Advances in Mathematics 170, 56–70 (2002)
doi:10.1006/aima.2001.2066

A Lower Bound in an Approximation Problem Involving the Zeros of the Riemann Zeta Function

Jean-François Burnol[1]

Université de Nice - Sophia Antipolis, Laboratoire J.-A. Dieudonné, Parc Valrose, F-06108 Nice Cedex 02, France
E-mail: burnol@math.unice.fr

Communicated by Alain Connes

Received March 23, 2001; accepted November 22, 2001

We slightly improve the lower bound of Báez-Duarte, Balazard, Landreau and Saias in the Nyman–Beurling formulation of the Riemann Hypothesis as an approximation problem. We construct Hilbert space vectors which could prove useful in the context of the so-called "Hilbert–Pólya idea".
Key Words: Riemann zeta function; Nyman–Beurling criterion; Hilbert–Pólya idea.

1. INTRODUCTION

The general framework of this paper is the Nyman–Beurling formulation of the Riemann hypothesis as an approximation problem [4, 17] (see also [3] and references therein). We have been especially motivated and inspired by recent theorems and conjectures by Báez-Duarte *et al.* [2]. The work of these authors is an indication that the potential of the Nyman–Beurling idea may not have been explored to its fullest yet. Another strong motivation is the so-called Hilbert–Pólya idea that the zeros of the Riemann zeta function are susceptible to a Hilbert space theoretical interpretation. We discuss this first, and will return later in this Introduction to the Nyman–Beurling problem per se.

In [11] and the subsequent paper [12], Connes gave a rather intrinsic construction of a Hilbert space intimately associated with the zeros of the Riemann zeta function on the critical line. But the zeros having multiplicities higher than a certain level (which is a parameter in Connes's construction) have (if they at all exist) their contributions limited to that level, and not to the extent given by their natural multiplicities. Thus, subsists the problem

[1]Current address: Université des Sciences et Technologies de Lille, Mathématiques, Cité Scientifique, F-59655 Villeneuve d'Ascq Cedex, France. E-mail: burnol@agat.univ-lille1.fr.

0001-8708/02 $35.00

of a natural definition of the so-called "Hilbert–Pólya space", with orthonormal basis indexed by the zeros ρ of ζ and integers k varying from 0 to $m_\rho - 1$ where m_ρ is the multiplicity of ρ. We do not solve that problem here but we do propose a rather natural construction of Hilbert space vectors $X^\lambda_{\rho,k}$, $\zeta(\rho) = 0$, $k < m_\rho$, which in the limit when the parameter λ goes to 0 become perpendicular (when they correspond to distinct zeros. The vectors corresponding to a multiple root ρ are independent but need to be orthogonalized). As in Connes's constructions these vectors live in a quotient space. Controlling the limit $\lambda \to 0$ to obtain the so-called Hilbert–Pólya space probably involves considerations from mathematical scattering theory (we have previously studied in [6, 7] some connections with the problems of L-functions; and in [9] we consider this theme further).

We now turn to the Nyman–Beurling formulation of the Riemann Hypothesis as an approximation problem ([4, 17]; see [3] and references therein for an introduction to this circle of ideas). Let $K = L^2(]0,\infty[, dt)$ (over the complex numbers), let χ be the indicator function of the interval $]0,1]$, and let ρ be the function "fractional part" (the letter ρ is also used to refer to a zero of the Riemann zeta function, hopefully no confusion will arise). Let $0 < \lambda < 1$ and let $\mathscr{B}_\lambda$ be the sub-vector space of K consisting of the finite linear combinations of the functions $t \mapsto \rho(\frac{\theta}{t})$, for $\lambda \leqslant \theta \leqslant 1$.

THEOREM 1.1 (Nyman [17], Beurling [4]). *The Riemann Hypothesis holds if and only if*

$$\chi \in \overline{\bigcup_{0<\lambda<1} \mathscr{B}_\lambda}.$$

Actually, we are following [2] here in using a slight variant of the original Nyman–Beurling formulation. It is a disappointing fact that this theorem can be proven without leading to any new information whatsoever on the zeros lying on the critical line (basically, what is at work is the factorization of functions belonging to the Hardy space of a half-plane [15]). The following is thus rather remarkable:

THEOREM 1.2 (Báez-Duarte *et al.* [2]). *Let us write $D(\lambda)$ for the Hilbert-space distance $\inf_{f\in\mathscr{B}_\lambda} ||\chi - f||$. We have*

$$\liminf_{\lambda\to 0}\ D(\lambda)\sqrt{\log\left(\frac{1}{\lambda}\right)} \geqslant \sqrt{\sum_\rho \frac{1}{|\rho|^2}}.$$

If the Riemann Hypothesis fails this result is true but trivial as the left-hand side then takes the value $+\infty$. The sum on the right-hand side is over all non-trivial zeros ρ of the zeta function, counted *only once* independent of their multiplicities m_ρ. We prove the following:

THEOREM 1.3. *We have*

$$\liminf_{\lambda\to 0} D(\lambda)\sqrt{\log\left(\frac{1}{\lambda}\right)} \geqslant \sqrt{\sum_\rho \frac{m_\rho^2}{|\rho|^2}}.$$

So the zeros are counted according to the *square* of their multiplicities. Let us discuss what this says about a conjecture by Báez-Duarte *et al.* [2]. It is known unconditionally that $\sum_\rho m_\rho/(\rho(1-\rho)) = 2+\gamma-\log(4\pi)$. Let us write C for $2+\gamma-\log(4\pi)$. The Riemann Hypothesis thus implies (and by an elementary lemma is itself in turn implied by) $\sum_\rho m_\rho/|\rho|^2 = C$. Báez-Duarte *et al.* [2] conjecture that $D(\lambda) \sim \sqrt{(C/|\log(\lambda)|)}$, and this, by the Nyman–Beurling criterion, implies the truthfulness of the Riemann Hypothesis. We see that our result implies, unconditionally, $\liminf_{\lambda\to 0} D(\lambda)\sqrt{|\log(\lambda)|} \geqslant \sqrt{C}$, and that from the conjecture of Báez-Duarte, Balazard, Landreau and Saias, if true, follows not only the truthfulness of the Riemann Hypothesis, but also the simplicity of all the zeros. We must confess though that we do not see in our proofs anything that could be argued to provide additional support to the (seemingly generally believed) hypothesis that the zeros are simple.

To prove our lower bound we construct remarkable Hilbert space vectors $X^\lambda_{\rho,k}$, $\zeta(\rho)=0$, $k<m_\rho$ and use them to control $D(\lambda)$. The following "toy-model" gives us reasons to expect that the lower bound, in fact, gives the exact order of decrease of $D(\lambda)$:

THEOREM 1.4. *Let* $Q(z) = \prod_\alpha (1-\bar{\alpha}\cdot z)^{m_\alpha}$ *be a polynomial of degree* $q \geqslant 1$ *with all its roots* α *on the unit circle* (*the root* α *having multiplicity* m_α). *Let* $P(z)$ *be an arbitrary polynomial. Let*

$$E(N,P) := \inf_{\deg(A)\leqslant N} \int_{S^1} |P(z) - Q(z)A(z)|^2 \frac{d\theta}{2\pi}.$$

We have as N goes to infinity,

$$\lim NE(N,P) = \sum_\alpha m_\alpha^2 |P(\alpha)|^2.$$

The association of Hilbert spaces and vectors to zeros of the Riemann zeta function, or more generally to the zeros of Abelian L-functions, is a matter which we have also pursued (unconditionally) in [9]. From now on in this paper, the Riemann Hypothesis is assumed to hold true.

2. THE PREDICTION ERROR FOR A SINGULAR MA(Q)

As motivation for our result we first consider a simpler approximation problem, in the context of the Hardy space of the unit disc rather than the Hardy space of a half-plane. Let $Q(z) = \prod_\alpha (1 - \bar{\alpha} z)^{m_\alpha}$ be a polynomial of degree $q \geqslant 1$ with all its roots α on the unit circle (the root α having multiplicity m_α so that $q = \sum_\alpha m_\alpha$). Let us define

$$E(N) := \inf_{\deg(A) \leqslant N} \int_{S^1} |1 - Q(z)A(z)|^2 \frac{d\theta}{2\pi}.$$

The measure $\frac{d\theta}{2\pi}$ is the rotation-invariant probability measure on the circle S^1, with $z = \exp(i\theta)$. The minimum is taken over all complex polynomials $A(z)$ with degree at most N. We are guaranteed that $\lim_{N \to \infty} E(N) = 0$ as $Q(z)$ is an outer factor [15]. More precisely:

THEOREM 2.1. *As N goes to infinity we have*

$$\lim NE(N) = \sum_\alpha m_\alpha^2.$$

Note 1. In case $Q(z)$ has a root in the open unit disc then $E(N)$ is bounded below by a positive constant. In case $Q(z)$ has all its roots outside the open unit disc, then the result above holds but only the roots on the unit circle contribute. Finally if all its roots are outside the closed unit disc then the decrease is exponential: $E(N) = O(c^N)$, with $c < 1$.

Note 2. The theorem, although not stated explicitly there, is easily extracted from the work of Grenander and Rosenblatt [14]. They state an $O(1/N)$ result, in a much wider set-up than the one considered here (which is limited to simple-minded qth order moving averages). Unfortunately, the $O(1/N)$ bound is now believed not to be systematically true under their hypotheses (as is explained in [16]; I thank Professor W. Van Assche for pointing out this fact to me). Nevertheless, their technique of proof goes through smoothly in the case at hand and yields the exact asymptotic result as stated above. We only sketch briefly the idea, as nothing beyond the tools used in [14] is needed.

Proof of Theorem 2.1. We point out in passing that it is, of course, possible to express $E(N)$ explicitly in terms of the Toeplitz determinants for the measure $d\mu = |Q(\exp(i\theta))|^2 \, d\theta/2\pi$. But already for an MA(2) this gives rise to unwieldy computations... . Rather, let $\mathscr{P}_N$ be the vector space of polynomials of degrees at most $N + q$, let $\mathscr{V}_N$ be the subspace of polynomials divisible by $Q(z)$, and let $\mathscr{W}_N$ be its q-dimensional orthogonal complement. Then $E(N)$ is the squared norm of the orthogonal projection of the constant function 1 to $\mathscr{W}_N$. A spanning set in $\mathscr{W}_N$ is readily identified: to each root α one associates $Y^N_{\alpha,0}, Y^N_{\alpha,1}, \dots, Y^N_{\alpha,m_\alpha-1}$ defined as

$$Y^N_{\alpha,0} := 1 + \bar\alpha z + \cdots + \bar\alpha^{N+q} z^{N+q},$$

$$Y^N_{\alpha,1} := z + 2\bar\alpha z^2 + \cdots + (N+q)\bar\alpha^{N+q-1} z^{N+q}$$

and similarly for $k = 2, \dots, m_\alpha - 1$. We can then express $E(N)$ using a Gram formula in terms of (the inverse) of the positive matrix (of fixed size $q \times q$ but depending on N) built with the scalar products of the Y's. It turns out that in the limit when N goes to infinity and after the rescaling $Y^N_{\alpha,k} \mapsto X^N_{\alpha,k} := N^{-k-1/2} Y^N_{\alpha,k}$ the Gram matrix decomposes into Cauchy blocks $(1/(i+j+1))_{0 \leqslant i,j < m_\alpha}$ of size m_α, one for each root α. It is known from Cauchy [10, Vol. XII, p. 177] that the top-left element of the inverse matrix is m_α^2. This is how $\sum_\alpha m_\alpha^2/N$ arises, after keeping track of the scalar products $(1, X^N_{\alpha,k})$. Instead of the constant polynomial 1 we could have looked at the approximation rate to an arbitrary polynomial $P(z)$. The proof just sketched applies identically and gives Theorem 1.4 from the Introduction. ∎

3. INVARIANT ANALYSIS AND A CONSTRUCTION OF BÁEZ-DUARTE

The Mellin transform $f(t) \mapsto \hat f(s) = \int_{t>0} f(t) t^{s-1}\, \mathrm{d}t$ establishes the Plancherel isometry between $K = L^2(]0,\infty[, dt)$ and $L^2(s = \frac{1}{2} + i\tau, \frac{d\tau}{2\pi})$, with inverse $F(s) \mapsto \int_{s=1/2+i\tau} F(s) t^{-s}\, d\tau/2\pi$. Let $a(s)$ be a measurable function of s (as a rule when using the letter s we implicitly assume $\mathrm{Re}(s) = 1/2$. We will use letters w and z for general complex numbers). If $a(s)$ is essentially bounded then $F(s) \mapsto a(s)F(s)$ defines a bounded operator on K which commutes with the unitary group $D_\theta : f(t) \mapsto f(t/\theta)/\sqrt{\theta}$, and all bounded operators commuting with the D_θ $(0 < \theta < \infty)$ are obtained in such a manner. More generally, all *closed* invariant operators are associated to a measurable multiplier $a(s)$ (finite almost everywhere, but not necessarily essentially bounded). For the details of this technical statement, see [8].

For example, the Hardy averaging operator $M : f(t) \mapsto (\int_{]0,t]} f(u)\, du)/t$ corresponds to the spectral multiplier $1/(1-s)$. The operator $1 - M$ corresponds to the spectral multiplier $s/(s-1)$ and is thus unitary. Another (see [5]) remarkable invariant operator is the (even) "Gamma" operator $\Gamma_+ = \mathscr{F}_+ I$. Here I is the inversion $f(t) \mapsto f(1/t)/t$ and $\mathscr{F}_+$ is the additive Fourier transform as applied to even functions (the cosine transform formally given as $\mathscr{F}_+(f)(u) = 2\int_0^\infty \cos(2\pi ut) f(t)\, dt$). The multiplier associated to Γ_+ is

$$\gamma_+(s) = \pi^{\frac{1}{2}-s} \frac{\Gamma(s/2)}{\Gamma((1-s)/2)} = \frac{\zeta(1-s)}{\zeta(s)} = 2^{1-s}\pi^{-s} \cos\left(\frac{\pi s}{2}\right)\Gamma(s).$$

This is closely related with the identity of distributions ($0 < \mathrm{Re}(s) < 1$):

$$\mathscr{F}_+(|t|^{s-1}) = \gamma_+(s)|u|^{-s}$$

which in [18] is a special case of the Tate functional equations on local fields. Let us pick a decreasing sequence of non-negative Schwartz functions $\phi_n(u)$, converging pointwise to the indicator function of $[-1, 1]$. When applied to ϕ_n the right-hand side converges to $2\gamma_+(s)/(1-s)$. The Fourier transforms of the ϕ_n converge pointwise to $2\sin(2\pi t)/(2\pi t)$, and they will be uniformly $O(1/t)$ if we choose the $\phi_n(u)$, as may be done, with the L^1 norm of their derivatives bounded. This gives the well-known identity ($0 < \mathrm{Re}(s) < 1$),

$$\gamma_+(s) = (1-s)\int_0^\infty t^{s-1}\frac{\sin(2\pi t)}{\pi t}\,dt.$$

A further invariant operator is the operator U introduced by Báez-Duarte [1] in connection with the Nyman–Beurling formulation of the Riemann Hypothesis: its spectral multiplier is $(s/(1-s))(\zeta(1-s)/\zeta(s))$, so $U = (M-1)\mathscr{F}_+ I = \mathscr{F}_+ I(M-1)$.

From the results recalled above on invariant operators, we see that invariant orthogonal projectors correspond to indicator functions of measurable sets on the critical line. So a function $f(t)$ is such that its multiplicative translates $D_\theta(f)$ $(0 < \theta < \infty)$ span K if and only if $F(s) = \hat{f}(s)$ is almost everywhere non-vanishing (Wiener's L^2-Tauberian Theorem). In that case, the phase function

$$U_f(s) = \frac{\overline{F(s)}}{F(s)}$$

is almost everywhere defined and of modulus 1. It thus corresponds to an invariant unitary operator, also denoted U_f.

Let us introduce the *anti-unitary* "time-reversal" operator J acting on K as $g \mapsto \overline{I(g)}$. The operator U_f commutes with the contractions–dilations, is unitary, and sends f to $J(f)$. We call this the *Báez-Duarte construction* as it appears in [1] (up to some non-essential differences) in relation with the Nyman–Beurling problem (the phase function arises in other contexts, especially in scattering theory).

To relate this with the operator $U = (M-1)\mathscr{F}_+ I$, one needs the formula

$$\frac{\zeta(s)}{s} = -\int_0^\infty \rho\left(\frac{1}{t}\right)t^{s-1}\,dt$$

which is fundamental in the Nyman–Beurling context. This formula shows that U is the phase operator associated with $\rho(1/t)$.

Generally speaking, the operators U_f are related to the Hardy spaces $\mathbb{H}^2 = L^2(]0,1], dt)$ and $\mathbb{H}^{2^\perp} = L^2([1,\infty[, dt)$ (we will also use the notation $\mathbb{H}^2$ for the Mellin transform of $L^2(]0,1], dt)$). Indeed, the time-reversal J is an isometry (anti-unitary) between $\mathbb{H}^2$ and $\mathbb{H}^{2^\perp}$. Let us assume that the function f belongs to $\mathbb{H}^2$. The operator U_f has the same effect as J on f, but contrarily to J is an *invariant* operator. This puts the space $\mathscr{B}_\lambda(f)$ (of finite linear combinations of contractions $D_\theta(f)$ for $\lambda \leqslant \theta \leqslant 1$) isometrically in a new light as a subspace of $L^2([\lambda,\infty[, dt)$. The marvelous thing is that in this new incarnation it appears to be sometimes possible to find vectors orthogonal to $\mathscr{B}_\lambda(f)$ and thus to get some control on $\mathscr{B}_\lambda(f)$ as λ decreases (as in the Grenander–Rosenblatt method).

4. THE VECTORS $Y^\lambda_{s,k}$

To get started on this we first replace the L^2 function $-\zeta(s)/s$ with an element of $\mathbb{H}^2$. This is elementary:

THEOREM 4.1 (Burnol [7], Ehm [13]). *The function $Z(s) = (s-1)\zeta(s)/s^2$ belongs to $\mathbb{H}^2$. Its inverse Mellin transform $A(t)$ is given by the formula*

$$A(t) = \left[\frac{1}{t}\right]\log(t) + \log\left(\left[\frac{1}{t}\right]!\right) + \left[\frac{1}{t}\right].$$

One has (for $0 < t \leqslant 1$) $A(t) = \log(1/t)/2 + O(1)$.

The Báez-Duarte construction will then associate to $A(t)$ the operator V with spectral multiplier

$$V(s) = \left(\frac{s}{1-s}\right)^3 \frac{\zeta(1-s)}{\zeta(s)},$$

so that

$$V = (1-M)^2 U.$$

This last representation will prove useful as it allows to use the formulae related to U from [1, 2]. Let $\mathscr{C}_\lambda$ $(0 < \lambda < 1)$ be the sub-vector space of $\mathbb{H}^2$ of linear combinations of the contractions $D_\theta(A)$ for $\lambda \leqslant \theta \leqslant 1$. The function $(s-1)/s1/s = 1/s - 1/s^2$ is the Mellin transform of $\chi_1(t) := (1 + \log(t))\chi(t)$. The quantity $D(\lambda)$ considered by Báez-Duarte, Balazard, Landreau and Saias is thus the Hilbert space distance between $\chi_1(t)$ and $\mathscr{C}_\lambda$. To bound it from below we will exhibit remarkable Hilbert space vectors $X^\lambda_{\rho,k}$ indexed by the zeros of the Riemann zeta function and perpendicular to $\mathscr{C}_\lambda$. We then compute the exact asymptotics of the orthogonal projection of χ_1 to the vector spaces spanned by the $X^{\lambda}_{\rho,k}$, for a finite set of roots, exactly as in the Grenander–Rosenblatt method.

To each complex number w and natural integer $k \geqslant 0$ we associate the function $\psi_{w,k}(t) = (\log(1/t))^k t^{-w}\chi(t)$ on $]0,\infty[$. For $\mathrm{Re}(w) < 1$ it is integrable, for $\mathrm{Re}(w) < 1/2$ it is in K. Let Q_λ be the orthogonal projector from K onto $L^2([\lambda,\infty[)$. The main point of this paper is the following:

THEOREM AND DEFINITION 4.2. *For each $0 < \lambda \leqslant 1$, each s on the critical line, and each integer $k \geqslant 0$ the L^2-limit in K of $V^{-1}Q_\lambda V(\psi_{w,k})$ exists as w tends to s from the left half-plane*:

$$Y^{\lambda}_{s,k} := \lim_{\substack{w \to s \\ \mathrm{Re}(w) < 1/2}} V^{-1} Q_\lambda V(\psi_{w,k}).$$

For each $\lambda \leqslant \theta \leqslant 1$ the scalar products between $D_\theta(A)$ and the vectors $Y^{\lambda}_{s,k}$ are

$$\lambda \leqslant \theta \leqslant 1 \Rightarrow (D_\theta(A), Y^{\lambda}_{s,k}) = \left(-\frac{d}{ds}\right)^k \theta^{s-\frac{1}{2}} Z(s).$$

Note 3. The proof shows the existence of an analytic continuation in w across the critical line. This is exploited further in [9].

Clearly, one has the following statement as an immediate consequence:

COROLLARY 4.3. *Let $0 < \lambda < 1$. The vector $Y^{\lambda}_{s,k}$ is perpendicular to $\mathscr{C}_\lambda$ if and only if $\zeta^{(j)}(s) = 0$ for all $j \leqslant k$, if and only if s is a zero ρ of the zeta function and $k < m_\rho$.*

Note 4. Our scalar products (f, g) are (complex) linear in the first factor and conjugate-linear in the second factor.

Note 5. The operator $\frac{d}{ds}$ when applied to a not necessarily analytic function on the critical line is defined to act as $1/i\, d/d\tau$ (where $s = 1/2 + i\tau$).

Proof of Corollary 4.3. The proof of existence will be given later. Here we check the statement involving the scalar product, assuming existence. The following holds for $\lambda \leqslant \theta \leqslant 1$ and $\mathrm{Re}(w) < 1/2$:

$$\begin{aligned}(V^{-1}Q_\lambda V(\psi_{w,k}), D_\theta(A)) &= (Q_\lambda V(\psi_{w,k}), VD_\theta(A))\\ &= (Q_\lambda V(\psi_{w,k}), D_\theta V(A))\end{aligned}$$

$$
\begin{aligned}
&= (V(\psi_{w,k}), Q_\lambda D_\theta J(A))\\
&= (V(\psi_{w,k}), D_\theta V(A))\\
&= (\psi_{w,k}, D_\theta(A))\\
&= \left(\frac{d}{dw}\right)^k (\psi_{w,0}, D_\theta(A))\\
&= \left(\frac{d}{dw}\right)^k (D_\theta^{-1}(t^{-w}\chi(t)), A)\\
&= \left(\frac{d}{dw}\right)^k (\theta^{1/2-w} t^{-w}\chi(\theta t), A)\\
&= \left(\frac{d}{dw}\right)^k \theta^{1/2-w} \int_{]0,1]} t^{-w}\overline{A(t)}\, dt.
\end{aligned}
$$

Taking the limit when $w \to s$ gives

$$
\begin{aligned}
(Y^\lambda_{s,k}, D_\theta(A)) &= \left(\frac{d}{ds}\right)^k \theta^{1/2-s} \int_{]0,1]} t^{-s}\overline{A(t)}\, dt\\
&= \left(\frac{1}{i}\frac{d}{d\tau}\right)^k \theta^{-i\tau} \int_{]0,1]} t^{-\frac{1}{2}-i\tau}\overline{A(t)}\, dt
\end{aligned}
$$

Taking the complex conjugate:

$$
\begin{aligned}
(D_\theta(A), Y^\lambda_{s,k}) &= \left(i\frac{d}{d\tau}\right)^k \theta^{i\tau} \int_{]0,1]} t^{-\frac{1}{2}+i\tau} A(t)\, dt\\
&= \left(-\frac{d}{ds}\right)^k \theta^{s-\frac{1}{2}} \int_{]0,1]} t^{s-1} A(t)\, dt\\
&= \left(-\frac{d}{ds}\right)^k \theta^{s-\frac{1}{2}} Z(s)
\end{aligned}
$$

which completes the proof (assuming existence). ∎

To prove the existence we will use in an essential manner the key Lemma 6 from [2]. We have seen that $V = (1-M)^2 U$ where M is the Hardy averaging operator and U the Báez-Duarte operator. The spectral function $U(s)$ extends to an analytic function $U(w)$ in the strip $0 < \mathrm{Re}(w) < 1$. We need pointwise expressions for $V(\psi_{w,k})(t)$, $t > 0$ (at first only $\mathrm{Re}(w) < 1/2$ is allowed here). Thanks to the general study of U given in [1], we know that for $\mathrm{Re}(w) < 1/2$ the vector $U(\psi_{w,k})$ in K is given as the following limit in square mean:

$$\lim_{\delta\to 0}\int_\delta^1 \left(\log\left(\frac{1}{v}\right)\right)^k v^{-w}\,\frac{d}{dv}\frac{\sin(2\pi t/v)}{\pi t/v}\,dv.$$

Following [2], with a slight change of notation, we now study for each complex number w with $\mathrm{Re}(w)<1$ (and each integer $k\geqslant 0$) the *pointwise* limit as a function of $t>0$ for $\delta\to 0$:

$$\varphi_{w,k}(t) := \lim_{\delta\to 0}\int_\delta^1 \left(\log\left(\frac{1}{v}\right)\right)^k v^{-w}\,\frac{d}{dv}\frac{\sin(2\pi t/v)}{\pi t/v}\,dv.$$

THEOREM 4.4 (Báez-Duarte *et al.* [2]). *Let* $k=0$. *For each* $t>0$ *and* $\mathrm{Re}(w)<1$ *the pointwise limit defining* $\varphi_{w,0}(t)$ *exists. It is holomorphic in* w *for each fixed* t. *When* w *is restricted to a compact set in* $\mathrm{Re}(w)<1$, *one has uniformly in* w *the bound* $\varphi_{w,0}(t)=O(1/t)$ *on* $[1,\infty[$. *Uniformly with respect to* w *satisfying* $0<\mathrm{Re}(w)<1$, *one has* $\varphi_{w,0}(t)=U(w)t^{-w}+O(1)$ *on* $0<t\leqslant 1$.

Proof. Everything is either stated explicitly in [2, Lemmas 6 and 4], or follows from their proofs. We will give more details for $k\geqslant 1$ as this is not treated in [2]. ∎

COROLLARY 4.5. *For each* w *in the critical strip* $0<\mathrm{Re}(w)<1$ *the Hardy operator* $M : f(t)\to(\int_0^t f(v)\,dv)/t$ *can be applied arbitrarily many times to* $\varphi_{w,0}(t)$. *The functions* $M^L(\varphi_{w,0})$ $(L\in\mathbb{N})$ *are* $O((1+\log(t))^L/t)$ *on* $[1,\infty[$, *uniformly with respect to* w *when it is restricted to a compact subset of the open strip, and satisfy on* $t\in\,]0,1]$ *the estimate* $M^L(\varphi_{w,0})(t)=(1/(1-w))^L U(w)t^{-w}+O(1)$, *uniformly with respect to* w.

Proof. A simple recurrence. ∎

We thus obtain:

COROLLARY 4.6. *The vectors* $Y^\lambda_{s,0}$ *exist* (*for* $\mathrm{Re}(s)=1/2$). *One has the estimates*:

$$V(Y^\lambda_{s,0})(t)=O\left(\frac{(1+\log(t))^2}{t}\right)\quad(t\in[1,\infty[),$$

$$V(Y^\lambda_{s,0})(t)=V(s)t^{-s}+O(1)\quad(\lambda<t\leqslant 1),$$

$$V(Y^\lambda_{s,0})(t)=0\quad(0<t<\lambda)$$

uniformly with respect to s *when its imaginary part is bounded.*

THEOREM 4.7. *Let $k \geqslant 1$. For each $t > 0$ and $\mathrm{Re}(w) < 1$ the pointwise limit defining $\varphi_{w,k}(t)$ exists. It is holomorphic in w for each fixed t. When w is restricted to a compact set in $\mathrm{Re}(w) < 1$, one has uniformly in w the bound $\varphi_{w,k}(t) = O(1/t)$ on $[1, \infty[$. Uniformly for $0 < \mathrm{Re}(w) < 1$ one has $\varphi_{w,k}(t) = (d/dw)^k (U(w)t^{-w}) + O(1)$ on $0 < t \leqslant 1$.*

Proof. The formula defining $\varphi_{w,k}(t)$ is equivalent to (after integration by parts and the change of variable $u = 1/v$)

$$\varphi_{w,k}(t) = \lim_{\Lambda \to \infty} \frac{1}{\pi t} \int_1^{\Lambda} (k + w \log(u))(\log(u))^{k-1}\, u^{w-1} \sin(2\pi t u) \frac{du}{u}.$$

This proves the existence of $\varphi_{w,k}(t)$, its analytic character in w, and the uniform $O(1/t)$ bound on $[1, \infty[$. The formula can be re-written as

$$\varphi_{w,k}(t) = \left(\frac{d}{dw}\right)^k \frac{w}{\pi t} \int_1^{\infty} u^{w-1} \sin(2\pi t u) \frac{du}{u}.$$

When w is in the critical strip the integral $\int_0^\infty u^{w-1} \sin(2\pi t\, u)\, du/u$ is absolutely convergent and its value is $t^{1-w} \int_0^\infty u^{w-1} \sin(2\pi u)\, du/u = \pi t^{1-w} \gamma_+(w)/(1-w)$, so that

$$\varphi_{w,k}(t) = \left(\frac{d}{dw}\right)^k \left(\frac{w}{1-w} \gamma_+(w) t^{-w} - \frac{w}{\pi t} \int_0^1 u^{w-1} \sin(2\pi t u) \frac{du}{u}\right).$$

The first term is $(d/dw)^k (U(w)t^{-w})$ and the second term can be explicitly evaluated using the series expansion of $\sin(2\pi t u)$ with the final result

$$\varphi_{w,k}(t) = \left(\frac{d}{dw}\right)^k (U(w)t^{-w}) + 2(-1)^k k! \sum_{j \geqslant 1} (-1)^j \frac{(2\pi t)^{2j}}{(2j+1)!} \frac{2j}{(w+2j)^{k+1}}$$

which shows $\varphi_{w,k}(t) = (d/dw)^k (U(w)t^{-w}) + O(1)$, on $0 < t \leqslant 1$, uniformly for $0 < \mathrm{Re}(w) < 1$. ∎

As was the case for $k = 0$ we then deduce that the Hardy operator can be applied arbitrarily many times to $\varphi_{w,k}$ for $0 < \mathrm{Re}(w) < 1$. The existence of the $Y_{s,k}^{\lambda}$ follows.

THEOREM 4.8. *Let* $k \geqslant 0$. *The vectors* $Y^{\lambda}_{s,k}$ *exist* (*for* $\mathrm{Re}(s) = 1/2$). *One has the estimates*

$$V(Y^{\lambda}_{s,k})(t) = O\left(\frac{(1+\log(t))^2}{t}\right) \quad (t \in [1,\infty[),$$

$$V(Y^{\lambda}_{s,k})(t) = \left(\frac{d}{ds}\right)^k (V(s)t^{-s}) + O(1) \quad (\lambda < t \leqslant 1),$$

$$V(Y^{\lambda}_{s,k})(t) = 0 \quad (0 < t < \lambda),$$

the implied constants are independent of λ *and are uniform with respect to* s *when its imaginary part is bounded.*

Proof. Clearly a corollary to Theorem 4.7. ∎

5. THE VECTORS $X^{\lambda}_{\rho,k}$ AND COMPLETION OF THE PROOF

DEFINITION 1. Let $0 < \lambda < 1$. To each zero ρ of the Riemann zeta function on the critical line, of multiplicity m_ρ, and each integer $0 \leqslant k < m_\rho$ we associate the Hilbert space vector

$$X^{\lambda}_{\rho,k} := \left(\log\left(\frac{1}{\lambda}\right)\right)^{-\frac{1}{2}-k} Y^{\lambda}_{\rho,k},$$

where $Y^{\lambda}_{\rho,k} = \lim_{w \to s} V^{-1} Q_\lambda V(\psi_{w,k})$, V is the unitary operator $(M-1)^3 \mathscr{F}_+ I$, Q_λ is the orthogonal projection to $L^2([\lambda,\infty[, dt)$, and $\psi_{w,k}(t) = (\log(1/t))^k\, t^{-w}\chi(t)$.

Note 6. Of course, there is no reason except psychological to allow only zeros of the Riemann zeta function at this stage.

THEOREM 5.1. *As* λ *decreases to* 0 *one has*

$$\lim_{\lambda \to 0} (X^{\lambda}_{\rho_1,k}, X^{\lambda}_{\rho_2,l}) = 0 \quad (\rho_1 \neq \rho_2),$$

$$\lim_{\lambda \to 0} (X^{\lambda}_{\rho,k}, X^{\lambda}_{\rho,l}) = \frac{1}{k+l+1}.$$

Proof. To establish this we first consider, for $\mathrm{Re}(s_1) = \mathrm{Re}(s_2) = 1/2$:

$$\int_{\lambda}^{1} \left(\log\left(\frac{1}{t}\right)\right)^{j_1} t^{-s_1} \left(\log\left(\frac{1}{t}\right)\right)^{j_2} t^{-(1-s_2)}\, dt.$$

If $s_1 \neq s_2$ an integration by parts shows that it is $O(\log(1/\lambda))^{j_1+j_2}$. On the other hand, when $s_1 = s_2$ its exact value is $(\log(1/\lambda))^{j_1+j_2+1}/(j_1+j_2+1)$. With this information the theorem follows directly from Theorem 4.8 as (for example) the leading divergent contribution as $\lambda \to 0$ to $(V(Y^\lambda_{s,k}), V(Y^\lambda_{s,l}))$ is $V(s)\overline{V(s)} \int_\lambda^1 (d/ds)^k t^{-s} (d/ds)^l t^{-s}\, dt$ which gives $(\log(1/\lambda))^{k+l+1}/(k+l+1)$. The rescaling $Y \mapsto X$ is chosen so that a finite limit for $(X^\lambda_{\rho,k}, X^\lambda_{\rho,l})$ is obtained. As the scalar products involving distinct zeros have a smaller divergency, the rescaling let them converge to 0. ∎

THEOREM 5.2. *Let $\chi_1(t) = (1+\log(t))\chi(t)$. As λ decreases to* 0 *one has*

$$\lim_{\lambda\to 0} \sqrt{\log\left(\frac{1}{\lambda}\right)}(\chi_1, X^\lambda_{\rho,k}) = 0 \quad (k \geqslant 1),$$

$$\lim_{\lambda\to 0} \sqrt{\log\left(\frac{1}{\lambda}\right)}(\chi_1, X^\lambda_{\rho,0}) = \frac{\rho-1}{\rho^2}.$$

Proof. We have $(1-M)\chi_1 = \chi$, and $V = (1-M)^2 U$ so $V\chi_1 = (1-M)U\chi$. From [1] we know that $U\chi$ is $\sin(2\pi t)/(\pi t)$ so $V\chi_1$ is the function $\sin(2\pi t)/(\pi t) - (\int_0^t \sin(2\pi v)/(\pi v)\, dv)/t$. It is thus $0(t^2)$ as $t \to 0$, and from Theorem 4.8 we then deduce that the scalar products $(\chi_1, Y^\lambda_{\rho,k})$ admit finite limits as $\lambda \to 0$. This settles the case $k \geqslant 1$. For $k = 0$, one uses the uniformity with respect to w in Theorem 4.4 to get

$$\lim_{\lambda\to 0} (\chi_1, Y^\lambda_{\rho,0}) = \lim_{w\to\rho} (\chi_1, \varphi_{w,0})$$

which gives $\lim_{w\to\rho} \int_0^1 (1+\log(t)) t^{w-1}\, dt = 1/\rho - 1/\rho^2 = (\rho-1)/\rho^2$. ∎

We can now conclude the proof of our estimate.

THEOREM 5.3. *We have*

$$\liminf_{\lambda\to 0} D(\lambda)\sqrt{\log\left(\frac{1}{\lambda}\right)} \geqslant \sqrt{\sum_\rho \frac{m_\rho^2}{|\rho|^2}}.$$

Proof. Let R be a non-empty finite set of zeros. We showed that $D(\lambda)$ is the Hilbert space distance from χ_1 to $\mathscr{C}_\lambda$, and that the vectors $X^\lambda_{\rho,k}$ for $0 \leqslant k < m_\rho$ are perpendicular to $\mathscr{C}_\lambda$. So $D(\lambda)$ is bounded below by the norm of the orthogonal projection of χ_1 to the finite-dimensional vector space H_R spanned by the vectors $X^\lambda_{\rho,k}, 0 \leqslant k < m_\rho, \rho \in R$. As is well known this is given exactly as the square root of the matrix product

$$[(\chi_1, X^\lambda_{\rho,k})]\, \mathrm{Gram}(X^\lambda_{\rho,k})^{-1} [(X^\lambda_{\rho,k}, \chi_1)]^t.$$

From Theorem 5.1 the Gram matrix (built with the scalar products of the $X^{\lambda}_{\rho,k}$) converges to diagonal blocks, one for each zero, given by Cauchy matrices $(1/(i+j+1))_{i,j\geqslant 0}$ of sizes $m_\rho \times m_\rho$. From Cauchy [10, Vol. XII, p. 177] we know that the top-left element of the inverse matrix is m_ρ^2. Combining this with the scalar products evaluated in Theorem 5.2 we get that the squared norm of the orthogonal projection of χ_1 to H_R is asymptotically equivalent as $\lambda \to 0$ to $(\sum_{\rho\in R} m_\rho^2/|\rho|^2)/\log(1/\lambda)$. The proof is complete. ∎

We can apply our strategy to a fully singular MA(q) on the unit circle. The relevant Báez-Duarte phase operator will then be (up to a non-important constant of modulus 1) the operator of multiplication by z^{-q} and it is apparent that this leads to a proof equivalent to the one we gave in our previous discussion, inspired by Grenander and Rosenblatt [14]. In the case of the Nyman–Beurling approximation problem for the zeta function, we expect in the quotient of $\mathbb{H}^2$ by $\overline{\mathscr{C}_\lambda}$ a "continuous spectrum" in addition to the "discrete spectrum" provided by the (projection to $\mathbb{H}^2$ of the) $X^{\lambda}_{\rho,k}$'s, $\zeta(\rho) = 0$, $k < m_\rho$. It is tempting to speculate that the rescaling will kill this continuous part as $\lambda \to 0$, so that in the end only subsists the so-called "Hilbert–Pólya" space. This would appear to require Theorem 5.3 to give the exact order of decrease of the quantity $D(\lambda)$ and the numerical explorations reported by Báez-Duarte *et al.* [2] seem to support this.

ACKNOWLEDGMENTS

I had been looking for the vectors $X^{\lambda}_{\rho,k}$, and with Lemma 6 of [2] they were suddenly there. I thank Michel Balazard for giving me copies of [1, 2] in preprint form.

REFERENCES

1. L. Báez-Duarte, A class of invariant unitary operators, *Adv. in Math.* **144** (1999), 1–12.
2. L. Báez-Duarte, M. Balazard, B. Landreau, and E. Saias, Notes sur la fonction ζ de Riemann 3, *Adv. in Math.* **149** (2000), 130–144.
3. M. Balazard and E. Saias, The Nyman–Beurling equivalent form for the Riemann hypothesis, *Exposition Math.* **18** (2000), 131–138.
4. A. Beurling, A closure problem related to the Riemann Zeta-function, *Proc. Nat. Acad. Sci.* **41** (1955), 312–314.
5. J.-F. Burnol, Sur les formules explicites I: Analyse invariante, *C. R. Acad. Sci. Paris Sér.* I **331** (2000), 423–428.
6. J.-F. Burnol, Scattering on the p-adic field and a trace formula, *Internat. Math. Res. Notes* **2000**, No. 2 (2000), 57–70.
7. J.-F. Burnol, An adelic causality problem related to abelian L-functions, *J. Number Theory* **87**, No. 2 (2001), 253–269.

8. J.-F. Burnol, Quaternionic gamma functions and their logarithmic derivatives as spectral functions, *Math. Res. Lett.* **8**, Nos. 1–2 (2001), 209–223.
9. J.-F. Burnol, Sur certains espaces de Hilbert de fonctions entières, liés à la transformation de Fourier et aux fonctions L de Dirichlet et de Riemann, *C. R. Acad. Sci. Paris Sér.* I **333** (2001), 201–206.
10. A. Cauchy, "Oeuvres complètes d'Augustin Cauchy," Publiées sous la direction scientifique de l'Académie des Sciences, IIe Série, Gauthier-Villars, Paris.
11. A. Connes, Formule de trace en géométrie non-commutative et hypothèse de Riemann, *C. R. Acad. Sci. Paris Sér.* I **323** (1996), 1231–1236.
12. A. Connes, Trace formula in non-commutative geometry and the zeros of the Riemann zeta function, *Selecta Math.* (*N.S.*) **5**, No. 1 (1999), 29–106.
13. W. Ehm, A family of probability densities related to the Riemann zeta function, *in* "Algebraic Methods in Probability and Statistics" (M. Viana and D. Richards, Eds.), Contemporary Mathematical Series, Vol. 287, Amer. Math. Soc., Providence, RI, 2001.
14. U. Grenander and M. Rosenblatt, An extension of a theorem of G. Szegö and its application to the study of stochastic processes, *Trans. Amer. Math. Soc.* **76** (1954), 112–126.
15. K. Hoffman, "Banach Spaces of Analytic Functions," Prentice-Hall, Inc., Englewood Cliffs, NJ, 1962 (Dover Pub., New York, 1988).
16. P. Nevai, Géza Freud, orthogonal polynomials and Christoffel functions. A case study, *J. Approx. Theory* **48** (1986), 3–167.
17. B. Nyman, "On the One-Dimensional Translation Group and Semi-Group in Certain Function Spaces," Thesis, University of Uppsala, 55p, 1950.
18. J. Tate, Fourier analysis in number fields and Hecke's zeta function, Princeton, 1950, *in* "Algebraic Number Theory" (J.W.S. Cassels and A. Fröhlich, Eds.), Academic Press, New York, 1967.