

RESEARCH



Cotangent sums, quantum modular forms, and the generalized Riemann hypothesis

John Lewis¹ and Don Zagier^{2*}

*Correspondence:
Don.Zagier@mpim-
bonn.mpg.de
²Max Planck Institute for
Mathematics, Bonn, Germany
Full list of author information is
available at the end of the article

Abstract

We show that an asymptotic property of the determinants of certain matrices whose entries are finite sums of cotangents with rational arguments is equivalent to the GRH for odd Dirichlet characters. This is then connected to the existence of certain quantum modular forms related to Maass Eisenstein series.

1 Introduction and main results

To answer immediately a question that the reader may be asking, we should say from the outset that the generalized Riemann hypothesis (henceforth, just GRH) of the title refers only to the L -series associated with odd real Dirichlet characters and that we do not aim to prove it, but merely to show its equivalence to an asymptotic statement about the determinants of certain matrices whose entries $c_{m,n}$ are given by finite sums of cotangents of rational multiples of π . To answer a second natural question, since it may seem unusual for an author to contribute to the proceedings volume of a conference partly intended to celebrate his own birthday, we should also say a few words about the genesis of this joint paper. Both the definition of the cotangent sums considered and their relation to GRH were found by the first author and presented in his talk at the conference in question. In the course of that talk he mentioned the homogeneity property $c_{\ell m, \ell n} = \ell c_{m,n}$, leading the second author to ask whether the function $C : \mathbb{Q} \rightarrow \mathbb{R}$ defined by $C(m/n) = c_{m,n}/n$ might be related to a quantum modular form in the sense of [16]. The answer to this question turned out to be positive. Furthermore, and unexpectedly, this quantum modular aspect turned out also to be related to the earlier joint work of the authors [13] on the analogs of classical period polynomials for Maass waveforms. A glance at the title of the final section of this paper will show why it seemed natural, and even irresistible, to publish both parts of the story in these proceedings.

We now describe the main results of the paper in more detail. We denote by

$$\chi(n) = \begin{cases} (-1)^{(n-1)/2} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even,} \end{cases} \quad L(s) = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots. \quad (1)$$

the primitive Dirichlet character of conductor 4 and its associated L -series. (In Sect. 5 we will describe the generalization to other odd primitive real Dirichlet characters.) For positive integers m and n we define a real number $c_{m,n}$ by

© The Author(s) 2018. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

$$c_{m,n} := \frac{4}{\pi} \sum_{j,k>0} \frac{\chi(j)\chi(k)}{\max(j/m, k/n)}, \tag{2}$$

where the conditionally convergent sum is to be interpreted as the limit for $N \rightarrow \infty$ of the sum over the rectangle $1 \leq j \leq mN, 1 \leq k \leq nN$. These numbers are studied in more detail in Sect. 2, where we will see that they are algebraic numbers (in fact, algebraic integers) that can be expressed as finite sums of cotangents of odd multiples of $\pi/4m$ or $\pi/4n$, a typical example being

$$c_{3,4} = \cot\left(\frac{\pi}{12}\right) + \cot\left(\frac{5\pi}{12}\right) + \cot\left(\frac{9\pi}{16}\right) + \cot\left(\frac{13\pi}{16}\right) = 6 - 2\sqrt{2 + \sqrt{2}}.$$

For each $N \in \mathbb{N}$ we define the two symmetric matrices

$$C_N := \begin{pmatrix} c_{1,1} & \cdots & c_{1,N} \\ \vdots & \ddots & \vdots \\ c_{N,1} & \cdots & c_{N,N} \end{pmatrix}, \quad \widehat{C}_N := \left(\begin{array}{c|c} C_N & \begin{matrix} 1 \\ \vdots \\ N \end{matrix} \\ \hline 1 \cdots N & 4N/\pi \end{array} \right), \tag{3}$$

which we will show later are positive definite, and define a function $R : \mathbb{N} \rightarrow \mathbb{R}$ by

$$R(N) := N \frac{\det C_N}{\det \widehat{C}_N}. \tag{4}$$

The first main result of this paper is then the following.

Theorem 1 *The following two statements are equivalent:*

- (a) *The function $R(N)$ is unbounded.*
- (b) *The L -series $L(s)$ has no zeros in the half-plane $\sigma := \Re(s) > \frac{1}{2}$ (GRH).*

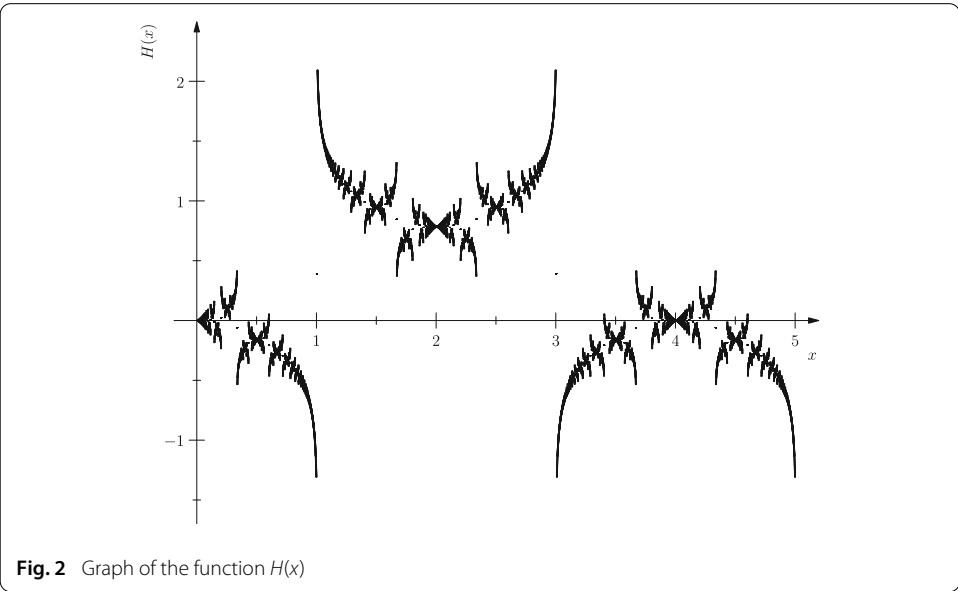
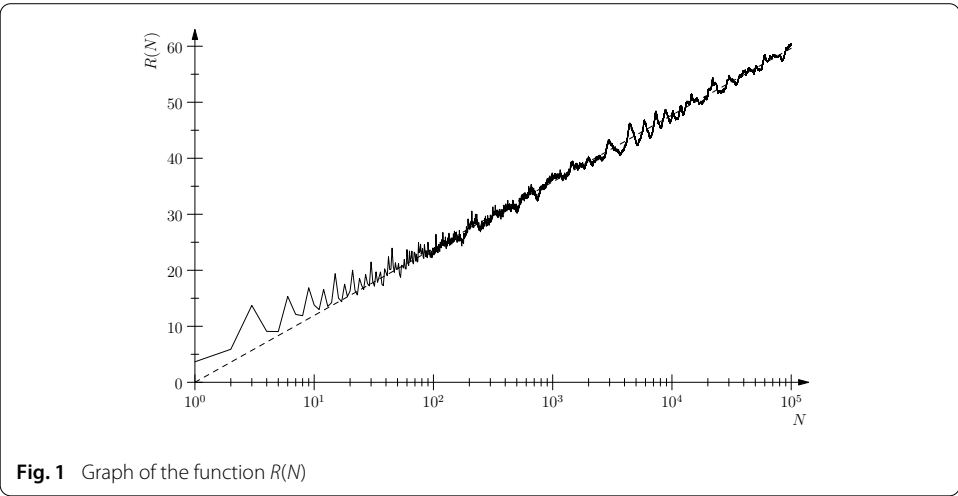
The proof of this theorem will be given in Sect. 3. The direction (a) \Rightarrow (b) is fairly elementary, while the reverse direction uses functional analysis and techniques coming from old work of Beurling. The result that we actually prove (Theorem 3 in §3) is in fact stronger than Theorem 1 in both directions: if GRH holds, then $R(N)$ is not merely unbounded, but actually tends to infinity as $N \rightarrow \infty$, and if $L(s)$ has a zero ρ with $\Re(\rho) > \frac{1}{2}$, then the function $R(N)$ is bounded by a number depending explicitly on ρ . Of course we believe that the two statements in Theorem 1 are both true rather than both false, the evidence being on the one hand the generally held belief in the validity of (b) and on the other hand the plot of the function $R(N)$ for $N \leq 100,000$ shown in Fig. 1, which suggests that $R(N)$ tends to infinity and perhaps even grows roughly linearly with $\log N$. (The straight-line fit shown in the picture is the graph of the function $5.18 \log N$.)

The second main result concerns the quantum modular nature of certain functions related to the coefficients $c_{m,n}$. On the one hand, the numbers $c_{m,n}$ defined by (2) have the obvious homogeneity property $c_{\ell m, \ell n} = \ell c_{m,n}$ and can be rewritten in the form

$$c_{m,n} = \frac{4n}{\pi} C\left(\frac{m}{n}\right), \quad C(x) := \sum_{j,k>0} \chi(j)\chi(k) \min\left(\frac{x}{j}, \frac{1}{k}\right) \quad (x \in \mathbb{Q}_{>0}). \tag{5}$$

On the other hand, by splitting the sums in (2) or (5) into two pieces depending on which its two arguments realize the “max” or “min,” we obtain the decompositions

$$c_{m,n} = h_{m,n} + h_{n,m}, \quad C(x) = H(x) + xH\left(\frac{1}{x}\right), \tag{6}$$



where the function $H : \mathbb{Q}_{>0} \rightarrow \mathbb{R}$ and numbers $h_{m,n} \in \mathbb{R}$ are defined by

$$H(x) := \sum_{0 < j \leq kx}^* \frac{\chi(k) \chi(j)}{k}, \quad h_{m,n} := \frac{4n}{\pi} H\left(\frac{m}{n}\right). \tag{7}$$

(Here and later an asterisk on a summation sign means that terms with equality—in this case, those with $j = kx$ —are to be counted with multiplicity $1/2$.) We will show in Sect. 2 that $h_{m,n}$ (or its sum with $1/2$ if mn is odd) is a sum of cotangents of odd multiples of $\pi/4n$ and hence is an algebraic integer, refining the corresponding properties for $c_{m,n}$ given above.

We can now come to the quantum modular form property mentioned in the opening paragraph. The graph of the function H on \mathbb{Q} is shown in Fig. 2.

This graph suggests that the function H is only naturally defined on \mathbb{Q} and is not well behaved as a function on \mathbb{R} , having a possibly dense set of discontinuities and certainly not differentiable at generic points. The graph of the function C shown in Fig. 3, on the other hand, suggests that it is everywhere continuous and possibly even differentiable except

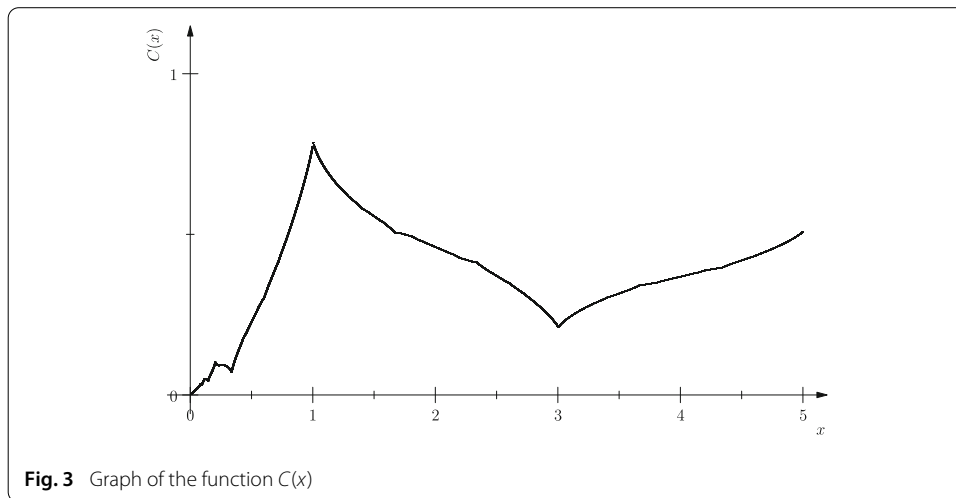


Fig. 3 Graph of the function $C(x)$

at rational points with odd numerator and denominator, where it seems to be left and right differentiable but with a discontinuous derivative. Our second main result, proved in Sect. 4, refutes the last of these statements and confirms the others.

Theorem 2 *The function $H(x)$ is discontinuous at all rational points x with odd numerator and denominator and is continuous but not left or right differentiable at other rational points. The function $C(x)$ is continuous everywhere, is not left or right differentiable at rational points with odd numerator and denominator, but is differentiable at all other rational points.*

The similarity between the continuity and differentiability statements for the two functions $H(x)$ and $C(x)$ in this theorem is not coincidental, since we will see in Sect. 4 that the derivative of $C(x)$ is equal to $H(1/x)$.

We now discuss the “quantum modular form” aspect of Theorem 2. Note first that the function H is periodic of period 4 and hence extends to a periodic function on all of \mathbb{Q} . This function is even, so we could also extend H from $\mathbb{Q}_{>0}$ to \mathbb{Q} by $H(x) := H(|x|)$. One might then think of extending C to negative values of the arguments by setting $C(x) = H(x) + xH(1/x)$ for all $x \neq 0$, but if we did that then $C(x)$ for $x < 0$ would exhibit oscillatory behavior of the same sort as $H(x)$. Instead, we extend C to all of \mathbb{Q} as an even function, i.e., we set $C(x) = C(|x|)$ for $x \neq 0$ and extend this by continuity to $C(0) = 0$. Then the relationship between C and H becomes

$$H(x) + |x|H(1/x) = C(x) \quad (x \neq 0). \tag{8}$$

On the other hand, the 4-periodicity property of $H(x)$ can be strengthened to the statement

$$H(x) + H(x + 2) = \frac{\pi}{4}, \tag{9}$$

saying that up to sign and up to a constant H actually has period 2. Also, since H is even, we can replace the term $H(1/x)$ in (8) by $H(-1/x)$. But the two matrices $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $T^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ generate the subgroup Γ_ϑ of index 3 in the full modular group $SL(2, \mathbb{Z}) = \langle S, T \rangle$ consisting of matrices congruent to $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ modulo 2 (this is the so-called theta group, under which the Jacobi theta function $\vartheta(z) = \sum_n e^{\pi i n^2 z}$ transforms like a modular form of weight 1/2), so Eqs. (8) and (9) can be combined to the following statement:

Corollary *The function $H : \mathbb{Q} \rightarrow \mathbb{R}$ satisfies the transformation property*

$$\varepsilon(\gamma) |cx + d| H\left(\frac{ax + b}{cx + d}\right) = H(x) + (\text{continuous function})$$

for every matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_\vartheta$, where $\varepsilon : \Gamma_\vartheta \rightarrow \{\pm 1\}$ is the homomorphism sending both generators $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ of Γ_ϑ to -1 .

This corollary tells us precisely that the function H is a quantum modular form in the sense of [16] (except that there is an absolute value sign here that was not present there). Recall that a “quantum modular form” is by definition a function on \mathbb{Q} that does not extend nicely to \mathbb{R} but which modulo (the restriction from \mathbb{R} to \mathbb{Q} of) better-behaved (here, continuous) functions on \mathbb{R} transforms like a modular form with respect to some subgroup of finite index of $SL(2, \mathbb{Z})$ (here the theta group Γ_ϑ) with some weight (here -1) and character (here ε). In Sect. 6 of this paper we will show that the quantum modularity property of H is related to an actual modular form (specifically, a Maass Eisenstein series of eigenvalue $1/4$) on the group Γ_ϑ .

The authors would like to thank Steven Johnson (MIT) and Alexander Weisse (MPI) for invaluable help with the numerical aspects of this paper.

2 Properties of the numbers $c_{m,n}$ and $h_{m,n}$

In this section we give an integral representation for the numbers $c_{m,n}$ that will be used later both to extend the function $C(x)$ continuously from \mathbb{Q} to \mathbb{R} and as the key to the proof in Sect. 3 of the relation between the numbers $c_{m,n}$ and the zeros of the Dirichlet L -series $L(s)$. We will also prove the formulas for $c_{m,n}$ and $h_{m,n}$ as algebraic numbers (finite sums of cotangents).

The first step is to rewrite the definition of H given in (7) more compactly in the form

$$H(x) = \sum_{k=1}^{\infty} \frac{\chi(k)}{k} S(kx) \quad (x \in \mathbb{Q}), \tag{10}$$

where $S(x)$ is the function defined on $\mathbb{R}_{>0}$ by

$$S(x) := \sum_{0 < j \leq x}^* \chi(j). \tag{11}$$

The function S , like H itself, is periodic of period 4, and we can use this property to extend it to all of \mathbb{R} . (Equivalently, we can extend it by $S(-x) = S(x)$ and $S(0) = 0$, or by observing that the summation condition in (11) can be replaced by $J < j \leq x$ for any integer $J < x$ divisible by 4, which works for all real x .) The function S is then the step function given by

$$S(x) = \begin{cases} 1 & \text{if } 4k + 1 < x < 4k + 3 \text{ for some } k \in \mathbb{Z}, \\ 0 & \text{if } 4k - 1 < x < 4k + 1 \text{ for some } k \in \mathbb{Z}, \\ \frac{1}{2} & \text{if } x \text{ is an odd integer,} \end{cases} \tag{12}$$

as pictured in Fig. 4. Note that the sum in (10) converges (conditionally) for $x \in \mathbb{Q}$ by summation by parts, because the function $\chi(k)S(kx)$ is then an odd periodic function of k and hence has average value zero, so that the sums $\sum_{k=1}^K \chi(k)S(kx)$ are bounded.

We use the bounded periodic function S to prove the following formula for the $c_{m,n}$:

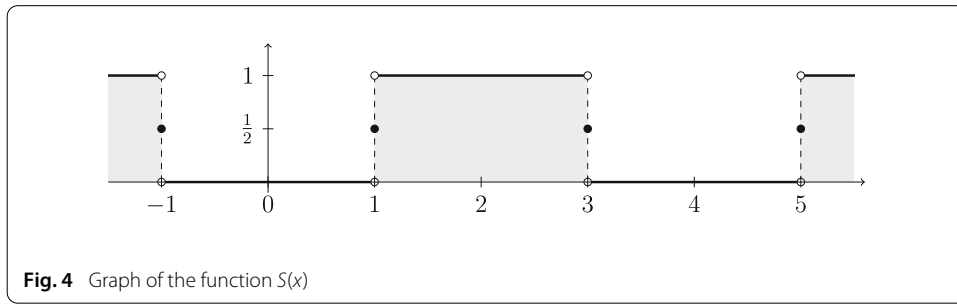


Fig. 4 Graph of the function $S(x)$

Proposition 1 *The numbers $c_{m,n}$ defined by (2) have the integral representation*

$$c_{m,n} = \frac{4}{\pi} \int_0^\infty S(mt) S(nt) \frac{dt}{t^2} \quad (m, n \in \mathbb{N}). \tag{13}$$

Proof From the definition of S we find

$$\int_0^N S(mt) S(nt) \frac{dt}{t^2} = \int_0^N \sum_{\substack{0 < j \leq mt \\ 0 < k \leq nt}}^* \chi(j)\chi(k) \frac{dt}{t^2} = \sum_{\substack{0 < j \leq mN \\ 0 < k \leq nN}} \left(\frac{\chi(j)\chi(k)}{\max(j/m, k/n)} - \frac{\chi(j)\chi(k)}{N} \right),$$

and since the second sum is identically zero for N divisible by 4, Eq. (13) follows by letting N tend to infinity in $4\mathbb{Z}$. \square

Our next result expresses $c_{m,n}$ as a finite sum of algebraic cotangent values indexed by lattice points in the rectangle $[0, m/2] \times [0, n/2]$ satisfying a certain “selection rule.”

Proposition 2 *Let m and n be positive integers. Then the number defined by (2) is given as the following integral linear combination of cotangents of rational multiples of π :*

$$c_{m,n} = \sum_{\substack{0 \leq j \leq m/2 \\ 0 \leq k \leq n/2}} \begin{cases} \cot(\pi a) - \cot(\pi b) & \text{if } a := \max\left(\frac{4j+1}{4m}, \frac{4k+1}{4n}\right) < b := \min\left(\frac{4j+3}{4m}, \frac{4k+3}{4n}, \frac{1}{2}\right), \\ 0 & \text{if } a \geq b. \end{cases}$$

Corollary *The numbers $c_{m,n}$ are algebraic integers for all $m, n > 0$.*

Using this proposition we can calculate any $c_{m,n}$ in closed form as an algebraic number. For example, the first 4×4 matrix of values [cf. Eq. (3)] is given by

$$C_4 = \begin{pmatrix} 1 & 2 - \sqrt{2} & 2 - \sqrt{3} & 2 - 2\sqrt{2 - \sqrt{2}} \\ 2 - \sqrt{2} & 2 & \sqrt{2} & 4 - 2\sqrt{2} \\ 2 - \sqrt{3} & \sqrt{2} & 3 & 6 - 2\sqrt{2 + \sqrt{2}} \\ 2 - 2\sqrt{2 - \sqrt{2}} & 4 - 2\sqrt{2} & 6 - 2\sqrt{2 + \sqrt{2}} & 4 \end{pmatrix}.$$

Proof Since $S(x)$ is even and periodic of period 4, we can rewrite the integral representation (13) using Euler’s formula $\sum_{k \in \mathbb{Z}} (k + t)^{-2} = \pi^2 / \sin^2(\pi t)$ as

$$c_{m,n} = \pi \int_0^{1/2} S(4mx) S(4nx) \frac{dx}{\sin^2(\pi x)}.$$

But from the definition of $S(4mx)$ and $S(4nx)$ as characteristic functions, it follows that the restriction of $S(4mx)S(4nx)$ to $[0, \frac{1}{2}]$ is the characteristic function of the union of all intervals $[a, b]$ with $a < b$ as given in the proposition. This proves the proposition because $\pi \int_a^b \csc^2(\pi x) dx = \cot(\pi a) - \cot(\pi b)$. The corollary then follows because by elementary

algebraic number theory the number $\cot(\pi\alpha)$ is an algebraic integer for every rational number α having an even denominator. \square

We next give the corresponding formula for the numbers $h_{m,n}$, which is in fact simpler.

Proposition 3 *Let m and n be positive integers. Then the number defined by (7) is given as the following half-integral linear combination of cotangents of odd multiples of $\pi/4n$:*

$$h_{m,n} = \sum_{k=1}^{2n-1} \chi(k) S\left(\frac{km}{n}\right) \cot\left(\frac{\pi k}{4n}\right). \tag{14}$$

Proof Since the summand in (10) is an even function of k , we can rewrite the sum as a sum over all nonzero integral values of k and then use the periodicity (with period at most $4n$) of the factor $\chi(k)S(km/n)$ together with Euler’s formula $\sum_{n \in \mathbb{Z}} (x+n)^{-1} = \pi \cot(\pi x)$ (with the sum being interpreted as a Cauchy principal value) to get

$$\begin{aligned} h_{m,n} &= \frac{2n}{\pi} \sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{\chi(k)}{k} S\left(\frac{km}{n}\right) = \frac{2n}{\pi} \sum_{k \pmod{4n}, k \neq 0} \chi(k) S\left(\frac{km}{n}\right) \sum_{r \in \mathbb{Z}} \frac{1}{4nr + k} \\ &= \frac{1}{2} \sum_{0 < k < 4n} \chi(k) S\left(\frac{km}{n}\right) \cot\left(\frac{\pi k}{4n}\right) = \sum_{0 < k < 2n} \chi(k) S\left(\frac{km}{n}\right) \cot\left(\frac{\pi k}{4n}\right) \end{aligned}$$

as claimed. \square

Remark One can also deduce the formula for $c_{m,n}$ in Proposition 2 from (14); this is left as an exercise to the reader. We also observe that the terms on the right-hand side of (14) (assuming, as we can, that m and n are coprime) are integral multiples of cotangents of odd multiples of $\pi/4n$, and hence are algebraic integers (in some cyclotomic field), except for the $k = n$ term if m and n are odd, which is half-integral because then $S(km/n) = \frac{1}{2}$ and $\cot(\pi k/4n) = 1$. Hence $h_{m,n}$ is an algebraic integer unless both m and n are odd, in which case it equals an algebraic integer plus $1/2$, and the sum $c_{m,n} = h_{m,n} + h_{n,m}$ is always an algebraic integer.

3 The connection with the generalized Riemann hypothesis

In the first subsection of this section we outline the basic strategy for proving Theorem 1. The easier direction (a) \Rightarrow (b) of this theorem is then proved in Sect. 3.2 and the more difficult converse direction in Sect. 3.3, using ideas from two classical papers by A. Beurling. As already stated in introduction, we will actually prove the following stronger result:

Theorem 3 *Let $L(s)$ be the L-series defined in (1) and $R(N)$ be the function defined in (4).*

- (a) *If $L(s)$ has no zeros with $\Re(s) > \frac{1}{2}$, then $R(N)$ tends to infinity as N tends to infinity.*
- (b) *If $L(\rho) = 0$ for some $\rho \in \mathbb{C}$ with $\Re(\rho) > \frac{1}{2}$, then $R(N)$ satisfies the bound*

$$R(N) \leq \frac{\pi}{8} \frac{|\rho|^2}{\Re(\rho) - \frac{1}{2}} \quad (\forall N \geq 1). \tag{15}$$

3.1 The square-wave function G and its multiplicative shifts

Set $L^2 = L^2(0, 1)$ and let ξ be the characteristic function of the interval $(0, 1]$ and $G : \mathbb{R}_{>0} \rightarrow \mathbb{R}$ the function

$$G(x) = S(1/x) = \begin{cases} 1 & \text{if } \frac{1}{4k+3} < x < \frac{1}{4k+1} \text{ for some } k \in \mathbb{Z}_{\geq 0}, \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

(and $G(x) = 1/2$ if $1/x$ is an odd integer, but we are working in L^2). Clearly both ξ and G belong to L^2 , since they are bounded with support in the interval $(0, 1]$. On the other hand, if we identify L^2 with the space of square-integrable functions on $\mathbb{R}_{>0}$ that vanish on $(1, \infty)$, then we have bounded operators τ_a on L^2 defined by

$$\tau_a f(x) = f(x/a) \quad (0 < a \leq 1, x \in \mathbb{R}_{>0})$$

We set $G_a = \tau_a G$ and for any $\mathcal{D} \subseteq \mathbb{R}$ and $N \in \mathbb{N}$ define subspaces $\mathcal{V}_{\mathcal{D}}$ and \mathcal{V}_N of L^2 by

$$\mathcal{V}_{\mathcal{D}} = \langle G_a | a \in \mathcal{D} \cap (0, 1] \rangle, \quad \mathcal{V}_N = \langle G_{n/N} | 1 \leq n \leq N \rangle, \tag{17}$$

where $\langle S \rangle$ denotes the linear span of a subset S of a linear space. Finally, we denote by $d(\xi, \mathcal{V}_N)$ the L^2 distance from $\xi \in L^2$ to the subspace \mathcal{V}_N , viz.

$$d(\xi, \mathcal{V}_N) = \|\xi - \mathcal{V}_N\| = \inf_{x \in \mathbb{R}^N} \left\| \xi - \sum_{n=1}^N x_n \tau_{n/N} G \right\|. \tag{18}$$

The key connection is then given by the following proposition.

Proposition 4 *The distance $d(\xi, \mathcal{V}_N)$ is related to the number $R(N)$ defined in (4) by*

$$d(\xi, \mathcal{V}_N)^2 = \frac{\pi/4}{R(N)} \quad (N \in \mathbb{N}). \tag{19}$$

The proof of this, which is a straightforward computation, will be given in Sect. 3.2. On the other hand, a quite elementary argument, also given in Sect. 3.2, proves the following

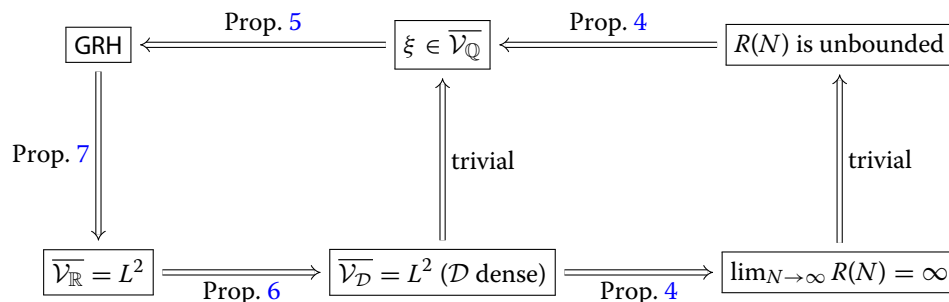
Proposition 5 *If $L(\rho) = 0$ with $\Re(\rho) > \frac{1}{2}$ then $d(\xi, \mathcal{V}_N)^2 \geq \frac{2\Re(\rho)-1}{|\rho|^2}$ for all $N \in \mathbb{N}$.*

Together these two results prove statement (b) of Theorem 3. The harder statement (a) follows from the two following results, proved in Sects. 3.2 and 3.3, respectively.

Proposition 6 *The space $\mathcal{V}_{\mathcal{D}}$ is dense in $\mathcal{V}_{\mathbb{R}}$ for any set \mathcal{D} that is dense in $(0, 1)$.*

Proposition 7 *If $L(s)$ has no zeros in $\sigma > \frac{1}{2}$, then $\mathcal{V}_{\mathbb{R}}$ is dense in L^2 .*

The whole argument is summarized by the following diagram, where the flow of the double arrows shows that the six boxed statements are mutually equivalent.



(The implications labeled “Prop. 4” follow in one direction because each \mathcal{V}_N is contained in $\mathcal{V}_{\mathbb{Q}}$ and in the other because if $R(N)$ were bounded for some infinite sequence of integers $\{N_i\}$ then the set $\cup_i \mathcal{V}_{N_i}$ would be dense in $\mathcal{V}_{\mathbb{R}}$ but have a nonzero distance from ξ .)

3.2 Proof of Propositions 4, 5 and 6

Recall from linear algebra that the distance $d(v, V)$ from a vector v in a (real) Hilbert space to a finite-dimensional vector space V with basis $\{v_1, \dots, v_N\}$ is given by the formula $d(v, V)^2 = \det(\widehat{G})/\det(G)$, where G is the $N \times N$ Gram matrix with (m, n) -entry $\langle v_m, v_n \rangle$ (scalar product of the basis elements v_m and v_n) and \widehat{G} the augmented $(N + 1) \times (N + 1)$ matrix obtained from G by adding a row and column with entries $\langle v, v_n \rangle$ and a diagonal entry $\langle v, v \rangle$. We apply this to the vector $v = \xi \in L^2$ and the space \mathcal{V}_N defined as above as the span of the vectors $G_{n/N} \in L^2$ ($n = 1, \dots, N$). (Here we mean the real rather than the complex span, i.e., we work in the real Hilbert space of real-valued square-integrable functions on $(0, 1)$.) We also denote by $\widehat{\mathcal{V}}_N$ the span of \mathcal{V}_N and ξ . We will show in a moment that the vectors $G_{n/N}$ and ξ are linearly independent, so $\dim \widehat{\mathcal{V}}_N = \dim \mathcal{V}_N + 1 = N + 1$. We can then compute the entries of the Gram matrices as

$$\langle \xi, \xi \rangle = \int_0^1 dx = 1, \quad \langle \xi, G_{n/N} \rangle = \int_0^1 S\left(\frac{n}{Nx}\right) dx = \frac{n}{N} L(1) = \frac{\pi n}{4N}$$

[the first is trivial and the second a special case of Eq. (21)] and

$$\langle G_{m/N}, G_{n/N} \rangle = \int_0^1 S\left(\frac{m}{Nx}\right) S\left(\frac{n}{Nx}\right) dx = \frac{m}{N} C\left(\frac{m}{n}\right) = \frac{\pi}{4N} c_{m,n}$$

[combine Eqs. (25) and (5)]. Proposition 4 follows.

To find the connection with the L -series $L(s)$ and its complex zeros, we next consider the “complex monomial functions” p_s on $(0, 1]$ defined by

$$p_s(t) = t^{s-1} \xi(t) = t^{s-1} \quad (0 < t \leq 1, s \in \mathbb{C}).$$

Note the key fact that $p_s \in L^2$ if and only if $\sigma > \frac{1}{2}$, and also that for $f \in L^2$ and $\sigma > \frac{1}{2}$, we have $\langle p_s, f \rangle = \widetilde{f}(s)$, where $\widetilde{f}(s)$ denotes the Mellin transform

$$\widetilde{f}(s) = \int_0^\infty f(x) x^{s-1} dx = \int_0^1 f(x) x^{s-1} dx.$$

(Here we are again identifying $L^2 = L^2(0, 1)$ with the space of functions $L^2(\mathbb{R}_{>0})$ supported on $(0, 1]$.) Then obviously

$$\langle p_s, \xi \rangle = \int_0^1 x^{s-1} dx = \frac{1}{s}, \tag{20}$$

and the Mellin transform of G_a is almost equally easy to calculate using (12):

$$\langle p_s, G_a \rangle = \int_0^\infty S\left(\frac{a}{x}\right) x^{s-1} dx = \frac{1}{s} \sum_{k=0}^\infty \left(\frac{a^s}{(4k+1)^s} - \frac{a^s}{(4k+3)^s} \right) = \frac{a^s}{s} L(s), \tag{21}$$

the calculation being valid for all s in $\Re(s) > 0$ since the integral then converges absolutely. One consequence of this is that the functions G_a ($0 < a \leq 1$) are all linearly independent, since if some (finite) linear combination of them vanished then the vanishing of its Mellin transform would imply that the product of $L(s)$ with a finite Dirichlet series vanished identically (first for $\Re(s) > 0$, and then for all s by analytic continuation). In particular, the space \mathcal{V}_N has dimension exactly N . In fact, almost the same argument also shows that the functions G_a and ξ are linearly independent, since a linear relation among them would imply that the product of $L(s)$ with a finite Dirichlet series was constant, which is obviously impossible. It follows that the extended space $\widehat{\mathcal{V}}_N$ spanned by the functions $G_{n/N}$ ($1 \leq n \leq N$) and ξ has dimension $N + 1$, as claimed above.

The proof of the direction (a) \Rightarrow (b) of Theorem 1 now follows easily. If there exists ρ with $\text{Re}(\rho) > \frac{1}{2}$ and $L(\rho) = 0$ then by (21) we have $p_\rho \perp \mathcal{V}_N$ for all N . Since \mathcal{V}_N is

finite-dimensional, the infimum in Eq. (18) is attained, so that $R(N)$ is unbounded (and hence $d(\xi, \mathcal{V}_N)$ has \liminf equal to 0) if and only if there is a sequence of functions $h_N \in \mathcal{V}_N$ converging in L^2 to ξ , and this would contradict $\mathfrak{p}_\rho \perp \mathcal{V}_N$ for all N since $\langle \mathfrak{p}_\rho, \xi \rangle = \rho^{-1} \neq 0$.

This argument actually proves the stronger statement given in Proposition 5 and in part (b) of Theorem 3. Indeed, if ρ is a zero of $L(s)$ with $\Re(\rho) > \frac{1}{2}$, then on the one hand \mathfrak{p}_ρ belongs to L^2 as already noted, with L^2 norm

$$\|\mathfrak{p}_\rho\|^2 = \int_0^1 |x^{\rho-1}|^2 dx = \frac{1}{2\Re(\rho) - 1},$$

and on the other hand $\langle \mathfrak{p}_\rho, f \rangle = 0$ for any $f \in \mathcal{V}_N$ by virtue of (21). We apply this to the function $f = f_N \in \mathcal{V}_N$ attaining the minimum distance to ξ . Then applying the Cauchy–Schwarz inequality to the scalar product $\langle \mathfrak{p}_\rho, \xi - f_N \rangle$ we get [using (20)]

$$\frac{1}{|\rho|^2} = |\langle \mathfrak{p}_\rho, \xi \rangle|^2 = |\langle \mathfrak{p}_\rho, \xi - f_N \rangle|^2 \leq \|\mathfrak{p}_\rho\|^2 \|\xi - f_N\|^2 = \frac{d(\xi, \mathcal{V}_N)^2}{2\Re(\rho) - 1},$$

and (15) now follows immediately from (19).

Finally, we prove Proposition 6 (“density lemma”). We first note that the L^2 norm $\|G_a - G_b\|$ tends to 0 as a tends to b , since by the calculation $\langle G_a, G_b \rangle = aC(a/b)$ given above (for rational numbers, but true in general by the same argument) we have

$$\|G_a - G_b\|^2 = \langle G_a, G_a \rangle + \langle G_b, G_b \rangle - 2\langle G_a, G_b \rangle = (a + b)C(1) - 2aC(b/a),$$

which tends to 0 as $b \rightarrow a$ because $C(x)$ is a continuous function. It follows immediately that if \mathcal{D} is dense in $(0, 1)$ then the vector $G_a \in L^2$ for any $a < 1$ lies in the closure of $\mathcal{V}_{\mathcal{D}}$, and since $\mathcal{V}_{\mathbb{R}}$ is spanned by such vectors, this shows that $\overline{\mathcal{V}_{\mathbb{R}}} = \overline{\mathcal{V}_{\mathcal{D}}}$ as claimed.

3.3 Proof of Proposition 7

It remains to show that GRH implies that $\overline{\mathcal{V}_{\mathbb{R}}} = L^2$.

Let $H^2 = H^2(\mathcal{R})$ denote the Hardy space of the right half-plane $\mathcal{R} = \{s \mid \Re(s) > \frac{1}{2}\}$, that is, the space of all holomorphic functions $F(s)$ on \mathcal{R} that are square-integrable on the lines $\Re(s) = c$ for all $c > \frac{1}{2}$ and with these L^2 norms $\|\cdot\|_c$ bounded in c , and then equipped with the Hilbert space norm $\|F\| = \sup\{\|F\|_c \mid c > \frac{1}{2}\}$. We quote some results from the standard analysis of these spaces, citing as we go the relevant sections in the books by Hoffman [11] and Garnett [8] for the details.

The Paley–Wiener theorem [11, pp. 131–132] gives the fact that the Mellin transform is (up to scalar multiple) an isometry between the spaces L^2 and H^2 . Applying this to the function G implies that its Mellin transform $\tilde{G}(s) = L(s)/s$ lies in the space H^2 , a fact which is also easy to see directly from the definition of H^2 . Next, we note that if we assume that GRH holds, so that $\tilde{G}(s)$ has no zeros in \mathcal{R} , then the factorization theorem for functions in the space H^2 [11, pp. 132–133] applied to the function $\tilde{G}(s)$ gives the implication

$$L(s) \text{ has no zeros in } \sigma > \frac{1}{2} \implies \tilde{G}(s) \text{ is an outer function.} \tag{22}$$

Indeed, the factorization theorem represents any element F of H^2 as the product of a “Blaschke product,” a “singular function” and an “outer function.” Here we can omit the definitions of all three of these since the Blaschke product is defined as a product over the zeros of F in \mathcal{R} and hence is equal to 1 if F has no such zeros, and the singular function factor is also constant because $\tilde{G}(s)$ both continues across the line $\Re(s) = \frac{1}{2}$ and has slow rate of convergence to zero as $s \rightarrow \infty$ on the positive reals (cf. 3.14 in [1] and Chapter II,

Theorem 6.3, of [8]). This leaves only the outer function factor of $\tilde{G}(s)$, and here we can apply a corollary of a theorem of Beurling on shift-invariant subspaces in H^2 ([4], [11, pp. 99–101]), which tells us that a function $F \in H^2$ is outer if and only if the space

$$\mathcal{W}(F) := \mathbb{C}[1/s]F(s) = \langle s^{-n}F(s) \mid n \geq 0 \rangle$$

is dense in H^2 . (Notice that $\mathcal{W}(F)$ is contained in H^2 because the definition of the H^2 norm implies that $\|s^{-1}F(s)\|_{H^2} \leq 2\|F\|_{H^2} < \infty$ for $F \in H^2$, so $s^{-1}H^2 \subset H^2$.)

It follows that GRH implies that $\mathcal{W}(\tilde{G})$ is dense in H^2 . We can now apply the inverse Mellin transform from H^2 to L^2 to obtain the isometrically equivalent picture in L^2 . An easy calculation shows that $s^{-1}\tilde{f}(s)$ is the Mellin transform of If for any $f \in L^2$, where

$$If(x) := \int_x^\infty f(t) \frac{dt}{t},$$

and hence that the space $\mathcal{W}(\tilde{f})$ is the Mellin transform of the space

$$W(f) := \mathbb{C}[I]f = \langle I^n f \mid n \geq 0 \rangle \quad (f \in L^2).$$

It is not actually obvious that I preserves the space L^2 , but this follows from the Paley–Wiener theorem together with the fact just shown that multiplication by s^{-1} preserves H^2 . A more direct argument is that the operator I is the adjoint A^* of the averaging operator

$$Ah(x) = \frac{1}{x} \int_0^x h(t) dt,$$

as is easily checked, and since A is a bounded operator on $L^2(0, \infty)$ (cf. [10, Example 5.4, p. 23]), its adjoint I is also bounded on $L^2(0, \infty)$. But I also preserves the condition of having support in $(0, 1)$, so it is a bounded operator on L^2 .

As a consequence implication (22) translates on the L^2 side to the statement

$$L(s) \text{ has no zeros in } \sigma > \frac{1}{2} \implies W(G) \text{ is dense in } L^2. \tag{23}$$

Proposition 7 is then obtained by applying to $f = G$ the following lemma, in which $\mathcal{V}(f)$ for any $f \in L^2$ denotes the span of the multiplicative translates $\tau_a f$ ($0 < a \leq 1$).

Lemma For $f \in L^2$ we have $W(f) \subset \overline{\mathcal{V}(f)}$.

Proof We first note that it suffices just to show that $If \in \overline{\mathcal{V}(f)}$ for all $f \in L^2$, because then $I^n f \in \overline{\mathcal{V}(f)}$ for all $n \geq 1$ by induction and the remark that $g \in \overline{\mathcal{V}(f)} \implies \overline{\mathcal{V}(g)} \subseteq \overline{\mathcal{V}(f)}$. Since $(W^\perp)^\perp = \overline{W}$ for all subspaces W , we need to show that

$$\langle h, \tau_a f \rangle = 0 \quad \text{for all } a \in (0, 1] \implies \langle If, h \rangle = 0. \tag{24}$$

But since $a\tau_{a^{-1}}$ is the adjoint of τ_a and A is the adjoint of I , this statement just says that any function that is orthogonal to all $h(ax)$ with $0 < a \leq 1$ is also orthogonal to Ah , and this is obvious because $Ah(x) = \int_0^1 h(ax) da$. □

4 Analytic properties of $C(x)$ and $H(x)$

In this section we study the analytic properties of $C(x)$ and $H(x)$, discussing in particular the question of their extendability from \mathbb{Q} to \mathbb{R} (which we answer completely only in the case of $C(x)$) and describing the asymptotic properties in neighborhoods of rational points needed to establish Theorem 2 and its corollary (quantum modularity). We also discuss the relation of these two functions to the Dirichlet series $L(s)$. It should be mentioned in passing that much of the material in this section is similar in spirit to the papers [2, 3] of Bettin and Conrey, with χ replaced by the trivial character and $L(s)$ by the Riemann zeta function.

4.1 Sum and integral representations of $C(x)$ and $H(x)$

Proposition 8 gives a number of formulas for these two functions. However, the word “function” has a slightly different meaning in the two cases. We have already defined them both as functions on \mathbb{Q} . Part (i) of the proposition says that C extends continuously to \mathbb{R} , and parts (iv) and (v) give further properties of that function. But in the case of H we do not know to what extent it can be defined as a function on \mathbb{R} . (This question is discussed in Sect. 4.2, though not in great detail since this is not our main interest.) However, it can certainly be defined as a distribution on \mathbb{R} , in fact even as the derivative of a continuous function, since by integrating (10) we get the absolutely convergent sum representation

$$\int_0^x H(x') dx' = \frac{\pi x}{8} - \frac{1}{2} \sum_{k=1}^{\infty} \frac{\chi(k)}{k^2} A\left(\frac{kx}{2}\right)$$

for its integral, where A is the periodic continuous function on \mathbb{R} defined by $A(n + \epsilon) = (-1)^n \epsilon$ for $n \in \mathbb{Z}$ and $|\epsilon| \leq \frac{1}{2}$. Parts (ii) and (iii) of the proposition are then to be interpreted in the sense of distributions, one giving the Fourier expansion of the even periodic distribution H and the other its relation, already mentioned in introduction, with the derivative of C .

Proposition 8 (i) *The function $C(x)$ extends continuously from \mathbb{Q} to \mathbb{R} via the formula*

$$C(x) = \int_0^{\infty} S(t) S(xt) \frac{dt}{t^2} \quad (x \in \mathbb{R}). \tag{25}$$

(ii) *The distribution H has the Fourier expansion*

$$H(x) = \frac{\pi}{8} - \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\chi(n) d(n)}{n} \cos\left(\frac{\pi nx}{2}\right), \tag{26}$$

where $d(n)$ denotes the number of positive divisors of n .

(iii) *The function $C(x)$ and the distribution $H(x)$ are related by the equation*

$$C'(x) = H(1/x) \quad (x > 0). \tag{27}$$

(iv) *The function $C(x)$ is also given by the sum*

$$C(x) = \frac{\pi}{8} + x \sum_{n=1}^{\infty} \chi(n) d(n) J\left(\frac{\pi nx}{2}\right) \quad (x > 0), \tag{28}$$

where $J(x)$ denotes the modified cosine integral

$$J(x) = - \int_x^{\infty} \frac{\cos(t)}{t^2} dt \quad (x \in \mathbb{R}_{>0}). \tag{29}$$

(v) *The Mellin transform of $C(x)$ is given by*

$$\tilde{C}(s) = - \frac{L(-s) L(s+1)}{s(s+1)} \quad (-1 < \sigma < 0), \tag{30}$$

where $L(s)$ is the analytic continuation of the L -series defined in (1).

Proof Formula (25) for $x \in \mathbb{Q}_{>0}$ is precisely the statement of Proposition 1 after an obvious change of variables, and the convergence and continuity of the integral for all x in $\mathbb{R}_{>0}$ are easy consequences of the facts that S is supported on $[1, \infty)$ and is locally constant

and bounded. To prove (ii), we insert the (standard) Fourier expansion

$$S(x) = \frac{1}{2} - \frac{2}{\pi} \sum_{m=1}^{\infty} \frac{\chi(m)}{m} \cos\left(\frac{\pi mx}{2}\right), \tag{31}$$

of the even periodic function $S(x)$ into sum (10), combine the double sum into a single one and use Leibniz’s formula $L(1) = \pi/4$. For (iii), we insert (11) into (25) to get

$$C(x) = \int_1^{\infty} \sum_{0 < k < xt} \chi(k) S(t) \frac{dt}{t^2} = \sum_{k=1}^{\infty} \chi(k) \int_{k/x}^{\infty} \frac{S(t)}{t^2} dt, \tag{32}$$

from which Eq. (27) follows immediately by differentiating and using (10). (The easy justification of these steps in the sense of distributions by integrating against smooth test functions is left to the reader.) Finally, to prove (iv) we combine (27) and (6) to get

$$-x^2 \frac{d}{dx} \left(\frac{C(x)}{x} \right) = C(x) - x C'(x) = C(x) - x H\left(\frac{1}{x}\right) = H(x)$$

and then insert Fourier development (26) and integrate term-by-term. This gives (28) up to an additive term λx that can be eliminated by noting that C is bounded. Note that the sum in (28) is uniformly and absolutely convergent since a simple integration-by-parts argument shows that $J(x) = O(x^{-2})$ as $x \rightarrow \infty$. Finally, for (v) we note first that the Mellin transform $\tilde{S}(s) := \int_0^{\infty} S(x) x^{s-1} dx$ of S is given according to Eq. (21) by $\tilde{S}(s) = -L(-s)/s$ for $\sigma = \Re(s) < 0$, and then use (25) to get

$$\tilde{C}(s) = \int_0^{\infty} S(t) \frac{\tilde{S}(s)}{t^s} \frac{dt}{t^2} = \tilde{S}(s) \tilde{S}(-s-1) = -\frac{L(-s)L(s+1)}{s(s+1)}$$

for $-1 < \sigma < 0$, as claimed. (The Mellin transform exists in this strip because $C(x) = O(\min(|x|, 1))$ as an easy consequence of (25).) In particular, $\tilde{C}(s)$ extends meromorphically to all of \mathbb{C} and is invariant under $s \mapsto -s - 1$. We also note that by the well-known functional equation of $L(s)$, Eq. (30) can be written in the alternative form

$$\tilde{C}(s) = -\frac{(2/\pi)^{s+1} \Gamma(s) \cos(\pi s/2)}{s+1} L(s+1)^2.$$

We can then use this to give a second derivation of (28), not making use of the distribution H , by first applying the Mellin inversion formula to write $C(x)$ as an integral over a vertical line $\sigma = c$ with $-1 < c < 0$ and then shifting the path of integration to the right, picking up the term $\pi/8$ in (28) from the residue at $s = 0$ and permitting us to use the convergent Dirichlet series representation $L(s)^2 = \sum_{n=1}^{\infty} \chi(n)d(n)n^{-s}$ for $\sigma > 1$. (We omit the details of this calculation.) □

4.2 The function H on the real line

In this subsection we give two one-parameter generalizations of $H(x)$ and discuss possible ways to define this function at irrational arguments.

We first generalize the definitions of $H(x)$ and $C(x)$ and the assertions of Proposition 8 to families depending on a complex parameter s . Specifically, we can generalize (10) to

$$H_s(x) := \sum_{k=1}^{\infty} \frac{\chi(k)}{k^s} S(kx) \quad (x \in \mathbb{R}, \sigma = \Re(s) > 1), \tag{33}$$

which now converges absolutely. The same calculation as for $H(x)$ then gives the Fourier expansion

$$H_s(x) = \frac{L(s)}{2} - \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\chi(n) \sigma_{s-1}(n)}{n^s} \cos\left(\frac{\pi nx}{2}\right),$$

where $L(s)$ is the Dirichlet L -function defined in (1) and σ_v is the sum-of-divisors function $\sigma_v(n) = \sum_{k|n, k>0} k^v$. Similarly, we define $C_s(x) = x^s C_s(1/x)$ for all $s \in \mathbb{C}$ with $\sigma > 0$ by

$$C_s(x) = s \int_0^\infty S(t) S(xt) \frac{dt}{t^{s+1}} = \sum_{j, k > 0} \frac{\chi(j) \chi(k)}{\max(j/x, k)^s}, \tag{34}$$

where the equality of the two expressions is proved by the same calculation as for the proof of (25). Then $C_s(x)$ and $H_s(x)$ are related by

$$C_s(x) = H_s(x) + x^s H_s\left(\frac{1}{x}\right), \quad C'_s(x) = s x^{s-1} H_s\left(\frac{1}{x}\right) \quad (x > 0) \tag{35}$$

by the same calculations as in the special case $s = 1$ and one has an expansion like (28) but with $J(x)$ replaced by $-\int_x^\infty t^{-s-1} \cos(t) dt$. As long as $\sigma > 1$ all of the sums and integrals involved are absolutely convergent and all steps are justified.

A different possible way to regularize $H(x)$ for $x \in \mathbb{R}$ is to define

$$H(x) = \lim_{\varepsilon \searrow 0} \left(\frac{8}{\pi} - \frac{2}{\pi} T(x, \varepsilon) \right), \quad \text{where} \quad T(x, \varepsilon) = \sum_{n=1}^\infty \frac{\chi(n) d(n)}{n} e^{-n\varepsilon} \cos(n\pi x/2). \tag{36}$$

Using the identity $\sum_{n>0} \frac{d(n)}{n} x^n = \sum_{k>0} \frac{1}{k} \log\left(\frac{1}{1-x^k}\right)$ and standard trigonometric identities, we find after a short calculation that

$$T(x, \varepsilon) = \frac{1}{2} \sum_{k=1}^\infty \frac{\chi(k)}{k} \arctan\left(\frac{\cos(k\pi x/2)}{\sinh(k\varepsilon)}\right).$$

Both this series and the original one defining T converge exponentially fast for any positive ε , giving us another possible approach to the analytic properties of the limiting function H .

Summarizing this discussion, we have at least five potential definitions of $H(x)$ for $x \in \mathbb{R}$:

Definition 1 Define $H(x)$ by series (10), if this sum converges.

Definition 2 Define $H(x)$ by Fourier series (26), if this series converges.

Definition 3 Define $H(x)$ as the limit for $s \searrow 1$ of series (33), if this limit exists.

Definition 4 Define $H(x)$ as the limit in (36), if this limit exists.

Definition 5 Define $H(x)$ as the limit of $H(x')$ as $x' \in \mathbb{Q}$ tends to x , if this limit exists.

We can then ask—but have not been able to answer—the question whether any or all of these definitions converge for irrational values of x or, if they do, whether they give the same value. The last definition is in the sense the strongest one, since any definition of H on \mathbb{R} or a subset of \mathbb{R} that agrees with the original definition on \mathbb{Q} must coincide with the value in Definition 5 at any argument x at which this function is continuous. In any case, we pose the explicit question:

Question Are the five definitions given above convergent and equal to one another for all irrational values of x , or for all x belonging to some explicit set of measure 1?

It seems reasonable to expect the answer to the second question to be affirmative with the set of measure 1 being the complement of some set of irrational numbers having extremely good rational approximations, like the well-known “Brjuno numbers” in dynamics.

4.3 Asymptotic behavior of $H(x)$ and $C(x)$ near rational points

The results of Sect. 4.1 prove the continuity of $C(x)$, which is part of the statement of Theorem 2 in Introduction. We now discuss the remaining statements there, concerning the behavior of C and H near rational points.

We begin by looking numerically at the asymptotic properties of $C(x)$ near $x = 1$. Computing the values of $C(1 \pm \frac{1}{n})$ for $1 \leq n \leq 1000$ and using a numerical interpolation technique that is explained elsewhere (see, e.g., [9]), we find empirically an expansion of the form

$$C\left(1 \pm \frac{1}{n}\right) \sim c_0 + \frac{1}{n} \left(-\frac{1}{4} \log n + c_1^\pm\right) \pm \frac{c_2}{n^2} + \frac{c_3}{n^3} \pm \frac{c_4}{n^4} + \frac{c_5}{n^5} \pm \dots \tag{37}$$

with the first few coefficients having numerical values given by

$$c_0 = C(1) = \frac{\pi}{4}, \quad c_1^+ = c_1^- + c_0 = -0.23528274848426799887\dots, \quad c_2 = -\frac{1}{8},$$

$$c_3 = 0.058801396529669\dots, \quad c_4 = \frac{1}{48} - c_3, \quad c_5 = 0.01927655829\dots$$

These numerics become clearer if we work with $H(x)$ instead, where we find the simpler expansion

$$H\left(1 \pm \frac{1}{n}\right) \sim \pm \left(\frac{1}{4} \log n + h_0^\pm + \frac{h_2}{n^2} + \frac{h_4}{n^4} + \dots\right) \tag{38}$$

with no odd powers of $1/n$ and with coefficients given numerically by

$$h_0^+ = h_0^- + c_0 = 0.7706809118817\dots, \quad h_2 = 0.01713472986\dots,$$

$$h_4 = -0.00345272385\dots,$$

which then gives expansion (37) with $c_1^\pm = -h_0^\mp - \frac{1}{4}$, $c_3 = h_2 + \frac{1}{24}$, and more generally $c_r = (-1)^r \sum_{0 < i < r/2} \binom{r-1}{2i-1} h_{2i}$ for $r > 1$. Moreover, on calculating the next few values of h_{2i} numerically to high precision we are able to guess the closed formulas

$$h_0^\pm = \pm \frac{\pi}{8} + \frac{\gamma}{4} + \frac{1}{4} \log \frac{8}{\pi}, \quad h_{2i} = (-1)^{i-1} \frac{(2^{2i-1} - 1)^2 B_{2i}^2}{2i(2i)!} \pi^{2i} \quad (i \geq 1) \tag{39}$$

for the coefficients in (38), where γ is Euler’s constant and B_n the n th Bernoulli number. In fact, it is not difficult to prove (38) [and hence also (37)] with these values of h_i by applying a twisted Euler–Maclaurin formula (giving the asymptotics as $\varepsilon \rightarrow 0$ of sums over an interval of $\chi(n)f(n\varepsilon)$ for a smooth function f) to the closed formula

$$H\left(1 \pm \frac{1}{n}\right) = \frac{4n}{\pi} \sum_{\substack{0 < k < 2n \\ k \equiv 1 \pmod{4}}} \cot \frac{\pi k}{4n}$$

which follows easily from (14), though we will not carry this out here.

The “log n ” terms in Eqs. (37) and (38) already show that $C(x)$ is not differentiable at $x = 1$ and $H(x)$ is not continuous at $x = 1$, but from these formulas one might imagine that the functions $C(x) - |x - 1| \log(|x - 1|)/4$ and $H(x) - \log(|x - 1|)/4$ are C^∞ to both the right and the left of this point. However, this is not the case, because both Eqs. (37) and (38) are valid only when n is an integer and change in other cases. For instance, if n tends to infinity in $\mathbb{N} + \frac{1}{2}$ rather than \mathbb{N} , then $H(1 \pm 1/n)$ has an expansion of the same form as (38) and with the same constants h_0^\pm , but with $h_2 = \pi^2/576$ replaced by $-7\pi^2/1152$, $h_4 = -49\pi^4/1,382,400$ replaced by $127\pi^4/2,764,800$, etc. This statement, which again

can be proved using an appropriate twisted version of the Euler–Maclaurin formula, is a typical phenomenon of quantum modular forms.

If we look at the asymptotics of $C(x)$ and $H(x)$ as x tends to any rational number α with odd numerator and denominator, then we find a similar behavior, which is not surprising since any such α is Γ_ϑ -equivalent to 1 and both H and C have transformation properties, modulo functions with better smoothness properties, under the action of Γ_ϑ . More precisely, if we write $\alpha = a/c$ with a and $c > 0$ odd and coprime and complete $\begin{pmatrix} a & \\ c & \end{pmatrix}$ to a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$, then we find asymptotic expansions

$$H\left(\frac{an+b}{cn+d}\right) \sim \pm \frac{\log|cn+d|}{4c} + \sum_{i=0}^{\infty} \frac{h_{i,\alpha}^\pm}{(cn+d)^i},$$

$$C\left(\frac{an+b}{cn+d}\right) \sim C\left(\frac{a}{c}\right) \pm \frac{\log|cn+d|}{4(cn+d)} + \sum_{i=1}^{\infty} \frac{c_{i,\alpha}^\pm}{(cn+d)^i}$$

as $n \rightarrow \pm\infty$ with $n \in \mathbb{Z}$, and similar expansions with other coefficients $h_{i,\alpha,\beta}^\pm$ and $c_{i,\alpha,\beta}^\pm$ if $n \rightarrow \pm\infty$ with $n \in \mathbb{Z} + \beta$ for some fixed rational number β .

Similar statements hold for x tending to rational numbers α with an even numerator or denominator, but now without the log term, the simplest case being

$$C\left(\frac{1}{n}\right) \sim \frac{\pi}{8n} + \chi(n) \frac{\pi^2}{16n^2} + \chi(n-1) \frac{\pi^3}{64n^3} + \dots$$

$$+ \frac{\chi(n-k+1)}{4} \frac{A_k^2}{k!} \left(\frac{\pi}{2n}\right)^{k+1} + \dots$$

where A_k is the number of “up-down” permutations of $\{1, \dots, k\}$ (permutations π where $\pi(i) - \pi(i-1)$ has sign $(-1)^i$ for all i), which is also equal to the coefficient of $x^k/k!$ in $\tan x + \sec x$. (Note that $A_{k-1} = (4^k - 2^k)|B_k|/k$ for k even, so that the coefficients here are closely related to those in (39).) The difference between the two cases is due to the fact that the action of Γ_ϑ on $\mathbb{P}^1(\mathbb{Q})$ has two orbits (cusps), and comes from the relation that we will see in the final section between the quantum modularity properties of C and H and a specific Maass modular form on Γ_ϑ that is exponentially small at the cusp ∞ but has logarithmic growth at the cusp 1.

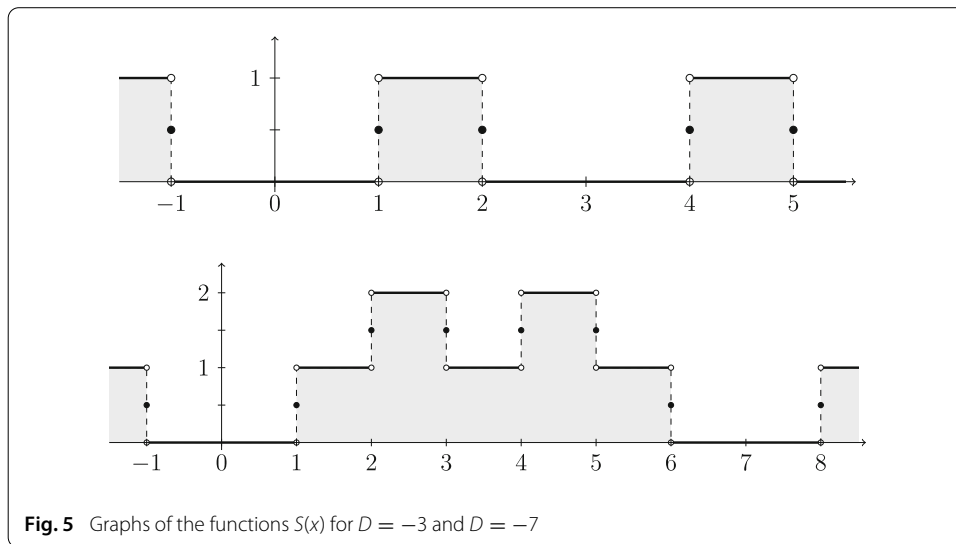
5 Generalization to other odd Dirichlet characters

In this section we discuss the changes that are needed when the character defined in (1) is replaced by an arbitrary odd primitive Dirichlet character. Suppose that $D < 0$ is the discriminant of an imaginary quadratic field K (or equivalently, that D is either square-free and congruent to 1 mod 4 or else equal to 4 times a square-free number not congruent to 1 mod 4), and let $\chi = \chi_D = \left(\frac{D}{\cdot}\right)$ be the associated Dirichlet character, the case studied up to now corresponding to $K = \mathbb{Q}(i)$ and $D = -4$. The Dirichlet L -series $L(s, \chi)$ is defined as $\sum_{n>0} \chi(n)n^{-s}$ for $\sigma > 1$ and by analytic continuation otherwise, and is equal to the quotient of the Dedekind zeta function of K by the Riemann zeta function. Its value at $s = 1$ is well known to be nonzero and given by

$$L(1, \chi) = \frac{\pi}{\sqrt{|D|}} h'(D),$$

where $h'(D)$ is $1/3$ or $1/2$ if $D = -3$ or $D = -4$ and otherwise is the class number of K .

We now define $S(x)$ by the same formula (11) as before, with $\chi = \chi_D$. It is still periodic (with period $|D|$) and even and hence bounded, but now has average value $h'(D)$ rather



than $1/2$ and also no longer takes on only the values 0 and 1 , as one sees in the pictures of the graphs of this function for $D = -3$ and $D = -7$ shown in Fig. 5.

We define $H(x) = H_D(x)$ and $C(x) = C_D(x)$ for $x \in \mathbb{Q}$ by the same formulas (8) and (10) as before, but we define $h_{m,n}$ and $c_{m,n}$ by (5) and (7) with a factor $|D|/\pi$ instead of $4/\pi$ in order to get integral linear combinations of cotangents. Fourier expansions (26) and (31) are then replaced by

$$S(x) = h'(D) - \frac{\sqrt{|D|}}{\pi} \sum_{m=1}^{\infty} \frac{\chi(m)}{m} \cos\left(\frac{2\pi mx}{|D|}\right), \tag{40}$$

$$H(x) = \frac{\pi h'(D)^2}{\sqrt{|D|}} - \frac{\sqrt{|D|}}{\pi} \sum_{n=1}^{\infty} \frac{\chi(n) d(n)}{n} \cos\left(\frac{2\pi nx}{|D|}\right), \tag{41}$$

and the formula corresponding to (14) now reads

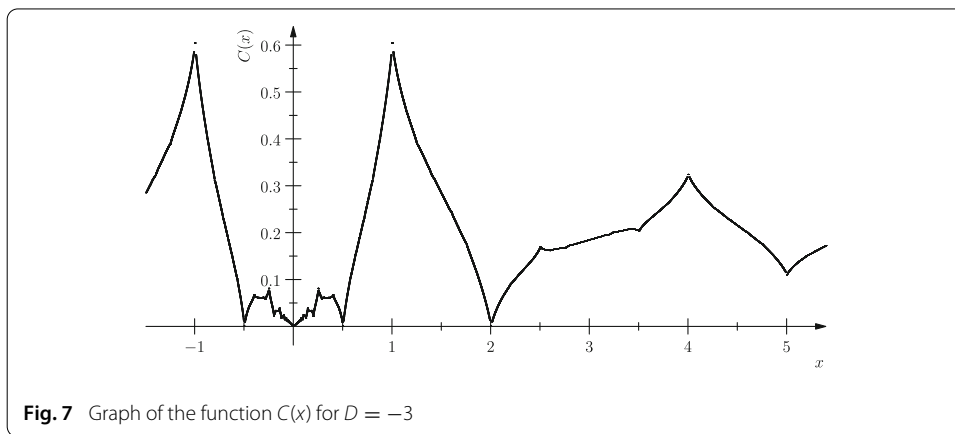
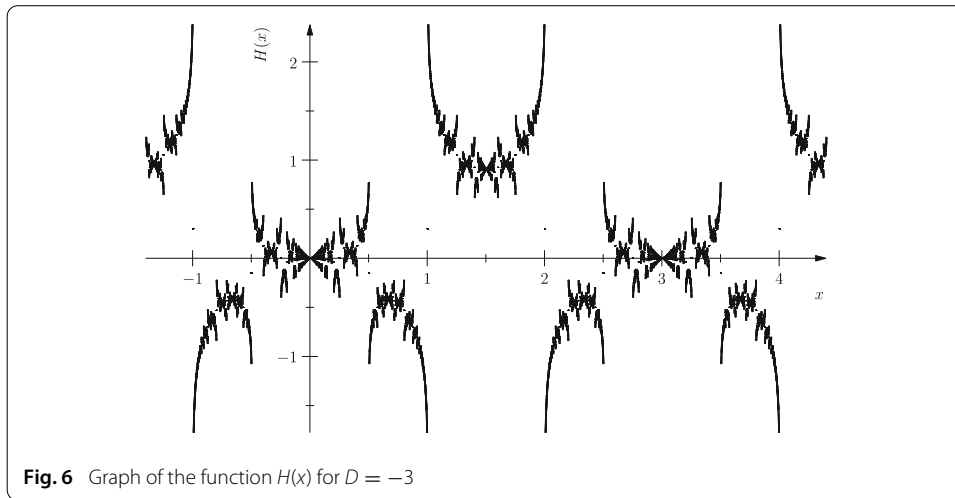
$$h_{m,n} = \sum_{0 < k < |D|n/2} \chi(k) S\left(\frac{km}{n}\right) \cot\left(\frac{\pi k}{|D|n}\right) \tag{42}$$

by the same calculations as before. Finally we also define the matrices C_N and \widehat{C}_N exactly as we did in the special case $D = -4$ (except for replacing the lower right-hand entry in \widehat{C}_N in (3) by $|D|N/\pi$), and then from the scalar product calculations

$$\langle \xi, \xi \rangle = 1, \quad \langle \xi, G_a \rangle = L(1) a, \quad \langle G_a, G_b \rangle = a C(a/b),$$

which are proved exactly as before, we deduce the same connection as in Theorems 1 and 3 between the unboundedness of the function $R(N)$ and the Riemann hypothesis for $L(s, \chi)$.

The only real difference with the case $D = -4$ is in the argument for quantum modularity. The function $H(x)$ on \mathbb{Q} is still even and periodic of period $|D|$ (and also anti-periodic up to a constant with period $|D|/2$ if D is even), it again has discontinuities at infinitely many rational points by an argument similar to the one given in Sect. 4.3, and by the analog of Proposition 8 the function $C : \mathbb{Q} \rightarrow \mathbb{R}$ again has a continuous extension to \mathbb{R} and is therefore much better behaved analytically than $H(x)$, as illustrated by the following graphs of these functions for $D = -3$ (Figs. 6, 7), which look qualitatively much like their $D = -4$ counterparts in Figs. 2 and 3.



The difference is that this longer suffices to prove the quantum modularity of H because the subgroup of $SL(2, \mathbb{Z})$ generated by the matrices $T^{|D|}$ (or $T^{|D|/2}$ if D is even) and S now no longer has finite index. Instead, we need a statement like the corollary to Theorem 2 with Γ_ϑ replaced by the congruence subgroup $\Gamma_{(D)} = \Gamma_0(D) \cup S\Gamma_0(D)$ of $SL(2, \mathbb{Z})$ (or by $\Gamma_{(D/2)}$ if D is even, which is indeed Γ_ϑ if $D = -4$). This statement is given in the following theorem.

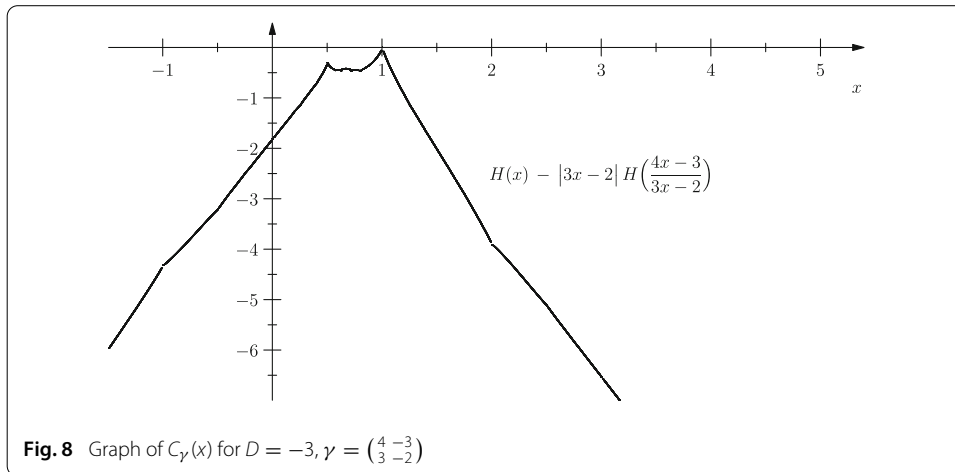
Theorem 4 *Let $D < 0$ and $H(x) = \sum_{k>0} \chi_D(k)S_D(kx)/k$ be as above. Then the function*

$$C_\gamma(x) := H(x) - \varepsilon(\gamma) |cx + d| H\left(\frac{ax + b}{cx + d}\right) \quad (x \in \mathbb{Q}) \tag{43}$$

extends continuously to \mathbb{R} for all matrices $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_{(D)}$, where $\varepsilon : \Gamma_{(D)} \rightarrow \{\pm 1\}$ is the homomorphism mapping $\Gamma_{(D)}$ to 1 and S to -1 .

This theorem is illustrated for $D = -3$ and a typical element of $\Gamma_{(-3)}$ in Fig. 8.

We will deduce Theorem 4 as a consequence of the following proposition, which is a generalization to arbitrary elements $\gamma \in \Gamma_{(D)}$ of Eq. (32) for $D = -4$ and $\gamma = S$.



Proposition 9 Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_{(D)}$ with $c \neq 0$. Then

$$C_\gamma(x) = \begin{cases} H(-d/c) - |c| \varepsilon(\gamma) \sum_{k=1}^{\infty} \chi(k) \int_{k\gamma x}^{\infty} \frac{S(t) dt}{(ct - ak)^2} & \text{if } x < -d/c, \\ H(-d/c) - |c| \varepsilon(\gamma) \sum_{k=1}^{\infty} \chi(k) \int_{-\infty}^{k\gamma x} \frac{S(t) dt}{(ak - ct)^2} & \text{if } x > -d/c. \end{cases} \tag{44}$$

To see that this implies the continuity, we note that we can rewrite the right-hand side by interchanging the summation and integration in the form [generalizing (25)]

$$C_\gamma(x) = H(-d/c) - |c| \varepsilon(\gamma) \int_{-\infty}^{\infty} S(t) \left(\sum_{k>0, k\gamma x \geq t} \frac{\chi(k)}{(ak - ct)^2} \right) dt \tag{45}$$

for $x \geq -d/c$. The right-hand side of this formula is continuous because $S(t)$ is piecewise continuous and bounded and has support in $|t| \geq 1$, while the second factor of the integrand is piecewise continuous and bounded by a constant times t^{-2} , as one sees easily by partial summation. Another argument is that (44) is equivalent to the equality of distributions

$$\begin{aligned} C'_\gamma(x) &= -\varepsilon(\gamma) |c| \operatorname{sgn}(x + d/c) \sum_{k=1}^{\infty} \chi(k) \frac{S(k\gamma x)}{k} \\ &= -\varepsilon(\gamma) c \operatorname{sgn}(cx + d) H\left(\frac{ax + b}{cx + d}\right), \end{aligned} \tag{46}$$

[generalizing (27)], and we already know that the distribution H is locally the derivative of a continuous function.

For the proof of (44) we will restrict to the case when the matrix γ belongs to $\Gamma_{(D)}$, which is sufficient for the proof of Theorem 4 because $\gamma \mapsto C_\gamma$ is a cocycle and we already know the continuity of C_γ ($= C$) for the special case $\gamma = S$ representing the nontrivial coset of $\Gamma_{(D)}$ in $\Gamma_{(D)}$. To avoid distracting case distinctions, we consider only the case $x + d/c < 0$ and $\gamma x > 0$. The argument is the same in principle in all other cases, but we are actually working with ordered tuples of points on the 1-manifold $\mathbb{P}^1(\mathbb{R})$, on which the group $\Gamma_{(D)}$ acts in an orientation-preserving way, and it is notationally simpler to fix the positions of the points occurring with respect to the point at infinity, so that we can work on \mathbb{R} instead.

We may also assume (by replacing γ by $-\gamma$ if necessary, but then remembering that γ may be congruent to $-\mathbf{1}_2$ rather than $\mathbf{1}_2$ modulo D) that $c > 0$. Finally, to be able to work with absolutely convergent sums and to reorder the terms freely, we will replace $H(x)$ by the function $H_s(x)$ ($s > 1$) introduced in Sect. 4.2 and C_γ by the corresponding cocycle

$$C_{\gamma,s}(x) = H_s(x) - \varepsilon(\gamma) |cx + d|^s H_s\left(\frac{ax + b}{cx + d}\right), \tag{47}$$

and only set $s = 1$ at the end. The equation we have to prove then becomes

$$C_{\gamma,s}(x) = H_s(-d/c) - cs \sum_{k=1}^{\infty} \chi(k) \int_{k\gamma x}^{\infty} \frac{S(t) dt}{(ct - ak)^{s+1}}. \tag{48}$$

We denote the last term (without the minus sign) by A . Since the variable t in the integral is always positive (because we are considering the case $\gamma x > 0$), we can replace $S(t)$ by its definition (11) and interchange the summation and integration to get

$$\begin{aligned} A &= \sum_{k,j>0} \chi(j) \chi(k) \int_{\max(j,k\gamma x)}^{\infty} \frac{cs dt}{(ct - ak)^{s+1}} = \sum_{k,j>0} \frac{\chi(j) \chi(k)}{\max(cj - ak, k/|cx + d|)^s} \\ &= \sum_{0 < j \leq k\gamma x}^* \frac{\chi(j) \chi(k)}{(k/|cx + d|)^s} + \sum_{k>0, j \geq k\gamma x}^* \frac{\chi(j) \chi(k)}{(cj - ak)^s}. \end{aligned}$$

In the second sum we replace the vector $\begin{pmatrix} j \\ k \end{pmatrix}$ by the vector $-\gamma \begin{pmatrix} j \\ k \end{pmatrix} = \begin{pmatrix} -aj - bk \\ -cj - dk \end{pmatrix}$. This does not change the product $\chi(j)\chi(k)$ because γ is congruent to plus or minus the identity modulo the period $|D|$ of χ , changes the expression $cj - ak$ in the denominator to k , and changes the inequalities $k > 0, j \geq k\gamma x$ to $kx \leq j < -kd/c$ (which imply $k > 0$). Hence

$$\begin{aligned} A &= |cx + d|^s \sum_{k=1}^{\infty} \frac{\chi(k)}{k^s} S(k\gamma x) + \sum_{k=1}^{\infty} \frac{\chi(k)}{k^s} (S(-kd/c) - S(kx)) \\ &= H_s(-d/c) - C_{\gamma,s}(x) \end{aligned}$$

completing the proof of Eq. (48) and hence of the proposition and theorem.

6 Modular forms are everywhere

As one can already see from the examples in the original article [16] where this notion was introduced, quantum modular forms are sometimes related to actual modular forms. These modular forms may be either holomorphic or Maass forms, with the quantum modular form in the latter case being related to the “periods” of Maass forms in the sense developed in [7, 13]. The quantum modular form $H(x)$ that we have been studying in this paper turns out to be of this latter type, with the associated Maass form being an Eisenstein series with eigenvalue $1/4$ for the hyperbolic Laplace operator. In this final section we explain how this works, first for the case $D = -4$ studied in the first four sections of this paper. In that case the relevant modular group $\Gamma_{\mathfrak{y}}$ is generated by a translation and an inversion, so that for both the period theory of the Maass form u and the quantum modularity of H one needs only the functional equation of the associated L -series. Then at the end we indicate how the quantum modularity of H for general D follows from the full theory of periods of Maass forms.

We start with the case $D = -4$, so that χ is the character given by (1). The relevant modular form here is the Maass Eisenstein series

$$u(z) = y^{1/2} \sum_{n=1}^{\infty} \chi(n) d(n) K_0\left(\frac{\pi ny}{2}\right) \sin\left(\frac{\pi nx}{2}\right) \quad (z = x + iy \in \mathfrak{H}), \tag{49}$$

where \mathfrak{H} denotes the upper half-plane and K_0 the usual K -Bessel function of order 0. This is an eigenfunction with eigenvalue $1/4$ with respect to the hyperbolic Laplace operator $\Delta = -y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)$ and is a modular function with character ε for the group Γ_ϑ defined in Sect. 1, meaning that $u(\gamma z) = \varepsilon(\gamma)u(z)$ for all $\gamma \in \Gamma_\vartheta$ or, more explicitly, that

$$u(z + 2) = -u(z) = u(-1/z) \quad (z \in \mathfrak{H}). \tag{50}$$

To see this, we observe that $u(z)$ is proportional to $E\left(\frac{z+1}{4}, \frac{1}{2}\right) - E\left(\frac{z-1}{4}, \frac{1}{2}\right)$, where $E(z, s)$ is the usual nonholomorphic Eisenstein series of weight 0 and eigenvalue $s(1-s)$ for the Laplace operator with respect to the full modular group $\Gamma_1 = \text{SL}(2, \mathbb{Z})$. (This follows easily from the well-known Fourier expansion of $E(z, \frac{1}{2})$ as a linear combination of the three functions \sqrt{y} , $\sqrt{y} \log y$ and $\sqrt{y} \sum_{n \neq 0} d(n)K_0(2\pi|n|y)e^{2\pi inx}$.) Transformation equations (50) then follow from the $\text{SL}(2, \mathbb{Z})$ -invariance of $E(z, \frac{1}{2})$, the first one trivially since $E(z + 1, s) = E(z, s)$ and the second by using the invariance of $E(z, s)$ under $\begin{pmatrix} 1 & 0 \\ \mp 4 & 1 \end{pmatrix}$ to get

$$E\left(\frac{-1/z \pm 1}{4}, \frac{1}{2}\right) = E\left(\frac{(-1/z \pm 1)/4}{\mp 4((-1/z \pm 1)/4) + 1}, \frac{1}{2}\right) = E\left(\frac{z \mp 1}{4}, \frac{1}{2}\right).$$

We now associate to $u(z)$ the periodic holomorphic function f on $\mathbb{C} \setminus \mathbb{R} = \mathfrak{H}^+ \cup \mathfrak{H}^-$ (where $\mathfrak{H}^\pm = \{z \in \mathbb{C} \mid \pm \Im(z) > 0\}$) having the same Fourier coefficients as u , i.e.,

$$f(z) = \sum_{n=1}^{\infty} \chi(n) d(n) q^{\pm n/4} \quad (z \in \mathfrak{H}^\pm, q = e^{2\pi iz}). \tag{51}$$

Proposition 10 *The period function $\psi(z)$ defined by*

$$\psi(z) = f(z) + \frac{1}{z} f\left(-\frac{1}{z}\right) \quad (z \in \mathbb{C} \setminus \mathbb{R}) \tag{52}$$

extends holomorphically from $\mathbb{C} \setminus \mathbb{R}$ to $\mathbb{C}' = \mathbb{C} \setminus (-\infty, 0]$.

Proof We follow the proof of the corresponding result given in Chapter 1 of [13] for Maass cusp forms on the full modular group Γ_1 . (See Theorem on p. 202 of [13], which also gives a converse statement characterizing cusp forms in terms of holomorphic functions ψ on \mathbb{C}' satisfying a certain functional equation.) That proof required only the functional equation of the L -series associated with the Maass cusp form, which worked because the group Γ_1 is generated by the translation T and the inversion S , and can be applied here because $\Gamma_\vartheta = \langle S, T^2 \rangle$ has a similar structure. Our situation is also a little different because our function $u(z)$ is an Eisenstein series rather than a cusp form, but since Fourier expansion (49) of $u(z)$ at infinity has no constant term this has no effect on the proof.

The argument in [13] was first to write the L -series of the Maass form $u(z)$ as the Mellin transform of the restriction of u (or of its normal derivative in the case of an odd cusp form) to the imaginary axis, multiplied by a suitable gamma factor, and to deduce from this relationship and the S -invariance of u a functional equation for the L -series. One then observed that the Mellin transforms of the restrictions to the positive or negative imaginary axis of both the associated periodic holomorphic function f and the associated period function ψ were also equal, up to different gamma factors, to the same L -series, and the functional equation of this L -series combined with Mellin inversion then led to a formula for ψ that applied in all \mathbb{C}' rather than just on $\mathbb{C} \setminus \mathbb{R}$. Here the first step, which would require the normal derivative since the restriction of function (49) to $i\mathbb{R}_+$ vanishes, can be skipped since the L -series $L(u; s) = \sum \chi(n)d(n)n^{-s}$ is simply the square of

Dirichlet L -series (1) and hence has a known functional equation. The Mellin transforms of the restrictions to f and ψ to the positive or negative imaginary axis are then given by

$$\begin{aligned} \tilde{f}_{\pm}(s) &:= \int_0^{\infty} f(\pm iy) y^{s-1} dy = \frac{\Gamma(s)}{(\pi/2)^s} L(s)^2 \quad (\Re(s) > 0), \\ \tilde{\psi}_{\pm}(s) &= \int_0^{\infty} \psi(\pm iy) y^{s-1} dy = \int_0^{\infty} \left(f(\pm iy) \mp \frac{i}{y} f(\pm i/y) \right) y^{s-1} dy \\ &= \tilde{f}_{\pm}(s) \mp i \tilde{f}_{\pm}(1-s) = \frac{\Gamma(s) L(s)^2}{(\pi/2)^s \cos(\pi s/2)} e^{\mp i\pi s} \quad (0 < \Re(s) < 1), \end{aligned}$$

where in the last line we have used the functional equation of $L(s)$ and standard identities for gamma functions. By the Mellin inversion formula we deduce that

$$\psi(\pm iy) = \frac{1}{2\pi i} \int_{\Re(s)=c} \frac{\Gamma(s) L(s)^2}{(\pi/2)^s \cos(\pi s/2)} (\pm iy)^{-s} ds \quad (y > 0)$$

for any $c \in (0, 1)$, and analytic continuation from $i\mathbb{R} \setminus \{0\}$ to $\mathbb{C} \setminus \mathbb{R}$ then gives the formula

$$\psi(z) = \frac{1}{2\pi i} \int_{\Re(s)=c} \left[\frac{\Gamma(s) L(s)^2}{(\pi/2)^s \cos(\pi s/2)} \right] z^{-s} ds$$

for all $z \in \mathbb{C} \setminus \mathbb{R}$. But the factor in square brackets is bounded by a power of s times $e^{-\pi|s|}$ for $|s| \rightarrow \infty$ on the vertical line $c + i\mathbb{R}$, so the integral on the right-hand side of this equation is absolutely convergent for all $z \in \mathbb{C}$ with $|\arg(z)| < \pi$, i.e., for all $z \in \mathbb{C}'$. \square

We now show how to connect the functions $f(z)$ and $\psi(z)$ to the functions $H(x)$ and $C(x)$, respectively, and to deduce the continuity of C —and hence the quantum modularity of H —from Proposition 10 for the period function ψ . The first step is easy, since from Fourier expansions (51) and (26) and standard integral formulae we find the relation

$$f(z) \doteq \int_{-\infty}^{\infty} \frac{H(t)}{(t-z)^2} dt \quad (z \in \mathbb{C} \setminus \mathbb{R})$$

between the periodic function $f(z)$ and the periodic distribution $H(x)$. (Here and in what follows the symbol \doteq denotes equality up to easily computed scalar factors whose values are irrelevant for the argument.) Replacing t by $1/t$ in the integral, we obtain

$$f(z) \doteq \int_{-\infty}^{\infty} \frac{C'(t)}{(1-tz)^2} dt \doteq z \int_{-\infty}^{\infty} \frac{C(t)}{(1-tz)^3} dt, \tag{53}$$

where we have used (27) to get the first equality and integration by parts for the second. But now replacing t by $1/t$ and using the functional equation $C(t) = |t|C(1/t)$ we find

$$\frac{1}{z} f\left(-\frac{1}{z}\right) = \frac{1}{z} f\left(\frac{1}{z}\right) \doteq -z \int_{-\infty}^{\infty} \frac{\operatorname{sgn}(t) C(t)}{(1-tz)^3} dt \tag{54}$$

with the same proportionality constant as in (53), and adding these two equations gives the relation

$$\psi(z) \doteq z \int_{-\infty}^0 \frac{C(t)}{(1-tz)^3} dt \doteq z \int_{-\infty}^0 \frac{C(t)}{(t-z)^3} dt \quad (z \in \mathbb{C} \setminus \mathbb{R}) \tag{55}$$

between the period function of the Maass waveform u and the function $C(x)$. This establishes the desired connection between the continuity of $C(x) = H(x) + |x|H(1/x)$, which expresses the quantum modularity of H , and Proposition 10, which expresses the modularity of u : in one direction, if we know that $C(x)$ is continuous (and bounded by $\min(1, |x|)$),

then (55) immediately gives the analytic continuation of $\psi(z)$ from $\mathbb{C} \setminus \mathbb{R}$ to \mathbb{C}' , and conversely, the inversion formula for the Stieltjes transform given in [14] lets us invert (55) to get

$$C(t) \doteq t \int_{\mathcal{C}} \psi(tz) \frac{1+z}{z} dz, \tag{56}$$

where \mathcal{C} is any contour with endpoints at $z = -1$ which encloses the origin, so that the continuity of C is a direct consequence of the holomorphy of ψ in \mathbb{C}' .

We observe that the entire discussion given here applies in an essentially unchanged form to the more general functions C_s and H_s discussed in Sect. 4.2, with Maass form (49) replaced by the form $E\left(\frac{z+1}{4}, \frac{s}{2}\right) - E\left(\frac{z-1}{4}, \frac{s}{2}\right)$ with spectral parameter $s/2$ instead of $1/2$.

Finally, we consider the case when the character χ in (1) is replaced by an arbitrary primitive odd Dirichlet character χ_D . Then all of the above calculations still go through: the function u is defined by Eq. (49) with χ replaced by χ_D and $\pi/2$ by $2\pi/|D|$, which is again a Maass form (with spectral parameter $\frac{1}{2}$ and character ε on the group $\Gamma_{(D)}$) because it is proportional to $\sum_{r \pmod{D}} \chi(r) E\left(\frac{z+r}{|D|}, \frac{1}{2}\right)$; the associated periodic function f and period function ψ are defined by (51) (with the new χ and with $q^{n/4}$ replaced by $q^{n/|D|}$) and (52) (with no change at all); and Proposition 10 remains true with the same proof. The difference, however, is that this proposition is no longer equivalent to the modularity of u , but only to its invariance (up to sign) under the transformations S and T^D (or $T^{D/2}$ if D is even), which in general generate a subgroup of $\Gamma_{(D)}$ of infinite order, as already discussed in Sect. 5. To get the full modularity (of this u or any other potential Maass form on $\Gamma_{(D)}$), we need to generalize Proposition 10 to the statement that for any matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_{(D)}$ the function

$$\psi_{\gamma}(z) = f(z) - \frac{\varepsilon(\gamma)}{cz + d} f\left(\frac{az + b}{cz + d}\right)$$

extends holomorphically from $\mathbb{C} \setminus \mathbb{R}$ to $\mathbb{C} \setminus (-\infty, -d/c]$ if $c > 0$ or $\mathbb{C} \setminus [-d/c, \infty)$ if $c < 0$. This statement can be proved in several ways and can be linked by a discussion similar to the one above to the continuity property of $C_{\gamma}(x)$ stated in Theorem 4. But the simplest approach is to relate the functions $C_{\gamma}(x)$ directly to the invariant distribution associated with u in the sense developed in [7, 12, 13]. Specifically, this theory says that an eigenfunction u of Δ with spectral parameter s and Fourier expansion $u(x + iy) = \sqrt{y} \sum_{n \neq 0} A_n K_{s-1/2}(\lambda|n|y) e^{i\lambda nx}$ is invariant (possibly with character) under the action of a Fuchsian group Γ if and only if the associated distribution $U(x) = \sum_{n \neq 0} |n|^{s-1/2} A_n e^{i\lambda nx}$ on $\mathbb{P}^1(\mathbb{R})$ is invariant (with the same character) with respect to the group action $(U|\gamma)(x) = |cx + d|^{-2s} U(\gamma x)$ for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$. (Here the word ‘‘distribution’’ must be interpreted correctly, namely, as a functional on the space of test functions $\phi(x)$ that are smooth on \mathbb{R} and for which $|x|^{-2s} \phi(1/x)$ is smooth near $x = 0$.) In our case the distribution U associated with u is given by $U(x) = \sum_{n=1}^{\infty} \chi(n) d(n) \sin(2\pi nx/|D|)$ and is related to distribution (41) by $U(x) \doteq H'(x)$. But that means that if we differentiate definition (43) of $C_{\gamma}(x)$ to get

$$C'_{\gamma}(x) = H'(x) - \frac{\varepsilon(\gamma)}{|cx + d|} H'(\gamma x) - \varepsilon(\gamma) |c| \operatorname{sgn}(x + d/c) H\gamma x,$$

then the first two terms on the right cancel and we recover Eq. (46).

Author details

¹Massachusetts Institute of Technology, Cambridge, MA, USA, ²Max Planck Institute for Mathematics, Bonn, Germany.

Acknowledgements

Open access funding provided by Max Planck Society.

Received: 19 May 2018 Accepted: 30 August 2018 Published online: 15 November 2018

References

1. Balazard, M., Saias, E.: The Nyman–Beurling equivalent form for the Riemann hypothesis. *Expo. Math.* **18**, 131–138 (2000)
2. Bettin, S., Conrey, J.B.: A reciprocity formula for a cotangent sum. *Int. Math. Res. Not. IMRN* **24**, 5709–5726 (2013)
3. Bettin, S., Conrey, J.B.: Period functions and cotangent sums. *Algebra Number Theory* **7**, 215–242 (2013)
4. Beurling, A.: On two problems concerning linear transformations in Hilbert space. *Acta Math.* **81**, 239–255 (1949)
5. Beurling, A.: A closure problem related to the Riemann zeta-function. *Proc. PNAS* **41**, 312–314 (1955)
6. Bruggeman, R., Lewis, J.: Function theory related to the group $PSL(2, \mathbb{R})$. In: Farkas, H.M., Gunning, R.C., Knopp, M.I., Taylor, B.A. (eds.) *From Fourier Analysis and Number Theory to Radon Transforms and Geometry*. In Memory of Leon Ehrenpreis. *Developments in Mathematics*, vol. 28, pp. 107–201. Springer, Berlin (2013)
7. Bruggeman, R., Lewis, J., Zagier, D.: Period functions for Maass wave forms and cohomology. *Mem. Am. Math. Soc.* **237**, xii+132 (2015)
8. Garnett, J.B.: *Bounded Analytic Functions*. Springer, Berlin (2007)
9. Grünberg, D., Moree, P.: Sequences of enumerative geometry: congruences and asymptotics, with an appendix by D. Zagier. *Exp. Math.* **17**, 409–426 (2008)
10. Halmos, P., Sunder, V.: *Bounded Integral Operators on L^2 -Spaces*. Springer, Berlin (1978)
11. Hoffman, K.: *Banach Spaces of Analytic Functions*. Prentice Hall, Upper Saddle River (1962)
12. Lewis, J.: Eigenfunctions on symmetric spaces with distribution-valued boundary forms. *J. Func. Anal.* **29**, 287–307 (1978)
13. Lewis, J., Zagier, D.: Period functions for Maass wave forms I. *Ann. Math.* **153**, 191–258 (2001)
14. Schwarz, J.H.: The generalized Stieltjes transform and its inverse. *J. Math. Phys.* **46**, 013501 (2005)
15. Zagier, D.: Eisenstein series and the Riemann zeta function. In: *Automorphic Forms, Representation Theory and Arithmetic*, pp. 275–301. Springer-Verlag, Berlin, Heidelberg, New York (1981)
16. Zagier, D.: Quantum modular forms. In: Étienne, B., David, E., Masoud, K., Matilde, M., Henri, M., Sorin, P. (eds.) *Quanta of Maths: Conference in Honor of Alain Connes*, *Clay Mathematics Proceedings*, vol. 11, pp. 659–675. AMS, Providence and Clay Mathematics Institute, Peterborough (2010)