ON THE NYMAN-BEURLING CRITERION FOR THE RIEMANN HYPOTHESIS

Laurent Habsieger

Abstract: The Nyman-Beurling criterion states that the Riemann hypothesis is equivalent to the density in $L^2(0, +\infty; t^{-2}dt)$ of a certain space. We introduce an orthonormal family in $L^2(0, +\infty; t^{-2}dt)$, study the space generated by this family and reformulate the Nyman-Beurling criterion using this orthonormal basis. We then study three approximations that could lead to a proof of this criterion.
Keywords: Riemann hypothesis, Nyman-Beurling criterion.

1. Introduction

Let $\mathcal{H}$ denote the Hilbert space $L^2(0, +\infty; t^{-2}dt)$, endowed with the scalar product

$$\langle f, g \rangle = \int_0^{+\infty} f(t)\bar{g}(t)\frac{dt}{t^2}\,.$$

It contains the functions $\rho_a(t) = \{t/a\}$, $a > 0$. The Nyman-Beurling [5,7] criterion states that the Riemann hypothesis is equivalent to the density in $L^2(1, +\infty; t^{-2}dt)$ of the space spanned by the functions $a\rho_a - \rho_1$, $a \geqslant 1$. As a matter of fact, the condition may be replaced by the following one: the characteristic function of the interval $[1, +\infty[$ is the limit in $\mathcal{H}$ of linear combinations of the ρ_a, $a \geqslant 1$. Báez-Duarte [2] even proved that one may assume $a \in \mathbb{N}$ in this last statement.

In this paper, we introduce an orthonormal family $(E_n)_{n \in \mathbb{N}^*}$ in $\mathcal{H}$, and study approximations properties with respect to this family. In Section 2 we describe the space $\mathcal{HP}$ spanned by this family: the functions in $\mathcal{HP}$ are those of $\mathcal{H}$ that enjoy two simple symmetry properties. This Hilbert space is smaller than $\mathcal{H}$, and its explicit orthonormal basis makes it easy to manipulate.

In Section 3, we relate the Hilbert space $\mathcal{HP}$ to the Riemann hypothesis: the Riemann hypothesis is true if and only E_1 is the limit (in $\mathcal{HP}$) of a finite linear combination of the functions $e_q(t) = \{qt\} - \{t\}$, $q \in \mathbb{N}$. In Section 4 we

compute the coordinates of the functions e_q in the orthonormal basis $(E_n)_{n \in \mathbb{N}^*}$ and present other useful computational lemmas.

The last three sections are devoted to the study of explicit approximations. In section 5, we find a linear combination of the n functions $(e_2, \ldots, e_{n+1})$ whose first n coordinates in the basis $(E_n)_{n \in \mathbb{N}^*}$ are precisely $(1, 0, \ldots, 0)$. The coefficients of $(e_2, \ldots, e_{n-1})$ do not depend on n. It is therefore natural to consider the approximation given by these coefficients, and we study it in Section 6. Finally, we slightly modify the natural approximation by introducing a smoothing factor. For each approximation, we compute the coordinates of the linear combination in the orthonormal basis $(E_n)_{n \in \mathbb{N}^*}$, and we prove that each of these coordinates has the requested limit when n goes to infinity. By Lebesgue's dominated convergence theorem, it is enough to find uniform bounds whose series of squares converges, for the coordinates of a subsequence, in order to get a convergent subsequence. We also show how to reduce the computation of these uniform bounds. These three approximations lead to three weaker criterions for the Riemann hypothesis.

2. Orthonormality

For $n, t \in \mathbb{R}$, let us define the function E_n of $\mathcal{H}$ by

$$E_n(t) = \frac{\sin(2\pi n t) - \sin(2\pi (n-1)t)}{\pi}.$$

We shall need the following lemma to compute scalar products involving these functions.

Lemma 1. *For $p, q > 0$, we have*

$$\int_0^{+\infty} \frac{\sin(2\pi p x) \sin(2\pi q x)}{x^2} \mathrm{d}x = \pi^2 \min(p, q).$$

Proof. We get

$$\int_0^{+\infty} \frac{\sin(2\pi p x) \sin(2\pi q x)}{x^2} \mathrm{d}x = \int_0^{+\infty} \frac{\cos(2\pi (p-q)x) - \cos(2\pi (p+q)x)}{2x^2} \mathrm{d}x$$

$$= \int_0^{+\infty} \frac{-2\pi(p-q)\sin(2\pi(p-q)x) + 2\pi(p+q)\sin(2\pi(p+q)x)}{2x} \mathrm{d}x$$

$$= \frac{\pi^2}{2} (p + q - |p - q|),$$

since

$$\int_0^{+\infty} \frac{\sin(2\pi r x)}{x} \mathrm{d}x = \begin{cases} \int_0^{+\infty} \frac{\sin(2\pi x)}{x} \mathrm{d}x = \pi/2 & \text{if } r > 0, \\ 0 & \text{if } r = 0, \\ \int_0^{+\infty} \frac{\sin(-2\pi x)}{x} \mathrm{d}x = -\pi/2 & \text{if } r < 0. \end{cases}$$

The lemma easily follows.   ∎

The next proposition motivates the introduction of the functions E_n.

Proposition 1. *The family* $(E_n)_{n\in\mathbb{N}^*}$ *is an orthonormal family of* $\mathcal{H}$.

Proof. For $n, m \in \mathbb{N}^*$, $n \geqslant m$, we use Lemma 1 to obtain

$$\langle E_n, E_m \rangle = \min(n,m) + \min(n-1,m-1) - \min(n,m-1) - \min(n-1,m)$$

$$= m - \min(n-1,m) = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{if } n > m. \end{cases}$$

This proves the result, using the symmetry in m and n. ∎

However this orthonormal family is not an orthonormal basis of $\mathcal{H}$, and one can ask for a description of the space spanned by the E_n's. The answer to this question will be given by the next theorem.

Let $\mathcal{HP}$ denote the subspace of $\mathcal{H}$ whose elements are the functions f that satisfy to the following two conditions:
   i) $f(x) = f(x+1)$ for almost all $x > 0$,
   ii) $f(x) = -f(1-x)$ for almost all $x \in]0, 1[$.

The functions E_n clearly belong to $\mathcal{HP}$ for every positive integer n. As a matter of fact, we even have the following fundamental result.

Theorem 1. *The family* $(E_n)_{n\in\mathbb{N}^*}$ *is an orthonormal basis of* $\mathcal{HP}$.

Proof. By Proposition 1, it is enough to show that the family $(E_n)_{n\in\mathbb{N}^*}$ spans a vector space which is dense in $\mathcal{HP}$. We shall prove this by getting back to a known space.

Let $f, g \in \mathcal{HP}$. We set

$$\psi'(x) = \sum_{n \geqslant 0} \frac{1}{(x+n)^2}$$

so that

$$\psi'(x) + \psi'(1-x) = \sum_{n \in \mathbb{Z}} \frac{1}{(x+n)^2} = \left(\frac{\pi}{\sin(\pi x)} \right)^2 .$$

We then find

$$\langle f, g \rangle = \int_0^{+\infty} f(x)\bar{g}(x)\frac{\mathrm{d}x}{x^2} = \int_0^1 f(x)\bar{g}(x)\psi'(x)\mathrm{d}x$$

$$= \frac{1}{2} \int_0^1 \left(f(x)\bar{g}(x) + f(1-x)\bar{g}(1-x) \right) \psi'(x)\mathrm{d}x$$

$$= \frac{1}{2} \int_0^1 f(x)\bar{g}(x) \left(\psi'(x) + \psi'(1-x) \right) \mathrm{d}x$$

$$= \frac{1}{2} \int_0^1 f(x)\bar{g}(x) \left(\frac{\pi}{\sin(\pi x)} \right)^2 \mathrm{d}x$$

$$= \frac{1}{4} \int_0^2 f(x)\bar{g}(x) \left(\frac{\pi}{\sin(\pi x)} \right)^2 \mathrm{d}x$$

$$= \frac{1}{2} \int_0^1 f(2x)\bar{g}(2x) \left(\frac{\pi}{\sin(2\pi x)} \right)^2 \mathrm{d}x .$$

Let us introduce the function $\tilde{f}$ defined by

$$\tilde{f}(x) = \frac{\pi}{\sqrt{2}\sin(2\pi x)} f(2x).$$

We then have $\tilde{f} \in L^2(0,1;\mathrm{d}t)$ and $\langle f,g \rangle = \int_0^1 \tilde{f}(t)\tilde{g}(t)\mathrm{d}t$. Let us expand $\tilde{f}$ in Fourier series. From conditions i) and ii), we get $\tilde{f}(1-x) = \tilde{f}(x)$ for almost all $x \in]0,1[$, and $\tilde{f}(x) = -\tilde{f}(1/2+x)$ for almost all $x \in]0,1/2[$. We deduce that

$$\int_0^1 \tilde{f}(t)\sin(2\pi nt)\mathrm{d}t = \int_0^1 \tilde{f}(t)\cos(4\pi nt)\mathrm{d}t = 0$$

for every integer n. This in turn implies that

$$\|f\|^2 = \int_0^1 |\tilde{f}(t)|^2\mathrm{d}t = \sum_{n=1}^{\infty}\left|\int_0^1 \tilde{f}(t)\sqrt{2}\cos(2\pi(2n-1)t)\mathrm{d}t\right|^2 .$$

One easily checks that $\tilde{E}_n(t) = \sqrt{2}\cos(2\pi(2n-1)t)$, from which we get the equalities

$$\|f\|^2 = \sum_{n=1}^{\infty}\left|\int_0^1 \tilde{f}(t)\overline{\tilde{E}_n(t)}\mathrm{d}t\right|^2 = \sum_{n=1}^{\infty}|\langle f,E_n\rangle|^2 ,$$

and the proof of the theorem is complete.   ∎

This proof even gives more information. Let HP denote the function space

$$HP = \left\{ f \in L^2(0,1;\mathrm{d}t) : \begin{cases} f(1-x) = f(x) & \text{for almost every } x \in]0,1[, \\ f(x) = -f(1/2+x) & \text{for almost every } x \in]0,1/2[. \end{cases} \right\}$$

The application from $\mathcal{HP}$ to HP that maps f to $\tilde{f}$ is an isomorphism between Hilbert spaces; the image of the basis $(E_n)_{n \in \mathbb{N}^*}$ is the usual canonical basis $(\sqrt{2}\cos(2\pi(2n-1)t))_{n \in \mathbb{N}^*}$.

3. Density results and the Riemann hypothesis

Let D_a denote the dilatation defined on $\mathcal{H}$ by $D_a f(x) = f(ax)$. Let $\mathcal{H}_0$ be the vector space spanned by $\{D_a E_1 : a > 0\}$. We have the following density result.

Proposition 2. $\overline{\mathcal{H}_0} = \mathcal{H}$.

Proof. It is enough to prove that continuous function f with compact support in $\mathcal{H}$ is the limit of a sequence of elements in $\overline{\mathcal{H}_0}$. Let M be an upper bound of f and T be a real number such that $supp(f) \subset [0,T]$. Let us define $g_T :]0,+\infty[\to \mathbb{R}$

by $g_T(t) = f(2T\{x\}) - f(2T\{-x\})$. By construction, we know that g_T satisfies to the conditions i) and ii) that occur in the definition of $\mathcal{HP}$. Moreover we have

$$
\int_0^{+\infty} g_T^2(x)\frac{\mathrm{d}x}{x^2} = \int_0^{+\infty} \left(f^2(2T\{x\}) + f^2(2T\{-x\})\right)\frac{\mathrm{d}x}{x^2} = \int_{-\infty}^{+\infty} f^2(2T\{x\})\frac{\mathrm{d}x}{x^2}
$$

$$
= \int_0^1 f^2(2Tx)\left(\sum_{n\in\mathbb{Z}} \frac{1}{(x+n)^2}\right)\mathrm{d}x
$$

$$
= \int_0^{1/2} f^2(2Tx)\left(\sum_{n\in\mathbb{Z}} \frac{1}{(x+n)^2}\right)\mathrm{d}x
$$

$$
\leqslant 2\int_0^{1/2} f^2(2Tx)\left(\sum_{n\in\mathbb{N}} \frac{1}{(x+n)^2}\right)\mathrm{d}x
$$

$$
\leqslant 2(1+\zeta(2))\int_0^{1/2} f^2(2Tx)\frac{\mathrm{d}x}{x^2} = 4(1+\pi^2/6)T\|f\|^2,
$$

which shows that $g_T \in \mathcal{HP}$. By Theorem 1, this implies that g_T belongs to the space spanned by $\{E_n : n \in \mathbb{N}^*\}$. Since $E_n = D_n E_1 - D_{n-1} E_1$, we already know that $g_T \in \overline{\mathcal{H}_0}$ and thus $D_a g_T \in \overline{\mathcal{H}_0}$ for $a > 0$. The property

$$
\int_0^{+\infty} \left(g_T\left(\frac{x}{2T}\right) - f(x)\right)^2 \frac{\mathrm{d}x}{x^2} = \int_T^{+\infty} g^2\left(\frac{x}{2T}\right)\frac{\mathrm{d}x}{x^2} \leqslant M^2 \int_T^{+\infty} \frac{\mathrm{d}x}{x^2} = \frac{M^2}{T}
$$

shows that $\|f - D_{1/(2T)} g_T\|$ tends to zero as T goes to infinity. The result follows. ∎

We can also deduce this proposition from a more general result of Wiener [8]: the space spanned by $\{D_a E : a > 0\}$ is dense in $\mathcal{H}$ if and only if the Mellin transform of the function E only vanishes on a set of zero measure. Since the Mellin transform of the function E_1 is $(2\pi)^{-s}\pi^{-1}\sin(\pi s/2)\Gamma(s)$, this condition is fulfilled. However the proof given for Proposition 2 describes a sequence of approximations and the rate of convergence.

Let us put $e_q(t) = \{qt\} - \{t\}$ for $q \in \mathbb{N}^*$ and $t \in \mathbb{R}$. One easily checks that the function e_q belongs to $\mathcal{HP}$. The link to the Riemann hypothesis is given by the following theorem.

Theorem 2. *The following two propositions are equivalent.*
   i) *E_1 is the limit (in $\mathcal{HP}$) of a finite linear combination of the e_q's.*
   ii) *The Riemann hypothesis is true.*

Proof. From the work of Nyman [6] and Beurling [5], we know that the Riemann hypothesis is true if the characteristic function of the interval $[1, +\infty[$ is a limit of linear combinations of the functions $\{1/(qt)\}$, $q \in \mathbb{N}$, in the space $L^2(0, +\infty; \mathrm{d}t)$. Recently Báez-Duarte [2] showed that this condition is indeed necessary. Moreover, Báez-Duarte, Balazard, Landreau and Saias [1,3] studied an isometry of

$L^2(0,+\infty;dt)$ which maps the characteristic function of the interval $[1,+\infty[$ on the function $E_1(t)/t$, and the functions $\{1/(qt)\}$ on the functions $\{qt\}/(qt)$. Therefore the Riemann hypothesis is true if and only if the function $E_1(t)/t$ is a limit of linear combinations of the functions $\{qt\}/(qt)$ in $L^2(0,+\infty;dt)$. Equivalently, the Riemann hypothesis is true if and only if the function $E_1(t)$ is a limit of linear combinations of the functions $\{qt\}$ in $\mathcal{H}$. Since a finite linear combination of the e_q's is also a finite linear combination of the functions $\{qt\}$, we still have to show that E_1 is the limit of a finite linear combination of the e_q's (in $\mathcal{HP}$) as soon as E_1 is the limit of a finite linear combination of the $\{qt\}$'s (in $\mathcal{H}$).

Assume that the functions $R_n(t) = E_1(t) - \sum_{1\leqslant q\leqslant n} c_{q,n}\{qt\}$ tend to zero in $\mathcal{H}$. Noticing that $R_n(t) + R_n(1-t) = -\sum_{1\leqslant q\leqslant n} c_{q,n}$ for $t \in \mathbb{R}\setminus\mathbb{Q}$, we find

$$\frac{3}{2}\left(\sum_{1\leqslant q\leqslant n} c_{q,n}\right)^2 \leqslant \int_{1/3}^{2/3} (R_n(t) + R_n(1-t))^2 \frac{dt}{t^2}$$

$$\leqslant 2\int_{1/3}^{2/3}\left(\frac{R_n(t)^2}{t^2} + \frac{R_n(1-t)^2}{t^2}\right)dt$$

$$\leqslant 2\int_{1/3}^{2/3}\left(\frac{R_n(t)^2}{t^2} + 4\frac{R_n(1-t)^2}{(1-t)^2}\right)dt$$

$$= 10\int_{1/3}^{2/3}\frac{R_n(t)^2}{t^2}dt \leqslant 10\|R_n\|^2.$$

Put $R'_n(t) = R_n(t) + \sum_{1\leqslant q\leqslant n} c_{q,n}\{t\} = E_1(t) - \sum_{2\leqslant q\leqslant n} c_{q,n}e_q(t)$. We obtain

$$\|R'_n\| \leqslant \|R_n\| + \left|\sum_{1\leqslant q\leqslant n} c_{q,n}\right| \|\{.\}\| \leqslant \left(1 + \sqrt{20/3}\|\{.\}\|\right)\|R_n\|,$$

which proves that R'_n also tends to zero in $\mathcal{H}$. This shows that E_1 is also the limit of a finite linear combination of the e_q's (in $\mathcal{HP}$) and the proof of the theorem is complete. ∎

4. Computational lemmas

Let us first compute the coordinates of the function e_q in the orthonormal basis given in Theorem 1.

Lemma 2. *For $p,q \in \mathbb{N}^*$, we have*

$$\langle E_p, e_q\rangle = \log q - \sum_{p/q\leqslant k<p}\frac{1}{k}.$$

Proof. We find

$$
\int_0^{+\infty} \sin(2\pi px)\{qx\}\frac{\mathrm{d}x}{x^2} = q\int_0^{+\infty} \sin(2\pi px)\frac{\mathrm{d}x}{x} - \sum_{k=1}^{\infty} k \int_{k/q}^{(k+1)/q} \sin(2\pi px)\frac{\mathrm{d}x}{x^2}
$$

$$
= \frac{\pi q}{2} - \sum_{k=1}^{\infty} \int_{k/q}^{+\infty} \sin(2\pi px)\frac{\mathrm{d}x}{x^2}
$$

$$
= \frac{\pi q}{2} - \sum_{k=1}^{\infty} \frac{1}{k} \int_{1/q}^{+\infty} \sin(2\pi pkx)\frac{\mathrm{d}x}{x^2}
$$

$$
= \frac{\pi q}{2} - \int_{1/q}^{+\infty} \left(\sum_{k=1}^{\infty} \frac{\sin(2\pi pkx)}{k} \right) \frac{\mathrm{d}x}{x^2}
$$

$$
= \frac{\pi q}{2} - \pi \int_{1/q}^{+\infty} \left(\frac{1}{2} - \{px\} \right) \frac{\mathrm{d}x}{x^2} = \pi \int_{1/q}^{+\infty} \{px\}\frac{\mathrm{d}x}{x^2}\, .
$$

We deduce that

$$
\langle E_p, e_q \rangle = \int_{1/q}^1 (\{px\} - \{(p-1)x\})\frac{\mathrm{d}x}{x^2} = \log q + \int_{1/q}^1 ([(p-1)x] - [px])\frac{\mathrm{d}x}{x^2}
$$

$$
= \log q - \sum_{p/q \leqslant k < p} \int_{k/p}^{k/(p-1)} \frac{\mathrm{d}x}{x^2} = \log q - \sum_{p/q \leqslant k < p} 1/k\, .
$$

$\blacksquare$

One can also compute scalar products of the e_q's.

Lemma 3. *For $p, q \in \mathbb{N}^*$, we have*

$$
\langle e_p, e_q \rangle = \frac{p-1}{2}\log q + \frac{q-1}{2}\log p
$$

$$
- \frac{\pi}{2}\left(V\left(\frac{p}{\gcd(p,q)}, \frac{q}{\gcd(p,q)} \right) + V\left(\frac{q}{\gcd(p,q)}, \frac{q}{\gcd(p,q)} \right) - V(1,p) - V(1,q) \right),
$$

where $V(p,q)$ denotes Vassiounine's sum:

$$
V(p,q) = \sum_{k=1}^{q-1} \left\{ \frac{kp}{q} \right\} \cot \frac{k\pi}{q}\, .
$$

Proof. The result easily follows from the following form [4, Proposition 5] of Vassiounine's result [9]:

$$
\int_0^{+\infty} \{px\}\{qx\}\frac{\mathrm{d}x}{x^2} = \frac{p+q}{2}(\log(2\pi) - \gamma) - \frac{p-q}{2}\log\frac{p}{q}
$$

$$
- \frac{\pi}{2}\left(V\left(\frac{p}{\gcd(p,q)}, \frac{q}{\gcd(p,q)} \right) + V\left(\frac{q}{\gcd(p,q)}, \frac{q}{\gcd(p,q)} \right) \right). \quad \blacksquare
$$

By Theorem 2, we look after linear combinations of the e_q's which converge to E_1. In order to use Lebesgue's dominated convergence theorem, we seek linear combinations of the e_q's, whose coordinates are uniformly bounded by terms of a convergent series and converge to those of E_1. There are several candidates to provide good approximations, and we shall study them in the following subsections. For these approximations, we shall show the convergence of the coordinates and upper bound the coordinates of large index with terms of a convergent series.

Let $n \geqslant 2$ be an integer. We shall extensively use the notations

$$S_n = \sum_{q=1}^{n-1} \frac{\mu(q)}{q} \quad \text{and} \quad S'_n = \sum_{q=1}^{n-1} \frac{\mu(q)}{q} \log q \,.$$

We shall also need to estimate these sums.

Lemma 4. *For any $A > 0$, we have*

$$S_n = O\left((\log n)^{-A}\right) \quad \text{and} \quad S'_n = -1 + O\left((\log n)^{-A}\right) \,.$$

Proof. It is well-known that

$$\sum_{1 \leqslant q \leqslant x} \mu(q) = O\left(x(\log x)^{-A}\right)$$

for any $A > 0$ (see [6, formula (2.48)] for instance). Since the prime number theorem implies the evaluations

$$\sum_{q=1}^{\infty} \frac{\mu(q)}{q} = 1 + \sum_{q=1}^{\infty} \frac{\mu(q)}{q} \log q = 0 \,,$$

the lemma follows from partial summation. ∎

Finally define the harmonic sum

$$h_x = \sum_{1 \leqslant k < x} 1/k \,.$$

5. The triangular approximation

It is natural to seek a linear combination of the n functions $(e_2, \ldots, e_{n+1})$ whose first n coordinates in the basis $(E_1, E_2, \ldots)$ are precisely $(1, 0, \ldots, 0)$. This linear combination is given in the next proposition and requires further notations. Put

$$\lambda_{q,n} = \begin{cases} \frac{\mu(q)}{q} & \text{for } 1 \leqslant q \leqslant n-1, \\ \frac{1 + S'_n - S_n \log(n+1)}{\log(1+1/n)} & \text{for } q = n, \\ \frac{-1 - S'_n + S_n \log n}{\log(1+1/n)} & \text{for } q = n+1. \end{cases}$$

Proposition 3. *We have the expansion*

$$\sum_{q=2}^{n+1} \lambda_{q,n} e_q + E_1 = \sum_{p=n+1}^{\infty} x_{n,p} E_p \,,$$

with

$$x_{n,p} = \sum_{q=1}^{n-1} \frac{\mu(q)}{q} \left(h_{p/q} - h_{p/n} - \log \frac{n}{q} \right) + \lambda_{n+1,n} \left(\log \left(1 + 1/n \right) - h_{p/n} + h_{p/(n+1)} \right)$$

for $p \geqslant n+1$.

Proof. By Theorem 1 and Lemma 2, we get

$$\sum_{q=2}^{n+1} \lambda_{q,n} e_q = \sum_{p=1}^{\infty} x_{n,p} E_p$$

with

$$x_{n,p} = \sum_{q=2}^{n+1} \lambda_{q,n} \left(\log q - \sum_{p/q \leqslant k < p} \frac{1}{k} \right) \,.$$

For $p = 1$ we find

$$x_{n,1} = S'_n + \lambda_{n,n} \log n + \lambda_{n+1,n} \log(n+1) \,.$$

For $2 \leqslant p \leqslant n$, we first notice that

$$\sum_{q=1}^{n-1} \frac{\mu(q)}{q} \sum_{1 \leqslant k < p/q} \frac{1}{k} = \sum_{1 \leqslant m \leqslant p-1} \frac{1}{m} \sum_{q|m} \mu(q) = 1 \,.$$

We thus obtain

$$x_{n,p} = \sum_{q=2}^{n-1} \frac{\mu(q)}{q} \left(\log q - \sum_{p/q \leqslant k < p} \frac{1}{k} \right) + \lambda_{n,n} \langle E_p, e_n \rangle + \lambda_{n+1,n} \langle E_p, e_{n+1} \rangle$$

$$= \sum_{q=1}^{n-1} \frac{\mu(q)}{q} \left(\log q - \sum_{p/q \leqslant k < p} \frac{1}{k} \right) + \lambda_{n,n} (\log n - h_p) + \lambda_{n+1,n} (\log(n+1) - h_p)$$

$$= 1 + \sum_{q=1}^{n-1} \frac{\mu(q)}{q} (\log q - h_p) + \lambda_{n,n} (\log n - h_p) + \lambda_{n+1,n} (\log(n+1) - h_p)$$

$$= 1 + S'_n + \lambda_{n,n} \log n + \lambda_{n+1,n} \log(n+1) - (S_n + \lambda_{n,n} + \lambda_{n+1,n}) h_p \,.$$

The pair $(\lambda_{n,n}, \lambda_{n+1,n})$ is a solution of the system

$$\begin{cases} 1 + S'_n + \lambda_{n,n} \log n + \lambda_{n+1,n} \log(n+1) = 0\,, \\ S_n + \lambda_{n,n} + \lambda_{n+1,n} = 0\,, \end{cases}$$

which implies that

$$x_{n,p} = \begin{cases} -1 & \text{if } p = 1, \\ 0 & \text{if } 2 \leqslant p \leqslant n. \end{cases}$$

For $p > n$, we notice that

$$x_{n,p} = \sum_{q=1}^{n+1} \lambda_{q,n} \left(\log q - h_p + h_{p/q} \right)$$

$$= \sum_{q=1}^{n+1} \lambda_{q,n} \left(h_{p/q} - h_{p/n} - \log \frac{n}{q} \right) + (S_n + \lambda_{n,n} + \lambda_{n+1,n}) \left(h_{p/n} - h_p + \log n \right)$$

$$= \sum_{q=1}^{n-1} \frac{\mu(q)}{q} \left(h_{p/q} - h_{p/n} - \log \frac{n}{q} \right) + \lambda_{n+1,n} \left(\log(1 + 1/n) - h_{p/n} + h_{p/(n+1)} \right)$$

and the formula is proved. ∎

By construction, it is obvious that $\lim_{n \to \infty} x_{n,p} = 0$. In order to use Lebesgue's dominated convergence theorem, we seek an upper bound $|x_{n,p}| \leqslant u_p$, where u_p is the general term of a convergent series. In the next proposition, we explicit u_p for $p \geqslant n^4$.

Proposition 4. *When n and p goes to infinity, we have*

$$x_{n,p} = O\left(\frac{n^2}{p \log n} \right),$$

and therefore $|x_{n,p}| \ll 1/(\sqrt{p} \log p)$ *for* $p \geqslant n^4$.

Proof. Euler-Maclaurin formula ensures that

$$\sum_{p/q \leqslant k < p/n} \frac{1}{k} = \log \frac{n}{q} + O\left(\frac{q}{p} \right).$$

We get this way

$$\sum_{q=1}^{n-1} \frac{\mu(q)}{q} \left(h_{p/q} - h_{p/n} - \log \frac{n}{q} \right) = \sum_{q=1}^{n-1} O\left(\frac{n}{qp} \right) = O\left(\frac{n \log n}{p} \right).$$

Similarly we have $\log(1 + 1/n) - h_{p/n} + h_{p/(n+1)} = O(n/p)$ and Lemma 4 implies $\lambda_{n+1,n} = O(n/\log n)$. Finally we obtain

$$x_{n,p} = O\left(\frac{n \log n}{p} + \frac{n^2}{p \log n} \right) = O\left(\frac{n^2}{p \log n} \right).$$

The second estimate is an easy consequence of this one. ∎

We deduce from these two propositions a sufficient condition for the Riemann hypothesis to be true.

Criterion 1. *If there exists a sequence* (u_p) *with* $\sum_p u_p^2 < \infty$ *such that*

$$\forall p \in [n+1, n^4], \quad |x_{n,p}| \leqslant u_p$$

for infinitely many n's, then the Riemann hypothesis is true.

6. The natural approximation

From the last subsection, it is natural to consider the linear combination

$$\sum_{q=1}^{n} \frac{\mu(q)}{q} e_q := \sum_{p=1}^{\infty} y_{n,p} E_p,$$

which enjoy the following properties.

Proposition 5. *We have*

$$\lim_{n \to \infty} y_{n,p} = \begin{cases} -1 & if \ \ p = 1 \\ 0 & otherwise \end{cases}.$$

Moreover $y_{n,p} = O(n/p)$ and therefore $|y_{n,p}| \ll 1/(\sqrt{p} \log p)$ for $p \geqslant n^2 \log n$.

Proof. By Lemma 2, we get

$$y_{n,p} = \sum_{q=1}^{n} \frac{\mu(q)}{q} \left(\log q - \sum_{p/q \leqslant k < p} \frac{1}{k} \right)$$

$$= S'_{n+1} - S_{n+1} \sum_{1 \leqslant k < p} \frac{1}{k} + \sum_{q=1}^{n} \frac{\mu(q)}{q} \sum_{1 \leqslant k < p/q} \frac{1}{k}.$$

Note that

$$\sum_{q=1}^{n} \frac{\mu(q)}{q} \sum_{1 \leqslant k < p/q} \frac{1}{k} = \sum_{1 \leqslant m < p} \frac{1}{m} \sum_{\substack{q | m \\ 1 \leqslant q \leqslant n}} \mu(q) = \begin{cases} 0 & if \ \ p = 1, \\ 1 & if \ \ n+1 \geqslant p \geqslant 2. \end{cases}$$

By Lemma 4, this implies that

$$\lim_{n \to \infty} y_{n,p} = \begin{cases} -1 & if \ \ p = 1 \\ 0 & otherwise \end{cases}.$$

As in the proof of Proposition 4, we get $\langle e_q, E_p \rangle = O(q/p)$, from which the estimate $y_{n,p} = O(n/p)$ easily follows. The second estimate is a straightforward consequence of the first one. ■

We can also state another criterion.

Criterion 2. *If there exists a sequence (u_p) with $\sum_p u_p^2 < \infty$ such that*

$$\forall p \in [1, n^2 \log n], \quad |y_{n,p}| \leqslant u_p$$

for infinitely many n's, then the Riemann hypothesis is true.

The range $[1, n^2 \log n[$ is smaller than the range from the previous subsection, namely $[n+1, n^4[$. However the sum of the $y_{n,p}^2$'s over the range $[1, n+1[$ can be quite large, as shown by the next proposition.

Proposition 6. *We have*

$$\sum_{2 \leqslant p \leqslant n+1} y_{n,p}^2$$

$$= n \left(1 + S'_{n+1} - \frac{n+1}{n} (h_{n+2} - 1) S_{n+1}\right)^2 + \left(n - \frac{n+1}{n} h_{n+1}^2 + h_{n+1}\right) S_{n+1}^2.$$

Proof. From the proof of the proposition 5, we know that $y_{n,p} = 1 + S'_{n+1} - h_p S_{n+1}$ for $2 \leqslant p \leqslant n+1$. We compute

$$\sum_{2 \leqslant p \leqslant n+1} h_p = \sum_{1 \leqslant k < p \leqslant n+1} \frac{1}{k} = \sum_{1 \leqslant k \leqslant n+1} \frac{n+1-k}{k} = (n+1)(h_{n+2} - 1)$$

and

$$\sum_{2 \leqslant p \leqslant n+1} h_p^2 = \sum_{1 \leqslant k, l < p \leqslant n+1} \frac{1}{kl} = \sum_{1 \leqslant k, l \leqslant n} \frac{n+1 - \max(k, l)}{kl}$$

$$= (n+1)h_{n+1}^2 - 2 \sum_{1 \leqslant k < l \leqslant n} \frac{l}{kl} - \sum_{1 \leqslant k \leqslant n} \frac{k}{k^2}$$

$$= (n+1)h_{n+1}^2 - 2n(h_{n+1} - 1) - h_{n+1}.$$

This gives

$$\sum_{2 \leqslant p \leqslant n+1} y_{n,p}^2 = n(1 + S'_{n+1})^2 - 2(n+1)(h_{n+2} - 1)(1 + S'_{n+1})S_{n+1}$$

$$+ \left((n+1)h_{n+1}^2 + 2n - (2n+1)h_{n+1}\right) S_{n+1}^2$$

$$= n \left(1 + S'_{n+1} - (1 + 1/n)(h_{n+2} - 1) S_{n+1}\right)^2 + \delta_n S_{n+1}^2,$$

with

$$\delta_n = (n+1)h_{n+1}^2 + 2n - (2n+1)h_{n+1} - (n+1)^2/n \left(h_{n+1} - n/(n+1)\right)^2$$

$$= n - (1 + 1/n)h_{n+1}^2 + h_{n+1}. \qquad \blacksquare$$

It follows from this proposition and Criterion 2 that we can restrict our study to sequences of integers n for which there exists an absolute constant C such that we simultaneously have $|S_{n+1}| \leqslant C/\sqrt{n}$ and $|1 + S'_{n+1} - S_{n+1} \log n| \leqslant C/\sqrt{n}$. These two conditions are quite restrictive, and it seems difficult to find infinite sequences for which they hold together. Indeed, it is well-known [8] that $\sum_{1 \leqslant q \leqslant n} \mu(q) = \Omega(\sqrt{n})$.

7. Smooth approximations

Let us slightly modify the natural approximation by introducing a smoothing factor:

$$\sum_{q=1}^{n} \frac{\mu(q)}{q^{1+\epsilon}} e_q := \sum_{p=1}^{\infty} y_{n,p}(\epsilon) E_p .$$

Proposition 7. Assume $\epsilon_n \geqslant 2 \log \log n / \log n$ and $\epsilon_n = o(1)$. Then we have

$$\lim_{n \to \infty} y_{n,p}(\epsilon_n) = \begin{cases} -1 & \text{if } p = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover $y_{n,p}(\epsilon_n) = O(n/p)$ and therefore $|y_{n,p}(\epsilon_n)| \ll 1/(\sqrt{p} \log p)$ for $p \geqslant n^2 \log n$.

Proof. We proceed as in the proof of Proposition 5:

$$\sum_{q=1}^{n} \frac{\mu(q)}{q^{1+\epsilon_n}} \sum_{1 \leqslant k < p/q} \frac{1}{k} = \sum_{1 \leqslant m < p} \frac{1}{m} \sum_{\substack{q|m \\ 1 \leqslant q \leqslant n}} \frac{\mu(q)}{q^{\epsilon_n}} = \sum_{1 \leqslant m < p} \frac{1}{m} \prod_{p|m} \left(1 - \frac{1}{p^{\epsilon_n}}\right)$$

$$= \begin{cases} o(1) & \text{if } p = 1, \\ 1 + o(1) & \text{if } n + 1 \geqslant p \geqslant 2. \end{cases}$$

Therefore we get

$$y_{n,p}(\epsilon_n) = -\langle E_1, E_p \rangle + 1 + \sum_{1 \leqslant q \leqslant n} \frac{\mu(q) \log q}{q^{1+\epsilon_n}} - h_p \sum_{1 \leqslant q \leqslant n} \frac{\mu(q)}{q^{1+\epsilon_n}} + o(1)$$

$$= -\langle E_1, E_p \rangle + 1 - (1/\zeta)'(1 + \epsilon_n) - h_p (1/\zeta)(1 + \epsilon_n)$$

$$- \sum_{q>n} \frac{\mu(q) \log q}{q^{1+\epsilon_n}} + h_p \sum_{q>n} \frac{\mu(q)}{q^{1+\epsilon_n}} + o(1)$$

$$= -\langle E_1, E_p \rangle - \sum_{q>n} \frac{\mu(q) \log q}{q^{1+\epsilon_n}} + h_p \sum_{q>n} \frac{\mu(q)}{q^{1+\epsilon_n}} + o(1) .$$

We now use direct estimates for $n \geqslant 4$

$$\left| \sum_{q>n} \frac{\mu(q) \log q}{q^{1+\epsilon_n}} \right| \leqslant \sum_{q>n} \frac{\log q}{q^{1+\epsilon_n}} \leqslant \int_{n}^{\infty} \frac{\log x}{x^{1+\epsilon_n}} \, \mathrm{d}x = \frac{\epsilon_n \log n + 1}{\epsilon_n^2 n^{\epsilon_n}} = O\left(\frac{1}{\log \log n}\right)$$

and

$$\left| \sum_{q>n} \frac{\mu(q)}{q^{1+\epsilon_n}} \right| \leqslant \sum_{q>n} \frac{1}{q^{1+\epsilon_n}} \leqslant \int_n^\infty \frac{1}{x^{1+\epsilon_n}} \mathrm{d}x = \frac{1}{\epsilon_n n^{\epsilon_n}} \leqslant \frac{1}{2 \log n \log \log n}$$

to complete the proof of first part the proposition. The second part is a straight-forward generalisation of Proposition 5. ∎

We still have another criterion.

Criterion 3. *Assume there exists a sequence* $(\epsilon_n)_{n\in\mathbb{N}}$, *with* $\epsilon_n \geqslant 2 \log \log n / \log n$ *and* $\epsilon_n = o(1)$, *and a sequence* (u_p) *with* $\sum_p u_p^2 < \infty$ *such that*

$$\forall p \in [1, n^2 \log n], \quad |y_{n,p}(\epsilon_n)| \leqslant u_p$$

for infinitely many n's, *then the Riemann hypothesis is true.*

This criterion is quite efficient. Indeed, Balazard [2, Proposition 3.1] proved that, under the Riemann hypothesis, the criterion is fulfilled with $\epsilon_n = c / \log \log n$, for some positive constant c; he even upper bounded the distance between this approximation and its limit.

References

[1] L. Báez-Duarte, *A class of invariant unitary operators* Adv. Math. **144** (1999), 1–12.

[2] L. Báez-Duarte, *A strengthening of the Nyman-Beurling criterion for the Riemann hypothesis*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **14** (2003), 5–11.

[3] L. Báez-Duarte, M. Balazard, B. Landreau, É. Saias, *Notes sur la fonction ζ de Riemann*, 3, Adv. Math. **149** (2000), 130–144.

[4] L. Báez-Duarte, M. Balazard, B. Landreau, É. Saias, *Étude de l'autocorrélation multiplicative de la fonction "partie fractionnaire"*, The Ramanujan Journal **9** (2005), 215–240.

[5] A. Beurling, *A closure problem related to the Riemann Zeta-function*, Proc. Nat. Acad. Sci. **41** (1955), 312–314.

[6] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, vol. 53, American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 2004.

[7] B. Nyman, *On some groups and semigroups of translations*, Ph. D. thesis, Uppsala, 1950.

[8] E.C. Titchmarsh, *The theory of the Riemann zeta-function*, second edition, edited and with a preface by D.R. Heath-Brown The Clarendon Press, Oxford University Press, New York, 1986.

[9] V.I. Vassiounine, *On a biorthogonal system associated with the Riemann hypothesis*, Algebra i Analiz **7** (1995), 118–135; translation in *St. Petersburg Math. J.* **7** (1996), 405–419.

[10] N. Wiener, *Tauberian thoerems*, Annals of Math. **33** (1932), 1–100.

Address: Institut Camille Jordan CNRS UMR 5208, Université de Lyon Université Lyon 1, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne Cedex, FRANCE

E-mail: `Laurent.Habsieger@math.univ-lyon1.fr`