Bedford, M. Keane, C. Series (eds.), *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces* (Oxford University Press, Oxford, 1991)
6. M.V. Berry, Distribution of modes in fractal resonators, in *Structural Stability in Physics*, ed. by W. Güttinger, H. Eikemeier. Graduate Texts in Mathematics, vol. 125 (Springer, Berlin, 1979), pp. 51–53
7. M.V. Berry, Some geometric aspects of wave motion: wavefront dislocations, diffraction catastrophes, diffractals, in *Geometry of the Laplace Operator*. Proceedings of Symposia in Pure Mathematics, vol. 36 (American Mathematical Society, Providence, RI, 1980), pp. 13–38
8. A.S. Besicovitch, S.J. Taylor, On the complementary intervals of a linear closed set of zero Lebesgue measure. J. Lond. Math. Soc. **29**, 449–459 (1954)
9. H. Bohr, Zur Theorie der Riemannschen ZetaFunktion im kritischen Streifen. Acta Math. **40**, 67–100 (1915)
10. H. Bohr, Über eine quasi-periodische Eigenschaft Dirichletscher Reihen mit Anwendung auf die Dirichletschen L-Funktionen. Math. Ann. **85**, 115–122 (1922)
11. H. Bohr, R. Courant, Neue Anwendungen der Theorie der diophantischen Approximationen auf die Riemannsche Zetafunktion. J. Reine Angew. Math. **144**, 249–274 (1914)
12. H. Bohr, E. Landau, Über das Verhalten von $\zeta(s)$ und $\zeta^{(k)}(s)$ in der Nähe der Geraden $\sigma = 1$. Nachr. Ges. Wiss. Göttingen Math. Phys. **K1**, 303–330 (1910)
13. H. Bohr, E. Landau, Ein Satz über Dirichletsche Reihen mit Anwendung auf die ζ-Funktion und die L-Funktionen. Rend. di Palermo **37**, 269–272 (1914)
14. J.-B. Bost, A. Connes, Produit eulérien et facteurs de type III. C. R. Acad. Sci. Paris Sér. I Math. **315**, 279–284 (1992)
15. J.-B. Bost, A. Connes, Hecke algebras, type III factors and phase transitions with spontaneous symmetry breaking in number theory. Sel. Math. N.S. **1**, 411–457 (1995)
16. G. Bouligand, Ensembles impropres et nombre dimensionnel. Bull. Sci. Math. **52**(2), 320–344/361–376 (1928)
17. H. Brezis, *Analyse Fonctionnelle*: *Théorie et Applications* (Masson, Paris, 1983). Expanded English version: *Functional Analysis, Sobolev Spaces and Partial Differential Equations* (Springer, New York, 2011)
18. J. Brossard, R. Carmona, Can one hear the dimension of a fractal? Commun. Math. Phys. **104**, 103–122 (1986)
19. P. Cartier, A mad day's work: from Grothendieck to Connes and Kontsevich. The evolution of concepts of space and symmetry (English transl. of the French original). Bull. Am. Math. Soc. (N.S.) **38**, 389–408 (2001)
20. E. Christensen, C. Ivan, M.L. Lapidus, Dirac operators and spectral triples for some fractal sets built on curves. Adv. Math. **217**(1), 42–78 (2008). Also: e-print, `arXiv:math.MG/0610222v2`, 2007
21. D.L. Cohn, *Measure Theory* (Birkhäuser, Boston, 1980)
22. A. Connes, *Noncommutative Geometry* (Academic, New York, 1994)
23. R. Courant, D. Hilbert, *Methods of Mathematical Physics*, vol. I (Interscience Publishers Inc., New York, 1953)

24. C. Deninger, Lefschetz trace formulas and explicit formulas in analytic number theory. J. Reine Angew. Math. **441**, 1–15 (1993)
25. C. Deninger, Evidence for a cohomological approach to analytic number theory, in *Proceedings of First European Congress of Mathematics*, Paris, July 1992, vol. I, ed. by A. Joseph et al. (Birkhäuser-Verlag, Basel and Boston, 1994), pp. 491–510
26. G. Derfel, P. Grabner, F. Vogl, The zeta function of the Laplacian on certain fractals. Trans. Amer. Math. Soc. **360**, 881–897 (2008)
27. J.D. Dollard, C.N. Friedman, *Product Integration, with Application to Differential Equations*. Encyclopedia of Mathematics and Its Applications, vol. 10 (Addison-Wesley, Reading, 1979)
28. N. Dunford, J.T. Schwartz, *Linear Operators*. Wiley Classics Library, Parts I–III (Wiley, Hoboken, 1971/1988)
29. H.M. Edwards, *Riemann's Zeta Function* (Academic, New York, 1974)
30. K.E. Ellis, M.L. Lapidus, M.C. Mackenzie, J.A. Rock, Partition zeta functions, multifractal spectra, and tapestries of complex dimensions, in *Benoît Mandelbrot: A Life in Many Dimensions*, ed. by M. Frame, N. Cohen. The Mandelbrot Memorial Volume (World Scientific, Singapore, 2015), pp. 267–322. Also: e-print, `arXiv:1007.1467v2[math-ph]`, 2011; IHES preprint, IHES/M/12/15, 2012
31. K.J. Falconer, On the Minkowski measurability of fractals. Proc. Amer. Math. Soc. **123**, 1115–1124 (1995)
32. K.J. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, 3rd edn. (Wiley, Chichester, 2014). First and second editions: 1990 and 2003
33. H. Federer, *Geometric Measure Theory* (Springer, New York, 1969)
34. J. Fleckinger, D. Vassiliev, An example of a two-term asymptotics for the "counting function" of a fractal drum. Trans. Amer. Math. Soc. **337**, 99–116 (1993)
35. G.B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd edn. (Wiley, New York, 1999)
36. M. Fukushima, T. Shima, On a spectral analysis for the Sierpinski gasket. Potential Anal. **1**, 1–35 (1992)
37. R. Garunkštis, J. Steuding, On the roots of the equation $\zeta(s) = \alpha$. Abh. Math. Seminar Univ. Hamburg **84**, 1–15 (2014). Also: e-print, `arXiv:1011.5339v2 [math.NT]`, 2014
38. J. Gerling, Untersuchungen zur Theorie von Weyl–Berry–Lapidus. Graduate thesis (Diplomarbeit), Department of Physics, Universität Osnabrück, Osnabrück, 1992
39. J. Gerling, H.-J. Schmidt, Self-similar drums and generalized Weierstrass functions. Physica A **191**(1–4), 536–539 (1992)
40. P.B. Gilkey, *Invariance Theory, the Heat Equation, and the Atiyah-Singer Index Theorem,* 2nd edn. (Publish or Perish, Wilmington, 1984). New revised and enlarged edition in *Studies in Advanced Mathematics* (CRC Press, Boca Raton, 1995)
41. J.A. Goldstein, *Semigroups of Linear Operators and Applications*. Oxford Science Publications, Oxford Mathematical Monographs (Oxford University Press, Oxford and New York, 1985)
42. B. M. Hambly, Brownian motion on a random recursive Sierpinski gasket. Ann. Probab. **25**, 1059–1102 (1997)
43. B.M. Hambly, On the asymptotics of the eigenvalue counting function for random recursive Sierpinski gaskets. Probab. Theory Relat. Fields **117**, 221–247 (2000)
44. B.M. Hambly, M.L. Lapidus, Random fractal strings: their zeta functions, complex dimensions and spectral asymptotics. Trans. Amer. Math. Soc. **358**(1), 285–314 (2006)
45. G.H. Hardy, E.M. Wright, *An Introduction to the Theory of Numbers*, 6th edn. (Oxford University Press, Oxford, 2008)
46. R. Harvey, J. Polking, Removable singularities of solutions of linear partial differential equations. Acta Math. **125**, 39–56 (1970)
47. C.Q. He, M.L. Lapidus, Generalized Minkowski content, spectrum of fractal drums, fractal strings and the Riemann zeta-function. Mem. Amer. Math. Soc. **127**(608), 1–97 (1997)
48. H. Hedenmalm, B. Korenblum, K. Zhu, *Theory of Bergman Spaces*. Graduate Texts in Mathematics, vol. 199 (Springer, New York, 2000)

49. H. Herichi, M.L. Lapidus, Riemann zeros and phase transitions via the spectral operator on fractal strings. J. Phys. A Math. Theor. **45**, 374005, 23 pp. (2012). Also: e-print, arXiv:1203.4828v2[math-ph], 2012; IHES preprint, IHES/M/12/09, 2012
50. H. Herichi, M.L. Lapidus, Fractal complex dimensions, Riemann hypothesis and invertibility of the spectral operator, in *Fractal Geometry and Dynamical Systems in Pure and Applied Mathematics I: Fractals in Pure Mathematics*, ed. by D. Carfi, M.L. Lapidus, E.P.J. Pearse, M. van Frankenhuijsen. Contemporary Mathematics, vol. 600 (American Mathematical Society, Providence, RI, 2013), pp. 51–89. Also: e-print, arXiv:1210.0882v3[math.FA], 2013; IHES preprint, IHES/M/12/25, 2012
51. H. Herichi, M.L. Lapidus, Truncated infinitesimal shifts, spectral operators and quantized universality of the Riemann zeta function. Annales de la Faculté des Sciences de Toulouse **23**(3), 621–664 (2014). Special issue in honor of Christophe Soulé. Also: e-print, arXiv:1305.3933v2[math-NT], 2015; IHES preprint, IHES/M/13/12, 2013
52. H. Herichi, M.L. Lapidus, Quantized Riemann zeta functions: its operator-valued Dirichlet series, Euler product and analytic continuation (2015, in preparation)
53. H. Herichi, M. L. Lapidus, *Quantized Number Theory, Fractal Strings and the Riemann Hypothesis*: *From Spectral Operators to Phase Transitions and Universality*. Research Monograph (World Scientific Publ., Singapore, 2016, to appear). Approx. 240 pp.
54. E. Hille, R.S. Phillips, *Functional Analysis and Semi-Groups*. American Mathematical Society Colloquium Publications, vol. XXXI, revised edn. (American Mathematical Society, RI, 1957)
55. L. Hörmander, The spectral function of an elliptic operator. Acta Math. **121**, 193–218 (1968)
56. L. Hörmander, *The Analysis of Linear Partial Differential Operators*, vols. II–IV (Springer, Berlin, 1983/1985)
57. L. Hörmander, *The Analysis of Linear Partial Differential Operators*. Distribution Theory and Fourier Analysis, vol. I, 2nd edn. (of the 1983 edn.) (Springer, Berlin, 1990)
58. A.E. Ingham, *The Distribution of Prime Numbers*, 2nd edn. (reprinted from the 1932 edn.) (Cambridge University Press, Cambridge, 1992)
59. A. Ivic, *The Riemann Zeta-Function: The Theory of the Riemann Zeta-Function with Applications* (Wiley, New York, 1985)
60. V.Ja. Ivrii, Second term of the spectral asymptotic expansion of the Laplace-Beltrami operator on manifolds with boundary. Funct. Anal. Appl. **14**, 98–106 (1980)
61. V.Ja. Ivrii, *Precise Spectral Asymptotics for Elliptic Operators Acting in Fiberings over Manifolds with Boundary*. Lecture Notes in Mathematics, vol. 1100 (Springer, New York, 1984)
62. V.Ja. Ivrii, *Microlocal Analysis and Precise Spectral Asymptotics* (Springer, Berlin, 1998)
63. G.W. Johnson, M.L. Lapidus, *The Feynman Integral and Feynman's Operational Calculus*. Oxford Science Publications, Oxford Mathematical Monographs (Oxford University Press, Oxford and New York, 2000). Corrected printing and paperback edition, 2002
64. G.W. Johnson, M.L. Lapidus, L. Nielsen, *Feynman's Operational Calculus and Beyond*: *Noncommutativity and Time-Ordering*. Oxford Science Publications, Oxford Mathematical Monographs (Oxford University Press, Oxford and New York, 2015); ISBN 978-0-19-870249-8. Approx. 400 pp.
65. M. Kac, Can one hear the shape of a drum? Amer. Math. Monthly (Slaught Memorial Papers, No. 11) **73**(4), 1–23 (1966)
66. A.A. Karatsuba, S.M. Voronin, *The Riemann Zeta-Function*. De Gruyter, Expositions in Mathematics (Walter de Gruyter, Berlin, 1992)
67. T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1995)
68. J. Kigami, *Analysis on Fractals* (Cambridge University Press, Cambridge, 2001)
69. J. Kigami, M.L. Lapidus, Weyl's problem for the spectral distribution of Laplacians on p.c.f. self-similar fractals. Commun. Math. Phys. **158**, 93–125 (1993)
70. J. Kigami, M.L. Lapidus, Self-similarity of volume measures for Laplacians on p.c.f. self-similar fractals. Commun. Math. Phys. **217**, 165–180 (2001)

71. P.T. Lai, Meilleures estimations asymptotiques des restes de la fonction spectrale et des valeurs propres relatifs au laplacien. Math. Scand. **48**, 5–38 (1981)
72. N. Lal, M.L. Lapidus, Hyperfunctions and spectral zeta functions of Laplacians on self-similar fractals. J. Phys. A Math. Theor. **45**, 365205, l4 pp. (2012). Also: e-print, `arXiv:12O2.4126v2[math-ph]`, 2012; IHES preprint, IHES/M/12/14, 2012
73. N. Lal, M.L. Lapidus, The decimation method for Laplacians on fractals: spectra and complex dynamics, in *Fractal Geometry and Dynamical Systems in Pure and Applied Mathematics II: Fractals in Applied Mathematics*, ed. by D. Carfi, M.L. Lapidus, E.P.J. Pearse, M. van Frankenhuijsen. Contemporary Mathematics, vol. 601 (American Mathematical Society, Providence, RI, 2013), pp. 227–249. Also: e-print, `arXiv:1302.4007v2[math-ph]`, 2014; IHES preprint, IHES/M/12/31, 2012
74. M.L. Lapidus, Fractal drum, inverse spectral problems for elliptic operators and a partial resolution of the Weyl–Berry conjecture. Trans. Amer. Math. Soc. **325**, 465–529 (1991)
75. M.L. Lapidus, Spectral and fractal geometry: from the Weyl–Berry conjecture for the vibrations of fractal drums to the Riemann zeta-function, in *Differential Equations and Mathematical Physics*, ed. by C. Bennewitz. Proceedings of Fourth UAB International Conference, Birmingham, March 1990 (Academic, New York, 1992), pp. 151–182
76. M.L. Lapidus, Vibrations of fractal drums, the Riemann hypothesis, waves in fractal media, and the Weyl–Berry conjecture, in *Ordinary and Partial Differential Equations*, ed. by B.D. Sleeman, R.J. Jarvis. Vol. IV, Proceedings of Twelfth International Conference (Dundee, Scotland, UK, June 1992), Pitman Research Notes in Math. Series, vol. 289 (Longman Scientific and Technical, London, 1993), pp. 126–209
77. M.L. Lapidus, Analysis on fractals, Laplacians on self-similar sets, noncommutative geometry and spectral dimensions. Topol. Methods Nonlinear Anal. **4**, 137–195 (1994). Special issue dedicated to Jean Leray
78. M.L. Lapidus, Towards a noncommutative fractal geometry? Laplacians and volume measures on fractals, in *Harmonic Analysis and Nonlinear Differential Equations: A Volume in Honor of Victor L. Shapiro*. Contemporary Mathematics, vol. 208 (American Mathematical Society, Providence, RI, 1997), pp. 211–252
79. M.L. Lapidus, *In Search of the Riemann Zeros: Strings, Fractal Membranes and Noncommutative Spacetimes* (American Mathematical Society, Providence, RI, 2008)
80. M.L. Lapidus, Towards quantized number theory: spectral operators and an asymmetric criterion for the Riemann hypothesis. Philos. Trans. Royal Soc. Ser. A No. 2047, **373**, 24 pp. (2015); doi:10.1098/rsta.2014.0240. Special issue titled "Geometric concepts in the foundations of physics". (Also: e-print, arXiv:1501.05362v2 [math-ph], 2015; IHES preprint, IHES/M/15/12, 2015.)
81. M.L. Lapidus, Riemann hypothesis, weighted Bergman spaces and quantized Riemann zeta function (tentative title) (2015, in preparation)
82. M.L. Lapidus, Quantized Weil conjectures, spectral operators and Pólya–Hilbert operators (tentative title) (2015, in preparation)
83. M.L. Lapidus, H. Lu, Nonarchimedean Cantor set and string. J. Fixed Point Theory Appl. **3**, 181–190 (2008). Special issue dedicated to the Jubilee of Vladimir I. Arnold, vol. I
84. M.L. Lapidus, H. Lu, Self-similar p-adic fractal strings and their complex dimensions. p-adic Numbers Ultrametric Anal. Appl. (Springer & Russian Academy of Sciences, Moscow) **1**(2), 167–180 (2009). Also: IHES preprint, IHES/M/08/42, 2008
85. M.L. Lapidus, H. Lu, The geometry of p-adic fractal strings: a comparative survey, in *Advances in Non-Archimedean Analysis*, ed. by J. Araujo, B. Diarra, A. Escassut. Proceedings of 11th International Conference on p-Adic Functional Analysis (Clermont-Ferrand, France, July 2010). Contemporary Mathematics, vol. 551 (American Mathematical Society, Providence, RI, 2011), pp. 163–206. Also: e-print, `arXiv:1105.2966v1 [math.MG]`, 2011
86. M. L. Lapidus, H. Maier, Hypothèse de Riemann, cordes fractales vibrantes et conjecture de Weyl–Berry modifiée. C. R. Acad. Sci. Paris Sér. I Math. **313**, 19–24 (1991)
87. M.L. Lapidus, H. Maier, The Riemann hypothesis and inverse spectral problems for fractal strings. J. Lond. Math. Soc. **52**(2), 15–34 (1995)

88. M.L. Lapidus, R. Nest, Fractal membranes as the second quantization of fractal strings. (preliminary) (2015, preprint)
89. M.L. Lapidus, M.M.H. Pang, Eigenfunctions of the Koch snowflake drum. Commun. Math. Phys. **172**, 359–376 (1995)
90. M.L. Lapidus, E.P.J. Pearse, A tube formula for the Koch snowflake curve, with applications to complex dimensions. J. Lond. Math. Soc. **74**(2), 397–414 (2006). Also: e-print, arXiv:math-ph/0412029v2, 2005
91. M.L. Lapidus, E.P.J. Pearse, Tube formulas for self-similar fractals, in *Analysis on Graphs and Its Applications*, ed. by P. Exner et al. Proceedings of Symposia in Pure Mathematics, vol. 77 (American Mathematical Society, Providence, RI, 2008), pp. 211–230. Also: e-print, arXiv:math.DS/0711.0173, 2007; IHES preprint, IHES/M/08/28, 2008
92. M.L. Lapidus, E.P.J. Pearse, Tube formulas and complex dimensions of self-similar tilings. Acta Appl. Math. **112**(1), 91–137 (2010). Springer Open Access: doi:10.1007/S10440-010-9562-x. Also: e-print, arXiv:math.DS/0605527v5, 2010; IHES preprint, IHES/M/08/27, 2008
93. M.L. Lapidus, C. Pomerance, Fonction zêta de Riemann et conjecture de Weyl–Berry pour les tambours fractals. C. R. Acad. Sci. Paris Sér. I Math. **310**, 343–348 (1990)
94. M.L. Lapidus, C. Pomerance, The Riemann zeta-function and the one-dimensional Weyl–Berry conjecture for fractal drums. Proc. Lond. Math. Soc. **66**(1), 41–69 (1993)
95. M.L. Lapidus, C. Pomerance, Counterexamples to the modified Weyl–Berry conjecture on fractal drums. Math. Proc. Camb. Philos. Soc. **119**, 167–178 (1996)
96. M.L. Lapidus, J.A. Rock, Towards zeta functions and complex dimensions of multifractals. Complex Variables Elliptic Equ. **54**(6), 545–560 (2009). Also: e-print, arXiv:math-ph/0810.0789, 2008
97. M.L. Lapidus, J.A. Rock, *An Invitation to Fractal Geometry*: *Dimension Theory, Zeta Functions and Applications* (2015, in preparation)
98. M.L. Lapidus, J.J. Sarhad, Dirac operators and geodesic metric on the harmonic Sierpinski gasket and other fractal sets. J. Noncommutative Geometry **8**(4), 947–985 (2014). doi:10.4171/JNCG/174. Also: e-print, arXiv:1212:0878v2 [math.MG], 2014; IHES preprint, IHES/M/12/32, 2012
99. M.L. Lapidus, M. van Frankenhuijsen, *Fractal Geometry and Number Theory: Complex Dimensions of Fractal Strings and Zeros of Zeta Functions* (Birkhäuser, Boston, 2000)
100. M.L. Lapidus, M. van Frankenhuijsen, *Fractal Geometry, Complex Dimensions and Zeta Functions: Geometry and Spectra of Fractal Strings*. Springer Monographs in Mathematics (Springer, New York, 2006)
101. M.L. Lapidus, M. van Frankenhuijsen, *Fractal Geometry, Complex Dimensions and Zeta Functions*: *Geometry and Spectra of Fractal Strings*, 2nd revised and enlarged edition (of the 2006 edn., [100]). Springer Monographs in Mathematics (Springer, New York, 2013)
102. M.L. Lapidus, J.W. Neuberger, R.J. Renka, C.A. Griffith, Snowflake harmonics and computer graphics: numerical computation of spectra on fractal domains. Int. J. Bifurcation Chaos **6**, 1185–1210 (1996)
103. M.L. Lapidus, J. Lévy-Véhel, J.A. Rock, Fractal strings and multifractal zeta functions. Lett. Math. Phys. **88**(1), 101–129 (2009). Springer Open Access, doi:10.1007/s1105-009-0302-y. Also: e-print, arXiv:math-ph/0610015v3, 2009
104. M.L. Lapidus, E.P.J. Pearse, S. Winter, Pointwise tube formulas for fractal sprays and self-similar tilings with arbitrary generators. Adv. Math. **227**, 1349–1398 (2011). Also: e-print, arXiv:1006.3807v3 [math.MG], 2011
105. M.L. Lapidus, E.P.J. Pearse, S. Winter, Minkowski measurability results for self-similar tilings and fractals with monophase generators, in *Fractal Geometry and Dynamical Systems in Pure and Applied Mathematics I: Fractals in Pure Mathematics*, ed. by D. Carfi, M.L. Lapidus, E.P.J. Pearse, M. van Frankenhuijsen. Contemporary Mathematics, vol. 600 (American Mathematical Society, Providence, RI, 2013), pp. 185–203. Also: e-print, arXiv:1104.1641v3 [math.MG], 2012; IHES preprint, IHES/M/12/33, 2012

106. M.L. Lapidus, J.A. Rock, D. Žubrinić, Box-counting fractal strings, zeta functions, and equivalent forms of Minkowski dimension, in *Fractal Geometry and Dynamical Systems in Pure and Applied Mathematics I: Fractals in Pure Mathematics*, ed. by D. Carfi, M.L. Lapidus, E.P.J. Pearse, M. van Frankenhuijsen. Contemporary Mathematics, vol. 600 (American Mathematical Society, Providence, RI, 2013), pp. 239–271. Also: e-print, arXiv:1207.6681v2 [math-ph], 2013; IHES preprint, IHES/M/12/22, 2012
107. M.L. Lapidus, H. Lu, M. van Frankenhuijsen, Minkowski measurability and exact fractal tube formulas for p-adic self-similar strings, in *Fractal Geometry and Dynamical Systems in Pure and Applied Mathematics I : Fractals in Pure Mathematics*, ed. by D. Carfi, M.L. Lapidus, E.P. J. Pearse, M. van Frankenhuijsen. Contemporary Mathematics, vol. 600 (American Mathematical Society, Providence, RI, 2013), pp. 185–203. Also: e-print, arXiv:1209.6440v1[math.MG], 2012; IHES preprint, IHES/M/12/23, 2012
108. M.L. Lapidus, G. Radunović, D. Žubrinić, Fractal tube formulas and a Minkowski measurability criterion for compact subsets of Euclidean spaces (2015). Also: e-print, arXiv:1411.5733v2[math-ph], 2015; IHES preprint, IHES/M/15/17, 2015
109. M.L. Lapidus, G. Radunović, D. Žubrinić, Fractal zeta functions and complex dimensions of relative fractal drums, survey article. J. Fixed Point Theory Appl. **15**(2), 321–378 (2014). Festschrift issue in honor of Haim Brezis' 70th birthday. doi:10.1007/s11784-014-0207-y. Also: e-print, arXiv:1407.8094v3[math-ph], 2014; IHES preprint, IHES/M/15/14, 2015
110. M.L. Lapidus, G. Radunović, D. Žubrinić, Distance and tube zeta functions of fractals and arbitrary compact sets (2015, preprint). Also: e-print, arXiv:1506.03525v2 [math-ph], 2015; IHES preprint, IHES/M/15/15, 2015
111. M.L. Lapidus, G. Radunović, D. Žubrinić, Complex dimensions of fractals and meromorphic extensions of fractal zeta functions (2015, preprint). Also: e-print, arXiv:1508.04784v1 [math-ph], 2015
112. M.L. Lapidus, G. Radunović, D. Žubrinić, Zeta functions and complex dimensions of relative fractal drums: theory, examples and applications (2015, preprint)
113. M.L. Lapidus, G. Radunović, D. Žubrinić, Fractal tube formulas for compact sets and relative fractal drums, with application to a Minkowski measurability criterion (2015, preprint)
114. M.L. Lapidus, H. Lu, M. van Frankenhuijsen, Minkowski dimension and explicit tube formulas for p-adic fractal strings (2015, preprint)
115. M.L. Lapidus, G. Radunović, D. Žubrinić, Fractal zeta functions and complex dimensions: a general higher-dimensional theory, survey article, in *Geometry and Stochastics V*, ed. by C. Bandt, K. Falconer, M. Zähle. Proceedings of Fifth International Conference (Tabarz, Germany, March 2014). Progress in Probability (Birkhäuser, Basel, 2015, pp. 229–257); doi:10.1007/978-3-319-18660-3_13. Based on a plenary lecture given by the first author at that conference. Also: e-print, arXiv:1502.00878v3[math.CV], 2015; IHES preprint, IHES/M/15/16, 2015
116. M.L. Lapidus, G. Radunović, D. Žubrinić, *Fractal Zeta Functions and Fractal Drums: Higher-Dimensional Theory of Complex Dimensions*. Research Monograph (Springer, New York, 2016, to appear). Approx. 625 pp.
117. A. Laurincikas, *Limit Theorems for the Riemann Zeta-Function* (Kluwer Academic Publishers, Dordrecht, 1996)
118. T. Lei (ed.), *The Mandelbrot Set, Theme and Variations*. London Mathematical Society Lecture Notes Series, vol. 274 (Cambridge University Press, Cambridge, 2000)
119. J. Lévy-Véhel, F. Mendivil, Multifractal and higher-dimensional zeta functions. Nonlinearity **24**(1), 259–276 (2011)
120. J.L. Lions, E. Magenes, *Non-homogeneous Boundary Value Problems and Applications*, vol. I, English transl. (Springer, Berlin, 1972)
121. B.B. Mandelbrot, *The Fractal Geometry of Nature*, revised and enlarged edition (of the 1977 edn.) (W. H. Freeman, New York, 1983)
122. P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability* (Cambridge University Press, Cambridge, 1995)
123. V.G. Maz'ja, *Sobolev Spaces* (Springer, Berlin, 1985)

124. R.B. Melrose, Weyl's conjecture for manifolds with concave boundary, in *Geometry of the Laplace Operator*. Proceedings of Symposia in Pure Mathematics, vol. 36 (American Mathematical Society, Providence, RI, 1980), pp. 254–274
125. R.B. Melrose, *The Trace of the Wave Group*. Contemporary Mathematics, vol. 27 (American Mathematical Society, Providence, RI, 1984), pp. 127–167
126. G. Métivier, Théorie spectrale d'opérateurs elliptiques sur des ouverts irréguliers, Séminaire Goulaic-Schwartz, No. 21 (Ecole Polytechnique, Paris, 1973)
127. G. Métivier, Etude asymptotique des valeurs propres et de la fonction spectrale de problèmes aux limites. Thèse de Doctorat d'Etat, Mathématiques, Université de Nice, 1976
128. G. Métivier, Valeurs propres de problèmes aux limites elliptiques irréguliers. Bull. Soc. Math. France Mém. **51–52**, 125–219 (1977)
129. S. Molchanov, B. Vainberg, On spectral asymptotics for domains with fractal boundaries. Commun. Math. Phys. **183**, 85–117 (1997)
130. G. Mora, J.M. Sepulcre, T. Vidal, On the existence of exponential polynomials with prefixed gaps. Bull. Lond. Math. Soc. **45**(6), 1148–1162 (2013)
131. A.N. Parshin, I.R. Shafarevich (eds.), *Number Theory, vol. II, Algebraic Number Fields*. Encyclopedia of Mathematical Sciences, vol. 62 (Springer, Berlin, 1992). Written by H. Koch.
132. A.N. Parshin, I.R. Shafarevich (eds.), *Number Theory, vol. I, Introduction to Number Theory*. Encyclopedia of Mathematical Sciences, vol. 49 (Springer, Berlin, 1995). Written by Yu. I. Manin and A. A. Panchishkin.
133. S.J. Patterson, *An Introduction to the Theory of the Riemann Zeta-Function* (Cambridge University Press, Cambridge, 1988)
134. E.P.J. Pearse, Canonical self-affine tilings by iterated function systems. Indiana Univ. Math. J. **56**(6), 3151–3169 (2007). Also: e-print, `arXiv:math.MG/0606111`, 2006
135. E.P.J. Pearse, S. Winter, Geometry of canonical self-similar tilings. Rocky Mountain J. Math. **42**, 1327–1357 (2012). Also: e-print, `arXiv:0811.2187`, 2009
136. Ch. Pommerenke, *Boundary Behavior of Conformal Maps* (Springer, New York, 1992)
137. A.G. Postnikov, *Tauberian Theory and its Applications*. Proceedings of the Steklov Institute of Mathematics (English transl., issue 2, 1980), vol. 144, 1979 (American Mathematical Society, Providence, RI, 1980)
138. C.R. Putnam, On the non-periodicity of the zeros of the Riemann zeta-function. Am. J. Math. **76**, 97–99 (1954)
139. C.R. Putnam, Remarks on periodic sequences and the Riemann zeta-function. Am. J. Math. **76**, 828–830 (1954)
140. G. Radunović, Fractal analysis of unbounded sets in Euclidean spaces and Lapidus zeta functions. Ph.D. thesis, University of Zagreb, Zagreb, Croatia, 2015
141. R. Rammal, Spectrum of harmonic excitations on fractals. J. Phys. **45**, 191–206 (1984)
142. R. Rammal, G. Toulouse, Random walks on fractal structures and percolation cluster. J. Phys. Lett. **44**, L13–L22 (1983)
143. J. Rataj, S. Winter, On volume and surface area of parallel sets. Indiana Univ. Math. J. **59**, 1661–1685 (2010)
144. J. Rataj, S. Winter, Characterization of Minkowski measurability in terms of surface area. J. Math. Anal. Appl. **400**, 120–132 (2013). Also: e-print, `arXiv:1111.1825v2 [math.CA]`, 2012
145. M. Reed, B. Simon, *Methods of Modern Mathematical Physics*. Fourier Analysis, Self-Adjointness, vol. II (Academic, New York, 1975)
146. M. Reed, B. Simon, *Methods of Modern Mathematical Physics*. Analysis of Operators, vol. IV (Academic, New York, 1979)
147. M. Reed, B. Simon, *Methods of Modern Mathematical Physics*. Functional Analysis, vol. I, revised and enlarged edition (of the 1975 edn.) (Academic, New York, 1980)
148. A. Reich, Universelle Wertevereteilung von Eulerprodukten. Nachr. Akad. Wiss. Göttingen Math.-Phys. **Kl. II**(1), 1–17 (1977)
149. A. Reich, Wertverteilung von Zetafunktionen. Arch. Math. **34**, 440–451 (1980)

150. B. Riemann, Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse. Monatsb. der Berliner Akad. pp. 671–680, 1858/1860. English transl. in [29], Appendix, pp. 229–305
151. W. Rudin, *Real and Complex Analysis*, 3rd edn. (McGraw-Hill, New York, 1987)
152. W. Rudin, *Functional Analysis*, 2nd edn. (of the 1973 edn.) (McGraw-Hill, New York, 1991)
153. C. Sabot, Integrated density of states of self-similar Sturm-Liouville operators and holomorphic dynamics in higher dimension. Ann. Inst. Henri Poincaré Probab. Stat. **37**, 275–311 (2001)
154. C. Sabot, Spectral properties of self-similar lattices and iteration of rational maps Mém. Soc. Math. Fr. (New Series) **92**, 1–104 (2003)
155. C. Sabot, Spectral analysis of a self-similar Sturm-Liouville operator. Indiana Univ. Math. J. **54**, 645–668 (2005)
156. B. Sapoval, Th. Gobron, A. Margolina, Vibrations of fractal drums. Phys. Rev. Lett. **67**, 2974–2977 (1991)
157. P. Sarnak, *L*-functions, in *Proceedings of International Congress of Mathematicians*, Berlin, 1998, ed. by G. Fischer, U. Rehmann, vol. I, pp. 453–465 (1998). Documenta Mathematica Journal DMV (Extra Volume ICM 98)
158. M. Schechter, *Operator Methods in Quantum Mechanics* (Dover Publications, Mineola, 2003)
159. M.R. Schroeder, *Fractal, Chaos, Power Laws: Minutes From an Infinite Paradise* (W. H. Freeman, New York, 1991)
160. L. Schwartz, *Méthodes Mathématiques pour les Sciences Physiques* (Hermann, Paris, 1961)
161. L. Schwartz, *Théorie des Distributions*, revised and enlarged edition (of the 1951 edn.) (Hermann, Paris, 1996)
162. R.T. Seeley, *Complex Powers of Elliptic Operators*. Proceedings of Symposia in Pure Mathematics, vol. 10 (American Mathematical Society, Providence, RI, 1967), pp. 288–307
163. R.T. Seeley, A sharp asymptotic remainder estimate for the eigenvalues of the Laplacian in a domain of $\mathbb{R}^3$. Adv. Math. **29**, 244–269 (1978)
164. R.T. Seeley, An estimate near the boundary for the spectral counting function of the Laplace operator. Amer. J. Math. **102**, 869–902 (1980)
165. J.-P. Serre, *A Course in Arithmetic*, English transl. (Springer, Berlin, 1973)
166. T. Shima, On eigenvalue problems for Laplacians on p.c.f. self-similar sets. Jpn. J. Ind. Appl. Math. **13**, 1–23 (1996)
167. M. Shishikura, The Hausdorff dimension of the boundary of the Mandelbrot set and Julia sets. Ann. Math. **147**, 225–267 (1998)
168. B. Simon, *Functional Integration and Quantum Physics* (Academic, New York, 1979)
169. L.L. Stachó, On the volume function of parallel sets. Acta Sci. Math. **38**, 365–374 (1976)
170. J. Steuding, Universality in the Selberg class, Special Activity in Analytic Number Theory and Diophantine Equations, in *Proceedings of a Workshop Held at the Max Planck-Institut in Bonn*, ed. by R.B. Heath-Brown, B. Moroz (2002). Bonner Math. Schriften **360**, 2003
171. J. Steuding, *Value-Distribution and L-Functions*. Lecture Notes in Mathematics, vol. 1877 (Springer, Berlin, 2007)
172. A. Teplyaev, Spectral zeta functions of symmetric fractals, in *Progress in Probability*, vol. 57 (Birkhäuser-Verlag, Basel, 2004), pp. 245–262
173. A. Teplyaev, Spectral zeta functions of fractals and the complex dynamics of polynomials. Trans. Amer. Math. Soc. **359**, 4339–4358 (2007). Also: e-print, `arXiv:math.SP/0505546`, 2005
174. E.C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd edn. (revised by D.R. Heath-Brown) (Oxford University Press, Oxford, 1986)
175. C. Tricot, *Curves and Fractal Dimension* (Springer, Berlin, 1995)
176. M. van den Berg, P.B. Gilkey, A comparison estimate for the heat equation with an application to the heat content of the *s*-adic von Koch snowflake. Bull. Lond. Math. Soc. **30**(4), 404–412 (1998)
177. S.M. Voronin, The distribution of the non-zero values of the Riemann zeta function. Izv. Akad. Nauk. Inst. Steklov **128**, 131–150 (1972) (Russian)

178. S.M. Voronin, Theorem on the 'universality' of the Riemann zeta-function. Izv. Akad. Nauk. SSSR Ser. Matem. **39**, 475–486 (1975) (Russian). Math. USSR Izv. **9** (1975), 443–445
179. H. Weyl, *Hermann Weyl: Gesammelte Abhandlungen* (Collected Works) (Springer, Berlin/New York, 1968)
180. H. Weyl, Über die Abhängigkeit der Eigenschwingungen einer Membran von deren Begrenzung. J. Reine Angew. Math. **141**, 1–11 (1912). Reprinted in [179, vol. I, pp. 431–441]
181. H. Weyl, Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen. Math. Ann. **71**, 441–479 (1912). Reprinted in [179, vol. I, pp. 393–430]
182. K.G. Wilson, Renormalization group and critical phenomena, I & II. Phys. Rev. **B4**, 3174–3183 & 3184–3205 (1971)
183. D. Žubrinić, Minkowski content and singular integrals. Chaos Solitons Fractals **17**(1), 169–177 (2003)
184. D. Žubrinić, Analysis of Minkowski contents of fractal sets and applications. Real Anal. Exch. **31**(2), 315–354 (2005/2006)

Sums of Two Squares in Short Intervals

James Maynard

Dedicated to Helmut Maier on the occasion of his 60th birthday

Abstract We show that there are short intervals $[x, x+y]$ containing $\gg y^{1/10}$ numbers expressible as the sum of two squares, which is many more than the average when $y = o((\log x)^{5/9})$. We obtain similar results for sums of two squares in short arithmetic progressions.

1 Introduction

Let X be large and $y \in (0, X]$. The prime number theorem shows that *on average* for $x \in [X, 2X]$ we have

$$\pi(x+y) - \pi(x) \approx \frac{y}{\log x}. \tag{1}$$

The approximation (1) is known ([10], improving on [12]) to hold in the sense of asymptotic equivalence for *all* $x \in [X, 2X]$ provided that $X^{7/12} \leq y \leq X$, whilst the celebrated result of Maier [13] shows that (1) does *not* hold in the sense of asymptotic equivalence for all $x \in [X, 2X]$ when y is of size $(\log X)^A$, for any fixed $A > 0$. He showed that for $y = (\log X)^A$, there exists $x \in [X, 2X]$ such that

$$\pi(x+y) - \pi(x) > (1+\delta_A)\frac{y}{\log x} \tag{2}$$

for a positive constant δ_A depending only on A (and he also obtained similar results for intervals containing fewer than the average number of primes).

J. Maynard (✉)
Magdalen College, Oxford OX1 4AU, UK
e-mail: james.maynard@magd.ox.ac.uk

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_15

It is trivial that in the range $y \ll \log X$, (1) cannot hold in the sense of asymptotic equivalence, and the extent to which this approximation fails is closely related to the study of gaps between primes. Maier and Stewart [15] combined the Erdős–Rankin construction for large gaps between primes with the work of Maier to get show the existence of intervals containing significantly fewer primes than the average. It follows from the recent advances in the study of small gaps between primes ([16, 18, 21] and unpublished work of Tao) that there are also short intervals containing significantly more primes than the average.

The situation is similar for the set $\mathscr{S}$ of integers representable by the sum of two squares. Hooley [11] has shown that for almost all $x \in [X, 2X]$ one has

$$\frac{\mathfrak{S}y}{\sqrt{\log X}} \ll \sum_{n \in \mathscr{S} \cap [x, x+y]} 1 \ll \frac{\mathfrak{S}y}{\sqrt{\log X}}, \tag{3}$$

provided $y/\sqrt{\log X} \to \infty$ (here $\mathfrak{S} > 0$ is an absolute constant so that the above sum is $(1 + o(1))\mathfrak{S}y/\sqrt{\log X}$ on average over $x \in [X, 2X]$).

By adapting Maier's method, Balog and Wooley [1] showed that for any fixed $A > 0$ and $y = (\log X)^A$, there exists $\delta'_A > 0$ and an $x \in [X, 2X]$ such that

$$\sum_{n \in \mathscr{S} \cap [x, x+y]} 1 \geq (1 + \delta'_A) \frac{\mathfrak{S}y}{\sqrt{\log X}}, \tag{4}$$

and so Hooley's result cannot be strengthened to an asymptotic which holds for all $x \in [X, 2X]$ when y is of size $(\log x)^A$.

As a result of the "GPY sieve", Graham, Goldston, Pintz and Yıldırım [6] have shown that for any fixed m there exists $x \in [X, 2X]$ such that $[x, x + O_m(1)]$ contains m numbers which are the product of two primes $\equiv 1 \pmod 4$, and so m elements of $\mathscr{S}$. By modifying the GPY sieve to find short intervals containing many elements of $\mathscr{S}$, we show the existence of intervals $[x, x + y]$ containing a higher order of magnitude than the average number when y is of size $o((\log X)^{5/9})$. Thus in this range we see that the upper bound in Hooley's result cannot hold for all $x \in [X, 2X]$. (It is a classical result that the lower bound cannot hold for all x in such a range; the strongest such result is due to Richards [20] who has shown there exists $x \in [X, 2X]$ such that there are no integers $n \in \mathscr{S}$ in an interval $[x, x + O(\log x)]$.)

2 Intervals with Many Primes or Sums of Two Squares

Let the function $g : (0, \infty) \to \mathbb{R}$ be defined by

$$g(t) = \sup_{u \geq t} e^{\gamma} \omega(u), \tag{5}$$

where γ is the Euler constant and $\omega(u)$ is the Buchstab function.[1] Clearly g is decreasing, and it is known that $g(t) > 1$ for all $t > 0$.

We first recall the current state of knowledge regarding short intervals and short arithmetic progressions containing unusually many primes

Theorem 2.1. *Fix $\epsilon > 0$. There is a constant $c_0(\epsilon) > 0$ such that we have the following:*

1. *Let X, y satisfy $c_0(\epsilon) \le y \le X$. Then there exist at least $X^{1-1/\log\log X}$ values $x \in [X, 2X]$ such that*

$$\pi(x+y) - \pi(x) \ge \Big(g\Big(\frac{\log y}{\log\log x}\Big) - \epsilon + c\frac{(\log x)(\log y)}{y}\Big)\frac{y}{\log x}.$$

2. *Fix $a \in \mathbb{N}$, and let Q, x satisfy $c_0(\epsilon) \le Q \le x/2$. Then there exist at least $Q^{1-1/\log\log Q}$ values $q \in [Q, 2Q]$ such that*

$$\pi(x; q, a) \ge \Big(g\Big(\frac{\log x/q}{\log\log x}\Big) - \epsilon + c\frac{(\log x)(\log x/q)}{x/q}\Big)\frac{x}{\phi(q)\log x}.$$

3. *Let q, x satisfy $c_0(\epsilon) \le q \le x/2$, and let q have no prime factors less than $\log x/\log\log x$. Then there exist at least $q^{1-1/\log\log q}$ integers $a \in [1, q]$ such that*

$$\pi(x; q, a) \ge \Big(g\Big(\frac{\log x/q}{\log\log x}\Big) - \epsilon + c\frac{(\log x)(\log x/q)}{x/q}\Big)\frac{x}{\phi(q)\log x}.$$

Here $c > 0$ is an absolute constant.

We do not claim Theorem 2.1 is new; all the statements of Theorem 2.1 follow immediately from the work [5, 13] and [16].

When the number of integers in the short interval or arithmetic progression is large compared with $\log x \log\log x$, the first terms in parentheses dominate the right-hand sides, and we obtain the results of Maier [13] and Friedlander and Granville [5] that there are many intervals and arithmetic progressions which contain more than the average number of primes by a constant factor. When the number of integers is small compared with $\log x \log\log x$, the final terms dominate, and we see that there are many intervals and arithmetic progressions which contain a number of primes which is a higher order of magnitude than the average number.

Part 3 of Theorem 2.1 is one of the key ingredients in recent work of the author on large gaps between primes [17]. The previous best lower bound for the largest gaps between consecutive primes bounded by x was due to Pintz [19] who explicitly indicated that one obstruction to improving his work was the inability of previous techniques to show the existence of many primes in such a short arithmetic

[1] The Buchstab function $\omega(u)$ is defined by the delay-differential equation

$$\omega(u) = u^{-1} \quad (0 < u \le 2), \qquad \frac{\partial}{\partial u}(u\omega(u)) = \omega(u-1) \quad (u \ge 2).$$

progression with prime modulus (this issue was also clear in the key work of Maier and Pomerance [14] on this problem). The independent improvement on this bound by Ford, Green, Konyagin and Tao [3] used the Green–Tao technology to show there are many primes q for which there were many primes in an arithmetic progression modulo q, although it appears this approach does not give improvements to Theorem 2.1.

The aim of this paper is to establish an analogous statement to Theorem 2.1 for the set $\mathscr{S}$ of numbers representable as the sum of two squares. We first require some notation. We define the counting functions $S(x)$ and $S(x;q,a)$, the constant $\mathfrak{S}$ and the multiplicative function $\phi_{\mathscr{S}}$ by

$$S(x) = \#\{n \le x : n \in \mathscr{S}\}, \tag{6}$$

$$S(x;q,a) = \#\{n \le x : n \in \mathscr{S}, n \equiv a \pmod q\}, \tag{7}$$

$$\mathfrak{S} = \frac{1}{\sqrt{2}} \prod_{p \equiv 3 \pmod 4} \Big(1 - \frac{1}{p^2}\Big)^{-1/2}, \tag{8}$$

$$\phi_{\mathscr{S}}(p^e) = \begin{cases} p^e, & p \equiv 1 \pmod 4, \\ p^{e+1}/(p+1), & p \equiv 3 \pmod 4, \\ 2^{e-1}, & p = 2 \text{ and } e \ge 2, \\ 2, & p^e = 2. \end{cases} \tag{9}$$

With this notation, the average number $S(x+y) - S(x)$ of elements of $\mathscr{S}$ in a short interval $[x, x+y]$ is $\sim \mathfrak{S} y/\sqrt{\log x}$ and the average number $S(x;q,a)$ of elements of $\mathscr{S}$ in the arithmetic progression $a \pmod q$ for $(a,q) = 1$ and $a \equiv 1 \pmod{(4,q)}$ which are less than x is $\sim \mathfrak{S} x/(\phi_{\mathscr{S}}(q)\sqrt{\log x})$.

Finally, we let F be the sieve function occurring in the upper bound half dimensional sieve, that is the function defined by the delay-differential equations

$$F(s) = \frac{2e^{\gamma/2}}{\pi^{1/2} s^{1/2}} \text{ for } 0 \le s \le 2, \qquad (s^{1/2}F(s))' = \frac{1}{2} s^{-1/2} f(s-1) \text{ for } s > 2,$$

$$f(s) = 0 \text{ for } 0 \le s \le 1, \qquad (s^{1/2}f(s))' = \frac{1}{2} s^{-1/2} F(s-1) \text{ for } s > 1. \tag{10}$$

It is known that $1 < F(s) < 1 + O(e^{-s})$ for all $s > 0$.

Theorem 2.2. *Fix $\epsilon > 0$, and let $\mathscr{S}$ denote the set of integers representable as the sum of two squares. There is a constant $c_0(\epsilon) > 0$ such that we have the following:*

1. *Let X, y satisfy $c_0(\epsilon) \le y \le X$. Then there exists at least $X^{1-1/\log\log X}$ values $x \in [X, 2X]$ such that*

$$S(x+y) - S(x) \ge \Big(F\Big(\frac{\log y}{\log\log x}\Big) - \epsilon + c\frac{(\log x)^{1/2}}{y^{9/10}}\Big)\frac{\mathfrak{S} y}{(\log x)^{1/2}}.$$

2. *Fix $a \in \mathscr{S}$, and let Q, x satisfy $c_0(\epsilon) \leq Q \leq x/2$. Then there exists at least $Q^{1-1/\log\log Q}$ values $q \in [Q, 2Q]$ such that*

$$S(x; q, a) \geq \Big(F\Big(\frac{\log x/q}{\log\log x}\Big) - \epsilon + c\frac{(\log x)^{1/2}}{(x/q)^{9/10}}\Big)\frac{\mathfrak{S}x}{\phi_{\mathscr{S}}(q)(\log x)^{1/2}}.$$

3. *Let q, x satisfy $c_0(\epsilon) \leq q \leq x/2$, and let q have no prime factors less than $\log x/\log\log x$. Then there exists at least $q^{1-1/\log\log q}$ integers $a \in [1, q]$ such that*

$$S(x; q, a) \geq \Big(F\Big(\frac{\log x/q}{\log\log x}\Big) - \epsilon + c\frac{(\log x)^{1/2}}{(x/q)^{9/10}}\Big)\frac{\mathfrak{S}x}{\phi_{\mathscr{S}}(q)(\log x)^{1/2}}.$$

Here $c > 0$ is an absolute constant.

Theorem 2.2 is our equivalent of Theorem 2.1 for numbers expressible as the sum of two squares. When the number of integers in the short interval or arithmetic progression is large compared with $(\log x)^{5/9}$, we find that there more than the average number of integers representable as the sum of two square, which is a result of Balog and Wooley [1], based on Maier's method. When the number of integers is small compared with $(\log x)^{5/9}$, then the number of integers representable as the sum of two squares is of a higher order of magnitude than the average. We note that unlike Theorem 2.1, Theorem 2.2 improves on the Maier matrix bounds in a range beyond the trivial region by a full positive power of $\log x$.

The constant $9/10$ appearing in the exponent in the denominators of the final terms of Theorem 2.2 is certainly not optimal; with slightly more effort, one could certainly improve this. We satisfy ourselves with a weaker bound for simplicity here, the key point of interest being one can obtain a constant less than 1, and so there are intervals of length considerably larger than the average gap between elements of $\mathscr{S}$ which contain a higher order of magnitude than the average number of elements.

Analogously to the work of Granville and Soundararajan [8], we believe the phenomenon of significant failures of equidistribution in short intervals should hold in rather more general "arithmetic sequences" which are strongly equidistributed in arithmetic progressions (in the sense of a weak Bombieri–Vinogradov type theorem). We hope to return to this in a future paper.

3 Overview

We give a rough overview of the methods, emphasizing the similarities between Maier's matrix method and the GPY method. For simplicity we concentrate on the case when we are looking for primes in short intervals; the arguments for arithmetic progressions are similar, and when looking for sums of two squares one wishes to "sieve out" only primes congruent to 3 (mod 4).

We estimate a weighted average of the number of primes in a short interval, and wish to show that this weighted average is larger than what one would obtain if the primes were evenly distributed. More specifically, we consider the ratio

$$\Big(\sum_{X<n<2X} \#\big\{p \text{ prime} \in [n, n+h]\big\} w_n\Big) \Big/ \Big(\sum_{X<n<2X} \frac{h}{\log X} w_n\Big), \tag{11}$$

where w_n are some non-negative weights chosen to be large when the interval $[n, n+h]$ contains many numbers with no small prime factors, and small otherwise.

If the ratio (11) is greater than some constant c_0, then it is clear that there must be at least one $n \in [X, 2X]$ such that $[n, n+h]$ contains at least $c_0 h/\log x$ primes (i.e. c_0 times the average number). Thus we aim to choose w_n to make this ratio as large as possible.

In the range $h = o(\log X)$, we choose the weights w_n to be concentrated on integers n such that the interval $[n, n+h]$ *only*[2] contains numbers with no small prime factors.[3] To achieve this we choose w_n to mimic weights of Selberg's Λ^2-sieve

$$w_n \approx \Big(\sum_{\substack{d_1,\dots,d_h \\ d_i \mid n+i}} \lambda_{d_1,\dots,d_h}\Big)^2, \tag{12}$$

where $\lambda_{d_1,\dots,d_h}$ are real numbers chosen later to optimize the final result.

In the range $h \approx (\log x)^\lambda$ for some constant $\lambda > 1$ it is unavoidable that a large number of elements of $[n, n+h]$ will have small prime factors, and so we require a slightly different choice of the weights w_n. Instead we observe that if n is a multiple of primes less than $z = (\log x)^{\lambda'}$ (where $\lambda' < 1$), then the probability that a random element of $[n, n+h]$ is coprime to all these primes fluctuates depending on the size of λ/λ', and can be more than the average for general n. More specifically, we take

$$w_n \approx \begin{cases} 1, & n \equiv 0 \pmod{\prod_{p<z} p}, \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

The number of elements of $[n, n+h]$ which have no prime factors less than z is given by

$$\#\{j \in [1,h] : (j,p) = 1 \forall p < z\} \sim e^\gamma \omega\Big(\frac{\log h}{\log z}\Big) h \prod_{p<z} \frac{p-1}{p} \tag{14}$$

[2]We ignore the elements of $[n, n+h]$ which are a multiple of some prime $p \le h$, which is unavoidable.

[3]Here a "small prime" refers to one bounded by roughly $\exp(h^{-1+o(1)} \log x)$.

for $z \to \infty$ and $z < h^{1-\epsilon}$ (here ω is the Buchstab function). Thus, given h we can choose z so that the number of elements coprime to small primes is larger than the expected number in a random interval by a factor $e^{\gamma}\omega(\log h/\log z) > 1$.

Given this choice of weights, to estimate the numerator in (11) we rewrite the count of the primes as a sum of the indicator function of the primes, and swap the order of summation. This gives

$$\sum_{j=1}^{h} \sum_{X<n\leq 2X} \mathbf{1}_{\mathbb{P}}(n+j)w_n. \tag{15}$$

(Here, and throughout this paper, $\mathbf{1}_{\mathscr{A}}$ will denote the indicator function of a set $\mathscr{A}$, in this case the set $\mathbb{P}$ of primes.) Our choice of w_n means that we can estimate this expression using knowledge of the distribution of primes in arithmetic progressions.[4] The inner sum vanishes or has a simple asymptotic expression depending on the arithmetic structure of j. In the case of the Maier weights, this leads us to (14), which gives our asymptotic for the numerator, and hence the ratio (11). In the range $h = o(\log X)$, we obtain a quadratic form in the coefficients $\lambda_{d_1,\dots,d_h}$ for the numerator, and similarly a quadratic form for the denominator. We then make a choice of these $\lambda_{d_1,\dots,d_h}$ which (approximately) maximizes the ratio of these quadratic forms.

Remark 3.1. One can heuristically investigate Selberg-sieve weights (12) for larger h, and (assuming various error terms are negligible) conclude that an approximately optimal choice of weights $\lambda_{d_1,\dots,d_k}$ should be given by $\lambda_{d_1,\dots,d_k} \approx \mu(d_1 \dots d_k)$ if $\prod_{i=1}^{k} d_i$ has all of its prime factors less than a small multiple of $\log x$ (and 0 otherwise). This corresponds precisely to Maier's choice of weights (13).

4 Admissible Sets of Linear Functions

The bounds involving the sieve function F in Theorem 2.2 follow from the argument of Balog and Wooley [1], and by making minor adaptions analogous to the work of Friedlander–Granville [5]. We sketch such arguments in Sect. 5.

Therefore the main task of this paper is to establish the bounds coming from the final terms in parentheses of Theorem 2.2. We will do this by an adaptation of the GPY sieve method. Indeed, we will actually prove a rather stronger result, given by Theorem 4.3 below, which is roughly analogous to the main theorem of Maynard [16].

[4] One needs to be slightly careful about the possible effect of Siegel zeros here, but this is a minor technical issue.

To ease notation, we let $\langle P_1\rangle$ denote the set of integers composed only of primes congruent to 1 (mod 4). Similarly let $\langle P_3\rangle$ those composed of primes congruent to 3 (mod 4).

Definition 4.1 ($\langle P_3\rangle$-Admissibility). We say a set $\mathscr{L} = \{L_1, \dots, L_k\}$ of distinct linear functions $L_i(n) = a_i n + b_i$ with integer coefficients is *$\langle P_3\rangle$-admissible* if for every prime $p \equiv 3 \pmod 4$ there is an integer n_p such that $(\prod_{i=1}^{k} L_i(n_p), p) = 1$, and if $a_i, b_i > 0$ for all i.

Theorem 4.3. *Let x be sufficiently large and $k \le (\log x)^{1/5}$. Then there is a prime $p_0 \gg \log\log x$ such that the following holds.*

Let $\mathscr{L} = \{L_1, \dots, L_k\}$ be a $\langle P_3\rangle$-admissible set of linear functions such that the coefficients of $L_i(n) = a_i n + b_i$ satisfy $0 < a_i \le (\log x)^{1/3}$, $(2p_0, a_i) = 1$, and $0 < b_i < x$. There exists an absolute constant $C > 0$ such that

$$\#\{n \in [x, 2x] : \text{at least } k^{1/2}/C \text{ of } L_1(n), \dots, L_k(n) \text{ are in } \mathscr{S}\} \ge x\exp(-(\log x)^{1/2}).$$

We remark that the conditions $k \le (\log X)^{1/5}$, $2 \nmid a_i$, $a_i < (\log X)^{1/3}$ and $b_i \le x$ are considerably weaker than what our method requires, but simplify some of the later arguments slightly and are sufficient for our application.

In the region where y, x/Q and x/q are small compared with $(\log x)^{5/9}$, the bounds in Theorem 2.2 follow easily from Theorem 4.3. For sums of squares in short intervals, we consider $y \le (\log x)^{5/9+\epsilon}$. We take $k = y^{1/5}$, and let $\mathscr{L} = \{L_1, \dots, L_k\}$ be the $\langle P_3\rangle$-admissible set of linear functions $L_i(n) = n + h_i$, where $h_1, \dots, h_k$ are the first k positive integers in $\langle P_1\rangle$ which are coprime with p_0. (This is $\langle P_3\rangle$-admissible, since $L_i(0)$ is coprime to all primes $p \equiv 3 \pmod 4$ for all i.) We note that $h_k \ll k(\log k)^{1/2} \le y$. It then follows immediately from Theorem 4.3 that there are $\gg X\exp(-(\log X)^{1/2})$ values of $n \in [X, 2X]$ such that $\gg k^{1/2} = y^{1/10}$ of the $L_i(n)$ are in $\mathscr{S}$, and hence $\gg X^{1-1/\log\log X}$ values of $x \in [X, 2X]$ such that $[x, x + y]$ contains at least $y^{1/10}$ elements of $\mathscr{S}$.

Parts 2 and 3 of Theorem 2.2 follow analogously, defining $k = (x/Q)^{1/5}$ and $L_i(n) = a + h_i n$ when $Q > x(\log x)^{-5/9-\epsilon}$ for part 2, and defining $k = (x/q)^{1/10}$ and $L_i(n) = n + h_i q$ when $q > x(\log x)^{-5/9-\epsilon}$ for part 3. We note that it follows from Theorem 2.2 that actually there are many intervals $[x, x + y]$ containing $\gg y^{1/2}(\log y)^{-1/4}$ elements of $\mathscr{S}$ (rather than $\gg y^{1/10}$), for the smaller range $1 \le y \le (\log x)^{1/5}$.

5 Irregularities from the Maier Matrix Method

As previously mentioned, the bounds in Theorem 2.2 when y, x/Q or x/q large compared with $(\log x)^{5/9}$ follow from minor adaptions of the argument of Balog and Wooley [1]. In particular, the relevant bound for part 2 of Theorem 2.2 follows from the proof of [1, Theorem 1]. We give brief outline of the argument for parts 2 and 3, leaving the details to the interested reader.

We first consider part 2 of Theorem 2.2. We assume that $(\log x)^{5/9-\epsilon} \le x/Q \le (\log X)^{O(1)}$ since otherwise either the result follows from Theorem 4.3, or the result is trivial. For simplicity we shall assume that a is odd; the case for even a is entirely analogous. We let $P = \prod_{p\equiv 3 \pmod 4, p\le z} p^{\alpha_p}$, where α_p is the least odd integer such that $p^{\alpha_p} \ge a(4x/Q+1)$ and $z = \log x/(\log\log x)^2$. We note that $P = Q^{o(1/\log\log Q)}$. Finally, we let $\mathscr{Q} = \{q \in [(1-\delta)Q + 4a, Q + 4a] : q \equiv 4a \pmod{4P}\}$ for some small fixed constant $\delta = \delta(\epsilon) > 0$. We have

$$\sum_{q\in\mathscr{Q}} S(x;q,a) \ge \sum_{r\le x/Q-1} \sum_{q\in\mathscr{Q}} \mathbf{1}_{\mathscr{S}}(a+rq)$$
$$= \sum_{r\le x/Q-1} \sum_{m\in[(1-\delta)Q/4P, Q/4P]} \mathbf{1}_{\mathscr{S}}(a(4r+1)+4mrP). \tag{16}$$

The inner sum is counting elements of $\mathscr{S}$ in an arithmetic progression to modulus $4rP$. We note that by construction of P we must have $(P,a) = e^2$ for some $e \in \langle P_3\rangle$ since $a \in \mathscr{S}$, and similarly if the inner sum is non-empty we must have $(P, a(4r+1)) = e^2d^2$ since $a(4r+1)+4mrP \in \mathscr{S}$. Moreover $P/(e^2d^2)$ and P are composed of the same prime factors, since each prime occurs in P with odd multiplicity. Thus, if r is such that $(P, 4r+1) = d^2$, the inner sum is

$$(1+o(1))\frac{\mathfrak{S}\delta rQ/(e^2d^2)}{\phi_{\mathscr{S}}(4rP/(d^2e^2))\sqrt{\log rQ/d^2e^2}} \sim \frac{\mathfrak{S}\delta Q}{2\phi_{\mathscr{S}}(P)\sqrt{\log x}}. \tag{17}$$

Therefore, letting $ud^2 = 4r+1$, we obtain

$$\sum_{q\in\mathscr{Q}} S(x;q,a) \ge \frac{(1+o(1))\mathfrak{S}\delta Q}{2\phi_{\mathscr{S}}(P)\sqrt{\log x}} \sum_{d^2|P} \sum_{\substack{u<(4x/Q+1)/d^2 \\ u\equiv 1 \pmod 4 \\ (u,P)=1}} 1. \tag{18}$$

This double sum is exactly the sum $\mathscr{R}^+$ which is estimated in [1, Proof of Lemma 4.3]. In particular, they show that

$$\sum_{d^2|P} \sum_{\substack{u<(4x/Q+1)/d^2 \\ u\equiv 1 \pmod 4 \\ (u,P)=1}} 1 \sim \frac{\phi_{\mathscr{S}}(P)}{P} F\Big(\frac{\log x/Q}{\log z}\Big). \tag{19}$$

Recalling that $z = \log x/(\log\log x)^2$, one arrives at (for δ sufficiently small)

$$\sum_{q\in\mathscr{Q}} \left(S(x;q,a) - \Big(F\Big(\frac{\log x/q}{\log\log x}\Big) - \epsilon\Big)\frac{\mathfrak{S}x}{\phi_{\mathscr{S}}(q)\sqrt{\log x}}\right) \gg \frac{x}{P\sqrt{\log x}}. \tag{20}$$

This implies the relevant bound in part 2 of Theorem 2.2.

We now consider part 3 of Theorem 2.2. We let $\tilde{P} = \prod_{p\equiv 3 \pmod 4, p\le z} p^{\beta_p}$ where β_p is the least odd integer such that $p^{\beta_p} \ge 4x/q$, and let $\mathscr{A} = \{a \in [(1-\delta)q, q] : a \equiv a_0 \pmod P\}$. We see that

$$\begin{aligned}\sum_{a\in\mathscr{A}} S(x;q,a) &\ge \sum_{r\le (x-q)/q} \sum_{a\in\mathscr{A}} \mathbf{1}_{\mathscr{S}}(a+rq)\\ &= \sum_{r\le x/q-1} \sum_{m\in[((1-\delta)q-a_0)/P,(q-a_0)/P]} \mathbf{1}_{\mathscr{S}}(a_0+rq+mP). \end{aligned} \tag{21}$$

We see the inner sum counts elements of $\mathscr{S}$ in an arithmetic progression modulo P. We choose $a_0 \equiv q/4 \pmod P$, so that $(a_0+rq, P) = (4r+1, P)$ since, by assumption, q has no prime factors in common with P. Thus $(4r+1,P) = d^2$ for some $d \in \langle P_3\rangle$ by the same argument as above, and we obtain

$$\sum_{a\in\mathscr{A}} S(x;q,a) \ge \frac{(1+o(1))\mathfrak{S}\delta q}{\phi_{\mathscr{S}}(P)\sqrt{\log x}} \sum_{d^2\mid P} \sum_{\substack{u<(4x/q-3)/d^2\\ u\equiv 1 \pmod 4\\ (u,P)=1}} 1. \tag{22}$$

Again, using (19), we obtain

$$\sum_{a\in\mathscr{A}} \left(S(x;q,a) - \left(F\Big(\frac{\log x/q}{\log\log x}\Big) - \epsilon\right)\frac{\mathfrak{S}x}{\phi_{\mathscr{S}}(q)\sqrt{\log x}}\right) \gg \frac{x}{P\sqrt{\log x}}, \tag{23}$$

which implies the relevant bound for part 3.

6 Setup for Theorem 4.3

The improvements of the GPY sieve in the author's work [16, 18] are not significant for the application of finding sums of two squares in short intervals (the improvement would be $y^{o(1)}$ in the final term for part 1 of Theorem 2.2, for example), and so we will use a uniform version of the simplest GPY sieve.

In order to get a result that applies in the larger range, however, it is necessary to modify the sieve to the application; a uniform version of the argument in [6] (i.e. counting numbers with two prime factors both congruent to 1 (mod 4)) would only give non-trivial results in the region $y \ll (\log x)^{1/2}\log\log x$, for example.

We follow a similar argument to [6] (and attempt to keep the notation similar), but make modifications to allow for uniformity (similar to those in [7] and [16], although it is simpler in this context) and to specialize so as to remove only primes congruent to 3 (mod 4) (so we will apply sieves of "dimension" $k/2$ and $(k+1)/2$ instead of dimension k and $k+1$).

In order to state our setup, we require the following lemma.

Lemma 6.1. *Let $x > 10$ and $\epsilon > 0$. There exists a prime $p_0 \in [(\log\log x)/2, x]$ such that*

$$\sum_{\substack{q<x^{1/2-\epsilon}\\(q,p_0)=1}} \sup_{\substack{(a,q)=1\\x'\le x}} \left|\pi(x';q,a)-\frac{\pi(x')}{\phi(q)}\right| \ll_\epsilon x\exp(-c_0\sqrt{\log x}),$$

for some absolute constant $c_0 > 0$.

Proof. This follows from known results on the repulsion of zeros of L-functions. Following [2, Chap. 28], we have (for a suitable constant $c > 0$)

$$\sum_{\substack{q<x^{1/2-\epsilon}\\(q,p_0)=1}} \sup_{\substack{(a,q)=1\\x'\le x}} \left|\pi(x';q,a)-\frac{\pi(x')}{\phi(q)}\right| \ll x\exp(-c\sqrt{\log x}) + \log x \sum_{\substack{q<\exp(2c\sqrt{\log x})\\(q,p_0)=1}} \sup_{x'\le x} \sideset{}{^*}\sum_{\chi \pmod q} \frac{|\psi'(x',\chi)|}{\phi(q)}. \tag{24}$$

However, by the Landau–Page theorem [2, Chap. 20], we have that if c is sufficiently small then

$$\phi(q)^{-1}\sideset{}{^*}\sum_{\chi}|\psi'(x',\chi)| \ll x\exp(-3c\sqrt{\log x}) \tag{25}$$

for all $q < \exp(2c\sqrt{\log x})$ and $x' \le x$, except possibly those which are a multiple of an exceptional modulus q_1. Such an exceptional modulus must satisfy $\log x \ll q_1(\log q_1)^4 \ll x$, and must be square-free apart from a factor of at most 4. Taking p_0 to be the largest prime factor of q_1 then gives the result. □

Let $\mathscr{L} = \{L_1, \dots, L_k\}$ be the $\langle P_3\rangle$-admissible set of Theorem 4.3, p_0 be the prime given by Lemma 6.1, and let

$$W = \prod_{\substack{p\le 2(\log X)^{1/3}\\p\equiv 3 \pmod 4\\p\ne p_0}} p. \tag{26}$$

It will be convenient to choose w_n to be 0 unless $n \equiv v_0 \pmod W$, where v_0 is chosen such that $(L_j(v_0), W) = 1$ for each j. (Such a v_0 exists by the $\langle P_3\rangle$-admissibility of $\{L_1, \dots, L_k\}$ and the Chinese Remainder Theorem.) This allows us to ignore n for which one of the $L_i(n)$ has a small prime factor congruent to 3 (mod 4).

We then define our weights w_n by

$$w_n = \begin{cases} \left(\sum_{d|\prod_{i=1}^k L_i(n)} \lambda_d\right)^2, & \text{if } n \equiv v_0 \pmod W, \\ 0, & \text{otherwise,} \end{cases} \tag{27}$$

for some coefficients λ_d (which we will choose explicitly later) satisfying

$$\lambda_d = 0 \quad \text{if } d > X^{1/10} \text{ or } \mu^2(d) = 0 \text{ or } d \notin \langle P_3 \rangle \text{ or } (d, p_0 W) \neq 1. \tag{28}$$

We recall that $\langle P_1 \rangle$ and $\langle P_3 \rangle$ denote the sets of integers composed only of prime factors $p \equiv 1 \pmod 4$ and $p \equiv 3 \pmod 4$, respectively. Finally, we define the function $\nu = \nu_{\mathscr{L}}$ on primes by

$$\nu(p) = \#\{1 \leq n < p : \prod_{i=1}^k L_i(n) \equiv 0 \pmod p\}. \tag{29}$$

We note that since $\{L_1, \dots, L_k\}$ is $\langle P_3 \rangle$-admissible, $\nu(p) < p$ for any prime $p \equiv 3 \pmod 4$.

7 Proof of Theorem 4.3

Lemma 7.2. *Assume the hypotheses of Theorem 4.3 and that* $\sup_d |\lambda_d| \ll (\log X)^{k/2}$. *Then*

$$\sum_{\substack{X < n \leq 2X \\ n \equiv v_0 \pmod W}} w_n = \frac{X}{W} \sum_{d,e} \lambda_d \lambda_e \prod_{p|de} \frac{\nu(p)}{p} + O\left(X^{1/5+\epsilon}\right).$$

Proof. From the definition of w_n, we have

$$\sum_{\substack{X < n \leq 2X \\ n \equiv v_0 \pmod W}} w_n = \sum_{d,e} \lambda_d \lambda_e \sum_{\substack{X < n \leq 2X \\ n \equiv v_0 \pmod W \\ d,e | \prod_{i=1}^k L_i(n)}} 1. \tag{30}$$

We concentrate on the inner sum. For each prime $p_1 | de$, we must have $L_i(n) \equiv 0 \pmod{p_1}$ for some $1 \leq i \leq k$, and so n can lie in one of $\nu(p_1)$ residue classes. Thus, using the Chinese remainder theorem, we can rewrite the inner sum as a sum of $\prod_{p|de} \nu(p)$ sums of n in a single residue class $\pmod{W[d,e]}$, where $[d,e]$ is the least common multiple of d and e (since $\lambda_d = 0$ if $(d, W) \neq 1$ or d is not square-free). The number of integers $X < n \leq 2X$ counted in any such residue class is then $X/([d,e]W) + O(1)$. Putting this together gives

$$\sum_{d,e}\lambda_d\lambda_e \sum_{\substack{X<n\le 2X\\ n\equiv v_0 \pmod W\\ d,e\mid\prod_{i=1}^k L_i(n)}} 1 = \sum_{d,e}\lambda_d\lambda_e\Big(\prod_{p\mid de}\nu(p)\Big)\Big(\frac{X}{W[d,e]}+O(1)\Big). \tag{31}$$

We see that the main term gives the corresponding main term in the statement of the lemma. Since $\nu(p)\le k$, the contribution from the error term is

$$\begin{aligned}\ll \sup_d|\lambda_d|^2\Big(\sum_{d<X^{1/10}}\prod_{p\mid d}k\Big)^2 &\le \sup_d|\lambda_d|^2\Big(X^{1/10}\sum_{d<X}\frac{\prod_{p\mid d}k}{d}\Big)^2\\ &\ll \sup_d|\lambda_d|^2X^{1/5}(1+\log X)^{2k}.\end{aligned} \tag{32}$$

By assumption of the Lemma, $k\le(\log X)^{1/3}$ and $\sup_d|\lambda_d|\ll(\log X)^{k/2}$, so this contribution is $O(X^{1/5+\epsilon})$. □

Lemma 7.3. *Assume the hypotheses of Theorem 4.3 and that* $\sup_d|\lambda_d|\ll(\log X)^{k/2}$. *Then*

$$\begin{aligned}\sum_{\substack{X<n\le 2X\\ n\equiv v_0 \pmod W}}\#\Big\{i: L_i(n)\in\mathscr{S}\Big\}w_n \gg{}& \frac{Xk}{\phi(W)(\log X)^{1/2}}\sum_{d,e}\lambda_d\lambda_e\prod_{p\mid de}\frac{\nu(p)-1}{p-1}\\ &+O(X\exp(-c\sqrt{\log X}))\end{aligned}$$

for some absolute constant $c>0$.

Proof. We first obtain a lower bound by only counting a subset of $\#\{1\le i\le k: L_i(n)\in\mathscr{S}\}$ which will be more convenient for our manipulations (we lose an unimportant constant factor in doing this). Let $\mathscr{S}'\subseteq\mathscr{S}$ be given by

$$\mathscr{S}'=\{n\in\langle P_1\rangle : n=mp \text{ for some } m<X^{1/3} \text{ and some prime } p\}. \tag{33}$$

We obtain a lower bound by only counting elements of $\mathscr{S}'$. This gives

$$\begin{aligned}\sum_{\substack{X<n\le 2X\\ n\equiv v_0 \pmod W}}\#\Big\{i: L_i(n)\in\mathscr{S}\Big\}w_n &\ge \sum_{\substack{X<n\le 2X\\ n\equiv v_0 \pmod W}}\#\Big\{1\le i\le k: L_i(n)\in\mathscr{S}'\Big\}w_n\\ &=\sum_{j=1}^k\sum_{\substack{X<n\le 2X\\ n\equiv v_0 \pmod W}}\mathbf{1}_{\mathscr{S}'}(L_j(n))\Big(\sum_{d\mid\prod_{i=1}^k L_i(n)}\lambda_d\Big)^2.\end{aligned} \tag{34}$$

By definition, if $L_j(n)=a_jn+b_j\in\mathscr{S}'$, then $L_j(n)$ can be written uniquely as mp for some $m<X^{1/3}$ composed only of primes congruent to 1 (mod 4) and

with $(m, a_j) = 1$, and some prime $p \equiv 1 \pmod 4$. Making this substitution and rearranging the sums, we are left to estimate

$$\sum_{j=1}^{k} \sum_{\substack{m\le X^{1/3}\\ m\in\langle P_1\rangle\\ (m,a_j)=1}} \sum_{\substack{d,e<X^{1/10}\\ d,e\in\langle P_3\rangle}} \lambda_d\lambda_e \sum_{\substack{(a_jX+b_j)/m<p<(2a_jX+b_j)/m\\ p\equiv 1 \pmod 4\\ p\equiv \overline{m}(a_jv_0+b_j) \pmod{a_jW}\\ [d,e]\mid \prod_{i=1}^{k} L_i((pm-b_j)/a_j)}} 1. \tag{35}$$

For each prime $p_1 | de$ with $p_1 > (\log x)^{1/3}$, there are $\nu(p) - 1$ possible primitive residue classes for $p \pmod{p_1}$ such that $\prod_{i=1}^{k} L_i((mp - b_j)/a_j) \equiv 0 \pmod{p_1}$. (All $\nu(p)$ residue classes for which $\prod_{i=1}^{k} L_i(n) \equiv 0 \pmod{p_1}$ correspond to a primitive residue class for p except for $n \equiv -\overline{a_j}b_j \pmod{p_1}$, and these all exist since $a_i < (\log X)^{1/3} < p_1$ and $p_1 \equiv 3 \pmod 4$ so $p_1 \nmid m$.) Thus, by the Chinese remainder theorem, we can rewrite the inner sum as a sum of $\prod_{p|de}(\nu(p) - 1)$ different sums of primes in arithmetic progressions to modulus $4a_jW\prod_{p|de} p$ (since $2 \nmid a_j$). Thus the inner sum is

$$\frac{1}{\phi(4a_jW)}\Big(\pi\Big(\frac{2a_jX + b_j}{m}\Big) - \pi\Big(\frac{a_jX + b_j}{m}\Big)\Big)\prod_{p|de}\frac{\nu(p) - 1}{p - 1} + O\Big(E\Big(\frac{3a_jX}{m}; 4a_jW[d, e]\Big)\prod_{p|de} k\Big), \tag{36}$$

where

$$E(x; q) = \sup_{\substack{(a,q)=1\\ x'\le x}} \Big|\pi(x'; q, a) - \frac{\pi(x')}{\phi(q)}\Big|, \qquad [d, e] = \prod_{p|de} p. \tag{37}$$

We first consider the error term of (36). Letting $r = 4a_jW[d, e] \ll X^{1/4}$ (noting that $(r, p_0) = 1$) and $\lambda_{max} = \sup_d |\lambda_d|$, this contributes at most

$$\begin{aligned} &k\lambda_{\max}^2 \sum_{m<X^{1/3}} \sum_{\substack{r\ll X^{1/4}\\ (r,p_0)=1}} E\Big(\frac{3a_jX}{m}; r\Big)\prod_{p|r} 3k\\ &\le k\lambda_{\max}^2 \sum_{m<X^{1/3}} \Big(\frac{3a_jX}{m} \sum_{r<4WX^{1/5}} \frac{\prod_{p|r} 9k^2}{r}\Big)^{1/2}\Big(\sum_{\substack{r<4WX^{1/5}\\ (r,p_0)=1}} E\Big(\frac{3a_jX}{m}; r\Big)\Big)^{1/2}\\ &\ll k\lambda_{max}^2\Big(a_jX(9k^2 + \log X)^{9k^2}\Big)^{1/2}\Big(a_jX\exp(-c_0\sqrt{\log X^{2/3}})\Big)^{1/2}\sum_{m<X^{1/3}}\frac{1}{m}. \end{aligned} \tag{38}$$

Here we have used Cauchy–Schwarz and the trivial bound $E(X,q) \ll X/q$ for $q < X$ in the first line, and Lemma 6.1 in the second. Thus, since $k \le (\log X)^{1/5}$, $a_j \le (\log X)^{1/3}$ and $\lambda_{\max} \ll (\log X)^{k/2}$ by assumption of the lemma, the contribution from (38) is $O(X\exp(-\frac{c_0}{4}\sqrt{\log X}))$.

We now consider the contribution from the main term of (36). The main term factorizes, simplifying to

$$\frac{1}{\phi(4W)}\Big(\prod_{p|a_j,p\nmid W}\frac{p}{p-1}\Big)\Big(\sum_{\substack{m<X^{1/3}\\ m\in\langle P_1\rangle\\ (m,a_j)=1}}\frac{X+o(X)}{m\log\frac{X}{m}}\Big)\Big(\sum_{d,e}\lambda_d\lambda_e\prod_{p|de}\frac{\nu(p)-1}{p-1}\Big). \tag{39}$$

The second term is straightforward to evaluate. For $t > X^{\epsilon}$ we have that

$$\sum_{\substack{m<t\\ m\in\langle P_1\rangle\\ (m,a_j)=1}} 1 \gg t(\log t)^{-1/2}\prod_{\substack{p|a_j\\ p\equiv 1 \pmod 4}}\frac{p-1}{p}, \tag{40}$$

and so by partial summation we see that

$$\sum_{\substack{m<X^{1/3}\\ m\in\langle P_1\rangle\\ (m,a_j)=1}}\frac{X+o(X)}{m\log\frac{X}{m}} \gg X(\log X)^{-1/2}\prod_{\substack{p|a_j\\ p\equiv 1 \pmod 4}}\frac{p-1}{p}. \tag{41}$$

Noting that the products over prime divisors of a_j cancel, this completes the proof of the lemma. □

Lemma 7.4. *Fix $A > 0$. Let $1/A \le \kappa \le (\log R)^{1/3}$ and let $f : [0,1] \to \mathbb{R}$ be a smooth non-negative function. Let $\gamma(p)$ satisfy $0 \le \gamma(p) \le \min(A\kappa, (1-1/A)p)$ for all $p < R$, and*

$$-L \le \sum_{w<p<z}\frac{\gamma(p)\log p}{p} - \kappa\log z/w \le A$$

for all $2 \le w \le z < R$. Then

$$\sum_{r<R}\mu^2(r)\Big(\prod_{p|r}\frac{\gamma(p)}{p-\gamma(p)}\Big)f\Big(\frac{\log r}{\log R}\Big) = \mathfrak{S}_{\gamma}\frac{(\log R)^{\kappa}}{\Gamma(\kappa)}\int_0^1 f(t)t^{\kappa-1}dt\Big(1+O_{A,f}\Big(\frac{\kappa L+\kappa^2\log\log R}{\log R}\Big)\Big),$$

where

$$\mathfrak{S}_\gamma = \prod_{p<R}\Big(1-\frac{\gamma(p)}{p}\Big)^{-1}\Big(1-\frac{1}{p}\Big)^{\kappa}.$$

Proof. For $f = 1$, this is a version of [9, Lemma 5.4], with the dependence on κ kept explicit. In particular, this is obtained in the proof of [4, Lemma 5.1]. The case of $f \neq 1$ then follows immediately by partial summation. □

Lemma 7.5. *Let $k \leq (\log X)^{1/5}$, and let λ_d be defined by*

$$\lambda_d = \mu(d)\Big(\prod_{p|d}\frac{p}{\nu(p)}\Big)\sum_{d|r} y_r\Big(\prod_{p|r}\frac{\nu(p)}{p-\nu(p)}\Big),$$

for $(d, p_0W) = 1$, where

$$y_r = \begin{cases} \mu(r)^2, & \text{if } r \in \langle P_3\rangle,\ (r, Wp_0) = 1,\ r < X^{1/10}, \\ 0, & \text{otherwise.}\end{cases}$$

Then

$$\sum_{d,e}\lambda_d\lambda_e\prod_{p|de}\frac{\nu(p)-1}{p-1} \gg k^{-1/2}(\log X)^{1/2}\frac{\phi(W)}{W}\sum_{d,e}\lambda_d\lambda_e\prod_{p|de}\frac{\nu(p)}{p} \gg 1.$$

Proof. We first note that the y_r variables diagonalize the second quadratic form. Substituting the expression for λ_d in terms of y_r given by then lemma, we obtain

$$\begin{aligned}\sum_{d,e}\lambda_d\lambda_e\prod_{p|de}\frac{\nu(p)}{p} &= \sum_{r,s}y_ry_s\Big(\prod_{p|r}\frac{\nu(p)}{p-\nu(p)}\Big)\Big(\prod_{p|s}\frac{\nu(p)}{p-\nu(p)}\Big) \\ &\quad\times\sum_{d,e:d|r,e|s}\mu(d)\mu(e)\Big(\prod_{p|d}\frac{p}{\nu(p)}\Big)\Big(\prod_{p|e}\frac{p}{\nu(p)}\Big)\Big(\prod_{p|de}\frac{\nu(p)}{p}\Big). \qquad (42)\end{aligned}$$

The sum over d, e vanishes unless $r = s$, and so we find that

$$\sum_{d,e}\lambda_d\lambda_e\prod_{p|de}\frac{\nu(p)}{p} = \sum_r y_r^2\prod_{p|r}\frac{\nu(p)}{p-\nu(p)}. \qquad (43)$$

We now introduce new variables y_r^* to perform an analogous diagonalization for the first quadratic form. Let y_r^* be zero unless $r \in \langle P_3\rangle$, $(r, Wp_0) = 1$, $r < X^{1/10}$ and $\nu(p) > 1$ for all $p|r$. In this case, let y_r^* be given by

$$y_r^* = \mu(r)\Big(\prod_{p|r}\frac{p-\nu(p)}{\nu(p)-1}\Big)\sum_{d:r|d}\lambda_d\Big(\prod_{p|d}\frac{\nu(p)-1}{p-1}\Big). \qquad (44)$$

We see from this that for any d satisfying the same conditions, we have

$$\begin{aligned}\mu(d)\lambda_d &= \Big(\prod_{p|d}\frac{p-1}{\nu(p)-1}\Big)\sum_{e:d|e}\lambda_e\Big(\prod_{p|e}\frac{\nu(p)-1}{p-1}\Big)\sum_{r:d|r,\,r|e}\mu(r)\\ &= \Big(\prod_{p|d}\frac{p-1}{\nu(p)-1}\Big)\sum_{r:d|r}y_r^*\Big(\prod_{p|r}\frac{\nu(p)-1}{p-\nu(p)}\Big).\end{aligned} \tag{45}$$

This gives λ_d in terms of y_r^*. Performing the analogous computation to (43) with y_r^* in place of y_r, we obtain

$$\sum_{d,e}\lambda_d\lambda_e\prod_{p|de}\frac{\nu(p)-1}{p-1} = \sum_r (y_r^*)^2\prod_{p|r}\frac{\nu(p)-1}{p-\nu(p)}. \tag{46}$$

Substituting the expression for λ_d in terms of y_r from the statement of the lemma in (44) gives (for r such that $y_r^* \neq 0$)

$$\begin{aligned}y_r^* &= \mu(r)\Big(\prod_{p|r}\frac{p-\nu(p)}{\nu(p)-1}\Big)\sum_{s:r|s}y_s\Big(\prod_{p|s}\frac{\nu(p)}{p-\nu(p)}\Big)\sum_{d:r|d,\,d|s}\mu(d)\Big(\prod_{p|d}\frac{p(\nu(p)-1)}{\nu(p)(p-1)}\Big)\\ &= \mu(r)^2 r\sum_{s:r|s}\frac{y_s}{\phi(s)}.\end{aligned} \tag{47}$$

We can now use Lemma 7.4 and definition of y_r from the statement of the lemma to evaluate this sum. We take $\gamma_1(p) = 1$ if $p \equiv 3 \pmod 4$ and $(p, rp_0W) = 1$, and take $\gamma_1(p) = 0$ otherwise. We see that for any $2 \le w \le z \le X^{1/10}$ we have

$$-\frac{1}{2}\sum_{p|rp_0W}\frac{\log p}{p} + O(1) \le \sum_{w<z<X^{1/10}}\frac{\gamma_1(p)\log p}{p} - \frac{1}{2}\log z/w \ll 1. \tag{48}$$

The sum on the left hand side is $O(\log\log X)$. Thus, by applying Lemma 7.4 to the sum (47), we have for squarefree $r < X^{1/10}$ with $r \in \langle P_3\rangle$ and $(r, Wp_0) = 1$ and $\prod_{p|r}(\nu(p)-1) > 0$ that

$$\begin{aligned}y_r^* &= \frac{r}{\phi(r)}\sum_{t<X^{1/10}/r}\mu^2(t)\prod_{p|t}\frac{\gamma_1(p)}{p-\gamma_1(p)}\\ &= \frac{r}{\phi(r)\Gamma(3/2)}\left(\Big(\log\frac{X^{\frac{1}{10}}}{r}\Big)^{\frac{1}{2}} + O\Big(\frac{\log\log X}{(\log X)^{\frac{1}{2}}}\Big)\right)\prod_{p<X}\Big(1-\frac{\gamma_1(p)}{p}\Big)^{-1}\Big(1-\frac{1}{p}\Big)^{\frac{1}{2}}.\end{aligned} \tag{49}$$

Since $r, W \in \langle P_3\rangle$ and $\phi(p_0)/p_0 = 1 + o(1)$, the product simplifies to give

$$\frac{r}{\phi(r)}\prod_{p<X}\Big(1-\frac{\gamma_1(p)}{p}\Big)^{-1}\Big(1-\frac{1}{p}\Big)^{1/2} = (1+o(1))\frac{\phi(W)}{W}\prod_{p<X}\Big(1-\frac{1}{p}\Big)^{\chi(p)/2}, \tag{50}$$

where χ is the non-trivial character modulo 4.

We can now evaluate both of the expressions (43) and (46) using Lemma 7.4. For (43), we take $\gamma_2(p) = \nu(p)$ if $p \equiv 3 \pmod 4$ and $(p, Wp_0) = 1$, and $\gamma_2(p) = 0$ otherwise. We see that

$$-\frac{k}{2}\sum_{p\mid p_0 W\prod_{i\neq j}(a_ib_j-b_ja_i)}\frac{\log p}{p} + O(1) \geq \sum_{w<p<z}\frac{\gamma_2(p)\log p}{p} - \frac{k}{2}\log\frac{z}{w} \ll 1. \tag{51}$$

Thus, we obtain (with $\kappa = k/2$ and $L \ll k\log\log X$)

$$\sum_{r<X^{1/10}} y_r^2\prod_{p\mid r}\frac{\nu(p)}{p-\nu(p)}$$
$$= \frac{(\log X^{1/10})^{k/2}}{\Gamma(1+k/2)}\Big(1+O\Big(\frac{k^2\log\log X}{\log X}\Big)\Big)\prod_{p<X}\Big(1-\frac{\gamma_2(p)}{p}\Big)^{-1}\Big(1-\frac{1}{p}\Big)^{k/2}. \tag{52}$$

Similarly, for (46) we take $\gamma_3(p) = (\nu(p)-1)p/(p-1)$ for $p \equiv 3 \pmod 4$ and $(p, Wp_0) = 1$, and take $\gamma_3(p) = 0$ otherwise. We obtain (with $\kappa = (k-1)/2$ and $L \ll k\log\log X$)

$$\sum_{r<X^{1/10}}\mu^2(r)\Big(\log\frac{X^{1/10}}{r} + O(\log\log X)\Big)\prod_{p\mid r}\frac{\gamma_3(p)}{p-\gamma_3(p)}$$
$$= \frac{(\log X^{1/10})^{\frac{k+1}{2}}}{\Gamma((k+3)/2)}\Big(1+O\Big(\frac{k^2\log\log X}{\log X}\Big)\Big)\prod_{p<X}\Big(1-\frac{\gamma_3(p)}{p}\Big)^{-1}\Big(1-\frac{1}{p}\Big)^{\frac{k-1}{2}}. \tag{53}$$

We note that if $\gamma_3(p) \neq 0$, then $1-\gamma_3(p)/p = (1-\gamma_2(p)/p)(1-1/p)^{-1}$. Thus, since $W \in \langle P_3\rangle$ and $\phi(p_0)/p_0 = 1 + o(1)$, we have

$$\prod_{p<X}\Big(1-\frac{\gamma_3(p)}{p}\Big)^{-1}\Big(1-\frac{1}{p}\Big)^{\frac{k-1}{2}} \sim \frac{W}{\phi(W)}\prod_{p<X}\Big(1-\frac{\gamma_2(p)}{p}\Big)^{-1}\Big(1-\frac{1}{p}\Big)^{\frac{k-\chi(p)}{2}}. \tag{54}$$

Putting the estimates (49), (50), (52)–(54) together, we obtain

$$\sum_r (y_r^*)^2\prod_{p\mid r}\frac{\nu(p)-1}{p-\nu(p)} \sim \frac{\phi(W)\Gamma(1+k/2)(\log X^{1/10})^{1/2}}{W\Gamma(3/2)^2\Gamma((k+3)/2)}\prod_{p<X}\Big(1-\frac{1}{p}\Big)^{\chi(p)/2}$$
$$\times\sum_r y_r^2\prod_{p\mid r}\frac{\nu(p)}{p-\nu(p)}. \tag{55}$$

Stirling's formula shows that $\Gamma(x) \sim \sqrt{2\pi x} x^x e^{-x}$, and so for k sufficiently large $\Gamma(1+k/2)/\Gamma((k+3)/2) \gg k^{-1/2}$. Since the product over primes is convergent as $X \to \infty$, this gives

$$\sum_r (y_r^*)^2 \prod_{p|r} \frac{\nu(p)-1}{p-\nu(p)} \gg \frac{\phi(W)}{W} k^{-1/2} (\log X)^{1/2} \sum_r y_r^2 \prod_{p|r} \frac{\nu(p)}{p-\nu(p)}. \tag{56}$$

Finally, from (52) and $k \le (\log X)^{1/5}$, it follows that we have the crude bound that this is $\gg 1$. □

Lemma 7.6. *Let $k \le (\log X)^{1/5}$ and λ_d be as given by Lemma 7.5. Then*

$$|\lambda_d| \ll (\log X)^{k/2}.$$

Proof. This is an immediate application of Lemma 7.4. We take $\gamma_4(p) = \nu(p)$ if $p \equiv 3 \pmod 4$ and $(p, dWp_0) = 1$, and $\gamma_4(p) = 0$ otherwise. By Lemma 7.4 (taking $\kappa = k/2$ and $L \ll k \log\log X$) we have

$$\begin{aligned}
|\lambda_d| &= \Big(\prod_{p|d} \frac{p}{p-\nu(p)}\Big) \sum_{\substack{r < X^{1/10}/d \\ (r, dWp_0)=1 \\ r \in \langle P_3 \rangle}} \mu^2(r) \prod_{p|r} \frac{\nu(p)}{p-\nu(p)} \\
&\ll \frac{(\log X)^{k/2}}{\Gamma(1+k/2)} \prod_{p|d} \Big(1-\frac{\nu(p)}{p}\Big)^{-1} \prod_p \Big(1-\frac{\gamma_4(p)}{p}\Big)^{-1} \Big(1-\frac{1}{p}\Big)^{k/2} \\
&\ll \frac{(\log X)^{k/2}}{\Gamma(1+k/2)} \prod_{\substack{(\log X)^{1/3} < p < X \\ p \equiv 3 \pmod 4}} \Big(1-\frac{k}{p}\Big)^{-1} \Big(1-\frac{1}{p}\Big)^{k/2} \prod_{\substack{(\log X)^{1/3} < p < X \\ p \equiv 1 \pmod 4}} \Big(1-\frac{1}{p}\Big)^{k/2}.
\end{aligned} \tag{57}$$

Recalling that $k \le (\log X)^{1/5}$, we see that this is

$$\begin{aligned}
&\ll \frac{(\log X)^{k/2}}{\Gamma(1+k/2)} \exp\Big(\frac{k}{2} \sum_{\substack{(\log X)^{1/3} < p < X \\ p \equiv 3 \pmod 4}} \frac{1+O(p^{-1})}{p} - \frac{k}{2} \sum_{\substack{(\log X)^{1/3} < p < X \\ p \equiv 1 \pmod 4}} \frac{1+O(p^{-1})}{p}\Big) \\
&\ll \exp(O(k)) k^{-k/2} (\log X)^{k/2} \ll (\log X)^{k/2}.
\end{aligned} \tag{58}$$

□

Proof (Proof of Theorem 4.3). By Lemmas 7.5 and 7.6, there is a choice of coefficients λ_d such that $|\lambda_d| \ll (\log X)^{k/2}$ and

$$\sum_{d,e} \lambda_d \lambda_e \prod_{p|de} \frac{\nu(p)-1}{p-1} \gg k^{-1/2} (\log X)^{1/2} \frac{\phi(W)}{W} \sum_{d,e} \lambda_d \lambda_e \prod_{p|de} \frac{\nu(p)}{p} \gg 1.$$

Thus, by Lemmas 7.2 and 7.3, this choice of λ_d corresponds to a choice of $w_n \geq 0$ such that

$$\sum_{\substack{X<n\leq 2X\\ n\equiv v_0 \pmod W}} \#\Big\{1\leq i\leq k : L_i(n)\in \mathscr{S}\Big\} w_n \gg k^{1/2} \sum_{\substack{X<n\leq 2X\\ n\equiv v_0 \pmod W}} w_n \gg \frac{X}{W\log X}. \tag{59}$$

Lemma 7.6 shows $|\lambda_d| \ll (\log X)^{k/2}$, so we have

$$w_n \ll (\log X)^k \Big(\sum_{\substack{d\mid \prod L_i(n)\\ \lambda_d\neq 0}} 1\Big)^2. \tag{60}$$

Thus

$$\begin{aligned}\sum_{X\leq n\leq 2X} w_n^2 &\ll (\log X)^{2k} \sum_{d_1,d_2,d_3,d_4<X^{1/10}} \Big(\prod_{i=1}^4 \mu^2(d_i)\Big) \sum_{\substack{X\leq n\leq 2X\\ d_1,d_2,d_3,d_4\mid \prod_{i=1}^k L_i(n)}} 1\\ &\ll X(\log X)^{2k} \sum_{d_1,d_2,d_3,d_4<X^{1/10}} \Big(\prod_{i=1}^4 \mu^2(d_i)\Big) \prod_{p\mid d_1d_2d_3d_4} \frac{k}{p}\\ &\ll X(\log X)^{2k} \sum_{r<X^{2/5}} \frac{(15k^2)^{\nu(r)}}{r} \ll X(\log x)^{16k^2}.\end{aligned} \tag{61}$$

By Cauchy–Schwarz, we have that

$$\begin{aligned}&\#\{n\in[X,2X] : \#\{L_1(n),\dots,L_k(n)\}\cap \mathscr{S}\geq ck^{1/2}\}\\ &\geq \Big(k^{-1}\sum_{n\in[X,2X]} \Big(\#\{i : L_i(n)\in\mathscr{S}\} - ck^{1/2}\Big)w_n\Big)^2 \Big(\sum_{n\in[X,2X]} w_n^2\Big)^{-1},\end{aligned} \tag{62}$$

provided the sum being squared is positive. Therefore, by (59) and (61), we see that for c sufficiently small we have

$$\begin{aligned}&\#\{n\in[X,2X] : \#\{L_1(n),\dots,L_k(n)\}\cap \mathscr{S}\geq ck^{1/2}\}\\ &\geq \frac{1}{X(\log X)^{16k^2}} \Bigg(\sum_{\substack{n\in[X,2X]\\ n\equiv v_0 \pmod W}} \Big(\#\Big\{i : L_i(n)\in\mathscr{S}\Big\} w_n - ck^{1/2}\Big)w_n\Bigg)^2\\ &\gg \frac{X}{W^2(\log X)^{16k^2+2}} \gg X\exp(-(\log X)^{1/2}).\end{aligned} \tag{63}$$

Here we have used the fact that $k\leq (\log X)^{1/5}$ and $W\ll \exp((\log X)^{2/5})$ in the last line. This completes the proof of Theorem 4.3. □

Acknowledgements The work in this paper was started whilst the author was a CRM-ISM Postdoctoral Fellow and the Université de Montréal, and was completed whilst he was a Fellow by Examination at Magdalen College, Oxford.

References

1. A. Balog, T.D. Wooley, Sums of two squares in short intervals. Can. J. Math. **52**(4), 673–694 (2000)
2. H. Davenport, *Multiplicative Number Theory*. Graduate Texts in Mathematics (Springer, New York, 2000)
3. K. Ford, B. Green, S. Konyagin, T. Tao, Large gaps between consecutive primes (2014). Preprint, http://arxiv.org/abs/1408.4505
4. K. Ford, S.V. Konyagin, F. Luca, Prime chains and Pratt trees. Geom. Funct. Anal. **20**(5), 1231–1258 (2010)
5. J. Friedlander, A. Granville, Limitations to the equi-distribution of primes. III. Compos. Math. **81**(1), 19–32 (1992)
6. D.A. Goldston, S.W. Graham, J. Pintz, C.Y. Yıldırım, Small gaps between products of two primes. Proc. Lond. Math. Soc. (3) **98**(3), 741–774 (2009)
7. D.A. Goldston, J. Pintz, C.Y. Yıldırım, Primes in tuples. II. Acta Math. **204**(1), 1–47 (2010)
8. A. Granville, K. Soundararajan, An uncertainty principle for arithmetic sequences. Ann. Math. (2) **165**(2), 593–635 (2007)
9. H. Halberstam, H.E. Richert, *Sieve Methods*. London Mathematical Society Monographs (Academic, London, 1974)
10. D.R. Heath-Brown, The number of primes in a short interval. J. Reine Angew. Math. **389**, 22–63 (1988)
11. C. Hooley, On the intervals between numbers that are sums of two squares. IV. J. Reine Angew. Math. **452**, 79–109 (1994)
12. M.N. Huxley, On the difference between consecutive primes. Invent. Math. **15**, 164–170 (1972)
13. H. Maier, Primes in short intervals. Mich. Math. J. **32**(2), 221–225 (1985)
14. H. Maier, C. Pomerance, Unusually large gaps between consecutive primes. Trans. Am. Math. Soc. **322**(1), 201–237 (1990)
15. H. Maier, C.L. Stewart, On intervals with few prime numbers. J. Reine Angew. Math. **608**, 183–199 (2007)
16. J. Maynard, Dense clusters of primes in subsets (2014). Preprint, http://arxiv.org/abs/1405.2593
17. J. Maynard, Large gaps between primes (2014). Preprint, http://arxiv.org/abs/1408.5110
18. J. Maynard, Small gaps between primes. Ann. of Math. (2) **181**(1), 383–413 (2015)
19. J. Pintz, Very large gaps between consecutive primes. J. Number Theory **63**(2), 286–301 (1997)
20. I. Richards, On the gaps between numbers which are sums of two squares. Adv. Math. **46**(1), 1–2 (1982)
21. Y. Zhang, Bounded gaps between primes. Ann. Math. (2) **179**(3), 1121–1174 (2014)

Infinite Sumsets with Many Representations

Melvyn B. Nathanson

To Helmut Maier on his 60th birthday

Abstract Let A be an infinite set of nonnegative integers. For $h \geq 2$, let hA be the set of all sums of h not necessarily distinct elements of A. Suppose that $\ell \geq 2$, and that every sufficiently large integer in the sumset hA has at least ℓ representations. If $\ell = 2$, then $A(x) \geq (\log x/\log h) - w_0$, where $A(x)$ counts the number of integers $a \in A$ such that $1 \leq a \leq x$. Lower bounds for $A(x)$ are also obtained for $\ell \geq 3$.

1 Representation Functions of Sumsets

Let A be a set of integers and let $h \geq 2$ be an integer. The *counting function* $A(x)$ counts the number of positive integers in the set A that do not exceed x. The h-fold sumset hA is the set of all integers n that can be written as sums of h not necessarily distinct elements of A. For every integer n, the *representation function* $r_{A,h}(n)$ counts the number of h-tuples $(a_1, a_2, \ldots, a_h) \in A^h$ such that

$$a_1 \leq a_2 \leq \cdots \leq a_h$$

and

$$a_1 + a_2 + \cdots + a_h = n.$$

A *Sidon set* is a set A of nonnegative integers such that every element in the sumset $2A$ has a unique representation, that is, $r_{A,2}(n) = 1$ for all $n \in 2A$. More generally, for positive integers h and s, a $B_{h,s}$-set is a set A of nonnegative integers such that $r_{A,h}(n) \leq s$ for all $n \in hA$. Sets whose sumsets have few representations have been studied intensively (cf. Halberstam–Roth [2], O'Bryant [4]).

M.B. Nathanson (✉)
Department of Mathematics, Lehman College (CUNY), Bronx, NY 10468, USA
e-mail: melvyn.nathanson@lehman.cuny.edu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_16

In this paper we consider sets whose h-fold sumsets have many representations. A basic result is that if $h \geq 2$ and A is an infinite set of nonnegative integers with $r_{A,h}(n) \geq 2$ for all sufficiently large integers $n \in hA$, then

$$A(x) \gg \log x.$$

In the special case $h = 2$, Balasubramanian and Prakesh [1] proved that there is a number $c > 0$ such that, if A is an infinite set of nonnegative integers with $r_{A,2}(n) \geq 2$ for all sufficiently large integers $n \in 2A$, then

$$A(x) \geq c \left(\frac{\log x}{\log \log x} \right)^2$$

for $x \geq x_0$. This improved the previous result of Nicolas et al. [3], who also proved the existence of an infinite set A of nonnegative integers with $r_{A,2}(n) \geq 2$ for all sufficiently large integers $n \in 2A$ such that

$$A(x) \ll (\log x)^2.$$

It is an open problem to extend these results to h-fold sumsets for $h \geq 3$.

2 Growth of Sets with Many Representations

Let $[u, v)$ denote the interval of integers i such that $u \leq i < v$. Let $|X|$ denote the cardinality of the set X.

Theorem 2.1. *Let $h \geq 2$ be an integer, and let A be an infinite set of nonnegative integers. If $r_{A,h}(n) \geq 2$ for all sufficiently large integers $n \in hA$, then there is a positive number $k_0 = k_0(A)$ such that*

$$A(x) > \frac{\log x}{\log h} - k_0$$

for all $x \geq h$.

Proof. For every positive integer k, let

$$I_k = \left[h^{k-1}, h^k \right)$$

and

$$A_k = A \cap I_k.$$

The sets $\{A_k : k = 1, 2, \ldots\}$ partition $A \setminus \{0\}$.

There exists a positive integer n_0 such that, if $n \geq n_0$ and $n \in hA$, then $r_{A,h}(n) \geq 2$. Because A is infinite, there exists $a_0 \in A$ with $ha_0 \geq n_0$. Choose k_0 such that $a_0 \in A_{k_0}$.

Suppose that $k \geq k_0$ and $A_k \neq \emptyset$. Let

$$a_k^* = \max(A_k).$$

Then

$$h^{k-1} \leq a_k^* < h^k$$

and

$$A \cap \left[a_k^* + 1, h^k\right) = \emptyset.$$

Consider the integer

$$ha_k^* \in hA.$$

Because $a_k^* \geq a_0$, we have $ha_k^* \geq ha_0 \geq n_0$, and so $r_{A,h}(ha_k^*) \geq 2$. It follows that the set A contains nonnegative integers $a_1, \ldots, a_h$ such that

$$a_1 < a_h \tag{1}$$

$$a_1 \leq a_2 \leq \cdots \leq a_h \tag{2}$$

and

$$a_1 + a_2 + \cdots + a_h = ha_k^*.$$

Because A is a set of nonnegative integers, we have

$$a_h \leq ha_k^*.$$

Inequalities (1) and (2) imply that

$$ha_k^* = a_1 + a_2 + \cdots + a_h < ha_h$$

and so

$$a_k^* < a_h \leq ha_k^* < h^{k+1}.$$

Therefore,

$$a_k^* + 1 \leq a_h < h^{k+1}.$$

Equivalently,

$$a_h \in \left[a_k^* + 1, h^{k+1}\right) = \left[a_k^* + 1, h^k\right) \cup A_{k+1}.$$

The maximality of a_k^* implies that $A \cap \left[a_k^* + 1, h^k\right) = \emptyset$, and so $a_h \in A_{k+1} \neq \emptyset$. It follows by induction that $A_k \neq \emptyset$ for all $k \geq k_0$.

Let $x \geq h$, and choose the positive integer t such that

$$h^t \leq x < h^{t+1}.$$

Because

$$A \setminus \{0\} = \bigcup_{k=1}^{\infty} A_k$$

it follows that

$$A(x) \geq A\left(h^t\right) \geq t - (k_0 - 1) > \frac{\log x}{\log h} - k_0$$

for all $x \geq h$. This completes the proof.

Theorem 2.2. *Let $\ell \geq 3$ be an integer, and let A be an infinite set of nonnegative integers. If $r_{A,2}(n) \geq \ell$ for all sufficiently large integers $n \in 2A$, then there is a positive number $k_1 = k_1(A)$ such that*

$$A(x) > \frac{(\ell - 1)\log x}{\log 2} - k_1$$

for all $x \geq 2$.

Proof. For every positive integer k, let

$$I_k = \left[2^{k-1}, 2^k\right)$$

and

$$A_k = A \cap I_k.$$

The sets $\{A_k : k = 1, 2, \ldots\}$ partition $A \setminus \{0\}$.

There exists a positive integer n_0 such that if $n \geq n_0$ and $n \in 2A$, then $r_{A,2}(n) \geq \ell$. Because A is infinite, there exists $a_0 \in A$ with $2a_0 \geq n_0$. Choose k_0 such that $a_0 \in A_{k_0}$.

Suppose that $k \geq k_0$ and $A_k \neq \emptyset$. Let

$$a_k^* = \max(A_k).$$

Then

$$2^{k-1} \leq a_k^* < 2^k$$

and

$$A \cap \left[a_k^* + 1, 2^k\right) = \emptyset.$$

Consider the integer

$$2a_k^* \in 2A.$$

Because $a_k^* \geq a_0$, we have $2a_k^* \geq 2a_0 \geq n_0$, and so $r_{A,2}(2a_k^*) \geq \ell$. It follows that the set A contains a subset

$$\{a_{i,j} : i = 1, 2 \text{ and } j = 1, \ldots, \ell - 1\}$$

such that, for $j = 1, 2, \ldots, \ell - 1$, we have

$$a_{1,1} < a_{1,2} < \cdots < a_{1,\ell-1} < a_k^* < a_{2,\ell-1} < \cdots < a_{2,2} < a_{2,1}$$

and

$$a_{1,j} + a_{2,j} = 2a_k^*.$$

Moreover, $a_{1,j} \geq 0$ implies that

$$a_{2,j} \leq 2a_k^*$$

for $j = 1, 2, \ldots, \ell - 1$.

Because $0 \leq a_{1,j} < a_{2,j}$, we have

$$2a_k^* = a_{1,j} + a_{2,j} < 2a_{2,j}$$

and so

$$a_k^* < a_{2,j} \leq 2a_k^* < 2^{k+1}.$$

Therefore,

$$a_k^* + 1 \leq a_{2,j} < 2^{k+1}.$$

Equivalently,

$$a_{2,j} \in \left[a_k^* + 1, 2^{k+1}\right) = \left[a_k^* + 1, 2^k\right) \cup A_{k+1}$$

for $j = 1, 2, \ldots, \ell - 1$. Because $A \cap \left[a_k^* + 1, 2^k\right) = \emptyset$, we see that $a_{2,j} \in A_{k+1}$ for $j = 1, 2, \ldots, \ell - 1$, and so $|A_{k+1}| \geq \ell - 1$. It follows by induction that $|A_k| \geq \ell - 1$ for all $k \geq k_0 + 1$.

Let $x \geq 2$, and choose the positive integer t such that

$$2^t \leq x < 2^{t+1}.$$

Because

$$A \setminus \{0\} = \bigcup_{k=1}^{\infty} A_k$$

it follows that

$$\begin{aligned} A(x) &\geq A\left(2^t\right) \geq (\ell - 1)(t - k_0) \\ &> (\ell - 1)\left(\tfrac{\log x}{\log 2} - k_0 - 1\right) \\ &= \tfrac{(\ell-1)\log x}{\log 2} - k_1 \end{aligned}$$

where $k_1 = (\ell - 1)(k_0 + 1)$. This completes the proof.

Theorem 2.3. *Let $h \geq 3$, $\ell \geq 3$, and $s \geq 1$ be integers, and let A be an infinite $B_{h-1,s}$ set of nonnegative integers. If $r_{A,h}(n) \geq \ell$ for all sufficiently large integers $n \in hA$, then there is a positive number $k_2 = k_2(A)$ such that*

$$A(x) > \frac{(\ell - 1)\log x}{s \log h} - k_2$$

for all $x \geq h$.

Proof. For every positive integer k, let

$$I_k = \left[h^{k-1}, h^k\right)$$

and

$$A_k = A \cap I_k.$$

The sets $\{A_k : k = 1, 2, \ldots\}$ partition $A \setminus \{0\}$.

There exists a positive integer n_0 such that, if $n \geq n_0$ and $n \in hA$, then $r_{A,h}(n) \geq \ell$. Because A is infinite, there exists $a_0 \in A$ with $ha_0 \geq n_0$. Choose k_0 such that $a_0 \in A_{k_0}$.

Suppose that $k \geq k_0$ and $A_k \neq \emptyset$. Let

$$a_k^* = \max(A_k).$$

Then

$$h^{k-1} \leq a_k^* < h^k$$

and

$$A \cap \left[a_k^* + 1, h^k\right) = \emptyset.$$

Consider the integer

$$ha_k^* \in hA.$$

Because $a_k^* \geq a_0$, we have $ha_k^* \geq ha_0 \geq n_0$, and so $r_{A,h}(ha_k^*) \geq \ell$. It follows that the set A contains a subset

$$\{a_{i,j} : i = 1, \ldots, h \text{ and } j = 1, \ldots, \ell - 1\}$$

such that, for $j = 1, 2, \ldots, \ell - 1$,

$$a_{1,j} < a_{h,j} \tag{3}$$

$$a_{1,j} \leq a_{2,j} \leq \cdots \leq a_{h-1,j} \leq a_{h,j} \tag{4}$$

$$a_{1,j} + a_{2,j} + \cdots + a_{h-1,j} + a_{h,j} = ha_k^*$$

and

$$\left(a_{1,j}, a_{2,j}, \ldots, a_{h-1,j}, a_{h,j}\right) \neq \left(a_{1,j'}, a_{2,j'}, \ldots, a_{h-1,j'}, a_{h,j'}\right)$$

for $1 \leq j < j' \leq \ell - 1$. Moreover, for $i = 1, \ldots, h-1$ and $j = 1, \ldots, \ell - 1$, the inequality $a_{i,j} \geq 0$ implies that

$$a_{h,j} \leq ha_k^*$$

Let $b \in A$ and let J be a subset of $\{1, \ldots, \ell - 1\}$ such that $a_{h,j} = b$ for all $j \in J$. If $j \in J$, then

$$a_{1,j} + a_{2,j} + \cdots + a_{h-1,j} = ha_k^* - a_{h,j} = ha_k^* - b$$

and so

$$r_{A,h-1}(ha_k^* - b) \geq |J|.$$

Because A is a $B_{h-1,s}$-set, we have

$$r_{A,h-1}(ha_k^* - b) \leq s$$

and so $|J| \leq s$. The pigeonhole principle implies that

$$\left|\left\{a_{h,j} : j = 1, \ldots, \ell - 1\right\}\right| \geq \frac{\ell - 1}{s}.$$

It follows from inequalities (3) and (4) that

$$ha_k^* = a_{1,j} + a_{2,j} + \cdots + a_{h,j} < ha_{h,j}$$

and so

$$a_k^* < a_{h,j} \leq ha_k^* < h^{k+1}.$$

Therefore,

$$a_k^* + 1 \leq a_{h,j} < h^{k+1}.$$

Equivalently,

$$a_{h,j} \in \left[a_k^* + 1, h^{k+1}\right) = \left[a_k^* + 1, h^k\right) \cup A_{k+1}$$

for $j = 1, 2, \ldots, \ell - 1$. Because $A \cap \left[a_k^* + 1, h^k\right) = \emptyset$, we see that

$$\{a_{h,j} : j = 1, 2, \ldots, \ell - 1\} \subseteq A_{k+1}$$

and so

$$|A_{k+1}| \geq \frac{\ell - 1}{s}.$$

It follows by induction that $|A_k| \geq (\ell - 1)/s$ for all $k \geq k_0 + 1$.

Let $x \geq h$, and choose the positive integer t such that

$$h^t \leq x < h^{t+1}.$$

Because

$$A \setminus \{0\} = \bigcup_{k=1}^{\infty} A_k$$

it follows that

$$\begin{aligned} A(x) &\geq A\left(h^t\right) \geq \left(\tfrac{\ell-1}{s}\right)(t-k_0) \\ &> \left(\tfrac{\ell-1}{s}\right)\left(\tfrac{\log x}{\log h} - k_0 - 1\right) \\ &= \tfrac{(\ell-1)\log x}{s \log h} - k_2 \end{aligned}$$

where $k_2 = (\ell - 1)(k_0 + 1)/s$. This completes the proof.

Acknowledgements I thank Michael Filaseta for bringing these problems to my attention, and Quan-Hui Yang for the reference to the paper of Balasubramanian and Prakesh.

References

1. R. Balasubramanian, G. Prakash, On an additive representation function. J. Number Theory **104**, 327–334 (2004)
2. H. Halberstam, K.F. Roth, *Sequences*, vol. 1 (Oxford University Press, Oxford, 1966); Reprinted by Springer, Heidelberg, in 1983
3. J.-L. Nicolas, I.Z. Ruzsa, A. Sárközy, On the parity of additive representation functions. J. Number Theory **73**, 292–317 (1998) [with an appendix in French by J.-P. Serre]
4. K. O'Bryant, A complete annotated bibliography of work related to Sidon sequences. Electron. J. Comb. Dyn. Surv. **11**, 39 (2004)

On the Ratio of Consecutive Gaps Between Primes

János Pintz

Dedicated to Helmut Maier on the occasion of his 60th birthday

Abstract In the present work we prove a common generalization of Maynard–Tao's recent result about consecutive bounded gaps between primes and of the Erdős–Rankin bound about large gaps between consecutive primes. The work answers in a strong form a 60-year-old problem of Erdős, which asked whether the ratio of two consecutive primegaps can be infinitely often arbitrarily small, and arbitrarily large, respectively. This is proved in the paper in a stronger form that not only $d_n = p_{n+1} - p_n$ can be arbitrarily large compared to d_{n+1} but this remains true if d_{n+1} is replaced by the maximum of the k differences $d_{n+1}, \dots, d_{n+k}$ for arbitrary fix k. The ratio can reach $c(k)$ times the size of the classical Erdős–Rankin function with a constant $c(k)$ depending only on k.

1 Introduction

The difference between the consecutive primes, the expression

$$d_n = p_{n+1} - p_n, \tag{1}$$

where $\mathscr{P} = \{p_i\}_{i=1}^{\infty}$ denotes the set of primes, has been investigated probably since the time of the Greeks. The Twin Prime Conjecture asserts

$$d_n = 2 \text{ infinitely often.} \tag{2}$$

This conjecture settles at the conjectural level the small values of d_n. Concerning the large values even the suitable conjecture is not completely clear. However, it seems to be that Cramér's conjecture [7, 8]

J. Pintz (✉)
Rényi Mathematical Institute of the Hungarian Academy of Sciences,
Reáltanoda u. 13–15, H-1053 Budapest, Hungary
e-mail: pintz.janos@renyi.mta.hu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_17

$$\limsup_{n\to\infty} \frac{d_n}{\log^2 n} = C_0 = 1 \tag{3}$$

is near to the truth. Granville suggested [28, 29], just based on the famous matrix method of Helmut Maier that the correct value of C is instead of 1 slightly larger

$$C \geq 2e^{-\gamma} > 1. \tag{4}$$

However, most mathematicians agree that the correct maximal order of d_n should be $(\log n)^{2+o(1)}$.

In the present work we give a short overview of the older and the recent new breakthrough results of small gaps between primes, of the present state of art of long differences. Our final result will be a common generalization of them (apart from the very recent ones proved in the last 6 months) which deals with a problem raised first by Erdős and Turán about small and large values of

$$\frac{d_{n+1}}{d_n}. \tag{5}$$

As it will be clear from the history and from the proof of our new result we will often refer to important and ingenious results of Helmut Maier, which often (but not exclusively) use his famous matrix method.

Since there were more than 40 papers discussing small and large values of d_n (or their relations) we will be relatively brief and try to concentrate on the most important developments.

Due to the prime number theorem,

$$\frac{1}{N}\sum_{n=1}^{N} d_n = \frac{p_{N+1}-2}{N} \sim \log N, \tag{6}$$

we have trivially

$$\Delta := \liminf_{n\to\infty} \frac{d_n}{\log n} \leq 1 \leq \limsup_{n\to\infty} \frac{d_n}{\log n} =: \lambda. \tag{7}$$

The present work appeared in its first version at the beginning of July 2014 at arXiv:1406.2658 [40], so it is based on our knowledge at that time. The very recent dramatic new developments are mentioned in Chapter (A*) of the Introduction.

(A) Large Gaps Between Primes (Until August 2014)

The first non-trivial result, $\lambda > 1$, was reached in 1929, when Backlund [1] proved $\lambda \geq 2$. One year later Brauer and Zeitz [5] improved this to $\lambda \geq 4$. Another year later E. Westzynthius [49] showed already $\lambda = \infty$, more precisely

$$\limsup_{n\to\infty} \frac{d_n/\log n}{\log_3 n/\log_4 n} \geq 2e^{\gamma}, \tag{8}$$

where $\log_\nu n$ denotes the ν-fold iterated logarithmic function.

Erdős [13] improved this in 1935 to

$$\limsup_{n\to\infty} \frac{d_n/\log n}{\log_2 n/(\log_3 n)^2} > 0. \tag{9}$$

Finally, 3 years later Rankin [45] succeeded in showing the function which is even currently, after more than 75 years the best result, apart from the constant:

$$\limsup_{n\to\infty} \frac{d_n/\log n}{\log_2 n \log_4 n/(\log_3 n)^2} \geq \frac{1}{3}. \tag{10}$$

Since, apart from the improvements of the constant $1/3$ to e^{γ} (in works of Ricci, Rankin and Schönhage between 1952 and 1963), the inequality (10) remained the strongest, Erdős offered in 1979, at a conference in Durham a prize of USD 10,000 for a proof of (10) with an arbitrarily large constant C in place of $1/3$; the highest prize ever offered by him for a mathematical problem. However, this did not help either in the past 35 years. Nevertheless, strong analytic methods introduced by Helmut Maier and Carl Pomerance, combined with the original Erdős–Rankin sieve procedure helped to prove [32] in 1990 (10) with $C = 1.3126\ldots e^{\gamma}$.

Finally, the current best result, the relation (10) with $1/3$ replaced by

$$C = 2e^{\gamma} \tag{11}$$

was reached by J. Pintz [36]. The methods to reach it involve besides the classical sieve methods and the large sieve used by Helmut Maier and Carl Pomerance, a pure graph-theoretical result, which, however, is proved in [36] using a so-called semi-random method of E. Szemerédi.

(A) Very Recent Developments Concerning Large Gaps Between Primes*

It was a great surprise when in August 2014 simultaneously and independently K. Ford, B. Green, S. Konyagin, T. Tao [21] and J. Maynard [34] showed the 76-year-old conjecture of Erdős mentioned in (A). In other words, this means that there exists some function $h(n) \to \infty$ as $n \to \infty$ such that

$$\limsup_{n\to\infty} \frac{d_n/\log n}{h(n)\log_2 n \log_4 n/(\log_3 n)^2} > 0. \tag{12}$$

Four months later in a joint work they showed that (K. Ford, B. Green, S. Konyagin, J. Maynard, T. Tao [22]) $h(n)$ can be chosen as $\log_3 n$, that is

$$\limsup_{n\to\infty} \frac{d_n/\log n}{\log_2 \log_4 n/\log_3 n} > 0. \tag{13}$$

(B) Chains of Large Gaps Between Consecutive Primes

In 1949 Erdős [16] proposed the problem whether k consecutive prime gaps d_{n+1}, $\ldots$, d_{n+k} can be simultaneously much larger than its mean value, i.e. whether

$$\limsup_{n\to\infty} \frac{\min(d_{n+1},\ldots,d_{n+k})}{\log n} = \infty, \tag{14}$$

and succeeded in showing this for $k = 2$ in the same work.

Thirty years later Helmut Maier [30] introduced his famous matrix method and showed this conjecture even in a stronger form when $\log n$ is replaced by the Erdős–Rankin function (cf. (10)). He proved for any k

$$\limsup_{n\to\infty} \frac{\min(d_{n+1},\ldots,d_{n+k})/\log n}{\log_2 n \log_4 n/(\log_3 n)^2} > 0. \tag{15}$$

(C) Small Gaps Between Consecutive Primes

Unlike the quick progress in the case of large gaps an analogue of the early result $\lambda \geq 4$ of Brauer–Zeitz from 1930, i.e. $\Delta \leq 1/4$ was essentially the best result still even 75 years later before 2005, and it was reached by Helmut Maier [31] also by his matrix method in 1985.

The first, although unpublished and conditional, result was reached by Hardy and Littlewood (see [46]) in 1926 who showed that the Generalized Riemann Hypothesis (GRH) implies $\Delta \leq 2/3$. However, the first non-trivial unconditional result was proved by Erdős [14],

$$\Delta \leq 1 - c_1, \tag{16}$$

with an unspecified explicitly calculable constant $c_1 > 0$.

After much work, calculating and improving c_1, the next breakthrough came by the large sieve of Bombieri [3] and Vinogradov [48] which enabled Bombieri and Davenport [4] to eliminate GRH and (incorporating also Erdős' ideas into their work) to show unconditionally

$$\Delta \leq (2+\sqrt{3})/8 = 0.466\ldots. \tag{17}$$

After five further improvements of Piltjai, Huxley and Fouvry–Grupp this was reduced to 0.4342. The next big step was the mentioned result of Helmut Maier [31], the inequality

$$\Delta \leq 0.2486. \tag{18}$$

Twenty years later D. Goldston, J. Pintz and C. Yıldırım succeeded in reaching the optimal value

$$\Delta = 0 \tag{19}$$

(see Primes in Tuples I and III in [24, 25]). Further they showed that d_n can be much smaller than $\log n$, in fact it was proved in [27] that

$$\liminf \frac{d_n}{(\log n)^c} = 0 \text{ for any } c > 1/2. \tag{20}$$

This was further improved by Pintz [39] for any $c > 3/7$ which in some sense was the limit of the original GPY method as proved by Farkas et al. [20].

Furthermore already the first work [24], proving $\Delta = 0$ in details contained also a conditional theorem. In order to formulate it we will describe the notion of admissibility for a k-tuple $\mathscr{H} = \{h_i\}_{i=1}^k$ of distinct integers.

Definition 1.1. $\mathscr{H} = \{h_i\}_{i=1}^k$ ($h_i \neq h_j$ for $i \neq j$) is called admissible if for any prime p the set $\mathscr{H}$ does not occupy all residue classes mod p.

This is equivalent with the formulation that $\prod_{i=1}^{k}(n+h_i)$ has no fixed prime divisor.

Remark 1.1. Although we usually suppose that $h_i \geq 0$ this is clearly not necessary in view of the consequence which is translation invariant.

We also introduced two connected conjectures which were named after Dickson, Hardy and Littlewood and which represent actually a weaker form of Dickson's prime k-tuple conjecture [11].

Conjecture DHL(k, k_0). *If $\mathscr{H}_k$ is admissible of size k, then the translated sets $n + \mathscr{H}_k$ contain for infinitely many n values at least k_0 primes if $N > N_0(k, k_0)$.*

The special case $k_0 = 2$ had special attention because of his connection with the

Bounded Gap Conjecture. $\liminf_{n\to\infty} d_n \leq C$ *with a suitable absolute constant C.*

The mentioned connection is the simple

Proposition 1.1. *If Conjecture* DHL$(k, 2)$ *is true for some k, then the Bounded Gap Conjecture is true.*

To formulate our conditional result we still need the following

Definition 1.2. A number ϑ is called a level of distribution of primes if for any $A, \varepsilon > 0$ we have

$$\sum_{q \le X^{\vartheta-\varepsilon}} \max_{\substack{a \\ (a,q)=1}} \left| \sum_{\substack{p \le X \\ p \equiv a (\mathrm{mod}\ q)}} \log p - \frac{X}{\varphi(q)} \right| \le \frac{C(A,\varepsilon)X}{(\log X)^A}. \tag{21}$$

The result that (21) holds with $\vartheta = 1/2$ is the famous Bombieri–Vinogradov theorem [3, 48], while Elliott–Halberstam [12] conjectured that even $\vartheta = 1$ is a level of distribution.

We can introduce for any $\vartheta \in (1/2, 1]$ the

Conjecture EH(ϑ). (21) *is true for* ϑ*, i.e.* ϑ *is a level of distribution of the primes.*

We showed in our original work

Theorem 1.1 ([24]). *If* EH(ϑ) *is true for any* $\vartheta > 1/2$*, then* DHL$(k, 2)$ *is true for* $k > C_1(\vartheta)$ *and consequently,* $\liminf d_n \le C_2(\vartheta)$.

Soon after this, Motohashi and Pintz [35] (MR 2414788 (2009d:1132), arXiv: math/0602599, Feb 27, 2006) showed in the work entitled "A Smoothed GPY sieve" that (21) can be substituted with the weaker condition that it holds for smooth moduli of q in (21) (by that we mean that for moduli q having all their prime factors below q^b—or even X^b—with an arbitrarily fixed b) and the maximum taken over all residue classes a with $(a, q) = 1$ can be reduced to those satisfying $\prod_{i=1}^{k} (a + h_i) \equiv 0$ (mod q) which are trivially the only cases appearing in the proof.

Finally, Zhang [50] showed that for $b = \frac{1}{292}$ the above condition holds with $\vartheta = \frac{1}{2} + \frac{1}{584}$.

This led to the very recent result.

Theorem 1.2 ([50]). DHL$(3.5 \cdot 10^6, 2)$ *is true and consequently* $\liminf\limits_{n\to\infty} d_n \le 7 \cdot 10^7$.

Some months later the joint effort of many mathematicians showed a stronger form of Zhang's theorem in the Polymath 8A project led by T. Tao [42].

Theorem 1.3 (Polymath 8A). DHL$(632, 2)$ *is true and consequently* $\liminf\limits_{n\to\infty} d_n \le 4680$.

Another improved version, proved independently by another refinement of the GPY method (in spirit closer to an elementary version of the method worked out in collaboration with S. W. Graham, see [26] and which was also close to the first attempt of Goldston and Yıldırım [23] which finally led to $\Delta \le 1/4$ in their version) was reached by Maynard [34], which showed the even stronger

Theorem 1.4 ([34]). DHL$(105, 2)$ *is true and consequently* $\liminf\limits_{n\to\infty} d_n \le 600$.

T. Tao used the same approach independently and simultaneously with Maynard and with help of his Polymath Project 8B [45] (involving further new theoretical ideas and a huge number of computations) they showed

Theorem 1.5 (Polymath 8B). *DHL*$(50, 2)$ *is true and consequently* $\liminf_{n\to\infty} d_n \leq 246$.

It is interesting to note that the above methods did not need the results or ideas of Motohashi–Pintz and Zhang, neither any weaker form of a result of type $\vartheta > 1/2$.

In complete contrast to this, the Maynard–Tao method shows the existence of infinitely many bounded gaps (although with weaker numerical bounds than 600 or 246) with any fixed positive distribution level of the primes. The first result of this type (that is, a positive distribution level of primes) was proved by Rényi in 1947–1948 [47].

(D) Chains of Bounded Gaps Between Consecutive Primes

The original GPY method [24] furnished only under the very strong original Elliott–Halberstam Conjecture EH(1) the bound

$$\Delta_2 := \liminf_{n\to\infty} \frac{p_{n+2} - p_n}{\log p_n} = 0. \tag{22}$$

Thus it failed to show on EH even DHL$(k, 3)$ for some k. The additional ideas of Motohashi–Pintz and Zhang helped neither.

Already Erdős mentioned [16] as a conjecture

$$\liminf_{n\to\infty} \frac{\min(d_n, d_{n+1})}{\log n} < 1 \tag{23}$$

which was proved in the mentioned work of Helmut Maier [31]. In the same work he has shown

$$\Delta_r = \liminf_{n\to\infty} \frac{p_{n+r} - p_n}{\log n} \leq e^{-\gamma} \left(r - \frac{5}{8} + o(1) \right) \quad (r \to \infty). \tag{24}$$

This was improved in [25] to

$$\Delta_r \leq e^{-\gamma} \left(\sqrt{r} - 1 \right)^2. \tag{25}$$

However, as mentioned already, even (22) was open unconditionally. In this aspect the Maynard–Tao method was much more successful, yielding

Theorem 1.6 (Maynard–Tao). $\liminf_{n\to\infty}(p_{n+r} - p_n) \le Ce^{4r}$ *for any* r *with an absolute constant* C.

2 Consecutive Values of d_n: Problems of Erdős, Turán, and Pólya

The questions discussed in Sect. 1 referred to single or consecutive small values of d_n or to analogous problems dealing with solely large values of d_n. In 1948 Erdős and Turán [19] showed that $d_{n+1} - d_n$ changes sign infinitely often. After this, still in the same year, Erdős [15] proved that

$$\liminf \frac{d_{n+1}}{d_n} < 1 < \limsup \frac{d_{n+1}}{d_n}. \tag{26}$$

He mentioned 60 years ago [17]: "One would of course conjecture that

$$\liminf_{n\to\infty} \frac{d_{n+1}}{d_n} = 0 \text{ and } \limsup_{n\to\infty} \frac{d_{n+1}}{d_n} = \infty, \tag{27}$$

but these conjectures seem very difficult to prove." Based on a generalization of the method of Zhang [50] the author proved (27) in [37].

In the mentioned work of Erdős and Turán [19] they also asked for a necessary and sufficient condition that

$$\sum_{i=1}^{k} a_i p_{n+i} \tag{28}$$

should change sign infinitely often as $n \to \infty$. They observed that the condition

$$\sum_{i=1}^{k} a_i = 0 \tag{29}$$

is clearly necessary. Using (29) one can reformulate the problem and ask for infinitely many sign changes of

$$\sum_{i=1}^{\ell} \alpha_i d_{n+i} = -\sum_{i=1}^{k} a_i p_{n+i} \tag{30}$$

if we use the notation

$$\alpha_j = \sum_{i=1}^{j} a_i \quad (j = 1, 2, \dots, k-1), \quad \ell = k-1. \tag{31}$$

This form shows an observation of Pólya (see [18]) according to which α_j ($j = 1, \dots, \ell$) cannot all have the same sign if (28) has infinitely many sign changes. Erdős [18] writes: "It would be reasonable to conjecture that Pólya's condition is necessary and sufficient for (30) to change sign infinitely often. Unfortunately the proof of this is not likely to succeed at the present state of science."

After this Erdős showed [18] that (28), i.e. (30) changes sign infinitely often if

$$\sum_{i=1}^{\ell} \alpha_i = 0, \quad \alpha_\ell \neq 0. \tag{32}$$

The author announced in [38] the proof of the following

Theorem 2.7 ([38]). *The sum* (28)*, i.e.* (30)*, changes sign infinitely often if at least one of the following conditions holds* ($\ell \geq 2$)

(i) $$\left|\sum_{i=1}^{\ell} \alpha_i\right| \leq c_0(\ell) \sum_{i=1}^{\ell} |\alpha_i|,$$
with a sufficiently small explicitly calculable constant $c_0(\ell)$ *depending on* ℓ*;*

(ii) *if* $\exists j \in [1, \ell]$ *such that*
$$\sum_{i=1, i\neq j}^{\ell} |\alpha_i| < |\alpha_j|, \quad sgn\, \alpha_i \neq sgn\, \alpha_j, \quad i \in [1, \ell] \setminus j; \tag{33}$$

(iii) *if the Hardy–Littlewood prime k-tuple conjecture is true for* $k = \ell$.

Now, (iii) shows that the mentioned conjecture of Erdős, namely that Pólya's trivial necessary condition (i.e., that all α_j cannot have the same sign) is probably really a necessary and sufficient condition for (30) to change sign infinitely often.

3 Results

In the present work we will show a kind of improvement of the result (10), which can be considered also as a common generalization of our result (27) and Maynard–Tao's mentioned theorem about the strongest known estimates of chains of small gaps between consecutive primes. The result will also give an improvement on the recent proof of the author [37] which solved the 60-year-old problem (27) of Erdős and improves the result (14) for $k = 2$ due to Erdős [16].

The proof will also show a simple method (already observed in [37] by the author after the proof of Zhang [50]) which makes the producing of bounded gaps (or chains of bounded gaps in case of Maynard and Tao) effective, since the original versions used a Bombieri–Vinogradov type theorem which again made use of the ineffective Siegel–Walfisz theorem.

Remark 3.2. The recent work [2] shows implicitly a way to make the Maynard–Tao theorem effective.

Finally we mention that concerning the Theorem [37] of Sect. 2 we can further give a very simple sufficient condition for (28), i.e. (30) to change sign infinitely often. To simplify the problem we can clearly suppose $\alpha_1 \neq 0$, $\alpha_\ell \neq 0$. In this case the sufficient condition is simply:

$$\operatorname{sgn} \alpha_1 \neq \operatorname{sgn} \alpha_\ell. \tag{34}$$

(If we do not make the trivial supposition $\alpha_1 \neq 0$, $\alpha_\ell \neq 0$ we can clearly formulate it in the way that the first and last non-zero elements of the sequence $\{\alpha_i\}_{i=1}^{\ell}$ should have opposite sign.)

Summarizing, we will prove the following results, using the basic notation $d_n = p_{n+1} - p_n$ of (1), $\log_\nu n$ for the ν-fold iterated logarithmic function.

Theorem 3.8. *Let k be an arbitrary fixed natural number, $c(k)$, $N(k)$ suitable positive, explicitly calculable constants depending only on k. Then for any $N > N(k)$ there exists an $n \in [N, 2N]$ such that*

$$\frac{d_{n+1}}{\max(d_n, \dots, d_{n-k+1})} > \frac{c(k) \log N \log_2 N \log_4 N}{(\log_3 N)^2}. \tag{35}$$

Essentially the same proof gives the analogous result:

Theorem 3.8′. *Under the conditions of Theorem 3.8 we have*

$$\frac{d_{n-k}}{\max(d_n, \dots, d_{n-k+1})} > \frac{c(k) \log N \log_2 N \log_4 N}{(\log_3 N)^2}. \tag{36}$$

Theorem 3.9. *Let k_0 be an arbitrary fixed natural number, $c(k_0)$, $N(k_0)$ suitable positive, explicitly calculable constants depending only on k_0. Then we have a $k \geq k_0$ such that for any $N > N(k_0)$ there exists an $n \in [N, 2N]$ such that*

$$\frac{\min(d_{n-k}, d_{n+1})}{\max(d_n, \dots, d_{n-k+1})} > \frac{c(k_0) \log N \log_2 N \log_4 N}{(\log_3 N)^2}.$$

Theorem 3.10. *Let $\ell \geq 2$ be an arbitrary integer, $\alpha_1, \dots, \alpha_\ell$ real numbers with $\alpha_1 \neq 0$, $\alpha_\ell \neq 0$. Then the expression*

$$\sum_{i=1}^{\ell} \alpha_i d_{n+i} \tag{37}$$

changes sign infinitely often as $n \to \infty$ if

$$sgn\, \alpha_1 \neq sgn\, \alpha_\ell. \tag{38}$$

We remark that Theorem 3.10 trivially follows from Theorems 3.8 and 3.8′.

4 Proofs

The proof requires a combination of the methods of Erdős–Rankin and that of Maynard–Tao. Since Theorem 3.10 follows from Theorems 3.8 and 3.8′, further Theorem 3.9 implies Theorems 3.8 and 3.8′, it is sufficient to prove Theorem 3.9. Concerning Theorem 3.9 we will show that we will have infinitely many cases when a block of at least k_0 consecutive primes in a bounded interval are preceded and followed by two primegaps of length at least

$$\frac{c_0(k_0)\log N\log_2 N\log_4 N}{(\log_3 N)^2} := c_0(k_0)\log Nf(\log N) \tag{39}$$

each, where the corresponding primes p'_νs satisfy $\nu \in [N, 2N]$. To be more specific we will choose an arbitrary set of m different primes $\{h_i\}_{i=1}^m$ with the property that for any $i, j, t \in [1, m]$, $i \neq j$:

$$m < C_3(m) < h_1 < h_2 < \cdots < h_m < C_4(m),\ \ m = \left\lceil C_5 e^{C_6 k_0}\right\rceil,\ \ h_t \nmid h_i - h_j \tag{40}$$

with the absolute constants C_5, C_6 and the constants $C_3(m)$ and $C_4(m)$ depending on m to be chosen later. Denoting

$$M = \prod_{p_i \le R,\ p_i \neq h_j\ (j=1,\ldots,m)} p_i, \tag{41}$$

we try to determine a residue class $z\ (\mathrm{mod}\ M)$ with the property

$$(z + h_i, M) = 1 \quad (1 \le i \le m), \tag{42}$$

$$(z \pm \nu, M) > 1 \ \text{ for all } \ \nu \in [0, c_0(k_0)Rf(R)] \setminus \{h_i\}_{i=1}^m. \tag{43}$$

We remark that we will have $Rf(R) \asymp \log Nf(\log N)$.

To describe the procedure briefly we will look for a block of at least k_0 primes among the numbers

$$\ell + h_1, \ldots, \ell + h_m, \quad \ell \equiv z\ (\mathrm{mod}\ M) \tag{44}$$

by the Maynard–Tao method, using the refinements in [2]. This will contain at least k_0 consecutive primes in a bounded block, while the two intervals

$$[\ell - c_0(k_0)Rf(R), \ell - h_m) \ \text{ and } \ (\ell + h_m, \ell + c_0(k_0)Rf(R)] \tag{45}$$

will certainly not contain any prime due to (43). Since (cf. [2]) we can choose M as large as a suitable small power of N, this will prove our theorem as by the Prime Number Theorem we will have

$$\log N \asymp \log M \sim R, \ \ f(\log N) \sim f(R) \tag{46}$$

and consequently $Rf(R) \asymp \log Nf(\log N)$.

Let us use parameters $0 < v < w < R/2$ to be chosen later, and let

$$P_1 := \prod_{p \le v,\ p \ne h_j\ (1 \le j \le m)} p, \tag{47}$$

$$P_2 := \prod_{v < p \le w} p, \tag{48}$$

$$P_3 := \prod_{w < p \le R/2} p, \tag{49}$$

$$P_4 := \prod_{R/2 < p \le R} p, \tag{50}$$

$$v = \log^3 R, \ U = c_0(k_0)Rf(R), \ P := P(R) = P_1P_2P_3P_4, \tag{51}$$

with a suitably small $c(k_0)$, chosen at the end. If we have a Siegel-zero, that is, a real primitive character $\chi_1 \mod r$, $r \le N$, with a zero β of the corresponding $\mathscr{L}$-function $\mathscr{L}(s, \chi)$ satisfying

$$\beta > 1 - \frac{c_7}{\log N}, \tag{52}$$

then this character and zero are uniquely determined by the Landau–Page theorem (see, e.g., [9, p. 95]) if c_7 is chosen as a suitably small explicitly calculable positive absolute constant. If χ_1 is real primitive $\mod r$, then r has to be square-free apart from the possibility that the prime 2 appears on the second or third power in r. Let us denote the greatest prime factor of r by q. Since we have

$$1 - \beta \gg \frac{\log^2 r}{\sqrt{r}} \tag{53}$$

(see [9, p. 96]) where the implied constant is explicitly calculable we have by (52)

$$r \gg (\log N)^{3/2}. \tag{54}$$

Consequently, by the Prime Number Theorem and the "almost square-free" property of r we have

$$q \gg \log r \gg \log_2 N \to \infty \text{ as } N \to \infty. \tag{55}$$

If there is no Siegel-zero, then we will let $q = 1$. Now we change slightly the definition of P_i and P as to exclude from their defining product the possible divisor q. Let

$$P_i^* = P_i/q \text{ if } q \mid P_i \ (1 \le i \le 4), \text{ otherwise } P_i^* = P_i,\ P^* = \prod_{i=1}^{4} P_i^*. \tag{56}$$

This change is not necessary for finding z with (42)–(43) but to assure uniform distribution of primes $\equiv z \pmod M$ in the Maynard–Tao procedure (and to make later the whole procedure effective as well). This means also that in contrast with the rough description at the beginning of this section the role of $M = P$ will be played actually by

$$M^* = P^* = P/q \text{ if } q \mid P \text{ (otherwise } M^* = M = P = P^*). \tag{57}$$

Returning with this change to our problem of choosing $z \bmod P^*$ let

$$z \equiv 0 \pmod{P_1^* P_3^*} \tag{58}$$

while we will choose $z \pmod{P_2^* P_4^*}$ suitably later. The letter p will denote in the following an unspecified prime. The choices of v and U in (51) guarantee that

$$(z \pm n, P_1^* P_3^*) = 1 \quad (0 < n \le U,\ n \ne 1) \tag{59}$$

if and only if either

$$n = pq^{\alpha} \prod_{i=1}^{m} h_i^{\alpha_i} \ (\alpha \ge 0,\ \alpha_i \ge 0) \text{ and } p \ge R/2 \tag{60}$$

or

$$n \text{ is composed only of primes } p \mid P_2^* q \prod_{i=1}^{m} h_i. \tag{61}$$

Our first (and main) goal in finding $z \bmod P^*$ will be to find residues α_p with $p \mid P_2^*$ so that choosing

$$z \equiv \alpha_p \pmod p \text{ for } p \mid P_2^* \tag{62}$$

the condition $(z^* \pm h_i, P^*) = 1$ (cf. (42)) should remain true and simultaneously we should have for as many as possible numbers n of the form (60) $(z+n, P^*) > 1$. By a suitable choice of the parameter w we can make the whole set of n's satisfying (61) relatively small, due to de Bruijn's result [10] which we use in a weaker form, proved by Rankin [45]. The present form is Lemma 5 of [30].

Lemma 4.1. *Let $\Psi(x,y)$ denote the number of positive integers $n \le x$ which are composed only of primes $p \le y$. For $y \le x$ and y approaching to infinity with x, we have*

$$\Psi(x,y) \le x \exp\left[-\frac{\log_3 y}{\log y}\log x + \log_2 y + O\left(\frac{\log_2 y}{\log_3 y}\right)\right]. \tag{63}$$

Since $\Psi(x,y)$ is clearly monotonically increasing in x we can estimate the number of n's in (61) from above by

$$(\log U)^{m+1}\Psi(U,w). \tag{64}$$

Now, choosing

$$w = \exp\left[\alpha(\log R \log_3 R/\log_2 R)\right] \tag{65}$$

with a constant α to be determined later, the quantity in (64) is by (51) clearly

$$\begin{aligned} &\ll U\exp\left[-\frac{(1+o(1))\log_3 R\log R}{\alpha\log R\log_3 R/\log_2 R} + (1+o(1))\log_2 R + (m+1)\log_2 R\right] \\ &\ll \frac{U}{\log^2 U} = o\left(\frac{R}{\log R}\right) \end{aligned} \tag{66}$$

if we chose, e.g.,

$$\alpha = \frac{1}{m+5}. \tag{67}$$

This bound will be completely satisfactory for us. So we have to concentrate on the numbers n in (60).

The simplest strategy in choosing $\alpha_p \pmod p$ in (62) would be to sieve out for the smallest $p > v$ the largest residue class $\bmod p$ among the N_0 numbers of the form (60) and repeat this consecutively for all primes in P_2^*. What makes our task more complicated, is the fact that we have to preserve the condition $(z+h_i, M) = 1$ in (42), so we have to choose for any $\tilde{p}_j \in P_2^*$

$$\alpha_{\tilde{p}_j} \not\equiv -h_i \pmod{\tilde{p}_j} \quad \text{for} \quad i = 1,2,\dots,m. \tag{68}$$

First we observe that an easy calculation shows that the Prime Number Theorem implies

$$\begin{aligned} N_0 &= \sum_{\alpha,\alpha_1,\ldots,\alpha_m\geq 0} \pi\left(\frac{U}{q^\alpha \prod_{i=1}^m h_i^{\alpha_i}}\right) - \pi\left(\frac{R}{2q^\alpha \prod_{i=1}^m h_i^{\alpha_i}}\right) \\ &\sim \frac{U}{\log U}\left(1+\frac{1}{q}+\frac{1}{q^2}+\ldots\right)\prod_{i=1}^m\left(1+\frac{1}{h_i}+\frac{1}{h_i^2}+\ldots\right) \\ &\sim \frac{U}{\log U}\left(1-\frac{1}{q}\right)^{-1}\prod_{i=1}^m\left(1-\frac{1}{h_i}\right)^{-1} \leq \frac{2U}{\log U} \end{aligned} \tag{69}$$

if $C_3(m)$ in (40) was chosen sufficiently large (taking into account also (55)).

We will choose the residue classes α_p one by one for all primes from v to w and consider at the ith step the arising situation. We will be left at the jth step with $N_j < N_0$ remaining values of n's with (60). If N_j is at any stage of size $\leq \dfrac{R}{5\log R}$, then we are ready. Thus we can suppose $N_j > R/(5\log R)$.

Otherwise, if α and all α_i $(1 \leq i \leq m)$ are determined, then we have trivially in case of

$$q^\alpha \prod_{i=1}^m h_i^{\alpha_i} > \sqrt{U} \tag{70}$$

at any rate altogether at most

$$O\left(\sqrt{U}(\log U)^{m+1}\right) \tag{71}$$

numbers of the given form (60).

On the other hand, if (70) is not true, then we have by the Brun–Titchmarsh theorem altogether at most

$$\frac{3mU}{\varphi\left(\tilde{p}_j q^\alpha \prod_{i=1}^m h_i^{\alpha_i}\right)\log\sqrt{U}} \leq \frac{8Um}{\tilde{p}_j q^\alpha\left(\prod_{i=1}^m h_i^{\alpha_i}\right)\log U} \tag{72}$$

numbers n of form (60) which lie in one of the m bad residue classes $\{h_i\}_{i=1}^m \bmod \tilde{p}_j$. Adding (72) up for all possible non-negative combination of $\alpha, \alpha_1, \ldots, \alpha_m$ we get altogether (cf. (69)) at most

$$\frac{16mU}{\tilde{p}_j \log U} < \frac{16mc_0(k_0)Rf(R)}{\tilde{p}_j \log R} < \frac{N_j f(R)}{\tilde{p}_j} < \frac{N_j}{4\log^2 R} \tag{73}$$

bad n values of the form (60) by $\tilde{p}_j > w$ and $N_j > R/(5\log R)$. This means that choosing from the remaining $\tilde{p}_j - m$ residue classes that one which sieves out the most elements from the remaining set of size N_j we obtain for the size of the new remaining set

$$\begin{aligned} N_{j+1} &< N_j - \frac{N_j\left(1-\frac{1}{4\log^2 R}\right)}{\tilde{p}_j - m} \\ &< N_j\left(1 - \frac{1-\frac{1}{4\log^2 R}}{\tilde{p}_j}\right) \\ &< N_j\left(1-\frac{1}{\tilde{p}_j}\right)^{1-\log^{-2} R}. \end{aligned} \tag{74}$$

This means that by Mertens' theorem we obtain a final residual set of cardinality at most

$$\begin{aligned} N^* &< N_0 \prod_{\substack{v<p<w \\ p\neq q}} \left(1-\frac{1}{p}\right)^{1-\log^{-2} R} \sim N_0\left(\frac{\log v}{\log w}\right)^{1-\log^{-2} R} \\ &= N_0\left(\frac{3\log_2^2 R(1+o(1))}{\alpha \log R \log_3 R}\right)^{1-\log^{-2} R} \\ &\le 4mN_0 \frac{\log_2^2 R}{\log R \log_3 R} \le \frac{8mU}{\log U}\cdot\frac{1}{f(R)} \\ &= \frac{8mc_0(k_0)R}{\log U} < \frac{8mc_0(k_0)R}{\log R} < \frac{\pi(R)-\pi(R/2)}{4} \end{aligned} \tag{75}$$

if the value of $c_0(k_0)$ in (39) was chosen sufficiently small.

Formulas (66) and (69) mean that after the sieving with suitably chosen $\alpha_{\tilde{p}_j} \pmod{\tilde{p}_j}$, $\tilde{p}_j \in P_2^*$ and taking into account that the set satisfying (61) was at any rate small, all elements of (60) and (61) can be sieved out by one third of the primes from P_4 choosing for every remaining $n^* = pq^\alpha \prod_{i=1}^{m} h_i^{\alpha_i}$ in (60) and (61) a separate $p^* \in P_4$, $p^* \neq p, q$ and $z \equiv -n^* \pmod{p^*}$, that is, $\alpha_{p^*} \equiv -n^* \pmod{p^*}$. This will assure $(z+n^*, M^*) > 1$. With an other third of $p^* \mid P_4$ we can similarly sieve out the remaining negative n^* values. The only problem is that for any $p^* \in P_4$ used above, simultaneously with $z + n^* \equiv 0 \pmod{p^*}$ we have to assure $z + h_i \not\equiv 0 \pmod{p^*}$. This is equivalent to that $n^* - h_i \equiv 0 \pmod{p^*}$ is not allowed if we want to sieve out $z + n^*$ with p^*. However, for every number $|n^*| \le U$ and for every $i \in [1, m]$ the number $n^* - h_i$ has at most one prime divisor $> R/2 > \sqrt{U}$. This means that for every n^* we have at most m forbidden primes to use to sieve out n^*. Thus, as the number of still "abundant" primes is at every step at least one third of all primes

between $(R/2, R]$ we can for all (positive and negative) values of n^* consecutively choose the primes $p^* \mid P_4^*$ nearly freely, just avoiding at most m forbidden primes at each step, so that we would have, step by step for each $p^* \mid P_4^*$, n^*

$$z + n^* \equiv 0 \ (\mathrm{mod}\ p^*), \quad z + h_i \not\equiv 0 \ (\mathrm{mod}\ p^*) \quad (i = 1, 2, \ldots, m). \tag{76}$$

At the end of this procedure some $p^* \mid P_4$ will remain. We can choose for these primes α_p^* freely with the only condition that

$$\alpha_p^* \not\equiv -h_i \ (\mathrm{mod}\ p^*) \quad (i = 1, 2, \ldots, m) \tag{77}$$

should hold, in order to assure for the remaining primes $p^* \mid P_4$ also

$$(z + h_i, p^*) = 1. \tag{78}$$

Finally to determine z uniquely $\mathrm{mod}\ M/q$ we choose for the primes $p = h_i\ \alpha_p$ again essentially freely with the only condition

$$\alpha_p \not\equiv -h_j \ (\mathrm{mod}\ p) \quad (j = 1, 2, \ldots, m) \tag{79}$$

which is again possible by $h_1 > m$. In this way we will have finally for any $i = 1, 2, \ldots, m$ with $M^* = P^* = M/q = P/q$

$$(z + h_i, P^*) = 1 \tag{80}$$

and this means that the relations (42)–(43) will be satisfied with M^* in place of M.

Now, we continue with the Maynard–Tao proof which we take over from [34] with the additional changes executed in [2]. We will look for primes among the numbers

$$\ell + h_1, \ldots, \ell + h_m, \quad \ell \equiv z \ (\mathrm{mod}\ W), \quad W = M^* = P^* \tag{81}$$

with the value z found in the previous procedure, satisfying with the notation $\mathscr{H}_m = \{h_i\}_{i=1}^m$

$$(z + h_i, W) = 1 \quad (i = 1, 2, \ldots, m) \tag{82}$$

$$(z + \nu, W) > 1 \ \text{ if } \ |\nu| \in \left[0, \log Lf(\log L)\right] \setminus \mathscr{H}_m. \tag{83}$$

We will show that for $L > L(k_0)$ we will find numbers

$$\ell \in [L, 2L], \quad \ell \equiv z \ (\mathrm{mod}\ W), \quad \#\left\{\ell + h_i \in \mathscr{P}\ (1 \le i \le m)\right\} \ge k_0 \tag{84}$$

with (81)–(83) which will prove our Theorem 3.9.

Remark 4.3. If there are additional primes among $\ell - h_i$ $(1 \le i < m)$ this does not change anything since $h_i \le C_4(m) \le C_8(k_0)$. The introduction of the new variables ℓ and L is only necessary since we look for primes around ℓ instead of investigating p_n and d_n, so essentially $\ell \sim n \log n$, $L \sim N \log N$, $\log L \sim \log N$, so the size of gaps (39) remains unchanged if we substitute N by L. We remark that (40) and (55) assure

$$p \Bigg| \prod_{\substack{i,j=1 \\ i \ne j}}^{m} (h_i - h_j) \Longrightarrow p = \mathcal{O}_m(1),\ p \ne h_t\ (1 \le t \le m) \Longrightarrow p \nmid W, \tag{85}$$

analogously to (65)–(69) of Banks et al. [2], which is actually the Maynard–Tao theorem.

We need also the modified Bombieri–Vinogradov theorem, Theorem 4.1 of Banks et al. [2], which is somewhat similar to, but stronger than Theorem 6 of [27]. The greatest prime factor of W is $\ll \log L$, so the level of smoothness of W is much better than that (L^{ε}) required by the condition of Theorem 10 of Chang [6]. So the whole Theorem 4.1 of Banks et al. [2] will remain true. Further, we can leave out from the actual sieving procedure the possibly existing q and its multiples. This causes just a negligible change of size $\left(1 - \frac{1}{q}\right)^{O(1)} = 1 + O\left(\frac{1}{\log L}\right)$ in the weighted number of primes and in the sum of weights. If we choose in (41)

$$R \le c_9(k_0) \log L \tag{86}$$

with a sufficiently small $c_9(k_0) > 0$, then we will have by the Prime Number Theorem

$$W \le L^{c_9(k_0)(1+o(1))} \tag{87}$$

so the whole Maynard–Tao procedure will remain valid, as in [2]. (The variable R is here completely different from that in [34] or [2].)

Summarizing, we will obtain at least k_0 bounded prime gaps in intervals of type

$$[\ell - h_m, \ell + h_m] \tag{88}$$

and around them two intervals of size $(1 + o(1))c_0(k_0) \log L \log_2 L \log_4 L/(\log_3^2 L)$ containing only composite numbers (see (45)) which prove Theorem 3.9.

Finally, the fact that we left out the largest prime factor q (and its multiples) of the possibly existing exceptional modulus r yields an effective modified Bombieri–Vinogradov theorem and so an effective final Theorem 3.9, consequently Theorems 3.8, 3.8′, and 3.10 are effective too.

Acknowledgements The author would like to express his sincere gratitude to Imre Z. Ruzsa, who called his attention that a possible combination of the methods of Erdős–Rankin and Zhang–Maynard–Tao might lead to stronger results about the ratio of consecutive primegaps than those proved by the author in his earlier work [37].

This work was supported by OTKA Grants NK104183, K100291 and ERC-AdG. 321104.

References

1. R.J. Backlund, Über die Differenzen zwischen den Zahlen, die zu den ersten n Primzahlen teilerfremd sind. Commentationes in honorem E. L. Lindelöf. Ann. Acad. Sci. Fenn. **32**(Nr. 2), 1–9 (1929)
2. W.D. Banks, T. Freibert, J. Maynard, On limit points of the sequence of normalized prime gaps. arXiv: 1404.5094v1 [math. NT] 21 April 2014
3. E. Bombieri, On the large sieve. Mathematika **12**, 201–225 (1965)
4. E. Bombieri, H. Davenport, Small differences between prime numbers. Proc. R. Soc. Ser. A **293**, 1–18 (1966)
5. A. Brauer, H. Zeitz, Über eine zahlentheoretische Behauptung von Legendre. Sber. Berliner Math. Ges. **29**, 116–125 (1930)
6. M.C. Chang, Short character sums for composite moduli. J. Anal. Math. **123**, 1–33 (2014)
7. H. Cramér, Prime numbers and probability. Skand. Math. Kongr. Stockholm **8**, 107–115 (1934)
8. H. Cramér, On the order of magnitude of the difference between consecutive prime numbers. Acta Arith. **2**, 23–46 (1936)
9. H. Davenport, *Multiplicative Number Theory*, 2nd edn. (Springer, New York, 1980). Revised by H.L. Montgomery
10. N.G. de Bruijn, On the number of positive integers $\leq x$ and free of prime factors $> y$. Indag. Math. **13**, 50–60 (1951)
11. L.E. Dickson, A new extension of Dirichlet's theorem on prime numbers. Messenger Math. **33**(2), 155–161 (1904)
12. P.D.T.A. Elliott, H. Halberstam, A conjecture in prime number theory, in *Symposia Mathematica*, vol. 4, INDAM, Rome, 1968/1969 (Academic, London, 1970), pp. 59–72
13. P. Erdős, On the difference of consecutive primes. Q. J. Math. Oxford Ser. **6**, 124–128 (1935)
14. P. Erdős, The difference of consecutive primes. Duke Math. J. **6**, 438–441 (1940)
15. P. Erdős, On the difference of consecutive primes. Bull. Am. Math. Soc. **54**, 885–889 (1948)
16. P. Erdős, Problems and results on the differences of consecutive primes. Publ. Math. Debrecen **1**, 33–37 (1949)
17. P. Erdős, Some problems on the distribution of prime numbers. C.I. M.E., Teoria dei Numeri, p 8 (1955)
18. P. Erdős, Some problems on consecutive prime numbers. Mathematika **19**, 91–95 (1972)
19. P. Erdős, P. Turán, On some new questions on the distribution of prime numbers. Bull. Am. Math. Soc. **54**, 371–378 (1948)
20. B. Farkas, J. Pintz, S. Révész, On the optimal weight function in the Goldston–Pintz–Yıldırım method for finding small gaps between consecutive primes, in *Paul Turán Memorial Volume: Number Theory, Analysis and Combinatorics* (de Gruyter, Berlin, 2014), pp. 75–104
21. K. Ford, B. Green, S. Konyagin, T. Tao, Large gaps between consecutive prime numbers. (2014)
22. K. Ford, B. Green, S. Konyagin, J. Maynard, T. Tao, Long gaps between primes. (2014)
23. D.A. Goldston, C. Yıldırım, Higher correlations of divisor sums related to primes. III. Small gaps between primes. Proc. Lond. Math. Soc. **95**(3), 653–686 (2007)
24. D.A. Goldston, J. Pintz, C. Yıldırım, Primes in tuples III: on the difference $p_{n+\nu} - p_n$. Funct. Approx. Comment. Math. **35**, 79–89 (2006)
25. D.A. Goldston, J. Pintz, C. Yıldırım, Primes in tuples I. Ann. Math. **170**(2), 819–862 (2009)

26. D.A. Goldston, S.W. Graham, J. Pintz, C.Y. Yıldırım, Small gaps between primes or almost primes. Trans. Am. Math. Soc. **361**(10), 5285–5330 (2009)
27. D.A. Goldston, J. Pintz, C. Yıldırım, Primes in tuples II. Acta Math. **204**(1), 1–47 (2010)
28. A. Granville, Unexpected irregularities in the distribution of prime numbers, in *Proceedings of the International Congress of Mathematicians* (Zürich, 1994), vols. 1, 2, (Birkhäuser, Basel, 1995), pp. 388–399
29. A. Granville, Harald Cramér and the distribution of prime numbers. Scand. Actuarial J. **1995**(1), 12–28 (1995)
30. H. Maier, Chains of large gaps between consecutive primes. Adv. Math. **39**(3), 257–269 (1981)
31. H. Maier, Small differences between prime numbers. Mich. Math. J. **35**, 323–344 (1988)
32. H. Maier, C. Pomerance, Unusually large gaps between consecutive primes. Trans. Am. Math. Soc. **322**, 201–237 (1990)
33. J. Maynard, Large gaps between primes. (2014)
34. J. Maynard, Small gaps between primes. Ann. of Math. **181**(1), 383–413 (2015)
35. Y. Motohashi, J. Pintz, A smoothed GPY sieve. Bull. Lond. Math. Soc. **40**(2), 298–310 (2008)
36. J. Pintz, Very large gaps between consecutive primes. J. Number Theory **63**, 286–301 (1997)
37. J. Pintz, Polignac numbers, conjectures of Erdős on gaps between primes and the bounded gap conjecture. arXiv: 1305.6289 [math.NT]. 27 May 2013
38. J. Pintz, Paul Erdős and the difference of primes, in *Erdős Centennial*, ed. by L. Lovász, I.Z. Ruzsa, V.T. Sós. Bolyai Society Mathematical Studies, vol. 25 (Springer, Berlin, 2013), pp. 485–513
39. J. Pintz, Some new results on gaps between consecutive primes, in *Paul Turán Memorial Volume: Number Theory, Analysis and Combinatorics* (de Gruyter, Berlin, 2014), pp. 261–278
40. J. Pintz, On the ratio of consecutive gaps between primes. (2014)
41. D.H.J. Polymath, A new bound for gaps between primes (Preprint)
42. D.H.J. Polymath (in preparation)
43. Polymath, D.H.J.; Castryck, Wouter; Fouvry, Étienne; Harcos, Gergely; Kowalski, Emmanuel; Michel, Philippe; Nelson, Paul; Paldi, Eytan; Pintz, János; Sutherland, Andrew V.; Tao, Terence; Xie, Xiao-Feng New equidistribution estimates of Zhang type. Algebra Number Theory **8**(9), 2067–2199 (2014)
44. D.H.J. Polymath, Variants of the Selberg sieve, and bounded intervals containing many primes, (2014) arXiv:1407.4897v1
45. R.A. Rankin, The difference between consecutive prime numbers. J. Lond. Math. Soc. **13**, 242–244 (1938)
46. R.A. Rankin, The difference between consecutive prime numbers. II. Proc. Camb. Philos. Soc. **36**, 255–266 (1940)
47. A. Rényi, On the representation of an even number as the sum of a single prime and a single almost-prime number. Dokl. Akad. Nauk SSSR **56**, 455–458 (1947) [Russian]
48. A.I. Vinogradov, The density hypothesis for Dirichlet *L*-series. Izv. Akad. Nauk. SSSR **29**, 903–934 (1965) [Russian]. Corr.: ibidem, **30**, 719–720 (1966)
49. E. Westzynthius, Über die Verteilung der Zahlen, die zu der *n* ersten Primzahlen teilerfremd sind. Commun. Phys. Math. Helsingfors **25**(5), 1–37 (1931)
50. Y. Zhang, Bounded gaps between primes. Ann. Math.(2) **179**(3), 1121–1174 (2014)

Remarks on Fibers of the Sum-of-Divisors Function

Paul Pollack

To Professor Helmut Maier on his 60th birthday

Abstract Let σ denote the usual sum-of-divisors function. We show that every positive real number can be approximated arbitrarily closely by a fraction m/n with $\sigma(m) = \sigma(n)$. This answers in the affirmative a question of Erdős. We also show that for almost all of the elements v of $\sigma(\mathbf{N})$, the members of the fiber $\sigma^{-1}(v)$ all share the same largest prime factor. We describe an application of the second result to the theory of L.E. Dickson's amicable tuples, which are a generalization of the ancient notion of an amicable pair.

1 Introduction

Let $\sigma(n) := \sum_{d|n} d$ be the familiar sum-of-divisors function. In this paper, we record two theorems concerning the fibers of σ. The first of these answers in the affirmative a 1959 question of Erdős (see [5, p. 172]).

Theorem 1.1. *Let $\beta > 0$. For every $\epsilon > 0$, one can find integers m and n with $\sigma(m) = \sigma(n)$ and $|\frac{m}{n} - \beta| < \epsilon$.*

The primary tool in the proof will be the remarkable recent theorem of Yitang Zhang approximating the prime k-tuples conjecture.

Our second result concerns the multiplicative structure of elements belonging to a typical fiber. Building on work of Maier and Pomerance [8], Ford [6] developed an extensive theory of σ-preimages and used it to answer a number of delicate questions about the distribution of φ and σ-values. For example, he showed that the count of σ-values in $[1, x]$ is

$$\asymp \frac{x}{\log x} \exp\left(C(\log_3 x - \log_4 x)^2 + D\log_3 x - (D + \frac{1}{2} - 2C)\log_4 x\right)$$

P. Pollack (✉)
University of Georgia, Boyd Graduate Studies Building, Athens, GA 30602, USA
e-mail: pollack@uga.edu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_18

for certain constants $C \approx 0.8178146$ and $D \approx 2.1769687$, and that precisely the same estimate holds for the φ-function. Here we show how Ford's methods can be adapted to prove the following theorem.

Theorem 1.2. *For asymptotically* 100 % *of the values* v *in the image of the* σ*-function, all of the elements of the set* $\sigma^{-1}(v)$ *share the same largest prime factor.*

"Asymptotically 100 %" means that the density of such v, relative to $\sigma(\mathbf{N})$, is 1.

Theorem 1.2 has an amusing consequence for a problem of L.E. Dickson. Recall that m and n form an *amicable pair* if $\sigma(m) = \sigma(n) = m + n$. Dickson [1] (cf. [2, p. 50]) proposed the following generalization: Say that $n_1, \dots, n_k$ form an *amicable* k*-tuple* if $\sigma(n_i) = n_1 + n_2 + \ldots + n_k$ for each $i \in \{1, 2, \dots, k\}$. (Below, we refer to v as the *common* σ*-value* of the amicable tuple.) Dickson gave a handful of examples with $k = 3$. For several others, with k as large as 6, see [9, 10, 15].

The distribution of amicable tuples remains mysterious. For example, we do not know if there are infinitely many amicable tuples, even if all k are considered simultaneously. In the case $k = 2$ (the amicable pair case), some progress has been made. In 1955, Erdős [4] showed that the set of natural numbers belonging to an amicable pair has density zero. This result has been steadily sharpened over the years [3, 12–14]. However, when $k > 2$, we still do not know if the set of numbers belonging to an amicable k-tuple has density zero. Thus, the following corollary of Theorem 1.2 seems of some interest.

Corollary 1.1. *Asymptotically* 0 % *of the elements in the range of the* σ*-function appear as the common* σ*-value of an amicable tuple.*

1.1 Notation

We reserve the letters p and q for primes. We let $P^+(n)$ denote the largest prime factor of n, with the convention that $P^+(1) = 1$. We use $\omega(n)$ for the number of distinct primes dividing n and $\Omega(n)$ for the number of prime factors of n counted with multiplicity. We write $\Omega(n, U, T)$ for the number of primes dividing n with $U < p \le T$, again counted with multiplicity. We write $\log_k$ for the kth iterate of the natural logarithm.

2 Proof of Theorem 1.1

Let $\mathbf{L}$ be the closure of the set $\{\log \frac{m}{n} : \sigma(m) = \sigma(n)\}$. Theorem 1.1 amounts to the claim that $\mathbf{L} = \mathbf{R}$. Since $\log \frac{m}{n} = -\log \frac{n}{m}$, the set $\mathbf{L}$ is symmetric about 0, and so it is enough to show that $\mathbf{L}$ contains all nonnegative real numbers.

For any finite set of primes $\mathscr{P}$, define $\mathbf{L}^{\mathscr{P}}$ in the same way as $\mathbf{L}$ but with m and n restricted to be divisible by none of the primes in $\mathscr{P}$. The next lemma is fundamental.

Lemma 2.1. *There is a natural number K for which the following holds: Let $\mathscr{P}$ be a finite set of primes. Let $\alpha_1 < \ldots < \alpha_K$ be any K real numbers. Then $\alpha_j - \alpha_i \in \mathbf{L}^{\mathscr{P}}$ for some pair of i and j with $1 \le i < j \le K$.*

The proof requires that we recall Dickson's prime k-tuples conjecture and the spectacular recent progress towards it made by Zhang [18]. A collection of linear polynomials $a_1x + b_1, \ldots, a_kx + b_k \in \mathbf{Z}[x]$, each with positive leading coefficient, is called *admissible* if $\gcd\{\prod_{i=1}^{k}(a_in + b_i)\}_{n\in\mathbf{Z}} = 1$. Dickson conjectured that for any admissible collection, there are infinitely many n for which all of the $a_in + b_i$ are simultaneously prime.

Zhang's breakthrough result in this direction was the following:

Proposition 2.1. *There is a natural number K for which the following holds: Suppose that $a_1x + b_1$, $a_2x + b_2$, $\ldots$, $a_Kx + b_K$ is admissible. For some pair of i and j with $1 \le i < j \le K$, the expressions $a_in + b_i$ and $a_jn + b_j$ simultaneously represent primes for infinitely many natural numbers n.*

Zhang states his result only in the case when all $a_i = 1$. A stronger version of the proposition—explicitly stated for general linear forms—appears in recent work of Maynard (see [11, Theorem 3.1]).

Proof (Lemma 2.1). Assuming Dickson's conjecture, Schinzel and Sierpiński [16, p. 193] showed that there are solutions to $\sigma(m) = \sigma(n)$ with the ratio m/n arbitrarily large. We modify their approach to demonstrate Lemma 2.1. Our proof will show that if K is acceptable in Proposition 2.1, then it is also acceptable in Lemma 2.1.

The set $\{\log \frac{\sigma(n)}{n} : \gcd(n, \prod_{p\in\mathscr{P}} p) = 1\}$ is dense in $[0, \infty)$. Indeed, if f is any nonnegative additive function for which (1) $\sum_p f(p)$ diverges, and (2) $f(p) \to 0$ along primes p, then the values of $f(n)$, with n squarefree, are dense in $[0, \infty)$. (This follows from a straightforward application of the greedy algorithm.) We apply this general fact with f any additive function having $f(p) = 0$ for $p \in \mathscr{P}$ and $f(p) = \log \frac{\sigma(p)}{p}$ for $p \notin \mathscr{P}$.

We can assume all $\alpha_i > 0$, by replacing each α_i with $\alpha_i + \alpha_0$ for a suitably large α_0. For each $1 \le i \le K$, fix a sequence $\{A_j^{(i)}\}_{j=1}^{\infty}$ of integers coprime to $\prod_{p\in\mathscr{P}} p$ satisfying

$$\lim_{j\to\infty} \log \frac{\sigma(A_j^{(i)})}{A_j^{(i)}} = \alpha_i.$$

For each $j \in \mathbf{N}$, we apply Proposition 2.1 to the collection $\{\sigma(A_j^{(i)})x - 1\}_{i=1}^{K}$. This is an admissible collection, since the product of the polynomials at $x = 0$ is ± 1. By Proposition 2.1, we can find a natural number n_j, and integers $1 \le a_j < b_j \le K$, for which $p_j := \sigma(A_j^{(a_j)})n_j - 1$ and $q_j := \sigma(A_j^{(b_j)})n_j - 1$ are simultaneously

prime. Moreover, we can assume that p_j and q_j are both larger than j, larger than $A_j^{(1)}, \dots, A_j^{(K)}$, and larger than any element of $\mathscr{P}$. Observe that

$$\frac{p_j+1}{q_j+1} = \frac{\sigma(A_j^{(a_j)})}{\sigma(A_j^{(b_j)})}, \quad \text{so that} \quad \sigma(p_j A_j^{(b_j)}) = \sigma(q_j A_j^{(a_j)});$$

also, both $p_j A_j^{(b_j)}$ and $q_j A_j^{(a_j)}$ are prime to $\prod_{p\in\mathscr{P}} p$. There are only $\binom{K}{2}$ possibilities for the pair (a_j, b_j), and so some choice (a, b) must be taken on for infinitely many j. As $j \to \infty$ through corresponding values,

$$\log \frac{q_j A_j^{(a)}}{p_j A_j^{(b)}} = \log \frac{(q_j+1)A_j^{(a)}}{(p_j+1)A_j^{(b)}} + o(1) = \log \frac{\sigma(A_j^{(b)})}{A_j^{(b)}} - \log \frac{\sigma(A_j^{(a)})}{A_j^{(a)}} + o(1),$$

which gives that $\alpha_b - \alpha_a \in \mathbf{L}^{\mathscr{P}}$.

Proof (Theorem 1.1). We show that $\mathbf{R}_{\geq 0}$ is contained in $\mathbf{L}$. Take any $\alpha \geq 0$.

Let $\epsilon > 0$. Apply Lemma 2.1 with $\mathscr{P} = \emptyset$ and $\alpha_1, \dots, \alpha_k$ chosen as $0, \frac{1}{K}\epsilon, \dots, \frac{K-1}{K}\epsilon$. We find that one of $\frac{1}{K}\epsilon, \frac{2}{K}\epsilon, \dots, \frac{K-1}{K}\epsilon \in \mathbf{L}$; in particular, $\mathbf{L} \cap (\frac{\epsilon}{2K}, \epsilon) \neq \emptyset$. Thus, we can choose m_1 and n_1 with $\sigma(m_1) = \sigma(n_1)$ and $\frac{\epsilon}{2K} < \log \frac{m_1}{n_1} < \epsilon$. Suppose we have already defined m_j and n_j. Apply Lemma 2.1 with the same $\alpha_1, \dots, \alpha_k$ but with $\mathscr{P}$ the set of primes dividing $\prod_{i=1}^{j} m_i n_i$. We find that there are natural numbers m_{j+1} and n_{j+1} with $\sigma(m_{j+1}) = \sigma(n_{j+1})$, $\gcd(m_{j+1}n_{j+1}, \prod_{i=1}^{j} m_i n_i) = 1$, and $\frac{\epsilon}{2K} < \log \frac{m_{j+1}}{n_{j+1}} < \epsilon$. We continue ad infinitum to produce infinite sequences $\{m_j\}$ and $\{n_j\}$.

Since each $\log \frac{m_j}{n_j} \geq \frac{\epsilon}{2K}$, we may choose J with $\sum_{j=1}^{J} \log \frac{m_j}{n_j} \geq \alpha$. Moreover, if J is chosen minimally, then $\alpha \leq \sum_{j=1}^{J} \log \frac{m_j}{n_j} < \alpha + \epsilon$. With $m := \prod_{j=1}^{J} m_j$ and $n := \prod_{j=1}^{J} n_j$, we see that $\sigma(m) = \sigma(n)$ and $0 \leq \log \frac{m}{n} - \alpha < \epsilon$.

Since $\epsilon > 0$ was arbitrary, $\alpha \in \mathbf{L}$.

Remark 2.1. Ford's methods (see the proof of the lower bound in [6]) show that given any fiber $\sigma^{-1}(v) = \{n_1, \dots, n_k\}$, a positive proportion of all fibers $\sigma^{-1}(w)$ have the form $\{dn_1, \dots, dn_k\}$ for some natural number d. Clearly, dilating by a factor of d does not change the ratios between elements of a set. Thus, not only is every $\beta > 0$ well approximable by a ratio m/n, where $\sigma(m) = \sigma(n)$, but it is not so unusual to see a ratio close to β. For example, a positive proportion of $v \in \sigma(\mathbf{N})$ have two preimages m and n with $|\frac{m}{n} - \pi| < 10^{-10}$.

Remark 2.2. It seems interesting to observe that in the statement of Theorem 1.1, m and n can be taken to be coprime. Indeed, the argument given above produces squarefree integers m and n. Since $\sigma(m) = \sigma(n)$, if we write $m/n = m'/n'$ in lowest terms, then also $\sigma(m') = \sigma(n')$.

Remark 2.3. With obvious modifications, our argument will show that Theorem 1.1 also holds with φ replacing σ. In the same source [5] where Erdős mentions the problem for σ, he claims that this φ-variant can be handled in an elementary fashion.

3 Proof of Theorem 1.2

3.1 Overview of the Basic Method

The proof follows the method of Ford and Pollack [7] for showing that most φ-values are not σ-values, and vice versa. We review the strategy of that argument here. For $f \in \{\varphi, \sigma\}$, let $V_f(x)$ be the number of f-values belonging to $[1, x]$. As already alluded to in the introduction, one knows [6, Theorem 14] that $V_\varphi(x) \asymp V_\sigma(x)$ for $x \geq 1$; thus, the main result of Ford and Pollack [7] follows if it is shown that the number of common φ, σ values in $[1, x]$ is $o(V_\varphi(x) + V_\sigma(x))$.

To this end, one begins by constructing [7, Sect. 3] sets $\mathscr{A}_\varphi$ and $\mathscr{A}_\sigma$ with the property that almost all f-values in $[1, x]$ have all their f-preimages in $\mathscr{A}_f$. The precise definition of the sets $\mathscr{A}_f$ [7, p. 1679] is quite intricate and incorporates both "anatomical" and "structural" constraints. By "anatomical," we mean multiplicative constraints of the sort that often arise in elementary number theory. For example, we insist that for $a \in \mathscr{A}_f$, neither a nor $f(a)$ has an extraordinarily large squarefull divisor or "too many" prime divisors. Chief among the anatomical constraints is the requirement that every prime p dividing an element of $\mathscr{A}_f$ be a *normal* prime, meaning that the prime divisors of both $p - 1$ and $p + 1$ are roughly uniformly distributed on a double-logarithmic scale.

By "structural," we mean that extensive use is made of the results of Ford [6] describing the fine structure of typical f-values and their preimages. As an example, precise inequalities are imposed on the prime divisors of elements of $\mathscr{A}_f$; the ordered list of such primes, after a double-logarithmic rescaling, must (up to a small error) correspond to a point in the *fundamental simplex* of Ford [6, Sect. 3]. In addition, we require—and this is the main innovation of Ford and Pollack [7]—that a particular linear combination of renormalized prime factors be slightly less than 1 (this is condition (8) below in the definition of the set $\mathscr{A}_\sigma$). This ensures that sieve bounds (such as those that feature in Lemma 3.2 below) eventually yield a nontrivial estimate.

Having constructed such sets $\mathscr{A}_f$, it is enough to study how many common φ, σ values appear as solutions to an equation of the form

$$\varphi(a) = \sigma(a'), \quad \text{where} \quad a \in \mathscr{A}_\varphi, a' \in \mathscr{A}_\sigma. \tag{1}$$

Write $a = p_0 p_1 p_2 \cdots$ and $a' = q_0 q_1 q_2 \cdots$, where the sequences of primes p_i and q_j are nonincreasing. The normality condition in the definition of the sets $\mathscr{A}_f$ implies that for small values of i, we have $p_i \approx q_i$, at least on a double logarithmic scale.

We classify the primes p_i and q_i dividing a and a' into three categories: "large," "small," and "tiny" (as described in [7, §5A]). Then (1) gives rise to an equation of the form

$$(p_0 - 1)(p_1 - 1)\cdots(p_{k-1} - 1)fd = (q_0 + 1)(q_1 + 1)\cdots(q_{k-1} + 1)e. \tag{2}$$

Here $p_0, \dots, p_{k-1}$ and $q_0, \dots, q_{k-1}$ are the large primes in a and a' (respectively), f is the contribution to $\varphi(a)$ of the small primes, d is the contribution to $\varphi(a)$ of the tiny primes, and e is the contribution to $\sigma(a')$ of both the small and tiny primes.

To finish the argument, we require an estimate for the number of solutions to equations of the form (2). We prove a lemma ([7, Lemma 4.1], cf. Lemma 3.2 below) counting the number of solutions $p_0, \dots, p_{k-1}, q_0, \dots, q_{k-1}, e, f$ to possible equations of the form (2), given d and given intervals encoding the rough location of the primes p_i and q_i. (The phrase "possible equations" means that there are many further technical hypotheses in the lemma, but that these hypotheses are automatically satisfied because of our choice of the sets $\mathscr{A}_f$.) Finally, we sum the estimate of the lemma over all possible values of d and all possible selections of intervals; this allows us to show [7, p. 1695] that

$$\#\{\varphi(a) : (a, a') \in \mathscr{A}_\varphi \times \mathscr{A}_\sigma \text{ and } \varphi(a) = \sigma(a')\} \ll \frac{x}{\log x} \exp\left(-\frac{1}{4}(\log_2 x)^{1/2}\right),$$

which is $o(V_\varphi(x) + V_\sigma(x))$ with much room to spare.

3.2 Definition of the Set $\mathscr{A}_\sigma$

Theorem 1.2 will be proved by modifying the above procedure. We start by giving a careful definition of the set $\mathscr{A}_\sigma$ appearing above. The set $\mathscr{A}_\varphi$ can be defined in an entirely similar manner, but we will not need this.

Put

$$F(z) = \sum_{n=1}^{\infty} a_n z^n, \quad \text{where each} \quad a_n = \int_n^{n+1} \log t\, dt. \tag{3}$$

The series defines F as a continuous, increasing function of z on $(0, 1)$. Moreover, $F(z) \to \infty$ as $z \uparrow 1$. Hence, there is a unique $\varrho \in (0, 1)$ with $F(\varrho) = 1$; numerically, $\varrho \approx 0.5426$. We let $C = \frac{1}{2|\log \varrho|}$, which is ≈ 0.8178. (We met this constant already in the introduction.)

Given a natural number n, write $n = p_0(n)p_1(n)p_2(n)\dots$, where $p_0(n) \geq p_1(n) \geq p_2(n) \geq \dots$ are the primes dividing n (with multiplicity). Put

$$x_i(n; x) = \begin{cases} \log_2 p_i(n)/\log_2 x & \text{if } i < \Omega(n) \text{ and } p_i(n) > 2, \\ 0 & \text{otherwise.} \end{cases}$$

For each real number L and each $\boldsymbol{\xi} = (\xi_0, \ldots, \xi_{L-2})$, we let $\mathscr{S}_L(\boldsymbol{\xi})$ be the set of $(x_1, \ldots, x_L) \in \mathbf{R}^L$ with $0 \le x_L \le x_{L-1} \le \ldots \le x_1 \le 1$, and satisfying the system of inequalities

$$\begin{aligned} a_1x_1 + a_2x_2 + \ldots + a_Lx_L &\le \xi_0 \\ a_1x_2 + \ldots + a_{L-1}x_L &\le \xi_1x_1 \\ &\vdots \\ a_1x_{L-1} + a_2x_L &\le \xi_{L-2}x_{L-2}. \end{aligned}$$

The region corresponding to $\xi = (1, 1, \ldots, 1)$ is called the L-dimensional *fundamental simplex*. Let

$$L_0(x) = \lfloor 2C(\log_3 x - \log_4 x) \rfloor;$$

when x is clear from context, we abbreviate $L_0(x)$ to L_0. One of the key observations of Ford [6] is that if the components of $\boldsymbol{\xi}$ are slightly larger than 1, and L is a little smaller than L_0, then the n for which $(x_1(n;x), \ldots, x_L(n;x)) \in \mathscr{S}_L(\boldsymbol{\xi})$ generate almost all σ-values in $[1, x]$. (See condition (5) below in the definition of $\mathscr{A}_\sigma$ for one way of making this precise.)

Let $S \ge 2$. We say that a prime p is *S-normal* if $\Omega(p+1, 1, S) \le 2\log_2 S$ and, for every pair of real numbers (U, T) with $S \le U < T \le p+1$,

$$|\Omega(p+1, U, T) - (\log_2 T - \log_2 U)| < \sqrt{\log_2 S \log_2 T}.$$

In other words, there are not exorbitantly many prime divisors of $p+1$ up to S, and the larger prime divisors of $p+1$ are roughly uniformly distributed on a double logarithmic scale. (In [7], the definition of S-normal required the same constraints also on $p-1$, but for this paper working with $p+1$ is sufficient.)

We can now give a precise definition of the set $\mathscr{A}_\sigma$. Fix $\epsilon \in (0, 1/2)$ and assume throughout that $x \ge x_0(\epsilon)$. Let

$$S = \exp((\log_2 x)^{36}), \quad \omega = (\log_2 x)^{-\frac{1}{2}+\frac{\epsilon}{2}},$$
$$L = \lfloor L_0 - 2\sqrt{\log_3 x} \rfloor, \quad \text{and} \quad \xi_i = 1 + \frac{1}{10(L_0 - i)^3}.$$

Then $\mathscr{A}_\sigma = \mathscr{A}_\sigma(\epsilon, x)$ is the set of $n = p_0(n)p_1(n)\ldots$ with $\sigma(n) \le x$ that satisfy all of

(0) $n \ge x/\log x$,
(1) every squarefull divisor m of n or of $\sigma(n)$ has $m \le (\log x)^2$,
(2) all of the primes $p_j(n)$ are S-normal,
(3) $\Omega(\sigma(n)) \le 10\log_2 x$ and $\Omega(n) \le 10\log_2 x$,

(4) if $d \parallel n$ and $d \geq \exp((\log x)^{1/2})$, then $\Omega(\sigma(d)) \leq 10 \log_2 \sigma(d)$,
(5) $(x_1(n;x), \ldots, x_L(n;x)) \in \mathscr{S}_L(\boldsymbol{\xi})$,
(6) n has at least $L+1$ odd prime divisors (counted with multiplicity),
(7) $P^+(p_0+1) \geq x^{1/\log_2 x}$, $p_1(n) < x^{1/(100\log_2 x)}$,
(8) $a_1 x_1(n;x) + \ldots + a_L x_L(n;x) \leq 1 - \omega$.

The following statement appears as [7, Lemma 3.2].

Proposition 3.2. *The number of σ-values in $[1,x]$ which have a preimage $n \notin \mathscr{A}_\sigma$ is*

$$\ll V_\sigma(x)(\log_2 x)^{-\frac{1}{2}+\epsilon}.$$

This makes precise our claim in the overview that the elements of $\mathscr{A}_\sigma$ generate almost all σ-values in $[1,x]$.

3.3 The Proof Proper

Say that the σ-value v is *exceptional* if it possesses two preimages a and a' for which $P^+(a) \neq P^+(a')$. We wish to show that the count of exceptional σ-values in $[1,x]$ is $o(V_\sigma(x))$, as $x \to \infty$. In view of Proposition 3.2, we may assume v is such that all of its preimages belong to $\mathscr{A}_\sigma$. Pick preimages a and a' with $P^+(a) \neq P^+(a')$, and write

$$a = p_0 p_1 p_2 \ldots, \quad \text{and} \quad a' = q_0 q_1 q_2 \ldots,$$

where p_i and q_i are nonincreasing sequences.

3.3.1 Rewriting $\sigma(a) = \sigma(a')$

Following the overview presented above, our first task is to deduce from the equation $\sigma(p_0 p_1 p_2 \cdots) = \sigma(q_0 q_1 q_2 \cdots)$ an auxiliary equation of the shape

$$(p_0+1)(p_1+1)\cdots(p_{k-1}+1)fd = (q_0+1)(q_1+1)\cdots(q_{k-1}+1)e. \tag{4}$$

We select k—our cutoff between "large" and "small" primes—in exact parallel with how k is selected in [7, Sect. 5A]. In other words, we choose k_0 as the smallest index for which

$$\log_2 P^+(p_{k_0}+1) \leq (\log_2 x)^{1/2+\epsilon/10},$$

and we take $k = k_0$ unless p_{k_0} and p_{k_0-1} are too close together in a certain technical sense, in which case we take $k = k_0 - 1$.[1] As explained on [7, p. 1689], the properties of $\mathscr{A}_\sigma$ imply that

$$k \sim (1/2 - \epsilon/10)L. \tag{5}$$

(This asymptotic formula is essentially [7, Lemma 5.3]; one has only to change the $p - 1$ to a $p + 1$ in its proof.) Moreover, if $0 \le i < k$, then

$$\log_2 p_i > (\log_2 x)^{1/2+\epsilon/10} \quad \text{and} \quad \log_2 q_i > (\log_2 x)^{1/2+\epsilon/11}.$$

(For this, see again [7, p. 1689].) These last estimates, along with condition (1) in the definition of $\mathscr{A}_\sigma$, guarantee that $p_i^2 \nmid a$ and $q_i^2 \nmid a'$. Thus, the first k factors on the left- and right-hand sides of (4) represent the contribution to $\sigma(a)$ and $\sigma(a')$ from the "large" primes $p_0, \dots, p_{k-1}$ and $q_0, \dots, q_{k-1}$, respectively.

To make the right-hand side of (4) coincide with $\sigma(a')$, it suffices to define

$$e = \sigma(q_k q_{k+1} q_{k+2} \cdots).$$

The choices of f and d are slightly more delicate. If $p_{L-1} \ne p_L$, then we put

$$f = \sigma(p_k p_{k+1} \cdots p_{L-1}), \quad d = \sigma(p_L p_{L+1} \cdots).$$

(In the language of the overview, L is what one thinks of as the cutoff between small and tiny primes.) In the general case, we let A be the largest divisor of a supported on the primes $p_k, \dots, p_{L-1}$, and we put $f = \sigma(A)$ and $d = \sigma(a/(p_1 \cdots p_{k-1} A))$. Then (4) holds. Note that by assumption, $p_0 \ne q_0$.

3.3.2 The Key Sieve Lemma

To continue, we need a tool that allows us to count solutions to (4). We use the following variant of Ford and Pollack [7, Lemma 4.1], which was proved by repeated application of the upper bound sieve. As the required changes to the proof of Ford and Pollack [7, Lemma 4.1] are little more than typographical, we omit the demonstration.

Lemma 3.2. *Let y be large, $k \ge 1$, $l \ge 0$, $30 \le S \le \dots \le v_0 = y$, and $u_j \le v_j$ for $0 \le j \le k - 1$. Suppose that $1 \le B \le y^{1/10}$, and put $\delta = \sqrt{\log_2 S / \log_2 y}$.*

[1] Precisely: With $\eta := 10L\sqrt{\log_2 S / \log_2 x}$, we choose $k = k_0$ unless $x_{k_0-1}(n; x) - x_{k_0}(n; x) < 20\eta$. This becomes relevant for verifying that the intervals $[u_i, v_i]$ selected later in the proof satisfy the conditions of Lemma 3.2 below. Since we will refer to [7] for the selection of u_i and v_i and this verification, it does not make sense here to go into more detail.

Set $v_j = \log_2 v_j / \log_2 y$ *and* $\mu_j = \log_2 u_j / \log_2 y$. *Suppose that* d *is a natural number for which* $P^+(d) \le v_k$. *Moreover, suppose that both of the following hold:*

(a) For $2 \le j \le k-1$, *either* $(\mu_j, v_j) = (\mu_{j-1}, v_{j-1})$ *or* $v_j \le \mu_{j-1} - 2\delta$. *Also,* $v_k \le \mu_{k-1} - 2\delta$.
(b) For $1 \le j \le k-2$, *we have* $v_j > v_{j+2}$.

The number of solutions of

$$(p_0+1)\cdots(p_{k-1}+1)fd = (q_0+1)\cdots(q_{k-1}+1)e \le y/B,$$

in $p_0, \dots, p_{k-1}, q_0, \dots, q_{k-1}, e, f$ *satisfying*

(i) $\gcd(\prod_{i=0}^{k-1} p_i, \prod_{j=0}^{k-1} q_j) = 1$;
(ii) p_i *and* q_i *are S-normal primes;*
(iii) $u_i \le P^+(p_i+1), P^+(q_i+1) \le v_i$ *for* $0 \le i \le k-1$;
(iv) no p_i+1 *or* q_i+1 *is divisible by* r^2 *for a prime* $r \ge v_k$;
(v) $P^+(ef) \le v_k$; $\Omega(f) \le 4l \log_2 v_k$;
(vi) p_0+1 *has a divisor* $\ge y^{1/2}$ *which is composed of primes* $\ge v_1$;

is

$$\ll \frac{y}{dB}(c\log_2 y)^{6k}(k+1)^{\Omega(d)}(\log v_k)^{8(k+l)\log(k+l)+1}(\log y)^{-2+\sum_{i=1}^{k-1} a_i v_i + E},$$

where $E = \delta \sum_{i=2}^{k}(i \log i + i) + 2\sum_{i=1}^{k-1}(v_i - \mu_i)$. *Here* c *is an absolute positive constant, and the* a_i *are as defined in* (3).

Remark 3.4. The condition (i) is not present in [7, Lemma 4.1]. The explanation is that applying the upper bound sieve requires a linear independence condition on the linear forms. This condition is automatic when treating Eq. (2), because the left-hand shifted prime factors are shifted by -1 whereas the right-hand shifts are by $+1$. In (4), the primes are shifted by $+1$ on both sides, forcing us to assume (i).

3.3.3 Capturing Solutions with Lemma 3.2: Attempt #1

Given a pair $(a, a') \in \mathscr{A}_\sigma \times \mathscr{A}_\sigma$ satisfying

$$\sigma(a) = \sigma(a') \quad \text{and} \quad P^+(a) \ne P^+(a'), \tag{6}$$

we described in Sect. 3.3.1 how to construct a solution to (4). By a "solution," we mean the values of $p_0, \dots, p_{k-1}, q_0, \dots, q_{k-1}, d, e$, and f. On the other hand, given a solution to (4) that arose this way, we can recover the common value of $\sigma(a)$ and $\sigma(a')$ by computing either side of (4). So we can bound the number of exceptional σ-values having all preimages in $\mathscr{A}_\sigma$ by the number of solutions to (4) that arise—in the manner detailed in Sect. 3.3.1—from $(a, a') \in \mathscr{A}_\sigma \times \mathscr{A}_\sigma$ satisfying (6). Henceforth, when we speak of a "solution" to (4), we always mean a solution that arose this way.

We group solutions to (4) according to the value of k, the value of d, and the "rough positions" of the primes $P^+(p_i+1)$ and $P^+(q_i+1)$. (In view of the fact that each p_i, q_i is normal, this is essentially the same as grouping by the positions of the p_i and q_i themselves, but turns out to be technically more convenient.) Our hope is to apply Lemma 3.2 to bound the number of solutions in each group, and then sum over all the groups.

Suppose we start with a solution to (4) and want to place it in a group. What does it mean precisely to specify "the rough positions" of the primes $P^+(p_i+1)$ and $P^+(q_i+1)$? We will interpret this to mean that we specify $u_0, \dots, u_{k-1}$ and $v_0, \dots, v_k$ so that, taking $y := x$,

- $30 \le S \le v_k \le v_{k-1} \le \dots \le v_0 = y$
- $u_j \le v_j$ for $0 \le j \le k-1$,
- (a) and (b) in Lemma 3.2 hold,
- (iii) holds.

A systematic way of choosing u_i and v_i to satisfy these criteria is described in detail in [7, Sect. 5B].[2]

Moreover, if one selects the u_i and v_i by that recipe, and takes

$$B = 1 \quad \text{and} \quad l = L - k,$$

then our solution to (4) satisfies not only (iii) but in fact every condition of Lemma 3.2 except possibly condition (i). That is, $P^+(d) \le v_k$ and all of (ii)–(vi) hold. This follows mutatis mutandis from the corresponding proofs in [7, Sect. 5C]. The only point that merits further discussion is the verification that $P^+(ef) \le v_k$ (as claimed in (v)) and the related point that $P^+(d) \le v_k$. Here the more wild behavior of σ on prime powers, vis-à-vis φ, complicates matters.

Let $r := P^+(e)$. Since $k < L$ and a' has at least $L+1$ distinct odd prime divisors, $e = \sigma(q_k q_{k+1} \cdots) > 1$ and so $r > 1$. Choose a prime power R exactly dividing $q_k q_{k+1} q_{k+2} \cdots$ for which $r \mid \sigma(R)$. If R is a proper prime power, then (1) in the definition of $\mathscr{A}_\sigma$ implies that $R \le (\log x)^2$ and so

$$r \le \sigma(R) < 2(\log x)^2 < S \le v_k.$$

So we can assume that R is a prime divisor of $q_k q_{k+1} \cdots$. Then $r \mid R+1$, and $r \le P^+(R+1) \le \max\{3, R\} \le q_k$. From the fifth display on [7, p. 1692],

$$\log_2 q_k / \log_2 x \le \log_2 p_k / \log_2 x + (2k+1)\sqrt{\log_2 S / \log_2 x} \le \log_2 v_k / \log_2 x.$$

Thus, $r \le q_k \le v_k$. An entirely similar argument shows that $P^+(\sigma(p_k p_{k+1} \cdots)) \le v_k$, so that both $P^+(f) \le v_k$ and $P^+(d) \le v_k$.

[2] Since we are working with solutions to (4) instead of (2), one should read [7, Sect. 5B] mentally replacing each expression of the form $p-1$ with $p+1$.

If (i) were to always hold, it would be clear how we ought to finish the proof of Theorem 1.2. In that case, every solution to (4) would fit in a group of the sort counted by Lemma 3.2. Summing the estimate of the lemma over the possible k, d, and u_i, v_i (that is, over all possible groups of solutions) would give us an upper bound on the count of all solutions. However, there is no reason for (i) to always hold. It could well be that the list $p_1, \dots, p_{k-1}$ overlaps with the list $q_1, \dots, q_{k-1}$. So we must work a bit harder before Lemma 3.2 can be applied.

3.3.4 Attempt #2

There is an easy fix for the problem we have just run into: Do not attempt to apply Lemma 3.2 until after canceling factors arising from the common large primes! Given a solution to (4), put

$$m = \gcd(p_0 \cdots p_{k-1}, q_0 \cdots q_{k-1}).$$

By assumption, $p_0 \neq q_0$. It follows that neither p_0 nor q_0 can divide m. Indeed, from conditions (3) and (7) in our definition of $\mathscr{A}_\sigma$, both $p_0, q_0 > x^{1/2}$ while each $p_i, q_i \leq x^{\frac{1}{100 \log_2 x}}$ for $i \geq 1$.

For each prime p dividing m, cancel the factors of $p+1$ from both sides of (4). Relabeling, we obtain an equation of the form

$$(\tilde{p}_0 + 1) \cdots (\tilde{p}_{K-1} + 1) df = (\tilde{q}_0 + 1) \cdots (\tilde{q}_{K-1} + 1) e \tag{7}$$

where $K = k - \omega(m)$ and the common value of both sides of (7) is at most $x/\sigma(m)$. We may assume that the $\tilde{p}_i$ and $\tilde{q}_i$ are in nonincreasing order. Since $\gcd(m, p_0 q_0) = 1$, we have $K \geq 1$, $\tilde{p}_0 = p_0$, and $\tilde{q}_0 = q_0$. Write each

$$\tilde{p}_i = p_{j_i}, \quad \text{and} \quad \tilde{q}_i = q_{j_i'},$$

where the indices i and i' satisfy $i \leq j_i, j_i' < k$ for $0 \leq i < K$.

In the last section, our choices of parameters in Lemma 3.2 possibly failed to capture the solution $p_0, \dots, p_{k-1}, q_0, \dots, q_{k-1}, d, e, f$ to (4). We now describe how to capture the solution $\tilde{p}_0, \dots, \tilde{p}_{K-1}, \tilde{q}_0, \dots, \tilde{q}_{K-1}, d, e, f$ to (7) by making slightly different choices of these parameters.

We continue to assume that u_i and v_i are chosen as in the preceding section. Since hypothesis (a) of Lemma 3.2 holds, for $i \geq 1$ the intervals $[u_i, v_i]$ and $[u_{i+1}, v_{i+1}]$ either coincide or are disjoint. Now appealing to (iii)—which was also satisfied for our choices of u_i, v_i—we see that

$$u_{j_i} = u_{j_i'} \leq P^+(p_{j_i} + 1), P^+(q_{j_i'} + 1) \leq v_{j_i} = v_{j_i'}$$

for every $1 \leq i \leq K-1$. Put

$$\tilde{u}_i = u_{j_i} \text{ and } \tilde{v}_i = v_{j_i} \text{ for } 0 \leq i \leq K-1, \quad \text{and put} \quad \tilde{v}_K = v_k.$$

From the second half of condition (7) in the definition of $\mathscr{A}_\sigma$ and the estimate $k < L = O(\log_3 x)$,

$$\sigma(m) \le m^2 \le (p_1 \cdots p_{k-1})^2 \le x^{O(\log_3 x/\log_2 x)}.$$

Hence,

$$\sigma(m) < x^{1/10}$$

for large x.

We will apply Lemma 3.2 with

$$y = x, \quad K \text{ playing the role of } k \text{ in the lemma}, \quad \tilde{u}_i, \tilde{v}_i \text{ in place of } u_i, v_i,$$
$$B = \sigma(m), \quad d \text{ as before}, \quad l \text{ as before (i.e., } l = L - k \text{ for our original } k).$$

Since all of the previous hypotheses except (possibly) (i) held for our solution to (4), all of the statements in Lemma 3.2 are satisfied for our solution to (7). That is, with this choice of parameters, Lemma 3.2 succeeds in capturing our solution to (7).

Now given m, one can recover the original solution to (4) from the solution to the canceled form (7). So to count solutions to (4), it is enough to sum the upper bound of the lemma over possible values of the parameters d, m (which determines $B = \sigma(m)$), k (which determines $l = L - k$), K, the $\tilde{u}_i$, and $\tilde{v}_i$.

3.3.5 Finishing Up

Making analogous calculations to those on [7, pp. 1693–1694], the upper bound arising from a single application of Lemma 3.2 is seen to be

$$\ll \frac{x}{\sigma(m)\log x} \exp(-\frac{1}{3}(\log_2 x)^{1/2+\epsilon/2}) \frac{L^{\Omega(d)}}{d}.$$

(compare with [7, Eq. (5–13)]). It remains to sum on d, k, K, m, $\tilde{u}_i$, and $\tilde{v}_i$.

Since each of k and K is bounded by L, there are only $O((\log_3 x)^2)$ possibilities for the pair (k, K). Reasoning as in [7, Eq. (5–14)], the number of possibilities for the $\tilde{u}_i$ and $\tilde{v}_i$ is bounded by $\exp(O((\log_3 x)^2))$.

To handle the sum on d, we first establish the uniform bound

$$\Omega(d) \ll (\log_2 x)^{1/2}.$$

Recall that we defined d so that $d = \sigma(h)$ for some unitary divisor h of a supported on primes $\le p_L$. If $h \le \exp((\log_2 x)^{1/2})$, then the desired bound on $\Omega(d)$ follows from the estimate $\Omega(d) \ll \log d$. Otherwise, conditions (3) and (4) in the definition of $\mathscr{A}_\sigma$ give us

$$\Omega(d) \le 10 \log_2 \sigma(h) \ll \log_2 h \le \log_2 p_L^{10 \log_2 x},$$

and this is also $O((\log_2 x)^{1/2})$, by the calculation in the final display of Ford and Pollack [7, p. 1694]. It follows that

$$L^{\Omega(d)} \le \exp(O((\log_2 x)^{1/2} \log_4 x)).$$

Recall that $P^+(d) \le v_k$. Our choice of v_k now yields $\log_2 P^+(d) \le (\log_2 x)^{1/2+\epsilon/5}$ [7, Eq. (5–12)]. Hence,

$$\sum \frac{1}{d} \ll \exp((\log_2 x)^{1/2+\epsilon/5}).$$

Assembling the preceding estimates, we find that the number of solutions to (4) corresponding to a given value of $m = \gcd(p_0 \dots p_{k-1}, q_0 \cdots q_{k-1})$ is at most

$$\frac{x}{\sigma(m) \log x} \exp\left(-(\log_2 x)^{1/2}\right).$$

It remains finally to treat the sum over m. Since m divides $p_1 \cdots p_{k-1}$, we have that m is squarefree and (recalling (5))

$$\omega(m) < k < \frac{1}{2}L < 0.9 \log_3 x.$$

Hence,

$$\sum_m \frac{1}{\sigma(m)} \le \sum_{j \le 0.9 \log_3 x} \frac{1}{j!} \left(\sum_{p \le x} \frac{1}{p+1} \right)^j < \exp((\log_3 x)^2),$$

by a short calculation using Mertens's estimate $\sum_{p \le x} p^{-1} = \log_2 x + O(1)$. We conclude that the total number of solutions to (4) is bounded by (say)

$$\frac{x}{\log x} \exp\left(-\frac{1}{2}(\log_2 x)^{1/2}\right).$$

As discussed before, this is also an upper bound on the number of exceptional σ-values in $[1, x]$ all of whose preimages belong to $\mathscr{A}_\sigma$. Since this upper bound is certainly $o(V_\sigma(x))$, the proof of Theorem 1.2 is complete.

Remark 3.5. Recall that $\epsilon > 0$ can be taken arbitrarily small in the definition of $\mathscr{A}_\sigma$. From the above argument and Proposition 3.2, it follows that the number of exceptional σ-values in $[1, x]$ is at most $V_\sigma(x)/(\log_2 x)^{1/2+o(1)}$, as $x \to \infty$.

Remark 3.6. Using the preceding remark and further ideas from [6, 7], one can establish the following strengthening of Theorem 1.2: *For each fixed K, almost all σ-values in $[1, x]$ are such that all of their preimages share the same largest $K+1$ prime factors*. One could even extend this to certain functions $K = K(x) \to \infty$, but we do not pursue this possibility here.

Remark 3.7. Theorem 1.2 as well as the comments in the preceding remarks all hold with σ replaced by Euler's φ-function, by essentially the same proofs.

4 Proof of Corollary 1.1

Proof. Suppose that $v \le x$ is the common σ-value of some amicable tuple. Then there are $n_1, \dots, n_k$ with $\sum_{i=1}^{k} n_i = v$ and each $\sigma(n_i) = v$. Clearly, each $n_i \le x$. By Theorem 1.2, we can assume that all the n_i have the same largest prime factor P. We can also assume that $P > z$, where $z := x^{1/(4 \log\log x)}$. Otherwise, a crude upper bound on the count of smooth numbers (such as [17, Theorem 1, p. 359]) shows that each n_i is restricted to a set of size $\ll x/(\log x)^2 = o(V_\sigma(x))$, which would mean that $v = \sigma(n_1)$ is also so restricted. For the remaining values of v, observe that P divides $\sum_{i=1}^{k} n_i = v = \sigma(n_1)$. Write $n_1 = Pm$. If $P \mid m$, then n_1 is divisible by the square of a prime exceeding z, leaving $\ll x \sum_{p>z} p^{-2} \ll x/z = o(V_\sigma(x))$ possibilities for n_1. So assume that $P \nmid m$. Since P divides $\sigma(n_1) = (P+1)\sigma(m)$ and $P = P^+(n_1)$, there must be a proper prime power R dividing m for which $P \mid \sigma(R)$. Since P divides $\sigma(R)$ and $\sigma(R) < 2R$, we have that $R > P/2 \ge z/2$. Thus, n_1 possesses a squarefull divisor exceeding $z/2$, which restricts n_1—and also $v = \sigma(n_1)$—to a set of size $\ll x/\sqrt{z} = o(V_\sigma(x))$.

Acknowledgements The author thanks Kevin Ford for guiding him through the arguments of Ford [6]. He also thanks the referee for useful feedback that led to improvements in the exposition. Much of this work was conducted during the Fall 2009 semester, while the author was visiting the Institute for Advanced Study. He thanks the IAS for providing an ideal working environment. While at the IAS, the author was supported by National Science Foundation award DMS-0802970; currently, he is supported by NSF DMS-1402268.

References

1. L.E. Dickson, Amicable number triples. Am. Math. Mon. **20**, 84–92 (1913)
2. L.E. Dickson, *History of the Theory of Numbers. Volume I: Divisibility and Primality* (Chelsea Publishing Co., New York, 1966)
3. P. Erdős, G.J. Rieger, Ein Nachtrag über befreundete Zahlen. J. Reine Angew. Math. **273**, 220 (1975)
4. P. Erdős, On amicable numbers. Publ. Math. Debrecen **4**, 108–111 (1955)
5. P. Erdős, Remarks on number theory. II. Some problems on the σ function. Acta Arith. **5**, 171–177 (1959)
6. K. Ford, The distribution of totients. Ramanujan J. **2**, 67–151 (1998). Revised version available as arXiv:1104.3264 [math.NT]
7. K. Ford, P. Pollack, On common values of $\varphi(n)$ and $\sigma(m)$, II. Algebra Number Theory **6**, 1669–1696 (2012)
8. H. Maier, C. Pomerance, On the number of distinct values of Euler's φ-function. Acta Arith. **49**, 263–275 (1988)

9. A. Mąkowski, On some equations involving functions $\varphi(n)$ and $\sigma(n)$. Am. Math. Mon. **67**, 668–670 (1960)
10. T.E. Mason, On amicable numbers and their generalizations. Am. Math. Mon. **28**, 195–200 (1921)
11. J. Maynard, Dense clusters of primes in subsets. Preprint online as arXiv:1405.2593 [math.NT] (2014)
12. C. Pomerance, On the distribution of amicable numbers. J. Reine Angew. Math. **293/294**, 217–222 (1977)
13. C. Pomerance, On the distribution of amicable numbers. II. J. Reine Angew. Math. **325**, 183–188 (1981)
14. C. Pomerance, On amicable numbers (2014). To appear in a Springer volume in honor of H. Maier
15. P. Poulet, *La chasse aux nombres. I: Parfaits, amiables et extensions* (Stevens, Bruxelles, 1929)
16. A. Schinzel, W. Sierpiński, Sur certaines hypothèses concernant les nombres premiers. Acta Arith. **4**, 185–208 (1958). Erratum **5**, 259 (1959)
17. G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*. Cambridge Studies in Advanced Mathematics, vol. 46 (Cambridge University Press, Cambridge, 1995)
18. Y. Zhang, Bounded gaps between primes. Ann. Math. **179**, 1121–1174 (2014)

On Amicable Numbers

Carl Pomerance

To Professor Helmut Maier on his 60th birthday

Abstract Let $\mathscr{A}(x)$ denote the set of integers $n \le x$ that belong to an amicable pair. We show that $\#\mathscr{A}(x) \le x/e^{\sqrt{\log x}}$ for all sufficiently large x.

Keywords Amicable number

Mathematics Subject Classification: 11A25, 11N25

1 Introduction

Let σ denote the sum-of-divisors function and let $s(n) = \sigma(n) - n$. Two different positive integers a, b with $s(a) = b$ and $s(b) = a$ are said to form an amicable pair. This concept is attributed to Pythagoras and has been studied over the millennia since both by numerologists and number theorists. The first example of an amicable pair is 220 and 284. About 12 million pairs are now known, but we don't have a proof of their infinitude.

Say a positive integer is amicable if it belongs to an amicable pair and let $\mathscr{A}$ denote the set of amicable numbers. Kanold [9] was the first to consider $\mathscr{A}$ from a statistical viewpoint, announcing in 1954 that $\mathscr{A}$ has upper density smaller than 0.204. Soon after, Erdős [5] showed they have asymptotic density 0. In the period 1973–1981 there were several papers getting successively better upper bounds for $\#\mathscr{A}(x)$, where $\mathscr{A}(x) = \mathscr{A} \cap [1, x]$. Somewhat simplifying the expressions, these upper bounds have progressed as follows:

$$\frac{x}{(\log\log\log\log x)^{1/2}}, \quad \frac{x}{\log\log\log x}, \quad \frac{x}{\exp\left((\log\log\log x)^{1/2}\right)}, \quad \frac{x}{\exp\left((\log x)^{1/3}\right)},$$

C. Pomerance (✉)
Dartmouth College, Hanover, NH 03755, USA
e-mail: carl.pomerance@dartmouth.edu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_19

see [7, 14–16], respectively. In this paper we are able to replace the exponent $1/3$ in the last estimate with $1/2$. In particular, we prove the following theorem.

Theorem 1.1. *As $x \to \infty$, we have*

$$\#\mathscr{A}(x) \le x/\exp\left(\left(\tfrac{1}{2} + o(1)\right)\sqrt{\log x \log\log\log x}\right).$$

The proof largely follows the plan in [15], but with some new elements. In particular, a separate argument now handles the case when an amicable number n is divisible by a very large prime p. Thus, assuming the largest prime factor p of n is not so large, it can be shown that, usually, a fairly large prime divides $\sigma(n/p)$. The argument to handle the case of p large is reminiscent of the paper [13] which deals with Lehmer's problem on Euler's function φ, also see the newer paper [1]. In addition, we use a result in [2] to streamline the argument that $\sigma(n/p)$ has a large prime factor.

The argument in [15] was subsequently used to estimate the distribution of numbers n with $\varphi(n) = \varphi(n+1)$, and some similar equations; see [6, 8]. These results were since improved in [17]. However, it is not clear if the method of Yamada [17] can be used for the distribution of amicable numbers.

We record some of the notation used. Let $P(n)$ denote the largest prime factor of $n > 1$ and let $P(1) = 1$. We say an integer n is squarefull if for each prime $p \mid n$ we have $p^2 \mid n$. We use the notation (a, b) for the greatest common divisor of the positive integers a, b. We write $d \parallel n$ if $d \mid n$ and $(d, n/d) = 1$.

Note that if n is large we have $s(n) < 2n \log\log n$.

2 A Lemma

Let $\Phi(x, y)$ denote the number of integers $n \in [1, x]$ with $P(\varphi(n)) \le y$. In [2, Theorem 3.1] it is shown that for any fixed $\epsilon > 0$, we have $\Phi(x, y) \le x\exp(-(1 + o(1))u \log\log u)$ as $u \to \infty$, where $u = \log x/\log y$ and $(\log\log x)^{1+\epsilon} \le y \le x$. At first glance one might think there is a typographical error here, since the corresponding result of de Bruijn [4] counts $n \in [1, x]$ with $P(n) \le y$, and the upper bound is $x\exp(-(1 + o(1))u \log u)$, which is supported by a corresponding lower bound, see [3]. However, a heuristic argument indicates that it is likely that for $\Phi(x, y)$, replacing $\log u$ with $\log\log u$ is correct (see [2, Sect. 8] and [10]). Until we have a corresponding lower bound, we will not know for sure.

In this paper we will require an estimate for the number of $n \in [1, x]$ with $P(\sigma(n)) \le y$. The function $\sigma(n)$ closely resembles the function $\varphi(n)$; we see "$p+1$" in $\sigma(n)$, where we would see "$p - 1$" in $\varphi(n)$. However, it is not as simple as this, since σ treats higher powers of primes differently than φ. It is possible to overcome this difference since most numbers n are not divisible by a large squarefull number. But to keep things simple, we restrict our numbers n in the following result to squarefree numbers. Let $\Sigma(x, y)$ denote the number of squarefree numbers $n \in [1, x]$ with $P(\sigma(n)) \le y$.

Lemma 2.1. *For each fixed $\epsilon > 0$, we have $\Sigma(x,y) \le x\exp(-(1+o(1))u\log\log u)$ as $u \to \infty$, where $u = \log x/\log y$ and $(\log\log x)^{1+\epsilon} \le y \le x$.*

As indicated above, the proof of this result follows from small cosmetic changes to the proof of the corresponding result on $\Phi(x,y)$ in [2].

3 Proof of Theorem 1.1

Let x be large and let

$$L = L(x) = \exp\left(\tfrac{1}{2}\sqrt{\log x \log\log\log x}\right).$$

(i) *For $n \in \mathscr{A}(x)$, we may assume that $n > x/L$ and $s(n) > x/L$.*

The first assertion is obvious and the second follows from the fact that each $n \in \mathscr{A}(x)$ is determined by $s(n)$.

(ii) *For $n \in \mathscr{A}(x)$ we may assume that the largest divisor d of n with $P(d) \le L^2$ has $d \le x^{1/3}$, and similarly for $s(n)$. In particular, we may assume that $P(n) > L^2$ and $P(s(n)) > L^2$.*

Indeed, by [4], for $z \ge x^{1/3}$, the number of integers $d \le z$ with $P(d) \le L^2$ is $O(z/L)$, so by partial summation, the number of $n \in \mathscr{A}(x)$ divisible by such a number $d > x^{1/3}$ is $O(x(\log x)/L)$. A parallel argument holds for $s(n)$. The assertions about $P(n), P(s(n))$ now follow from (i).

(iii) *For $n \in \mathscr{A}(x)$ we may assume the largest squarefull divisor of n is at most L^2, and the same for $s(n)$.*

Since the number of squarefull numbers $d \le z$ is $O(\sqrt{z})$ for all $z \ge 1$, partial summation implies that the number of integers $n \le x$ divisible by a squarefull number $d > L^2$ is $O(x/L)$. A similar estimate holds for $s(n)$.

Note that if $n \in \mathscr{A}(x)$, then (ii) and (iii) imply that $P(n) \parallel n$ and $P(s(n)) \parallel s(n)$. For the remainder of the proof, for $n \in \mathscr{A}(x)$ we write:

$$n = pm,\ p = P(n) \nmid m, \quad s(n) = n' = p'm',\ p' = P(n') \nmid m'.$$

Note too that for x large we have $n' < 2x\log\log x$.

(iv) *For $n \in \mathscr{A}(x)$, we may assume that $P((n, s(n))) \le L$.*

Suppose $n \in \mathscr{A}(x)$, $r = P((n,s(n)))$, and that $r > L$. Then $r \mid \sigma(n)$, so there is a prime power $q^j \mid n$ with $r \mid \sigma(q^j)$. We use an elementary inequality found in the proofs of [11, Lemma 3.6], [12, Lemma 3.3]: for any positive integer d,

$$\sum_{\substack{q^j \le x \\ d \mid \sigma(q^j)}} \frac{1}{q^j} \ll \frac{(\log x)^2}{d}. \tag{1}$$

We Apply (1) with $d = r$, getting that the number of $n \le x$ with $r \mid n$ and $r \mid \sigma(n)$ is $O(x(\log x)^2/r^2)$, and so the number of n which violate (iv) is at most a constant times

$$x(\log x)^2 \sum_{r>L} \frac{1}{r^2} \ll \frac{x(\log x)^2}{L}.$$

This estimate shows that we may assume (iv).

(v) *For $n \in \mathscr{A}(x)$, we may assume that $mm' > x/L$.*

Suppose $n \in \mathscr{A}(x)$. We have

$$p'm' = s(n) = \sigma(pm) - pm = ps(m) + \sigma(m), \tag{2}$$

$$p'\sigma(m') + \sigma(m') = \sigma(p'm') = \sigma(n) = p\sigma(m) + \sigma(m). \tag{3}$$

Multiplying (2) by $\sigma(m)$, (3) by $s(m)$ and subtracting to eliminate p, we have

$$p'm'\sigma(m) - p'\sigma(m')s(m) - \sigma(m')s(m) = \sigma(m)^2 - \sigma(m)s(m) = m\sigma(m).$$

Thus,

$$p'(m'\sigma(m) - \sigma(m')s(m)) = \sigma(m')s(m) + m\sigma(m). \tag{4}$$

Since the right side of (4) is positive, we see that m, m' determine p', and by symmetry, they also determine p. So the number of cases for which (v) fails is at most

$$\sum_{mm' \le x/L} 1 = O(x(\log x)/L).$$

(vi) *For $n \in \mathscr{A}(x)$, we may assume that $p, p' \le x^{3/4}L$.*

Suppose $n \in \mathscr{A}(x)$ and $p > x^{3/4}L$. Then $m < x^{1/4}/L$, so by (v), $m' > x^{3/4}$. Since $n' < 2x \log\log x$, we have $p' < 2x^{1/4} \log\log x$. Write the prime factorization of n' as $p_1 p_2 \dots p_t$, where $p_1 = p'$ and $p_1 \ge p_2 \ge \dots \ge p_t$, and for $j = 1, 2, \dots, t$, let $D_j = p_1 p_2 \dots p_j$. By (i), some $D_j > x^{1/2}$; let D be the least such divisor of n'. Since $p_1 < 2x^{1/4} \log\log x$, we have

$$x^{1/2} < D \le 2x^{3/4} \log\log x. \tag{5}$$

Further, by (ii) and (iii), D is squarefree, $D \parallel n'$, and every prime dividing D is larger than L^2. From the identity

$$\begin{aligned} s(m)\sigma(n') &= s(m)\sigma(n) = \sigma(m)\sigma(n) - \sigma(n)m = \sigma(m)\sigma(n) - \sigma(m)(p+1)m \\ &= \sigma(m)\sigma(n) - \sigma(m)(n+m) = \sigma(m)n' - \sigma(m)m, \end{aligned}$$

we have

$$s(m)\sigma(D)\sigma(M) = \sigma(m)DM - m\sigma(m).$$

Reading this equation as a congruence modulo $\sigma(D)$, we have

$$\sigma(m)DM \equiv m\sigma(m) \pmod{\sigma(D)}.$$

The number of choices for $M < 2x(\log\log x)/D$ which satisfy this congruence is at most

$$1 + \frac{2x\log\log x}{D\sigma(D)/(\sigma(m)D, \sigma(D))} \le 1 + \frac{2x\sigma(m)(D, \sigma(D))\log\log x}{D^2}.$$

Since $(D, \sigma(D)) \mid (n, n')$ and every prime dividing D exceeds $L^2 > L$, (iv) implies that $(D, \sigma(D)) = 1$. So, for a given choice of m, D, the number of choices for M is at most

$$1 + \frac{4xm(\log\log x)^2}{D^2}.$$

We sum this expression for D satisfying (5) and $m < x^{1/4}/L$ and so get that the number of choices for $n \in \mathscr{A}(x)$ with $p > x^{3/4}L$ is $O(x(\log\log x)/L)$. A similar argument holds if $p' > x^{3/4}L$. We conclude that the number of cases where (vi) fails is negligible.

For $n = pm \in \mathscr{A}(x)$ we write $m = m_0m_1$ where m_1 is the largest squarefree number with $m_1 \parallel m$, and we similarly write $m' = m_0'm_1'$.

(vii) *For $n \in \mathscr{A}(x)$, we may assume that $P(\sigma(m_1)) > L$ and $P(\sigma(m_1')) > L$.*

Assume for $n \in \mathscr{A}(x)$ that the first condition in (vii) fails. (The argument for the second condition will follow similarly.) By (iii) and (vi), $pm_0 \le x^{3/4}L^3$. For given choices of p and m_0 we count the number choices of squarefree integers $m_1 \le x/pm_0$ with $P(\sigma(m_1)) \le L$. For this, we use Lemma 2.1. Let $u = \log(x/pm_0)/\log L$. Since $pm_0 \le x^{3/4}L^3$, we have

$$u \ge \frac{\log\left(x^{1/4}/L^3\right)}{\log L} = \left(\tfrac{1}{2} + o(1)\right)\sqrt{\frac{\log x}{\log\log\log x}},$$

so that

$$u \log\log u = \left(\tfrac{1}{2} + o(1)\right) \sqrt{\log x \log\log\log x} = (1 + o(1)) \log L$$

as $x \to \infty$. Hence, we uniformly have that the number of choices for m_1 is at most

$$\frac{x}{pm_0 L^{1+o(1)}}$$

as $x \to \infty$. We now sum on p, m_0 getting that the number of $n \in \mathscr{A}(x)$ for which $P(\sigma(m_1)) \le L$ is at most $x/L^{1+o(1)}$ as $x \to \infty$.

We now turn to the conclusion of the argument. We suppose that $n \in \mathscr{A}(x)$ and that (i)–(vii) hold. At the cost of doubling our count and letting n run up to $2x \log\log x$, we may assume that $p > p'$. (That $p \ne p'$ can be seen from (ii), (iv).) By (vii), there is a prime $r \mid \sigma(m_1)$ with $r > L$. Thus, there is a prime $q \mid m$ with $q \equiv -1 \pmod r$. This implies that $q > L$, so that by (iv), $q \nmid n'$. But $\sigma(n) = \sigma(n')$, so there is a prime power $\ell^j \mid n'$ with $\ell \ne q$ and $r \mid \sigma(\ell^j)$. Note that, by (1), $\sum 1/\ell^j \ll (\log x)^2/r$. We have

$$n' = s(n) = ps(m) + \sigma(m) \equiv 0 \pmod{\ell^j}. \tag{6}$$

Say $\ell^i = (\ell^j, s(m))$, so that p is in a residue class $a(m) \pmod{\ell^{j-i}}$. Also, (6) implies that $\ell^i \mid \sigma(m)$, so $\ell^i \mid m$. Further, using (ii), (iii), and $p > \ell$, we may assume that $p > \ell^j$. The number of such numbers $n \le 2x \log\log x$ is at most

$$\begin{aligned}
&\sum_{r>L} \sum_{\substack{q<x \\ q \equiv -1 \pmod r}} \sum_{\substack{\ell^j < x \\ r \mid \sigma(\ell^j)}} \sum_{i \le j} \sum_{\substack{m<x \\ q\ell^i \mid m}} \sum_{\substack{p \le 2x\log\log x/m \\ p \equiv a(m) \pmod{\ell^{j-i}} \\ p > \ell^j}} 1 \\
&\quad \le \sum_r \sum_q \sum_{\ell^j} \sum_i \sum_m \frac{2x \log\log x}{\ell^{j-i} m} \ll \sum_r \sum_q \sum_{\ell^j} \sum_i \frac{x \log x \log\log x}{q\ell^j} \\
&\quad \ll \sum_r \sum_q \sum_{\ell^j} \frac{x(\log x)^2 \log\log x}{q\ell^j} \ll \sum_r \sum_q \frac{x(\log x)^4 \log\log x}{rq} \\
&\quad \ll \sum_r \frac{x(\log x)^5 \log\log x}{r^2} \ll \frac{x(\log x)^5 \log\log x}{L},
\end{aligned}$$

where we treated r, q, ℓ, p as integer variables. This calculation along with the previous cases finishes the proof.

Acknowledgements The author would like to thank Hanh Nguyen, Paul Pollack, and Lola Thompson for their interest in this work.

References

1. A. Anavi, P. Pollack, C. Pomerance, On congruences of the form $\sigma(n) \equiv a \pmod{n}$. IJNT **9**, 115–124 (2012)
2. W.D. Banks, J.B. Friedlander, C. Pomerance, I.E. Shparlinski, Multiplicative structure of values of the Euler function, in *High Primes and Misdemeanours: Lectures in Honour of the 60th Birthday of Hugh Cowie Williams*, ed. by A. van der Poorten (American Mathematical Society, Providence, 2004). Fields Inst. Commun. **41**, 29–47 (2004)
3. E.R. Canfield, P. Erdős, C. Pomerance, On a problem of Oppenheim concerning 'Factorisatio Numerorum'. J. Number Theory **17**, 1–28 (1983)
4. N.G. de Bruijn, On the number of integers $\leq x$ and free of prime factors $> y$. Nederl. Akad. Wetensch. Proc. Ser. A **54**, 50–60 (1951)
5. P. Erdős, On amicable numbers. Publ. Math. Debr. **4**, 108–111 (1955)
6. P. Erdős, C. Pomerance, A. Sárközy, On locally repeated values of certain arithmetic functions, II. Acta Math. Hungar. **49**, 251–259 (1987)
7. P. Erdős, G.J. Rieger, Ein Nachtrag über befreundete Zahlen. J. Reine Angew. Math. **273**, 220 (1975)
8. S.W. Graham, J.J. Holt, C. Pomerance, On the solutions to $\varphi(n) = \varphi(n+k)$, in *Number Theory in Progress*, ed. by K. Gyory, H. Iwaniec, J. Urbanowicz, vol. 2 (de Gruyter, Berlin/New York, 1999), pp. 867–882
9. H.J. Kanold, Über die Dichte der Mengen der vollkommenen und der befreundeten Zahlen. Math. Z. **61**, 180–185 (1954)
10. Y. Lamzouri, Smooth values of the iterates of the Euler phi-function. Can. J. Math. **59**, 127–147 (2007)
11. P. Pollack, L. Thompson, Practical pretenders. Publ. Math. Debr. **82**, 651–667 (2013)
12. P. Pollack, J. Vandehey, Some normal numbers given by arithmetic functions Canad. Math. Bull. **58**, 160–173 (2015)
13. C. Pomerance, On composite n for which $\varphi(n) \mid n-1$, I. Acta Arith. **28**, 387–389 (1976)
14. C. Pomerance, On the distribution of amicable numbers. J. Reine Angew. Math. **293/294**, 217–222 (1977)
15. C. Pomerance, On the distribution of amicable numbers, II. J. Reine Angew. Math. **325**, 183–188 (1981)
16. G.J. Rieger, Bemerkung zu einem Ergebnis von Erdős über befreundete Zahlen. J. Reine Angew. Math. **261**, 157–163 (1973)
17. T. Yamada, On equations $\sigma(n) = \sigma(n+k)$ and $\varphi(n) = \varphi(n+k)$. arXiv:1001.2511

Trigonometric Representations of Generalized Dedekind and Hardy Sums via the Discrete Fourier Transform

Michael Th. Rassias and László Tóth

To Professor Helmut Maier on his 60th birthday

Abstract We introduce some new higher dimensional generalizations of the Dedekind sums associated with the Bernoulli functions and of those Hardy sums which are defined by the sawtooth function. We generalize a variant of Parseval's formula for the discrete Fourier transform to derive finite trigonometric representations for these sums in a simple unified manner. We also consider a related sum involving the Hurwitz zeta function.

Keywords Dedekind sums • Hardy sums • Bernoulli polynomials and functions • Hurwitz zeta function • Discrete Fourier transform

2010 Mathematics Subject Classification: 11F20, 11L03

1 Introduction

The classical Dedekind sum is defined for $h \in \mathbb{Z}$, $k \in \mathbb{N} := \{1, 2, \ldots\}$ by

$$s(h,k) := \sum_{a \,(\mathrm{mod}\, k)} \left(\left(\frac{a}{k}\right)\right)\left(\left(\frac{ah}{k}\right)\right),$$

adopting the usual notation

M.Th. Rassias (✉)
Department of Mathematics, ETH-Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544-1000, USA
e-mail: michail.rassias@math.ethz.ch; michailrassias@math.princeton.edu

L. Tóth
Department of Mathematics, University of Pécs, Ifjúság útja 6, 7624 Pécs, Hungary
e-mail: ltoth@gamma.ttk.pte.hu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_20

$$((x)) := \begin{cases} \{x\} - \frac{1}{2}, & \text{if } x \notin \mathbb{Z}, \\ 0, & \text{if } x \in \mathbb{Z}, \end{cases}$$

where $\{x\} := x - \lfloor x \rfloor$ stands for the fractional part of x (cf. [9, 13]). If $\gcd(h,k) = 1$, then $s(h,k)$ can be represented as

$$s(h,k) = \frac{1}{4k} \sum_{a=1}^{k-1} \cot\left(\frac{\pi a}{k}\right) \cot\left(\frac{\pi a h}{k}\right) \tag{1}$$

and

$$s(h,k) = \frac{1}{2\pi} \sum_{\substack{r=1 \\ r \not\equiv 0 \pmod{k}}}^{\infty} \frac{1}{r} \cot\left(\frac{\pi r h}{k}\right). \tag{2}$$

The identities (1) and (2) were derived in 1933 by Rademacher [12] in order to obtain a simple direct proof of the reciprocity formula for the Dedekind sums. See also [13, pp. 18–25]. According to [1, p. 347], (1) was obtained earlier, in 1923 by Mellin. The identity (1) is also the starting point for various generalizations of $s(h,k)$. See, e.g., the papers of Beck [2], Dieter [8], Zagier [16].

It is known that (1) is a direct consequence of a variant of Parseval's formula for the discrete Fourier transform (DFT). See the paper of Almkvist [1, Sect. 6] and the book by Beck and Robins [4, Chap. 7]. More specifically, consider a function $f : \mathbb{Z} \to \mathbb{C}$, which is k-periodic (periodic with period k), where $k \in \mathbb{N}$. We define the DFT of f as the function $\hat{f} = \mathscr{F}(f)$, given by

$$\hat{f}(n) := \sum_{a \pmod{k}} f(a) e^{-2\pi i a n/k} \quad (n \in \mathbb{Z}).$$

Furthermore, if f_1 and f_2 are k-periodic functions, then their inner product is

$$\langle f_1, f_2 \rangle := \sum_{a \pmod{k}} f_1(a) \overline{f_2(a)},$$

having the property

$$\langle f_1, f_2 \rangle = \frac{1}{k} \langle \widehat{f_1}, \widehat{f_2} \rangle,$$

or equivalently,

$$\sum_{a \pmod{k}} f_1(a) f_2(-a) = \frac{1}{k} \sum_{a \pmod{k}} \widehat{f_1}(a) \widehat{f_2}(a). \tag{3}$$

Now, (1) follows by applying (3) to the functions

$$f_1(a) = \left(\left(\frac{a}{k}\right)\right), \ f_2(a) = \left(\left(\frac{ah}{k}\right)\right)$$

and using the fact that the DFT of the sawtooth function is essentially the cotangent function.

It is the aim of this paper to exploit this idea in order to deduce similar finite trigonometric representations for certain new generalized Dedekind and Hardy sums, in a simple unified manner. Our results are direct applications of a higher dimensional version of the identity (3), included in Theorem 2.1. We derive in this way Zagier-type identities for new higher dimensional generalizations of the Dedekind sums associated with the Bernoulli functions and of those Hardy sums which are defined by the sawtooth function. Note that all finite trigonometric representations we obtain contain only the tangent and cotangent functions, and are special cases of the Dedekind cotangent sums investigated by Beck [2]. Therefore the reciprocity law proved in [2, Theorem 2] can be applied for each sum.

Furthermore, we consider a related sum, studied by Mikolás [11], involving the Hurwitz zeta function. We remark that (3) was used to evaluate some finite trigonometric and character sums, but not Dedekind and related sums, by Beck and Halloran [3]. We point out that the identity (2) can be obtained from (1) using another general result (Lemma 4.2 in Sect. 4).

2 Properties of the DFT

We will apply the following general result.

Theorem 2.1. *Let $f_1, \ldots, f_m$ be arbitrary k-periodic functions and let $h_j \in \mathbb{Z}$, $\gcd(h_j, k) = 1$ $(1 \le j \le m)$, where $m, k \in \mathbb{N}$. Then*

$$\sum_{\substack{a_1,\ldots,a_m \ (mod\ k) \\ a_1+\ldots+a_m \equiv 0 \ (mod\ k)}} f_1(a_1h_1)\cdots f_m(a_mh_m) = \frac{1}{k} \sum_{a \ (mod\ k)} \widehat{f_1}(ah_1')\cdots\widehat{f_m}(ah_m'),$$

where h_j' is the multiplicative inverse of h_j (mod k), that is $h_jh_j' \equiv 1$ (mod k) with $1 \le j \le m$.

Proof. We only need some simple well-known facts concerning the DFT. See, for instance, [15, Chap. 2], [4, Chap. 7]. The Cauchy convolution of the k-periodic functions f_1 and f_2 is defined by

$$(f_1 \otimes f_2)(n) := \sum_{\substack{a_1,a_2 \ (\mathrm{mod}\ k) \\ a_1+a_2 \equiv n \ (\mathrm{mod}\ k)}} f_1(a_1)f_2(a_2) = \sum_{a \ (\mathrm{mod}\ k)} f_1(a)f_2(n-a) \quad (n \in \mathbb{Z}),$$

which is associative and commutative. Also,

$$\widehat{f_1 \otimes f_2} = \widehat{f_1}\widehat{f_2}.$$

More generally, if $f_1, \ldots, f_m$ are k-periodic functions, then

$$\mathscr{F}(f_1 \otimes \cdots \otimes f_m) = \mathscr{F}(f_1) \cdots \mathscr{F}(f_m).$$

Recalling that

$$\mathscr{F}(\mathscr{F}(f))(n) = kf(-n) \quad (n \in \mathbb{Z}),$$

which is valid for every k-periodic f, this yields

$$k(f_1 \otimes \cdots \otimes f_m)(-n) = \mathscr{F}(\mathscr{F}(f_1) \cdots \mathscr{F}(f_m))(n),$$

that is

$$\sum_{\substack{a_1,\ldots,a_m \pmod k \\ a_1+\ldots+a_m \equiv -n \pmod k}} f_1(a_1) \cdots f_m(a_m) = \frac{1}{k} \sum_{a \pmod k} \widehat{f_1}(a) \cdots \widehat{f_m}(a) e^{-2\pi i a n/k} \quad (n \in \mathbb{Z}).$$

For $n = 0$ we obtain

$$\sum_{\substack{a_1,\ldots,a_m \pmod k \\ a_1+\ldots+a_m \equiv 0 \pmod k}} f_1(a_1) \cdots f_m(a_m) = \frac{1}{k} \sum_{a \pmod k} \widehat{f_1}(a) \cdots \widehat{f_m}(a). \tag{4}$$

Now the result follows from (4) by showing the following property: Let f be a k-periodic function and let $h \in \mathbb{Z}$ such that $\gcd(h, k) = 1$. Then the DFT of the function g defined by $g(n) = f(nh)$ $(n \in \mathbb{Z})$ is $\hat{g}(n) = \hat{f}(nh')$ $(n \in \mathbb{Z})$, where h' is the multiplicative inverse of $h \pmod k$.

Indeed,

$$\hat{g}(n) = \sum_{a \pmod k} g(a) e^{-2\pi i a n/k} = \sum_{a \pmod k} f(ah) e^{-2\pi i a h h' n/k},$$

and since $\gcd(h, k) = 1$, if a runs through a complete system of residues $\pmod k$, then so does $b = ah$. Therefore,

$$\hat{g}(n) = \sum_{b \pmod k} f(b) e^{-2\pi i b h' n/k} = \hat{f}(nh').$$

□

Corollary 2.1. *Let f_1 and f_2 be k-periodic functions ($k \in \mathbb{N}$) and let $h_1, h_2 \in \mathbb{Z}$, $\gcd(h_1, k) = \gcd(h_2, k) = 1$. Then*

$$\sum_{a \pmod k} f_1(ah_1) f_2(ah_2) = \frac{1}{k} \sum_{a \pmod k} \widehat{f_1}(-ah_2)\widehat{f_2}(ah_1). \tag{5}$$

Proof. Apply Theorem 2.1 for $m = 2$. We deduce that

$$\begin{aligned}\sum_{a \pmod k} f_1(ah_1) f_2(-ah_2) &= \frac{1}{k} \sum_{a \pmod k} \widehat{f_1}(ah_1')\widehat{f_2}(ah_2') \\ &= \frac{1}{k} \sum_{a \pmod k} \widehat{f_1}(ah_1'h_2')\widehat{f_2}(ah_1'h_2'h_1) \\ &= \frac{1}{k} \sum_{b \pmod k} \widehat{f_1}(bh_2)\widehat{f_2}(bh_1),\end{aligned}$$

by using the fact that if a runs through a complete system of residues (mod k), then so does $b = ah_1'h_2'$, since $\gcd(h_1h_2, k) = 1$. This gives (5) by setting $h_2 := -h_2$. □

For $h_1 = 1, h_2 = -1$ from (5) we derive (3).

Corollary 2.2. *Let f_1 and f_2 be k-periodic functions ($k \in \mathbb{N}$) and assume that f_1 or f_2 is odd (resp. even). Let $h_1, h_2 \in \mathbb{Z}$, $\gcd(h_1, k) = \gcd(h_2, k) = 1$. Then*

$$\sum_{a \pmod k} f_1(ah_1) f_2(ah_2) = \frac{(-1)^s}{k} \sum_{a \pmod k} \widehat{f_1}(ah_2)\widehat{f_2}(ah_1), \tag{6}$$

where $s = 1$ if f_1 or f_2 is odd, $s = 0$ if f_1 or f_2 is even.

Note that if the function f is odd (resp. even), then $\hat{f}$ is also odd (resp. even). If one of the functions f_1, f_2 is odd and the other one is even, then both sides of (6) are zero.

In this paper we will use the following DFT pairs of k-periodic functions.

Lemma 2.1. *(i) Let $k \in \mathbb{N}$. The DFT of the k-periodic odd function $f(n) = \left(\left(\frac{n}{k}\right)\right)$ ($n \in \mathbb{Z}$) is*

$$\hat{f}(n) = \begin{cases} \frac{i}{2} \cot\left(\frac{\pi n}{k}\right), & \text{if } k \nmid n, \\ 0, & \text{if } k \mid n. \end{cases}$$

(ii) Let $k \in \mathbb{N}$ and let $\overline{B_r}$ ($r \in \mathbb{N}$) be the Bernoulli functions (cf. Sect. 3.1). The DFT of the k-periodic function $f(n) = \overline{B}_r\left(\frac{n}{k}\right)$ ($n \in \mathbb{Z}$) is

$$\hat{f}(n) = \begin{cases} rk^{1-r}\left(\frac{i}{2}\right)^r \cot^{(r-1)}\left(\frac{\pi n}{k}\right), & \text{if } k \nmid n, \\ B_r k^{1-r}, & \text{if } k \mid n, \end{cases}$$

where B_r is the r-th Bernoulli number and $\cot^{(m)}$ is the m-th derivative of the cotangent function.

(iii) Let $k \in \mathbb{N}$ be even. The DFT of the k-periodic odd function $f(n) = (-1)^n\left(\left(\frac{n}{k}\right)\right)$ $(n \in \mathbb{Z})$ is

$$\hat{f}(n) = \begin{cases} -\frac{i}{2}\tan\left(\frac{\pi n}{k}\right), & \text{if } n \not\equiv \frac{k}{2} \pmod{k}, \\ 0, & \text{if } n \equiv \frac{k}{2} \pmod{k}. \end{cases}$$

(iv) Let k be odd and let $n \pmod{k} = n - k\lfloor n/k \rfloor$ be the least nonnegative residue of $n \pmod{k}$. The DFT of the k-periodic odd function

$$f(n) = \begin{cases} (-1)^{n \pmod{k}}, & \text{if } k \nmid n, \\ 0, & \text{if } k \mid n \end{cases}$$

is

$$\hat{f}(n) = i\tan\left(\frac{\pi n}{k}\right) \quad (n \in \mathbb{Z}).$$

(v) Let $k \in \mathbb{N}$. Let $F(s,x)$, $\zeta(s,x)$, and $\zeta(s)$ be the periodic zeta function, the Hurwitz zeta function, and the Riemann zeta function, respectively (cf. Sect. 3.3). For $\Re s > 1$ the DFT of the k periodic function $f(n) = F(s, \frac{n}{k})$ $(n \in \mathbb{Z})$ is

$$\hat{f}(n) = \begin{cases} k^{1-s}\zeta\left(s, \left\{\frac{n}{k}\right\}\right), & \text{if } k \nmid n, \\ k^{1-s}\zeta(s), & \text{if } k \mid n. \end{cases}$$

Here (i) and (iv) are well known. They follow, together with (iii) and (v), by easy computations from the definition of the DFT. For (ii) we refer to [2, Lemma 6]. See also [1, Sect. 6].

3 Applications

3.1 Generalized Dedekind Sums

We first derive the following higher dimensional generalization of the identity (1), first deduced by Zagier [16, Th. p. 157], in a slightly different form, by applying some other arguments.

Theorem 3.2. *Let $k \in \mathbb{N}$, $m \in \mathbb{N}$ be even and let $h_j \in \mathbb{Z}$, $\gcd(h_j, k) = 1$ $(1 \le j \le m)$. Then*

$$\sum_{\substack{a_1,\dots,a_m \pmod k \\ a_1+\dots+a_m \equiv 0 \pmod k}} \left(\left(\frac{a_1 h_1}{k}\right)\right) \cdots \left(\left(\frac{a_m h_m}{k}\right)\right)$$

$$= \frac{(-1)^{m/2}}{2^m k} \sum_{a=1}^{k-1} \cot\left(\frac{\pi a h_1'}{k}\right) \cdots \cot\left(\frac{\pi a h_m'}{k}\right). \tag{7}$$

Proof. Apply Theorem 2.1 for $f_1 = \dots = f_m = f$, where $f(n) = \left(\left(\frac{n}{k}\right)\right)$ $(n \in \mathbb{Z})$. Use Lemma 2.1/(i). □

Note that if m is odd, then both sides of (7) are zero.

Corollary 3.3. *Assume that $m = 2$. Let $k \in \mathbb{N}$, $h_1, h_2 \in \mathbb{Z}$, $\gcd(h_1, k) = \gcd(h_2, k) = 1$. Then*

$$\sum_{a=1}^{k-1} \left(\left(\frac{a h_1}{k}\right)\right)\left(\left(\frac{a h_2}{k}\right)\right) = \frac{1}{4k} \sum_{a=1}^{k-1} \cot\left(\frac{\pi a h_1}{k}\right) \cot\left(\frac{\pi a h_2}{k}\right). \tag{8}$$

Identity (8) is the homogeneous version of (1) and is equivalent to (1).
Now consider the Bernoulli polynomials $B_r(x)$ $(r \ge 0)$, defined by

$$\frac{t e^{xt}}{e^t - 1} = \sum_{r=0}^{\infty} \frac{B_r(x)}{r!} t^r.$$

Here

$$B_1(x) = x - 1/2,\ B_2(x) = x^2 - x + 1/6,\ B_3(x) = x^3 - 3x^2/2 + x/2 \text{ and } B_r := B_r(0)$$

are the Bernoulli numbers. The Bernoulli functions $x \mapsto \overline{B}_r(x)$ are given by

$$\overline{B}_r(x) = B_r(\{x\}) \quad (x \in \mathbb{R}).$$

Note that

$$\overline{B}_1(x) = ((x)) \text{ for } x \notin \mathbb{Z},$$

but

$$\overline{B}_1(x) = -1/2 \ne 0 = ((x)) \text{ for } x \in \mathbb{Z}.$$

For $r_1, \ldots, r_m \in \mathbb{N}$, $h_1, \ldots, h_m \in \mathbb{Z}$ we define the higher dimensional Dedekind–Bernoulli sum by

$$s_{r_1,\ldots,r_m}(h_1,\ldots,h_m;k) := \sum_{\substack{a_1,\ldots,a_m \pmod k \\ a_1+\ldots+a_m \equiv 0 \pmod k}} \overline{B}_{r_1}\left(\frac{a_1h_1}{k}\right)\cdots\overline{B}_{r_m}\left(\frac{a_mh_m}{k}\right). \tag{9}$$

In the case when $m = 2$ and by setting $h_2 := -h_2$ we obtain the sum

$$s_{r_1,r_2}(h_1,-h_2;k) := \sum_{a \pmod k} \overline{B}_{r_1}\left(\frac{ah_1}{k}\right)\overline{B}_{r_2}\left(\frac{ah_2}{k}\right), \tag{10}$$

first investigated by Carlitz [7] and Mikolás [11]. See the paper of Beck [2] for further historical remarks.

Theorem 3.3. *Let $k, m, r_j \in \mathbb{N}$ be such that $A := r_1 + \ldots + r_m$ is even and $h_j \in \mathbb{Z}$, $\gcd(h_j, k) = 1$ $(1 \le j \le m)$. Then*

$$s_{r_1,\ldots,r_m}(h_1,\ldots,h_m;k) = \frac{B_{r_1}\cdots B_{r_m}}{k^{A-m+1}}$$
$$+ \frac{(-1)^{A/2}r_1\cdots r_m}{2^A k^{A-m+1}} \sum_{a=1}^{k-1} \cot^{(r_1-1)}\left(\frac{\pi ah_1'}{k}\right)\cdots\cot^{(r_m-1)}\left(\frac{\pi ah_m'}{k}\right). \tag{11}$$

Note that if A is odd, then the sum in (11) vanishes. If A is odd and there is at least one j such that $r_j \ge 3$, then $B_{r_j} = 0$ and the sum (9) vanishes as well.

Proof. Apply Theorem 2.1 and Lemma 2.1/(ii) to the functions

$$f_j(n) = \overline{B}_{r_j}\left(\frac{n}{k}\right) \quad (1 \le j \le m).$$

□

Corollary 3.4 ([2, Corollary 7]). *Let $k, r_1, r_2 \in \mathbb{N}$, $h_1, h_2 \in \mathbb{Z}$ be such that $r_1 + r_2$ is even and $\gcd(h_1, k) = \gcd(h_2, k) = 1$. Then*

$$\sum_{a \,(\mathrm{mod}\, k)} \overline{B}_{r_1}\left(\frac{ah_1}{k}\right)\overline{B}_{r_2}\left(\frac{ah_2}{k}\right)$$
$$= \frac{B_{r_1}B_{r_2}}{k^{r_1+r_2-1}} + \frac{(-1)^{(r_1-r_2)/2}r_1r_2}{2^{r_1+r_2}k^{r_1+r_2-1}} \sum_{a=1}^{k-1} \cot^{(r_1-1)}\left(\frac{\pi ah_1}{k}\right)\cot^{(r_2-1)}\left(\frac{\pi ah_2}{k}\right).$$

3.2 Generalized Hardy Sums

The Hardy sums (known also as Hardy–Berndt sums) are defined for $h \in \mathbb{Z}$, $k \in \mathbb{N}$ as follows:

$$S(h,k) := \sum_{a \pmod k} (-1)^{a+1+\lfloor ah/k \rfloor},$$

$$s_1(h,k) := \sum_{a \pmod k} (-1)^{\lfloor ah/k \rfloor} \left(\left(\frac{a}{k}\right)\right),$$

$$s_2(h,k) := \sum_{a \pmod k} (-1)^{a} \left(\left(\frac{a}{k}\right)\right)\left(\left(\frac{ah}{k}\right)\right),$$

$$s_3(h,k) := \sum_{a \pmod k} (-1)^{a} \left(\left(\frac{ah}{k}\right)\right),$$

$$s_4(h,k) := \sum_{a \pmod k} (-1)^{\lfloor ah/k \rfloor},$$

$$s_5(h,k) := \sum_{a \pmod k} (-1)^{a+\lfloor ah/k \rfloor} \left(\left(\frac{a}{k}\right)\right).$$

Berndt and Goldberg [6] derived finite and infinite series representations for the above sums. These identities were also obtained later by Sitaramachandrarao [14] by using some different arguments. One could see [2, 6, 8, 14] for the history of these sums as well as for further results on the Hardy sums, including reciprocity formulas.

We define the following generalization of $s_2(h,k)$:

$$A(h_1,\ldots,h_m;k) := \sum_{\substack{a_1,\ldots,a_m \pmod k \\ a_1+\ldots+a_m \equiv 0 \pmod k}} (-1)^{a_1} \left(\left(\frac{a_1h_1}{k}\right)\right)\cdots\left(\left(\frac{a_mh_m}{k}\right)\right).$$

Theorem 3.4. *Let $k,m \in \mathbb{N}$ be even, $h_1,\ldots,h_m \in \mathbb{Z}$, h_1 odd, $\gcd(h_j,k) = 1$ $(1 \le j \le m)$. Then*

$$A(h_1,\ldots,h_m;k) = \frac{(-1)^{m/2-1}}{2^m k} \sum_{\substack{1 \le a \le k-1 \\ a \ne k/2}} \tan\left(\frac{\pi a h_1'}{k}\right)\cot\left(\frac{\pi a h_2'}{k}\right)\cdots\cot\left(\frac{\pi a h_m'}{k}\right).$$

Proof. Let $f_1(n) = (-1)^n \left(\left(\frac{n}{k}\right)\right)$ and $f_j(n) = \left(\left(\frac{n}{k}\right)\right)$ $(2 \le j \le m)$. Apply Theorem 2.1 and Lemma 2.1/(i),(iv). □

Corollary 3.5. *Assume that* $m = 2$. *Let* $k \in \mathbb{N}$ *be even,* $h_1, h_2 \in \mathbb{Z}$, h_1 *odd, and* $\gcd(h_1, k) = \gcd(h_2, k) = 1$. *Then*

$$\sum_{a=1}^{k-1}(-1)^a \left(\left(\frac{ah_1}{k}\right)\right)\left(\left(\frac{ah_2}{k}\right)\right) = -\frac{1}{4k}\sum_{\substack{1\le a\le k-1\\ a\neq k/2}} \tan\left(\frac{\pi a h_2}{k}\right)\cot\left(\frac{\pi a h_1}{k}\right).$$

In the case $m = 2$, $h_1 = 1$, $h_2 = h$ we obtain the following corollary, cf. [6, Eq. (14)], [14, Eq. (7.3)].

Corollary 3.6. *If* $k \in \mathbb{N}$ *is even,* $h \in \mathbb{Z}$, $\gcd(h, k) = 1$, *then*

$$s_2(h,k) = -\frac{1}{4k}\sum_{\substack{1\le a\le k-1\\ a\neq k/2}} \tan\left(\frac{\pi a h}{k}\right)\cot\left(\frac{\pi a}{k}\right).$$

Next, we define the following common generalization of $s_1(h,k)$, $s_3(h,k)$ and $s_5(h,k)$, as follows.

$$B(h_1,\dots,h_m;k) := \sum_{\substack{a_1,\dots,a_m \pmod k\\ a_1\not\equiv 0 \pmod k\\ a_1+\dots+a_m\equiv 0 \pmod k}} (-1)^{a_1h_1+k\lfloor a_1h_1/k\rfloor}\left(\left(\frac{a_2h_2}{k}\right)\right)\cdots\left(\left(\frac{a_mh_m}{k}\right)\right).$$

Theorem 3.5. *Let* $k \in \mathbb{N}$ *be odd,* $m \in \mathbb{N}$ *be even,* $h_j \in \mathbb{Z}$, $\gcd(h_j, k) = 1$ *(*$1 \le j \le m$*). Then*

$$B(h_1,\dots,h_m;k) = \frac{(-1)^{m/2}}{2^{m-1}k}\sum_{a=1}^{k-1}\tan\left(\frac{\pi a h_1'}{k}\right)\cot\left(\frac{\pi a h_2'}{k}\right)\cdots\cot\left(\frac{\pi a h_m'}{k}\right).$$

Proof. Apply Theorem 2.1 to the following functions:

$$f_1(n) = \begin{cases}(-1)^{n \pmod k}, & \text{if } k \nmid n,\\ 0, & \text{if } k \mid n,\end{cases}$$

$$f_j(n) = \left(\left(\frac{n}{k}\right)\right) \quad (2 \le j \le m)$$

and also Lemma 2.1/(iv). □

Corollary 3.7. *Assume that* $m = 2$. *Let* $k \in \mathbb{N}$ *be odd,* $h_1, h_2 \in \mathbb{Z}$, h_1 *odd,* $\gcd(h_1, k) = \gcd(h_2, k) = 1$. *Then*

$$\sum_{a=1}^{k-1}(-1)^{a+\lfloor ah_1/k\rfloor}\left(\left(\frac{ah_2}{k}\right)\right) = \frac{1}{2k}\sum_{a=1}^{k-1}\tan\left(\frac{\pi a h_2}{k}\right)\cot\left(\frac{\pi a h_1}{k}\right).$$

For $m = 2$ in the special cases $h_1 = 1$, $h_2 = h$ and $h_1 = h$, $h_2 = 1$, respectively, we obtain the identities, cf. [6, Eqs. (15), (17)], [14, Eqs. (7.4), (7.6)], as follows:

Corollary 3.8. *If $k \in \mathbb{N}$ is odd, $h \in \mathbb{Z}$, $\gcd(h,k) = 1$, then*

$$s_3(h,k) = \frac{1}{2k} \sum_{a=1}^{k-1} \tan\left(\frac{\pi ah}{k}\right) \cot\left(\frac{\pi a}{k}\right). \tag{12}$$

If $k \in \mathbb{N}$ is odd, $h \in \mathbb{Z}$ is odd, $\gcd(h,k) = 1$, then

$$s_5(h,k) = \frac{1}{2k} \sum_{a=1}^{k-1} \tan\left(\frac{\pi a}{k}\right) \cot\left(\frac{\pi ah}{k}\right).$$

Corollary 3.9. *Assume that $m = 2$. Let $k \in \mathbb{N}$ be odd, $h_1, h_2 \in \mathbb{Z}$, h_1 even, $\gcd(h_1,k) = \gcd(h_2,k) = 1$. Then*

$$\sum_{a=1}^{k-1} (-1)^{\lfloor ah_1/k \rfloor} \left(\left(\frac{ah_2}{k}\right)\right) = \frac{1}{2k} \sum_{a=1}^{k-1} \tan\left(\frac{\pi ah_2}{k}\right) \cot\left(\frac{\pi ah_1}{k}\right).$$

For $m = 2$ in the special case $h_1 = h$ even, $h_2 = 1$ we obtain the identity, cf. [6, Eq. (13)], [14, Eq. (7.2)], as follows.

Corollary 3.10. *If $k \in \mathbb{N}$ is odd, $h \in \mathbb{Z}$ is even, $\gcd(h,k) = 1$, then*

$$s_1(h,k) = \frac{1}{2k} \sum_{a=1}^{k-1} \tan\left(\frac{\pi a}{k}\right) \cot\left(\frac{\pi ah}{k}\right). \tag{13}$$

Note that the Hardy sums $S(h,k)$ and $s_4(h,k)$ can also be treated with the DFT in the case when k is odd. For example, applying Corollary 2.2 to the functions

$$f_1(n) = f_2(n) = \begin{cases} (-1)^{n \pmod k}, & \text{if } k \nmid n, \\ 0, & \text{if } k \mid n \end{cases}$$

we obtain the following representation.

Corollary 3.11. *If $k \in \mathbb{N}$ is odd, $h_1, h_2 \in \mathbb{Z}$, $\gcd(h_1,k) = \gcd(h_2,k) = 1$, then*

$$\sum_{a=1}^{k-1} (-1)^{a(h_1+h_2) \pmod k} = \frac{1}{k} \sum_{a=1}^{k-1} \tan\left(\frac{\pi ah_1}{k}\right) \tan\left(\frac{\pi ah_2}{k}\right). \tag{14}$$

If $h_2 = 1$ and $h_1 = h$ is odd, then the left-hand side of (14) is exactly $s_4(h,k)$. See [6, Eq. (16)], [14, Eq. (7.5)]. If $h_1 = h_2 = 1$, then (14) provides the following classical identity, valid for $k \in \mathbb{N}$ odd, cf. [3, Proposition 3.1]:

$$\sum_{a=1}^{k-1} \tan^2\left(\frac{\pi a}{k}\right) = k^2 - k.$$

Remark 3.1. For k odd, h even, the formula [6, Eq. (13)] receives the following representation:

$$s_1(h,k) = -\frac{1}{2k} \sum_{\substack{j=1 \\ j\neq(k+1)/2}}^{k} \cot\left(\frac{\pi h(2j-1)}{2k}\right) \cot\left(\frac{\pi(2j-1)}{2k}\right),$$

which can easily be transformed into

$$s_1(h,k) = \frac{1}{k} \sum_{j=1}^{(k-1)/2} \tan\left(\frac{\pi j}{k}\right) \cot\left(\frac{\pi hj}{k}\right)$$

that is equal to the right-hand side of (13). Similar considerations are valid for the corresponding formulas on the Hardy sums $s_4(h,k)$, $s_5(h,k)$, and $S(h,k)$.

Remark 3.2. The finite sum identities (7.1)–(7.3), (7.5), and (7.6) from the paper [14] contain some misprints. Namely, in formulas (7.1) and (7.5) the sum $\sum_{r=1}^{k-1}$ should be $\sum_{r=1}^{k}$, while in (7.2) and (7.6) the sum $\sum_{r=1,r\neq(k+1)/2}^{k-1}$ should be $\sum_{r=1,r\neq(k+1)/2}^{k}$, the missing terms being nonzero. Furthermore, in formula (7.3) the sum $\sum_{r=1}^{k-1}$ should be $\sum_{r=1,r\neq k/2}^{k-1}$, the term for $r = k/2$ (namely $\tan(\pi/2)$) being not defined.

One could possibly investigate some further higher dimensional generalizations and analogues of the Hardy sums involving the Bernoulli functions, however we do not discuss this in the present paper.

3.3 Sums Involving the Hurwitz Zeta Function

Theorem 2.1 and its corollaries can be applied in several other situations as well. For example, let

$$\zeta(s,x) := \sum_{n=0}^{\infty} \frac{1}{(n+x)^s}$$

be the Hurwitz zeta function, where $0 < x \leq 1$ and $\zeta(s,1) = \zeta(s)$ is the Riemann zeta function. The function

$$D(h,k) := \sum_{a=1}^{k-1} \zeta\left(s_1, \left\{\frac{ah_1}{k}\right\}\right) \zeta\left(s_2, \left\{\frac{ah_2}{k}\right\}\right),$$

investigated by Mikolás [11], is an analogue of the Dedekind sum (10), taking into account that

$$B_n(x) = -n\zeta(1-n,x) \quad (n \in \mathbb{N},\ 0 < x \le 1).$$

Let

$$F(s,x) := \sum_{n=1}^{\infty} \frac{e^{2\pi inx}}{n^s} \quad (x \in \mathbb{R}) \tag{15}$$

be the periodic zeta function, which converges for $\Re s > 0$ if $x \notin \mathbb{Z}$ and for $\Re s > 1$ if $x \in \mathbb{Z}$.

Applying Corollary 2.1 to the functions

$$f_j(n) = F\left(s_j, \frac{n}{k}\right) \quad (n \in \mathbb{Z},\ 1 \le j \le 2),$$

we deduce by Lemma 2.1/(v) the next new result.

Theorem 3.6. *Let $k \in \mathbb{N}$, $h_1, h_2 \in \mathbb{Z}$, $\gcd(h_1,k) = \gcd(h_2,k) = 1$ and let $s_1, s_2 \in \mathbb{C}$, $\Re s_1, \Re s_2 > 1$. Then*

$$D(h,k) = (k^{s_1+s_2-1} - 1)\zeta(s_1)\zeta(s_2) + k^{s_1+s_2-1} \sum_{a=1}^{k-1} F\left(s_1, \frac{ah_2}{k}\right) F\left(s_2, -\frac{ah_1}{k}\right).$$

4 Some Further Remarks

The following simple and useful result can be applied to obtain infinite series representations for the Dedekind and Hardy sums.

Lemma 4.2. *If $f : \mathbb{N} \to \mathbb{C}$ is a k-periodic ($k \in \mathbb{N}$) odd function, then*

$$S(f) := \sum_{r=1}^{\infty} \frac{f(r)}{r} = \frac{\pi}{2k} \sum_{r=1}^{k-1} f(r) \cot\left(\frac{\pi r}{k}\right) \tag{16}$$

$$= -\frac{\pi i}{k^2} \sum_{r=1}^{k-1} r\hat{f}(r). \tag{17}$$

For the Dedekind sum, (2) is a direct consequence of identities (16) and (1). As another example, (16) and (12) imply that for k odd, $\gcd(h,k)=1$, one has

$$s_3(h,k)=\frac{1}{\pi}\sum_{r=1}^{\infty}\frac{1}{r}\tan\left(\frac{\pi rh}{k}\right).$$

One could see [6, Theorem 1], [14, Theorem 7.1] for the above formula as well as for similar representations regarding Hardy sums.

Identity (16) of Lemma 4.2 was proved in [14, Lemma 2.1] by applying results of Lehmer [10] on the generalized Euler constants $\gamma(r,k)$ associated with the infinite arithmetic progression $r, r+k, r+2k, \ldots$ $(1\le r\le k)$, where

$$\gamma(r,k):=\lim_{x\to\infty}\left(\sum_{\substack{1\le n\le x\\ n\equiv r \pmod k}}\frac{1}{n}-\frac{1}{k}\log x\right).$$

Berndt [5] deduced (17) by contour integration (with a different definition of the DFT). The fact that the finite sums (16) and (17) are equal provides another simple consequence of Corollary 2.2, applied to the odd functions $f:\mathbb{N}\to\mathbb{C}$ and $n\mapsto\left(\left(\frac{n}{k}\right)\right)$.

Furthermore, according to [10, Theorem 8], if f is a k-periodic function, then

$$S(f)=\sum_{r=1}^{k}f(r)\gamma(r,k),$$

provided that $\sum_{r=1}^{k}f(k)=0$, representing a necessary and sufficient condition for the convergence of the series $S(f)$ (which holds if f is a k-periodic and odd function).

We note that the DFT of the k-periodic function $r\mapsto\gamma(r,k)$ is

$$\hat{\gamma}(r,k)=\begin{cases}F\left(1,-\frac{r}{k}\right), & \text{if } k\nmid r,\\ \gamma, & \text{if } k\mid r\end{cases}$$

(cf. [10, p. 127]), where $F(s,x)$ is the periodic zeta function defined by (15) and $\gamma:=\gamma(0,1)$ is Euler's constant. Therefore, we deduce by Corollary 2.1 the next identity.

Corollary 4.12. *If $f:\mathbb{N}\to\mathbb{C}$ is a k-periodic ($k\in\mathbb{N}$) odd function, then*

$$S(f)=-\frac{1}{k}\sum_{r=1}^{k-1}\hat{f}(r)F\left(1,-\frac{r}{k}\right).$$

Acknowledgements The authors would like to thank the referee for useful remarks which helped improve the presentation of the paper.

References

1. G. Almkvist, Asymptotic formulas and generalized Dedekind sums. Exp. Math. **7**, 343–359 (1998)
2. M. Beck, Dedekind cotangent sums. Acta Arith. **109**, 109–130 (2003)
3. M. Beck, M. Halloran, Finite trigonometric character sums via discrete Fourier analysis. Int. J. Number Theory **6**, 51–67 (2010)
4. M. Beck, S. Robins, *Computing the Continuous Discretely: Integer-Point Enumeration in Polyhedra*. Undergraduate Texts in Mathematics (Springer, New York, 2007)
5. B.C. Berndt, Solution to problem E2719. Am. Math. Mon. **86**, 786–788 (1979)
6. B.C. Berndt, L.A. Goldberg, Analytic properties of arithmetic sums arising in the theory of the classical theta-functions. SIAM J. Math. Anal. **15**, 143–150 (1984)
7. L. Carlitz, Some theorems on generalized Dedekind sums. Pac. J. Math. **3**, 513–522 (1953)
8. U. Dieter, Cotangent sums, a further generalization of Dedekind sums. J. Number Theory **18**, 289–305 (1984)
9. H. Iwaniec, E. Kowalski, *Analytic Number Theory*, vol. 53 (American Mathematical Society Colloquium Publications, Providence, 2004)
10. D.H. Lehmer, Euler constants for arithmetical progressions. Acta Arith. **27**, 125–142 (1975)
11. M. Mikolás, On certain sums generating the Dedekind sums and their reciprocity laws. Pac. J. Math. **7**, 1167–1178 (1957)
12. H. Rademacher, Egy reciprocitásképletről a modulfüggvények elméletéből. Mat. Fiz. Lapok **40** (1933), 24–34; *Collected Papers II* (MIT, Cambridge, 1974), pp. 26–37
13. H. Rademacher, E. Grosswald, *Dedekind Sums*. Carus Mathematical Monograph, vol. 16 (Mathematical Association of America, Washington, 1972)
14. R. Sitaramachandrarao, Dedekind, and Hardy sums. Acta Arith. **48**, 325–340 (1987)
15. A. Terras, *Fourier Analysis on Finite Groups and Applications*. London Mathematical Society Student Texts, vol. 43 (Cambridge University Press, Cambridge, 1999)
16. D. Zagier, Higher dimensional Dedekind sums. Math. Ann. **202**, 149–172 (1973)

On Arithmetic Properties of Products and Shifted Products

Joël Rivat and András Sárközy

Dedicated to Helmut Maier on the occasion of his 60th birthday

Abstract Let $\mathscr{A}$ and $\mathscr{B}$ be large subsets of $\{1\dots,N\}$. Arithmetic properties of the products ab, resp. of the shifted products $ab+1$ with $a\in\mathscr{A}$, $b\in\mathscr{B}$ are studied. In particular, it is shown that the sum of digits of the products ab is well distributed, and the size of the subsets $\mathscr{A},\mathscr{B}\in\{1\dots,N\}$ with the property that $ab+1$ is always squarefree is estimated.

Keywords Product of sequences • Shifted products • Digit properties • Squares • Multiplicative characters • Large sieve

2010 Mathematics Subject Classification: Primary 11N25. Secondary 11K45, 11N36.

1 Introduction

Throughout this paper we will use the following notations: $\mathbb{N}$ and $\mathbb{R}$ denote the set of positive integers, resp. real numbers. $\omega(n)$ denotes the number of distinct prime factors of n, $\varphi(n)$ is Euler's function, and $\pi(x)$ is the number of primes not exceeding x. The notation $\sum_{\chi \bmod m}$ means that χ runs over the (multiplicative)

J. Rivat (✉)
Université d'Aix-Marseille, Institut de Mathématiques de Marseille, CNRS UMR 7373, 163 avenue de Luminy, Case 907, 13288 Marseille Cedex 9, France
e-mail: rivat@iml.univ-mrs.fr

A. Sárközy
Department of Algebra and Number Theory, Eötvös Loránd University, H-1117 Budapest, Pázmány Péter sétány 1/c, Hungary
e-mail: sarkozy@cs.elte.hu

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_21

characters modulo m, while $\sum_{\chi \bmod m}^{*}$ denotes that χ runs over the primitive characters modulo m. $\|\dots\|$ denotes the distance to the nearest integer and $\mathrm{e}(t) = \exp(2i\pi t)$.

In [1] we presented a possibly complete list of the papers written on arithmetic properties of sumsets, and we studied several further problems of this type. This paper is the multiplicative analogue of Balog et al. [1]: our goal is to present the possibly complete list of the papers written on arithmetic properties of products and shifted products. Note that when we are looking for the multiplicative analogue of a problem on a certain arithmetic property of sums $a + b$ with a, b taken from "large" sets $\mathscr{A}$, $\mathscr{B}$ of integers, then the most natural way is to replace $a + b$ in the given problem by the product ab. However, if the arithmetic property studied is of multiplicative nature, then the problem obtained in this way is usually trivial; in such a case, one may help by taking *shifted products* $ab + 1$ instead of the products ab. For example, Erdős et al. [11] proved that if $\mathscr{A}$, $\mathscr{B}$ are large subsets of $\{1, \dots, N\}$ then the sums $a + b$ with $a \in \mathscr{A}$, $b \in \mathscr{B}$ satisfy an Erdős–Kac type theorem, i.e., the distribution of $(\omega(a + b) - \log\log N)/(\log\log N)^{1/2}$can be approximated by the normal distribution (see also [8] and [38]). If we replace $\omega(a + b)$ by $\omega(ab)$, then we do not get an interesting problem, since because of the additivity of the ω function we may separate a and b in $\omega(ab)$: $\omega(ab) = \omega(a) + \omega(b)$ if $(a, b) = 1$. On the other hand, considering $\omega(ab + 1)$ instead of $\omega(ab)$, again we get an interesting problem; indeed, Elliott and Sárközy [9] showed that the shifted products $ab + 1$ also satisfy an Erdős–Kac type theorem.

The first problem on arithmetic properties of (shifted) products was studied by Diophantus more than 2000 years ago. He asked the following question: how many positive integers $a_1 < a_2 < \cdots < a_k$ can be given so that $a_i a_j + 1$ is a square for all $1 \leqslant i < j \leqslant k$? Fermat and Euler also studied this problem, and finally Dujella [7] settled this problem in 2004. In the last 50 years many further papers have been written on arithmetic properties of products and shifted products: [2–4, 9, 12, 15–17, 17–22, 27–37].

In this paper we will study two further problems on arithmetic properties of products and shifted products. First in Sects. 2 and 3 we will show that the sum of digits of the products ab with $a \in \mathscr{A}, b \in \mathscr{B}$ (where $\mathscr{A}$ and $\mathscr{B}$ are large subsets of $\{1 \dots, N\}$) is well distributed, and then in Sects. 4 and 5 we will study the following question: how large can be the subsets $\mathscr{A}$, $\mathscr{B}$ of $\{1, \dots, N\}$ with the property that $ab + 1$ is squarefree for every $a \in \mathscr{A}, b \in \mathscr{B}$?

2 The Sum of Digits of Products *ab*

Let $q \geqslant 2$ be an integer. Every integer $n \geqslant 0$ can be written uniquely in base q:

$$n = \sum_{k \geqslant 0} n_k q^k$$

where the digits $n_k \in \{0, \dots, q-1\}$. The sum of digits function $\mathrm{s}_q(n)$ is then defined by

$$\mathrm{s}_q(n) = \sum_{k \geqslant 0} n_k.$$

The sum of digits function appears in many mathematical problems (see [23] for a survey of these questions). Gelfond [14] proved that the sum of digits function is well distributed over arithmetic progressions and conjectured that a similar result holds true over the prime numbers and over polynomial sequences. Recently Mauduit and Rivat proved Gelfond's conjecture for any basis $q \geqslant 2$ both for the prime numbers [25] and for the squares [24], while Drmota, Mauduit and Rivat [6] managed to prove Gelfond's conjecture for polynomials for a sufficiently large basis q (depending on the degree of the polynomial).

In this paper our aim is to prove the analogue of these results for products ab with $a \in \mathscr{A} \subseteq \{1, \dots, N\}$ and $b \in \mathscr{B} \subseteq \{1, \dots, N\}$, namely

Theorem 2.1. *For any integer $q \geqslant 2$ there exists an explicit constant $c_q > 0$ [see (4)] such that for any subsets $\mathscr{A}$, $\mathscr{B}$ of $\{1, \dots, N\}$ we have*

$$\left| \sum_{a \in \mathscr{A}} \sum_{b \in \mathscr{B}} \mathrm{e}(\alpha\, \mathrm{s}_q(ab)) \right| \ll_q N^{2(1-c_q \|(q-1)\alpha\|^2)} \log N \tag{1}$$

uniformly for $\alpha \in \mathbb{R}$, where $\| \cdot \|$ denotes the distance to the nearest integer and $\mathrm{e}(t) = \exp(2i\pi t)$.

(It follows from this theorem that if m grows "slowly" in terms of q and N, then the numbers $\mathrm{s}_q(ab)$ are well distributed modulo m.)

3 Proof of Theorem 2.1

It follows from Proposition 3.2 of [5] (which is a variant of Proposition 1 of [25]) that there exists $\delta > 0$ such that, uniformly for all complex numbers a_m and b_n with $|a_m| \leqslant 1$ and $|b_n| \leqslant 1$, and all $\alpha \in \mathbb{R}$ we have

$$\left| \sum_{q^{\mu-1} < m \leqslant q^{\mu}} \sum_{q^{\nu-1} < n \leqslant q^{\nu}} a_m\, b_n\, \mathrm{e}(\alpha s_q(mn)) \right| \ll_q (\mu+\nu)\, q^{(\mu+\nu)(1-c_q\|(q-1)\alpha\|^2)}, \tag{2}$$

whenever μ and ν are positive integers satisfying

$$\frac{1}{3} - \delta \leqslant \frac{\mu}{\mu+\nu} \leqslant \frac{2}{3} + \delta, \tag{3}$$

where

$$c_q := \frac{1}{28}\min\left(4\,\omega_q, \frac{\pi^2}{12}\frac{q-1}{(q+1)\log q}\right). \tag{4}$$

with

$$\omega_2 = 1 - \frac{\log(2+\sqrt{2})}{2\log 2}; \quad \omega_q = \left(\frac{3}{2} - \frac{\log 5}{\log 3}\right)\frac{\log 2}{\log q} \quad \text{for } q \geqslant 3.$$

If $\mathscr{A}$, $\mathscr{B}$ are subsets of $\{1,\dots,N\}$, we can define

$$\mathscr{A}' = \mathscr{A} \cap [N^{1-\frac{c_q}{4}}, N], \quad \mathscr{B}' = \mathscr{B} \cap [N^{1-\frac{c_q}{4}}, N].$$

Then

$$\sum_{a\in\mathscr{A}}\sum_{b\in\mathscr{B}} \mathrm{e}(\alpha\, \mathrm{s}_q(ab)) = \sum_{a\in\mathscr{A}'}\sum_{b\in\mathscr{B}'} \mathrm{e}(\alpha\, \mathrm{s}_q(ab)) + O(N^{2-\frac{c_q}{4}}).$$

Observing that $\|(q-1)\alpha\|^2 \leqslant \frac{1}{4}$ we deduce that in (1) we may replace $\mathscr{A}$ by $\mathscr{A}'$ and $\mathscr{B}$ by $\mathscr{B}'$ at the price of a negligible error term. Therefore from now on we may assume that

$$\mathscr{A} \subset [N^{1-\frac{c_q}{4}}, N], \quad \mathscr{B} \subset [N^{1-\frac{c_q}{4}}, N]. \tag{5}$$

For positive integers μ and ν we define

$$\mathscr{A}_\mu = \mathscr{A} \cap [q^{\mu-1}+1, q^\mu], \quad \mathscr{B}_\nu = \mathscr{B} \cap [q^{\nu-1}+1, q^\nu].$$

We have

$$\mathscr{A} = \bigcup_{\mu=\mu_0}^{\mu_1} \mathscr{A}_\mu, \quad \mathscr{B} = \bigcup_{\nu=\nu_0}^{\nu_1} \mathscr{B}_\nu$$

with

$$\mu_0 = \nu_0 = 1 + \left\lfloor \left(1 - \frac{c_q}{4}\right)\frac{\log N}{\log q} \right\rfloor, \quad \mu_1 = \nu_1 = \left\lceil \frac{\log N}{\log q} \right\rceil.$$

Hence for N large enough

$$\frac{1}{3} < \frac{\mu_0}{\mu_1+\nu_1} \leqslant \frac{\mu}{\mu+\nu} \leqslant \frac{\mu_1}{\mu_0+\nu_0} < \frac{2}{3}$$

which implies that condition (3) is satisfied (actually for any $\delta > 0$). It follows that we can apply (2) with

$$a_m = \begin{cases} 1 \text{ if } m \in \mathscr{A}_\mu, \\ 0 \text{ otherwise.} \end{cases}; \quad b_n = \begin{cases} 1 \text{ if } n \in \mathscr{B}_\nu, \\ 0 \text{ otherwise.} \end{cases}$$

We obtain

$$\left| \sum_{a \in \mathscr{A}_\mu} \sum_{b \in \mathscr{B}_\nu} \mathrm{e}(\alpha\, \mathrm{s}_q(ab)) \right| \ll_q (\mu + \nu)\, q^{(\mu+\nu)(1-c_q\|(q-1)\alpha\|^2)}. \tag{6}$$

We have for all $\theta \in (0, 1)$

$$\sum_{\mu=\mu_0}^{\mu_1} q^{\mu\theta} \leqslant q^{(1+\mu_1)\theta} \ll_q N^\theta, \quad \sum_{\nu=\nu_0}^{\nu_1} q^{\nu\theta} \leqslant q^{(1+\nu_1)\theta} \ll_q N^\theta,$$

and $\mu + \nu \leqslant \mu_1 + \nu_1 \ll \log N$. Using these estimates with $\theta = 1 - c_q\|(q-1)\alpha\|^2$ together with (6), we can sum non-trivially over μ and ν and obtain

$$\begin{aligned} \left| \sum_{a \in \mathscr{A}} \sum_{b \in \mathscr{B}} \mathrm{e}(\alpha\, \mathrm{s}_q(ab)) \right| &\leqslant \sum_{\mu=\mu_0}^{\mu_1} \sum_{\nu=\nu_0}^{\nu_1} \left| \sum_{a \in \mathscr{A}_\mu} \sum_{b \in \mathscr{B}_\nu} \mathrm{e}(\alpha\, \mathrm{s}_q(ab)) \right| \\ &\ll_q N^{2(1-c_q\|(q-1)\alpha\|^2)} \log N. \end{aligned}$$

which is (1).

4 Squarefree Shifted Products

Erdős and Sárközy [10] studied the following problem: how large can a subset $\mathscr{A} \subset \{1, \ldots, N\}$ be with the property that $a + a'$ is squarefree for all $a \in \mathscr{A}$, $a' \in \mathscr{A}$? They proved the following results:

Theorem A *For $N > N_0$ there exists a subset $\mathscr{A} \subset \{1, \ldots, N\}$ such that*

$$|\mathscr{A}| > \frac{1}{248} \log N$$

and $a + a'$ is squarefree for all $a \in \mathscr{A}$, $a' \in \mathscr{A}$.

Theorem B *If $N > N_1$, $\mathscr{A} \subset \{1, \ldots, N\}$ and $a + a'$ is squarefree for all $a \in \mathscr{A}$, $a' \in \mathscr{A}$, then we have*

$$|\mathscr{A}| < 3N^{3/4} \log N. \tag{7}$$

Next we will study the multiplicative analogue of this problem in the more general form that we take *two* subsets $\mathscr{A}, \mathscr{B} \subset \{1, \ldots, N\}$ such that the shifted products $ab + 1$ are squarefree for all $a \in \mathscr{A}$, $b \in \mathscr{B}$, and the question is that how large is $M(N) = \max_{\mathscr{A},\mathscr{B}} |\mathscr{A}|\,|\mathscr{B}|$for subsets $\mathscr{A}$, $\mathscr{B}$ with these properties? We will prove the following theorem in Sect. 5:

Theorem 4.2. *If $N > N_2$, $\mathscr{A}, \mathscr{B} \subset \{1, \ldots, N\}$ and $ab + 1$ is squarefree for all $a \in \mathscr{A}$, $b \in \mathscr{B}$, then we have*

$$\min(|\mathscr{A}|, |\mathscr{B}|) < 2N^{3/4} \log N. \tag{8}$$

Note that for $\mathscr{A} = \mathscr{B}$ this gives

$$|\mathscr{A}| < 2N^{3/4} \log N.$$

(Compare this with (7) in the additive case.)

It is stated in [10] (without proof) that for $N \to \infty$ there exist sets $\mathscr{A}, \mathscr{B} \subset \{1, \ldots, N\}$ such that $|\mathscr{A}|\,|\mathscr{B}|/N \to \infty$ and the sums $a + b$ with $a \in \mathscr{A}$, $b \in \mathscr{B}$ are all squarefree. Indeed, with some work one can construct such sets $\mathscr{A}$, $\mathscr{B}$ with $|\mathscr{A}| \gg N$, $|\mathscr{B}| \gg \log N$. In the case of shifted products we have not been able to prove such a result. If we take $\mathscr{A} = \{n : \ 1 \leqslant n \leqslant N, \ n + 1 \text{ is squarefree}\}$ and $\mathscr{B} = \{1\}$, then we have $|\mathscr{A}|\,|\mathscr{B}| = (1 + o(1))\frac{6}{\pi^2}N$ and $ab + 1$ is squarefree for all $a \in \mathscr{A}$, $b \in \mathscr{B}$, thus trivially we have

$$(1 + o(1))\frac{6}{\pi^2}N < M(N).$$

On the other hand, by Theorem 4.2 we have

$$M(N) < 2\, N^{7/4} \log N.$$

There is a large gap between these lower and upper bounds. We think that the trivial lower bound is closer to the value of $M(N)$ and, perhaps, there are $c_1, c_2 > 0$ with

$$N(\log N)^{c_1} < M(N) < N(\log N)^{c_2}.$$

5 Proof of Theorem 4.2

Theorem B was proved by using the large sieve with prime square moduli. Instead, here we will use Gallagher's (multiplicative) character version of the large sieve [13]. Using the large sieve with optimal constants [26], Gallagher's sieve becomes

Lemma 5.1. *For $t \geqslant 1$, $M \in \mathbb{Z}$, $N \in \mathbb{N}$ and any complex numbers $c_{M+1}, c_{M+2}, \dots, c_{M+N}$ we have*

$$\sum_{k \leqslant t} \frac{k}{\varphi(k)} \sideset{}{^*}\sum_{\chi \bmod k} \left| \sum_{n=M+1}^{M+N} c_n \chi(n) \right|^2 \leqslant (N - 1 + t^2) \sum_{n=M+1}^{M+N} |c_n|^2 .$$

Now assume that contrary to the statement of Theorem 4.2 we have

$$\min(|\mathscr{A}|, |\mathscr{B}|) \geqslant 2N^{3/4} \log N, \tag{9}$$

however, $ab + 1$ is squarefree for all $a \in \mathscr{A}$, $b \in \mathscr{B}$.

Then for every prime p we have

$$\begin{aligned} 0 &= \left|\{(a, b) : a \in \mathscr{A},\ b \in \mathscr{B},\ ab + 1 \equiv 0 \bmod p^2\}\right| \\ &= \frac{1}{\varphi(p^2)} \sum_{\chi \bmod p^2} \overline{\chi}(-1) \sum_{a \in \mathscr{A}} \sum_{b \in \mathscr{B}} \chi(ab), \end{aligned}$$

so that

$$\overline{\chi_0}(-1) \sum_{a \in \mathscr{A}} \sum_{b \in \mathscr{B}} \chi_0(ab) = - \sum_{\substack{\chi \bmod p^2 \\ \chi \neq \chi_0}} \overline{\chi}(-1) \sum_{a \in \mathscr{A}} \sum_{b \in \mathscr{B}} \chi(ab),$$

i.e.

$$|\{(a, b),\ a \in \mathscr{A},\ b \in \mathscr{B},\ (ab, p) = 1\}| = - \sum_{\substack{\chi \bmod p^2 \\ \chi \neq \chi_0}} \overline{\chi}(-1) \sum_{a \in \mathscr{A}} \chi(a) \sum_{b \in \mathscr{B}} \chi(b),$$

whence

$$|\{(a, b),\ a \in \mathscr{A},\ b \in \mathscr{B},\ (ab, p) = 1\}| \leqslant \sum_{\substack{\chi \bmod p^2 \\ \chi \neq \chi_0}} \left| \sum_{a \in \mathscr{A}} \chi(a) \right| \left| \sum_{b \in \mathscr{B}} \chi(b) \right|$$

(for every p). Write $Q = N^{1/4}$. Then consider this inequality for every p with $Q/2 < p \leqslant Q$, and add these inequalities. Denoting the sum on the left-hand side by S, we get

$$\begin{aligned} S &= \sum_{Q/2 < p \leqslant Q} |\{(a, b),\ a \in \mathscr{A},\ b \in \mathscr{B},\ (ab, p) = 1\}| \qquad (10) \\ &\leqslant \sum_{Q/2 < p \leqslant Q} \sum_{\substack{\chi \bmod p^2 \\ \chi \neq \chi_0}} \left| \sum_{a \in \mathscr{A}} \chi(a) \right| \left| \sum_{b \in \mathscr{B}} \chi(b) \right|. \end{aligned}$$

By (9) and the prime number theorem we have

$$S = \sum_{Q/2<p\leqslant Q} |\{(a,b),\ a\in\mathscr{A},\ b\in\mathscr{B},\ (ab,p)=1\}| \tag{11}$$

$$\geqslant \sum_{Q/2<p\leqslant Q} |\{(a,b),\ a\in\mathscr{A},\ b\in\mathscr{B}\}|$$

$$- |\{(a,b),\ a\in\mathscr{A},\ b\in\mathscr{B},\ p\mid b\}| - |\{(a,b),\ a\in\mathscr{A},\ b\in\mathscr{B},\ p\mid a\}|$$

$$\geqslant |\mathscr{A}|\,|\mathscr{B}|\,(\pi(Q)-\pi(Q/2)) - \sum_{Q/2<p\leqslant Q}\left(|\mathscr{A}|\,\frac{N}{p}+|\mathscr{B}|\,\frac{N}{p}\right)$$

$$\geqslant |\mathscr{A}|\,|\mathscr{B}|\,(\pi(Q)-\pi(Q/2)) - (|\mathscr{A}|+|\mathscr{B}|)\,\frac{N}{Q/2}(\pi(Q)-\pi(Q/2))$$

$$= \left(|\mathscr{A}|\,|\mathscr{B}| - 2\,(|\mathscr{A}|+|\mathscr{B}|)\,N^{3/4}\right)(\pi(Q)-\pi(Q/2))$$

$$= (1+o(1))\,|\mathscr{A}|\,|\mathscr{B}|\,\frac{Q}{2\log Q/2} = (2+o(1))\,|\mathscr{A}|\,|\mathscr{B}|\,\frac{N^{1/4}}{\log N}.$$

On the other hand, it follows from (10) by the Cauchy–Schwarz inequality that

$$S \leqslant \left(\sum_{Q/2<p\leqslant Q}\ \sum_{\substack{\chi \bmod p^2\\ \chi\neq\chi_0}}\left|\sum_{a\in\mathscr{A}}\chi(a)\right|^2\right)^{1/2}\left(\sum_{Q/2<p\leqslant Q}\ \sum_{\substack{\chi \bmod p^2\\ \chi\neq\chi_0}}\left|\sum_{b\in\mathscr{B}}\chi(b)\right|^2\right)^{1/2} \tag{12}$$

$$\leqslant S_{\mathscr{A}}^{1/2} S_{\mathscr{B}}^{1/2}$$

where

$$S_{\mathscr{A}} = \sum_{p\leqslant Q}\ \sum_{\substack{\chi \bmod p^2\\ \chi\neq\chi_0}}\left|\sum_{a\in\mathscr{A}}\chi(a)\right|^2$$

and $S_{\mathscr{B}}$ is defined similarly with $\mathscr{B}$ in place of $\mathscr{A}$. A non-principal character χ modulo p^2 is not primitive character modulo p^2 if and only if it is a primitive character modulo p. Thus we may rewrite and estimate the sum $S_{\mathscr{A}}$ in the following way:

$$S_{\mathscr{A}} = \sum_{p\leqslant Q}\left(\sideset{}{^*}\sum_{\chi \bmod p^2}\left|\sum_{a\in\mathscr{A}}\chi(a)\right|^2 + \sideset{}{^*}\sum_{\chi \bmod p}\left|\sum_{a\in\mathscr{A}}\chi(a)\right|^2\right)$$

$$\leqslant \sum_{k\leqslant Q^2}\frac{k}{\varphi(k)}\sideset{}{^*}\sum_{\chi \bmod k}\left|\sum_{a\in\mathscr{A}}\chi(a)\right|^2.$$

This sum can be estimated by using Lemma 5.1 (with $c_n = 1$ if $n \in \mathscr{A}$ and $c_n = 0$ if $n \notin \mathscr{A}$). Then we get that

$$S_{\mathscr{A}} \leqslant (N - 1 + Q^4) \sum_{a \in \mathscr{A}} 1 < 2N\,|A|\,, \tag{13}$$

and in the same way we get

$$S_{\mathscr{B}} < 2N\,|B|\,. \tag{14}$$

It follows from (11), (12), (13) and (14) that

$$(2 + o(1))\,|\mathscr{A}|\,|\mathscr{B}|\,\frac{N^{1/4}}{\log N} \leqslant S \leqslant 2N\,(|A|\,|B|)^{1/2}\,,$$

whence

$$(|A|\,|B|)^{1/2} \leqslant (1 + o(1))N^{3/4}\log N,$$

which for N large enough contradicts our indirect assumption (9), and this completes the proof of Theorem 4.2.

Acknowledgements Research partially supported by the Hungarian National Foundation for Scientific Research, Grants No K100291 and NK104183, the French-Hungarian Balaton exchange program FR-33/2009 and the Agence Nationale de la Recherche grant ANR-10-BLAN 0103 MUNUM.

References

1. A. Balog, J. Rivat, A. Sárközy, On arithmetic properties of sumsets. Acta Math. Hung. **144**(1), 18–42 (2014)
2. Y. Bugeaud, On the greatest prime factor of $(ab+1)(bc+1)(ca+1)$. Acta Arith. **86**(1), 45–49 (1998)
3. Y. Bugeaud, F. Luca, A quantitative lower bound for the greatest prime factor of $(ab + 1)(bc + 1)(ca + 1)$. Acta Arith. **114**(3), 275–294 (2004)
4. P. Corvaja, U. Zannier, On the greatest prime factor of $(ab+1)(ac+1)$. Proc. Am. Math. Soc. **131**(6), 1705–1709 (2003)
5. M. Drmota, C. Mauduit, J. Rivat, Primes with an average sum of digits. Compositio **145**(2), 271–292 (2009)
6. M. Drmota, C. Mauduit, J. Rivat, The sum-of-digits function of polynomial sequences. J. Lond. Math. Soc. (2) **84**(1), 81–102 (2011)
7. A. Dujella, There are only finitely many Diophantine quintuples. J. Reine Angew. Math. **566**, 183–214 (2004)
8. P.D.T.A. Elliott, A. Sárközy, The distribution of the number of prime divisors of sums $a + b$. J. Number Theory **29**(1), 94–99 (1988)

9. P.D.T.A. Elliott, A. Sárközy, The distribution of the number of prime divisors of numbers of form $ab + 1$, in *New Trends in Probability and Statistics* (Palanga, 1996), vol. 4 (VSP, Utrecht, 1997), pp. 313–321
10. P. Erdős, A. Sárközy, On divisibility properties of integers of the form $a + a'$. Acta Math. Hung. **50**(1–2), 117–122 (1987)
11. P. Erdős, H. Maier, A. Sárközy, On the distribution of the number of prime factors of sums $a + b$. Trans. Am. Math. Soc. **302**(1), 269–280 (1987)
12. P. Erdős, A. Sárközy, V.T. Sós, On product representations of powers. I. Eur. J. Comb. **16**(6), 567–588 (1995)
13. P.X. Gallagher, The large sieve. Mathematika **14**, 14–20 (1967)
14. A.O. Gelfond, Sur les nombres qui ont des propriétés additives et multiplicatives données. Acta Arith. **13**, 259–265 (1967/1968)
15. K. Gyarmati, On a problem of Diophantus. Acta Arith. **97**(1), 53–65 (2001)
16. K. Gyarmati, A polynomial extension of a problem of Diophantus. Publ. Math. Debr. **66**(3–4), 389–405 (2005)
17. K. Gyarmati, A. Sárközy, C.L. Stewart, On shifted products which are powers. Mathematika **49**(1–2), 227–230 (2002)
18. K. Győry, A. Sárközy, On prime factors of integers of the form $(ab + 1)(bc + 1)(ca + 1)$. Acta Arith. **79**(2), 163–171 (1997)
19. K. Győry, A. Sárközy, C.L. Stewart, On the number of prime factors of integers of the form $ab + 1$. Acta Arith. **74**(4), 365–385 (1996)
20. S. Hernández, F. Luca, On the largest prime factor of $(ab + 1)(ac + 1)(bc + 1)$. Bol. Soc. Mat. Mex. (3) **9**(2), 235–244 (2003)
21. H. Iwaniec, A. Sárközy, On a multiplicative hybrid problem. J. Number Theory **26**, 89–95 (1987)
22. K. Matomäki, On the greatest prime factor of $ab + 1$. Acta Math. Hung. **124**(1–2), 115–123 (2009)
23. C. Mauduit, Multiplicative properties of the Thue-Morse sequence. Period. Math. Hung. **43**(1–2), 137–153 (2001)
24. C. Mauduit, J. Rivat, La somme des chiffres des carrés. Acta Math. **203**, 107–148 (2009)
25. C. Mauduit, J. Rivat, Sur un problème de Gelfond: la somme des chiffres des nombres premiers. Ann. Math. **171**(3), 1591–1646 (2010)
26. H.L. Montgomery, The analytic principle of the large sieve. Bull. Am. Math. Soc. **84**(4), 547–567 (1978)
27. C. Pomerance, A. Sárközy, On products of sequences of integers. Number theory. Vol. I. Elementary and analytic, Proc. Conf., Budapest/Hung. 1987. Colloq. Math. Soc. János Bolyai **51**, 447–463 (1990)
28. J. Rivat, A. Sárközy, A sequences analog of the Piatetski-Shapiro problem. Acta Math. Hung. **74**(3), 245–260 (1997)
29. J. Rivat, A. Sárközy, A Turán-Kubilius type inequality on shifted products. Publ. Math. Debr. **79**(3–4), 637–662 (2011)
30. A. Sárközy, Hybrid problems in number theory, in *Number Theory, New York 1985–88*. Lecture Notes in Mathematics, vol. 1383 (Springer, Berlin, 1989), pp. 146–169
31. A. Sárközy, On sums $a + b$ and numbers of the form $ab + 1$ with many prime factors. Grazer Math. Ber. **318**, 141–154 (1992)
32. A. Sárközy, On the average value for the number of divisors of numbers of form $ab + 1$. Acta Math. Hung. **66**(3), 223–245 (1995)
33. A. Sárközy, C.L. Stewart, On prime factors of integers of the form $ab + 1$. Publ. Math. Debr. **56**(3–4), 559–573 (2000). Dedicated to Professor Kálmán Győry on the occasion of his 60th birthday
34. C.L. Stewart, On the greatest prime factor of integers of the form $ab + 1$. Period. Math. Hung. **43**(1–2), 81–91 (2001)

35. C.L. Stewart, On prime factors of integers which are sums or shifted products, in *Anatomy of Integers*. CRM Proceedings and Lecture Notes, vol. 46 (American Mathematical Society, Providence, 2008), pp. 275–287
36. C.L. Stewart, On sets of integers whose shifted products are powers J. Comb. Theory Ser. A **115**(4), 662–673 (2008)
37. C.L. Stewart, R. Tijdeman, On the greatest prime factor of $(ab + 1)(ac + 1)(bc + 1)$. Acta Arith. **79**(1), 93–101 (1997)
38. G. Tenenbaum, Facteurs premiers de sommes d'entiers. Proc. Am. Math. Soc. **106**, 287–296 (1989)

Narrow Progressions in the Primes

Terence Tao and Tamar Ziegler

To Helmut Maier on his 60th birthday

Abstract In a previous paper of the authors, we showed that for any polynomials $P_1, \dots, P_k \in \mathbb{Z}[\mathbf{m}]$ with $P_1(0) = \dots = P_k(0)$ and any subset A of the primes in $[N] = \{1, \dots, N\}$ of relative density at least $\delta > 0$, one can find a "polynomial progression" $a + P_1(r), \dots, a + P_k(r)$ in A with $0 < |r| \leq N^{o(1)}$, if N is sufficiently large depending on $k, P_1, \dots, P_k$ and δ. In this paper we shorten the size of this progression to $0 < |r| \leq \log^L N$, where L depends on $k, P_1, \dots, P_k$ and δ. In the linear case $P_i = (i-1)\mathbf{m}$, we can take L independent of δ. The main new ingredient is the use of the densification method of Conlon, Fox, and Zhao to avoid having to directly correlate the enveloping sieve with dual functions of unbounded functions.

1 Introduction

1.1 Previous Results

We begin by recalling the well-known theorem of Szemerédi [18] on arithmetic progressions, which we phrase as follows:

Theorem 1.1 (Szemerédi's Theorem). *Let $k \geq 1$ and $\delta > 0$, and suppose that N is sufficiently large depending on k, δ. Then any subset A of $[N] := \{n \in \mathbb{Z} : 1 \leq n \leq N\}$ with cardinality $|A| \geq \delta N$ will contain at least one arithmetic progression $a, a + r, \dots, a + (k-1)r$ of length k, with $r > 0$.*

T. Tao (✉)
UCLA Department of Mathematics, 405 Hilgard Ave, Los Angeles, CA 90095, USA
e-mail: tao@math.ucla.edu

T. Ziegler
Einstein Institute of Mathematics, Edmond J. Safra Campus, Givat Ram, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
e-mail: tamarz@math.huji.ac.il

C. Pomerance, M.Th. Rassias (eds.), *Analytic Number Theory*,
DOI 10.1007/978-3-319-22240-0_22

In fact, by partitioning $[N]$ into intervals of some sufficiently large but constant size $L = L(k, \delta)$ and using the pigeonhole principle, one can ensure that the progression described above is "narrow" in the sense that $r \leq L(k, \delta)$.

In [10], Szemerédi's theorem was relativized to the primes $\mathbb{P} = \{2, 3, 5, 7, \ldots\}$:

Theorem 1.2 (Szemerédi's Theorem in the Primes). *Let $k \geq 2$ and $\delta > 0$, and suppose that N is sufficiently large depending on k, δ. Then any subset A of $[N] \cap \mathbb{P}$ with $|A| \geq \delta |[N] \cap \mathbb{P}|$ will contain at least one arithmetic progression $a, a+r, \ldots, a+(k-1)r$ of length k, with $r > 0$.*

In particular, the primes contain arbitrarily long arithmetic progressions.

The proof of Theorem 1.2 does not place a bound on r beyond the trivial bound $r \leq N$. In contrast with Theorem 1.1, one cannot hope here to make the step size r of the progression as short as $L(k, \delta)$. Indeed, as observed in [1], if N is large enough, then from [8, Theorem 3] one can find a subset A of $[N] \cap \mathbb{P}$ with $|A| \geq \frac{1}{2}|[N] \cap \mathbb{P}|$ (say) such that the gap between any two consecutive elements of A is $\geq c \log N$ for some absolute constant $c > 0$, which of course implies in this case that r must be at least $c \log N$ as well. Indeed, one can improve this lower bound to $\log^{k-1} N$ by a small modification of the argument:

Proposition 1.1. *Let $k \geq 2$, let $\varepsilon > 0$ be sufficiently small depending on k, and suppose that N is sufficiently large depending on k. Then there exists a subset A of $[N] \cap \mathbb{P}$ with $|A| \gg |[N] \cap \mathbb{P}|$ such that A does not contain any arithmetic progression $a, a + r, \ldots, a + (k-1)r$ of length k with $0 < r < \varepsilon \log^{k-1} N$.*

Proof. For any $0 < r < \varepsilon \log^{k-1} N$, let N_r denote the number of $a \in [N]$ such that $a, a + r, \ldots, a + (k-1)r$ are all prime. By the union bound and the prime number theorem, it will suffice to show that

$$\sum_{0 < r < \varepsilon \log^{k-1} N} N_r \leq \frac{1}{2} \frac{N}{\log N}$$

(say), since one can then form A by removing all the elements a in $[N] \cap \mathbb{P}$ that are associated with one of the N_r. But from standard sieve theoretic bounds (see, e.g., [12, Theorem 5.7]) we have

$$N_r \leq C_k \mathfrak{G}(k, r) \frac{N}{\log^k N}$$

where C_k depends only on k and $\mathfrak{G}(k, r)$ is the singular series[1]

$$\mathfrak{G}(k, r) = \prod_p \left(1 - \frac{1}{p}\right)^k \left(1 - \frac{\nu_p(r)}{p}\right) \tag{1}$$

[1] All sums and products over p in this paper are understood to be ranging over primes.

and $\nu_p(r)$ is the number of residue classes $a \in \mathbb{Z}/p\mathbb{Z}$ such that at least one of $a, a + r, \dots, a + (k-1)r$ is equal to 0 mod p, so it suffices to show that

$$\sum_{0<r<M} \mathfrak{G}(k,r) \leq C_k M$$

for any $M \geq 1$ (we allow C_k to represent a different k-dependent constant from line to line, or even within the same line). One could obtain precise asymptotics on the left-hand side using the calculations of Gallagher [7], but we can obtain a crude upper bound that suffices as follows. From (1) we see that

$$\mathfrak{G}(k,r) \leq C_k \exp\left(C_k \sum_{p|r} \frac{1}{p}\right)$$

and hence by [19, Lemma E.1]

$$\mathfrak{G}(k,r) \leq C_k \sum_{p|r} \frac{\log^{C_k} p}{p}$$

and thus

$$\sum_{0<r<M} \mathfrak{G}(k,r) \leq C_k \sum_{p} \frac{\log^{C_k} p}{p} \frac{M}{p} \leq C_k M$$

as required.

In the converse direction, if we use the "Cramér random model" of approximating the primes $\mathbb{P} \cap [N]$ by a random subset of $[N]$ of density $1/\log N$, we can asymptotically almost surely match this lower bound, thanks to the work of Conlon-Gowers [5] and Schacht [17]:

Proposition 1.2. *Let $k \geq 2$ and $\delta, \varepsilon > 0$, let $C > 0$ be sufficiently large depending on δ, k, and suppose that N is sufficiently large depending on k, δ, ε. Let $P \subset [N]$ be chosen randomly, such that each $n \in [N]$ lies in P with an independent probability of $1/\log N$. Then with probability at least $1 - \varepsilon$, every subset A of P with $|A| \geq \delta |P|$ will contain an arithmetic progression $a, a + r, \dots, a + (k-1)r$ of length k with $0 < r \leq C \log^{k-1} N$.*

We remark that a modification of the argument in Proposition 1.1 shows that we cannot replace the large constant C here by an arbitrarily small constant $c > 0$.

Proof. We partition $[N]$ into intervals $I_1, \dots, I_m$ of length between $\frac{C}{2} \log^{k-1} N$ and $C \log^{k-1} N$, thus $m \leq \frac{2N}{C \log^{k-1} N}$. For each interval I_i, we see from [17, Theorem 2.2] or [5, Theorem 1.12] that with probability at least $1 - \frac{\delta\varepsilon}{10}$, every subset A_i of $P \cap I_i$

with $|A_i| \geq \frac{\delta}{2}|P \cap I_i|$ will contain an arithmetic progression of length at least k. Call an interval I_i *bad* if this property does not hold, thus each I_i is bad with probability at most $\delta\varepsilon/10$. By linearity of expectation followed by Markov's inequality, we conclude that with probability at least $1 - \varepsilon$, at most $\frac{\delta N}{5C \log^{k-1} N}$ of the I_i are bad. Then if $A \subset P$ is such that $|A| \geq \delta|P|$, then at most $\frac{\delta}{2}|P|$ of the elements of A are contained in bad intervals, so from the pigeonhole principle there exists a good interval I_i such that $|A \cap I_i| \geq \frac{\delta}{2}|P \cap I_i|$, and the claim follows.

It is thus natural to conjecture that in Theorem 1.2 one can take r to be as small as $\log^{k-1+o(1)} N$.

Remark 1.1. If one seeks progressions inside the full set $\mathbb{P}$ of primes, rather than of dense subsets of the primes, then the Hardy–Littlewood prime tuples conjecture [13] predicts that one can take r to be of size $O_k(1)$; indeed, one should be able to take r to be the product of all the primes less than or equal to k. In the case $k = 2$, the claim that one can take $r = O(1)$ amounts to finding infinitely many bounded gaps between primes, a claim that was only recently established by Zhang [20]. For higher k, the claim $r = O_k(1)$ appears to currently be out of reach of known methods; the best known result in this direction, due to Maynard [15] (and also independently in unpublished work of the first author), shows that for any sufficiently large $R > 1$, there exist infinitely many intervals of natural numbers of length R that contain $\geq c \log R$ primes for some absolute constant $c > 0$, but this is too sparse a set of primes to expect to find length k progressions for any $k \geq 3$.

We now consider generalizations of the above results, in which arithmetic progressions $a, a + r, \ldots, a + (k-1)r$ are replaced by "polynomial progressions" $a + P_1(r), \ldots, a + P_k(r)$. More precisely, let $\mathbb{Z}[\mathbf{m}]$ denote the ring of polynomials of one indeterminate variable $\mathbf{m}$ with integer coefficients. Then Bergelson and Leibman [3] established the following polynomial version of Theorem 1.1:

Theorem 1.3 (Polynomial Szemerédi's Theorem). *Let $k \geq 1$, let $P_1, \ldots, P_k \in \mathbb{Z}[\mathbf{m}]$ be such that $P_1(0) = \ldots = P_k(0)$, let $\delta > 0$, and suppose that N is sufficiently large depending on $k, P_1, \ldots, P_k, \delta$. Then any subset A of $[N]$ with cardinality $|A| \geq \delta N$ will contain at least one polynomial progression $a + P_1(r), a + P_2(r), \ldots, a + P_k(r)$ with $r > 0$.*

Of course, Theorem 1.1 is the special case of Theorem 1.3 when $P_i = (i-1)\mathbf{m}$. As with Theorem 1.1, a partitioning argument shows that we may take $r \leq L(k, P_1, \ldots, P_k, \delta)$ for some quantity L depending on the indicated parameters. The polynomial analogue of Theorem 1.2 was established by the authors in [19]:

Theorem 1.4 (Polynomial Szemerédi's Theorem in the Primes). *Let $k \geq 2$, let $P_1, \ldots, P_k \in \mathbb{Z}[\mathbf{m}]$ be such that $P_1(0) = \ldots = P_k(0)$, $\varepsilon, \delta > 0$, and suppose that N is sufficiently large depending on $k, P_1, \ldots, P_k, \delta, \varepsilon$. Then any subset A of $[N] \cap \mathbb{P}$ with $|A| \geq \delta|[N] \cap \mathbb{P}|$ will contain at least one polynomial progression $a + P_1(r), a + P_2(r), \ldots, a + P_k(r)$ with $0 < r < N^{\varepsilon}$.*

In particular, this implies Theorem 1.2 with a bound $0 < r \leq N^{o(1)}$.

Remark 1.2. The condition $P_1(0) = \ldots = P_k(0)$ in Theorem 1.3 can be relaxed to the property of *intersectivity* (that $P_1, \ldots, P_k$ have a common root in the profinite integers $\hat{\mathbb{Z}} = \lim_{\infty \leftarrow m} \mathbb{Z}/m\mathbb{Z}$); see [4]. It is not yet known if Theorem 1.4 can similarly be relaxed to intersective polynomials, except in the $k = 2$ case which was established in [14]. We will not pursue this matter here.

1.2 Main New Result

Our main result is to improve the bound on the r parameter in Theorem 1.4 to be polylogarithmic in size:

Theorem 1.5 (Short Polynomial Progressions in the Primes). *Let $k \geq 2$, let $P_1, \ldots, P_k \in \mathbb{Z}[\mathbf{m}]$ be such that $P_1(0) = \ldots = P_k(0)$, $\varepsilon, \delta > 0$, and suppose that N is sufficiently large depending on $k, P_1, \ldots, P_k, \delta, \varepsilon$. Then any subset A of $[N] \cap \mathbb{P}$ with $|A| \geq \delta |[N] \cap \mathbb{P}|$ will contain at least one polynomial progression $a + P_1(r), a + P_2(r), \ldots, a + P_k(r)$ with $0 < r < \log^L N$, where L depends only on $k, P_1, \ldots, P_k, \delta$.*

In particular, there are infinitely polynomial progressions $a + P_1(r), a + P_2(r), \ldots, a + P_k(r)$ consisting entirely of primes with $0 < r \ll \log^L a$, with L now depending only on $k, P_1, \ldots, P_k$. This is new even in the case of arithmetic progressions $a, a + r, \ldots, a + (k-1)r$.

A modification of the proof of Proposition 1.1 shows that some power of $\log N$ is needed in the upper bound on r in Theorem 1.5. However, we do not know what the optimal value of L is; our argument for general $P_1, \ldots, P_k$ uses the PET induction method [2], and as such L will grow rapidly with the degrees of the $P_1, \ldots, P_k$. The dependence of L on δ occurs for technical reasons, and we conjecture that one can in fact select L to be independent of δ; we can verify this conjecture in the arithmetic progression case $P_i = (i-1)\mathbf{m}$, and in fact we can take the explicit value $L := Ck2^k$ in this case for some fixed constant C (actually $C = 3$ would already suffice). We discuss this explicit variant of Theorem 1.5 in Sect. 7. Propositions 1.1, 1.2 suggest that we should in fact be able to set $L = k - 1 + \varepsilon$ in these cases, although our methods do not seem strong enough to achieve this, even in the $k = 3$ case.

It is possible that one might be able to directly modify the arguments in [19] (which are in turn based on those in [10]) to establish Theorem 1.5; the main technical difficulties in doing so are the need of having to correlate the enveloping sieve with dual functions of unbounded functions and verifying the "correlation condition" required in that argument, in the setting when r is as small as $\log^L N$ (and the argument appears to have no chance of working when L is independent of δ). On the other hand, the need for the analogous correlation and boundedness conditions

in [10] to prove Theorem 1.2 were recently removed[2] by Conlon et al. [6], using a new method which they refer to as "densification". We will be able to combine the densification method with the arguments in [19] to establish Theorem 1.5. As a consequence, we also obtain a slightly different proof of Theorem 1.4 than the one in [19], in which the (rather complicated) verification of the correlation condition is no longer necessary, but the densification arguments of Conlon, Fox, and Zhao are inserted instead.

2 Preliminary Reductions

We now begin the proof of Theorem 1.5. We use the following asymptotic notation. We let N' be an asymptotic parameter tending to infinity along some sequence $N' = N'_j$ of natural numbers. All mathematical objects in this paper are implicitly permitted to depend on N', unless they are explicitly designated to be *fixed*, in which case they are independent of N'. We use $X = O(Y)$, $X \ll Y$, or $Y \gg X$ to denote the estimate $|X| \leq CY$ for some fixed C, and $X = o(Y)$ to denote the estimate $|X| \leq c(N')Y$ where $c(N') \to 0$ as $N' \to \infty$. Our statements will be implicitly restricted to the regime in which N' is sufficiently large depending on all fixed parameters.

Suppose for the sake of contradiction that Theorem 1.5 failed. Carefully negating the quantifiers (and relabeling N as N', for reasons that will be clearer later), we conclude that we may find a fixed $k \geq 2$ and fixed polynomials $P_1, \dots, P_k \in \mathbb{Z}[\mathbf{m}]$ with $P_1(0) = \dots = P_k(0)$, fixed $\delta > 0$, a sequence $N' = N'_j$ of natural numbers going to infinity, and a set $A = A_{N'} \subset [N', 2N'] \cap \mathbb{P}$ such that

$$|A| \geq \delta |[N'] \cap \mathbb{P}|$$

and such that for any fixed $L > 0$, there are no polynomial progressions $a + P_1(r), a + P_2(r), \dots, a + P_k(r)$ in A with $0 < r < \log^L N'$ (recall that we assume N' sufficiently large depending on fixed quantities such as L).

The first few reductions are essentially the same to those in [19]. We begin with the "W-trick" from [10] to eliminate local irregularities modulo small primes. We let $w = w_{N'}$ be a sufficiently slowly growing function of N'; for instance, we could take

[2]The "dual function condition" that the dual function of the enveloping sieve is bounded already failed in the arguments in [19], which was a significant cause of the complexity of that paper due to the need to find substitutes for this condition (in particular, the correlation condition became significantly more difficult to even state, let alone prove). But the arguments in [6] do not require any version of the dual function condition at all, leading to some simplifications in the current argument over those in [19].

$w := \frac{1}{10} \log\log\log N'$ as in [19] for the sake of concreteness, although the precise value of w is unimportant. We then define the quantity

$$W := \prod_{p<w} p$$

and the natural number

$$N := \left\lfloor \frac{N'}{W} \right\rfloor.$$

From the prime number theorem[3] we have

$$|A| \gg \frac{NW}{\log N}$$

and all elements of A larger than $\sqrt{N}$ (say) are coprime to W. Thus, by the pigeonhole principle, we may find $b \in [W]$ coprime to W (and depending on N' of course) such that

$$|\{n \in [N] : nW + b \in A\}| \gg \frac{NW}{\phi(W)\log N},$$

where ϕ denotes the Euler totient function.

The domain $[N]$ is not quite translation invariant. In order to eliminate this (minor) difficulty, we will follow [19] and work instead in the cyclic group $X := \mathbb{Z}/N\mathbb{Z}$ using the obvious embedding $\iota : [N] \to X$. We give this space the uniform Haar probability measure, thus

$$\int_X f := \frac{1}{N} \sum_{x \in X} f(x)$$

for any $f : X \to \mathbb{R}$. We also define shift maps $T^h f : X \to \mathbb{R}$ for any $h \in \mathbb{Z}$ by

$$T^h f(x) := f(x + h);$$

clearly we have the identities

$$\begin{aligned} T^h T^k f &= T^{h+k} f \\ T^h(fg) &= (T^h f)(T^h g) \\ \int_X T^h f &= \int_X f \end{aligned}$$

for any $h, k \in \mathbb{Z}$ and $f, g : X \to \mathbb{R}$. We will use these identities frequently without further comment in the sequel.

[3] Actually, the weaker lower bound $\pi(x) \gg \frac{x}{\log x}$ of Chebyshev would suffice here.

We will need a fixed quantity $\epsilon_0 > 0$ (depending only on $k, P_1, \ldots, P_k$) to be chosen later. We define the function $f : X \to \mathbb{R}$ by the formula

$$f(\iota(n)) = \frac{\epsilon_0}{10} \frac{\phi(W) \log N}{W} 1_{[N-\sqrt{N}]\setminus[\sqrt{N}]}(n) 1_A(nW + b) \tag{2}$$

for $n \in [N]$, where 1_A denotes the indicator function of A; the reason for the normalizing factor $\frac{\epsilon_0}{10}$ is so that f can be dominated by an enveloping sieve ν, to be defined later. We then have

$$\int_X f \gg 1, \tag{3}$$

where we allow the implied constants here to depend on the fixed quantities δ and ϵ_0.

Now let $L > 0$ be a sufficiently large fixed quantity (depending on $k, P_1, \ldots, P_k, \delta$) to be chosen later. We will need the "coarse scale"

$$M := \log^L N \tag{4}$$

which will basically be the domain of interest for the polynomials $P_1, \ldots, P_k$, and the "fine scale"

$$H := \log^{\sqrt{L}} N,$$

which will basically be the scale used for various applications of the van der Corput inequality. Note in particular that any quantity of size $O(H^{O(1)}\sqrt{M})$ will also be $o(M)$ if the implied constants do not depend on L, and L is large enough.

By hypothesis, if N (or N') is large enough, then the set A contains no polynomial progressions of the form $a + P_1(Wm), \ldots, a + P_k(Wm)$ with $m \in [M]$. In particular, we see that

$$\Lambda(f, \ldots, f) = 0 \tag{5}$$

where Λ is the k-linear form

$$\Lambda(f_1, \ldots, f_k) := \mathbb{E}_{m\in[M]} \int_X T^{P_1(Wm)/W} f_1 \ldots T^{P_k(Wm)/W} f_k \tag{6}$$

for k functions $f_1, \ldots, f_k : X \to \mathbb{R}$, and we use the averaging notation

$$E_{a\in A} f(a) := \frac{1}{|A|} \sum_{a\in A} f(a).$$

Note that there are no "wraparound" issues caused by the embedding into $\mathbb{Z}/N\mathbb{Z}$, due to our removal in (2) of the elements $n \in [N]$ that are less than $\sqrt{N}$ or larger than $N - \sqrt{N}$.

On the other hand, using the (multi-dimensional) polynomial Szemerédi theorem of Bergelson and Leibman [3], we have the following quantitative version of Theorem 1.3:

Theorem 2.6 (Quantitative Polynomial Szemerédi Theorem). *Let $\delta > 0$ be fixed, and let $g : X \to \mathbb{R}$ obey the pointwise bounds*

$$0 \leq g \leq 1$$

(i.e., $0 \leq g(x) \leq 1$ for all $x \in X$), as well as the integral lower bound $\int_X g \geq \delta - o(1)$. Then we have

$$\Lambda(g, \ldots, g) \geq c(\delta) - o(1)$$

where $c(\delta) > 0$ depends only on δ, k, and $P_1, \ldots, P_k$.

Proof. See [19, Theorem 3.2]. In that theorem, the common value $P_1(0) = \ldots = P_k(0)$ of the P_i was assumed to be zero, but the general case follows from this case by a simple change of variables.

This theorem cannot be directly applied to control $\Lambda(f, \ldots, f)$, because f is not uniformly bounded. However, Theorem 1.5 will now be a consequence of the following claim.

Theorem 2.7 (Approximation by Bounded Function). *Suppose that $\epsilon_0 > 0$ is sufficiently small (depending on $k, P_1, \ldots, P_k$). Let $\varepsilon > 0$ be fixed. Suppose that L is a fixed quantity which is sufficiently large depending on $k, P_1, \ldots, P_k, \varepsilon, \epsilon_0$. Let f be as in (2). Then there exists $g : X \to \mathbb{R}$ with the pointwise bounds*

$$0 \leq g \leq 1,$$

such that

$$\left| \int_X f - \int_X g \right| \ll \varepsilon + o(1) \tag{7}$$

and

$$|\Lambda(f, \ldots, f) - \Lambda(g, \ldots, g)| \ll \varepsilon + o(1), \tag{8}$$

where the implied constants in the $\ll$ notation do not depend on ε or L.

Let us assume Theorem 2.7 for now, and see how it implies Theorem 1.5. Let $\epsilon_0 > 0$ be small enough for Theorem 2.7 to apply, $\varepsilon > 0$ be a sufficiently small

fixed quantity (depending on δ, ϵ_0) to be chosen later, and let L be large enough depending on $k, P_1, \dots, P_k, \varepsilon, \epsilon_0$ (in particular, L will depend on δ). Let g be as in Theorem 2.7, let M be defined by (4), and Λ defined by (6) (in particular, Λ depends on L). From (3), (7), and the triangle inequality we have

$$\int_X g \gg 1$$

if ε is small enough depending on ϵ_0, δ. Applying Theorem 2.6, we have

$$\Lambda(g, \dots, g) \gg c_0 - o(1)$$

for some $c_0 > 0$ depending on ϵ_0, δ but not on ε, L. By (8) and the triangle inequality, we thus have

$$\Lambda(f, \dots, f) > 0$$

if ε is small enough (depending on ϵ_0, δ) and N is large enough, contradicting (5); and Theorem 1.5 follows.

It remains to establish Theorem 2.7. This is the focus of the remaining sections of the paper.

Remark 2.3. The above arguments show that if one could remove the dependence of L on ε in Theorem 2.7, then one could also remove the dependence of L on δ in Theorem 1.5.

3 The Enveloping Sieve

As in [10, 19], we view f as a "positive density fraction" of a well-controlled *enveloping sieve* ν, which is defined by the explicit formula

$$\nu(\iota(n)) := \frac{\phi(W)\log R}{W}\left(\sum_{d|Wn+b} \mu(d)\chi\left(\frac{\log d}{\log R}\right)\right)^2$$

for $n \in [N]$, where

$$R := N^{\epsilon_0}, \tag{9}$$

μ is the Möbius function, and $\chi : \mathbb{R} \to \mathbb{R}$ is a fixed smooth even function supported on $[-1, 1]$ with the normalization

$$\int_0^1 |\chi'(t)|^2\, dt = 1$$

(where χ' is the derivative of χ) and such that $\chi(0) \geq 1/2$ (say). By construction and (2), we have the pointwise bound

$$0 \leq f \leq \nu, \tag{10}$$

that is to say that $0 \leq f(n) \leq \nu(n)$ for all $n \in X$.

From [19, Corollary 10.5] we have the crude bound

$$\int_X \prod_{j=1}^{J} T^{h_j} \nu = 1 + o(1) + O\left(\mathrm{Exp}\left(O\left(\sum_{1 \leq j < j' \leq J} \sum_{w \leq p \leq R^{\log R} : p | h_j - h_{j'}} \frac{1}{p}\right)\right)\right) \tag{11}$$

for any fixed J and any integers $h_1, \ldots, h_J = O(\sqrt{N})$ (not necessarily distinct), assuming that ϵ_0 is sufficiently small depending on J, and where $\mathrm{Exp}(x) := e^x - 1$, where the implied constant in the $O()$ exponent only depends on J. Here, of course, we use $p|n$ to denote the assertion that p divides n. In particular, we have the mean bound

$$\int_X \nu = 1 + o(1) \tag{12}$$

and the crude bound

$$\int_X T^{h_1} \nu \ldots T^{h_J} \nu \ll \log^{O(1)} N \tag{13}$$

for any fixed J and any integers $h_1, \ldots, h_J = O(\sqrt{N})$ (not necessarily distinct), assuming ϵ_0 is sufficiently small depending on J, and where the implied constant in the $O(1)$ exponent depends only on J.

One can use (11) to establish the following fundamental pseudorandomness property:

Proposition 3.3 (Polynomial Forms Condition). *Let $J, d, D \geq 1$ be fixed natural numbers. Suppose that ϵ_0 is sufficiently small depending on J, d, D, and that L is sufficiently large depending on J, d, D. Then for any polynomials $Q_1, \ldots, Q_J \in \mathbb{Z}[\mathbf{m}_1, \ldots, \mathbf{m}_d]$ of degree at most D, with coefficients of size $O(W^{O(1)})$, and with $Q_j - Q_{j'}$ non-constant for every $1 \leq j < j' \leq d$, and any convex body $\Omega \in [-M^2, M^2]^d$ of inradius[4] at least $H^{1/2}$, we have the asymptotic*

$$\mathbb{E}_{\mathbf{h} \in \Omega \cap \mathbb{Z}^d} \int_X \prod_{j=1}^{J} T^{Q_j(\mathbf{h})} \nu = 1 + o(1). \tag{14}$$

[4]The *inradius* of a convex body is the radius of the largest open ball one can inscribe inside the body.

This proposition was established in [19, Theorem 3.18] in the case when M is a small power of N; the point is that M can be lowered to essentially $\log^L N$ for some large L. However, note that the number J of polynomials involved cannot be arbitrarily large depending on ϵ_0. The main obstruction to reducing the size of the coarse scale M is that we need the "diagonal" contributions to (14) (such as those coming from the terms where one of the $Q_j(\mathbf{h})$ vanishes) to be negligible when compared to the remaining terms. The sieve ν (or powers thereof, such as ν^2) tends to have size $\log^{O(1)} N$ on the average, and using this one expects to control diagonal contributions to (14) by something like $\log^{O(1)} N/M$, which will be negligible when M is a sufficiently large power of $\log N$.

Proof. We repeat the arguments from [19, §11]. For each $\mathbf{h}$, we see from (11) that

$$\int_X \prod_{j=1}^{J} T^{Q_j(\mathbf{h})}\nu = 1 + o(1) + O\left(\mathrm{Exp}\left(O\left(\sum_{1\le j<j'\le J}\ \sum_{w\le p\le R^{\log R}: p|Q_j(\mathbf{h})-Q_{j'}(\mathbf{h})} \frac{1}{p}\right)\right)\right).$$

Thus it suffices to show that

$$\mathbb{E}_{\mathbf{h}\in\Omega\cap\mathbb{Z}^d}\,\mathrm{Exp}\left(O\left(\sum_{1\le j<j'\le J}\ \sum_{w\le p\le R^{\log R}: p|Q_j(\mathbf{h})-Q_{j'}(\mathbf{h})} \frac{1}{p}\right)\right) = o(1).$$

Using the elementary bound $\mathrm{Exp}(a+b) \ll \mathrm{Exp}(2a)+\mathrm{Exp}(2b)$ repeatedly, it suffices to show that

$$\mathbb{E}_{\mathbf{h}\in\Omega\cap\mathbb{Z}^d}\,\mathrm{Exp}\left(O\left(\sum_{w\le p\le R^{\log R}: p|Q_j(\mathbf{h})-Q_{j'}(\mathbf{h})} \frac{1}{p}\right)\right) = o(1).$$

for each $1 \le j < j \le d'$.

We first dispose of the "globally bad" primes, in which p divides the entire polynomial $Q_j - Q_{j'}$. As $Q_j - Q_{j'}$ is non-constant and has coefficients $O(W^{O(1)})$, we see that the product of all such primes is $O(W^{O(1)})$. In [19, Lemma E.3], it is shown that

$$\sum_{p\ge w: p|n} \frac{1}{p} = o(1)$$

for any $n = O(W^{O(1)})$. Thus the contribution of such primes in the above sum is negligible.

In [19, Lemma E.1], it is shown that

$$\mathrm{Exp}\left(O\left(\sum_{p\in A}\frac{1}{p}\right)\right) \ll \sum_{p\in A}\frac{\log^{O(1)} p}{p}$$

for any set A of primes. Thus it suffices to show that

$$\sum_{w\le p\le R^{\log R}} \frac{\log^{O(1)} p}{p} \mathbb{E}_{\mathbf{h}\in\Omega\cap\mathbb{Z}^d} 1_{p|Q_j(\mathbf{h})-Q_{j'}(\mathbf{h});p|Q_j-Q_{j'}} = o(1).$$

From [19, Lemma D.3], we conclude that if p does not divide $Q_j - Q_{j'}$, then the average $1_{p|Q_j(\mathbf{h})-Q_{j'}(\mathbf{h})}$ on any cube in $\mathbb{Z}^d$ of sidelength $1 \le K \le p$ is $O(\frac{1}{K})$. By [19, Corollary C.2], [19, Lemma C.4] and the inradius hypothesis we then have

$$\mathbb{E}_{\mathbf{h}\in\Omega\cap\mathbb{Z}^d} 1_{p|h_j-h_{j'}} \ll \frac{1}{p} + H^{-1/2}.$$

Thus we reduce to showing that

$$\sum_{w\le p\le R^{\log R}} \frac{\log^{O(1)} p}{p^2} + H^{-1/2}\frac{\log^{O(1)} p}{p} = o(1),$$

but this follows easily from Mertens' theorem (or the prime number theorem) and the definition of M and w, if L is large enough.

4 Averaged Local Gowers Norms

As in [19], we will control the left-hand side of (8) using some local Gowers norms, which we now define. Given any $d \ge 2$ and integers $a_1, \dots, a_d$ and any scale $S \ge 1$, we define the *local Gowers uniformity norms* $U_S^{a_1,\dots,a_d}$ by the formula

$$\|f\|_{U_S^{a_1,\dots,a_d}}^{2^d} := \mathbb{E}_{m_1^{(0)},\dots,m_d^{(0)},m_1^{(1)},\dots,m_d^{(1)}\in[S]} \int_X \prod_{(\omega_1,\dots,\omega_d)\in\{0,1\}^d} T^{m_1^{(\omega_1)}a_1+\dots+m_d^{(\omega_d)}a_d} f \tag{15}$$

for $f : X \to \mathbb{R}$. Next, for any $t \ge 2$ and any d-tuple $\mathbf{Q} = (Q_1, \dots, Q_d)$ of polynomials $Q_i \in \mathbb{Z}[\mathbf{h}_1, \dots, \mathbf{h}_t, \mathbf{W}]$ in $t+1$ variables, we define the *averaged local Gowers uniformity norms* $U_S^{\mathbf{Q}([H]^t,W)}$ on functions $f : X \to \mathbb{R}$ by the formula

$$\|f\|_{U_S^{\mathbf{Q}([H]^t,W)}}^{2^d} := \mathbb{E}_{\mathbf{h}\in[H]^t} \|f\|_{U_S^{Q_1(\mathbf{h},W),\dots,Q_d(\mathbf{h},W)}}^{2^d}. \tag{16}$$

These are indeed norms; see [19, Appendix A]. One can extend these norms to complex-valued functions by inserting an alternating sequence of conjugation symbols in the product, but we will not need to use such an extension here. We remark that these expressions may also be defined for $d = 1$, but are merely seminorms instead of norms in that case.

From the Gowers–Cauchy–Schwarz inequality (see, e.g., [11, Appendix B]) and Hölder's inequality, we record the useful inequality

$$\left|\mathbb{E}_{m_1^{(0)},\dots,m_d^{(0)},m_1^{(1)},\dots,m_d^{(1)}\in[S]}\int_X \prod_{\omega\in\{0,1\}^d} T^{m_1^{(\omega_1)}a_1+\dots+m_d^{(\omega_d)}a_d} f_\omega\right| \leq \prod_{\omega\in\{0,1\}^d} \|f_\omega\|_{U_S^{a_1,\dots,a_d}}^{2^d} \tag{17}$$

for any functions $f_\omega : X \to \mathbb{R}$ for $\omega \in \{0,1\}^d$ where we write $\omega := (\omega_1,\dots,\omega_d)$.

In a similar spirit, we have the following basic inequality:

Theorem 4.8 (Polynomial Generalized von Neumann Theorem). *Suppose that $\epsilon_0 > 0$ is a fixed quantity which is sufficiently small depending on $k, P_1,\dots,P_k$, and that L is a fixed quantity which is sufficiently large depending on $k, P_1,\dots,P_k$. Then there exists fixed $t \geq 0, d \geq 2$ and a fixed d-tuple $\mathbf{Q} = (Q_1,\dots,Q_d)$ of polynomials $Q_i \in \mathbb{Z}[\mathbf{h}_1,\dots,\mathbf{h}_t,\mathbf{W}]$, none of which are identically zero, and which are independent of ϵ_0, L, such that one has the inequality*

$$|\Lambda(g_1,\dots,g_k)| \ll \min_{1\leq i\leq k} \|g_i\|_{U^{\mathbf{Q}([H]^t,W)}_{\sqrt{M}}}^c + o(1)$$

for some fixed $c > 0$ and all $g_1,\dots,g_k : X \to \mathbb{R}$ obeying the pointwise bound $|g_i| \leq \nu + 1$ for all $i = 1,\dots,d$.

Proof. This is essentially [19, Theorem 4.5] (which was proven by a combination of PET induction, the Cauchy–Schwarz inequality, and the polynomial forms condition), with the only difference being that H and M are now polylogarithmic in N, rather than polynomial in N. However, an inspection of the proof of [19, Theorem 4.5] shows that this does not affect the arguments (after replacing the polynomial forms condition used there with Proposition 3.3, of course); the key relationships between H, M, N that are used in the proof are that $(HW)^{O(1)}\sqrt{M} = o(M)$ and that $(HWM)^{O(1)} = o(\sqrt{N})$, where the implied constants depend only on $k, P_1,\dots,P_k$ (and in particular are independent of ϵ_0, L), and these properties are certainly obeyed for the choice of H and M used here. (The bound (13) is sufficient to deal with all the error terms arising from use of the van der Corput inequality in this regime.)

Setting $g_2 = \dots = g_k = 1$, we obtain in particular that

$$\left|\int_X g_1\right| \ll \|g_1\|_{U^{\mathbf{Q}([H]^t,W)}_{\sqrt{M}}}^c + o(1).$$

(In fact, one can take $c = 1$ in this inequality by the standard monotonicity properties of the Gowers norms, see [19, Lemma A.3], but we will not need this improvement here.)

In view of this theorem, Theorem 2.7 is now a consequence of the following claim (after replacing ε with ε^c):

Theorem 4.9 (Approximation by Bounded Function, Again). *Let $d \geq 2$ and $t \geq 0$ be fixed, and let $\mathbf{Q} = (Q_1, \dots, Q_d)$ be a fixed d-tuple of polynomials $Q_i \in \mathbb{Z}[\mathbf{h}_1, \dots, \mathbf{h}_t, \mathbf{W}]$, not identically zero. Let $\epsilon_0 > 0$ be a fixed quantity that is sufficiently small depending on $d, t, \mathbf{Q}$. Let $\varepsilon > 0$ be fixed, and let L be a fixed quantity that is sufficiently large depending on $d, t, \mathbf{Q}, \epsilon_0, \varepsilon$. Let ν be as above, and let $f : X \to \mathbb{R}$ obey the pointwise bound*

$$0 \leq f \leq \nu.$$

Then there exists $g : X \to \mathbb{R}$ with the pointwise bound

$$0 \leq g \leq 1,$$

such that

$$\|f - g\|_{U^{\mathbf{Q}([H]^t, W)}_{\sqrt{M}}} \ll \varepsilon + o(1). \tag{18}$$

It remains to establish Theorem 4.9. This is the objective of the remaining sections of the paper.

Remark 4.4. As before, if one could remove the dependence of L on ε in Theorem 4.9, then one could remove the dependence of L on δ in Theorem 1.5. Also, from this point on the number k of polynomials $P_1, \dots, P_k$ in Theorem 1.5 no longer plays a role, and we will use the symbol k to denote other (unrelated) natural numbers.

5 The Dense Model Theorem

Let $d, t, \mathbf{Q}$ be as in Theorem 4.9. The averaged local Gowers norm $\|f\|_{U^{\mathbf{Q}([H]^t, W)}_{\sqrt{M}}}$ of a function $f : X \to \mathbb{R}$ can then be expressed by the identity

$$\|f\|^{2^d}_{U^{\mathbf{Q}([H]^t, W)}_{\sqrt{M}}} = \int f \mathscr{D} f \tag{19}$$

where the *dual function* $\mathscr{D} f = \mathscr{D}^{\mathbf{Q}([H]^t, W)}_{\sqrt{M}} f : X \to \mathbb{R}$ is defined by the formula

$$\begin{aligned} \mathscr{D} f := \mathbb{E}_{\mathbf{h} \in [H]^t} \mathbb{E}_{m_1^{(0)}, \dots, m_d^{(0)}, m_1^{(1)}, \dots, m_d^{(1)} \in [\sqrt{M}]} \\ \prod_{(\omega_1, \dots, \omega_d) \in \{0,1\}^d \setminus \{0\}^d} T^{\sum_{i=1}^d (m_i^{(\omega_i)} - m_i^{(0)}) Q_i(\mathbf{h})} f \end{aligned} \tag{20}$$

More generally, we define

$$\mathscr{D}(f_\omega)_{\omega\in\{0,1\}^d\backslash\{0\}^d} := \mathbb{E}_{\mathbf{h}\in[H]^t}\mathbb{E}_{m_1^{(0)},\dots,m_d^{(0)},m_1^{(1)},\dots,m_d^{(1)}\in[\sqrt{M}]} \prod_{(\omega_1,\dots,\omega_d)\in\{0,1\}^d\backslash\{0\}^d} T^{\sum_{i=1}^d (m_i^{(\omega_i)}-m_i^{(0)})Q_i(\mathbf{h})} f_\omega \tag{21}$$

for any tuple of functions $f_\omega : X \to \mathbb{R}$ for $\omega \in \{0,1\}^d\backslash\{0\}^d$.

Theorem 4.9 is then an immediate consequence of combining the following two theorems (with the function f appearing in Theorem 5.11 being replaced by the function $f - g$ appearing in Theorem 5.10).

Theorem 5.10 (Weak Approximation by Bounded Function). *Let $d \geq 2$ and $t \geq 0$ be fixed, and let $\mathbf{Q} = (Q_1,\dots,Q_d)$ be a fixed d-tuple of polynomials $Q_i \in \mathbb{Z}[\mathbf{h}_1,\dots,\mathbf{h}_t,\mathbf{W}]$, not identically zero. Let $\epsilon_0 > 0$ be a fixed quantity that is sufficiently small depending on $d,t,\mathbf{Q}$. Let $\varepsilon > 0$ be fixed, and let L be a fixed quantity that is sufficiently large depending on $d,t,\mathbf{Q},\epsilon_0,\varepsilon$. Let ν be as above, and let $f : X \to \mathbb{R}$ obey the pointwise bound*

$$0 \leq f \leq \nu.$$

Then there exists $g : X \to \mathbb{R}$ with the pointwise bound

$$0 \leq g \leq 1,$$

such that

$$\left|\int_X (f-g)\mathscr{D}(F_\omega)_{\omega\in\{0,1\}^d\backslash\{0\}^d}\right| \ll \varepsilon + o(1) \tag{22}$$

for all functions $F_\omega : X \to \mathbb{R}$ for $\omega \in \{0,1\}^d\backslash\{0\}^d$ with the pointwise bounds $-1 \leq F_\omega \leq 1$.

Theorem 5.11 (Densification). *Let $d \geq 2$ and $t \geq 0$ be fixed, and let $\mathbf{Q} = (Q_1,\dots,Q_d)$ be a fixed d-tuple of polynomials $Q_i \in \mathbb{Z}[\mathbf{h}_1,\dots,\mathbf{h}_t,\mathbf{W}]$, not identically zero. Let $\epsilon_0 > 0$ be a fixed quantity that is sufficiently small depending on $d,t,\mathbf{Q}$. Let $\varepsilon > 0$ be fixed, and let L be a fixed quantity that is sufficiently large depending on $d,t,\mathbf{Q},\epsilon_0$. Let ν be as above, and let $f : X \to \mathbb{R}$ obey the pointwise bound*

$$|f| \leq \nu + 1,$$

and suppose that

$$\left|\int_X f\mathscr{D}(F_\omega)_{\omega\in\{0,1\}^d\backslash\{0\}^d}\right| \ll \varepsilon + o(1) \tag{23}$$

for all functions $F_\omega : X \to \mathbb{R}$ *with the pointwise bounds* $-1 \le F_\omega \le 1$. *Then we have*

$$\|f\|_{U^{\mathbf{Q}([H]^t,W)}_{\sqrt{M}}} \ll \varepsilon^c + o(1)$$

for some fixed $c > 0$ *(independent of* ε*).*

We prove Theorem 5.10 in this section, and Theorem 5.11 in the next section.

To prove Theorem 5.10, we invoke the dense model theorem, first established implicitly in [10] and then made more explicit in [9, 16, 19]. We use the formulation from [16, Theorem 1.1]:

Theorem 5.12 (Dense Model Theorem). *For every* $\varepsilon > 0$, *there is* $K = (1/\varepsilon)^{O(1)}$ *and* $\varepsilon' > 0$ *such that, whenever* $\mathscr{F}$ *is a set of bounded functions from* X *to* $[-1, 1]$, *and* $\nu : X \to \mathbb{R}^+$ *obeys the bound*

$$|\int_X (\nu - 1)F_1 \dots F_{K'}| \le \varepsilon'$$

for all $0 \le K' \le K$ *and* $F_1, \dots, F_{K'} \in \mathscr{F}$, *and every function* $f : X \to \mathbb{R}$ *with* $0 \le f \le \nu$, *one has a function* $g : X \to [0, 1]$ *such that*

$$|\int_X (f - g)F| \le \varepsilon$$

for all $F \in \mathscr{F}$.

This reduces Theorem 5.10 to the following calculation:

Theorem 5.13 (Orthogonality to Dual Functions). *Let* $d \ge 2$ *and* $t \ge 0$ *be fixed, and let* $\mathbf{Q} = (Q_1, \dots, Q_d)$ *be a fixed* d*-tuple of polynomials* $Q_i \in \mathbb{Z}[\mathbf{h}_1, \dots, \mathbf{h}_t, \mathbf{W}]$, *not identically zero. Let* $\epsilon_0 > 0$ *be a fixed quantity that is sufficiently small depending on* $d, t, \mathbf{Q}$. *Let* $K > 0$ *be a fixed integer, and let* L *be a fixed quantity that is sufficiently large depending on* $d, t, \mathbf{Q}, \epsilon_0, K$. *Let* ν *be as above. Then one has*

$$\int_X (\nu - 1) \prod_{k=1}^{K} \mathscr{D}(F_{k,\omega})_{\omega \in \{0,1\}^d \backslash \{0\}^d} = o(1) \tag{24}$$

for all functions $F_{k,\omega} : X \to \mathbb{R}$ *for* $k = 1, \dots, K$, $\omega \in \{0, 1\}^d \backslash \{0\}^d$ *with the pointwise bounds* $-1 \le F_{k,\omega} \le 1$.

Remark 5.5. The fact that L depends on K here is the sole reason why L depends on δ in Theorem 1.5 (note that no parameter related to δ or K appears in Theorem 5.11).

We now prove Theorem 5.13. Let $d, t, \mathbf{Q}, \epsilon_0, K, L, \nu, F_{k,\omega}$ be as in that theorem. We expand out the left-hand side of (24) as the average of

$$\int_X (\nu - 1)\mathbb{E}_{m_{i,k}^{(\omega)} \in [\sqrt{M}] \forall i=1,\dots,d; k=1,\dots,K; \omega=0,1} \prod_{(\omega_1,\dots,\omega_d)\in\{0,1\}^d\setminus\{0\}^d} \prod_{k=1}^{K} T^{\sum_{i=1}^d (m_{i,k}^{(\omega_i)} - m_i^{(0)})Q_i(\mathbf{h}_k)} F_{k,\omega} \tag{25}$$

as $\mathbf{h}_1, \dots, \mathbf{h}_K$ ranges over $[H]^t$.

We first deal with the degenerate cases in which $Q_i(\mathbf{h}_k) = 0$ for some i, k. By the Schwarz–Zippel lemma (see, e.g., [19, Lemma D.3]), the number of tuples $(\mathbf{h}_1, \dots, \mathbf{h}_K)$ with this degeneracy is $O(H^{Kt-1})$. Meanwhile, from (12) and the boundedness of the $F_{k,\omega}$, each expression (25) is $O(1)$. Thus the total contribution of this case is $O(H^{-1})$, which is acceptable.

It thus suffices to show that the expression (25) is $o(1)$ uniformly for all $\mathbf{h}_1, \dots, \mathbf{h}_K$ with none of the $Q_i(\mathbf{h}_k)$ vanishing.

The next step is to "clear denominators" (as in [19]). Fix $\mathbf{h}_1, \dots, \mathbf{h}_K$, and write $D_i := \prod_{k=1}^K |Q_i(\mathbf{h}_k)|$ for $i = 1, \dots, d$. Then we have $1 \le D_i \ll O(HW)^{O(K)}$, and we can write

$$D_i = Q_i(\mathbf{h}_k) r_{i,k}$$

for each $i = 1, \dots, d$, $k = 1, \dots, K$, and some $r_{i,k} = O(HW)^{O(K)}$.

Let $n_1^{(0)}, \dots, n_d^{(0)}, n_1^{(1)}, \dots, n_d^{(1)}$ be elements of $[M^{1/4}]$. Then if we shift each variable $m_{i,k}^{(\omega)}$ by $r_{i,k} n_i^{(\omega)}$, we can rewrite (25) as

$$\int_X (\nu - 1)\mathbb{E}_{m_{i,k}^{(\omega)} \in [\sqrt{M}] - r_{i,k} n_i^{(\omega)} \forall i=1,\dots,d; k=1,\dots,K; \omega=0,1} \prod_{(\omega_1,\dots,\omega_d)\in\{0,1\}^d\setminus\{0\}^d} \prod_{k=1}^{K} T^{\sum_{i=1}^d (m_{i,k}^{(\omega_i)} - m_i^{(0)})Q_i(\mathbf{h}_k) + (n_i^{(\omega_i)} - n_i^{(0)})D_i} F_{k,\omega}. \tag{26}$$

The shifted interval $[\sqrt{M}] - r_{i,k} n_i^{(\omega)}$ differs from $[\sqrt{M}]$ by shifts by a set of cardinality $O(M^{1/4}(HW)^{O(K)})$, and so by (13) one can replace the former by the latter after accepting an additive error of $O((\log^{O(1)} N) M^{1/4} (HW)^{O(1)} / \sqrt{M})$, where the implied constants in the exponents depend on K. It is at this point that we crucially use the hypothesis that L be large compared with K, to ensure that this error is still $o(1)$. Thus (26) can be written as

$$\int_X (\nu - 1)\mathbb{E}_{m_{i,k}^{(\omega)} \in [\sqrt{M}] \forall i=1,\dots,d; k=1,\dots,K; \omega=0,1} \prod_{(\omega_1,\dots,\omega_d)\in\{0,1\}^d\setminus\{0\}^d} \prod_{k=1}^{K} T^{\sum_{i=1}^d (m_{i,k}^{(\omega_i)} - m_i^{(0)})Q_i(\mathbf{h}_k) + (n_i^{(\omega_i)} - n_i^{(0)})D_i} F_{k,\omega} + o(1).$$

Averaging over all such $n_i^{(\omega)}$, we obtain

$$\mathbb{E}_{m_{i,k}^{(\omega)} \in [\sqrt{M}] \forall i=1,\dots,d; k=1,\dots,K; \omega=0,1} \int_X (\nu-1) \mathbb{E}_{n_1^{(0)},\dots,n_d^{(0)},n_1^{(1)},\dots,n_d^{(1)} \in [M^{1/4}]}$$
$$\prod_{(\omega_1,\dots,\omega_d) \in \{0,1\}^d \setminus \{0\}^d} \prod_{k=1}^K T^{\sum_{i=1}^d (m_{i,k}^{(\omega_i)} - m_i^{(0)}) Q_i(\mathbf{h}_k) + (n_i^{(\omega_i)} - n_i^{(0)}) D_i} F_{k,\omega}$$
$$+ o(1).$$

Shifting the integral $\int_X$ by $\sum_{i=1}^d (n_i^{(0)}) D_i$ and then using the Gowers–Cauchy–Schwarz inequality (17) (and the boundedness of the functions $\prod_{k=1}^K T^{\sum_{i=1}^d (m_{i,k}^{(\omega_i)} - m_i^{(0)}) Q_i(\mathbf{h}_k)} F_{k,\omega}$), we may bound this by

$$\|\nu - 1\|_{U^{D_1,\dots,D_d}_{M^{1/4}}} + o(1).$$

But from expanding out the Gowers norm (15) and using Proposition 3.3 to estimate the resulting 2^{2^d} terms (cf. [10, Lemma 5.2]), we see that

$$\|\nu - 1\|^{2^d}_{U^{D_1,\dots,D_d}_{M^{1/4}}} = o(1) \tag{27}$$

and Theorem 5.13 follows.

6 Densification

Now we prove Theorem 5.11. It will suffice to establish the following claim.

Proposition 6.4. *Let the notation and hypotheses be as in Theorem 5.11. Then one has*

$$\left| \mathbb{E}_{\mathbf{h} \in [H]^t} \mathbb{E}_{m_1^{(0)},\dots,m_d^{(0)},m_1^{(1)},\dots,m_d^{(1)} \in [\sqrt{M}]} \int_X \prod_{\omega \in \{0,1\}^d} T^{m_1^{(\omega_1)} Q_1(\mathbf{h}) + \dots + m_d^{(\omega_d)} Q_d(\mathbf{h})} f_\omega \right| \tag{28}$$
$$\ll \varepsilon^c + o(1)$$

for some fixed $c > 0$ (independent of ε), whenever $(f_\omega)_{\omega \in \{0,1\}^d}$ is a tuple of functions $f_\omega : X \to \mathbb{R}$ (with $\omega := (\omega_1, \dots, \omega_d)$), such that one of the f_ω is equal to f, and each of the remaining functions f_ω in the tuple either obey the pointwise bound $|f_\omega| \leq 1$ or $|f_\omega| \leq \nu$.

Indeed, given the above proposition, then by triangle inequality and decomposition we may replace the bounds $|f_\omega| \leq 1$ or $|f_\omega| \leq \nu$ with $|f_\omega| \leq \nu + 1$, and then by setting $f_\omega = f$ for every ω and using (19), we obtain the claim.

It remains to prove the proposition. We induct on the number of factors f_ω for which one has the bound $|f_\omega| \leq \nu$ instead of $|f_\omega| \leq 1$. First suppose that there are no such factors, thus $|f_\omega| \leq 1$ for all ω except for one ω, for which $f_\omega = f$. By permuting the cube $\{0,1\}^d$, we may assume that it is $f_{\{0\}^d}$ that is equal to f, with all other f_ω bounded in magnitude by 1. But then the expression in (28) may be rewritten as

$$\int_X f \mathscr{D}(f_\omega)_{\omega \in \{0,1\}^d \setminus \{0\}},$$

and the claim follows from the hypothesis (23).

Now suppose that at least one of the f_ω (other than the one equal to f) is bounded in magnitude by ν rather than 1. By permuting the cube we may assume that $|f_{\{0\}^d}| \leq \nu$. We then write the left-hand side of (28) as

$$\int_X f_{\{0\}^d} \mathscr{D}\mathbf{f}$$

where $\mathbf{f} := (f_\omega)_{\omega \in \{0,1\}^d \setminus \{0\}}$. By Cauchy–Schwarz, it thus suffices to show that

$$\int_X \nu (\mathscr{D}\mathbf{f})^2 \ll \varepsilon^c + o(1)$$

for some fixed $c > 0$. We will split this into two estimates,

$$\left| \int_X (\nu - 1)(\mathscr{D}\mathbf{f})^2 \right| = o(1) \tag{29}$$

and

$$\int_X (\mathscr{D}\mathbf{f})^2 \ll \varepsilon^c + o(1). \tag{30}$$

We set aside (29) for now and work on (30). Bounding all the components of $\mathbf{f}$ in magnitude by $\nu + 1$ and using Proposition 3.3, we see that

$$\int_X (\mathscr{D}\mathbf{f})^4 \ll 1,$$

so by Hölder's inequality, it suffices to show that

$$\int_X |\mathscr{D}\mathbf{f}| \ll \varepsilon^c + o(1)$$

(for a possibly different fixed $c > 0$). It thus suffices to show that

$$|\int_X g \mathscr{D}\mathbf{f}| \ll \varepsilon^c + o(1)$$

whenever $g : X \to \mathbb{R}$ is such that $|g| \leq 1$. But this expression is of the form (28) with $f_{\{0\}}$ replaced by g, and the claim then follows from the induction hypothesis.

It thus remains to show (29). We can rewrite

$$(\mathscr{D}\mathbf{f})^2 = \mathscr{D}_{\sqrt{M}}^{\mathbf{Q}\oplus\mathbf{Q}([H]^t,W)}\mathbf{f}_2$$

where $\mathbf{Q} \oplus \mathbf{Q}$ is the $2d$-tuple

$$\mathbf{Q} \oplus \mathbf{Q} = (Q_1, \dots, Q_d, Q_1, \dots, Q_d)$$

and $\mathbf{f}_2 = (f_{2,\omega})_{\omega\in\{0,1\}^{2d}\setminus\{0\}^d}$ is defined by setting

$$f_{2,\omega\oplus\{0\}^d} := f_\omega$$

and

$$f_{2,\{0\}^d\oplus\omega} := f_\omega$$

for $\omega \in \{0,1\}^d\setminus\{0\}^d$, and

$$f_{2,\omega\oplus\omega'} := 1$$

for $\omega, \omega' \in \{0,1\}^d\setminus\{0\}^d$. Applying the Gowers–Cauchy–Schwarz inequality (17), we may thus bound the left-hand side of (29) by

$$\|\nu - 1\|_{U_{\sqrt{M}}^{\mathbf{Q}\oplus\mathbf{Q}([H]^t,W)}} \prod_{\omega\in\{0,1\}^d\setminus\{0\}} \|f_\omega\|_{U_{\sqrt{M}}^{\mathbf{Q}\oplus\mathbf{Q}([H]^t,W)}}^2.$$

Bounding f_ω by ν or 1 and using Proposition 3.3, we can bound

$$\|f_\omega\|_{U_{\sqrt{M}}^{\mathbf{Q}\oplus\mathbf{Q}([H]^t,W)}} \ll 1$$

and further application of Proposition 3.3 (cf. (27)) gives

$$\|\nu - 1\|_{U_{\sqrt{M}}^{\mathbf{Q}\oplus\mathbf{Q}([H]^t,W)}} = o(1)$$

and the claim follows.

7 The Linear Case

We now explain why in the linear case $P_i = (i-1)\mathbf{m}$ of Theorem 1.5, one may take L to be independent of δ. In the linear case, one can replace the averaged local Gowers norm $U^{\mathbf{Q}([H]^t,W)}_{\sqrt{M}}$ in Theorem 4.8 with the simpler norm $U^{1,\dots,1}_{\sqrt{M}}$, where 1 appears $d = k-1$ times; this follows by repeating the proof of [10, Proposition 5.3], after replacing some global averages with local ones. (In fact one could replace $\sqrt{M}$ here by $M^{1-\sigma}$ for any fixed $\sigma > 0$.) As such, we can ignore the H parameter and the $\mathbf{h}$ averaging, and just prove Theorem 5.13 in the case when $Q_1 = \dots = Q_d = 1$. Here, the "clearing denominators" step is unnecessary, and so L does not need to be large depending on K, which by Remark 5.5 ensures that the final L is independent of δ.

Remark 7.6. A more careful accounting of exponents (in particular, replacing (11) with a more precise asymptotic involving a singular series similar to that in (1)) allows one to take L as small as $Ck2^k$ for some absolute constant C; we omit the details.

Acknowledgements The first author is supported by a Simons Investigator grant, the James and Carol Collins Chair, the Mathematical Analysis & Application Research Fund Endowment, and by NSF grant DMS-1266164. Part of this research was performed while the first author was visiting the Institute for Pure and Applied Mathematics (IPAM), which is supported by the National Science Foundation. The second author is supported by ISF grant 407/12. The second author was on sabbatical at Stanford while part of this work was carried out; she would like to thank the Stanford math department for its hospitality and support. Finally, the authors thank the anonymous referee for a careful reading of the paper.

References

1. J. Benatar, The existence of small prime gaps in subsets of the integers. Int. J. Number Theory **11**(3), 801–833 (2015)
2. V. Bergelson, Weakly mixing PET. Ergodic Theory Dyn. Syst. **7**(3), 337–349 (1987)
3. V. Bergelson, A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems. J. Am. Math. Soc. **9**(3), 725–753 (1996)
4. V. Bergelson, A. Leibman, E. Lesigne, Intersective polynomials and the polynomial Szemerédi theorem. Adv. Math. **219**(1), 369–388 (2008)
5. D. Conlon, T. Gowers, Combinatorial theorems in sparse random sets. Preprint
6. D. Conlon, J. Fox, Y. Zhao A relative Szemerédi theorem. Geom. Funct. Anal. **25**(3), 733–762 (2015)
7. P.X. Gallagher, On the distribution of primes in short intervals. Mathematika **23**, 4–9 (1976)
8. D. Goldston, J. Pintz, C. Yildirim, Primes in tuples IV: density of small gaps between consecutive primes. Acta Arith. **160**(1), 37–53 (2013)
9. W.T. Gowers, Decompositions, approximate structure, transference, and the Hahn-Banach theorem. Bull. Lond. Math. Soc. **42**(4), 573–606 (2010)
10. B. Green, T. Tao, The primes contain arbitrarily long arithmetic progressions. Ann. Math. (2) **167**(2), 481–547 (2008)
11. B. Green, T. Tao, Linear equations in primes. Ann. Math. **171**, 1753–1850 (2010)

12. H. Halberstam, H.-E. Richert, *Sieve Methods* (Academic, New York, 1974)
13. 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**, 1–70 (1923)
14. T.H. Le, Intersective polynomials and the primes. J. Number Theory **130**(8), 1705–1717 (2010)
15. J. Maynard, Small gaps between primes. Ann. of Math. (2) **181**(1), 383–413 (2015)
16. O. Reingold, L. Trevisan, M. Tulsiani, S. Vadhan, Dense subsets of pseudorandom sets, in *Proceedings of 49th IEEE FOCS, Electronic Colloquium on Computational Complexity*, 2008
17. M. Schacht, Extremal results for random discrete structures. Preprint
18. E. Szemerédi, On sets of integers containing no k elements in arithmetic progression. Acta Arith. **27**, 299–345 (1975)
19. T. Tao, T. Ziegler, The primes contain arbitrarily long polynomial progressions. Acta Math. **201**(2), 213–305 (2008)
20. Y. Zhang, Bounded gaps between primes. Ann. of Math. (2) **179**(3), 1121–1174 (2014)